Heliyon

Received:
18 April 2018
Revised:
16 June 2018
Accepted:
5 October 2018

Cite as:
Huay Woon You. Optimal estimated process parameters side sensitive group runs chart based on expected average run length.
Heliyon 4 (2018) e00848.
doi: 10.1016/j.heliyon.2018.e00848

Optimal estimated process parameters side sensitive group runs chart based on expected average run length

Huay Woon You*

Pusat PERMATApintar Negara, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

*Corresponding author.
E-mail address: hwyou@ukm.edu.my (H.W. You).

Abstract

The side sensitive group runs (SSGR) chart is better than both the Shewhart and synthetic charts in detecting small and moderate process mean shifts. In practical circumstances, the process parameters are seldom known, so it is necessary to estimate them from in-control Phase-I samples. Research has discovered that a large number of in-control Phase-I samples are needed for the SSGR chart with estimated process parameters to behave similarly to a chart with known process parameters. The common metric to evaluate the performance of the control chart is average run length (ARL). An assumption for the computation of the ARL is that the shift size is assumed to be known. In reality however, the practitioners may not know the following shift size in advance. In light of this, the expected average run length (EARL) will be considered to measure the performance of the SSGR chart. Moreover, the standard deviation of the ARL (SDARL) will be studied, which is used to quantify the between-practitioner variability in the SSGR chart with estimated process parameters. This paper proposes the optimal design of the estimated process parameters SSGR chart based on the EARL criterion. The application of the optimal SSGR chart with estimated process parameters is demonstrated with actual data taken from a manufacturing company.

Keywords: Applied mathematics, Industry

https://doi.org/10.1016/j.heliyon.2018.e00848

1. Introduction

Customer satisfaction is essential in the business world. Consequently, enhancing quality is a primary factor for a successful business. To improve the quality of a product or service, statistical process control (SPC) is implemented to maintain and enhance the stability and capability of manufacturing and service processes. Among these SPC techniques, control charts are frequently used to detect shifts in a process. Due to its simplicity, the Shewhart chart is widely applied to detect large shifts in the process mean. The main setback of the Shewhart chart is its lack of sensitivity to small and moderate process mean shifts.

Since then, research in the area of control charts has focused on methods to enhance the efficiency of the Shewhart chart. In recent years, Gadre and Rattihalli [1] introduced the group runs (GR) chart, which integrates the Shewhart sub-chart and an extended version of the conforming run length (CRL) sub-chart. Moreover, Gadre and Rattihalli [2] suggested the side sensitive GR (SSGR) chart by using the side sensitive feature. Combined with an additional rule, it is shown that the SSGR chart is more efficient compared to the Shewhart, synthetic and GR charts.

Control charts are usually designed assuming that the in-control process parameters are known. However, in most practical applications, the assumption of known process parameters is usually violated and hence must be estimated from in-control Phase-I samples. Due to the variability of the estimators, the performance of the control chart with estimated process parameters will differ from the corresponding chart with known process parameters [3]. It has been proven that an extremely large number of m in-control Phase-I samples is required to attain a performance similar to the chart with known process parameters, as claimed in Quesenberry [4], Jones et al. [5], Psarakis et al. [6], Jones-Farmer et al. [7], to name a few. Therefore, the impact of parameter estimation must be considered when designing the control chart.

A common performance measure is the average run length (ARL). The ARL is defined as the average number of sample results plotted on the control chart before the process signals out-of-control [8]. The computation of the ARL requires the shift size to be known. In a practical situation, the next shift size is not known in advance. To address this problem, the shift size should be considered as random. Hence, the expected ARL (EARL) will be employed as a performance measure when the shift size is unknown. Related studies on the EARL can be found in Celano [9], Castagliola et al. [10], and Yeong et al. [11]. Recently, You et al. [12] considered the SSGR chart when estimating the process parameters based on minimising ARL, which has motivated the current research in this paper. To the best of the author's knowledge, research investigating the application of the EARL is very limited. Hence, in this article, optimal design is developed for the

https://doi.org/10.1016/j.heliyon.2018.e00848

SSGR chart with estimated process parameters by minimising the out-of-control EARL for unknown shift sizes.

When the process parameters are estimated from m in-control Phase-I samples, the control chart performance will vary among practitioners. This is because practitioners use different m Phase-I samples to estimate the process parameters, which result in different parameter estimates. This variation has been named as practitioner-to-practitioner variability. In light of this, standard deviation of the ARL (SDARL) will be considered as a measure of the amount of practitioner-to-practitioner variability in the performance of the SSGR chart. The SDARL performance measure has been employed to determine the necessary amount of Phase-I samples for control charts with estimated process parameters. For an overview of SDARL, see Saleh et al. [3], Aly et al. [13] and Saleh et al. [14].

This paper is structured as follows. Section 2 briefly reviews the SSGR chart. The run length properties of the SSGR chart with known and estimated process parameters are given in Section 3. Section 4 outlines the optimal design of the SSGR chart with estimated process parameters by minimising the EARL. A performance comparison between the SSGR chart with known and estimated process parameters, in terms of the EARL, ARL and SDARL, is discussed in Section 5. The application of the SSGR chart with estimated process parameters is illustrated using actual industrial data in Section 6. Finally, concluding remarks are given in Section 7.

2. Background

2.1. The SSGR chart

The SSGR chart combines the Shewhart sub-chart and an extended version of the CRL sub-chart [2]. Eq. (1) shows the upper control limit (UCL) and lower control limit (LCL) of the Shewhart sub-chart when process parameters, i.e., the in-control mean, μ_0 and in-control standard deviation, σ_0 are known:

$$\mathrm{UCL/LCL} = \mu_0 \pm \frac{K}{\sqrt{n}}\sigma_0, \tag{1}$$

with K as the charting constant. The extended version of the CRL sub-chart is an attribute chart with one lower limit, i.e., L. The rth CRL value, CRL_r, for $r = 1, 2, \ldots$, is defined as the number of samples between two consecutive non-conforming samples from the Shewhart sub-chart (inclusive of the ending non-conforming sample). The SSGR chart has an implicit inclusion of headstart features, where the sample is out-of-control at the start of process monitoring [15]. The operation of the SSGR chart is as follows:

Step 1: Set the charting constants LCL, UCL and L.

https://doi.org/10.1016/j.heliyon.2018.e00848

Step 2: Take a sample of n items and compute the sample mean, $\overline{X}$.

Step 3: If $\overline{X} \in [\text{LCL}, \text{UCL}]$, the sample is considered as conforming and the control flow returns to Step 2. Otherwise, the sample is considered as non-conforming and the control flow moves to Step 4.

Step 4: (i) For the first CRL, if $\text{CRL}_1 \leq L$, the process is classified as out-of-control and the control flow moves to Step 5. Otherwise, the process is in-control and the control flow returns to Step 2.

