Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Robustification of CUSUM control structure for monitoring location shift of skewed distributions based on modified one-step M-estimator

Ayu Abdul Rahman, Sharipah Soaad Syed Yahaya & Abdu Mohammed Ali Atta

To cite this article: Ayu Abdul Rahman, Sharipah Soaad Syed Yahaya & Abdu Mohammed Ali Atta (2019): Robustification of CUSUM control structure for monitoring location shift of skewed distributions based on modified one-step M-estimator, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2018.1532001

To link to this article: https://doi.org/10.1080/03610918.2018.1532001

Published online: 21 Jan 2019.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2018.1532001

Check for updates

Robustification of CUSUM control structure for monitoring location shift of skewed distributions based on modified one-step M-estimator

Ayu Abdul Rahman[a], Sharipah Soaad Syed Yahaya[a], and Abdu Mohammed Ali Atta[b]

[a]College of Arts and Sciences, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia; [b]Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha, Qatar

ABSTRACT

Including three existing charts, a new approach employing a modified one-step M-estimator (MOM) with Cumulative Sum (CUSUM) control structure were evaluated and compared for their Phase II performances based on the average run length (ARL) under various skewed distributions. The primary focus was on the robustness of the CUSUM charts in two separate cases: (i) when the process parameters are known and (ii) when the process mean is unknown and estimated from an in-control Phase I sample. The simulation and real data analysis showed the proposed technique is comparable or sometimes better than the existing charts.

ARTICLE HISTORY
Received 23 October 2017
Accepted 27 September 2018

KEYWORDS
CUSUM chart; Average run length; Skewed distributions; Process location; MOM

1. Introduction

Two common aspects help to drive the application of utilizing a control chart as a statistical tool for monitoring process stability; economic budget and statistical power. Accordingly, the tool is hugely popular in many areas of quality control such as in the manufacturing and services industries. Moreover, it is a visual statistical tool typically structured based on three horizontal lines; the Upper Control Limit (UCL), Center Line (CL) and the Lower Control Limit (LCL), plotting the quality characteristics of a process. While the description of the control chart appears to be reasonably straightforward, it is difficult to challenge when the presence of special causes need to be highlighted. By utilizing a control chart, the existence of special causes can be detected when a point or more is observed to be outside the control limits, which indicates statistically, an out-of-control process. In this state or condition, the process is considered to be unstable as it may potentially operate with problems. Therefore, a thorough investigation along with corrective action is required to identify the cause of such problems. This is merely an indication of what can be accomplished using a control chart. On the other hand, it can provide an estimation of process capability and be extremely useful in providing diagnostic information based on the graphic pattern of the plots, if any.

CONTACT Ayu Abdul Rahman 1ayuabdul27@gmail.com College of Arts and Sciences, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia.

In an ideal state, a process is operating under statistical control. Graphically, this is viewed as all points contained within the UCL and LCL, without exhibiting a systematic or nonrandom pattern. However, an erroneous conclusion could also be presented in declaring either the statistical state or condition of the process (in-control or out-of-control). In this case, the likelihood of making an incorrect decision is determined by Type I and Type errors. Associated with the false alarm rate, the Type I error can be described as the probability of claiming that the process is out-of-control in the ideal state. On the contrary, the Type II error occurs if the process is claimed to be in-control in the presence of special causes. A set up for the UCL and LCL of a control chart fundamentally relies on these two likelihoods in order to attain a low probability of occurrence. Therefore, this constitutes a good control chart. However, the prospect of attaining a small value of a Type I error means that the chances of a Type II error occurring may increase. The statistical design charting procedure usually involves a series of steps and setting the false alarm rate with control limits to minimize the risk of a Type II error under normality. Furthermore, linking with the likelihoods mentioned above is also an alternative measurement for control chart performance, which is identified as the average run length (ARL). Notably, it is the expected value of the run length (RL) of a control chart and is the most frequently used parameter in assessing control chart performance (Woodall 2000). However, the computation of the ARL relies on normality and an outlier-free environment in Phase I and Phase II of the process. When these conditions are violated, the calculated ARL may no longer be approaching the true value. Also, deterioration in the ARL performance may further result when employing the standard design charting procedure with the estimated parameter(s) substituted for the known process parameter(s) (Jensen et al., 2006; Jones 2002; Jones, Champ, and Rigdon 2001, 2004). However, non-normality is the norm presently used in the industry as often the values of the process parameters are not readily available. Therefore, whether the usual variables charting method is appropriate in practice should be seriously considered.

Many researchers have attempted to propose a more resilient control structure to be adapted with a slight deviation from normality. With the aim to retain stable in-control RL distribution, proposals were forthcoming via heuristic approaches to design control charts (Atta, Shoraim, and Yahaya 2014; Castagliola and Khoo 2009; Khoo, Atta and Wu 2009; Khoo, Wu, and Atta 2008). Also, to construct the limits with unknown probabilistic distribution, researchers have mainly focused on non-parametric charts (Oprime et al. 2016; Riaz and Abbasi 2016; Yang and Cheng 2011; Yang, Lin, and Cheng 2011). However, the search to determine the optimal solution is ongoing to fill this void left by the existing methods. As a result, numerous recommendations on the alternative location and/or dispersion estimators with control chart structures have been proposed. Nazir et al. (2013) suggested CUSUM charts based on the median, Hodges-Lehmann (HL) and trimean to control the location parameters of a process. Furthermore, they concluded that the CUSUM control chart issued from trimean would be the preferred choice to achieve that when the observations are sampled from normal, non-normal, special cause and outlier environments. However, the estimation effects of the parameters in Phase I were not considered in the study. The effects, however, were later examined by Nazir et al. (2016) where they conducted a comparison study on the performance of three types of memory control charts. These included the CUSUM, Exponentially Weighted Moving Average (EWMA) and mixed EWMA-CUSUM control charts in the Phase II process based on median, mid-range, HL, trimean, and trimmed mean estimators under normal and mixed-normal distributions. The finding showed that none of the estimators behaved well in all environments, but the EWMA control chart offered the best performance from among the proposed methods based on the median estimator. Furthermore, it was difficult to assess the in-control robustness of the proposed charts in the study since the Phase I and Phase II distributions were dissimilar. In this case, their finding was directed towards the sensitivity of a control chart to detect a change in the Phase II process. Nazir et al. (2016) assumed that estimation is always conducted in ideal circumstances such as outlier-free, where there is no contamination or for other data anomalies which substantiated the use of the sample mean and sample standard deviation in the estimation procedure. However, Janacek and Meikle (1997) argued that ideal Phase I circumstances are rarely encountered in practical situations. Thus, the continual use of the sample mean and sample standard deviation in the estimation process may reduce the capability of a control chart in monitoring the Phase II process. Notwithstanding, to mitigate the problem, Zwetsloot, Schoonhoven, and Does (2016) recommended several robust Phase I location estimators based on the median function and the trimean, to be paired with a robust standard deviation estimator, namely a variant of the biweight. The authors also monitored the performance of the traditional EWMA control chart in the Phase II process. Nonetheless, all resulting Phase II EWMA control charts were ARL-biased, meaning that the out-of-control ARL was observed to be larger than the in-control ARL when Phase I contained contamination. To reduce the bias, especially when the data are seriously contaminated, Zwetsloot, Schoonhoven, and Does (2014) suggested several estimation approaches including the retrospective use of a control chart to obtain observations which are representative of the process. Accordingly, a two-step procedure was recommended, namely a robust estimator based on the median of sample averages $M(\overline{X})$ to estimate the location in Phase I control charting, and an efficient estimator based on the grand sample average $\overline{\overline{X}}$ for post-screening estimation. The finding showed an improvement regarding the bias for a large sample size ($n = 10$) via this approach. Unfortunately, however, for a smaller sample size ($n = 5$), the problem persisted, that is, all EWMA control charts were still ARL-biased. Similarly, significant work has been carried out in the direction of robust EWMA control charting; for example, see Zwetsloot, Schoonhoven, and Does (2015) and Khoo and Sim (2005).

