Fault Diagnosis in Multivariate Statistical Process Monitoring

André George Mostert

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
in the Department of Statistical Sciences at the
University of Cape Town
March 2021

Supervisors:
Professor S Lubbe
Department of Statistics and Actuarial Science
University of Stellenbosch

Professor RLJ Coetzer
Department of Statistics
North-West University &
Department of Mathematical Statistics and Actuarial Science
University of the Free State

Declaration

I, Andre George Mostert, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature Signed by candidate Date 15 March 2021

Abstract

The application of multivariate statistical process monitoring (MSPM) methods has gained considerable momentum over the last couple of decades, especially in the processing industry for achieving higher throughput at sustainable rates, reducing safety related events and minimizing potential environmental impacts. Multivariate process deviations occur when the relationships amongst many process characteristics are different from the expected. The fault detection ability of methods such as principal component analysis (PCA) and process monitoring has been reported in literature and demonstrated in selective practical applications. However, the methodologies employed to diagnose the reason for the identified multivariate process faults have not gained the anticipated traction in practice. One explanation for this might be that the current diagnostic approaches attempt to rank process variables according to their individual contribution to process faults. However, the lack of these approaches to correctly identify the variables responsible for the process deviation is well researched and communicated in literature. Specifically, these approaches suffer from a phenomenon known as fault smearing.

In this research it is argued, using several illustrations, that the objective of assigning individual importance rankings to process variables is not appropriate in a multivariate setting. A new methodology is introduced for performing fault diagnosis in multivariate process monitoring. More specifically, a multivariate diagnostic method is proposed that ranks variable pairs as opposed to individual variables. For PCA based MSPM, a novel fault diagnosis method is developed that decomposes the fault identification statistics into a sum of parts, with each part representing the contribution of a specific variable pair. An approach is also developed to quantify the statistical significance of each pairwise contribution.

In addition, it is illustrated how the pairwise contributions can be analysed further to obtain an individual importance ranking of the process variables. Two methodologies are developed that can be applied to calculate the individual ranking following the pairwise contributions analysis. However, it is advised that the individual rankings should be interpreted together with the pairwise contributions. The application of this new approach to PCA based MSPM and fault diagnosis is illustrated using a simulated data set.

Due to the complexity of industrial chemical processes, it is often required that non-linear methods such as kernel PCA be employed for MSPM. Existing fault diagnosis methods used for kernel PCA based MSPM follow the same approach of assigning individual rankings to process variables. It is illustrated that this diagnosis approach is not ideal for the non-linear setting. Therefore, a new fault diagnosis methodology is developed in

this research, which calculates pairwise variable contributions to diagnose non-linear multivariate process deviations. The pairwise contribution of variables i and j is defined as the distance from the reference region to the new observation for these two variables. In this approach each pair of variables is considered in isolation. However, this methodology is dependent on specifying suitable distance measures. It is illustrated how techniques such as kernel PCA, Support Vector Machine (SVM) classification and One-Class SVM can be applied to specify appropriate distance measures for accurate fault diagnosis. Furthermore, like PCA based MSPM, the pairwise contributions can be analysed further to specify an individual ranking of process variables. The application of this new fault diagnosis methodology for non-linear MSPM is demonstrated using simulated data.

The application and value of the newly developed fault diagnosis methods are illustrated using data from a benchmark simulation and a commercial chemical plant. The practical application illustrates how the new methodology of diagnosing multivariate process faults provides a multivariate answer. Furthermore, it is illustrated how the developed methods assist engineers to interpret why the process is statistically identified as displaying behaviour that is multivariately different from ideal operating conditions. Therefore, the improved interpretation of the statistical deviation will assist engineers to identify the appropriate corrective approach required to guide the process back into multivariate control.

Contents

1 Introduction **1**

1.1 Background 1
1.2 Review of statistical process control 3
1.2.1 Univariate fault diagnosis 6
1.3 Multivariate statistical process monitoring 7
1.3.1 Simulation study setup 7
1.3.2 Univariate SPC 11
1.3.3 Multivariate statistics 17
1.3.4 Multivariate fault diagnosis 20
1.4 Nonlinear multivariate statistical process monitoring 25
1.5 Summary 28
1.6 Research focus and objectives 29
1.7 Layout of research 30

2 Principal Component Analysis based Fault Diagnosis **32**

2.1 Introduction 32
2.2 Principal Component Analysis 32
2.3 Fault detection statistics 35
2.4 Fault identification methods 37
2.5 Fault smearing 40
2.6 Process fault types 42
2.7 Mathematical fault diagnosis-ability 45
2.7.1 Complete decomposition contribution 46
2.7.2 Partial decomposition contribution 46
2.7.3 Reconstruction based contribution 47
2.8 PCA fault diagnosis simulation study 48
2.8.1 Objective 48
2.8.2 Data definition 48
2.8.3 PCA results 49

2.8.4 Fault diagnosis results . . . 51
2.9 Recent developments in fault diagnosis analysis . . . 63
2.10 Summary . . . 65

3 New Methodology for PCA Fault Diagnosis 67
3.1 Introduction . . . 67
3.2 Motivation . . . 67
3.3 Pairwise variable contribution analysis . . . 72
3.4 Pairwise contribution importance . . . 76
3.5 Variable importance in pairwise contributions . . . 77
3.6 Deployment . . . 80
3.7 PCA fault diagnosis simulation study continued (Part 1) . . . 82
3.7.1 Objective . . . 82
3.7.2 Pairwise contribution results . . . 82
3.8 PCA fault diagnosis simulation study continued (Part 2) . . . 89
3.8.1 Objective . . . 89
3.8.2 Results . . . 89
3.8.3 Single sensor fault . . . 91
3.8.4 Multiple sensor fault . . . 92
3.8.5 Multivariate sensor fault . . . 94
3.9 Summary . . . 95

4 Kernel Principal Component Analysis based Fault Diagnosis 98
4.1 Introduction . . . 98
4.2 Kernel principal component analysis . . . 98
4.3 Fault detection statistics . . . 107
4.4 Fault diagnosis methods . . . 109
4.5 Kernel PCA Simulation Study . . . 111
4.5.1 Objective . . . 111
4.5.2 Data definition . . . 111
4.5.3 PCA and kernel PCA results . . . 114
4.5.4 Kernel PCA fault diagnosis simulation study . . . 122
4.6 Summary . . . 131

5 New Methodology for Kernel PCA Fault Diagnosis 133
5.1 Introduction . . . 133
5.2 Pairwise variable importance analysis . . . 134
5.3 Univariate variable ranking in pairwise importance . . . 137

5.4 Distance measure definition . . . 138
5.4.1 Kernel Principal Component Analysis . . . 139
5.4.2 Support Vector Machine Discriminant Analysis . . . 141
5.4.3 One-Class Support Vector Machine Discriminant Analysis . . . 152
5.5 Kernel PCA Simulation Study continued (Part 1) . . . 155
5.5.1 Objective . . . 155
5.5.2 Kernel PCA . . . 155
5.5.3 SVM . . . 160
5.5.4 One-Class SVM . . . 164
5.6 Kernel PCA Simulation Study continued (Part 2) . . . 167
5.6.1 Objective . . . 167
5.6.2 Results . . . 168
5.7 Summary . . . 177

6 Practical application **180**
6.1 Introduction . . . 180
6.2 Case study: The Tennessee Eastman Process . . . 181
6.2.1 Background and Objective . . . 181
6.2.2 Variables and Data . . . 181
6.2.3 PCA based fault diagnosis . . . 185
6.2.4 Kernel PCA based fault diagnosis . . . 201
6.3 Case study: Coal Fired Steam Boiler . . . 213
6.3.1 Background and Objective . . . 213
6.3.2 Data . . . 214
6.3.3 PCA MSPM . . . 215
6.3.4 Kernel PCA MSPM . . . 234
6.4 Summary . . . 240

7 Conclusions and Future Research **244**
7.1 Conclusions . . . 244
7.2 Future Research . . . 246

A R code **248**
A.1 R code for PCA based MSPM . . . 248
A.2 R code for kernel PCA based MSPM. . . . 262

Chapter 1

Introduction

1.1 Background

The way in which manufacturing organisations create value has exponentially evolved for the better in recent time. This evolutionary trajectory is characterised by the so-called industrial revolutions. The first industrial revolution identifies the period when the usage of human effort and animals as primary energy sources was replaced by the mechanical power induced by the use of fossil fuels. Improvements in electricity distribution, the identification of new power generating methods and the synthesis of ammonia describe the second industrial revolution. The most recent evolution, the third industrial revolution, characterises the advancements made in data generation, work-up and communication as a result of the implementation of advanced digital infrastructure.

Currently we are at the advent of the fourth industrial revolution. This revolution largely builds on the third, expanding the impact of digitalization on the manufacturing process. This is made possible by the increased ability of computer infrastructure, both physical (e.g. data storage, processing power) and virtual (e.g. software, algorithms), to ingest and digest large data sources.

It is widely argued that in order to maximise the impact of the fourth industrial revolution in the manufacturing process it is required that the process should have a digital twin. A digital twin is defined by Bolton et al. (2018) as 'a dynamic virtual representation of a physical object or system across its life-cycle, using real-time data to enable understanding, learning and reasoning'. In order to obtain this virtual replica of a physical system, it is required that a larger array of process characteristics be measured and digitized. Advances in sensory devices have made it possible to extend the reach of the internet into physical devices. This extension of the internet, termed the Internet of Things (IoT), will assist to enable the realisation of digital twin images of physical processes.

Given the digital twins of processes implies that large quantities of real time data will be accumulated. Providing manufacturing engineers with an information rich data source. A source of information that will dramatically change industrial processes to be more efficient as a result of the improved process understanding that it enables.

The valuable process knowledge hidden inside the digital twin needs to be thoughtfully extracted in order to attain an optimised factory. Recent advances in computing power have made it possible to expand on and improve the statistical tools that can be used to extract value from data. Specifically, significant advancements have been made to the statistical tools required to analyse the large volumes of data generated by commercial processes (see He and Wang (2018)).

An important subset of the library of statistical methods, called statistical process monitoring (SPM), will be the area of focus of this research document. Statistical process monitoring is a collection of data driven techniques used to monitor process characteristics relative to the ideal operating conditions. Process conditions that are statistically different from the targeted operating ranges are flagged as faulty operation. Processes that are statistically similar to the ideal state are referred to as being statistically in-control while those conditions that are not represents an out-of-control state. The identification of the in-control and out-of-control data points is called fault detection. Given the out-of-control condition, it is required that a diagnostic analysis be done to statistically quantify the difference between the desirable and undesirable states. The objective is to assist process engineers to identify, within the physical boundaries of the process, what needs to be done in order to attain an in-control process. This analysis is often referred to as fault diagnosis.

Statistical process monitoring (SPM) can be grouped into two sets of analyses. Univariate statistical process monitoring and multivariate statistical process monitoring. He and Wang (2018) classify the development of SPM into a roadmap of three generations. The first generation identifies the era when statistical process control (SPC) was introduced as a method of quality control in manufacturing. Multivariate statistical process monitoring (MSPM) defines the second generation, which addresses the shortcomings of SPC. The definition of the third generation is a set of techniques that addresses the limitations identified in the first two generations. He and Wang (2018), however, state that there exists no specific name to label the third generation of SPM. A possible reason why it is difficult to categorise the third generation is due to the variety of possible techniques available that can be applied for SPM in this generation.

1.2 Review of statistical process control

Statistical process control is a univariate analysis technique, first introduced in the early 1920's by Walter Shewhart, a Bell Laboratories employee (Deming, 1968). The objective of SPC charts is to provide a statistically objective tool to distinguish between 'common cause' variability and 'assignable cause' variability. Development of the SPC method is based on the observation that any process operating under normal conditions exhibit variability that can be considered to be common to the process. Any uncommon variability observed, in the quality characteristic being monitored, should be explainable by an assignable cause. SPC control charts are characterised by control limits, a central line and a set of decision rules that are used to visually identify when a process is operating abnormally. The xbar chart and R chart are used most frequently in manufacturing applications (Leavenworth and Grant, 2000). Collectively these two charts are often referred to as Shewart control charts. Analysis of the mean value of a process variable is done using the xbar chart. The R chart is used to analyse the variability of the process characteristic being monitored. Combined, these two charts can be used to identify if the process display a 'constant system of chance causes' (Leavenworth and Grant, 2000).

Typical Shewhart chart examples are displayed by Figure 1.1 and Figure 1.2. Figure 1.1 is an example of a xbar chart and Figure 1.2 is an example of a R chart. This example is taken from the qcc R package for statistical process control introduced in Scrucca (2004). In this particular example the manufacturing process considered is a forging process that produce piston rings for an automotive engine. The quality characteristic monitored is the inside diameter of the piston rings produced.

In Figures 1.1 and 1.2, each circular point on the graph represents a statistic, e.g the mean, calculated on a subgroup of five piston rings. In the construction of the Shewhart chart it is therefore required that measurements be grouped into subgroups. The size of which needs to be specified upfront. Leavenworth and Grant (2000) provide some guidelines and industry norms for the selection of subgroup sizes. In Figure 1.1, the xbar chart, each point represents the mean of five inside diameter piston ring measurements. While in Figure 1.2, the R chart, each point represents the difference between the largest and smallest value of five measurements i.e., the range of the subgroup of measurements. The top horizontal broken line, in Figures 1.1 and 1.2, labelled UCL, represents the upper control limit and the bottom broken line, labelled LCL, represents the lower control limit. For practical considerations it can be assumed that the process is displaying a constant source of variability, common to the process, if the subgroup means are observed inside these bounds. The solid horizontal line in the center of the charts, labelled CL, is called the center line. The CL line, in Figure 1.1, is the mean of all the subgroup means. In

Figure 1.2 the CL line is calculated as the mean of all the subgroup ranges.

In Figures 1.1 and 1.2 the first 25 subgroups are identified as calibration data while the remaining 15 subgroups are referred to as the new data. The data in the calibration set are used, in this example, to statistically quantify the common cause variability of the forging process i.e., to calculate the lines CL, LCL and UCL for each chart. A normal distribution is assumed in specifying the lines on the control chart. Therefore, the calibration information is used to approximate a normal distribution through the estimation of the mean and standard deviation. The mean constitutes the CL line in the xbar chart, while, leveraging the assumption of normality, the standard deviation is estimated using the data range information of the calibration set. The control limits are functions of the calibration mean and standard deviation e.g., the xbar chart limits are $UCL = CL + 3 \times (standard\,deviation)$ and $LCL = CL - 3 \times (standard\,deviation)$. The Shewhart control charts can now be used to determine if the variability observed in future diameter measurements exhibit variability that is indigenous to the forging process. In this example the new data plays the role of future observed data. From Figure 1.1 it can be observed that the penultimate three points of the new data set display uncommon variability for which an assignable cause should be investigated. The R chart, Figure 1.2, does not indicate any deviation from common cause variability.

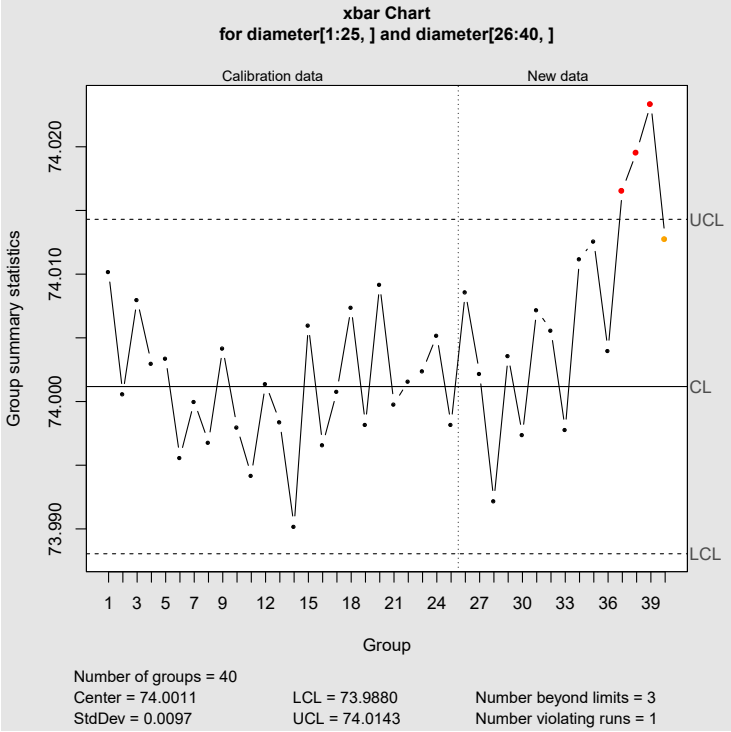


Figure 1.1: Example of Shewharts X bar chart.

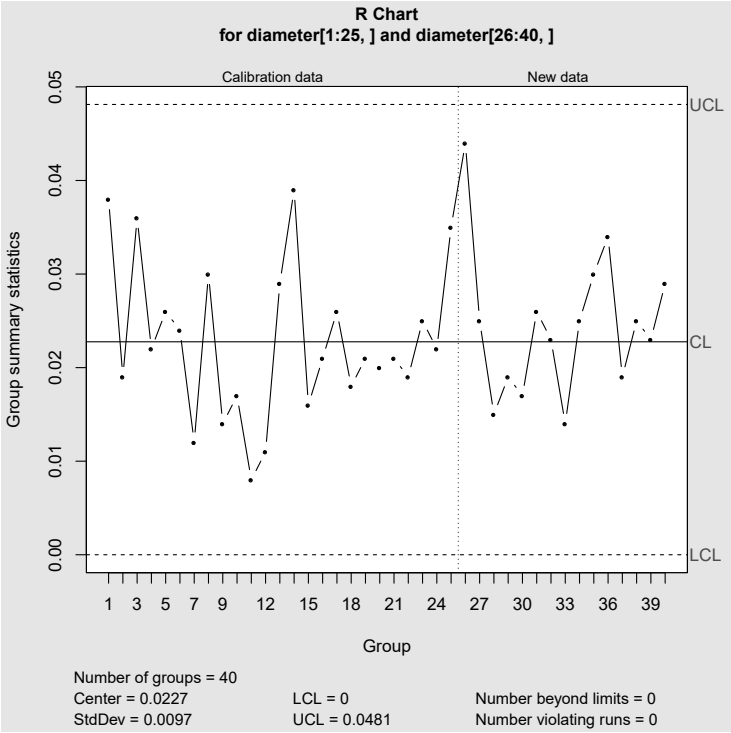


Figure 1.2: Example of Shewharts R chart.

It should be noted that the decision rules, that accompany the Shewhart charts, which are used to determine if the process is out-of-control are not limited to identifying if the process is operating beyond the upper and lower control limits. Various other rules have been developed to identify if assignable cause variability is present in a system. These rules, first introduced by Nelson (1984) are referred to as 'Nelson rules' or 'run rules'. A central objective of these rules is to interrogate the point to point variability in the Shewart charts to identify if the process is displaying non-random behaviour. Some example rules are:

- A specified number (e.g. seven or more) of points is observed in succession on the same side of the central line. This could indicate a bias in the measurements.
- Points are observed in sequence to continuously increase or decrease. If six or more points have this behaviour then there is evidence of a undesirable trend.
- Values in the chart are observed to alternate in sequence between measurements that are above the CL line and below the CL line. Fourteen or more such consecutive points would indicate non-random oscillation.

An example of a run rule violation is indicated by the yellow point in Figure 1.1. This specific violation has been triggered because seven consecutive points are observed above the CL line.

The information captured by the Shewhart chart is therefore numerous and valuable to engineers for process optimisation and understanding.

1.2.1 Univariate fault diagnosis

It is required that an indication be given to the engineer as to which process characteristics he or she needs to focus on, if the Shewhart charts indicate that the process is statistically out-of-control i.e., that assignable cause variability is present. The answer is rather obvious in SPC, since only one variable is monitored in the Shewhart charts. The engineer should apply process understanding to determine how to get the deviating plant characteristic back in-control.

It is often the case in practice that multiple process variables are monitored simultaneously using the SPC method. The possibility therefore exists that a situation might arise where more than one of these charts will flag that the specific characteristic it is monitoring is out-of-control. A plausible question that the engineer may ask is: which one of the deviations is more severe? Alternatively, it might be of interest to know if there is a ranking that can be specified in which the out-of-control variables should be considered. These type of questions have not been thoroughly addressed in the SPC literature. A rather naive suggestion, if one of the control limits have been exceeded for a process variable x_i, would be to rank the variables according to their deviating percentage:

$$100 * \frac{|f_{spc}(x_i) - L_f|}{UCL_f - LCL_f}, \tag{1.1}$$

where $f_{spc}(x_i)$ is the control chart evaluation of the specific measurement on process variable x_i and with UCL_f and LCL_f the upper and lower control limits of the chart respectively. The variable L_f represents the control limit closest to $f_{spc}(x_i)$ i.e., it is equal to UCL_f or LCL_f depending on whether the observed value is greater than UCL_f or smaller than LCL_f respectively.

Therefore, relative to the normal operating range, equation (1.1) quantifies the magnitude of the deviation for each process variable x_i. The engineer would need to combine this ranking of variables with process understanding. He or she needs to be cognisant that a small percent deviation in one process characteristic might be more important than a larger percentage deviation in another. Weights may be applied to the variables in this situation. The statistical ranking should only be used as a guide. It is not obvious how one would rank the importance of variables when more than one control chart indicate a run rule violation. Further research is therefore required to address the question of ranking variables in univariate SPC for diagnostic analysis.

The SPC methodology consists of a simple yet powerful set of statistical tools. It should be emphasised that prior to its introduction, no model based approach existed in manufacturing to objectively monitor the performance of processes. Many successful contributions of SPC to the manufacturing process have been recorded previously, most notably the big impact that SPC have had on the Japanese industrial sector (Leavenworth and Grant, 2000). Regardless of the many successes, it is the opinion of the author that the most important contribution made by Shewhart is the introduction of statistical thinking in process monitoring.

1.3 Multivariate statistical process monitoring

Multivariate statistical process monitoring was introduced to address limitations in the univariate SPC approach for SPM (MacGregor and Kourti, 1995). Due to the inter-dependent nature of industrial processes, it is oftentimes required that multiple process characteristics simultaneously adhere to certain conditions, in addition to their univariate limits, to ensure that process optimality is achieved. This multivariate nature of manufacturing plants is not accounted for by univariate SPC as is highlighted by MacGregor and Kourti (1995). It is therefore argued by MacGregor and Kourti (1995) that as a result of deviations in the relationships between variables it is possible for the process to yield sub-optimal results. Therefore, the limitations of univariate SPC together with the solutions provided by the MSPM methodology will be presented in this section.

1.3.1 Simulation study setup

A simulation study, building on the motivation provided by MacGregor and Kourti (1995) will now be discussed. In addition, computer code from the R programming language (R Core Team, 2018) will be provided to illustrate the simulation. It should be noted that, it is not known to the author, if a simulation study has explicitly been reported previously which evaluates the performance of SPC in the multivariate setting. The motivation of MacGregor and Kourti (1995) only provides an intuitive understanding of why it can be expected that the fault detection ability of univariate SPC methods will be suboptimal. Here the argument of MacGregor and Kourti (1995) is therefore extended to illustrate, in no uncertain terms, the shortcomings of SPC.

In this simulation example a fictitious manufacturing process will be considered. It will be assumed that, for this process, there exist two quality characteristics that need to be monitored in order to ensure that the process is operating optimally. These two process variables will be labelled *var1* and *var2* respectively. It is known from the fundamental process understanding that *var1* and *var2* are positively correlated. Using historical

process information, it is found that the magnitude of the correlation amounts to approximately 0.9. Furthermore, the bivariate distribution of the process characteristics is a multivariate normal (MVN) distribution,

$$MVN\Big(\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 2.00 & 2.26 \\ 2.26 & 4.00 \end{bmatrix}\Big). \tag{1.2}$$

The R code, used to simulate 250 observations, representing typical operational data from this imaginary process is as follows:

```
library(MASS)
xmean1 <- c(0,0) #bivariate mean
p <- 0.9 #correlation between variables
s11 <- 2
s22 <- 4
s12 <- p * sqrt(s11)*sqrt(s22)
s21 <- s12
covmat <- matrix(c(s11,s21,s12,s22),2,2) #covariance matrix
N <- 250 #number of samples
#simulate from multivariate normal distribution
X1 <- mvrnorm(n=N,mu=xmean1,Sigma=covmat)
colnames(X1) <- c("var1","var2") #label variables
```

The data set, $\mathbf{X}_1 : 250 \times 2$, represents the reference operating conditions and therefore captures the common cause variability inherent to the process. The black dots in Figure 1.3 depict the simulated data for $var1$ and $var2$. The vertical and horizontal broken lines represent the univariate bounds in which each variable is expected to operate. Also included in Figure 1.3 are contours representing the $(1-\alpha)\%$ confidence ellipses of (1.2), the distribution used to generate the typical operational data vectors. These contour lines, therefore, give a general sense of where one can expect observations on $var1$ and $var2$ to lie. It is clear that, while the process is operating inside the univariate limits, if an observation is recorded in the 'white-space' surrounding the specified $(1-\alpha)\%$ confidence ellipse, it can be assumed that assignable cause variability is present. An example of an atypical observation of the simulated plant is depicted in Figure 1.3 by the crossed red circle $\oplus$.

The objective of the current simulation is to generate new samples that are foreign to the process and to evaluate the behaviour of the SPC charts. Various changes can be made to the underlying mechanism (1.2) used to generate the normal operating characteristics that will give rise to measurements on $var1$ and $var2$ to occupy the low density 'white-

space' regions.

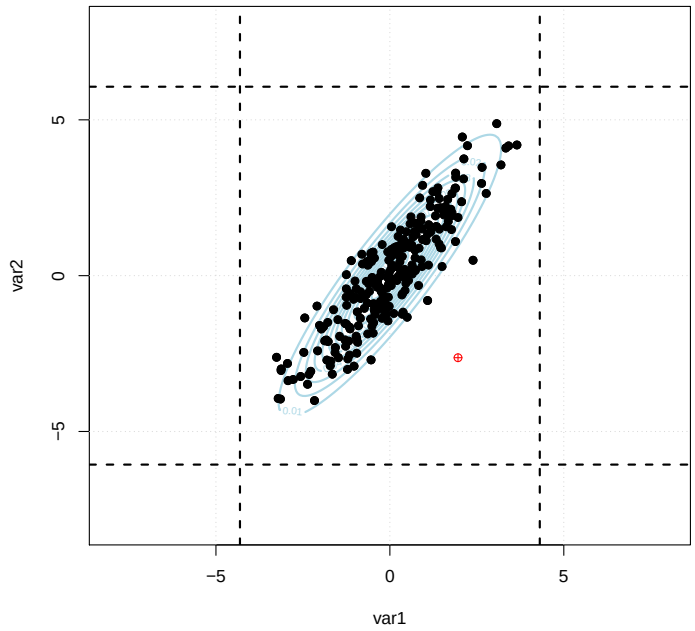


Figure 1.3: Bivariate scatter plot of $var1$ versus $var2$ under normal operating conditions (solid black points). The crossed red circle represents an atypical observation.

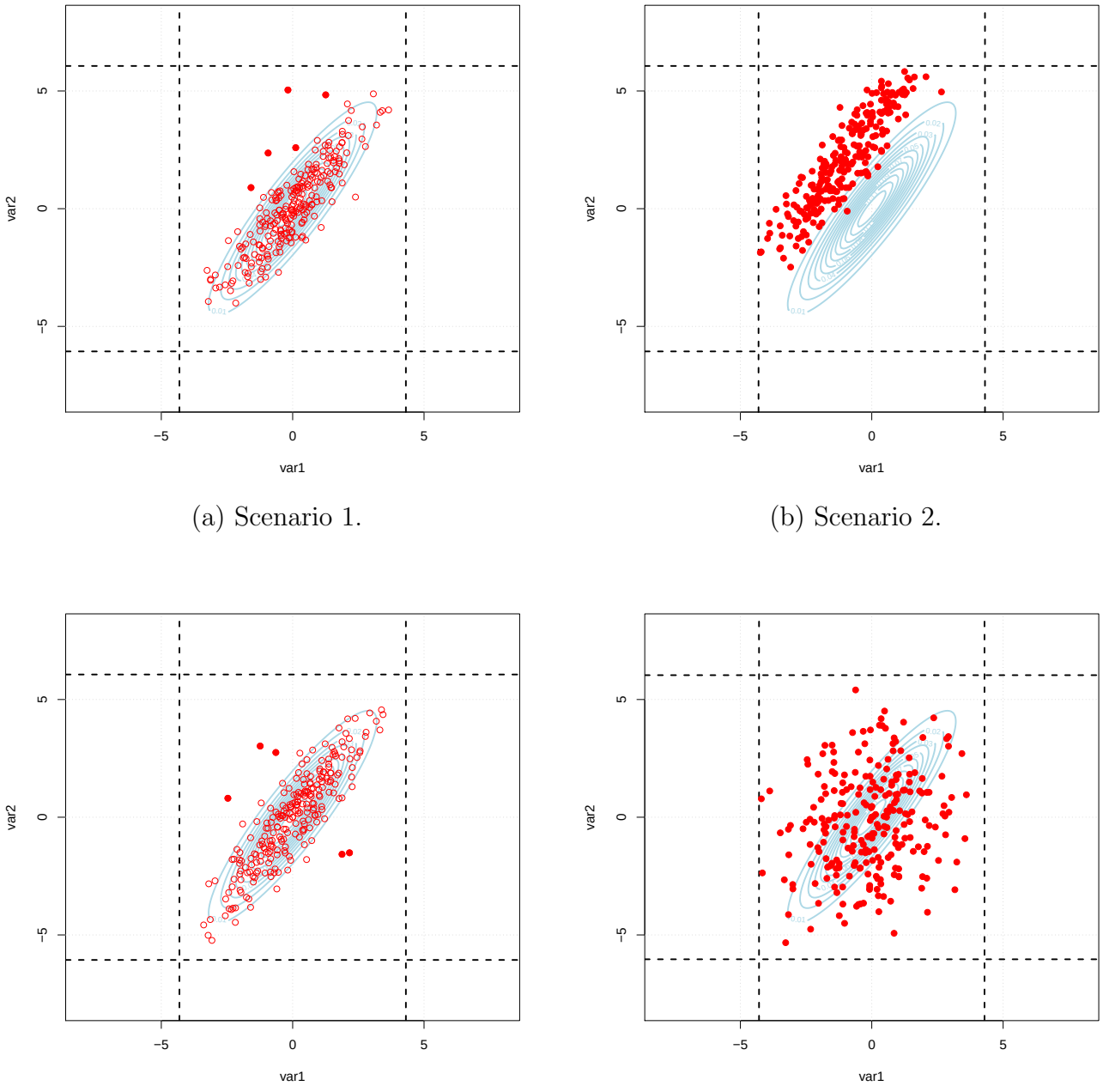
Specifically, the following scenarios will be explored:

Scenario 1 In the multivariate normal distribution (1.2) random changes in its mean vector are introduced. Observations from this modified distribution are made randomly over time. The process therefore does not completely shift to the new distribution i.e., the assignable cause variation is not present continuously but occurs at random over time.

Scenario 2 In the multivariate normal distribution (1.2) a sudden delta change in its mean vector is introduced. Observations from this modified distribution are made consistently over time. The process therefore completely shifts to the new distribution at a specified point in time.

Scenario 3 In the multivariate normal distribution (1.2) random changes in its covariance matrix are introduced. Observations from this modified distribution are made randomly over time. The process therefore does not completely shift to the new distribution i.e. the assignable cause variation is not present continuously but occurs at random over time.

Scenario 4 In the multivariate normal distribution (1.2) a sudden delta change in its covariance matrix is introduced. Observations from this modified distribution are made consistently over time. The process therefore completely shifts to the new distribution at a specified point in time.



(a) Scenario 1.

(b) Scenario 2.

(c) Scenario 3.

(d) Scenario 4.

Figure 1.4: Data observed on $var1$ and $var2$ for the different simulated scenarios.

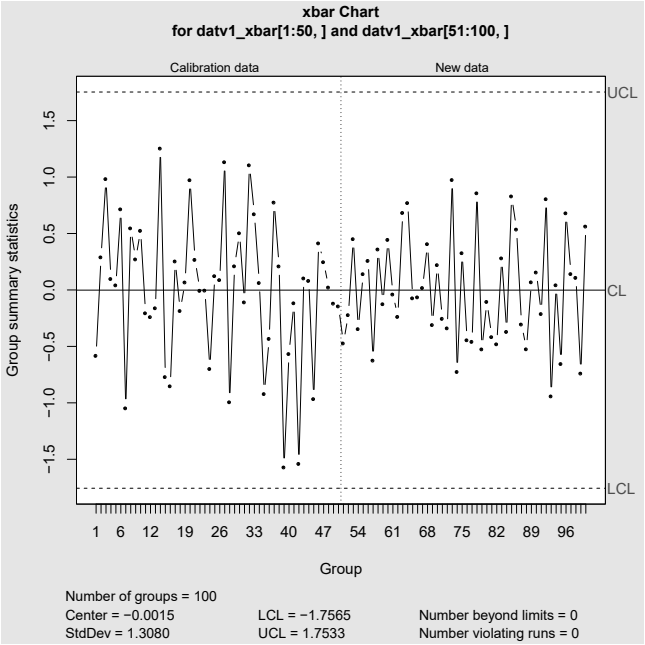
1.3.2 Univariate SPC

In the first two scenarios the change in mean resulted from the following multivariate normal distribution:

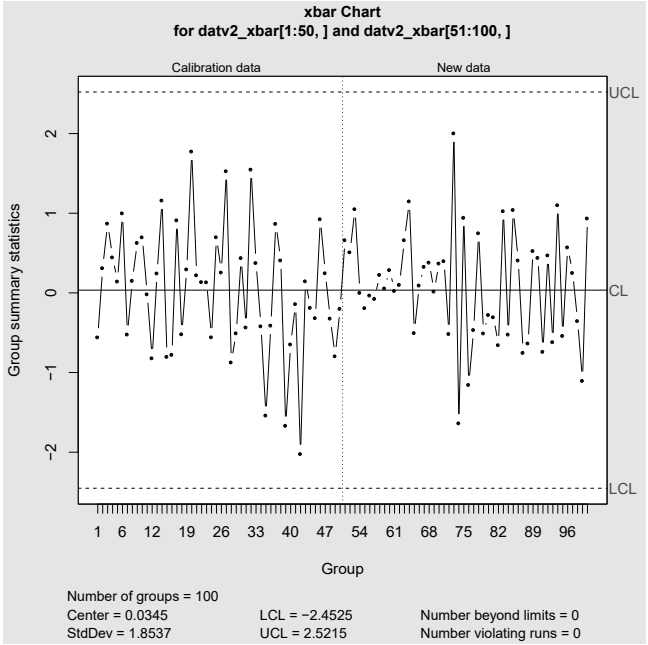
$$MVN\Big(\boldsymbol{\mu} = \begin{bmatrix} -1.02 \\ 2.08 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 2.00 & 2.26 \\ 2.26 & 4.00 \end{bmatrix}\Big). \tag{1.3}$$

For the first scenario the situation where the new samples are randomly observed from distribution (1.3) is considered. However, the majority of the samples have variability that is common to the process. To achieve this, 250 samples were drawn using both the reference distribution (1.2) and the mean shifted distribution (1.3). Specifically, 245 samples were drawn from (1.2) and five from (1.3). The positions in time where the five samples from (1.3) were observed in the sequence of 250 samples were selected at random with the constraint that there should at least be a five time point separation between them. This is to ensure that the observations from (1.3) are not recorded within the same subgroup. The top left graph in Figure 1.4 is a bivariate plot of the simulated data displayed as red circles for Scenario 1, together with the blue contours that specify the common cause variability contained in the 95% confidence region. The five observations drawn from (1.3) are represented as solid circles, while empty circles are used to identify the samples produced using (1.2). As expected, it is clear that the samples simulated randomly over time from the mean shifted distribution are located in the low density white space region. The SPC results for this scenario are illustrated by the xbar and R-chart graphs in Figure 1.5. Note the control limits are specified using the univariate SPC rules for the data from (1.2). It is clear from these graphs that the SPC analysis did not detect any measurement that can be considered to be different from normal operation or common cause variability.

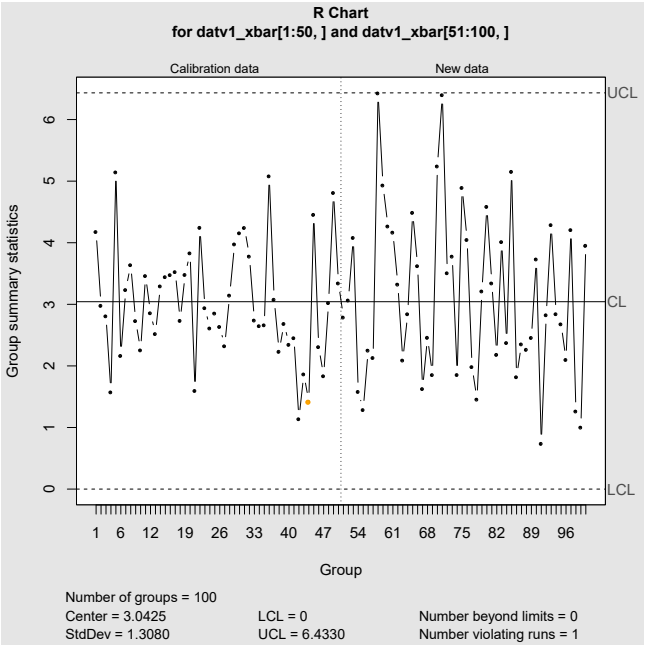
The second scenario is also based on the mean shifted distribution used in the first i.e., (1.3). It will be assumed, in this setting, that the process completely shifts and produce data that are generated by (1.3) at a specified point in time. Therefore, the new 250 data points are expected to have assignable cause variability. In Figure 1.4, the data simulated for this case study are represented by the top right graph. Solid red circles are again used to represent the new data simulated from (1.3). As constructed, it is observed that most of the new samples are located in the assignable cause white space region. Figure 1.6 reports the results of the SPC analysis. Although all of the new samples are drawn from (1.3), very few samples are observed beyond the control limits on all of the univariate control graphs.



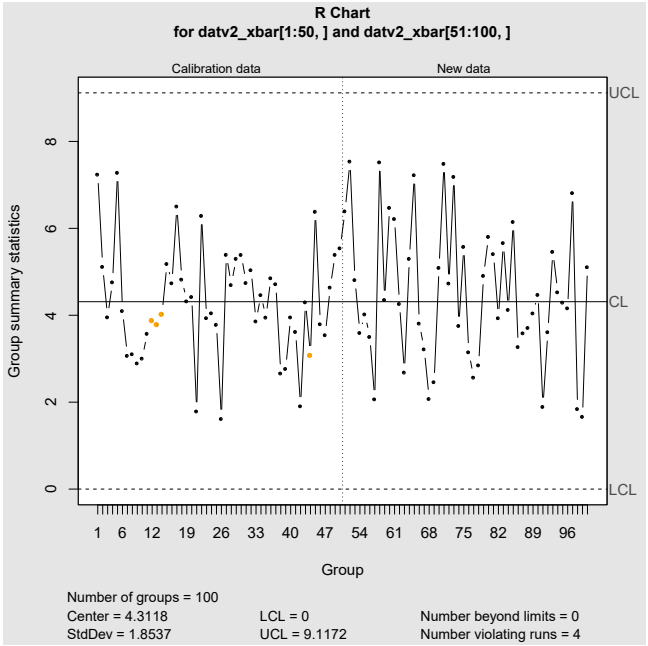
(a) Xbar chart of $var1$.



(b) Xbar chart of $var2$.

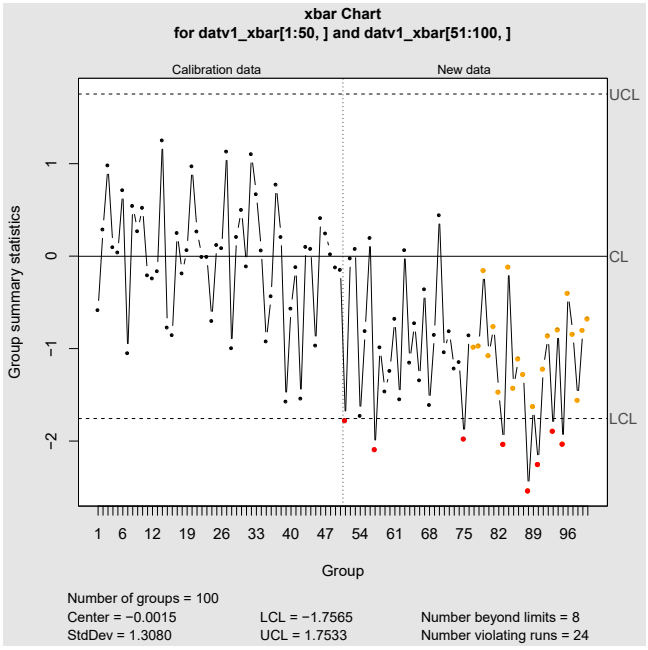


(c) R chart of $var1$.

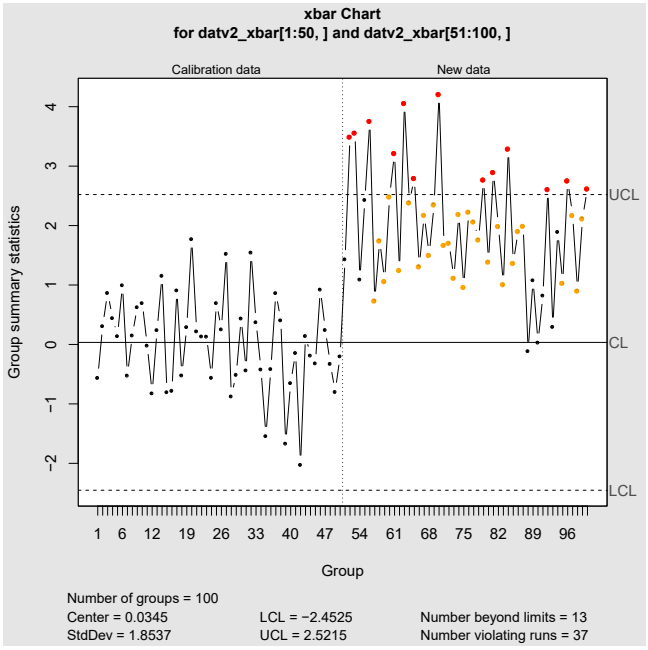


(d) R chart of $var2$.

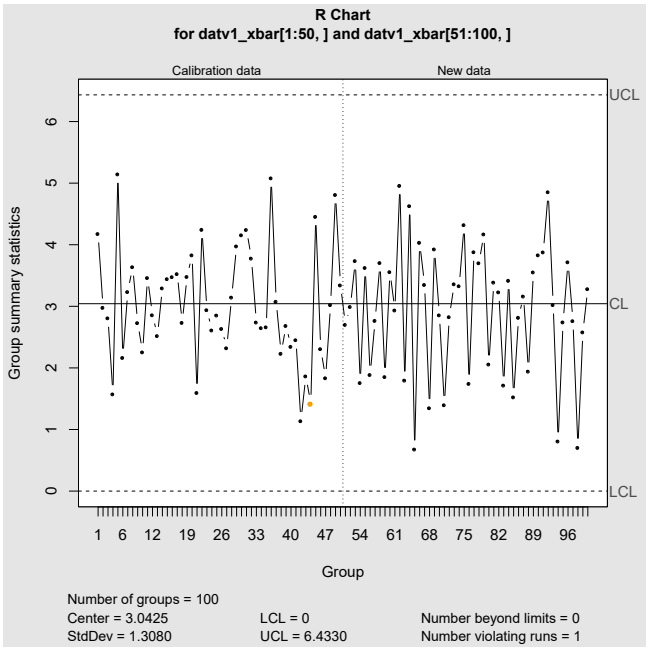
Figure 1.5: SPC analysis of samples generated in Scenario 1. Vertical dotted lines indicate the point in time when the new data are collected.



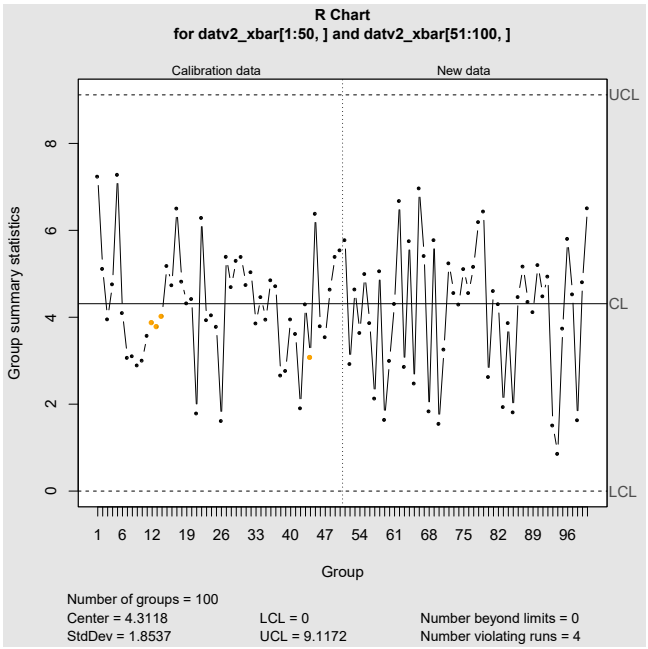
(a) Xbar chart of $var1$.



(b) Xbar chart of $var2$.



(c) R chart of $var1$

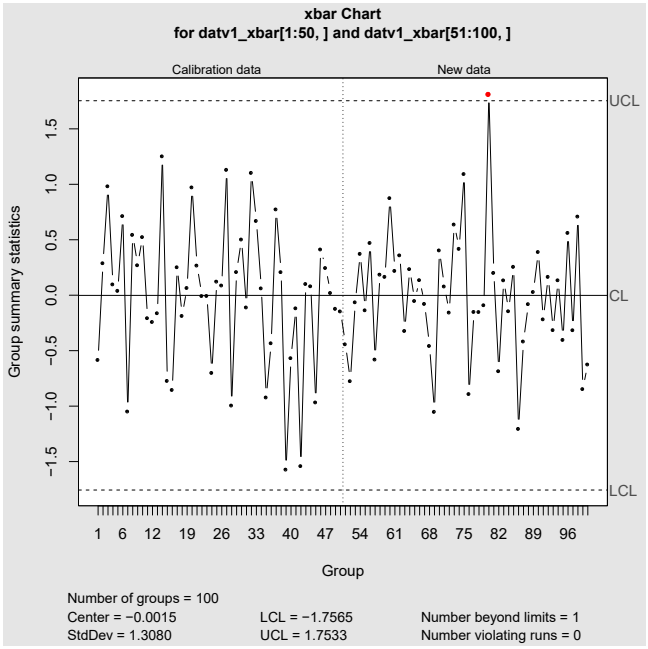


(d) R chart of $var2$

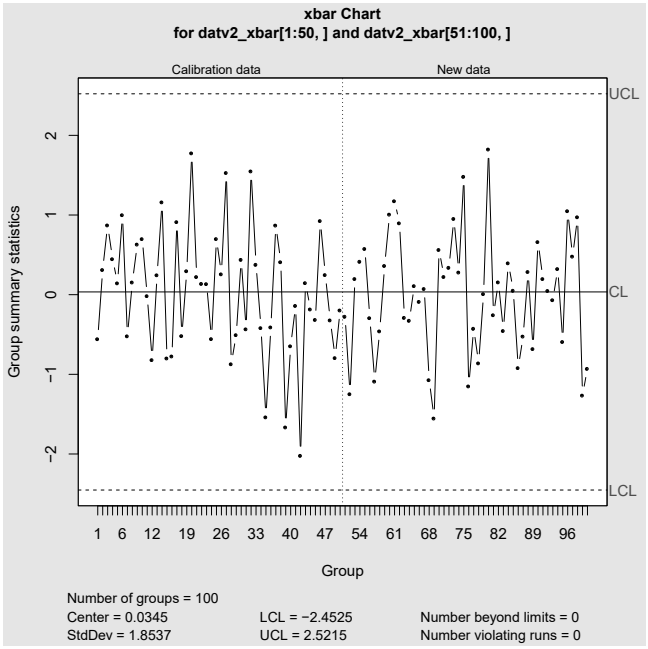
Figure 1.6: SPC analysis of samples generated in Scenario 2. Vertical dotted lines indicate the point in time when the new data are collected.

It is striking to note the characteristic change exhibited by the xbar charts for both of the process variables. In both graphs the subgroup means are clustered on one side of the central line. The xbar chart run rules are also triggered for most of those observations. Therefore, even though few observations are found to violate the control limits, a clear shift is observed in the xbar charts. The change in the grand mean of the distribution

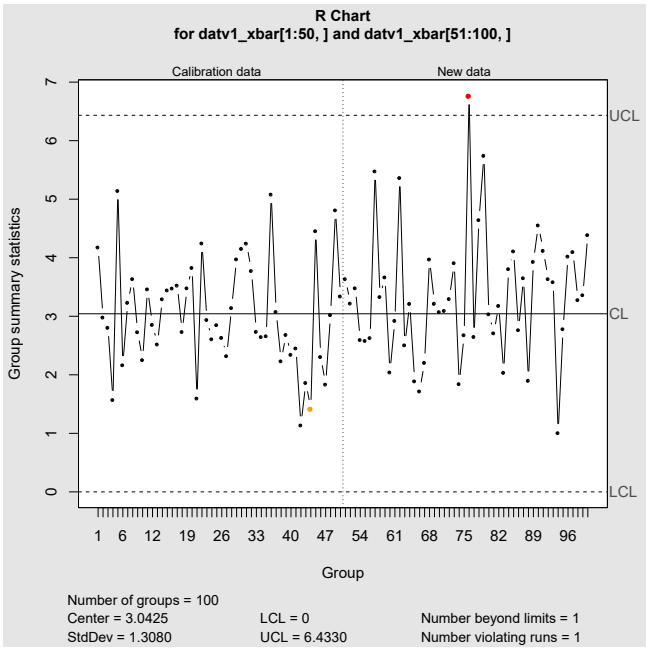
used to generate the process variability has therefore, to some degree, been detected by the SPC analysis. Note the R-chart remains in-control since the variance did not change.



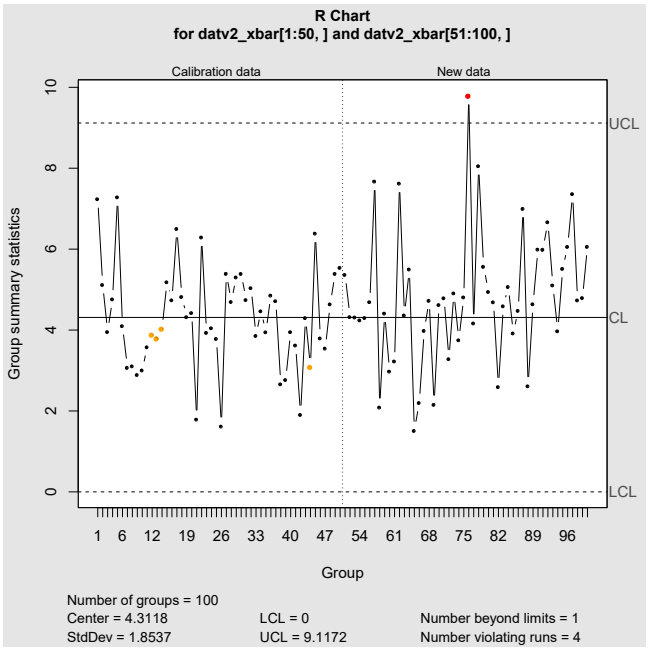
(a) Xbar chart of $var1$.



(b) Xbar chart of $var2$.

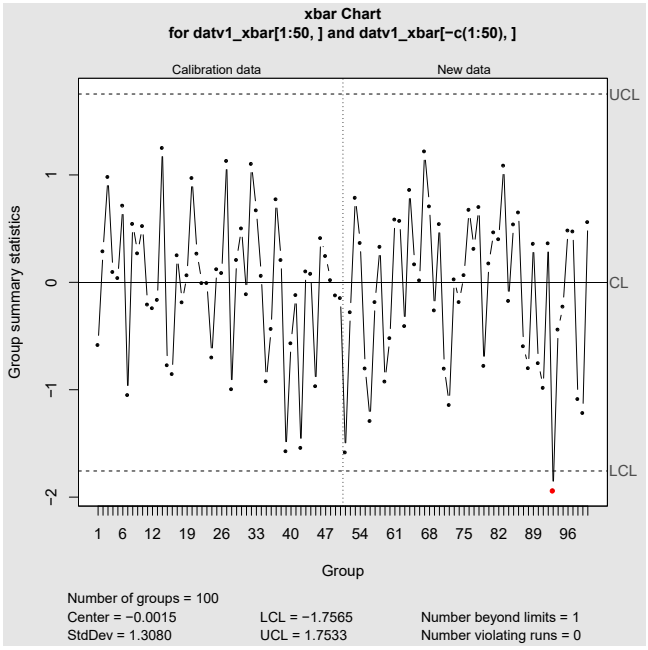


(c) R chart of $var1$.

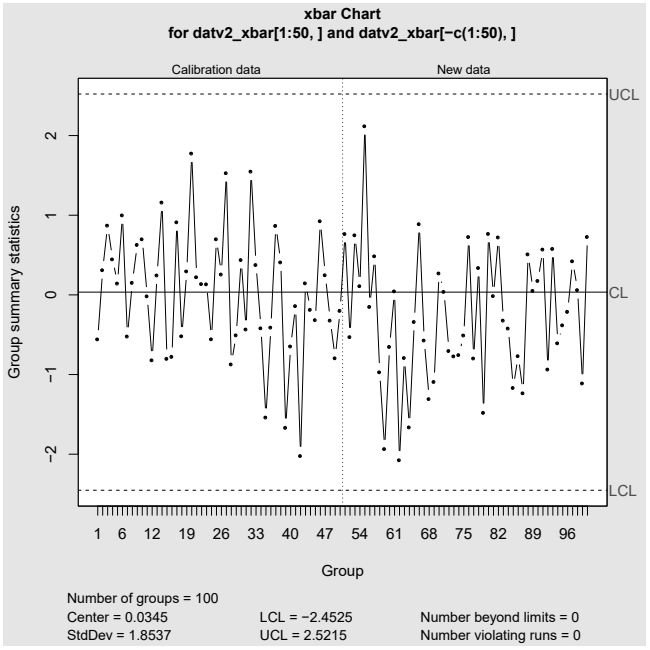


(d) R chart of $var2$.

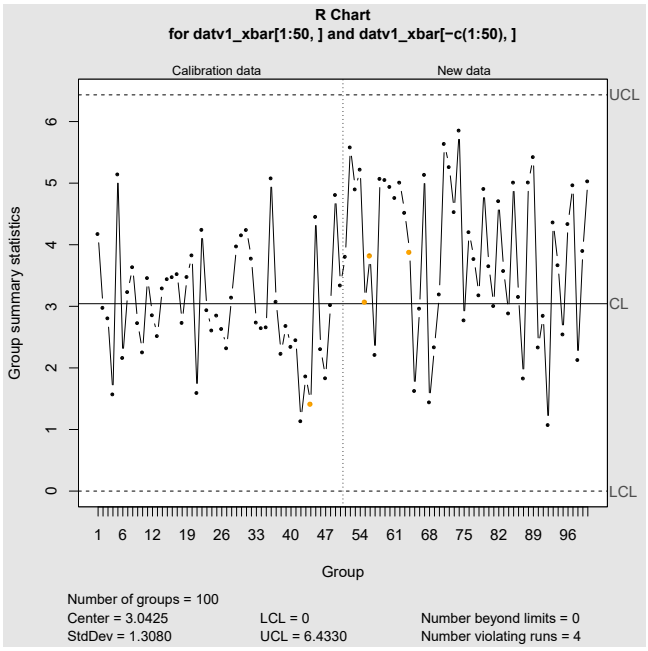
Figure 1.7: SPC analysis of samples generated for Scenario 3. Vertical dotted lines indicate the point in time when the new data are collected.



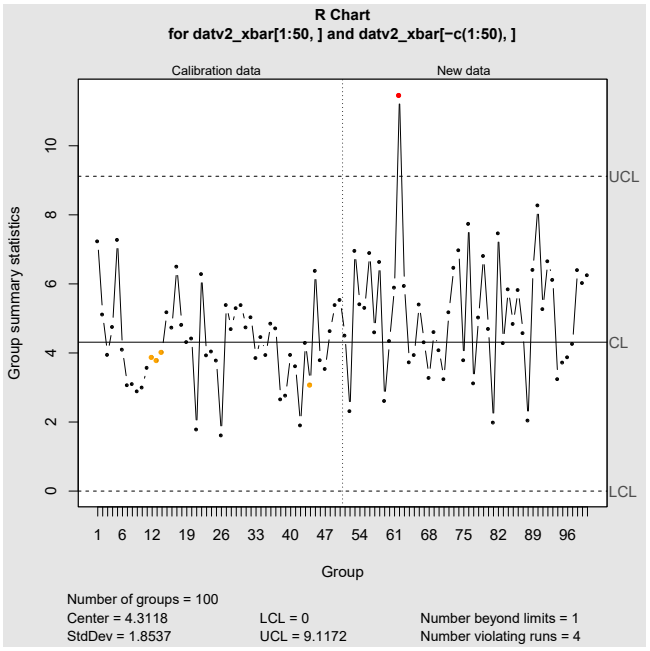
(a) Xbar chart of $var1$.



(b) Xbar chart of $var2$.



(c) R chart of $var1$.



(d) R chart of $var2$.

Figure 1.8: SPC analysis of samples generated for Scenario 4. Vertical dotted lines indicate the point in time when the new data are collected.

In the third and fourth scenarios, the change in covariance matrix resulted from the following multivariate normal distribution:

$$MVN\Big(\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 2.00 & 0.28 \\ 0.28 & 4.00 \end{bmatrix}\Big). \tag{1.4}$$

Similar to the first scenario, the third scenario attempts to mimic the situation when the process produces samples from (1.4) at random over time. The specific procedure used to achieve this is therefore the same as that used for the first scenario with the only difference being that (1.4) is used instead of (1.3) to simulate the new observations. The simulated example is displayed in the bottom left graph of Figure 1.4. It is again noted that the five samples simulated from (1.4) do occupy the white space region. Figure 1.7 depicts the results of the SPC analysis. It is again clear that the univariate analysis does not alert us that five of the samples were drawn from a distribution that is different to the one used to simulate typical operation. It should be noted that in three of the graphs some of the subgroups are identified as being different. However, the particular subgroups are not representative of the simulated faulty samples.

For the fourth scenario the process in its entirety shifts to the covariance altered distribution (1.4). The simulated 250 samples are displayed in the bottom right graph of Figure 1.4. Since (1.4) has the same mean as (1.2), it is observed that many of the observations are contained inside the 95% confidence region of the training data. However, as a result of the modified covariance between the variables, it is ensured that many of the samples are recorded in the low density region. For this example it is important to note that the SPC analysis graphs in Figure 1.8 did not identify that the process generating mechanism has completely changed.

It is understood that although the simulation study presented here is very crude, valuable information is gathered to improve the understanding of the limitations of SPC in multivariate process monitoring. Therefore, a summary of the findings of the simulated scenarios are:

- SPC predominantly fails to detect assignable cause variability when the observations happened at random points in time. This holds true for each of the distributions tested i.e., for both the mean altered and covariance altered distributions. Refer back to scenarios 1 and 3.

- SPC was able to detect the multivariate change in the underlying mechanism when the mean was shifted at a specified point in time. The xbar charts displayed drifting behaviour inside the univariate control limits. Xbar chart run rules were triggered for many of those samples. Refer back to scenario 2. Personal experience in the chemical manufacturing industry informs us that the type of process change simulated here

is the most commonly observed.

- SPC was not capable of alerting us to the multivariate change in the process simulating distribution when the covariance matrix was changed. Refer back to scenario 4.

By means of simulation it was clearly illustrated how, in some instances, the SPC methodology might not be suitable to detect multivariate process deviations. It is informative to note that the SPC approach is not completely void of usefulness in the multivariate setting as was demonstrated by the second scenario. An improved method of statistical process monitoring is therefore required to detect assignable cause variability in the multivariate setting.

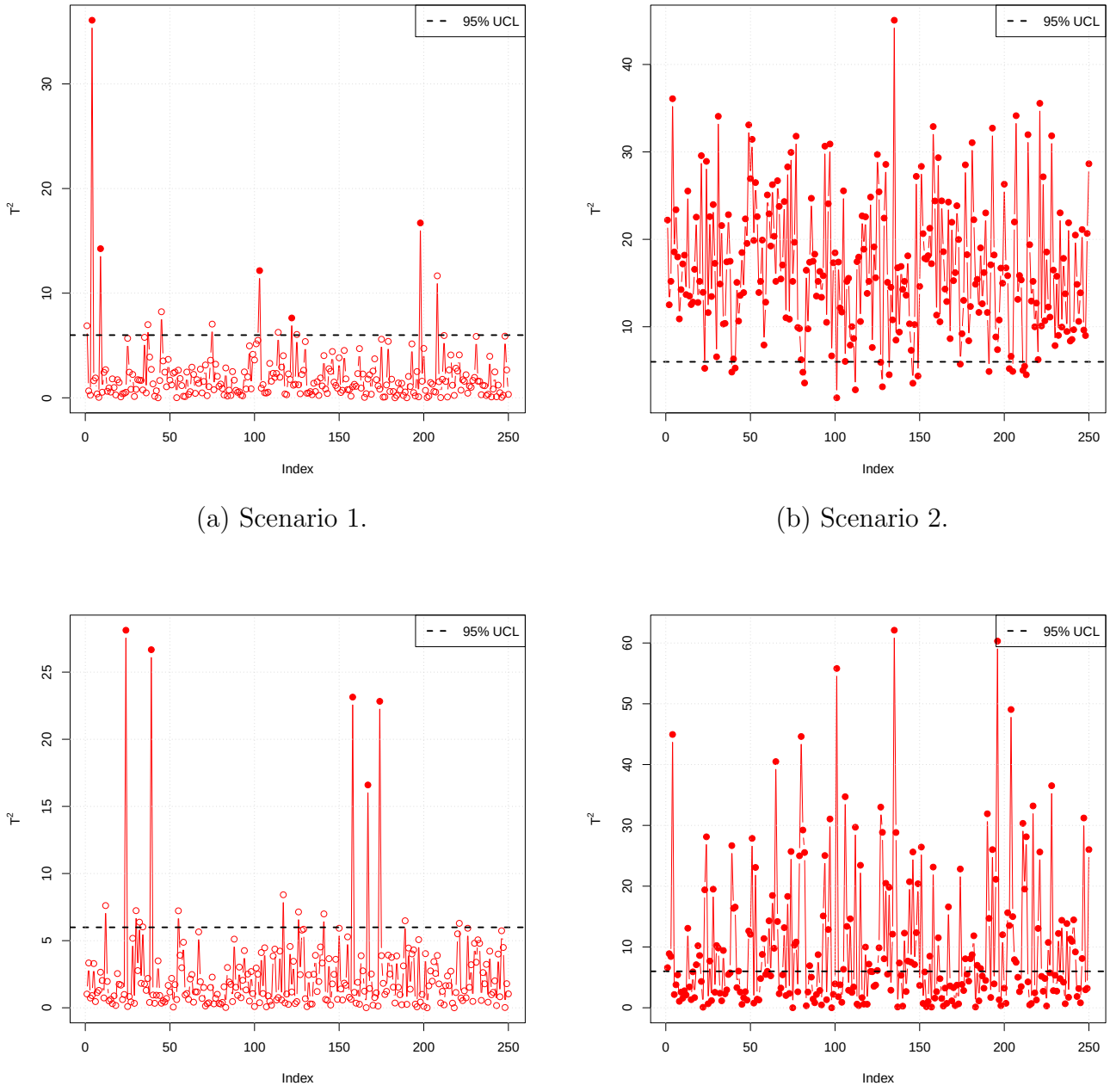
1.3.3 Multivariate statistics

Various statistical techniques have been proposed to account for the fault detection limitations inherent to the SPC techniques when applied in the multivariate process environment. The T^2 statistic developed by Hotelling (1947) is one method that can be employed as a natural extension of Shewhart's xbar chart. Suppose that for the manufacturing process considered there are p process variables that need to be monitored. Let the observations on these variables be recorded by the vector $\mathbf{x} : p \times 1 = \begin{bmatrix} x_1 & x_2 & \dots & x_p \end{bmatrix}^T$. Assume that under common cause conditions the process variability can be described by a multivariate normal distribution with mean $\boldsymbol{\mu} : p \times 1$ and covariance $\boldsymbol{\Sigma} : p \times p$. The T^2 statistic evaluation of a new observation, $\mathbf{x}^{new} : p \times 1$, from the process is then defined as

$$T^2(\mathbf{x}^{new}) \quad = \quad (\mathbf{x}^{new} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}^{new} - \boldsymbol{\mu}). \tag{1.5}$$

During normal operating conditions the $T^2(\mathbf{x}^{new})$ evaluations will follow a central χ^2 distribution with p degrees of freedom. Special or assignable cause variability will be present when the T^2 values are calculated to lie in the tail areas of this distribution. In a similar fashion to the xbar chart the T^2 control chart is presented as a graph showing the calculated $T^2(\mathbf{x}^{new})$ values over time. Also included are horizontal control limit lines, where the LCL = 0 and the UCL = $\chi^2_\alpha(p)$ with α an appropriately selected significance level. Note that in practical applications the parameters of the underlying multivariate normal distribution i.e., $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ would not be known and should be estimated using historical data common to the process. This will result in the $\chi^2(p)$ distribution used to describe T^2 values under normal conditions being replaced by an F-distribution. The data of Figure 1.4 that were simulated in the previous section will again be considered to illustrate the

ability of the T^2 statistic to detect special cause variability. Figure 1.9 depicts the results of the T^2 statistic analyses for all four scenarios.



(a) Scenario 1.

(b) Scenario 2.

(c) Scenario 3.

(d) Scenario 4.

Figure 1.9: T^2 statistic evaluations for the different simulated scenarios.

In each of the graphs the T^2 values of the newly generated samples can be read off on the y-axis. The order in which the observations were made is indicated by the x-axis. Solid points are again used to identify the observations that were generated by

a distribution that is different from the one common to the process. Note that each observation point represents a single observation and not a subgroup of points as in SPC. A broken horizontal line is used to display the UCL, which is calculated as the 95^{th} percentile of the χ^2_2 distribution. Examination of Figure 1.9 reveals that the points located in the white space regions of Figure 1.4 are mostly located beyond the upper control limit for scenarios one to four. The points that were simulated to represent samples that are inherent to the process are predominantly located below the UCL line. Therefore, by using the information that the process variables have a multivariate relationship, correct detection is enabled, using the T^2 statistic, when deviations occurred inside the univariate limits of the process. The T^2 statistic detects deviations in both the mean and covariance matrix. The latter being very important to detect changes in the correlation structure amongst many process variables.

In the above, the discussion of the T^2 statistic made use of an unusually small number of variables. Due to large numbers of variables up for consideration in most manufacturing applications, dimension reducing techniques are often utilized to reduce the variable count to a manageable few. These newly statistically calculated variables are referred to as latent or canonical variables and are calculated as weighted sums of the original process variables. Principal component analysis (PCA) is a popular dimension reduction technique employed in practice, see for example Qin (2012). Therefore, the latent variables are used instead of the original process variables to calculate (1.5). A high level summary of PCA based MSPM will now be provided. A more detailed discussion on PCA is given in Chapter 2.

Conveniently in the PCA setting, the T^2 calculation can be expressed, in a similar fashion to (1.5), as a function of the original process variables as

$$T^2(\mathbf{x}^{new}) \quad = \quad (\mathbf{x}^{new} - \boldsymbol{\mu}^{pca})^T (\boldsymbol{\Sigma}^{pca})^{-1} (\mathbf{x}^{new} - \boldsymbol{\mu}^{pca}), \tag{1.6}$$

with $\boldsymbol{\mu}^{pca}$: $p \times 1$ and $\boldsymbol{\Sigma}^{pca}$: $p \times p$ the PCA estimations of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively. A convention that is often followed for mathematical convenience is to pre-process the data such that $\boldsymbol{\mu}^{pca} = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix}^T$. This zero centering simplifies (1.6) to,

$$T^2(\mathbf{z}^{new}) \quad = \quad (\mathbf{z}^{new})^T (\boldsymbol{\Sigma}^{pca})^{-1} (\mathbf{z}^{new}), \tag{1.7}$$

where $\mathbf{z}$: $p \times 1$ represents the observation vector post preprocessing and $\boldsymbol{\Sigma}^{pca}$ is the corresponding covariance matrix. It should be noted that when PCA is employed, there exists another fault identification statistic called the Squared Prediction Error (SPE) (Qin, 2012) which is popular amongst statistical process monitoring practitioners. This

statistic is defined as the squared residual,

$$SPE(\mathbf{x}^{new}) \quad = \quad (\mathbf{x}^{new} - \hat{\mathbf{x}}^{pca})^T(\mathbf{x}^{new} - \hat{\mathbf{x}}^{pca}), \tag{1.8}$$

where $\hat{\mathbf{x}}^{pca} : p \times 1$ is the PCA approximation of $\mathbf{x}^{new} : p \times 1$. An intuitive interpretation of the SPE index is that the PCA model will be inaccurate in estimating samples that are different from the data on which it was build, and will therefore reflect as a high SPE value. The distributional properties of the SPE statistic is reported by Jackson and Mudholkar (1979).

1.3.4 Multivariate fault diagnosis

As a result of their design, multivariate indices, such as the T^2 statistic, are ideally suited to detect assignable cause variability when the relationship between process characteristics are different from the norm. However, in contrast to univariate SPC, the multivariate approach to SPM does not provide an obvious way of determining which characteristics of the process caused the plant to drift away from the expected. A diagnostic analysis, in addition to the multivariate fault detection step, is therefore required to assist in understanding as to why assignable cause variability is being detected. An ideal approach would be to construct a fault classification statistical model that relates process measurements to a library of previously observed fault classes. If a process deviation is observed, this model can then be used to identify the relevant fault class. However, this approach is highly dependent on having a history of all possible deviations that can occur. Developing a fault classification model is therefore not feasible in most practical cases.

Diagnostic approaches that consider the deviating process observation in isolation is therefore required. Contribution analysis, first introduced by Miller and Swanson (1998), is a commonly employed diagnostic methodology that interrogates the fault identification statistic of the violating sample in order to obtain an importance ranking of the process variables. The calculated relative ranking, or contribution, of the variables indicates the sequence in which the engineers should consider the variables in order to correct the abnormal process behaviour observed. This specific interrogation entails decomposing the fault identification statistic into a sum of parts. Each part of the decomposed statistic represents the contribution of a specific process variable to the faulty condition. There exist more than one way of factoring the T^2- and SPE statistics into different parts, as a consequence of their quadratic form. The contribution analysis methodology is therefore not uniquely defined for a specific statistic. A summary of the different methodologies can be found in Alcala and Qin (2011). Three popular decompositions are identified:

- Complete decomposition (CD) contributions (Miller and Swanson (1998) and Wise et al. (2006)),
- Partial decomposition (PD) contributions (Nomikos, 1997) and
- Reconstruction-based (RB) contributions (Alcala and Qin, 2009).

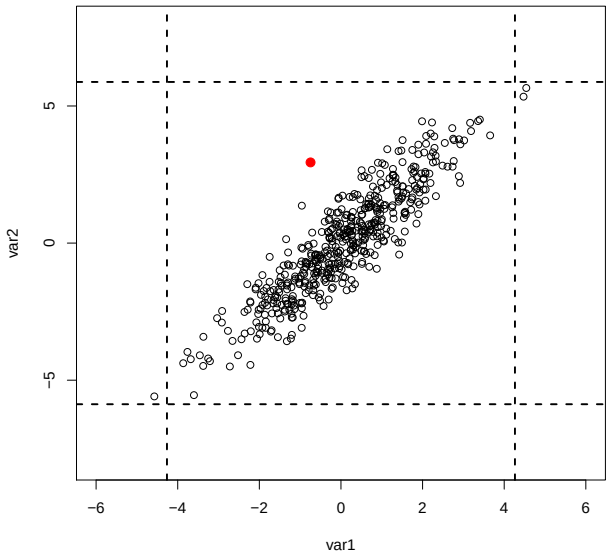
A detailed exploration of the different decomposition strategies will be presented in Chapter 2.

1.3.4.1 Contribution analysis simulation

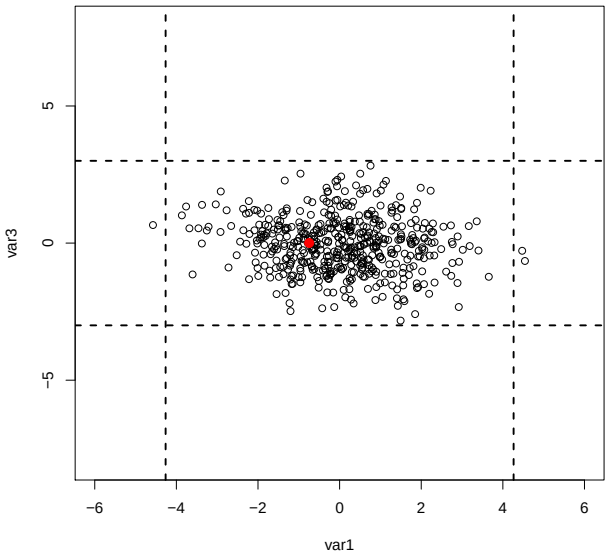
A typical scenario will now be simulated to illustrate how contribution analysis can be applied in practice. In this simulation a process will be considered with three process characteristics $var1, var2$ and $var3$ that need to be monitored simultaneously. Furthermore, it will be assumed that the common cause variability of the manufacturing process being monitored is described by the following distribution,

$$MVN\Big(\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 2.00 & 2.55 & -0.14 \\ 2.55 & 4.00 & 0.60 \\ -0.14 & 0.60 & 1.00 \end{bmatrix}\Big). \tag{1.9}$$

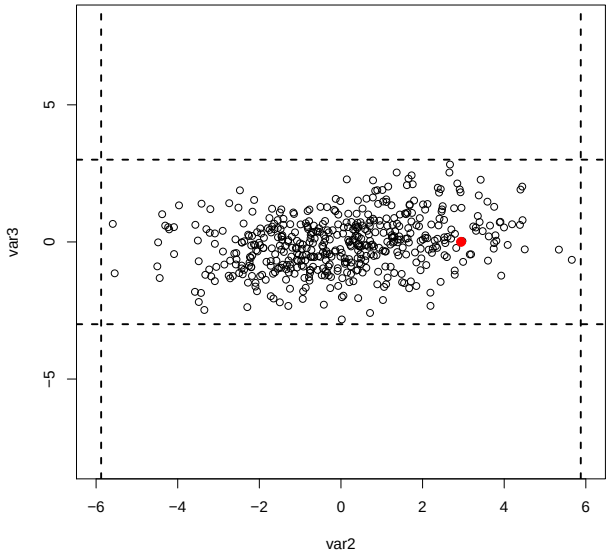
A total of 500 observations was simulated to represent typical historical operational measurements. It is difficult to visualise the data in three dimensions. Therefore, in Figure 1.10, the data is represented using pairwise scatterplots. In each of the graphs in Figure 1.10 the black circular points represent the historical common cause realisations. Dotted lines are used to indicate the univariate control limits. The regions defined by the black points are therefore where one would expect the process to operate. Following the discussions in the earlier sections, it is expected that if a measurement is made in the white space region, within the univariate control limits, of any of the pairwise plots that such an observation would be identified as being different. The solid red dot in Figure 1.10 is an example of one such observation. Visually, it clearly violates the common cause variability that describes the relationship between process variables $var1$ and $var2$. Multivariate fault detection statistics must be employed to determine if this observation statistically resembles an observation that is uncommon to the process.



(a) Process variable $var1$ versus $var2$.

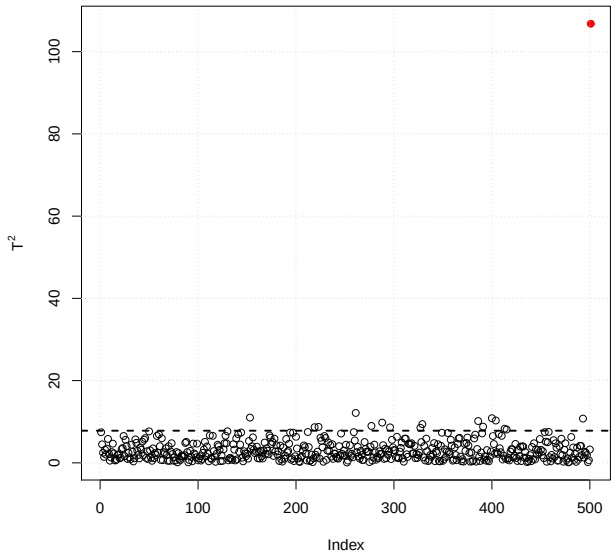


(b) Process variable $var1$ versus $var3$.

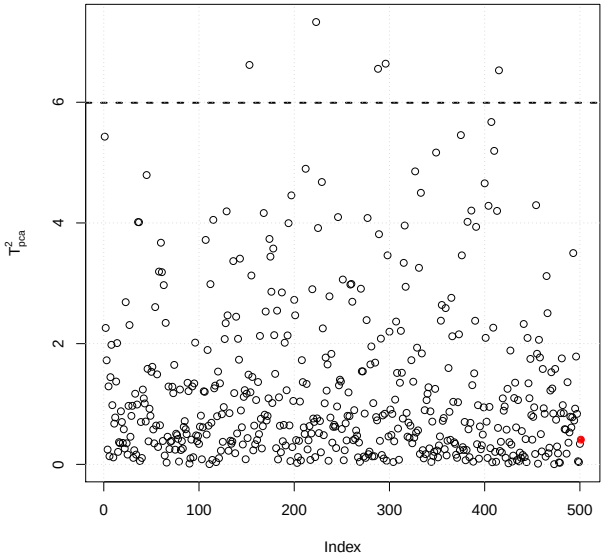


(c) Process variable $var2$ versus $var3$.

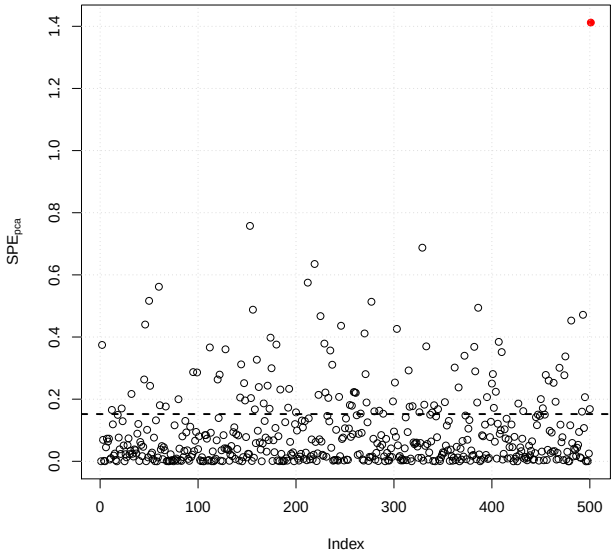
Figure 1.10: Simulated three dimensional observations from a fictitious process. Empty black circles represent observations made during normal operating conditions. Broken lines are used to indicate the univariate operating control limits. The red solid dot represents the values of a newly observed sample.



(a) T^2 statistic evaluation.

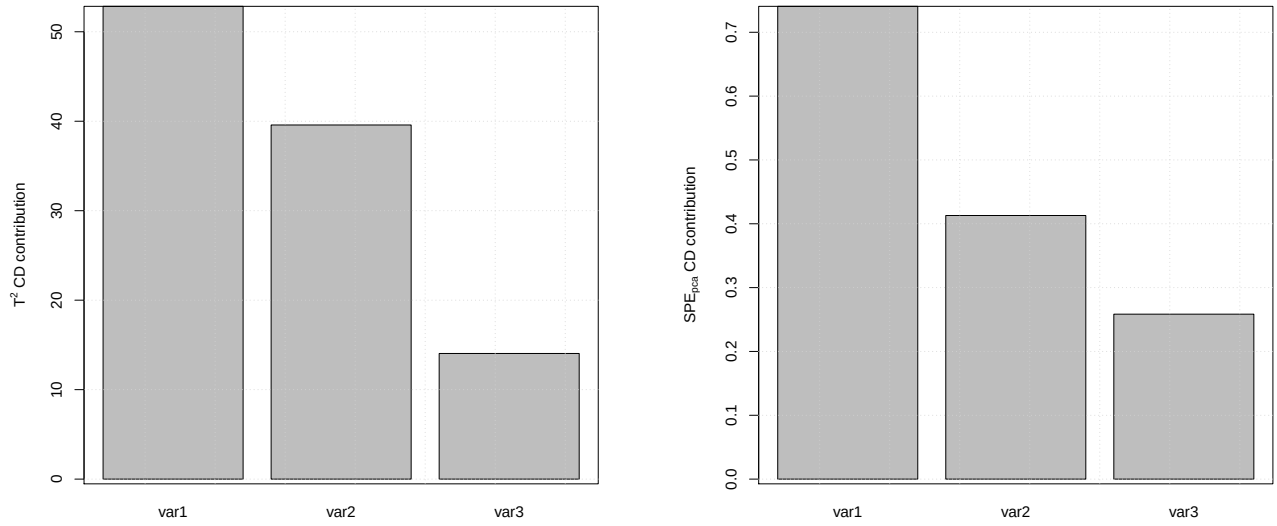


(b) PCA based T^2 statistic evaluation.



(c) PCA based SPE statistic evaluation.

Figure 1.11: Multivariate statistic evaluation of simulated data.



(a) CD contribution of T^2 statistic.

(b) CD contribution of PCA SPE statistic.

Figure 1.12: Complete decomposition contribution analysis based on out-of-control statistics of the simulated faulty observation.

The fault detection results are reported in Figure 1.11. Black points are used in each graph of Figure 1.11 to represent the multivariate statistic evaluations of the common cause variability observations. The corresponding evaluation of the solid red point observation in Figure 1.10 is again represented by a solid red point in Figure 1.11. Horizontal broken lines are used to specify the 95% upper control limit of each statistic. For this analysis three fault detection statistics are considered. The first statistic is the T^2 measure in (1.5) where the calculation is done in three dimensional space. The other two statistics, defined by (1.7) and (1.8), are based on a PCA of the common cause data. The specifics of the PCA results are beyond the scope of this chapter A detailed review of the PCA method will be given in Chapter 2. Two PCA latent variables were constructed which accounted for 99% of the common cause variability. Figure 1.11 indicates that the new observation is identified as being different by two of the statistics i.e., the T^2 statistic and the PCA based SPE statistic. Interestingly, the PCA based T^2 statistic does not identify the sample as having foreign variability.

A complete decomposition contribution analysis was done, based on the statistics that recorded an out-of-control evaluation of the new observation, to determine why the statistics flagged assignable cause variability to be present. The results of the contribution analysis are reported in Figure 1.12. A barplot is used to graphically represent the relative size of each process variable's contribution to the out-of-control statistic. It is clear from both graphs in Figure 1.12 that the contributions of $var1$ and $var2$ are the greatest.

The engineers should therefore investigate process variables $var1$ and $var2$ to determine what needs to done in order to get the process back in control. This result is in line with what is expected, since the new observation was simulated to be in the white space region of the $var1$ versus $var2$ scatterplot (see top left graph of Figure 1.10).

It has been argued in literature (see e.g., Van den Kerkhof, Vanlaer, Gins and Impe (2013)) that the variable contribution analysis method is not the ideal solution for fault diagnosis in multivariate statistical process monitoring. The most notable problem, first highlighted by Westerhuis et al. (2000), is that of fault smearing. The fault smearing phenomenon refers to the fact that individual variable contributions, as calculated by existing methods, do not truly capture the contribution of a specific variable. In the calculation of the contribution for a specific process variable information on other variables are included. This will result in variables that will appear to be important when they may not be. Fault smearing will be discussed in Chapter 2.

The specific problem of fault diagnosis in multivariate statistical process monitoring will be the main focus of this research document.

1.4 Nonlinear multivariate statistical process monitoring

It has been illustrated in the preceding section that MSPM provides considerably more information than univariate SPC when process monitoring requires that the multivariate relationships between variables should be considered. The SPC approach should not be ignored in favour of MSPM, since it may still provide valuable information. It should also be remembered that MSPM adds additional value when measurements are made inside the univariate bounds of the process variables (Kourti and Macgregor, 1995). The two statistical process monitoring techniques should therefore be used in combination.

Building on SPC, the MSPM methodology has introduced an important way of thinking about statistical process monitoring. Central to the methodology is the realisation that, if normally distributed variables are correlated then the measurements will be contained in an ellipsoidal region in p-dimensional space, see for example Figure 1.3. There is therefore a region in the p-dimensional hypercube, defined by the univariate limits of the process, where one would expect the observations to naturally cluster. Measurements that fall outside of the ellipsoidal region can therefore be considered to be the result of assignable cause variability being present. The MSPM fault detection statistics have therefore been developed to summarise the behaviour of the measurements that fall inside the common cause variability ellipsoidal region. If the MSPM fault identification statistic results in an extreme value then the process is identified as being out-of-control.

