UvA-DARE (Digital Academic Repository)

On the Design of Control Charts with Guaranteed Conditional Performance under Estimated Parameters

Does, R.J.M.M.; Goedhart, R.; Woodall, W.H.

Citation for published version (APA):

WILEY

On the design of control charts with guaranteed conditional performance under estimated parameters

Ronald J.M.M. Does[1] | Rob Goedhart[1] | William H. Woodall[2]

[1]Department of Operations Management, Amsterdam Business School, University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Statistics, Virginia Tech, Blacksburg, Virginia, USA

Correspondence
Ronald J. M. M. Does, Department of Operations Management, Amsterdam Business School, University of Amsterdam, 1001NL Amsterdam, The Netherlands.
Email: r.j.m.m.does@uva.nl

Abstract

When designing control charts the in-control parameters are unknown, so the control limits have to be estimated using a Phase I reference sample. To evaluate the in-control performance of control charts in the monitoring phase (Phase II), two performance indicators are most commonly used: the average run length (ARL) or the false alarm rate (FAR). However, these quantities will vary across practitioners due to the use of different reference samples in Phase I. This variation is small only for very large amounts of Phase I data, even when the actual distribution of the data is known. In practice, we do not know the distribution of the data, and it has to be estimated, along with its parameters. This means that we have to deal with model error when parametric models are used and stochastic error because we have to estimate the parameters. With these issues in mind, choices have to be made in order to control the performance of control charts. In this paper, we discuss some results with respect to the in-control guaranteed conditional performance of control charts with estimated parameters for parametric and nonparametric methods. We focus on Shewhart, exponentially weighted moving average (EWMA), and cumulative sum (CUSUM) control charts for monitoring the mean when parameters are estimated.

KEYWORDS

cumulative sum, CUSUM control chart, EWMA control chart, exponentially weighted moving average, nonparametric, parameter estimation, Phase I, Phase II, Shewhart control chart

1 | INTRODUCTION

Statistical process monitoring (SPM) is an area, which provides statistical tools for monitoring data streams, with the goal of detecting changes in the underlying processes. In 1924, Walter A. Shewhart proposed the first statistical tool to address these problems (cf. Shewhart[1]). He named it the control chart, which is 'a form which might be modified from time to time, in order to give at a glance the greatest amount of accurate information'. The control chart is used to

This paper has been based on the Box Medal speech given by the first author.

detect changes in the parameters of the underlying distribution of the process characteristic. Shewhart proposed the well-known control chart based on three-sigma limits as a baseline.

Historically, a fundamental component of SPM research has been the control of the false alarm rate (FAR) and the in-control average run length (ARL), that is, the average of the number of samples that must be taken before the control chart gives a false alarm. Other metrics that reflect the occurrence of false alarms have also been considered. False alarms can lead to unneeded distraction, wasted time and effort, and increased variability. An excessive number of false alarms can lead practitioners to ignore signals from the monitoring method altogether.

The choice of three-sigma limits implies that the FAR equals 0.0027 for normally distributed data with known in-control parameters. The corresponding in-control ARL equals 370.4, because of the relation ARL = 1/FAR. In practice, the in-control parameters μ_0 and σ_0 of the normal distribution should be estimated based on a Phase I sample, which consists according to the standard procedure of about $m = 20 - 40$ subgroups of about $n = 5$ units each. Most often, the estimator of the mean is the overall sample mean and of the standard deviation the mean of the subgroup ranges divided by an unbiasing constant. The estimated control limits are read as follows: $\bar{\bar{X}} \pm A_2 \bar{R}$, where the constant A_2 may found in Montgomery.[2] Remarkably, the constant A_2 only depends on the subgroup size (n) and not on the number of subgroups (m). In particular, $A_2 = 3/(d_2(n)\sqrt{n})$, where $d_2(n)$ is an unbiasing constant for the sample range (cf. Does and Schriever[3] for a discussion on these constants).

It was noted by Quesenberry[4] that when parameters are estimated, the unconditional run length (RL) distribution is not geometric, as was previously assumed. Quesenberry[4] concluded that more samples, compared with the standard advice of $m = 20 - 40$ subgroups of about $n = 5$ units each, are required in order to let the control chart behave on average as if parameters are known. He concluded that around $m = 300$ observations are required when $n = 1$ and around $m = 400/(n-1)$ subgroups for $n > 1$. Similar recommendations were found in Chen,[5] among others. However, for this recommendation, Quesenberry[4] did not consider the sampling variation between practitioners. Since different practitioners obtain different Phase I samples, their estimates and corresponding control limits will vary. As a consequence, the control chart performance in terms of the FAR or ARL will also vary. This effect is extensively discussed by Saleh et al.[6] and is referred to as sampling or practitioner-to-practitioner variability. This variation becomes less severe when the number of samples and their size increase, but the original sample size suggestions from Quesenberry[4] are not sufficient by a wide margin to guarantee a performance equivalent to that of the known parameters situation. For S^2 and S control charts, for example, Epprecht et al.[7] concluded that the total sample size requirements are often closer to several thousands instead of several hundreds.

In many situations, the sample size requirements such as indicated in Epprecht et al.[7] may not be feasible. Albers and Kallenberg[8–10] suggested two criteria for adjusting the control limits. The first they called the bias criterion which focuses on the unconditional RL distribution properties. Conditional on the estimates of the control limits, the conditional RL properties are again geometric with parameter equal to the conditional FAR (CFAR). The properties of the conditional RL distribution for these estimates can then be calculated as a function g of CFAR. For example, the conditional in-control ARL (CARL) is equal to CARL = 1/CFAR (i.e., $g(x) = 1/x$). Another option for the function g is the conditional standard deviation of the RL defined by $SDCARL = \sqrt{1-CFAR}/CFAR$ (i.e., $g(x) = \sqrt{1-x}/x$). The unconditional equivalent of these properties is then equal to the expectation of these measures. While it may seem logical to aim for a certain performance in expectation, one should be aware that there could still be a large probability of an unsatisfactory control chart performance for individual practitioners due to practitioner-to-practitioner variation.

