Estimation methods for statistical process control

Schoonhoven, M.

Publication date
2011

Chapter 3

A Robust Standard Deviation Control Chart

3.1 Introduction

In this chapter 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 (see Chapter 2). For diffuse outliers, we think that an individuals chart would detect outliers more quickly. We therefore include an estimator based on the individuals chart. In order to measure the variability within and not between subgroups, we correct for differences in the location between subgroups. Moreover, we present an algorithm that combines the last two approaches. The performance of the estimators is evaluated by assessing their MSE under normality and in the presence of several types of contaminations. Finally, 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 chapter 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, summarize simulation results and draw some final conclusions.

3.2 Proposed Phase I estimators

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.

3.2.1 Standard deviation estimators

Recall that X_{ij}, $i = 1, 2, ..., k$ and $j = 1, 2, ..., n$, denote 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 $\tilde{S}$ (see (2.1)). 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.

Second, we evaluate a robust estimator proposed by Tatum (1997) (see (2.11)). This estimator is denoted by $D7$ as in Tatum (1997).

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

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

where R_i is the range of the i-th 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 (1986), 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 $\widehat{UCL} = U_n\bar{R}/d_2(n)$ and $\widehat{LCL} = L_n\bar{R}/d_2(n)$. For simplicity, 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 3.1 shows the factors for $n = 4, 5, 9$ as well as the constants added to obtain unbiased estimates once the data have been screened. Note that the factors and the constants are the same for $k = 20, 50, 100$. The resulting estimator is denoted by $\overline{R}^s$.

In addition, we evaluate an adaptive trimmer where the estimate of σ is obtained by the mean subgroup average deviation from the median instead

	$n=4$			$n=5$			$n=9$		
$\hat{\sigma}$	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
$\overline{MD}^s$	2.321	0.170	0.998	2.305	0.172	1	1.950	0.330	1
$\overline{MD}^i$	3	-3	0.990	3	-3	0.975	3	-3	0.986
$\overline{MD}^{i,s}$	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 3.1: Factors for Phase I control limits for $k = 20, 50, 100$

of $\bar{R}$. The mean subgroup average deviation from the median is given by

$$\overline{MD} = \frac{1}{k}\sum_{i=1}^{k} MD_i,$$

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

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

An unbiased estimator of σ is $\overline{MD}/t_2(n)$, where $t_2(n)$ equals $E(\overline{MD}/\sigma)$. Since it is difficult to obtain $E(\overline{MD})$ 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 (see Chapter 2). The resulting estimator is denoted by $\overline{MD}^s$. The values used for the Phase I control limits and the constants necessary to obtain unbiased estimates once the data have been screened are given in Table 3.1.

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 the standard deviation σ is estimated by $\overline{MD}$. To simplify, the factors for the individuals chart are 3 and -3 (see Table 3.1). 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. The resulting estimates of σ are slightly biased under normality. The constants necessary to obtain an unbiased estimate can be found in Table 3.1. The unbiased estimator is denoted by $\overline{MD}^i$.

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 $\overline{MD}$. 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 3.1. The subgroup screening is continued until all IQR's fall within the limits. The resulting estimates of σ are unbiased and are used to screen observations with an individuals control chart (the procedure used to derive $\overline{MD}^i$). The final estimates of σ are slightly biased. The constants necessary to obtain an unbiased estimate can be found in Table 3.1. The unbiased estimator is denoted by $\overline{MD}^{i,s}$.

3.2.2 Real data example

In this section we demonstrate the estimation of σ in Phase I. Our dataset was supplied by Wadsworth et al. (2001, pp. 235-237). The operation concerns the melt index of a polyethylene compound. The data consist of 20 subgroups of size 4 (Table 3.2).

The factors used for the $n = 4, k = 20$ case are presented in Table 3.1. Note that $d_2(4) = 2.06$, $c_4(4) = 0.92$, $t_2(4) = 0.66$ and $d_{IQR}(4) = 0.59$. 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 $\overline{MD}^s$ 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

Sample	Observations				$R/d_2(4)$	$S/c_4(4)$	$MD/t_2(4)$	$IQR/d_{IQR}(4)$
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

Table 3.2: Melt index measurements

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 $\overline{MD}^s$ 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 $\overline{MD}^i$ chart, we use a procedure based on the individuals control chart for the residuals. The residuals are calculated by substracting the subgroup median from each observation in the corresponding subgroup (see Table 3.3). 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

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 3.3: Residuals of melt index measurements

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.