The assumptions made by the MSPM method restricts us to the consideration of data that are independent and identically distributed measurements from a multivariate normal distribution. This restriction implies that MSPM considers only common cause variability clusters that have an ellipsoidal shape. However, it is common in practical cases that the ideal operating region of the process is not ellipsoidal due to process characteristics not being normally distributed. For example, consider the ideal operating measurements on two simulated process variables, $var1$ and $var2$, which are represented by the circular black points in the top left graph of Figure 1.13. The common cause variability displayed by these two variables is therefore, visually, not shaped as an ellipse and therefore possibly not distributed multivariate normal. Also included are red solid points, generated to lie in the white space that do not conform to the shape that describes common variability. The red points are therefore simulated to have variability that are uncommon to normal or ideal operation.

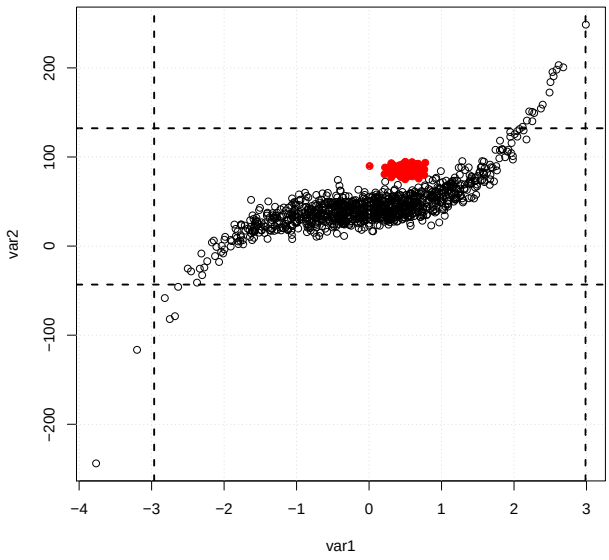
The specific manner in which the two simulated variables are related is as follows

$$var2 \quad = \quad 10 + 5(var1) + 5(var1)^2 + 6(var1)^3 + \epsilon, \tag{1.10}$$

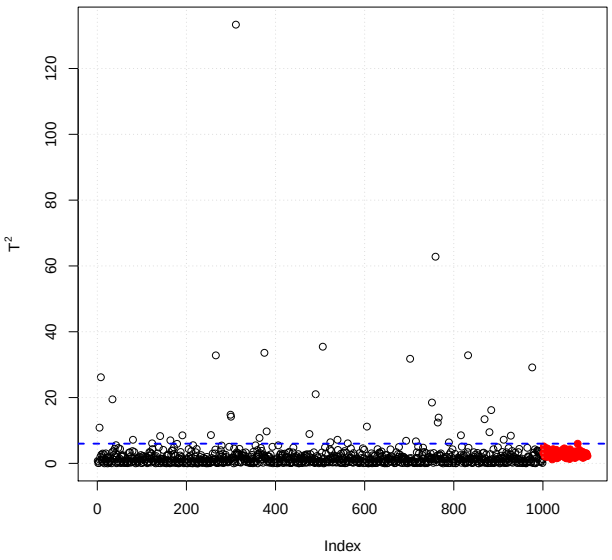
where $var1 \sim N(0, 1)$ and $\epsilon \sim N(30, 10)$. It should be emphasised that in practice, it is mostly not known what the mathematical relationships are amongst process variables. The assignable cause variability represented by the red points is simulated from

$$MVN\Big(\begin{bmatrix} 0.50 \\ 85.00 \end{bmatrix}, \begin{bmatrix} 0.02 & 0.00 \\ 0.00 & 20.00 \end{bmatrix} \Big). \tag{1.11}$$

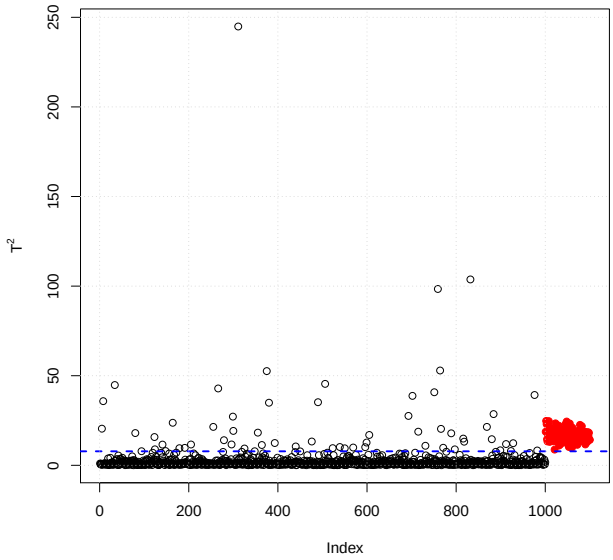
As it has been illustrated earlier, it is clear that univariate SPC will not be ideal to detect the deviations as simulated in this example, since the points are mostly contained inside the univariate limits of the process. Therefore, considering only the techniques discussed thus far, using a multivariate approach would be more appropriate. The T^2 statistic evaluations of the two dimensional observations in graph (a) of Figure 1.13 are represented by graph (b) of the same figure. A notable observation from this analysis is that none of the assignable cause samples, represented by the red solid circles, are detected by the T^2 statistic as being simulated from a distribution different from the one that governs normal operation. However, instead of basing the T^2 statistic calculation on the two dimensional vectors, i.e. $\begin{bmatrix} var1 & var2 \end{bmatrix}^T$, consider the situation where information of the non-linearity is included in the analysis. Specifically, if the higher dimensional vectors $\begin{bmatrix} var2 & var1^2 & var1^3 \end{bmatrix}^T$ are considered, it is observed that the resulting T^2 values indicate



(a) $var1$ versus $var2$.



(b) T^2 statistic evaluation.



(c) T^2 statistic evaluation of transformed data.

Figure 1.13: Nonlinear simulated data.

that the red points are indeed different from the black points as should be expected, see graph (c) of Figure 1.13. Therefore, by including information on the specific non-linearities the observations that fall in the assignable cause white space region are correctly detected.

The shortcomings of the multivariate analysis are overcome if knowledge of the non-linear relationship is incorporated. It is immediately clear that in practical cases it is not always known to process engineers what the specific mathematical forms of the underlying relationships between variables are. Different approaches have been proposed to account for non-linear relationships in SPM. A list of the different approaches can be found in He and Wang (2018). A popular methodology, that is of interest, uses kernel functions (Schölkopf et al., 1998) which implicitly projects the process vectors into a higher dimensional feature space. Multivariate statistical process monitoring is then performed in the high dimensional space, similar to the idea portrayed by the simulated example of Figure 1.13. Kernel PCA refers to the principal component analysis of the high dimensional vectors that were obtained using the kernel mapping. The MSPM techniques and methodologies used for fault identification and fault diagnosis are therefore also used in kernel PCA, see Cho et al. (2005) . However, the shortcomings of the contribution analysis based diagnostic techniques, highlighted by Van den Kerkhof, Vanlaer, Gins and Impe (2013), are inherited by kernel PCA. A second focus of this research document is therefore to also address fault diagnosis in kernel PCA based MSPM.

Note that the imperfections of multivariate process monitoring exposed by nonlinear relationships are not the only ones observed in practice. For example, it can be illustrated that MSPM will yield undesirable results when process characteristics change as a function of time. See He and Wang (2018) for a comprehensive record of instances when methods alternative to traditional MSPM should be used for SPM. In this document attention will be restricted to the non-linear case.

1.5 Summary

In this Chapter the importance and relevance of statistical process monitoring in modern manufacturing have been discussed. A broad overview was given of the techniques utilized in statistical process monitoring.

The important contributions made by Walter Shewhart in the 1920's have enlightened the manufacturing environment of the importance of statistical thinking in the understanding and subsequent optimisation of their processes. As illustrated by the Shewhart control charts, the idea that valuable information on process behaviour are locked away inside manufacturing data have enticed modern plants to increase the amount of data recorded. The large influx of measurements on process characteristics has introduced new

ways in which the behaviour of common cause variability can be statistically defined.

Building on the ideas of Shewhart, new statistical methodologies have been developed to address process monitoring challenges of current manufacturing factories. For the most part, it has been illustrated that these newly developed techniques have exceptional ability to detect when assignable cause variability is present. However, due to complex relationships amongst process variables, it is often found that it is difficult to objectively determine why process deviations occur.

1.6 Research focus and objectives

When utilising statistical methodologies for real life applications, it is often found that there are unanticipated shortcomings inherent to the techniques. This was the experience of the author after applying multivariate statistical process monitoring techniques to analyse data from chemical manufacturing plants. Specifically, it was identified for the approaches used to diagnose multivariate faults that:

- The univariate ranking of process variables can lead to the interpretation of multivariate faults that are confusing. This ranking has the unintended consequence that each variable will be considered on an individual basis, when all the variables should be evaluated as a collective. No clear guidelines are given on how the ranked variables should be further explored in order to explain why the process is statistically classified as being multivariately out-of-control.

- The effect of fault smearing is concerning. Variables that contain no information of the observed multivariate fault are ranked as important contributors. This makes the statistical interpretation of multivariate deviations more challenging.

Therefore, the shortcomings inherent to the diagnostic step of MSPM have led to the work captured in this research.

Research focus

The main focus of this research document is on the diagnostic philosophies employed to interpret multivariate process deviations that are identified using the PCA method. It was purposely decided to investigate PCA based MSPM since it is the most popular technique used by practitioners. Also, the detection and diagnostic principles utilised by PCA based MSPM are present in other MSPM techniques. Therefore, the idea is that if any improvements are made to the diagnostic step of PCA MSPM it will naturally extend, in some way or form, to most of the other MSPM methodologies.

It is known from the practical experience of the author and that of others (Lee et al., 2004), that there exists a great need for non-linear MSPM in the chemical industry. Specifically, it was decided to focus on kernel PCA based MSPM and the accompanying fault diagnostic approaches. It is the author's view, from a practical stance, that kernel PCA is relatively easy to implement in practice due to its close relation to PCA, and holds great potential value for industry.
It is noted that MSPM is but one aspect used in the operation of a complex manufacturing process. The focus of this research will be on the mathematical and statistical developments of the analysis of multivariate process data. Other facets of process monitoring and control, such as causality analysis, falls outside the scope of this research.

Research objectives

The objectives of the research are:

- Document the shortcomings of the traditional techniques used for fault diagnosis in MSPM through a theoretical evaluation of the methods, as well as with simulation studies.
- Develop new statistical analysis for fault diagnosis in PCA based MSPM.
- Extend the new statistical analysis for fault diagnosis in kernel PCA based MSPM.
- Illustrate the fault diagnosis performance of the newly developed methods using published data and data from a commercial chemical process. The performance will be evaluated based on the techniques ability to clearly communicate why the process is statistically identified as being different from multivariate common cause variability.

1.7 Layout of research

- The first chapter of this research document provides a broad overview of the popular techniques used for statistical process monitoring.
- In Chapter 2 a detailed account of principal component analysis based monitoring is provided. A comprehensive analysis of the diagnostic methods in PCA based MSPM is given to clarify the shortcomings.
- In Chapter 3 a new methodology is developed for fault diagnosis in PCA. This methodology introduces a new way of performing fault diagnosis that is fundamentally different to the existing approaches.

- In Chapter 4 non-linear MSPM is presented using the kernel PCA method. Shortcomings of the diagnostic methods employed for kernel PCA based fault interrogation will be emphasised.
- Chapter 5 provides new diagnostic methods that are developed based on the fault diagnosis methodology introduced in Chapter 3.
- A benchmark data set and data from the chemical industry is analysed in Chapter 6. It is illustrated how the new approach to diagnostic analysis assist engineers to statistically interpret multivariate process deviations.

Chapter 2

Principal Component Analysis based Fault Diagnosis

2.1 Introduction

The Principal Component Analysis technique is commonly used for multivariate statistical process monitoring of industrial chemical plants (MacGregor and Kourti, 1995). It can be used to statistically quantify if the process, that is being monitored, is operating at a condition that can be considered normal or otherwise. Additionally, principal component analysis can also be utilised as a diagnostic tool to provide some degree of understanding as to why the plant is operating at abnormal settings. Various techniques, based on the contribution analysis methodology, are employed in practice for diagnostic purposes (Alcala and Qin, 2011).

The objective of this chapter is to provide a review of multivariate statistical process monitoring based on Principal Component Analysis. A specific focus will be to give a detailed exploration i.e., theoretical and through simulation studies, of the different techniques used for fault diagnosis and to identify areas for improvement.

2.2 Principal Component Analysis

The multivariate statistical process control methodology is well suited for the monitoring of complex industrial processes, because it models the correlations that may exist between the process performance variables. In the current study an industrial process refers to one single process unit, such as a coal-fired boiler, a chemical reactor or a distillation unit. The process unit can have many variables including controllable and observational variables, and we are not limiting the discussions to any specific types or sets of variables. Furthermore, the MSPM methods discussed are not necessarily developed for implementation

within an advanced process control (APC) program since that requires specific knowledge of the control variables of the unit, how those variables are and may be manipulated, and what is feasible and practical within the APC program. Therefore, the MSPM methods discussed and developed are for multivariate process monitoring in general, and the variables to be used depend on the problem which is formulated in collaboration with the subject matter experts.
Most popular amongst the multitude of multivariate techniques employed in industry is Principal Component Analysis (PCA) (Qin, 2012). PCA is a technique often employed in areas such as chemometrics to analyse high dimensional data matrices. PCA creates a smaller set of new uncorrelated variables called latent variables or principal components. These variables, which are weighted linear combinations of the original variables, are selected such that the covariance structure present in the original data matrix is well approximated by the newly calculated latent variables. PCA is therefore a dimension reduction technique since it reduces the study of many variables to the analysis of a few representative latent variables.

PCA is usually introduced as an exploratory method that can be used to get a general understanding of the data being analysed, which will then provide insight into which other statistical techniques should be employed to analyse the data in more detail. PCA, however, has proven in itself to be an adequate statistical technique with which meaningful insights can be obtained. This is especially evident in areas such as statistical process monitoring (see Jackson and Mudholkar (1979)).

For statistical process monitoring, using PCA, it is required that subject matter experts (e.g. plant engineers) provide a dataset that consists of historical data describing the ideal operating conditions (IOC) as well as the process variables to consider for the specific process. Examples of IOC datasets might be the process conditions observed over time when the process being monitored

- is operated at conditions which yield maximum production, or
- demonstrated process behaviour at the desired stability characteristics.

This IOC collection of process observations will act as a reference against which future observations will be compared. The objective of PCA statistical process monitoring is to specify an IOC null distribution, based on some statistic, that can be referenced to determine if a new observation is similar or otherwise to ideal conditions.

Let $\mathbf{X} : n \times p$ denote the IOC reference matrix, where n is the number of observations and p the number of process variables to be monitored. Before PCA is performed, the reference matrix, $\mathbf{X}$ is preprocessed to have zero mean and unit variance i.e., mean centering

and variance scaling is applied. The weights used to calculate the PCA latent variables are then obtained by considering the eigen-decomposition of the sample covariance matrix,

$$\mathbf{S}_{p\times p} = \frac{1}{n-1}\mathbf{X}^T\mathbf{X}, \tag{2.1}$$

as follows

$$\mathbf{S} = \bar{\mathbf{V}}\bar{\mathbf{\Lambda}}\bar{\mathbf{V}}^T, \tag{2.2}$$

where $\bar{\mathbf{\Lambda}} : p \times p$ is the diagonal matrix of eigenvalues i.e.,

$$\bar{\mathbf{\Lambda}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}, \tag{2.3}$$

with $\lambda_i \geqslant 0$ for $i = 1, \ldots, p$, and $\bar{\mathbf{V}} : p \times p$ is the matrix of orthogonal eigenvectors. The elements of these eigenvectors are used as weights to construct the principal component latent variables. Usually only the first $r < p$ latent variables, represented by $\mathbf{V} : p \times r$, are used in the principal component analysis if it is found that the resulting estimating covariance matrix

$$\hat{\mathbf{S}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{2.4}$$

is a good approximation of $\mathbf{S}$. More specifically, the covariance can be decomposed into the estimate $\hat{\mathbf{S}}$ and residual

$$\tilde{\mathbf{S}} = \tilde{\mathbf{V}}_{p\times(p-r)}\tilde{\mathbf{\Lambda}}_{(p-r)\times(p-r)}\tilde{\mathbf{V}}^T_{(p-r)\times p)}, \tag{2.5}$$

matrices as follows

$$\begin{aligned}\mathbf{S} &= \begin{bmatrix} \mathbf{V}_{p\times r} & \tilde{\mathbf{V}}_{p\times(p-r)} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{r\times r} & \mathbf{0}_{r\times(p-r)} \\ \mathbf{0}_{(p-r)\times r} & \tilde{\mathbf{\Lambda}}_{(p-r)\times(p-r)} \end{bmatrix} \begin{bmatrix} \mathbf{V}^T_{r\times p} \\ \tilde{\mathbf{V}}^T_{(p-r)\times p} \end{bmatrix} \\ &= \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T + \tilde{\mathbf{V}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{V}}^T \\ &= \hat{\mathbf{S}} + \tilde{\mathbf{S}}. \end{aligned} \tag{2.6}$$

The first r principal component latent variables are then defined by the n-dimensional column vectors of

$$\mathbf{T}_{n\times r} = \mathbf{X}\mathbf{V}. \tag{2.7}$$

This matrix, $\mathbf{T}$, consisting of the n observations made on the r latent variabels, is often referred to as the PCA score matrix. PCA therefore reduces the high dimensional IOC matrix $\mathbf{X} : n \times p$ to a more manageable low dimensional matrix $\mathbf{T} : n \times r$. The objective of the PCA process monitoring methodology is to use the score matrix as a benchmark against which future observations, $\mathbf{x}_{new} : p \times 1$, can be compared. The linear relationships established in the construction of the IOC latent variables will be used to obtain an equivalent representation of the new observations i.e.,

$$\mathbf{t}_{new} : r \times 1 \;=\; \mathbf{V}^T\mathbf{x}_{new}, \tag{2.8}$$

where the preprocessing applied to $\mathbf{X}$ has also been performed on $\mathbf{x}_{new}$. The PCA monitoring approach defines different statistical measures or indices that can be used to determine whether the distribution that gave rise to $\mathbf{t}_{new}$ is similar to the one underlying the IOC score matrix $\mathbf{T}$.

2.3 Fault detection statistics

Performance statistics based on the PCA latent variables are used to generate null distributions against which new observations are compared to determine if current process conditions are different from IOC. The performance statistics most often employed are the squared prediction error (SPE) and Hotelling's T-squared (T^2) statistic. Consider a new observation $\mathbf{x}_{new} : p \times 1$ centered and scaled using the IOC matrix information. The SPE fault detection statistic for this new observation is defined by Jackson and Mudholkar (1979) as:

$$\begin{aligned} SPE_{new} &= \tilde{\mathbf{t}}_{new}^T \tilde{\mathbf{t}}_{new} \\ &= \mathbf{x}_{new}^T \tilde{\mathbf{V}} \tilde{\mathbf{V}}^T \mathbf{x}_{new} \\ &= \mathbf{x}_{new}^T \tilde{\mathbf{C}} \mathbf{x}_{new}, \end{aligned} \tag{2.9}$$

where $\tilde{\mathbf{C}} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^T$. The process is considered different from IOC if SPE_{new} is an extreme value of the set of SPE values calculated using the IOC data. Statistically the observed SPE_{new} value is considered extreme, with confidence level $(1-\alpha)100\%$, if it is larger than the following control limit

$$\delta^2 = k\chi_\alpha^2(df_{spe}), \tag{2.10}$$

where $df_{spe} = \frac{\theta_1^2}{\theta_2}$ and $k = \frac{\theta_2}{\theta_1}$ with $\theta_1 = \sum_{i=1}^{p-r} \tilde{\lambda}_i$ and $\theta_2 = \sum_{i=1}^{p-r} \tilde{\lambda}_i^2$.

The T^2 fault detection statistic is defined in Westerhuis et al. (2000) as:

$$\begin{aligned} T_{new}^2 &= \mathbf{t}_{new}^T \mathbf{\Lambda}^{-1} \mathbf{t}_{new} \\ &= \mathbf{x}_{new}^T \mathbf{V} \mathbf{\Lambda}^{-1} \mathbf{V}^T \mathbf{x}_{new} \\ &= \mathbf{x}_{new}^T \mathbf{D} \mathbf{x}_{new}, \end{aligned} \tag{2.11}$$

where $\mathbf{D} = \hat{\mathbf{S}}^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T$. The calculated T_{new}^2 is an extreme value, with confidence level $(1-\alpha)100\%$, if

$$T_{new}^2 > \tau^2, \tag{2.12}$$

where $\tau^2 = \chi_\alpha^2(r)$.

These performance statistics can also be combined into one single index. Yue and Qin (2001) defines the *Combined* performance statistic as follows:

$$\varphi_{new} = \mathbf{x}_{new}^T \mathbf{\Phi} \mathbf{x}_{new}, \tag{2.13}$$

where

$$\mathbf{\Phi} = \frac{\tilde{\mathbf{C}}}{\delta^2} + \frac{\mathbf{D}}{\tau^2}. \tag{2.14}$$

The control limit for this statistic is calculated as

$$\zeta^2 = g^{\varphi}\chi^2_{\alpha}(h^{\varphi}), \tag{2.15}$$

with confidence $(1-\alpha)100\%$ and where

$$g^{\varphi} = \frac{r/\tau^4 + \theta_2/\delta^4}{r/\tau^2 + \theta_1/\delta^2} \tag{2.16}$$

and

$$h^{\varphi} = \frac{(r/\tau^2 + \theta_1/\delta^2)^2}{r/\tau^4 + \theta_2/\delta^4}. \tag{2.17}$$

All of the above performance statistics are quadratic forms. The statistics can therefore be summarized in the following general form

$$Index(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}, \tag{2.18}$$

where $\mathbf{A}$ depends on the specific fault index.

2.4 Fault identification methods

If any of the performance statistics for a new process observation is found to be extreme, the process is considered to be out-of-control. It is therefore important to determine which of the process variables are responsible for the drift in process conditions. In multivariate statistical process control this is often referred as the diagnostic step. The use of contribution plots, first introduced by MacGregor and Kourti (1995), have become the method of choice for out-of-control diagnosis in PCA monitoring applications. Variable contribution analysis is the decomposition of the fault detection statistic into a sum of terms, such that each term corresponds to one of the process variables. These terms are called variable contributions. Variables with large contributions are considered responsible for the process drift. Hence, suppose it is possible to decompose the fault index (2.18) as the sum of

p parts as follows

$$Index(\mathbf{x}) \quad = \quad \sum_{i=1}^{p} contrib_i. \tag{2.19}$$

If each part can be associated with a specific process variable, using some predefined criterion, then $contrib_i$ constitutes the contribution of variable i to the statistic. Consequently the i^{th} process variable that corresponds to the largest $contrib_i$ is identified as the variable most likely to be the cause of the process deviation.

Therefore, given an abnormal fault detection statistic, contribution analysis tries to analyse this statistic to ascertain which variables caused this out-of-control behaviour. A relative importance ranking, called variable contribution, is then provided. Process engineers can therefore use this relative ranking of process variables in combination with fundamental understanding to identify what needs to be done in order to again operate within the IOC state.

When considering the quadratic nature of the fault identification statistic, as summarized by the expression in (2.18), it is clear that there is no unique way of factorising the fault statistic into a sum of independent terms, each related to a single process variable. Hence, as a result, many different methods have been proposed for variable contribution calculation. A detailed summary of these techniques is given by Alcala and Qin (2011). Three popular variable contribution methodologies for consideration are

- complete decomposition contributions,
- partial decomposition contributions and
- reconstruction based contributions.

The complete decomposition (CD) contribution of the i^{th} variable is given in Alcala and Qin (2011) as

$$CD_i \quad = \quad \left(\boldsymbol{\xi}_i^T \mathbf{A}^{\frac{1}{2}} \mathbf{x}\right)^2, \tag{2.20}$$

where $\boldsymbol{\xi}_i^T$ is the i^{th} row of the identity matrix $\mathbf{I} : p \times p$. These contributions are therefore always positive.

The definition of partial decomposition (PD) contribution is reported in Alcala and Qin (2011) for the i^{th} variable as

$$PD_i \quad = \quad \mathbf{x}^T \mathbf{A} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{x}. \tag{2.21}$$

Different from CD contributions, PD contributions can take on negative values. Even though different factorizing methodologies are employed by CD and PD contributions, it can be shown that both add up to the value of the corresponding statistic (Alcala and Qin, 2011).

The factorizations that define the complete and partial decomposition contributions can be combined into a generalized decomposition (GD) contribution. The following definition is provided by Alcala and Qin (2011),

$$GD_i = \mathbf{x}^T \mathbf{A}^{1-\varrho} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{A}^{\varrho} \mathbf{x} \tag{2.22}$$

for $0 \leqslant \varrho \leqslant 1$. Therefore if $\varrho = \frac{1}{2}$ then GD contribution equals CD contribution and when $\varrho = 0$ or $\varrho = 1$ then PD contribution is obtained.

Reconstruction-based (RB) contributions are different from CD and PD contributions. RB contributions determine how much each variable needs to be reconstructed in order to minimize the value of the performance statistic. The RB method of calculating variable contributions is therefore an optimization analysis of the statistic. It is therefore not related to a specific decomposition of the fault identification statistic. RB contributions, developed by Alcala and Qin (2011), considers, for each of the process variables $i = 1, \ldots, p$, the reconstructed vector

$$\mathbf{z}_i = \mathbf{x} - f_i \boldsymbol{\xi}_i, \tag{2.23}$$

with f_i representing a real constant value. The corresponding fault identification index value of the i^{th} reconstruction $\mathbf{z}_i$ equals

$$\mathbf{z}_i^T \mathbf{A} \mathbf{z}_i = (\mathbf{x} - f_i \boldsymbol{\xi}_i)^T \mathbf{A} (\mathbf{x} - f_i \boldsymbol{\xi}_i) \tag{2.24}$$

$$= \|\mathbf{x} - f_i \boldsymbol{\xi}_i\|_{\mathbf{A}}^2. \tag{2.25}$$

The objective of RB contribution analysis is to find the value of f_i such that (2.24) is as small as possible. After differentiating (2.24) w.r.t f_i and equating it to zero it is found that the minimizing value of f_i is

$$\hat{f}_i = \left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i \right)^{-1} \boldsymbol{\xi}_i^T \mathbf{A} \mathbf{x}. \tag{2.26}$$

Using equations (2.24) and (2.26), Alcala and Qin (2011) define the RB contribution of

the i^{th} variable as

$$\begin{aligned} RB_i &= \|\hat{f}_i \boldsymbol{\xi}_i\|^2_{\mathbf{A}} \\ &= \|\boldsymbol{\xi}_i \left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i\right)^{-1} \boldsymbol{\xi}_i^T \mathbf{A}\mathbf{x}\|^2_{\mathbf{A}} \\ &= \left[\boldsymbol{\xi}_i \left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i\right)^{-1} \boldsymbol{\xi}_i^T \mathbf{A}\mathbf{x}\right]^T \mathbf{A} \left[\boldsymbol{\xi}_i \left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i\right)^{-1} \boldsymbol{\xi}_i^T \mathbf{A}\mathbf{x}\right] \\ &= \left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i\right)^{-2} \mathbf{x}^T \mathbf{A} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{A}\mathbf{x} \\ &= \frac{\left(\boldsymbol{\xi}_i^T \mathbf{A}\mathbf{x}\right)^2}{\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i}. \end{aligned} \tag{2.27}$$

From (2.27) it is clear that RB contributions are always positive, but it does not necessarily add up to the performance statistic.

2.5 Fault smearing

Westerhuis et al. (2000) were the first to point out the contamination that is present in variable contributions based on the complete decomposition. The contamination refers to the contribution of one variable that is present in the contributions of other variables. Contributions of non-faulty variables therefore appear larger because of the influence of faulty variables. The effect of the fault can therefore be seen as being smeared out across all of the variables. Hence this contamination is often referred to as the smearing effect. Consider again the CD contributions as defined in (2.20). Van den Kerkhof, Vanlaer, Gins and Impe (2013) demonstrates that the CD contribution of the i^{th} variable can be written as a function of a linear combination of all the variables. The CD contribution based on the T^2 statistic can be rewritten as

$$CD_i^{T^2} = \left(\beta_i^{T^2} x_i + \sum_{j=1, j\neq i}^{p} \eta_{ij}^{T^2} x_j\right)^2, \tag{2.28}$$

where

$$\beta_i^{T^2} = \textstyle\sum_{k=1}^{r} \lambda_k^{-\frac{1}{2}} v_{ik}^2 \quad \text{and} \quad \eta_{ij}^{T^2} = \textstyle\sum_{k=1}^{r} \lambda_k^{-\frac{1}{2}} v_{ik} v_{jk}. \tag{2.29}$$

Similarly the SPE based CD contribution can be rewritten as

$$CD_i^{SPE} = \Big(\beta_i^{SPE} x_i + \sum_{j=1, j\neq i}^{p} \eta_{ij}^{SPE} x_j\Big)^2, \tag{2.30}$$

with

$$\beta_i^{SPE} = 1 - \textstyle\sum_{k=1}^{r} v_{ik}^2 \quad \text{and} \quad \eta_{ij}^{SPE} = -\textstyle\sum_{k=1}^{r} v_{ik} v_{jk}. \tag{2.31}$$

The presence of variables x_j, $j \neq i$ in the CD contribution calculations, given in (2.28) and (2.30), is a clear indication of fault smearing. The role of correlation in CD contribution fault smearing can now be understood when looking at (2.28) and (2.30). Recall from (2.4) that the PCA approximated correlation between the variables i and j can be written as

$$\sum_{k=1}^{r} \lambda_k v_{ik} v_{jk}. \tag{2.32}$$

The coefficients $\eta_{ij}^{T^2}$ and η_{ij}^{SPE} in the CD variable contribution calculation are therefore closely related to the correlation between variables. Note that the coefficient of variable j in the variable contribution calculation of variable i is the same as that of variable i in the contribution calculation of variable j. Therefore, if the correlation is large it can be expected that both variable i and j will have relatively large variable contributions. Van den Kerkhof, Vanlaer, Gins and Impe (2013) also give expressions similar to (2.28) and (2.30) for PD and RB contributions. The T^2 based PD variable contribution (2.21) is rewritten as

$$PD_i^{T^2} = x_i\Big(\gamma_i^{T^2} x_i + \sum_{j=1, j\neq i}^{p} \tau_{ij}^{T^2} x_j\Big), \tag{2.33}$$

with

$$\gamma_i^{T^2} = \textstyle\sum_{k=1}^{r} \lambda_k^{-1} v_{ik}^2 \quad \text{and} \quad \tau_{ij}^{T^2} = \textstyle\sum_{k=1}^{r} \lambda_k^{-1} v_{ik} v_{jk}. \tag{2.34}$$

The PD contribution based on the SPE statistic is rewritten as

$$PD_i^{SPE} = x_i\Big(\beta_i^{SPE}x_i + \sum_{j=1,j\neq i}^{p}\eta_{ij}^{SPE}x_j\Big). \tag{2.35}$$

Contributions calculated using the reconstruction approach are rewritten as follows

$$RB_i^{T^2} = \Big(\gamma_i^{T^2}x_i + \sum_{j=1,j\neq i}^{p}\tau_{ij}^{T^2}x_j\Big)^2/\gamma_i^{T^2}, \tag{2.36}$$

$$RB_i^{SPE} = \Big(\beta_i^{SPE}x_i + \sum_{j=1,j\neq i}^{p}\eta_{ij}^{SPE}x_j\Big)^2/\beta_i^{SPE}. \tag{2.37}$$

The reconstruction based contribution was originally introduced by Alcala and Qin (2009) to address the smearing effect. However, from Van den Kerkhof, Vanlaer, Gins and Van Impe (2013) and the above it is learned that the smearing effect is present in CD, PD and RB contribution analysis. The smearing effect will therefore incorrectly inflate the importance of variables that do not carry information on the process fault for all traditional fault diagnosis techniques. Fault smearing poses a serious obstacle in the correct interpretation of multivariate process faults (Liu et al., 2014).

2.6 Process fault types

Given the complexity of the processes often monitored using PCA, it implies that the characteristics of the faults encountered are not simplistic. As illustrated earlier, multivariate statistical process control is employed as an analysis that is additional to univariate statistical process control, with the objective of identifying process conditions that violate the correlation structure in process variables. That is, the process drifts are observed when the variables are univariately in-control. The deviations in process variables that give rise to these multivariate out-of-control states can be classified into three general process variable fault classes called single, multiple and multivariate sensor faults (Van den Kerkhof, Vanlaer, Gins and Van Impe, 2013). Note that this terminology should not be confused with terms often used in control engineering. A sensor in this setting refers to a process variable that is considered in the PCA model. The three sensor faults are defined as follows:

1. **Single sensor fault**
   Single sensor faults are observed as a result of a constant deviation in one of the

process variables. This is the most elementary of fault types. Let $\mathbf{x}_{ioc} : p \times 1$ represent an observation recorded when the process being monitored had the IOC characteristics. For a real value $f \in \mathbb{R}$, the single sensor fault, induced by the i^{th} process variable, is defined as

$$\mathbf{x}^*_{single} = \mathbf{x}_{ioc} + f\boldsymbol{\xi}_i. \tag{2.38}$$

Note that $\mathbf{x}^*_{single} : p \times 1$ can only be considered faulty if f is large enough such that at least one of the fault detection statistics (2.9),(2.11) or (2.13) quantifies it to be different from the IOC samples.

2. **Multiple sensor fault**
   Multiple sensor faults are realised when more than one single sensor fault occurs at the same time. To illustrate, consider the case when two single sensor faults simultaneously cause a process upset. Let

$$\mathbf{x}^{*(i)}_{single} = \mathbf{x}_{ioc} + f_1\boldsymbol{\xi}_i, \tag{2.39}$$

   represent the first single sensor fault where a constant $f_1 \in \mathbb{R}$ has been added to the i^{th} position of $\mathbf{x}_{ioc}$. Similarly let the second single sensor fault be given by

$$\mathbf{x}^{*(j)}_{single} = \mathbf{x}_{ioc} + f_2\boldsymbol{\xi}_j, \tag{2.40}$$

   with a fault size $f_2 \in \mathbb{R}$ added to the $j \neq i^{th}$ variable. In isolation both $\mathbf{x}^{*(i)}_{single}$ and $\mathbf{x}^{*(j)}_{single}$ will be out-of-control according to the PCA monitoring. Notice that both these single sensor faults have deviated from the same underlying IOC vector $\mathbf{x}_{ioc}$. The following vector,

$$\mathbf{x}^*_{multiple} = \mathbf{x}_{ioc} + f_1\boldsymbol{\xi}_i + f_2\boldsymbol{\xi}_j, \tag{2.41}$$

   is defined as a multiple sensor fault if it is out-of-control for at least one of the PCA fault identification indices subject to the vectors $\mathbf{x}_{ioc}+f_1\boldsymbol{\xi}_i$ and $\mathbf{x}_{ioc}+f_2\boldsymbol{\xi}_j$ being single sensor faults. This exact definition of multiple sensor faults has not been explicitly defined previously in the MSPM literature. In previous discussions (e.g. Van den

Kerkhof, Vanlaer, Gins and Impe (2013)) the authors ignored the requirement that the single sensor faults should independently yield the faulty condition. They simply define a multiple sensor fault as a sum of two constants $f_a \in \mathbb{R}$ and $f_b \in \mathbb{R}$ in positions i and j,

$$\mathbf{x}^*_{multiple^{old}} = \mathbf{x}_{ioc} + f_a\boldsymbol{\xi}_i + f_b\boldsymbol{\xi}_j, \tag{2.42}$$

for $i \neq j$. This definition will be referred to as the old multiple sensor fault definition. Given this definition, the possibility exists that either

$$\mathbf{x}_{ioc} + f_a\boldsymbol{\xi}_i, \tag{2.43}$$

or

$$\mathbf{x}_{ioc} + f_b\boldsymbol{\xi}_j, \tag{2.44}$$

will not result in a faulty condition. If one of these single sensor expressions is in-control, then the specific single sensor's role in (2.42) can be replaced with an in-control vector $\mathbf{x}^*_{ioc}$. Expression (2.42) can therefore, for example, be expressed as

$$\mathbf{x}^*_{multiple^{old}} = \mathbf{x}^*_{ioc} + f_b\boldsymbol{\xi}_j, \tag{2.45}$$

where $\mathbf{x}^*_{ioc} = \mathbf{x}_{ioc} + f_a\boldsymbol{\xi}_i$. The old multiple sensor fault definition will therefore, in some instances, actually take on the form of the single sensor fault.

3. **Multivariate sensor fault**

   To the best of the author's knowledge, no definition has previously been provided in the MSPM literature to describe the following fault type. The multivariate sensor fault is defined as the occurrence of multiple deviations, where if these deviations happened independently, none of them would yield a single sensor fault. That is,

$$\mathbf{x}^*_{multivariate} = \mathbf{x}_{ioc} + f_i\boldsymbol{\xi}_i + f_j\boldsymbol{\xi}_j, \tag{2.46}$$

with $i \neq j$ and $f_i, f_j \in \mathbb{R}$, is defined as a multivariate fault if $\mathbf{x}^*_{multivariate}$ is out-of-control, subject to both

$$\mathbf{x}_{ioc} + f_i \boldsymbol{\xi}_i \tag{2.47}$$

and

$$\mathbf{x}_{ioc} + f_j \boldsymbol{\xi}_j \tag{2.48}$$

being in-control.

It is obvious that, in practice, it is not known what the fault type associated with an out-of-control observation is. However, it is important to define these fault classes in order to evaluate the fault diagnosis ability of the different fault identification methods. This is usually quantified using mathematical or simulation analysis.

2.7 Mathematical fault diagnosis-ability

Alcala and Qin (2011) state that a diagnosis method can only be considered for application in practice if it can correctly diagnose the most simplest of faults. The single sensor fault

$$\mathbf{x}_{faulty} = \mathbf{x}_{ioc} + f\boldsymbol{\xi}_j, \tag{2.49}$$

with the IOC vector $\mathbf{x}_{ioc}$ equal to the mean of the IOC data. Recall that the mean of the IOC dataset is equal to zero due to the preprocessing performed. The single sensor fault can therefore be written as

$$\begin{aligned} \mathbf{x}_{faulty} &= \bar{\mathbf{x}}_{ioc} + \boldsymbol{\xi}_j f \\ &= \boldsymbol{\xi}_j f. \end{aligned} \tag{2.50}$$

The ability of the contribution techniques to diagnose this simple single sensor fault will now be presented. Please refer to Alcala and Qin (2011) for a more comprehensive discussion on the diagnosability requirement.

2.7.1 Complete decomposition contribution

The CD contribution, defined by (2.20), of variable i for the fault vector $\mathbf{x}_{faulty}$ (2.50) equals

$$\begin{aligned} CD_i &= \left(\boldsymbol{\xi}_i^T \mathbf{A}^{\frac{1}{2}} \mathbf{x}_{faulty}\right)^2 && (2.51) \\ &= \boldsymbol{\xi}_j^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_j f^2, && (2.52) \end{aligned}$$

where $\mathbf{x}_{faulty} = \boldsymbol{\xi}_j f$ for $i \neq j$ and the contribution of the variable position where the fault occurs is given by

$$CD_j = \boldsymbol{\xi}_j^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_j \boldsymbol{\xi}_j^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_j f^2, \quad (2.53)$$

for $i = j$. Correct fault identification is achieved when the contribution calculated in position j is the largest amongst all i.e., when

$$\begin{aligned} CD_j &\geqslant CD_i \\ \left(\boldsymbol{\xi}_j^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_j\right)^2 &\geqslant \left(\boldsymbol{\xi}_i^T \mathbf{A}^{\frac{1}{2}} \boldsymbol{\xi}_j\right)^2. && (2.54) \end{aligned}$$

$\forall j \neq i$. The above inequality does not always hold true. CD contribution analysis therefore fails the simple single sensor fault diagnosis test.

2.7.2 Partial decomposition contribution

Variable i's PD contribution for $\mathbf{x}_{faulty}$ is obtained by substituting (2.50) into (2.21)

$$\begin{aligned} PD_i &= \boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \boldsymbol{\xi}_j f^2 \\ &= 0, && (2.55) \end{aligned}$$

since $\boldsymbol{\xi}_i^T \boldsymbol{\xi}_j = 0$ for $i \neq j$. The contribution of the faulty variable is obtained by setting $i = j$

$$PD_j = \boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_j \boldsymbol{\xi}_j^T \boldsymbol{\xi}_j f^2. \quad (2.56)$$

The correct identification of variable j as the faulty variable is realised when

$$\begin{aligned} PD_j &\geqslant PD_i \\ \boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_j &\geqslant 0, \end{aligned} \tag{2.57}$$

$\forall j \neq i$. In other words correct faulty variable isolation is obtained when the matrix $\mathbf{A}$ is positive definite. In the case of the T^2 and SPE statistics, the matrix $\mathbf{A}$ is known to be positive definite. It can therefore be concluded that PD contributions will always identify the correct variables in the simple one sensor fault case.

2.7.3 Reconstruction based contribution

Substituting (2.50) into (2.27) gives the RB contributions for $\mathbf{x}_{faulty}$ as

$$RB_i = \begin{cases} \frac{\left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_j\right)^2}{\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i} f^2 & \text{for } i \neq j \\ \frac{\left(\boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_j\right)^2}{\boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_j} f^2 & \text{for } i = j. \end{cases} \tag{2.58}$$

Correct fault isolation is achieved when

$$\begin{aligned} RB_j &\geqslant RB_i \\ \boldsymbol{\xi}_j^T \mathbf{A} \boldsymbol{\xi}_j &\geqslant \frac{\left(\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_j\right)^2}{\boldsymbol{\xi}_i^T \mathbf{A} \boldsymbol{\xi}_i}. \end{aligned} \tag{2.59}$$

It follows from the positive semidefinite properties of $\mathbf{A}$ that the above inequality is true. Therefore, given the fault (2.50), it follows that RB contribution analysis will always detect the correct fault.

Diagnosis-ability analysis is a quick analysis that can be done to determine if a certain fault diagnosis technique is worth further investigation. The above analysis shows that only partial decomposition and reconstruction based contribution analysis pass the diagnosis-ability test.

2.8 PCA fault diagnosis simulation study

2.8.1 Objective

In this section a simulation study will be formulated to evaluate the diagnosis abilities of the methods described in Section 2.4. This will be done by defining an IOC dataset that will represent the reference used to construct the PCA model. A large array of fault examples will then be simulated from the three fault classes discussed in Section 2.6. The contribution analysis diagnosis techniques will then be used to scrutinise these simulated faults to determine their ability to correctly identify the faulty sensors. Hence, a general view of the impact of fault smearing using simulated data will be given. The results obtained using this simulation only apply to the simulated data set. However, it is envisioned that this simulation study will provide additional insight into the expected performance of the different diagnosis techniques in practice.

2.8.2 Data definition

The process under consideration is identified by seven variables $\mathbf{x}_{7\times1} = \begin{bmatrix} x_1 & x_2 & \dots & x_7 \end{bmatrix}^T$. The following PCA model will be used to simulate example observations from this process when operating under the ideal conditions:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_7 \end{bmatrix} = \mathbf{V}_{7\times4}\mathbf{t}_{4\times1} + \mathbf{e}_{7\times1} \tag{2.60}$$

$$= \begin{bmatrix} 0.55 & 0.31 & -0.02 & -0.10 \\ -0.22 & 0.43 & 0.61 & 0.13 \\ 0.24 & -0.40 & 0.54 & 0.30 \\ -0.47 & -0.13 & 0.04 & 0.59 \\ -0.36 & 0.30 & 0.37 & -0.53 \\ -0.22 & 0.53 & -0.41 & 0.33 \\ -0.44 & -0.42 & -0.17 & -0.38 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_7 \end{bmatrix} \tag{2.61}$$

where t_i for $i = 1, \ldots, 4$ is distributed normal as follows:

$$t_1 \sim N(0, 1.54) \tag{2.62}$$
$$t_2 \sim N(0, 1.13) \tag{2.63}$$
$$t_3 \sim N(0, 1.08) \tag{2.64}$$
$$t_4 \sim N(0, 0.89). \tag{2.65}$$

The residual vector $\mathbf{e} : 7 \times 1 = \begin{bmatrix} e_1 & e_2 & \ldots & e_7 \end{bmatrix}^T$ is distributed multivariate normal as

$$\mathbf{e} \sim MN(\mathbf{0}, \mathbf{\Sigma}_e), \tag{2.66}$$

with $\mathbf{0} : 7 \times 1$ a vector of zeros and

$$\mathbf{\Sigma}_e : 7 \times 7 = \begin{bmatrix} 0.15 & -0.01 & -0.01 & 0.11 & 0.03 & -0.03 & 0.07 \\ -0.01 & 0.20 & -0.09 & -0.08 & -0.20 & -0.05 & 0.11 \\ -0.01 & -0.09 & 0.25 & -0.10 & 0.10 & 0.24 & 0.07 \\ 0.11 & -0.08 & -0.10 & 0.19 & 0.08 & -0.14 & -0.07 \\ 0.03 & -0.20 & 0.10 & 0.08 & 0.20 & 0.06 & -0.09 \\ -0.03 & -0.05 & 0.24 & -0.14 & 0.06 & 0.25 & 0.09 \\ 0.07 & 0.11 & 0.07 & -0.07 & -0.09 & 0.09 & 0.17 \end{bmatrix}. \tag{2.67}$$

This model structure has been derived from a PCA model applied to an actual data set from a commercial process. Therefore, although the data are simulated, it is representative of data observed on a real life process plant. Due to intellectual property constraints the actual data cannot be provided. A similar simulation data model was provided by Alcala and Qin (2011). It is the author's opinion that their simulated data set is overly simplified. For example, in Alcala and Qin (2011), it is assumed that the error vector can be replaced by a normally distributed variable.

2.8.3 PCA results

Using this model, 2000 observations were simulated to construct the IOC matrix $\mathbf{X}$: 2000×7. This matrix was then centered and scaled before calculating the PCA information. Figure 2.1 give histograms to convey the distributional properties of the values observed for the individual simulated variables.

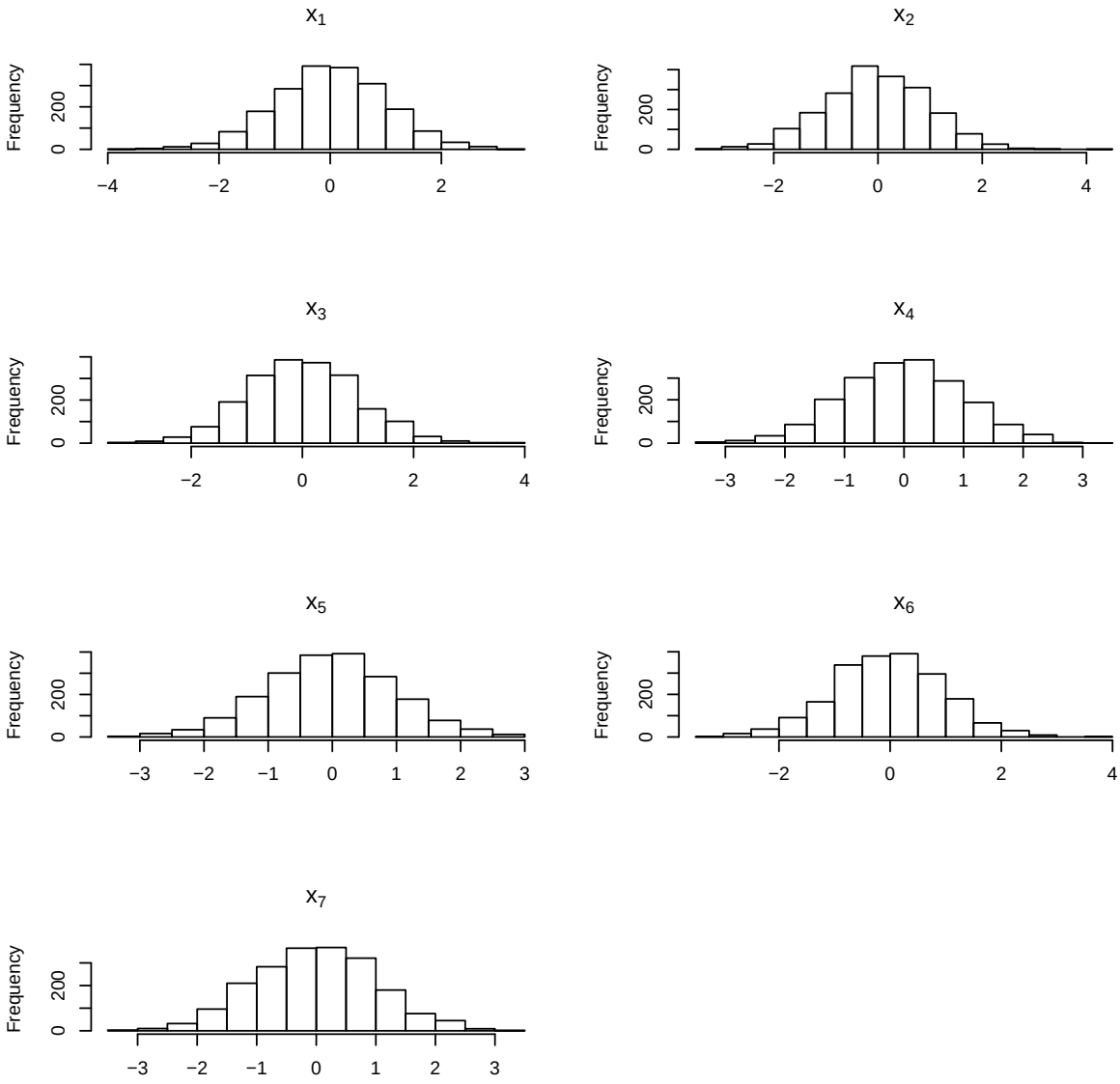


Figure 2.1: Distributions of simulated process variables.

The PCA model fitted is based on four principal component vectors. The amount of variability captured by this model is equal to 80.45% which is considered sufficient given the process of interest. The critical values of the three fault identification statistics are listed in Table 2.1 for this PCA model.

Table 2.1: Critical values of fault identification statistics for the PCA model fitted on simulated data.

Statistic	Critical value
SPE	3.73
T^2	9.49
$Combined$	1.62

2.8.4 Fault diagnosis results

Faulty vector examples will now be constructed using the definitions specified in Section 2.6. The ability of the contribution analysis techniques to detect the process faults correctly will now be evaluated.

2.8.4.1 Single sensor fault

The first fault type that will be considered is the single sensor fault type. That is, faulty vectors were constructed based on the following expression

$$\mathbf{x}_{faulty} = \mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i, \tag{2.68}$$

for a specific variable i, explained in Section 2.6. The fault free part in this expression was established in three ways i.e., the vectors representing IOC were determined using the three different fault detection statistics. Therefore, three sets of single sensor fault examples were created, each corresponding to one of the fault identification statistics. For each of these three sets of fault data the following was done:

1. For every fault size f in the set of fault magnitudes $0.1, 0.2, 0.3, \ldots, 5$,
2. generate 5000 IOC samples that are in-control based on a particular fault detection statistic,
3. for each IOC sample randomly select a process variable j,
4. if the selected IOC sample is part of the first 2500 samples, then add f to the selected process variable,
5. otherwise subtract f from the j^{th} variable in the selected IOC sample.

In doing this 250000 possible single sensor faulty examples are obtained, for each statistic. These observation vectors were then filtered to identify only those that are out-of-control based on the corresponding statistic's critical value. Additionally, as motivated in Section 2.6, the data were filtered and only those vectors that are univariately inside the IOC ranges as given by Figure 2.1 are retained. The number of single sensor faulty data vectors is reported in Table 2.2.

Table 2.2: The number of single sensor faults simulated.

Statistic	# single sensor faults
SPE	47786
T^2	14283
$Combined$	49576

Figure 2.2 gives frequencies of the fault identification statistics observed for the simulated single sensor faults. The fault statistic sizes are therefore skewed towards the smaller magnitudes.

The contribution analysis techniques of Section 2.4 were then applied to these single sensor vectors to identify the faulty variables. The results are summarised in Figures 2.3, 2.4 and 2.5. The percentage of variables that was correctly identified as being the cause of the deviation, by the different variable contribution methods, is plotted for each of the absolute fault magnitudes. As an example, in Figure 2.3 it is illustrated that for a fault size of three the RB method correctly identified the faulty variable in 60% of the cases. In comparison, the PD method achieves about 90% accuracy for the same fault size.

Even though single sensor faults are the simplest type of fault that can occur, it is clear from Figure 2.3 that none of the variable contribution techniques, applied to the SPE statistic, achieve 100% correct diagnosis across the range of fault sizes. Of the three variable contribution techniques, PD contribution analysis is the only technique that attains a correct diagnosis rate that is close to 100% for a fault size close to five. The highest correct diagnosis is achieved at the larger fault magnitudes (> 4). The deductions that can be made on the relative performance of the different contribution analyses in Figures 2.4 and 2.5, for the T^2 and $Combined$ statistics respectively, are similar to those from Figure 2.3.

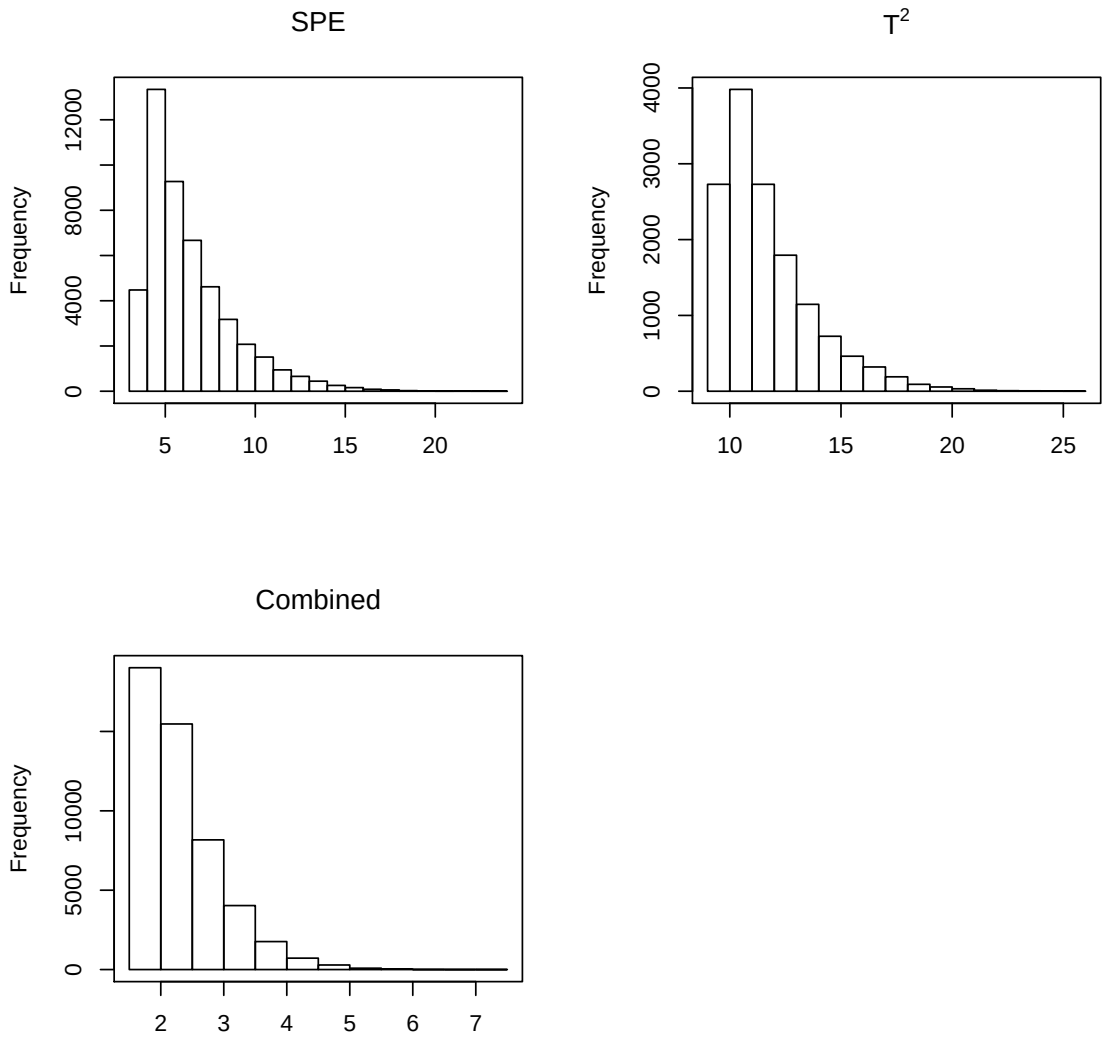


Figure 2.2: Distributions of fault identification statistic sizes observed for the simulated single sensor faults.

It should be re-emphasised that all of the samples evaluated in Figures 2.3, 2.4 and 2.5 were out-of-control based on the corresponding 95% control limit. Therefore, it is equally important across all fault sizes that the diagnosis results are acceptable. It is therefore concerning to observe the poor behaviour of the different methods for the smaller fault magnitudes. In most cases the reported correct diagnosis rate does not exceed 40% at fault sizes smaller than 0.5. This is not acceptable for application in practice. The results therefore suggest that the contribution analysis techniques are not ideal for the diagnosis of the most simplistic fault type i.e., single sensor faults. Therefore, it is expected that the performance of these techniques will be much worse when more complex fault types are considered.

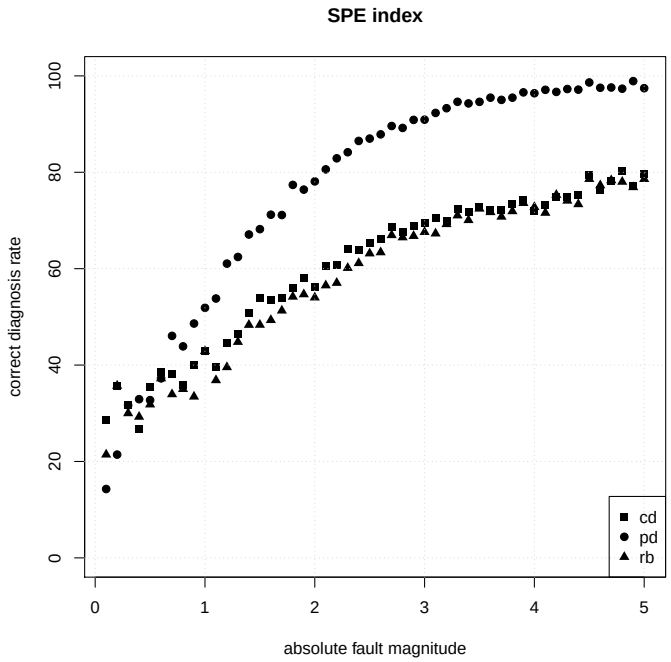


Figure 2.3: Correct diagnosis of faulty variables expressed relative to the magnitude of the fault induced. Contributions were calculated using the SPE index and corresponding PCA model.

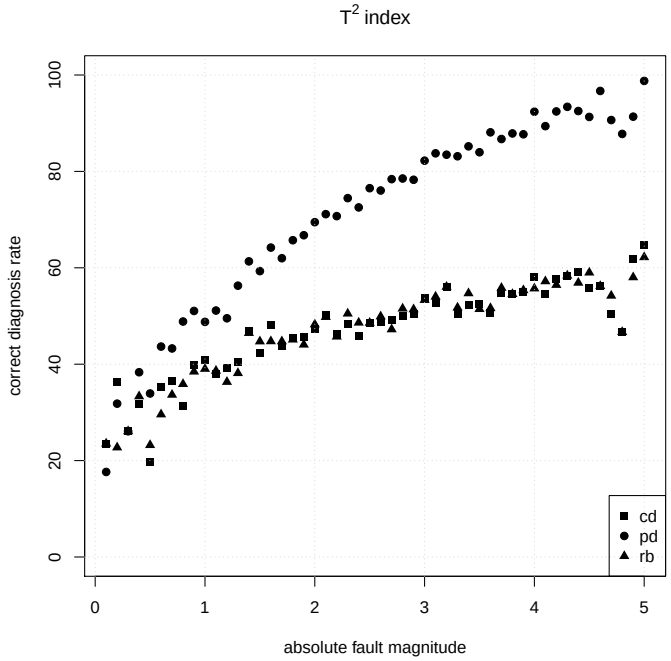


Figure 2.4: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the T^2 index and corresponding PCA model.

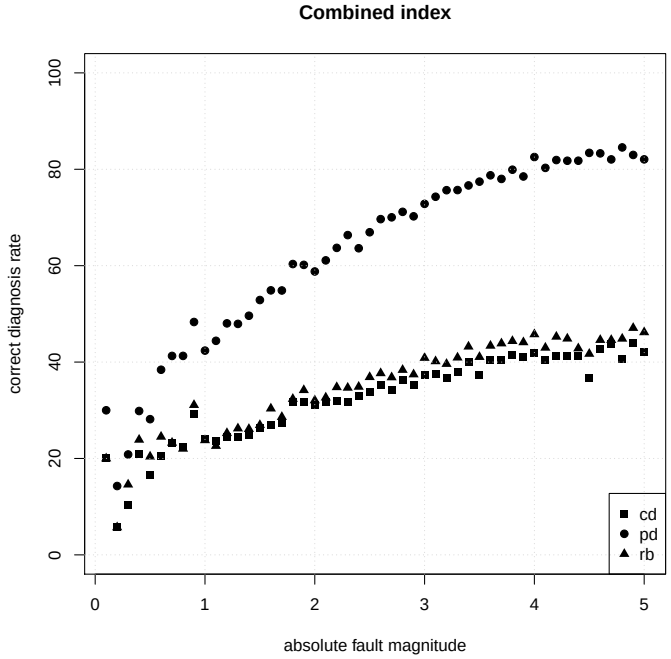


Figure 2.5: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the *Combined* index and corresponding PCA model.

2.8.4.2 Multiple sensor fault

The diagnostic performance of the contribution analysis methodology for the multiple sensor fault type is now evaluated. Two dimensional multiple sensor fault examples were generated using the following expression

$$\mathbf{x}_{faulty} \quad = \quad \mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i \pm f\boldsymbol{\xi}_j, \tag{2.69}$$

for $i \neq j$ and with $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i$ and $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_j$ larger than the control limit of the chosen fault detection index. Note that this is a simplified version of the multiple sensor fault where the faults added to the i^{th} and j^{th} process variables are identical. This simplification was made in order to limit the number of potential multiple sensor faulty samples that can be simulated. It will also make it easier to interpret the simulation results.

Potential multiple sensor faulty vectors were created as follows:

1. For every fault size f in the set of fault magnitudes $0.1, 0.2, 0.3, \ldots, 5$,

2. generate 20000 IOC samples, that are in-control based on a particular fault detection index,

3. for each IOC sample, $\mathbf{x}_{ioc}$, randomly select two process variables i and j with $i \neq j$,
4. if the selected IOC sample is part of the first 10000, then add f to the selected process variables,
5. otherwise subtract f from the i^{th} and j^{th} position in selected IOC sample,
6. check if the single sensors $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i$ and $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_j$ are both out-of-control,
7. if the two single sensors are faulty then create a multiple sensor fault,
8. otherwise return to step 3.

The potential multiple sensor fault vectors were then filtered to identify only those that are out-of-control based on the corresponding statistic's critical value. The number of faulty vectors identified for each of the fault detection indices is listed in Table 2.3. Figure 2.6 gives frequencies of the fault identification statistic sizes observed for the simulated multiple sensor faults.

Table 2.3: The number of multiple sensor faults simulated.

Statistic	# single sensor faults
SPE	55626
T^2	9010
$Combined$	62808

The diagnostic ability of the contribution analysis techniques introduced in Section 2.4 was evaluated using these multiple sensor fault vectors. The results are summarised in Figures 2.7, 2.8 and 2.9. These graphs depict the percentage of the vectors for which the correct variable pair was identified as being the cause of the multiple sensor fault, by the different variable contribution methods.

For example, the contribution analysis results for the SPE statistic, in Figure 2.7, shows that none of the techniques achieve a correct fault isolation rate higher than 80%. The PD technique, relative to the other two methods, display the best results. It is important to note the poor diagnostic performance at the smaller fault sizes in all three graphs i.e., Figures 2.7, 2.8 and 2.9. In many of the cases the correct diagnosis rate does not exceed 30% for fault sizes smaller than 0.5. The simulation results therefore suggest that the contribution analysis techniques are not well suited for the diagnosis of multiple sensor faults.

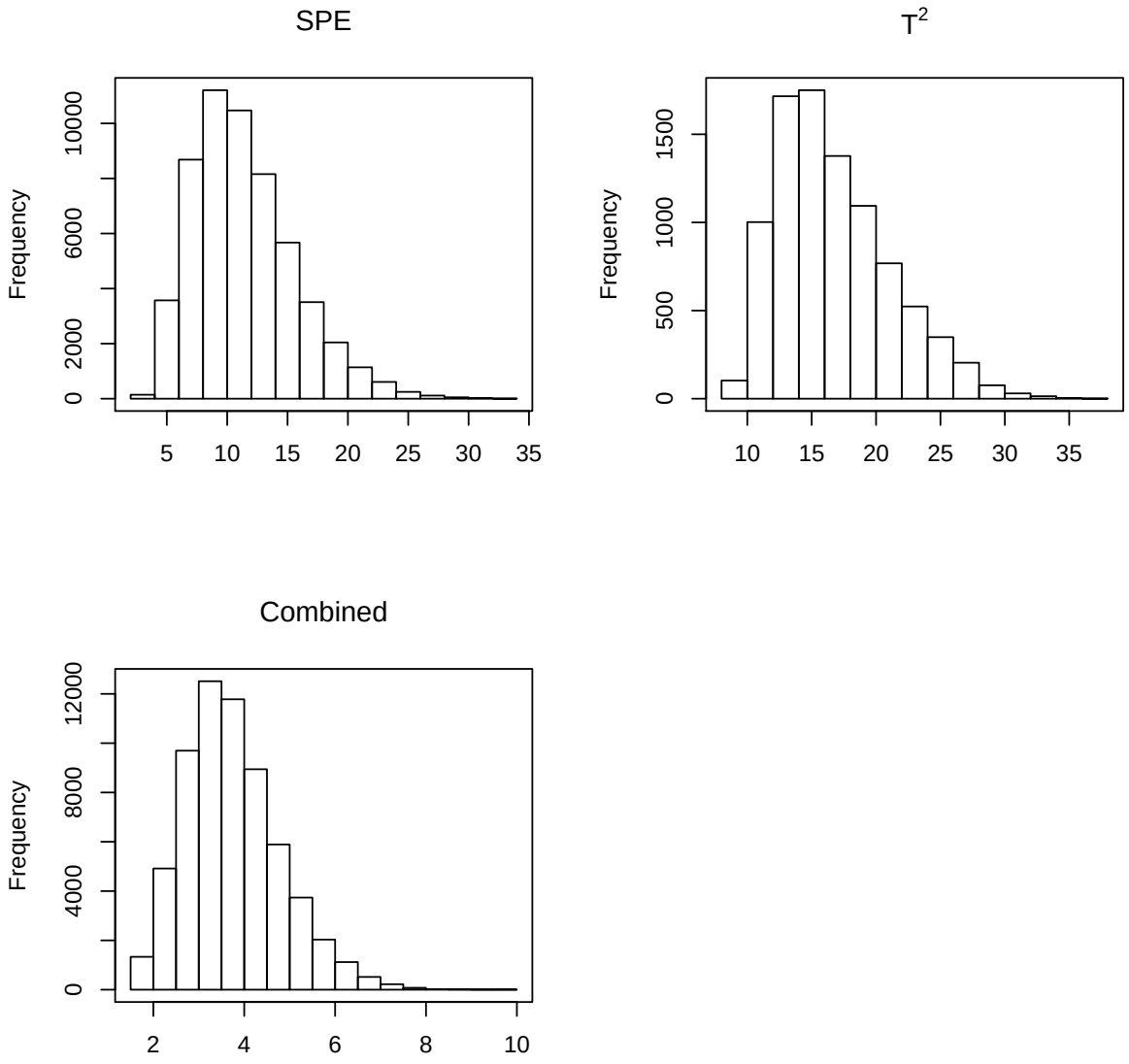


Figure 2.6: Distributions of fault identification statistic sizes observed for the simulated multiple sensor faults.

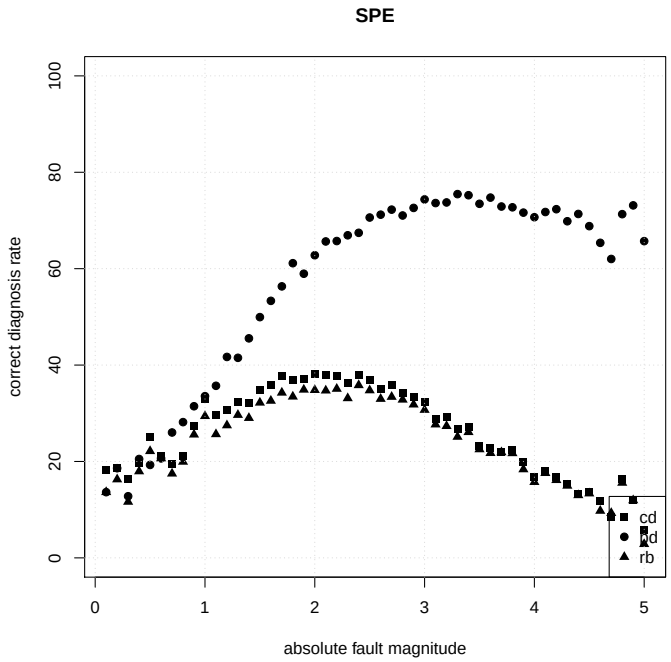


Figure 2.7: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the SPE index and corresponding PCA model.

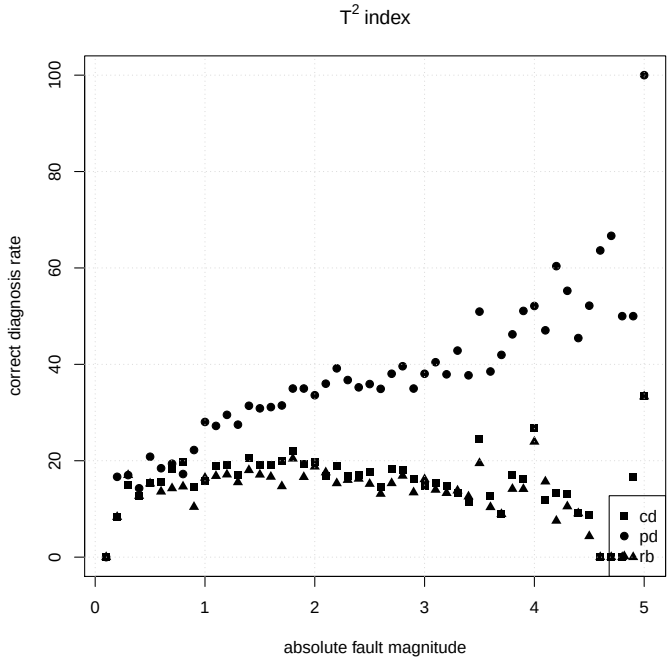


Figure 2.8: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the T^2 index and corresponding PCA model.

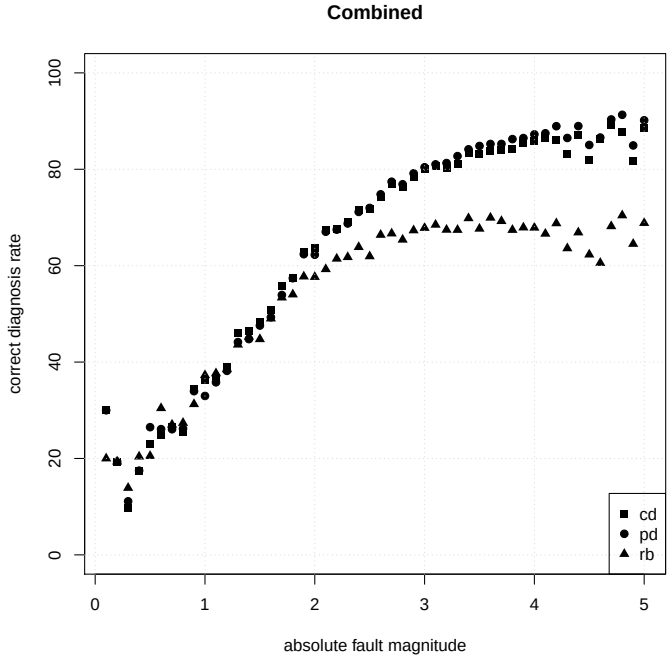


Figure 2.9: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the *Combined* index and corresponding PCA model.

2.8.4.3 Multivariate sensor fault

The third fault type considered is the multivariate sensor fault. Potential multivariate sensor faults, with deviations in two of the process variables, were generated using the following equation

$$\mathbf{x}_{faulty} \quad = \quad \mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i \pm f\boldsymbol{\xi}_j, \tag{2.70}$$

for $i \neq j$ and with the vectors $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i$ and $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_j$ both representing in-control vectors. This represents a simplified version of the multivariate sensor fault where the faults added to the i^{th} and j^{th} process variables are identical. This simplification was made in order to limit the number of potential multivariate sensor faulty samples that can be simulated. It will also make it easier to interpret the simulation results.

Potential multivariate sensor faulty vectors were created as follows:

1. For every fault size f in the set of fault magnitudes $0.1, 0.2, 0.3, \ldots, 5$,

2. generate 20000 IOC samples, that are in-control based on a particular fault detection index,

3. for each IOC sample, $\mathbf{x}_{ioc}$, randomly select two process variables i and j with $i \neq j$,
4. if the selected IOC sample is part of the first 10000, then add f to the selected process variables,
5. otherwise subtract f from the i^{th} and j^{th} position in selected IOC sample,
6. check if the single sensors $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_i$ and $\mathbf{x}_{ioc} \pm f\boldsymbol{\xi}_j$ are both in-control,
7. if the two single sensors are not faulty then create a multivariate sensor fault,
8. otherwise return to step 3.

The potential faulty vectors were then filtered, using the fault identification indices, to identify only those which are multivariately out-of-control. The sample sizes of the faulty vectors, for each of the fault indices, are summarised in Table 2.4. Figure 2.10 gives the distributions of the fault statistic evaluations of these faulty observations.

Table 2.4: The number of multivariate sensor faults simulated.

Statistic	# single sensor faults
SPE	12519
T^2	66232
$Combined$	21259

Similar to the simulation studies of the single and multiple sensor faults, the diagnostic ability of the contribution analysis techniques will now be demonstrated using these simulated multivariate sensor fault examples. The results are displayed in Figures 2.11, 2.12 and 2.13. Simulation results of the contribution analysis based on the SPE statistic, in Figure 2.11, again show that the PD contribution is to be preferred as it achieves 100% correct diagnosis at a fault size of about 3.5. Figure 2.12 illustrates, for the T^2 based contribution analysis, that PD is the only method that attains 100% correct diagnosis. The simulation outcome for the contribution analysis of the Combined statistic, in Figure 2.13, is much more favourable, compared to other two statistics. All three methods report percentages that can be considered to be identical. A 100% correct diagnostic rate was achieved for a fault magnitude of three and larger. It is again important to note that for the smaller fault sizes the diagnostic performance is not acceptable. It can therefore be concluded that the contribution analysis approaches are not suitable to diagnose multivariate sensor faults.

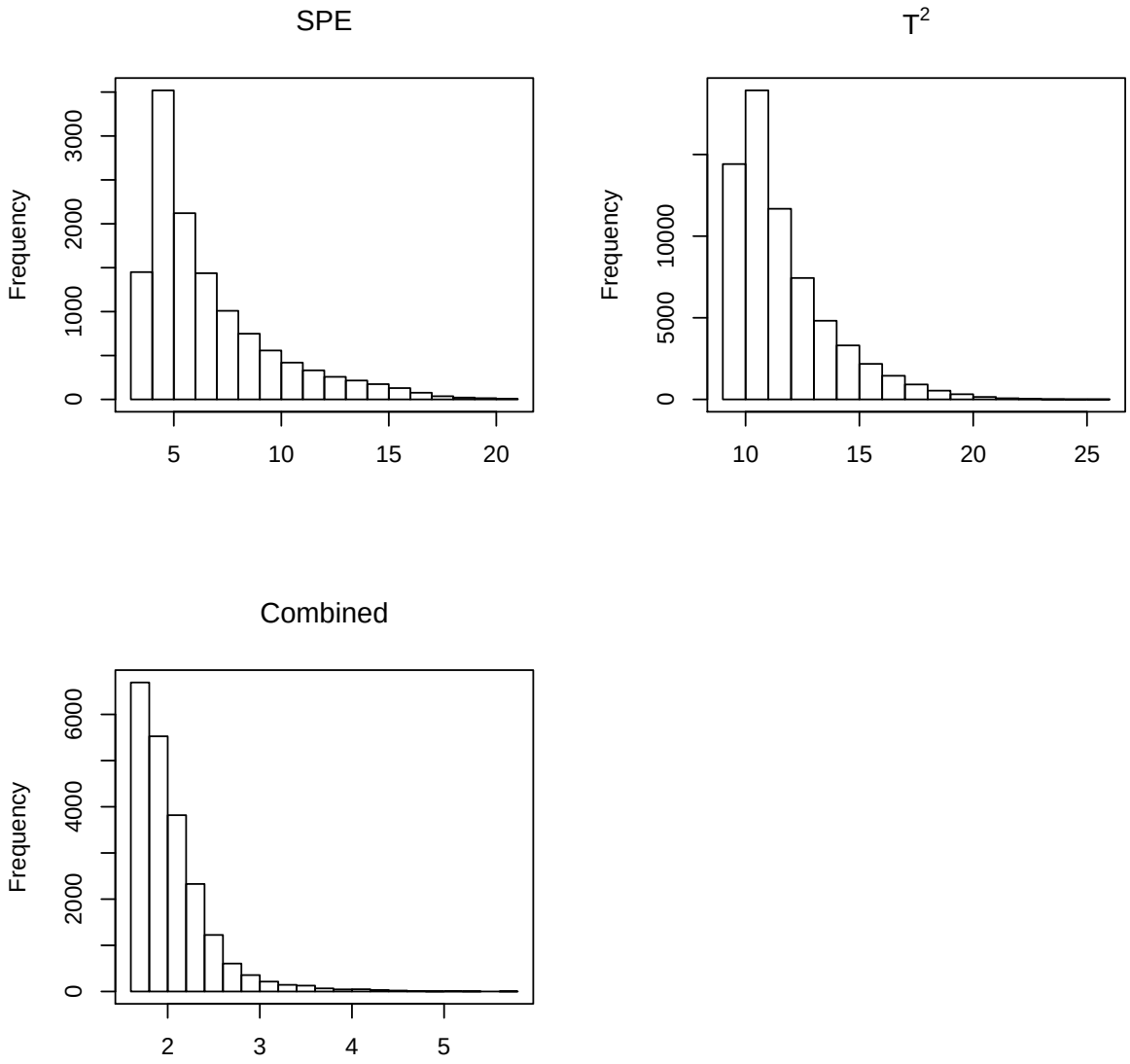


Figure 2.10: Distributions of fault identification statistic sizes observed for the simulated multivariate sensor faults.

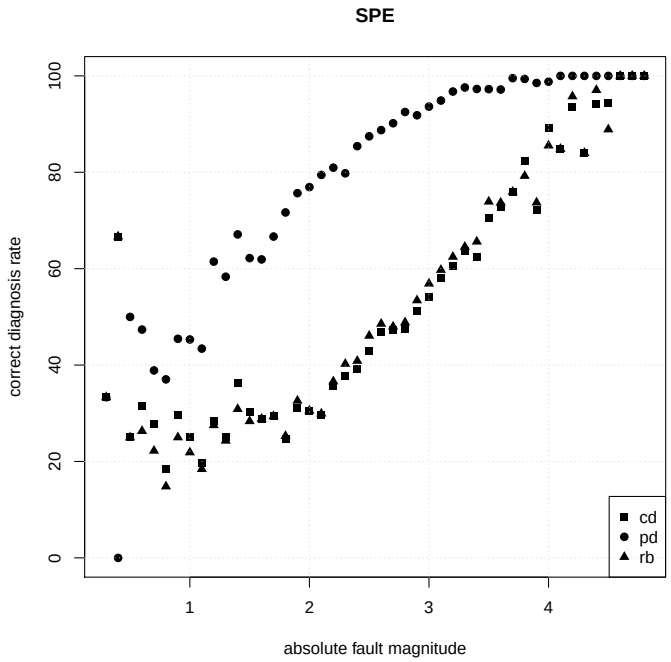


Figure 2.11: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the SPE index and corresponding PCA model.

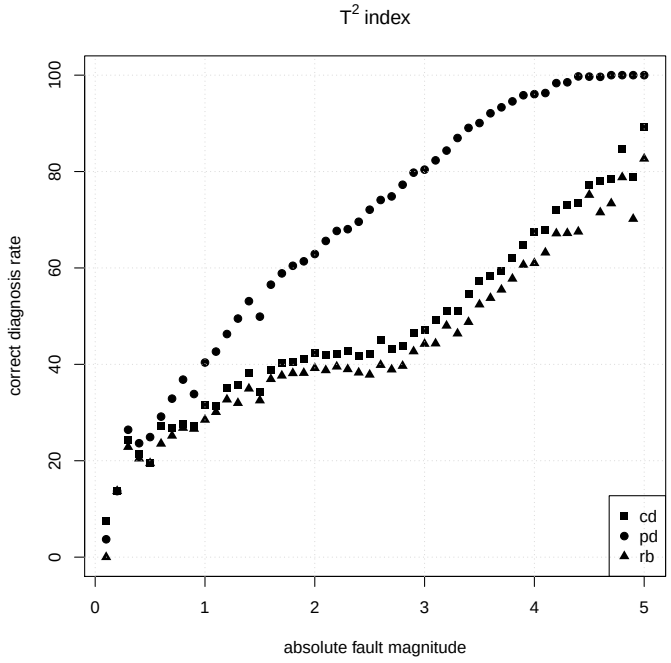


Figure 2.12: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the T^2 index and corresponding PCA model.

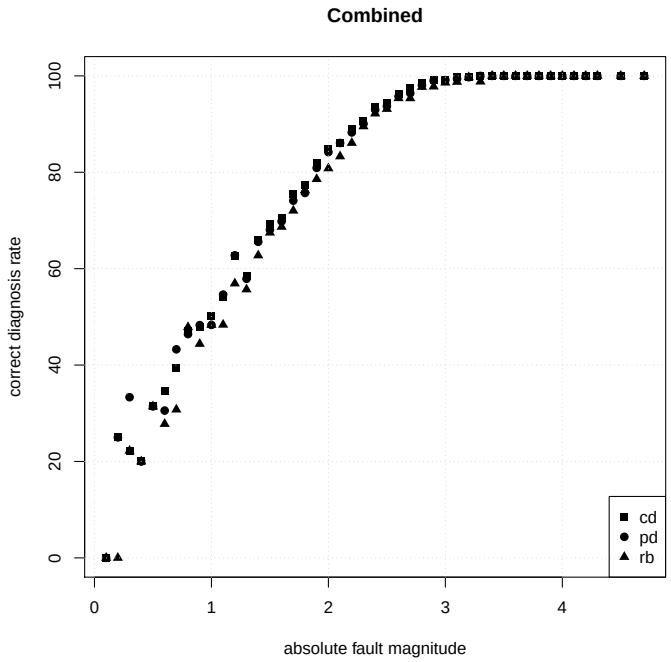


Figure 2.13: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the *Combined* index and corresponding PCA model.

2.9 Recent developments in fault diagnosis analysis

In this section interesting developments in fault diagnosis research will be listed. Many researchers agree that fault smearing is an important shortcomming of the traditional diagnostic methodologies that needs to be addressed. See for example Liu (2012), Liu et al. (2014) and Ji et al. (2018). The RB contribution (Alcala and Qin, 2011) was one of the first attempts to alleviate fault smearing. However, it is still present as was demonstrated in the simulation study of the previous section. Shang et al. (2016) showed that RB contributions give diagnosis results that are misleading. Ji et al. (2016) indicated that RB contributions have diminished effect when the latent variable dimension equals one. Liu et al. (2012) introduced the reconstruction based block (RBB) contribution to quantify the importance of grouped process variables. Since RBB contributions are based on the methodology of RB contribution it still suffers from the smearing effect.

Recently most of the published research is motivated by the use of sparse learning for improved fault diagnosis. Examples are Yan and Yao (2015), Liu et al. (2019) and Shang et al. (2019). It is theorised that only a few process variables are responsible for most multivariate deviations (Dunia and Joe Qin, 1998). The fault reconstruction problem is reformulated as the search for a constrained subset of reconstructed variables that will

minimize the out-of-control fault detection index. Branch and bound (B&B) algorithms have been implemented to solve this nonlinear optimization problem. See Kariwala et al. (2010) and He et al. (2012). However, the high computational cost of the B&B algorithm has made online implementation difficult (Yan and Yao, 2015). A fault isolation strategy based on variable selection is put forward by Yan and Yao (2015) to address this limitation. They transform the reconstruction problem into a quadratic programming problem that can be solved using the least absolute shrinkage and selection operator (LASSO) (Hastie et al., 2001).