When the process is in-control, large FAR values (or low ARL values) are undesirable, as that would mean that the control chart produces many false signals. Thus, the bias criterion aims to adjust the control limits such that the control chart provides a specified in-control RL property (such as the FAR or ARL) in expectation. As an alternative design criterion, Albers and Kallenberg[8–10] proposed the exceedance probability criterion, which aims to provide a specified minimum in-control performance with a specified large probability by focusing on the conditional performance. In our view, this criterion is much more useful. In this paper, we discuss some results with respect to this criterion because it much better reflects the desired in-control performance of a control chart. It means that for some prespecified threshold ARL_0 and probability p, we require $P(CARL < ARL_0) = p$ or equivalently $P(CARL > ARL_0) = 1 - p$. Note that in practice, we take p to be relatively small, such as 0.05 or 0.1 so that with high probability, we have at least an in-control ARL of ARL_0 for the control chart when we adjust the control limits accordingly. Our paper will be restricted to the Shewhart, exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) control charts for monitoring the mean when parameters are estimated.

This paper is organized as follows. In Section 2, we give an overview of the results obtained for the Shewhart control chart. The following two sections describe the results for the EWMA and the CUSUM control charts, respectively. In

Section 5, we discuss a head to head comparison of the control charts discussed in this paper. In Section 6, we present some other developments in the literature of control charts with estimated parameters. The paper concludes with a summary and some recommendations for further research.

2 | SHEWHART CONTROL CHARTS

Consider the standard Shewhart control chart to monitor the mean of a variable, which uses both a lower control limit (LCL) and an upper control limit (UCL). If we assume normality of the data with in-control process parameters equal to μ_0 for the mean and σ_0 for the standard deviation, then the control limits are defined by

$$UCL = \mu_0 + K\frac{\sigma_0}{\sqrt{n}}$$

$$LCL = \mu_0 - K\frac{\sigma_0}{\sqrt{n}},$$

where K represents the control limit coefficient (e.g., $K = 3$ corresponds to three-sigma limits). In practice, the values of μ_0 and σ_0 are unknown and need to be replaced by some estimates $\hat{\mu}_0$ and $\hat{\sigma}_0$, respectively. This results in the estimated control limits

$$U\hat{C}L = \hat{\mu}_0 + \frac{K\hat{\sigma}_0}{\sqrt{n}}$$

$$L\hat{C}L = \hat{\mu}_0 - \frac{K\hat{\sigma}_0}{\sqrt{n}}.$$

Many choices of estimators are available, with the overall sample mean being the common estimator for location, and the pooled sample standard deviation and average moving range common estimators for dispersion. If the Phase I sample contains outliers, it can lead to a poor estimation of the in-control process behavior. An approach to address this problem is to use robust estimators, as described in Schoonhoven et al.[11–13]

As discussed in Section 1, the choice of the control limit constant K influences the control chart performance. Due to the practitioner-to-practitioner variation, this performance may vary for different practitioners. Albers and Kallenberg[8–10] provided two criteria that can be used to determine the required adjusted value for K: the bias criterion and the exceedance probability criterion as introduced in Section 1. They also provided solutions to obtain these required values for one-sided control charts based on various Phase I estimators. Both criteria have been adopted by other researchers as well. Results of control limit adjustments for Shewhart type control charts based on the bias criterion can be found in Tsai et al.[14], Goedhart et al.[15] and Diko et al[16].

The exceedance probability criterion has gained much more attention in the recent literature and has been promoted recently by various authors, such as Jones and Steiner,[17] Gandy and Kvaløy[18], Faraz et al.[19], Saleh et al.[6,20], and Goedhart et al[21–24]. When considering a standard Shewhart control chart for individual observations, the exceedance probability criterion leads to a statistical tolerance interval (see e.g., Krishnamoorthy and Mathew[25]). This property has also been used to develop and compare the methods in Goedhart et al.[21,23,24] This means that, for many different data distributions, existing tolerance interval literature can be used to provide analytical solutions for the required control limit coefficient in order to match the exceedance probability criterion. Numerical solutions are also available, such as the bootstrap method in Gandy and Kvaløy.[18] They provided a general bootstrap approach that has been applied to Shewhart, CUSUM, and EWMA control charts, as well as to quite a few other control charts.

Many methods for meeting the exceedance probability criterion are based on the assumption of normally distributed data. For example, it is obvious that the symmetrical control limits considered earlier will not work well for highly skewed data. This means that more general parametric and/or nonparametric methods are required that allow for different kinds of data distributions. However, the most appropriate data distribution is generally unknown and has to be estimated along with its parameters. This issue was discussed by Albers and Kallenberg,[26] who divided the total

estimation error in two distinct factors: the *model error* and the *stochastic error*. The first is caused by incorrect model assumptions, such as distributional form, while the latter is caused by estimation uncertainty and is as such related to the sample size and number of Phase I samples under consideration.

