



香港城市大學
City University of Hong Kong

專業 創新 胸懷全球
Professional · Creative
For The World

CityU Scholars

A New Robust Multivariate EWMA Dispersion Control Chart for Individual Observations

Ajadi, Jimoh Olawale; Zwetsloot, Inez Maria; Tsui, Kwok-Leung

Published in:
Mathematics

Published: 01/05/2021

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY

Publication record in CityU Scholars:
[Go to record](#)

Published version (DOI):
[10.3390/math9091038](https://doi.org/10.3390/math9091038)

Publication details:
Ajadi, J. O., Zwetsloot, I. M., & Tsui, K-L. (2021). A New Robust Multivariate EWMA Dispersion Control Chart for Individual Observations. *Mathematics*, 9(9), Article 1038. Advance online publication.
<https://doi.org/10.3390/math9091038>

Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy

Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Article

A New Robust Multivariate EWMA Dispersion Control Chart for Individual Observations

Jimoh Olawale Ajadi ¹, Inez Maria Zwetsloot ^{2,*} and Kwok-Leung Tsui ³

¹ Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong; ajadi_jimoh2001@yahoo.com

² Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong

³ Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061-0002, USA; klttsui@vt.edu

* Correspondence: i.m.zwetsloot@cityu.edu.hk

Abstract: A multivariate control chart is proposed to detect changes in the process dispersion of multiple correlated quality characteristics. We focus on individual observations, where we monitor the data vector-by-vector rather than in (rational) subgroups. The proposed control chart is developed by applying the logarithm to the diagonal elements of the estimated covariance matrix. Then, this vector is incorporated in an exponentially weighted moving average (EWMA) statistic. This design makes the chart robust to non-normality in the underlying data. We compare the performance of the proposed control chart with popular alternatives. The simulation studies show that the proposed control chart outperforms the existing procedures when there is an overall decrease in the covariance matrix. In addition, the proposed chart is the most robust to changes in the data distribution, where we focus on small deviations which are difficult to detect. Finally, the compared control charts are applied to two case studies.

Keywords: individual observations; covariance matrix; non-normality; multivariate dispersion chart; EWMA



Citation: Ajadi, J.O.; Zwetsloot, I.M.; Tsui, K.-L. A New Robust Multivariate EWMA Dispersion Control Chart for Individual Observations. *Mathematics* **2021**, *9*, 1038. <https://doi.org/10.3390/math9091038>

Academic Editor: Timothy M. Young

Received: 26 March 2021

Accepted: 28 April 2021

Published: 3 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The quality of a process can often be characterized by multiple correlated variables. For example, when monitoring a patient's systolic and diastolic blood pressure, it is more appropriate to apply one multivariate control chart than two univariate charts. Hotelling [1] introduced the first multivariate control chart. Many researchers have improved on this chart. For example, Lowry et al. [2] proposed a multivariate exponentially weighted moving average (MEWMA), also Crosier [3] developed a multivariate cumulative sum (MCUSUM) chart. Some other researchers introduced multivariate control charts that monitor the variability of a process. For example, Alt [4] proposed a generalized variance chart (GVC). A GVC uses the determinant of the estimated covariance matrix of grouped observations as the monitoring statistics. For details on multivariate dispersion charts for grouped observations, see the review articles by Yeh et al. [5] and Bersimis et al. [6].

Recently, Ajadi and Zwetsloot [7] compared the performance of a multivariate EWMA chart proposed by Huwang et al. [8] for monitoring process variability of individual observations with various charts based on grouped observations. The authors concluded that monitoring methods based on individual observations are quickest in detecting sustained shifts in the process variability. Reynolds Jr and Stoumbos [9] and Reynolds Jr and Stoumbos [10] also advised monitoring with individual observations in the univariate setup.

Ajadi et al. [11] gave an overview of the existing multivariate parametric and nonparametric control charts for monitoring process dispersion of individual observations. Next we briefly introduce some noteworthy examples. Yeh et al. [12] proposed a MaxMEWMV chart.

This chart is developed by computing the difference between the estimated covariance matrix derived from the EWMA statistic and the identity matrix. Then, the L_2 -norms of both the diagonal elements and the upper triangular off-diagonal elements are computed. Finally, a single statistic is derived from the two distance-measures to detect changes in the process variability. Huwang et al. [8] proposed the multivariate exponentially weighted mean squared deviation (MEWMS) and the multivariate exponentially weighted moving variance (MEWMV) charts; these charts are based on the trace of the estimated covariance matrix obtained from the EWMA statistic. According to the authors, the MEWMS chart is designed to monitor only changes in the process variability while the MEWMV chart is designed to simultaneously monitors shifts in both the process mean vector and the covariance matrix. Hawkins and Maboudou-Tchao [13] proposed the multivariate exponentially weighted moving covariance matrix (MEWMC) chart. The authors applied the Alt's likelihood ratio statistic to compare the estimated EWMA covariance matrix with the identity matrix. The multivariate control charts discussed above are parametric charts. Li et al. [14] employed a multivariate spatial sign test and EWMA statistic to develop a nonparametric control chart for monitoring shape parameters of the underlying data. The authors named the chart as the multivariate nonparametric shape EWMA (MNSE).

Except for the nonparametric control charts, the other multivariate charts for monitoring the process variability of individual observations are based on the assumption that process data are normally distributed. Stoumbos and Sullivan [15] investigated the robustness of the MEWMA control chart for monitoring process mean vector. The authors showed that the MEWMA control chart is very robust to non-normality even for a highly skewed and heavy-tailed multivariate distributions. Testik et al. [16] extended the work of Borror et al. [17] to examine the robustness properties of multivariate EWMA control charts to non-normal data. Zwetsloot and Ajadi [18] showed that the performance of the univariate EWMA control chart based on logarithm of sample variance is the most robust to normality assumption among three compared univariate EWMA dispersion charts. The authors recommended the chart to be used in practice since it can be difficult to distinguish between normally distributed data, and data that slightly deviates from normality. We have not seen any research that investigate the robustness of the multivariate dispersion charts for individual observations.

Ajadi et al. [11] identified some research directions. One of these is that the robustness of multivariate dispersion charts for individual observations should be investigated. This is one of the motivations for the current study. The objective of this article is to introduce a monitoring method for detecting changes in the process dispersion using individual observations. This method should be robust to slight deviation from the normality assumption. The chart is designed using the logarithm of diagonal elements of the sample covariance matrix estimated using the EWMA statistic. This makes the chart robust to non-normality. In Section 2, we briefly discuss the MaxMEWMV, the MEWMS and a nonparametric control chart. In Section 3, the proposed control chart is discussed in detail. Next, in Section 4, we compare the performance of the proposed chart with the other existing control charts discussed in Section 2. An example is used in Section 5 to support the comparison. Finally, we provide conclusions and recommendations in Section 6.

2. Multivariate EWMA Dispersion Control Charts for Individual Observations

Assume that the data follow a multivariate normal distribution, i.e., $X_i \sim N_p(\mu, \Sigma)$, where the process mean vector and covariance matrix of a p correlated quality characteristics are denoted by μ and Σ , respectively. Next, X_i is standardized by transforming it to Y_i as

$$Y_i = \Sigma_0^{-\frac{1}{2}}(X_i - \mu_0) \text{ for } i = 1, 2, 3, \dots \quad (1)$$

Thus, Y_i follows a multivariate normal distribution $N(\mu_Y, \Sigma_Y)$, where $\mu_Y = \Sigma_0^{-\frac{1}{2}}(\mu - \mu_0)$ and $\Sigma_Y = \Sigma_0^{-\frac{1}{2}}\Sigma\Sigma_0^{-\frac{1}{2}}$. Note that μ_0 and Σ_0 , are the in-control values of μ and Σ ,

respectively. When the process is in-control, $Y_i \sim N_p(\mathbf{0}, I_p)$, where I_p is a $p \times p$ identity matrix. To compute Y_i , we need estimates of Σ_0 and μ_0 . These are usually obtained in a Phase I analysis (see Section 3.2 for details).

Next, we discuss two popular parametric (MaxMEWMV, MEWMS) and one non-parametric (MNSE) control charts for monitoring the covariance matrix of individual observations.

