A robust standard deviation control chart

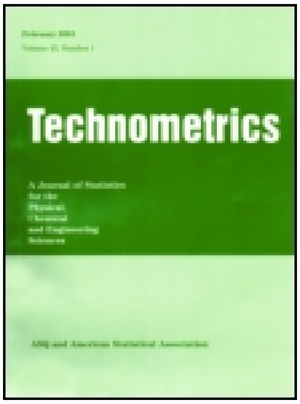
Schoonhoven, M.; Does, R.J.M.M.

DOI
10.1080/00401706.2012.648869

Publication date
2012

Document Version
Final published version

Published in
Technometrics



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

A Robust Standard Deviation Control Chart

Marit Schoonhoven [a] & Ronald J.M.M. Does [a]
[a] Institute for Business and Industrial Statistics, University of Amsterdam, 1018 TV, Amsterdam, The Netherlands

Available online: 20 Dec 2011

A Robust Standard Deviation Control Chart

Marit Schoonhoven and Ronald J.M.M. Does

Institute for Business and Industrial Statistics, University of Amsterdam
1018 TV Amsterdam, The Netherlands
(m.schoonhoven@uva.nl; r.j.m.m.does@uva.nl)

This article studies the robustness of Phase I estimators for the standard deviation control chart. A Phase I estimator should be efficient in the absence of contaminations and resistant to disturbances. Most of the robust estimators proposed in the literature are robust against either diffuse disturbances, that is, outliers spread over the subgroups, or localized disturbances, which affect an entire subgroup. In this article, we compare various robust standard deviation estimators and propose an algorithm that is robust against both types of disturbances. The algorithm is intuitive and is the best estimator in terms of overall performance. We also study the effect of using robust estimators from Phase I on Phase II control chart performance. Additional results for this article are available online as Supplementary Material.

KEY WORDS: Average run length; Mean squared error; Phase I; Phase II; Statistical process control.

1. INTRODUCTION

The performance of a process depends on the stability of its location and dispersion parameters and any change in these parameters should be detected as soon as possible. To monitor these parameters, Shewhart (1931) introduced the idea of control charts in the 1920s. The dispersion parameter is controlled first, followed by the location parameter. The present article focuses on control charts for monitoring the standard deviation.

We assume that in the design of such charts, the in-control standard deviation (σ) is unknown. Therefore, σ must be estimated from subgroups taken when the process is assumed to be in control. This stage in the control charting process is denoted as Phase I (cf. Woodall and Montgomery 1999). Control limits are calculated from the estimated σ to monitor the process standard deviation in Phase II. The Phase I and Phase II data are arranged in subgroups indexed by i. We denote by X_{ij}, $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n$, the Phase I data and by Y_{ij}, $i = 1, 2, \ldots$ and $j = 1, 2, \ldots, n$, the Phase II data. The X_{ij}'s are assumed to be independent and $N(\mu, \sigma^2)$ distributed and the Y_{ij}'s are assumed to be independent and $N(\mu, (\lambda\sigma)^2)$ distributed, where λ is a constant. When $\lambda = 1$, the standard deviation is in control; otherwise it has changed. Let $\hat{\sigma}$ be an unbiased estimate of σ based on the X_{ij}'s, and let $\hat{\sigma}_i$ be an unbiased estimate of $\lambda\sigma$ based on the ith subgroup Y_{ij}, $j = 1, 2, \ldots, n$. The process standard deviation can be monitored in Phase II by plotting $\hat{\sigma}_i$ on a Shewhart-type control chart with limits

$$\widehat{\mathrm{UCL}} = U_n\hat{\sigma}, \ \widehat{\mathrm{LCL}} = L_n\hat{\sigma}, \qquad (1)$$

where U_n and L_n are chosen so that the desired control chart behavior is achieved when the process is in control. When $\hat{\sigma}_i$ falls within the control limits, the process is deemed to be in control. We define F_i as the event that $\hat{\sigma}_i$ falls beyond the limits, $P(F_i) = p$ as the probability of that event and RL as the run length, that is, the number of subgroups until the first $\hat{\sigma}_i$ falls beyond the limits. When the limits are known, F_s and F_t ($s \neq t$) are independent and therefore RL is geometrically distributed with parameter p. Hence, the average run length (ARL) is given by $1/p$ and the standard deviation of the run length (SDRL) by $\sqrt{1-p}/p$. It is common practice to use $p = 0.0027$ and so ARL $= 370.4$ and SDRL $= 369.9$ when σ is known.

When the standard deviation is estimated, the conditional run length—the run length given an estimate of σ—has a geometric distribution. However, the unconditional RL distribution—the run length distribution averaged over all possible values of the estimated σ—is not geometric. Quesenberry (1993) showed that for the $\overline{X}$ and X control charts, the unconditional ARL as well as the unconditional p are higher than in the (μ,σ)-known case. Chen (1998) studied the unconditional run length distribution of the standard deviation control chart and showed that the situation is somewhat better than for the $\overline{X}$ control chart. To achieve the intended unconditional in-control performance when the limits are estimated, one could derive U_n and L_n by controlling either the in-control p or ARL, or a percentile point of the in-control RL distribution. An advantage of using the ARL is its intuitive interpretation. A drawback, however, is that the ARL is strongly determined by the occurrence of extremely long runs. Hillier (1969) and Yang and Hillier (1970) derived correction factors for the range (R) and standard deviation (S) control charts by controlling p.

Jensen et al. (2006) conducted a literature survey of the effects of parameter estimation on control chart properties and identified issues for future research. Their suggestion on page 360 is the subject of the present article. More specifically, we will find robust estimators for Phase I data and we will study the performance of these robust estimators during Phase II monitoring.

Rocke (1989) proposed standard deviation control charts based on the mean or the trimmed mean of the subgroup ranges or subgroup interquartile ranges. Moreover, he studied a two-stage procedure whereby the initial chart is constructed first and then groups that seem to be out of control are excluded. The control limits are recomputed from the remaining subgroups. Rocke (1992) provided the practical details for the

TECHNOMETRICS, FEBRUARY 2012, VOL. 54, NO. 1
DOI: 10.1080/00401706.2012.648869

construction of these charts. Tatum (1997) explained the difference between diffuse and localized disturbances: diffuse disturbances are equally likely to perturb any observation, whereas localized disturbances affect all observations in a subgroup. He proposed a method, constructed around a variant of the biweight A estimator, that is resistant to both diffuse and localized disturbances. Finally, Davis and Adams (2005) proposed a diagnostic technique for monitoring data that might be contaminated with outliers to react to signals that indicate a true process shift only.

In this article, we investigate robust Phase I estimators for the subgroup standard deviation control chart. The estimators considered are the pooled standard deviation, the robust biweight A estimator of Tatum (1997), and several adaptive trimmers. Additionally, we look at an adaptive trimmer based on the mean deviation from the median, a statistic more resistant to diffuse outliers (cf. Schoonhoven, Riaz, and Does 2011). For diffuse outliers, we think that a control chart for individual observations would detect outliers more quickly. We therefore include an estimator based on the individuals chart. To measure the variability within and not between subgroups, we correct for differences in the location between the subgroups. Finally, we present an algorithm that combines the last two approaches. The performance of the estimators is evaluated by assessing their mean squared error (MSE) under normality and in the presence of several types of contaminations. Moreover, we derive factors for the Phase II limits of the standard deviation control chart and assess the performance of the control charts by means of a simulation study.