Several possibilities have been considered in the literature to reduce the model error. One example, already mentioned, is to apply for individual observations the statistical tolerance intervals for other distributions (see Krishnamoorthy and Mathew[25]). Another example is based on the application of the central limit theorem (CLT), as discussed in Huberts et al.[27] When considering subgroup averages instead of individual observations, it is expected that the normality assumption is more appropriate. However, the authors show that convergence to normality takes much longer than anticipated for some data distributions, especially when they are highly skewed or heavy tailed. Other parametric alternatives are to consider data transformations such as the ones suggested in Box and Cox[28] or Chou et al.[29] or to use other general parametric models such as the Pearson system of distributions. However, as shown in Goedhart et al.[24], these methods yield similar conclusions as the application of CLT, due to the interest of SPM in using control limits that fall into the far end of the tails of the distribution. To this end, several researchers have considered nonparametric control charts. For such a control chart, the model error vanishes, and the total estimation error is determined by the stochastic error.

For a general literature overview of nonparametric control charts, we refer to Qiu.[30] When it comes to the exceedance probability criterion, recall that the tolerance interval literature can be used to construct a nonparametric control chart satisfying this criterion. Such an interval is available from Krishnamoorthy and Mathew.[25] However, the downside of this interval is that it is discrete in nature due to the use of order statistics. Because of this, the actual coverage probability of an interval may be much greater than the minimum desired coverage, and the minimum required amount of data may be quite large. Young and Mathew[31] provided an alternative method based on interpolated or extrapolated order statistics to circumvent this issue and provided tolerance intervals with a coverage probability closer to the desired value. This theory can be applied to construct nonparametric control charts satisfying the exceedance probability criterion as well, as was done in Goedhart et al.[24] A major advantage of nonparametric control charts is that they can be applied to any monitoring statistic of interest, by treating subgroup statistics as individual observations. In that way, the proposed control chart can be applied to X, X-bar, R, and S control charts or even other statistics, regardless of the distribution under consideration.

3 | EWMA CONTROL CHARTS

The EWMA control chart was introduced by Roberts.[32] The charting statistic W_i of the EWMA chart is defined as

$$W_i = \lambda Y_i + (1-\lambda)W_{i-1} \qquad (3.1)$$

where $0 < \lambda \leq 1$ is the weight assigned to Y_i, which is the mean of ith sample of Phase II. The chart's initial starting value W_0 is usually set at μ_0, when we using the EWMA control chart for monitoring the mean. Furthermore, note that for $\lambda = 1$, the Shewhart control chart is a special case of the EWMA control chart.

In this section, we use the estimators $\hat{\mu}_0 = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}$, the overall sample mean, and $\hat{\sigma}_0 = \sqrt{\frac{1}{m}\sum_{i=1}^{m} S_i^2} = S_p$, the pooled standard deviation estimator, where S_i^2 denotes the variance of the ith Phase I sample, and m is the number of samples. Robust estimators are discussed in Zwetsloot et al.[33] Among the commonly used estimators for σ_0, the pooled standard deviation estimator can provide the lowest values of the mean squared error (Mahmoud et al.[34]). In addition in Diko et al.[16], the corresponding unbiased version $\hat{\sigma}_0 = S_p/c_4(m(n-1)+1)$ (see Montgomery[2]) is equivalent, because, $m(n-1)$ is typically quite large in our applications, and hence, the constant $c_4(m(n-1)+1)$ is almost indistinguishable from 1. For simplicity, we use the asymptotic (steady state) control limits:

$$\mu_0 \pm L\sqrt{\frac{\lambda}{n(2-\lambda)}}\sigma_0, \qquad (3.2)$$

where L is the charting constant to be specified. For example, for a given value of λ and a nominal ARL_0, assuming that the parameters are known, the L values can be found in Crowder[35] or by using the function *xewma.crit(λ,ARL_0, sided = 'two')* from the R-package *spc* (cf. Knoth[36]). Often, these parameters-known L values are used to construct the Phase II EWMA chart when estimated parameters are used in the control limits. It is recognized in the literature (cf. Jensen et al.[37]) that this is a problem in the sense of having many more false alarms than nominally expected, particularly when the amount of Phase I data is small to moderately large.

The performance of the EWMA control chart with estimated parameters was first investigated by Jones et al.[38], who derived the RL distribution of the chart. They showed that the EWMA chart performance deteriorates substantially when parameters are estimated, particularly with small amounts of Phase I data. The next step was to solve this problem. In Saleh et al.[20], the work of Jones et al.[38] was extended by evaluating the performance of the EWMA chart with estimated parameters while considering the practitioner-to-practitioner variability using the standard deviation of the ARL (SDARL) metric. Because it has been shown that the standard deviation estimator has a strong effect on control chart performance (Saleh et al.[6]), they also evaluated several estimators. The estimator based on the pooled standard deviation turned out to be the best. Regarding the recommendations of the required amount of Phase I data, they suggested to adjust the control limits to guarantee a minimal performance of the in-control ARL value of an EWMA chart. The (conditional) performance metrics were calculated using the Markov chain approach (cf. Brook and Evans[39] and Lucas and Saccucci,[40] among others).

Quite a few authors have used the SDARL as a metric for determining the necessary amount of Phase I data for control charts with estimated parameters (see e.g., Jones and Steiner,[17] Keefe et al.[41], Saleh et al.[6], and Faraz et al.[19]). These studies invariably show that impractically large amounts of Phase I data are needed for a practitioner to have confidence that her/his in-control ARL is near the desired value. This was the reason that Albers and Kallenberg[8–10] introduced the exceedance probability criterion to guarantee a minimum performance in terms of the CARL.

Saleh et al.[20] used the bootstrap-based design approach of Jones and Steiner[17] and Gandy and Kvaløy,[18] which was proposed for controlling the probability of the in-control CARL being at least a specified value. Their results show that adjusting the EWMA control limits accordingly can result in a highly skewed in-control CARL distribution. However, such increases in the in-control CARL did not have much of an effect on the out-of-control performance of the chart. In their opinion, this bootstrap design approach is very promising and should be considered while evaluating and comparing control charts.