For the $\overline{MD}^{i,s}$ 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 $\overline{MD}^{i,s}$ chart is 6.87.

The final estimates for σ are presented in Table 3.4. 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.

Chart	$\hat{\sigma}$
$\tilde{S}$	10.14
$D7$	6.59
$\bar{R}^s$	7.31
$\overline{MD}^s$	7.03
$\overline{MD}^i$	6.55
$\overline{MD}^{i,s}$	6.87

Table 3.4: Final estimates of σ

3.2.3 Efficiency of proposed estimators

In order to evaluate Phase I performance we now assess the MSE of the proposed Phase I estimators. Recall that the MSE is estimated as

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

where $\hat{\sigma}^i$ is the value of the unbiased estimate in the i-th simulation run and N is the number of simulation runs. Comparisons are made under normality and four types of disturbances (cf. Tatum (1997), see Section 2.2.2) but with an error rate of 6% in each case. The number of simulation runs N is equal to 50,000. (Note that Tatum (1997) used 10,000 simulation runs.)

Figure 3.1 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 individuals control chart to identify individual outliers, i.e. $\overline{MD}^i$ and $\overline{MD}^{i,s}$, coincide and perform best, followed by $D7$. The estimators based on only subgroup screening, namely $\bar{R}^s$ and $\overline{MD}^s$, 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 $\overline{MD}^s$, because $\bar{R}^s$ uses $\bar{R}$ (rather than $\overline{MD}$) 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 3.2), the results are comparable to the situation with diffuse symmetric disturbances: $\overline{MD}^i$ and $\overline{MD}^{i,s}$ coincide and perform best, followed by $D7$ and $\overline{MD}^s$. Note that in this situation $\tilde{S}$ and, for $n = 9$, $\bar{R}$ perform badly.

Figure 3.3 shows the results in situations with localized disturbances. The estimators incorporating subgroup screening ($\bar{R}^s$ and $\overline{MD}^s$) perform best. The estimator $\overline{MD}^{i,s}$ performs better than $D7$ in this situation. Finally, $\overline{MD}^i$ does not perform as well in this case because it does not take into account information on the subgroup spread.

The results for diffuse mean disturbance are shown in Figure 3.4. We can conclude that $\tilde{S}$ and $\bar{R}^s$ coincide for $n = 9$ and perform far worse than the other estimators. $\overline{MD}^s$ 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 $\overline{MD}^s$ is less capable of detecting such outliers. The estimators $\overline{MD}^{i,s}$ and $\overline{MD}^i$ coincide and perform best in this situation.

