A Rank Based Multivariate CUSUM Procedure

Peihua Qiu and Douglas Hawkins
School of Statistics
University of Minnesota
313 Ford Hall
224 Church St. SE
Minneapolis, MN 55455

Abstract

We consider statistical process control when measurements are multivariate. A CUSUM procedure is suggested in detecting a shift in the mean vector of the measurements, which is based on the cross-sectional antiranks of the measurements. At each time point, the measurements are ordered and their antiranks, which are the indices of the order statistics, are recorded. When the process is in-control and the joint distribution of the multivariate measurements satisfies some regularity conditions, the antirank vector at each time point has a given distribution. This distribution changes to some other distribution when the process is out-of-control and the components of the shift in the mean vector of the process are not all the same. This CUSUM can therefore detect shifts in all directions except the one in which the components of the shift in the mean vector are all the same but not zero. The shift with equal components, however, can be easily detected by another univariate CUSUM. The former CUSUM procedure is distribution-free in the sense that all its properties depend on the distribution of the antirank vector only.

Key Words: Antiranks; CUSUM; Discrete Measurements; Distribution-Free; Goodness-Of-Fit Test; Multivariate Measurements; Statistical Process Control; Ties.

1 Introduction

In modern quality control, it is becoming common to monitor several quality characteristics of a process simultaneously by taking various measurements with on-line computers and other advanced data-acquisition equipment. A number of multivariate CUSUM procedures have been proposed in the literature for detecting a shift in the mean vector of the measurements (see e.g., Mason, Champ, Tracy, Wierda, and Young 1997). This paper suggests an alternative multivariate CUSUM

procedure based on the cross-sectional antiranks of the measurements which are defined by the indices of the order statistics of the measurements at each time point. One major feature of this new procedure is that it is distribution-free in the sense that all its properties depend on the distribution of the antirank vector only.

Suppose that $\mathbf{X}(i) = (X_1(i), X_2(i), \cdots, X_p(i))'$ denotes p measurements of a process at the i-th time point. When the process is in-control, $\mathbf{X}(i)$ is assumed to be independent and identically distributed across different time points with joint distribution function $F(\mathbf{x})$. For simplicity, let us assume that its mean vector $\mu = \mathbf{0} = (0, 0, \cdots, 0)'$, the process becomes out-of-control in the sense that the mean vector μ has a shift from $\mathbf{0}$ after an unknown time point, and other properties of the process are assumed to be unchanged after the shift. The performance of a CUSUM procedure is often measured by the average run length (ARL), which is the average number of samples needed for the procedure to signal the shift.

Hotelling's (1947) control chart signals a shift in the mean vector when $T^2 = \mathbf{X}(i)'\Sigma_{\mathbf{X}}^{-1}\mathbf{X}(i) > h$ where h is a control limit and $\Sigma_{\mathbf{X}}$ denotes the covariance matrix of $\mathbf{X}(i)$. As some authors already pointed out (e.g., Alt 1985; Jackson 1985), this procedure is not sensitive to small and moderate shifts because it uses only the current observation in detecting the shift. Woodall and Ncube (1985) suggested detecting the shift by monitoring CUSUMs of individual measurements simultaneously. This procedure may also be ineffective because individual measurements are often not as sensitive to a shift as the entire measurement vector. When the direction of a shift is known, Healy (1987) suggested a CUSUM procedure based on a linear combination of individual measurements. Hawkins (1991) extended Healy's method by proposing a CUSUM procedure based on regression adjustment among p individual measurements to detect a shift in several known directions. Pignatiello and Runger (1990) also proposed a procedure to detect a shift by using several aimed CUSUMs.

Other multivariate CUSUM procedures in the literature include two procedures suggested by Crosier (1988): one is based on a CUSUM of T and the other signals a shift when $S_i'\Sigma_{\mathbf{X}}^{-1}S_i > h$, where $S_i = 0$ if $C_i \leq k$ and $S_i = (S_{i-1} + \mathbf{X}(i))(1 - k/C_i)$ otherwise, $C_i = \{(S_{i-1} + \mathbf{X}(i))'\Sigma_{\mathbf{X}}^{-1}(S_{i-1} + \mathbf{X}(i))\}^{1/2}$, $S_0 = 0$, and k is a constant. (The control limit h appearing at several different places may represent different constants. We use the same notation for simplicity.) Based on a simulation study, Crosier (1988) pointed out that his second procedure performed better than the first one in many cases.

Lowry, Woodall, Champ, and Rigdon (1992) extended the univariate exponentially weighted moving average control procedure (e.g., Lucas and Saccucci 1990) to the multivariate case. Their procedure signals a shift when $Z_i'\Sigma_{Z_i}^{-1}Z_i > h$, where $Z_i = R\mathbf{X}(i) + (I - R)Z_{i-1}, Z_0 = 0, R = diag(r_1, r_2, \cdots, r_p)$, and $0 < r_i \leq 1$ for all $i = 1, 2, \cdots, p$. A generalization of this procedure was discussed by Liu (1996).

The statistical properties of these multivariate CUSUM procedures are often based on the assumption that the measurement vector has a normal distribution, which is true in many situations but might not be true in many others (e.g., when the measurements are counts or when the measurements can only take positive values such as the economic indices; see Section 4 for a real example). When the true distribution of the measurement vector is not normal, the actual in-control ARL value is generally not equal to the specified in-control ARL value assuming normality (see Section 3 for numerical examples). If the actual in-control ARL value is smaller than the specified in-control ARL value, for example, then the chance that the CUSUM procedure signals because of randomness would be higher than expected and therefore it is harder to distinguish signals due to real shifts from signals due to randomness. One way to avoid this difficulty is to transform the measurement vector to a normal distribution vector before applying the conventional procedures. In multivariate cases, however, it is often difficult to find such transformations. As a matter of fact, the statistical theory to describe multivariate distributions other than the multivariate normal distributions is very limited in the literature (cf. Johnson 1987). To partially overcome this difficulty, we suggest a multivariate CUSUM procedure which does not depend on the normal distribution assumption in this paper. We first notice that the null hypothesis involved in multivariate process control

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_p = 0 \tag{1}$$

is equivalent to the combination of the following two hypotheses:

$$H_0^{(1)} : \mu_1 = \mu_2 = \cdots = \mu_p \tag{2}$$

and

$$H_0^{(2)} : \sum_{j=1}^{p} \mu_j = 0. \tag{3}$$

The hypothesis $H_0^{(1)}$ is related to the order of magnitude of $\{\mu_j, j = 1, 2, \cdots, p\}$. When $H_0^{(1)}$ is true, $H_0^{(2)}$ is related to the magnitudes of $\{\mu_j, j = 1, 2, \cdots, p\}$. For a mean vector $\mu = (\mu_1, \mu_2, \cdots, \mu_p)'$ with fixed Euclidean length, violation of H_0 implies violation of either $H_0^{(1)}$ or $H_0^{(2)}$ or both.

Furthermore if μ does not satisfy H_0, then the closer it is to $H_0^{(1)}$ (it is therefore more difficult to detect by a CUSUM designed for $H_0^{(1)}$), the farther it is from $H_0^{(2)}$ and consequently it is easier to detect by a CUSUM designed for $H_0^{(2)}$, and vice versa. Hence a combination of two CUSUM procedures designed for $H_0^{(1)}$ and $H_0^{(2)}$, respectively, will be able to detect all kinds of shifts in the mean vector of the process.

Shifts that violate $H_0^{(2)}$ can be detected by a univariate CUSUM based on $\sum_{j=1}^{p} X_j(i)$. In Section 2, we will introduce a multivariate CUSUM procedure to detect shifts that violate $H_0^{(1)}$. This procedure is based on the antiranks of the p measurements. It is distribution-free in the sense that its ARL properties depend on the joint distribution of the p measurements through the distribution of the antirank vector. In other words, as long as the distribution of the antirank vector is determined, the ARL properties of the procedure are uniquely determined, regardless of the joint distribution of the p original measurements and whether or not that distribution is continuous. A numerical study is presented in Section 3 regarding the ARL properties of the procedure. We apply this method to a real data set in Section 4. Some remarks conclude the article in Section 5.

2 An Antirank Based CUSUM

As mentioned in Section 1, hypothesis $H_0^{(1)}$ defined in (2) is related to the order of magnitude of $\{\mu_j, j = 1, 2, \cdots, p\}$. It is therefore natural to expect that some rank based procedure can detect shifts that violate $H_0^{(1)}$. Let $\mathbf{A}(i) = (A_1(i), A_2(i), \cdots, A_p(i))'$ be the antirank vector of $\mathbf{X}(i)$. That is, $\mathbf{A}(i)$ is a permutation of $(1, 2, \cdots, p)'$ such that $X_{A_1(i)}(i) \leq X_{A_2(i)}(i) \leq \cdots \leq X_{A_p(i)}(i)$ are the order statistics of $\{X_j(i), j = 1, 2, \cdots, p\}$.

Let us first consider $A_1(i)$. For $1 \leq j \leq p$, define

$$\xi_{1,j}(i) = \begin{cases} 1 & \text{if } A_1(i) = j \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

and $\boldsymbol{\xi}_{\mathbf{1}}(i) := (\xi_{1,1}(i), \xi_{1,2}(i), \cdots, \xi_{1,p}(i))'$. That is, $\xi_{1,j}(i)$ is defined as an indicator of the event that the j-th measurement takes the smallest value among p measurements at time point i. Under $H_0^{(1)}$, suppose that $E(\xi_{1,j}(i)) = g_{1,j}$, for $j = 1, 2, \cdots, p$. Then the probability distribution of $A_1(i)$ is $\{g_{1,j}, j = 1, 2, \cdots, p\}$. After a shift in the mean vector of the process, the corresponding distribution of $A_1(i)$ is denoted by $\{g_{1,j}^*, j = 1, 2, \cdots, p\}$.

Theorem 1 Let ν_F be the probability measure defined by the joint distribution function $F(\mathbf{x})$ of the p measurements under H_0. If $\nu_F(O) > 0$ for any open set $O \in R^p$ which includes the origin in its closure and has positive Lebesgue area, then $\{g^*_{1,j}, j = 1, 2, \cdots, p\}$ is different from $\{g_{1,j}, j = 1, 2, \cdots, p\}$ when $H_0^{(1)}$ is violated by the shift in the mean vector of the process. (Proof is given in Appendix A.)

In Theorem 1, the condition on the distribution function $F(\mathbf{x})$ basically says that $F(\mathbf{x})$ has a positive support around the mean vector of the process, which is true for most continuous distribution functions. If some or all of the measurements are discrete, then the corresponding result is stated in the following theorem.

