The Pennsylvania State University

The Graduate School

The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering

THE CUSCORE AND HIGH-DIMENSIONAL CONTROL CHARTS FOR

STATISTICAL MONITORING OF AUTOCORRELATED PROCESS DATA

A Thesis in

Industrial Engineering

by

Shuohui Chen

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2006

The thesis of Shuohui Chen was reviewed and approved* by the following:

Harriet Black Nembhard
Associate Professor and Bashore Career Professor of Industrial and
Manufacturing Engineering
Thesis Advisor
Co-Chair of Committee

Paul H. Cohen
Distinguished Professor of Industrial and Manufacturing Engineering
Co-Chair of Committee

Runze (Richard) Li
Associate Professor of Statistics

Ling Rothrock
Assistant Professor of Industrial and Manufacturing Engineering

Richard J. Koubek
Professor of Industrial and Manufacturing Engineering
Head of the Department of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School.

ABSTRACT

In many statistical process control (SPC) applications, the quality of a process or a product can be characterized by a single variable. A vast body of research in SPC charts, such as Shewhart, EWMA, Cusum, and Cuscore control charts, has been developed in order to control univariate quality characteristics. However, rapid developments in modern technologies such as high-speed computing, Internet, sensors, and nanotechnology, etc., have provided many industries with data-rich environments in which observations of multiple quality characteristics are simultaneously available for analysis. In response, traditional SPC charts have been extended to the multivariate environment in the forms of Hotelling T^2, multivariate EWMA, multivariate Cusum control charts, and the like, and now play an important role in controlling the quality of processes and products. Furthermore, in many practical situations, it is assumed that the quality characteristics of processes or products can be better represented and summarized by a large collection of data points or variables, which can in turn be represented by a relationship or a function between a response variable and a number of explanatory variables. After each sampling stage, profile data is observed which is subsequently fitted to either a linear or a nonlinear regression model.

Traditional control charts assume that the sample data are sequentially independent over time and follow normal or multi-normal distribution. As a direct result, high false alarm rates will occur if these traditional charts are applied to monitor processes with high autocorrelation, as experience has shown in some industrial processes. One of the approaches in contending with the difficulties presented by

autocorrelation is to use so-called engineering process control (EPC) techniques to remove process autocorrelation, and then implement SPC control charts to monitor the adjusted processes. Indeed, this approach is frequently justified, as the integration of SPC and EPC has often proved effective for monitoring and controlling some autocorrelated processes.

In this research, the goal is to develop statistical monitoring methods to address both univariate and multivariate autocorrelated processes, and processes whose quality can be characterized or represented by profile data, which can be represented by a function or a linear/nonlinear regression model with response and independent variables. In pursuing such goals, three major objectives are achieved:

1. To develop the likelihood-based Cuscore control approach to monitor the mean shift of a univariate autocorrelated process adjusted by a generalized feedback control scheme. The methodology is applied to monitor an industrial valve system for leakage.

2. To extend the Cuscore control chart to monitor the mean shift of an autocorrelated multivariate process and to compare its performance to those of multivariate Cusum control charts. The methodology is applied to monitor a simulated reactive ion etching (RIE) process in semiconductor manufacturing.

3. To develop the high-dimensional control chart to monitor a process when its quality characteristics can be represented by a profile. The Fourier transform and adaptive Neyman test are used in this approach for detecting a change in the mean function of profiles. The impact of stationary noise on the performance of the

control chart is analyzed and the methodology is shown to be able to monitor both

linear and nonlinear profiles with good performance. The approach is applied to a

simulated woodboard manufacturing process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude and respect to my advisor, Dr. Harriet Black Nembhard. Without her support and supervision, this thesis could not have been accomplished. To me, she is not only my supervisor in my research, but also my good friend and mentor.

I am deeply grateful to Dr. Soundar R. T. Kumara, Dr. Paul H. Cohen, Dr. Runze Li and Dr. Ling Rothrock for being or having been a part of my doctoral committee. I appreciate their time and comments on my research.

I would also like to thank Dr. Dennis K. J. Lin for his enlightening teaching and discussions with me on the research.

I also should express my thanks to my fellow graduate students in the QUEST lab in the Department of Industrial Engineering of both the University of Wisconsin-Madison and the Penn State University, in particular, Navinchandra Acharya, Mehmet Aktan, Wenny Chandra, Pannapa Changpetch and René Valverde-Ventura

I would like to thank the staff and faculty in the Department of Industrial and Manufacturing Engineering of the Penn State University for their excellent teaching and student services.

Finally, I would like to express my appreciation to my parents and my brother, especially to my mother who passed away in China during my Ph.D. study. If there is a heaven, I wish she is there and glad to see that her son is presenting this dissertation to her as a precious gift.

To my mother

Chapter 1

Introduction

Originating with the Shewhart Chart in 1930s, statistical process control (SPC) charts have played a significant role in process quality control and improvement in business and manufacturing. With rapid developments in technology and business, modern processes have exhibited important characteristics requiring advanced SPC methodologies. Some of these characteristics and the resulting tasks for SPC in the development of modern processes will be introduced in this chapter, and the research objectives of this thesis will be briefly summarized.

1.1 Characteristics of Modern Processes

A *process* is a series of continuous actions or operations leading to an end result. In a manufacturing process, the series of actions or operations lead to a product with certain quality characteristics, such as performance, reliability, durability, etc.; in a service process, a series of actions is taken to perform or maintain some functions or services in order to meet customers' satisfaction.

The revolutionary development of modern technology, such as high computing, Internet, sensors and nanotechnology, etc., has been revolutionizing modern industries in large-size data collection and high data analysis capacity, and related aspects. In both manufacturing and nonmanufacturing applications, some processes characteristics

emerge as particularly significant and beckon for effective SPC approaches. Two important process characteristics are summarized below:

- High autocorrelation. Physically, all manufacturing processes are driven by inertial elements, and when the interval between samples is small, the observations on the process can be autocorrelated over larger intervals of time. Naturally, the shorter the interval, the higher the autocorrelation. For instance, with the use of advanced sensors and sampling techniques, intervals between neighboring observations can be in the centi-second range for vibration signals from the gear system test (Suh et al., 1999, Kamarthi et al., 2000). In such cases, the sequential observation data over time are highly correlated.

- High dimensionality. The quality of most practical processes or products cannot be adequately characterized by a single variable. In many such cases, observations from more than one variable are collected simultaneously and the cross-correlation among variables must be considered. Furthermore, for some processes or products, a large number of variables are needed in order to adequately characterize their quality. For instance, 313 measurements were collected to characterize a woodboard density profile (Walker and Wright, 2002) and 24 such profile data are plotted in Figure 1.1. For such profile data, changes in either all or some of variables or measurements are likely to cause it to go out of control.
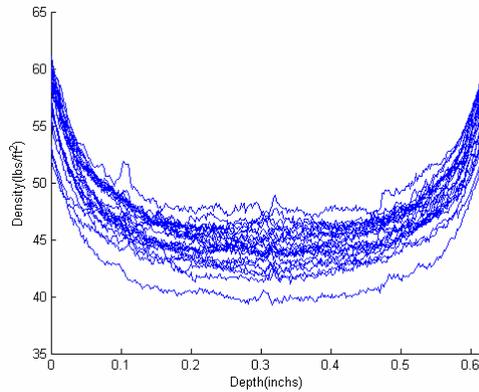
Figure 1.1: Twenty-four vertical density profiles of woodboards.

1.2   Tasks for Statistical Process Control in Autocorrelated Processes

Fundamentally, any process or nonquantum-mechanical system is physically governed by the second law of thermodynamics, which implies that if left to itself, the entropy or disorganization of any system not governed by quantum mechanics can never decrease and will eventually increase. Therefore, efforts of process control are needed in order to cancel the effect of the law and maintain the process quality characteristics within acceptable ranges.

The control chart is one of the major techniques for statistical process control (SPC) which is used to detect, actuate, and compensate for abnormal or fault signals that represent unacceptable disturbances within the process. Traditional control charts, such as Shewhart charts, EWMA charts and Cusum charts are based on the assumption that the samples are sequentially independent. However, high false alarm rates usually occur when they are used for monitoring autocorrelated processes. Therefore, advanced approaches are necessary to deal with process autocorrelation.

3

Process autocorrelation enables the prediction of future observations based on historical observations as long as the autocorrelation structure can be properly identified. If so, corresponding adjustments can be applied to compensate for the disturbance and control the process. Techniques of combining process monitoring and adjustment are more effective than either of the two approaches used separately. A likelihood-based cumulative score (Cuscore) control chart has proven to be an effective complementary approach in monitoring univariate autocorrelated process, and its integration with feedback-adjusted techniques is under investigation.

Multivariate process control has long been a challenging and rewarding research topic. However, much of the previous research in this topic ignores process autocorrelation and cross-correlation and is, consequently, unrealistic and unreliable for most practical applications. Based on the success of the Cuscore control chart for monitoring univariate autocorrelated processes, it seems both logical and promising to extend the Cuscore chart to monitor multivariate autocorrelated processes.

1.3   Tasks for Statistical Process Control in High-Dimensional Processes

The quality characteristics of some processes or products can be represented by a function or a linear/nonlinear regression model with response and independent variables. This type of data is called profile data, a branch of functional data. The technique for monitoring the profile data is known as profile monitoring. Figure 1.2 illustrates the representation of profile data in matrix form, where the number of measurements, n, is called profile dimensionality, which is usually larger than the number of profile samples, N.

4

$$
\begin{array}{c}
\text{Measurements}\\[4pt]
\begin{array}{ccccc}
 & 1 & 2 & \text{L} & n
\end{array}\\[2pt]
\begin{array}{r}
\text{Profile}\\
\text{Sample}\\
\text{Number}
\end{array}
\begin{array}{c}
1\\2\\\text{M}\\N
\end{array}
\begin{bmatrix}
x_{11} & x_{12} & \text{L} & x_{1n}\\
x_{21} & x_{22} & \text{L} & x_{2n}\\
\text{M} & \text{M} & \text{O} & \text{M}\\
x_{N1} & x_{N2} & \text{L} & x_{Nn}
\end{bmatrix}
\end{array}
$$

Figure 1.2: Representation of profile samples in matrix form.

Equation (1.1) gives the model form for the jth profile with the independent variable X and the response variable Y

$$
Y_{jk} = m(X_k, \boldsymbol{\beta}_j) + \varepsilon_{jk}, \; j = 1,2,\ldots, N \text{ and } k = 1,2,\ldots, n, \tag{1.1}
$$

where $m(X_k, \boldsymbol{\beta}_j)$ is the mean function which indicates the signal component of the jth profile, $\boldsymbol{\beta}_j$ is the parameter vector, and ε_{jk} is the noise component of the jth profile.

An application example of profiles is illustrated in Figure 1.1, in which 24 woodboard vertical density profiles are plotted. Each profile consists of 313 paired measurements of the depth and the density. The relationship between these two variables is essential for characterizing the quality of the woodboard products.

Using the control chart approach for profile monitoring is viewed as "the most promising area of research in statistical process control" by Woodall et al. (2004). Unlike most applications of univariate or multivariate SPC charts, in which the quality characteristics can be adequately represented by a single random variable or by a vector of multivariate distributed variables, the control chart for profiles is designed for monitoring more complex objects, such as curves and images, which normally consist of a relatively large number of observations or measurements. Techniques based on

dimensionality reduction and feature extraction play pivotal roles in all high-dimensional mathematical problems, and have many significant applications in biology, health studies, financial engineering and risk management, and machine learning and data mining (Fan and Li, 2006). The development of high-dimensional control charts for profile monitoring represents a challenging endeavor for statistical process monitoring and control.

1.4 Research Significance and Objectives

The goal of this thesis is to develop effective statistical monitoring methods for detecting the mean shift of an autocorrelated process whose quality may be characterized by a single variable, several jointly distributed variables or even a large number of variables. This work mainly focuses on the Phase II analysis, and is dependent on the assumption that the process has been properly characterized and modeled in Phase I. The research significance and objectives focus on three aspects:

1. To develop the likelihood-based Cuscore control chart to monitor the mean shift of a univariate autocorrelated process adjusted by a generalized feedback control scheme. The methodology is applied to a valve leakage detection process and its effectiveness is illustrated.

2. To extend the Cuscore control chart to monitor the mean shift of an autocorrelated multivariate process and to compare its performance to that of multivariate Cusum control charts. The approach is applied to monitor a simulated autocorrelated reactive ion etching (RIE) process in semiconductor manufacturing.

6

3. To develop the high-dimensional control chart to monitor a process whose quality characteristics can be represented by a profile, which generally consists of a signal component and a noise component. The approach is illustrated by monitoring two simulated woodboard density profiles data sets.

The major body of the investigator's Ph.D. research consists of work to meet these three objectives.

1.5   Thesis Organization

This thesis is organized as follows. Chapter 2 reviews the previous research on the topics covered in this thesis. Chapter 3 develops an approach for the integration of statistical process control (SPC) and engineering process control (EPC) for univariate autocorrelated process, with focus on cumulative score (Cuscore) chart and generalized minimum variance (GMV) control scheme. Chapter 4 extends the univariate Cuscore control chart to the multivariate environment and applies it to monitor the shift signal of process mean. Chapter 5 introduces the high-dimensional control chart for monitoring the mean function of profiles. Chapter 6 concludes the thesis and proposes some topics of future research. Appendices after Chapters 3 and 4 provide the derivations of some formulae and the source codes for the simulation programs in each chapter.

Literature Review

The review of the literature begins with univariate SPC charts and its integration with EPC. It then moves to multivariate control charts, and finally to the SPC approaches for profile data.

2.1   SPC Approaches for Univariate Autocorrelated Processes

A univariate process is a process whose quality can be characterized by single variable. Autocorrelation among observations exists in most practical processes, as explained by Montgomery (2005):

> "All manufacturing processes are driven by inertial elements,
> and when the interval between samples becomes small relative to these
> forces, the observations on the process will be correlated over time."

However, traditional SPC control charts, such as Shewhart, Cusum, and exponentially weighted moving-average (EWMA) charts, assume that the process observations are independent and identically distributed (i.i.d.). Therefore, high false alarm rates often occur when the traditional SPC control charts are used to monitor highly autocorrelated processes (Alwan and Roberts, 1988; Harris and Ross, 1991).

Some SPC approaches have been proposed for monitoring and controlling autocorrelated processes. Four major approaches will be reviewed in the following subsections. Montgomery (2005) put them into two categories: model-based approaches

and model-free approaches which differ in the need of an explicit time-series model. For the four approaches summarized below, the first two are model-based and the last two are model-free.

2.1.1   Time-Series Models

Time-series models are commonly used for modeling an autocorrelated process. Equation (2.1) gives an expression of the ARMA time-series model:

$$\Phi(B)X_t = \Theta(B)\varepsilon_t \,, \tag{2.1}$$

where $\Phi(B)$ and $\Theta(B)$ are the ARMA polynomials in the backshift operator with roots outside the unit circle (e.g., for an ARMA(1,1) model, $\Phi(B) = 1 - \phi B$ and $\Theta(B) = 1 - \theta B$), ε_t represents white noise with mean 0 and standard deviation σ_a.

Ideally, if a proper time-series model is available for an autocorrelated process, the residuals can be approximately uncorrelated and monitored by traditional SPC methods (Harris and Ross, 1991; Loredo et al. 2002). However, when a signal occurs in the form of process mean shift as represented in Equation (2.2),

$$X_t = \frac{\Theta(B)}{\Phi(B)}\varepsilon_t + \gamma f(t) \,, \tag{2.2}$$

where γ is the size of a signal, and $f(t)$ is the function that indicates the nature of the signal, then the residual is no longer white noise and change patterns of mean shift signals need to be considered. Hu and Roan (1996) illustrated the change patterns of a process mean shift signal in the residual.

The drawback of this approach are that a time-series model may not be readily available or that the model identified in Phase I does not adequately describe the process. Therefore the process cannot be effectively decorrelated and false alarm rate is still higher than expected.

2.1.2 Integration of SPC and EPC

In some dynamic processes, controllers or adjustment variables are available for the application of a time-series-based engineering process control (EPC) scheme. When this is the case, the traditional SPC approaches can be applied to effectively monitor the quality characteristics of the outputs (Montgomery et al. 1994; Del Castillo, 2002; Box et al., 1994; Nembhard and Valverde-Ventura, 2003, 2006).

The most commonly used EPC strategy is minimum mean square error (MMSE) feedback control, which seeks to minimize the variability of the output error in a manner that is cost-insensitive to controllable factors in the input. In Chapter 3 it will be shown that the MMSE controller performs like an inverse ARMA filter on both the noise and the signal and removes the process autocorrelation completely.

The approach of integrating SPC and EPC is more effective in detecting mean shift signals in an autocorrelated process than traditional SPC methods. However, as is the case with to the time-series model approach, its performance relies heavily on the fitted time-series model.

Moreover, in some practical cases, an MMSE control strategy may not be economical if frequent, large, and/or costly adjustments are needed. In these cases, suboptimal control strategies, such as constrained MMSE (CMMSE) and generalized

minimum variance (GMV) control (Clarke and Gawthrop, 1975), are more suitable for transferring some variability from the controllable input variables to the output quality characteristic. Our review of the literature suggests that effective integration of SPC methods with suboptimal EPC schemes has not been reported.

2.1.3 EWMA and Moving Center-line EWMA (MCEWMA) Control Charts

Some researchers credit the EWMA and MEWMA approaches with overcoming the influence of an improperly identified time-series model on the performance of SPC charts. Montgomery (2005) categorized these model-based approaches. In order to avoid unnecessary complexity, we prefer to classify these approaches as model-free because no explicit time-series model is needed.

Montgomery and Mastrangelo (1991) and Mastrangelo and Montgomery (1995) proposed a moving center-line EWMA (MCEWMA) approach to fit EWMA statistics to the observation values in order to minimize the one-step ahead prediction error, and thus combined the information of statistical control and process dynamics on a single control chart. They estimated the standard deviation of the one-step-ahead errors or model residuals by the mean absolute deviation (MAD). The MCEWMA approach is ideal for decorrelating IMA stochastic processes.

Lu and Reynolds (1999a, b) used EWMA control charts to monitor the mean and/or variance of autocorrelated process. They discussed extensively their study of both the EWMA applied directly to the data and the EWMA of the residuals. They also compared the performance of several types of control charts, such as EWMA and

Shewhart charts, and combinations of control charts in monitoring autocorrelated processes.

2.1.4 Batch-Means Control Charts

Runger and Willemain (1995) proposed batch-means control charts to decorrelate the data by averaging the observation values in nonoverlapping windows. The window size is selected such that the means in each window are approximately uncorrelated. Thus, higher correlation will require a longer window size. Significant limitations to this approach are that it cannot detect a shift in a shorter period than the window length, and the window size may have to be determined empirically.

2.1.5 Multiscale SPC (MSSPC) Control Charts

Conceptually, traditional control charts are all single scale monitoring approaches which are suited to detect specific types of process disturbances, e.g., a Shewhart chart is suited for detecting large shifts at small scales, whereas EWMA and CUSUM charts are more suited for small shifts at large scales. Traditional single scale monitoring approaches have limitation when being used for monitoring autocorrelated processes (Ganesan et al., 2004).

Based upon favorable properties of wavelets, an autocorrelated process can be decomposed into uncorrelated wavelet coefficients (Vidakovic, 1999). Bakshi (1998) proposed a multiscale SPC (MSSPC) approach that transforms a process into multi-scales using wavelets, and then monitors the wavelet coefficients at each scale with a Shewhart

chart. Using this method for univariate processes, Aradhye et al. (2003) examined the performance of MSSPC in comparison to the methods of weighted batch means, MCEWMA, and the time-series model for a stationary autocorrelated process. This study was repeated for a nonstationary process using MCEWMA. The multiscale performance has been shown to be in between the weighted batch means and the residuals in the stationary case, and better than MCEWMA at large shifts for the nonstationary scenario.

2.1.6   Wavelet Scalogram

The wavelet scalogram provides measures of signal energy at various frequency bands. Theoretically, when a large transient mean shift fault signal occurs, energy will leak from higher levels of wavelet coefficients to lower levels. This phenomenon can be calculated in terms of energy and represented using wavelet scalogram (Vidakovic, 1999). Combining this property with the decorrelation capacity of wavelets for autocorrelated processes, a wavelet scalogram can be applied for detecting large transient shifts in autocorrelated or cyclic processes which is illustrated in Kamarthi et al. (2000) for detecting broken teeth of a gearbox.

Jeong (2004) and Jeong et al. (2003, 2006) extended the scalogram's capacity in handling noisy and possibly massive data based on fast wavelet transformation. The authors investigated the asymptotic properties of thresholded scalograms and derived point-wise confidence intervals of thresholded scalograms to construct control limits for fault detection and classification. However, some unsolved problems, such as the determination of the maximum depth of wavelet decomposition and the width of the testing window, limit the effectiveness of this approach.

2.2   SPC Approaches for Multivariate Autocorrelated Processes

Traditional univariate SPC control charts have been extended to fit multivariate scenarios because of the high demand for monitoring and controlling multivariate processes. Dealing with process autocorrelation in multivariate processes is much more complicated than in univariate processes.

2.2.1   Traditional Multivariate Control Charts for Monitoring Process Mean Shifts

A review of multivariate time series models and some major traditional multivariate control procedures for monitoring the process mean shift is summarized in this section. Their details will be presented in Chapter 4.

As the extension of univariate framework, the multivariate time series modeling techniques can be used to remove the autocorrelation structure in observations collected from a multivariate process. Montgomery et al. (1990) and Hamilton (1994) introduced the multivariate time series model in vector ARMA (VARMA) form and related it to the state-space model. Noorossana and Vaghefi (2005) illustrated the use of the first-order vector AR(1) time-series model to decorrelate the multivariate process and the application of multivariate control chart to monitor the residuals.

Hotelling's T^2 control chart is one of the earliest techniques in multivariate process control (Hotelling, 1947). It gives a general metric for turning measurement vectors into scalars that retain the essential information about whether the mean vector is indeed in control or not. A useful extension of this basic approach is discussed in Hawkins and Olwell (1998).

Healy (1987) derived a multivariate Cusum (MCusum) procedure based on the sequential likelihood ratio test of multivariate variables on the scale of Hotelling's T^2 statistic. Crosier (1988) proposed a MCusum procedure that accumulates on the scale of the observations $\mathbf{X}$, instead of $\mathbf{X}^2$ or T^2. Noorossana and Vaghefi (2005) applied the MCusum control chart to monitor the residuals from a vector AR(1) time series model.

Lowry et al. (1992) presented a multivariate EWMA (MEWMA) control chart procedure as a logical extension of the univariate EWMA. Kramer and Schmid (1997) applied the MEWMA to the residuals from a vector AR(1) time series model. Mastrangelo and Forrest (2002) developed an adaptive approach to monitor autocorrelated processes using a MEWMA residual.

2.2.2   Multiscale SPC Control Chart

As the number of process variables increases, traditional SPC approaches are not practical and they lack multivariate extensions. Principal component analysis (PCA) is an important technique for reducing the dimension of multivariate processes. Based on the wavelet transform and PCA, Bakshi (1998) proposed a multiscale SPC (MSSPC) approach for monitoring multivariate processes at multi-scales. He illustrated that a MSSPC can deal with autocorrelated observations and multivariate problems.

2.3 Control Chart Approaches for Profile Monitoring

Profile monitoring is a new and rapidly developing area in statistical process control, and little research has been developed in using SPC charts to monitor profile data. Woodall et al. (2004) gave a thorough review of this topic. In this section, the focus is on using multivariate control charts and nonparametric approaches for profile monitoring.

2.3.1 Multivariate Control Charts for Profile Monitoring

In many applications, the structure of the profile model is determined based on scientific research or engineering knowledge. Therefore, traditional multivariate control charts can be applied to monitor the regression parameters for the fitted profile model.

Linear profile monitoring has attracted the most research attention in recent years. Kang and Albin (2000) proposed using two approaches for monitoring linear profiles. The first is to apply a bivariate T^2 chart to monitor the slope and the intercept; the second is to use a EWMA chart to monitor the residual averages and an R chart to monitor the variance of the residuals along with the regression line. The approaches were recommended for both Phase I and Phase II monitoring.

Kim et al. (2003) proposed a method for monitoring the linear profile data in Phase II. They transformed the estimators of intercept and slope to the coded independent variables and used three Shewhart control charts to monitor the intercept, the slope, and the variance of the deviations about the regression line respectively in Phase II. They

showed that their method had better performance than that of Kang and Albin (2000) in terms of average run length (ARL).

With the idea of comparing k regression lines collected in Phase I, Mahmoud and Woodall (2004) introduced k-1 indicator variables and constructed a multiple regression model to test whether the kth regression line is statistically significant based on F-test. They also recommended two Shewhart charts to monitor the coded intercept and slope variables by Kim et al. (2003) for fault diagnose purposes. They showed their approach had better performance by comparing it with three other approaches by using simulation, and illustrated its use with the real data from a calibration process.

The multivariate control chart can be applied to monitor the fitted regression parameters of nonlinear profiles if the structure of the parametric model is known. For example, the "bathtub" function was suggested to model the mean function of the profiles in Figure 1.1 by William et al. (2003):

$$f(x_i, \boldsymbol{\beta}) = \begin{cases} a_1 (x_i - d)^{b_1} + c & x_i > d \\ a_2 (-x_i + d)^{b_2} + c & x_i \le d \end{cases} \tag{2.3}$$

where a_1 and a_2 are the width parameter of the "bathtub", b_1 and b_2 are the flatness parameters, c is the bottom and d is the center of the "bathtub". William et al. (2003, 2004) gave a thorough discussion on monitoring such nonlinear profiles using Hotelling's T^2 statistics of the regression parameters with the comparison of three estimators of the covariance matrix.

2.3.2 Profile Monitoring Using Nonparametric Regression

In monitoring complicated profiles that are not smooth and have an unknown mean profile function $m(X_k, \boldsymbol{\beta})$ as in Equation (1.1), nonparametric regression methods, such as Fourier transforms, wavelets, splines and local polynomial regression, are valuable in smoothing the profiles for comparison.

Walker and Wright (2002) used an additive model to fit the profile data for assessing the source of variation in the vertical density profile data. The additive model contains a B-spline component for smoothing each profile data and a parametric portion for incorporating other sources of variation. However, such methods cannot explicitly be used for setting up control charts for process monitoring.

Jin and Shi (1999, 2001) proposed using wavelet modeling to fit complicated profiles that have sharp corners containing the most useful information. They relied on engineering knowledge as a prior or "oracle" to determine the local segment for fault diagnosis purposes. Fan (1996) introduced two hypothesis testing techniques for high-dimensional data, wavelet thresholding and adaptive Neyman's (AN) truncation of Fourier coefficients. These two approaches provide the statistical basis for setting up SPC charts for monitoring high-dimensional data. Jeong et al. (2004) applied wavelet thresholding techniques to monitor complicated profiles by automatically selecting the significant variables for tests. His research was based on one of the approaches to test significance proposed by Fan (1996). Further, Fan and Lin (1998) illustrated how the two procedures can be applied to test the differences between two sets of curves with i.i.d. noise or even stationary noise fitted by an ARMA model by capitalizing on the fact that the impact of the stationary errors on the null distribution is asymptotically negligible.

Spitzner and Woodall (2003) compared classical multivariate testing approaches with the AN test of Fan and Lin (1998). They applied the AN method to the Fourier coefficients of the vertical density profile data and thickness profile data for silicone nitride film in Gardner et al. (1997).

Many profile monitoring approaches are based on the techniques for checking the distribution of a dataset or comparing two datasets in terms of their distribution model. The research by Fan (1996) is in this category. Wang and Tsung (2004) proposed a profile monitoring technique by comparing the slope and intercept of the Q–Q plot to monitor the distributions of the residuals from different profiles.

Cuscore Control Charts for Generalized Feedback Control Systems

By its design, the cumulative score (Cuscore) statistic is able to "resonate" with deviations in signals of an *expected* type. When a process signal subject to feedback control occurs, it results in a *fault signature* in the output error. In this chapter, Cuscore statistics are designed to monitor process parameters and characteristics measured by a generalized minimum variance (GMV) feedback control system sensitive to the fault signature of a spike, step, and bump signal.

The GMV control system considered in this investigation is a first-order dynamic system with auto-regressive moving average (ARMA) noise. It is shown theoretically that the performance of Cuscore charts is independent of the amount of variability transferred from the output quality characteristic to the adjustment actions in the GMV control system. Simulation is used to test the performance of using the Cuscore charts.

Generally, the Cuscore can detect signals over a broad range of system parameter values. However, areas of low detection capability occur for certain fault signatures. In these cases, a tracking signal test is combined with the Cuscore statistics to improve detection performance. This investigation provides several illustrations of the underlying behavior and shows how the methodology developed can be easily applied in practice.

3.1   Introduction

The aim of conventional statistical process control (SPC) is to monitor a process to detect aberrant behavior. Deming (1986) called such aberrant behavior *special causes* that are suggested by data patterns that indicate the existence of systematic *signals*. The timing, nature, size, and other information about these signals can help identify and potentially eliminate signaling factor(s) by interfacing with an out-of-control action plan. Conventional Shewhart charts provide this capability in cases where the detected signal is an *unexpected* spike in white noise. By contrast, in many situations certain process signals are *anticipated* because they are characteristic of a system or operation. In these cases the Cuscore chart is the appropriate approach; it can be devised to be especially sensitive to deviations or signals of an expected type.

In general, after working with a particular process, engineers and operators have often experienced how a process will falter. Despite repeated experience, however, the problem seldom announces its time and location in advance. Consider, for example, a process where a valve is used to maintain pressure in a pipeline. Because the valve will experience wear over time, it must be periodically replaced. However, in addition to the usual wear, engineers are concerned that the valve may fatigue or fail more rapidly than normal. The Cuscore can be used to incorporate this working knowledge (failure mode analysis) and experience into the statistical monitoring function. This concept often has a lot of intuitive appeal for industry practitioners.

To improve quality engineering, many researchers have studied integrating SPC to monitor a process with engineering process control (EPC) to make process

adjustments. For example, Box and Luceño (1997) suggested that "process monitoring and process adjustment are two complementary strategies in the maintenance and improvement of quality and productivity and the integration of SPC and EPC can result in major improvements in industrial efficiency." Traditionally, a key assumption for using SPC is that the successive values of the quality characteristics are independent and identically distributed (i.i.d.). Therefore if SPC charts are misapplied to a highly autocorrelated process, their successful performance on detection rates and false alarm rates is likely to decrease. On the other hand, process adjustment using EPC actually relies on autocorrelated process noise in order to minimize output variability by adjusting compensatory processing variables.

The EPC strategy of minimum mean square error (MMSE) feedback control is to minimize the variability of the output error in a manner that is cost-insensitive to controllable factors in the input. It is shown that the MMSE controller performs like an inverse ARMA filter on both the noise and the signal and removes the process autocorrelation completely citation. The drawback of MMSE control is that excessive manipulation of the input control actions often occurs in order to achieve the minimized output errors. To avoid this problem, constrained MMSE (CMMSE) control was proposed (see, Box et al., 1994). In this approach, the mean square error (MSE) of the output is minimized subject to a constraint on the variance of the controllable input factor. Some variability can thus be transferred from the output quality characteristic to the controllable input factor by the operation. Clarke and Gawthrop (1975) proposed a simpler approach to constrain the variability in the controllable factor, called generalized minimum variance (GMV) control.

When a MMSE controller or GMV controller is used to minimize the autocorrelation in the noise component of a disturbance given by an ARMA time-series model, the signal component of the disturbance exhibits a pattern in the residual which is serially correlated and temporally variable. This change pattern is referred to as the *fault signature* (Apley and Shi, 1999). Hu and Roan (1996) studied the fault signatures for two types of signals, the spike and the step, by using the inversed AR(1), ARMA(1,1) and ARMA(2,1) filters which are equivalent to MMSE controllers. Other authors, including Tsung and Tsui (2003), Apley and Shi (1999) and Luceño (2004), have also considered the fault signatures for step or spike signals in the output processes adjusted by MMSE controllers.

In a first-order dynamic system with autocorrelated noise, minimum mean square error (MMSE) feedback control can filter the noise so that the adjusted output errors become independent and identically distributed (i.i.d.) white noise series. However, in a generalized minimum variance (GMV) control system, the adjusted output error includes white noise plus some of variability in the output quality characteristic transferred to the controllable input factor. A review of the literature showed that the fault signatures for signals in the GMV-controlled system were not addressed.

In integrating SPC and EPC, traditional residual-based control charts, such as the Shewhart, Cusum, and EWMA charts, are unable to benefit from the dynamic properties of the fault signature. However, it is possible to achieve higher detection capabilities by incorporating the property information of the noise and the signals using time series-based control charts or methods such as Cuscore charts (Capilla et al., 1999; Luceño, 1999, 2004; Nembhard and Valverde-Ventura, 2003, 2006; Tsung and Tsui, 2003 and

Nembhard and Changpetch, 2005) and generalized likelihood ratio tests (GLRTs) (Apley and Shi, 1999 and Runger and Testik, 2003).

Motivated by the Cuscore control chart's ease of use and sensitivity to the appearance of transient disturbances in an industrial process, Shao (1998) examined the feasibility of using the Cuscore control chart as an interface for integrating EPC and SPC techniques. He demonstrated the effective incorporation of Cuscore and EPC techniques and used simulation to confirm the superiority of integrated SPC and EPC schemes over control schemes employing EPC alone. He also suggested the superiority of the Cuscore/EPC scheme to other SPC/EPC integration approaches in certain circumstances. However, he only explored the case of MMSE-controlled non-stationary IMA(1,1) noise with ramp and step signals in a steady state system.

In this chapter, the Cuscore chart is integrated with GMV feedback control in order to develop a more general model for SPC/EPC for systems with anticipated signals. This investigation specifically considers spike, step, and bump signals in a ARMA(1,1) noise model and analyzes the fault signatures for the signals in differently weighted GMV control processes. It was determined that Cuscore charts performed independently of the amount of variability transferred from the output quality characteristic to the adjustment actions in the GMV control system. However, for some patterns of noise, the Cuscore chart has a low detection capability which suggests the addition of forecasting methods.

In Section 3.2, the background on monitoring and control is briefly reviewed with a focus on the GMV controller and Cuscore chart. In Section 3.3, the approach for integrating these two concepts is presented. In Section 3.4, simulation is used to analyze

performance of the integrated Cuscore monitoring/GMV control for the spike, step, and bump signals. Section 3.5 illustrates how to use the methodology to detect leakage in a valve pipeline system. Concluding remarks are in Section 3.6.

3.2   Background on Monitoring and Control

Figure 3.1 provides a diagram of integrating SPC/EPC where the control chart is applied to the quality characteristic. This working model shows a dynamic system that has controllable input X_t and quality characteristic output y_t. In a dynamic system, there will be a time delay between when X is changed and when that change is realized in the output y. (By contrast, in a responsive (or steady-state, zero-ordered) system, a change in X is realized immediately in y.) In addition, this working model shows that both a controller, represented by the control equation, and a monitor, represented by the Cuscore monitoring chart, use the information available at the output to test and adjust the input.