(ii) From the second CRL onwards, if $\text{CRL}_r \leq L$, for $r = 2, 3, \ldots$, the process is not immediately classified as out-of-control, and the control flow returns to Step 2. However, if $\text{CRL}_r \leq L$ and $\text{CRL}_{r+1} \leq L$, for $r = 2, 3, \ldots$, and both have shifts on the same side of the Shewhart sub-chart, the process is regarded as out-of-control and the control flow proceeds to Step 5.

Step 5: Immediate corrective actions are taken to eliminate the assignable cause(s). Then, the control flow returns to Step 2.

3. Theory/calculation

3.1. Run length properties of the SSGR chart with known and estimated process parameters

Let P_1 denote the probability of a non-conforming sample on the Shewhart sub-chart, i.e., [2]:

$$P_1 = 1 - \Pr\left(\overline{X} \in [\text{LCL}, \text{UCL}] \middle| \overline{X} \sim N\left(\mu_0 + \delta\sigma_0, \frac{\sigma_0^2}{n}\right)\right), \tag{2}$$

where LCL and UCL are stated in Eq. (1). From Eq. (2), it can then be shown that:

$$P_1 = 1 - \left[\Phi\left(K - \delta\sqrt{n}\right) - \Phi\left(-K - \delta\sqrt{n}\right)\right]. \tag{3}$$

whereby δ represents the size of a standardised mean shift where a quick detection is required.

Additionally, the probability of an event $\text{CRL}_r \leq L$ is [2]:

$$A = 1 - (1 - P_1)^L \tag{4}$$

and the conditional probability $h = \Pr(\overline{X} > \text{UCL} | \overline{X} \notin [\text{LCL}, \text{UCL}])$, taking into account of the side sensitivity feature:

$$\begin{aligned} h &= \frac{\Pr\left(\overline{X} > \text{UCL}\right)}{P_1} \\ &= \frac{1 - \Phi(K - \delta\sqrt{n})}{P_1}. \end{aligned} \tag{5}$$

https://doi.org/10.1016/j.heliyon.2018.e00848

The ARL of the SSGR chart with known process parameters is:

$$\mathrm{ARL} = \frac{1 - h(1-h)A^2}{P_1 A^2 [1 + h(1-h)(A-2)]}, \tag{6}$$

where the P_1, A and h can be obtained from Eqs. (3), (4), and (5), respectively.

When the exact shift size δ is unknown, it is essential to consider the EARL as the alternative performance measure. The EARL is an overall range of shifts $(\delta_{\min}, \delta_{\max})$, where $\delta_{\min}$ and $\delta_{\max}$ represent the lower and upper bounds of the mean shift, respectively. The EARL of the SSGR chart is:

$$\mathrm{EARL} = \int_{\delta_{\min}}^{\delta_{\max}} \mathrm{ARL} f(\delta) d\delta, \tag{7}$$

where ARL is given in Eq. (6) and $f(\delta)$ is the probability density function (pdf) of the shift size, δ. As it is difficult to estimate the actual shape of $f(\delta)$, Castagliola et al. [10] and Tang et al. [16] assumed that the process mean shift follows a uniform distribution $U(\delta_{\min}, \delta_{\max})$.

When the parameters of a process, i.e., the in-control mean, μ_0, and in-control standard deviation, σ_0, are unknown, they are estimated from in-control Phase-I samples which consist of m samples, each of size n. Let the in-control Phase-I samples be $\{Y_{i,1}, Y_{i,2}, ..., Y_{i,n}\}$, where $i = 1, 2, \ldots, m$. An estimator of μ_0 is [12]:

$$\widehat{\mu}_0 = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} Y_{i,j} \tag{8}$$

and an estimator of σ_0 is:

$$\widehat{\sigma}_0 = \sqrt{\frac{1}{m(n-1)} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(Y_{i,j} - \overline{Y}_i\right)^2}, \tag{9}$$

where $\overline{Y}_i$ is the sample mean for the ith sample, i.e., $\overline{Y}_i = \frac{1}{n} \sum_{j=1}^{n} Y_{i,j}$.

After estimating μ_0 and σ_0 from the in-control Phase-I samples, the monitoring of a Phase-II process begins. The sample means, $\overline{X}_i$, for $i = \{1, 2, \ldots\}$ are plotted on the Shewhart sub-chart in the Phase-II process monitoring. Based on the estimated process parameters, i.e., $\widehat{\mu}_0$ and $\widehat{\sigma}_0$, from the Phase-I parameter estimation, the LCL and UCL of the Shewhart sub-chart are:

$$\widehat{\mathrm{UCL}}/\widehat{\mathrm{LCL}} = \widehat{\mu}_0 \pm \frac{K'}{\sqrt{n}} \widehat{\sigma}_0, \tag{10}$$

where K' is the charting constant of the Shewhart sub-chart with estimated process parameters.

https://doi.org/10.1016/j.heliyon.2018.e00848

Then, the conditional probability that a sample on the Shewhart sub-chart is non-conforming is given as:

$$\begin{aligned}\widehat{P}_1 &= 1 - \Pr\left(\overline{X}_i \in [\mathrm{LCL}, \mathrm{UCL}] | \widehat{\mu}_0, \widehat{\sigma}_0\right) \\ &= 1 - \Pr\left(\widehat{\mu}_0 - \frac{K'}{\sqrt{n}}\widehat{\sigma}_0 \leq \overline{X}_i \leq \widehat{\mu}_0 + \frac{K'}{\sqrt{n}}\widehat{\sigma}_0\right).\end{aligned} \tag{11}$$

Since $\overline{X}_i \sim N\left(\mu_0 + \delta\sigma_0, \frac{\sigma_0^2}{n}\right)$, subtracting $\mu_0 + \delta\sigma_0$ followed by multiplying $\frac{\sqrt{n}}{\sigma_0}$ on both sides of the inequality in Eq. (11), then simplifying will yield [12]:

$$\widehat{P}_1 = 1 - \left[\Phi\left(U + \frac{K'}{\sqrt{n}}V - \delta\sqrt{n}\right) - \Phi\left(U - \frac{K'}{\sqrt{n}}V - \delta\sqrt{n}\right)\right], \tag{12}$$

by denoting $U = (\widehat{\mu}_0 - \mu_0)\frac{\sqrt{n}}{\sigma_0}$ and $V = \frac{\widehat{\sigma}_0\sqrt{n}}{\sigma_0}$.

