



## Correction factors for Shewhart and control charts to achieve desired unconditional ARL

Rob Goedhart, Marit Schoonhoven & Ronald J.M.M. Does

To cite this article: Rob Goedhart, Marit Schoonhoven & Ronald J.M.M. Does (2016) Correction factors for Shewhart and control charts to achieve desired unconditional ARL, International Journal of Production Research, 54:24, 7464-7479, DOI: [10.1080/00207543.2016.1193251](https://doi.org/10.1080/00207543.2016.1193251)

To link to this article: <http://dx.doi.org/10.1080/00207543.2016.1193251>



Published online: 05 Jun 2016.



Submit your article to this journal [↗](#)



Article views: 73



View related articles [↗](#)



View Crossmark data [↗](#)

## Correction factors for Shewhart $X$ and $\bar{X}$ control charts to achieve desired unconditional ARL

Rob Goedhart\*, Marit Schoonhoven and Ronald J.M.M. Does

*Department of Operations Management, Institute for Business and Industrial Statistics, University of Amsterdam, Amsterdam, The Netherlands*

*(Received 22 December 2015; accepted 15 May 2016)*

In this paper we derive correction factors for Shewhart control charts that monitor individual observations as well as subgroup averages. In practice, the distribution parameters of the process characteristic of interest are unknown and, therefore, have to be estimated. A well-known performance measure within Statistical Process Monitoring is the expectation of the average run length (ARL), defined as the unconditional ARL. A practitioner may want to design a control chart such that, in the in-control situation, it has a certain expected ARL. However, accurate correction factors that lead to such an unconditional ARL are not yet available. We derive correction factors that guarantee a certain unconditional in-control ARL. We use approximations to derive the factors and show their accuracy and the performance of the control charts – based on the new factors – in out-of-control situations. We also evaluate the variation between the ARLs of the individually estimated control charts.

**Keywords:** average run length; parameter estimation; process control; process monitoring; SPC; statistical methods

### 1. Introduction

Shewhart  $X$  and  $\bar{X}$  control charts are common tools to monitor process means. In practice, the true process parameters are often unknown, and need to be estimated using a Phase I reference sample. However, since different practitioners have different Phase I data, the estimates of the process parameters will vary across practitioners. Because of this the estimated control limits, and consequently the control chart performance, are actually random variables. This variation in control chart performance has been addressed by several researchers, and is of great interest in current SPC literature. An overview of this literature is given by [Jensen et al. \(2006\)](#) and [Psarakis, Vyniou, and Castagliola \(2014\)](#).

The performance of a control chart is generally measured in terms of the false alarm rate (FAR) or the average run length (ARL). However, for example as shown by [Quesenberry \(1993\)](#) and [Chen \(1997\)](#), the amount of Phase I data required for the control chart to perform properly is substantially larger than initially thought. Although increasing the amount of Phase I data generally improves the accuracy of the parameter estimates, the required amounts are not always feasible in practice. Therefore, another solution is to apply corrections to the control chart limits to take parameter estimation into account. For example, [Hillier \(1969\)](#), [Does and Schriever \(1992\)](#) and [Schoonhoven, Riaz, and Does \(2009\)](#) provide corrections leading to a desired value of the FAR in expectation. Next to that, more recent works (e.g. [Albers and Kallenberg 2004, 2005](#); [Chakraborti 2006](#); [Faraz, Woodall, and Heuchenne 2015](#); [Saleh et al. 2015](#)) also consider other performance measures such as the ARL.

Performance of control charts based on estimated parameters can be evaluated for a given set of parameter values (known as conditional performance), or in expectation (known as unconditional performance). In this paper we derive new corrections for the Shewhart  $X$  and  $\bar{X}$  control charts to achieve a desired unconditional ARL. Although [Albers and Kallenberg \(2005\)](#) derived correction factors for the same objective function, they have only considered the individuals chart. Therefore, we extend their approach for the Shewhart  $\bar{X}$  control chart. In addition, we derive new correction factors and compare their performance. It will be shown that our newly proposed correction factors outperform those of [Albers and Kallenberg \(2005\)](#).

The next section discusses the considered model and notation. In Section 3, the performance of the correction factors of [Albers and Kallenberg \(2005\)](#) is assessed. Next, in Section 4 we describe the derivation of the newly proposed correction factors, and compare their performance. In Section 5, we discuss the implications of using the proposed corrections in in-control and out-of-control situations. In Section 6, the implementation of the control chart is illustrated with a real-life example. Finally, in Section 7 we provide concluding remarks.

---

\*Corresponding author. Email: [R.Goedhart2@uva.nl](mailto:R.Goedhart2@uva.nl)

## 2. Model and approach

The Shewhart control chart for monitoring the mean with known parameters has control limits

$$\text{UCL} = \mu + K\sigma/\sqrt{n}, \quad \text{LCL} = \mu - K\sigma/\sqrt{n}, \quad (1)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the process characteristic of interest, respectively, and where  $K$  is the constant used to achieve the desired probability of a false signal ( $\alpha$ ). Let  $Y_{ij}$  denote the  $j$ th observation in sample  $i$  ( $i = 1, 2, \dots$  and  $j = 1, 2, \dots, n$ ) and let  $Y_{ij}$  be i.i.d.  $N(\mu + \delta\sigma/\sqrt{n}, \sigma)$  random variables, where  $\delta = 0$  corresponds to the in-control situation, and where the process is considered out-of-control for  $\delta \neq 0$ . Then,  $K = \Phi^{-1}(1 - \alpha/2)$  with  $\Phi^{-1}$  the inverse of the standard normal cumulative distribution function. The run length (RL), i.e. the number of samples before the control chart gives a signal, has a geometric distribution with parameter  $\alpha$ . As a consequence, the in-control ARL is given by  $1/\alpha$ .

In practice,  $\mu$  and  $\sigma$  are not known and therefore have to be estimated. To this end,  $m$  samples of  $n$  measurements on the process characteristic are collected when the process is considered to be in control. Let  $X_{ij}$  (the  $j$ th observation in the  $i$ th subgroup) be i.i.d.  $N(\mu, \sigma)$  random variables ( $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ ). Note that the  $X_{ij}$ 's correspond to Phase I, and the  $Y_{ij}$ 's correspond to Phase II.

For the subgroup control chart,  $\mu$  is usually estimated by the grand sample mean

$$\bar{\bar{X}} = \frac{1}{m} \sum_{i=1}^m \left( \frac{1}{n} \sum_{j=1}^n X_{ij} \right), \quad (2)$$

and  $\sigma$  is usually estimated by the pooled standard deviation

$$\tilde{S} = \left( \frac{1}{m} \sum_{i=1}^m S_i^2 \right)^{1/2}, \quad (3)$$

where  $S_i$  is the  $i$ th sample 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 of  $\sigma$  is  $\tilde{S}/c_4(m(n-1)+1)$  (see e.g. [Schoonhoven and Does \(2012\)](#) or [Saleh et al. \(2015\)](#)), where  $c_4(k)$  is defined by

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

For the individuals chart ( $n = 1$ ),  $\mu$  is usually estimated by the sample average

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad (4)$$

where  $X_i$  is the  $i$ th observation ( $i = 1, 2, \dots, m$ ). When  $n = 1$ ,  $\sigma$  is usually estimated by the average moving range

$$\bar{\text{MR}} = \frac{1}{m-1} \sum_{i=1}^{m-1} |X_{i+1} - X_i|. \quad (5)$$

An unbiased estimator of  $\sigma$  is  $\bar{\text{MR}}/d_2(2)$ , where  $d_2(2) = \frac{2}{\sqrt{\pi}}$ . Hence we use the most common estimators of  $\sigma$  in both situations. However, it is possible to use any other desired estimator in our approach.

