WILEY

The effect of continuously updating control chart limits on control chart performance

Leo C.E. Huberts | Marit Schoonhoven | Ronald J.M.M. Does

Department of Operations Management, University of Amsterdam, Amsterdam, The Netherlands

Correspondence
Leo C.E. Huberts, Department of Operations Management, University of Amsterdam, Plantage Muidergracht 12, 1018TV, Amsterdam, The Netherlands.
Email: L.c.e.huberts@uva.nl

Abstract

An open topic within statistical process monitoring is the effect on control chart properties of updating the control chart limits during the monitoring period. The challenge is to use the correct data for updating the control limits as in-control data could be incorrectly classified as out of control and therefore not used for re-estimating the parameters, and out-of-control data could be classified as in control and therefore used for re-estimating. In the present article, we study the effect of updating the Shewhart, cumulative sum, and exponentially weighted moving average control chart limits. We simulate different scenarios: the monitoring data could be in or out of control, and the practitioner may or may not be able to find out whether the process is indeed out of control when the control chart gives a signal. The results reveal that the variation in the performance of the conditional control charts decreases significantly as a result of updating the control chart limits when the updating data are in control and also when the updating data are out of control and the practitioner is able to classify correctly data samples that produce a signal. However, when a practitioner is not able to classify a signal correctly, the advisability of updating depends on the type of control chart and the level of data contamination.

KEYWORDS

big data, conditional, CUSUM, EWMA, Shewhart, unconditional

1 | INTRODUCTION

The control chart proposed by Shewhart,[1] the cumulative sum (CUSUM) control chart proposed by Page,[2] and the exponentially weighted moving average (EWMA) control chart proposed by Roberts[3] are tools that are applied for monitoring quality indicators. The advantage of the CUSUM and EWMA compared with the Shewhart control chart is that they can detect smaller sustained shifts in the process mean earlier. In order to construct the control chart limits, the distributional parameters of the process characteristic have to be estimated. In the literature, it is often assumed that the parameters are estimated only once during the so-called Phase I stage, and that the control chart remains fixed during the monitoring period, a stage which is usually referred to as Phase II (see, eg, Vining[4]).

A control chart design is usually evaluated by determining the performance in expectation, ie, the unconditional control chart performance, and the distributional properties of performance metrics of the conditional control charts. Jones et al[5,6] and Saleh et al[7] studied the unconditional and conditional performance of the EWMA, CUSUM, and Shewhart control

charts, respectively. All conclude that the variation within the performance of the conditional control charts is very large when the parameters are estimated.

Jensen et al[8] surveyed the literature on estimated control charts and identified 13 issues for future research. Issue 6 of Jensen et al[8] suggests reestimating the limits during the monitoring period: "Related to the previous research question is the effect on control chart properties when the control limits are updated in some future time that is not necessarily during a start-up period. If the process is in control, it would be reasonable to use the data to update control limits during Phase II and not continue to use the original limits indefinitely. It is not clear how control chart performance is impacted, but it seems that making use of earlier Phase II data would lead to better control charts." This issue is the subject under study in the present paper. In this context, we should also mention the self-starting CUSUM and Shewhart control chart designs proposed by Hawkins[9] and Quesenberry,[10] respectively. These designs can already be used when just a few samples are available. The performance of these charts was studied by Keefe et al[11] using a simulation procedure. However, they did not study the effect of out-of-control data and reset the parameters to the initial Phase I estimates when the conditional control chart gives a signal.

In the present article, we study the effect of updating for different scenarios: the updating data may be in or out of control, signals may or may not be correctly classified (depending on the scenario), and when the control chart signals the parameters are re-estimated, and the updating continues.

The article is structured as follows. Section 2 gives the relevant control chart designs. Section 3 describes the scenarios that will be included in the simulation procedure, which is described in Section 4. Then, the results during and after updating are given in Sections 5 and 6, respectively, and Section 7 offers conclusions and recommendations.

2 | CONTROL CHART DESIGNS

This section gives the Shewhart, CUSUM, and EWMA control chart designs.

Let X_{ij} denote the jth observation in sample i ($i = 1, 2, ...$ and $j = 1, 2, ...n$), and let $\boldsymbol{X}_i$ denote the vector containing the n observations of sample i. Further, let m_I denote the number of Phase I samples for initial parameter estimation and m_u the number of samples within the updating period. We assume that the m_I samples in Phase I are in control (in this stage the practitioner determines the in-control state). The m_u monitoring/updating samples may or may not be out of control. Further, we assume that the observations X_{ij} in the first m_I samples are independently and identically $N(\mu, \sigma^2)$ distributed and that the observations in the next m_u samples are independently and identically $N(\mu + \delta, \sigma^2)$ distributed with probability P, where the values of δ and P depend on the scenario. Finally, let t be the time stamp during monitoring ($t = m_I + 1, m_I + 2, \ldots, m_I + m_u$), and let $\boldsymbol{X}_{t-1}^{ic}$ denote the samples that are classified as in control up to and including time $t - 1$. The number of samples within $\boldsymbol{X}_{t-1}^{ic}$ is denoted by m_{t-1} (so $m_I \leqslant m_{t-1} \leqslant m_I + m_u$).