Based on the literature mentioned above, the continual use of robust statistics in control charting is worth pursuing. However, a delicate balance between stable in-control performance and the quick detection of out-of-control status is still considered to be challenging even within the realm of robust estimation. The use of robust statistics to control the location parameter of control charting was claimed to offer improved if not better protection against outliers or non-normally distributed process data concerning false alarm rates. Indeed, one of the acclaimed robust statistics frequently employed using this approach has been the median estimator. The charting structure of a usual sample median, however, has drawbacks which is mainly because the median control charts are more outlier-resistant than the mean charts, as they yield less efficiency than the latter (Ahmad et al. 2014; Sheu and Yang 2006; Yang, Pai, and Wang 2010). An extension of this approach based on memory-type control charts was undertaken by Nazir et al. (2013, 2016). Although, as mentioned earlier, their approaches require further discussion and examination. The current study extends the existing literature in the form of the following contribution by suggesting a highly robust location estimator known as a modified one-step M-estimator (MOM) to be applied to the CUSUM charting structure. The aim in this case is to provide a robust structure of control charting under slight and severe non-normality with minimal loss in sensitivity to the actual location shift. Moreover, if this can be achieved, the proposed chart will be a good substitute for a median chart when small location shifts are of interest under the limitation to fulfilling the normality assumption.

As a median trimmed based estimator, MOM comprises good qualities that are usually found in the median and trimmed means. First, the MOM estimator possesses the best possible breakdown point (the proportion of extreme values that the estimator can tolerate without completely breaking down), which is 50%. Indeed, this has been made possible via the outlier detection rule established in MOM. The rule is constructed based on the usual median and median absolute deviation (MAD) estimators, thus resulting in the highest possible value of the breakdown point. Second, MOM employs a trimming approach to handle non-normality. This approach, as employed in any trimmed mean, allows dealing directly with extreme values, or more specifically the outliers. The trimming approach in MOM, however, takes into consideration the distributional shape of the data which resolves a practical concern that deals with symmetric trimming when using methods based on trimmed means. Additionally, the flexibility of the amount of trimming to be used in MOM eliminates the stipulation of a predetermined amount of trimming in the usual trimmed means.

In this study, the primary goal is to explore the effect of MOM on CUSUM chart performance under normal and non-normal distributions and also, to reflect on the improvement that could be offered by this new chart over existing charts. This is narrowed down to three existing CUSUM charts based on the mean, mid-range and median estimators which identifies a total of four CUSUM control charts, including the proposed MOM method in the study. Next, the effects of non-normality on the performance of these CUSUM charts is addressed for skewed distributions with a collection of different skewness coefficients on two different cases. For Case 1, the in-control mean, and standard deviation of the process are assumed to be known. For Case 2, the mean is unknown and estimated from an in-control Phase I sample. In the latter case, the in-control standard deviation is treated and assumed to be known to isolate the effect of estimating the location parameter.

The article is structured into the following sections. The properties and formulas for the location estimators employed in this study are given in Sec. 2, following a brief explanation of the design structure of the CUSUM control chart. Section 3 presents the set-up and simulation outcomes for the two case studies followed by Sec. 4 describing the results of the simulation. In Sec. 5, an illustrative example based on a real data set is presented, followed by the conclusion of the study in Sec. 6.

2. A two sided-CUSUM control structure

In this section, a basic outline on the CUSUM control structure is reported, along with the Phase I and Phase II location estimators that have been applied in the present study. Mathematical notation for the probability density of the chosen distributions (with their means and variances) are included for clarity.

Page (1954) proposed the idea to measure the accumulative sum of deviation of data from the in-control process location in two different plotting statistics; the upper ($C_{U,j}$) and the lower ($C_{L,j}$) part. The idea in this case, is to simultaneously monitor the upward and downward shifts of the process based on the charting statistics as given in Eq. (1).

$$C_{U,j} = \max\left\{0, C_{U,j-1} + \left(Z_{U,j} - k_U\right)\right\} \quad (1)$$
$$C_{L,j} = \min\left\{0, C_{L,j-1} + \left(Z_{L,j} + k_L\right)\right\}$$

where j is the sample number, $C_{U,0}$ *and* $C_{L,0}$ are the initial values; typically set at 0. The standardized statistics ($Z_{U,j},\ Z_{L,j}$) and the reference values (k_U, k_L) are defined as Eqs. (2) and (3), respectively.

$$Z_{U,j} = Z_{L,j} = \frac{\widehat{\theta}_j - \theta_0}{\sigma_0 / \sqrt{n}} \quad (2)$$

$$k_U = k_L = \frac{\delta_{opt}}{2} \quad (3)$$

where $\widehat{\theta}$ is the location estimator used to monitor a shift from the assumed in-control process parameter, θ_0, σ_0 is the standard deviation of $\widehat{\theta}$, δ_{opt} is the standardized shift in the location where fast detection is required, and n is the sample size. An out-of-control signal is detected when $C_{U,j} > h$ or $C_{L,j} < -h$, where h is the decision limit.

When θ_0 is unknown and requires to be estimated from the Phase I samples, formulas for the charting statistics and the reference values remain the same. But, the standardized statistics are now defined as given in Eq. (4).

$$Z_{U,j} = Z_{L,j} = \frac{\widehat{\theta}_j - \widehat{\theta}_0}{\sigma_0 / \sqrt{n}} \quad (4)$$

Previous mathematical notations involve θ_0 and σ_0. Consider, $X_{ij},\ i = 1, 2, ..., n$ and $j = 1, 2, ..., m$ denote Phase I data, when the process is in-control and let $Y_{ij},\ i = 1, 2, ..., n$ and, $j = 1, 2, ...,$ denote Phase II data. Both X_{ij} and Y_{ij} are assumed to be independent and identically distributed from the same target distribution $\mathcal{F}$ with mean θ_0 and variance ${\sigma_0}^2$. Here, $\mathcal{F}$ may result from one of these four distributions; a standard normal distribution, a Weibull distribution, a lognormal distribution, and a gamma distribution. Further, their probability density function (pdf's) as well as an expression for their means and variances are given in Table 1. While the formulas can be used to attain θ_0 and σ_0 for the chosen distribution, the application is limited to the mean estimator and indeed, is mathematically incorrect for MOM, median and mid-range. Because of this, the values (θ_0 and ${\sigma_0}^2$) are simulated for all estimators for further use in the charting procedure. Meanwhile, if the location parameter θ_0 is unknown, $\widehat{\theta}_0$ from the mean of the selective location parameter is found and calculated as given in Eq. (5).

Table 1. Distributions used in the study.

