USING EXTERNAL INFORMATION FOR STATISTICAL PROCESS CONTROL

By

SEONGKYU YOON, M.Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University

Doctor of Philosophy (2001) McMaster University
(Chemical Engineering) Hamilton, Ontario

TITLE: Using external information for statistical process control

AUTHOR: Seongkyu Yoon, M.Sc, Korea Advanced Institute of Science and Technology

SUPERVISOR: Professor John F. MacGregor

NUMBER OF PAGES: xvi, 183

Abstract

This research presents practical solutions for several problems in the case of multivariate statistical process control (MSPC) for process monitoring and fault diagnosis of industrial processes. Various types of external information are used with multivariate statistical models to deal with fault detection and isolation (FDI) problems. The external information considered in this work includes fault knowledge based on historical data or first principles, spectral characteristics of process variations and faults, and process knowledge on variables and observations.

The first part of the research deals with the fault isolation problem. It is based on the fact that MSPC approaches have been found to be very useful for fault detection, but less powerful for fault isolation because of the non-causal nature of the data. To improve fault isolation, an approach is proposed that uses additional data on past faults to supplement existing contribution plot methods. This approach extracts steady-state fault signatures of faults in both the correlation model space and the residual space. The vector direction of the fault signature is compared to the corresponding directions of known faults. The fault isolation is performed with a joint plot of angles between the direction of the current fault and those of the known faults. A simulated continuous stirred tank reactor (CSTR) with feedback control is used to demonstrate the proposed method. The method is shown to be effective in isolating both simple and complex faults.

The use of transient fault trajectory or initial fault signature would minimize detection delay and false isolation of a fault. An indirect usage of process dynamic information to deal with FDI problems is proposed in the second part of the work. Since process signals represent the cumulative effects of many underlying process phenomena, multiresolution analysis via wavelet transformations is used to decompose signals. The decomposed process measurements are rearranged according to their scales, and generalized principal component analysis (GPCA) is applied to these multiscale data to

capture process variable correlations occurring at different scales. The conventional PCA as well as multiscale PCA (MSPCA) models are shown to be limiting cases of the proposed GPCA model. Based on the FDI procedure proposed, a comparison study is done through Monte Carlo simulation. The new method can significantly enhance FDI performance when the frequency content of a fault effect is confined to a narrow frequency band.

The third part of this thesis examines both the fundamental and the practical differences between the causal and statistical model-based approaches to FDI. The causal-model-based approach is based on causal state variable or parity relation models developed from theory or identified from plant test data. Faults are then detected and isolated with structured or directional residuals from these models. MSPC approaches are based on non-causal models built with multivariate latent variable methods using historical process data. Faults are then detected by referencing future data against these covariance models, and isolation is attempted through examining contributions to the breakdown of the covariance structure. The differences in the nature of the models used (arising from the nature of the data used to build them) are shown to be responsible for the very different approaches to FDI. This model difference is also shown to be responsible for major strengths and weaknesses of the two approaches.

Most processes are subject to change of the operating conditions such as feed rate and composition, product grade, controller status, and so on. Sometimes these large *common-cause* variations disguise or distort the relevant information to faults. This difficulty can be minimized if one builds a correlation model that includes only the process variations relevant to FDI. By incorporating various types of prior knowledge into the empirical model building process, one can estimate a hybrid correlation model that includes both raw data and prior knowledge. Application of the hybrid correlation model for process monitoring and fault diagnosis is proposed and used for analyzing a real industrial dataset. The hybrid correlation model can increase the sensitivity for process monitoring, and outperform the regular correlation model in analyzing the process abnormality.

Acknowledgements

I would like to give my sincere thanks to Dr. John F. MacGregor, who has supervised this research and made it one of the most enjoyable and stimulating experiences for me. His guidance and encouragement will be remembered at every turn of my life.

I would like to thank Dr. Tom E. Marlin and Dr. Mo J. Elbestawi on supervisory committee, and Dr. Tim Davidson, for their inspiring discussions and advice. I am also grateful to Koji Muteki (Mitsubishi Chemical Co., Japan) for the industrial data in this work, and Dr. Theodora Kourti and Dr. Paul A. Taylor for their interest and advice.

My old friends Robert Kranz and Dr. Gerardo Mijares, and my mentor Dr. Sunwon Park from Korean Advanced Institution of Science and Technology deserve my special thanks. They guided and encouraged me to open a new path of my academic career.

To be a member of the process control group at the department of Chemical Engineering has been an asset through my years at McMaster. My fellow researchers and the department staff made me feel truly part of the group, with a lot of unforgettable memories. Joonhong, Ron, Youngmin, Jeongae, Shinjoon, and other members of the Korean mafia in Hamilton also have my heartful thanks.

For all the love and care from my brother- and sister-in-law Changhyun and Hyunsook, and my sister Mison in Korea, I will never be able to put in words how much they mean to me.

Finally, I would like to give thanks and love to my son Jungwoo and wife Hyosuk for their untiring love, patience, and understanding during the study period.

Table of Contents

List of Tables

List of Figures

1. Introduction

Process supervision and monitoring for consistent operation of processes is an important concern for all types of industries. The objective of process supervision and monitoring is prompt detection and isolation of faults at an early stage. The most common approach to fault detection and isolation (FDI) is to use mathematical process models. The estimated signals called analytical redundancy are calculated from the causal model and compared with measured quantities. This method can be used to obviate the approach based on hardware redundancy. However, the causal-model-based approach is built upon a number of rather idealized assumptions, one of which is that the mathematical model used is a reliable replica of a process. This is, of course, difficult to achieve in practice since an accurate and complete mathematical description of the process is rarely available. For other applications, the process parameters may not be fully known or may be partially known over a limited range of the process operation. Another idealized assumption is that the true characteristics (e.g., spectral shape, variance, mean, stationarity, etc) of disturbances and noise signals acting upon the process are already known. It is also very difficult to know this information completely. All these assumptions make it hard to implement the causal-model-based approach. On the other hand, this approach works efficiently if those assumptions are satisfied.

The blind spots of the causal-model-based approach particularly hold for large and complex processes where reliable measurement of process variables is usually complicated, expensive, or in many cases simply impossible. A knowledge based approach using qualitative models associated with heuristic reasoning can be used as a

complementary scheme to the causal-model-based approach. In many practical situations, a combination of both the causal-model-based and the knowledge-based approaches may be the most appropriate solution to FDI (Frank, 1990). However, formulating behavioral and causal descriptions, and diagnostic hierarchy with probabilistic or fuzzy rules of knowledge based method, may be just as difficult and time-consuming as the causal–model-based approach. Further, in case of an unsteady-state application, there are many drawbacks such as the lack of a statistical basis for interpreting data and classifying results, and the complexity of the approaches when dealing with more than a few variables (Nomikos and MacGregor, 1994).

The multivariate statistical methods have been used as a reasonable approach to industrial FDI problems. They require fewer resources for successful implementation. It is generally recognized that this approach has practical benefits in large complex processes and can handle such problems as collinearity between process variables and missing process data. The resource for the multivariate statistical methods is normal operation data for building the correlation model. The correlation model is based on *common-cause* variations present in a process. Since the correlation models can be more easily estimated than the causal relationships, the multivariate statistical models have been popularly used for FDI for large or complex processes. FDI approaches using these correlation models fall under the category of multivariate statistical process control (MSPC).

However, this approach has mostly been applied to static systems. Some modifications are necessary when one uses it in dynamic systems. Furthermore, this approach does not provide a diagnostic feature for all faults because the correlation models do not generally provide a causal structure compatible with the underlying physical reality. The process engineer must follow up by doing a deeper analysis. This limitation is caused by the nature of the data for modeling rather than the methods themselves. The correlation models are also limited in that they should be used only within the range of the operating conditions such as feed rates, feed compositions, product specifications, and control strategies. For example, when there is a change in

operating conditions, or when the empirical correlation model is used outside the base-case conditions, the correlation models could give poor model predictions and hence they should be estimated again.

Since each of these approaches has both limitations and advantages, reliable and efficient FDI methods can be obtained by combining the best features of each while using all available FDI resources appropriately. This research is aimed at extending MSPC by using external information, and thus providing practical solutions for process monitoring and fault diagnosis of industrial processes. This research proposes how to enlarge the practical applicability of the conventional MSPC methods by using various types of process knowledge not used for building correlation models. The process knowledge includes fault information on faults based on historical data or first principles (chapter 2), spectral characteristics on process variations and faults (chapters 3 and 4); process information on process variables and observations (chapter 6). The issues to be studied in each chapter are: enhancement of fault isolation under MSPC framework (chapter 2); FDI of dynamic system (chapters 3 and 4); comparison of causal and statistical model based FDI approaches (chapter 5), incorporation of external information into principal component analysis (PCA)/projection to latent structures (PLS) models (chapter 6). Brief descriptions of each problem as well as a discussion of the objectives and contributions of this work are presented below.

1.1 Fault isolation using SS fault signature under MSPC

The multivariate statistical methods have been recently proposed and successfully implemented in many industries. The major benefit is that there is no need for a fundamental or causal model of the system. These approaches only require a good database of normal historical data. They can also easily handle very large problems involving hundreds of variables. They are insensitive to noise in the measured variables, and, very importantly, handle missing data in a very efficient and easy manner. On the other hand, the major deficiency is that it does not provide a causal description of the

process that can be used as the basis for fault isolation (diagnosis). The basic tools used for fault isolation are contribution plots. These plots are usually very powerful in isolating simple sensor or actuator faults, but much less successful at providing a clear isolation of complex faults.

The primary objective of this work is to develop a method that can supplement these contribution plots, and provide much better fault isolation when some additional information of the faults is available. It is assumed that additional information on various faults is available from mechanistic knowledge, or from historical data on past occurrences of these faults, or from data obtained from plant tests. Fault signatures are then developed for these faults based on the PCA or PLS model built from normal *common-cause* data. The fault signature bank, or catalog would simply be the collection of these fault signatures that are available to date. Once a new fault occurs its signature is compared with those in the fault bank in order to identify the most likely cause. The new methods are illustrated using a non-isothermal continuous stirred tank reactor (CSTR) system with feedback control

The contribution of this work is to provide an enhanced methodology to diagnose faults under the MSPC framework by using additional fault information available. It has been shown that the proposed method is very effective in isolating both simple and complex faults. A version of this chapter has been presented at the AIChE annual meeting, Miami, Florida, 1998 Nov., and published in the *Journal of Process Control* (Yoon and MacGregor, 2001a).

1.2 Unifying PCA and multiscale approaches to FDI

The usage of steady-state fault signatures causes an isolation delay. False isolation may arise if the transient fault directionality is very different from the steady-state one. Thus, the use of transient fault trajectory or initial fault signature may help fault detection and isolation.

In order to handle process dynamics under the multivariate statistical framework,

it has been proposed that the dynamic model be obtained by augmenting lagged variables in the measurement matrix and applying the classical method for it (Kourti and MacGregor, 1995). This is equivalent to a time series modeling approach (Box and Jenkins, 1976). As an explicit way of dynamic modeling, one can use latent variables and subspace methods (Shi and MacGregor, 2000). With these dynamic models, various FDI methods can be used. However, the explicit dynamic models and their FDI applications are usually limited due to process dimension and data characteristics.

As an alternative, indirect usage of process dynamic information for fault detection and diagnosis is proposed in this work. It is based on the multiresolution analysis via wavelet transformation. One can decompose process signals such that all the contributing events are approximately discriminated according to their scale contents and a fault is distinguished from other events. A unifying framework that generalizes multiscale approaches (Multiscale principal component analysis: MSPCA) and PCA is presented in chapter 3. The potential aspects of the proposed method (Generalized principal component analysis: GPCA) are scrutinized by comparing it with existing approaches such as regular PCA and MSPCA. In chapter 4, the application of the proposed method to fault detection and isolation problem is addressed. The modeling method of chapter 3 is further extended to develop on-line and off-line implementation of FDI methods. Based on the analogy between the multiblock and the multiscale methods, a new algorithm to calculate the GPCA model is also presented. The existing FDI method based on PCA is extended to the multiscale case. Performance of FDI using GPCA is assessed using a simulated CSTR system. The FDI performance of the GPCA based method is compared with that of the regular PCA based method via Monte Carlo simulation. Also discussed is the detection and isolation of a complex fault whose signature becomes confused with the multiscale nature of the process signals.

The contribution of this work is to provide a unifying framework of PCA and multiscale approaches to FDI problem, and introduce contributions plots into multiscale analysis to enhance the isolation of faults. Through the assessment based on the Monte Carlo simulations, the types of faults that can be most effectively treated with the GPCA

model are clarified. This work has been presented at three conferences: CSChE conference, Montreal, Quebec, 2000 Oct.; AIChE annual meeting, Los Angeles, CA, 2000 Nov.; 6th IFAC symposium on dynamics and control of process systems (DYCOPS-6), Chejudo, Korea, June 2001 (Yoon and MacGregor, 2001b).

1.3 Statistical and causal model based approaches to FDI

The multivariate statistical approaches are much easier to develop because of the ready availability of routine operating data and their ability to handle a very large number of measured variables. They are capable of detecting almost any type of fault. However, their major weakness lies in the ambiguous isolation of faults arising from the non-causal nature of the models. On the other hand, the causal-model-based FDI methods can both detect and isolate faults as long as causal models are developed for all the variables and the fault structures identified *a priori*. However, the need for mechanistic or identified causal models makes the approach more difficult to develop and hence limited to smaller systems.

This research performs a comprehensive comparative study in which the differences and the complementary strengths/weaknesses of the statistical and the causal-model-based approaches to FDI problems are assessed via several simulation studies. The parity relation approach is selected to represent the causal-model-based methods and a PCA based method to represent the MSPC approaches. Practical differences of these approaches are illustrated using simulated data from a continuous stirred tank reactor process.

The contribution of this work is an in-depth examination of the differences in the models, the data required to build models, the processes to which they are applicable, and the assumptions behind the methods. Through this work, various fundamental aspects are compared, critical differences are assessed, and their strengths and weaknesses are highlighted. The contribution to academia is to provide a benchmarking study to the research that involves the ways of combining two approaches in a manner that utilizes the

strengths of both. This work has been presented at IFAC international symposium on advanced control of chemical processes (ADCHEM 2000), Pisa, Italy, June 2000, and published in the *AIChE Journal* (Yoon and MacGregor, 200a, b, c).

1.4 Incorporation of external information into multivariate statistical models

Since the causal- and statistical-model-based methods have complementary strengths as well as significant differences, a potential FDI approach would be utilizing the strengths of both. In many cases, complete causal process models are rarely available due to their dimensions and complexities. The partly available process knowledge combined with the correlation models would be useful in interpreting questionable events occurring in processes.

This research assesses the issues of using various types of prior process knowledge with the multivariate statistical methods to efficiently use all available resources. These methods of incorporating prior knowledge into the multivariate statistical correlation models define a hybrid correlation modeling. This work examines various types of prior knowledge used for the hybrid correlation modeling, various methods of incorporating external information into the correlation models, and process monitoring and fault diagnosis using the hybrid correlation model. The focus is on integrating the available resources under the MSPC framework, and using the resulting hybrid correlation model for FDI. The existing correlation methods using prior knowledge are assessed from the viewpoint of the proposed methods. The key features of the hybrid correlation modeling are illustrated. A real industrial dataset is analyzed using the proposed scheme.

The contribution of this work is to provide a way of incorporating external information into multivariate statistical model, and systemize various correlation modeling approaches using prior knowledge. This work has been presented at 4th IFAC workshop on on-line fault detection and supervision in the chemical process industries,

Chejudo, Korea, 2001 June (Yoon and MacGregor, 2001c).

1.5 Thesis outline

This thesis consists of 7 chapters, the first one being the current introduction. Chapters 2-6 form the core of this thesis. Fault isolation problem of MSPC with additional information on various faults is discussed in chapter 2. PCA model of multiscale data and its application to FDI problem with prior knowledge on spectral characteristics of process variations and faults are presented in chapters 3 and 4. Chapter 5 compares the causal- and statistical-model-based FDI methods. Chapter 6 focuses on the incorporation of external information into multivariate statistical model. Finally, the results obtained in the thesis are summarized in Chapter 7, where some conclusions are drawn and future work is discussed.

2. Fault isolation using steady-state fault signature under MSPC

2.1 Introduction

The purpose of fault detection is to determine the occurrence of an abnormal event in a process, and that of fault isolation is to identify its reason or source. Traditional fault detection and isolation (FDI) methods have been based on a mechanistic model of the process, or an identified causal model of the process. In addition to the process model, relationships between model states, or parameters and faults are required. The estimated states and model parameters are the basis of these approaches (Willsky, 1976; Isermann, 1984; Frank, 1990; Patton, 1995; Gertler, 1995a). Most approaches use state observers or Kalman filters, or consistency (parity) relations that make direct/indirect uses of a state-space or causal input/output model of the monitored process. However, it may often be both difficult and time-consuming to develop a good mechanistic model. Knowledge-based methods such as the expert system have been developed as complementary approaches to the causal model based methods in the last several years. This gives some potential for using diagnostic rules and qualitative knowledge of a process instead of a mechanistic model. However, considerable effort is also required for building these knowledge-based diagnostic systems for complex and large processes. Neural networks and clustering based methods for FDI have been reported for providing a solution to these difficulties of the knowledge-based methods. However, their applicability is limited by the requirements of having comprehensive and excessive plant data on all the faults, or very detailed and good simulation models (Venkatasubramanian and Chan, 1989;

Hoskins *et al.*, 1991; Fan *et al.*, 1993; Lennox *et al.*, 1998).

Alternative approaches based on multivariate statistical methods have been proposed (Kresta *et al.*, 1991; Wise and Ricker, 1991). In many ways, these approaches, based on the philosophy of multivariate statistical process control (MSPC), take a quite different and complementary approach as compared to the conventional FDI methods. The most fundamental difference is that they are based on non-causal empirical correlation models built from normal plant operating data only when *common-cause* variation is present. By using multivariate latent variable methods such as principal component analysis (PCA) and projections to latent structures (PLS), they are able to project most of the important information in these data down into low dimensional latent variable spaces where regions of normal and abnormal operation are easily defined. As new data comes available they are compared with the normal or *common-cause* variation captured by these low dimensional models and statistical tests used to detect any abnormal behavior. The major benefits of this approach are that there is no need for a fundamental or causal model of the process. The approach only requires a good database of normal historical data, and the models are easily and quickly built from this. They can also easily handle very large problems involving hundreds of variables, they are insensitive to noise in the measured variables, and, very importantly, they handle missing data in a very efficient and easy manner. On the other hand, the major deficiency of these multivariate statistical approaches is that, unlike the conventional state variable model or parity equation approaches, they do not provide a causal description of the process that can be used as the basis for fault isolation (diagnosis). The basic tools used for fault isolation in the multivariate statistical approaches are contribution plots - *i.e.* plots of the contributions of each variable to the deviation of the current operating point from that defined by the PCA or PLS model for normal *in-control* operation. These plots are usually very powerful at isolating simple sensor or actuator faults, but much less successful at providing a clear isolation of complex faults. However, without further information about causal effects or about the nature of various faults, these contribution plots appear to be the best that one can do with methods based only on historical *in-*

control data.

The objective of this chapter is to develop a method that can supplement these contribution plots, and provide much better fault isolation when some additional information about these complex faults is available. It is assumed that additional information on various faults is available from mechanistic knowledge, or from historical data on past occurrences of these faults, or from data obtained from plant tests. Fault signatures are then developed for these faults based on the PCA or PLS model built from normal *common-cause* data. A fault signature consists of the directions of movement of the process in both the PCA/PLS model space and in the orthogonal residual space during the period immediately following its detection of the fault. The fault signature bank, or catalog would simply be the collection of these fault signatures that are available to date. There will be many unknown fault signatures, which can be added after detecting a fault. Once a new fault occurs its signature is compared with those in the fault bank in order to identify the most likely cause.

An outline of the chapter is as follows. The next section introduces the multivariate statistical process control approaches based on PCA and PLS, and their basis for fault detection and isolation. Recent literature involving attempts to enhance the fault isolation aspects of these methods using additional information is then discussed. The main topic of the chapter involving an improvement on these approaches using fault signatures lying both in and orthogonal to PCA/PLS model space is then presented. Finally the new methods are illustrated using a non-isothermal continuous stirred tank reactor (CSTR) simulation with feedback control.

2.2 Multivariate statistical approaches to FDI

2.2.1 Fault detection using PCA/PLS model

The use of PCA/PLS to build low dimensional models for the analysis and

monitoring of process operations is now well established (Kresta *et al.*, 1991; Wise and Ricker, 1991; Jackson, 1991; MacGregor *et al.*, 1994; Nomikos and MacGregor, 1994; Kourti and MacGregor, 1995), and many industrial applications exist (Kourti and MacGregor, 1996; Kosanovich *et al.*, 1996; Neogi and Schlags, 1998). Provided with historical data ($\mathbf{X}$) collected during normal process operation, most of the *common-cause* variations in the process can be expressed in terms of a small number of the principal components (PCs). A cross-validation stopping criterion is usually used to determine the number of principal components (A) that capture most of the relevant variations:

$$\mathbf{X} = \mathbf{T}_A\mathbf{P}_A^T + \mathbf{E} = \sum_{i=1}^{A} \mathbf{t}_i\mathbf{p}_i^T + \mathbf{E}, \tag{2.1}$$

where $\hat{\mathbf{X}} = \mathbf{T}_A\mathbf{P}_A^T$ represents the *common-cause* variations lying in the A-dimensional subspace spanned by the first A principal components, and $\mathbf{E}$ is the orthogonal space of the residuals. The residual space can also be expressed as $\mathbf{E} = \tilde{\mathbf{T}}\tilde{\mathbf{P}}^T = \sum_{i=A+1}^{m} \mathbf{t}_i\mathbf{p}_i^T$ where $\tilde{\mathbf{P}}$ is the matrix of loadings associated with the smallest singular values of $\mathbf{X}$.

The principal component loadings ($\mathbf{p}_i$) are an orthogonal set of basis vectors for the extracted features called scores ($\mathbf{t}_i$). They span the subspace (representational space, $\hat{\mathbf{X}} = \mathbf{T}_A\mathbf{P}_A^T$) of the predictable variations in the training data. The first A principal components define the maximum variance directions in the score space. They give a set of reduced dimension features which are uncorrelated and account for as much of the total data variance as possible. Since the training set is assumed fault-free, any future score vectors should fall in this reduced space if the process is still fault-free. The remaining principal components, which correspond to the small eigenvalues, span the orthogonal complement ($\mathbf{E}$) of the representation subspace. This is referred to as the residual subspace. Note that any true variable is orthogonal to the residual subspace under *common-cause* variations if there are no errors in model, and thus $\tilde{\mathbf{P}}^T\mathbf{x}_k^o = \mathbf{0}$.

If one has both process data $\mathbf{X}(n\times m)$, and key quality and productivity data $\mathbf{Y}(n\times l)$

then, rather than PCA, one can use PLS to obtain alternative estimates of the latent variable space (t_1, t_2, ...). PLS will focus more on capturing the high variance directions in the **X** space that are most correlated with the **Y** space. The process monitoring approaches would remain unchanged. In the case of dynamic processes each row of **X** includes lagged values of all measured variables instead of steady-state values. In fact, the monitoring of dynamic batch and semi-batch processes using PCA or PLS models developed from the time histories of the variable deviations from their average trajectories has become a well-accepted industrial approach to FDI (Nomikos and MacGregor, 1994, 1995; Kourti *et al.*, 1996; Kosanovich *et al.*, 1996; Neogi and Schlags, 1998).

Having established a PCA model, future behavior can be referenced against this *in-control* model (MacGregor and Kourti, 1995). That is, new observations ($\mathbf{x}_{new}$) are projected onto the plane defined by the loading vectors to obtain their scores ($\mathbf{t}_{A,new} = \mathbf{P}_A^T\mathbf{x}_{new}$) and the residuals ($\mathbf{e}_{new}=\mathbf{x}_{new}- \hat{\mathbf{x}}_{new}$) where $\hat{\mathbf{x}}_{new} = \mathbf{P}_A\mathbf{t}_{A,new}=\mathbf{P}_A\mathbf{P}_A^T\mathbf{x}_{new}$). Multivariate control charts based on Hotelling's T^2 can be plotted based on the first A principal components as follows:

$$T^2 = \sum_{i=1}^{A} \frac{t_i^2}{s_{t_i}^2}, \tag{2.2}$$

where $s_{t_i}^2$ is the estimated variance of t_i. An upper control limit on this chart is obtained using the F-distribution (Kourti and MacGregor, 1996). This control chart will only detect variation in the plane of the first A principal components that is greater than what can be explained by the *common-cause* variations. When a new type of special event occurs which was not present in the *in-control* PCA model, the new observations will move off the plane. This type of event can be detected by computing the squared prediction error (SPE) of the residual for new observations (Kresta *et al.*, 1991). This statistic is also called the Q-Statistic (Jackson and Mudholkar, 1979), or distance to model (DModX). It is defined as:

$$\mathrm{SPE} = \sum_{i=1}^{m}\left(x_{i,new} - \hat{x}_{i,new}\right)^2 = (\mathbf{x}_{new} - \hat{\mathbf{x}}_{new})^{\mathrm{T}}(\mathbf{x}_{new} - \hat{\mathbf{x}}_{new}). \tag{2.3}$$

When the process is *in-control*, this SPE statistic represents unstructured residuals that cannot be accounted for by the PCA model. When an unusual event occurs that results in a change in the process mean or covariance structure, it will be detected by a high value of this statistic. The upper control limits for this statistic can be computed based on the reference data (Jackson and Mudholkar, 1979, Nomikos and MacGregor, 1995).

2.2.2 Fault isolation using contribution plots

Once a fault is detected by the above multivariate tests, the main approach to fault isolation using PCA/PLS models is the use of contribution plots (MacGregor *et al.*, 1994; Miller *et al.*, 1993; Kourti and MacGregor, 1996). Although the MSPC monitoring charts are very effective for fault detection, fault isolation is more difficult with this approach. When the SPE statistic violates its upper control limit, the contributions of the individual variables ($x_{i,new} - \hat{x}_{i,new}$) can be plotted and those variables having large contributions examined to indicate possible causes. Similarly, if the variation in the model space (T^2, or t_1, t_2, ...) becomes large, then contribution of each variable (x_i) to a large value of the j-th score (t_j) is given by:

$$Contribution(x_i) = p_{ij}\Delta x_i\,, \tag{2.4}$$

where p_{ij} is the weight of the i-th variable (x_i) in the j-th latent variable ($\mathbf{p}_j$) and Δx_i is the change in x_i over the time period in question. More details can be found in literature (Kourti and MacGregor, 1996). Since these contribution plots come from an underlying correlation model, which does not provide a causal relationship among the variables, they do not provide direct fault isolation. They only show which group of variables are highly

correlated with the fault, and it is up to the engineer to use his process insight to provide feasible interpretations. In the case of simple actuator or sensor faults these contribution plots usually can clearly isolate the fault since it is mainly just that one variable whose correlation structure with the other variables has changed (Kourti *et al.*, 1996). In the case of complex or process faults, the isolation is often not clear. However, these plots have been found to be very useful in many applications, because they provide a fault signature that focuses one's attention on a small subset of the large number of process variables, and so greatly restricts the number of fault possibilities that the engineer may have to consider. Given that the only information used in this approach is normal process operating data, this is the best one can expect to do with fault isolation. To do more, one needs to have additional information. This is discussed in the following sections of the chapter. How well fault isolation based on simple contribution plots will perform depends on the fault characteristics largely.

2.2.3 Fault characteristics

All components of processes which consist of actuators, process dynamics and sensors, can lead to an alarm or a false alarm as one form of additive or multiplicative faults. *Additive measurement faults* are discrepancies between the measured and true values of process outputs or inputs. Such faults best describe sensor biases. They can also describe actuator malfunction which is the discrepancy between the intended controller output and its realization by the actuator. *Additive process faults* are unmeasured inputs acting on the plant, which are normally zero and cause a shift in process outputs, independent of the measured inputs. They are considered deterministic such as a constant bias or drift, or semideterministic which is jumps occurring at random intervals with random amplitudes. Such faults best describe process leaks, loads, etc. *Multiplicative process faults* are abrupt or gradual changes of the process parameters. Such faults best describe the deterioration of process equipment, surface contamination, etc.

However, these fault categories may not describe some practical failure situations

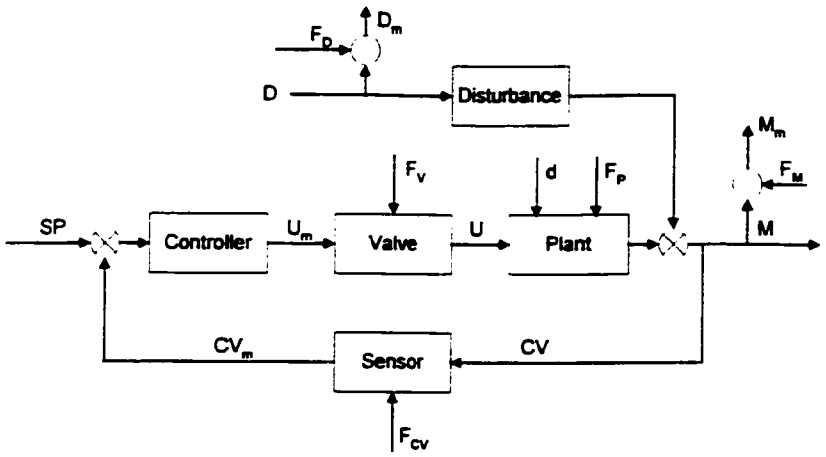


Figure 2.1 Faults in process (Measurements : CV_m, U_m, D_m, M_m; Instrument Faults : F_{CV}, F_V, F_D, F_M; Process Fault : F_P; External Changes : SP(Setpoint), D(measured disturbance), d(unmeasured disturbance))

in a natural way. A complete sensor failure would have to be described either as a variable bias or as a multiplicative fault where some parameter becomes zero. Many disturbances are not completely additive: the intensity of a leak, for example, may depend on pressure that is a process variable. An accurate description of this situation would require nonlinear modeling. Furthermore, actuator failure might be more naturally described as an additive or multiplicative process fault, depending on the nature of the failure. Typically, abrupt faults play a role in safety-relevant systems where hard-failures have to be detected early enough so that catastrophic consequences can be avoided by early system recognition. In case of maintenance problems, where early detection of worn equipment is required, incipient faults are of major concern. In this case, the faults are small and not easy to detect, but the detection time is less important.

Figure 2.1 illustrates the basic configuration of a process and possible faults. In general, faults can occur either in the actuators or sensors, or in the functional units inside of a process. They are classified into two groups depending upon their characteristics.

One is a *simple fault* that occurs in a specific fault source and its effect is not propagated into other variables. The fault signature coming from the contribution plots would therefore be dominated by one large contribution in that sensor or actuator

Table 2.1 Steady state values for bias faults under PI and without control (No dynamics on sensor and valve; CV_m: controlled variable; M_m: its correlated measurement; U_m: controller output; D_m: disturbance measurement; K_P: process gain of controlled variable; K_M: process gain of correlated variable; $f=K_M/K_P$)

Step changes	Fault Signature, $[CV_m\ U_m\ D_m\ M_m]$	
	Under PI Control	W/O Control
Controlled Variable (F_{CV})	$[0\ -F_{CV}/K_P\ 0\ -fF_{CV}]$	$[F_{CV}\ 0\ 0\ 0]$
Valve (F_V)	$[0\ -F_V\ 0\ 0]$	$[0\ F_V\ 0\ 0]$
Disturbance (F_D)	$[0\ 0\ F_D\ 0]$	$[0\ 0\ F_D\ 0]$
Correlated Measurement(F_M)	$[0\ 0\ 0\ F_M]$	$[0\ 0\ 0\ F_M]$
Setpoint (ΔSP)	$[\Delta SP\ \Delta SP/K_P\ 0\ f\Delta SP]$	$[0\ 0\ 0\ 0]$
Real disturbance (ΔD)	$[0\ -\Delta D/K_P\ \Delta D\ 0]$	$[0\ 0\ \Delta D\ f\Delta D]$
Process Fault (F_P)	$[0\ -F_P/K_P\ 0\ -fF_P]$	$[F_PK_P\ 0\ 0\ fF_P]$
Unmeasured disturbance(Δd)	$[0\ -\Delta d/K_P\ 0\ -f\Delta d]$	$[\Delta dK_P\ 0\ 0\ f\Delta d]$

measurement in which the fault occurred. This results because only this variable has broken its correlation pattern with all the remaining variables in the model. Therefore, in the case of a simple sensor failure, fault isolation by simple contribution plots is relatively easy (Kourti *et al.*, 1996). In Figure 2.1, this type of fault would arise due to individual sensor or actuator faults arising in the measured disturbance (F_D), or the monitoring variable (F_M).

The other type is a *complex fault* that arises from some fundamental change in the process such as fouling of heat exchanger surfaces or contaminating impurities entering a reactor system (e.g., d or F_P in Figure 2.1). These complex faults affect many measured variables. In the case of processes with feedback and feedforward control systems, even sensor and actuator faults become complex faults since their effects are propagated into many other variables by the actions of the control system (e.g., F_{CV} and F_V in Figure 2.1).

To illustrate the differences between fault isolation in the open and closed-loop situations, Table 2.1 shows the steady-state value changes caused by bias changes occurring in all possible sensors/actuators in the system shown in Figure 2.1 for both

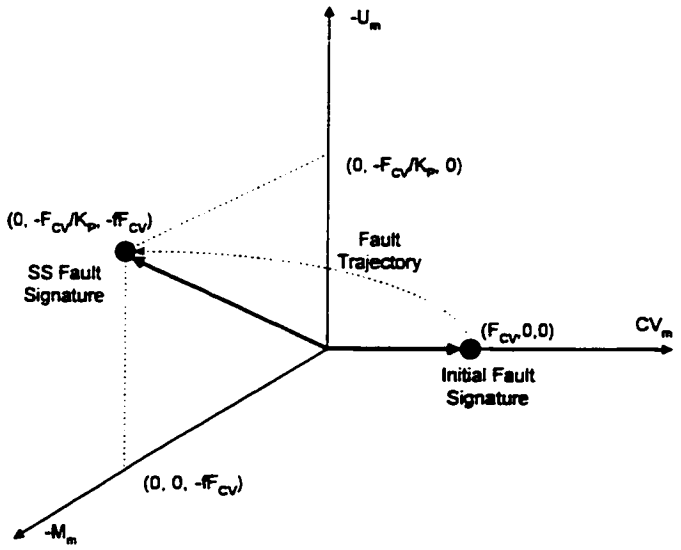


Figure 2.2 Fault trajectory and fault signatures (CV_m: controlled variable; M_m: its correlated measurement; U_m: controller output; K_P: process gain of controlled variable; K_M: process gain of correlated variable; $f=K_M/K_P$)

open and closed-loop operations. Except for the bias changes on the disturbance (F_D) and the correlated measurement (F_M), all the fault signatures in the closed loop are different from those in the open loop operation. The differences are caused by the controller actions to remove the deviation on the controlled variable. As a further illustration, consider a bias fault on the controlled variable in Figure 2.1. This would show up in the process in the manner illustrated in Figure 2.2. An initial deviation (the bias) would appear in the controlled variable, but this would be followed by a transient period in which the fault effect would follow a trajectory determined by the dynamics of the process and controller, and finally reach a new steady-state having biases in the controller output (U_M) and the correlated measurement (M_m) but none in the controlled variable (CV_M)

Hence in complex faults arising either from fundamental process changes or from the propagation of simple faults by control systems, the fault signature or contribution plots will contain large deviations in many variables and fault isolation becomes much more difficult. To overcome this problem more information about faults or causal

relationships are needed. Methods that combine this fault information with the multivariate PCA/PLS models for normal or *common-cause* information can then be developed to provide improved fault isolation. Recent publications in this area are discussed in the next section, and a new approach is outlined and illustrated in the remainder of the chapter.

2.3 Fault decomposition and angle measures for isolation

2.3.1 Fault isolation approaches

Before the proposed scheme is presented, previous research on fault isolation is reviewed. It is assumed that a fault detection scheme is in place, which detects the onset of abnormal plant operation, and there exists a database of reference patterns, each corresponding to a known fault. In this situation, the fault detection is performed by testing whether the behavior of current measurement data is consistent with past *in-control* behavior captured by a PCA or PLS model, and the fault isolation is implemented by referencing signatures of the current fault against a database of the reference fault signatures. Existing fault diagnosis methods differ in the type of signatures used to characterize the faults, and in the manner of comparing them against the reference signature bank.

Wold (1976) used the principal component models in pattern recognition with the emphasis on applications in chemistry and biology. By representing the objects in separate classes by completely separate models using the principal components, it was possible to get a simple and powerful method of pattern recognition. The total model for a number of classes consequently consists of a collection of disjoint PC models; one model for each class. By means of data observed on objects with known classifications, the parameters in the separate models are estimated. Unclassified objects are then fitted to all the parameterized class models and classified according to which model they fit best.

A fault, or disturbance can be treated as one class, and one can apply the method for the fault isolation.

Ku *et al.* (1995) developed a disturbance isolation approach, which was based on PCA models, T^2 and Q charts, and a model bank of possible disturbances. Q and T^2 plots of data characterizing different disturbances projected on the PCA model of normal operating data were used for the disturbance isolation. They incorporated dynamic behavior of the process into the PCA model by using the standard method of including time-lagged data

Zhang *et al.* (1995) proposed an isolation approach which used only the first principal component of fault data sets as the extracted features. They assumed that data covering a fault incidence is polarized such that variances in the data are mainly represented by the first principal component. When a new abnormal data set is obtained, the new principal component is calculated. The cosines of angles between the first principal component of a new measurement data set and the loading vectors of the known faults are then used for fault diagnosis. A fault that gives maximum value of the cosine is identified as an assignable one for the new data set. A cosine value of one implies the colinearity between the new fault direction and one of the known fault directions. Although it is questionable that the characteristic feature of a fault can be captured with only the first principal component, the methodology using the fault directionality is a very simple and convenient way of diagnosis. This concept will be used in the method proposed in this chapter.

The common feature of the above approaches is the use of separate PCA models for each fault or disturbance. In this case, there exists an increased complexity as the number of faults or disturbances increase. On the other hand, one can simplify the problem by using only the PCA model already built for fault detection based on *common-cause* variation. Fault isolation can then be based on the projection of the fault history onto this model (scores), and its movement in the orthogonal residual (SPE or Q) space. As long as the projected scores of a fault form a data cluster in the principal component space, one can apply a various pattern recognition methods. For the diagnosis of

abnormalities leading to an out-of-control process behavior, Raich and Cinar (1997) presented a method based on the distance matrices and the angles between data clusters. They used the discriminants in both the model and residual spaces. The angle in their approach was calculated to look at similarities between objects with respect to an original or neutral state. However, since the vertex for the angle calculation was positioned at the mean (center) of the normal operation region, there is a possibility of wrong diagnosis when the same abnormal event starts from different points that are still within the normal region.

After a fault is detected with the correlation model under the framework of the multivariate statistical methods, the analysis of the scores and residuals leads to classification of the abnormal variation from the PCA model into two parts. One is a variation component lying within the plane of the *common-cause* PCA model and is defined by the variation in the projected score values (t_1, t_2, ...). The second is a variation component orthogonal to the PCA model plane leading to large SPE or Q statistics. Most faults will give rise to variations in both of these spaces. Therefore, once a fault is detected, any method for isolation should exercise the movement in both of these spaces. On the other hand, one needs to investigate only the residual space for the fault diagnosis when a process mechanistic model is used. This is one difference between fault diagnosis methods based on causal vs. non-causal methods of the process.

As seen in Figure 2.2, a complex fault has an initial fault signature, a time varying trajectory and a steady-state fault signature. The initial fault signature may provide a good and prompt source for the fault diagnosis since it is not affected by the fault propagation. However, the initial fault signature can be easily missed by any fault detection scheme, and thus the transient behavior and the final steady-state vector of the measurements after a fault occurrence are generally used as an alternative to the initial fault signature to characterize the complex fault signature.

Wise and Ricker proposed a fault isolation method based on reconstructing each variable using PCA/PLS models that used the remaining variables (Wise and Ricker, 1991). Any variable whose reconstruction error was large (above a threshold value) was

considered a faulty sensor. However, a fundamental assumption behind this approach is that the fault only affects the one variable being reconstructed, and does not affect any of the other variables being used to reconstruct it. It effectively limits this approach to the detection of simple faults. Complex faults cannot in general be isolated by this reconstruction approach based on projection models built from *in-control* or *common-cause* data. Such models are non-causal and have no ability to account for the propagated causal effect of the fault into the other variables, and the change that these would have on the predictions made by the *in-control* model.

Again, based on the reconstruction concept, Dunia *et al.* proposed a sensor validity index (SVI) to isolate a faulty sensor (Dunia *et al.*, 1996). Later they presented a unified approach to process and sensor fault detection, identification and reconstruction via principal component analysis (Dunia and Qin, 1998a; 1998b). Due to the assumption that the fault effect is not propagated into the other variables, the use of sensor validity index is again limited to the simple sensor fault situation. The approach also examines only the behavior of the fault in the residual space, and does not consider the movement that is also included in the PCA model space.

Although these PCA models built from normal operating data cannot be used for reconstruction of complex faults, they can be used to develop signatures of past faults from the final steady-state or the transient trajectories arising from the faults, and then isolate new faults by comparing their signatures against those stored in bank of known faults. In this chapter, we focus on the use of the steady-state fault signatures. By using steady-state fault signatures an isolation delay is inevitable, and a false isolation may arise if the transient fault directionality is very different from the steady-state one. However, the resulting methods appear to work well in many applications even during the transient periods, and provide a basis for future enhancements of the approach, which can accommodate the transient fault trajectory in an explicit manner. The next section presents a framework for the fault diagnosis while satisfying the required features of the fault diagnosis using the process correlation model.

2.3.2 Decomposition using PCA models

Based on the PCA model, measurements $\mathbf{x}(m\times 1)$ can be decomposed into the modeled ($\hat{\mathbf{x}}$) and unmodeled parts ($\tilde{\mathbf{x}}$) of measurements ($\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}}$). They are expressed in terms of principal components ($\hat{\mathbf{x}} = \mathbf{Pt} = \mathbf{PP}^{\mathrm{T}}\mathbf{x} = \mathbf{Cx}$ and $\tilde{\mathbf{x}} = (\mathbf{I}-\mathbf{PP}^{\mathrm{T}})\mathbf{x} = (\mathbf{I}-\mathbf{C})\mathbf{x} = \tilde{\mathbf{C}}\mathbf{x}$). $\mathbf{P}(m\times A)$ is orthonormal matrix whose columns are the principal component loadings. Each column is orthogonal to the others. $\hat{\mathbf{x}}$ and $\tilde{\mathbf{x}}$ are orthogonal ($\hat{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{x}} = 0$) due to this orthogonality. By using these relationships, faults and disturbances can be decomposed into two vectors, which explain the fault effects in both the principal component model space and the residual space.

2.3.3 Fault signature

A fault signature can be obtained from steady-state information via plant test, a mechanistic model, or historical fault data. Since it is assumed that historical fault data is available, the fault signatures such as ones shown in Table 2.1, can be obtained. Let the sample vector of measurements for normal operation condition just prior to a fault be denoted by $\mathbf{x}^n$. In the presence of a fault, the sample vector ($\mathbf{x}$) can be represented using an additive fault vector, $\mathbf{f}_i$, ($\mathbf{x} = \mathbf{x}^n + \mathbf{f}_i$). Then a vector for the i-th fault can be decomposed using the PCA model into two components; one ($\hat{\mathbf{f}}_i$) lying in the model space, and the other ($\tilde{\mathbf{f}}_i$) lying in the orthogonal (residual) space as follows:

$$\mathbf{f}_i = \hat{\mathbf{f}}_i + \tilde{\mathbf{f}}_i = \mathbf{C}\mathbf{f}_i + (\mathbf{I}-\mathbf{C})\mathbf{f}_i . \tag{2.5}$$

These vector components are normalized to be insensitive to their magnitudes. That is,

$$\tilde{\mathbf{f}}_i^o = \frac{\tilde{\mathbf{f}}_i}{\|\tilde{\mathbf{f}}_i\|} \quad \hat{\mathbf{f}}_i^o = \frac{\hat{\mathbf{f}}_i}{\|\hat{\mathbf{f}}_i\|} , \tag{2.6}$$

where it is assumed that $\|\tilde{\mathbf{f}}_i\| \neq 0$ and $\|\hat{\mathbf{f}}_i\| \neq 0$. A fault signature library consists of all known fault signature vectors ($i = 1, 2, \ldots, r$) as follows:

$$\hat{\mathbf{F}} = \begin{bmatrix} \hat{\mathbf{f}}_1^o & \hat{\mathbf{f}}_2^o & \cdots & \hat{\mathbf{f}}_r^o \end{bmatrix}, \quad \tilde{\mathbf{F}} = \begin{bmatrix} \tilde{\mathbf{f}}_1^o & \tilde{\mathbf{f}}_2^o & \cdots & \tilde{\mathbf{f}}_r^o \end{bmatrix}. \tag{2.7}$$

These two fault signature matrices include all known fault information in both the modeled and the unmodeled spaces about the r faults.

2.3.4 Angle measures and joint plot for isolation

When a new measurement vector ($\mathbf{x}_{new}$) is available it is decomposed into two components ($\hat{\mathbf{x}}_{new}$ and $\tilde{\mathbf{x}}_{new}$), and then the two components are normalized as follows;

$$\hat{\mathbf{x}}_{new} = \mathbf{C}\mathbf{x}_{new} \quad \tilde{\mathbf{x}}_{new} = (\mathbf{I} - \mathbf{C})\mathbf{x}_{new}, \tag{2.8}$$

$$\hat{\mathbf{x}}^o = \frac{\hat{\mathbf{x}}_{new}}{\|\hat{\mathbf{x}}_{new}\|} \quad \tilde{\mathbf{x}}^o = \frac{\tilde{\mathbf{x}}_{new}}{\|\tilde{\mathbf{x}}_{new}\|}. \tag{2.9}$$

Then, angle measures between the known fault signatures (*i.e.* the columns of $\hat{\mathbf{F}}$ and $\tilde{\mathbf{F}}$) and the new measurement vector signature ($\hat{\mathbf{x}}^o$ and $\tilde{\mathbf{x}}^o$) are used for the fault isolation. Figure 2.3 shows the relationship between the new measurement vector ($\mathbf{x}_{new}$) and an existing fault signature vector ($\mathbf{f}_1$) for a simple fault in variable x_1. The cosines of the angle between the new measurement vector and one of the known fault signatures gives the relative measure of colinearity between them. For instance, the angle measure in the residual space is:

$$\cos\tilde{\theta}_1 = \tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{f}}_1^o \quad (-1 \leq \tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{f}}_1^o \leq 1), \tag{2.10}$$

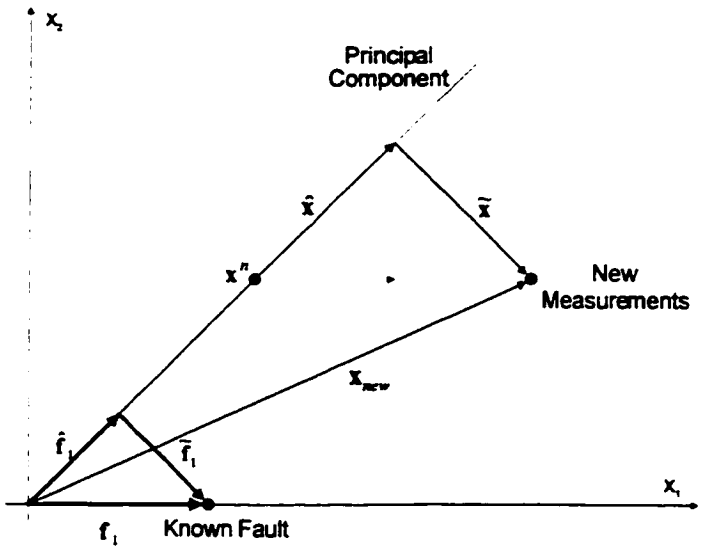


Figure 2.3 Decomposition of the known fault and new faulty measurement vectors

where $\tilde{\mathbf{x}}^o$ and $\tilde{\mathbf{f}}_1^o$ are the normalized residual components of the new measurement vector and the known fault signature for variable x_1, respectively. When the cosine value is close to one, it means that the new measurement vector is nearly collinear to the fault direction. Therefore, once the T^2 and Q charts detect a fault, the fault can be tentatively isolated as the one whose cosine value is close to one, or is the maximum among the row vector cosine components as follows;

$$\max_i(\tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{F}}) = \max_i\left[\tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{f}}_1^o \quad \tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{f}}_2^o \quad \cdots \quad \tilde{\mathbf{x}}^{o\mathrm{T}}\tilde{\mathbf{f}}_r^o\right]. \tag{2.11}$$

The model component of the angle measure is similarly calculated. However, the starting point of the model component of the fault signature must not be taken as the origin point of the score space, but rather as the normal operating point ($\mathbf{x}^n$) just before the fault is detected in Figure 2.3. This is also illustrated in Fig. 2.4 where two identical faults are shown to originate at different operating points. When the model components

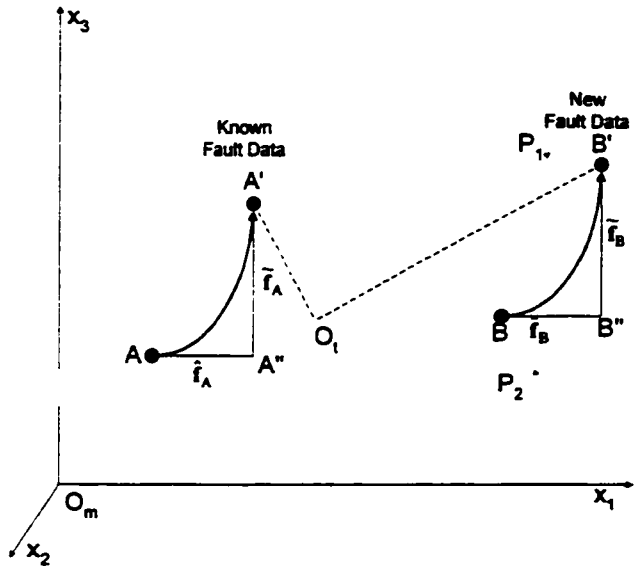


Figure 2.4 Similarity of fault direction in different operation points

of their fault directions are considered with respect to the origin (average *in-control* operating point), $\overline{O_tA"}$ and $\overline{O_tB"}$, respectively, they look different even though both faults are the same. This is due to the starting point of the fault vectors. Therefore, the starting point should not be the origin of score space coordinate, but the points where the faults are initiated, that is, points A and B in Figure 2.4. Thus the model component directions of the faults are $\overline{AA"}$ and $\overline{BB"}$ respectively which indeed look similar. Except for the vertex of the angle measure, all other things are handled in the same manner as with the residual components of fault.