Theorem 2 Suppose that $X_j(i), j = 1, \cdots, r$, are discrete measurements and $X_j(i), j = r+1, \cdots, p$, are continuous measurements where $0 \leq r \leq p$ is a positive integer. Under H_0, $X_j(i)$ is assumed to take values of $v_j + k_j w_j$ for $k_j = 0, \pm 1, \cdots, \pm C_j$ and $j = 1, \cdots, r$ where v_j and $w_j > 0$ are constants and $0 \leq C_j \leq \infty$ is an integer. Assume that the mean vector of $\mathbf{X}(i)$ changes to $\mu = (\mu_1, \mu_2, \cdots, \mu_p)'$ from $\mathbf{0}$ after a shift. Let $\mu_{(1)} = \min(\mu_1, \mu_2, \cdots, \mu_p)$ and $\mu_{(p)} = \max(\mu_1, \mu_2, \cdots, \mu_p)$. If there exists a discrete measurement $X_{j_1}(i)$ with $1 \leq j_1 \leq r$ which has a shift $\mu_{(1)}$ in its mean, then it is further assumed that $\mu_{(p)} - \mu_{(1)} \geq w_{j_1}$. Then $\{g^*_{1,j}, j = 1, 2, \cdots, p\}$ is different from $\{g_{1,j}, j = 1, 2, \cdots, p\}$ if $\nu_F(O \bigcup \mathcal{S}) > 0$ for any open set $O \in R^p$ which satisfies (i) O includes the origin in its closure; (ii) it has positive Lebesgue area; and (iii) $O \bigcup \mathcal{S} \neq \emptyset$, where $\mathcal{S}$ denotes the sample space of $\mathbf{X}(i)$ before the shift. (Proof is given in Appendix B.)

For discrete measurements, if the shifts in their means are so small that their orders are not changed, then the distribution of $A_1(i)$ is not changed either. This explains the reason why the assumption $\mu_{(p)} - \mu_{(1)} \geq w_{j_1}$ is imposed in Theorem 2. If the discrete measurements are counts and the possible shifts in their means are multiples of the increments of their values, then this assumption is often satisfied because $\mu_{(1)}$ and $\mu_{(p)}$ could be practically assumed to have opposite signs (0 could be counted as either positive or negative but not both for this purpose). Shifts in which $\mu_{(1)}$ and $\mu_{(p)}$ have the same sign could be easily detected by a univariate CUSUM for $H_0^{(2)}$.

2.1 The Antirank-Based CUSUM

Under the conditions of Theorem 1 or Theorem 2, testing $H_0^{(1)}$ is equivalent to testing

$$H_0^{(1)*}: \text{ the distribution of } A_1(i) \text{ is } \{g_{1,j}, j = 1, 2, \cdots, p\}. \tag{5}$$

By combining the goodness-of-fit test of a distribution function (see e.g., Read and Cressie 1988) and Crosier's (1988) idea for constructing a multivariate CUSUM procedure, we suggest the following procedure for testing $H_0^{(1)*}$. Let

$$\begin{cases} \mathbf{S}_n^{(1)} = \mathbf{0} & \\ \mathbf{S}_n^{(2)} = \mathbf{0}, & \text{if } C_n \le k_1 \\ \mathbf{S}_n^{(1)} = (\mathbf{S}_{n-1}^{(1)} + \xi_{\mathbf{1}}(n))(C_n - k_1)/C_n & \\ \mathbf{S}_n^{(2)} = (\mathbf{S}_{n-1}^{(2)} + \mathbf{g_1})(C_n - k_1)/C_n, & \text{if } C_n > k_1, \end{cases}$$

$$C_n = [(\mathbf{S}_{n-1}^{(1)} - \mathbf{S}_{n-1}^{(2)}) + (\xi_{\mathbf{1}}(n) - \mathbf{g_1})]' diag((S_{n-1,1}^{(2)} + g_{1,1})^{-1}, \cdots, (S_{n-1,p}^{(2)} + g_{1,p})^{-1})[(\mathbf{S}_{n-1}^{(1)} - \mathbf{S}_{n-1}^{(2)}) + (\xi_{\mathbf{1}}(n) - \mathbf{g_1})],$$

where $\mathbf{g_1} = (g_{1,1}, g_{1,2}, \cdots, g_{1,p})'$, $\mathbf{S}_0^{(1)} = \mathbf{S}_0^{(2)} = \mathbf{0}$ and $k_1 \ge 0$ is a constant. Define

$$y_n = (\mathbf{S}_n^{(1)} - \mathbf{S}_n^{(2)})' diag(1/S_{n,1}^{(2)}, \cdots, 1/S_{n,p}^{(2)})(\mathbf{S}_n^{(1)} - \mathbf{S}_n^{(2)}). \tag{6}$$

Then

$$y_n > h_1 \tag{7}$$

signals a shift where $h_1 > 0$ is a control limit. When $k_1 = 0$, $\mathbf{S}_n^{(1)}$ is a frequency vector with its j-th element denoting the observed count of $(A_1(i) = j)$ as of time point n, for $j = 1, 2, \cdots, p$; and $\mathbf{S}_n^{(2)}$ equals $(ng_{1,1}, ng_{1,2}, \cdots, ng_{1,p})'$ where $ng_{1,j}$ is the expected count of $(A_1(i) = j)$ (expected under $H_0^{(1)*}$) as of time point n, for $j = 1, 2, \cdots, p$. Therefore $y_n = \sum_{j=1}^{p} (S_{n,j}^{(1)} - S_{n,j}^{(2)})^2 / S_{n,j}^{(2)}$ is the conventional Pearson's X^2 statistic which measures the difference between the observed counts of events $(A_1(i) = j)$ as of time point n, for $j = 1, 2, \cdots, p$, and the expected counts of the same events under $H_0^{(1)*}$. It is well-known that this statistic has an approximate χ^2 distribution with degrees of freedom $p - 1$ when n is large enough. This approximate distribution should work well as long as all expected counts are larger than or equal to 1 (Read and Cressie 1988, Chapter 5). When $H_0^{(1)*}$ does not hold, the value of y_n tends to be large. Therefore large value of y_n indicates a possible violation of $H_0^{(1)*}$. In applications (cf. the numerical examples in Section 3), the value of h_1 needs to be determined by simulation.

2.2 Some Technical Remarks

From the definition of the CUSUM procedure, it can be seen that (see Appendix C for explanation)

$$y_n = \max(0, C_n - k_1). \tag{8}$$

Therefore the parameter k_1 is used to repeatedly restart the CUSUM procedure when there is no evidence of violation of $H_0^{(1)*}$ such that the procedure can react to a real shift promptly. To achieve some specific in-control ARL property, k_1 should be chosen from the interval

$$[0, \max_{\ell=1}^{p} \frac{\sum_{j\neq\ell} g_{1,j}}{g_{1,\ell}}). \tag{9}$$

If k_1 is chosen outside this interval, then the CUSUM procedure will be restarted at each time point and consequently the in-control ARL property can not be achieved (see Appendix D for detailed explanation).

It is apparent that both the in-control and out-of-control properties of the CUSUM procedure (6)-(7) are determined uniquely by the distribution of $A_1(i)$ for fixed k_1 and h_1. The null distribution of $A_1(i)$, $\{g_{1,j}, j = 1, 2, \cdots, p\}$, could be determined from preliminary data by computing the relative frequencies of $(A_1(i) = j)$ for $j = 1, 2, \cdots, p$. If the joint distribution $F(\mathbf{x})$ of the p measurements is known under H_0, then $\{g_{1,j}, j = 1, 2, \cdots, p\}$ could also be computed algebraically or numerically from $F(\mathbf{x})$. As long as $\{g_{1,j}, j = 1, 2, \cdots, p\}$ is determined, the in-control property of the procedure could be obtained by simulating a series of i.i.d. multinomial random variables $\{A_1(i)\}$ (cf. Section 3). In that sense, the procedure (6)-(7) is *distribution-free*.

In some cases the p measurements are exchangeable before the shift in their mean vector in the sense that their joint distribution function remains unchanged if their order changes. For example, the thickness measurements across a sheet product made by a rolling operation could be assumed exchangeable in some circumstances (Crosier 1988). In such cases, $\{A_1(i)\}$ has a uniform distribution $(1/p, 1/p, \cdots, 1/p)'$ under H_0.

To test $H_0^{(1)}$, another natural idea is to regard $A_1(i)$ as a univariate random variable and then to construct a univariate CUSUM based on $A_1(i)$ as usual. However, this kind of CUSUM can not distinguish the mean vector $(0, -1, 0)'$ from $(0, 0, 0)'$ (in the case of $p = 3$) when the null distribution of the measurement vector is symmetric because $E(A_1(i)) = 2$ in both cases. Consequently it is not appropriate for testing $H_0^{(1)}$.

When all or $p-1$ measurements are continuous, the chance of ties among the p measurements is negligible for all practical purposes. Therefore $A_1(i)$ is well defined. When two or more measurements are discrete and these discrete measurements can take the same values, however, ties among the p measurements are possible. For example, suppose that the measurement vector takes the value of $(-1,-1,0)'$ (in the case of $p=3$) at time point i_0. Then $X_1(i_0)$ and $X_2(i_0)$ are tied, and the value of $A_1(i_0)$ is not well defined. To overcome this difficulty, there are several existing proposals in the literature (see e.g., Gibbons and Chakraborti 1992). One such proposal is to randomly assign 1 or 2 to $A_1(i_0)$ with probability 0.5 for each number. To do that, we first need to identify the ties which would affect the value of $A_1(i_0)$ and then to have a supplementary mechanism to randomly assign values for $A_1(i_0)$. Unfortunately results from such a random mechanism are not reproducible. In our simulation study presented in Section 3, we overcome the difficulty caused by ties by modifying the definition of $\xi_{1,j}(i)$ in (4) as follows. Suppose that $X_{j_1}(i), X_{j_2}(i), \cdots, X_{j_k}(i)$ form a tie and their values reach the minimum among all p measurements at time point i. Then we define

$$\xi_{1,j}(i) = \begin{cases} 1/k & \text{if } j \in \{j_1, j_2, \cdots, j_k\} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

By using this definition, no information about the antiranks is lost and moreover the results are reproducible.

The multivariate CUSUM (6)-(7) is based on one component of the antirank vector. It is possible to construct a similar CUSUM based on several components of the antirank vector. Let $(A_{j_1}(i), A_{j_2}(i), \cdots, A_{j_q}(i))'$ be any q components of the antirank vector with $j_1 < j_2 < \cdots < j_q$. Suppose that the distribution of $(A_{j_1}(i), A_{j_2}(i), \cdots, A_{j_q}(i))'$ when $H_0^{(1)}$ is true has been determined in sample space:

$$S(j_1, j_2, \cdots, j_q) = \{(i_1, i_2, \cdots, i_q) : i_1, i_2, \cdots, i_q \text{ are } q \text{ different integers in } (1, 2, \cdots, p)\}.$$

The sample space $S(j_1, j_2, \cdots, j_q)$ has $P_{q,p} = p(p-1)\cdots(p-q+1)$ elements. Similar to (4), we can define a $P_{q,p}$-dimensional random vector $\xi_{j_1,j_2,\cdots,j_q}(i)$ with its j-th component equal to 1 when $(A_{j_1}(i), A_{j_2}(i), \cdots, A_{j_q}(i))'$ takes the value of the j-th element in $S(j_1, j_2, \cdots, j_q)$ and 0 otherwise. Then a multivariate CUSUM could be constructed similar to (6)-(7) with $\xi_1(i)$ replaced by $\xi_{j_1,j_2,\cdots,j_q}(i)$ and p by $P_{q,p}$. Based on the consideration addressed at the end of this section and the numerical examples presented in Section 3, a CUSUM procedure based on the first and last components of the antirank vector is highly recommended.