For each of the three control charts, μ and σ have to be estimated, and we use the same estimators for each chart. The parameter μ is estimated by

$$\overline{\overline{X}}_{t-1} = \frac{1}{m_{t-1}} \sum_{i=1}^{t-1} \left(\frac{1}{n} \sum_{j=1}^{n} X_{ij} \right) \mathbb{1}_{\boldsymbol{X}_i \in \boldsymbol{X}_{t-1}^{ic}}, \tag{1}$$

with $\mathbb{1}$ as the indicator function. Further, σ is estimated by

$$\tilde{S}_{t-1} = \left(\frac{1}{m_{t-1}} \sum_{i=1}^{t-1} S_i^2 \mathbb{1}_{\boldsymbol{X}_i \in \boldsymbol{X}_{t-1}^{ic}} \right)^{1/2} / c_4(m_{t-1}(n-1) + 1), \tag{2}$$

where S_i is the ith sample standard deviation defined by

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

and

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

The following sections outline the Shewhart, CUSUM, and EWMA control chart designs.

2.1 | Shewhart control chart

The estimated Shewhart control limits based on the samples in X_{t-1}^{ic}, used for monitoring at time t, $t = m_I + 1, m_I + 2, \ldots, m_I + m_u$, are given by

$$\widehat{UCL}_{t-1} = \overline{\overline{X}}_{t-1} + 3\tilde{S}_{t-1}/\sqrt{n},$$
$$\widehat{LCL}_{t-1} = \overline{\overline{X}}_{t-1} - 3\tilde{S}_{t-1}/\sqrt{n}, \tag{3}$$

where $\overline{\overline{X}}_{t-1}$ is given by (1) and $\tilde{S}_{t-1}$ by (2), while 3 is the traditional constant used for known parameters to ensure that the false alarm probability (average run length) of the chart is equal to 0.0027 (370)*.

2.2 | CUSUM control chart

As the Crosier CUSUM control chart outperforms the classical CUSUM chart (see Crosier[14]), we will apply this design.

Let C_t, $t = m_I + 1, m_I + 2, \ldots, m_I + m_u$, denote the standardized statistic defined as

$$C_t = |V_{t-1} + \widehat{Q}_t|, \tag{4}$$

where $\widehat{Q}_t = \frac{\overline{X}_t - \overline{\overline{X}}_{t-1}}{\tilde{S}_{t-1}/\sqrt{n}}$ and $V_{m_I} = 0$. The monitoring statistic then becomes

$$V_t = \begin{cases} 0 & \text{if} \quad C_t \leqslant k \\ (V_{t-1} + \widehat{Q}_t)(1 - \frac{k}{C_t}) & \text{if} \quad C_t > k, \end{cases} \tag{5}$$

where k is the reference value that determines the point at which V_t accumulates deviations from the target value $\overline{\overline{X}}_{t-1}$ and is commonly set at $k = \frac{1}{2}$. The chart signals if $|V_t| > H$, where H is the CUSUM control chart limit whose value depends on desired chart performance. To achieve an in-control average run length (ARL) of 370 for $k = \frac{1}{2}$ and known parameters, H should be set at 4.3904 (see also Crosier[14]).

2.3 | EWMA control chart

The EWMA control chart is an extension of the CUSUM chart, adding weights to the cumulative sum of observations. The EWMA statistic is defined as

$$Z_t = \lambda \overline{X}_t + (1 - \lambda)Z_{t-1}, \tag{6}$$

where $0 < \lambda \leqslant 1$ and Z_{m_I} equals the mean estimate $\overline{\overline{X}}_{m_I}$. For $\lambda = 1$, the EWMA control chart is equal to the Shewhart control chart. The EWMA control limits for monitoring the process at time $t = m_I + 1, m_I + 2, \ldots, m_I + m_u$ are

$$\widehat{UCL}_{t-1} = \overline{\overline{X}}_{t-1} + L\frac{\tilde{S}_{t-1}}{\sqrt{n}}\sqrt{\frac{\lambda}{2 - \lambda}[1 - (1 - \lambda)^{2t}]},$$
$$\widehat{LCL}_{t-1} = \overline{\overline{X}}_{t-1} - L\frac{\tilde{S}_{t-1}}{\sqrt{n}}\sqrt{\frac{\lambda}{2 - \lambda}[1 - (1 - \lambda)^{2t}]}, \tag{7}$$

where λ and L determine the in-control performance. We set $\lambda = 0.1$ and $L = 2.703$ because for these settings, the in-control ARL is 370 when the distributional parameters are known and the chart is able to detect small mean shifts quickly (cf. Lucas and Saccucci[15]). When Z_t falls above (below) $\widehat{UCL}_{t-1}$ ($\widehat{LCL}_{t-1}$), the process is considered out of control.

3 | SIMULATION SCENARIOS

In this section, we describe the various scenarios included in the simulation. For each scenario and control chart type, we simulate 100 000 conditional control charts. The scenarios that we consider are presented in Table 1. The case $m_u = 0$ is used to investigate how control chart performance improves when updating ($m_u > 0$) compared with not updating

---

*The reason why we use the traditional constant is that we want to make a comparison with the CUSUM and EWMA control charts with traditional constants. We have, however, also studied the performance of the updated Shewhart chart with the constants for estimated parameters of Goedhart et al,[12,13] but the results are similar.

TABLE 1 Scenarios

	m_I	m_u	Phase I: IC or OOC?	Updating: IC or OOC?	Signal: Reason Known?
Reference 1	5	0	IC	N/A	N/A
Reference 2	200	0	IC	N/A	N/A
Reference 3	2000	0	IC	N/A	N/A
Scenario 1	5	1995	IC	IC	No
Scenario 2	200	1800	IC	IC	No
Scenario 3	5	1995	IC	IC	Yes
Scenario 4	200	1800	IC	IC	Yes
Scenario 5	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	No
Scenario 6	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	No
Scenario 7	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	Yes
Scenario 8	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	Yes
Scenario 9	5	1995	IC	OOC: $P = 0.01; \delta = 2$	No
Scenario 10	200	1800	IC	OOC: $P = 0.01; \delta = 2$	No
Scenario 11	5	1995	IC	OOC: $P = 0.01; \delta = 2$	Yes
Scenario 12	200	1800	IC	OOC: $P = 0.01; \delta = 2$	Yes
Scenario 13	5	1995	IC	OOC: $P = 0.1; \delta = 3$	No
Scenario 14	200	1800	IC	OOC: $P = 0.1; \delta = 3$	No
Scenario 15	5	1995	IC	OOC: $P = 0.1; \delta = 3$	Yes
Scenario 16	200	1800	IC	OOC: $P = 0.1; \delta = 3$	Yes

Abbreviations: IC, in control; OOC, out of control.

