A Distribution-Free Multivariate Phase I Location Control Chart for Subgrouped Data from Elliptical Distributions

Article *in* Technometrics · January 2013

DOI: 10.1080/00401706.2013.879264

CITATIONS
24

READS
286

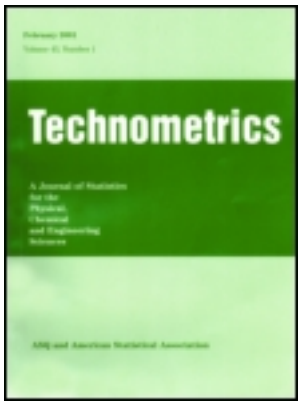
3 authors, including:

L.A. Jones-Farmer
Miami University
83 PUBLICATIONS **3,929** CITATIONS

SEE PROFILE

Nedret Billor
Auburn University
52 PUBLICATIONS **1,269** CITATIONS

SEE PROFILE



Technometrics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/utch20

A Distribution-Free Multivariate Phase I Location Control Chart for Subgrouped Data from Elliptical Distributions

Richard C. Bell Jr. [a] , L. Allison Jones-Farmer [b] & Nedret Billor [c]

[a] Office of the Deputy Chief of Staff, G-8, Headquarters, Department of the Army , Washington , DC , 20310

[b] Department of Aviation and Supply Chain Management , Auburn University , Auburn , AL , 36849

[c] Department of Mathematics and Statistics , Auburn University , Auburn , AL , 36849

Accepted author version posted online: 16 Jan 2014.Published online: 16 Jan 2014.

To cite this article: Technometrics (2014): A Distribution-Free Multivariate Phase I Location Control Chart for Subgrouped Data from Elliptical Distributions, Technometrics, DOI: 10.1080/00401706.2013.879264

To link to this article: http://dx.doi.org/10.1080/00401706.2013.879264

A Distribution-Free Multivariate Phase I Location Control Chart for Subgrouped Data from Elliptical Distributions

Richard C. BELL, JR.

Office of the Deputy Chief of Staff, G-8

Headquarters, Department of the Army

Washington, DC 20310

(*richard.bell1@us.army.mil*)

L. Allison JONES-FARMER

Department of Aviation and Supply Chain Management

Auburn University

Auburn, AL 36849

(*joneall@auburn.edu*)

Nedret BILLOR

Department of Mathematics and Statistics

Auburn University

Auburn, AL 36849

(*billone@auburn.edu*)

In quality control, a proper Phase I analysis is essential to the success of Phase II monitoring. A literature review reveals no distribution-free Phase I multivariate techniques in existence. This research develops a Phase I location control chart for multivariate elliptical processes. The resulting in-control reference sample can then be used to estimate the parameters for Phase II monitoring. Using Monte Carlo simulation, the proposed method is compared to the Hotelling's T^2 Phase I chart. Although Hotelling's T^2 chart is preferred when the data are multivariate normal, the proposed method is shown to perform significantly better under non-normality. This paper has supplementary material online.

KEY WORDS: Data Depth, Mahalanobis Depth, Mean Rank Chart, Nonparametric, Outliers, Retrospective Analysis, Robust Estimators

1. INTRODUCTION

Multivariate quality control charting has grown in both popularity and relevance since its introduction by Hotelling (1947). In a review of statistical process control research issues, Woodall and Montgomery (1999) pointed out the rise in multivariate quality control research due to increased measurement capability and computing power. Montgomery (2009) noted that the advent of larger manufacturing databases has greatly increased the use of multivariate quality control methods in recent years. Another important application area for multivariate control charts is in the monitoring of profiles, where a quality characteristic can be expressed as a modeled function of one or more explanatory variables. Woodall et al. (2004) give an overview and literature review of control charts used in profile monitoring.

A control charting application is typically divided into two distinct phases. Phase I involves the analysis of a data set to establish the in-control (IC) state of the process and identify a baseline reference sample. Upon completion of Phase I, the IC reference sample can be used to establish control limits for Phase II, the monitoring stage of a control charting application. In Phase II, process observations are prospectively compared to the control limits to identify significant departures from the IC state.

The purpose of this paper is to introduce a Phase I method to detect either isolated or sustained shifts in the location vector of a multivariate process. The control chart method we introduce is based on the concept of ranking data, similar to the mean rank chart introduced by Jones-Farmer et al. (2009) for univariate Phase I location analysis, extended to multivariate processes. Because ranking observations in multivariate space may be unfamiliar to the reader, in Section 2 we introduce an area of computational geometry known as data depth. In Section 3,

we give a brief background on multivariate control charts in Phase I. We introduce a Phase I location chart for multivariate elliptical processes in Section 4. In Section 5, we provide a Monte Carlo simulation study comparing the performance of our proposed method to the Phase I Hotelling's T^2 chart. An example application of the proposed method is given in Section 6. Finally, in Section 7 we offer advice to practitioners for moving from a Phase I to a Phase II analysis.

2. BACKGROUND ON DATA DEPTH

The concept of data depth originated with Tukey (1975), and there are many types of data depth functions, with many different mathematical properties and differing levels of computational complexity. A *data depth* measures how deep (or central) a point $\boldsymbol{x} \in \mathbb{R}^p$ is with respect to a certain probability distribution *F* or a given data cloud $\mathbb{X}_n = \{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$ in $\mathbb{R}^p$. A data depth is computed by applying one of many known data depth functions to a multivariate data point, thus reducing it from a *p*-vector to a univariate depth value. Assuming unimodality of the data, a large depth value indicates centrality and a small depth value suggests outlyingness of a given point. The point of maximal depth is considered the center of the data and is referred to as the multivariate median.

A data depth function may be visualized in *p*-dimensional space as a series of nested contours around the multivariate median, where each contour represents the set of *p*-dimensional points with equal depth values. Some depth functions force contours of a particular geometric form and are suitable for use with elliptical distributions, whereas others allow contours to follow the actual geometric shape of the data.

Elliptical distributions are generalizations of the multivariate normal distribution. These distributions are defined by their characteristic functions and differ from the multivariate normal in the fourth and higher moments. Example univariate elliptical distributions and their multivariate extensions include the normal, Laplace, *t*-, Cauchy, and logistic distributions (Jensen, 1985). There are methods suggested for testing elliptical symmetry of a multivariate density by Beran (1979), Huffer and Park (2007), and Li et al. (1997). Data depth methods are often designed and evaluated with respect to either elliptical or non-elliptical distributions. In general, the depth functions designed for use with non-elliptical distributions are more computationally expensive.

Data depth facilitates the extension of order statistics to higher dimensions because depth values can be ranked from largest to smallest to produce a center-outward ordering of the data. This rank-based perspective makes data depth potentially very useful as a distribution-free method in multivariate analysis. For a detailed review of several depth functions, see Liu et al. (1999), Zuo and Serfling (2000a,b), Mosler (2002) and the references therein.

For a data depth function to serve most effectively as an analytical tool, the following four properties are required [Liu (1990), Zuo and Serfling (2000a)].

- Property 1: Affine invariance. The depth of a point $\boldsymbol{x} \in \mathbb{R}^p$ should not depend on the underlying coordinate system or the scales of the underlying measurements.
- Property 2: Maximality at center. For a distribution having a uniquely defined "center", the depth function should attain maximum value at this center.
- Property 3: Monotonicity relative to deepest point. As a point $\boldsymbol{x} \in \mathbb{R}^p$ moves away from the "center", the depth at **x** should decrease monotonically.

- Property 4: Vanishing at infinity. The depth of a point, **x,** should approach zero as $\|\mathbf{x}\| \to \infty$, where $\|\mathbf{x}\|$ is the Euclidean norm of **x**.

According to Zuo and Serfling (2000a), depth functions which satisfy these four properties are particularly well suited for nonparametric multivariate inference; thus, these properties will serve as a useful basis for selecting particular data depth functions for implementation in this research.