In Shang et al. (2019) it is argued that the individual evaluation of process variables for fault diagnosis makes plant-wide interpretation difficult. It is therefore motivated that contribution analysis should be performed on groups of variables. Where the variables are typically grouped according to their physical location in the plant. Shang et al. (2019) extends the ideas of Yan and Yao (2015) by defining a generalised grouped contribution that is based on the group Lasso (Yuan and Lin, 2006). This methodology will identify a specific group of variables as being important to diagnose the multivariate deviation. Additionally it also provides individual importance measures for the variables inside the identified group. The proposed diagnosis methodology is therefore hierarchical.

Other interesting developments in fault diagnosis are:

- Liu (2012) proposed an approach that construct contribution plots without smearing effect on non-faulty variables. In this method a subspace of variables is identified that can be used to interpret the deviation. The subspace of variables is identified using a missing variable approach (Qin et al., 1997). Each process variable in this subset have individual importances assigned to them.

- Zhang et al. (2018) developed a new methodology for improved fault detection in the residual subspace. An associated fault diagnosis strategy based on contribution analysis is suggested. In this diagnosis strategy the extreme scores are identified. Variable importances are then calculated as the total contribution across the deviating scores.

All of the research referenced in this section are different ways of calculating individual variable importance measures that can be used to interpret multivariate process deviations. In Chapter 3 it will be motivated that individual variable contributions can potentially be ineffective in explaining the nature of multivariate process deviations. A new methodology based on pairwise variable contributions will be introduced.

2.10 Summary

In this chapter the most popular method employed in industrial applications for multivariate statistical process monitoring i.e., Principal Component Analysis (PCA), has been reviewed. It was illustrated how, based on historical data representing ideal operating conditions (IOC), PCA can be used to identify if future process characteristics are statistically similar or otherwise to the IOC data. The formulation of the three statistical measures commonly used to do this identification was discussed. Two fault types are widely accepted as deviations that can occur in practice, single and multiple sensor faults. It was determined that the formulation of the multiple sensor fault to be incomplete since in the definition it allows for realisation of single sensor faults. The definition has therefore been updated to clearly distinguish between single and multiple sensor faults. A new fault type referred to as a multivariate sensor fault was also specified. In addition to the identification of the deviation, it is also important to diagnose the process abnormality. In other words, it is critical to understand which variables convey meaningful information that can be used to interpret the multivariate deviation. This diagnosis of process faults was the main focus of this chapter.

The well established contribution analysis methodology for fault diagnosis was summarised. A specific focus was on the three main contribution analysis methods i.e., Complete decomposition (CD), Partial decomposition (PD) and Reconstruction-based (RB) contribution analysis. The objective of these approaches is to calculate an individual importance for each process variable that quantify how much each variable contribute to the multivariate deviation. It is well documented in literature that contribution analysis suffers from a condition known as fault smearing. This condition makes it difficult to isolate the individual contribution of a process variable on the deviation observed.

A simulation study, similar to the one used in literature (Van den Kerkhof, Vanlaer, Gins and Impe, 2013), was presented to illustrate the diagnostic performance of the contribution analysis techniques. The formulation of the simulation study was made to be more stringent compared to those found in literature which are too simplistic and not a good representation of reality. The structure of the proposed simulation model has been derived from a PCA model that was constructed using actual process data. A large number of potential samples from the three fault classes was created for different fault sizes. Only samples that are univariately inside the IOC bounds were considered in the specification of these vectors. Note that this constraint is omitted in the definition of the simulation studies found in literature (Alcala and Qin (2011)). A subset of the potential samples is identified which are out-of-control based on the different fault identification statistics. The contribution analysis techniques were then utilised to diagnose the faulty vectors.

The diagnostic performance was reported as the number of samples correctly diagnosed for a specific fault size. A fault was correctly diagnosed if the process variable on which the sensor fault was induced attains a relatively high contribution value.

The results of this simulation study suggest that the PD contribution analysis method should be preferred, since its diagnostic results consistently outperform the CD and RB methods. Given that all of the samples used in the simulation study are multivariately out-of-control at a specified confidence level, it is important that the diagnostic performance is acceptable for all of the fault sizes considered. This is a sensible requirement since in practice the size and dimension of the sensor fault induced is not known. For the smaller fault sizes, poor diagnostic performance was consistently observed throughout the simulation study i.e., for CD, PD and RB contribution analysis across all fault types and fault detection statistics. From this observation it is concluded that the traditional diagnostic techniques i.e., CD, PD and RB contribution analysis, are not suitable for application in industry. A section that highlights recent research that aims to improve on traditional contribution analysis is provided. Common to all of these techniques is the attempt to assign an individual importance to process variables. In Chapter 3 it will be demonstrated that the individual ranking of variables can lead to multivariate fault interpretations that are misleading.

The information captured in this chapter act as the motivation that a new methodology is required for the diagnosis of multivariate process faults.

Chapter 3

New Methodology for PCA Fault Diagnosis

3.1 Introduction

The fault identification step in multivariate statistical process monitoring is of little value to plant engineers if it is not accompanied by a diagnostic analysis that provide a clear interpretation of why the process is identified as being statistically different from multivariate common cause variability. In the previous chapter, PCA was introduced as the most popular technique used for multi-variable process fault detection and diagnosis. The discussion around the contribution analysis methodology, often employed for diagnostic purposes, illustrated that a different method is required to assist in providing insight into multivariate process faults.

The purpose, therefore, of this chapter is to present an alternative methodology for PCA fault diagnosis. It will be argued that the question which traditional diagnostic approaches attempt to answer is not appropriate. The new diagnostic method motivated is accompanied with a new way of thinking when it comes to the diagnosis of PCA process faults.

3.2 Motivation

The common thread in the PCA fault diagnosis methodologies discussed so far is that each of the techniques attempts to calculate an individual importance measure that can be used to identify and rank the variables responsible for the process drift. It was learnt from the underlying theory and simulation study that correlation has a noticeable effect on the existing fault diagnosis methodologies. For example, in the simulation study presented in Chapter 2 it was observed that the current fault diagnosis techniques confound variables with one another that are highly correlated. This is especially clear when the

fault sizes were relatively small.

Another reason why these approaches are not appropriate for fault diagnosis can be illustrated when considering Figure 3.1. In this illustration two highly correlated variables, x_1 and x_2 are plotted. It will be assumed that these two variables fully describe the process being monitored. The data points representing the IOC data are grey in color and are bounded by the red 95% T^2 confidence ellipse. The blue point, $\mathbf{x}_{faulty}$ is an example of an out-of-control observation. From previous discussions, the out-of-control point can be decomposed as the sum of an in control point ($\mathbf{x}_{ioc}$) plus some error vector ($\boldsymbol{\epsilon}$) i.e.,

$$\mathbf{x}_{faulty} = \mathbf{x}_{ioc} + \boldsymbol{\epsilon}. \tag{3.1}$$

In the bottom graph of Figure 3.1 it is observed that the faulty blue point is generated as follows

$$\mathbf{x}_{faulty} = \mathbf{x}_{ioc}^{\blacktriangle} + \begin{bmatrix} f & 0 \end{bmatrix}^T, \tag{3.2}$$

where $\mathbf{x}_{ioc}^{\blacktriangle}$ represents the in control point marked with a triangle (▲) in the graph. The size of the fault f is represented by the horizontal broken line for x_1. Similarly the top graph in Figure 3.1 shows that it is possible to arrive at the blue point using the following configuration

$$\mathbf{x}_{faulty} = \mathbf{x}_{ioc}^{\blacksquare} + \begin{bmatrix} 0 & f^* \end{bmatrix}^T, \tag{3.3}$$

with $\mathbf{x}_{ioc}^{\blacksquare}$ representing the in control observation marked with a square (■) and f^* represents the fault size given by the vertical broken line for x_2. Therefore, Figure 3.1 illustrates that it is possible to arrive at the same out-of-control point by inducing faults on different variables. Hence, with information on the out-of-control point only, it is impossible to attach high importance to either variable, x_1 or x_2, as being responsible for the process deviation. The contribution and reconstruction methods discussed so far can therefore never be expected to correctly identify the process variables that were the cause of the out-of-control state. A new way of thinking and approach are therefore required in the diagnosis part of PCA based multivariate statistical process monitoring.

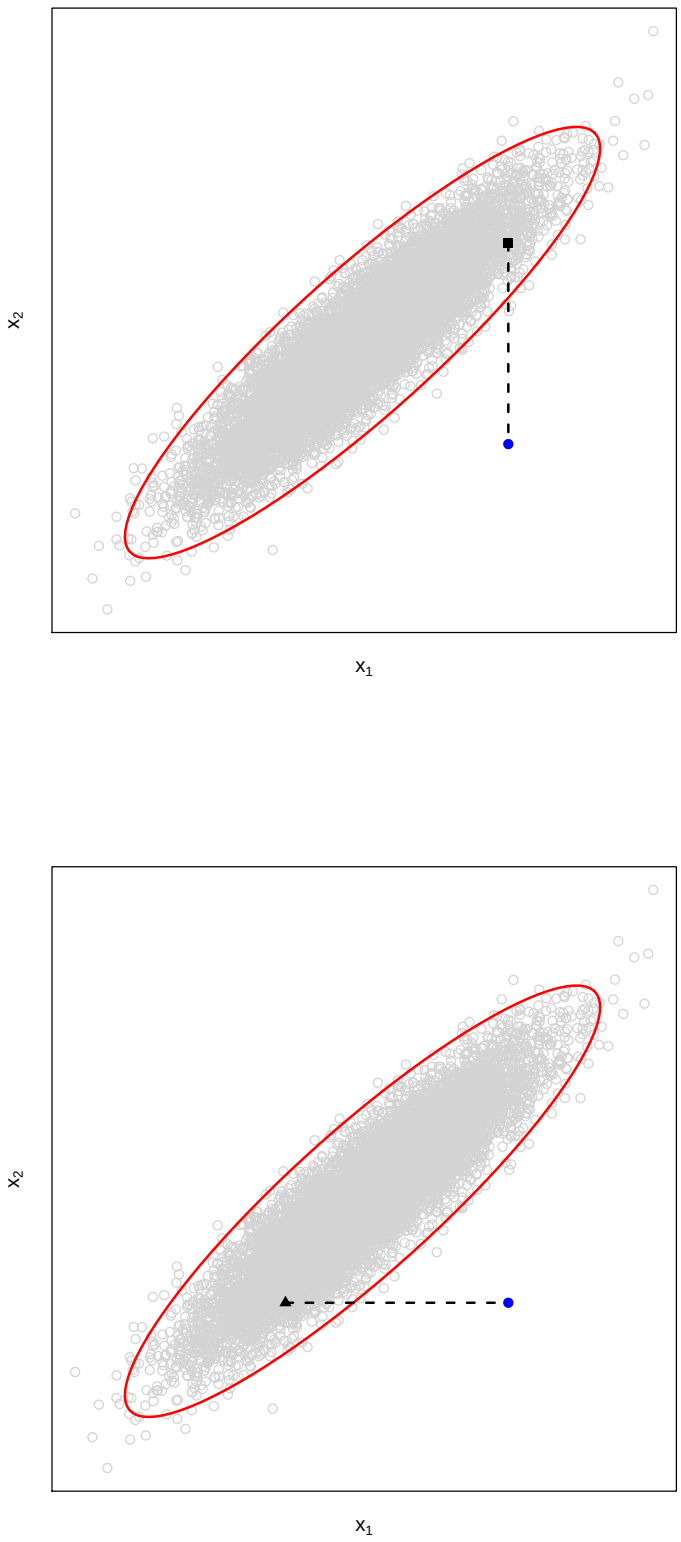


Figure 3.1: Illustration in two dimensions of why it is incorrect to attribute multivariate process deviations to only one process variable.

This motivational illustration is made in two dimensions. It can be demonstrated that it holds true for higher dimensional spaces. An example will now be given to illustrate that similar insight can be gained using the simulated seven dimensional IOC data of Chapter 2. The general process of identifying a single sensor fault example that can be generated from two different in-control samples was:

1. Consider a single sensor fault,

$$\mathbf{x}^{0f} = \mathbf{x}^0 + f^0\boldsymbol{\xi}_k, \tag{3.4}$$

   where $\mathbf{x}^0 : p \times 1$ is a multivariate in-control vector and f^0 is the fault size that is added in the k^{th} position.

2. Search for a vector $\mathbf{x}^*$ that is equal to $\mathbf{x}^0$ except for the k^{th} and j^{th} position, $j \neq k$. The k^{th} element of $\mathbf{x}^*$ equals $x_k^0 + f^0$. The j^{th} element is selected such that $\mathbf{x}^*$ is an in-control vector.

3. Determine the magnitude f^* that can be added to the j^{th} element of $\mathbf{x}^*$ such that the result equals the j^th element of $\mathbf{x}^{0f}$.

This implies that, if such an $\mathbf{x}^*$ can be found, then the single sensor fault $\mathbf{x}^{0f}$ can be generated using both $\mathbf{x}^0$ and $\mathbf{x}^*$. Using this approach resulted in the following single sensor fault vector based on the simulated IOC data of Chapter 2:

$$\mathbf{x}^{0f} = \begin{bmatrix} -2.0895 & 0.0373 & -2.1604 & -1.1621 & 1.1006 & 0.7166 & 0.5563 \end{bmatrix}^T. \tag{3.5}$$

This single sensor fault is identified as being out-of-control using the SPE statistic. The in-control vectors that can be used to construct this single sensor fault are:

$$\mathbf{x}^0 = \begin{bmatrix} -2.0895 & 0.0373 & -2.1604 & 1.7379 & 1.1006 & 0.7166 & 0.5563 \end{bmatrix}^T \tag{3.6}$$

and

$$\mathbf{x}^* = \begin{bmatrix} 0.8573 & 0.0373 & -2.1604 & -1.1621 & 1.1006 & 0.7166 & 0.5563 \end{bmatrix}^T. \tag{3.7}$$

Adding $f^0 = -2.9$ to the fourth entry of the first in-control sample will yield the identified single sensor fault. The single sensor fault is also obtained by adding $f^* = -2.9468$ to the first entry of the second in-control sample. The discussion around Figure 3.1 has

therefore been demonstrated to hold true in this seven dimensional simulation data set. It is therefore assumed that this problem will be present for any amount of process variables.

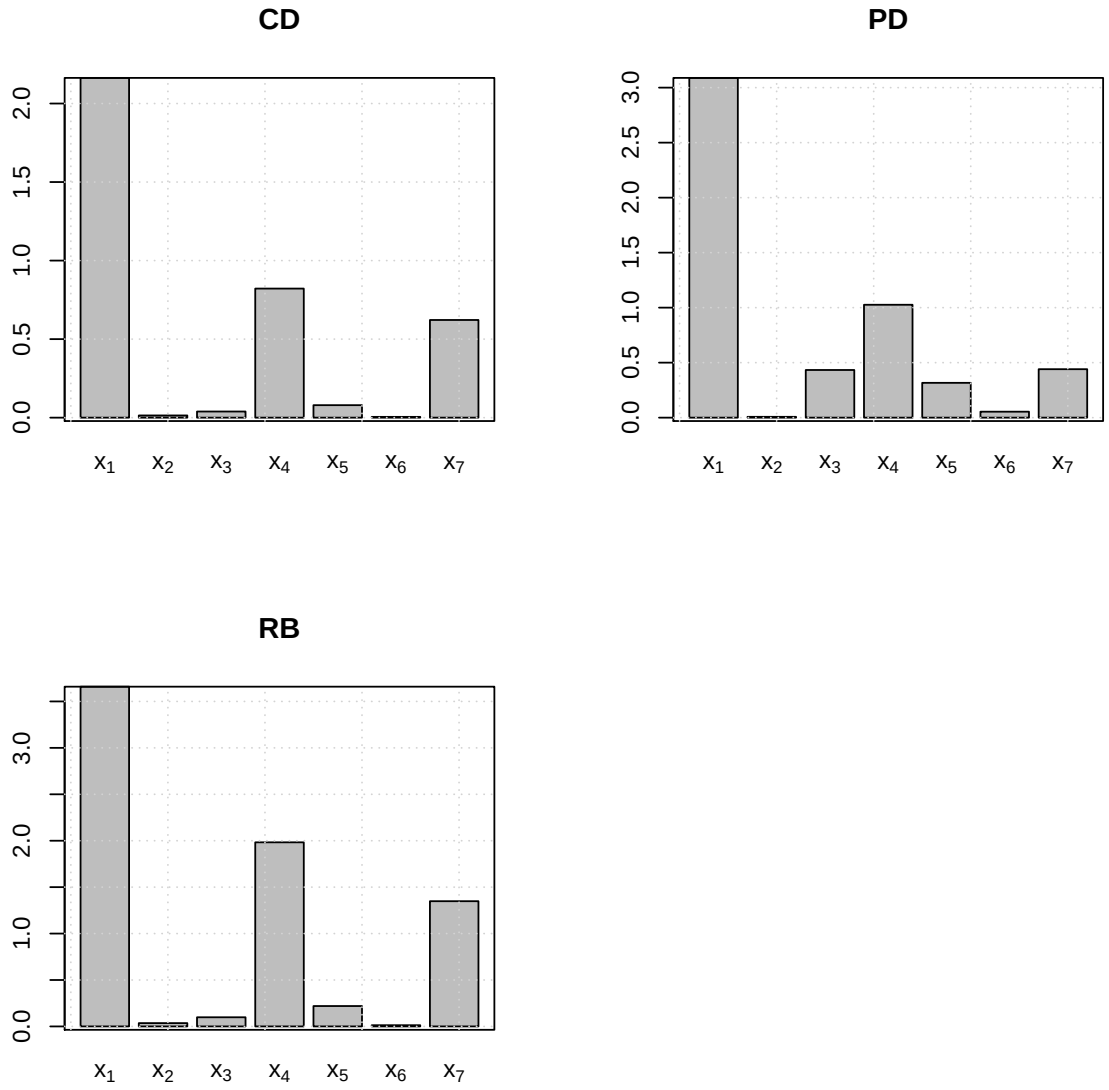


Figure 3.2: Contribution analysis of the single sensor fault example in (3.5) that can be constructed by inducing single sensor faults in different process variables. Complete decomposition, Partial decomposition and Reconstruction based contributions are presented.

The traditional contribution analysis techniques, that were discussed in Chapter 2, were applied to diagnose the single sensor fault in (3.5). The results of this analysis is summarised by the bar charts in Figure 3.2. In all of the graphs, for the three contribution analysis techniques, it is observed that the contribution of the first process variable is approximately double that of the second most important variable. Therefore, the contribution analysis will yield confusing results if the single sensor fault (3.5) was constructed

by inducing a deviation in the fourth variable of (3.6).
Ideally in this situation it is sensible to expect that the contribution analysis technique should produce a result that considers the first and fourth variables as equally important. Since it is impossible to deduce from the limited information which variable was altered to induce the fault. Traditional contribution analysis techniques will always consider one variable as being more important than the other. In the next section pairwise variable contributions are introduced to address this shortcoming.

3.3 Pairwise variable contribution analysis

Motivated by the discussion around Figure 3.1 it is proposed, as an alternative diagnostic approach, that the fault diagnosis methodologies should focus on ranking the importance of pairs of variables instead of individual variables. Consider the following 3 dimensional example. Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$ represent an in-control or IOC sample and let $\mathbf{z} = \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}^T$ be an out-of-control observation. Further suppose that the out-of-control state was achieved by a fault of size $f > 0$ in the first position of $\mathbf{x}$ i.e.,

$$\begin{aligned} \mathbf{z} &= \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}^T & (3.8) \\ &= \begin{bmatrix} x_1 + f & x_2 & x_3 \end{bmatrix}^T. & (3.9) \end{aligned}$$

The vector $\mathbf{z}$ is therefore a single sensor deviation from $\mathbf{x}$. It will now be assumed that the T^2 statistic, as an example, was used to identify that $\mathbf{z}$ was extreme. Therefore, from equation (2.11), the extreme observed T^2 index value for $\mathbf{z}$ equals

$$\begin{aligned} T^2_{\mathbf{z}} &= \mathbf{z}^T \mathbf{V} \mathbf{\Lambda^{-1}} \mathbf{V}^T \mathbf{z} & (3.10) \\ &= \mathbf{z}^T \mathbf{D} \mathbf{z} & (3.11) \\ &> \chi^2_3, & (3.12) \end{aligned}$$

where

$$\begin{aligned} \mathbf{D} &= \hat{\mathbf{S}}^{-1} & (3.13) \\ &= \mathbf{V} \mathbf{\Lambda^{-1}} \mathbf{V}^T, & (3.14) \end{aligned}$$

for notational convenience. In a fashion similar to the variable contribution techniques, the T^2 value of $\mathbf{z}$ is decomposed into a sum of parts, each part uniquely associated with a variable pair. This is achieved by using the quadratic form as follows:

$$
\begin{aligned}
T^2_{\mathbf{z}} &= \mathbf{z}^T\mathbf{D}\mathbf{z} && (3.15)\\
&= \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} && (3.16)\\
&= \sum_{i=1}^{3}\sum_{j=1}^{3} d_{ij} z_i z_j && (3.17)\\
&= \sum_{i=1}^{3} \left(d_{i1} z_i z_1 + d_{i2} z_i z_2 + d_{i3} z_i z_3 \right) && (3.18)\\
&= d_{11}z_1z_1 + d_{12}z_1z_2 + d_{21}z_2z_1 + d_{22}z_2z_2 + \\
&\quad d_{11}z_1z_1 + d_{13}z_1z_3 + d_{31}z_3z_1 + d_{33}z_3z_3 + \\
&\quad d_{22}z_2z_2 + d_{23}z_2z_3 + d_{32}z_3z_2 + d_{33}z_3z_3 - \\
&\quad d_{11}z_1z_1 - d_{22}z_2z_2 - d_{33}z_3z_3 && (3.19)\\
&= \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} z_1 & z_3 \end{bmatrix} \begin{bmatrix} d_{11} & d_{13} \\ d_{31} & d_{33} \end{bmatrix} \begin{bmatrix} z_1 \\ z_3 \end{bmatrix} + \\
&\quad \begin{bmatrix} z_2 & z_3 \end{bmatrix} \begin{bmatrix} d_{22} & d_{23} \\ d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} z_2 \\ z_3 \end{bmatrix} - \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix} \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} && (3.20)\\
&= \sum_{i<j}^{2}\sum^{3} \mathbf{z}^T_{(i,j)} \mathbf{D}(i,j) \mathbf{z}_{(i,j)} - \mathbf{z}^T diag(\mathbf{D}) \mathbf{z}, && (3.21)
\end{aligned}
$$

where $\mathbf{z}_{(i,j)} = \begin{bmatrix} z_i & z_j \end{bmatrix}^T$ with $i \neq j$ and

$$
\mathbf{D}(i,j) = \begin{bmatrix} d_{ii} & d_{ij} \\ d_{ji} & d_{jj} \end{bmatrix}, \qquad (3.22)
$$

is a 2×2 sub matrix of the inverse of the PCA approximation of the covariance matrix $\mathbf{S}$. This decomposition therefore implies that each of the $\binom{3}{2}$ variable pairs can be considered separately. It can be assumed that each pair represents a newly created variable. An intuitive way of defining the contribution of the variable pairs to the $T^2_{\mathbf{z}}$ is given by the Mahalanobis-like distance

$$T^2_{\mathbf{z}_{(i,j)}} = \mathbf{z}^T_{(i,j)}\mathbf{D}(i,j)\mathbf{z}_{(i,j)}. \tag{3.23}$$

Note, from (3.21) that the contribution of these newly created variables i.e., $\mathbf{z}_{(i,j)}$, is not smeared with information on the other new variables.

On closer inspection of these contributions, it is observed that the term

$$\begin{aligned} T^2_{\mathbf{z}(2,3)} &= \mathbf{z}^T_{(2,3)}\mathbf{D}(2,3)\mathbf{z}_{(2,3)} && (3.24)\\ &= \mathbf{x}^T_{(2,3)}\mathbf{D}(2,3)\mathbf{x}_{(2,3)}, && (3.25) \end{aligned}$$

represents a contribution that contain only information of the non-faulty variables 2 and 3. This term will therefore contribute little to the size of the extreme $T^2_{\mathbf{z}}$ value. The size of $T^2_{\mathbf{z}}$ will therefore largely be determined by the size of

$$T^2_{\mathbf{z}(1,2)} = \mathbf{z}^T_{(1,2)}\mathbf{D}(1,2)\mathbf{z}_{(1,2)}, \tag{3.26}$$

and

$$T^2_{\mathbf{z}(1,3)} = \mathbf{z}^T_{(1,3)}\mathbf{D}(1,3)\mathbf{z}_{(1,3)}, \tag{3.27}$$

i.e., the contributions that contain the faulty observation, z_1, on variable 1. It is also useful to note, from (3.21), that

$$\begin{aligned} T^2_{\mathbf{z}} &> \chi^2_3 && (3.28)\\ \sum_{i<j}^{2}\sum^{3}\mathbf{z}^T_{(i,j)}\mathbf{D}(i,j)\mathbf{z}_{(i,j)} - \mathbf{z}^T diag(\mathbf{D})\mathbf{z} &> \chi^2_3 && (3.29)\\ \Longrightarrow \sum_{i<j}^{2}\sum^{3}\mathbf{z}^T_{(i,j)}\mathbf{D}(i,j)\mathbf{z}_{(i,j)} &> \chi^2_3. && (3.30) \end{aligned}$$

Equation (3.30) follows from the fact that the diagonal entries of $\mathbf{D} = \mathbf{S}^{-1}$ are positive, which implies that the quadratic form $\mathbf{z}^T diag(\mathbf{D})\mathbf{z}$ is greater than zero. Therefore, the diagnosis can be performed by only focussing on the $T^2_{\mathbf{z}(i,j)}$ distances. The fault diagnosis procedure is now similar to the fault detection phase which is an analysis of quadratic form distance measures. Up till now the focus was on the three dimensional example for

ease of illustration. The logic also holds true for the p-dimensional case e.g., in $p > 2$ dimensions equation (3.21) becomes

$$T^2_{\mathbf{z}} = \sum_{i<j}^{p-1}\sum^{p}\mathbf{z}^T_{(i,j)}\mathbf{D}(i,j)\mathbf{z}_{(i,j)} - (p-2)\mathbf{z}^T diag(\mathbf{D})\mathbf{z}. \tag{3.31}$$

It should also be noted that the T^2 statistic was used as an example, the decomposition can be applied to any of the PCA fault detection indices described in Section 2.3.

To summarise, the following definition is developed to define pairwise contribution analysis for PCA based MSPM.

Definition 3.1 *The pairwise contribution of process variable i and j for $j \neq i$ is defined as*

$$\mathbf{z}^T_{(i,j)}\mathbf{A}(i,j)\mathbf{z}_{(i,j)}, \tag{3.32}$$

where

- $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T$ *for the T^2 statistic (see (2.11)),*
- $\mathbf{A} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^T$ *for the SPE statistic (see (2.9)),*
- $\mathbf{A} = \mathbf{\Phi}$ *, as defined in (2.13), for the Combined statistic*

and

$$\mathbf{A}(i,j) = \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}, \tag{3.33}$$

is the 2×2 dimensional sub matrix of $\mathbf{A}$. The value a_{ij} represents the ij^{th} entry of $\mathbf{A}$.

Using good process understanding together with the relative sizes of the pairwise distances associated with each pair of variables will guide the engineer to understand why the process being monitored is out-of-control. This approach to fault diagnosis also addresses an interpretation limitation that accompanies classical single variable PCA contribution analysis. It was argued in Section 2 that multivariate statistical process control techniques are applied to identify process characteristic changes when all the process variables are

operating inside their univariate bounds. Hence, when diagnosing a multivariate fault it should be assumed that univariately the process is in-control.

The results of the diagnostic analysis should therefore not provide a result that only gives univariate importance to process variables. Practically a process engineer would not be able to use this univariate ranking of process variables to fix a multivariate problem. The pairwise contribution technique gives a multivariate answer to a multivariate diagnostic problem that will make understanding the process deviation easier. In the 3-dimensional toy illustration, the engineer would therefore consider the relationships within the larger contributing process variable pairs $(1, 2)$ and $(1, 3)$ to identify what needs to be done, inside the univariate limits of these variables, to aid the process in moving back to being within control. In the next two sections guidance and statistical methods will be provided on how these pairwise contributions can be utilised to enable an in-control process.

3.4 Pairwise contribution importance

Even though pairwise contribution analysis presents a more sensible way of performing fault diagnostic investigation it still has some shortcomings. One obvious disadvantage is that instead of analysing p variable contributions, it is now required that $\binom{p}{2}$ contributions be analysed. Analysing these pairwise contributions therefore becomes impractical when the number of variables monitored is large e.g., when $p > 4$.

One way to simplify the interpretation of the result is to give importance only to the contributions that are determined to be statistically significant. A quantification of what constitute a significant contribution can be obtained by considering the fault detection statistic evaluations of the IOC observations. Building on the discussion so far, the T^2 statistic will again be used as an example. Let

$$T^2_{\mathbf{x}^k_{(i,j)}} = (\mathbf{x}^k_{(i,j)})^T \mathbf{D}(i,j)\mathbf{x}^k_{(i,j)}, \tag{3.34}$$

represent the contribution of the $(i, j)^{th}$ variable pair for the k^{th} IOC sample. The individual contributions of the $(i, j)^{th}$ variable combination for all of the IOC samples can therefore be collected into one vector. Density estimation can then be applied to this vector of observed IOC contributions to summarise their distribution under ideal process conditions. This distribution can then be used as reference against which the contributions calculated for future observations are compared. This method will therefore greatly minimize the number of variable pairs that need to be considered for a specific out-of-control observation.

3.5 Variable importance in pairwise contributions

It was illustrated earlier that it is impossible to isolate the faulty variable using the two dimensional example in Figure 3.1. From this example it was concluded that providing a univariate ranking of process variables is not appropriate or optimal in the multivariate setting. However, a ranking of process variables might still prove useful in combination with the important pairwise contributions to assist in the diagnosis step.

A relative ranking of process variables can be obtained when the complete set of pairwise contributions, for a specific out-of-control point, is analysed simultaneously. For the moment again consider the three dimensional example discussed thus far. In this example the problem was constructed so that variable one is faulty. It is therefore expected that the variable combinations (1,2) and (1,3) have the largest contributions. From this information, using the frequency of the number of times that variable one is present in the significant pairwise variable combinations, it can be deduced that variable one most likely played an important role in the process drift. Therefore, the correct diagnosis is made by using the 'interaction' of variable one with a number of the other variables. It is again obvious that analysing the multiple variable pairs in this manner becomes tedious when the number of pairs is large and when the fault dimension is high.

The simultaneous analysis of all of these pairwise contributions can be achieved by realising that this problem can be reformulated as the analysis of a distance matrix. Consider the following symmetric matrix:

$$\boldsymbol{\Delta} : p \times p \quad = \quad \begin{bmatrix} 0 & T^2_{\mathbf{z}(1,2)} & T^2_{\mathbf{z}(1,3)} & \dots & T^2_{\mathbf{z}(1,p)} \\ T^2_{\mathbf{z}(2,1)} & 0 & T^2_{\mathbf{z}(2,3)} & \dots & T^2_{\mathbf{z}(2,p)} \\ T^2_{\mathbf{z}(3,1)} & T^2_{\mathbf{z}(3,2)} & 0 & \dots & T^2_{\mathbf{z}(3,p)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T^2_{\mathbf{z}(p,1)} & T^2_{\mathbf{z}(p,2)} & T^2_{\mathbf{z}(p,3)} & \dots & 0 \end{bmatrix}, \tag{3.35}$$

with $T^2_{\mathbf{z}(i,j)}, i \neq j$ the pairwise contribution calculated for the variable pair (i, j), (3.23). The diagonal entries of this matrix are zero. The process is therefore operating at the ideal operating conditions if all of the entries in $\boldsymbol{\Delta}$ are relatively small in some way i.e., as identified using the pairwise contribution importance method described in the previous section.

Given this matrix, assume now that it represents the dissimilarity matrix of a set of pseudo variables $1^*, 2^*, \dots, p^*$. Each row/column identifies a specific pseudo variable and each off-diagonal matrix entry specifies the dissimilarity between the corresponding pseudo variables. That is, assume that the pairwise contribution of variable i and j i.e.,

$T^2_{\mathbf{z}(i,j)}, i \neq j$, represents the dissimilarity between pseudo variables i^* and j^*. If the dissimilarities between these pseudo variables are small i.e., if the pseudo variables are similar to each other, then the process will be operating at the IOC's. A graphical exploration of this matrix can now be performed by using the multidimensional scaling (MDS) method (Cox and Cox, 2000) to observe the similarities between these pseudo measures. Specifically, metric multidimensional scaling can be utilised to obtain a 2-dimensional display consisting of a set of p^* points, each representing one of the p^* pseudo variables. The MDS visualisation is optimally constructed such that the Euclidean distances between the points closely resemble the original dissimilarities $T^2_{\mathbf{z}(i,j)}, i \neq j$. More precisely, the distances between the points approximate the original dissimilarities as follows

$$d_{i^*,j^*} \approx f(T^2_{\mathbf{z}(i,j)}), \tag{3.36}$$

where the function f is continuous, parametric and monotonic. Classical scaling is a special case of metric multidimensional scaling where the function f is equal to the identity function. Also the calculated Euclidean distances exactly equals the original dissimilarities. Given a dissimilarity matrix $\mathbf{\Delta}$, (3.35), classical scaling treats the original dissimilarities as if it is Euclidean. The coordinates of the points representing the pseudo variables are obtained as follows:

- Double center the squared dissimilarity matrix

$$\mathbf{B^{\Delta}} : p^* \times p^* \;=\; -\frac{1}{2}\mathbf{J\Delta J}, \tag{3.37}$$

  where $\mathbf{J} : p^* \times p^* = \mathbf{I}_{p^*} - \frac{1}{p^*}\mathbf{1}\mathbf{1}^T$. The double centering ensures that the final multidimensional scaling display is centered about zero.

- Calculate the spectral decomposition of $\mathbf{B^{\Delta}}$ i.e.,

$$\mathbf{B^{\Delta}} \;=\; \mathbf{Q^{\Delta}\Lambda^{\Delta}}(\mathbf{Q^{\Delta}})^T. \tag{3.38}$$

- Consider only the eigenvector eigenvalue pairs that correspond to the positive eigenvalues. Let $\mathbf{\Lambda^{\Delta}_+} : m \times m$ represent the positive diagonal eigenvalue matrix and let $\mathbf{Q^{\Delta}_+} : p^* \times m$ represent the matrix with the corresponding eigenvectors. The classical scaling coordinate matrix $\mathbf{X^{\Delta}} : p^* \times m$ representing the multidimensional scaling points is then obtained as

$$\mathbf{X}^{\mathbf{\Delta}} = \mathbf{Q}_{+}^{\mathbf{\Delta}}(\mathbf{\Lambda}_{+}^{\mathbf{\Delta}})^{\frac{1}{2}}. \tag{3.39}$$

Note that since the dissimilarities captured in (3.35) are not Euclidean, it is not necessarily ensured that the distances between the points in the display will be Euclidean. In order for this to be guaranteed, it is required that $\mathbf{B}^{\mathbf{\Delta}}$, in (3.37), should be positive semi-definite. If this is not the case then it is required that the off-diagonal entries in (3.35) should be appropriately altered using the method prescribed in Cox and Cox (2000). This will result in a positive semi-definite matrix.

In the multidimensional scaling (MDS) display, the coordinates, representing the pseudo variables, that are more remote from the center of the display are indicative of the corresponding process variables that are part of the variable pairs that have significant pairwise contributions. That is, if pseudo variable j^* is far away from the center then process variable j will be important in the understanding of the process fault. Therefore, by evaluating the MDS representation of the dissimilarity matrix of a faulty process observation the observer can visually determine a ranking of the process variables.

Another approach that can be used to obtain a ranking of the process variables, that is more direct, is to simply consider the row (or column) sums of $\mathbf{\Delta}$, (3.35). It was argued that the process being monitored is operating ideally if all of the pseudo variables can be considered similar in some way. Therefore, if all pseudo variables are similar, the sum of a specific variable's dissimilarities over all other variables should also be relatively small. A pseudo variable i^* with a larger than expected row sum would therefore indicate that process variable i should have a high ranking. Formally, a relative ranking of process variables based on the row sums of the distance matrix (3.35) is defined as follows:

- Process variable i will achieve the highest ranking if its corresponding pseudo variable i^* solves

$$\arg\max_{i^*} \quad T^2_{\mathbf{z}_{(i^*,\cdot)}}, \tag{3.40}$$

  with $T^2_{\mathbf{z}_{(i^*,\cdot)}} = \sum_{j^*=1}^{p^*} T^2_{\mathbf{z}_{(i^*,j^*)}}$ for $i^* = 1, \ldots, p$ the sum of the $(i^*)^{th}$ row elements.

- To identify the second most important variable, it is required, for obvious reasons, that the information on the previous identified process variable be removed from (3.35). Therefore, after removing row i^* and column i^*, the second most important variable is identified as the one that corresponds with the maximum row sum calculated using the reduced matrix i.e., the pseudo variable $k^* \neq i^*$ that solves

$$\arg\max_{k^* \neq i^*} \quad \sum_{j^* \neq i^*}^{p} T^2_{\mathbf{z}_{(k^*,j^*)}}, \tag{3.41}$$

will identify process variable k as the second most important.

- This logic is repeated until a ranking is obtained for $p-2$ of the process variables. Note that the final two process variables will attain the same ranking.

As in the case of the size quantification of the pairwise contributions, the IOC data can be used to specify null distributions for each row of the distance matrix. The process of identifying important variables could therefore be stopped, at any point, when no one of the row sums in the matrix is significantly large.

It is important that these relative rankings are not reported in isolation. Interpretation of the out-of-control condition should be done using both the significant pairwise contributions and the relative rankings of the process variables.

3.6 Deployment

The following are some considerations when implementing the pairwise methodology in practice:

- **Model construction:** The model construction step is performed off-line. This typically involves:
  - Specifying the process variables of interest to the problem.
  - Reference set definition to capture IOC common cause variability using historic data.
  - PCA model specification. Information are captured that are required to specifiy the fault detection statistics. The associated control limtis are also calculated.
  - Calculation of pairwise control limits. For each observation vector in the IOC reference data the corresponding distance matrix representation i.e., the matrix with off diagonal entries equal to (3.32). The entries in all of the matrices that correspond to a specific variable pair are collected into one set. Empirical analysis is then performed to define the control limit for each pair of variables.
  - Calculation of univariate control limits. For each of the reference distance matrices calculate the row (or column) sums. Collect the row sum, from each distance matrix, for a specific row into a combined vector. Use this vector to

define the reference row sum distribution for each process variable. This can be used to specify a control limit to identify the top contributing variable. It is suggested that the specification of the control limits used to specify the second and lower ranked variables be performed on-line. Since the reference row sum distributions of the remaining variables are based on knowing which row and column was removed in the previous step.

- **On-line implementation:** The on-line implementation of the diagnosis methodology involves the following:
  - PCA model information are used to evaluate new observations. Preprocessing is performed and fault identification statistic values are calculated.
  - If the fault identification statistic evaluation of a new observation is extreme then the distance matrix representation of this observation is obtained.
  - The entries of this matrix are compared to the pairwise control limits that were calculated off-line.
  - The pairs of variables that are identified to be significant are reported.
  - Use the reference row sum information to identify an individual ranking of process variables.

The pairwise fault diagnosis approach suggested boils down to the calculation of the symmetric $p \times p$ distance matrix, for each observation, with each entry defined by the quadratic form given by equation (3.32). There is no obvious way of defining this matrix using matrix multiplication. Therefore, the lower (or upper) off diagonal entries of this matrix needs to be iteratively populated. The calculation of the quadratic form in (3.32) is not computationally demanding since it is constructed as a subset of known information. Therefore, computation time is only effected by the number of entries in the matrix. Most of the distance matrix creation will happen off-line during the model building step to determine the reference matrices. Therefore, it is expected that this will also not be computationally straining, since measures such as parallel computation is easily employed off-line. In the on-line setting the calculation of the distance matrices is limited to the out-of-control observations. The cost of computation is therefore not expected to be an obstacle to the on-line implementation.

3.7 PCA fault diagnosis simulation study continued (Part 1)

3.7.1 Objective

The purpose of the discussion presented in this section is to showcase via simulation how this new methodology of using pairwise contributions can be used to diagnose a multivariate out-of-control state. This simulation will again use the information constructed in Section 2.8. The results in this section is limited to the simulation data used. It is envisioned that the illustration will give a good understanding of how to implement the methodology and demonstrate the potential value that it can add to multivariate fault interpretation.

3.7.2 Pairwise contribution results

In the following a variable will be interchangeably referred to as variable i or x_i. Consider the following multivariate sensor fault vector,

$$\mathbf{x}^m = \begin{bmatrix} 2.95 & -1.18 & 2.73 & -0.02 & -0.77 & -0.68 & -2.35 \end{bmatrix}^T, \tag{3.42}$$

generated using the definition given in Section 2.6. The faulty variables for this example are variables 1 and 4. To re-emphasize, please note that all of the variables in this vector satisfy the univariate control limits of the IOC data, refer back to Figure 2.1. This vector was identified by the SPE fault detection index as being different from the ideal conditions. The SPE value for this observation was determined to be equal to 4.39, which is about 0.66 SPE units above the control limit. The SPE pairwise contributions of this faulty observation,

$$SPE_{\mathbf{x}^m_{(i,j)}} = (\mathbf{x}^m_{(i,j)})^T \tilde{\mathbf{C}} \mathbf{x}^m_{(i,j)}, \tag{3.43}$$

where $\tilde{\mathbf{C}} = \tilde{\mathbf{V}}\tilde{\mathbf{V}}^T$ and $i \neq j = 1, \ldots, 7$, was calculated and captured in the following symmetric distance matrix

$$\begin{bmatrix} 0.00 & 5.42 & 8.29 & 4.99 & 4.96 & 5.23 & 3.06 \\ 5.42 & 0.00 & 4.44 & 0.54 & 0.09 & 0.58 & 4.22 \\ 8.29 & 4.44 & 0.00 & 2.98 & 2.55 & 1.78 & 3.34 \\ 4.99 & 0.54 & 2.98 & 0.00 & 0.19 & 0.16 & 2.46 \\ 4.96 & 0.09 & 2.55 & 0.19 & 0.00 & 0.44 & 2.09 \\ 5.23 & 0.58 & 1.78 & 0.16 & 0.44 & 0.00 & 3.18 \\ 3.06 & 4.22 & 3.34 & 2.46 & 2.09 & 3.18 & 0.00 \end{bmatrix}, \tag{3.44}$$

i.e., a construction similar to (3.35) for the SPE statistic. Each matrix element therefore provides the pairwise contribution of the variable pair associated with the specific row and column. In order to get an appreciation of the size of each element in the matrix, it is required that the elements are compared to their reference distributions. In Section 3.4 it was explained how the IOC data can be used to obtain a reference distribution for each of the pairwise contributions. The pairwise reference contribution density estimates, for this example, are displayed in Figure 3.3. By using the contribution reference distributions, an empirically determined p-value can be associated with each of the off-diagonal elements of this distance matrix. This occurrence probability information of the observed pairwise contributions was calculated and is captured in the following matrix

$$\begin{bmatrix} 0.0000 & 0.0042 & 0.0005 & 0.0000 & 0.0058 & 0.0074 & 0.0068 \\ 0.0042 & 0.0000 & 0.0063 & 0.5005 & 0.6465 & 0.4489 & 0.0047 \\ 0.0005 & 0.0063 & 0.0000 & 0.0358 & 0.0263 & 0.0400 & 0.0111 \\ 0.0000 & 0.5005 & 0.0358 & 0.0000 & 0.7550 & 0.7861 & 0.0558 \\ 0.0058 & 0.6465 & 0.0263 & 0.7550 & 0.0000 & 0.5411 & 0.0638 \\ 0.0074 & 0.4489 & 0.0400 & 0.7861 & 0.5411 & 0.0000 & 0.0269 \\ 0.0068 & 0.0047 & 0.0111 & 0.0558 & 0.0638 & 0.0269 & 0.0000 \end{bmatrix}. \tag{3.45}$$

The smaller the p-value of a specific contribution the higher its importance in diagnosing the process deviation. The highest pairwise SPE contributions for this faulty observation are summarized by Figure 3.4. These are the contributions which had a p-value less than 0.01. These high ranking pairwise contributions should therefore form the basis of the investigation to eventually understand why the process has deviated. Note that the pairwise contribution of the two variables, 1 and 4, that were manipulated to induce the multivariate fault has the highest importance. In order to correct the process fault, it is suggested that the process engineer consider the pairwise scatter plot of these variables. Figure 3.5 plots the IOC data of x_1 and x_4 against each other. From this graphical

representation, it can visually be deduced that the new observation, represented by the red dot, is different from most of the IOC data points. The engineer should therefore determine, by using fundamental knowledge, how to interpret the pairwise deviations and how the process conditions could be altered to enable x_1 and x_4 to operate within their bivariate IOC reference region.

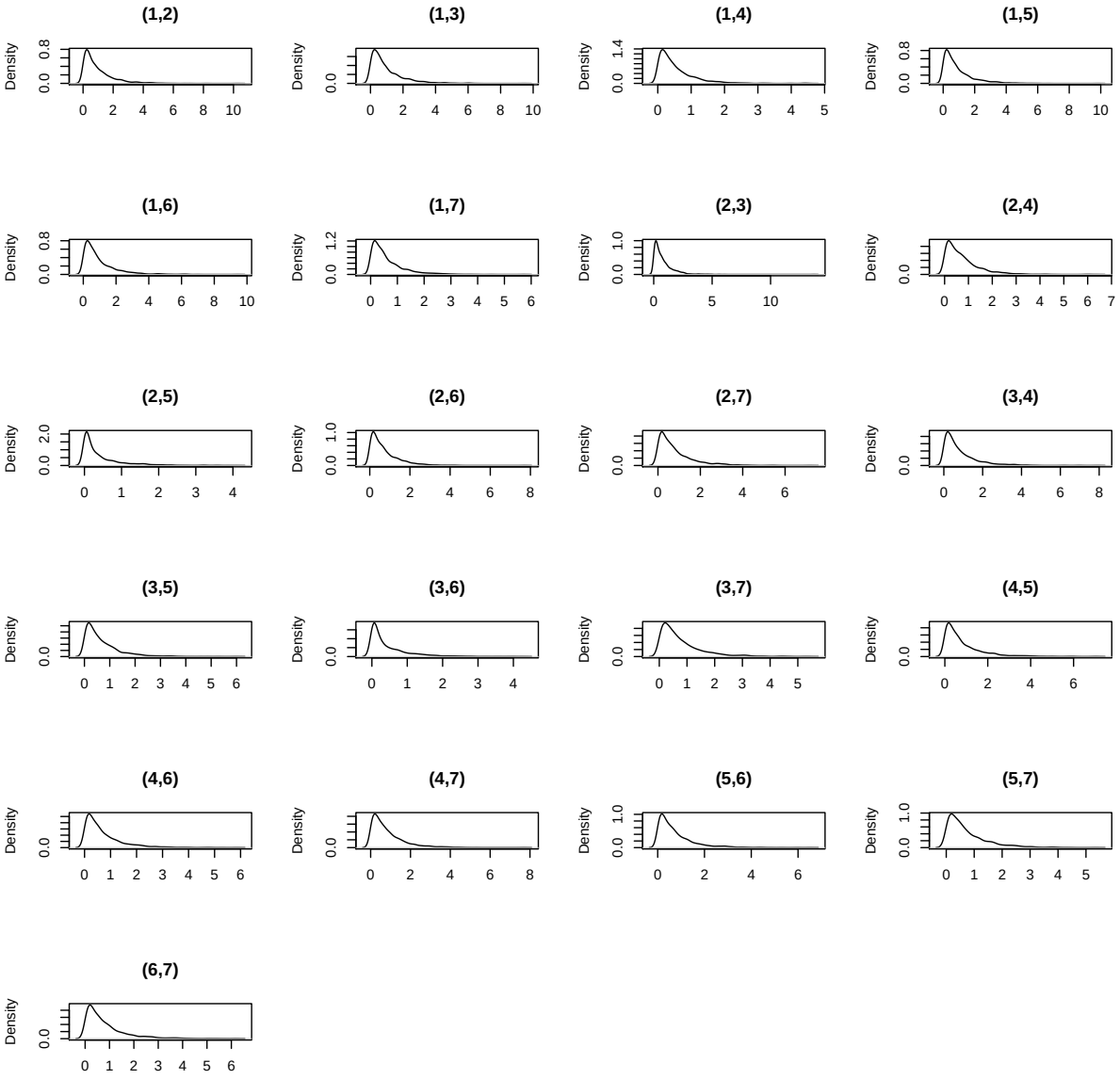


Figure 3.3: Reference distributions for the pairwise SPE contributions.

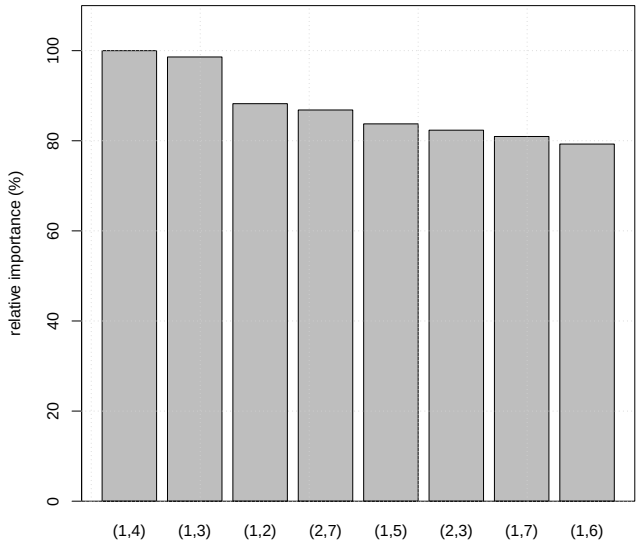


Figure 3.4: Relative importance of the pairwise SPE contributions calculated as $(1-p\text{-}value) \times 100\%$

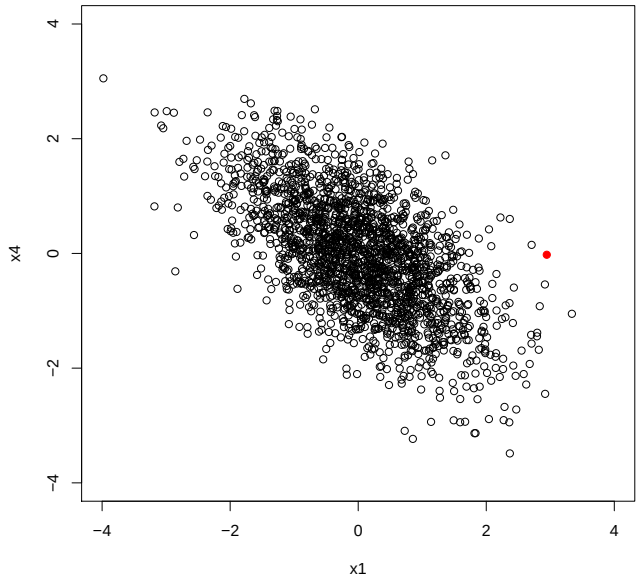


Figure 3.5: Scatter plot of x_1 against x_4. IOC data are represented by open circles. The solid red circle represents the faulty observation.

To further assist in the fault diagnosis, a ranking of the individual process variables has been determined. This was done in two ways. The first approach is by using the MDS method. The 2-dimensional MDS display of the matrix in (3.44) is depicted in Figure 3.6.

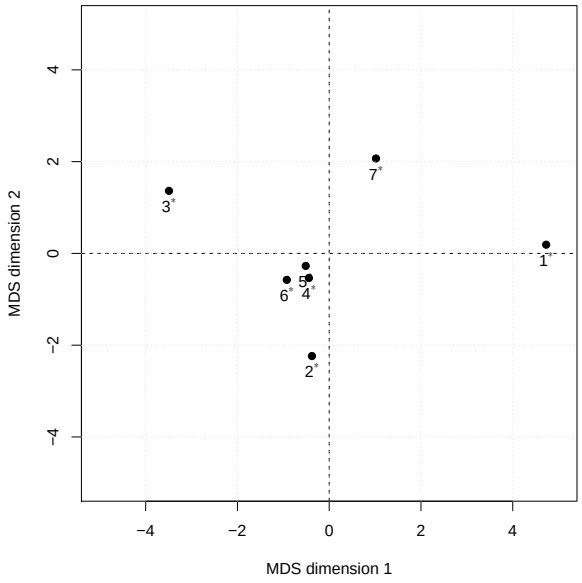


Figure 3.6: MDS display of the SPE distance matrix in (3.44)

In this MDS representation it is observed that pseudo variable 1^* is the most remote from the center of the display. Therefore, process variable 1 is the most important. Other process variables that are identified using this approach are variables 3 and 7.

The second approach in specifying an individual variable importance ranking is by using the row sums method. A step by step illustration will be given to illustrate exactly how this is done. First, to identify the highest ranking process variable, the row sums of the complete 7×7 distance matrix (3.44) is calculated as

$$\begin{bmatrix} 1^* & 2^* & 3^* & 4^* & 5^* & 6^* & 7^* \\ 31.95 & 15.30 & 23.37 & 11.32 & 10.33 & 11.37 & 18.35 \end{bmatrix}. \tag{3.46}$$

The observed row sums are then compared to the IOC row sum distributions. The IOC row sum density estimates for the first iteration are displayed by Figure 3.7. The empirical derived p-values were calculated and is equal to

$$\begin{bmatrix} 1^* & 2^* & 3^* & 4^* & 5^* & 6^* & 7^* \\ 0.002 & 0.014 & 0.002 & 0.047 & 0.054 & 0.040 & 0.009 \end{bmatrix}. \tag{3.47}$$

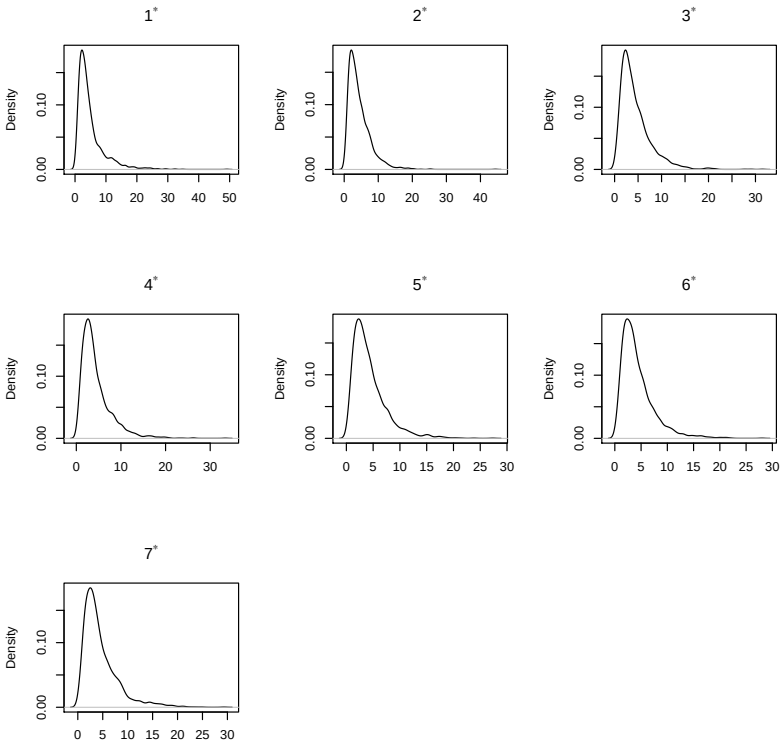


Figure 3.7: Reference row sum densities for the complete distance matrix.

The first and third process variables achieved a tie for the lowest p-value. In this instance x_1 is selected to have the highest rank since it was part of the most important pairwise contribution. To identify the variable with the second highest ranking, the row sums is calculated of the matrix (3.44) with the first row and column removed. Row sums for this reduced matrix equalled

$$\begin{bmatrix} 2^* & 3^* & 4^* & 5^* & 6^* & 7^* \\ 9.88 & 15.08 & 6.34 & 5.37 & 6.14 & 15.29 \end{bmatrix}. \tag{3.48}$$

Figure 3.8 depicts the corresponding IOC densities for the 6×6 distance matrix. P-values for these row sums were calculated as

$$\begin{bmatrix} 2^* & 3^* & 4^* & 5^* & 6^* & 7^* \\ 0.004 & 0.001 & 0.033 & 0.026 & 0.017 & 0.003 \end{bmatrix}. \tag{3.49}$$

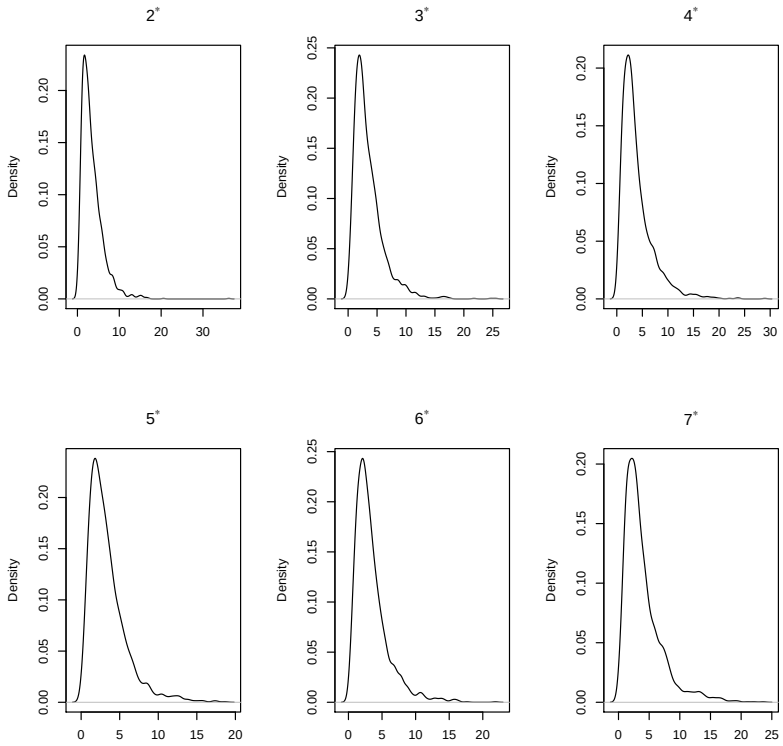


Figure 3.8: Reference row sum densities for the distance matrix with row and column 1 removed.

Process variable x_3 is therefore the second most important. After repeating this process for the remaining variables, the final ranking had the following order:

$$\begin{bmatrix} x_1 & x_3 & x_7 & x_2 & x_6 & x_4 & x_5 \end{bmatrix}, \tag{3.50}$$

with the variable importance level decreasing from left to right. This relative ranking of the process variables can now be combined with the importance ranking of the variable pairs to better understand the drift in the plant conditions. Observe that the faulty variables i.e., x_1 and x_4, are correctly identified by pairwise contribution analysis in Figure 3.4. However, in the individual ranking only x_1 obtain a high rank. This make sense, since the pairwise contributions of x_4 with the other variables i.e., (x_4, x_2), (x_4, x_3), (x_4, x_5), (x_4, x_6) and (x_4, x_7), are of acceptable magnitude.

It is informative to compare the classical contribution analysis results to the variable ranking specified in (3.50). The rankings obtained using the variable contribution techniques yielded the following importance orderings:

$$CD: \quad \begin{bmatrix} x_5 & x_2 & x_1 & x_4 & x_3 & x_6 & x_7 \end{bmatrix}, \tag{3.51}$$

$$PD: \quad \begin{bmatrix} x_1 & x_3 & x_2 & x_5 & x_6 & x_7 & x_4 \end{bmatrix}, \tag{3.52}$$

and

$$RB: \quad \begin{bmatrix} x_5 & x_2 & x_4 & x_1 & x_3 & x_6 & x_7 \end{bmatrix}. \tag{3.53}$$

Interestingly the top PD contributions closely resemble the ranking of the new method. In contrast, the CD and RB rankings are completely different. From this example it is observed that by simply selecting the top contributing variables and manually forming variable pairs to identify the significant ones will be misleading. For example, the RB contribution analysis identifies x_5 and x_2 as the top contributing variables. However, the variable pair (x_2, x_5) is not ranked amongst the significant pairs. Similarly, the pair (x_1, x_4) is identified as important by the pairwise approach, however the PD analysis ranked x_4 as the least important. It is therefore not clear how one would arrive at the identification of (x_1, x_4) as being important from the PD analysis ranking. To conclude, individual rankings on their own lead to interpretations that do not clearly explain the multivariate deviation, and it must be interpreted in combination with the significant pairwise rankings.

3.8 PCA fault diagnosis simulation study continued (Part 2)

3.8.1 Objective

For comparison to the individual variable contribution ranking results of Section 2.8, the univariate diagnosis performance of the new method will now be provided for all of the fault samples generated in the simulation study in Chapter 2. This will be a superficial performance illustration, where only the ranking of the variables as identified by the new method will be considered. It is not recommended that the univariate importance ranking of process variables be considered in isolation, in this manner, when performing fault diagnosis, as motivated in this chapter.

3.8.2 Results

The results of this study are illustrated in Figures 3.9 to 3.17. The following is observed:

- Single sensor faults. For the SPE statistic, in Figure 3.9, it is observed that the performance of the pairwise method is similar to PD contribution analysis and better than CD and RB contribution analysis. For the T^2 statistic, in Figure 3.10, a similar observation is made when compared to that of the SPE statistic. For the $Combined$ statistic, in Figure 3.11, the pairwise technique outperforms the contribution analysis techniques.

- Multiple sensor faults. For all the fault identification statistics i.e., SPE, T^2 and $Combined$, it is observed, in Figures 3.12 to 3.14, that the new method outperforms the traditional techniques.

- Multivariate sensor faults. For the SPE and $Combined$ statistics, in Figures 3.15 and 3.17, the performance of the new technique is slightly inferior to that of the best performing traditional method. Comparable performance, to that of the top performing method, is observed for the T^2 statistic in Figure 3.16.

In summary, the new method either outperforms or yields comparable performance to the best performing traditional technique i.e., partial decomposition contribution. Although the PD method has the highest correct diagnosis rate for multivariate faults, the pairwise contribution analysis developed provides a multivariate interpretation of multivariate deviations, in addition to the univariate ranking of the variables, which has reduced fault smearing in most cases. It is interesting to note that in each of these graphs that the correct diagnostic rate improves with an increase in fault magnitude for the new method. Conceptually, this is expected. If the fault size increase than more variable pairs that include the faulty variable will start to deviate. This will increase the row sum of the faulty variable and increase the chance to detect the specific variable.

3.8.3 Single sensor fault

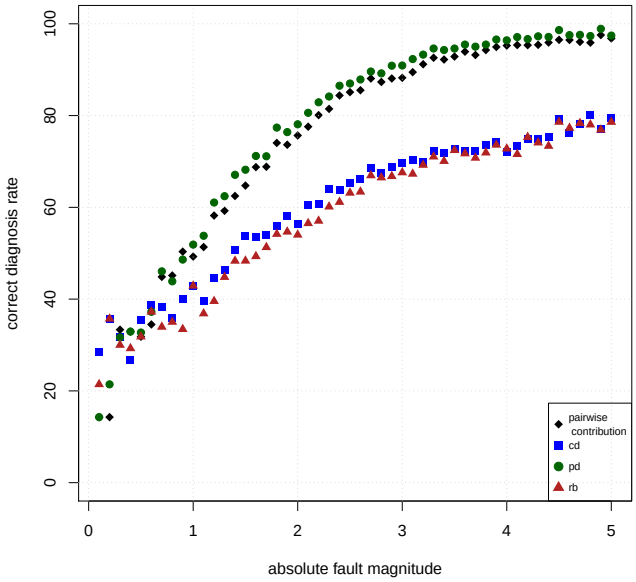


Figure 3.9: Ranking based on SPE pairwise decomposition correct diagnosis rate. Single sensor fault.

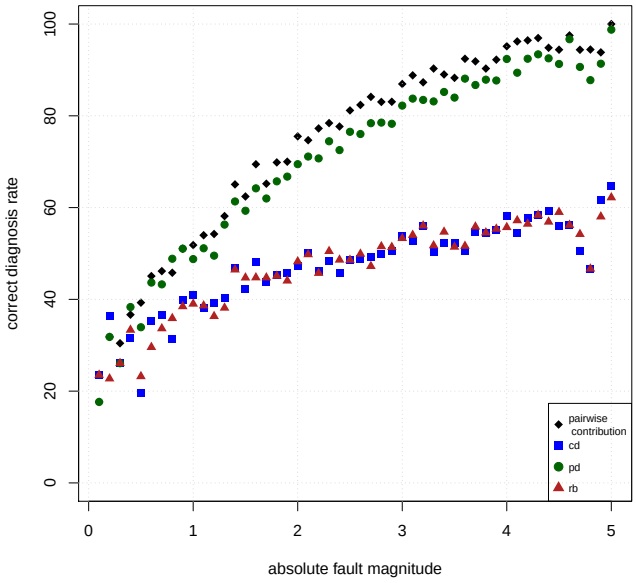


Figure 3.10: Ranking based on T^2 pairwise decomposition correct diagnosis rate. Single sensor fault.

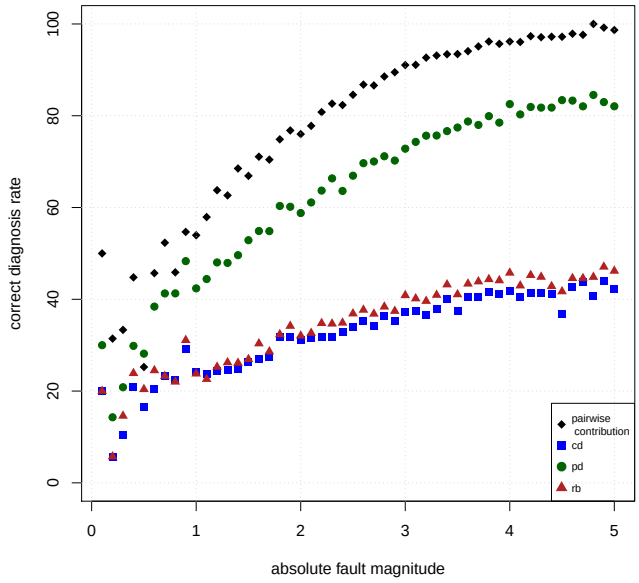


Figure 3.11: Ranking based on Combined index pairwise decomposition correct diagnosis rate. Single sensor fault.

3.8.4 Multiple sensor fault

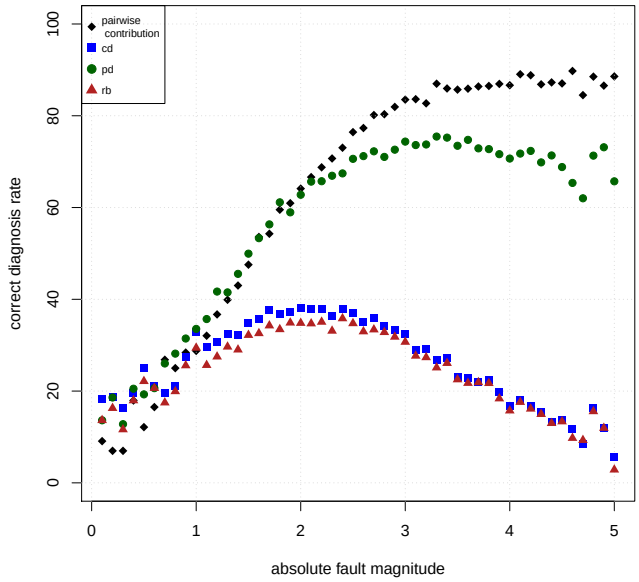


Figure 3.12: Ranking based on SPE pairwise decomposition correct diagnosis rate. Multiple sensor fault.

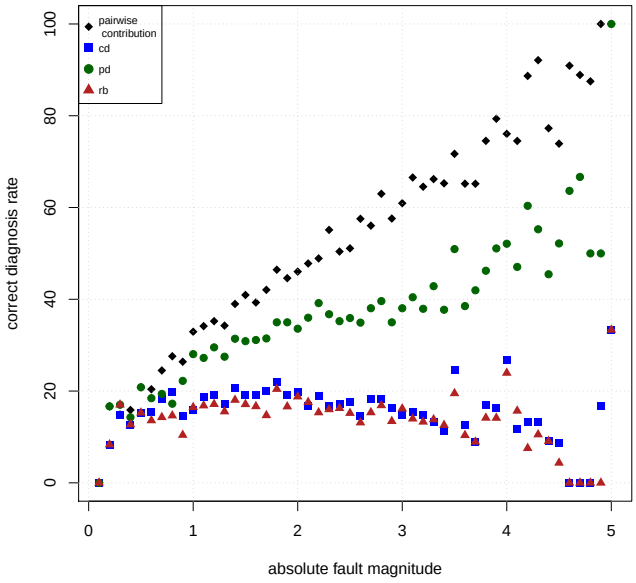


Figure 3.13: Ranking based on T^2 pairwise decomposition correct diagnosis rate. Multiple sensor fault.

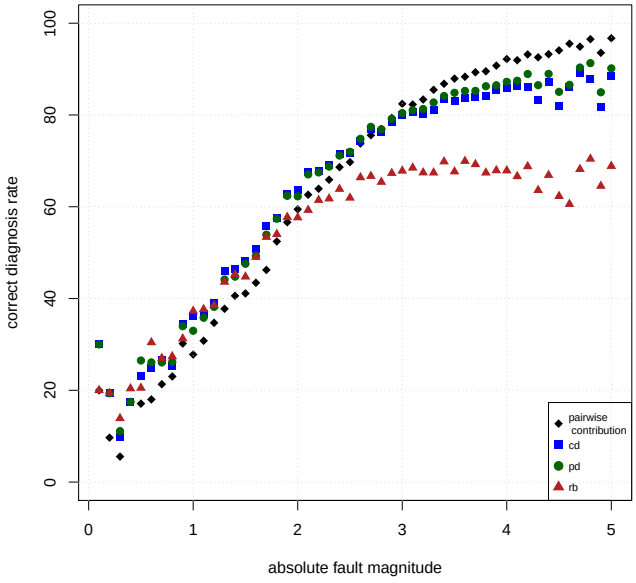


Figure 3.14: Ranking based on Combined index pairwise decomposition correct diagnosis rate. Multiple sensor fault.

3.8.5 Multivariate sensor fault

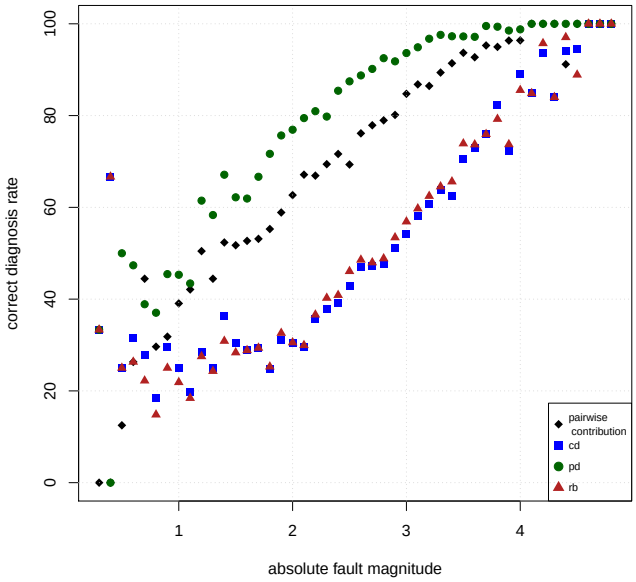


Figure 3.15: Ranking based on SPE pairwise decomposition correct diagnosis rate. Multivariate sensor fault.

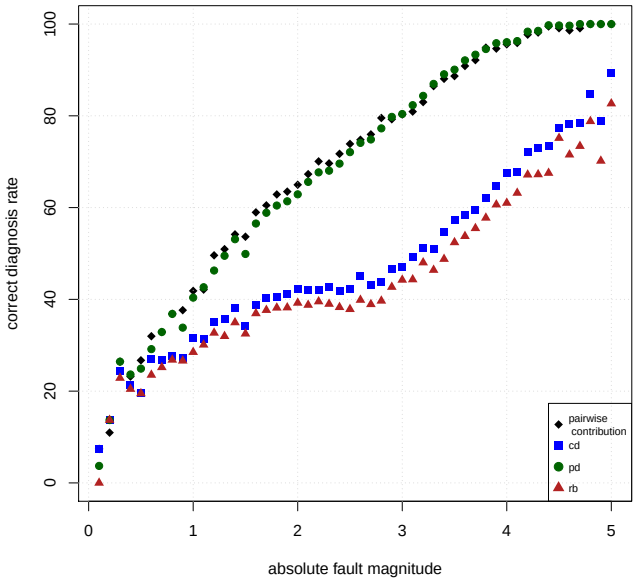


Figure 3.16: Ranking based on T^2 pairwise decomposition correct diagnosis rate. Multivariate sensor fault.

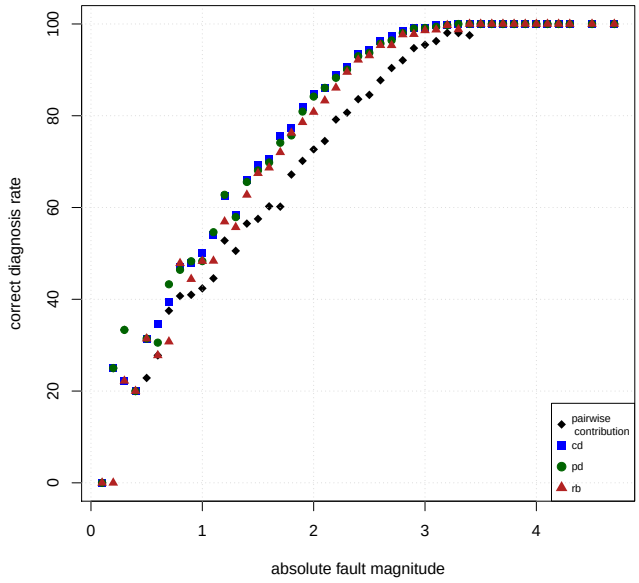


Figure 3.17: Ranking based on Combined index pairwise decomposition correct diagnosis rate. Multivariate sensor fault.

3.9 Summary

It has been illustrated in this chapter that the objective of traditional multivariate fault diagnostic techniques to attach an individual importance ranking to process variables is not achievable. Through illustration it was argued, in two dimensions, that it is possible to attain the same faulty condition by inducing sensor faults in different variables. The motivation provided can easily be extended to show that it holds true in any dimension larger than two. An example from the seven-dimensional data set of Chapter 2 has been given. In the illustration provided it was easy to demonstrate this observation for small fault sizes. This is corroborated by the poor diagnostic performance exhibited by the traditional contribution analysis techniques for the smaller fault magnitudes, in Chapter 2. Another shortcoming of the existing diagnostic techniques is that it is not clear how the individual ranking of process variables can be used by an engineer to interpret why the process variability is statistically different from multivariate common cause variability. It is known from practical experience in the chemical industry that the univariate ranking will naturally prompt the engineer to consider the process variables on an individual basis, which in turn will not alert the engineer to the multivariate nature of the problem, since univariately the process is in-control.

Multivariate statistical process monitoring seeks to identify process abnormalities inside the univariate bounds of the process variables. In PCA this is achieved using fault

identification statistics that are based on the covariance structure of the data. It therefore makes more sense, similar to multivariate fault detection, that the fault diagnosis step should also provide a multivariate answer. Providing a multi-variable diagnostic solution will let the engineer focus on the multivariate characteristics of the process deviation and that the relationships amongst variables should be considered when attempting to correct the process.

Therefore, the suggestion is to specify a ranking of the pairwise importances of the process variables. This proposal is intuitive since the covariance matrix, on which PCA is based, is build using bivariate information. A novel approach is introduced that can be used to decompose the PCA fault identification statistics into a sum of terms, with each term uniquely associated with a pair of variables. Therefore, by using a contribution analysis approach the contribution of each variable pair to the out-of-control fault detection statistic can be quantified. The pairwise contributions calculated in this manner was mathematically motivated. Specifically, it was illustrated how the propagation of a single sensor fault is only observed in a subset of the pairwise contributions. An approach has also been developed that can be used to identify statistically significant pairwise contributions.

Novel methodologies, based on the pairwise contribution method, were also developed that can be used to determine an individual ranking of process variables. Populating the pairwise contributions into a symmetric matrix allowed for the identification of two new approaches that can be used to achieve an individual variable ranking.

In the first approach the specified matrix is viewed as a dissimilarity matrix. This enabled us to use MDS to obtain a two dimensional graphical representation consisting of p points. Each point, indirectly referring to one of the p process variables. Variables associated with points that are remote from the center of the display are identified to have high individual importance for fault diagnosis. In the second approach the row sums of the symmetric matrix are analysed. The process variable corresponding to the row with the highest sum is identified as the highest contributor. Information containing the highest contributor is then removed from the matrix. The second most important variable is then identified as the one associated with the highest row sum in the reduced matrix. This process is then repeated to obtain the final ranking of the variables. The significance level of the ranked variables can be quantified by the empirical comparison to the reference row sum information. The approach based on the row sums calculation is preferred since it is more exact compared to the visual MDS approach.

For comparison, the univariate diagnostic performance of the new method was evaluated using the simulation study of Chapter 2. Reduced fault smearing is observed in some cases, for the new method, when compared to the results of the traditional contribution techniques. However, this technique also exhibit poor performance at the smaller fault

sizes, as expected for the multivariate problem. Fault smearing leading to confusing interpretation is therefore also present, although to a much lesser extent, in the individual ranking obtained using the pairwise approach. It is strongly advised that the individual ranking of process variables should be analysed in addition to the information provided by the pairwise contribution analysis. The value of the new approach is captured by the interpretation value that is conveyed by the pairwise information.

The use of the newly developed methodology was illustrated using a simulated data set. It was specifically demonstrated how a multivariate sensor fault can be correctly diagnosed using pairwise contribution analysis. In Chapter 6 this methodology will be further evaluated using benchmark data and data from a commercial chemical plant.

Chapter 4

Kernel Principal Component Analysis based Fault Diagnosis

4.1 Introduction

The discussion thus far has been centred around the application of principal component analysis (PCA) for multivariate statistical process monitoring. Although being widely used there are instances when PCA is not the appropriate tool to use. This is especially true when the data representing ideal operating conditions (IOC's) are nonlinear. One popular technique used to account for nonlinearity in process data is kernel principal component analysis (KPCA)(Lee et al., 2004). The KPCA approach is a generalisation of the PCA approach with a close resemblance in both fault detection and diagnosis. Although the computational burden of KPCA is known as the sample size increase, it will be illustrated here that it works well for non-linear relationships. In the presence of non-linearity, an approach such as PCA is suboptimal and necessitates the increased complexity and computational burden which must be assessed and handled on a case-by-case basis when applied to industrial processes.
In this chapter it will be illustrated how KPCA is employed for multivariate fault detection and diagnosis.

4.2 Kernel principal component analysis

The general steps of how KPCA is utilised for multivariate statistical process monitoring is identical to how PCA is used i.e.,

1. a reference data set representing IOC's is analysed,

2. to develop fault identification statistics,

3. which are analysed to identify the important process variables responsible for process deviations.