Distribution	pdf, mean and variance
Standard normal	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} x \quad x \in \mathbb{R}; \theta_0 = 0, \ \sigma_0{}^2 = 1$
Weibull with parameters $(\lambda, k) = (1{,}0)$, (1,0.5), (1,1), (1,1.5), (1,2), (1,3)	$f(x) = \begin{cases} [(k/\lambda)(x/\lambda)]^{k-1} e{-}(x/\lambda)^k & x \geq 0 \\ 0x{<}0 \end{cases};$ $\theta_0 = \lambda\ \Gamma(1 + 1/k),\ \sigma_0{}^2 = \ \lambda\ ^2\Gamma(1 + 2/k) - \ (\Gamma(1 + 1/k))^2$
Lognormal with parameters $(u,\ \sigma) = (0{,}0)$, (0,0.5), (0,1), (0,1.5), (0,2), (0,3)	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(lnx-u)^2}{2\sigma^2}} x\ ;\ \theta_0 = e^{\left(\mu + \frac{\sigma^2}{2}\right)},\ \ \sigma_0{}^2 = e^{(\sigma^2 - 1)} e^{\left(2\mu + \sigma^2\right)}$
Gamma with parameters $(k, \theta) = (0{,}1)$, (0.5,1), (1,1), (1.5,1), (2,1), (3,1)	$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-x/\theta} \quad x \in \mathbb{R}; \theta_0 = k\theta,\ \ \sigma_0{}^2 = k\theta^2$

$$\widehat{\theta}_0 = \sum_{j=1}^{m} \widehat{\theta}_j / m \tag{5}$$

While the sample mean $\overline{Y} = \frac{\sum_i^n Y_i}{n}$ is used to replace $\widehat{\theta}$ in Eqs. (1)–(5) in Phase II, it has a breakdown point (BP) of 0, which means that $\overline{Y}$ is highly receptive to the outliers. Despite the failing, $\overline{Y}$ possesses the smallest standard error without comparison under normality, making it the most efficient estimator in this case. To achieve the same positive impact, other researchers have also considered the sample mid-range $\mathrm{MR} = \frac{Y_{(1)} + Y_{(n)}}{2}$, where $Y_{(1)}$ and $Y_{(n)}$ are the minimum and maximum order of statistics, respectively, for $\widehat{\theta}$. While its dominant efficiency is noted in platykurtic distribution, the design structure of MR, nonetheless, takes on the extreme values. Thus, it suffers from a breakdown point of 0. In contrast, an outlier-resistant statistic can offer a more substantial value of *BP*. In this case, the sample median $\tilde{Y}$ has the best possible BP, which is 50% (Rousseeuw and Croux 1993). Therefore, even in extreme situations when close to half of the data points are replaced by arbitrary numbers, $\tilde{Y}$ remains bounded. Indeed, its supreme efficiency is also in motion when the distribution moves far from normal. A study by Figueiredo and Gomes (2004) revealed a good measure of $\tilde{Y}$ efficiency when the tailweight and skewness coefficients of the symmetric and asymmetric distributions, respectively, increase. More importantly, in this case, the measurement supersedes the efficiency of $\overline{Y}$. The computation of the sample median $\tilde{Y}$ involves the separation between the upper and the lower half of the ordered data. Specifically, $\tilde{Y}$ is defined as the average of the two middle order statistics, $\frac{1}{2}\left[Y_{\left(\frac{n}{2}-1\right)} + Y_{\left(\frac{n}{2}+1\right)}\right]$ for even sample sizes or simply, the value of the mid order statistic, $Y_{\left(\frac{n}{2}+\frac{1}{2}\right)}$ for the odd sample sizes. Alternatively, the trimmed based estimator could be considered for $\widehat{\theta}$ if one desires to have better statistical efficiency than that offered by the sample median under normal or light-tailed distribution. Further, it can be inferred from the Oẓdemir (2010) simulation study that estimators based on trimming clearly offer a substantial improvement over the efficiency of $\overline{Y}$, if outliers are present. The list includes the modified one step M-estimator (MOM). As suggested by Wilcox and Keselman (2003b), the MOM is defined as $\widehat{Y} = \frac{\sum_{i=i_1+1}^{n-i_2} Y_{(i)}}{n-i_1-i_2}$, where $Y_{(i)}$ is the ith ordered observation. The calculation of $\widehat{Y}$ involves the outlier detection rule that is embedded inside i_1 and i_2. In the trimming process, the rule is used to flag outliers and is constructed based on two outlier-resistant statistics; the sample median $\tilde{Y}$ and the median absolute deviation about the median ($\mathrm{MAD_n}$). Accordingly, both have BP of 0.5, which makes the outlier detection rule in $\widehat{Y}$ extremely robust, despite the simplicity of the formula. Therefore, employing this rule, an outlier is detected if $(Y_i - \tilde{Y}) < -K(\mathrm{MAD_n})$ or $(Y_i - \tilde{Y}) > K(\mathrm{MAD_n})$ and the sum of all observations in Y_i that satisfies the lower constraint is noted as i_1. Meanwhile, for the upper constraint, it is denoted as i_2 and constant K is adjusted to 2.24 for a reasonably good small standard error under normality (Wilcox and Keselman 2003b).

3. Performance evaluation and simulation procedure

Some studies have shown that robust statistics usually possess a smaller standard error than the sample mean in non-normal distributions (see, e.g. Wilcox (1998) and Wilcox and Keselman (2003a). However, it remains to be revealed that this will translate into improved CUSUM control chart performance, specifically under the selective distributions from Table 1. To demonstrate this, the run length (RL) behavior of CUSUM control charting is compared. The four control charts examined include; CUSUM-MOM, CUSUM-mean, CUSUM-MR, and CUSUM-median. The remainder of this paper will refer to each of these charts by their respective location estimator; MOM, mean, MR, and median.

Two frameworks were used in performing the simulation studies: i) when the mean and variance of the underlying distribution are known and ii) when the mean is unknown and therefore, requires to be estimated from the in-control Phase I samples. The performance of MOM, mean, MR, and median in both studies were evaluated and compared using the expected value of the RL distribution. It is also important to recall that a signal is triggered when a charting statistic plots outside the decision limit. Thus, RL is a random variable that represents the number of plotted CUSUM statistics before a signal is observed. The average run length (ARL) gives the expected value of this random variable. The rule of thumb in this instance is to have a large ARL when the process is in-control and a small ARL when the process is out-of-control. The in-control ARL will be referred to as ARL_0 and the out-of-control ARL by ARL_1 henceforth.

The design of the Phase II CUSUM control charts require the values of k_U, k_L, and h. Since $k_U = k_L$ (refer to Eq. 3) these will be simply denoted as, k. The nominal ARL_0 is set to 500, which is a choice widely adopted. Different estimators are employed for Phase II plotting statistics for $k = 0.5$ as an optimal constant in order to detect a shift of magnitude 1.0, that is a shift of $1\sigma_0$. Next, the factors h, are simulated for MOM, mean, MR, and median by considering two sample sizes, $n = 5$ and $n = 10$ from a standard normal distribution with the desired value of ARL_0. This gives the factors h in Case 1 as illustrated in Table 2.

In Case 2, the Phase I process location parameter θ_0 needs to be estimated. To obtain the factors h, $m = 50$ subgroups of a sample size $n = 10$ from a standard normal distribution are considered. The values of these factors are provided in Table 3.

Table 2. Factors of the CUSUM charts for Case 1 under standard normal distribution at ARL_0.

n	MOM	Mean	MR	Median
5	5.0949	5.0717	5.086	5.0625
10	5.1075	5.0733	5.118	5.0569

The simulation procedure regarding the robustness and sensitivity of the CUSUM chart in Case 1 and Case 2 are discussed separately in the following subsections. Both outcomes were attained using SAS software.

3.1. Case 1: θ_0 and σ_0 are known

The RL distribution of the CUSUM control chart was obtained via two series of Monte Carlo simulations. The first series was used to determine the mean value and the standard error for all estimators. Using 1,000,000 samples of size $n = 5$ and 10 from the chosen distribution (Table 1), the mean and standard deviation (i.e., standard error) of the sampling distribution of MOM, mean, mid-range and median estimators were obtained. In the second series, 15,000 of Phase II samples of size n were used. The observations were generated from the chosen distribution and the charting statistics, and $C_{U,i}$ and $C_{L,i}$ were calculated according to Eq. (1), with $C_{U,0} = C_{L,0} = 0$. If $C_{U,j} < h$ or $C_{L,j} > -h$, the run length counter was incremented. The steps were repeated until the record showed that either $C_{U,j} > h$ or $C_{L,j} < -h$. When this occurred, a signal was given, and the corresponding run length equals j. The calculations were attained for different shifts of size $\delta\sigma_0$ in the mean, considering $\delta =$ 0, 0.1, 0.15, 0.2, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3. The process was repeated for 10,000 simulation runs which provided a total of 10,000 RLs. The ARL was computed by averaging over the total 10,000 RLs. The results are presented in Tables 4 and 5.

