Mathematical Sciences (2018) 12:249–262
https://doi.org/10.1007/s40096-018-0265-1

ORIGINAL RESEARCH

Asymmetric control limits for range chart with simple robust estimator under the non-normal distributed process

Derya Karagöz[1]

Received: 25 May 2018 / Accepted: 27 September 2018 / Published online: 4 October 2018

Abstract
This paper aims to modify Shewhart, the weighted variance and skewness correction methods in industrial statistical process control. The robust and asymmetric control limits of range chart are constructed to use in contaminated and skewed distributed process. The way of construction of control limits is simple and corresponds to three methods in which sample range estimator is replaced with the robust interquartile range. These three modified methods are evaluated in terms of their type I risks and average run length by using simulation study. The performance of the proposed range charts is assessed when the Phases I and II data are uncontaminated and contaminated. The Weibull, gamma and lognormal distributions are chosen since they can represent a wide variety of shapes from nearly symmetric to highly skewed.

Keywords Skewed distributions · Shewhart method · Weighted variance method · Skewness correction method · Robust estimator

Introduction

When the quality variable has a skewed distribution, it might be misleading to observe the process by using the Shewhart $\bar{X}$ and R control charts. The usage of Shewhart control charts in skewed distributions causes an increase in type I risk (p) when the skewness increases because of the variability in population. For this reason, three methods which use the asymmetric control limits were considered as an alternative to the classical method [13]. The first one is the weighted variance (WV) method proposed by Choobineh and Ballard [6], which is based on the semivariance approximation of Choobineh and Branting [5]. They obtained the asymmetric control limits of $\bar{X}$ and R charts for skewed distributions based on the standard deviation of sample means and ranges. Bai and Choi [2] also proposed a simple heuristic method of constructing $\bar{X}$ and R charts by using the WV method. The second one is the weighted standard deviations (WSD) proposed by Chang and Bai [4] to obtain control limits by decomposing the standard deviation into two parts. The last one is a skewness correction (SC) method proposed by Chan and Cui [3] for constructing $\bar{X}$ and R chart by taking into consideration the degree of skewness of the process distribution, with no assumptions on the distribution. Karagöz and Hamurkaroğlu [13] worked on $\bar{X}$ and R control charts for skewed distributions which are Weibull, gamma and lognormal. Classical methods of estimating parameters of the distribution of quality characteristic may be affected by the presence of outliers. In order to overcome such situation, robust estimators, which are less affected by the extreme values or small departures from the model assumptions, are introduced in industrial application. Abu-Shawiesh [1] presented a simple approach to robust estimation of the process standard deviation based on a very robust scale estimator, namely, the median absolute deviation (MAD) from the sample median. The proposed method provides an alternative to the Shewhart S control chart. Schoonhove et al. [19] studied design schemes for the standard deviation control charts with estimated parameters. Different estimators of the standard deviation were considered, and the effect of the estimator on the performance of the control charts under non-normality was investigated.

Jensen et al. [12] conducted a literature survey of the effects of parameter estimation on control chart properties and identified several issues for future research. The effect of using robust or other alternative estimators has not been studied thoroughly. Most evaluations of performance have

✉ Derya Karagöz
deryacal@hacettepe.edu.tr

[1] Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey

considered standard estimators based on the sample mean and standard deviation and have used the same estimators for both Phases I and II. However, in Phase I applications, it seems more appropriate to use an estimator that will be robust to outliers, step changes and other data anomalies. Examples of paper discussing robust estimation methods in Phase I control charts include [7, 16, 17, 25, 26]. One of Jensen et al. [12] their recommendations is to consider the effect of using these robust estimators on Phase II performance. By considering this recommendation, Schoonhove et al. [19] study the impact of these estimators on the Phase II performance of standard deviation control chart.

Recently works on control charts: Sukparungsee [23] studied the robustness of the asymmetric Tukey's control chart for skew and non-skew distributions as Lognormal and Laplace distributions. The results found that the asymmetric performs better than symmetric Tukey's control chart for both cases of skew and non-skew process observation. Sindhumol et al. [22] introduced a modification to trimmed standard deviation to increase its efficiency and it is used in controlling process dispersion. Authors constructed a Phase I control chart derived from standard deviation of trimmed mean, which is robust. Wei-Heng et al. [11] proposed a new control chart for monitoring the standard deviation of a lognormal process based on the methodologies studied in Tang and Yeh [24]. The fundamental assumption in deriving the approximate confidence intervals in Tang and Yeh [24] was that the variance of the log-transformed normal distribution is less than 1. If the variance is larger than 1, they further derived an approximate confidence interval and develop the control chart accordingly. The proposed chart was compared to the existing charts based on the average run length (ARL), where the run length is defined as the number of samples taken before the first out-of-control signal shows up on a control chart. Duclos and Pillet [8] proposed the use of a control chart (L chart) build with a minimum variance estimator whose performances have been compared to those of the average in term of variance and distribution shape. They studied this estimator in the case of data incoming from a Multi -generator process. Koyuncu and Karagöz [14] proposed to construct the mean control chart limits based on Shewhart, weighted variance and skewness correction methods using simple random sampling, ranked set sampling, median ranked set sampling and neoteric ranked set sampling designs. The performance of the proposed control charts based on neoteric ranked set sampling designs is compared with their counterparts in ranked set sampling, median ranked set sampling and simple random sampling by Monte Carlo simulation.

In this paper, we consider this recommendation to construct asymmetric control limits of R charts under non-normality and contamination. We propose to modify the Shewhart, WV and SC methods by using the interquartile range estimator of the standard deviation. And we called them modified Shewhart (MS), modified weighted variance (MWV) and modified skewness correction (MSC) methods, respectively. We study on the effect of the robust estimator on control chart performance under non-normality for moderate sample size (30 subgroups of 5–10). The considered standard estimator is interquartile range. The performance of the estimator is evaluated by assessing their root mean squared error (RMSE) under skewed distribution and in the presence of several types of contamination. Moreover, we derive factors of range control chart for each modified method. The modified robust methods are evaluated in terms of their type I risks and average run length and then compared with the modified Shewhart method. By using Monte Carlo simulation, the p and ARL values of proposed R control charts are compared based on classic and robust estimators. The performance of the proposed robust range charts is assessed when the Phases I and II data are uncontaminated and contaminated skewed distributed process. The Weibull, gamma and lognormal distributions are chosen since they can represent a wide variety of shapes from nearly symmetric to highly skewed. Khodabin and Ahmadabadi [10] was introduced the generalized gamma (GG) distribution that is a flexible distribution in statistical literature, and has exponential, gamma, and Weibull as subfamilies, and lognormal as a limiting distribution.

The remainder of the paper is structured as follows. The next section presents the design schemes and gives the methods. In the subsequent "Measuring estimator's efficiency" section, the efficiency of measuring estimators is described and the control chart constants are given in "Determination of control charts constants" section. The performance of methods is evaluated in "The performance of the modified methods" section by considering simulation study. "Results" section evaluates the results of the study. Finally, a conclusion of this study is given in "Conclusion" section.

Skewed distributions, estimators and modified methods

The main interest of this section is to give all mathematical details by regarding the robust R control charts for skewed distributions. Firstly, the skewed distributions are discussed in "Skewed distributions" section. Secondly, the classic and robust estimators are given in "Classic and robust estimators" section. We propose to modify Shewhart, WV and SC methods by replacing the mean of the subgroup ranges with the mean of the subgroup interquartile ranges. And finally, the modified methods based on robust estimator for skewed distributions are given in "Modified methods" section.

Skewed distributions

The Weibull, gamma and lognormal distributions are chosen as skewed distributions since they can represent a wide variety of shapes from nearly symmetric to highly skewed.