In Diko et al.[42], a different approach was used to guarantee a minimum performance of the conditional in-control ARL. Their method to adjust the control chart limits according to the exceedance probability criterion is a better alternative to parametric bootstrapping. Like the bootstrap method, their method involves, in a different way, simulations to obtain the empirical distribution of the in-control CARL, which is accompanied by a search algorithm for finding the adjusted constant L in (3.2). However, unlike the bootstrap method, the search algorithm is bounded. This makes the method faster than parametric bootstrapping. Based on their new method, tables and graphs of the adjusted constants were provided over a wide range of chart parameters (cf. Diko et al.[42]). This will help with the implementation of the exceedance probability criterion based adjusted charts in practice, since these constants are currently unavailable in SPM textbooks or in computer software programs. An in-control and out-of-control performance evaluation with these adjusted limits for the EWMA control chart was also presented in Diko et al.[42] It was shown that the adjusted constants guarantee (with high probability) a specified minimum in-control CARL performance at a marginal cost of worse out-of-control performance. They showed how the in-control and out-of-control performance trade-offs can be done. Moreover, because of a different way that they run the simulations, the method of Diko et al.[42] produces more accurate and reliable results compared with the bootstrap method. The bootstrap method results in different adjusted limits with each run and thus the practitioner should run the bootstrap approach say 100 times and average the results (cf. Hu and Castagliola[43]).

4 | CUSUM CONTROL CHARTS

Page[44] introduced the CUSUM control chart for monitoring the process mean. The tabular form of the CUSUM chart was developed by Ewan and Kemp.[45] Like the EWMA control chart, the CUSUM control chart is also based on past and current information. Due to this feature, CUSUM and EWMA control charts are more efficient to detect small and

moderate sized shifts than Shewhart charts (cf. Hawkins and Wu[46]). The charting statistics of the two-sided CUSUM control chart are given by

$$D_i^- = min\left(0, D_{i-1}^- + Y_i - \mu_0 + k^- \sigma_0\right)$$
$$D_i^+ = max\left(0, D_{i-1}^+ + Y_i - \mu_0 - k^+ \sigma_0\right)$$

(4.1)

where D_0^+ and D_0^- are the starting values, while k^+ and k^- are the reference values. The starting values are assumed to be equal to 0. The variable Y_i is the ith sample statistic of Phase II. Note that in this case, Y_i may be the mean of sample i or the ith individual observation of Phase II. Of course when Y_i is the sample mean, we have to divide σ_0 by $\sqrt{n}$. The control limits for the two-sided CUSUM chart are given as

$$LCL = -h\sigma_0 \text{ and } UCL = h\sigma_0,$$

(4.2)

where h is the charting constant to be specified. In practice, the parameters μ_0 and σ_0 have to be estimated based on a Phase I sample. We will use the same estimators for μ_0 and σ_0 as we have used for the EWMA control chart. When we are using individual observations (i.e., sample size equals 1), then we use the average moving range estimator as the estimator for σ_0 (cf. Montgomery[2]). Robust estimators are discussed in Nazir et al.[47]

With these estimates, the monitoring Phase II can start. We assume again that the Phase II data Y_i are normally distributed. Traditionally, the conditional in-control ARL values of CUSUM control charts have been calculated by simulations (cf. Nazir et al.[47]) and Markov Chains (cf. Brook and Evans,[39] Crosier,[48] and Woodall[49]). For the parameters known case, the h values can also be found by specifying ARL$_0$ and k in the function *xcusum.crit(k,ARL$_0$, sided = 'two')* of the R-package *spc* (cf. Knoth[36]). Note that, for the CUSUM chart there exists an accurate ARL approximation formula (cf. Reynolds,[50] Siegmund[51]). Recently, Jeske[52] presented a modified Siegmund formula for approximating the conditional in-control ARL of the upper one-sided CUSUM chart. In Diko et al.,[53] this formula is extended to the two-sided CUSUM chart, and a comparison was made with the results from the well-known Markov Chain method (cf. Jones et al.[54]). It was shown that the generalization of the Siegmund formula to obtain the CARL for the two-sided CUSUM control chart is very accurate and practical when the CUSUM chart is used to detect small and moderate shifts.

In Saleh et al.,[55] the work of Jones et al.[54] was generalized by evaluating the performance of the CUSUM control chart with estimated parameters, while considering the practitioner-to-practitioner variability using the standard deviation of the conditional in-control ARL metric. This standard deviation indicates the amount of spread of the conditional in-control ARL values around the average of these values. However, a high value could be due to either lower than expected conditional in-control ARL values or due to higher than expected conditional in-control ARL values. Ideally, the number of low ARL values is minimized and the number of high ARL values is maximized. Therefore, in Diko et al.[53] another approach has been used based on the exceedance probability criterion. In Section 3 we have reviewed that in Diko et al.[42] the conditional in-control ARL distribution of the two-sided EWMA chart was approximated by an empirical distribution. This empirical distribution was obtained by generating many Phase I subgroups, calculating the corresponding conditional in-control ARL values by using the Markov Chain method, and then ordering them in ascending order. In Diko et al.,[53] this approach has been adopted for the two-sided CUSUM chart. However, they use the Siegmund formula instead of the Markov Chain method to calculate the conditional in-control ARL values (cf. Siegmund[51]). Once the empirical distribution has been found, its percentiles can be calculated. These percentiles can be used as an upper prediction bound, which can be compared with the nominal ARL$_0$. Recall that ARL$_0$ is the specified value that must be exceeded with high probability $1 - p$, according the exceedance probability criterion. Based on these prediction bounds, it was seen that even more Phase I data are required than previously recommended using the standard deviation criterion in Saleh et al.[55]

Because unrealistically large amounts of Phase I data are needed for practitioners to be confident of having an in-control ARL value near a specified value, Saleh et al.[55] recommended designing the charts with the bootstrap approach of Jones and Steiner[17] and Gandy and Kvaløy.[18] They illustrated the usefulness of this approach when designing the CUSUM charts to monitor the mean of a process under the assumption of normality. However, because the conditional in-control ARL is controlled to a greater extent than with the standard approach, the chart will be a little slower in detection out-of-control situations. In Diko et al,[53] a more accurate method is introduced to adjust the control limits according to the exceedance probability criterion. Comprehensive tables of the adjusted charting constants are given to facilitate implementation of the two-sided CUSUM control chart for the mean.

