Quality Technology & Quantitative Management
Vol. 10, No. 1, pp. 115-132, 2013

QTQM

A Directional Multivariate Sign EWMA Control Chart

Xuemin Zi[1], Changliang Zou[2], Qin Zhou[3] and Jingshen Wang[2]

[1]School of Science, Tianjin University of Technology and Education, Tianjin, China
[2]Department of Statistics, Nankai University, Tianjin, China
[3]School of Mathematical Sciences, Jiangsu Normal University, Xuzhou, China
(*Received* April 2012*, accepted* October 2012)

Abstract: In many applications the shift directions of observation vectors are limited, which allows focusing detection power on a limited subspace with improved sensitivity. This paper develops a new multivariate nonparametric statistical process control chart for monitoring location parameters, which is based on integrating a directional multivariate spatial-sign test and exponentially weighted moving average control scheme to on-line sequential monitoring. The computation speed of the proposed scheme is fast with a similar computation effort to its parametric counterpart, regression-adjusted control charts. It has a distribution-free property over a broad class of population models, which implies the in-control run length distribution can attain or is always very close to the nominal one when using the same control limit designed for a multivariate normal distribution. This proposed control chart possesses some other appealing features. Simulation studies show that it is efficient in detecting small or moderate shifts, when the process distribution is heavy-tailed or skewed. Finally, a specific SPC example, multistage process control, is also presented to demonstrate the effectiveness of our method.

Keywords: Distribution-free, multivariate regression-adjusted chart, nonparametric procedure, robustness, spatial sign test, statistical process control.

1. Introduction

Statistical process control (SPC) has been widely used to monitor various industrial processes. In many applications, the quality of a product is often related to several correlated quality characteristics. In modern SPC, it becomes common to monitor several quality characteristics of a process simultaneously in the areas but not limited to signal processing, network security, image processing, genetics, stock marketing and other economic problems. This is called multivariate statistical process control (MSPC) in the literature, which is also the focus of this paper. See [2, 31, 37, 40] for discussion. With newly developed advancement in data acquisition systems and computing technologies, MSPC can and should play a greater role in monitoring and improving manufacturing and service processes.

One of the tasks of MSPC is to detect the change in a multivariate process location vector of parameters μ (mean, median or some percentile of the distribution) as quickly as possible. To be more specific, it is usually assumed that there are m_0 independent and identically distributed (i.i.d.) historical (reference) observations, $\mathbf{x}_{-m_0+1},\ldots,\mathbf{x}_0 \in R^p$, for some integer, $p \geq 1$, and the i th future observation, $\mathbf{x}_i$, is collected over time from the following multivariate change-point model

$$\mathbf{x}_i \overset{\text{i.i.d.}}{\sim} \begin{cases} F_0(\mathbf{x}-\mu_0), & for \quad i=-m_0+1,\ldots,0,1,\ldots,\tau, \\ F_0(\mathbf{x}-\mu_1), & for \quad i=\tau+1,\ldots, \end{cases} \tag{1}$$

where τ is the unknown change point and $\mu_0 \neq \mu_1$. Methods for accomplishing these tasks are usually derived under the assumption that the observed measurement vectors $\mathbf{x}_i$ are distributed $N_p(\mu_0, \Sigma)$ for $i = 1, 2, \ldots, \tau$, and $N_p(\mu_1, \Sigma)$ for $i = \tau + 1, \ldots, n$, with μ_0 and Σ known. Throughout we take $\mu_0 = 0$ without loss of generality. Then a portmanteau test for detecting a mean shift is based on testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$, using the likelihood ratio test (LRT) statistic $\mathbf{x}_i{}^T \Sigma^{-1} \mathbf{x}_i \equiv \| \mathbf{x}_i \|^2_{\Sigma^{-1}}$. This procedure assumes that the covariance matrix does not change. Replacing Σ with some sample covariance matrix, $\mathbf{S}$, results in the Hotelling's T^2 statistic. Based on such test statistics, several MSPC control charts have been proposed, in the framework of cumulative sum (CUSUM) or exponentially weighted moving average (EWMA). Most charting statistics take quadratic forms of the related test statistics; for instance, [6, 11, 19, 25-26, 38].

MSPC control charts with quadratic charting statistics are powerful when one is interested in detecting shifts that occur in majority components of μ. In practice, however, shifts often occur in only a few of the mean components. In such cases, more powerful control charts are possible. Hawkins [8-9] suggested a multivariate control chart, the regression-adjusted scheme, and showed that this chart was more effective than the one based on T^2 when the potential shift occurs in only a few measurement components. Zou and Qiu [41] showed that the regression-adjusted chart is essentially equivalent to the chart based on the LRT test for the case when only one measurement component shifts but the component index is unknown (see Section 2 for a related discussion). Hereafter, the charting techniques based on integrating several known directions will generically be termed as directional control charts. Besides general applications in MSPC, the idea of directional control charts has been applied to several SPC fields; see [43, 45] for the monitoring of multistage processes, [14] for change detection in autocorrelated processes and [35] for multi-sensor change-point detection problem.

All these works are based on an assumption that the process distribution is completely known with multivariate normal distribution. However, it is well recognized that, in many applications, the underlying process distribution is unknown and not multivariate normal, so that statistical properties of these control charts, designed to perform best under the normal distribution, could potentially be (highly) affected. The problem of performance deterioration due to the non-normality is severe with small samples, particularly individual observation cases (c.f., [22]) since the cental limit theorem is no longer (approximately) valid. Nonparametric or robust charts would be useful in such situations. In the last several years, univariate nonparametric control charts have attracted much attention from researchers and a nice overview of this topic was presented by [4]. Some effort has been devoted to multivariate nonparametric control schemes, such as the control schemes based on data-depth ([18]), support vector machines ([34]), antiranks ([27]) or bivariate sign ([7]). Recently, Zou and Tsung [44] developed a multivariate control chart, multivariate sign EWMA (MSEWMA), based on integrating spatial-sign invariant test of [28] and EWMA control scheme. Their method can be viewed as a nonparametric T^2 -type charting technique (c.f., Section 2).

Motivated by [44] use of spatial sign, this paper develops a new multivariate SPC methodology for monitoring location parameters. This methodology adapts directional spatial-sign test to on-line sequential monitoring by incorporating the EWMA scheme. Our proposed new chart has the following positive features: (i) It has distribution-free property over a broad class of population models in the sense that the IC run length distribution can attain or is always very close to the nominal one when using the same control limit designed for a multivariate normal distribution; (ii) Its computation speed is fast with a similar computation effort to the regression-adjusted chart; (iii) It can be easily designed and

constructed because only the multivariate median and the transformation matrix need to be specified from the reference data set before monitoring; (iv) Compared with Hawkins's regression-adjusted chart, it is efficient in detecting small or moderate process shifts when the process distribution is heavy-tailed or skewed. In comparison with MSEWMA, the proposed scheme is more effective when potential changes occur in only very few components of a high-dimensional monitoring system. The rest of this paper is organized as follows: our proposed methodology is described in detail in Section 2. Its numerical performance is thoroughly investigated in Section 3. In Section 4, we demonstrate the method using an example from manufacturing processes. Conclusions and extensions are given in Section 5.

2. Methodology

2.1. Multivariate Directional Tests Using Spatial-Sign

The monitoring problem is closely related to parametric and nonparametric statistical tests of hypotheses for the one-sample location problem in the context of multivariate statistical analysis. Hence, to facilitate the derivation of the proposed charting statistic, we start by assuming that $\mathbf{x}_1,\ldots,\mathbf{x}_n$ are i.i.d from $F(\mathbf{x}-\theta)$, where $F(\cdot)$ represents a continuous p-dimensional distribution "located" at the vector θ. We want to test the null hypothesis, H_0, that $\theta=\theta_0$ against H_1 that $\theta\neq\theta_0$. Without loss of generality, we assume that $\theta_0=0$. Otherwise, we substitute $\mathbf{x}_i-\theta_0$ in place of $\mathbf{x}_i$.