- The probability density function of the Weibull distribution is defined as

  $$f(x|\beta,\lambda)=\beta\lambda^{\beta}x^{\beta-1}\exp(-x\lambda)^{\beta}$$

  for $x>0$, where β is a shape parameter and λ is a scale parameter.
- The probability density function of the gamma distribution is defined as

  $$f(x|\alpha,\beta)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}\exp\left(-\frac{x}{\beta}\right)$$

  for $x>0$, where α is a shape parameter and β is a scale parameter.
- The probability density function of the lognormal distribution is defined as

  $$f(x|\sigma,\mu)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)$$

  for $x>0$, where σ is a scale parameter and μ is a location parameter.

Classic and robust estimators

The process is assumed to be in control (i.e., in Phase I) with given $\hat{\sigma}$. The process parameters μ and σ are estimated from samples, and the resulting estimates are used to monitor the process in Phase II. We define $\hat{\mu}$ and $\hat{\sigma}$ as unbiased estimates of μ and σ, respectively, based on the number of sample k.

The first scale estimator is the mean of the sample range

$$\bar{R}=\frac{1}{k}\sum_{i=1}^{k}R_i \tag{2.1}$$

where R_i is the range of the ith sample. An unbiased estimator of σ is $\bar{R}/d_2(n)$. We also consider the mean of the sample interquartile ranges since the mean of the sample range is not robust against to outliers. The mean of the sample interquartile ranges (IQRs) is defined by

$$\mathrm{I\bar{Q}R}=\frac{1}{k}\sum_{i=1}^{k}\mathrm{IQR}_i \tag{2.2}$$

where IQR_i is the interquartile range of sample i

$$\mathrm{IQR}_i=Q_{75,i}-Q_{25,i} \tag{2.3}$$

where $Q_{r,i}$ is the rth percentile of the values in sample i.

Modified methods

In this section, we construct the control limits of R control chart by considering modification in the Shewhart, WV and SC methods. The control limits are derived by assuming that the parameters of the process are unknown. What actually we do is to use simple robust estimator in these three models under the contaminated skewed process. These proposed models are called the MS, MWV and MS methods. When the control limits of MS are symmetric for normal distributed process, the control limits of MWV and MSC are asymmetric for the skewed distributed process.

The MS method

The conventional control charts when the distribution is normal are the Shewhart control charts. We first consider the Shewhart method proposed by Montgomery [15]. The Shewhart R control chart limits are given as follows:

$$\mathrm{UCL}_{\mathrm{Shewhart}}=\left(1+\frac{3d_3}{d_2}\right)\bar{R}, \tag{2.4}$$

$$\mathrm{LCL}_{\mathrm{Shewhart}}=\left(1-\frac{3d_3}{d_2}\right)\bar{R} \tag{2.5}$$

where d_2 and d_3 are constants that depend on the subgroup size n, and are calculated when the distribution is normal [15].

The MS R control chart limits are derived by replacing the range with the interquartile range as follows:

$$\mathrm{UCL}_{\mathrm{MS}}=\left(1+\frac{3d_3^Q}{d_2^Q}\right)\mathrm{I\bar{Q}R}, \tag{2.6}$$

$$\mathrm{LCL}_{\mathrm{MS}}=\left(1-\frac{3d_3^Q}{d_2^Q}\right)\mathrm{I\bar{Q}R} \tag{2.7}$$

where d_2^Q and d_3^Q are constants that depend on the subgroup size n, and are calculated when the distribution is skewed.

The MWV method

The WV method was proposed by Choobineh and Ballard [6]. The WV method decomposes the skewed distribution into two parts at its mean, and both parts are considered symmetric distributions which have the same mean and different standard deviation. In this method, μ_R is normally estimated using the mean of the subgroup ranges $\bar{R}$.

When the parameters of the process are unknown, the WV R control chart limits are defined by Bai and Choi [2] as follows:

$$\mathrm{UCL_{WV}} = \bar{R}\left[1 + 3\frac{d_3^*}{d_2^*}\sqrt{2\hat{P}_x}\right] \tag{2.8}$$

$$\mathrm{LCL_{WV}} = \bar{R}\left[1 - 3\frac{d_3^*}{d_2^*}\sqrt{2\left(1-\hat{P}_x\right)}\right] \tag{2.9}$$

where d_2^* and d_3^* are the control chart constants for R chart based on WV. These constants which are defined as the mean and standard deviation of relative range $\left(\frac{R}{\sigma}\right)$ have been obtained under the non-normality assumption. These values can be computed via numerical integration once the distribution is specified. In Eq. (2.9) P_X indicates the probability that can be estimated by using the number of observations less than or equal to

$$\bar{\bar{X}} : \hat{P}_X = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}\delta\left(\bar{\bar{X}} - X_{ij}\right)}{nk} \tag{2.10}$$

where k and n are the number of samples and the number of observations in a subgroup, and $\delta(X) = 1$ for $X \geq 0, 0$ otherwise. Usually, μ_x is estimated by the grand mean of the subgroup means $\bar{\bar{X}}$ and μ_R is estimated by the mean of the subgroup ranges $\bar{R}$ [2].

In this paper, we propose the MWV method in which the mean of the subgroup ranges is replaced by the mean of the subgroup interquartile ranges. If the parameters of the process are unknown, the MWV R control chart limits are given by

$$\mathrm{UCL_{MWV}} = \bar{\mathrm{IQR}}\left[1 + 3\frac{d_3^Q}{d_2^Q}\sqrt{2\hat{P}_x^Q}\right] \tag{2.11}$$

$$\mathrm{LCL_{MWV}} = \bar{\mathrm{IQR}}\left[1 - 3\frac{d_3^Q}{d_2^Q}\sqrt{2\left(1-\hat{P}_x^Q\right)}\right] \tag{2.12}$$

where d_2^Q and d_3^Q are the control chart constants of MWV R control charts. These constants which are defined as the mean and standard deviation of interquartile range $\left(\frac{\mathrm{IQR}}{\sigma}\right)$ have been obtained under the non-normality assumption, see in "Measuring estimator's efficiency" and "Determination of control charts constants" sections. In this paper, this constant based on classic and robust estimators is obtained via simulation for each skewed distribution, because of the difficulty in numerical integration. Equation (2.12) allows the probability to be estimated from

$$\hat{P}_X^Q = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}\delta\left(\bar{\mathrm{TM}}_\alpha - X_{ij}\right)}{nk} \tag{2.13}$$

where k and n are the number of samples and the number of observations in a subgroup, respectively, and $\delta(X) = 1$ for $X \geq 0, 0$ otherwise. In Eq. (2.13), $\bar{\mathrm{TM}}_\alpha$ is the mean of the sample trimmed means, defined by

$$\bar{\mathrm{TM}}_\alpha = \frac{1}{k}\sum_{i=1}^{k}\bar{\mathrm{TM}}_{(v)_i} \tag{2.14}$$

where $\mathrm{TM}_{(v)_i}$ denotes the vth ordered value of the sample trimmed means defined by

$$\bar{\mathrm{TM}}_{(v)_i} = \frac{1}{n - 2\lceil n\alpha\rceil}\left[\sum_{j=\lceil n\alpha\rceil+1}^{n-\lceil n\alpha\rceil} X_{(ij)}\right] \tag{2.15}$$

where α denotes the percentage of samples to be trimmed and $\lceil n\alpha\rceil$ denotes the ceiling function, i.e., the smallest integer not less than $n\alpha$.

The MSC method

The last method being considered is the SC method proposed by Chan and Cui [3]. They proposed to construct the $\bar{X}$ and R control charts limits for SC method under the skewed distributions. It's asymmetric control limits are obtained by taking into consideration the degree of skewness estimated from subgroups and making no assumptions about distributions.