The article is structured as follows. The next section introduces the standard deviation estimators, demonstrates the implementation of the estimators by means of a real-world example, and assesses their MSE. Next, we present the design schemes for the standard deviation control chart and derive the Phase II control limits. We then describe the simulation procedure and simulation results. The article ends with some concluding remarks.

2. PROPOSED PHASE I ESTIMATORS

In practice, the same statistic is generally used to estimate both the in-control standard deviation σ in Phase I and the standard deviation $\lambda\sigma$ in Phase II. Since the requirements for the estimators differ between the two phases, this is not always the best choice. In Phase I, an estimator should be efficient in uncontaminated situations and robust against disturbances, whereas in Phase II, the estimator should be sensitive to disturbances (cf. Jensen et al. 2006). In this section, we present six Phase I estimators, demonstrate the implementation of the estimators by means of a real data example, and assess the efficiency of the estimators in terms of their MSE.

2.1 Estimators of the Standard Deviation

Recall that X_{ij}, $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n$, denotes the Phase I data with n the subgroup size and k the number of subgroups.

The first estimator of σ is based on the pooled subgroup standard deviation

$$\tilde{S} = \left(\frac{1}{k}\sum_{i=1}^{k} S_i^2\right)^{1/2}, \tag{2}$$

where S_i is the ith subgroup standard deviation defined by

$$S_i = \left(\frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2\right)^{1/2}.$$

An unbiased estimator is given by $\tilde{S}/c_4(k(n-1)+1)$, where $c_4(m)$ is defined by

$$c_4(m) = \left(\frac{2}{m-1}\right)^{1/2}\frac{\Gamma(m/2)}{\Gamma((m-1)/2)}.$$

This estimator provides a basis for comparison under normality when no contaminations are present. Mahmoud et al. (2010) showed that this estimator is more efficient than the mean of the subgroup standard deviations and the mean of the subgroup ranges when the data are normally distributed.

We also evaluate a robust estimator proposed by Tatum (1997). This approach is applicable for $n \geq 4$. The method begins by calculating the residuals in each subgroup, which involves subtracting the subgroup median from each value: $\mathrm{res}_{ij} = X_{ij} - M_i$. If n is odd, then in each subgroup, one of the residuals will be zero and is dropped. As a result, the total number of residuals is $m' = nk$ when n is even and $m' = (n-1)k$ when n is odd. Tatum's estimator is given by

$$S_c^* = \frac{m'}{(m'-1)^{1/2}}\frac{\left(\sum_{i=1}^{k}\sum_{j:|u_{ij}|<1}\mathrm{res}_{ij}^2\left(1-u_{ij}^2\right)^4\right)^{1/2}}{\left|\sum_{i=1}^{k}\sum_{j:|u_{ij}|<1}\left(1-u_{ij}^2\right)\left(1-5u_{ij}^2\right)\right|}, \tag{3}$$

where $u_{ij} = h_i\mathrm{res}_{ij}/(cM^*)$, M^* is the median of the absolute values all residuals,

$$h_i = \begin{cases} 1 & E_i \leq 4.5, \\ E_i - 3.5 & 4.5 < E_i \leq 7.5, \\ c & E_i > 7.5, \end{cases}$$

and $E_i = \mathrm{IQR}_i/M^*$. IQR_i is the interquartile range of subgroup i and is defined as the difference between the second-smallest and the second-largest observation for $4 \leq n \leq 7$, and as the difference between the third-smallest and the third-largest observation for $8 \leq n \leq 11$. The constant c is a tuning constant. Each value of c leads to a different estimator. Tatum studied the behavior of the estimator for $c = 7$ and $c = 10$ and showed that $c = 7$ gives an estimator that loses some efficiency in the absence of disturbances but gains efficiency in the presence of disturbances. We apply this value of c in our simulation study. Note that we have $h(i) = E_i - 3.5$ for $4.5 < E_i \leq 7.5$ instead of $h(i) = E_i - 4.5$, as presented by Tatum (1997, p. 129). This was a typographical error, resulting in too much weight on localized disturbances and thus an overestimation of σ. An unbiased estimator of σ is given by $S_c^*/d^*(c, n, k)$, where $d^*(c, n, k)$ is a normalizing constant. During the implementation of the estimator, we discovered that for odd values of n, the values of $d^*(c, n, k)$ given by Table 1 in Tatum (1997) are incorrect. We use the corrected values, which are presented in Table 1. The resulting estimator is denoted by $D7$, as in Tatum (1997).

Table 1. Normalizing constants $d^*(c, n, k)$ for Tatum's estimator (S_c^*)

	$c = 7$			$c = 10$		
n	$k = 20$	$k = 30$	$k = 40$	$k = 20$	$k = 30$	$k = 40$
5	1.070	1.069	1.068	1.054	1.053	1.053
7	1.057	1.056	1.056	1.041	1.040	1.040
9	1.052	1.051	1.050	1.034	1.034	1.033
11	1.047	1.046	1.046	1.029	1.029	1.028
13	1.044	1.044	1.043	1.026	1.025	1.025
15	1.041	1.041	1.041	1.023	1.023	1.023

We also include other procedures to obtain $\hat{\sigma}$. The first is a variant of Rocke (1989). Rocke's procedure first estimates σ by the mean subgroup range

$$\bar{R} = \frac{1}{k}\sum_{i=1}^{k} R_i, \tag{4}$$

where R_i is the range of the ith subgroup. An unbiased estimator of σ under normality is $\bar{R}/d_2(n)$, where $d_2(n)$ is the expected range of a random $N(0, 1)$ subgroup of size n. Values of $d_2(n)$ can be found in Duncan (1974, table M). Any subgroup that exceeds the Phase I control limits is deleted and $\bar{R}$ is recomputed from the remaining subgroups. Our approach is similar but continues until all subgroup ranges fall between the Phase I control limits. These are set at $\widehat{\mathrm{UCL}} = U_n\bar{R}/d_2(n)$ and $\widehat{\mathrm{LCL}} = L_n\bar{R}/d_2(n)$. We derive the factors U_n and L_n from the 0.99865 and 0.00135 quantiles of the distribution of $\bar{R}/d_2(n)$. Table 2 shows the factors for $n = 4, 5, 9$ as well as the constants added to obtain unbiased estimates from the screened data. The factors as well as the constants are obtained by simulation. Note that the factors and the constants are the same for $k = 20, 50, 100$. The resulting estimator is denoted by $\bar{R}^s$.