The performance of a control chart design based on estimated parameters can be evaluated for a given set of parameter values, known as conditional performance, or in expectation, known as unconditional performance. The conditional ARL (henceforth denoted as CARL) of a control chart design can be determined as follows. We define  $E_i$  ( $i = 1, 2, \dots$ ) as the event that  $\bar{Y}_i$  falls outside the control limits. Then

$$P(E_i | \hat{\mu}, \hat{\sigma}) = P(\bar{Y}_i < \widehat{\text{LCL}}) + P(\bar{Y}_i > \widehat{\text{UCL}}), \quad (6)$$

where  $\widehat{\text{LCL}}$  and  $\widehat{\text{UCL}}$  are the control limits as defined in (1) but with  $\mu$  and  $\sigma$  replaced by their unbiased estimates  $\hat{\mu}$  and  $\hat{\sigma}$ , respectively. Given a pair of control limits, the events  $E_i$  are independent so that the conditional RL is geometrically

distributed with parameter  $P(E_i|\hat{\mu}, \hat{\sigma})$ . The CARL is then given by  $1/P(E_i|\hat{\mu}, \hat{\sigma})$  and the unconditional ARL (henceforth denoted as EARL) by  $E(1/P(E_i|\hat{\mu}, \hat{\sigma}))$ . The unconditional probability of a signal can be determined by  $E(P(E_i|\hat{\mu}, \hat{\sigma}))$ .

For the Shewhart  $X$  and  $\bar{X}$  control charts with estimated parameters, we propose control limits of the form

$$\widehat{UCL} = \hat{\mu} + \tilde{K}\hat{\sigma}/\sqrt{n}, \quad \widehat{LCL} = \hat{\mu} - \tilde{K}\hat{\sigma}/\sqrt{n}, \quad (7)$$

with  $\widehat{UCL}$  and  $\widehat{LCL}$  the respective upper and lower control chart limits based on the unbiased estimators  $\hat{\mu}$  and  $\hat{\sigma}$  of  $\mu$  and  $\sigma$ , respectively, and  $\tilde{K} = K + c$  the factor used to achieve a desired in-control EARL, which is denoted by  $ARL_0$ .

### 3. Albers and Kallenberg's correction factors

In this section, we assess the performance of the correction factors derived by [Albers and Kallenberg \(2005\)](#). First, we extend their work to subgroup control charts and then, based on simulations, we assess the accuracy of the proposed correction factors. The final subsection gives an overview of the accuracy of the correction factors.

#### 3.1 Derivation of the correction factors

[Albers and Kallenberg \(2005\)](#) derived correction factors for the individuals  $X$  chart in a multiplicative form of  $K(1 + c)$  rather than the additive form  $K + c$  that we use. However, their corrections can easily be transformed into a comparable form since  $K(1 + c) = K + Kc = K + c_{AK}$ . Hence, we can simply multiply their proposed corrections  $c$  by  $K$  to make a comparison. For the individuals  $X$  chart with  $n = 1$ , their proposed correction  $c_{AK}$  is equal to

$$c_{AK} = -\frac{K + K^3\tau^2}{2m}, \quad (8)$$

where  $\tau^2 = \lim_{m \rightarrow \infty} \left[ m \text{var} \left( \frac{\hat{\sigma}}{E\hat{\sigma}} \right) \right]$ . Although [Albers and Kallenberg \(2005\)](#) considered estimators of  $\sigma$  based on grouped observations (such as the pooled standard deviation), they did not derive corrections for the Shewhart  $\bar{X}$  chart when  $n > 1$ . However, following their derivations (see Appendix A.1) we arrive at the following corrections

$$c_{AK} = -\frac{nK + K^3\tau^2}{2mn} \quad (9)$$

where  $\tau^2 = \lim_{mn \rightarrow \infty} \left[ mn \text{var} \left( \frac{\hat{\sigma}}{E\hat{\sigma}} \right) \right]$ . Note that this more general correction is equal to their proposed correction for  $n = 1$ . The resulting correction factors are given in Table 1.

#### 3.2 Simulation procedure

The accuracy of the correction factor  $c_{AK}$  is assessed for several combinations of parameter values, namely  $n \in \{1, 3, 5, 7\}$ ,  $m \in \{20, 30, 40, 50, 75, 100\}$  and  $\alpha \in \{0.001, 0.0027, 0.005, 0.01\}$ .

For each combination of parameter values, we calculate the EARL as follows. Without loss of generality, we generate numerous Phase I data-sets (at least 1,000,000 data-sets consisting of  $m$  samples of size  $n$ ) from a  $N(0, 1)$  distribution in order to obtain the CARL values. For each Phase I data-set, we determine the control limits  $\widehat{UCL}$  and  $\widehat{LCL}$  according to (7). Let  $\bar{Y}_i$  be the average of the  $i$ th subgroup in Phase II ( $i = m + 1, m + 2, \dots$ ) and let  $E_i$  denote the event that  $\bar{Y}_i$  falls above  $\widehat{UCL}$  or below  $\widehat{LCL}$ . We define  $P(E_i|\hat{\mu}, \hat{\sigma})$  as the probability that subgroup  $i$  generates a signal conditional on  $\hat{\mu}$  and  $\hat{\sigma}$ , i.e.

$$P(E_i|\hat{\mu}, \hat{\sigma}) = P(\bar{Y}_i < \widehat{LCL}) + P(\bar{Y}_i > \widehat{UCL}).$$

Conditional on  $\hat{\mu}$  and  $\hat{\sigma}$ , the distribution of the RL is geometric with parameter  $P(E_i|\hat{\mu}, \hat{\sigma})$ . Hence, the CARL is given by

$$\text{CARL} = E(RL|\hat{\mu}, \hat{\sigma}) = \frac{1}{P(E_i|\hat{\mu}, \hat{\sigma})}. \quad (10)$$

When we take the expectation over all Phase I samples, we obtain the EARL

$$\text{EARL} = E \left( \frac{1}{P(E_i|\hat{\mu}, \hat{\sigma})} \right).$$

This expectation is obtained by simulation: for each Phase I data-set,  $E(RL|\hat{\mu}, \hat{\sigma})$  is computed. The number of Phase I data-sets was at least 1,000,000 for each parameter combination. For a few combinations, additional data-sets are generated so that the maximum relative standard error of the EARL is below 5% for  $n = 1$ , and below 1% for  $n > 1$  for all cases.

Table 1. Correction factors  $c_{AK}$  for various values of  $\alpha$ ,  $n$  and  $m$ .

$n$	$m$	ARL <sub>0</sub> = 1000	ARL <sub>0</sub> = 370	ARL <sub>0</sub> = 200	ARL <sub>0</sub> = 100
		$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	-0.8180	-0.6325	-0.5269	-0.4173
	30	-0.5453	-0.4217	-0.3513	-0.2782
	40	-0.4090	-0.3163	-0.2635	-0.2087
	50	-0.3272	-0.2530	-0.2108	-0.1669
	75	-0.2181	-0.1687	-0.1405	-0.1113
	100	-0.1636	-0.1265	-0.1054	-0.0835
3	20	-0.3049	-0.2437	-0.2084	-0.1712
	30	-0.2033	-0.1625	-0.1389	-0.1141
	40	-0.1525	-0.1219	-0.1042	-0.0856
	50	-0.1220	-0.0975	-0.0834	-0.0685
	75	-0.0813	-0.0650	-0.0556	-0.0457
	100	-0.0610	-0.0487	-0.0417	-0.0342
5	20	-0.1936	-0.1594	-0.1393	-0.1178
	30	-0.1291	-0.1062	-0.0929	-0.0785
	40	-0.0968	-0.0797	-0.0696	-0.0589
	50	-0.0774	-0.0637	-0.0557	-0.0471
	75	-0.0516	-0.0425	-0.0371	-0.0314
	100	-0.0387	-0.0319	-0.0279	-0.0236
7	20	-0.1565	-0.1312	-0.1163	-0.1000
	30	-0.1043	-0.0875	-0.0775	-0.0667
	40	-0.0782	-0.0656	-0.0581	-0.0500
	50	-0.0626	-0.0525	-0.0465	-0.0400
	75	-0.0417	-0.0350	-0.0310	-0.0267
	100	-0.0313	-0.0262	-0.0233	-0.0200

### 3.3 Accuracy of correction factors

For a wide range of parameter values, the realised values of the EARL are given in Table 2. It can be observed that these values are still quite different from the desired values  $ARL_0$ , especially for small sample sizes. This difference appears to be larger for larger values of  $ARL_0$ . For larger sample sizes, the uncorrected control limits even perform better than the corrected control limits, as the EARL of the uncorrected chart are then closer to the desired value. It is clear that the proposed correction factors by [Albers and Kallenberg \(2005\)](#) are not sufficient: the EARL still deviates heavily from its desired value. For this reason, we derive new correction factors in the next section.

## 4. New correction factors

In this section, we derive new correction factors which give a numerical example and present the accuracy of the proposed correction factors.