5 | HEAD-TO-HEAD COMPARISON OF CONTROL CHARTS ADJUSTED FOR PARAMETER ESTIMATION

In Sections 2, 3, and 4, we discussed three commonly used types of control charts, namely, the Shewhart, EWMA, and CUSUM control charts. Each of these charts has their own characteristics, which make them more or less applicable to detect certain types of shifts. For example, the Shewhart control chart is better suited to detect large shifts, while the CUSUM and EWMA control charts yield better detection capabilities against small sustained shifts. These aspects increase the demand for comparative studies, where the control chart capabilities are evaluated based on different scenarios. All existing comparisons are under normal theory.

Hawkins and Wu[46] performed a comparative study of the CUSUM and EWMA control charts. They found that the CUSUM outperforms the EWMA in the case that the actual shift to be expected is (approximately) known in size. However, their comparison is based on the assumption of known in-control process parameters. In practice, these parameters have to be estimated using a Phase I reference sample. As noted before, different practitioners obtain different Phase I samples, and thus, their parameter estimates will differ, leading to different estimated control limits. The performance of the control chart is then conditional on these obtained estimates. This effect of Phase I estimation has received attention in recent literature.

Zwetsloot and Woodall[56] performed a comparative study on the conditional performance of the Shewhart, CUSUM, and EWMA control charts, where they compare the effect of estimation error across these charts. They studied this question by considering both the overall (marginal) as well as the CARL performance. The in-control CARL performance of control charts in Phase II was studied conditioning on the Phase I in-control parameter estimates. The effect of Phase I parameter estimation on control charts is known to be significant. Generally, it can lead to more frequent false alarms and a loss in the ability to detect shifts quickly, as already discussed in our sections on the Shewhart, CUSUM, and EWMA control charts.

Zwetsloot and Woodall[56] concluded that the Shewhart chart is greatly affected by estimation error and that the EWMA and CUSUM charts behave quite differently when evaluated on conditional performance. They evaluated and compared the effect of estimation across these three charts by evaluating the conditional in-control performance for all charts pairwise for a specific set of Phase I estimates. Their simulation study showed that the effect of estimated parameters on performance variation is a lot larger for the Shewhart control chart than for the CUSUM or EWMA control charts. However, somewhat unexpectedly, it also shows that the EWMA and CUSUM control charts behave quite differently. The conditional in-control ARL of the CUSUM control chart is generally higher than the conditional in-control ARL of the EWMA control chart. Furthermore, the relative performance depends on the amount of data in Phase I. Their advice to consider the conditional performance further led to the results of Diko et al.[57]

Diko et al.[57] performed a comparative study between these control charts with estimated parameters. Because the exceedance probability criterion provides a specified in-control performance, they focused mainly on the out-of-control detection capability of the control charts. The CARL values for the Shewhart, CUSUM, and EWMA control charts show that the in-control CARL values are in general larger when the mean is estimated more accurately. This is a natural consequence of the control limit adjustments in order to guarantee a minimum in-control ARL with a specified (large) probability, which generally widens the control limits for given Phase I estimates and thus leads to larger ARL values. Note that large in-control CARL values are actually favorable, while for the out-of-control situation, smaller values mean quicker detection on average. If the in-control mean is overestimated in Phase I, detection of increases in the process mean become more difficult to detect. Likewise, underestimation of the in-control mean leads to slower detection of decreases in the process mean.

When considering the results for the Shewhart control chart only, we observe that estimation of the mean has substantially less impact on the CARL than estimation of the standard deviation. This result was also obtained in Zwetsloot and Woodall.[56] Note that an overestimation of the mean will increase both the UCL and the LCL, which will increase the in-control probability of the plotting statistic falling below the LCL, but decrease that probability for the UCL. On the other hand, overestimation of the standard deviation increases the value of the UCL and lowers the value of the LCL, thus decreasing the FAR on both sides of the control chart.

The results of the CUSUM and EWMA control charts are quite similar. Compared with the Shewhart control chart, these charts are much more affected by estimation error in the mean. This is especially so for the charts designed to detect small shifts. Without adjustments, small estimation errors can lead to low in-control CARL values due to estimation error (see e.g., Saleh et al.[6]). In order to correct for this, substantial control limit adjustments are required. When designed to detect small shifts, this required correction is especially large, as small estimation errors would otherwise

quickly be detected as process changes. For CUSUM and EWMA control charts designed for larger shifts, the impact of estimation error of the mean is more similar to that of the Shewhart control charts. We note again that the Shewhart control chart is actually a special case of the EWMA when $\lambda = 1$. The impact of estimation error of the standard deviation on the CUSUM and EWMA control charts is quite similar to that of the Shewhart control chart. Overestimation of the standard deviation leads to larger CARL values, both in control and out of control.