($m_u = 0$). The reference scenario $m_I = 2000$ and $m_u = 0$ is included to determine the difference in performance by starting with a small sample and then updating (using $m_I + m_u = 2000$ samples in total) compared with directly constructing a control chart based on $m_I = 2000$ samples. Hence, we consider three reference scenarios: $m_I = 5, 200, 2000$, all with $m_u = 0$.

In scenarios 1 to 4, the monitoring/updating data are in control, and m_I is 5 or 200, and m_u is 1995 (for $m_I = 5$) or 1800 (for $m_I = 200$). In scenarios 5 to 8, 9 to 12, and 13 to 16, the monitoring/updating data may be out of control, with varying levels of contamination. The contamination is modeled as follows: each sample has a probability P that the observations are drawn from a $N(\mu + \delta, \sigma^2)$ distribution. We assume that, when the process is out of control, then the process remains out of control until the control chart gives a signal.

Moreover, we investigate the effect of the incorrect classification of a signal. In scenarios 1, 2, 5, 6, 9, 10, 13, and 14, it is assumed that the operator cannot classify a signal correctly and therefore relies on the control chart. Thus, samples that give no signal are considered as in control and used for re-estimating the control limits, and data that give a signal are classified as out of control and not used for reestimation. In these scenarios, it is thus not possible to trace back the start of the out-of-control situation and filter out previous out-of-control samples. On the other hand, in scenarios 3, 4, 7, 8, 11, 12, 15, and 16, we assume that the operator is able to identify the cause of a signal; so data samples that give a false signal are included, and we assume that in this case, it is possible to retrace the start of the out-of-control situation and exclude previous out-of-control samples from estimation.

4 | SIMULATION PROCEDURE

Below, we describe the simulation procedure.

Step 1: Generate conditional control chart
In this step, we construct the initial conditional control chart based on the Phase I dataset of m_I (5, 200) samples of size 5. We estimate μ and σ with the estimators $\overline{\overline{X}}_{t-1}$ and $\tilde{S}_{t-1}$ ($t - 1 = m_I$) given by (1) and (2), respectively, and determine the Shewhart or EWMA control limits using (3) or (7), respectively. The limit for the CUSUM control chart is 4.3904.

Step 2: Use conditional control chart for monitoring and update the chart

For each $t = m_I + 1, m_I + 2, \ldots, m_I + m_u$: for scenarios 1 to 4, draw a sample X_t from $N(\mu, \sigma^2)$, and for scenarios 5 to 16, when the process is in control, draw a sample with probability $1 - P$ from $N(\mu, \sigma^2)$ and otherwise from $N(\mu + \delta, \sigma^2)$, with the values of P and δ depending on the scenario (see Table 1). Calculate the test statistics ($\overline{\overline{X}}_t$, V_t, or Z_t).

When the test statistic falls between $\widehat{UCL}_{t-1}$ and $\widehat{LCL}_{t-1}$, the dataset of classified in-control samples, X_{t-1}^{ic}, is augmented with X_t and denoted by X_t^{ic}. The process parameters are recalculated using (1) and (2), and the control limits $\widehat{UCL}_t$ and $\widehat{LCL}_t$ are determined with the new parameter estimates.

When the test statistic falls outside the control limits $\widehat{UCL}_{t-1}$ and $\widehat{LCL}_{t-1}$, the next step depends on the scenario. For scenarios 1, 2, 5, 6, 9, 10, 13, and 14 (the reason-unknown scenarios) the current data sample X_t is considered out of control and therefore not added to the dataset of classified in-control samples. The limits are not recalculated. Thus, $\widehat{UCL}_t = \widehat{UCL}_{t-1}$, $\widehat{LCL}_t = \widehat{LCL}_{t-1}$, $\overline{\overline{X}}_t = \overline{\overline{X}}_{t-1}$ and $\tilde{S}_t = \tilde{S}_{t-1}$. For the CUSUM and EWMA control charts, we set $V_t = 0$ and $Z_t = \overline{\overline{X}}_{m_I}$.

In contrast, when the test statistic falls outside the control limits, but the process is in control, and the operator is able to determine the cause of the signal—as is the case in scenarios 3, 4, 7, 8, 11, 12, 15, and 16—X_t will be used for updating the control chart. The dataset of classified in-control samples, X_{t-1}^{ic}, is augmented with X_t and is denoted by X_t^{ic}. The process parameters are recalculated using (1) and (2), and the control limits $\widehat{UCL}_t$ and $\widehat{LCL}_t$ are determined with the new parameter estimates.