As a simple and convenient isolation tool, a joint angle plot is introduced. The horizontal and vertical axes of the joint plot consist of the residual and model components of the angle measures, respectively. A correctly isolated fault will be one that goes to the upper right or lower left corner of the plot (*i.e.* cosines between the new fault and a reference fault become (+1,+1) or (-1,-1) in the two spaces). Due to noises on all measurements some filtering (e.g., a 1st order filter) is applied to remove unnecessary random variations in the angle measures for both the components of fault signatures. The smaller the filter constant, the higher the random variation of the angle measure is

obtained. The effectiveness of the joint angle plots in isolating faults will be illustrated in the next section.

2.4 Numerical examples

A nonisothermal continuous stirred tank chemical reactor model (Marlin, 1995, pages 90-92) is used to illustrate the proposed method of fault isolation. Detailed model information and simulation parameters are provided in the Appendix. With the base-case condition, the normal operating data of the 9 process variables for 200 minute were collected. After mean-centering and scaling all variables to unit variance, a PCA was estimated. Based on cross-validation, 3 principal components were identified.

2.4.1 Simple vs. Complex faults

Consider a bias fault that suddenly develops in the sensor of one of the measured

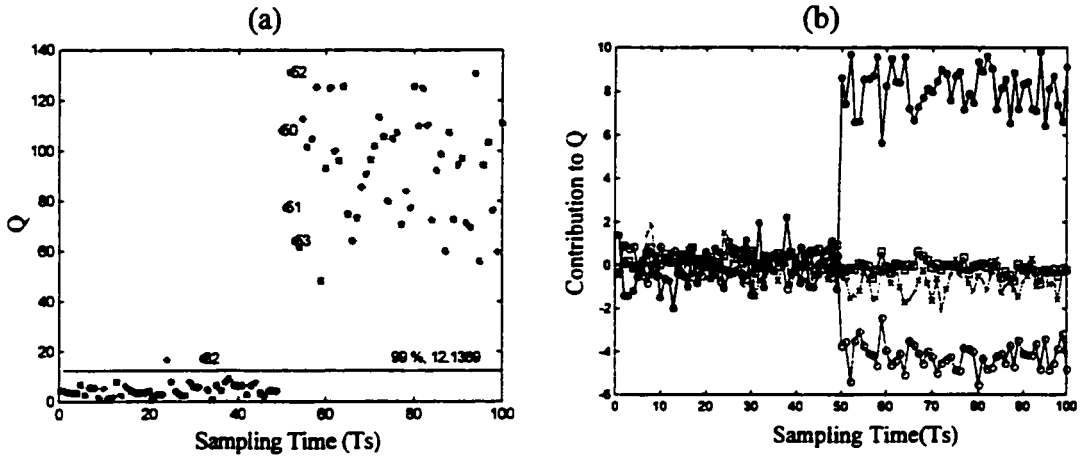


Figure 2.5 Isolation of simple fault by using variable contribution (Bias fault in inlet temperature measurement (T_O) at 50 min with 0.5 °C magnitude; PCA model under only the outlet temperature control); (a) Q monitoring plot with 99% limit; (b) Trend of the variable contributions to Q (T:□; T_O:∗; C_{AA}:○; C_A:×)

disturbances. The inlet temperature (T_O) is one of the process disturbances and its sensor fault is a typical example of a simple fault. A new data set of 100 samples was generated, but this time a bias of 0.5 °C was added to the measurement of the inlet temperature starting at the sampling time of 50. Figure 2.5(a) shows the Q statistic for the new data set along with the 99% control limit and Figure 2.5(b) shows the trends of the variable contributions to this Q statistic. As expected, since the other process variables are unaffected by this simple sensor fault the most significant variable contribution results from the inlet temperature measurement as seen in Figure 2.5(b). However, the contribution of the inlet concentration (C_{AI}) is also large. Thus, even with this simple fault the contribution plots have not given an unambiguous isolation. On the other hand, it has narrowed the fault diagnosis problem down to one of deciding on one of two confounded faults, a positive bias on the inlet temperature sensor or a negative on the inlet concentration sensor. With more measurements such as one would have on real processes, such ambiguities are often resolved and simple faults can easily be isolated with these contribution plots.

Under the same conditions, the bias fault on the measurement of the outlet

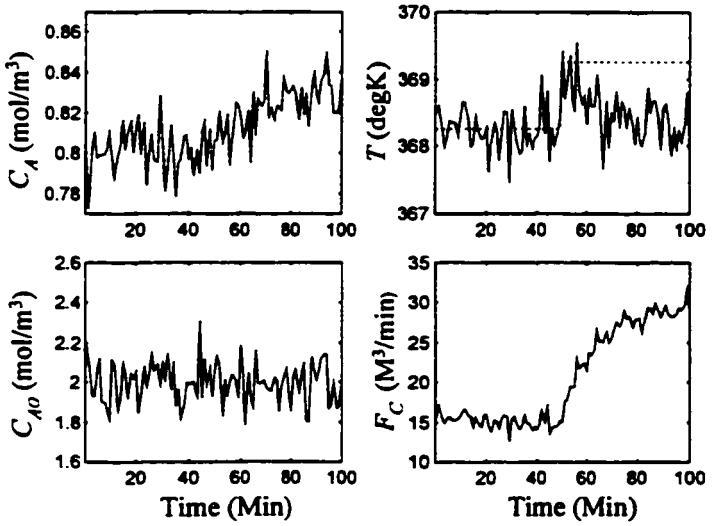


Figure 2.6 Bias fault on outlet temperature measurement under temperature only control (Bias with the magnitude of 1 °C at 50 min, measurement fault (--))

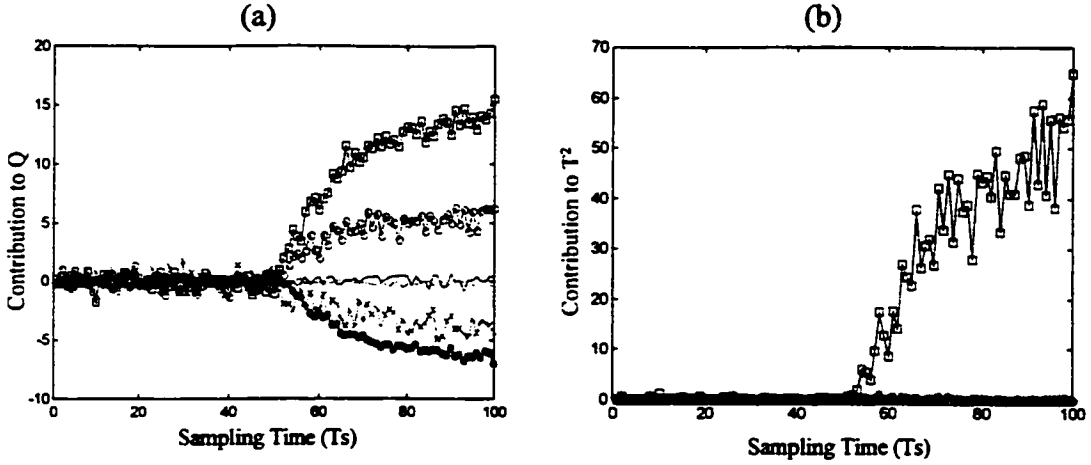


Figure 2.7 Trend of variable contributions for sensor bias fault on outlet temperature; (a) Contribution trends to Q; (b) Contribution trend to Hotelling's T^2 (T:∗; F_C:□; T_C:×; C_A:○)

temperature is simulated as seen in Figure 2.6. Since the outlet temperature is the controlled variable of the process, the bias fault is removed because of the controller action after a short period and instead its effect is distributed into the other variables. The propagated fault effects are shown in the cooling water flow rate and the outlet concentration. The inlet concentration calculated from the flow rates and concentrations of both streams has not been affected because the controller of the outlet concentration was not active in this study.

The variable contributions to the bias fault of the outlet temperature measurement are shown in Figure 2.7. The variable contributions to the Hotelling's T^2 and Q statistic have the transient responses until the process settles to a new steady-state. Note that the fault effects are shown in both spaces (T^2 and SPE). The steady-state values of the variable contributions do not properly indicate the root source of the fault, but instead show the propagated effects of the faults in the other variables. Thus, it is difficult to conclude from these contribution plots that the fault resulted from the measurement of the outlet temperature.

2.4.2 Isolation using the joint angle plot

In this section we illustrate the isolation power of the proposed joint plots of the angles in the residual and model spaces between the vectors of a new fault and those of existing faults in the fault library. Consider a positive bias fault on the F_A measurement occurring after 50 measurements. Figure 2.8(a) is the joint plot for this fault. The angle components in both the residual and model spaces between this fault direction and the directions of the faults in the library are shown. Note that the bias fault on F_A is clearly and unambiguously isolated. The cosines of the angles between this new fault vectors and the F_A fault vectors in the library go to the (+1,+1) corner of the plot indicating essentially perfect colinearity with a F_A fault in both spaces. On the other hand, it is likely that one would be misled if the angle in only one of the spaces were examined. In the residual direction, the new F_A fault has a vector that is also collinear with that of a T_C bias. However, a T_C bias does not align with the new fault vector in the model space. Similarly, the new fault aligns with a C_A bias fault in the model space but is in the wrong direction in the residual space. This confirms that both spaces need to be jointly examined for

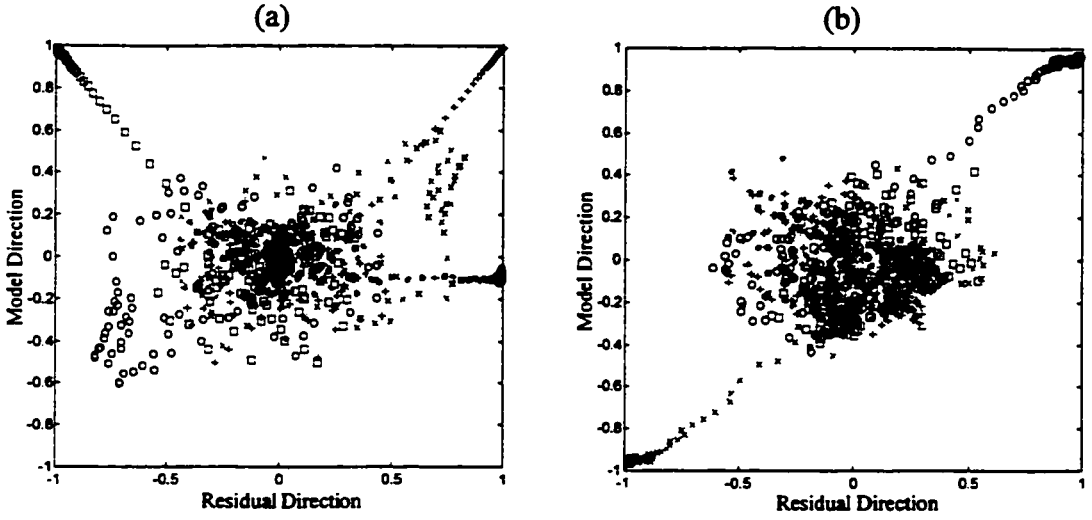


Figure 2.8 Fault isolation using joint plot; (a) Bias fault on F_A with magnitude of 0.02 mol/m^3 at 50 min; (b) Bias fault on T with 1 °C at 50 min (T:○; F_C:×; F_A:+; T_C:∗; C_A:□)

effective fault isolation.

Figure 2.8(b) shows the joint angle plot for a bias fault in the sensor for the outlet temperature (T). This is a typical example of a simple sensor fault becoming a complex fault due to the feedback control system. The apparent error in the measured value of the controlled variable (T) is eventually eliminated by the PI control action using the manipulated variable, namely the coolant flow rate (F_C). Clearly the joint angle plot isolates a bias error in T as a possible expectation for the fault. However, because the joint angles between the current fault and a bias in F_C goes to the (-1,-1) corner, another possible explanation for the fault is a negative bias error in F_C. In fact, these two faults are indistinguishable in this case because they are almost perfectly correlated with the opposite signs as a result of the feedback control.

In practice such ambiguities can be reduced through the use of more process measurements. This is often the case in large industrial systems. Additional measurements that are affected differently by faults in T and F_C would break this ambiguity, and the joint plot would then correctly isolate the bias in T as the fault. In effect, the power of this fault isolation approach with multivariate projection models is greatly enhanced by the use of a larger number of correlated variables.

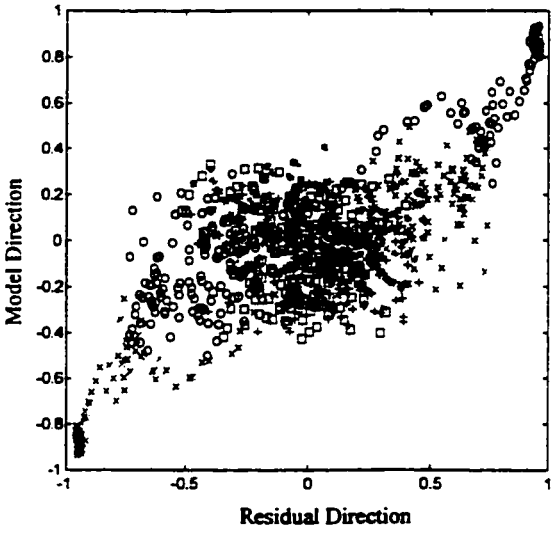


Figure 2.9 Isolation of drift fault on the outlet temperature measurement with the magnitude of 1 °C at 50 min (T:○, F_C:×, F_A:+, T_C:∗, C_A:□)

With most methods, detection and isolation of a drift fault is relatively more difficult task than that of a bias fault. Drift has a small magnitude of fault signature at the start time and it gradually gets bigger. However, the angle measures may give a good isolation of a drift fault if its directionality is consistent over the period. Figure 2.8(b) is for a step bias fault on the outlet temperature measurement and Figure 2.9 is for the drift fault on the same variable. It is shown that the proposed methodology can be successfully used for the isolation of a drift fault as well. However, it is noted that the isolation could be misleading if the initial direction of the fault is very different from its final direction due to nonlinearity.

2.4.3 Effect of the number of principal components

There is sometimes an ambiguity in determining the number of principal components to include in the model developed from the normal operating data. It is shown here that this will not result in any difficulty in the fault isolation step as long as the fault signatures in the model and residual space are both used in the joint angle plots.

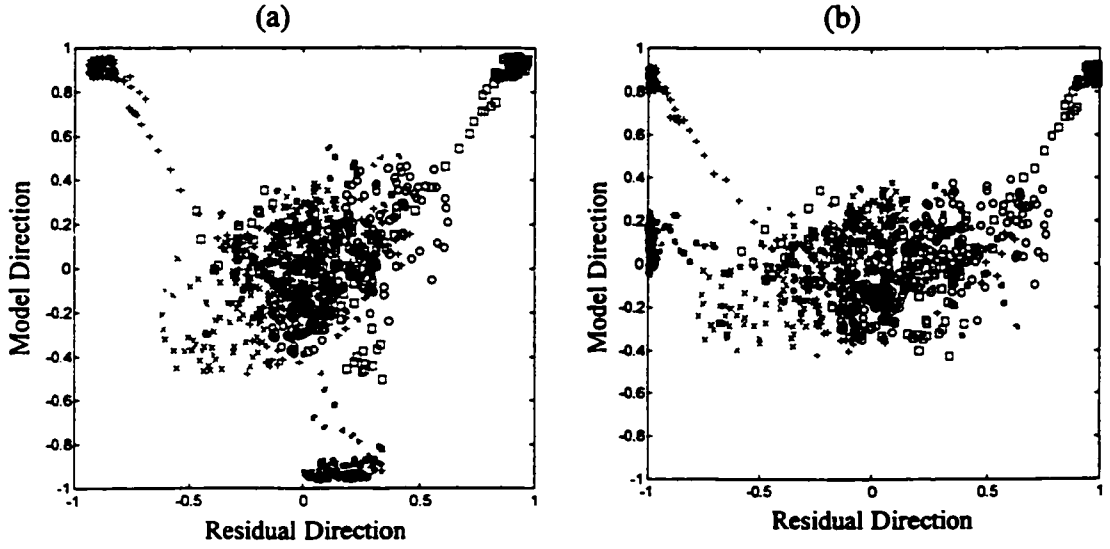


Figure 2.10 Fault isolation with different number of principal components models; (a) Bias fault isolation of C_A with 2 PCs model; (b) Bias fault isolation of C_A with 3 PCs

This arises because the sum of the model and residual spaces is always equal to the full measurement space, and by adding more PCs to the model space, these directions are removed from the residual space. Figure 2.10 shows joint angle plots for the isolation of a bias fault on the outlet concentration (C_A) using the two and three component PCA models. In both cases, the isolation of the C_A fault is clear. The only difference is that the high colinearity of the new fault with a T_C bias has just been transferred from the model space to the residual space with no loss in isolability.

2.5 Conclusion

Multivariate statistical approaches to fault detection based on historical operating data have been shown, in previous literature and in industrial practice, to be very useful when there are large number of measured variables and when causal models are not available. For fault isolation or diagnosis, they have been less powerful because of the non-causal nature of the data on which they are based. Contribution plots in the model space and residual space have been the primary tools for fault diagnosis. These are very effective for isolating simple faults, but much less useful for diagnosing complex faults.

To improve the fault isolation with these methods additional data on past faults has been used to supplement the models. This chapter has provided a critical review of these fault isolation methods, and shown that many of them are only suitable for dealing with simple faults. An approach to fault isolation that is capable of handling both simple and complex faults is developed. This approach extracts fault signatures that are the vectors of movement of the fault in the model space and the residual space. The directions of these vectors are compared with the corresponding vector directions of known faults in a fault library. Isolation is then based on a joint plot of the angles between the vectors of the current fault and those of the known faults. The method is demonstrated using a simulated continuous stirred tank reactor system with feedback control, and shown to be very effective at isolating both simple and complex faults.

The fault signatures used in this chapter have been based on the steady-state

information obtained from a model, from a plant test, or from historical data of the known faults. However, the methodology is applied to the dynamic data obtained from the process as soon as a fault has been detected. The methodology also does not assume the plant is at steady state prior to the fault occurring. But it does assume that prior to and during the fault development, the process is subject to the same time varying *common-cause* disturbances that are present during normal operation. However, the use of only SS fault signatures can lead to a time delay in the fault isolation. In order to reduce the time delay in isolating a fault and to identify it more accurately it is necessary to use dynamic data collected during the evolution of the fault. This can be included in PCA models using time-lagged data, and may allow one to perform better diagnosis of evolving faults. However, in treating dynamically evolving faults, a major problem is how to align them with respect to time against the library of known dynamic faults, and how to scale the data for different magnitudes of faults. These issues will be studied in a subsequent chapter.

3. PCA model of multiscale data

3.1. Introduction

The fault isolation using steady-state fault signature in both PCA model and residual spaces may cause a delay if the transient period is substantial. A false isolation may also arise if the transient fault directionality is different from the steady-state fault signature. This difficulty results from the use of steady-state fault signatures. The use of transient fault trajectory or initial fault signature may thus help fault isolation. In particular, the information on process dynamics obtained by lagging process measurement data and applying MSPC methods would be useful. As an explicit way of dynamic modeling, one can use latent variables and subspace methods (Shi and MacGregor, 2000). With these dynamic models, a variety of causal-model-based fault detection and isolation (FDI) methods can be used. However, the explicit dynamic models and their FDI applications are usually limited by process dimension and data characteristics (Yoon and MacGregor, 2001a). As an alternative, indirect usage of process dynamics for the FDI is considered in this chapter.

Process signals usually represent the cumulative effects of many underlying process phenomena such as process dynamics, measurement noise, external disturbances, process degradation, etc. Each of the phenomena manifests on a different scale (a reciprocal of frequency). Faults occurring at different locations, times and frequencies in a process, would be usually confounded with multiscale natures of the process signals. One can decompose the process signals such that all the contributing events are approximately discriminated according to their frequency contents and hence a fault is

distinguished from other events. The decomposed signals called multiscale data provide information on frequency content that is unavailable with original data. The decomposition can be done with multiresolution analysis via wavelet transformation that offers an efficient characterization of signals resulting in representation on different time and scales (Mallat, 1989a, b; Daubechies, 1988, 1990). Their time-frequency (or scale) localization is an appealing property of wavelets from a signal analysis perspective. Basis functions that are distinguished in both time and scale domains facilitate simultaneous incorporation of time and frequency considerations. Thus, the multiscale analysis based on signal decomposition may simplify the FDI. The main benefits of the signal decomposition are based on discriminating all contributing events of a process and providing a better analysis framework for the FDI application.

Recently, multiscale PCA (MSPCA) has been formulated such that one simultaneously extracts process variable correlations and accounts for auto-correlation within sensor data (Bakshi, 1994, 1998, 1999). It involves decomposing the variables on a selected family of wavelets and developing separate PCA models at each scale. The models at important scales are then combined in an efficient scale-recursive manner to yield the model for all scales together. The MSPCA formulation makes it suitable to work with process signals having a multiscale nature presented by, for example, measurement noises, process disturbances and faults. One may be able to better discriminate these contributing variations distributed at different frequencies with the decompositions of signals. By applying separate MSPC monitoring on each scale one potentially has more power to *detect* faults that occur dominantly in that scale. Fault detection has been successfully done with the MSPCA concept, but fault isolation has not been dealt with. Similarly, Luo *et al.* (1999) proposed sensor fault detection via multiscale analysis and dynamic PCA. In their work, high-frequency information was tested and classified as noise. Low-frequency information was presumed to represent the effects of process dynamics on the sensor, and sensor malfunctions. What was left to analyze was the middle frequency ranges, and PCA was applied on these data. Kosanovich and Piovoso (1997) also used PCA of wavelet transformed data for

monitoring. They pre-filtered the data with a finite impulse response median hybrid filter (FMH), and developed the calibration model from wavelet coefficients obtained from a Haar wavelet transform (HWT) of the FMH-filtered data. They claimed that the proposed method produced a calibration model with better classification features. The calibration model was developed by employing PCA on the HWT coefficients, but not on the FMH filtered data. Various applications in chemical engineering using wavelet transformation are well reported in literature (Motard and Joseph, 1994).

In this chapter, a unified framework that generalizes MSPCA as well as PCA is presented, and its promising features for the FDI are addressed. The unified PCA model (GPCA) is obtained by applying wavelet transforms on measurement signals, rearranging the decomposed measurement matrices side by side in the order of ascending scale, and applying PCA on the rearranged fat matrix. The choice of orthonormal wavelets is necessary in order to have block orthogonality in the way that the principal components of the GPCA are the decomposition of those of the PCA. Due to the generalizing property of the GPCA for the PCA as well as the MSPCA, the resulting method inherits a few beneficial aspects of both methods. The scale contributions available from the GPCA model can provide additional diagnosing information for the fault isolation.

In section 3.2, it is shown how muiltiscale data of a finite length signal is generated. The multiresolution analysis and its implementation issues are then addressed. The GPCA and the various properties are presented in section 3.3. In section 3.4, several illustrating examples are considered using a simulated CSTR system with feedback control. The promising aspects of GPCA are scrutinized by comparing it to PCA and MSPCA. The comprehensive application study to FDI problem using GPCA model will be considered in the subsequent chapter.

3.2 Decomposition of a finite length signal

3.2.1 Signal decomposition

Relevant process variations usually can be characterized by their frequency

contents. Thus, one may more implicitly describe the process variations with decomposed components of a signal containing many such process variations. The signal decomposition according to their scale (inverse of frequency) contents can be implemented with multi-channel filter banks built by cascading two-channel banks, which have the property of splitting a signal into two lower-resolution versions. A two-channel filter bank splits a signal into a lowpass, or coarse resolution component, and a highpass component. This decomposition is recursively applied on the lowpass version. This leads a hierachy of multiresolution decomposition. Long time intervals are considered for analysis of high scale contents and shorter regions for analysis of low scale contents. The key concept is to write a signal as a limit of successive approximations, each of which is a smoother version of a signal. The successive approximations correspond to different resolutions. A given signal will be represented by a coarse approximation plus added details. The coarse and detail subspaces are orthogonal to each other since the detail signal is the difference between the original, or coarse signal at the next higher level and the coarse version of the signal.

In the early 1980s, Morlet first proposed the wavelet transform as a tool for signal analysis. In the context of signal processing, he investigated an alternative to local Fourier analysis based on a single prototype function, and its scales and shifts. The modulation by complex exponentials in the Fourier transform is replaced by a scaling operation, and the notion of scale replaces that of frequency. This led to the discovery of wavelets, which form orthonormal bases for square-integrable and other function spaces by Meyer (1993) and Daubechies (1992), and others. A formalization of such constructions by Mallat (1989b) and Meyer (1992) created a framework for multiresolution analysis, and established links with methods used in other fields. One of the first links between wavelet theory and signal processing was due to Daubechies (1988) and Mallat (1989b). The pyramid coding technique, which builds up a signal from its lower resolution version plus a sequence of details, is closely related to wavelet theory and multiresolution analysis, and filter banks or subband coding schemes can be used for the computation of wavelet decompositions. For complete overview on the history and

the theoretical study to applications, refer to literature (Vetterli and Herley, 1992; Vetterli and Kovacevic, 1995; Strang and Nguyen, 1996). In the following, the discussion is limited to the discrete-time wavelet series and its filter bank realization under the multiresolution analysis.

In the discrete version of the multiresolution analysis, it is assumed that the initial, discrete data, which is a sequence, already represents an approximation at a certain scale that is related to the sampling interval. By convention, this scale is fixed at j=0. This is a finest resolution, associated with the space V_0 (Rioul, 1993). One cannot refine the signal further. Then, a finite number of decomposition steps J leads to a coarsest resolution associated with V_J. By definition, a multiresolution analysis consists of a sequence of embedded closed spaces ($V_J \subset \cdots \subset V_2 \subset V_1 \subset V_0$). The space V_0 is the space of all square-summable sequences. That is, $V_0 = l_2\{Z\}$. The orthogonal complement of V_{j+1} in V_j will be denoted by W_{j+1}, and thus $V_j = V_{j+1} \rho W_{j+1}$ with $V_{j+1} \zeta W_{j+1}$, where ρ denotes the direct sum. Assuming that there exists a sequence $g_0[n] \in V_0$ such that $\{g_0[n-2k]\}_{k \in Z}$ is a basis for V_1, then it can be shown that there exists a sequence $g_1[n] \in V_0$ such that $\{g_1[n-2k]\}_{k \in Z}$ is a basis for W_1. Such a sequence is given by $g_1[n] = (-1)^n g_0[-n+1]$. In other words, $\{g_0[n-2k], g_1[n-2k]\}_{k \in Z}$ is an orthonormal basis for V_0. This splitting can be iterated on V_1. Therefore, V_0 can be decomposed by simply iterating the decomposition J times and $V_0 = W_1 \rho W_2 \rho \cdots \rho W_J \rho V_J$

The input is decomposed into a very coarse resolution which exists in V_J and added details which exist in the spaces W_j, j=1, ..., J, where V_j's are called approximation spaces and W_j's detail spaces. The sum of the coarse version and all the added details yields back the original signal. One starts with its lower-resolution version belonging to V_J, and adds up the details until the final resolution is reached.

3.2.2 Multiscale representation of a signal

A general filter bank with low and high pass analysis filters $h_0[n]$, $h_1[n]$ and

synthesis filters $g_0[n]$, $g_1[n]$, implements a biorthogonal expansion of discrete-time signals. Such a system is called a perfect reconstruction filter bank. A multichannel (Tree-Structured) filter bank, constructed by cascading two-channel banks, implements a discrete-time biorthogonal wavelet series. If the two-channel filter bank is orthonormal, then it implements a discrete time orthonormal wavelet series. In the special sampling used in the discrete-time wavelet series, each subsequent channel is downsampled by 2 with respect to the previous one and has a bandwidth that is reduced by 2 as well. This is called a dyadic sampling.

Consider a two-channel filter bank with filters $h_0[n]$, $h_1[n]$ $g_0[n]$, and $g_1[n]$. Repeating the filter bank J times, one obtains the discrete time wavelet series over J octaves, plus the final octave containing the lowpass version. In Figure 3.1(a), the multiresolution analysis is shown with multichannel filter banks with J stages. Decomposition spaces V_j and W_j are indicated. Analysis and synthesis filters are represented by $H_i(z)$ and $G_i(z)$ where $i = 1$ for a high pass and $i = 0$ for a low pass filter. A signal sequence, $x[n]$ is expressed as

$$x[n] = \sum_{j=1}^{J} \sum_{k \in Z} X^{(j)}[2k+1] g_1^{(j)}[n - 2^j k] + \sum_{k \in Z} X^{(J)}[2k] g_0^{(J)}[n - 2^J k], \tag{3.1}$$

where

$$X^{(j)}[2k+1] = \left\langle h_1^{(j)}[2^j k - l], x[l] \right\rangle, \quad j = 1, \cdots, J, \tag{3.2a}$$

$$X^{(J)}[2k] = \left\langle h_0^{(J)}[2^J k - l], x[l] \right\rangle. \tag{3.2b}$$

These are the convolutions of the input with $h_0[n]$ and $h_1[n]$ evaluated at even indexed $2^j k$. In these equations, $h_i^{(1)}[n] = h_i[n]$ and $g_i^{(1)}[n] = g_i[n]$. Because any input sequence can be decomposed as above, the family of functions $\{g_1^{(j)}[2^j k - n], g_0^{(J)}[2^J k - n]\}, j = 1, \cdots, J$, and k, $n \in Z$ is an orthogonal basis for $l_2(Z)$. The basis functions $g_{(i)}^{(j)}[n]$ are defined as the time domain versions of $G_0^{(j)}(z)$ and $G_1^{(j)}(z)$. That is, $g_0^{(2)}[n]$ is the time-domain version of

$G_0^{(2)}(z) = G_0(z)G_0(z^2)$ while $g_1^{(2)}[n]$ is the time-domain version of $G_1^{(2)}(z) = G_0(z)G_1(z^2)$. Given the general filters $g_i[n]$ and a number of states J, the equivalent synthesis filters (basis functions) in an octave-band filter bank with J stages are given by

$$G_0^{(J)}(z) = G_0^{(J-1)}(z)G_0(z^{2^{J-1}}) = \prod_{K=0}^{J-1} G_0(z^{2^K}), \tag{3.3}$$

$$G_1^{(j)}(z) = G_0^{(j-1)}(z)G_1(z^{2^{j-1}}) = G_1(z^{2^{j-1}})\prod_{K=0}^{j-2} G_0(z^{2^K}), \tag{3.4}$$

where $G_0^{(1)}(z) = G_0(z)$, $G_1^{(1)}(z) = G_1(z)$, and j=1, ..., J. The expressions for analysis filters, $H_0^{(J)}(z)$ and $H_1^{(J)}(z)$ for j=1, ..., J can be similarly obtained. Using these expressions and the multirate identity for filtering by sampling, the octave-band filter banks in Figure 3.1(a) can be shown as in Figure 3.1(b). It implies that the multiresolution analysis consists of several filter banks which are the combination of analysis and synthesis filters, downsamplers, and upsamplers. If $h_i[n]$ is an orthogonal filter, $h_i[-n] = g_i[n]$ and thus $h_i^{(J)}[n] = g_i^{(J)}[-n]$. This structure implements an orthogonal discrete time wavelet series expansion. Given a mother wavelet and a number of states J, J+1 filter banks are thus expressed in terms of moving average filters. Figure 3.1(c) illustrates the wavelet series expansion in terms of impulse responses of time series filter banks in case of Daubechies-5 mother wavelet and J=4.

Applying filter banks merely entails doing a linear expansion of a signal. The filter banks can be written as an infinite matrix operation. The action of a two-channel analysis filter on an infinite signal column vector $\mathbf{x}$ is represented by using the analysis filter matrix, $\mathbf{F}_A^T = \begin{bmatrix}\mathbf{H}_1^T & \mathbf{H}_0^T\end{bmatrix}^T$ where the premultiplication by the operator, $\mathbf{H}_i$ has the effect of filtering the signal by $H_i(z)$ and subsampling by 2 (represented by the shift by 2 in $\mathbf{H}_i$). $\mathbf{H}_i$ is defined as:

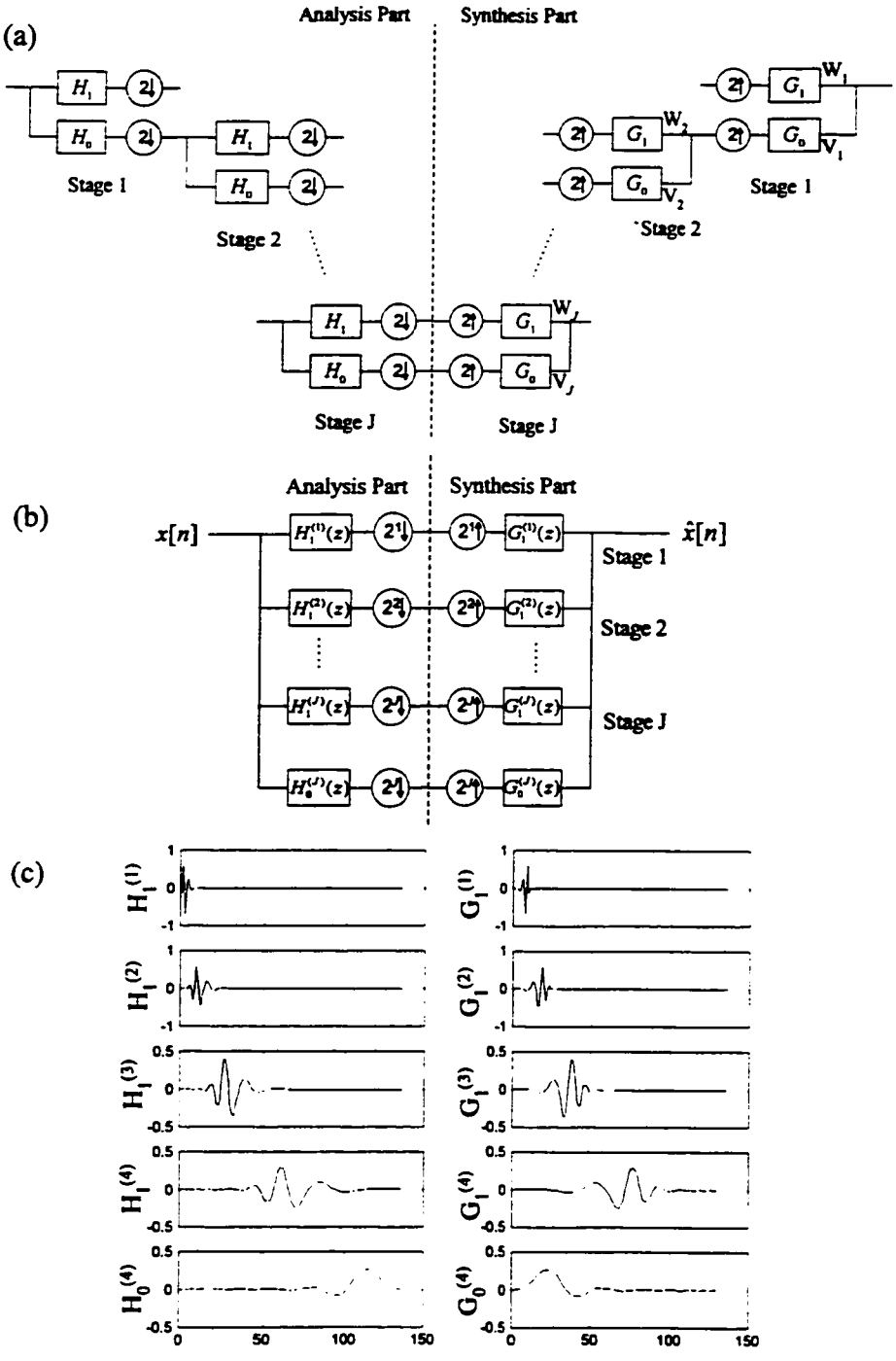


Figure 3.1 Multiresolution analysis; (a) Cascade representation of filter banks with J stages; (b) Block representation of filter banks with J stages; (c) Impulse responses of filter banks with Daubechies-5 and 4 levels

$$\mathbf{H}_i = \begin{pmatrix} \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \\ \cdots & h_i[L-1] & h_i[L-2] & h_i[L-3] & \cdots & h_i[0] & 0 & 0 & \cdots \\ \cdots & 0 & 0 & h_i[L-1] & \cdots & h_i[2] & h_i[1] & h_i[0] & \cdots \\ & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \end{pmatrix}, \tag{3.5}$$

where each line is an even shift of the impulse response of $h_i[n]$ and L is the filter length. The analysis filter outputs are hence represented by the approximation, $\mathbf{y}_0=\mathbf{H}_0\mathbf{x}$ and the detail at the first level, $\mathbf{y}_1=\mathbf{H}_1\mathbf{x}$. The projections on the low-pass component are recursively decomposed to obtain coarser approximations. The projections on the high-pass component that is wavelets at each scale contain the finer details. Thus,

$$\mathbf{x}_a^{(j)} = \mathbf{H}_0\mathbf{x}_a^{(j-1)}, \text{ and } \mathbf{x}_d^{(j)} = \mathbf{H}_1\mathbf{x}_a^{(j-1)}. \tag{3.6}$$

After J decompositions, the transformation of $\mathbf{x}$, or $\mathbf{x}_a^{(0)}$ is expressed in terms of $\mathbf{x}_a^{(J)}$ and $\mathbf{x}_d^{(J)}$;

$$\mathbf{W}_A\mathbf{x}_a^{(0)} = \left[(\mathbf{x}_d^{(1)})^{\mathrm{T}} \quad (\mathbf{x}_d^{(2)})^{\mathrm{T}} \quad \cdots \quad (\mathbf{x}_d^{(J)})^{\mathrm{T}} \quad (\mathbf{x}_a^{(J)})^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{3.7}$$

where, $\mathbf{x}_a^{(j)}$ and $\mathbf{x}_d^{(j)}$ are column vectors whose lengths are $n/2^j$. Assuming n is a relatively large number, and the analysis transformation matrix, $\mathbf{W}_A$ is expressed as;

$$\begin{aligned} \mathbf{W}_A &= \underbrace{Diag\{\overbrace{\mathbf{I},\cdots,\mathbf{I}}^{2^{(J-1)}-1},\mathbf{F}_A\}\cdots Diag\{\mathbf{I},\mathbf{I},\mathbf{I},\mathbf{F}_A\}Diag\{\mathbf{I},\mathbf{F}_A\}}_{J-1}\mathbf{F}_A \\ &= \left[\mathbf{H}_1 \quad \mathbf{H}_1\mathbf{H}_0 \quad \cdots \quad \mathbf{H}_1\underbrace{\mathbf{H}_0\cdots\mathbf{H}_0}_{J-1} \quad \underbrace{\mathbf{H}_0\cdots\mathbf{H}_0}_{J}\right]^{\mathrm{T}} = \left[\mathbf{H}_1^{(1)} \quad \mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{H}_1^{(J)} \quad \mathbf{H}_0^{(J)}\right]^{\mathrm{T}}, \end{aligned} \tag{3.8}$$

where, $\mathbf{H}_1^{(j)}$ is the matrix containing wavelet filter coefficients corresponding to scale j and $\mathbf{H}_0^{(J)}$ is the matrix of scaling function filter coefficients at the coarsest scale. It is

noted that the dimension of the component matrix, $\mathbf{H}_i$ in $\mathbf{H}_1^{(J)}$ and $\mathbf{H}_0^{(J)}$ is not the same as the original dimension of $\mathbf{H}_i$ $(n/2^1, n)$ For example,

$$\mathbf{H}_0^{(J)} = \mathbf{H}_0(n/2^J, n/2^{J-1})\cdots\mathbf{H}_0(n/2^2, n/2^1)\,\mathbf{H}_0(n/2^1, n)\,, \tag{3.9}$$

$$\mathbf{H}_1^{(j)} = \mathbf{H}_1(n/2^j, n/2^{j-1})\mathbf{H}_0(n/2^{j-1}, n/2^{j-2})\cdots\mathbf{H}_0(n/2^2, n/2^1)\,\mathbf{H}_0(n/2^1, n)\,. \tag{3.10}$$

$\mathbf{H}_1^{(j)}$ and $\mathbf{H}_0^{(J)}$ are expressed in terms of successive filters, $\mathbf{H}_0$ and $\mathbf{H}_1$, espectively. Thus, $\mathbf{W}_A\mathbf{x}$ is the same size as the original data sequence, but due to the wavelet decomposition, the low-frequency component is concentrated in a relatively small number of coefficients in $\mathbf{x}_a^{(J)}$, while the other components in each variable are approximately decorrelated and spread over all components in $\mathbf{x}_d^{(j)}$, $j = 1,...,J$, according to its power spectrum.

The decomposed matrix at each scale in (3.7) $\mathbf{x}_a^{(j)}$ and $\mathbf{x}_d^{(j)}$ does not have the same number of rows. One can reconstruct all the terms with the synthesis filter such that they have the same number of rows. Based on the synthesis filter matrix, $\mathbf{F}_S^T = \begin{bmatrix}\mathbf{G}_1^T & \mathbf{G}_0^T\end{bmatrix}^T$, the transformation matrix for the reconstruction $\mathbf{W}_S$ is obtained. The reconstruction step is required to have all the decomposed signals in their original units. The reconstruction algorithm is similar to the decomposition algorithm except the steps are now reversed. Finer details captured by the wavelets are added to the coarser approximations to reconstruct the original signal. Therefore, the reconstructed signal is

$$\begin{aligned}\hat{\mathbf{x}} &= \mathbf{W}_S\mathbf{W}_A\mathbf{x} = \mathbf{W}_S\begin{bmatrix}(\mathbf{x}_d^{(1)})^T & (\mathbf{x}_d^{(2)})^T & \cdots & (\mathbf{x}_d^{(J)})^T & (\mathbf{x}_a^{(J)})^T\end{bmatrix}^T \\ &= \mathbf{G}_1^{(1)}\mathbf{x}_d^{(1)} + \mathbf{G}_1^{(2)}\mathbf{x}_d^{(2)} + \cdots + \mathbf{G}_1^{(J)}\mathbf{x}_d^{(J)} + \mathbf{G}_0^{(J)}\mathbf{x}_a^{(J)} \\ &= \mathbf{G}_1\mathbf{x}_d^{(1)} + \mathbf{G}_0\mathbf{G}_1\mathbf{x}_d^{(2)} + \cdots + \underbrace{\mathbf{G}_0\mathbf{G}_0\cdots\mathbf{G}_0}_{J-1}\mathbf{G}_1\mathbf{x}_d^{(J)} + \underbrace{\mathbf{G}_0\mathbf{G}_0\cdots\mathbf{G}_0}_{J}\mathbf{x}_a^{(J)}\end{aligned}\,, \tag{3.11}$$

In terms of the analysis and synthesis filter banks, and the original signal, the reconstructed signal is

$$\hat{\mathbf{x}} = (\mathbf{G}_1^{(1)}\mathbf{H}_1^{(1)} + \mathbf{G}_1^{(2)}\mathbf{H}_1^{(2)} + \cdots + \mathbf{G}_1^{(J-1)}\mathbf{H}_1^{(J-1)} + \mathbf{G}_1^{(J)}\mathbf{H}_1^{(J)} + \mathbf{G}_0^{(J)}\mathbf{H}_0^{(J)})\mathbf{x}, \qquad (3.12)$$

where, the last term is the contribution of the coarsest approximation of **x**, and the remaining terms are the contributions of the details from the coarsest to the finest levels.

3.2.3 Boundary corrected filters for a finite length signal

The filter bank necessarily operates over infinite signals (the matrix is infinite along both dimensions) and the filters do not change with time. The application to finite lengths will generally involve either distortion at the boundary or the introduction of some redundancy. A typical way around this difficulty has been to treat the finite signal as a segment of an infinite one formed by replicating boundary values or periodizing, or padding with zeros (Karlsson and Vetterli, 1989). However, some distortion remains at the boundary. A symmetric periodic extension of the signal and process with linear phase filters will give less distortion or redundancy, but the filters are restricted to be symmetric. If a finite unitary matrix has the same block structure as the infinite unitary matrix **T** that will be explained below, one can get a square unitary matrix for any size for a given filter set. It was shown how this problem could be overcome in the case of two-channel orthogonal filter banks, by using boundary filters (Herley *et al.*, 1993; Herley and Vetterli, 1994; Herley, 1995). In this section, how an orthogonal filter can be obtained for a finite signal is presented.

Consider an orthogonal wavelet. The coefficients for the analysis filters, $\mathbf{H}_1$ and $\mathbf{H}_0$, and the synthesis filters $\mathbf{G}_1$ and $\mathbf{G}_0$, are determined by $H_1(z) = z^{-(L-1)}H_0(-z^{-1})$ and $G_i(z) = H_i(z^{-1})$, where L is the filter length. This is known as the orthogonal solution and the coefficients, $h_1[n]$ and $h_0[n]$ are orthogonal with respect to even shifts

$$\begin{aligned} \langle h_0(n), h_0(n-2k)\rangle &= \langle h_1(n), h_1(n-2k)\rangle = \delta(k) \\ \langle h_1(n), h_0(n-2k)\rangle &= \langle h_0(n), h_1(n-2k)\rangle = 0 \end{aligned} \qquad (3.13)$$

Because of the time reversal, the synthesis filters are represented by $\mathbf{H}_1^*$ and $\mathbf{H}_0^*$, where * denotes Hemitian transpose. In all cases we shall be interested only in filters with real coefficients so that $\mathbf{H}_i^* = \mathbf{H}_i^T$. The perfect reconstruction requires $\mathbf{F}_A^T\mathbf{F}_A = \mathbf{I}$ since $\hat{\mathbf{x}} = \mathbf{F}_A^T\mathbf{F}_A\mathbf{x} = \mathbf{x}$. Its columns are orthogonal and so are its rows: $\mathbf{F}_A\mathbf{F}_A^T = \mathbf{I}$. In block form this means that

$$\begin{bmatrix}\mathbf{H}_1^T & \mathbf{H}_0^T\end{bmatrix}\begin{bmatrix}\mathbf{H}_1 \\ \mathbf{H}_0\end{bmatrix} = \mathbf{H}_1^T\mathbf{H}_1 + \mathbf{H}_0^T\mathbf{H}_0 = \mathbf{I}, \tag{3.14}$$

$$\begin{bmatrix}\mathbf{H}_1 \\ \mathbf{H}_0\end{bmatrix}\begin{bmatrix}\mathbf{H}_1^T & \mathbf{H}_0^T\end{bmatrix} = \begin{bmatrix}\mathbf{H}_1\mathbf{H}_1^T & \mathbf{H}_1\mathbf{H}_0^T \\ \mathbf{H}_0\mathbf{H}_1^T & \mathbf{H}_0\mathbf{H}_0^T\end{bmatrix} = \begin{bmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\end{bmatrix}. \tag{3.15}$$

That the system gives perfect reconstruction is shown by the fact that the sum of the two paths in an identity (3.14), $\mathbf{H}_1^T\mathbf{H}_1 + \mathbf{H}_0^T\mathbf{H}_0 = \mathbf{I}$. The orthogonality properties above in the time domain, (3.14) and (3.15) imply orthogonality properties of the matrices, $\mathbf{H}_1\mathbf{H}_1^T = \mathbf{H}_0\mathbf{H}_0^T = \mathbf{I}$ and $\mathbf{H}_1\mathbf{H}_0^T = \mathbf{H}_0\mathbf{H}_1^T = \mathbf{0}$.

