International Journal for Quality Research 13(3) 641–654
ISSN 1800-6450

Moustafa Omar Ahmed AbuShawiesh[1]
Hayriye Esra Akyüz
Hatim Solayman Ahmed Migdadi
B.M. Golam Kibria

Article info:
Received 12.01.2019
Accepted 12.03.2019

UDC – 005.311.121
DOI – 10.24874/IJQR13.03-09

ROBUST ALTERNATIVES TO THE TUKEY'S CONTROL CHART FOR THE MONITORING OF THE STATISTICAL PROCESS MEAN

Abstract:*Control Charts are one of the most powerful tools used to detect aberrant behavior in industrial processes. A valid performance measure for a control chart is the average run length (ARL); which is the expected number of runs to get an out of control signal. At the same time, robust estimators are of vital importance in order to estimate population parameters. Median absolute deviation (MAD) and quantiles are such estimators for population standard deviation. In this study, alternative control charts to the Tukey control chart based on the robust estimators are proposed. To monitor the control chart's performance, the ARL values are compare for many symmetric and skewed distributions. The simulation results show that the in-control ARL values of proposed control charts are higher than Tukey's control chart in all cases and more efficient to detect the process mean. However, the out- of- control ARL values for the all control charts are worse when the probability distribution is non-normal. As a result, it is recommended to use control chart based on the estimator Q_n for the process monitoring performance when data are from normal or non-normal distribution. An application example using real-life data is provided to illustrate the proposed control charts, which also supported the results of the simulation study to some extent.*

Keywords:*Average run length (ARL), Box plot, Robust estimator, Statistical process control, Tukey's control chart*

1. Introduction

Statistical process control (SPC) charts are widely used for monitoring, measuring, controlling and improving the quality of production in many areas of application, such as, in industry and manufacturing, finance and economics, epidemiology and health care, environmental sciences and other fields (Sukparungsee, 2013). The Shewhart control chart, used for monitoring industrial processes is the most popular tool in SPC, developed under the assumption that the observations from a process are independent, identically distributed and from a normal distribution (Sindhumol et al., 2016). On the other hand, the Tukey's control chart (TCC) proposed first by Alemi (2004), based on the principles of the box plot and has no assumptions about data with unknown distribution so the TCC can be used with all probability distributions. The TCC is also applied to small data and it is robust to data with outliers. Furthermore, it is easy and simple to setup control limits because the TCC does not use the standard

Corresponding author: Moustafa Omar Ahmed AbuShawiesh
Email: mabushawiesh@hu.edu.jo

deviation or the mean to construct the control chart limits, but uses only quartiles (Mekparyup et al., 2014). The problem with the TCC is that it assigns the same weight for both upper and lower control limits, which is acceptable for symmetrical distributions but becomes a problem with asymmetric distribution (Tercero-Gomez et al. 2012).

Alemi (2004) presented the Tukey's control chart (TCC) which applied the principle of box plot to obtain the control limits. Borckardt et al. (2005) suggested that TCC does not perform well in the presence of serial dependence. However, for independent measurement, TCC performed well even for small samples of 12 observations. Torng and Lee (2008) compared the performance of the TCC with Shewhart's control chart for both symmetric and asymmetric distributions and concluded that TCC is not sensitive to shift detection when the process violates the normality assumption. Lee (2011) extended the TCC with asymmetrical control limits so called ACL-Tukey's control chart to detect a change in parameter of skew population. Tercero-Gomez et al., 2012) proposed a modified version of the TCC to consider skewness with minimum assumptions on the underlying distribution of observations. Sukparungsee (2012) studied the performance of TCC for detecting a change in parameter when observations are from skewed distributions such as exponential and Laplace. Also, the performance of TCC compared with Shewhart and Exponentially Weighted Moving Average (EWMA) charts. Sukparungsee (2013) determined the robustness of asymmetric TCC for skew and symmetric distributions such as Lognormal and Laplace distributions.

Khaliq et al. (2015) compared TCC with individual / moving range (XmR) control charts and found that TCC was more efficient than XmR in the sense of average run length, extra quadratic loss, median run length, standard deviation run length, performance comparison index, and relative average run length. Mekparyup et al. (2014) proposed the Adjusted Tukey's Control Chart (ATCC) as an improvement of the TCC by substitution median of absolute deviation from the sample median (MAD) instead of the interquartile range (IQR). They concluded that the ATCC is more efficient to detect the process when it is in control.

It is noted that the modification of TCC based on the robust estimators, namely MAD, sample quantile (S_n) and sample quantile (Q_n) are limited in literature. Therefore, in this paper, a modification of the TCC based on these robust estimators is presented as an alternative procedure with independent and identically distributed observations. These modified control charts are named as MAD-TCC, Sn-TCC and Qn-TCC.

The rest of paper is organized as follows: Section 2 discusses the robust estimators of IQR and defined Tukey's control charts. In addition to, the ARL and design of TCC are explained in Section 2. The design and the ARL for the proposed modified-TCCs are given in section 3. In Section 4, a Monte-Carlo simulation study is conducted to evaluate the performance of the proposed modified methods. In Section 5, an application example based on real life dataset is presented to illustrate the implementation of the proposed modified-TCCs. Finally, Section 6 concludes and summarizes the findings and outcomes of the paper.

2. Robust estimators and control charts

A robust estimator is an estimator which performs well in presence of outliers and when data are not drawn from a normal distribution. There are many robust estimators used for location and scale in literature, see for example, Rousseeuw and Croux (1993), Tiku and Akkaya (2004), Akyüz and Gamgam (2017), and Akyüz et al. (2017). In this study, we give the definition, properties and information about three robust estimators of scale parameter. They are the MAD and the two robust measures of scale proposed by Rousseeuw and Croux (1993) as alternatives to MAD, the first estimator is Sn and the second estimator is Q_n. The three robust estimators are used in this paper to construct the proposed modified-TCCs.