The theory describing KPCA will now be summarised (Schölkopf et al., 1998). Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ be the p-dimensional rows of the matrix $\mathbf{X} : n \times p$ that represents the IOC historical information. The IOC matrix is preprocessed using mean centering and unit variance scaling. Consider the following nonlinear mapping function

$$\phi : \mathbf{x} \in \mathbb{R}^p \longrightarrow \phi(\mathbf{x}) \in \mathbb{R}^h, \tag{4.1}$$

where $p < h < \infty$, which maps a p-dimensional vector into a higher dimensional space $\mathbb{R}^h$ often referred to as the feature space. Note that the vectors in the feature space, $\phi(\mathbf{x})$, can be of infinite dimensionality. An example of a feature mapping function is $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3, (x_1, x_2) \rightarrow (x_1^2, \sqrt{2}x_1x_2, x_2^2)$. The principal component analysis of these high dimensional feature vectors was first introduced by Schölkopf et al. (1998) as kernel PCA. Kernel PCA is therefore performed by calculating the eigenvectors of the sample covariance matrix in feature space

$$\mathbf{S}^{\phi}_{h\times h} = \frac{1}{n-1}\sum_{i=1}^{n}(\phi(\mathbf{x}_i) - \tilde{\phi}(\mathbf{x})(\phi(\mathbf{x}_i) - \tilde{\phi}(\mathbf{x}))^T, \tag{4.2}$$

with

$$\tilde{\phi}(\mathbf{x})_{h\times 1} = \frac{1}{n}\sum_{i=1}^{n}\phi(\mathbf{x}_i), \tag{4.3}$$

the sample mean in feature space. Note that $\mathbf{S}^{\phi} : h \times h$ is maybe of infinite dimensionality, which makes it difficult to perform the feature space PCA using conventional methods. Schölkopf et al. (1998) illustrate that it is possible to calculate the eigenvectors of $\mathbf{S}^{\phi}$ by making use of the so-called kernel trick. Let the mean centered feature vector be denoted by

$$\bar{\phi}(\mathbf{x}_i) \quad = \quad \phi(\mathbf{x}_i) - \tilde{\phi}(\mathbf{x}). \tag{4.4}$$

The feature space principal components are then obtained by calculating all nonzero vectors $\mathbf{v} : h \times 1$ such that

$$\begin{aligned}\lambda\mathbf{v} &= \mathbf{S}^{\phi}\mathbf{v} && (4.5)\\ &= \frac{1}{n-1}\sum_{i=1}^{n}\bar{\phi}(\mathbf{x}_i)\bar{\phi}(\mathbf{x}_i)^T\mathbf{v}\\ &= \frac{1}{n-1}(\mathbf{X}^{\phi})^T\mathbf{X}^{\phi}\mathbf{v}, && (4.6)\end{aligned}$$

where $\mathbf{X}^{\phi} : n \times h = \begin{bmatrix}\bar{\phi}(\mathbf{x}_1)^T & \bar{\phi}(\mathbf{x}_2)^T & \dots & \bar{\phi}(\mathbf{x}_n)^T\end{bmatrix}^T$. The presence of the infinite dimensional vectors $\bar{\phi}(\mathbf{x}_i)$ and $\mathbf{v}$ in this eigenvalue problem (4.5) makes it difficult to solve. It is useful to observe that if both sides of (4.5) are multiplied by $\mathbf{X}^{\phi}$ then the eigenvalue problem is fully specified by finite dimensional matrices and vectors

$$\begin{aligned}\lambda\mathbf{v} &= \mathbf{S}^{\phi}\mathbf{v}\\ \lambda\mathbf{X}^{\phi}\mathbf{v} &= \mathbf{X}^{\phi}\mathbf{S}^{\phi}\mathbf{v}\\ \lambda\mathbf{X}^{\phi}\mathbf{v} &= \frac{1}{n-1}\Big[\mathbf{X}^{\phi}(\mathbf{X}^{\phi})^T\Big]\mathbf{X}^{\phi}\mathbf{v}\\ \lambda\mathbf{w} &= \mathbf{K}\mathbf{w} && (4.7)\end{aligned}$$

where

$$\mathbf{w}_{n\times 1} = \mathbf{X}^{\phi}\mathbf{v} \qquad (4.8)$$

and

$$\begin{aligned}\mathbf{K}_{n\times n} &= \frac{1}{n-1}\mathbf{X}^{\phi}(\mathbf{X}^{\phi})^T && (4.9)\\ &= \frac{1}{n-1}\begin{bmatrix}\bar{\phi}(\mathbf{x}_1)^T\\ \bar{\phi}(\mathbf{x}_2)^T\\ \vdots\\ \bar{\phi}(\mathbf{x}_n)^T\end{bmatrix}\Big[\bar{\phi}(\mathbf{x}_1) \quad \bar{\phi}(\mathbf{x}_2)\dots\bar{\phi}(\mathbf{x}_n)\Big] && (4.10)\\ &= \frac{1}{n-1}\begin{bmatrix}\bar{\phi}(\mathbf{x}_1)^T\bar{\phi}(\mathbf{x}_1) & \bar{\phi}(\mathbf{x}_1)^T\bar{\phi}(\mathbf{x}_2) & \cdots & \bar{\phi}(\mathbf{x}_1)^T\bar{\phi}(\mathbf{x}_n)\\ \bar{\phi}(\mathbf{x}_2)^T\bar{\phi}(\mathbf{x}_1) & \bar{\phi}(\mathbf{x}_2)^T\bar{\phi}(\mathbf{x}_2) & \cdots & \bar{\phi}(\mathbf{x}_2)^T\bar{\phi}(\mathbf{x}_n)\\ \vdots & \vdots & \ddots & \vdots\\ \bar{\phi}(\mathbf{x}_n)^T\bar{\phi}(\mathbf{x}_1) & \bar{\phi}(\mathbf{x}_n)^T\bar{\phi}(\mathbf{x}_2) & \cdots & \bar{\phi}(\mathbf{x}_n)^T\bar{\phi}(\mathbf{x}_n)\end{bmatrix} && (4.11)\end{aligned}$$

The solution of the eigenvalue problem in feature space, represented by (4.5), can therefore

be obtained by considering solutions to the equivalent eigen problem expressed by (4.7). Notice that after calculating the n dimensional principal component weights $\mathbf{w}$ it is possible to obtain the corresponding h dimensional vector $\mathbf{v}$ of the original problem. This is achieved by multiplying equation (4.8) with $(\mathbf{X}^{\phi})^T$ as follows

$$(\mathbf{X}^{\phi})^T\mathbf{w} = (\mathbf{X}^{\phi})^T\mathbf{X}^{\phi}\mathbf{v}, \tag{4.12}$$

from equation (4.6) this then simplifies to

$$(\mathbf{X}^{\phi})^T\mathbf{w} = (n-1)\lambda\mathbf{v}, \tag{4.13}$$

which implies that

$$\mathbf{v} = \frac{1}{\lambda(n-1)}(\mathbf{X}^{\phi})^T\mathbf{w}. \tag{4.14}$$

In PCA the latent variable weight vectors are restricted to have unit length. That is,

$$\mathbf{v}^T\mathbf{v} = 1 \tag{4.15}$$

$$\Big[\frac{1}{\lambda(n-1)}\Big]^2\mathbf{w}^T\mathbf{X}^{\phi}(\mathbf{X}^{\phi})^T\mathbf{w} = 1 \tag{4.16}$$

$$(n-1)\Big[\frac{1}{\lambda(n-1)}\Big]^2\mathbf{w}^T\mathbf{K}\mathbf{w} = 1 \tag{4.17}$$

$$\lambda(n-1)\Big[\frac{1}{\lambda(n-1)}\Big]^2\mathbf{w}^T\mathbf{w} = 1, \tag{4.18}$$

which implies that the eigenvectors $\mathbf{w}$ should be scaled such that

$$\mathbf{w}^T\mathbf{w} = \lambda(n-1). \tag{4.19}$$

Let $\tilde{\mathbf{w}}$ represent the scaled loading vectors i.e., set

$$\tilde{\mathbf{w}} = \sqrt{\lambda(n-1)}\mathbf{w}, \tag{4.20}$$

such that $\tilde{\mathbf{w}}^T\tilde{\mathbf{w}} = \lambda(n-1)$. As a result of this scaling the relationship between the two

sets of loadings changes from (4.14) to

$$\mathbf{v} = \frac{1}{\lambda(n-1)}(\mathbf{X}^{\phi})^T\tilde{\mathbf{w}}. \tag{4.21}$$

The main advantage to writing the eigenvalue problem as (4.7) is that the matrix $\mathbf{K}$ is constructed using only the dot products between feature vectors i.e., the $(i,j)^{th}$ element of $\mathbf{K}$ is

$$K_{i,j} = \frac{1}{n-1}\bar{\phi}(\mathbf{x}_i)^T\bar{\phi}(\mathbf{x}_j). \tag{4.22}$$

It is well known that the dot product between two vectors can be used as a measure to quantify their similarity (Johnson and Wichern, 2002). The matrix $\mathbf{K}$ can therefore be considered a similarity matrix. The different matrix entries $K_{i,j}$ can therefore be calculated if the pairwise similarities between the n feature space vectors are known. Hence negating the need to know the feature mapping function. Schölkopf et al. (1998) illustrate how the similarity between feature vector pairs can be estimated by using kernel functions. Only the information on the original IOC data vectors $\mathbf{x}$ is required to evaluate the kernel function. The elements of $\mathbf{K}$ are estimated as follows

$$(n-1)K_{i,j} = \bar{\phi}(\mathbf{x}_i)^T\bar{\phi}(\mathbf{x}_j) \tag{4.23}$$

$$\approx k(\mathbf{x}_i,\mathbf{x}_j). \tag{4.24}$$

The approximation of the dot products between feature space vectors using kernel functions is known as the kernel trick. Let $\tilde{\mathbf{K}}$ represent the approximating matrix containing the kernel function evaluations. A large library of kernel functions exists that can be used to estimate the similarities in this fashion (Schölkopf and Smola, 2002). The Gaussian kernel function is a kernel that is popular amongst practitioners (Hastie et al., 2001). It is defined as follows

$$k(\mathbf{x}_i,\mathbf{x}_j) = \exp\Big[-\frac{(\mathbf{x}_i-\mathbf{x}_j)^T(\mathbf{x}_i-\mathbf{x}_j)}{2\sigma^2}\Big], \tag{4.25}$$

for $i,j=1,\ldots,n$ and with $\sigma>0$ a hyperparameter that needs to be pre-specified.
In order for a kernel function to be used in the kernel trick it must adhere to Mercer's theorem (Schölkopf and Smola, 2002). It is required that for a selected kernel function that

the resulting kernel matrix $\tilde{\mathbf{K}}$ has the positive semidefinite property which characterizes the original similarity matrix $\mathbf{K}$. Another property that the approximating similarity matrix $\tilde{\mathbf{K}}$ should have is that the implied feature vectors comprising the inner products, which are being estimated, should be centered around zero. Consider the inner product of two zero mean feature vectors

$$\begin{aligned}
\bar{\phi}(\mathbf{x}_i)^T\bar{\phi}(\mathbf{x}_j) &= \Big[\phi(\mathbf{x}_i) - \tilde{\phi}(\mathbf{x})\Big]^T\Big[\phi(\mathbf{x}_j) - \tilde{\phi}(\mathbf{x})\Big] && (4.26)\\
&= \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j) - \tilde{\phi}(\mathbf{x})^T\phi(\mathbf{x}_j) - \phi(\mathbf{x}_i)^T\tilde{\phi}(\mathbf{x}) + \tilde{\phi}(\mathbf{x})^T\tilde{\phi}(\mathbf{x}) && (4.27)\\
&= \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j) - \frac{1}{n}\sum_{m=1}^{n}\phi(\mathbf{x}_m)^T\phi(\mathbf{x}_j) - \frac{1}{n}\sum_{m=1}^{n}\phi(\mathbf{x}_i)^T\phi(\mathbf{x}_m) + \\
&\quad \tilde{\phi}(\mathbf{x})^T\tilde{\phi}(\mathbf{x}). && (4.28)
\end{aligned}$$

Therefore, if it can be assumed that the feature vectors inferred by the kernel function entries in $\tilde{\mathbf{K}}$ are uncentered i.e. $K_{i,j} = \frac{1}{n-1}\phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$ then

$$\begin{aligned}
\frac{1}{n-1}\bar{\phi}(\mathbf{x}_i)^T\bar{\phi}(\mathbf{x}_j) &= \frac{1}{n-1}\Big[k(\mathbf{x}_i,\mathbf{x}_j) - \frac{1}{n}\sum_{m=1}^{n}k(\mathbf{x}_m,\mathbf{x}_j) - \frac{1}{n}\sum_{m=1}^{n}k(\mathbf{x}_i,\mathbf{x}_m) + \\
&\quad \frac{1}{n^2}\sum_{l=1}^{n}\sum_{u=1}^{n}k(\mathbf{x}_l,\mathbf{x}_u)\Big]. && (4.29)
\end{aligned}$$

This implies that in order to ensure that the feature vectors are centered, it is required that each element of $\tilde{\mathbf{K}}$ should be translated by $\frac{1}{n^2}\sum_{l=1}^{n}\sum_{u=1}^{n}K_{l,u} - \frac{1}{n}\sum_{m=1}^{n}K_{m,j} - \frac{1}{n}\sum_{m=1}^{n}K_{i,m}$. In matrix form the translation is achieved by the following expression

$$(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T)\tilde{\mathbf{K}}(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T), \tag{4.30}$$

where $\mathbf{I}_n$ represents the $n \times n$ identity matrix and $\mathbf{1}_n : n \times 1$ is a vector equalling one in every position.

Considering that kernel PCA is just a PCA analysis in feature space, only the first few loading vectors $\mathbf{v}_i$ for $i = 1, \ldots, h^*$ with $h^* < h$ are required. Note that since only the eigenvectors of $\tilde{\mathbf{K}}$ are calculated, the ultimate dimension h^* will be less than n. Let the columns of $\mathbf{V} : h \times h^*$ contain the feature loading vectors i.e.

$$\mathbf{V} = \left[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \ldots \quad \mathbf{v}_{h^*}\right] \tag{4.31}$$

$$= \left[\tfrac{1}{\lambda_1(n-1)}(\mathbf{X}^\phi)^T\tilde{\mathbf{w}}_1 \quad \tfrac{1}{\lambda_2(n-1)}(\mathbf{X}^\phi)^T\tilde{\mathbf{w}}_2 \quad \ldots \quad \tfrac{1}{\lambda_{h^*}(n-1)}(\mathbf{X}^\phi)^T\tilde{\mathbf{w}}_{h^*}\right], \tag{4.32}$$

from (4.21). This equation can be written more compactly in matrix form as

$$\mathbf{V} = \frac{1}{n-1}(\mathbf{X}^\phi)^T\tilde{\mathbf{W}}\mathbf{\Lambda}^{-1}, \tag{4.33}$$

with $\mathbf{\Lambda} : h^* \times h^* = diag(\lambda_1, \lambda_2, \ldots, \lambda_{h^*})$ and

$$\tilde{\mathbf{W}}_{n\times h^*} = \left[\tilde{\mathbf{w}}_1 \quad \tilde{\mathbf{w}}_2 \quad \ldots \quad \tilde{\mathbf{w}}_{h^*}\right], \tag{4.34}$$

the matrix containing the scaled eigenvectors of $\tilde{\mathbf{K}}$. The score representation, $\mathbf{t}^\phi : h^* \times 1$, of an appropriately centered feature vector $\bar{\phi}(\mathbf{x}) : h \times 1$ can therefore be calculated as follows

$$\mathbf{t}^\phi = \mathbf{V}^T\bar{\phi}(\mathbf{x}) \tag{4.35}$$

$$= \frac{1}{n-1}\mathbf{\Lambda}^{-1}\tilde{\mathbf{W}}^T\mathbf{X}^\phi\bar{\phi}(\mathbf{x}) \tag{4.36}$$

$$= \frac{1}{n-1}\mathbf{\Lambda}^{-1}\tilde{\mathbf{W}}^T\begin{bmatrix}\bar{\phi}(\mathbf{x}_1)^T\\ \bar{\phi}(\mathbf{x}_2)^T\\ \vdots\\ \bar{\phi}(\mathbf{x}_n)^T\end{bmatrix}\bar{\phi}(\mathbf{x}) \tag{4.37}$$

$$= \mathbf{\Lambda}^{-1}\tilde{\mathbf{W}}^T\begin{bmatrix}\frac{1}{n-1}\bar{\phi}(\mathbf{x}_1)^T\bar{\phi}(\mathbf{x})\\ \frac{1}{n-1}\bar{\phi}(\mathbf{x}_2)^T\bar{\phi}(\mathbf{x})\\ \vdots\\ \frac{1}{n-1}\bar{\phi}(\mathbf{x}_n)^T\bar{\phi}(\mathbf{x})\end{bmatrix}. \tag{4.38}$$

Notice that in the expression above it is also conveniently not required to know $\bar{\phi}(\mathbf{x})$ for calculating $\mathbf{t}^\phi$. Information on the inner product similarity of $\bar{\phi}(\mathbf{x})$ with the feature IOC vectors are needed only i.e., $\bar{\phi}(\mathbf{x}_i)^T\bar{\phi}(\mathbf{x})$ for $i = 1, \ldots, n$. The kernel function used to define $\tilde{\mathbf{K}}$ can therefore be used to approximate these inner products

$$k(\mathbf{x}_i, \mathbf{x}) \approx \bar{\phi}(\mathbf{x}_i)^T \bar{\phi}(\mathbf{x}), \tag{4.39}$$

$i = 1, \ldots, n$. The score representation (4.36) therefore simplifies to

$$\mathbf{t}^{\phi} = \mathbf{\Lambda}^{-1} \tilde{\mathbf{W}}^T \begin{bmatrix} \frac{1}{n-1} k(\mathbf{x}_1, \mathbf{x}) \\ \frac{1}{n-1} k(\mathbf{x}_2, \mathbf{x}) \\ \vdots \\ \frac{1}{n-1} k(\mathbf{x}_n, \mathbf{x}) \end{bmatrix}. \tag{4.40}$$

Note that it is again important to ensure that after applying the kernel function that the implied feature vectors are centered. Each term in the vector containing the kernel approximations in (4.40) should therefore be centered as follows

$$\frac{1}{n-1} k(\mathbf{x}_i, \mathbf{x}) - \frac{1}{n(n-1)} \sum_{m=1}^{n} k(\mathbf{x}_m, \mathbf{x}) - \frac{1}{n} \sum_{m=1}^{n} K_{i,m} + \frac{1}{n^2} \sum_{l=1}^{n} \sum_{u=1}^{n} K_{l,u}. \tag{4.41}$$

It can be illustrated that the final solution of kernel PCA is highly dependent on the values chosen for the hyperparameters describing the shape of the kernel functions (Alam and Fukumizu, 2014). This will therefore result in different interpretations for different hyperparameter readings. A statistically sound and consistent methodology is therefore required to select the ideal hyperparameter setting. In the paper by Alam and Fukumizu (2014) it is proposed that the hyperparameters be selected in a manner which ensures that the so-called pre-images of the IOC samples are close to their kernel PCA approximations. Let the kernel PCA estimate of a feature vector $\bar{\phi}(\mathbf{x})$ be represented by

$$\hat{\bar{\phi}}(\mathbf{x}) = \mathbf{V}\mathbf{V}^T \bar{\phi}(\mathbf{x}). \tag{4.42}$$

A vector $\mathbf{z} : p \times 1$ is then identified as the pre-image of $\mathbf{x}$ if

$$\rho(\mathbf{z}) = \|\hat{\bar{\phi}}(\mathbf{x}) - \bar{\phi}(\mathbf{z})\|^2, \tag{4.43}$$

is the smallest. The pre-image objective function $\rho(\mathbf{z})$ can equivalently be expressed as

$$
\begin{aligned}
\rho(\mathbf{z}) &= \|\hat{\bar{\phi}}(\mathbf{x}) - \bar{\phi}(\mathbf{z})\|^2 && (4.44)\\
&= \left[\hat{\bar{\phi}}(\mathbf{x}) - \bar{\phi}(\mathbf{z})\right]^T \left[\hat{\bar{\phi}}(\mathbf{x}) - \bar{\phi}(\mathbf{z})\right] && (4.45)\\
&= \bar{\phi}(\mathbf{z})^T\bar{\phi}(\mathbf{z}) - 2\bar{\phi}(\mathbf{z})^T\hat{\bar{\phi}}(\mathbf{x}) + \hat{\bar{\phi}}(\mathbf{x})^T\hat{\bar{\phi}}(\mathbf{x}) && (4.46)\\
&= \bar{\phi}(\mathbf{z})^T\bar{\phi}(\mathbf{z}) - 2\bar{\phi}(\mathbf{z})^T\mathbf{V}\mathbf{V}^T\bar{\phi}(\mathbf{x}) + \hat{\bar{\phi}}(\mathbf{x})^T\hat{\bar{\phi}}(\mathbf{x}) && (4.47)\\
&= \bar{\phi}(\mathbf{z})^T\bar{\phi}(\mathbf{z}) - 2(\mathbf{t}^{\phi(\mathbf{z})})^T\mathbf{t}^{\phi(\mathbf{x})} + \hat{\bar{\phi}}(\mathbf{x})^T\hat{\bar{\phi}}(\mathbf{x}), && (4.48)
\end{aligned}
$$

where $\mathbf{t}^{\phi(\mathbf{z})} = \mathbf{V}^T\bar{\phi}(\mathbf{z})$. Observe that the inner product term $\hat{\bar{\phi}}(\mathbf{x})^T\hat{\bar{\phi}}(\mathbf{x})$ does not include $\mathbf{z}$. Therefore the minimizing pre-image of $\rho(\mathbf{z})$ can be obtained by finding the vector $\mathbf{z}$ which minimizes

$$
k(\mathbf{z},\mathbf{z}) - 2(\mathbf{t}^{\phi(\mathbf{z})})^T\mathbf{t}^{\phi(\mathbf{x})}, \tag{4.49}
$$

with $k(\mathbf{z},\mathbf{z}) \approx \bar{\phi}(\mathbf{z})^T\bar{\phi}(\mathbf{z})$, which needs to be appropriately centered as previously stated. The equivalent form of the pre-image objective function is therefore fully expressible using kernel approximations. Notice that when the Gaussian kernel is used the kernel similarity of $\mathbf{z}$ with itself will be constant. Therefore when using the Gaussian kernel, the search for a pre-image $\mathbf{z}$ can be reduced to the maximization of

$$
(\mathbf{t}^{\phi(\mathbf{z})})^T\mathbf{t}^{\phi(\mathbf{x})}. \tag{4.50}
$$

A number of optimization searching strategies have been proposed in literature to solve the pre-image problem. For example in Alam and Fukumizu (2014) it is suggested that the fixed-point and gradient descent algorithms be used. It is also suggested by Alam and Fukumizu (2014) that leave one out cross-validation be used to identify the hyperparameter setting that optimises the pre-image problem. The method of identifying this optimum, for kernel PCA based on the Gaussian kernel, is explained by Algorithm 1.

Algorithm 1 Kernel PCA - Gaussian kernel hyperparemter selection.

Require: $\mathbf{X}$: $n \times p$ the IOC data and $(\sigma_1, h_1^*), \ldots, (\sigma_S, h_S^*)$ a list of S possible paired hyperparameters and PCA dimension values.

1: Set $s = 1$ where $s \leqslant S$
2: Set $i = 1$
3: **while** $i < n$ **do**
4: Fit the (σ_s, h_s^*) kernel PCA model with i^{th} sample $\mathbf{x}_i : p \times 1$ removed.
5: Calculate the approximating pre-image $\mathbf{z}_i$ of the removed sample $\mathbf{x}_i$.
6: Record the reconstruction error $err_{si} = \|\mathbf{z}_i - \mathbf{x}_i\|^2$.
7: Set $i = i + 1$.
8: If $i > n$ go to 9 else go to 4.
9: Calculate the average error $err_s = \frac{1}{n}\sum_{i=1}^{n} err_{si}$.
10: Set $s = s + 1$
11: If $s > S$ go to step 12 else go to step 2.
12: Identify the condition $s = s^*$ for which err_s is the smallest.
13: **Output:** (σ_{s*}, h_{s*}^*)

4.3 Fault detection statistics

The methodologies developed for process deviation detection using PCA naturally extend to kernel PCA. The PCA detection statistics discussed in Chapter 2 are directly applied to the feature space vectors. Movements from ideal operating conditions will therefore be identified by evaluating the score representations of new feature vectors to that of the reference features.

From Westerhuis et al. (2000) the T^2-statistic of a new feature space observation $\bar{\phi}(\mathbf{x})$ is defined as

$$T^2 = \bar{\phi}(\mathbf{x})^T \mathbf{V} \mathbf{\Lambda}^{-1} \mathbf{V}^T \bar{\phi}(\mathbf{x}) \tag{4.51}$$

$$= (\mathbf{t}^\phi)^T \mathbf{\Lambda}^{-1} \mathbf{t}^\phi. \tag{4.52}$$

Using the PCA definition for the SPE statistic as provided by Jackson and Mudholkar (1979), yields the following definition in feature space

$$SPE = \bar{\phi}(\mathbf{x})^T \tilde{\mathbf{V}} \tilde{\mathbf{V}}^T \bar{\phi}(\mathbf{x}), \tag{4.53}$$

where $\tilde{\mathbf{V}} : h \times (h - h^*) = \begin{bmatrix} \mathbf{v}_{h^*+1} & \mathbf{v}_{h^*+2} & \ldots & \mathbf{v}_h \end{bmatrix}$ is the matrix containing the residual loading vectors. Notice that since $\tilde{\mathbf{V}}$ is potentially of infinite dimensionality it is difficult to evaluate the above expression for the SPE statistic. This difficulty can be circumvented

by using the orthogonality property of PCA loading vectors. Let $\bar{\mathbf{V}} : h \times h$ represent the full set of feature loading vectors. Orthogonality implies that

$$\begin{aligned}
\bar{\mathbf{V}}\bar{\mathbf{V}}^T &= \mathbf{I}_{h\times h} && (4.54)\\
\begin{bmatrix}\mathbf{V}_{h\times h^*} & \tilde{\mathbf{V}}_{h\times(h-h^*)}\end{bmatrix}\begin{bmatrix}\mathbf{V}^T_{h\times h^*}\\ \tilde{\mathbf{V}}^T_{h\times(h-h^*)}\end{bmatrix} &= \mathbf{I}_{h\times h} && (4.55)\\
\mathbf{V}_{h\times h^*}\mathbf{V}^T_{h\times h^*} + \tilde{\mathbf{V}}_{h\times(h-h^*)}\tilde{\mathbf{V}}^T_{h\times(h-h^*)} &= \mathbf{I}_{h\times h} && (4.56)\\
\implies \tilde{\mathbf{V}}_{h\times(h-h^*)}\tilde{\mathbf{V}}^T_{h\times(h-h^*)} &= \mathbf{I}_{h\times h} - \mathbf{V}_{h\times h^*}\mathbf{V}^T_{h\times h^*}. && (4.57)
\end{aligned}$$

The feature space definition of SPE therefore becomes

$$\begin{aligned}
SPE &= \bar{\phi}(\mathbf{x})^T(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\bar{\phi}(\mathbf{x}) && (4.58)\\
&= \bar{\phi}(\mathbf{x})^T\bar{\phi}(\mathbf{x}) - \bar{\phi}(\mathbf{x})^T\mathbf{V}\mathbf{V}^T\bar{\phi}(\mathbf{x}) && (4.59)\\
&= \bar{\phi}(\mathbf{x})^T\bar{\phi}(\mathbf{x}) - (\mathbf{t}^\phi)^T\mathbf{t}^\phi && (4.60)\\
&= k(\mathbf{x},\mathbf{x}) - (\mathbf{t}^\phi)^T\mathbf{t}^\phi, && (4.61)
\end{aligned}$$

where the kernel approximation $k(\mathbf{x},\mathbf{x}) \approx \bar{\phi}(\mathbf{x})^T\bar{\phi}(\mathbf{x})$ needs to be centered by adding the following factor

$$\frac{n-1}{n^2}\sum_{l=1}^{n}\sum_{u=1}^{n}K_{l,u} - \frac{2}{n}\sum_{i=1}^{n}\phi(\mathbf{x}_i)^T\phi(\mathbf{x}). \tag{4.62}$$

This centering factor is obtained using the motivation presented for the derivation of (4.30). The SPE definition is therefore now fully expressed using known matrices.
Similar to traditional PCA, the combined statistic defined by Yue and Qin (2001) can also be calculated

$$\varphi = \frac{T^2}{\tau^2} + \frac{SPE}{\delta^2}, \tag{4.63}$$

with τ^2 and δ^2 representing the control limits of the T^2 and SPE statistics. By using the results of Box (1954), Alcala and Qin (2010) derived control limits for the aforementioned feature space control statistics. These control limits are reported in Table 4.1. It is important to note that these control limits are based on the assumption that the feature vectors are normally distributed. Normality cannot always be guaranteed, we therefore

Table 4.1: Kernel PCA control limit definitions.

Statistic	T^2	SPE	φ
Limit	τ^2	δ^2	ς^2
Value	$\chi^2_\alpha(h^*)$	$\frac{\sum_{i=h^*+1}^{h}\lambda_i^2}{\sum_{i=h^*+1}^{h}\lambda_i}\chi^2_\alpha\left(\frac{(\sum_{i=h^*+1}^{h}\lambda_i)^2}{\sum_{i=h^*+1}^{h}\lambda_i^2}\right)$	$\frac{h^*/\tau^4+\sum_{i=h^*+1}^{h}\lambda_i^2/\delta^4}{h^*/\tau^2+\sum_{i=h^*+1}^{h}\lambda_i/\delta^2}\chi^2_\alpha\left(\frac{h^*/\tau^2+\sum_{i=h^*+1}^{h}\lambda_i/\delta^2}{h^*/\tau^4\sum_{i=h^*+1}^{h}\lambda_i^2/\delta^4}\right)$

suggest that these control limits be estimated using non-parametric density estimation (Hastie et al., 2001). The practical relevance of theoretical control limits compared to non-parametric determined limits will be illustrated in Section 4.5.

4.4 Fault diagnosis methods

If the kernel PCA fault detection statistics identify a new feature vector as being significantly different from the IOC data set, the next step is then to identify the fault inducing process variables. It is clear from the formulation of the detection indices that it is difficult to use the PCA decomposition contribution methods discussed in Alcala and Qin (2011) for fault diagnosis.

Cho et al. (2005) propose a variable contribution calculation methodology based on a variable selection methodology used in machine learning to address this difficulty. They consider the identification of a faulty variable as the process variable which has the greatest influence on the out-of-control fault detection statistics. Their method is based on the definition of the derivative of a kernel function proposed by Rakotomamonjy (2003). The derivative definition makes use of an imaginary scaling vector $\boldsymbol{\omega}^T = \Big[\omega_1, \ldots \omega_p\Big]$ such that

$$k(\mathbf{x}_i, \mathbf{x}_j) \quad = \quad k(\boldsymbol{\omega} \cdot \mathbf{x}_i, \boldsymbol{\omega} \cdot \mathbf{x}_j), \tag{4.64}$$

where $\boldsymbol{\omega} \cdot \mathbf{x} = \Big[\omega_1 x_1, \ldots \omega_p x_p\Big]^T$ and $\omega_i = 1$ for $i = 1, \ldots, p$. The kernel function gradient with respect to the i^{th} scaling variable, ω_i, is then obtained as

$$\frac{\partial k(\mathbf{x}_i, \mathbf{x}_j)}{\partial \omega_i} \quad = \quad \left.\frac{\partial k(\boldsymbol{\omega} \cdot \mathbf{x}_i, \boldsymbol{\omega} \cdot \mathbf{x}_j)}{\partial \omega_i}\right|_{\omega_i = 1}, \tag{4.65}$$

for $i = 1, \ldots, p$. The gradient is evaluated at $\omega_i = 1$ with the other ω_j scaling values for $i \neq j$ set equal to one. Using this definition Cho et al. (2005) then define the contribution of the i^{th} variable to the T^2 and SPE statistics respectively as

$$C_i^{T^2} = \left| \frac{\partial T^2}{\partial \omega_i} \right| \tag{4.66}$$

and

$$C_i^{SPE} = \left| \frac{\partial SPE}{\partial \omega_i} \right|. \tag{4.67}$$

Similar to variable contributions calculated in PCA, large values indicate the variables most likely to contain information on why the process was identified to have drifted from IOC's.

The reconstruction-based (RB) contribution methodology developed by Alcala and Qin (2009) for PCA can be applied to kernel PCA to obtain variable contributions to process deviations. As the name suggests, the RB method searches for the process variables that need to be reconstructed the most in order to minimize a specific control statistic. The following reconstructed vector

$$\mathbf{z}_i = \mathbf{x} - f_i \boldsymbol{\xi}_i, \tag{4.68}$$

with $f_i \in \mathbb{R}$ and $\boldsymbol{\xi}_i$ a vector of zeros except with a 1 in the i^{th} position, is considered for each of the process variables. With the objective of finding, for each variable, the value of f_i that minimizes the detection statistic calculated for $\mathbf{z}_i$. The RB contribution of the i^{th} process variable is therefore defined as

$$RB_i = \arg\min_{f_i} Index(\mathbf{z}_i), \tag{4.69}$$

with $Index$ referencing a specific fault detection statistic.

The kernel PCA diagnostic techniques are therefore in principle the same as the contribution techniques employed in linear PCA based process monitoring. Both sets of approaches seek to attach importance to individual variables. Therefore, this manner of thinking about multivariate fault diagnosis is not ideal as illustrated in Chapter 3. An illustration will be provided in Chapter 5 to demonstrate the inappropriateness of calculating individual variable contributions in the non-linear setting.

4.5 Kernel PCA Simulation Study

4.5.1 Objective

A simulation study will now be presented to illustrate how the different methodologies summarized in Section 4.4 are typically applied for fault diagnosis in the non-linear setting. A definition of a simulated data set will be provided to represent samples observed during IOC's. This will be a new definition that is a variant of well established simulation strategies, see for example Cho et al. (2005). PCA and kernel PCA models will then be fitted using the IOC data. A case study will then be used for illustration purposes to motivate why kernel PCA would be preferred in this setting. Specifically, the fault detection ability of PCA and kernel PCA will be discussed for one process fault example. It will then be illustrated how this particular example would be diagnosed using the methodologies introduced for kernel PCA by Cho et al. (2005) and Alcala and Qin (2009). A larger array of process faults will then be simulated to evaluate the interpretation ability of these fault diagnosis methods. That is, similar to the simulation that was presented in Chapter 2, this evaluation will be done for different fault magnitudes and fault types. It is envisioned that the results presented will give a good indication of the performance that can be expected in real life applications, even though the results presented are specific to this simulation study.

4.5.2 Data definition

The data set generated in this section to represent the IOC samples will form the basis of all future discussions. In order to generate a data set that calls for the application of kernel PCA instead of linear PCA, it is required that the p-dimensional data vectors are optimally summarized by a non-linear surface instead of a linear one. Known examples, see for example Cho et al. (2005), will be used as a guideline to construct such a simulated data matrix. A common motif in these examples is to generate the simulation process variables as non-linear functions of some shared underlying variable. Examples will also be given to illustrate how the simulation strategy can be related to known process use cases.

Therefore, based on the author's experience in the chemical industry and the knowledge of published simulation studies, the following four dimensional simulated data set is proposed that will represent historic IOC data. It will be assumed that each of the simulated process variables has some non-linear relationship with the same unobservable underlying variable u. Many examples of this unobservable variable can be identified when considering for example an industrial chemical process. Examples of these unobservable variables in

practice might be catalyst age, catalyst activity and fluctuations in ambient temperature. The definitions of the four simulated process variables are as follows:

$$x_1 = \begin{cases} 1 + N(1, 10), & \text{if } 1 \leqslant u < 30 \\ -60 + 2u + N(0, 10), & \text{if } 30 \leqslant u < 60 \\ 55 + N(0, 10), & \text{otherwise} \end{cases}, \tag{4.70}$$

$$x_2 = \begin{cases} 5 - 0.2u + N(0, 0.5), & \text{if } 1 \leqslant u < 30 \\ -8 + 0.2u + N(0, 0.5), & \text{if } 30 \leqslant u < 60 \\ 28 - 0.4u + N(0, 0.5), & \text{otherwise} \end{cases}, \tag{4.71}$$

$$x_3 = \begin{cases} 80 + N(0, 10), & \text{if } 1 \leqslant u < 30 \\ 20 + N(0, 10), & \text{if } 30 \leqslant u < 60 \\ 50 + N(0, 10), & \text{otherwise} \end{cases}, \tag{4.72}$$

and

$$x_4 = \begin{cases} 10 + \cos(0.5u) + N(0, .5), & \text{if } 1 \leqslant u < 60 \\ 34 - 0.4u + N(0, 0.5), & \text{otherwise} \end{cases}, \tag{4.73}$$

where it is assumed that the underlying variable u is distributed uniformly in the interval $[1, 100]$. The simulated data set of the four process variables was constructed using 2000 realizations of the unobservable variable u. The simulated variables are graphed against u in Figure 4.1 to illustrate the nonlinear relationships. Motivational examples from practice readily exist that can be used to illustrate how healthy process profiles similar to that of these four variables might come about. Consider for example the first process variable x_1. A similar IOC profile can be constructed if x_1 represents the vibration signal emitted by an optimally operating wind turbine. In this example the unobservable variable that is difficult to measure might be the effect of wind. Different vibration signals describing a healthy wind turbine can be obtained as a function of wind speed, or the direction at which the wind is hitting the wind turbine. Therefore, by evaluating the vibration signal profile at different wind speeds, for example, a profile similar to the top left graph of Figure 4.1 can possibly arise. It should be noted that the wind turbine example is just to

explain how process variables can be influenced by unobservable variables, it was not used to generate variable x_1, and is not intended for generalisation to all industrial processes. The scale of the simulated process variables is unit free.

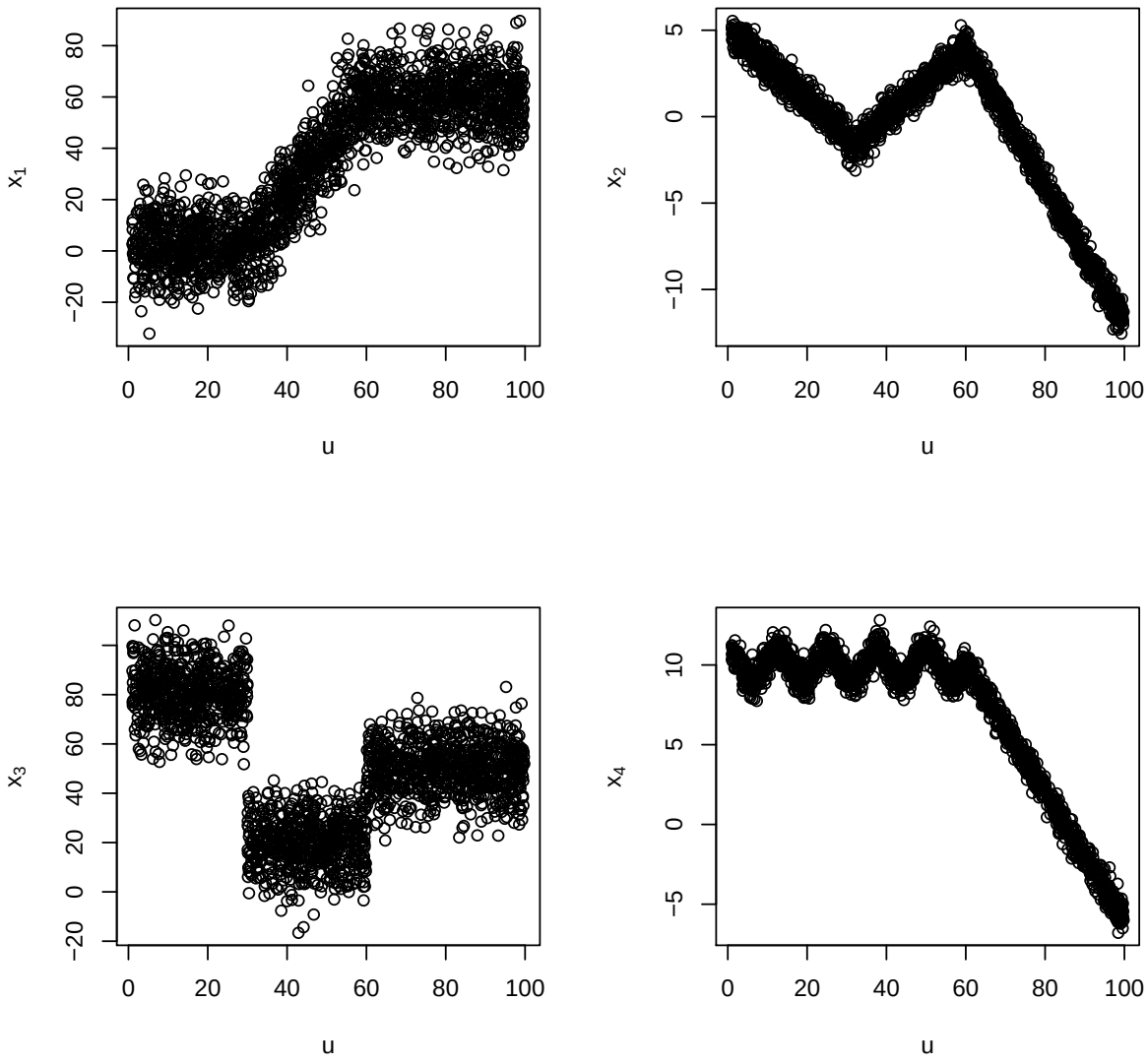


Figure 4.1: Simulated process variables plotted against the unobserved variable u.

Note that for the purposes of this simulation study it will be assumed that no knowledge, or measurements, exists of the underlying variable that influences the process being monitored. In the wind turbine example the speed at which the wind blows the moment it hits the wind turbine might be difficult to measure. There is also no order in which the u values are generated i.e., the process phases observed as a function of u are random.

In other words, it is assumed, for example that when process variable x_1 is observed, as a function of time, it randomly takes on a value between -20 and 80. This is not an unusual assumption. In the wind turbines case a gusting wind might have been observed i.e., the speed of the wind can randomly fluctuate, which means that the vibration signal will correspondingly appear to be random. It is obvious that if measurements of the underlying variable, which are not process specific, were available, the task of monitoring and diagnosis would be considerably easier. In this simulation study, for example, the IOC data can be grouped based on the range of u. Separate linear PCA models can therefore be fitted and used to monitor new data relative to each subset of IOC samples.

4.5.3 PCA and kernel PCA results

Before giving a demonstration of the diagnostic application of the above mentioned established methodologies, it will first be illustrated why PCA is not the appropriate tool to use in this setting. Firstly, one can be nudged in the direction of using kernel PCA when evaluating the scatter plot matrix in Figure 4.2. As previously mentioned, the objective of PCA, in general, is to specify low dimensional surfaces, such that when the data vectors are projected onto these surfaces the variance of the projections is optimal. If the simulation study was only based on two process variables, for example x_1 and x_2, then clearly from Figure 4.2, the surface which will visually result in the maximum sample projection variance is non-linear. That is, linear PCA will not be optimal. The second approach to illustrate the non-applicability of linear PCA is to observe its performance when evaluating a known process fault. Linear PCA will therefore be used to summarize the simulated IOC data set and the resulting fault detection indices will be used to identify if the new faulty sample is different from IOC conditions.

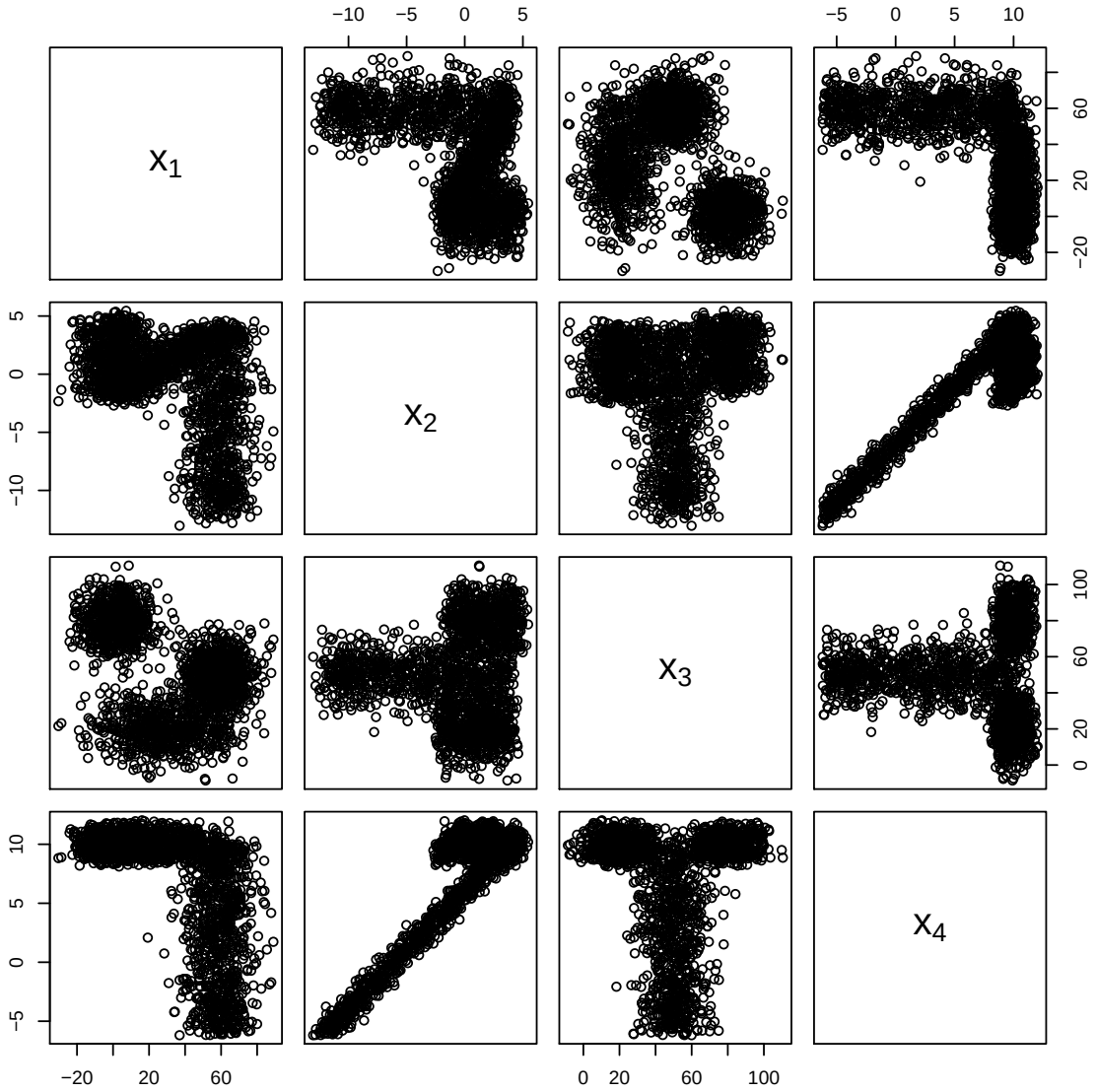


Figure 4.2: Scatter plot matrix of simulated process variables.

A known fault example can be realised by using information on the underlying variable u. The fault that will be simulated is a single sensor fault induced on the first process variable. Consider the IOC process conditions when the unobservable u is larger than 60. Process variable x_1 would therefore be faulty if it does not produce a value that falls within the range induced by $55 + N(0, 10)$. The following four dimensional vector,

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T = \begin{bmatrix} 10.58 & -4.97 & 50 & 2.29 \end{bmatrix}^T, \tag{4.74}$$

can therefore be considered different from the IOC data. All of the process variables ex-

cluding x_1 fall within their respective ranges for $u > 60$, see Figure 4.1. It should again be emphasised that the variable u is not known in practice. In this simulation example it is convenient to use the knowledge of u to simulate known faulty samples.

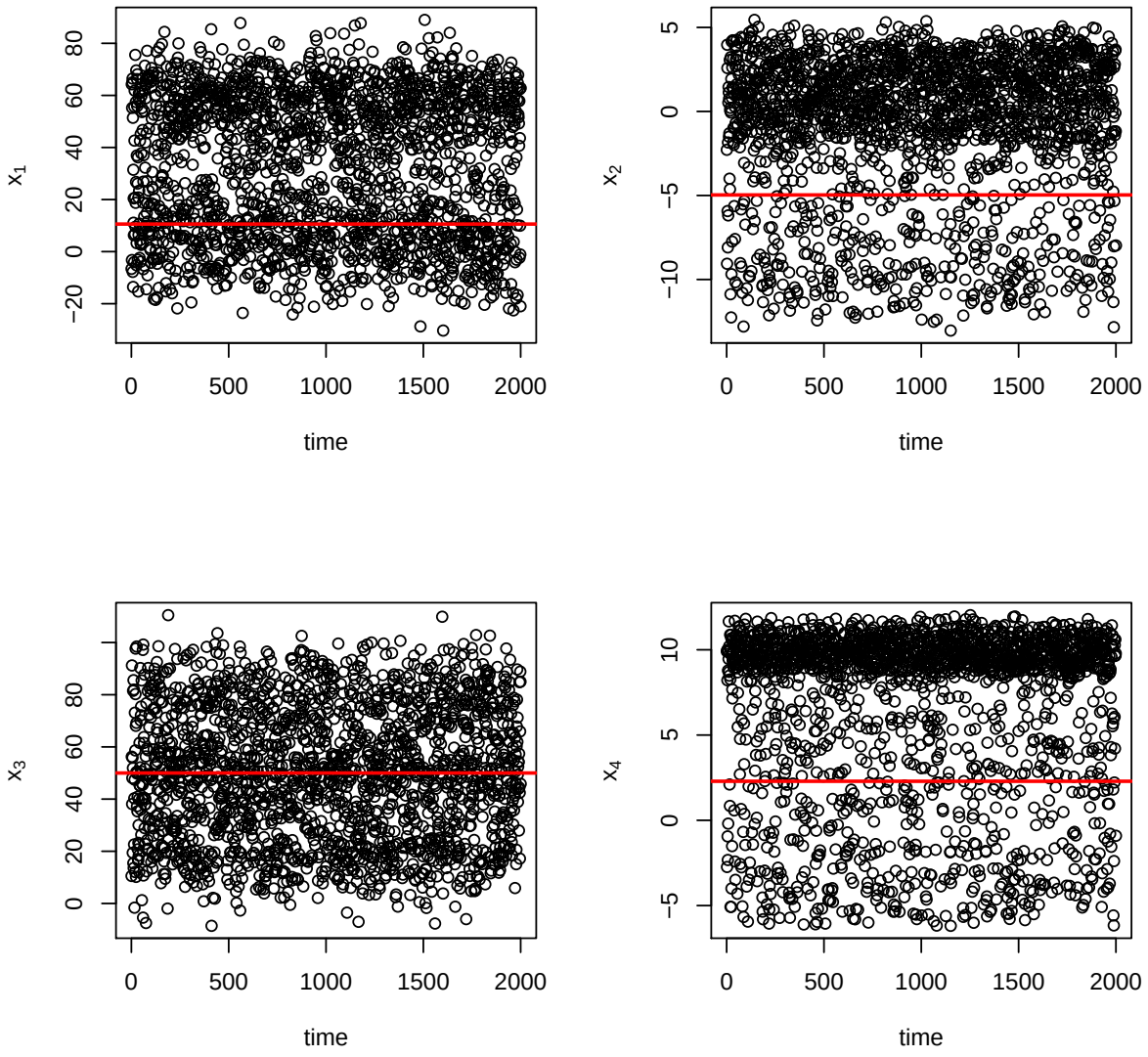


Figure 4.3: Simulated process variables plotted against time. The horizontal lines display an example of a faulty observation.

Consider a typical analysis performed by an engineer. In practice the engineer will typically do a univariate analysis of these process variables on the control panel. Since the engineer is not aware of u, he won't be able to observe the deviation in the one dimensional setting. Figure 4.3 displays time plots of the process variables with the values of the faulty vector illustrated as horizontal red lines in each graph. It is clear from these

graphs that it is difficult to deduce that the process has deviated and specifically that x_1 was the cause. Univariate fault detection and diagnosis would have been possible if u was measured. This simulated example will therefore again reaffirm the importance and relevance of multivariate statistical process monitoring, whether linear or non-linear.
The principal component analysis of this data set resulted in the 2-dimensional score plot given by Figure 4.4. The first two principal components account for 87.04% of the variability in the reference data set. The open circles in Figure 4.4 represent the IOC samples and the solid triangle represents the model evaluation of the faulty sample.

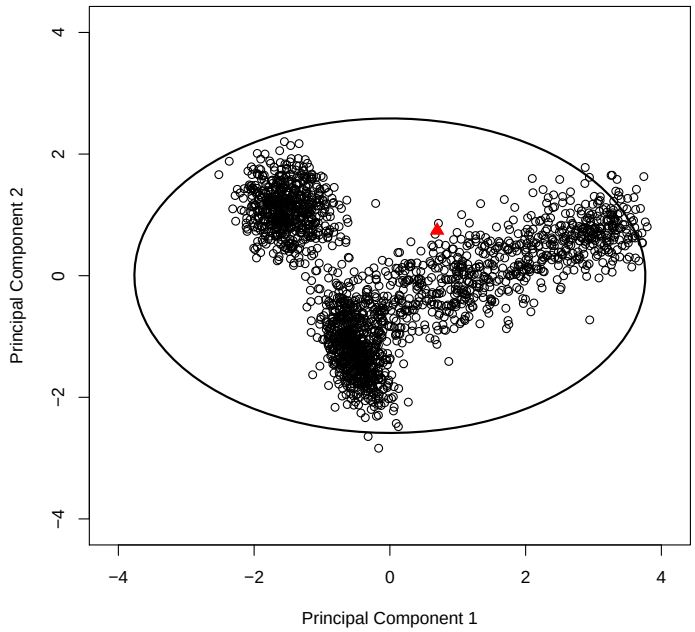


Figure 4.4: PCA score plot of first two principal components. Open circles represent the IOC cases. The solid triangle is the projection of the faulty example. The ellipse identifies the 2-dimensional T^2 95% control region.

From visual inspection, it is clear from Figure 4.4 that the score representation of the fault lies to the edge of the point cloud representing the IOC scores, suggesting that the faulty score is different. When the PCA fault detection statistics are used to quantify the difference it is found that there is not enough evidence to confirm the visual inclination. Also included in Figure 4.4 is the 95% T^2 control ellipse. From this it is clear that the new sample falls inside the control ellipse and can therefore not be considered to be different from the IOC samples. In this example it is therefore illustrated that linear PCA is not the optimal method to use.
In contrast to PCA, the fault detection performance of PCA applied in high dimensional feature space, using kernel approximations, will now be presented for this example. The

Gaussian kernel function, as defined in (4.25), was used to estimate the feature space inner products of the pairwise IOC samples. Optimal hyperparameter and dimension settings for this kernel PCA model was calculated using the pre-image based methodology of Alam and Fukumizu (2014). Using crossvalidation, as described by Algorithm 1 in Section 4.2, the ideal value for σ was determined to equal 0.82 while the number of kernel principal components selected h^* was found to be 45. The likeness of the artificially generated faulty sample to the IOC data vectors can now be calculated using (4.52), (4.61) and (4.63) for the T^2, SPE and combined statistics respectively. It was stated earlier that the control limits of these statistics are calculated under the assumption of normality. Since normality is often violated in practice, it is proposed that the null distribution of the fault indices be non-paramatically estimated using the IOC data. Specifically, the observed value of the fault detection statistic can be calculated for each of the IOC data points. The observed distribution of these IOC detection statistic values can then be used as the null distribution against which future observations can be compared. In Figure 4.5 the observed IOC densities of the respective fault identification statistics are displayed. The vertical red lines in Figure 4.5 represent the theoretically calculated control limits of the different fault indices. It is the opinion of the author that these limits are too stringent and therefore suggests that the empirical percentiles of the IOC fault statistic values be used as control limits. These data driven control limits are represented by the blue vertical lines in Figure 4.5. The numerical values of the control limits are captured in Table 4.2.

Table 4.2: Observed kernel PCA control limits.

Fault Statistic	Theoretical CL - 90%	Theoretical CL - 95%	90% Percentile	95% Percentile
T^2	57.51	61.66	77.66	94.55
SPE	0.051	0.054	0.080	0.120
$Combined$	17.31	18.37	2.79	3.67

Table 4.3: Fault statistic values for faulty sample.

T^2	SPE	$Combined$
92.83	0.74	15.02

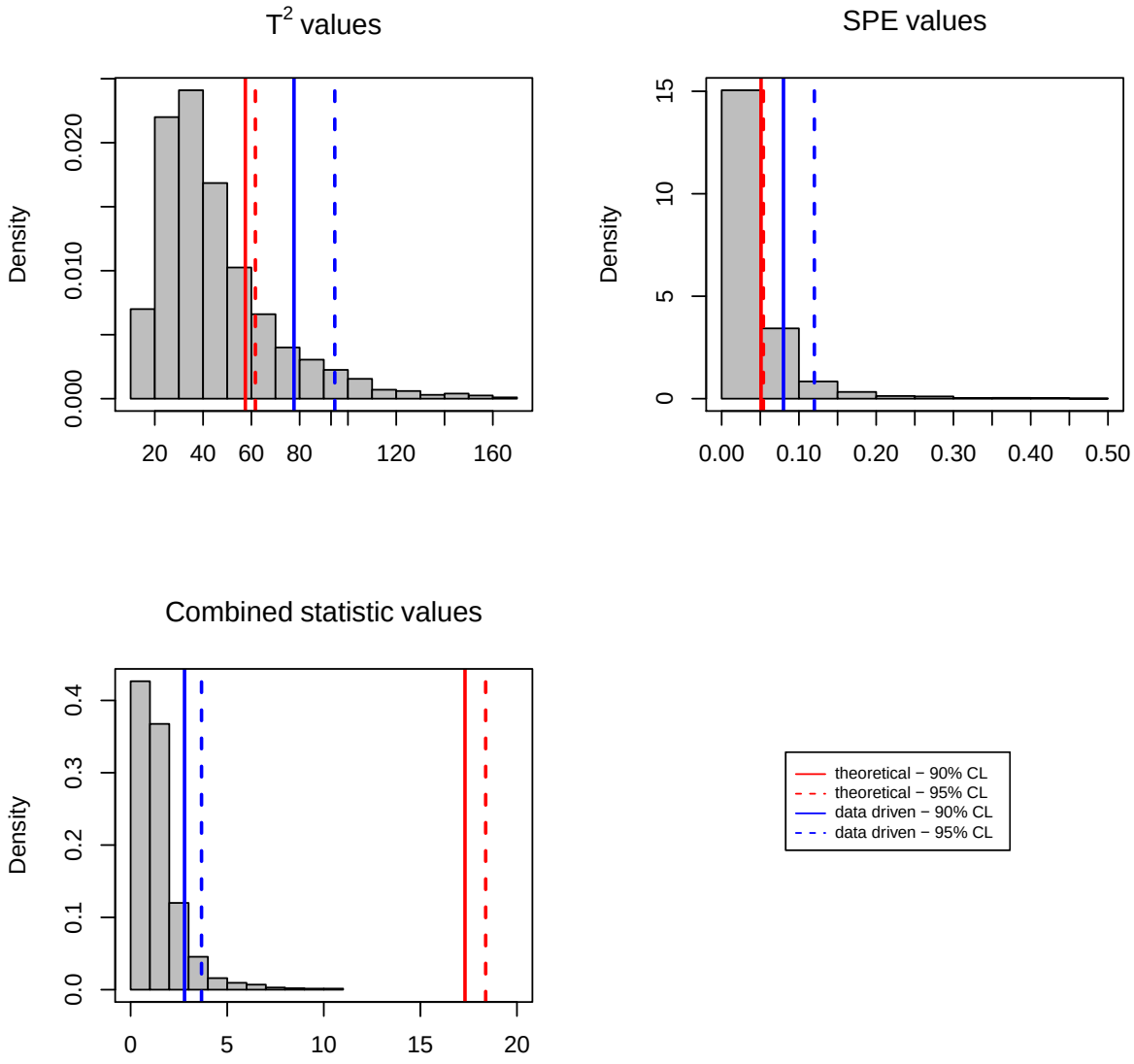


Figure 4.5: Observed distributions of the fault detection statistics together with the control limits.

The fault statistic evaluations for the constructed faulty sample are reported in Table 4.3. It can be observed that the faulty sample is correctly reported as being different from the IOC samples by the SPE and $Combined$ statistics using the non-parametric percentiles. However, the T^2 statistic, based on the non-parametric control limit, identifies the sample as not being different at the 95% confidence level. It should be noted that the T^2 value of the faulty sample is relatively close to the control limit. The sample is identified as being extreme when using the 90% CL. It can therefore be concluded in this example, in comparison to PCA, that the kernel PCA process control method displays the desired fault detection performance.

Kernel PCA based fault diagnosis can therefore now be performed by applying the contribution analysis strategies of Section 4.4 to the out-of-control statistic values. The recommendations made by the different kernel PCA fault diagnosis methods are given in Figures 4.6 and 4.7. In Figure 4.6 the variable contributions as developed by Cho et al. (2005) are displayed for the fault statistics of the faulty sample. Note that the contribution results of the T^2 statistic has been included even though it was not out-of-control based on the 95% control limit. The relative contribution of variable x_1 to all of the fault statistics is correctly ranked as being the largest. The method of Cho et al. (2005) therefore correctly diagnosed the single sensor fault in the first process variable.

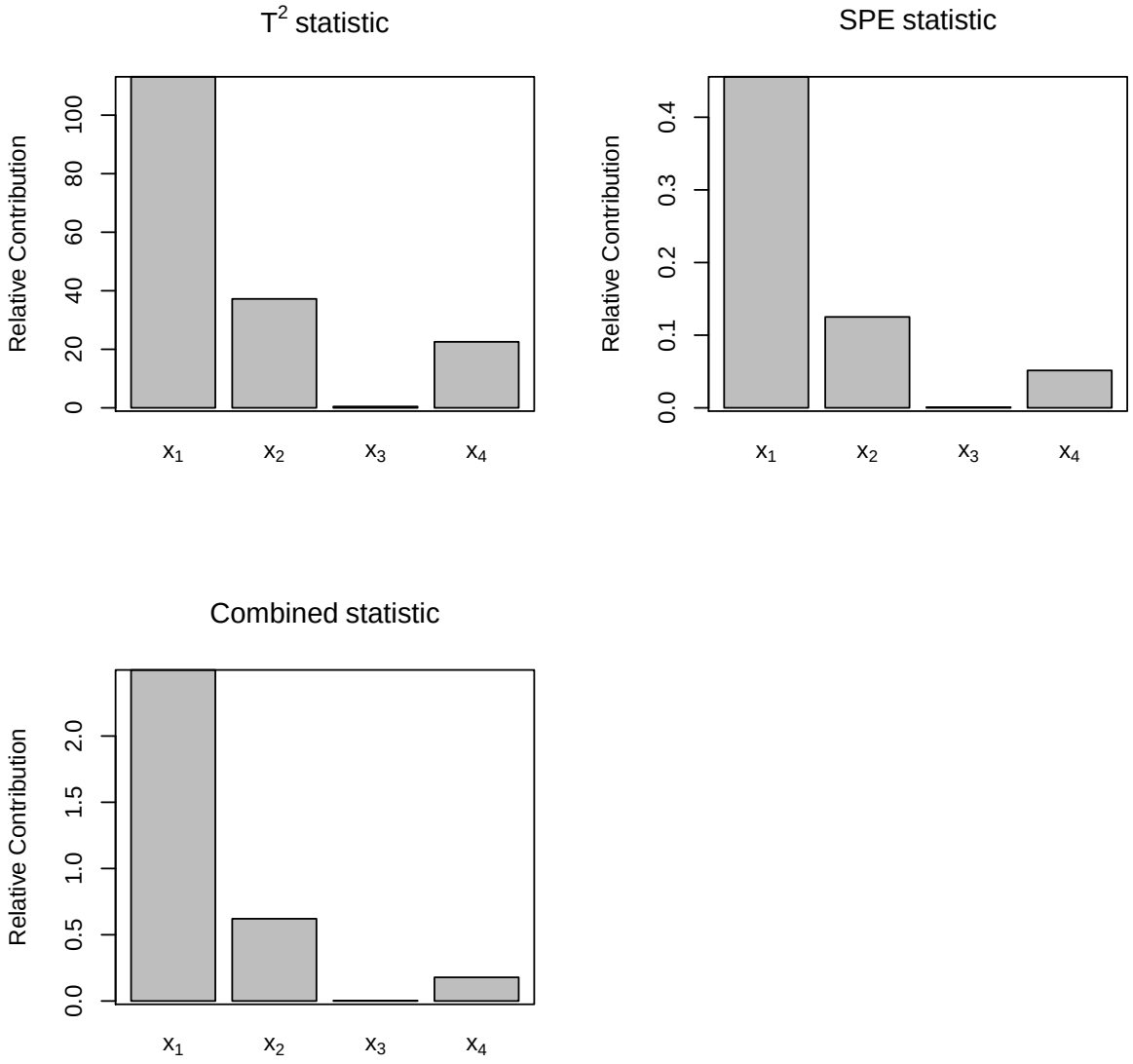


Figure 4.6: Kernel PCA variable contributions calculated for the simulated faulty sample using the method of Cho et al. (2005).

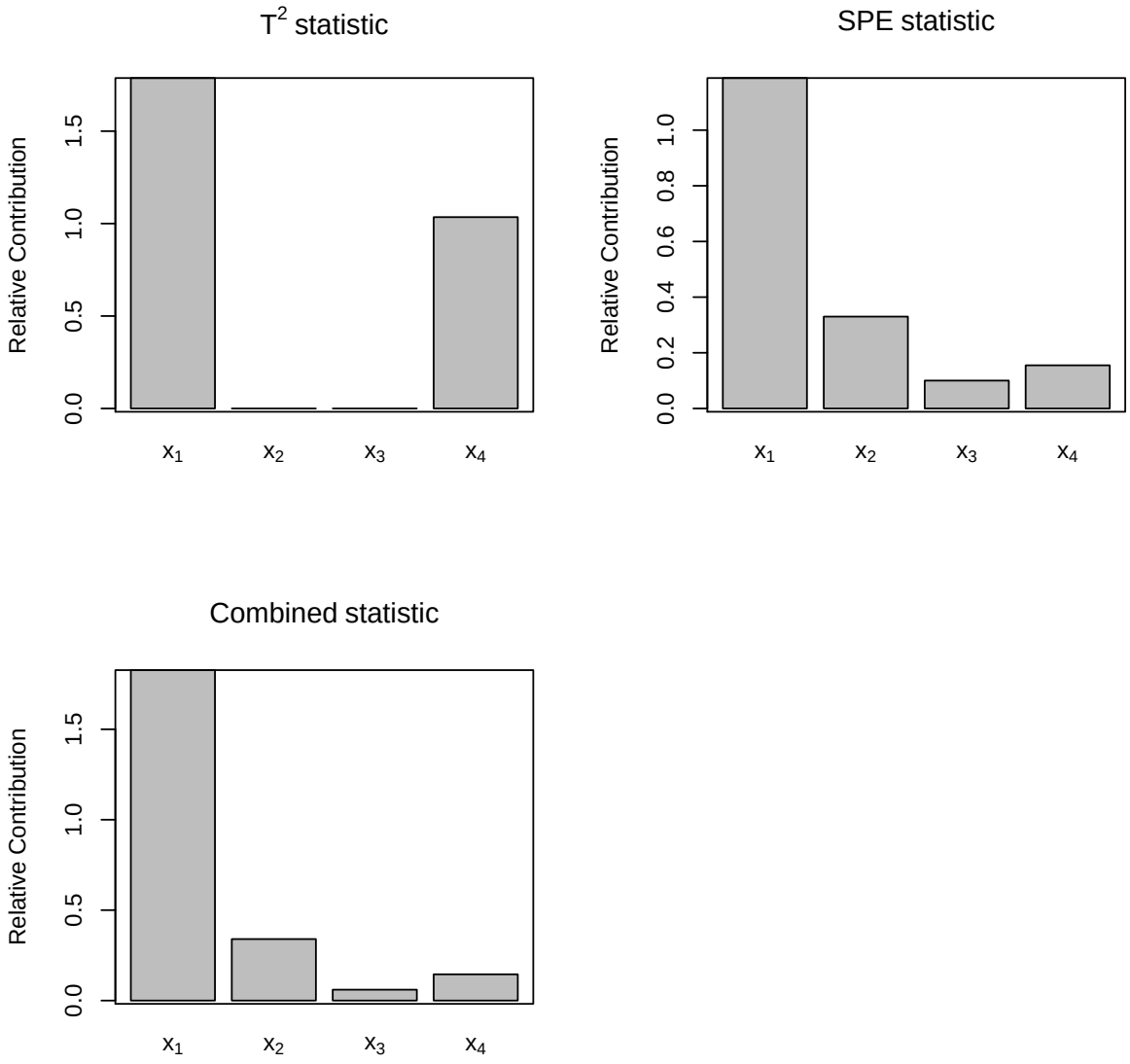


Figure 4.7: Kernel PCA variable contributions calculated for the simulated faulty sample using the method of Alcala and Qin (2010).

The results of the reconstruction optimisation diagnostic approach developed by Alcala and Qin (2010), see Section 4.4, are captured in Figure 4.7. This technique also correctly identifies variable x_1 as being responsible for the process deviation.

Although the fault diagnosis techniques correctly identified the process variable that was altered to have assignable cause variability, it is not clear how an engineer would use this information i.e., Figures 4.6 and 4.7 to correct the multivariate deviation. The univariate ranking of the process variables will prompt the engineer to consider the graphs in Figure 4.3. Specifically, the engineer will look at the univariate graph for x_1, in Figure 4.3, and will not know why this process variable is supposedly important. The existing diagnostic

techniques therefore do not give practical guidance as to what needs to be done in order to achieve multivariate common cause variability.

4.5.4 Kernel PCA fault diagnosis simulation study

In this section the simulation study presented in the previous section will be extended in order to evaluate the fault diagnostic ability of the methods introduced by Cho et al. (2005) and Alcala and Qin (2010). The objective is to do a study similar to the one presented in Chapter 2. That is, the correct diagnosis rates of single, multiple and multivariate sensor faults will be presented for various fault magnitudes. Refer back to Section 2.6 for the definition of these fault types. In this simulation the fault free part of the different sensor faults was again determined using the three different fault detection statistics i.e., (4.52), (4.61) and (4.63). The results presented here will give a general indication of the applicability of the diagnosis methods in practice.

4.5.4.1 Single sensor fault

Single sensor faulty vectors were constructed using the formulation given by (2.68). For each of the fault free data sets the following were done to simulate the faulty vectors:

1. For each IOC sample select a process variable j in turn,
2. add the fault magnitude $\pm f$ to the selected process variable,
3. where the fault size f is one of 9 equally spaced fault magnitudes between 0.3 and three.

These observation vectors were then filtered to identify only those that are out-of-control based on the corresponding statistic's critical value. In addition the data were filtered to retain only those vectors that are univariately inside the IOC ranges. Note that the way in which the faulty vectors are obtained is slightly different to the approach that was followed in Chapter 2. This was done to limit the number of possible single sensor faults, due to the kernel PCA evaluations being computationally expensive. The fault diagnosis techniques presented in Section 4.4 were applied to the final set of single sensor vectors to identify the faulty variables.
The percentage of samples that were correctly diagnosed by the two diagnostic methods is plotted for each of the absolute fault magnitudes. These results are presented in Figures 4.8, 4.9 and 4.10 for each of the three faults statistics respectively.
From Figure 4.8, the results for the SPE statistic, indicate that the gradient based method

of Cho et al. (2005) resulted in a better correct diagnosis rate when compared to the results for the reconstruction based method of Alcala and Qin (2010). It is also observed that the correct diagnosis rate never exceeds 80%.

The simulation results based on the T^2 index are displayed in Figure 4.9. It is observed that for most of the fault sizes the methods of Cho et al. (2005) and Alcala and Qin (2010) yield comparable results. However, for fault magnitudes larger than two, the reconstruction based method consistently outperforms the gradient based method. The best correct diagnosis rate equals approximately 98% and is observed at a fault size equal to three.

For the *Combined* statistic, in Figure 4.10, it is observed that the gradient based method yield better results as compared to the reconstruction based method for each fault size. The maximum correct diagnosis rate equals 80% which is observed at a fault magnitude that approximately equals 1.45.

Collectively the results presented in Figures 4.8, 4.9 and 4.10 suggest that the fault diagnosis methods of Cho et al. (2005) and Alcala and Qin (2010) would not be ideal to diagnose single sensor faults in practice.

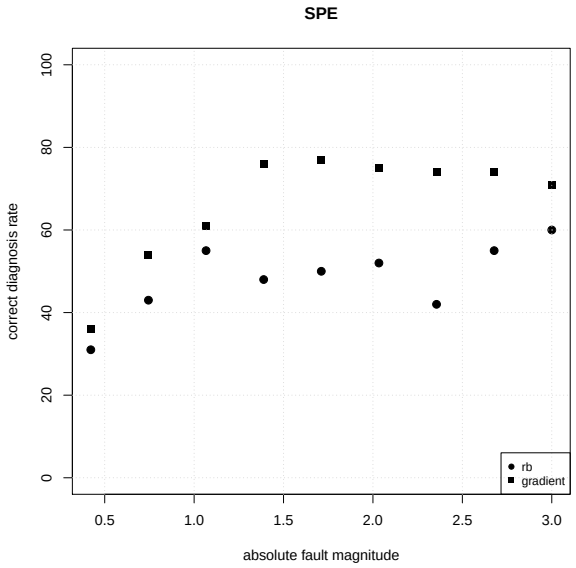


Figure 4.8: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

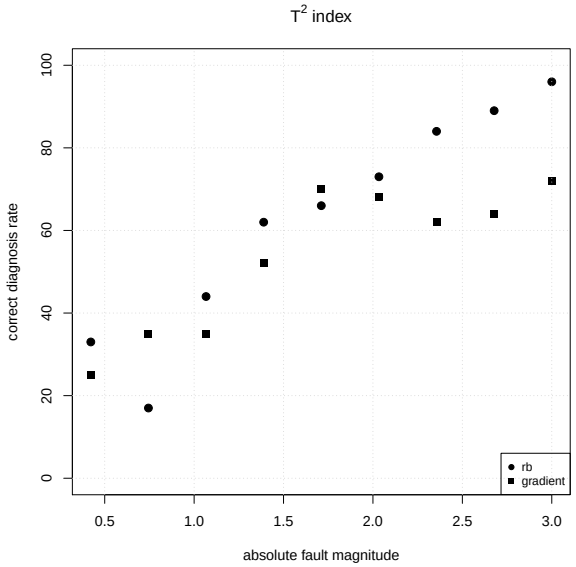


Figure 4.9: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

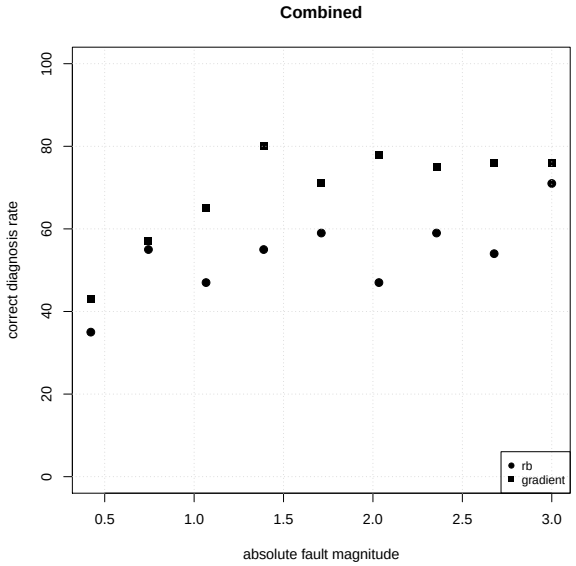


Figure 4.10: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

4.5.4.2 Multiple sensor fault

The performance of the faults diagnosis methods prescribed by Cho et al. (2005) and Alcala and Qin (2010) will now be evaluated using multiple sensor faults. We will employ the definition used in the simulation section of Chapter 2 i.e., equation (2.69), to construct multiple sensor examples in this section. The following was done to simulate the multiple sensor faulty vectors for each of the fault free data sets:

1. For each IOC sample select process variable pair (i, j) in turn,
2. add the fault magnitude $\pm f$ to the selected process variable pair using the formulation of (2.69),
3. where the fault size f is one of 9 equally spaced fault magnitudes between 0.3 and 1.5.

The collection of possible multiple sensor fault vectors obtained were then filtered to identify only the ones that are out-of-control based on the corresponding statistic's critical value. Note that the multiple sensor conditions that are specified in Chapter 2 i.e., that the single sensor vectors that define a multiple sensor fault should also be out-of-control, were also checked. Filtering was also applied to remove samples that violated the univariate bounds of the IOC data. The fault diagnosis methodologies of Cho et al. (2005) and Alcala and Qin (2010) were then utilised to analyse the final multiple sensor fault vectors. The results of the analysis are depicted in Figures 4.11, 4.12 and 4.13 for each of the respective fault detection indices. In each graph the percentage of variable pairs that were correctly identified by the two diagnosis methods are plotted against each of the corresponding absolute fault magnitudes.

In Figure 4.11 the results for the SPE statistic are reported. An holistic view of Figure 4.11 leads us to conclude that the fault diagnosis techniques yielded results that are comparable. The highest correct diagnosis rate, equalling 90%, is observed for the gradient based method at a fault magnitude of 0.3. However, most of the correct diagnosis rates recorded are below 50%.

In Figure 4.12 it is observed, for the T^2 statistic, that none of the techniques reported results that exceeded a correct diagnosis rate of 40%. The performance of the gradient based method is slightly better than the reconstruction based method.

For the *Combined* index, in Figure 4.13, it is observed that the results presented indicate that the performance of the two methods are comparable. The highest correct diagnosis rate of 90% is observed at a fault magnitude of 0.3. While the lowest correct diagnosis rate equalled approximately 30% and was observed at a fault size of 1.38.

Collectively the results presented in Figures 4.11, 4.12 and 4.13 indicate that the fault

diagnosis methods of Cho et al. (2005) and Alcala and Qin (2010) would not be ideal to diagnose multiple sensor faults in practice.

It is interesting to observe the general downward trend of the correct diagnosis rates in Figures 4.11 and 4.13. One would have expected that the correct diagnosis rate should improve with increased fault magnitude. However, this behaviour can be expected due to the non-linear characteristics of the data. Consider the two dimensional scatter plot of the first two simulated variables, illustrated in Figure 4.14, as an example. The red square represents an out-of-control observation, which is a single sensor deviation from the blue square in the x_1 direction. The horizontal broken line represents the magnitude of the single sensor fault. The size of this deviation is relatively large when compared to the range of x_1. It is observed from Figure 4.14 that the out-of-control red square is not remarkably different from the IOC data, even though the fault magnitude is large. A contrasting observation is made when the out-of-control observation represented by the red triangle in Figure 4.14 is considered. This observation is a single sensor deviation in the x_2 direction from the in-control blue triangle. The size of the deviation is represented by the vertical broken blue line. It is observed from Figure 4.14 that even though the vertical fault magnitude is smaller than the horizontal fault size, the red triangle is more remote from the IOC data than the red square. Therefore, this conceptual illustration argues that large fault sizes do not necessarily determine the severity of the out-of-control deviation and the subsequent complexity of the fault diagnosis. The graphs representing correct diagnosis rate as a function of fault size should therefore be interpreted with this in mind for the non-linear case.

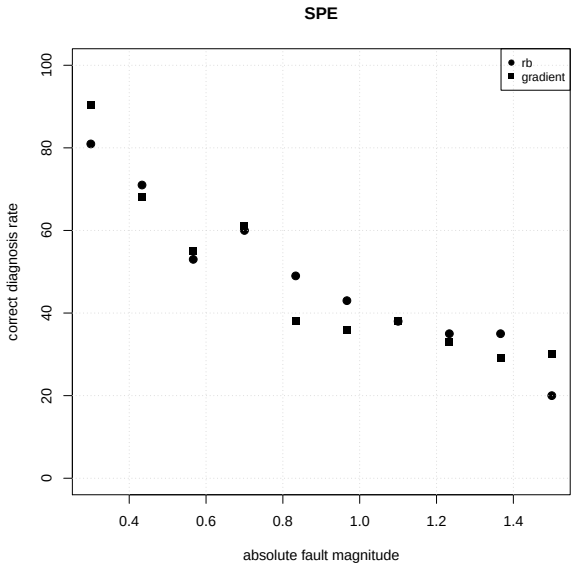


Figure 4.11: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

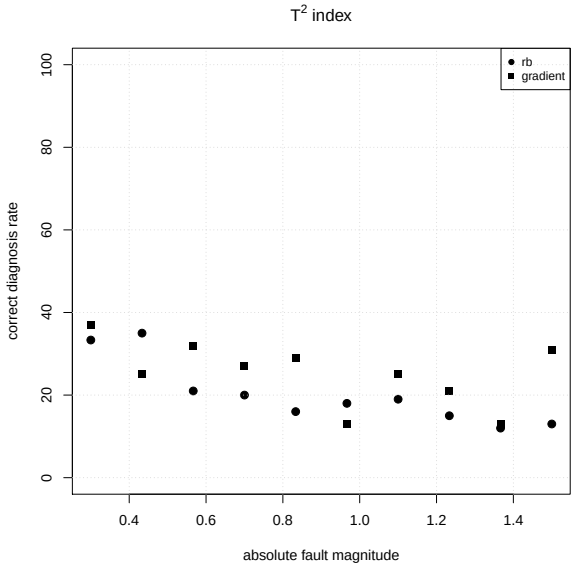


Figure 4.12: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

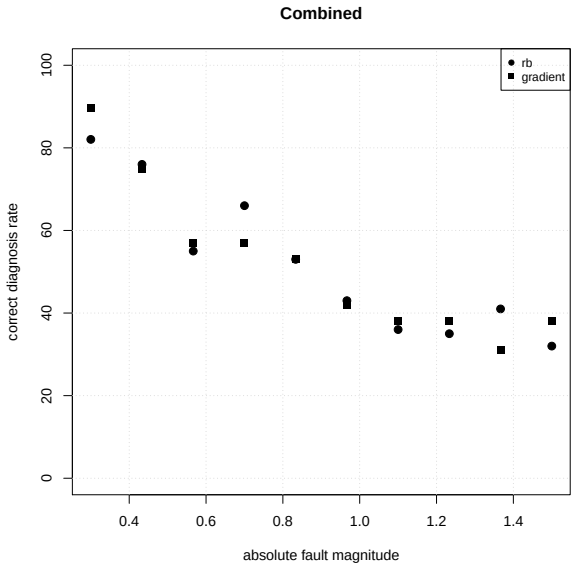


Figure 4.13: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

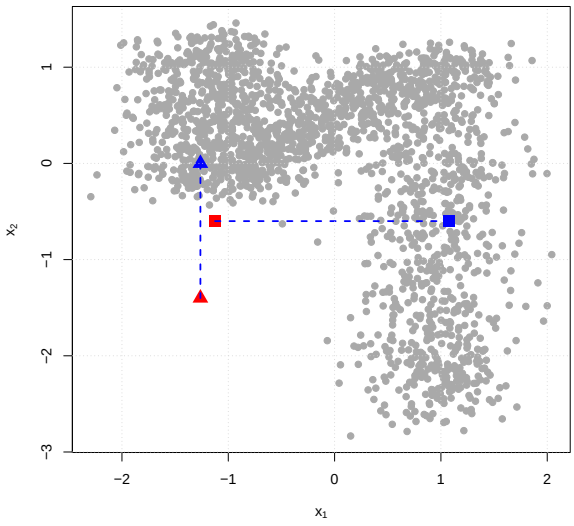


Figure 4.14: Scatter plot of x_1 versus x_2 used to demonstrate that a large fault size does not necessarily imply a large multivariate deviation in the non-linear setting.

4.5.4.3 Multivariate sensor fault

The gradient based fault diagnosis method of Cho et al. (2005) and the reconstruction based method of Alcala and Qin (2010) will now be evaluated using simulated multivariate

sensor faults. The simplified definition of a multivariate sensor fault used in the simulation study of Chapter 2 i.e., equation (2.70), will also be used in this simulation. The following steps were followed to generate multivariate sensor faulty vectors for each of the three fault free data sets:

1. For each IOC sample select process variable pair (i, j) in turn,
2. add the fault magnitude $\pm f$ to the selected process variable pair using the formulation of (2.70),
3. where the fault size f is one of 9 equally spaced fault magnitudes between 0.3 and 1.5.

The appropriate statistic's critical value was now used to filter the collection of possible multivariate sensor fault vectors in order to identify the out-of-control samples. Note that the multivariate sensor conditions that are specified in Chapter 2 i.e., that the single sensor vectors that define a multivariate sensor fault should be in-control, were also checked. Filtering was also applied to remove samples that violated the univariate bounds of the IOC data. Figures 4.15, 4.16 and 4.17 visualise a summary of the simulation results for each fault detection index. For the SPE index, it is observed in Figure 4.15 that the reconstruction based method mostly reported a correct diagnosis rate that is higher than that of the gradient based method. The correct diagnosis rate for the reconstruction based method are predominantly above 80% while the gradient based method recorded results that for the most part fell below 60%.
In Figure 4.16 it is observed, for the T^2 index, that none of the diagnosis methods reported results that exceeded a correct diagnosis rate of 70%. The reconstruction based method recorded slightly better results when compared to the gradient based method. There are instances where the gradient based method yielded correct diagnosis rates that equalled 0%.
The graph for the *Combined* index in Figure 4.17 indicate that the reconstruction based method yielded results that are preferred to that of the gradient based approach. A correct diagnosis rate greater than 80% is prevalent for the reconstruction based method while the gradient based method report results that are predominantly below 60%.
Collectively the results presented in Figures 4.15, 4.16 and 4.17 indicate that the fault diagnosis methods of Cho et al. (2005) and Alcala and Qin (2010) would not be ideal to diagnose multivariate sensor faults in practice.

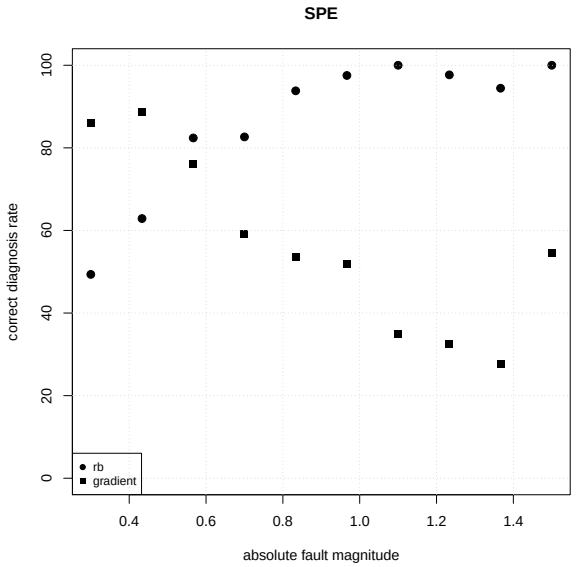


Figure 4.15: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

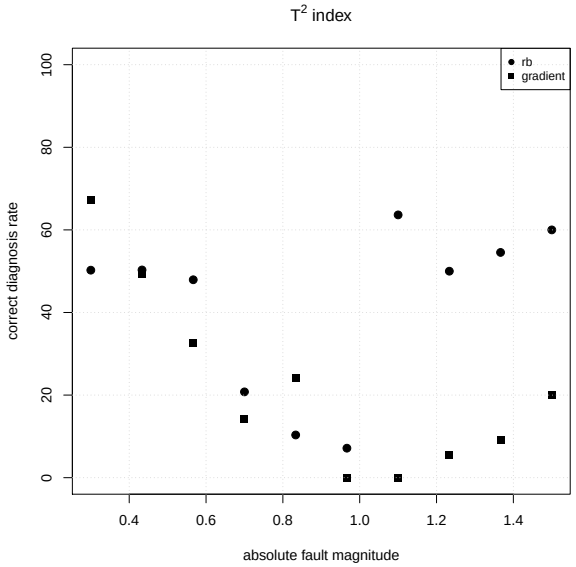


Figure 4.16: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

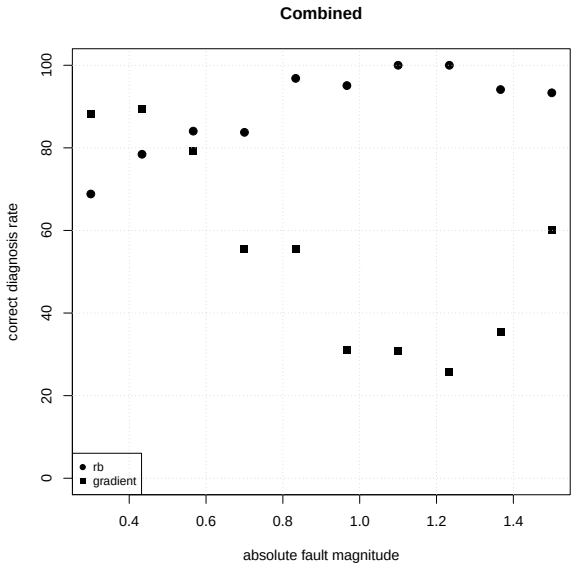


Figure 4.17: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles.

4.6 Summary

Modern process plants are fitted with equipment that record data on a large number of performance characteristics. This growth in information on manufacturing processes has made statisticians realise that the relationships that exist amongst process variables are oftentimes highly non-linear (see Cho et al. (2005)). Many examples therefore exist where multivariate process monitoring based on linear PCA is not suited to model non-linear common cause variability. Alternative techniques have therefore been developed to account for the complex relationships between process characteristics.

In this chapter it was illustrated how kernel PCA is currently employed for multivariate statistical process monitoring. Specifically, in addition to a summary of kernel PCA theory, a new simulation example was constructed to illustrate how kernel PCA can be used in practice for fault detection and diagnosis. A discussion was provided to motivate how a data set similar to the simulated one might realise in practical situations. A typical fault was constructed to illustrate the fault identification ability of kernel PCA compared to linear PCA. In this example the fault identification statistics based on kernel PCA were able to detect that the simulated faulty sample was statistically different from common cause variability. The linear PCA based fault identification statistics did not identify that the simulated faulty sample is foreign to the process. The need for non-linear techniques

such as kernel PCA have therefore been motivated.