An equivalent representation can be found by interleaving the rows of $\mathbf{H}_1$ and $\mathbf{H}_0$. With this ordering of the rows, a matrix for decomposition is

$$\mathbf{T} = \begin{bmatrix}
\ddots & & & & & & & & & & \\
\cdots & \mathbf{0} & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \\
\cdots & \mathbf{0} & \mathbf{0} & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} & \mathbf{0} & \mathbf{0} & \cdots \\
\cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} & \mathbf{0} & \cdots \\
 & & & & & & & & & & \ddots
\end{bmatrix}, \tag{3.16}$$

where

$$\mathbf{A}_i = \begin{bmatrix}h_0(2K-2i-1) & h_0(2K-2i-2) \\ h_1(2K-2i-1) & h_1(2K-2i-2)\end{bmatrix}. \tag{3.17}$$

Since an orthogonal filter bank necessarily implies that T is unitary, the rows of T form

an orthonormal basis for $l^2(z)$, the space of square summable sequences. $\mathbf{T}$ is a doubly infinite block Toeplitz matrix with blocks of size 2×2 blocks since the filters in the two-channel case are always of even length, there are $K=L/2$ such 2×2 blocks. For perfect reconstruction, $\mathbf{T}\mathbf{T}^T=\mathbf{T}^T\mathbf{T}=\mathbf{I}$. However, the unitary matrix is based on infinite signal and so its properties are not valid for finite signal. Define $\mathbf{T}_t$, a finite unitary matrix as

$$\mathbf{T}_t=\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & & & \cdots & & & & \cdots & & & \vdots \\ \vdots & & & \cdots & & & & \cdots & & & \vdots \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \cdots & \mathbf{A}_{K-1} \end{bmatrix}. \tag{3.18}$$

The truncated matrix contains p block rows of the infinite analysis matrix. The dimensions are then $2p\times 2(p+K-1)$. Thus, $\mathbf{T}_t\mathbf{T}_t^T=\mathbf{I}$ where $\mathbf{I}$ is the $2p\times 2p$ identity matrix, but $\mathbf{T}_t^T\mathbf{T}_t\neq\mathbf{I}$. $\mathbf{T}_t$ operates on a vector of length $2(p+K-1)$, but produces only $2p$ expansion coefficients. Then it has no left inverse since the analysis and synthesis matrices are not mutually inverse and square. To meet the requirement, one needs to add $2(K-1)$ new rows to $\mathbf{T}_t$ and $2(K-1)$ new columns to $\mathbf{T}_t^T$. These new rows and columns should preserve the mutual orthogonality relations. Then $\mathbf{T}_t^T$ and $\mathbf{T}_t$ are mutually inverse. Since $\mathbf{T}_t\mathbf{T}_t^T=\mathbf{I}$, $(\mathbf{T}_t^T\mathbf{T}_t)^2=\mathbf{T}_t^T\mathbf{T}_t$ and $\mathbf{P}_t=\mathbf{T}_t^T\mathbf{T}_t$ is a projection. For any column $\mathbf{x}$, $\mathbf{T}_t^T\mathbf{T}_t\mathbf{x}$ is the projection of $\mathbf{x}$ onto the columnspace of $\mathbf{T}_t^T$, and $\mathbf{x}^T\mathbf{T}_t^T\mathbf{T}_t$ is the projection of $\mathbf{x}^T$ onto the rowspace of $\mathbf{T}_t$. Any row orthogonal to all the rows of $\mathbf{T}_t$ will have nonzero values in the first and last $2(K-1)$ positions only and is of the form

$$\mathbf{h}_{(i)}=\mathbf{x}_{(i)}^T\left(\mathbf{I}-\mathbf{T}_t^T\left(\mathbf{T}_t\mathbf{T}_t^T\right)^{-1}\mathbf{T}_t\right), \tag{3.19}$$

$\mathbf{h}_{(i)}$ are orthonormal to all columns of $\mathbf{T}_t^T$ and suitable to add as new rows to $\mathbf{T}_t$. Based on $\mathbf{T}_t$, $\mathbf{H}_1$ and $\mathbf{H}_0$ for a finite length signal are obtained (Herley, 1995).

3.3 Generalized principal component analysis

3.3.1 Generation of multiscale data

In the previous section, it was shown how a data sequence is decomposed and expressed in terms of its contributions at different scales. The same transformation can be applied to all the measurement data. Let **X** be the observation matrix of dimension, $n \times m$, where X_{ij} is the value of the jth measurement at time i. All the columns of **X** can be decomposed into the details at all levels and the approximation at the coarsest level.

$$\begin{aligned}\mathbf{W}_S\mathbf{W}_A\mathbf{X} &= \left(\mathbf{G}_1^{(1)}\mathbf{H}_1^{(1)} + \mathbf{G}_1^{(2)}\mathbf{H}_1^{(2)} + \cdots + \mathbf{G}_1^{(j)}\mathbf{H}_1^{(j)} + \cdots + \mathbf{G}_1^{(J)}\mathbf{H}_1^{(J)} + \mathbf{G}_0^{(J)}\mathbf{H}_0^{(J)}\right)\mathbf{X} \\ &= \mathbf{G}_1\mathbf{H}_1\mathbf{X} + \mathbf{G}_0\mathbf{G}_1\mathbf{H}_1\mathbf{H}_0\mathbf{X} + \cdots + \underbrace{\mathbf{G}_0\cdots\mathbf{G}_0}_{j-1}\mathbf{G}_1\mathbf{H}_1\underbrace{\mathbf{H}_0\cdots\mathbf{H}_0}_{j-1}\mathbf{X} + \\ &\cdots + \underbrace{\mathbf{G}_0\cdots\mathbf{G}_0}_{J-1}\mathbf{G}_1\mathbf{H}_1\underbrace{\mathbf{H}_0\cdots\mathbf{H}_0}_{J-1}\mathbf{X} + \underbrace{\mathbf{G}_0\cdots\mathbf{G}_0}_{J}\underbrace{\mathbf{H}_0\cdots\mathbf{H}_0}_{J}\mathbf{X}\end{aligned}, \tag{3.20}$$

where each term represents a scale contribution of the original data matrix in terms of reconstructed values. All the terms have the same number of rows and columns. Since we are interested in scale contributions, all the above terms are split and rearranged. One can characterize the original data by three parameters of variable(m), sample time(n), and scale(J). Figure 3.2 shows how to decompose the original data, reconstruct, and rearrange the reconstructed scale components in the orthogonal wavelet case. All the scale contributions are split and laid side by side to produce a two-dimensional matrix of size $[n \times (J+1)m]$. For instance, the first m columns correspond to all the variables over all the sampled times at scale J=1. Define the rearranged matrix in Figure 3.3 as $\mathbf{X}_G$

$$\mathbf{X}_G = \mathbf{V}Diag(\mathbf{X})_{J+1}, \tag{3.21}$$

where $\mathbf{V} = \left[\mathbf{G}_1^{(1)}\mathbf{H}_1^{(1)} \quad \mathbf{G}_1^{(2)}\mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{G}_1^{(j)}\mathbf{H}_1^{(j)} \quad \cdots \quad \mathbf{G}_1^{(J)}\mathbf{H}_1^{(J)} \quad \mathbf{G}_0^{(J)}\mathbf{H}_0^{(J)}\right]$ and $Diag(\mathbf{X})_{J+1}$

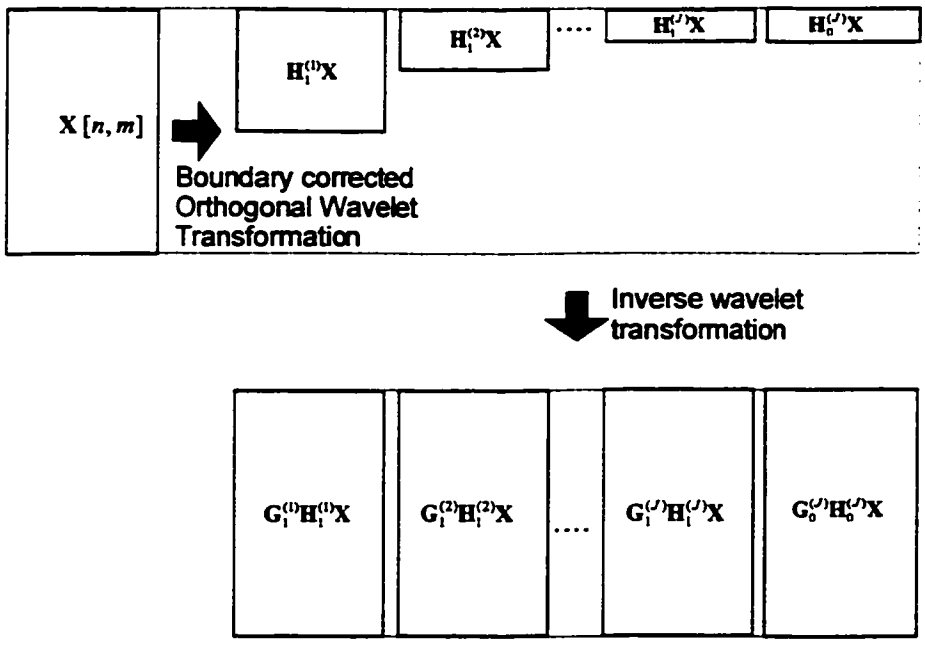


Figure 3.2 Data decomposition and reconstruction

is a diagonal matrix whose diagonal component is **X** and off-diagonal terms are zero matrices with the same dimension as **X**. Its dimension becomes $[(J+1)n, (J+1)m]$. If one chooses an orthogonal mother wavelet, $\mathbf{G}_0=\mathbf{H}_0^*$, $\mathbf{G}_1=\mathbf{H}_1^*$, then $\mathbf{G}_0^{(J)} = [\mathbf{H}_0^{(J)}]^T$, $\mathbf{G}_1^{(J)} = [\mathbf{H}_1^{(J)}]^T$. In this case, **V** is simplified in term of the analysis filter matrices as follows

$$\mathbf{V} = \left[\mathbf{H}_1^{(1)T}\mathbf{H}_1^{(1)} \quad \mathbf{H}_1^{(2)T}\mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{H}_1^{(J)T}\mathbf{H}_1^{(J)} \quad \cdots \quad \mathbf{H}_1^{(J)T}\mathbf{H}_1^{(J)} \quad \mathbf{H}_0^{(J)T}\mathbf{H}_0^{(J)} \right]. \qquad (3.22)$$

3.3.2 GPCA model

Since PCA model is concerned only with relative differences between objects, the averages of each variable are first subtracted from each column in data matrix. Then, the mean-centered data is rescaled by dividing each column by its standard deviation. This gives each variable equal weight and makes the variation the same for all variables along

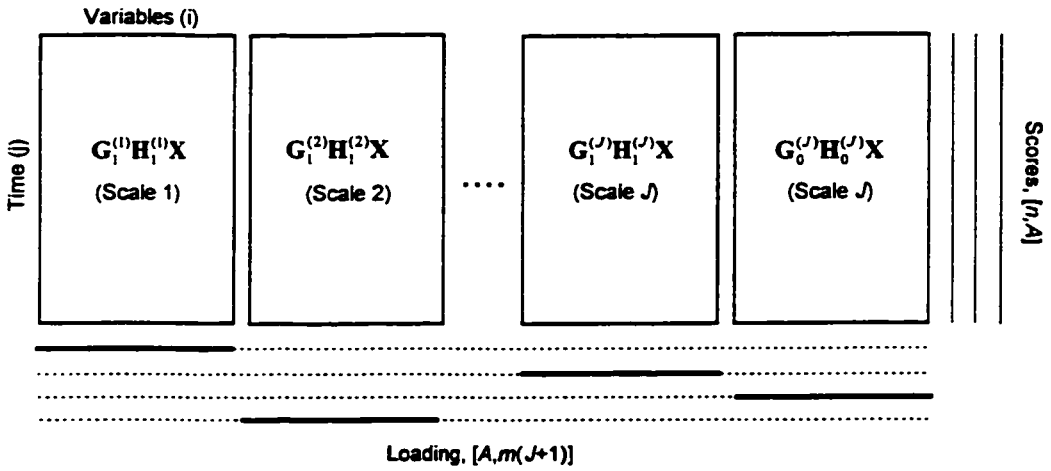


Figure 3.3 GPCA model of multiscale data

their axes (Wold *et al.*, 1989).

In case of the GPCA model, there are two extreme cases of the scaling. The first is to scale **X** rather than $\mathbf{X}_G$. The GPCA model that is now estimated with the unscaled **X** reveals the scale contributions on the total process variations. Then, the orders of the principal components are determined according to the scale contributions as well as the magnitudes of the process correlations. This type of scaling would be appropriate for denoising and detrending since one can selectively include significant principal components for the user's applications. The other way is to scale $\mathbf{X}_G$. This type of scaling will equally weigh all the variables of each scale as is done in the conventional PCA modeling. But one will lose information on the scale contribution to the total process variations. It is likely to distort the true process variations over the examined scales.

Therefore, one needs to carefully select an appropriate scaling method. When prior information about the importance of the variables in a given problem is available, this should be used in the scaling (Wold *et al.*, 1989). For example, frequency characteristics of all the process variations can be used for the scaling. In case of the fault detection and isolation, Pareto type of scaling can be used if frequency characteristics of the faults are known.

With the carefully determined scaling methods, the GPCA model is then

estimated with $\mathbf{X}_G$ (Figure 3.3). In the following example, we examine the differences among PCA, MSPCA and GPCA. By estimating all three models with auto-correlated data, their modeling properties on process dynamics are compared. Then the scaling effect on the GPCA model is addressed.

Illustration

Signals from most processes, unless they are sampled slowly relative to the process time constants, are auto-correlated due to the effects of process dynamics, disturbances, etc. Consider an autoregressive time series process of 1st order, $y_t=\phi y_{t-1}+e_t$ where $e_t \sim N(0,1)$. Frequency characteristic of a signal depends on the coefficient (ϕ) that changes the degree of autocorrelation. By changing the model coefficient, one has the effect of changing a process time constant. For low value of the model coefficient ϕ, power spectral density (PSD) is evenly distributed over frequencies but its low frequency part becomes dominant as ϕ increases as shown in Figure 3.4(a). For data generation, the driving force, σ_e^2 is adjusted to keep $\sigma_{y_t}^2$ constant. The generated data is augmented with the previous observations, $\mathbf{X}=[y_t\ y_{t-1}]$. Then this augmented matrix is scaled to unit variance and used for the estimation of the correlation models.

Consider three cases of 0.1, 0.5, and 0.9 for ϕ values. The eigenvectors of the covariance matrix of $\mathbf{X}$ are the same for all three cases as shown in Figure 3.4(b). The first eigenvector corresponds to the positive correlation between y_t and y_{t-1}, and the second to the negative correlation. Note that the ratio of the first to the second eigenvalue increases as ϕ increases. This ratio represents the magnitude of the process autocorrelation. This tendency is confirmed by the 99 % ellipsoidal areas of the data for the three cases in Figure 3.4(b).

Now consider PCA, MSPCA and GPCA model with the same data $\mathbf{X}$ scaled to unit variance. The PCA model gives the same information as obtained above since it is a different expression of the eigenvalue/eigenvector of the covariance matrix of $\mathbf{X}$.

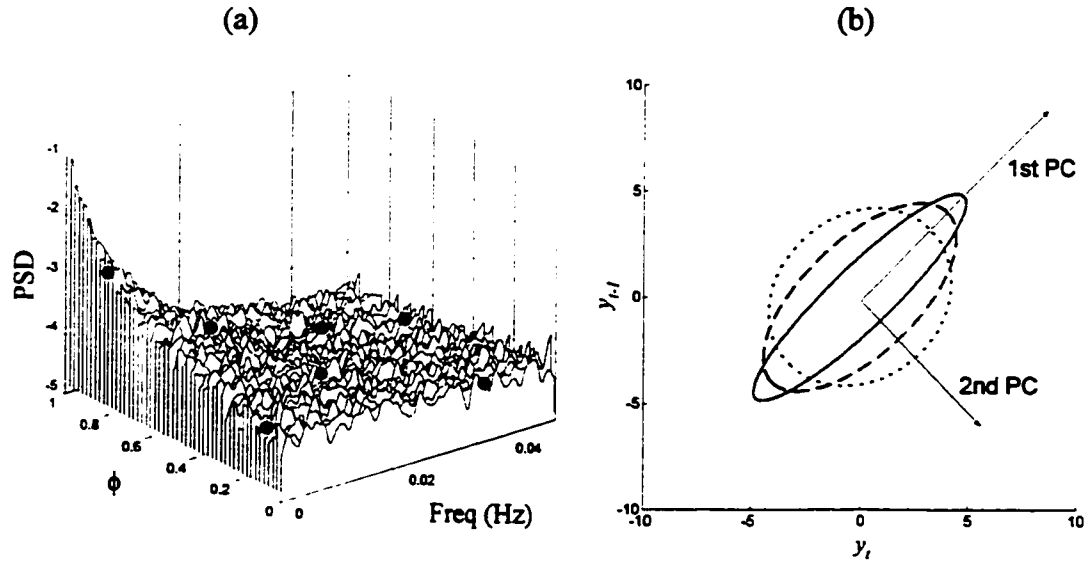


Figure 3.4 (a) Effect of AR(1) coefficient on frequency content (Power Spectral Density) of auto-correlated signal; (b) Eigenvectors and 99 % confidence regions of augmented matrix, $\mathbf{X}=[y_t\ y_{t-1}]$ for three cases of N, 0.1, 0.5 and 0.9 (No. of observations: 360; dotted: $N=0.1$; dashed: $N=0.5$; solid: $N=0.9$)

Consider two cases of ϕ=0.1 as a nearly auto-correlated process and ϕ=0.9 as a severe auto-correlated case that the most significant variation occurs at low frequency band. In Figure 3.5(a) and (b), it is shown that the first principal component corresponding to the positive correlation between y_t and y_{t-1} explains 82.3 % of the total process variations in the severely auto-correlated case. On the other hand, the first principal component only explains 54.3 % of the total process variations in the nearly auto-correlated case. This observation is consistent with the simulation conditions given. PCA thus can capture the process auto-correlation with the time lagged measurement data.

X is now decomposed with Daubechies-5 mother wavelet into 4 details and one approximation. Two principal components in each block are identified. From the MSPCA model shown in Figure 3.5(c) and (d), it is shown that the negative correlation between y_t and y_{t-1} is dominant at D_1, the finest detail block while the positive correlations between y_t and y_{t-1} are dominant at all the other blocks. This correlation distribution over scales is

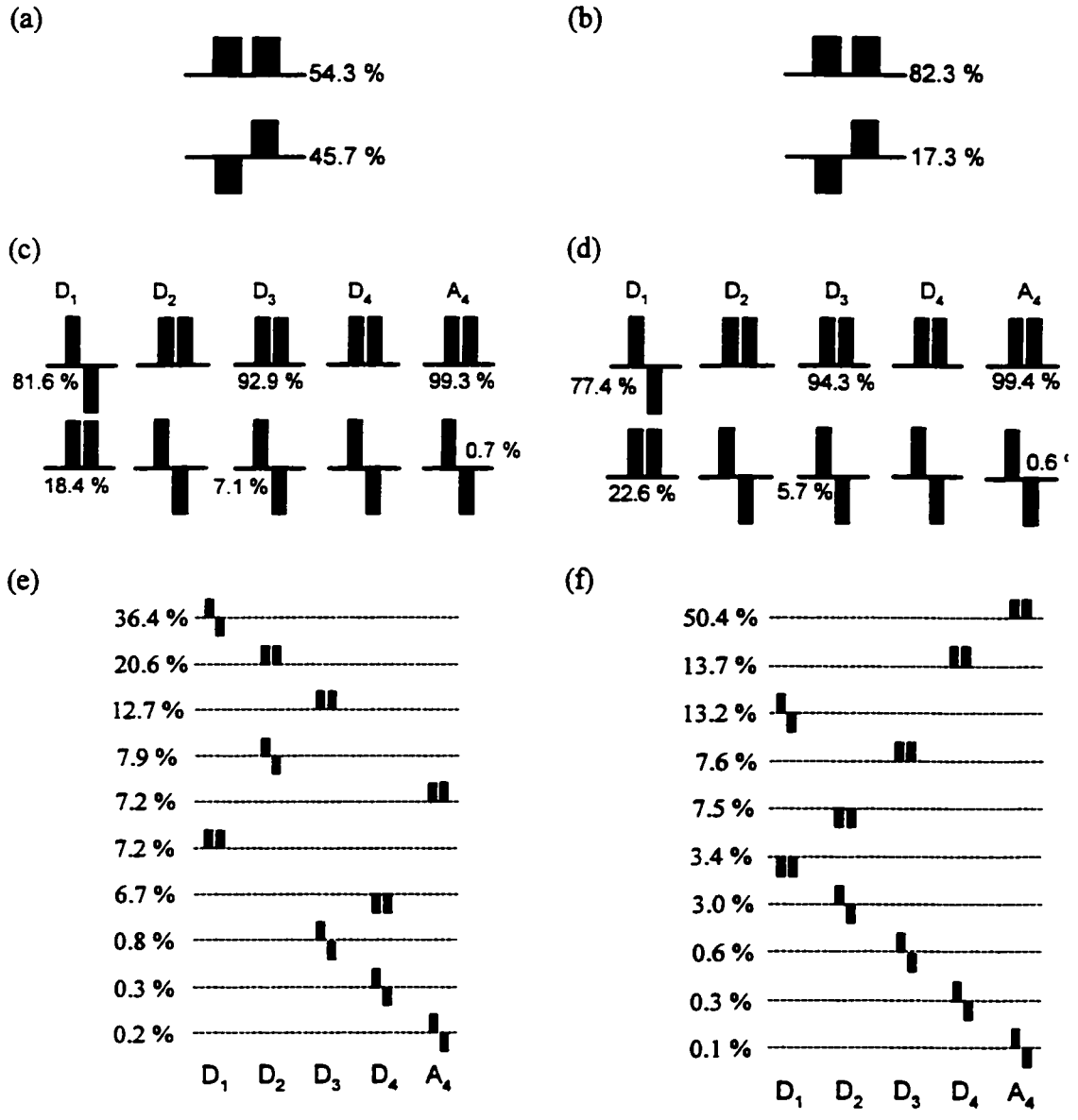


Figure 3.5 Correlation models of auto-correlated data $\mathbf{X}=[y_t\ y_{t-1}]$; (a) PCA model of **X** for $N = 0.1$; (b) PCA model of **X** for $N = 0.9$; (c) MSPCA model of **X** for $N = 0.1$ (daubechies-5 mother wavelet; Scales: 4); (d) MSPCA model of **X** for $N = 0.9$; (e) GPCA model of **X** for $N = 0.1$; (f) GPCA model of **X** for $N = 0.9$.

additional information of the MSPCA. However, one can hardly distinguish the difference in the two cases with the MSPCA model since almost the same information is obtained from the two cases. It is because one is looking at all the decomposed data

blocks separately.

As in (3.21), $\mathbf{X}_G$ is formulated and the PCA model of $\mathbf{X}_G$ is estimated. Figures 3.5(e) and (f) show the GPCA models for the two cases, respectively. When $N = 0.1$, the first principal component corresponds to the negative correlation caused by the random error function at the highest frequency level. The two orthogonal variations are shown as evenly distributed over scales. On the other hand, when $N = 0.9$, the positive correlation becomes dominant as shown with the 1st, 2nd, 4th, 5th, and 6th principal components. It is obvious that the correlation distribution of the highly positively auto-correlated case is clearly different from that of the nearly uncorrelated case. It is because the positive correlation between y_t and y_{t-1} becomes more significant than the negative correlation as N increases. The process dynamic changes made by changing N have been clearly reflected in the GPCA model. This is because GPCA model orders all the principal components with their magnitudes and their frequency characteristics, whereas MSPCA simply presents the PCA component at each scale without regard to the ordering.

Scaling effect

With the two data sets used in the above example, the eigenvalue magnitudes of the two principal components at each scale are shown as filled and empty bars in Figure 3.6(a) - (d). Each combined bar represents the scale contribution to the total variation. Figure 3.6(a) and (b) correspond to the case of unit variance scaling on $\mathbf{X}$, and Figure 3.6(c) and (d) to the case of unit variance scaling on $\mathbf{X}_G$. Based on the eigenvalues which are arranged in descending order of the magnitudes, a stopping criteria is applied to capture significant process correlations by choosing the corresponding eigenvectors. The GPCA model consists of these eigenvectors.

Figure 3.6(a) and (b) show the very different pattern in the eigenvalue arrangements for the two cases of different autocorrelations (N=0.1and 0.9 respectively) and unit scaling on the original $\mathbf{X}$. The original scale contributions to the process variations can be inferred with the GPCA model. This scaling would be useful for denoising or detrending since one can keep the information on the scale contributions to

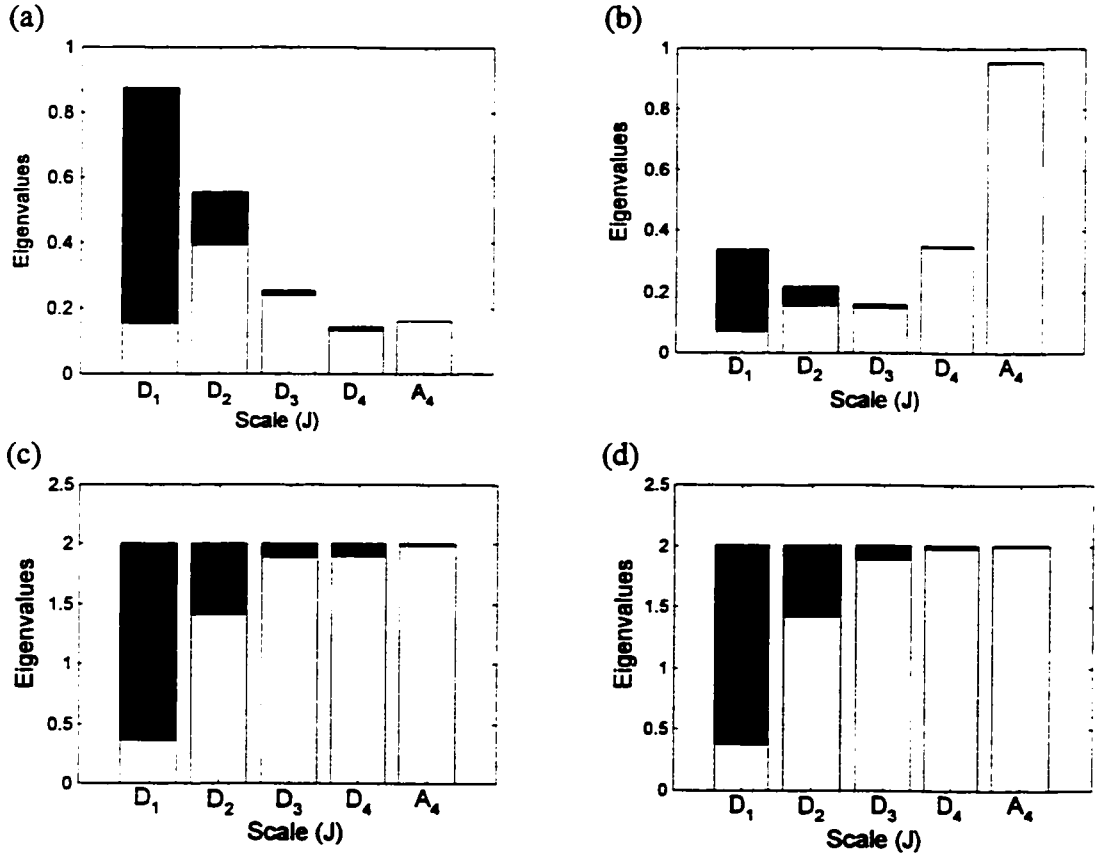


Figure 3.6. Scaling effect on multiscale correlation models of auto-correlated data $\mathbf{X}=[y_t\ y_{t-1}]$; (a) Eigenvalues at each scale for $N = 0.1$ and scaling on $\mathbf{X}$; (b) $N = 0.9$ and scaling on $\mathbf{X}$; (c) $N = 0.1$ and scaling on $\mathbf{X}_G$; (d) $N = 0.9$ and scaling on $\mathbf{X}_G$ (Daubechies-5 mother wavelet; 4 decomposition levels; filled bar: negative correlation; empty bar: positive correlation)

the total process variation with the GPCA model. On the other hand, Figure 3.6(c) and (d) show the similar pattern in the eigenvalue arrangements for the two cases of the different autocorrelations ($N = 0.1$ and 0.9) and unit scaling on the original $\mathbf{X}_G$. It is shown that the original scale contributions to the process variations cannot be inferred with the GPCA model. The scaling on $\mathbf{X}_G$ gives each decomposed variable of $\mathbf{X}_G$ equal weight and makes the variances the same for all decomposed variables. This scaling may be useful for fault detection and isolation in which all the decomposed variables are needed.

3.3.3 Model properties

The usage of an orthogonal mother wavelet retains the orthogonality among decomposed blocks. The GPCA model that uses the rearranged multiscale data, $\mathbf{X}_G$ obtained with the orthogonal wavelet, inherits a few useful properties from the wavelet orthonormality and the related filter properties.

Block orthogonality

Due to the orthogonality, $\mathbf{H}_1^{(i)}\mathbf{H}_0^{(J)\mathrm{T}} = \mathbf{H}_0^{(J)}\mathbf{H}_1^{(i)\mathrm{T}} = \mathbf{0}$ for i=1, ..., J and $\mathbf{H}_1^{(i)}\mathbf{H}_1^{(j)\mathrm{T}} = \mathbf{0}$ for $i \neq j$, the block off-diagonal terms of $\mathbf{X}_G{}^{\mathrm{T}}\mathbf{X}_G$ have all zero values. The covariance matrix of $\mathbf{X}_G{}^{\mathrm{T}}\mathbf{X}_G$ becomes a block diagonal matrix. Therefore, there will be no interactions between blocks, and the principal component will be a function of variables within only a scale block.

Proof:

$$\mathbf{H}_0^{(J)}\mathbf{H}_1^{(J)\mathrm{T}} = \underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{J}\underbrace{\mathbf{H}_0^{\mathrm{T}} \cdots \mathbf{H}_0^{\mathrm{T}}}_{J-1}\mathbf{H}_1^{\mathrm{T}} = \underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{J-1}\underbrace{\mathbf{H}_0^{\mathrm{T}} \cdots \mathbf{H}_0^{\mathrm{T}}}_{J-2}\mathbf{H}_1^{\mathrm{T}} = \mathbf{H}_0\mathbf{H}_1^{\mathrm{T}} = \mathbf{0} \quad \left(\because \mathbf{H}_0\mathbf{H}_0^{\mathrm{T}} = \mathbf{I}\right)$$

$$\mathbf{H}_1^{(i)}\mathbf{H}_1^{(j)\mathrm{T}} = \mathbf{H}_1\underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{i-1}\underbrace{\mathbf{H}_0^{\mathrm{T}} \cdots \mathbf{H}_0^{\mathrm{T}}}_{j-1}\mathbf{H}_1^{\mathrm{T}} = \mathbf{H}_1\underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{i-2}\underbrace{\mathbf{H}_0^{\mathrm{T}} \cdots \mathbf{H}_0^{\mathrm{T}}}_{j-2}\mathbf{H}_1^{\mathrm{T}} = \mathbf{H}_1\underbrace{\mathbf{H}_0^{\mathrm{T}} \cdots \mathbf{H}_0^{\mathrm{T}}}_{j-i}\mathbf{H}_1^{\mathrm{T}} = \mathbf{0}.$$

Generalization of PCA and Multiscale approaches

The block diagonal terms of $\mathbf{X}_G{}^{\mathrm{T}}\mathbf{X}_G$ have non-zero values. Their block diagonal terms are the same as those of individual blocks in MSPCA, and the sum of the block diagonal terms becomes an identity matrix. When one considers the zero level of decomposition, GPCA becomes PCA. MSPCA is a special case of GPCA without the information on the percentages of process variations explained by the principal components. Thus the GPCA model unifies PCA as well as MSPCA.

Proof: Since $\mathbf{H}_0^{(J)}\mathbf{H}_0^{(J)^T} = \underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{J}\underbrace{\mathbf{H}_0^T \cdots \mathbf{H}_0^T}_{J} = \underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{J-1}\underbrace{\mathbf{H}_0^T \cdots \mathbf{H}_0^T}_{J-1} = \mathbf{H}_0\mathbf{H}_0^T = \mathbf{I}$ and

$\mathbf{H}_1^{(J)}\mathbf{H}_1^{(J)^T} = \mathbf{H}_1\underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{j-1}\underbrace{\mathbf{H}_0^T \cdots \mathbf{H}_0^T}_{j-1}\mathbf{H}_1^T = \mathbf{H}_1\underbrace{\mathbf{H}_0 \cdots \mathbf{H}_0}_{j-2}\underbrace{\mathbf{H}_0^T \cdots \mathbf{H}_0^T}_{j-2}\mathbf{H}_1^T = \mathbf{H}_1\mathbf{H}_1^T = \mathbf{I}$, the multiplication matrix of the analysis and synthesis filters in one block becomes an identify matrix. Thus,

$$\mathbf{H}_0^{(J)}\mathbf{H}_0^{(J)^T} = \mathbf{H}_1^{(J)}\mathbf{H}_1^{(J)^T} = \cdots = \mathbf{H}_1^{(j)}\mathbf{H}_1^{(j)^T} = \cdots \mathbf{H}_1^{(2)}\mathbf{H}_1^{(2)^T} = \mathbf{H}_1^{(1)}\mathbf{H}_1^{(1)^T} = \mathbf{I}_n$$

Then, the covariance of $\mathbf{X}_G$ is simplified as follows

$$\mathbf{X}_G^T\mathbf{X}_G = \begin{bmatrix} \mathbf{X}^T\mathbf{H}_1^{(1)^T}\mathbf{H}_1^{(1)}\mathbf{X} & & & & & \\ & \mathbf{X}^T\mathbf{H}_1^{(2)^T}\mathbf{H}_1^{(2)}\mathbf{X} & & \mathbf{0} & & \\ & & \cdots & & & \\ & & & \mathbf{X}^T\mathbf{H}_1^{(j)^T}\mathbf{H}_1^{(j)}\mathbf{X} & & \\ & & & & \cdots & \\ & \mathbf{0} & & & \mathbf{X}^T\mathbf{H}_1^{(J)^T}\mathbf{H}_1^{(J)}\mathbf{X} & \\ & & & & & \mathbf{X}^T\mathbf{H}_0^{(J)^T}\mathbf{H}_0^{(J)}\mathbf{X} \end{bmatrix}$$

$$= \mathbf{X}^T Diag\left(\mathbf{H}_1^{(1)^T}\mathbf{H}_1^{(1)} \quad \mathbf{H}_1^{(2)^T}\mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{H}_1^{(j)^T}\mathbf{H}_1^{(j)} \quad \cdots \quad \mathbf{H}_1^{(J)^T}\mathbf{H}_1^{(J)} \quad \mathbf{H}_0^{(J)^T}\mathbf{H}_0^{(J)}\right)\mathbf{X} \quad (3.23)$$

The summation of all components in the above diagonal matrix is the n dimensional identity matrix because $\mathbf{I} = \mathbf{H}_0^{(J)^T}\mathbf{H}_0^{(J)} + \sum_{j=1}^{J}\mathbf{H}_1^{(j)^T}\mathbf{H}_1^{(j)}$. Since $\mathbf{H}_1^{(J)^T}\mathbf{H}_1^{(J)} + \mathbf{H}_0^{(J)^T}\mathbf{H}_0^{(J)} = \mathbf{H}_0^{(J-1)^T}\mathbf{H}_0^{(J-1)}$, the following holds

$$\mathbf{H}_0^{(J)^T}\mathbf{H}_0^{(J)} + \sum_{j=1}^{J}\mathbf{H}_1^{(j)^T}\mathbf{H}_1^{(j)} = \mathbf{H}_0^{(J-1)^T}\mathbf{H}_0^{(J-1)} + \sum_{j=1}^{J-1}\mathbf{H}_1^{(j)^T}\mathbf{H}_1^{(j)}. \quad (3.24)$$

One can recursively apply the same identity relation, and it ends up with an identity matrix. This shows that $\mathbf{X}_G$ includes the same information as $\mathbf{X}$. The GPCA model of multiscale data thus includes information equivalent to the one that decomposes the principal components of the PCA model of the single scale data, $\mathbf{X}$. This decomposed

model captures the *common-cause* variations at different scales. As a result, the PCA model can be considered an extreme case of the GPCA model when no scale decomposition is executed. Right singular vectors (Principal components) of $\mathbf{X}_G$ are equal to the collections of those of all scale components of **X**. The same principal components obtained with the PCA model of multiscale data thus can be estimated by separately applying PCA on all the scale components of **X** (Bakshi, 1998). The difference is that the GPCA model orders all these principal components according to the magnitudes of their contributions to the total variances. On the other hand, MSPCA based correlation models are obtained by individually applying PCA on all scale components. With the PCA model of multiscale data, one can consider all process variations at different scales at the same time, and arrange all the principal components in the order of their contributions to the total variance, while MSPCA does not provide relative comparison of process variations distributed at different scales.

Clustering of multiscale data

When all the scale characteristics of dominant events in a process are not shown in the orders of $2^j T_S$(Sampling time) and are non-dyadically spaced from lower to higher scales, one may not need to investigate all the decomposed scales. One can then cluster the several scales at which the similar correlations are shown. This will significantly reduce the number of scales to be analyzed and simplify the multiscale analysis. On the other hand, when all the decomposed scale blocks show the same correlations at all the level, the PCA model of multiscale data does not give any additional information other than obtained by PCA model. In such an occasion, one does not need to decompose the original data and apply PCA models for process analysis. By using this property, one can confirm if multiscale analysis is required.

Detrending or denoising

Once a multiscale data based PCA model has been built, one can identify a scale

or a principal component which is not needed for process monitoring. One can then remove one or all the principal components within the suspected scale. Removing all the principal components within a scale has an effect of using a bandpass filter. But one cannot selectively filter one of the principal components at a given scale by using a bandpass filter. Thus the GPCA based modeling method provides a flexible framework in building the sensitive model relevant for the process monitoring. For example, assume that the lowest frequency component is the day and night effect on the cooling water temperature in a process. Several variables affected by temperature changes of the cooling water will have the same correlation. Depending on the magnitudes of other causes of process variation, the intensity of the cooling water correlation shown in the other variables differs. If one determines that this component is not required for further analysis and wants to remove it, one can remove the corresponding principal component from the further analysis. Whether one removes all the principal components of a scale, or just one of them, depends on how many of them are related with unnecessary process variations. Then the related effect will not be considered.

3.4 Numerical examples

In this section, a GPCA modeling on multivariate process is considered, and the essential differences among GPCA, PCA and MSPCA are illustrated through the comparison. Using the non-isothermal CSTR model explained in Appendix A, the training data was collected for 200 minutes under routine operation when no faults are present. The data collection interval (T_S) was 10 seconds.

In the simulation system, all the disturbances were modeled as 1st order auto-regressive processes, $x_t = \phi x_{t-1} + \sigma_x e_t$ where $e_t \sim N(0,1)$ and σ_x^2 is a variance of x_t. Thus, the disturbances introduce different frequency characteristics into the process depending on their value of ϕ and the location where they enter the process. This fact is confirmed in Figure 3.7 in which the power spectral densities (PSD) of all the process variables are

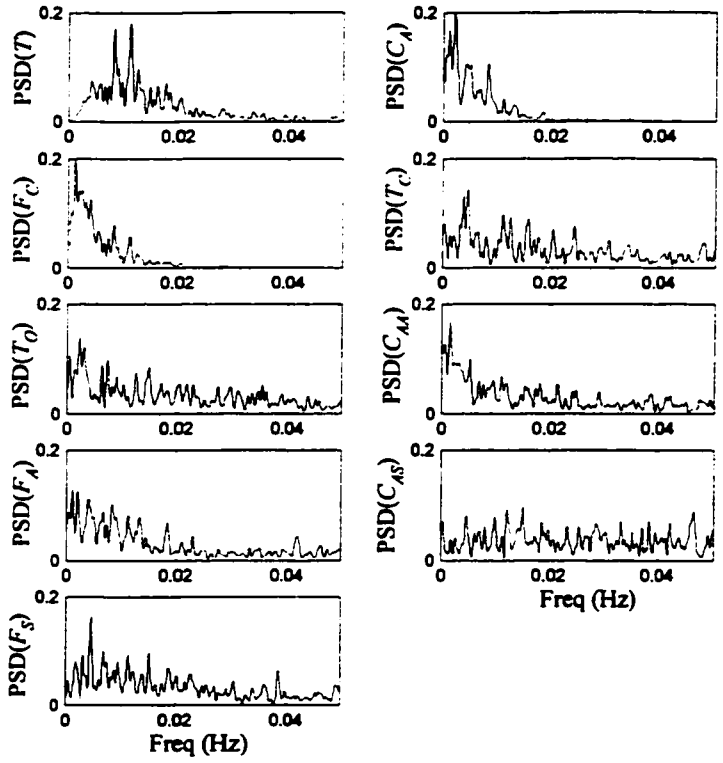


Figure 3.7 Power spectral densities of base-case simulation data

shown. The PSD of T is very different from those of the others. The PSD of F_C and C_A show similar patterns. Their dominant frequency contents are shown at the frequencies lower than 0.015 Hertz, which corresponds to a period of $8T_S$. F_S, T_C, T_O and F_A form a group showing different frequency characteristics from those of F_C and C_A. This group has quite intensive frequency characteristics between the critical (0.05 Hz) and lower frequency range, and has similar patterns as those of F_C and C_A at the lower frequency region. C_{AA} and C_{AS} show evenly distributed frequency contents over all frequency ranges. These different frequency characteristics of all the process variables may validate the multiscale analysis for the CSTR simulation data.

3.4.1 Comparison to PCA and MSPCA

The collected data was decomposed into 4 detail and one approximation blocks

corresponding to $2T_S$, $4T_S$, $8T_S$, $16T_S$, and the frequencies lower than $16T_S$. Daubechies-5 wavelet was used as a mother wavelet to generate the multiscale data. Then, PCA and MSPCA/GPCA models were calculated with the single-scale and multiscale data, respectively. Figure 3.8 shows principal component loadings of the three models. Each cell represents the effects of all 9 variables- T, C_A, F_C, T_C, T_O, C_{AA}, F_A, C_{AS}, and F_S, respectively.

The PCA model consists of 3 principal components. The first principal component explains the control action. It shows that the cooling water flow (F_C), the reactor outlet concentration (C_A) and the reactor outlet temperature (T) are positively correlated. The second shows a negative correlation between the solvent flow (F_S) and

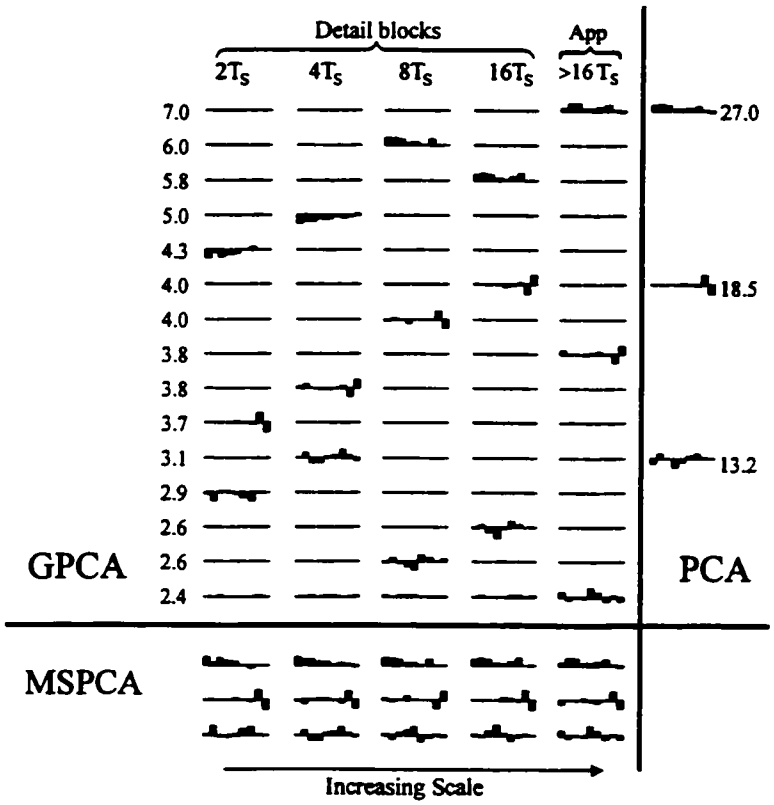


Figure 3.8 Comparison of loadings for PCA, MSPCA and GPCA with db-5 and 4 decomposition levels; data (Base-case with T only control; X=[T C_A F_C T_C T_O C_{AA} F_A C_{AS} F_S], T_S = 10 seconds)

the solvent concentration (C_{AS}). The third one seems to be a combination of several effects. The PCA model reasonably describes the relevant process correlations of the CSTR system.

Since the single-scale data was dyadically decomposed into 5 blocks of 4 details and 1 approximation, MSPCA model consists of PCA models of the 5 scale blocks. From left to right in the bottom of Figure 3.8, each column corresponds to the PCA model of the decomposed block of $2T_S$, $4T_S$, $8T_S$, $16T_S$, and the lower frequencies than $16T_S$. The PCA models at all 5 scales show almost the same behavior as that of the single-scale PCA model. This means that the same process correlations existed over all scales.

GPCA model consists of 15 principal components since there are 5 decomposed blocks and each block has 3 principal components. Each principal component of the GPCA model is expressed with 45 variables. It results from the fact that there are 5 decomposed blocks and each block is described with 9 process variables. However, each of the principal components of GPCA model are actually described with 9 variables at only one scale due to orthogonality among scale blocks. In the GPCA model of Figure 3.8, it is shown the 1^{st} principal component is a function of variables at 4^{th} approximation block and describes a positive correlation among the cooling water flow, the reactor outlet concentration, and the reactor outlet temperature. This correlation is based on the control action at the bandwidth beyond 16 T_S. The effect of the feedback control is most strongly shown at the frequency band lower than 16 T_S. The second principal component indicates a similar correlation among the process variables at the frequency band of 8 T_S.

The process correlations shown at different scales with the GPCA model are almost the same as those shown by the regular PCA model. The first 5 principal components of the GPCA model are very similar to the first principal component of the regular PCA model. The 6^{th} -10^{th} principal components of the GPCA model in Figure 3.8 appear to the second principal component of the regular PCA model. The GPCA model thus consists of the decomposed principal components of the regular PCA model. When the two models convey the same information, the multiscale analysis may not be necessarily required. With the GPCA model, one can judge if a multiscale analysis is

needed for process modeling.

3.4.2 Disturbance structure change

In this example, it is illustrated how the GPCA model can capture disturbances that occur in a narrow frequency band. The simulation conditions are the same as for the previous case study with the exception that the cooling water temperature (T_C) is assumed to have a sinusoidal variation and the inlet temperature (T_O) a positive correlation to T_C. That is,

$$T_{C,t} = T_{C,init} + \sqrt{\left(\frac{\sigma_{T_C}^2}{1-\phi_{T_C}^2}\right)}\sin\left(\frac{2\pi}{80}t\right),$$

$$T_{O,t} = T_{O,init} + \phi_{T_O}\left(T_{O,t-1} - T_{O,init}\right) + \sigma_{T_O}\left((1-\gamma)e_{T_O} + \gamma\sin\left(\frac{2\pi}{80}t\right)\right),$$

where $\gamma = 0.9$, $e_{T_O} \sim N(0,1)$. This frequency corresponds to a period of $8T_S$ where T_S is the sampling time. Since the changes of T_C and T_O affect on the reactor outlet temperature via feedback, the controller of the reactor outlet temperature is turned off in this example. In Figure 3.9(a), PSD of all measurements are shown. PSDs of T, T_O and T_C show the intensive signal contents at 0.0125 Hz, which is the boundary between the second and third detail blocks. PSD of C_A also shows a quite strong effect at the same frequency.

Using a Daubechies-5 wavelet, 4 levels of decomposition have been done dyadically. Figure 3.9(b) shows the GPCA and the PCA models estimated with the decomposed data. The first principal component of the PCA model only shows the correlation among T, T_O, and T_C. It is because there is no feedback control on the reactor outlet temperature and the process dynamics among those correlated measurements. The 2nd principal component of the PCA model shows the same patterns as that of Figure 3.8.

The first principal component of the GPCA model also shows a positive correlation among T, C_A, T_O, and T_C in the third detail block corresponding to the $8T_S$ scale as

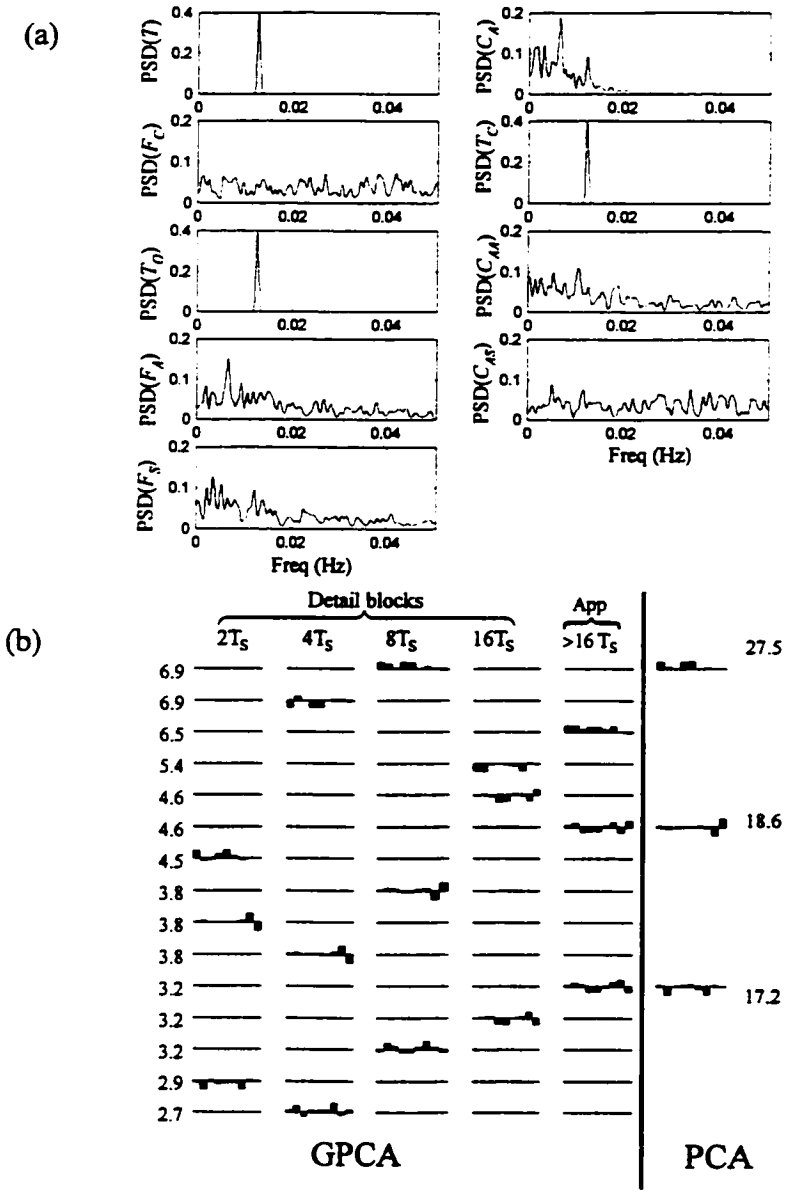


Figure 3.9 GPCA/PCA models with positive correlation between T_C and T_O, and no feedback control on reactor outlet temperature (db5; L=4; T_S = 10(sec); Frequencies on sinusoidal variations on T_C and T_o: 1/80sec, **X**=[T C_A F_C T_C T_O C_{AA} F_A C_{AS} F_S]); (a) PSD of base case simulation; (b) GPCA/PCA models

expected. The positive correlation resulted from both the process dynamics and the correlation among T, T_O, and T_C. Since the increases in T_C and T_O cause the increase of the reactor outlet temperature, it is shown as confounded. This variable correlation is

shown to be consistent with the fact that the oscillation frequencies of T_O and T_C are 80 seconds. In addition, the first principal component also indicates that the positive correlation includes the effects of C_A and F_A. It is because the reactor outlet temperature and the concentration increase together when F_A increases. These two correlations are shown in the first principal component of the GPCA model.

