Journal of Engineering and Applied Sciences 13 (10): 3423-3429, 2018
ISSN: 1816-949X
© Medwell Journals, 2018

A Robust Cumulative Sum Control Chart for Monitoring the Process Mean based on a High Breakdown Point Scale Estimator

Ayu Abdul Rahman, Sharipah Soaad Syed Yahaya and Abdu Mohammed Ali Atta
College of Arts and Sciences, Universiti Utara Malaysia (UUM), 06010 Sintok, Kedah, Malaysia

Abstract: Unlike traditional Shewhart Chart, Cumulative Sum (CUSUM) chart is more sensitive to small and moderate shifts. Nonetheless, its reliability in monitoring the mean shifts is usually hampered by the underlying distribution of the data. Although, apparent cause of non-normality is owed to outliers, their presence may simply be a genuine part of the process rather than attributing to the special causes. To set these occasional outliers apart from the real distributional shifts, numerous extensions of the CUSUM charts have been suggested. One possible way is via robust estimation. This paper proposes a simple, yet effective way to make the chart highly effective for detecting small sustained shifts. A very robust scale estimator, namely Median Absolute Deviation about the median (MADn) is used as an estimate for dispersion. The performance evaluation of the proposed chart for monitoring mean shift is compared with the standard CUSUM chart using several aspects of the run length distribution the Average Run Length (ARL), Standard Deviation of the Run Length (SDRL) and percentile run length. The simulation results indicate the robust CUSUM chart efficiency in detecting small magnitude of shifts in both normal and outlier-prone data.

Key words: Average Run Length (ARL), Standard Deviation of the Run Length (SDRL), percentile of the run length, contaminated normal distribution, CUSUM control chart, MADn

INTRODUCTION

Several statistical techniques have been proposed to identify variation due to special causes. One of them is through Shewhart type chart. Though, simple and relatively easy to construct the control structure, the chart possesses some limitations. Termed as memoryless chart, the design of Shewhart control structure only uses current information of the process, thus loses its practical use when small disturbances are of interest. In this plight, control charts with more advanced scheme are compelled to be used. Such control scheme is Cumulative Sum (CUSUM) chart which incorporates historical data along with the current one in setting up the control structure. Due to this feature, CUSUM control chart is classed under the memory-type chart and lend itself to small and moderate shifts.

CUSUM control chart constructed from the sample mean was first pioneered by Page (1954). However, this chart is formed based on normality assumption. On that note, computation of the Average Run Length (ARL) is also hinged on the assumption. The ARL is used to gauge how responsive is the chart towards special causes if these variations occur in phase 2. When normality assumption is violated, the chart is expected to signal more frequently than its nominal ARL would suggest. In general, this translated to unnecessary process adjustment and loss of confidence in any chart as monitoring tools (Chang and Bai, 2004). Thus, continual efforts on the improvement of CUSUM control chart have been seen about, since, it was first introduced. The idea is to ensure that its authoritative form in phase 2 process monitoring remain indubitable, reflected by a long in-control ARL and much smaller out-of-control ARL even under disturbances to normality. One possible approach is to employ some robust techniques into CUSUM structure.

To exploit the functional use of CUSUM design in the presence of outliers, Rocke (1992) proposed to replace $\bar{X}$ with trimmed-mean in the CUSUM statistic. The resultant CUSUM chart aptly named as trimmed-mean CUSUM, showed a promising result in the departure of normality. Similar conclusion was arrived by Midi *et al*. (2004) under Student-t distribution with varying degree of freedom as well as Chi-square distributions. Following that, Yang *et al*. (2010) used median to replace $\bar{X}$ in the

Corresponding Author: Ayu Abdul Rahman, College of Arts and Sciences, Universiti Utara Malaysia (UUM), 06010 Sintok, Kedah, Malaysia

CUSUM statistic. Subsequently, they studied the performance of median-CUSUM chart in the presence of outliers. Nazir *et al.* (2013) extended Yang *et al.* (2010) research by proposing another three additional location estimators including two of the robust ones namely Hodges-Lehmann (HL) and trimean. Testing against various conditions such as contaminated-normal and non normal distributions, trimean-CUSUM chart was found to produce the best ARL performance.