2.1. MaxMEWMV Chart

The first multivariate EWMA control chart for monitoring the process variability of individual observations was proposed by Yeh et al. [12]. This chart is referred to as the MaxMEWMV. As Yeh et al. [12] suggested we first transform the data as

$$S_i^E = \lambda Y_i Y_i' + (1 - \lambda) S_{i-1}^E, \quad (2)$$

where $0 < \lambda < 1$ is a smoothing constant and $S_0^E = Y_1 Y_1'$. The monitoring statistic for the MaxMEWMV chart is defined as

$$MaxD_i = \max \left(\frac{D_{1i} - E(D_{1i})}{\sqrt{Var(D_{1i})}}, \frac{D_{2i} - E(D_{2i})}{\sqrt{Var(D_{2i})}} \right)$$

where $D_i = S_i^E - I_p$. Note that D_{1i} is the L_2 -norm of the diagonal elements of D_i and D_{2i} is the L_2 -norm of the upper triangular off-diagonal elements of D_i . Where D_{1i} is used to detect shifts in the process variance and D_{2i} to detect changes in the correlation structure. The expectation and variance of D_{1i} and D_{2i} are derived analytically by Yeh et al. [12]. A signal is given when $MaxD_i$ exceeds an upper control limit (UCL).

2.2. MEWMS Chart

Later in 2007, Huwang et al. [8] proposed two multivariate EWMA dispersion charts. One of the charts (MEWMS) is designed to detect only changes in the process variability. The MEWMS control chart applies the trace of the estimated covariance matrix obtained from the EWMA statistic (S_i^E) as the monitoring statistic.

$$tr(S_i^E) = \sum_{k=1}^p s_{i,kk}^2$$

where $s_{i,kk}^2$ is the k -th diagonal element of S_i^E . The monitoring statistic, $tr(S_i^E)$, is compared with the upper and lower control limits (UCL and LCL);

$$LCL/UCL = p \pm L\sqrt{2pc_i}$$

where $c_i = \frac{\lambda}{2-\lambda} + \frac{2-2\lambda}{2-\lambda}(1-\lambda)^{2(i-1)}$ and L is the control charting constant. The chart signals when $tr(S_i^E) > UCL$ or $tr(S_i^E) < LCL$.

2.3. MNSE Chart

Li et al. [14] incorporated a multivariate spatial sign test with an exponentially weighted moving average (EWMA) scheme to develop a distribution-free control chart for monitoring shape parameters. The authors named the chart the multivariate nonparametric shape EWMA (MNSE) chart. They showed that the MNSE chart is robust to non-normal distributions of the data, in the sense that the in-control performance is not affected. The mean vector (θ_0) and covariance matrix (A_0) estimators of the MNSE chart can be obtained simultaneously through Equations (3) and (4).

$$\frac{1}{m} \sum_{j=1}^m U(A_0(X_j - \theta_0)) = 0 \quad (3)$$

$$\frac{1}{m} \sum_{j=1}^m U\left(A_0(X_j - \theta_0)\right) U' \left(A_0(X_j - \theta_0)\right) = \frac{I_p}{p} \tag{4}$$

where m is the size of the historical dataset. The algorithm to compute $\hat{\theta}_0$ and \hat{A}_0 estimates from Equations (3) and (4) is available in the supplementary file by Zou and Tsung [19]. Next, Phase II observations, i.e., $X_i, i = 1, 2, 3, \dots$, can be standardized and transformed by multivariate spatial sign, i.e., $Z_i = U\left(A_0(X_i - \theta_0)\right)$, where $U(x) = \frac{x}{\|x\|_2}$ is the spatial sign vector. Then, the EWMA statistic can be computed as

$$S_i = \lambda Z_i Z_i' + (1 - \lambda) S_{i-1},$$

where $S_0 = I_p/p$. Finally, the monitoring statistic, Q_i in Equation (5), is compared with an upper control limit (UCL).

$$Q_i = \sqrt{\frac{2 - \lambda}{\lambda} \text{tr}(p S_i - I_p)^2} \tag{5}$$

3. The Proposed Robust EWMA Control Chart

In this section, we introduce our new multivariate dispersion chart for individual observations. The proposed chart is based on the logarithm of the diagonal elements of Y_i and referred to as REWMV, where R is short for robust. We focus on robustness to deviations from the normality assumption that are difficult to identify. As we believe that this is a common situation in practice, data may look approximately normally distributed however it is impossible to verify if they could have slightly heavy tails or if the distribution is a little skewed.

3.1. Guidelines to Implementation of the REWMV Chart

The REWMV control chart incorporates the logarithm of the estimated covariance matrix in the multivariate EWMA statistic. There are three reasons why the logarithm makes the chart more robust. Firstly, as mentioned by Box et al. [20] (Chapter 5.4), the logarithmic transformation of the sample variance is approximately more normally distributed than the sample variance itself. As the logarithm transforms skewed data to approximately normally distributed data. Secondly, Crowder and Hamilton [21] showed that the logarithmic transformation changes the model from a variance shift to a location shift model. This is beneficial as most charts are designed to detect mean shifts quickly. Finally, the variance of the logarithm of sample variance is independent of σ^2 and only depends on sample size.

This idea has been used in univariate monitoring tools quite extensively; Crowder and Hamilton [21] suggested the use of an EWMA scheme to the logarithm transformation of the sample variance for monitoring an increase in the process variability. Shu and Jiang [22] also mentioned that using the logarithm transformation makes the control limits for a two-sided EWMA control chart nearly symmetrical. Zwetsloot and Ajadi [18] compared the performance of three univariate EWMA dispersion charts based on estimated parameters under normal and non-normally distributed data. The authors showed that the EWMA control chart based on logarithm of the sample variance is more robust to non-normality than the other compared charts.

First, we integrate the logarithm of diagonal elements of the estimated covariance matrix into the multivariate EWMA statistic as

$$S_i^{logE} = \lambda \log\{diag(Y_i Y_i')\} + (1 - \lambda) S_{i-1}^{logE}, \tag{6}$$

where $0 < \lambda < 1$, is the smoothing parameter, $S_0^{logE} = E[S_i^{logE}]$ and $diag(A)$ is a vector of the diagonal elements of the matrix A .

Our objective for this study is to detect both increases and decreases in the process variability for multivariate individual observations. Next, we discuss the procedure for the upper and lower-sided monitoring with the proposed REWMV control chart.

The monitoring statistic is reset to a reflection boundary whenever the statistic is less than the boundary for the upper-sided monitoring; or greater than the boundary for the lower-sided monitoring. The reflection boundary $E[S_i^{logE}]$ is the expected value of S_i^{logE} , and it is obtained as

$$E[S_i^{logE}] = [-1.27 \quad -1.27 \quad \dots \quad -1.27]_{p \times 1},$$

under the assumption that $Y_i \sim N(0, I_p)$.

We apply the statistics in Equations (7) and (8) to detect respectively increases and decreases in the process variability. The *max* and *min* in these equations are respectively the element-wise maximum and minimum values of the two vectors.

$$S_i^{log+} = \sum_{k=1}^p \max \{ S_i^{logE}, -1.27 \}. \tag{7}$$

$$S_i^{log-} = \sum_{k=1}^p \min \{ S_i^{logE}, -1.27 \}. \tag{8}$$

There are many ways to develop summary statistics to monitor the change of the covariance parameters for multivariate observations. We tried various methods such as taking EWMA statistic before applying logarithm to the estimated covariance matrix; we also employed the mahalanobis distance (a standardized distance measure incorporating the variance of the data) of the estimated EWMA statistic as the summary statistic, among other methods. However, the present method works the best in detecting the considered ranges and types of shifts.

Finally, S_i^{log+} and S_i^{log-} are compared against the upper and lower control limits (*UCL* and *LCL*) to detect upward and downward shifts in the process dispersion respectively.

