Guaranteed In-Control Performance for the Shewhart X and $\bar{X}$ Control Charts

ROB GOEDHART, MARIT SCHOONHOVEN, and RONALD J. M. M. DOES

University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands

When in-control parameters are unknown, they have to be estimated using a reference sample. Due to the use of different reference samples in phase I, the control chart performance in phase II will vary across practitioners. This variation is especially large for small sample sizes. To prevent low in-control average run lengths, new corrections for Shewhart control charts are proposed that guarantee a minimum in-control performance with a specified probability. However, a minimum in-control performance guarantee generally lowers the out-of-control performance. To balance the tradeoff between in-control and out-of-control performance, the minimum performance threshold and specified probability can be adjusted as desired. The corrections are given in a closed form so that the bootstrap method, which has recently been suggested, is no longer required. The performance of our proposed correction is illustrated by simulating some practical situations. Furthermore, a comparison is made with tolerance intervals and self-starting control charts.

Key Words: Average Run Length; Conditional Distribution; Parameter Estimation; Self-Starting Control Charts; Statistical Process Control; Tolerance Intervals.

1. Introduction

The Shewhart X- and $\bar{X}$-charts are commonly used to monitor the process mean. Because the in-control values of the process mean and variance are generally unknown, they have to be estimated using a reference sample. However, due to the use of different phase-I data, the estimated control limits vary across practitioners. Consequently, the performance of the chart in phase II will vary across practitioners as well. This may result in a low in-control average run length (ARL) or high false alarm rate (FAR) because of an insufficient amount of phase-I data. The practitioner-to-practitioner variability has already been pointed out by multiple researchers. See, e.g., Quesenberry (1993) and Chen (1997), who investigated the average ARL when parameters are estimated, and Saleh et al. (2015a) who consider this problem more generally. An extensive overview of the current literature on the effects of parameter estimation is given in Jensen et al. (2006) and Psarakis et al. (2014).

In order to take the uncertainty of estimation into account, several researchers (e.g., Nedumaran and Pignatiello (2001), Albers and Kallenberg (2004a, 2004b, 2005), Tsai et al. (2004, 2005) and, more recently, Gandy and Kvaløy (2013)) have suggested applying corrections to the control-chart limits. These corrections are intended to make sure that the performance of the control chart satisfies a minimum requirement, which is relevant for practitioners. Generally, this performance is measured in terms of the ARL, FAR, or by matching specific percentiles of the run-length distribution. Related to these measures, Albers and Kallenberg (2004a, 2004b) introduced the *exceedance probability criterion*. This criterion aims to correct the control limits such that a certain value of an in-control performance characteristic (e.g., FAR, ARL) is guaranteed with a specified probability. In other words, the performance of a con-

Mr. Goedhart is a PhD student at the Department of Operations Management, Faculty of Economics and Business and a consultant at the Institute for Business and Industrial Statistics (IBIS UvA). His email address is r.goedhart2@uva.nl.

Dr. Schoonhoven is an Associate Professor in the Department of Operations Management, Faculty of Economics and Business and a senior consultant at IBIS UvA. Her email address is m.schoonhoven@uva.nl.

Dr. Does is a Professor and the Head of the Department of Operations Management and the Director of IBIS UvA. He is a Fellow of the ASQ. His email address is r.j.m.m.does@uva.nl.

trol chart is at least equal to a certain value with a specified probability.

This exceedance probability criterion is of great interest in current SPC literature and similar approaches are recommended by Jones and Steiner (2012), Gandy and Kvaløy (2013), and Saleh et al. (2015a, 2015b). While Albers and Kallenberg (2005) give relatively simple closed-form correction terms for different performance measures, a bootstrap approach to determine the required correction factors is becoming more common (cf., Jones and Steiner (2012), Gandy and Kvaløy (2013), and Saleh et al. (2015a, 2015b)).

In this study, we do not use the bootstrap method but determine an analytical correction term for the Shewhart X- and $\overline{X}$-charts under normal theory in a very general setup. Our proposed term is applicable to both one- and two-sided control charts and can make use of different estimators for location and spread. Also, the specifications of the correction can be adjusted by specifying the parameters. This gives the freedom to adapt the correction to the circumstances as desired by the practitioner. It turns out that our corrections perform better than the correction terms of Albers and Kallenberg (2005). Next to that, because our correction terms are analytical expressions, the computation time is negligible.

Furthermore, we compare our approach with the construction of two-sided tolerance intervals for normal populations. This problem has a long history of academic interest (cf., Wald and Wolfowitz (1946), Weissberg and Beatty (1960), Gardiner and Hull (1966), Howe (1969), and Krishnamoorthy and Mathews (2009)). In those publications, various approximations and practical guidelines were developed along with rigorous justification for the guarantee of tolerance probability. Finally, we compare our limits with self-starting control charts introduced by Hawkins (1989) and Quesenberry (1991), which also guarantee the conditional and unconditional in-control performance in an exact way. In all comparisons, we show that our corrections perform much better and/or are more general than the existing ones.

In the next section, we present our model and approach, which includes some baseline terminology and derivations. In Section 3, we derive our proposed correction terms, illustrate the performance, and compare the results with the case of uncorrected limits. In Section 4, we address the out-of-control performance of the proposed method. In Section 5, we illustrate the performance and consequences of the proposed correction terms, compared with existing methods. Concluding remarks and recommendations are given in Section 6.

2. Model and Approach

In phase I, we have m samples of size n, so that sample i, for $i = 1, \ldots, m$, consists of the observations $X_{i1}, \ldots, X_{in}$. Let X_{ij} be i.i.d. $N(\mu, \sigma)$-distributed random variables. In phase II, the average of a subsample ($\overline{X}_{m+1}$, $\overline{X}_{m+2}$, etc) is used to determine whether the process is in-control or not. Note that $\overline{X}_{m+1}$ is distributed as $N(\mu, \sigma^2/n)$. We consider a standard Shewhart $\overline{X}$ control chart with both an upper control limit (UCL) and a lower control limit (LCL), as well as the invividuals Shewhart X chart for the case $n = 1$. If the parameters μ and σ are known, the control limits equal

$$\begin{aligned} \text{UCL} &= \mu + K_\alpha \frac{\sigma}{\sqrt{n}} \\ \text{LCL} &= \mu - K_\alpha \frac{\sigma}{\sqrt{n}}, \end{aligned} \tag{1}$$

where α is the desired FAR and $K_\alpha = \Phi^{-1}(1 - \alpha/2)$ (henceforth $K_\alpha = K$), where $\Phi^{-1}(x)$ denotes the inverse of the standard normal CDF. In practice, however, parameters are generally unknown and need to be estimated. Using unbiased estimators $\hat{\mu}$ and $\hat{\sigma}$ for μ and σ, respectively, estimated control limits $\widehat{\text{UCL}}$ and $\widehat{\text{LCL}}$ equal

$$\begin{aligned} \widehat{\text{UCL}} &= \hat{\mu} + K \frac{\hat{\sigma}}{\sqrt{n}} \\ \widehat{\text{LCL}} &= \hat{\mu} - K \frac{\hat{\sigma}}{\sqrt{n}} \end{aligned} \tag{2}$$

and are unbiased estimators of UCL and LCL, respectively. The true probability of a false alarm is a function of these estimated control limits and is thus dependent on K, m, n, $\hat{\mu}$, and $\hat{\sigma}$. Given the estimators $\hat{\mu}$ and $\hat{\sigma}$, it is possible to calculate the true FAR conditional on these estimators. This FAR, denoted as $P_{mn}(K; \hat{\mu}, \hat{\sigma})$, is equal to

$$\begin{aligned} &P_{mn}(K; \hat{\mu}, \hat{\sigma}) \\ &\quad = 1 - P\left(\widehat{\text{LCL}} < \overline{X}_{m+1} < \widehat{\text{UCL}}\right) \\ &\quad = 1 - P\left(\overline{X}_{m+1} < \hat{\mu} + K\frac{\hat{\sigma}}{\sqrt{n}}\right) \\ &\qquad + P\left(\overline{X}_{m+1} < \hat{\mu} - K\frac{\hat{\sigma}}{\sqrt{n}}\right) \end{aligned}$$

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

where $\Phi(x)$ denotes the standard normal cumulative distribution function (CDF). $P_{mn}(K;\hat{\mu},\hat{\sigma})$ can be rewritten as

$$P_{mn}(K;Z,W) = 1 - \Phi\left(\frac{Z}{\sqrt{m}} + KW\right) + \Phi\left(\frac{Z}{\sqrt{m}} - KW\right), \quad (3)$$

where $Z = (\hat{\mu}-\mu)/(\sigma/\sqrt{mn})$ and $W = \hat{\sigma}/\sigma$. Note that Z and W are random variables, of which the distributions depend on the estimators $\hat{\mu}$ and $\hat{\sigma}$, respectively. Consequently, the unconditional FAR (i.e., before $\hat{\mu}$ and $\hat{\sigma}$ are obtained) is a random variable as well. Because practitioners use different phase-I samples, their estimations will vary, resulting in a different FAR for their control charts. This variation is known as practitioner-to-practitioner variability; cf., Saleh et al. (2015a).

