GLR Control Charts for Monitoring the Mean Vector or the Dispersion of a Multivariate Normal Process

Sai Wang

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Statistics

Marion R. Reynolds, Jr., Chair
Dong-Yun Kim
Eric P. Smith
William H. Woodall

February 1, 2012
Blacksburg, Virginia

Keywords: Change point; Covariance; Generalized likelihood ratio; Multivariate CUSUM; Multivariate EWMA; Quality control; Statistical process control; Surveillance

GLR Control Charts for Monitoring the Mean Vector or the Dispersion of a Multivariate Normal Process

Sai Wang

(ABSTRACT)

In many applications, the quality of process outputs is described by more than one characteristic variable. These quality variables usually follow a multivariate normal (MN) distribution. This dissertation discusses the monitoring of the mean vector and the covariance matrix of MN processes.

The first part of this dissertation develops a statistical process control (SPC) chart based on a generalized likelihood ratio (GLR) statistic to monitor the mean vector. The performance of the GLR chart is compared to the performance of the Hotelling χ^2 chart, the multivariate exponentially weighted moving average (MEWMA) chart, and a multi-MEWMA combination. Results show that the Hotelling χ^2 chart and the MEWMA chart are only effective for a small range of shift sizes in the mean vector, while the GLR chart and some carefully designed multi-MEWMA combinations can give similarly better overall performance in detecting a wide range of shift magnitudes. Unlike most of these other options, the GLR chart does not require specification of tuning parameter values by the user. The GLR chart also has the advantage in process diagnostics: at the time of a signal, estimates of change-point and out-of-control mean vector are immediately available to the user. All these advantages of the GLR chart make it a favorable option for practitioners. For the design of the GLR chart, a series of easy to use equations are provided to users for calculating the control limit to achieve the desired in-control performance. The use of this GLR chart with a variable sampling interval (VSI) scheme has also been evaluated and discussed.

The rest of the dissertation considers the problem of monitoring the covariance matrix. Three GLR charts with different covariance matrix estimators have been discussed. Results show that the GLR chart with a multivariate exponentially weighted moving covariance (MEWMC) matrix estimator is slightly better than the existing method for detecting any general changes in the covariance matrix, and the GLR chart with a constrained maximum likelihood estimator (CMLE) gives much better overall performance for detecting a wide range of shift sizes than the best available options for detecting only variance increases.

Acknowledgments

I am heartily grateful to my advisor Dr. Marion Reynolds for his immeasurable help and guidance throughout this research. I am also grateful to the other professors on the committee: Dr. Dong-Yun Kim, Dr. Eric Smith and Dr. William Woodall, for their insightful advices. My gratitude also goes to the Advanced Research Computing at Virginia Tech and my peer researchers in the Reynolds QC group, including Wandi Huang, Yiming Peng, Lei Sun, Ning Wang, Pei Xiao and Liaosa Xu.

I owe my most special thanks to my parents and my brothers for always being there for me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Statistical Process Control (SPC) charts have been widely used in many applications, where quality measurements on the output from a process are monitored in real-time. Although the initial and the most frequent applications of control charts are in the manufacturing industries, the uses of control charts in many other areas, such as the health care and Internet industries, are expanding rapidly. For simplicity of explanation, the wording in this dissertation is based on a context of a manufacturing process, but the methods and ideas apply to any other types of process that fits the problem structure assumed here.

Suppose a manufacturing process operates continuously and produces a fixed number of items in each unit time interval. The quality of a product is determined by the measurement of characteristic(s) that describe(s) how well this product conforms to some specific requirements. Use p to denote the number of quality characteristics being measured. Despite the fact that advanced modern technology has greatly improved most manufacturing processes, items produced from a process at different times will not be exactly identical. The sources of variation in output quality are usually classified as two types: common cause and assignable cause (Montgomery 2009). Even under ideal operating conditions, the quality measurements of products from a process will vary from time to time due to allowable variations in worker skills, the precision limits of machinery, etc. These factors of variation

are inevitably inherent in the process, and are referred to as common cause variation. As the process runs over time, it is very likely that some incident will occur at some point of time, which will affect the quality of the output, for example a component in the machinery could wear out and impact the precision. Assignable causes refer to these sources of variation that are not natural to the process and can be eliminated if identified. The consequences of assignable causes on the distribution of quality measurements could be a change in the mean and/or a change in the variance. A process is defined as in control if the common cause is the only source of variation. In the presence of any assignable causes, a process is said to be out of control. Quality control charts are graphical tools used to distinguish between the in-control and out-of-control status of the process, where samples of product are inspected and values of some statistics based on these samples are plotted over time.

The earliest stage of SPC focused on the problem of monitoring in the univariate case, where the number of quality characteristics is $p = 1$. The Shewhart charts proposed by Shewhart (1931) are the first statistical control charts, and because of the ease of use in practice, they are still the most widely used ones. For Shewhart charts the statistics being calculated at each time point depend only on the current sample, which therefore results in good performance in detecting very large changes in the process. However, these charts are not as effective for detecting small changes due to the inability to gather information from past samples.

In order to quickly detect changes of small or intermediate magnitudes, two widely recognized alternative techniques are the cumulative sum (CUSUM) chart developed by Page (1954), and the exponentially weighted moving average (EWMA) chart introduced by Roberts (1959). These two alternatives allow information from past observations to be accumulated over time. Optimal performance at a specific size of change that is of interest can be obtained by the appropriate choice of the tuning parameters of the CUSUM and EWMA charts, however their abilities to detect other sizes of change will be compromised. These methods were originally developed for the problem of monitoring the process mean, and have been extended later to monitor other process parameters such as the process variance; see

Wortham (1971) and MacGregor and Harris (1993) for example.

The Shewhart, CUSUM and EWMA charts are perhaps the most widely known and used control charts. Targeting at different types of changes in the process, various modifications to these control charts have been extensively discussed in the literature. There is another family of control charts, the GLR chart, which is based on a generalized likelihood ratio statistic and a change-point model. The GLR chart has not received as much attention as the three charts that have already been mentioned, but it actually provides very attractive advantages; see for example, Willsky and Jones (1976), Lai (2001), and Hawkins, Qiu, and Kang (2001). Reynolds and Lou (2010) recently evaluated a control chart based on the GLR statistic to detect changes in the mean parameter of a univeriate normal process. They show that without the requirement for users to specify any tuning parameter values, the performance of the GLR chart is only slightly inferior to that of the CUSUM chart for shifts that the CUSUM is specifically tuned to detect, but much better for a wide range of shifts. Reynolds, Lou, Lee and Wang (2011) also discuss a GLR chart for detecting changes in both the mean and/or the variance, and results show an improvement in performance over exsiting competing methods.

In many SPC applications, the quality of an item from a process is characterized by more than one variable ($p > 1$), for example, the inner and outer diameters of an alloy bearing, and the length, width and height of a plastic rectangular component. As pointed out in Woodall and Montgomery (1999), Stoumbos, Reynolds, Ryan and Woodall (2000) and Woodall (2000), with the fast growing data acquisition technology and computing power, the problem of monitoring multivariate processes is an important research area for modern quality control.

A standard assumption made about processes with multiple continuous quality characteristics is that the vector $\boldsymbol{X}$ of quality variables of a randomly sampled item from the process follows a multivariate normal (MN) distribution with a mean vector $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$. Denote the parameters for the distribution of $\boldsymbol{X}$ when the process is in control as

$\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$, and use $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_1$ to denote these parameters for an out-of-control process. In this dissertation, we concentrate on Phase-II type control charts, where it is assumed that the in-control parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are either known or have been precisely estimated from Phase-I data. When an assignable cause occurs to the process, it could change either $\boldsymbol{\mu}$ or $\boldsymbol{\Sigma}$. In the first half of this dissertation, we consider the problem of monitoring $\boldsymbol{\mu}$ where it is assumed that only $\boldsymbol{\mu}$ can change while $\boldsymbol{\Sigma}$ always remains in-control (i.e. $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0$). The second half considers the problem of monitoring $\boldsymbol{\Sigma}$ with the assumption that only $\boldsymbol{\Sigma}$ can change while $\boldsymbol{\mu}$ stays unchanged (i.e. $\boldsymbol{\mu}_1 = \boldsymbol{\Sigma}_0$).

The objective of this dissertation is to evaluate the performance and explore the properties of Phase-II control charts based on the generalized likelihood ratio (GLR) statistics for monitoring the mean vector $\boldsymbol{\mu}$ or the covariance matrix $\boldsymbol{\Sigma}$ of a multivariate normal process. More specifically, the goals of this dissertation include the following:

1. For monitoring the mean vector $\boldsymbol{\mu}$, develop a GLR chart. Based on the performance of the GLR chart for the problem of univariate monitoring, we expect this multivariate GLR chart to have similar attractive properties, such as effectiveness in detecting a broad range of shift sizes, and no requirement for sophisticated tuning from the user. We will study the performance and properties of this GLR chart, and discuss how it can be used in practical applications.

2. Applications of the monitoring of MN dispersion can be categorized into two types based on which kind of change they are designed to detect: one for detecting any general changes in the covariance matrix $\boldsymbol{\Sigma}$, and the other for detecting only increases in the variability. In the second part of the dissertation, we develop GLR charts to monitor $\boldsymbol{\Sigma}$ for each type of change. We will evaluate the performance and properties of these charts, and discuss their applications.

This dissertation is presented with the following structure:

1. From Chapter 2 to Chapter 6 we discuss the monitoring of the mean vector. In Chapter

2 we provide the definition of the process of interest, explain the metric that is used to measure the performance of control charts, and then define the commonly used traditional multivariate control charts. In Chapter 3 we outline the derivation of the GLR chart for monitoring the mean vector, show the proof that this GLR chart is directionally invariant to the shift of the out-of-control mean, introduce the modified GLR chart with the restriction of a moving window, study the effect of the window size, and finally provide guidance that helps the user to choose the control limit corresponding to a desired in-control performance. In Chapter 4, the performance of the GLR chart for monitoring $\boldsymbol{\mu}$ is compared to the Hotelling χ^2 chart, individual MEWMA charts, as well as a scheme that simultaneously uses a combination of multiple MEWMA charts. In Chapter 5, we show that the GLR chart is equivalent to an infinite set of MCUSUM charts and discuss the approximation of the GLR chart with a finite set of MCUSUM charts. In Chapter 6, we discuss the use of the GLR chart with a variable sampling interval, compare the performance and provide guidance for users to choose the charting design parameters.

2. The monitoring of dispersion is discussed from Chapter 7 to Chapter 9. In Chapter 7 we define the process of interest and review existing methods. In Chapter 8 the GLR chart for monitoring $\boldsymbol{\Sigma}$ is derived, three variations of the GLR chart with different estimators are discussed: two for monitoring the general change and one for monitoring only increase in the variance, and the standardization of the GLR charts are discussed. In Chapter 9, we first define the settings for the out-of-control process, then compare the performance of the GLR charts to existing methods.

3. In Chapter 10, we draw summarizing conclusions and discuss some potential future research topics.

Chapter 2

Background on the Monitoring of $\boldsymbol{\mu}$

2.1 Definition of the Process

Suppose that the quality of products from the process of interest is characterized by p variables. Use the $p \times 1$ random vector $\boldsymbol{X} = [X_1, X_2, \ldots, X_p]'$ to denote the p quality measurements of an item. Suppose that when the process is in control, $\boldsymbol{X}$ follows a multivariate normal (MN) distribution with some known mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$.

Suppose that a fixed-rate sampling plan is adopted to monitor the process of interest. At each evenly spaced discrete sampling time point, an independent random sample of n items is taken from the production line and the p quality variables are measured immediately. Assume that the time interval between samples is one time unit. Then $\boldsymbol{X}_{k1}, \boldsymbol{X}_{k2}, \ldots, \boldsymbol{X}_{kn}$ are the n observations at the k^{th} sampling time point, with $\boldsymbol{X}_{ki} = [X_{ki1}, X_{ki2}, \ldots, X_{kip}]'$, where X_{kij} is the observation of the j^{th} variable on item i in sample k. The sample mean $\bar{\boldsymbol{X}}_k = \sum_{i=1}^{n} \boldsymbol{X}_{ki}/n$ is the sufficient statistic for monitoring the mean vector of a multivariate normal process. Monitoring the mean vector of a process with $\boldsymbol{X} \sim \text{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ using sample means $\bar{\boldsymbol{X}}_k$ is equivalent to monitoring a process of $\boldsymbol{Y} = \bar{\boldsymbol{X}} \sim \text{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ using single observations $\boldsymbol{Y}_k = \bar{\boldsymbol{X}}_k$. Therefore, without loss of generality, assume for now that $n = 1$

so that only one item is sampled at each inspection time. Later in the dissertation, when considering the problem of monitoring $\boldsymbol{\Sigma}$, we will consider the case of $n \geq 1$.

In this dissertation, we focus primarily on *sustained shifts* which are described as follows in the case of a shift in $\boldsymbol{\mu}$. Assume that at some random change-point $\tau_c \geq 0$, an assignable cause occurs in the process, causing the mean vector to shift from the in-control value $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$, while the covariance matrix $\boldsymbol{\Sigma}$ remains unchanged. Until this assignable cause is detected and removed from the process, the mean vector will remain at $\boldsymbol{\mu}_1$. The noncentrality parameter $\delta = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}$ is the Mahalanobis distance that measures the size of shift from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$. For any $\tau_c \geq 0$, there are two consecutive time points τ and $\tau + 1$ that satisfy $\tau \leq \tau_c < \tau + 1$. Items randomly sampled at or before time point τ are produced from the in-control process, while those sampled after τ come from the out-of-control process. For example, if $\tau_c = 5.6$ then $\tau = 5$, and $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_5 \sim \text{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, while $\boldsymbol{X}_6, \boldsymbol{X}_7, \cdots \sim \text{MN}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0)$. Note that the value of τ_c cannot be observed.

For some processes, the change in $\boldsymbol{\mu}$ induced by an assignable cause could be in a manner that is different from a *sustained shift*. In the SPC literature several different types of out-of-control situations have been considered, such as a *drift* where $\boldsymbol{\mu}$ gradually drifts away from $\boldsymbol{\mu}_0$ at a fixed rate, and a *transient shift* where $\boldsymbol{\mu}$ only goes out-of-control for a fixed amount of time and then returns to $\boldsymbol{\mu}_0$.

2.2 Literature Review

2.2.1 A General Review

As the first multivariate control chart, Hotelling (1947) extended the univariate Shewhart chart to the Hotelling χ^2 chart (the Hotelling T^2 chart if in-control parameters are unknown) to monitor the mean vector of multivariate normal processes. The Hotelling χ^2 chart signals if the chi-squared statistic for a random sample is greater than a predetermined control

limit. The control limit is chosen so that some specific in-control false alarm rate is achieved. Similar to its univariate version, this Shewhart-type multivariate control chart utilizes information only from the most recent sample, and thus is able to detect large departures from the in-control mean vector quickly, but is very ineffective for small or intermediate shifts.

Healy (1987) proposed a multivariate cumulative sum (MCUSUM) procedure derived from a set of sequential probability ratio tests (SPRT) for detecting a shift in $\boldsymbol{\mu}$ to some particular out-of-control mean vector $\boldsymbol{\mu}_c$ that the user wants to detect. The MCUSUM chart is only effective for detecting shifts along the pre-specified direction determined by $\boldsymbol{\mu}_c$, so it can be ineffective if $\boldsymbol{\mu}$ shifts to an unanticipated direction. Crosier (1988) discussed two multivariate CUSUM schemes, one based on the scalar-valued chi-squared sample statistics and the other based on the actual vectors of sample measurements, and the performance of both schemes only depends on the noncentrality parameter that measures the size of shift from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$. It has been shown that the vector-based multivariate CUSUM scheme has better performance than the scalar-based one. Pignatiello and Runger (1990) also proposed two alternative directionally invariant multivariate CUSUM procedures, one accumulates information from sample vectors and the other is based on the chi-squared statistics, and the former procedure was shown to be better. The performance and properties of both the vector-based multivariate CUSUM procedures discussed by Crosier (1988) and Pignatiello and Runger (1990) are similar. All of these directional invariant CUSUM-type charts are shown to be better than the Hotelling multivariate control chart in detecting shifts of small or intermediate size.

Lowry, Woodall, Champ and Rigdon (1992) presented a multivariate exponentially weighted moving average (MEWMA) scheme that is an extension of the univariate EWMA chart. The performance of the MEWMA chart was shown to be at least as good as the multivariate CUSUM charts of Crosier (1988) and Pignatiello and Runger (1990). All of these multivariate CUSUM charts and MEWMA chart resemble their univariate counterparts in the way that good performance for a particular size of shift can be obtained by carefully choosing the tuning parameter values.

The Hotelling χ^2 chart, the MEWMA chart, and different versions of the multivariate CUSUM charts are the most commonly used multivariate control charts. The MEWMA chart and CUSUM charts can be carefully designed and tuned for fast detection of some specific sizes or directions of the mean shift. However in practice, the information on how the mean vector is going to change when an assignable cause occurs in the process is rarely available ahead of time. Therefore, a control chart that works reasonably well for a wide range of possible shifts is more useful in most practical applications. To obtain good performance over a wide range of possible shifts, one option is to use multiple control charts together as a combination. This approach has been thoroughly explored in the literature of univariate process monitoring. Applications of this idea to multivariate SPC have also been investigated. For example, in their paper investigating the use of a combination of multivariate Shewhart and MEWMA charts for monitoring $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ simultaneously, Reynolds and Stoumbos (2008) showed that for the problem of monitoring $\boldsymbol{\mu}$, a combination of two charts performs much better for a wide range of shift sizes than either chart being used alone.

Zamba and Hawkins (2006) discussed a multivariate GLR-type control chart for monitoring $\boldsymbol{\mu}$, but they concentrated on short-run processes for which estimates of the in-control parameters based on a large amount of Phase-I data are not available. They compared their GLR chart to the Hotelling T^2 chart with estimated parameters and concluded that the GLR chart outperforms the latter for detecting medium and large size shifts.

A lot of other work has been done in the multivariate SPC literature; reviews of the most commonly recognized multivariate control procedures are available in Lowry and Montgomery (1995), Mason, Champ, Tracy, Wierda and Young (1997), Bersimis, Psarakis and Panaretos (2007), and Montgomery (2009).

For monitoring $\boldsymbol{\mu}$, we will consider the Hotelling χ^2 chart, the MEWMA chart and a Multi-MEWMA procedure that uses a combination of more than one MEWMA charts as the benchmarks when we evaluate the performance of our proposed control chart. Details about these control charts will be discussed in the following subsections.

2.2.2 The Hotelling χ^2 Chart

In Hotelling (1947), two forms of Shewhart-type control charts were proposed: one assumes that the in-control parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are both known, and the other assumes both to be unknown. The Hotelling T^2 chart is designed for the situation in which $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are both unknown or have to be estimated with non-negligible error. If $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are both known or have been precisely estimated from Phase I data, the control chart is called the Hotelling χ^2 chart and the control chart statistic at time k is $\chi_k^2 = (\boldsymbol{X}_k - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_k - \boldsymbol{\mu}_0)$. A signal will be given at time k if χ_k^2 exceeds the control limit h_{χ^2}.

The in-control distribution of χ_k^2 is a chi-square distribution with p degrees of freedom, while the out-of-control distribution is a noncentral chi-square distribution with the noncentrality parameter δ and p degrees of freedom. The Hotelling χ^2 chart is directionally invariant because the value of χ_k^2 only depends on $\boldsymbol{\mu}_1$ through the value of the noncentrality parameter δ. The control limit h_{χ^2} associated with any desired in-control ATS can be easily obtained by using the inverse cumulative density function (CDF) of the chi-square distribution, and the out-of-control ATS at different shift sizes can be calculated using the CDF of the noncentral chi-square distribution.

Note that for known $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$, the value of χ_k^2 at time k depends only on the current sampled observations $\boldsymbol{X}_k$. The fact that the Hotelling χ^2 chart does not accumulate any information from past samples means that it is only capable of fast detection of large shift sizes, but not small or intermediate shift magnitudes.

2.2.3 The MEWMA Chart

Coherent to the rationale of the EWMA procedure, Lowry et al. (1992) first defined an MEWMA vector at time k as $\boldsymbol{Z}_k = \lambda(\boldsymbol{X}_k - \boldsymbol{\mu}_0) + (1-\lambda)\boldsymbol{Z}_{k-1}$, where $0 < \lambda \leq 1$ and $\boldsymbol{Z}_0 = \boldsymbol{0}$. Then the charting statistic M_k is the quadratic form of this MEWMA vector, i.e. $M_k = \boldsymbol{Z}_k' \boldsymbol{\Sigma}_{0\boldsymbol{Z}_k}^{-1} \boldsymbol{Z}_k$, where $\boldsymbol{\Sigma}_{0\boldsymbol{Z}_k}$ is the covariance matrix of $\boldsymbol{Z}_k$. The authors have shown

that the exact form of the covariance matrix of $\boldsymbol{Z}_k$ is $\boldsymbol{\Sigma}_{0\boldsymbol{Z}_k} = \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2k}]\boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is the known covariance matrix of $\boldsymbol{X}_k$. Using this exact covariance matrix can give a fast initial response when the process goes out of control soon after the monitoring begins (when k is small). However it is more likely that the process will first stay in-control for some time before a shift occurs, therefore in most applications the asymptotic covariance matrix $\boldsymbol{\Sigma}_{0\boldsymbol{Z}_k} \cong \frac{\lambda}{2-\lambda}\boldsymbol{\Sigma}_0$ is used. In this dissertation we also use the asymptotic covariance, which leads to the MEWMA charting statistic $M_k = \frac{2-\lambda}{\lambda}\boldsymbol{Z}_k^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{Z}_k$. A signal is given as soon as $M_k > h_M$ where h_M is the pre-specfied control limit to give a desirable in-control performance.

The value of λ determines how weights should be allocated among all samples. To see this more clearly, express the recursive formula for $\boldsymbol{Z}_k$ as an explicit function of the weight parameter λ and all past samples $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_k$: $\boldsymbol{Z}_k = \sum_{i=1}^{k}\lambda(1-\lambda)^{k-i}(\boldsymbol{X}_i - \boldsymbol{\mu}_0)$ for $k = 1, 2, \ldots$ and $\boldsymbol{Z}_0 = \boldsymbol{0}$. The weight of sample $\boldsymbol{X}_i$ in the MEWMA vector $\boldsymbol{Z}_k$ decays at a constant rate of $1-\lambda$ as the sample time i goes back from k to 1. For large λ values, the weights of recent samples dominate those of older samples; whereas for small λ values, the weights are allocated more evenly on past observations, allowing the MEWMA vector to accumulate more information from past samples. Therefore MEWMA charts with large λ values are only effective for detecting large shifts, while those with small λ give better performance at small shift sizes. Sometimes in applications, fast detection at some particular shift magnitude is of more interest to the user or the out-of-control mean shift is known to be around some particular size, then the MEWMA chart can be tuned to give good performance by selecting an appropriate value of λ. Note that when $\lambda = 1$, the MEWMA vector $\boldsymbol{Z}_k = \boldsymbol{X}_k - \boldsymbol{\mu}_0$ and the MEWMA control chart becomes the Hotelling χ^2 chart.

Ridgon (1995a, 1995b) used integral equations to analyze the in-control and out-of-control performance of the MEWMA chart. Runger and Prabhu (1996) introduced a Markov chain method to approximate the average run length performance of the MEWMA chart. But neither method will work for the combination of two or more MEWMA charts which will be discussed in the following section. Therefore simulation is used here to find control

limits and to assess out-of-control performance.

2.2.4 The Multi-MEWMA Combination

The MEWMA chart can be tuned to give better performance at some specific shift sizes by choosing an appropriate value for the tuning parameter λ. However in most practical applications, the user does not have much information on what shift sizes should be expected. A control charting scheme that works generally well for a wide range of shift sizes is usually preferable to the one that is only optimal for a specific shift magnitude but worse elsewhere.

One intuitive approach to obtaining good performance for a wide range of shift sizes is to use a combination of multiple MEWMA control charts to monitor the process simultaneously; each individual chart in the combination should be sensitive to a different region of shift sizes. Several papers have mentioned the possibility of using this scheme, and Reynolds and Stoumbos (2008) used combinations of multivariate Shewhart and MEWMA charts to monitor both the mean vector and/or the covariance matrix, but no dedicated work has been done to guide the use of multiple MEWMA charts simultaneously for monitoring $\boldsymbol{\mu}$. Here we consider using two or three MEWMA control charts with different λ values. We will refer to the general combination of multiple MEWMA charts as the multi-MEWMA chart, and 2-MEWMA/3-MEWMA specifically for those with the number of components being two/three.

Intuitively the value of λ for each MEWMA component should be remarkably different from each other, so that for any particular shift magnitude, there is an individual chart that is more sensitive to this size of shift than the other components. For each individual MEWMA chart in the combination, a designated control limit is necessary. The global multi-MEWMA scheme signals as soon as any individual chart signals. Given any specific false alarm rate for the overall multi-MEWMA scheme, there is infinite number of combinations of individual control limits that can be found. As an arbitrary example, consider the 3-MEWMA combination with $(\lambda_1, \lambda_2, \lambda_3) = (0.05, 0.2, 0.9)$; the three MEWMA components

are presumably sensitive to small, medium and large shift sizes. Suppose that two sets of control limits $\boldsymbol{h}_1 = (a, b, c)$ and $\boldsymbol{h}_2 = (d, e, f)$ both give an in-control ATS of 800 for the global 3-MEWMA combination, while the in-control ATS values when using each MEWMA component alone with these corresponding control limits are $(2000, 2000, 2000)$ for (a, b, c) and $(1000, 2000, 3000)$ for (d, e, f). It is clear that $(a > d, b = e, c < f)$. Then the 3-MEWMA combination using control limits $\boldsymbol{h}_1 = (a, b, c)$ will be worse for small shifts and better in detecting large shifts than using $\boldsymbol{h}_2 = (d, e, f)$. Therefore, besides the choice of λ values, the choice of control limits can also serve as tuning parameters that will affect the performance of the multi-MEWMA scheme. For simplicity, here we only consider choosing the set of control limits so that the overall in-control ATS is satisfied and each MEWMA component has approximately the same in-control ATS when used by itself. Extensive simulations have to be performed to find the appropriate set of control limits.

2.3 Performance Measurements of Control Charts

Control charts are used in process monitoring, so that a change in the process can be detected as soon as possible after its occurrence. When a control chart signals, it suggests that there might be some assignable causes that have driven the process out of control. But the signal is not always correct. When the process is in control, a signal could be triggered by the common cause variation in the process. Analogous to the type I error in hypothesis testing, a signal of a control chart when the process is actually in control is called a false alarm. In applications, when a signal is triggered, whether it is correct or false is not immediately known to the user. An excessive number of false alarms not only destroys the user's confidence but also creates unnecessary cost associated with process diagnosis and possible production interruption. If the process is out of control, an earlier signal reduces the costs for producing non-conforming products. Therefore, it is desirable for a control chart to signal as soon as possible after the process goes out of control, and to have a low rate of false alarms while the process stays in control.