Although our Phase I control chart can be used with any data depth procedure, we limit our focus to Mahalanobis depth and simplicial depth which have been shown to satisfy the four desirable properties stated above. We consider Mahalanobis depth because it balances desirable mathematical properties with computational efficiency. We consider simplicial depth because it was first applied to Phase II quality control by Liu (1995) and Liu et al. (2004) and studied by Stoumbos and Jones (2000) and Stoumbos et al. (2000).

2.1 Mahalanobis Depth

The *Mahalanobis depth* (*MD*) of a point ***x*** in $\mathbb{R}^p$ with respect to a distribution *F* in $\mathbb{R}^p$ is defined as

$$MD(\mathbf{x}; F) = [1 + d^2_{\boldsymbol{\Sigma}}(\mathbf{x}, \boldsymbol{\mu})]^{-1},$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the location vector and covariance matrix of F, respectively and $d^2_{\mathbf{M}}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})'\mathbf{M}^{-1}(\mathbf{x} - \mathbf{y})$ is the Mahalanobis distance [Mahalanobis (1936)] between two points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^p$ with respect to a positive definite *p* x *p* matrix $\mathbf{M}$. The sample version is obtained by replacing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ with their sample estimators, $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$**,** respectively. Thus, the sample version of the depth function is given by

$$MD(\mathbf{x}; F_n) = [1 + d^2_{\hat{\Sigma}}(\mathbf{x}, \hat{\boldsymbol{\mu}})]^{-1}, \quad (1)$$

where F_n denotes the sample empirical distribution function. Large values of $d^2_{\hat{\Sigma}}(\mathbf{x}, \hat{\boldsymbol{\mu}})$ will result in small values of $MD(\mathbf{x}; F_n)$; thus, points that are outlying with respect to the sample mean will result in very small Mahalanobis depth values. The Mahalanobis depth function is most appropriate when the underlying distribution, *F,* is elliptical because it produces elliptical contours of equal depth.

The Mahalanobis depth function can be used in conjunction with robust estimators, and will be referred to as robust Mahalanobis depth (RMD) when $\hat{\boldsymbol{\mu}}$ and $\hat{\Sigma}$ are robust location and/or scale estimators. There are numerous robust estimation methods from which to choose. Because of its computational efficiency, we will use the blocked adaptive computationally efficient outlier nominators (BACON) method of Billor et al. (2000) for robust parameter estimation. Two versions of the BACON method are available: Version 1 has a replacement breakdown point (RBP) of approximately 20% , while Version 2 of the BACON method has a RBP exceeding 40% due to a robust starting point. Although Version 2 has higher RBP, the robust starting point is not affine equivariant and the method is computationally expensive. However since the subsequent iterations are affine equivariant, the overall algorithm based on Version 2 tends to be nearly affine equivariant (Billor et al., 2000). Version 1 offers an affine equivariant robust starting point and is computationally less expensive than Version 2. Therefore, when computing RMD, we used Version 1 of the BACON method with a Type I error probability of 0.10 to estimate the mean vector for applications with 20% or fewer OC observations, and Version 2 when more than 20% of the observations were OC. We estimated

the covariance matrix with $\hat{\Sigma} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{S}_i$, where $\boldsymbol{S}_i$ is the sample covariance matrix of the i^{th} subgroup, because it is robust to changes in location.

2.2 Simplicial Depth

Introduced by Liu (1990), the *simplicial depth* (SD) of a point $\mathbf{x}$ in $\mathbb{R}^p$ with respect to a distribution F in $\mathbb{R}^p$ is defined as the probability that $\mathbf{x}$ belongs to a random simplex in $\mathbb{R}^p$, formally stated as

$$SD(\mathbf{x}; F) = P_F(\mathbf{x} \in S[\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}]),$$

where $\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}$ are independent observations from F and $S[\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}]$ denotes the p-dimensional simplex with vertices $\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}$, or the set of all points in $\mathbb{R}^p$ that are convex combinations of $\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}$. When F is unknown, a reference sample, $\{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$, from F can be used to compute the sample simplicial depth which is defined to be

$$SD(\mathbf{x}; F_n) = \binom{n}{p+1}^{-1} \sum_{1 \le l_1 < \ldots < l_{p+1} \le n} I(\mathbf{x} \in S[\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}]),$$

where F_n denotes the empirical distribution function of $\{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$, $n \ge p+1$, the summation runs over all subsets of *(p+1)* points from $\{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$, and

$$I(\mathbf{x} \in S[\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}]) = \begin{cases} 1, & \text{if } \mathbf{x} \in S[\boldsymbol{X}_1, ..., \boldsymbol{X}_{p+1}] \\ 0, & \text{otherwise.} \end{cases}$$

For example, if $\mathbf{x} \in \mathbb{R}^2$, $SD(\mathbf{x}; F_n)$ is the proportion of closed triangles among the triangles formed from all possible triplets of points from $\{\boldsymbol{X}_1, ..., \boldsymbol{X}_n\}$. If the assumption of elliptical structure of the data is in doubt or no particular distributional information in the data is known, one should use a geometric depth such as simplicial depth. Rousseeuw and Ruts (1996) provided

an algorithm for computing simplicial depth in the bivariate case. For dimensions greater than two, computing simplicial depth remains an open problem.

3. BACKGROUND ON PHASE I METHODS

3.1 Multivariate Phase I and Nonparametric Control Charts

There are a number of control chart methods developed for use in Phase I, though they are mostly variations of Hotelling's T^2 control chart based on the assumption of a multivariate normally distributed process. The T^2 control chart can be applied to individual data in Phase I using control limits outlined by Tracy et al. (1992). Nedumaran and Pignatiello (2000) addressed the issue of constructing T^2 control chart limits for retrospective testing when the parameters of a subgrouped multivariate normally distributed process are unknown. Later we will show that Hotelling's T^2 chart does not perform well when the process distribution is not multivariate normal.

Phase II nonparametric, distribution-free, and robust multivariate control charting methods have also been developed [see, e.g., Hayter and Tsui (1994) , Qiu and Hawkins (2001), Chou et al. (2001), Qiu and Hawkins (2003), Sun and Tsung (2003), Thissen et al. (2005), and Qiu (2008)]. Some nonparametric multivariate Phase II control charts have been proposed using data depth. Liu (1995) suggested using transformed simplicial depth-based ranks instead of raw univariate data to compute the control statistics for X, $\overline{X}$, and cumulative sum (CUSUM) charts. Liu et al. (2004) later introduced a simplicial data depth-based moving average (DDMA) control chart.

Zarate (2004) used principal components analysis, and then used a data depth-based control chart to monitor some of the principal components in Phase II. Beltran (2006) used Liu's

(1995) *r* chart in conjunction with the simplicial depth ranks of the first and last principal components of the process. Messaoud et al. (2008) proposed a data depth-based exponentially weighted moving average (EWMA) control chart for multivariate observations. Hamurkaroglu et al. (2004) proposed the use of Liu's (1995) *r* and *Q* charts using the ranks of Mahalanobis depth rather than simplicial depth. Stoumbos and Jones (2000) and Stoumbos et al. (2000) investigated the performance of Phase II control charts based on simplicial depth, and concluded that the IC reference sample size requirements were large, and depended on the underlying process distribution.

While there are many nonparametric control charts developed for Phase II, they all require an IC reference sample. Little advice has been given as to how to attain the IC sample when the distribution is unknown or the multivariate normality assumption cannot be met.

3.2 Phase I Control Chart Performance Measures

Typical performance metrics, such as average run length (ARL) and average time to signal (ATS), are not relevant for designing, evaluating, or comparing Phase I methods. IC performance of Phase I charts is usually characterized by the false alarm probability (FAP), which is the probability of observing a signal when the process is IC. We will use the FAP to design and evaluate the IC performance of the Phase I methods.