As mentioned before, the CUSUM procedure (6)-(7) can detect shifts in all directions except the one in which the components of the mean vector are all the same but not zero. For many problems, this procedure might be adequate since this situation is unlikely. If we do care about this situation, however, it can be detected easily by a univariate CUSUM based on $SX(i) = \sum_{j=1}^{p} X_j(i) / \sum_{j_1=1}^{p} \sum_{j_2=1}^{p} \sigma_{j_1,j_2}$, where σ_{j_1,j_2} is the (j_1, j_2)-th element of $\Sigma_{\mathbf{X}}$. Define

$$\begin{cases} C_0^+ = 0 \\ C_0^- = 0 \\ C_n^+ = \max(0, C_{n-1}^+ + SX(n) - k_2) \\ C_n^- = \min(0, C_{n-1}^- + SX(n) + k_2), \end{cases} \tag{11}$$

where k_2 is a constant. Then this univariate CUSUM signals a shift when

$$C_n^+ > h_2 \text{ or } C_n^- < -h_2, \tag{12}$$

where h_2 is a control limit. This univariate CUSUM is well studied in the literature. The relationship between its in-control ARLs and the parameters k_2 and h_2 has been tabulated by several authors (see e.g., Hawkins and Olwell 1998, Chapter 3). Therefore a CUSUM which can detect all kinds of shift in the mean vector of a multivariate process can be obtained by combining (6)-(7) and (11)-(12), which signals a shift whenever (7) or (12) holds.

Sometimes monitoring the variance of the process is also our interest (e.g., Gan 1995; Hawkins and Olwell 1998, Chapter 4). Increased variability increases costs and usually decreases quality and therefore needs to be detected promptly. On the other hand, reduced variability increases quality and offers the opportunity for reduced costs. As a by-product, as long as the change in the variance of the process changes the distribution of the antirank vector, it could be detected by the CUSUM procedure (6)-(7) also (see Section 3 for a numerical example).

2.3 Rank or antirank?

One question that may occur is why we use the antirank rather than the more familiar rank of each measurement within the vector. If we were interested in a univariate control chart for a particular component of $\mathbf{X}(i)$, then the rank of that component within the vector might indeed be a relevant distribution-free summand. However this problem is not common in multivariate SPC. Instead, we are faced with the detection of shifts in some unknown component (or perhaps components). For

this problem, the first antirank is particularly effective in detecting a downward shift in a single arbitrary and unknown component – this component will start to figure prominently in the first antirank. Similarly the last antirank is attractive for detecting an upward shift in the mean of one of the components.

3 Simulation Study

We present some numerical results regarding the in-control and out-of-control properties of our CUSUM procedure. First, we focus on the antirank-based procedure (6)-(7). Second, we study the combined procedure (6)-(7) and (11)-(12). Because the univariate CUSUM (11)-(12) has been well discussed in the literature (e.g., Hawkins and Olwell 1998), it will not be discussed separately here.

We first assume that $p = 4$, the in-control distribution of $A_1(i)$ is uniform distribution $(1/4, 1/4, 1/4, 1/4)$ and the in-control ARL of the procedure is fixed at 200. The corresponding control limit values (searched with step 0.001) for several measurement distributions are presented in Table 1 when k_1 is chosen to be 0.5, 1.0 and 1.5, respectively. The first row (labelled as "distribution-free") corresponds to the case when the CUSUM procedure is constructed directly from a series of random vectors with uniform distribution $(1/4, 1/4, 1/4, 1/4)$. The actual in-control ARL values are presented in parentheses along with their standard errors based on 10000 replications. The next four rows correspond to four measurement distributions: $N(\mathbf{0}, I_4), N(\mathbf{0}, \Sigma_{CS5})$ where Σ_{CS5} denotes a covariance matrix with all diagonal elements equal to 1 and all off-diagonal elements equal to 0.5, $IndExp(1.0)$ in which the four measurements are independent of each other and identically distributed as exponential with mean 1.0, and $IndPoisson(1.0)$ in which the four measurements are independent and identically distributed as Poisson with mean 1.0. In the cases of $IndExp(1.0)$ and $IndPoisson(1.0)$, the measurements are also normalized to have mean 0 in our simulation to be consistent with the assumptions made in Section 1 although this will not change any results. All results presented in this section are averages of 10,000 replications. For the case of $IndPoisson(1.0)$, the modified definition of $\xi_{1,j}(i)$ in (10) is used. The random numbers used in this section are generated by random number generators in *RANLIB*, a library of Fortran routines to generate random numbers. These random number generators can be downloaded from the web page http://www.netlib.org/.

Table 1: The control limit values of the procedure (6)-(7) for several measurement distributions when its in-control ARL is fixed at 200. In each of the 15 tabled entries, the actual in-control ARL value is 200 and the standard error is 2, based on 10,000 replications.

	$k_1 = 0.5$	$k_1 = 1.0$	$k_1 = 1.5$
distribution-free	8.053 (200,2)	6.840 (200,2)	5.180 (200,2)
$N(\mathbf{0}, I_4)$	8.029 (200,2)	6.842 (200,2)	5.171 (200,2)
$N(\mathbf{0}, \Sigma_{CS5})$	8.043 (200,2)	6.777 (200,2)	5.168 (200,2)
$IndExp(1.0)$	8.044 (200,2)	6.833 (200,2)	5.181 (200,2)
$IndPoisson(1.0)$	8.033 (200,2)	6.849 (200,2)	5.173 (200,2)

From Table 1, the control limit values of the procedure (6)-(7) for several different measurement distributions are almost the same (the slight differences among them could be explained by simulation errors). Therefore the procedure is indeed distribution-free in that its ARL property depends on the distribution of $A_1(i)$ only. Consequently the control limit value of the procedure could be determined by a simulation in which the CUSUM is constructed from a series of random vectors with multinomial distribution $\{g_{1,j}, j = 1, 2, \cdots, p\}$ (these vectors could be generated by a multinomial random number generator) without knowing the true distribution of the measurement vector.

As a comparison, the next example shows that the conventional multivariate CUSUM procedures indeed require the normal distribution assumption on the measurement vector. Suppose $p = 4$ and the assumed joint distribution of the measurements is $N(\mathbf{0}, I_4)$. We use Crosier's (1988) procedure as an example here because it is one of the bases of our antirank based procedure. In Crosier's procedure, the control limit and the constant k are chosen to be 4.503 and 1.0, respectively, so its in-control ARL equals 200 when the measurement vector is normally distributed (see Section 1 for a description of Crosier's procedure). If the four measurements are actually independent of each other and each has a Poisson distribution with mean parameter λ where $\lambda = 0.01, 0.1, 0.5, 1.0, 1.5, 2.0$ and 5.0, then the true in-control ARL values are presented in Figure 1(a) by the small diamonds (the measurements are normalized to have mean 0 and standard deviation 1 in the simulation). It can be seen that the true in-control ARL values differ from the assumed in-control ARL value in all cases. When λ is small, they are not even close to each other. One idea to overcome this difficulty is to transform the measurement vector to a normal distribution vector. A commonly used transformation for Poisson distribution variable is the square root transformation. The corresponding true in-control ARL values of the root measurements are presented in Figure 1(b) (the

transformed measurements are also normalized to have mean 0 and standard deviation 1 in the simulation). Again, the true in-control ARL values are far from the assumed in-control ARL value in most situations. In reality, the statistical theory to describe multivariate distributions other than the multivariate normal distributions is limited and it is also hard to find an appropriate transformation to transform a non-Gaussian random vector to a Gaussian vector especially when p is large.

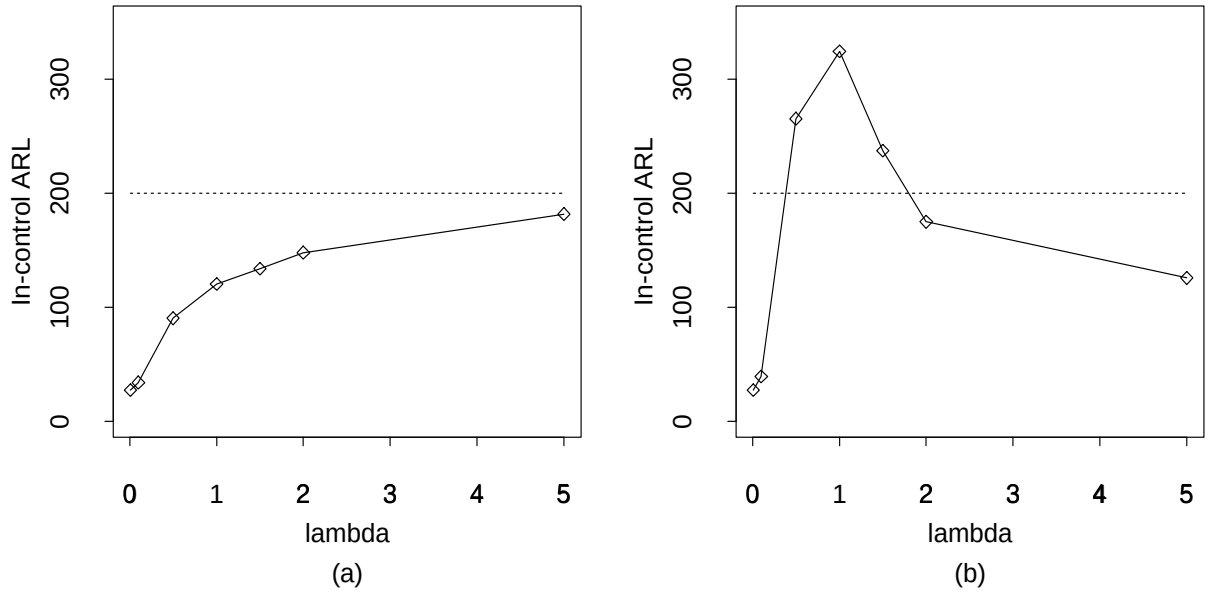

Figure 1: (a) The small diamonds denote the true in-control ARL values of Crosier's procedure when $p = 4$, the measurements are independent of each other and each measurement has a Poisson distribution with mean parameter λ where $\lambda = 0.01, 0.1, 0.5, 1.0, 1.5, 2.0$ and 5.0. (b) The small diamonds denote the true in-control ARL values of Crosier's procedure after the measurements used in (a) are square root transformed. In Crosier's procedure, the control limit and the constant k are chosen to be 4.503 and 1.0, respectively, such that its in-control ARL equals 200 when the joint distribution of the measurement vector is assumed to be $N(\mathbf{0}, I_4)$. The dotted line in each plot denotes the assumed in-control ARL value 200.