3.2. Phase I Estimation

To implement our and the other charts, we need estimates of $\hat{\mu}_0$ and $\hat{\Sigma}_0$ which we obtain in a Phase I analysis. Many covariance matrix estimators have been employed in the literature for estimating parameters in Phase I. The sample covariance matrix of the pooled observations is the most commonly used covariance matrix, and it is defined as

$$\hat{\Sigma}_0 = \frac{\sum_{j=1}^m (X_j - \hat{\mu}_0)(X_j - \hat{\mu}_0)'}{(m - 1)}, \tag{9}$$

where $\hat{\mu}_0 = \frac{\sum_{j=1}^m X_j}{m}$ and the observations, X_j , are observed at times $j = 1, 2, 3, \dots, m$. Note that m is Phase I sample size. For each observation of the MEWMS, the MaxMEWMV and the REWMV charts in Phase II, we standardize the data based on the Phase I estimates as

$$Y_i = \hat{\Sigma}_0^{-\frac{1}{2}}(X_i - \hat{\mu}_0). \tag{10}$$

Note that, the Phase I estimators for the MNSE chart are obtained through Equations (3) and (4).

4. Performance Comparison

In this section, we compare the performance of the proposed REWMV control chart with the existing competing charts (MaxMEWMV, MEWMS, and MNSE charts). We investigate the robustness of the selected charts to data that is slightly non-normal. We believe this is an important aspect of robustness. Because, as we will show in Section 4.3 it

can be very difficult to distinguish normally distributed data from slightly non-normally distributed data. Hence a monitoring method should perform consistently, however the existing monitoring methods will show many additional false alarms if this slight deviation occurs, as we will show in Section 4.5. But first we discuss the selected performance measures (Section 4.1) and the set-up of our simulation experiments (Section 4.2). Code for obtaining the simulation results is available on github.com/ajadi1982.

4.1. Performance Measures and Control Limits

The performance of the charts is defined in terms of the average of the conditional ARL values (AARL). The conditional ARL is defined as the average run length (RL) conditional on a specific Phase I estimate, i.e., $ARL[\widehat{\Sigma}_0, \widehat{\mu}_0] = E[RL|\widehat{\Sigma}_0, \widehat{\mu}_0]$. We evaluate the conditional ARL for each of the charts with 100 Monte Carlo simulation of the run lengths based on the Phase I parameters. Then, we calculate the mean of 10,000 conditional ARL simulations to obtain the AARL. We adjust the control limits for different Phase I sample sizes such that the $AARL_0 = 200$ for $p = 2$, $p = 5$ and $p = 25$. The adjusted control limits for each of the charts at specific values of p , λ and m are provided in Table 1. They are obtained under the assumption that the data are normally distributed.

Table 1. Control limits and control limit constants for the REWMV, MEWMS, the MaxMEWMV, and the MNSE control charts at $AARL_0 = 200$ under normally distributed data.

λ	p	m	REWMV		MEWMS	MaxMEWMV	MNSE
			LCL	UCL	L	UCL	UCL
0.1	2	50	-4.108	-1.220	2.550	5.200	3.034
		200	-4.230	-1.175	2.475	4.830	2.900
		2000	-4.255	-1.166	2.475	4.600	2.830
	5	50	-8.840	-3.542	3.170	7.900	7.210
		200	-9.173	-3.787	2.638	5.055	6.594
		2000	-9.260	-3.874	2.470	4.210	6.332
	25	50	-34.720	-8.820	20.200	84.270	35.500
		200	-38.970	-21.100	4.930	8.000	28.340
		2000	-40.500	-23.400	2.550	3.700	26.920
0.3	2	50	-6.450	0.060	3.338	7.170	2.700
		200	-6.540	0.076	3.380	6.907	2.660
		2000	-6.570	0.080	3.380	6.740	2.637
	5	50	-12.380	-1.235	3.908	9.240	6.675
		200	-12.700	-1.570	3.260	6.620	6.427
		2000	-12.800	-1.684	3.076	5.900	6.360
	25	50	-42.200	-0.700	17.120	59.00	31.120
		200	-48.000	-13.890	4.670	8.000	27.750
		2000	-49.680	-16.400	2.910	4.700	27.160

4.2. Description of Experiments

Our in-control model is $\mu_0 = (0\ 0\ \dots\ 0)_{1 \times p}$ and $\Sigma_0 = I_p$. For the out-of-control model, we set a general form Σ_1 and Σ_2 for overall and sparse shifts, respectively, as

$$\Sigma_1 = \delta \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-2} & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-3} & \rho^{p-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{p-4} & \rho^{p-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \rho^{p-2} & \rho^{p-3} & \rho^{p-4} & \dots & 1 & \rho \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & \rho & 1 \end{pmatrix}, \tag{11}$$

and

$$\Sigma_2 = \begin{pmatrix} \delta & \rho & \rho^2 & \dots & \rho^{p-2} & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-3} & \rho^{p-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{p-4} & \rho^{p-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \rho^{p-2} & \rho^{p-3} & \rho^{p-4} & \dots & 1 & \rho \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & \rho & 1 \end{pmatrix}. \quad (12)$$

We consider the following four scenarios for overall or sparse shifts in the process:

- Case 1: overall increase shifts in the variance-covariance matrix
We increase all the diagonal elements of the covariance matrix (i.e., $\delta > 1$) and/or the correlation coefficients as given in Equation (11).
- Case 2: Sparse increase shifts in the variance-covariance matrix
We consider increasing shifts in only the first element of the diagonal matrix where $\rho = 0$ and $\rho = 0.3$ as given in Equation (12).
- Case 3: overall decrease shifts in the variance-covariance matrix
We decrease all the diagonal elements of the covariance matrix (i.e., $\delta < 1$) where $\rho = 0$ and $\rho = 0.3$ as given in Equation (11).
- Case 4: Sparse decrease shifts in the variance-covariance matrix
We consider decreasing shifts in only the first element of the diagonal matrix where $\rho = 0$ and $\rho = 0.3$ as given in Equation (12).

In addition, we investigate the robustness of the proposed chart and its counterparts. We consider the multivariate t -distribution to model high kurtosis (kurtosis measures the thickness of the tails of a distribution), $X_i \sim t_p(\nu, \Sigma)$, where ν is the number of degrees of freedom. We use $\nu = 30$ degrees of freedom. The in-control process mean vector and covariance matrix of X_i are μ_0 and $\frac{\nu}{(\nu-2)}\Sigma_0$ respectively. Since the focus of this study is to monitor the changes in the process variability, the process mean vector remains constant in the out-of-control model while the covariance matrix changes to $\frac{\nu}{(\nu-2)}\Sigma_1$ or $\frac{\nu}{(\nu-2)}\Sigma_2$.

In addition, we employ the multivariate gamma distribution to model high skewness, $X_i \sim \Gamma_p(\alpha, \beta, \Sigma)$, where α and β are the shape and scale parameters of the distribution respectively. Throughout this article, we use $\beta = 1$. Note that, for gamma distributed data, the in-control covariance matrix, $\alpha\Sigma_0$, changes to $\alpha\Sigma_1$ (since $\beta = 1$) in the out-of-control model. We employed $\alpha = 16$. Details of the multivariate t and gamma distributions can be found in the appendix of Stoumbos and Sullivan [15].

4.3. How Non-Normal Are the Data?

In this section, we test for the normality of Phase I data simulated for different models using the Henze and Zirkler multivariate normality test. We run 10000 simulations of the test for different values of m ; and compute the percentage of the tests that do not reject normality. Table 2 shows the results. We require Phase I sample of $m \geq 500$, $m \geq 1000$ and $m \geq 5000$ for t distributed data with 10, 15 and 30 degrees of freedom respectively to have confidence that the data is non-normal. Moreover, we need $m > 30$, $m > 100$ and $m > 500$ for $\Gamma(1, 1)$, $\Gamma(4, 1)$ and $\Gamma(16, 1)$ respectively to be certain that the skewed data fail the multivariate normality test.

Hence, we see that for small and moderate size Phase I samples it is very challenging and sometimes impossible to distinguish between normal and slight non-normally distributed data. Hence there is a need for control charts that are robust to deviations from the normality assumption. We recommend the use of robust charts when the number Phase I data is less than 500, since testing cannot distinguish between data which is normally or non-normally distributed with high accuracy.