When evaluating the out-of-control (OC) performance of Phase I methods, most authors have used alarm probabilities, or the probability of observing one or more alarms when the process is OC [see, e.g., Sullivan and Woodall (1996), Vargas (2003), Woodall et al. (2004), Jensen et al. (2007), Alfaro and Ortega (2008), Jobe and Pokojovy (2009), Jones-Farmer et al. (2009), and Jones-Farmer and Champ (2010)]. In OC situations, it seems important to also

evaluate a Phase I method based on its ability to correctly identify *the right* observations as outliers as recommended by Chen et al. (2013). The OC observations will either be correctly classified as OC, or incorrectly classified as IC by the control chart. We will use the percent of correctly classified OC events as our primary measure to evaluate and compare the OC performance of Phase I methods. Fraker et al. (2008) provides an excellent discussion of performance measures for control charts used in surveillance schemes, including those useful for retrospective analysis.

4. THE MULTIVARIATE MEAN-RANK (MMR) CHART

The chart introduced here is the multivariate analog of Jones-Farmer et al.'s (2009) Phase I mean-rank chart for *univariate* processes. The mean-rank chart is similar to the $\bar{X}$ chart for univariate data, but it uses the standardized average subgroup rank rather than the average of raw subgroup data values as a control statistic. The use of ranks rather than actual data values renders the method distribution free, since the IC distribution of ranks is the same regardless of the underlying process distribution. The mean-rank chart modified for use with ranked data depth values from a multivariate process will be hereafter referred to as the multivariate mean-rank (MMR) chart.

4.1 The MMR Control Statistic

Consider a reference sample consisting of *m* subgroups of size *n* from a *p*-dimensional multivariate process in which all variables are continuous. Let the random vector $\mathbf{x}_{ij}$ represent the 1 x *p* row vector containing the j^{th} observation from the i^{th} subgroup. Treating the observations from the *m* mutually independent samples of size *n* as a single sample of size $N = n \text{ x } m$, a data depth function is applied to each $\mathbf{x}_{ij}$, resulting in a corresponding depth

value, $D(\mathbf{x}_{ij};F_N)$, where F_N denotes the empirical distribution function of the pooled reference sample.

Next, integer ranks R_{ij} = 1, 2,..., N are assigned to each $D(\mathbf{x}_{ij};F_N)$ in the pooled sample of size N, beginning with the largest depth value and continuing in descending order. The depth function, $D(\mathbf{x}_{ij};F_N)$, can be any data depth function, but we recommend using one that satisfies the four desirable properties described by Zuo and Serfling (2000a).

When the process is IC, the mean of the random variable R_{ij} is $E(R_{ij}) = (N+1)/12$ and the variance is $Var(R_{ij}) = (N-1)(N+1)/12$. In the event of a tie, the midrank method can be used as a correction without affecting the mean and variance of the random variable R_{ij}. The average of the ranks in each subgroup i, is denoted by $\bar{R}_i = \frac{\sum_{j=1}^{n} R_{ij}}{n}$. If a process is IC, the ranks should be distributed evenly throughout the m subgroups, resulting in approximately equal $\bar{R}_i$ for each subgroup. The chart statistic for the MMR control chart is the standardized subgroup mean rank, given by $Z_i = \left(\bar{R}_i - E(\bar{R}_i)\right)/\sqrt{Var(\bar{R}_i)}$, where $E(\bar{R}_i) = (N+1)/2$, and $Var(\bar{R}_i) = (N-n)(N+1)/12n.$

4.2 Control Limits for the MMR Chart

Like many multivariate control charts, the MMR chart has only an upper control limit (UCL). Recall that large depth values indicate centrality and will receive low ranks; thus, a near zero or negative value of Z_i indicates centrality with respect to the p-dimensional data cloud and

a likely IC point. Conversely, an extremely positive value of Z_i is realized when a subgroup has an extremely large mean rank, which may be indicative of an OC condition.

It is possible to use asymptotic control limits for the MMR chart that depend only on the number of subgroups, (m). Details on the asymptotic control limits for the MMR chart are given as in Section S.1 of the supplementary materials. Table 1 provides empirically derived control limits for several sample (m) and subgroup sizes (n). The empirical limits for the MMR chart depend on m and n, but do not depend on the dimension, p. Due to the discrete nature of the mean-rank distribution, simulated FAP values do not exactly match the desired FAP values; however, conservative limits were chosen. Because the empirical FAP values more closely match the desired values than those derived from the asymptotic limits, we recommend use of the empirically derived limits for the MMR chart. The MATLAB code used to generate the empirically derived control limits is provided as supplementary material to this paper.

5. THE PERFORMANCE OF THE MMR CHART

5.1 The Simulation Study Design

Because there are few distribution-free multivariate Phase I Shewhart-type control chart methods to serve as a basis of comparison for the MMR chart, the performance of the MMR chart based on RMD and simplicial depth (MMR-RMD and MMR-SD, respectively) will be compared to the Hotelling's T^2 chart adjusted for Phase I analysis.

For both the MMR-RMD and T^2 charts, we use the BACON estimator to robustly estimate the mean vector. When the data contain less than 20% contaminated points, we use Version 1 of Billor et al.'s (2000) BACON estimator, and Version 2 of the BACON estimator when

contamination rates exceed 20%. For all simulations, we use a Type I error probability of 0.10 for the BACON estimator, and estimate the covariance matrix using $\hat{\Sigma} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{S}_i$. Limits for the MMR-RMD and MMR-SD charts were computed empirically according to the method presented in Section 4.2.

When the IC process parameters are unknown, the T^2 chart statistic is given by

$$T_i^2 = n\left(\overline{\mathbf{X}}_i - \hat{\boldsymbol{\mu}}\right)' \hat{\Sigma}^{-1}\left(\overline{\mathbf{X}}_i - \hat{\boldsymbol{\mu}}\right),$$

for $i = 1, ..., m$. When the process observations follow a *p*-variate normal distribution, the control limits are computed according to

$$UCL_{T^2} = C(m, n, p)F_{\alpha, p, mn-m-p+1}, \quad (2)$$

where $C(m, n, p) = p(m-1)(n-1)/(mn-m-p+1)$, $F_{\alpha, p, mn-m-p+1}$ represents the (1 - α)$^{\text{th}}$ percentile of the *F* distribution with *p* and (*mn* - *m* - *p* + 1) degrees of freedom, and α is the desired IC FAP for each *individual* subgroup (Alt, 1976). Nedumaran and Pignatello (2000) showed that a retrospective T^2 chart with control limits computed according to (2) using a Bonferroni error adjustment performed nearly as well as computationally cumbersome exact retrospective limits. In order to achieve a desired *overall* IC FAP for all *m* subgroups in a reference data set, α is selected as $\alpha = 1 - \left(1 - \alpha_{overall}\right)^{1/m}$, where $\alpha_{overall}$ is the desired IC FAP of the chart.

Both IC and OC performance of the MMR and Hotelling's T^2 charts will be compared using observations from multivariate normal and heavy-tailed elliptical distributions. Multivariate normal data are generated from the standard multivariate normal distribution $\left(\Sigma = I\right)$, and heavy-tailed data are represented by the multivariate *t*-distribution with 10 and 3

degrees of freedom. All results were simulated using MATLAB code from the MathWorks Statistics Toolbox at http://www.mathworks.com/help/toolbox/stats/. Cases based on simplicial depth were computed using the 'depth' package in R [Genest et al. (2013)].

An isolated shift in the mean is defined as a location shift occurring in a single subgroup of size *n*, and takes place in the first subgroup of each simulated data set. A sustained shift in the mean is defined as a location shift that occurs in several subgroups in a row for a sustained period. Similar to the methods used by Sullivan and Woodall (1996), Vargas (2003), and others, sustained shifts were induced in the last 5%, 15%, 20%, 25%, and 30% of the subgroups.

The magnitude of shifts is measured by the noncentrality parameter, $\lambda = \sqrt{\boldsymbol{\delta}\boldsymbol{\delta}'}$, where the process mean vector shifts from $\boldsymbol{\mu}_o$ to $\boldsymbol{\mu}_o + \boldsymbol{\delta}$. Because the direction of a shift does not affect control chart performance with elliptically symmetric distributions [see, Stoumbos and Sullivan (2002, pp. 274-275)], shifts will be fixed in the direction of $\boldsymbol{e}_1 = [1, 0, ..., 0]$. OC control chart performance is quantified in terms of the percent of OC events correctly classified as OC.