In addition, we evaluate an adaptive trimmer where the estimate of σ is obtained by the mean subgroup average deviation from the median instead of $\bar{R}$. The mean subgroup average deviation from the median is given by

$$\overline{\mathrm{MD}} = \frac{1}{k}\sum_{i=1}^{k} \mathrm{MD}_i, \tag{5}$$

where MD_i is the average absolute deviation from the median M_i of subgroup i defined by

$$\mathrm{MD}_i = \sum_{j=1}^{n} |X_{ij} - M_i|/n.$$

An unbiased estimator of σ is $\overline{\mathrm{MD}}/t_2(n)$, where $t_2(n)$ equals $E(\overline{\mathrm{MD}}/\sigma)$. Since it is difficult to obtain $E(\overline{\mathrm{MD}})$ analytically, it is obtained by simulation. Extensive tables for $t_2(n)$ can be found in Riaz and Saghir (2009). The advantage of this estimator is that it is less sensitive to outliers than R (cf. Schoonhoven et al. 2011). The resulting estimator is denoted by $\overline{\mathrm{MD}}^s$. The values used for the Phase I control limits and the constants necessary to obtain unbiased estimates from the screened data are given in Table 2. Both are obtained by simulation.

For subgroup control charts, only adaptive trimming methods based on the subgroup averages or subgroup standard deviations have been proposed in the literature so far. For diffuse outliers, however, an individuals chart should detect outliers more quickly. We therefore propose a screening method based on an individuals chart. The algorithm first calculates the residuals by subtracting the subgroup median from each observation in the corresponding subgroup. This ensures that the variability is measured within and not between subgroups. Next, an individuals chart of the residuals is constructed. The location of the chart (μ) is estimated by the mean of the subgroup medians, which is zero because the subgroup medians have been subtracted from the observations, and σ is estimated by $\overline{\mathrm{MD}}$. For simplicity, the factors for the individuals chart are 3 and –3 (see Table 2). The residuals that fall outside the control limits are excluded from the dataset. Then the procedure is repeated: the median values of the adjusted subgroups are determined, the residuals are calculated, and the control limits of the individuals chart are computed. The residuals that now exceed the limits are removed. This continues until all residuals fall within the control limits. Simulation revealed that the resulting estimates of σ are slightly biased under normality. The constants necessary to obtain an unbiased estimate can be found in Table 2. The unbiased estimator is denoted by $\overline{\mathrm{MD}}^i$.

The above procedure does not use the spread of the subgroups. Therefore, we finally propose an algorithm that combines the use of an individuals chart with subgroup screening. First, an initial estimate of σ is obtained via $\overline{\mathrm{MD}}$ (see (5)). This estimate is then used to construct a standard deviation control chart so that the subgroups can be screened. Adopting R as a charting statistic will result in the exclusion of many subgroups, including many uncontaminated observations, when diffuse disturbances are present. For this reason, we employ IQR for screening purposes. The constants required to obtain an unbiased estimate of σ based on IQR are 0.594 for $n = 4$, 0.990 for $n = 5$, and 1.144 for $n = 9$. The values chosen for the Phase I control limits are presented in Table 2. The subgroup screening is continued until all IQRs fall within the limits. The resulting estimates of σ are

Table 2. Factors for Phase I control limits for $k = 20, 50, 100$

	$n = 4$			$n = 5$			$n = 9$		
Chart	U_n	L_n	Constant	U_n	L_n	Constant	U_n	L_n	Constant
$\bar{R}^s$	2.321	0.170	1	2.305	0.172	1	1.950	0.330	1
$\overline{\mathrm{MD}}^s$	2.321	0.170	0.998	2.305	0.172	1	1.950	0.330	1
$\overline{\mathrm{MD}}^i$	3	–3	0.990	3	–3	0.975	3	–3	0.986
$\overline{\mathrm{MD}}^{i,s}$	4.703	0.0018	1	3.225	0.035	1	2.485	0.142	1
	3	–3	0.988	3	–3	0.975	3	–3	0.986

Table 3. Melt index measurements

Sample	Observations				$R/d_2(4)$	$S/c_4(4)$	$\text{MD}/t_2(4)$	$\text{IQR}/d_{\text{IQR}}(4)$
1	218	224	220	231	6.31	6.23	6.46	6.73
2	238	236	247	234	6.31	6.23	5.70	3.37
3	280	228	228	221	28.65	29.70	22.42	0
4	210	249	241	246	18.94	19.51	16.72	8.42
5	243	240	230	230	6.31	7.33	8.74	16.84
6	225	250	258	244	16.03	15.26	14.82	10.10
7	240	238	240	243	2.43	2.24	1.90	0
8	244	248	265	234	15.06	14.02	13.30	6.73
9	238	233	252	243	9.23	8.80	9.12	8.42
10	228	238	220	230	8.74	8.03	7.60	3.37
11	218	232	230	226	6.80	6.72	6.84	6.73
12	226	231	236	242	7.77	7.43	7.98	8.42
13	224	221	230	222	4.37	4.38	4.18	3.37
14	230	220	227	226	4.86	4.55	4.18	1.68
15	224	228	226	240	7.77	7.80	6.84	3.37
16	232	240	241	232	4.37	5.35	6.46	13.47
17	243	250	248	250	3.40	3.59	3.42	3.37
18	247	238	244	230	8.26	8.14	8.74	10.10
19	224	228	228	246	10.68	10.69	8.36	0
20	236	230	230	232	2.91	3.07	3.04	3.37

unbiased and are used to screen observations with an individuals control chart (the procedure used to derive $\overline{\text{MD}}^i$). Simulation revealed that the final estimates of σ are slightly biased. The constants necessary to obtain an unbiased estimate can be found in Table 2. The unbiased estimator is denoted by $\overline{\text{MD}}^{i,s}$.

2.2 Real Data Example

In this section, we demonstrate the estimation of σ in Phase I. Our dataset was supplied by Wadsworth, Stephens, and Godfrey (2001, pp. 235–237). The operation concerns the melt index of a polyethylene compound. The data consist of 20 subgroups of size 4 (Table 3).

The factors used for the $n = 4$, $k = 20$ case are presented in Table 2. Note that $d_2(4) = 2.06$, $c_4(4) = 0.92$, $t_2(4) = 0.66$, and $d_{\text{IQR}}(4) = 0.59$. The estimates of σ obtained by $\tilde{S}$ and $D7$ are determined in one iteration and are 10.14 and 6.59, respectively. The values obtained by $\bar{R}^s$ and $\overline{\text{MD}}^s$ incorporate subgroup screening. The initial value of $\bar{R}^s$ is 8.96 and the respective upper and lower control limits are 20.80 and 1.52. The unbiased estimate of the range (i.e., $\bar{R}/d_2(4)$) of subgroup 3 falls above the control limit and so this subgroup is deleted. The second estimate of $\bar{R}^s$ equals 7.92 and the corresponding Phase I upper and lower control limits are 18.38 and 1.35. Now subgroup 4 does not meet the Phase I upper control limit and is removed. The third estimate of $\bar{R}^s$ is 7.31 and the control limits are 16.97 and 1.24. There are no further subgroups whose $R/d_2(4)$ exceeds the upper control limit. The resulting unbiased estimate of σ is 7.31. The $\overline{\text{MD}}^s$ procedure works in a similar way. In this case, subgroups 3 and 4 are again deleted. The final unbiased estimate is 7.03.