Concerns are being raised on some of the abovementioned robust approaches. While able to retain a long in-control ARL under non-normality median chart is also probable to experience detection delay when shifts commence, irrespective of the data scenario. On top of that, each and every one of the deliberated robust charts (trimmed-mean, median, HL and trimean) require that data consist of rational subgroups an issue that was brought up by Hawkins (1993) in his study. Considering this limitation, Hawkins (1993) as well as Lucas and Crosier (1982) recommended a more general approach that could be applied on either individual or grouped data. Their proposals were made in the context of unknown parameters wherein the winsorization approach was adopted to attain robust parameter values. It was claimed to successfully maintain the ARL performance of the chart as the effect of occasional outliers could now be curtailed. Such outliers only affect individual reading and lead to a transient shift in the process. By limiting their effect, the functional use of CUSUM chart is safely reverted to detect small sustained shifts only. Following the same flow of thought, we proposed an alternative way to achieve similar favorable outcome in the ARL calculation. This study investigates a new way to curb the effect of occasional outliers if presence when CUSUM chart is used to monitor a process mean. Ultimately, this is attained by constructing a CUSUM chart using robust estimate of the process standard deviation, known as the average of the subgroup median absolute deviation about the median $(\overline{MAD})n$ rather than the average of the subgroup standard deviations $(\bar{S})$. While keeping the in-control ARL relatively close to the nominal value, the newly proposed chart also maintains sensitivity towards shifts in the mean. This is shown by conditioning the data on normal and mixed normal distributions. Although, rational subgroup concept is exercised in this study, the procedure can be easily implemented along the same line for sample size 1.

The standard cusum control chart: The standard CUSUM control chart is based on the following statistics (Nazir *et al*., 2013):

$$C_{u,i} = \max\{0, C_{u,\, i-1}+\left(\bar{X}_i-\mu_o\right)-K\}$$
$$C_{L,i} = \min\{0, C_{L,\, i-1}+\left(\bar{X}_i-\mu_o\right)+K \qquad (1)$$

Where:

I = Defines the subgroup number
$\bar{X}_i$ = The subgroup sample mean
μ_0 = The in-control process mean (or sometimes taken as the target mean) and
K = The reference value
$C_{u,0}$ and $C_{L,0}$ = The initial values typically set at 0

Here, the subscripts U and L denote the upper and lower part of the CUSUM statistics, respectively. To monitor the process mean, plots of both upper and lower statistics will be compared against the decision limit, H. These statistics measure the cumulative sums of deviation of data from the in-control mean, indicating an upward shift when $C_{u,\,i}$>H and vice-versa when $C_{L,\,i}$<-H. As the procedure is highly sensitive to the selection of H as well as K, extra care needs to be exercised while selecting the parameters. K and H are defined as follows (Montgomery, 2009):

$$K = k\sigma_{\bar{X}},\ H = h\sigma_{\bar{X}} \qquad (2)$$

Where:

$$\sigma\bar{x} = \frac{\sigma}{\sqrt{n}}$$

is the standard deviation of the sampling distribution of the mean estimator (i.e., the standard error of the X estimator) and n is the sample size. Let:

$$Z_i = \frac{\bar{X}_i-\mu_0}{\sigma\sqrt{n}} \qquad (3)$$

Thus, the standardized CUSUM is defined as follows.

$$C_{u,\,i} = \max\{0, C_{u,\,i-1}+\left(Z_i-k\right)\}$$
$$C_{L,\,i} = \min\{0, C_{L,\,i}+\left(Z_i-k\right) \qquad (4)$$

Where:

$$k = \frac{\delta_{opt}}{2} \qquad (5)$$

and σis the magnitude of a shift where a quick detection is required. Following the standardized CUSUM, the decision limit is now at h. When process parameters are unknown and require to be computed from an in-control phase 1 data set, then μ_0 can be estimated in practice by the average of the sample mean:

$$\bar{\bar{X}} = \frac{1}{m}\sum_{i=1}^{m}\bar{X}_i \quad (6)$$

and σ can be estimated by $\bar{s}/c_{4,n}$ where is the average of sample standard deviation defined as follows:

$$\bar{S} = \frac{1}{m}\sum_{i=1}^{m}S_i \quad (7)$$

Where:

m = The total number of subgroups being employed in the stud each with size

n and $c_{4,n}$ = A constant which only depends on the sample size n

A list of values for the factor $c_{4,n}$ could be referred by Montgomery (2009) for $n\in\{2, ..., 25\}$.