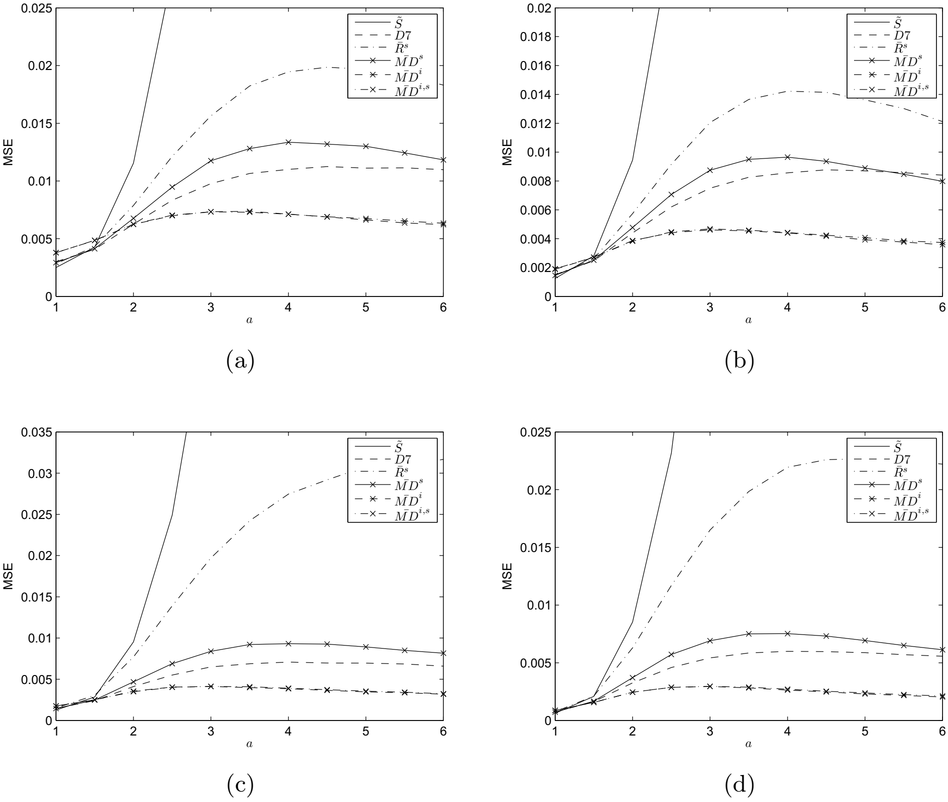

Figure 3.1: MSE of estimators when symmetric diffuse variance disturbances are present. (a) $n = 5, k = 50$ (b) $n = 5, k = 100$ (c) $n = 9, k = 50$ (d) $n = 9, k = 100$

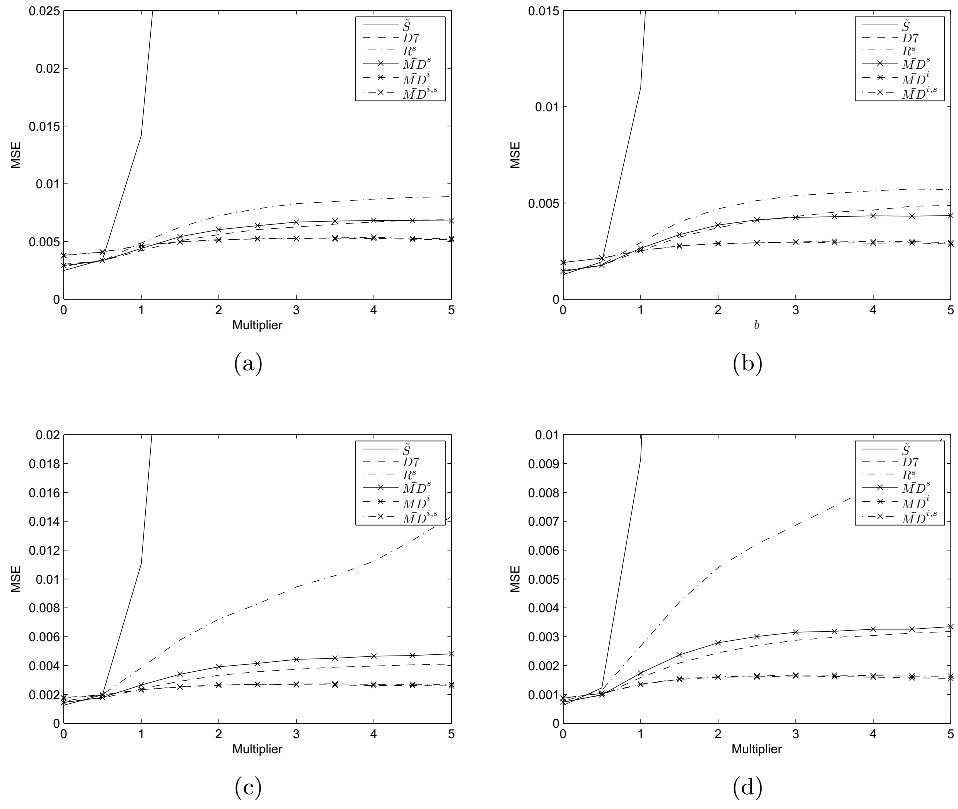

Figure 3.2: MSE of estimators when asymmetric diffuse variance disturbances are present. (a) $n = 5, k = 50$ (b) $n = 5, k = 100$ (c) $n = 9, k = 50$ (d) $n = 9, k = 100$