The second principal component of the GPCA model also shows the similar correlation except for the negative correlation between T and C_A. It is because the oscillation frequency is at the boundary between the second and third detail blocks. One can infer the first principal component of PCA model resulted from the first and second ones of the GPCA model. The GPCA model compared to the PCA provides additional information on the process variation.

3.4.3 Detrending long-term effect

It is examined how one can use GPCA model to remove a long-term disturbance effect. The concept is based on the modeling property of the GPCA in section 3.3. The simulation conditions are the same as for the base-case with the exception that the reactant concentration in the feed (C_{AA}) is assumed to have a sinusoidal variation of 2 hour period as;

$$C_{AA,t} = C_{AA,init} + \phi_{C_{AA}}\left(C_{AA,t-1} - C_{AA,init}\right) + \sigma_{C_{AA}} \sin\left(\frac{2\pi}{7200}t\right).$$

Due to the effect of feedback control on the reactor outlet temperature and the process dynamics, the process trends of C_A and F_C will show long-term sinusoidal variations the same as that of C_{AA} while the reactor outlet temperature is maintained at its setpoint value. This long-term variation deteriorates the sensitivity on fault detection. By removing this, one may enhance the detection sensitivity.

Using a Daubechies-5 wavelet, 3 levels of decomposition have been done dyadically. Figure 3.10(a) shows the GPCA estimated with the decomposed data. The

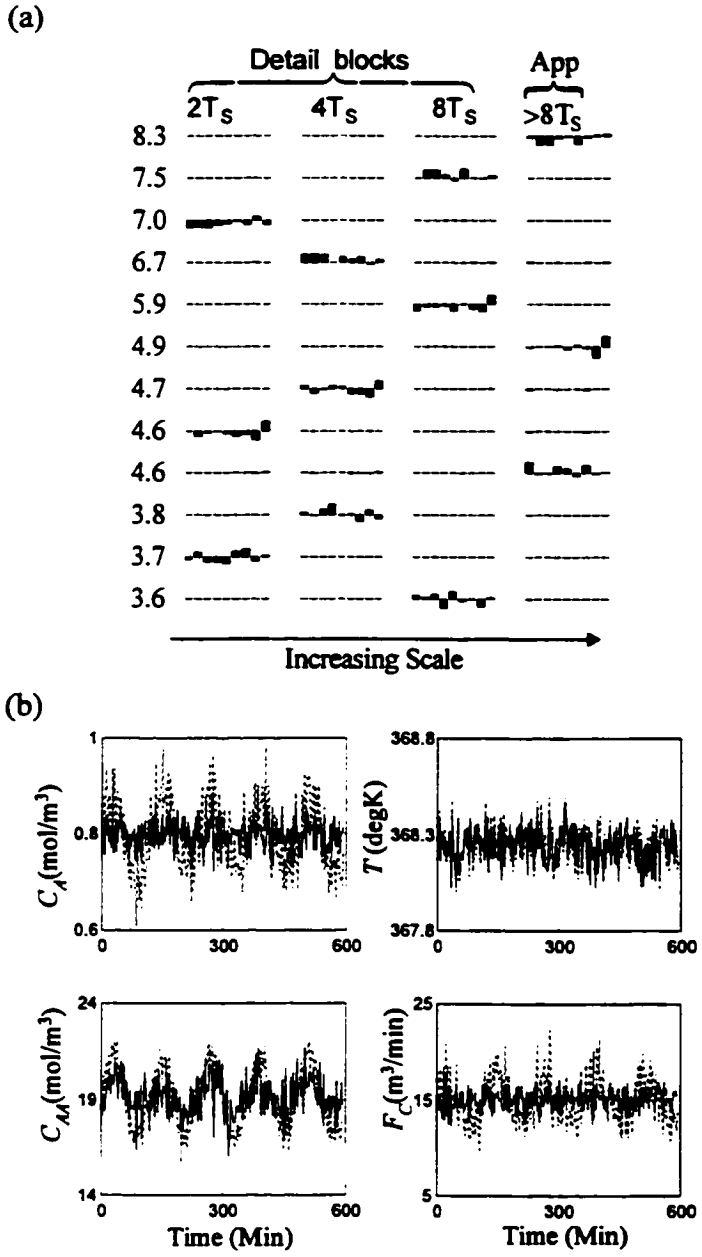


Figure 3.10 GPCA models of case with sinusoidal variation of 2 hour period on C_{AA} (db5; L=3; T_S = 2 minutes); (a) GPCA model; (b) Original vs. filtered process trends with GPCA model (Original:--; Filtered: —)

multiscale blocks from left to right correspond to the details at periods of 4, 8, and 16 minutes, and the approximation at 16 minutes since sampling time is 2 minutes. The first

principal component shows a positive correlation among C_A, F_C, and C_{AA}. This correlation results from the sinusoidal change on C_{AA}, and the feedback control on T. The second principal component shown at the third detail block conveys almost the same as the first principal component. It directly resulted from the sinusoidal variation on C_{AA}. The third principal component represents the positive correlation among T, C_A, F_C, T_C, and T_O. The fourth principal component also represents the similar process variation to the third principal component caused by the feedback control, but includes the effect of the reactant flow rate. Thus the first and second principal components are mainly responsible for the effect of long term variation. The detrending can be done by removing the process variation caused by these two principal components. Since the original matrix can be decomposed into three terms;

$$X = t_1 p_1^T + t_2 p_2^T + \tilde{X}.$$

PCA modeling for process monitoring is done not on X, but on $\tilde{X}$. In Figure 3.10(c), the detrended process trends are shown. All the variations related with the sinusoidal change are removed and thus one can obtain the correlation model whose sensitivity to fault detection would be enhanced. On the other hand, one can use a bandpass filter for detrending purpose. Removing all the principal components of the third detail and approximation blocks will accomplish this. However, it is likely that one would remove process variations which are needed for process monitoring. However, one can more clearly identify the process correlation to be removed by using the GPCA model, and detrend process observations by removing the corresponding principal components. The same concept can be used to remove an effect of significant process condition change such as feed/product flows, feed compositions, and so on.

3.5 Conclusion

When process signals represent the cumulative effects of many underlying process

phenomena and each of them manifests on a different scale, one can explicitly decompose those process effects over scale levels, and capture the process correlation among variables and scales. Measurement signals are decomposed by wavelet transformation. In this study, GPCA that is PCA of the multiscale data has been proposed as a modeling method for data showing multiscale features. Its potential features have been presented and examined through case studies. The modeling properties have been demonstrated through the comparisons to the PCA and the MSPCA with data generated by the CSTR simulation model.

In this study, the discrete wavelet transformation with an orthonormal mother wavelet for finite signal was used. By using orthonormal wavelets, the orthogonality among scale blocks is retained. The scale block orthonormality gives good model interpretation such that the principal components of GPCA become a function of the variables only in one scale block. Thus, one can explicitly decompose the effect of process correlations over different scale levels. It also enables one to remove specific frequency effects of process events from further analysis, and cluster several scale blocks showing similar correlation to efficiently summarize the multiscale process correlations. One can also confirm if the multiscale analysis is needed.

The proposed scheme unifies the existing methods since PCA and MSPCA are limiting cases of GPCA. The principal components of the GPCA decompose those of the PCA model. The GPCA model gives more information than the MSPCA since it orders all the principal components of the MSPCA according to their significance. The GPCA model thus not only conveys all the information contents shown in both MSPCA and PCA, but also reveals the relative significance of all the principal components over scales.

However, it is noted that multiscale analysis may be better suited for extracting information from measured data when each process variation has its dominant frequency content at different scale from the others. It is not always guaranteed that the multiscale analysis will give better results than the single scale analysis. For example, when all signals have their dominant frequency contents at similar frequency band, there would be little difference between the single scale analysis and multiscale analysis.

As a primary application of GPCA, fault detection and isolation are considered in the following chapter. The existing FDI method based on PCA will extend to the multiscale case based on GPCA model.

4. GPCA for process monitoring and fault diagnosis

4.1 Introduction

Underlying process phenomena such as measurement noise, disturbances, and process degradation are cumulatively reflected in process measurements. The process measurements, unless they are slowly sampled relative to the process time constants, are auto-correlated due to the effects of process dynamics and disturbances. A continuous dynamic process is usually characterized with these auto-correlated measurements. Detecting and diagnosing faults that would change the serial correlation as well as the contemporaneous correlation among the process variables in the continuous dynamic processes is problematic.

The common MSPC approach to the continuous dynamic processes has been to use principal component analysis (PCA) or projections to latent structures (PLS) to build a model that captures contemporaneous correlations among the variables and, ignore the serial correlation existing in data during normal operation. In this approach, the resulting monitoring plots then exhibit serial correlation. In spite of many statements to the contrary, this way of process monitoring is still a valid use of PCA modeling for the continuous dynamic processes, as the Shewhart chart is still valid for serially correlated data. Because of the serial correlation, the control limits of T^2 and squared prediction error (SPE) plots in this case, will be larger than those if no serial correlation existed. However, as long as these limits are determined from normal operating data with this serial correlation being present these multivariate statistical process control (MSPC)

schemes will provide valid monitoring schemes.

As a more explicit method to account for the auto-correlated nature of process data in dynamic systems, one can expand the X and Y matrices by including lagged values of all the variables (Wold *et al.*, 1984; MacGregor *et al.* 1991; Wise *et al.*, 1991; Ku *et al.*, 1995). Each row of the measurement matrix is augmented with the previous l observations and PCA or PLS is applied to the augmented data matrix. In this way PCA or PLS will model both contemporaneous and serial correlations among all the variables. The number of principal components required may now be larger, but the MSPC scheme can handle both types of correlations within the model. Then this model will be able to detect any changes in the serial correlation of the variables with ease. This is analogous to the use of time series models to develop SPC charts on individual variables (Montgomery, 1985; Alwan and Roberts, 1988; Harris and Ross, 1991).

An alternative approach to account for the dynamic aspects of the data in MSPC is to use MRA (Multi-Resolution Analysis) via wavelet decomposition (Daubechies, 1990; Mallat, 1989b). In this way, the individual signals are decomposed into different scales (or frequencies). Data in each decomposed scale is used for MSPC. This is an indirect way of handling process dynamics. As an application of this approach to fault detection, Multiscale PCA (MSPCA) has been proposed by Bakshi (1998). MSPCA enables one to simultaneously extract process correlations across data and account for auto-correlation within sensor data. By using the MSPCA, one can take into account process signals having a multiscale nature. It also captures process correlations made by various events occurring at different scales. Aradhye *et al.* (2000) presented the theoretical analysis of MSPCA and addressed several properties from a statistical process control viewpoint. It was shown that the existing statistical process control (SPC) methods would be best suited for situations where the scale of the signal features that represent abnormal operation is known in advance. If the nature of abnormal features cannot be predicted *a priori*, the multiscale approach provides better performance. Misra *et al.* (1999) also proposed a similar scheme. Several PCA models built with approximation and detail blocks of decomposed observations were used for FDI. Teppola

and Minkkinen (2000) proposed wavelet-PLS regression models for data analysis and process monitoring. They combined PLS and multiresolution analysis (MRA). In the wavelet-PLS based approach, periodic seasonal fluctuation and long-term drifting shown as low frequency variations were filtered by removing the low frequency blocks. The wavelet-PLS model was constructed based on the filtered measurements. The wavelet based multiscale approach has also been used in other method fields such as data compression (Misra *et al.*, 2000), sensor validation (Luo *et al.*, 1999) data rectification (Nounou and Bakshi, 1999), as well as process monitoring and diagnosis (Vedam and Venkatasubramanian, 1997; Bakshi, 1998; Kasashima *et al.*, 1995; Xiaoli, 1999)

In the previous chapter, it was shown that one could effectively decompose the process signals by using the generalized principal component analysis (GPCA), and approximately discriminate all the contributing events according to their scale (or frequency) contents. The GPCA model enables one to use additional scale information for fault diagnosis. As a primary application of GPCA, fault detection and isolation (FDI) problems will be investigated in this chapter. On-line and off-line implementation schemes of GPCA for FDI are proposed. Based on the analogy between the multiblock MSPC methods and the multiscale methods, a new algorithm to calculate the GPCA model is also presented. Fault isolation methods based on contribution plots are extended to multi-scale methods. They are shown to greatly facilitate the isolation of faults that occur in different scales. Performance of FDI using GPCA is assessed using a simulated continuous stirred tank reactor (CSTR) system. The FDI performance of the GPCA based method is compared to that of the regular PCA based method via Monte Carlo simulation. The types of faults that can be most effectively treated with the GPCA model are clarified. The detection and isolation of a complex fault is also dealt with whose signature becomes confused with the multiscale nature of the process signals.

4.2 GPCA for process monitoring and fault diagnosis

4.2.1 Off-line GPCA model

Let **X** be the observation matrix of dimension, $n \times m$. Using the transformation matrix that is obtained by the boundary corrected wavelet filters, all the columns of **X** can be decomposed into the details at all levels and the approximation at the coarsest level. And the decomposed block matrices are reconstructed such that all the columns have the same number of rows as in the case of the original observation matrix.

$$\mathbf{W}_S\mathbf{W}_A\mathbf{X} = \mathbf{G}_1^{(1)}\mathbf{H}_1^{(1)}\mathbf{X} + \mathbf{G}_1^{(2)}\mathbf{H}_1^{(2)}\mathbf{X} + \ldots + \mathbf{G}_1^{(j)}\mathbf{H}_1^{(j)}\mathbf{X} + \ldots + \mathbf{G}_1^{(J)}\mathbf{H}_1^{(J)}\mathbf{X} + \mathbf{G}_0^{(J)}\mathbf{H}_0^{(J)}\mathbf{X}, \quad (4.1)$$

where each term represents the scale contribution of the original data matrix in terms of reconstructed values. Once we have chosen an orthogonal mother wavelet, $\mathbf{G}_0 = \mathbf{H}_0^*$, $\mathbf{G}_1 = \mathbf{H}_1^*$, then $\mathbf{G}_0^{(J)} = \mathbf{H}_0^{(J)^{\mathrm{T}}}$, $\mathbf{G}_1^{(j)} = \mathbf{H}_1^{(j)^{\mathrm{T}}}$ and $\mathbf{W}_S\mathbf{W}_A = \mathbf{I}$. All the products of the analysis and the synthesis in (4.1) are simplified in terms of the analysis filter matrices. Thus, (4.1) is expressed as;

$$\begin{aligned}\mathbf{X} &= \mathbf{H}_1^{(1)^{\mathrm{T}}}\mathbf{H}_1^{(1)}\mathbf{X} + \mathbf{H}_1^{(2)^{\mathrm{T}}}\mathbf{H}_1^{(2)}\mathbf{X} + \ldots + \mathbf{H}_1^{(j)^{\mathrm{T}}}\mathbf{H}_1^{(j)}\mathbf{X} + \ldots + \mathbf{H}_1^{(J)^{\mathrm{T}}}\mathbf{H}_1^{(J)}\mathbf{X} + \mathbf{H}_0^{(J)^{\mathrm{T}}}\mathbf{H}_0^{(J)}\mathbf{X} \\ &= \mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_j + \ldots + \mathbf{X}_J + \mathbf{X}_{J+1}\end{aligned}. \quad (4.2)$$

One can characterize the original data in terms of three parameters: variables (m), observations (n), and scales (J). To investigate the variable correlations over frequency, we split all the scale contributions in (4.2) and lay them side-by-side to produce a two-dimensional matrix of size $[n \times (J+1)m]$. The two dimensional matrix is defined as follows;

$$\mathbf{X}_G = \left[\mathbf{X}_1 \quad \mathbf{X}_2 \quad \cdots \quad \mathbf{X}_j \quad \cdots \quad \mathbf{X}_J \quad \mathbf{X}_{J+1}\right], \quad (4.3)$$

where $\mathbf{X}_G$ has the dimension of $[n \times (J+1)m]$, n is the number of observations, m the number of total variables and J the number of scales. One can express $\mathbf{X}_G$ in terms of **V**

and $Diag(\mathbf{X})_{J+1}$ as follows;

$$\mathbf{X}_G = \mathbf{V}Diag(\mathbf{X})_{J+1}, \tag{4.4}$$

where $Diag(\mathbf{X})_{J+1}$ is a diagonal matrix whose diagonal component is $\mathbf{X}$ and off-diagonal terms are zero matrices with the same dimension as $\mathbf{X}$. The dimension of $Diag(\mathbf{X})_{J+1}$ becomes $[(J+1)n, (J+1)m]$. $\mathbf{V}$ is defined as follows;

$$\mathbf{V} = \left[\mathbf{H}_1^{(1)\mathrm{T}}\mathbf{H}_1^{(1)} \quad \mathbf{H}_1^{(2)\mathrm{T}}\mathbf{H}_1^{(2)} \quad \cdots \quad \mathbf{H}_1^{(j)\mathrm{T}}\mathbf{H}_1^{(j)} \quad \cdots \quad \mathbf{H}_1^{(J)\mathrm{T}}\mathbf{H}_1^{(J)} \quad \mathbf{H}_0^{(J)\mathrm{T}}\mathbf{H}_0^{(J)} \right] \tag{4.5}$$

The data matrix is then mean-centered and scaled with the variances. An off-line GPCA model is obtained by applying the regular PCA on $\mathbf{X}_G$. One will capture the variable correlations over several scales. For further details, refer to the chapter 3. If a testing dataset is available, one can project it to the off-line GPCA model and perform the process analysis. The procedure outlined here can be used for the off-line analysis with the assumption that both the training and testing data sets are available.

4.2.2 On-line GPCA model

A crucial part of the on-line GPCA model is the on-line decomposition of the observations. The off-line decomposition procedure thus needs to be modified. Consider a sequence of variable observations, $x[n]$ expressed as;

$$x[n] = \sum_{j=1}^{J}\sum_{k\in Z} X^{(j)}[2k+1]g_1^{(j)}[n-2^j k] + \sum_{k\in Z} X^{(J)}[2k]g_0^{(J)}[n-2^J k], \tag{4.6}$$

where

$$X^{(j)}[2k+1] = \left\langle h_1^{(j)}[2^j k - l], x[l] \right\rangle, \quad j = 1,\cdots,J,$$

$$X^{(J)}[2k] = \left\langle h_0^{(J)}[2^J k - l], x[l] \right\rangle.$$

These are the convolutions of the input with $h_0[n]$ and $h_1[n]$ evaluated at evenly indexed $2^j k$. In these equations, $h_i^{(1)}[n] = h_i[n]$ and $g_i^{(1)}[n] = g_i[n]$. If $h_i[n]$ is an orthogonal filter, $h_i[-n] = g_i[n]$ and $h_i^{(J)}[n] = g_i^{(J)}[-n]$. This structure implements an orthogonal discrete time wavelet series expansion. The basis functions $g_i^{(J)}[n]$ are defined as the time domain versions of $G_0^{(J)}(z)$ and $G_1^{(J)}(z)$ given by

$$G_0^{(J)}(z) = G_0^{(J-1)}(z) G_0(z^{2^{J-1}}) = \prod_{K=0}^{J-1} G_0(z^{2^K}), \tag{4.7}$$

$$G_1^{(J)}(z) = G_0^{(J-1)}(z) G_1(z^{2^{J-1}}) = G_1(z^{2^{J-1}}) \prod_{K=0}^{J-2} G_0(z^{2^K}), \tag{4.8}$$

where $G_0^{(1)}(z) = G_0(z)$, $G_1^{(1)}(z) = G_1(z)$, and j=1,..., J. The expressions for analysis filters, $H_0^{(J)}(z)$ and $H_1^{(J)}(z)$ for j=1,..., J. can be similarly obtained. Using these expressions and the multirate identity for filtering, the octave-band filter banks that perform the multiresolution analysis can be expressed in terms of J+1 filter banks that are the combination of analysis and synthesis filters, downsamplers, and upsamplers. The filter banks at coarser scales need more samples. This introduces time delay in calculating the wavelet transformation. The time delay is due to the dyadic down sampling used in wavelet transformation.

Since wavelet filters are in general non-causal in nature and require future measured data for calculating the current wavelet coefficient, special wavelets at the edges that eliminate boundary errors while being orthogonal to the other wavelets are required. These boundary corrected filters are causal and require no information about the future observations to compute wavelet coefficients at the signal ends (Cohen *et al.*, 1993; Nounou and Bakshi, 1999).

Thus, the on-line implementation scheme requires that one decomposes the

measured data within a window of dyadic length using a causal boundary corrected wavelet filter and retains only the last data point of the reconstructed signal for on-line use. When new measured data are available, one moves the window in time to include the most recent measurement while maintaining the maximum dynamic window length. Nounou and Bakshi (1999) used a similar scheme for on-line data rectification. It is noted that the data transformation is done on a row as the new measured observation is obtained. Thus on-line transformation does not guarantee the scale block orthogonality. However, one can obtain the block orthogonal PCA model by applying multiblock principal component analysis (MBPCA) (Cheng and McAvoy, 1997) or consensus principal component analysis (CPCA) (Wold *et al.*, 1987) algorithm on the data transformed by the on-line scheme. The algorithm is summarized in 4.2.3.

4.2.3 GPCA algorithm

a. Perform regular PCA on each scale block $\mathbf{X}_j$ to obtain loadings, $\mathbf{P}_j$ and eigenvalues, λ_j

b. Select the largest eigenvalue among all the existing values and the corresponding loading.

c. Obtain a GPCA loading $\mathbf{P}_{G,k}$ with the selected loading of a scale block, j. For example, if the selected eigenvalue is i-th eigenvalue of j-th scale block, then the corresponding GPCA loading $\mathbf{P}_{G,k}$ is;

$$\mathbf{P}_{G,k}^T = \left[\underbrace{\mathbf{0}(1,m) \quad \cdots \quad \mathbf{0}(1,m)}_{(j-1)m} \quad \mathbf{P}_{ij}^T \quad \underbrace{\mathbf{0}(1,m) \quad \cdots \quad \mathbf{0}(1,m)}_{(J+1-j)m} \right].$$

d. The scores can be obtained in two ways. One can use the block score according to the loading selected, or use $\mathbf{X}_G$ and $\mathbf{P}_{G,k}$, the GPCA loading

$$\mathbf{T}_{G,k} = \mathbf{X}_G \mathbf{P}_{G,k}.$$

e. Deflate residuals

$$\mathbf{E}_G = \mathbf{X}_G - \mathbf{T}_{G,k}\mathbf{P}_{G,k}^T$$

g. Return to step b and repeat the procedure until one obtains all the GPCA loadings with all block loadings. In step b, the eigenvalue is selected among the remaining eigenvalues except for the one used in the previous step.

4.2.4 Process monitoring and fault diagnosis

With the notion of GPCA, one can estimate a scale contribution to Hotelling's T^2 for fault diagnosis;

$$T^2 = \sum_{i=1}^{A} \frac{t_i^2}{s_{t_i}^2} \approx \sum_{i=1}^{A} \frac{1}{s_{t_i}^2} \left(\sum_{k=1}^{m(J+1)} (p_{ki} x_{G,k})^2 \right) = \sum_{k=1}^{m(J+1)} x_{G,k}^2 \left(\sum_{i=1}^{A} \left(\frac{p_{ki}}{s_{t_i}} \right)^2 \right) = \sum_{k=1}^{m(J+1)} x_{G,k}^2 w_k , \qquad (4.9)$$

where p_{ki} is k-th variable weight in i-th latent variable, m is the number of variables, and J is the number of scales. The overall T^2 is partitioned into a contribution from each scale as follows;

$$T_j^2 = \sum_{i=1}^{m} x_{G,(j-1)m+i}^2 w_{(j-1)m+i} \quad j = 1, \ldots, J+1 . \qquad (4.10)$$

This is the expression for the scale contribution of T^2 to a fault. Once the contributing scale for a fault is identified, one can consult with the variable contributions of the contributing scale. It can be expressed as follows;

$$T_{jk}^2 = x_{G,(j-1)m+k}^2 w_{(j-1)m+k} . \qquad (4.11)$$

The above equation is the scale variable contribution to T^2 relative to zero. The expression for contribution to a change in T^2 over some time interval is;

$$\Delta T_{jk}^2 = T_{jk}^2(t) - T_{jk}^2(t-1) = \Delta x_{G,(j-1)m+k}^2 w_{(j-1)m+k}, \tag{4.12}$$

where j=1, ..., J+1, k=1, ..., m. The upper control limits for interval for T^2, T_j^2, and T_{jk}^2 for a given confidence 1-α are given by:

$$T_{UCL}^2 = \frac{(n-1)(n+1)q}{n(n-q)} F_\alpha(q, n-q), \tag{4.13}$$

where $F_\alpha(q, n\text{-}q)$ is the upper 100α % critical point of the F-distribution with q and n-q degrees of freedom (Tracy *et al.*, 1992; MacGregor and Kourti, 1995; Nomikos and MacGregor, 1995). In case of T^2, T_j^2, and T_{jk}^2, the degrees of freedom of the numerator terms (q) are $A(J+1)$, A, and 1, respectively. One needs to carefully determine the significance level, α of T_j^2 because the monitoring task using the multiple scale T_j^2 plots becomes a multiple statistical test and then the actual type I error will be $1-(1-\alpha)^{J+1}$. It has an effect of lowering the upper critical point and increasing the sensitivity of the fault detection. Thus, the α value of the test using multiple T_j^2 has to be modified such that the overall type I error would be the same as the given type I error. Otherwise, comparison results between the regular PCA/PLS and the multiway, or multiscale methods may lead to an incorrect assessment of the sensitivity of the fault detection (Misra *et al.*, 1999; Qin *et al.*, 1999). Further detail of the Bonferroni adjustment is given in the next section.

Similar expressions can be obtained for the SPE that is;

$$\text{SPE} = \sum_{i=1}^{m(J+1)} \left(x_{G,i} - \hat{x}_{G,i}\right)^2. \tag{4.14}$$

From this expression, the scale SPE that is the scale contribution to the overall SPE is

given by,

$$\mathrm{SPE}_j = \sum_{i=1}^{m}\left(x_{G,(j-1)m+i} - \hat{x}_{G,(j-1)m+i}\right)^2 \quad , j = 1,\cdots,J+1, \tag{4.15}$$

where $\hat{x}_{G,k} = (\mathbf{P}_A\mathbf{P}_A^T)_k x_{G,k}$ and the j variable contribution to the SPE at that scale SPE_{jk} is

$$\mathrm{SPE}_{jk} = \left(x_{G,(j-1)m+k} - \hat{x}_{G,(j-1)m+k}\right)^2 \quad j = 1,\cdots,J+1,\ k = 1,\cdots,m. \tag{4.16}$$

The upper control limits for SPE, SPE_j, and SPE_{jk} for a given confidence $1-\alpha$ can be calculated by a weighted chi-squared distribution (Box, 1954) as $\delta_{\mathrm{SPE},\alpha} \equiv g\chi_\alpha^2(h)$, where the weight ($g$) and the degree of freedom (h) are both functions of the eigenvalues of the covariance matrix. The mean and variance of the $g\chi_\alpha^2(h)$ distribution ($\mu = gh$, $\sigma^2 = 2g^2h$) are equated to the sample mean (m) and variance (v) of the SPE sample (Nomikos and MacGregor, 1995). This is a quick and reliable way to estimate g and h provided that the number of SPE observations is sufficiently large. Thus the upper control limit on the SPE at the given significance level α is given by

$$\delta_{\mathrm{SPE},\alpha} \equiv \left(\frac{v}{2m}\right)\chi_\alpha^2\left(\frac{2m^2}{v}\right). \tag{4.17}$$

This is the upper $100\alpha\%$ critical point of the chi-squared variable weighted with $v/2m$ and with $2m^2/v$ degrees of freedom at the significance level, α (Tracy *et al.*, 1992; MacGregor and Kourti, 1995; Nomikos and MacGregor, 1995). Based on this expression, the upper control limit for the scale SPE and the variable contribution to the scale SPE can be obtained. On the other hand, the upper control limit can also be computed from historical data using approximate results for the distribution of quadratic forms (Jackson

and Mudholkar, 1979; Nomikos and MacGregor, 1995, MacGregor and Kourti, 1995). Jackson and Mudholkar's approximate distribution is very close to the one given by Box. Nomikos and MacGregor proved that both methods give the same result when one uses the Wilson-Hilferty equation (Nomikos and MacGregor, 1995). In case of the SPE statistics, one also has to modify the significance level, α according to the Bonferroni adjustment due to the same reason explained above.

The process monitoring and fault diagnosis procedure based on the GPCA can be summarized as follows;

a. Obtain GPCA model with the training data set and calculate the control limits for T^2, T^2_j, SPE, and SPE_j based on the training data set.
b. For a new sample, calculate, T^2, T^2_j, SPE, and SPE_j
c. If either T^2 or SPE, or both of them are out of their limits, a fault is detected.
d. The scale j whose T^2_j or SPE_j violates its control limit is deemed to be related for the fault.
e. Any of the variable contributions T^2_{jk} and SPE_{jk} for $k = 1, 2, \ldots, m$, that are large compared to the others in the responsible scale blocks are then diagnosed to be related to the fault.

It is noted that the steps c and d would be done together assuming that the control limits for the T^2, T^2_j, SPE, and SPE_j are adjusted by Bonferroni corrections. Depending upon the fault characteristics, only a few values of T^2_j and SPE_j in the step d, T^2_{jk} and SPE_{jk} in the step e should be affected. In general, one can approximately recognize a fault type with T^2_j/SPE_j plots. When the fault is occurring at a narrow frequency band, one can further investigate a root cause by examining the variable contribution to T^2_j/SPE_j. It is recommended to use the overall GPCA statistics (T^2 and SPE) at first as overall monitoring tools, and the scale GPCA statistics (T^2_j and SPE_j) for further diagnosis. Since GPCA allows one to use information on scale localization in addition to time localization, the fault investigation thus may provide a suitable framework for the fault isolation.

Figure 4.1 summarizes the FDI procedure using the GPCA model. The proposed scheme provides a systematic approach for diagnosing abnormalities in processes. In the following section of the case study, it is illustrated how the procedure can be used for a different type of fault with different frequency characteristics.

4.2.5 Fault detection and isolation assessment

Performance measures

The objective of a fault detection technique is to be robust to data independent of the training set, sensitive to all the possible faults of the process, and prompt to the detection of the faults (Russel *et al.*, 2000). A fault detectability of a method is thus measured with the robustness, sensitivity and promptness of SPE and T^2 statistics with

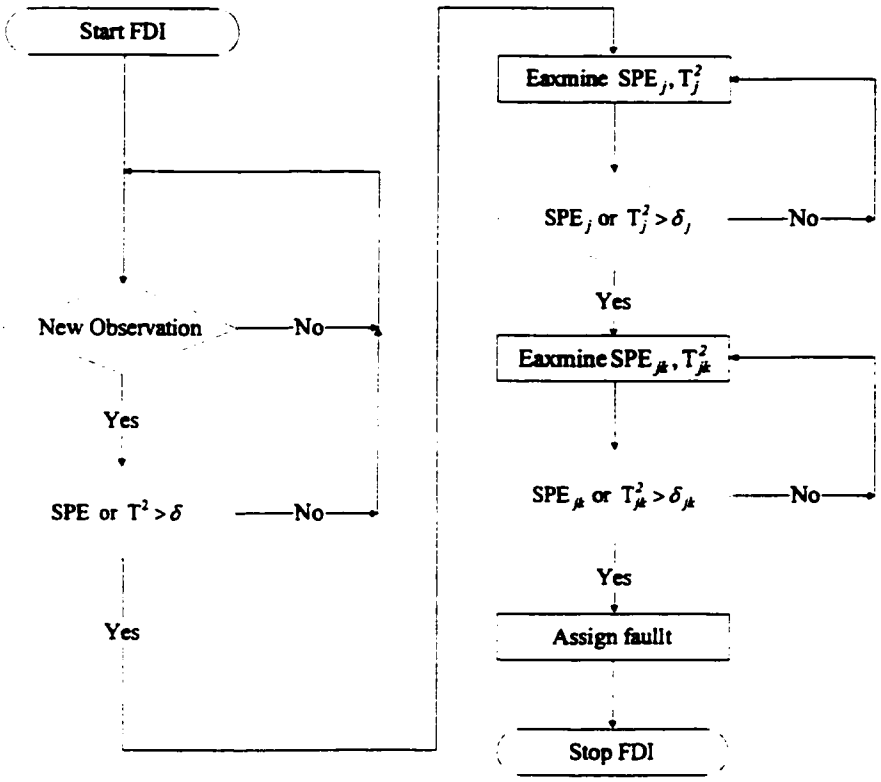


Figure 4.1 Flowchart of fault detection and isolation using GPCA

respect to various types of faults. With following null and alternative hypotheses;

$$H_0 : \text{No fault}, H_1 : \text{fault}, \tag{4.18}$$

type I error is defined as a probability of rejecting the null hypothesis that is true (Mongomery and Runger, 1994), and presented by the α risk;

$$\alpha = p(\text{reject } H_0 \mid H_0 \text{ is true}). \tag{4.19}$$

The robustness of each statistic can be determined by calculating the false alarm rate during normal operating conditions for the testing set and comparing it against the level of significance upon which the threshold is based. By comparing the type I error with data sets to the significance level, one can measure the robustness of each method. The discrepancy between the calculated type I error and the significance level, indicates deviation from the assumed distributions on the statistics of T^2 and SPE.

Type II error is a probability of accepting the null hypothesis that is false, and presented by the β risk;

$$\beta = p(\text{accept } H_0 \mid H_0 \text{ is false}). \tag{4.20}$$

The type II error thus represents the rate of missed alarm. When the chance of making a type I error on an individual test is reduced, the chance of making a type II error is increased. The promptness defined as the ratio of the samples violating the threshold limits to the total samples since a fault occurs, can be quantified by $1-\beta$.

When the type I error is far from the specified significance level, it indicates more deviation from the assumed distributions of the statistics. Thus, one has to adjust the significance level such that the calculated values are consistent with the specified significance level. It would not be fair to directly compare the fault detection methods in

terms of missed detection rates when they have such widely varying false alarm rates (Russel *et al.*, 2000). The adjustment of thresholds provides a fairer basis for the comparison of the sensitivities of the statistics.

The average run length (ARL) is the average number of samples required to detect a shift (Mongomery and Runger, 1994). When the magnitude of shift is zero, the corresponding ARL value is a reciprocal of the probability of false alarms, and is referred to as the in-control run lengths ($ARL_o = 1/\alpha$). It is desirable to have the same in-control run length for calculating the consistent ARL values for non-zero mean shifts. This approach provides a standard way of comparing the relative performance of different SPC techniques. For example, when plotted against the magnitude of the shift, the ARL curve is expected to be non-increasing and typically converges to 1 as the magnitude of shift tends to infinity (Aradhye *et al.*, 2000). By calculating ARL, one can quantify the detection sensitivity. In the conventional PCA, the T^2 and SPE statistics include all the events occurring at all the scale bandwidths. Their sensitivities with respect to abnormality, which may occur at one scale bandwidth, become less sharp. On the other hand, one may enhance the sensitivities of T^2 and SPE statistics by using the multiscale approaches.

We use SPE and T^2 statistics and their scale contributions, SPE_j and T_j^2 with respect to various values of the fault to signal ratio (FSR) for the performance evaluation of the fault detection. Based on those statistics, we calculate the type I and II errors, and ARL. In practice, one can look at other SPC charts such as cumulative sum (CUSUM), Shewart, and exponentially weighted moving average (EWMA). Confidence limits of each case are determined based on several parameters of each chart. Depending on the type of fault, one can determine a suitable chart for different purposes. It is important that one has the same basis when comparing two different methods. In the study, the T^2 and SPE plots are used together without any smoothing effect.

In the MSPCA proposed by Bakshi (1998) the multiple threshold plots based on the wavelet coefficients were used, and the scales whose thresholds were violated, were selected. The selected scales were used for reconstructing the monitoring plots. Misra *et*

al. (1999) used several T^2_j/SPE_j plots for the fault detection. In this case, the upper control limits should be adjusted with Bonferroni correction to give correct α or ARL_0.

Similarly, one can address a fault isolability of a method in terms of the robustness, the sensitivity and promptness. The scale contributions to the SPE and T^2 (SPE_j and T^2_j) and the variable contributions to SPE_j and T^2_j statistics (SPE_{jk} and T^2_{jk}) are the main tools for the fault isolation under the assumption that a fault is detected. However, the measures outlined above will be limited to the simple type of fault due to the fault propagation caused by process dynamics and feedback controls. All the methods are so different in the ways of isolating faults that it would be more appropriate to highlight potential features of each method rather than comparing their performance with respect to the same faults.

Bonferroni adjustment

The significance level (α) is the chance taken to incorrectly declare a fault because of the type I error. In the case of more than one statistical test, the chance of finding at least one test statistically significant due to chance fluctuation increases. Thus, multiple tests increase the chance to incorrectly declare a fault as significant. Since the overall significance level using all tests should result in the same value of the original significance level intended, the significance level of each test should be adjusted downward to consider the case of more than one test in a particular study. For a set of independent tests, the Bonferroni adjustment (Alt *et al.*, 1988) can be used. When α is set at 0.05, the test will show 'something' in one out of twenty statistical tests while in fact there is nothing. With five independent tests, the chance of finding at least one fault significant due to chance fluctuation equals 0.23 ($=1-0.95^5$). In ten tests, this chance becomes 0.40 ($=1-0.95^{10}$), which is about one in two. Using the Bonferroni method, α of each individual test is adjusted downwards to ensure that the overall risk for a number of tests remains 0.05. For example, α in case of five tests will be 0.01 ($=1-0.95^{1/5}$). Then the risk of incorrectly finding a fault continues to be 0.05. Thus, in case of GPCA and

MSPCA with J levels of decomposition, the significance levels of the scale statistics, T_j^2 and SPE_j, are adjusted as $\alpha_j = 1 - (1-\alpha_{given})^{1/(J+1)}$. Then, the risk of incorrectly finding a fault continuous to be 0.05, providing each of the test statistics is statistically independent.

The Type I error of T_j^2/SPE_j is based on the threshold value adjusted with the Bonferroni rule. This adjustment is based on the assumption that one uses the scale monitoring plots, T_j^2/SPE_j as main fault detection tools (Bakshi, 1998; Misra *et al.*, 1999). On the other hand, adjusting the threshold with the Bonferroni rule is not appropriate in the case that the scale monitoring plots are used only for the purpose of fault diagnosis since it is assumed that a fault is already detected.

4.2.6 Analogy between multiblock and multiscale methods

In very large processes involving many processing units with many variables in each unit, the number of potential errors or faults can be very large, making the diagnosis more difficult. Schemes for process monitoring using multivariate statistical projection methods such as PCA and PLS can be extended to situations where the processes can be naturally blocked into several subsections (MacGregor *et al.*, 1994). The multiblock projection methods allow one to establish monitoring charts for the individual process subsections as well as for the entire process. Multiblock data analysis methods have their origins in path analysis and path modeling in the fields of sociology and econometrics. (Wold, 1987; Wagen and Kowalski, 1988). Westerhuis *et al.* (1998) presented a unifying view of all the multiblock PCA and PLS algorithm.

The main advantage of such blocking is to allow for easier interpretation of the data by looking at smaller meaningful blocks and the relationship between blocks. The choice of blocking depends upon engineering judgments. The block should correspond to distinct units of the process where all the variables within a block or process unit may be highly coupled, but where there is minimal coupling among variables in different blocks. Variables associated with streams that leave one block and enter another (feed or recycle streams) should generally be included in both blocks.

If a special event or fault occurs in a certain section of the process, then the multiblock methods can generally detect the event earlier and reveal the subsection within which the event has occurred (MacGregor *et al.*, 1994). This is because a process is monitored by the deviations of the block T^2/SPE statistics relative to normal variations in the variables of the corresponding process subsection, but not with respect to variations in all the process variables. However, if a fault occurs that affects all sections of the plant almost simultaneously, then the superblock (or consensus) scores and SPE should detect the event earlier, because it is simultaneously using information on the fault gathered from all blocks of the process. Note that the superscores level in MBPCA/MBPLS is equivalent to the conventional (single block) PCA/PLS.

The analogy between the multiblock methods and multiscale methods comes from the fact that both methods decompose the overall monitoring statistics and provide detailed frameworks for diagnosing process abnormalities. Table 4.1 summarizes the analogy between the multiscale and multiblock methods. The multiblock methods block the information according to variables while the multiscale methods block the data with respect to frequency characteristics. When a fault occurs in one subsection of the process, or one frequency band of the whole frequency ranges of the process variations, either the multiblock or multiscale method would be better than the regular methods in detecting

Table 4.1 Comparison between multiblock and multiscale approaches

	Multiblock	Multiscale
Decomposition Methods	Process units or Variables CPCA (Wold et al., 1987); MBPCA (Cheng and McAvoy, 1997);	Scales (Frequencies) MSPCA (Bakshi et al., 1998; Misra et al., 1998); GPCA (Yoon and MacGregor, 2001b)
Tools for FDI	T^2/SPE, T_j^2 /SPE$_j$, T_{jk}^2 /SPE$_{jk}$	T^2/SPE, T_j^2 /SPE$_j$, T_{jk}^2 /SPE$_{jk}$
Requirements	Location information of process variations and faults	Frequency characteristics of process variations and faults
Applications	Large processes	Small but Complicated processes

and isolating faults. However, it is questionable if the multiblock and multiscale methods would result in better performance when the fault effect is spread over more than one subsection of the process or one frequency range of the process variations. Thus, a suitable monitoring method to give the best detection and isolation of the faults may be determined on the fault characteristics. Information on the faults is very crucial.

4.3 Numerical examples

4.3.1 Precision degradation with a localized frequency contents

Unstable contact or an old instrument can cause an excessive oscillatory action on the measurement while its real value remains unchanged. Assume that the inlet temperature measurement (T_O) of the CSTR model in appendix A (Marlin, 1995) is showing this type of behavior. The fault is simulated with a trigonometric function whose oscillation period is 25 seconds. The fault (μ) is assumed proportional to the square root of the signal variance. Then,

$$\mu = \sigma(x_{t,meas}) \sin\left(\frac{2\pi t}{25}\right), \tag{4.21}$$

where $\sigma(x_{t,meas})$ is square root of a measurement variance. In the CSTR system, the inlet disturbance is modeled as an autoregressive time series, $x_{t,meas} = x_t + \sigma_m m_t$ where $x_t = \phi x_{t-1} + \sigma_e e_t$, and m_t and $e_t \sim N(0,1)$. Thus the measurement variance is;

$$\sigma^2(x_{t,meas}) = \frac{\sigma_e^2}{1-\phi^2} + \sigma_m^2. \tag{4.22}$$

Using the CSTR model and the base-case condition, a set of observations was collected for 200 minutes under normal operation. Sampling time was 10 seconds. Testing data was generated with the fault introduced at 51 minutes into the run as shown

Table 4.2 Type I errors and upper control limits

	α	T^2_{UCL}	SPE_{UCL}
T^2/SPE (PCA)	0.01	11.3	12.1
T^2/SPE (GPCA)	0.01	32.6	35.4
T_1^2/SPE_1	0.002	18.1	16.4
T_2^2/SPE_2	0.002	18.6	19.1
T_3^2/SPE_3	0.002	20.3	18.3
T_4^2/SPE_4	0.002	16.1	15.4
T_5^2/SPE_5	0.002	18.4	16.3

in Figure 4.2(a). PCA and GPCA models were estimated with the training data. The scaling for the GPCA model was done on X_G. Table 4.2 summarizes the significance levels of the T^2/SPE statistics and the corresponding control limits. The significance levels of T_j^2/SPE_j statistics with the Bonferroni adjustments show the lowest values. The adjusted significance level is 0.0020 ($=1-(1-\alpha)^{1/5}$). The upper control limits of GPCA T^2/SPE statistics are higher than those of the PCA statistics.

While the monitoring plots of the regular PCA in Figure 4.2(b) do not clearly indicate the fault at 51 minutes, the GPCA monitoring plots in Figure 4.2(c) detects the fault more clearly. One can also detect the fault with the scale monitoring plots using T_j^2/SPE_j. In Figure 4.2(d), it is shown that the low scale block (D_1) is responsible for the fault. To diagnose a root cause of the fault, the variable contributions to SPE_{D1} were examined. Figure 4.2(e) clearly indicates that the inlet temperature measure is a plausible root cause.

The above analysis is based on a case of a single simulation. One cannot generalize the analysis results due to many uncertainties involved. To fairly assess the performance of the proposed method, the analysis was done using Monte Carlo simulation (Sobol, 1974). 500 sets of training and testing data were generated at each value of various fault to signal ratio (FSR) on the inlet temperature measurement. With each data, the ARL values for the fault detection and isolation were calculated. The Type I error of T_j^2/SPE_j is based on the threshold value adjusted with the Bonferroni rule. With

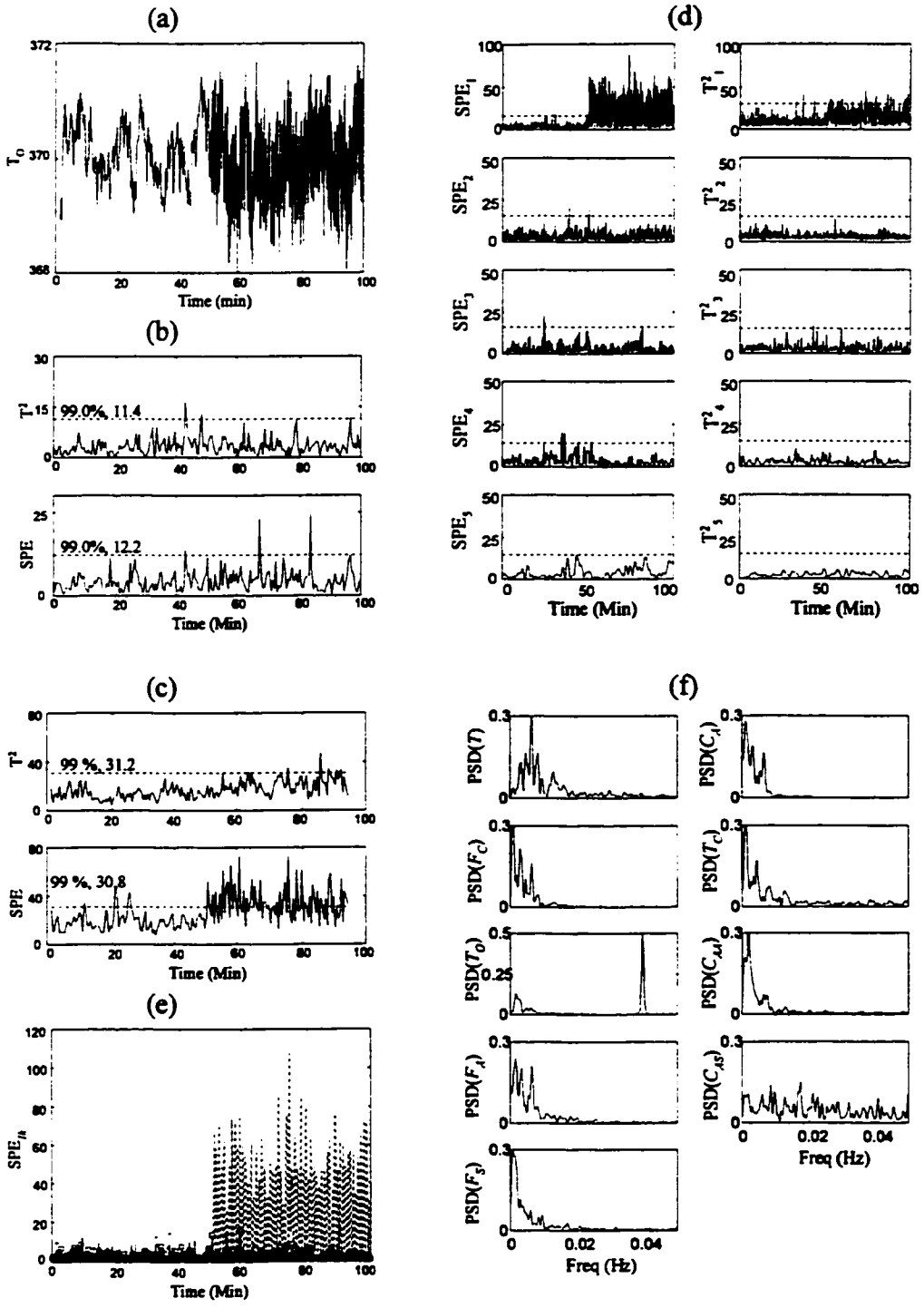


Figure 4.2 FDI of T_O sensor precision degradation; (a) Process trend; (b) T^2/SPE (PCA); (c) T^2/SPE (GPCA); (d) T_j^2/SPE_j; (e) SPE_{1k}(T:- ; F_C:□; T_O:--; F_A:*); (f) Frequency characteristics (FSR = 2 at 51 min)

these monitoring statistics, the overall ARL based on the T^2 and SPE plots is calculated. The individual control limits for the T^2 and SPE plots are then slightly adjusted such that the overall control limit is 0.01 in all cases. Since 0.01 value of the type I error statistically means that one has false alarm once every 100 samples, the ARL is obtained by taking the average values of 500 runs such that one can minimize the uncertainty of the type I error.