Since $\widehat{\mu}_0 \sim N\left(\mu_0, \frac{\sigma_0^2}{mn}\right)$, it can be deduced that $U \sim N\left(0, \frac{1}{m}\right)$. Then, the pdf of U can be represented as:

$$f_U(u|m) = f_N\left(u \middle| 0, \frac{1}{m}\right), \tag{13}$$

where f_N is the pdf of a normal distribution having a mean of 0 and variance of $\frac{1}{m}$. Moreover, using $\frac{\widehat{\sigma}_0^2}{\sigma_0^2} \sim \gamma\left(\frac{m(n-1)}{2}, \frac{2}{m(n-1)}\right)$, Zhang et al. [17] showed that $V^2 = \frac{\widehat{\sigma}_0^2 n}{\sigma_0^2} \sim \gamma\left(\frac{m(n-1)}{2}, \frac{2n}{m(n-1)}\right)$, which is a gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2n}{m(n-1)}$. Consequently, the pdf of V is:

$$f_V(v|m,n) = 2v f_\gamma\left(v^2 \middle| \frac{m(n-1)}{2}, \frac{2n}{m(n-1)}\right), \tag{14}$$

where f_γ is the pdf of the gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2n}{m(n-1)}$.

In a similar manner, the probability of an event $\mathrm{CRL}_r \leq L'$ is:

$$\widehat{A} = \Pr(\mathrm{CRL}_r \leq L' | \widehat{\mu}_0, \widehat{\sigma}_0) = 1 - \left(1 - \widehat{P}_1\right)^{L'}, \tag{15}$$

where L' is the lower limit of the CRL sub-chart with estimated process parameters. In addition, the SSGR chart considers the location where the sample falls on the Shewhart sub-chart, which influences the signalling event of the SSGR chart. Therefore, the conditional probability, $\Pr(\overline{X}_i > \mathrm{UCL} | \overline{X}_i \notin [\mathrm{LCL}, \mathrm{UCL}], \widehat{\mu}_0, \widehat{\sigma}_0)$, i.e.:

$$\widehat{h} = \frac{1 - \Phi\left(U + \frac{K'}{\sqrt{n}}V - \delta\sqrt{n}\right)}{\widehat{P}_1}. \tag{16}$$

Note that for complete and detailed derivation, the reader can refer to You et al. [12].

Consequently, the ARL of the SSGR chart with estimated process parameters being equal to:

https://doi.org/10.1016/j.heliyon.2018.e00848

$$\mathrm{ARL} = \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \frac{1-\widehat{h}\left(1-\widehat{h}\right)\widehat{A}^2}{\widehat{P}_1\widehat{A}^2\left(1+\widehat{h}\left(1-\widehat{h}\right)\left(\widehat{A}-2\right)\right)} f_U(u|m) f_V(v|m,n)\,dv\,du \tag{17}$$

where $\widehat{P}_1$, $f_U(u|m)$, $f_V(v|m,n)$, $\widehat{A}$ and $\widehat{h}$ are obtained from Eqs. (12), (13), (14), (15), and (16), respectively. In addition, the SDARL of the SSGR chart when the process parameters are estimated is

$$\mathrm{SDARL} = \left[E\left(\mathrm{ARL}^2\right) - [\mathrm{ARL}]^2\right]^{1/2}. \tag{18}$$

Moreover, the EARL of the SSGR chart with estimated process parameters is:

$$\mathrm{EARL} = \int_{\delta_{\min}}^{\delta_{\max}} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \frac{1-\widehat{h}\left(1-\widehat{h}\right)\widehat{A}^2}{\widehat{P}_1\widehat{A}^2\left(1+\widehat{h}\left(1-\widehat{h}\right)\left(\widehat{A}-2\right)\right)} f(\delta) f_U(u|m) f_V(v|m,n)\,dv\,du\,d\delta. \tag{19}$$

4. Design

4.1. Optimal design of the estimated process parameters SSGR chart

The performance of the SSGR chart with estimated process parameters differs significantly from the same chart with known process parameters. This is due to the existence of variability when process parameters are estimated, whereas variability does not exist when process parameters are known. Thus, adopting optimal charting parameters of the chart with known process parameters when process parameters are estimated will ignore the existence of variability. Moreover, in many applications, the magnitude of mean shift δ is seldom known in advance. In light of this, the optimal design to obtain optimal charting parameters (K', L') of the SSGR chart with estimated process parameters is necessary to overcome the circumstances. The optimization of the SSGR chart with estimated process parameters consists of finding the chart's parameters (K', L') which satisfy the following conditions:

$$(K', L') = \underset{(K,\,L)}{\mathrm{argmin}} \mathrm{EARL}(m, n, K, L, \delta_{\min}, \delta_{\max}). \tag{20}$$

subject to $\mathrm{ARL}(m, n, K, L, \delta = 0) = \mathrm{ARL}_0$, where $\mathrm{EARL}(m, n, K, L, \delta_{\min}, \delta_{\max})$ is computed using Eq. (19). Hence, the optimal charting parameters (K', L') fulfil the condition such that:

(i) When $\delta = 0$, $\mathrm{ARL}_0 = \mathrm{EARL}_0$, where EARL_0 is the intended in-control EARL value. In this paper, the $\mathrm{EARL}_0 = 370.4$ is used.

https://doi.org/10.1016/j.heliyon.2018.e00848

(ii) When $\delta \neq 0$, the optimal charting parameters (K', L') yield the smallest EARL_1 value.

The ScicosLab software version 4.4.2. (www.scicoslab.org) is used to compute the optimal charting parameters (K', L') of the SSGR chart with estimated process parameters. The procedure to compute the optimal charting parameters (K', L') is described as follows:

Step 1: Specify m, n, $\delta_{\min}$, $\delta_{\max}$ and EARL_0.

Step 2: Initialize L' as unity.

Step 3: Compute K' using a nonlinear equation solver so that EARL_0 specified in Step (1) is achieved.

Step 4: Compute the EARL_1 using Eq. (19), based on the current charting parameters (K', L').

Step 5: If $L' = 1$ or "$L' > 1$ and EARL_1 has been reduced", increase L' by one and return to Step (3). Otherwise, proceed to the next step.

Step 6: Select the charting parameters (K', L') that give the smallest EARL_1 as the optimal charting parameters.

Note that a similar approach is used to obtain the optimal charting parameters (K', L') of the SSGR chart with estimated process parameters for minimising out-of-control ARL (ARL_1) by replacing EARL_1 in Steps 1−6 with ARL_1.

5. Results & discussion

5.1. Performance analysis of the SSGR chart

Since the performance of the SSGR chart with estimated process parameters and known process parameters is significantly different, it is essential to determine the minimum number of in-control Phase-I samples, i.e., m^*, so that the SSGR chart with estimated process parameters has approximately the same EARL performance as the chart with known process parameters. For this purpose, Table 1 displays the value of m^* to satisfy the condition of (i) $\Delta < 0.01$ and (ii) $\Delta < 0.05$, where $\Delta = \left|\mathrm{EARL}_{1,m^*} - \mathrm{EARL}_{1,\infty}\right| / \mathrm{EARL}_{1,\infty}$. Note that EARL_{1,m^*} is the out-of-control EARL for the SSGR chart with estimated process parameters computed using the optimal charting parameters (K, L), (that corresponding to the known process parameters chart given in the last column (i.e., $m = +\infty$) of Table 2) based on sample m^*. Moreover, $\mathrm{EARL}_{1,\infty}$ is calculated using Eq. (7), based on the optimal charting parameters (K, L).