Fault diagnosis techniques employed for kernel PCA are in principle similar to the methods developed for PCA, as discussed in Chapter 2. Two fault diagnosis approaches are reviewed. The first approach is the gradient based method of Cho et al. (2005) and the second is the reconstruction based method of Alcala and Qin (2010). Similar to PCA contribution analysis, the kernel PCA diagnosis techniques attempt to attach individual importance to process variables. In the simulation example, it was illustrated how this is typically done. Additionally, a simulation study, similar to that presented in Chapter 2, was performed to evaluate the performance of the traditional diagnosis techniques proposed for kernel PCA. The correct diagnosis rates for single, multiple and multivariate sensor faults were evaluated. It is concluded from the results obtained that these diagnosis approaches are not suitable for application in the practical setting.

The concerns that were highlighted for the PCA diagnosis techniques, in Chapter 3, extend to the approaches used in kernel PCA. Individual importances of process variables do not provide the engineer with an intuitive understanding of how to interpret the, statistically identified, non-linear and multivariate deviation. A different approach to fault diagnosis in kernel PCA based monitoring is therefore required.

Chapter 5

New Methodology for Kernel PCA Fault Diagnosis

5.1 Introduction

A novel way of performing the fault diagnosis step in kernel PCA based statistical process control will be motivated in this chapter. The methodology that will be proposed is an analysis that is completely independent from the kernel PCA model used to identify process drift. A new contribution to fault diagnosis is therefore suggested which decouples the diagnostic step from the fault identification step. That is, in contrast to the diagnostic methods described in Section 4.4, the out-of-control fault identification statistics will not be further analysed during the diagnostic process. The general principle that motivated the development of the proposed analysis is an extension of the methods developed in Chapter 3 for PCA fault diagnosis.

The reasons provided in Chapter 3, to motivate the change in diagnostic philosophy in PCA fault investigations, hold true for the kernel PCA setting as well. As an illustration, consider the pairwise scatter plot in Figure 5.1. The IOC point cloud displayed in Figure 5.1 represents a typical non-ellipsoidal shape that can be expected when non-linear relationships are present. Similar to the discussion in Chapter 3 it is illustrated that it is possible to simulate the same single sensor fault in different ways. In certain settings the single sensor fault is caused by a deviation in the first process variable, while in other settings abnormalities in the second process variable are the cause. The traditional methodology of attaching importance to individual process variables is therefore not appropriate in the kernel PCA setting.

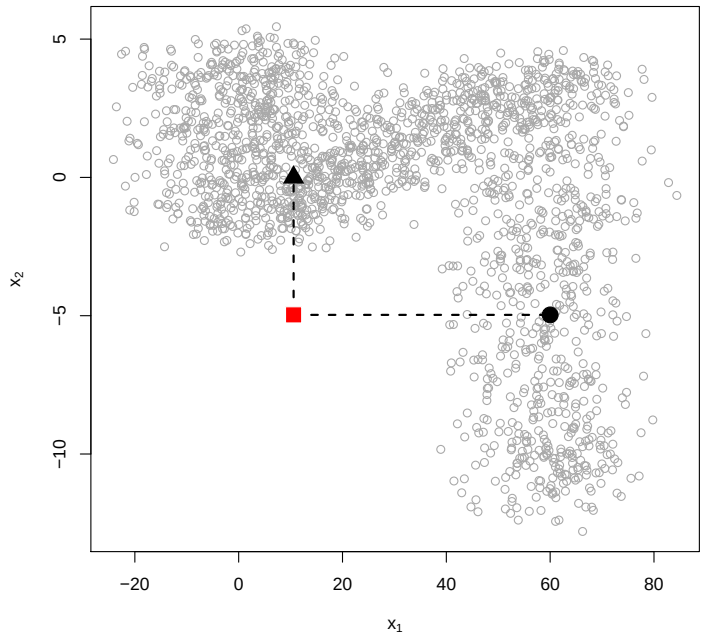


Figure 5.1: Illustrative example, in two dimensions, of a single sensor fault, red square, simulated from two different IOC points. Using the first IOC point, the black circle, to derive the out-of-control point means that only x_1 is altered. For the second derivation, the fault is realised by changing only x_2 in the IOC point represented with the black triangle.

5.2 Pairwise variable importance analysis

In this new diagnosis methodology it is proposed that the pairwise relationships between variables should be monitored and analysed in order to properly understand process deviations. That is, similar to Chapter 3, a method of calculating pairwise variable contributions will be developed. Note that there is no method available of deriving the kernel PCA equivalent of the pairwise variable contribution technique developed in Chapter 3. A new method of ranking variable pairs is derived by considering the scatter plot matrix, of the data simulated in Chapter 4, that is displayed in Figure 5.2.

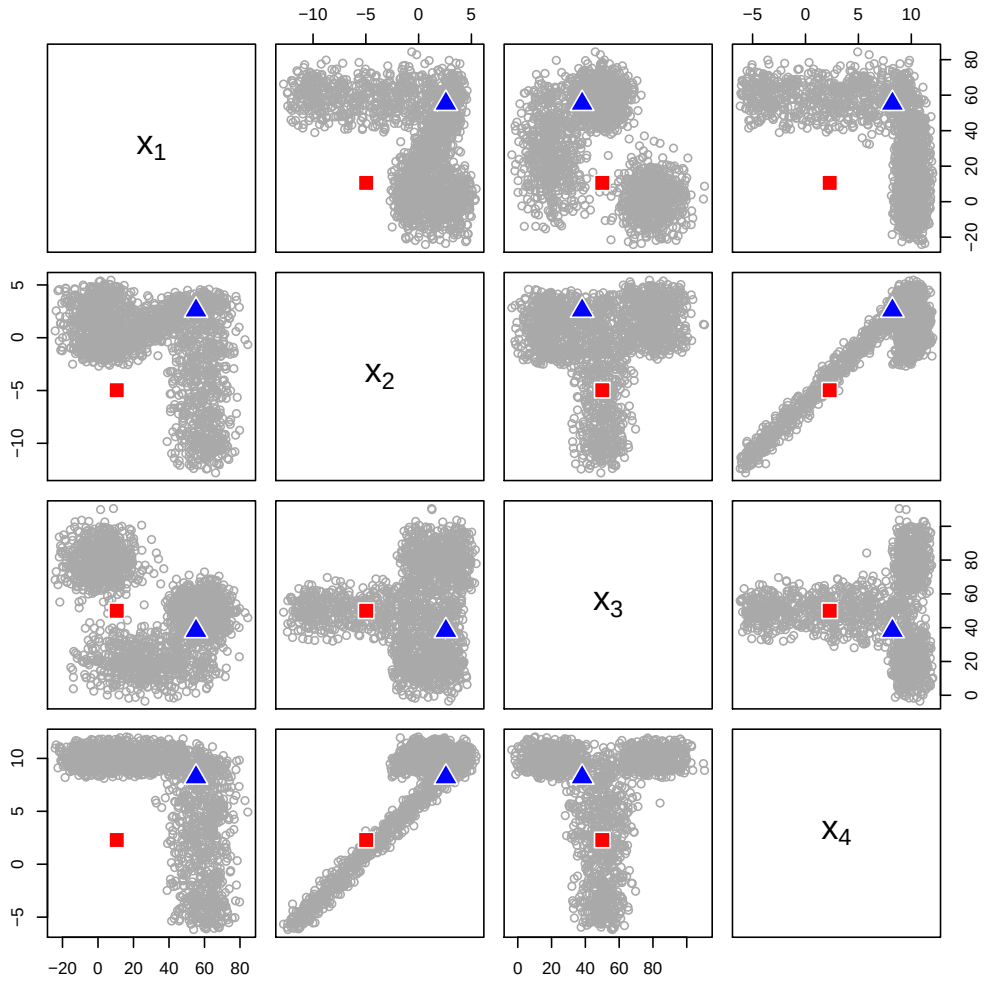


Figure 5.2: Scatter plot matrix of the four simulated process variables of Section 4.5. Grey points are used to display the IOC data. The blue triangle is an example of a process observation that were determined by the kernel PCA T^2 statistic to be in-control. The red square is an example of a sample that was identified to be out-of-control.

The grey circular points represent a subset of the ideal operating condition simulated data set discussed in Section 4.5. The subset of points is those that were determined to be in-control based on the T^2 kernel PCA fault detection statistic at the 95% confidence level. It is therefore clear that during IOC periods that the pairwise relationships between process variables take on a specific shape as described by the point clouds. The bivariate relationships should then simultaneously be adhered to for the process to be considered multivariately in-control. Consider the case where a new observation is determined by kernel PCA to have common cause variability. If this observation is plotted on the IOC scatter plot matrix, it can intuitively be expected that the new point will be 'close' to the IOC points in each of the graphs. The blue triangular point in Figure 5.2 is an example

of a observation that is considered to be in-control based on kernel PCA. Conversely, consider the case where kernel PCA identifies that a new measurement was observed due to assignable cause variability. It is logical to expect that when this observation is plotted on the IOC scatter plot matrix that this point will be 'remote' from the IOC points in at least one of the graphs. The red square shaped point in Figure 5.2 is an example of an observation that was identified to be out-of-control based on kernel PCA.
Let $\mathbf{x}^* : p \times 1 = \begin{bmatrix} x_1^* & , \ldots , & x_p^* \end{bmatrix}^T$ represent the new observation from the process with $\mathbf{x}^*_{(i,j)} = \begin{bmatrix} x_i^* & x_j^* \end{bmatrix}^T$ the bivariate measurement on variables i and j where $i \neq j = 1, \ldots, p$. Define $\mathbf{X}_{(i,j)} : n^c \times 2$ as the matrix of n^c data vectors describing the bivariate IOC samples for variables i and j. In addition let $d_{ij} \geqslant 0$ represent the value of a consistently applied distance measure, f, used to quantify 'closeness' in each of the $\binom{p}{2}$ scatter plots. This distance measure is therefore a function of the IOC data and the new observation i.e.,

$$d_{(i,j)} \quad = \quad f(\mathbf{X}_{(i,j)}, \mathbf{x}^*_{(i,j)}), \tag{5.1}$$

for $i \neq j = 1, \ldots, p$. The distance measure is selected such that it is small when the new point is similar to the IOC cloud of points and large when it is dissimilar.

Definition 5.1 *The distance (5.1) is defined as the contribution of the $(i,j)^{th}$ variable pair to the process deviation.*

Defining pairwise contributions in this manner provide us with a sense of what is expected of any corrective intervention in order to achieve a process that is multivariately in-control. That is, for the variable pairs with large contributions, the engineer should focus on implementing appropriate process adjustments to the controllable variables that will yeild measurements that are 'close' to the corresponding point clouds. Note some variable pairs may include uncontrollable or reponse variables, which although not controllable, may provide additional process understanding and cause and effect relationships.
The significance of the size of a specific pairwise contribution can be quantified by comparing it to the empirical distribution of the distances observed under ideal conditions. Specifically, let

$$d^k_{(i,j)} \quad = \quad f(\mathbf{X}_{(i,j)}, \mathbf{x}^k_{(i,j)}), \tag{5.2}$$

for $k = 1, \ldots, n^c$, represent the distance of the $(i,j)^{th}$ variable pair for the k^{th} in control observation. The n^c distances calculated for the $(i,j)^{th}$ variable pair can then be pooled to represent the IOC conditions against which future distances can be compared. The

distribution of these reference distances can be estimated using empirical strategies such as kernel density estimation (Hastie et al., 2001). For new observations, the variable pairs that correspond to distances that have a low probability of occurring, when compared to their reference distribution, should be analysed further to diagnose process drifts. Using this method to identify only the variable pairs with significantly large distances allows for the consideration of a few variable pairs instead of the complete set of $\binom{p}{2}$ combinations.

5.3 Univariate variable ranking in pairwise importance

In addition to the importance ranking of the variable pairs, the method allows for the ranking of individual variables. This is achieved, similar to the method employed in Chapter 3, by considering a matrix of pairwise contributions. The pairwise distances (5.1) of a specific observation can be captured in a symmetric distance matrix,

$$\mathbf{D} = \begin{bmatrix} 0 & d_{1,2} & \cdots & d_{1,p} \\ d_{2,1} & 0 & \cdots & d_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ d_{p,1} & d_{p,2} & \cdots & 0 \end{bmatrix}, \tag{5.3}$$

with the diagonal elements $d_{ii} = 0$ and off-diagonal entries $d_{i,j} \geqslant 0$ for $i \neq j = 1, \ldots, p$. The distance from the $(i,j)^{th}$ IOC point cloud calculated for the new observation can be interpreted as a dissimilarity between process variables i and j if the following definition is applied:

Definition 5.2 *Process variables i and j are considered similar if the new observation vector $\mathbf{x}^*_{(i,j)}$ is 'close' to the bivariate IOC point cloud $\mathbf{X}_{(i,j)}$ and dissimilar if the new point is 'remote' from the IOC region.*

When viewing the distance matrix in (5.3) as containing the dissimilarity measures between process variables, a natural method of ranking individual variables for fault diagnosis is identified. The variable that will rank high is the one that is the most dissimilar to the other variables. The Multidimensional scaling method, described in Chapter 3, can again be applied to the distance matrix in (5.3) to obtain a graphical representation of the relative distances between variables. A relative ranking is obtained by observing the distances of the points representing the variables to the center of the display. Variables that are the furthest away from the the center, are the ones that are the most dissimilar

to the rest, and will attain the highest contribution ranking.
Another approach that can be used to identify a ranking of the process variables, is to evaluate the row sums of (5.3). The process variable with the highest ranking can be identified as the variable that is the furthest away from the other variables by specifying the variable that corresponds to the row in (5.3) with the largest sum i.e., where

$$\sum_{j=1}^{p} d_{i,j} \tag{5.4}$$

for $i = 1, \ldots, p$, is the largest. The second most important variable to consider as the likely cause for the drift in process conditions is the variable that is the most different in the residual set of process variables. If variable i^* was identified as the most important then the distance matrix of the remaining variables is the one with row and column i^* removed. The variable with the largest row sum in the residual set

$$\sum_{j \neq i^*}^{p-1} d_{i,j} \tag{5.5}$$

for $i \neq i^*$ and $i = 1, \ldots, (p-1)$, is the second most important. By employing this logic, a relative importance ranking can be obtained for $p-2$ of the variables. The importance of the remaining two process variables can be quantified using a future process observation after the fault in the initial set of faulty variables has been addressed. Note that empirical distributions of the reference row sums can be calculated to identify row sums that are significantly large. Refer back to Chapter 3.
As argued previously, it is important to use the individual ranking of process variables in combination with the important contributing pairs to diagnose the faulty state. Notice that this diagnostic procedure, using both pairwise and individual rankings, cannot be used when the process that is being monitored is described by only two variables. These processes are rare and should be easily diagnosable using fundamental process knowledge.

5.4 Distance measure definition

It is clear from the discussion up to this point that this distance based diagnostic methodology is highly dependent on the distance measure used to define the 'closeness' or 'remoteness' of a new observation $\mathbf{x}^*_{(i,j)}$ to its corresponding IOC samples $\mathbf{X}_{(i,j)}$. An obvious requirement is that the distance measure must be the same for all the variable pairs. The way in which it is applied should be consistent to ensure that the resulting distances are

comparable across the different point cloud comparisons.
The author is aware that a large number of strategies can be motivated for the selection of a distance measure that is suitable for application in the new fault diagnosis methodology. This discussion will be limited to examples of possible distance measures that are based on three statistical method classes. The first method can be classified under unsupervised statistical learning. Supervised statistical modelling can be used to classify the second method, while the third method's description is identified under novelty detection. The three examples from these classes, that will be used to develop potential distance measures, are

- Kernel PCA,
- Support vector machine classification and
- One-class support vector machine classification.

The development of distance measures using these methods will now be explored.

5.4.1 Kernel Principal Component Analysis

The distance measure promoted in this section for use in the diagnostic methodology discussed will be based on the analysis used in the fault identification step i.e., kernel principal component analysis. The theory describing kernel PCA was discussed in much detail in Chapter 4. An understanding of kernel PCA is therefore assumed. Furthermore, the notation of Chapter 4 will be used in this discussion.
As defined previously, let $\mathbf{X}_{(i,j)} : n^c \times 2$ represent the cleaned bivariate IOC data set for process variables i and j for $i < j = 1, \ldots, p$. **After modelling this data using kernel PCA, it is suggested that the resulting fault identification statistics be used as distance measures in the proposed diagnostic methodology**. That is, three possible distance measures are identified. The first distance measure is the T^2-statistic,

$$T^2(\mathbf{x}^*_{(i,j)}) = \bar{\phi}(\mathbf{x}^*_{(i,j)})^T \mathbf{V}_{(i,j)} \mathbf{\Lambda}^{-1}_{(i,j)} \mathbf{V}^T_{(i,j)} \bar{\phi}(\mathbf{x}^*_{(i,j)}) \tag{5.6}$$

$$T^2(\mathbf{x}^*_{(i,j)}) = (\mathbf{t}^\phi(\mathbf{x}^*_{(i,j)}))^T \mathbf{\Lambda}^{-1}_{(i,j)} \mathbf{t}^\phi(\mathbf{x}^*_{(i,j)}), \tag{5.7}$$

where

$$\mathbf{t}^{\phi}(\mathbf{x}^{*}_{(i,j)})_{h^{*}\times 1} = \frac{1}{n^{c}-1}\mathbf{\Lambda}^{-1}_{(i,j)}\tilde{\mathbf{W}}^{T}_{(i,j)}\begin{bmatrix} k(\mathbf{x}_{(i,j)_1},\mathbf{x}^{*}_{(i,j)}) \\ k(\mathbf{x}_{(i,j)_2},\mathbf{x}^{*}_{(i,j)}) \\ \vdots \\ k(\mathbf{x}_{(i,j)_{n^c}},\mathbf{x}^{*}_{(i,j)}) \end{bmatrix} \tag{5.8}$$

is the kernel PCA score representation of $\mathbf{x}^{*}_{(i,j)}$. The matrices $\mathbf{\Lambda}_{(i,j)} : h^{*}\times h^{*}$, $\mathbf{V}_{(i,j)} : h \times h^{*}$ and $\tilde{\mathbf{W}}_{(i,j)} : n^{c} \times h^{*}$ are those which define the h^{*}-dimensional kernel PCA performed on $\mathbf{X}_{(i,j)}$ and are based on the kernel function k. As discussed in Chapter 4, the kernel function is assumed to have been appropriately centered.
Secondly, the SPE-statistic can be utilized as a candidate distance measure,

$$SPE(\mathbf{x}^{*}_{(i,j)}) = k(\mathbf{x}^{*}_{(i,j)},\mathbf{x}^{*}_{(i,j)}) - (\mathbf{t}^{\phi}(\mathbf{x}^{*}_{(i,j)}))^{T}(\mathbf{t}^{\phi}(\mathbf{x}^{*}_{(i,j)})), \tag{5.9}$$

and a third possible measure that is identified is the *Combined*-statistic,

$$\varphi_{(i,j)}(\mathbf{x}^{*}_{(i,j)}) = \frac{T^{2}(\mathbf{x}^{*}_{(i,j)})}{\tau^{2}_{(i,j)}} + \frac{SPE(\mathbf{x}^{*}_{(i,j)})}{\delta^{2}_{(i,j)}}, \tag{5.10}$$

with $\tau^{2}_{(i,j)}$ and $\delta^{2}_{(i,j)}$ the theoretical control limits of $T^{2}_{(i,j)}$ and $SPE_{(i,j)}$ respectively.
Note that by employing this method $\binom{p}{2}$ different kernel PCA models are fitted. Each model is based on a different set of reference points $\mathbf{X}_{(i,j)}$ and can be fitted with different kernel functions. The distances calculated of a new observation to the point clouds are therefore not comparable between the different pairs of variables. In order to have distances that are comparable it is suggested that the selected kernel PCA distance measure be evaluated on a fine grid of points that cover the combined range of variables i and j. The computed distances for the $(i,j)^{th}$ grid points can then be used to identify the smallest and largest distances that are possible. Let f^{kpca} denote the selected kernel PCA distance measure and identify the minimum and maximum distance by

$$f^{kpca}_{min_{(i,j)}} \quad \text{and} \quad f^{kpca}_{max_{(i,j)}}, \tag{5.11}$$

for the $(i,j)^{th}$ pair. The distance evaluation of a new observation, $f^{kpca}(\mathbf{x}^{*}_{(i,j)})$, can now be scaled to a value in the interval $[0,1]$ as follows

$$\frac{f^{kpca}(\mathbf{x}^*_{(i,j)}) - f^{kpca}_{min_{(i,j)}}}{f^{kpca}_{max_{(i,j)}} - f^{kpca}_{min_{(i,j)}}}. \tag{5.12}$$

All of the distances calculated of the new sample to the $(i,j)^{th}$ point cloud is therefore comparable after applying this scaling.

5.4.2 Support Vector Machine Discriminant Analysis

In this section it will be motivated how Support Vector Machine (SVM) discriminant analysis can be used to define a distance measure suitable for the above fault diagnostic methodology.

A short summary, based on the discussion in Hastie et al. (2001), will now be given to explain the main ideas behind SVM for classification. As previously defined, let the row vectors of $\mathbf{X}_{(i,j)} : n^c \times 2$ be the IOC data points and let the rows of $\mathbf{X}^r_{(i,j)} : n^r \times 2$ be the complement bivariate vectors in the range of variables i and j that fall outside the region occupied by the IOC samples. A multivariate density is fitted to the data set and all variable combinations above a cut-off such as the 10^{-8}th quantile is deemed to be IOC data points. Throughout this section it will be assumed that $i, j \in [1, 2, \ldots, p]$ and that $i < j$. The k^{th} element of the group indicator variable $\mathbf{y} : (n^c + n^r) \times 1$ can now be defined as

$$y_k \quad = \quad \begin{cases} +1, & \text{if } \mathbf{x}_k \in \mathbf{X}_{(i,j)} \\ -1, & \text{otherwise} \end{cases} \tag{5.13}$$

for $k = 1, \ldots, (n^c + n^r)$. Each of the observations in the rows of $\mathbf{X}_{(i,j)}$ and $\mathbf{X}^r_{(i,j)}$ can therefore be associated with a value in $\mathbf{y}$ i.e., the following set of data pairs can be constructed: $(y_1, \mathbf{x}_1)$, $(y_2, \mathbf{x}_2), \ldots, (y_{(n^c+n^r)}, \mathbf{x}_{(n^c+n^r)})$. This data set of paired observations, often referred to as the training data, can now be analysed using SVM discriminant analysis to identify a function that can be used to model the difference in process conditions that led to possitive and negative y values. Specifically, the objective of SVM is to find a hyperplane

$$f(\mathbf{x}) \quad = \quad \beta_0 + \mathbf{x}^T \boldsymbol{\beta} = 0, \tag{5.14}$$

with $\boldsymbol{\beta} : 2 \times 1 = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix}^T$ and $\beta_0, \beta_1, \beta_2 \in \mathbb{R}$, that can be implemented as a classification rule where an observation $\mathbf{x}$ is identified as being part of the IOC samples if $f(\mathbf{x}) > 0$ and different from IOC samples if $f(\mathbf{x}) < 0$. In order to obtain a decision boundary that generalizes well for new data, the hyperplane is selected such that it has the largest margin between the in-control and out-of-control data points. The Point-plane or signed distance of the closest observation from either of the two data sets to the hyperplane defines the margin. The Point-plane distance of each training sample $\mathbf{x}_k$ for $k = 1, \ldots, (n^c + n^r)$ equals

$$\frac{f(\mathbf{x}_k)}{\|f'(\mathbf{x}_k)\|} = \frac{\beta_0 + \mathbf{x}_k^T \boldsymbol{\beta}}{\|\boldsymbol{\beta}\|} \tag{5.15}$$

$$= \frac{\beta_0 + \mathbf{x}_k^T \boldsymbol{\beta}}{\sqrt{\beta_1^2 + \beta_2^2}}. \tag{5.16}$$

If it can be assumed for the moment that the two sets of data are linearly separable, then if C is the distance from the plane to the nearest point, the search for the hyperplane with maximum margin amounts to finding a solution for the constrained optimization problem

$$\max_{\beta_0, \boldsymbol{\beta}} \quad C \tag{5.17a}$$

$$\text{subject to} \quad \frac{y_k}{\|\boldsymbol{\beta}\|}(\beta_0 + \mathbf{x}_k^T \boldsymbol{\beta}) \geqslant C, \tag{5.17b}$$

where the constraint (5.17b) combine two conditions in one. With the first constraint being that it is required that the model correctly classify the training data,

$$y_k(\beta_0 + \mathbf{x}_k^T \boldsymbol{\beta}) > 0, \tag{5.18}$$

and secondly it is required that all of the signed distances must be at least C units from the hyperplane,

$$\frac{1}{\|\boldsymbol{\beta}\|}(\beta_0 + \mathbf{x}_k^T \boldsymbol{\beta}) \geqslant C. \tag{5.19}$$

Notice that the solution set of weights $\beta_0, \beta_1, \beta_2$ that solves this maximization problem can be multiplied by a constant yielding another possible solution i.e., the hyperplane is not uniquely defined. In order to ensure that the solution is unique it will be required that

the hyperplane should be canonical. The hyperplane is said to be canonical with respect to the training input data $\mathbf{x}_1, \ldots, \mathbf{x}_{(n^c+n^r)}$ if it is scaled such that

$$\min_{k=1,\ldots,(n^c+n^r)} \quad |\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}| = 1. \tag{5.20}$$

Hence if this requirement is enforced the distance of the closest point to the hyperplane simplifies to

$$C \;=\; \frac{1}{\|\boldsymbol{\beta}\|}, \tag{5.21}$$

with the constraint (5.17b) becoming $y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) \geqslant 1$. That is, the optimization problem can be rephrased as the search for a set of weights $\boldsymbol{\beta}$ with minimal length such that $y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) \geqslant 1$.
For mathematical convenience, the optimization problem is written in the following equivalent form

$$\begin{aligned}
&\min_{\beta_0,\boldsymbol{\beta}} && \frac{1}{2}\|\boldsymbol{\beta}\|^2 && \text{(5.22a)}\\
&\text{subject to} && y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) \geqslant 1, && \text{(5.22b)}\\
&\text{for all} && k = 1, \ldots, (n^c + n^r). && \text{(5.22c)}
\end{aligned}$$

It has been assumed up till now that the two sets of points are linearly separable. Consider now the case where there is overlap between the groups and that it is not possible to find a hyperplane that perfectly separate the groups. This implies that there will be observations that will be wrongly identified as being out-of-control when it is in-control and the other way around i.e., in-control when it is out-of-control. The SVM algorithm accounts for this by the introduction of so-called slack variables $\boldsymbol{\xi} = \left[\xi_1, \ldots, \xi_{(n^c+n^r)}\right]^T$. These slack variables are introduced to allow some of the observations to be inside or on the opposite side of the margin. This is achieved by making the constraint (5.22b) less stringent

$$y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) \geqslant 1 - \xi_k, \tag{5.23}$$

for $k = 1, \ldots, (n^c + n^r)$, $\xi_k \geqslant 0$ and $\sum_{k=1}^{n^c+n^r} \xi_k \leqslant \tau$. The sum constraint of size τ put on the slack variables is introduced to avoid trivial solutions. The value of which needs to be pre-specified. The SVM optimization problem can now be restated for the overlapping

case as

$$\begin{aligned} \min_{\beta_0,\boldsymbol{\beta}} \quad & \frac{1}{2}\|\boldsymbol{\beta}\|^2 + \tau \sum_{k=1}^{n^c+n^r} \xi_k && (5.24a) \\ \text{subject to} \quad & \xi_k \geqslant 0, y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) \geqslant 1 - \xi_k, && (5.24b) \\ \text{for all} \quad & k = 1, \ldots, (n^c + n^r). && (5.24c) \end{aligned}$$

This final form of the SVM minimization problem is in the shape of a convex optimization problem. Lagrange (or Karush-Kuhn-Tucker) multipliers can therefore be used to obtain a quadratic programming solution for this problem. The Lagrangian expression that needs to be minimized with respect to $\beta_0, \boldsymbol{\beta}$ and $\boldsymbol{\xi}$ equals

$$\begin{aligned} \mathcal{L}(\beta_0, \boldsymbol{\beta}, \boldsymbol{\xi}) &= \frac{1}{2}\|\boldsymbol{\beta}\|^2 + \tau \sum_{k=1}^{n^c+n^r} \xi_k - \sum_{k=1}^{n^c+n^r} \alpha_k \Big(y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) - (1 - \xi_k) \Big) \\ &\quad - \sum_{k=1}^{n^c+n^r} \alpha_k^0 \xi_k, \end{aligned} \tag{5.25}$$

with Karush-Kuhn-Tucker conditions

$$\begin{aligned} \alpha_k \Big(y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) - (1 - \xi_k) \Big) &= 0, && (5.26) \\ y_k(\beta_0 + \mathbf{x}_k^T\boldsymbol{\beta}) - (1 - \xi_k) &\geqslant 0, && (5.27) \\ \text{and } \alpha^0 \xi_k &= 0, && (5.28) \end{aligned}$$

for $k = 1, \ldots, (n^c + n^r)$. Setting each of the partial derivatives $\frac{\partial \mathcal{L}(\beta_0,\boldsymbol{\beta},\boldsymbol{\xi})}{\partial \beta_0}$, $\frac{\partial \mathcal{L}(\beta_0,\boldsymbol{\beta},\boldsymbol{\xi})}{\partial \boldsymbol{\beta}}$ and $\frac{\partial \mathcal{L}(\beta_0,\boldsymbol{\beta},\boldsymbol{\xi})}{\partial \xi_k}$ equal to zero result in the following set of equations

$$\begin{aligned} 0 &= \sum_{k=1}^{n^c+n^r} \alpha_k y_k, && (5.29) \\ \boldsymbol{\beta} &= \sum_{k=1}^{n^c+n^r} \alpha_k y_k \mathbf{x}_k, && (5.30) \\ \text{and } \alpha_k &= \tau - \alpha_k^0, && (5.31) \end{aligned}$$

for $\alpha_k, \alpha_k^0 \geqslant 0$. Substituting these results in (5.25) gives the following quadratic programming optimization problem:

$$\max_{\boldsymbol{\alpha}} \mathcal{L}(\boldsymbol{\alpha}) = \sum_{k=1}^{n^c+n^r} \alpha_k - \sum_{k=1}^{n^c+n^r} \sum_{k'=1}^{n^c+n^r} \alpha_k \alpha_{k'} y_k y_{k'} \mathbf{x}_k^T \mathbf{x}_{k'} \tag{5.32a}$$

$$\text{subject to} \quad 0 \leqslant \alpha_k \leqslant \tau \text{ for } k = 1, \ldots, (n^c + n^r) \tag{5.32b}$$

$$\text{and} \quad \sum_{k=1}^{n^c+n^r} \alpha_k y_k = 0. \tag{5.32c}$$

Writing the objective function of (5.25) into the form of the maximization problem (5.32a) provides an easier optimization problem that can be solved using well known techniques. The conditions at which this problem attains its maximum $\hat{\boldsymbol{\alpha}} = \left[\hat{\alpha}_1, \ldots, \hat{\alpha}_{n^c+n^r}\right]$ can now be used to obtain the corresponding optimal values of $\beta_0, \boldsymbol{\beta}$ and $\boldsymbol{\xi}$. The optimal settings for $\boldsymbol{\beta}$ can be obtained from (5.30) as

$$\hat{\boldsymbol{\beta}} = \sum_{k=1}^{n^c+n^r} \hat{\alpha}_k y_k \mathbf{x}_k. \tag{5.33}$$

Notice that only a subset of the $\hat{\boldsymbol{\alpha}}$ weights will be non-zero. These are the alpha values for which the conditions of (5.26) and (5.27) are simultaneously met. The observations $\mathbf{x}_k$ which correspond to the non-zero α_k values are referred to as the support vectors. Some of these supporting vectors will define the margin i.e., when $\hat{\xi}_k = 0$ and $0 < \alpha_k < \tau$. The other support vectors will lie inside the margin, to the side of the hyperplane, when $\hat{\xi}_k > 0$, and correspond with $\hat{\alpha}_k = \tau$.

An estimate for the intercept, β_0, can be derived when considering (5.26) for cases when $\hat{\alpha}_k > 0$ and $\hat{\xi}_k = 0$ i.e., when the support vectors lie on the margin. Multiple support vectors can lie on the margin therefore it is advised that an average be taken of the possible $\hat{\beta}_0$ solutions.

For a given value of τ the optimal hyperplane parameters $\hat{\beta}_0$ and $\hat{\boldsymbol{\beta}}$ can be calculated. This fitted hyperplane

$$\hat{f}(\mathbf{x}) = \hat{\beta}_0 + \mathbf{x}^T \hat{\boldsymbol{\beta}}, \tag{5.34}$$

can therefore be used to determine if a new point $\mathbf{x}^*$ is close to the IOC cloud of points ($\hat{f}(\mathbf{x}^*) > 0$) or whether it can be considered out-of-control ($\hat{f}(\mathbf{x}^*) < 0$).

The application of the SVM model in its current linear form is limited, given the nonlinear characteristics of the data considered in kernel PCA based statistical process monitoring. Similar to the discussion in Chapter 4, this problem can be addressed by projecting the input training data into a high dimensional feature space. Let $\phi(\mathbf{x}_k) : h \times 1$ with $2 < h \leqslant \infty$

represent the high, possibly infinite, dimensional feature vector versions of the input training data. The training data pairs therefore become $(y_1, \phi(\mathbf{x}_1)), \ldots, (y_{n^c+n^r}, \phi(\mathbf{x}_{n^c+n^r}))$. The SVM algorithm discussed for finding the separating hyperplane can now directly be applied to the feature vectors. This implies that the objective function of the SVM Lagrange maximization problem becomes

$$\mathcal{L}(\boldsymbol{\alpha}) = \sum_{k=1}^{n^c+n^r} \alpha_k - \sum_{k=1}^{n^c+n^r} \sum_{k'=1}^{n+n^r} \alpha_k \alpha_{k'} y_k y_{k'} \phi(\mathbf{x})_k^T \phi(\mathbf{x}_{k'}). \tag{5.35}$$

Using the solution to this redefined problem, the estimate of the hyperplane weights is

$$\hat{\boldsymbol{\beta}} = \sum_{k=1}^{n^c+n^r} \hat{\alpha}_k y_k \phi(\mathbf{x}_k). \tag{5.36}$$

The same principal of using the average information of the margin points can be used to estimate $\hat{\beta}_0$. The SVM decision boundary is therefore defined by

$$\begin{aligned} \hat{f}(\mathbf{x}) &= \hat{\beta}_0 + \phi(\mathbf{x})^T \hat{\boldsymbol{\beta}} && (5.37) \\ &= \hat{\beta}_0 + \sum_{k=1}^{n^c+n^r} \hat{\alpha}_k y_k \phi(\mathbf{x})^T \phi(\mathbf{x}_k). && (5.38) \end{aligned}$$

It can be observed that both the feature space SVM problem definition (5.35) and the solution (5.38) are functions of the inner product of the feature vectors. The kernel trick (see Chapter 4), can therefore be applied to approximate these inner products using kernel functions. Using kernel functions, the SVM formulation in feature space can therefore be approximated as

$$\mathcal{L}(\boldsymbol{\alpha}) = \sum_{k=1}^{n^c+n^r} \alpha_k - \sum_{k=1}^{n^c+n^r} \sum_{k'=1}^{n^c+n^r} \alpha_k \alpha_{k'} y_k y_{k'} k(\mathbf{x}_k, \mathbf{x}_{k'}) \tag{5.39}$$

and

$$\hat{f}(\mathbf{x}) = \hat{\beta}_0 + \sum_{k=1}^{n^c+n^r} \hat{\alpha}_k y_k k(\mathbf{x}, \mathbf{x}_k). \tag{5.40}$$

Note that by using kernel functions, additional hyperparameters, in addition to τ, are introduced that need to be pre-specified before solving the SVM problem. In practice a grid of possible values are considered for these parameters. Multiple models are fitted, for each of the parameter combinations, using cross-validation. The model with the lowest cross-validated misclassification rate is then selected as the final model.
As mentioned earlier, the final model can now be used to classify a new observation as being similar to or different from the IOC samples by considering the sign of $\hat{f}(\mathbf{x})$. It would however be more informative to have the model prediction in the form of a posterior probability

$$P(\mathbf{x} \in IOC|\hat{f}(\mathbf{x})), \tag{5.41}$$

which will add a degree of confidence to the final prediction. It is motivated by Platt (1999) that these posterior probabilities should be modelled directly using the observed distribution of the model predictions made on the training data. The sigmoidal function of the form

$$P(y_k = 1|\hat{f}(\mathbf{x}_k)) \quad = \quad \frac{1}{1 + \exp(A_1\hat{f}(\mathbf{x}_k) + A_2)}, \tag{5.42}$$

with A_1 and A_2 pre-selected constants, is identified as a good approximation of the distribution of the observed predictions $\hat{f}(\mathbf{x}_k)$ for $k = 1, \ldots, (n^c + n^r)$. The optimal values of the constants A_1 and A_2 can be estimated from training data. Consider the training data pairs $(y_1^*, \hat{f}(\mathbf{x}_1)), \ldots, (y_{(n^c+n^r)}^*, \hat{f}(\mathbf{x}_{(n^c+n^r)}))$ where the response variable has been rescaled to be binary i.e., $y_k^* \in [0, 1]$, as follows:

$$y_k^* \quad = \quad \frac{y_k + 1}{2}. \tag{5.43}$$

The negative log likelihood of the training data can now be obtained as

$$-\sum_{k=1}^{n^c+n^r} \Big[y_k^* \log(p_k) + (1 - y_k^*) \log(1 - p_k) \Big], \tag{5.44}$$

with

$$p_k = \frac{1}{1+\exp(A_1\hat{f}(\mathbf{x}_k)+A_2)}. \tag{5.45}$$

Traditional optimization techniques can now be used to find the constants $\hat{A}_1$ and $\hat{A}_2$ that define the minimum of this negative log likelihood function. A note of caution against the over-fitting of these posterior probabilities is issued by Platt (1999), and it is therefore suggested that cross-validation be used. More specifically, Platt (1999) proposes the use of three fold cross-validation. The training data set is randomly divided into three parts of approximately equal size. The SVM model is then fitted in turn on two of the data sets and tested on the third. Each iteration therefore produce a set of predictions that can be pooled into a combined set which is used to fit the posterior probability distribution.
To summarize, let the following define the SVM classification model fitted for $(i,j)^{th}$ process variable pair. The model fitted using the two data sets $\mathbf{X}_{(i,j)}$ and $\mathbf{X}^r_{(i,j)}$ has the following form

$$\hat{f}(\mathbf{x}) = \hat{\beta}_0 + \sum_{k=1}^{n^c+n^r} \hat{\alpha}_k y_k k(\mathbf{x},\mathbf{x}_k). \tag{5.46}$$

The model predictions as class probabilities

$$P(\mathbf{x} \in IOC|\hat{f}(\mathbf{x})) = \frac{1}{1+\exp(\hat{A}_1\hat{f}(\mathbf{x})+\hat{A}_2)}, \tag{5.47}$$

and

$$P(\mathbf{x} \notin IOC|\hat{f}(\mathbf{x})) = 1-\frac{1}{1+\exp(\hat{A}_1\hat{f}(\mathbf{x})+\hat{A}_2)}. \tag{5.48}$$

The calculated probability of a new sample not being part of the IOC distribution i.e., (5.48), can be used as a distance measure in the new fault diagnosis methodology. That is, the distance matrix of a new observation $\mathbf{x}^*$: $p \times 1$ will have $(i,j)^{th}$ off-diagonal element equal to

$$d_{ij} = P(\mathbf{x}^*_{(i,j)} \notin IOC|\hat{f}(\mathbf{x}^*_{(i,j)})). \tag{5.49}$$

This distance measure naturally adheres to the properties required of a distance measure in the methodology developed. Specifically, it is always positive $d_{i,j} \geqslant 0$ and will be small when the sample is 'close' to the IOC cloud, and big when the sample is 'remote' from the IOC region. The methodology used to calculate these distances can be consistently applied across process variable pairs and since it is a probability, implies that the distances are comparable for the different paired evaluations.

Support vector machine classification therefore provides a dissimilarity measure that ideally fits the new fault diagnosis methodology. However, this discrimination technique is supervised. In the training process it is required that samples representing 'close' and 'remote' observations are included. The statistical process monitoring fault detection step, based on kernel PCA, is unsupervised i.e. information is only available on the in-control samples. Therefore, in order to use the SVM technique, it is required that information be provided on the complementary data set $\mathbf{X}^r_{(i,j)}$.

It is suggested that the point vectors defining the 'empty white spaces' in the combined range of variables i and j be used to define $\mathbf{X}^r_{(i,j)}$. This is a sensible choice since in multivariate statistical process monitoring out-of-control examples are univariately in-control. The 'remote' points can therefore be limited to lie inside the range of the data. One possible way to identify the points occupying the white spaces is to make use of density estimation. An estimate of the bivariate density describing the IOC cloud of points can be obtained with kernel density estimation (Duong, 2007). A fine grid of points can then be constructed to cover all possible bivariate vectors in the range of variables i and j. Each of these grid points can then be evaluated using the IOC kernel density estimate. All of the points that have a very small density can then be used to define the out-of-control region. In Figures 5.3 to 5.6 this proposed method of estimating the complementary data set is illustrated. The data used in this illustration are again taken from the simulated data set created in Chapter 4. The subset of the IOC cloud of points, for variable 1 and 2, determined by the kernel PCA monitoring process to be in-control, using the T^2 statistic at the 95% confidence level, is plotted in Figure 5.3. In Figure 5.4 the bivariate kernel density estimate of these points is plotted. A grid of points were created to take on all of the possible combined values of variable 1 and 2. This grid points were then evaluated using the IOC density estimate. Figure 5.5 displays the grid points in grey which were identified to have low density in the IOC region. The two data sets are simultaneously displayed in Figure 5.6.

A novel strategy is therefore introduced of how to utilise SVM classification in order to specify a distance measure that is suitable for the new fault diagnosis methodology.

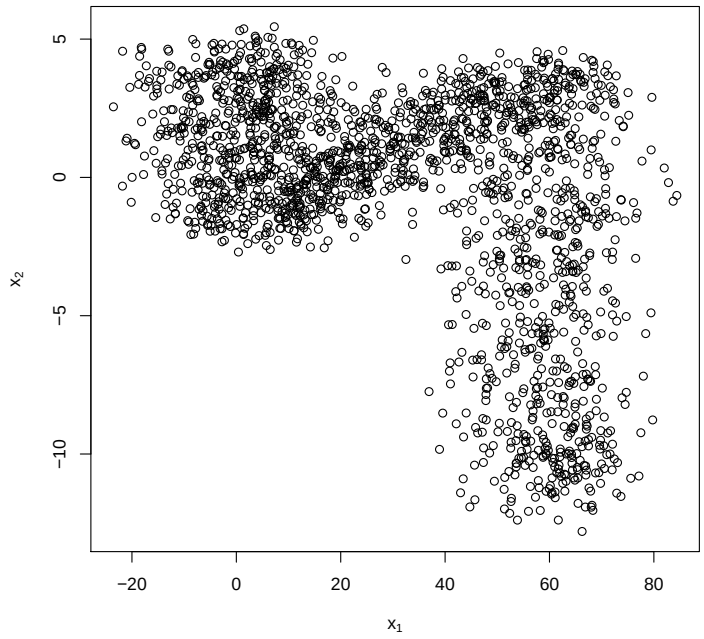


Figure 5.3: Scatter plot of variable 1 versus variable 2. The data are extracted from the four dimensional vectors that were considered to be in-control from the kernel PCA model.

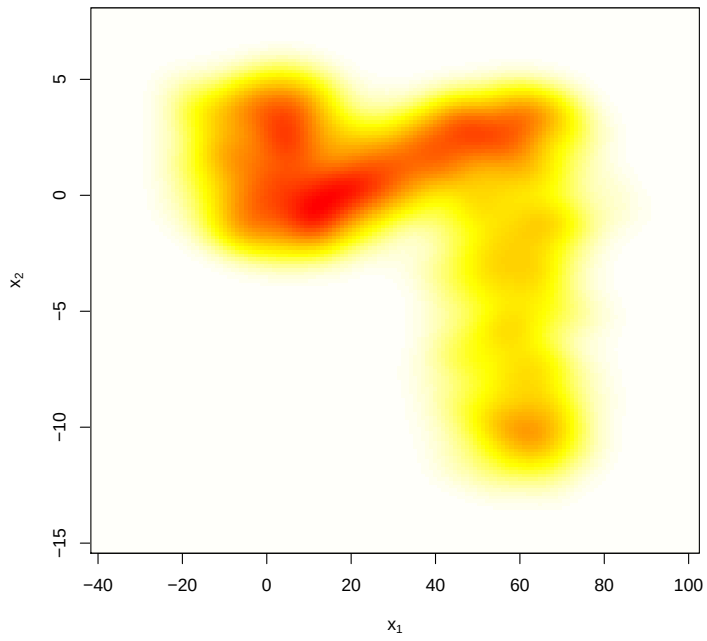


Figure 5.4: Bivariate density estimate of variable 1 and variable 2.

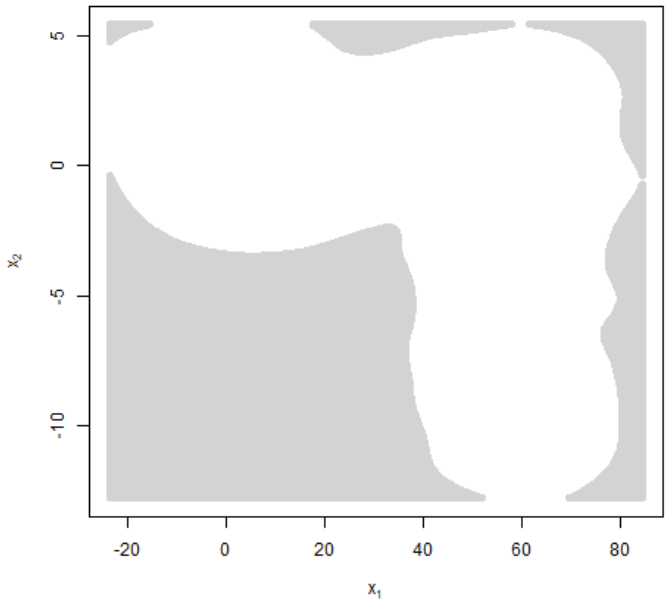


Figure 5.5: Grey points in the range of variable 1 and 2 determined by the density estimate to have low density.

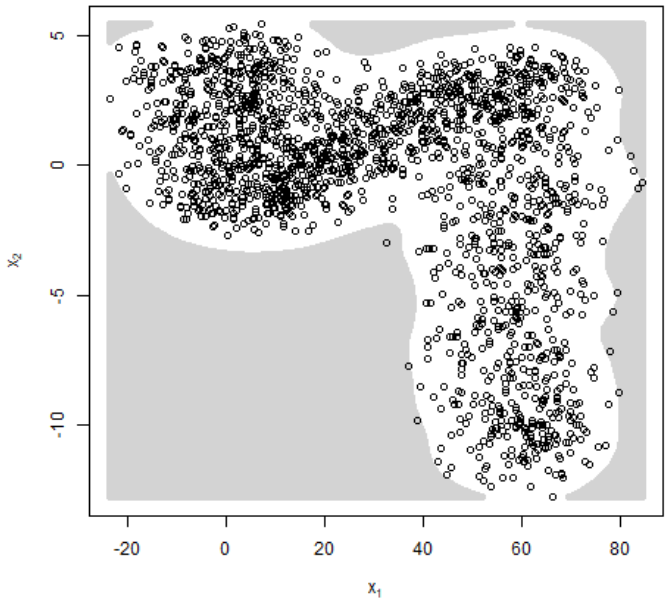


Figure 5.6: Grey points in the range of variable 1 and 2 determined by the density estimate to have low density. In control circular data points included.

5.4.3 One-Class Support Vector Machine Discriminant Analysis

The final potential distance measure, for use in the pairwise variable importance methodology, is developed using the one class classification (OCC) technique called One-class support vector machine discriminant analysis (OCSVM).

In OCC problems the objective is to learn the characteristics of a specific class using historical observations from that class. This collection of historical observations is referred to as the target or reference set. The set of rules identified to summarise the properties of the class is then applied to determine if new observations belong to that class or not. Schölkopf et al. (2001) introduced an extended version of two class SVM classification as an approach to model OCC problems. The strategy employed by Schölkopf et al. (2001) is to project the target data vectors into a feature space and to maximally separate the resulting vectors from the origin using a hyperplane. The target data lie on the side of the hyperplane that is opposite to the side containing the origin. Similar to two class SVM, the hyperplane is used to define a function f that predicts $+1$ if an observation is part of the target class and -1 if it is not.

The objective of identifying if a new observation is 'close' or 'remote' from the corresponding target point cloud for each of the variable pairs is therefore in-line with the solution provided by OCSVM. Using previous notation, let the rows, $\mathbf{x}^k_{(i,j)}$ for $k = 1, \ldots, n^c$, of $\mathbf{X}_{(i,j)} : n^c \times 2$ represent the bivariate IOC data set of the process variable pair (i, j) for $i < j = 1, \ldots, p$. Furthermore assume that there exists a feature space mapping function ϕ such that the inner product of the vectors in feature space can be approximated by a kernel function k i.e., for any target vectors $\mathbf{x}^*_{(i,j)}$ and $\mathbf{x}^+_{(i,j)}$

$$\phi(\mathbf{x}^*_{(i,j)})^T \phi(\mathbf{x}^+_{(i,j)}) \approx k(\mathbf{x}^*_{(i,j)}, \mathbf{x}^+_{(i,j)}). \tag{5.50}$$

Using most of the formulations established in Section 5.4.1 the realisation of the maximal separating hyperplane function,

$$f(\mathbf{x}) \quad = \quad \phi(\mathbf{x})^T \boldsymbol{\beta} - \beta_0, \tag{5.51}$$

is achieved by solving the following quadratic programming problem,

$$\min_{\beta_0,\boldsymbol{\beta},\boldsymbol{\xi}} \quad \frac{1}{2}\|\boldsymbol{\beta}\|^2 + \frac{1}{n^c\nu}\sum_{k=1}^{n^c}\xi_k \tag{5.52a}$$

$$\text{subject to} \quad \xi_k \geqslant 0, \phi(\mathbf{x})_k^T\boldsymbol{\beta} - \beta_0 \geqslant -\xi_k, \tag{5.52b}$$

$$\text{for all} \quad k = 1, \ldots, n^c, \tag{5.52c}$$

with ν a regularisation constant that needs to be specified upfront. Similar to the methods utilised in SVM, this optimisation problem can be reformulated using Lagrange multipliers yielding the following minimization problem

$$\min_{\boldsymbol{\alpha}} \quad \frac{1}{2}\sum_{k=1}^{n^c}\sum_{k'=1}^{n^c}\alpha_k\alpha_{k'}k(\mathbf{x}_k, \mathbf{x}_{k'}) \tag{5.53a}$$

$$\text{subject to} \quad \sum_{k}^{n^c}\alpha_k = 1 \text{ and } 0 \leqslant \alpha_k \leqslant \frac{1}{\nu n^c}, \tag{5.53b}$$

where the kernel trick has been leveraged to approximate the feature vector inner products using an appropriate kernel function. The optimal settings $\hat{\boldsymbol{\alpha}}$ for this quadratic programming problem are then used, in a similar fashion to SVM classification, to define the hyperplane. Therefore with $\hat{f}$ representing the fitted hyperplane for the $(i,j)^{th}$ variable pair, the final functional evaluation of a new observation $\mathbf{x}^*_{(i,j)}$ is calculated as

$$\hat{f}(\mathbf{x}^*_{(i,j)}) \quad = \quad \sum_{k=1}^{n^c}\hat{\alpha}_k k(\mathbf{x}^*_{(i,j)}, \mathbf{x}_k) - \hat{\beta}_0. \tag{5.54}$$

Note that again only a subset of the elements in $\hat{\boldsymbol{\alpha}}$ will be non-zero, with the corresponding target vectors representing the support vectors. The estimation of the offset β_0 is calculated as

$$\hat{\beta}_0 \quad = \quad \sum_{k=1}^{n^c}\hat{\alpha}_k k(\mathbf{x}_l, \mathbf{x}_k), \tag{5.55}$$

with $\mathbf{x}_l$ any target sample that has a corresponding $\hat{\alpha}_l$ estimate which adheres to the constraint $0 < \hat{\alpha}_l < \frac{1}{\nu n^c}$. This maximally fitted hyperplane function can now be used to determine if a new observation is local or alien to the target data. An observation will be classified as being 'close' to the specific IOC data cloud if $\hat{f}(\mathbf{x}^*_{(i,j)}) > 0$ and 'remote' if $\hat{f}(\mathbf{x}^*_{(i,j)}) < 0$.

In Section 5.4.1 the SVM derived distance measure suitable for the pairwise methodology was naturally identified as the posterior probability

$$P(\mathbf{x}^*_{(i,j)} \notin IOC | \hat{f}(\mathbf{x}^*_{(i,j)})). \tag{5.56}$$

The calculation of the posterior probability is attainable in SVM classification since it is modelling well defined classes. However, in the OCC setting the class representing the complement of the target class is not established. It is required that $\hat{f}$ be evaluated over samples from all classes in order to estimate the posterior probability. In most cases this proves infeasible, since the class opposite to the target set is of infinite dimensionality. It is still possible in this problem setting to obtain a good representation of both classes, due to the constraint imposed by multivariate process monitoring, which limits the data for consideration to the univariate ranges of the process variables. Hence, posterior probabilities can be obtained from the model evaluations made on a fine grid of points in the range of the i^{th} and j^{th} process variables. Let the fine grid of N points, for the $(i,j)^{th}$ variable combination be represented by the vectors $\mathbf{x}^1_{(i,j)}, \ldots, \mathbf{x}^N_{(i,j)}$. Note that a large enough N should be selected in order to capture most of the possible values of variables i and j. For each of these grid points, the OCSVM model evaluation is calculated as $\hat{f}(\mathbf{x}^k_{(i,j)})$ for $k = 1, \ldots, N$. The associated class membership prediction $\hat{y}_k$ can then be determined as

$$\hat{y}_k = \begin{cases} +1, & \text{if } \hat{f}(\mathbf{x}^k_{(i,j)}) \geqslant 0 \\ -1, & \text{otherwise,} \end{cases} \tag{5.57}$$

for $k = 1, \ldots, N$. The following pairs of grid data evaluations can therefore be created $(\hat{y}^*_1, \hat{f}(\mathbf{x}^1_{(i,j)})), \ldots, (\hat{y}^*_k, \hat{f}(\mathbf{x}^N_{(i,j)}))$ with

$$\hat{y}^*_k \quad = \quad \frac{\hat{y}_k + 1}{2}. \tag{5.58}$$

Therefore, using the ideas of Platt (1999), the negative log likelihood of these data pairs

$$-\sum_{k=1}^{N} \Big[\hat{y}^*_k \log(p_k) + (1 - \hat{y}^*_k) \log(1 - p_k) \Big], \tag{5.59}$$

with

$$p_k = \frac{1}{1+\exp(A_1\hat{f}(\mathbf{x}^k_{(i,j)})+A_2)}, \tag{5.60}$$

can be optimized to identify the settings $\hat{A}_1$ and $\hat{A}_2$. These optimal settings can therefore now be used to specify the posterior complement class membership probability

$$P(\mathbf{x}^*_{(i,j)} \notin IOC|\hat{f}(\mathbf{x}^*_{(i,j)})) = 1 - \frac{1}{1+\exp(\hat{A}_1\hat{f}(\mathbf{x}^*_{(i,j)})+\hat{A}_2)}. \tag{5.61}$$

A complement class membership probability function for OCSVM is developed which is analogous to the one used in the SVM setting. **Therefore, (5.61) is identified as a suitable distance measure for the new pairwise importance methodology**. A novel strategy is therefore introduced of how to utilise OCSVM in order to specify a distance measure that is suitable for the new fault diagnosis methodology.

5.5 Kernel PCA Simulation Study continued (Part 1)

5.5.1 Objective

The fault diagnosis ability of the new pairwise diagnostic methodology for non-linear MSPM will now be demonstrated. The simulation study that was introduced in Chapter 4 will again be considered. First, the methodology based on each of the proposed distance measures will be implemented to diagnose the single fault example that was simulated in Section 4.5. The objective is to demonstrate how the new methodologies can typically be applied.

5.5.2 Kernel PCA

In this section the new fault diagnosis methodology will utilise the kernel PCA related distance measure proposed in Section 5.4.1. Specifically, it will be demonstrated how the distance calculated using the T^2 statistic (5.7) can be used in this methodology. The methodology based on the other statistics i.e., (5.9) and (5.10), will follow a similar work flow. The following steps will be followed:

1. Clean the IOC training data set using the kernel PCA model used in the fault identification step i.e., remove all of the samples that are out-of-control based on one of the fault detection statistics.

2. Fit a kernel PCA model for each of the $\binom{4}{2}$ process variable combinations.

3. Calculate the T^2 statistic value of the new sample.

4. Determine the range of possible T^2 statistic values i.e., evaluate kernel PCA model over a grid of possible values for the two process variables. Here a 10×10 grid was used, but it can be adapted to the specific data set.

5. Scale the T^2 value of the new sample to be in the range $[0, 1]$ using (5.12).

These steps will now be illustrated for the process variable combination (x_1, x_2). The reference data for (x_1, x_2) are displayed using the black points in Figure 5.7. The red dot in Figure 5.7 represents the (x_1, x_2) measurement that was recorded for the simulated faulty sample. A kernel PCA model, using the Gaussian kernel function, was fitted on the reference data of (x_1, x_2). The steps summarised by Algorithm 1, in Chapter 4, was utilised to determine the optimal hyperparameter settings for the model. The number of kernel principal components selected was 10 and the value of the Gaussian kernel parameter i.e., σ equalled 0.1. The T^2 statistic value of the simulated faulty observation was then calculated using the fitted model. In Figure 5.8, the distribution of the T^2 values for the reference data are displayed. Also included in Figure 5.8 is the empirically determined T^2 95% control limit, identified by the vertical broken red line, which equals 27.66. The value of the T^2 statistic for the simulated faulty vector equalled 74.41 and is represented by the vertical blue line in Figure 5.8. It can be concluded from Figure 5.8 that the new observation is considered to be remote from the reference region for (x_1, x_2). A grid is now constructed that contains all possible values that the vector $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ can take on. The kernel PCA model is then used to obtain the T^2 statistic evaluations of these grid vectors. Figure 5.9 depicts the distribution of all possible T^2 values. It was determined that the possible T^2 values are approximately in the range $(2.99, 1234.40)$. This information can now be used to scale the observed T^2 value of the simulated faulty sample to be comparable to the other T^2 distances calculated for the other variable combinations. Therefore, using (5.12) the scaled T^2 value equals

$$\frac{(74.41 - 2.99)}{(1234.40 - 2.99)} = 0.05786311. \tag{5.62}$$

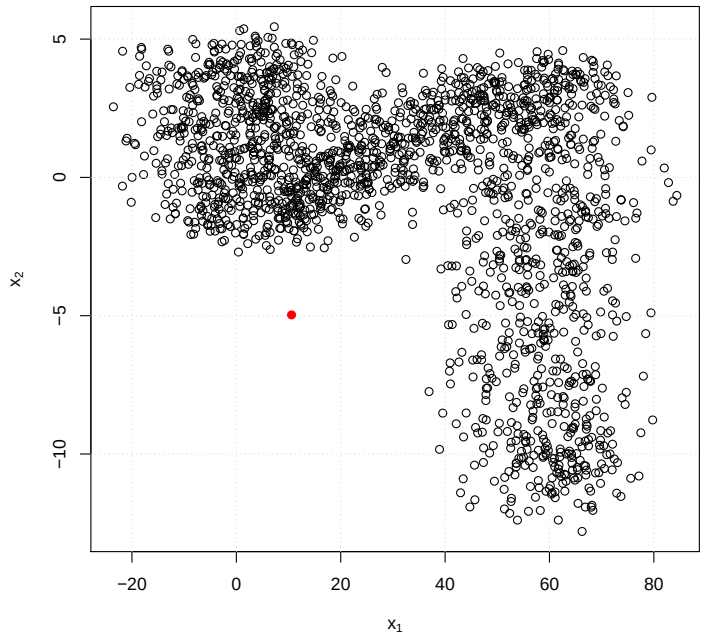


Figure 5.7: Black dots in scatter plot represent the reference data of (x_1, x_2). Red dot represent the readings recorded for the simulated faulty sample.

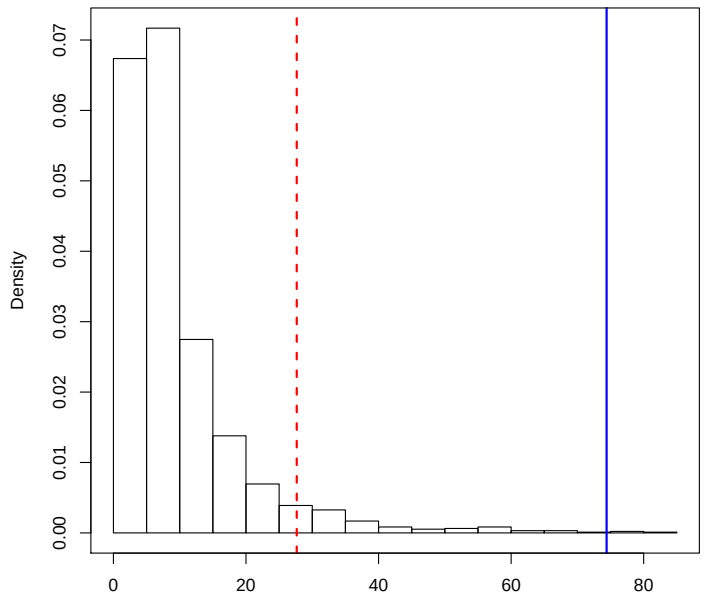


Figure 5.8: Distribution of the kernel PCA T^2 values for the reference data of (x_1, x_2).

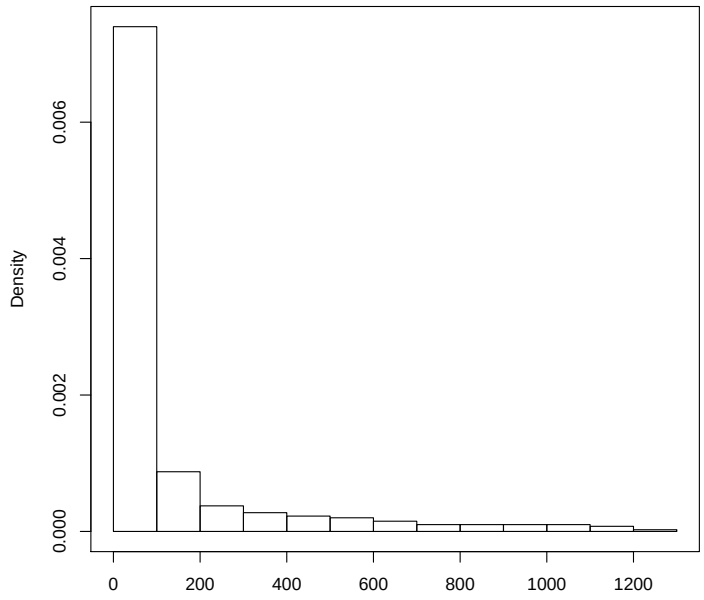


Figure 5.9: Distribution of the kernel PCA T^2 values for the grid values of (x_1, x_2).

The complete set of $\binom{4}{2}$ scaled T^2 distances for the sample under consideration was calculated and are captured in the following matrix

$$\begin{bmatrix} 0.0000 & 0.0579 & 0.0360 & 0.0888 \\ 0.0579 & 0.0000 & 0.0030 & 0.0104 \\ 0.0360 & 0.0030 & 0.0000 & 0.0016 \\ 0.0888 & 0.0104 & 0.0016 & 0.0000 \end{bmatrix}. \tag{5.63}$$

It can, for example, be observed from (5.63) that the largest distance and therefore pairwise contribution is reported by (x_1, x_4). The scatter plot for this variable combination is displayed by Figure 5.10. In this figure it can visually be confirmed that the new measurement is remote from the IOC region. The calculated T^2 distance is therefore sensible. Figure 5.11 displays the magnitude of the pairwise contributions in a bar plot.

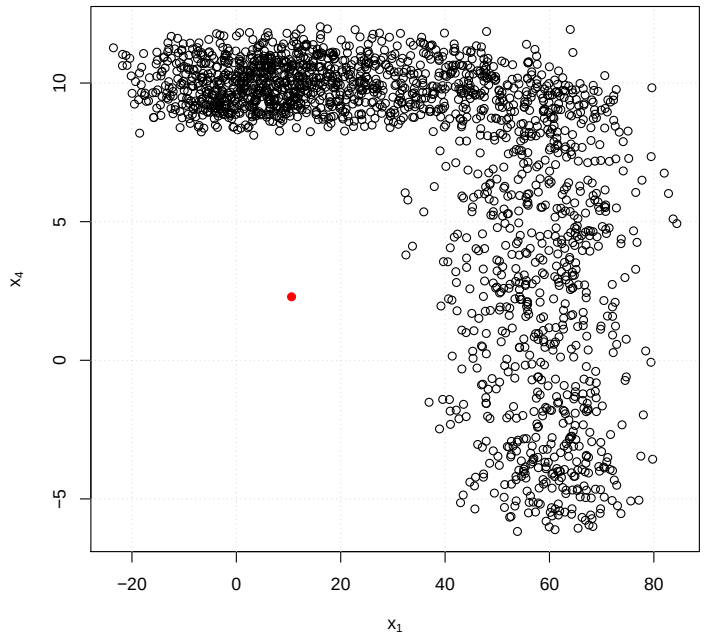


Figure 5.10: Black dots in scatter plot represent the reference data of (x_1, x_4). Red dot represents the measurement recorded for the simulated faulty sample.

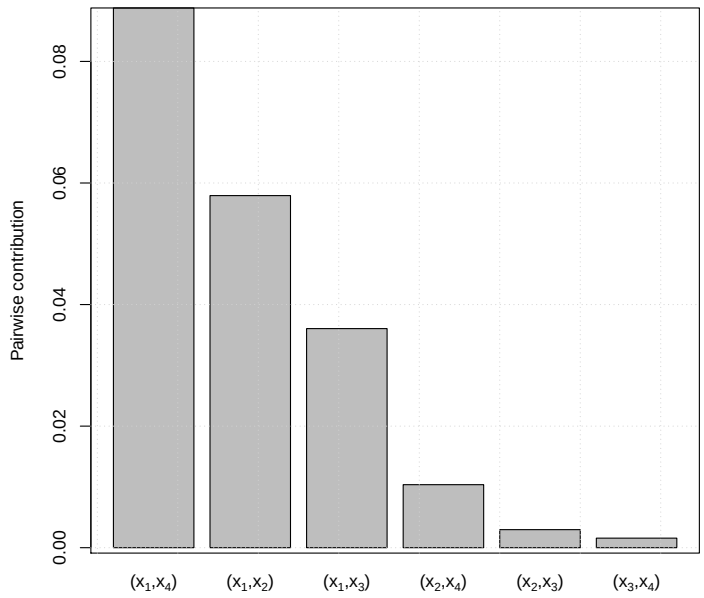


Figure 5.11: Ranking of pairwise contributions that were calculated using the kernel PCA T^2 statistic as distance measure.

Each of these calculated T^2 values has an associated quantification of its significance. For example, as demonstrated above, it was found that the T^2 distance calculated for the new observation on (x_1, x_2) was significantly large. This information can therefore be used to identify significant pairwise contributions. The variable combinations with significantly large contributions are: (x_1, x_2), (x_1, x_3) and (x_1, x_4). These variable pairs represent the combinations on which the engineer should focus on in order to get the process back into multivariate control.
The row sums method, discussed in Section 5.3, was used to identify an individual ranking of the process variables that can be used to supplement the information portrayed by the significant pairwise contributions. Only process variable x_1 was determined to be significant in the identified ranking. Recall from Section 4.5 that this is the process variable that was altered to construct the simulated faulty observation.

5.5.3 SVM

The diagnostic methodology based on distances calculated using the SVM approach i.e., (5.49), will now be used to diagnose the fault simulated in Section 4.5. The following steps will be followed:

1. Clean the IOC data set using the kernel PCA model used in the fault identification step i.e., remove all of the samples that are out-of-control based on one of the fault detection statistics.

2. For each of the $\binom{4}{2}$ process variable combinations use the density estimation approach to identify the complementary data set.

3. Fit a SVM model for each of the $\binom{4}{2}$ process variable combinations.

4. Calculate the class membership probability (5.49) of the new sample.

It will now be illustrated how the SVM based pairwise contribution can be calculated for one of the process variable pairs. Therefore, consider the scatter plot of (x_1, x_3) in Figure 5.12. The black dots again represent the reference region and the red dot the new observation. Bivariate density estimation was used to specify the complementary data set. This data set is identified by the points occupying the grey region in Figure 5.13. A Gaussian kernel based SVM model was fitted using the reference and complementary data sets. A grid search was performed to identify the optimal model settings for τ and σ, the SVM hyperparameters. Model performance was evaluated based on the crossvalidated misclassification rate. The optimal settings identified were $\tau = 188$ and $\sigma = 1.27$. This model was now used to predict the probability that the red dot in Figure 5.12 is part of

the grey region in Figure 5.13. The predicted probability equalled 0.9889. It can therefore be concluded that the new observation is statistically remote from the reference data for (x_1, x_3). A visual confirmation is made when Figure 5.12 is considered.

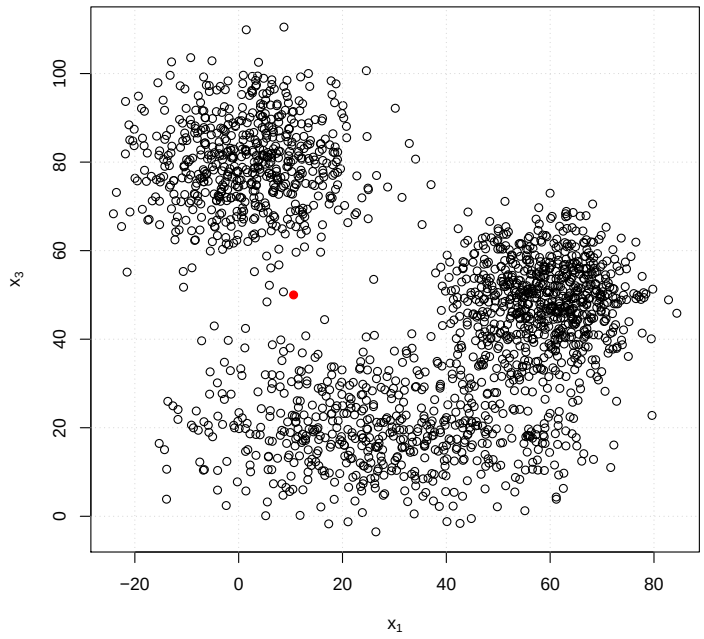


Figure 5.12: Black dots in scatter plot represent the reference data of (x_1, x_3). Red dot represents the readings recorded for the simulated faulty sample.

The complete SVM based pairwise variable importance distance matrix for this faulty vector was calculated to equal

$$\begin{bmatrix} 0.0000 & 1.0000 & 0.9889 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0232 \\ 0.9889 & 0.0000 & 0.0000 & 0.0000 \\ 1.0000 & 0.0232 & 0.0000 & 0.0000 \end{bmatrix}. \tag{5.64}$$

Figure 5.14 displays the pairwise importances in order.

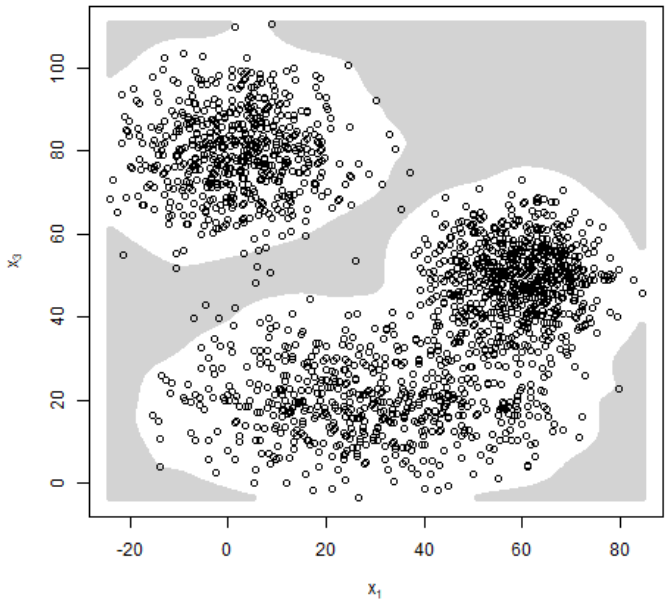


Figure 5.13: Grey points in the range of variable x_1 and x_3 determined by the density estimate to have low density.

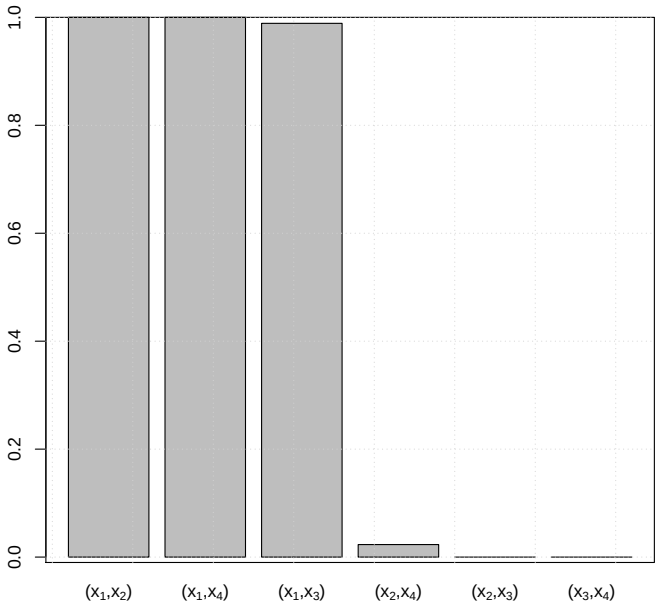


Figure 5.14: Ranking of pairwise importances calculated using SVM classification.

Note that since class probabilities are used to quantify distances, it is not required that null distributions be calculated for each of the variable pairs in order to determine the

significance of their observed distances. The natural class cut-off probability of 0.5 can be used to identify distances that are significantly large. The pairwise importance of the variable combinations (x_1, x_2), (x_1, x_3) and (x_1, x_4) is therefore identified to be significantly large. These pairs of variables should therefore be the focus of the investigation to gain insight into the process deviation. Figure 5.15 displays the pairwise information. In each sub plot, black points are used to represent the reference data and a red dot is used to identify the measurements made on the simulated fault. It is clear from Figure 5.15 why the variable pairs (x_1, x_2), (x_1, x_3) and (x_1, x_4) were identified as important contributors. The engineering team should therefore focus their efforts to get the red observation close to the black points for the identified variable pairs.

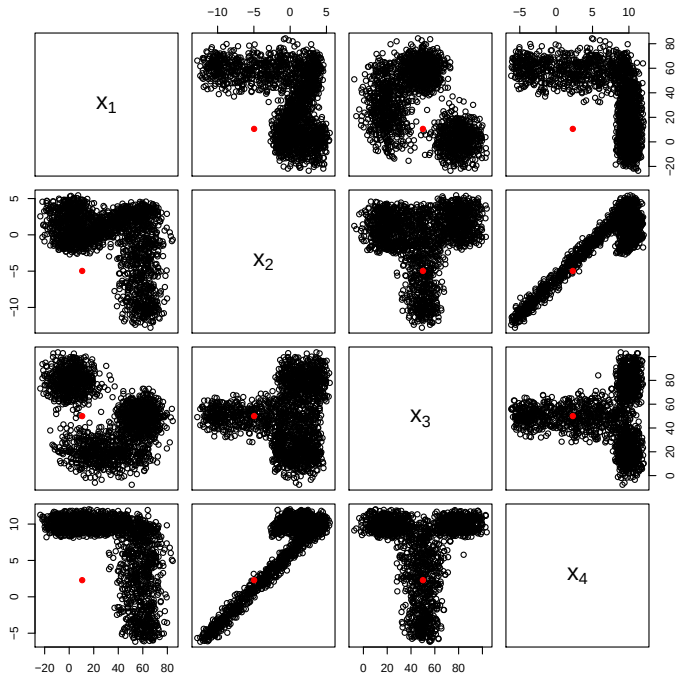


Figure 5.15: Scatterplot matrix of the reference data with the measurements of the simulated faulty measurement included. Black points represent reference data and the red points are the faulty data.

The individual ranking of the process variables using the pairwise distances was also calculated. For the first iteration of the row sums method, the empirically calculated p-values were

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0.000000 & 0.000526 & 0.005783 & 0.000526 \end{bmatrix}. \tag{5.65}$$

The first process variable therefore obtains the highest ranking. After removing row 1 and column 1 from (5.64) the empirical p-values of the resulting row sums were again calculated

$$\begin{bmatrix} 2 & 3 & 4 \\ 0.057308 & 0.936383 & 0.066246 \end{bmatrix}. \tag{5.66}$$

None of the remaining row sums are significantly large and the ranking procedure can therefore be stopped. The first process variable is therefore the only one that needs to be considered together with the high ranking variable pairs i.e., (x_1, x_2), (x_1, x_3) and (x_1, x_4), to understand the process fault.

5.5.4 One-Class SVM

In this section One-Class SVM discriminant analysis will form the basis of the new fault diagnosis methodology. The diagnosis of the simulated fault in Section 4.5 will be demonstrated using the following steps:

1. Clean the IOC data set using the kernel PCA model used in the fault identification step i.e., remove all of the samples that are out-of-control based on one of the fault detection statistics.

2. Fit a OCSVM model for each of the $\binom{4}{2}$ process variable combinations.

3. For each of the variable combinations create a fine grid of points that cover all bivariate combinations of values in the data.

4. Evaluate the bivariate grid vector using the fitted OCSVM model to obtain the decision values (5.54).

5. Use the method of Platt (1999) to fit a model, based on the grid evaluations, that can be used to convert a decision value (5.54) into a class membership probability (5.61).

6. Obtain the model evaluations (5.54) of the simulated faulty sample.

7. Convert the model evaluation into a probability of not being part of the reference region i.e., (5.61).

One process variable combination will be selected to illustrate these steps. In Figure 5.10 the reference data (black points) together with the measurements of the faulty sample (red

dot) are displayed for (x_1, x_4). A OCSVM model, with regularisation parameter $\nu = 0.05$, was fitted using the reference data in Figure 5.10. The Gaussian kernel function, with hyperparameter $\sigma = 0.94$, was used to approximate the inner products in feature space. Note that crossvalidation was used to prevent overfitting. A fine grid was then created over the combined range of (x_1, x_4). The model was then used to evaluate the bivariate vectors specified in the grid. Figure 5.16 graphically summarises the result of the grid evaluation. The grey area in Figure 5.16 is where the model evaluation is greater than zero i.e., points in this region are predicted to be close to the reference data. Similarly the points in the light blue area are predicted to be remote from the reference region i.e., the model evaluations reported values that are less than zero for these points. The method of Platt (1999) was now used to specify the posterior complement class probability function (5.61) for these grid evaluations. In Figure 5.17 the calculated probabilities are plotted against the model evaluations. The model evaluation (5.54) of the simulated faulty sample equals -10.466. This is indicated by the vertical blue line in Figure 5.17. The faulty observation is therefore considered to be remote from the reference region since it has a zero probability of being close. Therefore, the pairwise contribution of (x_1, x_4) for the simulated faulty sample equals one.

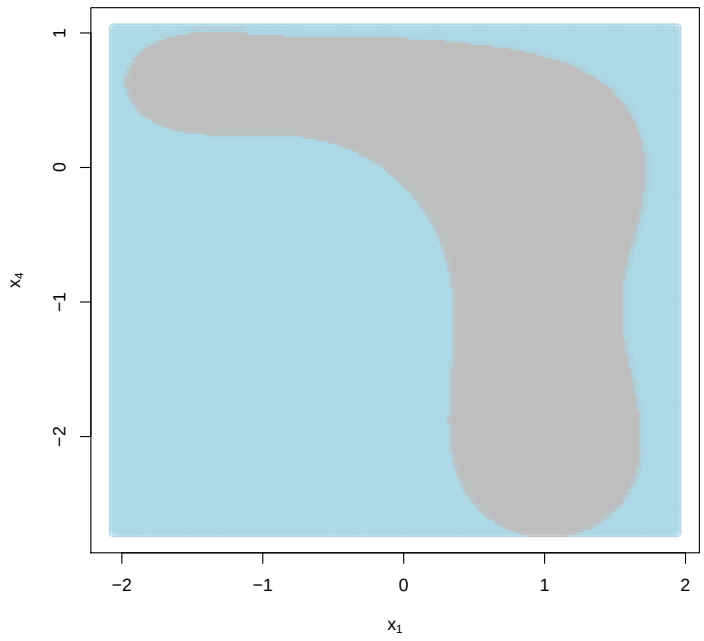


Figure 5.16: OCSVM model evaluation of grid vectors in the range of (x_1, x_4). Blue area is where the model predicted that the corresponding bivariate vectors are remote from the reference data. The vectors in the grey area are predicted to be close to the reference region.

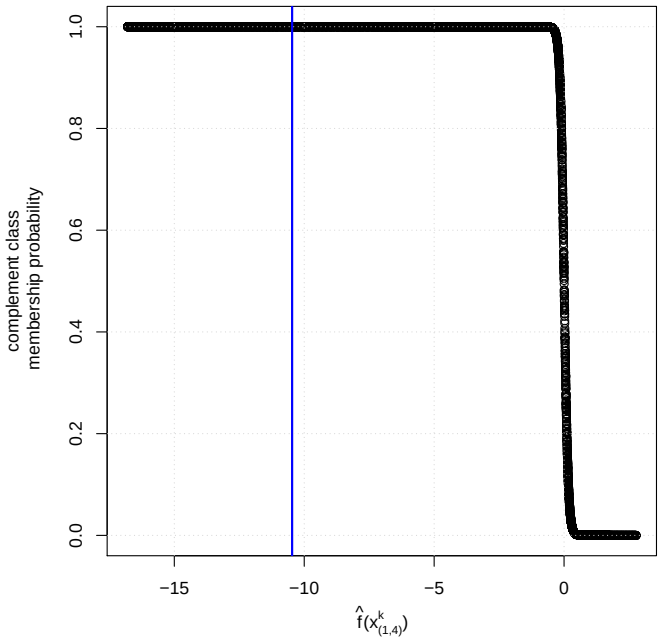


Figure 5.17: Estimated probability function of the OCSVM model fitted for (x_1, x_4). The vertical blue line identifies the model prediction for the simulated faulty sample.

The matrix containing the full set of pairwise contributions, for the faulty sample, calculated using the OCSVM approach equals

$$\begin{bmatrix} 0.0000 & 1.0000 & 0.8802 & 1.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.8802 & 0.0000 & 0.0000 & 0.0000 \\ 1.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}. \tag{5.67}$$

The relative sizes of the pairwise contributions are depicted in Figure 5.18. A contribution is considered significant if a value greater than 0.5 is recorded, similar to the approach followed for the SVM method. Therefore, the pairwise contributions of (x_1, x_2), (x_1, x_3) and (x_1, x_4) are significant. The row sums method was employed to identify an individual importance ranking of the process variables. Similar to the previous results, only variable x_1 was determined to be important.

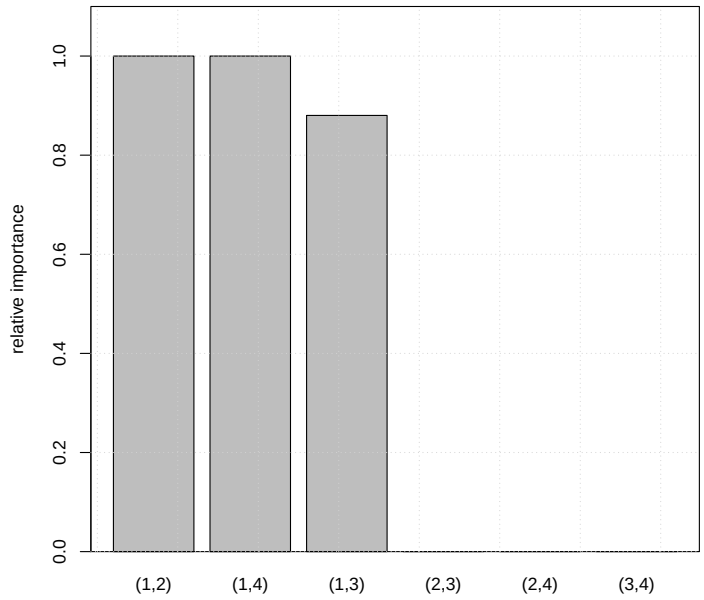


Figure 5.18: Ranking of pairwise importances calculated using OCSVM classification method.