In creating tests for this problem, different levels of assumption have been proposed for the distribution of the $\mathbf{x}_i$'s. Under the multinormality assumption, the classical parametric test, the generalized likelihood ratio test (GLRT; [1]) for testing if the mean $\theta=0$ results in the Hotelling's T^2 statistic as mentioned in Section 1. Sometimes we do have some knowledge about the alternative hypothesis, such as its specific direction, $\theta=\delta\mathbf{d}$, where the direction, $\mathbf{d}$, is a known vector but the scale, δ, is an unknown constant. In such a case, the hypothesis test would be $H_0:\theta=0$ versus $H_1:\theta=\delta\mathbf{d}$. For this problem, the likelihood ratio test is considered to be the uniformly most powerful unbiased test ([17]), which assumes that the underlying population is p-variate normal. It rejects H_0 in favor of H_1 if $V_{\mathbf{d}}\equiv n(\mathbf{d}^T\Sigma^{-1}\bar{\mathbf{x}})^2/\|\mathbf{d}\|^2_{\Sigma^{-1}}$ is large enough.

Furthermore, in a more practical situation, we may consider the following hypothesis test,

$$H_0:\theta=0\leftrightarrow H_1:\theta=\delta\mathbf{d}_1 or\theta=\delta\mathbf{d}_2\cdots or\theta=\delta\mathbf{d}_r, \tag{2}$$

where the alternative hypothesis has several possible directions, $\mathbf{d}_1,\mathbf{d}_2,\ldots,\mathbf{d}_r$, with known vectors. To consider the case r is finite, the generalized likelihood ratio test (under normality) will reject the null hypothesis if the test statistic,

$$\max_{i=1,\cdots,r} V_{\mathbf{d}_i} \tag{3}$$

is larger than a pre-specified critical value. The efficiency of this test can be expected to be superior to T^2-type tests due to its more sharply focused rejection region. Zou *et al.* [45] establish some good asymptotic superiority of the directional GLRT over the invariant GLRT under certain conditions. It can be easily checked that $\sqrt{V_{\mathbf{d}_i}}$'s are just the regression-adjusted variables defined in [8], where $\mathbf{d}_i$ has a single nonzero component 1 located at the ith position. Hawkins [8] suggested an MSPC control chart using $\max_{i=1,\ldots,p}\sqrt{V_{\mathbf{d}_i}}$ as a charting statistic.

As is well known, the definitions of univariate signs and ranks are based on the ordering of the data. However, a natural ordering of the data points does not exist in the multivariate case. The multivariate concepts of a spatial sign and a spatial rank have been developed accordingly in the literature, see a recent book [23] for a comprehensive introduction. Some key points are given in the following. In one dimension, the sign of an observation is basically its direction ($+1$ or -1) from the origin. In higher dimensions, in this spirit, the spatial sign function is defined as

$$U(\mathbf{x}) = \begin{cases} \| \mathbf{x} \|^{-1} \mathbf{x}, & \mathbf{x} \neq 0, \\ 0, & \mathbf{x} = 0, \end{cases}$$

where $\| \mathbf{x} \| = (\mathbf{x}^T \mathbf{x})^{1/2}$ is the Euclidean length of the vector $\mathbf{x}$. The function value is just a direction (a point in the unit p-sphere) whenever $\mathbf{x} \neq 0$. Applying the spatial sign function to the empirical distribution given by data points $\mathbf{x}_1, ..., \mathbf{x}_n$ produces the so-called spatial sign vectors $\mathbf{u}_i = U(\mathbf{x}_i)$. Clearly, in the univariate case, it reduces to the classical sign statistic $\text{sgn}(x_i)$, where $\text{sgn}(\cdot)$ is the sign function.

Additional published work has investigated location testing problem in the nonparametric setting in the literature. A nice overview on this topic and related references can be found in [24]. Specially, Randles [28] develops a simple multivariate sign test based on the transformation proposed by Tyler [36]. Tyler's transformation, in a population version Γ_0, is defined by the solutions of the following equation:

$$E\left[U(\Gamma \mathbf{x}) U^T (\Gamma \mathbf{x}) \right] = \frac{1}{p} \mathbf{I}_p, \tag{4}$$

where $\mathbf{I}_p$ denotes the $p \times p$ identity matrix. Such $\mathbf{W} \equiv \Gamma_0^T \Gamma_0$ is unique as Tyler showed. Thus, Γ_0 is unique if we define it is a $p \times p$ upper triangular positive-definite matrix with a one in the upper left-hand element. Randles [28] proposes to use $Q_I = np \left\| \bar{\mathbf{v}} \right\|^2$ as an invariant test statistic and H_0 is rejected for large values, where $\mathbf{v}_i = U(\Gamma_0 \mathbf{x}_i)$ and $\bar{\mathbf{v}} = \frac{1}{n} \sum_{i=1}^n \mathbf{v}_i$. Intuitively speaking, the equation (4) aims to make the second moment of the transformed random vector $\mathbf{x}_i$ match that of elliptical distributions. To be more specific, as well known, under elliptical distributions of $\mathbf{x}_i$ and assuming $Cov(\mathbf{x}_i) \propto \mathbf{I}_p$ where $\propto$ indicates that the vectors are the same except for a constant, $E[U(\mathbf{x}_i)] = 0$ and $E[U(\mathbf{x}_i) U^T (\mathbf{x}_i)] = p^{-1} \mathbf{I}_p$. The equation in (4) is to transform $\mathbf{x}_i$ to $\Gamma_0 \mathbf{x}_i$ so that the corresponding spatial-sign vector $U(\Gamma_0 \mathbf{x}_i)$ would perform like a uniform sign vector under null hypotheses (from the viewpoint of the first two moments).

In light of (3), we will propose a directional sign test statistic which allows focusing detection power on a limited subspace with improved sensitivity to the invariant sign test. To this end, we firstly state a useful result as follows.

Result 1. Suppose the underlying distribution is elliptically symmetric with density function $f(\mathbf{x}) = k_p \mid \Sigma_0 \mid^{-\frac{1}{2}} h(\left\| \mathbf{x} - \theta \right\|_{\Sigma^{-1}}^2)$. We have $E(\mathbf{v}_i) \propto \Gamma_0 \theta$.

By [36], we know $\mathbf{W} \propto \Sigma^{-1}$. Then, this proposition can be readily seen by using integration by parts.

Consider the hypothesis test $H_0 : \theta = 0$ versus $H_1 : \theta = \delta \mathbf{d}$. By Result 1, $n^{-1} \sum_{i=1}^n \mathbf{v}_i$ is roughly an estimate of shift direction $\Gamma_0 \mathbf{d} / \| \mathbf{d} \|_{\mathbf{W}}$. By mimicking the directional GLRT, we naturally use

$$Q_d = np (\mathbf{d}^T \Gamma_0^T \bar{\mathbf{v}})^2 / \left\| \mathbf{d} \right\|_{\mathbf{W}}^2 \tag{5}$$

as a directional test statistic. Actually, the Q_d^2 can be regarded as a score test statistic with a specific loss function. Consider the following L_1 distance

$$D(\delta) = \sum_{i=1}^{n} \| \mathbf{x}_i - \delta \mathbf{d} \|_{\mathbf{W}} .$$

Clearly, $\partial D / \partial \delta = \sum_{i=1}^{n} \mathbf{d}^T \mathbf{W}(\mathbf{x}_i - \delta \mathbf{d}) / \| \mathbf{x}_i - \delta \mathbf{d} \|_{\mathbf{W}}$. As a consequence, the score statistic is $S = (\partial D / \partial \delta)_{|_{\delta=0}} = n\mathbf{d}^T \Gamma_0^T \overline{\mathbf{v}}$, where we assume $\theta_0 = 0$ again. By noting $\mathrm{Cov}(\overline{\mathbf{v}}) = (np)^{-1}$, we can obtain the test statistic Q_d after normalizing the score statistic S. In a similar spirit, we can view the invariant test statistic Q_I proposed by [28] as a Wald-type statistic based on the L_1 loss function $D(\cdot)$.