A common metric for evaluating the performance of a control chart is the average run length (ARL), which is the average number of samples before a signal. Therefore the ARL is also called the average number of samples to signal (ANSS). Another widely used performance measurement is the average time to signal (ATS). If samples are taken at evenly spaced time points with intervals of fixed length d apart, then ATS $= d(ANSS)$. The in-control ATS measures the rate of false alarms while the out-of-control ATS measures how long it takes on average for a control chart to detect the change in the process. For the out-of-control performance metric, some research papers use the initial state average time to signal (ISATS), which simply assumes that the process is already out of control when control charts start to be used, with the control statistics in their initial states. But a more realistic scenario is that the process is in control when the control chart starts being implemented and remains in control for a while, allowing the control chart statistic to reach its steady state distribution, then later on at some random time τ_c which is uniformly distributed on the interval of $[\tau, \tau + 1)$, the process goes out of control due to some assignable cause. The steady state average time to signal (SSATS) is commonly used to provide more reasonable comparisons of out-of-control performance. In this dissertation, the SSATS will be used for most comparisons.

To compare the performance of different control charts, the idea is to first specify the control limit for each control chart so that they all have the same in-control ATS, and then compare their SSATS values under different out-of-control situations. For the same in-control ATS, at a given shift size the control chart with the smaller SSATS value is better in terms of the ability to quickly detect shifts of this magnitude.

Chapter 3

The GLR Chart for Monitoring $\boldsymbol{\mu}$

3.1 Derivation of the GLR Chart

Let f denote the probability density function (pdf) of a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e.,

$$f(\boldsymbol{X}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = (2\pi)^{-\frac{p}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right\}.$$

At a sampling time point k, a series of random vectors $(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_k)$ has been observed. If the process is still in control at the k^{th} sampling time point ($k \leq \tau$), then all the k samples come from the in-control distribution and the likelihood is

$$L_{k,\infty}(\boldsymbol{\mu}_0) = \prod_{i=1}^{k} f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0).$$

If there has been a change in the process mean by the time of k where $k \geq \tau + 1$, then samples taken at or before τ have the in-control mean of $\boldsymbol{\mu}_0$ while those taken after τ come

from the out-of-control distribution. The likelihood of the mean vector is then

$$L_{k,\tau}(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = \prod_{i=1}^{\tau} f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \times \prod_{i=\tau+1}^{k} f(\boldsymbol{X}_i|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0).$$

For a known $\tau < k$, the maximum likelihood estimator (MLE) of $\boldsymbol{\mu}_1$ is

$$\hat{\boldsymbol{\mu}}_{1,k,\tau} = \frac{\sum_{i=\tau+1}^{k} \boldsymbol{X}_i}{k-\tau}.$$

However, τ has to be estimated as well since the change-point τ_c is unobservable. Under the assumption that the process mean has changed prior to time k, the MLE of τ can be obtained by maximizing the profile likelihood

$$\begin{aligned}\hat{\tau} &= \arg\max_{0\le t<k} L_{k,t}(\boldsymbol{\mu}_0, \hat{\boldsymbol{\mu}}_{1,k,t}) \\ &= \arg\max_{0\le t<k} \left\{ \prod_{i=1}^{t} f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \times \prod_{i=t+1}^{k} f(\boldsymbol{X}_i|\hat{\boldsymbol{\mu}}_{1,k,t}, \boldsymbol{\Sigma}_0) \right\}.\end{aligned}$$

Then $(\hat{\tau}, \hat{\boldsymbol{\mu}}_{1,k,\hat{\tau}})$ is the MLE of $(\tau, \boldsymbol{\mu}_1)$.

Consider the log likelihood-ratio for testing whether there has been a change in the process prior to the current time point, i.e. H_0: the process is in-control versus H_1: there has been a shift in the mean. The test statistic at time point k is

$$\begin{aligned}R_k &= \log \frac{L_{k,\hat{\tau}}(\boldsymbol{\mu}_0, \hat{\boldsymbol{\mu}}_{1,k,\hat{\tau}})}{L_{k,\infty}(\boldsymbol{\mu}_0)} \\ &= \log \frac{\max_{0\le t<k} \left\{ \prod_{i=1}^{t} f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \times \prod_{i=t+1}^{k} f(\boldsymbol{X}_i|\hat{\boldsymbol{\mu}}_{1,k,t}, \boldsymbol{\Sigma}_0) \right\}}{\prod_{i=1}^{k} f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)} \\ &= \max_{0\le t<k} -\frac{1}{2} \sum_{i=t+1}^{k} \left\{ (\boldsymbol{X}_i - \hat{\boldsymbol{\mu}}_{1,k,t})' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \hat{\boldsymbol{\mu}}_{1,k,t}) - (\boldsymbol{X}_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) \right\} \\ &= \max_{0\le t<k} \frac{k-t}{2} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0).\end{aligned}$$

If the value of R_k exceeds some pre-specified control limit h_{GLR}, then the GLR control chart signals at time k and declares that at some time between $\hat{\tau}$ and $\hat{\tau}+1$, the mean vector of the process shifted to $\hat{\boldsymbol{\mu}}_{1,k,\hat{\tau}}$. If no signal is triggered then conclude that there is not significant evidence indicating that $\boldsymbol{\mu}$ has deviated from $\boldsymbol{\mu}_0$ and continue the sampling scheme at succeeding time points.

3.2 The Directional Invariance Property

In this section, we show that the GLR control chart is invariant to the direction of the out-of-control mean vector $\boldsymbol{\mu}_1$ and the form of covariance matrix $\boldsymbol{\Sigma}_0$. For a given dimensionality p, the performance for the GLR chart only depends on $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_0$ through the value of the noncentrality parameter $\delta = \sqrt{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)}$.

Without loss of generality, assume that the in-control mean vector $\boldsymbol{\mu}_0 = \mathbf{0}$. Then the GLR statistic at time k is

$$R_k = \max_{0\le t<k} \frac{k-t}{2}\hat{\boldsymbol{\mu}}_{1,k,t}^{'}\boldsymbol{\Sigma}_0^{-1}\hat{\boldsymbol{\mu}}_{1,k,t}, \text{ where } \hat{\boldsymbol{\mu}}_{1,k,t} = \frac{\sum_{i=t+1}^{k}\boldsymbol{X}_i}{k-t}.$$

Consider a full-rank transformation of the data vector $\boldsymbol{X}^* = \boldsymbol{M}\boldsymbol{X}$, where $\boldsymbol{M}$ is a $p\times p$ full-rank matrix. Note that $\boldsymbol{X}^*$ still follows a multivariate normal distribution with $\mathrm{E}(\boldsymbol{X}^*) = \boldsymbol{M}\mathrm{E}(\boldsymbol{X}) = \boldsymbol{M}\boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{X}^*) = \boldsymbol{M}\mathrm{Var}(\boldsymbol{X})\boldsymbol{M}^{'} = \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^{'}$. When the process is in control, the parameters for $\boldsymbol{X}^*$ are $\mathrm{E}(\boldsymbol{X}^*) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{X}^*) = \boldsymbol{M}\boldsymbol{\Sigma}_0\boldsymbol{M}^{'}$. For a given $0\le t<k$, the estimated out-of-control mean vector calculated from $\boldsymbol{X}_i^*$ is $\hat{\boldsymbol{\mu}}_{1,k,t}^* = \frac{\sum_{i=t+1}^{k}\boldsymbol{X}_i^*}{k-t} = \boldsymbol{M}\hat{\boldsymbol{\mu}}_{1,k,t}$. Then the GLR statistic calculated from $\boldsymbol{X}^*$ is

$$R_k^* = \max_{0\le t<k} \frac{k-t}{2}\hat{\boldsymbol{\mu}}_{1,k,t}^{*'}\boldsymbol{\Sigma}_{0\boldsymbol{X}^*}^{-1}\hat{\boldsymbol{\mu}}_{1,k,t}^* = \max_{0\le t<k} \frac{k-t}{2}\hat{\boldsymbol{\mu}}_{1,k,t}^{'}\boldsymbol{M}^{'}\boldsymbol{M}^{'-1}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{M}^{-1}\boldsymbol{M}\hat{\boldsymbol{\mu}}_{1,k,t} = R_k.$$

This means that value of the GLR statistic calculated from the transformed data vector $\boldsymbol{X}^*$ is exactly the same as that being calculated using $\boldsymbol{X}$.

Crosier (1988) has shown that the performance of the two multivariate CUSUM schemes that he proposed depends on $\boldsymbol{\mu}_1$ and $\boldsymbol{\Sigma}_0$ only through the noncentrality parameter. His results can be applied here to show the directional invariance of the GLR chart.

For two different mean vectors $\boldsymbol{\mu}_a \neq \boldsymbol{\mu}_b$ satisfying that $\boldsymbol{\mu}_a^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_a = \boldsymbol{\mu}_b^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b$, there exists a nonsingular matrix $\boldsymbol{M}$ such that $\boldsymbol{\mu}_a = \boldsymbol{M}\boldsymbol{\mu}_b$. To find $\boldsymbol{M}$, first consider an orthogonal transformation $\boldsymbol{P}$ that transforms $\boldsymbol{X}$ to its principal components, i.e. $\boldsymbol{P}^{'}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}^{'} = \boldsymbol{I}$ and $\text{Var}(\boldsymbol{P}\boldsymbol{X}) = \boldsymbol{P}\boldsymbol{\Sigma_0}\boldsymbol{P}^{'} = \boldsymbol{D}$ where $\boldsymbol{D}$ is a diagonal matrix of eigenvalues of $\boldsymbol{\Sigma}_0$ and $\boldsymbol{I}$ is the identity matrix whose diagonal elements are all 1 and off-diagonal elements are all 0. Then $\boldsymbol{D}^{-1/2}\boldsymbol{P}$ is the transformation matrix that transforms $\boldsymbol{X}$ to its principal components whose variances are scaled to 1, i.e. $\text{Var}(\boldsymbol{D}^{1/2}\boldsymbol{P}\boldsymbol{X}) = \boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\Sigma_0}\boldsymbol{P}^{'}\boldsymbol{D}^{-1/2} = \boldsymbol{I}$. Denote these principal components as $\boldsymbol{Y} = \boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{X}$. Let two $p \times 1$ vectors $\boldsymbol{\nu}_a = \boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\mu}_a$ and $\boldsymbol{\nu}_b = \boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\mu}_b$. Then it follows directly that $\boldsymbol{\nu}_a^{'}\boldsymbol{\nu}_a = \boldsymbol{\nu}_b^{'}\boldsymbol{\nu}_b$ because $\boldsymbol{\nu}_a^{'}\boldsymbol{\nu}_a = \boldsymbol{\mu}_a^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_a$, $\boldsymbol{\nu}_b^{'}\boldsymbol{\nu}_b = \boldsymbol{\mu}_b^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b$, and $\boldsymbol{\mu}_a^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_a = \boldsymbol{\mu}_b^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b$. Note that from linear algebra, there always exists an orthogonal matrix $\boldsymbol{Q}$ that rotates $\boldsymbol{\nu}_b$ to $\boldsymbol{\nu}_a$, i.e. $\boldsymbol{\nu}_a = \boldsymbol{Q}\boldsymbol{\nu}_b$. From definitions of $\boldsymbol{\nu}_a$ and $\boldsymbol{\nu}_b$, we have $\boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\mu}_a = \boldsymbol{Q}\boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\mu}_b$. Therefore, for two different mean vectors $\boldsymbol{\mu}_a$ and $\boldsymbol{\mu}_b$ of the same shift size δ, there exists a nonsingular matrix $\boldsymbol{M} = \boldsymbol{P}^{'}\boldsymbol{D}^{1/2}\boldsymbol{Q}\boldsymbol{D}^{-1/2}\boldsymbol{P}$, such that $\boldsymbol{\mu}_a = \boldsymbol{M}\boldsymbol{\mu}_b$.

Consider two random variables $\boldsymbol{X}_a$ and $\boldsymbol{X}_b$ where $\boldsymbol{X}_a \sim \text{MN}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_0)$ and $\boldsymbol{X}_b \sim \text{MN}(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_0)$. Also consider a transformed random variable $\boldsymbol{X}_b^* = \boldsymbol{M}\boldsymbol{X}_b$ using the transformation matrix given in the previous paragraph. We have shown that the values of the GLR statistic when calculated from $\boldsymbol{X}_b^*$ are the same as those based on $\boldsymbol{X}_b$. Note that $\text{E}(\boldsymbol{X}_b^*) = \boldsymbol{M}\boldsymbol{\mu}_b = \boldsymbol{\mu}_a$ and $\text{Var}(\boldsymbol{X}_b^*) = \boldsymbol{M}\boldsymbol{\Sigma}_0\boldsymbol{M}^{'} = \boldsymbol{P}^{'}\boldsymbol{D}^{1/2}\boldsymbol{Q}\boldsymbol{D}^{-1/2}\boldsymbol{P}\boldsymbol{\Sigma}_0\boldsymbol{P}^{'}\boldsymbol{D}^{-1/2}\boldsymbol{Q}^{'}\boldsymbol{D}^{1/2}\boldsymbol{P} = \boldsymbol{\Sigma}_0$ (by substituting $\boldsymbol{P}\boldsymbol{\Sigma}_0\boldsymbol{P}^{'} = \boldsymbol{D}$). Therefore $\boldsymbol{X}_b^*$ has the same distribution as $\boldsymbol{X}_a$, which is $\text{MN}(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_0)$. So the distributions of the GLR statistic for $\boldsymbol{X}_a$ and $\boldsymbol{X}_b$ are the same when $\boldsymbol{\mu}_a^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_a = \boldsymbol{\mu}_b^{'}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b$.

These results mean that for the same p, the performance of the GLR chart does not depend on the direction of $\boldsymbol{\mu}_1$ nor the structure of $\boldsymbol{\Sigma}_0$. As long as the noncentrality

parameter δ is the same, the performance is the same. This implies that when evaluating the performance of the GLR chart through simulation, we can simply use $\boldsymbol{\Sigma}_0 = \boldsymbol{I}$ and $\boldsymbol{\mu}_1 = [\delta\ 0 \dots 0]'$.

3.3 The GLR Control Chart With a Moving Window

Note that at a given sampling time point k, the calculation of R_k requires the calculation of $L_{k,t}(\boldsymbol{\mu}_0, \hat{\boldsymbol{\mu}}_{1,k,t})$ for all possible change-points, i.e. $t \in \{0, 1, \dots, k-1\}$. This means that all the past observations have to be kept in memory too. With current computing technology, using the GLR chart in this form in applications will not be very difficult unless an extremely long sequence of samples is anticipated. However when studying the properties and evaluating the performance of the GLR chart with computer simulations, this computational burden stands out. Instead of evaluating $L_{k,t}(\boldsymbol{\mu}_0, \hat{\boldsymbol{\mu}}_{1,k,t})$ at all past time points, an alternative way is to adopt a moving window which restricts the evaluations to within $[\max(0, k-m), k)$, where m is the window size. If $k \leq m$, evaluate $L_{k,t}(\boldsymbol{\mu}_0, \hat{\boldsymbol{\mu}}_{1,k,t})$ at all past time points, i.e. $t \in \{0, 1, \dots, k-1\}$. If $k > m$, evaluate only at the m most recent time points, i.e. $t \in \{k-m, k-m+1, \dots, k-1\}$. This means that the user only has to record the quality measurements for at most m of the most recent samples. The modified control statistic now becomes

$$R_{k,m} = \max_{\max(0,k-m)\leq t<k} \frac{k-t}{2} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0).$$

In the following section where the effect of windows size m is discussed, it can be observed that m also serves as a tuning parameter for this multivariate GLR control chart. But the purpose of introducing this modification is to allow a faster evaluation through simulations and to overcome some possible data storage problems. In practical applications, the GLR control chart without the window limitation or the GLR control chart with a very large window size are recommended because these options give the best overall performance.

3.4 Effect of the Window Size

In this section, the effect of window size m in the performance of the GLR control chart is studied. Simulation results shown here are for the case of $p = 4$ dimensions, but similar patterns can be observed for other dimensions.

In Table 3.1, columns labeled [1],. . . , [7] give the ATS and SSATS values for GLR charts with different window sizes $m \in\{25, 50, 100, 200, 400, 600 \text{ and } 800\}$. For each choice of the window size m, the control limit h_{GLR} is chosen based on simulation so that the in-control ($\delta = 0$) ATS is approximately 800. For out-of-control situations, SSATS values at a wide range of 16 different shift sizes $\delta \in\{0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2, 2.5, 3, 4, 5, 8, \text{ and } 12\}$ are provided. When simulating SSATS values for the GLR control chart with window size m at a specific shift size δ, a sequence of $\tau = 400$ random in-control observations was first generated and used for monitoring. The idea is to allow the GLR control chart statistic $R_{k,m}$ to reach its steady state distribution at time $k = 400$, conditioned on the fact that no false alarm was present. At a random time τ_c between samples 400 and 401 (assuming the distribution of τ_c is uniform on this interval), the process mean shifted to $\boldsymbol{\mu}_1 = [\delta\ 0\ 0\ 0]'$. Then the SSATS of this control chart for shifts of size δ is the average time to signal from the true change-time: $\text{E}(T - \tau_c)$, where T is the time when the GLR chart signals. Results shown here are based on one million iterations.

First we compare the control limits of the GLR charts. The moving window restricts the evaluation of the likelihood ratio to within the m recent samples. Given the same data sequence, the GLR statistic with a larger window size will always be greater or equal to that with a smaller m, which means that for the same control limit h_{GLR} the GLR chart with a larger window size will signal no later than the one with a smaller m. Therefore, for each GLR chart to achieve the same in-control ATS, the control limit goes up as the window size m increases. Then compare the performances of the GLR charts at very small shift sizes, say at δ =0.2 and 0.4. The SSATS values of the GLR charts with small window sizes are much

Table 3.1: Effect of the GLR Window Size m on the SSATS ($p = 4$)

$m =$	25	50	100	200	400	600	800	12000
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
0.0	799.99	799.99	800.02	800.10	799.97	800.00	799.97	800.02
0.2	499.35	428.38	356.78	295.03	256.73	247.49	245.53	245.02
0.4	169.63	116.28	88.73	80.70	80.05	79.95	80.05	80.06
0.6	58.37	43.02	39.42	39.11	39.12	39.15	39.20	39.21
0.8	26.83	23.52	23.29	23.42	23.40	23.41	23.40	23.41
1.0	16.01	15.57	15.61	15.64	15.66	15.66	15.66	15.67
1.2	11.21	11.18	11.25	11.26	11.26	11.27	11.27	11.29
1.4	8.42	8.45	8.51	8.50	8.53	8.53	8.53	8.53
1.6	6.59	6.65	6.67	6.68	6.70	6.70	6.70	6.70
1.8	5.33	5.37	5.40	5.39	5.40	5.41	5.41	5.41
2.0	4.39	4.43	4.45	4.46	4.46	4.46	4.46	4.46
2.5	2.93	2.94	2.96	2.96	2.96	2.96	2.96	2.96
3.0	2.09	2.09	2.10	2.11	2.11	2.11	2.11	2.11
4.0	1.20	1.21	1.21	1.22	1.22	1.22	1.22	1.22
5.0	0.75	0.76	0.76	0.76	0.76	0.76	0.76	0.76
8.0	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
12.0	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
$h_{\text{GLR}} =$	10.7590	10.8444	10.8847	10.9044	10.9104	10.9122	10.9138	10.9139

larger than those with large windows. The reason is that when the shift size is very small, the GLR chart needs a large number of out-of-control observations to trigger a signal; if the window size m of the GLR chart is too small, then a big portion of the useful information will be lost. Therefore the larger the window size m is, the better ability of the GLR chart to detect small shifts. On the other hand when the shift sizes are large, say at shifts of size $\delta \geq 2$, the GLR chart with the smallest window size $m = 25$ gives the smallest SSATS values. The reason is that, if the shift size is large, only a few out-of-control samples would be enough to trigger a signal; the control limit is lower for the GLR chart with the smallest window size, allowing for faster detection of large shifts.

The window size m thus serves as a tuning parameter for the GLR chart. By choosing different window sizes, the GLR chart can be tuned to give different performance at some particular shift magnitudes. However, after a closer look at the SSATS values, it appears that

for the GLR charts with large window sizes the SSATS values at large shifts are only slightly smaller than those with small window sizes; whereas the performance at small shifts is much better than the GLR charts with small m. As the window size m increases, the performance converges to the GLR chart with no moving window restriction. The SSATS values of the GLR chart with $m = 600$ are very close to those of $m = 800$, even at the smallest shift size that we explored. Therefore in Chapter 4 where we compare the performance of the GLR chart with other multivariate control charts, the window size of $m = 600$ will be used.

In Table 3.1, the control limits for each choice of window size are chosen so that all of the GLR charts have almost the same in-control ATS value, which allows for a fair comparison of the out-of-control performances of the GLR charts with different window sizes. Another perspective for evaluating the effect of the window size m is to compare the ATS differences for the GLR charts with different window sizes for the same control limit under both in-control and out-of-control situations. In Table 3.2, we tabulate the ATS values of GLR charts with two different window sizes ($m_1 = 600$ and $m_2 = 12000$) for a range of different control limits $h_{\text{GLR}} \in \{7, 7.5, \ldots, 14.5\}$ under the in-control ($\delta = 0$) situation and several out-of-control ($\delta > 0$) scenarios. The results provided in this table are also for the case of $p = 4$, but similar patterns can be observed for other dimensions. To provide easier comparisons, two rows of ATS values are stacked together at each control limit value, where the upper row corresponds to the GLR with $m_1 = 600$ and the lower one is for $m_2 = 12000$. For the two window sizes, there are only differences in ATS values when the process is in-control or when there is a shift of very small magnitude ($\delta = 0.2$); other than shifts of size $\delta = 0.2$, the out-of-control ATS values of the two GLR charts are almost identical. For $h_{\text{GLR}} \leq 11$, the in-control ATS values are less than 1000. For such false alarm rates, a window size of 12000 is much larger than necessary, and the performance approximates that of the GLR chart with no moving window restriction. When comparing the ATS values of the GLR chart with $m_1 = 600$ to those of $m_2 = 12000$ at these control limits, only small differences can be found when $\delta = 0$ or 0.2. As the control limit h_{GLR} increases, the ATS values also increase, as well as the differences between the ATS values of the two window

Table 3.2: ATS Comparison of GLR chart with $m_1 = 600$ and $m_2 = 12000$ ($p = 4$)

h_{GLR}	$\delta = 0.0$	0.2	0.4	0.6	0.8	1.0	2.0	3.0	5.0	8.0	
7	47.8	40.1	26.7	17.6	12.2	8.9	3.2	1.8	1.1	1.0	m_1
	47.8	39.9	26.8	17.5	12.1	8.9	3.2	1.8	1.1	1.0	m_2
7.5	67.1	53.5	32.8	20.5	13.8	9.9	3.5	1.9	1.1	1.0	m_1
	66.9	53.4	32.9	20.4	13.8	9.9	3.5	1.9	1.1	1.0	m_2
8	94.6	70.9	39.7	23.6	15.5	11.0	3.7	2.0	1.1	1.0	m_1
	94.3	70.9	39.7	23.5	15.4	11.0	3.7	2.0	1.1	1.0	m_2
8.5	134.4	92.8	47.1	26.8	17.2	12.0	4.0	2.1	1.1	1.0	m_1
	134.4	92.9	47.1	26.7	17.1	12.0	3.9	2.1	1.1	1.0	m_2
9	192.0	119.8	54.9	30.0	18.9	13.1	4.2	2.2	1.1	1.0	m_1
	191.8	120.1	54.9	30.0	18.9	13.1	4.2	2.2	1.1	1.0	m_2
9.5	276.1	151.7	63.0	33.3	20.6	14.1	4.4	2.3	1.2	1.0	m_1
	274.8	151.8	63.1	33.2	20.6	14.1	4.4	2.4	1.2	1.0	m_2
10	400.1	187.9	71.3	36.5	22.3	15.2	4.7	2.5	1.2	1.0	m_1
	398.8	187.3	71.3	36.5	22.3	15.2	4.7	2.5	1.2	1.0	m_2
10.5	582.6	227.6	79.5	39.8	24.1	16.3	4.9	2.6	1.2	1.0	m_1
	579.9	226.4	79.6	39.7	24.0	16.3	4.9	2.6	1.2	1.0	m_2
11	855.4	269.8	87.6	43.1	25.8	17.3	5.2	2.7	1.3	1.0	m_1
	848.8	266.8	87.7	42.9	25.8	17.3	5.2	2.7	1.3	1.0	m_2
11.5	1261.3	313.4	95.6	46.2	27.5	18.4	5.4	2.8	1.3	1.0	m_1
	1251.4	307.8	95.7	46.2	27.5	18.4	5.5	2.8	1.3	1.0	m_2
12	1874.0	357.6	103.4	49.4	29.2	19.4	5.7	2.9	1.4	1.0	m_1
	1861.6	348.5	103.5	49.4	29.2	19.5	5.7	2.9	1.4	1.0	m_2
12.5	2791.2	402.2	111.1	52.5	30.9	20.5	5.9	3.0	1.4	1.0	m_1
	2780.5	388.1	111.2	52.5	30.9	20.5	6.0	3.0	1.4	1.0	m_2
13	4186.1	447.5	118.6	55.6	32.5	21.5	6.2	3.1	1.4	1.0	m_1
	4150.7	425.9	118.7	55.6	32.5	21.5	6.2	3.1	1.4	1.0	m_2
13.5	6305.6	493.0	125.9	58.7	34.2	22.6	6.5	3.2	1.5	1.0	m_1
	6242.0	461.8	126.0	58.6	34.2	22.6	6.5	3.2	1.5	1.0	m_2
14	9549.4	539.6	133.1	61.7	35.8	23.6	6.7	3.3	1.5	1.0	m_1
	9435.5	496.0	133.2	61.7	35.8	23.6	6.7	3.4	1.5	1.0	m_2
14.5	14486.0	588.1	140.2	64.7	37.5	24.7	7.0	3.5	1.6	1.0	m_1
	14349.5	528.4	140.2	64.7	37.5	24.7	7.0	3.5	1.6	1.0	m_2

sizes at $\delta = 0$ and 0.2. For example, when $h = 14.5$, the in-control ATS of m_1 is 14486 while that of m_2 is 14349, but the difference of 137 is only 0.95% of the in-control ATS. This fact justifies the use of the GLR chart with window size $m = 600$ in the following section where we try to estimate the relationship between the control limit h_{GLR} and the in-control ATS.

Our findings on the choice of m here can be justified by the asymptotic properties of the GLR chart. It has been shown in Lai (1998) that the window-restricted GLR chart is asympototically optimal as $h_{\text{GLR}} \to \infty$ if the window size satisfies $m = O(\log(\text{in-control ATS})/I_{\text{KL}})$, where I_{KL} is the Kullback-Leibler information number corresponding to the smallest shift that is of interest. For the case of a mean shift in the MN distribution, it can be shown that $I_{\text{KL}} = \delta^2$. From Table 3.1, it appears that for the performance of the window-restricted GLR chart to be almost identical to the performance of the GLR chart without the window for detecting shifts of size $\delta = 0.2$, m has to be approximately 800, which is about 4.79 times the value of $\log 800/0.2^2$. Then for shifts of size 0.4, 0.6, 0.8, and 1, the necessary window size should be about $4.79 * \log 800/0.4^2 \approx 200$, $4.79 * \log 800/0.6^2 \approx 89$, $4.79 * \log 800/0.8^2 \approx 50$, and $4.79 * \log 800/1^2 \approx 32$, respectively. These values seem to agree with what we would conclude from Table 3.1.