The simulation algorithm follows:

1) Establish the UCL for the MMR or Hotelling's T^2 chart.
2) Simulate *m* subgroups of size *n* from a *p*-dimensional multivariate normal or *t* distribution.
3) For IC comparisons, go to step #4. For OC comparisons, add isolated or sustained location shifts to the desired subgroups.
4) Compute control chart statistics for each subgroup and compare to the UCL.
5) Repeat steps 2 - 4 a total of 10,000 times.

This process is repeated for all combinations of m, n, p, process distribution, shift type, and control chart. **Equation Section (Next)**Using 10,000 iterations gives a standard error of the estimated FAP near 0.003 for the cases considered. Results were generated for m = 20, 50(50)200 subgroups of size n = 5(5)20, with the desired IC FAP set to 0.10. Similar results were obtained for desired IC FAP values of 0.05. We considered dimensions of p = 2, 5, and 10, but due to space constraints, we limit the results we present to those that best represent the overall story of the MMR chart's relative performance.

5.2 IC Chart Performance

The fundamental advantage of a distribution-free control chart is its ability to maintain a desired IC FAP for any process distribution. Figure 1 shows the IC performance of the bivariate MMR-RMD and Hotelling's T^2 charts under a variety of distribution, sample, and subgroup size combinations. Because the MMR-SD chart performs nearly identically to the MMR-RMD chart in the IC conditions, the MMR-SD cases are not shown. In each case depicted in Figure 1, the MMR-RMD chart maintains the desired IC FAP of 0.10. Although it is difficult to see in Figure 1, Hotelling's T^2 chart also maintains the desired IC FAP for the bivariate normal process; however, the performance of the T^2 chart becomes progressively worse as the distribution deviates from normality and the number of subgroups increases. For a bivariate $t(3)$ process, the IC FAP for Hotelling's T^2 chart ranges from approximately 30% when m = 20 to over 90% when m = 200. The poor performance of the T^2 chart as the number of subgroups increases is due to the division of the desired FAP into an increasing number of Phase I comparisons. The tail behavior of the multivariate t-distribution deviates from that of the multivariate normal, and this is especially evident in the extreme tails of the distribution.

Figure 2 illustrates the effects of dimensionality on the IC control chart performance using a $t(3)$ process. The MMR-RMD chart consistently maintains the desired IC FAP for all cases considered. The performance of Hotelling's T^2 chart worsens in higher dimensions, reaching empirical IC FAPs near 100% for some scenarios when $p = 10$.

5.3 OC Chart Performance for Isolated Shifts

For all OC conditions considered, in the case of multivariate normality, the T^2 chart limits are computed according to (2); however, in the cases of non-normality, the T^2 limits are computed via simulation to achieve the desired IC FAP for each (m,n) combination and process distribution studied. The use of simulated limits for the T^2 chart requires knowledge of the process distribution and is impractical outside of a simulation environment; however, it is necessary here in order to accurately compare the OC chart performance since the IC performance of the T^2 chart differs according to the process distribution and number of subgroups.

MMR-RMD and Hotelling's T^2 chart performance comparisons for isolated shifts in $p = 5$ dimensions for (m,n) combinations (50,5), (100,5), and (200,5) are given in Figure 3. Under multivariate normality, Hotelling's T^2 correctly identifies isolated shifts of a given size with a higher probability than the MMR-RMD. It is common for parametric methods to perform better than their nonparametric counterparts when the distributional assumptions hold.

For heavy-tailed process data following a multivariate $t(3)$ distribution, the MMR-RMD chart is both significantly better and much more consistent than Hotelling's T^2 chart at correctly classifying OC observations. The performance of Hotelling's T^2 chart declines dramatically as m

is increased to 200, whereas MMR chart performance is far less affected when the number of subgroups is increased. For example, in the case of multivariate $t(3)$ data with an isolated shift of magnitude $\lambda = 3.5$, for $m = 50$, 100, and 200, the MMR-RMD chart correctly classifies the OC observations in more than 97% of the cases, compared to 95%, 60%, and 3%, respectively, for Hotelling's T^2 chart.

Figure 4 shows the performance of the MMR-RMD and T^2 charts when the process observations originate from a heavy-tailed [$t(3)$], 10-dimensional multivariate distribution. The MMR chart maintains its ability to detect isolated shifts in higher dimensions and has better performance than Hotelling's T^2 chart. Although it is difficult for both charts to detect smaller shifts in this case, the MMR-RMD shows a performance edge over the T^2 chart for $\lambda > 2$.

5.4 OC Chart Performance for Sustained Shifts

In Figure 5, the percent of OC conditions correctly classified as OC is given for the MMR-RMD and MMR-SD charts under several sustained shifts using data generated from a bivariate $t(3)$ distribution. It is clear from Figure 5 that the more contaminated the reference sample, the more difficult it is to for the charts to retrospectively detect the OC conditions. In the case of 30% sustained shifts, neither the MMR-RMD nor the MMR-SD chart performs very well in terms of correctly identifying the OC condition. Interestingly, the MMR chart using robust Mahalanobis depth correctly classifies a higher percentage of sustained location shifts of a given size and length than the MMR chart based on simplicial depth. For example, when $\lambda =$

3.0, the MMR-RMD chart correctly classifies 94.5% of the 5% sustained shift conditions. The MMR-SD chart correctly classifies the same size and length sustained shift in only 55.6% of the simulated cases. This is consistent with Stoumbos and Jones (2000), who pointed out that control charts for individuals based on simplicial depth exhibit difficulty in distinguishing extreme points in a retrospective analysis. Because of its computational simplicity and performance benefits, we recommend the use of the MMR chart based on RMD when the process distribution is elliptically symmetric. In Section 6, we illustrate a method to evaluate elliptical symmetry of a data set. Phase I charts for non-elliptically symmetric data is a topic of current research by the authors.

Figure 6 compares the OC performance of the MMR-RMD and T^2 charts for 5% and 15% sustained shifts in the mean. In the case of a 5% sustained shift, the MMR-RMD chart's performance surpasses that of the T^2 chart by an increasing margin as m is increased from 50 to 200. In the 15% sustained shift scenario, the MMR-RMD chart outperforms the T^2 chart in correctly identifying small shifts, but the T^2 chart gains a slight edge over the MMR-RMD chart for larger shifts.

To further illustrate the chart performance for highly contaminated samples, Figure 7 compares the MMR-RMD chart and T^2 charts with 15%, 20%, and 25% sustained shifts. The MMR-RMD chart outperforms the T^2 chart in correctly identifying a 15% sustained shift of size $\lambda < 3.5$; however, the performance of the MMR-RMD deteriorates as the contamination level in the samples increases to 20% and 25%. Although this suggests that the T^2 chart has some

advantages over the MMR-RMD for detecting large sustained shifts in highly contaminated samples, the results should be interpreted with caution. For the simulation study, the limits of the T^2 chart were artificially widened to account for the underlying process distribution, which is known in a simulated environment. In practice, the underlying process distribution is usually unknown, and it would not be possible to properly specify the control limits of the T^2 chart to attain the desired IC chart performance. Thus, as discussed in Section 5.2, it would be difficult for a T^2 chart to distinguish IC from OC situations when the underlying process is non-normal.

The deterioration in the OC performance of the MMR-RMD under sustained shifts of increasing duration occurs because the ranks are affected by the contamination levels in the sample. The more the sample is contaminated with OC observations, the more the pooled ranks will be affected, and the more it difficult it will be for rank-based charts to distinguish IC and OC points. More details regarding the affect of contamination levels on the pooled ranks is provided in Section S.2 of the supplementary information.