More specifically, the transfer function that relates the change in X_t to the response in y_t is

$$y_t = \frac{B_1(B)B^k}{A_1(B)} X_t \ ,$$

(3.1)
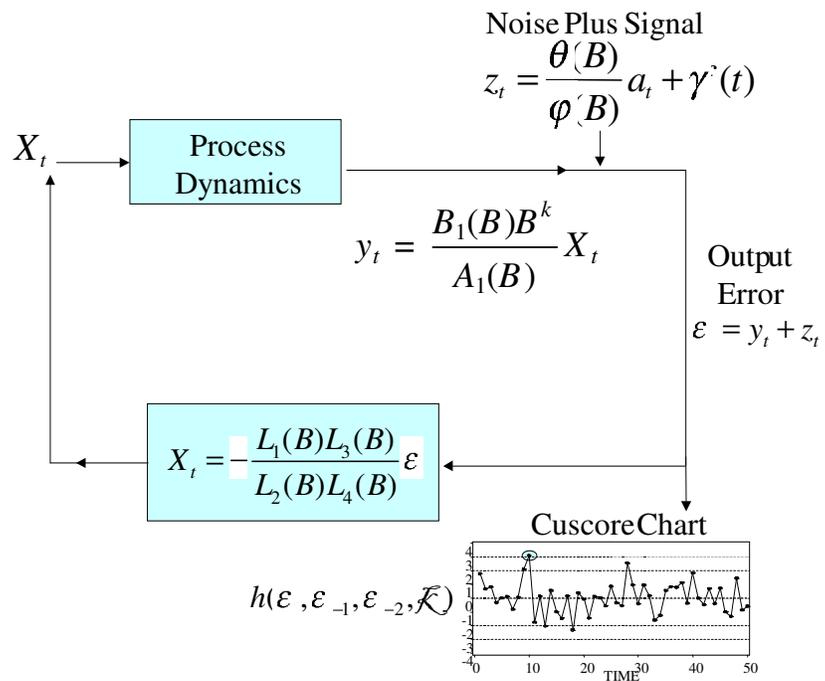
Figure 3.1: A block diagram showing the input, output, and disturbance components and the relationship between feedback control and Cuscore monitoring of an anticipated signal.

where B is the backshift operator, A_1 and B_1 are polynomials in B with roots outside the unit circle to assure stability and invertibility of the process, and k is an integer indicating the delay of a change in X to induce a change in y (in a responsive system, $k = 0$; in a dynamic system, $k \geq 1$) (Box, Jenkins, and Reinsel, 1994). Independently of this, a disturbance is given by the noise described by a time-series model *plus* an anticipated *signal* that could appear at some time, namely

$$z_t = \frac{\theta(B)}{\phi(B)} a_t + \gamma f(t), \tag{3.2}$$

where $\phi(B)$ and $\theta(B)$ are the ARMA polynomials in the backshift operator with roots outside the unit circle (e.g., for an ARMA(1,1) model, $\phi(B) = 1 - \phi B$ and $\theta(B) = 1 - \theta B$), a_t represents white noise with standard deviation σ_a, γ is the size of a signal, and $f(t)$ is the function that indicates the nature of the signal. Together, $\gamma f(t)$ corresponds to an assignable cause of a specific type that could occur in the system at any time. The combination of the process dynamics and the disturbance gives the output error:

$$\varepsilon_t = \frac{B_1(B)}{A_1(B)} X_{t-k} + \frac{\theta(B)}{\phi(B)} a_t + \gamma f(t). \tag{3.3}$$

The output error is then used in two ways: (1) to control the input by specifying the next level of X_t; and (2) to construct the Cuscore chart to monitor the process output. A discussion of each of these concepts follows.

3.2.1 Feedback Control and Fault Signatures

The MMSE feedback control scheme has considerable theoretical appeal because it is an optimization procedure to minimize variability. However, it is often impractical since it is cost-insensitive (Del Castillo, 2002). Consider a feedback control system with a nonstationary IMA(1,1) disturbance. The adjusted process tends to diverge as the system runs for a long time. In practice, suboptimal schemes, such as CMMSE control, can be used so that fewer compensatory adjustments are required at the expense of a small increase in MSE. The CMMSE control schemes can be obtained by minimizing

$$J_1 = MSE(\varepsilon_t) + \lambda_0 MSE(X_t), \tag{3.4}$$

where λ_0 can be regarded as an undetermined multiplier that allocates the relative quadratic costs of variations of output error ε_t and input adjustments X_t (see Box et al., 1994). To minimize the objective J_1 in a simpler way while keeping the same constraint effect, Clarke and Gawthrop (1975) proposed to minimize

$$J_2 = \hat{\varepsilon}_{t+k|t}^2 + \lambda_0 X_t^2, \tag{3.5}$$

where $\hat{\varepsilon}_{t+k|t}$ is the estimated output error at time $t+k$ given that the errors are known before time t (see Del Castillo, 2002 and Capilla et al., 1999). For applying the control law, Equation (3.3) (apart from the signal) can be expressed as

$$A_1(B)\phi(B)\varepsilon_t = B_1(B)\phi(B)X_{t-k} + \theta(B)A_1(B)a_t,$$

which can be reduced to the ARMA exogenous variable (ARMAX) form:

$$A(B)\varepsilon_t = B(B)X_{t-k} + C(B)a_t, \tag{3.6}$$

where $A(B)$, $B(B)$ and $C(B)$ are the new time series model polynomials (Del Castillo, 2002). The GMV controller for the model in Equation (3.6) is

$$X_t = -\frac{G(B)}{B(B)F(B) + (\lambda_0 / b_0)C(B)}\varepsilon_t, \tag{3.7}$$

where λ_0 is a weight that measures the variability transferred to the adjustment actions from the quality characteristic (and so it is referred as the "variance transfer parameter"), b_0 is the first coefficient in the $B(B)$ polynomial, and $F(B)$ and $G(B)$ are polynomials from forecasting (see Del Castillo, 2002, for details). When λ_0 is 0, the GMV control scheme reduces to MMSE control.

Hu and Roan (1996) discussed the change patterns fault signatures for a spike and a step signal after being filtered using AR(1), ARMA(1,1) and ARMA(2,1) models. They illustrated that out-of-control process data can be decomposed into two parts: a stationary time-series process and a change function. When the process is filtered using a time series model, the result is a sum of the response due to the stationary time series, which is white noise (for a properly fit ARMA model), and the response due to the change function, which is the fault signature. Similarly, the output of a feedback controlled process is this same type of summation. However, in a first-order dynamic system, the response due to the time series is white noise only in the special case of MMSE control. The response due to the time series in the case of GMV feedback control may be called "colored noise".

3.2.2 Process Monitoring

A proven approach in monitoring autocorrelated data is to create the correlative structure with an appropriate time series model, then use that model to remove the autocorrelation, and finally apply SPC techniques to the residuals (see e.g., Alwan and Roberts, 1988). These residuals may be viewed as MMSE-adjusted output error values composed of two parts: a white noise and a fault signature that can be monitored with residual-based SPC charts, such as the Shewhart, EWMA, or CUSUM chart. However, such SPC charts neglect the dynamic properties in the fault signature when the noise parameters can be obtained or estimated before making feedback adjustment.

Box and Ramírez (1992) proposed a Cuscore chart based on the likelihood ratio test to identify suspected deviations known to be characteristic of (or peculiar to) a monitored system. They suggested that if a model can be expressed in the form

$$a_i = a_i(y_i, X_i, \gamma) \; i = 1, 2, ..., t \,, \tag{3.8}$$

where the y_i are observations (on the quality characteristic), the X_i are known (controllable input) quantities, γ is some unknown parameter (of the signal), the a_i's are independent normal random variables with mean 0 and variance σ_a^2 (i.e., white noise). If σ_a is known and does not depend on γ, then the logarithmic likelihood is given by

$$l = -\frac{1}{2\sigma^2} \sum_{i=1}^{t} a_i^2 + c, \tag{3.9}$$

where c is a constant that does not depend on γ.

Following Fisher (1925), the efficient score statistic is obtained from Equation (3.9) by differentiating with respect to γ at $\gamma = \gamma_0$. Thus

$$\left.\frac{\partial l}{\partial \gamma}\right|_{\gamma=\gamma_0} = \frac{1}{\sigma^2} \sum_{i=1}^{t} a_{i0} r \quad \text{with } r_i = -\left.\frac{a_i}{\gamma}\right|_{\gamma=\gamma_0}, \qquad (3.10)$$

where the null values, a_{i0}, are obtained by setting $\gamma = \gamma_0$ in Equation (3.10), and the series of r_i's is referred to as the *detector*. Thus

$$Q = \sum_{i=1}^{t} a_{i0} r \qquad (3.11)$$

is the Cuscore associated with the parameter value $\gamma = \gamma_0$.

In some cases, two-sided Cuscore statistics are preferred to monitor the bias of signals in a process. Let the superscript $+$ and $-$ denote positive and negative biases respectively. The formula can be written as

$$\begin{aligned} Q_t^+ &= \max(0, Q_{t-1}^+ + a_{t0} r_t), \\ Q_t^- &= \min(0, Q_{t-1}^- + a_{t0} r_t), \\ Q_0^+ &= Q_0^- = 0. \end{aligned} \qquad (3.12)$$

As shown in Box and Luceño (1997) and Nembhard (2005), the Cuscore chart designed for detecting a spike signal in white noise is equivalent to a traditional Shewhart chart; the Cuscore chart designed for designed for detecting a step signal in white noise is equivalent to a CUSUM chart; and the Cuscore chart designed for detecting a bump signal in white noise is equivalent to an arithmetic moving average (AMA) chart.

3.3 Cuscore Statistics for GMV Feedback Controlled Processes

As stated in the chapter introduction, the approach used here is to integrate the Cuscore chart with a GMV control to develop a more general model for SPC/EPC for systems with anticipated signals. Specifically, the objective is to monitor the output characteristic from a GMV feedback control system when known signals appear. To do this we consider a noise disturbance described by the ARMA(1,1) model with a signal in a first-order dynamic process as described in Equation (3.3) and controlled by a GMV control scheme as described in Equation (3.7). The output is composed of two parts, the control action filtered by the first-order dynamic transfer function, and the noise and signal:

$$\varepsilon_t = \frac{g}{1-\delta B} X_{t-1} + \frac{1-\theta B}{1-\phi B} a_t + \gamma f(t) , \qquad (3.13)$$

where γ is the steady-state system gain. The GMV controller for this system is

$$X_t = \frac{(\theta - \phi)(1-\delta B)}{g(1-\phi B) + \dfrac{\lambda_0}{g}(1-\theta B)(1-\delta B)} \varepsilon_t . \qquad (3.14)$$

Applying the GMV controller to the system by substituting Equation (3.14) into Equation (3.13), then rearranging gives

$$\varepsilon_t = \frac{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)} \left(a_t + \frac{1-\phi B}{1-\theta B} \gamma f(t) \right) . \qquad (3.15)$$

We note that for MMSE control, which is equivalent to GMV control with $\lambda_0 =$ 0, Equation (3.15) reduces to

$$\varepsilon_t = a_t + \frac{1 - \phi B}{1 - \theta B} \gamma f(t) = a_t + \gamma \tilde{f}(t), \tag{3.16}$$

where $\gamma \tilde{f}(t)$ is the fault signature for the signal $\gamma f(t)$. Equation (3.16) shows that the fault signature is dependent on the ARMA parameters ϕ and θ but independent of the system dynamics parameter δ for an MMSE feedback control scheme. The expectation of Equation (3.16) is

$$E[\varepsilon_t] = E[a_t + \gamma \tilde{f}(t)] = E[a_t] + E[\gamma \tilde{f}(t)] = E[\gamma \tilde{f}(t)] = \gamma \tilde{f}(t). \tag{3.17}$$

Equation (3.17) shows that the expected output from a MMSE feedback control system is the fault signature of fault $f(t)$ which is equivalent to the change patterns of a signal filtered by an ARMA(1,1) model as discussed in Hu and Roan (1996) with examples of spike and step signals. However, in the GMV control scheme, the fault signature is different due to the λ_0 parameter. Expanding Equation (3.17) shows that the output error is composed of two parts: the inflated noise and the fault signature

$$\varepsilon_t = \frac{g^2(1 - \phi B) + \lambda_0(1 - \phi B)(1 - \delta B)}{g^2(1 - \phi B) + \lambda_0(1 - \theta B)(1 - \delta B)} a_t + \frac{g^2(1 - \phi B) + \lambda_0(1 - \phi B)(1 - \delta B)}{g^2(1 - \phi B) + \lambda_0(1 - \theta B)(1 - \delta B)} \gamma \tilde{f}(t). \tag{3.18}$$

In order to align with common system disruptions, it seems most useful to consider monitoring and control for the cases for a spike, step, and bump signal. For each of these three signals, we show the form of the Cuscore in terms of the parameters of the

control model and in terms of the fault signature (the derivations are in the Appendices 3A - 3C).

Spike signal:

Form in terms of the parameters of the control model:

$$Q_t = \sum_A a_{t0} r = a_{t0} r_{t0} = a_{t0} = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t \qquad (3.19)$$

Form in terms of the fault signature:

$$Q_t = a_t + \frac{1-\phi B}{1-\theta B} \gamma f(t) \qquad (3.20)$$

Step signal:

Form in terms of the parameters of the control model:

$$Q_t = \sum_A a_{t0} r = \sum_A \left[\frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t \right] \left[1 + \frac{\theta - \phi}{1-\theta} \right] \qquad (3.21)$$

Form in terms of the fault signature:

$$Q_t = \sum_A \left[a_t + \frac{1-\phi B}{1-\theta B} \right] \left[1 + \frac{\theta - \phi}{1-\theta} \right] \qquad (3.22)$$

Bump signal (duration = 3 periods):

Form in terms of the parameters of the control model:

$$Q_t = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} (\varepsilon_t r_{t0} + \varepsilon_{t-1} r_{t0-1} + \varepsilon_{t-2} r_{t0-2}) \qquad (3.23)$$

where

$$r_{t_0} = 1 + (\theta - \phi)(1 + \theta)$$
$$r_{t_0-1} = 1 + \theta - \phi$$
$$r_{t_0-2} = 1$$

Form in terms of the fault signature:

$$Q_t = \left(a_t + \frac{1-\phi B}{1-\theta B} \gamma f(t) \right) r_{t_0} + \left(a_{t-1} + \frac{1-\phi B}{1-\theta B} \gamma f(t-1) \right) r_{t_0-1} + \left(a_{t-2} + \frac{1-\phi B}{1-\theta B} \gamma f(t-2) \right) r_{t_0-2} \quad (3.24)$$
$$= (a_t + \gamma r_{t_0}) r_{t_0} + (a_{t-1} + \gamma r_{t_0-1}) r_{t_0-1} + (a_{t-2} + \gamma r_{t_0-2}) r_{t_0-2}$$

Equations (3.20), (3.22), and (3.24) show that the Cuscore statistics used in monitoring the three signals in the GMV feedback control system are composed of two parts: the white noise and the fault signature of the signals. The latter is independent of the GMV parameter λ_0 and the system dynamic parameter δ.

One can build an understanding of the behavior of the control and monitoring approach by first considering the fault signature of the signals when subject only to MMSE control, then considering the fault signature of the signals when subject only to GMV control, and finally by considering a side-by-side comparison of the fault signatures and Cuscore statistics for a process adjusted by a GMV controller.

Figures 3.2, 3.3, and 3.4 show the fault signatures of a spike, a step, and a bump that occur at time $t = 10$. (To facilitate comparison among the different scales, the figures show ±3 unit upper/lower lines; they are not defined as formal control limits.) It can be observed that in Figure 3.2 and Figure 3.4 the size of the first observation after the change time is constantly equal to the spike size. The fault signatures then gradually

converge to zero in 5 to 20 time intervals. In Figure 3.3, the first observation after the change time is also constantly equal to the step size. Then, after a few intervals, some fault signatures remain constant (cases (a), (c), (d), and (e)), some converge or get very close to zero (cases (f), (g), and (h)), while one diverges (case (b)).

Figures 3.5, 3.6, and 3.7 illustrate the fault signatures for spike, step, and bump signals in a GMV controlled process. It can be observed that the more λ_0 increases, the faster the fault signatures converge.

Figures 3.8, 3.9, and 3.10 show the Cuscore statistics and output errors for a spike, a step, and a bump signal in a process adjusted by a GMV controller. It can be seen that for each type of signal, the output errors are different for different λ_0 values, but the Cuscore statistics remain same.
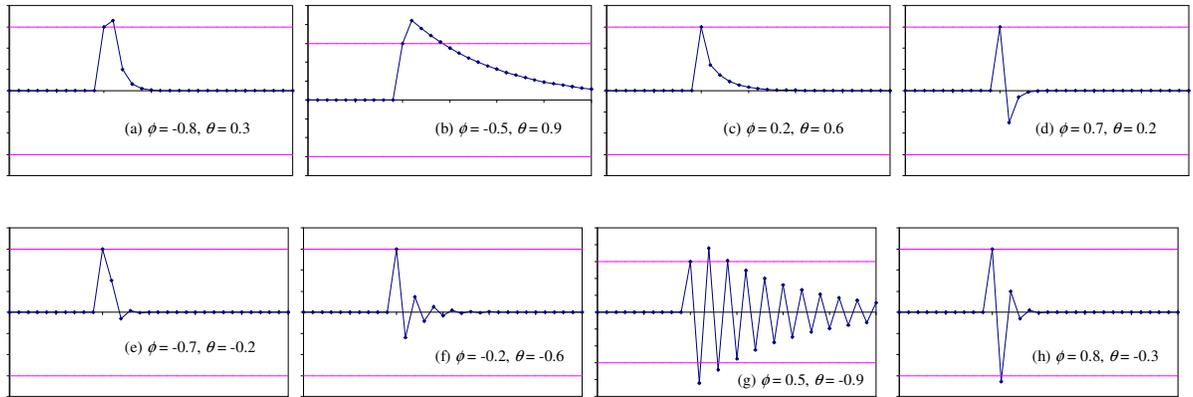
Figure 3.2: Fault signature of a spike signal in a MMSE controlled process (spike size = 3, $\delta = 0.5$).
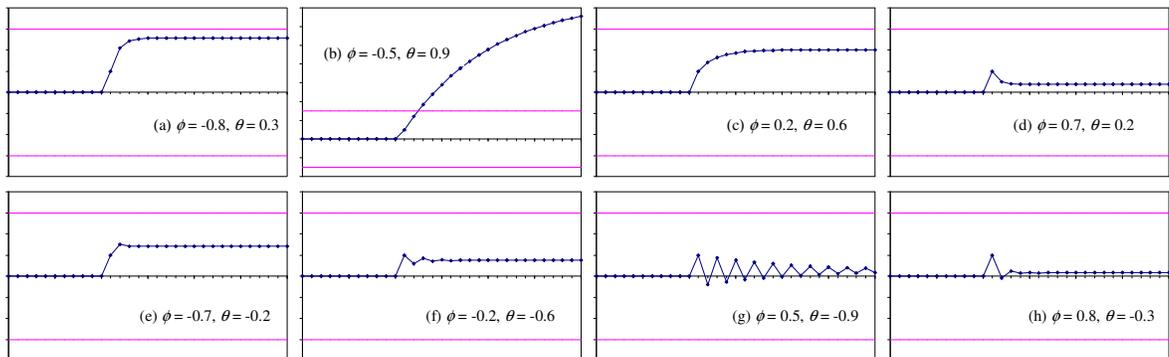
Figure 3.3: Fault signature of a step signal in a MMSE controlled process (step size = 1, $\delta = 0.5$).
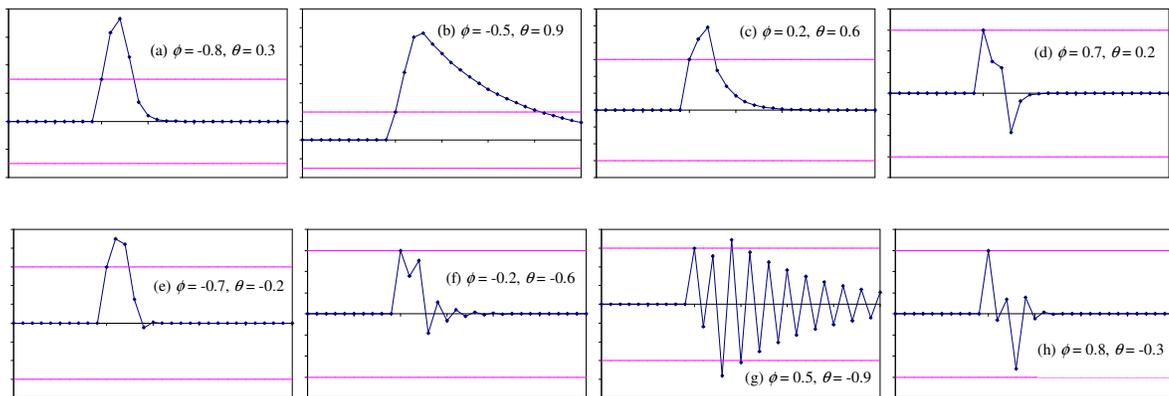
Figure 3.4: Fault signature of a three-period bump signal in a MMSE controlled process (bump size = 3, $\delta = 0.5$).

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$ (c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
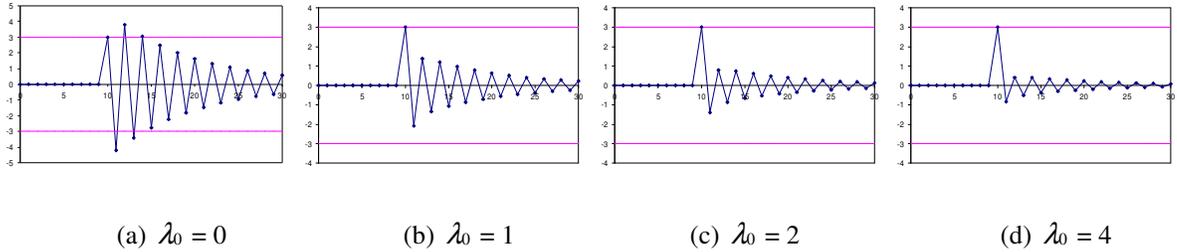
Figure 3.5: Fault signature of a spike signal in a GMV controlled process (spike size = 3, start at $t = 10$, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$ (c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
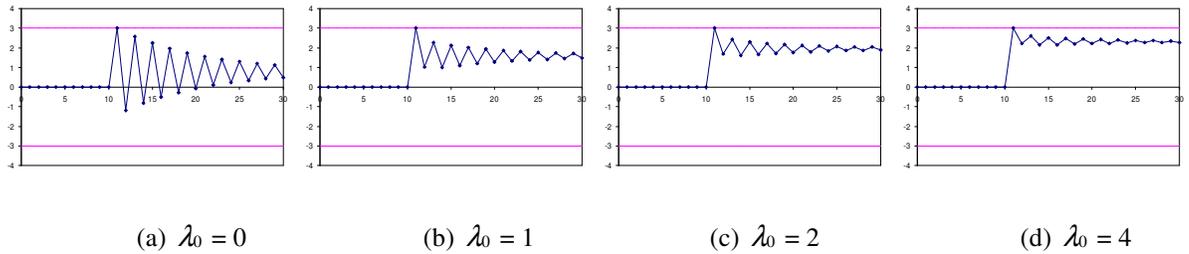
Figure 3.6: Fault signature of a step signal in a GMV controlled process (step size = 3, start at $t = 10$, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$ (c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
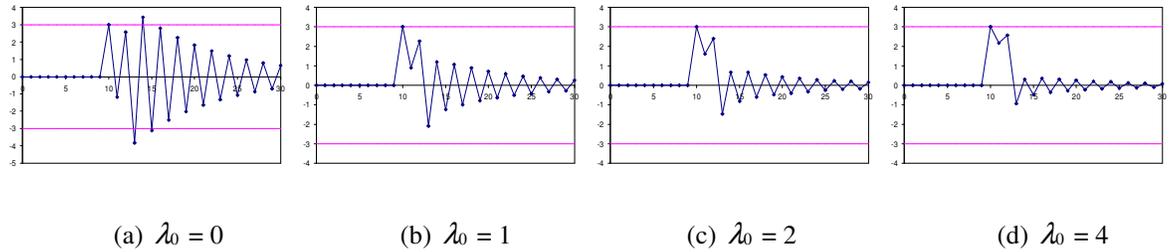
Figure 3.7: Fault signature of a bump signal for three time periods in a GMV controlled process (bump size = 3, start at $t = 10$, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

38

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$

(c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
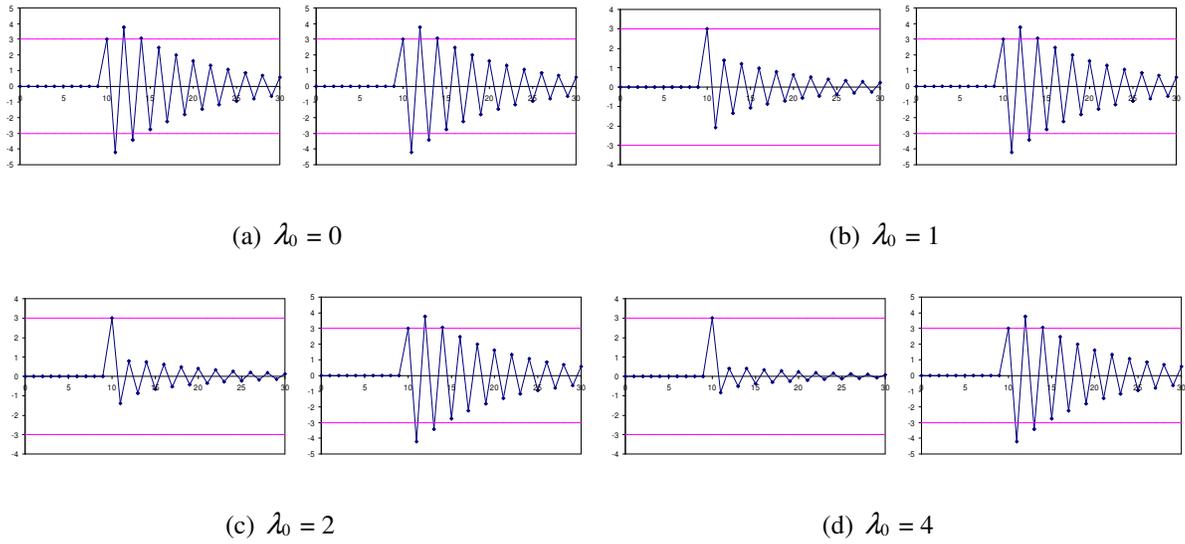
Figure 3.8: The fault signatures (in column 1) and Cuscore statistics (in column 2) for a spike signal adjusted by a GMV feedback controller in a first order dynamic process (spike size $= 3$, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$

(c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
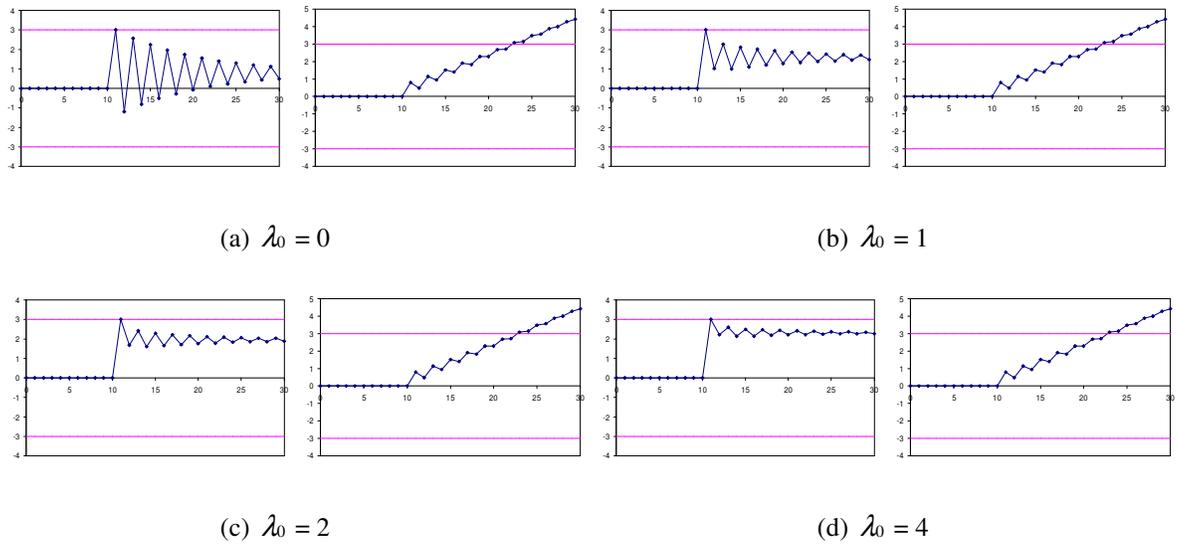
Figure 3.9: The fault signatures (in column 1) and Cuscore statistics (in column 2) for a step signal adjusted by a GMV feedback controller in a first order dynamic process (step size $= 3$, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

(a) $\lambda_0 = 0$ (b) $\lambda_0 = 1$

(c) $\lambda_0 = 2$ (d) $\lambda_0 = 4$
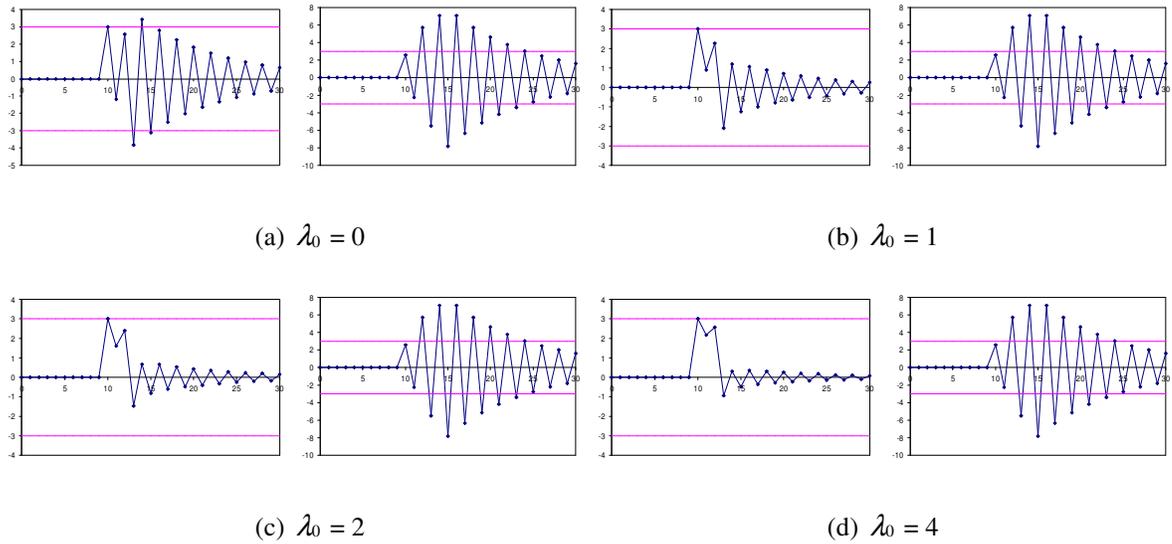
Figure 3.10: The fault signatures (in column 1) and Cuscore statistics (in column 2) for a step signal adjusted by a GMV feedback controller in a first order dynamic process (step size = 3, $\phi = 0.5$, $\theta = -0.9$, $\delta = 0.5$, $g = 1$).

3.4 Performance Results

In this section, we use Monte-Carlo simulation to investigate the performance of the integrated Cuscore/GMV control approach developed in Section 3.3. The focus will be on the spike, step, and bump signals, with an intention to establish the performance of the Cuscore to detect these signals in a GMV controlled process. The advantage of the Cuscore chart over traditional charts is its adaptability of detecting the fault signature of these of signals from the output error in a GMV feedback control system.

3.4.1 Performance for Detecting a Spike

The objective of this investigation is to detect the spike signal from output errors using the Cuscore chart with plus and minus three sigma control limits. In order to do

this, we consider a system of a first-order dynamic process with $\delta = 0.5$ where the noise disturbance is modeled by an ARMA(1,1) process over a range of values for the of θ and ϕ parameters. A spike signal of size 3 units occurs in the noise at time $t = 5$ and the process is adjusted by using GMV control with varying λ_0 values.

A MATLAB program simulated 10,000 runs with 20 observations for each run and computed the detection rates of the Cuscore chart for this controlled system. The results provided in the first row of Table 3.1 show detection rates that vary over a wide range depending upon the ARMA parameters. For the values of θ and ϕ in cases (c), (d), (e) and (f), the detection rates are about 0.5; in approximately half of the runs, the signal was not detected. For cases (b) and (g), the rate is 1; the signal never went undetected. For cases (a) and (h), the rate is over 0.8, a fair performance. The results also show that the GMV control parameter λ_0 has no significant effect on the detection rate which is consistent with the theoretical result in Equation (3.20).

Obviously the detection performance for cases (c), (d), (e) and (f) are inadequate since the signal was undetected in nearly half of the runs. Following Montgomery and Mastrangelo (1991) and Mastrangelo and Montgomery (1995), who used the supplementary tracking signal test to enhance the detection performance of control charts, tracking signals were applied to the Cuscore chart. Tracking signals use forecast errors to monitor and control a forecasting process (Montgomery et al., 1990). This present investigation used the smoothed error tracking signal (SETS)

$$T_s(t) = \left| \frac{Q(t)}{\hat{\Delta}(t)} \right| \tag{3.25}$$

where $Q(t) = \alpha e_t + (1-\alpha)Q(t-1)$ and $\hat{\Delta}(t) = \alpha|e_t(t)| + (1-\alpha)\hat{\Delta}(t-1)$; α is a smoothing constant, typically selected between 0.05 and 0.15; and e_t is the forecast error at time t. The value of the tracking signal statistic is compared to a constant K (Montgomery et al., 1990) suggest a value between 0.2 and 0.5) in order to test the hypothesis that the expected forecast error is zero.

A second set of simulation runs used a tracking signal with $\alpha = 0.1$ and $K = 0.5$ to monitor the Cuscore statistics in combination with the Cuscore charts. These results are shown in the second rows of Table 3.1. This combination method ideally improves the detection rate for cases (c), (d), (e) and (f) to 1. However, as with any test of hypothesis procedure, one should be cautious in using the tracking signal in practice because there will be higher false alarm rate as the trade-off for enhanced detection performance.

Table 3.1: Detection Rates for the Cuscore Charts and Combined Tracking Signal Methods on the Output Error by a GMV Controller

λ_0	(a) $\phi=-0.8$ $\theta=0.3$	(b) $\phi=-0.5$ $\theta=0.9$	(c) $\phi=0.2$ $\theta=0.6$	(d) $\phi=0.7$ $\theta=0.2$	(e) $\phi=-0.7$ $\theta=-0.2$	(f) $\phi=-0.2$ $\theta=-0.6$	(g) $\phi=0.5$ $\theta=-0.9$	(h) $\phi=0.8$ $\theta=0.3$
0.0	0.818	1.000	0.547	0.551	0.559	0.542	1.000	0.826
	(0.999)	(1.000)	(0.987)	(0.956)	(0.980)	(0.951)	(0.999)	(0.977)
0.2	0.825	1.000	0.546	0.549	0.552	0.537	1.000	0.824
	(0.998)	(1.000)	(0.990)	(0.956)	(0.980)	(0.948)	(1.000)	(0.977)
0.5	0.821	1.000	0.541	0.547	0.545	0.538	1.000	0.820
	(0.999)	(1.000)	(0.987)	(0.953)	(0.980)	(0.950)	(1.000)	(0.975)
0.8	0.816	1.000	0.550	0.554	0.560	0.546	1.000	0.821
	(0.998)	(1.000)	(0.989)	(0.957)	(0.982)	(0.952)	(1.000)	(0.979)
1.0	0.825	1.000	0.542	0.556	0.552	0.550	1.000	0.818
	(0.999)	(1.000)	(0.989)	(0.953)	(0.978)	(0.951)	(1.000)	(0.978)
1.5	0.823	1.000	0.540	0.564	0.549	0.546	1.000	0.821
	(0.999)	(1.000)	(0.984)	(0.960)	(0.981)	(0.951)	(1.000)	(0.977)
2.0	0.813	1.000	0.544	0.555	0.560	0.544	1.000	0.834
	(0.998)	(1.000)	(0.987)	(0.949)	(0.981)	(0.953)	(1.000)	(0.979)

3.4.2   Performance for Detecting a Step Shift

Again consider a system of a first-order dynamic process with $\delta = 0.5$ where the noise disturbance is modeled by an ARMA(1,1) process with a range of parameters for θ and ϕ. Here, a step signal of size 2 occurs in the noise at $t = 0$. The process is adjusted by using a GMV control scheme with varying λ_0 values. The system gain is 1. The control limits are chosen at $\pm 4\sigma_a$.

A MATLAB program simulated 10,000 runs with 100 observations for each run and computed the detection rates of the Cuscore chart for this controlled system. The simulation results listed in Table 3.2 show that the out-of-control ARL values vary over a wide range depending upon the ARMA parameters. For the values of ϕ and θ in cases (a), (b), (c), (e), and (f), the ARL is short which indicates the Cuscore chart has good detection capacity. For cases (d) and (g), the ARL is fair. For case (h), the ARL is large, failing to detect the step signal until 77 time periods on average. The results also show that the out-of-control ARLs are independent of the GMV control parameter λ_0, consistent with the theoretical results in Equation (3.22).