3.2. Case 2: θ_0 are unknown

Because the estimate is substituted for the unknown process parameter θ_0, it is useful to examine the impact of estimation on the Phase II RL distribution and hence, the robustness of the proposed chart. To achieve the target outcome, Phase I and Phase II distributions are assumed to be the same. Later, a shift in the process of the same distribution was introduced to model the out-of-control situation, wherein the outcome is then used to evaluate the sensitivity of the CUSUM control charts.

Notably as a reminder, $\widehat{\theta}_0$ was computed using in-control Phase I samples based on four location estimators; MOM, mean, mid-range and median. For a fair comparison of the effect imparted by these location estimators, the in-control standard deviation σ_0 was treated as known. Thus, the observed performance of the charts is purely based on the effect of the location parameter. Also, it is crucial to note the distinction between the conditional and unconditional RL distribution for control chart performance when the process parameter(s) is estimated. Jones, Champ and Rigdon (2004) provided some interesting thoughts regarding this particular issue. The conditional RL distribution is attained based on a specific value of process parameters. In the case, it would be $\widehat{\theta}_0$, in which the estimate depends on the set-up of the in-control Phase I data. Thus, this conditional distribution of RL only gives the conditional performance of the CUSUM

Table 3. Factors of the CUSUM charts for Case 2 under standard normal distribution at $ARL_0 = 500$.

n	MOM	Mean	MR	Median
10	5.4853	5.4484	5.4934	5.444

control chart. In order to obtain the overall performance, the unconditional RL distribution can be used, also considering this unconditional distribution as the average of the RL distribution over all possible values of $\widehat{\theta}_0$, which gives the estimate of the unconditional ARL.

To attain the ARL, the following simulation procedure was applied. By first generating $m = 50$ samples of size $n = 10$, this gives a total of 500 observations. These were the in-control Phase I dataset, drawn from one of the distributions in Table 1, used to compute $\widehat{\theta}_0$. In Phase II, a total of 15,000 samples of size n were generated from the same distribution as in Phase I. From each sample, $\widehat{\theta}_j$ was computed thereby applying the value on the CUSUM charting statistics; $C_{U,j}$ and $C_{L,j}$ with $C_{U,0} = C_{L,0} = 0$. The run length counter was then recorded. If $C_{U,j} < h$ or $C_{L,j} > -h$, the counter was increased. The steps were repeated until the record showed that either $C_{U,j} > h$ or $C_{L,j} < -h$, which triggers the signal. The corresponding run length equals j. As in Case 1, the calculations were also made for $\delta = 0, 0.1, 0.15, 0.2, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3$; giving the in-control and the out-of-control RLs. Next, to obtain the ARL, the entire procedures were repeated 10,000 times. The value of ARL was computed by averaging the RLs over all 10,000 repetitions. The outcomes are presented in Tables 5 and 6.

4. Simulation results

The outcomes of the simulation studies in Case I and Case 2 are discussed separately. Firstly, taking a closer look at the in-control and out-of-control RL distributions of the CUSUM charts in Case I is undertaken, followed by discussing the result in detail for Case 2.

4.1. Results of case I: θ_0 and σ_0 are known

The simulation procedure described in Sec. 3.1 give the findings as summarized in Tables 4 and 5. All three skewed distributions in Table 1 are considered when the process is in-control, i.e. $\delta = 0$. Ultimately, there are four CUSUM control charts for three skewed distributions that are compared based on the skewness coefficient of $\alpha_3 = \{0, \ 0.5, 1.0, 1.5, 2, 3\}$ and sample size $n = \ \{5, 10\}$. In an attempt to assess robustness, these CUSUM charts were designed using the indicated values shown for factor h in Table 2, which technically are only appropriate for normally distributed process data. Notably, for sample size 10, and the ARL_0 values of MOM are much larger than the mean, MR, and median, especially when $\alpha_3 = 1.5$. In this particular case, the in-control average run lengths of MOM are approximately 95% of the nominal ARL value but then declined by 14% when $n = 5$. For $n = 5$ or α_3 greater than 2, the ARL_0 values of MOM, mean, MR, and median are much shorter than 500, thus the number of false alarm rates in these cases will far exceed the number initially anticipated. However, examining the ARL_0 values under the lognormal distribution, the in-control average run

Table 4. In-control ARLs for known parameters (Case 1). In the table, β, ω and α are the shape parameters that characterize the degree of asymmetric (α_3) of a Weibull, lognormal, and gamma distributions.

	Distribution	α_3	$n=5$ MOM	$n=5$ Mean	$n=5$ MR	$n=5$ Median	$n=10$ MOM	$n=10$ Mean	$n=10$ MR	$n=10$ Median
	Normal	0.0	500.02	500.05	500.02	500.03	500.09	500.09	500.11	500.14
	Weibull									
	3.6286	**0.0**	482.49	513.61	549.79	502.10	498.96	508.41	534.39	494.19
	2.2266	**0.5**	495.42	497.69	489.87	498.77	**513.98**	499.08	488.81	492.72
β	**1.5688**	**1.0**	448.16	457.51	417.40	416.01	**514.25**	479.29	422.61	455.53
	1.2123	**1.5**	390.71	406.63	340.64	345.64	**475.89**	448.85	343.60	402.23
	0.9987	**2**	320.84	349.40	282.21	282.36	405.26	**409.49**	288.97	344.58
	0.7637	**3**	226.56	262.92	206.97	208.91	299.51	**340.07**	222.31	273.50
	Lognormal									
	0.0010	**0.0**	534.76	530.83	491.87	536.40	497.79	563.21	492.93	474.01
	0.1656	**0.5**	480.16	**513.26**	482.08	467.53	472.74	463.86	459.23	**498.67**
ω	**0.3170**	**1.0**	**445.04**	**449.63**	372.62	438.06	**494.31**	470.51	353.98	449.08
	0.4484	**1.5**	**403.81**	385.66	286.65	371.16	**468.31**	443.69	281.49	432.93
	0.5593	**2**	**359.11**	334.37	241.23	331.91	**431.33**	399.67	238.04	396.18
	0.7315	**3**	**292.71**	259.64	194.64	269.42	**376.25**	328.96	199.60	342.58
	Gamma									
	38000	**0.0**	481.981	503.754	502.165	492.014	499.971	503.523	501.900	498.161
	15.4	**0.5**	**468.545**	**483.240**	462.160	470.160	**503.553**	**493.920**	446.754	486.662
α	**3.913**	**1.0**	**443.104**	**450.724**	393.832	415.054	**498.352**	**484.285**	380.292	455.851
	1.788	**1.5**	**399.440**	**406.809**	332.593	351.262	**471.593**	**447.736**	324.226	403.221
	0.983	**2**	315.455	347.443	281.490	280.109	**399.558**	**410.945**	290.442	345.624
	0.442	**3**	196.903	268.735	227.826	189.147	248.781	**336.180**	247.261	239.261

lengths of MOM for the said case are significantly larger than its counterparts. Moreover, for the very skewed distributions ($\alpha_3 = 2$ and $\alpha_3 = 3$), the ARL_0 values of MR are much smaller than the rest of the charts. This is particularly noted when the underlying process data actually follow the lognormal distribution. Furthermore, it is also interesting to note at this point, that several ARL_0 values are slightly above 500, mostly attributed by MOM. For $n = 5$ and for α_3 1.5 or less, the in-control average run lengths of the MOM are comparable to the mean. In general, it can be seen that ARL_0 of MOM is significantly better than the median.