In summary, in Diko et al.[57], the performance of Shewhart, CUSUM, and EWMA control charts are compared when they are adjusted for parameter estimation. The adjustments are made to provide a specified minimum in-control performance with a specified probability. This is in line with adjustments advocated recently by many authors, in order to reduce the risk of large FARs when parameters are estimated. They found that CUSUM control charts provide the fastest out-of-control detection of sustained shifts for almost all shift sizes and estimation errors. Since the control limit adjustments are designed to control the in-control performance, the out-of-control detection speed is considered the most important aspect in the comparison. The EWMA control chart is only slightly behind the CUSUM chart, while the Shewhart control chart is only able to compete with the other two schemes for very large shifts.

6 | OTHER DEVELOPMENTS

In the previous sections, we discussed some results on Shewhart, EWMA, and CUSUM control charts when one applies the exceedance probability criterion. While many methods are designed for normally distributed data, some of the discussed methods apply a more general setting, for example, through the use of nonparametric estimation. Of course, next to parametric assumptions, other aspects of control chart design can be considered that are potential limitations to the control chart performance.

Many methods are based on the assumption of serially independent data. There are various control chart procedures available that provide adjustments for such data, see, for example, Jensen et al.[37] However, little research is done on autocorrelated data in combination with the exceedance probability criterion. One example is Weiss et al.[58], who apply the bootstrap procedure of Gandy and Kvaløy[18] in an AR(1) setting.

Another decision in the design of a control chart with estimated parameters is whether the parameter estimates should be updated in Phase II or not. Obviously, when more in-control data is available and used for estimations, the estimation uncertainty will decrease. Hawkins[59] introduced self-starting control charts, which allow the Phase II monitoring to start based on a relatively small Phase I sample due to such parameter updates. Keefe et al.[41] evaluated the conditional in-control performance of such charts. However, a pitfall of such self-starting methods is the possibility of incorporating out-of-control observations into the parameter estimates. Huberts et al.[60] investigated the effect of continuously updating control chart limits on the control chart performance under various out-of-control scenarios as well. They find indeed that updating of parameters should be done with caution. To that end, Capizzi and Masarotto[61] proposed a new approach to use parameter updates in the setting of the exceedance probability criterion for Shewhart, CUSUM, and EWMA control charts, which is based on a delay in the time to update. This delay reduces the chance of out-of-control observations to be included in the parameter estimates, while still benefitting from more accurate parameter estimates.

The exceedance probability criterion is obviously not the only approach to adjust control charts for parameter estimation, and countless other criteria are imaginable. For example, Albers and Kallenberg[9,10] also introduced the bias criterion, which aims to provide a specified control chart performance in expectation. Some other recent ideas on control chart design and evaluation can be found in Testik et al.[62], Atalay et al[63], and Woodall and Faltin,[64] among others.

7 | SUMMARY AND RECOMMENDATIONS

We strongly encourage practitioners to use the exceedance probability approach, which was introduced by Albers and Kallenberg,[9,10] in designing control charts. This reduces the rate of occurrence of false alarms with minimal effects on detecting process changes. In recent literature, one may note that a lot of researchers already use this recommendation for alternative situations.

When designing Shewhart control charts with a large amount of Phase I data, we recommend the nonparametric approach of Goedhart et al.[24] With smaller amounts of Phase I data, we recommend the approach of Goedhart et al.[23]

for normal data or to apply for individual observations the statistical tolerance intervals for other distributions (see Krishnamoorthy and Mathew[25]). For designing EWMA and CUSUM control charts for monitoring the mean, we recommend the approaches of Diko et al.[42,53], respectively.

We have the following thoughts for future research:

1. Woodall and Faltin[64] argued that it is important to ensure the practical significance of control chart signals, not just statistical significance. The importance of specified process changes should be specified by the practitioner, not measured in units of the standard error as is the standard in practice. Their approach leads naturally to CUSUM charts with potentially larger reference values. The effect of estimation error in such CUSUM charts with large reference values would be expected to be less than for smaller reference values, but study of this issue is required.

2. It has been shown that estimation error has a significant effect on the conditional performance for profile monitoring methods (see, e.g., Aly et al[65]). The exceedance probability criterion has yet to be applied for profile monitoring applications.

3. The method of Goedhart et al.[24] provides a nonparametric alternative using the exceedance probability criterion for Shewhart control charts. Nonparametric approaches based on this criterion are not yet available for CUSUM and EWMA control charts.

4. The reviews on the effect of estimation error on control chart performance by Jensen et al.[37] and Psarakis et al.[66] are out of date due to the extensive amount of recent research on the use of the exceedance probability approach. An updated review of the literature would be useful and help to identify topics that merit investigation.