If the parameters of the process are unknown, the SC R control chart limits are defined by Chan and Cui [3] as follows:

$$\mathrm{UCL_{SCR}} = \left[1 + \left(3 + d_4^*\right)\frac{d_3^*}{d_2^*}\right]\bar{R} \tag{2.16}$$

$$\mathrm{LCL_{SCR}} = \left[1 + \left(-3 + d_4^*\right)\frac{d_3^*}{d_2^*}\right]^+\bar{R} \tag{2.17}$$

where d_4^* is the control chart constant that is obtained as follows:

$$d_4^* = \frac{\frac{4}{3}k_3(R)}{1 + 0.2k_3^2(R)} \tag{2.18}$$

where $k_3(R)$ is the skewness of the subgroup range R [3].

In this paper, we propose MSC method in which the mean of the subgroup ranges is replaced by the mean of the subgroup interquartile ranges. If the parameters of the process are unknown, the MSC R control chart limits are defined as follows:

$$\mathrm{UCL}_{\mathrm{MSCR}} = \left[1 + \left(3 + d_4^Q\right)\frac{d_3^Q}{d_2^Q}\right]\bar{\mathrm{IQR}} \tag{2.19}$$

$$\mathrm{LCL}_{\mathrm{MSCR}} = \left[1 + \left(-3 + d_4^Q\right)\frac{d_3^Q}{d_2^Q}\right]^+ \bar{\mathrm{IQR}} \tag{2.20}$$

where d_4^Q are the control chart constant which is obtained for the MSC method as follows:

$$d_4^Q = \frac{\frac{4}{3}k_3(\mathrm{IQR})}{1 + 0.2k_3^2(\mathrm{IQR})} \tag{2.21}$$

where k_3 (IQR) is the skewness of the subgroup interquartile ranges.

Simulation study

The considered standard deviation estimator is interquartile range. The performance of the estimator is evaluated by assessing their RMSE under skewed distribution and in the presence of several types of contamination. The simulation studies evaluate the efficiency of measuring estimator in "Measuring estimator's efficiency" section, the control chart constants in "Determination of control charts constants" section and the performance of modified methods in "The performance of the modified methods" section.

Measuring estimator's efficiency

In this section, we evaluate the effect of outliers on the accuracy of the conventional and proposed robust estimators by means of a Monte Carlo simulation. ($M = 50{,}000$) simulation runs of 30 ($k = 30$) subgroups each of size n = 5, 10 are performed to generate data under the skewed distributions. The generated data are Weibull, lognormal and gamma distributions with different parameters as presented in Table 1. The process dispersion is estimated by both classic and robust methods. We consider four models in the case of no outliers and outliers like [9],

- **Model 1**: The reference distribution parameters are selected with respect to skewness of distribution that is given in Table 1.
- **Model 2**: The case of 10% replacement outliers coming from another Weibull distribution with a different scale parameter ($\lambda_1 = 0.2$) and a shape parameter ($\beta_1 = 0.2 * \beta$) , another lognormal distribution with a different location parameter ($\mu_1 = 0.2$) and a scale parameter ($\sigma_1 = 2 * \sigma$) and another gamma distribution with a different shape parameter ($\alpha_1 = 2\alpha$) and a scale parameter ($\beta_1 = 0.2$).
- **Model 3**: The case of 10% replacement outliers from a uniform distribution on [0, 20].
- **Model 4**: The more extreme case of 10% of outliers placed at 50. We replace 10% of observations from the data with extreme values such as 50 to create a outliers in the data.

We thus allow that some observations come from a different skewed population, and in the last two models, we allow for the occurrence of gross errors.

We run the simulation $M = 50{,}000$ times and generate $k = 30$ samples of size $n = 5, 10$ according to different simulation schemes and compute the scale estimate $\hat{\sigma}_j$ for each sample for $j = 1, \dots, M$. For each simulation setting and for estimators, we compute the RMSE of the scale estimator

$$\mathrm{RMSE}_\sigma = \sqrt{\frac{1}{M}\sum_{j=1}^{M}(\hat{\sigma}_j - \sigma)^2}.$$

where $\hat{\sigma}_j$ is the robust estimation of the standart deviation $\hat{\sigma}$.

The results for Weibull, lognormal and gamma distributions are reported in Table 2. The conclusions from the study are as follows:

(i) When there is no contamination for small sample size, the efficiency of the classic and robust estimators is more or less similar. However, for the large sample size, the robust estimator of scale performs better than the classic estimator when no contamination is present.

(ii) Contamination by extreme outliers causes a large increase in the RMSE of the classical estimator, especially for large samples $n = 10$ and a much smaller increase in the RMSE of the robust alternative. The

Table 1 Values of the skewness and the parameters of distributions

	k_3	0.50	1.00	1.50	2.00	2.50	3.00
Weibull	β	2.15	1.57	1.20	1.00	0.86	0.77
Lognormal	σ	0.16	0.32	0.44	0.54	0.66	0.72
Gamma	α	16.00	4.00	1.80	1.00	0.64	0.44

Table 2 RMSE of the scale (σ) estimator under the skewed distributions for $n = 5, 10$

Model/k_3	$n = 5$						$n = 10$					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
Weibull distribution												
Model 1												
Classic	0.0293	0.0446	0.0712	0.1048	0.1520	0.2036	0.1363	0.1899	0.2742	0.3741	0.5095	0.6592
Robust	0.0337	0.0491	0.0738	0.1042	0.1460	0.1924	0.0290	0.0409	0.0623	0.0934	0.1453	0.2117
Model 2												
Classic	0.0315	0.0492	0.0799	0.1176	0.1707	0.2291	0.1467	0.2093	0.3067	0.4230	0.5791	0.7509
Robust	0.0382	0.0578	0.0901	0.1298	0.1836	0.2432	0.0347	0.0501	0.0686	0.0873	0.1154	0.1511
Model 3												
Classic	1.5846	1.5864	1.5966	1.6086	1.6191	1.6268	2.5950	2.6168	2.6653	2.7317	2.8020	2.8895
Robust	0.7251	0.7575	0.8178	0.8898	0.9804	1.0721	0.0591	0.0893	0.1281	0.1664	0.2067	0.2429
Model 4												
Classic	8.4085	8.5658	8.8463	9.1533	9.5031	9.8179	13.3728	13.6361	14.1158	14.6384	15.2350	15.7725
Robust	3.6826	3.8093	4.0516	4.3354	4.6874	5.0316	0.0621	0.0942	0.1386	0.1826	0.2323	0.2821
Lognormal distribution												
Model 1												
Classic	0.0314	0.0752	0.1247	0.1839	0.2902	0.3610	0.1481	0.3266	0.5037	0.6972	1.0141	1.2212
Robust	0.0352	0.0798	0.1263	0.1777	0.2675	0.3269	0.0298	0.0685	0.1118	0.1679	0.2716	0.3461
Model 2												
Classic	0.1640	0.1279	0.1631	0.2442	0.4417	0.6190	0.3744	0.4560	0.6217	0.8501	1.3030	1.6448
Robust	0.1018	0.1089	0.1473	0.2064	0.3297	0.4311	0.0394	0.0669	0.1034	0.1528	0.2474	0.3152
Model 3												
Classic	1.3094	1.2203	1.1497	1.0904	1.0209	0.9931	2.1632	2.1307	2.1254	2.1454	2.2356	2.3170
Robust	0.6167	0.6405	0.6767	0.7213	0.7938	0.8367	0.0519	0.1131	0.1655	0.2109	0.2725	0.3057
Model 4												
Classic	8.0412	8.0930	8.1701	8.2778	8.4516	8.5472	12.8626	13.0311	13.2564	13.5063	13.9106	14.1228
Robust	3.5883	3.7202	3.8852	4.0817	4.4061	4.6041	0.0557	0.1308	0.2028	0.2763	0.3863	0.4468
Gamma distribution												
Model 1												
Classic	0.2813	0.1570	0.1203	0.1048	0.0966	0.0927	1.3256	0.6772	0.4724	0.3748	0.3227	0.2938
Robust	0.3183	0.1698	0.1241	0.1045	0.0946	0.0896	0.2693	0.1425	0.1062	0.0940	0.0941	0.1023
Model 2												
Classic	0.8501	0.1626	0.1228	0.1089	0.1004	0.0953	2.3528	0.7093	0.4487	0.3469	0.2969	0.2681
Robust	0.6610	0.1803	0.1261	0.1098	0.0998	0.0939	0.3639	0.1428	0.1130	0.1143	0.1192	0.1267
Model 3												
Classic	0.5513	0.8855	1.3340	1.6147	1.8183	2.0093	1.9356	1.8544	2.3800	2.7347	3.0083	3.2867
Robust	0.4404	0.5739	0.7623	0.8931	1.0037	1.1248	0.2813	0.2239	0.1900	0.1663	0.1454	0.1270
Model 4												
Classic	4.9489	7.6274	8.4630	9.1127	9.7395	10.4095	8.7237	12.5101	13.6815	14.6295	15.5769	16.6415
Robust	2.5862	3.6174	3.9768	4.3167	4.7097	5.1889	0.5141	0.2969	0.2233	0.1815	0.1535	0.1318

fact that the best performing estimator is robust one, when diffuse outlier disturbances is present for large sample sizes.