Figure 4.3(a) shows the ARL values of several statistics for fault detection of the precision degraded T_O sensor. Those of the GPCA based on T^2/SPE, and the scale contribution based on T_j^2/SPE_j are shorter than those of the regular PCA statistics. This is due to the characteristic of the sensor precision degradation whose effect is concentrated in a single frequency range. But the result may not be the same for other types of faults since their frequency characteristics are different. The isolation ARL is also calculated with the testing data under the assumption that the fault is detected. In the case of GPCA, the variable contributions to T_j^2/SPE_j statistics, T_{jk}^2/SPE_{jk} are used while the variable contributions T_k^2/SPE_k are used for PCA. From Figure 4.3(b), it is obvious that one can diagnose the sensor precision degradation better with the statistics of the GPCA than with those of PCA over all ranges of fault magnitudes. The GPCA using the scale contribution

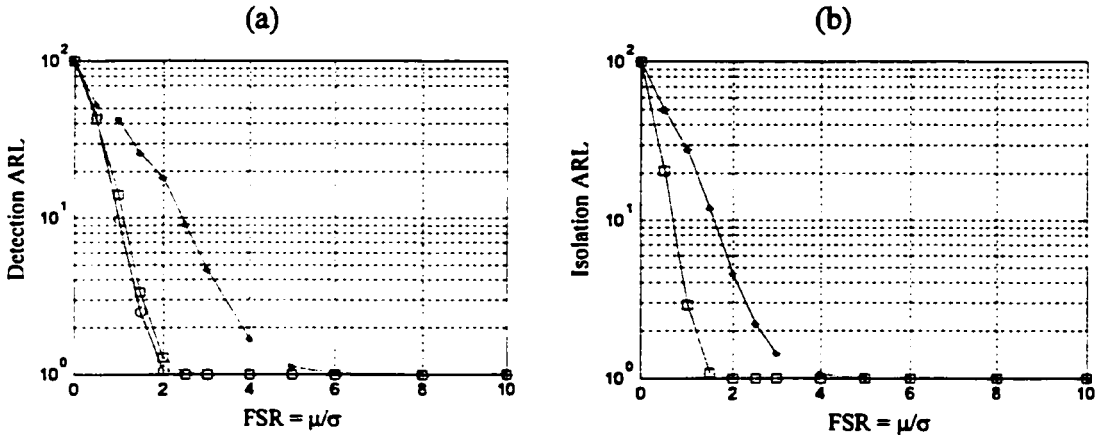


Figure 4.3 Monte Carlo simulation of T_O sensor precision degradation; (a) Detection ARL; (b) Isolation ARL (# of simulations at one FSR: 500; α=0.01; Scaling on X_G; ∗: T^2/SPE(PCA); □:T^2/SPE(GPCA); O: T_j^2/SPE_j)

outperforms the regular PCA in diagnosing the sensor precision degradation.

In Figure 4.2(f), the frequency characteristics of all the process variables are shown with the testing data. Note that the power spectral density (PSD) of the inlet temperature measurement, T_O was restricted to 0.04 Hz and not propagated into the other variables. PSD(T_O) plot clearly indicates that the precision degradation appears at the high frequency-low scale band and its effect is not spread over the other frequencies. Thus, the narrow frequency distribution of the fault may explain why the GPCA based method outperforms the regular PCA based method. On the other hand, this enhancement may result from the scaling effect. Unit scaling of each scale block would put equal weights on all the process correlations of the decomposed scales regardless of the scale contributions to the overall process variation. If the scale contribution of the high frequency block is small, the unit scaling might unreasonably increase the small contribution. The issue on the scaling will be treated in the example 6.

4.3.2 Sensor drift (Incipient fault)

The fault effect of the sensor drift is also confined in one scale band but at low frequency – high scale bandwidth. A seasonal variation or long-term aging of reactor catalyst may show this type of characteristics. Figure 4.4(a) shows a drift on the inlet temperature measurement. The drifting sensor is simulated to start at 51 minute and continue until the end of the testing simulation. The fault magnitude is $4\sigma_{T_O}^2$ where $\sigma_{T_O}^2$ is the total variance of the inlet temperature measurement. The frequency characteristic of the sensor drift is shown in Figure 4.4(b). The fault effect becomes dominant as approaching to the low frequency end.

Monte Carlo simulations were done for the sensor drift. All the simulation conditions were the same as those for the precision degradation except for the type of fault. 500 simulations were made for each magnitude of the drift fault. Figure 4.4(c) and (d) summarize the detection and isolation ARL for the PCA, GPCA, and the scale contribution with respect to various magnitudes of the fault to signal ratio. Both plots show that the detection and isolation performance of the three methods is almost the same.

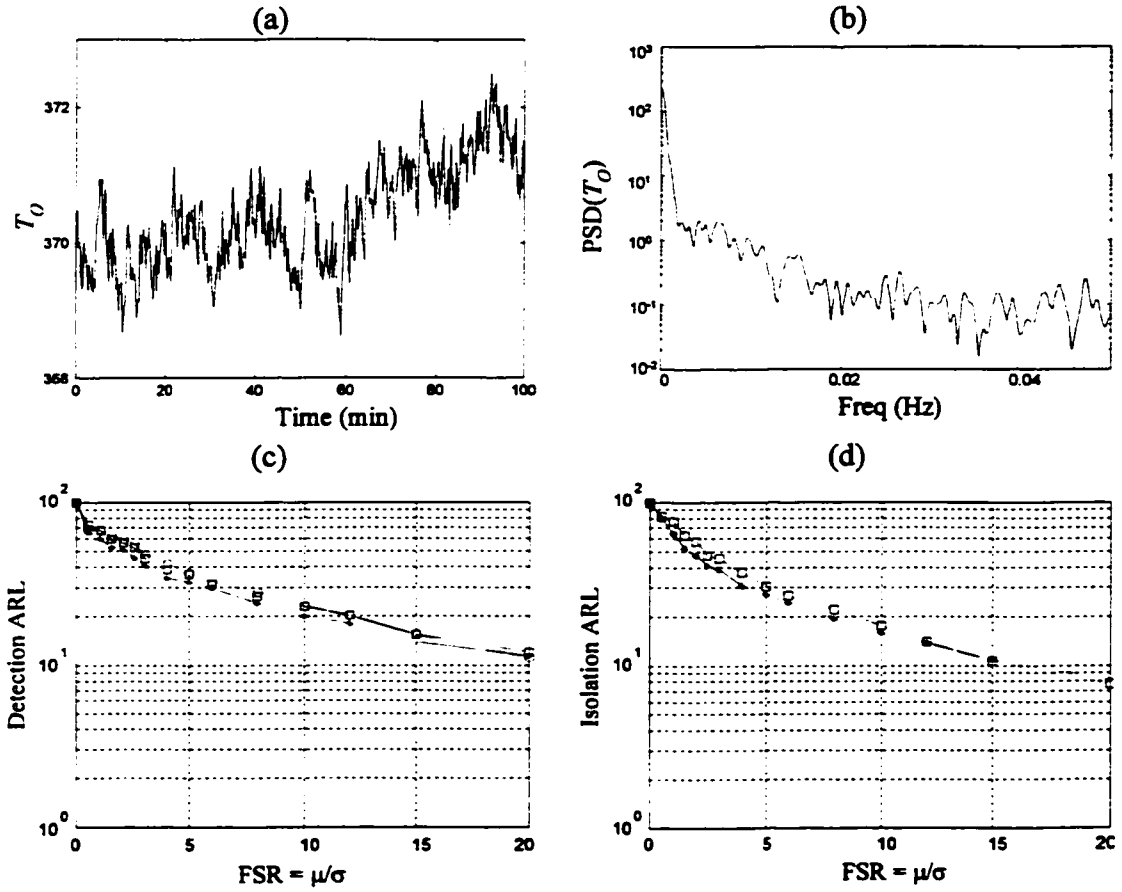


Figure 4.4 FDI of T_O sensor drift; (a) Process trend with $4\sigma_{To}^2$ (σ_{To}^2 =0.5026); (b) Frequency characteristics of T_O sensor; (c) Detection ARL based on Monte Carlo simulation; (d) Isolation ARL based on Monte Carlo simulation (# of simulations at one FSR: 500; α=0.01; Scaling on X_G; ∗: T^2/SPE(PCA); □:T^2/SPE(GPCA); O: T_j^2/SPE_j)

It results from the fact that the fault scale effect is exponentially distributed over the wide frequency ranges even though the effect of the sensor drift shows a strong intensity at low frequency. The FDI using the GPCA and its scale statistics does not provide an enhanced performance if the fault effect is confined in low frequency region

4.3.3 Sensor bias

A sensor bias of the inlet temperature measurement (T_O) is simulated as in Figure

4.5(a). All the other simulation conditions are the same as the above case. Figure 4.5(a) shows the process trend of the inlet temperature. Even though this type of fault is a high frequency event, the high frequency feature is easily missed and instead other effects such as measurement noise and measurement dynamic characterize the fault signature. Figure 4.5(b) confirms this fact.

500 sets of training and testing data were generated with different magnitudes of the bias on the inlet temperature measurement. With each data, ARL values were calculated with respect to different FSR. Figure 4.5(c) and 4.5(d) show the fault detection

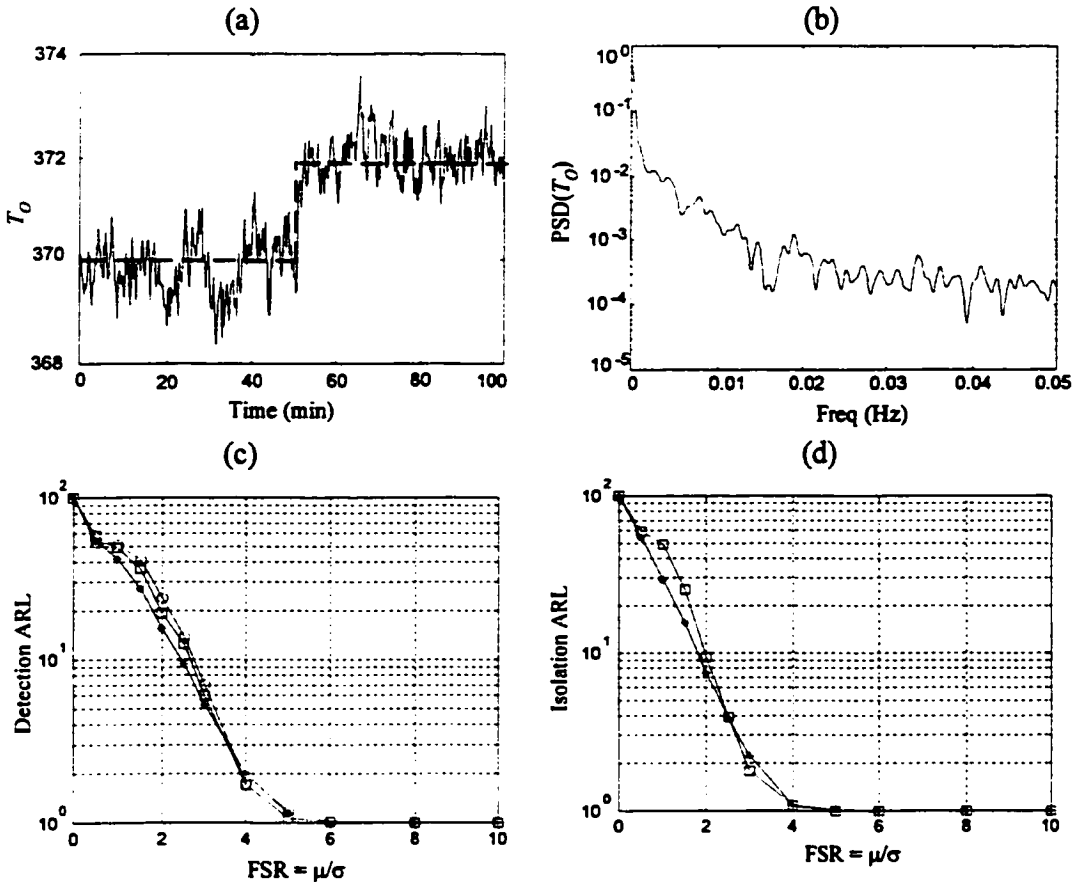


Figure 4.5 FDI of T_O sensor bias; (a) Process trend with $4\sigma_{To}^2$ (σ_{To}^2 =0.5026); (b) Frequency characteristics of T_O sensor; (c) Detection ARL based on Monte Carlo simulation; (d) Isolation ARL based on Monte Carlo simulation (# of simulations at one FSR: 500, α=0.01; Scaling on X_G; ∗: T^2/SPE(PCA); □:T^2/SPE(GPCA); O: T_j^2/SPE$_j$)

and isolation ARL for the sensor bias. The detection and isolation ARL of the GPCA based on T^2/SPE, and the scale contribution based on T_j^2/SPE_j show almost the same results as those of the regular PCA. This is due to the characteristic of the sensor bias whose high frequency effect is not captured and instead the effect is spread over wide frequency ranges. The simulation result confirms that the multiscale model does not significantly outperform the regular PCA in such a case.

4.3.4 Sensor spike

Figure 4.6(a) shows a spike on the inlet temperature measurement. The fault is simulated to occur at the 51st minute. All the simulation conditions are the same as before except for the type of fault. The magnitude of the fault shown is $4\sigma_{To}^2$ where σ_{To}^2 is the total variance of the inlet temperature measurement. The fault, a sudden change in the measured value of the temperature occurs at a higher frequency band. Since it is an instantaneous action, it is not clearly shown in PSD plot of the sensor spike. Instead, the signal content of the sensor spike at low frequency shows a strong intensity.

Monte Carlo simulations were done for the sensor spike case to generalize the result obtained. All the simulation conditions were the same as those for the previous simulation except for the type of fault. 500 simulations were made for each magnitude of the fault. Figure 4.6(c) and (d) show the ARL of several statistics for fault detection and isolation of the T_O sensor spike. Both plots show that the detection and isolation performance is almost the same in the three cases. It results from the fact that the effect of the sensor spike is easily missed when its magnitude is small while the effect of the sensor spike is shown clearly at all the frequency ranges when its magnitude is large. The FDI using the GPCA and its scale statistics does not provide an enhanced performance. It is also noted that the overall performance of the FDI with respect to the sensor spike is worse than those with respect to the sensor bias.

4.3.5 Complex fault

In contrast with the simple faults simulated above, a complex fault is the one

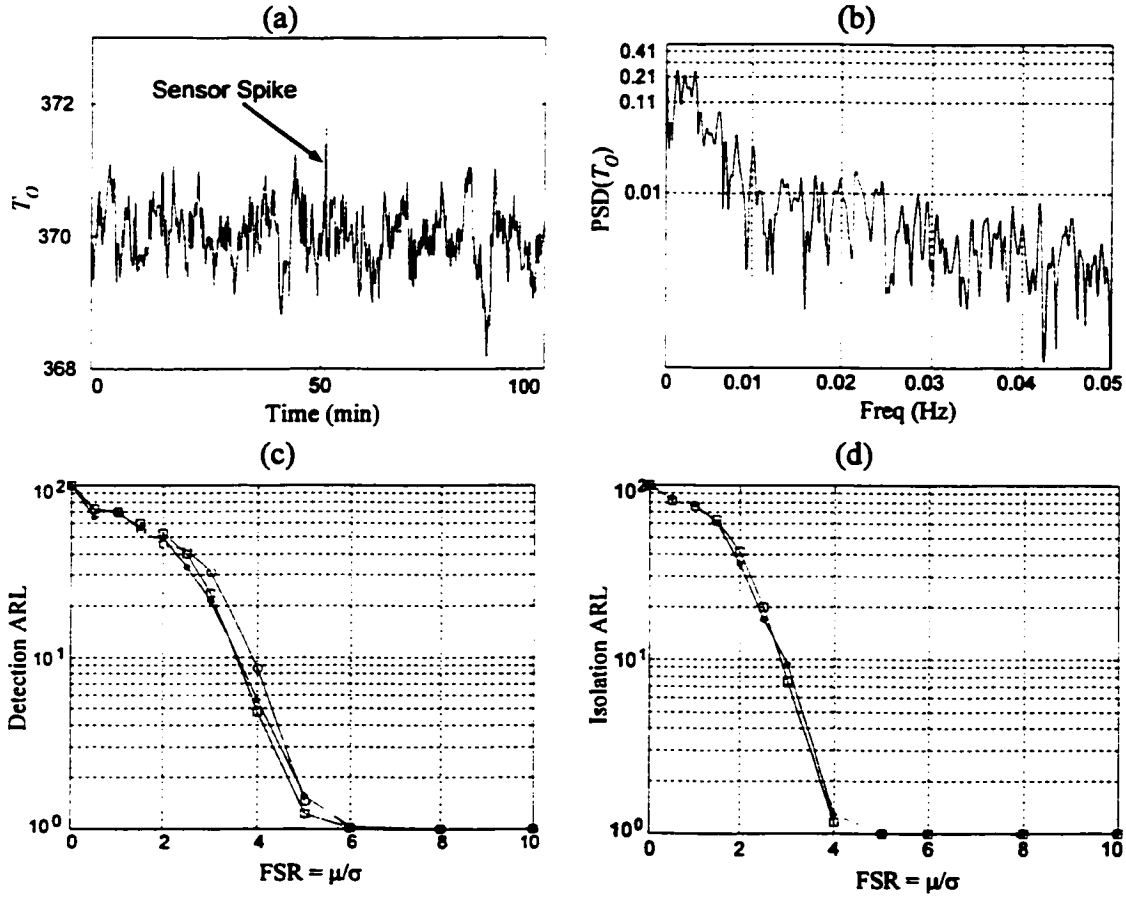


Figure 4.6 FDI of T_O sensor spike; (a) Process trend with $4\sigma_{To}^2$ (σ_{To}^2 =0.5026); (b) Frequency characteristics of T_O sensor (c) Detection ARL based on Monte Carlo simulation; (d) Isolation ARL based on Monte Carlo simulation (# of simulations at one FSR: 500; α=0.01; Scaling on $\mathbf{X}_G$; ∗: T^2/SPE(PCA); □:T^2/SPE(GPCA); O: T_j^2/SPE_j)

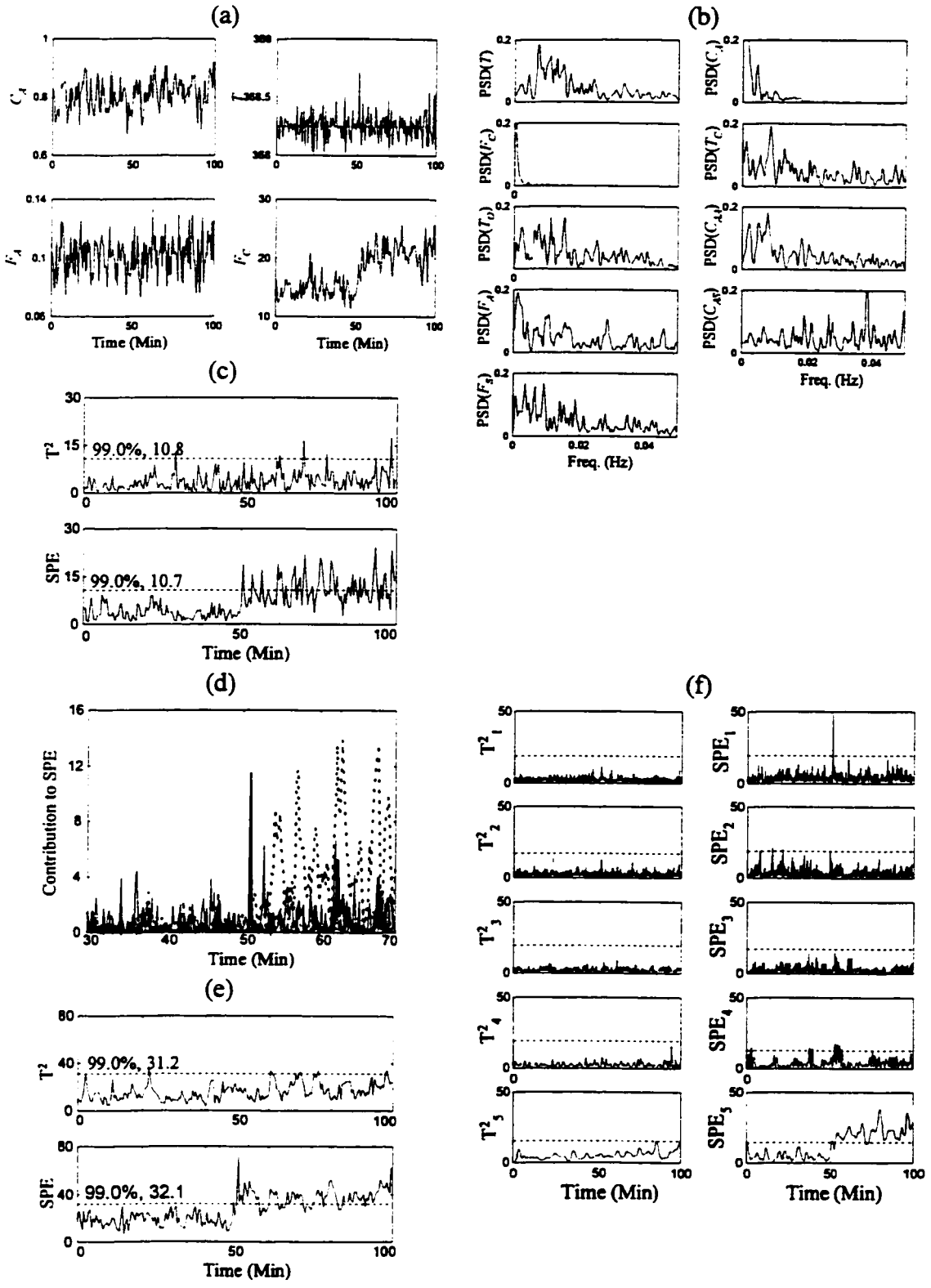
whose effects are propagated into other measurements. Complex faults occur in forms of bias, drift changes in process parameters, or controlled variables by which many measurements are confounded. The isolation of complex faults becomes complicated and problematic since process dimension and complexity increase fault propagation. The complex fault has an initial fault signature, a time varying trajectory and a steady-state fault signature. The initial fault signature may provide a reliable and prompt source for fault isolation since it is not affected by fault propagation. However, the initial fault

signature can be easily missed by any fault detection scheme. The transient behavior and final steady-state vector of the measurements are thus generally used as an alternative to the initial fault signature. It is highly demanding in industry to conduct rigorous analysis of the complex faults and hence enhance its fault isolation.

Assume a measurement bias of −1 °C on the reactor outlet temperature shown in Figure 4.7(a). Since the reactor outlet temperature is the controlled variable, the effect of measurement bias is propagated into other variables. The coolant flow rate that is the manipulated variable is changed to remove the deviation of the reactor outlet temperature from its setpoint, and this causes a change in the reactor outlet concentration.

Figure 4.7(b) is PSD plots for the testing date set. PSD plots of all the other variables except for the coolant flow rate show similar features to the base-case shown in Figure 3.7 of the previous chapter. Low frequency content of the coolant flow rate is shown to be significant. This change is an action at a low frequency level compared to the random variation in normal case. On the other hand, the fault, a sudden change in the temperature measurement occurs at a higher frequency band. Since it is an instantaneous action, it is not clearly shown in PSD plot of the fault case.

In the conventional PCA, the T^2 and SPE statistics include all the events occurring at all scale bandwidths. Figure 4.7(c) of the PCA T^2/SPE statistics clearly detects the fault and the variable contribution plot in Figure 4.7(d) indicates that the outlet temperature and the coolant flow rate are associated with the fault. Similarly, the GPCA based T^2/SPE monitoring plots in Figure 4.7(e) detect the fault successfully. The scale contributions to T^2 and SPE in Figure 4.7(f) were examined. It is shown that both the lowest and highest scale blocks are responsible for the fault. In the variable contribution plot to the SPE_1 of the D_1 scale in Figure 4.7(g), the outlet temperature measurement is clearly associated with the fault. The variable contribution plot to the SPE_5 of the A_4 scale, highest approximation block in Figure 4.7(h) indicates that the measurement of the cooling water flow rate is associated with the fault. Based on these two plots, one can conclude that the measurement of the outlet temperature may be a root source of the fault and the cooling water flow rate as an affected variable. This example illustrates how GPCA does help in


(a)
Time (Min)
Time (Min)
(b)
Freq. (Hz)
Freq. (Hz)
(c)
99.0%, 10.8
99.0%, 10.7
Time (Min)
(d)
Contribution to SPE
Time (Min)
(e)
99.0%, 31.2
99.0%, 32.1
Time (Min)
(f)
Time (Min)
Time (Min)

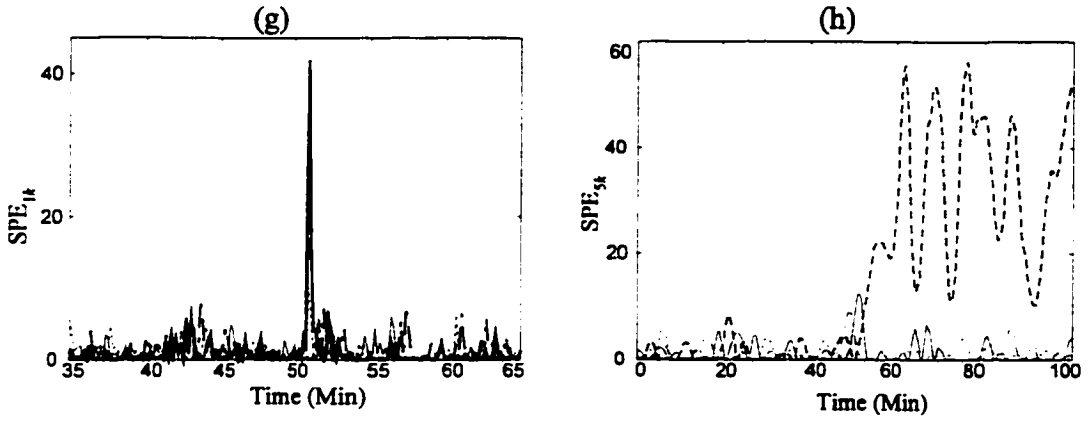


Figure 4.7 Detection and isolation of a complex fault with ΔT=-1 at 51 minutes; (a) Process trend; (b) Frequency characteristics; (c) T^2/SPE(PCA); (d) Variable contribution to T^2; (e) T^2/SPE(GPCA); (f) T_j^2/SPE_j; (g) SPE_{1k}; (h) SPE_{5k} (T: —; C_{AO}: ⋯; F_C: --; T_O:— ; F_A: — •)

the fault isolation through contribution plots in specific scales even though GPCA doesn't improve the detection of the fault.

4.3.6 Scaling effect

In order to examine the scaling effect on the FDI performance, the Monte Carlo simulation has been done for the sensor precision degradation with the scaling on **X** instead of $\mathbf{X}_G$. The simulation conditions used in the previous examples were used.

Figure 4.8(a) thus shows the detection ARL values for the T_O sensor precision degradation. ARL has been calculated with T^2 and SPE statistics of the regular PCA, the GPCA, and the scale contribution of the GPCA model (MSPCA). Compared to Figure 4.3(a) the detection ARL of GPCA in this example is degraded. It is because one would put much greater weight on the faults occurring in high frequency by $\mathbf{X}_G$ scaling while **X** scaling will weight the less dominant low frequencies more. Thus, the sensor precision degradation with the same magnitude is better detected in $\mathbf{X}_G$ scaling case than in **X** scaling case. As a result, the detection ARL is shown to be shorter in the $\mathbf{X}_G$ scaling case.

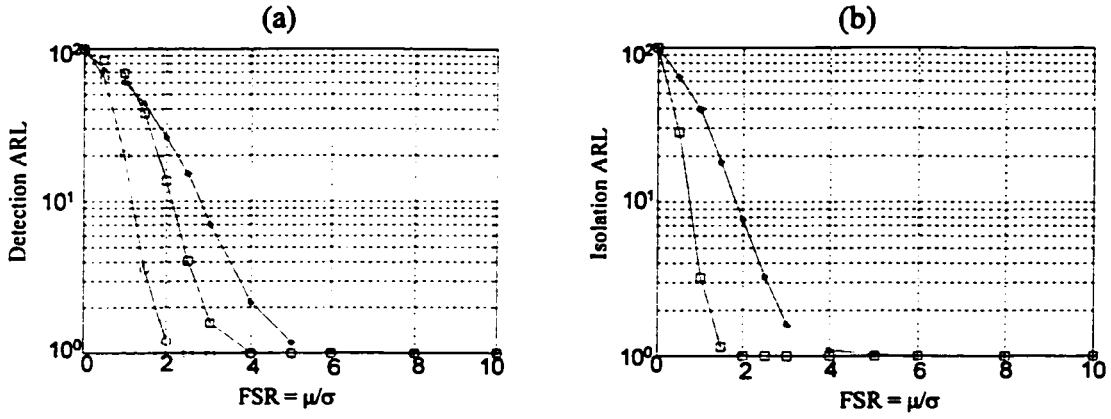


Figure 4.8 Scaling effect: Monte Carlo simulation of T_O sensor precision degradation; (a) Detection ARL; (b) Isolation ARL (# of simulations at one FSR: 500; α=0.01; Scaling on X; ∗: T^2/SPE(PCA); □:T^2/SPE(GPCA); O: T_j^2/SPE_j);

On the other hand, the isolation ARL values in Figure 4.8(b) show the same result as given in Figure 4.3(b) regardless of the scaling since the same statistics T_{jk}^2/SPE_{jk} are used in both cases.

4.4 Conclusion

In the previous chapter, GPCA was proposed as a modified version of PCA with multiscale data. It has been addressed in this chapter how to use the GPCA method for fault detection and diagnosis. A procedure for both fault detection and isolation is presented. Contributions of the various scales to the overall T^2 and SPE, and contributions of the variables to each scale have been proposed as additional tools for fault isolation. Due to the usage of scale information, FDI performance can be significantly enhanced.

FDI performance using the GPCA has been assessed through Monte Carlo simulation with the CSTR system for several types of faults and the complex fault problem. When an abnormal event occurs at one scale bandwidth, the GPCA based FDI method will be more effective than that based on the regular PCA model in detecting and

isolating faults. However, it appears to provide little improvement if the fault effect is spread over more than one frequency band, or the fault effect in one scale is minor compared to the other scale components. Thus, a monitoring method that gives the best detection and isolation of faults will depend upon the fault characteristics. An appropriate fault detection and isolation method has to be determined by considering fault characteristics.

A well-known benefit of the wavelet methodology is the time-frequency localization. This feature might be useful to explicitly characterize fault signatures for fault diagnosis. Thus, the fault isolation method presented in the chapter 2 can be used together with fault signatures based on GPCA model. In addition, GPCA method can also be used as an approximate dynamic modeling method. Research on these issues were not explored further, but would be worthwhile as a continuing research work.

5. Statistical and Causal model based approaches to FDI

5.1 Introduction

Various approaches to fault detection and isolation (FDI) can be classified into three categories: (i) methods based on causal models, (ii) methods based on qualitative knowledge, and (iii) statistical methods based on correlation models (Figure 5.1). The causal process models are obtained from theory, or identified empirically from designed experiments. Most common FDI frameworks are based on them. However, obtaining a complete and robust causal model is difficult due to process complexity and dimension. Therefore this approach is generally limited to processes with a small number of variables.

On the other hand, statistical correlation models developed using PCA and PLS can easily handle large number of variables (hundreds), since they are built from routine

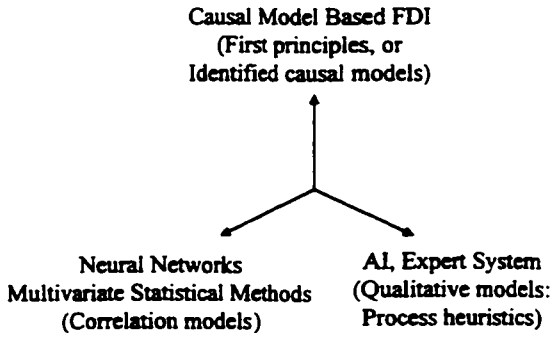


Figure 5.1 FDI approaches and resources

operating data in historical databases. These data are readily available and require no plant testing or fundamental knowledge. FDI approaches using these models fall under the category of multivariate statistical process control (MSPC). Although they are very powerful for fault detection, their main limitation lies in their ability to isolate or diagnose faults. It inevitably suggests that combining the approaches may resolve the many of difficulties inherent in each of them individually. However, before this can be attempted the fundamental differences between these approaches must be clearly understood.

This chapter presents the fundamental differences between the causal-model-based methods and the multivariate statistical methods to FDI problems, and illustrates them via several simulation studies. The parity relation approach is selected to represent the causal-model-based methods and a PCA (Principal Component Analysis)-based-method to represent the MSPC approaches. By examining the assumptions, model characteristics, and FDI performances to various example problems, the two approaches are overviewed, critical differences are assessed, and their strengths and weaknesses are highlighted. Recognizing the differences between the two approaches should help in the selection of appropriate approaches and provide insight into how the approaches can be combined.

An outline of this chapter is as follows. In the next section, the FDI using the causal models are reviewed, and the parity relation and PCA based FDI approaches are introduced. The fundamental differences and relationships between two methods are then addressed. By examining the differences in the models, the data required to build models, the processes to which they are applicable, and the assumptions behind the methods, various fundamental aspects of FDI system design using two approaches are compared. Practical differences in these approaches are illustrated using simulated data from a continuous stirred tank reactor process.

5.2 Causal model based method

The general procedure of causal-model-based methods consists of two steps. The first step is the generation of analytical redundancy given by the mechanistic model. The second step is detection and isolation of the faults. This approach requires that the residual generator perform the validation of nominal relationships of the system, using the actual input and output. The redundancy relations to be evaluated can simply be interpreted as input-output relations of processes. If a fault occurs, the redundancy relations are no longer satisfied and nonzero residual occurs. The residuals are then used to form appropriate decision functions.

The basis for the decision on the fault occurrence is a fault signature. The fault signature is obtained from the effects associated with the fault. FDI is completed when fault location and fault time are identified in most applications. In special cases, however, one needs a deeper insight into the situation by knowing fault type, size, and cause that can be acquired by a subsequent fault diagnosis. For this purpose, deeper knowledge about the nature of processes such as degree of aging, operational environment, history of operation and maintenance, and fault statistics, is required. This task is, therefore, commonly tackled with the aid of an expert system as a complementary scheme to the model based methods.

There are two ways of generating fault signals using analytical redundancy: state estimation techniques, and parameter estimation approaches. Reviews of model-based FDI methods can be found in survey papers by Willsky (1976), Isermann (1984), Basseville (1988), Frank (1990); Patton (1995) and Gertler (1995).

The methods based on the residual generation include the direct consistency (parity) relations, the diagnostic observers and the Kalman filters relations. They use the state-space or input/output models of processes. Parity equation approach appeared in the late 70s and early 80s, and has been used in various fields (Chow and Willsky, 1984; Gertler and Singer, 1985). Dynamic consistency relation was used by Chow and Willsky (1984), and extended by several researchers (Gertler, 1998; Gertler and Singer, 1990). The key idea is to check the parity of the mathematical equations of a system that is the analytical redundancy relations by using actual measurements. A fault is declared to have

occurred once preassigned error bounds are surpassed.

Diagnostic observers appeared in the early 70s, and have been actively studied. The Kalman filter approach was first proposed by Mehra and Peschon (1971) and further developed by Willsky and Jones (1976). Many authors have approached FDI by directly starting with a single or a bank of Luenberger observers or Kalman filters (Willsky, 1976). The basic idea of these two approaches is to reconstruct the outputs of processes from the measurements, or subsets of the measurements, with the aid of observers or Kalman filters. This is the residual for the detection and isolation of the faults. Mehra and Peshon (1971) used a single Kalman filter driven by the full order vector. They used knowledge that the residual is white noise with zero mean and known covariance when no fault occurs. The occurrence of a fault is monitored by statistical tests of whiteness, mean, and covariance. The fault isolation is carried out on the basis of different fault hypotheses. The multiple hypothesis testing can be carried out using Bayesian decision theory (Willsky, 1976). One can use linear or nonlinear, full or reduced-order state observers in the deterministic case, or Kalman filters in the stochastic case.

As an alternative to the methods based on the state estimation, the parameter identification approach uses the fact that faults of processes are reflected in the physical parameters such as friction, mass, viscosity, resistance, capacitance, inductance, etc. The idea of the parameter identification approach is to detect faults via estimation of the parameters of the mathematical model (Isermann, 1984). This approach may be particularly useful for the detection of incipient faults. Recently Fathi *et al.* (1993) incorporated state and parameter estimation modules within the diagnostic reasoning of a knowledge-based system to overcome the deficiencies of both approaches and to increase the diagnostic ability of the system.

Parity equations and diagnostic observers have proved to lead to identical residual sets (Magni and Mouyon, 1994). Relationships between parity equation and parameter identification methods (Gertler, 1995a), and between observer based and parameter-identification-based methods were studied (Garcia and Frank, 1996)

All causal-model-based methods remarked above are built upon a number of rather

idealized assumptions. One of which is that the mathematical model used is a faithful replica of the plant dynamics. The term "model", here means a mechanistic, causal process model based on first principles and usually consists of material or energy balances. This is often not possible in practice. The system parameters may not be fully known or may be partially known over a limited range of the plant's operation. Another idealized assumption is that the true characteristics (e.g., spectral shape, variance, mean, stationarity, etc) of disturbance and noise signals acting upon system are known. Thus, all faults must be defined at the design stage and properly reflected in the fault isolation scheme. This method, therefore, would be easier to apply for well-defined processes like electrical and mechanical processes.

Along with the above assumptions, FDI methods must be generally sensitive to the appearance of faults, but insensitive (robust) to other changes like noises, modeling errors, operating conditions, etc. Because these requirements often contradict each other, various trade-offs exist. The trade-offs are on i) size of fault vs. detection time, ii) speed of fault appearance vs. detection time, or process response time, iii) size and speed of fault vs. speed of process parameter changes, iv) type I vs. type II error rate, etc. Thus, as the threshold for the detection of faults under the model based methods is set higher, the possibility of false alarm detection is reduced, but FDI system becomes less sensitive. These sensitivity and robustness issues are a critical concern in implementing the model based FDI methods.

5.3 Parity relation approach

For the comparison purpose, the parity relation approach is chosen. The design methodologies and techniques have been extensively studied and published by Gertler and coworkers (Gertler, 1998). Parity relations are rearranged and transformed forms of the input-output model equations of processes. They are based on linear process dynamic models in state-space form or in an equivalent input/output form (Gertler and Singer, 1990; Gertler and Monajemy, 1995):

$$\mathbf{y}_k^\circ = \mathbf{M}^\circ(z)\mathbf{u}_k^\circ, \tag{5.1}$$

where $\mathbf{y}_k^\circ$ and $\mathbf{u}_k^\circ$ are the true process outputs and inputs acting on the true process, $\mathbf{M}^\circ(z)$. The transfer function model of a process, $\mathbf{M}(z)$ can be obtained either from first principles of a process or from identification using designed experiments satisfying all the identifiability conditions (Ljung, 1999). It can also be derived from the state space model of the process as $\mathbf{M}(z) = [\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}+\mathbf{D}]$ where **A**, **B**, **C**, and **D** are state-space model matrices. The residuals between the true process outputs, $\mathbf{y}_k^\circ$ and their model estimates, $\hat{\mathbf{y}}_k$ can be defined as:

$$\mathbf{r}_k = \mathbf{y}_k^\circ - \hat{\mathbf{y}}_k = \begin{bmatrix}\mathbf{I} & -\mathbf{M}(z)\end{bmatrix}\begin{bmatrix}\mathbf{y}_k^\circ & \mathbf{u}_k^\circ\end{bmatrix}^T. \tag{5.2}$$

By introducing faults ($\mathbf{x}_k = \mathbf{x}_k^\circ + \Delta\mathbf{x}_k$) on the true process variables, the parity relations are then obtained from the residual equations in terms of the actual measurements ($\mathbf{y}_k$ and $\mathbf{u}_k$) and the faults as follows:

$$\begin{aligned}\mathbf{O}_k &= \begin{bmatrix}\mathbf{I} & -\mathbf{M}(z)\end{bmatrix}\begin{bmatrix}\mathbf{y}_k & \mathbf{u}_k\end{bmatrix}^T \\ &= \begin{bmatrix}\mathbf{I} & -\mathbf{M}(z)\end{bmatrix}\begin{bmatrix}\Delta\mathbf{y}_k & \Delta\mathbf{u}_k\end{bmatrix}^T, \\ &= \mathbf{F}\Delta\mathbf{x}_k\end{aligned} \tag{5.3}$$

where the first line in (5.3) is called the external parity relation. It is based on the actual measurements and is used for computing all the residuals. The second line in (5.3) is called the internal parity relation. It shows how the observed residuals depend upon the unknown faults. The parity relationship becomes non-zero if there is any measurement fault in the variables contained in the model or if there is any process fault which affects the parity relations. Therefore, one can detect faults by checking these primary parity relations. However, fault diagnosis, or isolation tasks are more complicated because non-

zero parity relations can result from any faults and so one has to try to isolate the root fault by interrogating all parity relations. To facilitate isolation, parity relations are usually enhanced so that the residual set has one of the following properties (Gertler and Kunwer, 1995).

- *Structured residual sets*. In response to a particular fault, only a fault specific subset of the residuals becomes non-zero.
- *Fixed direction residuals*. In response to a particular fault, the residual vector is confined to a fault-specific direction. Isolating a fault amounts to determining to which of the pre-defined directions the observed residuals lie the closest.

With such residuals, the design of FDI systems using parity relationships involves generating additional parity relations, or rearranging the parity relations such that FDI system has the above properties and fault isolation can be easily done. The directional residuals are usually used when $f=m$ and structural residuals when $f \geq m$, where f and m are the number of specified faults and the number of independent measurements, respectively. To deal with large processes, one can build many FDI systems, one for each individual unit.

When equivalent parity relations are obtained from empirically identified process models using designed plant tests, all identifiability requirements must be satisfied in performing the tests (Ljung, 1999). As with all linear models they are valid only around the operating conditions at which the nonlinear process is linearly approximated. The parity relation approach thus cannot be easily applied to batch, or semi-batch processes where operating conditions vary continuously. To handle well-defined nonlinear systems, instead of obtaining the structured residuals by algebraic operations on the primary parity equations, the desired nonlinear residual structures can be directly identified (Gertler, 1998).

In this FDI approach, it is usually assumed that there are no process uncertainties such as modeling errors or unmeasured disturbances, and that the explicit model can explain all

faults. If any of these assumptions is violated, the performance of the FDI system will be degraded, and the FDI system must be redesigned in a proper manner to account for the assumption violations.

5.4 Statistical model based method – MSPC approach

The detail explanation of this method is given in section 2.2 and thus it is not repeated in this section.

5.5 Comparisons

Table 5.1 provides a comparison between the statistical and causal model-based approaches under various headings. The major difference between the approaches is the nature of the models used, and the types of data required to build them. The parity relation/state estimator approaches usually require models that define the causal effects of all inputs on all outputs. For this one needs mechanistic models or empirical models obtained from identification studies using designed experiments. The data from these identification studies must be persistently excited and satisfy all identifiability conditions (Ljung, 1999; Soderstrom and Stoica, 1989). On the other hand, for the statistical approaches a causal model is not even desirable. They require a model for the covariance structure among all measured variables when only *common-cause* variations are present. Data for building such models is readily available in databases from periods where the plant was operating in a normal and well-behaved manner.

The major strengths and weaknesses of the two different approaches then arise from this difference. The causal model approaches are generally limited to well-defined systems with a small number of variables, while the statistical approaches can easily handle very large and ill-defined processes. On the other hand, the parity equation approach allows for much more direct isolation of known faults through knowledge of

Table 5.1 Comparison on FDI approaches

	Causal model-based method	Statistical model based method
Model and Data	First principles model or an empirical model identified from designed experiment Model provides causal effects for all inputs on all outputs	A correlation model built from undesigned normal operating data a Model provides no causal effects. It only models the covariance structure among all measured variables under normal operating
Technologies	State estimator, Kalman filter, Observer, or input/output parity equations Parameter estimation	Latent variable methods - PCA - PLS
Tools	Structured residuals Directional residuals	Hotelling's T^2, score, SPE plots Contribution plots Directional residuals of prior faults
Size and Applications	Small number of variables (10-20) Incorporates only variables for which one has causal models Well defined processes where causal models available, e.g., Electrical & mechanical processes	Large number of variables (hundreds) Incorporates all measured variables Less-well defined processes e.g., Petrochemical, Resources (steel, pulp & paper and semi-conductor industries)
Detection	Easily done Approach: Fault breaks causal parity relationship	Easily done Approach: Fault breaks variable covariance structure existing under normal operations
Isolation	Causal model allows for direct isolations of simple and well modeled faults Not possible for unmodeled faults Can handle multiple faults with special design	Contribution plots provide for easy isolation of simple faults but ambiguous for complex and multiple faults Need additional information on causal effects or on past fault histories to isolate complex Faults
Remarks	Detection and isolation done together Assumes predefined faults and known disturbances	Detection and isolation performed sequentially Needs representative normal operating data containing all sources of *common-cause* variations Often uses only steady-state models, but readily applied to dynamic models

the causal structure, while in the statistical approaches the isolation is much more indirect because of the absence of causal information. The statistical approaches must be

supplemented with some causal information or prior fault knowledge to provide less ambiguous isolation.

Recently Gertler *et al.*, (1999) proposed an "isolation enhanced PCA" approach. It was shown that, with data from designed experiments, one could use the last principal component relationships directly as parity relations ($\varepsilon_k = \tilde{\mathbf{P}}^T \mathbf{x}_k = \tilde{\mathbf{P}}^T \Delta \mathbf{x}_k$ since $\tilde{\mathbf{P}}^T \mathbf{x}_k^\circ = \mathbf{0}$) and apply the same transformation procedures to get isolation properties. Alternatively, one could compute the explicit causal model (5.3) from them ($\tilde{\mathbf{P}}^T = \mathbf{WF}$, where $\mathbf{W}$ is a full-rank square matrix) and apply the parity equation methods directly. However, this approach is clearly still that of the causal model-based approach. Their models are built from designed experiments and provide causal relationships among all the inputs and outputs. The only difference is that PCA is used as the identification method for the causal models rather than the more traditional prediction error methods (Ljung, 1999; Soderstrom and Stoica, 1989). This use of PCA to identify linear causal dependences among the variables from the eigenvectors associated with the smallest eigenvalues is based on literature on the last principal component methods for identification (Ku *et al.*, 1995; Negiz and Cinar, 1997) and the literature on total least squares. With no noise, the principal components corresponding to zero eigenvalues would define the space of exact linear relationships among the variables. This approach to identification works well in deterministic situations, but generally is poor compared to prediction error approaches when noise and unmeasured disturbances are present (Negiz and Cinar, 1997).

In the statistical approaches to FDI, latent variable methods such as PCA and PLS are employed in a very different manner. They are used to obtain information on the dominant directions of variations (eigenvectors associated with the largest eigenvalues) that are present in normal operating data. Under these conditions, it is usual to have only a few (2 to 5) dominant directions that explain most of the process variations arising from these *common-causes*. The statistical approaches to FDI then rely upon referencing future behavior against this normal behavior defined by the low dimensional PCA model of the hyper-plane. Since the MSPC approach incorporates the effect of all normal disturbances into the PCA model, it will also automatically account for these disturbances in the

detection and isolation steps. Clearly the two FDI approaches are completely different even though PCA is used (albeit for different purposes) in both.

5.6 Numerical examples

In this section, simulation studies on a nonisothermal continuous stirred-tank reactor (CSTR) model (Marlin, 1995, page 90-92) explained in Appendix A are performed to illustrate the two FDI approaches and point out critical differences. In all studies, the true process model was used to obtain the linearized parity relations for the reactor outlet concentration (C_A) and the temperature (T). The PCA model for the MSPC approach was obtained with a set of observations collected for 200 minutes from the process under routine operation when no faults were present. (In real processes, much more data would be required to capture all sources of *common-cause* variations.)

Before investigating fault detection and isolation by the two approaches, we examine the fault characteristics of the process. Based on (5.3), the parity relations at any sampling interval are as follows;

$$\begin{aligned}\mathbf{O} &= \mathbf{F}[C_A \quad T \quad C_{AS} \quad C_{AA} \quad F_S \quad F_A \quad T_O \quad T_C \quad F_C]^T \\ &= \mathbf{F}[\Delta C_A \quad \Delta T \quad \Delta C_{AS} \quad \Delta C_{AA} \quad \Delta F_S \quad \Delta F_A \quad \Delta T_O \quad \Delta T_C \quad \Delta F_C]^T\end{aligned}$$

where $\mathbf{O} = [O(C_A) \quad O(T)]^T$, and the parity relation matrix, **F**, obtained from the mechanistic model of the CSTR process is as follows:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & -0.3254 & -0.0362 & 0.2618 & -6.6072 & 0.0049 & 0.0264 & -0.0034 \\ 0 & 1 & -1.1779 & -0.1309 & 0.6224 & -24.2450 & -0.1678 & -0.8965 & 0.1143 \end{bmatrix}$$

As seen in **F**, the parity residual for C_A, $O(C_A)$ is expressed in terms of all process variables except for T since $F_{12} = 0$. Similarly, the primary parity residual for T, $O(T)$

is expressed in terms of all process variables except for C_A since $\mathbf{F}_{21}$=0. Thus, in principle, one can isolate only the faults of C_A and T measurements with the primary parity relations. To isolate faults of the other measurements one needs to generate secondary parity relations by taking linear combinations of the primary parity relations.