From Table 1, we can observe that the value of m^* required by $\Delta < 0.01$ is larger than that required by $\Delta < 0.05$. This finding is attributed to the fact that the performance of the SSGR chart with estimated process parameters is closer to the

https://doi.org/10.1016/j.heliyon.2018.e00848

Table 1. Minimum m (represented as m^*), required by the estimated process parameters SSGR chart, for different combinations of $(n, \delta_{\min}, \delta_{\max})$, based on the optimal charting parameters (K, L) corresponding to n, $\delta_{\min}$ and $\delta_{\max}$ of the chart with known process parameters, achieving $\Delta < 0.01$ and $\Delta < 0.05$, where $\Delta = \left|\mathrm{EARL}_{1,m^*} - \mathrm{EARL}_{1,\infty}\right| / \mathrm{EARL}_{1,\infty}$.

Δ	$\delta_{\min}$	$\delta_{\max}$	n			
			3	**4**	**5**	**6**
$\Delta < 0.01$	0.2	1.0	2362	1693	1357	1147
	1.0	2.0	120	54	33	20
$\Delta < 0.05$	0.2	1.0	490	351	282	238
	1.0	2.0	28	14	9	6

corresponding chart with known process parameters for $\Delta < 0.01$ compared to $\Delta < 0.05$. For example, for $n = 4$, $\delta_{\min} = 0.2$ and $\delta_{\max} = 1.0$, the m^* for $\Delta < 0.01$ is 1693, compared to 351 for $\Delta < 0.05$. In addition, the value of m^* decreases when n, $\delta_{\min}$ and $\delta_{\max}$ increase, regardless of whether $\Delta < 0.01$ or $\Delta < 0.05$. Table 1 is useful for practitioners who want to implement optimal charting parameters (K, L) based on known process parameters on the chart with estimated process parameters yet to obtain similar performance as the chart with known process parameters. However, the value of m^* can be quite large when n, $\delta_{\min}$, $\delta_{\max}$ and Δ are small. For instance, when $n = 3$, $\delta_{\min} = 0.2$, $\delta_{\max} = 1.0$ and $\Delta < 0.01$, the value of m^* is 2362.

Table 2. Optimal charting parameters (K', L') and the corresponding EARL_1s, for $n = \{3, 4, 5, 6\}$, with different combinations of $(m, \delta_{\min}, \delta_{\max})$ when $\mathrm{EARL}_0 = 370.4$.

n	$\delta_{\min}$	$\delta_{\max}$	m						
			10	**25**	**30**	**40**	**50**	**80**	$+\infty$
3	0.2	1.0	(2.1075, 92)	(2.2241, 39)	(2.2312, 35)	(2.2336, 30)	(2.2305, 27)	(2.2316, 24)	(2.2284, 20)
			50.07	34.10	32.42	30.31	29.03	27.10	23.84
	1.0	2.0	(1.6783, 5)	(1.7416, 4)	(1.7530, 4)	(1.6895, 3)	(1.6961, 3)	(1.7055, 3)	(1.7185, 3)
			1.44	1.43	1.43	1.42	1.42	1.42	1.41
4	0.2	1.0	(2.1862, 57)	(2.2230, 28)	(2.2235, 26)	(2.2156, 23)	(2.2161, 22)	(2.2086, 20)	(2.1886, 17)
			38.91	25.69	24.23	22.43	21.35	19.77	17.19
	1.0	2.0	(1.6417, 3)	(1.6970, 3)	(1.7021, 3)	(1.7078, 3)	(1.7109, 3)	(1.7148, 3)	(1.7185, 3)
			1.24	1.22	1.22	1.21	1.21	1.21	1.20
5	0.2	1.0	(2.2120, 41)	(2.2122, 23)	(2.2126, 22)	(2.2027, 20)	(2.1972, 19)	(2.1928, 18)	(2.1735, 16)
			31.17	19.95	18.74	17.28	16.42	15.17	13.18
	1.0	2.0	(1.6767, 3)	(1.7117, 3)	(1.7144, 3)	(1.5979, 2)	(1.5982, 2)	(1.5981, 2)	(1.5953, 2)
			1.14	1.12	1.12	1.12	1.12	1.11	1.11
6	0.2	1.0	(2.2210, 33)	(2.1991, 20)	(2.1942, 19)	(2.1892, 18)	(2.1796, 17)	(2.1697, 16)	(2.1401, 14)
			25.48	15.93	14.95	13.76	13.08	12.09	10.54
	1.0	2.0	(1.6984, 3)	(1.6026, 2)	(1.6029, 2)	(1.6025, 2)	(1.6020, 2)	(1.6004, 2)	(1.5953, 2)
			1.08	1.07	1.07	1.07	1.07	1.06	1.06

https://doi.org/10.1016/j.heliyon.2018.e00848

In reality, it is impractical to accumulate a large m^* value in a Phase-I process due to cost and time constraints. Hence, the SSGR chart with estimated process parameters is optimally designed and the optimal charting parameters (K', L') are computed based on the procedure illustrated in Section 4 and Eq. (20). The computed optimal charting parameters (K', L') will enable practitioners to have almost the same performance as the known process parameters based SSGR chart with a reasonable number of Phase-I samples, m. In Table 2, the optimal charting parameters (K', L') and the corresponding EARL_1 for each $(\delta_{\min}, \delta_{\max})$ are displayed in the first and second rows, respectively. The results in Table 2 have been verified using the Monte-Carlo simulation. From Table 2, for the same n, $\delta_{\min}$ and $\delta_{\max}$, the value of EARL_1 generally decreases with an increase in m. This is because the performance of the SSGR chart with estimated process parameters will behave more closely to the corresponding chart with known process parameters when the value of m increases. Note that a decreasing EARL_1 value indicates better performance.

Moreover, the proposed optimal design of the SSGR chart with estimated process parameters for computing optimal charting parameters (K', L') is crucial, as it allows the performance of the SSGR chart with estimated process parameters to approach that of the chart with known process parameters. For illustration, consider $m = 30$, $n = 3$, $\delta_{\min} = 0.2$ and $\delta_{\max} = 1.0$, where the optimal charting parameters $(K', L') = (2.2312, 35)$ yields $\text{EARL}_1 = 32.42$. However, instead of these optimal charting parameters (K', L') for the estimated process parameters based SSGR chart being used, we consider the optimal charting parameters $(K, L) = (2.2284, 20)$ (see the last column of Table 2) for the SSGR chart with known process parameters. $\text{EARL}_1 = 69.55$ is obtained using the ScicolsLab programme. Obviously, $\text{EARL}_1 = 32.42$ is closer to $\text{EARL}_1 = 23.84$ (see the last column of Table 2) compared to $\text{EARL}_1 = 69.55$. In light of this, using the optimal charting parameters for the SSGR chart with estimated process parameters is recommended.