The proposed cusum control chart: Analogous to the standard CUSUM chart, the proposed method also incorporates every observed sample mean into the CUSUM statistics. Thus, Eq. 1-5 remain unchanged. However, rather than assuming Phase 1 data set is truly the representative of the process, we recognize the possibility of some contaminations in phase 1. This situation is quite common in practice as alluded by Janacek and Meikle (1997). Under this quandary, the standard estimates given in Eq. 6 and 7 can be easily perturbed by outlying values and consequently, affect the phase 2 performance of the chart. To cope with outliers, step changes and other data anomalies (Jensen *et al*., 2006) recommended studying robust or alternative estimators for μ_0 and σ. This subject is being addressed in our study. Rather than estimating σ by the most commonly used estimator $\bar{s}$, we propose an alternative with the highest possible Breakdown Point (BP) known as MADn. Also, identified as Median Absolute Deviation (MAD), the estimator has made a mark in modern robust statistical methods due to its 50% BP as well as its bounded influence function with the sharpest possible bound among scale estimators (Rousseeuw and Croux, 1993). These merits occasionally outweigh a duo setback experienced by MAD 37% efficiency at Gaussian data and less suitable for asymmetric distributions. Therefore, the effect of outliers could still be minimized on the scale estimates upon application. MAD is defined by Rousseeuw and Croux (1993):

$$MAD = med_i\left|x_i - med_j x_j\right| \quad (8)$$

Common approach is to multiply some constants say b with the output (Eq. 8) to make the estimator consistent for the parameter of interest.

$$MAD_n = bMAD \quad (9)$$

In this study, b is set at 1/0.6745 to make MAD consistent for σ at normally distributed data. To isolate the effect of using an alternate dispersion measurement on CUSUM chart performance, we retain the application of sample mean for the location parameter. Thus, μ_0 is still estimated by $\bar{\bar{x}}$ as Eq. 6 but σ is now computed as follows:

$$MAD_n = \frac{1}{m}\sum_{I=1}^{M}(\overline{MAD_n})_i \quad (10)$$

Performance evaluation and simulation: The performance of the proposed and standard control charts is evaluated using ARL, defined as the average number of points plotted on the chart until an out-of-control condition is signaled. This criterion is used to assess an in-control and an out-of-control state of the process. The in-control ARL is denoted by ARL_0 and the out-of-control ARL is referred by ARL_1. A high value of ARL_0, accompanied by low values of ARL_1 is always desirable as they signify a good control chart. To improve the overall assessment, we also incorporate additional criterion to supplement the ARL. This approach has been recommended by many researchers. For an example, Woodall (2000), Jensen *et al*. (2006) and Chakraborti (2007). The suggestion was brought up due to the extreme right-skewness frequently observed in the run length distribution when process parameters are being estimated. From this standpoint, statistical meaning of the ARL would be undermined. Thus, we compute and examine SDRL along with 10, 50, 75 and 90th percentiles of both states of the run length distribution in addition to the ARL.

The ARL values are estimated by means of simulation using SAS Version 9.4 Software. The procedures are taken as follows. First, samples of 50 in-control phase 1 of size $n\in\{5, 7, 10, 15\}$ are generated from the chosen distribution (Table 1). Then, $\hat{\upsilon}$ and $\hat{\sigma}$ are computed from the data. Note that in this study, we assume data are independent and identically distributed. We further generate 15,000 phase 2 samples of size n from the selected distribution (which is the same as phase 1 distribution) under in-control state and apply them on the charts. The chart statistics are computed and recorded whether they are within decision limits or not. The respective sample number (when either $C_{u,I}$>or $C_{L,i}$<-h) is noted as the in-control run length. The process is repeated for 10,000 simulation runs. With that, the ARL_0 is attained by averaging the value over the total runs. Ultimately, a shift is commenced. A shift in the process mean is referred by $\mu_1 = \mu_0+\delta\sigma$ where δ is the magnitude of the shift and μ_1 is the shifted mean value. We set

Table 1: Distributions applied in the study

Variables	Distribution	Percentage of outliers	Description
Normal	N (μ_0, σ^2)	0	Ideal
CN1	(0.95)N(μ_0,σ^2)+(0.05)N(μ_0, 9σ^2)	5	Mild contamination (Inflated variance)
CN2	(0.95)N(μ_0,σ^2)+(0.05)N(μ_0, 4σ^2)	5	Mild contamination (Inflated mean)

δ for {0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.75, 1, 1.5, 2.0, 2.5, 3.0}. Analogous to ARL_0, identical steps are taken to procure ARL_1 with respect to the δ.

In this study, the ideal condition is assumed to be N (μ_0, σ_2). Also, a duo of contaminated normal distribution is used to illustrate the presence of outliers in the process. The distributions are described in Table 1 where we set $\mu_0 = 0$ and $\sigma = 1$ without loss of generality.