6. AN EXAMPLE PHASE I ANALYSIS

We illustrate the application of the MMR chart to a dataset from a white wine production process. This dataset can be downloaded from *http://archive.ics.uci.edu/ml/datasets/Wine+Quality* and was also examined by Zou et al. (2011). The data were collected between May 2004 and February 2007, and contain 4,898 observations on eleven continuous physiochemical measurements that are possible indicators of wine quality. The dataset also includes one ordinal variable based on sensory analysis indicating wine quality ratings ranging from 0 (very bad) to 10 (excellent). According to Cortez et al. (2009), the goal

of the analysis is to model the ordinal wine quality ratings using the physiochemical measures. Zou et al. (2011) considered this data in a statistical process control context of sequential monitoring of the wine production process, and used only the data for the wines falling into the ordinal rating category "seven" (LV7), which represented a standard level of quality. They used a LASSO method to detect the variables responsible for a quality shift from the quality standard (LV7) to subsequent observations of a known lesser quality (LV6). They found that a significant change occurred in three variables (chlorides, density, and alcohol) before and after the change-point between the two known quality levels.

Following the recommendation of Cortez et al. (2009) and Zou et al. (2011), we will assume that the quality rating of LV7 is the standard and investigate only these 880 observations as a potential reference sample to establish a monitoring scheme for the three variables of interest (chlorides, density, and alcohol). An investigation of the appropriateness of the assumption of elliptical symmetry of this data is provided in Section S.3 of the supplementary material.

For the purposes of the Phase I analysis we will group the observations into $m = 176$ subgroups of size $n = 5$. Table 2 gives the RMD values and ranks for each observation in the first four subgroups as well as the first four chart statistics.

The control chart and the control limit are shown in Figure 8. The UCL was computed using the empirical procedure described in Section 4.2 by modifying the number of subgroups and the subgroup size in the MATLAB code included as supplemental material to this article. Subgroups 75, 86, 151, and 155 had MMR-RMD values above the UCL. These subgroups should be investigated further to determine if assignable causes to a true OC condition can be

associated with them. If so, the OC observations should be eliminated, and a new UCL should be determined based on the remaining subgroups. Comparisons of the remaining subgroups to the new limits may reveal other OC signals; thus, the Phase I analysis should continue in an iterative manner until all subgroups are deemed IC.

7. CONCLUDING REMARKS

The MMR chart has been shown to perform comparably to the T^2 chart under the assumption of multivariate normality, and to detect isolated shifts more quickly than the T^2 chart in most cases of heavy-tailed observations in Phase I. When detecting sustained shifts of less than 15% of the observations, the MMR chart also outperformed the T^2 chart in most cases of heavy-tailed observations. However, the MMR chart did not detect sustained shifts occurring in more than 15% of the sample as quickly as the T^2 chart. The reader is cautioned that the IC performance of the T^2 chart is unpredictable unless the underlying process follows a multivariate normal distribution. The MMR chart offers distribution-free performance when the process is IC, because the UCL for a given application depends only on the number of subgroups, the size of each subgroup, and the desired IC FAP without regard to the form of the underlying process distribution. On the other hand, when the multivariate normality assumption is violated, the limits for the T^2 chart must be simulated according to the exact process distribution when in order to attain the desired FAP. This may be impossible in the early phase of process control when little is known about the process distribution. For a given OC scenario involving non-normally distributed data, the MMR chart performance is far less dependent on the number of

subgroups, *m*, than Hotelling's T^2 chart with empirical UCL, making it even more attractive as a distribution-free alternative.

Once an IC reference sample has been determined through a successful Phase I analysis using the MMR chart, the resultant reference sample can be used in conjunction with an appropriate Phase II method to monitor future observations for any departures from the IC state. Since more is known about a process at the conclusion of a Phase I analysis, the form of a Phase II control chart does not necessarily have to match the form of a Phase I control chart. Thus, the Phase II charts do not have to be limited to rank-based charts, or even charts designed for subgrouped data. It may be that a practitioner chooses a different sampling plan in Phase II and uses a robust chart designed for monitoring individual observations.

Much work remains to be done in the field of distribution-free Phase I multivariate quality control. With regard to the MMR chart using Mahalanobis depth-based measures, the use of certain estimators may affect the invariance of the method to certain covariance structures. Future studies addressing this method under various estimated parameters and conditions is warranted, especially studies that compare the performance of the computationally efficient RMD method to more computationally expensive depth measures that may not rely on estimated process parameters.

Although it is believed that the fundamental structure of the MMR chart is sound, potential refinements include further exploration of the use of robust estimators and experimentation with alternative data depth functions which may enhance MMR chart performance. Additionally, since the MMR chart is designed to detect location changes in subgrouped multivariate data from elliptical distributions during Phase I, an equivalent

distribution-free chart for detecting scale changes is needed. Finally, Phase I distribution-free charts for detecting both location and scale changes in small subgroups ($n < 5$) and individual multivariate observations ($n = 1$) should be considered.

SUPPLEMENTARY MATERIALS

The MATLAB code used to generate the empirically derived control limits in Section 4.2 and the MATLAB code used to evaluate the elliptical symmetry of the sample data are included in a zip file as supplementary materials to this paper. In addition, we provide a pdf file with supplementary material to the text of the article. Section S.1 contains information regarding asymptotic control limits for the MMR chart. Section S.2 contains a discussion of the effect of sample contamination on the ranks. Section S.3 gives an evaluation of elliptical symmetry of the sample data.

ACKNOWLEDGEMENTS

The authors wish to thank the editor, associate editor and anonymous referees for their guidance and helpful suggestions that greatly improved the presentation of this research. We also thank Sidi Niu, Northeastern University, for allowing us to use his MATLAB code to evaluate sample elliptical symmetry and to include this code as supplemental material for the readers.

REFERENCES

Alfaro, J.L., and Ortega, J.F. (2008). A Robust Alternative to Hotelling's T^2 Control Chart Using Trimmed Estimators. *Quality and Reliability Engineering International*, 24, 601-611.

Alt, F.B. (1976). Small Sample Probability Limits for the Mean of a Multivariate Normal Process. *ASQC Technical Conference Transactions*, 170-176.

Beran, R. (1979). Testing for Ellipsoidal Symmetry for a Multivariate Density. *The Annals of Statistics*, 7, 150-162.

Beltran, L.A. (2006). *Nonparametric Multivariate Statistical Process Control Using Principal Component Analysis and Simplicial Depth*. Dissertation, University of Central Florida.

Billor, N., Hadi, A.S., and Velleman, P.F. (2000). BACON: Blocked Adaptive Computationally Efficient Outlier Nominators. *Computational Statistics & Data Analysis*, 34, 279-298.

Chen, Y., Birch, J.B., and Woodall, W.H. (2013). Cluster-Based Profile Monitoring in Phase I Analysis. Submitted for publication.

Cortez, P., Cerdeira, A., Almeida, F., Matos, T., and Reis, J. (2009). Modeling Wine Preferences by Data Mining from Physicochemical Properties. *Decision Support Systems*, 47(4), 547-553.

Chou, Y.M., Mason, R.L., and Young, J.C. (2001). The Control Chart for Individual Observations from a Multivariate Non-Normal Distribution. *Communications in Statistics – Theory and Methods*, 30(8), 1937-1949.

Fraker, S.E., Woodall, W.H., and Mousavi, S. (2008). Performance Metrics for Surveillance Schemes. *Quality Engineering*, 20, 451-464.

Genest, M., Masse, J.-C., and Plante, J.-F. (2013). depth: Depth Functions Tools for Multivariate Analysis, 2.0-0. http://cran.r-project.org/web/packages/depth.

Hamurkaroglu, C., Mert, M., and Saykan, Y. (2004). Nonparametric Control Charts Based on Mahalanobis Depth. *Hacettepe Journal of Mathematics and Statistics*, 33, 57-67.

Hayter, A.J., and Tsui, K. (1994). Identification and Quantification in Multivariate Quality Control Problems. *Journal of Quality Technology*, 26, 197-208.

Hotelling, H. (1947). Multivariate Quality Control – Illustrated By the Air Testing of Sample Bombsights. *Techniques of Statistical Analysis*, C. Eisenhart, M.W. Hastay, and W.A. Wallis (eds.), New York: McGraw-Hill, pp. 111-184.