Table 2. Percentage of Henze and Zirkler tests for multivariate normality test that do not reject at a significance level of 0.05.

m	Normal	t₁₀	t₁₅	t₃₀	Γ(1, 1)	Γ(4, 1)	Γ(16, 1)
30	95	86	91	94	2	67	91
50	95	77	87	93	0	39	86
70	95	68	84	92	0	20	81
100	95	58	83	91	0	6	73
150	95	35	72	90	0	0	61
200	95	26	64	89	0	0	48
500	95	0	26	79	0	0	7
1000	95	0	2	62	0	0	0
2000	95	0	0	32	0	0	0
5000	95	0	0	2	0	0	0
6000	95	0	0	0	0	0	0

4.4. Analysis of Phase I Estimators on Different Distributions

In this section, we analyze how the Phase I sample size affects the covariance matrix estimators under different distributions (normal, gamma, t) when the process is in-control. The variability of the trace of the covariance matrix estimates for different values of *m* is presented in Table 3. We observe that the variability decreases as *m* increases. The multivariate normal distribution has the least variability when estimating $\hat{\Sigma}_0$. The non-parametric covariance matrix estimator (\hat{A}_0) has similar variability irrespective of the distribution.

Table 3. Variability of the trace for the covariance matrix estimates when *p* = 5.

m	Var(tr($\hat{\Sigma}_0$))			Var(tr(\hat{A}_0))		
	Normal	Gamma	t	Normal	Gamma	t
10	1.113	1.294	1.371	19.762	22.925	18.020
25	0.417	0.486	0.522	0.206	0.211	0.207
50	0.202	0.243	0.259	0.031	0.032	0.032
75	0.136	0.159	0.171	0.012	0.013	0.012
100	0.101	0.120	0.127	0.007	0.007	0.007
125	0.080	0.095	0.102	0.004	0.004	0.004
150	0.067	0.080	0.086	0.003	0.003	0.003
200	0.050	0.060	0.064	0.001	0.002	0.002

4.5. Performance of the Dispersion Charts

We simulated the *AARL* values for each of the four competing control charts under the different scenarios discussed in Section 4.2. Results are displayed in Tables 4–7 where gray highlighted values correspond to in-control performance and bold face indicates the best performing chart, i.e., the lowest *AARL*₁ value. The *AARL* values in italics show the most robust charts to non-normality under in-control performance. For Tables 6 and 7, which show results for the scenarios with decreasing shifts, we set the maximum run length to 10,000 in the simulations for the MEWMS and MaxMEWMV charts. Because these charts are not designed to detect decrease shifts and therefore have extremely long run lengths. This maximum ensures reasonable simulation time. We use an asterisk to indicate any *AARL*-value which is influenced by this maximization and is therefore a underestimate of the actual *AARL*. In addition, we represent *AARL* values greater than 1000 by > 1000.

4.5.1. Case 1: Overall Increase Shifts in the Variance-Covariance Matrix

For case 1, the results for different number of correlated quality characteristics (*p* = 2, 5, 25) under multivariate normal, gamma and t distributions when *m* = 200 are reported

in Table 4. For small and moderate increases in the process variance, we notice that the MEWMS chart has the best performance among the competing charts. For example, when $p = 25$ and $\delta = 1.1$, the MEWMS chart has the lowest $AARL_1$ value of 42, while the $AARL_1$ of the MaxMEWMV, the MNSE and the REWMV control charts are respectively 63, 203 and 55. In addition, the parametric charts (REWMV, MEWMS, and the MaxMEWMV charts) have almost similar performance at high increase shift ($\delta = 2$) in the process especially for a high value of p .

Table 4. $AARL$ values for an overall increase shift in the process dispersion when $\lambda = 0.3$, and $m = 200$ under multivariate normal, t and gamma distributions. Grey highlighted rows indicate in-control data.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
normal	0	1.0	203	201	201	200	198	200	197	200	204	196	202	198
		1.1	108	119	201	119	82	102	197	98	42	63	203	55
		1.2	66	77	202	78	43	59	197	56	16	27	203	25
		1.3	44	53	202	55	26	38	197	37	9	14	203	14
		1.5	25	30	201	33	13	19	198	20	4	6	202	8
		2.0	10	13	203	15	5	7	198	9	2	3	205	4
	0.3	1.0	159	118	130	165	134	113	95	152	132	127	79	152
		1.1	91	78	131	101	65	65	94	81	35	48	79	48
		1.2	59	55	131	68	37	42	94	49	15	23	79	23
		1.3	41	41	131	50	24	29	94	33	9	13	79	14
		1.5	23	26	131	31	13	17	94	19	5	6	79	8
		2.0	10	12	131	15	5	7	94	9	2	3	78	4
gamma	0	1.0	119	120	203	216	110	87	198	220	115	52	201	233
		1.1	75	80	202	126	56	54	197	106	32	26	202	63
		1.2	51	56	204	83	33	36	197	61	14	15	201	27
		1.3	37	42	201	58	22	26	197	39	8	10	202	15
		1.5	23	26	201	35	12	15	197	21	4	5	201	8
		2.0	11	13	203	16	5	7	198	9	2	2	200	4
	0.3	1.0	106	88	131	174	84	66	97	163	79	44	80	175
		1.1	68	62	132	108	47	43	94	86	27	23	80	54
		1.2	48	45	131	73	30	30	94	52	13	14	80	25
		1.3	36	35	130	53	21	23	94	35	8	9	81	15
		1.5	22	24	132	32	12	14	95	20	4	5	79	8
		2.0	11	12	130	15	6	7	94	9	2	2	79	4
t	0	1.0	134	134	201	203	101	99	198	182	45	43	202	98
		1.1	81	87	201	123	54	60	197	94	21	24	203	40
		1.2	54	60	204	81	33	40	198	56	12	15	203	21
		1.3	39	45	204	58	22	28	200	38	8	10	202	14
		1.5	23	27	201	35	12	17	197	21	5	6	202	8
		2.0	11	13	199	16	5	7	199	9	2	3	204	4
	0.3	1.0	116	92	132	168	82	71	96	143	41	37	79	85
		1.1	73	63	130	105	47	46	95	79	20	21	78	37
		1.2	50	47	131	72	30	33	95	50	12	14	79	21
		1.3	37	37	132	52	21	24	95	34	8	10	78	13
		1.5	23	24	132	32	12	15	94	20	5	6	79	8
		2.0	11	12	132	16	6	7	93	9	2	3	79	4

When the process variances are unchanged (i.e., $\delta = 1$) but the correlation increases ($\rho = 0.3$), the MaxMEWMV chart shows the best performance for $p = 2$ while the MNSE control chart outperforms the other compared charts for $p = 5$ and $p = 25$. When $p = 2$ the MaxMEWMV control chart perform best for simultaneous shifts in both process variances and the correlation. However, the MEWMS control chart has the best performance among the compared charts for $p = 5$ and $p = 25$.

We also observed that $AARL$ values of the REWMV and the MNSE control charts under normal, gamma and t distributions are almost equal for each of the process shifts we consider, i.e. they are robust to deviations from the normality assumption. However, the disadvantage of the MNSE chart is that it does not respond to the overall increase shifts in the process variance; see that all values are approximately 200. The MEWMS and MaxMEWMV control charts are seriously affected by slight deviation from the normality assumption. For instance, $AARL_0$ values of the MEWMS and the MaxMEWMV control charts are respectively 119 and 120 with gamma distributed data. The MEWMS and the MaxMEWMV charts still have the least $AARL_1$ values under the non-normal distributions but note that their in-control performance have the highest rate of false alarms with t and gamma distributions.

4.5.2. Case 2: Sparse Increase Shifts in the Variance-Covariance Matrix

Next, we consider a scenario where only one variable increases. Without loss of generality, we select the first diagonal element of the covariance matrix to changes where $\rho = 0$ and $\rho = 0.3$. Here, we simulated for $\delta = 1.1, 1.2, 1.3, 1.5$ and 2 under multivariate normal, gamma and t distributions, results are reported in Table 5.