(iii) For the scale estimation, the interquartile range estimator performs for large sample size better than the small sample size, especially in contamination by extreme outliers for all considered distributions.

(iv) In the presence of outliers, the classic scale estimator has the highest RMSE of all skewed distributions.

(v) For three skewed distributions, the robust scale estimator has a lower RMSE than the classical in all contaminated cases considered. So it is seen that the robust estimator is more efficient than the classic estimator.

Determination of control charts constants

The constants d_2, d_3 and d_4 are considered under non-normality to correct the control chart limits. The corrected constants are determined such that the expected value of the statistic divided by the constant is equal to the true value of σ. The WV method constants d_2^* and d_3^* were calculated by taking the mean and standard deviation of range $\left(\frac{R}{\sigma}\right)$, respectively. In this study, we consider the MS and MWV methods constants d_2^Q and d_3^Q which are calculated by taking the mean and standard deviation of interquartile range $\left(\frac{\text{IQR}}{\sigma}\right)$, respectively. The SC method constant d_4^* is calculated by using Eq. (2.18). We consider the MSC method constant d_4^Q which is calculated by using Eq. (2.21).

In this paper, these constants based on the classic and robust estimators are obtain via simulation for each skewed distribution, because of the difficulty of numerical integration. These all constants are obtained for three skewed distributions via simulation. We obtain $E(\bar{\text{IQR}})$ by simulation: we generate 100,000 times k samples of size n, compute IQR for each instance and take the average of the values. The results of the constants for the Shewhart, WV and SC methods are presented in Table 3 for $k = 30$ and $n = 5, 10$. Moreover, the results of the constants for the MS, MWV and MSC methods are presented in Table 4 for $k = 30$ and $n = 5, 10$.

Table 3 Values of constants for the skewed distributions

k_3	Weibull			Lognormal			Gamma		
	d_2^*	d_3^*	d_4^*	d_2^*	d_3^*	d_4^*	d_2^*	d_3^*	d_4^*
$n = 5$									
0.50	2.3088	0.8493	0.5553	2.3092	0.8948	0.6919	2.3089	0.8889	0.6738
1.00	2.2559	0.9377	0.8193	2.2575	0.9843	0.9825	2.2595	0.9629	0.9108
1.50	2.1702	1.0690	1.0564	2.1974	1.0771	1.1630	2.1827	1.0661	1.0873
2.00	2.0831	1.1859	1.1998	2.1346	1.1644	1.2666	2.0827	1.1852	1.1991
2.50	1.9903	1.2950	1.2955	2.0423	1.2765	1.3456	1.9758	1.3023	1.2729
3.00	1.9102	1.3822	1.3501	1.9911	1.3315	1.3675	1.8621	1.4120	1.3275
$n = 10$									
0.50	3.0213	0.7667	0.5034	3.0640	0.8442	0.6495	3.0587	0.8335	0.6241
1.00	2.9709	0.8902	0.7733	3.0225	0.9786	0.9696	3.0050	0.9374	0.8780
1.50	2.8990	1.0706	0.9946	2.9701	1.1162	1.1463	2.9258	1.0759	1.0360
2.00	2.8301	1.2342	1.1294	2.9145	1.2490	1.2502	2.8287	1.2337	1.1303
2.50	2.7530	1.3933	1.2305	2.8300	1.4182	1.3278	2.7323	1.3870	1.1928
3.00	2.6842	1.5247	1.2943	2.7806	1.5061	1.3525	2.6348	1.5346	1.2445

Table 4 Values of robust constants for the skewed distributions

k_3	Weibull			Lognormal			Gamma		
	d_2^Q	d_3^Q	d_4^Q	d_2^Q	d_3^Q	d_4^Q	d_2^Q	d_3^Q	d_4^Q
$n = 5$									
0.50	1.3332	0.5686	0.7739	1.3094	0.5682	0.8198	1.3122	0.5684	0.8144
1.00	1.2921	0.5918	0.8998	1.2665	0.5860	0.9534	1.2773	0.5880	0.9326
1.50	1.2201	0.6244	1.0481	1.2166	0.6029	1.0670	1.2218	0.6161	1.0529
2.00	1.1459	0.6504	1.1545	1.1656	0.6177	1.1583	1.1457	0.6502	1.1552
2.50	1.0672	0.6709	1.2399	1.0932	0.6346	1.2500	1.0587	0.6821	1.2339
3.00	1.0003	0.6846	1.2970	1.0539	0.6416	1.2868	0.9635	0.7088	1.2968
$n = 10$									
0.50	1.3415	0.4825	0.5630	1.2928	0.4733	0.6166	1.2982	0.4747	0.6129
1.00	1.2892	0.4920	0.6993	1.2350	0.4756	0.7484	1.2572	0.4829	0.7313
1.50	1.1929	0.5037	0.8752	1.1698	0.4760	0.8701	1.1893	0.4948	0.8742
2.00	1.0930	0.5096	1.0089	1.1030	0.4738	0.9702	1.0926	0.5092	1.0098
2.50	0.9875	0.5084	1.1177	1.0113	0.4665	1.0768	0.9786	0.5218	1.1172
3.00	0.8980	0.5027	1.1923	0.9617	0.4608	1.1237	0.8504	0.5284	1.2127

The performance of the modified methods

When the parameters of the process are unknown, control charts can be applied in a two-phase procedure. In Phase I, control charts are used to define the in-control state of the process and to assess process stability for ensuring that the reference sample is representative of the process. The parameters of the process are estimated from Phase I sample, and control limits are estimated for using in Phase II. In Phase II, samples from the process are prospectively monitored for departures from the in-control state. The p indicates the probability of a subgroup range falling outside the control limits. The ARL is the number of points plotted within the control limits before one exceeds the limits. The ARL is the most common measure of control chart performance, and much of it is popularity is due to it is intuitively appealing and more widely applicable.