Table 3.2: Out-of-Control ARL for the Cuscore Chart on the Output Error from a GMV Controller

λ_0	(a) $\phi=-0.8$ $\theta=0.3$	(b) $\phi=-0.5$ $\theta=0.9$	(c) $\phi=0.2$ $\theta=0.6$	(d) $\phi=0.7$ $\theta=0.2$	(e) $\phi=-0.7$ $\theta=-0.2$	(f) $\phi=-0.2$ $\theta=-0.6$	(g) $\phi=0.5$ $\theta=-0.9$	(h) $\phi=0.8$ $\theta=0.3$
0	1.328	1.044	1.525	13.279	1.852	4.041	14.497	76.611
0.2	1.329	1.042	1.524	13.216	1.864	4.033	18.147	76.797
0.5	1.325	1.045	1.519	13.156	1.856	4.039	21.078	76.402
0.8	1.335	1.043	1.529	13.19	1.865	4.035	22.217	76.852
1	1.326	1.042	1.524	13.223	1.859	4.043	22.91	76.833
1.5	1.334	1.043	1.527	13.17	1.861	4.04	23.049	76.241
2	1.331	1.04	1.515	13.215	1.858	4.06	22.359	76.556

Table 3.3: Detection Rates for the Cuscore Chart on the Output Error from a GMV Controller (with Control Limits of the Cuscore Charts Given for Each Case)

λ_0 \ CL	(a) $\phi=-0.8$ $\theta=0.3$	(b) $\phi=-0.5$ $\theta=0.9$	(c) $\phi=0.2$ $\theta=0.6$	(d) $\phi=0.7$ $\theta=0.2$	(e) $\phi=-0.7$ $\theta=-0.2$	(f) $\phi=-0.2$ $\theta=-0.6$	(g) $\phi=0.5$ $\theta=-0.9$	(h) $\phi=0.8$ $\theta=0.3$
	9.97	13.35	7.02	3.54	6.70	4.27	4.11	3.10
0	1.000	1.000	1.000	0.765	1.000	0.924	1.000	0.729
0.2	1.000	1.000	1.000	0.770	1.000	0.916	1.000	0.739
0.5	1.000	1.000	1.000	0.769	1.000	0.922	1.000	0.735
0.8	1.000	1.000	1.000	0.764	1.000	0.921	1.000	0.726
1	1.000	1.000	1.000	0.763	1.000	0.917	1.000	0.730
1.5	1.000	1.000	1.000	0.769	1.000	0.919	1.000	0.728
2	1.000	1.000	1.000	0.769	1.000	0.928	1.000	0.727

3.4.3 Performance for Detecting a Bump

Again consider a system of a first-order dynamic process with $\delta = 0.5$ where the noise disturbance is modeled by an ARMA(1,1) process with a range of parameters for θ and ϕ. Here, a three-time-period bump signal of size 3 units occurs in the noise at $t = 5, 6$ and 7. The process is adjusted by using a GMV control scheme with λ_0 varying from 0 to 2. The system gain is again 1.

We first simulated 2,000 runs with 2,000 observations for each run without signals to find a the control limits which give an in-control ARL0 = 370 for each Cuscore chart designed for each ARMA(1,1) model. The control limits are listed in the CL row of Table 3.3.

A MATLAB Program simulated 10,000 runs with 30 observations for each run and computed the detection rates of the Cuscore chart for this controlled system. The simulation results are listed in Table 3.3. The results show that the detection rates are high for all cases except case (d) and (h). The results also illustrate that the detection rates are independent of the GMV control parameter λ_0, which is consistent with the theoretical results in Equation (3.24). For example, in Table 3.3, when $\theta = 0.3$ and $\phi = $ -0.8, the standard deviation of the ARLs corresponding to the varying λ_0 values is 0.004, which can be neglected in comparison with the mean of the ARLs 1.33.

3.5   Leakage Detection across a Valve

Consider a feedback control process applied to a fluid pipeline system in which the interest is in detecting the flow leakage across a check valve as illustrated in Figure 3.11. The quality characteristic in this system is the flow pressure which is adjusted by controlling the gap of the valve. It can be assumed that the sequential differentiated pressure fluctuates around a non-zero mean and follows Gaussian distribution due to a variety of factors, such as the inherent variance of sensors and fluid compressibility (Tu et al., 2003). The pressure across the valve is periodically measured by two sensors or meters at each side of the valve. In practice, correlation usually exists between two successive pressure measurements at small sampling intervals (Wang et al., 1993). Therefore, we applied a first-order autoregressive (AR) model to the pipeline pressure data series.

Furthermore, we assume that the effect of the valve adjustment action on the flow pressure can be modeled by a first-order dynamic process with a delay of one time period between the control actions and the output pressure. In order to achieve a small variance on the outlet pressure while keeping the variance of the valve control actions low as well, the GMV feedback controller is applied to the pipeline system. Suppose at time t, a leakage occurs in the gate valve and exists until it has been fixed. Therefore, a steady pressure drop signal is assumed to affect the inlet pressures series and it is important to detect the signal in a small run length by monitoring the controlled outlet pressure series.
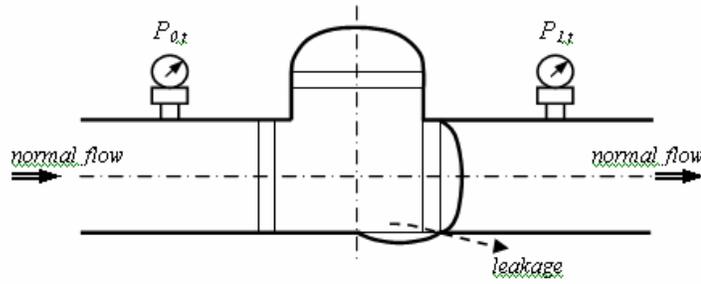
Figure 3.11: Schematic of valve flow and leakage.

In this example, the pressure set point $T = 100$ psi. $P_{0, t}$ and $P_{1, t}$ denote the inlet and outlet pressure deviations from the set point at time t respectively. We simulate an AR(1) model for the inlet pressure series:

$$P_{0, t} = 0.7\, P_{0, t-1} + a_t$$

where the initial pressure deviation $P_{0, 0} = 0$ psi, $\{a_t\}_{t \geq 0}$ is a white noise sequence, $N(0$ psi, 1 psi^2), and the first-lag correlation coefficient is $\phi = 0.7$. Suppose a leakage causes a 1% pressure drop occurs at the 50th observation in the inlet pressure and lasts for a sufficiently long period to enable detection. This downward step shift is modeled as

$$p_t = \begin{cases} 0 \text{ psi} & \text{at t} < 50 \\ -1 psi & \text{at t} \geq 50 \end{cases}.$$

We assume a first-order dynamic process for the inlet and the adjusted outlet pressure deviations, and a dynamic parameter $\delta = 0.3$ and a system gain $g = 0.95$:

$$P_{1, t} = 0.3\, P_{1, t-1} + 0.95\, X_{t-1}$$

where X_t is the control action at time t. Thus the adjusted outlet pressure deviation at time t can be written as

$$P_{1,t} = \frac{0.95}{1-0.3B} X_{t-1} + \frac{1}{1-0.7B} a_t + p_t .$$

Using a variance transfer parameter $\lambda_0 = 1.5$, the GMV controller at time t, X_t, can be obtained using Equation (3.14):

$$X_t = \frac{-0.7(1-0.3B)}{0.95(1-0.7B) + \dfrac{1.5}{0.95}(1-0.3B)} P_{1,t} = \frac{-0.7+0.21B}{2.53-1.14B} P_{1,t} .$$

Rearranging terms gives

$$X_t = 0.45X_{t-1} - 0.28\, P_{1,\, t} + 0.083\, P_{1,\, t-1}.$$

Suppose that pressure measurements taken by the sensors at both sides of the valve start at $t_0 = 0$ and observations are taken every 5 minutes. The interest of this research is to consider the actual pressure values which combine the deviations with the set point. Figure 3.12(a) shows the plot of the AR(1) inlet pressure series plus pressure drop signal. Figure 3.12(b) illustrates the outlet controlled pressure series and Figure 3.12(c) shows the control actions series. Shewhart control charts with $\pm 3\sigma$ control limits are applied to the time series in Figure 3.12(a-c) and no signal is fired in the three charts.

Alternatively, we can use the two-sided Cuscore control chart to monitor the outlet pressure deviations. The null model for the Cuscore statistics can be obtained by using Equation (3B.4) with an initial value of $a_{0,\,0} = 0$:

$$a_{t,0} = \frac{0.95^2(1-0.7B)+1.5(1-0.7B)(1-0.3B)}{0.95^2(1-0.7B)+1.5(1-0.3B)}P_{1,t}$$

$$= \frac{2.40-2.13B+0.315B^2}{2.40-1.08B}P_{1,t}$$

$$= P_{1,t} - 0.89P_{1,t-1} + 0.13P_{1,t-2} + 0.45a_{t-1,0}.$$

The detector can be obtained by using Equation (3B.5):

$$r_t = 1+(-0.7) = 0.3.$$

According to Equation (3.12), the two-sided Cuscore control chart is:

$$Q_t^+ = \max(0, Q_{t-1}^+ + 0.3(P_{1,t} - 0.89P_{1,t-1} + 0.13P_{1,t-2} + 0.45a_{t-1,0})),$$
$$Q_t^- = \min(0, Q_{t-1}^- + 0.3(P_{1,t} - 0.89P_{1,t-1} + 0.13P_{1,t-2} + 0.45a_{t-1,0})),$$
$$Q_0^+ = Q_0^- = 0.$$

Simulation is used to find the appropriate control limits for the Cuscore charts on the actual pressure values (10,000 runs and 1,000 observations for each run) and also find that an UCL of 108.0 psi and LCL of 92.0 psi gives an in-control ARL_0 of 370. Using these control limits, we apply the two-sided Cuscore chart to the outlet pressure series to detect the step signal caused by the leakage. It can be seen in Figure 3.12(d) that the leakage signal is detected in about 23 observations, at time $t = 73$.
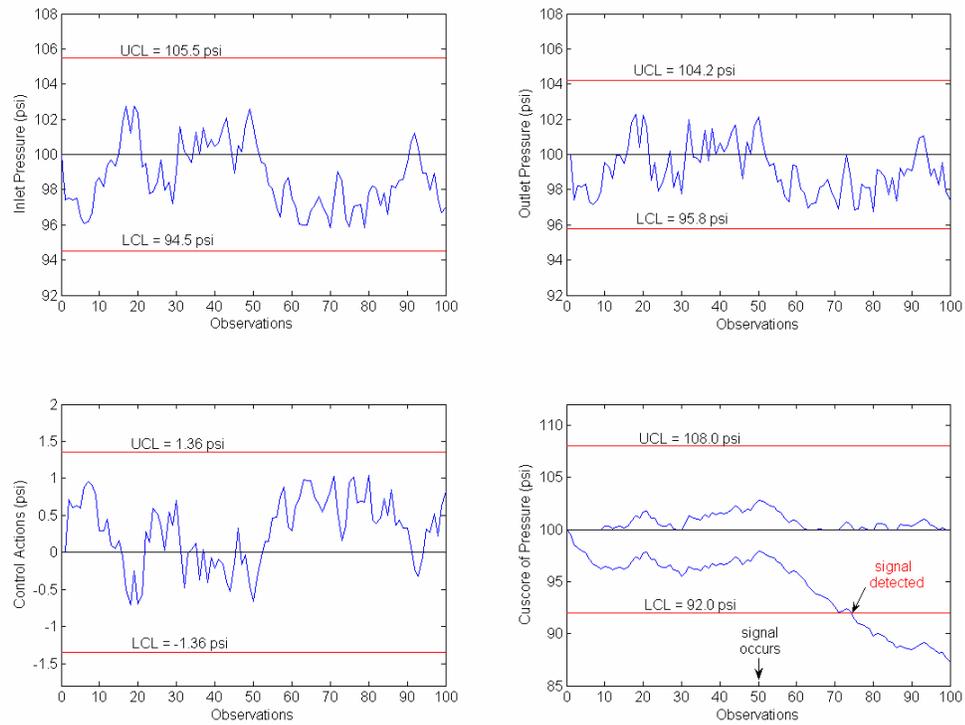
Figure 3.12: (a) AR(1) inlet pressure series; (b) Outlet pressure series; (c) Control action series; (d) Cuscore chart of the outlet pressure time series.

3.6   Conclusions

Cuscore charts can be used effectively to monitor the output of a GMV feedback control system for the presence of a signal. The approach developed in this investigation uses the fault signatures of the signal to derive appropriate statistics for their detection in an ARMA noise process. We show theoretically that the performance of Cuscore charts is independent of the amount of variability transferred from the output quality characteristic to the adjustment actions in the GMV control system. This property is potentially very useful when detecting signals in a GMV system with high variability transfer parameter values.

Simulation was used to explore the performance of using the Cuscore charts to monitor a ARMA(1,1) noise in detecting the spike, the step and the bump signal in a GMV control system. In general, the Cuscore has the ability to detect signals over a broad range of system parameter values. However, this investigation did identify some areas of low detection capability for certain fault signatures. In these cases, a tracking signal test is used in combination with the Cuscore statistics to give a better detection performance.

We believe that this approach has considerable theoretical and practical appeal. We provided several illustrations of the underlying behavior and showed how the methodology developed in this chapter is easily applied in a practical case of valve leakage detection.

Appendix 3A: Derivation of the Cuscore Statistics for a First-Order Dynamic Process and ARMA(1,1) Disturbance without Delay in GMV Control System and a Spike Signal

The transfer function model is given by

$$Y_t = \frac{g}{1 - \delta B} X_{t-1}. \tag{3A.1}$$

The time series model for the disturbance is given by

$$z_t = \frac{1 - \theta B}{1 - \phi B} a_t. \tag{3A.2}$$

The spike signal is given by

$$s_t = \gamma f(t). \tag{3A.3}$$

where γ is the size of the signal and $f(t)$ indicates the nature of the signal

52

$$f(t) = \begin{cases} 1 & t = t_0 \\ 0 & t \neq t_0 \end{cases}. \tag{3A.4}$$

The output error at time t can be written as

$$\varepsilon_t = \frac{g}{1-\delta B} X_{t-1} + \frac{1-\theta B}{1-\phi B} a_t + \gamma f(t). \tag{3A.5}$$

The GMV controller is given by

$$X_t = \frac{(\theta - \phi)(1-\delta B)}{g(1-\phi B) + \frac{\lambda_0}{g}(1-\theta B)(1-\delta B)} \varepsilon_t = K\varepsilon_t. \tag{3A.6}$$

Substituting Equation (3A.6) into Equation (3A.5) and writing ε_t in terms of a_t gives

$$\varepsilon_t = \frac{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)} \left(a_t + \frac{1-\phi B}{1-\theta B} \gamma f(t) \right). \tag{3A.7}$$

Rearranging items in Equation (3A.7) gives the discrepancy model

$$a_t = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t - \frac{1-\phi B}{1-\theta B} \gamma f(t). \tag{3A.8}$$

Obtaining the null model by letting $\gamma = \gamma_0 = 0$ in Equation (3A.8) gives

$$a_{t0} = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t. \tag{3A.9}$$

From Equation (3A.8) we obtain the detector for a spike signal

$$r_t = -\frac{a_t}{\gamma}\bigg|_{\gamma=\gamma_0} = \frac{1-\phi B}{1-\theta B} f(t) = f(t_0) = 1. \tag{3A.10}$$

For $t \neq t_0$, $r_t = 0$. Finally, we have the Cuscore statistics

$$Q_t = \sum_A a_{t0} r_t = a_{t0} r_{t0} = a_{t0} = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t. \tag{3A.11}$$

Substituting Equation (3A.7) into Equation (3A.11) gives

$$Q_t = a_t + \frac{1-\phi B}{1-\theta B}\gamma f(t).$$ (3A.12)

Equation (3A.12) shows that the Cuscore is equivalent to the Shewhart for in monitoring white noise when there is no spike signal. However the Cuscore statistics have an extra term produced by the filtered signal through an inversed ARMA(1,1) filter when the discrepancy model is valid. Note that the GMV control parameter λ_0 is not included in the Cuscore statistics.

The Cuscore statistics in Equation (3A.12) can also be used for the responsive system and/or IMA(1,1) noise model simply by letting the parameter $\delta = 0$ and/or $\phi = 1$.

Appendix 3B: Derivation of the Cuscore Statistics for a First-order Process and ARMA(1, 1) Disturbance without Delay in GMV Control System and a Step Signal

For this process, the output error at time t can be written as

$$\varepsilon_t = \frac{g}{1-\delta B}X_{t-1} + \frac{1-\theta B}{1-\phi B}a_t + \gamma f(t),$$ (3B.1)

where γ is the size of the step signal and $f(t)$ indicates the nature of the signal, which for a step is:

$$f(t) = \begin{cases} 1 & t < t_0 \\ 0 & t \geq t_0 \end{cases}.$$ (3B.2)

The discrepancy model and null model are the same as those for the spike signal case

$$a_t = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}\varepsilon_t - \frac{1-\phi B}{1-\theta B}\gamma f(t),$$ (3B.3)

and

$$a_{t0} = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}\varepsilon_t.$$　(3B.4)

The detector can be obtained from Equation (3B.3) by giving the step signal from time $t = t_0$

$$
\begin{aligned}
r_t &= \frac{1-\phi B}{1-\theta B}f(t)\\
&= f(t) - \phi f(t-1) + \theta\, r_{t-1}\\
&= f(t) + (\theta-\phi)f(t-1) + \theta(\theta-\phi)f(t-2) + \theta^2(\theta-\phi)f(t-3) + \mathrm{L} + \theta^{t-2}(\theta-\phi)f(1)\\
&= 1 + (\theta-\phi) + \theta(\theta-\phi) + \theta^2(\theta-\phi) + \mathrm{L} + \theta^{t-2}(\theta-\phi)\\
&= 1 + (\theta-\phi)\frac{1-\theta^{t-1}}{1-\theta}\\
&= 1 + \frac{\theta-\phi}{1-\theta}\qquad (t \gg 1).
\end{aligned}
$$　(3B.5)

We have the Cuscore statistics

$$Q_t = \sum_A a_{t0} r_t = \sum_A \left[\frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}\varepsilon_t\right]\left(1+\frac{\theta-\phi}{1-\theta}\right).$$　(3B.6)

Substituting ε_t in Equation (A7) into Equation (3B.6) gives

$$
\begin{aligned}
Q_t &= \sum_A \left[a_t + \frac{1-\phi B}{1-\theta B}\gamma f(t)\right]\left(1+\frac{\theta-\phi}{1-\theta}\right)\\
&= \sum_A \left[a_t + \frac{1-\phi B}{1-\theta B}\gamma\right]\left(1+\frac{\theta-\phi}{1-\theta}\right).
\end{aligned}
$$　(3B.7)

Equation (3B.7) shows that the Cuscore statistic is affected by the ARMA(1,1) parameters ϕ and θ and the step size γ. Note that the GMV control parameter λ_0 is not a part of the Cuscore statistic.

Appendix 3C: Derivation of the Cuscore Statistics for a First-order process and ARMA(1, 1) Disturbance without Delay in GMV Control System and a Three-Period Bump Signal

For this process, the output error at time t can be written as

$$\varepsilon_t = \frac{g}{1-\delta B} X_{t-1} + \frac{1-\theta B}{1-\phi B} a_t + \gamma f(t), \qquad (3C.1)$$

where γ is the size of the bump signal and $f(t)$ indicates the nature of the signal, which

for a three-period bump signal is

$$f(t) = \begin{cases} 1 & t = t_0, t-1 \text{ and } t_0 - 2 \\ 0 & \text{otherwise} \end{cases}. \qquad (3C.2)$$

The discrepancy model and null model are the same as those for the spike signal case

$$a_t = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t - \frac{1-\phi B}{1-\theta B} \gamma f(t), \qquad (3C.3)$$

The discrepancy model and null model are the same as those for the spike signal case

$$a_{t0} = \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)} \varepsilon_t. \qquad (3C.4)$$

The detector can be obtained from Equation (3C.3) by giving the bump signal

from time $t = 1$,

$$\begin{aligned} r_t &= \frac{1-\phi B}{1-\theta B} f(t) \\ &= f(t) - \phi f(t-1) + \theta r_{t-1} \\ &= f(t) + (\theta - \phi) f(t-1) + \theta(\theta - \phi) f(t-2). \end{aligned}$$

Thus,

$$\begin{aligned} r_{t_0} &= 1 + (\theta - \phi)(1 + \theta), \\ r_{t_0-1} &= 1 + \theta - \phi, \\ r_{t_0-2} &= 1. \end{aligned} \qquad (3C.5)$$

and $r_t = 0$ for all other time t. We have the Cuscore statistics

$$
\begin{aligned}
Q_t &= a_{t0}r_{t0} + a_{t_0-1}r_{t_0-1} + a_{t_0-2}r_{t_0-2} \\
&= \frac{g^2(1-\phi B) + \lambda_0(1-\phi B)(1-\delta B)}{g^2(1-\phi B) + \lambda_0(1-\theta B)(1-\delta B)}(\varepsilon_t r_{t0} + \varepsilon_{t-1}r_{t_0-1} + \varepsilon_{t-2}r_{t_0-2}).
\end{aligned}
\tag{3C.6}
$$

Expressing t, t-1 and t-2 respectively using Equation (3A.7) and substituting them

into Equation (3C.6) gives

$$
\begin{aligned}
Q_t &= a_{t_0}r_{t_0} + a_{t_0-1}r_{t_0-1} + a_{t_0-2}r_{t_0-2} \\
&= \left(a_t + \frac{1-\phi B}{1-\theta B}\gamma f(t)\right)r_{t_0} + \left(a_{t-1} + \frac{1-\phi B}{1-\theta B}\gamma f(t-1)\right)r_{t_0-1} + \left(a_{t-2} + \frac{1-\phi B}{1-\theta B}\gamma f(t-2)\right)r_{t_0-2}.
\end{aligned}
\tag{3C.7}
$$

Equation (3C.7) shows that the Cuscore statistics are affected by the ARMA(1,1)

parameters ϕ and θ and the signal size γ, but independent of GMV control parameter

λ_0 and dynamic system parameter δ.

Appendix 3D: The MATLAB Codes for Simulating the Detection Rates of the Cuscore Chart in Detecting a Spike Signal in the GMV Controller

```matlab
function cuscore_gmvc_spike_rate

n = 20;
y = zeros(1, n); z = zeros(1, n); a = zeros(1, n); e = zeros(1, n); x = zeros(1, n);

g = 1;  % the system gain
r = 10000;  h = 3;
theta = 0.3; %[0.3 0.9 0.6 0.2 -0.2 -0.6 -0.9 0.3];
phi = -0.8; %[-0.8 -0.5 0.2 0.7 -0.7 -0.2 0.5 0.8];
delta = 0.5; spike_size = 3; spike_here = 5;

lambda = [0 0.2 0.5 0.8 1.0 1.5 2];
m = length(lambda);

for i = 1:1:m
   lambda0 = lambda(i);
   count = 0;
```

```matlab
% calculate the Cuscore coefficients
ec0 = g^2 + lambda0;
ec1 = -g^2 * phi - lambda0 * phi - lambda0 * delta;
ec2 = lambda0 * phi * delta;
qc0 = ec0;
qc1 = -g^2 * phi - lambda0 * theta - lambda0 * delta;
qc2 = lambda0 * theta * delta;

for rep = 1:1:r

    % generate ARMA(1,1) noise
    a = randn(1, n);
    z(1) = a(1);
    for t = 2:1:n
        z(t) = phi * z(t-1) + a(t) - theta * a(t-1);
    end

    z(spike_here) = z(spike_here) + spike_size;

    % put the noise+signal over the first-order dynamic process
    y(1) = 0;
    e(1) = 0;
    x(1) = -(g * phi - g * theta) * e(1) / (g^2 + lambda0);
    y(2) = delta * y(1) + g * x(1);
    e(2) = z(1) + y(2);
    x(2) = (-g * (phi - theta) * e(2) + g * (phi - theta) * delta * e(1) + (-qc1) * x(1)) / ...
      (qc0);

    for t = 3:1:n
        y(t) = delta * y(t-1) + g * x(t-1);
        e(t) = y(t) + z(t-1);
        x(t) = (-g * (phi - theta) * e(t) + g * (phi - theta) * delta * e(t-1) + (-qc1) * x(t-1) + ...
          (-qc2) * x(t-2)) / (qc0);
    end

    % calculate Cuscore statistic
    CS(1) = e(1);
    CS(2) = ((ec0 * e(2) + ec1 * e(1)) - qc1 * CS(1)) / qc0;
    for t = 3:1:n
        CS(t) = ((ec0 * e(t) + ec1 * e(t-1) + ec2 * e(t-2)) - (qc1 * CS(t-1) + qc2 * ...
      CS(t-2))) / qc0;

        if abs(CS(t)) > h
            count = count + 1;
            break;
        end
    end
end
rate(i) = double(count/r);  % detection rate
end

% write the simulation results to a data file
```

```matlab
fid = fopen('dataout1.txt','w');
fprintf(fid, 'GMVC_lambda\trate\n');
for i = 1:1:m
   fprintf(fid,'%f\t%f\n',lambda(i), rate(i));
end
fclose(fid);
```

Appendix 3E: The MATLAB Codes for Simulating the ARL of the Cuscore Chart in Detecting a Mean Shift Signal in the GMV Controller

```matlab
function cuscore_gmvc_step_arl

n = 100; r = 10000;
y = zeros(1, n); z = zeros(1, n); a = zeros(1, n); e = zeros(1, n); x = zeros(1, n);

h = 4;
theta = 0.3; %[0.3 0.9 0.6 0.2 -0.2 -0.6 -0.9 0.3];
phi = -0.8; %[-0.8 -0.5 0.2 0.7 -0.7 -0.2 0.5 0.8];
g = 1; delta = 0.5;
step_size = 2; step_start = 1;

lambda = [0 0.2 0.5 0.8 1.0 1.5 2 3];
m = length(lambda);

%ve = zeros(1,m);
arl = zeros(1,m);

% give values to M
for t = 1:1:n
   M(t) = 1 + (theta - phi) / (1 - theta);
end

% do iteration to compute arl
for i = 1:1:m
   lambda0 = lambda(i);
   count = 0;
   for rep = 1:1:r

      % generate ARMA(1,1) noise + signal
      a = randn(1, n);
      z(1) = a(1);
      for t = 2:1:n
         z(t) = phi * z(t-1) + a(t) - theta * a(t-1);
      end

      for j = step_start:1:n
         z(j) = z(j) + step_size;
      end
```

```matlab
    % put the noise+signal over the first-order dynamic process
    y(1) = 0;
    e(1) = 0;
    x(1) = -(g * phi - g * theta) * e(1) / (g^2 + lambda0);
    y(2) = delta * y(1) + g * x(1);
    e(2) = z(1) + y(2);
    x(2) = (-g * (phi - theta) * e(2) + g * (phi - theta) * delta * e(1) + (g^2 * phi +
      lambda0 * theta + lambda0 * delta) * x(1)) / (g^2 + lambda0);


    for t = 3:1:n
        y(t) = delta * y(t-1) + g * x(t-1);
        e(t) = y(t) + z(t-1);
        x(t) = (-g * (phi - theta) * e(t) + g * (phi - theta) * delta * e(t-1) + (g^2 * phi +
    lambda0 * theta + lambda0 * delta) * x(t-1) + (-lambda0 * theta * delta) *
    x(t-2)) / (g^2 + lambda0);
    end


    % calculate Cuscore statistic
    ec0 = g^2 + lambda0;
    ec1 = -g^2 * (phi + theta) - lambda0 * (phi + theta + delta);
    ec2 = lambda0 * (phi * theta + phi * delta + theta * delta) + g^2 * phi * theta;
    ec3 = -lambda0 * phi * theta * delta;
    qc0 = ec0;
    qc1 = -g^2 * (theta + phi) - lambda0 * (delta + 2 * theta);
    qc2 = g^2 * theta * phi + lambda0 * (2 * theta * delta + theta^2);
    qc3 = -delta * theta^2 * lambda0;


    CS(1) = e(1) * M(1);
    CS(2) = ((ec0 * e(2) * M(2) + ec1 * e(1) * M(1)) - (qc1 - qc0) * CS(1)) / qc0;
    CS(3) = ((ec0 * e(3) * M(3) + ec1 * e(2) * M(2) + ec2 * e(1) * M(1)) - ((qc1 - qc0)
    * CS(2) + (qc2 - qc1) * CS(1))) / qc0;
    CS(4) = ((ec0*e(4)*M(4) + ec1*e(3)*M(3) + ec2*e(2)*M(2) + ec3*e(1)*M(1)) -
((qc1 - qc0) * CS(3) + (qc2 - qc1) * CS(2) + (qc3 - qc2) * CS(1))) / qc0;


    for t = 1:1:n
        if t > 4
            M(t) = 1 + (theta - phi) * (1 - theta^(t-1)) / (1 - theta);
            CS(t) = ((ec0 * e(t) * M(t) + ec1 * e(t-1) * M(t-1) + ec2 * e(t-2) * M(t-2) +
    ec3 * e(t-3) * M(t-3)) - ((qc1 - qc0) * CS(t-1) + (qc2 - qc1) * CS(t-2) +
    (qc3 - qc2) * CS(t-3) - qc3 * CS(t-4))) / qc0;
        end

        % calculate ARL
        if CS(t) > h | t >= n
            arl(i) = arl(i) + (t - step_start);
            break;
        end
    end
end


arl(i) = double(arl(i)/r);  % ARL
end
```

```matlab
% write the simulation results to a data file
fid = fopen('dataout.txt','w');
for i = 1:1:m
    fprintf(fid,'%f\n',arl(i));
end
fclose(fid);
```

Multivariate Cuscore Control Charts for Monitoring the Mean Vector in Autocorrelated Processes

Many modern industries operate in a data-rich environment where observations of more than one process variable can be collected simultaneously and used to characterize the process for quality control purposes. Additionally, the effect of sequential autocorrelation on process monitoring and control is increasingly important as the intervals between neighboring observations become shorter. Consequently, such sequential autocorrelation often leads to a high false alarm rate for traditional multivariate control charts, such as the multivariate cumulative sum (MCusum) chart or the multivariate exponentially weighted moving average (MEWMA) chart. In this chapter, a multivariate Cuscore (MCuscore) procedure based on the likelihood ratio test and fault signature analysis is proposed for monitoring the mean vector of an autocorrelated multivariate process. Simulation is performed to illustrate that the MCuscore chart outperforms the traditional residual-based MCusum control chart for detecting a mean vector shift in an autocorrelated multivariate process. An example of monitoring two autocorrelated process variables of a reactive ion etching (RIE) process used in semiconductor manufacturing demonstrates the application of the MCuscore chart.

4.1 Introduction

With the rapid development of modern technology and increasing complexity in modern industry, many manufacturing and business processes are more adequately represented by more than one quality variable. In many cases, historical data can be easily collected on several variables simultaneously but the data are often characterized by large size, high correlation, missing measurements, and low content of information due to low signal-to-noise ratios (Kourti, 2002). Meanwhile, there may also be a level of working knowledge or understanding about the inherent nature of processes that have been in operation over time.

The control chart is one of the primary techniques of statistical process control (SPC). Depending on the number of variables that it is designed to monitor, a control chart can be categorized as a univariate control chart or multivariate control chart. In general, most well-designed control charts consist of two phases. In Phase I, historical data are used to analyze the properties of the process, estimate the relevant parameters of process models, and construct the control limits or criteria of the control charts. In Phase II, control charts are applied to monitor the process and perform the tasks of fault detection and fault diagnosis.

Traditional univariate statistical process analysis characteristically ignores the correlation among the process variables and constructs a separate chart for each of them. This approach often leads to inaccurate control limits and poor detection performance in monitoring mean shift signals for correlated variables. By assuming an independent and identically distributed (i.i.d.) multivariate Gaussian process with constant covariance

matrix among sequential observations, Shewhart, Cusum, and EWMA charts can be extended to the multivariate environment to monitor the process mean (e.g., Jackson, 1985; Healy, 1987; Lowry et al., 1992). However, the false alarm rates are usually high when the multivariate processes are actually sequentially autocorrelated.

To remedy this problem, some multivariate SPC approaches have been proposed for monitoring the mean vector of the autocorrelated multivariate process, such as Mastrangelo and Forrest (2002), Noorossana and Vaghefi (2005), and Bakski (1998). However, some important issues, such as the analysis of *fault signature*, the change pattern of the fault signal imposed on the output residual in removing the process autocorrelation by filtering, and its combination with time series model identification, have not been fully addressed.

The univariate Cuscore statistic has been used in control charts to monitor for anticipated process signals (Box and Ramirez, 1992; Box and Luceño, 1997; Nembhard and Changpetch, 2006; Nembhard, 2006; Nembhard and Chen, 2006; and Nembhard and Valverde-Ventura, 2003, 2006, Shu et al., 2002). This research extends the univariate Cuscore statistic to the multivariate environment and provides the theoretical derivation of bivariate Cuscore statistics. The multivariate Cuscore (MCuscore) control chart is applied to monitor the mean shift in two simulated autocorrelated process variables in a reactive ion etching (RIE) process in semiconductor manufacturing. For comprehensiveness, some background is provided on the multivariate fault signature before developing the MCuscore procedure.

In Section 4.2, a brief literature review of multivariate control charts is provided, along with an introduction to the multivariate time series model, in the form of a vector

auto-regressive and moving average (VARMA) process. In Section 4.3, the VARMA is used as a basis for discussing the fault signature of a mean shift signal in the output of the inverse VARMA filter, especially for the vector autoregressive (VAR) process. In Section 4.4, the MCuscore statistic is derived with the aid of fault signatures in a VARMA model. In Section 4.5, the performance of the MCuscore chart is investigated and compared with the residual-based MCusum control chart, and the robustness of the control chart is briefly discussed. In Section 4.6, the MCuscore chart approach is applied to a simulated application example for monitoring the mean shift of two autocorrelated RIE process variables. We present a discussion on the diagnosis of out-of-control signals for the MCuscore chart in Section 4.7 and conclude the chapter in Section 4.8.

4.2   A Brief Review of Multivariate Control Charts

In this section, the multivariate time series model is introduced as an important tool to characterize autocorrelated multivariate processes. Then three major types of multivariate control charts to monitor the process mean vector shift – Hotelling's T^2 control chart, multivariate Cusum (MCusum) control chart and multivariate EWMA (MEWMA) control chart – are briefly reviewed.

4.2.1   Multivariate Time Series Model

Time series modeling techniques are commonly used in univariate processes to remove any autocorrelation structure in the observations. In almost all cases it is assumed

that the Box-Jenkins model can be used to describe the behavior of observations. To extend the univariate framework of time series modeling to multivariate observations, suppose x_{1t}, x_{2t}, ..., x_{nt} are an output series of a stable process. Following Montgomery et al. (1990) and Hamilton (1994), a multivariate time series model can be written in the VARMA form:

$$\mathbf{\Phi}_p(B)\nabla^d \mathbf{X}_t = \mathbf{\Theta}_q(B)\mathbf{\varepsilon}_t, \qquad (4.1)$$

where $\mathbf{X}_t = (x_{1t}, x_{2t}, ..., x_{nt})'$ represents the vector of the time series of interest, $\mathbf{\varepsilon}_t$ is a sequence of independent multivariate normal random vectors with mean zero and variance-covariance matrix $\mathbf{\Sigma}$, $\mathbf{\Phi}_p(B)$ and $\mathbf{\Theta}_q(B)$ are polynomials in B with $\mathbf{\Phi}_p(B) = \mathbf{\Theta}_q(B) = \mathbf{I}$, ∇ is the difference operator, and d indicates the required differencing to make the time series stationary.

Note that VARMA(1,1) model is reduced to an i.i.d. multivariate Gaussian model if the $\mathbf{\Phi}$ matrix is reduced to a diagonal matrix and the $\mathbf{\Theta}$ matrix is a zero matrix.

Practically, it is much more difficult to fit a proper VARMA model than an ARMA model because actual multivariate processes are often very complicated and highly correlated. Tools such as SAS *Proc Statespace* can be used to perform the multivariate time series model identification. Theoretical particulars of such model identification are discussed in Box et al. (1994) and Del Castillo (2002).

4.2.2 Hotelling's T2 Control Chart

One of the earliest approaches to multivariate process control was Hotelling's T^2 statistics (Hotelling, 1947; Johnson and Wichern, 2002). Consider a sequence of multivariate observations represented by $\mathbf{X}_t$, $t = 1, 2, \mathrm{K}, n$, where each observation $\mathbf{X}_t$ is a $p \times 1$ random vector whose jth element represents the jth quality characteristic of the process. Assume $\mathbf{X}_t$, $t = 1, 2, \mathrm{K}, n$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Suppose the objective is to test whether $\mathbf{X}_t$, $t = 1, 2, \mathrm{K}, n$ come from a multivariate normal distribution with a "good" mean $\boldsymbol{\mu}_0$ or a "bad" mean $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_0 + \boldsymbol{\delta}$. The covariance matrix $\boldsymbol{\Sigma}$ is assumed constant for both cases. Of clear relevance is a test of the null hypothesis $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against the alternative hypothesis $H_a : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. Multivariate analysis shows that the most powerful test statistic for H_0 against H_a rejects the null hypothesis if the value of $T_t^2 = (\mathbf{X}_t - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma} (\mathbf{X}_t - \boldsymbol{\mu}_0)$ is sufficiently large. In distribution theory, T_t^2 follows a χ_p^2 distribution with p degrees of freedom, $T_t^2 \sim \chi_p^2$. Thus, given a Type-I error α for the null hypothesis test, a T^2 control chart with the control limit of $\chi_p^2(\alpha)$ can be established.