Finally, we would like to reiterate that the purpose of introducing this moving window restriction is to make the evaluation of the GLR chart with millions of simulation repetitions feasible. In a practical application there is only one run of the GLR chart, therefore the practitioner should use the GLR chart without the window to get the best performance.

3.5 Control Limit of the GLR Chart

When specifying the control limit of a control chart, people choose the value so that some desired in-control ATS can be obtained. In this section we provide results for finding the control limit value of the GLR chart for p=1, 2, ..., up to 30 dimensional processes. For each p, we simulate the in-control ATS values for a sequence of control limit values with a lower

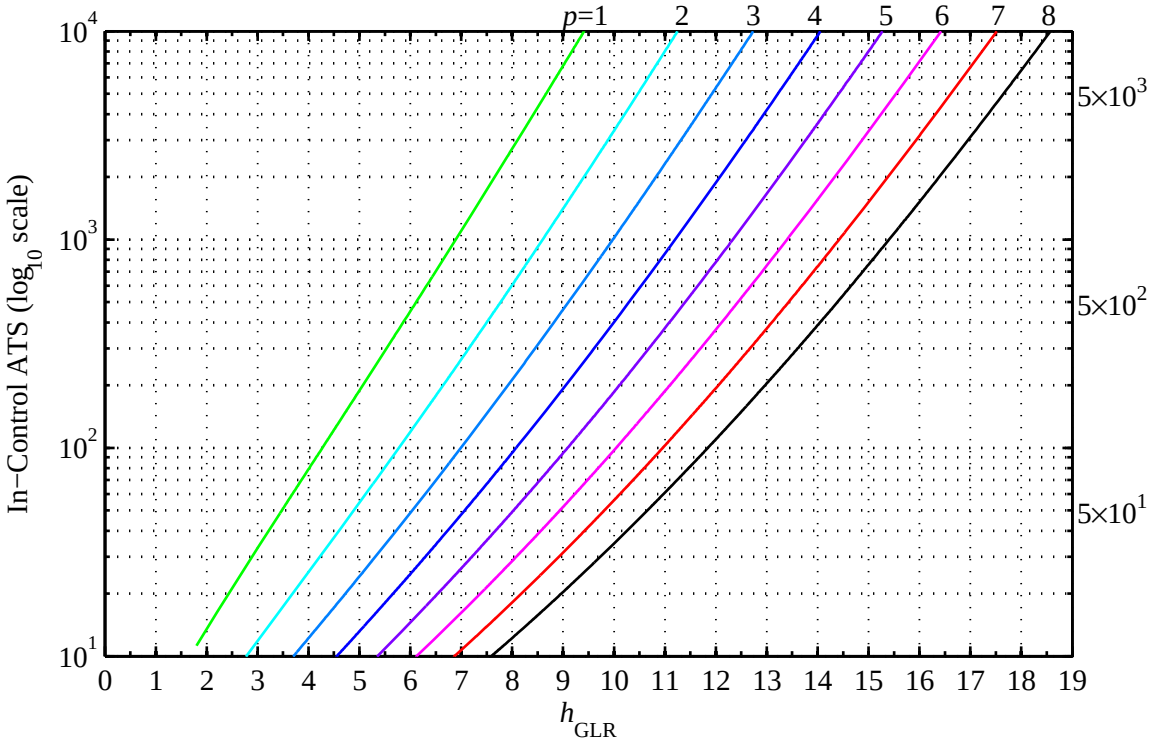
boundary whose in-control ATS is roughly 10, an upper boundary whose in-control ATS is at least 12000, and an increment size of 0.01.

Note that for each dimensionality p, a simulation run with 1 million iterations for the GLR chart with a window size $m = 600$ to give an in-control ATS of about 15000 takes approximately 3 days (using Matlab 2009b on a desktop with Dual-socket quad-core Intel Nehalem processors), whereas using the window size $m = 12000$ will take about 10 times longer. However, as was shown in Table 3.2 in the previous section, for a very large control limit the in-control ATS of the GLR chart with $m = 600$ is still acceptably close to that of the GLR chart with $m = 12000$. Therefore in simulations of these in-control ATS values, the GLR chart with window size $m = 600$ was used for all h_{GLR} and p. In practice when an extremely large in-control ATS is of interest (say 10000), the user sould also use a large window size (say 5000) so that very small shifts can be detected slightly faster. The control limit that we suggest here for a GLR chart with window size $m = 600$ to satisfy that desired false alarm rate will not give an in-control ATS of exactly 10000 when used with the GLR chart with window size $m = 5000$, but the error would be so small that it can be ignored in practical applications.

In Figure 3.1, we plot the in-control ATS values on a base 10 log scale against the control limits h_{GLR} for p=1, 2, ..., 8. Results for higher dimensional processes with $p > 8$ follow the same pattern, but we only plot $p \leq 8$ to show better details. It turns out that the relationships between $\log_{10}$(in-control ATS) and the control limit h_{GLR} are approximately linear for all dimensions. This property is very useful for practical applications, because it allows for the use of linear interpolation to obtain the control limit corresponding to any specific in-control ATS, as long as this ATS value is within the available range.

However, tabulating all in-control ATS values at each evaluation point of the control limit for all dimension could be quite messy. As an alternative to linear interpolation, a more straightforward approach is to use a polynomial function of the base 10 log of the in-control ATS to approximate the corresponding control limit. Since the relationship be-

Figure 3.1: Plot of $\log_{10}$(In-Control ATS) versus h_{GLR}



tween the control limit h_{GLR} and $\log_{10}$(in-control ATS) is almost linear, a cubic polynomial function is sufficient to model reasonably accurate control limits. The fitted functions of

$\log_{10}$(in-control ATS) for the control limit h_{GLR} are given as follows:

$$h_{\text{GLR}} = \begin{cases} -1.146630 + 2.747351 * L - 0.010303 * L^2 - 0.004151 * L^3 & \text{if } p = 1 \\ -0.596310 + 3.482806 * L - 0.165768 * L^2 + 0.008854 * L^3 & \text{if } p = 2 \\ 0.003872 + 3.923609 * L - 0.243645 * L^2 + 0.014615 * L^3 & \text{if } p = 3 \\ 0.481699 + 4.389605 * L - 0.342118 * L^2 + 0.023314 * L^3 & \text{if } p = 4 \\ 0.964141 + 4.786985 * L - 0.422579 * L^2 + 0.030090 * L^3 & \text{if } p = 5 \\ 1.542762 + 5.037944 * L - 0.459168 * L^2 + 0.032459 * L^3 & \text{if } p = 6 \\ 2.028680 + 5.356360 * L - 0.521003 * L^2 + 0.037537 * L^3 & \text{if } p = 7 \\ 2.533318 + 5.635085 * L - 0.574358 * L^2 + 0.042067 * L^3 & \text{if } p = 8 \end{cases} \tag{3.1}$$

where $L = \log_{10}$(desired in-control ATS).

With the provided equations, the practitioner can easily calculate the control limit corresponding to any desired false alarm rate. For example, suppose we have a production process with $p = 3$ quality variables, and one item from the process is inspected every hour. For a desired in-control ATS of 1200, which corresponds to on average 1 false alarm per $1200/24 = 50$ days, the control limit of the GLR control chart can be calculated as: $h_{\text{GLR}} = 0.003872 + 3.923609 * \log_{10}(1200) - 0.243645 * \log_{10}(1200)^2 + 0.014615 * \log_{10}(1200)^3 = 10.2020$. For the validation purpose, a simulation run with 1 million iterations was carried out for the GLR control chart with control limit $h_{\text{GLR}} = 10.2020$ and window size $m = 600$ for $p = 3$ dimensions, and the simulated in-control ATS is 1198.34, which is extremely close to the user specified value of 1200. Even for GLR charts with very large window sizes, the provided equations work quite well; for example, the simulated in-control ATS of the GLR chart with $m = 12000$ for a process with $p = 5$ is 11974.03 when using the control limit calculated for a desired in-control ATS of 12000.

In Table 3.3, we summarize the fitted coefficients of modeling h_{GLR} as a 3^{rd} order

Table 3.3: Coefficients for Approximating the Control Limit (p=1,...,30)

$h_{\mathrm{GLR}} = b_0 + b_1 L + b_2 L^2 + b_3 L^3$

p	b_0	b_1	b_2	b_3
1	-1.146630	2.747351	-0.010303	-0.004151
2	-0.596310	3.482806	-0.165768	0.008854
3	0.003872	3.923609	-0.243645	0.014615
4	0.481699	4.389605	-0.342118	0.023314
5	0.964141	4.786985	-0.422579	0.030090
6	1.542762	5.037944	-0.459168	0.032459
7	2.028680	5.356360	-0.521003	0.037537
8	2.533318	5.635085	-0.574358	0.042067
9	2.885007	6.062278	-0.682014	0.052750
10	3.511159	6.169922	-0.678480	0.050852
11	3.934768	6.482659	-0.748144	0.057226
12	4.495027	6.632962	-0.763477	0.057766
13	4.942616	6.907536	-0.825718	0.063686
14	5.468849	7.089404	-0.857188	0.066242
15	6.009139	7.240899	-0.876229	0.067119
16	6.591087	7.329358	-0.872083	0.065277
17	6.962787	7.659058	-0.957590	0.073940
18	7.388556	7.922730	-1.022227	0.080456
19	7.821077	8.166863	-1.077721	0.085678
20	8.331837	8.329834	-1.108168	0.088312
21	8.874439	8.444377	-1.120913	0.088973
22	9.280314	8.715785	-1.192093	0.096431
23	9.852142	8.783807	-1.184984	0.094246
24	10.359749	8.923336	-1.207162	0.095861
25	10.801726	9.134351	-1.255272	0.100432
26	11.322027	9.266821	-1.280327	0.102813
27	11.890537	9.327733	-1.274931	0.101177
28	12.398875	9.466466	-1.300497	0.103341
29	13.001436	9.480035	-1.278705	0.099845
30	13.487674	9.638554	-1.313540	0.103280

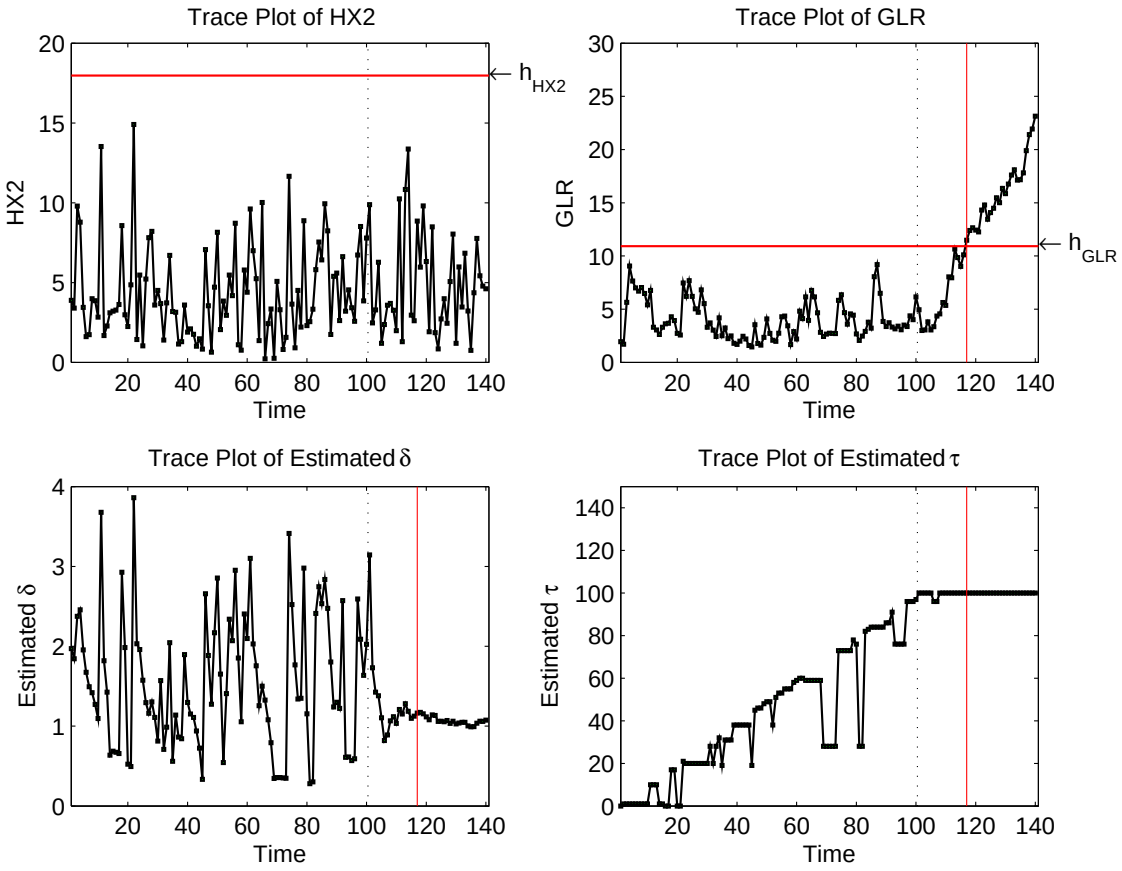
polynomial function of L for processes with up to 30 quality variables. Note that for each p there are about 1000 pairs of h_{GLR} and L values, the minimum coefficient of determination (model R^2) of all 30 models is greater than 0.9999, and the largest residual standard error is 0.0063. These results should cover most cases encountered in applications. If a user has a process with the number of quality variables p greater than 30, the control limit for a specific in-control ATS would need to be evaluated using simulation.

3.6 Plotting the GLR Chart

To show what the multivariate GLR control chart will look like, a set of GLR example plots using simulated observations are presented here. In this example, suppose we have a process with $p = 4$ quality variables, and the in-control mean vector is $\boldsymbol{\mu}_0 = [0\ 0\ 0\ 0]'$ and the covariance matrix $\boldsymbol{\Sigma}_0 = \boldsymbol{I}$. To demonstrate the pattern of the GLR chart while the process is in control, 100 in-control samples were first generated from the $\text{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ distribution. Suppose that some incident occurs at τ_c between the 100^{th} and 101^{th} sampling time points, causing the process mean to shift to the out-of-control mean vector $\boldsymbol{\mu}_1 = [-0.0696\ -0.4086\ -0.6768\ 0.6084]'$. The size of shift in this example is $\delta = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} = 1$. To illustrate the GLR control chart under the out-of-control situation, 40 out-of-control samples were then generated from the $\text{MN}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0)$ distribution.

In Figure 3.2, we present plots of the Hotelling χ^2 statistic (denoted by HX2), the GLR statistic, the estimated shift size $\hat{\delta}$, and the estimated change-point $\hat{\tau}$ from the GLR chart. The HX2 chart is presented because it not only can be used as a comparison to the GLR chart, but also provides the distance that measures how far each observation deviates from the in-control mean. In the trace plots of HX2 and GLR, the solid horizontal lines are the control limits corresponding to an in-control ATS of 800, and the dotted vertical lines near time 100 indicate the actual change-point. For the first 100 in-control observations, neither the HX2 chart nor the GLR chart had any false alarms. When the process was in control,

Figure 3.2: GLR Plot Example

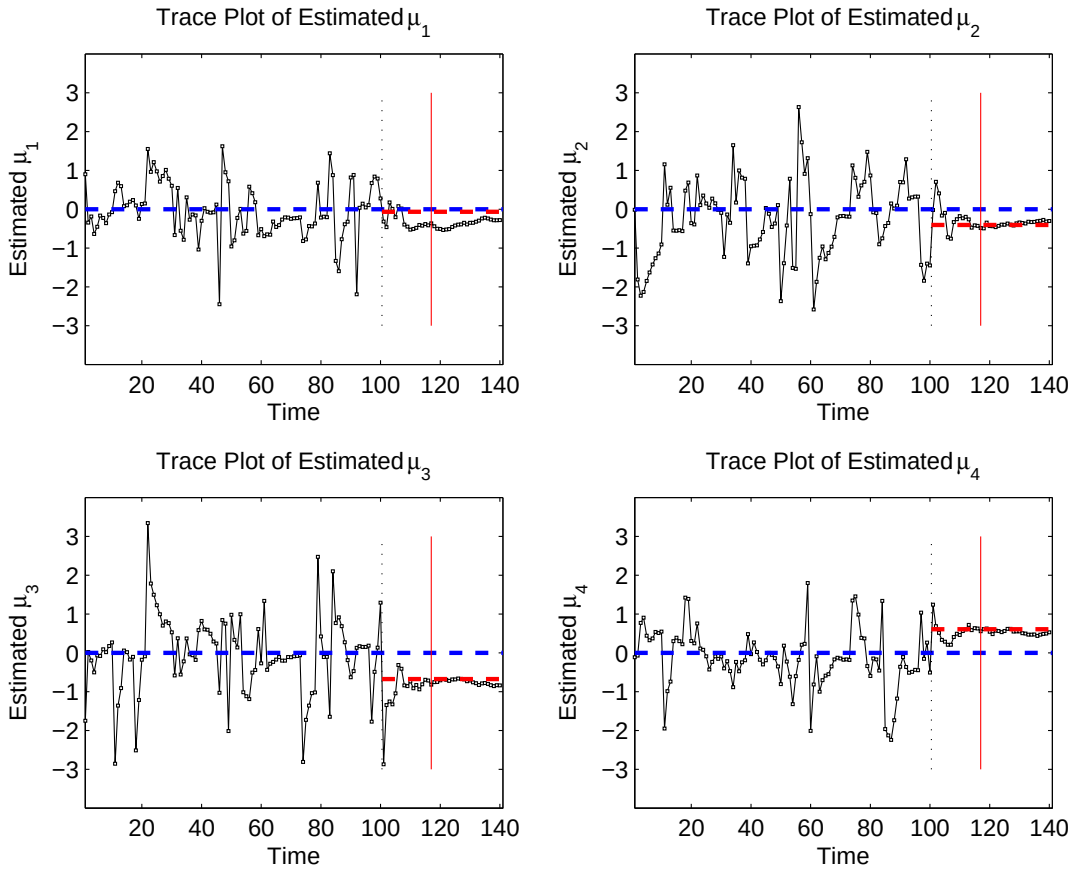


both of the control chart statistics varied randomly. When the process went out of control, as the number of out-of-control samples that had been observed increased, the GLR statistic showed an increasing trend, whereas the HX2 chart did not follow the same trend due to its inability to accumulate information from past observations. Eventually the GLR chart gave the first signal at time $k = 117$ with $R_{117} = 11.4865$, while the HX2 chart failed to detect the shift within 40 time units after the shift occurred. In the GLR chart, the ML estimates of the shift magnitude and change-point are immediately available after the control chart statistic R_k is calculated. In addition to plotting the GLR statistic R_k after each sample being taken, the estimated shift size $\hat{\delta}$ and change-point $\hat{\tau}$ can be plotted to aid the user to

diagnose the process if a signal is given. Solid vertical lines at $k = 117$ are drawn on these plots to indicate the time when the GLR chart gave the first signal. Note the differences between the trace plots in the example when the process was in control and those when the process was out of control. For sampling times $k \leq 100$, $\hat{\delta}$ and $\hat{\tau}$ exhibited a large amount of variation from time to time, and $\hat{\tau}$ tended to get close to the sampling time k. This suggests that if the process is in control, the likelihood ratio is usually maximized by a few 'large' (in terms of the Mahalanobis distance from the sample to $\boldsymbol{\mu}_0$) recent samples. When the process goes out of control, the estimation of $\hat{\boldsymbol{\mu}}_1$ will eventually converge toward the true $\boldsymbol{\mu}_1$, but the estimation error at the time of the signal depends on the number of out-of-control samples being used in the estimation. Therefore, it might be better to suppress the plots of these estimates until a signal is triggered, so that users unfamiliar with the GLR chart will not be misled.

In this example the GLR chart signaled at $k = 117$ when 17 out-of-control samples had been inspected, the estimated change-point was $\hat{\tau} = 100$, the estimated shift size was $\hat{\delta} = 1.1625$, and the estimated mean vector was $\hat{\boldsymbol{\mu}}_{1,k=117,\hat{\tau}=100} = [-0.3633 \ -0.4843 \ -0.8223 \ 0.5556]'$. Here $\hat{\delta}$ is just the 1-dimensional representation of the mean vector estimates. After a signal is given, the trace plots of $\hat{\mu}_{11}, \hat{\mu}_{12}, \ldots, \hat{\mu}_{1p}$ can be useful for the user to identify possible assignable causes that have triggered the signal. In Figure 3.3, we present such plots for this example. The longer dashed horizontal line at 0 represents the in-control mean value; the shorter dashed horizontal line between time 100 and 140 represent the true out-of-control mean value for each dimension; and the solid vertical line at time 117 marks the time when the GLR chart gave the first signal. We can see that except for $\hat{\mu}_{11}$, $\hat{\boldsymbol{\mu}}_1$ is very close to the true mean $\boldsymbol{\mu}_1$. When the user sees this figure after the signal at time $k = 117$, he/she can conclude that there are probably some decreases in the means of X_1, X_2 and X_3 and an increase in the mean of X_4, and the change has probably occurred between time 100 and 101. This information might be helpful for the user to identify and remove the assignable causes.

Figure 3.3: Plots of Estimated Mean Vector



Chapter 4

Performance Comparisons of Control Charts for Monitoring μ

In this chapter, we first compare the performance of the GLR chart to the performance of the traditional Hotelling's χ^2 chart and MEWMA charts with different choices of the tuning parameter λ. Then we compare the GLR chart to multi-MEWMA charts that use a combination consisting of 2 or 3 MEWMA charts with different weights and separate control limits run simultaneously. Results show that the GLR chart gives the best overall performance for detecting a wide range of shift sizes. The MEWMA chart is only effective for a small particular region of shift magnitudes. The multi-MEWMA chart with 3 different weights can give performance similar to the GLR chart, but the choice of weights and control limits requires sophisticated study.

The simulation setup in this chapter is identical to that used in Section 3.4 where the effect of the window size for the GLR chart was studied. Recall that when simulating the SSATS, at each iteration $\tau = 400$ in-control samples are first generated to allow the control chart statistics to reach their steady state, then at a random time $\tau_c \sim \text{Uniform}(400, 401)$ the mean vector changes from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$, where the size of the change is δ. The SSATS of a control chart at the shift size δ is then the average length of time from τ_c until the signal.

Results shown in this chapter are for the process with $p = 4$, but similar patterns hold for other dimensions.

4.1 The GLR Chart, the Hotelling χ^2 Chart and MEWMA Charts

In this section, the performance of the GLR chart is compared to the Hotelling χ^2 chart and the MEWMA chart. Table 4.1 gives the SSATS values of these control charts. The column labeled [1] corresponds to the GLR chart with window size $m = 600$, columns labeled [2]-[6] correspond to MEWMA charts with values of the tuning parameter $\lambda = 0.025$, 0.05, 0.1, 0.2 and 0.4 respectively, and column [7] corresponds to the Hotelling χ^2 chart (or equivalently the MEWMA chart with $\lambda = 1$).

First compare the performance of the MEWMA charts in columns [2]-[7]. The effect of the weighting parameter on the performance of the MEWMA chart has been well studied: a large λ allocates most of the weights to recent samples, resulting in good performance for large shifts; on the other hand, a small λ allocates relatively larger weights to past samples, resulting in good performance for small shifts. Results in the table support these conclusions. For small shift magnitudes, say $\delta < 1$, the SSATS values of MEWMA charts in columns [2] and [3] with small λ are smaller than those with larger λ; for large shift sizes where $\delta > 2.5$, the MEWMA charts with large λ in columns [6] and [7] have the best performance. For each MEWMA chart, there is a small range of shift sizes where its performance is optimal. As λ increases from 0.025 to 1, the optimal region of the MEWMA chart moves from small δ to large δ. For shift sizes at which the MEWMA chart is not optimal, the performance can be very poor. For example, at $\delta = 0.2$ the SSATS values of the MEWMA chart with λ =0.4 and 1 are 603.05 and 739.34, respectively, while the SSATS given by the MEWMA chart with $\lambda = 0.025$ is only 184.30.

Table 4.1: SSATS for the GLR Chart, the Hotelling χ^2 Chart and MEWMA Chart ($p = 4$)

	GLR	MEWMA					HX2
$\lambda =$	-	0.025	0.05	0.1	0.2	0.4	1
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]
0.0	800.00	799.06	799.99	799.89	799.21	799.79	800.00
0.2	247.49	184.30	245.92	347.60	475.11	603.05	739.34
0.4	79.95	61.27	69.77	100.05	173.43	312.47	592.19
0.6	39.15	34.24	33.61	40.22	66.70	143.35	428.52
0.8	23.41	23.53	21.28	22.12	31.21	67.85	291.91
1.0	15.66	17.88	15.47	14.75	17.69	34.81	190.31
1.2	11.27	14.40	12.09	10.89	11.68	19.93	123.78
1.4	8.53	12.06	9.96	8.61	8.51	12.44	80.25
1.6	6.70	10.35	8.43	7.12	6.63	8.50	52.79
1.8	5.41	9.07	7.30	6.06	5.39	6.24	34.70
2.0	4.46	8.06	6.44	5.25	4.53	4.83	23.33
2.5	2.96	6.32	4.96	3.94	3.23	2.99	9.46
3.0	2.11	5.19	4.04	3.16	2.50	2.13	4.33
4.0	1.22	3.82	2.95	2.26	1.72	1.34	1.31
5.0	0.76	3.03	2.32	1.76	1.33	0.95	0.66
8.0	0.50	1.86	1.42	1.07	0.66	0.51	0.50
12.0	0.50	1.24	0.90	0.56	0.50	0.50	0.50
$h =$	10.9122	13.5854	15.1256	16.3752	17.2855	17.8032	17.9715

Now compare the GLR chart in column [1] to the MEWMA charts. It can be seen that, for a wide range of shift sizes where $\delta \geq 1$, the GLR chart is either optimal or very close to optimal in terms of the SSATS value. When the shift sizes δ are small, the SSATS values of the GLR chart are slightly larger than those of the MEWMA charts with small λ, but they are still better than many of the other MEWMA charts. For example when $\delta = 0.2$, the smallest SSATS of 184.30 is given by the MEWMA chart with $\delta = 0.025$, while the SSATS of the GLR chart is 248.37; but the GLR chart still signals much faster than the MEWMA charts in columns [4]-[7].

It can be concluded that the GLR chart has the best overall performance in detecting a wide range of shift sizes. For a process whose out-of-control shift size is known and constant, the MEWMA chart can be specifically tuned to quickly detect this shift. But in applications,

it is more likely that the size of shift will be unknown and random, and then the GLR chart will give better performance.