In the process control, the R, S and S^2 control charts are widely used tools to monitor process variability. Let X_{ij}, $i = 1, 2, 3, \ldots$ and $j = 1, \ldots, n$ denote independent random samples of size n taken in sequence on the process variable of interest; let $\hat{\sigma}_i$ denote an estimate of the process standard deviation σ based on the ith sample. The control limits are

$$\hat{\mathrm{UCL}} = U_n\hat{\sigma} \qquad \hat{\mathrm{LCL}} = L_n\hat{\sigma}$$

where U_n and L_n are chosen based on the skewness for this study so that the desired control chart limits are constructed when the process is in control. When the $\hat{\sigma}_i$ falls with in the control limits, the process is called in control. Let E_i denote the event that the ith sample standard deviation is beyond the limits. Further, denote by $P(Ei|\hat{\sigma})$ the conditional probability that is given for $\hat{\sigma}$; the sample standard deviation $\hat{\sigma}_i$ is beyond the control limits

$$P(Ei|\hat{\sigma}) = P(\hat{\sigma}_i < \mathrm{LCL} \quad or \quad \hat{\sigma}_i > \mathrm{UCL}|\hat{\sigma}) \tag{3.1}$$

The RL as the run length is the number of subgroups until the first $\hat{\sigma}_i$ falls beyond the limits. Given $\hat{\sigma}$, when the E_s and $E_t(s = t)$ are independent, and therefore, the distribution of the run length is geometric with parameter $P(Ei|\hat{\sigma})$. The mean of the geometric distribution is given by $1 / p$. Consequently, the conditional ARL is given by

$$E(RL|\hat{\sigma}) = \frac{1}{P(Ei|\hat{\sigma})} \tag{3.2}$$

Table 5 Results of p and ARL values for the R control chart under the skewed distributions for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Weibull												
Shewhart	0.0090	0.0110	0.0136	0.0157	0.0177	0.0191	150.8978	98.1740	72.3275	59.6196	53.4702	47.9846
WV	0.0077	0.0084	0.0096	0.0107	0.0118	0.0125	181.8512	132.5381	102.8912	88.8415	80.8865	72.4061
SC	0.0023	0.0028	0.0036	0.0045	0.0053	0.0064	427.7160	351.2469	274.0477	223.9642	190.4399	156.0549
Lognormal												
Shewhart	0.0084	0.0127	0.0157	0.0177	0.0195	0.0204	119.0660	78.9634	63.8602	56.6409	51.2424	49.1268
WV	0.0073	0.0102	0.0122	0.0134	0.0145	0.0151	136.2565	98.2000	82.1112	74.4801	68.8478	66.4143
SC	0.0028	0.0041	0.0049	0.0055	0.0065	0.0071	350.9141	243.1315	202.5481	181.6233	153.6098	139.9835
Gamma												
Shewhart	0.0051	0.0090	0.0143	0.0166	0.0183	0.0181	197.3515	111.4107	70.1326	60.1030	54.5786	55.2337
WV	0.0042	0.0066	0.0099	0.0112	0.0121	0.0117	240.1825	150.7409	100.6765	89.3152	82.8363	85.5498
SC	0.0014	0.0023	0.0036	0.0044	0.0053	0.0054	715.5123	429.7194	274.0777	226.0551	189.4191	186.1851
$n = 10$												
Weibull												
Shewhart	0.0062	0.0094	0.0124	0.0151	0.0169	0.0183	160.6942	106.6553	80.4959	66.0153	59.0493	54.6209
WV	0.0052	0.0069	0.0085	0.0099	0.0108	0.0116	192.9012	144.7807	117.8134	101.0816	92.3702	86.3707
SC	0.0035	0.0039	0.0041	0.0044	0.0048	0.0053	283.9296	259.4707	243.6647	229.2526	208.3333	188.4659
Lognormal												
Shewhart	0.0080	0.0122	0.0150	0.0169	0.0186	0.0194	125.7182	82.1970	66.7111	59.3246	53.6579	51.5730
WV	0.0069	0.0097	0.0115	0.0127	0.0137	0.0141	145.4376	102.7485	86.6371	78.9011	72.9442	70.9441
SC	0.0040	0.0054	0.0058	0.0058	0.0062	0.0066	252.2513	185.4599	171.0864	171.8715	161.6501	150.9799
Gamma												
Shewhart	0.0040	0.0076	0.0127	0.0148	0.0168	0.0171	252.8957	131.0547	78.9004	67.5466	59.5334	58.5888
WV	0.0032	0.0055	0.0086	0.0097	0.0107	0.0107	312.2171	180.9431	115.7823	103.4148	93.3315	93.8183
SC	0.0019	0.0031	0.0045	0.0043	0.0048	0.0051	539.2580	327.1502	223.8489	234.6702	207.3183	196.7420

When the standard deviation is estimated, the conditional runlength—the run length given an estiamte of σ—has a geometric distribution. However, the unconditional run length distribution the run length distribution averaged overall possible values of the estimated σ—is not geometric [20].

In contrast with the conditional RL distirbution, the marginal RL distribution takes into account the random variability introduced into the charting procedure through parameter estimation. It can be obtained by averaging the conditional RL distribution over all possible values of the parameter estimates. The unconditional p and unconditional average run length are given in [19] as, respectively

$$p = E(P(Ei|\hat{\sigma})) \tag{3.3}$$

$$\text{ARL} = E(E(\text{RL}|\hat{\sigma})) = E\left(\frac{1}{P(Ei|\hat{\sigma})}\right). \tag{3.4}$$

These expectations are simulated by generating 10,000 times k data samples of size n: numerous datasets are generated from the contaminated skewed distributions and computing for each data set the conditional value $(Ei|\hat{\sigma})$. By averaging these values, we obtain the unconditional values over the data sets. Note that for the calculation of the control limits in Phase I the process is considered to be in control [18].

In this section, we consider design schemes for the R control chart for non-contaminated and contaminated skewed distributed data. We use the mean and the trimmed mean estimators of mean and the range and the interquartile range estimators of the standard deviation for considered methods. To evaluate the control chart performance, we obtain p and ARL for moderate sample size (30 subgroups of 3–10) for each skewed distribution. Control charts can be applied in a two-stage procedure, when the parameters of a quality characteristic of the process are unknown. In Phase I, control charts are used to study a historical data set and determine

Table 6 Results of p and ARL values for the R control chart under the contaminated skewed distributions for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Weibull												
Shewhart	0.3638	0.2773	0.1483	0.0537	0.0101	0.0016	0.0003	0.0004	0.0007	0.0019	0.0099	0.0631
WV	0.7537	0.6602	0.6187	0.6005	0.3984	0.0357	0.0001	0.0002	0.0002	0.0002	0.0003	0.0028
SC	0.8663	0.6513	0.1025	0.0006	0.0000	0.0000	0.0001	0.0002	0.0010	0.1595	9.9010	9.0909
Lognormal												
Shewhart	0.3569	0.2587	0.0935	0.0198	0.0014	0.0002	0.0000	0.0000	0.0000	0.0001	0.0007	0.0052
WV	0.7537	0.6283	0.5957	0.3833	0.0062	0.0001	0.0000	0.0000	0.0000	0.0000	0.0002	0.0179
SC	0.7775	0.4556	0.1713	0.0175	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	1.2500	0.3333
Gamma												
Shewhart	0.2829	0.2531	0.1458	0.0544	0.0126	0.0018	0.0000	0.0000	0.0001	0.0002	0.0008	0.0056
WV	0.6596	0.6300	0.6142	0.6004	0.5560	0.2455	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SC	0.3019	0.3242	0.1533	0.0007	0.0001	0.0000	0.0000	0.0000	0.0001	0.0152	0.1420	1.4493
$n = 10$												
Weibull												
Shewhart	0.4183	0.3774	0.0004	0.0000	0.0000	0	0.0000	0.0000	0.0245	0.7752	1.5873	Inf
WV	0.3715	0.3687	0.3686	0.3687	0.3668	0.3175	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SC	0.3827	0.3691	0.3686	0.3594	0.2119	0.0051	0.0000	0.0000	0.0000	0.0000	0.0000	0.0019
Lognormal												
Shewhart	0.4050	0.0033	0.0000	0.0000	0.0000	0.0000	0.0000	0.0030	0.2833	Inf	8.3333	2.0833
WV	0.3689	0.3686	0.3686	0.3590	0.1266	0.0075	0.0000	0.0000	0.0000	0.0000	0.0001	0.0013
SC	0.3713	0.3687	0.3658	0.3292	0.1363	0.0268	0.0000	0.0000	0.0000	0.0000	0.0001	0.0004
Gamma												
Shewhart	0.3292	0.0404	0.0004	0.0000	0.0000	0.0000	0.0003	0.0025	0.2684	5.1546	Inf	Inf
WV	0.3690	0.3687	0.3687	0.3685	0.3686	0.3686	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
SC	0.3687	0.3684	0.3679	0.3592	0.2221	0.0000	0.0003	0.0003	0.0003	0.0003	0.0005	Inf

*ARL values of Weibull multiplies with $1.0e + 04$ for $n = 5$, $1.0e + 05$ for $n = 10$

*ARL values of lognormal multiplies with $1.0e + 6$ for $n = 5$ and $1.0e + 05$ for $n = 10$

*ARL values of gamma multiplies with $1.0e + 05$ for $n = 5$ and $1.0e + 04$ for $n = 10$

the samples that are out of control. On the basis of the resulting reference sample, the process parameters are estimated and control limits are calculated for Phase II. In Phase II, control charts are used for real-time process monitoring [21].