Table 5. $AARL$ values for an overall increase shift in the process dispersion when $\lambda = 0.3$, and $m = 200$ under multivariate normal, t and gamma distributions. Grey highlighted rows indicate in-control data.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
normal	0	1.0	203	201	201	200	198	200	197	200	204	196	202	198
		1.1	145	152	200	153	162	170	196	172	187	185	203	187
		1.2	107	115	193	121	134	142	191	150	173	172	200	176
		1.3	81	89	186	99	111	119	184	131	159	159	199	165
		1.5	52	56	165	71	77	81	166	105	138	127	191	151
		2.0	23	25	120	39	38	36	116	68	92	69	164	123
	0.3	1.0	159	118	130	165	134	113	95	152	132	127	79	152
		1.1	119	98	136	129	114	100	96	132	124	122	79	145
		1.2	94	81	135	105	98	89	96	119	115	116	79	138
		1.3	74	67	134	88	84	77	95	106	108	108	79	131
		1.5	48	48	126	66	63	59	92	87	97	91	78	119
		2.0	23	24	102	38	35	31	74	59	71	56	72	100

Table 5. Cont.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
gamma	0	1.0	119	120	203	216	110	87	198	220	115	52	201	233
		1.1	93	96	200	164	93	78	195	189	107	49	201	219
		1.2	74	78	195	129	80	69	191	164	100	48	200	207
		1.3	60	63	184	105	70	61	185	143	95	46	198	195
		1.5	42	45	166	75	53	48	166	112	82	40	193	177
		2.0	22	24	120	41	31	28	115	73	58	29	161	142
	0.3	1.0	106	88	131	174	84	66	97	163	79	44	80	175
		1.1	86	75	137	138	74	60	95	144	76	43	79	166
		1.2	69	63	135	112	66	55	97	127	72	41	81	157
		1.3	56	54	134	93	58	49	96	114	68	39	80	148
		1.5	40	40	127	68	46	40	92	93	60	36	79	136
		2.0	22	23	101	40	29	25	75	64	46	27	73	113
t	0	1.0	134	134	201	203	101	99	198	182	45	43	202	98
		1.1	102	106	198	155	88	88	197	157	43	41	203	94
		1.2	81	85	195	124	77	78	191	139	41	41	202	90
		1.3	64	69	184	103	66	69	184	122	40	39	199	87
		1.5	44	48	168	74	52	54	164	100	38	36	192	81
		2.0	22	24	121	42	31	30	114	67	31	28	163	70
	0.3	1.0	116	92	132	168	82	71	96	143	41	37	79	85
		1.1	91	78	137	134	73	65	97	128	39	36	79	82
		1.2	75	66	136	110	64	59	95	114	38	35	79	80
		1.3	60	56	136	92	58	54	95	103	37	34	79	77
		1.5	42	42	130	69	46	43	92	86	34	32	78	72
		2.0	22	23	101	40	29	26	73	60	29	25	72	63

The MEWMS control chart performs best while the MNSE chart shows the worst performance when detecting only shifts in the process variances. For instance, when $p = 5$ at $\delta = 1.2$ under normally distributed data, the $AARL_1$ for the MEWMS, the MaxMEWMV are respectively 134 and 142, while that of the MNSE and the REWMV control charts are 191 and 150 respectively.

In addition, the MaxMEWMV chart has the best performance for $p = 2$ when the process variances are unchanged but the correction increases ($\rho = 0.3$), while the MNSE control chart outperforms the other compared charts for $p = 5$ and $p = 25$. The MaxMEWMV chart perform best when $p = 2$ and $p = 5$ for simultaneous shifts in the process variance and correlation while the MNSE chart outperforms the other charts for $p = 25$.

Overall, the out-of-control performance of the REWMV control chart is the most robust to non-normality for a sparse increase shift in the process variance among the compared parametric charts.

4.5.3. Case 3: Overall Decrease Shifts in the Variance-Covariance Matrix

We simulated for the overall decrease shifts ($\delta = 0.2, 0.3, 0.4, 0.6, \text{ and } 0.8$) in the covariance matrix when $\rho = 0$ or $\rho = 0.3$ (case 3). The simulation results for this scenario when $m = 200$ for $\lambda = 0.3$ under multivariate normal, gamma and t distributions are presented in Table 6. The MEWMS and MaxMEWMV control charts are inefficient in detecting decrease shifts in the variance as their $AARL_1$ values in most cases are greater

than 1000. We notice that the REWMV chart has the best performance in detecting overall decrease shifts in the process dispersion.

The MNSE control chart does not react to downward shifts in the process. For example, when $p = 2$ at $\delta = 0.2$ and $\rho = 0$, the $AARL_1$ value of the MNSE chart is 201 and that of the MEWMS and the MaxMEWMV charts are greater than 1000 but the proposed REWMV control chart has the least $AARL_1$ value of 12. In addition, with the t and gamma distributed data, we observe that the REWMV chart is robust to non-normality for the out-of-control performance.

Table 6. $AARL$ values for an overall decrease shift in the process dispersion (case 3) when $\lambda = 0.3$ and $m = 200$ under normal, t and gamma distributions. Grey highlighted rows indicate in-control data.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
normal	0	1.0	203	201	201	200	198	200	197	200	204	196	202	198
		0.8	> 1000	887 *	202	125	> 1000	> 1000	197	92	> 1000	> 1000	204	41
		0.6	> 1000	> 1000	202	71	> 1000	> 1000	200	39	> 1000	> 1000	201	11
		0.4	> 1000	> 1000	204	34	> 1000	> 1000	197	14	11	> 1000	198	4
		0.3	> 1000	> 1000	202	21	> 1000	> 1000	196	9	5	> 1000	194	3
		0.2	> 1000	> 1000	201	12	41	> 1000	193	5	4	> 1000	184	2
	0.3	0.8	799 *	380	131	122	> 1000	522 *	93	88	> 1000	> 1000	78	38
		0.6	> 1000	> 1000	132	68	> 1000	> 1000	92	37	> 1000	> 1000	77	10
		0.4	> 1000	> 1000	135	33	> 1000	> 1000	94	14	10	> 1000	79	4
		0.3	> 1000	> 1000	133	21	711 *	> 1000	94	9	5	> 1000	78	3
0.2		> 1000	> 1000	131	12	37	> 1000	95	5	4	> 1000	75	2	
gamma	0	1.0	119	120	203	216	110	87	198	220	115	52	201	233
		0.8	466	384	202	129	822 *	335	196	93	> 1000	331	199	40
		0.6	> 1000	> 1000	204	70	> 1000	> 1000	199	38	> 1000	> 1000	202	11
		0.4	> 1000	> 1000	203	35	> 1000	> 1000	194	15	10	> 1000	190	4
		0.3	> 1000	> 1000	202	22	> 1000	> 1000	195	9	5	> 1000	180	3
		0.2	> 1000	> 1000	198	13	44	> 1000	193	6	4	> 1000	156	2
	0.3	0.8	353	237	132	123	445 *	221	94	89	> 1000	268	78	39
		0.6	> 1000	> 1000	134	68	> 1000	> 1000	95	37	> 1000	> 1000	77	11
		0.4	> 1000	> 1000	133	34	> 1000	> 1000	93	15	10	> 1000	76	4
		0.3	> 1000	> 1000	136	21	> 1000	> 1000	95	9	5	> 1000	71	3
0.2		> 1000	> 1000	134	13	39	> 1000	91	5	4	> 1000	62	2	
t	0	1.0	134	134	201	203	101	99	198	182	45	43	202	98
		0.8	523	412	199	113	604 *	372	197	77	557	221	205	29
		0.6	> 1000	> 1000	199	64	> 1000	> 1000	196	34	462	> 1000	199	9
		0.4	> 1000	> 1000	200	31	> 1000	> 1000	198	13	10	> 1000	198	4
		0.3	> 1000	> 1000	204	20	604 *	> 1000	200	8	5	> 1000	193	3
		0.2	> 1000	> 1000	200	12	35	> 1000	197	5	4	> 1000	184	2
	0.3	0.8	413	241	132	109	396	223	95	74	420	186	79	29
		0.6	> 1000	> 1000	130	63	> 1000	> 1000	96	33	337	> 1000	78	9
		0.4	> 1000	> 1000	134	30	> 1000	> 1000	93	13	9	> 1000	80	4
		0.3	> 1000	> 1000	131	20	> 1000	> 1000	96	8	5	> 1000	78	3
0.2		> 1000	> 1000	134	12	32	> 1000	98	5	4	> 1000	77	2	

4.5.4. Case 4: Sparse Decrease Shifts in the Variance-Covariance Matrix

Consider case 4, where only the first variance decreases, results are displayed in Table 7. The proposed REWMV control chart outperforms the other charts irrespective of the values of p when the sparse decrease shifts are introduced under multivariate normal, gamma and t distributions and $\rho = 0$. The MEWMS and the MaxMEWMV charts show the worst performance among the compared charts. For instance, when $p = 5$ at $\delta = 0.3$, the $AARL_1$ values of the MEWMS, the MaxMEWMV are respectively 714 and 319 under normal distribution; while the REWMV control chart has the lowest $AARL_1$ value of 76. The MNSE chart performs best when there is simultaneously shifts in both process variances and correlations.