Quesenberry (1993) and Chen (1997) already illustrated the effect of estimated limits for $\overline{X}$ and X control charts by means of the ARL and the standard deviation of the run length. We consider the practitioner variability by considering the average ARL (AARL), which is the average ARL over all practitioners, and its variation. Note that, conditional on the FAR, the run length follows a geometric distribution with parameter FAR, such that the ARL equals 1/FAR. The AARL then equals the unconditional expectation of the ARL. As illustrated by Quesenberry (1993), Chen (1997), and others, this unconditional expectation of the ARL using estimated parameters generally does not equal the desired value of $1/\alpha$. To deal with this, the use of correction terms for the control limits is suggested; cf., Nedumaran and Pignatiello (2001), Albers and Kallenberg (2004a, 2004b), and Tsai et al. (2004, 2005).

If the distribution of $\hat{\mu}$ is symmetric (such as is the case for $\hat{\mu} = \overline{\overline{X}} = (1/m)\sum_{i=1}^{m} \overline{X}_i = (1/mn) \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}$), the corrections for one-sided control charts with either a UCL or an LCL are identical. However, one cannot simply apply these corrections separately to two-sided control charts. For the equivalence of the one-sided corrections, note that the FAR, when using only a UCL, is equal to the first part of Equation (3) and can be written as

$$1 - \Phi\left(\frac{Z}{\sqrt{m}} + KW\right) = \Phi\left(-\frac{Z}{\sqrt{m}} - KW\right). \quad (4)$$

If $\hat{\mu}$ follows a symmetric distribution around μ, consequently this also holds for Z around 0. That means that the distribution of Equation (4) is in that case equal to that of $\Phi(Z/\sqrt{m} - KW)$, which is the FAR for a control chart using an LCL only. Therefore, both one-sided charts require the same correction. However, applying this one-sided correction to both sides of a two-sided chart is essentially the same as correcting for twice the FAR of one side, which is distributed as

$$2\Phi\left(\frac{Z}{\sqrt{m}} - KW\right) = 1 - \Phi\left(-\frac{Z}{\sqrt{m}} + KW\right) + \Phi\left(\frac{Z}{\sqrt{m}} - KW\right). (5)$$

Note that the expression in Equation (5) is quite different from the actual two-sided FAR as defined in Equation (3). The intuition behind it is that, in the two-sided case, an underestimation of the mean increases $P(\overline{X}_{m+1} > \widehat{\text{UCL}})$, but at the same time decreases $P(\overline{X}_{m+1} < \widehat{\text{LCL}})$ (and vice versa for an overestimation of the mean). This effect is not taken into account when one simply applies one-sided corrections to a two-sided control chart. For this reason, one-sided and two-sided control charts should be treated separately.

We consider a correction term c that is added to K. This changes our estimated control limits of Equation (2) into

$$\widehat{\text{UCL}} = \hat{\mu} + (K+c)\frac{\hat{\sigma}}{\sqrt{n}}$$
$$\widehat{\text{LCL}} = \hat{\mu} - (K+c)\frac{\hat{\sigma}}{\sqrt{n}}. \quad (6)$$

If we denote $\tilde{K} = K + c$, then our conditional FAR when using the correction term c is calculated by $P_{mn}(\tilde{K};\hat{\mu},\hat{\sigma})$. A common choice is to correct the control-chart limits such that the expectation of a chosen performance measure (e.g., FAR or ARL) is equal to a desired value. This would ensure that the control chart performs as it should on average over all practitioners. However, this does not take the variability between practitioners into account. We suggest a correction term that is based on the exceedance probability criterion, introduced by Albers and Kallenberg (2004b) and which is also used in some recent research (e.g., Gandy and Kvaløy (2013)). The approach aims to correct the control-chart limits such that at least a specified minimum in-control performance is obtained with a specified probability. More specifically, for the FAR, the correction term aims to obtain a FAR that is equal to $(1+\epsilon)\alpha$ or smaller with a probability of $1-p$. Here, ϵ

determines our minimum threshold, proportional to the nominal level, while p represents the probability of obtaining a control-chart performance lower than the specified minimum. In our view, p should be small (e.g., 0.05 or 0.10) and ϵ may be slightly larger (e.g., 0.2 to 0.5). This is because $1-p$ represents the probability for the practitioners with which the minimum performance is satisfied, while ϵ determines our *minimum* performance threshold. Our choices of p and ϵ are the same as in recent research (e.g., Albers and Kallenberg (2005)) and for $\epsilon = 0$ (see also Saleh et al. (2015b)). Hence, for the FAR, we derive a correction term such that

$$P\left(P_{mn}(\tilde{K}; Z, W) < (1+\epsilon)\alpha\right) = 1-p. \quad (7)$$

Note that, for the ARL in the in-control situation, large ARLs are preferred. The criterion would then be

$$P\left(\frac{1}{P_{mn}(\tilde{K}; Z, W)} > (1-\epsilon)\frac{1}{\alpha}\right) = 1-p. \quad (8)$$

The required correction term for the ARL can be obtained through solving the term for the FAR while replacing ϵ by $\tilde{\epsilon} = \epsilon/(1-\epsilon)$. This is because the left-hand side of Equation (8) is equivalent to

$$\begin{aligned} &P\left(P_{mn}(\tilde{K}; Z, W) < \frac{1}{1-\epsilon}\alpha\right) \\ &= P\left(P_{mn}(\tilde{K}; Z, W) < \left(1+\frac{\epsilon}{1-\epsilon}\right)\alpha\right) \\ &= P\left(P_{mn}(\tilde{K}; Z, W) < (1+\tilde{\epsilon})\,\alpha\right). \end{aligned}$$

Because of this equivalence, our further derivations are based on the FAR.

In order to determine the correction term c, we require information on the distribution of $P_{mn}(\tilde{K}; Z, W)$, as we need to find the value c for which the $1-p$ percentile of the distribution of $P_{mn}(\tilde{K}; Z, W)$ equals $(1+\epsilon)\alpha$. Although the exact distribution of $P_{mn}(K; Z, W)$ for an arbitrary K is unknown, it is possible to calculate its moments using integrals, similar to Chen (1997). The first and second moment can be calculated as by

$$\begin{aligned} &E\left(P_{mn}(K; Z, W)\right) \\ &= \int_{-\infty}^{\infty}\int_{0}^{\infty} P_{mn}(K; z, w) f(z) f(w)\, dw dz \quad (9) \end{aligned}$$

and

$$\begin{aligned} &E(P_{mn}^2(K; Z, W)) \\ &= \int_{-\infty}^{\infty}\int_{0}^{\infty} P_{mn}^2(K; z, w) f(z) f(w)\, dw dz, \quad (10) \end{aligned}$$

respectively. Here, $f(z)$ equals the probability density of $Z = (\hat{\mu}-\mu)/(\sigma/\sqrt{mn})$ and $f(w)$ is the probability density of $W = \hat{\sigma}/\sigma$. The variance of $P_{mn}(K; Z, W)$ can be calculated through

$$\begin{aligned} &\mathrm{Var}(P_{mn}(K; Z, W)) \\ &= E(P_{mn}^2(K; Z, W)) - E^2(P_{mn}(K; Z, W)). \quad (11) \end{aligned}$$

Using these moments, we approximate the distribution of $P_{mn}(K; Z, W)$ by a $a\chi_b^2/b$ distribution (for a detailed motivation, we refer to Appendix A). Because $\mathrm{E}(a\chi_b^2/b) = a$ and $\mathrm{Var}(a\chi_b^2/b) = 2a^2/b$, we have

$$a = E(P_{mn}(K; Z, W)) \quad (12)$$

and

$$b = \frac{2a^2}{\mathrm{Var}(P_{mn}(K; Z, W))} = \frac{2E^2(P_{mn}(K; Z, W))}{\mathrm{Var}(P_{mn}(K; Z, W))}. \quad (13)$$

Note that both the expectation and variance depend on K and, consequently, after implementing the correction term through $\tilde{K} = K + c$, also on c.

3. Correction Terms

3.1. Two-sided Control Limits

As mentioned in the previous section, both a and b, as in Equations (12) and (13), depend on K. This means that changing K to $\tilde{K} = K + c$ changes the $a\chi_b^2/b$ distribution as well. The correction factor c has to be determined such that, for $\tilde{K} = K + c$, the $(1-p)$'th percentile of this distribution lies at $(1+\epsilon)\alpha$.