For the $\overline{\text{MD}}^i$ chart, we use a procedure based on the individuals control chart for the residuals. The residuals are calculated by subtracting the subgroup median from each observation in the corresponding subgroup (see Table 4). The initial value of σ is 8.26 and the control limits of the individuals chart are 24.78 and –24.78. One residual in subgroup 3 and one residual in subgroup 4 fall outside the control limits. The corresponding observations are deleted from the dataset. The subgroup medians are determined from the remaining observations and the residuals are recalculated. The second estimate of σ is 6.82 and the control limits are now 20.47 and –20.47. One residual in subgroup 6 falls below the lower control limit and so one observation is removed. Again, the medians are determined from the remaining observations and the residuals are recomputed. The third estimate of σ is 6.49 and the control chart has limits at 19.47 and –19.47. There are now no residuals that fall outside the control limits. The resulting unbiased estimate is 6.55.

Table 4. Residuals of melt index measurements

Sample	Residuals			
1	−4.0	2.0	−2.0	9.0
2	1.0	−1.0	10.0	−3.0
3	52.0	0.0	0.0	−7.0
4	−33.5	5.5	−2.5	2.5
5	8.0	5.0	−5.0	−5.0
6	−22.0	3.0	11.0	−3.0
7	0.0	−2.0	0.0	3.0
8	−2.0	2.0	19.0	−12.0
9	−2.5	−7.5	11.5	2.5
10	−1.0	9.0	−9.0	1.0
11	−10.0	4.0	2.0	−2.0
12	−7.5	−2.5	2.5	8.5
13	1.0	−2.0	7.0	−1.0
14	3.5	−6.5	0.5	−0.5
15	−3.0	1.0	−1.0	13.0
16	−4.0	4.0	5.0	−4.0
17	−6.0	1.0	−1.0	1.0
18	6.0	−3.0	3.0	−11.0
19	−4.0	0.0	0.0	18.0
20	5.0	−1.0	−1.0	1.0

Table 5. Summary of estimates of σ and data deletions

Chart	$\hat{\sigma}$	Deleted subgroup	Deleted observation
$\tilde{S}$	10.14		
$D7$	6.59		
$\bar{R}^s$	7.31	3; 4	
$\overline{\mathrm{MD}}^s$	7.03	3; 4	
$\overline{\mathrm{MD}}^i$	6.55		3:1; 4:1; 6:1
$\overline{\mathrm{MD}}^{i,s}$	6.87	3; 7; 19	4:1; 6:1

For the $\overline{\mathrm{MD}}^{i,s}$ chart, the first part of the procedure screens the subgroup IQR. The respective upper and lower control limits of the IQR chart are 38.86 and 0.015. The IQR of subgroups 3, 7, and 19 are 0 and so these subgroups are deleted. It is not necessary to delete any further subgroups. Next, individual observations are screened. The estimate of σ is 7.81 and the upper and lower control limits for the residuals are 23.45 and –23.45. An outlier in subgroup 4 is deleted. The next estimate of σ is 7.18 with corresponding control limits 21.55 and –21.55. The outlier in subgroup 6 is removed from the dataset. Now σ is set at 6.79 with corresponding control limits 20.37 and –20.37. No further deletions are required. The unbiased estimate for the $\overline{\mathrm{MD}}^{i,s}$ chart is 6.87.

The final estimates for σ as well as the data deletions are presented in Table 5. The estimate based on $\tilde{S}$ is higher than the other estimates. This is because $\tilde{S}$ is more sensitive to outliers than the other estimators. Note, however, that the question of which estimator has done the best job cannot be resolved from such a limited dataset.

2.3 Efficiency of the Proposed Estimators

To evaluate Phase I performance, we now assess the MSE of the proposed Phase I estimators. The MSE is estimated as

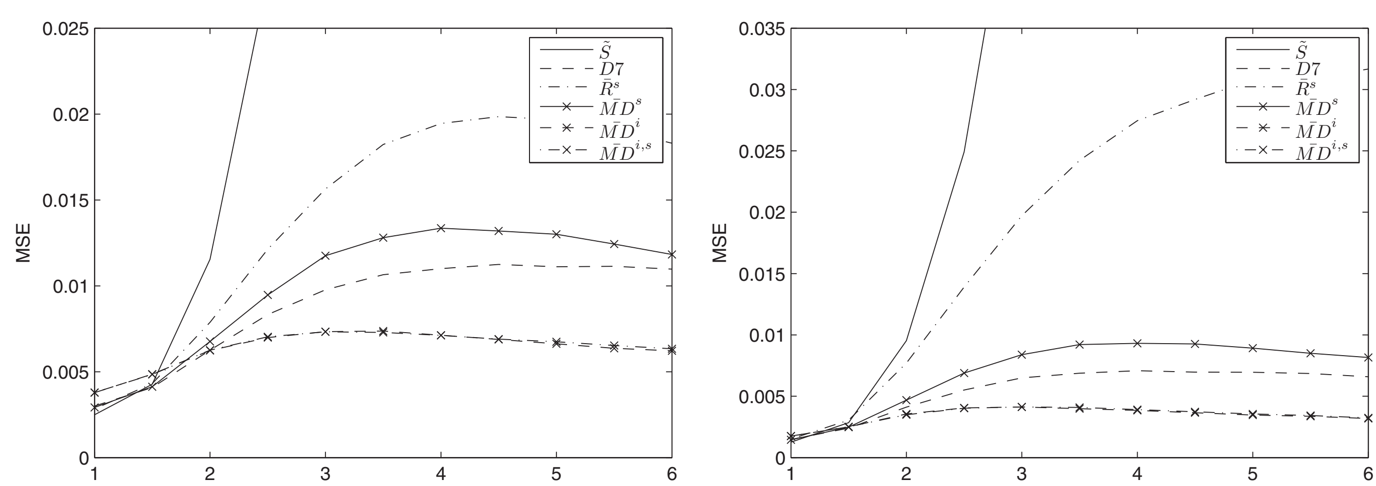
$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\sigma}^i - \sigma)^2, \tag{6}$$

where $\hat{\sigma}^i$ is the value of the unbiased estimate in the ith simulation run and N is the number of simulation runs. Comparisons are made under normality and four types of disturbances (cf. Tatum 1997), but with an error rate of 6% in each case. In general, we expect that a higher error rate would result in more pronounced differences between the estimators. The four disturbances are captured in:

1. A model for diffuse symmetric disturbances in which each observation has a 94% probability of being drawn from the $N(0, 1)$ distribution and a 6% probability of being drawn from the $N(0, a)$ distribution, with $a = 1.5, 2.0, \ldots, 5.5, 6.0$.
2. A model for diffuse asymmetric variance disturbances in which each observation is drawn from the $N(0, 1)$ distribution and has a 6% probability of having a multiple of a χ^2_1 variable added to it, with the multiplier equal to $0.5, 1.0, \ldots, 4.5, 5.0$.
3. A model for localized variance disturbances in which all observations in three (when $k = 50$) or six (when $k = 100$) subgroups are drawn from the $N(0, a)$ distribution, with $a = 1.5, 2.0, \ldots, 5.5, 6.0$.
4. A model for diffuse mean disturbances in which each observation has a 94% probability of being drawn from the $N(0, 1)$ distribution and a 6% probability of being drawn from the $N(b, 1)$ distribution, with $b = 0.5, 1.0, \ldots, 9.0, 9.5$.