2.1. The average run length (ARL) calculation

The average run length (ARL) is the expected number of points plotted on the control chart until an out of signal indicated and it is used to evaluate the process behavior. If x_i is an observation, $P(\delta)$ represents the probability of an out-of-control process with shift size, $\delta = \frac{\mu_1 - \mu_0}{\sigma}$, where μ_0 is initial process mean, σ is population standard deviation and $\mu_1 = \mu_0 + \sigma\delta$, then $P(\delta)$ can be calculated as follows:

$$\begin{aligned} P(\delta) &= 1 - P(LCL \le x \le UCL \mid \mu = \mu_1 = \mu_0 + \delta\sigma) \\ &= 1 - P(LCL - \delta\sigma \le x \le UCL - \delta\sigma \mid \mu = \mu_0) \\ &= 1 - \int_{LCL-\delta\sigma}^{UCL-\delta\sigma} f(x)\,dx. \end{aligned}$$

In this case, $ARL(\delta)$ is obtained as follows:

$$ARL(\delta) = \frac{1}{P(\delta)}$$

The $P(\delta)$ will give the probability of Type I error (α) or the probability of the occurrence of false alarm. The larger in-control average run length (ARL_0) value shows that the probability of occurrence of false alarm is small. In the in-control process, the smaller α the larger ARL_0 value (Davis, 2004). In contrast, if the process is out of control or mean shift occurs ($\delta \neq 0$), then $P(\delta \neq 0)$ will give the power, $(1-\beta)$ corresponding with type II error probability (β) and the out-of-control average run length (ARL_1) is used to monitor the out-of-control process. The small ARL_1 value illustrates that the power of detecting the process shift. In the out-of-control process, the smaller the value of β the smaller the value of ARL_1. There is no regulation for what value of $ARL(\delta)$ should be either the process is in control or out of control. For the Shewhart's control chart for individual observations, ARL_0 = 143.338 for 3-sigma control limits with $P(\delta = 0) = 0.0027$. Wheeler and Chambers (1992) suggested that ARL_0 = 92 could be accepted when the process is in control.

2.2. The Tukey's Control Chart (TCC)

Tukey's control chart (TCC) is simple and easy to use. It has an effective charting structure that exhibits robustness for the skewed distribution. The TCC was first proposed by Alemi (2004) who applies the principles of box plot. The lower control limit (LCL) and the upper control limit (UCL) for this control chart are constructed as follows:

$$LCL = Q_1 - k\,IQR$$
$$UCL = Q_3 + k\,IQR$$

where $Q_1 = F^{-1}(0.25)$, $Q_3 = F^{-1}(0.75)$ and $IQR = Q_3 - Q_1$. Under the normal distribution with population mean μ and population standard deviation σ, the parameter k is usually set as k = 1.5 (Ryan, 2000).

2.3. The ARL Calculation for the Tukey's Control Chart

When the control limits of the TCC have been set, the in-control and out-of-control average run lengths can be calculated. Assuming that the process is in control, then the ARL_0 for the TCC can be calculated as follows:

$$ARL_0 = \frac{1}{1 - \int_{F^{-1}(0.25) - k\,IQR}^{F^{-1}(0.75) + k\,IQR} f(x)\,dx} = \frac{1}{\alpha}$$

If the process is out of control or the mean shift of $\delta\sigma$ occurs, then the ARL_1 for the TCC can be calculated as follows:

$$ARL_1 = \frac{1}{1 - \int_{F^{-1}(0.25) - (k\,IQR) - \delta\sigma}^{F^{-1}(0.75) + (k\,IQR) - \delta\sigma} f(x)\,dx} = \frac{1}{1-\beta}$$

where α is the probability of Type I error, β is the probability of Type II error, f(x) is the probability density function (pdf) and F(x) is its cumulative distribution function (CDF) of the random variable X.

3. The proposed modified Tukey's control charts

In this section, we will introduce the design and the ARL for the three modified-TCCs methods based on the three robust measures of scale as an alternative to the TCC. The proposed modified control charts are computationally simple, easy to use and therefore analytically a more desirable method. These methods are known as MAD-TCC, Sn-TCC and Qn-TCC.

3.1. The MAD-TCC control chart

The first method we propose in this paper is called the MAD-TCC control chart, which is a modification of the TCC and based on the MAD. The MAD was first introduced by Hampel (1974) who attributed it to Gauss, as a robust scale estimator alternative to the sample standard deviation (S). It is often used as an initial value for the computation of more efficient robust estimators. The MAD is widely used in various applications as an alternative to sample standard deviation (S). Abu-Shawiesh (2008), for example, used the MAD for constructing a simple robust dispersion control chart. For a random sample X1, X2, …,Xn with a sample median (MD), the MAD is defined as follows:

$$MAD = 1.4826\, MD\left\{ \left| X_i - MD \right| \right\};\ i = 1, 2, 3, \dots, n$$

The constant, 1.4826 in (6) adjusts the scale for maximum efficiency when the data comes from a normal distribution. The MAD has the highest breakdown point possible which is 50% and this value is twice as much as the IQR. Furthermore, the influence function of MAD is bounded but not smooth. The efficiency of MAD compare to IQR is 37% at normal distribution. The control limits of MAD-TCC are obtained as follows:

$$LCL = Q_1 - k\,MAD$$

$$UCL = Q_3 + k\,MAD$$

Under the normal distribution with population mean μ and population standard deviation σ, the IQR equals 1.34898 σ and the MAD equals 0.67449 σ so the MAD is twice as much as IQR, then the parameter k is set as k = 3 in order to equal the control limits of the TCC.