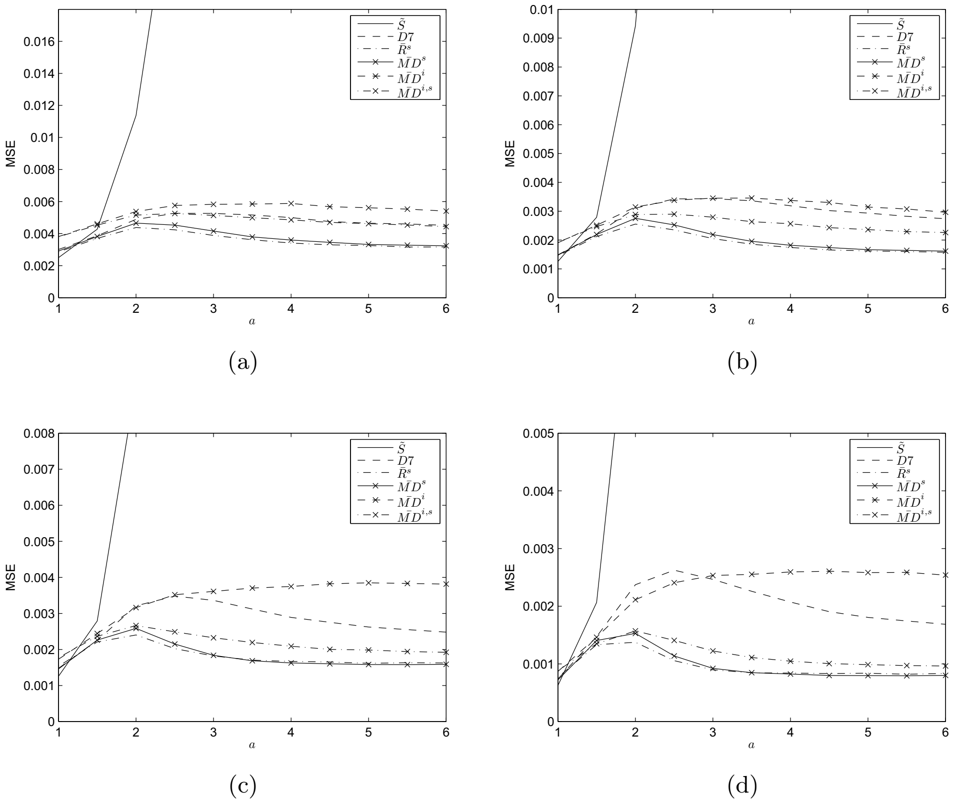

Figure 3.3: MSE of estimators when localized variance disturbances are present. (a) $n = 5, k = 50$ (b) $n = 5, k = 100$ (c) $n = 9, k = 50$ (d) $n = 9, k = 100$