To further illustrate the implementation of the proposed optimal charting parameters, Table 3 provides the optimal charting parameters (K', L') and the corresponding ARL_1 for different combinations of (m, n, δ) based on minimising ARL_1. Note that the ARL_1s for the SSGR chart with estimated process parameters are computed using Eq. (17). For comparison purposes, $\delta \in \{0.2, 0.5, 0.9, 1.2, 1.6, 2.0\}$ is considered. Here, $\delta \in \{0.2, 0.5, 0.9\}$ and $\delta \in \{1.2, 1.6, 2.0\}$ are included in $(\delta_{\min}, \delta_{\max}) = (0.2, 1.0)$ and $(\delta_{\min}, \delta_{\max}) = (1.0, 2.0)$, respectively. In Table 2, when $m = 40$, $n = 3$, $\delta_{\min} = 0.2$ and $\delta_{\max} = 1.0$, the $\text{EARL}_1 = 30.31$ is obtained using the optimal charting parameters $(K', L') = (2.2336, 30)$. By employing the same optimal charting parameters for $\delta = 0.5$ (i.e. $\delta \in (\delta_{\min}, \delta_{\max})$), $\text{ARL}_1 = 19.03$ is yielded using the ScicolsLab programme. It is observed that the ARL_1 value is almost the same, i.e., $\text{ARL}_1 = 18.83$, when using the optimal charting parameters $(K', L') = (2.1694, 22)$ based on minimising ARL_1 (see Table 3). This reveals that the optimal charting parameters obtained based on minimising EARL_1 can be implemented as long as $\delta \in (\delta_{\min}, \delta_{\max})$,

https://doi.org/10.1016/j.heliyon.2018.e00848

Table 3. Optimal charting parameters (K', L') and the corresponding ARL_1s, for $n = \{3, 4, 5, 6\}$, with different combinations of (m, δ) when $ARL_0 = 370.4$.

n	*δ*	*m*					
		10	**30**	**40**	**50**	**80**	**+ ∞**
3	0.2	(2.2087, 227)	(2.3998, 90)	(2.4120, 76)	(2.4157, 68)	(2.4193, 58)	(2.4125, 44)
		240.01	179.39	168.54	161.43	149.79	127.88
	0.5	(2.0982, 85)	(2.1659, 25)	(2.1694, 22)	(2.1658, 20)	(2.1662, 18)	(2.1574, 15)
		34.18	19.82	18.83	18.27	17.46	16.11
	0.9	(1.8302, 12)	(1.8904, 7)	(1.8709, 6)	(1.8809, 6)	(1.8954, 6)	(1.8660, 5)
		3.52	3.33	3.30	3.28	3.25	3.19
	1.2	(1.6783, 5)	(1.7530, 4)	(1.7669, 4)	(1.7749, 4)	(1.7862, 4)	(1.7185, 3)
		1.7477	1.7407	1.74	1.74	1.73	1.72
	1.6	(1.5764, 3)	(1.5665, 2)	(1.5751, 2)	(1.5799, 2)	(1.5866, 2)	(1.5953, 2)
		1.18	1.18	1.18	1.18	1.17	1.17
	2.0	(1.4872, 2)	(1.5665, 2)	(1.5751, 2)	(1.5799, 2)	(1.5866, 2)	(1.5953, 2)
		1.04	1.04	1.04	1.04	1.03	1.03
4	0.2	(2.3101, 141)	(2.4026, 64)	(2.4048, 57)	(2.4042, 53)	(2.4034, 48)	(2.3853, 39)
		213.99	150.52	139.23	132.00	120.58	100.53
	0.5	(2.1479, 44)	(2.1318, 17)	(2.1192, 15)	(2.1117, 14)	(2.1056, 13)	(2.1009, 12)
		22.02	13.30	12.61	12.22	11.66	10.73
	0.9	(1.8271, 7)	(1.8389, 5)	(1.8474, 5)	(1.8521, 5)	(1.7968, 4)	(1.8025, 4)
		2.59	2.40	2.37	2.36	2.33	2.28
	1.2	(1.6417, 3)	(1.7021, 3)	(1.7078, 3)	(1.7109, 3)	(1.7148, 3)	(1.7185, 3)
		1.43	1.40	1.39	1.39	1.38	1.37
	1.6	(1.5422, 2)	(1.5865, 2)	(1.5902, 2)	(1.5921, 2)	(1.5943, 2)	(1.5953, 2)
		1.08	1.07	1.07	1.07	1.07	1.06
	2.0	(1.5422, 2)	(1.5865, 2)	(1.5902, 2)	(1.5921, 2)	(1.5943, 2)	(1.5953, 2)
		1.01	1.01	1.01	1.01	1.01	1.01
5	0.2	(2.3521, 100)	(2.3933, 52)	(2.3931, 48)	(2.3886, 45)	(2.3807, 41)	(2.3606, 35)
		191.59	126.67	115.83	109.08	98.73	81.44
	0.5	(2.1401, 27)	(2.0926, 13)	(2.0825, 12)	(2.0659, 11)	(2.0719, 11)	(2.0537, 10)
		15.15	9.54	9.09	8.83	8.44	7.81
	0.9	(1.7994, 5)	(1.7942, 4)	(1.7981, 4)	(1.8000, 4)	(1.8021, 4)	(1.8025, 4)
		2.04	1.90	1.88	1.87	1.86	1.82
	1.2	(1.6767, 3)	(1.7144, 3)	(1.5979, 2)	(1.5982, 2)	(1.5981, 2)	(1.5953, 2)
		1.25	1.23	1.22	1.22	1.22	1.21
	1.6	(1.5714, 2)	(1.5967, 2)	(1.5979, 2)	(1.5982, 2)	(1.5981, 2)	(1.5953, 2)
		1.03	1.03	1.03	1.03	1.03	1.03
	2.0	(1.5714, 2)	(1.5967, 2)	(1.5979, 2)	(1.5982, 2)	(1.5981, 2)	(1.5953, 2)
		1.00	1.00	1.00	1.00	1.00	1.00
6	0.2	(2.3693, 78)	(2.3866, 46)	(2.3786, 42)	(2.3741, 40)	(2.3647, 37)	(2.3400, 32)
		171.79	107.34	97.34	91.26	82.17	67.53
	0.5	(2.1162, 19)	(2.0417, 10)	(2.0467, 10)	(2.0225, 9)	(2.0252, 9)	(1.9948, 8)
		10.96	7.25	6.94	6.76	6.49	6.05
	0.9	(1.7714, 4)	(1.7219, 3)	(1.7227, 3)	(1.7228, 3)	(1.7223, 3)	(1.7185, 3)
		1.71	1.61	1.60	1.59	1.58	1.55
	1.2	(1.5895, 2)	(1.6029, 2)	(1.6025, 2)	(1.6020, 2)	(1.6004, 2)	(1.5953, 2)
		1.15	1.13	1.13	1.12	1.12	1.12
	1.6	(1.5895, 2)	(1.6029, 2)	(1.6025, 2)	(1.6020, 2)	(1.6004, 2)	(1.5953, 2)
		1.01	1.01	1.01	1.01	1.01	1.01
	2.0	(1.5895, 2)	(1.6029, 2)	(1.6025, 2)	(1.6020, 2)	(1.6004, 2)	(1.5953, 2)
		1.00	1.00	1.00	1.00	1.00	1.00

https://doi.org/10.1016/j.heliyon.2018.e00848

i.e., when the practitioners do not have prior knowledge to determine the exact process shift size.