Hawkins and Olwell (1998) discussed an extension of this basic result. Instead of studying a single observation vector $\mathbf{X}_t$, they considered a rational group of size m and computed the sample mean vector $\bar{\mathbf{X}}_m$, so that $\bar{\mathbf{X}}_m \sim MVN\left(\boldsymbol{\mu}, \dfrac{\boldsymbol{\Sigma}}{m}\right)$, and the optimal test

statistic becomes $T_t^2 = m(\mathbf{X}_t - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}(\mathbf{X}_t - \boldsymbol{\mu}_0) \sim \chi_p^2$. In this form, the T^2 control chart can be constructed.

4.2.3 Multivariate Cusum Control Chart

Healy (1987) derived a multivariate Cusum (MCusum) procedure based on the sequential likelihood ratio test of multivariate variables. The MCusum statistic for detecting a specific shift in the process mean vector $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ is

$$S_n = \max(S_{n-1} + \mathbf{a}'\mathbf{X}_n - K, 0) > H, \tag{4.2}$$

where $S_0 = 0$, $\mathbf{X}_n$ is the sample mean vector at time n, H is a fixed threshold, $\mathbf{a}$ is an $m \times 1$ vector of constants defined as

$$\mathbf{a}' = \frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}}{\sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}}, \tag{4.3}$$

and

$$K = \frac{(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}{2\sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}}. \tag{4.4}$$

Specifically, $\mathbf{a}'\mathbf{X}_n$ has a univariate normal distribution with variance 1 and mean $\mathbf{a}'\boldsymbol{\mu}_0$ when the process operates in the good state and mean $\mathbf{a}'\boldsymbol{\mu}_1$ when the process has shifted to the bad state. By defining the non-centrality parameter

$D = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}$, the MCusum statistics can be written in the form of Equation (4.2) using $K = D/2$.

Noorossana and Vaghefi (2005) applied the MCusum control chart to monitor the residuals from a vector AR(1) time series model for a mean vector shift by assuming that the parameters of the model were estimated accurately by using the historical data. They showed by simulation that the average run length (ARL) properties of MCusum control charts can be improved considerably if the residuals from a time series model were used instead of the original data.

Crosier (1988) proposed a MCusum procedure that accumulates on the scale of $\mathbf{X}_n$, instead of $\mathbf{X}_n$ or T^2. The MCusum vector $\mathbf{S}_n$ is initialized to a zero vector, and then recursively updated

$$\mathbf{S}_n = \begin{cases} 0 & \text{for } C_n \leq k \\ (\mathbf{S}_{n-1} + \mathbf{X}_n - \boldsymbol{\mu}_0)/(1 - k/C_n) & \text{for } C_n \geq k \end{cases}, \tag{4.5}$$

where $C_n = (\mathbf{S}_{n-1} + \mathbf{X}_n - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\mathbf{S}_{n-1} + \mathbf{X}_n - \boldsymbol{\mu}_0)$. The MCusum control chart will signal if $\mathbf{S}_n' \boldsymbol{\Sigma}^{-1} \mathbf{S}_n > H$, where H is the control threshold. Hawkins and Olwell (1998) stated that this MCusum is intuitively attractive because it accumulates on the $\mathbf{X}$ scale rather than a quadratic scale. It has the same property as the scalar Cusum of resetting to zero when there seems little evidence that the process is off target, and its final decision uses the T^2 metric.

4.2.4 Multivariate EWMA Control Chart

Lowry et al. (1992) presented a multivariate EWMA procedure:

$$\mathbf{Z}_t = \mathbf{R}\mathbf{X}_t + (\mathbf{I} - \mathbf{R})\mathbf{X}_{t-1},$$
(4.6)

where $\mathbf{X}_t$ is a vector of the process variables at time t, $\mathbf{R}$ is a diagonal matrix, $\mathbf{I}$ is the identity matrix. This procedure signals a shift when $T_{Z_i}^2 = \mathbf{Z}_i' \mathbf{\Sigma}_Z^{-1} \mathbf{Z}_i > H$, where $\mathbf{\Sigma}_Z$ is the exact covariance of $\mathbf{Z}_t$. Kramer and Schmid (1997) applied the MEWMA to the residuals from a vector AR(1) time series model. Mastrangelo and Forrest (2002) developed an adaptive approach to monitor autocorrelated processes using a MEWMA residual.

4.3 Fault Signatures of Mean Vector Shifts

For a univariate autocorrelated process, Hu and Roan (1996) showed that removing the autocorrelation using a time series model is equivalent to inverse filtering the data using the time series model. In the case of a mean shift signal in the original autocorrelated data series, the residual from the inverse filter is composed of two parts, a white noise series and a change pattern of the mean shift that often behaves as a transient response. The magnitude and the pattern of signal in the residual depend on the structure of the ARMA model and the model parameters. In some process control literature, the change pattern of a fault signal is called *fault signature* because the controller or filter

gives a signature of the fault signal to the output series (e.g., Apley and Shi, 1999; Yoon and MacGregor, 2001; Nembhard and Changpetch, 2006).

Equivalently, for a multivariate correlated process, a properly fitted multivariate time series model can serve as an inverse filter to remove the autocorrelation in the data. The fault signal of a mean vector shift in the original data takes the form of a transient response and is imposed on the multivariate white noise residuals. For example of a first-order vector auto-regressive and moving average (VARMA(1,1)) process, the model can be written as

$$\mathbf{X}_t = \mathbf{\Phi}\mathbf{X}_{t-1} + \mathbf{\epsilon}_t - \mathbf{\Theta}\mathbf{\epsilon}_{t-1}, \text{ or } \mathbf{X}_t = \frac{\mathbf{I} - \mathbf{\Theta}B}{\mathbf{I} - \mathbf{\Phi}B}\mathbf{\epsilon}_t, \qquad (4.7)$$

where $\mathbf{\epsilon}_t$ follows a multinormal distribution, $\varepsilon_m \sim MVN(\mathbf{\mu}, \mathbf{\Sigma})$.

As is the case for univariate time series analysis, stationarity and invertibility are two important properties for a VARMA time series model. Stationarity requires that neither the process mean vector nor the covariance matrix depends on time t, and invertibility requires that a VMA(1) model can be inverted to a VAR(∞) model. Statistically, for an VARMA(1,1) model such as Equation (4.7), stationarity requires that all elements λ_i of the root vector $\mathbf{\lambda}$ for

$$\det(\mathbf{I} - \mathbf{\lambda}\mathbf{\Phi}) = 0 \qquad (4.8)$$

are inside the unit circle, or $|\lambda_i| < 1$, and invertibility requires all roots $|\lambda_i| < 1$ for

71

$$\det(\mathbf{I} - \lambda\boldsymbol{\Theta}) = 0 \tag{4.9}$$

are also inside the unit circle. Refer to Hamilton (1994) and Brockwell and Davis (1991) for a detailed discussion on stationarity and invertibility of vector time series models.

The VARMA(1,1) model with a mean shift can be represented by

$$\mathbf{X}_t^* = \frac{\mathbf{I} - \boldsymbol{\Theta}B}{\mathbf{I} - \boldsymbol{\Phi}B}\boldsymbol{\varepsilon}_t + \boldsymbol{\gamma}\boldsymbol{\Gamma}_t, \tag{4.10}$$

where $\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$ is the mean shift signal, $\boldsymbol{\gamma}$ indicates the vector of the shift size, and $\boldsymbol{\Gamma}_t$ is the diagonal pattern matrix indicating the type of mean shift signals, such as steady shifts, ramp shifts, and so on as follows

$$\boldsymbol{\Gamma}_t = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & t < t_0 \\[2mm] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & t \geq t_0 \text{ (steady shift)} \\[2mm] \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} & t \geq t_0 \text{ (ramp)} \\[2mm] \begin{bmatrix} e^{-it} & 0 \\ 0 & e^{-it} \end{bmatrix} & t \geq t_0 \text{ (sinusoidal)} \\[2mm] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} t = t_0 \text{ and } \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & t \geq t_0 \text{ (transient shift)} \end{cases} \tag{4.11}$$

Applying an inverse VARMA filter (VARMA^{-1}) to Equation (4.10) gives

$$\frac{(\mathbf{I} - \mathbf{\Phi}B)}{(\mathbf{I} - \mathbf{\Theta}B)}\mathbf{X}_t^* = \boldsymbol{\varepsilon}_t + \frac{(\mathbf{I} - \mathbf{\Phi}B)}{(\mathbf{I} - \mathbf{\Theta}B)}\gamma\mathbf{\Gamma}_t. \tag{4.12}$$

It can be seen in Equation (4.12) that the filtered data consists of the white noise $\boldsymbol{\varepsilon}_t$ and the fault signature

$$\boldsymbol{\delta} = \frac{(\mathbf{I} - \mathbf{\Phi}B)}{(\mathbf{I} - \mathbf{\Theta}B)}\gamma\mathbf{\Gamma}_t. \tag{4.13}$$

Figure 4.1 graphically illustrates the composition of the mean shift signal and the vector time series, and the filtering of an inverse VARMA model.

As in the univariate process analysis, the structure of the VARMA model and its parameters can affect the magnitude and the pattern of the fault signature. We use a bivariate VAR(1) model to illustrate the impact of the $\mathbf{\Phi}$ matrix on the fault signatures. In the simulation, we assume the white noise term $\boldsymbol{\varepsilon}_t$ in the VAR(1) model has a multivariate normal distribution with $\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$. The length of process is 20 and the mean shift vector $\mathbf{\Gamma} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ starts at the 10$^{\text{th}}$ observation.

We first study the effect of ϕ_1 on the fault signature by decreasing the value of ϕ_1 from 0.7 to -0.7 while keeping the other three parameters fixed (Table 4.1) and Figure 4.2). Then we study the effect of ϕ_{12} on the fault signature by increasing ϕ_{12} from -1.2 to 1.6 while keeping others fixed (Table 4.2 and Figure 4.3). The values are selected to meet the stationarity condition of a VAR(1) model (Equation (4.8)).
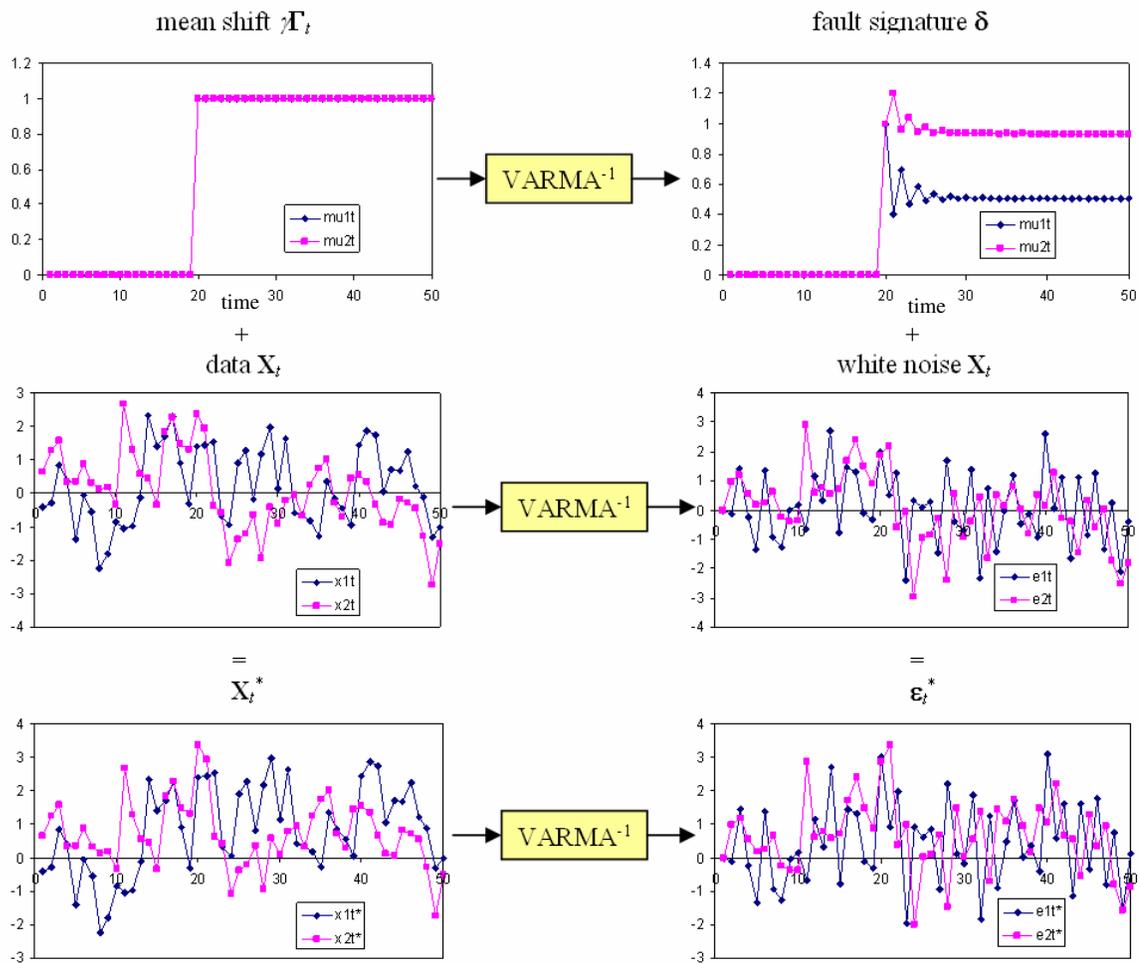
Figure 4.1: Composition and filtering using the VARMA(1,1) process.

Note that the size and pattern of the fault signature can be affected by many factors, such as the dimension of the process, the size and pattern of the mean shift signal, the structure of VARMA model, the parameter matrixes of Σ, Φ and Θ, thus are very complicated. An extensive study on the pattern of the fault signatures for multivariate time series models is beyond the scope of this chapter. In the next few steps, we will limit our discussion to the effects of the Φ and Θ matrixes on fault signatures of the mean

shift vector with unit elements in a bivariate process with fixed covariance matrix $\mathbf{\Sigma}$ for the multivariate normal distribution of the white noise.

Table 4.1: Selection of $\mathbf{\Phi}$ Matrix to Vary ϕ_1

	ϕ_1	ϕ_{12}	ϕ_{21}	ϕ_2
a	*0.7*	0.1	0.2	0.5
b	*0.4*	0.1	0.2	0.5
c	*0.1*	0.1	0.2	0.5
d	*-0.1*	0.1	0.2	0.5
e	*-0.4*	0.1	0.2	0.5
f	*-0.7*	0.1	0.2	0.5
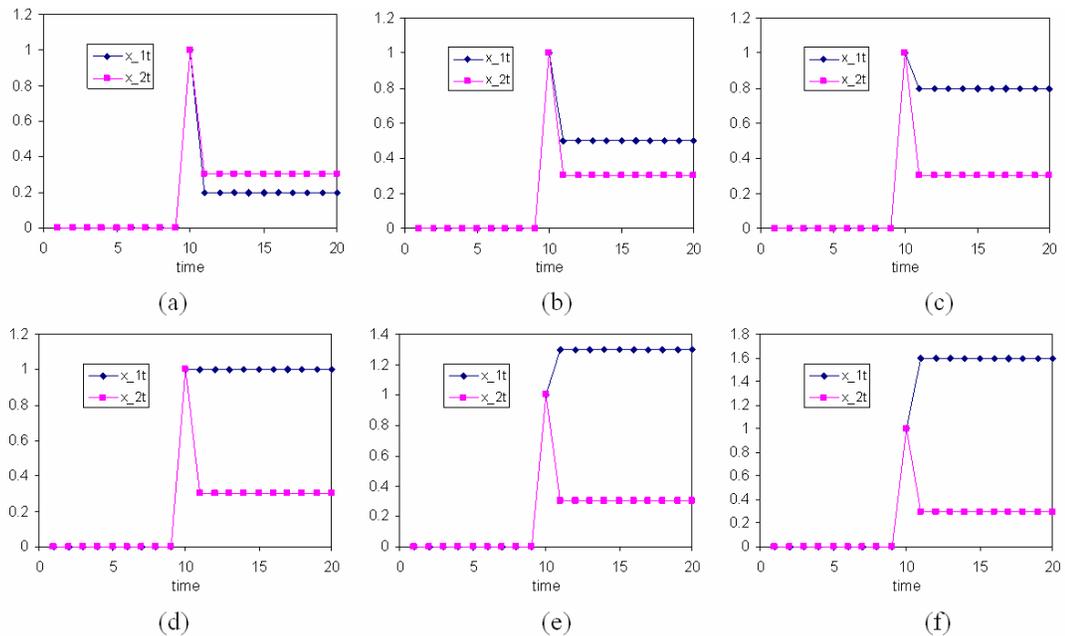
Figure 4.2:   Fault signatures corresponding to the respective values in Table 4.1.

Table 4.2: Selection of **Φ** Matrix to Vary ϕ_{12}

	ϕ_1	ϕ_{12}	ϕ_{21}	ϕ_2
a	-0.7	*-1.2*	0.2	0.5
b	-0.7	*-0.8*	0.2	0.5
c	-0.7	*-0.4*	0.2	0.5
d	-0.7	*0.0*	0.2	0.5
e	-0.7	*0.4*	0.2	0.5
f	-0.7	*0.6*	0.2	0.5
g	-0.7	*0.8*	0.2	0.5
h	-0.7	*1.2*	0.2	0.5
i	-0.7	*1.6*	0.2	0.5
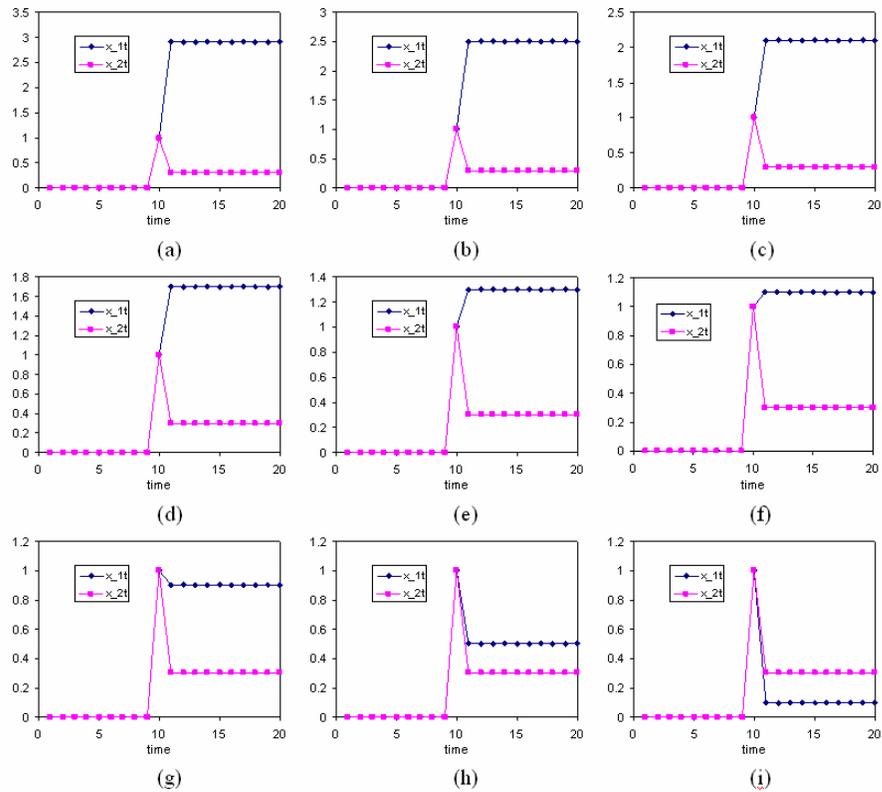
Figure 4.3: Fault signatures corresponding to the respective values in Table 4.2.

To further study the impact of the parameters of the VARMA model on the fault signature of the mean shift vector $(1, 1)'$, we illustrate the fault signatures using three $\boldsymbol{\Phi}$ matrixes by varying only the ϕ_1 value, and each $\boldsymbol{\Phi}$ matrix corresponds to six $\boldsymbol{\Theta}$ matrixes (see Table 4.3). The $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ matrixes are selected to meet the stationarity and invertibility condition of the VARMA(1,1) model (Equations (4.8) and (4.9)).

For comparison, we first plot the fault signatures of the mean shift vector $(1, 1)'$ for a VMA(1)$^{-1}$ filter in Figure 4.4. Each panel corresponds to the respective value in Table 4.3 with $\boldsymbol{\Theta}$ matrix from row 1-6 respectively. It can be seen that the first three fault signatures steadily increase to a constant level higher than the original level $(1, 1)'$ by different amounts, and the last three have some upside-down patterns around their final constant levels. Figures 4.5-4.7 illustrate the fault signature of mean shift vector $(1, 1)'$ in different VARMA(1,1)$^{-1}$ filters. It is common that the first set of values of the fault signatures for both variables remain the same as the original mean and then the rest of the fault signatures display different shift patterns.

Table 4.3: Selection of $\boldsymbol{\Phi}$ Matrix to Vary ϕ_1, and $\boldsymbol{\Theta}$ Matrix to Vary θ_{12}

	ϕ_1	ϕ_{12}	ϕ_{21}	ϕ_2		θ_1	θ_{12}	θ_{21}	θ_2
a	0.7	*0.1*	0.2	0.5	1	0.3	*0.4*	0.6	0.3
b	-0.7	*0.1*	0.2	0.5	2	0.3	*0.2*	0.6	0.3
c	-0.7	*0.6*	0.2	0.5	3	0.3	*0.0*	0.6	0.3
					4	0.3	*-0.2*	0.6	0.3
					5	0.3	*-0.4*	0.6	0.3
					6	0.3	*-0.6*	0.6	0.3
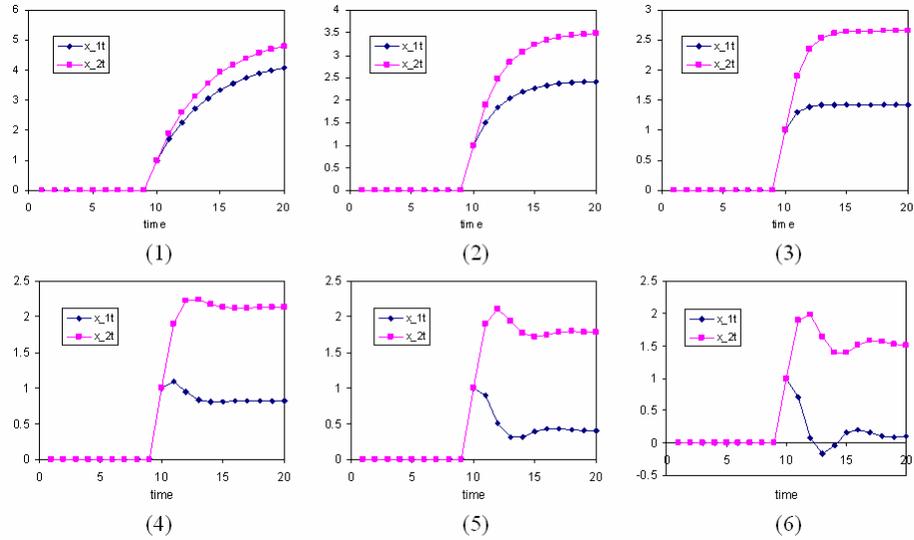
Figure 4.4: Fault signatures corresponding to the respective values in Table 4.3 with Φ matrix from 1-6 respectively.

Figure 4.5: Fault signatures corresponding to the respective values in Table 4.3 with Φ matrix from a and Θ matrix from 1-6 respectively.

Figure 4.6: Fault signatures corresponding to the respective values in Table 4.3 with Φ matrix from *b* and Θ matrix from 1-6 respectively.

Figure 4.7: Fault signatures corresponding to the respective values in Table 4.3 with Φ matrix from *c* and Θ matrix from 1-6 respectively.

4.4   The Multivariate Cuscore Procedure

Box and Ramírez (1992) proposed a Cuscore chart based on the likelihood ratio test to identify suspected deviations known to be characteristic of (or peculiar to) a monitored system. They assumed that a model can be expressed in the form

$$a_i = a(y_i, X_i, \gamma), \ i = 1, 2, ..., t,$$ (4.14)

where the y_i are observations (on the quality characteristic), the X_i are known (controllable input) quantities, γ is the size of the mean shift signal, either transient or steady, the a_i's are independent normal random variables with mean 0 and variance σ^2 (i.e., white noise). Further, they assumed σ is known and does not depend on γ, and the size and starting time of the signal are known. Then the log-likelihood function is

$$l = -\frac{1}{2\sigma^2} \sum_{i=1}^{t} a_i^2 + c,$$ (4.15)

where c is a constant independent of γ.

Following Fisher (1925), the efficient score statistic is obtained from Equation (4.15) by differentiation with respect to γ at $\gamma = \gamma_0$. Thus

$$Q = \sum_{i=1}^{t} a_{i0} r$$ (4.16)

is the Cuscore associated with the parameter value $\gamma = \gamma_0$.

Two-sided Cuscore statistics are easily constructed. Let the superscripts + and –
denote positive and negative biases respectively. The formulae can be written as

$$
\begin{aligned}
Q_t^+ &= \max(0, Q_{t-1}^+ + a_{t0} r_t), \\
Q_t^- &= \min(0, Q_{t-1}^- + a_{t0} r_t), \\
Q_0^+ &= Q_0^- = 0.
\end{aligned}
\tag{4.17}
$$

As shown in Box and Luceño (1997) and in Nembhard (2006), the Cuscore chart
designed for detecting a spike signal in white noise is equivalent to a traditional Shewhart
chart; the Cuscore chart designed for designed for detecting a step signal in white noise is
equivalent to a CUSUM chart; and the Cuscore chart designed for detecting a bump
signal in white noise is equivalent to an arithmetic moving average (AMA) chart.

The Cuscore chart can be used to monitor mean shift signals, either steady or
transient, in an autocorrelated process by transforming the detector according to the
autocorrelation structures (Box and Luceño, 1997; Nembhard, 2006; Nembhard and
Chen, 2006; Nembhard and Valverde-Ventura, 2006).

In a manner similar to the derivation process for the univariate Cuscore statistics,
we derive the MCuscore statistics with a specific bivariate VAR(1) as follows. Consider a
bivariate VAR(1) model

$$
\mathbf{X}_t = \mathbf{\Phi} \mathbf{X}_{t-1} + \mathbf{a}_t,
\tag{4.18}
$$

which can be rewritten as a null model

$$\mathbf{a}_t = (\mathbf{I} - \mathbf{\Phi}B)\mathbf{X}_t. \tag{4.19}$$

Assume $\mathbf{a}_t$ follows a bivariate normal distribution with $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and constant

$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_1^2 \end{pmatrix}$. Further, assume the mean shift size $\boldsymbol{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$ and pattern matrix

$\boldsymbol{\Gamma}_t = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & t < t_0 \\ \begin{bmatrix} f_{1t} & 0 \\ 0 & f_{2t} \end{bmatrix} & t \geq t_0 \end{cases}$ of the shift signal $\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$ are known. Thus the discrepancy model

of the MCuscore statistic is

$$\mathbf{a}_t = (\mathbf{I} - \mathbf{\Phi}B)\mathbf{X}_t - (\mathbf{I} - \mathbf{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma}. \tag{4.20}$$

In the derivation of the MCuscore statistics, we expand Equation (4.20) into the scalar form and take derivatives with respect to γ_1 and γ_2. After rearranging the terms into vector and matrix forms and using the expression of the fault signature $\boldsymbol{\delta}_t = (\mathbf{I} - \mathbf{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$, we can obtain the two-sided MCuscore statistics

$$S_t = \begin{cases} \max(S_{t-1} + \boldsymbol{\delta}_t'\boldsymbol{\Sigma}^{-1}\mathbf{x}_t - \frac{1}{2}\boldsymbol{\delta}_t'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_t,\ 0) > H \\ \min(S_{t-1} + \boldsymbol{\delta}_t'\boldsymbol{\Sigma}^{-1}\mathbf{x}_t - \frac{1}{2}\boldsymbol{\delta}_t'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}_t,\ 0) < -H \end{cases}. \tag{4.21}$$

The derivation of Equation (4.21) is based on the sequential probability ratio test (SPRT) and is given in Appendix 4A.

82

It can be seen that the form of Equation (4.21) is similar to that of the MCusum statistic in Equation (4.2), except that the mean shift term is the fault signature $\boldsymbol{\delta}$ in Equation (4.21) instead of the constant mean shift vector in Equation (4.2). However, for an *i.i.d.* bivariate Gaussian process which is equivalent to a VAR(1) process with $\boldsymbol{\Phi} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, the fault signature has no difference from the constant mean shift, and thus the MCuscore statistic is actually the MCusum statistic.

We note that in the residual-based MCusum procedure proposed by Noorossana and Vaghefi (2005), the VARMA filter and MCusum control chart are performed separately. The proposed MCuscore procedure is actually equivalent to the procedure of combining the inverse VARMA filter with residual-based MCusum charts.

The MCuscore statistics for detecting a mean shift signal in a VARMA(1,1) noise are similar to those in Equation (4.21). The difference, however, is that the fault signature term is no longer $\boldsymbol{\delta}_t = (\mathbf{I} - \boldsymbol{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$; it takes a more complicated form. For example, for the VARMA(1,1) model, $\boldsymbol{\delta} = \dfrac{(\mathbf{I} - \boldsymbol{\Phi}B)}{(\mathbf{I} - \boldsymbol{\Theta}B)}\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$. Furthermore, the derived MCuscore procedure is not confined to bivariate process. It can be used for monitoring processes with $p > 2$ variables, and correspondingly, the vectors $\mathbf{x}_t$, $\boldsymbol{\mu}_t$, $\boldsymbol{\gamma}$ and matrixes $\boldsymbol{\Gamma}_t$, $\boldsymbol{\Phi}$, $\boldsymbol{\Theta}$ are all p dimension.

4.5   Performance Evaluation and Robustness Analysis

In this section, we use Monte Carlo simulation to evaluate the performance of the MCuscore chart in detecting the mean shift vector in the VAR(1) and the VARMA(1,1) processes. By varying some of the elements in the Φ and Θ matrixes, we also analyze the robustness of the MCuscore control chart.

First, we investigate the out-of-control ARL of the MCuscore chart in monitoring the bivariate VAR(1) models. The white noise term in each inverse VAR(1) filter is assumed to follow a bivariate normal distribution with constant $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$. We compare the performance of the MCuscore and the residual-based MCusum control chart in monitoring the mean shift faults. The Φ matrixes are chosen from Table 4.1 and Table 4.2 and satisfy the stationary condition of a VAR(1) time series.

We first simulate 5,000 runs with 1,000 observations in each run to determine the control limits of the MCuscore chart and residual-based MCusum chart based on $ARL_0 = 200$, see Table 4.4. Note that, the control limits are constant for each residual-based MCusum chart at equal ARL_0 because the mean shift vector in calculating the sequential MCusum statistics is constant; however, for the MCuscore chart, the control limit varies for different Φ matrixes because the mean vectors and the reference values in the sequential MCuscore statistics are not constant. Therefore, we must re-identify the control limits for MCuscore when the Φ matrix has been changed.

Two mean shift vectors $(1.0, 1.0)'$ and $(1.0, 0.5)'$ are examined for each VAR(1) model respectively. It can be seen in Table 4.4 that the MCuscore chart outperforms the residual-based MCusum chart in terms of the out-of-control ARLs. The same conclusion can also be drawn from the simulation results in Table 4.5 which compares the two control charts in detecting the mean shift vector $(1.0, 1.0)'$ in an VARMA(1,1) process whose parameters are from Table 4.3.

For model-based control charts like the residual-based MCusum and MCuscore charts, the accuracy of the estimated parameters has an effect on performance. Table 4.4 and Table 4.5 briefly illustrate how different model parameters affect the performance of the control charts.

It can be observed that in both Table 4.4 and Table 4.5 the perturbation in the parameter of the Φ or Θ matrix can usually cause either a smaller out-of-control ARL at the cost of a higher false alarm rate, or a larger out-of-control ARL with the benefit of a smaller false alarm rate. For an example of MCuscore chart with the Φ matrix in Table 4.4, if the ϕ_1 value perturbs from -0.1 to 0.1 with the other three ϕ's values fixed, we can obtain an out-of-control ARL of 9.18 with the control limit 4.17. This ARL is higher than the expected value of 8.44, but has a smaller false alarm rate than 1/200 because the actual control limit is smaller than the simulated limit of 5.01.

From the perspective of economical design of control charts, such as Duncan (1956) and Woodall (1986), the costs of false alarms and finding an assignable cause are among the major components of the cost model. Such cost model can still be used for the design of new control charts, such as the Cuscore chart, because the purpose of the control chart remains essentially unchanged since it is originally proposed. In practice,

different weights can be chosen accordingly to balance the cost components of the ARL and the false alarms rate in order to minimize the total cost for a controlled process. Therefore, two approaches can be adopted in determining the control limits for designing a Cuscore control chart for monitoring practical processes. If the cost of false alarms is relatively high and enough prior knowledge of the process can be obtained, the control limits of the Cuscore control chart can be designed to generate a large ARL values or a small false alarm rate. If the cost of the false alarm is relatively low, a set of constant and relatively tight control limits can be used which may give an economically robust control chart at the costs of high false alarm rates.

Table 4.4: ARL Properties of the MCuscore and the Residual-based MCusum Chart for VAR(1) Models with Mean Shift $\boldsymbol{\gamma}$

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1 & \phi_{12} \\ \phi_{21} & \phi_2 \end{bmatrix}$$

| | ϕ_1 | ϕ_{12} | ϕ_{21} | ϕ_2 | \multicolumn{3}{c}{$\boldsymbol{\gamma} = (1.0,\ 1.0)'$} | \multicolumn{3}{c}{$\boldsymbol{\gamma} = (1.0,\ 0.5)'$} |

	ϕ_1	ϕ_{12}	ϕ_{21}	ϕ_2	MCuscore H	MCuscore ARL$_1$	Res-MCusum (H = 3.70) ARL$_1$	MCuscore H	MCuscore ARL$_1$	Res-MCusum (H = 4.17) ARL$_1$
(a)	0.7	0.1	0.2	0.5	9.25	46.31	67.94	9.92	55.25	75.12
	0.4	0.1	0.2	0.5	6.96	23.95	34.56	6.20	18.57	23.28
	0.1	0.1	0.2	0.5	5.01	12.16	18.73	4.45	9.44	11.11
	-0.1	0.1	0.2	0.5	4.17	8.44	14.06	3.71	6.97	8.09
	-0.4	0.1	0.2	0.5	3.3	5.53	9.81	2.94	4.81	5.93
	-0.7	0.1	0.2	0.5	2.65	4.10	7.46	2.43	3.66	4.80
(b)	-0.7	-1.2	0.2	0.5	1.30	2.03	4.05	1.69	2.47	3.57
	-0.7	-0.8	0.2	0.5	1.59	2.36	4.67	1.90	2.77	3.83
	-0.7	-0.4	0.2	0.5	1.95	2.82	5.49	2.09	3.07	4.19
	-0.7	0	0.2	0.5	2.43	3.70	7.06	2.36	3.52	4.67
	-0.7	0.4	0.2	0.5	3.22	5.61	9.89	2.64	4.15	5.25
	-0.7	0.8	0.2	0.5	2.62	4.04	15.88	3.43	6.06	6.25
	-0.7	1.2	0.2	0.5	6.82	23.52	33.67	3.57	6.46	7.58
	-0.7	1.6	0.2	0.5	9.13	45.24	82.20	4.25	8.84	10.26

Table 4.5: ARL Properties of the MCuscore and the Residual-based MCusum Chart for Detecting Mean Shift $(1, 1)'$ in VARMA(1,1) Models

$\Theta = \begin{bmatrix} \theta_1 & \theta_{12} \\ \theta_{21} & \theta_2 \end{bmatrix}$				$\Phi = \begin{bmatrix} 0.7 & 0.1 \\ 0.2 & 0.5 \end{bmatrix}$			$\Phi = \begin{bmatrix} -0.4 & 0.1 \\ 0.2 & 0.5 \end{bmatrix}$			$\Phi = \begin{bmatrix} -0.7 & 1.6 \\ 0.2 & 0.5 \end{bmatrix}$		
					MCuscore	Res-MCusum (H = 3.7)		MCuscore	Res-MCusum (H = 3.7)		MCuscore	Res-MCusum (H = 3.7)
θ_1	θ_{12}	θ_{21}	θ_2	H	ARL$_1$	ARL$_1$	H	ARL$_1$	ARL$_1$	H	ARL$_1$	ARL$_1$
0.3	0.4	0.6	0.3	3.22	5.68	6.05	0.44	1.55	5.08	1.25	2.33	3.84
0.3	0.2	0.6	0.3	4.65	9.78	10.21	1.11	2.06	5.16	2.03	3.34	4.23
0.3	0.0	0.6	0.3	5.74	14.85	21.71	1.67	2.70	5.18	2.74	4.44	4.94
0.3	-0.2	0.6	0.3	6.49	19.38	42.81	2.14	3.26	5.16	3.46	6.03	6.27
0.3	-0.4	0.6	0.3	6.98	22.55	70.24	2.62	4.06	5.10	4.12	8.05	8.42
0.3	-0.6	0.6	0.3	7.24	23.93	92.39	3.07	5.02	5.11	4.64	10.13	11.96

4.6 Application Example

To illustrate the MCuscore control chart procedure, we apply it to monitor an RIE process in the semiconductor manufacturing industry. Figure 4.8 illustrates a diagram of a typical RIE setup. An RIE consists of two electrodes that create an electric field in accelerating ions toward the surface of the samples. The plasma contains both positively and negatively charged ions in equal quantities. These ions are generated from the gas pumped into the chamber. In this example, we suppose that O_2 and CF_4 gasses are used for etches. When CF_4 has been pumped into the chamber, a plasma is made with fluorine

(F^-) ions. Then, the fluorine ions are accelerated in the electric field and collide into the surface of the sample to cause the etching.
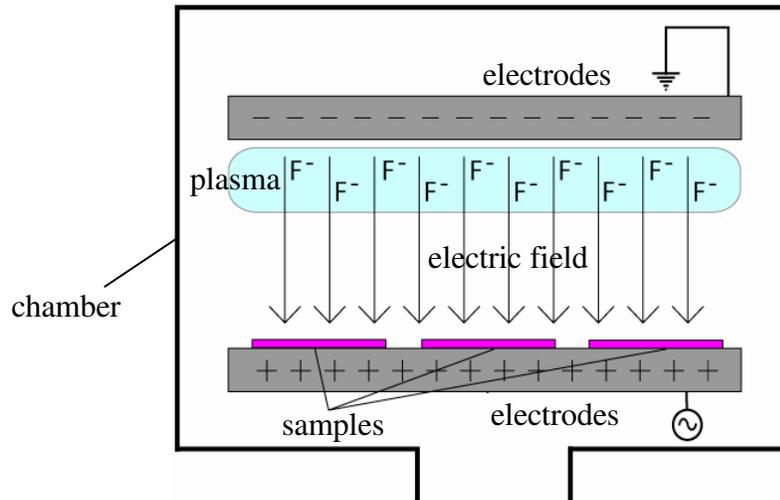
Figure 4.8: Diagram of a typical RIE setup.