With the help of this heuristic derivation, the difference and connection between Q_d and its parametric counterpart, the regression adjusted statistic V_d, would be more apparent. As well known, V_d is a score test statistic, under multivariate normality assumption or equivalently under the L_2 norm $\sum_{i=1}^{n} \| \mathbf{x}_i - \delta \mathbf{d} \|_{\Sigma^{-1}}^2$. In contrast, Q_d is a spatial-sign version of score test statistic in the sense that the L_1 norm is used instead. Now, consider the testing problem (3). Under H_0, $Q_{d_j} \to \chi_1^2$ as $n \to \infty$. Hence, like the test in (4), it is natural to combine all the r score test statistics Q_{d_j} through taking their maximum, say,

$$M_r = \max_{1 \le j \le r} Q_{d_j} . \tag{6}$$

Corresponding, $\hat{k} = \mathrm{argmax}_{1 \le j \le r} Q_{d_j}$ is an ideal estimate of the index of true shift direction. Similar to their univariate counterparts, the spatial signs based methods are expected and have been shown to be quite robust for various distributions since those methods use the direction of observations from the origin rather than the original magnitudes of observations ([23]). Therefore, we are interested in tackling the monitoring problem (1) using the test statistics M_r.

2.2. A Directional Multivariate Sign EWMA Control Chart

Firstly, we elaborate on the individual observation model, which is an advantage of our proposed control scheme because it is able to handle the case when the sample size is one. The extension to the group case is presented at the end of this subsection. Although the monitoring problem (1) is closely related to the standard hypothesis tests in Section 2.1, they are completely different and distinguished by the fundamental difference between on-line and off-line decision issues (c.f., [37]).

Following the idea of [44], the proposed control scheme contains two steps. The first step is to establish the baseline based on the reference sample, that is to say, to extract information from the sample of size m_0 by obtaining a multivariate center θ_0, and a transformation matrix, Γ_0. This step is similar to that of constructing traditional control charts in which μ_0 and Σ_0 are estimated from the historical data before monitoring. We recommend using [12]'s affine equivariant multivariate median which serves sign-based testing purpose and the by-product of finding such median is just the desired transformation matrix ([44]). The affine equivariant multivariate median, θ_0, and the associated transformation matrix, Γ_0, are defined by the solutions of the following equations:

$$E[U(\Gamma(\mathbf{x}_i - \theta))] = 0, \quad E[U(\Gamma(\mathbf{x}_i - \theta))U^T(\Gamma(\mathbf{x}_i - \theta))] = p^{-1}\mathbf{I}_p, \tag{7}$$

and the corresponding sample version, $(\hat{\theta}_0, \hat{\Gamma}_0)$, is defined by the solution of the sample equations based on m_0 historical observations,

$$\frac{1}{m_0}\sum_{i=-m_0+1}^{0} U(\Gamma(\mathbf{x}_i-\theta))=0,\quad \frac{1}{m_0}\sum_{i=-m_0+1}^{0} U(\Gamma(\mathbf{x}_i-\theta))U^T(\Gamma(\mathbf{x}_i-\theta))=p^{-1}\mathbf{I}_p, \tag{8}$$

where Γ is a $p\times p$ upper triangular positive-definite matrix with a one in the upper left-hand element. In a multivariate normal distribution with mean vector μ_0 and variance-covariance matrix Σ_0, it is easily seen that $\theta_0=\mu_0$ and $\Gamma_0^T\Gamma_0=p^{-1}tr(\Sigma_0)\Sigma_0^{-1}$. In what follows, we use (θ_0,Γ_0) rather than $(\hat{\theta}_0,\hat{\Gamma}_0)$ unless indicated otherwise, as an SPC Phase II convention. It should be emphasized that the simultaneous equations (7) aim to make the first two moments of the transformed random vector $\mathbf{x}_i$ match those of elliptical distributions. In other words, one transforms $\mathbf{x}_i$ to $\Gamma_0(\mathbf{x}_i-\theta_0)$ so that the corresponding spatial-sign vector $U(\Gamma_0(\mathbf{x}_i-\theta_0))$ would perform like a uniform sign vector (from the viewpoint of the first two moments). It does not use its distance from the origin.

In light of (6), after (θ_0,Γ_0) is specified or estimated, for on-line collected observations $\mathbf{x}_i, i=1,2,\ldots$, the second step is to standardize and transform them to obtain the unit vector $\mathbf{v}_i$, i.e., the multivariate spatial sign, through $\mathbf{v}_i=U(\Gamma_0(\mathbf{x}_i-\theta_0))$. With this choice, the unit vectors of the transformed data have a variance-covariance structure like that of a random variable that is uniform on the unit p-sphere when the process is IC. Then, we define an EWMA sequence based on spatial-sign vectors with current individual observation as

$$\mathbf{w}_i=(1-\lambda)\mathbf{w}_{i-1}+\lambda\mathbf{v}_i, \tag{9}$$

where the initial vector, $\mathbf{w}_0$, is usually taken to be $E(\mathbf{w}_0)$ and thus should be 0 due to our definition in (7) and $0<\lambda\le 1$ is the smoothing constant. Finally, similar to M_r, the proposed control chart issues a signal if

$$M_i=\frac{(2-\lambda)p}{\lambda}\max_{1\le j\le r}(\mathbf{d}_j^T\Gamma_0^T\mathbf{w}_i)^2/\|\mathbf{d}_j\|^2_{\Gamma_0^T\Gamma_0}>L, \tag{10}$$

where $L>0$ is a control limit chosen to achieve a specific IC Average Run Length (ARL) (ARL_0). Note that the weighted average (9) reflects the relevance of the data: the more recent observations are more informative for detecting the change and thus getting the larger weights. Hereafter, this chart is referred to as the multivariate directional sign EWMA (MDSE) chart.

After the MDSE chart gives an OC signal at the k th observation, we can use $\mathbf{w}_i^2$ directly to implement a diagnostic procedure. Namely, we can find ζ^*

$$\zeta^*=\arg\max_{1\le j\le r}(\mathbf{d}_j^T\Gamma_0^T\mathbf{w}_i)^2/\|\mathbf{d}_j\|^2_{\Gamma_0^T\Gamma_0} \tag{11}$$

to specify the shifted component. As a side note, when a group of g observations, say $\{\mathbf{x}_{i1},\ldots,\mathbf{x}_{ig}\}$ are taken sequentially from the process at each time point, the MDSE chart can be readily defined in a similar way to (10) by using $g^{-1}\sum_{j=1}^{g}U(\Gamma_0(\mathbf{x}_{ij}-\theta_0))$ instead of $\mathbf{v}_i$ in (9).

2.3. Design and Implementation

Estimating (θ_0,Γ_0) involves iterative routines for MDSE and it is a little more complicated than estimating (μ_0,Σ_0) for traditional parametric schemes. However, by using some efficient algorithms ([12]), the convergence of (θ_0,Γ_0) from the historical data with any practical p and m_0 is guaranteed and is usually quite fast. The detailed algorithm is provided in the Appendix of [44]. Please also refer to that paper for some detailed discussions on computation.