5.6 Kernel PCA Simulation Study continued (Part 2)

5.6.1 Objective

The kernel PCA simulation study will now be extended to evaluate the ability of the newly developed techniques to diagnose single, multiple and multivariate process faults. A comparison will be done against the traditional kernel PCA diagnosis techniques, presented in Chapter 4, using the fault data that was used in Section 4.5.4. Therefore, this Section is a continuation of Section 4.5.4. Note that the results obtained in this simulation study is for information purposes since it is an evaluation of the individual importances only. It is advised that the pairwise ranking of process variables should accompany the individual importances when interpreting multivariate process deviations. The purpose of this simulation study is to give a general idea of how the different techniques would compare when it is applied for the calculation of individual process variable rankings. It was decided to only base the distance calculations on the SVM analysis for ease of illustration. For single sensors the row sum approach was employed to identify the high ranking process variables. The highest ranking variable pair based on the pairwise distances was used in the multiple and multivariate sensor fault diagnosis.

5.6.2 Results

The results of this study are illustrated in Figures 5.19 to 5.27. The following is observed:

- Single sensor faults. In Figure 5.19, for the SPE fault index, it is observed that the new approach recorded relatively better correct diagnosis rates, for all fault sizes, compared to the traditional methods. For the T^2 index, in Figure 5.20, it is observed that the pairwise distance method showed improved performance for fault magnitudes less than two. The reconstruction based method of Alcala and Qin (2010) recorded better results for fault sizes greater than two. In Figure 5.21, the results for the *Combined* index report that the pairwise distance approach was the best performing method across the range of fault magnitudes.

- Multiple sensor faults. For the SPE fault index, in Figure 5.22, it is observed that the new approach based on pairwise distances reported improved performance, relative to the other methods, for all fault sizes. In Figure 5.23, for the T^2 statistic, it is also observed that the pairwise method recorded better correct diagnosis rates for all fault magnitudes. A similar observation is made for the *Combined* index in Figure 5.24.

- Multivariate sensor faults. In Figure 5.25 the results for the SPE index indicate that the new approach recorded correct fault diagnosis rates that are better than the traditional techniques for all fault sizes. For the T^2 fault detection index, in Figure 5.26, it is observed that the pairwise distance based approach outperformed the reconstruction and gradient based methods. In Figure 5.27, a similar conclusion is reached for the *Combined* index. More over, the new pairwise distance approach yielded about 100% correct diagnosis rates over all fault sizes for all the fault detection indices.

It is noted that there are some interesting trends that are exhibited by the different diagnosis techniques, as a function of the fault sizes. For example, in Figure 5.19, it is observed that the correct diagnosis rate of the new approach increases as fault size increases. However, an attempt to interpret these behaviours will not be made, since individual importances are being evaluated. It has been motivated earlier that pairwise information would do a better job of explaining multivariate deviation than the individual ranking of variables. The individual information should be used to supplement the pairwise analysis.

5.6.2.1 Single sensor fault

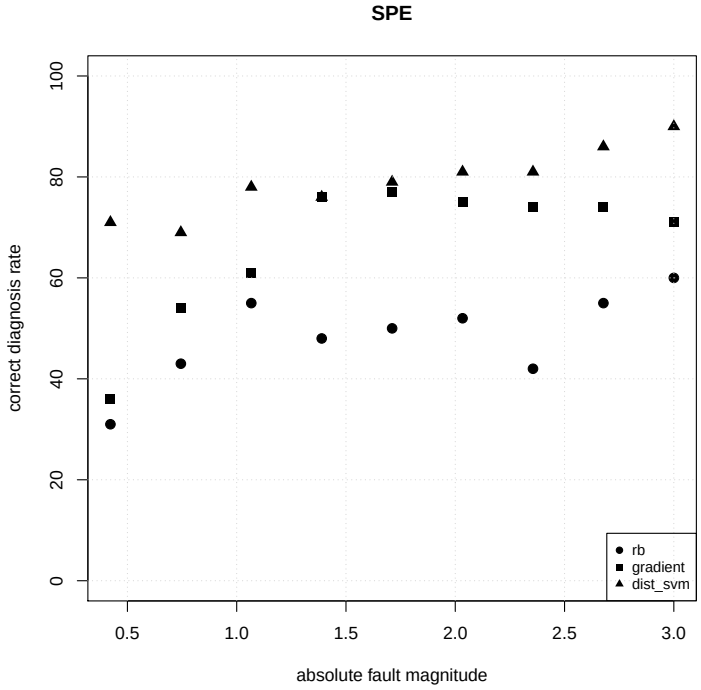


Figure 5.19: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

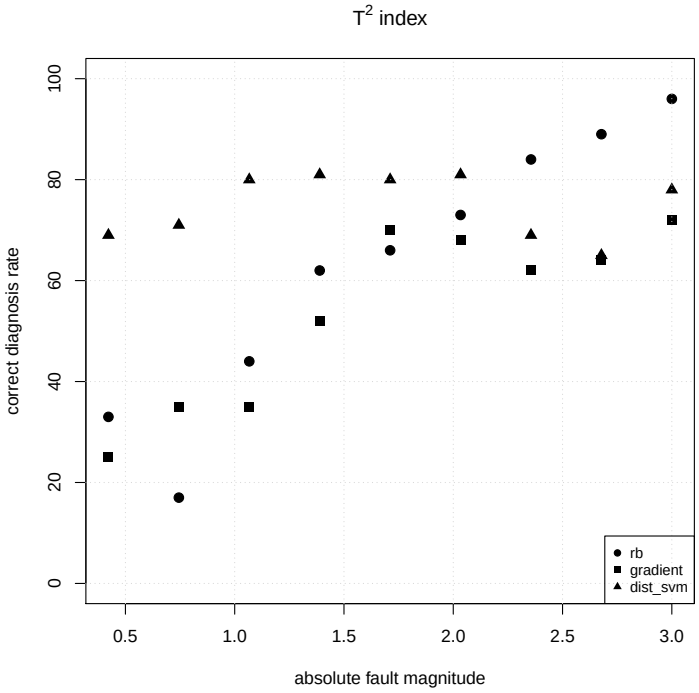


Figure 5.20: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

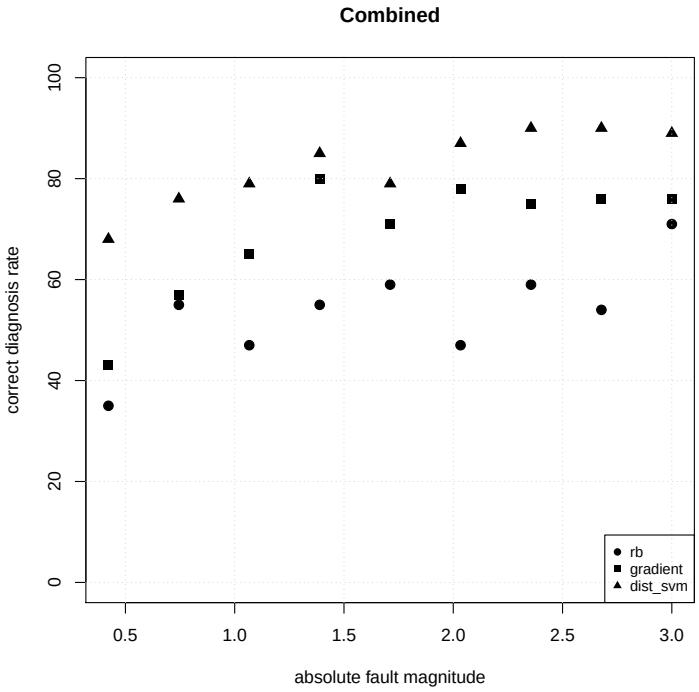


Figure 5.21: Correct diagnosis of faulty variables expressed relative to the magnitude of the single sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

5.6.2.2 Multiple sensor fault

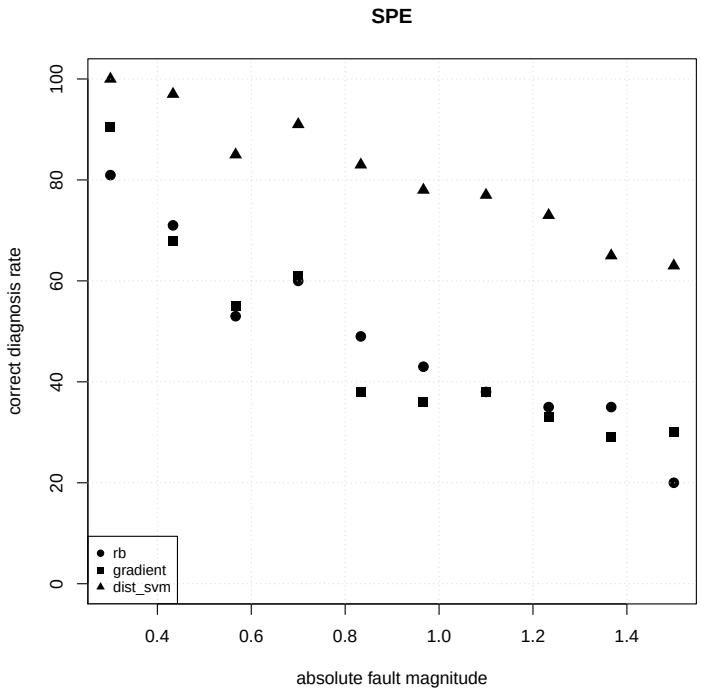


Figure 5.22: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

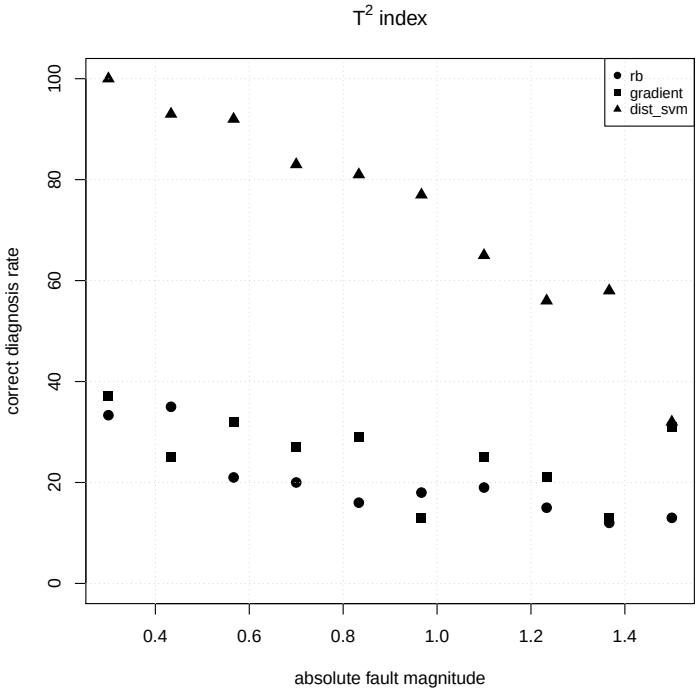


Figure 5.23: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

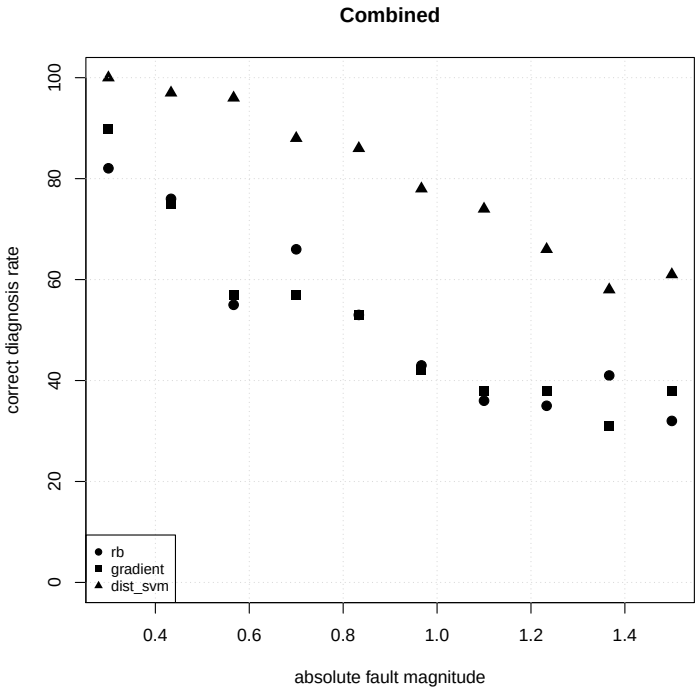


Figure 5.24: Correct diagnosis of faulty variables expressed relative to the magnitude of the multiple sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

5.6.2.3 Multivariate sensor fault

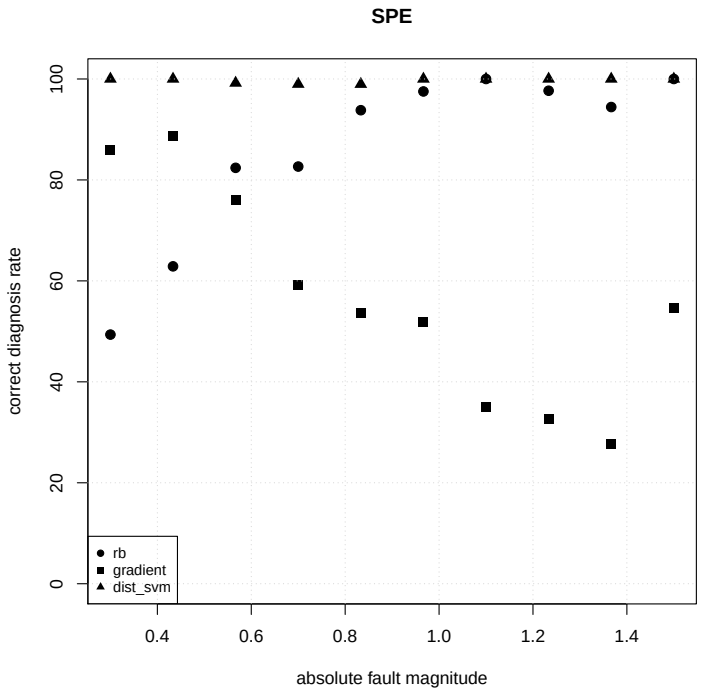


Figure 5.25: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the SPE index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

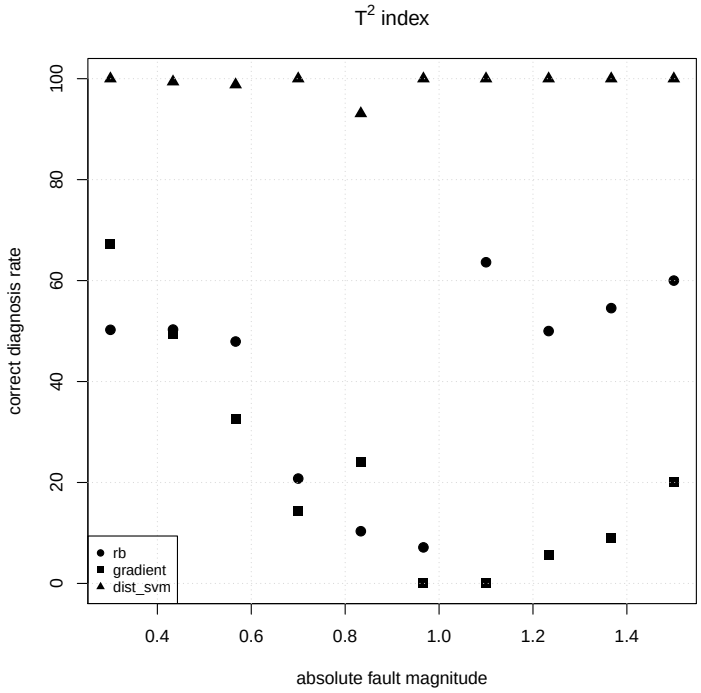


Figure 5.26: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the T^2 index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

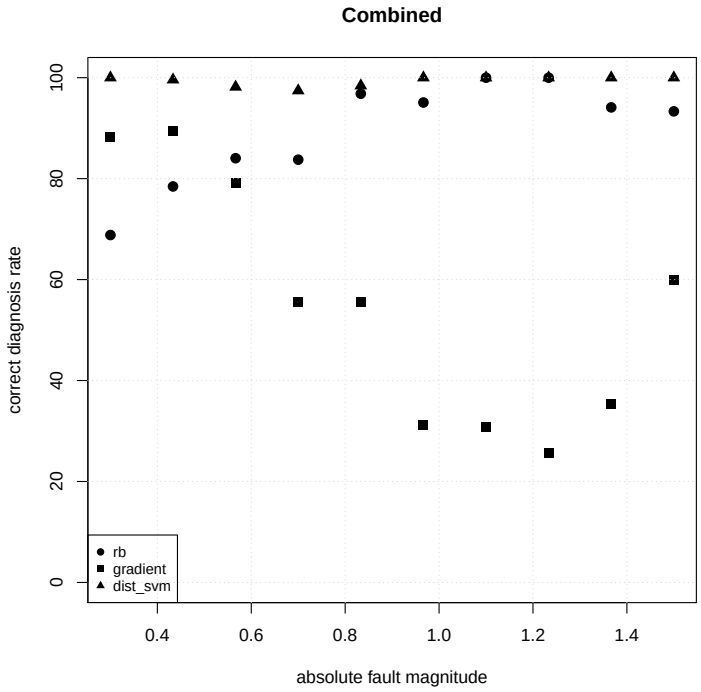


Figure 5.27: Correct diagnosis of faulty variables expressed relative to the magnitude of the multivariate sensor fault induced. Contributions were calculated using the *Combined* index and corresponding kernel PCA model. Solid squares are used to represent the results of the gradient based method of Cho et al. (2005). The results of the reconstruction based method of Alcala and Qin (2010) are indicated using solid circles. Triangles are used to identify the results generated by the new pairwise distance methodology. These distances are calculated using SVM based pairwise models.

5.7 Summary

Similar to the discussion in Chapter 3, it has been illustrated for kernel PCA that the notion of attempting to assign an individual importance to process variables is not the correct question that needs to be answered. Existing techniques employed to diagnose process faults identified by kernel PCA attempt to assign individual importance rankings to process variables. These diagnostic techniques are therefore fundamentally flawed. It also does not provide practical guidance to engineers on what the required outcome should be of any corrective measure in order to achieve a process that is multivariately in-control. In addition, assigning individual importance to process variables provides limited insight to engineers on corrective actions required to achieve a process that is multivariate in-control. The new methodology proposed provides greater insight into fault diagnosis since pairwise importance ranking or process variables is provided.

It is proposed that the methodology developed in Chapter 3 be leveraged to develop a diagnostic method for kernel PCA, which produces a pairwise importance ranking of the

process variables. It is not clear how one would go about decomposing the kernel PCA fault identification statistics into a sum of parts with each part related to a specific variable pair. Therefore, the approach developed in Chapter 3 cannot be applied in the kernel PCA setting. In order to obtain a pairwise contribution for each process variable combination, it was proposed to consider each variable pair in isolation. The pairwise contribution of a specific variable combination is then defined as the distance which quantifies how remote an observation is from the bivariate IOC point cloud for the two variables. This method of calculating pairwise contributions is therefore not based on a decomposition of the fault identification statistics. Significant pairwise contributions are identified using the reference set density estimation approach introduced in Chapter 3. In a similar fashion to Chapter 3, the pairwise contributions can be analysed to obtain an individual importance ranking of the process variables.

The applicability of this methodology is dependent on appropriately specifying a distance function that can be used to quantify the similarity of a new observation to the IOC region. Conditions have been specified that a distance measure must comply with in order for it to be suitable for use in the diagnosis methodology. Three different approaches have been discussed to demonstrate the process of specifying acceptable distance measures. The first approach utilises the scaled kernel PCA fault identification statistics as suitable distance measures. A distance measure based on SVM classification is employed in the second example. In the third approach a distance measure based on OCSVM is developed.

The simulation study introduced in Chapter 4 was again considered to illustrate the application of the new diagnosis methodology. Specifically, it was demonstrated how this methodology, based on the three different distance measures proposed, can be used to diagnose the fault simulated in Section 4.5. It is clear from the simulation results that the calculated pairwise contributions are sensible i.e., it provides information that clearly indicate to engineers why the process is classified as being multivariately out-of-control. Any corrective process change, implemented by the engineer to remedy the multivariate digression, should therefore address the deviations observed in the identified variable pairs. It was also demonstrated how the individual ranking can be calculated using the pairwise contributions. For this illustration the correct process variable was identified as the most important. Recall from Section 4.5 that the first process variable was altered to induce the faulty observation.

A more comprehensive evaluation of individual variable importance ranking, using the pairwise approach, is presented in Section 5.6. Correct diagnosis rates are calculated using the single, multiple and multivariate sensor faults that were simulated in Section 4.5.4. The performance of the new approach is therefore compared against the results obtained for the traditional diagnosis techniques that are introduced in Chapter 4. It is observed,

in most cases, that the new approach recorded correct diagnosis rates that were higher than that of the gradient and reconstruction based methods. It is again emphasised that individual importances will not sufficiently explain to engineers why the process is considered to be multivariately out-of-control. The bivariate diagnosis information should be communicated together with the individual rankings. Therefore, the correct diagnosis rate simulation results should only be viewed as an attempt to compare the new approach to existing techniques.

The practical application of this methodology will be further evaluated in Chapter 6 using a benchmark data set and data from a commercial chemical plant.

Chapter 6

Practical application

6.1 Introduction

The true value and applicability of the new fault diagnosis methodology motivated in this research document can only be assessed by applying it to a real life manufacturing example. Therefore, in this Chapter the multivariate monitoring and subsequent fault diagnosis of two commercial chemical plants will be considered. The first case study will focus on the diagnosis of faults simulated in the well known Tennessee Eastman process (Downs and Vogel, 1993). In the second case study process data produced by a commercial coal fired steam boiler will be analysed.

In this research a new methodology has been introduced for fault diagnosis in multivariate statistical process monitoring. The objective here is to evaluate the behaviour of this new methodology when observed in practice. New methods were developed for fault diagnosis in PCA and kernel PCA based MSPM. Therefore, the commercial plants will be statistically monitored using PCA and kernel PCA models. The newly developed methods will then be applied to diagnose process faults identified by the two models. The pairwise methodologies will be evaluated based on their ability to clearly communicate why the process conditions are statistically identified as displaying behaviour that is different from multivariate common cause variability.

In addition to quantifying the diagnostic value of the new methods in the practical setting, the added objective of this Chapter is to practically illustrate, by means of R code in Appendix A, how to operationalise the methodology in practice.

6.2 Case study: The Tennessee Eastman Process

6.2.1 Background and Objective

The Eastman Chemical company introduced the Tennessee Eastman (TE) process (Downs and Vogel, 1993) as a simulation of a realistic industrial process for the evaluation of process monitoring methodologies. Therefore, many of the literature on fault diagnosis use the TE process as a consistent benchmark to quantify the performance of new approaches. Examples are Li et al. (2009), Jiang et al. (2015), Wang et al. (2017) and Gharahbagheri et al. (2017). It is therefore appropriate that the methods introduced in this thesis are also evaluated using the data simulated by the TE process. The TE process will be used to evaluate the diagnosis ability of both the linear and non-linear pairwise contribution methodologies introduced in this research. Diagnosis results will be reported in a statistical manner that follows the perspective of the research that was presented i.e., no attempt will be made to interpret the results from an engineering viewpoint.
Five main operational units are used to describe the TE process. A reactor, a condenser, a vapour-liquid separator, a recycle compressor and a product stripper identifies the different operations. The TE process flow diagram is depicted in Figure 6.1. A total of 52 process measurements are recorded. There are 22 continuous variables (Table 6.1), 19 compositional measurements (Table 6.2) and 11 manipulated variables (Table 6.3).

6.2.2 Variables and Data

Historical data representing ideal operating conditions were simulated and are provided by Rieth et al. (2017). Process measurement data are conveniently made available as R data frames. The IOC samples are recorded in the data frame called 'fault_free_training'. A total of 500 samples were taken at three minute intervals i.e., this data frame represents 25 hours of operation. Note that the 'fault_free_training' data set consists of 250000 rows, with each of the 500 consecutive subsets of 500 rows representing one replicated simulation run. In this illustration, it is assumed that three of the simulations are sufficiently representative of the IOC conditions and will therefore only focus on the first 1500 measurements.
Rieth et al. (2017) also provide TE simulation measurements observed after specific process faults were introduced. A list of the 20 different process faults that were explored are recorded in Table 6.4. A separate simulation run is recorded for each of the faults, with each fault being introduced into the simulation after 60 minutes from the start of the run. Measurements are recorded every three minutes for a duration of 25 hours. The faulty simulation run data are provided in the R data frame called 'faulty_training'. Note that similar to the 'fault_free_training' data frame, 'faulty_training' consists of multiple

replicated simulation runs. For each of the 20 process faults 500 replicated runs are performed. In this illustration we will only focus on the first replication, i.e., the first 500 observations, for a selected process fault.

Three simulated TE faults will be selected to demonstrate the diagnostic capability of the methods presented in this research. These are fault one, 12 and 17 in Table 6.4. Downs and Vogel (1993) provide a classification of the fault type that can be used to describe how the different TE faults were simulated. These categorisation are also included in Table 6.4. The different faults that will be examined are therefore from different fault type categories.

Table 6.1: The 22 continuous Tennessee Eastman process variables.

Process Variables	**Description**	**Unit**
x_1	A feed (stream 1)	km^3h^{-1}
x_2	D feed (stream 1)	kgh^{-1}
x_3	E feed (stream 1)	kgh^{-1}
x_4	Total feed (stream 4)	km^3h^{-1}
x_5	Recycle flow (stream 8)	km^3h^{-1}
x_6	Reactor feed rate (stream 6)	km^3h^{-1}
x_7	Reactor pressure	kPa
x_8	Reactor level	%
x_9	Reactor temperature	$^\circ C$
x_{10}	Purge rate (stream 9)	km^3h^{-1}
x_{11}	Product separator temperature	$^\circ C$
x_{12}	Product separator level	%
x_{13}	Product separator pressure	kPa
x_{14}	Product separator underflow (stream 10)	m^3h^{-1}
x_{14}	Product separator underflow (stream 10)	m^3h^{-1}
x_{15}	Stripper level	%
x_{16}	Stripper pressure	kPa
x_{17}	Stripper underflow (stream 11)	m^3h^{-1}
x_{18}	Stripper temperature	$^\circ C$
x_{19}	Stripper stream flow	kgh^{-1}
x_{20}	Compressor work	kW
x_{21}	Reactor cooling water outlet temperature	$^\circ C$
x_{22}	Condenser cooling water outlet temperature	$^\circ C$

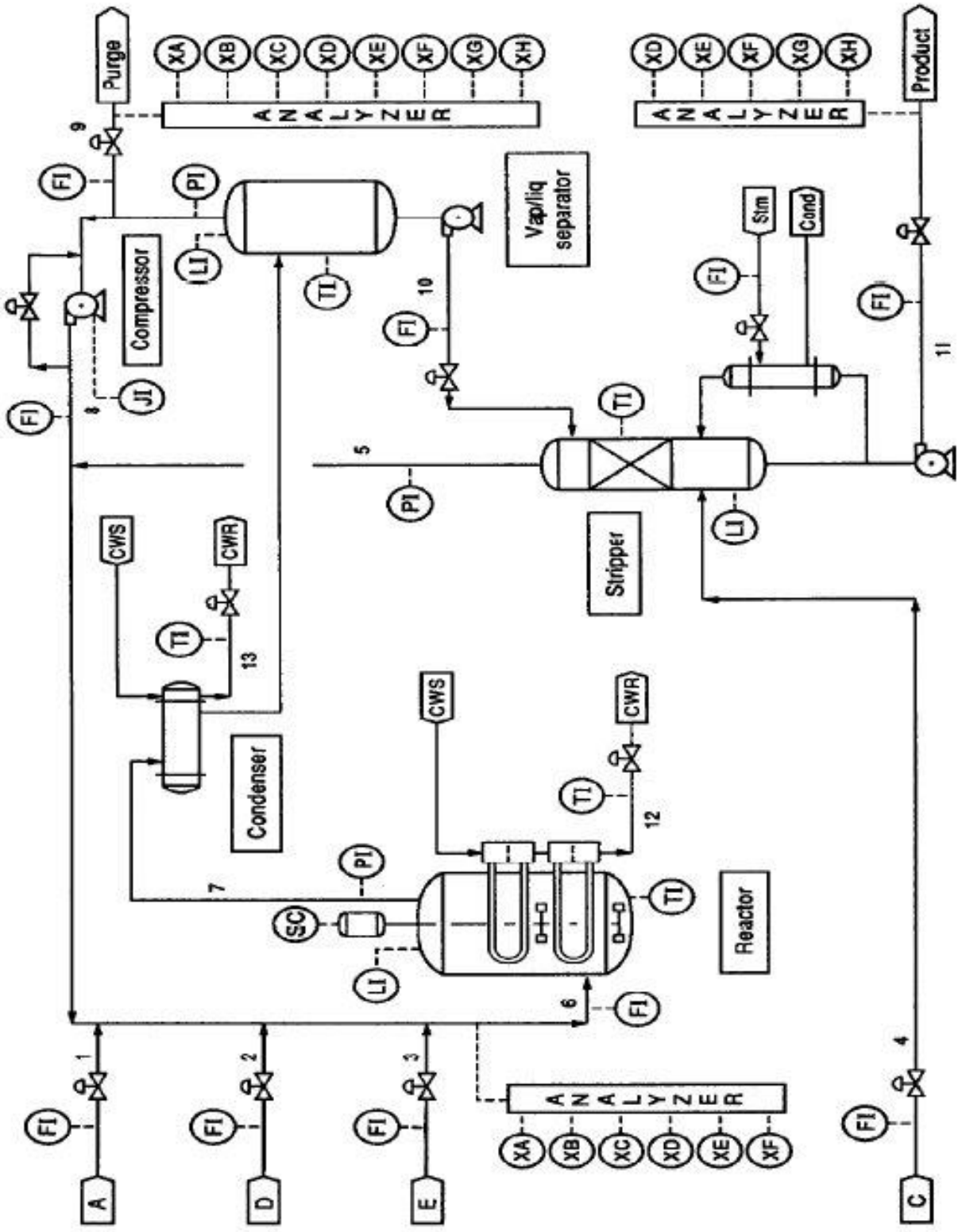


Figure 6.1: Tennessee Eastman process flow diagram (Downs and Vogel, 1993).

Table 6.2: The 19 compositional Tennessee Eastman process measurements.

Process Variables	Description	Unit
x_{23}	Component A (stream 6)	$mol\%$
x_{24}	Component B (stream 6)	$mol\%$
x_{25}	Component C (stream 6)	$mol\%$
x_{26}	Component D (stream 6)	$mol\%$
x_{27}	Component E (stream 6)	$mol\%$
x_{28}	Component F (stream 6)	$mol\%$
x_{29}	Component A (stream 9)	$mol\%$
x_{30}	Component B (stream 9)	$mol\%$
x_{31}	Component C (stream 9)	$mol\%$
x_{32}	Component D (stream 9)	$mol\%$
x_{33}	Component E (stream 9)	$mol\%$
x_{34}	Component F (stream 9)	$mol\%$
x_{35}	Component G (stream 9)	$mol\%$
x_{36}	Component H (stream 9)	$mol\%$
x_{37}	Component D (stream 11)	$mol\%$
x_{38}	Component E (stream 11)	$mol\%$
x_{39}	Component F (stream 11)	$mol\%$
x_{40}	Component G (stream 11)	$mol\%$
x_{41}	Component H (stream 11)	$mol\%$

Table 6.3: The 11 manipulated Tennessee Eastman process variables.

Process Variables	Description	Unit
x_{42}	D feed flow (stream 2)	%
x_{43}	E feed flow (stream 3)	%
x_{44}	A feed flow (stream 1)	%
x_{45}	A and C feed flow (stream 4)	%
x_{46}	Compressor recycle valve	%
x_{47}	Purge valve (stream 9)	%
x_{48}	Separator pot liquid flow (stream 10)	%
x_{49}	Stripper liquid product flow (stream 11)	%
x_{50}	Stripper steam valve	%
x_{51}	Reactor cooling water flow	%
x_{52}	Condenser cooling water flow	%

Table 6.4: Process faults simulated in the Tennessee Eastman process.

Fault id	Fault description	Fault type
1	A/C feed ratio, B composition constant	Step change
2	B composition, A/C feed ratio constant	Step change
3	D feed temperature (stream 2)	Step change
4	Reactor cooling water inlet temperature	Step change
5	Condenser cooling water inlet temperature	Step change
6	A feed loss (stream 1)	Step change
7	C header pressure loss-reduced availability (stream 4)	Step change
8	A, B, C feed composition (stream 4)	Random fluctuation
9	D feed temperature (stream 2)	Random fluctuation
10	C feed temperature (stream 4)	Random fluctuation
11	Reactor cooling water inlet temperature	Random fluctuation
12	Condenser cooling water inlet temperature	Random fluctuation
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16	Unknown	Unknown fault type
17	Unknown	Unknown fault type
18	Unknown	Unknown fault type
19	Unknown	Unknown fault type
20	Unknown	Unknown fault type

6.2.3 PCA based fault diagnosis

In this section the fault detection and diagnosis will be based on PCA. Specifically, the pairwise decomposition approach developed in Chapter 3 will be used to diagnose fault one, 12 and 17. Diagnosis results pertaining to the traditional methodologies of Chapter 2 will also be reported.

6.2.3.1 Diagnosis of TE Fault 1.

In this section the ability of the PCA based methods to detect and diagnose the first simulated TE process fault will be evaluated . The evaluation will be performed using the R programming language. R code used to conduct the analysis is displayed in Appendix A. The following can be used to summarise the steps used to fit the PCA model:

1. Selection of the IOC data samples from the 'fault_free_training' R data frame. Use only three of the simulation replicates i.e., the first 1500 observations. Use all 52 variables in the analysis.

2. Zero mean centering and unit variance scaling of the IOC data set.

3. Defining the covariance matrix and subsequent eigen-decomposition.
4. Selecting the appropriate number of principal components.
5. Calculate the matrices used to define the PCA fault identification statistics.
6. Determine the critical value for each of the fault identification statistics.

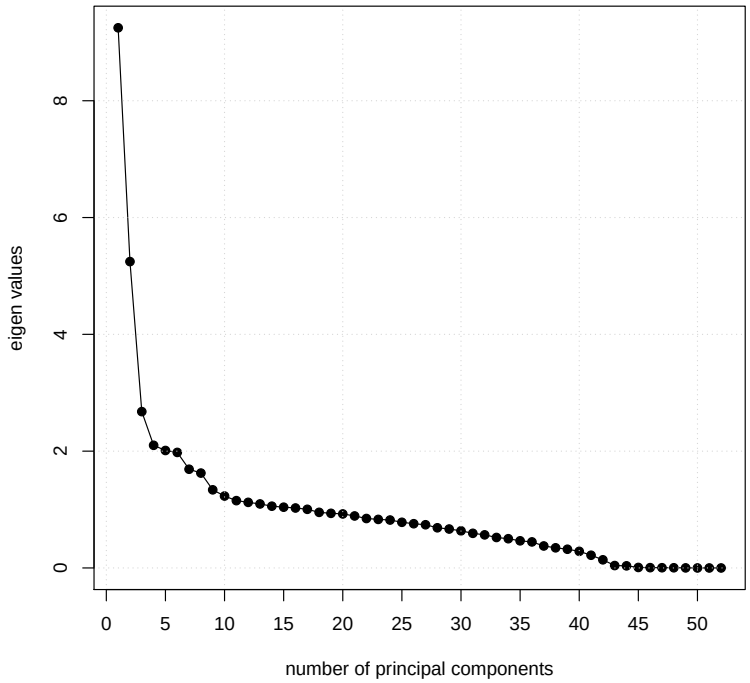


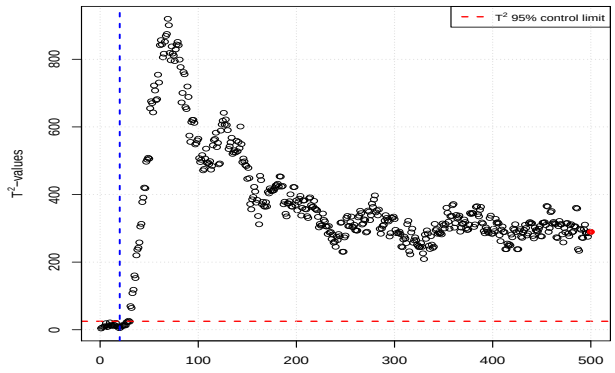
Figure 6.2: PCA Scree plot. The eigenvalues are graphed as a function of the number of principal components. A general rule is to select the number of components as the point on the graph where the relative change in eigenvalues start to stabilise.

The number of principal components were selected by visual inspection of the PCA scree plot displayed in Figure 6.2. From this representation it was determined that 15 components should be sufficient. Selecting 15 components imply that the PCA model explains about 67% of the variability. It is noted that other techniques can be employed to select the number of principal components (Josse and Husson, 2012). For this illustration it was found that 15 components are adequate to detect the induced process faults that will be presented.
The following describe the process that was followed to obtain the fault identification statistic evaluations of the new faulty simulated data:

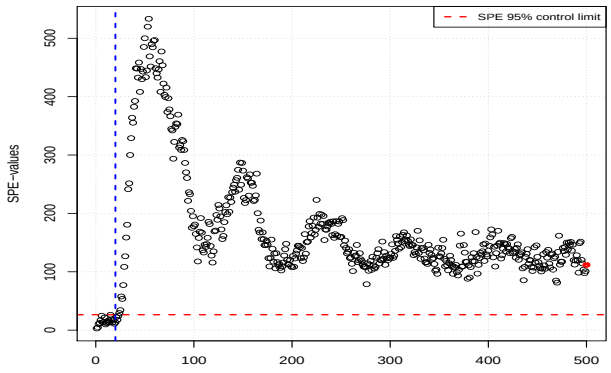
1. Identify the data that correspond to the first simulated process fault in the 'faulty_training' data set. Select only one of the replicated simulations i.e., the first 500 observations.

2. Center and scale the selected faulty data using the mean and variance information of the IOC data.

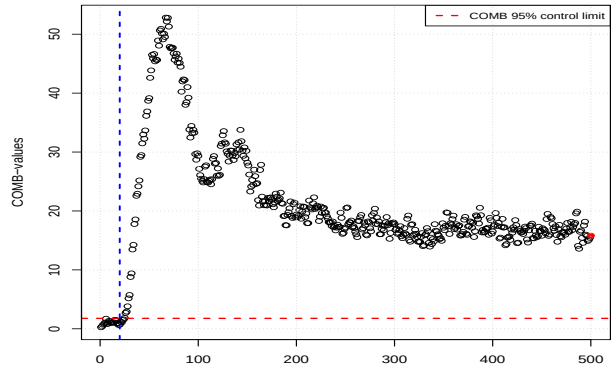
3. Calculate the PCA fault identification statistic evaluations i.e., the T^2, SPE and *Combined* values.



(a) T^2 statistic evaluations.



(b) SPE statistic evaluations.



(c) *Combined* statistic evaluations.

Figure 6.3: PCA fault identification statistic evaluations of the TE Fault 1 simulation data. Evaluations are displayed for the T^2, SPE and $Comb$ statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.

The calculated fault identification statistic values for the faulty simulated data of the first TE process fault are depicted in Figure 6.3. Horizontal red broken lines are used in Figure 6.3 to indicate the 95% control limit for each of the respective statistics. Observations recorded above this line indicate that the process is statistically different from IOC's. It

can therefore be observed from Figure 6.3 that the data observed for the first simulated fault are mostly above the control limit, for all of the statistics. The blue vertical broken line in Figure 6.3 indicates the one hour point at which the fault was introduced in the simulation. It can be observed from Figure 6.3 that all three fault identification statistics displayed out-of-control evaluations directly after the fault was induced. In this example the SPE statistic was the quickest to identify the change in simulation conditions. The comparable shape of the observed fault statistic evaluations over time, for all statistics, is informative. It can be observed in Figure 6.3 that for the initial period after the fault has been introduced i.e., for the period prior to the 300^{th} observation, that the three statistics indicate erratic process behaviour. After this period the observations seem to become less erratic, which indicate that the effect of the process fault on the TE simulation has stabilised. In this illustration we will demonstrate the fault diagnostic analysis using the faulty data observed during the stabilised period. Specifically, the PCA fault diagnosis techniques will be used to diagnose the solid red points in Figure 6.3.

Before the diagnosis step is commenced it is noted for this specific fault i.e., the first fault in Table 6.4, that it cannot be considered to be a multivariate process fault. It is observed that a number of process variables have violated their univariate ranges, as is specified by the normal operating data. Consider the scatter plot of x_1 in Figure 6.4 as an example. In this figure the IOC data (hollow black points) and the faulty data (hollow green points) for x_1 are plotted in sequence. It is clear that the univariate ideal operating range, which approximately equals $(0.16, 0.38)$, is different from the range in which the faulty data are recorded. This therefore indicates that the multivariate fault diagnosis methods are not strictly required. The diagnosis can simply be performed by identifying the process variables that violate the univariate IOC ranges.

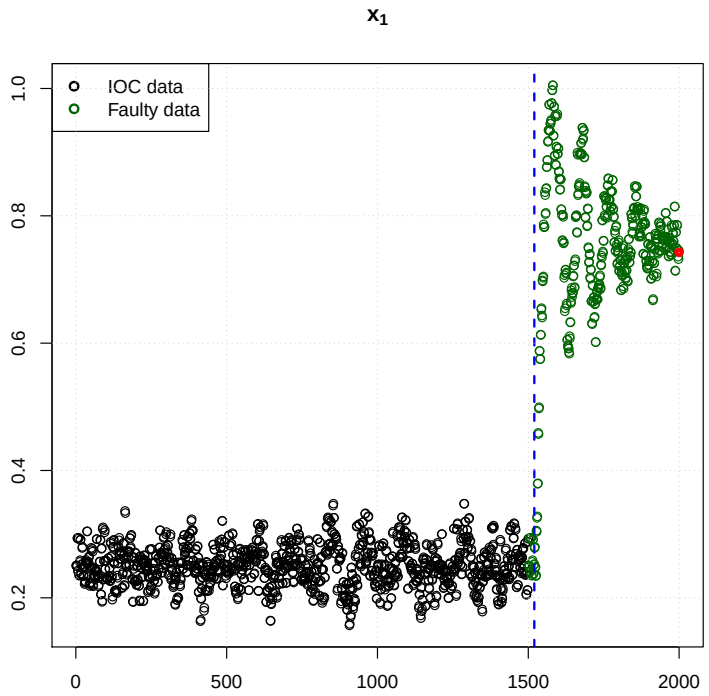


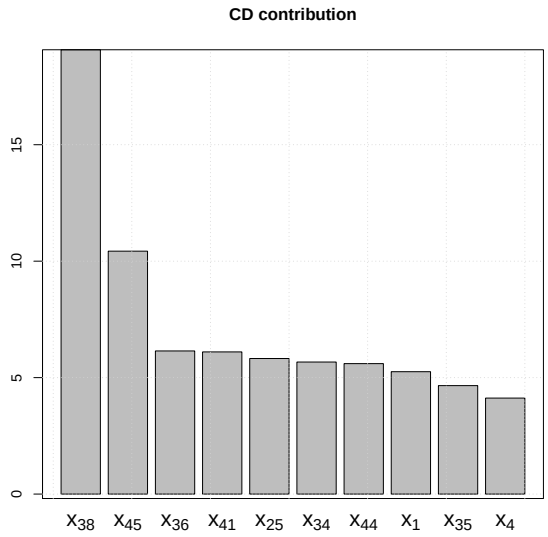
Figure 6.4: Measurements recorded for process variable x_1 plotted in sequence. The black points represent the values captured during the IOC simulations. Green points are used to indicate the data measured during the Fault 1 simulation. The vertical broken blue line indicates the point in time when the fault was introduced.

However, for illustration purposes and for comparison to published research e.g., Jiang et al. (2015), the diagnostic methods discussed in this research will be applied to diagnose the red points selected in Figure 6.3. Specifically, attention will be focussed on the diagnosis of the fault based on the SPE statistic. Note that a reasonable minimum requirement would be that the various fault diagnosis techniques correctly identify the variables that are univariately different from IOC's. The process variables corresponding to the red dot in graph (b) of Figure 6.3 that recorded measurements outside of the range observed during IOC's are:

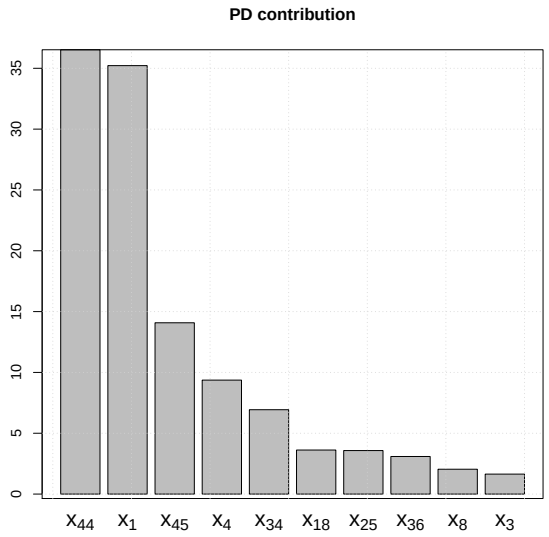
$$\begin{bmatrix} x_1 & x_4 & x_{18} & x_{19} & x_{34} & x_{44} & x_{45} & x_{50} \end{bmatrix}. \tag{6.1}$$

First the traditional PCA fault diagnosis techniques i.e., CD, PD and RB contribution analysis will be used to diagnose the observation that correspond to the red dot in graph (b) of Figure 6.3. The results of this analysis are depicted in Figure 6.5, where the top 10 contributing variables are displayed. It can be observed from graph (a) of Figure 6.5 that the CD contribution analysis only identified five of the eight process variables that are considered to be univariately different. The PD approach identified six variables and

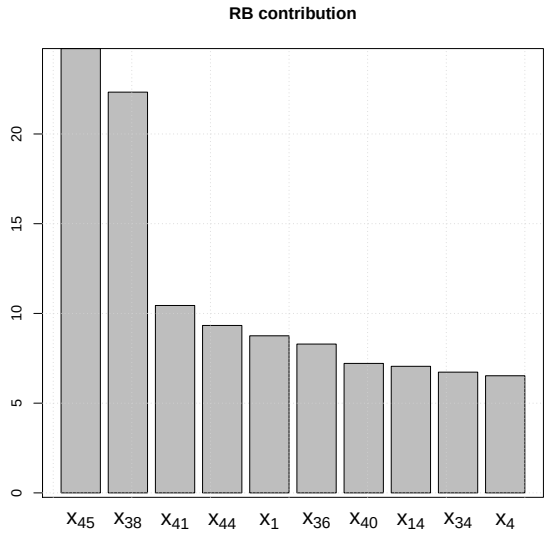
the RB analysis five.



(a) Complete decomposition.



(b) Partial decomposition.



(c) Reconstruction based decomposition.

Figure 6.5: Complete decomposition, Partial decomposition and Reconstruction based contribution analysis results for the red sample selected in graph (b) of Figure 6.3. Bar graphs are used to represent the SPE statistic based contribution of the top 10 ranked process variables for each method.

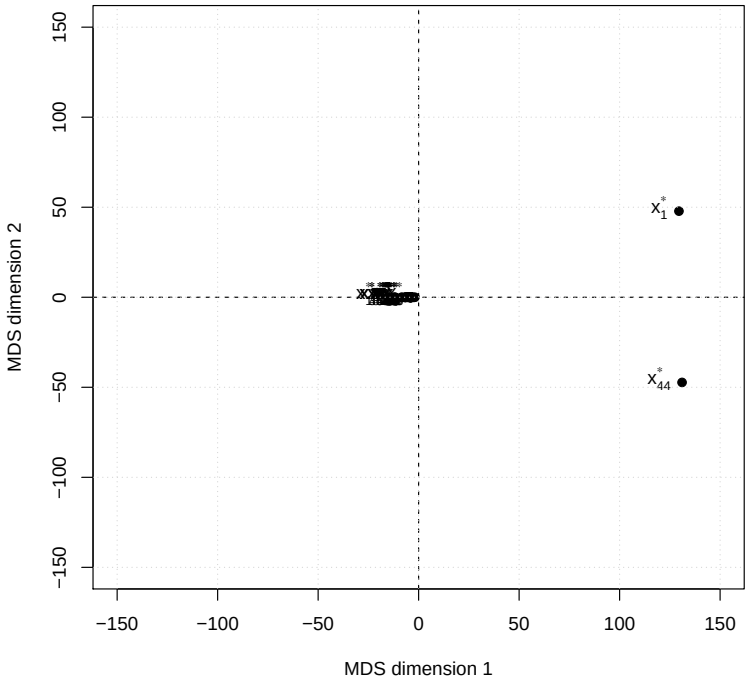


Figure 6.6: MDS display of the red SPE point selected in Figure 6.3. Each point on the graph represents a row or column of the SPE based distance matrix calculated for the point selected.

The new methodology developed in Chapter 3 will now be employed to diagnose the selected process fault. The corresponding R code that are listed in Appendix A follow the next general steps:

1. Decompose the SPE evaluation of the selected out-of-control sample into a sum of parts, as described in Chapter 3, and convert into the distance matrix representation.

2. Similarly obtain the distance matrix representation based on the SPE statistic for each of the IOC samples. This will define the reference distance matrices.

3. For each process variable collect the corresponding row in each of the reference distance matrices into a separate matrix.

4. Iteratively compare the row sums of the distance matrix calculated for the out-of-control point to that of the reference row sums, as was illustrated in Chapter 3. It was decided to label a specific row sum as abnormally large if it recorded a value that is larger than the 95^{th} percentile of the corresponding reference row sum distribution.

Chapter 3 also describes a graphical method based on MDS that can be used to visually diagnose process faults. This representation for the selected out-of-control sample is displayed in Figure 6.6. Interpretation of Figure 6.6 concludes that variables x_1 and x_{44} are

important to the interpretation of the large SPE value that was recorded. These variables were both identified to have univariate deviating measurements. It is clear from Figure 6.6 that the visual interpretation is not optimal to fully diagnose the fault. Therefore, the more precise row sum method was utilised to identify the statistically significant variables that define the large SPE value observed. The variables identified are,

$$\begin{bmatrix} x_{44} & x_1 & x_4 & x_{18} & x_{50} & x_{45} & x_{34} & x_{19} & x_{28} \end{bmatrix}, \tag{6.2}$$

ranked in the order of importance. The variables identified include all of the process variables that were identified to deviate univariately. One additional process variable, x_{28}, was also identified. In Figure 6.7 the scatterplot representation of x_{28} is given, similar to the graph presented for x_1 in Figure 6.4. Also included in Figure 6.7 is a horizontal broken firebrick red line that indicate the 99^{th} percentile of IOC data for x_{28}. It can be observed that the red solid dot is recorded beyond the 99^{th} percentile. The general upward movement of the green points in Figure 6.7 is also insightful. An argument can therefore be made from a statistical view point that the identification of x_{28} is sensible. In the research presented by Jiang et al. (2015) a fault diagnosis methodology based on Canonical variate analysis is introduced. During the validation of the approach Jiang et al. (2015) listed $x_1, x_4, x_{18}, x_{19}, x_{34}, x_{44}, x_{45}, x_{50}$ as variables that contribute to the first TE fault. These correspond to variables that are identified to have univariate deviations. The pairwise contribution approach therefore identifies x_{28} as an additional process variable that needs to be considered.

The variable pairs with significant pairwise contributions were also identified. However, in this setting the pairwise contribution analysis is not sensible, since each variable that deviate univariately will have a deviating pairwise relationship with all of the other process variables. See for example Figure 6.8. It is sensible that this relationship is identified as deviating. However, the deviation is observed due to both x_{18} and x_{19} recording a measurement outside the IOC univariate range.

Based on the results presented, it is concluded that the information provided by the pairwise diagnosis analysis can be used to supplement fundamental engineering understanding in order to assist with the practical interpretation of the multivariate deviations observed during the first TE process fault simulation. In this example the multivariate abnormality was identified due to univariate changes in the identified process variables. Note that in this analysis the univariate rankings obtained from the pairwise information captured more information regarding the multivariate deviation since the process moved beyond its univariate bounds.

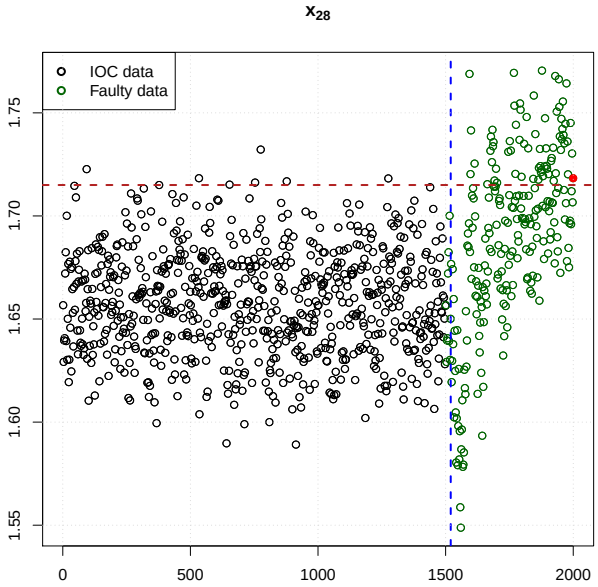


Figure 6.7: Measurements recorded for process variable x_{28} plotted in sequence. The black points represent the values captured during the IOC simulations. Green points are used to indicate the data measured during the Fault 1 simulation. The vertical broken blue line indicate to the point in time when the fault was introduced.

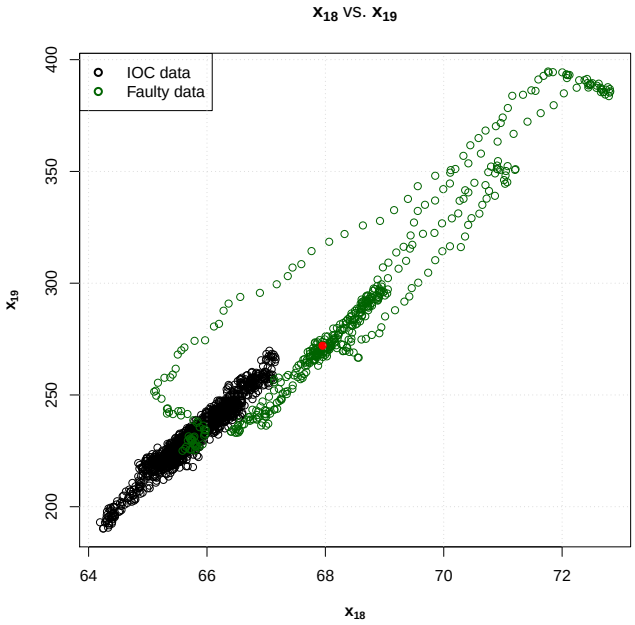
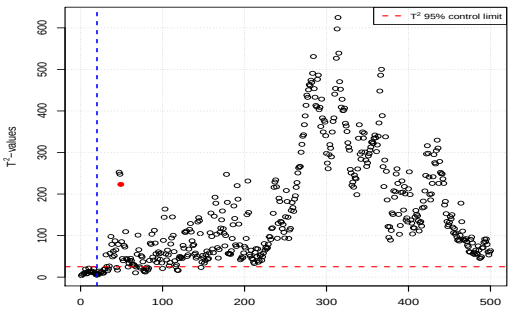


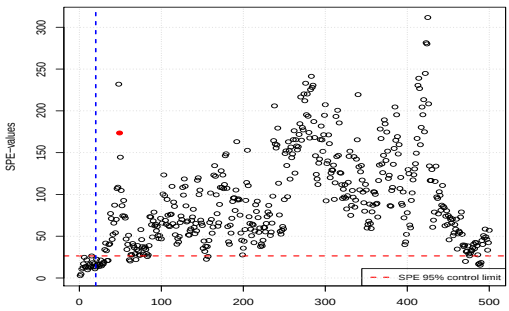
Figure 6.8: Scatter plot of x_{18} versus x_{19}. Black points are used to identify the samples recorded during the IOC simulations. Green points were observed during the simulation of Fault 1. The red point corresponds to the red point selected in Figure 6.3.

6.2.3.2 Diagnosis of TE Fault 12.

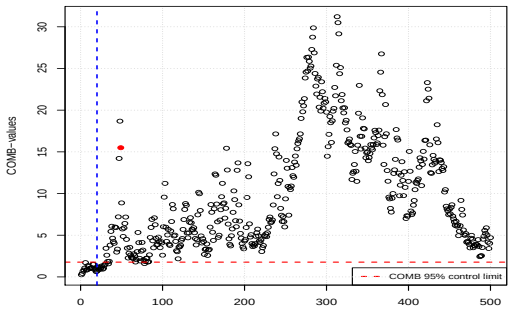
In this section the detection and diagnosis of the 12^{th} TE process fault in Table 6.4 will be considered. This will again be based on the PCA model that was established in the previous section. Therefore, the exact same steps discussed previously will be used to evaluate the data observed from the 12^{th} TE fault simulation. R code used in this illustration will again be presented in Appendix A. The calculated fault identification statistic values for the simulated data of the 12^{th} TE fault are presented in Figure 6.9. It can be observed from Figure 6.9 that all of the statistics were able to identify abnormal process behaviour soon after the fault was introduced. In this example we will employ the different diagnostic methods to interpret the process deviation identified by the red solid point in graph (c) of Figure 6.9. That is, the diagnosis will be based on the *Combined* statistic.



(a) T^2 statistic evaluations.



(b) SPE statistic evaluations.



(c) *Combined* statistic evaluations.

Figure 6.9: PCA fault identification statistic evaluations of the 12^{th} TE simulated fault. Evaluations are displayed for the T^2, SPE and $Comb$ statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.

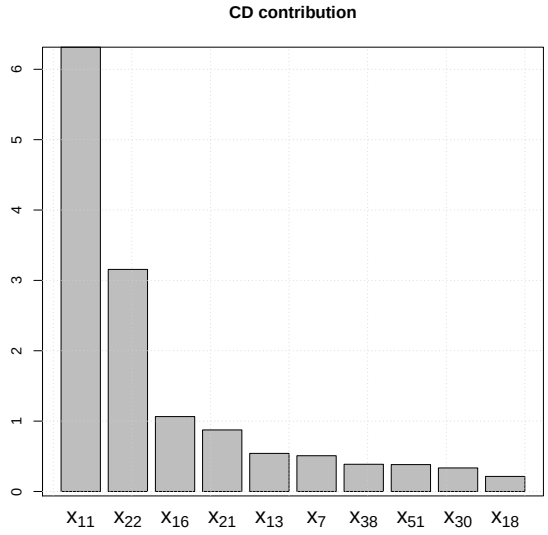
It is again observed, similar to the first TE process fault, that this fault was simulated by altering process variables such that measurements are recorded beyond the univariate ranges established during normal operation. Fault diagnosis can therefore be performed by identifying the variables that are univariately different from IOC's. The following variables are identified

$$\begin{bmatrix} x_7 & x_9 & x_{11} & x_{13} & x_{16} & x_{18} & x_{21} & x_{22} & x_{30} & x_{51} \end{bmatrix}. \tag{6.3}$$

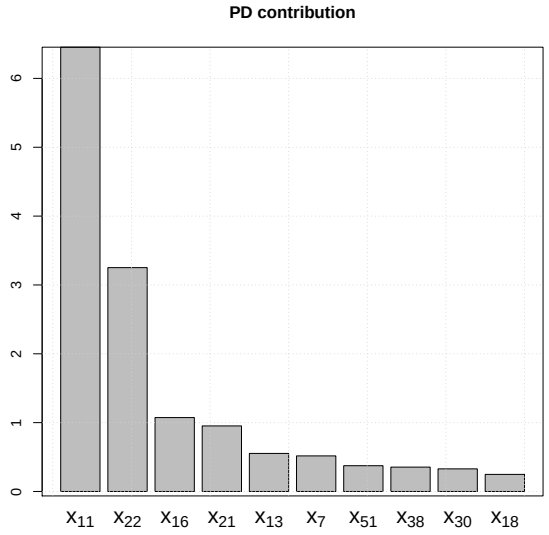
As with the analysis of Fault 1, it is required that the multivariate fault diagnosis analysis should at least identify these variables as having high contributions.
The diagnosis results based on the traditional methodologies are summarised in Figure 6.10 for the red dot in graph (c) of Figure 6.9. It is observed that most of the univariate deviating variables are listed in the top 10 contributions of the traditional techniques. Only process variable x_9 is not listed under the high contributing variables.
Univariate fault diagnosis based on the pairwise methodology introduced in Chapter 3 identified the following variables as having significantly high row sums:

$$\begin{bmatrix} x_{11} & x_{22} & x_{21} & x_{16} & x_{51} & x_{13} & x_9 & x_7 & x_{17} & x_{30} & x_{38} & x_{18} & x_{52} \end{bmatrix}. \tag{6.4}$$

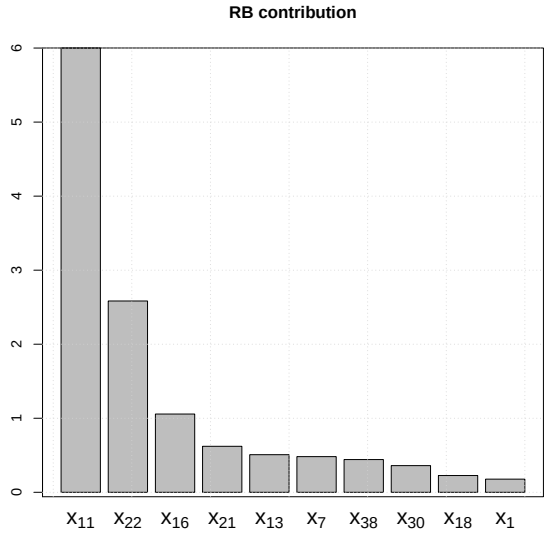
This set of variables include all of the variables which were identified to deviate univariately. Additionally, variables x_{17}, x_{38} and x_{52} are also identified. In Figure 6.11 it can be observed that these variables recorded measurements that were at the edge of the IOC range. The identification of these variables is therefore sensible. In Figure 6.12 the corresponding visual representation of the distance matrix is displayed. It is again difficult to interpret due to the large number of variables. The more exact approach of analysing the row sums is therefore preferred.
It is again concluded that the new pairwise fault diagnosis methodology is capable of explaining why the observation that was selected is identified as being multivariately out-of-control. In this example the multivariate abnormality was identified due to univariate changes in the identified process variables.



(a) Complete decomposition.



(b) Partial decomposition.



(c) Reconstruction based decomposition.

Figure 6.10: PCA Complete decomposition, Partial decomposition and Reconstruction based contribution analysis results for the red sample selected in graph (c) of Figure 6.9. Bar graphs are used to represent the *Combined* statistic based contribution of the top 10 ranking process variables for each method.

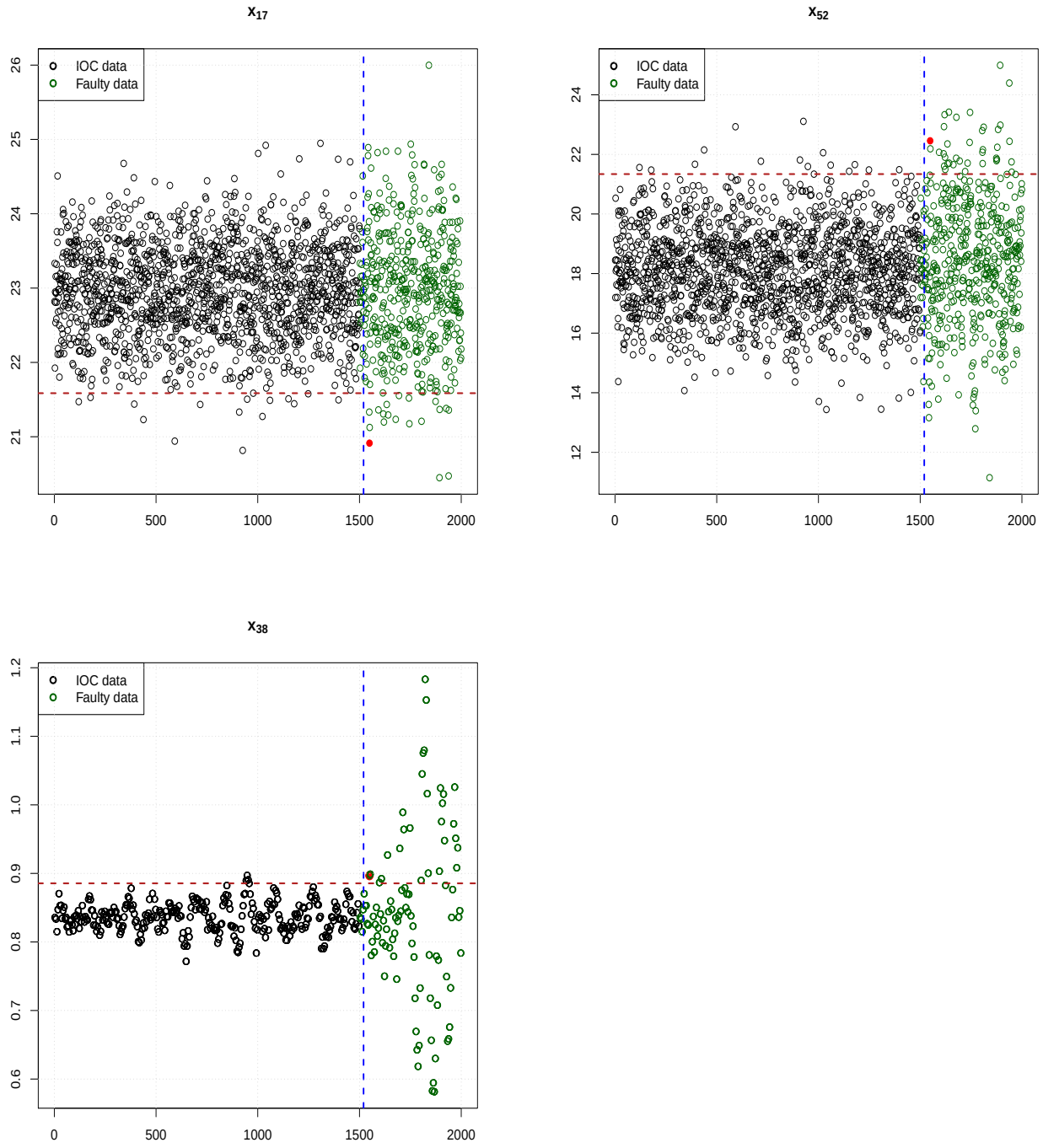


Figure 6.11: Measurements recorded for process variables x_{17}, x_{52} and x_{38} plotted in sequence. The red dot represent the measurement that correspond to the deviating *Combined* statistic observation that was selected in Figure 6.9. Horizontal broken red lines represent the 1^{st} or 99^{th} percentile of the IOC data.

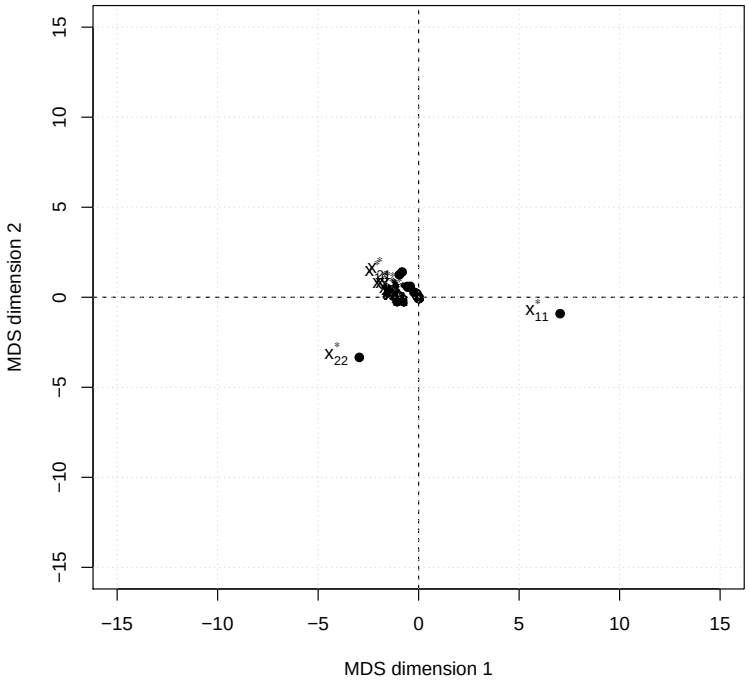


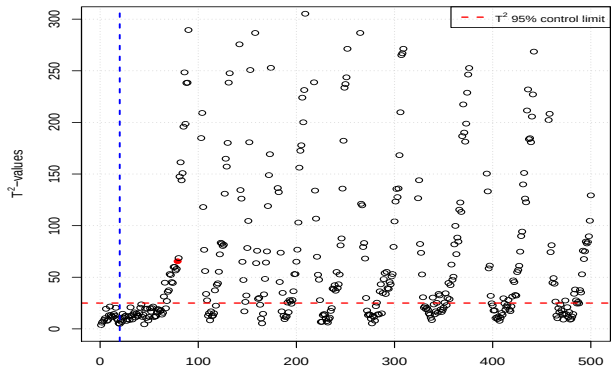
Figure 6.12: PCA *Combined* pairwise contribution MDS display of the red *Combined* statistic point selected in Figure 6.9.

6.2.3.3 Diagnosis of TE Fault 17.

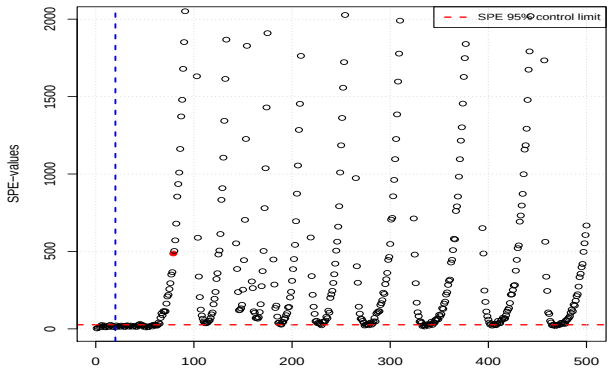
Fault 17 in Table 6.4 will be considered in the final example. Using the PCA model fitted in Section 6.2.3.1 to examine the simulated data recorded for Fault 17, resulted in the fault identification statistic evaluations presented in Figure 6.13. In this set of graphs it can be observed that the different fault identification statistics are capable in interpreting that the data observed during the Fault 17 simulation are foreign to the normal TE process. The out-of-control red point identified by the T^2 statistic in graph (a) of Figure 6.13 will be selected to demonstrate the diagnostic step. Again we observe, for the selected point, that the measurements of certain process variables are outside the ranges that were recorded during IOC's. These variables are:

$$\begin{bmatrix} x_{21} & x_{51} \end{bmatrix}. \tag{6.5}$$

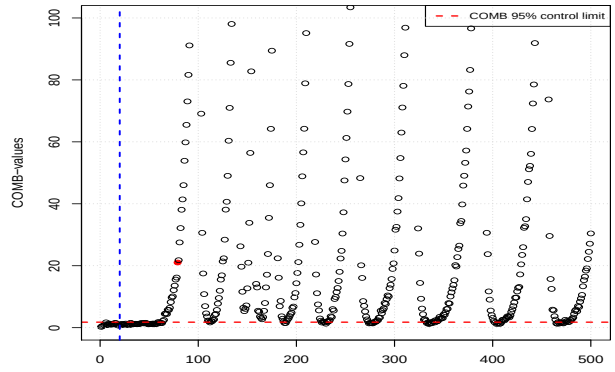
Therefore, a reasonable minimum requirement of the MSPM fault diagnosis would be that it must identify these variables as high contributors.
PCA T^2 contribution analysis results for the red point selected in the T^2 graph of Figure 6.13 are given in Figure 6.14. It is observed, for all of the bar graphs in Figure 6.14, that variables x_{21} and x_{51} are ranked in the top three contributors.



(a) T^2 statistic evaluations.



(b) SPE statistic evaluations.



(c) *Combined* statistic evaluations.

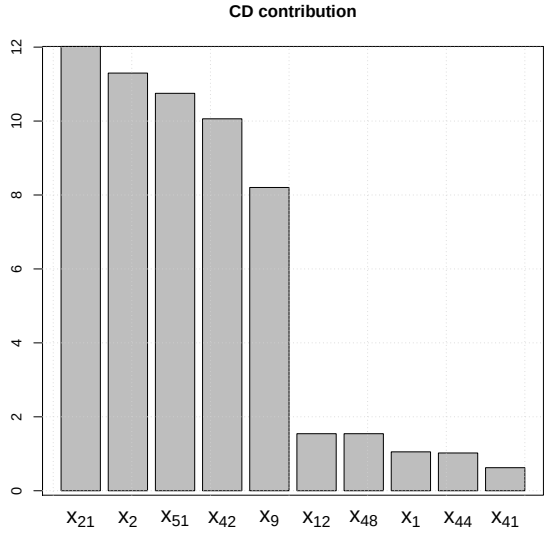
Figure 6.13: PCA fault identification statistic evaluations of the 17^{th} TE simulated fault. Evaluations are displayed for the T^2, SPE and $Comb$ statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.

Univariate contribution analysis based on the new distance based methodology, introduced in Chapter 3, identified the following variables as having significant contributions:

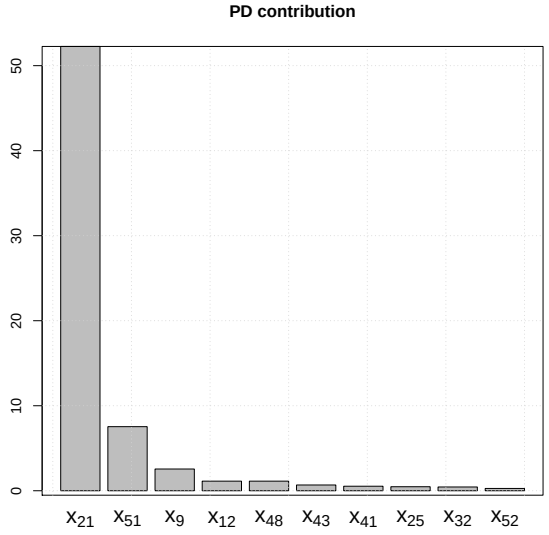
$$\begin{bmatrix} x_{21} & x_{51} \end{bmatrix}, \tag{6.6}$$

i.e., the variables that are univariately different from the IOC's ranges. The MDS graph of the T^2 distance matrix for the point selected in graph (a) of Figure 6.13 is displayed in Figure 6.15. This visual representation also identifies x_{21} and x_{51} as the leading variables that contribute to the faulty T^2 categorisation. It is again noted that the pairwise con-

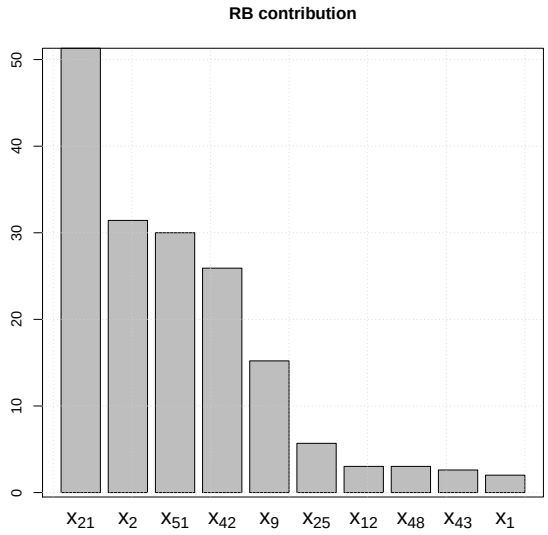
tribution analysis did not yield meaningful results for this example. All of the significant contributing pairs that were identified can be attributed to the univariate violations observed for x_{21} and x_{51}. The R code used in this section are presented in Appendix A.The results presented illustrate that the new methodology can be used to practically interpret the selected multivariate process deviation.



(a) Complete decomposition.



(b) Partial decomposition.



(c) Reconstruction based decomposition.

Figure 6.14: PCA T^2 contribution analysis results for the red sample selected in graph (a) of Figure 6.13.

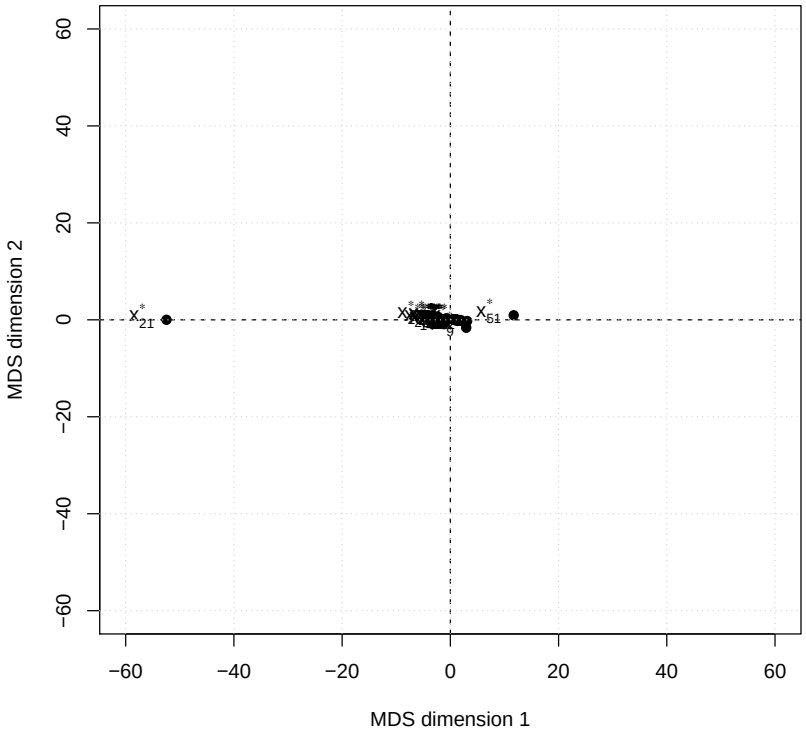


Figure 6.15: PCA T^2 pairwise contribution MDS display of the red T^2 point selected in Figure 6.13

6.2.4 Kernel PCA based fault diagnosis

In this section faults detected using kernel PCA will be considered for diagnosis purposes. Specifically, the pairwise contribution approach developed in Chapter 5 will be used to diagnose fault one, 12 and 17. Results based on the traditional diagnosis approaches of Chapter 4 will also be given.

6.2.4.1 Diagnosis of TE Fault 1

In this section the ability of kernel PCA based methods to detect and diagnose the first TE process fault will be evaluated. The R programming language will again be used to perform the analysis. Steps used to fit the kernel PCA model can be summarised as follows:

1. Selection of the IOC data samples from the 'fault_free_training' R data frame. Use only one of the simulation replicates i.e., the first 500 observations.

2. Zero mean centering and unit variance scaling of the IOC data set.

3. Define the double centered Gaussian kernel matrix and subsequent eigen-decomposition.

4. Use the crossvalidation based algorithm of Alam and Fukumizu (2014) explained in Algorithm 1 of Chapter 4 to select the optimal hyperparameter setting and the dimension of the kernel PCA model.

5. Calculate the matrices used to define the kernel PCA fault identification statistics.

6. Empirically determine the critical value for each of the fault identification statistics.

The algorithm of Alam and Fukumizu (2014) identified that the optimal dimension of the kernel PCA model should be 80 and that the Gaussian hyperparameter should be set to 0.003.
The fault identification statistic evaluations of the new faulty simulated data was obtained as follows:

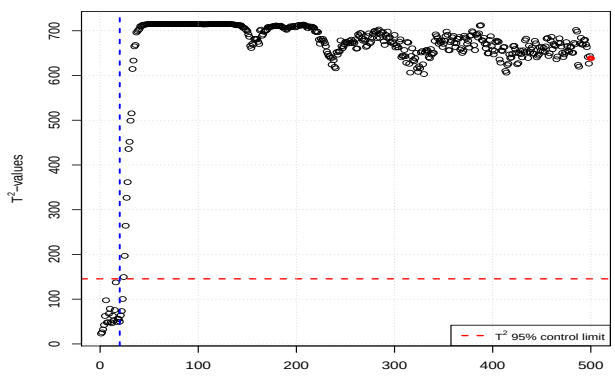
1. Identify the data that correspond to the first simulated process fault in the 'faulty_training' data set. Select only one of the replicated simulations i.e., the first 500 observations.

2. Center and scale the selected faulty data using the mean and variance information of the IOC data.

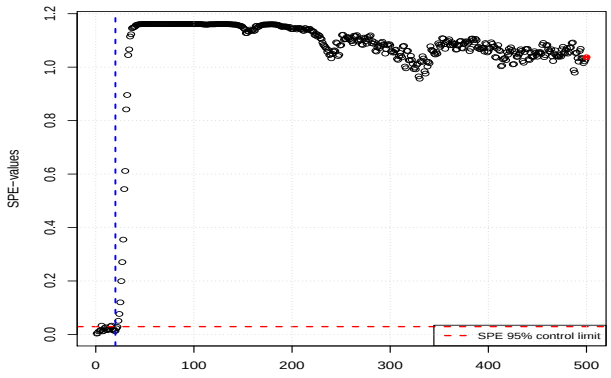
3. Calculate the PCA fault identification statistic evaluations i.e., T^2, SPE and $Combined$ values.

The fault identification statistic evaluations of the simulated data for the first TE process fault are displayed in Figure 6.16. The 95% control limit for each of the respective statistics is represented using a horizontal red broken line. If measurements are recorded above this line then it indicates that the process is statistically different from IOC's. It can be observed from Figure 6.16 that the data observed for the first simulated fault are mostly above the control limit, for all of the three fault identification indices. The blue vertical broken line in Figure 6.16 indicates the one hour point at which the fault was introduced in the simulation. It is observed from Figure 6.16 that all of the statistics displayed out-of-control evaluations directly after the fault was introduced in the simulation. Similar to the PCA results, it is observed that the SPE statistic was the quickest to identify the faulty condition. In this illustration the fault diagnosis analysis will be demonstrated using the solid red point in graph (b) of Figure 6.16. This is the exact same point that was used in the diagnosis based on the PCA model. The results presented in this section can therefore be compared to the PCA based diagnosis results of the first TE process fault. It was established in Section 6.2.3.1 that a number of process variables violate their univariate bounds for the first TE fault. Therefore, it is again expected that the diagnosis results

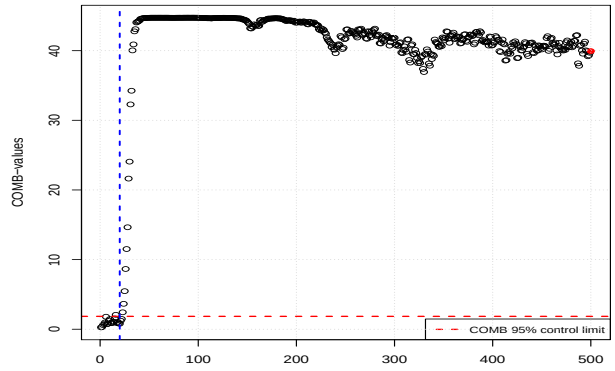
should at least identify these variables.



(a) T^2 statistic evaluations.



(b) SPE statistic evaluations.

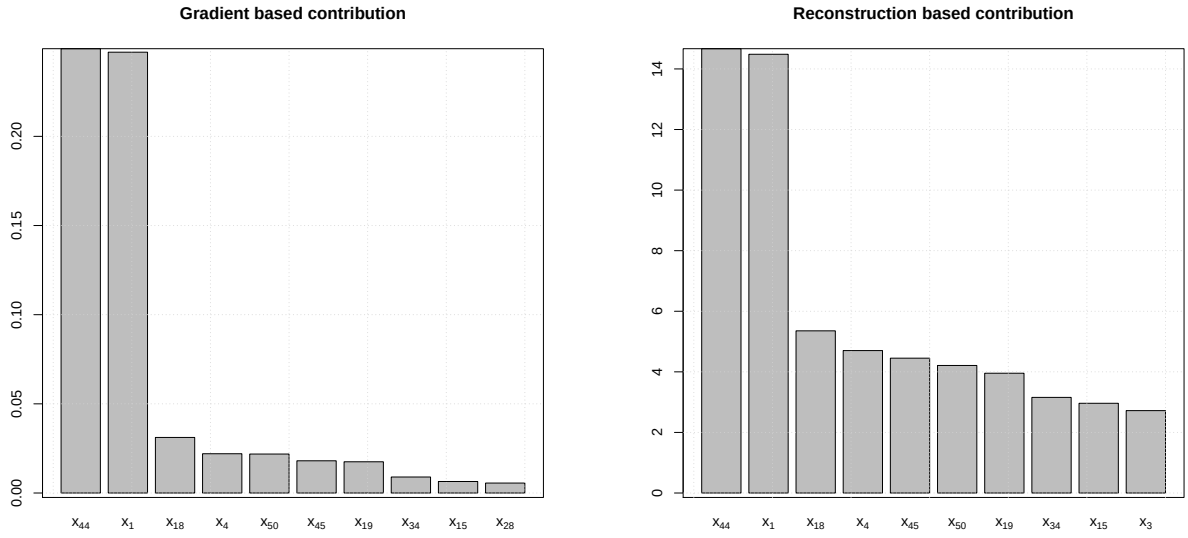


(c) *Combined* statistic evaluations.

Figure 6.16: Kernel PCA fault identification statistic evaluations of the TE Fault 1 simulation data. Evaluations are displayed for the T^2, SPE and *Combined* statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.