When the control limits of MAD-TCC have been set, the ARL_0 and ARL_1 can be calculated. Assuming that the process is in control, then the ARL_0 and ARL_1 for the MAD-TCC are obtained respectively as follows:

$$ARL_0 = \frac{1}{1 - \int_{F^{-1}(0.25) - k\,MAD}^{F^{-1}(0.75) + k\,MAD} f(x)\,dx} = \frac{1}{\alpha}$$

and

$$ARL_1 = \frac{1}{1 - \int_{F^{-1}(0.25) - (k\,MAD) - \delta\sigma}^{F^{-1}(0.75) + (k\,MAD) - \delta\sigma} f(x)\,dx} = \frac{1}{1-\beta}$$

3.2. The Sn-TCC control chart

The second method we propose is called the Sn-TCC control chart, which is a modification of the TCC. This method is based on the S_n estimator, which is proposed by Rousseeuw and Croux (1993). The S_n estimator is a powerful alternative to the

MAD. For a random sample $X_1, X_2, \ldots, X_n$ the S_n estimator is defined as follows:

$$S_n = 1.1926\, MD_i \left\{ MD_j \left| X_i - X_j \right| \right\};\ i, j = 1, 2, 3, \ldots, n$$

For each iwe compute the median of $\left\{ \left| X_i - X_j \right|;\ j = 1, 2, \ldots, n \right\}$. This yields n numbers, the median of which multiplied by the factor 1.1926 gives our final estimate S_n. The S_n has the highest breakdown possible points which is 50% and the influence function of it is also bounded. The efficiency of the S_n estimator compare to the IQRis 58.23% for normal distribution which is better than that of MAD. The control limits of Sn-TCC are obtained as follows:

$$LCL = Q_1 - k\, S_n$$

$$UCL = Q_3 + k\, S_n$$

Under the normal distribution with population mean μ and population standard deviation σ, the IQR equals 1.34898 σ and S_n equals 0.83850 σ so the S_n is 1.6 as much as IQR, then the parameter k is set as k = 2.4 in order to equal the control limits of the TCC.

When the control limits of Sn-TCC have been set, the ARL_0 and ARL_1 can be calculated. Assuming that the process is in control, the ARL_0 and ARL_1 for the Sn-TCC are obtained respectively as follows:

$$ARL_0 = \frac{1}{1 - \int_{F^{-1}(0.25) - k S_n}^{F^{-1}(0.75) + k S_n} f(x)\, dx} = \frac{1}{\alpha}$$

and

$$ARL_1 = \frac{1}{1 - \int_{F^{-1}(0.25) - (k S_n) - \delta \sigma}^{F^{-1}(0.75) + (k S_n) - \delta \sigma} f(x)\, dx} = \frac{1}{1 - \beta}$$

3.3. The Qn-TCC control chart

The third method we propose in this paper is called the Qn-TCC control chart, which is a modification of the TCC. This method is based on the Qn estimator, which is proposed by Rousseeuw and Croux (1993). The Qn estimator is simple, easy to compute and does not need any location estimator. It is also a powerful alternative to the MAD. For a random sample X1,X2,…,Xn the Qn estimator is defined as follows:

$$Q_n = 2.2219 \left\{ \left| X_i - X_j \right|;\ i < j \right\}_{(g)};\ i, j = 1, 2, 3, \ldots, n$$

where $g = \binom{h}{2} = \frac{h(h-1)}{2}$, $h = \left\lceil \frac{n}{2} \right\rceil + 1$

and $\left\lceil \frac{n}{2} \right\rceil$ is the integer part of the fraction $\frac{n}{2}$

Here the symbol (.) represents the combination. The $Q_n$estimator is 2.2219 times the *g-th*order statistic of the $\binom{n}{2}$ distances between data points and the factor 1.1926 is for consistency. The $Q_n$estimator has the highest breakdown possible points which is 50% and the influence function of it is smooth, bounded and has no discrete part. The efficiency of the Q_n estimator compare to IQR is 82% at normal distribution which is better than that of MAD and S_n estimators. It is found that for a small sample sizes, the S_n estimator performs better than the Q_n estimator (Rousseeuw and Croux, 1993). The lower and upper control limits of the Qn-TCC are obtained as follows:

$$LCL = Q_1 - k\, Q_n$$

$$UCL = Q_3 + k\, Q_n$$

Under the normal distribution with population mean μ and population standard deviation σ, the IQR equals 1.34898 σ and Qn equals 0.45062 σ so the Qn is 3 as much as IQR, then the parameter k is set as k = 4.5 in order to equal the control limits of the TCC.

When the control limits of Qn-TCC have been set, the ARL0 and ARL1 can be calculated. Assuming that the process is in control, the ARL0 and ARL1 for the Qn-TCC are obtained respectively as follows:

$$ARL_0 = \frac{1}{1 - \int_{F^{-1}(0.25) - kQ_n}^{F^{-1}(0.75) + kQ_n} f(x)\,dx} = \frac{1}{\alpha}$$

and

$$ARL_1 = \frac{1}{1 - \int_{F^{-1}(0.25) - (kQ_n) - \delta\sigma}^{F^{-1}(0.75) + (kQ_n) - \delta\sigma} f(x)\,dx} = \frac{1}{1-\beta}$$

Table 1, lists the value of the parameter k under the normal distribution for all proposed modified-TCCs methods in this paper.

Table 1. The values of the parameter k

Control Chart	k
TCC	1.5
MAD-TCC	3
Sn-TCC	2.4
Qn-TCC	4.5

4. Performance analysis

In this section, we carry out performance analysis of the proposed modified methods and the competing TCC. There are different measures based on run length (RL) that are used to evaluate the performance of control charts. The most common measure of these is the average run length (ARL). In this study, the ARL measure will be used as a criterion in evaluating the performance of all considered control charts followed by comparative analysis.