4.2 The GLR Chart and Multi-MEWMA Charts

In this section, we compare the performance of the GLR chart to the performance of multi-MEWMA charts with the number of individual MEWMA components being 2 or 3. SSATS values for the GLR chart and 2-MEWMA combinations are given in Table 4.2. SSATS values for the single MEWMA chart with $\lambda = 0.025$ are also included in this table for easier comparisons. Column [1] corresponds to the GLR chart with window size $m = 600$, column [2] corresponds to the MEWMA chart with $\lambda = 0.025$, and columns [3]-[8] correspond to 2-MEWMA combinations with the values of tuning parameter specified in rows labeled λ_1 and λ_2.

Using 2 different MEWMA control charts simultaneously will impair the performance at shifts that the individual MEWWA charts are specifically tuned to detect, but will gain significant improvement over a large range of other shift magnitudes. For example, compare the SSATS values of the individual MEWMA chart in column [2] and the 2-MEWMA chart in column [3]. The MEWMA chart with $\lambda = 0.025$ has the smallest time to signal at very small shift sizes (δ =0.2 or 0.4), but its performance is bad for large shift sizes. When used together with the MEWMA chart with $\lambda = 0.1$, the performance of the 2-MEWMA combination for small shifts (δ =0.2 or 0.4) is compromised a little bit, but the SSATS values at all other shifts are greatly reduced. By comparing any 2-MEWMA combination to the single MEWMA chart that makes up the combination, it can be concluded that the use of two MEWMA charts simultaneously can significantly improve the performance in detecting a wide range of shift sizes at the cost of minor deteriorations only for a small region of shift magnitudes.

But when comparing all of the 2-MEWMA combinations, it appears that there are

Table 4.2: SSATS for the GLR Chart and 2-MEWMA Combinations ($p = 4$)

	GLR	MEWMA	2-MEWMA					
$\lambda_1 =$	-	0.025	0.025	0.025	0.025	0.1	0.1	0.2
$\lambda_2 =$	-	-	0.1	0.2	1	0.2	1	1
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
0.0	800.00	799.06	800.23	799.64	800.26	800.48	800.53	800.03
0.2	247.49	184.30	211.54	221.29	231.79	383.28	300.22	456.82
0.4	79.95	61.27	63.62	65.76	69.06	108.64	97.08	172.88
0.6	39.15	34.24	33.86	35.16	37.17	43.12	40.58	69.44
0.8	23.41	23.53	21.65	22.93	25.08	23.32	22.60	32.82
1.0	15.66	17.88	15.26	16.09	18.89	15.04	15.02	18.66
1.2	11.27	14.40	11.46	11.80	15.11	10.82	11.07	12.22
1.4	8.53	12.06	9.09	8.96	12.55	8.33	8.72	8.81
1.6	6.70	10.35	7.48	7.08	10.66	6.69	7.10	6.81
1.8	5.41	9.07	6.36	5.78	9.16	5.53	5.97	5.52
2.0	4.46	8.06	5.52	4.86	7.92	4.68	5.07	4.59
2.5	2.96	6.32	4.15	3.44	5.53	3.35	3.54	3.13
3.0	2.11	5.19	3.31	2.66	3.72	2.60	2.48	2.26
4.0	1.22	3.82	2.36	1.82	1.50	1.78	1.20	1.21
5.0	0.76	3.03	1.84	1.40	0.72	1.37	0.66	0.70
8.0	0.50	1.86	1.12	0.71	0.50	0.69	0.50	0.50
12.0	0.50	1.24	0.58	0.50	0.50	0.50	0.50	0.50
$h_1 =$	10.9122	13.5854	14.8917	15.2115	15.4704	17.2391	18.0537	18.8187
$h_2 =$	-	-	17.5344	18.6587	19.4641	18.1118	19.4500	19.4216

still significant ranges of shift sizes where each 2-MEWMA chart is relatively ineffective. For example, the 2-MEWMA chart in column [5] works well for small and large shifts, but it is apparently worse than the other combinations for medium shifts where $1 \leq \delta \leq 3$. It seems that using a small and an intermediate λ should be preferable to using two close values of λ. The GLR chart does not give the best performance at all shift magnitudes, but its SSATS values are very close to the best. We conclude that the GLR chart is better than the 2-MEWMA chart for detecting a wide range of shift sizes.

In Table 4.3, SSATS values for the GLR chart and some 3-MEWMA combinations are provided. The SSATS values for three of the 2-MEWMA combinations from Table 4.2 are

also included in this table for easier comparisons. Column [1] corresponds to the GLR chart with window size $m = 600$, columns [2]-[4] correspond to the 2-MEWMA combinations, and columns [5]-[7] correspond to the 3-MEWMA combinations with values of tuning parameter specified in rows λ_1, λ_2 and λ_3.

Table 4.3: SSATS for the GLR Chart, 2-MEWMA and 3-MEWMA Combinations ($p = 4$)

	GLR	2-MEWMA			3-MEWMA		
$\lambda_1 =$	-	0.025	0.025	0.1	0.025	0.025	0.1
$\lambda_2 =$	-	0.1	0.2	0.2	0.1	0.1	0.2
$\lambda_3 =$	-	-	-	-	0.2	1	1
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]
0.0	800.00	800.23	799.64	800.48	800.64	799.68	799.87
0.2	247.49	211.54	221.29	383.28	228.89	249.02	454.16
0.4	79.95	63.62	65.76	108.64	65.58	69.64	127.69
0.6	39.15	33.86	35.16	43.12	34.67	36.08	48.34
0.8	23.41	21.65	22.93	23.32	22.07	22.93	25.40
1.0	15.66	15.26	16.09	15.04	15.27	15.95	16.05
1.2	11.27	11.46	11.80	10.82	11.25	12.00	11.50
1.4	8.53	9.09	8.96	8.33	8.70	9.44	8.79
1.6	6.70	7.48	7.08	6.69	6.99	7.73	7.03
1.8	5.41	6.36	5.78	5.53	5.78	6.50	5.79
2.0	4.46	5.52	4.86	4.68	4.90	5.57	4.85
2.5	2.96	4.15	3.44	3.35	3.48	3.95	3.33
3.0	2.11	3.31	2.66	2.60	2.69	2.86	2.40
4.0	1.22	2.36	1.82	1.78	1.84	1.42	1.27
5.0	0.76	1.84	1.40	1.37	1.41	0.74	0.71
8.0	0.50	1.12	0.71	0.69	0.72	0.50	0.50
12.0	0.50	0.58	0.50	0.50	0.50	0.50	0.50
$h_1 =$	10.9122	14.8917	15.2115	17.2391	15.5598	16.1987	18.4874
$h_2 =$	-	17.5344	18.6587	18.1118	18.1234	18.7054	19.3198
$h_3 =$	-	-	-	-	18.9635	20.0624	19.8845

Without too much surprise, by comparing the 3-MEWMA chart in column [5] to those 2-MEWMA combinations in columns [2]-[4] that are composed of subsets of the charts in the former, it can be observed that the performance of the 3-MEWMA combination is further balanced over a broader range of shift sizes, with only slight decreases in performance for a

small number of shifts. The comparison for all three 3-MEWMA combinations suggests again that when using the multi-MEWMA scheme, the values of the tuning parameters should be remarkably different from each other so that the effective regions of all the charts cover as wide a range of shift magnitudes as possible. When comparing the 3-MEWMA combinations in column [5] and [6] to the GLR chart, it is hard to say which one outperforms the other two; there are regions where one is better and also places where it is worse. But in general it can be concluded that the 3-MEWMA scheme can be designed to give similar performance to the GLR chart.

It is unlikely that we can find a control chart that uniformly outperforms all the other existing control charts. In practice, the user can use a single MEWMA chart if only a particular small range of shift sizes is likely to occur in the process. If fast detection of a wide range of shifts is of interest, the user could use either the 3-MEWMA scheme or the GLR chart. However the multi-MEWMA approach requires sophisticated determination of values for the weighting parameter λ and the set of appropriate control limits for each individual component so that one has the desired global in-control ATS. Whereas on the other hand, for the GLR chart all the user has to do is to use the formula provided in section 3.5 to calculate the control limit for a desired in-control ATS.

Chapter 5

Relationship between the GLR Chart and the MCUSUM Chart

5.1 Equivalence of the GLR Chart to an Infinite Set of MCUSUM Charts

In their paper on monitoring the mean of a univariate process, Reynolds and Lou (2010) discussed the equivalence of the GLR chart to a set consisting of an infinite number of CUSUM charts. Here in this chapter we check whether this equivalence still holds for multivariate normal processes. There are many different forms of multivariate CUSUM charts that have been developed and discussed. Here we consider the Healy (1987) MCUSUM chart which is derived directly from a set of SPRTs for testing a shift in $\boldsymbol{\mu}$ from the in-control $\boldsymbol{\mu}_0$ to a specific out-of-control value $\boldsymbol{\mu}_c$.

The preliminary form of the MCUSUM statistic at time k is the difference between the accumulated log-likelihood ratio of samples by time k and the minimum of the accumulated

log-likelihood ratio of samples by time $j \leq k$:

$$S_k = \sum_{i=1}^{k} \log \frac{f(\boldsymbol{X}_i|\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_0)}{f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)} - \min_{j \leq k} \sum_{i=1}^{j} \log \frac{f(\boldsymbol{X}_i|\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_0)}{f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)},$$

or equivalently

$$S_k = \max_{1 \leq j \leq k} \sum_{i=j}^{k} \log \frac{f(\boldsymbol{X}_i|\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_0)}{f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)}.$$

The MCUSUM chart signals if S_k is greater than a pre-specified control limit. For observation $\boldsymbol{X}_i$, the increment that the MCUSUM accumulates is

$$\begin{aligned} \log \frac{f(\boldsymbol{X}_i|\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_0)}{f(\boldsymbol{X}_i|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)} &= -\frac{1}{2}[(\boldsymbol{X}_i - \boldsymbol{\mu}_c)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_c) + (\boldsymbol{X}_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0)] \\ &= (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0). \end{aligned}$$

Then the MCUSUM statistic becomes

$$\begin{aligned} S_k &= \max_{1 \leq j \leq k} \sum_{i=j}^{k} \left\{ (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \right\} \\ &= \max_{0 \leq t < k} \sum_{i=t+1}^{k} \left\{ (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \right\}. \end{aligned}$$

In Section 3.1 the GLR statistic R_k is presented in a simplified form that is easy to understand. It can be easily shown that

$$\begin{aligned} R_k &= \max_{0 \leq t < k} \frac{k-t}{2} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0) \\ &= \max_{0 \leq t < k} \sum_{i=t+1}^{k} \left\{ (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0) \right\}. \end{aligned}$$

Now it becomes obvious that the GLR chart and the MCUSUM chart are only different in that the hypothesized out-of-control mean in the MCUSUM chart is the user specified $\boldsymbol{\mu}_c$, whereas in the GLR chart the hypothesized out-of-control mean is estimated from the

observed data, which can take any possible value in the parameter space. Therefore the GLR chart is equivalent to an infinite set of MCUSUM charts with $\boldsymbol{\mu}_c$ spanning over the space of $\boldsymbol{\mu}$.

In applications of the MCUSUM chart, the recursive formula of R_k is more commonly used: $S_k = \max(S_{k-1}, 0) + (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{X}_k - \boldsymbol{\mu}_0) - \frac{1}{2}(\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)$. Unlike the GLR chart, the performance of the MCUSUM chart depends on the direction of the shift. It only works well if the out-of-control mean $\boldsymbol{\mu}_1$ shifts in the same direction as the specified $\boldsymbol{\mu}_c$.

5.2 Approximation of the GLR Chart with a Finite Set of MCUSUM Charts

Reynolds and Lou (2010) also discussed the approximation of the univariate GLR chart using a finite set of CUSUM charts with carefully chosen values of the tuning parameter. They have shown that an upper bound exists for the value of the CUSUM chart tuning parameter when approximating the GLR chart. In the next paragraph, we show that this upper bound also exists for the multivariate case.

Given a sequence of observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$, the GLR chart without the moving window restriction signals at time k if there is a $t \in \{0, 1, \ldots, k-1\}$ satisfying

$$\frac{k-t}{2}(\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0) = \frac{k-t}{2}\hat{\delta}_{k,t}^2 > h_{\mathrm{GLR}},$$

where $\hat{\delta}_{k,t}$ is the estimated shift size using the $k-t$ most recent samples. The signal given by the GLR chart with $\hat{\delta}_{k,t} > \sqrt{\frac{2h_{\mathrm{GLR}}}{k-t}}$ is equivalent to the signal by the MCUSUM chart with

$\boldsymbol{\mu}_c = \hat{\boldsymbol{\mu}}_{1,k,t}$ such that

$$\sum_{i=t+1}^{k} \left\{ (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \right\} > h_{\mathrm{GLR}}.$$

Now consider another sequence of observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{X}^*_{t+1}, \ldots, \boldsymbol{X}^*_k$, such that $\hat{\boldsymbol{\mu}}^*_{1,k,t} = \frac{\sum_{i=t+1}^{k} \boldsymbol{X}^*_i}{k-t} = a(\hat{\boldsymbol{\mu}}_{1,k,t} - \boldsymbol{\mu}_0) + \boldsymbol{\mu}_0$, where $a > 1$. The GLR chart will also signal because $\hat{\delta}^*_{k,t} = a\hat{\delta}_{k,t} > \sqrt{\frac{2h_{\mathrm{GLR}}}{k-t}}$. For the MCUSUM chart with tuning parameter $\boldsymbol{\mu}_c = \hat{\boldsymbol{\mu}}_{1,k,t}$, we have

$$\begin{aligned}
&\sum_{i=t+1}^{k} \left\{ (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}^*_i - \boldsymbol{\mu}_0) - \frac{1}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \right\} \\
&= (k-t)(\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\hat{\boldsymbol{\mu}}^*_{1,k,t} - \boldsymbol{\mu}_0) - \frac{k-t}{2} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \\
&= (2a-1) \frac{k-t}{2} \hat{\delta}^2_{k,t} > h_{\mathrm{GLR}}.
\end{aligned}$$

This means that if the GLR chart signals because of some $\hat{\boldsymbol{\mu}}_{1,k,t}$ of size $\hat{\delta}_{k,t} > \sqrt{\frac{2h_{\mathrm{GLR}}}{k-t}}$, then the MCUSUM chart with $\boldsymbol{\mu}_c$ of size $\sqrt{\frac{2h_{\mathrm{GLR}}}{k-t}}$ in the same direction of $\hat{\boldsymbol{\mu}}_{1,k,t}$ will also signal. Therefore, for a given control limit h_{GLR}, the GLR chart is equivalent to the set of MCUSUM charts with $\boldsymbol{\mu}_c \in \left\{ \boldsymbol{\mu}_c : (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) = \frac{2h_{\mathrm{GLR}}}{k-\tau}, (k-t) = 1, 2, \ldots \right\}$. The upper bound for the necessary shift sizes for the set of MCUSUM charts is $\sqrt{2h_{\mathrm{GLR}}}$, which corresponds to the minimum $\hat{\delta}_{k,k-1}$ for a single observation to trigger a signal. For example if the control limit of the GLR chart is $h_{\mathrm{GLR}} = 8$, then it is sufficient to use a set of MCUSUM charts with $\boldsymbol{\mu}_c \in \left\{ \boldsymbol{\mu}_c : (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) \leq 16 \right\}$.

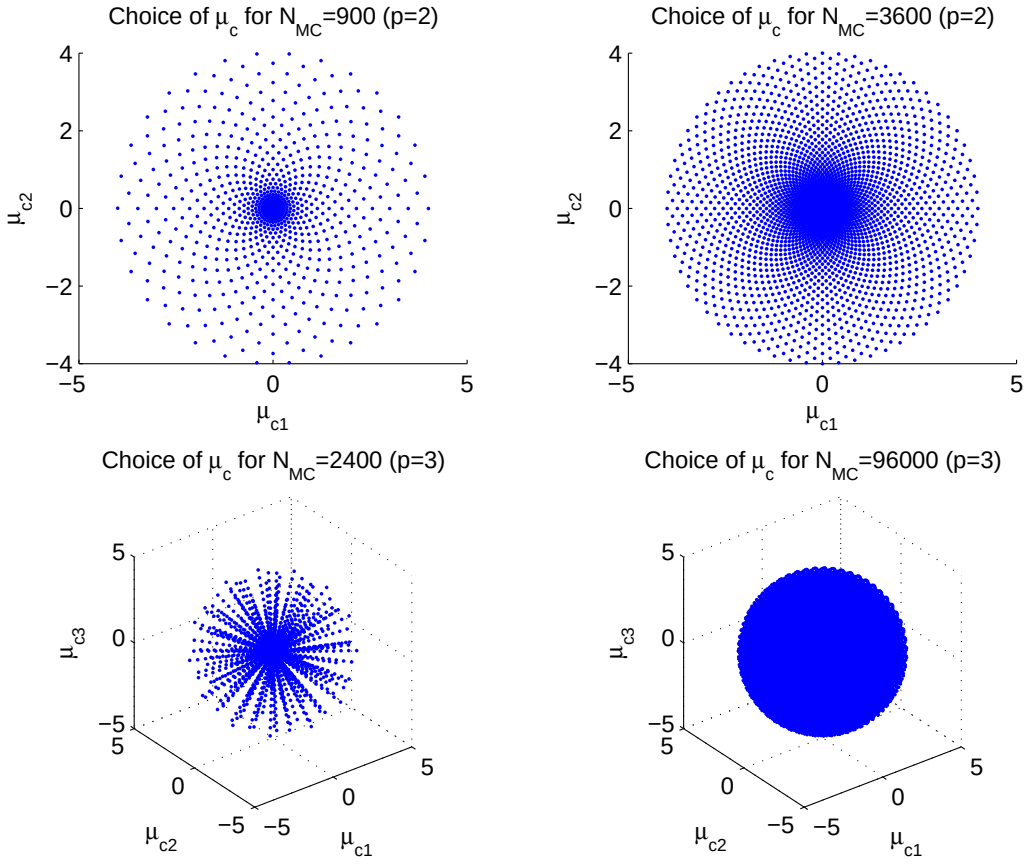
Although the range of shift sizes in the set of MCUSUM charts has been reduced, the number of directions of $\boldsymbol{\mu}_c$ is still infinite, which makes the approximation of the GLR chart with a set of finite MCUSUM charts difficult. As an example, in Table 5.1 we compare the ATS values for the GLR chart and 4 sets of a finite number of MCUSUM charts for $p = 2$ and $p = 3$. In this example assume that the in-control mean is $\mathbf{0}$ and the covariance matrix $\boldsymbol{\Sigma}_0$ is the identity matrix $\boldsymbol{I}$. The control limits h_{GLR} are arbitrarily chosen to be 8 for $p = 2$ and 9 for $p = 3$. Note that initial state ATS values are used here for out-of-control situations for

simplicity. Each set of MCUSUM charts consists of $N_{\text{MC}} = N_\delta * N_{\text{dir}}$ different values of tuning parameters $\boldsymbol{\mu}_c$, where N_δ is the number of different sizes of $\boldsymbol{\mu}_c$ (denoted as δ_c) and N_{dir} is the number of different directions of $\boldsymbol{\mu}_c$ at each δ_c. For $p = 2$, the set of MCUSUM charts in column [1] has $(N_\delta, N_{\text{dir}}) = (30, 30)$ and the other set in column [2] has $(N_\delta, N_{\text{dir}}) = (60, 60)$; for $p = 3$, these numbers are (40,60) and (80,1200) for the sets in columns [4] and [5]. When selecting the N_δ δ_c values for the set of MCUSUM charts, we let $\delta_{c,i} = \sqrt{2h_{\text{GLR}}}\frac{i^2}{N_\delta^2}, i = 1, \ldots, N_\delta$. For each $\delta_{c,i}$, the N_{dir} different values of $\boldsymbol{\mu}_c$ are selected evenly along the ellipse(for $p = 2$)/ellipsoid(for $p = 3$) defined by $\boldsymbol{\mu}_c : (\boldsymbol{\mu}_c - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\mu}_c - \boldsymbol{\mu}_0) = \delta_{c,i}$. In Figure 5.1 the scatter plots of $\boldsymbol{\mu}_c$ are given to help understand the choice of the tuning parameters in the sets of MCUSUM charts.

Table 5.1: ATS for GLR Charts and Sets of MCUSUM Charts

	$p = 2$			$p = 3$		
	MCUSUM	MCUSUM	GLR	MCUSUM	MCUSUM	GLR
$N_{\text{MC}} =$	900	3600	-	2400	96000	-
δ	[1]	[2]	[3]	[4]	[5]	[6]
0.0	617.6	605.7	602.9	539.7	466.1	459.6
0.2	200.8	199.7	197.6	206.2	191.1	189.4
0.4	67.7	67.4	67.3	72.4	69.2	68.9
0.6	33.6	33.5	33.5	36.3	35.1	35.0
0.8	20.4	20.3	20.3	22.1	21.4	21.4
1.0	13.8	13.8	13.8	14.9	14.5	14.5
1.2	10.1	10.1	10.0	10.9	10.6	10.6
1.4	7.7	7.7	7.7	8.3	8.2	8.1
1.6	6.2	6.2	6.2	6.6	6.5	6.5
1.8	5.1	5.1	5.0	5.4	5.3	5.3
2.0	4.3	4.3	4.2	4.6	4.5	4.5
2.5	3.0	3.0	3.0	3.2	3.1	3.1
3.0	2.3	2.2	2.2	2.4	2.4	2.4
4.0	1.5	1.5	1.5	1.6	1.6	1.5
5.0	1.1	1.1	1.1	1.2	1.2	1.2
8.0	1.0	1.0	1.0	1.0	1.0	1.0
12.0	1.0	1.0	1.0	1.0	1.0	1.0
$N_\delta =$	30	60	-	40	80	-
$N_{\text{dir}} =$	30	60	-	60	1200	-
$h_{\text{GLR}} =$	8	8	8	9	9	9

Figure 5.1: Scatter Plot of $\boldsymbol{\mu}_c$ in Table 5.1



Comparing the ATS values for the GLR charts and the sets of MCUSUM charts, it appears that the ATS values of the set of MCUSUM charts approach those of the GLR chart as the number of MCUSUM charts in the set increases. For $p = 2$: the ATS values of the set of 900 MCUSUM charts in column [1] are not very different from those of the GLR chart; if the number of values of $\boldsymbol{\mu}_c$ is increased to 3600, the ATS values do get closer to those of the GLR chart. For the higher dimension $p = 3$, the direction of a shift is in the three-dimensional space. Therefore for each δ_c, a large number of different $\boldsymbol{\mu}_c$ are necessary so that the performance of the set of MCUSUM charts can be close to that of the GLR chart. In the table, for most shift sizes δ, the ATS values of the set of MCUSUM charts

with $N_{\text{MC}} = 2400$ in column [4] are remarkably larger than those of the GLR chart. For the set of $N_{\text{MC}} = 96000$ MCUSUM charts in column [5], the performance is much closer to the performance of the GLR chart. But there are still notable differences in in-control ATS values and out-of-control ATS values for shifts of small size.

Reynolds and Lou (2010) mentioned the main advantage of approximating the GLR chart with a finite set of CUSUM charts, which is that the simulation speed can be much faster if using the set of CUSUM charts. In the multivariate situation, for processes with dimension $p = 2$, we can also benefit in terms of simulation speed from using a set of an adequate number of MCUSUM charts. But as the dimension p gets larger, the approximation quickly becomes infeasible because we will need an extremely large number of $\boldsymbol{\mu}_c$ values to span over the parameter space.

Chapter 6

Applying the GLR Chart with Variable Sampling Intervals

Prior to this chapter, it has been assumed that a standard fixed sampling interval (FSI) sampling scheme is being used. However as the technology advances, more and more computer-controlled automatic sampling instruments are available in many practical applications. Much work in the SPC literature has shown that the performance of control charts that are capable of accumulating information can be dramatically improved if used with variable sampling intervals (VSI) (see, for example, Reynolds (1995), Runger and Pignatiello (1991) and Reynolds and Cho (2011)). In this chapter we discuss the use of the GLR chart for monitoring $\boldsymbol{\mu}$ with VSI.

6.1 Definition of VSI

It has been previously assumed that a FSI sampling scheme is used, where independent samples of a constant sample size n are randomly collected from the process at evenly spaced sampling time points. Note that as long as the sample size is constant, $n = 1$ can still be

assumed without loss of generality. Use d to denote the duration between samples. Then $d = d_0$ constantly for the FSI scheme, and it can be assumed without loss of generality that $d_0 = 1$.

Now consider the VSI sampling scheme. The idea is to employ time-varying intervals between samples based on the most recent control charting statistic. If the statistic at the current sample shows some evidence of a possible shift even though not strong enough to produce a signal, then the next sample will be taken sooner. On the other hand, if the statistic shows little evidence of a shift, then one should wait a little bit longer before the next sample is taken. Think about the time before the next sample as a function of the current value of the charting statistic, then presumably this function can be as simple as a step function or as complicated as a continuous function. In Stoumbos, Mittenthal and Runger (2001), the authors discussed an efficient numerical optimization algorithm for constructing VSI control charts with continuous sampling interval functions to give the optimal SSATS performance. Their results also show that control charts with dual-sampling-interval can be designed to give performance comparable to those SSATS-optimal control charts with complicated sampling interval functions. For the ease of implementation, they recommended the dual-sampling-intervals scheme be used in practical applications.

Use d_{S} and d_{L} ($d_{\mathrm{S}} < d_{\mathrm{L}}$) to denote respectively the shorter interval and the longer interval. Recall that for the GLR chart, a larger value of the GLR statistic indicates stronger evidence that the process is out-of-control. Consider a warning threshold $w_{\mathrm{GLR}} < h_{\mathrm{GLR}}$. At sampling time point k, if $R_k > h_{\mathrm{GLR}}$ then a signal is triggered and the process is interrupted for the diagnosis and the removal of the cause of the signal; if $h_{\mathrm{GLR}} \geq R_k > w_{\mathrm{GLR}}$, then it indicates some evidance of an out-of-control situation therefore the $(k+1)^{\mathrm{th}}$ sample should be taken after d_{S} time units; if $R_k \leq w_{\mathrm{GLR}}$, then the next sample is taken after a time interval of d_{L}.

A fair assumption when comparing different sampling schemes is that the rate of the average number of items being sampled is held constant when the process is in control,

say being the same as that of the FSI scheme where $d_0 = 1$. For a given pair of h_{GLR} and w_{GLR}, denote the proportion of sample statistics falling within the warning zone as $\pi_{\mathrm{S}} = \mathrm{Prop}\,(R_k > w_{\mathrm{GLR}} | R_k \leq h_{\mathrm{GLR}})$ and that of those falling below the warning limit as $\pi_{\mathrm{L}} = \mathrm{Prop}\,(R_k \leq w_{\mathrm{GLR}} | R_k \leq h_{\mathrm{GLR}})$ when the process is in-control. Then it is easy to see that there are three constraints so that the constant sampling rate can be satisfied: $d_{\mathrm{S}} < 1$, $d_{\mathrm{L}} > 1$, and $d_{\mathrm{S}}\pi_{\mathrm{S}} + d_{\mathrm{L}}\pi_{\mathrm{L}} = d_0 = 1$.