To compare the performance of the charts, both standard and robust charts are designed to attain the same nominal ARL_0. Following that, a chart with the smaller value of ARL_1 for a specific size of shift outweighs its competitor. Occasionally called as the optimal shift, we denote this shift as δ_{opt}.

Design and derivation of limits of the standard and robust CUSUM control charts: The design of the standard and the proposed CUSUM control charts involves a derivation of optimal pair of parameters, k and h. This optimal pair is derived in such way that we obtain the nominal ARL_0 under normality. A step by step approach to get them is explained as follows:

- Fixed n
- Fixed ARL_0 when the process ~ N (μ_{00}, σ^2)
- Set δ_{opt}
- Set the optimal constant, k accordingly where $k = \delta_{opt}/2$
- Based on the designated value of k, determine h such that the CUSUM chart produces the nominal ARL_0

RESULTS AND DISCUSSION

The results of our simulation study are summarized in Table 2-5. We first give the ARL, followed by their respective SDRL in the parentheses. Table 6 illustrates the outcome for the percentile. To accomplish that, we set the nominal $ARL_0 = 370$ and $\delta_{opt} = 1.0$. The associated factor (h) can be referred in Table 2.

First, consider the situation where the distribution of the data is assumed ideal (Table 3). We see that, the ARL_1 of the robust CUSUM chart are significantly smaller than the standard chart for $\delta \leq 0.4$ and n = 5, 7. Although, the performance of the standard chart seems to improve when n>7 the robust CUSUM chart still maintains a faster detection ability for particularly small shifts resulting in a powerful control chart when small shifts are of interest. In addition, the proposed method can be considered as an alternative to the standard chart in the presence of large shifts. For $\delta \geq 1.5$, the out-of-control performance of the robust chart trails closely behind the standard chart.

Table 2: Factor (h) of CUSUM charts under N (0, 1) at $ARL_0 = 370$

n	Standard CUSUM	Robust CUSUM
5	4.994	6.996
7	5.048	6.266
10	5.072	5.917
15	5.098	5.588

The strength of the proposed CUSUM chart is further enlightened when there are contaminations in phase 1 data. From Table 4, we observe the best performance for relatively small magnitude of shifts is achieved by the robust CUSUM chart. This confirms that it is advisable to use a robust scale estimator if there are special causes in Phase I. The robust estimation results in a powerful chart for monitoring the mean shift. The CUSUM chart based on $\overline{MAD}_n$ possesses much smaller ARL_1 than its competitor when $\delta<0.5$. We discover the recurring pattern for the ARL_1, similar to the N(0, 1) distribution when both CUSUM charts are conditioned on large magnitude of shifts. On a different note, the in-control performance for both charts is quite unpredictable. Studied against different n, the ARL_0 fluctuates to a certain degree, greater than the nominal value (370). The standard CUSUM chart demonstrates a slightly better in-control robustness than its competitor for a small sample size (n = 5). But the outcome is reverted when n increases. In particular, the CUSUM chart based on $\overline{MAD}_n$ possesses a remarkable in-control robustness when n = 10 whereby the $ARL_0 \cong 370$. Since, the comparison between Table 4 and 5 shows a similar trend, it is sufficed to discuss and draw conclusion based on the preceding comparison.

The tabulated result for the SDRL is quite conspicuous wherein all the in-control SDRL is noticeably larger than the ARL_0. This clearly indicates that an extra variability is introduced into the CUSUM statistics when both mean and standard deviation are being estimated. However, the impact is predominantly dwindled as large n is used to compute $\hat{\upsilon}_0$ and $\hat{\sigma}$. Both normal and mixed normal data scenarios show smaller values of the in-control SDRL when n = 10, 15 in contrast to n = 5. On a side note, more variability is captured when CUSUM chart is constructed using $\overline{MAD}$, rather than $\bar{S}/c_{4,n}$. This is particularly evident when n = 5, conditioned on CN1. Confined under this setting, the proposed method