The MSE is obtained for $n = 5, 9$ and $k = 50, 100$. The number of simulation runs N is equal to 50,000. Note that Tatum (1997) used 10,000 simulation runs. Below we only present the results for $k = 50$ because the conclusions for $k = 100$ are very similar. The figures comparing $k = 50$ and $k = 100$ are available as Supplementary Material.

Figure 1 shows the MSE values when diffuse symmetric variance disturbances are present. The y-intercepts show that the pooled standard deviation ($\tilde{S}$) has the lowest MSE when no disturbances are present. However, when the size of the disturbance (a) increases, the MSE increases quickly. The other estimators are more robust against outliers of this type. Those that use an



(a) (b)

Figure 1. MSE of estimators when symmetric diffuse variance disturbances are present: (a) $n = 5, k = 50$ and (b) $n = 9, k = 50$.

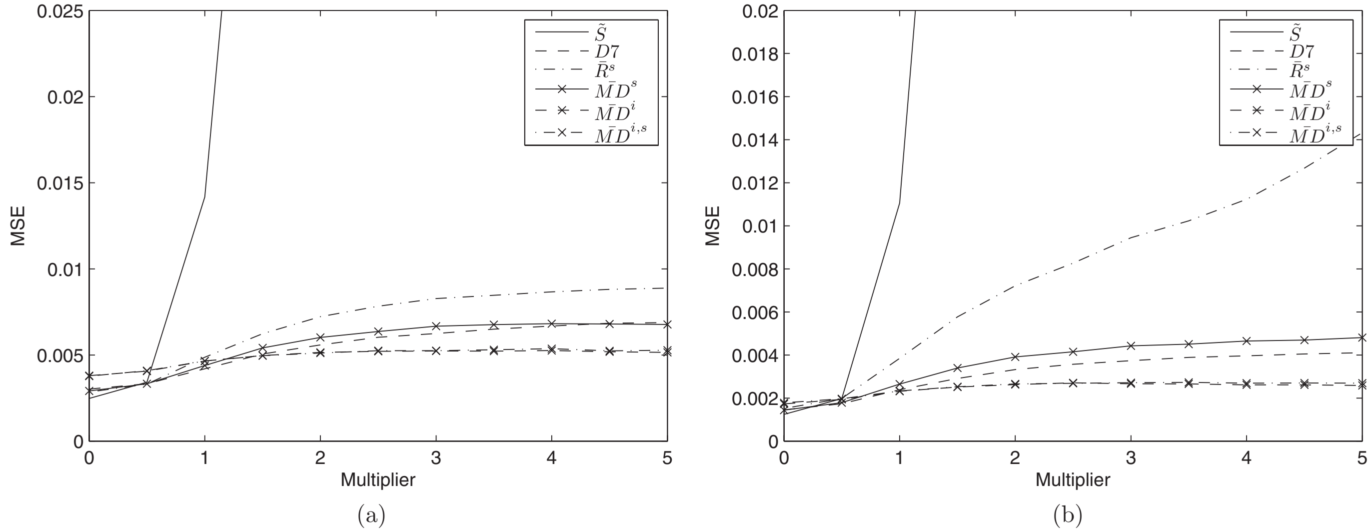


Figure 2. MSE of estimators when asymmetric diffuse variance disturbances are present: (a) $n = 5$, $k = 50$ and (b) $n = 9$, $k = 50$.

individuals control chart to identify individual outliers, that are, $\overline{\text{MD}}^i$ and $\overline{\text{MD}}^{i,s}$, coincide and perform best, followed by $D7$. The estimators based on only subgroup screening, namely $\bar{R}^s$ and $\overline{\text{MD}}^s$, turn out to perform less well in this situation. The reason is that they screen subgroup dispersion and ignore individual outliers. Note that $\bar{R}^s$ falls far short of $\overline{\text{MD}}^s$, because $\bar{R}^s$ uses $\bar{R}$ (rather than $\overline{\text{MD}}$) to estimate σ. As $\bar{R}$ is more sensitive to outliers, the Phase I limits are broader, making it more difficult to detect outliers. This effect is particularly significant for $n = 9$, because a larger subgroup is more likely to be infected with an outlier.

When asymmetric diffuse disturbances are present (Figure 2), the results are comparable to the situation with diffuse symmetric disturbances: $\overline{\text{MD}}^i$ and $\overline{\text{MD}}^{i,s}$ coincide and perform best, followed by $D7$ and $\overline{\text{MD}}^s$. Note that in this situation, $\tilde{S}$ and, for $n = 9$, $\bar{R}$ perform badly.

Figure 3 shows the results in situations with localized disturbances. The estimators incorporating subgroup screening ($\bar{R}^s$ and $\overline{\text{MD}}^s$) perform best. The estimator $\overline{\text{MD}}^{i,s}$ performs better than $D7$ in this situation. Finally, $\overline{\text{MD}}^i$ does not perform as well in this case because it does not take into account information on the subgroup spread.

The results for the fourth type of disturbance are shown in Figure 4. We can conclude that $\tilde{S}$ and $\bar{R}^s$ coincide for $n = 9$ and perform far worse than the other estimators. $\overline{\text{MD}}^s$ performs better but not as well as $D7$ and not as well as the estimators using an individuals chart to identify individual outliers. The reason is that $\overline{\text{MD}}^s$ is less capable of detecting such outliers. The estimators $\overline{\text{MD}}^{i,s}$ and $\overline{\text{MD}}^i$ coincide and perform best in this situation.

Out main conclusion from the above results is that the estimator $\overline{\text{MD}}^{i,s}$ performs best overall.