The above assumptions and constraints mean that the control limit h_{GLR} for the GLR chart used VSI is the same as that with FSI, therefore the formulas in Equation 3.1 for obtaining the control limit corresponding to a desired in-control ATS can still be used. The values of d_{S} and d_{L} can be carefully choosen according to what performance the practitioner desires. It is recommended in previous research on control charts using VSI that d_{S} should be chosen as small as is feasible, while d_{L} should be chosen depending on the emphasis of shift sizes: a d_{L} slightly larger than 1 leads to faster detection of large shifts and a larger d_{L} is more effective for detecting smaller shift sizes.

6.2 A Comparison Study

In this section we compare the performances of the GLR chart, the MEWMA chart and the Hotelling χ^2 chart used with VSI. The main intention of this comparison is to provide the user with an impression of to what extent the VSI sampling scheme can boost the performance of the GLR chart, and how well it works when compared to other VSI charts.

In the simulation, we still use a $p = 4$ dimesional process and set control limits of the control charts to give in-control ATS values of 800. Note that the way the SSATS values are calculated for the VSI charts is slightly different from that used for the FSI results in Chapter 4. For the FSI sampling scheme, a sample is taken at each integer time point and the change-point τ_c when the process mean shifts from $\boldsymbol{\mu}_0$ to some $\boldsymbol{\mu}_1$ is assumed to be uniform on the interval between time 400 and 401. For the VSI sampling scheme, this assumption

is no longer appropiate because the sampling time points are random and there may not be a sample at time 400 or 401. Therefore we assume here that the change-point τ_c is uniform on the time interval between the 400^{th} and 401^{th} sampling time points, and the SSATS is the expected time to signal after τ_c. Regarding the choice of the VSI parameters, two set of sampling intervals are considered: $(d_{\text{S}}, d_{\text{L}}) = (0.1, 1.25)$ and $(0.1, 1.9)$.

Table 6.1: SSATS for the GLR Chart, the Hotelling χ^2 Chart and MEWMA Chart with VSI ($p = 4$)

	GLR			MEWMA				HX2	
$d_{\text{S}} =$	-	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$d_{\text{L}} =$	-	1.25	1.90	1.25	1.90	1.25	1.90	1.25	1.90
$\lambda =$	-	-	-	0.026	0.026	0.11989	0.11989	1	1
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
0.0	800.0	800.0	800.0	800.0	800.0	799.7	799.9	800.0	800.0
0.2	247.5	163.2	141.8	154.2	136.4	353.4	335.2	732.9	729.9
0.4	80.0	49.6	45.3	41.7	37.5	87.8	74.3	578.3	569.3
0.6	39.2	24.0	22.7	22.3	21.5	27.8	22.4	407.3	393.0
0.8	23.4	14.3	13.8	15.3	15.4	12.9	10.9	266.1	250.1
1.0	15.7	9.5	9.4	11.8	12.0	8.0	7.3	166.5	151.3
1.2	11.3	6.9	6.9	9.5	9.9	5.9	5.6	101.4	88.8
1.4	8.5	5.2	5.4	8.0	8.4	4.6	4.5	61.0	51.4
1.6	6.7	4.1	4.3	7.0	7.3	3.9	3.9	36.6	29.6
1.8	5.4	3.3	3.5	6.1	6.5	3.3	3.4	22.0	17.2
2.0	4.5	2.8	2.9	5.5	5.8	2.9	3.0	13.3	10.2
2.5	3.0	1.9	2.0	4.3	4.6	2.2	2.4	4.0	3.3
3.0	2.1	1.4	1.5	3.6	3.9	1.8	2.0	1.7	1.6
4.0	1.2	0.8	1.1	2.7	2.9	1.3	1.5	0.7	1.0
5.0	0.8	0.7	0.9	2.1	2.3	1.1	1.2	0.6	0.9
8.0	0.5	0.6	0.9	1.3	1.5	0.7	1.0	0.6	0.9
12.0	0.5	0.6	0.9	0.9	1.1	0.6	0.9	0.6	0.9
$h =$	10.9122	10.9122	10.9122	13.6858	13.6858	16.6546	16.6546	17.9715	17.9715
$w =$	-	5.9091	4.5416	5.3448	3.1329	5.6585	3.3050	5.7530	3.3527

In Table 6.1, SSATS values of these control charts used with the VSI scheme are presented. Rows d_{S} and d_{L} denote the value of the short interval and the long interval for

charts used with VSI. The last row of w is the value of the warning limit, such that together with the control limi h, d_S and d_L, the constraint $d_S\pi_S + d_L\pi_L = d_0 = 1$ is satisfied.

The column labeled [1] is the GLR chart used with the standard FSI sampling scheme, and columns [2] and [3] are the GLR chart used under the VSI scheme for with two sets of sampling interval parameters. It is easy to see that the VSI sampling has greatly improved the performance of the GLR chart except for detecting extremely large shift sizes. In conformance with the general conclusion about the effect of the choice of d_L, the GLR chart in column [3] with a larger d_L is better for detecting small and intermediate shifts, while column [2] is slightly better for large shifts. Columns [4]-[7] correspond to the MEWMA charts with λ =0.026 and 0.11989, and column [8]-[9] correspond to the the Hotelling χ^2 chart. These results are from Reynolds and Cho (2011). The same conclusions about the effect of the tuning parameter λ and the long interval d_L can be made. When compared to the MEWMA charts and the Hotelling χ^2 chart under the same VSI sampling parameters, the GLR chart provides a better overall performance for detecting a wide range of shift sizes.

6.3 The Design of the VSI GLR Chart

The GLR chart used with FSI sampling does not require the user to specify any tuning parameter values. However as shown in the previous section, if used with a VSI scheme, the performance of the GLR chart can be greatly improved. If the sampling instrument in the application is capable of using the VSI scheme, then we should make efficient use of it.

To apply the GLR chart for monitoring $\boldsymbol{\mu}$ of a MN process with a VSI scheme, there are 4 parameters that need to be specified: the desirable in-control ATS value (h_{GLR}), the shorter sampling interval d_S, the longer sampling interval d_L and the warning limit w_{GLR}. Note that d_0, the average number of items inspected per unit time, can always be assumed to be 1 by adjusting the time unit; when we talk about values of d_S and d_L, they in fact represent the short and long intervals as factors of d_0. Because of the constraint that

$d_S\pi_S + d_L\pi_L = d_0 = 1$, the control limit h_{GLR} for the VSI chart is the same as that for the FSI chart. Given a desirable in-control ATS, the control limit can be obtained using Equation 3.1 for $p = 1, \ldots, 8$ dimensional processes.

Reynolds and Cho (2011) evaluates two settings of the VSI parameters (d_S, d_L), (0.10, 1.25) for slightly better ability to detect large shifts and (0.10, 1.90) for faster detection of small and intermediate shifts. For some applications $d_S = 0.10$ is too small to be feasible, so the authors also consider another two settings (0.25, 1.25) and (0.25, 1.90). We also treat these four VSI settings as the standard options, and recommend general users to pick from these settings because we provide a method here to determine the warning limit w_{GLR} for these settings.

Once h_{GLR}, d_S and d_L are determined, the value of w_{GLR} is uniquely determined because $\pi_S = \Pr(R_k > w_{GLR} | R_k \leq h_{GLR}) = (d_L - 1)/(d_L - d_S)$. When presenting the results about how to find the warning limit, we describe the four VSI settings (d_S, d_L) =(0.10, 1.25), (0.25, 1.25), (0.10, 1.90) and (0.25, 1.90) in terms of π_S =0.2174, 0.25, 0.50, and 0.5455 respectively. Results provided in this section can also be applied to any other combinations of (d_S, d_L) yielding the same π_S.

For each $p \in \{1, \ldots, 8\}$, we first simulate the in-control ATS values for a sequence of h_{GLR} values with a lower boundary whose in-control ATS is roughly 10, an upper boundary whose in-control ATS is at least 12000 and an increment size of 0.01; then for each h_{GLR} value we scan over the interval $(0, h_{GLR})$ with a step size of 0.0001 to find the corresponding w_{GLR} values to to give π_S of 0.2174, 0.25, 0.50 and 0.5455 respectively. The number of simulation runs is one million for each value of p.

In Figures 6.1, 6.2, 6.3 and 6.4, the scatter-plots with connected lines of the warning limits and the control limits for the GLR chart versus the $\log_{10}$(In-Control ATS) values are presented for $p = 1, \ldots, 8$ dimensional processes under the four VSI settings. The user can visually find the approximate w_{GLR} from these figures based on the desired in-control ATS and VSI setting, and use w_{GLR} together with the control limit h_{GLR} obtained either from

Figure 6.1: Plot of $\log_{10}$(In-Control ATS) versus w_{GLR} for $\pi_{\mathrm{S}} = 0.2174$

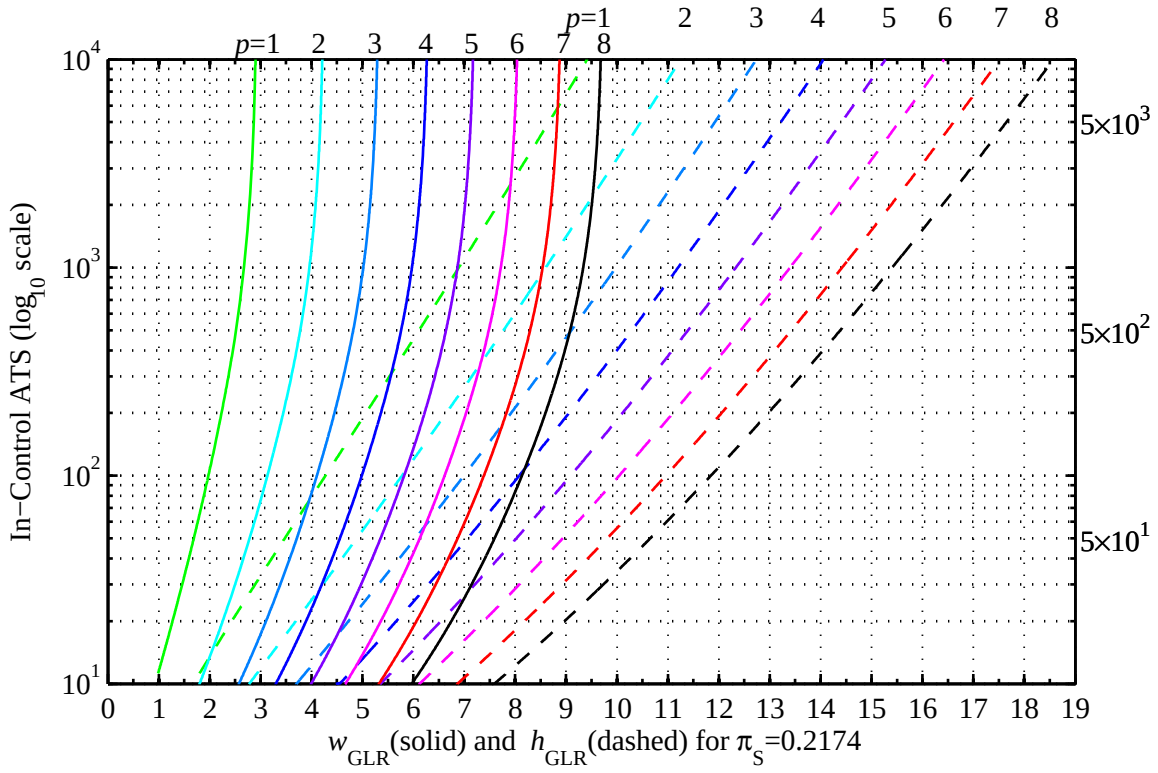


Figure 6.2: Plot of $\log_{10}$(In-Control ATS) versus w_{GLR} for $\pi_{\mathrm{S}} = 0.25$

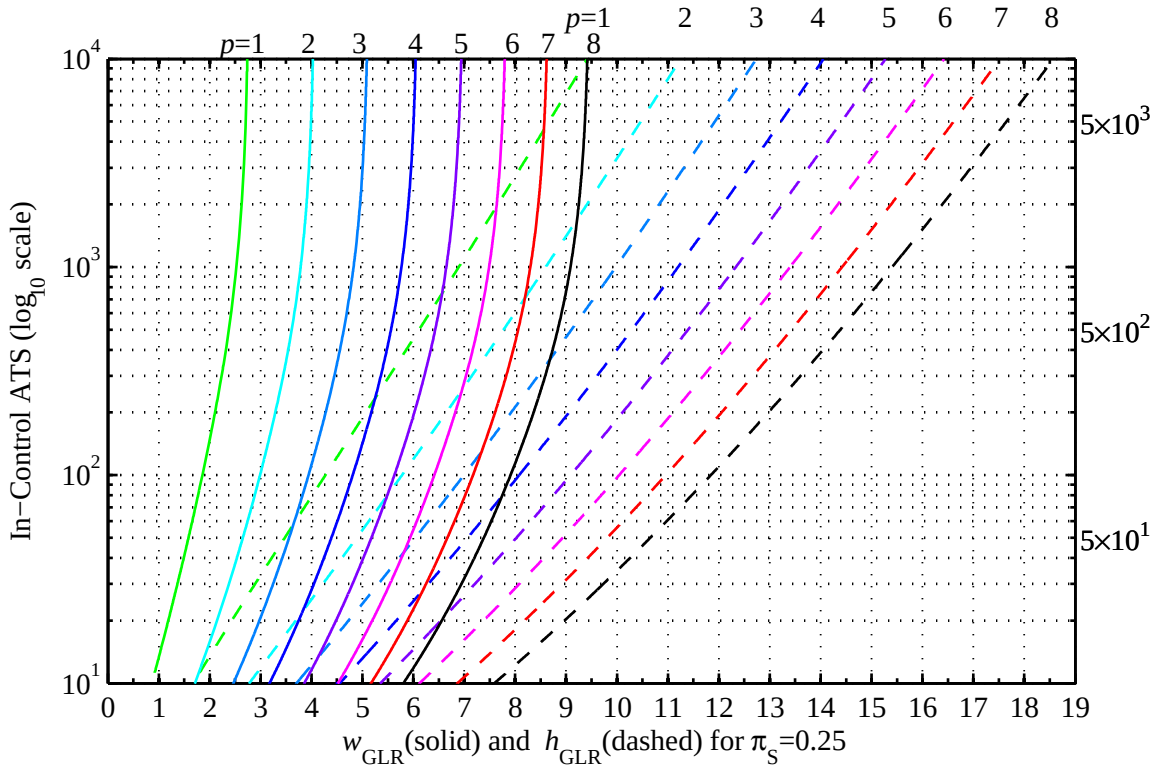


Figure 6.3: Plot of $\log_{10}$(In-Control ATS) versus w_{GLR} for $\pi_{\mathrm{S}} = 0.50$

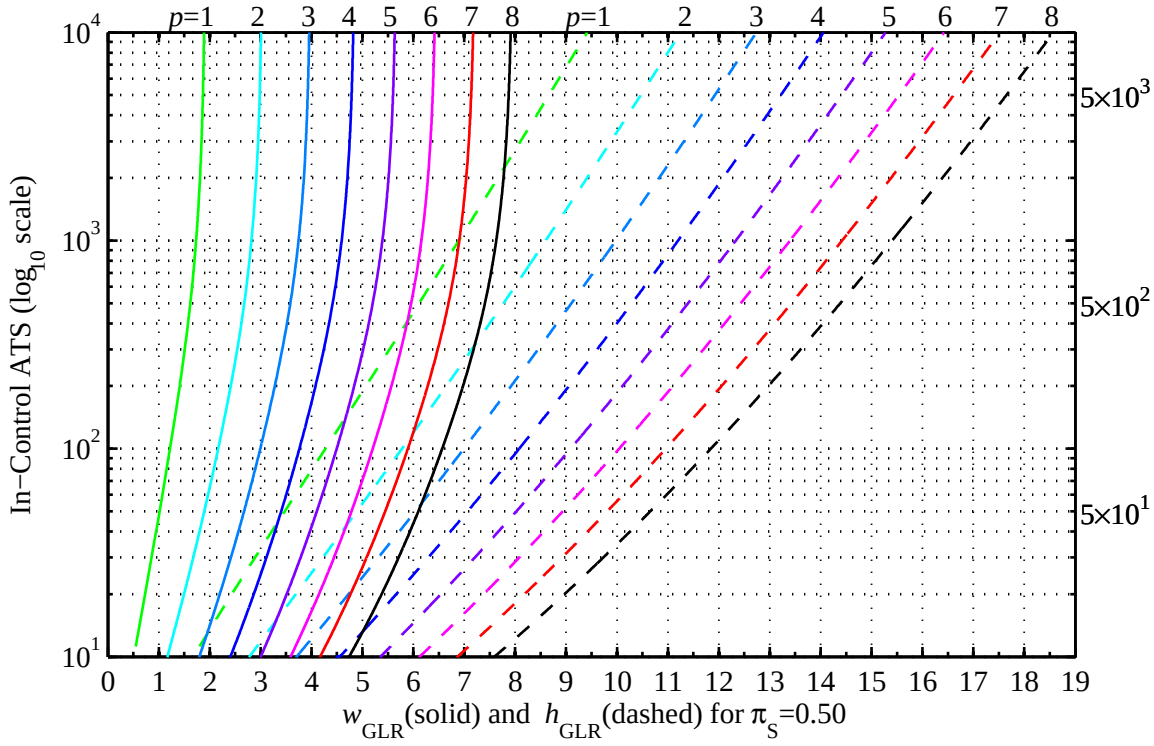
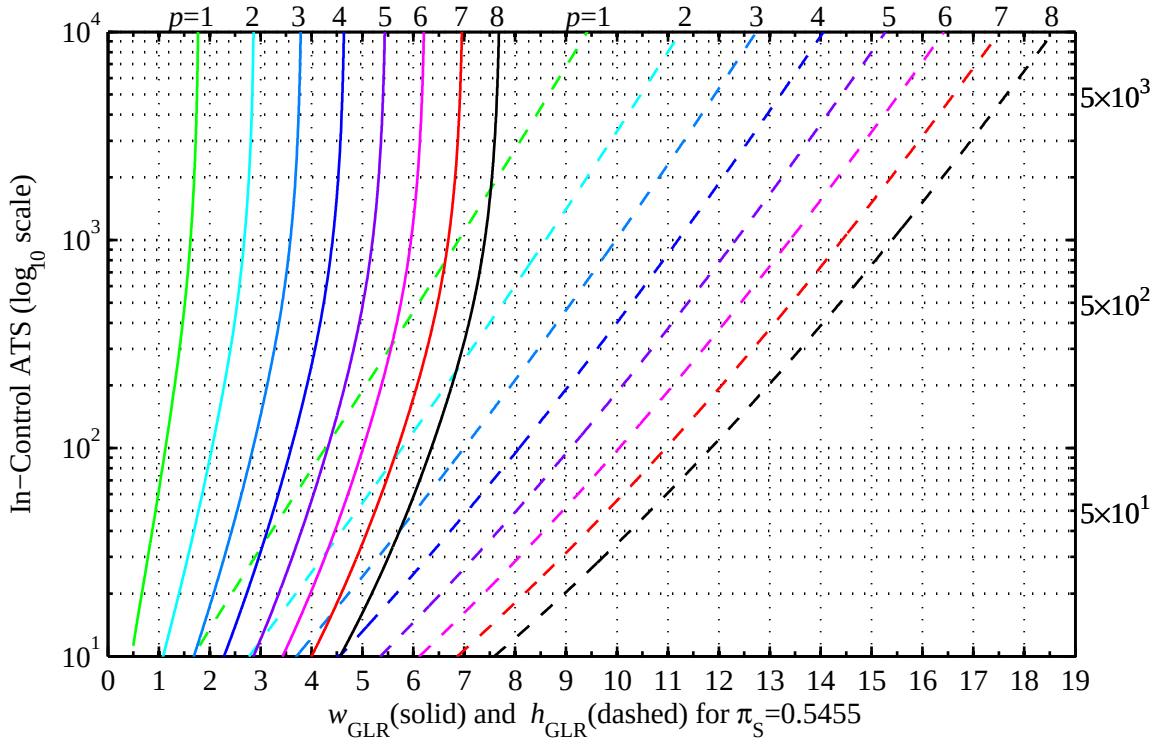


Figure 6.4: Plot of $\log_{10}$(In-Control ATS) versus w_{GLR} for $\pi_{\mathrm{S}} = 0.5455$



direct reading off the figures or from Equation 3.1.

For more accurate results, the fitted models of w_{GLR} as a polynomial function of $L = \log_{10}(\text{In-Control ATS})$ are also provided. Since the relationship between w_{GLR} and L is not as linear as that between h_{GLR} and L, we use 4^{th} order polynomials to approximate the underlying function. Note that there are about 900 data points for each model, the minimum coefficient of determination (model R^2) is greater than 0.9999, and the largest residual standard error is 0.0081. The estimated coefficients are summarized in Table 6.2. Users can use this table to calculate the w_{GLR} value for the desirable in-control ATS value and a corresponding VSI setting in a similar manner as for calculating h_{GLR} in Section 3.5. For example, suppose a user has a $p = 4$ dimensional process and a feasible VSI setting of $\pi_{\text{S}} = 0.50$, and he/she wishes to have an in-control ATS of 800, then $L = \log_{10}(800)$ and the control limit can be first found from Equation 3.1 as $h_{\text{GLR}} = 0.481699 + 4.389605 * L - 0.342118 * L^2 + 0.023314 * L^3 = 10.9122$, and the warning limit can be calculated from Table 6.2 as $w_{\text{GLR}} = 2.531777 - 1.206315 * L + 1.749727 * L^2 - 0.512865 * L^3 + 0.046645 * L^4 = 4.5412$. These values are very close to those limits in Table 6.1 ($h_{\text{GLR}} = 10.9122$ and $w_{\text{GLR}} = 4.5416$) which were obtained by careful trial and error.

Table 6.2: Coefficients for Approximating the Warning Limit

$w_{\mathrm{GLR}} = b_0 + b_1 L + b_2 L^2 + b_3 L^3 + b_4 L^4$

π_{S} $(d_{\mathrm{S}}, d_{\mathrm{L}})$	p	b_0	b_1	b_2	b_3	b_4
	1	1.001127	-0.911671	1.317204	-0.375570	0.033231
	2	1.682051	-0.746746	1.436484	-0.427418	0.038623
	3	2.286187	-0.569135	1.479645	-0.453198	0.041449
$\pi_{\mathrm{S}} = 0.2174$	4	2.893911	-0.475116	1.550064	-0.484058	0.044724
(0.10,1.25)	5	3.275747	-0.079409	1.440430	-0.470701	0.044109
	6	4.306317	-0.653768	1.835617	-0.572468	0.053190
	7	4.525649	-0.035203	1.573437	-0.518091	0.048710
	8	5.198855	-0.118254	1.690348	-0.552946	0.051947
	1	0.998274	-0.983215	1.322364	-0.374252	0.033080
	2	1.660136	-0.835076	1.452128	-0.428611	0.038684
	3	2.252023	-0.669294	1.502386	-0.456069	0.041648
$\pi_{\mathrm{S}} = 0.25$	4	2.854630	-0.591269	1.581939	-0.489107	0.045110
(0.25,1.25)	5	3.230367	-0.206338	1.478346	-0.477166	0.044614
	6	4.226259	-0.746655	1.855306	-0.574701	0.053335
	7	4.447934	-0.149286	1.605692	-0.523491	0.049143
	8	5.144976	-0.280362	1.747691	-0.564027	0.052851
	1	0.977118	-1.358123	1.334915	-0.360485	0.031474
	2	1.466018	-1.264527	1.505234	-0.424664	0.037873
	3	2.002739	-1.225516	1.631284	-0.470105	0.042343
$\pi_{\mathrm{S}} = 0.50$	4	2.531777	-1.206315	1.749727	-0.512865	0.046645
(0.10,1.90)	5	2.923088	-0.983335	1.737370	-0.522333	0.047956
	6	3.739188	-1.383071	2.044085	-0.604281	0.055393
	7	3.893872	-0.804397	1.810374	-0.557499	0.051608
	8	4.620450	-1.081856	2.034271	-0.617403	0.056973
	1	0.976580	-1.409641	1.333490	-0.357225	0.031096
	2	1.436882	-1.329285	1.513102	-0.423631	0.037675
	3	1.955793	-1.301252	1.646497	-0.470818	0.042287
$\pi_{\mathrm{S}} = 0.5455$	4	2.477862	-1.301553	1.776720	-0.516388	0.046825
(0.25,1.90)	5	2.861469	-1.092744	1.773045	-0.527925	0.048308
	6	3.652857	-1.478210	2.073083	-0.608408	0.055622
	7	3.809554	-0.923817	1.853566	-0.565028	0.052128
	8	4.531529	-1.213526	2.084984	-0.626767	0.057651

Chapter 7

Background on the Monitoring of Σ

Previously we have focused on the problem of monitoring the mean vector. For the remaining part of this dissertation we will consider the problem of monitoring the dispersion of MN processes.

7.1 Definition of the Process

For the discussion of monitoring the dispersion of multivariate processes, we continue to use the $p \times 1$ random vector $\boldsymbol{X} = [X_1, X_2, \ldots, X_p]'$ to denote the p quality measurements of an item, with the standard assumption that $\boldsymbol{X}$ follows a multivariate normal (MN) distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$. The in-control parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are assumed to be known or can be estimated with Phase I data with neglectable error.

In the previous chapter, we discussed the application of the GLR chart for $\boldsymbol{\mu}$ with the variable sample interval scheme. For the problem of monitoring $\boldsymbol{\Sigma}$, only the fixed sampling interval scheme is used. After each time interval of length d, a group of n independent items is randomly sampled from the process and the measurements of the p quality characteristic variables are taken immediately. Without loss of generality, it is assumed that $n = d$ so that

on average there is one item to be inspected after each unit-length time interval. Use the $p \times 1$ vector $\boldsymbol{X}_{ij} = [X_{ij}^{(1)}, \cdots, X_{ij}^{(p)}]'$ to denote the quality measurements of the j^{th} item in the i^{th} sample. It was pointed out that the sample mean is a sufficient statistic for monitoring $\boldsymbol{\mu}$ with the assumption that $\mathbf{\Sigma}$ does not change, thus the sample size can be assumed to be $n = 1$ without loss of generality. However this is not true for the monitoring of $\mathbf{\Sigma}$.

For the type of change in the process, we consider a sustained abrupt change in the covariance matrix. Assume that at some random time between two consecutive sampling points τ and $\tau + 1$, some incident occurs to the process, causing $\mathbf{\Sigma}$ to shift from its in-control value $\mathbf{\Sigma}_0$ to some unknown $\mathbf{\Sigma}_1$, while $\boldsymbol{\mu}$ remains unchanged. Samples taken at or before the time point τ are from the in-control distribution $\text{MN}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$, and those obtained at or after $\tau + 1$ come from the out-of-control distribution $\text{MN}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_1)$.