If we denote the expectation and variance of $P_{mn}(K; Z, W)$ in Equations (9) and (11) by E and V, respectively, and their derivatives with respect to K by dE/dK and dV/dK, respectively, then the correction term can be expressed by

$$c = \frac{\Phi^{-1}(1-p) - Y(K)}{Y'(K)}, \quad (14)$$

where

$$Y(K) = \sqrt[3]{(1+\epsilon)\alpha}\frac{3E^{2/3}}{\sqrt{V}} - \frac{3E}{\sqrt{V}} + \frac{\sqrt{V}}{3E} \quad (15)$$

$$\begin{aligned} Y'(K) = &\sqrt[3]{(1+\epsilon)\alpha}\frac{2E^{-1/3}\sqrt{V}\frac{dE}{dK} - \frac{3E^{2/3}}{2\sqrt{V}}\frac{dV}{dK}}{V} \\ &- \frac{3\frac{dE}{dK}\sqrt{V} - \frac{3E}{2\sqrt{V}}\frac{dV}{dK}}{V} \\ &+ \frac{\frac{3E}{2\sqrt{V}}\frac{dV}{dK} - 3\frac{dE}{dK}\sqrt{V}}{9E^2} \end{aligned} \quad (16)$$

and where

$$\frac{dE}{dK} = \int_{-\infty}^{\infty}\int_{0}^{\infty}\frac{dP_{mn}(K;z,w)}{dK}f(z)f(w)dwdz$$
$$= \int_{-\infty}^{\infty}\int_{0}^{\infty} -w\left[\phi\left(\frac{z}{\sqrt{m}}+Kw\right) + \phi\left(\frac{z}{\sqrt{m}}-Kw\right)\right] \times f(z)f(w)dwdz \quad (17)$$

$$\frac{dV}{dK} = \frac{d(E^2)}{dK} - 2E\frac{dE}{dK} \quad (18)$$

with

$$\frac{d(E^2)}{dK} = \int_{-\infty}^{\infty}\int_{0}^{\infty}\frac{dP^2_{mn}(K;z,w)}{dK}f(z)f(w)dwdz$$
$$= -2\int_{-\infty}^{\infty}\int_{0}^{\infty} wP_{mn}(K;z,w) \times \left[\phi\left(\frac{z}{\sqrt{m}}+Kw\right) + \phi\left(\frac{z}{\sqrt{m}}-Kw\right)\right] \times f(z)f(w)dwdz. \quad (19)$$

For a detailed derivation of the correction term, we refer to Appendix B.

Note that, instead of a minimum performance threshold that is relative (through ϵ) to α (or $1/\alpha$), one could also chose to specify an absolute minimum performance threshold. This can be done by setting $\alpha = \mathrm{FAR}_0 = 1/\mathrm{ARL}_0$ and $\epsilon = 0$, where FAR_0 and ARL_0 are the desired threshold values for the FAR or ARL, respectively. However, in order to make a better comparison between the corrected and uncorrected charts, we consider relative thresholds. The same approach has been used by other authors (e.g., Albers and Kallenberg (2005)).

We use simulation to evaluate the performance of the proposed correction term. In order to do this, we need to define our estimators and determine the corresponding distributions $f(z)$ and $f(w)$ of $Z = (\hat{\mu}-\mu)/(\sigma/\sqrt{mn})$ and $W = \hat{\sigma}/\sigma$, respectively. We consider $\hat{\mu} = \overline{\overline{X}}$ as the estimator for location, in which case $f(z)$ equals the standard normal probability density. For the spread, we consider two estimators. The first estimator, which we use in the case that $n > 1$ (groups), is based on the pooled standard deviation (s_p) and is equal to

$$\hat{\sigma}_1 = \frac{s_p}{c_4(m(n-1)+1)}, \quad (20)$$

where $c_4(m(n-1)+1)$ is such that $\hat{\sigma}_1$ is an unbiased estimator of σ. For this estimator, $W = \hat{\sigma}_1/\sigma$ is distributed as $\tau\chi_\nu/\sqrt{\nu}$ with $\tau = 1/c_4(m(n-1)+1)$ and $\nu = m(n-1)$. The probability density of $\tau\chi_\nu/\sqrt{\nu}$ is then

$$f(w;\tau,\nu) = \left(\frac{2}{\tau}\right)\frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}\left(\frac{w}{\tau}\right)^{\nu-1}\exp\left(-\frac{\nu}{2}\left(\frac{w}{\tau}\right)^2\right) \quad (21)$$

(see also Chen (1997)). The second estimator, which is used in the case $n = 1$ (individuals), is based on the average moving range $\overline{\mathrm{MR}}$ and is equal to

$$\hat{\sigma}_2 = \frac{\overline{\mathrm{MR}}}{d_2(2)}, \quad (22)$$

where $d_2(2) = 2/\sqrt{\pi}$, which yields that $\hat{\sigma}_2$ is an unbiased estimator of σ. Although the exact distribution of $\overline{\mathrm{MR}}$ is not easy to obtain, the distribution of $W = \hat{\sigma}_2/\sigma$ can be approximated by $\beta\chi_\gamma/\sqrt{\gamma}$, where

$$\beta = \sqrt{\mathrm{Var}(W)+1}$$
$$\gamma = \frac{1}{2}\left(1+\frac{1}{\mathrm{Var}(W)}\right). \quad (23)$$

See for example Roes et al. (1993) for this approximation, or Patnaik (1950) for a similar one. The variance of $W = \hat{\sigma}_2/\sigma$ is investigated by Cryer and Ryan (1990) and can be approximated by

$$\mathrm{Var}\left(\frac{\hat{\sigma}_2}{\sigma}\right) = \left[\frac{0.8264m - 1.082}{(m-1)^2}\right]. \quad (24)$$

Note that it is possible to use any other estimator for location and spread by applying their corresponding probability functions $f(z)$ and $f(w)$. In Table 1 and Table 2, several estimators for μ and σ, respectively, are listed. For the estimators of μ, the (approximated) probability density function $f(z)$ is tabulated. For the mean, we have an exact distribution and, for the median, the distribution has been approximated (cf., Johnson and Kotz (1970)). For es-

TABLE 1. Estimators for Location, Including the (Approximated) Probability Density Function $f(z)$ of $Z = (\hat{\mu} - \mu)/(\sigma/\sqrt{mn})$

Estimator	$f(z)$
Average ($\overline{\overline{X}}$)	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$
Median ($\tilde{X}$)	$\frac{1}{\pi}\exp\left(-\frac{z^2}{\pi}\right)$

TABLE 2. Estimators for Spread, Including the (Approximated) Probability Density Function $f(w)$ of $W = \hat{\sigma}/\sigma$

Estimator	ζ and λ	var(W)
Individuals $(n = 1)$		
Average moving range, $\hat{\sigma} = \dfrac{\overline{\text{MR}}(\mathbf{X})}{d_2(2)}$	$\zeta = \sqrt{\text{Var}(W)+1}$, $\lambda = \dfrac{1}{2}\left(1 + \dfrac{1}{\text{Var}(W)}\right)$	$\dfrac{0.8264m - 1.082}{(m-1)^2}$
Interquartile range $\hat{\sigma} = \dfrac{\text{IQR}(\mathbf{X})}{1.349}$	$\zeta = \sqrt{\text{Var}(W)+1}$, $\lambda = \dfrac{1}{2}\left(1 + \dfrac{1}{\text{Var}(W)}\right)$	$\dfrac{2.46}{1.820m}$
Sample standard deviation $\hat{\sigma} = \dfrac{s}{c_4(m)}$	$\zeta = 1/c_4(m)$, $\lambda = m - 1$	$\dfrac{1 - c_4^2(m)}{c_4^2(m)}$
Groups $(n > 1)$		
Average sample standard deviation $\hat{\sigma} = \dfrac{\bar{s}}{c_4(n)}$	$\zeta = \sqrt{\text{Var}(W)+1}$, $\lambda = \dfrac{1}{2}\left(1 + \dfrac{1}{\text{Var}(W)}\right)$	$\dfrac{1 - c_4^2(n)}{mc_4^2(n)}$
Pooled sample standard deviation $\hat{\sigma} = \dfrac{s_p}{c_4(m(n-1+1))}$	$\zeta = \dfrac{1}{c_4(m(n-1)+1)}$, $\lambda = m(n-1)$	$\dfrac{1 - c_4^2(m(n-1)+1)}{c_4^2(m(n-1)+1)}$

timators of σ, the distribution of $W = \hat{\sigma}/\sigma$ is either exact or approximated by a $\zeta\chi_\lambda/\sqrt{\lambda}$ distribution, so that $f(w)$ is given by Equation (21). The approximation is similar to that of $\overline{\text{MR}}$, by determining the corresponding values of ζ and λ. The required values of ζ and λ for the considered estimators are listed in Table 2. We refer to Roes et al. (1993) for the approximations and to Albers and Kallenberg (2005) for explicit expressions of the listed estimators of σ.