4.1. The performance of modified- TCCs and effects of outliers

To evaluate the performance of the proposed modified methods and the TCC, we conductedaMonte-Carlo simulation study to estimate ARL_0 and ARL_1. We assume that the in-control process follows a normal distribution with population mean, μ , and population variance σ^2, and the out-of-control process is normally distributed with population mean, μ , and population variance $\delta\sigma^2$ where δ refers to the amount of shift. We vary sample sizes n = 10, 20, 30, 50 and n = 100. The amount of shift (δ) values is assumed as $\{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$. The statistical software MATLAB R2016a was used in the data analysis. The Monte-Carlo simulation results gives an average of 300.000 simulated run lengths. The Newton-Raphson method was used to determine the values of k.

Table 2 summaries the simulations results for TCC. According to the results; every 118 samples will expect a false alarm for n = 10 and k=1.5, that is, $ARL(\delta = 0) = 117.7723$. This means that until the 118 samples are taken, the process is still under control. If $\delta = 0.5$, the process variation will be detected after sampling 54 times, that is, $ARL(\delta = 0.5) = 54.2548$. If only 10 observations are obtained for some reason, the value is about 82% of the theoretical value.

Tables 3-5 show ARLs values of MAD-TCC for various sample sizes.

Table 2. ARLs values of Tukey's control chart (TCC) for various sample sizes

Sample Size (n)	Shift (δ)						
	0	0.5	1.0	1.5	2.0	2.5	3.0
10	117.7723	54.2548	18.6112	7.7808	3.7205	2.2007	1.5657
20	120.4980	63.5718	21.1055	8.2182	4.0016	2.2986	1.6025
30	126.8320	66.2819	20.4156	8.3490	4.0161	2.3199	1.5997
50	138.2633	67.2616	21.7122	8.4745	4.0367	2.3110	1.6110
100	144.1205	68.0924	22.3167	8.5035	4.1172	2.3438	1.6125

Table 3. ARLs values of MAD-TCC for various sample sizes

Sample Size (n)	Shift (δ)						
	0	0.5	1.0	1.5	2.0	2.5	3.0
10	119.5074	61.5596	20.3545	8.1800	3.8475	2.2730	1.5979
20	126.8424	66.7574	21.7084	8.4969	4.0888	2.3558	1.6242
30	135.8967	69.5906	21.2512	8.5963	4.1176	2.3557	1.6149
50	142.4278	68.7830	22.3068	8.6286	4.1041	2.3318	1.6212
100	142.4475	68.3872	22.5721	8.5598	4.1479	2.3538	1.6152

Table 4. ARLs values of Sn-TCC for various sample sizes

Sample Size (n)	Shift (δ)						
	0	0.5	1.0	1.5	2.0	2.5	3.0
10	64.1884	32.5730	12.3883	5.4431	2.8740	1.8418	1.3892
20	89.9623	46.6643	16.2633	6.7298	3.4379	2.0814	1.4907
30	101.3785	53.4901	17.4659	7.2814	3.6337	2.1600	1.5245
50	118.4101	58.2513	19.3304	7.7786	3.7764	2.2125	1.5594
100	126.9907	62.0208	20.7892	8.0780	3.9470	2.2775	1.5818

Table 5. ARLs values of Qn-TCC for various sample sizes

Sample Size (n)	Shift (δ)						
	0	0.5	1.0	1.5	2.0	2.5	3.0
10	141.7237	66.4338	22.0997	8.5691	4.0850	2.3382	1.6219
20	141.9912	68.2538	21.8990	8.5191	4.1244	2.3612	1.6229
30	142.3076	67.7821	21.3017	8.5779	4.1099	2.3558	1.6178
50	142.8497	68.5882	22.1042	8.6571	4.1063	2.3383	1.6183
100	143.0513	68.3275	22.5187	8.5957	4.1456	2.3621	1.6194

On the other hand, when the observation value is 30, the $ARL(\delta = 0)$ value becomes 88.5% of its theoretical value.

From Table 3, it can be deduced that the performance of MAD-TCC is more efficient than that for the TCC. On average, every 120 samples will expect a false alarm for n = 10 and k=3, that is, $ARL(\delta = 0) = 119.5074$

If the process mean shifts to $\delta = 0.5$, then the process variation will be identified after sampling 62 times, for example, $ARL(\delta = 0.5) = 61.5596$. When the observation value is 30, the $ARL(\delta = 0)$ value becomes 95% of its theoretical value. Thus, for a better performance of the MAD-TCC, it may be recommended that at least 30 observations should be collected.

From Table 4, it can be concluded that the performance of Sn-TCC is less efficient than the TCC and MAD-TCC. If n = 10 and k = 2.4are selected, in average one will expect a false alarm in every 64 times of sampling, that is, $ARL(\delta = 0) = 64.1884$. If the process mean shifts to $\delta = 0.5$, then the process variation will be detected after sampling 33

times, that is, $ARL(\delta = 0.5) = 32.5730$

When the number of observations increases to 100, the $ARL(\delta = 0)$ value becomes 89% of its theoretical value, that is, we need more than 100 observations.

As seen in Table 5, Qn-TCC is more efficient than TCC, MAD-TCC and Sn-TCC, the $ARL(\delta = 0) = 141.7237$ for n = 10 and k = 4.5. This illustrates that the process is still in control through 142 samples are drawn. Thus, we can suggest that at least 10 observations should be collected, if one expects a better performance from Qn-TCC. Therefore, the Q_n Tukey's control chart (Qn-TCC) is the best one among the four methods considered in this study.