The simulation consists of two phases is run by using MATLAB R2013. The steps of each phase are described as follows.

Phase I:

1.a. Generate n i.i.d. Weibull $(\beta, 1)$, gamma $(\alpha, 1)$ and log-normal $(0, \sigma)$ varieties for $n = 3, 5, 7, 10$.
1.b. Repeat step 1.a 30 times ($k = 30$).
1.c. By using classic estimators, compute the control limits for Shewhart, the WV and the SC methods. By using robust estimators, compute the control limits for the MS, the MWV and the MSC methods.

Phase II:

2.a. Generate n i.i.d. Weibull $(\beta, 1)$, gamma $(\alpha, 1)$ and log-normal $(0, \sigma)$ varieties using the procedure of step 1.a.
2.b. Repeat step 2.a 100 times ($k = 100$).
2.c. Compute the sample statistics for R chart for the Shewhart, WV and SC methods. Compute the robust estimator interquartile range IQR for the MS, MWV and MSC methods.
2.d. Record whether or not the sample statistics calculated in step 2.c are within the control limits of step 1.c. for all methods.
2.e. Repeat steps 1.a through 2.d, 100.000 times and obtain p and ARL values for each method.

In the simulation study, we consider non-contaminated and contaminated data set in Phases I and II. We consider the 20% trimmed mean, which trims the six smallest and the six largest sample trimmed means when $k = 30$.

Table 7 Results of p and ARL values for the modified R control chart under the skewed distributions for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Weibull												
MS	0.0068	0.0101	0.0140	0.0166	0.0190	0.0207	110.8697	91.0349	73.5456	63.6217	56.3784	52.3714
MWV	0.0056	0.0074	0.0097	0.0111	0.0125	0.0135	129.1222	118.8241	103.7151	93.4588	84.5559	79.7550
MSC	0.0023	0.0027	0.0034	0.0040	0.0048	0.0054	441.3647	370.3978	295.8667	249.5882	207.9910	185.1269
Lognormal												
MS	0.0097	0.0119	0.0140	0.0156	0.0174	0.0183	103.0460	84.1085	71.2550	64.1420	57.3076	54.6397
MWV	0.0086	0.0096	0.0109	0.0117	0.0128	0.0133	116.2601	104.1873	92.1328	85.2123	77.9460	75.2417
MSC	0.0025	0.0031	0.0037	0.0043	0.0050	0.0054	402.9496	321.8228	266.9585	231.3958	199.0129	183.6311
Gamma												
MS	0.0096	0.0115	0.0136	0.0158	0.0176	0.0195	104.0832	86.7965	73.3003	63.3136	56.7032	51.1593
MWV	0.0085	0.0091	0.0099	0.0108	0.0114	0.0122	118.3124	109.7020	100.6867	92.8902	87.9825	82.0371
MSC	0.0025	0.0029	0.0035	0.0041	0.0046	0.0055	406.3389	342.9120	289.1093	245.2784	215.5172	183.3954
$n = 10$												
Weibull												
MS	0.0065	0.0083	0.0107	0.0129	0.0149	0.0163	154.6671	120.3891	93.1359	77.2630	67.2079	61.2985
MWV	0.0058	0.0069	0.0084	0.0098	0.0110	0.0119	171.5031	144.8834	118.9061	102.0169	90.9579	84.0555
MSC	0.0022	0.0026	0.0030	0.0034	0.0039	0.0043	454.5661	387.3717	332.7123	293.6082	257.9580	230.4466
Lognormal												
MS	0.0072	0.0091	0.0108	0.0122	0.0141	0.0149	139.3942	109.9989	93.0103	81.9618	71.1187	67.1686
MWV	0.0066	0.0078	0.0090	0.0099	0.0112	0.0118	151.5129	127.4551	111.7069	100.5935	89.1353	84.8097
MSC	0.0024	0.0029	0.0033	0.0037	0.0041	0.0044	418.7079	342.8415	299.8950	268.6583	243.4808	229.8586
Gamma												
MS	0.0072	0.0087	0.0108	0.0128	0.0149	0.0170	139.8562	114.4505	92.8402	78.0616	66.9035	58.7492
MWV	0.0066	0.0074	0.0086	0.0097	0.0108	0.0118	152.3693	134.6928	116.2966	103.2439	92.7756	84.7214
MSC	0.0024	0.0027	0.0031	0.0033	0.0038	0.0045	418.3750	367.4984	324.8019	299.8591	261.7116	223.7086

- **Non-contaminated case:** The reference distribution parameters are selected with respect to skewness of distribution given in Table 1.
- **Contaminated case:** The more extreme case of 10% of outliers placed at 50. We consider the contamination in Phases I and II.

The simulation results of p and ARL for the R control chart for non-contaminated data under skewed distributions are given in Tables 5 and 7. The results of p and ARL for the R control chart for contaminated Weibull, lognormal and gamma distrubuted data are given in Tables 8, 9 and 10, respectively.

Results

In this section, the performance of design schemes is evaluated. When the process in control, it is expected that p is to be as low as possible and ARL is to be as high as possible. First we consider the design scheme where the process follows skewed distribution and the Phase I data are non-contaminated. Tables 5 and 7 present the p and the ARL values for the R control chart based on classic and robust estimators under the skewed distributions. The tables shows that :

- The results for the uncontaminated case based on classic estimator are given in Table 5 as follows:

  When the distribution is approximately symmetric ($k_3 = 0.5$), then the p of SC, WV and Shewhart method are comparable, while the SC method has a noticeable smaller p values. When the skewness increases, the ARL values decrease for all design schemes while the

Table 8 Results of p and ARL values for the R control chart for contaminated Weibull distrubuted data for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Model 1												
MS	0.0073	0.0085	0.0103	0.0119	0.0135	0.0147	137.6349	118.2033	96.6828	84.2950	74.0351	68.2058
MWV	0.0063	0.0064	0.0072	0.0078	0.0086	0.0092	159.7393	155.6735	139.6980	127.8462	115.6109	108.1210
MSC	0.0016	0.0017	0.0021	0.0025	0.0030	0.0035	631.7918	572.4098	473.1040	399.7921	332.4468	288.5087
Model 2												
MS	0.1577	0.1276	0.0835	0.0494	0.0297	0.0220	6.3413	7.8361	11.9811	20.2308	33.6828	45.4339
MWV	0.2869	0.1092	0.0454	0.0236	0.0143	0.0111	3.4851	9.1608	22.0294	42.4196	69.9208	89.9612
MSC	0.0034	0.0045	0.0056	0.0061	0.0068	0.0076	294.1176	222.2222	178.5714	163.9344	147.0588	131.5789
Model 3*												
MS	0.2591	0.1619	0.0616	0.0217	0.0092	0.0063	0.0039	0.0062	0.0162	0.0461	0.1083	0.1596
MWV	0.6400	0.6168	0.5944	0.4955	0.2309	0.0315	0.0016	0.0016	0.0017	0.0020	0.0043	0.0317
MSC	0.0033	0.0041	0.0054	0.0061	0.0071	0.0076	303.0303	243.9024	185.1852	163.9344	140.8451	131.5789
$n = 10$												
Model 1												
MS	0.0052	0.0065	0.0084	0.0102	0.0119	0.0132	190.8543	152.9847	119.2521	98.3362	84.3633	75.7714
MWV	0.0047	0.0054	0.0065	0.0075	0.0085	0.0093	211.7164	185.8322	154.4879	132.9770	117.0412	107.0652
MSC	0.0016	0.0018	0.0020	0.0022	0.0026	0.0030	607.2014	551.0856	498.0080	446.3489	385.8769	335.7057
Model 2												
MS	0.0088	0.0117	0.0153	0.0177	0.0198	0.0211	114.1657	85.3446	65.3104	56.6200	50.6201	47.4404
MWV	0.0076	0.0095	0.0118	0.0133	0.0145	0.0153	132.4468	105.5086	84.5931	75.2950	68.7753	65.1831
MSC	0.0033	0.0042	0.0050	0.0056	0.0062	0.0066	304.5995	237.2198	198.2043	179.3207	162.3034	151.6139
Model 3												
MS	0.0084	0.0117	0.0156	0.0185	0.0210	0.0228	118.8736	85.4672	64.1997	53.9476	47.5563	43.8416
MWV	0.0072	0.0094	0.0119	0.0139	0.0155	0.0166	139.2641	106.7954	83.7318	72.1511	64.6914	60.4186
MSC	0.0032	0.0042	0.0052	0.0059	0.0068	0.0075	316.3456	239.8772	191.7068	168.1775	146.5159	133.5381