Table 4 displays the in-control SDARL, i.e. when $\delta = 0$ for different values of m, ranging from 30 to 5000, with sample sizes $n = 3$, 4, 5 and 6. The SDARL is computed using Eq. (18). Similar charting parameters $(K, L) = (1.3712, 1)$ have been used to compute the in-control SDARL under different values of m. When the process parameters are known, the charting parameters $(K, L) = (1.3712, 1)$ produce $\text{ARL}_0 = 370.4$. The SDARL has been employed to measure the amount of practitioner-to-practitioner variability in the performance of the SSGR chart with estimated process parameters. According to Zhang et al. [18], an SDARL within 10% of the ARL_0 is reasonable, although still reflecting a significant amount of variation. In view of this, from Table 4, a practitioner would need about 3000, 3000, 800 and 700 Phase-I samples of size $n = 3$, 4, 5 and 6, respectively, to obtain SDARL values of no more than 37.04 (10% of 370.4). We can observe that the number of m Phase-I samples decreases when the sample size increases.

6. Example

6.1. A real-life application

This section demonstrates the construction of the proposed optimal SSGR chart with estimated process parameters using an actual dataset for the flow width measurements (in microns) of the hard-bake process. The dataset for the flow width measurements of the hard-bake process is taken from Montgomery [19]. From prior experience, the proposed SSGR chart is designed to identify a process mean shift size between $\delta_{\min} = 0.2$ and $\delta_{\max} = 1.0$, where $\text{EARL}_0 = 370.4$ is desired.

In this example, $m = 25$ Phase-I samples, each sample with a size of $n = 5$ wafers, are used to estimate the process parameters. A Bonferroni-type adjustment, i.e., Bonferroni-adjusted $\overline{X}$ and S charts, is performed to ensure that these Phase-I samples are in-control before reliable estimated process parameters are computed. The control limits of the Bonferroni-adjusted $\overline{X}$ and S charts are calculated with:

$$\text{UCL}_{\overline{X}}/\text{LCL}_{\overline{X}} = \overline{\overline{X}} \pm Z_{FAP/(2m)}\frac{(\overline{S}/c_4)}{\sqrt{n}} \tag{21}$$

and

$$\text{UCL}_S/\text{LCL}_S = \overline{S} \pm Z_{FAP/(2m)}\sqrt{1-c_4^2}\left(\overline{S}/c_4\right), \tag{22}$$

respectively, where $\overline{\overline{X}} = 1.5056$ is the sample grand mean, and $\overline{S} = 0.1316$ is the average sample standard deviation. Here, Z_ς is the $(1-\varsigma)100$th percentile of the standard normal distribution, c_4 is an unbiased constant, and FAP = 0.1264 is

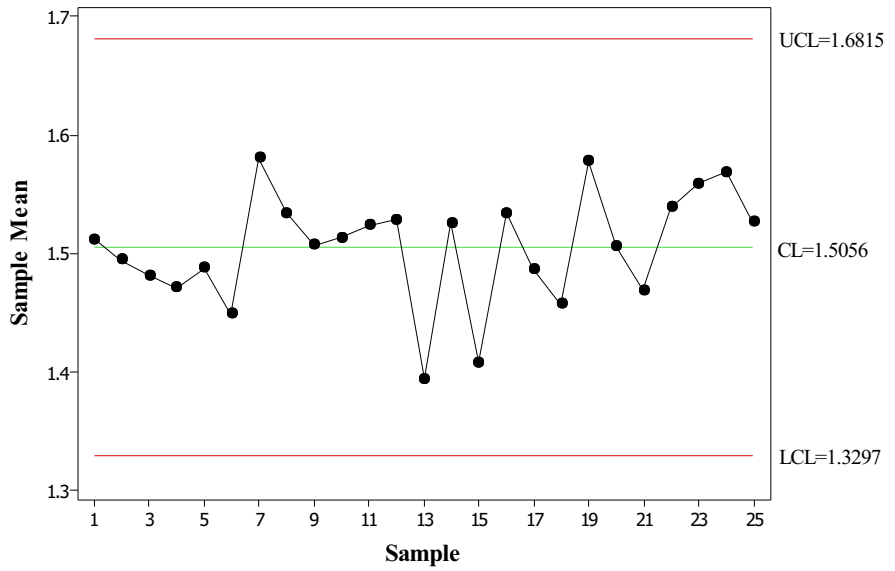
https://doi.org/10.1016/j.heliyon.2018.e00848

Table 4. In-control SDARL based on $(K, L) = (1.3712, 1)$ when m Phase-I samples with sample sizes, $n = 3, 4, 5$ and 6 are used to estimate the in-control process parameters.

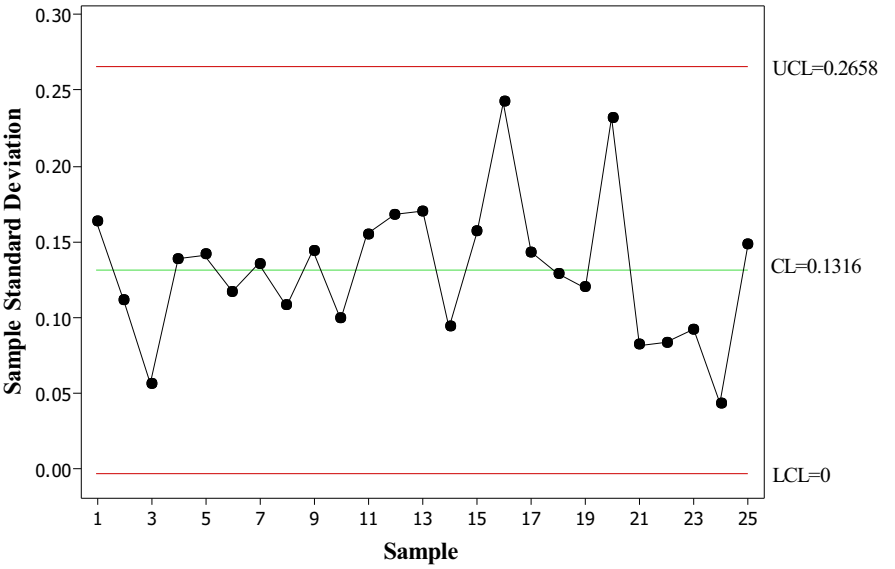
n	*m*														
	30	**50**	**100**	**200**	**300**	**400**	**500**	**600**	**700**	**800**	**900**	**1000**	**3000**	**5000**	$+\infty$
3	366.83	244.42	156.74	106.01	85.33	73.38	65.36	59.50	54.98	51.35	48.36	45.83	26.32	20.36	0.00
4	256.68	183.82	123.38	85.11	68.92	59.44	53.03	48.33	44.69	41.76	39.35	37.31	21.47	16.62	0.00
5	208.48	154.06	105.44	73.31	59.49	51.35	45.84	41.79	38.65	36.13	34.05	32.28	18.59	14.39	0.00
6	180.82	135.87	93.91	65.50	53.19	45.93	41.00	37.38	34.57	32.32	30.45	28.88	16.63	12.87	0.00