The RIE process is complicated because many variables, such as the generator power, chamber pressure, chamber temperature, and so on, are involved in the process and may be correlated, Furthermore shifts in certain variables may cause the process to go out of control in certain runs. For the sake of illustration, assume that in a process setup, the generator power (P_G) and the fluorine line intensity (I_F) are identified as two most significant process variables that affect the process quality, and they follow a bivariate normal distribution with mean vector and variance-covariance matrix

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_{PG} \\ \mu_{IF} \end{pmatrix} = \begin{pmatrix} 850W \\ 703.7nm \end{pmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_1^2 \end{pmatrix} = \begin{pmatrix} 25W^2 & 40W\,gnm \\ 40W\,gnm & 100nm^2 \end{pmatrix}. \text{ Due to the inertia}$$

in the two variables, the collected samples are sequentially correlated and cross-

correlated with correlation coefficient matrix $\Phi = \begin{pmatrix} \phi_1 & \phi_{12} \\ \phi_{21} & \phi_2 \end{pmatrix} = \begin{pmatrix} -0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix}$. Therefore, a

VAR(1) model can be used to describe the process.

Assume at time $t_0 = 10$, a long-term steady mean shift signal with size

$\gamma = \begin{pmatrix} \gamma_{PG} \\ \gamma_{IF} \end{pmatrix} = \begin{pmatrix} 2.5W \\ 5.0nm \end{pmatrix}$ is introduced in the process, and 30 observations are collected

(Figure 4.9(a)). Therefore, the pattern matrix is $\Gamma_t = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & t < 10 \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & t \geq 10 \end{cases}$ and the shift signal

$\Gamma_t \gamma = \begin{pmatrix} 2.5W \\ 5.0nm \end{pmatrix}$ for $t \geq 10$. Assume the Σ matrix is constant during the process, and to

facilitate analysis, the means of the two variables are transformed to zeros while keeping

the Σ matrix unchanged. We apply the MCuscore control chart to monitor the process. To

determine the control limit of the MCuscore chart, we simulated 5000 runs of VAR(1)

processes and each has 1000 observations, and found the control limits ± 5.16 give an in-

control ARL 200. Using these control limits in the MCuscore chart, the mean shift signal

is detected in 6 observations (Figure 4.9(c)).

In comparison, we also apply the residual-based MCusum control chart to monitor

the process. The control limits of ± 3.35 are determined by simulation to give an in-

control ARL 200 in the MCusum chart. The mean shift signal is detected in 15

observations (Figure 4.9(d)).
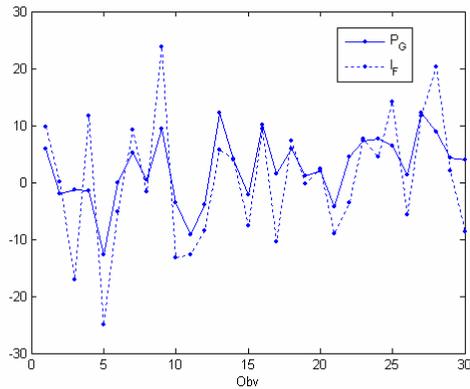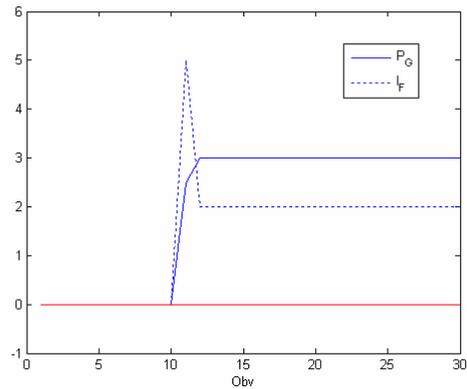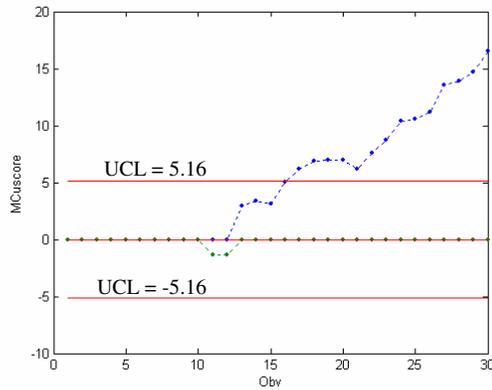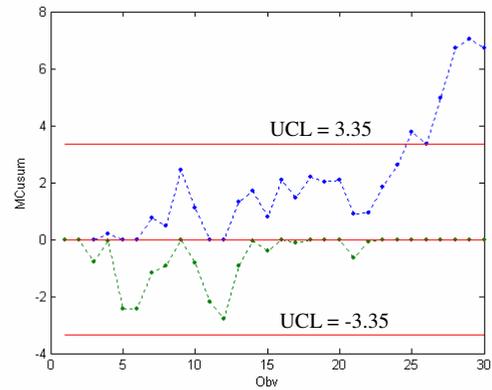
Figure 4.9: (a) Time series plot of process variables P_G and I_F; (b) Fault signature of P_G and I_F; (c) The MCuscore chart for P_G and I_F; (c) The residual-based MCusum chart for P_G and I_F.

4.7 Diagnosis of Out-of-Control Signals

For multivariate process control, the diagnosis of out-of-control signals is crucial because continuously maintaining the multivariate process at its optimal condition relies on the early and accurate isolation of the fault variables. There has been much previous discussion in the literature on this issue with regard to traditional multivariate control

charts, such as the T^2 chart, MEWMA chart, and MCusum chart (see Montgomery, 2005; Yoon and MacGregor, 2001; Mason et al., 1997; Sullivan and Woodall, 1996; and Venkatasubramanian et al., 2003). For the T^2-based multivariate control charts, including the MCuscore chart, fault diagnosis is essentially the problem of decomposing the T^2 statistics into components and evaluating their individual contributions.

Theoretically, the diagnosis step is not necessary when using the MCusocre chart to detect mean shift signals because one of the fundamental assumptions for the MCuscore chart is that the size of the mean shift vector is known or has been estimated based on the historical data. However, for some multivariate processes with a relatively high false alarm rate, the alarms should be evaluated to determine whether or not it is due to a real signal in practice. In this case, the MCuscore chart can serve to check the fired alarms by using the detected signal as prior knowledge and comparing the new detection results with the previously detected ones, or simply comparing the detection results with the earliest system knowledge. As such, fault diagnosis may be useful for verifying the specific responsible variables that cause the signals.

For a sequentially independent process monitored by the MCuscore chart, the approach suggested by Runger, Alt, and Montgomery (1996) can be used for fault diagnosis. This approach is to decompose the T^2 statistics by using an indicator to denote the relative contribution of the ith variable to the overall statistic, and to use $\chi^2_{\alpha,1}$ as the approximate control limit with α the significance level. For example, $\chi^2_{\alpha,1} = 6.63$ for $\alpha = 0.01$. For the MCuscore chart, the indicator can be expressed as

$$d_i = \Delta S - \Delta S_{(i)}, \tag{4.23}$$

where ΔS is the current score of the MCuscore statistic, and $\Delta S_{(i)}$ is the score for all process variables except the ith one. Therefore, the ith variable whose d_i value is out-of-control limit is suspected to be the fault variable. However, this approach is not recommended for autocorrelated processes in that the covariance matrix may shift significantly due to process autocorrelation, and thus cannot be explained accurately by the d_i statistic for individual samples.

4.8   Conclusions

The Cuscore control chart is a powerful statistical process monitoring tool when there is prior knowledge about a process shift. It is also effective for monitoring autocorrelated processes when the process autocorrelation has been estimated from historical data. In this chapter, the multivariate Cuscore approach based on the likelihood ratio test and fault signature analysis is introduced for monitoring the mean vector shift in an autocorrelated multivariable process. A bivariate time series model is used to establish the theory and application of the MCuscore chart which can also be used to monitor processes with higher dimensions. Simulation is used to show that the MCuscore chart outperforms the traditional residual-based MCusum control chart in detecting a mean vector shift signal. An example of monitoring the mean shift of two process variables of an RIE process illustrates the use of the MCuscore chart and demonstrates that the MCuscore chart has better performance than the residual-based MCusum chart in monitor

autocorrelated multivariate process when certain information on the process and the signal is known *a priori*.

Appendix 4A: Derivation of the Multivariate Cuscore Statistics

Consider the bivariate VAR(1) time series model with the fault signature of mean vector shift signal:

$$\mathbf{X}_t = (\mathbf{I} - \mathbf{\Phi}B)^{-1}\boldsymbol{\varepsilon}_t + \mathbf{\Gamma}_t\boldsymbol{\gamma},$$

or,

$$(\mathbf{I} - \mathbf{\Phi}B)\mathbf{X}_t = \boldsymbol{\varepsilon}_t + (\mathbf{I} - \mathbf{\Phi}B)\mathbf{\Gamma}_t\boldsymbol{\gamma},\tag{4A.1}$$

where $\boldsymbol{\gamma}$ = the size of mean vector shift, $\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$, $\mathbf{\Gamma}_t = \begin{pmatrix} f_{1t} & f_{12t} \\ f_{21t} & f_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$,

$\mathbf{\Phi} = \begin{pmatrix} \phi_1 & \phi_{12} \\ \phi_{21} & \phi_2 \end{pmatrix}$, and $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

The fault signature of mean vector shift signal can be written in the following scalar form

$$(\mathbf{I} - \mathbf{\Phi}B)\mathbf{\Gamma}_t\boldsymbol{\gamma} = \begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & t < t_0 \\[2ex] \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} & t = t_0 \\[2ex] \begin{pmatrix} \gamma_1(1-\phi_1) - \gamma_2\phi_{12} \\ \gamma_2(1-\phi_2) - \gamma_1\phi_{21} \end{pmatrix} & t > t_0 \end{cases}.\tag{4A.2}$$

The probability density function (PDF) of a multivariate normal distribution is

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right], \qquad (4A.3)$$

where $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_1^2 \end{pmatrix}$ for bivariate normal distribution.

To derive the bivariate Cuscore statistic, we start with the sequential probability ratio test (SPRT) for H_0: $\boldsymbol{\mu}_{0t} = \mathbf{0}$ against H_1: $\boldsymbol{\mu}_{1t} = (\mathbf{I}-\boldsymbol{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$

$$
\begin{aligned}
LR_k &= \prod_{t=1}^{k} \ln \frac{f_1(\mathbf{x}_t)}{f_2(\mathbf{x}_t)} \\
&= -\frac{1}{2}\sum_{t=1}^{k}\left[(\mathbf{x}_t - \boldsymbol{\mu}_{1t})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_{1t}) - (\mathbf{x}_t - \boldsymbol{\mu}_{0t})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_{0t})\right].
\end{aligned}
\qquad (4A.4)
$$

Because the term $(\mathbf{x}_t - \boldsymbol{\mu}_{1t})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_{1t})$ is a function of $\boldsymbol{\Gamma}_t\boldsymbol{\gamma}$, we rewrite the above equation as

$$
\begin{aligned}
LR_k &= -\frac{1}{2}\sum_{t=1}^{k}\left[f(\boldsymbol{\Gamma}_t\boldsymbol{\gamma}) - f(\mathbf{0})\right] \\
&= -\frac{1}{2}\sum_{t=1}^{k}\left[(\mathbf{x}_t - (\mathbf{I}-\boldsymbol{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_t - (\mathbf{I}-\boldsymbol{\Phi}B)\boldsymbol{\Gamma}_t\boldsymbol{\gamma}) - (\mathbf{x}_t - \boldsymbol{\mu}_{0t})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_t - \boldsymbol{\mu}_{0t})\right].
\end{aligned}
\qquad (4A.5)
$$

For the bivariate case, rewriting $f(\boldsymbol{\Gamma}_t\boldsymbol{\gamma})$ as a scalar expression with two variables γ_1 and γ_2, and implementing a Taylor expansion of $f(\boldsymbol{\Gamma}_t\boldsymbol{\gamma})$ with respect to $\boldsymbol{\gamma} = \mathbf{0}$ gives

$$f(\mathbf{\Gamma}_t\boldsymbol{\gamma}) = (\mathbf{x}_t - (\mathbf{I} - \mathbf{\Phi}B)\mathbf{\Gamma}_t\boldsymbol{\gamma})'\mathbf{\Sigma}^{-1}(\mathbf{x}_t - (\mathbf{I} - \mathbf{\Phi}B)\mathbf{\Gamma}_t\boldsymbol{\gamma})$$

$$= f(\gamma_1, \gamma_2)$$

$$= f(\mathbf{0}) + \gamma_1 \frac{\partial f}{\partial \gamma_1}\Big|_{\gamma_1=0} + \gamma_2 \frac{\partial f}{\partial \gamma_2}\Big|_{\gamma_2=0} + \frac{1}{2}\gamma_1^2 \frac{\partial^2 f}{\partial \gamma_1^2}\Big|_{\gamma_1=0} + \frac{1}{2}\gamma_2^2 \frac{\partial^2 f}{\partial \gamma_2^2}\Big|_{\gamma_2=0} + \gamma_1\gamma_2 \frac{\partial^2 f}{\partial \gamma_1 \partial \gamma_2}\Big|_{\gamma_1=0,\gamma_2=0} + \mathrm{L} \ .$$

$$\tag{4A.6}$$

For an **i.i.d.** multivariate Gaussian process, $\mathbf{\Phi} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, thus

$$\frac{\partial f}{\partial \gamma_1}\Big|_{\gamma_1=0} = 2(x_1\sigma_2^2 - x_2\sigma_{12}), \ \ \frac{\partial f}{\partial \gamma_2}\Big|_{\gamma_2=0} = 2(x_2\sigma_1^2 - x_1\sigma_{12}),$$

and

$$\frac{\partial^2 f}{\partial \gamma_1^2}\Big|_{\gamma_1=0} = 2\sigma_2^2, \ \frac{\partial^2 f}{\partial \gamma_2^2}\Big|_{\gamma_2=0} = 2\sigma_1^2, \ \text{and} \ \frac{\partial^2 f}{\partial \gamma_1 \partial \gamma_2}\Big|_{\gamma_1=0,\gamma_2=0} = -2\sigma_{12}.$$

Substituting the above equations into Equation (4A.6) and rearranging terms gives

$$LR_k = \sum_{n=1}^{k}\left[\boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}x_n - \frac{1}{2}\boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}\boldsymbol{\gamma} \right]. \tag{4A.7}$$

Equation (4A.7) can be rewritten in the MCuscore form:

$$S_k = \begin{cases} \max(S_{k-1} + \boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}\mathbf{x}_k - \frac{1}{2}\boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}\boldsymbol{\gamma}, \ 0) > H \\[2ex] \min(S_{k-1} + \boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}\mathbf{x}_k - \frac{1}{2}\boldsymbol{\gamma}'\mathbf{\Sigma}^{-1}\boldsymbol{\gamma}, \ 0) < -H \end{cases}, \tag{4A.8}$$

which can be rewritten in MCusum form (Healy, 1987)

$$S_k = \begin{cases} \max(S_{k-1} + \dfrac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{x}_k}{D} - \dfrac{1}{2}D,\ 0) > H \\ \min(S_{k-1} + \dfrac{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{x}_k}{D} - \dfrac{1}{2}D,\ 0) < -H \end{cases}, \tag{4A.9}$$

where $D = \sqrt{\boldsymbol{\gamma}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}$.

For an **autocorrelated** bivariate process fitted by a VAR(1) model, $\boldsymbol{\Phi} = \begin{pmatrix} \phi_1 & \phi_{12} \\ \phi_{21} & \phi_2 \end{pmatrix}$, a procedure similar to deriving the MCuscore statistics for i.i.d.

processes can be used to obtain the transformed log likelihood ratio statistics in the form

$$LR_k = \sum_{t=1}^{k} \left[\boldsymbol{\mu}_t'\boldsymbol{\Sigma}^{-1}\mathbf{x}_t - \frac{1}{2}\boldsymbol{\mu}_t'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_t \right]. \tag{4A.10}$$

It can still be represented in the generic form of MCusum statistics as

$$S_k = \begin{cases} \max(S_{k-1} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{x}_k - \dfrac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu},\ 0) > H \\ \min(S_{k-1} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{x}_k - \dfrac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu},\ 0) < -H \end{cases}, \tag{4A.11}$$

or the form in Equation (4A.9).

Appendix 4B: SAS Simulation Code for Determining the ARL$_0$ for the RIE Application Example

```
proc iml;
/* Using simulation to find the arl0 for the MCuscore chart.*/

theta = {0 0,0 0};
phi = {-0.4 0.1, 0.2 0.5};
mu0 = {0, 0};
mud = {2.5, 5};
sig = {25 40, 40 100};
siginv = inv(sig);

N = 1000;
r = 5000;
mu = j(2,N,0);
yt = j(2,N,0);
S1 = j(N,1,0);
S2 = j(N,1,0);
flag = 0;
arl0 = 0;
fault = j(2,N,0);
D = j(N,1,0);

h = {5.1 5.12 5.13 5.14 5.15 5.16 5.17 5.18 5.19};
len = 10;
arl0 = j(len,1,0);

do i = 2 to N;
     mu[,i] = mud;
end;

do k = 1 to len;
     do i = 2 to N;
          fault[,i] = mu[,i] – phi*mu[,i–1]+ theta*fault[,i–1];
     end;

     do j = 1 to r;
          call vnormal(xt, mu0, sig, N);
          yt = xt`;
          do i = 2 to N;
               D[i] = sqrt(fault[,i]`*siginv*fault[,i]);
               tmp = fault[,i]`*siginv*yt[,i]/D[i];

               if flag = 0 then do;
                    S1[i] = max(S1[i–1]+tmp–D[i]/2, 0);
                    S2[i] = min(S2[i–1]+tmp+D[i]/2, 0);

                    if S1[i] > h[k] | S2[i] < –h[k] then do;
                         arl0[k] = arl0[k] + i;
                         flag = 1;
                    end;
                    if i = N then arl0[k] = arl0[k] + N;
```

```
                            end;

                    end;
                    flag = 0;
            end;
end;

arl0 = arl0/r-1;
print arl0;

quit;
```

Appendix 4C: SAS Simulation Code for Determining the ARL$_1$ for the RIE Application Example

```
proc iml;
/* Using simulation to find the arl1 for the residual-based MCusum
chart and MCuscore chart. */

theta = {0 0,0 0};
phi = {-0.4 0.1, 0.2 0.5};
mu0 = {0, 0};
mud = {2.5, 5};
sig = {25 40, 40 100};
siginv = inv(sig);

N = 30;
r = 1;
mu = j(2,N,0);
yt = j(2,N,0);
S1 = j(N,1,0);
S2 = j(N,1,0);
MS1 = j(N,1,0);
MS2 = j(N,1,0);
D= j(N,1,0);

arl0 = 0;
fault = j(2,N,0);

hcusum = 3.35;
hcuscore = 5.16;

do i = 11 to N;
      mu[,i] = mud;
end;

do i = 2 to N;
      fault[,i] = mu[,i] - phi*mu[,i-1]+ theta*fault[,i-1];
end;
```

```
ff = fault`;
print ff;

do j = 1 to r;
      call vnormal(xt, mu0, sig, N);
      yt = xt`+fault;

      /*MCusum*/
      DD = sqrt(mud`*siginv*mud);
      do i = 2 to N;
            tmp = mud`*siginv*yt[,i]/DD;

            MS1[i] = max(MS1[i-1]+tmp-DD/2, 0);
            MS2[i] = min(MS2[i-1]+tmp+DD/2, 0);

            if MS1[i] > hcusum | MS2[i] < -hcusum then do;
                  arl0_cusum = i-20;
            end;
      end;

      /*MCuscore*/
      do i = 2 to N;
            D[i] = sqrt(fault[,i]`*siginv*fault[,i]);
            tmp = fault[,i]`*siginv*yt[,i]/D[i];

                  S1[i] = max(S1[i-1]+tmp-D[i]/2, 0);
                  S2[i] = min(S2[i-1]+tmp+D[i]/2, 0);

                  if S1[i] > hcuscore | S2[i] < -hcuscore then do;
                        arl0_cuscore =  i-20;
                  end;
      end;
end;
arl0 = arl0/r-1;

t = 1:N;
t = t`;
yt1 = yt[1,]`;
yt2 = yt[2,]`;
mat = yt1||yt2||S1||S2||MS1||MS2||t;
create newset1 var{yt1,yt2,S1,S2,MS1,MS2,t};
append from mat;
close newset1;

print DD;
print mat;
quit;

proc gplot data=newset1;
      symbol1 width=1 i=join height=1;
      plot yt1*t = 1 yt2*t =2 / overlay;
      plot S1*t = 3 S2*t = 4 / overlay vref=5.16 vref=-5.16 vref=0;
      plot MS1*t = 3 MS2*t = 4 / overlay vref=3.35 vref=-3.35 vref=0;
run;
```

100

Chapter 5

A High-Dimensional Control Chart for Profile Monitoring

Profile monitoring is an important and rapidly emerging area of statistical process control (SPC). In many industries, the quality of processes or products can be characterized by a profile that describes a relationship or a function between a response variable and one or more independent variables, and each profile may consist of a large number of paired observations of such response and independent variables. A change in the profile relationship can indicate a change in the quality characteristic of the process or product and, therefore, needs to be monitored for control purposes. The techniques used in monitoring such quality characteristics are categorized as profile monitoring. We propose a high-dimensional (HD) control chart approach for profile monitoring that is based on the adaptive Neyman test statistic for the coefficients of discrete Fourier transform of profiles. We investigate both linear and nonlinear profiles, and we study the robustness of the HD control chart for monitoring profiles associated with stationary noise. The application of the control chart is illustrated on two simulated woodboard vertical density profile data sets.

5.1 Introduction

One of the major objectives of statistical process control (SPC) charts is to monitor the quality characteristics of a process over time or space in order to detect out-

of-control signals. For some processes, a single variable, with its value in the form of scalar, is used to represent the quality characteristics of a process. For some others, multiple variables are used to adequately characterize the quality of a process, where each data sample is represented in the form of the vector, and the process variance is represented by the covariance matrix. Consequently, departures from the in-control values for any element or element group in the sample vector and the sample covariance matrix suggest the presence of signals. However, due to the rapid development of advanced data acquisition techniques in some systems or processes, samples can be collected in the forms of profiles or curves over time or space and therefore, the quality characteristics consist of observations or measurements taken from the same order of locations or time points for each sample.

We assume that each profile has the same number of the observations/design points and they are equally distributed on the interval of a profile. We further assume N random samples are collected in a historical data set and there are n observations in each sample. We refer to n as the dimensionality of a profile. For the jth sample collected over time or space, there are a sequence of paired observations (x_1, y_{jk}), $j = 1,2,\ldots, N$ and $k = 1,2,\ldots, n$. Suppose a model can be found to relate the independent variable x to the response variable y for each sample, and it is represented by

$$Y_{jk} = m(X_k, \boldsymbol{\beta}_j) + \varepsilon_{jk}, \ j = 1,2,\ldots, N \text{ and } k = 1,2,\ldots, n, \tag{5.1}$$

where $m(X_k, \boldsymbol{\beta}_j)$ is a mean function, or the signal component of a profile, $\boldsymbol{\beta}_j$ is the parameter vector, and ε_{jk} is the noise component.

102

Most of the published research on profile monitoring stemmed from the study of linear profiles associated with white noise component; see Kang and Albin (2000), Kim et al. (2003), and Mahmoud and Woodall (2004). A linear profile can be modeled by

$$Y_{jk} = \beta_{j0} + \beta_{j1} X_k + \varepsilon_{jk}, \; j = 1,2,\ldots, N \text{ and } k = 1,2,\ldots, n, \tag{5.2}$$

Some profiles can take complicated structure, such as the linear regression model

$$\mathbf{Y}_j = \mathbf{X}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \; j = 1,2,\ldots, N \text{ and } k = 1,2,\ldots, n, \tag{5.3}$$

where $\mathbf{Y}_j$ is a n by 1 vector of responses for the jth profile, $\mathbf{X}_j$ is a n by p matrix of the independent or regressor variables, $\boldsymbol{\beta}_j$ is the p by 1 parameter vector, and $\boldsymbol{\varepsilon}_j \sim MVN(\mathbf{0}, \mathbf{I})$ is the n by 1 vector of errors or noises where $\mathbf{I}$ is the identity matrix. Some profiles have more complicated nonlinear structures, in which some sophisticated parametric models, such as the dose-response model in William et al. (2004), William et al. (2006) and Jensen and Birch (2006), or nonparametric regression model, such as wavelet analysis and additive models, see Jin and Shi (2001), Jeong et al. (2004) and Walker and Wright (2003), are used to represent the profile characteristics.

In the profile models expressed by Equations (5.1-5.3), the noise component $\{\varepsilon_{j1}, \varepsilon_{j2}, \ldots, \varepsilon_{jn}\}$ for profile j is not i.i.d. and white noise, but a stationary noise modeled by a autoregressive (AR) time series, which can be equivalently stated as an infinite order moving-average (MA(∞)) time series model (Del Castillo, 2002). For the kth observation in the jth profile, the noise component can be represented by a MA(k-1) model

$$\varepsilon_{jk} = \sum_{l=0}^{k-1} \phi_j^l a_{k-l} = a_k + \phi_j a_{k-1} + \phi_j^2 a_{k-2} + \mathrm{K} + \phi_j^{k-1} a_1, \ k = 1, 2, \ldots, n, \qquad (5.4)$$

where $\{a_k, \ k = 1, 2, \ldots, n\}$ is i.i.d. with distribution $N(0,1)$.

In general, the techniques used in statistically monitoring the process or product profiles are known as *profile monitoring* (Woodall et al., 2004). Effective SPC approaches for profile monitoring have many significant and practical uses in manufacturing. Changes in profile structure introduced by either local or global abnormal observations may suggest an out-of-control or significant profile. Figure 5.1 illustrates two practical examples of profiles. Figure 5.1(a) shows 10 dose-response curves or profiles from a pharmaceutical drug discovery process (William et al., 2004), and Figure 5.1(b) illustrates 24 vertical density profiles of particle boards manufactured in the forest products industry (Walker and Wright, 2002). Profile monitoring provides a useful and convenient approach to solve the important practical problem of statistically identifying the significant profile(s) from each group.
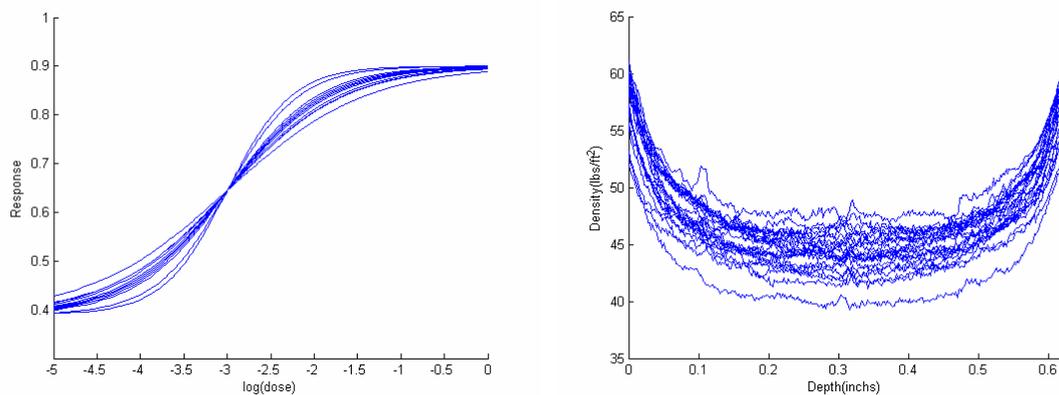
Figure 5.1: (a) 10 dose-response profiles of a drug; (b) 24 vertical density profiles of particle boards.

Traditional Hotelling's T^2-based multivariate SPC charts, such as the multivariate Cusum (MCusum) chart or the multivariate EWMA (MEWMA) chart, put equal weight on each observation and treat the profile as a long measurement vector (the number of observations in a profile can vary from five to hundreds). This approach often cannot effectively detect significantly different profile samples caused by a structure change or a relatively small portion of observation changes (Fan and Lin, 1998).

In this chapter, we propose a profile monitoring approach that is based on the high-dimensional hypothesis testing problem. The adaptive Neyman (AN) test for coefficients and the discrete Fourier transform (DFT) of the profiles are used to develop the model. Our approach can be used to monitor a broad category of profiles, either linear or nonlinear, by using the historical in-control profiles to estimate the profile mean function and variance function in Phase I for monitoring the profiles in Phase II. For some complicated profiles whose model structures are not known, our approach provides a convenient means to detect significantly different profile samples. We also investigate the effect of a stationary noise component on the performance of our profile monitoring approach. Because the dimensionality of the profiles investigated in this research is generally high, from five to hundreds, we refer to our control chart approach for profile monitoring as the high-dimensional (HD) control chart.

The remainder of this chapter is organized as follows. We review the background of profile monitoring in Section 5.2. In Section 5.3 we introduce the adaptive Neyman hypothesis test and its relationship to the high-dimensional control chart. The discrete Fourier transform and its effect on signal compression and decorrelation are illustrated in Section 5.4. We illustrate the HD control chart procedure for profile monitoring in

Section 5.5. The performance evaluation results of the HD control chart are presented in Section 5.6 by simulation, and its application is illustrated using simulated woodboard vertical density profiles. In Section 5.7, we conclude the chapter.

5.2   Background on Profile Monitoring

From an SPC viewpoint, there are three key papers on profile monitoring. Kang and Albin (2000) used two approaches for monitoring linear profiles. The first was to apply a bivariate T^2 chart to monitor jointly distributed slope and intercept variables, and the second was to use a EWMA chart to monitor the residual averages and an R chart to monitor the variance of the residuals along with the regression line. Kang and Albin (2000) recommended using their approaches to monitor both Phase I and Phase II profiles.

Kim et al. (2003) proposed a method for monitoring a linear profile in Phase II. They transformed the estimators of the intercept and the slope variables to be independent by using coded variables, and then recommended the use of three Shewhart control charts in Phase II respectively to monitor the intercept, the slope and the variance of the deviations about the regression line. They showed that their method had better performance than the Kang and Albin (2000) approach in terms of the average run length (ARL).

With the idea of comparing k regression lines collected in Phase I, Mahmoud and Woodall (2004) introduced k-1 indicator variables and constructed a multiple regression model to test whether the kth regression line is statistically significant based on the F-test.

106

They also proposed the use of two Shewhart charts to monitor the coded intercept and slope variables by Kim et al. (2003) for diagnosing the fault variables in a profile. They illustrated their approach by using real data from a calibration process.

One of the key issues with the high-dimensional data monitoring is dimension reduction or feature extraction. Principal component analysis (PCA) (Jackson, 1991) and independent component analysis (ICA) (Hyvarinen et al. 2001) are two popular approaches for dimension reduction. A recent paper by Ding et al. (2006) presented a strategy by using ICA for data-reduction and data-separation in detecting single and multiple shifts in nonlinear profiles. They focused on Phase I analysis. However, applying their method to highly nonlinear profiles, such as those with high frequencies and sharp corners, has not been investigated.

In terms of a statistical organization, profile monitoring is in the scope of functional data analysis (Ramsay and Silverman, 2005; Li and Chow, 2005). It is from this standpoint that many researchers have relied on nonparametric regression or data-driven techniques, such as wavelet thresholding, spline, and local polynomial, etc. for monitoring nonlinear or complicated profiles. Walker and Wright (2002) used an additive model to assess the sources of variation active on vertical density profile data. Their models contained a B-spline to smooth the profile data, and a parametric portion to incorporate other sources of variation.

Jin and Shi (1999, 2001) proposed using wavelet modeling to fit complicated profiles that have sharp corners that contain the most useful information. They relied on engineering knowledge as a prior or "oracle" to determine the local segment for fault diagnosis purposes. Fan (1996) introduced two hypothesis testing techniques for high-

dimensional data, wavelet thresholding and adaptive Neyman's (AN) truncation of Fourier coefficients. These two approaches provide the statistical basis for setting up SPC charts for monitoring high-dimensional data. Jeong et al. (2004) applied wavelet thresholding techniques to monitor complicated profiles by automatically selecting the significant variables for tests. His research was based on one of the approaches to test for significance proposed by Fan (1996). Further, Fan and Lin (1998) illustrated how the two procedures can be applied to test the differences between two sets of curves with i.i.d. noise or even stationary noise fitted by an ARMA model by capitalizing on the fact that the impact of the stationary errors on the null distribution is asymptotically negligible.

Spitzner and Woodall (2003) compared classical multivariate testing approaches with the AN test of Fan and Lin (1998). They noted that high-dimensional profile monitoring departs from classical multivariate testing problem in at least two ways. First, the dimensionality of digitized functional measurements is completely determined by the resolution of the measurement instruments; therefore, the dimensionality of the data can be larger than the available number of observations, which is a violation of the assumption of traditional multivariate testing. Secondly, different weights are given to different dimensions in order to achieve high statistical power, which is different from multivariate testing where equal consideration is placed on every dimension. They applied the AN method to the Fourier coefficients of the vertical density profile dataset and the thickness profile data for silicone nitride film in Gardner et al. (1997).

Woodall et al. (2004) gave a comprehensive review of using control chart to monitor process and product quality profiles. Many research issues on profile monitoring were summarized and new research topics were recommended.

5.3 The Adaptive Neyman Test for Control Charts

Profile monitoring follows the basic approach of functional data analysis that the collection of observed data for all the process variables are treated as a single profile sample, rather than as merely a sequence of individual observations (Ramsay and Silverman, 2005). For example, in the jth dose-response sample in Figure 5.1(a), the sequential values of doses are taken as process variables X_k, $k = 1,2,\ldots,$ n, and the response values corresponding to each dose are the dependent variables Y_{jk}. The relation function between X_k and Y_{jk} is referred to as the profile. In this section, our purpose is to connect the high-dimensional hypothesis testing problem to profile monitoring and combine it with the AN test.

5.3.1 Hypothesis Test and SPC Chart for Profile Monitoring

Similar to the univariate and multivariate SPC chart, the SPC chart for profile monitoring can be viewed as a sequence of tests of hypothesis on the mean and/or variance where the goal is to reject the profile sample which is significantly different from the in-control or historical profiles.