As the ARL is the most commonly used performance measure of control charts, we evaluate the performance of our proposed correction terms based on that. The corresponding criterion function is given in Equation (8). To calculate the required correction terms, we use the model as described in the previous section, thus replacing ϵ by $\tilde{\epsilon} = \epsilon/(1-\epsilon)$ in the correction terms for the FAR. We have calculated the correction terms and simulated their performance for a wide range of parameter values. For each combination of parameter values 1,000,000 simulation runs are performed. The relative standard errors of all reported AARL values in Tables 3–8 are less than 1%. Table 3 and Table 4 illustrate, for multiple combinations of n and m, the correction term and its performance compared with an uncorrected chart, each for a different set of p, ϵ, and α. The performance is measured by the exceedance probability as in Equation (8) while, for the comparisons, the AARLs are also given. The results illustrate that, after implementing our suggested correction term, the exceedance probability is very close to the desired level. Especially for small sample sizes, the differences between the corrected and the uncorrected chart are large. Note that this is in agreement with the sample size recommendations by, e.g., Quesenberry (1993), which states that about $400/(n-1)$ samples of size n are required in phase I for the $\overline{X}$ chart to behave properly on average. Furthermore, Saleh et al. (2015a) show that far larger amounts of phase-I data are required when also taking the practitioner-to-practitioner variability into account. As more data are available in phase I, the required correction becomes smaller. At some point, the exceedance probability of the uncorrected chart is already below the desired level, meaning that the correction term becomes negative to increase the exceedance probability. The use of a negative correction term in this case is questionable, as it will de-

TABLE 3. Correction Terms c for the $\overline{X}$-Chart, Including the Corresponding Exceedance Probability and AARL for the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.0027$, $p = 0.05$, and $\epsilon = 0.2$

m	n	Correction c	Exc. Pr. (cor)	Exc. Pr. (unc)	AARL (cor)	AARL (unc)
25	3	0.5687	0.0516	0.4836	7261	569
	5	0.3970	0.0478	0.4715	1890	418
	9	0.2822	0.0473	0.4534	984	364
50	3	0.3532	0.0483	0.4275	1721	449
	5	0.2311	0.0494	0.3956	879	389
	9	0.1541	0.0501	0.3595	610	362
75	3	0.2615	0.0495	0.3908	1078	418
	5	0.1651	0.0507	0.3451	674	382
	9	0.1034	0.0512	0.2922	514	364
100	3	0.2097	0.0501	0.3617	850	405
	5	0.1302	0.0511	0.3132	585	376
	9	0.0756	0.0519	0.2434	469	365
150	3	0.1540	0.0504	0.3239	662	389
	5	0.0884	0.0521	0.2555	503	374
	9	0.0447	0.0519	0.1752	425	366
200	3	0.1202	0.0512	0.2893	579	384
	5	0.0647	0.0520	0.2120	463	373
	9	0.0274	0.0522	0.1289	402	367
250	3	0.0980	0.0516	0.2600	531	381
	5	0.0491	0.0521	0.1792	438	372
	9	0.0160	0.0523	0.0971	388	368

crease the in-control performance in that situation. However, the out-of-control performance im proves while keeping the in-control performance at a desired level.

As mentioned earlier, this correction term is applicable for any estimator. Previously, we have shown the results for the $\overline{X}$-chart using an estimator for σ based on the pooled standard deviation (see Equation (20)). In this section, we also show the results for the individuals X-chart, by using an unbiased version of the average moving range as an estimator of σ, as in Equation (22). Similarly as for the $\overline{X}$-chart, Table 5 and Table 6 indicate the correction term and its performance compared with the uncorrected chart for multiple values of m (and $n = 1$), each for a different set of p, ϵ, and α. Even though we use an approximation of the density of $\hat{\sigma}_2$, the correction term still performs well, with the exceedance probabilities close to the desired level.

3.2. One-sided Control Limits

A fairly straightforward change in the previous derivations leads to the correction term in case we are dealing with either a UCL or an LCL only (one-sided). As the correction term for the UCL and LCL is the same, we consider the case of the UCL. The main difference lies in the change of Equation (3) to

$$P_{kn}(u; Z, W) = 1 - \Phi\left(\frac{z}{\sqrt{k}} + uw\right), \qquad (25)$$

where u is now equal to $\Phi^{-1}(1-p)$. Then this formula should be used to calculate the expectation and variance of $P_{kn}(u; Z, W)$ in Equations (9) and (11). The expressions in Equations (17) and (19) will change

TABLE 4. Correction Terms c for the $\overline{X}$-Chart, Including the Corresponding Exceedance Probability and AARL for the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.01$, $p = 0.1$, and $\epsilon = 0.4$

m	n	c	Exc. pr. (cor)	Exc. pr. (unc)	AARL (cor)	AARL (unc)
25	3	0.2325	0.0975	0.3049	275	123
	5	0.1216	0.0967	0.2352	151	104
	9	0.0507	0.0946	0.1615	111	96
50	3	0.0875	0.0975	0.1956	144	110
	5	0.0124	0.0987	0.1150	105	101
	9	−0.0349	0.0988	0.0518	88	97
75	3	0.0289	0.0991	0.1347	116	106
	5	−0.0305	0.0994	0.0601	92	101
	9	−0.0688	0.1004	0.0172	80	98
100	3	−0.0040	0.0996	0.0946	103	104
	5	−0.0529	0.1005	0.0337	86	100
	9	−0.0875	0.1017	0.0059	77	98
150	3	−0.0390	0.1005	0.0511	91	102
	5	−0.0802	0.1008	0.0103	79	100
	9	−0.1081	0.1027	0.0007	73	99
200	3	−0.0606	0.1005	0.0272	85	102
	5	−0.0957	0.1016	0.0033	76	100
	9	−0.1197	0.1032	0.0001	71	99
250	3	−0.0749	0.1012	0.0152	82	101
	5	−0.1060	0.1020	0.0011	74	100
	9	−0.1273	0.1035	0.0000	69	99

as

$$\frac{dE}{du} = \int_{-\infty}^{\infty}\int_{0}^{\infty} \frac{dP_{kn}(u;z,w)}{du} f(z)f(w)dwdz = \int_{-\infty}^{\infty}\int_{0}^{\infty} -w\phi\left(\frac{z}{\sqrt{k}} + uw\right) f(z)f(w)dwdz \quad (26)$$

$$\frac{d(E^2)}{du} = \int_{-\infty}^{\infty}\int_{0}^{\infty} \frac{dP_{kn}^2(u;z,w)}{du} f(z)f(w)dwdz = -2\int_{-\infty}^{\infty}\int_{0}^{\infty} wP_{kn}(u;z,w)\phi\left(\frac{z}{\sqrt{k}} + uw\right) \times f(z)f(w)dwdz, \quad (27)$$

which can be used to obtain the corresponding expression in (18):

$$dV/du = d(E^2)/du - 2E(dE/du). \quad (28)$$

The corresponding correction term c in Equation (14) for the one-sided control limit can be obtained by implementing the resulting expressions in Equations (15) and (16).

4. Out-of-Control Performance

Correcting the control-chart limits in order to guarantee a minimum performance clearly has an advantage in the in-control situation. However, this inevitably leads to a deterioration of the out-of-control performance. In the previous section, more specifically in Tables 3–6, the AARLs of the corrected and uncorrected charts are given. It can already be seen that there is a large difference between the corrected and uncorrected control limits for small sample sizes. As more information becomes available, and as a consequence the correction term becomes smaller, this difference becomes smaller. For the out-

TABLE 5. Correction Terms c for the Individuals X-Chart, with the Corresponding Exceedance Probability and AARL for the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.0027$, $p = 0.05$, and $\epsilon = 0.2$

m	c	Exc. pr. (cor)	Exc. pr. (unc)	AARL (cor)	AARL (unc)
50	0.6930	0.0563	0.4723	78,131	1,086
75	0.5510	0.0492	0.4492	9,902	697
100	0.4596	0.0471	0.4308	4,156	580
150	0.3495	0.0470	0.4041	1,916	491
200	0.2852	0.0475	0.3819	1,317	455
250	0.2425	0.0483	0.3633	1,053	436
500	0.1419	0.0502	0.2979	655	401
1,000	0.0760	0.0516	0.2191	498	385

of-control situation, this behavior is very similar. To illustrate this, we have simulated the AARL of the corrected and uncorrected $\overline{X}$-charts for different sizes of shifts. We consider $p = 0.05$, $\epsilon = 0.2$, and $\alpha = 0.0027$ and have again simulated for a wide range of values for m and n. For the out-of-control situation, we consider a shift of $\overline{X}$ from a $N(\mu, \sigma^2/n)$ to a $N(\mu + \delta(\sigma/\sqrt{n}), \sigma^2/n)$ distribution, with δ equal to 0.5, 1, or 2. The results are listed in Table 7.

We find that, in this situation, for small sample sizes, the differences in AARLs between the corrected and uncorrected charts are substantial. Note that we have chosen a rather strict set of parameters, as we guarantee an in-control ARL of at least 296 with a probability of 90%. The out-of-control performance becomes better as α, ϵ, and p increase. More specifically, increasing the value of ϵ and/or α results in a lower minimum in-control performance threshold and, consequently, in a lower AARL. On the other hand, a larger value of p means that we allow a larger proportion of the in-control ARLs to be below the minimum performance threshold. This has the consequence that the AARL will be smaller. Thus, increasing any of the parameters α, ϵ, and/or p leads to lower AARL values, which is beneficial for the out-of-control situation, but of course not for the in-control situation. This tradeoff between in-control and out-of-control performance is inherent to control charts. The advantage of our proposed method is that the parameters can be easily adjusted in order to balance the performance of the chart as desired by the practitioner. In addition, Table 8 illustrates the corrections, exceedance probabilities, and AARL values of the corrected and uncorrected charts for various combinations of m and n, for $\alpha = 0.01$, $\epsilon = 0.2$, and $p = 0.1$.