https://doi.org/10.1016/j.heliyon.2018.e00848

the false alarm probability. Note that FAP is computed using the formula FAP $= 1 - (1 - \alpha)^{2m}$, where $\alpha = 0.0027$ is used in this example. Consequently, FAP $= 1 - (1 - 0.0027)^{50} = 0.1246$, which is the probability of at least one false alarm out of every 50 samples. Then, $Z_{\mathrm{FAP}/2m} = Z_{0.1264/50} = Z_{0.002528} = 2.81$ is obtained from the standard normal cdf table. The $c_4 = 0.9400$ for $n = 5$ is obtained from Statistical Quality Control text books. Therefore, from Eqs. (21) and (22), the control limits of the Bonferroni-adjusted $\overline{X}$ and S charts are:



(a)



(b)

Fig. 1. (a) Bonferroni-adjusted $\overline{X}$ chart (b) Bonferroni-adjusted S chart.

https://doi.org/10.1016/j.heliyon.2018.e00848

$$\mathrm{UCL}_{\overline{X}}/\mathrm{LCL}_{\overline{X}} = 1.5056 \pm 2.81\frac{(0.1316/0.94)}{\sqrt{5}} = 1.6815/1.3297 \tag{23}$$

and

$$\mathrm{UCL}_S/\mathrm{LCL}_S = 0.1316 \pm 2.81\sqrt{1-0.94^2}(0.1316/0.94) = 0.2658/0, \tag{24}$$

respectively.

From Fig. 1(a) and (b), it is shown that the Phase-I samples are in-control based on the control limits from Eqs. (23) and (24). In light of this, the $\widehat{\mu} = \overline{\overline{X}} = 1.5056$ and $\widehat{\sigma} = 0.1391$ are computed using Eqs. (8) and (9), respectively. The optimal charting parameters of the SSGR chart with estimated process parameters for $m = 25$, $n = 5$, $\delta_{\min} = 0.2$ and $\delta_{\max} = 1.0$ are computed using the ScicosLab programme as $(K', L') = (2.2122, 23)$. By using Eq. (10), the UCL and LCL of the Shewhart sub-chart for the SSGR chart with estimated process parameters are:

$$\widehat{\mathrm{UCL}}/\widehat{\mathrm{LCL}} = \widehat{\mu}_0 \pm K'\frac{\widehat{\sigma}_0}{\sqrt{n}} = 1.5056 \pm 2.2122\frac{0.1391}{\sqrt{5}} = 1.6432/1.3680. \tag{25}$$

In Phase-II hard-bake process monitoring, 20 additional samples with size $n = 5$ wafers are taken. The Shewhart sub-chart with the control limits from Eq. (25) and the corresponding CRL sub-chart for Phase-II SSGR chart with estimated process parameters are illustrated in Figs. 2 and 3, respectively. From the Shewhart sub-chart, the first non-conforming sample occurs at sample 14. Hence, $\mathrm{CRL}_1 = 14$ is obtained. $\mathrm{CRL}_1 \leq L' = 23$, the first out-of-control signal, is detected at sample 14. Immediate actions are performed to investigate and remove the assignable cause(s). Other non-conforming samples occur at samples 16, 18 and 20 so that

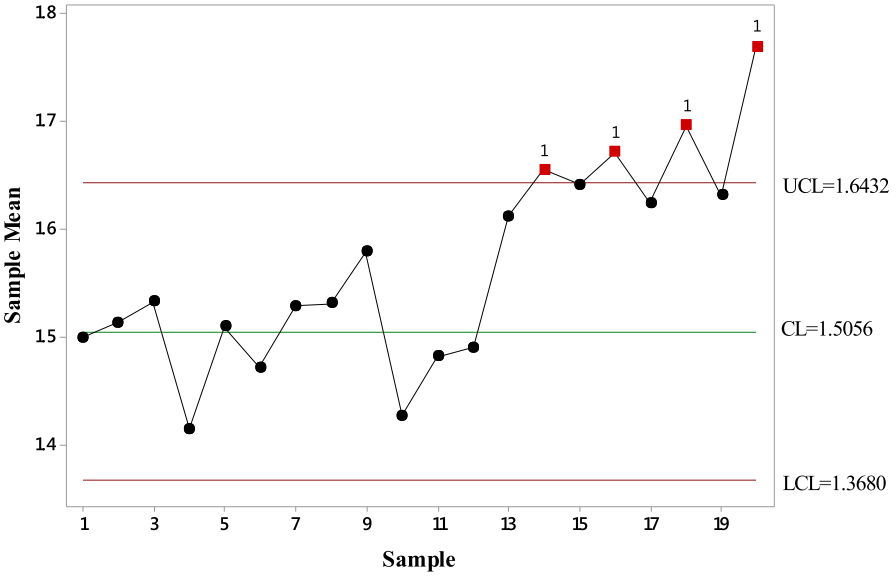


Fig. 2. The Shewhart sub-chart.

https://doi.org/10.1016/j.heliyon.2018.e00848

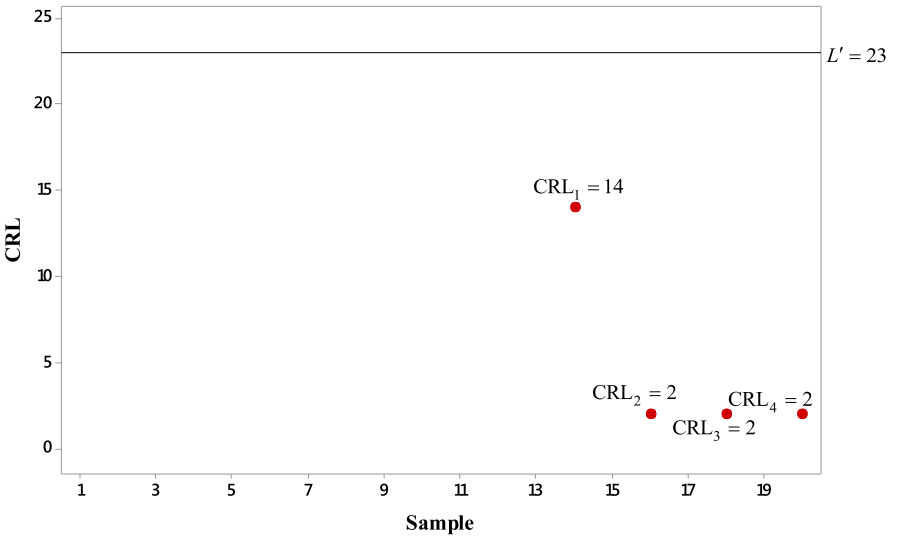


Fig. 3. The CRL sub-chart.