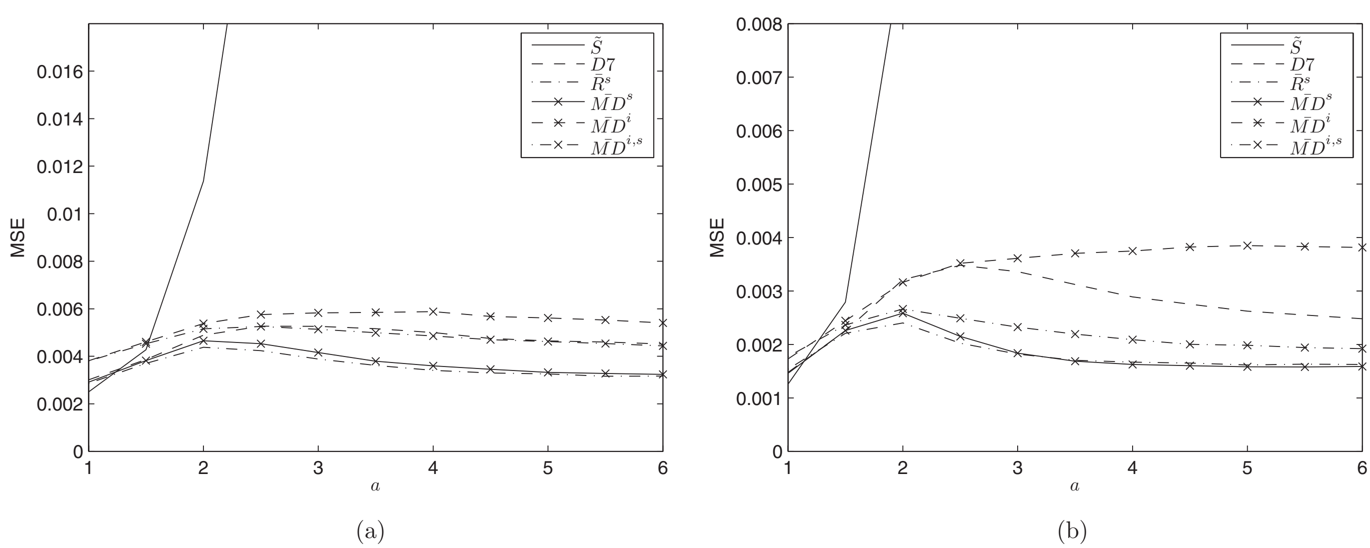
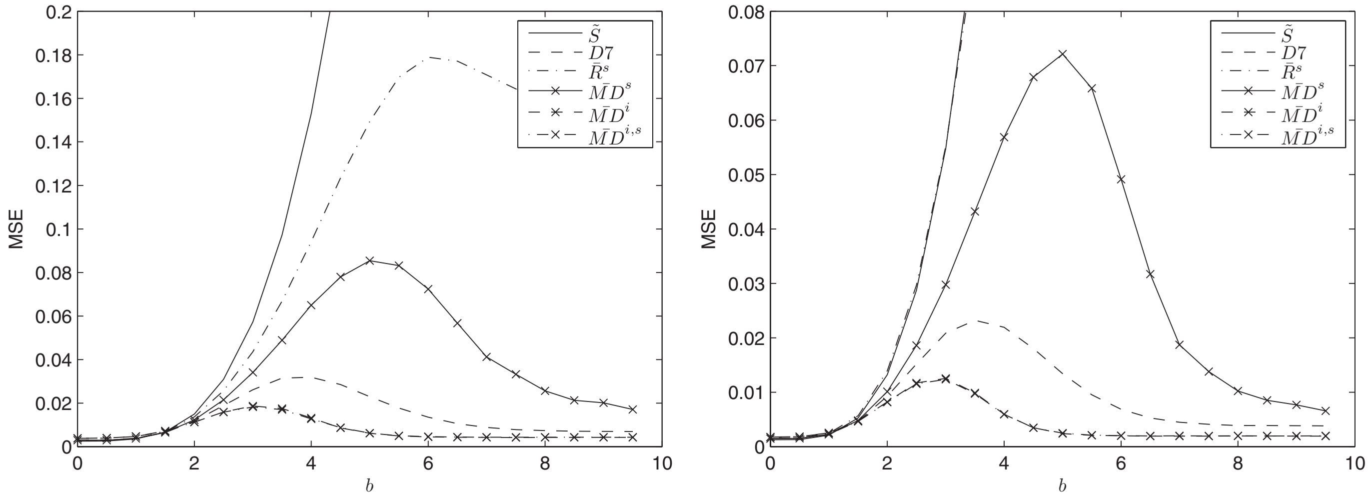


Figure 3. MSE of estimators when localized variance disturbances are present: (a) $n = 5$, $k = 50$ and (b) $n = 9$, $k = 50$.



(a) (b)

Figure 4. MSE of estimators when diffuse mean disturbances are present: (a) $n = 5$, $k = 50$ and (b) $n = 9$, $k = 50$.

3. DERIVATION OF THE PHASE II CONTROL LIMITS

Equation (1) gives control limits for the standard deviation control chart with σ estimated in Phase I. We estimate $\lambda\sigma$ in Phase II by $S/c_4(n)$ for all charts. One of the criteria used to assess Phase II performance is the ARL. To allow comparison, U_n and L_n are chosen such that the unconditional ARL equals 370 and, for each chart, the ARLs for the upper and lower control limits are similar. U_n and L_n cannot be obtained easily in analytic form and are obtained from 50,000 simulation runs. Table 6 presents U_n and L_n for $n = 5, 9$ and $k = 50, 100$.

4. CONTROL CHART PERFORMANCE

We now evaluate the effect of the proposed estimators on the Phase II performance of the standard deviation control chart. We consider the same Phase I estimators as those used to asses the MSE with a, b and the multiplier equal to 4 to simulate the contaminated cases (see Section 2.3).

Table 6. Factors U_n and L_n to determine Phase II control limits for $n = 5$ and $n = 9$

		$k = 50$		$k = 100$	
n	$\hat{\sigma}$	U_n	L_n	U_n	L_n
5	$\tilde{S}$	2.230	0.163	2.236	0.169
	$D7$	2.225	0.162	2.236	0.167
	$\bar{R}^s$	2.226	0.163	2.236	0.168
	$\overline{\text{MD}}^s$	2.226	0.162	2.237	0.167
	$\overline{\text{MD}}^i$	2.217	0.160	2.333	0.166
	$\overline{\text{MD}}^{i,s}$	2.217	0.160	2.333	0.166
9	$\tilde{S}$	1.832	0.343	1.835	0.347
	$D7$	1.830	0.341	1.834	0.346
	$\bar{R}^s$	1.831	0.342	1.835	0.347
	$\overline{\text{MD}}^s$	1.830	0.342	1.835	0.346
	$\overline{\text{MD}}^i$	1.829	0.341	1.833	0.346
	$\overline{\text{MD}}^{i,s}$	1.829	0.341	1.833	0.346

The performance of the control charts is assessed in terms of the unconditional ARL and SDRL. We compute these run length characteristics in an in-control situation and several out-of-control situations. We consider different shifts in the standard deviation $\lambda\sigma$, setting λ equal to 0.6, 1, 1.2, and 1.4. The performance characteristics are obtained by simulation. The next section describes the simulation method, followed by the results for the control charts constructed in the uncontaminated situation and various contaminated situations.