We assume that the process profiles are sequentially independent. Let $f_j(X)$ be the regression function for the jth profile and $\hat{m}(X)$ be the estimated mean profile for population mean profile $m(X)$ in Equation (5.1). To test whether the profile fitted by $f_j(X)$ is statistically different from the estimated mean profile, we set up the null and alternative hypotheses:

$$H_0 : \hat{m}(X) = f_j(X)$$
$$H_1 : \hat{m}(X) \neq f_j(X)$$
(5.5)

Note that we do not confine the forms or structures of the functions $\hat{m}(X)$ and $f_j(X)$. Actually, they can take the form of both linear and nonlinear models, such as $f(X) = \beta_0 + \beta_1 X$ and $f(X) = \beta_0 / \left(1 + \beta_1 \exp(\beta_2 X)\right)$ for β_0, β_1 and β_2. Although it is beyond the scope of this chapter, many parametric or nonparametric regression methods can be applied to fit the profiles (see Fan and Gijbels, 1996 and Ruppert, 2002).

Let $\hat{\varepsilon} = \{\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_n\}$ be the resulting residual vector with the kth element $\hat{\varepsilon}_k = Y_k - \hat{m}(X_k) = [f(X_k) - \hat{m}(X_k)] + \varepsilon_k$. Therefore, if the null hypothesis is not rejected for the model in Equation (5.5), $\hat{\varepsilon}_k$ is nearly i.i.d. distributed and the residual vector $\hat{\varepsilon} \sim MVN(\mathbf{d}, \mathbf{I})$, where $\mathbf{d} = (d_1, d_2, \mathrm{K}, d_n)'$ is the mean vector with $d_k = f(X_k) - \hat{m}(X_k)$ (Fan and Huang, 2001). Consequently, the hypothesis test in Equation (5.5) becomes

$$H_0 : \mathbf{d} = \mathbf{0}$$
$$H_1 : \mathbf{d} \neq \mathbf{0} .$$
(5.6)

Therefore, the hypothesis test for the difference of two profiles is transformed to a problem of testing the difference between a mean vector and the zero vector. Our goal is to find a proper distribution model for $\mathbf{d}$ and establish the control limit for the SPC chart on profile monitoring, given the significance level or the probability of type I error, α.

5.3.2　The Adaptive Neyman Test and Control Limits

Distribution-based approaches such as the Kolmogorov-Smirnov test (KS test) and Cramer-Von Mises test (CVM test) are traditionally used for the hypothesis test in Equation (5.6). However, they suffer low power in detecting densities containing high-frequency components in high-dimensional space (Fan, 1996). To address this issue, Fan (1996) developed an approach to adaptively monitor the high-dimensional process based on the Neyman test (Neyman, 1937; Hart, 1997).

The Neyman test focuses on only the first m-dimensional sub-problem if there is prior knowledge that most of nonzero elements lie on the first m dimensions. In such cases, the test has a χ^2 distribution with m degrees of freedom. As a significant extension of Neyman's earlier work, Fan (1996) proposed the adaptive Neyman test statistic

$$T_{AN}^* = \max_{1 \le p \le n} \frac{1}{\sqrt{2p}} \sum_{k=1}^{p} \left(Z_k^2 - 1 \right) \tag{5.7}$$

where $\mathbf{Z} \sim MVN(0, \mathbf{I}_n)$ is an n-dimensional normal random vector. The approach works by *adaptively* selecting the first p coefficients for testing the hypothesis. Fan et al. (2001) proved that the statistic is equivalent to rejecting H_0 when

$$T_{AN} = \sqrt{2 \log \log n} T_{AN}^* - \{2 \log \log n + 0.5 \log \log \log n - 0.5 \log(4\pi)\}$$

is significantly large.

We assume the observed profile data in Phase I and II are random samples from two models respectively given by

$$Y_{1jk} = f_{1j}(X_k) + \varepsilon_{1jk}, \quad k = 1, 2, \text{K}, n, \quad j = 1, 2, \text{K}, N_1, \tag{5.8}$$

and

$$Y_{2jk} = f_{2j}(X_k) + \varepsilon_{2jk}, \quad k = 1, 2, \text{K}, n, \quad j = 1, 2, \text{K}, N_2, \tag{5.9}$$

where Y_{ijk} indicates the kth observation value in the jth profile of Phase i ($i = 1$ and 2), $f_{ij}(X_k)$ indicates the kth fitted value in the jth estimated profile of Phase i, the random variables $\varepsilon_{1jk} \sim N(0, \sigma_{1 \cdot k}^2)$ and $\varepsilon_{2jk} \sim N(0, \sigma_{2 \cdot k}^2)$ are assumed to be independent heteroscedastic errors for all j, N_1 is the number of profile data in Phase I, and N_2 is the number of profiles data considered in a particular subgroup in Phase II.

Fan and Lin (1998) recommended standardizing profile differences by

$$Z_k = \frac{\overline{Y}_{1 \cdot k} - \overline{Y}_{2 \cdot k}}{\sqrt{\hat{\sigma}_{1 \cdot k}^2 / N_1 + \hat{\sigma}_{2 \cdot k}^2 / N_2}}, \tag{5.10}$$

where

$$\overline{Y}_{i \cdot k} = (1 / N_i) \sum_{j=1}^{N_i} Y_{ijk} \tag{5.11}$$

and

$$\hat{\sigma}_{i \cdot k}^2 = (N_i - 1)^{-1} \sum_{j=1}^{N_i} \left(Y_{ijk} - \overline{Y}_{i \cdot k} \right) \tag{5.12}$$

for $i = 1$ and 2. When N_1 and N_2 are reasonably large, the vector $\mathbf{Z} = (Z_1, \ldots, Z_n)'$ converges to a multivariate normal distribution $MVN(\mathbf{d}, \mathbf{I})$ under the null hypothesis in Equation (5.6). Therefore, the adaptive Neyman test statistic in Equation (5.7) can be

applied to **Z**. The resulting values of T_{AN} are used as the statistic on the high-dimensional control chart.

To construct the control limits, we use the asymptotic distribution of T_{AN} under H_0 illustrated by Fan and Lin (1998)

$$P\{T_{AN} \leq x\} \rightarrow \exp\{-\exp(-x)\}$$ (5.13)

The CDF and PDF of the distribution of T_{AN} are plotted in Figure 5.2.
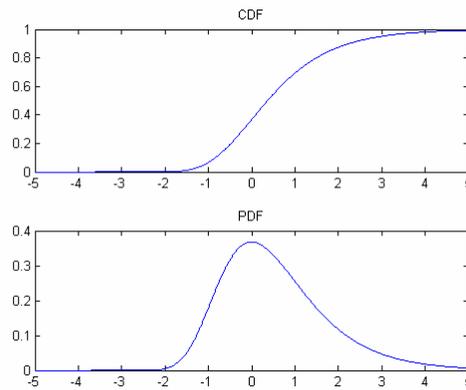
Figure 5.2: CDF and PDF of the distribution of T_{AN}.

Table 5.1: $\alpha = 0.005$ Upper Quartile of the Distribution T_{AN} (Fan and Lin, 1998).

n	5	10	20	30	40	50	60	70	80	90	100	120	140	160	180	200	∞
CL	5.97	6.77	7.16	7.29	7.41	7.43	7.51	7.55	7.57	7.65	7.65	7.65	7.66	7.69	7.77	7.72	5.30

Table 5.1 gives the finite sample distribution of T_{AN} for significance level $\alpha = 0.005$ upper quartile provided by Fan and Lin (1998), which can be used as the control limits of the SPC chart for profile monitoring at certain dimensionality, or the number of observations in the profile. For example, if the dimensionality $n = 100$ and the

in-control average run length (ARL) is 200 (which is equivalent to a significant level $\alpha = 0.005$) then the control limit is CL $= 7.65$.

5.4 Discrete Fourier Transform for the Adaptive Neyman Test

We have discussed that the adaptive Neyman test statistics are maximized by automatically selecting the residuals of the first m variables. Therefore, the order of the residuals is important for the adaptive Neyman test statistics. To order the residuals, we were motivated to use a Fourier transform. We found that it also transformed stationary autocorrelated noise (if any) in the residuals to approximately independent noise.

5.4.1 Discrete Fourier Transform

Fourier analysis provides an approach to transform a continuous function from time domain to its counterpart in the frequency domain for analysis. Since real processes are discrete given sufficiently small sampling intervals, the discrete Fourier transform (DFT) is more useful in analyzing practical processes. The DFT is a projection method that is used to analyze the frequencies contained in a vector of a discrete data set $\mathbf{Y} = (Y_1, Y_2, \mathrm{K}, Y_n)'$ by projecting it to the Fourier basis

$$\mathbf{\Psi}(\omega_t) = \left(1, e^{i\omega_t}, e^{i2\omega_t}, \mathrm{K}, e^{i(n-1)\omega_t}\right), \ t = 1, 2, \mathrm{K}, n,$$

where $i = \sqrt{-1}$ is the imaginary unit, $\omega_t = 2\pi(t-1)/n$ are Fourier frequencies on the interval $[0, 2\pi)$.

The Fourier basis can be written in the design matrix form

$$\Psi(\omega_t) = \begin{pmatrix} 1 & 1 & 1 & L & 1 \\ 1 & e^{i\omega_2} & e^{i2\omega_2} & L & e^{i(n-1)\omega_2} \\ 1 & e^{i\omega_3} & e^{i2\omega_3} & L & e^{i(n-1)\omega_3} \\ M & M & O & & M \\ 1 & e^{i\omega_n} & e^{i2\omega_n} & L & e^{i(n-1)\omega_n} \end{pmatrix}. \tag{5.14}$$

Then the DFT of $\mathbf{Y}$, denoted by a vector of complex number $\mathbf{Y}^*$, can be written as

$$\mathbf{Y}^* = \bar{\Psi}(\omega_t)\mathbf{Y} = \left(\sum_{k=1}^{n} Y_k \quad \sum_{k=1}^{n} Y_k e^{-i(k-1)\omega_2} \quad K \quad \sum_{k=1}^{n} Y_k e^{-i(k-1)\omega_n} \right)', \tag{5.15}$$

where $\bar{\Psi}(\omega_t)$ is the complex conjugate of the Fourier basis $\Psi(\omega_t)$.

The Fourier basis functions can also be rewritten in a matrix of triangle functions by representing the basis $e^{ik\omega_t}$ in the polar form, $e^{i\omega_t} = \cos(k\omega_t) + i\sin(k\omega_t)$. Then, the Fourier basis functions in Equation (5.15) can be written in the triangle function basis. When n is even,

$$\Psi(\omega_t) = \begin{pmatrix} 1 & 1 & 1 & L & 1 & 1 & 1 \\ 1 & \cos(\omega_2) & \sin(\omega_2) & L & \cos(\frac{n}{2}\omega_2) & \sin(\frac{n}{2}\omega_2) & \cos((\frac{n}{2}+1)\omega_2) \\ 1 & \cos(\omega_3) & \sin(\omega_3) & L & \cos(\frac{n}{2}\omega_3) & \sin(\frac{n}{2}\omega_3) & \cos((\frac{n}{2}+1)\omega_3) \\ M & M & M & O & M & M & M \\ 1 & \cos(\omega_n) & \sin(\omega_n) & L & \cos(\frac{n}{2}\omega_n) & \sin(\frac{n}{2}\omega_n) & \cos((\frac{n}{2}+1)\omega_n) \end{pmatrix},$$

(5.16)

and when n is odd,

$$\Psi(\omega_t) = \begin{pmatrix} 1 & 1 & 1 & L & 1 & 1 \\ 1 & \cos(\omega_2) & \sin(\omega_2) & L & \cos(\frac{n+1}{2}\omega_2) & \sin(\frac{n+1}{2}\omega_2) \\ 1 & \cos(\omega_3) & \sin(\omega_3) & L & \cos(\frac{n+1}{2}\omega_3) & \sin(\frac{n+1}{2}\omega_3) \\ M & M & M & O & M & M \\ 1 & \cos(\omega_n) & \sin(\omega_n) & L & \cos(\frac{n+1}{2}\omega_n) & \sin(\frac{n+1}{2}\omega_n) \end{pmatrix}.$$

(5.17)

Theoretically, the time series Y_k, $k = 1, 2, \ldots, n$, can be expressed as a sum of weighted Fourier basis in the frequency domain, e.g., if n is odd, it can be expressed by

$$Y_k = A_0 + \sum_{t=1}^{(n-1)/2} \left[A_t \cos(k\omega_t) + B_t \sin(k\omega_t) \right], \, k = 1, 2, \ldots, n,$$

(5.18)

where the A_t's and B_t's are Fourier coefficients. See Brockwell and Davis (1991), Del Castillo (2002), and Boggess and Narcowich (2002) for more details on the use of DFT in time series analysis.

116

5.4.2 Using the Fourier Transform for Data Ordering Prior to the Application of the Adaptive Neyman Test

The Fourier transform provides an effective approach for compressing useful signals into low frequencies. In general, when smoothing a data set by a function $f(X)$, the smoother the function $f(X)$ is, the more significant are the Fourier coefficients at low frequencies. Therefore, the adaptive Neyman test is more powerful on a discrete data set that is preprocessed by the Fourier transform because the transform yields a smoother function $f(X)$ that has more significant coefficients at low frequencies.

For example, the jth profile consists of a discrete paired signal data set (X_1, Y_1), (X_2, Y_2), …, (X_{100}, Y_{100}) with X_k equally distributed on the interval [1, 100], and the data are smoothed by a log function

$$Y_{jk} = \beta_0 + \beta_1 \log(X_k), \text{ for } j = 1,2,\ldots, N, \text{ and } k = 1,2,\ldots, n, \tag{5.19}$$

where β_0 and β_1 represent the function parameters. Without loss of generality, we let $\beta_0 = 1$ and $\beta_1 = 1$ to illustrate the Fourier transform. The Fourier coefficients can be obtained by projecting the discrete data set $\{Y_k, k = 1,2,\mathrm{K}, n\}$ onto the basis function in Equation (5.14). A plot of the function and its DFT coefficients over $\{Y_k, k = 1,2,\mathrm{K}, n\}$ are shown in Figure 5.3. It can be observed in Figure 5.3(b) that most of the Fourier coefficients decrease exponentially along with the increase of the frequency levels. In particular, the constant parameter a affects only the first Fourier coefficient, which is the summation of Y_k. For the rest of the Fourier coefficients, β_0 has no effect. This property will be used to monitor a constant mean shift among profiles in the later part of this chapter.

<center>(a)   (b)</center>
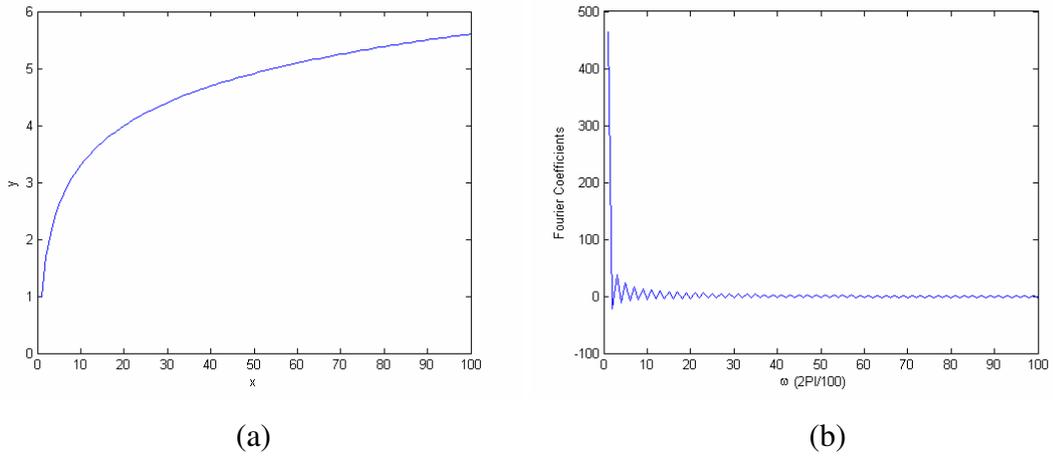
Figure 5.3: (a) Plot of function $y = 1+\log(x)$; (b) Plot of the DFT of function $y = 1+\log(x)$.

5.4.3   Using the Fourier Transform to Decorrelate Stationary Noise Prior to the Application of the Adaptive Neyman Test

In some cases, the profile data may contain stationary noise or error, ε_t. We will illustrate in this section that the adaptive Neyman test is still applicable by preprocessing the profile data via the Fourier transform. The Fourier transform can be used to convert stationary errors into approximately independent Gaussian errors (Brockwell and Davis, 1991 and Fan and Lin, 1998).

For example, suppose the noise $\{\varepsilon_k, k = 1, 2, \mathrm{K}, n\}$ follows an AR(1) model with parameter ϕ, and it can be transformed into an MA(k-1) model represented in Equation (5.4). The DFT of the stationary noise vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \mathrm{K}, \varepsilon_n)'$ on the Fourier basis (Equation (5.14)) can be represented in the vector form

118

$$\boldsymbol{\varepsilon}^{*}(\omega_{t}) = \overline{\boldsymbol{\Psi}}(\omega_{t}) \cdot \boldsymbol{\varepsilon} = \frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n}\varepsilon_{k}e^{-i(k-1)\omega_{1}}\ \mathrm{K}\ \sum_{k=1}^{n}\varepsilon_{k}e^{-i(k-1)\omega_{n}}\right)', \ t = 1,2,\mathrm{K}, n. \tag{5.20}$$

where $\overline{\boldsymbol{\Psi}}(\omega_{t})$ is the complex conjugate of the Fourier basis $\boldsymbol{\Psi}(\omega_{t})$, and $\{\omega_{t} = 2\pi(t-1)/n, t = 1,2,\ldots, n\}$ are Fourier frequencies on the interval $[0, 2\pi)$.

Brockwell and Davis (1991) proved that the DFT of stationary noise $\varepsilon(x_{k})$ is approximately an uncorrelated sequence when n is relatively large

$$\varepsilon^{*}(\omega_{t}) = \left(\sum_{k=1}^{n}\phi^{(k-1)}e^{-i(k-1)\omega_{t}}\right)a^{*}(\omega_{t}), \ t = 1,2,\ldots, n, \tag{5.21}$$

where $a^{*}(\omega_{t})$ is the DFT of white noise $a(x_{k})$ and it is still a white noise with mean zero and variance scaled by K, $N(0, K)$ (Fan, 1996). Note that the $\left(\sum_{k=1}^{n}\phi^{(k-1)}e^{-i(k-1)\omega_{t}}\right)$ term is equivalent to the power spectrum of the stationary noise. The power spectrum shows how the total variance of a stochastic process is distributed across all possible frequencies. It can be used to explain theoretically the pattern of Fourier coefficients for the stationary noise, see Brockwell and Davis (1991) and Del Castillo (2002).

The DFT coefficient sequence of an AR(1) time series with $\phi = 0.5$ and length 200 is plotted in Figure 5.4. Figure 5.4(a) illustrates that larger DFT coefficients tend to distribute to lower frequencies. In addition, Figure 5.4(b) illustrates that the DFT coefficients are approximately normally distributed, and Figure 5.4(c) and (d) illustrate that the autocorrelation among the DFT coefficients is negligible.

A more interesting observation can be made from the DFT coefficients of a negative uncorrelated AR(1) time series. For example, Figure 5.5 illustrates the DFT coefficients of a AR(1) time series with $\phi = -0.5$ and length 200. It can be observed that larger coefficients tend to distribute to lower frequencies, although the autocorrelation in the observations of a profile is removed in the DFT coefficients which approximately follow a normal distribution.

Depending on the amount of autocorrelation in a stationary process, the distribution of their DFT coefficients varies significantly. Figure 5.6 compares the plots of DFT for AR(1) models with different ϕ values. It can be observed that the distribution of large DFT coefficients moves from high frequencies to low ones as ϕ changes from large negative value to large positive value within the range of [-1, 1] in order for the time series to be stationary.
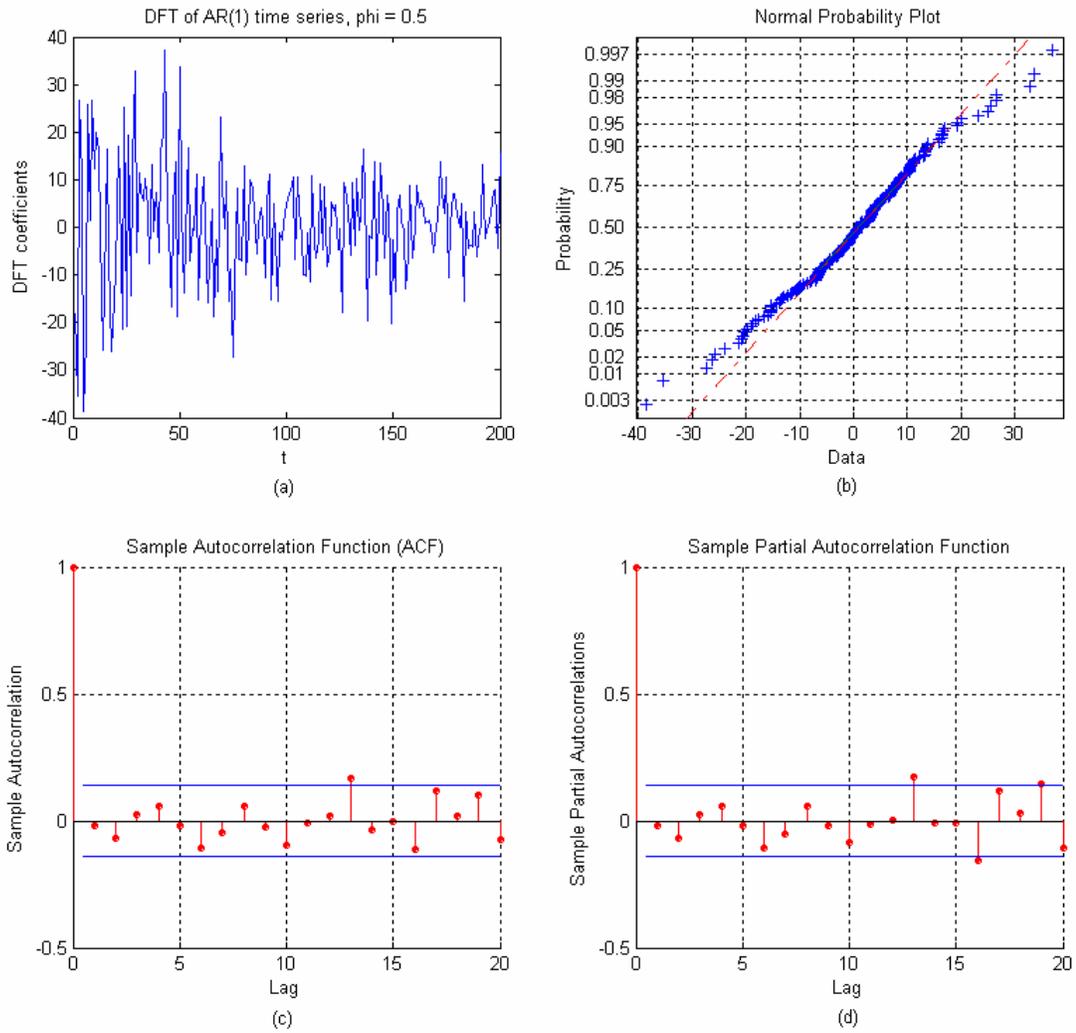
Figure 5.4: (a) The DFT of AR(1) time series with $\phi = 0.5$. (b) The normality plot of the DFT. (c) The autocorrelation function of the DFT. (d) The partial autocorrelation function of the DFT.
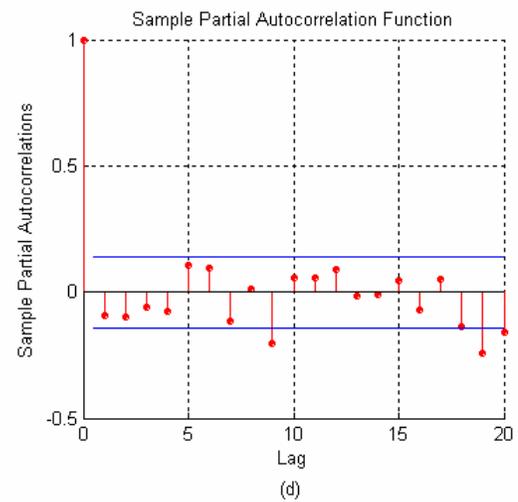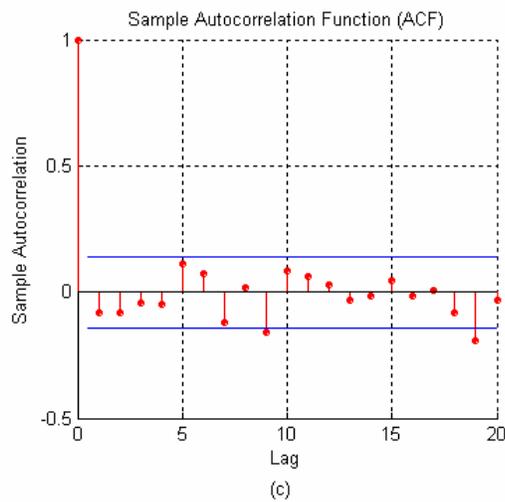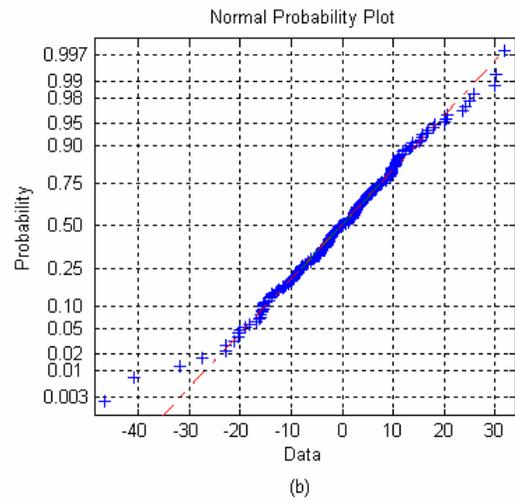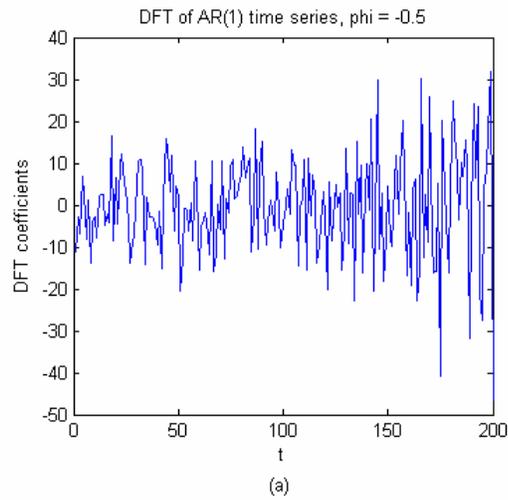
Figure 5.5: (a) The DFT of AR(1) time series with $\phi = -0.5$. (b) The normality plot of the DFT. (c) The autocorrelation function of the DFT. (d) The partial autocorrelation function of the DFT.
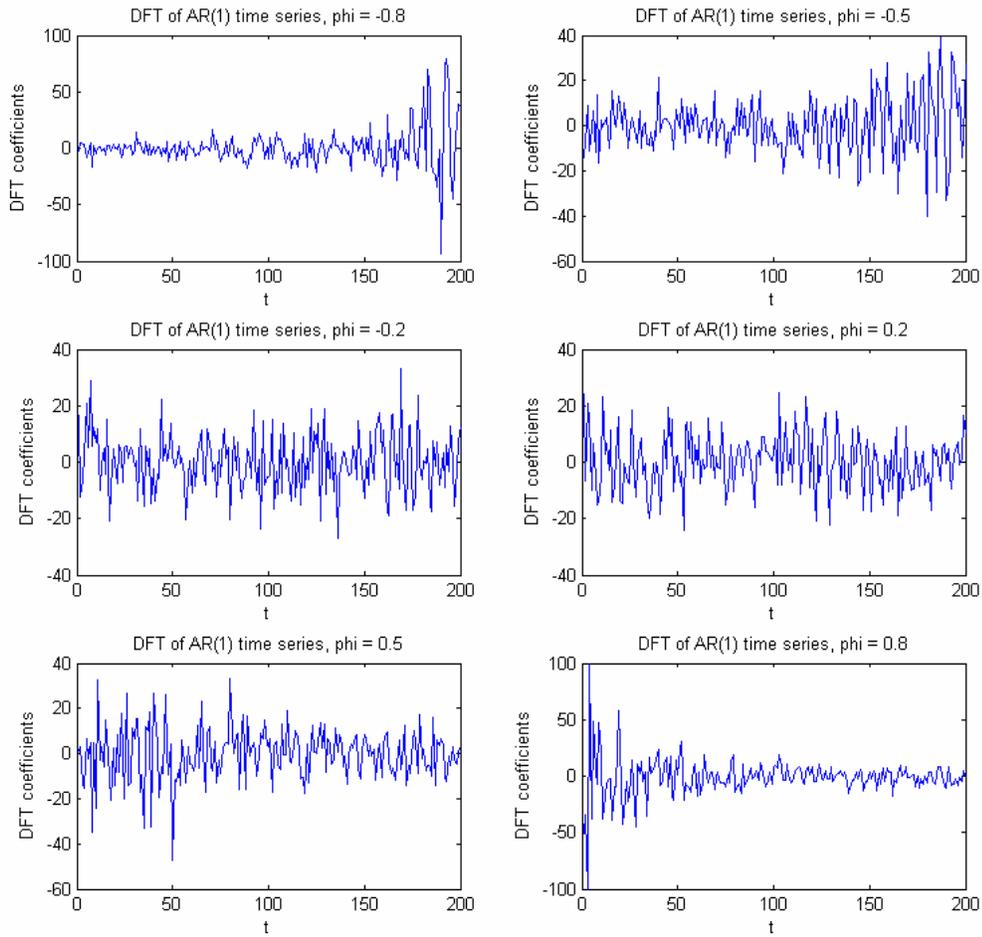
Figure 5.6: The DFT of AR(1) time series with ϕ = -0.8, -0.5, -0.2, 0.2, 0.5 and 0.8.

5.5 The HD Control Chart Procedure for Monitoring Profiles with Stationary Noise

In this section, we present the control chart procedure for monitoring profiles based on the AN test and the discrete Fourier transform. This control chart approach can be used to monitor profiles with its signal component represented by either a linear or nonlinear function, and its noise component represented by either white noise or

stationary AR(1) noise. Ideally, the nonlinear function is fairly smooth so that its Fourier

coefficients are more likely to be concentrated on the low frequency levels.

5.5.1 The HD Control Chart Procedure Based on the DFT and AN Test for Monitoring Profiles with Stationary Noise

We assume that a sufficiently large number of independent profiles can be

collected in Phase I, where the process variance can be assumed to be only due to the

common causes by the noise components of profiles.

Applying the DFT to the noise component of each profile using Equation (5.16)

or (5.17), we then estimate the mean and variance of the tth DFT coefficients of profiles

by

$$\bar{\varepsilon}_{gt}^{*} = (1/N) \sum_{j=1}^{N} \varepsilon_{jt}^{*}, \tag{5.22}$$

and

$$\hat{\sigma}_{\varepsilon_{gt}^{*}}^{2} = (N-1)^{-1} \sum_{j=1}^{N} \left(\varepsilon_{jt}^{*} - \bar{\varepsilon}_{gt}^{*} \right)^{2}, \tag{5.23}$$

where ε_{jt}^{*} is the tth DFT coefficient of the residuals from j^{th} profile, $t = 1,2,\ldots, n$, and $j =$

$1,2,\ldots, N$, and N is the number of profiles in Phase I.

As is discussed in the previous section, when $\left\{ \varepsilon_{j1}, \varepsilon_{j2}, \ldots, \varepsilon_{jn} \right\}$ is a stationary noise

series, its DFT coefficients are skewed to either low or high frequency levels depending

on the autoregressive coefficients ϕ. Intuitively, the variance function of the DFT coefficients of profile noise is also skewed in the same direction.

We summarize the control chart procedure based on the AN test and discrete Fourier transform (DFT) for profile monitoring in both Phase I and II.

For a Phase I process:

1. Take the average of the profile samples and obtain the estimated point-wise mean profile function $\{\hat{m}(X_k), \ k = 1, 2, K, n\}$.

2. Obtain the residuals for each profile sample by subtracting $\hat{m}(X_k)$.

3. Apply the DFT to the residuals of each profile sample.

4. Obtain the standardized DFT coefficients of each profile residuals by using

$$ Z_{jt}^* = \frac{\varepsilon_{jt}^* - \bar{\varepsilon}_g^*}{\sqrt{\hat{\sigma}_{\varepsilon_g^*}^{*2} / N}}, \tag{5.24} $$

where $\bar{\varepsilon}_g^*$ and $\hat{\sigma}_{\varepsilon_g^*}^2$ are from Equation (5.22) and Equation (5.23) respectively.

5. Obtain the AN statistics for the standardized DFT coefficients for each profile and plot them sequentially on the control chart.

6. Based on the dimensionality n, or the number of observations in each profile, select proper control limit for the control chart from Table 5.1. Use interpolation to find the corresponding control limit for any n that does not appear in the table. For n larger than 200, use a control limit of 5.3.

7. If any out-of-control AN statistics are detected, remove the corresponding profile samples from the Phase I data set, and apply steps 2-5 iteratively to the rest Phase I profiles until no out-of-control profile samples are detected.

8. Obtain the final point-wise estimated mean profile by taking average of the remaining Phase I profile samples.

For Phase II process:

1. Standardize the Phase II profile samples by using Equation (5.22-5.24), where the point-wise mean profile is estimated from Phase I.

2. Obtain the AN test statistics for the standardized DFT coefficients for each profile and plot them on the control chart sequentially.

3. If any out-of-control AN statistic is detected, report a profile outlier.

This control chart procedure for profile monitoring in both Phase I and II is illustrated by the diagram in Figure 5.7. Its performance will be evaluated using simulation in Section 5.6.
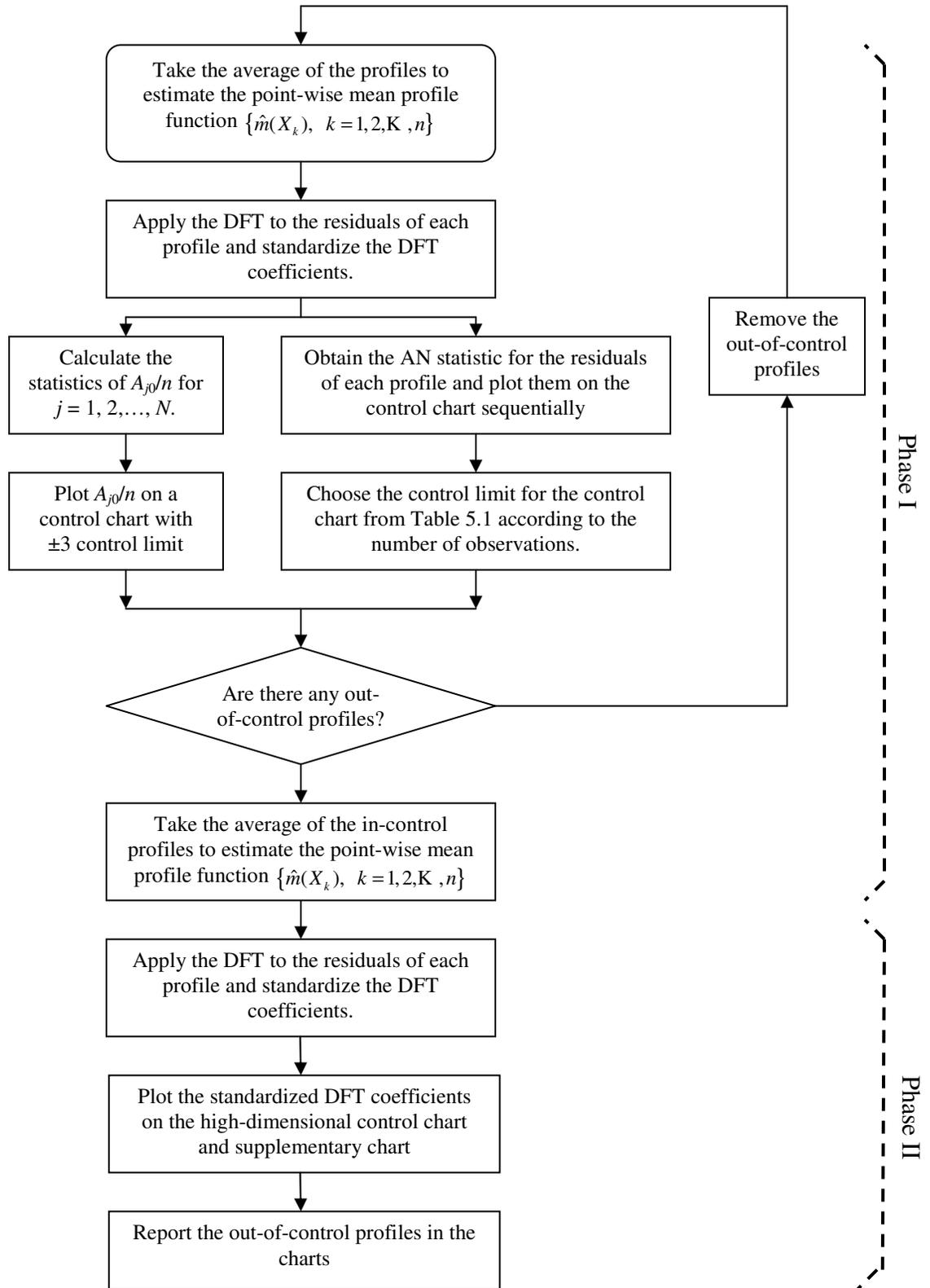
Figure 5.8: The diagram of the high-dimensional control chart approach

5.5.2 Using a Supplementary Chart to Monitor Profiles with Constant Global Mean Shifts

Theoretically, the DFT coefficients of a series of constant values are zeros except the first coefficient A_0 in Equation (5.15) and (5.18). Therefore, the parameter β_0 in the profile modeled by Equation (5.19) has no effect on DFT coefficients except A_0. Therefore, the HD control chart is expected to perform poorly in monitoring profiles with constant global mean shifts and requires a supplementary chart for this purpose.