Table 3: ARL (SDRL) values for the standard and robust CUSUM charts under N(0, 1) based on m = 50 and δ_{opt} = 1

	n = 15		n = 7		n = 10		n = 5	
Variables	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM
0.00	370.04(511.06)	370.09(640.35)	370.25(484.19)	369.69(570.58)	370.10(441.38)	369.92(514.86)	370.29(441.83)	370.07(483.47)
0.10	217.72(343.31)	210.07(415.97)*	184.37(287.48)	177.44(320.18)*	140.06(213.86)	132.11(227.90)*	96.28(157.47)	91.84(154.94)*
0.15	126.94(221.25)	117.50(280.05)*	93.44(173.67)	86.69(162.50)*	59.53(95.15)	55.91(93.60)*	37.06(51.98)	35.21(48.22)0.2*
	68.47(111.61)	61.22(115.59)*	44.95(66.94)	42.24(70.01)*	29.21(34.78)	27.83(32.73)*	18.47(16.16)	17.98(15.25)*
0.25	39.74(56.56)	34.95(57.95)*	26.08(29.58)	24.13(27.66)*	17.37(15.55)	17.10(15.16)*	11.81(7.99)	11.77(7.55)*
0.30	25.13(27.83)	22.76(23.78)*	17.33(15.80)	16.57(14.57)*	12.35(8.98)	12.04(8.37)	8.68(4.74)	8.67(4.61)0.4*
	13.77(10.89)	13.16(9.01)*	10.06(6.19)	9.94(5.75)*	7.57(3.78)*	7.65(3.75)	5.68(2.31)	5.73(2.27)
0.50	9.17(5.39)*	9.22(4.97)*	7.11(3.47)*	7.22(3.36)	5.54(2.24)*	5.64(2.21)	4.25(1.45)*	4.34(1.50)
0.75	5.02(1.96)*	5.33(1.94)*	4.12(1.40)*	4.27(1.39)	3.33(0.99)*	3.45(1.00)	2.70(0.72)*	2.76(0.73)
1.00	3.55(1.13)*	3.80(1.16)	2.97(0.085)*	3.13(0.87)	2.48(0.63)*	2.56(0.64)	2.09(0.42)*	2.12(0.43)1.5
	2.32(0.56)*	2.50(0.61)	2.04(0.41)*	2.13(0.43)	1.80(0.43)*	1.87(0.38)	1.41(0.49)*	1.49(0.50)
2.00	1.85(0.41)*	1.99(0.36)	1.59(0.49)	1.72(0.46)	1.23(0.42)*	1.32(0.47)	1.02(0.13)	1.03(0.16)
2.50	1.47(0.50)*	1.70(0.46)	1.15(0.36)*	1.26(0.44)	1.01(0.11)	1.03(0.16)	1.00(0.01)	1.00(0.01)3
	1.12(0.33)*	1.31(0.46)	1.01(0.10)	1.03(0.18)	1.00(0.01)	1.00(0.01)	1.00(0.00)	1.00(0.00)

Table 4: ARL (SDRL) values for the standard and robust CUSUM charts under CN1 distribution based on m = 50 and δ_{opt} = 1

	n = 15		n = 7		n = 10		n = 5	
Variables	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM
0.00	377.59(530.22)	391.51(705.41)	377.75(459.84)	379.75(576.51)	374.47(457.55)	369.35(493.95)	381.88(443.37)	378.27(484.67)
0.10	237.48(362.01)	236.35(444.90)*	203.83(310.51)	200.76(373.64)*	167.19(266.72)	157.72(262.36)*	119.00(183.99)	116.01(193.32)*
0.15	152.83(250.35)	142.55(328.30)*	111.26(177.69)	105.74(199.21)*	77.26(121.67)*	71.58(117.19)*	47.61(70.57)	47.41(78.99)*
0.20	83.87(141.52)	77.73(165.93)*	58.79(92.01)	53.23(87.92)*	37.58(48.68)	34.94(44.04)*	23.45 (24.40)	22.50(22.88)*
0.25	50.64(84.89)	44.51(76.75)*	32.94(42.40)	31.19(47.46)*	22.37(23.70)	21.23(22.08)*	14.39(11.26)	14.09(10.67)*
0.30	33.03(44.54)	29.46(41.97)*	21.40(21.60)	20.03(21.47)*	15.06(12.13)	14.65(11.57)*	10.41(6.58)	10.37(6.39)*
0.40	16.69(15.13)	15.75(12.67)*	11.99(8.39)	11.77(7.72)*	8.83(4.82)*	8.90(4.66)	6.51(2.91)*	6.63(2.93)
0.50	10.65(6.74)*	10.74(6.29)	8.19(4.35)*	8.21(4.06)*	6.31(2.81)*	6.42(2.78)	4.87(1.80)*	4.90(1.75)
0.75	5.71(2.43)*	5.97(2.32)	4.66(1.72)*	4.82(1.71)	3.78(1.22)*	3.88(1.20)	3.01(0.85)*	3.08(0.85)1
	4.00(1.36)*	4.30(1.39)	3.31(0.99)*	3.47(1.01)	2.75(0.75)*	2.84(0.75)	2.26(0.51)*	2.31(0.53)
1.50	2.55(0.66)*	2.78(0.71)	2.21(0.50)*	2.30(0.52)	1.93(0.38)*	1.99(0.36)	1.63(0.49)*	1.69(0.46)
	2.00(0.41)*	2.14(0.43)	1.78(0.43)*	1.88(0.38)	1.44(0.50)*	1.56(0.50)	1.08(0.28)*	1.12(0.32)
2.50	1.67(0.47)*	1.86(0.39)	1.35(0.48)*	1.49(0.50)	1.06(0.24)*	1.11(0.31)	1.00(0.03)	1.00(0.05)
3.00	1.30(0.46)	1.53(0.50)	S1.06(0.24)*	1.13(0.33)	1.00(0.05)	1.01(0.07)	1.00(0.00)	1.00(0.00)