5. Comparison with Existing Methods

In order to illustrate the performance of the proposed correction term, a comparison is made with the existing methods. First, we make a comparison with the methods of Albers and Kallenberg (2005)

TABLE 6. Correction Terms c for the Individuals X-Chart, with the Corresponding Exceedance Probability and AARL for the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.01$, $p = 0.1$, and $\epsilon = 0.4$

m	c	Exc. pr. (cor)	Exc. pr. (unc)	AARL (cor)	AARL (unc)
50	0.3176	0.0979	0.3279	658	174
75	0.2127	0.0959	0.2773	301	140
100	0.1512	0.0959	0.2393	212	127
150	0.0808	0.0966	0.1837	151	117
200	0.0407	0.0970	0.1440	127	112
250	0.0142	0.0984	0.1156	114	109
500	−0.0483	0.0993	0.0409	91	105
1,000	−0.0898	0.1011	0.0066	79	102

TABLE 7. Out-of-Control Performance of the $\overline{X}$-Chart for Shifts in the Mean of Size $\delta/\sqrt{n}$, for Both the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.0027$, $p = 0.05$, and $\epsilon = 0.2$

		Size of shift δ					
		0.5		1		2	
m	n	AARL (cor)	AARL (unc)	AARL (cor)	AARL (unc)	AARL (cor)	AARL (unc)
25	3	143	69	42	23	6	4
	5	84	60	27	21	5	4
	9	64	56	22	20	4	4
50	3	74	58	24	20	4	4
	5	57	55	19	19	4	4
	9	48	53	17	18	3	4
75	3	60	56	20	19	4	4
	5	49	53	17	18	4	4
	9	44	52	16	18	3	4
100	3	53	54	18	19	4	4
	5	46	52	16	18	3	4
	9	41	52	15	18	3	4
150	3	47	52	17	18	3	4
	5	42	51	15	18	3	4
	9	39	51	14	18	3	4
200	3	44	52	16	18	3	4
	5	40	51	15	18	3	4
	9	38	51	14	18	3	4
250	3	43	51	15	18	3	4
	5	39	51	14	18	3	4
	9	37	51	14	18	3	4

and Gandy and Kvaløy (2013). Next, we compare the proposed X chart with tolerance intervals for a normal distribution because these use an equivalent criterion for $n = 1$. Finally, we compare the proposed control chart with the self-starting Q chart of Quesenberry (1993).

5.1. Comparison of Shewhart X and $\overline{X}$ Control Charts

We consider the two-sided case, with $\hat{\mu} = \overline{\overline{X}}$ and $\hat{\sigma}$ as in Equation (20) for $n > 1$ and Equation (22) for $n = 1$. For this situation, Albers and Kallenberg (2005) proposed control limits of the form

$$\widehat{\mathrm{UCL}}_{\mathrm{AK}} = \overline{\overline{X}} + K(\hat{\sigma}/\sqrt{n})(1 + c_{\mathrm{AK}})$$

$$\widehat{\mathrm{LCL}}_{\mathrm{AK}} = \overline{\overline{X}} - K(\hat{\sigma}/\sqrt{n})(1 + c_{\mathrm{AK}}), \qquad (29)$$

where

$$c_{\mathrm{AK}} = \frac{\Phi^{-1}(1-p)\theta}{\sqrt{mn}} - \frac{\epsilon}{K^2}, \qquad (30)$$

where $\theta^2 = \lim_{mn\to\infty}[mn\mathrm{var}(\hat{\sigma}/(\mathrm{E}\hat{\sigma}))]$. For the estimators $\hat{\sigma}$ as in Equations (20) and (22), the value of θ^2 equals $n/[2(n-1)]$ and 0.826, respectively. Note that their proposed correction (c_{AK}) only depends on the initial phase-I sample through its size (m and n). This is in line with our proposed correction. For the bootstrap approach of Gandy and Kvaløy (2013), this is different, as the actual correction depends on the sample estimates. Therefore, in the comparison,

TABLE 8. Out-of-Control Performance of the $\overline{X}$-Chart for Shifts in the Mean of Size $\delta/\sqrt{n}$, for Both the Corrected (Cor) and Uncorrected (Unc) Control Limits. Parameter values are $\alpha = 0.01$, $p = 0.1$, and $\epsilon = 0.4$.

		Size of shift δ					
		0.5		1		2	
m	n	AARL (cor)	AARL (unc)	AARL (cor)	AARL (unc)	AARL (cor)	AARL (unc)
25	3	143	69	42	23	6	4
	5	84	60	27	21	5	4
	9	64	56	22	20	4	4
50	3	74	58	24	20	4	4
	5	57	55	19	19	4	4
	9	48	53	17	18	3	4
75	3	60	56	20	19	4	4
	5	49	53	17	18	4	4
	9	44	52	16	18	3	4
100	3	53	54	18	19	4	4
	5	46	52	16	18	3	4
	9	41	52	15	18	3	4
150	3	47	52	17	18	3	4
	5	42	51	15	18	3	4
	9	39	51	14	18	3	4
200	3	44	52	16	18	3	4
	5	40	51	15	18	3	4
	9	38	51	14	18	3	4
250	3	43	51	15	18	3	4
	5	39	51	14	18	3	4
	9	37	51	14	18	3	4

only the realized exceedance probabilities are shown. For the explicit bootstrapping procedure, we refer to Saleh et al. (2015b), who provide a simplification of the computations in Gandy and Kvaløy's (2013) approach for the Shewhart control chart. We performed the bootstrap procedure (based on 1001 bootstraps) for 10,000 simulated phase-I samples, in order to calculate the exceedance probability. The considered parameter values are $p = 0.1$, $\epsilon = 0$, and $\alpha = 0.0027$. The results of the simulations are listed in Table 9 for different values of m and $n = 5$. Results for other values of n are similar. It is clear that our proposed correction performs much better than the correction of Albers and Kallenberg (2005), as it is closer to the desired level of $p = 0.1$. The bootstrap procedure of Gandy and Kvaløy (2013) also has good performance. There is no real difference with our proposed method in the sense of performance. Also, the performance of our proposed method and the bootstrap method appears to be less sensitive to the value of ϵ, as becomes clear when changing its value. This is shown in Table 10, which indicates the performance of the three methods when implementing $\epsilon = 0.2$ while leaving the other parameters as before.

To illustrate the performance and consequences of the proposed methods graphically, the distributions of the ARLs of the different methods in the in-control ($\delta = 0$) and out-of-control situation ($\delta = 1$) are shown in Figures 1 and 2, respectively, for $m = 50$,

TABLE 9. Exceedance Probabilities of the Proposed Correction, Albers and Kallenberg (2005) (AK) Correction and Gandy and Kvaløy (2013) (bootstrap) Method for Different Values of *m*. Parameter values are $n = 5$, $\alpha = 0.0027$, $p = 0.1$. and $\epsilon = 0$

m	Proposed	AK	Bootstrap
25	0.0916	0.1822	0.1046
50	0.0934	0.1580	0.0993
75	0.0957	0.1462	0.0989
100	0.0947	0.1440	0.1010
150	0.0967	0.1350	0.1054
200	0.0989	0.1318	0.0983
250	0.0998	0.1263	0.1021

TABLE 10. Exceedance Probabilities of the Proposed Correction, Albers and Kallenberg (2005) (AK) Correction and Gandy and Kvaløy (2013) (bootstrap) Method for Different Values of *m*. Parameter values are $n = 5$, $\alpha = 0.0027$, $p = 0.1$, and $\epsilon = 0.2$

m	Proposed	AK	Bootstrap
25	0.0921	0.1918	0.1044
50	0.0936	0.1739	0.0992
75	0.0973	0.1649	0.0989
100	0.0982	0.1673	0.1010
150	0.1008	0.1665	0.1054
200	0.0999	0.1662	0.0983
250	0.1004	0.1708	0.1022

$n = 5$, $p = 0.1$, $\epsilon = 0.2$, and $\alpha = 0.0027$. The vertical line represents the desired threshold of the $100p$'th (in this case 10th) percentile of the in-control ARL distribution. The desired threshold with $\alpha = 0.0027$ and $\epsilon = 0.2$ is equal to 296. As can also be seen in Table 10, our proposed correction and the bootstrap method perform best, with the 10th percentile close to the desired level. It is gratifying to note that, with our correction term, no extensive bootstrapping is needed for the Shewhart control charts for location.

The tradeoff between in-control and out-of-control performance of the control chart also becomes clear from these figures, if we compare the proposed methods with the uncorrected chart. The corrected charts correspond with better in-control performance, but have a slower detection in the out-of-control situation.