As we will see in Section 3, our proposed MDSE chart is robust under IC with appropriate weights, $\lambda \in (0, 0.2]$, except for very skewed distributions and high dimensional cases. In general, a smaller λ leads to a quicker detection of smaller shifts (c.f., e.g., [20]), which is still valid with MDSE. Based on our simulation results, we suggest choosing $\lambda \in (0.025, 0.2]$ in practice, and using $\lambda \in (0.025, 0.1]$ when there is evidence that the underlying distribution is very skewed. In practice, one could get a rough picture about the departure of the underlying distribution from normality by looking at its marginal distributions. Some graphical tools like histogram or QQ-plot are usually helpful. Some multivariate normality tests can be used as well, such as Mardia's test (c.f., [21]). Very small p-values would indicate that the underlying distribution may be far away from multivariate normal distribution. Some engineering experience or prior knowledge about the process should also be taken into consideration.

In what follows, we summarize two useful properties of the MDSE chart which facilitate its design. The proof of these two results are straightforward based on the proof of Propositions 1 and 2 in [44] and omitted here.

Result 2. When the process is IC, the MDSE chart is location-invariant.

The invariant property here is in the sense that for any constant vector $\mathbf{b}$, the run-length distribution of the MDSE stays the same if the IC observations are distributed as $\mathbf{x} + \mathbf{b}$. This property is intuitively appealing since it ensures that the performance of MDSE is the same for any initial location.

Result 3. Given r fixed directions $\mathbf{d}_j$'s, the MDSE chart is distribution free for the class of distributions with elliptical directions in the sense that its IC run length distribution depends only on Γ_0 rather than the distribution.

In the class of distributions with elliptical directions, random variables are generated via $\mathbf{x}_i = r_i \mathbf{D}\mathbf{u}_i$, where the $\mathbf{u}_i$'s are i.i.d. uniform on the unit p sphere, $\mathbf{D}$ is a $p \times p$ nonsingular matrix, and the r_i's are positive scalars. The elliptical directions family contains all the elliptically symmetric distributions, such as multinormal and multivariate t distributions and certain skewed distributions. This result is particularly useful in determining the control limit, L, because, for any continuous process distribution with elliptical directions, it is determined by Γ_0 as achieving the desired IC run-length distribution. In other words, the control limits for distributions with elliptical directions are the same if their Γ_0's are the same. Hence, we can use the multivariate normal distribution with given Γ_0 (obtained through (8)) to find the control limits. The dependence of L on Γ_0 analogously raises when designing regression-adjusted control charts whose control limits depend on Σ even for multinormal distributions ([8]). This dependence is a particular disadvantage of directional charts relative to invariant charts because we cannot tabulate the control limits for general use. For the same reason, the Markov chain method for ARL approximation developed by [44] is not feasible for MDSE, and thus simulation procedures are necessary. Certainly, this would not produce any difficulty in constructing the MDSE chart by virtue of the massive computing and data storage capabilities of modern computers. Since the searching procedure is a one-time computation before the Phase II online process monitoring, it is convenient to accomplish. Fortran code for implementing the proposed procedure is available from the authors upon request.

As a Phase II SPC convention, it is usually assumed that the IC parameters are known or, equivalently, that they are estimated from a sufficiently large reference dataset. It should be pointed out that when m_0 is not large (say, $m_0 \le 1000$; see Table 3 in Section 3), there would be considerable uncertainty in the parameter estimation, which in turn would distort the IC

run length distribution of the MDSE control chart. From the results, we can see that, as long as $m_0 \geq 2000$, the ARL_0 values are quite stable in various cases. Therefore, we suggest collecting at least 2000 IC observations before Phase II process monitoring. To deal with the situation when a sufficiently large reference data set is unavailable, one possible method is to adjust the control limit of the chart properly by simulation to obtain the desired ARL_0 (cf., [16]). That is, we can generate a pseudo reference sample of size m_0 and then obtain the corresponding run-length with a given control limit. Repeat this procedure to approximate the ARL_0 and then use the bisection search method to find the control limit. Of course, the detection ability would still be severely compromised. This is essentially analogous to the estimated parameters problem in the context of parametric control charts (see [13] for an overview).

3. Numerical Performance

We present some simulation results in this section regarding the numerical performance of the proposed MDSE chart and compare it with its parametric counterpart, REWMA, originally proposed by [8] and adapted by [41]. To be more specific, the chart statistic of REWMA is given by

$$\max_{j=1,\cdots,r}\left\{\frac{(2-\lambda)}{\lambda}(\mathbf{d}_j^T\Sigma_0^{-1}\overline{\mathbf{z}}_i)^2 \,/\, \left\|\mathbf{d}_j\right\|^2_{\Sigma_0^{-1}}\right\},$$

where $\mathbf{z}_i = (1-\lambda)\mathbf{z}_{i-1} + \lambda(\mathbf{x}_i - \mu_0)$. In addition, the invariant control chart, MSEWMA, proposed by [44] is also included in this study for comparison use. The MEWMA chart ([19]) is not considered here for the following reason: its pros and cons relative to REWMA and MSEWMA have been addressed by [41] and [44] in detail, respectively. Hence, based on the results shown below, we can easily get a rough picture of the comparison between MDSE and MEWMA. At the end of this section, we will provide some general guideline for engineers to choose an appropriate chart according to their major concern.

We start by assuming that m_0 is sufficiently large, in this case 30,000. In all the underlying distributions considered, we first generate m_0 i.i.d. samples and then estimate (μ_0, Σ_0) and (θ_0, Γ_0). Control limits of the MDSE and REWMA charts are determined by simulations to attain the nominal ARL_0 under the standard multivariate normal distribution, while the Markov chain approximation is used for MSEWMA and MEWMA ([29, 44]). Since the zero-state and steady-state ARL (SSARL) comparison results are similar, only the SSARLs are provided. To evaluate the SSARL behavior of each chart, any series in which a signal occurs before the $(\tau+1)$ th observation is discarded (c.f., [10]). We only present the results when ARL_0 =200 and $\tau = 50$ for illustration because similar conclusions hold for other cases. All the ARL results in this section are obtained from 40,000 replications. Following the robustness analysis in [32]), we consider the following distributions: (i) multivariate normal, denoted as N_p; (ii) multivariate t with ζ degrees of freedom, denoted as $t_{p,\zeta}$; (iii) multivariate gamma with shape parameter ζ and scale parameter 1, denoted as $\Gamma_{p,\zeta}$. Details on the multivariate t and gamma distributions can be found in the Appendix of [32]. In addition, the following distribution is involved in the comparison: (iv) in each observed vector, the first $[p/2]$ measurement components are i.i.d. t distributed with ζ degrees of freedom and the other $p-[p/2]$ measurement components are i.i.d. chi-square distributed with $[\zeta/2]+1$ degrees of freedom. The reason for considering this distribution is that unlike (i)-(iii), its marginal distributions are not all the same.

Table 1. ARL_0 values of MDSE and REWMA charts under scenarios (ii)-(iv) when $p=10$ or 20.