When the test statistic falls outside the control limits, the process is out of control, and the operator is able to determine the cause of the signal—as is the case in scenarios 3, 4, 7, 8, 11, 12, 15, and 16—X_t will not be used for updating the control chart. Moreover, previous real out-of-control samples, the number denoted by N with $N \leqslant m_u$, are excluded from X_{t-1}^{ic}. Thus, $X_t^{ic} = X_{t-N}^{ic}$, $\widehat{UCL}_t = \widehat{UCL}_{t-N}$, $\widehat{LCL}_t = \widehat{LCL}_{t-N}$, $\overline{\overline{X}}_t = \overline{\overline{X}}_{t-N}$, and $\tilde{S}_t = \tilde{S}_{t-N}$. For the CUSUM and EWMA control charts, we set $V_t = 0$ and $Z_t = \overline{\overline{X}}_{t-N}$.

To assess the performance during updating, we determine for each conditional control chart the true alarm percentage (CTAP) and the false alarm percentage (CFAP) within each simulation run. Related measures were presented by Fraker et al,[16] Chakraborti et al,[17] and Frisén.[18] To this end, for 100 000 simulation runs, we count the number of correct signals, out-of-control samples, false alarms, and in-control samples.

Step 3: Assess the conditional control chart performance

To assess the performance of a conditional control chart during updating, we determine the CTAP and CFAP as follows:

$$CTAP = \frac{\#\text{correct signals}}{\#\text{out-of-control samples}},$$

and

$$CFAP = \frac{\#\text{false signals}}{\#\text{in-control samples}}.$$

To assess the conditional performance of a conditional control chart after updating, we determine the conditional false alarm probability (CFAR) and the conditional average run length (CARL). For the Shewhart chart, these values can be obtained by

$$CFAR = \Phi\left(\frac{\widehat{LCL}_t - \mu}{\sigma/\sqrt{n}}\right) + 1 - \Phi\left(\frac{\widehat{UCL}_t - \mu}{\sigma/\sqrt{n}}\right), \tag{8}$$

and

$$CARL = 1/CFAR. \tag{9}$$

For the CUSUM and EWMA control charts, CFAR is assessed by determining the number of false alarms on an interval of 100 000 samples, and CARL is determined by the Markov chain approaches given by Hany and Mahmoud[19] and Saleh et al.[20]

Step 4: Assess the overall control chart performance

To assess the unconditional control chart performance during updating, we determine the average true alarm percentage (ATAP) and the average false alarm percentage (AFAP) by averaging the CTAP and CFAP values for the R simulation runs.

The expected average run length after updating (*AARL*) and the expected false alarm rate after updating (*AFAR*) are determined by averaging the corresponding conditional values obtained in the R simulation runs. Moreover, we determine the 10th and 90th percentiles of the *CARL* and *CFAR* values, which are indicated by $CARL_{10}$, $CARL_{90}$, $CFAR_{10}$, and $CFAR_{90}$.

The next two sections present the performance results during and after updating.

5 | PERFORMANCE DURING UPDATING

In this section, we consider the chart performance during updating for either $m_u = 1995$ or $m_u = 1800$ updates. The *ATAP* and *AFAP* values for the three charts are reported in Table 2.

5.1 | Shewhart

The in-control behavior of scenarios 1 to 4 for the Shewhart chart is as expected. For the small mean deviations ($\delta = 0.5$) of scenarios 5 to 8, in the reason-unknown scenarios (5 to 6), the *ATAP* values are smaller than the known parameter detection probability (0.03), and the *AFAP* values are larger as well. In the reason-known scenarios (7 to 8), the *ATAP* and *AFAP* values are very close to the known detection probabilities for the Shewhart chart. For larger deviations $\delta = 2, 3$ in scenarios 9 to 16, the detection percentages of Table 2 show almost perfect detection, as expected for the Shewhart chart when large mean shifts occur.

5.2 | CUSUM and EWMA

As the results for the CUSUM and EWMA are comparable, we will consider them together in this section. The in-control *AFAP* performance is around 0.0027. As expected, the CUSUM and EWMA charts are more capable of detecting small mean shifts ($\delta = 0.5$) in scenarios 5 to 8 than the Shewhart chart. For $\delta = 2$, the *ATAP* values for both charts are around

TABLE 2 Performance during updating

	m_I	m_u	Phase I: IC or OOC?	Phase II: IC or OOC?	Signal: Reason Known?	*ATAP* (*AFAP*) Shewhart	CUSUM	EWMA
Reference 1	5	0	IC	N/A	N/A	NA (NA)	NA (NA)	NA (NA)
Reference 2	200	0	IC	N/A	N/A	NA (NA)	NA (NA)	NA (NA)
Reference 3	2000	0	IC	N/A	N/A	NA (NA)	NA (NA)	NA (NA)
Scenario 1	5	1995	IC	IC	No	NA (0.0027)	NA (0.0028)	NA (0.0025)
Scenario 2	200	1800	IC	IC	No	NA (0.0027)	NA (0.0027)	NA (0.0026)
Scenario 3	5	1995	IC	IC	Yes	NA (0.0027)	NA (0.0028)	NA (0.0025)
Scenario 4	200	1800	IC	IC	Yes	NA (0.0027)	NA (0.0027)	NA (0.0026)
Scenario 5	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	No	0.0107 (0.0086)	0.1106 (0.0054)	0.0962 (0.0091)
Scenario 6	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	No	0.0167 (0.0044)	0.1199 (0.0029)	0.113 (0.0028)
Scenario 7	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	Yes	0.0225 (0.0027)	0.1263 (0.0026)	0.1143 (0.0024)
Scenario 8	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	Yes	0.0277 (0.0027)	0.13 (0.0026)	0.1245 (0.0024)
Scenario 9	5	1995	IC	OOC: $P = 0.01; \delta = 2$	No	0.9296 (0.0033)	0.5901 (0.0027)	0.5477 (0.0027)
Scenario 10	200	1800	IC	OOC: $P = 0.01; \delta = 2$	No	0.932 (0.0027)	0.593 (0.0027)	0.5564 (0.0025)
Scenario 11	5	1995	IC	OOC: $P = 0.01; \delta = 2$	Yes	0.9297 (0.0027)	0.5993 (0.0026)	0.5638 (0.0024)
Scenario 12	200	1800	IC	OOC: $P = 0.01; \delta = 2$	Yes	0.9309 (0.0027)	0.5995 (0.0026)	0.5681 (0.0024)
Scenario 13	5	1995	IC	OOC: $P = 0.1; \delta = 3$	No	0.9999 (0.0027)	0.8803 (0.0024)	0.6971 (0.0074)
Scenario 14	200	1800	IC	OOC: $P = 0.1; \delta = 3$	No	0.9999 (0.0027)	0.8935 (0.0019)	0.745 (0.003)
Scenario 15	5	1995	IC	OOC: $P = 0.1; \delta = 3$	Yes	0.9999 (0.0027)	0.9095 (0.0018)	0.8191 (0.0015)
Scenario 16	200	1800	IC	OOC: $P = 0.1; \delta = 3$	Yes	0.9999 (0.0027)	0.91 (0.0017)	0.8176 (0.0015)