ORCID
Ronald J.M.M. Does https://orcid.org/0000-0003-3452-6441

REFERENCES
1. Shewhart WA. Quality control charts. *Bell Syst Tech J*. 1926;5(4):593-603.
2. Montgomery DC. *Introduction to statistical quality control: a modern introduction*. 7th ed. Hoboken: John Wiley & Sons; 2013.
3. Does RJMM, Schriever BF. Variables control chart limits and tests tor special causes. *Statistica Neerlandica*. 1992;46(4):229-245.
4. Quesenberry CP. The effect of sample size on the estimated limits of $\bar{X}$ and X control charts. *J Qual Technol*. 1993;25(4):237-247.
5. Chen G. The mean and standard deviation of the run length distributions of $\bar{X}$ charts when control limits are estimated. *Stat Sin*. 1997;7(3):789-798.
6. Saleh NA, Mahmoud MA, Keefe MJ, Woodall WH. The difficulty in designing Shewhart X-bar and X control charts with estimated parameters. *J Qual Technol*. 2015;47(2):127-138.
7. Epprecht EK, Loureiro LD, Chakraborti S. Effect of the amount of phase I data on the phase II performance of S^2 and S control charts. *J Qual Technol*. 2015;47(2):139-155.
8. Albers W, Kallenberg WCM. Are estimated control charts in control? *Stat*. 2004;38(1):67-79.
9. Albers W, Kallenberg WCM. Estimation in Shewhart control charts: effects and corrections. *Metrika*. 2004;59(3):207-234.
10. Albers W, Kallenberg WCM. New corrections for old control charts. *Qual Eng*. 2005;17(3):467-473.
11. Schoonhoven M, Nazir HZ, Riaz M, Does RJMM. Robust location estimators for the $\bar{X}$ control chart. *J Qual Technol*. 2011;43(4):363-379.
12. Schoonhoven M, Riaz M, Does RJMM. Design and analysis of control charts for standard deviation with estimated parameters. *J Qual Technol*. 2011;43(4):307-333.
13. Schoonhoven M, Does RJMM. A robust standard deviation control chart. *Technometrics*, 2012;54(1):73-82.
14. Tsai TR, Lin JJ, Wu SJ, Lin HC. On estimating control limits of X-bar chart when the number of subgroups is small. *Int J Adv Manuf Technol*. 2005;26(11–12):1312-1316.
15. Goedhart R, Schoonhoven M, Does RJMM. Correction factors for Shewhart X and X-bar control charts to achieve desired unconditional ARL. *Int J Prod Res*. 2016;54(24):7464-7479.
16. Diko MD, Goedhart R, Chakraborti S, Epprecht EK, Does RJMM. Phase II control charts for monitoring dispersion when parameters are estimated. *QuaL Eng*. 2017;29(4):605-622.
17. Jones MA, Steiner SH. Assessing the effect of estimation error on the risk-adjusted CUSUM chart performance. *International J Qual Health Care*. 2012;24(2):176-181.
18. Gandy A, Kvaløy JT. Guaranteed conditional performance of control charts via bootstrap methods. *Scandanavian J Stat*. 2013;40(4):647-668.

19. Faraz A, Woodall WH, Heuchenne C. Guaranteed conditional performance of the S^2 control chart with estimated parameters. *Int J Prod Res*. 2015;53(14):4405-4413.

20. Saleh NA, Mahmoud MA, Jones-Farmer LA, Zwetsloot IM, Woodall WH. Another look at the EWMA control chart with estimated parameters. *J Qual Technol*. 2015;47(4):363-382.

21. Goedhart R, Schoonhoven M, Does RJMM. Guaranteed in-control performance for the Shewhart X and X-bar control charts. *J Qual Technol*. 2017;49(2):155-171.

22. Goedhart R, da Silva MM, Schoonhoven M, et al. Shewhart control charts for dispersion adjusted for parameter estimation. *IISE Trans*. 2017;49(8):838-848.

23. Goedhart R, Schoonhoven M, Does RJMM. Letter to the editor: on guaranteed in-control performance for the Shewhart X and X-bar control charts. *J Qual Technol*. 2018;50(1):130-132.

24. Goedhart R, Schoonhoven M, Does RJMM. Nonparametric control of the conditional performance in statistical process monitoring. *J Qual Technol* (to appear). 2019. https://doi.org/10.1080/00224065.2019.1611352

25. Krishnamoorthy K, Mathew T. *Statistical tolerance regions: theory, applications, and computation*. New York: Wiley; 2009.

26. Albers W, Kallenberg WCM, Nurdati S. Parametric control charts. *J Stat Plann Infer*. 2004;124(1):159-184.

27. Huberts LCE, Schoonhoven M, Goedhart R, Diko MD, Does RJMM. The performance of $\bar{X}$ control charts for large non-normally distributed datasets. *Qual Reliab Eng Int*. 2018;34(6):979-996.

28. Box GEP, Cox DR. An analysis of transformations. *J Royal Stat Soc, Ser B (Methodol)*. 1964;26(2):211-252.

29. Chou YM, Polansky AM, Mason RL. Transforming nonnormal data to normality in statistical process control. *J Qual Technol*. 1998;30(2):133-141.

30. Qiu P. Some perspectives on nonparametric statistical process control. *J Qual Technol*. 2018;50(1):49-65.

31. Young DS, Mathew T. Improved nonparametric tolerance intervals based on interpolated and extrapolated order statistics. *J Nonparametr Stat*. 2014;26(3):415-432.

32. Roberts SW. Control chart tests based on geometric moving averages. *Technometrics*,1959;1(3):239-250.

33. Zwetsloot IM, Schoonhoven M, Does RJMM. A robust estimator of location in Phase I based on an EWMA chart. *J Qual Technol*. 2014;46(4):302-316.

34. Mahmoud MA, Henderson GR, Epprecht EK, Woodall WH. Estimating the standard deviation in quality-control applications. *J Qual Technol*. 2010;42(4):348-357.

35. Crowder SV. Design of exponentially weighted moving average schemes. *J Qual Technol*. 1989;21(3):155-162.

36. Knoth, W. *spc*: Statistical process control—calculation of ARL and other control chart performance measures. R package version 0.6.2. https://CRAN.R-project.org/package=spc 2019

37. Jensen WA, Jones-Farmer LA, Champ CW, Woodall WH. Effects of parameter estimation on control chart properties: a literature review. *J Qual Technol*. 2006;38(4):349-364.

38. Jones LA, Champ CW, Rigdon SE. The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*, 2001;43(2):156-167.

39. Brook D, Evans DA. An approach to the probability distribution of CUSUM run length. *Biometrika*. 1972;59(3):539-549.

40. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 1990;32(1):1-12.

41. Keefe MJ, Woodall WH, Jones-Farmer LA. The conditional in-control performance of self-starting control charts. *Qual Eng*. 2015;27(4):488-499.