		MDSE						REWMA					
		$p=10$			$p=20$			$p=10$			$p=20$		
	$\zeta \setminus \lambda$	0.1	0.05	0.025	0.1	0.05	0.025	0.1	0.05	0.025	0.1	0.05	0.025
	3	201	200	198	200	202	200	103	153	209	86.0	137	197
	4	200	200	201	201	199	201	105	159	190	90.1	139	182
	5	202	201	199	200	199	201	119	162	192	100	146	185
(ii)	7	199	201	200	200	201	200	139	175	192	123	164	193
	10	199	200	201	201	200	199	159	184	197	147	177	195
	15	200	200	200	201	202	200	171	191	196	163	184	200
	1	110	167	195	64.7	113	166	52.2	96.4	155	36.3	72.1	126
	2	139	183	197	87.7	138	183	73.2	125	181	52.0	98.6	151
	3	151	187	199	105	153	186	87.5	142	180	65.7	114	165
(iii)	5	167	192	200	126	168	193	107	158	184	83.7	137	176
	10	182	197	200	153	183	200	138	176	192	117	163	189
	15	188	197	200	167	190	201	150	187	195	134	173	192
	3	131	168	189	83.0	135	172	61.2	110	160	43.1	87.1	126
	4	140	169	181	102	143	172	81.2	130	172	58.3	101	153
	5	141	168	184	116	144	173	89.6	140	178	67.3	104	161
(iv)	7	148	170	189	134	150	177	108	156	185	84.0	115	177
	10	161	171	190	150	161	181	128	166	190	106	130	183
	15	165	182	192	160	168	185	143	180	193	121	141	185

3.1. Comparisons Between MDSE and REWMA

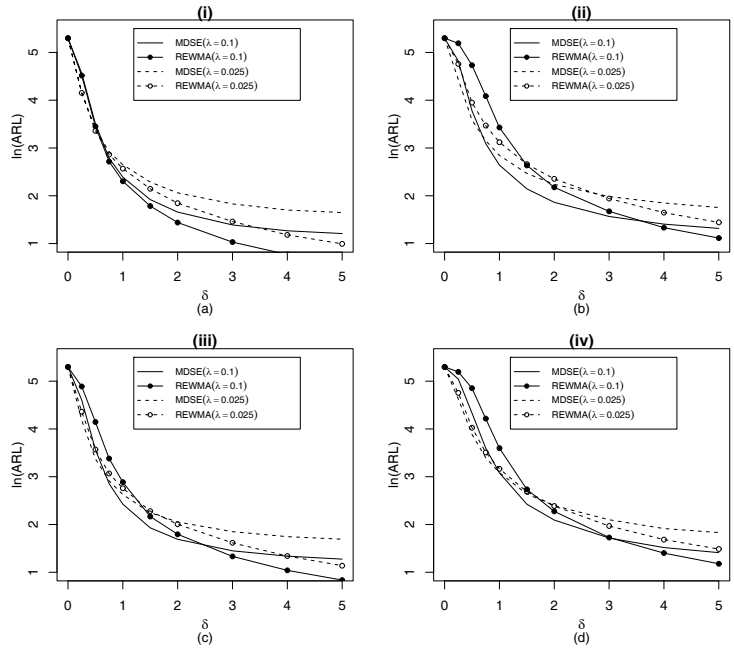

Figure 1. OC ARL comparison of the MDSE and REWMA charts using $\lambda=0.025$ and 0.1 with a shift in the first component and $p=10$ under: (a) multivariate normal distribution; (b) multivariate t distribution; (c) multivariate Gamma distribution; (d) multivariate mixed-components distributions.

In this subsection, the MDSE and REWMA charts are compared in terms of IC and OC ARLs. The number and variety of covariance matrices and shift directions are too large to allow a comprehensive, all-encompassing comparison. Our goal is to show the effectiveness, robustness and sensitivity of the MDSE chart, and thus we only choose certain representative models for illustration. Specifically, for the first three distribution cases, the covariance matrix $\Sigma_0 = (\sigma_{ij})$ is chosen to be $\sigma_{ii} = 1$ and $\sigma_{ij} = 0.5^{|i-j|}$, for $i, j = 1, 2, \ldots, p$. As both the MDSE and REWMA are directional charts, for brevity, a shift of size δ in only the first component is used, i.e., $\mathbf{x}_i + \delta \mathbf{e}_1$ with $\mathbf{e}_1 = (1, 0, \ldots, 0)^T$. We conducted some other simulations with various shift types and the results show that the general conclusions given below do not change.

We firstly evaluate the ARL_0 values under distribution scenarios (ii)-(iv). The control limits of both charts are found by simulation under normality assumption. The ARL_0s of MDSE and REWMA under various cases with different combinations of dimensionality, λ and degrees of freedom ζ are tabulated in Table 1. From this table, we can see that the MDSE is more robust than REWMA to the heavy-tailed and skewed distributions even when the distribution is not in the class of elliptical direction such as the distributions (iii) and (iv). Its ARL_0 is always not far away from the nominal one even for the extremely non-normal and high-dimensional distributions. The REWMA usually has a larger bias in the ARL_0 and the degradation becomes more pronounced as the dimensionality increases.

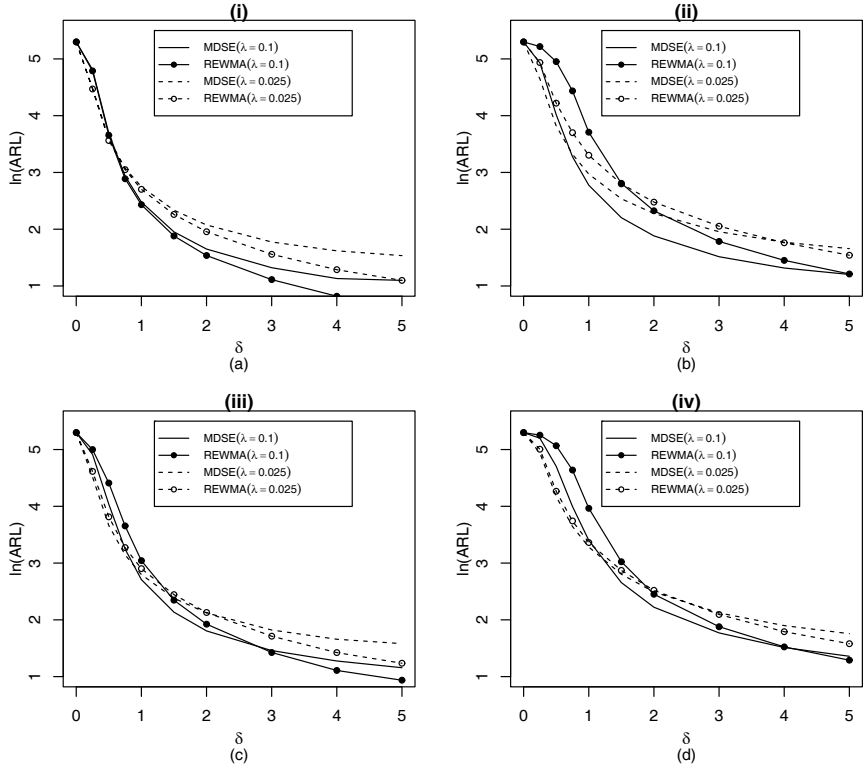

Figure 2. OC ARL comparison of the MDSE and REWMA charts using $\lambda = 0.025$ and 0.1 with a shift in the first component and $p = 20$ under: (a) multivariate normal distribution; (b) multivariate t distribution; (c) multivariate Gamma distribution; (d) multivariate mixed-components distributions.

Next, we turn to Figures 1 and 2, which give OC ARL values (in log-scale) when $p=10$ and $p=20$, respectively. The ζ values in the scenarios (ii)-(iv) are chosen as 4, 2 and 4, respectively. To appreciate the pros and cons of MDSE with respect to REWMA under OC models, we consider a bias-corrected OC ARL comparison in the sense that the control limits of both charts are adjusted to make their IC ARLs equal to the nominal one. The corresponding ARL results with $\lambda=0.1$ and 0.025 are summarized in those two figures. Note that such charting schemes with adjustment are only for comparison use in our simulations but not applicable in practical applications since the error distribution is usually unknown as we claimed before.