5.2 Tolerance Intervals

The literature of tolerance intervals considers a criterion that is closely related to the proposed corrections in this paper. From Krishnamoorthy and Mathew (2009), we cite "a tolerance interval is expected to capture a certain proportion or more of the population, with a given confidence". The tolerance intervals are based on the sample average ($\overline{X}$) and the sample standard deviation (s) and are in the form of $\overline{X} \pm k_2 s$, where k_2 is determined such that

$$P(P(\overline{X} - k_2 s \leq X \leq \overline{X} + k_2 s \mid \overline{X}, s) \geq 1 - \alpha) = 1 - p. \quad (31)$$

Using the tolerance limits as control limits guarantees an in-control FAR (ARL) that is smaller (larger) than α ($1/\alpha$) with probability $1-p$. Thus, k_2 is equivalent to our $\tilde{K}$ in the case that $\epsilon = 0$ and $n = 1$, when

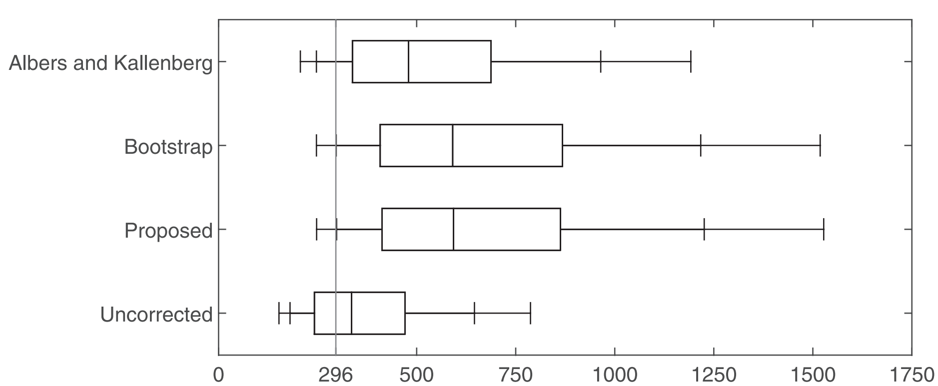

FIGURE 1. In-Control ARL Distribution when $m = 50$, $n = 5$, $\alpha = 0.0027$, $p = 0.1$, and $\epsilon = 0.2$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the distributions. The vertical line represents the desired threshold level of the ARL (296).

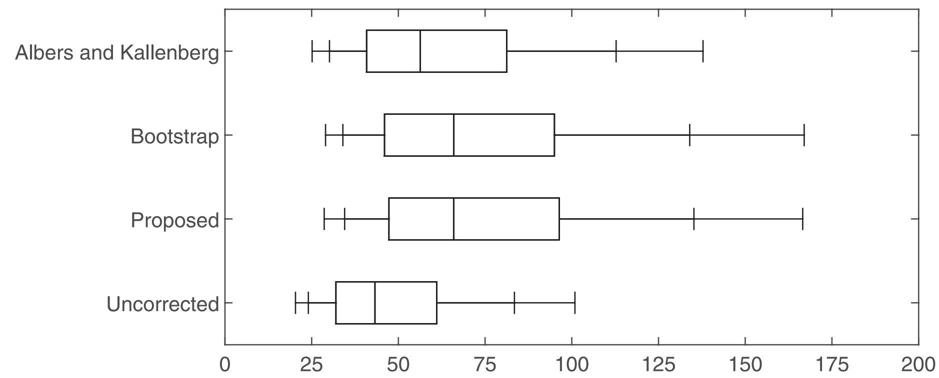

FIGURE 2. Out-of-Control ARL Distribution for $\delta = 1$ when $m = 50$, $n = 5$, $\alpha = 0.0027$, $p = 0.1$, and $\epsilon = 0.2$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the distributions.

using $\overline{X}$ and s as phase-I estimators. Only in this case, the objective is exactly identical to our approach. Note that the tolerance limit approach can be applied when $n > 1$, by treating $\overline{X}$ as an individual variable. However, this would mean that standard deviation of $\overline{X}$ should be used to estimate σ rather than the standard deviation of X. This means that the within-subgroup variation is ignored in the estimation. The performance of the tolerance-interval approach is then equal to the performance when $n = 1$. Because our approach does consider the within-subgroup variation, it has less uncertainty in parameter estimation, leading to less variation in ARLs for $n > 1$.

Because of the arguments mentioned above, we have compared our proposed corrections with the approximated tolerance factors as in Krishnamoorthy and Mathew (2009) for the case $n = 1$. Although they provide tolerance factors K_{KM} instead of corrections c_{KM}, we can determine their 'correction' by subtracting K from K_{KM}. We consider $\overline{X}$ as estimator of μ and s as estimator of σ. Their proposed correction c_{KM} then equals

$$c_{\mathrm{KM}} = \left(\frac{(n-1)\chi^2_{[1,1-\alpha]}(1/n)}{\chi^2_{[n-1,p]}} \right)^{1/2} - K, \qquad (32)$$

where $\chi^2_{[d,q]}$ represents the q-quantile of a chi-squared distribution with d degrees of freedom, and $\chi^2_{[d,q]}(\theta)$ represents the q-quantile of a noncentral chi-squared distribution with d degrees of freedom and noncentrality parameter θ.

For different combinations of α and p, we have calculated the corrections and their performance. The results are shown in Table 11 for $\alpha = 0.0027$ and $p = 0.05$ and in Table 12 for $\alpha = 0.01$ and $p = 0.1$. As can be seen, there appears to be no significant difference in performance because the resulting exceedance probabilities are close to the desired value p for both approximations. Note that, for other estimators of σ for $n = 1$, like the average moving range, the results of Krishnamoorthy and Mathew (2009) are not straightforward. Furthermore, note that, for one-sided control charts, for normally distributed data, exact solutions are available for the tolerance bounds.

5.3. Self-Starting Control Charts

There are also other control chart designs that lead to a desirable in-control performance. In particular, for normally distributed data, self-starting control charts by Hawkins (1987) and Quesenberry (1991) can guarantee a good in-control performance

TABLE 11. Corrections and Exceedance Probabilities of the Proposed Correction and the Approximated Tolerance from Krishnamoorthy and Mathew (2009) for Different Values of m and $n = 1$ (individuals). Parameter values are $\alpha = 0.0027$, $p = 0.05$, and $\epsilon = 0$

m	c	Exc. pr. (c)	c_{KM}	Exc. pr. (c_{KM})
50	0.6286	0.0507	0.6403	0.0484
75	0.4990	0.0514	0.4966	0.0521
100	0.4215	0.0503	0.4174	0.0519
150	0.3322	0.0473	0.3293	0.0486
200	0.2812	0.0458	0.2794	0.0469
250	0.2475	0.0505	0.2465	0.0514
500	0.1683	0.0515	0.1686	0.0512
1,000	0.1159	0.0529	0.1165	0.0523

TABLE 12. Corrections and Exceedance Probabilities of the Proposed Correction and the Approximated Tolerance from Krishnamoorthy and Mathew (2009) for Different Values of m and $n = 1$ (individuals). Parameter values are $\alpha = 0.01$, $p = 0.1$, and $\epsilon = 0$

m	c	Exc. pr. (c)	c_{KM}	Exc. pr. (c_{KM})
50	0.4130	0.1013	0.4249	0.0957
75	0.3257	0.1071	0.3304	0.1032
100	0.2753	0.1026	0.2781	0.1006
150	0.2179	0.0997	0.2196	0.0984
200	0.1851	0.1010	0.1865	0.0982
250	0.1634	0.1067	0.1646	0.1044
500	0.1118	0.1000	0.1127	0.0981
1,000	0.0773	0.0996	0.0779	0.0979

in the long run as well. However, the major drawback is that, because of the continuous updating of the control chart limits, there is a risk that out-of-control data influence the process estimates. A small change in the process mean can therefore slowly change the control limits with it, making the out-of-control situations harder to detect. This can result in larger out-of-control ARLs. To illustrate this, we have simulated ARLs for both the self-starting Q chart in Quesenberry (1991) and our proposed corrections for both the in-control and out-of-control situations (with $\delta = 1$). As can be observed in Figure 3, the self-starting Q chart indeed has a long out-of-control ARL, much longer than the proposed corrections. Next to that, although the in-control performance of the Q chart is very stable, our proposed correction yields a much longer in-control ARL, which is highly beneficial for practitioners.

6. Concluding Remarks and Recommendations

To deal with the effect of parameter estimation, we propose new correction terms for the Shewhart X and $\overline{X}$ control charts. As model parameters are generally unknown, they have to be estimated using a reference sample. Because different practitioners use different samples, their control chart limits and, consequently, their control-chart performance, will vary. The newly proposed correction terms are in line with the idea introduced by Albers and Kallenberg (2004), which corrects the control limits to guarantee a specified minimum performance of the control chart with a specified probability. The new correction terms are shown to be very accurate in achieving this. The performance of the proposed method is much better than Albers and Kallenberg (2005) and similar to the bootstrap method of Gandy and Kvaløy (2013). However, no bootstrapping is required, as the proposed correction only depends on the initial phase-I sample through its size rather than its parameter estimates. In this paper, also comparisons are made with tolerance intervals and self-starting control charts. The conclusions are that our corrections behave very well for the individuals Shewhart X control chart and outperform the self-starting control chart in both in- and out-of-control situations.