### 4.1 Derivation of new correction factor

The idea behind the derivation of the factor is as follows. We want to derive a correction factor  $c$  such that the EARL, given by  $E(1/P(E_i|\hat{\mu}, \hat{\sigma}))$ , equals a desired value, namely  $ARL_0 = 1/\alpha$ . In order to do this, we derive  $P(E_i|\hat{\mu}, \hat{\sigma})$  and solve  $c$  from the equation  $E(1/P(E_i|\hat{\mu}, \hat{\sigma})) = 1/\alpha$ .

Table 2. EARL of corrected (based on  $c_{AK}$ ) and, between brackets, uncorrected control charts.

$n$	$m$	ARL <sub>0</sub> = 1000	ARL <sub>0</sub> = 370	ARL <sub>0</sub> = 200	ARL <sub>0</sub> = 100
		$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	554 (1,151,800)	269 (191,160)	160 (7858)	85 (968)
	30	807 (45,575)	321 (3632)	178 (1079)	91 (326)
	40	878 (11,698)	337 (1828)	185 (589)	94 (210)
	50	914 (5458)	346 (1084)	188 (442)	95 (173)
	75	948 (2532)	355 (699)	192 (320)	97 (139)
	100	963 (1945)	359 (581)	194 (280)	98 (128)
3	20	616 (2378)	251 (648)	142 (301)	75 (134)
	30	734 (1693)	287 (520)	160 (259)	83 (117)
	40	796 (1444)	306 (472)	169 (238)	87 (112)
	50	835 (1335)	318 (447)	175 (231)	89 (110)
	75	887 (1201)	335 (415)	183 (219)	93 (106)
	100	913 (1148)	343 (405)	187 (213)	94 (105)
5	20	624 (1328)	249 (433)	141 (224)	74 (106)
	30	727 (1184)	283 (406)	158 (214)	81 (103)
	40	786 (1127)	302 (391)	167 (209)	86 (101)
	50	822 (1095)	314 (389)	173 (204)	88 (101)
	75	876 (1053)	331 (384)	181 (202)	92 (100)
	100	898 (1036)	338 (377)	184 (201)	93 (100)
7	20	624 (1102)	248 (387)	140 (200)	74 (98)
	30	724 (1055)	282 (379)	157 (199)	81 (98)
	40	781 (1038)	300 (374)	166 (200)	85 (98)
	50	819 (1018)	312 (371)	172 (199)	88 (99)
	75	867 (1005)	328 (369)	180 (197)	91 (99)
	100	897 (1004)	338 (370)	184 (199)	93 (98)

Note that the conditional probability of signaling ( $P(E_i|\hat{\mu}, \hat{\sigma})$ ) is given by

$$\begin{aligned}
P(E_i|\hat{\mu}, \hat{\sigma}) &= 1 - P(\widehat{\text{LCL}} < \bar{Y}_i < \widehat{\text{UCL}}) \\
&= 1 - \left[ P\left(\frac{\bar{Y}_i - \mu}{\sigma/\sqrt{n}} < \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}}{\sigma}\right) - P\left(\frac{\bar{Y}_i - \mu}{\sigma/\sqrt{n}} < \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}}{\sigma}\right) \right] \\
&= 1 - \left[ \Phi\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + \tilde{K} \frac{\hat{\sigma}}{\sigma}\right) - \Phi\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} - \tilde{K} \frac{\hat{\sigma}}{\sigma}\right) \right] \\
&= [1 - \Phi(\tilde{K} + \Delta_1(\tilde{K}))] + [1 - \Phi(\tilde{K} + \Delta_2(\tilde{K}))] \\
&= \bar{\Phi}(\tilde{K} + \Delta_1(\tilde{K})) + \bar{\Phi}(\tilde{K} + \Delta_2(\tilde{K}))
\end{aligned} \tag{11}$$

with  $\tilde{K} = K + c$ ,  $\bar{\Phi}(x) = 1 - \Phi(x)$  and

$$\Delta_1(\tilde{K}) = \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}}{\sigma} - 1\right)$$

and

$$\Delta_2(\tilde{K}) = -\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + \tilde{K} \left(\frac{\hat{\sigma}}{\sigma} - 1\right).$$

This notation is similar to that of [Albers and Kallenberg \(2004\)](#), who derived corrections for the one-sided individuals control chart. From (11), we can write for any function  $g$  of  $P(E_i|\hat{\mu}, \hat{\sigma})$

$$g(P(E_i|\hat{\mu}, \hat{\sigma})) = h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) = h(x, y).$$

Table 3. Derivatives of the function  $h$ , general at point  $(x, y)$  or specifically at point  $(K, K)$  for  $g(\alpha) = 1/\alpha$ .

Point	$h$	$h_x$	$h_{xx}$		$h_{xy}$
$(x, y)$	$\frac{1}{\bar{\Phi}(x) + \bar{\Phi}(y)}$	$\frac{\phi(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^2}$	$\frac{2\phi^2(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^3}$	$-\frac{x\phi(x)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^2}$	$\frac{2\phi(x)\phi(y)}{(\bar{\Phi}(x) + \bar{\Phi}(y))^3}$
$(K, K)$	$\frac{1}{2\bar{\Phi}(K)}$	$\frac{\phi(K)}{4\bar{\Phi}^2(K)}$	$\frac{\phi^2(K)}{4\bar{\Phi}^3(K)}$	$-\frac{K\phi(K)}{4\bar{\Phi}^2(K)}$	$\frac{\phi^2(K)}{4\bar{\Phi}^3(K)}$

where we use  $h(x, y)$  as general notation, which is required for the next steps in the derivation. Thus, for a specific value of  $\tilde{K}$  we can write  $h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) = h(x_0 + \Delta x, y_0 + \Delta y)$  where  $(x_0, y_0) = (\tilde{K}, \tilde{K})$  and  $(\Delta x, \Delta y) = (\Delta_1(\tilde{K}), \Delta_2(\tilde{K}))$ . Next, we approximate the probability  $P(E_i|\hat{\mu}, \hat{\sigma})$  by using a two-step Taylor expansion of  $h(x, y) = h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K}))$  around  $(x_0, y_0) = (\tilde{K}, \tilde{K})$ .

$$\begin{aligned}
 &h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\
 &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K})\Delta_1(\tilde{K}) + h_y(\tilde{K}, \tilde{K})\Delta_2(\tilde{K}) \\
 &\quad + \frac{1}{2}[h_{xx}(\tilde{K}, \tilde{K})\Delta_1^2(\tilde{K}, \tilde{K}) + 2h_{xy}(\tilde{K}, \tilde{K})\Delta_1(\tilde{K})\Delta_2(\tilde{K}) + h_{yy}(\tilde{K}, \tilde{K})\Delta_2^2(\tilde{K})], \tag{12}
 \end{aligned}$$

where  $h_x(x, y)$  denotes the first-order partial derivative with respect to  $x$ ,  $h_{xx}(x, y)$  denotes the second-order partial derivative with respect to  $x$  (and likewise for  $y$ ) and  $h_{xy}(x, y)$  denotes the cross-partial derivative with respect to  $x$  and  $y$ .

When parameters are estimated,  $Eg(P(E_i|\hat{\mu}, \hat{\sigma})) \neq g(\alpha)$ . Note that in (12) we have  $h_x(\tilde{K}, \tilde{K}) = h_y(\tilde{K}, \tilde{K})$  and  $h_{xx}(\tilde{K}, \tilde{K}) = h_{yy}(\tilde{K}, \tilde{K})$ . To ensure that  $Eg(P(E_i|\hat{\mu}, \hat{\sigma})) = g(\alpha)$ , we want to find a correction factor  $c$  such that for  $\tilde{K} = K + c$  the following holds (cf. (12)):

$$\begin{aligned}
 Eg(P(E_i|\hat{\mu}, \hat{\sigma})) &= Eh(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\
 &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) [E\Delta_1(\tilde{K}) + E\Delta_2(\tilde{K})] \\
 &\quad + \frac{1}{2}h_{xx}(\tilde{K}, \tilde{K}) [E\Delta_1^2(\tilde{K}) + E\Delta_2^2(\tilde{K})] + h_{xy}(\tilde{K}, \tilde{K})E [\Delta_1(\tilde{K})\Delta_2(\tilde{K})] \\
 &= g(\alpha) = h(K, K). \tag{13}
 \end{aligned}$$

In Appendix A.2 we show that, for unbiased estimators  $\hat{\mu}$  and  $\hat{\sigma}$ , we can rewrite this formula as

$$c = -\frac{h_{xx}(K, K)E\Delta_1^2(K) + h_{xy}(K, K)E [\Delta_1(K)\Delta_2(K)]}{2h_x(K, K)}. \tag{14}$$

The functions  $h(x, y)$ ,  $h_x(x, y)$ ,  $h_{xx}(x, y)$  and  $h_{xy}(x, y)$  depend on the function  $g(\alpha)$  that is being considered. For the ARL,  $g(\alpha) = 1/\alpha$ . The derivatives for this function at point  $(K, K)$  are given in Table 3. Note that the approach can easily be adapted for any other function  $g(\alpha)$  by adapting the corresponding derivatives. Commonly considered functions, apart from the ARL, are the FAR ( $g(\alpha) = \alpha$ ) and the probability that the RL is at most a specified value  $j$  ( $g(\alpha) = 1 - (1 - \alpha)^j$ , see e.g. Tsai et al. (2005)).