The results in the two figures are similar: On one hand, the MDSE chart is more efficient in detecting the small and moderate shifts than the REWMA chart in the sense that with the same λ values, the OC ARLs of the MDSE are generally smaller than those of the REWMA in most of cases. Even for normal distribution (the scenario (i)), the MDSE chart performs quite satisfactorily for small and moderate shifts and the difference between MDSE and REWMA is small, although the REWMA chart has superior efficiency as we would expect since the parametric hypothesis is the correct one in this case. Clearly, with an appropriate value of λ, the MDSE chart is able to outperform the REWMA chart in detecting small and moderate shifts even when the elliptical assumption is not valid (in scenarios (iii) and (iv)), and in some cases the advantage is prominent especially when p is large; On the other hand, the REWMA is more efficient in detecting large shifts, such as $\delta \geq 3.0$. This is understandable because the MDSE, which is essentially based on signs rather than distances, shares a similar drawback as those sign- or rank-based charts for univariate processes. That is, even though the shift is quite large, the ranks of the observations may not be able to grow larger.

3.2. Comparisons Between MDSE and MSEWMA

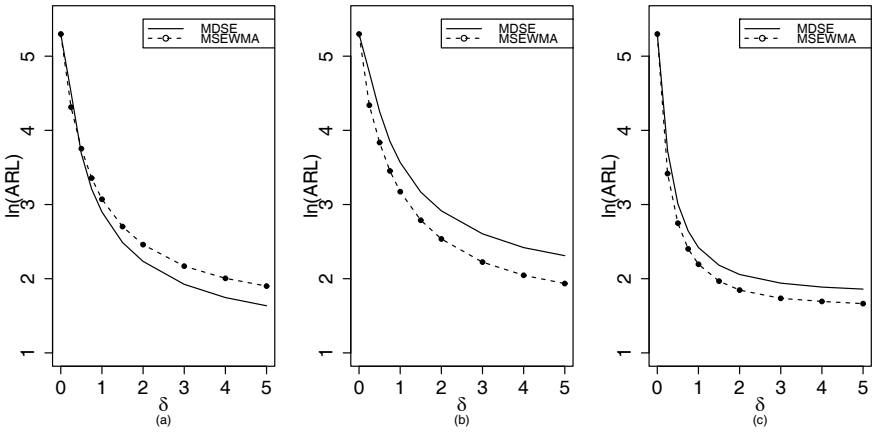

Figure 3. OC ARL comparison of the MDSE and MSEWMA charts using $\lambda=0.025$ under $p=20$ scenario (ii) with: (a) shift in only the first component; (b) the first $[p/2]$ components have a common shift size; (c) linearly increasing shifts in the first $[p/2]$ components

Although the major benefit of MDSE is to be in the situation that the hypothesis (2) is correct and the distribution is far away from normal, it is also interest to compare the MDSE and MSEWMA charts which can provide better understanding of the performance of MDSE and the difference between directional and invariant charting schemes. As mentioned before, both the MSEWMA and MDSE charts are based on spatial-signs and

thus their difference lies in the shift pattern rather than the distribution. So, we only consider to use multivariate t distribution (the scenario (ii)) for illustration. Because a similar conclusion holds for other cases, we only present the results with $\lambda = 0.025$ and $p = 20$ for brevity. Figure 3 shows the OC ARL comparison between MDSE and MSEWMA under the $t_{20,4}$ distributions with the nominal ARL_0 is fixed as 200. As an illustration, three scenarios are considered: (1) shift in only the first component; (2) equal shift, i.e, the first p_a components have a common shift size δ; (3) linearly increasing allocation, say $\delta_i = i\beta$ for $i = 1,\ldots,p_a$. Note that the last two scenarios correspond to the situation the post-change distributions are misspecified for MDSE. To make the ARL performance comparable among the configurations of various scenarios, we set $\sum_{i=1}^{p_a}\delta_i^2 = \delta^2$.

Figure 3 presents the log-ARL curves of the MDSE and MSEWMA charts with p_a being chosen as $[p/2]$. Clearly, MDSE is more effective than MSEWMA when potential changes occur in only one component as we assume in (2). On the other hand, MDSE would be outperformed by MSEWMA by a large margin, as when the change occurs in moderate or large number of components as shown in Figure 3(b)-(c). These findings are consistent with the properties of LRT or score tests with misspecified alternatives ([45]). We conducted other simulations with various correlation structures, p and ARL_0, to see if the above conclusions would change. These simulation results (not reported here but available from authors upon request) show that the conclusion on the performance of MDSE relative to MSEWMA and REWMA would generally hold for other choices of p and ARL_0.

3.3. Diagnostic Performance Analysis

We now compare the proposed diagnostic procedure (11) with the diagnostic method using REWMA. Table 2 summarizes the simulation results under scenarios (i)-(iv) in which the parameters are chosen as the same as those in Figure 2. In this table, the observed frequencies that the diagnostic procedures identify shifted measurement components correctly are reported. Except for a very small shift, such as $\delta = 0.25$, both estimating procedures perform quite well from the viewpoint of correct probabilities. It can be easily observed that MDSE is more accurate than REWMA for the non-normal distributions as we would expect, especially when δ is not large enough. After taking into account its computational effort, we think that the MDSE approach provides a reasonable diagnosis tool for MSPC.

3.4. Effect of m_0

Finally, we study the effect of m_0 on the performance of MDSE and REWMA because in all the foregoing numerical analysis, it is assumed that the IC parameters are estimated from a sufficiently large reference data set (equivalently assume that the IC distribution is known). To this end, we use the multivariate normal and multivariate gamma distributions with two degrees of freedom. Only the case $p = 10$ and $\mathrm{ARL}_0 = 200$ is considered. Table 3 shows the IC ARL and standard deviation of the run length (SDRL) values of MDSE and REWMA when the IC parameters (θ_0, Γ_0) for MDSE and (μ_0, Σ_0) for REWMA are computed from an IC data set with various historical sample sizes, m_0. From this table, it can be seen that (i) when m_0 is relatively small, the actual ARL_0 and SDRL values of the two charts are both quite far away from the nominal level of 200; (ii) when m_0 increases, such biases decrease; (iii) the biases in ARL_0 of MDSE are smaller than REWMA with the same λ (by the results of $N_{10}(0, \Sigma_0)$), although it appears that the chart with the smaller λ has a little larger bias in ARL_0.

Table 2. Diagnosis comparison between MDSE and REWMA.

	(i)				(ii)			
	MDSE		REWMA		MDSE		REWMA	
$\delta \setminus \lambda$	0.1	0.025	0.1	0.025	0.1	0.025	0.1	0.025
0.25	0.40	0.55	0.40	0.56	0.35	0.52	0.11	0.34
0.50	0.78	0.80	0.79	0.80	0.77	0.79	0.34	0.69
0.75	0.90	0.88	0.88	0.88	0.89	0.88	0.65	0.83
1.00	0.94	0.91	0.93	0.91	0.94	0.92	0.82	0.87
1.50	0.96	0.94	0.95	0.93	0.97	0.96	0.93	0.92
2.00	0.98	0.95	0.96	0.94	0.98	0.97	0.96	0.95
3.00	0.99	0.97	0.98	0.96	0.99	0.98	0.98	0.96
4.00	0.99	0.98	0.98	0.97	1.00	0.99	0.98	0.98
5.00	0.99	0.98	0.99	0.98	1.00	0.99	0.99	0.98

	(iii)				(iv)			
	MDSE		REWMA		MDSE		REWMA	
$\delta \setminus \lambda$	0.1	0.025	0.1	0.025	0.1	0.025	0.1	0.025
0.25	0.37	0.26	0.57	0.52	0.14	0.10	0.35	0.29
0.75	0.89	0.89	0.82	0.87	0.75	0.83	0.52	0.80
1.00	0.94	0.93	0.91	0.91	0.87	0.89	0.75	0.87
1.50	0.98	0.96	0.95	0.94	0.94	0.94	0.91	0.91
2.00	0.98	0.97	0.98	0.96	0.97	0.96	0.95	0.94
3.00	0.99	0.99	0.98	0.98	0.99	0.98	0.97	0.97
4.00	1.00	0.99	0.99	0.98	0.99	0.98	0.98	0.97
5.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.98

Table 3. IC ARL and SDRL values with various Phase I sample sizes, m_0. Numbers in parentheses are SDRL values.