In Table 6, we consider 100 observations including one outlier (4.7198) to see the effect of outlier on the propsed control charts. The mean and standard deviation of these 100 observations are 0.0254 and 0.94396, respectively.

We define the outlier by a multiple of the standard deviation from mean. The ARL values for all charts are presented in Table 7.

In comparison the values of ARL in Table 7 with the values in Tables 2-5, it is easily seen that the ARL values for all considered methods in this study are very close and the performance of the modified-TCCs was not significantly affected by the an outlier value.

Table 6. The simulated observations including an outlier

0.5845	-0.3338	1.6470	-2.3649	0.4557	0.5049	-1.2036	0.8380	-1.3949	-1.1676
0.4048	-0.0522	-0.3085	0.7709	0.3745	0.2313	-0.0203	0.1264	-0.1522	-1.1676
0.0254	0.6836	0.2931	-1.3363	-1.2002	-0.3090	-0.3024	-0.7600	1.2369	-0.6596
-0.6518	0.2076	0.1172	-0.6530	-1.4289	-0.1111	-1.1891	0.6093	-0.0735	1.0924
-0.3165	0.2048	2.1862	-0.6970	0.7834	1.4233	1.6389	0.3906	0.3396	0.2167
-1.6342	-1.9518	1.4545	2.0922	-0.0307	-0.5869	0.2466	1.1835	0.6617	-0.1922
-1.1149	-1.0058	-1.3140	-0.6667	0.2847	1.5557	0.1800	-1.7435	0.4577	1.3295
0.7864	0.1887	1.6068	1.1374	0.5274	-0.6592	-0.0351	0.0102	0.3445	0.5089
-0.4693	0.0196	0.3070	1.6963	1.6514	1.4499	-0.8026	0.4402	1.0632	0.6232
1.1516	-0.1727	-0.6873	-0.0110	0.8203	-0.3850	-0.7876	-0.5618	-1.3261	4.7198

Table 7. Effect of outlier value on the ARL values

Control Chart	Shift (δ)						
	0	0.5	1.0	1.5	2.0	2.5	3.0
TCC	155.5771	75.3560	24.2966	9.2745	4.3983	2.4582	1.6631
MAD-TCC	158.3895	76.1870	24.5102	9.3548	4.4276	2.4725	1.6684
Sn-TCC	145.4152	70.2645	22.9234	8.8988	4.2495	2.4139	1.6404
Qn-TCC	168.6374	80.1944	25.5155	9.6944	4.5375	2.5322	1.6922

4.1. The performance of modified- TCCs and effects of outliers

In this section, we examine the effect of skewness and kurtosis on the ARL values for all control charts. In this regard, following Stoumbos and Reynolds (2000), Calzada and Scariano (2001), Lin and Chou (2007), Abu-Shawiesh (2008), and Torng and Lee (2008) the data are generated from different symmetric and skewed probability distributions. The Newton–Raphson method was used to determine k values.

We select four Student-t distributions of with degrees of freedom, $v = 30, 20, 10, 4$, Logistics distribution with a = 0 and b= 1, Laplace distribution with a = 0 and b = 1 and Gamma distribution with b = 1, a = 4, 1to examine the ARL values of TCC and the modified-TCCs. For comparing the performance of detecting mean shift between the TCC and modified-TCCs, the ARL_0, was adjusted to 143.34 and the parameter kwas chosen corresponding to the ARL_0. The ARL values are tabulated in Tables 8 - 11 for TCC, MAD-TCC, Sn-TCC and Qn-TCC respectively.

From Tables (8-11), it was indicated that the performance of detecting the process of the TCC and modified-TCCs, was worse when the distribution was away from a normal distribution. Obviously, this is especially true for cases of t (4), Laplace (0, 1) and Gamma (1,1).

Table 8. The ARL values of Tukey's control chart (TCC) for nonnormal distributions

Distribution		Shift (δ)						
	k	0	0.5	1.0	1.5	2.0	2.5	3.0
t (4)	2.956	143.32	123.98	83.54	46.86	23.03	10.34	4.55
t (10)	1.924	143.32	96.01	41.51	16.74	7.18	3.51	2.04
t (20)	1.680	143.32	82.36	30.46	11.78	5.26	2.80	1.78
t (30)	1.623	143.32	77.62	27.42	10.58	4.82	2.63	1.72
Logistic (0,1)	2.074	143.26	99.89	46.29	19.71	8.58	4.06	2.24
Laplace (0,1)	3.082	143.26	113.65	65.77	33.86	16.88	8.34	4.12
Gamma (1,1)	4.124	143.54	128.56	88.67	36.86	18.96	9.04	6.83
Gamma (4,1)	2.604	143.32	70.56	25.79	12.86	6.78	3.36	1.86

Table 9. The ARL values of MAD-TCC for nonnormal distributions

Distribution		Shift (δ)						
	k	0	0.5	1.0	1.5	2.0	2.5	3.0
t (4)	4.559	143.32	123.98	83.54	46.86	23.03	10.34	4.55
t (10)	3.578	143.32	96.01	41.51	16.74	7.18	3.51	2.04
t (20)	3.278	143.32	82.36	30.46	11.78	5.26	2.80	1.78
t (30)	3.177	143.32	77.62	27.42	10.58	4.82	2.63	1.72
Logistic (0,1)	3.720	143.26	99.89	46.29	19.71	8.58	4.06	2.24
Laplace (0,1)	4.480	143.26	113.65	65.77	33.86	16.88	8.34	4.12
Gamma (1,1)	5.120	143.54	128.56	88.67	36.86	18.96	9.04	6.83
Gamma (4,1)	4.012	143.32	70.56	25.79	12.86	6.78	3.36	1.86