The MNSE control chart, again, does not detect any change as the $AARL_0$ values are approximately 200 under multivariate t and gamma distributions. The REWMV chart is the most robust to normality assumption among the compared parametric charts for the out-of-control performance. For instance, the AARL values for the REWMV chart under normal, gamma, and t distributions are respectively 56, 57, and 49 for $p = 5$, $\delta = 0.2$ and $\rho = 0$ while that of MaxMEWMV chart are respectively 319, 117, and 155. However, the excessive rate of false alarm of the MEWMS and MaxMEWMV control charts with the non-normality affect the out-of-control performance of the charts such that their $AARL_1$ are lower than that of the MNSE and the REWMV control charts when $p = 25$.

Table 7. AARL values for a sparse shift in the process variance when $\lambda = 0.3$ and $m = 200$ under normal, t and gamma distributions. Grey highlighted rows indicate in-control data.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
normal	0	1.0	203	201	201	200	198	200	197	200	204	196	202	198
		0.8	432	354	187	155	298	265	189	169	241	208	203	179
		0.6	896 *	592 *	150	114	426	304	167	135	266	213	196	160
		0.4	> 1000	784 *	89	75	622 *	314	128	98	321	227	186	135
		0.3	> 1000	817 *	60	55	730 *	319	109	76	349	227	179	115
		0.2	> 1000	807 *	36	37	847 *	319	88	56	360	228	173	96
	0.3	0.8	297	172	115	149	182	131	88	159	150	135	77	163
		0.6	542 *	253	84	107	251	153	76	124	173	141	75	149
		0.4	943 *	401	46	69	320	173	58	89	192	148	71	125
		0.3	> 1000	550	31	50	356	181	50	69	199	150	70	108
		0.2	> 1000	711 *	18	33	407	185	40	50	220	145	65	89
		gamma	0	1.0	119	120	203	216	110	87	198	220	115	52
0.8	215			191	190	156	146	104	188	168	133	54	200	175
0.6	358			277	151	118	193	116	165	136	145	55	191	158
0.4	522			325	90	76	229	118	129	99	162	56	183	133
0.3	588 *			329	60	57	259	121	110	77	174	56	177	115
0.2	660 *			331	37	39	290	117	88	57	179	54	174	97
0.3	0.8		166	125	113	151	105	76	89	160	86	45	78	161
	0.6		280	170	83	109	130	87	76	127	98	45	75	147
	0.4		405	244	46	69	153	92	59	88	102	47	72	123
	0.3		456	269	31	51	170	90	51	71	104	47	69	107
	0.2		529	322	18	34	182	93	40	51	115	47	67	88

Table 7. Cont.

Distribution	ρ	δ	$p = 2$				$p = 5$				$p = 25$			
			MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV	MEWMS	MaxMEWMV	MNSE	REWMV
t	0	1.0	134	134	201	203	101	99	198	182	45	43	202	98
		0.8	232	210	192	140	137	120	190	142	48	44	202	107
		0.6	434	320	149	105	177	134	166	113	51	47	194	98
		0.4	733 *	424	91	68	226	149	129	81	55	47	184	82
		0.3	908 *	463	60	51	261	151	108	67	58	48	179	74
	0.2	> 1000	462	36	35	287	155	85	49	59	50	175	63	
	0.3	0.8	189	123	114	136	106	83	86	131	42	39	77	101
		0.6	315	173	80	99	130	92	75	104	46	40	76	90
		0.4	500	257	46	62	160	102	60	76	49	41	72	79
		0.3	616 *	312	30	47	179	110	50	59	52	42	69	70
0.2		751 *	420	18	31	200	115	40	44	53	42	67	60	

4.6. Further Investigation on Robustness

In this section, we want to investigate the robustness of the charts to non-normality under the in-control performance, results are displayed in Figure 1. We employed t distribution with different degree of freedom (15 up to 50) to model moderate to high kurtosis for the 4 considered multivariate EWMA dispersion charts. We notice that the proposed chart has the most robust performance among the three parametric charts because its $AARL_0$ is closest to 200. Moreover, its performance is almost similar to the MNSE chart.

In addition, we observe that for all the values of α we considered for the gamma distributed data, the proposed REWMV and the MNSE charts have approximately $AARL_0 = 200$. In general, the MaxMEWMV and the MEWMS charts show the worst performance in terms of their robustness to non-normally distributed data.

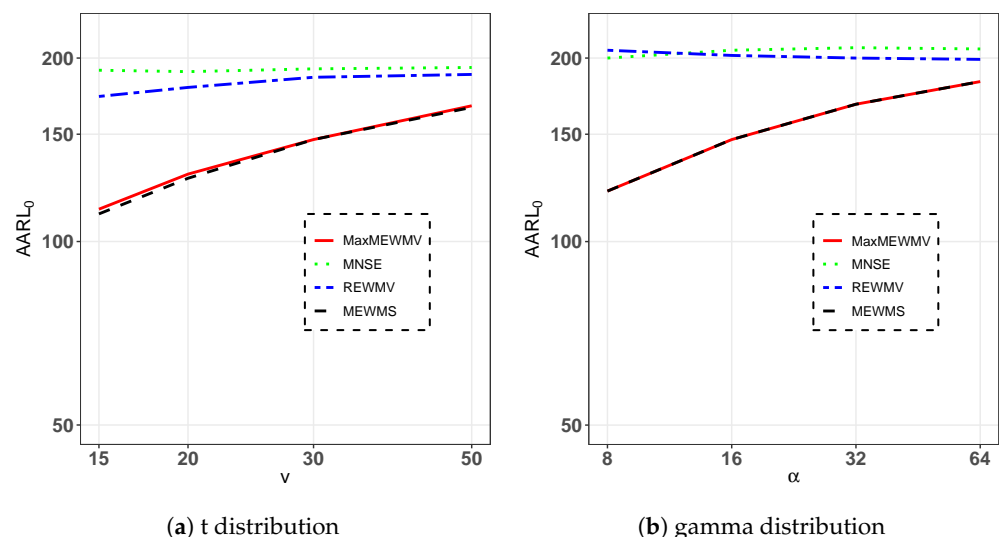


Figure 1. In-control AARL for each of the compared multivariate EWMA dispersion charts for individual observations versus degree of freedom in (a) and shape parameter in (b). Charts are designed with $\lambda = 0.1$, $m = 50$ for $p = 2$ and an UCL from Table 1 using normally distributed data.

Overall, in Tables 4–7 and Figure 1, we notice that the REWMV and MNSE control charts consistently show the most robust performance to non-normality in the in-control. However, the MNSE shows very poor performance for overall shifts in the process variance.

The MEWMS and MaxMEWMV control charts have the best performance for increase shifts in the process variances, only if the data is perfectly normally distributed. However, the proposed REWMV control chart shows the best performance for the overall decrease shifts in the process variances and comparable performance to the MEWMS and MaxMEWMV charts for increases. In addition it shows robust performance if the data show small deviation from normality.

5. Examples

In this section, we demonstrate how to apply the proposed multivariate dispersion charts to real applications. We consider two different datasets; one healthcare and one industrial application.