Abbreviations: *AFAP*, average false alarm percentage; *ATAP*, average true alarm percentage; CUSUM, cumulative sum; EWMA, exponentially weighted moving average; IC, in control; OOC, out of control.
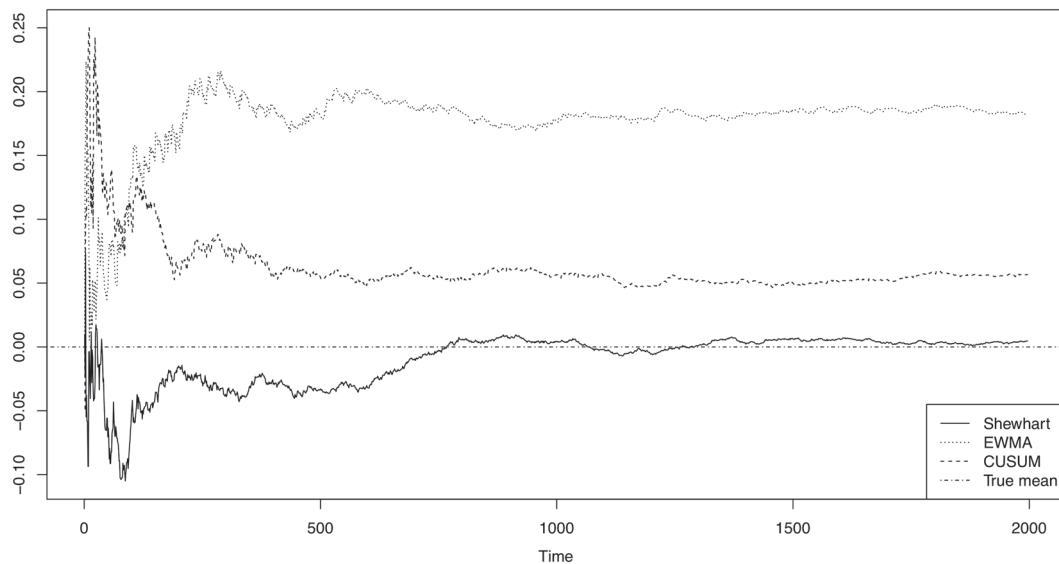
FIGURE 1 Mean estimates during updates for scenario 13

0.5 to 0.6, detecting more than half of the out-of-control observations. Furthermore, the *AFAP* values for both charts in Table 2 are acceptably close to 0.0027 in scenarios 9 to 12. However, in scenarios 13 to 16, for large and frequent deviations ($\delta = 3, P = 0.1$), the *ATAP* values indicate that the CUSUM and EWMA charts still miss a significant portion of the out-of-control observations. The CUSUM (EWMA) chart detects during updating only about 90% (70% to 82%) of the out-of-control observations.

The *ATAP* values are affected by the (contaminated) parameter estimates. Consider the mean estimates during updating for scenario 13 in Figure 1 for the three charts. For the first 100 updates, all three charts show very inaccurate estimation of the mean because of (mostly) the effects of an inaccurate Phase I mean estimate. Over time, the Shewhart chart improves to a very accurate mean estimate of around 0 as it perfectly detects all out-of-control observations. The CUSUM and EWMA chart converge to higher estimates of the mean, as because of their lower detection probability, they include contaminated samples in the mean estimate.

6 | PERFORMANCE AFTER UPDATING

This section presents the performance results of the control charts after updating. The tables that correspond to this section are Tables 3 (Shewhart), 4 (CUSUM), and 5 (EWMA).

6.1 | Shewhart

As we can see in Table 3, updating the chart after a limited initial Phase I dataset (scenarios 1 to 4 and 7 to 16) can result in a control chart performance that is similar to directly estimating limits from a very large Phase I dataset (reference scenario 3). More specifically, for these scenarios, a chart estimated on $m_I = 5$ or $m_I = 200$ samples, which is then used for monitoring and updating during $m_u = 1995$ or $m_u = 1800$ samples, respectively, results in a chart that has the same performance as charts estimated directly from a large Phase I dataset ($m_I = 2000$ and $m_u = 0$). This means that in these scenarios, monitoring can start quickly on a limited dataset. It should also be noted that a chart that is initially estimated on a small dataset ($m_I = 5$), followed by monitoring and updating, leads to a similar performance as of a chart for which a larger Phase I dataset is used ($m_I = 200$) before the monitoring and updating starts.