Table 5 presents the ARL_1 for MOM, mean, MR, and median when the underlying process data follow various Weibull distributions. The table was designed using $ARL_0 = 370$, $n = 10$ and $\delta_{opt} = 1.0$ for a shift in the process mean ranging from the smallest magnitude ($0.1\sigma_0$) to the largest magnitude ($3\sigma_0$). Also, it was found that the average run lengths for all charts using process data from the Weibull distribution with zero degree of skewness ($\alpha_3 = 0$) and the normal distribution are similar. Further, due to the space limitation, results from the Weibull were only provided, and later used the average run lengths values when $\alpha_3 = 0$ as the basis for comparison under the normal theory value. For $n = 10$, the average run lengths of MOM using skewed process data ($0 < \alpha_3 < 2$) and the non-skewed process data ($\alpha_3 = 0$) were also found to be comparable. Thus, violating the normality assumption does not dampen the ability of the MOM to detect location shifts for this range of α_3. Also, the average run lengths of various CUSUM charts are comparable when the underlying process data actually follow a normal distribution (or a Weibull with $\alpha_3 = 0$) and for the very skewed Weibull distributions ($\alpha_3 = 2$ and $\alpha_3 = 3$). For a shift in the mean of $0.1\sigma_0$ and for $\alpha_3 = 0.5$, 1 1.5, the average run lengths of MR and median are significantly smaller than the MOM and the mean. Hence, the former two charts are more efficient.

Table 5. Out-of-control ARLs for known parameters (Case 1). In the table, β is the shape parameter that characterizes the degree of asymmetric (α_3) of a Weibull.

			$n=5$				$n=10$			
β	α_3	δ	MOM	Mean	MR	Median	MOM	Mean	MR	Median
3.6286	**0.0**	**0.1**	178.17	**171.75**	180.09	176.52	99.93	98.18	98.24	**97.79**
		0.15	**86.68**	**86.41**	88.93	**86.33**	43.49	43.46	43.83	**42.73**
		0.2	49.69	49.56	50.33	49.03	24.20	23.93	24.23	**23.79**
		0.25	31.61	31.26	31.44	31.46	15.87	15.70	15.86	15.66
		0.5	8.87	8.82	8.83	8.80	5.45	5.45	5.48	5.42
		0.75	5.04	5.03	5.05	5.03	3.35	3.34	3.36	3.33
		1	3.57	3.56	3.57	3.56	2.49	2.46	2.50	2.46
		1.5	2.34	2.34	2.34	2.33	1.82	1.80	1.82	1.80
		2	1.89	1.88	1.89	1.88	1.23	1.22	1.25	1.23
		2.5	1.50	1.49	1.50	1.48	1.01	1.01	1.01	1.01
		3	1.13	1.14	1.13	1.12	1.00	1.00	1.00	1.00
2.2266	**0.5**	**0.1**	169.09	161.08	**155.44**	159.59	98.83	95.87	**93.71**	**93.70**
		0.15	85.81	83.07	**82.73**	82.35	44.48	43.04	44.36	**42.57**
		0.2	49.50	48.88	**48.18**	48.81	24.62	24.23	25.05	**24.00**
		0.25	31.76	31.17	31.34	30.96	15.98	15.86	16.18	15.79
		0.5	8.97	8.99	8.94	8.98	5.44	5.44	5.51	5.44
		0.75	5.06	5.05	5.10	5.09	3.35	3.33	3.37	3.34
		1	3.59	3.57	3.60	3.57	2.49	2.48	2.49	2.46
		1.5	2.34	2.34	2.34	2.34	1.82	1.80	1.81	1.80
		2	1.89	1.88	1.89	1.87	1.24	1.23	1.25	1.22
		2.5	1.51	1.50	1.52	1.50	1.01	1.01	1.00	1.00
		3	1.13	1.12	1.13	1.13	1.00	1.00	1.00	1.00
1.5688	**1.0**	**0.1**	151.86	150.17	**141.13**	142.28	94.90	92.70	**89.26**	90.29
		0.15	82.04	81.23	80.42	**79.83**	**43.76**	**43.15**	44.49	**43.53**
		0.2	48.37	48.53	48.41	**47.53**	**24.58**	**24.49**	25.05	**24.37**
		0.25	31.73	31.25	32.17	31.12	16.01	15.86	16.36	16.02
		0.5	9.01	8.99	9.03	8.98	5.47	5.49	5.54	5.48
		0.75	5.12	5.07	5.12	5.08	3.35	3.34	3.37	3.34
		1	3.60	3.57	3.57	3.58	2.47	2.48	2.50	2.46
		1.5	2.35	2.35	2.33	2.33	1.82	1.81	1.82	1.80
		2	1.88	1.88	1.88	1.86	1.24	1.23	1.25	1.24
		2.5	1.54	1.52	1.54	1.52	1.01	1.00	1.00	1.00
		3	1.13	1.12	1.12	1.12	1.00	1.00	1.00	1.00
1.2123	**1.5**	**0.1**	140.23	140.12	**127.89**	129.50	92.37	90.62	**86.80**	**86.74**
		0.15	81.14	79.99	**77.69**	79.08	**43.85**	**43.70**	44.87	**43.06**
		0.2	49.32	**48.35**	**48.87**	49.16	24.85	24.67	25.72	24.79
		0.25	31.89	31.96	32.21	32.72	16.24	16.00	16.85	16.32
		0.5	9.13	9.11	9.11	9.19	5.55	5.47	5.60	5.52
		0.75	5.16	5.11	5.14	5.08	3.37	3.36	3.38	3.35
		1	3.61	3.58	3.59	3.58	2.47	2.48	2.50	2.47
		1.5	2.34	2.33	2.35	2.33	1.81	1.81	1.83	1.80
		2	1.87	1.87	1.87	1.86	1.26	1.24	1.25	1.24
		2.5	1.55	1.54	1.55	1.55	1.00	1.00	1.00	1.00
		3	1.12	1.12	1.10	1.10	1.00	1.00	1.00	1.00
0.9987	**2**	**0.1**	125.63	130.01	**119.16**	**119.35**	86.88	87.23	**84.99**	**84.97**
		0.15	78.99	78.55	**76.16**	**76.10**	44.19	44.08	44.73	**42.82**
		0.2	49.10	48.73	48.97	49.38	25.36	**24.88**	26.12	**24.96**
		0.25	32.56	32.48	33.10	33.26	16.69	16.39	16.98	16.38
		0.5	9.38	9.17	9.34	9.26	5.56	5.49	5.60	5.52
		0.75	5.17	5.08	5.16	5.15	3.38	3.36	3.38	3.36
		1	3.61	3.59	3.59	3.59	2.49	2.47	2.50	2.47
		1.5	2.34	2.34	2.34	2.33	1.81	1.81	1.83	1.81
		2	1.87	1.87	1.87	1.86	1.26	1.24	1.25	1.23
		2.5	1.57	1.55	1.57	1.57	1.00	1.00	1.00	1.00
		3	1.11	1.11	1.09	1.08	1.00	1.00	1.00	1.00
0. 7637	**3**	**0.1**	109.83	116.91	**106.88**	107.62	**82.18**	83.44	**82.36**	79.64
		0.15	74.56	75.26	74.46	**73.66**	45.18	**43.68**	46.23	**43.81**
		0.2	50.60	49.14	50.52	50.15	25.82	25.37	27.93	26.08
		0.25	34.78	33.05	34.52	34.57	17.01	16.63	17.57	16.87

(*continued*)

Table 5. Continued.