Huffer, F.W., and Park, C. (2007). A Test for Elliptical Symmetry. *Journal of Multivariate Analysis*, 98, 256–281.

Jensen, D.R. (1985). Multivariate Distributions. *Encyclopedia of Statistical Sciences*, 6, S. Kotz, N.L. Johnson, and C.B. Read (eds.), New York: John Wiley & Sons, pp. 43-55.

Jensen, W.A., Birch, J.B., and Woodall, W.H. (2007). High Breakdown Estimation Methods for Phase I Multivariate Control Charts. *Quality and Reliability Engineering International*, 23(5), 615-629.

Jobe, J.M., and Pokojovy, M. (2009). A Multistep, Cluster-Based Multivariate Chart for Retrospective Monitoring of Individuals. *Journal of Quality Technology*, 41(4), 323-339.

Jones-Farmer, L.A., and Champ, C. W. (2010). A Distribution-Free Phase I Control Chart for Subgroup Scale. *Journal of Quality Technology*, 42(4), 373-387.

Jones-Farmer, L.A., Jordan, V., and Champ, C.W. (2009). Distribution-Free Phase I Control Charts for Subgroup Location. *Journal of Quality Technology*, 41(3), 304-316.

Li, R., Fang, K-T. and Zhu, L-X. (1997). Some Q-Q Plots to Test Spherical and Elliptical Symmetry. *Journal of Computational and Graphical Statistics*, 6, 435-450.

Liu, R.Y. (1990). On a Notion of Data Depth Based on Random Simplices. *The Annals of Statistics*, 18, 405-414.

Liu, R.Y. (1995). Control Charts for Multivariate Processes. *Journal of the American Statistical Association*, 90, 1380-1388.

Liu, R.Y., Parelius, J.M., and Singh, K. (1999). Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference. *The Annals of Statistics*, 27, 783-840.

Top of Form

Liu, R.Y., Singh, K., and Teng, J.H. (2004). DDMA-Charts: Nonparametric Multivariate Moving Average Control Charts Based on Data Depth. *Allgemeines Statistisches*, 88, 235-258.

Mahalanobis, P. C. (1936). On the Generalized Distance in Statistics. *Proceedings of the National Institute of Science of India*, 12, 49-55.

Messaoud, A., Weihs, C., and Hering, F. (2008). Detection of Chatter Vibration in a Drilling Process Using Multivariate Control Charts. *Computational Statistics & Data Analysis*, 52(6), 3208-3219.

Montgomery, D.C. (2009). *Introduction to Statistical Quality Control* (6th ed.). Hoboken, NJ: John Wiley & Sons.

Mosler, K. (2002). Multivariate Dispersion, Central Regions and Depth: The Lift Zonoid Approach. *Lecture Notes in Statistics*, 165, New York: Springer.

Nedumaran, G., and Pignatiello, J.J., Jr. (2000). On Constructing T^2 Control Charts for Retrospective Examination. *Communications in Statistics – Simulation and Computation*, 29(2), 621-632.

Qiu, P. (2008). Distribution-Free Multivariate Process Control Based on Log-Linear Modeling. *IIE Transactions*, 40(7), 664-691.

Qiu, P., and Hawkins, D.M. (2001). A Rank-Based Multivariate CUSUM Procedure. *Technometrics*, 43(2), 120-132.

Qiu, P., and Hawkins, D.M. (2003). A Nonparametric Multivariate Cumulative Sum Procedure for Detecting Shifts In All Directions. *The Statistician*, 52(2), 151-164.

Rousseeuw, P.J., and Ruts, I. (1996). Algorithm AS 307: Bivariate Location Depth. *Applied Statistics*, 45, 516-526.

Stoumbos, Z.G., and Jones, L. A. (2000). On the Properties and Design of Individuals Control Charts Based on Simplicial Depth. *Nonlinear Studies*, 7(2), 147-178.

Stoumbos, Z.G., Jones, L.A., Woodall, W.H., and Reynolds, M.R., Jr. (2000). On Nonparametric Multivariate Control Charts Based on Data Depth. *Frontiers in Statistical Quality Control*, 6, H.-J. Lenz, and P.-T. Wilrich (eds.), Heidelberg, Germany: Springer-Verlag, pp. 208-228.

Stoumbos, Z.G., and Sullivan, J.H. (2002). Robustness to Non-Normality of the Multivariate EWMA Control Chart. *Journal of Quality Technology*, 34(3), 260-276.

Sullivan, J.H., and Woodall, W.H. (1996). A Comparison of Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 28(4), 398-408.

Sun, R., and Tsung, F. (2003). A Kernel-Distance-Based Multivariate Control Chart Using Support Vector Methods. *International Journal of Production Research*, 41(13), 2975-2989.

Thissen, U., Swierenga, H., de Weijer, A.P., Wehrens, R., Melssen, W.J., and Buydens, L.M.C. (2005). Multivariate Statistical Process Control Using Mixture Modelling [*sic*]. *Journal of Chemometrics*, 19, 23-31.

Tracy, N.D., Young, J.C., and Mason, R.L. (1992). Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 24, 88-95.

Tukey, J.W. (1975). Mathematics and Picturing Data. In R. James (ed.), *Proceedings of the 1974 International Congress of Mathematicians* (Vol. 2, pp. 523-531). Vancouver, BC.

Vargas, J.A. (2003). Robust Estimation in Multivariate Control Charts for Individual Observations. *Journal of Quality Technology*, 35(4), 367-376.

Woodall, W.H., and Montgomery, D.C. (1999). Research Issues and Ideas in Statistical Process Control. *Journal of Quality Technology*, 31(4), 376-386.

Woodall, W.H., Spitzner, D.J., Montgomery, D.C., and Gupta, S. (2004). Using Control Charts to Monitor Process and Product Quality Profiles. *Journal of Quality Technology*, 36(3), 309-320.

Zarate, P.B. (2004). *Design of Nonparametric Control Chart for Monitoring Multivariate Processes Using Principal Components Analysis and Data Depth*. Dissertation, University of South Florida.

Zou, C., Jiang, W., and Tsung, F. (2011). A LASSO-Based Diagnostic Framework for Multivariate Statistical Process Control. *Technometrics*, 53(3), 297-309.

Zuo, Y., and Serfling, R. (2000a). General Notions of Statistical Depth Functions. *The Annals of Statistics,* 28(2), 461-482.

Zuo, Y., and Serfling, R. (2000b). Structural Properties and Convergence Results for Contours of Sample Statistical Depth Functions. *The Annals of Statistics,* 28(2), 483-499.

Table 1. Empirical FAPs using Empirical control limits for the MMR chart. Monte Carlo results are based on 100,000 iterations. The MMR chart was computed using RMD as defined in (1). and limits were computed for a desired FAP value of 0.10.

Sample Size		Simulated Limits		Sample Size		Simulated Limits	
m	**n**	**UCL**	**Empirical FAP**	**m**	**n**	**UCL**	**Empirical FAP**
20	5	2.476	0.0941	100	5	2.854	0.0982
20	10	2.519	0.0984	100	10	2.972	0.0974
20	15	2.54	0.0979	100	15	3.013	0.0982
20	20	2.544	0.0973	100	20	3.03	0.0982
50	5	2.702	0.0983	200	5	2.985	0.0981
50	10	2.787	0.0981	200	10	3.144	0.0985
50	15	2.818	0.0968	200	15	3.188	0.0984
50	20	2.829	0.0979	200	20	3.214	0.0984

Table 2. The RMD values, pooled ranks, and MMR-RMD statistics for the first four example subgroups in the first round of the analysis. Version 1 of the BACON estimator was used to compute RMD values for this example.