When the joint distribution of the measurements is normal, we would expect that the antirank based procedure is less efficient than Crosier's procedure for detecting a shift in the mean vector of the measurements because it uses only the order information of the measurements (instead of the measurements themselves). This conjecture is confirmed by the following example. Suppose $p = 4$, the null distribution of the measurements is $N(\mathbf{0}, I_4)$, and a shift $(a, a, 0, 0)'$ in the mean vector of the measurements occurs at the initial time point. Both Crosier's procedure and our antirank based procedure (6)-(7) are used to detect the shift. The parameters in the two procedures are chosen such that their in-control ARL values both equal 200 ($h = 4.503$ and $k = 1.0$ in Crosier's

procedure; $h_1 = 6.842$ and $k_1 = 1.0$ in (6)-(7)). Their out-of-control ARL values are presented in Figure 2 by the small triangles and small diamonds, respectively, when $a = -0.1, -0.5, -1.0, -1.5$ and -2.0. It can be seen that Crosier's procedure is more sensitive to the shift than the procedure (6)-(7) as expected.

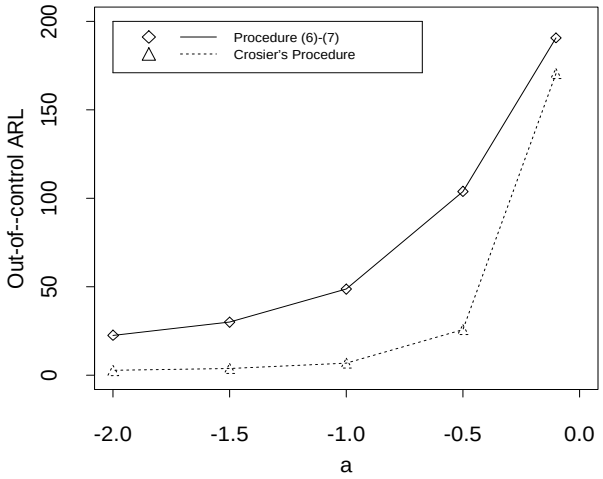

Figure 2: The small triangles and small diamonds denote the out-of–control ARL values of Crosier's procedure and the antirank based procedure (6)-(7), respectively, when $p = 4$, the null joint distribution of the measurements is $N(\mathbf{0}, I_4)$ and the mean vector of the measurements has a shift $(a, a, 0, 0)'$ at the initial time point where $a = -0.1, -0.5, -1.0, -1.5$ and -2.0. The parameters in the two procedures are chosen such that their in-control ARL values are both 200 ($h = 4.503$ and $k = 1.0$ in Crosier's procedure; $h_1 = 6.842$ and $k_1 = 1.0$ in (6)-(7)).

Because the CUSUM procedure (6)-(7) is "distribution-free", its ARL properties could be studied based directly on the distribution of $A_1(i)$. When the null distribution of $A_1(i)$ is $(1/4, 1/4, 1/4, 1/4)$ and the in-control ARL is 200, the out-of-control ARLs and their standard errors based on 10,000 replications for $k_1 = 0, 0.5, 1.0, 1.5$ and for various out-of-control distributions of $A_1(i)$ are presented in Table 2. It can be seen that the out-of-control ARL values are roughly determined by $Q = \sum_{j=1}^{p}(g^*_{1,j} - g_{1,j})^2/g_{1,j}$ which is a measure of the discrepancy between the in-control and out-of-control distributions of $A_1(i)$. Apparently Q is a major but not the only factor affecting the out-of-control property of the procedure. In Table 2, the shift could occur at the initial time point or at the 500th time point. The latter is used as an approximation to the "steady state start" after which the distribution of the CUSUM statistic y_n approaches some "steady state distribution" that does not depend on n (see e.g., Hawkins and Olwell 1998, Chapter 3). Results in these two different cases are close to each other when $k_1 = 0.5, 1.0$ and 1.5. When $k_1 = 0$, the out-of-control ARL values are small in the initial start case because all observations are fully used instead of

being weighted by the factor $(C_n - k_1)/C_n$ in the definition of the CUSUM procedure (6)-(7). In the case when the shift occurs at the 500th time point, however, the corresponding out-of-control ARL values are much larger than the values when k_1 equals to 0.5, 1.0 or 1.5 because the first 500 observations are also involved in the CUSUM criterion y_n and they make the procedure harder to signal the shift. This example explains the reason why a parameter such as k_1 is often used in CUSUM procedures. We also tried the cases where the shift started at the 50th time point and the 1000th time point, respectively. The results are close to those in the case when the shift starts at the 500th time point and therefore are not presented here. We would like to mention that the case when $(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$ has one or more zero components is not likely to happen in reality. They are included in Table 2 for showing the performance of the CUSUM procedure when some components of $(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$ are small (please note that a positive shift in the mean of a measurement component will make the corresponding component of $(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$ small).

As with most existing CUSUM procedures, the optimal selection of the parameters in our antirank based CUSUM procedure depends on the target shift (see e.g., Crosier 1988; Hawkins and Olwell 1998). Table 3 presents some results regarding the optimal selection of the parameters k_1 and h_1 of the CUSUM procedure (6)-(7). As before the in-control ARL is fixed at 200. For a specific shift, the optimal k_1 and h_1 ($k_{1,min}$ and $h_{1,min}$ in Table 3) are defined by the values of these two parameters which make the procedure reach the shortest out-of-control ARL (ARL_{min} in the table). The optimal k_1 and h_1 values are found by a search in steps 0.01 and 0.0001, respectively, and we investigate values of the dimension $p = 2, 3, 4, 5$. The null distribution of $A_1(i)$ is assumed to be uniform in each case. For each p, the first two components of the distribution of $A_1(i)$ are assumed to be $(1.0, 0.0), (0.9, 0.1), (0.8, 0.2), (0.7, 0.3)$ and $(0.6, 0.4)$, respectively, after the shift occurs at the 500th time point. The remaining components are assumed to be 0. The five shifts are ordered from the largest to the mildest. This table provides us the following guidelines: (1) For a given p, a milder shift dictates a smaller choice for k_1. Consequently h_1 needs to be chosen larger to achieve the fixed in-control ARL value. (2) Both k_1 and h_1 should be chosen larger for larger p (which is consistent with what Crosier (1988) found in his simulation study). From Table 3, $k_{1,min}$ and $h_{1,min}$ follow roughly the following linear relationships: $h_{1,min} = 6.38 - 5.73k_{1,min}$ when $p = 2$ ($R^2 = 0.9917$); $h_{1,min} = 7.94 - 3.77k_{1,min}$ when $p = 3$ ($R^2 = 0.9993$); $h_{1,min} = 9.04 - 2.52k_{1,min}$ when $p = 4$ ($R^2 = 0.9792$); $h_{1,min} = 10.56 - 2.23k_{1,min}$ when $p = 5$ ($R^2 = 0.9915$). Therefore when p is between 2 and 5, the optimal k_1 and h_1 could be searched around these lines for a specific

Table 2: The out-of-control ARL values and their standard errors (numbers in parentheses) based on 10,000 replications for $k_1 = 0, 0.5, 1.0$ or 1.5 and for various out-of-control distributions of $A_1(i)$. The null distribution of $A_1(i)$ is $(1/4, 1/4, 1/4, 1/4)$ and the in-control ARL is fixed at 200. $Q = \sum_{j=1}^{p}(g_{1,j}^* - g_{1,j})^2/g_{1,j}$ measures the discrepancy between the in-control and out-of-control distributions of $A_1(i)$.

		shift at the initial time point			
$(g_{1,1}^*, g_{1,2}^*, g_{1,3}^*, g_{1,4}^*)$	Q	$k_1 = 0$	$k_1 = 0.5$	$k_1 = 1.0$	$k_1 = 1.5$
(1,0,0,0)	3.00	2.00 (.00)	4.00 (.00)	4.00 (.00)	4.00 (.00)
(.9,.1,0,0)	2.28	2.40 (.01)	4.98 (.02)	5.27 (.02)	5.41 (.02)
(.8,.2,0,0)	1.72	2.74 (.01)	6.35 (.03)	7.30 (.04)	7.77 (.05)
(.8,.1,.1,0)	1.64	2.95 (.01)	6.64 (.03)	7.29 (.04)	7.66 (.05)
(.7,.3,0,0)	1.32	3.00 (.01)	8.02 (.03)	10.54 (.07)	11.94 (.09)
(.7,.2,.1,0)	1.16	3.52 (.02)	9.28 (.06)	10.87 (.08)	11.77 (.09)
(.7,.1,.1,.1)	1.08	3.87 (.02)	9.62 (.06)	10.86 (.08)	11.75 (.09)
(.6,.4,0,0)	1.08	3.19 (.01)	9.50 (.04)	15.12 (.10)	18.61 (.16)
(.6,.3,.1,0)	.84	3.99 (.02)	13.31 (.09)	18.02 (.15)	19.77 (.17)
(.6,.2,.2,0)	.76	4.35 (.03)	15.17 (.11)	18.60 (.16)	20.20 (.18)
(.6,.2,.1,.1)	.68	4.88 (.03)	15.24 (.12)	18.37 (.15)	20.21 (.18)
(.5,.5,0,0)	1.00	3.24 (.01)	10.04 (.04)	17.34 (.11)	22.56 (.20)
(.5,.4,.1,0)	.68	4.19 (.02)	16.76 (.11)	27.89 (.24)	30.23 (.27)
(.5,.3,.2,0)	.52	4.89 (.03)	26.33 (.22)	34.41 (.31)	36.37 (.34)
(.5,.3,.1,.1)	.44	5.80 (.05)	25.66 (.21)	33.84 (.31)	36.81 (.34)
(.5,.2,.2,.1)	.36	6.70 (.06)	29.56 (.26)	35.21 (.32)	38.83 (.36)
(.4,.4,.2,0)	.44	5.04 (.03)	36.14 (.31)	49.55 (.47)	51.92 (.48)
(.4,.4,.1,.1)	.36	6.14 (.05)	34.08 (.29)	48.47 (.46)	51.13 (.48)
(.4,.3,.3,0)	.36	5.45 (.03)	62.50 (.61)	65.80 (.64)	66.83 (.65)
(.4,.3,.2,.1)	.20	8.22 (.08)	60.98 (.57)	71.16 (.69)	75.50 (.73)
(.4,.2,.2,.2)	.12	11.91 (.16)	71.03 (.69)	81.34 (.79)	87.07 (.85)
(.3,.3,.3,.1)	.12	10.23 (.12)	117.04 (1.16)	117.96 (1.15)	118.87 (1.17)
(.3,.3,.2,.2)	.04	19.42 (.35)	143.14 (1.40)	149.05 (1.46)	151.30 (1.47)
		shift at the 500th time point			
$(g_{1,1}^*, g_{1,2}^*, g_{1,3}^*, g_{1,4}^*)$	Q	$k_1 = 0$	$k_1 = 0.5$	$k_1 = 1.0$	$k_1 = 1.5$
(1,0,0,0)	3.00	24.20 (.42)	4.87 (.04)	4.56 (.03)	4.44 (.03)
(.9,.1,0,0)	2.28	27.03 (.51)	5.97 (.08)	5.80 (.09)	5.81 (.09)
(.8,.2,0,0)	1.72	31.32 (.64)	7.31 (.11)	7.54 (.15)	7.97 (.18)
(.8,.1,.1,0)	1.64	33.86 (.63)	7.73 (.13)	8.03 (.16)	8.53 (.20)
(.7,.3,0,0)	1.32	36.08 (.69)	9.17 (.14)	11.08 (.25)	12.45 (.33)
(.7,.2,.1,0)	1.16	37.39 (.80)	10.33 (.21)	11.49 (.28)	12.33 (.33)
(.7,.1,.1,.1)	1.08	40.81 (.88)	10.53 (.22)	11.50 (.28)	12.14 (.34)
(.6,.4,0,0)	1.08	40.14 (.78)	10.94 (.14)	15.70 (.37)	18.91 (.57)
(.6,.3,.1,0)	.84	45.66 (.97)	14.64 (.31)	18.12 (.51)	20.11 (.59)
(.6,.2,.2,0)	.76	47.11 (.94)	16.06 (.39)	18.01 (.51)	19.68 (.58)
(.6,.2,.1,.1)	.68	50.38 (1.02)	16.65 (.41)	17.86 (.51)	20.07 (.60)
(.5,.5,0,0)	1.00	40.30 (.82)	11.08 (.14)	17.70 (.40)	23.13 (.71)
(.5,.4,.1,0)	.68	50.02 (1.01)	18.03 (.40)	28.66 (.87)	31.95 (1.03)
(.5,.3,.2,0)	.52	57.39 (1.18)	27.35 (.81)	34.27 (1.12)	35.35 (1.18)
(.5,.3,.1,.1)	.44	63.09 (1.32)	26.86 (.80)	34.18 (1.11)	35.32 (1.14)
(.5,.2,.2,.1)	.36	73.80 (1.60)	31.23 (.92)	35.91 (1.13)	38.43 (1.16)
(.4,.4,.2,0)	.44	60.98 (1.23)	36.07 (1.08)	50.46 (1.70)	53.05 (1.85)
(.4,.4,.1,.1)	.36	69.56 (1.41)	34.26 (1.02)	49.86 (1.63)	54.15 (1.88)
(.4,.3,.3,0)	.36	70.68 (1.40)	61.51 (1.94)	63.71 (2.09)	64.84 (2.27)
(.4,.3,.2,.1)	.20	96.91 (2.16)	60.71 (2.03)	68.16 (2.31)	72.61 (2.54)
(.4,.2,.2,.2)	.12	124.02 (2.99)	71.83 (2.46)	81.90 (2.72)	82.53 (2.76)
(.3,.3,.3,.1)	.12	122.47 (2.87)	114.78 (4.01)	116.99 (4.15)	121.44 (4.51)
(.3,.3,.2,.2)	.04	220.72 (6.27)	132.27 (4.51)	146.03 (5.28)	151.37 (5.40)