The i-th column of parity relation matrix $\mathbf{F}$ defines the fault direction of the i-th sensor fault $\Delta\mathbf{x}_i$ in (3). That is, the columns of $\mathbf{F}$ correspond to the effects of faults in the measurements of C_A, T, C_{AS}, C_{AA}, F_S, F_A, T_O, T_C and F_C on the primary residuals $O(C_A)_k$ and $O(T)_k$, respectively. Fault isolability is largely dependent on the fault directions. The larger the difference in fault directions among the various faults, the better is the ability to isolate them.

Figure 5.2 shows the normalized fault directions of the CSTR process. The normalized fault directions ($\mathbf{O}_i^\circ = \mathbf{F}_i^\circ$ assuming $\Delta\mathbf{x}_j$=1 if j=i, otherwise 0) are obtained by scaling all columns of $\mathbf{F}$ to unit length. From the columns of the normalized $\mathbf{F}^0$ and their plots in Figure 5.2, one can see that the fault directions of F_C, T, T_O and T_C are either perfectly collinear or nearly collinear, and hence cannot be isolated from one another.

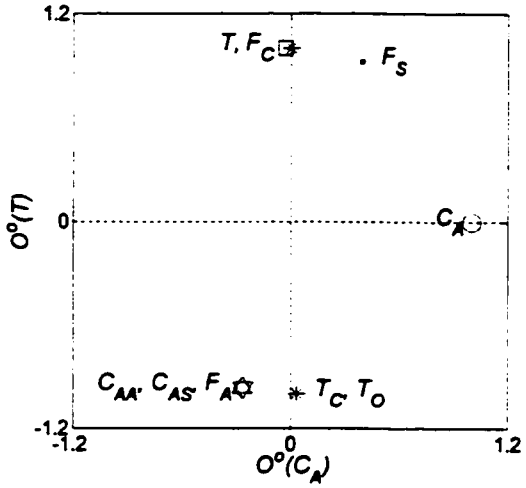


Figure 5.2 Normalized fault directions ($\mathbf{O}^\circ$) of CSTR system (C_A:○; T:✱; C_{AS}:△; C_{AA}:◇; F_S:●; F_A:▽; T_O:✚; T_C:✕; F_C:□; $\mathbf{O}_i^\circ = \mathbf{F}_i^\circ$)

$$\mathbf{F}^\circ = \begin{bmatrix} 1 & 0 & -0.2662 & -0.2662 & 0.3876 & -0.2629 & 0.0295 & 0.0295 & -0.0295 \\ 0 & 1 & -0.9639 & -0.9639 & 0.9218 & -0.9648 & -0.9996 & -0.9996 & 0.9996 \end{bmatrix}$$

The fault directions of C_{AS}, C_{AA}, and F_A are also perfectly collinear and cannot be isolated from one another. Furthermore, the angles between the fault directions of the (F_C, T, T_O, T_C) group, that of the (C_{AS}, C_{AA}, F_A) group, and that of F_S are small implying that they could be difficult to isolate from one another in the presence of noises and disturbances. Only a fault in C_A is clearly isolatable from the rest. These exact or near linear dependences between columns of **F** cause ill-conditioning in the residual equations used in the fault isolation and make the corresponding faults impossible or difficult to distinguish. For this reason only the residuals in C_A and T obtained from the primary parity equations are used in the following FDI analysis. Deriving residuals for all the other variables (that are collinear or nearly collinear with T) through algebraic transformations of these two primary parity equations would yield few useful results in this case. In order to enhance the fault isolability, one needs additional measurements which will break the correlation between the sensor fault directions.

5.6.1 FDI of simple faults

This case study is given to show how the two FDI approaches work for the detection and isolation of simple faults occurring at the normal operating condition. In the first simulation a reactor outlet concentration sensor bias (ΔC_A = 0.2) occurs at 51 minutes. The fault in C_A is representative of a *simple output sensor fault*. Its effect is not propagated into other process variables (Yoon and MacGregor, 2000a). The fault detection and isolation should be relatively easy compared to those of a complex fault.

Figure 5.3(a) shows the trends of the primary parity residuals. They clearly indicate that a fault has occurred in the C_A measurement at 51 minute. Figure 5.3(b) shows two monitoring plots for the MSPC approach. Both the SPE and T^2 plots clearly indicate that there is an unusual event around 51 minute. The SPE contribution plot in Figure 5.3(c) reveals that the measurements contributing most to the abnormal event at 51 minute are C_A and the coolant flow rate (F_C). In this particular example, faults in these two sources are completely confounded in routine operating data, and so they are difficult to isolate

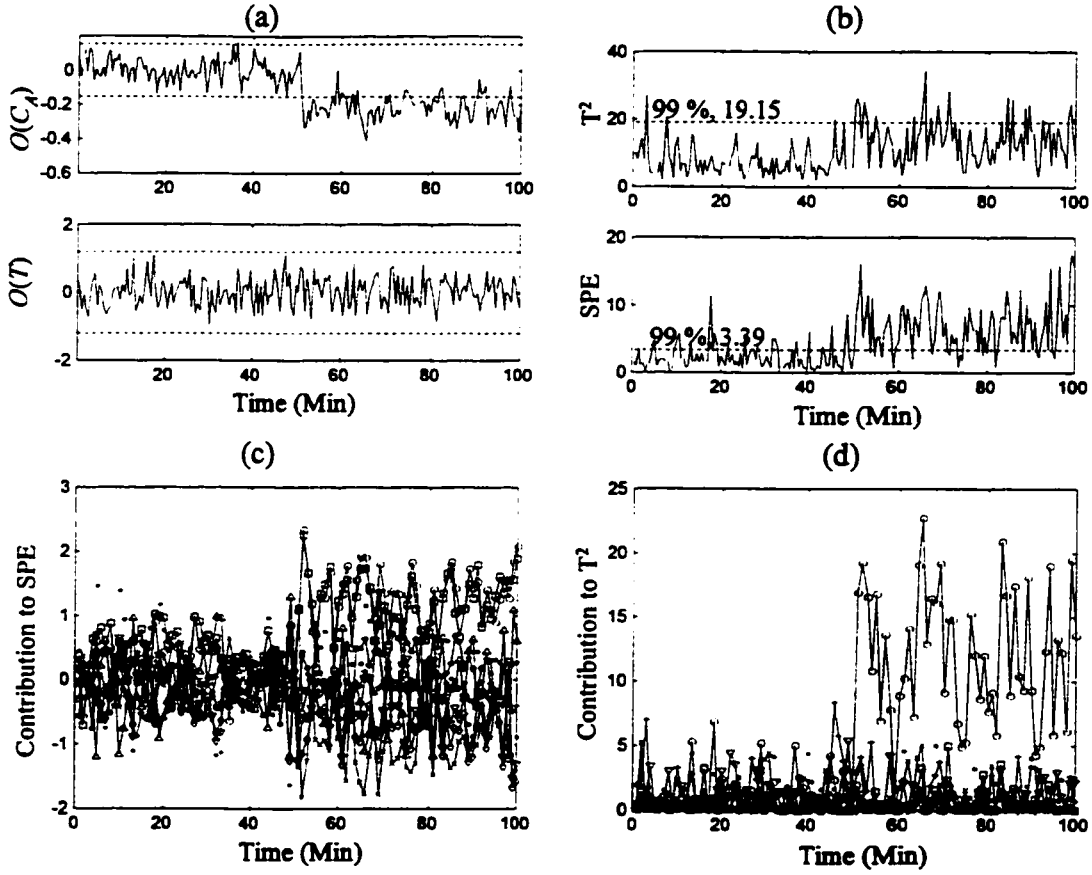


Figure 5.3 FDI of simple bias fault (ΔC_A = 0.2 at 51 min); (a) Parity relation plots; (b) SPE/T^2 plots; (c) SPE contribution plot; (d) T^2 contribution plot (C_A:O; T:✱; C_{AS}:△; C_{AA}:◇; F_S:•; F_A:▽; T_O:+; T_C:×; F_C:□)

using contribution plots. One can interpret the increase in SPE as a breakdown of the *common-cause* correlation that is usually present between C_A, F_C, and the remaining variables. On the other hand, the T^2 contribution plot of Figure 5.3(d) more clearly isolates the fault as being the C_A sensor because of the large increase in the contribution of C_A starting at 51 minutes. As a result a simple bias fault has been detected and isolated through a breakdown of the normal correlation among the variables (SPE plot) and a larger variation than normal magnitude of contribution of variation in a variable (T^2 plot). In real industrial processes where many additional correlated variables are measured, contribution plots usually can clearly isolate simple sensor faults (e.g., Kourti *et al.*,

1996).

As a second example, a reactant flow rate bias (ΔF_A = -0.015) is simulated to occur at 51 minutes. This is also a simple sensor fault, but one in an input sensor. The fault should be detectable using the two primary parity residual equations, $O(C_A)$ and $O(T)$ since these equations will no longer hold with the fault in F_A being present. The residual plots in Figure 5.4(a) show that there is a significant shift in $O(C_A)$, but not in $O(T)$ around 51 minutes. Thus the fault is detected, but cannot be successfully isolated using only the primary parity residuals. One might conclude that the fault is on the C_A measurement. Generating a secondary parity relation (by a transformation on the primary

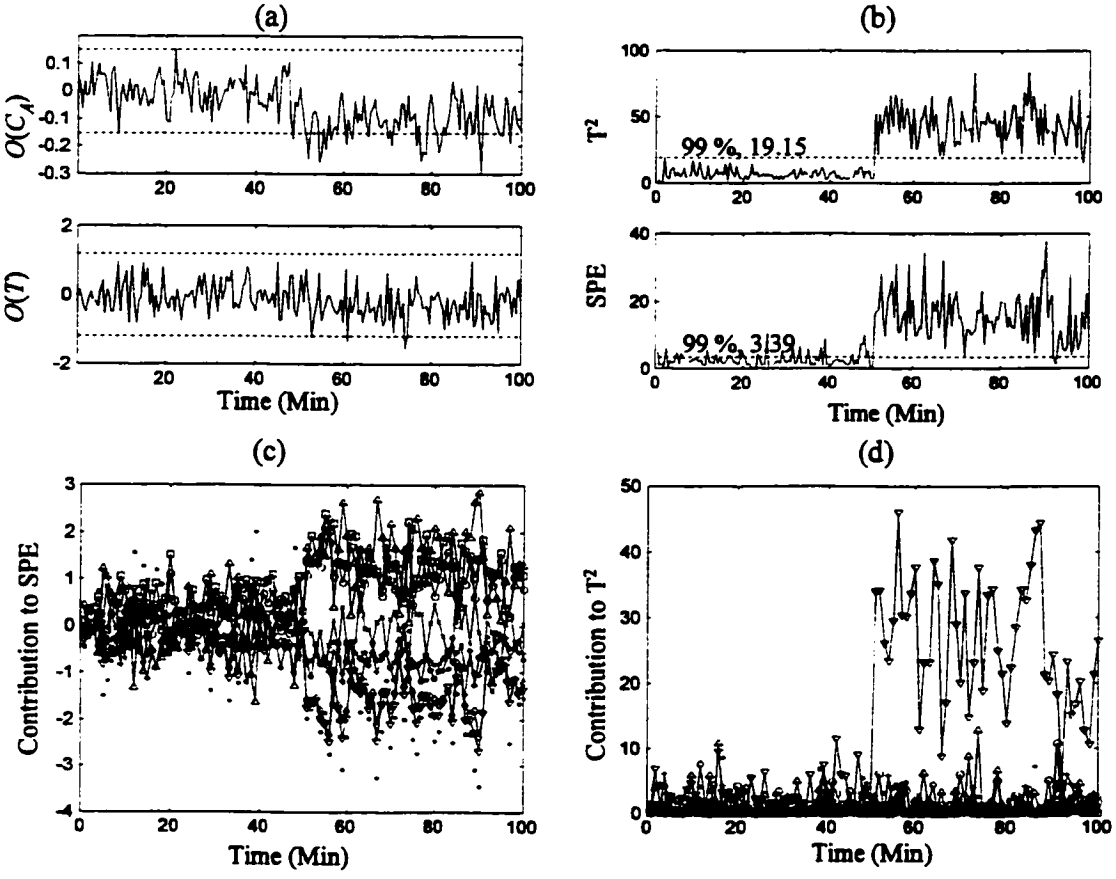


Figure 5.4 FDI of simple bias fault (ΔF_A = -0.015 at 51 min); (a) Parity relation plots; (b) SPE/T^2 plots; (c) SPE contribution plot; (d) T^2 contribution plot (C_A:O; T:✱; C_{AS}:△; C_{AA}:◇; F_S:●; F_A:▽; T_O:+; T_C:×; F_C:□)

relations) to specifically isolate the F_A fault also does not help in the isolation in this case. This is because the near collinearity of the fault directions (Figure 5.2) leads to insufficient sensitivity and a magnification of the noise and disturbance components.

The SPE/T^2 plots, based on the PCA model developed from normal plant operating data, are shown in Figure 5.4(b). Both the SPE and T^2 plots clearly detect a fault occurring around 51 minutes. (Note that it is only important that one of the SPE or T^2 plots alarms in order to detect a statistically significant event) The SPE contribution plot in Figure 5.4(c) provides no isolation information since many of the variables show increased contributions after 51 minutes. However, the T^2 contribution plot of Figure 5.4(d) clearly isolates the fault as being the F_A sensor because of the large increase in the contribution of F_A starting at 51 minutes and the negligible contributions from the other variables. The MSPC approach thus clearly detects and isolates this fault. This results from the fact that the fault in F_A significantly breaks the *common-cause* correlation structure between F_A and the other variables. Another contributing factor is that the effects of the natural disturbances in the system are accounted for more effectively in the MSPC scheme. Their effect on the covariance structure of all the variables is automatically contained in the PCA model and hence in the SPE and T^2 tests.

5.6.2 FDI of complex sensor faults

Many faults in a process are complex faults. Isolation of the complex faults usually needs the causal relationship between the root source of the fault and the affected variables. Since MSPC is not based on a causal model, it will not give a clear and unambiguous isolation of the fault. On the other hand, the causal model-based method can promptly isolate complex faults as long as they are modeled. The object of this example is to show the differences between two FDI methods for complex sensor faults.

A fault is assumed to occur at 51 minute in the sensor of the reactor outlet temperature (T), which is under closed loop control. Since the parity relations are based on causal relationships between variables that do not change with the addition of

feedback loops, the parity residuals will not be affected by the presence of the feedback controller on T. In Figure 5.5(a) this is illustrated by showing that the complex sensor fault on T is clearly detected and isolated by the parity residuals. Figure 5.5(b) shows the MSPC monitoring plots for the same fault. The abnormality is detected more slowly since it takes time for the feedback controller to propagate the effect of the fault in T into other variables. The corresponding contribution plots, Figure 5.5(c) and 5.5(d), indicate that the coolant flow rate (F_C) and the reactor outlet concentration (C_A) contribute most to the detected fault. This result occurs because the sensor bias in the measured value of the controlled variable (T) is eventually eliminated by the PI control action using the manipulated variable, namely the coolant flow rate (F_C). The change of the coolant flow

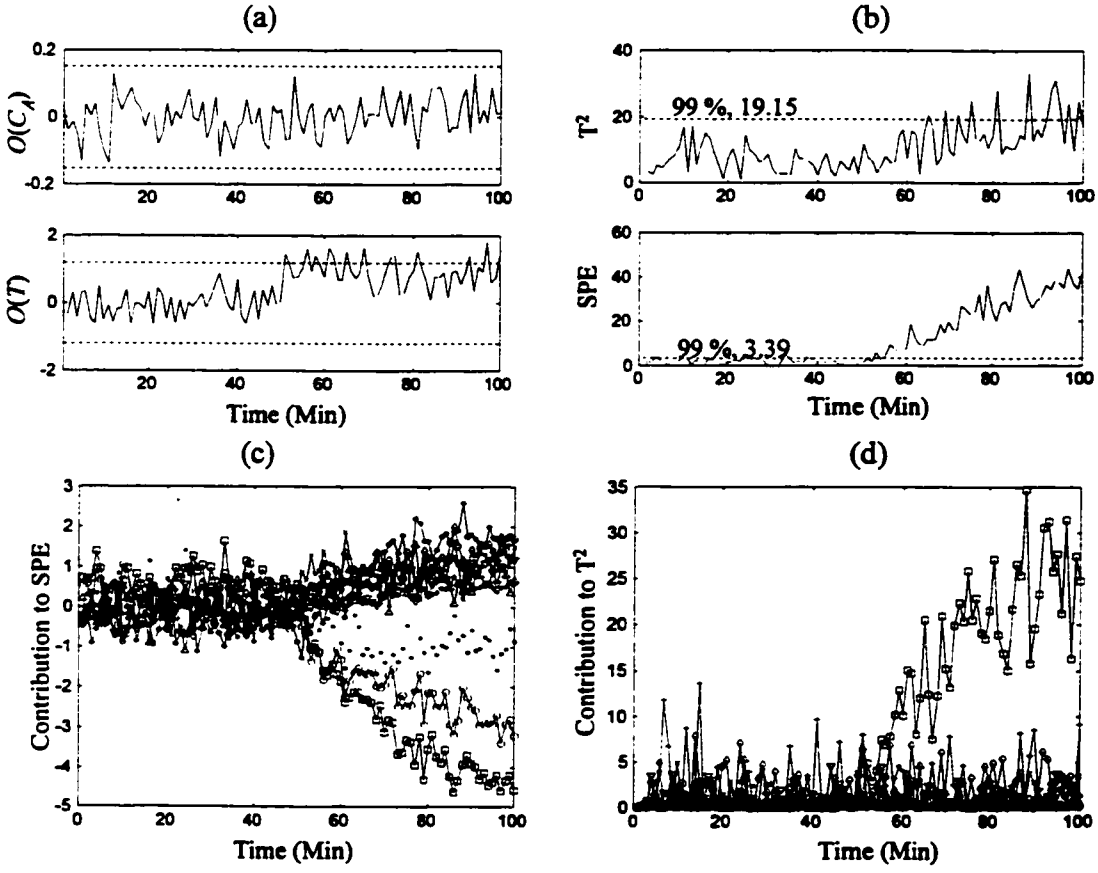


Figure 5.5 FDI of complex fault (ΔT=-1°C at 51 min); (a) Parity relation plots; (b) SPE/T^2 plots; (c) SPE contribution plot; (d) T^2 contribution plot (C_A:○; T:✱; C_{AS}:△; C_{AA}:◇; F_S:●; F_A:▽; T_O:✚; T_C:✕; F_C:□)

rate then lowers the rate of reaction resulting in an increase in the reactor outlet concentration (C_A).

This example illustrates the difficulty in isolating complex faults with the MSPC approach. The contribution plots only show how the covariance structure that existed under normal operation has been broken when a fault occurs. In this case, they reveal that the relationship among nearly all variables has changed somewhat, but for the reasons explained above, the relationship of the other variables with C_A and F_C have changed the most. To more clearly isolate the fault using the MSPC approach causal knowledge or past fault information on the effect of a fault in T could be used (Yoon and MacGregor, 2000).

5.6.3 FDI of process faults

Process faults look like complex sensor faults because they affect many variables and make them difficult to diagnose. The ability to handle process faults is an important property of FDI methods. In this example, it is shown how the two methods work for a process fault and what additional steps are required.

In this study, a sudden degradation of the heat exchanger performance is simulated as a decrease in the heat transfer coefficient at 51 minutes. However, the causal model does not include the equation that relates the heat exchanger fouling to the process measurements, or the other parameters in the model. Figure 5.6(a) shows how the causal model-based FDI works for this unmodeled process fault. The causal model-based method does detect that a fault has occurred. The $O(C_A)$ residual increases, but not enough to violate the limit. However, the temperature residual $O(T)$ drops and violates its lower limit at many points. This detection results from the fact that, with the change in heat transfer coefficient, the parity equations no longer exactly hold. However, without a model for the effect of heat exchanger fouling on the system, isolation is not possible. In order to provide the diagnostic capability on a parametric fault, one would have to generate additional parity relations to incorporate terms for all suspected process faults

using special design procedures (Gertler and Kunwer, 1995). Although any number of process faults can be incorporated into the model the number of independent fault directions (columns in **F**) and independent measurements will limit the number of faults that can be isolated.

Figure 5.6(b) shows the SPE/T^2 plots for the same process fault. The corresponding contribution plots, Figure 5.6(c) and 5.6(d), show that the main contributor to the T^2 and the SPE deviations is the coolant flow rate (F_C). This results from the fact that a drop in the heat transfer coefficient causes the temperature controller to increase F_C in order to

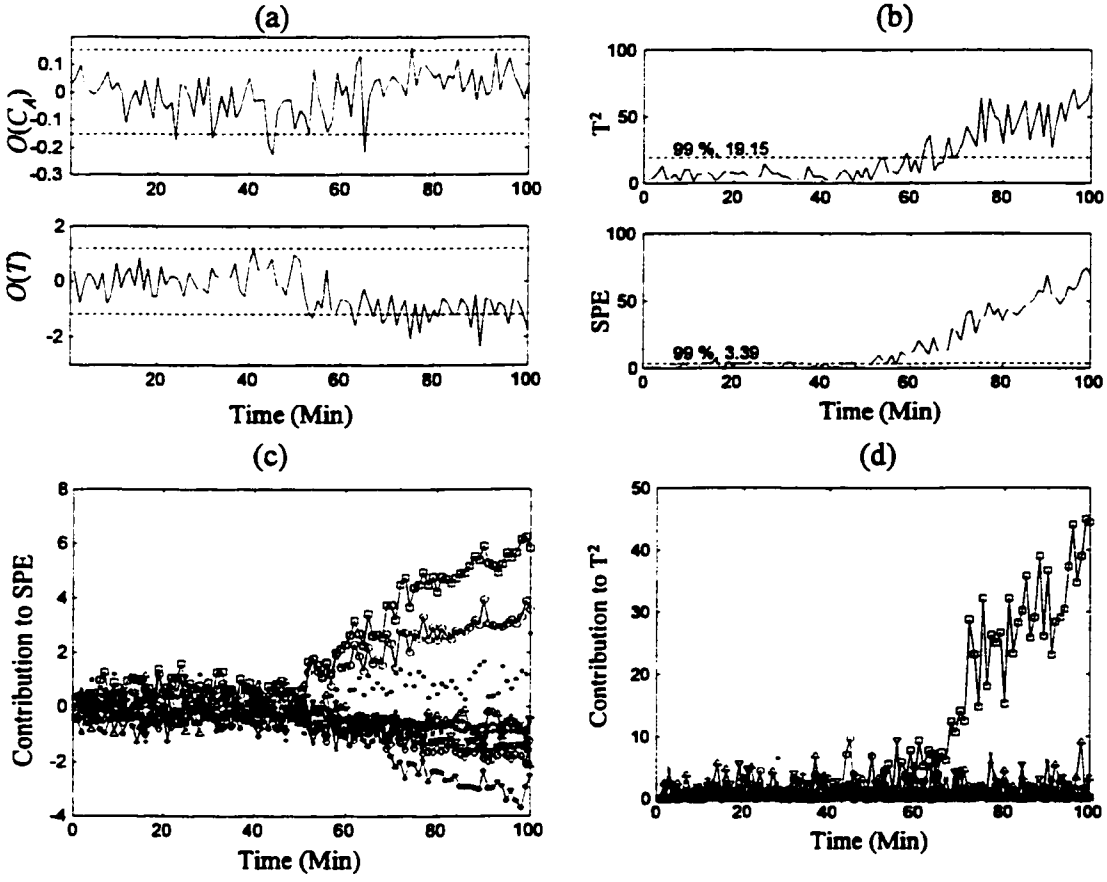


Figure 5.6 FDI of process fault (Increase in heat exchanger fouling, $\Delta UA = -0.3$ at 51 min); (a) Parity relation plots; (b) SPE/T^2 plots; (c) SPE contribution plot; (d) T^2 contribution plot (C_A:○; T :✱; C_{AS} :△; C_{AA} : ◇; F_S :•; F_A :▽; T_O :+; T_C :×; F_C :□)

keep the reactor outlet temperature (T) at its setpoint. Most of the remaining variables are unaffected and hence the contribution plots show that it is mainly F_C whose covariance structure differs from what existed during normal operation. Obviously, the isolation of the process fault is ambiguous since this same pattern could easily have resulted from a simple actuator or sensor fault on F_C. However, the MSPC approach clearly detected the unexpected process fault and the contribution plots clearly reduced the number of fault sources that would have to be investigated. One way to enhance the isolability of process faults is to use fault signatures based on past data as discussed earlier.

5.6.4 Effects of data and identification method on model-based FDI

The purpose of this example is to illustrate the effects that the model identification approach used to obtain the parity relations can have on the behavior of the causal model-based FDI scheme. In one case, an equation error identification method (Ljung, 1999) is used to obtain the explicit parity model. This simply corresponds to using a least squares method on the residuals of the parity equations to estimate the causal model parameters. The residual equations are then obtained from these. In the other case, the last principal component approach (Gertler *et al.*, 1999) is used to directly obtain parity relations for residual generation. In all comparisons, the same data is used and generated using random binary sequences (RBS) in all inputs. To simplify the study only two of the variables (F_C and F_A) were used as inputs. The steady-state gains between these inputs and outputs are only estimated after data is mean-centered and auto-scaled. Then the real gains are obtained by unscaling the identified parameters and compared against those using an exact linearization of the true fundamental model equations. The performance of the FDI scheme whose parity relations are formulated with the identified gains will be compared with one another and with that of the MSPC approach.

Figure 5.7 compares the process gains estimated by the two approaches for a wide range of magnitudes used for the RBS's. The RBS magnitudes range between 0.2 and 4.0σ where σ is the standard deviation of the inputs during normal operation period. Each set of gains is estimated with 400 measurements. For large magnitude RBS's the signal to

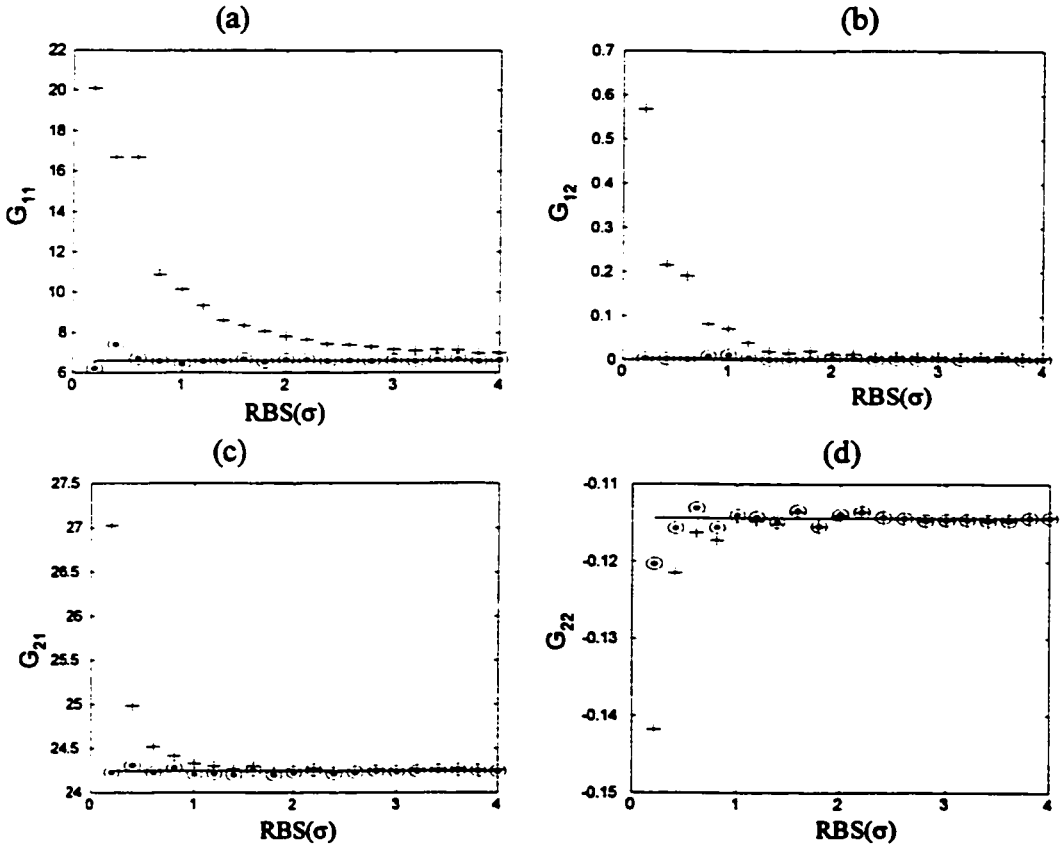


Figure 5.7 Process gain estimations by least squares method and last principal component method with full rank RBS as inputs but noisy outputs; (a) G_{11}; (b) G_{12}; (c) G_{21}; (d) G_{22} (Least Squares Method:⊙; Last Principal Components Method: +; Solid lines represent the true gains)

noise ratio is large enough that both methods identify the correct gains (shown by the solid lines), but at lower signal to noise ratios the last principal component approach provides very poor model compared to the least squares method. This confirms the results shown by Negiz and Cinar (1997).

To illustrate the impact on the FDI scheme, Figure 5.8(a) and 5.8(b) contrasts the residual plots and thresholds for the parity relations based on the models identified using the least squares on the parity equations and the last principal component method for a RBS of magnitude 0.4σ. The models identified by the least squares provided almost identical behavior to that of the theoretical model shown in Figure 5.3(a) in clearly

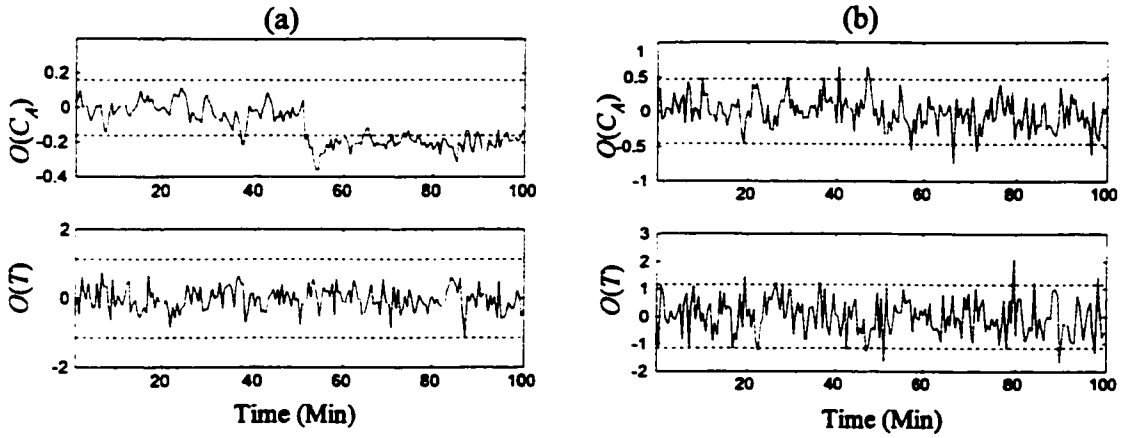


Figure 5.8 FDI performance of parity relation approach with mechanistic models and identified models for a bias fault of C_A measurement ((ΔC_A=0.2); (a) FDI with the identified model under RBS magnitude of 0.4σ and least square regression; (b) FDI with the identified model under RBS magnitude of 0.4σ and by using last principal components associated with the smallest two eigenvalues (σ: standard deviation of input signal during normal operation)

detecting and isolating the C_A sensor fault. The parity equations based on the last principal components failed to even detect the fault. The poorly identified parity relations resulting from the latter approach also leads to much larger residuals when no fault is present and hence the need for much wider threshold limits in the FDI scheme. Table 5.2 compares the FDI thresholds (99%) required for the C_A parity relation, $O(C_A)$ identified by the different approaches as a function of the RBS magnitude used in generating the data. Clearly, these limits get very large for the models from the last principal component method when the signal to noise ratio in the data is low, while they remain almost unchanged and close to those for the theoretical model when the models from the least squares identification are used. This example illustrates that the last principal component method of identifying residual relationship must be used with great caution, especially when there is measurement noise and disturbances present in the data used for identification.

The FDI results for the same fault in C_A (ΔC_A = 0.2) using the MSPC approach were previously shown in Figure 5.3(b), where the fault was also clearly detected at 51

Table 5.2 99% thresholds for C_A parity relation, $O(C_A)$ from models identified by least squares and last principal component method for different RBS magnitudes used during the identification (Threshold value based on the mechanistic model is 0.156)

RBS Magnitude	Least Squares	Last PCs
0.2σ	0.1612	4.0925
0.4σ	0.1483	0.4629
0.6σ	0.1557	0.3262
0.8σ	0.1498	0.2190
1.0σ	0.1659	0.2121

minutes in the SPE plot and isolated by using the SPE/T^2 contribution plots.

5.7 Conclusion

The fundamental differences between the causal model-based FDI approaches (as represented by the parity relation approach) and the multivariate statistical process control approaches based on non-causal models developed from historical data using PCA/PLS have been examined. The differences in the nature of the models used (arising from the nature of the data used to build them) are shown to be responsible for the very different approaches to fault detection and isolation employed by the two methods. This model difference is also shown to be responsible for their major strengths and weaknesses.

The multivariate statistical approaches are much easier to develop because of the ready availability of routine operating data, and they can handle a very large number of measured variables. They are capable of detecting almost any type of fault. Their major weakness lies in the ambiguous isolation of faults arising from the non-causal nature of the models. On the other hand, the causal model-based FDI methods can both detect and isolate faults as long as causal models are developed for all the variables and the fault structures identified *a priori*. However, this need for theoretical or identified causal

models makes the approach more difficult to develop and usually limits it to smaller systems. A more complete comparison was made under a number of different headings (Table 5.1). Some of these differences and the complementary strengths and weaknesses of the methods were illustrated using a simulated CSTR process.

Because the approaches have such different, but complementary strengths, a potentially fruitful area of research involves ways of combining them in a manner that utilize the strengths of both. We hope that by pointing out major differences between these approaches and discussing their strengths and weaknesses, this chapter will serve to foster this research.

6. Incorporation of external information into multivariate statistical models

6.1 Introduction

Multivariate statistical process control (MSPC) methods provide an alternative to the causal-model-based fault detection and isolation (FDI). Using the correlation models, one can summarize the *common-cause* variations with a few principal components. Process operation data for building correlation models is collected only when *common-cause* variations are present. The *common-cause* variations may include changes in the operating conditions such as feed conditions, product grades, controller status, and so on. When these operating conditions change, they affect many other variables inducing variations into the process data. These variations are usually described with a few principal components. This correlation model, which mainly explains the operating condition changes and the related major variations, is likely to be insensitive to faults if the sufficient number of principal components is not used. When the faults are correlated or their magnitudes are small compared to these operating variations, the correlation model may lead to poor fault detection and isolation.

However, one can avoid or minimize this difficulty by preprocessing data. For example, process variation caused by feed flowrate changes can be removed by regressing out the feed flowrate effect from the operation data. The same concept can be used to remove any disturbing effect only if its measurement is available. Note that these measurements are a type of prior knowledge not used by the conventional correlation models. On the other hand, one may have causal models for parts of the process even

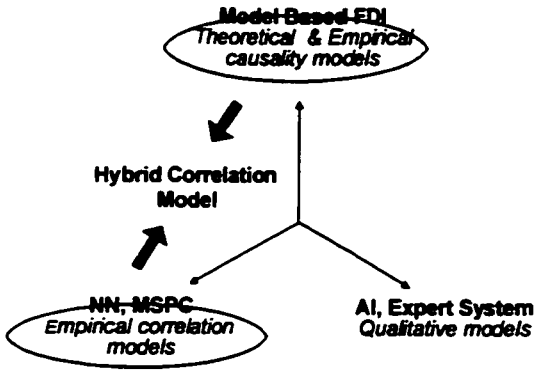


Figure 6.1 Hybrid correlation modeling

when complete causal models are unavailable. Even partially available causal models combined with the correlation models may be useful in interpreting questionable events occurring in the process. Therefore, an approach that combines available causal process information with correlation models would offer significant benefits and make it possible to avoid or minimize the limitations of the correlation models.

These methods of incorporating prior knowledge into the multivariate statistical correlation models define hybrid correlation modeling as shown in Figure 6.1. This chapter examines various types of prior knowledge used for the hybrid correlation models, various methods of incorporating external information into the correlation models, and process monitoring and fault diagnosis using the hybrid correlation model. The focus is on integrating available resources under an MSPC framework, and using the hybrid correlation model for FDI.

An outline of the chapter is as follows. In section 6.2, various types of prior knowledge are discussed and the existing correlation methods using prior knowledge are reviewed. In the following sections, several hybrid correlation methods are presented: the decomposition based method; the augmentation based method; blocking and transformation based method; the hybrid model which combines theoretical and empirical relationships. The existing correlation methods using prior knowledge are assessed from the viewpoint of the proposed methods. The key features of the hybrid correlation modeling are illustrated and its potential applications are addressed. In section 6.6, a real

industrial data is analyzed using the proposed scheme.

6.2 Preliminaries

A common feature of the hybrid correlation approaches is to use various types of external information to enhance the interpretability of a correlation model. In this section, the different types of hybrid model structures and corresponding FDI methods are briefly addressed. Then, types of prior knowledge and external information that can be used for the hybrid correlation models are addressed.

6.2.1 Hybrid model structures

When one has mechanistic models of the process, correlation models may be used in parallel or as a complementary scheme to the models. Figure 6.2 shows several hybrid model structures that combine the mechanistic models and the empirical correlation models to describe processes. Figure 6.2(a) shows an approach referred to as serial hybrid model structure. The correlation empirical model is tandem with the mechanistic models, which have a fixed structure derived from mechanistic models with parametric uncertainty. The empirical model component estimates the unknown parameters which are fed into the mechanistic equations. In a parallel approach as shown in Figure 6.2(b), the outputs of the empirical correlation and mechanistic models are combined to determine the total model outputs. The empirical correlation models are trained on the residual between the data and the mechanistic model to compensate for any uncertainties that arise from the inherent process complexity. Thus, a parallel structure of the hybrid model allows the empirical model component to capture the model mismatch between the process data and the predictions based on the mechanistic models. In Figure 6.2(c) which explains the portion of the data that is left unexplained by the mechanistic models, the residual of the mechanistic model is fed as an additional process input to the empirical

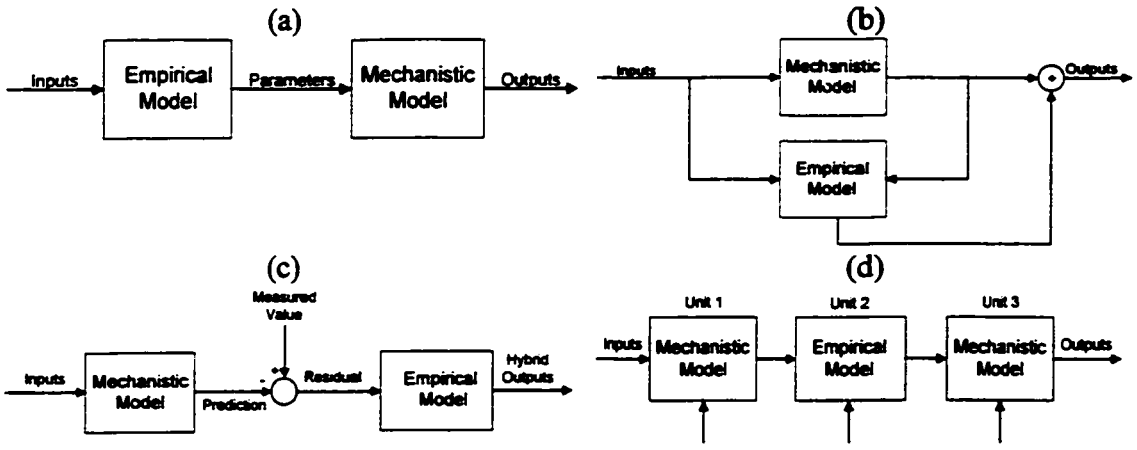


Figure 6.2 Hybrid model structures; (a) Serial approach (type 1) forces the output to be consistent with mechanistic model; (b) Parallel approach uses mechanistic model as a guide to assist the empirical model; (c) Serial approach (type 2) uses empirical model on residual from theoretical model; (d) Serial approach (type 3)

model. Building an empirical model on the residual then allows one to model and monitor those parts not explained by the mechanistic models. Wachs and Lewin (1998) presented a model based PCA that accounted for the nonlinearity of a process. The residuals estimated from the differences between the measured values and the calculated values based on the mechanistic models were used for correlation modeling. They addressed the fact that the model based PCA could deal with parametric and structural uncertainty of the models. Rotem *et al.*(2000) used the model based PCA for fault monitoring of an ethylene compressor which operates under a significant periodic disturbance caused by the ambient temperature. The serial structure shown in Figure 6.2(d) could be used for processes that are only partly understood (Thompson and Kramer, 1994).

The integration of the mechanistic models and the empirical causal models has been an important issue because of its potential benefits. One can avoid the time-consuming development of full mechanistic models for poorly/partly understood processes or non-linear parts of processes by using all the *a priori* process knowledge available (Tulleken, 1993). This concept has been popularly used by grey-box models by

which model structure is mainly mechanistic but is augmented with stochastic elements, empirical relations and parameters.

The combination of mechanistic models and neural networks (NN) has received significant attention in the literature. Johanson and Foss (1992) developed a hybrid model that used NN only in the operation regimes where the mechanistic model is not accurate enough. Thompson and Kramer (1994) discussed the serial and parallel hybrid structures used for integrating different prior knowledge with NN. *A priori* mechanistic model is used as a starting point. An empirical correlation model compensates for the mismatch between the prediction of the inaccurate mechanistic model and the process data. The model is expressed in terms of equality and inequality constraints. These constraints arise from mass and energy balances, and physical restrictions imposed on processes by equipment limitations. Anderson *et al.* (2000) applied this approach for the control of wastewater systems. A different approach was taken by Psichogios and Ungar (1992) that uses NN to estimate parameters of the mechanistic models. A mechanistic model structure containing some unknown internal variables is assumed to be available. The typical unknown internal variables are reaction rates or thermodynamic properties that depend on a number of other model variables in a complex manner. NN is used as an empirical model of these dependencies. Su *et al.* (1992) integrated recurrent NN with the mechanistic models to fit unmodeled process dynamics. Wilson and Zorzetto (1997) used a hybrid model for a generalized technique of the on-line state estimation. In their method, NN elements were exclusively applied to modeling of more complex non-linear rate relationships.

6.2.2 Hybrid fault detection and isolation

Little work has been done for FDI based on the hybrid models. Becraft and Lee (1993) used an integrated NN and expert system (ES) approach for fault diagnosis. Isermann (1993) combined a parameter estimation technique with heuristic knowledge processing for fault diagnosis. Zhang *et al.*, (1995) compared several knowledge based

methods with a CSTR model. Gertler and McAvoy (1997), and Gertler *et al.* (1999) proposed an isolation enhanced PCA. Yoon and MacGregor (2000a) compared the strengths and weaknesses of the statistical and causal model based FDI approaches. Based on the comparison, it is argued that the requirements of hybrid models and FDI methodology are to minimize the limitations caused by unavailability of mechanistic models, and to obtain the best attainable FDI performances for faults in processes.

The basic concept of the hybrid FDI is to choose a method with which one can achieve the best attainable FDI performance with limited resources. The hybrid FDI can be implemented in three ways: the causal model based methods, the multivariate statistical methods, and the combined methods as main FDI frameworks. When a mechanistic model is unavailable only for a part or unit of a process, one can use an empirical correlation model instead of the causal empirical model for the part or unit. One needs to be careful when using this combination because one obtains a model only for the correlation structure that was present when the data was collected. When most of the mechanistic models are unavailable, one has to use the correlation model based methods for FDI. When complete mechanistic models and historical fault data are available, either of the two methods can be used as a main FDI framework. Or both methods can be used complementarily to each other. For instance, the first step of the fault isolation is done by the multivariate statistical methods and the next detail step by the causal model based methods.

6.2.3 Types of prior knowledge/external information

Since the main resource for the multivariate statistical models is process operation data, any types of prior process knowledge or information readily available in a process are defined as external information which is distinguished from the process operation data used for PCA/PLS. The external information thus includes partly available mechanistic models, detailed information on process data, and so on. However, qualitative knowledge that is the main modeling resource of Artificial Intelligence

(AI)/Expert system is excluded here. The external information available at different levels includes:

- Prior knowledge on variables: *i.e.* which are disturbances, which are independent manipulated variables, and so on,
- Additional information such as controller setpoints, mode, controller output and so on
- Mass and energy balances that may be partially or fully known,
- Real relations among variables: for example, the fact that mole fractions to sum to one, the requirement of non-negative inputs and outputs, etc.,
- Calculated parameters such as reaction kinetic parameters, heat exchanger fouling factors,
- Spectra of pure chemical species,
- Process structure, or frequency characteristics of process variations.

6.3 Incorporation of external information by augmentation of X matrix

A simple approach to incorporate prior knowledge into the correlation model is to augment key calculated variables to the observation matrix and estimate the correlation model with the augmented matrix, *i.e.*,

$$\mathbf{X}_{Aug} = [\,\mathbf{X} \mid \mathbf{X}_C\,], \tag{6.1}$$

where, $\mathbf{X}$ and $\mathbf{X}_C$ are the columns of the original observation and the calculated quantity, respectively. Nomikos and MacGregor (1994) included the estimates of the total conversion and the instantaneous rate of energy release into the observation matrix, and estimated improved multiway principal component analysis (MPCA) models for batch processes. Those calculated variables, whose direct measurements were unavailable, but

which would be calculated from other variables using mass and energy balance equations, were shown to be dominant contributions in the first two principal components. Similarly, Hodouin *et al.* (1993) augmented the observation matrix from mineral flotation process with additional calculated variables such as a flotatability index, etc. Some of them were direct combinations of raw measurements, and other variables were not measured but calculated. They mentioned that this approach enriched the data with prior knowledge, but it could also have drawbacks. First, prior knowledge can distort the data if the underlying assumptions for the calculated variables are not verified. Second, because of the redundancy of the raw data, there are a number of different paths to calculate the created variables. Thus, preliminary data reconciliation was required to remove the inconsistency of the calculated variables.

The same method can be used when a stochastic variable such as the heat exchanger fouling factor, or reaction impurity, is estimated from the available mechanistic models. This approach cannot only overcome the difficulty in isolating multiplicative faults, but also minimize the necessity of process expertise in interpreting the contribution plots of the multiplicative faults. This is because the corresponding contribution of a more meaningful multiplicative parameter is highlighted instead of a group of measurements for the parametric fault. Controller output or setpoint values can also be augmented to observation matrix and the PCA model is estimated with the augmented matrix. The resulting model may have better diagnostic properties to test controller and actuator problems.

6.4 Using knowledge on process structure and signal frequencies

In very large processes involving several processing units with many variables in each unit, the number of potential errors or faults can be very large, making the diagnosis more difficult. Schemes for process monitoring using multivariate statistical projection methods such as PCA and PLS can be extended to situations where the processes can be

naturally blocked into several subsections (MacGregor *et al.*, 1994). The multiblock projection methods allow one to establish monitoring charts for the individual process subsections as well as for the entire process. The main advantage of such blocking is to allow for easier interpretation of the data by looking at smaller meaningful blocks and the relationship between blocks. The choice of blocking depends upon engineering judgment (prior knowledge). The block should correspond to distinct units of the process where all the variables within a block or process unit may be highly coupled, but where there is minimal coupling among variables in different blocks. Variables associated with streams that leave one block and enter another (feed or recycle streams) should generally be included in both blocks.

Similarly, one can decompose the process variations over several frequency (scale) ranges and estimate the PCA models for each scale, or all scales together. For detail, refer to chapters 3 and 4. Frequency characteristics of the faults are very crucial for the implementation of multiscale decomposition approaches. This frequency information constitutes prior knowledge that can lead to enhanced FDI models.

When either the frequency contents of process variations or information on the process blocking is available, one can split or decompose the observation matrix, and estimate the corresponding correlation model. Such multiscale and multiblock PCA models can provide better performances of the fault detection and isolation.

6.5 Decomposition of X matrix with row and column constraints

Principal component analysis (PCA) is often used to explore structures in multivariate data. Simple PCA on the data matrix ($\mathbf{X}$) may not be an appropriate method to apply when additional information about variables and observations is available. The external information can be heuristically used to aid process trend analysis. Alternatively, the external information can be directly incorporated into the formal analysis. The original data matrix can first be decomposed into several components: those that can be

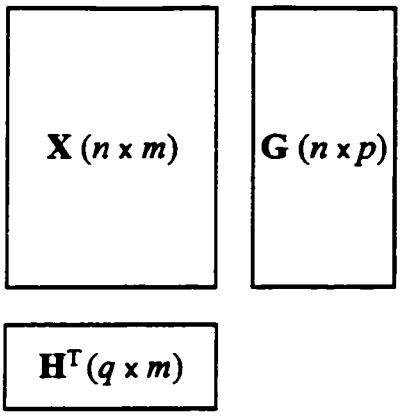


Figure 6.3 Row and column information

explained, and those that cannot be explained, by the external information. PCA is then applied to each component separately or to the combined components (Takane and Shibayama, 1991).