We propose to setup a control chart for the first DFT coefficient A_0 alone in order to monitor the β_0 shift in the diagram of Figure 5.7. As is shown in Equation (5.15), A_{j0} for the jth profile residuals is the summation of the values of the jth profile residuals. Dividing A_{j0} by the number of observations, n, gives the mean of the jth profile residuals. Then take standardization of the statistics by dividing them by the standard deviation of A_{j0}, $j = 1, 2, \ldots, n$, and plot them on the control chart sequentially. Actually, the control chart for monitoring such standardized statistics with $\pm 3\sigma$ control limits is actually a $\bar{x}$ chart for the residuals of the jth profile. We note that an alternate approach is to monitor the average of the residuals for linear profile samples by using a EWMA chart as in Kang and Albin (2000).

5.6   Performance Evaluation of the HD Control Chart

In this section, we use simulation to study the performance of the HD control chart approach in monitoring both linear and nonlinear profiles with either white noise or stationary noise. The performance is evaluated by using both the average run length (ARL) and the 25th, 50th, and 75th percentile run lengths. We also use simulated woodboard vertical density profiles to illustrate the application of the HD control chart and assess its ability to detect an outlier.

5.6.1   The Run Length Distribution for HD Control Charts

The run length of a control chart is a discrete random variable that is defined as the number of plotted statistics before an out-of-control point is observed on the chart. The ARL is the expected value of this random variable and it is the most popular criteria for evaluating the performance of control charts in Phase II monitoring.

Theoretically, the run length follows a geometric distribution for Shewhart charts that assumes a known process mean and variance. However, for non-Shewhart control charts or charts in which the process parameters are unknown and require estimation, the run lengths are no longer geometrically distributed and take a complicated form (Del Castillo, 2002; Jones et al., 2001, 2004; and Shu et al., 2004).

For the HD control chart approach, it basically follows a Shewhart chart model since the profile samples are assumed to be independent and the plotted AN statistics are not cumulatively considered. We use Equations (5.22) and (5.23) to estimate the sample

point-wise mean function and squared standard deviation function based on the in-control profiles. For the HD control chart, the mean profile function and variance function are unknown and estimated from historical profile samples. Therefore, its run length does not follow a geometric distribution.

In this section, we obtain the run length distribution of the HD control charts by using simulation, and present it in ARLs, 25th, 50th and 75th percentiles of the distribution. We assume there are shifts of parameters β in three different underlying profile models, one nonlinear and two linear.

Example 1. *Nonlinear profile with stationary AR(1) noise*

Consider the profile whose signal component is modeled by Equation (5.19) repeated here for convenience

$$Y_{jk} = \beta_0 + \beta_1 \log(X_k), \text{ for } j = 1,2,\ldots, N, \text{ and } k = 1,2,\ldots, n, \tag{5.25}$$

and whose noise component ε is modeled by an AR(1) time series with autoregressive coefficient ϕ. Without loss of generality, we let $\beta_0 = 0$ and study the impact of shifts with the parameter β_1, $\Delta\beta_1$, from 0 to 0.1, on the run lengths of the control chart. Meanwhile, the autoregressive coefficient ϕ varies from -0.8 to 0.8 with step size 0.3. We also study impact of the profile dimensionality n on the run lengths by varying it from 5 to 50. Upper control limits are chosen from Table 5.1 for significance level $\alpha = 0.005$, or in-control ARL = 200. Two thousand runs with 1000 profiles in each run are simulated.

The simulation results of ARLs and quartiles of run lengths are summarized in Table 5.2. For $\Delta\beta_1 = 0$ and ϕ on the interval [-0.5, 0.5], the ARLs are close to the in-control ARL value of 200 which suggests that the control chart is robust for Phase I monitoring when no shifted signal is presented. For Phase II monitoring with $\Delta\beta_1 > 0$, the robustness to noise autocorrelation of the control chart is not ideal even for a very small ϕ interval [-0.3, 0.3]. In addition, for $\Delta\beta_1 > 0$, the ARLs decrease but at higher rates for large dimensionality n and small ϕ, and at smaller rates for the opposite values of n and ϕ. The dimensionality n has a significant impact on the run lengths because the increase with n (when $n > 5$) causes the decrease of run length. The run lengths are not affected obviously by small dimensionality $n = 5$.

It can also be noted that larger ϕ's correlate with higher run lengths. The explanation is that small standardized statistics are more likely to concentrate at the low frequency levels. A more fundamental reason can be traced to the skewed distribution towards the low frequency levels for the variance of DFT coefficients of a profile with positive autocorrelation, see Figure 5.6. Therefore, by the standardization in Equation (5.24), the statistics at low frequency levels are likely to be smaller than those at high frequency levels, and are unfavorable for profile comparison by using the AN test. Likewise, the skewed distribution towards the high frequency levels for the DFT coefficients of negative autocorrelated noise leads to small run lengths on the control charts. Therefore, we recommend cautious use of the HD control chart for monitoring profiles with highly autocorrelated noise.

Table 5.2: (a). Run Lengths of the HD Control Chart for the Nonlinear Profile Equation (5.19) with Both $\Delta\beta_1$ (0 and 0.025) and ϕ Shifts

$\Delta\beta_1$		0.000							0.025						
n\$\Delta\phi$		-0.8	-0.5	-0.3	0	0.3	0.5	0.8	-0.8	-0.5	-0.3	0	0.3	0.5	0.8
5	ARL	122.4	169.1	181.5	201.6	197.0	177.1	119.5	119.8	168.0	178.6	209.2	197.1	177.4	116.3
	25th	39.0	51.5	57.0	58.0	54.0	45.5	37.0	34.0	57.0	60.0	61.0	57.0	50.0	35.0
	50th	91.0	123.0	137.0	145.0	128.5	116.0	83.0	83.0	125.0	136.0	147.0	135.0	117.0	79.0
	75th	171.0	253.0	271.0	280.0	261.0	237.5	160.5	174.0	246.5	259.0	278.0	262.0	236.0	157.5
10	ARL	138.6	206.9	215.2	201.7	195.0	179.6	115.0	127.8	178.3	190.7	187.9	196.4	182.3	112.9
	25th	42.0	61.0	58.0	55.0	62.0	55.0	35.0	35.0	52.0	54.0	64.0	65.0	58.0	33.0
	50th	98.0	143.0	140.0	134.0	139.0	128.0	81.0	86.0	120.5	132.0	137.0	146.0	132.0	78.0
	75th	188.0	269.5	276.5	281.0	277.0	259.0	161.0	172.0	243.0	255.0	260.0	272.0	257.0	155.0
20	ARL	158.0	192.2	186.5	198.6	183.6	179.6	139.1	111.5	135.0	145.9	160.6	167.8	171.1	136.3
	25th	48.0	53.0	60.0	59.0	58.0	56.0	43.0	34.0	41.0	46.0	49.0	52.0	51.0	41.0
	50th	113.0	139.0	140.0	138.0	133.0	136.5	102.0	80.0	100.0	107.0	113.0	122.5	128.0	98.0
	75th	221.0	275.0	261.5	278.0	264.5	258.0	199.5	155.0	187.0	201.0	231.0	237.5	255.0	197.0
30	ARL	164.1	182.8	181.7	196.7	191.6	192.5	146.2	77.7	100.2	114.8	135.8	156.9	174.6	152.1
	25th	49.0	61.0	57.0	59.0	59.5	61.0	44.0	24.0	30.0	34.0	42.0	46.0	53.0	47.0
	50th	121.5	137.0	134.0	140.0	138.0	141.0	104.0	55.0	71.0	81.0	99.0	112.0	128.0	107.0
	75th	234.0	265.0	267.0	275.0	264.5	271.0	203.0	108.0	140.0	163.0	190.0	214.0	244.0	214.0
40	ARL	182.8	195.3	188.9	209.7	189.5	190.3	159.3	56.3	71.9	87.0	115.6	144.7	160.4	159.6
	25th	56.0	61.0	58.0	61.0	60.0	64.0	52.5	16.0	22.0	27.0	35.0	41.0	51.0	50.0
	50th	132.0	143.0	135.0	160.5	143.0	150.0	122.0	40.0	51.0	62.0	84.0	108.0	120.0	118.0
	75th	258.0	278.0	265.0	300.0	291.0	279.0	225.0	80.0	99.0	124.5	162.0	213.0	229.5	231.0
50	ARL	180.0	175.5	177.8	193.0	197.5	191.3	176.7	38.6	53.6	68.8	94.0	134.0	159.8	171.9
	25th	59.0	58.0	54.5	57.0	63.0	55.0	52.0	11.0	16.0	22.0	29.0	41.0	53.5	50.0
	50th	135.0	132.0	132.5	133.0	141.0	141.0	122.5	26.5	38.0	50.0	69.0	97.0	112.0	120.0
	75th	264.0	257.0	265.0	264.5	273.0	272.0	255.0	55.0	75.5	99.0	131.0	187.0	226.0	245.0

Table 5.2: (b). Run Lengths of the HD Control Chart for the Nonlinear Profile Equation (5.19) with Both $\Delta\beta_1$ (0.05 and 0.10) and ϕ Shifts

$\Delta\beta_1$		0.050							0.100						
n\ $\Delta\phi$		-0.8	-0.5	-0.3	0	0.3	0.5	0.8	-0.8	-0.5	-0.3	0	0.3	0.5	0.8
5	ARL	114.8	163.1	174.6	202.4	194.2	174.2	114.4	96.8	138.6	147.4	180.2	174.5	164.4	112.4
	25th	35.0	52.5	57.0	61.0	59.0	51.0	33.0	29.0	41.0	46.5	54.0	51.0	48.0	33.0
	50th	82.0	122.0	134.0	136.0	133.0	116.0	81.5	68.0	99.5	108.0	128.0	117.0	114.0	78.0
	75th	162.0	238.5	258.0	264.0	263.0	230.0	157.5	134.5	196.0	209.0	241.5	232.0	225.0	152.0
10	ARL	102.8	140.2	154.4	168.2	174.7	161.3	111.7	51.5	65.8	80.4	107.2	135.9	148.9	107.3
	25th	30.0	39.0	45.0	49.0	54.0	46.0	35.0	17.0	21.0	24.0	33.0	44.0	46.0	33.0
	50th	70.5	96.5	107.0	120.0	126.0	112.5	79.0	36.0	45.0	58.0	78.0	98.0	109.0	75.0
	75th	144.0	195.0	209.0	241.0	246.0	227.5	150.0	72.5	94.0	113.5	151.0	187.5	211.0	151.0
20	ARL	49.7	66.3	82.6	104.3	135.9	150.0	135.2	10.2	15.8	21.7	37.5	69.3	100.8	120.5
	25th	15.0	20.0	26.0	31.0	39.0	47.0	41.0	3.0	5.0	6.0	12.0	21.0	32.0	37.0
	50th	35.0	46.0	60.5	75.0	97.0	109.5	100.0	7.0	12.0	15.0	26.0	50.5	71.0	88.0
	75th	71.0	93.0	117.0	143.0	189.0	209.0	190.0	14.0	21.5	31.0	53.0	96.0	139.0	171.0
30	ARL	23.1	34.3	44.5	70.0	103.0	132.4	141.7	3.6	5.6	8.4	16.1	36.1	64.3	117.2
	25th	7.0	10.0	13.0	22.0	30.0	39.0	43.0	1.0	2.0	3.0	5.0	10.5	19.0	36.0
	50th	17.0	23.0	31.0	50.0	73.0	95.0	100.5	3.0	4.0	6.0	11.0	25.0	47.0	85.0
	75th	32.0	47.0	62.0	97.0	142.0	190.0	199.0	5.0	7.0	11.5	22.0	52.0	90.0	163.0
40	ARL	12.3	18.5	25.9	44.2	80.6	114.5	147.2	1.8	2.8	4.0	8.5	20.6	42.7	107.9
	25th	4.0	6.0	8.0	14.0	23.0	33.0	45.0	1.0	1.0	1.0	3.0	7.0	12.0	33.0
	50th	9.0	13.0	18.0	32.0	56.0	84.0	108.0	1.0	2.0	3.0	6.0	15.0	29.0	76.0
	75th	17.0	25.0	36.0	61.0	113.0	164.0	214.0	2.0	4.0	5.0	11.0	28.0	60.0	152.0
50	ARL	6.9	11.2	15.9	30.8	58.5	93.9	149.6	1.3	1.7	2.4	4.9	13.0	29.6	103.0
	25th	2.0	4.0	5.0	9.0	17.0	28.0	46.0	1.0	1.0	1.0	2.0	4.0	9.0	29.0
	50th	5.0	8.0	11.0	22.0	42.0	67.0	105.0	1.0	1.0	2.0	3.0	9.0	21.0	71.0
	75th	10.0	15.0	22.0	44.0	83.0	132.0	205.0	1.0	2.0	3.0	7.0	18.0	40.0	146.0

Example 2. *Linear profile with stationary AR(1) noise and slope shift*

In this example, the HD control chart is applied to monitor linear profiles modeled by

$$y_{jk} = \beta_0 + \beta_1 x_{jk} + \varepsilon_{jk}, \text{ for } j = 1,2,\ldots, N, \text{ and } k = 1,2,\ldots, n, \quad (5.26)$$

where β_0 and β_1 are intercept and slope respectively, and the random variables ε_{jk} are stationary and represented by a AR(1) model with autoregressive coefficient ϕ for each j. We use this example to investigate the robustness of the HD control chart in monitoring the shift in β_1 with stationary noise on the linear profile. The profile dimensionality n varies from 5 to 30 (most of the run lengths are very close to 1 for monitoring our-of-control profiles with higher dimensionalities), ϕ shifts from -0.8 to 0.8 with step size 0.3, and β_0 is fixed at 0. Two thousand runs with 1000 profiles each are simulated for each combination of n and $\Delta\beta_1$. The control limits are chosen from Table 5.1 for significance level $\alpha = 0.005$, or in-control ARL = 200.

The simulation results of ARLs and quartiles of run lengths are summarized in Table 5.3. Similar findings can be obtained from the table except that the shift $\Delta\beta_1$ has a much larger impact on the ARLs than that of the nonlinear profile monitoring in previous example, and the impact of n is larger too.

As a comparison, we list in the first row of Table 5.4 the simulation results by Kang and Albin (2000) in monitoring the slope shift of a linear profile with white noise. The dimensionality of the profiles considered in Kang and Albin (2000) is unclear. It can be seen that the ARLs of Kang and Albin (2000) are overall smaller than those of the HD

134

control chart for dimensionality $n = 5$, and overall larger than those of the HD control chart for $n = 10$ or more. We note that the approach by Kang and Albin (2000) relies on prior knowledge of the structure of the linear profiles or high-order polynomial profiles. Therefore, it cannot be applied to applications where such prior knowledge is unavailable. In comparison, the HD control chart approach can be still be used to monitor the profiles directly without such prior knowledge and still gives fairly good performance, especially for the linear profiles with high dimensions, $n \geq 10$.

Table 5.3: (a) Run Lengths of the HD Control Chart for Monitoring Linear Profile with Both $\Delta\beta_1$ (= 0 and 0.025) and ϕ Shifts

$\Delta\beta_1$		0.000							0.025						
n\ $\Delta\phi$		-0.8	-0.5	-0.3	0	0.3	0.5	0.8	-0.8	-0.5	-0.3	0	0.3	0.5	0.8
5	ARL	**123.0**	**164.4**	**181.8**	**188.7**	**177.5**	**162.1**	**113.9**	**104.2**	**145.5**	**162.9**	**173.5**	**175.9**	**156.3**	**111.6**
	25th	39.0	53.0	61.5	62.5	53.0	50.0	36.0	33.0	46.0	54.0	56.5	57.0	48.0	33.0
	50th	88.0	123.0	138.5	151.5	132.0	123.0	81.0	76.0	107.0	122.0	133.0	135.0	119.0	80.0
	75th	171.5	243.0	277.0	287.5	267.0	240.0	160.5	146.5	204.0	239.0	261.5	257.5	225.0	157.0
10	ARL	**134.2**	**174.8**	**185.8**	**184.7**	**183.5**	**173.1**	**115.4**	**57.8**	**79.1**	**89.8**	**118.7**	**145.3**	**144.1**	**109.7**
	25th	38.5	57.0	59.0	63.0	58.0	55.0	35.0	17.5	24.0	28.0	36.0	46.0	44.0	33.5
	50th	96.0	134.0	148.0	143.0	136.0	133.0	83.0	40.0	57.0	62.0	82.0	109.0	106.0	77.0
	75th	197.0	254.0	281.0	269.5	275.5	253.5	161.0	82.0	110.0	125.5	164.5	212.0	213.0	154.0
20	ARL	**148.7**	**183.9**	**184.9**	**181.9**	**186.8**	**187.3**	**132.5**	**5.5**	**8.3**	**12.2**	**23.1**	**45.5**	**78.1**	**105.2**
	25th	47.0	59.0	60.0	59.0	59.0	63.0	41.0	2.0	3.0	4.0	7.0	13.0	23.0	32.5
	50th	113.0	144.0	146.0	141.0	148.0	143.0	97.0	4.0	6.0	9.0	16.0	33.0	56.0	72.0
	75th	216.0	278.0	279.0	271.0	278.0	284.5	190.0	7.0	12.0	17.0	32.0	63.0	110.0	151.5
30	ARL	**165.7**	**190.5**	**183.3**	**181.6**	**178.1**	**178.7**	**156.5**	**1.3**	**1.7**	**2.3**	**4.5**	**12.5**	**27.2**	**86.1**
	25th	51.0	67.0	57.0	58.0	60.0	60.0	48.0	1.0	1.0	1.0	2.0	4.0	8.0	27.0
	50th	125.0	148.0	141.0	141.0	138.0	140.0	116.0	1.0	1.0	2.0	3.0	9.0	19.0	62.0
	75th	236.0	291.5	276.0	273.0	264.0	266.0	236.0	1.0	2.0	3.0	6.0	17.0	37.0	121.0

Table 5.3: (b) Run Lengths of the HD Control Chart for Monitoring Linear Profile with Both $\Delta\beta_1$ (= 0.05 and 0.10) and ϕ Shifts

$\Delta\beta_1$		0.050							0.100						
n\ $\Delta\phi$		-0.8	-0.5	-0.3	0	0.3	0.5	0.8	-0.8	-0.5	-0.3	0	0.3	0.5	0.8
5	ARL	**78.1**	**102.4**	**122.8**	**138.2**	**150.6**	**144.4**	**109.9**	**33.1**	**37.0**	**47.5**	**68.9**	**92.3**	**98.1**	**89.6**
	25th	22.0	32.0	40.0	43.0	45.0	44.0	34.0	10.0	11.0	14.0	20.0	29.0	31.0	29.0
	50th	55.0	75.0	87.0	104.0	111.0	105.0	79.0	24.0	26.0	34.0	50.0	67.0	68.0	64.0
	75th	108.0	142.0	174.0	195.5	215.5	208.5	154.0	46.0	53.0	65.0	94.5	130.5	136.5	126.0
10	ARL	**14.3**	**19.9**	**25.9**	**43.9**	**72.0**	**92.0**	**93.5**	**2.2**	**2.9**	**4.1**	**8.0**	**18.4**	**34.2**	**56.5**
	25th	4.0	6.0	8.0	13.0	23.0	27.0	30.0	1.0	1.0	2.0	3.0	6.0	11.0	17.0
	50th	10.0	14.0	18.0	32.0	51.0	66.0	68.0	2.0	2.0	3.0	6.0	13.0	25.0	39.0
	75th	20.0	28.0	37.0	62.0	101.0	126.0	130.0	3.0	4.0	5.0	11.0	25.0	49.0	77.5
20	ARL	**1.2**	**1.4**	**1.8**	**3.4**	**8.7**	**20.7**	**65.2**	**1.0**	**1.0**	**1.0**	**1.0**	**1.5**	**2.9**	**18.5**
	25th	1.0	1.0	1.0	1.0	3.0	7.0	20.0	1.0	1.0	1.0	1.0	1.0	1.0	6.0
	50th	1.0	1.0	1.0	2.0	6.0	15.0	48.0	1.0	1.0	1.0	1.0	1.0	2.0	13.0
	75th	1.0	2.0	2.0	4.0	12.0	28.0	91.0	1.0	1.0	1.0	1.0	2.0	4.0	25.0
30	ARL	**1.0**	**1.0**	**1.0**	**1.1**	**1.8**	**4.2**	**31.2**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.1**	**5.3**
	25th	1.0	1.0	1.0	1.0	1.0	2.0	9.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0
	50th	1.0	1.0	1.0	1.0	1.0	3.0	22.0	1.0	1.0	1.0	1.0	1.0	1.0	4.0
	75th	1.0	1.0	1.0	1.0	2.0	6.0	43.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0

Table 5.4: Comparison of ARLs for Control Charts for Linear Profiles with the Slope Shift $\Delta\beta_1$ and White Noise

Chart	$\Delta\beta_1$									
	0.025	0.05	0.075	0.1	0.125	0.15	0.175	0.2	0.225	0.25
EWMA/R	119	43.9	19.8	11.3	7.7	5.8	4.7	3.9	3.4	3
T^2	168	106.5	60.7	34.5	19.9	12.3	7.8	5.2	3.7	2.7
HD, $n = 5$	178.9	147.1	98.6	72.5	48.1	32.4	22.6	16.2	11.7	8.9
HD, $n = 10$	116.4	45.5	18.0	8.1	4.3	2.6	1.8	1.4	1.2	1.1

Example 3. *Linear profile with shifts in the standard deviation of the noise*

To study the impact of variance shift on the HD control chart performance, we consider the linear profile model in Equation (5.26) with white noise ε_{jk} i.i.d. and $\varepsilon_{jk} \sim N(0,1)$. The profile dimensionality n varies from 5 to 100 and the standard deviation of the noise σ varies from 1.2 to 3 while the values of β_0 and β_1 are fixed. Two thousand runs with 500 profiles in each run are simulated. The control limits are chosen from Table 5.1 for significance level $\alpha = 0.005$, or in-control ARL = 200.

The run lengths of the HD control chart are listed in Table 5.5. It can be observed that the ARLs decrease sharply as either σ or n increases. Until a certain level, for example, $\sigma = 2.4$ and $n = 60$, the ARL is close to 1 which indicates an immediate detection once σ shift occurs.

To aide comparison, we put the corresponding results from Kang and Albin (2000) in the 2nd and 3rd rows of Table 5.5. Although the profile dimensionality they

discussed is unclear, it can be seen that the ARLs of our control chart are smaller for any listed dimensionality. In Figure 5.8, the ARLs vs. dimensionality n are plotted for comparison. It visually illustrates that the HD control chart approaches have overall smaller ARLs.

Table 5.5:   Run Lengths of the HD Control Chart for Linear Profiles with Variance Shifts $\gamma\sigma$ of the White Noise

n\γ		1.0	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
EWMA/R		-	34.3	12	6.1	3.9	2.9	2.3	1.9	1.7	1.5	1.4
T^2		-	39.2	14.9	7.9	5.1	3.8	3	2.5	2.2	2	1.8
5	ARL	**178.6**	**33.9**	**11.1**	**5.0**	**3.2**	**2.2**	**1.8**	**1.5**	**1.4**	**1.3**	**1.2**
	25th	64	10	4	2	1	1	1	1	1	1	1
	50th	133.5	24	8	4	2	2	1	1	1	1	1
	75th	262	45	15	7	4	3	2	2	2	2	2
10	ARL	**189.6**	**28.1**	**8.0**	**3.3**	**2.0**	**1.6**	**1.3**	**1.2**	**1.1**	**1.1**	**1.0**
	25th	59	8	3	1	1	1	1	1	1	1	1
	50th	147.5	19	6	2	1	1	1	1	1	1	1
	75th	288.5	40	11	4	3	2	1	1	1	1	1
20	ARL	**181.5**	**22.6**	**4.7**	**2.0**	**1.4**	**1.1**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	25th	55	7	2	1	1	1	1	1	1	1	1
	50th	140	17	3	1	1	1	1	1	1	1	1
	75th	279	31	6	2	2	1	1	1	1	1	1
30	ARL	**192.0**	**17.8**	**3.6**	**1.5**	**1.1**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	25th	58.5	6	1	1	1	1	1	1	1	1	1
	50th	145.5	13	3	1	1	1	1	1	1	1	1
	75th	301.5	25	5	2	1	1	1	1	1	1	1
40	ARL	**186.1**	**17.4**	**2.7**	**1.3**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	25th	58.5	5	1	1	1	1	1	1	1	1	1
	50th	142	12	2	1	1	1	1	1	1	1	1
	75th	281.5	24	3	1	1	1	1	1	1	1	1
50	ARL	**183.5**	**14.7**	**2.2**	**1.1**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	25th	58	5	1	1	1	1	1	1	1	1	1
	50th	145	10	2	1	1	1	1	1	1	1	1
	75th	274.5	21	3	1	1	1	1	1	1	1	1
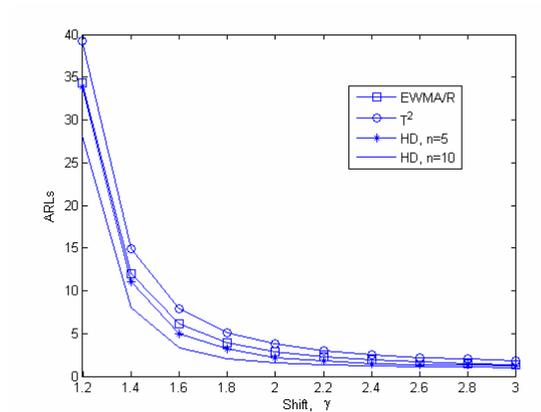
Figure 5.8: Comparisons of ARLs for control charts with standard deviation shifts.

5.6.2   HD Control Chart and Supplemental Chart Application

In this section, we apply our HD control chart to monitor the process of simulated nonlinear woodboard vertical density profiles (VDP). The measurements of VDP are normally collected from a particleboard manufacturing operation. The density is an important quality characteristic for manufacturing the engineered woodboard because it determines its machinability. The density is measured by using a profilometer, which is a laboratory measuring instrument that measures a feature's length or depth, usually in the micrometer or nanometer level. In this example, it works by taking measurements at fixed depth across the thickness of the board. The VDP of the board is therefore formed by the measurements on a sample (usually a 2×2 inch piece). Each VDP consists of 314 measurements taken 0.002 inches apart.

Williams et al. (2003) fitted the following nonlinear "bathtub" function for a typical woodboard profile

$$f(x_i, \boldsymbol{\beta}) = \begin{cases} a_1 (x_i - d)^{b_1} + c & x_i > d \\ a_2 (-x_i + d)^{b_2} + c & x_i \le d \end{cases}, \tag{5.26}$$

where x_i is the ith regressor variable value, and parameters a_1 and a_2 determine the width of the "bathtub", b_1 and b_2 determine the flatness, c and d determine the bottom and the center respectively. The values of the parameters for one representative in-control profile function are listed in Table 5.6.

Table 5.6:   Parameters of Estimated Nonlinear Profile of One Representative Profile

a_1	5708	a_2	3921
b_1	5.14	b_2	4.87
c	46.0	d	0.313

In this example, two profile data sets are simulated by using the parameter values in Table 5.6. There are 100 in-control and 50 out-of-control profiles in each dataset. For the first data set, we assume c has a normal distribution of $N(46.0, 0.05 \bullet \sqrt{46})$ and other parameters have constant values as in Table 5.6. Three standard deviation shifts are introduced to the mean of parameter c in Phase II. For the second, we assume parameters b_1 and b_2 have normal distributions of $N(5.14, 0.05 \bullet \sqrt{5})$ and $N(4.87, 0.05 \bullet \sqrt{5})$ and keep other parameters' values constant at in Table 5.6. Again, three standard deviation shifts are introduced to the mean of parameters b_1 and b_2 in Phase II. In addition, a AR(1) stationary noise with $\phi = 0.8$ is added to each simulated profile for both Phase I and II, which is in line with the stationary noise of the practical woodboard profile data.

For illustration, five simulated profiles are plotted in Figure 5.9. The two profiles in red are outliers with mean shifts on the parameters b and c respectively.

Using the procedure illustrated in Figure 5.7, two high-dimensional control charts are constructed in Figure 5.10(a) and (c) for the two simulated data sets. The control limits for both charts are 7.72 which approximates that for the dimensionality $n = 200$ in Table 5.1. For both cases, no out-of-control profile is detected and removed in Phase I, and the out-of-control profiles are detected in 11 observations for the first dataset and in 3 observations for the second dataset. The supplementary control charts are applied to each of the simulated datasets in Figure 5.10(b) and (c) and no outlier is found in the two charts.
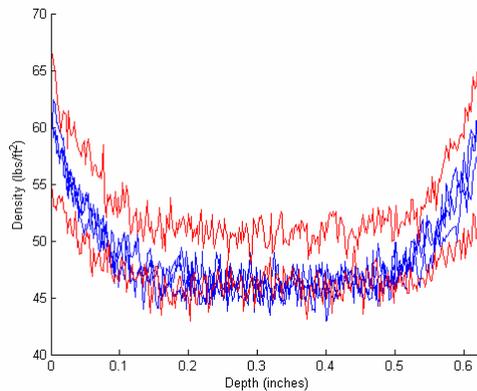
Figure 5.9: Five simulated vertical density profiles (the two profiles in red are outliers)
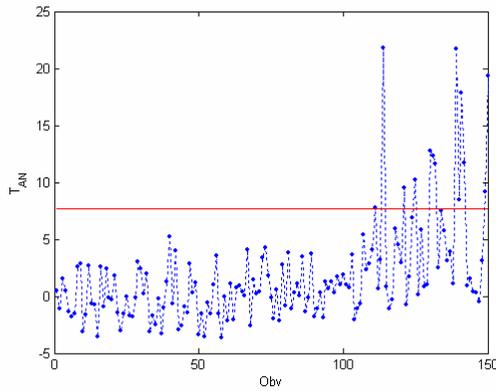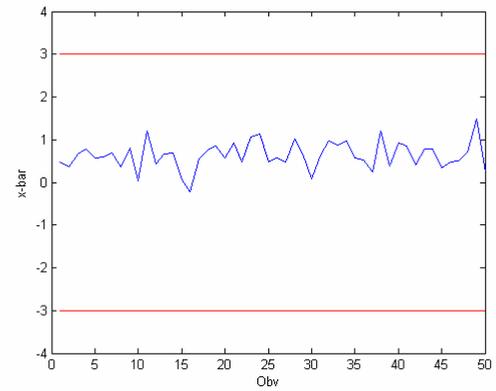
(a)   (b)
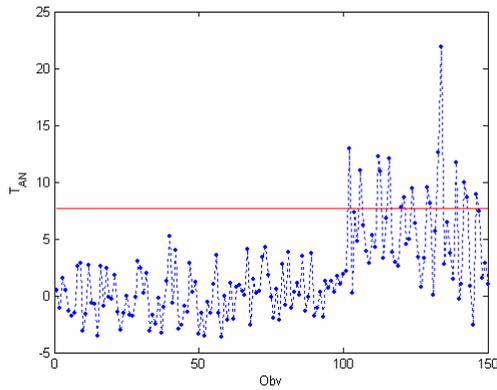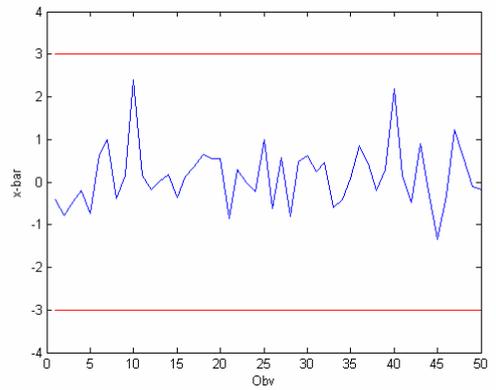
(c)   (d)

Figure 5.10: (a, b) The HD control chart and the supplementary $\overline{x}$ chart for the first simulated profile dataset; (c, d) The HD control chart and the supplementary $\overline{x}$ chart for the second simulated profile dataset.

5.7   Conclusions

In this chapter, a high-dimensional (HD) control chart approach is proposed to monitor processes or products whose quality can be characterized by profiles or functions between a response variable and one or more independent variables. The profile dimensionality, or the number of paired values for the response variable and the

independent variables, is normally large which makes it impractical to apply a traditional Hotelling's T^2 control chart. This HD control chart approach relies on the discrete Fourier transform (DFT) to decorrelate the profile noise, either stationary or i.i.d., and to compress the profile signal to low frequency levels. The adaptive Neyman (AN) test is then used to automatically select the number of coefficients at low frequencies to maximize the AN statistics.

A construction procedure for the HD control chart based on the combination of DFT and AN test is presented, and its performance in monitoring both linear and nonlinear profiles with either i.i.d. or autocorrelated stationary noise is evaluated by simulation. A comparison with other approaches in monitoring linear profiles is provided, and the advantages of the HD chart approach are summarized as follows:

First, it can be directly used to monitor profiles without prior knowledge of their structures when enough historical profile data can be obtained to estimate the profile mean and variance function. Based on the data-driven techniques of the AN test, the HD chart approach can effectively test whether a profile is significantly different from the estimated process mean profile or not, and it has good adaptability in monitoring either linear or nonlinear profiles.

Second, it can be used to monitoring profiles with a stationary noise component. Our results showed that the impact of the noise autocorrelation can be neglected in this approach if the autoregressive coefficient ϕ is in the range of [-0.5, 0.5] for both linear and nonlinear profiles.

Graphical control charts for the HD approach were used to monitor simulated vertical density profiles in a woodboard manufacturing process.