Because of the guarantee of minimum performance, the corrected chart performs better than an uncorrected chart in the in-control situation. This inevitably leads to a deterioration of the out-of-control

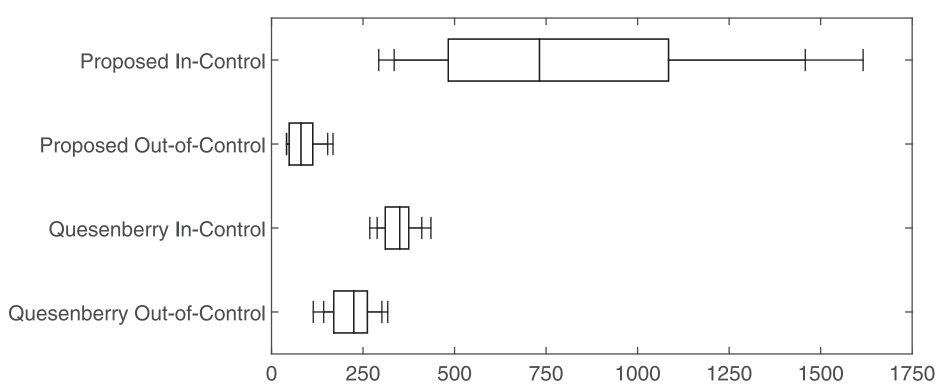

FIGURE 3. In-Control ($\delta = 0$) and Out-of-Control ($\delta = 1$) ARL Distribution of the Self-Starting Q Chart in Quesenberry (1991) and Our Proposed Correction when $m = 50$, $n = 5$, $\alpha = 0.0027$, $p = 0.1$, and $\epsilon = 0$. The boxplots indicate the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles of the distributions.

performance. However, the strictness of the correction can be easily adapted by changing α, ϵ, and p as desired. The choice of parameters should be based on the context. As costs of a false alarm and costs of running a process out-of-control are very dependent on the application, it is recommended to take this into account when making the tradeoff between in-control and out-of-control performance. An advantage of the proposed method is that this tradeoff can be taken into account when setting up the control limits rather than being disregarded.

The present paper considers the Shewhart X and $\overline{X}$ control charts under normal theory. In other settings (e.g., the CUSUM and EWMA control charts or autocorrelated processes), alternative methods to guarantee a minimum performance to practitioners are the subject for further investigation.

Appendix A

First note that $P_{mn}(K; Z, W)$ can be written according to Equation (3). This can be rewritten as

$$P_{mn}(K; \hat{\mu}, \hat{\sigma}) = \bar{\Phi}(K + \Delta_1(K)) + \bar{\Phi}(K + \Delta_2(K)), \tag{A.1}$$

with $\bar{\Phi}(x) = 1 - \Phi(x)$ and

$$\Delta_1(K) = \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + K\left(\frac{\hat{\sigma}}{\sigma} - 1\right) \tag{A.2}$$

and

$$\Delta_2(K) = -\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} + K\left(\frac{\hat{\sigma}}{\sigma} - 1\right). \tag{A.3}$$

Hence, for any function $g(\alpha)$, we can write

$$\begin{aligned} g(P_{mn}(K; \hat{\mu}, \hat{\sigma})) &= h(K + \Delta_1(K), K + \Delta_2(K)) \\ &= h(x, y). \end{aligned} \tag{A.4}$$

Using a two-step Taylor expansion, this is approximately equal to

$$\begin{aligned} g(P_{mn}) &= h(K + \Delta_1(K), K + \Delta_2(K)) \\ &\approx h(K,K) + h_x(K,K)\Delta_1(K) \\ &\quad + h_y(K,K)\Delta_2(K) \\ &\quad + \frac{1}{2}[h_{xx}(K,K)\Delta_1^2(K) \\ &\quad + 2h_{xy}(K,K)\Delta_1(K)\Delta_2(K) \\ &\quad + h_{yy}(K,K)\Delta_2^2(K)], \end{aligned} \tag{A.5}$$

where $h_x(K,K)$ and $h_y(K,K)$ are the first-order partial derivatives of $h(x,y)$ with respect to x and y, respectively, $h_{xx}(K,K)$ and $h_{yy}(K,K)$ are the second-order partial derivatives of $h(x,y)$ with respect to x and y, respectively, and $h_{xy}(K,K)$ equals the cross partial derivative of $h(x,y)$ with respect to x and y. Note that $h_x(K,K) = h_y(K,K)$ and $h_{xx}(K,K) = h_{yy}(K,K)$. Taking this into account, we can simplify (A.5) into

$$\begin{aligned} g(P_{mn}) &\approx h(K,K) + h_x(K,K)\left[\Delta_1(K) + \Delta_2(K)\right] \\ &\quad + \frac{1}{2}h_{xx}(K,K)\left[\Delta_1^2(K) + \Delta_2^2(K)\right] \\ &\quad + h_{xy}(K,K)\Delta_1(K)\Delta_2(K). \end{aligned} \tag{A.6}$$

Using Equations (A.2) and (A.3), we can rewrite this into

$$\begin{aligned} g(P_{mn}) &\approx h(K,K) + h_x(K,K)2K\left(\frac{\hat{\sigma}}{\sigma} - 1\right) \\ &\quad + (h_{xx}(K,K) - h_{xy}(K,K))\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}}\right)^2 \\ &\quad + (h_{xx}(K,K) + h_{xy}(K,K))K^2\left(\frac{\hat{\sigma}}{\sigma} - 1\right)^2. \end{aligned} \tag{A.7}$$

For X_{i_j} i.i.d. $N(\mu, \sigma)$-distributed random variables and $\hat{\mu} = \overline{\overline{X}}$, we know that

$$\left(\frac{\hat{\mu} - \mu}{\sigma/\sqrt{mn}}\right)^2$$

has a chi-squared distribution. Common estimators of the standard deviation, such as the (pooled) sample standard deviation, follow a scaled chi distribution, while even the distribution of the average moving range can be approximated as such (cf., Roes et al. (1993)). This means that $\hat{\sigma}/\sigma$ and $\hat{\sigma}^2/\sigma^2$ generally follow a scaled chi and scaled chi-squared distribution, respectively. Hence, we may conclude that $P_{mn}(K; Z, W)$ is approximately a combination of scaled chi and chi-squared distributed random variables. Note that the distribution of $P_{mn}(K; Z, W)$ should be approximated such that it not only gives an accurate description but also such that it is possible to obtain the required correction term. The difficulty in obtaining this correction term lies in the sense that the correction term c not only changes the expectation but also the variance of the distribution. As the currently obtained approximation is still of rather complicated form and, because the scaled chi-squared part appears to be dominant, we approximate the distribution of $P_{mn}(K; Z, W)$ by a $a\chi_b^2/b$ distribution, where we use the first two central moments of $P_{mn}(K; Z, W)$ to identify a and b.

Appendix B

In order to obtain the required correction term, we use the Wilson–Hilferty transformation (Wilson and Hilferty (1931)), which states that, for $X \sim \chi^2_b$, we have

$$\sqrt[3]{X/b} \overset{\text{approx}}{\sim} N(1-2/(9b), 2/(9b)).$$

This transformation is quite accurate, which was shown recently by Inglot (2010). Henceforth, we abbreviate $P_{mn}(K; Z, W)$ as P_{mn}. Then, in our case, when

$$P_{mn} \overset{\text{approx}}{\sim} a\chi^2_b/b$$

or similarly

$$\frac{b}{a}P_{mn} \overset{\text{approx}}{\sim} \chi^2_b,$$

we obtain

$$\sqrt[3]{\frac{P_{mn}}{a}} \overset{\text{approx}}{\sim} N(1-2/(9b), 2/(9b)).$$

This is equivalent to

$$\frac{\sqrt[3]{\frac{P_{mn}}{a}} - 1 + \frac{2}{9b}}{\sqrt{\frac{2}{9b}}} \overset{\text{approx}}{\sim} N(0,1).$$

We want to have $P(P_{mn} < (1+\epsilon)\alpha) = 1 - p$ (cf. Equation (7)). This is equivalent to

$$\begin{aligned} P(P_{mn} < (1+\epsilon)\alpha) &= P\left(\frac{\sqrt[3]{\frac{P_{mn}}{a}} - 1 + \frac{2}{9b}}{\sqrt{\frac{2}{9b}}} < \frac{\sqrt[3]{\frac{(1+\epsilon)\alpha}{a}} - 1 + \frac{2}{9b}}{\sqrt{\frac{2}{9b}}}\right) \\ &\approx \Phi\left(\frac{\sqrt[3]{\frac{(1+\epsilon)\alpha}{a}} - 1 + \frac{2}{9b}}{\sqrt{\frac{2}{9b}}}\right) = 1 - p, \end{aligned} \quad \text{(B.1)}$$

which in turn, by using the inverse of the standard normal CDF (denoted Φ^{-1}), leads to the following equation that needs to be solved:

$$\frac{\sqrt[3]{\frac{(1+\epsilon)\alpha}{a}} - 1 + \frac{2}{9b}}{\sqrt{\frac{2}{9b}}} = \Phi^{-1}(1-p). \quad \text{(B.2)}$$