The above approach can have significant advantages over the regular correlation modeling methods since it allows for the activation of a larger portion of the available prior knowledge. Takane and Shibayama (1991) described a method of including external information on the observations (or objects) and/or variables in a bilinear model for the two-way data. The key concept of the approach is the separation of sources of variation which brings greater clarification as to the role of the external information within the measured data. Using the method proposed by Takane and Shibayama, Gurden *et al.* (2001) incorporated external information on the spectroscopically active compounds into the latent components in the form of constraints. The total variations in the observation matrix were decomposed into three parts: one part of the model reflects known causes; the second part describes the systematic variation that is not explained by the known causes; the third part is the residual variation. They claimed that one might improve the interpretability of the model, examine the relative influence of sources of process variations, and increase the numerical stability of the modeling algorithm.

Various types of the external information can be represented in the form of either a column or a row associated with the observation matrix, **X**, as shown in Figure 6.3.

Given the data matrix **X** ($n \times m$), assume an n by p ($\leq n$) observation information matrix, **G**, and m by q ($\leq n$) variable information matrix, **H**. For example, **G** may be an n component vector indicating class membership of each object in **X**, a matrix of dummy variables (0's and 1's), or a matrix of continuous variables such as production rate characterizing the observations. **H** may contain m-component vectors of coefficients that capture relationships among the columns of **X**.

6.5.1 Decomposition of data matrix X with prior knowledge

Provided with **G** and **H** in addition to **X**, the data matrix can be represented as follows;

$$\mathbf{X} = \mathbf{GMH}^{\mathrm{T}} + \mathbf{BH}^{\mathrm{T}} + \mathbf{GC} + \mathbf{E}, \tag{6.2}$$

where the data may be raw, or preprocessed by standardization or other transformations. **M**($p \times q$), **B**($n \times q$), and **C**($p \times m$) are matrices of coefficients to be estimated, and **E**($n \times m$) a matrix of error components. The four terms in (6.2) explain the decomposed components of the original data matrix, **X**. The first term pertains to what can be explained by both **G** and **H** combined, the second term by **H** only, the third term by **G** only, and the fourth term by neither **G** nor **H**. The first three terms are explainable terms by prior knowledge while the last term is inexplicable. The least squares estimates of **M**, **B**, and **C** are;

$$\hat{\mathbf{M}} = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{XH}(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \tag{6.3}$$

$$\hat{\mathbf{B}} = (\mathbf{I} - \mathbf{P}_G)\mathbf{XH}(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \tag{6.4}$$

$$\hat{\mathbf{C}} = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{X}(\mathbf{I} - \mathbf{P}_H), \tag{6.5}$$

where $(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}$ and $(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}$ are pseudo inverses of $(\mathbf{G}^{\mathrm{T}}\mathbf{G})$ and $(\mathbf{H}^{\mathrm{T}}\mathbf{H})$, respectively. $\mathbf{P}_G$ and $\mathbf{P}_H$ are orthogonal projection operators onto spaces spanned by the column vectors of

G and **H**. They are calculated as follows;

$$\mathbf{P}_G = \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T, \tag{6.6}$$

$$\mathbf{P}_H = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T. \tag{6.7}$$

Using orthogonal complement projectors of $\mathbf{P}_G$ and $\mathbf{P}_H$, $\mathbf{Q}_G = \mathbf{I}-\mathbf{P}_G$ and $\mathbf{Q}_H = \mathbf{I}-\mathbf{P}_H$, the fourth term in (6.2) is simplified as follows:

$$\begin{aligned}\hat{\mathbf{E}} &= \mathbf{X} - \mathbf{G}\hat{\mathbf{M}}\mathbf{H}^T - \hat{\mathbf{B}}\mathbf{H}^T - \mathbf{G}\hat{\mathbf{C}} \\ &= \mathbf{X} - \mathbf{P}_G\mathbf{X}\mathbf{P}_H - \mathbf{Q}_G\mathbf{X}\mathbf{P}_H - \mathbf{P}_G\mathbf{X}\mathbf{Q}_H \\ &= \mathbf{Q}_G\mathbf{X}\mathbf{Q}_H\end{aligned}. \tag{6.8}$$

By substituting the least squares estimates for the corresponding parameters in (6.2), the data matrix, **X**, is decomposed as follows:

$$\begin{aligned}\mathbf{X} &= (\mathbf{P}_G + \mathbf{Q}_G)\mathbf{X}(\mathbf{P}_H + \mathbf{Q}_H) \\ &= \mathbf{P}_G\mathbf{X}\mathbf{P}_H + \mathbf{Q}_G\mathbf{X}\mathbf{P}_H + \mathbf{P}_G\mathbf{X}\mathbf{Q}_H + \mathbf{Q}_G\mathbf{X}\mathbf{Q}_H\end{aligned}. \tag{6.9}$$

The four terms in (6.9) are the estimates of the corresponding four terms in (6.2). Note that some of the terms in (6.9) may be zero when **G** or **H** is a square matrix of full rank (e.g., $\mathbf{G} = \mathbf{I}$ or $\mathbf{H} = \mathbf{I}$).

Once the data matrix is decomposed according to the external information, PCA may be applied to each component separately. In certain cases, some of the decomposed submatrices may be recombined for PCA. For example, the first and second terms in (6.9) may be combined, amounting to PCA of $\mathbf{X}\mathbf{P}_H$. Since the components analyzed are associated with specific meanings of their own, PCA of the components may be more readily interpretable than direct PCA of the original data matrix. In what follows, (6.9) is used as a general framework for hybrid correlation modeling.

6.5.2 Prior knowledge in the form of columns: $\mathbf{X} = \mathbf{P}_G\mathbf{X} + \mathbf{Q}_G\mathbf{X}$

If a given measured variable is having an undesired and disturbing influence on the process, its influence can be removed from **X** by data laundering through target rotation (Christie, 1996). The specific variable is referred to as the target variable. The model matrix, $\mathbf{P}_G\mathbf{X}$ reproduces the information in **X** specifically connected to the target variation. As the residual matrix is the difference between the model matrix and the observation matrix, the residual matrix contains the remaining information, which is obtained by subtracting the information specifically related to the target variable from the original observation matrix. The basic concept of this approach is to remove the effects of known and uninteresting variations from the data matrix. Christie (1996) applied this concept for oil field exploration. The residual matrix after the target rotations represented only 17% of the total variance of the standardized raw data, but it contained all the important information which meets the purpose of the study. In the following, this concept is generalized with the framework of (6.9).

When prior knowledge is available in the form of column(s) **G**, $\mathbf{H} = \mathbf{I}$, $\mathbf{P}_H = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T = \mathbf{I}$, $\mathbf{Q}_H = \mathbf{0}$ (zero matrix). Thus (6.9) becomes

$$\mathbf{X} = \mathbf{P}_G\mathbf{X} + \mathbf{Q}_G\mathbf{X}, \tag{6.10}$$

where the first term represents variation that can be explained by **G**, and the second term variation that cannot be explained by **G**, respectively. A PCA model can thus be estimated with $\mathbf{Q}_G\mathbf{X}$ rather than with the original data matrix. This PCA model excludes the variations correlated with the nuisance variables expressed by **G**. This scheme generalizes the data laundering via target rotation (Cristie, 1996).

In the data laundering, the laundered matrix is obtained by regressing/projecting out the effect of the nuisance variable on all the other variables in **X**: the regression coefficient on the target variable by regressing each column of **X** on the target variable, $\hat{\boldsymbol{\beta}} = \left(\mathbf{y}^T\mathbf{y}\right)^{-1}\mathbf{y}^T\mathbf{X}$; the laundered data matrix by subtracting $\mathbf{y}\hat{\boldsymbol{\beta}}$ from the original data matrix, $\mathbf{X}_l = \mathbf{X} - \mathbf{y}\hat{\boldsymbol{\beta}}$. The data laundering can be considered the case that the target

variable column, $\mathbf{y}$ is to be $\mathbf{G}$ in (6.9) and then $\mathbf{P}_y = \mathbf{y}(\mathbf{y}^T\mathbf{y})^{-1}\ \mathbf{y}^T$. The data matrix is decomposed into two terms as $\mathbf{X} = \mathbf{P}_y\mathbf{X} + \mathbf{Q}_y\mathbf{X}$, where $\mathbf{P}_y\mathbf{X}$ is just a different expression for $\mathbf{y}\hat{\boldsymbol{\beta}}$. The laundered data, $\mathbf{X}_l$ is nothing but $\mathbf{Q}_y\mathbf{X}$ that is a projection of the data matrix onto the orthogonal complement of the space spanned by the target variable. The data decomposition method used here can be applied to the case where there are more than one target variable. In such case, one can implement the multiple target rotation at the same time. If the target matrix $\mathbf{G}$ contains vectors of observations of many correlated variables then a least squares projection may not be desirable. Rather, a latent variable regression such as PLS may be used to project the information in $\mathbf{G}$ into a lower dimensional space, and then $\mathbf{X}$ can be decomposed as $\mathbf{X} = \mathbf{TP}^T + \mathbf{E}$ where $\mathbf{TP}^T$ is equivalent to $\mathbf{P}_G\mathbf{X}$ and $\mathbf{E}$ to $\mathbf{Q}_G\mathbf{X}$. Orthogonal signal correction in PLS (Wold *et al.*, 1998) can be understood as the case that one uses the response variable(s), $\mathbf{Y}$ as a target matrix. Then $\mathbf{X} = \mathbf{P}_Y\mathbf{X} + \mathbf{Q}_Y\mathbf{X}$. One can remove the information in $\mathbf{X}$ that is not predictive of, *i.e.*, orthogonal to $\mathbf{Y}$. The remaining step is to build the PLS model with $\mathbf{Y}$ and $\mathbf{P}_Y\mathbf{X}$.

Illustration

Combining signal components shown in Figure 6.4(a) generates an artificial signal. One can remove the effect of each component by the decomposition method. Figure 6.4(b) show the original signal (top) and the laundered signal (bottom) in which the effects of disturbance and equipment degradation are removed. It is shown that the sensor and equipment failure can be more clearly detected and diagnosed with the laundered signal than the original signal. The decomposition method can also be used for detrending. By setting C_{AA} measurements in the CSTR simulation system as $\mathbf{G}$ and decomposing $\mathbf{X}$ into $\mathbf{P}_{C_{AA}}\mathbf{X}$ and $\mathbf{Q}_{C_{AA}}\mathbf{X}$, one can remove the sinusoidal effect on all the other measurements. A PCA model is then estimated with $\mathbf{Q}_{C_{AA}}\mathbf{X}$. This PCA model does not include any effect of C_{AA}, and so this PCA model cannot be used for monitoring abnormalities related with C_{AA}. Thus, the decomposition based method is recommended only when there is obviously a variable that needs to be removed as in this example.

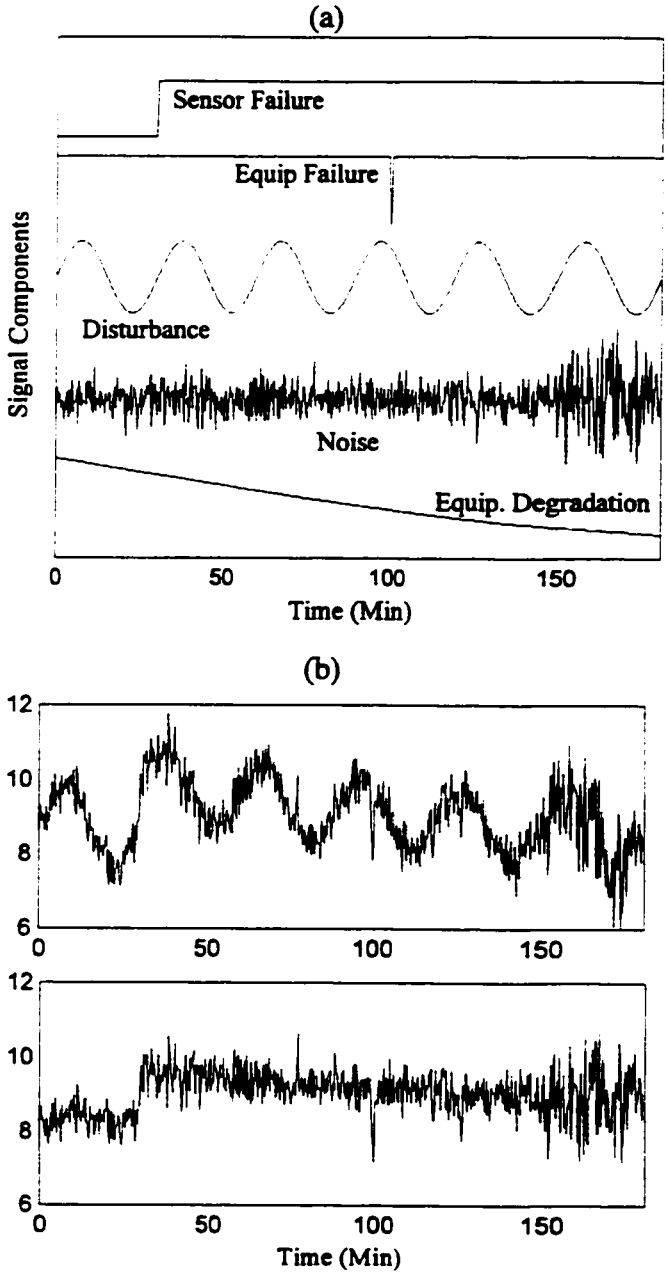


Figure 6.4 Data laundering using target projection; (a) Signal components; (b) Original vs. laundered signal without effects of disturbance and equipment degradation

Sometimes, normal operation data includes a subset experiencing different process feed or product flow. Due to the effects of these process disturbances, the

corresponding MSPC model built with this data may have a large control limit in some of the principal components. Such effects thus could degrade the sensitivity of the correlation model in monitoring the process. The data collected under different setpoint values of the controlled variable may also have similar characteristics. By setting the setpoint value as a target variable, the observation matrix is decomposed into two terms, $\mathbf{X} = \mathbf{P}_{SP}\mathbf{X} + \mathbf{Q}_{SP}\mathbf{X}$. To illustrate the concept, three datasets were generated with different setpoints of 368.3 °C, 367.3 °C and 369.3 °C on the reactor outlet temperature of the CSTR system. The PCA model was estimated with the data collected during the operation under 368.3 °C on the reactor outlet temperature. The other two sets of data were projected into the model space. As shown in Figure 6.5(a), the training sets are clustered in different locations from that of the testing data. All these datasets are part of the normal operation. When one includes all those for building a correlation model, the number of latent variables may be large in order to model all this known variation. To remove the effect of setpoint change, the vector of the setpoint values was set as **G**, and the training data was decomposed into two terms, $\mathbf{P}_{SP}\mathbf{X} + \mathbf{Q}_{SP}\mathbf{X}$. The PCA model was estimated only with $\mathbf{Q}_{SP}\mathbf{X}$. The other two datasets were also decomposed and projected onto the new PCA model. As shown in the Figure 6.5(b), the two sets of the projected

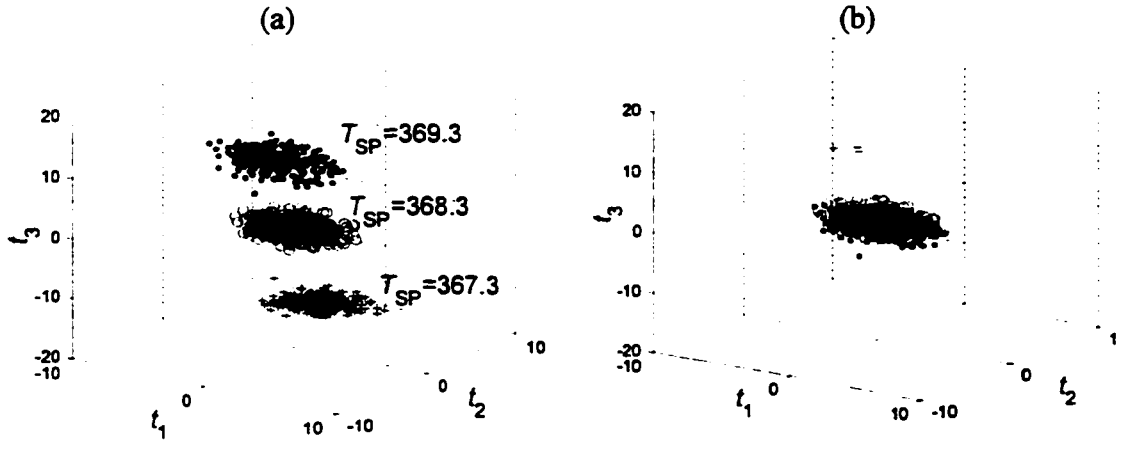


Figure 6.5 Effect of setpoint changes on reactor outlet temperature of CSTR system; (a) Data projection without laundering; (b) Data projection with laundering (Base setpoint: 368.25 °C; Setpoint changes: 367.3 & 369.3 °C)

data overlap within the ellipsoid made with the training dataset. Thus, the new PCA model obtained after the decomposition may be used for the process monitoring regardless of any future setpoint changes, as long as any future setpoint changes are also projected out.

The same concept has been used by Tembec Co. (Pulp and paper company in Canada) to deal with a large number of product grades (presented at the spring meeting of McMaster Advanced Control Consortium, 2001). More than 70 paperboard grades were classified into 9 families, and a single PCA model was developed for each of the 9 family grades. However, it is noted that this approach is limited by the assumption that the correlation structure of data in one family is consistent.

6.5.3 Prior knowledge in the form of rows: $\mathbf{X} = \mathbf{X}\mathbf{P}_H + \mathbf{X}\mathbf{Q}_H$

The key parameters used by Nomikos and MacGregor (1994), and Hodouin *et al.* (1993) were calculated from material or energy balance equations. Those equations can also be used for the generation of the residuals with which PCA model is estimated. Tong and Crowe (1995) applied this idea to data reconciliation using only the material balance equations and proposed principal component tests in identifying the variable in gross error.

Kettaneh *et al.* (1994) illustrated another type of the use of process knowledge for the multivariate design of process experiments (M-DOPE). They proposed selective-PLS algorithms that allow for the selection of important variables having large weight and the exclusion of variables having small loadings. Similarly, Roffel *et al.* (1989) reported a systematic approach which uses the PCA and the least squares regression for the selection of the most significant output variable subspace from the available measurements.

The common features of these methods assume that the relevant information on the target process variables is available. Assume row information on **X**, *i.e.* **H**, is available. Mass and energy balances belong to this type of information. Thus, $\mathbf{G} = \mathbf{I}$, $\mathbf{P}_G =$

$\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T = \mathbf{I}$, $\mathbf{Q}_G = \mathbf{0}$ (zero matrix), and (6.9) becomes

$$\mathbf{X} = \mathbf{X}\mathbf{P}_H + \mathbf{X}\mathbf{Q}_H, \tag{6.11}$$

where the first term represents the variations that can be explained by **H**, and the second term the variations that cannot be explained by **H**.

The data reconciliation proposed by Tong and Crowe (1995) uses the mass balance equations, **B**, where **B** is $r \times m$ with $r < m$. The residuals of reduced constraints for a linear steady-state process are defined in the matrix form, $\mathbf{R}=\mathbf{X}\mathbf{B}^T$. The PCA is estimated with the residuals rather than the original observation matrix. The gross errors are identified with a principal component test. The principal component test for the data reconciliation can be interpreted differently if one uses (6.9). Since the variable relationship, $\mathbf{H}=\mathbf{B}^T$ is available, the observation matrix is decomposed: $\mathbf{X}=\mathbf{X}\mathbf{P}_B+\mathbf{X}\mathbf{Q}_B$, where $\mathbf{P}_B=\mathbf{B}^T(\mathbf{B}\mathbf{B}^T)^{-1}\mathbf{B}$. Then the process variations are decomposed into two sources of what can be explained and what cannot be explained by the balance equations. The first term will explain the process variations based on the steady-state and deterministic variable relationships. The least squares projection of data onto **B** would be zero if the mass balances held exactly. The second term, $\mathbf{X}\mathbf{Q}_H$ contains all information orthogonal to mass balances, *i.e.*, all disturbances, energy terms, and so on.

The PCA provides the best fixed-rank approximation of the matrix in the least squares sense under any orthogonality invariant norm. Computationally, PCA amounts to the singular value decomposition (SVD) of a rectangular matrix. The ordinary SVD can be used if PCA is applied to the entire submatrices in (6.9). In some cases, only a portion of the decomposed submatrices may be meaningfully analyzed. In this case, the ordinary SVD does not apply and instead the generalized SVD (GSVD) is required (Takane and Shibayama, 1991). Although the ordinary SVD of **X** will find $\mathbf{X}^*$ that minimizes $SS(\mathbf{X}-\mathbf{X}^*)$, the same $\mathbf{X}^*$ does not minimize $SS(\mathbf{R}-\mathbf{R}^*) = SS[(\mathbf{X}-\mathbf{X}^*)\mathbf{B}^T]$ unless **B** is column wise orthonormal. In most cases, the GSVD of **X** with $\mathbf{B}\mathbf{B}^T$ as row metrics is required in finding $\mathbf{X}^*$ which minimizes $SS(\mathbf{R}-\mathbf{R}^*)$. The ordinary SVD of the residuals can be

approximately estimated with GSVD of **X** with $\mathbf{BB}^T$. Unless otherwise specified, **X** is not scale-invariant since the elements of **X** depend on $diag(\mathbf{BB}^T)$, but **X** is directly comparable only when the columns of **B** have comparable scales. The metric matrix allows for the comparability (Takane and Shibayama, 1991). Therefore, the PCA model used for the data reconciliation (Tong and Crowe, 1995) is nothing but a transformation of the PCA with the scaling information.

Theoretically, one can interpret the data decomposition using the balance equations (**B**) with the concept of the total PCA (Huang, 2001) that finds linear combinations of the column vectors. By splitting **X** into $\mathbf{XP}_H$ and $\mathbf{XQ}_H$, one can extract the variable dimensions corresponding to zero, or near-zero singular values from the total dimensions. All the variations not to be explained by the linear relationships will be captured with the PCA model on $\mathbf{XQ}_H$. This model can be used for the model based PCA (Wachs and Lewin, 1998) as well.

Alternatively, the relationships such as the balance equations can be used as constraints when one builds empirical correlation models as done in the grey-box (Tulleken, 1993) methodology. It would be possible to incorporate the process causality relationship into the multivariate statistical models. Burnham (1997) presented how to formulate the objective function for the estimation of the multivariate correlation models with the constraints.

Illustration

Consider the simple mass balance model used by Crowe *et al.* (1983). The process consists of three units in series, with an unknown leak in the second unit in Figure 6.6(a). Only total mass balances are considered.

$$\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \Sigma = \mathbf{I}, \mathbf{x}_{true} = \begin{bmatrix} 100 & 100 & 95 & 95 \end{bmatrix}$$

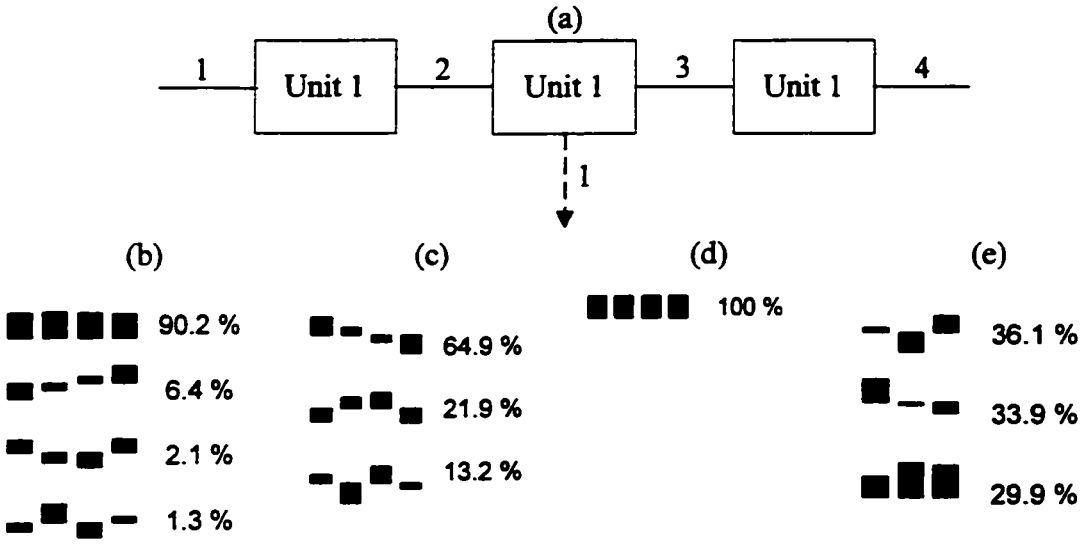


Figure 6.6 PCA model of a process with leak: (a) Process with leak; (b) Loadings for PCA(**X**); (c) Loadings for PCA($\mathbf{XP}_B$); (d) Loadings for PCA($\mathbf{XQ}_B$); (e) Loadings for PCA($\mathbf{R} = \mathbf{XB}'$) (Each cell of the loadings in (b)~(d) corresponds to x_1, x_2, x_3, and x_4; each cell of the loadings in (e) corresponds to r_1, r_2, and r_3)

With the random noises of the squared root of the variances $\sigma = diag(0.19\ 0.19\ 0.19\ 0.19)$, 100 samples were generated. Based on (6.9), **X** matrix is decomposed: $\mathbf{X}=\mathbf{XP}_B+\mathbf{XQ}_B$ where $\mathbf{P}_B=\mathbf{B}^T(\mathbf{BB}^T)^{-1}\mathbf{B}$. (b), (c), (d), and (e) in Figure 6.6 show the loadings from the PCA models of **X**, $\mathbf{XP}_B$, $\mathbf{XQ}_B$, and $\mathbf{R}=\mathbf{XB}^T$. In each figure, the number indicates the percent of variations explained by each principal component. In Figure 6.6(b), the 2-4th principal components correspond to the near-zero singular values of **X**. Those principal components explain the linear relationships among variables. As shown in Figure 6.6(c), the principal components of $\mathbf{XP}_B$ correspond to the 2~4th principal components of PCA(**X**). One can derive the following linear relationships from the principal components of $\mathbf{XP}_B$.

$$\mathbf{B}_{est} = \begin{bmatrix} 1 & -0.963 & 0 & 0 \\ 0 & 1 & -1.0018 & 0 \\ 0 & 0 & 1 & -1.0435 \end{bmatrix}.$$

The above estimated matrix confirms that $\mathbf{XP}_B$ explains the process variations based on the steady-state and deterministic variable relationships.

It is noted that the rank($\mathbf{XQ}_B$) = null($\mathbf{XP}_B$). Thus rank($\mathbf{X}$) = 4, rank($\mathbf{XP}_B$) = 3, and rank($\mathbf{XQ}_B$) = 1. PCA($\mathbf{XP}_B$) will be on the space spanned by the linear relationships, $\mathbf{B}$. PCA($\mathbf{XQ}_B$) explains all the other process variations except for the variation based on the $\mathbf{B}$-spanned space. PCA($\mathbf{XQ}_B$) will have a higher sensitivity that PCA($\mathbf{X}$) or PCA($\mathbf{R}$) with respect to abnormalities from the *common-cause* variation space.

6.5.4 Row and column information, both G and H are available

When both $\mathbf{G}$ and $\mathbf{H}$ information are available, all the terms in (6.9) become relevant. This is a common situation in industrial processes. Note that row information

Table 6.1 Hybrid Correlation Approaches

Framework	Knowledge type	Examples	Applications
Augmentation	Physical parameters Balance equations	• Batch modeling (Nomikos and MacGregor, 1994) • PCA (Houdouin *et al.*, 1993)	• Batch process monitoring
Splitting	Frequency characteristics, Knowledge on process blocking	• Multiscale PCA (Bakshi, 1998) • Multiblock PCA (Wold, 1995)	• Process monitoring and fault diagnosis
Decomposition	G (observations)	• Target rotation (Cristie, 1996) • Orthogonal signal correction (Wold *et al.*, 1998)	• Data laundering • Predictive modeling • Disturbance effects-free PCA model
	H (Variable relationships)	• Data reconciliation (Tong & Crowe, 1995) • Model based PCA (Wachs and Lewin, 1998)	• Process monitoring
	Both G and H	• Grey modeling (Gurden *et al.*, 2001)	• Robust batch modeling • Process control analysis

can be converted into column information, and *vice versa*. For example, the material balance over the variables can be represented as a residual and the calculated residual can be augmented into the observation matrix as a column, or used as G for the data decomposition. Thus, the decomposition of the data matrix and the analysis of the decomposed terms can be performed in a similar way to that in the previous two cases.

Table 6.1 summarizes several hybrid correlation approaches. The framework of hybrid modeling, knowledge type, examples, and potential application areas are listed.

6.6 Industrial application

In this section, a real industrial data is analyzed. Discussed are the issues on how to improve process monitoring and fault diagnosis while minimizing the effects of disturbances by using some of the proposed methods.

6.6.1 Process and problem description

A dataset was collected from a Styrene Monomer (SM) production plant. The full scale SM unit consists of a dehydrogenation section with styrene, product recovery system and a distillation section in which benzene-toluene, ethylbenzene (EB) and tars are separated from styrene. Ethylbenzene is dehydrogenated into styrene monomer in the presence of the iron oxide catalyst. This reaction is endothermic and the necessary heat is supplied by the addition of high temperature steam. Since there are two moles of product for each mole of reactant, the reduction of the partial pressure of the product accelerates the forward reaction. The partial pressure reduction is accomplished by reducing system pressure and by diluting the styrene monomer with steam. SHR (Steam to Hydrocarbon Ratio) is determined based on the characteristic of a catalyst. Thus, the reactor product produces about 70% of the crude SM and 28 % of the unreacted EB with the condensing

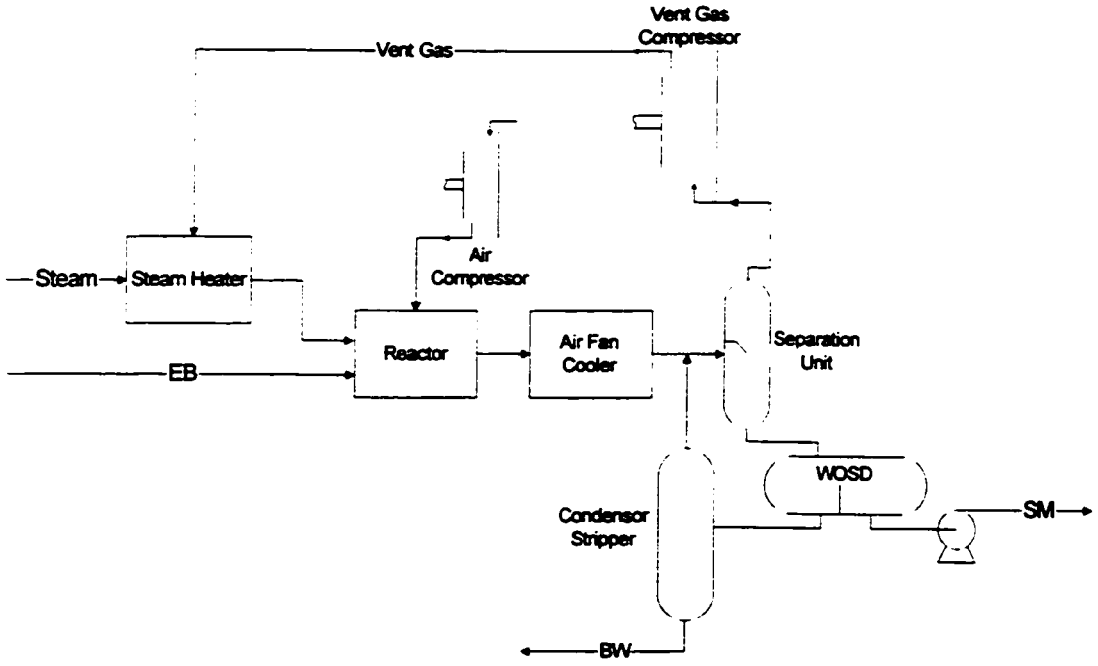


Figure 6.7 Simplified dehydrogenation unit of styrene monomer (SM) production plant

water. The condensing water from the oil/water settling drum (WOSD) is stripped, and turns into fresh water for recycling use in steam boilers. In this study, only the dehydrogenation section is considered and its process flow diagram is shown in Figure 6.7. Table 6.2 summarizes the key process variables.

Hourly average values of 105 process variables were collected over the period of 5 months from November 1999 to March 2000. The data includes both normal and abnormal operations. The plant operation was shut down due to the malfunctions of the interface level indicator of the WOSD and the SM product pump at 9 a.m. on 117^{th} day. The interface level of the WOSD had indicated approximately 1% lower value than the setpoint since 5 a.m. on the same day. The feedback control of the level controller was kept on for approximately 4 hours without operator intervention and resulted in a complete closure of the water disposal valve. In the mean time, the viscous material existing between the water and SM phases, which was considered polymer, overflowed to the SM side of WOSD. This caused a trip of the SM product pump since the viscous material blocked the pump suction line.

Table 6.2 Summary of selected process variables

Variable	Description
EB	EB flow rate
S1	Steam to 1st stage reactor
S2	Steam to 2nd stage reactor
S3	Steam to 3rd stage reactor
A2	Air to 2nd stage reactor
A3	Air to 3rd stage reactor
EBT	EB temperature
R1SiT	Steam temperature to 1st stage reactor
R2SAiT	Steam and air temperature to 2nd stage reactor
R3SAiT	Steam and air temperature to 3rd stage reactor
R1iT	1st stage reactor inlet temperature
R2iT	2nd stage reactor inlet temperature
R2oT	2nd stage reactor outlet temperature
R3iT	3rd stage reactor inlet temperature
R3oT	3rd stage reactor outlet temperature
RXo1T	Reactor outlet temperature 1
RXo2T	Reactor outlet temperature 2
RXo3T	Reactor outlet temperature 3
AFCoW1	Liquid phase temperature 1 of Air fan cooler
AFCoW2	Liquid phase temperature 2 of Air fan cooler
R1oT	1st stage reactor outlet temperature
AFCoV1T	Vapor phase temperature 1 of Air fan cooler
AFCoV2T	Vapor phase temperature 2 of Air fan cooler
StrTT	Stripper top temperature
CoT	Compressor downstream temperature
VGT	Vent gas temperature
EBP	EB pressure
R1iP	1st stage reactor inlet pressure
R1oP	1st stage reactor outlet pressure
R2oP	2nd stage reactor outlet pressure
R3oP	3rd stage reactor outlet pressure
RXoP	Reactor outlet pressure
StrTP	Stripper top pressure
CMPiP	Compressor inlet pressure
VGP	Vent gas pressure
AmbT	Ambient temperature
AFCDP	Air fan cooler delta pressure
StrWWV	Stripper water withdraw valve opening
SD1WV	Water phase valve of Separation drum 1
SD2WV	Water phase valve of Separation drum 2
WOSDOL	Oil level of Water/Oil separation drum
WOSDWL	Water level of Water/Oil separation drum

Through the troubleshooting, it was found that the malfunction of the water level sensor of WOSD was actually triggered by the accumulation of the viscous material in the level instrument. This viscous material is usually generated during the condensation of the reactor products at a high temperature. It was understood that looking at the pressure difference between the reactor outlet and the recycle stream from the condenser stripper could monitor the viscous material content. However, the indicator of the pressure difference was not shown sensitive enough to detect the polymer accumulation due to the noisy characteristic of the pressure sensor and the disturbance effects. Accordingly, the operator intervention was not properly made at an early stage of the abnormal process operation.

6.6.2 Fault detection and diagnosis using regular PCA model

Among 105 variables, 45 key variables were selected for the postmortem analysis. A few outliers occurring during routine maintenance period were removed and treated as missing data. The last several points were collected when the plant was experiencing the pump failure of the SM product line and thus most measurements showed unreasonable values. These measurements were also removed from the testing data set.

Table 6.3 Summary of PCA and hPCA models

	A	R2X	R2X(cum)	Eigenvalues	Q2	Q2(cum)
PCA	1	0.386	0.386	16.2	0.364	0.364
	2	0.196	0.582	8.2	0.294	0.551
	3	0.139	0.721	5.9	0.301	0.686
	4	0.070	0.791	2.9	0.218	0.755
	5	0.057	0.848	2.4	0.240	0.814
	6	0.031	0.879	1.3	0.054	0.824
Hybrid PCA	1	0.346	0.346	12.1	0.324	0.324
	2	0.264	0.610	9.2	0.380	0.581
	3	0.094	0.704	3.3	0.202	0.666
	4	0.050	0.754	1.8	0.066	0.688
	5	0.040	0.794	1.4	0.066	0.709
	6	0.036	0.830	1.3	0.054	0.724
	7	0.030	0.860	1.0	0.031	0.733

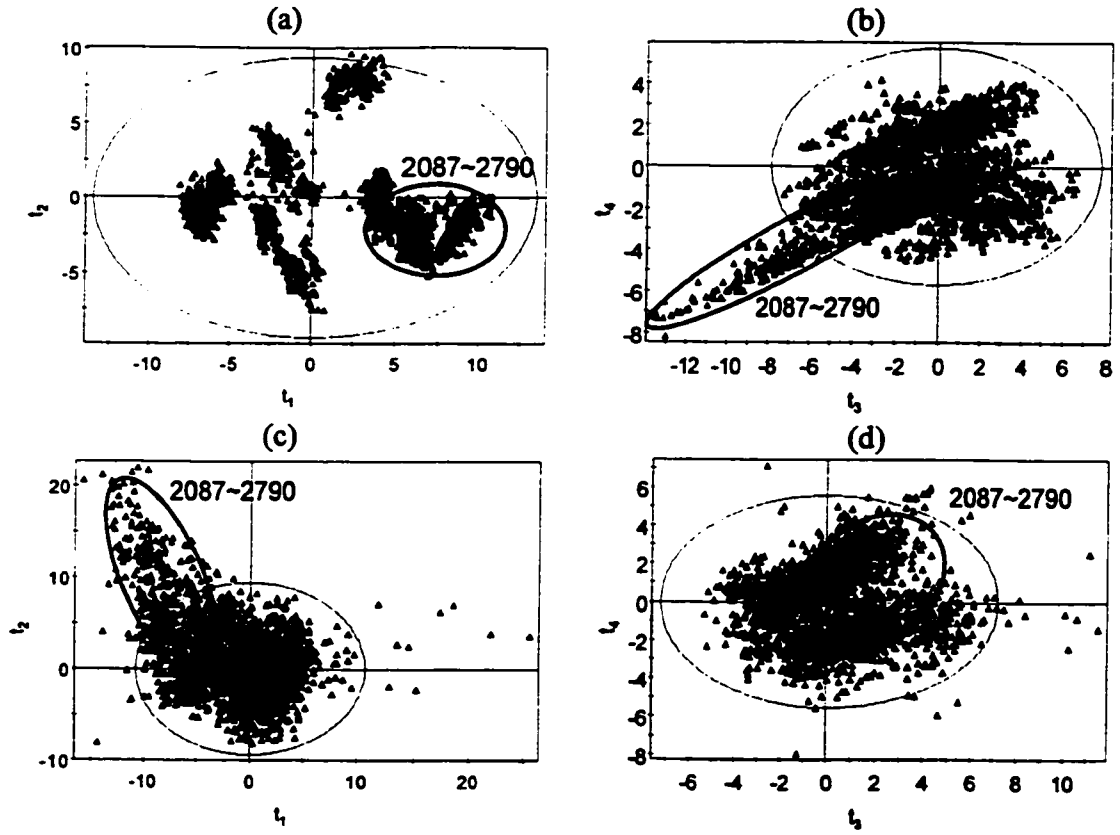
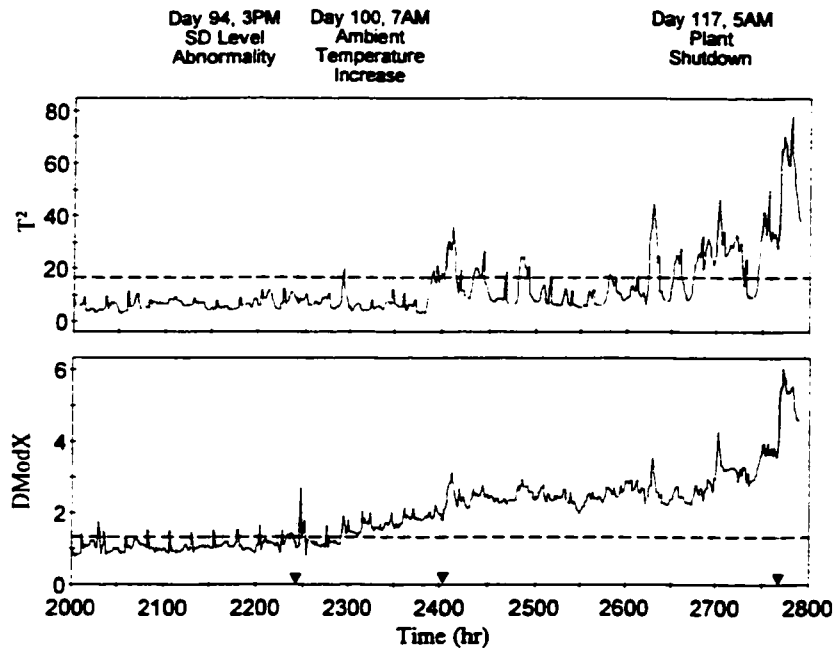


Figure 6.8 Score plots; (a) t_1 vs. t_2 of regular PCA model; (b) t_3 vs. t_4 of regular PCA model; (c) t_1 vs. t_2 of hybrid PCA model; (d) t_3 vs. t_4 of hybrid PCA model

With the training data, 15 principal components were identified through the use of the cross-validation as a stopping criterion. This correlation model explained 98.7 % of the total process variations. However, the 6 principal components corresponding to those with eigenvalues greater than 1 were used for the simplicity of the PCA model. They explained 87.9 % of the total variations. Table 6.3 summarizes the PCA model based on the training data. The first two principal components explain more than 56 % of the total variations. Figure 6.8 (a) and (b) show the score plots of t_1 vs. t_2, and t_3 vs. t_4 for the training (1~2086) and testing (2087-2790) data. The training and testing data on the t_1 vs. t_2 scores space formulate several clusters. Interestingly, the testing data corresponding to abnormal operation is found within the 99 % control region. On the other hand, the score plot of t_3 vs. t_4 in Figure 6.8(b) properly indicates the abnormality of the plant operation.

(a)



(b)

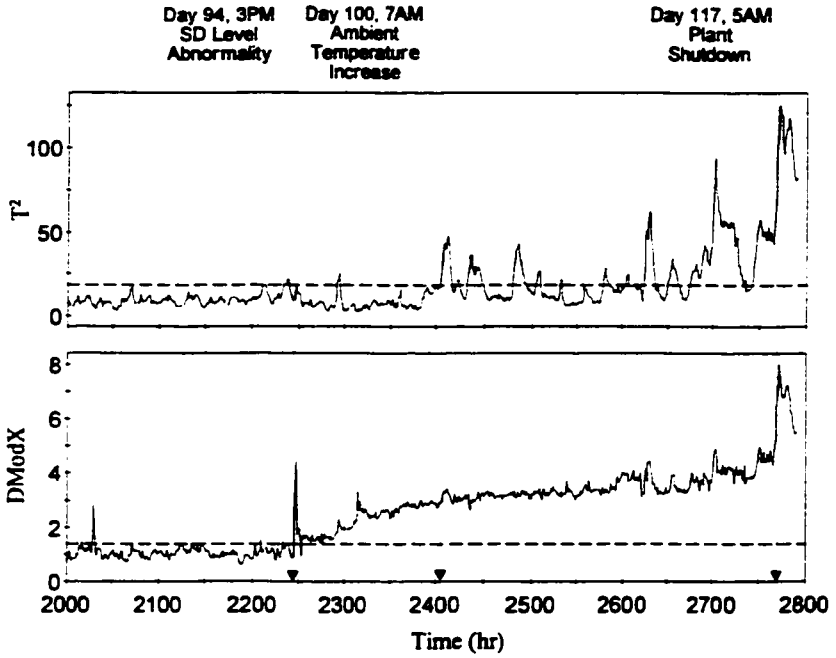


Figure. 6.9 Monitoring polymer accumulation with the regular and the hybrid PCA models; (a) T^2 and DModX plots based on regular PCA model (Dcrit = 1.30 with 6 PCs, 99% control limit); (b) T^2 and DModX plots based on hybrid PCA model (Dcrit = 1.34 with 7 principal components, 99% control limit)

Figure 6.9(a) shows the T^2 and DModX based on the PCA model (DModX statistic is just a normalized SPE. For the i^{th} observation in the **X** space, $DModX_i = \sqrt{SPE_i/(m-A)}$, where m is the number of variables in **X** and A is the number of latent variables. For detail, refer to Duchesne (2000). The DModX plot indicates the deviation of the *common-cause* variations around 2247 T_S. The variable contribution to DModX at 2247 T_S is shown in Figure 6.10(a). It was found that the levels of the WOSD were mainly responsible for the abnormality. Even though the level sensors seemed functioning properly at this time, they were about to deviate from the *common-cause* variations. This means that the polymer accumulation within the process and the instruments had already started. It is also observed that the fault detection was hindered by several substantial false alarms. One of the disturbing effects was a high peak occurring at every 24 hour.

However, the deviation of the level sensors from the *common-cause* variations does not directly explain how it resulted in the malfunction of the level sensor of WOSD waterside (WOSDWL) approximately one month later. One may find out a possible explanation from Figure 6.9(a), T^2 monitoring plot. It indicates that the 99 % control limit was violated around 2410 T_S. This corresponds to the data cluster outside of the control limit of the t_3 vs. t_4 score plot in Figure 6.8(b). The variable contributions to 3rd principal component between 2383 and 2412 T_S in Figure 6.10(b) indicates a group of variables being responsible for the T^2 deviation at this time. They are the temperature measurements around the air fan cooler (AFCoV1T, AFCoV2T), the return line temperature from the stripper (StrTT), the compressor outlet temperature (CoT), the bottom valve opening of the 1st separation drum (SD1WV), the water side level of the W/O settling drum (WOSDWL), the ambient temperature (AmbT), pressure measurements within the reactor (R1iP, R1oP, R2oP, R3oP, RxoP), and so forth. It is difficult to diagnose the root cause from this contribution plots.

In summary, the PCA monitoring plot indicates the initial symptom of the catastrophic plant shutdown about one month in advance. If this plant had been equipped with the plant monitoring system, the plant shutdown could have been avoided. On the

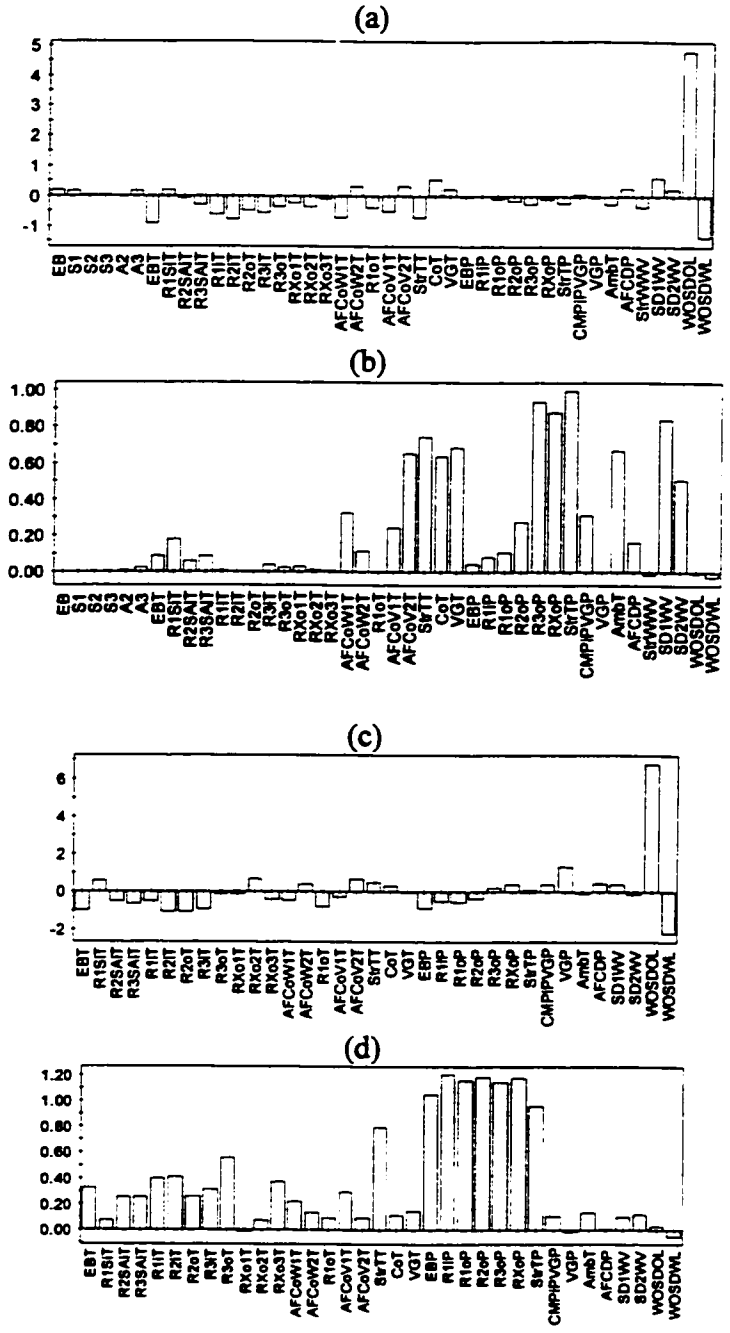


Fig. 6.10. Diagnosis of polymer accumulation; (a) Variable contribution to DModX at 2247 T_S based on PCA model; (b) Variable contribution to 3rd scores between 2383 and 2412 T_S based on PCA model; (c) Variable contribution to DModX at 2247 T_S based on hybrid PCA model; (d) Variable contribution to 1st scores between 2383 and 2412 T_S based on hybrid PCA model

other hand, the fault detection using the regular PCA is significantly affected by the fault alarm that may be caused by several process disturbances. The fault signature shown by the variable contributions in Figure 6.10(b) is also distorted by several blocking effects of the process disturbances.