42. Diko MD, Chakraborti S, Does RJMM. Guaranteed in-control performance of the EWMA chart for monitoring the mean. *Qual Reliab Eng Int*. 2019;35(4):1144-1160.

43. Hu X, Castagliola P. Guaranteed conditional design of the median chart with estimated parameters. *Qual Reliab Eng Int*. 2014;33(8):1873-1884.

44. Page ES. Continuous inspection schemes. *Biometrika*. 1954;41(1–2):100-114.

45. Ewan WD, Kemp KW. Sampling inspection of continuous processes with no autocorrelation between successive results. *Biometrika*. 1960;47(3–4):363-380.

46. Hawkins DM, Wu Q. The CUSUM and the EWMA head-to-head. *Qual Eng*. 2014;26(2):215-222.

47. Nazir HZ, Riaz M, Does RJMM, Abbas N. Robust CUSUM control charting. *Qual Eng*. 2013;25:211-224.

48. Crosier RB. A new two-sided cumulative sum quality control scheme. *Technometrics*, 1982;28:187-194.

49. Woodall WH. On the Markov Chain approach to the two-sided CUSUM procedure. *Technometrics*, 1984;26(1):41-46.

50. Reynolds MR Jr. Approximations to the average run length in cumulative sum control charts. *Technometrics*, 1975;17(1):65-71.

51. Jeske DR. Determining the reference sample size needed to control the accuracy of the conditional in-control ARL of a normal theory CUSUM. *Qual Reliab Eng Int*. 2016;32(7):2499-2504.

52. Siegmund D. *Sequential analysis: tests and confidence intervals*. Berlin: Springer; 1985.

53. Diko MD, Chakraborti S, Does RJMM. An alternative design of the two-sided CUSUM chart for monitoring the mean when parameters are estimated. *Comput Indust Eng*. 2019;137:106042.

54. Jones LA, Champ CW, Rigdon SE. The run length distribution of the CUSUM with estimated parameters. *J Qual Technol*. 2004;36(1):95-108.

55. Saleh NA, Zwetsloot IM, Mahmoud MA, Woodall WH. CUSUM charts with controlled conditional performance under estimated parameters. *Qual Eng*. 2016;28(4):402-415.

56. Zwetsloot IM, Woodall WH. A head-to-head comparative study of the conditional performance of control charts based on estimated parameters. *Qual Eng*. 2017;29(2):244-253.

57. Diko, MD, Goedhart, R & Does, RJMM A head-to-head comparison of the out-of-control performance of control charts adjusted for parameter estimation. Accepted for Publication *Quality Engineering* 2020.

58. Weiss CH, Steuer D, Jentsch C, Testik MC. Guaranteed conditional ARL performance in the presence of autocorrelation. *Comput Stat Data Anal*. 2018;128:367-379.

59. Hawkins DM. Self-starting CUSUM charts for location and scale. *J Royal Stat Soc: Ser D (the Statistician)*. 1987;36(4):299-316.

60. Huberts LCE, Schoonhoven M, Does RJMM. The effect of continuously updating control chart limits on control chart performance. *Qual Reliab Eng Int*. 2019;35(4):1111-1128.

61. Capizzi G, Masarotto G. Guaranteed in-control control chart performance with cautious parameter learning. *J Qual Technol*. 2020 (to appear). https://doi.org/10.1080/00224065.2019.1640096

62. Testik MC, Weiss CH, Koca Y, Testik OM. Effectiveness of phase I applications for identifying randomly scattered out-of-control observations and estimating control chart parameters. *Qual Reliab Eng Int*. 2018;34(1):78-92.

63. Atalay M, Caner Testik M, Duran S, Weiss CH. Guidelines for automating Phase I of control charts by considering effects on Phase-II performance of individuals control chart. *Qual Eng*. 2020 32(2): 223-243. https://doi.org/10.1080/08982112.2019.1641208

64. Woodall WH, Faltin FW. Rethinking control chart design and evaluation. *Qual Eng*. 2019;31(4):596-605.

65. Aly AA, Mahmoud MA, Woodall WH. A comparison of the performance of Phase II simple linear profile control charts when parameters are estimated. *Commun Stat Simul Comput*. 2015;44:1432-1440.

66. Psarakis S, Vyniou AK, Castagliola P. Some recent developments on the effects of parameter estimation on control charts. *Qual Reliab Eng Int*. 2014;30(8):1113-1129.

AUTHOR BIOGRAPHIES

Dr Ronald J. M. M. Does is a professor of Industrial Statistics at the University of Amsterdam, Director of the Institute for Business and Industrial Statistics, and Head of the Department of Operations Management at the Amsterdam Business School. He is the recipient of the Hunter Award (2008), Shewhart Medal (2019), and Box Medal (2019). He is a Fellow of the American Society for Quality and American Statistical Association, an elected member of the International Statistical Institute, and an Academician of the International Academy for Quality.

Dr Rob Goedhart is an assistant professor in the Department of Operations Management and senior consultant at the Institute for Business and Industrial Statistics of the University of Amsterdam, the Netherlands. He received his MS (cum laude) in econometrics in 2014 and his PhD in statistics in 2018, both from the University of Amsterdam.

Dr William H. Woodall is an emeritus professor in the Department of Statistics at Virginia Tech. He is a former editor of the Journal of Quality Technology (2001–2003). He is the recipient of the Box Medal (2012), Shewhart Medal (2002), Hunter Award (2019), Youden Prize (1995, 2003), Brumbaugh Award (2000, 2006), Bisgaard Award (2012), Nelson Award (2014), Ott Foundation Award (1987), and best paper award for IIE Transactions on Quality and Reliability Engineering (1997). He is a Fellow of the American Statistical Association, a Fellow of the American Society for Quality, and an elected member of the International Statistical Institute.