The derivations given in (11)–(14) are still in general form with regard to the used estimators. For  $\hat{\mu} = \bar{\bar{X}}$  and  $\hat{\sigma} = \tilde{S}/c_4(m(n - 1) + 1)$  (cf. (2) and (3)) we have (see Appendix A.3)

$$E [\Delta_1(K)\Delta_2(K)] \approx \frac{K^2}{2(m(n - 1) + 1)} - \frac{1}{m} \tag{15}$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx \frac{K^2}{2(m(n - 1) + 1)} + \frac{1}{m}. \tag{16}$$

Besides the application to the Shewhart  $\bar{X}$  control chart it can also be applied to the individuals control chart. Since a different estimator for the standard deviation is used in this case, namely  $\overline{MR}/d_2(2)$ , the explicit calculation of  $E [\Delta_1(K)\Delta_2(K)]$ ,  $E\Delta_1^2(K)$  and  $E\Delta_2^2(K)$  will be different here. In Appendix A.4, we show that in this case

$$E [\Delta_1(K)\Delta_2(K)] \approx K^2 \left[ \frac{0.8264m - 1.082}{(m - 1)^2} \right] - \frac{1}{m} \tag{17}$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx K^2 \left[ \frac{0.8264m - 1.082}{(m-1)^2} \right] + \frac{1}{m}. \quad (18)$$

#### 4.2 Examples of computation of the correction factors

In this section, we demonstrate how to calculate the correction factor for both the subgroup control chart and individuals control chart. In addition, we give an overview of correction factors for different values of  $\alpha$ ,  $n$  and  $m$ .

First, we give an example of the calculation of a correction factor for a subgroup control chart with  $\alpha = 0.0027$ ,  $n = 5$  and  $m = 50$ . In this case,  $K = \Phi^{-1}(1 - \alpha/2) = 3$ . We need to find the correction factor  $c$  in  $\tilde{K} = K + c$ . For the subgroup control chart,  $c$  is given by formula (14). To calculate  $c$ , we need to find  $h_{xx}(K, K)$ ,  $h_{xy}(K, K)$ ,  $h_x(K, K)$ ,  $E\Delta_1^2(K)$  and  $E[\Delta_1(K)\Delta_2(K)]$ . These are as follows.

$$\begin{aligned} h_x(K, K) &= \frac{\phi(3)}{4\bar{\Phi}^2(3)} = 608.03, \\ h_{xx}(K, K) &= \frac{\phi^2(3)}{4\bar{\Phi}^3(3)} - \frac{3\phi(3)}{4\bar{\Phi}^2(3)} = 172.13, \\ h_{xy}(K, K) &= \frac{\phi^2(3)}{4\bar{\Phi}^3(3)} = 1996.2, \\ E\Delta_1^2(K) &= \frac{3^2}{2(50(5-1)+1)} + \frac{1}{50} = 0.042, \\ E[\Delta_1(K)\Delta_2(K)] &= \frac{3^2}{2(50(5-1)+1)} - \frac{1}{50} = 0.0024. \end{aligned}$$

From this, we can derive  $c$  as

$$c = -\frac{172.13 \cdot 0.042 + 1996.2 \cdot 0.0024}{2 \cdot 608.03} = -0.0099.$$

Next, we give an example of the calculation of a correction factor for individuals control chart with  $\alpha = 0.005$ ,  $n = 1$  and  $m = 100$ . Note that in this case  $\sigma$  is estimated by  $\overline{MR}/d_2(2)$ . We now have  $K = \Phi^{-1}(1 - \alpha/2) = 2.81$ . Again, we need to find the correction factor  $c$  in  $\tilde{K} = K + c$  using (14). To calculate  $c$ , we need to find  $h_{xx}(K, K)$ ,  $h_{xy}(K, K)$ ,  $h_x(K, K)$ ,  $E\Delta_1^2(K, K)$  and  $E[\Delta_1(K)\Delta_2(K)]$ . These are as follows.

$$\begin{aligned} h_x(K, K) &= \frac{\phi(2.81)}{4\bar{\Phi}^2(2.81)} = 310.44, \\ h_{xx}(K, K) &= \frac{\phi^2(2.81)}{4\bar{\Phi}^3(2.81)} - \frac{3\phi(2.81)}{4\bar{\Phi}^2(2.81)} = 92.30, \\ h_{xy}(K, K) &= \frac{\phi^2(2.81)}{4\bar{\Phi}^3(2.81)} = 963.70, \\ E\Delta_1^2(K) &= 2.81^2 \left[ \frac{0.8264 \cdot 100 - 1.082}{(100-1)^2} \right] + \frac{1}{100} = 0.076, \\ E[\Delta_1(K)\Delta_2(K)] &= 2.81^2 \left[ \frac{0.8264 \cdot 100 - 1.082}{(100-1)^2} \right] - \frac{1}{100} = 0.056. \end{aligned}$$

The desired value of  $c$  is now given by

$$c = -\frac{92.30 \cdot 0.076 + 963.70 \cdot 0.076}{2 \cdot 310.44} = -0.098.$$

In Table 4, we present the correction factors for  $\alpha = 0.001$ ,  $\alpha = 0.0027$ ,  $\alpha = 0.005$  and  $\alpha = 0.01$  for various values of  $n$  and  $m$ . As can be seen from the table, a smaller  $n$  and  $m$  result in a larger correction. Similarly, the smaller the value of  $\alpha$ , the larger the correction needed. Finally, smaller values of  $\alpha$  and  $n$  result in a negative correction factor. The reason is that there is a larger probability of extreme limits resulting in high CARLs. These high CARLs have an impact on the EARL – calculated as the average of the CARL – so that the correction factor has to be negative. This is not the case for higher values of  $n$  and  $m$  or higher values of  $\alpha$ .



Table 4. Correction factors  $c$  for various values of  $\alpha$ ,  $n$  and  $m$ .

$n$	$m$	ARL <sub>0</sub> = 1000	ARL <sub>0</sub> = 370	ARL <sub>0</sub> = 200	ARL <sub>0</sub> = 100
		$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	-0.8022	-0.6116	-0.5032	-0.3910
	30	-0.5279	-0.4024	-0.3310	-0.2571
	40	-0.3934	-0.2998	-0.2466	-0.1915
	50	-0.3135	-0.2389	-0.1965	-0.1526
	75	-0.2079	-0.1585	-0.1303	-0.1012
	100	-0.1556	-0.1185	-0.0975	-0.0757
3	20	-0.1698	-0.1207	-0.0933	-0.0654
	30	-0.1146	-0.0815	-0.0631	-0.0443
	40	-0.0865	-0.0616	-0.0476	-0.0335
	50	-0.0694	-0.0494	-0.0383	-0.0269
	75	-0.0465	-0.0331	-0.0257	-0.0181
	100	-0.0350	-0.0249	-0.0193	-0.0136
5	20	-0.0453	-0.0241	-0.0126	-0.0013
	30	-0.0306	-0.0163	-0.0086	-0.0010
	40	-0.0231	-0.0123	-0.0065	-0.00084
	50	-0.0185	-0.0099	-0.0053	-0.00072
	75	-0.0124	-0.0067	-0.0035	-0.00051
	100	-0.0093	-0.0050	-0.0027	-0.00039
7	20	-0.0032	0.0087	0.0148	0.0204
	30	-0.0023	0.0057	0.0098	0.0135
	40	-0.0018	0.0042	0.0073	0.0101
	50	-0.0014	0.0033	0.0058	0.0081
	75	-0.00098	0.0022	0.0039	0.0054
	100	-0.00074	0.0016	0.0029	0.0040

### 4.3 Accuracy of correction factors

In this section, we assess the accuracy of the correction factors. In order to do so, we simulate the performance of control charts with these correction factors. The simulation procedure is described in Section 3.2.