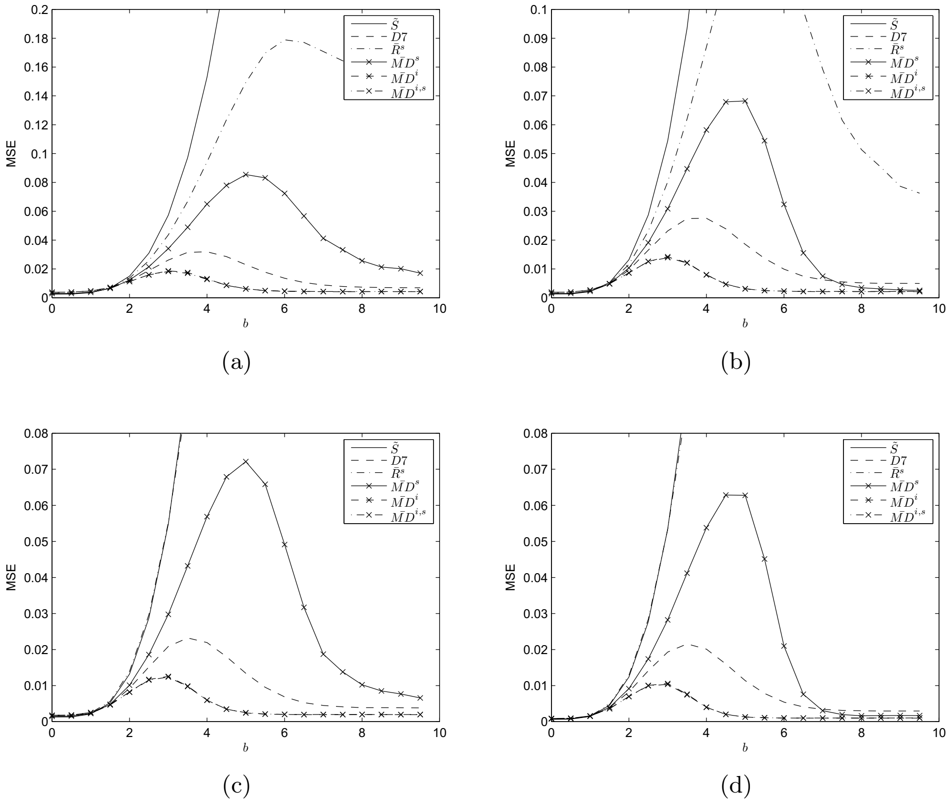

Figure 3.4: MSE of estimators when diffuse mean disturbances are present. (a) $n = 5, k = 50$ (b) $n = 5, k = 100$ (c) $n = 9, k = 50$ (d) $n = 9, k = 100$

3.3 Derivation of Phase II control limits

Equation (1.3) 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 ARL's for the upper and lower control limits are similar. U_n and L_n can not be obtained easily in analytic form and are obtained from 50,000 simulation runs. Table 3.5 presents U_n and L_n for $n = 5, 9$ and $k = 50, 100$.

		$k = 50$		$k = 100$	
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
	$\overline{MD}^s$	2.226	0.162	2.237	0.167
	$\overline{MD}^i$	2.217	0.160	2.333	0.166
	$\overline{MD}^{i,s}$	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
	$\overline{MD}^s$	1.830	0.342	1.835	0.346
	$\overline{MD}^i$	1.829	0.341	1.833	0.346
	$\overline{MD}^{i,s}$	1.829	0.341	1.833	0.346

Table 3.5: Factors U_n and L_n to determine Phase II control limits for $n = 5, 9$ and $k = 50, 100$

3.4 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 2.2.2).

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 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 two sections describe the simulation method and the results for the control charts constructed in the uncontaminated and contaminated situations.

3.4.1 Simulation procedure

We use the same procedure as in Section 2.4.1. Enough replications of the procedure were performed to obtain sufficiently small relative estimated standard errors for the ARL. The relative standard error never exceeds 0.76%.

3.4.2 Simulation results

The ARL and SDRL are obtained in the in-control situation and in the out-of-control situation. When the process is in control, i.e. when $\lambda = 1$, we want the ARL and SDRL to be close to their intended values, namely 370. In the out-of-control situation, i.e. when $\lambda \neq 1$, we want to detect changes in the standard deviation as soon as possible, so the ARL should be as low as possible.

Table 3.6 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 $\overline{MD}^i$ and $\overline{MD}^{i,s}$ charts.

Tables 3.7 and 3.8 show that, when there are disturbances in the Phase I data, the ARL values increase considerably. Thus, when the Phase I data are contaminated, changes in the process standard deviation are less likely to be detected. With diffuse disturbances (Table 3.7 and second half of Table 3.8), their impact is smallest for the charts based on $\overline{MD}^i$, $\overline{MD}^{i,s}$ and $D7$. When there are localized disturbances (second half of Table 3.8), the charts based on $\overline{R}^s$, $\overline{MD}^s$ and $\overline{MD}^{i,s}$ perform best, because these charts trim extreme subgroups. Note that in a number of cases the $\tilde{S}$, $\overline{R}^s$ and $\overline{MD}^s$ charts are ARL-biased: the in-control ARL is lower than the out-of-control ARL (cf. Jensen et al. (2006)).

Overall, the $\overline{MD}^{i,s}$ 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.