	$N_{10}(0,\Sigma_0)$				$\Gamma_{10,2}$			
	MDSE		REWMA		MDSE		REWMA	
$m_0 \setminus \lambda$	0.05	0.025	0.05	0.025	0.05	0.025	0.05	0.025
500	158 (140)	154 (126)	144 (128)	145 (116)	145 (128)	151 (122)	95 (88.9)	119 (101)
750	172 (154)	166 (134)	162 (147)	159 (128)	154 (138)	163 (134)	104 (96.8)	133 (115)
1000	179 (161)	176 (144)	172 (157)	169 (136)	163 (146)	173 (141)	108 (99.1)	143 (122)
1500	188 (171)	186 (156)	184 (167)	182 (153)	166 (150)	178 (149)	115 (105)	150 (129)
2000	194 (176)	192 (159)	186 (169)	188 (159)	170 (154)	183 (154)	115 (107)	156 (136)
3000	199 (183)	199 (168)	194 (176)	193 (162)	173 (154)	186 (156)	119 (108)	163 (143)
5000	201 (180)	200 (170)	202 (183)	198 (167)	181 (163)	196 (165)	121 (111)	167 (147)

4. An Industrial Application: Monitoring in Multistage Process Control

In this section, we present an industrial example in SPC practice, the monitoring in multistage process control, to demonstrate the performance of the proposed framework and compare it with some existing methods. As modern technologies become increasingly sophisticated, most manufacturing operations, such as semiconductor manufacturing and automotive body assembly, comprise multiple stages. Shi and Zhou [30] provide an extensive review of the multistage process control problems with many industrial examples. In these systems, it is often desirable and necessary to design an effective monitoring and

diagnosis approach for detecting, isolating and identifying the sources of a change by linking the current stage signal to information about earlier stages in the serial process. Zhou *et al.* [39] and many others discuss sensor allocations and fault diagnosability in multistage processes. Zou and Tsung [43] and Zou *et al.* [45] investigate multistage process monitoring and diagnosis problems in various settings.

Consider a common manufacturing process comprised of p stages. For the j *th* product collected, a two-level linear state-space model generated from a practical application is usually used to describe the quality measurement to the k *th* stage ([45]): for $k = 1,\ldots, p$, $j = 1,2,\ldots,$,

$$y_{kj} = C_k x_{kj} + v_{kj}$$
$$x_{kj} = A_k x_{k-1j} + w_{kj} + \delta \mathrm{I}_{\{k=\zeta^*, j>\tau\}},$$

where $x_{0j} \sim (a_0, \sigma_\varepsilon^2)$. The v_{kj} and w_{kj} are assumed to be independent from each other and $v_k \sim (\mu_k, \sigma_{v_k}^2)$, $w_k \sim (0, \sigma_{w_k}^2)$. The first level of the model involves the fitting of the quality measurement to the quality characteristic, and C_k is used to relate the unobservable process quality characteristic, x_k, to the observable quality measurement, y_k. The second level of the model involves modeling the transfer of the quality characteristic from the previous stage to the present stage, in which A_k denotes the transformation coefficient of the quality characteristic from stage $k-1$ to stage k. In multistage process applications, A_k and C_k are known constants (or matrices) that are usually derived or estimated from engineering knowledge (see [15] for details). The unknown magnitude, δ, reflects the difference before and after the change-point τ.

As illustrated by [43], under the model assumption above the process shift happens only in one of p known directions , $\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_p$, where

$$\mathbf{d}_1 \propto \begin{pmatrix} c_1 \\ c_2 a_2 \\ \vdots \\ c_p \prod_{i=2}^{p} a_i \end{pmatrix} \quad \mathbf{d}_2 \propto \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \prod_{i=3}^{p} a_i \end{pmatrix} \quad \ldots \mathbf{d}_p \propto \begin{pmatrix} 0 \\ 0 \\ \vdots \\ c_p \end{pmatrix}. \tag{12}$$

Hence, a directional test and accordingly a directional monitoring system is particularly desired in this problem. Under the normality assumption, say v_{kj}'s and w_{kj}'s are all normally distributed, Zou and Tsung [43] implemented an REWMA procedure to deal with the monitoring and diagnosis of multistage processes. In practice, we usually cannot have an assurance of normality, especially when p is large. In those situations, the MDSE scheme seems to be an ideal alternative taking robustness into account. Here, we present a numerical example, setting $p = 20$ and $(A_k, C_k) = (1.2, 0.8)$ which are consistent with the numerical examples in [45]. v_{kj}'s and w_{kj}'s are assumed to be standardized t-distributed with five degrees of freedom (so that they have unit variance). We consider a stage shift at $\zeta = 10$ and $\tau = 50$ with shift magnitude 2.0. To attain a nominal ARL_0 200, the control limits for MDSE and REWMA charts with $\lambda = 0.025$ under normality assumption are 9.59 and 9.61 respectively. The resulting ARL_0 values under the t-distributed assumption are 199 and 181, while the OC ARLs are 31.4 and 45.7 for MDSE and REWMA respectively. This result reflects that the MDSE would be more robust in detecting shifts than the REWMA chart in the sense that even when the value of ARL_0 is larger than that of the REWMA (closer to the nominal one), the OC ARL of the MDSE is still considerably smaller than that of the REWMA.

5. Concluding Remarks

In this paper, we propose a new detection scheme, MDSE, for monitoring multivariate processes with several directions. This procedure is derived from a spatial-sign-based score test and naturally integrates the directional information from processes with the EWMA scheme. With updating formulations, the proposed scheme is fast to compute with a similar computational effort to existing schemes. Compared with existing parametric methods, this scheme is not only more robust in IC performance, but it is also generally more sensitive to the small and moderate shifts in location parameters for skewed and heavy-tailed multivariate distributions. In many cases, the improvement is quite remarkable.

One drawback of this scheme, which is common to almost all rank-based (univariate) nonparametric schemes, is that it is not as efficient as parametric schemes for very large shifts because it only uses the direction of observations from the origin rather than the original magnitudes of observations. This disadvantage is mainly due to the trade-off between robustness and sensitivity. It is also worth pointing out that while a directional control chart offers a better performance against shift occurring along known, a priori given, directions, it can have zero, or almost zero, power against shifts happening along others, unforeseen, directions. Further, the diagnostic procedures described in Section 2 can give misleading results if the actual shift direction is not comprised in the assumed possible directions. The effectiveness of the proposed MDSE chart and associated diagnosis relies on the correct specification of shift directions, and hence some physical knowledge and experience from engineers are always necessary in practical applications.

To deal with the situation when a sufficiently large reference data set is unavailable, self-starting methods that handle sequential monitoring by simultaneously updating parameter estimates and checking for OC conditions have been developed accordingly. See, e.g., [3], [11], [33] under normality assumption and [42] for nonparametric settings. Some studies on the development of corresponding directional nonparametric self-starting charts are beyond the scope of this paper but warrant future research. In addition, one of our ongoing research is to consider the economic design of multivariate nonparametric charts ([5]).

Acknowledgements

The authors thank the editor, associate editor, and two anonymous referees for their many helpful comments that have resulted in significant improvements in the article. This research was supported by the NNSF of China Grants 11101306, 11001138, 11071128, 11131002, 11001083, 11201189, the RFDP of China Grant 20110031110002 and the Fundamental Research Funds for the Central Universities. Zou also thanks the support of the National Center for Theoretical Sciences, Math Division. Zhou's work is partially supported by PAPD of Jiangsu Higher Education Institutions; the scientific research support project for teachers with doctor's degree, Jiangsu Normal University under Grant No.11XLR32. Zi's work is partially supported by TJUTE Research Development Grants YY09003.