Table 10. The ARL values of Sn-TCC for nonnormal distributions

Distribution		Shift (δ)						
	k	0	0.5	1.0	1.5	2.0	2.5	3.0
t (4)	3.842	127.86	108.69	68.76	38.76	20.16	9.94	4.11
t (10)	2.856	127.86	82.69	39.45	14.69	6.63	3.37	1.98
t (20)	2.624	127.86	76.89	28.64	10.89	5.04	2.54	1.09
t (30)	2.563	127.86	65.73	25.36	9.56	3.97	1.81	1.00
Logistic (0,1)	2.564	127.64	86.45	40.25	18.11	7.56	3.67	1.79
Laplace (0,1)	3.964	127.64	110.01	58.93	30.17	14.93	7.64	3.58
Gamma (1,1)	4.563	127.96	100.16	80.89	34.79	17.56	8.93	5.12
Gamma (4,1)	3.664	127.86	65.26	23.86	10.98	5.56	2.99	1.79

Table 11. The ARL values of Qn-TCC for nonnormal distributions

Distribution		Shift (δ)						
	k	0	0.5	1.0	1.5	2.0	2.5	3.0
t (4)	5.964	143.35	123.98	83.56	46.85	23.03	10.34	4.56
t (10)	5.642	143.35	96.01	41.51	16.74	7.18	3.51	2.04
t (20)	4.869	143.35	82.36	30.46	11.78	5.26	2.80	1.78
t (30)	4.674	143.35	77.62	27.42	10.58	4.82	2.63	1.72
Logistic (0,1)	4.701	143.26	99.89	46.29	19.71	8.58	4.06	2.24
Laplace (0,1)	5.967	143.26	113.65	65.77	33.86	16.88	8.34	4.12
Gamma (1,1)	6.862	143.54	128.56	88.67	36.86	18.96	9.04	6.83
Gamma (4,1)	5.014	143.32	70.56	25.79	12.86	6.78	3.36	1.86

5. An application example

In this section we provide an application of the proposed methodology using a real dataset from Tercero-Gomez, et al., 2012). The data points represent the time between arrivals, was observed at the Texas Tech University (TTU) library. A brief description of the dataset is given here. The data was manually measured by one of the authors of the paper. At that time authors were interested in simulation modeling where data collection presents an issue as assignable causes might contaminate the samples, hence biasing the analysis.

The authors find that the control charts may be helpful for detecting the changes and therefore the proposed chart could be used as a way to spot truly atypical behaviors and helps to determine if the data was suitable for modeling. By avoiding mistakes such as (i) including too large times between arrivals, or (ii) mislabeling behaviors as atypical when they were actually typical, was helpful to determine demand levels for authors models. A batch of extremely large time between arrivals might also indicate the end of the rush hour, a period beyond the scope of some of author's simulation models. 100 observations was collected. The data points are presented in Table 12.

Table 12 The sample of collected time between arrivals at TTU library

15.18	7.82	3.60	24.32	6.30	61.94	17.55	10.32	7.12	20.00
6.64	46.18	11.13	9.46	4.20	12.78	14.63	41.44	17.10	13.38
4.84	17.10	44.28	53.89	13.41	27.60	12.60	6.88	16.35	8.49
5.34	6.25	19.79	8.51	5.08	13.81	11.22	10.47	62.56	4.11
19.95	3.07	14.1	41.16	3.11	7.75	3.48	18.82	18.17	3.79
4.71	40.28	6.99	4.29	7.14	28.68	26.33	10.14	48.08	12.72
7.44	14.26	16.75	50.29	12.65	37.37	3.32	21.70	18.25	4.17
17.22	3.59	17.03	9.52	18.06	4.25	8.55	10.49	2.90	4.91
9.52	41.14	8.61	5.61	7.47	24.63	4.43	19.56	13.94	6.18
9.84	5.02	3.80	9.50	18.25	8.87	12.04	11.32	4.24	3.84

Anderson-Darling statistic of 7.260 was calculated to test normal distribution of the data set with P-value < 0.005. As can be observed, Anderson-Darling test's P-value and normal probability plot showed that the data is not normally distributed. Additionally, the histogram shows a skewed distribution.

Also the Box Plot shows several outliers on the right tail. According to Tercero-Gomez et al. (2012), the data comes from exponentially distributed population. Using the 100 observations, the calculated values of scale robust estimators, the control limits values and the number of observations out of the control limits for both the TCC and the modified-TCCs are given in Table 13 and Table 14, respectively.

Table 13. Values of scale estimators

Robust Scale Estimator Value			
IQR	MAD	Sn	Qn
11.995	8.947	7.931	8.065

Table 14. The control charts comparison

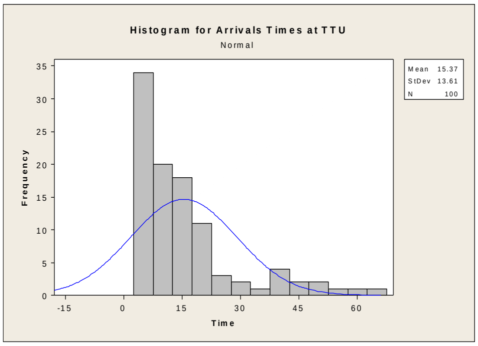
Control Chart	LCL	UCL	n
TCC	0	36.203	12
MAD-TCC	0	45.052	6
Sn-TCC	0	37.244	12
Qn-TCC	0	54.505	2

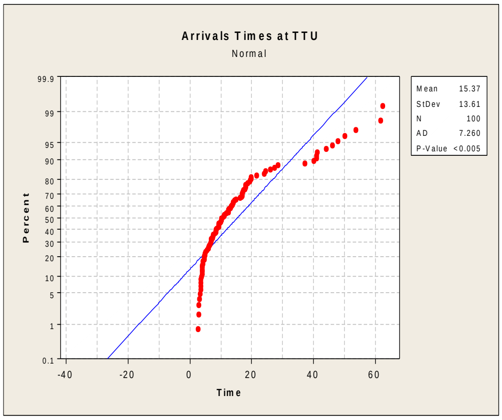
A summary with descriptive statistics from the data was obtained using Minitab® Release 14 (Minitab Inc., 2012) and the results are shown in Figure 1.