Table 5: ARL (SDRL) values for the standard and robust CUSUM charts under CN2 distribution based on m = 50 and δ_{opt} 1

	n = 15		n = 7		n = 10		n = 5	
Variables	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM	Standard CUSUM	Robust CUSUM
0.00	386.42(538.61)	388.48(685.86)	378.80(486.69)	370.68(573.70)	371.42(452.45)	358.64(488.13)	380.42(446.25)	378.14(481.75)
0.10	215.61(328.84)	210.55(441.85)*	180.95(288.23)	178.59(334.92)*	138.09(214.94)	133.94(241.61)*	97.62(163.28)	96.40(185.48)
0.15	127.02(216.53)	118.79(245.92)*	90.43(154.68)	83.49(158.26)*	62.51(95.65)	58.15(101.13)*	37.39(57.17)	36.36(56.94)
0.20	68.45(117.22)	61.16(117.77)*	45.98(73.10)	42.15(71.32)*	29.23(34.19)	27.84(35.26)*	18.65(18.17)	18.18(16.43)
0.25	39.05(57.42)	35.14(56.59)*	26.69(33.18)	24.39(29.51)*	17.92(16.28)	17.42(15.40)*	11.91(8.70)	11.82(7.59)*
0.30	25.38(30.63)	22.90(25.44)	17.36(15.42)	16.54(13.32)	12.25(8.96)	12.12(8.17)	8.68(4.76)	8.63(4.55)
0.40	13.58(10.73)	13.11(8.97)*	10.01(6.04)	9.93(5.68)*	7.58(3.81)*	7.60(3.69)	5.71(2.38)*	5.78(2.36)
0.50	9.07(5.14)*	9.17(4.70)	7.13(3.48)*	7.16(3.27)	5.50(2.24)*	5.63(2.22)	4.29(1.50)*	4.32(1.49)
0.75	5.06(1.97)*	5.38(2.00)	4.14(1.43)*	4.27(1.42)	3.34(0.99)*	3.45(1.01)	2.71(0.72)*	2.77(0.73)
1.00	3.54(1.11)	3.83(1.16)	2.97(0.84)	3.11(0.86)	2.49(0.63)	2.57(0.66)	2.09(0.42)	2.12(0.43)
1.50	2.32(0.56)*	2.52(0.61)	2.04(0.41)*	2.12(0.42)	1.80(0.42)*	1.87(0.38)	1.42(0.49)*	1.49(0.50)
2.00	1.86(0.41)*	1.99(0.37)	1.61(0.49)*	1.73(0.45)	1.23(0.42)*	1.33(0.47)	1.02(0.13)	1.03(0.17)
2.50	1.46(0.50)	1.69(0.46)	1.15(0.36)	1.27(0.45)	1.01(0.11)	1.03(0.16)	1.00(0.00)	1.00(0.00)
3.00	1.13(0.33)*	1.32(0.47)	1.01(0.10)	1.03(0.18)	1.00(0.00)	1.00(0.02)	1.00(0.00)	1.00(0.00)

ARL value *indicates a better out-of-control performance between the two charts for a specific sample size n and shift δ

captures a substantial amount of variability, far more than the standard method (Table 4: 750.41>530.22). The effect is identical on CN2 (Table 5). Overall, the out-of-control SDRL is larger than ARL_1 for $\delta \leq 0.2$ though the gap is drastically reduced as δ and n increase. This is noted for both charts.