The in-control EARL values are presented in Table 5. For comparison purposes, the in-control EARLs of the uncorrected charts are presented in brackets. The EARLs of the uncorrected charts deviate from the desired value ARL<sub>0</sub>. The smaller the value of  $\alpha$  and the smaller the values of  $m$  and  $n$ , the larger the deviation is. As we can conclude from Table 5, the correction factor performs relatively well: for sample sizes greater than  $n = 1$  and  $m = 40$ , the correction factor leads an EARL close to the desired value. As we will see in the next section, this leads to better out-of-control performance as well: when the in-control ARL is exceptionally high, the control chart is not able to detect changes in the mean very quickly.

## 5. Implications of proposed corrected charts

In this section we discuss the implications of using the proposed correct control charts. First, we discuss the out-of-control performance compared to uncorrected charts. Afterwards, we evaluate the variation between CARLs in both the in-control and out-of-control situation.

### 5.1 Out-of-control performance of corrected charts

To evaluate the out-of-control performance of the charts with the new correction factors, we determine the EARLs for two out-of-control situations. We use the same simulation procedure as described earlier but with the adjustment that the process mean has changed, i.e.  $Y_{ij} \sim N(\mu + \delta\sigma/\sqrt{n}, \sigma)$  with  $\delta \neq 0$ . Again, without loss of generality we set  $\mu = 0$  and  $\sigma = 1$  in the simulations.

Table 5. EARL of corrected and, between brackets, uncorrected control charts.

$n$	$m$	ARL <sub>0</sub> = 1000	ARL <sub>0</sub> = 370	ARL <sub>0</sub> = 200	ARL <sub>0</sub> = 100
		$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	534 (1,151,800)	307 (191,160)	179 (7858)	96 (968)
	30	864 (45,575)	337 (3632)	192 (1079)	101 (326)
	40	962 (11,698)	375 (1828)	197 (589)	100 (210)
	50	993 (5458)	367 (1084)	198 (442)	99 (173)
	75	986 (2532)	369 (699)	198 (320)	99 (139)
	100	1002 (1945)	370 (581)	199 (280)	100 (128)
3	20	1090 (2378)	398 (648)	212 (301)	107 (134)
	30	1046 (1693)	383 (520)	208 (259)	102 (117)
	40	1022 (1444)	378 (472)	203 (238)	101 (112)
	50	1017 (1335)	375 (447)	204 (231)	101 (110)
	75	1008 (1201)	371 (415)	202 (219)	101 (106)
	100	1008 (1148)	372 (405)	200 (213)	101 (105)
5	20	1111 (1328)	397 (433)	214 (224)	106 (106)
	30	1053 (1184)	383 (406)	208 (214)	103 (103)
	40	1034 (1127)	374 (391)	204 (209)	101 (101)
	50	1022 (1095)	376 (389)	201 (204)	101 (101)
	75	1007 (1053)	371 (384)	200 (202)	100 (100)
	100	1002 (1036)	370 (377)	199 (201)	100 (100)
7	20	1089 (1102)	399 (387)	210 (200)	105 (98)
	30	1047 (1055)	387 (379)	205 (199)	102 (98)
	40	1031 (1038)	379 (374)	205 (200)	101 (98)
	50	1013 (1018)	375 (371)	203 (199)	101 (99)
	75	1002 (1005)	370 (369)	200 (197)	100 (99)
	100	1001 (1004)	372 (370)	201 (199)	100 (98)

Table 6 shows the results for  $\delta = 0.5$  and Table 7 for  $\delta = 1$ . We can conclude from the tables that the corrected control charts are more powerful in detecting shifts in the process mean than the uncorrected charts, the reason being the high in-control ARL of some uncorrected charts. The positive effect of the correction factor shows up mainly for small sample sizes: for the individuals control chart ( $n = 1$ ) and for  $n = 3$ . In conclusion, for small sample sizes the corrected control charts lead to more stable in-control as well as out-of-control performance compared to the uncorrected control charts.

## 5.2 Variation in CARLs

We have derived correction factors that lead to the desired value of the in-control EARL. Note that a given correction factor leads only to a desired average value: the ARL values will vary across the conditional control charts used by each individual practitioner. To illustrate this, we have drawn boxplots for the CARL values obtained in the simulations. For each simulation run, we can calculate  $E(RL|\hat{\mu}, \hat{\sigma})$  from (10) and we can calculate the percentiles of the CARLs obtained by the simulations runs.

Figures 1 and 2 show the 10th, 25th, 50th, 75th and 90th percentiles of the in-control CARLs when  $\alpha = 0.0027$ , for  $n = 3, m = 50$  and  $n = 1, m = 100$ , respectively. As shown in the figures both the corrected and the uncorrected control charts display quite some variation within the CARL values. However, this variation is much less for the corrected charts than for the uncorrected charts. The same results are found in the out-of-control situation, as can be seen in Figures 3 and 4. These figures show the 10th, 25th, 50th, 75th and 90th percentiles of the out-of-control ( $\delta = 0.5$ ) CARLs when  $\alpha = 0.0027$ , for  $n = 3, m = 50$  and  $n = 1, m = 100$ , respectively.

Table 6. Corrected and, between brackets, uncorrected out-of-control EARL for  $\delta = 0.5$ .

$n$	$m$	$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	276 (1,759,500)	138 (16,449)	97 (7172)	52 (441)
	30	339 (11,595)	157 (1301)	97 (474)	53 (145)
	40	373 (3091)	160 (589)	95 (250)	52 (100)
	50	379 (1636)	162 (460)	95 (193)	52 (85)
	75	379 (871)	159 (276)	93 (142)	52 (70)
	100	377 (672)	160 (237)	93 (125)	51 (63)
3	20	511 (1112)	208 (326)	117 (161)	61 (75)
	30	473 (734)	191 (252)	107 (131)	57 (65)
	40	444 (608)	178 (217)	104 (120)	55 (60)
	50	428 (546)	175 (204)	101 (113)	55 (59)
	75	405 (474)	168 (186)	99 (106)	53 (55)
	100	399 (448)	165 (178)	96 (101)	52 (54)
5	20	546 (646)	210 (227)	119 (124)	62 (62)
	30	491 (547)	190 (201)	110 (113)	59 (59)
	40	452 (489)	184 (191)	104 (106)	56 (56)
	50	427 (454)	176 (182)	101 (103)	55 (55)
	75	412 (429)	169 (173)	98 (99)	53 (53)
	100	401 (413)	165 (167)	95 (96)	52 (52)
7	20	544 (550)	215 (209)	119 (114)	62 (59)
	30	486 (490)	191 (188)	111 (108)	58 (56)
	40	453 (455)	184 (181)	105 (103)	56 (55)
	50	431 (433)	178 (176)	103 (101)	55 (54)
	75	410 (412)	169 (168)	98 (97)	53 (52)
	100	410 (402)	165 (164)	96 (96)	52 (52)