*ARL values multiplies with $1.0e + 03$

ARL values of the Shewhart chart decrase too much and are quite lower than others. The ARL values based on Shewhart and WV methods are lower than the SC method. So the SC method performs better than the others, especially for skewness. According to the p and ARL values, there is no difference between the Weibull, gamma and lognormal distributions. It is seen from the results, in the case of skewness, the Shewhart charts does not perform well any more. So we can recommend to use asymmetric control charts based on WV and SC methods (see more details in [13]).

– The results for the contaminated case based on classic estimator are given in Table 6 as follows:

When we consider the contamination in the skewed distributed data, the WV and SC are effected so much from the outliers. So control charts based on WV and SC methods do not perform well any more. So we reccomend to use asymmetric control charts based on robust estimator.

– The results for the uncontaminated case based on robust estimator are given in Table 7 as follows:

As the skewness increases, the MWV method gives better results than the MS, MSC gives better results than the MS and WV. The MSC method works very well for all skewed distributions for small and large sample sizes for all skewed distributions, except gamma distribution for $n = 10$.

When the skewed data are uncontaminated, the performance of the control charts based on WV and SC methods using classic estimators is comparable with the modified control charts based on MWV and MSC methods using robust estimators.

– The results for the contaminated case based on robust estimator are given in Tables 8, 9 and 10 for Weibull, lognormal and gamma distributed data, respectively, as follows:

The p values for MSC method for gamma distribution are increasing when the number of the sample size

Table 9 Results of p and ARL for the R control chart for contaminated lognormal distrubuted data for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Model 1*												
MS	0.2332	0.0928	0.0355	0.0192	0.0134	0.0126	4.2883	10.7794	28.1769	52.0115	74.4718	79.6292
MWV	0.6302	0.5951	0.4382	0.1759	0.0108	0.0041	1.5867	1.6803	2.2818	5.6865	92.3898	244.5287
MSC	0.1151	0.0156	0.0037	0.0017	0.0019	0.0022	8.6847	64.1355	270.2995	588.0969	532.0847	452.7755
Model 2												
MS	0.2342	0.0918	0.0356	0.0193	0.0135	0.0126	4.2697	10.8977	28.0793	51.6994	74.0516	79.3789
MWV	0.6307	0.5954	0.4379	0.1757	0.0107	0.0041	1.5856	1.6796	2.2837	5.6921	93.4789	243.3623
MSC	0.1156	0.0155	0.0037	0.0017	0.0019	0.0022	8.6472	64.3488	266.8446	577.1673	528.3737	452.2840
Model 3												
MS	0.2339	0.0925	0.0353	0.0192	0.0134	0.0126	4.2753	10.8137	28.3280	52.0869	74.3732	79.4407
MWV	0.6305	0.5953	0.4379	0.1763	0.0108	0.0041	1.5860	1.6799	2.2835	5.6721	92.9964	242.6007
MSC	0.1154	0.0156	0.0037	0.0017	0.0019	0.0022	8.6678	64.0541	269.0125	587.8895	526.2327	455.1869
$n = 10$												
Model 1												
MS	0.0066	0.0084	0.0104	0.0124	0.0145	0.0157	151.0688	119.0023	95.7781	80.9042	68.8463	63.7357
MWV	0.0064	0.0074	0.0088	0.0102	0.0117	0.0125	156.2891	135.3034	113.8330	98.4581	85.6362	80.0109
MSC	0.0021	0.0026	0.0033	0.0038	0.0045	0.0048	467.3334	380.7058	307.4274	264.3614	224.6383	209.0957
Model 2												
MS	0.0096	0.0133	0.0159	0.0177	0.0193	0.0197	103.7700	75.4484	62.8460	56.5294	51.7582	50.8536
MWV	0.0085	0.0112	0.0130	0.0142	0.0152	0.0154	117.5834	89.6781	76.6554	70.1887	65.5832	64.9642
SC	0.0037	0.0050	0.0060	0.0065	0.0066	0.0065	273.8901	199.0723	167.1682	154.9571	150.6546	152.8608
Model 3												
MS	0.0097	0.0137	0.0170	0.0195	0.0223	0.0237	102.7485	72.8444	58.8433	51.1679	44.7507	42.2060
MWV	0.0085	0.0114	0.0137	0.0156	0.0175	0.0185	118.2033	87.9245	72.9325	64.2731	57.0145	54.0862
MSC	0.0037	0.0053	0.0065	0.0074	0.0084	0.0089	272.4796	189.8686	153.4425	134.7600	118.5354	111.9545

Table 10 Results of p and ARL for the R control chart for contaminated gamma distrubuted data for $n = 5, 10$

Method/k_3	p Values						ARL values					
	0.5	1.0	1.5	2	2.5	3.0	0.5	1.0	1.5	2	2.5	3.0
$n = 5$												
Model 1*												
MS	0.0126	0.0298	0.0337	0.0218	0.0096	0.0032	0.0080	0.0034	0.0030	0.0046	0.0105	0.0317
MWV	0.0042	0.2180	0.4449	0.4953	0.4653	0.3109	0.0237	0.0005	0.0002	0.0002	0.0002	0.0003
MSC	0.0023	0.0033	0.0015	0.0005	0.0001	0.0000	0.0434	0.0304	0.0646	0.2147	1.0460	8.4746
Model 2												
MS	0.0070	0.0108	0.0149	0.0175	0.0199	0.0215	142.3751	92.7154	67.3260	57.0002	50.2535	46.6105
MWV	0.0073	0.0085	0.0106	0.0118	0.0129	0.0134	136.3066	117.6277	93.9541	84.4466	77.6108	74.6001
MSC	0.0015	0.0027	0.0039	0.0048	0.0056	0.0063	676.1782	375.6574	257.1818	208.6202	178.3930	159.2433
Model 3												
MS	0.0099	0.0227	0.0394	0.0492	0.0506	0.0463	100.5257	44.1117	25.3756	20.3405	19.7551	21.5964
MWV	0.0096	0.0157	0.0223	0.0234	0.0206	0.0160	103.7990	63.6821	44.7950	42.8016	48.5765	62.6586
MSC	0.0025	0.0069	0.0107	0.0112	0.0097	0.0072	394.7888	144.8792	93.8738	89.2746	103.1481	139.5615
$n = 10$												
Model 1												
MS	0.0065	0.0077	0.0101	0.0137	0.0169	0.0188	153.6523	129.8634	99.3631	72.8810	59.2466	53.0729
MWV	0.0063	0.0067	0.0081	0.0103	0.0122	0.0132	159.9488	149.1892	123.9572	96.7455	82.1450	75.6630
MSC	0.0021	0.0023	0.0028	0.0037	0.0046	0.0053	475.3078	433.3131	354.9624	269.2877	216.7458	189.5735
Model 2												
MS	0.0070	0.0119	0.0154	0.0178	0.0194	0.0208	142.6595	84.3398	65.0830	56.2506	51.4279	48.1621
MWV	0.0066	0.0099	0.0121	0.0134	0.0140	0.0143	150.8978	100.9000	82.3621	74.7138	71.3353	69.8944
MSC	0.0023	0.0043	0.0053	0.0056	0.0059	0.0062	427.5514	233.7923	189.9407	177.8126	168.2737	160.1640
Model 3												
MS	0.0096	0.0128	0.0159	0.0185	0.0203	0.0217	103.8217	77.9065	62.7668	54.0728	49.3167	46.0259
MWV	0.0083	0.0105	0.0124	0.0139	0.0145	0.0151	120.0999	95.4399	80.4117	72.1975	68.9280	66.3883
MSC	0.0036	0.0048	0.0056	0.0060	0.0063	0.0067	276.7323	210.1370	180.0731	167.6868	158.6169	149.2270