The covariance matrix $\mathbf{\Sigma}$ can be expressed in terms of the variance parameters and the correlation parameters

$$\mathbf{\Sigma} = [\text{diag}\, \boldsymbol{\sigma}]\mathbf{R}[\text{diag}\, \boldsymbol{\sigma}],$$

where $\boldsymbol{\sigma} = [\sigma^{(1)}, \cdots, \sigma^{(p)}]$ is the vector of the standard deviations of the p quality variables, $[\text{diag}\, \boldsymbol{\sigma}]$ is the diagonal matrix with elements of $\boldsymbol{\sigma}$ on the diagonal, and $\mathbf{R}$ is the correlation matrix. The purpose of the monitoring of MN dispersion in many applications is to detect any form of change in the covariance matrix, while some applications only consider an increase in the variability as a quality deterioration thus the detection of increases in the variance parameters is the primary interest. In this dissertation we consider both types of changes in $\mathbf{\Sigma}$, and refer to control charts for detecting a general change in $\mathbf{\Sigma}$ as the covariance charts, and those for monitoring only the variance parameters as the variance charts.

7.2 Literature Review for Monitoring $\boldsymbol{\Sigma}$

7.2.1 A General Review

Generally speaking, the topic of monitoring the dispersion of multivariate processes has received much less attention than the mean vector. Reviews of existing work can be found in Lowry and Montgomery (1995), Montgomery (2009), Yeh, Lin and McGrath (2006) and Bersimis, Psarakis and Panaretos (2007). Like the charts for monitoring $\boldsymbol{\mu}$, based on the nature of the method, most of the control charts for monitoring $\boldsymbol{\Sigma}$ belong to three categories: the Shewhart-type charts, the CUSUM-type charts and the EWMA-type charts.

Alt (1984) discussed two Shewhart-type charting procedures for monitoring $\boldsymbol{\Sigma}$ based on samples of size $n \geq p$ (so that the sample estimate of the covariance matrix is of full rank). The first one is based the likelihood ratio test (LRT) for comparing the sample likelihood under the sample covariance matrix against that under the in-control covariance matrix $\boldsymbol{\Sigma}_0$. This well-known LRT statistic is known to follow a chi-squared distribution with $\frac{p(p+1)}{2}$ degrees of freedom asymptotically. The second approach is based on the sample generalized variance (the determinant of the covariance matrix) whose asymptotic distribution under the in-control process is known. But a common criticism for control charts only based on the generalized variance is that the statistic over-simplifies the covariance matrix. For example, it is possible that the determinants of two distinctively different covariance matrices can be the same. Many other Shewhart-type procedures for monitoring $\boldsymbol{\Sigma}$ have been discussed, and discussion about these methods can be found in the aforementioned review papers. Due to the nature of the Shewhart-type charts that the charting statistic is only based on the current sample, these procedures are not effective for detecting small or intermediate shift sizes.

Cumulative procedures such as the CUSUM and EWMA methods have been applied to the monitoring of $\boldsymbol{\Sigma}$, although not as extensively as for the detection of changes in $\boldsymbol{\mu}$. With a focus on monitoring the mean vector, Hawkins (1991) and Hawkins (1993) also men-

tioned the idea of monitoring $\mathbf{\Sigma}$ based on the regression-adjusted variables using multiple CUSUM or multiple EWMA control charts. Huwang, Yeh and Wu (2007) discussed the implementation of this idea, and also proposed two control charts based on the trace of the EWMA-type running estimate of the covariance matrix. Yeh, Huwang, and Wu (2004) discussed an EWMA chart based the likelihood ratio statistic for a two-sample hypothesis test of equal covariance matrices, one sample being the Phase-I sample, and the other is repeatedly sampled from the Phase-II process. Hawkins and Maboudou-Tchao (2008) discussed a multivariate exponentially weighted moving covariance (MEWMC) chart which applies Alt (1984)'s statistic to compare the difference between the in-control covariance and the running MEWMC estimate similar to that used in Huang, Yeh and Wu (2007). In Reynolds and Cho (2006), Reynolds and Stoumbos (2008), and Reynolds and Cho (2011), the authors discussed the use of the MEWMA chart along with several variations of control charts based on the squared-deviations from target to jointly monitor the mean vector and the covariance matrix, and concluded that the charts that they called the M_2RZ^2 and M_2RA^2 charts give the best performance for detecting increases in the variability for independent and correlated variables, respectively.

There are also work on GLR-type procedures in the literature for the monitoring of $\mathbf{\Sigma}$. Sullivan and Woodall (2000) and Zamba and Hawkins (2009) discussed two control charts using the GLR change-point formulation for monitoring both the mean vector and the covariance matrix, where the former is a Phase-I procedure and the latter is designed for real-time monitoring when there is no Phase-I.

In this dissertation, as the benchmark to which the performance of the proposed control charts are compared, we choose the MEWMC chart for the detection of any general changes in $\mathbf{\Sigma}$, and the M_2RZ^2 and M_2RA^2 charts for the detection of variability increases only. Details about theses charts are given in the following subsections.

7.2.2 The MEWMC Chart

Hawkins and Maboudou-Tchao (2008) presented their MEWMC chart based on the assumption that an individual ($n = 1$) observation is sampled at each sampling time point. Their control chart is based on the standardized random vectors $\boldsymbol{U}$, whose in-control distribution is $\text{MN}(\mathbf{0}, \mathbf{I})$. They first defined a running estimate of the covariance matrix (the MEWMC estimate) at time k,

$$\mathbf{S}_k = \lambda \boldsymbol{U}_k \boldsymbol{U}_k^{'} + (1 - \lambda)\mathbf{S}_{k-1},$$

where $\mathbf{S}_0 = \mathbf{I}$ and $0 < \lambda \leq 1$. Then they applied Alt's likelihood ratio statistic to compare the estimated covariance matrix $\mathbf{S}_k$ to the in-control matrix $\mathbf{I}$ to get the MEWMC chart statistic

$$C_k = \text{tr}(\mathbf{S}_k) - \log|\mathbf{S}_k| - p.$$

A signal is given at time k if C_k is greater than a pre-specified control limit. Unlike most other control charts that are only based on either the trace or the determinant, the MEWMC chart can detect any general changes in $\mathbf{\Sigma}$. The MEWMC chart is shown to generally outperform other current control charts for monitoring the covariance matrix.

7.2.3 The M_2RZ^2 Chart

Reynolds and Kim (2007) and Reynolds and Cho (2011) presented their M_2RZ^2 chart based on the marginally-standardized vectors. Let $\boldsymbol{Z}_{ij} = [Z_{ij}^{(1)}, \cdots, Z_{ij}^{(p)}]'$ be the j^{th} standardized observation in the i^{th} sample, where $Z_{ij}^{(q)} = \frac{X_{ij}^{(q)} - \mu_0^{(q)}}{\sigma_0^{(q)}}$ is the q^{th} marginally-standardized quality variable. Note that the covariance matrix of $\boldsymbol{Z}_{ij}$ is the in-control correlation matrix $\mathbf{R}_0$ when the process is in-control. The authors first defined the EWMA statistic of the squared standardized deviation with a reset for the q^{th} variable at sampling time k as

$$E_k^{Z(q)} = \lambda \frac{1}{n} \sum_{j=1}^{n} \left(Z_{kj}^{(q)} \right)^2 + (1 - \lambda) \max \left\{ E_{k-1}^{Z(q)}, 1 \right\}, \quad q = 1, \ldots, p,$$

where $E_0^{Z(q)} = 1$ and $0 < \lambda \leq 1$. The 'max' function resets the previous estimate to the in-control value 1 if it drops below 1. It has been shown that the approxiamte covariance matrix of the vector of EWMA statistics is $\frac{2\lambda}{n(2-\lambda)}\mathbf{R}_0^{<2>}$, where $^{<2>}$ is the elementwise square operator. The M_2RZ^2 charting statistic is the quadratic form for the vector of the EWMA statistics

$$M_k^Z = \left(\frac{2\lambda}{n(2-\lambda)}\right)^{-1} [E_k^{Z(1)}, \ldots, E_k^{Z(p)}] \left(\mathbf{R}_0^{<2>}\right)^{-1} [E_k^{Z(1)}, \ldots, E_k^{Z(p)}]'.$$

A signal is given at time k if $M_k^Z > h$.

7.2.4 The M_2RA^2 Chart

Reynolds and Kim (2007) and Reynolds and Cho (2011) also discussed a similar M_2RA^2 chart based on the vector of regression-adjusted variables (see Hawkins (1991) and Hawkins (1993)) to improve the performance when the original p quality variables are correlated. The regression-adjusted vector corresponding to $\boldsymbol{Z}_{ij}$ is $\boldsymbol{A}_{ij} = [\operatorname{diag}\mathbf{R}_0^{-1}]^{-\frac{1}{2}}\mathbf{R}_0^{-1}\boldsymbol{Z}_{ij}$. Then following the same logic as the M_2RZ^2 chart, the EWMA statistics of the squared standardized deviation for the regression-adjusted variables are

$$E_k^{A(q)} = \lambda\frac{1}{n}\sum_{j=1}^{n}\left(a_{kj}^{(q)}\right)^2 + (1-\lambda)\max\left\{E_{k-1}^{A(q)}, 1\right\}, \quad q = 1, \ldots, p,$$

where $E_0^{A(q)} = 1$ and $0 < \lambda \leq 1$. The approximate covariance matrix of the vector of the EWMA statistics is $\mathbf{\Sigma}_0^A = \frac{2\lambda}{n(2-\lambda)}\left(\left[\operatorname{diag}\mathbf{R}_0^{-1}\right]^{-\frac{1}{2}}\mathbf{R}_0^{-1}\left[\operatorname{diag}\mathbf{R}_0^{-1}\right]^{-\frac{1}{2}}\right)^{<2>}$. Then the M_2RA^2 chart statistic is the quadratic form

$$M_k^A = \left[E_k^{A(1)}, \ldots, E_k^{A(p)}\right]\left(\mathbf{\Sigma}_0^A\right)^{-1}\left[E_k^{A(1)}, \ldots, E_k^{A(p)}\right]'.$$

A signal is given as soon as the control chart statistic exceeds the control limit.

Chapter 8

The GLR Chart for Monitoring $\boldsymbol{\Sigma}$

8.1 Derivation of the GLR Chart

At sampling point k, a series of $k \times n$ observations $\{\boldsymbol{X}_{11}, \cdots, \boldsymbol{X}_{1n}, \cdots, \boldsymbol{X}_{k1}, \cdots, \boldsymbol{X}_{kn}\}$ has been collected. Consider the log-likelihood ratio test for testing the hypotheses of whether or not there has been a change in $\boldsymbol{\Sigma}$ prior to the current sampling point k:

$$\begin{aligned}
&\mathrm{H}_0 : \{\boldsymbol{X}_{ij} : i = 1, \cdots, k;\ j = 1, \cdots, n\} \sim \mathrm{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0);\\
&\mathrm{H}_1 : \exists\ \tau \in \{0, 1, \cdots, k-1\} \text{ s.t. } \{\boldsymbol{X}_{ij} : i = 1, \cdots, \tau;\ j = 1, \cdots, n\} \sim \mathrm{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)\\
&\qquad \text{and } \{\boldsymbol{X}_{ij} : i = \tau + 1, \cdots, k;\ j = 1, \cdots, n\} \sim \mathrm{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_1),\ \text{where } \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_0.
\end{aligned}$$

The likelihood for H_0 is

$$L_k^{\mathrm{H}_0}(\boldsymbol{\Sigma}_0) = \prod_{i=1}^{k} \prod_{j=1}^{n} f(\boldsymbol{X}_{ij} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0),$$

where $f()$ is the probability density function of the MN distribution. For the alternative hypothesis that the process has already been out of control, the likelihood function is

$$L_k^{\mathrm{H}_1}(\mathbf{\Sigma}_0, \tau, \mathbf{\Sigma}_1) = \prod_{i=1}^{\tau}\prod_{j=1}^{n} f(\boldsymbol{X}_{ij}|\boldsymbol{\mu}_0, \mathbf{\Sigma}_0) \times \prod_{i=\tau+1}^{k}\prod_{j=1}^{n} f(\boldsymbol{X}_{ij}|\boldsymbol{\mu}_0, \mathbf{\Sigma}_1).$$

However τ and $\mathbf{\Sigma}_1$ are both unknown and have to be estimated. The method of profile likelihood can be used to obtain these estimates: suppose we can find a good estimator of $\mathbf{\Sigma}_1$ (we will discuss three types of estimators in the following section), say $\hat{\mathbf{\Sigma}}_{1,t,k}$, assuming that a previous integer time point t is the sampling time point τ immediately before the shift; then the maximum likelihood estimator (MLE) for τ is $\hat{\tau} = \arg\max_{t\in\{0,\ldots,k-1\}} L_k^{\mathrm{H}_1}(\mathbf{\Sigma}_0, t, \hat{\mathbf{\Sigma}}_{1,t,k})$. We plug in these estimates to get the following GLR statistic of the log likelihood ratio test for comparing the two hypotheses:

$$R_k = \ln \frac{L_k^{\mathrm{H}_1}(\mathbf{\Sigma}_0, \hat{\tau}, \hat{\mathbf{\Sigma}}_{1,\hat{\tau},k})}{L_k^{\mathrm{H}_0}(\mathbf{\Sigma}_0)} = \max_{t=0,\ldots,k-1} r_{t,k} \tag{8.1}$$

where

$$r_{t,k} = -\frac{1}{2}\left[(k-t)n\left(\ln|\hat{\mathbf{\Sigma}}_{1,t,k}| - \ln|\mathbf{\Sigma}_0|\right) + \sum_{i=t+1}^{k}\sum_{j=1}^{n}(\boldsymbol{X}_{ij}-\boldsymbol{\mu}_0)'\hat{\mathbf{\Sigma}}_{1,t,k}^{-1}(\boldsymbol{X}_{ij}-\boldsymbol{\mu}_0) - \sum_{i=t+1}^{k}\sum_{j=1}^{n}(\boldsymbol{X}_{ij}-\boldsymbol{\mu}_0)'\mathbf{\Sigma}_0^{-1}(\boldsymbol{X}_{ij}-\boldsymbol{\mu}_0)\right]. \tag{8.2}$$

A signal is given at sampling point k if $R_k > h_{\mathrm{GLR}}$, where h_{GLR} is the control limit specified to satisfy a desirable in-control ATS.

8.2 The Window Restricted GLR Chart

As previously mentioned, the GLR-type charts are known in the literature to be computationally intensive. The computational burden of the GLR chart for monitoring $\mathbf{\Sigma}$ is even

worse that that for monitoring $\boldsymbol{\mu}$. Again we employ a moving window to restrict the calculation of the GLR statistic to only be based on the m most recent samples, and the window restricted GLR statistic is

$$R_k^m = \max_{t=\max\{0,k-m\},\ldots,k-1} r_{t,k} \ . \tag{8.3}$$

In some applications such as an autocorrelated process with feed back control (see Apley and Shi (1999)), m can be used as a tuning parameter of the GLR chart. However in the application to processes with independent observations, m is also perceived as a tuning parameter by some users. In the literature several decades ago when the computing power is not as advanced as today's, there were some published work using far-from-sufficient window sizes: a small window size discards past observations too soon and thus yields a lower control limit, which in turn can result in a slightly better performance for detecting large shift sizes but is much worse for small shifts. Recent studies of the asymptotic properties in Lai (1995), Lai (1998) and Lai (2001), the evaluations of the GLR charts in Reynolds and Lou (2010) and Reynolds et al. (2011), and our discussion of the effect of the window size for the GLR chart for $\boldsymbol{\mu}$ chart, all suggest that a sufficiently large window size should be used so that the GLR chart can detect small shifts effectively. Therefore we recommend that in practical applications, users should use the GLR chart without the window restriction or with a window at least as large as the average in-control ATS value.

8.3 Choices of the Covariance Matrix Estimator

In this section we discuss three variations of the GLR chart for monitoring $\boldsymbol{\Sigma}$, based on three choices of $\hat{\boldsymbol{\Sigma}}_{1,t,k}$ for the assumption that the change-point when the covariance matrix shifts from $\boldsymbol{\Sigma}_0$ to some $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_0$ is on the interval (t, t+1].

8.3.1 The Maximum Likelihood Estimator

A natural candidate for $\hat{\boldsymbol{\Sigma}}_{1,t,k}$ is the MLE, given as,

$$\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}} = \frac{1}{(k-t)n} \sum_{i=t+1}^{k} \sum_{j=1}^{n} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)(\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)'.$$

Note that $\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}}$ has to be positive definite, thus it is only defined when $(k-t)n \geq p$. Using $\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}}$, $r_{t,k}$ can be simplified as

$$r_{t,k}^{\mathrm{ML}} = -\frac{(k-t)n}{2} \left[\ln |\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}}| - \ln |\boldsymbol{\Sigma}_0| + p - \mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}} \boldsymbol{\Sigma}_0^{-1} \right) \right] \tag{8.4}$$

since

$$\begin{aligned}
&\sum_{i=t+1}^{k} \sum_{j=1}^{n} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)' (\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}})^{-1} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0) - \sum_{i=t+1}^{k} \sum_{j=1}^{n} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0) \\
&= \mathrm{tr}\left(\sum_{i=t+1}^{k} \sum_{j=1}^{n} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)(\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)' (\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}})^{-1} \right) - \mathrm{tr}\left(\sum_{i=t+1}^{k} \sum_{j=1}^{n} (\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)(\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} \right) \\
&= (k-t)n \,\mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}} (\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}})^{-1} \right) - (k-t)n \,\mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}} \boldsymbol{\Sigma}_0^{-1} \right) \\
&= (k-t)n \left[p - \mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{ML}} \boldsymbol{\Sigma}_0^{-1} \right) \right].
\end{aligned}$$

For a random sample of $(k-t)n$ observations from the p-variate MN distribution, with $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ given, $2r_{t,k}^{\mathrm{ML}}$ is the log likelihood ratio statistic for testing whether or not the true covariance matrix is $\boldsymbol{\Sigma}_0$. Under the in-control distribution, the asymptotic distribution of $r_{t,k}^{\mathrm{ML}}$ is $\frac{1}{2}\chi^2_{\frac{p(p+1)}{2}}$, as there are $\frac{p(p+1)}{2}$ free parameters to be estimated for the alternative hypothesis. The asymptotic distribution only works for large $(k-t)n$; when $(k-t)n$ is small, the expected value of $r_{t,k}^{\mathrm{ML}}$, $\mathrm{E}\left[r_{t,k}^{\mathrm{ML}}\right]$, is much higher than the asymptotic mean of

$\frac{1}{2}\frac{p(p+1)}{2}$. Following the same steps in Zamba and Hawkins (2009), it can be shown that

$$\mathrm{E}\left[r_{t,k}^{\mathrm{ML}}\right] = -\frac{(k-t)n}{2}\left(-p\ln\left(\frac{(k-t)n}{2}\right) + \sum_{i=1}^{p} f_{\psi}\left(\frac{(k-t)n-i}{2}\right)\right)$$

where $f_{\psi}()$ is the di-gamma function. As an illustration, in Figure 8.1, $\mathrm{E}\left[r_{t,k}^{\mathrm{ML}}\right]$ is plotted for $(k-t) \in \{4, \cdots, 60\}$ for a $p = 4$ dimensional MN process with $n = 1$ and $\mathbf{\Sigma}_0 = \mathbf{I}$. Note that the asymptotic mean for $r_{t,k}^{\mathrm{ML}}$ is $\frac{1}{2}\frac{4*5}{2} = 5$. As a result, simply plugging $r_{t,k}^{\mathrm{ML}}$ into the GLR

Figure 8.1: The expected value of $r_{t,k}^{\mathrm{ML}}$

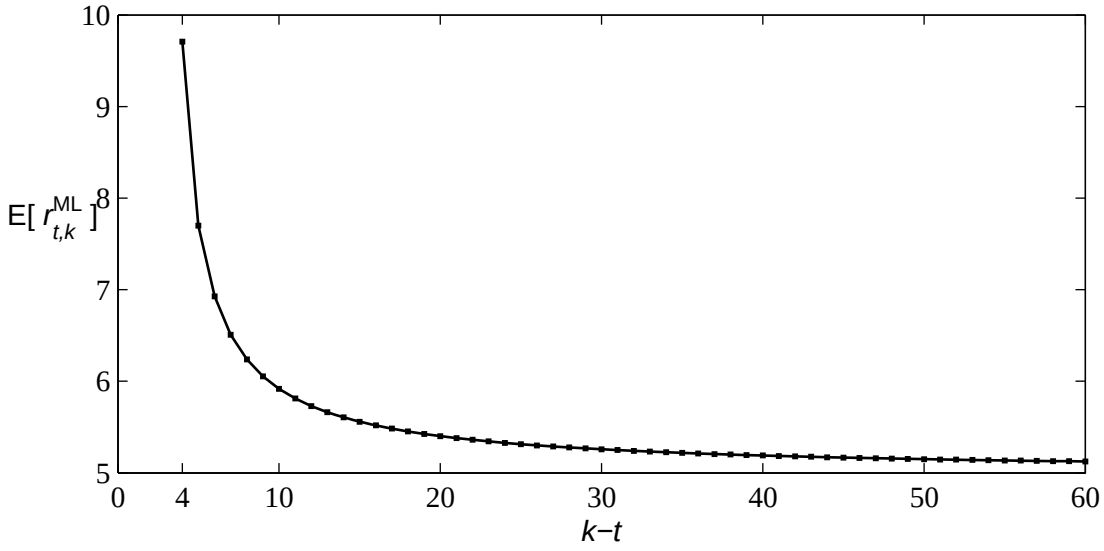


chart statistic in Equations 8.1 and 8.3 does not give satisfactory performances, because the control statistic is inflated by $r_{t,k}^{\mathrm{ML}}$ with small $(k-t)n$, and this requires a too-large value of h_{GLR}.

A remedy is to use a similar Bartlett Correction (BC) as in Zamba and Hawkins (2009). The idea is to multiply $r_{t,k}^{\mathrm{ML}}$ by $f_{\mathrm{BC}}(k-t,n,p)$ to make the expected value of the product constant across different values of $(k-t)n$, where

$$f_{\mathrm{BC}}(k-t,n,p) = \frac{p(p+1)/2}{\mathrm{E}\left[r_{t,k}^{\mathrm{ML}}\right]}\ , \tag{8.5}$$

and the corrected window restricted GLR statistic is

$$R_k^{\mathrm{BC}} = \max_{t=\max\{0,k-m\},\ldots,k-1} r_{t,k}^{\mathrm{ML}} f_{\mathrm{BC}}(k-t,n,p) \ . \tag{8.6}$$

We will refer to this GLR chart with Bartlett Correction as the GLR-BC chart.

8.3.2 The Multivariate Exponentially Weighted Moving Covariance Estimator

Instead of using the MLE, an alternative for $\hat{\boldsymbol{\Sigma}}_{1,t,k}$ is the multivariate exponentially weighted moving covariance (MEWMC) estimator

$$\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{E}} = \lambda \frac{1}{n} \sum_{j=1}^{n} (\boldsymbol{X}_{tj} - \boldsymbol{\mu}_0)(\boldsymbol{X}_{tj} - \boldsymbol{\mu}_0)' + (1-\lambda)\hat{\boldsymbol{\Sigma}}_{1,t,k-1}^{\mathrm{E}} \ ,$$

where $0 < \lambda \leq 1$ is the tuning parameter and $\hat{\boldsymbol{\Sigma}}_{1,t,t}^{\mathrm{E}} = \boldsymbol{\Sigma}_0$.

The procedures discussed in Huwang et al. (2006) and Hawkins and Maboudou-Tchao (2008) use similar MEWMC estimators: the former procedure signals when the trace of the MEWMC estimator falls in the tail area of its in-control distribution while the latter applies Alt (1984)'s LRT statistic to compare the running MEWMC estimate to the in-control value $\boldsymbol{\Sigma}_0$. Here in our procedure, we use a change-point formulation by plugging $\hat{\boldsymbol{\Sigma}}_{1,t,k}^{\mathrm{E}}$ into $r_{t,k}$ (denote as $r_{t,k}^{\mathrm{E}}$) and maximizing over all past sampling points to obtain the GLR statistic

$$R_k^{\mathrm{E}} = \max_{t=\max\{0,k-m\},\ldots,k-1} r_{t,k}^{\mathrm{E}} \ . \tag{8.7}$$

We will use GLR-E to denote this chart.

8.3.3 The Constrained Maximum Marginal Likelihood Estimator

The GLR-BC chart and the GLR-E chart are designed to detect any changes in the whole covariance matrix of the underlying distribution. However in some applications, when the process goes out of control, only the variance parameters are going to be affected, and usually only increases in the variance are considered to be a deterioration in quality. Here we consider an estimator that can be used with the GLR chart for effectively detecting only increases in the variance parameters.

First we rewrite the in-control covariance matrix as $\mathbf{\Sigma}_0 = [\text{diag}\,\boldsymbol{\sigma}_0]\mathbf{R}_0[\text{diag}\,\boldsymbol{\sigma}_0]$, where $\mathbf{R_0}$ is the in-control correlation matrix, $\boldsymbol{\sigma}_0 = [\sigma_0^{(1)}, \cdots, \sigma_0^{(p)}]'$ is the vector of the in-control standard deviations of the p quality variables and $[\text{diag}\,\boldsymbol{\sigma}_0]$ is a diagonal matrix with the elements of $\boldsymbol{\sigma}_0$ on the diagonal. Under the constraints that the correlation structure remains unchanged and the variances can only increase when the process is out of control, the constrained maximum likelihood estimator (CMLE) of the covariance matrix can be obtained as $\hat{\mathbf{\Sigma}}_{1,t,k}^{\text{CML}} = \text{diag}(\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CML}})\mathbf{R}_0\,\text{diag}(\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CML}})$, where $\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CML}}$ is the CMLE of the standard deviations. Generally there is no closed-form expression for $\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CML}}$, therefore it has to be solved using numerical optimization methods.

Alternatively, for the purpose of computational simplicity, we consider an estimator

$$\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CMML}} = \left[\hat{\sigma}_{1,t,k}^{(1)\text{CMML}}, \cdots, \hat{\sigma}_{1,t,k}^{(p)\text{CMML}}\right]', \tag{8.8}$$

where $\hat{\sigma}_{1,t,k}^{(q)\text{CMML}} = \max\left(\sigma_0^{(q)}, \sum_{i=t+1}^{k}\sum_{j=1}^{n}\frac{(X_{ij}^{(q)}-\mu_0^{(q)})^2}{(k-t)n}\right)$ for $q = 1, \cdots, p$ is the constrained maximum marginal likelihood estimator (CMMLE) of the standard deviation for the q^{th} dimension. Then it follows that the CMMLE for the covariance matrix is

$$\hat{\mathbf{\Sigma}}_{1,t,k}^{\text{CMML}} = \left[\text{diag}\,\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CMML}}\right]\mathbf{R}_0\left[\text{diag}\,\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CMML}}\right]. \tag{8.9}$$

Note that $\hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CMML}} = \hat{\boldsymbol{\sigma}}_{1,t,k}^{\text{CML}}$ when $\mathbf{R}_0 = \mathbf{I}$. Plug $\hat{\mathbf{\Sigma}}_{1,t,k}^{\text{CMML}}$ into the log-likelihood ratio statistic

$r_{t,k}$ and denote the result as $r^{\mathrm{V}}_{t,k}$ (the superscript 'V' is used because only the variance parameters are to be estimated). Then the GLR statistic using the CMMLE for detecting only increases in the variance parameters is

$$R^{\mathrm{V}}_k = \max_{t=\max\{0,k-m\},\ldots,k-1} r^{\mathrm{V}}_{t,k} \ . \tag{8.10}$$

We will refer to this chart as the GLR-V chart.