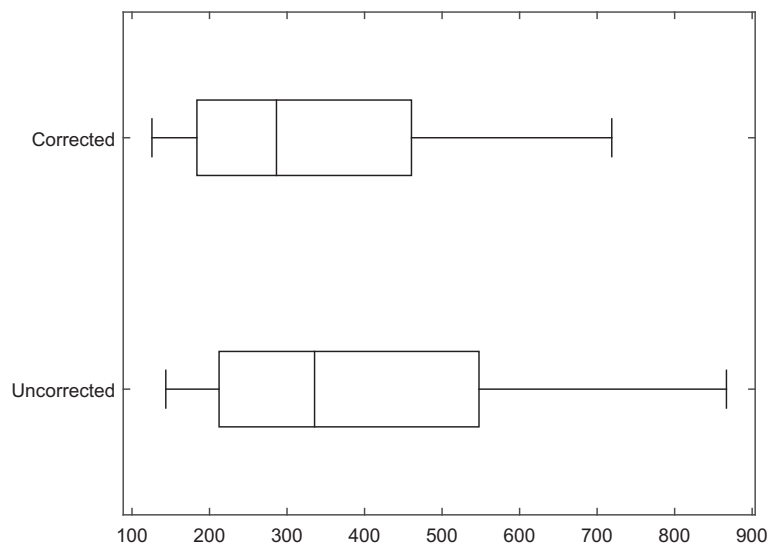
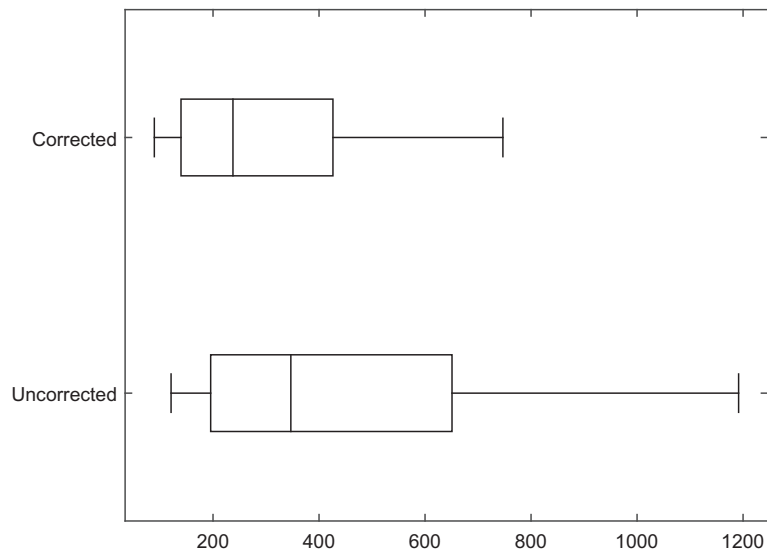


Figure 1. 10th, 25th, 50th, 75th and 90th percentiles of the in-control CARLs for  $n = 3$ ,  $m = 50$  and  $\alpha = 0.0027$ .

Table 7. Corrected and, between brackets, uncorrected out-of-control EARL for  $\delta = 1$ .

$n$	$m$	$\alpha = 0.001$ ( $K = 3.29$ )	$\alpha = 0.0027$ ( $K = 3.00$ )	$\alpha = 0.005$ ( $K = 2.81$ )	$\alpha = 0.01$ ( $K = 2.58$ )
1	20	60 (122,320)	35 (2451)	25 (307)	17 (94)
	30	74 (12,232)	40 (204)	27 (97)	17 (38)
	40	80 (410)	42 (121)	27 (58)	17 (29)
	50	84 (281)	42 (90)	28 (49)	17 (25)
	75	87 (171)	43 (68)	28 (39)	17 (22)
	100	89 (145)	43 (60)	28 (36)	17 (21)
3	20	122 (224)	57 (84)	35 (46)	21 (25)
	30	110 (158)	52 (66)	33 (39)	19 (21)
	40	105 (136)	50 (59)	31 (35)	19 (20)
	50	102 (125)	48 (54)	31 (34)	19 (20)
	75	97 (110)	47 (51)	30 (32)	18 (19)
	100	96 (106)	47 (50)	30 (31)	18 (19)
5	20	131 (151)	59 (63)	37 (38)	22 (22)
	30	115 (126)	53 (56)	34 (34)	20 (20)
	40	110 (117)	51 (53)	32 (33)	19 (19)
	50	105 (110)	49 (51)	31 (32)	19 (19)
	75	99 (103)	48 (48)	30 (30)	18 (18)
	100	96 (99)	46 (47)	29 (30)	18 (18)
7	20	135 (137)	62 (60)	37 (36)	22 (21)
	30	118 (119)	54 (53)	34 (33)	20 (20)
	40	110 (110)	51 (51)	33 (32)	19 (19)
	50	105 (106)	50 (49)	32 (31)	19 (19)
	75	100 (100)	48 (48)	30 (30)	18 (18)
	100	98 (98)	46 (46)	30 (30)	18 (18)

Figure 2. 10th, 25th, 50th, 75th and 90th percentiles of the in-control CARLs for  $n = 1$ ,  $m = 100$  and  $\alpha = 0.0027$ .

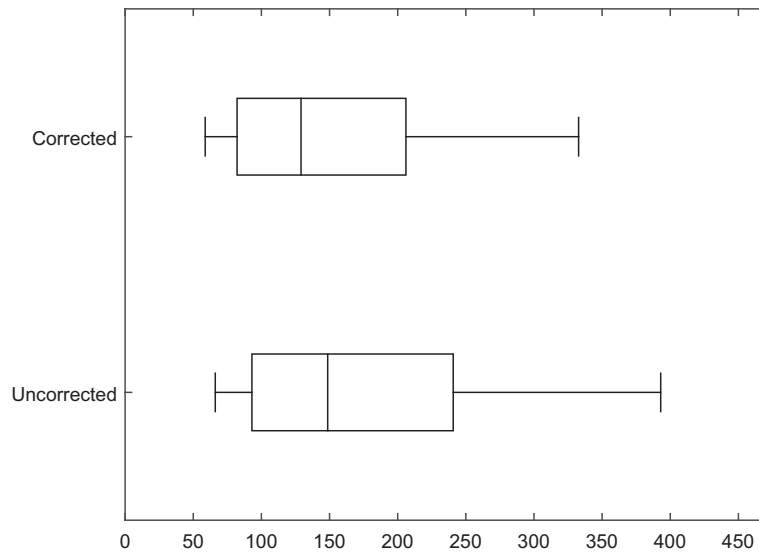


Figure 3. 10th, 25th, 50th, 75th and 90th percentiles of the out-of-control ( $\delta = 0.5$ ) CARLs for  $n = 3, m = 50$  and  $\alpha = 0.0027$ .

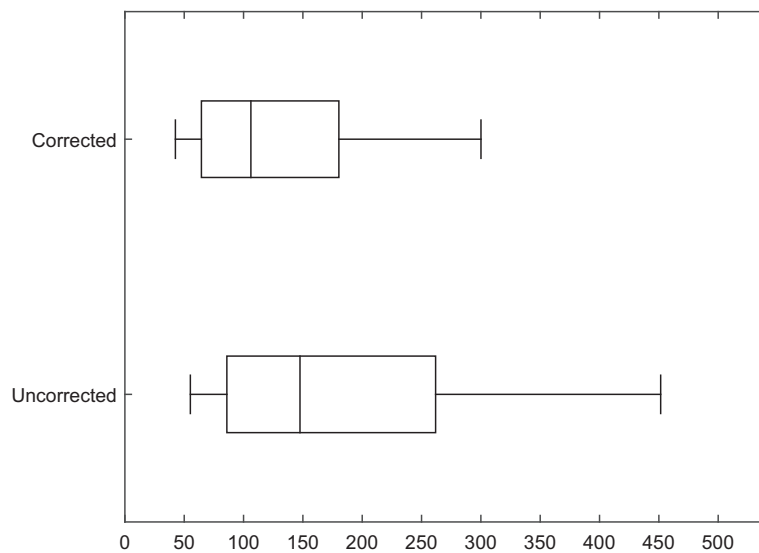


Figure 4. 10th, 25th, 50th, 75th and 90th percentiles of the out-of-control ( $\delta = 0.5$ ) CARLs for  $n = 1, m = 100$  and  $\alpha = 0.0027$ .

### 6. Application of the proposed control chart

In this section we demonstrate the application of the proposed control chart through a real-life example. Data is collected of the torque of Torque-to-Yield bolts that are used as fasteners in engines at a subsidiary of Paccar. The bolts are tightened at several different positions of the engines using a specific fastening procedure. At the end of this procedure, the torque is measured (in Newton-meter) for bolts at two specific positions. The measurements are performed by a process engineer for the purpose of process monitoring. Clearly, detecting trends or anomalies in the applied torque is very important, as this indicates problems with either the bolt(s) or the fastening procedure. For example, the performance of the used wrenches can deteriorate over time, which can result in fasteners being too tight or too loose. It is thus of major importance to detect such out-of-control situations.

An initial Phase I data-set of 40 observations, consisting of  $m = 20$  subgroups of size  $n = 2$ , is collected by the process engineer to construct the control limits (see Table 8). First, we have checked our data for normality. We found no reason to

Table 8. Phase I ( $X_{ij}$ 's) and Phase II ( $Y_{ij}$ 's) data of torque values.