4.1 Simulation Procedure

For each Phase I dataset of k subgroups of size n, we determine $\hat{\sigma}$ and the control limits $\widehat{\text{UCL}}$ and $\widehat{\text{LCL}}$. Let $\hat{\sigma}_i$ be an estimate of $\lambda\sigma$ based on the ith subgroup Y_{ij}, $j = 1, 2, \ldots, n$. Further, let F_i denote the event that $\hat{\sigma}_i$ is above $\widehat{\text{UCL}}$ or below $\widehat{\text{LCL}}$. We define $P(F_i|\hat{\sigma})$ as the probability that subgroup i generates a signal given $\hat{\sigma}$, that is,

$$P(F_i|\hat{\sigma}) = P\big(\hat{\sigma}_i < \widehat{\text{LCL}} \text{ or } \hat{\sigma}_i > \widehat{\text{UCL}}|\hat{\sigma}\big).$$

Given $\hat{\sigma}$, the distribution of the run length is geometric with parameter $P(F_i|\hat{\sigma})$. Consequently, the conditional ARL is given by

$$E(\text{RL}|\hat{\sigma}) = \frac{1}{P(F_i|\hat{\sigma})}.$$

When we take the expectation over the X_{ij}'s we get the unconditional ARL

$$\text{ARL} = E\frac{1}{P(F_i|\hat{\sigma})}.$$

This expectation is obtained by simulation: numerous datasets are generated from the normal distribution or contaminated normal distribution, and for each dataset, $E(\text{RL}|\hat{\sigma})$ is computed. By averaging these values, we obtain the unconditional value. The unconditional standard deviation is

Table 7. ARL and SDRL under normality for $k = 50$

		ARL				SDRL			
n	Chart	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$
5	$\tilde{S}$	131	378	69.5	17.5	136	412	87.0	19.6
	$D7$	135	373	69.6	17.4	141	414	90.0	20.0
	$\bar{R}^s$	132	369	68.9	17.3	137	407	88.4	19.7
	$\overline{\text{MD}}^s$	135	375	69.7	17.4	141	414	90.0	20.0
	$\overline{\text{MD}}^i$	143	371	70.0	17.3	151	421	94.6	20.6
	$\overline{\text{MD}}^{i,s}$	143	371	69.9	17.4	151	421	95.0	20.7
9	$\tilde{S}$	28.4	371	43.6	9.02	29.3	392	51.6	9.30
	$D7$	29.5	373	43.9	9.06	30.8	397	53.5	9.50
	$\bar{R}^s$	29.2	392	43.9	9.05	30.5	392	53.9	9.52
	$\overline{\text{MD}}^s$	29.1	369	43.9	9.02	30.3	393	53.4	9.45
	$\overline{\text{MD}}^i$	29.8	368	44.7	9.11	31.4	395	56.0	9.71
	$\overline{\text{MD}}^{i,s}$	29.9	366	44.4	9.05	31.6	394	56.4	9.65

determined by:

$$\text{SDRL} = \sqrt{\text{var(RL)}} = \sqrt{E(\text{var}(\text{RL}|\hat{\sigma})) + \text{var}(E(\text{RL}|\hat{\sigma}))}$$
$$= \sqrt{2E\left(\frac{1}{p(F_i|\hat{\sigma})}\right)^2 - \left(E\frac{1}{p(F_i|\hat{\sigma})}\right)^2 - E\frac{1}{p(F_i|\hat{\sigma})}}.$$

Enough replications of the above procedure were performed to obtain sufficiently small relative estimated standard errors for ARL. The relative standard error never exceeds 0.76%.

4.2 Simulation Results

The ARL and SDRL are obtained in the in-control situation ($\lambda = 1$) and in the out-of-control situation ($\lambda \neq 1$). When the process is in control, we want the ARL and SDRL to be close to their intended values, namely 370. In the out-of-control

Table 8. ARL and SDRL when contaminations are present in Phase I for $k = 50$

			ARL				SDRL			
	n	Chart	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$
$N(0, 1)$	5	$\tilde{S}$	43.9	297	425	303	50.8	337	452	391
&		$D7$	102	484	159	34.7	108	498	215	47.2
$N(0, 4)$		$\bar{R}^s$	92.2	464	206	50.1	99.9	479	286	87.2
(symm)		$\overline{\text{MD}}^s$	101	474	171	38.2	108	492	171	58.6
		$\overline{\text{MD}}^i$	123	443	114	25.3	131	479	164	34.2
		$\overline{\text{MD}}^{i,s}$	122	446	114	25.6	131	481	164	34.6
	9	$\tilde{S}$	5.47	114	317	286	5.92	148	354	351
		$D7$	19.9	433	109	17.0	20.8	440	144	20.1
		$\bar{R}^s$	14.3	336	238	47.5	16.2	368	311	94.5
		$\overline{\text{MD}}^s$	18.9	410	128	19.6	20.3	421	178	26.3
		$\overline{\text{MD}}^i$	23.8	420	75.2	12.9	25.3	433	102	15.0
		$\overline{\text{MD}}^{i,s}$	23.7	421	75.9	13.0	25.2	434	103	15.0
$N(0, 1)$	5	$\tilde{S}$	23.2	149	231	266	38.4	243	325	355
&		$D7$	112	461	121	26.9	118	483	164	34.2
$N(0, 4)$		$\bar{R}^s$	108	450	133	29.9	115	472	191	43.6
(asymm)		$\overline{\text{MD}}^s$	115	449	119	26.6	121	475	168	35.7
		$\overline{\text{MD}}^i$	130	419	92.7	21.7	138	461	131	27.5
		$\overline{\text{MD}}^{i,s}$	130	422	95.0	21.9	138	463	135	27.9
	9	$\tilde{S}$	2.34	33.3	93.6	165	3.27	79.5	186	263
		$D7$	22.7	435	80.0	13.5	23.7	443	104	15.2
		$\bar{R}^s$	18.8	399	140	22.6	20.6	415	207	39.7
		$\overline{\text{MD}}^s$	22.5	418	83.3	14.0	23.9	429	116	16.7
		$\overline{\text{MD}}^i$	26.0	407	61.1	11.2	27.5	425	80.5	12.6
		$\overline{\text{MD}}^{i,s}$	25.7	409	61.7	11.3	27.3	426	81.4	12.7

Table 9. ARL and SDRL when contaminations are present in Phase I for $k = 50$

			ARL				SDRL			
	n	Chart	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1.0$	$\lambda = 1.2$	$\lambda = 1.4$
$N(0, 1)$	5	$\tilde{S}$	42.8	293	436	308	48.3	330	459	390
&		$D7$	118	442	103	23.5	124	469	139	28.7
$N(0, 4)$		$\bar{R}^s$	129	384	77.1	18.7	134	422	103	22.2
(loc)		$\overline{\mathrm{MD}}^s$	131	391	78.5	19.0	137	431	106	22.9
		$\overline{\mathrm{MD}}^i$	127	428	100	22.9	136	468	143	29.5
		$\overline{\mathrm{MD}}^{i,s}$	135	404	85.9	20.4	144	451	122	25.7
	9	$\tilde{S}$	5.34	110	324	295	5.55	136	357	254
		$D7$	24.3	427	67.9	12.1	25.4	438	86.9	13.3
		$\bar{R}^s$	28.8	372	46.0	9.31	30.2	396	57.8	9.93
		$\overline{\mathrm{MD}}^s$	28.6	372	46.0	9.31	30.0	396	57.5	9.91
		$\overline{\mathrm{MD}}^i$	23.9	420	74.4	12.8	25.5	433	100	14.7
		$\overline{\mathrm{MD}}^{i,s}$	28.9	377	49.1	9.68	30.7	404	64.4	10.6
$N(0, 1)$	5	$\tilde{S}$	40.2	280	470	335	42.8	300	480	395
&		$D7$	80.3	471	286	72.8	86.7	483	359	117
$N(4, 1)$		$\bar{R}^s$	54.1	358	431	207	61.1	384	466	292
		$\overline{\mathrm{MD}}^s$	65.0	409	394	144	73.0	431	450	220
		$\overline{\mathrm{MD}}^i$	115	447	152	64.4	127	481	237	64.4
		$\overline{\mathrm{MD}}^{i,s}$	115	449	152	63.5	127	483	234	63.5
	9	$\tilde{S}$	5.04	101	325	321	4.87	114	348	363
		$D7$	14.4	358	230	35.1	15.3	380	291	51.4
		$\bar{R}^s$	5.59	117	341	292	5.84	144	367	353
		$\overline{\mathrm{MD}}^s$	9.53	231	370	99.0	10.4	266	404	151
		$\overline{\mathrm{MD}}^i$	22.8	409	93.6	15.2	25.0	425	142	20.9
		$\overline{\mathrm{MD}}^{i,s}$	22.8	410	92.1	15.2	25.0	425	140	21.0