Table 3: The optimal parameters $k_{1,min}$ and $h_{1,min}$ and the corresponding out-of-control ARL values and their standard errors (SE) for several shifts and dimensions.

$p=2$				$p=3$			
$(g^*_{1,1}, g^*_{1,2})$	$k_{1,min}$	$h_{1,min}$	$ARL_{min}(SE)$	$(g^*_{1,1}, g^*_{1,2}, g^*_{1,3})$	$k_{1,min}$	$h_{1,min}$	$ARL_{min}(SE)$
(1.0,0.0)	0.54	3.2199	6.79 (0.05)	(1.0,0.0,0)	0.88	4.6463	5.12 (0.03)
(0.9,0.1)	0.45	3.8499	10.36 (0.22)	(0.9,0.1,0)	0.47	6.1200	6.98 (0.10)
(0.8,0.2)	0.11	5.9321	16.88 (0.38)	(0.8,0.2,0)	0.28	6.8800	9.22 (0.16)
(0.7,0.3)	0.07	5.9478	32.32 (0.94)	(0.7,0.3,0)	0.18	7.2801	12.68 (0.18)
(0.6,0.4)	0.04	6.0108	85.68 (2.66)	(0.6,0.4,0)	0.16	7.3600	15.22 (0.20)
$p=4$				$p=5$			
$(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$	$k_{1,min}$	$h_{1,min}$	$ARL_{min}(SE)$	$(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4}, g^*_{1,5})$	$k_{1,min}$	$h_{1,min}$	$ARL_{min}(SE)$
(1.0,0.0,0,0)	1.70	4.6250	4.43 (0.03)	(1.0,0.0,0,0,0)	2.23	5.5022	3.96 (0.03)
(0.9,0.1,0,0)	0.80	7.3110	5.66 (0.08)	(0.9,0.1,0,0,0)	0.97	8.5900	5.02 (0.07)
(0.8,0.2,0,0)	0.47	8.0380	7.19 (0.11)	(0.8,0.2,0,0,0)	0.71	9.0800	6.39 (0.09)
(0.7,0.3,0,0)	0.28	8.1610	8.92 (0.13)	(0.7,0.3,0,0,0)	0.51	9.3500	7.48 (0.10)
(0.6,0.4,0,0)	0.27	8.1901	10.04 (0.13)	(0.6,0.4,0,0,0)	0.45	9.3800	8.19 (0.10)

target shift. In the same row of the table, the first two components of the distribution of $A_1(i)$ after the shift are kept the same. It can be seen that the CUSUM procedure (6)-(7) is more sensitive to the shift when the dimension p increases. This confirms the idea that the CUSUM procedure based on more measurements is often more sensitive to a shift in the mean vector of the process. If we do not have any target shift in mind, then k_1 is usually determined first. Based on our experience, it is often sufficient to choose k_1 from $[0, \frac{1}{2}\max_{\ell=1}^{p}\frac{\sum_{j\neq\ell} g_{1,j}}{g_{1,\ell}}]$ instead of from $[0, \max_{\ell=1}^{p}\frac{\sum_{j\neq\ell} g_{1,j}}{g_{1,\ell}})$ as specified by (9); a large (small) value of k_1 is good for detecting relatively large (small) shifts. We would like to mention that the optimal k_1 is 0 in all cases from our simulation when the shift occurs at the initial time point and the out-of-control ARL value is used as the only criterion to evaluate the performance of the CUSUM procedures. But the variability of the out-of-control ARL value is large in such cases.

The CUSUM procedure (6)-(7) is based on the first antirank of the measurements. As mentioned at the end of Section 2, the first and last antiranks are more sensitive to the shift than the remaining antiranks. Consequently CUSUM procedures based on them should be more powerful to detect the shift. The following example can demonstrate this point well. Suppose $p = 4$ and the joint distribution of the measurements is $N(\mathbf{0}, I_4)$. The in-control ARL is fixed at 200 and $k_1 = 1$. In Table 4, AR1 denotes the CUSUM procedure based on the first antirank of the measurements. AR12 denotes a procedure based on the first and second antiranks. The other notations initiated with "AR" are similarly defined. The control limit values are 6.842 and 15.6887 for AR1-AR4 and AR12-AR34, respectively. The numbers in the table are the out-of-control ARL values and their

standard errors based on 10,000 replications for several shifts. Without loss of generality, all mean components after the shift are assumed to be non-positive and each mean vector has at least one 0 component. (For a general mean vector, it can be transformed to the mean vector satisfying the conditions mentioned above by substracting its maximum component from each component of the vector and this transformation will not affect any property of the antirank vector.)

We first check the performance of AR1-AR4 which are based on a single antirank component. The first shift considered is $(\mu_1, \mu_2, \mu_3, \mu_4) = (-4, 0, 0, 0)$. In this case, there is only one measurement having shift in its mean, and the probability for this measurement to reach the minimum of the four measurements is large; consequently AR1 performs best. In the second case, shifts occur in two mean components, with one shift is larger in magnitude than the other one; still AR1 performs best. In the third situation, two shifts have the same magnitudes; both AR1 and AR4 perform slightly better than the other two. Each of the last four cases has shifts in three mean components: two components have shifts of the same magnitudes in the first two situations; all three components have different shifts in the third situation; and all three mean components have the same amount of shift in the last situation. It can be seen that at least one of AR1 and AR4 is better than the other two procedures. This example shows that AR1 and AR4 are indeed more sensitive to a shift than AR2 and AR3 in general. But we still need to choose between AR1 and AR4. This can be decided by the following rough guideline: to use AR1 if the number of mean components with smallest or close to smallest magnitudes is expected to be less than the number of mean components with largest or close to largest magnitudes and vice versa.

Next we check the performance of procedures AR12, AR13, $\cdots$, and AR34, all of which are based on two antirank components. Overall the out-of-control ARL values of these procedures are smaller than those of AR1-AR4 because they use more antirank information for shift detection. In the cases that AR1 performances well (e.g., when $(\mu_1, \mu_2, \mu_3, \mu_4) = (-4, 0, 0, 0)$), the procedures with the first antirank involved, namely AR12, AR13 and AR14, also perform reasonably well. This phenomenon is also true for AR4. Therefore we can expect that AR14 works reasonably well in all cases, which is confirmed by this example. We can also expect that CUSUM procedures based on more than two antirank components will be even more sensitive to shifts. But the procedures themselves will become more complicated especially with many measurement components, requiring a trade-off between the sensitivity of the procedure to a mean shift and the complexity of the procedure. When p is small, we recommend using more antirank components, and to use less

Table 4: Comparisons of the CUSUM procedures based on different antiranks of the p measurements where $p = 4$ and the measurements have joint distribution $N(\mathbf{0}, I_4)$. The in-control ARL is fixed at 200 and $k_1 = 1$. "AR1" denotes the CUSUM procedure based on the first antirank of the measurements. "AR12" denotes a procedure based on the first and the second antiranks. The other notations initiated with "AR" are similarly defined. The control limit values are 6.842 and 15.6887 for AR1-AR4 and AR12-AR34, respectively. The numbers in this table are the out-of-control ARL values and their standard errors (in parentheses) based on 10,000 replications.