Subgroup	RMD	Rank	MMR-RMD Statistic	Subgroup	RMD	Rank	MMR-RMD Statistic
1	0.240	500	-0.81	3	0.377	248	-1.83
	0.506	108			0.383	240	
	0.341	302			0.734	19	
	0.536	95			0.562	77	
	0.123	740			0.208	581.5	
2	0.057	876	1.07	4	0.519	102	-1.72
	0.106	791			0.208	581.5	
	0.286	387.5			0.312	340	
	0.201	592			0.682	27	
	0.440	164			0.431	175	

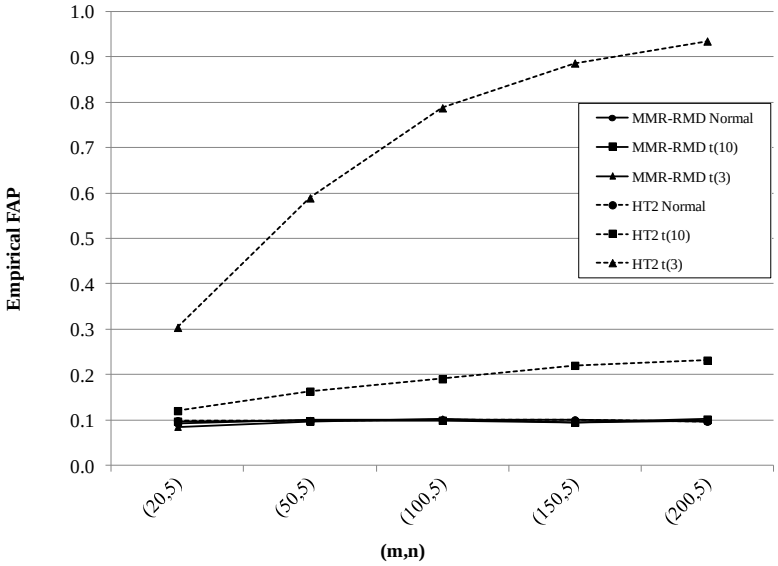


Figure 1. Empirical IC FAPs for MMR-RMD and Hotelling's T^2 charts based on the bivariate distributions indicated in the graph legend. Results are based on *m* subgroups of size *n*.

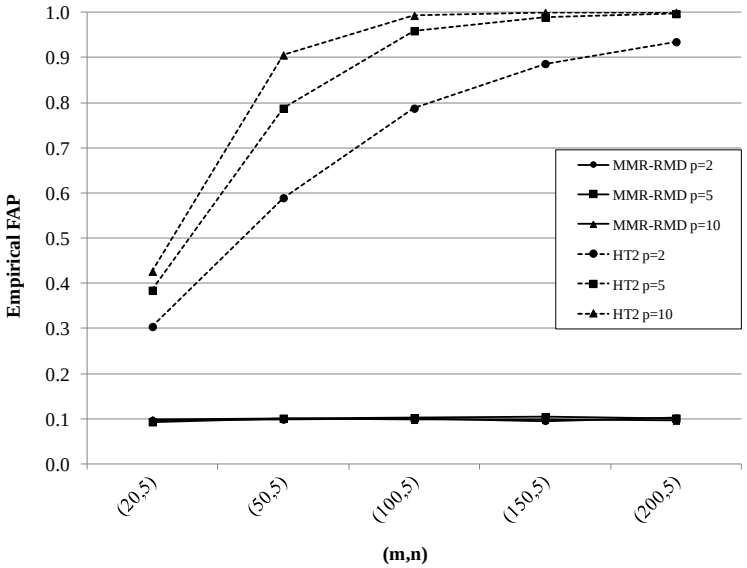


Figure 2. Empirical IC FAPs for MMR-RMD and Hotelling's T^2 charts based on the *p*-variate *t*-distribution with 3 degrees of freedom. Results are based on *m* subgroups of size *n*.

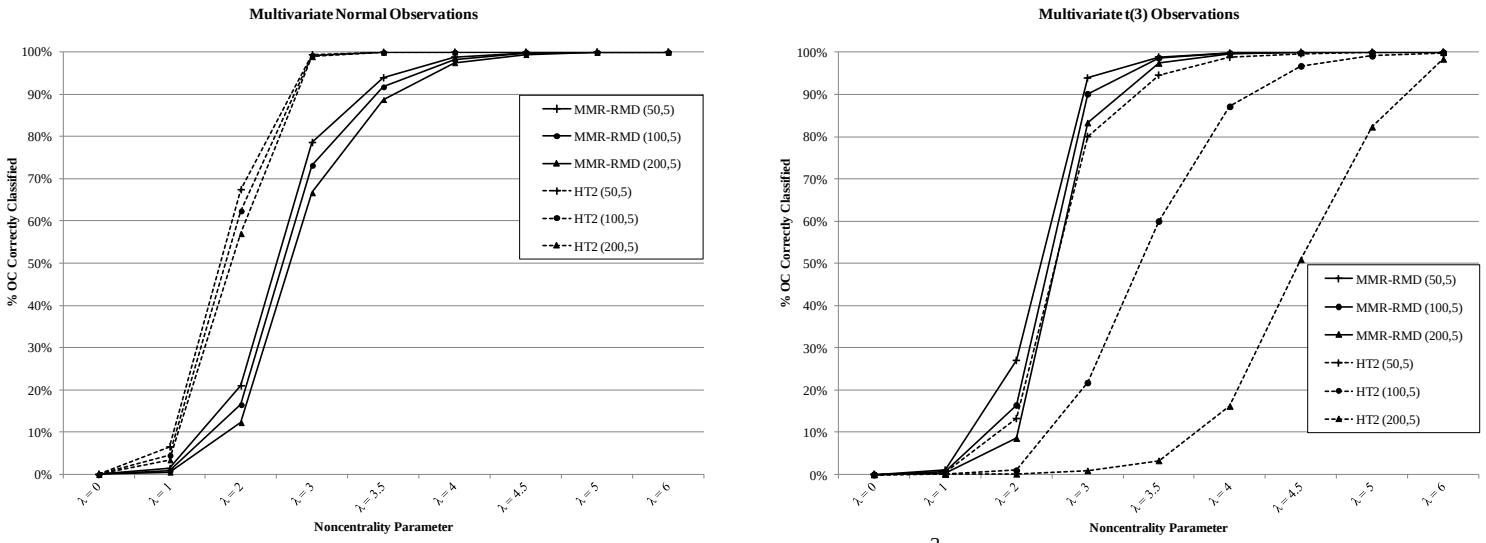


Figure 3. Performance of the MMR-RMD and Hotelling's T^2 charts when $p = 5$ under isolated shifts of size λ.

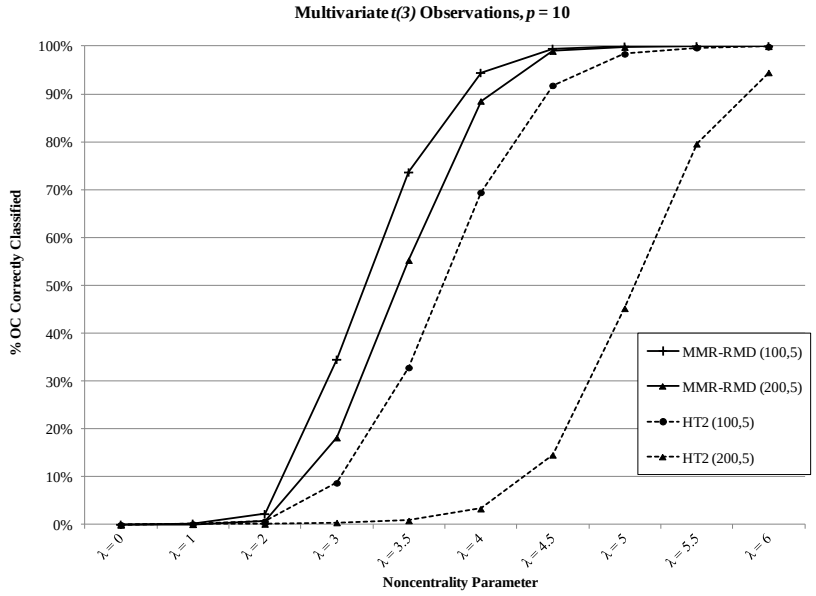


Figure 4. Performance of the MMR-RMD and T^2 charts when data originate from a multivariate $t(3)$ distribution under isolated shifts of size λ.