Sample	$X_{i1}$	$X_{i2}$	$Y_{i1}$	$Y_{i2}$
1	164.06	163.98	164.13	164.19
2	164.11	164.05	164.18	164.02
3	164.03	164.09	164.17	164.02
4	164.10	164.13	164.10	164.07
5	164.04	164.15	163.95	164.04
6	164.06	164.22	164.15	164.03
7	163.98	164.11	163.92	164.02
8	164.06	164.09	164.08	164.15
9	164.10	164.08	164.06	163.96
10	164.03	164.03	163.97	164.05
11	164.12	164.09	164.11	164.15
12	164.13	164.04	164.10	164.15
13	164.03	164.10	163.98	164.02
14	164.17	164.05	164.08	164.08
15	164.00	164.06	164.02	164.16
16	164.15	163.98	164.02	164.18
17	163.96	164.02	164.11	164.03
18	164.02	164.08	164.03	164.05
19	164.17	164.23	163.98	164.00
20	164.05	164.07	164.09	163.99
21	–	–	164.14	164.04
22	–	–	163.94	164.03
23	–	–	164.12	164.02
24	–	–	164.03	164.12
25	–	–	164.15	164.18
26	–	–	164.13	164.11
27	–	–	164.00	164.05
28	–	–	164.10	164.15
29	–	–	164.15	164.16
30	–	–	164.33	164.02
31	–	–	164.07	164.28

reject the normality assumption, because the p-value is larger than 0.1 (using the Shapiro-Wilk test for normality). Afterwards, the control limits are constructed and used to monitor the process mean. The Phase II data-set consists of 31 subgroups of size 2 (see also Table 8). The control limits are obtained using the following steps:

- (1) First, we determine our parameters. We have  $m = 20$ ,  $n = 2$  and choose  $\alpha = 0.0027$ . The latter is conventional for Shewhart control charts, and means that we have  $K = 3$ . As estimators of  $\mu$  and  $\sigma$  we use  $\hat{\mu} = \bar{\bar{X}}$  and  $\hat{\sigma} = \tilde{S}/c_4(m(n-1)+1)$ , respectively. As parameter estimates we find  $\hat{\mu} = 164.08$  and  $\hat{\sigma} = 0.0508$  from our Phase I data-set.
- (2) Next, we calculate the required correction  $c$ . This requires the calculations of  $h_{xx}(K, K)$ ,  $h_{xy}(K, K)$ ,  $h_x(K, K)$ ,  $E\Delta_1^2(K)$  and  $E[\Delta_1(K)\Delta_2(K)]$ . These are the same as calculated in Section 4.2, except for

$$E\Delta_1^2(K) = \frac{3^2}{2(20(2-1)+1)} + \frac{1}{20} = 0.2643,$$

$$E[\Delta_1(K)\Delta_2(K)] = \frac{3^2}{2(20(2-1)+1)} - \frac{1}{20} = 0.1643.$$

From this we can calculate  $c$  as

$$c = -\frac{172.13 \cdot 0.2643 + 1996.21 \cdot 0.1643}{2 \cdot 608.03} = -0.3071.$$

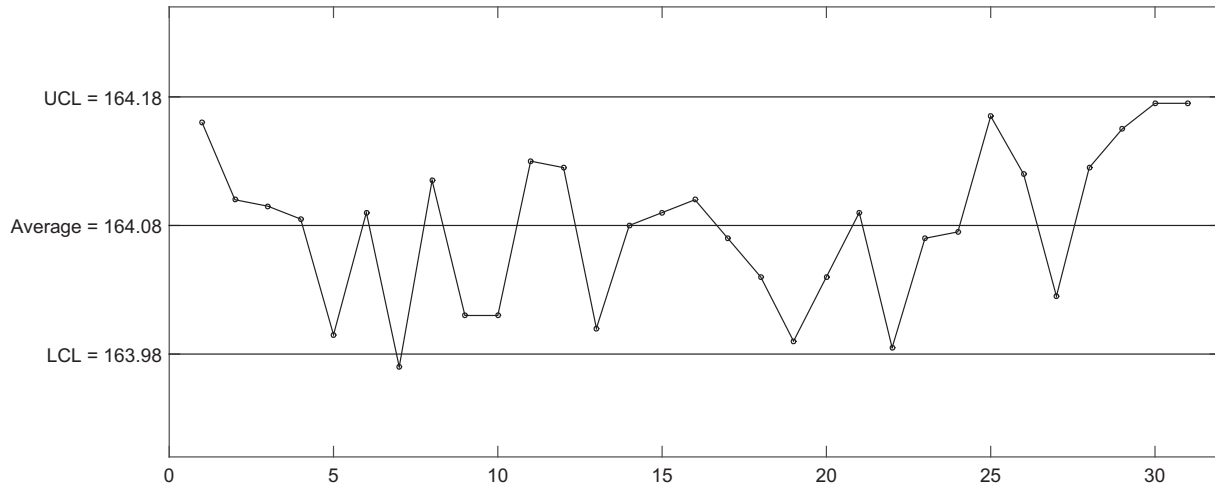


Figure 5. Control chart of the monitoring of torque values with the proposed control limits. The monitored Phase II sample consists of  $m = 31$  subgroups of size  $n = 2$  each.

(3) Finally, we calculate the control limits using  $\hat{\mu}$ ,  $\hat{\sigma}$  and  $\tilde{K} = 3 - 0.3071 = 2.6929$ , such that:

$$\widehat{UCL} = 164.08 + 2.6929 \cdot 0.0508/\sqrt{2} = 164.1767,$$

$$\widehat{LCL} = 164.08 - 2.6929 \cdot 0.0508/\sqrt{2} = 163.9833.$$

The Phase II data is now monitored using these control limits, as is illustrated in Figure 5. Note that the width of the control limits is quite small. As can be observed, an out-of-control signal is given by observation 7 in the data-set. Usually, when a signal is given, the wrenches should be recalibrated. However, after further investigation no action was required in this case, because the out-of-control signal turned out to be caused by measurement uncertainty. The control chart is used in practice to detect substantial anomalies and clear trends.

## 7. Concluding remarks

The ARL is a well-known metric to evaluate control chart performance. However, when the parameters of the process characteristic are estimated from a limited amount of data, the expected value of the in-control ARL differs substantially from the desired value.

In this paper, we have derived correction factors for the Shewhart  $\bar{X}$  and  $\bar{X}$  control charts that can be used to obtain the desired value of the in-control ARL in expectation. As is shown, the correction factor is accurate: it leads to the desired value of the in-control ARL for a broad range of parameter values ( $\alpha$ ,  $n$  and  $m$ ). Thus, using such a factor leads to a control chart with the desired in-control EARL. Moreover, the correction factor overcomes the problem of potentially having high CARLs in the out-of-control situation. The added value of the correction factor shows up predominantly when sample sizes are small, namely the individuals chart ( $n = 1$ ) and the subgroup control chart when  $n = 3$  or  $5$  and  $m = 20$ .

## Acknowledgements

The authors are very grateful to the editor, associate editor and anonymous referees. Their constructive comments have contributed to substantial improvements of this paper. Furthermore, the authors thank Marco Huisman for his valuable contribution in the case reported in this paper.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## References