	$(\mu_1, \mu_2, \mu_3, \mu_4)$						
	(-4,0,0,0)	(-4,-2,0,0)	(-4,-4,0,0)	(-4,-4,-2,0)	(-4,-2,-2,0)	(-4,-3,-1,0)	(-4,-4,-4,0)
AR1	4.06 (.00)	4.97 (.02)	17.48 (.11)	20.13 (.15)	5.82 (.03)	8.54 (.05)	79.60 (.79)
AR2	79.40 (.77)	7.54 (.04)	18.44 (.13)	32.52 (.30)	35.17 (.33)	12.14 (.09)	78.71 (.77)
AR3	76.73 (.75)	32.92 (.30)	18.59 (.13)	7.55 (.05)	35.38 (.33)	12.19 (.09)	79.34 (.78)
AR4	78.01 (.76)	20.52 (.16)	17.72 (.12)	4.98 (.02)	5.86 (.03)	8.65 (.05)	4.06 (.00)
AR12	4.09 (.02)	2.92 (.01)	3.03 (.01)	3.87 (.02)	4.12 (.02)	3.26 (.02)	11.06 (.08)
AR13	4.12 (.02)	4.29 (.02)	5.58 (.03)	4.70 (.03)	4.85 (.03)	4.83 (.03)	11.41 (.09)
AR14	4.11 (.02)	3.67 (.02)	5.45 (.03)	3.65 (.02)	3.22 (.02)	4.01 (.02)	4.07 (.02)
AR23	11.51 (.09)	4.71 (.03)	5.58 (.03)	4.70 (.03)	5.36 (.03)	4.95 (.03)	11.43 (.09)
AR24	11.49 (.09)	4.70 (.03)	5.57 (.03)	4.27 (.02)	4.85 (.03)	4.83 (.03)	4.11 (.02)
AR34	11.14 (.09)	3.90 (.02)	3.04 (.01)	2.90 (.01)	4.10 (.02)	3.28 (.02)	4.07 (.02)

antirank components otherwise. In most cases including that discussed in Section 4 below, the procedure with the first and last antirank components is highly recommended.

The CUSUM procedure (6)-(7) is designed to test $H_0^{(1)*}$. That is, it is expected to detect a change in the distribution of $A_1(i)$. The previous examples addressed a shift in the mean vector of the process which could result in a change in the distribution of $A_1(i)$. As mentioned in Section 2, sometimes our concern lies with a change in variability of the process. The quality of a process becomes worse if its variability increases and thus we would want to detect quickly such a change in the variance of the measurement vector. If variance changes affect the antirank distribution, then our method can be expected to detect these changes also, which is shown by the following example. In Figure 3, we assume that the null distribution of the measurement vector is $N(\mathbf{0}, I_4)$. As before, the in-control ARL of the CUSUM procedure (6)-(7) is fixed at 200 and $k_1 = 1$. The first measurement is assumed to have a shift in its variance from 1 to σ at the initial time point and σ changes from 0 to 15. The plot shows that the CUSUM procedure is able to detect such a change in variance.

Next we present a numerical example to test both $H_0^{(1)}$ and $H_0^{(2)}$ by combining the CUSUM procedure (6)-(7) with the univariate CUSUM procedure (11)-(12). As mentioned in the previous

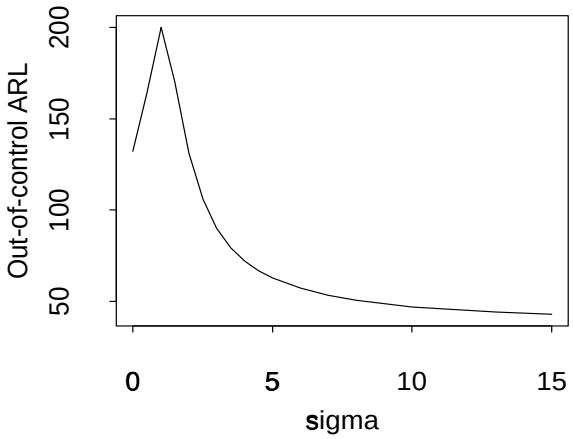

Figure 3: The out-of-control ARL values of the CUSUM procedure (6)-(7) when the variance of the first measurement component shifts from 1 to σ at the initial time point. The null distribution of the measurement vector is $N(\mathbf{0}, I_4)$. The in-control ARL is fixed at 200 and $k_1 = 1$.

sections, the two procedures are designed for two different shift detection purposes and thus need not be used together in applications. A main purpose of this example is to show that these two procedures can support each other in the sense that when one procedure has difficulty in signaling a shift, the other one can signal the same shift well (cf. the related discussion in Section 1). In Table 5, the null distribution of the measurement vector is $N(\mathbf{0}, I_4)$. At the initial time point, the mean vector shifts to μ which takes six possible vectors with unit Euclidean length. The first three mean vectors are ordered by the decreasing Euclidean distance from μ to $(-0.5, -0.5, -0.5, -0.5)$ (which results in a uniform distribution of $A_1(i)$) and by the decreasing numbers of 0 components. The last three vectors are similarly ordered except that the mean components of each vector have the same absolute magnitude 0.5. In the CUSUM procedures, both k_1 and k_2 are fixed at 1. The control limits h_1 and h_2 are searched with step 0.0001 such that the in-control ARL value is 200 and the out-of-control ARL reaches the minimum (denoted as $h_{1,min}$ and $h_{2,min}$, respectively, in the table). The minimum out-of-control ARL value and its standard error based on 10000 replications are listed in the column labeled by $CUSUM_{12}$. The corresponding out-of-control ARL values and their standard errors of the two individual procedures with $(k_1 = 1, h_1 = h_{1,min})$ and $(k_2 = 1, h_2 = h_{2,min})$ are presented in the columns labeled by $CUSUM_1$ and $CUSUM_2$, respectively. The values of $(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$, Q and $\sum_{j=1}^4 \mu_j$ are given in Table 6 for reference. It can be seen from the tables that (i) at least one of the two procedures performs reasonably well for a shift with fixed Euclidean length, and (ii) the combined procedure signals the shift reasonably fast in all cases.

Table 5: The numbers in the column labeled by $CUSUM_{12}$ denote the minimum out-of-control ARL values and their standard errors (in parentheses) of the combined CUSUM procedure (6)-(12) based on 10,000 replications. $h_{1,min}$ and $h_{2,min}$ are the corresponding control limits. The corresponding out-of-control ARL values and their standard errors (in parentheses) of the two individual procedures with $(k_1 = 1, h_1 = h_{1,min})$ and $(k_2 = 1, h_2 = h_{2,min})$ are presented in the columns labeled by $CUSUM_1$ and $CUSUM_2$, respectively.

$(\mu_1, \mu_2, \mu_3, \mu_4)$	$h_{1,min}$	$h_{2,min}$	$CUSUM_1$	$CUSUM_2$	$CUSUM_{12}$
(-1,0,0,0)	6.9900	3.7000	25.00 (.23)	254.65 (2.46)	23.78 (.21)
$(-\sqrt{2}/2, -\sqrt{2}/2, 0, 0)$	10.1300	2.2109	399.39 (3.94)	25.94 (.23)	24.51 (.22)
$(-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}, 0)$	12.5500	2.1885	3297.65 (27.10)	16.28 (.14)	16.18 (.13)
(-0.5,-0.5,0.5,0.5)	6.8500	4.7709	46.72 (.46)	4742.12 (56.72)	46.72 (.46)
(-0.5,-0.5,-0.5,0.5)	11.2500	2.1970	1262.59 (12.73)	49.45 (.46)	48.00 (.44)
(-0.5,-0.5,-0.5,-0.5)	12.7000	2.1886	4314.73 (37.31)	11.49 (.09)	11.49 (.09)

Table 6: The distribution of $A_1(i)$, Q and $\sum_{j=1}^{4} \mu_j$ for several shifts considered in Table 5

$(\mu_1, \mu_2, \mu_3, \mu_4)$	$(g^*_{1,1}, g^*_{1,2}, g^*_{1,3}, g^*_{1,4})$	Q	$\sum_{j=1}^{4} \mu_j$
(-1,0,0,0)	(0.55,0.15,0.15,0.15)	0.48	-1
$(-\sqrt{2}/2, -\sqrt{2}/2, 0, 0)$	(0.37,0.36,0.14,0.13)	0.22	$-\sqrt{2}$
$(-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3}, 0)$	(0.29,0.29,0.29,0.13)	0.08	$-\sqrt{3}$
(-0.5,-0.5,0.5,0.5)	(0.41,0.40,0.09,0.10)	0.38	0
(-0.5,-0.5,-0.5,0.5)	(0.31,0.31,0.31,0.07)	0.18	-1
(-0.5,-0.5,-0.5,-0.5)	(0.25,0.25,0.25,0.25)	0	-2

4 An Application

We illustrate the methods of the paper with a data set (kindly supplied to us by Len Homer) from an aluminum smelter. The data set contains 5 variables – the content of SiO_2, Fe_2O_3, MgO, CaO, and Al_2O_3 (labelled as $x1, x2, x3, x4$ and $x5$ below) in the charge. All these measures are relevant to the operation of the smelter. Stability of the alumina level and calcium oxide level is desirable. The silica, ferric oxide and magnesium oxide levels are affected by the raw materials and are potential covariates to be taken into account in a fully-fledged multivariate scheme. The measures are substantially correlated – both cross-correlated with each other, and autocorrelated.

The data set comprised 189 vectors. We used the first half (95 vectors) to calibrate the models and the second half to test. A calibration sample of this size is smaller than one would like to fully determine the in-control distribution, but suffices to illustrate the use of the method in a real-world setting.

Looking first at the autocorrelation, we modeled each measure by the following kth order

Table 7: Results from the autoregression modeling of the five measurements.

measurement	μ	k	$\alpha_1, \cdots, \alpha_k$
x_1	0.63	3	0.07,0.12,0.28
x_2	24.81	2	0.30,0.24
x_3	12.97	1	0.55
x_4	4.14	1	0.54
x_4	57.86	1	0.32

autoregression model using the S-plus function ar.yw():

$$x(i) - \mu = \alpha_1(x(i-1) - \mu) + \cdots + \alpha_k(x(i-k) - \mu) + \epsilon(i), \tag{13}$$

where $x(i)$ denotes the measure at the ith time point, μ is its mean, and $\epsilon(i)$ is a white noise process with zero mean and variance σ_ϵ^2. The default Akaike's Information Criterion (AIC) is used to determine the value of k. The results are summarized in Table 7. We used this identification to pre-whiten the test data sets into residual vectors ($\epsilon(i)$ in (13)) with cross-correlation, but with their autocorrelations removed (see Lu and Reynolds (1999) for discussions about the relationship between the original measurements and their residuals). The correlation matrix of the five residual vectors is

$$\begin{pmatrix} 1.0000 & 0.2287 & -0.1056 & -0.0277 & -0.3062 \\ 0.2287 & 1.0000 & -0.5702 & -0.0632 & 0.0676 \\ -0.1056 & -0.5702 & 1.0000 & -0.3362 & 0.2205 \\ -0.0277 & -0.0632 & -0.3362 & 1.0000 & -0.2826 \\ -0.3062 & 0.0676 & 0.2205 & -0.2826 & 1.0000 \end{pmatrix}.$$

Apparently the residuals are substantially cross-correlated (e.g., between the second and third residuals).

Figures 4 and 5 show the time series plots and density plots of the residuals. As Figure 5 makes abundantly clear, the marginals are not normal (e.g., the density of x_1 has a long right tail), and so the vector cannot be multivariate normal. It is therefore inadvisable to apply normal-based methods such as Crosier's CUSUM to these data. We used the calibration portion of the data to estimate the distribution of the antirank vector, and then applied our CUSUM to the test set.