5.1. Example to Monitor Elder Care Center Dataset

In this section, we employed the daily measurement of systolic blood pressure (SBP), see Zwetsloot and Ajadi [18] for details about the dataset. Zwetsloot and Ajadi [18] show that the data deviates slightly from the normality assumption. In this section, we arrange the dataset as multivariate individual observations where we select five elders. The first 30 multivariate observations (i.e., $m = 30$ for $p = 5$) are employed to estimate the process parameters. The Phase I estimates are computed based on Equation (9), we obtain the estimated process mean vector and sample covariance matrix ($\hat{\Sigma}_0$) respectively as

$$\hat{\mu}_0 = (121.9 \quad 128.1 \quad 111.8 \quad 124.7 \quad 133.1),$$

$$\hat{\Sigma}_0 = \begin{pmatrix} 113.706 & -34.301 & 9.037 & -19.175 & 55.147 \\ -34.301 & 187.099 & 6.395 & 33.294 & -46.246 \\ 9.037 & 6.395 & 55.013 & -3.444 & 13.361 \\ -19.175 & 33.294 & -3.444 & 94.202 & -16.671 \\ 55.147 & -46.246 & 13.361 & -16.671 & 237.995 \end{pmatrix}.$$

Moreover, for the MNSE chart, the estimated location and scatter matrix are obtained through Equations (3) and (4):

$$\hat{\theta}_0 = (121.1 \quad 129.7 \quad 111.2 \quad 124.9 \quad 132.5),$$

$$\hat{A}_{\rho 0} = \begin{pmatrix} 1 & 0.071 & -0.264 & 0.311 & -0.196 \\ 0 & 0.646 & 0.032 & -0.359 & 0.121 \\ 0 & 0.000 & 1.129 & 0.112 & -0.005 \\ 0 & 0.000 & 0.000 & 0.859 & 0.032 \\ 0 & 0.000 & 0.000 & 0.000 & 0.443 \end{pmatrix}.$$

Due to a change in the measurement system, for details see Zwetsloot and Ajadi [18] and Mahmood et al. [23], there is a downward shift in the process variance from 12 February 2018 onwards. We try to detect this change with the different charts.

We construct a two-sided REWMV chart and the other competing charts for Phase II monitoring where the in-control $AARL$ is fixed at 100 for $\lambda = 0.1$. We obtained the control limits for the MEWMS, the MaxMEWMV and the MNSE charts respectively as 3.17, 8.4, and 7.26. In addition, the control chart's constants for the REWMV chart in the upper- and lower-sided monitoring are respectively -2.964 and -9.13 , where the reflection boundary is -6.35 , i.e., $\sum_{i=1}^p E[\log D_i] = -1.27p$ for $p = 5$.

Figure 2 displays the charts. It is obvious from Figure 2a,b that there is no signal received for detecting decrease shifts in the covariance matrix with the MEWMS and MaxMEWMV control charts, which is inline with the simulation results. Figure 2d shows that the REWMV chart first signals (three out-of-control signals) for decrease shifts in the process during the shifted period of the SBP measurement. Moreover, we observe in Figure 2c that the MNSE chart triggers one signal on 2 March 2018. This supports our simulation results in Section 4.5 that the proposed REWMV control chart is very effective in detection of decrease shifts in the covariance matrix when data are normally distributed or possibly slightly skewed or heavy tailed.

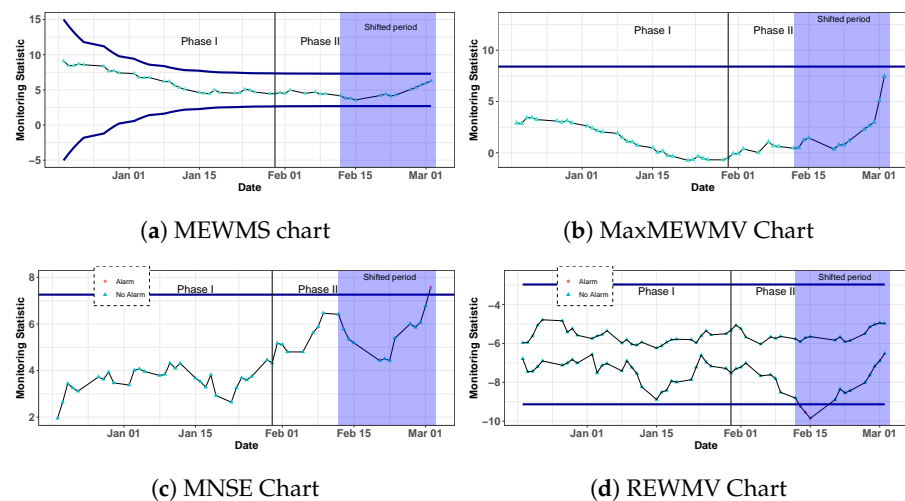


Figure 2. Monitoring SBP multivariate data with the four compared charts.

5.2. Industrial Application

We also employed an industrial dataset collected from a mechanical process to monitor process dispersion of the compared charts. The dataset is available from the MSQC package in R; see Santos-Fernández [24] for details. The Phase I dataset (mech1) contains 45 individual observations ($m = 45$) with seven correlated quality characteristics ($p = 7$). Based on this, we compute the phase I estimates; we also obtained the control limits for the charts respectively as 3.5, 8.75, and 9.56 for the MEWMS, the MaxMEWMV and the MNSE control charts. In addition, the control chart’s limits for the REWMV chart in the upper- and lower-sided monitoring are respectively -4.55 and -12.1 . These control limits are obtained based on $m = 45$ at $\lambda = 0.1$ that satisfy $ARL_0 = 100$.

Next, we applied the mech2 (Phase II) dataset for online monitoring. Figure 3 shows the charts applied to the Phase II data. We noticed that the MaxMEWMV chart detects no signal in the process variability where the MEWMS chart first detects increase shifts in the covariance matrix at observations 22 and 25. The MNSE chart detects a shift at observation 39 and the REWMV control chart signals for a decrease shift at 46. As we do not know if or when the data is out-of-control it is impossible to tell which chart is signalling a false alarm or which chart has signalled a shift quickest.

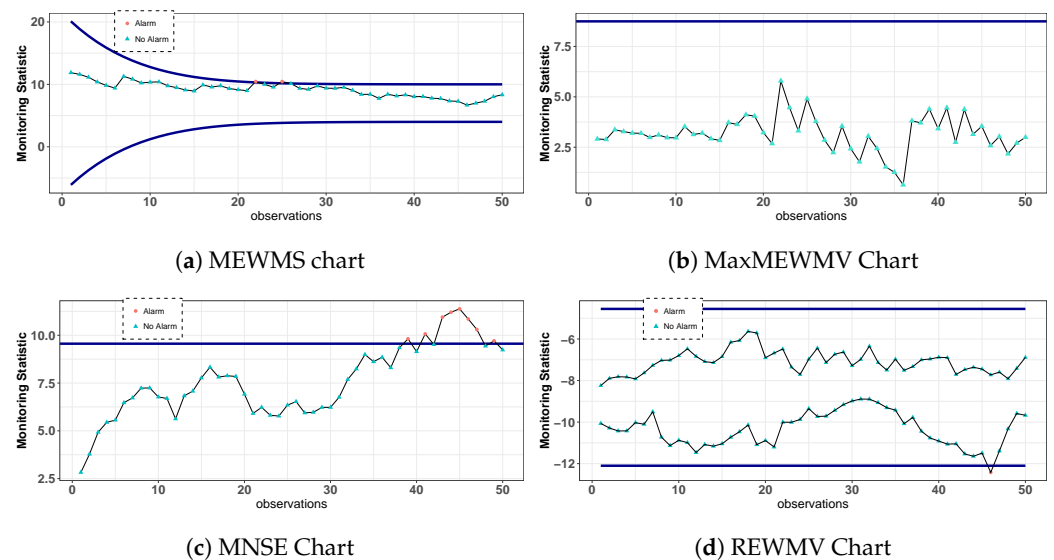


Figure 3. Monitoring SBP multivariate data with the four compared charts.