			ARL				*SDRL*			
n	k	Chart	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$
5	50	$\tilde{S}$	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
		$\overline{MD}^s$	135	375	69.7	17.4	141	414	90.0	20.0
		$\overline{MD}^i$	143	371	70.0	17.3	151	421	94.6	20.6
		$\overline{MD}^{i,s}$	143	371	69.9	17.4	151	421	95.0	20.7
	100	$\tilde{S}$	113	366	66.4	17.0	115	382	74.3	17.8
		$D7$	119	374	67.4	17.2	121	394	77.3	18.2
		$\bar{R}^s$	116	368	66.8	17.0	118	387	76.0	17.9
		$\overline{MD}^s$	119	375	67.3	17.3	121	394	76.8	18.2
		$\overline{MD}^i$	121	372	67.6	17.2	125	397	80.0	18.6
		$\overline{MD}^{i,s}$	122	371	67.8	17.2	125	396	80.2	18.5
9	50	$\tilde{S}$	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
		$\overline{MD}^s$	29.1	369	43.9	9.02	30.3	393	53.4	9.45
		$\overline{MD}^i$	29.8	368	44.7	9.11	31.4	395	56.0	9.71
		$\overline{MD}^{i,s}$	29.9	366	44.4	9.05	31.6	394	56.4	9.65
	100	$\tilde{S}$	26.0	373	42.2	8.90	26.2	382	45.7	8.77
		$D7$	26.4	375	42.8	8.98	26.7	386	47.1	8.92
		$\bar{R}^s$	26.3	369	42.3	8.93	26.6	380	46.8	8.89
		$\overline{MD}^s$	26.6	375	42.6	8.96	26.8	386	42.6	8.91
		$\overline{MD}^i$	26.7	369	42.2	8.92	27.1	382	47.3	8.94
		$\overline{MD}^{i,s}$	26.7	367	42.3	8.90	27.2	380	47.5	8.92

Table 3.6: ARL and SDRL under normality

				ARL				*SDRL*			
	n	k	Chart	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$
$N(0,1)$	5	50	$\tilde{S}$	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
(sym.)			$\overline{MD}^s$	101	474	171	38.2	108	492	171	58.6
			$\overline{MD}^i$	123	443	114	25.3	131	479	164	34.2
			$\overline{MD}^{i,s}$	122	446	114	25.6	131	481	164	34.6
		100	$\tilde{S}$	35.9	248	413	308	38.5	269	425	362
			$D7$	89.7	476	152	32.8	91.4	479	182	37.8
			$\bar{R}^s$	81.4	458	191	42.0	84.6	462	239	55.2
			$\overline{MD}^s$	89.0	470	159	34.4	89.0	470	159	42.0
			$\overline{MD}^i$	104	443	107	24.3	108	458	133	28.0
			$\overline{MD}^{i,s}$	104	444	108	28.0	108	458	134	28.0
	9	50	$\tilde{S}$	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	368	238	47.5	
			$\overline{MD}^s$	18.9	410	128	19.6	20.3	421	178	26.3
			$\overline{MD}^i$	23.8	420	75.2	12.9	25.3	433	102	15.0
			$\overline{MD}^{i,s}$	23.7	421	75.9	13.0	25.2	434	103	15.0
		100	$\tilde{S}$	4.86	95.6	305	302	4.75	110	331	345
			$D7$	17.9	420	104	16.5	18.0	423	122	17.6
			$\bar{R}^s$	12.9	323	225	36.0	12.9	343	276	53.4
			$\overline{MD}^s$	17.3	409	118	18.2	17.3	414	118	18.2
			$\overline{MD}^i$	21.3	420	70.0	12.4	21.7	424	82.2	13.1
			$\overline{MD}^{i,s}$	21.2	420	70.6	12.5	21.6	424	83.2	13.2
$N(0,1)$	5	50	$\tilde{S}$	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
(asym.)			$\overline{MD}^s$	115	449	119	26.6	121	475	168	35.7
			$\overline{MD}^i$	130	419	92.7	21.7	138	461	131	27.5
			$\overline{MD}^{i,s}$	130	422	95.0	21.9	138	463	135	27.9
		100	$\tilde{S}$	14.9	98.0	183	262	21.3	147	253	262
			$D7$	98.0	458	116	25.9	100	466	138	28.8
			$\bar{R}^s$	95.2	447	122	27.3	97.9	455	151	31.9
			$\overline{MD}^s$	101	448	111	25.1	103	458	111	28.5
			$\overline{MD}^i$	111	419	88.9	21.1	114	438	109	23.5
			$\overline{MD}^{i,s}$	110	421	89.8	21.2	114	440	109	23.7
	9	50	$\tilde{S}$	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
			$\overline{MD}^s$	22.5	418	83.3	14.0	23.9	429	116	16.7
			$\overline{MD}^i$	26.0	407	61.1	11.2	27.5	425	80.5	12.6
			$\overline{MD}^{i,s}$	25.7	409	61.7	11.3	27.3	426	81.4	12.7
		100	$\tilde{S}$	1.79	18.8	59.5	140	1.72	34.5	112	220
			$D7$	20.4	429	75.9	13.2	20.5	431	87.2	13.7
			$\bar{R}^s$	17.1	396	123	18.9	17.6	403	159	22.9
			$\overline{MD}^s$	20.6	424	77.4	13.4	20.6	428	92.7	14.3
			$\overline{MD}^i$	23.2	408	57.4	10.9	23.6	415	66.3	11.2
			$\overline{MD}^{i,s}$	23.1	411	67.4	11.4	23.4	417	67.4	11.4