First the gradient based approach of Cho et al. (2005) and the reconstruction based approach of Alcala and Qin (2010) will be used to diagnose the solid red observation in the SPE graph of Figure 6.16. Therefore, the diagnosis will be based on the SPE statistic. Figure 6.17 display the results of the analysis. For each technique the contribution of the top ten ranked variables are displayed. It is observed from Figure 6.17 that for both techniques the relative importance of x_{44} and x_1 greatly dwarf the contributions of the remaining process variables. However, it is observed that both techniques identified the univariate violating variables when all of the top ten ranked process variables are

considered.



(a) Gradient based. (b) Reconstruction based.

Figure 6.17: Kernel PCA Gradient based (Cho et al., 2005) and Reconstruction based (Alcala and Qin, 2010) contribution analysis results of the red sample selected in graph (b) of Figure 6.16. Bar graphs are used to represent the SPE statistic based contribution of the top 10 ranking process variables for each method.

The pairwise distance methodology described in Chapter 5 will now be employed to diagnose the selected red point in graph (b) of Figure 6.16. Note that the analysis will focus mostly on the univariate rankings that are obtained from the pairwise contributions since the data have univariate violations. Before the analysis can be performed it is required that the pairwise models be constructed that will be used to quantify the distances. It was decided to use the SVM approach to specify the pairwise distances. In this analysis there are 50 process variables which implies that $\binom{50}{2} = 1225$ SVM models had to be constructed. Note that two process variables, x_{48} and x_{49}, was removed from the analysis due them having 100% correlation with two other variables. This once-off model construction was completed in 98 minutes using parallel computing on a Windows 10 computer with 32 logical CPU cores and 32 gigabyte RAM.
The following steps were followed to identify the univariate ranking based on the pairwise information:

- Use the pairwise SVM models to construct the distance matrix representation of the out-of-control sample as described in Chapter 5.
- Iteratively compare the row sums of the distance matrix calculated for the out-of-control point to that of the reference row sums, as was discussed in Chapter 5.

The distance matrix representation of the out-of-control observation can be used to obtain a MDS graphical representation that can be used to visually diagnose the process fault. This representation for the selected out-of-control point is displayed in Figure 6.18. Visible from Figure 6.18 is that the following samples,

$$\begin{bmatrix} x_1 & x_{44} & x_{45} & x_{34} & x_4 & x_{18} & x_{19} & x_{50} & x_{28} \end{bmatrix}, \tag{6.7}$$

are remote from the center of the display. These variables are the ones that were identified to violate the univariate bounds. Additionally x_{28} is also identified as important. Therefore, the SVM based MDS analysis yield results that are similar to the PCA based pairwise approach in Section 6.2.3.1.

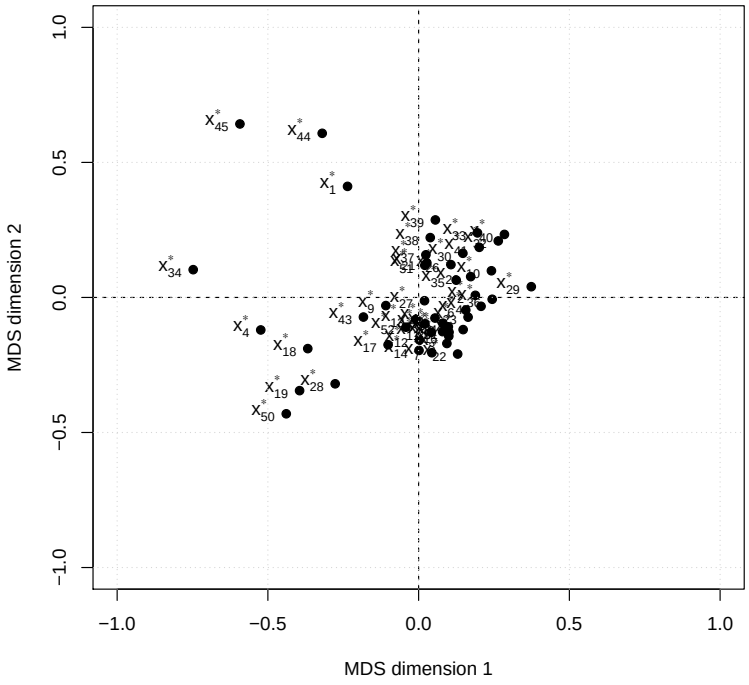


Figure 6.18: MDS display of the red SPE point selected in Figure 6.16. Each point on graph repesent a row or column of the SVM based distance matrix calculated for the point selected.

Note that the row sums approach was simplified by setting pairwise distances that are less than 0.5 equal to 0. This is possible since the distances calculated using SVM are probabilities. Therefore, it was possible to determine the ranking of process variables using only the maximum row sum during each iteration. The following univariate ranking

is obtained,

$$\begin{bmatrix} x_1 & x_{44} & x_{45} & x_{34} & x_{19} & x_{50} & x_4 & x_{28} & x_{18} & x_{10} & x_{47} \end{bmatrix}, \tag{6.8}$$

based on the SVM based pairwise distances. These include all of the variables that are univariately different. Additionally, x_{28}, x_{10} and x_{47} are also identified. It was already explained in Section 6.2.3.1 why it is sensible to include x_{28}. In Figure 6.19 graphs, similar to that for x_{28} in Figure 6.7, are depicted for x_{10} and x_{47}. The horizontal broken red line in Figure 6.19 are used to indicate the 5^{th} percentile of the IOC data. Therefore, it is observed from Figure 6.19 for both variables that the red point, corresponding to the out-of-control point, is recorded close to the 5^{th} percentile. Additionally it is observed that the green points close to the red point visually indicate a shift in the mean. Therefore, it is concluded that the identification of x_{10} and x_{47} by the SVM based pairwise distance method is sensible.

The results obtained using the SVM based pairwise distance method identified process variables that can be used to interpret the multivariate process deviation. All of the variables that were identified are measured beyond or close to the IOC process limits. It is also noted that the non-linear pairwise approach identified additional process variables i.e., x_{10} and x_{47}, that were not identified by the PCA based pairwise approach.

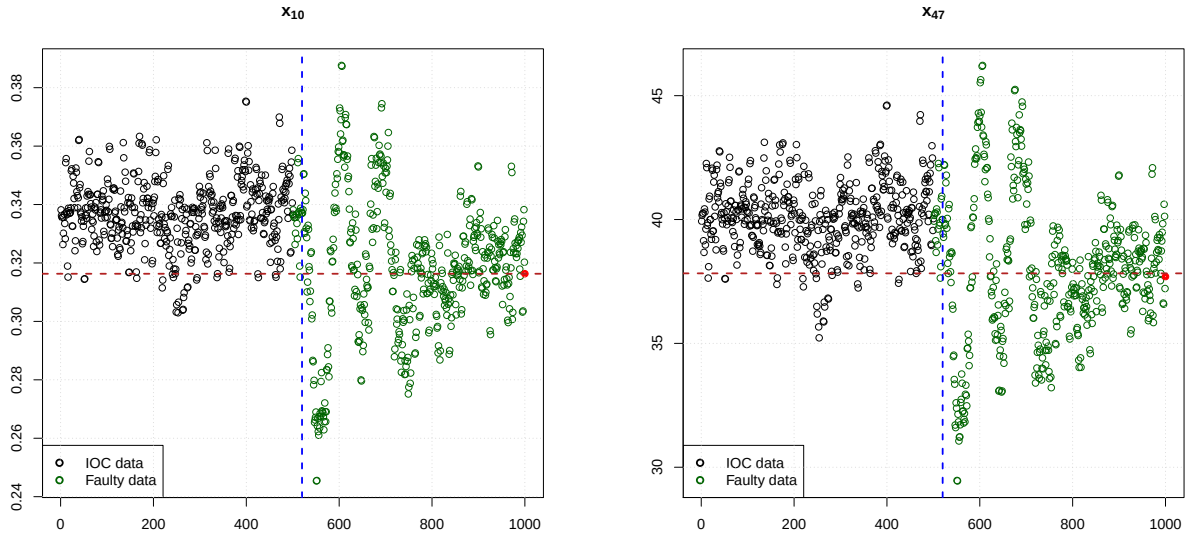


Figure 6.19: Measurements recorded for process variable x_{10} and x_{47} plotted in sequence. The black points represent the values captured during the IOC simulation. Green points are used to indicate the data measured during the Fault 1 simulation. The vertical broken blue line indicate to point in time when the fault was introduced. The red dot represent the measurement that correspond to the deviating SPE that were selected in Figure 6.16. Horizontal broken red lines represent the 5^{th} percentile of the IOC data.

6.2.4.2 Diagnosis of TE Fault 12.

The detection and diagnosis of the 12^{th} TE process fault will be considered in this section. The kernel PCA model developed in the previous section will again form the basis of the analysis. Therefore, the same detection and diagnosis steps will be followed to evaluate the data observed from the 12^{th} TE fault simulation. Also the results obtained in this section will be comparable to the PCA based analysis in Section 6.2.3.2.

The kernel PCA based fault detection statistic values are presented in Figure 6.20. It is observed from Figure 6.20 that the abnormal process behaviour is identified by all of the statistics soon after the fault was introduced. In this example, the different diagnostic methods will be employed to interpret the process deviation identified by the red solid point in graph (c) of Figure 6.20. That is, the diagnosis will be based on the *Combined* statistic.

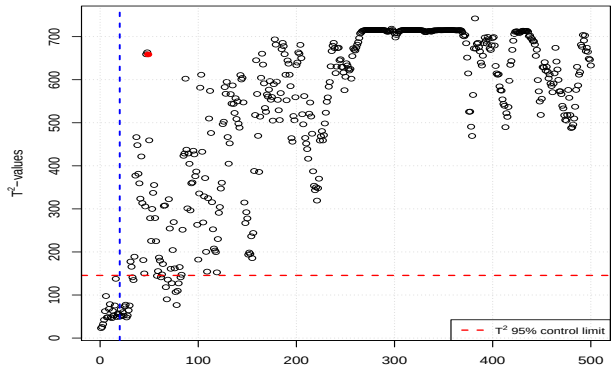
The results based on the traditional kernel PCA based diagnosis methods are summarised in Figure 6.21. It is observed that most of the univariate deviating variables are ranked amongst the top ten contributors of the traditional techniques. Only process variable x_{30} is not listed under the high contributing variables.

The univariate variable ranking based on the pairwise row sum methodology introduced in Chapter 5 is as follows:

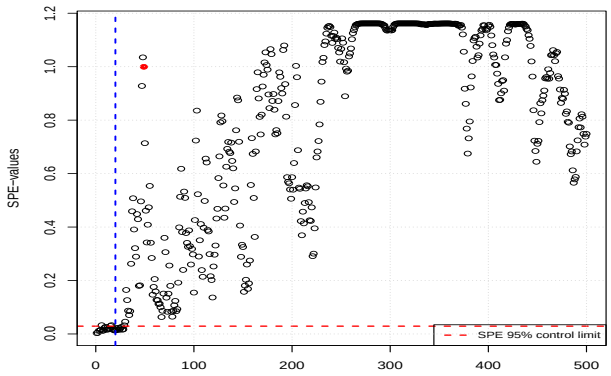
$$\begin{bmatrix} x_{17} & x_{52} & x_9 & x_{11} & x_{21} & x_{35} & x_{22} & x_{30} & x_{18} & x_{16} & x_{51} & x_7 & x_{38} & x_{13} & x_{24} & x_{14} \end{bmatrix}. \quad (6.9)$$

This set of variables include all of the variables that were identified to deviate univariately. The process variables x_{17}, x_{52}, x_{35}, x_{38}, x_{24} and x_{14} are also identified. It has already been indicated in Section 6.2.3.2 why the identification of x_{17}, x_{52} and x_{38} is sensible. In Figure 6.23 representations of x_{35}, x_{24} and x_{14} are given to motivate why the identification of these variables is sensible. In Figure 6.22 the MDS visual representation of the SVM based out-of-control distance matrix is displayed. It is observed in Figure 6.22 that all but two of the variables, x_{24} and x_{14}, identified by the row sum method are indicated by the visual approximation. The more exact approach of analysing the row sums is preferred.

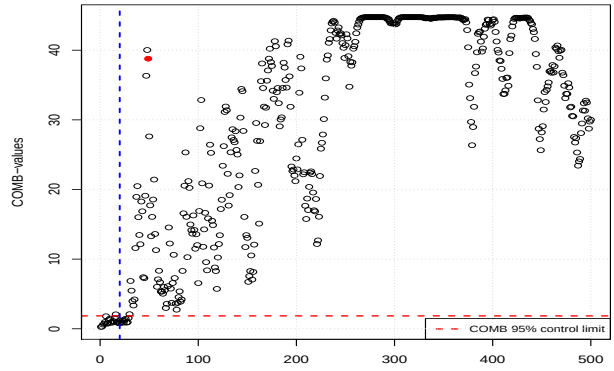
It is again observed that the univariate rankings obtained from the pairwise distance information correspond to variables that deviated univariately or recorded values that are close to the edge of the IOC range.



(a) T^2 statistic evaluations.

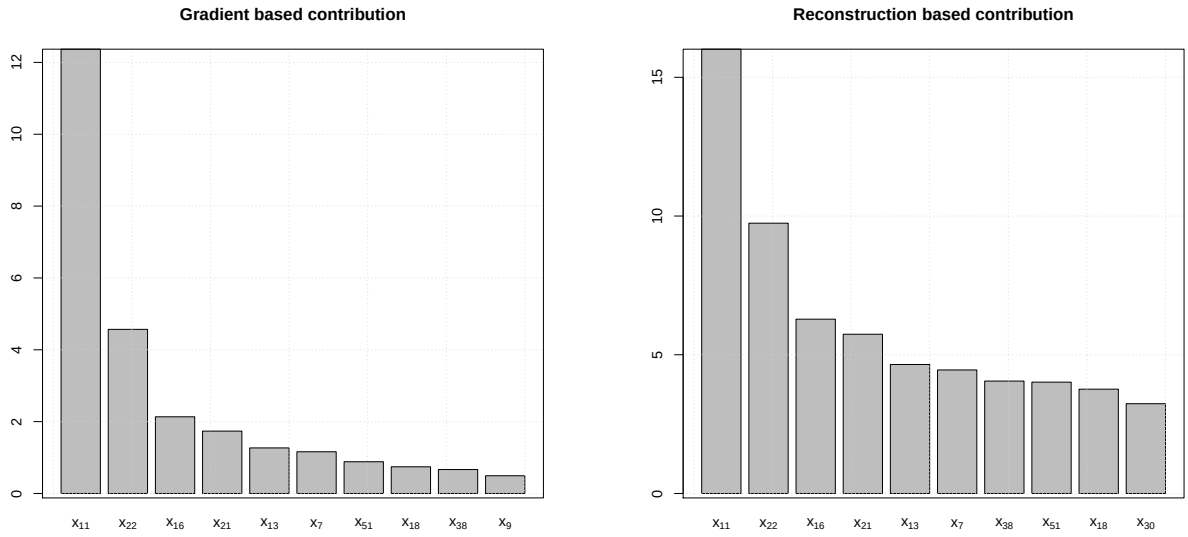


(b) SPE statistic evaluations.



(c) *Combined* statistic evaluations.

Figure 6.20: Kernel PCA fault identification statistic evaluations of the TE Fault 12 simulation data. Evaluations are displayed for the T^2, SPE and *Combined* statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.



(a) Gradient based. (b) Reconstruction based.

Figure 6.21: Kernel PCA Gradient (Cho et al., 2005) and Reconstruction based (Alcala and Qin, 2010) contribution analysis results of the red sample selected in graph (b) of Figure 6.20. Bar graphs are used to represent the *Combined* statistic based contribution of the top 10 ranking process variables for each method.

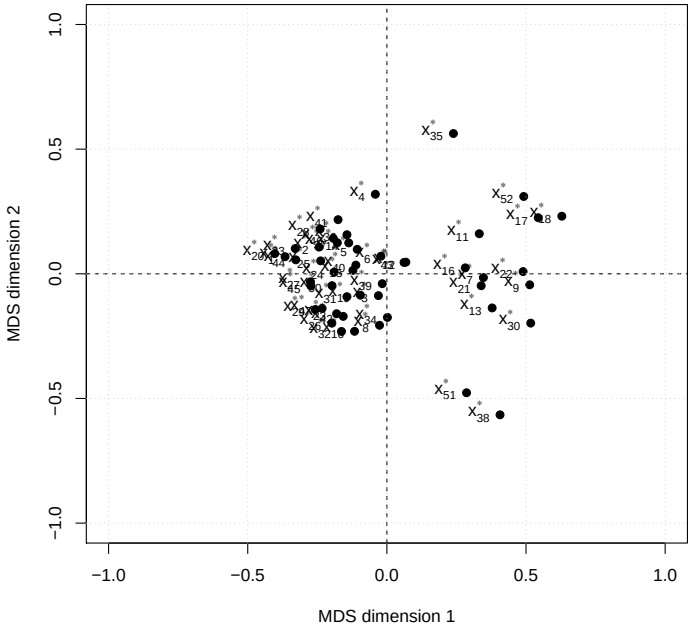


Figure 6.22: MDS display of the red *Combined* statistic point selected in Figure 6.20. Each point on the graph represent a row or column of the SVM based distance matrix calculated for the point selected.

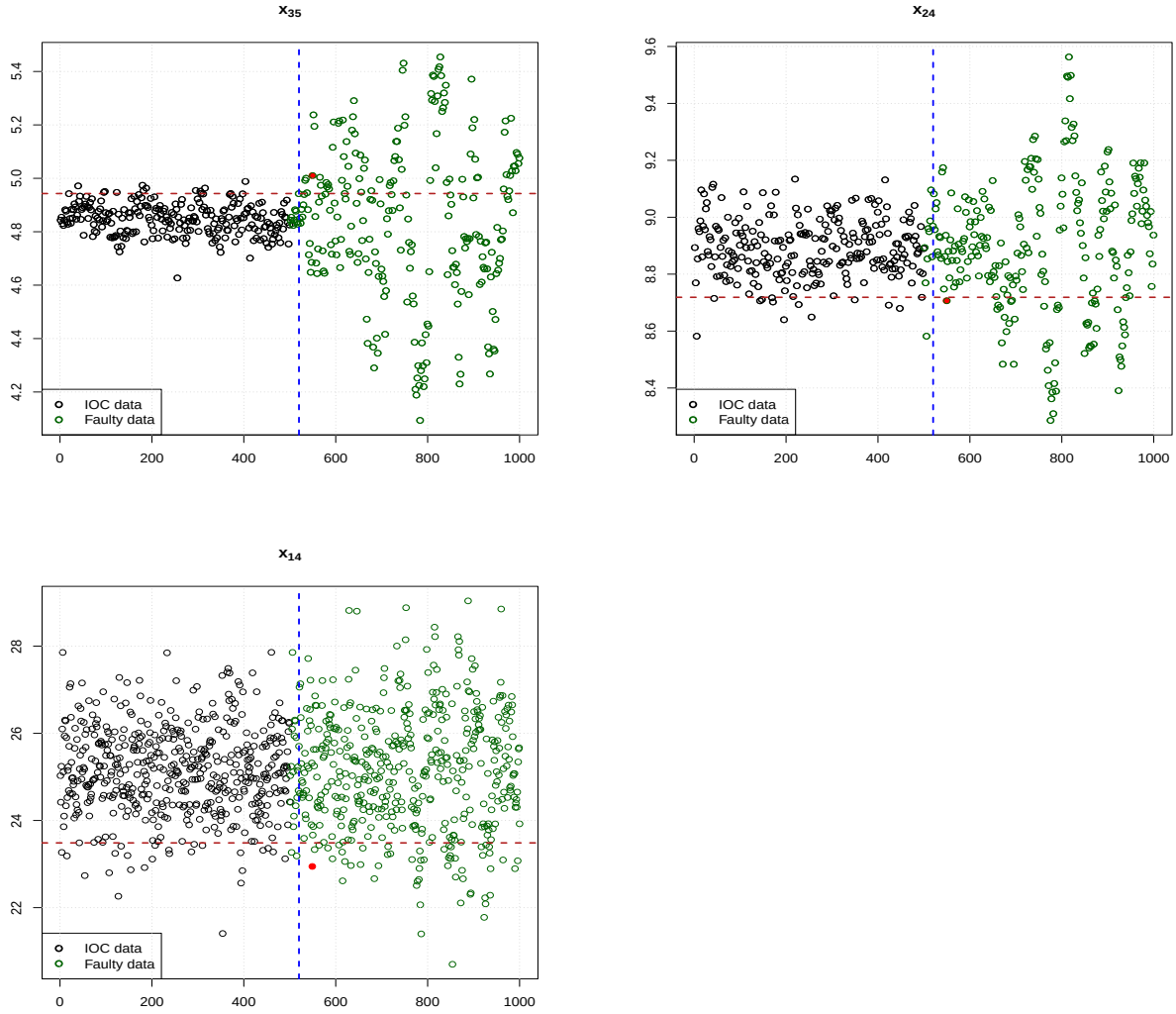
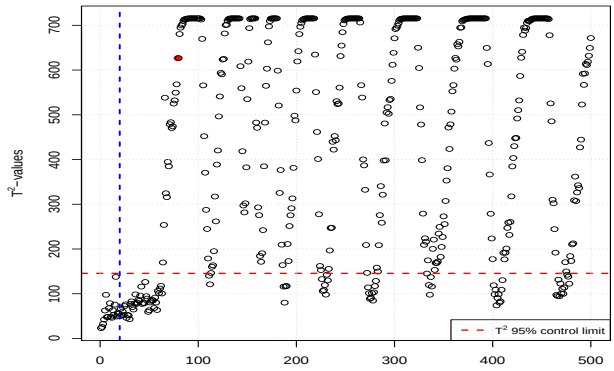


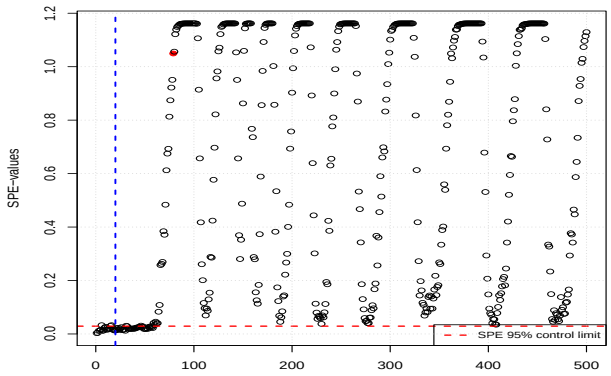
Figure 6.23: Measurements recorded for process variables x_{35}, x_{24} and x_{14} plotted in sequence. The red dot represent the measurement that correspond to the deviating *Combined* statistic observation that was selected in Figure 6.20. Horizontal broken red lines represent the 5^{th} or 95^{th} percentile of the IOC data.

6.2.4.3 Diagnosis of TE Fault 17.

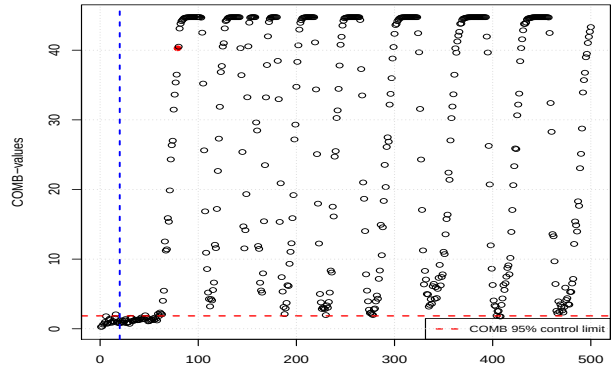
In the final example Fault 17 in Table 6.4 will be considered. Using the kernel PCA model fitted in Section 6.2.4.1 to evaluate the simulation data of the 17^{th} TE process fault yielded the fault statistic evaluations presented in Figure 6.24. It can be observed from Figure 6.24 that the different fault identification indices are able to recognise that the data recorded during the 17^{th} fault simulation are foreign to the normal TE process. The out-of-control red point identified by the T^2 statistic in graph (a) of Figure 6.24 will be selected to illustrate the diagnosis.



(a) T^2 statistic evaluations.



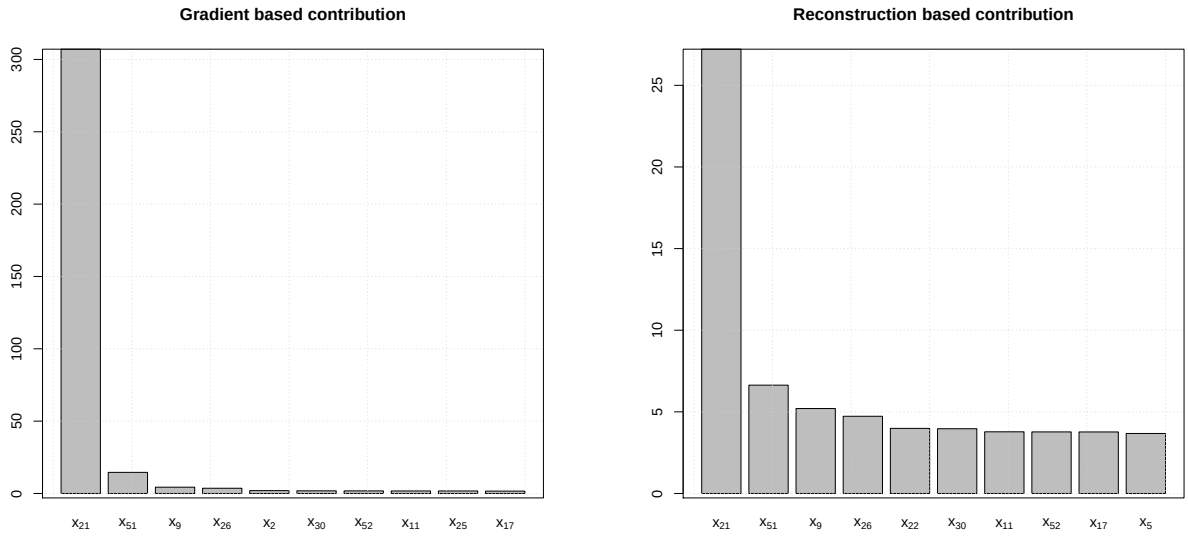
(b) SPE statistic evaluations.



(c) *Combined* statistic evaluations.

Figure 6.24: Kernel PCA fault identification statistic evaluations of the TE Fault 17 simulation data. Evaluations are displayed for the T^2, SPE and *Combined* statistics. Hollow black dots are used to indicate the evaluations at specific time points. The red dots are specially selected points that will be used for further investigation. Broken horizontal red lines identify the 95% control limits. Vertical broken blue lines indicate the point after one hour when the specific fault was introduced in the process.

The traditional gradient and reconstruction based analysis results, for the selected out-of-control observation, are given in Figure 6.25. It is observed for both methods that the univariate deviating variables, x_{21} and x_{51}, are ranked as the top two contributors.
The univariate ranking obtained from the pairwise SVM based distance matrix identified x_{21} and x_{51} as the only meaningful contributors. The corresponding visual MDS representation is given in Figure 6.26. This graphical display also identifies x_{21} and x_{51} as the leading variables that contribute to the faulty T^2 categorisation.
All of the non-linear MSPM diagnosis techniques therefore identified that the univariate deviating variables are key to the interpretation of the multivariate deviation detected by the T^2 statistic during the simulation of the 17^{th} TE process fault.



(a) Gradient based. (b) Reconstruction based.

Figure 6.25: Kernel PCA Gradient (Cho et al., 2005) and Reconstruction based (Alcala and Qin, 2010) contribution analysis results of the red sample selected in graph (b) of Figure 6.20. Bar graphs are used to represent the SPE statistic based contribution of the top 10 ranking process variables for each method.

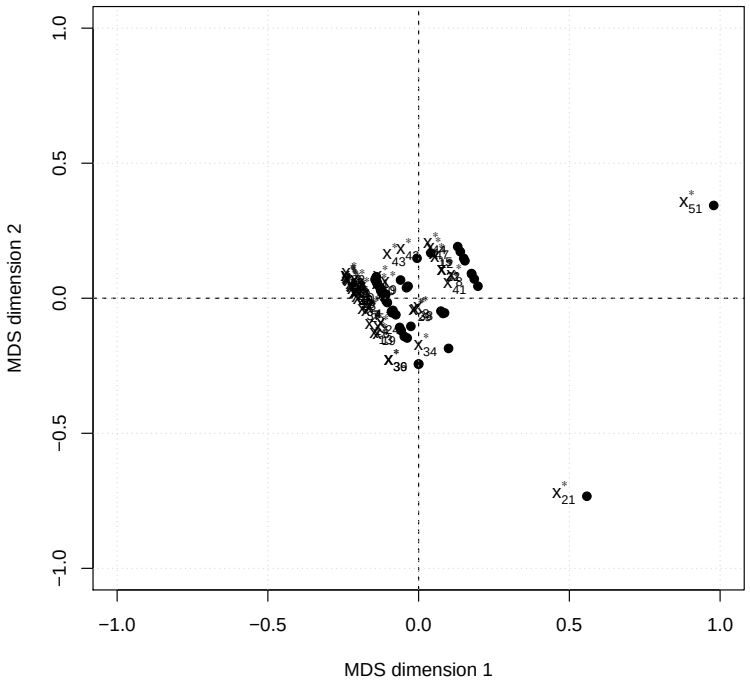


Figure 6.26: MDS display of the red SPE point selected in Figure 6.20. Each point on graph repesent a row or column of the SVM based distance matrix calculated for the point selected.

6.3 Case study: Coal Fired Steam Boiler

6.3.1 Background and Objective

Sasol South Africa is an energy and chemicals company that operates a large coal to liquids factory in the Mpumalanga province in the east of South Africa. A large array of products is produced from natural feedstock, such as coal and natural gas, using the Sasol proprietary Fischer-Tropsch related processes. The factory consists of a number of complex processing plants, each performing an important task in the complete product value chain. One of these processing plants will be the focus of this case study.
The steam processing plant is responsible for producing high pressured steam which is required for the operation of the other downstream units in the value chain. Common uses of steam are, for example,

- to assist in temperature regulating,
- to provide a driving force e.g., to drive turbines in electricity generation,
- to move product e.g., propelling liquid in pipes and
- to clean equipment.

The role of the steam plant is therefore important in the operation of the coal to liquids facility. Therefore, this plant needs to operate optimally, since other processes are reliant on the steam produced. The coal fired boiler is the primary steam generating equipment in the steam plant. Inside the boiler, heat energy is applied to water to produce steam. Specifically, coal is ground to a small particle size and mixed with air. This coal and air mixture defines the fuel which is burned in the combustion zone of the boiler to generate heat. A large network of tubes is used to transport water inside the boiler. The heat energy generated during combustion is used to heat up the boiler tubes. Consequently the fluids inside the tubes are heated and converted to steam.
The Sasol steam station consists of a total of 16 boilers, grouped into two sets of eight. Each set of boilers provides steam to the eastern and western side of the factory respectively. It is important that each of these boilers is online and operate optimally i.e., to ensure that steam is reliably produced to meet the demands of the downstream processes. Significant amounts of money are lost due to unplanned boiler downtime. A common reason for unplanned boiler downtime is the failure of boiler tubes. Early detection of process movement from ideal operating region will assist in planning downtimes better. In this study multivariate statistical process monitoring will be applied to one of the boilers

to detect and diagnose performance deviation from ideal boiler operation.

The objective of considering the boiler example is to demonstrate, using actual process data, how the various diagnosis techniques can be used to assist in interpreting multivariate process deviations. This case study will mimic a real life situation where the statistical information on the process deviation and corresponding diagnosis results are shared with the engineer. The quality of the diagnosis information will be judged on its ability to communicate why the statistical methods flagged the process to be multivariately different from IOC common cause variability.
The boiler case study will be analysed using both the PCA and kernel PCA MSPM techniques. Therefore, the diagnosis ability of the linear and non-linear pairwise methodologies will be evaluated. Note that the focus will be solely on the interpretation value of the diagnosis results. The objective is not to compare results of the two approaches i.e., PCA versus kernel PCA.

6.3.2 Data

For the current application six key boiler process characteristics were identified to be monitored in order to ensure the desired operating conditions. Information on the selected process variables are captured in Table 6.5. Note that the process variables are referred to in general terms, due to intellectual property constraints. Process engineers assisted in identifying a historical data set that can be used to characterise IOC's. The acceptable univariate ranges for each of the six process variables were also provided. These ranges are referred to as the operation within agreed parameter (OWAP) limits.
The specific approach of defining the IOC data set was:

1. Identify a boiler which had the least number of failures over its lifetime.
2. Collect ten minute aggregated data on the identified process variables for this boiler over a six year period.
3. Filter out data that fall outside the univariate OWAP bounds for each variable. More specifically, if, for a specific timestamp, one or more process conditions violate their operational ranges then remove the data on all variables for the specific timestamp.
4. Omit data on all process variables collected in the three months prior to a boiler failure. It is assumed that a boiler will start to exhibit unhealthy characteristics three months prior to a tube failure.

Therefore, this IOC data set represents the reference operating region against which all boilers can be compared for multivariate statistical process monitoring.

Table 6.5: Boiler process variable information.

Process Variable	Description
x_1	Gas flow control.
x_2	Feed pressure measurement.
x_3	Outlet temperature measurement.
x_4	Feed preheat temperature.
x_5	Feed flow measurement.
x_6	Gas feed flow.

6.3.3 PCA MSPM

Boiler operating fundamentals indicated that the multivariate statistical process monitoring (MSPM) methodology is required to adequately monitor process behaviour. The motivation for a MSPM analysis can also be made by observing the complicated correlation structure that exists amongst the process variables. The correlation matrix is given in (6.10). For example, it can be observed from (6.10) that variables x_1 and x_2 are highly correlated. Figure 6.27 displays a scatter plot of the historic IOC data for x_1 and x_2. A visually strong relationship exists between x_1 and x_2. Also included in Figure 6.27 are the univariate OWAP control limits for each process characteristic. Example multivariate violations can therefore occur if measurements on x_1 and x_2 result in observations that are located in the white space region of the inner most square of Figure 6.27. The relationship in Figure 6.27 is one of many relationships that need to be adhered to for the process to be multivariately in control. These relationships are visually captured by the IOC's data scatter plot matrix in Figure 6.28 for all process variables in Table 6.5.

$$\begin{bmatrix} 1.00 & -0.94 & 0.68 & 0.52 & 0.58 & 0.56 \\ -0.94 & 1.00 & -0.74 & -0.59 & -0.58 & -0.64 \\ 0.68 & -0.74 & 1.00 & 0.88 & 0.32 & 0.68 \\ 0.52 & -0.59 & 0.88 & 1.00 & 0.15 & 0.56 \\ 0.58 & -0.58 & 0.32 & 0.15 & 1.00 & 0.37 \\ 0.56 & -0.64 & 0.68 & 0.56 & 0.37 & 1.00 \end{bmatrix} \tag{6.10}$$

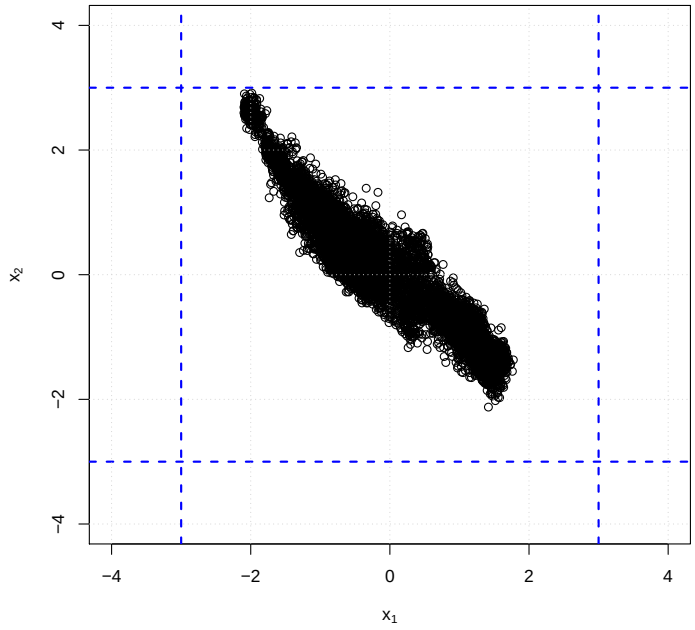


Figure 6.27: Scatter plot of the IOC data for boiler process variables x_1 and x_2. Dotted lines are OWAP limits. (x_1: Gas flow control, x_2: Feed pressure measurement)

Considering the multivariate nature of the problem and the correlation structure of the process variables, it was decided that PCA is suitable for the statistical monitoring of boiler behaviour. The general steps followed to fit the PCA model were:

1. Zero mean centering and unit variance scaling of IOC data.
2. Eigendecomposition of covariance matrix.
3. PCA dimension selection using the scree plot approach.
4. Definition of fault detection matrices and critical values.

The PCA model that was fitted contains information on the first 4 principal components and explains 97.8% of the variability. The boiler selected in this case study for performance monitoring is different from the one used to define the IOC data set. It is acceptable to compare the data of one boiler to the IOC data, because the boilers are considered to be identical. Process data, aggregated to ten minute means, were collected over a five month period from the plant historian for the specific boiler being monitored. This new data set was cleaned by removing observations that violated the univariate bounds specified by the OWAP limits. The resulting univariate acceptable behaviour for the process variables is depicted by the graphs in Figure 6.29. The horizontal broken blue lines in Figure 6.29 represent the univariate IOC limits. It is therefore visually clear from Figure 6.29 that

the measured process conditions mostly fall well inside the IOC univariate bounds. Figure 6.29 represents an example of a typical dashboard used by engineers to monitor boiler operation. This representation of measurements from the process will therefore not alert engineers to multivariate process deviations.

It should be reiterated that the objective of this study is to evaluate the new methodology developed for the diagnosis of multivariate process faults identified using PCA based MSPM. That is, only observations that were identified to have been observed as a result of multivariate assignable cause variability will be diagnosed.

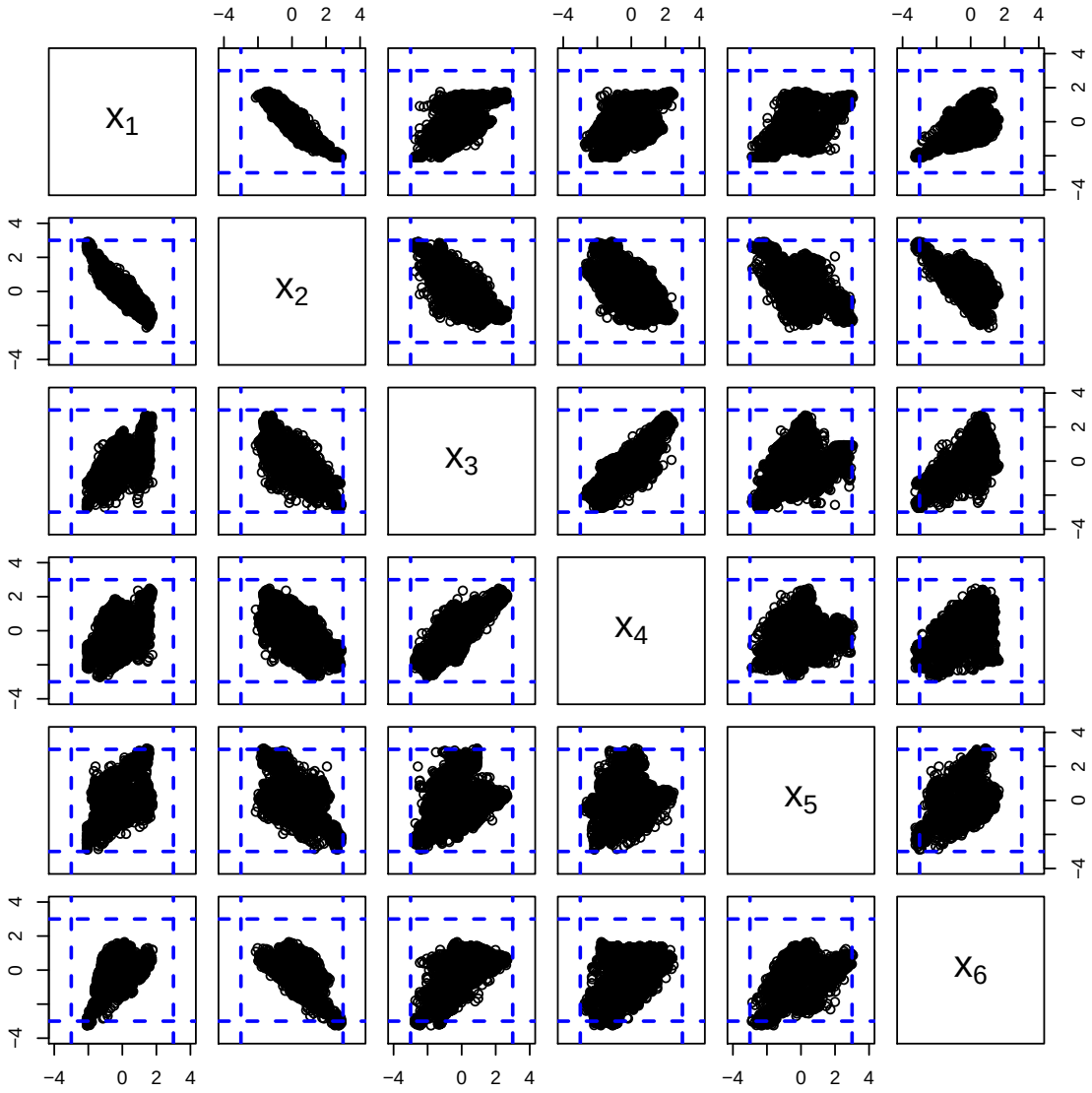
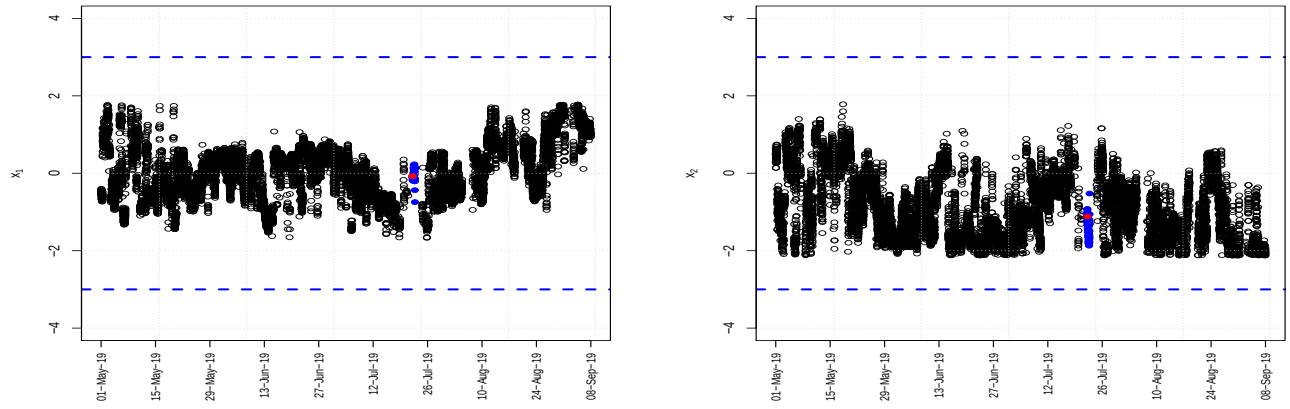
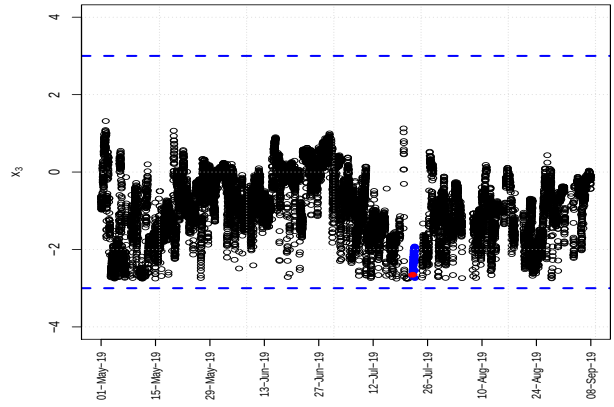


Figure 6.28: Scatter plot matrix of the IOC data for boiler process variables. Dotted lines are the OWAP limits.

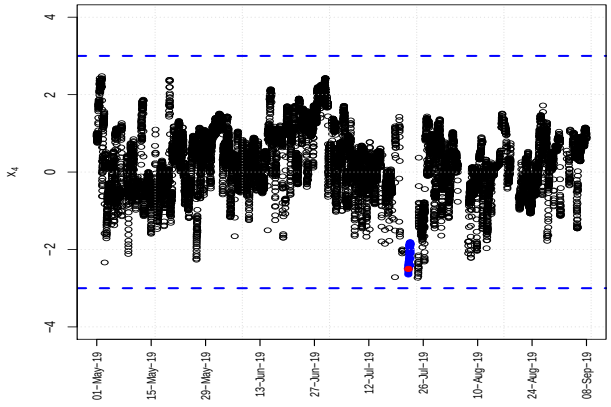


(a) x_1

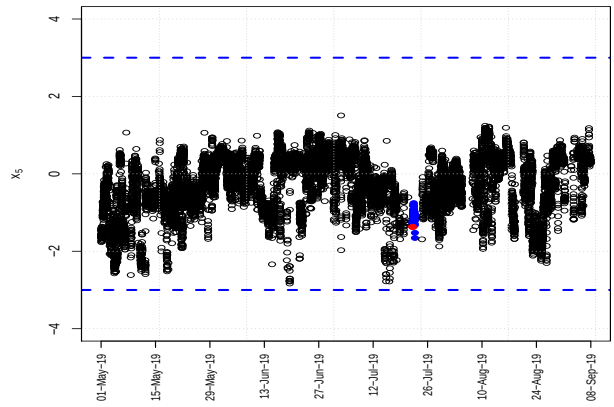
(b) x_2



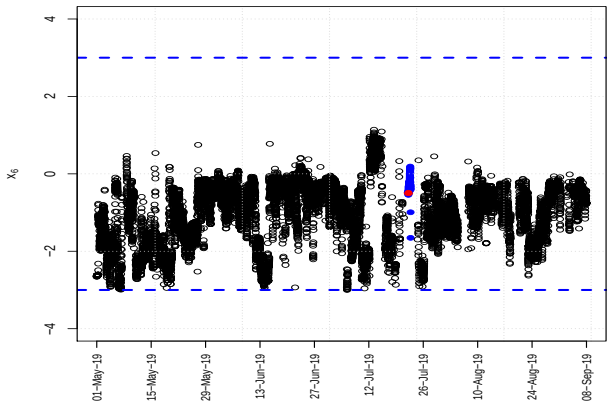
(c) x_3



(d) x_4



(c) x_5



(d) x_6

Figure 6.29: Univariate observations on the six boiler process variables. Black points represent measurements over time. Horizontal broken blue lines represent univariate OWAP control limits. The red and blue dots are specific examples selected that correspond to multivariate process deviations.

PCA based multivariate fault identification statistics were therefore employed to identify possible multivariate process abnormalities. Refer back to Chapter 2. For the moment only the T^2 statistic will be considered. The result of the T^2 evaluations over the selected time period are captured by Figure 6.30. Each hollow circular black point represents a T^2 statistic value for a specific point in time. The horizontal red broken line in Figure 6.30 represents the T^2 control limit. It can be observed from Figure 6.30 that for the majority of the time the performance of the boiler was multivariately in-control, based on the T^2 statistic. However, there is a large number of instances where the performance was multivariately out-of-control as is indicated by the points that are recorded above the control limit. The T^2 multivariate monitoring therefore provides additional information on boiler process conditions which are not highlighted by the univariate graphs in Figure 6.29.

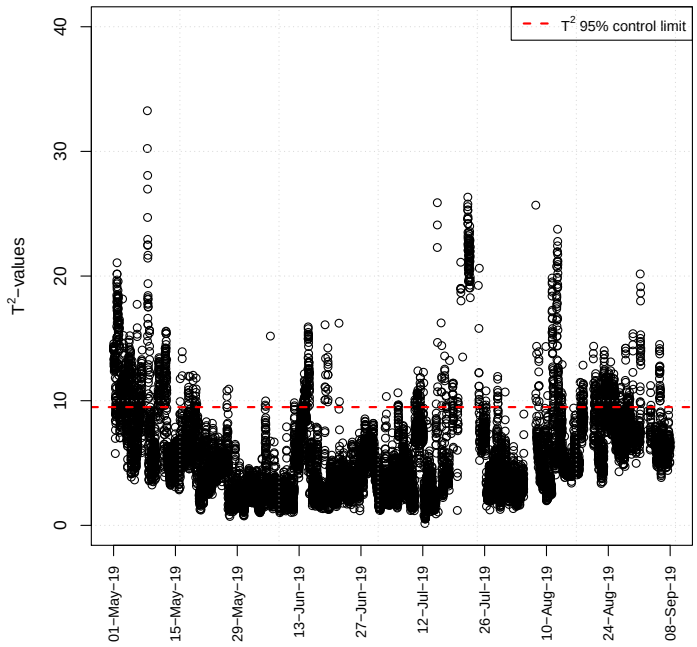


Figure 6.30: T^2-statistic evaluation of new data over time. Red broken horizontal line represents the 95% control limit.

However, multivariate control charts, in isolation, do not assist engineers much from a practical viewpoint. It only provides them with an indication of whether assignable cause variability is present or not. Diagnostic analysis is required to help engineers to interpret the instances when there are statistically defined deviations from multivariate common cause variability. The diagnostic process will be illustrated by focusing on specific samples that are out-of-control. It is interesting to observe, from Figure 6.30, that the boiler was operating mostly in-control for approximately a month from middle June 2019 to middle

July 2019. After this period there was a relatively big movement from common cause variability. It would therefore be informative to understand what caused the significant deviation after operating ideally for a relatively long period. One of these points will therefore be selected to demonstrate the diagnostic step. The specific sample selected is indicated by a solid red dot in Figure 6.31. Also included in Figure 6.31 are vertical broken black lines to identify the month of relative stable operation prior to observing the large deviation.

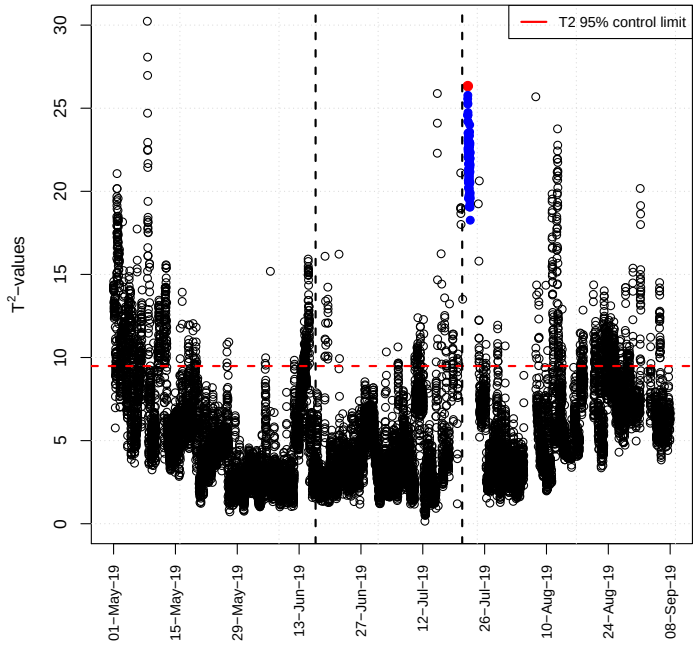
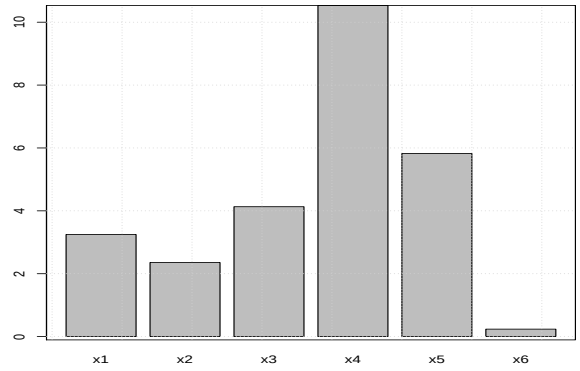


Figure 6.31: T^2 statistic evaluation of new data over time. Red and blue points indicate samples selected for diagnosis illustration. Vertical black broken lines identify the stable period before the deviation.

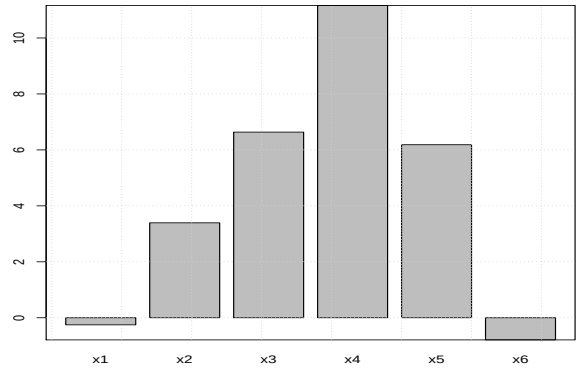
6.3.3.1 Traditional fault diagnosis

It will again be illustrated, in this practical setting, how traditional contribution diagnostic analysis do not intuitively assist engineers to interpret the statistically identified process deviation. Please consult Chapter 2 for a review of the traditional diagnostic approaches. The various traditional contribution analysis techniques discussed in this research document were calculated in R. The code is given in Appendix A. The results of the contribution analysis are captured by the bar graphs in Figure 6.32. The traditional approach to multivariate process monitoring would therefore provide the plant engineer with graphs similar to the ones in Figures 6.31 and 6.32. Given this information, the engineer is tasked to combine the statistical results with fundamental process understanding in

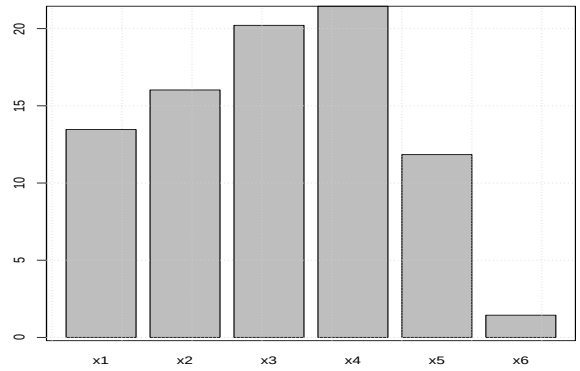
order to get the process back into multivariate statistical control. However, the statistical results provided do not practically give the engineer an intuitive understanding of why the process is flagged by the statistical tools as being multivariately out-of-control. Given the individual variable rankings provided in Figure 6.32 forces the engineer, to some extent, to think univariately i.e., to consider the boiler characteristics in isolation. The engineer would therefore, most likely, combine the information of Figure 6.32 with the univariate dashboard in Figure 6.29. The corresponding process data related to the red out-of-control point selected in Figure 6.31 are indicated as red points in Figure 6.29. It would then not be clear why the multivariate fault detection statistic is flagging, since univariately all of the process variables are inside their univariate IOC ranges.



(a) Complete decomposition



(b) Partial decomposition



(c) Reconstruction based

Figure 6.32: PCA T^2 contribution analysis results of the process fault indicated by the red point in Figure 6.31.

6.3.3.2 Pairwise fault diagnosis

The diagnostic approach will now be performed using the new pairwise contribution method developed in this research document. Please consult Chapter 3 for the definition of this method. The objective here is to only give information on how to practically implement and use the pairwise contribution method.
In order to determine the pairwise contributions, it is first required that additional information be calculated from the fitted PCA model. Therefore, the next step in setting up the PCA modelling infrastructure is to calculate the information required for the diagnostic part. First, the IOC distributions of the elements of the T^2 pairwise contribution matrix are calculated. The R code are presented in Appendix A. Note that in the R code it is only illustrated how the distributional information are collected for the T^2 statistic. The approach would be the same for the SPE and $Combined$ statistics. All that is required is to appropriately replace the distance matrix D. The R function used to calculate the pairwise contributions for a specific control statistic and out-of-control observation is also given in Appendix A.

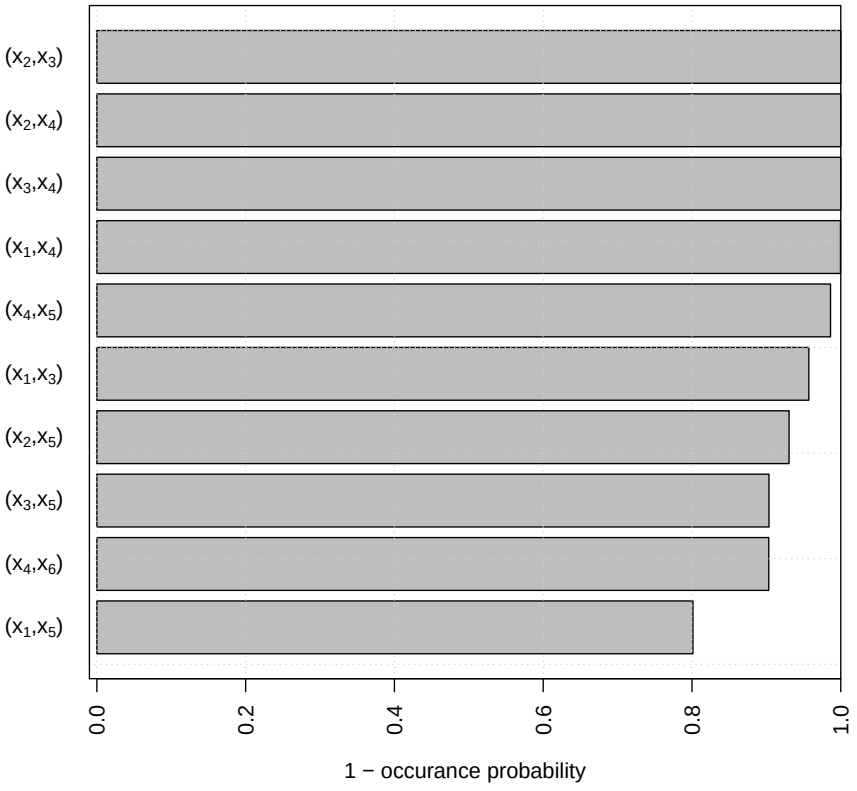


Figure 6.33: PCA pairwise contribution analysis based on the T^2-statistic for out-of-control sample selected in Figure 6.31. The top 10 contributing process variable pairs are listed.

The results of the pairwise diagnosis of the deviating red point selected in Figure 6.31 are reported by the bar chart in Figure 6.33. Each bar represents one minus the probability of observing the distance measured for a specific variable pair. In other words, for each variable pair the following question is asked: what is the probability of not observing the measured distance given the distribution specified using the reference set of distances? The deviation identified by Figure 6.31 is multivariate. Therefore, by providing information on the important variable pairs forces engineers to think multivariately about the boiler operation. In Figure 6.33 it is indicated, for example, that the variable pairs $(x_2, x_3), (x_2, x_4), (x_2, x_3)$ and (x_1, x_4) each reported distances that have a zero probability of being part of their reference distribution. Although many pairwise plots are to be considered, here one of these pairs illustrate the methodology. Focusing on the pair, e.g. (x_2, x_3), will prompt the engineer to visualise the relationship between the process variables. One advisable way of viewing the pairwise relationship is to construct a scatterplot using the IOC data together with the newly measured data point. Figure 6.34 is an example of such a graph for (x_2, x_3). The black points in Figure 6.34 represent the IOC data for (x_2, x_3). A blue point is used to represent the observed measurement on x_2 and x_3 for the T^2 observation selected in Figure 6.31. The process engineer, when looking at this graph, would understand, based on the relationship between x_2 and x_3, why the statistical analysis indicated that the process has deviated from multivariate common cause variability. Any corrective engineering solution should therefore adjust boiler operation in some way such that the (x_2, x_3) measurements are recorded inside or close to the point cloud i.e., the black points, which describes common cause variability in Figure 6.34.
The deviation identified in variables x_2 and x_3 has practical meaning to process engineers in this case. Thermodynamic fundamentals indicate that optimal pressure regulation in the boiler is indirectly reflected by the IOC relationship observed for (x_2, x_3) in Figure 6.34. Therefore, the multivariate deviation in this instance points to suboptimal pressure regulation in the boiler. Reasons for abnormalities in boiler pressure regulation are numerous. One possibility might be that water has been introduced into the boiler's burning zone due to small tube leakages. Therefore, the diagnostic results produced by the pairwise contribution analysis provide information that practically indicate why the multivariate out-of-control state is detected. An example is also given of how the statistical diagnosis information can be related to fundamental boiler understanding.

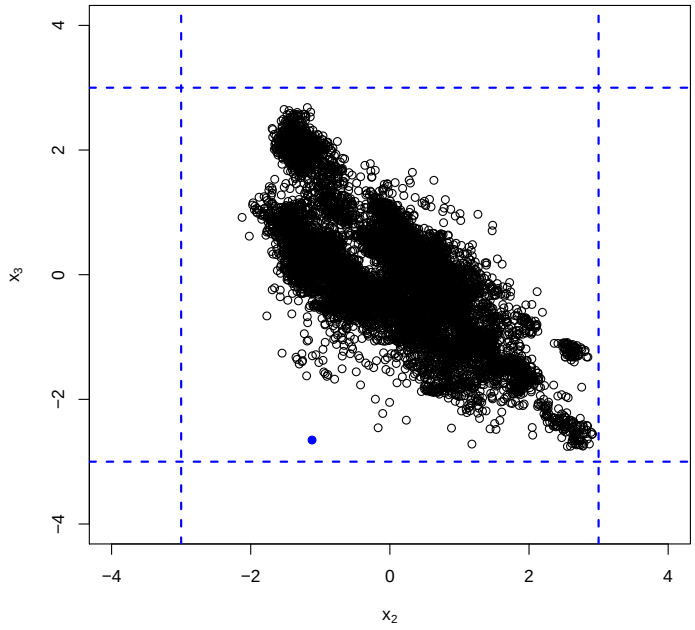


Figure 6.34: Evaluation of high T^2-statistic contributing pair (x_2, x_3) to out-of-control sample selected in Figure 6.31. This pair was identified using the PCA based pairwise contribution approach. Black points represent IOC data for x_2 and x_3. The blue point is the observed (x_2, x_3) data for the deviating point selected in Figure 6.31 according to the T^2 statistic. (x_2: Feed pressure measurement, x_3: Outlet temperature measurement)

It is observed in Figure 6.31 that a large number of observations close to the red point also display relatively large T^2 values. It would be informative to assess if these points consistently yield the same diagnostic information. Specifically, the additional points that will be diagnostically analysed are the 86 blue points in Figure 6.31. The selected points are also highlighted on the univariate dashboard in Figure 6.29. For each of these samples, variable pairs were isolated that reported a contribution which was significant at a 5% significance level. The frequency table of the variable pairs identified are reported in Table 6.6. It is observed from Table 6.6 that the variable pairs (x_2, x_3) and (x_2, x_4) were consistently identified to be important for all of the 86 samples analysed. Therefore, the pairwise contribution analysis of the out-of-control blue points in Figure 6.31 consistently identify the same variables as being problematic.

The observed pairwise data together with the IOC measurements for these variable pairs are displayed by Figures 6.35 and 6.36. It is clear from these graphs that the blue points, representing the new data, forms point clusters that are visually different from the regions which define the IOC observations i.e., the black points.

Table 6.6: PCA T^2 based pairwise contribution frequency analysis for the blue point samples selected in Figure 6.31.

Variable pair	Frequency
(x_2, x_3)	86
(x_1, x_4)	86
(x_2, x_4)	86
(x_3, x_4)	58
(x_4, x_5)	39
(x_1, x_3)	12
(x_2, x_5)	8
(x_3, x_5)	1

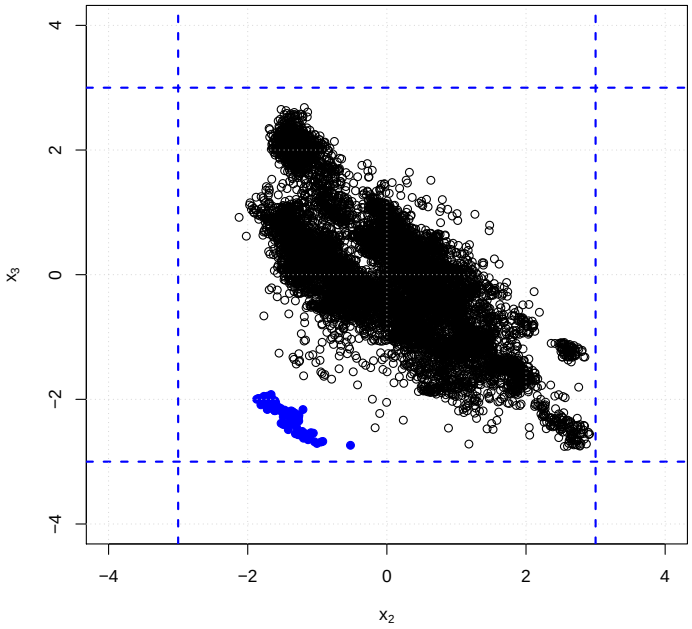


Figure 6.35: Evaluation of high T^2-statistic contributing pair (x_2, x_3) to the out-of-control samples selected in Figure 6.31. This pair was identified using the PCA based pairwise contribution approach. Black points represent IOC data for x_2 and x_3. Blue points are the corresponding (x_2, x_3) data for the out-of-control samples selected. (x_2: Feed pressure measurement, x_3: Outlet temperature measurement)

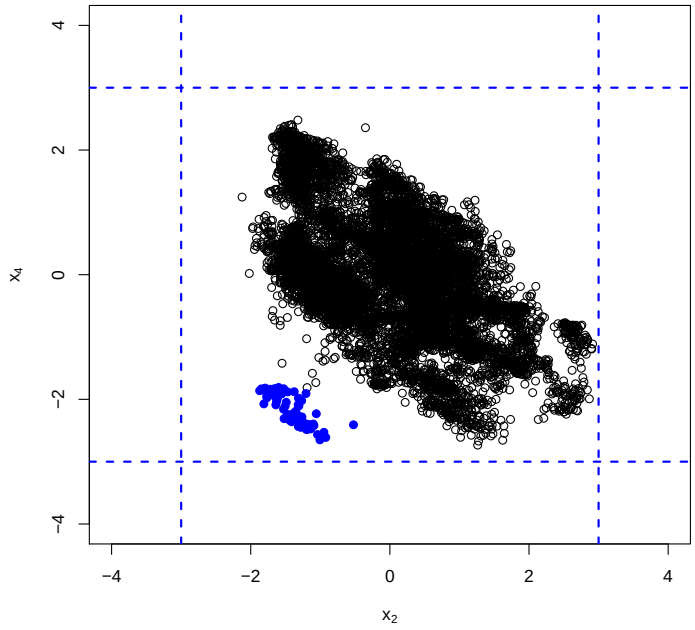


Figure 6.36: Evaluation of high T^2-statistic contributing pair (x_2, x_4) to the out-of-control samples selected in Figure 6.31. This pair was identified using the PCA based pairwise contribution approach. Black points represent IOC data for x_2 and x_4. Blue points are the corresponding (x_2, x_4) data for the out-of-control samples selected. (x_2: Feed pressure measurement, x_4: Feed preheat temperature)

It is informative to note that for the violating samples selected in Figure 6.31 that the SPE statistic also flagged the same data points as being out-of-control. The pairwise contribution analysis based on the SPE information was also performed. Results of the analysis are reported by Table 6.7 and Figure 6.37. It is striking to observe the similarities and dissimilarities between the T^2 and SPE pairwise contribution analysis results. Both analysis identified that (x_2, x_3) needs to be considered when attempting to correct the process deviation. However, the SPE analysis identified (x_1, x_2) as being important, which was not flagged by the T^2 analysis. The validity of the identification of the variable pair (x_1, x_2) is justified when one observe the scatterplot in Figure 6.37, displaying the observed data against the reference data. There is visually a clear separation between the observed and reference data. Similarly the variable pair (x_2, x_4) is identified by the T^2 based analysis as important but not by the SPE analysis. Therefore, this analysis illustrate that it is important to collectively use the information from the various fault detection indices to correctly diagnose process faults.

Table 6.7: PCA SPE based pairwise contribution frequency analysis for the blue point samples selected in Figure 6.31.

Variable pair	Frequency
(x_1, x_2)	86
(x_2, x_3)	86
(x_3, x_6)	75
(x_3, x_5)	60
(x_4, x_5)	51
(x_4, x_6)	41
(x_2, x_6)	34
(x_1, x_3)	28
(x_2, x_4)	7
(x_2, x_5)	5
(x_3, x_4)	1

The blue points selected in Figure 6.31 were also identified by the *Combined* statistic as being abnormal based on the 95% control limit. Pairwise contribution analysis based on the *Combined* statistic for these samples is summarised in Table 6.8. The frequencies agree, to a large extent, with the contributions calculated using the SPE statistic in Table 6.7. Interestingly, the *Combined* statistic reports a relatively high pairwise contribution frequency for (x_2, x_4), which was identified by the T^2 statistic, Table 6.6, as being important but not by the SPE statistic, Table 6.7. However, the *Combined* statistic contribution results do not agree with that of the T^2 statistic in identifying that the variable pair (x_1, x_4) is important. Therefore, this analysis illustrates that it is important to collectively use the information from the various fault detection indices on the pairwise contributions to correctly diagnose process faults.

Table 6.8: PCA *Combined* statistic based pairwise contribution frequency analysis for the blue point samples selected in Figure 6.31.

Variable pair	Frequency
(x_1, x_2)	86
(x_2, x_3)	86
(x_3, x_6)	74
(x_2, x_4)	70
(x_3, x_5)	65
(x_4, x_5)	63
(x_3, x_4)	43
(x_4, x_6)	40
(x_1, x_3)	31
(x_2, x_6)	29
(x_2, x_5)	27
(x_1, x_4)	1

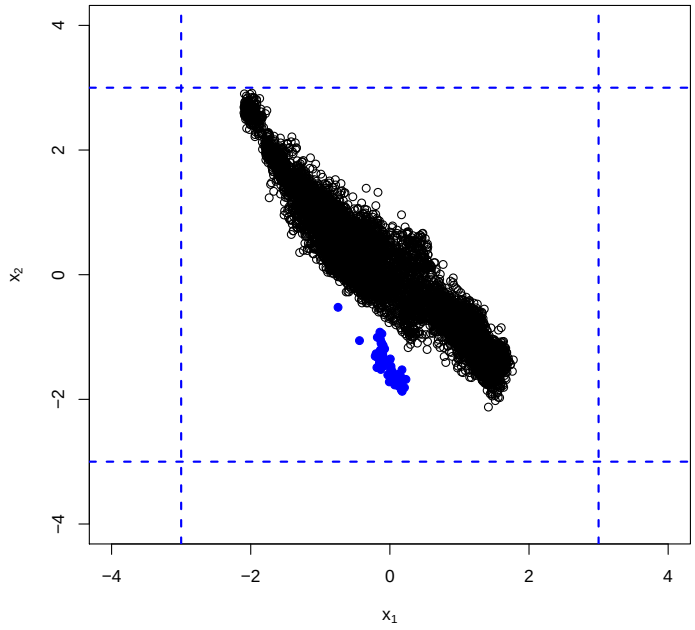


Figure 6.37: Evaluation of high SPE-statistic contributing pair (x_1, x_2) identified using PCA based pairwise contribution approach. Black points represent IOC data for x_1 and x_2. Blue points are the corresponding (x_1, x_2) data for the out-of-control samples selected. (x_1: Gas flow control, x_2: Feed pressure measurement)

Further research is required to better understand why the pairwise contribution analysis based on the different fault identification indices will yield different results. On an intuitive level it can be expected that the various indices will give different interpretations. This intuition comes from the fact that the T^2 statistic quantifies variability observed in the PCA scores and that the SPE statistic summarises variability in the residual space. Following from the discussion of PCA in Chapter 2, it is possible to decompose the IOC data set $\mathbf{X} : n \times p$ as follows,

$$\mathbf{X} \quad = \quad \hat{\mathbf{X}} + \tilde{\mathbf{X}}, \tag{6.11}$$

where $\hat{\mathbf{X}} = \mathbf{X}\mathbf{V}_{p\times r}\mathbf{V}_{r\times p}^T$ is the PCA approximation of $\mathbf{X}$ and $\tilde{\mathbf{X}} = \mathbf{X}\tilde{\mathbf{V}}_{p\times(p-r)}\tilde{\mathbf{V}}_{(p-r)\times p}^T$ is the residual matrix. This decomposition is similar to that of the covariance matrix as presented in Chapter 2.

A possible reason for the discrepancy of (x_1, x_2) being identified as important by the SPE based analysis and not by the T^2 based analysis can be obtained when the PCA based decomposition of (x_1, x_2) is considered. The decomposition of the (x_1, x_2) data in Figure 6.37 is given by Figures 6.38 and 6.39 respectively. In both graphs black points are

used to represent the reference data and blue points are used to indicate the out-of-control points selected in Figure 6.31. The PCA approximation, Figure 6.38, can be considered a view of how the PCA model summarises the variability between x_1 and x_2. It is therefore reasonable to deduce that any analysis based on the PCA model will, in this instance, conclude that the variability displayed by the blue points are likely to be common to the process, since there is no visual separation between the two groups of data. Using this thought process, it can therefore be expected that the T^2 based pairwise contribution results for (x_1, x_2) will therefore not be significant, since the T^2 statistic uses information generated from the PCA model.

A similar explanation can be given for the SPE based pairwise contribution results obtained in relation to (x_1, x_2). The variability in the PCA residuals are visually represented by Figure 6.39. There is clear separation between the blue points, representing the out-of-control points, and the black points, representing the reference conditions. It is therefore plausible that statistical analyses, based on the residual information, that quantifies the variability as displayed in Figure 6.39 will conclude that the blue points are statistically different from the reference data. The identification of (x_1, x_2) as being important by the SPE based pairwise contribution analysis, is therefore sensible, since the SPE statistic is based on the residual information.

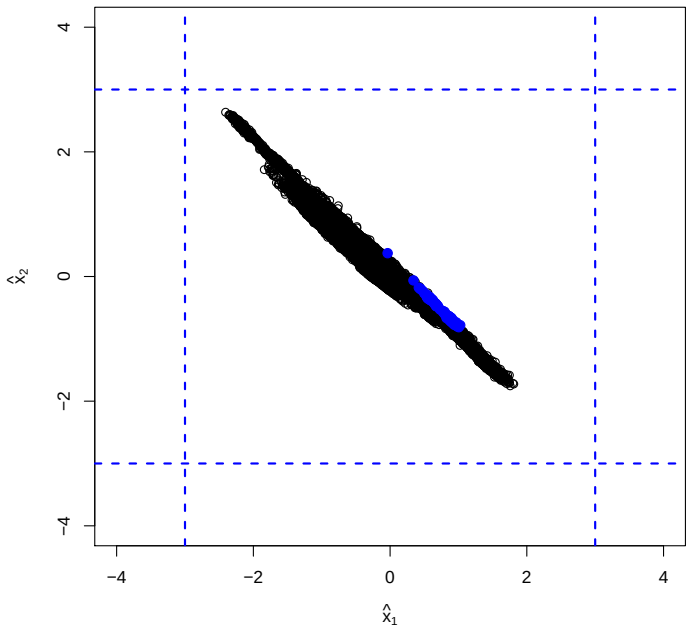


Figure 6.38: PCA approximation of the data for (x_1, x_2) represented in Figure 6.37. Black points represent the approximation of the IOC data for x_1 and x_2. Blue points are the corresponding estimation of (x_1, x_2) for the out-of-control samples selected. (x_1: Gas flow control, x_2: Feed pressure measurement)

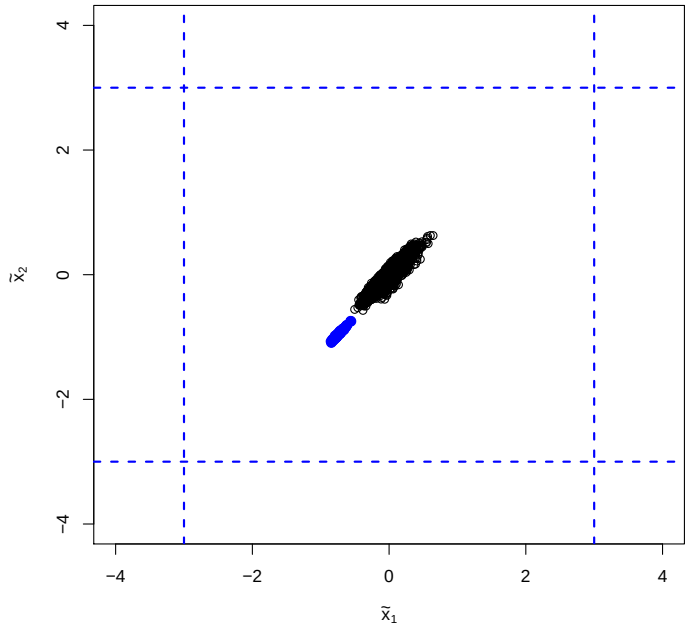


Figure 6.39: PCA residuals of the data for (x_1, x_2) represented in Figure 6.37. Black points represent the residuals of the IOC data for x_1 and x_2. Blue points are the corresponding residuals of (x_1, x_2) for the out-of-control samples selected. (x_1: Gas flow control, x_2: Feed pressure measurement)

It was illustrated in Chapter 3 that it is possible to identify an individual ranking of process variables. The following ranking was identified:

$$\begin{bmatrix} x_4 & x_3 & x_5 & x_1 & x_2 & x_6 \end{bmatrix}. \tag{6.12}$$

Therefore, variable x_4 is an important characteristic that should be considered in diagnosing the process fault. The engineer should use this ranking together with the pairwise contributions to diagnose the multivariate process deviation. It is interesting to note that all of the traditional variable contribution results in Figure 6.32 identify that process variable x_4 is the most important. However, the order of the other process variables is different. The ranking identified by the partial decomposition contribution method is the closest to the one identified by the new method.

A useful property of the new diagnostic method developed in Chapter 5, for the analysis of non-linear multivariate process faults, is that it can also be applied for the linear case. In the non-linear setting the IOC point cloud for each variable pair is irregularly shaped. For the linear case the IOC points clouds have a simpler ellipsoidal form when compared to the non-linear IOC regions. Therefore, the methodologies of Chapter 5 can be applied

for the diagnosis of process faults detected using linear PCA. It will now be illustrated how this method can be applied to diagnose the highlighted faults displayed by Figure 6.31. The same IOC data set that was used to construct the reference distributions of the T^2 pairwise contributions will be used i.e., the IOC samples that are in-control according to the 95% T^2 control limit. It was decided to employ the SVM class membership probabilities (5.49) to define the distance measures to be used in this illustration. The R code used to do the analysis are again presented in Appendix A. Note that the R function is highly reliant on the R library developed by Karatzoglou et al. (2004). The SVM based distance matrix can now be calculated for the red point in Figure 6.31. This yielded the distance matrix displayed by (6.13). It can, for example, be observed from (6.13) that there is only a six percent chance that the new bivariate measurement made on variable 1 and 3 is part of the corresponding IOC point cloud.

$$\begin{bmatrix} 0.00 & 1.00 & 0.94 & 1.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 1.00 & 1.00 & 0.93 & 1.00 \\ 0.94 & 1.00 & 0.00 & 1.00 & 0.91 & 1.00 \\ 1.00 & 1.00 & 1.00 & 0.00 & 1.00 & 1.00 \\ 0.00 & 0.93 & 0.91 & 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 1.00 & 1.00 & 0.00 & 0.00 \end{bmatrix} \tag{6.13}$$

An individual ranking of process variables can also be identified using (6.13). The following individual ranking of process variables was identified:

$$\begin{bmatrix} x_4 & x_2 & x_3 & x_1 & x_5 & x_6 \end{bmatrix}. \tag{6.14}$$

Process variable x_4 is again identified to have the highest importance. If the fourth row of (6.13) is examined, it is observed that all of the pairwise contributions that include x_4 is significant. The engineer should therefore investigate if there is something specific to x_4 that causes its relationship with the other process variables to be abnormal.

A frequency analysis was also performed to evaluate the pairwise contributions for the blue points in Figure 6.31. The results are captured in Table 6.9. It is clear, from Table 6.9, that the information from Tables 6.6, 6.7 and 6.8 are combined by the SVM based pairwise contribution analysis. For example, the variable pair (x_1, x_4) was identified by the T^2 based analysis, Table 6.6, as important but not by the SPE and *Combined* statistic pairwise analysis. Also, the variable combination (x_1, x_2) was identified by the pairwise contribution analysis based on the SPE and *Combined* statistics i.e., Tables 6.7 and 6.8,

but was not identified by the T^2 analysis. However, both sets of variables i.e., (x_1, x_4) and (x_1, x_2) are identified by the SVM based pairwise contribution analysis. Additional variable pairs are also identified in Table 6.9 which were not highlighted by the other pairwise contribution techniques. One such example is (x_2, x_6). The IOC data together with the new data points for (x_2, x_6) are displayed in Figure 6.40. It is clear that there is a visual difference between the IOC black points and the blue points representing the new data. Therefore, the identification of (x_2, x_6) is sensible.

Table 6.9: SVM based pairwise contribution frequency analysis for the PCA T^2 out-of-control blue point samples selected in Figure 6.31.

Variable pair	Frequency
(x_1, x_2)	86
(x_1, x_3)	86
(x_2, x_3)	86
(x_2, x_4)	86
(x_2, x_5)	86
(x_3, x_6)	86
(x_2, x_6)	85
(x_4, x_6)	85
(x_1, x_4)	55
(x_4, x_5)	44
(x_3, x_5)	42
(x_3, x_4)	22

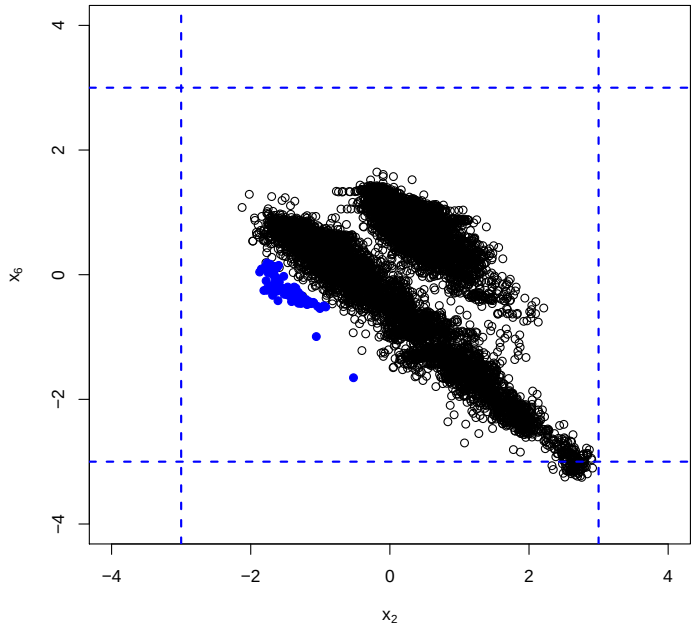


Figure 6.40: Evaluation of high contributing pair (x_2, x_6) identified using SVM based pairwise distance approach. Black points represent IOC data for x_2 and x_6. Blue points are the corresponding (x_2, x_6) data for the out-of-control PCA T^2 samples selected. (x_2: Feed pressure measurement, x_6: Gas feed flow)

6.3.3.3 PCA results summary

The PCA based diagnosis results presented in this section for the boiler data can be summarised as follows:

1. The pairwise contribution analysis results clearly communicate the reason why the process is classified as being multivariately out-of-control.

2. The results are fairly consistent for deviating observations that are recorded close to each other in time.

3. It is important that the pairwise diagnosis results based on the different fault identification statistics are collectively considered.

4. Non-linear pairwise contribution analysis can be applied to diagnose deviations identified using PCA. It is observed from this example that the non-linear diagnosis results combine the diagnosis results obtained from the various PCA based pairwise analyses.

6.3.4 Kernel PCA MSPM

The reader will observe from Figure 6.40 that the shape of the black point cloud, representing IOC's, can be grouped into two distinct clusters. This is clearly different from the ellipsoidal shape that is expected when the relationship between process variables are linear. Fundamental boiler understanding indicate that the reason for the two point clusters is due to soot blowing. The larger cluster represent the IOC relationship of x_2 and x_6 during normal operation. The smaller cluster is the IOC region when soot blowing is performed. Therefore, soot blowing has an effect on the characteristics of the process.
Soot is a byproduct of the combustion process that deposits on the furnace tubes. The presence of soot in the boiler can influence the efficiency of steam production. Therefore, soot blowing is performed to remove the unwanted deposits. In this specific case study the timestamps that identify periods of soot blowing are not electronically captured and stored on the plant historian. The soot blowing dates are currently hand written into the control room log books. These dates are also not available for the complete historical data set that is used to describe IOC's. Therefore, in the statistical monitoring of the boiler, the periods when soot blowing occur represent an unobservable process variable that influence the behaviour of the process. Based on this information and the discussions of Chapter 4, it is clear that a method such as kernel PCA is better suited to multivariately monitor boiler operation. In this section it will be illustrated, using R code, how kernel PCA based fault detection and diagnosis perform when applied to the boiler data.