8.4 The Standardization of the GLR Charts

In all the previous formula, we have developed the GLR charts based on the original vector of quality variables $\boldsymbol{X}_{ij}$ whose in-control distribution is $\mathrm{MN}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$. Since $\mathbf{\Sigma}_0$ is positive definite, there exists a spectral decomposition such that $\mathbf{\Sigma}_0 = \mathbf{PDP}'$, where $\mathbf{P}$ is a $p \times p$ orthogonal matrix and $\mathbf{D}$ is a $p \times p$ diagonal matrix with non-zero diagonal elements. Consider the transformation $\boldsymbol{U}_{ij} = \mathbf{PD}^{-\frac{1}{2}}\mathbf{P}'(\boldsymbol{X}_{ij} - \boldsymbol{\mu}_0)$. Then it follows that $\boldsymbol{U}_{ij} \sim \mathrm{MN}(\mathbf{0}, \mathbf{I})$ when the process is in-control. Now we show that monitoring $\boldsymbol{X}_{ij} \sim \mathrm{MN}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$ is equivalent to monitoring $\boldsymbol{U}_{ij} \sim \mathrm{MN}(\mathbf{0}, \mathbf{I})$ for the GLR-E chart. Note that we now use the notation $S(\boldsymbol{X})$ and $S(\boldsymbol{U})$ to differentiate the statistics S based on the original and the transformed variables.

First consider the MEWMC statistic. It can be easily shown that the MEWMC based on $\boldsymbol{U}$ is

$$\begin{aligned}
\hat{\mathbf{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{U}) &= \lambda \sum_{i=t+1}^{k} (1-\lambda)^{k-i} \frac{1}{n} \sum_{j=1}^{n} \boldsymbol{U}_{tj}\boldsymbol{U}'_{tj} + (1-\lambda)^{k-t}\mathbf{I} \\
&= \lambda \sum_{i=t+1}^{k} (1-\lambda)^{k-i} \frac{1}{n} \sum_{j=1}^{n} \mathbf{PD}^{-\frac{1}{2}}\mathbf{P}'(\boldsymbol{X}_{tj} - \boldsymbol{\mu}_0)(\boldsymbol{X}_{tj} - \boldsymbol{\mu}_0)'\mathbf{PD}^{-\frac{1}{2}}\mathbf{P}' \\
&\qquad\qquad + \quad (1-\lambda)^{k-t}\mathbf{PD}^{-\frac{1}{2}}\mathbf{P}'\mathbf{PDP}'\mathbf{PD}^{-\frac{1}{2}}\mathbf{P}' \\
&= \mathbf{PD}^{-\frac{1}{2}}\mathbf{P}' \left(\hat{\mathbf{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})\right) \mathbf{PD}^{-\frac{1}{2}}\mathbf{P}',
\end{aligned}$$

whose determinant and trace are

$$\begin{aligned}
|\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{U})| &= |\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'| \cdot |\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})| \cdot |\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'| \\
&= |\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'| \cdot |\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})| = \frac{|\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})|}{|\boldsymbol{\Sigma}_0|}, \\
\mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{U})\right) &= \mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'\mathbf{P}\mathbf{D}^{-\frac{1}{2}}\mathbf{P}'\right) = \mathrm{tr}\left(\hat{\boldsymbol{\Sigma}}^{\mathrm{E}}_{1,t,k}(\boldsymbol{X})\boldsymbol{\Sigma}_0^{-1}\right).
\end{aligned}$$

Plugging these quantities into Equation 8.2, and it can be seen that $r^{\mathrm{E}}_{t,k}(\boldsymbol{U}) = r^{\mathrm{E}}_{t,k}(\boldsymbol{X})$ and $R^{\mathrm{E}}_{k}(\boldsymbol{U}) = R^{\mathrm{E}}_{k}(\boldsymbol{X})$. The same property can also be shown for the GLR-BC chart following the same steps.

The standardizable property means that the GLR-BC and GLR-E charts based on the original vector of quality variables are equivalent to those based on the standardizing transformation. This suggests that when simulating the performances of these charts, we can simply use $\mathrm{MN}(\mathbf{0}, \mathbf{I})$ as the in-control distribution to simplify the computation. For the same dimension p, the control limits for these two charts obtained based on $\boldsymbol{U} \sim \mathrm{MN}(\mathbf{0}, \mathbf{I})$ can be applied to the monitoring of any $\boldsymbol{X} \sim \mathrm{MN}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$.

Unfortunately there is no such standardization for the GLR-V chart for detecting only the variance increases (the GLR-V statistic depends on the correlation matrix $\mathbf{R}_0$).

Chapter 9

Performance Comparison of Control Charts for Monitoring $\boldsymbol{\Sigma}$

9.1 Performance Measurement

To measure control chart performance in monitoring $\boldsymbol{\Sigma}$, we will use the average time to signal (ATS) as the main metric as was done previously for monitoring $\boldsymbol{\mu}$. When the process is in-control, the ATS is the expected time to the first false alarm since the start of the monitoring. For the out-of-control situation, we use the steady state ATS (SSATS), which is based on the assumption that the process will first remain in-control for some time to allow the the charting statistic to get into its steady state distribution, then at a random time between sampling points τ and $\tau + 1$, a sustained abrupt change in $\boldsymbol{\Sigma}$ occurs and the SSATS is the expected time to signal after the actual change-point. It is also assumed that there is no false alarm prior to the actual change-point.

For a fair performance comparison of different control charts, first we have to set the control limits so that each chart has the same in-control ATS value, then compare the SSATS values under different out-of-control settings. The control chart with the smallest SSATS

value is considered to be the best for detecting that specific change setting.

The objective of this dissertation is to find control charts with relatively good overall performance for detecting a wide range of shift sizes under different out-of-control settings. However it should be noted that long after the very early stage of statistical quality control, it is now almost impossible that we can find a new control chart that outperforms the current methods uniformly across all the different out-of-control scenarios. For some comparisons, it is fairly easy to draw a conclusion about which procedure is generally better; for some cases, it is not. Especially in the problem of monitoring the general covariance matrix, there are a vast variety of change types and change sizes that can happen; simply comparing the SSATS values will not give us an immediate conclusion. In Reynolds and Lou (2010), the authors used a one-number index, the extra quadratic loss, to summarize control chart performances. In a later session where we compare the GLR-E chart and the MEWMC chart, a similar index, the expected loss (EL) will be used.

First we assume a prior distribution for the out-of-control $\mathbf{\Sigma}_1$, $\pi(\mathbf{\Sigma}_1)$, and choose an appropriate loss function that quantifies the loss rate for the process to run under $\mathbf{\Sigma}_1$ rather than the in-control $\mathbf{\Sigma}_0$. A natural choice of the prior distribution is the Wishart distribution, as it is the multivariate analogy to the univariate chi-squared distribution whose population is positive definite matrices. Presumably the loss function should be invariant to the multi-standardization (because both charts are invariant to the standardization) and the loss rate should be small for $\mathbf{\Sigma}_1$ close to $\mathbf{\Sigma}_0$ and large for $\mathbf{\Sigma}_1$ distinctively different from the in-control covariance matrix. Here we use the Stein's loss function

$$l(\mathbf{\Sigma}_1, \mathbf{\Sigma}_0) = \left(\mathrm{tr}(\mathbf{\Sigma}_1^{-1}) - \log|\mathbf{\Sigma}_1^{-1}|\right) - \left(\mathrm{tr}(\mathbf{\Sigma}_0^{-1}) - \log|\mathbf{\Sigma}_0^{-1}|\right),$$

which is the same function as the MEWMC chart statistic in Hawkins and Maboudou-Tchao (2008). Then the EL is the expected loss before a chart detects a shift to $\mathbf{\Sigma}_1$ which follows

a prior distribution,

$$\text{EL} = \int_{\mathbf{\Sigma}_1} l(\mathbf{\Sigma}_1, \mathbf{I}) SSATS(\mathbf{\Sigma}_1) \pi(\mathbf{\Sigma}_1) d\mathbf{\Sigma}_1.$$

This quantity can be evaluated using Monte Carlo integration with simulation.

9.2 Parameter Choices

We presented our GLR charts for $\mathbf{\Sigma}$ in the general form that allows more than one observation to be measured at each sampling time, but here in the comparison study we assume that only one observation is taken ($n = 1$ and $d = 1$) since it is the accepted wisdom that single observation sampling is better than rational grouping for control charts that can accumulate information (see Hawkins and Olwell (1998), Reynolds and Stoumbos (2004a) and Reynolds and Stoumbos (2004b)). All the comparison results are based on a $p = 4$ dimensional process.

Each control chart requires the specification of the control limit according to a desirable in-control performance. In the comparison, the control limits are set so that each chart has an in-control ATS of approximately 800. Other than the control limit, many charts have a tuning parameter that also has to be specified: the exponential weight parameters for the EWMA-type estimators in the GLR-E chart, the MEWMC chart, the M_2RZ^2 chart and the M_2RA^2 chart. In the comparison, we evaluate a wide range of parameter choices: $\lambda \in \{0.01, 0.05, 0.10, 0.20\}$. For the purpose of computation efficiency, we are using the window size $m = 600$ for all the GLR charts. We have verified that for an in-control ATS of 800, this window size is sufficient for detecting the smallest changes that we are considering.

Since all the covariance charts are invariant to the multi-standardization, the simulation is based on the standardized vectors whose in-control distribution is $\text{MN}(\mathbf{0}, \mathbf{I})$. However such standardization does not apply to the variance charts. The variance charts depend on the in-control correlation structure and thus can only be marginally standardized. The

simulation of these variance charts are based on the marginally standardized vectors whose in-control distribution is $\mathrm{MN}(\mathbf{0}, \mathbf{R}_0)$. In the comparison of the variance charts, two special in-control covariance matrices are considered: one is the independent case with $\mathbf{\Sigma}_{01} = \mathbf{R}_{01} = \mathbf{I}$, and the other is the highly correlated case with

$$\mathbf{\Sigma}_{02} = \mathbf{R}_{02} = \begin{bmatrix} 1 & 0.9 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 & 0.9 \\ 0.9 & 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 0.9 & 1 \end{bmatrix}.$$

Note that in the remainder of this chapter, when we talk about the out-of-control covariance matrix $\mathbf{\Sigma}_1$, we are referring to the covariance matrix of the multi-standardized vectors for the covariance charts and the marginally standardized vectors for the variance charts.

Most of the numerical results in this chapter are based on simulation with 1 million iterations.

9.3 The Out-of-Control Settings

The effectiveness of a control chart to detect a change in the process is measured by the SSATS, which assumes that at a random time uniformly distributed on the interval between sampling points $\tau = 400$ and 401, a special incident causes a change of the covariance matrix to some $\mathbf{\Sigma}_1$. Since the covariance charts and the variance charts are designed for different monitoring purposes, their performances are evaluated under different shift patterns.

For the covariance charts, since there are a total number of $\frac{p(p+1)}{2}$ parameters in the covariance matrix and it is possible that any subset of the parameters can change by any magnitude, it is nearly impossible to tabulate the SSATS values under all out-of-control scenarios. Here we consider four particular shift patterns in $\mathbf{\Sigma}_1$ when the process goes out of control:

OOC#1: the variances increase and the correlations remain constant;

OOC#2: the variances decrease and the correlations remain constant;

OOC#3: the correlations change along with the variances increase/decrease;

OOC#4: the change only affects the parameters associated with only two quality variables: one variance increases, one variance decreases, and there is a change in the correlation between the two variables.

We rewrite the $\mathbf{\Sigma}_1$ in terms of the variance parameters and the correlation matrix: $\mathbf{\Sigma}_1 = [\text{diag}\,\boldsymbol{\sigma}_1]\mathbf{R}_1[\text{diag}\,\boldsymbol{\sigma}_1]$. Then OOC#1 corresponds to $\boldsymbol{\sigma}_1 > \mathbf{1}$ element-wise and $\mathbf{R}_1 = \mathbf{I}$. The size of the increase in $\boldsymbol{\sigma}_1$ is expressed in terms of

$$\psi = \sqrt{(\boldsymbol{\sigma}_1 - \mathbf{1})'(\boldsymbol{\sigma}_1 - \mathbf{1})} = \sqrt{\sum_{q=1}^{p} \left(\sigma_1^{(q)} - 1\right)^2}.$$

OOC#2 corresponds to the case $\mathbf{0} < \boldsymbol{\sigma}_1 < \mathbf{1}$ element wise and $\mathbf{R}_1 = \mathbf{I}$. We have used ψ with positive values to describe the size of the increase in $\boldsymbol{\sigma}_1$, here we recycle the notation and use ψ with a negative value to describe the size of the decrease:

$$\psi = -\sqrt{(\boldsymbol{\sigma}_1^{<-1>} - \mathbf{1})'(\boldsymbol{\sigma}_1^{<-1>} - \mathbf{1})} = -\sqrt{\sum_{q=1}^{p} \left(\frac{1}{\sigma_1^{(q)}} - 1\right)^2},$$

where $^{<-1>}$ is the element-wise inverse operator. OOC#3 corresponds to changes in the correlation structure

$$\mathbf{R}_1 = \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

so that all correlations becomes ρ when the process goes out of control, and possibly with/without variance increases/decreases. We use ρ to denote the size of the shift in $\mathbf{R}$ and use the same ψ as used in OOC#1/OOC#2 to denote the size of the increase/decrease in

$\boldsymbol{\sigma}$. OOC#4 corresponds to a covariance matrix of the form

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} (\sigma_1^{(1)})^2 & \rho\sigma_1^{(1)}\sigma_1^{(2)} & 0 & 0 \\ \rho\sigma_1^{(1)}\sigma_1^{(2)} & (\sigma_1^{(2)})^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where $\sigma_1^{(1)} > 1$ and $\sigma_1^{(2)} < 1$. We use ρ to denote the correlation change between the first two quality characteristic variables, and use ψ with $\pm$ prefixed values to indicate that there is one variance increase and one decrease, where the value of ψ is expressed in terms of

$$\psi = \pm\sqrt{\left(\sigma_1^{(1)} - 1\right)^2 + \left(\frac{1}{\sigma_1^{(2)}} - 1\right)^2}.$$

It should be noted that the covariance charts are not invariant to the direction of the shift in $\boldsymbol{\sigma}$, so the SSATS is averaged over random directions of $\boldsymbol{\sigma}_1$. In the simulation for OOC#1, OOC#2, and OOC#3, we first generate a random vector from the $4-$variate standard MN distribution and reflect any negative values to positive, then rescale this vector so that its Euclidean distance from $\mathbf{0}$ is 1. Denote this unit-distance positive vector as $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]'$. Then $\boldsymbol{\sigma}_1$ can be obtained as

$$\boldsymbol{\sigma}_1 = \begin{cases} [1 + \psi\beta_1, \ldots, 1 + \psi\beta_4]' & \text{if } \psi > 0 \\ \left[\frac{1}{1-\psi\beta_1}, \ldots, \frac{1}{1-\psi\beta_4}\right]' & \text{if } \psi < 0 \end{cases}.$$

In OOC#4 only two elements in $\boldsymbol{\sigma}$ are affected, so only a bivariate random direction vector $\boldsymbol{\beta} = [\beta_1, \beta_2]'$ is needed. Then the out-of-control variances can be obtained as $\left[\sigma_1^{(1)}, \sigma_1^{(2)}\right]' = \left[1 + |\psi|\beta_1, \frac{1}{1+|\psi|\beta_2}\right]'$. For each size of the shifts in $\boldsymbol{\sigma}$, the SSATS is obtained by averaging the time to signal from the true change-point for many random draws of $\boldsymbol{\sigma}_1$ with that ψ value.

For the variance charts, it is assumed that the correlations remain unchanged and the variances can only increase when the process is out of control. This means that $\boldsymbol{\Sigma}_1 =$

$[\text{diag}\,\boldsymbol{\sigma}_1]\mathbf{R}_0[\text{diag}\,\boldsymbol{\sigma}_1]$, where $\boldsymbol{\sigma}_1$ can be obtained in the same way as in OOC#1 for the covariance charts and the same ψ is used to denote the size of the increase.

9.4 Comparison of the Covariance Charts

The SSATS values of the covariance charts for the cases of OOC#1 and OOC#2 are presented in Table 9.1. The rows '$m =$', '$\lambda =$' and '$h =$' list the window size for the GLR charts, the tuning parameter values of the MEWMC chart and the MEWMC estimator for the GLR-E chart, and the control limits for each chart to give an in-control ATS of approximately 800. The size of the shift in the variances is indexed by ψ and the change of the correlation structure is indexed by ρ. The value of ρ is always 0 in Table 9.1, as it is assumed that the correlations remain unchanged under OOC#1 and OOC#2. The row with $\psi = 0$ and $\rho = 0$ corresponds to the in-control case, and the in-control ATS value is tabulated. The rows with a positive ψ value correspond to the OOC#1 case where all variance parameters increase, while those with negative values are for the OOC#2 case of decreases in the variances. Columns are labeled with numbers for easy reference.

The columns [1]-[4] are the MEWMC charts with different tuning parameter values. For detecting the variance increase: column [1] has the smallest (of the four) SSATS value for a shift of size 0.2; for small-to-intermediate shift sizes $\psi = 0.4$ and 0.6, column [2] has the shortest detection time; for intermediate-to-large and large variance increases, columns [3] and [4] are better. Each MEWMC chart has a particular range of size of increase in variance for which it is very effective, however it appears that for shift sizes outside that range, its performance can be relatively bad. The same pattern can be observed for the OOC#2 case of variance decreases. Therefore the effect of the tuning parameter value to the MEWMC chart agrees with our familiar property of the EWMA-type charts, a smaller λ value allows better accumulation of information from past sample and thus is effective for small shift magnitudes, while a larger λ allocates relatively more weight on new samples therefore giving

Table 9.1: SSATS Values of the Covariance Charts Under OOC#1 and OOC#2 ($p = 4, \tau = 400$)

			MEWMC				GLR-E				GLR-BC	GLR-V
		m=	-	-	-	-	600	600	600	600	600	600
		λ=	0.01	0.05	0.10	0.20	0.01	0.05	0.10	0.20	-	-
		h=	0.1028	0.6534	1.4268	3.1474	10.7533	13.1900	14.0265	14.8283	18.7163	7.0814
	ψ	ρ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
IC	0.0	0.0	801.73	799.94	799.90	800.06	800.00	799.98	800.02	800.02	799.90	799.98
	0.2	0.0	109.70	143.26	179.39	220.83	121.26	186.88	259.79	347.08	235.33	83.22
	0.4	0.0	44.62	41.29	49.89	67.90	48.71	50.88	64.39	95.41	76.88	28.68
	0.6	0.0	26.54	20.56	22.00	28.11	29.27	25.25	28.11	36.75	39.12	15.21
	0.8	0.0	18.28	13.00	12.84	14.95	20.29	16.11	16.52	19.40	24.43	9.75
OOC	1.0	0.0	13.63	9.26	8.70	9.43	15.21	11.58	11.32	12.39	17.12	6.96
#1	1.5	0.0	7.86	5.10	4.54	4.45	8.84	6.47	6.06	6.02	9.22	3.84
	2.0	0.0	5.26	3.38	2.96	2.79	5.92	4.31	3.98	3.84	6.04	2.57
	3.0	0.0	2.96	1.94	1.71	1.58	3.34	2.46	2.27	2.17	3.37	1.53
	4.0	0.0	1.99	1.36	1.21	1.12	2.23	1.68	1.57	1.50	2.25	1.12
	6.0	0.0	1.19	0.89	0.82	0.78	1.31	1.06	1.00	0.97	1.33	0.78
	9.0	0.0	0.80	0.66	0.63	0.61	0.86	0.74	0.71	0.70	0.87	0.62
	-0.2	0.0	162.09	402.05	887.27	1550.94	144.36	238.06	360.40	536.33	258.09	*
	-0.4	0.0	79.17	105.51	279.44	1231.82	70.58	70.86	102.16	204.37	95.73	*
	-0.6	0.0	56.59	52.76	98.37	548.94	51.46	38.92	46.76	83.53	54.07	*
	-0.8	0.0	46.21	36.72	50.63	227.51	42.58	28.37	29.20	44.65	36.87	*
OOC	-1.0	0.0	40.24	29.41	33.99	108.14	37.39	23.42	21.75	28.74	27.95	*
#2	-1.5	0.0	32.53	21.58	20.68	33.87	30.63	18.00	14.91	15.02	17.65	*
	-2.0	0.0	28.75	18.26	16.37	19.72	27.28	15.60	12.46	10.94	13.26	*
	-3.0	0.0	25.02	15.21	12.91	12.58	23.92	13.33	10.37	8.32	9.35	*
	-4.0	0.0	23.15	13.76	11.40	10.34	22.22	12.21	9.38	7.35	7.58	*
	-6.0	0.0	21.26	12.33	9.96	8.53	20.51	11.09	8.42	6.46	5.91	*
	-9.0	0.0	19.99	11.39	9.04	7.46	19.34	10.34	7.76	5.87	4.83	*

better performance for larger changes. Hawkins and Maboudou-Tchao (2008) mentioned that the MEWMC chart with a large λ can be ATS biased (the case when a control chart gives an out-of-control ATS larger than the in-control ATS, see discussion in Pignatiello, Ascosta-Mejia and Rao (1995) for details) when there is a small decrease in the variance. We can see that this occurs for the MEWMC charts in column [3] and [4], and the issue is quite severe for $\lambda = 0.20$.

The columns [5]-[8] give the SSATS values of the GLR-E charts with the same collection of tuning parameter values. Since the same MEWMC estimator is being used with the GLR change-point formulation, the effect of the tuning parameter value on the chart performance follows the same property as the EWMA-family. By comparing the SSATS values of the GLR-E charts to those corresponding MEWMC charts, it seems that the GLR-E charts are slightly worse for variance increases but somewhat better for variance decreases. Note that the GLR-E charts with $\lambda = 0.01$ and 0.02 are not ATS biased at those variance decrease sizes, whereas the MEWMC charts are biased. It seems that the GLR-E chart is more balanced toward the detection of both variance increases and decreases.

The column [9] is the GLR-BC chart. Unlike the GLR-E chart, the GLR-BC chart is based on the MLE of the out-of-control covariance matrix which connects more naturally to the likelihood ratio tests. However it turns out that the GLR-BC chart is not very effective for detecting most sizes of either variance increases or decreases. It has the fastest detection for very large variance decreases, but it is almost uniformly worse than the GLR-E chart in column [6] for any other variance changes. A possible explanation is that although the Bartlett Correction is used to correct the biasness in the mean of $r_{t,k}$ when $k - t$ is small, there is still a higher possibility to get a value of $r_{t,k}$ from the right tail area, thus the inflation problem still damages the control chart performance. This argument can be supported by comparing the control limits of the GLR-E charts and the GLR-BC chart: the charting statistics of both charts are on the same scale, but the control limit of the GLR-BC chart is considerably higher than those of the GLR-E charts.

The column [10] gives the SSATS values of the GLR-V chart just for reference. As what can be expected, the GLR-V chart specifically designed for only detecting variance increases is able to detect these increases much faster than all the covariance charts. For the OOC#2 case of variance decreases, the GLR-V chart is incapable of detecting any decreases, and those '*' in the table represent SSATS values considerably higher than 800.

It seems that a general conclusion from Table 9.1 is that if someone is primarily concerned with variance increases, then they should consider using the GLR-V chart (or possibly other specifically designed charts for variance increases) because it is so much better than those charts designed for general changes in $\mathbf{\Sigma}$. However, the penalty is that the GLR-V chart will not detect any variance decreases. One option for good overall performance might be to use the GLR-V chart together with the GLR-E chart. The GLR-V chart would quickly respond to variance increases, which may be the primary concern, and the GLR-E chart would provide protection against other changes in $\mathbf{\Sigma}$.

Table 9.2 continues the numerical results for the comparisons of these covariance charts under OOC#3 and OOC#4. Under OOC#3, all pairwise correlation efficient are assumed to change from the in-control independence to ρ, and possibly along with a increase/decrease in the variance parameters. OOC#4 assumes when a special cause occurs, it can only affect the the correlation between two quality variables and their variance parameters, with one increase, and one decrease.

It seems that the MEWMC charts and the GLR-E charts also have similar performances for detecting a change in the correlation structure. The GLR-BC chart is not as effective when compared to the two, and the GLR-V chart is very insensitive to the correlation change. When there is simultaneously an increase in the variance of one variable and a decrease in the other, the GLR-E charts are generally better than the MEWMC charts. And those charts with a not-too-large λ can still effectively detect the change.

If we combine our findings from the two tables, it is not difficult to see that the MEWMC chart in [2] and the GLR-E chart in [6] (both with $\lambda = 0.05$) probably give much

Table 9.2: SSATS Values of the Covariance Charts Under OOC#3 and OOC#4 ($p = 4, \tau = 400$)

			MEWMC				GLR-E				GLR-BC	GLR-V
		m=	-	-	-	-	600	600	600	600	600	600
		λ=	0.01	0.05	0.10	0.20	0.01	0.05	0.10	0.20	-	-
		h=	0.1028	0.6534	1.4268	3.1474	10.7533	13.1900	14.0265	14.8283	18.7163	7.0814
	ψ	ρ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
IC	0.0	0.0	801.73	799.94	799.90	800.06	800.00	799.98	800.02	800.02	799.90	799.98
	0.0	0.1	159.52	266.45	368.00	473.46	161.36	264.60	368.31	489.44	303.48	716.79
	0.5	0.1	31.32	25.71	28.75	37.71	34.38	31.27	36.05	48.97	48.00	20.22
	-0.5	0.1	60.08	58.94	116.00	565.80	54.53	43.39	54.43	100.27	60.64	*
OOC	0.0	0.3	43.24	39.54	50.40	79.50	43.97	38.02	44.86	68.09	54.15	401.68
#3	0.5	0.3	22.40	17.55	18.21	21.96	24.73	20.60	21.55	25.96	29.89	20.14
	-0.5	0.3	40.87	31.14	37.48	93.25	38.60	25.52	25.26	35.64	32.54	*
	0.0	0.5	24.89	18.59	18.99	24.68	25.81	18.26	17.17	19.94	22.58	199.95
	0.5	0.5	15.95	11.93	11.69	12.97	17.56	13.67	13.16	14.15	17.84	19.80
	-0.5	0.5	28.58	19.39	18.63	25.49	27.90	17.20	14.61	15.03	17.90	*
	$\pm$0.2	0.0	137.44	233.36	345.90	463.05	136.72	222.08	322.60	449.68	253.29	346.94
	$\pm$0.2	0.1	125.93	207.39	314.04	432.48	124.66	195.73	289.35	418.92	227.16	346.02
	$\pm$0.2	0.3	81.02	103.56	166.49	271.07	78.82	95.65	140.01	246.23	125.44	334.38
	$\pm$0.2	0.5	53.15	49.67	70.82	129.93	51.70	45.03	57.22	98.36	63.69	310.49
	$\pm$0.2	0.7	38.43	28.93	32.38	55.00	37.53	25.71	26.07	36.89	33.27	271.77
	$\pm$0.6	0.0	42.06	37.11	51.27	99.47	41.92	34.99	41.60	66.51	49.66	130.37
	$\pm$0.6	0.3	38.48	31.86	40.39	76.84	38.39	30.06	33.86	51.38	42.07	129.92
OOC	$\pm$0.6	0.5	33.85	25.90	29.24	50.66	33.78	24.35	25.06	34.61	32.42	129.48
#4	$\pm$0.6	0.7	29.12	20.67	20.59	29.21	29.12	19.29	17.64	20.81	22.65	126.29
	$\pm$1.0	0.0	26.30	19.52	21.03	35.66	26.65	18.93	18.59	23.77	24.42	80.34
	$\pm$1.0	0.3	25.42	18.52	19.31	30.55	25.74	17.99	17.19	21.13	22.58	80.33
	$\pm$1.0	0.7	22.28	15.31	14.31	17.19	22.61	14.84	12.89	13.10	15.74	79.58
	$\pm$2.0	0.0	14.41	9.90	9.12	10.45	14.86	10.05	8.76	8.51	10.56	41.33
	$\pm$2.0	0.7	13.74	9.26	8.28	8.28	14.17	9.43	8.02	7.12	8.70	41.14
	$\pm$4.0	0.0	8.02	5.49	4.88	4.72	8.32	5.73	4.95	4.39	5.25	21.40
	$\pm$4.0	0.7	7.90	5.41	4.78	4.48	8.24	5.64	4.84	4.24	4.80	21.29

better overall performances for monioring all types of changes than the other charts. However it is not an easy task to conclude which one of the two is the best due to the complexity of the covariance matrix structure: the MEWMC chart appears to be better for variance increases, while the GLR-E chart is better for other general changes. Here we use the metric EL that assesses the extra loss during the dectection delay to evaluate the overall performance of the two types of covariance charts.