Note again that both a and b are functions of K. Given the values of m, n, α, and ϵ, the left-hand side of Equation (B.2) is a function of K only, say $Y(K)$. Using Equations (12) and (13), we can write $Y(K)$ as

$$\begin{aligned} Y(K) = \sqrt[3]{(1+\epsilon)\alpha}\frac{3\mathrm{E}(P_{mn})^{2/3}}{\sqrt{\mathrm{Var}(P_{mn})}} \\ - \frac{3\mathrm{E}(P_{mn})}{\sqrt{\mathrm{Var}(P_{mn})}} + \frac{\sqrt{\mathrm{Var}(P_{mn})}}{3E(P_{mn})} \end{aligned} \quad \text{(B.3)}$$

In order to solve Equation (B.2,) we need to find c such that, for $\tilde{K} = K + c$, there holds $Y(\tilde{K}) = \Phi^{-1}(1-p)$. This value of c is found by a linear approximation of $Y(\tilde{K})$ as $Y(\tilde{K}) \approx Y(K) + c(dY(K)/dK)$.

If we denote the derivatives of $E(P_{mn})$ and $V(P_{mn})$ with respect to K by dE/dK and dV/dK, respectively, and $E(P_{mn})$ and $V(P_{mn})$ by E and V, respectively, we have

$$\begin{aligned} Y'(K) &= \frac{dY(K)}{dK} \\ &= \frac{2E^{-1/3}\sqrt{V}\frac{dE}{dK} - \frac{3E^{2/3}}{2\sqrt{V}}\frac{dV}{dK}}{V} \\ &\quad - \frac{3\frac{dE}{dK}\sqrt{V} - \frac{3E}{2\sqrt{V}}\frac{dV}{dK}}{V} \\ &\quad + \frac{\frac{3E}{2\sqrt{V}}\frac{dV}{dK} - 3\frac{dE}{dK}\sqrt{V}}{9E^2}. \end{aligned} \quad \text{(B.4)}$$

The values of dE/dK and dV/dK can be calculated (as can be seen from Equations (9), (11), and (10)) as

$$\begin{aligned} \frac{dE}{dK} &= \int_{-\infty}^{\infty}\int_0^{\infty} \frac{dP_{mn}(z,w)}{dK} f(z)f(w)dwdz \\ &= \int_{-\infty}^{\infty}\int_0^{\infty} -w\left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right)\right. \\ &\quad \left. + \phi\left(\frac{z}{\sqrt{m}} - Kw\right)\right] \\ &\quad \times f(z)f(w)dwdz \end{aligned} \quad \text{(B.5)}$$

and

$$dV/dK = d(E^2)/dK - 2E(dE/dK), \quad \text{(B.6)}$$

where

$$\begin{aligned} \frac{d(E^2)}{dK} &= \int_{-\infty}^{\infty}\int_0^{\infty} \frac{dP^2_{mn}(K;z,w)}{dK} f(z)f(w)dwdz \\ &= -2\int_{-\infty}^{\infty}\int_0^{\infty} wP_{mn}(K;z,w) \\ &\quad \times \left[\phi\left(\frac{z}{\sqrt{m}} + Kw\right)\right. \\ &\quad \left. + \phi\left(\frac{z}{\sqrt{m}} - Kw\right)\right] \\ &\quad \times f(z)f(w)dwdz. \end{aligned} \quad \text{(B.7)}$$

Getting back to the approximation $Y(\tilde{K}) \approx Y(K) + c(dY(K)/dK) = Y(K) + cY'(K)$, we thus obtain

$$c = [\Phi^{-1}(1-p) - Y(K)]/Y'(K). \quad \text{(B.8)}$$

Acknowledgments

The authors are very grateful for the remarks and suggestions made by the referees and the Editor. They have substantially improved the paper.

References

Albers, W. and Kallenberg, W. C. M. (2004a). "Are Estimated Control Charts in Control?" *Statistics* 38(1), pp. 67–79.

Albers, W. and Kallenberg, W. C. M. (2004b). "Estimation in Shewhart Control Charts: Effects and Corrections". *Metrika* 59, pp. 207–234.

Albers, W. and Kallenberg, W. C. M. (2005). "New Corrections for Old Control Charts". *Quality Engineering* 17, pp. 467–473.

Chen, G. (1997). "The Mean and Standard Deviation of the Run Length Distribution of $\overline{X}$ Charts when Control Limits Are Estimated". *Statistica Sinica* 7(3), pp. 789–798.

Cryer, K. D. and Ryan, T. P. (1990). "The Estimation of Sigma for an X Chart: $\overline{\mathrm{MR}}/d_2$ or S/c_4?" *Journal of Quality Technology* 22, pp. 187–192.

Gandy, A. and Kvaløy, J. T. (2013). "Guaranteed Conditional Performance of Control Charts via Bootstrap Methods". *Scandinavian Journal of Statistics* 40, pp. 647–668.

Gardiner, D. A. and Hull, N. C. (1966). "An Approximation to Two-Sided Tolerance Limits for Normal Populations". *Technometrics* 8(1), pp. 115–122.

Hawkins, D. M. (1987). "Self-Starting CUSUM Charts for Location and Scale". *The Statistician*, pp. 299–316.

Howe, W. G. (1969). "Two-Sided Tolerance Limits for Normal Populations—Some Improvements". *Journal of the American Statistical Association* 64(326), pp. 610–620.

Inglot, T. (2010). "Inequalities for Quantiles of the Chi-Square Distribution". *Probability and Mathematical Statistics* 30(2), pp. 339–351.

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(4), pp. 349–364.

Johnson, N. L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*, I. New York, NY: Wiley.

Jones, M. A. and Steiner, S. H. (2012). "Assessing the Effect of Estimation Error on the Risk-Adjusted CUSUM Chart Performance". *International Journal for Quality in Health Care* 24, pp. 176–181.

Krishnamoorthy, K. and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation*, I. New York, NY: Wiley.

Nedumaran, G. and Pignatiello, J. J., Jr. (2001). "On Estimating $\overline{X}$ Control Chart Limits". *Journal of Quality Technology* 33(2), pp. 206–212.

Patnaik, P. B. (1950). "The Use of Mean Range as an Estimator in Statistical Tests". *Biometrika* 37, pp. 78–87.

Psarakis, S.; Vyniou, A. K.; and Castagliola, P. (2014). "Some Recent Developments on the Effects of Parameter Estimation on Control Charts". *Quality and Reliability Engineering International* 30(8), pp. 641–650.

Quesenberry, C. P. (1991). "SPC Q Charts for Start-Up Processes and Short or Long Runs". *Journal of Quality Technology* 23(3), pp. 213–224.

Quesenberry, C. P. (1993). "The Effect of Sample Size on Estimated Limits for $\overline{X}$ and X Control Charts". *Journal of Quality Technology* 25(4), pp. 237–247.

Roes, K. C. B.; Does, R. J. M. M.; and Schurink, Y. (1993). "Shewhart-Type Control Charts for Individual Observations". *Journal of Quality Technology* 25(3), pp. 188–198.

Saleh, N. A.; Mahmoud, M. A.; Keefe, M. J.; and Woodall, W. H. (2015a). "The Difficulty in Designing Shewhart $\overline{X}$ and X Control Charts with Estimated Parameters". *Journal of Quality Technology* 47(2), pp. 127–138.

Saleh, N. A.; Mahmoud, M. A.; Jones-Farmer, L. A.; Zwetsloot, I.; and Woodall, W. H. (2015b). "Another Look at the EWMA Control Chart with Estimated Parameters". *Journal of Quality Technology* 47(4), pp. 363–382.

Tsai, T. R.; Wu, S. J.; and Lin, H. C. (2004). "An Alternative Control Chart Approach Based on Small Number of Subgroups". *International Journal of Information and Management Sciences* 15(4), pp. 61–73.

Tsai, T. R.; Lin, J. J.; Wu, S. J.; and Lin, H. C. (2005). "On Estimating Control Limits of X-Bar Chart when the Number of Subgroups Is Small". *International Journal of Advanced Manufacturing Technology* 26(11–12), pp. 1312–1316.

Wald, A. and Wolfowitz, J. (1946). "Tolerance Limits for a Normal Distribution". *The Annals of Mathematical Statistics* pp. 208–215.

Weissberg, A. and Beatty, G. H. (1960). "Tables of Tolerance-Limit Factors for Normal Distributions". *Technometrics* 2(4), pp. 483–500.

Wilson, E. B. and Hilferty, M. M. (1931). "The Distribution of Chi-Squared". *Proceedings of the National Academy of Sciences of the United States of America* 17(12), pp. 684–688.