situation, we want to detect changes in the standard deviation as soon as possible, so the ARL should be as low as possible. The results are obtained for $n = 5, 9$ and $k = 50, 100$. However, as was done for the MSE comparison, we only present the results for $k = 50$ because the conclusions for $k = 100$ are not all that different. The tables comparing $k = 50$ and $k = 100$ are available as Supplementary Material.

Table 7 shows the ARL and SDRL for the situation when the Phase I data are uncontaminated and normally distributed. The ARL is very similar across charts and the SDRL is slightly higher for the $\overline{\mathrm{MD}}^i$ and $\overline{\mathrm{MD}}^{i,s}$ charts.

Tables 8 and 9 show that when there are disturbances in the Phase I data, the ARL values increase (decrease) considerably for $\lambda > 1$ ($\lambda < 1$) relative to the normal situation. Thus, when the Phase I data are contaminated, changes in the process standard deviation are less likely to be detected when $\lambda > 1$, while there are more signals when $\lambda < 1$. With diffuse disturbances (Table 8 and second half of Table 9), their impact is smallest for the charts based on $\overline{\mathrm{MD}}^i$, $\overline{\mathrm{MD}}^{i,s}$, and $D7$. When there are localized disturbances (first half of Table 9), the charts based on $\bar{R}^s$, $\overline{\mathrm{MD}}^s$, and $\overline{\mathrm{MD}}^{i,s}$ perform best, because these charts trim extreme subgroups. Note that in a number of cases, the $\tilde{S}$, $\bar{R}^s$, and $\overline{\mathrm{MD}}^s$ charts are ARL-biased: the in-control ARL is lower than the out-of-control ARL (cf. Jensen et al. 2006).

Overall, the $\overline{\mathrm{MD}}^{i,s}$ chart performs best. Under normality, this chart almost matches the standard chart based on $\tilde{S}$, and in the presence of any contamination, the chart outperforms the alternatives.

5. CONCLUDING REMARKS

In this article, we consider several estimators of the standard deviation in Phase I of the control charting process. We have found that the performance of certain robust estimators is almost identical to the pooled subgroup standard deviation under normality, while the benefit of using such robust estimators can be substantial when there are disturbances. Following Rocke (1989, 1992), we have considered estimators that include a procedure for subgroup screening, but whereas Rocke used $\bar{R}$, we have used the average deviation from the median. This estimator performs better when there are localized disturbances and is much more robust against diffuse disturbances. However, when there are diffuse mean disturbances, the procedure loses efficiency.

To address this problem, we have proposed other algorithms based on a procedure that also screens for individual outliers. The algorithms remove the variation between subgroups so that only the variation within subgroups is measured. We have shown that these algorithms are very effective when there are diffuse disturbances. When there might also be localized disturbances, the method can be combined with subgroup screening based on the IQR. The latter procedure reveals a performance very similar to the robust estimator for the standard deviation control chart proposed by Tatum (1997). We think that this is a noteworthy outcome since the procedure is simple and intuitive. Moreover, it can be used to estimate σ in other practical applications.

SUPPLEMENTARY MATERIAL

Figures S1– S4 and Tables S1– S3.

ACKNOWLEDGMENTS

The authors are very grateful to the Editor, the Associate Editor, and a referee for their helpful comments, which led to improvements in this article.

[Received November 2010. Revised October 2011.]

REFERENCES

Chen, G. (1998), "The Run Length Distributions of the R, S and s^2 Control Charts When σ Is Estimated," *Canadian Journal of Statistics*, 26, 311–322. [73]

Davis, C. M., and Adams, B. M. (2005), "Robust Monitoring of Contaminated Data," *Journal of Quality Technology*, 37, 163–174. [73]

Duncan, A.J. (1974), *Quality Control and Industrial Statistics* (4th ed.), Homewood, IL: R.D. Irwin. [75]

Hillier, F. (1969), "$\overline{X}$ and R Chart Control Limits Based on a Small Number of Subgroups," *Journal of Quality Technology*, 1, 17–26. [73]

Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall, W. H. (2006), "Effects of Parameter Estimation on Control Chart Properties: A Literature Review," *Journal of Quality Technology*, 38, 349–364. [73,74,81]

Mahmoud, M. A., Henderson, G. R., Epprecht, E. K., and Woodall, W. H. (2010), "Estimating the Standard Deviation in Quality-Control Applications," *Journal of Quality Technology*, 42, 348–357. [74]

Quesenberry, C. P. (1993), "The Effect of Sample Size on Estimated Limits for $\overline{X}$ and X Control Charts," *Journal of Quality Technology*, 25, 237–247. [73]

Riaz, M., and Saghir, A. (2009), "A Mean Deviation-Based Approach to Monitor Process Variability," *Journal of Statistical Computation and Simulation*, 79, 1173–1193. [75]

Rocke, D. M. (1989), "Robust Control Charts," *Technometrics*, 31, 173–184. [73,74,81]

—— (1992), "$\bar{X}_Q$ and R_Q Charts: Robust Control Charts," *The Statistician*, 41, 97–104. [73,81]

Schoonhoven, M., Riaz, M., and Does, R. J. M. M. (2011), "Design and Analysis of Control Charts for Standard Deviation With Estimated Paramaters," *Journal of Quality Technology*, 44, 307–333. [74,75]

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*, Princeton, NJ: Van Nostrand. [73]

Tatum, L. G. (1997), "Robust Estimation of the Process Standard Deviation for Control Charts," *Technometrics*, 39, 127–141. [73,74,77,81]

Wadsworth, H. M., Stephens, K. S., and Godfrey, A. B. (2001), *Modern Methods for Quality Control and Improvement* (2nd ed.), New York: Wiley. [76]

Woodall, W.H., and Montgomery, D.C. (1999), "Research Issues and Ideas in Statistical Process Control," *Journal of Quality Technology*, 31, 376–386. [73]

Yang, C., and Hillier, F. (1970), "Mean and Variance Control Chart Limits Based on a Small Number of Subgroups," *Journal of Quality Technology*, 2, 9–16. [73]