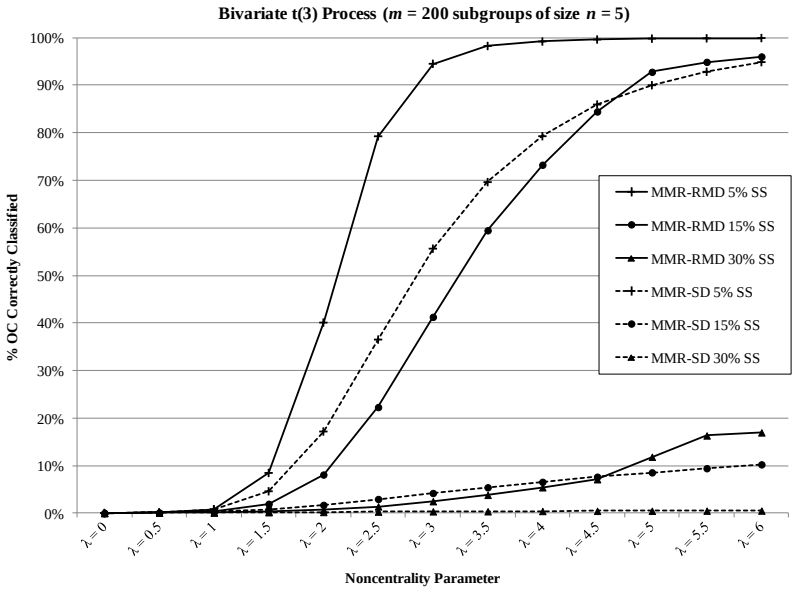


Figure 5. Performance comparison of MMR-RMD with MMR-SD when $p = 2$ under sustained shifts of various sizes.

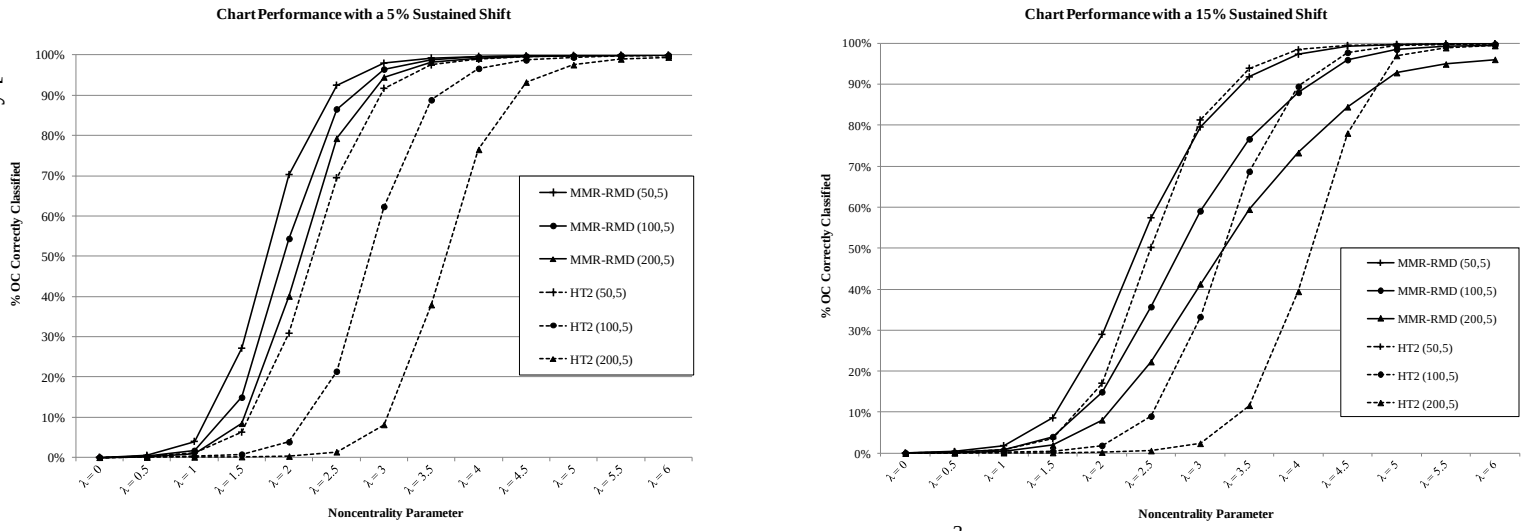


Figure 6. Performance of the MMR-RMD and Hotelling's T^2 charts when $p = 2$ under sustained shifts of size λ induced in the last 5% or 15% of the subgroups.

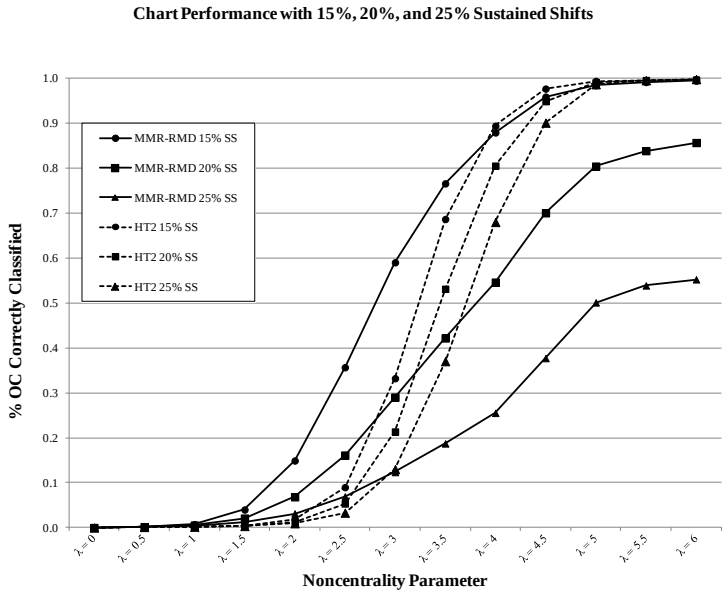


Figure 7. Performance of the MMR-RMD and Hotelling's T^2 charts when $p = 2$, $m = 100$, $n = 5$ under sustained shifts of size λ induced in the last 15%, 20%, or 25% of the subgroups.

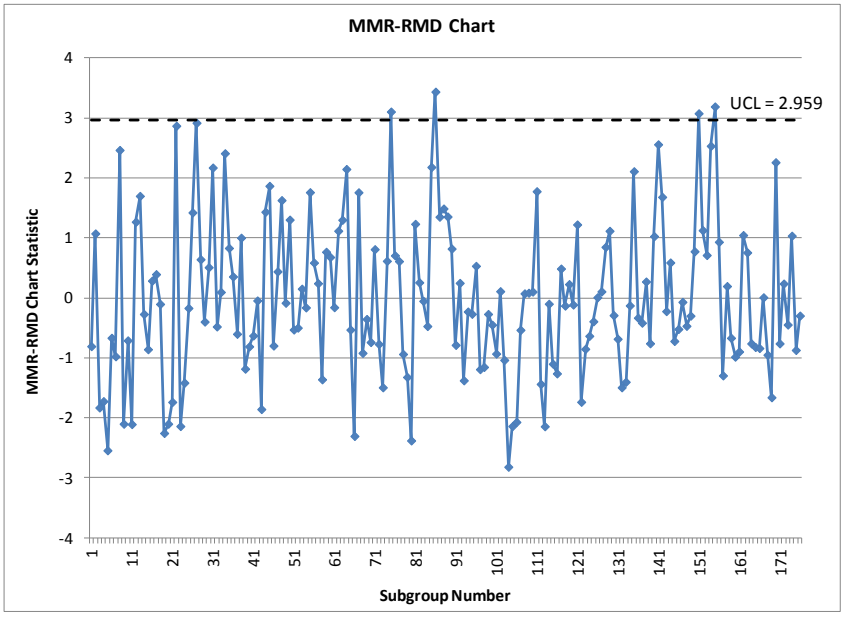


Figure 8. MMR-RMD control chart for three variables (chlorides, density, and alcohol) related to white wine quality.