For the prior distribution of $\mathbf{\Sigma}_1$, $\pi(\mathbf{\Sigma}_1)$, we consider the Wishart distribution for the population of positive definite matrices. We assume that the expected value of $\pi(\mathbf{\Sigma}_1)$ is $\mathbf{I}$ and consider several choices of degree of freedom parameter values: df =5, 10, 20, and 50. Therefore it is indeed assumed that $df\mathbf{\Sigma}_1 \sim$ Wishart $(\mathbf{I}, df)$.

In Figure 9.1, for each choice of df, the scatterplot of the standard deviations for the first two quality variables (after multi-standardization) and the histogram of the correlation coefficient between the two variables based on 10000 randomly sampled $\mathbf{\Sigma}_1$ values from the distribution are plotted. This figure illustrates that df is serving as a dispersion parameter that describes the variability of $\mathbf{\Sigma}_1$: the prior distribution with a small df corresponds to processes where the covariance matrix can vary dramatically from the in-control identity matrix when a shift occurs, while a large df means that $\mathbf{\Sigma}_1$ can only vary within a close neighborhood of $\mathbf{I}$. We will say that df = 5, 10, 20 and 40 each corresponds to large shifts, intermediate-large shifts, small-intermediate shifts, and small shifts, respectively.

Table 9.3: Expected Extra Loss for Selected Charts

	MEWMC			GLR-E		
m=	-	-	-	600	600	600
λ=	0.01	0.05	0.10	0.01	0.05	0.10
h=	0.1028	0.6534	1.4268	10.7533	13.1900	14.0265
df	[1]	[2]	[3]	[5]	[6]	[7]
5	4525.79	2148.90	1938.67	3368.84	1999.76	1693.23
10	83.72	57.83	57.10	81.98	51.84	45.83
20	33.28	27.60	35.69	32.28	24.40	27.30
50	16.66	19.88	32.67	16.05	17.63	24.08

Figure 9.1: The Prior Distributions of $\mathbf{\Sigma}_1$

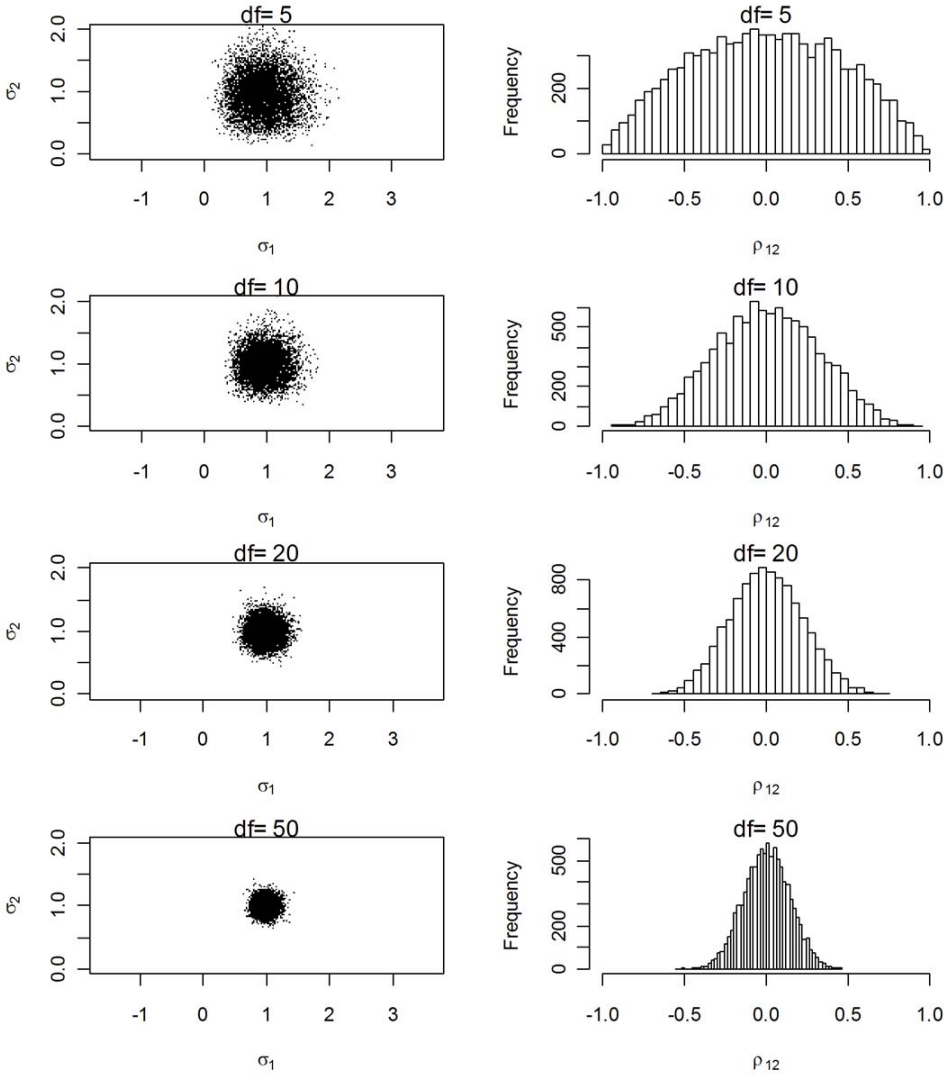


Table 9.3 gives the values of EL under the four choices of *df* for the prior distribution for the MEWMC charts and the GLR-E chart with λ =0.01, 0.05, and 0.10. Note that we continue to use the column labels that appeared in Tables 9.1 and 9.2. For each type of chart, $\lambda = 0.1$ is better for detecting large shifts, $\lambda = 0.05$ is better for intermediate shift and $\lambda = 0.01$ is the most effective for small shifts. In terms of the overall performance over a wide range of shift sizes, the two charts with $\lambda = 0.05$ give either the smallest or close to the smallest EL vaules. With the same tuning parameter values, the GLR-E charts always have a smaller EL than the MEWMC charts. In addtion to Stein's loss function, we have also considered another loss function, the multivariate generalization of the univariate quadratic loss function where $l(\mathbf{\Sigma}_1^{-1}, \mathbf{I}) = \text{tr}(\mathbf{\Sigma}_1^{-1} - \mathbf{I})^2$. Since the conclusion about the relative performance of the charts is the same as what was obtained using EL, the results are not shown here.

9.5 Comparison of the Variance Charts

In this section we evaluate the performance of the control charts designed for detecting only variance increases. Before comparing the GLR-V chart to other existing charts, we have to verify that the GLR-V chart using the CMMLE can approximate the GLR-V chart with the CMLE. Recall in Section 8.3 where we discussed the choices of the estimators for the GLR charts, conditioned on the assumption that the correlation matrix does not change and the variance can not decrease, a natural solution is the CMLE which requires the use of numerical methods to solve for the $\hat{\boldsymbol{\sigma}}_1$ that maximizes $r_{t,k}$. As an alternative, we are proposing to use the CMMLE which has a closed form formula. For the independent case when the correlation matrix is $\mathbf{I}$, the CMMLE is the CMLE; when $\mathbf{R}_0 \neq \mathbf{I}$, the two are different.

In Table 9.4, the SSATS values of the GLR-V chart with the CMMLE and the GLR-V chart with the CMLE for a $p = 4$ dimensional process with the highly correlated correlation

Table 9.4: SSATS Values of the GLR-V Chart with the CMMLE and the GLR-V chart with the CMLE (p=4, τ=200, $\mathbf{R}_0 = \mathbf{R}_{02}$)

	CMMLE	CMLE
m=	600	600
h=	6.3703	6.9813
ψ	[1]	[2]
0.0	400.00	399.81
0.1	110.14	108.85
0.2	43.83	42.54
0.4	15.33	14.82
0.6	8.26	7.90
0.8	5.41	5.19
1.0	3.94	3.74
1.5	2.30	2.19
2.0	1.64	1.56
3.0	1.09	1.03
4.0	0.86	0.81
6.0	0.67	0.64
9.0	0.57	0.56

structure $\mathbf{R}_{02}$ are presented. Note that due to the excessive computational burden for simulating the GLR-V chart with the CMLE, the in-control ATS is set to 400 and the results in this table are based on 50000 simulation runs. For the same data sequence, the statistic of the GLR-V chart with the CMMLE is always less or equal to that of the GLR-V chart with the CMLE, therefore for the two charts to give the same in-control ATS values, the control limit of the latter has to be higher than the former. Comparing the SSATS values, it turns out that the GLR-V chart with the CMLE is uniformly better than using the CMMLE for detecting variance increases.

Note that the computation burden of the GLR-V chart with the CMLE is a problem only when we have to evaluate its performance with a large number of simulations runs. Applying the chart in real-time monitoring is not a problem as there is only one run. Therefore in practical applications, if the user can afford the excessive simulation time to obtain the control limit corresponding to the desired in-control ATS, then the GLR-V chart with

the CMLE should be used to give a uniformly better performance. However considering the fact that the GLR-V chart with the CMMLE requires a significantly shorter simulation time than the GLR-V chart with the CMLE, the minor margin of performance deficiency can still be acceptable. The results in the comparison with other control charts for detecting variance increases are based on one million iterations, therefore the CMMLE is being used with the GLR-V chart.

Table 9.5: SSATS Values of the Variance Charts ($p = 4$, $\tau = 400$, $\mathbf{R}_0 = \mathbf{I}$)

	M_2RZ^2				GLR-V
m=	-	-	-	-	600
λ=	0.01	0.05	0.10	0.20	-
h=	527.4785	163.0868	114.2216	92.8350	7.0814
ψ	[1]	[2]	[3]	[4]	[5]
0.0	800.00	800.00	800.00	800.00	799.98
0.1	153.56	219.55	269.46	329.83	210.96
0.2	66.11	84.72	109.45	148.21	83.12
0.4	27.44	27.01	31.94	43.23	28.67
0.6	16.62	14.34	15.34	19.07	15.22
0.8	11.61	9.42	9.49	10.89	9.74
1.0	8.80	6.89	6.70	7.26	6.95
1.5	5.31	4.00	3.73	3.74	3.83
2.0	3.70	2.76	2.53	2.46	2.57
3.0	2.24	1.68	1.54	1.47	1.54
4.0	1.59	1.23	1.13	1.07	1.12
6.0	1.03	0.84	0.79	0.76	0.78
9.0	0.74	0.65	0.62	0.61	0.62

Next we compare the performance of the GLR-V chart with the CMMLE to other methods when the variables being monitored are uncorrelated. Among the existing methods, the M_2RZ^2 chart is considered the most effective chart when only increases in the variance are of interest. Table 9.5 gives the SSATS values of the GLR-V chart and those of the M_2RZ^2 charts for detecting variance increases when $p = 4$ and $\mathbf{R}_0 = \mathbf{I}$.

The columns labeled [1]-[4] are the M_2RZ^2 charts. Since the M_2RZ^2 chart is an EWMA-type chart, the effect of the tuning parameter value conforms to the property of the EWMA-

family: a small λ allows better information accumulation of past samples, therefore leads to a better ability for detecting small changes; on the other side, a large λ is more effective for large shifts. It seems that each λ has a small range of shift sizes such that the M_2RZ^2 chart gives the smallest SSATS values, but outside the range its SSATS values are relatively larger.

Column [5] gives the SSATS values of the GLR-V chart. Unlike the M_2RZ^2 chart, the GLR-V chart does not have a tuning parameter. It appears that for any shift sizes that we have considered, there is always a M_2RZ^2 chart giving a smaller SSATS value than the GLR-V chart. However considering a wide range of shift sizes, the GLR-V chart always gives the SSATS values close to the best.

Table 9.6: SSATS Values of the Variance Charts (p =4, τ=400, $\mathbf{R}_0 = \mathbf{R}_{02}$)

	M_2RA^2				GLR-V
m=	-	-	-	-	600
λ=	0.01	0.05	0.10	0.20	-
h=	411.2101	134.3700	98.8412	85.9258	7.2801
ψ	[1]	[2]	[3]	[4]	[5]
0.0	800.00	800.00	800.00	800.00	799.99
0.1	157.78	225.50	274.83	330.64	156.10
0.2	62.47	79.16	101.77	135.63	55.40
0.4	23.16	22.02	25.19	32.93	18.19
0.6	12.85	10.86	11.26	13.28	9.55
0.8	8.49	6.82	6.70	7.31	6.14
1.0	6.16	4.83	4.63	4.80	4.43
1.5	3.48	2.70	2.52	2.48	2.54
2.0	2.39	1.86	1.73	1.67	1.79
3.0	1.46	1.18	1.10	1.06	1.16
4.0	1.09	0.91	0.86	0.83	0.90
6.0	0.78	0.69	0.66	0.65	0.69
9.0	0.63	0.58	0.57	0.56	0.58

Next we compare the performance of the variance charts for the high correlation case. For correlated variables, the M_2RA^2 chart based on regression adjusted variables is known to outperform the M_2RZ^2 chart. In Table 9.6, the SSATS values of the GLR-V chart with

the CMMLE and the M_2RA^2 charts are presented for a $p = 4$ dimensional process with the correlation matrix $\mathbf{R}_{02}$ whose off diagonal coefficients are all 0.9.

The columns labeled [1]-[4] are the M_2RA^2 charts with the same tuning parameter values as in the previous table. The same conclusion about the effect of λ can be made. When comparing the results in this table to those in Table 9.5, it appears that the M_2RA^2 chart for detecting the variance increase with correlated variables is generally better than the M_2RZ^2 chart for uncorrelated variables.

Column [5] gives the SSATS values of the GLR-V chart with the CMMLE. It turns out the GLR-V chart gives the best performance for variance increases of size up to $\psi = 1$, for larger shifts, the SSATS values of the GLR-V chart are very close to the smallest. With a fast algorithm for solving CMLE, the GLR-V can be expected to give even better performance.

From the comparison of the variance charts when only the increase in the variability is of interest, we can conclude that the GLR-V chart provides a very attractive option to practitioners in that it does not require specifically choosing a tuning parameter value and that it can effectively detect a wide range of shift sizes.

Chapter 10

Conclusions and Discussion

In this dissertation we have shown that for monitoring the mean vector of a multivariate normal process, the GLR chart and some carefully designed multi-MEWMA combinations are effective for detecting a wide range of shift sizes. The multi-MEWMA scheme requires sophisticated determinations of tuning parameter values and control limits, whereas the GLR chart does not require tuning from the user. The GLR chart also has the advantage that estimates of the change-point and the out-of-control mean vector are immediately available from the calculation of the GLR statistic, and these estimates can be helpful with process diagnostics after a signal is given. A set of formulas are provided for calculation of the control limit of the GLR chart associated with the user specified in-control ATS for processes with $p = 1, \cdots, 30$ quality variables. Thus, the GLR control chart provides an easy-to-use option that will effectively detect a wide range of shifts in the mean of a multivariate normal process. The application the GLR chart with a variable sample interval has also been discussed. Guidance with the design of the VSI GLR chart has been provided.

This dissertation also discussed three GLR charts for detecting changes in the dispersion of multivariate normal processes based on three different estimators: the GLR-BC chart uses the MLE with a Bartlett Correction for small samples; the GLR-E chart is based on a MEWMC estimator; and the GLR-V chart is based on the constrained variance estimator

for the problem of detecting only increases in the variability. Performance comparisons show that the GLR-E chart is at least as good as existing methods for detecting general changes in the covariance matrix and the GLR-V chart provides a similar conclusion as in the case of monitoring the mean vector, in that the GLR-V chart is almost as effective as other existing methods for detecting change sizes which these charts are specifically tuned to detect, and outperforms these charts in terms of the overall performance over a wide range of shift sizes.

Current work in this dissertation only considers the monitoring of the mean vector and the covariance separately. Although intuitively the two types of GLR charts can be applied together to jointly monitor both parameters of the MN process, the evaluation and discussion of such procedures have to be addressed. Moreover, a single GLR chart can be developed to integrate the monitoring of both parameters. This dissertation assumes an ideal case that the in-control parameters are known or can be estimated without any error. However in many applications, this assumption is not always satisfied. An integration of both the Phase-I and Phase-II monitoring could be really helpful in practical applications.

Bibliography

[1] Alt, F. B. (1984). "Multivariate Quality Control". in S. Kotz, N. L. Johnson, and C. R. Read (Eds.), *The Encyclopedia of Statistical Sciences*, pp. 110-122. John Wiley, New York, NY.

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

[3] Basseville, M. and Nikiforov, I. (1993). "Detection of Abrupt Changes: Theory and Applications", Prentice Hall, Englewood cliffs, NJ.

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

[5] Crosier, R. B. (1988). "Multivariate Generalizations of Cumulative Sum Quality-Control Schemes". *Technometrics* 30, pp. 291-303.

[6] Hawkins, D. M. (1991). "Multivariate Quality Control Based on Regression-Adjusted Variables". *Technometrics* 33, pp. 61-75.

[7] Hawkins, D. M. (1993). "Regression Adjustment for Variables in Multivariate Quality Control". *Journal of Quality Technology* 25, pp. 170-182.

[8] Hawkins, D. M. and Olwell, D. H. (1998). "Cumulative Sum Control Charts and Charting for Quality Improvement". Springer-Verlag, New York, NY.

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

[10] Hawkins, D. M. and Maboudou-Tchao, E. M. (2008). "Multivariate Exponentially Weighted Moving Covariance Matrix". *Technometrics* 50, pp. 155-166.

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

[12] Hotelling, H. (1947). "Multivariate Quality Control - Illustrated by the Air Testing of Sample Bombsights". *Techniques of Statistical Analysis*. McGraw-Hill, New York, NY.

[13] Huwang, L.; Yeh, A. B. and Wu, C. W. (2007). "Monitoring Multivariate Process Variability for Individual Observations". *Journal of Quality Technology* 39, pp. 258-278.

[14] Lai, T. L. (1995). "Sequential Changepoint Detection in Quality Control and Dynamical Systems" (with discussion). *Journal of the Royal Statistical Society* Ser. B, 57, 613-658.

[15] Lai, T. L. (1998). "Information Bounds and Quick Detection of Parameter Changes in Stochastic Systems". *IEEE Transactions on Information Theory* 44, 2917-2929.

[16] Lai, T. L. (2001). "Sequential Analysis: Some Classical Problems and New Challenges" (with discussion). *Statistica Sinica* 11, pp. 303-408.

[17] Lowry, C. A.; Woodall, W. H.; Champ, C. W. and Rigdon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart". *Technometrics* 34, pp. 46-53.

[18] Lowry, C. A. and Montgomery, D. C. (1995). "A Review of Multivariate Control Charts". *IIE Transactions* 27, pp. 800-810.

[19] MacGregor, J. F. and Harris, T. J. (1993). "The Exponentially Weighted Moving Variance". *Journal of Quality Technology* 25, Issue 3, pp. 106-118.

[20] MacGregor, J. F. and Kourti, T. (1995). "Statistical Process Control of Multivariate Processes". *Control Engineering Practice* Volume 3, Issue 3, pp. 403-414

[21] 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, pp. 140-143.

[22] Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*, 6th edn, John Wiley & Sons, New York.

[23] Page, E. S. (1954). "Continuous Inspection Schemes". *Biometrika* 41, pp. 100-115.

[24] Pignatiello, J. J., Jr. and Runger, G. C. (1990). "Comparisons of Multivariate CUSUM Charts". *Journal of Quality Technology* 22, pp. 173-186.

[25] Pignatiello, J. J., Jr.; Ascosta-Mejia, C. A. and Rao, B. V. (1995). "The Performance of Control Charts for Monitoring Process Dispersion" in *Proceedings of the Fourth Industrial Engineering Research Conference, eds.* B. Schmeiser and B. Bidanda, Nashville, TN: Institute of Industrial Engineers, pp. 320-328.

[26] Reynolds, M. R., Jr. "Evaluating Properties of Variable Sampling Interval Control Charts". *Sequential Analysis* 14, pp. 59-97.

[27] Reynolds, M. R., Jr. and Cho, G. Y. (2006). "Multivariate Control Charts for Monitoring the Mean Vector and Covariance Matrix". *Journal of Quality Technology* 38, pp. 230-253.

[28] Reynolds, M. R., Jr. and Cho, G. Y. (2011). "Multivariate Control Charts for Monitoring the Mean Vector and Covariance Matrix with Variable Sampling Intervals". *Sequential Analysis* 30, pp. 1-40.

[29] Reynolds, M. R., Jr. and Kim, K. (2007). "Multivariate Control Charts for Monitoring the Process Mean and Variability using Sequential Sampling". *Sequential Analysis* 26, pp. 283-315.

[30] Reynolds, M. R., Jr. and Lou, J. (2010). "An Evaluation of a GLR Control Chart for Monitoring the Process Mean". *Journal of Quality Technology* 42, pp. 287-310.

[31] Reynolds, M. R., Jr.; Lou, J.; Lee, J. and Wang, S. (2011). "The Design of GLR Control Charts for Monitoring the Process Mean and Variance". *Journal of Quality Technology* submitted for publication.

[32] Reynolds, M. R., Jr. and Stoumbos, Z. G. (2004a). "Control Charts and the Efficient Allocation of Sampling Resources". *Technometrics* 46, pp. 200-214.

[33] Reynolds, M. R., Jr. and Stoumbos, Z. G. (2004b). "Should Observations be Grouped for Effective Process Monitoring?". *Journal of Quality Technology* 36, pp. 343-366.

[34] Reynolds, M. R., Jr. and Stoumbos, Z. G. (2005). "Should Exponentially Weighted Moving Average and Cumulative Sum Charts be Used with Shewhart Limits?". *Technometrics* 47, pp. 409-424.

[35] Reynolds, M. R., Jr. and Stoumbos, Z. G. (2008). "Combinations of Multivariate Shewhart and MEWMA Control Charts for Monitoring the Mean Vector and Covariance Matrix". *Journal of Quality Technology* 40, pp. 381-393.

[36] Rigdon, S. E. (1995a). "A Double-Integral Equation for the Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart". *Statistics and Probability Letters* 24, pp. 365-373.

[37] Rigdon, S. E. (1995b). "An Integral Equation for the In-Control Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart". *Journal of Statistical Computation and Simulation* 52, pp. 351-365.

[38] Roberts, S. W. (1959). "Control Chart Tests Based on Geometric Moving Averages". *Technometrics* 1, pp. 239-250.

[39] Runger, G. C. and Pignatiello, J. J., Jr. (1991). "Adaptive Sampling for Process Control". *Journal of Quality Technology* 23, pp. 135-155.

[40] Runger, G. C. and Prabhu, S. S. (1996). "A Markov Chain Model for the Multivariate Exponentially Weighted Moving Average Control Chart". *Journal of the American Statistical Association* 91, pp. 1701-1706.

[41] Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*, Van Nostrand, New York.

[42] Stoumbos, Z. G.; Mittenthal, J. and Runger, G. C. (2001) "Steady-State-Optimal Adaptive Control Charts Based on Variable Sampling Intervals". *Stochastic Analysis and Applications* 19, pp. 1025-1057.

[43] Stoumbos, Z. G.; Reynolds, M. R., Jr.; Ryan, T. P. and Woodall, W. H. (2000) "The State of Statistical Process Control as We Proceed into the 21st Century". *Journal of the American Statistical Association* 95, pp. 992-998.

[44] Sullivan, J. H. and Woodall, W. H. (2000) "Change-point detection of mean vector or covariance matrix shifts using multivariate individual observations". *IIE Transactions on Quality and Reliability Engineering* 32, pp. 537-549.

[45] Willsky, A. S. and Jones, H. L. (1976). "A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems". *IEEE Transactions on Automatic Control* 21, pp. 108-112.

[46] Wortham, A. W. and Ringer, L. J. (1971). "Control via Exponential Smoothing". *Transportation and Logistics Review* 7, pp. 33-39.

[47] Woodall, W. H. (2000). "Controversies and Contradictions in Statistical Process Control". *Journal of Quality Technology* 32, pp. 341-378.

[48] Woodall, W. H. and Montgomery, D. C. (1999). "Research Issues and Ideas in Statistical Process Control". *Journal of Quality Technology* 31, pp. 376-386.

[49] Woodall, W. H. and Ncube, M. M. (1985). "Multivariate CUSUM Quality Control Procedures". *Technometrics* 27, pp. 285-292.

[50] Yeh, A. B.; Huwang, L. and Wu. Y.-F. (2004). "A Likelihood Ratio Based EWMA Control Chart for Monitoring Variability of Multivariate Normal Process". *IIE Transactions on Quality and Reliability Engineering* 36, pp. 865-879.

[51] Yeh, A. B.; Lin, D. K.-J. and McGrath R. N. (2006). "Multivariate Control Charts for Monitoring Covariance Matrix: A Review". *Quality Technology and Quantitative Management* 3, pp. 415-436.

[52] Zamba, K. D. and Hawkins, D. M. (2006). "A Multivariate Change-Point Model for Statistical Process Control". *Technometrics* 48, pp. 539-549.

[53] Zamba, K. D. and Hawkins, D. M. (2009). "A Multivariate Change-Point Model for Change in Mean Vector and/or Covariance Structure". *Journal of Quality Technology.* 41, 285-303.