References

1. Anderson, T. W. (1984). *An Introduction of Multivariate Statistical Analysis, 2nd ed*. Wiely, New York.

2. Bersimis, S., Psarakis, S. and Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International*, 23, 517-543.

3. Capizzi, G. and Masarotto, G. (2010). Self-starting CUSCORE control charts for individual multivariate observations. *Journal of Quality Technology*, 42, 136-151.

4. Chakraborti, S., Van der Laan, P. and Bakir, S. T. (2001). Nonparametric control charts: an overview and some results. *Journal of Quality Technology*, 33, 304-315.

5. Cheng, S. W. and Mao, H. (2011). The economic design of multivariate MSE control chart. *Quality Technology and Quantitative Management*, 8, 75-85.

6. Crosier, R. B. (1988). Multivariate generations of cumulative sum quality control schemes. *Technometrics*, 30, 219-303.

7. Das, N. (2009). A new multivariate non-parametric control chart based on sign test. *Quality Technology and Quantitative Management*, 6, 155-169.

8. Hawkins, D. M. (1991). Multivariate quality control based on regression-adjusted variables. *Technometrics*, 33, 61-75.

9. Hawkins, D. M. (1993). Regression adjustment for variables in multivariate quality control. *Journal of Quality Technology*, 25, 170-182.

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

11. Hawkins, D. M. and Maboudou-Tchao, E. M. (2007). Self-starting multivariate exponentially weighted moving average control charting. *Technometrics*, 49, 199-209.

12. Hettmansperger, T. P. and Randles, R. H. (2002). A practical affine equivariant multivariate median. *Biometrika*, 89, 851-860.

13. Jensen, W. A., Jones-Farmer, L. A., Champ, C. W. and Woodall, W. H. (2006). Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality Technology*, 38, 349-364.

14. Jiang, W. (2004). Multivariate control charts for monitoring autocorrelated processes. *Journal of Quality Technology*, 36, 367-379.

15. Jin, J. and Shi, J. (1999). State space modeling of sheet metal assembly for dimensional control. *ASME Transactions, Journal of Manufacturing Science and Engineering*, 121, 756-762.

16. Jones, L. A. (2002). The statistical design of EWMA control charts with estimated parameters. *Journal of Quality Technology*, 34, 277-288.

17. Lehmann, E. L. (1986). *Testing Statistical Hypothesis, 2nd ed.* Wiley, New York.

18. Liu, R. (1995). Control charts for multivariate processes. *Journal of the American Statistical Association*, 90, 1380-1388.

19. Lowry, C. A., Woodall, W. H., Champ, C. W. and Rigdon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34, 46-53.

20. Lucas, J. M. and Saccucci, M. S. (1990). Exponentially weighted moving average control scheme properties and enhancements. *Technometrics*, 32, 1-29.

21. Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.

22. Montgomery, D. C. (2005). *Introduction to Statistical Quality Control, 6th edition*. John Wiley & Sons, New York.

23. Oja, H. (2010). *Multivariate Nonparametric Methods with R*. Springer Science+Business Media, New York.

24. Oja, H. and Randles, R. (2004). Multivariate Nonparametric Tests. *Statistical Science*, 19, 598-605.

25. Pignatiello, J. J. and Runger, G. C. (1990). Comparison of multivariate CUSUM charts. *Journal of Quality Technology*, 22, 173-186.

26. Prabhu, S. S. and Runger, G. C. (1997). Designing a multivariate EWMA control chart. *Journal of Quality Technology*, 29, 8-15.

27. Qiu, P. and Hawkins, D. M. (2001). A rank based multivariate CUSUM procedure. *Technometrics*, 43, 120-132.

28. Randles, R. H. (2000). A simpler, affine invariant, multivariate, distribution-free sign test. *Journal of the American Statistical Association*, 95, 1263-1268.

29. Runger, G. C. and Prabhu, S. S. (1996). A markov chain model for the multivariate exponentially weighted moving averages control chart. *Journal of the American Statistical Association*, 91, 1701-1706.

30. Shi, J. and Zhou, S. (2009). Quality control and improvement for multistage systems: a survey. *IIE Transactions*, 41, 744-753.

31. Stoumbos, Z. G., Reynolds, M. R., Ryan, T. P. and Woodall, W. H. (2000). The state of statistical process control as we proceed into the 21st century. *Journal of American Statistical Association*, 95, 992-998.

32. Stoumbos, Z. G. and Sullivan, J. H. (2002). Robustness to non-normality of the multivariate EWMA control chart. *Journal of Quality Technology*, 34, 260-276.

33. Sullivan, J. H. and Jones, L. A. (2002). A self-starting control chart for multivariate individual observations. *Technometrics*, 44, 24-33.

34. Sun, R. and Tsung, F. (2003). A kernel-distance-based control chart using support vector methods. *International Journal of Production Research*, 41, 2975-2989.

35. Tartakovsky, A. G., Rozovskii, B. L., Blazek, R. B. and Kim, H. (2006). Detection of intrusions in information systems by sequential change-point methods (with Discussion). *Statistical Methodology*, 3, 252-340.

36. Tyler, D. E. (1987). A distribution-free M-estimator of multivariate scatter. *The Annals of Statistics*, 15, 234-251.

37. Woodall, W. H. and Montgomery, D. C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31, 376-386.

38. Zamba, K. D. and Hawkins, D. M. (2006). A multivariate change-point model for statistical process control. *Technometrics*, 48, 539-549.

39. Zhou, S., Ding, Y., Chen, Y. and Shi, J. (2003). Diagnosability study of multistage manufacturing processes based on linear mixed-effects models. *Technometrics*, 45, 312-325.

40. Zou, C., Jiang, W. and Tsung, F. (2011). A LASSO-based SPC diagnostic framework for multivariate statistical process control. *Technometrics*, 53, 297-309.

41. Zou, C. and Qiu, P. (2009). Multivariate statistical process control using LASSO. *Journal of the American Statistical Association*, 104, 1586-1596.

42. Zou, C., Wang, Z. and Tsung, F. (2012). A spatial rank-based multivariate EWMA control chart. *Naval Research Logistic*, 59, 91-110.

43. Zou, C. and Tsung, F. (2008). Directional MEWMA schemes for multistage process monitoring and diagnosis. *Journal of Quality Technology*, 40, 407-427.

44. Zou, C. and Tsung, F. (2011). A multivariate sign EWMA control chart. *Technometrics*, 53, 84-97.

45. Zou, C., Tsung, F. and Liu, Y. (2008). A change point approach for phase I analysis in multistage processes. *Technometrics*, 50, 344-356.

Authors' Biographies:

Xuemin Zi is Associate Professor of the School of Science at Tianjin University of Technology and Education. Her research interests include statistical process control and design of experiments.

Changliang Zou is Associate Professor of School of Mathematical Sciences, Nankai University. He obtained his B.S, M.S, and Ph.D. in Statistics, from Nankai University, in 2003, 2006, and 2008, respectively. His primary research interests include statistical process control, nonparametric regression and dimension reduction, high-dimensional data analysis. His research has been published in various referred journals in Statistics and Industrial Engineering including Journal of the American Statistical Association, Annals of Statistics, Technometrics, Journal of Quality Technology, IIE Transactions, Statistica Sinica, Naval Research Logistic, Annals of Operation Research, etc.

Qin Zhou is Assistant Professor of School of Mathematical Sciences, Jiangsu Normal University. Her research interests include statistical process control and health-care system.

Jingshen Wang is undergraduate student of School of Mathematical Sciences, Nankai University. Changliang Zou is her final-year project supervisor. Her research interests include statistical process control, semiparametric regression and lack-of-fit test.