Next, we provide a brief explanation on the percentile of the run length. Here, we limit our discussion to n = 7, based on the ideal data condition. This sample size is regarded as an intermediate between small and large n, which gives a fair evaluation when we contrast the behavior between the two charts. Considering the result

Table 6: Percentile run length values for the CUSUM chart under N (0.1) distribution with n = 7 and $\delta_{opt} = 1$ at $ARL_0 = 370$

	δ											
Variales/Percentile	0	0.1	0.2	0.3	0.4	0.5	0.75	1	1.5	2	2.5	3
Standard CUSUM chart												
P_{10}	35	17	9	6	5	4	3	2	2	1	1	1
P_{25}	85	35	14	8	6	5	3	2	2	1	1	1
P_{50}	213	87	26	13	9	6	4	3	2	2	1	1
P_{75}	465	212	50	21	12	9	5	3	2	2	1	1
P_{90}	879	443	96	33	17	12	6	4	2	2	2	1
Robust CUSUM chart												
P_{10}	32	16	9	6	5	4	3	2	2	1	1	1
P_{25}	74	32	14	9	6	5	3	3	2	1	1	1
P_{50}	189	77	24	13	8	6	4	3	2	2	1	1
P_{75}	434	189	46	20	12	9	5	4	2	2	2	1
P_{90}	892	421	87	30	17	11	6	4	3	2	2	1

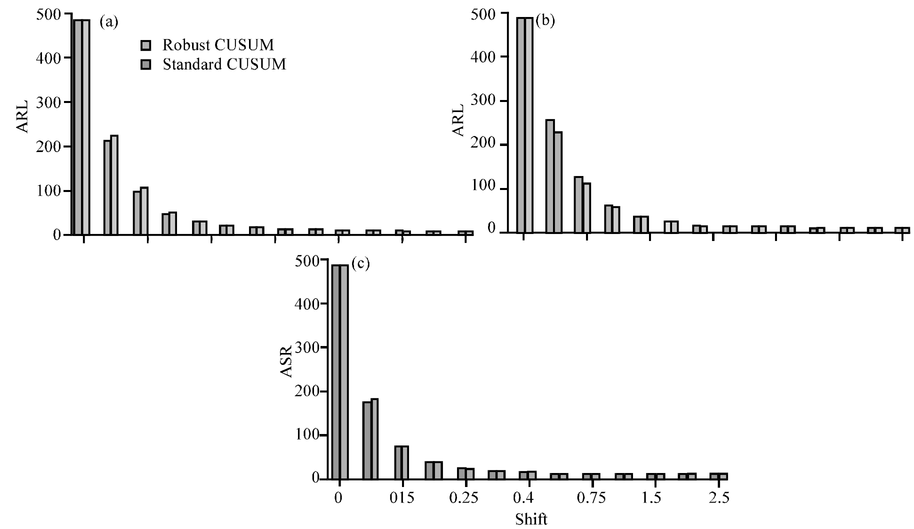

Fig. 1: Comparison between the robust CUSUM and the standard CUSUM chart in terms of ARL values; a) Output for N (0.1) distribution with n = 7; b) Output for CN1 distribution with n = 7 and c) Output for CN2 distribution with n = 7

provided in Table 6, we observe that the run length distribution of both CUSUM charts is influenced by shift in the process mean. It is highly right skewed when the process is in-control. The pattern follows through when the shift starts to commence. As the magnitude becomes significantly large, the shape of the run length distribution changes to nearly symmetrical. The analysis could be easily extended to the CNI and CN2 distributions. Although, the results are not showing here, the shape is still skewed to the right especially for $\delta = 0$, comparable to Table 6. Additionally, the percentile can be used to assess Median Run Length (MRL). Equitable to the 50th percentile (P_{50}), MRL is an alternative measure to compare performance of CUSUM charts.

The tabulated ARL validates the sensitivity of the proposed robust CUSUM chart under the influence of exceptionally small sustained shifts with respect to neither the data distributions nor the sample sizes. However, it remains to be shown that similar outcome could be achieved for a different fixed in-control ARL. To do so, we design both charts (the standard and the robust) for a nominal $ARL_0 = 500$ but keeping the δ_{opt} at 1.0. For m = 50 and n = 7, the resulting factors (h) are equal to 5.364 for the standard CUSUM chart and 6.656 for the robust chart. (Fig. 1a-c) illustrates the results.

Under normality, both charts behave as expected the $ARL_0 \cong 500$. For small magnitude of shifts ($0<\delta\leq 0.4$), the performance of our CUSUM chart triumphs the standard one. From there onwards, nor charts hold power over the other as their ARL are comparatively equivalent. This is demonstrated in Fig. 1a.