Visually, Figure 4 does not suggest step changes in the measures over the course of the data set, so it is interesting to see what multivariate methods can indicate. Figure 6 shows the CUSUM of the first antirank, and of the first-and-last combined antiranks of the residual vectors along with

Table 8: Distributions of $(A_1(i), A_5(i))$ before and after the detected shift in the mean vector.

	Before					After				
	A_5					A_5				
A_1	1	2	3	4	5	1	2	3	4	5
1	0.00	0.02	0.03	0.01	0.05	0.00	0.02	0.03	0.02	0.06
2	0.00	0.00	0.23	0.03	0.01	0.01	0.00	0.07	0.09	0.07
3	0.03	0.22	0.00	0.05	0.03	0.02	0.08	0.00	0.10	0.03
4	0.02	0.03	0.08	0.00	0.01	0.02	0.04	0.08	0.00	0.11
5	0.07	0.01	0.03	0.02	0.00	0.03	0.02	0.06	0.03	0.00

their decision intervals. In both cases, k_1 is fixed at 1 and h_1 is computed to be 11.85 and 34.89, respectively, such that the in-control ARL is 200. The first antirank CUSUM (AR1) broke out of its decision interval briefly around time 153. This brief excursion does not support a step change leading to a change in the first antirank. The first-and-last antirank CUSUM (AR15), however, broke through its decision interval convincingly, almost immediately into the test set, and remained well above for the remainder of the history. This result therefore shows a marked change, which seems to have occurred at or near the midpoint of the data set.

We have some idea of the nature of this change from the antirank distributions (estimated by the relative frequencies) before and after the change. These distributions are presented in Table 8. From the table, it can be seen that the probability $(A_1(i), A_5(i)) = (2, 3)$ is 0.23 before the detected shift while it is 0.07 after the shift. Most other cell probabilities also change. Another apparent difference between the two distributions is that the probabilities that the second and third measures are the largest among all five measures are 0.28 and 0.38 before the detected shift. They decrease to 0.17 and 0.23, respectively, after the shift. Similarly the probabilities that the fourth and fifth measures are the largest among all five measures change from 0.11 and 0.10 to 0.24 and 0.27, respectively. These changes are not at all obvious from the univariate plots.

5 Conclusions

The normal assumption is even more confining with multivariate data than with univariate data, as it is so much harder to escape through variable transformations, and the likelihood of a vector being multivariate normal diminishes with each additional dimension. Despite this situation, the normal distribution underlies the large majority of multivariate methods. Our approach seems particularly valuable as it is able to handle arbitrary data distributions.

Our antirank based procedure is based on the orders of the measurements instead of on the measurements themselves. It might be worth noting that, while the procedure does require extra computation, the added insurance against non-normal data is well worth the slight increase in computation needed to sort the components of a measurement (especially if dimension of measurement is not very large and fast sorts are available). The antirank based procedure can detect shifts in all directions except the one in which the components of the shift are all the same but not zero. If we are interested in detecting shifts in all directions, then combining the antirank CUSUM with a univariate CUSUM of the total of all variables is effective.

We illustrate the use of the approach with an application from on-line multivariate SPC where the multivariate normal is clearly inapplicable. The method is successful in locating an apparent step change that is not otherwise apparent.

Acknowledgements The authors are grateful to the Editor and three referees for helpful comments which greatly improved the article.

Appendix

A Proof of Theorem 1

Without loss of generality, we assume that $\mu_1 = \min(\mu_1, \mu_2, \cdots, \mu_p)$ and $\mu_1 < \mu_2$ when $H_0^{(1)}$ is violated after the shift. Then

$$g_{1,1} = \int \cdots \int_{D_1} F(x_1, x_2, \cdots, x_p) dx_1 \cdots dx_p,$$

and

$$g^*_{1,1} = \int \cdots \int_{D_1^*} F(x_1, x_2, \cdots, x_p) dx_1 \cdots dx_p,$$

where $D_1 = \{(x_1, x_2, \cdots, x_p) : x_1 \leq x_2; \cdots; x_1 \leq x_p\}$ and $D_1^* = \{(x_1, x_2, \cdots, x_p) : x_1 + \mu_1 \leq x_2 + \mu_2; \cdots; x_1 + \mu_1 \leq x_p + \mu_p\}$. Then

$$g^*_{1,1} - g_{1,1} \geq \int \cdots \int_{dD_1} F(x_1, x_2, \cdots, x_p) dx_1 \cdots dx_p,$$

where $dD_1 = \{(x_1, x_2, \cdots, x_p) : x_2 \leq x_1 \leq x_2 + (\mu_2 - \mu_1); x_1 \leq x_3; \cdots; x_1 \leq x_p\}$ which is a point set includes the origin and has a positive Lebesgue area. By the condition in Theorem 1, $g^*_{1,1} - g_{1,1} > 0$. Therefore we have the conclusion of the theorem.

B Proof of Theorem 2

We first assume that there exists a discrete measurement $X_{j_1}(i)$ which has a shift $\mu_{(1)}$ in its mean. Without loss of generality, assume that $j1 = 1$ and $\mu_2 = \max(\mu_1, \mu_2, \cdots, \mu_p)$. Then

$$\begin{aligned} g^*_{1,1} &= P(X_1(i) + \mu_1 \leq X_2(i) + \mu_2, X_1(i) + \mu_1 \leq X_3(i) + \mu_3, \cdots, X_1(i) + \mu_1 \leq X_p(i) + \mu_p) \\ &= P(X_1(i) \leq X_2(i) + (\mu_2 - \mu_1), X_1(i) \leq X_3(i) + (\mu_3 - \mu_1), \cdots, X_1(i) \leq X_p(i) + (\mu_p - \mu_1)) \\ &\geq P(X_1(i) \leq X_2(i) + (\mu_2 - \mu_1), X_1(i) \leq X_3(i), \cdots, X_1(i) \leq X_p(i)), \end{aligned}$$

and

$$g^*_{1,1} - g_{1,1} = P(X_2(i) < X_1(i) \leq X_2(i) + (\mu_2 - \mu_1), X_1(i) \leq X_3(i), \cdots, X_1(i) \leq X_p(i)) = \nu_F(dD_1),$$

where $dD_1 = \{(x_1, x_2, \cdots, x_p) : x_2 \leq x_1 \leq x_2 + (\mu_2 - \mu_1); x_1 \leq x_3; \cdots; x_1 \leq x_p\}$. In the theorem it is assumed that $\mu_{(p)} - \mu_{(1)} \geq w_{j_1}$. Therefore $dD_1 \bigcap \mathcal{S} \neq \emptyset$. Apparently dD_1 is an open set and has

a positive Lebesgue area. Therefore $g^*_{1,1} > g_{1,1}$ by using the assumption on ν_F. The theorem can be similarly proved in the case that there is a continuous measurement which has a shift $\mu_{(1)}$ in its mean.

C Derivation of (8)

If $C_n - k_1 > 0$, then it can be seen from the definition of the CUSUM that

$$\mathbf{S}^{(1)}_{n-1} + \xi_{\mathbf{1}}(n) = \frac{C_n}{C_n - k_1}\mathbf{S}^{(1)}_n$$

and

$$\mathbf{S}^{(2)}_{n-1} + \mathbf{g_1} = \frac{C_n}{C_n - k_1}\mathbf{S}^{(2)}_n.$$

Therefore

$$\begin{aligned} C_n &= \frac{C_n}{C_n - k_1}(\mathbf{S}^{(1)}_n - \mathbf{S}^{(2)}_n)'(\frac{C_n}{C_n - k_1})^{-1} diag(1/S^{(2)}_{n,1}, \cdots, 1/S^{(2)}_{n,p})\frac{C_n}{C_n - k_1}(\mathbf{S}^{(1)}_n - \mathbf{S}^{(2)}_n) \\ &= \frac{C_n}{C_n - k_1} y_n. \end{aligned}$$

So $y_n = C_n - k_1$ and consequently (8) is true in such case. It is obvious that (8) is true when $C_n - k_1 \leq 0$.

D Explanation of (9)

When $n = 1$,

$$\begin{aligned} C_1 &= (\xi_{\mathbf{1}}(1) - \mathbf{g_1})' diag(1/g_{1,1}, \cdots, 1/g_{1,p})(\xi_{\mathbf{1}}(1) - \mathbf{g_1}) \\ &= \sum_{j=1}^{p} (\xi_{1,j}(1) - g_{1,j})^2 / g_{1,j} \end{aligned}$$

because $\mathbf{S}^{(1)}_0 = \mathbf{S}^{(2)}_0 = \mathbf{0}$. By the definition of $\xi_{\mathbf{1}}(1)$ (cf. (4)), only one of its components equals 1 and the remaining components are all 0. Without loss of generality, let us assume that $\xi_{1,1}(1) = 1$ and $\xi_{1,2}(1) = \cdots = \xi_{1,p}(1) = 0$. Then

$$\begin{aligned} C_1 &= (1 - g_{1,1})^2 / g_{1,1} + g_{1,2} + \cdots + g_{1,p} \\ &= (1 - g_{1,1})^2 / g_{1,1} + (1 - g_{1,1}) \\ &= (1 - g_{1,1}) / g_{1,1} \\ &= \frac{\sum_{j \neq 1} g_{1,j}}{g_{1,1}} \end{aligned}$$

If $k_1 \geq \max_{\ell=1}^{p} \frac{\sum_{j \neq \ell} g_{1,j}}{g_{1,\ell}}$, then $C_1 \leq k_1$ and consequently $\mathbf{S}_1^{(1)} = \mathbf{S}_1^{(2)} = 0$. Similarly, it can be shown that $\mathbf{S}_n^{(1)} = \mathbf{S}_n^{(2)} = 0$ for any $n \geq 1$ when k_1 satisfies this condition. That is, the CUSUM procedure is restarted at each time point. To avoid this phenomenon, it is necessary to choose k_1 from the interval given by (9).

References

Alt, F.B. (1985), "Multivariate quality control," in *Encyclopedia of the Statistical Sciences* (N.L. Johnson, S. Kotz, and C.R. Read eds.), **vol. 6**, 111-122.

Crosier, R.B. (1988), "Multivariate generalizations of cumulative sum quality-control schemes," *Technometrics*, **30**, 291-303.

Gan, F.F. (1995), "Joint monitoring of process mean and variance using exponentially weighted moving average control charts," *Technometrics*, **37**, 446-453.

Gibbons, J.D., and Chakraborti, S. (1992), *Nonparametric Statistical Inference*, Marcel Dekker, Inc., New York.

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

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

Healy, J.D. (1987), "A note on multivariate CUSUM procedures," *Technometrics,* **29**, 409-412.

Hotelling, H. (1947), "Multivariate quality control - illustrated by the air testing of sample bomb sights," in *Techniques of Statistical Analysis* (C. Eisenhart, M.W. Hastay, and W.A. Wallis eds.), McGraw Hill, New York.

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

Johnson, M.E. (1987), *Multivariate Statistical Simulation*, John Wiley & Sons, New York.