It was decided to set the LCL for the TCC and the modified-TCCs equal to 0 since it is impossible to have a negative time between arrivals. Figure 2 shows a control chart showing the sampled data and the control limits for both the TCC and the modified-TCCs.Comparing the control limits as shown in Figure 2, it can be seen that, according toTCC, there are 12 observations out of the control limits. According to MAD-TCC, there are 6 observations out of the control limits. According to Sn-TCC, there are 12 observations out of the control limits and only 2 observations out of the control limits according to the Qn-TCC control chart, so the modified-TCCs was more efficient and have more capacious than the TCC.

Summary for Arrival Times

Descriptive Statistics	
Sample Size (n)	100
Mean	15.37
Standard Deviation	13.61
Variance	185.29
Skewness	1.78
Kurtosis	2.73
Minimum	2.90
1st Quartile (Q_1)	6.215
Median	10.81
3rdQuartile (Q_3)	18.21
Maximum	62.56





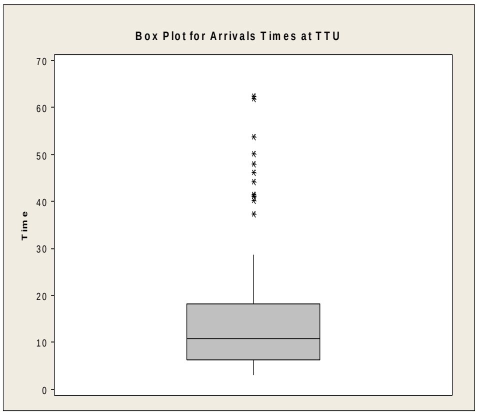


Figure 1. Descriptive statistics of the time between arrivals at a TTU library

6. Conclusion

This paper proposed some modified control limits of the Tukey Control Chart based on the robust estimators for IQR in order to reduce probability of type I error. The TCC method as well as the proposed modified methods use only information from the first three quartiles (Q_1, median, Q_3), the MAD, the S_n and the Q_n estimators to set up the control limits. A simulation study has been conducted to compare the performance of the control charts. Many symmetric and skewed distributions were selected to examine the performance of the proposed methods by comparing the ARL values. It is found that the Qn-TCC has the best process monitoring performance followed by the MAD-TCC.

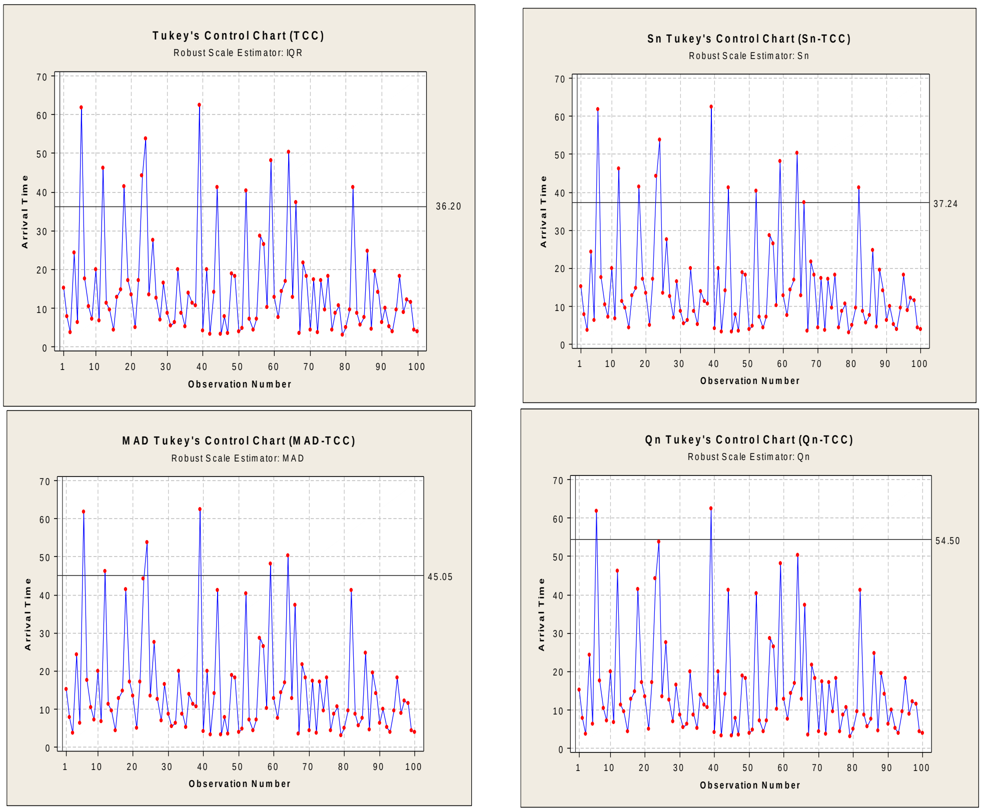


Figure 2. The TCC and modified-TCCs of the time between arrivals at a TTU library