β	α_3	δ	$n=5$ MOM	Mean	MR	Median	$n=10$ MOM	Mean	MR	Median
		0.5	9.60	9.41	9.49	9.48	5.61	5.57	5.63	5.53
		0.75	5.22	5.14	5.18	5.16	3.37	3.34	3.35	3.36
		1	3.61	3.57	3.59	3.59	2.49	2.48	2.50	2.46
		1.5	2.36	2.35	2.34	2.33	1.82	1.82	1.84	1.82
		2	1.88	1.87	1.88	1.88	1.25	1.24	1.24	1.22
		2.5	1.61	1.58	1.61	1.61	1.00	1.00	1.00	1.00
		3	1.02	1.07	1.04	1.02	1.00	1.00	1.00	1.00

Table 6. In-control ARLs for unknown parameters (Case 2) with $n=10$. In the table, β is the shape parameter that characterizes the degree of asymmetric (α_3) of a Weibull.

	Distribution	α_3	MOM	Mean	MR	Median
	Normal	**0.0**	500.20	500.27	499.98	500.19
	Weibull					
	3.6286	**0.0**	501.05	495.00	524.30	492.80
	2.2266	**0.5**	**510.85**	494.01	487.29	499.19
β	**1.5688**	**1.0**	**515.67**	488.27	449.14	471.07
	1.2123	**1.5**	**492.25**	469.49	376.29	435.75
	0.9987	**2**	437.79	438.66	333.27	387.85
	0.7637	**3**	344.78	**374.80**	264.36	315.07

For the very skewed Weibull distributions ($\alpha_3 = 2$ and $\alpha_3 = 3$) with $n=5$ and a shift in the mean of $0.1\sigma_0$, the average run lengths of the mean are slightly larger than its counterpart. Meanwhile. the ARL_1 values of MOM, MR, and median are comparable in this case. For a minimal location shift ($0.1\sigma_0$) and $n=5$, the ARL_1 of all the charts are significantly shorter than their corresponding normal theory values for all Weibull distributions considered. Indeed, it is interesting to note that the impact is less severe on MOM, when compared to the existing charts.

4.2. Results for case II: θ_0 is unknown

In this section, the focus is on the unconditional Phase II RL distribution based on the ARL as given in Tables 6 and 7. Based on the nominal in-control ARL of 500, the simultaneous effect of parameter estimation and non-normality on the four CUSUM control charts are examined. The ARL was obtained when both parameter estimation and the charting statistics were subjected to Weibull distribution for $\alpha_3 = \{0,\ 0.5, 1.0, 1.5, 2, 3\}$ and sample size $n=10$. First, the robustness of the CUSUM charts to a violation of the normality assumption is considered (Table 6). Thus, the MOM, mean, MR, and median were designed using the indicated values shown for factor h in Table 3. For $\alpha_3 = 0.5,\ \alpha_3 = 1$ and $\alpha_3 = 1.5$, the in-control average run lengths of MOM are significantly better than the mean, MR, and median. For this case, the in-control average run lengths of MOM are approximately 98% of the nominal ARL value. Indeed, it is interesting to note that the ARL_0 values of MOM are marginally above the nominal theory values for light non-normality. In general, ARL_0 is seen to be significantly smaller than 500 for the very skewed Weibull distributions ($\alpha_3 = 2$ and $\alpha_3 = 3$). Also, in the most extreme case ($\alpha_3 = 3$), it may result in false alarm rates that could be unacceptably high for many practical applications, especially when a CUSUM chart is constructed using the mid-range estimator.

Meanwhile, Table 7 provides the ARL_1 for MOM, mean, MR, and median when the underlying process data follow various Weibull distributions. As previous, the table was designed using $ARL_0 = 370$, $n = 10$ and $\delta_{opt} = 1.0$ for a shift in the process mean ranging from $0.1\sigma_0$ (the smallest magnitude) to $3\sigma_0$ (the largest magnitude). In this case, ARL_1 is excluded when observations are sampled from a normal distribution and subsequently set forth the ARL_1 using process data from the Weibull distribution with zero degree of skewness ($\alpha_3 = 0$) as a basis for corresponding normal theory values, since the two ARLs are very similar. In the cases where the Weibull distributions are very similar to the normal distribution, that is when $\alpha_3 = 0.5$ and $\alpha_3 = 1.0$, the average run lengths of MOM are comparable to the corresponding normal theory values. Thus, violating the normality assumption does not dampen the ability of the MOM to detect location shifts under slight non-normality. Also, the average run lengths of various CUSUM charts are comparable when the underlying process data actually follow a normal distribution (or a Weibull with $\alpha_3 = 0$). For a shift in the location of $0.1\sigma_0$ with $\alpha_3 > 0$, it is noted however, that the average run lengths of MOM are slightly larger than the mean, MR, and median. However, it is interesting to note that those three charts provide shorter ARL_1 than their corresponding normal theory values under the same set of conditions.

5. An example application

In this section, the practical application of the proposed method is demonstrated using a real data set on patient waiting time (in minutes) to undergo a colonoscopy procedure in a regional medical center and is based on the data from Jones-Farmer, Jordan, and Champ (2009). The waiting times for the procedure were measured for 30 subgroups of patients in a group of five (i.e. $n = 5$). The measure of ARL in the previous two sections recorded a comparable performance between the MOM and the mean regarding the false alarm, which was found to outweigh the performance of the median and MR further. Regarding the detection of the out-control status, both the MOM and the mean compete well with the MR and median at the shift for which each was designed, that is $\delta = 0.1$. Thus, only the MOM and mean are applied on the real data for further analysis.

The MOM and mean charts were designed using the indicated values shown for h in Table 3 for $k = 0.5$. Figure 1a displays patient waiting time for the respective subgroups. The presence of large values in subgroups 6, 15, 16, and 30 (as captured in the figure) should not be a seen as a direct cause for the CUSUM chart to signal since they do not directly represent a shift in the mean. From this perspective, the application of the mean chart will be exercised cautiously. The computation of the usual sample mean is easily perturbed by the outlying values and may indirectly influence the performance of the chart when used to detect a difference in the location shift.

Figure 1b–c display the output of the two charts, along with their upper decision limit (h) and lower decision limit ($-h$). General observations based on the two figures show an almost similar behavior between the performance of the two charts. Upon closer inspection, however, the mean detects the out-of-control points at subgroups 8,