Liu, Y. (1996), "An improvement for MEWMA in multivariate process control," *Computers & Industrial Engineering,* **31**, 779-781.

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.

Lu, C.W., and Reynolds, M.R., Jr. (1999), "Control charts for monitoring the mean and variance of autocorrelated processes," *Journal of Quality Technology*, **31**, 259-274.

Lucas, J.M., and Saccucci, M.S. (1990), "Exponentially weighted moving average control schemes: properties and enhancements," *Technometrics*, **32**, 1-12.

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**, 140-143.

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

Read, T., and Cressie, N. (1988), *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer, New York.

Woodall, W.H., and Ncube, M.M. (1985), "Multivariate CUSUM quality-control procedures," *Technometrics*, **27**, 285-292.

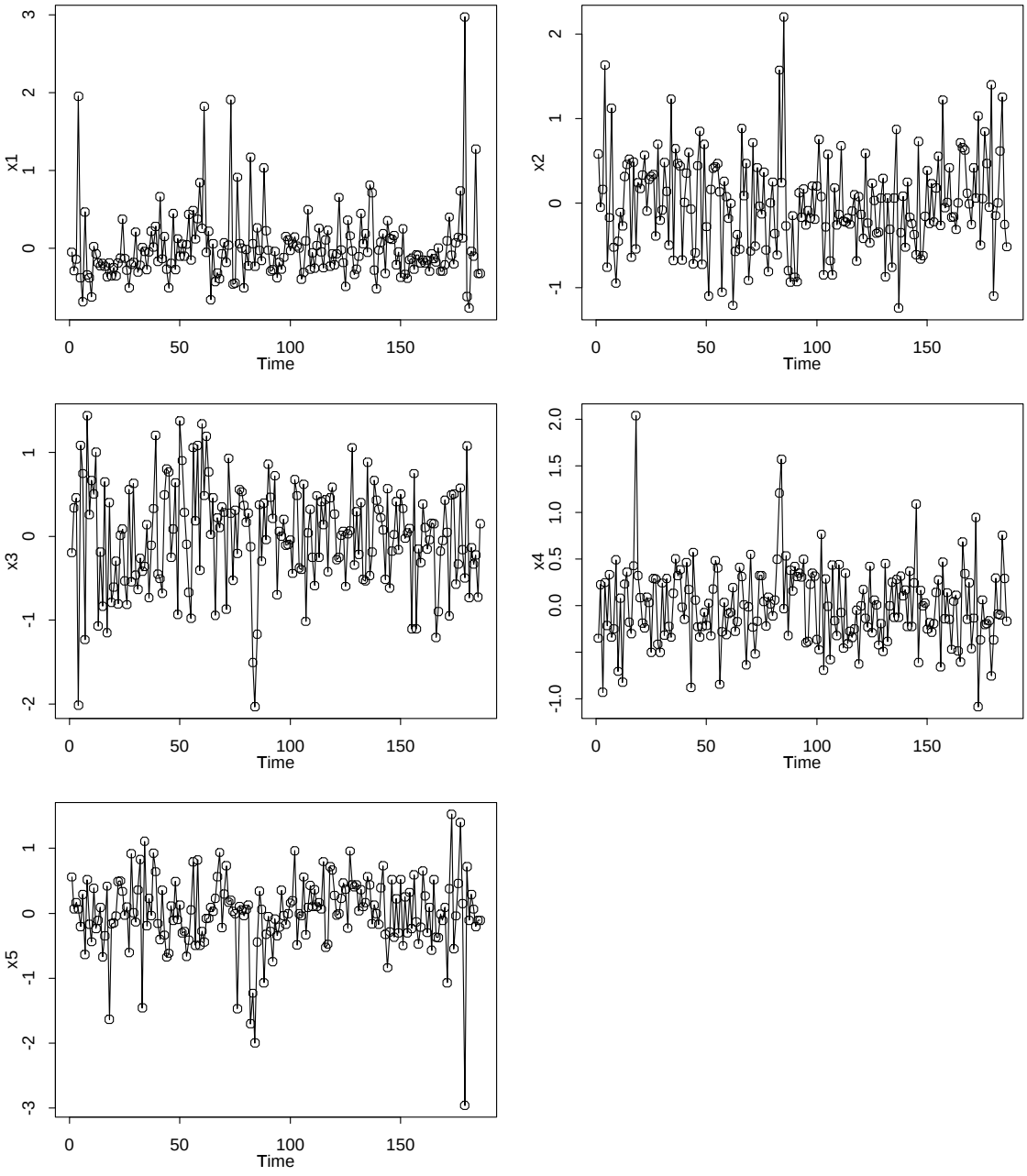

Figure 4: The data after autocorrelation in each measurement is excluded.

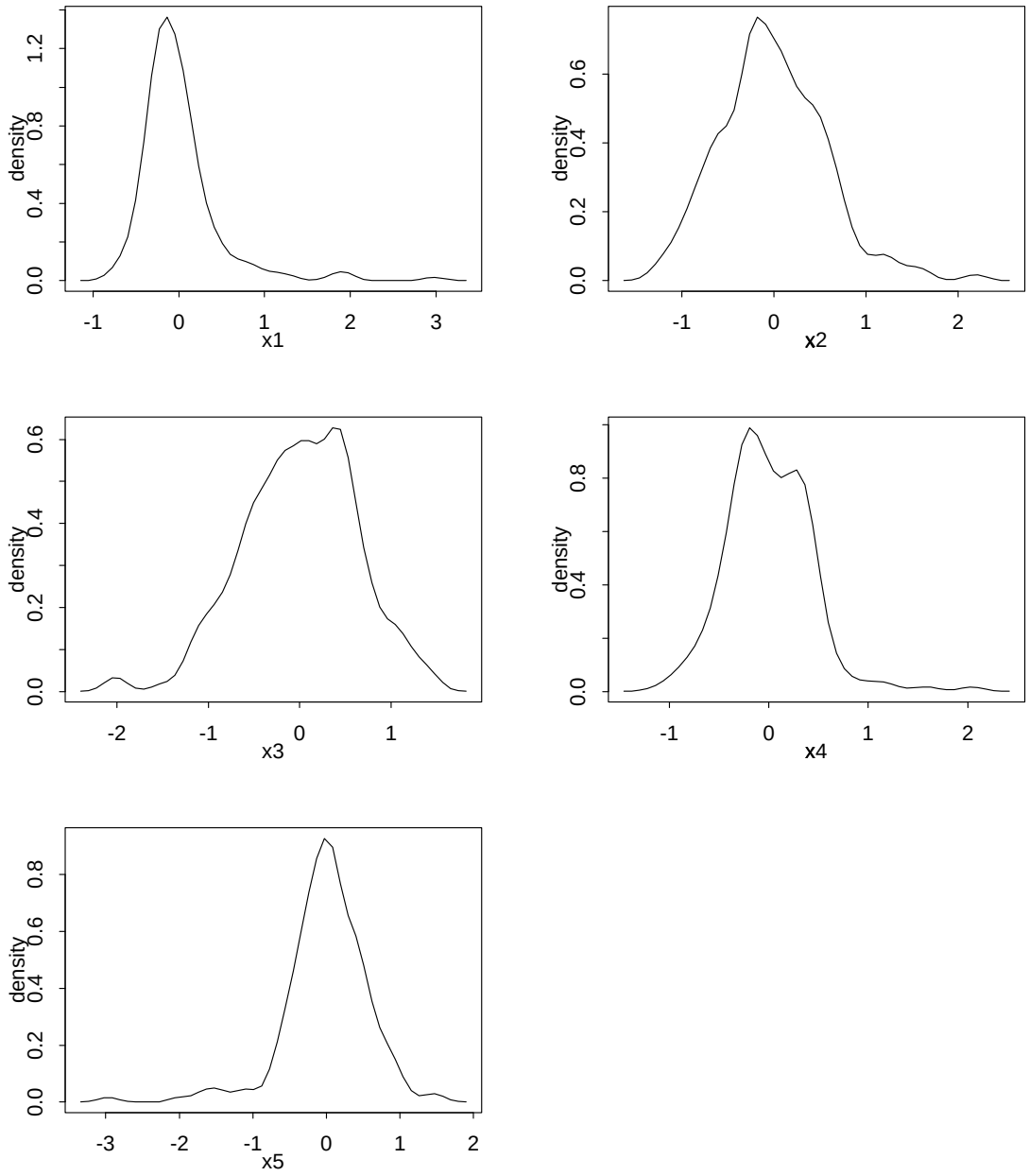

Figure 5: The density plots of five measurements after autocorrelation is excluded.

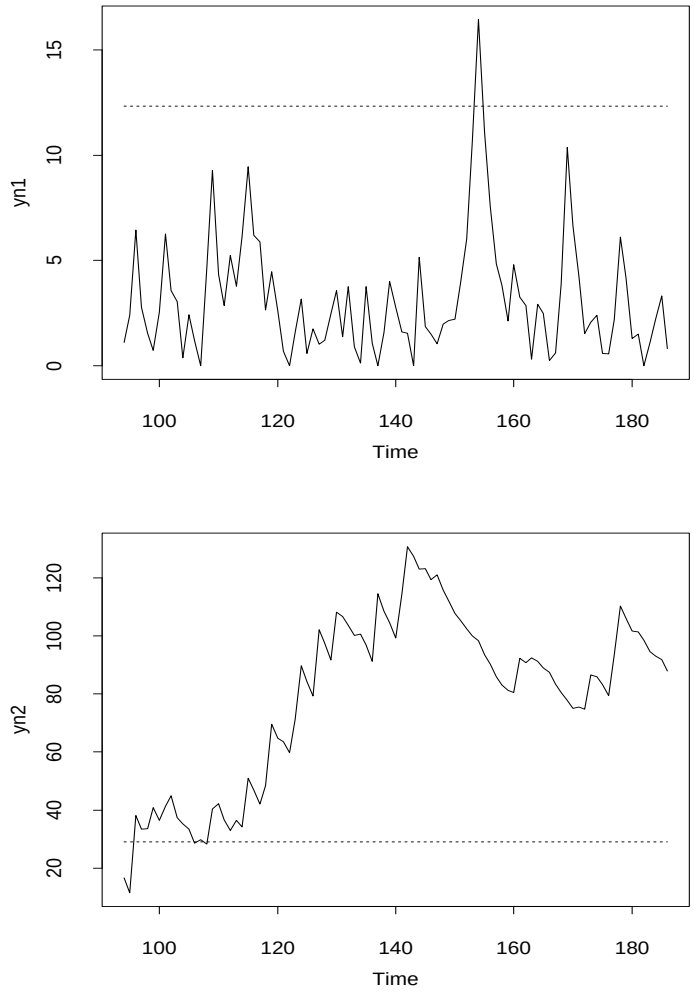

Figure 6: The above panel displays the CUSUM criterion y_n of the procedure AR1. The below panel displays the same criterion for AR15. The dotted lines indicate the control limits of the procedures such that their in-control ARL values are both 200.