The Sn-TCC has approximately the same process monitoring performance as the TCC. We also found that the ARL_1 performance, for both the TCC and the modified Tukey's control charts, are worse when the probability distribution is non-normal. Furthermore, it is observed that from the ARL_0 values, the performance of the modified Tukey's control charts has more capacious than the TCC with various numbers of observations. When the three proposed methods are compared with each other, we notice that the Qn-TCC produces the smaller ARL_1 and the sensitivity of it to detect a shift in the process when the shift occur is more than that for MAD-TCC and Sn-TCC. Both Sn-TCC and TCC produces approximately the same ARL_1. Furthermore, in the presence of outliers, the performance of TCC and the modified Tukey's control charts do not differ significantly. The proposed modified Tukey's control charts shows more robust behavior and detects the shifts more efficiently. The results of the numerical example supported the results of the simulation study to some extent. The superior detection ability depends on the magnitude of departure from normality. Finally, we hoped that the proposed modified Tukey's control charts, especially Qn-TCC and MAD-TCC, will serve as an attractive alternative to the Tukey's control chart (TCC) by quality control operators.

Acknowledgements: The authors will be thankful in advance to the editor in chief and anonymous reviewers for the time they give to consider and review thepaper.

References:

Abu-Shawiesh, M. (2008). A simple robust control chart based on MAD. *Journal of Mathematics and Statistics*, *4*(2), 102-107.

Akyüz, H. E., & Gamgam, H. (2017). Interval estimation for nonnormal population variance *with* kurtosis coeffificient based on trimmed mean. *Turkiye Klinikleri Journal of Biostatistics*, *9*(3), 213-221.

Akyüz, H. E., Gamgam, H., & Yalçınkaya, A. (2017). Interval estimation for the difference of *two* independent nonnormal population variances. *Gazi University Journal of Science*, *30*(3), 117-129.

Alemi, F. (2004). Tukey's control chart. *Quality Management in Health Care*, *13*(4), 216-221.

Borckardt, J. J., Nash, M. R., Hardesty, S., Herbert, J., Cooney, H., & Pelic, C. (2005). An *empirical* evaluation of Tukey’s control chart for use in health care and quality management applications. *Quality Managementin Healthcare*, *14*(2), 112-115.

Calzada, M. E., & Scariano, S. M. (2001). The robustness of the synthetic control chart to non-normality. *Communications in Statistics-Simulation and Computation*, *30*(2), 311-326.

Davis, R. B. (2004). Constructing control charts with average run length constraints. *In Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition*. Salt Lake City, Utah, pp. 1-9.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, *69*(436), 383-393.

Khaliq, Q. U. A., Riaz, M., & Alemi, F. (2015). Performance of Tukey’s and individual / moving range control charts. *Quality and Reliability Engineering International*, *31*(6), 1063-1077.

Lee, P. H. (2011). The effect of Tukey’s control chart with as asymmetrical control limits on monitoring of production processes. *African Journal of Business Management*, *5*(11), 4044-4050.

Lin, Y. C., & Chou, C. Y. (2007). Non-normality and the variable parameters X control charts. European *Journal of Operational Research*, *176*(1), 361-373.

Mekparyup, J., Kornpetpanee, S., & Saithanu, K. (2014). The performance of the adjusted Tukey's control chart under asymmetric distributions. *Global Journal of Pure and Applied Mathematics*, *10*(5), 719-724.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of* the *American Statistical Association*, *88*(424), 1273-1283.

Ryan, T. P. (2000). *Statistical methods for quality improvement* (2nd ed.). New York, NY: John Wiley and Sons.

Sindhumol, M. R., Srinivasan, M. M., & Gallo, M. (2016). Robust control charts based on modified trimmed standard deviation and Gini's mean difference. *Journal of Applied Quantitative Methods*, *11*(3), 18-30.

Stoumbos, Z. G. B., & Reynolds, M. R. (2000). Robustness to non-normality and autocorrelation of individuals control charts. *Journal of Statistical Computation and Simulation*, *66*(2), 145-187.

Sukparungsee, S. (2012). Robustness of Tukey's control chart in detecting a changes of parameter of skew distributions. *International Journal of Applied Physics and Mathematics*, *2*(5), 379-382.

Sukparungsee, S. (2013). Asymmetric Tukey's control chart robust to skew and non-skew process observation. *International Journal of Mathematical and Computational Sciences*, *7*(8), 1275-1278.

Tercero-Gomez, V., Ramirez-Galindo, J., Cordero-Franco, A., Smith, M., & Beruvides, M. (2012). Modified Tukey's control chart. *Communications in Statistics-Simulation and Computation*, *41*(9), 1566-1579.

Tiku, M. L., & Akkaya, A. D. (2004). *Robust estimation and hypothesis testing*. New Delhi: New Age International (P) Limited.

Torng, C. C., & Lee, P. H. (2008). ARL performance of the Tukey's control chart. *Communication in Statistics- Simulation and Computation*, *37(9)*, 1904-1913.

Wheeler, D. J., & Chambers, D. S. (1992). *Understanding statistical process control* (2nd ed.). Knowville, TN: SPC Press.

Moustafa Omar Ahmed AbuShawiesh
Department of Mathematics
Faculty of Science, The Hashemite University
Al-Zarqa, 13115,
Jordan
mabushawiesh@hu.edu.jo

Hayriye Esra Akyüz
Department of Statistics,
Faculty of Science and Arts
BitlisEren University,
Bitlis, 13000,
Turkey
heakyuz@beu.edu.tr

Hatim Solayman Ahmed Migdadi
Department of Mathematics
Faculty of Science, The Hashemite University
Al-Zarqa, 13115,
Jordan
hatims@hu.edu.jo

B.M. Golam Kibria
Department of Mathematics and Statistics, Florida International University,
Miami, FL 33199,
USA
kibriag@fiu.edu