6.6.3 Fault detection and diagnosis using hybrid PCA model

The process disturbances listed in Table 6.4 consist of the external input changes, intermediate process changes, and the external output changes. Among these disturbances, the ambient temperature caused a day and night load change to the air fan cooler (AFC) since the cooling capacity of the cooled air heavily depends on the AFC inlet temperature. The effect of temperature change was magnified in the condenser stripper as a 24-hour peak due to mass imbalance during daytime. The operating condition changes in feed, steam, and air flow rates, also resulted in significant effects on most process variables. Table 6.5 summarizes the percent variation extracted by laundering out each disturbance and its combinations. All the effects of the EB flow, the steam flow rates, the air flow

Table 6.4 Disturbances of SM production process

Disturbances	Variables
External inputs	
Feed	EB
Medium pressure steam / make-up	Measured
Low pressure steam / make-up	S1
Compressed air	A2/A3
Boiling Water	Measured
Intermediate changes	
Water flow from oil/water settling drum to stripper	WOSDWL
Discharge temperature of vent gas compressor	CoT
Ambient temperature	AmbT
24 hour peak caused by stripper level control	StrWWV
External outputs	
Product	SM
Waste water disposal	Measured
Recycle water from stripper	Measured
Vent gas flow	Measured

Table 6.5 Percent variations by each target variables

Laundered variables	Variations (%)
EB	32.4
Ambient temperature	16.9
24 Hour peak	4.2
Air flow	46.6
Steam flow	48.6
EB / Air flow	32.4 / 16.8 (49.2)
EB / Air / Steam flows	32.4 / 16.8 / 9.2 (58.4)
EB / Air / Steam flows / 24 hour peak	32.4 / 16.8 / 9.2 / 3.4 (61.8)

Note: The value in the parenthesis indicates the summation of all the % variations removed

rates, and a variable related to the 24 hour peak were considered the blocking factors that degrade the sensitivity of the correlation model. These four effects were, therefore, set as the target variables (G) and their effects were regressed from the original data to leave $\mathbf{Q}_G\mathbf{X}$. To take into account transient effects, the target variable column was augmented with its lagged values. The number of lags was 5.

With the residual data after subtracting the major disturbance effects from the original extracted data, 12 principal components were identified based on the cross-validation as a stopping criterion. They explained 95.3 % of the total process variations. Note that fewer principal components are required to explain all the process variations. Among the 12 principal components, 7 corresponding to the principal components with eigenvalues greater than 1 were used for the simplicity of the PCA model. They explained 86.0 % of the total variation. Figure 6.8(c) and (d) show the score plots of t_1 vs. t_2, and t_3 vs. t_4 for training (1~2086) and testing (2087-2790) data. The testing data on the t_1 vs. t_2 scores space clearly shows the abnormal operation whose data deviates from the 99 % control region.

In Figure 6.9(b), T^2 and DModX charts based on the laundered data ($\mathbf{Q}_G\mathbf{X}$) provide essentially the same fault detection by the regular PCA model as shown in Figure 6.9(a). It is, however, noted that many false alarms have been significantly reduced. The variable contribution to DModX at 2247 T_S is shown in Figure 6.10(c). The contribution

plot clearly indicates that the levels of the WOSD were responsible for the abnormality.

Figure 6.10(d) shows the variable contributions to the 1st score of the hybrid PCA model between 2383 and 2412 T_S. It indicates that a group of variables were also responsible for the abnormality. Most of the pressure measurements around the reactor and the temperature measurements around the AFC have high contributions. Note that the similar contribution pattern is shown by the regular PCA model in Figure 6.10(b), but the effects of the target variables in the case of the hybrid PCA model are not involved in the variable contributions. One can thus confirm that the disturbance variables used for the regression are not responsible for the abnormality at 2412 T_S.

An in-depth analysis indicates that the sudden ambient temperature increase is highly correlated with the situation that causes the styrene monomer to polymerize. Since the cooling capability of the air fan coolers is significantly affected by the ambient air temperature, the degraded cooling capability caused by the sudden ambient temperature increase results in the temperature increase of the reactor products including the condensing water at the downstream of the air-fan cooler. In addition, the ambient

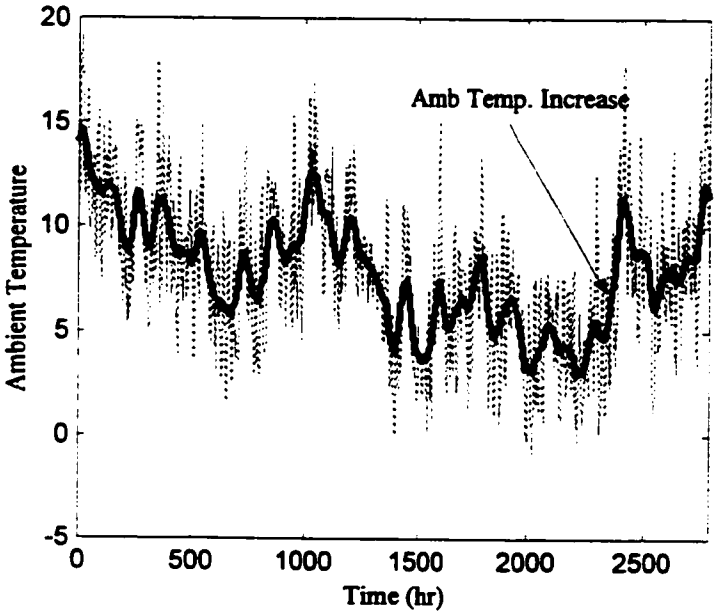


Figure 6.11 Original and filtered values of ambient temperature (dotted: original; solid: filtered)

temperature increase also caused the system pressure increase. It was found that the compressed air flow rates to the second and the third stage reactors had been increased at the same time. Consequently, the increased system pressure hindered the forward reactions and magnified the reactor outlet temperature as well. The combined effects accelerated the polymer accumulation. This analysis is consistent with the contribution plot shown in Figure 6.10(d). Figure 6.11 shows the filtered ambient temperature trend over 5 months, which confirms that there exists a sudden ambient temperature increase around 2412 T_S. As a result, the viscous polymer has been continuously accumulated in the process equipment including the level sensor of the oil/water-settling drum. After a while, the accumulated polymer resulted in the malfunction of the level sensor. Due to the faulty level sensor, the viscous polymer between water and oil layers overflowed to the oil side of the oil/water-settling drum. Finally, this overflowed viscous polymer blocked the suction line of the styrene momomer product pump.

On the other hand, since the data was composed of hourly averaged values, the enhancement on the fault detection sensitivity could not be clearly shown. To confirm how much the hybrid PCA model enhances the sensitivity of the fault detection, an

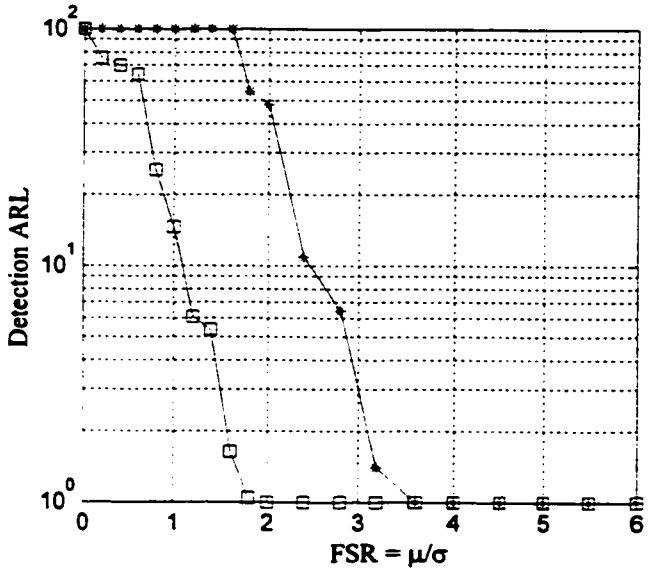


Figure 6.12 Monte Carlo simulation for sensor bias on 2nd stage outlet temperature (# of simulations at one FSR: 1000; α=0.01; ✳: regular PCA; □: hybrid PCA)

artificial sensor bias is added to the 3rd stage reactor outlet temperature sensor while changing the bias magnitude. At each magnitude of the sensor bias, 500 simulations have been done. As a fault detection measure, average run length (ARL) based on the T^2/SPE monitoring plots is used. The control limits of both monitoring plots were adjusted such that the overall significant level is 0.01. Figure 6.12 summarizes the result of Monte Carlo simulation. It is clearly shown that the hybrid PCA model significantly enhances the sensitivity of the fault detection.

It is noted that there may be potential improvement if one uses wavelet filtering to remove unnecessary fluctuation caused by the ambient temperature while keeping the non-stationary part of the temperature variation.

6.7 Discussion on the issues of process control

The effects of process control on the multivariate statistical methods have not been well studied. In what follows, issues on the use of hybrid correlation model approaches to system with process controllers are discussed.

Process control makes the fault detection and isolation difficult using the correlation models. Controller action distorts the initial fault signatures, and diminishes the duration of the initial fault signature since the controller action tries to eliminate a deviation between the controlled variable and its setpoint. This results in the propagation of a fault effect into the other measurements. In addition, changing the controller mode into the one that is different from the original model during the data collection would cause a different correlation. Also changing the control structure would result in a different correlation structure among process measurements. Even a set of different tuning parameters will probably cause a different correlation. This difficulty becomes intensified when the multivariable predictive controllers (MPC) are placed in the process. In MPC, many process variables are involved and, the controller status including modes, tuning and control structure is frequently varying for model updating, preventive maintenance on instruments, and so on. Different constraint sets of manipulated and

controlled variables become active in different times.

In principle, the control structure and the status should be consistent for both the normal operation and the predicted data. This requirement cannot be met in practice since the process includes many controllers. As a result, one may include all the operation data that includes different controller modes for modeling *common-cause* variations. The correlation model would then be less sensitive as more data covering the different controller modes is included in building the correlation model. Consequently, a very sluggish monitoring performance may occur.

In building a correlation model, it is usually assumed that data consists of measurements that were taken at times when the process had reached a steady-state, and the mode of operating process has remained consistent over time *i.e.* the covariance structure of the data is not time-varying. It is also assumed that the disturbances stay at a constant level over an extended period. When the disturbance values are not available and the data contains only the values of the manipulated variable and the output, the regression coefficient between the output and manipulated variables can be estimated by fitting a model to the data. The presence of feedback correlation entails misleading results if the estimated regression coefficient is interpreted as a causal relationship. This is based on the fact that there are two sources of variation in operating data: one source obviously comes from changes in the manipulated variables which in turn cause a variation in intermediate process variables and in output variable such as product quality. A second source of variation lies in disturbances, which also affect intermediate and output variables, which is not directly measurable. Some operating variables may have moved deliberately in order to lessen or completely counteract the effect of a disturbance on the output variable. In the course of feedback, the variation caused initially by the disturbance is transferred from the intermediate process variables and output variables into one or more manipulated variables that are changed in response to the disturbance effect. In a more general case of several manipulated variables, the difficulties caused by correlation with disturbance are even more complex (Jaeckle, 1998). The unmeasured disturbances and the feedback control to compensate for them are the key factors for

tilting the causal relationship among variables and generating the governing covariance structure.

A PCA model built with the data collected under one controller mode cannot be used for diagnosing those faults occurring under a different controller mode or a different set of active constraints. This is because operation data with different controller status usually has a different correlation structure. Based on the data decomposition using (6.9) and controller status information, the effect of controller status could be removed by regressing out the observation matrix with dummy variable which indicates controller status, $\mathbf{X} = \mathbf{P}_{Mode}\mathbf{X} + \mathbf{Q}_{Mode}\mathbf{X}$. The purpose of this data separation is to decompose the observation matrix into two terms which belong to the different controller modes. The model building would then be done with each term of $\mathbf{Q}_{Mode}\mathbf{X}$ and $\mathbf{P}_{Mode}\mathbf{X}$, respectively. This way of model building confirms that all the observations used for the estimation of each correlation model would be collected under same controller mode.

When the feedback control is perfect and all the disturbances are reflected into the manipulated variable space, the controller output will include all the information on the manipulated variable. In this case, one can set the controller output as a target variable to remove the controller effect. Laundering data with this target variable may remove all the control related variations from the observation matrix and minimize the fault propagation by control actions. That is, the observation matrix is decomposed with the controller output taken as G, $\mathbf{X} = \mathbf{P}_{OP}\mathbf{X} + \mathbf{Q}_{OP}\mathbf{X}$. However, when the effects of more than one controller are related to each other, the interpretation becomes complicated. Thus the usage of the decomposition-based method would be context-specific.

6.8 Conclusion

This research is aimed at coping with a practical situation in industries where external information is available. Issues of dealing with process knowledge for correlation modeling have been addressed. Several existing methods were explained under the framework of the hybrid correlation modeling. Types of prior knowledge used

for the correlation modeling were clarified.

Using least squares and prior knowledge, the observation matrix is decomposed into several terms. In general, it is decomposed into two terms: one including the effect of prior knowledge and the other term including the remaining effects. Depending upon the applications, one can use any component of the decomposed terms. It was shown that various types of prior knowledge would be used to enhance the FDI performance. However, caution needs to be exercised in using the prior knowledge for correlation modeling. When removing the effect of process disturbances, it has to be confirmed that the target variables would be clearly unnecessary process variations for monitoring.

The hybrid correlation modeling has been assessed with a few illustrating examples and the industrial data. It was shown that the proposed method provided improved fault detection and isolation results when available prior knowledge was properly used.

7. Summary and conclusions

The general objective of this thesis was to provide practical solutions for process monitoring and fault diagnosis of industrial processes. The research was aimed at extending the multivariate statistical process control methodology by using external information. The following four subtopics were addressed in this thesis: (i) enhancement of fault isolation using additional knowledge on faults based on historical data or first principles; (ii) fault detection and isolation of dynamic system through multiresolution analysis via wavelet transformation; (iii) comparison of causal and statistical model based approaches; (iv) incorporation of operational knowledge on process variables and observations into multivariate statistical models. In the following, each area of discussion is summarized, the contributions to the field are outlined and some conclusions are drawn.

7.1 Fault isolation using SS fault signatures under MSPC

Multivariate statistical approaches to fault detection based on historical operating data have been found to be useful in processes having a large number of measured variables. For fault isolation or diagnosis, however, they have been less powerful because of the non-causal nature of the data. The main contribution of this work is enhancing fault isolation of MSPC methods by using additional knowledge on faults.

To improve the fault isolation, additional data on past faults have been used to supplement the models. A critical literature review on fault isolation is given, and an improved approach capable of handling both simple and complex faults is presented. This

approach extracts fault signatures that are vectors of movement of the fault in both the model space and the residual space. The directions of these vectors are then compared with the corresponding vector directions of known faults in the fault library. Isolation is then performed based on a joint plot of the angles between the vectors of the current fault and those of the known faults. Although the fault signatures are based on steady-state information, the methodology assumes that it is applied to dynamic data. The method has been demonstrated using a simulated CSTR system with feedback control, and has been shown to be effective in isolating both simple and complex faults.

The fault signatures used in this proposed approach have been based on the steady-state information obtained from a model, from a plant test, or from historical data of the known faults. It was indicated that the use of steady-state fault signatures could lead to a time delay in fault isolation. In order to reduce the time delay in isolating a fault and to identify it more accurately, it is necessary to use dynamic data collected during the evolution of the fault. This can be included in PCA models using time-lagged data, and may allow one to perform better diagnosis of evolving faults.

7.2 Unifying PCA and multiscale approaches to FDI

The usage of steady-state fault signatures in the previous work can cause an isolation delay, and a false isolation may arise if the transient fault directionality is very different from the steady-state one. The use of a transient fault trajectory or an initial fault signature may help fault detection and isolation. In this study, indirect use of dynamic information in MSPC is proposed. Multiresolution analysis using wavelet transformation is used to decompose the data in frequency. A generalized PCA has been proposed as a modeling method for data showing multiscale features. Its potential features have been presented and examined through case studies. The modeling properties have been demonstrated through the comparison of the PCA and the MSPCA with data generated by the CSTR simulation model.

By using orthonormal wavelets, the orthogonality among scale blocks is retained.

The scale block orthonormality gives a good model interpretation such that the principal components of GPCA become a function of the variables only in one scale block. Thus, one can explicitly decompose the effect of process correlations over different scale levels. It also enables one to remove unnecessary process variation for further analysis, and cluster several scales showing similar correlation to efficiently summarize the multiscale process correlations.

The proposed scheme unifies the existing methods. The GPCA model not only conveys all the information contents shown in both MSPCA and PCA, but also reveals the relative significance of all the principal components over the scales. Multiscale analysis may be best suited for extracting information from measured data when each process variation has its dominant frequency content at different scale from the others. It is not always guaranteed that the multiscale analysis will give better results than the single scale analysis. For example, when all signals have their dominant frequency contents at similar frequency bands there would be little difference between the single scale analysis and multiscale analysis.

The on-line/off-line implementation procedures for both fault detection and isolation have been presented. It was shown that FDI performance could be significantly enhanced due to the usage of scale information. FDI performance using the GPCA has been assessed through Monte Carlo simulation with the CSTR system for several types of faults and the complex fault problem. When an abnormal event occurs at one scale bandwidth, the GPCA based FDI method will be more effective than other methods based on the regular PCA model in detecting and isolating faults. However, it is questionable if the fault effect is spread over more than one frequency band, or the fault effect in one scale is minor compared to the other scale components. An appropriate fault detection and isolation method has to be determined by considering fault characteristics. The fault isolation method presented in the chapter 2 can also be used together with fault signatures based on the GPCA model. In addition, the GPCA method can also be used as an approximate dynamic modeling method. Further research on these issues would be worthwhile.

7.3 Statistical and causal model-based approaches to FDI

The primary objective of this part of the research was to compare the fundamental differences between the causal-model-based and the statistical-model-based FDI methods to FDI problem, and thus suggest that combining the two approaches may resolve many of the difficulties inherent in each method individually. Recognizing the differences between the two approaches is critical in the selection of appropriate approaches, and providing insight into how the approaches can be combined. This research was the first to analyze and compare these two quite different approaches, thus resolving several misconceptions in the literature.

By examining the differences in the models, the data required to build models, the processes to which they are applicable, and the assumptions behind the methods, various fundamental aspects of FDI system design, and practical differences in these approaches were assessed. The differences in the nature of the models used (arising from the nature of the data used to build them) are shown to be responsible for the very different approaches to fault detection and isolation employed by the two methods. The model differences are also shown to be responsible for their major strengths and weaknesses.

Because the approaches have complementary strengths as well as differences, a potentially fruitful area of research involves how to combine the different methods in a manner that utilizes the strengths of both. We hope that this chapter will serve to foster such research by pointing out major differences between these approaches and discussing their strengths and weaknesses.

7.4 Incorporation of external information into multivariate statistical models

Most processes are subject to changes in the operating conditions such as feed rate and composition, product specification, controller status, and so on. Sometimes these

large *common-cause* variations disguise or distort the relevant information on faults. The FDI performance can then be significantly degraded. The primary objective of this work was to investigate how to incorporate various types of external information into multivariate statistical models to deal with the FDI problem. External information includes operational knowledge on process variables and observations, calculated parameters based on first principles, and so on.

The new model called a hybrid correlation model is estimated by incorporating prior knowledge into an empirical correlation model. Various approaches to incorporating different types of prior knowledge into the correlation models, and using them for process monitoring and fault diagnosis were presented in this study. The focus was given on integrating the available resources under an MSPC framework, and using the hybrid correlation model for FDI. Using least squares and prior knowledge, the observation matrix was decomposed into two terms: one including the effect of prior knowledge and the other term including the remaining effects. MSPC was then applied to either of these parts.

However, one needs to be careful in using the prior knowledge for correlation modeling. It has to be confirmed that the target variables would be clearly unnecessary process variations for monitoring. The clarification of the variable types is a key factor in the decomposition. The hybrid correlation modeling has been assessed with a few illustrating examples and an industrial dataset. Fault detection and isolation based on the proposed hybrid models has been also shown to be capable of outperforming the standard MSPC approaches when the available prior knowledge is properly used.

7.5 Presentation and publication

This section lists the presentations and the publications based on the thesis project. The presentations are as follows;

1. Yoon, S. and J. F. MacGregor (1998). Sensor fault diagnosis for dynamic systems

using multivariate statistical methods. *AIChE Annual meeting*, Nov. 10~15, Miami beach, FL, USA,

2. Yoon, S. and J. F. MacGregor (2000). Relationships between statistical and causal model based approaches to fault detection and isolation. *International symposium on advanced control of chemical processes (ADCHEM 2000)*, June 14-16, Pisa, Italy,
3. Yoon, S. and J. F. MacGregor (2000). Unifying principal component analysis and multiscale approaches. *50th CSChE conference*, Oct. 13~15, Montreal, Quebec, Canada,
4. Yoon, S. and J. F. MacGregor (2000). PCA of multiscale data and its application to fault diagnosis. *AIChE Annual meeting*, Nov. 11~17, Los Angeles, CA, USA,
5. Yoon, S. and J. F. MacGregor (2001). Unifying PCA and multiscale approaches to fault detection and isolation. *6th IFAC symposium on dynamics and control of process systems (DYCOPS-6)*, June 4~6, Chejudo, Korea,
6. Yoon, S. and J. F. MacGregor (2001). Incorporation of external information into multivariate PCA/PLS models. *4th IFAC workshop on on-line fault detection and supervision in the chemical process industries (CHEMFAS-4)*, June 7~8, Chejudo, Korea,

The refereed publications including the proceedings and journals are as follows;

1. Yoon, S. and J. F. MacGregor (2000a). Statistical and Causal Model-Based Approaches to Fault Detection and Isolation. *AIChE J.*, **46**(9), 1813-1824.
2. Yoon, S. and J. F. MacGregor (2000b). Reply to letter to the editor. *AIChE J.*, **46**(9), 1897-1899.
3. Yoon, S. and J. F. MacGregor (2000c). Relationships between statistical and causal model based approaches to fault detection and isolation. In: *Proc. of international symposium on advanced control of chemical processes (ADCHEM 2000)*, June 14-16, Pisa, Italy, 81-86.
4. Yoon, S. and J. F. MacGregor (2001a). Fault Diagnosis with Multivariate Statistical

Models, Part I: Using Steady-State Fault Signatures. *J. of Process Control*, **11**. 387-400.

5. Yoon, S. and J. F. MacGregor (2001b). Unifying PCA and multiscale approaches to fault detection and isolation. In: *Proc. of 6th IFAC symposium on dynamics and control of process systems (DYCOPS-6)*, June 4~6, Chejudo, Korea.
6. Yoon, S. and J. F. MacGregor (2001c). Incorporation of external information into multivariate PCA/PLS models. In: *Proc. of 4th IFAC workshop on on-line fault detection and supervision in the chemical process industries (CHEMFAS-4)*, June 7~8, Chejudo, Korea.
7. Yoon, S. and J. F. MacGregor (2001d). Principal Component Analysis of multiscale data. Submitted to *Journal of Chemometrics*.
8. Yoon, S. and J. F. MacGregor (2001e). Generalized Principal Component Analysis for process monitoring and fault diagnosis. Submitted to *AIChE Journal*.
9. Yoon, S. and J. F. MacGregor (2001f). Incorporation of external information into multivariate statistical methods. Submitted to *Journal of Chemometrics*.

References

Alt, F. B. and N. D. Smith (1988). Multivariate process control in: P. R. Krishnaish and C. R. Rao (eds), *Handbook of Statistics* 7, North-Holland, Amsterdam, 333-351.

Alwan, L. C. and H. V. Roberts. (1988). Time series modeling for statistical process control. *J. Econ. Bus. Stat.*, **6**, 87-95.

Anderson, J. S., T. J. McAvoy and O. J. Hao (2001). Use of hybrid models in wastewater systems. *Industrial & Engineering Chemistry Research*, **39**(6), 1694-1704.

Aradhye, H. B., B. R. Bakshi, R. A. Strauss, and J. F. David (2000). Multiscale statistical process control using wavelets – theoretical analysis and properties, Submitted to *Technometrics*.

Bakshi, B. R. and G. Stephanopoulos (1994). Representation of process trends-III. Multiscale extraction of trends from process data. *Computers and Chemical Engineering*, **18**(4), 267-302

Bakshi, B. (1998). Multiscale PCA with application to multivariate statistical process monitoring. *AIChE J.*, **44**(7), 1596-1610.

Bakshi, B. (1999). Multiscale analysis and modeling using wavelets. *J. Chemometrics*, **13**(3-4), 415-434.

Basseville, M. (1988). Detecting changes in signals and systems - A survey, *Automatica* **24**(3), 309-326.

Becraft, W. R. and P. L. Lee, (1993). An integrated neural network/expert system approach for fault diagnosis. *Computers and Chemical Engineering*, **17**(10), 1001-1014.

Box, G. E. P. (1954). Some theorems on quadratic forms applied on the study of analysis

of variance problems: Effect of inequality of variance in one-way classification. *The Annals of Mathematical Statistics*, **25**, 290-302.

Box, G. E. P. and G. M. Jenkins (1976). *Time series analysis, forecasting and control.* Prentice Hall, New Jersey.

Burnham, A. (1997). Multivariate latent variable regression: modeling and estimation. Ph.D thesis, McMaster University, Hamilton, Ontario, Canada, Department of Chemical Engineering.

Christie, O. H. J. (1996). Data laundering by target rotation in chemistry-based oil exploration. *Journal of Chemometrics*, **10**, 453-461.

Chow, E. Y., and A. S. Willsky (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. on Automatic Control*, **AC-29**, 603-614.

Cheng, G. and T. J. McAvoy (1997). Multi-block predictive monitoring of continuous processes. In: *IFAC ADCHEM 1997*, Banff, Canada

Cohen, A., I. Daubechies, and V. Pierre (1993). Wavelets on the interval and fast wavelet transforms. *Applied Computation and Harmonic Analysis*, **1**, 54-81.

Crowe, C. M., Y. A. G. Campos, and A. Hrymak (1983). Reconciliation of process flow rates by matrix projection. *AIChE J.*, **29**(6), 881-888.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communication on Pure and Appl. Math.*, **41**, 909-996.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Ttransactions on Information Theory*. **36**(5), 961-1005.

Daubechies, I. (1992). Ten lectures in wavelets, *SIAM*, Philadelphia, PA.

Duchesne, C. (2000). *Improvement of processes and product quality through multivariate data analysis.* Ph.D thesis, McMaster University, Hamilton, Canada.

Dunia, R., S. J. Qin, T. E. Edgar, and T. J. McAvoy (1996). Identification of faulty sensors Using Principal Component Analysis. *AIChE J.*, **42**(10), 2797-2812.

Dunia, R., and S. J. Qin. (1998a). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice* **6**(4), 457-469.

Dunia, R. and S. J. Qin. (1998b). Subspace approach to multidimensional fault identification and reconstruction. *AIChE J.*, **44**(8), 1813-1831.

Fan, J. Y., M. Nikolaou, and R. E. White. (1993). An approach to fault diagnosis of chemical processes via neural networks, *AIChE J* **39**(1), 82-88.

Fathi, Z. W., F. Ramirez, and J. Korbicz. (1993). Analytical and knowledge based redundancy for fault diagnosis in process plants. *AIChE J.*, **39**, 42-56.

Frank, P. M. (1990). Fault diagnosis in dynamic system using analytical and knowledge based redundancy - A survey and some new results, *Automatica*, **26**(3), 459-474.

Garcia, E. and E. Frank (1996). On the relationship between observer and parameter identification based approaches to fault detection. In: *Proc. Automatic Control Conference 13th Triennial World Congress*, San Francisco, USA

Gertler, J. (1995a). Towards a theory of dynamic consistency relations. In: *Proc. of IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries*. Newcastle, England.

Gertler, J. (1995b). Diagnosing parametric faults: from parameter estimation to parity relations. In: *Proc. of American Control Conference*, Seattle, USA.

Gertler, J. and T. J. McAvoy (1997). Principal component analysis and parity relations - a strong duality. In: *Proc. of IFAC Symp. on Fault Detection, Supervision and Safety for Technical Processes*, Hull, UK.

Gertler, J. and D. Singer (1990). A new structural framework for parity equation based failure detection and isolation. *Automatica*, **26**(2), 381-388.

Gertler, J. and M. K. Kunwer (1995). Optimal residual decoupling for robust fault diagnosis," *Int. J. Control*, **61**(2), 395-421.

Gertler, J. and R. Monajemy (1995). Generating directional residuals with dynamic parity

relations. *Automatica*, **31**(4), 627-635.

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

Gertler, J., W. Li, Y Huang and T. J. McAvoy (1999). Isolation enhanced principal component analysis. *AIChE J.*, **45**(2), 323-334.

Gurden, S. P., J. A. Westerhuis, S. Bijlsma, and A. K. Smilde (2001). Modelling of spectroscopic batch process data using grey models to incorporate external information. *J. Chemometrics*, **15**, 101-121.

Harris, T. J. and W. H. Ross (1991). Statistical process control procedures for correlated observations. *Can. J. Chem. Eng.*, **69**, 48-57.

Herley, C., J. Kovacevic, K. Ronchandranm and M. Vetterli (1993). Tilings of the time-frequency plane: construction of arbitrary orthogonal bases and fast tiling algorithms. *IEEE Transactions on Signal Processing*, **41**(12), 3341-3359.

Herley, C. and M. Vetterli. (1994). Orthogonal time-varying filter banks and wavelet packets. *IEEE Transactions on Signal Processing*, **42**(10), 2650-2663.

Herley, C. (1995). Boundary filters for finite-length signals and time-varying filter banks. *IEEE Transactions on Circuit and Systems II: Analog and Digital Signal Processing*, **42**(2), 102-114.

Hodouin, D., J. F. MacGregor, M. Hou and M. Franklin (1993). Multivariate statistical analysis of mineral processing plant data. *CIM Bulletin*, **86**(975).

Hoskins, J. C., K. M. Kaliyur, and D. M. Himmelblau (1991). Fault diagnosis in complex chemical plants using artificial neural networks. *AIChE J.*, **37**(4), 137-141.

Huang, B. (2001). Process identification based on last principal component analysis. *Journal of Process Control*, **11**,19-33.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods-a survey, *Automatica* **20**, 387-404.

Isermann, R. (1993). Fault diagnosis of machine via parameter estimation and knowledge processing-tutorial paper. *Automatica*, **29**(4), 815-835.

Jaeckle, C. A. (1998). *Product and process improvement using latent variable modethods.* Ph.D thesis, McMaster University, Hamilton, Canada.

Jackson, J. E. and G. S. Mudholkar (1979). Control procedure for residuals associated with principal component analysis. *Technometrics*, **21**(3), 341-349.

Jackson, J. E. (1991). *A User's Guide to Principal Components*, John Wiley and Sons Inc., New York.

Johanson T. A., and B. A. Foss (1992). Representation and learning unmodeled dynamics with neural network memories. In: *Proc. American Control Conference*, **3**, 3037-3042, Chicago

Karlsson, G and M. Vetterli. (1989). Subband coding of finite length signals. *Signal Processing*, **17**(2), 161-168.

Kasashima, N., K. Mori, R. G. Herrera, and N. Taniguchi (1995). On-line failure detection in face milling using discrete wavelet transform. *CIRP Annals - Manufacturing Technology*, **44**(1), 483-487.

Kettaneh, N., J. F. MacGregor, B. Dayal, and S. Wold (1994). Multivariate design of process experiments (M-DOPE). *Chemometrics and Intelligent Laboratory Systems*, **23**, 39-54.

Kosanovich, K. A., K. S. Dahl, and M. J. Piovoso (1996). Improved process understanding using multiway principal component analysis, *Industrial & Engineering Chemistry Research*. **35**(1), 138-146.

Kosanovich, K. A., and M. J. Piovoso (1997). PCA of wavelet transformed process data for monitoring. *Intelligent data analysis*. **1**, 85-99.

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

Kourti, T., J. Lee, and J. F. MacGregor (1996). Experiences with industrial applications of projection methods for multivariate statistical process control. *Computers and Chemical Engineering*, **20**, S747-S752.

Kourti, T. and J. F. MacGregor (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, **28**(4), 409-427.

Kresta, J., J. F. MacGregor, and T. E. Marlin (1991). Multivariate statistical monitoring of process operating performance. *Can. J. Chem. Eng.*, **69**, 35-47.

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

Lennox, B., P. Rhtherford, G. A. Montague, and C. Haughin. (1998). Case study investigating the application of neural networks for process modeling and condition monitoring. *Computers and Chemical Engineering*, **22**(11), 1573-1579.

Luo R., M. Misra, and D. M. Himmelblau (1999). Sensor fault detection via multiscale analysis and dynamic PCA. *Ind. Eng. Chem. Res.*, **38**, 1489-1495.

Ljung, L. (1999). *System identification: Theory for the User*, Prentice-Hall, Inc., New Jersey.

MacGregor, J. F., T. Kourti and J. V. Kresta (1991). Multivariate Identification: A Study of several Methods, In: *Proc. IFAC Symposium ADCHEM-91*, Toulouse, France

MacGregor, J. F., C. Jaeckle, C. Kiparissides and M. Koutoudi (1994). Process monitoring and diagnosis by multiblock PLS methods. *AIChE Journal*, **40**(5), 826-838.

MacGregor, J. F. and T. Kourti (1995). Statistical process control of multivariable processes. *Control Engineering Practice*, **3**(3), 403-414.

Magni J. and P. Mouyon (1994). On residual generation by observer and parity space approaches. *IEEE Trans. on Automatic Control*,. **39**(2), 441-456.

Mallat, S. G. (1989a). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **11**(7), 674-693.

Mallat, S. G. (1989b). Multiresolution approximations and wavelet orthonormal bases of $L_2(R)$. *Trans. Amer. Math. Soc.*, **315**, 69-87.

Marlin, T. E. (1995). *Process Control: Designing Processes and Control Systems for Dynamic Performances*. McGraw-Hill, New York.

Mehra, R. K., and I. Peshon (1971). An innovation approach to fault detection and diagnosis in dynamic systems. *Automatica*, **7**, 637-640.

Meyer, Y. (1992). *Ondeletters et operateurs, I:Ondelettes, II:Operateurs de Calderon-Zygmund, III:Operateurs multilineaires*. Hermann, Paris. English translation published by the Cambridge University Press.

Meyer, Y. (1993). Wavelets, algorithms and applications, *SIAM*, Philadelphia, PA.

Miller, P., R. E. Swanson, and C. E. Heckler (1993). Contribution Plots: The missing link in multivariate quality control. In: *Conf. ASQCC and ASA*, Milwaukee, WI.

Misra, M. H., Yue, S. J. Qin and C. Ling (1999). Multivariate process monitoring and fault diagnosis by multi-scale PCA. Submitted to *Computer & Chemical Engineering*.

Misra, M., S. J. Qin, S. Kumar, and D. Seemann (2000). On-line data compression and error analysis using wavelet technology. *AIChE J.*, **46**(1), 119-132.

Montgomery, D. C. (1985). *Introduction to Statistical Quality Control*, Wiley, New York.

Montgomery, D. C. and G. C. Runger (1994). *Applied Statistics and Probability for Engineers*, John Wiley & Sons, New York.

Motard, R., and B. Joseph (1994). *Wavelet Applications in Chemical Engineering*, Kluwer Academic Publishers, Boston.

Negiz, A. and A. Cinar (1997). Statistical monitoring of multivariable dynamic processes

with state-space models. *AIChE J.*, **43**, 2002-2020.

Neogi, D., and C. E. Schlags (1998). Multivariate statistical analysis of an emulsion batch process. *Ind. & Eng. Chem. Res.*, **37**(10), 3971-3979.

Nomikos, P., and J. F. MacGregor (1994). Monitoring batch processes using multiway principal component analysis. *AIChE J.*, **40**(8), 1361-1375.

Nomikos, P., and J. F. MacGregor (1995). Multivariate SPC charts for monitoring batch processes, *Technometrics*, **37**(1), 41-59.

Nounou, M. N. and B. R. Bakshi (1999). On-line multiscale filtering of random and gross errors without process models. *AIChE J.*, **45**(5), 1041-1058.

Patton, R. (1995). Robustness in model based fault diagnosis: the 1995 situation. In: *Proc. of IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries*. Newcastle, England.

Psichogios, D. C. and L. H. Ungar (1992). A hybrid neural network-first principles approach to process modeling. *AIChE J.* **38**(10) 1499-1512.

Qin, S. J., S. Valle and M. J. Piovoso (1999). On unifying multi-block analysis with application to decentralized process monitoring. Submitted to *J. Chemometrics*.

Raich, A. and A. Cinar (1997). Diagnosis of process disturbances by statistical distance and angle measures. *Computers and Chemical Engineering*, **21**(6), 661-673.

Rioul, O. (1993). Discrete-time multiresolution theory, *IEEE Trans. on Signal Processing*, **41**(8), 2591-2606.

Roffel, J. J., J. F. MacGregor and T. W. Hoffman (1989). The design and implementation of a multivariate internal model controller for a continuous polybutadiene polymerization train. In: *IFAC DYCORD+'89*, Maastricht, The Netherlands.

Rotem, Y., A. Wachs and D. R. Lewin (2000). Ethylene compressor monitoring using model based PCA. *AIChE J.*, **46**(9), 1825-1836.

Russel, E. L., L. H. Chiang, and R. D. Braatz (2000). Fault detection in industrial

processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **51**, 81-93.

Shi, R. and J. F. MacGregor (2000). Modeling of dynamic systems using latent variable and subspace methods. *J. Chemometrics*, **14**(5-6), 423-439.

Sobol, I. M. (1974). *The Monte Carlo Method*, The University of Chicago Press, Chicago and London.

Soderstrom, T. and P. Stoica (1989). *System Identification*, Prentice-Hall, Cambridge.

Strang G. and T. Nguyen (1996). *Wavelets and Filter Banks*, Wellesley-Cambridge, Wellesley, MA.

Su, H. T., N. Bhat, P. A. Minderman, and T. J. McAvoy (1992) Integrating neural networks with first principles models for dynamic modeling. In: *Proc. of IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes* (DYCOPS+), 285-290.

Takane, Y., and T. Shibayama (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika*, **56**(1), 97-120.

Teppola, P. and P. Minkkinen (2000). Wavelet-PLS regression models for both exploratory data analysis and process monitoring. *J. Chemometrics*, **14**, 383-399.

Thompson, M. and M. A. Kramer (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE J.*, **40**(8), 1328-1340

Tong, H and C. M. Crowe (1995). Detection of gross errors in data reconciliation by principal component analysis. *AIChE J.*, **41**(7), 1712-1722.

Tracy, N. D., J. C. Young, and R. L. Mason (1992). Multivariate control charts for individual observations, *Journal of Quality Technology*, **24**, 88-95.

Tulleken, H. J. A. F. (1993). Grey-box modeling and identification using physical knowledge and Bayesian techniques. *Automatica*, **29**(2), 285-308

Vedam, H and V. Venkatasubramanian (1997). Signed diagraph based multiple fault

diagnosis. *Computers and Chemical Engineering*, **21**(12), S655-S660.

Venkatasubramanian, V. and K. Chan (1989). Neural network methodology for process fault diagnosis. *AIChE J.*, **35**(12), 1993-2002.

Vetteri, M. and C. Herley (1992). Wavelets and Filter Banks: Theory and Design. *IEEE Transactions on Signal Processing*. **40**(9), 2207-2232.

Vetteri, M. and J. Kovacevic (1995). *Wavelets and Subband Coding*, Prentice Hall, Englewood Cliffs, New Jersey.

Wachs, A. and D. R. Lewin (1998). Process monitoring using model-based PCA. In: *Proc. IFAC Symp. On Dynamics and Control of Process Systems*, Corfu.

Wagon, L. E. and B. R. Kowalski (1988). A multiblock partial least squares algorithm for investigating complex chemical systems. *J. Chemometrics*, **3**, 3-20.

Westerhuis, J. A., T. Kourti, and J. F. MacGregor (1998). Analysis of multiblock and hierachical PCA and PLS models. *J. Chemometrics*, **12**, 301-321.

Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, **12**, 601-611.

Willsky, A. S. and H. L. Jones (1976). A generalized likelihood ratio approach to detection and estimation of jumps in linear system. *IEEE Trans. on Automatic Control*, **21**, 108-112

Wilson, J. A. and L. F. M. Zorzetto (1997). A generalized approach to process state estimation using hybrid artificial neural network/mechanistic models. *Computers & Chemical Engineering*, **21**(9), 951-963.

Wise, B.M. and N.L. Ricker (1991). Recent advances in multivariate statistical process control: Improving robustness and sensitivity, In: *IFAC Symposium, ADCHEM*, Toulouse, France, 125-130.

Wise, B. M., D. J. Veltkamp, N. L. Ricker, B. R. Kowalski, S. M. Barnes and V. Arakali (1991). Application of multivariate statistical process control (MSPC) to the west

valley slurry-red ceramic melter process, In: *Proc. Waste Management*, Tuscon, AZ.

Wold, S. C. (1976). Pattern recognition by means of disjoint principal components models. *Pattern Recognition*, **8**. 127-139.

Wold, S., C. Albano, W. J. Dunn III, O. Edlund, K. Esbensen, P. Geladi, S. Hellberg, E. Johansson, W. Lindberg, M. Sjostrom (1984). Multivariate data analysis in chemistry. In *Chemometrics, Mathematics and Statistics in Chemistry*, B.R. Kowalski, ed., 17-95, Reidel, Dordrecht, The Netherlands.

Wold, S., S. Hellberg, T. Lundstedt, M. Sjostrom, and H. Wold (1987). In: *Proc. Symp. on PLS Model Building Theory and Application*, Frankfurt, Germany.

Wold, S., C. Albano, W. J. Dunn III, K. Esbensen, P. Geladi, S. Hellberg, E. Johansson, W. Lindberg, M. Sjostrom, B. Skagerberg, C. Wikstrom and J. Ohman (1989). Multivariate data analysis: Converting chemical data tables to plots, *Intelligent Instruments & Computers*, 197-216.

Wold, S., N. Kettaneh, and K. Tjessem (1996). Hierachical multiblock PLS and PC models for easier model interpretation and as an alternative to variable selection. *J. Chemometrics*, **10**, 463-482.

Wold, S., H. Antti, F. Lindgren, and J. Ohman (1998). Orthogonal signal correction of near –infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, **44**, 175-185.

Xiaoli, Li (1999). On-line detection of the breakage of small diameter drills using current signature wavelet transform. *International Journal of Machine Tools & Manufacture*, **39**(1), 157-164.

Yoon, S. and J. F. MacGregor (2000a). Statistical and causal model-based approaches to fault detection and isolation. *AIChE J.*, **46**(9), 1813-1824.

Yoon, S. and J. F. MacGregor (2000b). Reply to letter to the editor. *AIChE J.*, **46**(9), 1897-1899.

Yoon, S. and J. F. MacGregor (2000c). Relationships between statistical and causal

model based approaches to fault detection and isolation. In: *Proc. Int. Symposium on Advanced Control of Chemical Processes (ADCHEM 2000),* June 14-16, Pisa, Italy, 81-86.

Yoon, S. and J. F. MacGregor (2001a). Fault diagnosis with multivariate statistical models, part I: using steady-state fault signatures. *J. Process Control*, **11**, 387-400.

Yoon, S. and J. F. MacGregor (2001b). Unifying PCA and multiscale approaches to fault detection and isolation. In: *Proc. of 6th IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS-6)*, June, Chejudo, Korea.

Yoon, S. and J. F. MacGregor (2001c). Incorporation of external information into multivariate PCA/PLS models. In: *Proc. of 4th IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries (CHEMFAS-4)*, June, Chejudo, Korea.

Zhang, J., E. Martin, and A. J. Morris (1995). Fault detection and classification through multivariate statistical techniques. In: *Proc. American Control Conference*, Washington, USA.

Appendix. Simulation model

A nonisothermal continuous stirred tank reactor model (Marlin, 1995, page 90-92) is used to illustrate the proposed methods and examine several concepts. The reaction is 1st order ($A \rightarrow B$) and the reactor system involves heat transfer and heat of reaction. It is assumed that the tank is well mixed and the physical properties are constant. The process has one feed stream that is the merged stream of the solvent and the reactant A, one product stream, and a cooling water flow to the coils (Figure A.1). The reactant A (F_A) and the cooling water (F_C) flows control the outlet concentration (C_A) and temperature (T), respectively. In the simulation, however, only the temperature controller is assumed to be active unless otherwise stated. The controller equation to be used is as follows:

$$\Delta OP = K_C\left[(E_t - E_{t-1}) + T_S E_t / T_I\right], \text{ where } E_t = SP - PV_t \tag{A.12}$$

Measured process disturbances are the inlet concentrations (C_{AS}, C_{AA}), the inlet temperature (T_O), the solvent flow (F_S), and the cooling water temperature. The simulation model consists of the component material balance on the reactant and the energy balance as follows:

$$\frac{dC_A}{dt} = \frac{F}{V}C_{A0} - \frac{F}{V}C_A - k_0 e^{\frac{E}{RT}}C_A, \tag{A.2}$$

$$V\rho C_P \frac{dT}{dt} = \rho C_P F(T_0 - T) - \frac{aF_C^{b+1}}{F_C + aF_C^b / 2\rho_C C_{PC}}(T - T_{C,in}) + (-\Delta H_{rxn})Vk_0 e^{\frac{E}{RT}}C_A, \tag{A.3}$$

where the inlet reactant concentration is obtained from the two feed streams, the reactant and the solvent ($C_{A0}F_0 = C_{AA}F_A + C_{AS}F_S$), and the heat transfer coefficient is related to the

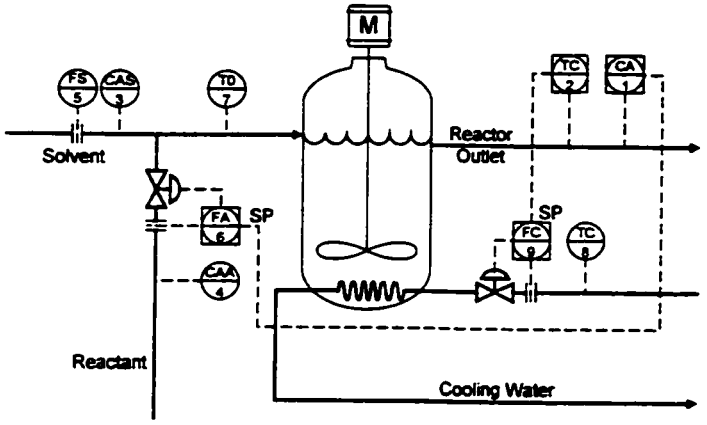


Figure A.1 Process flow diagram of CSTR system

coolant flow by the empirical relationship ($UA = aF_C^b$). In addition, it is assumed that the unmeasured stochastic disturbances arise from variations in reactive impurities that affect to the reaction and from fouling of the cooling coils. These latter two disturbances are also simulated as first order autoregressive behaviors in the reaction rate (k) and heat transfer constant (UA), respectively. These stochastic states are introduced as premultipliers (a_1 and a_2) in the reaction constant and heat transfer coefficient as follows;

$$k = a_1 k_0 \exp(-E/RT), \quad UA = a_2 a F_C^b . \tag{A.4}$$

All stochastic disturbances are modeled as first order processes;

$$x_t = \phi x_{t-1} + \sigma_e e_t , \tag{A.5}$$

where $e_t \sim N(0,1)$ and σ_e^2 is the disturbance variance. Measurement noise is added to all values of the process variables: $x_{t,meas} = x_t + \sigma_m m_t$, where σ_m^2 is the measurement error variance and $m_t \sim N(0,1)$. Table A.1 and A.2 show the simulation conditions of the base-case, the autoregressive parameters (ϕ), generating noise variances (σ_e^2) for the process disturbances, and the measurement noise variances (σ_m^2).

Table A.1 Simulation condition of base-case run

Simulation parameters	
$V = 1\ (m^3)$	$\rho = 10^6\ (g/m^3)$
$\rho_C = 10^6\ (g/m^3)$	$E/R = 8330.1\ (K)$
$C_P = 1\ (cal/gK)$	$C_{PC} = 1\ (cal/gK)$
$b = 0.5$	$k_0 = 10^{10}\ (m^3/kmole \times min)$
$a = 1.678 \times 10^6\ (cal/minK)$	$\Delta H_{rxn} = -1.3 \times 10^7\ (cal/kmole)$
Initial conditions	
$T_O = 370.0$ (K)	$T_C = 365.0$ (K)
$F_C = 15\ (m^3/min)$	$T = 368.25$ (K)
$F_S = 0.9\ (m^3/min)$	$F_A = 0.1\ (m^3/min)$
$C_A = 0.8\ (kmole/m^3)$	$C_{AS} = 0.1\ (kmole/m^3)$
$C_{AA} = 19.1\ (kmole/m^3)$	
Controller information	
$K_C(T) = -1.5$ %/K	$T_I(T) = 5.0$ (min)
$K_C(C_A) = 48.25$ %/$(kmole/m^3)$	$T_I(C_A) = 2.0$ (min)

Table A.2 Measurement noises and disturbances: $x_t = \phi x_{t-1} + \sigma_e e_t$; $x_{t,meas} = x_t + \sigma_m m_t$

	Measure. Noise (σ_m^2)	Process Noise (σ_e^2)	AR Coefficients (ϕ)
T	4×10^{-4}		
C_A	2.5×10^{-5}		
F_C	1.0×10^{-2}		
T_C	2.5×10^{-3}	0.475×10^{-1}	0.9
T_O	2.5×10^{-3}	0.475×10^{-1}	0.9
C_{AA}	1.0×10^{-2}	0.475×10^{-1}	0.9
F_A	4×10^{-6}		
C_{AS}	2.5×10^{-5}	1.875×10^{-3}	0.5
F_S	4.0×10^{-6}	0.19×10^{-2}	0.9
a_1		0.19×10^{-2}	0.9
a_2		0.0975×10^{-2}	0.95