One practical limitation, that was encountered, is that kernel PCA model development can be computationally slow when data sets are large. It was decided to work with a reduced data set to speed up the modelling step. The reduced data set was obtained by selecting 2000 samples using the k-means clustering method (Hastie et al. (2001)). Other techniques exist that can be used to address the computational expense associated with kernel PCA, see for example Bencheikh et al. (2020). Although downsampling will impact on the estimates based on the reduced sample, it was found that the approach of using clustering provided satisfactory results.
The Gaussian kernel was again used in the kernel PCA model definition. In Appendix A the R function used to fit the model is defined. Algorithm 1, in Chapter 4, was used to identify the optimal hyperparamters for the model. Optimal settings for the Gaussian kernel hyperparameter and kernel PCA dimension equalled 0.15 and 120 respectively.
New observations of the boiler being monitored will now be evaluated using the kernel PCA model. The kernel PCA T^2 statistic will be used as an example to illustrate the non-linear MSPM fault detection and diagnosis. The kernel PCA T^2 statistic values of the new observations are depicted in Figure 6.41. The dots represent the T^2 values over time and the horizontal red broken line is used to display the 95% control limit. The

control limit is empirically determined using the T^2 statistic values of the IOC data.

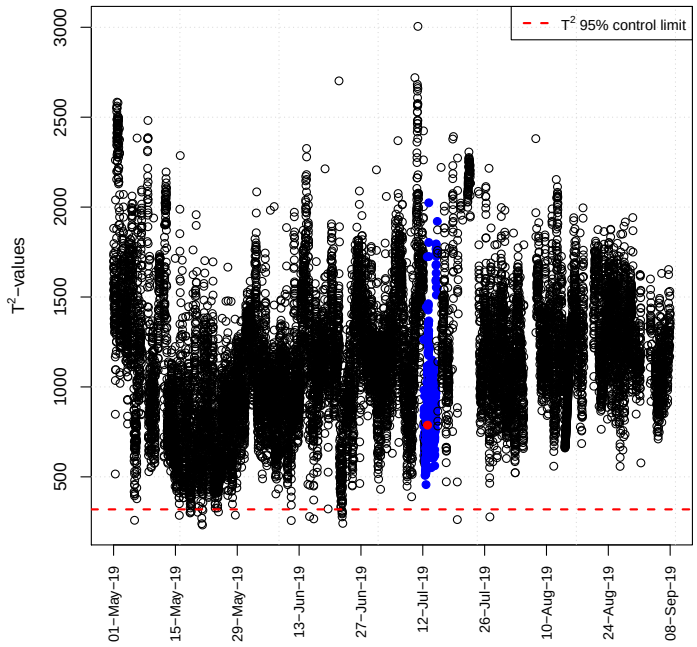


Figure 6.41: Kernel PCA T^2 statistic values over time for the new observations in Figure 6.29. The red and blue points indicate samples that will be used to demonstrate fault diagnosis.

.

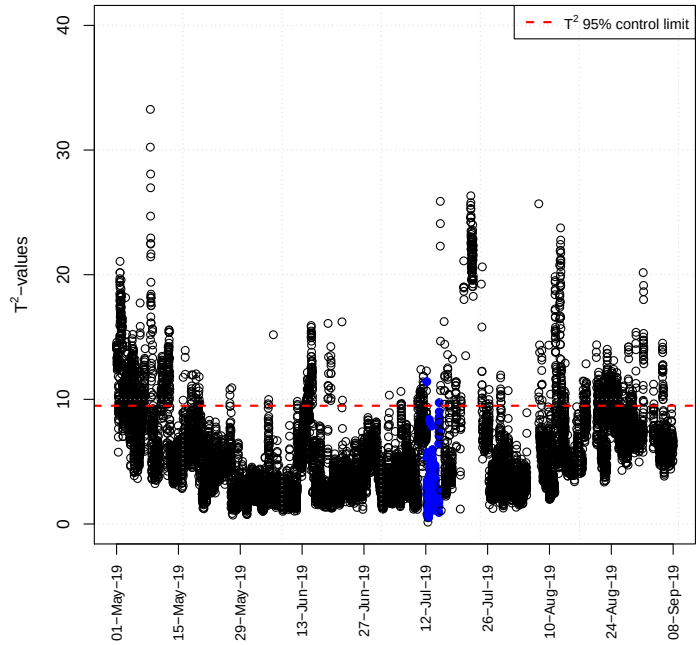
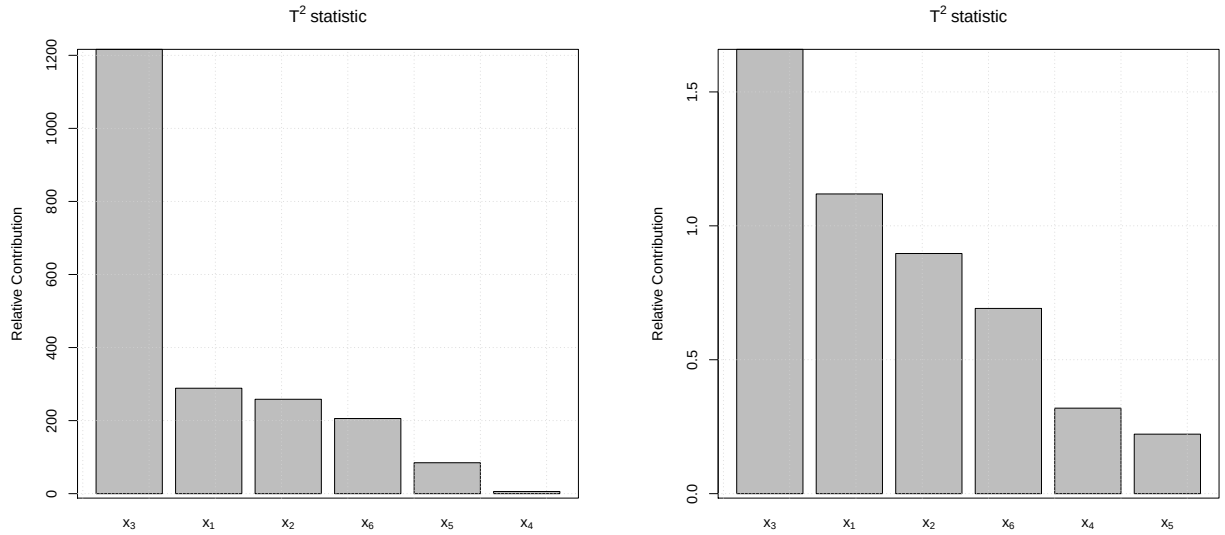


Figure 6.42: PCA T^2 statistic values over time for the new observations in Figure 6.29. Blue points are the corresponding points selected in Figure 6.41.

.



(a) Gradient based. (b) Reconstruction based.

Figure 6.43: Gradient based (Cho et al., 2005) and Reconstruction based (Alcala and Qin, 2010) contribution analysis results of the red sample selected in Figure 6.41. Bar graphs are used to represent the T^2 statistic based contributions.

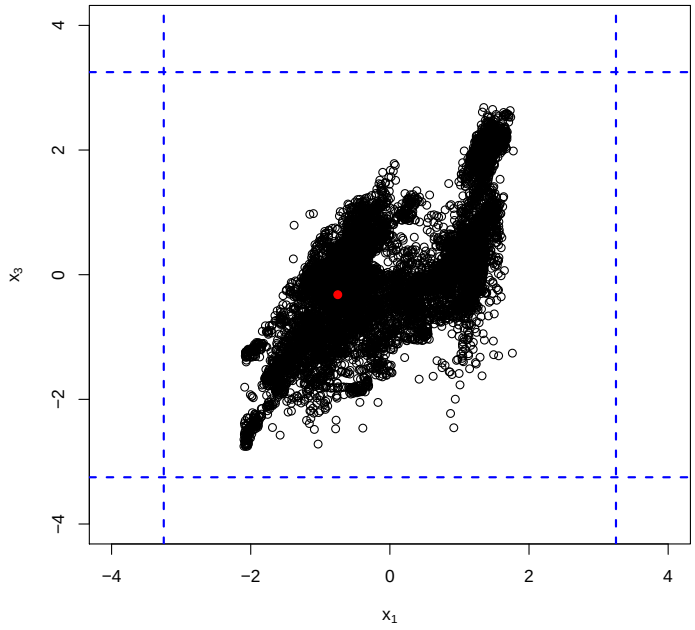


Figure 6.44: Evaluation of (x_1, x_3) identified using gradient and reconstruction based approach. Black points represent IOC data for x_1 and x_3. The red point is the corresponding (x_1, x_3) data for the out-of-control sample selected in Figure 6.41. (x_1: Gas flow control, x_3: Outlet temperature measurement)

It can be observed from Figure 6.41 that for the majority of the period under consideration the T^2 values are above the 95% control limit. Therefore, the boiler is considered to be out-of-control most of the time. The samples corresponding to the extreme T^2 values need to be analysed in order to diagnose why assignable cause variability is considered to be present. Diagnosis of the red point in Figure 6.41 will first be considered.

6.3.4.1 Traditional fault diagnosis

Results obtained by using the traditional diagnosis methods i.e., gradient and reconstruction based contribution analysis, are depicted in Figure 6.43. Both methodologies identify x_3 and x_1 as the top two contributors. It is not clear why these two variables are being highlighted. Given that the process is univariately inside the OWAP bounds, it was decided to collectively consider x_3 and x_1 in Figure 6.44. This representation also does not assist in interpreting the multivariate deviation since the new observation is located in the IOC point cloud. Therefore, in this example it proofs difficult to interpret the information produced by the traditional approaches.

6.3.4.2 Pairwise fault diagnosis

The set of SVM based pairwise models constructed in the previous subsection was now used to identify the variables with significant pairwise distances. The following distance matrix is obtained:

$$\begin{bmatrix} 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.91 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.91 & 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix} \tag{6.15}$$

Therefore, variable pairs (x_2, x_6) and (x_1, x_2) were identified to have SVM based distances that are larger than normal. In Figures 6.45 and 6.46 the pairwise scatter plots are given for (x_2, x_6) and (x_1, x_2) respectively. It is clear from these figures that the identification of (x_2, x_6) and (x_1, x_2) is sensible. For example, in Figure 6.45 the reason for the multivariate out-of-control categorisation can be interpreted by observing that the measurement is recorded between the two IOC point clusters.

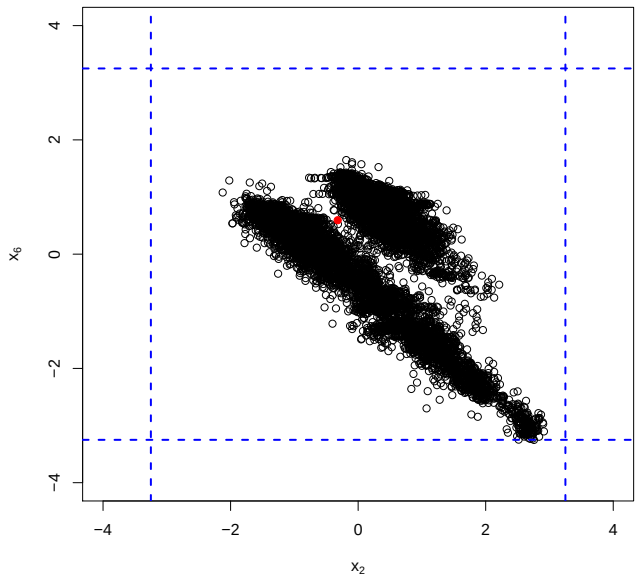


Figure 6.45: Evaluation of high contributing pair (x_2, x_6) identified using SVM based pairwise distance approach. Black points represent IOC data for x_2 and x_6. The red point is the corresponding (x_2, x_6) data for the out-of-control sample selected in Figure 6.41. (x_2: Feed pressure measurement, x_6: Gas feed flow)

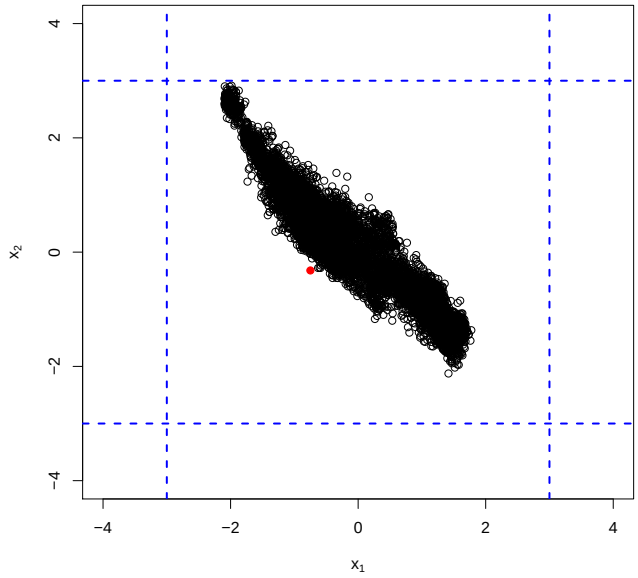


Figure 6.46: Evaluation of high contributing pair (x_1, x_2) identified using SVM based pairwise distance approach. Black points represent IOC data for x_1 and x_2. The red point is the corresponding (x_1, x_2) data for the out-of-control sample selected in Figure 6.41. (x_1: Gas flow control, x_2: Feed pressure measurement)

A few of these samples will now be considered, similar to the approach in Section 6.3.3. The 395 selected blue points in Figure 6.41 will be used to evaluate the frequency of the identified variable pairs. It should be noted that the selected blue points were not identified by the linear PCA T^2 statistic as being out-of-control, see Figure 6.42. The result of the diagnostic analysis is summarised by the frequencies captured in Table 6.10.

Table 6.10: SVM based pairwise contribution frequency analysis for the blue point samples selected in Figure 6.41.

Variable pair	Frequency
(x_2, x_6)	276
(x_1, x_2)	271
(x_2, x_3)	141
(x_3, x_4)	75
(x_3, x_6)	74
(x_5, x_6)	20
(x_2, x_5)	18
(x_1, x_3)	7
(x_1, x_5)	5
(x_3, x_5)	5
(x_4, x_5)	3

It can be observed from Table 6.10 that (x_2, x_6) was identified as an important variable

pair for 276 out of the 395 samples for diagnostic purposes. The IOC data together with the measurements of the blue samples selected in Figure 6.41 are depicted in Figure 6.47. It can be observed from Figure 6.47 that a large number of the new observations is located between the two clusters. Therefore, the identification of (x_2, x_6) as having high contribution is sensible. A possible interpretation might be that during soot blowing the process did not exhibit the ideal relationship between x_2 and x_6. The engineering team should investigate why this deviation is observed and appropriately implement measures to correct the process. Therefore, the methodology of Chapter 5 provides a clear indication of how to interpret the multivariate process deviation. Given a good understanding of why the kernel PCA based model is flagging will assist engineers to come up with the appropriate corrective action required to correct the deviation.

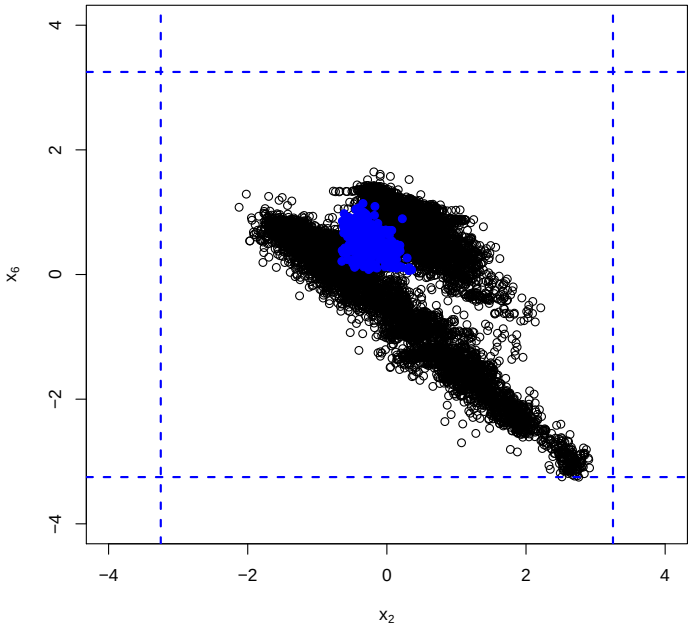


Figure 6.47: Evaluation of high contributing pair (x_2, x_6) identified using SVM based pairwise distance approach. Black points represent IOC data for x_2 and x_6. Blue points are the corresponding (x_2, x_6) data for the out-of-control samples selected in Figure 6.41. (x_2: Feed pressure measurement, x_6: Gas feed flow)

6.3.4.3 KPCA results summary

The kernel PCA based diagnosis results presented in this section for the boiler data can be summarised as follows:

1. The pairwise contribution analysis results clearly communicate the reason why the process is classified as being multivariately out-of-control.

2. The results are fairly consistent for deviating observations that are recorded close to each other in time.

3. It was demonstrated that traditional approaches does not clearly communicate why the multivariate out-of-control categorisation occurred.

6.4 Summary

In this chapter the value and applicability of the new fault diagnosis methodologies proposed in Chapters 3 and 5 were quantified. To this end, data from actual chemical process plants were analysed. The first case study focussed on diagnosing faults simulated in the well known Tennessee Eastman process. In the second case study process data produced by a coal fired steam boiler were analysed. Throughout the chapter, R code were also provided to illustrate how to implement the new methodologies in R. Using these case studies has again illustrated how common multivariate deviations are in practice. These type of process abnormalities are not identified when considering only the univariate graphs of process variables over time. The proper diagnosis of these assignable cause deviations is therefore important to ensure optimal process operation.
The Tennessee Eastman (TE) simulation (Downs and Vogel, 1993) produce data that represent an actual chemical process. This data are often utilised in literature to benchmark new MSPM methodologies. In this chapter the Tennessee Eastman dataset was used to evaluate the performance of the new fault diagnosis methodology introduced in Chapter 3. Data representing normal operating conditions were used to fit a PCA model and a kernel PCA model. The TE data include simulations of various faults that are common to the process. Observations selected from three of these fault simulations were used to demonstrate the diagnosis ability of the new method. It was observed for the selected faults that it is possible to perform the diagnosis by comparing the IOC OWAP ranges to the ranges observed during the fault simulation. However, the multivariate diagnosis techniques were still applied to evaluate the faults. A sensible minimum requirement of the MSPM diagnosis is that it must produce results that correspond with the univariate analysis. Note that due to the univariate differences the fault diagnosis mostly focussed on the univariate rankings obtained from the pairwise information.
The PCA diagnosis results based on the row sum method of Chapter 3 was observed to correspond with the univariate analysis for all of the selected faults. Additional variables were also identified that were found to be sensible identifications. However, the graphical approach, based on the MDS analysis, produced results that were difficult to visually interpret due to a large number of points on the MDS display. The method based on the analysis of the row sums were therefore preferred. It is concluded that the new methodol-

ogy of Chapter 3 is capable of statistically diagnosing the selected TE simulated faults. Process faults that were identified using kernel PCA was diagnosed using the SVM based pairwise distance approach. For each of the three faults it was observed that the MDS approach was able to visually indicate the variables that are univariately different. However, the row sums method identified additional variables that were found to operate close to OWAP bounds of the process. It is concluded that the non-linear pairwise methodology provided sensible diagnosis results for the selected TE simulated faults.

The section that discuss the coal fired steam boiler case study was predominantly divided into two parts. In the first part the focus was on MSPM using linear PCA. The second part was used to illustrated fault detection and diagnosis in the non-linear setting using kernel PCA.

For linear PCA, specific multivariate process faults were analysed to illustrate that the diagnosis approaches will practically assist engineers to statistically interpret the identified multivariate excursions.

The impracticality of traditional contribution analysis was briefly highlighted. Traditional individual ranking of process variables does not give the engineer a practical indication as to why the process is statistically categorised as displaying behaviour that is different from multivariate common cause variability.

The pairwise contributions identified by the new methodology of Chapter 3 provided practical information that can be used to supplement fundamental process understanding in order to better interpret the process deviation. For example, the diagnosis results might indicate that the relationship that exists between two variables are statistically different when compared to IOC observations. The engineer can then use fundamental understanding of the plant to interpret why the specific deviation might occur. An individual ranking of process variables was also calculated, based on the row sum approach, that can be interpreted in combination with the pairwise contributions.

It was also observed that the application of the new diagnosis method to real process data indicated that there exist instances when the fault identification statistics will collectively identify the same process conditions as being abnormal but will not necessarily provide the same reasons for the deviation. For example, the pairwise contributions calculated based on the T^2 statistic differed from those calculated using the SPE statistic. The contributions calculated based on the *Combined* statistic, to some extend, combined the results obtained from the T^2 and SPE statistic. However, there were still variable combinations identified by the T^2 statistic that were not identified by the *Combined* statistic. An attempt has been made to explain why this is happening. This explanation is based on the fact that the T^2 statistic is based on the PCA model data and that the SPE statistic

is based on the residual model data. However, further research is needed to better understand the diagnostic differences. Crucially, this observation illustrates that it is important to use all of the information provided by the PCA monitoring methods to attain optimal process fault detection and diagnosis.
It was illustrated that the new methodology developed in Chapter 5 for the diagnosis of non-linear multivariate process faults can also be used to evaluate process faults identified by linear PCA. It was informative to observe that the significant pairwise contributions calculated for the T^2, SPE and $Combined$ statistic by the method of Chapter 3 were also identified by the technique of Chapter 5. The non-linear diagnostic method also identified additional significant pairwise contributions not identified by the pairwise decomposition method of Chapter 3.
In the second part, of the boiler case study, a process specific reasoning was provided to motivate that the relationship between the process variables is non-linear. Therefore, a kernel PCA model was fitted to model the non-linear common cause variability of the process. Specific faulty observations that were identified using the kernel PCA T^2 statistic were selected to demonstrate the diagnostic ability of the new method developed in Chapter 5. The high ranking pairwise contributions identified by this method proved to be sensible when comparing the IOC data to the new measurements. Results of this method succeed in specifying to plant engineers why the observations were identified as being multivariately out-of-control. The real life coal fired steam boiler case study presented in this chapter clearly demonstrate that the methods developed in Chapters 3 and 5 are capable to diagnose multivariate process faults.

The results of the TE and boiler analysis suggest that the non-linear pairwise approach should be the preferred diagnosis methodology. In the boiler example it was observed that the non-linear pairwise approach identified more deviating pairs compared to the PCA based pairwise approach. Also for the univariate rankings in the TE example, the non-linear pairwise approach identified more variables that could be interpreted on an individual basis.

Chapter 7

Conclusions and Future Research

Modern chemical process plants are equipped with sensors that measure a vast amount of operational characteristics. This has led to an increase in the amount of process related data generated. Valuable information is therefore recorded that can assist engineers to better understand and ultimately optimise their processes. Statistical techniques are required to appropriately convert the high dimensional data sets into information that is practically interpretable by engineers. The statistical process monitoring (SPM) methodology is a popular set of techniques used to identify and diagnose when a process displays characteristics that are uncommon to the process. In this research linear and non-linear multivariate statistical process monitoring (MSPM) were considered. The main objective was to evaluate and improve on the methods used in Principal component analysis (PCA) and kernel PCA to diagnose process deviations accurately. The research objectives were:

- Document the shortcomings of the traditional techniques used for fault diagnosis in PCA and kernel PCA based MSPM through a theoretical evaluation of the methods, as well as with simulation studies.
- Develop a new methodology for fault diagnosis in PCA based MSPM.
- Develop a new methodology for fault diagnosis in kernel PCA based MSPM.
- Illustrate the fault diagnosis performance of the newly developed methods using data from a benchmark data set and a commercial chemical process.

7.1 Conclusions

A review was given of how PCA is typically utilised for MSPM. Definitions were provided for three different sensor fault types that can occur in practice. These are single, multiple and multivariate sensor faults. The specific definitions for multiple and multivariate sensor

faults have not previously been defined. It was demonstrated how PCA based MSPM can detect process deviations which are observed as a result of these fault types. It is important in multivariate MSPM that a fault diagnosis analysis is performed when an abnormal process behaviour is detected. The objective of the fault diagnosis analysis is to provide information that the process engineers can use to interpret the multivariate deviation. A detailed account of the traditional contribution analysis methodology for fault diagnosis was provided. The three main contribution analysis methods i.e., Complete decomposition (CD), Partial decomposition (PD) and Reconstruction-based (RB) contribution analysis were of interest. The objective of these approaches is to calculate an individual importance for each process variable that quantifies how much each variable contributes to the multivariate deviation. It is well documented in literature that contribution analysis suffers from a condition known as fault smearing. This condition makes it difficult to isolate the individual contribution of a process variable on the deviation observed. A simulation study, derived from a PCA model fitted on an actual chemical process, was performed to evaluate the fault diagnosis ability of the contribution analysis methodology. Based on the results of the simulation study and the mathematical results provided, it is concluded that traditional contribution analysis is not a suitable fault diagnosis methodology in a multivariate setting due to fault smearing.

For PCA based MSPM, it was argued, by means of an illustration, that the objective of traditional contribution analysis i.e., to assign individual importance to process variables, is not sensible. It was demonstrated that it is possible to arrive at the same faulty observation by inducing single sensor faults in different variables. The objective of ranking one process variable as being more important than another is therefore inappropriate.

Additionally, from a practical viewpoint, an individual ranking of process variables does not provide engineers with a clear explanation of why the process is identified by the statistical methods as being multivariately out-of-control. It is motivated that the diagnosis of multivariate faults should provide a multivariate answer to enable improved interpretation. To this end, it is proposed that a ranking should be determined that quantifies the importance of variable pairs. A novel methodology was developed that is used to obtain a pairwise variable importance ranking for faults detected using PCA based MSPM. Accompanying this methodology is an approach that can be used to identify statistically significant pairwise contributions. This methodology is also extended in order to obtain an individual ranking of process variables. It is advised that the pairwise and individual ranking of process variables should be collectively analysed. The diagnosis ability of this new methodology was demonstrated using a simulation example.

Chemical process plants that would benefit from non-linear MSPM are numerous. A review was provided to illustrate how kernel PCA is employed for MSPM. The practical

implementation of kernel PCA for fault detection and diagnosis was demonstrated using a simulated data set. The fault diagnosis techniques employed for kernel PCA are in principle similar to those methods employed for PCA. The kernel PCA diagnosis techniques attempt to attach individual importance to process variables. Therefore, the concerns raised for PCA based MSPM extends to kernel PCA. Individual importances of process variables do not provide the engineer with a clear understanding of why the statistical fault identification statistic classified the process as being different from multivariate common cause variability. A different approach to fault diagnosis in kernel PCA based monitoring is therefore required.

Similar to the discussion provided to promote the new methodology for fault diagnosis in PCA, it was illustrated for kernel PCA that the notion of attempting to assign an individual importance to process variables is not the correct question that needs to be addressed. It was therefore proposed that pairwise contributions should be calculated in kernel PCA based fault diagnosis. A new methodology was developed that can be used to calculate pairwise contributions in the non-linear setting. This methodology is dependent on the appropriate specification of distance measures. Different strategies were discussed on how to specify suitable distance measures. Specifically, distance measures were developed using kernel PCA, Support Vector Machine (SVM) classification and One-Class SVM classification. The implementation of the new diagnosis methodology for kernel PCA based fault diagnosis was illustrated using simulated data. It was demonstrated how this methodology can indicate to engineers why the process is statistically classified as being multivariately out-of-control.

The value and applicability of the new fault diagnosis methods were evaluated using data from the Tennessee Eastman benchmark simulation and from an actual chemical process plant. These case studies clearly demonstrate that the new fault diagnosis methods provide information that effectively assist with the statistical interpretation of multivariate process deviations.

7.2 Future Research

The following future research areas are identified:

- Investigate if the new fault diagnosis methodologies can be implemented for techniques other than PCA and kernel PCA based MSPM. The possible areas that should be explored are:

    1. Partial Least Squares (PLS) and kernel PLS.
       PLS (Wold et al., 1984) is a dimension reduction technique often employed for

MSPM (MacGregor et al., 1994). In Li et al. (2009) a total PLS based contribution plot is introduced for fault diagnosis. This contribution methodology is based on the decomposition of a fault detection index that assign individual importance values to process variables. In Van den Kerkhof, Vanlaer, Gins and Van Impe (2013) it is pointed out that the PLS based contribution analysis techniques also suffer from fault smearing. Kernel PLS (Rosipal and Trejo, 2002) the non-linear variant of PLS can be employed for non-linear statistical process monitoring (Wang and Shi, 2014). In Godoy et al. (2014) a gradient based approach is used for kernel PLS fault diagnosis.

2. Independent component analysis (ICA). ICA (Kano et al., 2003) is a technique that decomposes the data into statistically independent components (IC's). In Lee et al. (2006) a modified ICA approach is utilised for fault detection and diagnosis in MSPM. Contribution decomposition of the ICA T^2 and SPE statistics are used to rank process variables for fault diagnosis.

- The diagnostic methodology developed for the non-linear case can be applied to diagnose PCA based faults. It should be investigated if the new method developed for the diagnosis of non-linear process faults can be used as a generalisation of the new method used to diagnose linear PCA faults.

- In both the linear and non-linear pairwise diagnosis methodologies the possibility exist that a large number of process variable pairs can be identified as containing information on a specific fault. The consideration of many deviating pairs can potentially confuse interpretation. Strategies should be investigated to assist in the simultaneous interpretation of numerous significant pairwise contributions.

- In the pairwise contribution analysis method developed for the non-linear fault diagnosis, it is required that $\binom{p}{2}$ models should be fitted in order to determine the distance evaluations. Therefore, a large number of models need to be fitted when the number process variable p is large. The development of efficiently automated strategies are therefore required to optimally fit the models.

Appendix A

R code

The R code used in the analysis of the Tennessee Eastman (TE) data and boiler data is presented in this appendix. Note that code illustrating the detection and diagnosis of the first TE process fault will only be provided. The code can easily be adapted for the boiler data analysis and the other TE process faults.

A.1 R code for PCA based MSPM

TE PCA model definition.

```
#load IOC data
load('TEP_FaultFree_Training.RData')

#Define IOC data set
xmat <- fault_free_training[1:1500,4:55]

#IOC data dimensions
n <- nrow(xmat)
p <- ncol(xmat)

#Center and scale preprocessing
xmean <- apply(xmat,2,mean)
xsd <- apply(xmat,2,sd)
zmat <- scale(xmat,xmean,xsd)

#Covariance matrix
smat <- (1/(n-1))*t(zmat)%*%zmat

#Eigen decompositions
```

```
smat_eig <- eigen(smat)
lambda <- smat_eig$values

#PCA model dimension selection
#Scree plot
plot(lambda,type='l',ylab="eigen values",
 xlab = "number of principal components",axes=FALSE)
points(x = 1:length(lambda),y=lambda,pch=19)
axis(2)
axis(1,at = seq(0,50,5))
grid()
box()

pca_dim <- 15
var_exp <- sum(smat_eig$values[1:pca_dim])/sum(smat_eig$values)
#"Variance explained: 66.57 %"

#model loading vectors
vmat <- smat_eig$vectors[,1:pca_dim]
#residual loading vectors
vrmat <- smat_eig$vectors[,-c(1:pca_dim)]
#diagonal matrix of model eigenvalues
Lmat <- diag(lambda[1:pca_dim])

# 95 % Control limits of statistical control
# indices (T2, SPE and Combined)
#t2 critical value
t2_crit <- qchisq(p = 0.95,df = pca_dim)
#spe critical value
theta1 <- sum(lambda[-c(1:pca_dim)])
theta2 <- sum(lambda[-c(1:pca_dim)]^2)
k <- theta2/theta1
df_spe <- theta1^2/theta2
spe_crit <- k*qchisq(p = 0.95,df = df_spe)
#combined critical value
gcomb <- (pca_dim/t2_crit^2 + theta2/spe_crit^2)/
(pca_dim/t2_crit + theta1/spe_crit)
hcomb <- (pca_dim/t2_crit + theta1/spe_crit)^2/
(pca_dim/t2_crit^2 + theta2/spe_crit^2)
comb_crit <- gcomb*qchisq(p=0.95,df=hcomb)
```

```
#Square matrices of statistical control indices
T2mat <- vmat%*%solve(diag(lambda[1:pca_dim]))%*%t(vmat)
SPEmat <- vrmat%*%t(vrmat)
COMBmat <- T2mat/t2_crit + SPEmat/spe_crit
#Store results
pca_mod_obj <- list()
length(pca_mod_obj) <- 14
names(pca_mod_obj) <- c("xdat","xmean","xsd","zdat",
                        "cov_eig","pca_dim","vmat",
                        "vrmat","T2mat","SPEmat","COMBmat",
                        "t2_crit","spe_crit","comb_crit")
pca_mod_obj[["xdat"]] <- xmat
pca_mod_obj[["xmean"]] <- apply(xmat,2,mean)
pca_mod_obj[["xsd"]] <- apply(xmat,2,sd)
pca_mod_obj[["zdat"]] <- zmat
pca_mod_obj[["cov_eig"]] <- smat_eig
pca_mod_obj[["pca_dim"]] <- pca_dim
pca_mod_obj[["vmat"]] <- vmat
pca_mod_obj[["vrmat"]] <- vrmat
pca_mod_obj[["T2mat"]] <- T2mat
pca_mod_obj[["SPEmat"]] <- SPEmat
pca_mod_obj[["COMBmat"]] <- COMBmat
pca_mod_obj[["t2_crit"]] <- t2_crit
pca_mod_obj[["spe_crit"]] <- spe_crit
pca_mod_obj[["comb_crit"]] <- comb_crit
rm(list=setdiff(ls(), "pca_mod_obj"))
```

Evaluation of TE process fault one.

```
library(latex2exp)
#PCA model information
pca_mod_obj[["xdat"]] -> xmat
pca_mod_obj[["xmean"]] -> xmean
pca_mod_obj[["xsd"]] -> xsd
pca_mod_obj[["zdat"]] -> zmat
pca_mod_obj[["T2mat"]] -> T2mat
pca_mod_obj[["SPEmat"]] -> SPEmat
pca_mod_obj[["COMBmat"]] -> COMBmat
pca_mod_obj[["t2_crit"]] -> t2_crit
pca_mod_obj[["spe_crit"]] -> spe_crit
```

```
pca_mod_obj[["comb_crit"]] -> comb_crit

#Reference data information
n <- nrow(xmat)
p <- ncol(xmat)

vnmes <- colnames(xmat)

#variable names used in thesis

vnmes_xnmes <- data.frame(vnmes = vnmes,
                          xnmes =
                          paste0(paste('x',1:p,sep='_{'),'}'),
                          stringsAsFactors = FALSE)

rownames(vnmes_xnmes) <- vnmes_xnmes$vnmes

#load new data
load('TEP_Faulty_Training.RData')

#Select data corresponding to the first
#process fault
FAULT_ID <- 1
FAULT_ID_ind <- faulty_training$faultNumber == FAULT_ID
xnew <- faulty_training[FAULT_ID_ind,4:55][1:500,]

#Preprocess new data
znew <- scale(xnew,xmean,xsd)

#Calculate fault identification statistic
#values

#T2 index values
t2vals <- apply(znew,1,function(x,P){
  t(x)%*%P%*%x
},P=T2mat)

#SPE index values
spevals <- apply(znew,1,function(x,P){
  t(x)%*%P%*%x
```

```
},P=SPEmat)

#Combined index values
combvals <- apply(znew,1,function(x,P){
  t(x)%*%P%*%x
},P=COMBmat)

#Plot fault identification statistic
#evaluations

pchs <- c(rep(1,499),19)
cols <- c(rep('black',499),'red')

plot(t2vals,xlab="",axes = FALSE,
     ylab=TeX("$T^2$-values"),pch=pchs,col=cols)
abline(h=t2_crit,col="red",lwd=2,lty=2)
legend("topright",legend=TeX("$T^2$ 95% control limit"),
col="red",lwd=2,cex=0.8,lty=2)
grid()
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)

plot(spevals,xlab="",axes = FALSE,
     ylab="SPE-values",pch=pchs,col=cols)
abline(h=spe_crit,col="red",lwd=2,lty=2)
legend("topright",legend="SPE 95% control limit",
col="red",lwd=2,cex=0.8,lty=2)
grid()
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)

plot(combvals,xlab="",axes = FALSE,
     ylab="COMB-values",pch=pchs,col=cols)
abline(h=comb_crit,col="red",lwd=2,lty=2)
legend("topright",legend="COMB 95% control limit",
col="red",lwd=2,cex=0.8,lty=2)
```

```
grid()
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)

#Id of sample that will be diagnosed
SAMPLE_ID <- 500
znvec <- znew[SAMPLE_ID,]

#Identify variables that are observed
#beyond the range of the data

outvars <- NULL

for (varii in vnmes){
  qvalsi <- range(xmat[,varii])
  if (xnew[SAMPLE_ID,varii] < qvalsi[1] |
  xnew[SAMPLE_ID,varii] > qvalsi[2]){
    outvars <- c(outvars,varii)
  }
}

univ_out_vars <- vnmes_xnmes[outvars,'xnmes']

#Traditional contribution analysis

#Complete decomposition
o_cd <- CD_func(Amat=SPEmat,znvec)
o_cd_out <- order(o_cd,decreasing=TRUE)

#Partial decomposition
o_pd <- PD_func(Amat=SPEmat,znvec)
o_pd_out <- order(o_pd,decreasing=TRUE)

#Reconstruction based decomposition
o_rb <- RB_func(Amat=SPEmat,znvec)
o_rb_out <- order(o_rb,decreasing=TRUE)

#Select top 10 variables
```

```
cd_10names <- vnmes_xnmes$xnmes[o_cd_out[1:10]]
cd_10graphs_nms <- paste0('$',cd_10names,'$')

pd_10names <- vnmes_xnmes$xnmes[o_pd_out[1:10]]
pd_10graphs_nms <- paste0('$',pd_10names,'$')

rb_10names <- vnmes_xnmes$xnmes[o_rb_out[1:10]]
rb_10graphs_nms <- paste0('$',rb_10names,'$')

#Plot contribution analysis results

barplot(height = o_cd[o_cd_out[1:10]],
        names.arg = TeX(cd_10graphs_nms),
        cex.names = 1.5,main = "CD␣contribution")
box()
grid()

barplot(height = o_pd[o_pd_out[1:10]],
        names.arg = TeX(pd_10graphs_nms),
        cex.names = 1.5,main = "PD␣contribution")
box()
grid()

barplot(height = o_rb[o_rb_out[1:10]],
        names.arg = TeX(rb_10graphs_nms),
        cex.names = 1.5,main = "RB␣contribution")
box()
grid()

# Univariate contribution analysis using the new method
# row sums distributional analysis

#get distance matrix representation of new observation
#based on the SPE statistic

zndist <- dist_calc(x = znvec,varnames = colnames(znew),
                Dmat = SPEmat)

#MDS analysis of distance matrix
loc0 <- cmdscale(zndist)
```

```
lbls <- paste0("$x_{",1:p,"}^*$")
plot(loc0,xlim=c(-150,150),ylim=c(-150,150),pch=19,
xlab="MDS dimension 1",ylab="MDS dimension 2")
text(x=loc0[,1],y=loc0[,2],labels=TeX(lbls),pos=2)
abline(v=0,h=0,lty=2)
grid()

#identify row sums that are significantly high when compared
# to the reference information

#First we need to collect the reference distance matrix
# information

ref_dist_list <- ref_dist_matrices(xref = zmat,
        varnames = colnames(zmat),Dmat = SPEmat)

#iteratively calculate row sums and compare to corresponding
#reference rowsum
#remove row and column from new dist after identifying a
#significantly large rowsum
#remove corresponding column in reference data

#At each step check if the maximum row sum is significantly large
#when compared to the reference distribution

#Univariate ranking
vars_identified <- identify_univ_vars(xnew_dist = zndist,
                xref_dist = ref_dist_list)

#pairwise contribution analysis

#Pairwise IOC information based on SPE statistic
pairwise_ref_ecdf <- pairwise_ref_info(pca_mod_list = pca_mod_obj,
                        statistic='SPE')

#Pairwise contribution based on SPE statistic
out <- get_pairwise_contrib_stat(xvec=znvec,
                          varnames=vnmes,
                          pca_mod=pca_mod_obj,
                          reference_dists=pairwise_ref_ecdf,
```

```
                                stat = "SPE")

contrib_vec <- c(contrib_vec,colnames(out)[which(out < 0.01)])
```

Definition of functions used in the above code listing.

```
#Complete decomposition
CD_func <- function(Amat,x_vek){
  #Amat: positive definite matrix of fault stat
  #x_vek: observed sample
  x_vek <- matrix(x_vek,ncol=1,byrow=FALSE)
  p <- nrow(Amat)
  Imat <- diag(p)
  svdAmat <- svd(Amat)
  Amat05 <-  svdAmat$u%*%(diag(svdAmat$d)^0.5)%*%t(svdAmat$v)
  CDi <- matrix(NA,ncol=1,nrow=p)
  for(i in 1:p){
    xi_vek <- as.matrix(Imat[,i])
    CDi[i] <- (t(xi_vek)%*%Amat05%*%x_vek)^2
  }
  return(CDi)
}
#Partial decomposition
PD_func <- function(Amat,x_vek){
  #Amat: positive definite matrix of fault stat
  #x_vek: observed sample
  x_vek <- matrix(x_vek,ncol=1,byrow=FALSE)
  p <- nrow(Amat)
  Imat <- diag(p)
  PDi <- matrix(NA,ncol=1,nrow=p)
  for(i in 1:p){
    xi_vek <- as.matrix(Imat[,i])
    PDi[i] <- t(x_vek)%*%Amat%*%xi_vek%*%t(xi_vek)%*%x_vek
  }
  return(PDi)
}
#Reconstruction based contribution
RB_func <- function(Amat,x_vek){
  #Amat: positive definite matrix of fault stat
  #x_vek: observed sample
  x_vek <- matrix(x_vek,ncol=1,byrow=FALSE)
```

```
  p <- nrow(Amat)
  Imat <- diag(p)
  RBi <- matrix(NA,ncol=1,nrow=p)
  for(i in 1:p){
    xi_vek <- as.matrix(Imat[,i])
    RB1 <- (t(xi_vek)%*%Amat%*%x_vek)^2
    RB2 <- t(xi_vek)%*%Amat%*%xi_vek
    RBi[i] <- RB1/RB2
  }
  return(RBi)
}
#Pairwise distance matrix
dist_calc <- function(x,varnames,Dmat){
  #x: vector input data (appropriately scaled)
  #varnames: vector variable names
  #Dmat: PCA statistic matrix
  p <- length(varnames)
  #variable combinations
  var_combs <- t(combn(x = 1:p,m = 2))
  #distance for each variable pair
  dist_pairs <- apply(var_combs,1,function(xvar,xvec,D0){
    ii <- xvar[1]
    jj <- xvar[2]
    Dij <- sub_D(D = D0,i = ii,j = jj)
    return( t(xvec[c(ii,jj)])%*%Dij%*%xvec[c(ii,jj)] )
  },xvec = x,D0 = Dmat)

  var_dist_pairs <- cbind(var_combs,dist_pairs)
  #convert to a distance matrix
  out_dist <- to_dist_matrix_func(var_dist_pairs)
  colnames(out_dist) <- rownames(out_dist) <- varnames
  return(out_dist)
}
#get 2x2 submatrix
sub_D <- function(D,i,j){
  Dii <- D[i,i]
  Djj <- D[j,j]
  Dij <- D[i,j]
  out <- matrix(c(Dii,Dij,Dij,Djj),ncol=2,nrow=2,byrow = TRUE)
  return(out)
```

```
}
#convert pairwise distances to distance matrix
to_dist_matrix_func <- function(X){
  #X: contains all distances of all possible pairwise
  #variable combinations
  p <- max(X[,1:2])
  Dmat <- matrix(0,ncol = p,nrow = p)
  m <- nrow(X)
  for(i in 1:m){
    ii <- X[i,1]
    jj <- X[i,2]
    Dmat[ii,jj] <- X[i,3]
  }
  Dmat <- Dmat + t(Dmat)
  return(Dmat)
}
#Calculate distance matrix for each reference sample
#store each row of each calculated in a separate matrix
#i.e. collect variable i's information separately
ref_dist_matrices <- function(xref,varnames,Dmat){
  #xref: reference matrix
  #varnames: vector variable names
  #Dmat: PCA statistic matrix
  n <- nrow(xref)
  p <- ncol(xref)
  outlist <- list()
  length(outlist) <- p
  #assign empty matrix to each list position
  outlist <- lapply(X = outlist,FUN = function(x,vnms){
    o <- matrix(0,ncol=p,nrow=0)
    colnames(o) <- vnms
    return(o)
  },vnms = varnames)

  for(i in 1:n){
    disti <- dist_calc(x = xref[i,],varnames = varnames,Dmat = Dmat)
    for(k in 1:p){
      outlist[[k]] <- rbind(outlist[[k]],disti[k,])
    }
  }
```

```
  names(outlist) <- varnames
  return(outlist)
}
#Calculate univariate ranking
identify_univ_vars <- function(xnew_dist,xref_dist){
        #xnew_dist: distance matrix of new observation
        #xref_dist: reference distance information
  vars_identified <- NULL
  VARS_FOUND <- FALSE

  while(!VARS_FOUND){
    if(is.null(vars_identified)){
      initial_row_sum <- rowSums(xnew_dist)
      initial_vars <- rownames(xnew_dist)
    }else{
      znnames <- colnames(xnew_dist)
      id <- !(znnames %in% vars_identified)
      initial_row_sum <- rowSums(xnew_dist[id,id])
      initial_vars <- rownames(xnew_dist[id,id])
    }
    max_name <- names(which.max(initial_row_sum))
    max_value <- max(initial_row_sum)
    #reference row sums for the selected process variable
    refrowsum_max_name <- apply(xref_dist[[max_name]],1,sum)
    if (1-ecdf(refrowsum_max_name)(max_value) < 0.05){
      vars_identified <- c(vars_identified,max_name)
    }else{
      VARS_FOUND <- TRUE
    }
  }
  return(vars_identified)
}
#Pairwise reference information
pairwise_ref_info <- function(pca_mod_list,statistic='T2'){
  #pca_mod_list: pca model information
  xmat <- pca_mod_list[["xdat"]]
  zmat <- pca_mod_list[["zdat"]]
  vnmes <- colnames(xmat)
  p <- ncol(xmat)
  n <- nrow(xmat)
```

```
if(statistic == 'T2'){
  Dmat <- pca_mod_list[["T2mat"]]
  critical_value <- pca_mod_list[["t2_crit"]]
}else if(statistic == 'SPE'){
  Dmat <- pca_mod_list[["SPEmat"]]
  critical_value <- pca_mod_list[["spe_crit"]]
}else if(statistic == 'COMB'){
  Dmat <- pca_mod_list[["COMBmat"]]
  critical_value <- pca_mod_list[["comb_crit"]]
}

statvals_ref <- apply(zmat,1,function(x,Pmat){
  t(x)%*%Pmat%*%x
},Pmat=Dmat)
incontrol_ids <- statvals_ref <= critical_value
zref <- zmat[incontrol_ids,] #scaled reference data
xref <- xmat[incontrol_ids,]
ref_dists <- list()
length(ref_dists) <- nrow(zref)
varid_combs <- t(combn(1:p,2))
var_combs <- t(combn(vnmes,2))

for(ii in 1:nrow(zref)){
  zfii <- zref[ii,]
  dd <- apply(varid_combs,1,function(x,D,zfvec){
    ii <- x[1]
    jj <- x[2]
    Dij <- sub_D(D = D,i = ii,j = jj)
    t(zfvec[c(ii,jj)])%*%Dij%*%zfvec[c(ii,jj)]
  },D=Dmat,zfvec=zfii)
  out_dist <- to_dist_matrix_func(cbind(varid_combs,dd))
  colnames(out_dist) <- rownames(out_dist) <- vnmes
  ref_dists[[ii]] <- out_dist
}
#Calculate empirical distributions
pos_name <- apply(var_combs,1,function(x){
  paste0("(",paste0(x,collapse=","),")")
})
rownames(var_combs) <- pos_name
```

```
  rownames(varid_combs) <- pos_name
  varij_comb_list <- list()
  length(varij_comb_list) <- length(pos_name)
  names(varij_comb_list) <- pos_name

  for(ij in pos_name){
    k <- varid_combs[ij,1]
    l <- varid_combs[ij,2]
    varij_comb_list[[ij]] <- unlist(lapply(X=ref_dists,
                FUN=function(x,k,l){x[k,l]},k=k,l=l))
  }

  pair_ij_ecdf <- lapply(X=varij_comb_list,FUN=function(x){
    o <- ecdf(x)
    return(o)
  }
  )
  return(pair_ij_ecdf)
}
#Calculate pairwise contributions
get_pairwise_contrib_stat <- function(xvec,varnames,pca_mod,
                                      reference_dists,stat="T2"){
  #xvec: new observation on the original scale (vector)
  #varnames: variable names
  #pca_mod: list object containing pca model
  #information required for process monitoring
  #containing objects: "xdat","xmean","xsd","zdat","cov_eig",
  # "pca_dim","vmat","vrmat","T2mat","SPEmat",
  # "COMBmat","t2_crit","spe_crit","comb_crit"
  #reference_dists: empirical distribution
  #function of in-control distances for each index.
  #stat: fault identification statistic selected.

  p <- length(xvec)
  varid_combs <- t(combn(1:p,2))
  var_combs <- t(combn(varnames,2))

  #dist matrix position names
  pos_name <- apply(var_combs,1,function(x){
    paste0("(",paste0(x,collapse=","),")")
```

```
  })
  #preprocessing
  xvec <- matrix(xvec,ncol=1,nrow=p)
  zvec <- t(scale(x=t(xvec),center=pca_mod$xmean,
                  scale=pca_mod$xsd))

  #Calculate contributions
  #based on out-of-control index
  if(stat == "T2"){
    DMAT <- pca_mod$T2mat
  }else if(stat == "SPE"){
    DMAT <- pca_mod$SPEmat
  }else if(stat == "COMB"){
    DMAT <- pca_mod$COMBmat
  }
  contrib_by_stat <- matrix(0,ncol=choose(n=p,k=2),nrow=1)
  colnames(contrib_by_stat) <- pos_name

  #Calculate distance matrix
  dd <- apply(varid_combs,1,function(x,D,zf){
    ii <- x[1]
    jj <- x[2]
    Dij <- sub_D(D = D,i = ii,j = jj)
    t(zf[c(ii,jj)])%*%Dij%*%zf[c(ii,jj)]
  },D=DMAT,zf = zvec)
  out_dist <- to_dist_matrix_func(cbind(varid_combs,dd))
  colnames(out_dist) <- rownames(out_dist) <- varnames

  #Calculate occurance probabilities
  contribs <- eval_ecdfs(dmat=out_dist,ecdf_mods=reference_dists)
  contrib_by_stat[1,] <- contribs[lower.tri(contribs)]
  return(contrib_by_stat)
}
```

A.2 R code for kernel PCA based MSPM.

KPCA hyperparameter selection.

```
library(kernlab)
library(Rsolnp)
```

```
library(snowfall)

#load data
load('TEP_data/dataverse_files/TEP_FaultFree_Training.RData')
#specify IOC data
xdat <- fault_free_training[1:500,4:55]
xdat <- xdat[,!(colnames(xdat) %in% c("xmv_7","xmv_8"))]

n <- nrow(xdat)
p <- ncol(xdat)

#specify search grid
sigma_seq <- c(0.001,0.002,0.003,0.004,0.008)
dim_seq <- c(5,15,50,80,100)
sigma_dim_grid <- expand.grid(sigma_seq,dim_seq)

#samples that will be used to evaluate error
kmout <- kmeans(x=xdat,centers=5)
kmdat <- kmout$centers

#function to optimize
ff <- function(zpre,tz,zdat_i,Wtilde_i,L_i,ss_i){
  zpre_score <- pred_new_score(zpre,zdat_i,Wtilde_i,L_i,ss_i)
  o <- t(tz)%*%zpre_score
  return(-1*o[1,1])
}

sfInit(parallel=TRUE,cpus=25,type = 'SOCK')
sfLibrary("snowfall", character.only=TRUE)
sfLibrary("kernlab", character.only=TRUE)
sfLibrary("Rsolnp", character.only=TRUE)

sfExport(list = c('sigma_dim_grid','kmdat','xdat',
'kpca_model','pred_new_score','ff'))

cv_values <- NULL

for(i in 1:nrow(sigma_dim_grid)){

  hh_i <- sigma_dim_grid[i,1]
```

```
  dim_i <- sigma_dim_grid[i,2]

  cvout <- sfApply(x = kmdat,margin=1,fun = function(xi,hh,dim){

    p <- ncol(xdat)
    xi <- matrix(xi,ncol=p,nrow=1)
    zdat <- scale(xdat)
    o <- kpca_model(x = xdat,hp = hh,dim=dim)
    zi <- scale(xi,center = o$xmean,scale = o$xsd)
    tzi <- pred_new_score(zi,zdat,o$Wtilde,o$Lmat,o$hp)

    out <- solnp(pars = rep(0,p),fun = ff,LB = apply(zdat,2,min),
                 UB = apply(zdat,2,max),tz = tzi,zdat_i = zdat,
                 Wtilde_i = o$Wtilde, L_i = o$Lmat, ss_i = o$hp)

    optval <- tail(out$values,1)
    optval <- sum((zi - out$pars)^2)
    return(optval)
  },hh=hh_i,dim=dim_i)

  cv_values <- c(cv_values,mean(cvout))
}
sfStop()
#combine results
dat <- data.frame(sigma_dim_grid,cv=cv_values)
```

Fit KPCA model.

```
library(kernlab)
library(snowfall)
source("kpca_funcs.R")

#load data
load('TEP_data/dataverse_files/TEP_FaultFree_Training.RData')

xdat <- fault_free_training[1:500,4:55]
xdat <- xdat[,!(colnames(xdat) %in% c("xmv_7","xmv_8"))]

n <- nrow(xdat)
p <- ncol(xdat)
```

```
#preprocessing
xmean <- apply(xdat,2,mean)
xsd <- apply(xdat,2,sd)
zdat <- scale(xdat,xmean,xsd)

#kernel matrix and eigen-decomposition
SS <- 0.003
Kmat <- kernelMatrix(rbfdot(sigma = SS), zdat)
Kmat <- (1/(n-1))*Kmat
Kbar <- (diag(n) - (1/n)*onevec%*%t(onevec))%*%Kmat%*%(diag(n) -
 (1/n)*onevec%*%t(onevec))
Ksvd <- svd(Kbar)

#Loading matrix information
kpca_dim <- 80
Lmat <- diag(Ksvd$d[1:kpca_dim])
Lhalf <- diag(Ksvd$d[1:kpca_dim]^(.5))
Wmat <- Ksvd$v[,1:kpca_dim]
Wmat <- sqrt(n-1)*Wmat%*%Lhalf
Tmat <- t(solve(Lmat)%*%t(Wmat)%*%Kbar)

#Calculate critical values
#T2 statistic
t2vals <- apply(Tmat,1,function(x){
  x <- (as.matrix(x))
  t(x)%*%solve(Lmat)%*%x
})
t2_crit <- quantile(t2vals,0.95)

#SPE statistic
fInit(parallel=TRUE,cpus=30,type = 'SOCK')
sfLibrary("snowfall", character.only=TRUE)
sfLibrary("kernlab", character.only=TRUE)
sfExport(list=c('zdat','SS','pred_new_score',
'spe_value','Wmat','Kmat','Lmat'))

SPEvals <- sfApply(x=zdat,margin=1,fun=spe_value,zdat = zdat,
Kmat = Kmat,SS = SS,W_tilde = Wmat,L = Lmat)

sfStop()
```

```
spe_crit <- quantile(SPEvals,0.95)

#Combined statistic
cmb_vals <- t2vals/t2_crit + SPEvals/spe_crit
cmb_crit <- quantile(cmb_vals,0.95)

#Save model information
save(xdat,xmean,xsd,zdat,SS,Kmat,Kbar,
kpca_dim,Lmat,Lhalf,Wmat,Tmat,t2_crit,
spe_crit,cmb_crit,
file='tep_kpca_model_information.Rdata')
```

Fit SVM pairwise models.

```
library(ks)
library(kernlab)
library(snowfall)

source('fit_svm.R')
#Reference data
load('TEP_data/dataverse_files/TEP_FaultFree_Training.RData')
xdat <- fault_free_training[1:500,4:55]
xdat <- xdat[,!(colnames(xdat) %in% c("xmv_7","xmv_8"))]
#Fit models
tep_pairwise_svm_models <- fit_pairwise_svm_models(xdat=xdat)

save(tep_pairwise_svm_models,file='tep_pairwise_svm_models.Rdata')
```

TE Fault 1 evaluation.

```
library(xtable)
library(latex2exp)
library(kernlab)
library(numDeriv)
source('kpca_funcs.R')

#load model info
load('tep_kpca_model_information.Rdata')

#load pairwise model info
load('tep_pairwise_svm_models.Rdata')
```

```
#load fault data
load('TEP_data/dataverse_files/TEP_Faulty_Training.RData')

FAULT_ID <- 1
FAULT_IND <- faulty_training$faultNumber == FAULT_ID
xnew <- faulty_training[FAULT_IND,4:55][1:500,]
xnew <- xnew[,!(colnames(xnew) %in% c("xmv_7","xmv_8"))]
znew <- scale(xnew,center=xmean,scale=xsd)

#sample that will be diagnosed
SAMPLE_ID <- 500

#Fault detection index evaluations

#T2
sfInit(parallel=TRUE,cpus=30,type = 'SOCK')
sfLibrary("snowfall", character.only=TRUE)
sfLibrary("kernlab", character.only=TRUE)
sfExport(list=c('znew','zdat','Wmat','Lmat','SS','t2_value'))

t2vals_new <- t(sfApply(znew,1,function(x){

  xj <- t2_value(znew = x,zdat = zdat,W_tilde = Wmat,
                 L = Lmat,ss = SS)
  xj
}))
sfStop()

#SPE
sfInit(parallel=TRUE,cpus=30,type = 'SOCK')
sfLibrary("snowfall", character.only=TRUE)
sfLibrary("kernlab", character.only=TRUE)
sfExport(list=c('znew','zdat','Wmat','Lmat','SS',
                'pred_new_score','Kmat','spe_value'))

spe_evals <- t(sfApply(znew,1,function(x){

  xj <- spe_value(znew = x,
                  zdat = zdat,
                  Kmat = Kmat,
```

```
                    W_tilde = Wmat,
                    L = Lmat,
                    SS = SS)
  xj
}))
sfStop()

#Combined
sfInit(parallel=TRUE,cpus=30,type = 'SOCK')
sfLibrary("snowfall", character.only=TRUE)
sfLibrary("kernlab", character.only=TRUE)
sfExport(list=c('znew','zdat','Wmat','Lmat','SS',
                'spe_crit','t2_crit','COMB_value',
                'pred_new_score','Kmat'))
comb_evals <- t(sfApply(znew,1,function(x){

  xj <- COMB_value(znew = x,
                   zdat = zdat,
                   Kmat = Kmat,
                   W_tilde = Wmat,
                   L = Lmat,
                   th_t2_crit = t2_crit,
                   th_spe_crit = spe_crit,
                   SS = SS)
  xj
}))
sfStop()
#Plot fault identification statistic
#evaluations

pchs <- c(rep(1,499),19)
cols <- c(rep('black',499),'red')

#Combined
plot(comb_evals[1,],xlab="",axes = FALSE,
     ylab="COMB-values",pch=pchs,col=cols)
abline(h=cmb_crit,col="red",lwd=2,lty=2)
legend("bottomright",legend="COMB␣95%␣control␣limit",
col="red",lwd=2,cex=0.8,lty=2)
grid()
```

```
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)

#SPE
plot(spe_evals[1,],xlab="",axes = FALSE,
     ylab="SPE-values",pch=pchs,col=cols)
abline(h=spe_crit,col="red",lwd=2,lty=2)
legend("bottomright",legend="SPE 95% control limit",
col="red",lwd=2,cex=0.8,lty=2)
grid()
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)

#T2
plot(t2vals_new[1,],xlab="",axes = FALSE,
     ylab=TeX("$T^2$-values"),pch=pchs,col=cols)
abline(h=t2_crit,col="red",lwd=2,lty=2)
legend("bottomright",legend=TeX("$T^2$ 95% control limit"),
col="red",lwd=2,cex=0.8,lty=2)
grid()
axis(2)
axis(1)
box()
abline(v = 20,lwd=2,col='blue',lty=2)
#Traditional fault diagnosis
#Gradient based approach
cho_t2_contrib <- NULL
cho_spe_contrib <- NULL
cho_comb_contrib <- NULL

for(i in 1:ncol(znew)){

  ot2 <- grad(func = t2func,x = 1,ind=i,xnew=znew[SAMPLE_ID,],
              zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS,
              method = "simple")
  ospe <- grad(func = spefunc,ind = i,x = 1,xnew=znew[SAMPLE_ID,],
```

```
                zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS,
                method = "simple")
  ocomb <- grad(func = cmbfunc,ind = i,x = 1,xnew=znew[SAMPLE_ID,],
                zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS,t2lim=t2_crit,
                spelim=spe_crit,method = "simple")[1,1]

  cho_t2_contrib <- c(cho_t2_contrib,ot2)
  cho_spe_contrib <- c(cho_spe_contrib,ospe)
  cho_comb_contrib <- c(cho_comb_contrib,ocomb)

}

cho_t2_contrib <- abs(cho_t2_contrib)
cho_spe_contrib <- abs(cho_spe_contrib)
cho_comb_contrib <- abs(cho_comb_contrib)
names(cho_t2_contrib) <- colnames(xdat)
names(cho_spe_contrib) <- colnames(xdat)
names(cho_comb_contrib) <- colnames(xdat)

vnames <- sapply(c(1:47,50,51,52),function(i){
  xi <- paste0("x[",i,"]")
  xi <- parse(text = xi)
})

#Plot top ten contributors
barplot(cho_spe_contrib[order(cho_spe_contrib,decreasing=TRUE)],
        names.arg = vnames[order(cho_spe_contrib,decreasing=TRUE)],
        ylab="",main= 'Gradient␣based␣contribution')
grid()
box()

#Reconstruction based approach
out <- NULL
recon_t2_out <- NULL
recon_spe_out <- NULL
recon_comb_out <- NULL

for(i in 1:ncol(xnew)){

  ot2 <- optimise(f = reconT2func,
```

```
                    interval = apply(zdat,2,range)[,i],
                    ind = i,xnew=znew[SAMPLE_ID,],
                    zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS)
  ospe <- optimise(f = reconSPEfunc,
                     interval = apply(zdat,2,range)[,i],
                     ind = i,xnew=znew[SAMPLE_ID,],
                     zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS)
  ocomb <- optimise(f = reconcmbfunc,
                      interval = apply(zdat,2,range)[,i],
                      ind = i,xnew=znew[SAMPLE_ID,],
                      zdat=zdat,Wmat=Wmat,Lmat=Lmat,SS=SS,
                      t2lim=t2_crit,spelim=spe_crit)

  ot2 <- round(ot2$minimum,4)
  ospe <- round(ospe$minimum,4)
  ocomb <- round(ocomb$minimum,4)

  if(ot2 %in% vari_range){
    ot2 <- znew[SAMPLE_ID,i]
  }
  if(ospe %in% vari_range){
    ospe <- znew[SAMPLE_ID,i]
  }
  if(ocomb %in% vari_range){
    ocomb <- znew[SAMPLE_ID,i]
  }

  recon_t2_out <- c(recon_t2_out,ot2)
  recon_spe_out <- c(recon_spe_out,ospe)
  recon_comb_out <- c(recon_comb_out,ocomb)
}

alcala_t2_contrib <- abs(recon_t2_out - znew[SAMPLE_ID,])
alcala_spe_contrib <- abs(recon_spe_out - znew[SAMPLE_ID,])
alcala_comb_contrib <- abs(recon_comb_out - znew[SAMPLE_ID,])

#Plot top ten contributors
barplot(alcala_spe_contrib[order(alcala_spe_contrib,decreasing=TRUE)],
        names.arg = vnames[order(alcala_spe_contrib,decreasing=TRUE)],
        ylab="",main='Reconstruction␣based␣contribution')
```

```
box()
grid()

#Pairwise fault diagnosis

#calculate distance matrix
dd <- calc_dist_matrix_svm(znew=xnew[SAMPLE_ID,],
                           pairwise_models = tep_pairwise_svm_models)
rownames(dd) <- colnames(dd) <- c(1:47,50,51,52)
dd[dd < 0.5] <- 0

#MDS display
loc0 <- cmdscale(dd)
lbls <- paste0("$x_{",rownames(dd),"}^*$")
plot(loc0,xlim=c(-1,1),ylim=c(-1,1),pch=19,
     xlab="MDS␣dimension␣1",
     ylab="MDS␣dimension␣2")
text(x=loc0[,1],y=loc0[,2],labels=TeX(lbls),pos=2)
abline(v=0,h=0,lty=2)
grid()

#Individual rankings based on pairwise information
mm <- dd
univ_contrib <- NULL

while(nrow(mm) > 2 & any(mm > 0) & max(rowSums(mm)) > 1){
  univ_contrib_mm <- apply(mm,1,sum)
  xi <- names(which.max(univ_contrib_mm))
  univ_contrib <- c(univ_contrib,xi)
  sub_var_ids <- !(rownames(dd) %in% univ_contrib)
  mm <- dd[sub_var_ids,sub_var_ids]
}
print(univ_contrib)
```

Definition of functions used in the above code listing.

```
#KPCA T2 statistic values
t2_value <- function(znew,zdat,W_tilde,L,ss){
  #znew: new data vector centered and scaled
  #zdat: ioc data centered and scaled
  #W_tilde: scaled kpca eigen vectors
```

```
  #ss: gaussian kernel hyperparameter
  n <- nrow(zdat)
  z <- rbind(znew,zdat)
  Kmat <- kernelMatrix(rbfdot(sigma = ss), z)
  kz <- (as.matrix(Kmat[,1]))
  kz <- kz[-1,]
  kz <- (1/(n-1))*kz
  K <- kernelMatrix(rbfdot(sigma = ss), zdat)
  K <- (1/(n-1))*K
  onevec <- as.matrix(rep(1,n))
  kzbar <- (diag(n) - (1/n)*onevec%*%t(onevec))%*%
  (kz - (1/n)*K%*%onevec)
  tnew <- solve(L)%*%t(W_tilde)%*%kzbar
  tnew <- (as.matrix(tnew))
  t2val <- t(tnew)%*%solve(L)%*%tnew
  return(t2val[1,1])
}

#KPCA score calculation
pred_new_score <- function(znew,zdat,W_tilde,L,ss){
  #znew: new data vector centered and scaled
  #zdat: ioc data centered and scaled
  #W_tilde: scaled kpca eigen vectors
  #L: diagonal eigenvalue matrix
  #ss: gaussian kernel hyperparameter
  n <- nrow(zdat)
  z <- rbind(znew,zdat)
  Kmat <- kernelMatrix(rbfdot(sigma = ss), z)
  kz <- (as.matrix(Kmat[,1]))
  kz <- kz[-1,]
  kz <- (1/(n-1))*kz
  K <- kernelMatrix(rbfdot(sigma = ss), zdat)
  K <- (1/(n-1))*K
  onevec <- as.matrix(rep(1,n))
  kzbar <- (diag(n) - (1/n)*onevec%*%t(onevec))%*%
  (kz - (1/n)*K%*%onevec)
  tnew <- solve(L)%*%t(W_tilde)%*%kzbar
  return(tnew)
}
#KPCA SPE statistic values
```

```
spe_value <- function(znew,zdat,Kmat,W_tilde,L,SS){
  #znew: new data vector centered and scaled
  #zdat: ioc data centered and scaled
  #W_tilde: scaled kpca eigen vectors
  #L: diagonal eigenvalue matrix
  #ss: gaussian kernel hyperparameter
  #Kmat: kernel matrix
  tvec <- pred_new_score(znew = znew,zdat = zdat,
  W_tilde = W_tilde,L = L,ss=SS)
  n <- nrow(zdat)
  x <- rbind(znew,zdat)
  K <- kernelMatrix(rbfdot(sigma = SS), x)
  kz <- (as.matrix(K[,1]))
  kz <- kz[-1,]
  spe_val<-1+((-2/n)*sum(kz)+((n-1)/(n^2))*sum(Kmat))-t(tvec)%*%tvec
  return(spe_val[1,1])
}
#KPCA Combined statistic values
COMB_value <- function(znew,zdat,Kmat,W_tilde,L,SS,
                       th_t2_crit,th_spe_crit){
  #znew: new data vector centered and scaled
  #zdat: ioc data centered and scaled
  #W_tilde: scaled kpca eigen vectors
  #L: diagonal eigenvalue matrix
  #ss: gaussian kernel hyperparameter
  #th_t2_crit: T2 critical value
  #th_spe_crit: SPE critical value
  #Kmat: kernel matrix
  tvec <- pred_new_score(znew = znew,zdat = zdat,
        W_tilde = W_tilde,L = L,ss=SS)
  t2val <- t(tvec)%*%solve(L)%*%tvec
  n <- nrow(zdat)
  x <- rbind(znew,zdat)
  K <- kernelMatrix(rbfdot(sigma = SS), x)
  kz <- (as.matrix(K[,1]))
  kz <- kz[-1,]
  spe_val<-1+((-2/n)*sum(kz)+((n-1)/(n^2))*sum(Kmat))-t(tvec)%*%tvec
  comb_val <- t2val/th_t2_crit + spe_val/th_spe_crit
  return(comb_val)
}
```

```
#T2 function used to in gradient based contributions
t2func <- function(x,ind,xnew,zdat,Wmat,Lmat,SS){
  p <- length(xnew)
  v <- rep(1,p)
  v[ind] <- x
  xs <- v*xnew
  #Calc the score representation
  ts <- pred_new_score(znew = xs,zdat = zdat,
        W_tilde = Wmat,L = Lmat,ss = SS)
  ts <- matrix(ts,ncol=1)
  o <- t(ts)%*%solve(Lmat)%*%ts
  return(o[1,1])
}
#SPE function used to in gradient based contributions
spefunc <- function(x,ind,xnew,zdat,Wmat,Lmat,SS){
  p <- length(xnew)
  v <- rep(1,p)
  v[ind] <- x
  xs <- v*xnew
  o <- spe_value(znew = xs,zdat = zdat,
        W_tilde=Wmat,Kmat = Kmat,L=Lmat,SS = SS)
  return(o)
}
#Combined function used to in gradient based contributions
cmbfunc <- function(x,ind,xnew,zdat,Wmat,Lmat,SS,t2lim,spelim){
  p <- length(xnew)
  v <- rep(1,p)
  v[ind] <- x
  xs <- v*xnew
  speval <- spe_value(znew = xs,zdat = zdat,Kmat = Kmat,
        W_tilde=Wmat,L=Lmat,SS = SS)
  ts <- pred_new_score(znew = xs,zdat = zdat,
        W_tilde = Wmat,L = Lmat,ss = SS)
  ts <- matrix(ts,ncol=1)
  t2val <- t(ts)%*%solve(Lmat)%*%ts
  o <- speval/spelim + t2val/t2lim
  return(o)
}
#T2 function used in the reconstruction contributions
reconT2func <- function(k,ind,xnew,zdat,Wmat,Lmat,SS){
```

```
  xs <- xnew
  xs[ind] <- k
  #Calc the score representation
  ts <- pred_new_score(znew = xs,zdat = zdat,
        W_tilde = Wmat,L = Lmat,ss = SS)
  ts <- matrix(ts,ncol=1)
  o <- t(ts)%*%solve(Lmat)%*%ts
  return(o[1,1])
}
#SPE function used in the reconstruction contributions
reconSPEfunc <- function(k,ind,xnew,L,zdat,Wmat,Lmat,SS){
  xs <- xnew
  xs[ind] <- k
  o <- spe_value(znew = xs,zdat = zdat,Kmat = Kmat,
        W_tilde=Wmat,L=Lmat,SS = SS)
  return(o)
}

#Combined function used in the reconstruction contributions
reconcmbfunc <- function(k,ind,xnew,zdat,Wmat,Lmat,SS,t2lim,spelim){
  xs <- xnew
  xs[ind] <- k
  speval <- spe_value(znew = xnew,zdat = zdat,Kmat = Kmat,
                      W_tilde=Wmat,L=Lmat,SS = SS)
  ts <- pred_new_score(znew = xs,zdat = zdat,
        W_tilde = Wmat,L = Lmat,ss = SS)
  ts <- matrix(ts,ncol=1)
  t2val <- t(ts)%*%solve(Lmat)%*%ts
  o <- speval/spelim + t2val/t2lim
  return(o)
}
#SVM distance matrix calculation
calc_dist_matrix_svm <- function(znew,pairwise_models){
  #znew: new observation
  #pairwise_models: list of svm models
  p <- length(znew)
  pair_ids <- t(combn(1:p,2))
  Dmat <- matrix(0,ncol=p,nrow=p)
  #for each variable combination calculate
  #distance to IOC point cloud for new observation
```

```
  for(k in 1:nrow(pair_ids)){
    varij <- pair_ids[k,]#variable combination
    ii <- varij[1]
    jj <- varij[2]
    #select appropriate model
    modij <- pairwise_models[[paste0(ii,',',jj)]]
    dat <- data.frame(znew[varij])
    rownames(dat) <- NULL
    #evaluate new observation
    oo <- predict(modij,dat,type="probabilities")
    #store in matrix the probability of
    #being in the residual region
    Dmat[ii,jj] <- Dmat[jj,ii] <- oo[,2]
  }
  return(Dmat)
}
```

Bibliography

Alam, M. A. and Fukumizu, K. (2014), 'Hyperparameter selection in kernel principal component analysis', *Journal of Computational Science* **10**, 1139–1150.

Alcala, C. F. and Qin, S. J. (2009), 'Reconstruction-based contribution for process monitoring', *Automatica* **45**(7), 1593 – 1600.

Alcala, C. F. and Qin, S. J. (2010), 'Reconstruction-based contribution for process monitoring with kernel principal component analysis', *Industrial and Engineering Chemistry Research* **49**, 7849–7857.

Alcala, C. F. and Qin, S. J. (2011), 'Analysis and generalization of fault diagnosis methods for process monitoring', *Journal of Process Control* **21**(3), 322 – 330. Thomas McAvoy Festschrift.

Bencheikh, F., Harkat, M., Kouadri, A. and Bensmail, A. (2020), 'New reduced kernel PCA for fault detection and diagnosis in cement rotary kiln', *Chemometrics and Intelligent Laboratory Systems* **204**, 104091.

Bolton, R. N., McColl-Kennedy, J. R., Cheung, L., Gallan, A., Orsingher, C., Witell, L. and Zaki, M. (2018), 'Customer experience challenges: bringing together digital, physical and social realms', *Journal of Service Management* **29**(5), 776–808.

Box, G. E. P. (1954), 'Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification', *The Annals of Mathematical Statistics* **25**(2), 290–302.

Cho, J., Lee, J., Choi, S. W., Lee, D. and Lee, I. (2005), 'Fault identification for process monitoring using kernel principal component analysis', *Chemical Engineering Science* **60**(1), 279 – 288.

Cox, T. F. and Cox, M. A. (2000), *Multidimensional scaling*, Chapman and Hall/CRC.

Deming, W. E. (1968), 'Walter A. Shewhart, 1891-1967', *Revue de l'Institut International de Statistique / Review of the International Statistical Institute* **36**(3), 372–375.

Downs, J. and Vogel, E. (1993), 'A plant-wide industrial process control problem', *Computers & Chemical Engineering* **17**(3), 245 – 255.

Dunia, R. and Joe Qin, S. (1998), 'Subspace approach to multidimensional fault identification and reconstruction', *AIChE Journal* **44**(8), 1813–1831.

Duong, T. (2007), 'ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R', *Journal of Statistical Software, Articles* **21**(7), 1–16.

Gharahbagheri, H., Imtiaz, S. A. and Khan, F. (2017), 'Root cause diagnosis of process fault using KPCA and Bayesian network', *Industrial & Engineering Chemistry Research* **56**(8), 2054–2070.

Godoy, J. L., Zumoffen, D. A., Vega, J. R. and Marchetti, J. L. (2014), 'New contributions to non-linear process monitoring through kernel partial least squares', *Chemometrics and Intelligent Laboratory Systems* **135**, 76–89.

Hastie, T., Tibshirani, R. and Friedman, J. (2001), *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York Inc., New York, NY, USA.

He, B., Yang, X., Chen, T. and Zhang, J. (2012), 'Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach', *Journal of Process Control* **22**(7), 1228–1236.

He, Q. P. and Wang, J. (2018), 'Statistical process monitoring as a big data analytics tool for smart manufacturing', *Journal of Process Control* **67**, 35 – 43.

Hotelling, H. (1947), Multivariate quality control illustrated by air testing of sample bombsights, *in* C. Eisenhart, M. W. Hastay and W. A. Wallis, eds, 'Techniques of Statistical Analysis', McGraw Hil, New York, pp. 111–184.

Jackson, J. E. and Mudholkar, G. S. (1979), 'Control procedures for residuals associated with principal component analysis', *Technometrics* **21**(3), 341–349.

Ji, H., He, X., Shang, J. and Zhou, D. (2018), 'Exponential smoothing reconstruction approach for incipient fault isolation', *Industrial & Engineering Chemistry Research* **57**(18), 6353–6363.

Ji, H., He, X. and Zhou, D. (2016), 'On the use of reconstruction-based contribution for fault diagnosis', *Journal of Process Control* **40**, 24–34.

Jiang, B., Huang, D., Zhu, X., Yang, F. and Braatz, R. D. (2015), 'Canonical variate analysis-based contributions for fault identification', *Journal of Process Control* **26**, 17 – 25.

Johnson, R. and Wichern, D. (2002), *Applied multivariate statistical analysis*, 5th edn, Prentice Hall, Upper Saddle River, NJ.

Josse, J. and Husson, F. (2012), 'Selecting the number of components in principal component analysis using cross-validation approximations', *Computational Statistics & Data Analysis* **56**(6), 1869–1879.

Kano, M., Tanaka, S., Hasebe, S., Hashimoto, I. and Ohno, H. (2003), 'Monitoring independent components for fault detection', *AIChE Journal* **49**(4), 969–976.

Karatzoglou, A., Smola, A., Hornik, K. and Zeileis, A. (2004), 'kernlab – an S4 package for kernel methods in R', *Journal of Statistical Software* **11**(9), 1–20.

Kariwala, V., Odiowei, P., Cao, Y. and Chen, T. (2010), 'A branch and bound method for isolation of faulty variables through missing variable analysis', *Journal of Process Control* **20**(10), 1198–1206.

Kourti, T. and Macgregor, J. (1995), 'Process analysis, monitoring and diagnosis, using multivariate projection methods', *Chemometrics and Intelligent Laboratory Systems* **28**, 3–21.

Leavenworth, R. S. and Grant, E. L. (2000), *Statistical quality control*, International student edition, McGraw-Hill Education (India) Pvt Limited.

Lee, J.-M., Qin, S. J. and Lee, I.-B. (2006), 'Fault detection and diagnosis based on modified independent component analysis', *AIChE Journal* **52**(10), 3501–3514.

Lee, J., Yoo, C., Choi, S. W., Vanrolleghem, P. A. and Lee, I. (2004), 'Nonlinear process monitoring using kernel principal component analysis', *Chemical Engineering Science* **59**(1), 223 – 234.

Li, G., Qin, S., Ji, Y. and Zhou, D. (2009), 'Total PLS based contribution plots for fault diagnosis', *Acta Automatica Sinica* **35**(6), 759 – 765.

Liu, J. (2012), 'Fault diagnosis using contribution plots without smearing effect on non-faulty variables', *Journal of Process Control* **22**(9), 1609–1623.

Liu, J., Wong, D. S. H. and Chen, D. (2014), 'Bayesian filtering of the smearing effect: Fault isolation in chemical process monitoring', *Journal of Process Control* **24**(3), 1–21.

Liu, Q., Chai, T. and Qin, S. J. (2012), 'Fault diagnosis of continuous annealing processes using a reconstruction-based method', *Control Engineering Practice* **20**(5), 511–518.

Liu, Y., Zeng, J., Xie, L., Luo, S. and Su, H. (2019), 'Structured joint sparse principal component analysis for fault detection and isolation', *IEEE Transactions on Industrial Informatics* **15**(5), 2721–2731.

MacGregor, J. F., Jaeckle, C., Kiparissides, C. and Koutoudi, M. (1994), 'Process monitoring and diagnosis by multiblock PLS methods', *AIChE Journal* **40**(5), 826–838.

MacGregor, J. and Kourti, T. (1995), 'Statistical process control of multivariate processes', *Control Engineering Practice* **3**(3), 403 – 414.

Miller, P. and Swanson, R. (1998), 'Contribution plots: The missing link in multivariate quality control', *Applied Mathematics and Computer Science* **8**, 775–792.

Nelson, L. S. (1984), 'The Shewhart control chart—tests for special causes', *Journal of Quality Technology* **16**(4), 237–239.

Nomikos, P. (1997), Statistical monitoring of batch processes, *in* 'Preprints of Joint Statistical Meeting', Anaheim, CA.

Platt, J. C. (1999), Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, *in* 'Advances in Large Margin Classifiers', MIT Press, pp. 61–74.

Qin, S. J. (2012), 'Survey on data-driven industrial process monitoring and diagnosis', *Annual Reviews in Control* **36**(2), 220 – 234.

Qin, S. J., Yue, H. and Dunia, R. (1997), 'Self-validating inferential sensors with application to air emission monitoring', *Industrial & Engineering Chemistry Research* **36**(5), 1675–1685.

R Core Team (2018), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Rakotomamonjy, A. (2003), 'Variable selection using SVM based criteria', *Journal of Machine Learning Research* **3**, 1357–1370.

Rieth, C. A., Amsel, B. D., Tran, R. and Cook, M. B. (2017), 'Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation'.

Rosipal, R. and Trejo, L. J. (2002), 'Kernel partial least squares regression in reproducing kernel hilbert space', *J. Mach. Learn. Res.* **2**, 97–123.

Schölkopf, B., Platt, J. C., Shawe-Taylor, J. C., Smola, A. J. and Williamson, R. C. (2001), 'Estimating the support of a high-dimensional distribution', *Neural Computation* **13**(7), 1443–1471.

Schölkopf, B. and Smola, A. (2002), *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, USA.

Schölkopf, B., Smola, A. and Müller, K. (1998), 'Nonlinear component analysis as a kernel eigenvalue problem', *Neural Computation* **10**(5), 1299–1319.

Scrucca, L. (2004), 'qcc: an R package for quality control charting and statistical process control', *R News* **4/1**, 11–17.

Shang, C., Huang, B., Yang, F. and Huang, D. (2016), 'Slow feature analysis for monitoring and diagnosis of control performance', *Journal of Process Control* **39**, 21–34.

Shang, C., Ji, H., Huang, X., Yang, F. and Huang, D. (2019), 'Generalized grouped contributions for hierarchical fault diagnosis with group lasso', *Control Engineering Practice* **93**, 104193.

Van den Kerkhof, P., Vanlaer, J., Gins, G. and Impe, J. F. M. V. (2013), Contribution plots for statistical process control: Analysis of the smearing-out effect, *in* '2013 European Control Conference (ECC)', pp. 428–433.

Van den Kerkhof, P., Vanlaer, J., Gins, G. and Van Impe, J. F. (2013), 'Analysis of smearing-out in contribution plot based fault isolation for statistical process control', *Chemical Engineering Science* **104**, 285–293.

Wang, J., Ge, W., Zhou, J., Wu, H. and Jin, Q. (2017), 'Fault isolation based on residual evaluation and contribution analysis', *Journal of the Franklin Institute* **354**(6), 2591 – 2612. Special issue on recent advances on control and diagnosis via process measurements.

Wang, L. and Shi, H. (2014), 'Improved kernel PLS-based fault detection approach for nonlinear chemical processes', *Chinese Journal of Chemical Engineering* **22**(6), 657–663.

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**(1), 95 – 114.

Wise, B., Gallagher, N., Bro, R., Shaver, J., Windig, W. and Koch, R. (2006), 'PLS toolbox 4.0 for use with Matlab', *Eigenvector Research Inc., Manson, WA* .

Wold, S., Ruhe, A., Wold, H. and Dunn, III, W. J. (1984), 'The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses', *SIAM J. Sci. Stat. Comput.* **5**(3), 735–743.

Yan, Z. and Yao, Y. (2015), 'Variable selection method for fault isolation using least absolute shrinkage and selection operator (lasso)', *Chemometrics and Intelligent Laboratory Systems* **146**, 136–146.

Yuan, M. and Lin, Y. (2006), 'Model selection and estimation in regression with grouped variables', *Journal of the Royal Statistical Society, Series B* **68**, 49–67.

Yue, H. H. and Qin, S. J. (2001), 'Reconstruction-based fault identification using a combined index', *Industrial & Engineering Chemistry Research* **40**(20), 4403–4414.

Zhang, C., Gao, X., Xu, T., Li, Y. and Pang, Y. (2018), 'Fault detection and diagnosis strategy based on a weighted and combined index in the residual subspace associated with PCA', *Journal of Chemometrics* **32**(11), e2981.