6. Conclusions and Recommendation

In this paper, we introduced the REWMV control chart, a chart that is effective in detecting changes in the process dispersion when monitoring multivariate processes based on individual observations. The chart is robust to slight deviations from the normality assumption and guarantees in-control performance as expected and reasonable quick shift detection. It does not suffer from excessive false alarms like most method. This chart applies the logarithm to the covariance matrix. The REWMV is a new robust multivariate dispersion chart. We compared the performance of the REWMV control chart with some existing multivariate EWMA dispersion charts based on estimated parameters under normally and non-normally distributed data.

The REWMV chart outperforms the other competing charts when there is an overall decrease shift in the covariance matrix and shows comparable performance for increase shifts. In general, the MEWMS and the MaxMEWMV control charts have bad performance for a decrease in the process variances, while they show the best performance among the compared charts for an increase shifts in covariance matrix. Moreover, the performance of the non-parametric MNSE chart is bad for an overall increase or decrease in the process dispersion.

Next, we investigated the robustness of the REWMV control chart and the other multivariate EWMA dispersion charts for individual observations. We showed that the in-control performance of the MaxMEWMV and MEWMS control charts suffers from an excessive rate of false alarms with the gamma and t distributions. Both the MNSE and the REWMV control charts are robust to non-normality under in-control and out-of-control performance, where the MNSE chart is almost perfectly distribution free.

We apply only the constant control limits for monitoring the proposed REWMV chart by using two separate charts for detecting upward and downward shifts in the process dispersion. However, we recommend the use of probability limits for future work to enhance simultaneous monitoring of both increase and decrease shifts in the covariance matrix. In addition, in this article, we are only restricted to monitoring the process mean vector and covariance matrix, however, other shape parameters shifts may also be investigated for monitoring the REWMV charts.

The focus of this paper is on monitoring process variability with individual observations, however, the proposed chart can as well be extended to monitoring grouped observations where the sample covariance matrix of the grouped observations are incorporated in the EWMA statistic of Equation (6). Moreover, we can extend this study by investigating the robustness of some Phase I covariance matrix estimators to outliers with the proposed REWMV control chart. In addition, integrating two multivariate control charts to detect the process mean vector and covariance matrix simultaneously sometimes loses the ability to provide detailed diagnostic information when a process is out of control. Hence, the performance of the proposed REWMV chart can be investigated when monitoring such out-of-control scenario.

Author Contributions: Methodology, I.M.Z. and J.O.A.; Formal analysis, I.M.Z.; Supervision, K.-L.T.; Writing—original draft, J.O.A. All authors have read and agreed to the published version of the manuscript.

Funding: We are thankful to the Centre of Systems Informatics Engineering (CSIE) of City University of Hong Kong for their support and for supplying the data used as illustrating healthcare case in this paper. The data are part of a larger project partly supported by Research Grants Council Theme-Based Research Scheme [T32-102-14N] and City University of Hong Kong [9610406]. The work of I.M. Zwetsloot was partially support by the City University of Hong Kong [7005090,705567].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: The R code for the simulations in this article is available on the github page of the author with account ajadi1982.

References

1. Hotelling, H. Multivariate Quality Control Illustrated by the Air Testing of Sample Bomb Sights, Techniques of Statistical Analysis, Ch. II. In *Selected Techniques of Statistical Analysis for Scientific and Industrial Research, and Production and Management Engineering*; McGraw-Hill: New York, NY, USA, 1947; pp. 111–184.
2. Lowry, C.; Woodall, W.; Champ, C.; Rigdon, S.E. A multivariate exponentially weighted moving average control chart. *Technometrics* **1992**, *34*, 46–53. [[CrossRef](#)]
3. Crosier, R.B. Multivariate generalizations of cumulative sum quality-control schemes. *Technometrics* **1988**, *30*, 291–303. [[CrossRef](#)]
4. Alt, F.B. Multivariate quality control. *Encycl. Stat. Sci.* **1985**, *6*, 110–122.
5. Yeh, A.B.; Lin, D.K.; McGrath, R.N. Multivariate control charts for monitoring covariance matrix: A review. *Qual. Technol. Quant. Manag.* **2006**, *3*, 415–436. [[CrossRef](#)]
6. Bersimis, S.; Psarakis, S.; Panaretos, J. Multivariate statistical process control charts: An overview. *Qual. Reliab. Eng. Int.* **2007**, *23*, 517–543. [[CrossRef](#)]
7. Ajadi, J.O.; Zwetsloot, I.M. Should observations be grouped for effective monitoring of multivariate process variability? *Qual. Reliab. Eng. Int.* **2020**, *36*, 1005–1027. [[CrossRef](#)]
8. Huwang, L.; Yeh, A.B.; Wu, C.W. Monitoring multivariate process variability for individual observations. *J. Qual. Technol.* **2007**, *39*, 258–278. [[CrossRef](#)]
9. Reynolds, M.R., Jr.; Stoumbos, Z.G. Control charts and the efficient allocation of sampling resources. *Technometrics* **2004**, *46*, 200–214. [[CrossRef](#)]
10. Reynolds, M.R., Jr.; Stoumbos, Z.G. Should observations be grouped for effective process monitoring? *J. Qual. Technol.* **2004**, *36*, 343–366. [[CrossRef](#)]
11. Ajadi, J.O.; Wang, Z.; Zwetsloot, I.M. A review of dispersion control charts for multivariate individual observations. *Qual. Eng.* **2021**, *33*, 60–75. [[CrossRef](#)]
12. Yeh, A.; Huwang, L.; Chien-Wei, W.U. A multivariate EWMA control chart for monitoring process variability with individual observations. *IIE Trans.* **2005**, *37*, 1023–1035. [[CrossRef](#)]
13. Hawkins, D.M.; Maboudou-Tchao, E.M. Multivariate exponentially weighted moving covariance matrix. *Technometrics* **2008**, *50*, 155–166. [[CrossRef](#)]
14. Li, Z.; Zou, C.; Wang, Z.; Huwang, L. A multivariate sign chart for monitoring process shape parameters. *J. Qual. Technol.* **2013**, *45*, 149–165. [[CrossRef](#)]
15. Stoumbos, Z.G.; Sullivan, J.H. Robustness to non-normality of the multivariate EWMA control chart. *J. Qual. Technol.* **2002**, *34*, 260–276. [[CrossRef](#)]
16. Testik, M.C.; Runger, G.C.; Borrór, C.M. Robustness properties of multivariate EWMA control charts. *Qual. Reliab. Eng. Int.* **2003**, *19*, 31–38. [[CrossRef](#)]
17. Borrór, C.M.; Montgomery, D.C.; Runger, G.C. Robustness of the EWMA Control Chart to Non-Normality. *J. Qual. Technol.* **1999**, *31*, 309–316. [[CrossRef](#)]
18. Zwetsloot, I.M.; Ajadi, J.O. A comparison of EWMA control charts for dispersion based on estimated parameters. *Comput. Ind. Eng.* **2019**, *127*, 436–450. [[CrossRef](#)]
19. Zou, C.; Tsung, F. A multivariate sign EWMA control chart. *Technometrics* **2011**, *53*, 84–97. [[CrossRef](#)]
20. Box, G.E.P.; Hunter, W.; Hunter, J. *Statistics for Experimenters*, 1st ed.; John Wiley and Sons: New York, NY, USA, 1978.
21. Crowder, S.V.; Hamilton, M.D. An EWMA for Monitoring a Process Standard Deviation. *J. Qual. Technol.* **1992**, *24*, 12–21. [[CrossRef](#)]
22. Shu, L.; Jiang, W. A New EWMA Chart for Monitoring Process Dispersion. *J. Qual. Technol.* **2008**, *40*, 319–331. [[CrossRef](#)]
23. Mahmood, T.; Wittenberg, P.; Zwetsloot, I.M.; Wang, H.; Tsui, K.L. Monitoring data quality for telehealth systems in the presence of missing data. *Int. J. Med Inform.* **2019**, *126*, 156–163. [[CrossRef](#)]
24. Santos-Fernández, E. *Multivariate Statistical Quality Control Using R*; Springer Science & Business Media, Germany Science & Business Media: Berlin/Heidelberg, Germany, 2012; Volume 14.