- Albers, W., and W. C. M. Kallenberg. 2004. "Estimation in Shewhart Control Charts: Effects and Corrections." *Metrika* 59: 207–234.
- Albers, W., and W. C. M. Kallenberg. 2005. "New Corrections for Old Control Charts." *Quality Engineering* 17: 467–473.
- Chakraborti, S. 2006. "Parameter Estimation and Design Considerations in Prospective Applications of the X Chart." *Journal of Applied Statistics* 33 (4): 439–459.
- Chen, G. 1997. "The Mean and Standard Deviation of the Run Length Distribution of  $\bar{X}$  Charts when Control Limits are Estimated." *Statistica Sinica* 7 (3): 789–798.
- Cryer, J. D., and T. P. Ryan. 1990. "The Estimation of Sigma for an X Chart:  $MR/d_2$  or  $S/c_4$ ?" *Journal of Quality Technology* 22: 187–192.
- Does, R. J. M. M., and B. F. Schriever. 1992. "Variables Control Chart Limits and Tests for Special Causes." *Statistica Neerlandica* 46: 229–245.
- Faraz, A., W. H. Woodall, and C. Heuchenne. 2015. "Guaranteed Conditional Performance of the  $S^2$  Control Chart with Estimated Parameters." *International Journal of Production Research* 53 (14): 4405–4413.
- Hillier, F. 1969. "X and R Chart Control Limits Based on a Small Number of Subgroups." *Journal of Quality Technology* 1: 17–26.
- Jensen, W. A., L. A. Jones-Farmer, C. W. Champ, and W. H. Woodall. 2006. "Effects of Parameter Estimation on Control Chart Properties: A Literature Review." *Journal of Quality Technology* 38 (4): 349–364.
- Psarakis, S., A. K. Vyniou, and P. Castagliola. 2014. "Some Recent Developments on the Effects of Parameter Estimation on Control Charts." *Quality and Reliability Engineering International* 30 (8): 641–650.
- Quesenberry, C. P. 1993. "The Effect of Sample Size on Estimated Limits for  $\bar{X}$  and X Control Charts." *Journal of Quality Technology* 25 (4): 237–247.
- Saleh, N. A., M. A. Mahmoud, M. J. Keefe, and W. H. Woodall. 2015. "The Difficulty in Designing Shewhart  $\bar{X}$  and X Control Charts with Estimated Parameters." *Journal of Quality Technology* 47 (2): 127–138.
- Schoonhoven, M., and R. J. M. M. Does. 2012. "A Robust Standard Deviation Control Chart." *Technometrics* 54 (1): 73–82.
- Schoonhoven, M., M. Riaz, and R. J. M. M. Does. 2009. "Design Schemes for the  $\bar{X}$  Control Chart." *Quality and Reliability Engineering International* 25 (5): 581–594.
- Tsai, T. R., J. J. Lin, S. J. Wu, and H. C. Lin. 2005. "On Estimating Control Limits of  $\bar{X}$  Chart when the Number of Subgroups is Small." *International Journal of Advanced Manufacturing Technology* 26 (11–12): 1312–1316.

## Appendix 1.

### A.1 Albers and Kallenberg correction

Although Albers and Kallenberg (2005) derived corrections for the individuals X chart only, it is possible to apply their approach to the  $\bar{X}$  chart as well. First note that they derive correction factors  $\tilde{c}$  in the form  $K(1 + \tilde{c})$  rather than  $K + c$  as we do. Their approach for the individuals chart states that in the one-sided case, the bias is best removed when

$$\tilde{c} = \frac{1}{2} \left( \text{Var} \left( \frac{\bar{X}}{\sigma} \right) + K^2 \text{Var} \left( \frac{\hat{\sigma}}{\sigma} \right) \right). \quad (\text{A1})$$

Next, note that  $\bar{X}$  can essentially be treated as a  $N(\mu, \sigma/\sqrt{n})$  variable. Consequently, replacing  $\sigma$  and  $\hat{\sigma}$  in (A1) by  $\sigma/\sqrt{n}$  and  $\hat{\sigma}/\sqrt{n}$ , respectively, leads to the required corrections for the  $\bar{X}$  chart. This gives us  $\tilde{c} = \frac{1}{2} \left( \text{Var} \left( \frac{\bar{X}}{\sigma/\sqrt{n}} \right) + K^2 \text{Var} \left( \frac{\hat{\sigma}}{\sigma} \right) \right) = \frac{1}{2} \left( \frac{1}{m} + \frac{K^2 \tau^2}{mn} \right) = \frac{n + K^2 \tau^2}{2mn}$ . In order to obtain the required correction for the two-sided case, Albers and Kallenberg instruct to add a minus sign in front of the obtained correction, and to adjust the value of K from  $\Phi^{-1}(1 - \alpha)$  to  $\Phi^{-1}(1 - \alpha/2)$ . Finally, since their correction is in the form of  $K(1 + \tilde{c}) = K + K\tilde{c} = K + c_{AK}$ , we then construct  $c_{AK} = -\frac{nK + K^3 \tau^2}{2mn}$  to make our correction factors comparable. Note that for  $n = 1$  this result corresponds with the original proposal of Albers and Kallenberg (2005).

### A.2 Correction factor

Here, we show how to derive (14) from (13). As given in (13), we want to have

$$\begin{aligned} E h(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) &\approx h(\tilde{K}, \tilde{K}) + h_x(\tilde{K}, \tilde{K}) \left[ E \Delta_1(\tilde{K}) + E \Delta_2(\tilde{K}) \right] \\ &+ \frac{1}{2} h_{xx}(\tilde{K}, \tilde{K}) \left[ E \Delta_1^2(\tilde{K}) + E \Delta_2^2(\tilde{K}) \right] + h_{xy}(\tilde{K}, \tilde{K}) E \left[ \Delta_1(\tilde{K}) \Delta_2(\tilde{K}) \right] = h(K, K). \end{aligned}$$

Next, we apply a two-step Taylor approximation around  $K$  to the middle part and ignore relatively small factors such as  $c^2$ ,  $cE\Delta_1(K)$  and  $cE\Delta_2(K)$ . This essentially boils down to approximating  $h(\tilde{K}, \tilde{K})$  by  $h(K, K) + c[h_x(K, K) + h_y(K, K)]$  and replacing  $\tilde{K}$  by  $K$  in



the remaining parts. We then find that this equation is approximately equal to

$$\begin{aligned} & Eh(\tilde{K} + \Delta_1(\tilde{K}), \tilde{K} + \Delta_2(\tilde{K})) \\ & \approx h(K, K) + c[h_x(K, K) + h_y(K, K)] + h_x(K, K)[E\Delta_1(K) + E\Delta_2(K)] \\ & \quad + \frac{1}{2}h_{xx}(K, K)[E\Delta_1^2(K) + E\Delta_2^2(K)] + h_{xy}(K, K)E[\Delta_1(K)\Delta_2(K)] = h(K, K). \end{aligned}$$

Note that  $h_x(K, K) = h_y(K, K)$ , and that for unbiased estimators  $\hat{\mu}$  and  $\hat{\sigma}$  there holds  $E\Delta_1(K) = E\Delta_2(K) = 0$  and  $E\Delta_1^2(K) = E\Delta_2^2(K)$ . Then, solving this equation to  $c$  leads to

$$c = -\frac{h_{xx}(K, K)E\Delta_1^2(K) + h_{xy}(K, K)E[\Delta_1(K)\Delta_2(K)]}{2h_x(K, K)}.$$

### A.3 Subgroup control chart

For  $E[\Delta_1(K)\Delta_2(K)]$ , we have

$$E[\Delta_1(K)\Delta_2(K)] = -E\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}}\right)^2 + K^2E\left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2,$$

where for  $\hat{\mu} = \bar{X}$  and  $\hat{\sigma} = \tilde{S}/c_4(m(n-1) + 1)$  (cf. (2) and (3)) we have

$$-E\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}}\right)^2 = -\frac{1}{m},$$

and

$$\begin{aligned} & E\left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2 \\ & = \frac{1}{c_4^2(m(n-1) + 1)}E\left(\frac{\tilde{S}^2}{\sigma^2}\right) - 1 \\ & \approx \frac{1}{2(m(n-1) + 1)}, \end{aligned}$$

so that

$$E[\Delta_1(K)\Delta_2(K)] \approx \frac{K^2}{2(m(n-1) + 1)} - \frac{1}{m}$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) \approx \frac{K^2}{2(m(n-1) + 1)} + \frac{1}{m}.$$

### A.4 Individuals control chart

For the individuals control chart, the calculation of  $E[\Delta_1(K)\Delta_2(K)]$ ,  $E\Delta_1^2(K)$  and  $E\Delta_2^2(K)$  is different. For  $E[\Delta_1(K)\Delta_2(K)]$ , we have

$$\begin{aligned} E[\Delta_1(K)\Delta_2(K)] & = -E\left(\frac{\hat{\mu} - \mu}{\sigma}\right)^2 + K^2E\left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2 \\ & \approx -\frac{1}{m} + K^2\text{var}\left(\frac{\hat{\sigma}}{\sigma}\right). \end{aligned}$$

Cryer and Ryan (1990) showed that

$$\text{Var}(\bar{MR}/\sigma) \approx d_2^2(2) \left[ \frac{0.8264m - 1.082}{(m-1)^2} \right]$$

and thus

$$\text{Var}\left(\frac{\hat{\sigma}}{\sigma}\right) \approx \left[ \frac{0.8264m - 1.082}{(m-1)^2} \right].$$

Hence,

$$E[\Delta_1(K)\Delta_2(K)] = -\frac{1}{m} + K^2 \left[ \frac{0.8264m - 1.082}{(m-1)^2} \right]$$

and

$$E\Delta_1^2(K) = E\Delta_2^2(K) = \frac{1}{m} + K^2 \left[ \frac{0.8264m - 1.082}{(m-1)^2} \right].$$