$CRL_2 = CRL_3 = CRL_4 = 2$. Therefore, out-of-control signal is also declared at samples 18 and 20, as $CRL_r \leq L'$ and $CRL_{r+1} \leq L'$, for $r = 2$ and 3.

7. Conclusion

In reality, true process parameters are rarely known and are estimated from in-control Phase-I samples. In this paper, it was demonstrated that the performance of the estimated process parameters SSGR chart is significantly different compared to the case when process parameters are known, particularly when n and δ are small, unless the in-control Phase-I samples are large enough. Since accumulating a large number of in-control Phase-I samples is not feasible and costly, optimal design of the estimated process parameters SSGR chart by minimising the EARL is developed for unknown shift sizes. The optimal design based on EARL is capable of tackling the situation with random shift size when the process parameters are unknown. The optimal charting parameters provided in Table 2 would appeal to practitioners whose interest is to implement the estimated process parameters SSGR chart without taking a large number of in-control Phase-I samples. Moreover, the proposed optimal charting parameters based on minimising $EARL_1$ can be employed as long as $\delta \in (\delta_{\min}, \delta_{\max})$. In addition, the SDARL measure has been considered to quantify the variability between practitioners in the performance of the SSGR chart with estimated process parameters. The application of the SSGR chart with estimated process parameters using actual industrial data shows easy implementation. A future research area to be considered is the optimal design of the group run type charts based on median run length and expected median run length under known and estimated process parameters cases.

https://doi.org/10.1016/j.heliyon.2018.e00848

Declarations

Author contribution statement

Huay Woon You: Conceived and designed the analysis; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.

Funding statement

This work was supported by the Universiti Kebangsaan Malaysia, Geran Galakan Penyelidik Muda (GGPM-2017-062).

Competing interest statement

The authors declare no conflict of interest.

Additional information

The ScicosLab programs to compute the average run length and expected average run length of the side sensitive group runs chart can be requested from the author.

Acknowledgements

The author wishes to express sincere appreciation to the Editor and anonymous reviewers, for providing useful comments and suggestions.

References

[1] M.P. Gadre, R.N. Rattihalli, A group runs control chart for detecting shifts in the process mean, Econ. Qual. Contr. 19 (1) (2004) 29−43.

[2] M.P. Gadre, R.N. Rattihalli, A side sensitive group runs control chart for detecting shifts in the process mean, Stat. Methods Appl. 16 (1) (2007) 27−37.

[3] N.A. Saleh, M.A. Mahmoud, L.A. Jones-Farmer, I. Zwetsloot, W.H. Woodall, Another look at the EWMA control chart with estimated parameters, J. Qual. Technol. 47 (4) (2015) 363−382.

[4] C.P. Quesenberry, The effect of sample size on estimated limits for $\overline{X}$ and X control charts, J. Qual. Technol. 25 (4) (1993) 237−247.

[5] L.A. Jones, C.W. Champ, S.E. Rigdon, The performance of exponentially weighted moving average charts with estimated parameters, Technometrics 43 (2) (2001) 156−167.

[6] S. Psarakis, A.K. Vyniou, P. Castagliola, Some recent developments on the effects of parameter estimation on control charts, Qual. Reliab. Eng. Int. 30 (8) (2014) 1113−1129.

https://doi.org/10.1016/j.heliyon.2018.e00848

[7] L.A. Jones-Farmer, W.H. Woodall, S.H. Steiner, C.W. Champ, An overview of phase I analysis for process improvement and monitoring, J. Qual. Technol. 46 (3) (2014) 265−280.

[8] S. Chakraborti, Run length distribution and percentiles: the Shewhart $\overline{X}$ chart with unknown parameters, Qual. Eng. 19 (2) (2007) 119−127.

[9] G. Celano, Robust design of adaptive control charts for manual manufacturing/inspection workstations, J. Appl. Stat. 36 (2) (2009) 181−203.

[10] P. Castagliola, G. Celano, S. Psarakis, Monitoring the coefficient of variation using EWMA charts, J. Qual. Technol. 43 (3) (2011) 249−265.

[11] W.C. Yeong, M.B.C. Khoo, W.L. Teoh, P. Castagliola, A control chart for the multivariate coefficient of variation, Qual. Reliab. Eng. Int. 32 (3) (2016) 1213−1225.

[12] H.W. You, M.B.C. Khoo, P. Castagliola, Y. Ou, Side sensitive group runs $\overline{X}$ chart with estimated process parameters, Comput. Stat. 30 (4) (2015) 1245−1278.

[13] A.A. Aly, M.A. Mahmoud, W.H. Woodall, A comparison of the performance of phase II simple linear profile control charts when parameters are estimated, Commun. Stat. Simulat. Comput. 44 (6) (2015) 1432−1440.

[14] N.A. Saleh, M.A. Mahmoud, M.J. Keefe, W.H. Woodall, The difficulty in designing Shewhart and X control charts with estimated parameters, J. Qual. Technol. 47 (2) (2015) 127−138.

[15] S. Knoth, The case against use of synthetic control charts, J. Qual. Technol. 48 (2) (2016) 178−195.

[16] A. Tang, P. Castagliola, J. Sun, X. Hu, The effect of measurement errors on the adaptive EWMA $\overline{X}$ chart, Qual. Reliab. Eng. Int. 34 (4) (2018) 609−630.

[17] Y. Zhang, P. Castagliola, Z. Wu, M.B.C. Khoo, The synthetic $\overline{X}$ chart with estimated parameters, IIE Trans. 43 (9) (2011) 676−687.

[18] M. Zhang, F.M. Megahed, W.H. Woodall, Exponential CUSUM charts with estimated control limits, Qual. Reliab. Eng. Int. 30 (2) (2014) 275−286.

[19] D.C. Montgomery, Statistical Quality Control: a Modern Introduction, sixth ed., Wiley, New York, 2009.

https://doi.org/10.1016/j.heliyon.2018.e00848