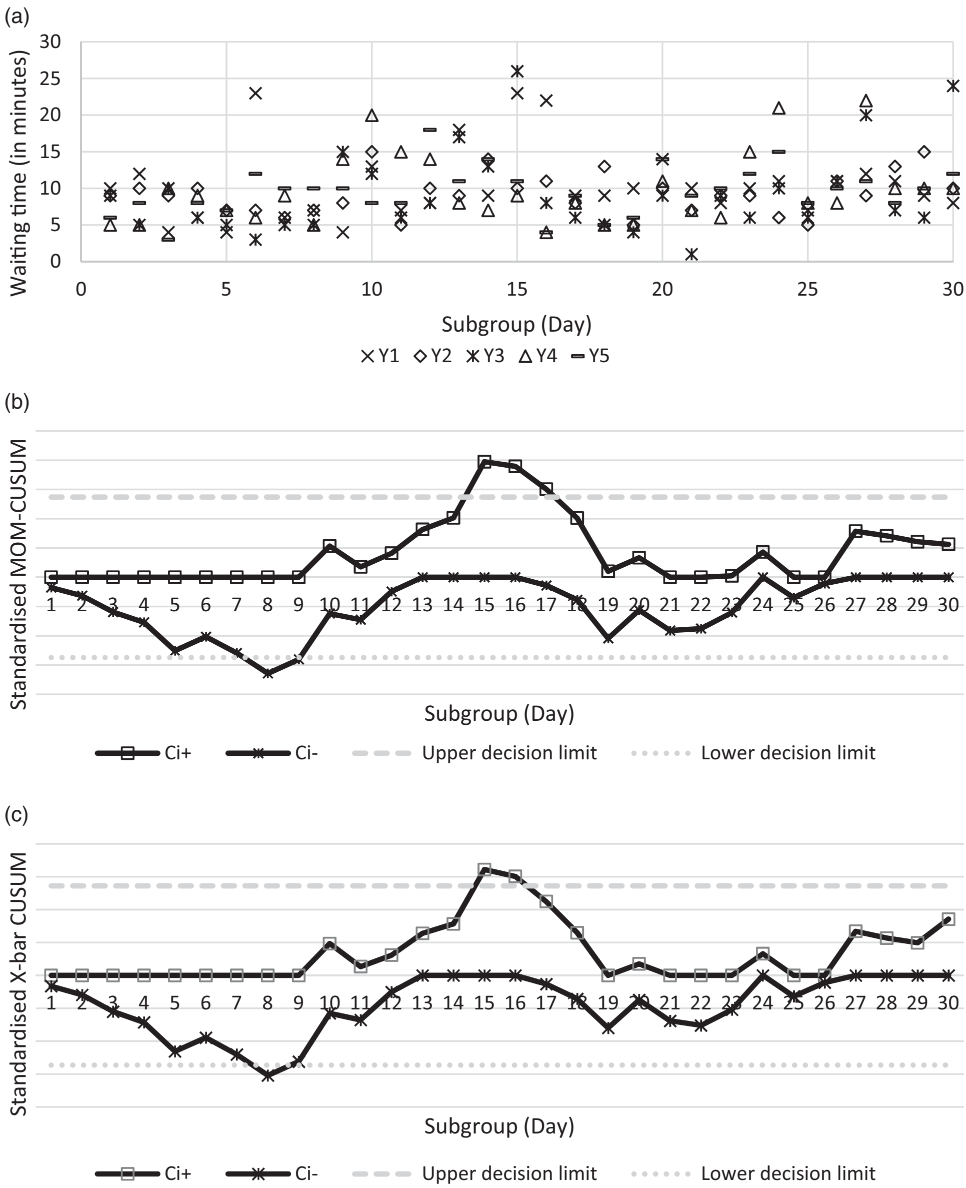

Figure 1. Real Data from Jones-Farmer, Jordan and Champ (2009) on the Colonoscopy Procedure from a Regional Medical Centre: (a) Patient Waiting Time (in minutes); (b) output of the MOM CUSUM chart; and (c) output of the mean CUSUM chart.

15, and 16 while the MOM detects two additional out-of-control points (i.e. at subgroups 9 and 17) in addition to the three points (at subgroups 8, 15, 16) signaled by the mean. From Figure 1b–c, a shift can be observed that is gradually decreasing from subgroup 1 until subgroup 8, followed by a subsequent increasing shift until subgroup 10. Here, both charts successfully detect the downward shifts, but only MOM can detect the upward shifts. Both charts successfully capture the following downward shift (from subgroups 15 to 19).

Table 7. Out-of-control ARLs for unknown parameters (Case 2), with $n = 10$. In the table, β is the shape parameter that characterizes the degree of asymmetric (α_3) of a Weibull.

β	α_3	δ	MOM	Mean	MR	Median
3.6286	**0.0**	**0.1**	180.99	**178.25**	180.31	180.10
		0.15	70.67	69.45	70.68	**68.17**
		0.2	33.11	**32.88**	33.85	**32.80**
		0.25	19.12	19.21	19.67	18.96
		0.5	5.87	5.86	5.87	5.84
		0.75	3.57	3.57	3.58	3.54
		1	2.62	2.61	2.64	2.61
		1.5	1.91	1.91	1.91	1.90
		2	1.37	1.36	1.38	1.35
		2.5	1.03	1.03	1.02	1.03
		3	1.00	1.00	1.00	1.00
2.2266	**0.5**	**0.1**	180.12	166.24	**160.36**	168.24
		0.15	72.08	68.31	**66.41**	68.19
		0.2	33.57	**32.72**	33.01	**32.79**
		0.25	19.46	19.48	19.85	19.10
		0.5	5.86	5.84	5.97	5.86
		0.75	3.56	3.57	3.59	3.55
		1	2.62	2.62	2.64	2.62
		1.5	1.91	1.90	1.90	1.89
		2	1.38	1.36	1.40	1.36
		2.5	1.03	1.02	1.02	1.02
		3	1.00	1.00	1.00	1.00
1.5688	**1.0**	**0.1**	168.80	158.80	**146.49**	153.18
		0.15	69.69	68.88	**65.97**	65.40
		0.2	34.09	**32.63**	33.16	**32.27**
		0.25	19.43	19.75	19.83	19.69
		0.5	5.93	5.89	6.01	5.87
		0.75	3.59	3.55	3.59	3.57
		1	2.64	2.61	2.63	2.62
		1.5	1.90	1.89	1.90	1.89
		2	1.38	1.37	1.41	1.38
		2.5	1.02	1.02	1.01	1.01
		3	1.00	1.00	1.00	1.00
1.2123	**1.5**	**0.1**	154.82	151.33	**127.88**	136.59
		0.15	65.64	67.27	**63.54**	65.26
		0.2	32.95	**31.79**	33.56	33.04
		0.25	20.02	19.94	20.76	19.76
		0.5	5.94	5.95	5.99	5.97
		0.75	3.60	3.56	3.60	3.56
		1	2.63	2.61	2.65	2.62
		1.5	1.89	1.89	1.89	1.88
		2	1.40	1.39	1.41	1.39
		2.5	1.01	1.01	1.00	1.01
		3	1.00	1.00	1.00	1.00
0.9987	**2**	**0.1**	136.51	143.16	**120.73**	126.71
		0.15	63.34	65.38	**62.14**	64.72
		0.2	33.64	32.54	33.56	32.87
		0.25	20.00	19.61	21.12	20.10
		0.5	6.00	5.94	6.02	5.97
		0.75	3.58	3.58	3.59	3.58
		1	2.63	2.63	2.64	2.61
		1.5	1.89	1.89	1.89	1.89
		2	1.41	1.39	1.43	1.40
		2.5	1.00	1.01	1.00	1.00
		3	1.00	1.00	1.00	1.00

(*continued*)

Table 7. Continued.

β	α_3	δ	MOM	Mean	MR	Median
0. 315.07	**3**	**0.1**	118.18	127.46	**110.36**	116.09
		0.15	61.12	62.71	62.29	**58.97**
		0.2	34.26	33.69	35.26	34.67
		0.25	20.75	20.31	21.12	20.51
		0.5	6.03	5.99	6.09	6.00
		0.75	3.60	3.56	3.60	3.58
		1	2.64	2.63	2.63	2.62
		1.5	1.89	1.88	1.90	1.88
		2	1.43	1.41	1.44	1.41
		2.5	1.00	1.00	1.00	1.00
		3	1.00	1.00	1.00	1.00

6. Conclusion and recommendation

The routine use of control charting techniques requires practitioners to adhere to normality assumption. However, in many practical instances, the underlying distribution of the output measurements may not always be confined to symmetric and bell-shaped distributions only. For example, one tail of the distribution could be considerably longer than the other, therefore giving what can best be described as a skewed distribution. Indeed, this circumstance is frequently encountered when measurements are taken from semiconductor and chemical processes, which may lead to some erroneous conclusions when the outputs are applied on the usual control charting techniques for some desired outcomes. For that very reason, ubiquitous studies on the means to fortify the control structures are seen in the quality control literature.

The need for a robust technique has become increasingly important when information for the in-control process parameter(s) is not readily available, which thereby force reliance on the estimates in the place of the known parameter(s). The variability introduced by the estimates may also have a consequence on the charts performance that indeed differs from the charts designed with known parameter(s). Intuitively, this would drive the search for a reasonable estimate without incurring unnecessary costs in computation time. Therefore, this study presented an alternative robust design structure for a CUSUM control chart and subsequently, compared its performance with three existing CUSUM charts; the mean, mid-range and median charts under various skewed distributions.