We can also conclude that, for a more serious level of contamination (scenarios 9 to 16), it does not matter whether the practitioner classifies signals correctly (scenarios 9, 10, 13, 14) or incorrectly (scenarios 11, 12, 15, 16). The reason is that these out-of-control conditions are quickly detected so that there are fewer out-of-control samples in the estimation dataset. Moreover, excluding from the estimation dataset any samples that give a false alarm does not result in a significantly worse performance, as we have also seen in scenarios 1 to 4.

TABLE 3 Performance results: Shewhart control chart

	m_I	m_u	Phase I: IC or OOC?	Phase II: IC or OOC?	Signal: Reason Known?	Shewhart: AARL ($CARL_{10}$; $CARL_{90}$)	AFAR ($CFAR_{10}$; $CFAR_{90}$)
Reference 1	5	0	IC	N/A	N/A	1434 (33; 1897)	0.01170 (0.00053; 0.02989)
Reference 2	200	0	IC	N/A	N/A	373 (264; 497)	0.00285 (0.00201; 0.00379)
Reference 3	2000	0	IC	N/A	N/A	371 (335; 409)	0.00271 (0.00245, 0.00299)
Scenario 1	5	1995	IC	IC	No	371 (335; 408)	0.00272 (0.00245; 0.00299)
Scenario 2	200	1800	IC	IC	No	371 (335; 409)	0.00271 (0.00245; 0.00299)
Scenario 3	5	1995	IC	IC	Yes	371 (334; 408)	0.00271 (0.00245; 0.00299)
Scenario 4	200	1800	IC	IC	Yes	370 (334; 409)	0.00272 (0.00244; 0.00299)
Scenario 5	5	1995	IC	OOC: $P = 0.01$; $\delta = 0.5$	No	143 (51; 270)	0.01009 (0.00371; 0.01978)
Scenario 6	200	1800	IC	OOC: $P = 0.01$; $\delta = 0.5$	No	208 (102; 311)	0.00578 (0.00322; 0.00978)
Scenario 7	5	1995	IC	OOC: $P = 0.01$; $\delta = 0.5$	Yes	369 (324; 419)	0.00285 (0.00239; 0.00309)
Scenario 8	200	1800	IC	OOC: $P = 0.01$; $\delta = 0.5$	Yes	371 (329; 415)	0.00272 (0.00241; 0.00304)
Scenario 9	5	1995	IC	OOC: $P = 0.01$; $\delta = 2$	No	371 (334; 410)	0.00271 (0.00244; 0.00299)
Scenario 10	200	1800	IC	OOC: $P = 0.01$; $\delta = 2$	No	371 (335; 410)	0.00271 (0.00244; 0.00299)
Scenario 11	5	1995	IC	OOC: $P = 0.01$; $\delta = 2$	Yes	371 (335; 408)	0.00271 (0.00245; 0.00299)
Scenario 12	200	1800	IC	OOC: $P = 0.01$; $\delta = 2$	Yes	370 (334; 407)	0.00272 (0.00246; 0.00299)
Scenario 13	5	1995	IC	OOC: $P = 0.1$; $\delta = 3$	No	371 (333; 410)	0.00272 (0.00244; 0.00300)
Scenario 14	200	1800	IC	OOC: $P = 0.1$; $\delta = 3$	No	371 (333; 410)	0.00271 (0.00244; 0.00300)
Scenario 15	5	1995	IC	OOC: $P = 0.1$; $\delta = 3$	Yes	370 (332; 410)	0.00272 (0.00244; 0.00301)
Scenario 16	200	1800	IC	OOC: $P = 0.1$; $\delta = 3$	Yes	371 (333; 410)	0.00272 (0.00244; 0.00300)

Abbreviations: AARL, average run length after updating; AFAR, expected false alarm rate after updating; CARL, conditional average run length; CFAR, conditional false alarm probability; IC, in control; OOC, out of control.

TABLE 4 Performance results: CUSUM control chart

	m_I	m_u	Phase I: IC or OOC?	Phase II: IC or OOC?	Signal: Reason Known?	CUSUM AARL ($CARL_{10}$; $CARL_{90}$)	AFAR ($CFAR_{10}$; $CFAR_{90}$)
Reference 1	5	0	IC	N/A	N/A	205 (5; 264)	0.02545 (0.00167; 0.06595)
Reference 2	200	0	IC	N/A	N/A	257 (106; 401)	0.00319 (0.00224; 0.00431)
Reference 3	2000	0	IC	N/A	N/A	347(299;389)	0.00281(0.00259,0.0031)
Scenario 1	5	1995	IC	IC	No	346 (298; 389)	0.00275 (0.00244; 0.00306)
Scenario 2	200	1800	IC	IC	No	346 (298; 387)	0.00274 (0.00243; 0.00306)
Scenario 3	5	1995	IC	IC	Yes	346 (297; 388)	0.00274 (0.00243; 0.00306)
Scenario 4	200	1800	IC	IC	Yes	346 (297; 389)	0.00274 (0.00243; 0.00306)
Scenario 5	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	No	189 (89; 306)	0.00500 (0.00271; 0.00433)
Scenario 6	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	No	226 (130; 332)	0.00314 (0.00263; 0.00373)
Scenario 7	5	1995	IC	OOC: $P = 0.01; \delta = 0.5$	Yes	344 (294; 389)	0.00276 (0.00242; 0.00308)
Scenario 8	200	1800	IC	OOC: $P = 0.01; \delta = 0.5$	Yes	344 (293; 389)	0.00275 (0.00243; 0.00307)
Scenario 9	5	1995	IC	OOC: $P = 0.01; \delta = 2$	No	309 (225; 376)	0.00283 (0.00249; 0.0032)
Scenario 10	200	1800	IC	OOC: $P = 0.01; \delta = 2$	No	316 (238; 379)	0.00281 (0.00247; 0.00316)
Scenario 11	5	1995	IC	OOC: $P = 0.01; \delta = 2$	Yes	346 (296; 389)	0.00274 (0.00244; 0.00307)
Scenario 12	200	1800	IC	OOC: $P = 0.01; \delta = 2$	Yes	345 (297; 388)	0.00275 (0.00244; 0.00306)
Scenario 13	5	1995	IC	OOC: $P = 0.1; \delta = 3$	No	178 (81; 295)	0.00356 (0.00276; 0.00452)
Scenario 14	200	1800	IC	OOC: $P = 0.1; \delta = 3$	No	219 (121; 329)	0.00317 (0.00265; 0.00378)
Scenario 15	5	1995	IC	OOC: $P = 0.1; \delta = 3$	Yes	344 (292; 390)	0.00275 (0.00242; 0.00308)
Scenario 16	200	1800	IC	OOC: $P = 0.1; \delta = 3$	Yes	344 (292; 389)	0.00274 (0.00242; 0.00307)