*ARL values multiplies with $1.0e + 04$

is increasing. So the modified models perform well for large size, except MSC for gamma distribution. MSC method has the lowest p values and the highest ARL values for all skewed distributions in all designs. So this modified method has the best performance in the case of contamination in Phases I and II for the skewed data .

When the simulation program is run for $n = 25$, the results are the same results as $n = 10$. So we can say that the results are same for large sample size.

We investigate the effect of non-normality on estimated limits under the contamination. The SC and MSC methods have the best performance for all design schemes, especially in the case of skewness.

Conclusion

Control charts are known to be effective tools for monitoring the quality of process and are applied in many industries. In this study, we consider the non-normality and the contamination for the R control charts. We propose to use the interquartile range estimator of the standard deviation to modify the methods. We study the effect of the estimator on control chart performance under non-normality for moderate sample size (30 subgroups of 5–10). To evaluate the control chart performance, we obtain p and ARL values of this control charts and the results used to compare the methods. We consider the design schemes where the Phase I and the Phase II data are non-contaminated and contaminated. The results are: The Shewhart chart has the worst performance for all design schemes, since the p values of the Shewhart chart are quite higher than others. As the skewness increases, the p values of the Shewhart chart increase too much and are effected by skewness. So the asymmetric control charts based on WV and SC methods can be used in the case of skewness. When the skewed data are uncontaminated, the performance of the control charts based on WV and SC methods using classic estimators is comparable with the modified control charts based on MWV and MSC methods using robust estimators. When there is no contamination, the SC and MSC methods work very well for all skewed distributions for small and large sample sizes for all skewed distributions. However, in the case of contamination, control charts based on WV and SC methods do not perform well any more. The MSC method has the lowest p values and the highest ARL values for all skewed distributions under contamination and so has the best performance. We reccomend to use asymmetric control charts based on MSC method for the skewed data in the case of contamination in Phases I and II.

As a future research, the proposed control chart can be extended using some other sampling schemes such as repetitive sampling, multiple dependent state sampling, ranked set sampling and neoteric rank set sampling.

As another future research, it is possible to consider other skewed distributions as heavy-tailed distributions.

Funding This study was not funded.

Compliance with ethical standards

Conflict of interest Derya Karagöz declares that she has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Abu-Shawiesh, M.O.A.: A simple robust control chart based on MAD. J. Math. Stat. **4**(2), 102–107 (2008)
2. Bai, D.S., Choi, I.S.: $\bar{X}$ and R control charts for skewed populations. J. Qual. Technol. **27**, 120–131 (1995)
3. Chan, L.K., Cui, H.J.: Skewness correction $\bar{X}$ and R charts for skewed distributions. Nav. Res. Logist. **50**, 1–19 (2003)
4. Chang, Y.S., Bai, D.S.: Control charts for positively skewed populations with weighted standard deviations. Qual. Reliab. Eng. Int. **17**, 397–406 (2001)
5. Choobineh, F., Branting, D.: A simple approximation for semivariance. Eur. J. Oper. Res. **27**, 364–370 (1986)
6. Choobineh, F., Ballard, J.L.: Control-limits of QC charts for skewed distributions using weighted variance. IEEE Trans. Reliab. **36**, 473–477 (1987)
7. Davis, C.M., Adams, B.M.: Robust monitoring of contaminated data. J. Qual. Technol. **37**, 163–174 (2005)
8. Duclos, E., Pillet, M.: An optimal control chart for non-normal process. In: 1^{st} IFAC Workshop on Manufacturing Systems: Manufacturing Modeling, Management and Control, Austria 12p (1997)
9. He, X., Fung, W.K.: Method of medians for lifetime data with Weibull models. Stat. Med. **18**, 1993–2009 (1999)
10. Khodabin, M., Ahmadabadi, A.: Some properties of generalized gamma distribution. Math. Sci. **4**(1), 9–28 (2010)
11. Huang, W.-H., Yeh, A.B., Wang, H.: A control chart for the lognormal standard deviation. Qual. Technol. Quant. Manag. (2017). https://doi.org/10.1080/16843703.2017.1304044
12. Jensen, W.A., Jones-Farmer, L.A., Champ, C.W., Woodall, W.H.: Effects of parameter estimations on control chart properties: a literature review. J. Qual. Technol. **38**(4), 349–364 (2006)
13. Karagöz, D., Hamurkaroğlu, C.: Control charts for skewed distributions: Weibull, Gamma, and Lognormal. Adv. Methodol. Stat. **9**(2), 95–106 (2012)
14. Koyuncu, N., Karagöz, D.: New mean charts for bivariate asymmetric distributions using different ranked set sampling designs. Qual. Technol. Quant. Manag. (2017). https://doi.org/10.1080/16843703.2017.1321220
15. Montgomery, D.C.: Introduction to Statistical Quality Control. Wiley, Hoboken (1997)
16. Rocke, D.M.: Robust control charts. Technometrics **31**(2), 173–184 (1989)
17. Rocke, D.M.: $\hat{X}_Q$ and R_Q charts: robust control charts. Statistician **41**(1), 97–104 (1992)
18. Schoonhoven, M., Does, R.J.M.M.: The $\bar{X}$ control chart under non-normality. Qual. Reliab. Eng. Int. **26**, 167–176 (2010)
19. Schoonhoven, M., Riaz, M., Does, R.J.M.M.: Design and analysis of control charts for standard deviation with estimated parameters. J. Qual. Technol. **43**, 307–333 (2011)
20. Schoonhoven, M., Does, R.J.M.M.: A robust standard deviation control chart. Technometrics **54**(1), 73–82 (2012)
21. Schoonhoven, M., Does, R.J.M.M.: A robust $\bar{X}$ control chart. Qual. Reliab. Eng. Int. **29**, 951–970 (2013)
22. Sindhumol, M.R., Srinivasan, M.R., Gallo, M.: A robust dispersion control chart based on modified trimmed standard deviation. Electron. J. Appl. Stat. Anal. **9**(1), 111–121 (2016)
23. Sukparungsee, S.: Asymmetric Tukey's control chart robust to skew and non-skew process observation. World Acad. Sci. Eng. Technol. Int. J. Math. Comput. Sci. **7**(8), 1275–1278 (2013)
24. Tang, S., Yeh, A.B.: Approximate confidence intervals for the lognormal standard deviation. Qual. Reliab. Eng. Int. **32**, 715–725 (2016)
25. Tatum, L.G.: Robust estimation of the process standard deviation for control charts. Technometrics **39**, 127–141 (1997)
26. Vargas, J.A.: Robust estimation in multivariate control charts for individual observations. J. Qual. Technol. **35**, 367–376 (2003)