The robustness of the CUSUM charts towards a violation in the normality assumption was also considered in this paper. Two situations were discussed, one situation in which the process parameters are known and the other situation when the mean of the process is unknown and estimated from the in-control Phase I samples. It was shown that in both cases, controlling the location parameter based on MOM will help to maintain a stable in-control performance of the CUSUM control chart when the underlying process data are from skewed distributions. Generally, for sample size $n = 10$, the chart is quite robust to violation of the normality assumption and provides higher in-control average run lengths than the existing charts. The minute variabilities between the out-of-control average run lengths from the normal and the moderately skewed distributions indicate that the ability of the CUSUM chart based on MOM to detect shifts in the location continues to endure when the underlying distributional assumption is violated. This makes the proposed method quite appealing even though the implementation of this chart is slightly more complicated than the existing charts. The estimated effect of the parameter (i.e. mean of the process) is substantial upon examining the efficiency of all charts to detect small shifts in the location ($1\sigma_0 < \delta < 0.5\sigma_0$). That is, the performance of the CUSUM charts in Case 2 are significantly slower, when compared to that in Case 1. Furthermore, the focus was primarily on Weibull distributions but there is no reason to doubt or distrust this substantive conclusion would carry over to other skewed distributions (lognormal and gamma) applied in this study.

Although, for a minimal location shift, that is $0.1\sigma_0$, the out-of-control average run lengths of MOM from the moderately skewed distributions is to some extent larger than the existing charts. Thus, if a given application demands monitoring a particularly small magnitude of the shift, it is recommended to obtain the standard error of the MOM estimate by means of the bootstrap method. This will provide a sound estimate with a possibly lower value than the usual formula for the standard error. For future research, this will be the primary focus.

Acknowledgments

The authors would like to acknowledge the work that led to this paper is supported Fundamental Research Grant Scheme (S/0 CODE 13578) of the Utara Malaysia and Ministry of Higher Education, Malaysia. The authors are grateful for the reviewer's careful reading of the manuscript and constructive comments for improvement.

References

Ahmad, S., M. Riaz, S. A. Abbasi, and Z. Lin. 2014. On efficient median control charting. *Journal of the Chinese Institute of Engineers* 37 (3):358–75.

Atta, A. M., M. H. Shoraim, and S. S. Yahaya. 2014. A multivariate EWMA control chart for skewed populations using weighted variance method. *International Research Journal of Science & Engineering* 2 (6):191–202.

Castagliola, P., and M. B. C. Khoo. 2009. A synthetic scaled weighted variance control chart for monitoring the process mean of skewed populations. *Communications in Statistics: Simulation & Computation* 38 (8):1659–74.

Figueiredo, F., and M. I. Gomes. 2004. The total median in statistical quality control. *Applied Stochastic Models in Business and Industry* 20 (4):339–53.

Janacek, G. J., and S. E. Meikle. 1997. Control charts based on medians. *The Statistician* 46 (1): 19–31.

Jensen, W. A., A. Jones-Farmer, C. W. Champ, and W. H. Woodall. 2006. Effects of parameter estimation on control chart properties: A literature review. *Journal of Quality Technology* 38 (4):349–64.

Jones-Farmer, L. A., V. Jordan, and C. W. Champ. 2009. Distribution-Free Phase I control charts for subgroup location. *Journal of Quality Technology* 41 (3):304–16.

Jones, L. A. 2002. The statistical design of EWMA control charts with estimated parameters. *Journal of Quality Technology* 34 (3):277–88.

Jones, L. A., C. W. Champ, and S. E. Rigdon. 2001. The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics* 43 (2):156–67.

Jones, L. A., C. W. Champ, and S. E. Rigdon. 2004. The run length distribution of the CUSUM with estimated parameters. *Journal of Quality Technology* 36 (1):95–108.

Khoo, M. B. C., A. M. A. Atta, and Z. Wu. 2009. A multivariate synthetic control chart for monitoring the process mean vector of skewed populations using weighted standard deviations. *Communications in Statistics: Simulation & Computation* 38 (7):1493–518.
Khoo, M. B. C., and S. Y. Sim. 2005. A robust exponentially weighted moving average control chart for the process mean. *Journal of Modern Applied Statistical Methods* 5 (2):464–74.
Khoo, M. B. C., Z. Wu, and A. M. A. Atta. 2008. A synthetic control chart for monitoring the process mean of skewed populations based on the weighted variance method. *International Journal of Reliability, Quality and Safety Engineering* 15 (03):217–45.
Nazir, H. Z., N. Abbas, M. Riaz, and R. J. M. M. Does. 2016. A comparative study of memory-type control charts under normal and contaminated normal environments. *Quality and Reliability Engineering International* 32 (4):1347–56.
Nazir, H. Z., M. Riaz, R. J. M. M. Does, and N. Abbas. 2013. Robust CUSUM control charting. *Quality Engineering* 25 (3):37–41.
Oprime, P. C., J. C. De Toledo, M. O. A. González, and S. Chakraborti. 2016. Method for determining the control limits of nonparametric charts for monitoring location and scale. *Gestao E Producao* 23 (1):146–64.
ƠZdemir. A. F. 2010. Comparing measures of location when the underlying distribution has heavier tails than normal. *_Istatistikc¸Iler Dergisi* 3:8–16.
Page, E. 1954. Continuous inspection schemes. *Biometrika* 41 (1):100–15.
Riaz, M., and S. A. Abbasi. 2016. Nonparametric double EWMA control chart for process monitoring. *Revista Colombiana de Estadística* 39 (2):167–84.
Rousseeuw, P. J., and C. Croux. 1993. Alternatives to the median absolute deviation. *Journal of the American Statistical Association* 88 (424):1273–83.
Sheu, S., and L. Yang. 2006. The generally weighted moving average control chart for monitoring the process median. *Quality Engineering* 18 (3):333–44.
Wilcox, R. 1998. The goals and strategies of robust methods. *British Journal of Mathematical and Statistical Psychology* 51 (1):1–39.
Wilcox, R. R., and H. J. Keselman. 2003. Modern robust data analysis methods: measures of Central tendency. *Psychological Methods* 8 (3):254–74.
Wilcox, R. R., and H. J. Keselman. 2003. Repeated measures one-way ANOVA based on a modified one-step M-estimator. *British Journal of Mathematical and Statistical Psychology* 56 (1): 15–25.
Woodall, W. H. 2000. Controversies and contradictions in statistical process control. *Journal of Quality Technology* 32 (4):341–50.
Yang, L., S. Pai, and Y. Wang. 2010. A Novel CUSUM Median Control Chart. Paper presented at the International Multiconference of Engineers and Computer Scientists, March 17–19, 2010, Hong Kong.
Yang, S. F., and S. W. Cheng. 2011. A new non-parametric CUSUM mean chart. *Quality and Reliability Engineering International* 27 (7):867–75.
Yang, S. F., J. S. Lin, and S. W. Cheng. 2011. A new nonparametric EWMA sign control chart. *Expert Systems with Applications* 38 (5):6239–43.
Zwetsloot, I. M., M. Schoonhoven, and R. J. M. M. Does. 2014. A robust estimator for location in phase I based on an EWMA chart. *Journal of Quality Technology* 46 (4):302–16.
Zwetsloot, I. M., M. Schoonhoven, and R. J. M. M. Does. 2015. A robust phase I exponentially weighted moving average chart for dispersion. *Quality and Reliability Engineering International* 31 (6):989–99.
Zwetsloot, I. M., M. Schoonhoven, and R. J. M. M. Does. 2016. Robust point location estimators for the EWMA control chart robust point location estimators for the EWMA control chart. *Quality Technology & Quantitative Management* 13 (1):29–38.