Abbreviations: AARL, average run length after updating; AFAR, expected false alarm rate after updating; CARL, conditional average run length; CFAR, conditional false alarm probability; CUSUM, cumulative sum; IC, in control; OOC, out of control.

TABLE 5 Performance results: EWMA control chart

	m_I	m_u	Phase I: IC or OOC?	Phase II: IC or OOC?	Signal: Reason known?	EWMA AARL ($CARL_{10}$; $CARL_{90}$)	AFAR ($CFAR_{10}$; $CFAR_{90}$)
Reference 1	5	0	IC	N/A	N/A	156 (5; 235)	0.03491 (0.00477; 0.08170)
Reference 2	200	0	IC	N/A	N/A	248 (96; 396)	0.00324 (0.00225; 0.00442)
Reference 3	2000	0	IC	N/A	N/A	346(292,391)	0.00279 (0.00257,0.00300)
Scenario 1	5	1995	IC	IC	No	345 (291; 390)	0.00272 (0.00242; 0.00303)
Scenario 2	200	1800	IC	IC	No	345 (290; 390)	0.00272 (0.00241; 0.00303)
Scenario 3	5	1995	IC	IC	Yes	345 (289; 390)	0.00271 (0.00241; 0.00303)
Scenario 4	200	1800	IC	IC	Yes	346 (293; 390)	0.00271 (0.00241; 0.00301)
Scenario 5	5	1995	IC	OOC: $P = 0.01$; $\delta = 0.5$	No	155 (40; 279)	0.00803 (0.00278; 0.00695)
Scenario 6	200	1800	IC	OOC: $P = 0.01$; $\delta = 0.5$	No	205 (107; 323)	0.00326 0.00266; 0.00398)
Scenario 7	5	1995	IC	OOC: $P = 0.01$; $\delta = 0.5$	Yes	343 (289; 391)	0.00281 (0.00241; 0.00305)
Scenario 8	200	1800	IC	OOC: $P = 0.01$; $\delta = 0.5$	Yes	344 (285; 390)	0.00272 (0.00241; 0.00304)
Scenario 9	5	1995	IC	OOC: $P = 0.01$; $\delta = 2$	No	286 (188; 371)	0.00291 (0.00250; 0.00330)
Scenario 10	200	1800	IC	OOC: $P = 0.01$; $\delta = 2$	No	300 (206; 375)	0.00283 (0.00247; 0.00322)
Scenario 11	5	1995	IC	OOC: $P = 0.01$; $\delta = 2$	Yes	346 (291; 392)	0.00271 (0.00241; 0.00303)
Scenario 12	200	1800	IC	OOC: $P = 0.01$; $\delta = 2$	Yes	345 (289; 390)	0.00271 (0.00241; 0.00303)
Scenario 13	5	1995	IC	OOC: $P = 0.1$; $\delta = 3$	No	23 (10; 40)	0.01414 (0.00674; 0.02353)
Scenario 14	200	1800	IC	OOC: $P = 0.1$; $\delta = 3$	No	44 (22; 72)	0.00756 (0.00476; 0.01089)
Scenario 15	5	1995	IC	OOC: $P = 0.1$; $\delta = 3$	Yes	342 (281; 390)	0.00272 (0.00240; 0.00304)
Scenario 16	200	1800	IC	OOC: $P = 0.1$; $\delta = 3$	Yes	343 (284; 391)	0.00272 (0.00240; 0.00304)

Abbreviations: $AARL$, average run length after updating; $AFAR$, expected false alarm rate after updating; $CARL$, conditional average run length; $CFAR$, conditional false alarm probability; EWMA, exponentially weighted moving average; IC, in control; OOC, out of control.

When there is a little contamination (scenarios 5 to 8), performance depends on the operator's ability to classify out-of-control signals correctly. When it is possible to determine the cause of the signal and trace back the start of the out-of-control situation, the resulting control chart performs well. On the other hand, when it is not possible to identify the cause and retrace the start of the signal, then many out-of-control samples will be left in the estimation sample (because smaller out-of-control levels are not detected quickly by the Shewhart control chart), resulting in less representative final control limits and worse performance.

6.2 | CUSUM and EWMA

As the conclusions for the CUSUM (Table 4) and EWMA (Table 5) control charts are very similar, we describe them together.