Both charts are expected to encounter detection delay when we inflate the variance of the process. When small shifts commence, the ARL_1 of both charts are noticeably larger than the normal case. A case in point of δ = 0.1 for CN1 data, yields ARL_1>250. For the same value of δ and under the ideal condition, the robust and the standard charts acquire ARL_1 of 220 and 230, respectively. Yet, as value of δ rises, the disparity in ARL_1 between the normal data (Fig. 1a) and CN1 data (Fig. 1b) slowly diminish. Even so, the robust chart is still potentially more efficient than the standard chart. This is further reinforced when we conditioned both charts on the CN2 distribution. We do note, however, that the standard chart seems to own a robust in-control performance, comparable to the robust CUSUM chart when outliers are presence. This seems to negate the prior simulation outcome based on the designated chart for ARL_0 370. However, bear in mind that this section of analysis is conducted on a solitary sample size (n = 7). Thus, rather than generalizing this result as a whole, we speculate that the performance of CUSUM may very well depend on the designated constants. This includes several factors such as the nominal value of in-control ARL, δ_{opt} and n.

CONCLUSION

In this study, we proposed a new robust CUSUM control chart for monitoring the prosess mean, constructed using $\overline{MAD_n}$. It is a new way to robustify a CUSUM chart, so that, it is less sensitive to occasional outliers while at the same time, still able to discern sustained distributional shifts. A highly robust scale estimator with 50% breakdown point is not easily perturbed by the outlying values in Phase I. Thus, it becomes a superior alternative to the $\overline{S}/c_4, n$ if outliers are presence. The application of $\overline{MAD_n}$ with CUSUM structure allows the chart to be reliable when a rather small magnitude of shifts is of interest. This concurs in both normal and outlier-prone data. To examine the in-control robustness of the CUSUM chart, the process is calibrated using contaminant data of the same type as are subsequently monitored. The goal is to keep the ARL relatively close to the nominal level. It was shown that for moderate n, this target can be met, quite well by the proposed robust method.

ACKNOWLEDGEMENT

This research is supported in part by the Fundamental Research Grant Scheme of the Ministry of Higher Education, Malaysia at Universiti Utara Malaysia.

REFERENCES

Chakraborti, S., 2007. Run length distribution and percentiles: The Shewhart chart with unknown parameters. Qual. Eng., 19: 119-127.

Chang, Y.S. and D.S. Bai, 2004. A multivariate T2 control chart for skewed populations using weighted standard deviations. Qual. Reliab. Eng. Intl., 20: 31-46.

Hawkins, D.M., 1993. Robustification of cumulative sum charts by winsorization. J. Qual. Technol., 25: 248-261.

Janacek, G.J. and S.E. Meikle, 1997. Control charts based on medians. J. R. Stat. Soc. Ser., 46: 19-31.

Jensen, W.A., L.A. Jones-Farmer, C.W. Champ and W.H. Woodall, 2006. Effects of parameter estimation on control chart properties: A literature review. J. Qual. Technol., 38: 349-364.

Lucas, J.M. and R.B. Crosier, 1982. Robust CUSUM: A robustness study for CUSUM quality control schemes. Commun. Stat. Theor. Methods, 11: 2669-2687.

Midi, H., A. Jaafar and H.A. Lim, 2004. [Comparison of standard min and min fastest in CUSUM chart (In Indonesian)]. Math., 20: 69-75.

Montgomery, D.C., 2009. Introduction to Statistical Quality Control. 6th Edn., John Wiley & Sons, New York, USA., ISBN:9780470233979, Pages: 734.

Nazir, H.Z., M. Riaz, R.J. Does and N. Abbas, 2013. Robust CUSUM control charting. Qual. Eng., 25: 211-224.

Page, E.S., 1954. Continuous inspection schemes. Biometrika, 41: 100-115.

Rocke, D.M., 1992. XQ and RQcharts: Robust control charts. Statistician, 41: 97-104.

Rousseeuw, P.J. and C. Croux, 1993. Alternatives to the median absolute deviation. J. Am. Statist. Assoc., 80: 1273-1283.

Woodall, W.H., R.W. Hoerl, A.C. Palm and D.J. Wheeler, 2000. Controversies and contradictions in statistical process control-discussion-response. J. Qual. Technol., 32: 341-350.

Yang, L., S. Pai and Y.R. Wang, 2010. A novel CUSUM median control chart. Proceedings of the 2010 International Multi-Conference on Engineers and Computer Scientists (IMECS'10), March 17-19, 2010, IAENG, Hong Kong, ISBN:978-988-18210-5-8, pp: 1-4.