Table 3.7: ARL and SDRL when contaminations are present in Phase I

				ARL				*SDRL*			
	n	k	Chart	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$	$\lambda = 0.6$	$\lambda = 1$	$\lambda = 1.2$	$\lambda = 1.4$
$N(0,1)$	5	50	$\tilde{S}$	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.)			$\overline{MD}^s$	131	391	78.5	19.0	137	431	106	22.9
			$\overline{MD}^i$	127	428	100	22.9	136	468	143	29.5
			$\overline{MD}^{i,s}$	135	404	85.9	20.4	144	451	122	25.7
		100	$\tilde{S}$	35.6	245	419	312	37.7	263	429	361
			$D7$	103	441	99.1	23.0	105	451	117	25.1
			$\bar{R}^s$	113	382	73.1	18.3	113	401	85.1	19.7
			$\overline{MD}^s$	115	392	74.6	18.5	117	411	87.9	20.0
			$\overline{MD}^i$	109	429	95.0	22.1	112	446	117	25.0
			$\overline{MD}^{i,s}$	115	403	81.6	19.7	118	425	99.6	21.9
	9	50	$\tilde{S}$	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
			$\overline{MD}^s$	28.6	372	46.0	9.31	30.0	396	57.5	9.91
			$\overline{MD}^i$	23.9	420	74.4	12.8	25.5	433	100	14.7
			$\overline{MD}^{i,s}$	28.9	377	49.1	9.68	30.7	404	64.4	10.6
		100	$\tilde{S}$	4.80	93.8	306	309	4.59	105	329	347
			$D7$	21.8	425	65.3	11.9	22.0	428	74.0	12.2
			$\bar{R}^s$	25.9	372	43.9	9.12	26.2	384	49.1	9.14
			$\overline{MD}^s$	26.2	380	44.1	9.19	26.5	391	49.0	9.20
			$\overline{MD}^i$	21.4	420	69.4	12.4	21.8	424	81.6	13.0
			$\overline{MD}^{i,s}$	25.8	379	46.1	9.42	26.3	391	52.8	9.59
$N(0,1)$	5	50	$\tilde{S}$	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
			$\overline{MD}^s$	65.0	409	394	144	73.0	431	450	220
			$\overline{MD}^i$	115	447	152	64.4	127	481	237	64.4
			$\overline{MD}^{i,s}$	115	449	152	63.5	127	483	234	63.5
		100	$\tilde{S}$	34.4	237	434	328	35.1	246	439	360
			$D7$	70.0	448	282	65.2	72.5	454	327	85.4
			$\bar{R}^s$	46.6	320	436	193	49.2	336	450	248
			$\overline{MD}^s$	55.9	378	399	129	58.9	392	429	173
			$\overline{MD}^i$	98.3	451	136	30.2	103	464	183	39.7
			$\overline{MD}^{i,s}$	98.0	451	136	30.2	103	464	182	39.4
	9	50	$\tilde{S}$	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
			$\overline{MD}^s$	9.53	231	370	99.0	10.4	266	404	151
			$\overline{MD}^i$	22.8	409	93.6	15.2	25.0	425	142	20.9
			$\overline{MD}^{i,s}$	22.8	410	92.1	15.2	25.0	425	140	21.0
		100	$\tilde{S}$	4.67	89.6	302	327	4.29	94.8	316	351
			$D7$	12.9	330	225	32.2	13.0	343	262	38.7
			$\bar{R}^s$	5.00	99.7	322	294	4.77	110	338	332
			$\overline{MD}^s$	8.59	207	387	88.6	8.64	226	403	117
			$\overline{MD}^i$	20.4	413	82.2	13.9	21.0	418	105	15.6
			$\overline{MD}^{i,s}$	20.4	414	81.7	13.8	21.0	418	104	15.4

Table 3.8: ARL and SDRL when contaminations are present in Phase I

3.5 Concluding remarks

In this chapter 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 variance 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.