Appendix 5A: Matlab Simulation Code for the HD Control Chart for Monitoring Nonlinear Profiles

```
% Out-of-control ARL of the high-dimensional control chart for monitoring
% nonlinear profiles, yt = b0 + b1*log(xt) + et, where et = phi*e_t-1 + at,
% by varying values of phi and b1.
%
clear;

start = clock;

dim = [5 10:10:50];    % dimensionality
Jn = [5.97 6.7700 7.1600 7.2900 7.4100 7.4300];
dphi = [-0.8 -0.5 -0.3 0 0.3 0.5 0.8]; % size of mean shift
dslope = [0 0.025 0.05 0.1];    % size of mean shift

run = 2000;
n = 1000;

for j = 1:1:length(dim)
   for jj = 1:1:length(dphi)
      for k = 1:1:length(dslope)
         if k == 1 &jj==4
            n = 2000;
         else
            n = 500;
         end
         for r = 1:1:run
            X0 = randn(dim(j),n);
            e = ARMA11(dim(j),n,dphi(jj),0);
            x = 1:1:dim(j);
            X = dslope(k)*log(x'); X = repmat(X,1,n);
            X = X + e;

            mdim = dim(j);

            % F-transform and standardization
            X = fft_coeff(X);
            X0 = fft_coeff(X0);
            Xavg = mean(X0')';
            Xvar = var(X0')';

            for i = 1:1:n
               X(:,i) = (X(:,i)-Xavg)./sqrt(Xvar);
            end
            %********************************%

            for i = 1:1:n
               for m = 1:1:mdim
                  T1(m,i) = sum(X(1:m,i).^2-1)/sqrt(2*m);
               end;
            end;
```

```matlab
        ANA1 = max(T1);
        TAN1 = sqrt(2*log(log(mdim)))*ANA1-(2*log(log(mdim)) + 0.5*log(log(log(mdim))) -
0.5*log(4*pi));

        % Evaluate Test Results
        rls = find(TAN1>Jn(j));
        if (length(rls) == 0)
           rl(r,j,jj,k) = n;
        else
           rl(r,j,jj,k) = rls(1);
        end
        %**********************%
      end
    end
  end
end

% generate the table of the results

% get the quartile of the run length
arl = mean(rl);  % arl is in the first row
p = 100*(0.25:0.25:0.75)'; % three quartiles
qrl = [arl;prctile(rl,p)]  % 25,50,75 percentile in the next rows

qrltable = zeros(4*length(dim),length(dphi),length(dslope));
for k = 1:1:length(dslope)
   for j = 1:1:length(dphi)
     for i = 1:1:length(dim)
        srow = (i-1)*4+1;
        qrltable(srow:srow+3,j,k)=qrl(:,i,j,k);
     end
   end
end

qrltable

endt = clock;
elapsetime = etime(endt,start)

save simu_nonlinear.mat qrl qrltable elapsetime;


function X = fft_coeff(ox)
% Generate the coefficients for fast Fourier transform used in Appendix 5A
% ox - a dimm-by-n matrix
% Row number is the order of dimensionality!
%
X = ox;
[m,n] = size(X);
dim = m;
xr = real(fft(X)); xi = imag(fft(X));

if (mod(dim,2) == 0)  % even
   X(1,:) = xr(1,:);
```

```
    for i = 2:1:dim/2
        X(i*2-2,:) = xr(i,:);
        X(i*2-1,:) = xi(i,:);
    end
    X(dim,:) = xr(dim/2+1,:);
else %odd
    X(1,:) = xr(1,:);
    for i = 2:1:(dim+1)/2
        X(i*2-2,:) = xr(i,:);
        X(i*2-1,:) = xi(i,:);
    end
end


function z = ARMA11(dim,n,phi,theta)
%% simulate n sets of ARMA(1,1) time series with length dim for each set,
%% used in Appendix 5A & 5B
%
z = zeros(dim,n);
for i = 1:1:n
    a = randn(dim,1);
    z(1,i) = a(1);
    for t = 2:1:dim
        z(t,i) = phi * z(t-1,i) + a(t) - theta * a(t-1);
    end
end
```

Appendix 5B: Matlab Simulation Code for the HD Control Chart for Monitoring Linear Profiles with Slope and Autoregressive Coefficient Changes

```
% Out-of-control ARL of the high-dimensional control chart for monitoring
% linear profiles, yt = b0 + b1*x_t + e_t, where e_t = phi*e_(t-1) + a_t,
% by varying values of phi and b1.

clear;

start = clock;

dim = 10:10:100;    % dimensionality
Jn = [6.7700 7.1600 7.2900 7.4100 7.4300 7.5100 7.5500 7.5700 7.6500 7.6500];
dphi = [-0.6 -0.3 0 0.3 0.6]; % size of mean shift
dslope = [0.02 0.04 0.08 0.1 0.2]; % size of mean shift

n = 100000;

for j = 1:1:length(dim)
    for jj = 1:1:length(dphi)
        for k = 1:1:length(dslope)
            X0 = randn(dim(j),n);
            e = ARMA11(dim(j),n,dphi(jj),0);
```

```matlab
        X = dslope(k)*(1:1:dim(j))'; X = repmat(X,1,n);  % profile signals
        X = X + e;

        mdim = dim(j);

        %**************** F-transform and standardization  *****************%
        X = real(fft(X));
        X0 = real(fft(X0));
        Xavg = mean(X0')';
        Xvar = var(X0')';

        for i = 1:1:n
           X(:,i) = (X(:,i)-Xavg)./sqrt(Xvar);
        end
        %**********************************%

        for i = 1:1:n
           for m = 1:1:mdim
              T1(m,i) = sum(X(1:m,i).^2-1)/sqrt(2*m);
           end;
        end;
        ANA1 = max(T1);
        TAN1 = sqrt(2*log(log(mdim)))*ANA1-(2*log(log(mdim)) + 0.5*log(log(log(mdim))) -
0.5*log(4*pi));

        %**************** Evaluate Test Results  *****************%
        count = 0;
        for i = 1:1:length(TAN1)
           if (TAN1(i) > Jn(j))
              count = count + 1;
           end
        end
        rl(j,jj,k) = count/n;
      end
   end
   clock
end
rl

endt = clock;
elapsetime = etime(endt,start)
```

Appendix 5C: Matlab Simulation Code for the HD Control Chart for Monitoring Linear Profiles with Variance Change

```matlab
% Out-of-control ARL of the high-dimensional control chart for monitoring
% linear profiles, yt = b0 + b1*x_t + e_t, where e_t = phi*e_(t-1) + a_t,
% by varying the variance of the errors
clear;
```

```
start = clock;

dim = 10:10:100;    % dimensionality
Jn = [6.7700 7.1600 7.2900 7.4100 7.4300 7.5100 7.5500 7.5700 7.6500 7.6500];
dvar = 1:0.2:3; % multiple of var shift
n = 100000;

for j = 1:1:length(dim)
   for k = 1:1:length(dvar)
      e = randn(dim(j),n);
      X = sqrt(dvar(k))*e;
      mdim = dim(j);

      %**************** F-transform and standardization  ****************%
      X = fft_coeff(X);
      e = fft_coeff(e);
      Xavg = mean(e')';
      Xvar = var(e')';
      for i = 1:1:n
         X(:,i) = (X(:,i)-Xavg)./sqrt(Xvar);
      end
      %****************

      for i = 1:1:n
         for m = 1:1:mdim
            T1(m,i) = sum(X(1:m,i).^2-1)/sqrt(2*m);
         end;
      end;
      ANA1 = max(T1);
      TAN1 = sqrt(2*log(log(mdim)))*ANA1-(2*log(log(mdim)) + 0.5*log(log(log(mdim))) -
0.5*log(4*pi));

      %**************** Evaluate Test Results  ****************%
      count = 0;
      for i = 1:1:length(TAN1)
         if (TAN1(i) > Jn(j))
            count = count + 1;
         end
      end
      rl(j,k) = count/n;
   end
   clock
end
rl = [dvar; rl];
rl

endt = clock;
duration = etime(endt,start)
```

Appendix 5D: Matlab Simulation Code for Monitoring the Woodboard Density Profiles

```
clear;

scale = 0.0025;
x1 = (1:1:156)'; x2 = (157:1:313)'; x = [x1; x2]';
x1 = x1*0.002; x2 = x2*0.002; x = x*0.002;
ma1 = 3921; mb1 = 4.87; mc1 = 46; ma2 = 5708; mb2 = 5.14; mc2 = 46;
va = 5000*0; vb = 5*scale; vc = 46*scale;
phi = 0.8;

%********** Phase I ***********%
r = 100;
y = zeros(313,r);
for i = 1:1:r
   e = randn(313,1);
   e = arma11(313,1,phi,0);

   a1 = randn(1)*sqrt(va)+ma1; a2 = randn(1)*sqrt(va)+ma2;
   b1 = randn(1)*sqrt(vb)+mb1; b2 = randn(1)*sqrt(vb)+mb2;
   c1 = randn(1)*sqrt(vc)+mc1; c2 = c1;

   y1 = a1*(-x1+0.313).^b1+c1;
   y2 = a2*(x2-0.313).^b2+c2;

   y(:,i) = [y1; y2]+e;
end
X = y(:,1:r);

%***** case I *****
% Phase II
r = 50;
y = zeros(313,1,r);
for i = 1:1:r
   e = randn(313,1);
   e = arma11(313,1,phi,0);

   a1 = randn(1)*sqrt(va)+ma1; a2 = randn(1)*sqrt(va)+ma2;
   b1 = randn(1)*sqrt(vb)+mb1; b2 = randn(1)*sqrt(vb)+mb2;
   c1 = randn(1)*sqrt(vc)+mc1+3*sqrt(vc); c2 = c1;

   y1 = a1*(-x1+0.313).^b1+c1;
   y2 = a2*(x2-0.313).^b2+c2;

   y(:,1,i) = [y1; y2]+e;
end

fft_ratio = 1;
mdim_ratio = 1;
aneyman_complex(X,y,fft_ratio,mdim_ratio);
```

```matlab
%***** case II *****
% Phase II
r = 50;
y = zeros(313,1,r);
for i = 1:1:r
   e = randn(313,1);
   e = arma11(313,1,phi,0);

   a1 = randn(1)*sqrt(va)+ma1; a2 = randn(1)*sqrt(va)+ma2;
   b1 = randn(1)*sqrt(vb)+mb1+3*sqrt(vb); b2 = randn(1)*sqrt(vb)+mb2+3*sqrt(vb);
   c1 = randn(1)*sqrt(vc)+mc1; c2 = c1;

   y1 = a1*(-x1+0.313).^b1+c1;
   y2 = a2*(x2-0.313).^b2+c2;

   y(:,1,i) = [y1; y2]+e;
end

fft_ratio = 1;
mdim_ratio = 1;
aneyman_complex(X,y,fft_ratio,mdim_ratio);


function aneyman_complex(X,Y,fft_ratio,mdim_ratio)
% Adaptive neyman test for vectors X and Y, used in Appendix 5B
%
% X is the input matrix for the first group:  T x n_1
% Y is the input matrix for the secod group:  T x n_2 x K

M = []; MM = [];
[mx nx] = size(X);
[my ny ky] = size(Y);
fftn = floor(fft_ratio*mx); mdim = floor(mdim_ratio*mx); n = mdim;
if (mx > my | mx < my)
   error('The dimension of X and Y should be equal!');
end;

% use both the real and imaginary parts for the fft
dim = mx;
X = fft_coeff(X);
Xavg = mean(X')'; Xvar = var(X')';
for i = 1:1:ky
   YY = Y(:,:,i);
   if (ny > 1)
      Yavg(:,i) = mean(fft_coeff(YY)')';
      Yvar(:,i) = var(fft_coeff(YY)')';
   else
      Yavg(:,i) = fft_coeff(YY);
   end
end

%% for constant estimator of variance
varx = mean(Xvar);
```

```
%***************  Phase I  *****************%
pos = zeros(nx,1);
for i = 1:1:nx
    Z(:,i) = (Xavg - X(:,i))./sqrt(Xvar/nx+Xvar);
    for m = 1:1:mdim
        T1(m,i) = sum(Z(1:m,i).^2-1)/sqrt(2*m);
    end;
    a = T1(:,i);
    pos(i) = find(a == max(a));
end;
ANA1 = max(T1);
TAN1 = sqrt(2*log(log(n)))*ANA1-(2*log(log(n)) + 0.5*log(log(log(n))) - 0.5*log(4*pi));

%****************  Phase II  *****************%
for i = 1:1:ky
    Z(:,i) = (Xavg - Yavg(:,i))./sqrt(Xvar/nx+Xvar/ny);
    for m = 1:1:mdim
        TT1(m,i) = sum(Z(1:m,i).^2-1)/sqrt(2*m);
    end;
    a = TT1(:,i);
    pos(i) = find(a == max(a));
end;
ANAA1 = max(TT1);
TANN1 = sqrt(2*log(log(n)))*ANAA1-(2*log(log(n)) + 0.5*log(log(log(n))) - 0.5*log(4*pi));

%***************  Combine Phase I and II in one chart  *****************%
alpha = 0.005;
dim = 10:10:100;    % dimensionality
Jn = [6.7700 7.1600 7.2900 7.4100 7.4300 7.5100 7.5500 7.5700 7.6500 7.6500];
if mx <=100
    UL = Jn(find(dim==mx));
else
    UL = 7.65;
end

NN = length(TAN1)+length(TANN1);
t = 1:1:NN;
TAN = [TAN1, TANN1];
Tp = ones(1,NN)*UL;
T2 = ones(1,NN)*chi2inv(1-alpha,mx);
figure; plot(t,TAN,':.', t,Tp,'-r');
title('Control Chart of T_A_N for Phase I&II'); xlabel('Obv'); ylabel('T_A_N');
```

Summary and Future Research

This thesis addresses the development of a unified Cuscore control chart to monitor the mean shift in univariate or multivariate autocorrelated process, and a high-dimensional control chart for monitoring the mean function of processes that can be represented by profile data. The work was broadly motivated by the need to gain better control over manufacturing processes. The contributions of this research are presented in the next section. Directions for the advancement of this research are presented in the subsequent section.

6.1 Research Contributions

6.1.1 Cuscore Control Charts for Generalized Feedback Control Systems

The first part of this research centered around monitoring a feedback controlled process. In Chapter 3, the Cuscore control chart was used to monitor the output of a GMV feedback control system for the presence of a signal. Appropriate statistics based on the fault signatures of the signal were derived for the detection of signals in an ARMA noise process. We showed theoretically that the performance of Cuscore charts is independent of the amount of variability transferred from the output quality characteristic to the adjustment actions in the GMV control system. Simulation was used to explore the performance of the Cuscore charts for monitoring an ARMA(1,1) noise in detecting a

154

spike, step and bump signal in a GMV control system. In general, the Cuscore chart has the ability to detect signals over a broad range of system parameter values. For certain ranges of system parameters where the Cuscore chart displays low detection capability for the fault signatures, a tracking signal test was used in combination with the Cuscore statistics to achieve satisfactory detection performance.

6.1.2 Multivariate Cuscore Control Charts for Monitoring Autocorrelated Processes

The Cuscore control chart is a powerful statistical process monitoring tool when there is prior knowledge about the process shift. In Chapters 3 and 4 it was also illustrated to be effective for monitoring an autocorrelated process when the process autocorrelation has been estimated in Phase I. In Chapter 4, the multivariate Cuscore approach based on the likelihood ratio test and fault signature analysis was introduced for monitoring the mean vector shift in an autocorrelated multivariable process. A bivariate time series model was used to illustrate the theory and application of the MCuscore chart. Simulation was used to show that the MCuscore chart outperforms the traditional residual-based MCusum control chart in detecting a mean vector shift signal in autocorrelated bivariate processes. An example of monitoring the mean shift of two process variables in an RIE process illustrated the use of the MCuscore chart and showed it to perform better than the MCusum chart in monitoring and autocorrelated multivariate process when *a priori* information on the process and the signal is available. In addition, the integration of fault diagnosis with MCuscore control chart was briefly discussed.

155

6.1.3   A High-Dimensional Control Chart for Profile Monitoring

In Chapter 5, a high-dimensional control chart approach was presented to monitor the process or product whose quality can be characterized by profiles. The dimensionality of the investigated profiles, or the number of paired values for the response variable and the independent variables, is normally large. This high-dimensional control chart approach relies on the discrete Fourier transform to decorrelate the profile noise and to compress the profile signal into low frequency levels. The adaptive Neyman test is then used to automatically select the number of large coefficients at the low frequency levels by maximizing the AN statistics. A construction procedure for the high-dimensional control chart based on the combination of DFT and AN test was presented, and its performance was evaluated by simulation for monitoring both linear and nonlinear profiles with either i.i.d. or autocorrelated stationary noise.

Simulation was also used to compare the HD control chart with other approaches in monitoring profile data. It showed that the high-dimensional chart has two main advantages. First, it can be directly used to monitor profiles without prior knowledge of their structures when enough historical profile data can be obtained to estimate the profile mean and variance function. Second, it can be used to monitor profiles with a stationary noise component. Our results showed that the impact of the noise autocorrelation can be neglected in this approach if the autoregressive coefficient ϕ is in the range of [-0.5, 0.5] for both the linear and nonlinear profiles.

Graphical control charts for the HD approach were used to monitor data profiles that were representative of an actual woodboard manufacturing process.

6.2   Future Work

We conclude with a brief discussion of some potential future research related to the Cuscore chart and profile monitoring approaches investigated in this thesis.

- Robustness of the Cuscore chart for unknown signal information

Although the Cuscore control chart has been illustrated to outperform many traditional control charts and it is often used as a powerful supplementary tool for detecting specific signals, some practical issues hinder the extensive use of Cuscore charts. One of them is lack of ready-to-use software for integrating the Cuscore chart with other traditional charts, and the other, which is more crucial, is the constraint of the fundamental assumption of the Cuscore statistics, which requires prior knowledge of the starting time and shift size of the signals being monitored.

To address the second issue, Nembhard and Changpetch (2006) investigated the robustness of the Cuscore chart for mismatched signals and applied the idea of fixed-sized moving-window or adaptive-sized detection window to the Cuscore chart. Their research can be extended to multivariate Cuscore chart. In addition, some other researchers have taken the approach of designing a "trigger" for the Cuscore chart so that it can react to the expected signal automatically, such as the Cusum-trigger Cuscore chart by Shu et al. (2002). The idea of using the change-point model (Hawkins et al., 2003; Hawkins and Zamba, 2005a, b) to trigger Cuscore is potentially a promising direction. It is expected that more research will be conducted with the aim of relaxating some of the assumptions of Cuscore statistics.

- Estimation of the covariance matrix

Estimation of covariance matrix and monitoring its change is an important issue for multivariate process control. The assumption of a fixed covariance matrix, as is used in the research of Chapter 4, may easily lead to a higher false alarm rate in the control chart when a relatively large mean shift signal occurs and causes a change in the covariance matrix. Many researchers have investigated this issue, such as Sullivan and Woodall (1996) and William et al. (2006), and it may be possible to extend the multivariate Cuscore chart to monitoring the changes in the covariance matrix, or to find a good estimate of the process covariance matrix using the multivariate Cuscore chart to monitor the process mean vector.

- Multivariate control charts for profile monitoring

Currently most of research on profile monitoring relies on fitting a parametric regression model, either linear or nonlinear, to the profile data, and then monitoring the parameter vector using multivariate process control approaches; see Kang and Albin (2000), Kim et al. (2003), Mahmoud and Woodall (2004) and William et al. (2006). Theoretically, if prior knowledge has been acquired about the size and time of a shift in the parameter vector from the historical data, a MCuscore chart can be designed for efficient profile monitoring by focusing on the multivariate time series of the parameters.

Profile monitoring is a rich research area and many other developed or on-going research topics can be related to it, such as the change-point approach, the non-parametric regression for curve fitting, and the monitoring of autocorrelated process whose quality characteristics are profiles. An advanced integration of the MCuscore chart with monitoring autocorrelated profiles may be a promising research direction.

References

Alwan, L. and Roberts, H. V. (1988). "Time-Series Modeling for Statistical Process Control," *Journal of Business & Economic Statistics*, 6, 87-95.

Apley, D. W. and Shi, J. (1999). "The GLRT for Statistical Process Control of Autocorrelated Processes," *IIE Transactions*, 31, 1123-1134.

Aradhye, H. B., Baskshi, B. R., Strauss, R. A. and Davis, J. F. (2003). "Multiscale SPC Using Wavelets: Theoretical Analysis and Properties," *AIChE Journal*, 49(4), 939-958.

Bakshi, B. R. (1998). ''Multiscale PCA with Application to Multivariate Statistical Process Monitoring,'' *AIChE Journal*, 44, 1596.

Boggess, A. and Narcowich, F. J. (2002). *A First Course in Wavelets with Fourier Analysis*. Prentice Hall, NJ.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*, 3rd ed. Prentice Hall, Englewood Cliffs, NJ.

Box, G. E. P. and Kramer, T. (1992). "Statistical Process Monitoring and Feedback Adjustment - A Discussion," *Technometrics*, 34, 251-285.

Box, G. E. P. and Luceño, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. Wiley, New York, NY.

Box, G. E. P. and Ramírez, J. (1992). "Cumulative Score Charts," *Quality and Reliability Engineering International* 8, 17-27.

Capilla, C., Ferrer, A., Romero, R., and Hualda, A. (1999). "Integration of Statistical and Engineering Process Control in a Continuous Polymerization Process," *Technometrics*, 41, 14-28.

Clark, D. W., and Gawthrop, P. J. (1975). "Self-Tuning Controller," *Proceedings of the Institution of Electrical Engineers*, 122, 9, 929-934.

Crosier (1988). "Multivariate Generalizations of Cumulative Sum Quality-control Schemes," *Technometrics*, 30, 3, 291-303.

Del Castillo, E. (1996). "Evaluation of Run Length Distribution for Charts with Unknown Process Variance," *Journal of Quality Technology*, 28, 116–122.

Del Castillo, E. (2002). *Statistical Process Adjustment for Quality Control*. Wiley, New York, NY.

Ding, Y., Zeng, L., and Zhou, S. (2006). "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes," *Journal of Quality Technology*, 38, 3, 199-216.

Duncan, A. J. (1956). "The Economic Design of $\bar{X}$ Charts Used to Maintain Current Control of a Process," *Journal of the American Statistical Association*, 51, 228-242.

Fan, J. (1996). "Test of Significance Based on Wavelet Thresholding and Neyman's Truncation," *Journal of American Statistical Association*, 91, 434, 674-688.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*. Chapman & Hall, New York, NY.

Fan, J. and Huang, L. S. (2001). "Goodness-of-Fit Tests for Parametric Regression Models," *Journal of the American Statistical Association*, 96, 454, 640-652.

Fan, J. and Li, R. (2006). "Statistical Challenges with High Dimensionality: Feature Selection in Knowledge Discovery," Proceedings of International Congress of Mathematicians (ICM) (Sanz-Solé, M., Soria, J., Varona, J. L., and Verdera, J. eds.), III, 595-622.

Fan, J. and Lin, S. K. (1998). "Test of Significance When Data Are Curves," *Journal of the American Statistical Association*, 93, 443, 1007-1021.

Fan, J., Zhang, C. and Zhang, J. (2001). "Generalized Likelihood Ratio Statistics and Wilks Phenomenon," *The Annals of Statistics*, 29, 1, 153-193.

Fisher, R. A. (1925). "Theory of Statistical Estimation," *Proceedings of the Cambridge Philosophical Society*, 22, 700-725.

Ganesan, R., Das, T. K. and Venkataraman V., (2004). "Wavelet-Based Multiscale Statistical Process Monitoring: A Literature Review," *IIE Transactions*, 36, 787-806.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

Harris, T. J., and Ross, W. H. (1991). ''Statistical Process Control Procedures for Correlated Observations,'' *Canadian Journal Chemical Engineering*, 69, 48.

Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer, New York, NY.

Hawkins, D. M. and Olwell D. H., (1998). *Cumulative Sum Charts and Charting for Quality Improvement,* Springer, New York, NY.

Hawkins, D. M., Qiu, P., and Kang, C. W. (2003). "The Changepoint Model for Statistical Process Control," *Journal of Quality Technology*, 35, 4, 355-366.

Hawkins, D. M., Zamba, K. D. (2005). "A Change Point Model for a Shift in Variance," *Journal of Quality Technology*, 37, 1, 21-31.

Hawkins, D. M., Zamba, K. D. (2005). "Statistical Process Control for Shifts in Mean or Variance Using a Changepoint Formulation," *Technometrics*, 47, 2, 164-173.

Healy, J. D. (1987). "A Note on Multivariate CUSUM Procedures," *Technometrics*, 29, 4, 409-412.

Hotelling, H. (1947). "Multivaraite Quality Control – Illustrated by the Air Testing of Sample Bomb Sights," *Techniques of Statistical Analysis* Ed. C. Eisenhart, M. W. Hastay and W. A. Wallis., McGraw Hill, New York.

Hu, S. J. and Roan, C. (1996). "Change Patterns of Time Series-Based Control Charts," *Journal of Quality Technology*, 28, 3, 302-312.

Hyvarinen, A., Harhunen, J., and Oja, E. (2001). *Independent Component Analysis.* Wiley, New York, NY.

Jackson, J. E. (1985). "Multivariate Quality Control," *Communications in Statistics - Theory and Methods*, 14, 2657-2688.

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

Jensen, W. A. and Birch, J. B. (2006). "Profile Monitoring via Nonlinear Mixed Models", Technical Report No. 06-4. (http://www.stat.org.vt.edu/dept/web-e/tech_reports/TechReport06-4.pdf), Virginia Tech.

Jeong, M. K. (2004). *Wavelet-Based Methodology in Data Mining for Functional Data*, Ph.D. Dissertation, Georgia Institute of Technology.

Jeong, M. K., Chen, D. and Lu, J. C., (2003). "Thresholded Scalogram and Its Applications in Process Fault Detection," *Applied Stochastic Models In Business and Industry*, 19, 231-244.

Jeong, M. K., Lu, J. C. and Wang, N. (2006). "Wavelet-Based SPC Procedure for Complicated Functional Data," *International Journal of Production Research*, 44(4), 729-744.

Jiang, W. and Tsui, K. L. (2002). "SPC Monitoring of MMSE- and PI-Controlled Processes," *Journal of Quality Technology*, 34, 4, 384-398.

Jin, J. and Shi, J. (1999). "Feature-Preserving Data Compression of Stamping Tonnage Information Using Wavelets," *Technometrics*, 41, 4, 327-339.

Jin, J. and Shi, J. (2001). "Automatic Feature Extraction of Waveform Signals for In-Process Diagnostic Performance Improvement," *Journal of Intelligent Manufacturing*, 12, 257-268.

Johnson, R. A. and Wichern, D. W. (2002). *Applied Multivariate Statistical Analysis, 5ed.* Person Education, New York, NY.

Jones, L. A., Champ, C. W., and Rigdon, S. E. (2004). "The Run Length Distribution of the CUSUM with Estimated Parameters". *Journal of Quality Technology*, 36, 1, 95-108.

Jones, L. A., Champ, C. W., and Rigdon, S. E. (2001). "The Performance of Exponentially Weighted Moving Average Charts with Estimated Parameters". *Technometrics*, 43, 156–167.

Kamarthi, S. V., Kumara, S. R. T., and Cohen, P. H. (2000). "Flank Wear Estimation in Turning Through Wavelet Representation of Acoustic Emission Signals," *Transactions of the ASME*, 122, Feb., 12-19.

Kang, L. and Albin, S. L. (2000). "On-Line Monitoring When the Process Yields a Linear Profile," *Journal of Quality Technology*, 32, 418-426.

Kao, M. S., (2001). *A Unified Framework for Statistical Process Control of a Dynamic-Stochastic System in Manufacturing Transitions*. Ph.D. Dissertation, University of Wisconsin-Madison.

Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003). "On the Monitoring of Linear Profiles," *Journal of Quality Technology*, 35, 317-328.

Kourti, T. (2002). "Process Analysis and Abnormal Situation Detection: From Theory to Practice," *IEEE Control Systems Magazine*, Oct, 10-25.

Kramer, H. and Schmid, W. (1997). "EWMA Charts for Multivariate Time Series," *Sequential Analysis*, 16, 2, 131-154.

Li, R. and Chow, M. (2005). "Evaluation of Reproducibility for Paired Functional Data," *Journal of Multivariate Analysis*, 93, 81-101.

Loredo, E. N., Jearkpaporn, D., and Borror, C. M. (2002). "Model-Based Control Chart for Autoregressive and Correlated Data," *Quality and Reliability Engineering International*, 18, 6, 489-496.

Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, 34, 1, 46-53.

Lu, C. W. and Reynolds, M. R. (1999a). "EWMA Control Charts for Monitoring the Mean of Autocorrelated Processes," *Journal of Quality Technology*, 31, 2, 189-206.

Lu, C. W. and Reynolds, M. R. (1999b). "Control Charts for Monitoring the Mean and Variance of Autocorrelated Processes," *Journal of Quality Technology*, 31, 3, 259-274.

Luceño, A. (1999). "Average Run Lengths and Run Length Probability Distributions for Cuscore Charts to Control Normal Mean," *Computational Statistics and Data Analysis*, 32, 177-195.

Luceño, A. (2004). "Cuscore Charts to Detect Level Shifts in Autocorrelated Noise," *Quality Technology and Quantitative Management*, 1, 1, 27-45.

Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J., and Young, J. C. (1997). "Assessment of Multivariate Process Control Techniques," *Journal of Quality Technology*, 29, 2, 140-143.

Mastrangelo, C. M. and Forrest, D. R. (2002). "Process Monitoring Using Multivariate EWMA Residuals," working paper.

Mastrangelo, C. M. and Montgomery, D. C. (1995). "SPC with Correlated Observations for the Chemical and Process Industries," *Quality and Reliability Engineering International*, 11, 79-89.

Mahmoud, M. A. and Woodall, W. H. (2004). "Phase I Analysis of Linear Profiles With Calibration Applications," *Technometrics*, 46, 4, 380-391.

Mahmoud, M. A., Parker, P. A., Woodall, W. H. and Hawkins, D. M. (2006). "A Change Point Method for Linear Profile Data," *Quality and Reliability Engineering International*, to be published.

Montgomery, D. C. (2005). *Introduction to Statistical Quality Control, 5th ed.* Wiley, New York, NY.

Montgomery, D. C. and Mastrangelo, C. M. (1991). "Some Statistical Process Control Methods for Autocorrelated Data," *Journal of Quality Technology*, 23, 3, 179-193.

Montgomery, D. C., Johnson, L. A., and Gardiner, J. S. (1990). *Forecasting & Time Series Analysis.* McGraw-Hill, New York, NY.

Montgomery, D. C., Keats, J. B., Runger G. C., and Messina, W. S (1994). "Integrating Statistical Process Control and Engineering Process Control," *Journal of Quality Technology*, 26, 2, 79-87.

Nembhard, H. B. (2006). "Cuscore Statistics, Directed Process Monitoring for Early Problem Detection," in *Springer Handbook of Engineering Statistics*, Ed H. Pham, Springer-Verlag, London.

Nembhard, H. B. and Changpetch, P. (2006). "Detected Monitoring Using Cuscore Charts for Seasonal Time Series," to appear in *Quality and Reliability Engineering International*.

Nembhard, H. B. and Chen, S., (2006). "Cuscore Control Charts for Generalized Feedback Control Systems," to appear in *Quality and Reliability Engineering International*.

Nembhard, H. B. and Chen, S. (2006). "Multivariate Cuscore Control Charts for Monitoring the Mean Vector in Autocorrelated Processes," Working Paper.

Chen, S. and Nembhard, H. B. (2006). "A High-Dimensional Control Chart for Profile Monitoring," working paper.

Nembhard, H. B. and Valverde-Ventura, R. (2003). "Integrating Experimental Design and Statistical Control for Quality Improvement," *Journal of Quality Technology*, 35, 4, 406-423.

Nembhard, H. B. and Valverde-Ventura, R. (2006). "Cuscore Statistics to Monitor a Nonstationary System," to appear in *Quality and Reliability Engineering International.*

Neyman, J. (1937). "Smooth Test for Goodness of Fit," *Journal of Skandinavisk Aktuarietidskrift.* 20, 149-199.

Noorossana, R. and Vaghefi, S. J. M. (2005). "Effect of Autocorrelation on Performance of the MCUSUM Control Chart," *Quality and Reliability Engineering International*, 22, 2, 191-197.

Ramsay, J. O. and Silverman, B. W., (2005), *Functional Data Analysis, 2nd.* Springer, New York, NY.

Reinsel, G. C. (1997). *Elements of Multivariate Time Series Analysis, 2nd.* Springer, New York, NY.

Runger, G. C., Alt, F. B. and Montgomery, D. C. (1996). "Contributions to a Multivariate Statistical Process Control Signal," *Communications in Statistics-Theory and Methods*, 25, 10, 2203-2213.

Runger, G. C. and Testik, M. C. (2003). "Control Charts for Monitoring Fault Signatures: Cuscore versus GLR," *Quality and Reliability Engineering International* 19, 387-396.

Runger, G. C. and Willemain, T. R. (1995). "Model-Based and Model-Free Control Autocorrelated Processes," *Journal of Quality Technology*, 27, 4, 283-292.

Ruppert, D. (2002). "Selecting the Number of Knots for Penalized Splines," *Journal of Computational and Graphical Statistics*, 11, 4, 735–757.

Schenk, H. J. and Jackson, R. B. (2002). "The Global Biogeography of Roots," *Ecological Monographs*, 72, 3, 2002, 311–328

Shao, Y. E. (1998). "Integrated Application of the Cumulative Score Control Chart and Engineering Process Control," *Statistica Sinica*, 8, 239-252.

Shu, L., Apley, D. W. and Tsung, F. (2002). "Autocorrelated Process Monitoring Using Triggered Cuscore Charts," *Quality and Reliability Engineering International*, 18, 411–421.

Shu, L., Tsung, F., and Tsui, K. L. (2004). "Run-Length Performance of Regression Control Charts with Estimated Parameters," *Journal of Quality Technology*, 36, 3, 280-292.

Spitzner, D. J. and Woodall, W. H. (2003). "High-dimensional Directed Testing for Monitoring Functional Profiles," *Proceedings of the American Statistical Association*, 225–236.

Suh, J. H., Kumara, S. R. T. and Mysore, S. P., (1999). "Machinery Fault Diagnosis and Prognosis: Application of Advanced Signal Processing Techniques," *CIRP Annals-Manufacturing Technology*, 48, 1, 317-320.

Sullivan, J. H. and Woodall, W. H. (1996). "A Comparison of Multivariate Control Charts for Individual Observations," *Journal of Quality Technology*, 28, 4, 398-408.

Tsung, F. and Tsui, K.-L. (2003). "A Mean Shift Pattern Study on Integration of SPC and APC for Process Monitoring," *IIE Transactions*, 35, 231-242.

Tu, Y., Qiu, D. and Zhang, D. (2003). "A Revised Statistical Deduction Approach to Pipeline Leak Detection," *IMTC2003 – Instrumentation and Measurement Technology Conference*, 1639-1642.

Van der Wiel, S. A. (1996). "Monitoring Processes that Wander using Integrated Moving Average Models," *Technometrics*, 38, 2, 139-151.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S. N. (2003). "A Review of Process Fault Detection and Diagnosis, Part I: Quantitative model-based methods," *Computers and Chemical Engineering*, 27, 293-311.

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Wiley, New York, NY.

Walker, E. and Wright, S. P. (2002). "Comparing Curves Using Additive Models," *Journal of Quality Technology*, 34, 118–129.

Wang, G., Dong, D. and Fang, C. (1993). "Leak Detection for Transport Pipelines Based on Autoregressive Modeling," *IEEE Transactions on Instrumentation and Measurement*, 42, 1, 68-71.

Wang, K. and Tsung, F. (2004). "Using Profile Monitoring Techniques for a Data-rich Environment with Huge Sample Size," *Quality and Reliability Engineering International*, 21, 677-688.

Williams, J. D., Woodall, W. H. and Birch, J. B., (2003). "Phase I Monitoring of Nonlinear Profiles", 2003 Quality & Productivity Research Conference, Yorktown Heights, NY.

Williams, J. D., Woodall, W. H., Birch, J. B., and Sullivan, J. H. (2004). "Appropriate Upper Control Limits for the T^2 Control Chart Based on the Successive Differences Covariance Matrix Estimator," The 48th Annual Fall Technical Conference, Roanoke, VA.

Williams, J. D., Woodall, W. H., and Ferry, N. M. (2006). "Statistical Monitoring of Heteroscedastic Dose-Response Profiles from High-throughput Screening," *Technical Report No. 06-1* (http://www.stat.org.vt.edu/dept/web-e/tech_reports/TechReport06-1.pdf), Virginia Tech.

Woodall, W. H., Lorenzen, T. J. and Vance, L. C. (1986). "Weaknesses of the Economic Design of Control Charts," *Technometrics*, 28, 4, 408-410.

Woodall, W. H. (2000). "Controversies and Contradictions in Statistical Process Control," *Journal of Quality Technology*, 32, 4, 341-350.

Woodall, W. H., Spitzener, D. J., Montgomery, D. C. and Gupta S. (2004). "Using Control Charts to Monitor Process and Product Quality Profiles," *Journal of Quality Technology*, 36, 3, 309-320.

Yoon, S. and MacGregor, J. F. (2001). "Fault Diagnosis with Multivariate Statistical Models Part I: Using Steady-state Fault Signatures," *Journal of Process Control*, 11, 4, 387-400.

VITA

Shuohui Chen was born in Jiayuguan, P.R. China. He received his B.S. degree in Hydraulic Engineering from Tsinghua University in P.R. China in 1998, his M.S. in Civil Engineering from the University of Michigan at Ann Arbor in 2002, and his Ph.D. in Industrial Engineering from the Pennsylvania State University at University Park in 2006. He worked as an assistant project manager, a civil engineer and a software R&D engineer in china for three years. His research interests focus on developing statistical process control methods for improving and controlling the quality of manufacturing processes. The research presented in this dissertation involves one accepted paper, one submitted paper and two paper to be submitted in refereed journals. He is currently a member of ASQ.