As with the Shewhart control chart, we can see that depending on the scenario, updating generates the same performance as direct estimation from a large Phase I dataset (reference scenario 3: $m_I = 2000$, $m_u = 0$). Mainly, in scenarios where the data are in control (scenarios 1 to 4) or where the operator is able to identify the cause of a signal and trace back the start (scenarios 7, 8, 11, 12, 15, and 16), control chart performance is similar to the performance of charts estimated on a large Phase I dataset.

When the operator is not able to identify the cause of a signal (and filter out previous out-of-control data samples), the performance of the updated control chart depends on the contamination level. Mainly, for major contamination (scenarios 13 and 14), out-of-control signals will not be detected quickly so that, until the chart gives a signal, many of the out-of-control samples are included in the estimation dataset, affecting the control chart limits.

7 | CONCLUSION

In the present article, we have simulated 16 scenarios, differing in the size of Phase I datasets ($m_I = 5, 200$), in the ability to determine the cause of a signal, and in the status (in or out of control) of the updating datasets. The charts' performances have been analyzed both during and after (at $t = 2000$) the updating period.

The results show improved chart behavior for updated limits when the updating dataset is in control, even when the size of the initial dataset is very small ($m_I = 5$). This holds for all three charts and means that excluding samples that give a false alarm from the estimation dataset does not affect control chart performance for these values of m_I and m_u.

For a low level of contamination (eg, $\delta = 0.5$), the limits can be updated safely for all three charts as long as the signal reason is known and the out-of-control data can be removed. When the signal reason is unknown, so that the origin of the out-of-control situation can not be retraced, the results show a decline in performance for all three charts. The Shewhart chart is especially vulnerable, as it has the least ability to quickly detect small shifts in the process mean.

For higher contamination levels (eg, $\delta = 2, 3$), it is safe to update the Shewhart chart control limits even if the signal reason is unknown. This is due to the chart's ability to detect all out-of-control signals, preventing that contaminated samples are included in the data set. For the CUSUM and EWMA charts, updating is only safe if the signal reason is known and the origin can be retraced. If this is not the case, the data and evolving statistic will get contaminated, and the performance of the CUSUM and EWMA charts is quite poor.

In summary, we definitely recommend updating control limits for the Shewhart, EWMA, and CUSUM charts as long as the reason for out-of-control signals is known and the origin can be retraced. If this is not the case, the best strategy depends on the size of the expected mean deviation. For large deviations, the Shewhart chart is safe to use, but the EWMA and CUSUM charts are not. For smaller deviations, the Shewhart chart fails, and performance of the CUSUM and EWMA charts is better. We suggest further research on the behavior of updating the limits for various subgroup sample sizes, as well as on performance for varying distributional assumptions.

REFERENCES

1. Shewhart WA. Quality control charts. *Bell Syst Tech J*. 1926;5:593-603.
2. Page ES. Continuous inspection schemes. *Biometrika*. 1954;41:100-115.
3. Roberts SW. Control chart tests based on geometric averages. *Technometrics*. 1959;42:97-101.
4. Vining G. Technical advice: phase I and phase II control charts. *Qual Eng*. 2009;21:478-479.
5. Jones LA, Champ CW, Rigdon SE. The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*. 2001;43:156-167.

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

7. Saleh NA, Mahmoud MA, Keefe MJ, Woodall WH. The difficulty in designing Shewhart $\overline{X}$ and X control charts with estimated parameters. *J Qual Technol*. 2015;47:127-138.

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

9. Hawkins DM. Self-starting CUSUM charts for location and scale. *Statistician*. 1987;18:299-315.

10. Quesenberry CP. SPC Q charts for start-up processes and short or long runs. *J Qual Technol*. 1991;23:213-224.

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

12. Goedhart R, Schoonhoven M, Does RJMM. Correction factors for Shewhart X and $\overline{X}$ control charts to achieve desired unconditional ARL. *Int J Prod Res*. 2016;54:7464-7479.

13. Goedhart R, Schoonhoven M, Does RJMM. On guaranteed in-control performance for the Shewhart X and $\overline{X}$ control charts. *J Qual Technol*. 2018;50:130-132.

14. Crosier RB. A new two-sided cumulative sum quality control scheme. *Technometrics*. 1986;28:187-194.

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

16. Fraker SE, Woodall WH, Mousavi S. Performance metrics for surveillance schemes. *Qual Eng*. 2008;20:451-464.

17. Chakraborti S, Human SW, Graham MA, Phase I. Statistical process control charts: an overview and some results. *Qual Eng*. 2009;21:52-62.

18. Frisén M. Optimal sequential surveillance for finance, public health, and other areas (with discussion). *Seq Anal Design Methods Appl*. 2009;28:310-337.

19. Hany M, Mahmoud MA. An evaluation of the Crosier's CUSUM control chart with estimated parameters. *Qual Reliab Eng Int*. 2016;32:1825-1835.

20. Saleh NA, Mahmoud MA, Abdel-Salam AG. The performance of the adaptive exponentially weighted moving average control chart with estimated parameters. *Qual Reliab Eng Int*. 2013;29:595-606.

AUTHOR BIOGRAPHIES

Leo C.E. Huberts is a PhD student in the Department of Operations Management and consultant at the Institute for Business and Industrial Statistics of the University of Amsterdam, the Netherlands. His current research topic is big data in statistical process monitoring.

Marit Schoonhoven is an associate professor at the Department of Operations Management and senior consultant at the Institute for Business and Industrial Statistics of the University of Amsterdam, the Netherlands. Her current research interests include control charting techniques and operations management methods.

Ronald J.M.M. Does is 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 a Fellow of the ASQ and ASA, an elected member of the ISI, and an Academician of the International Academy for Quality. His current research activities include the design of control charts for nonstandard situations, healthcare engineering, and operations management methods.