Arab J Sci Eng (2017) 42:4941–4955
DOI 10.1007/s13369-017-2699-4

RESEARCH ARTICLE - MECHANICAL ENGINEERING

The Efficacy of Process Capability Indices Using Median Absolute Deviation and Their Bootstrap Confidence Intervals

Muhammad Kashif[1] · Muhammad Aslam[1] · Chi-Hyuck Jun[2] · Ali Hussein Al-Marshadi[1] · G. Srinivasa Rao[3]

Received: 21 February 2017 / Accepted: 12 July 2017 / Published online: 1 August 2017

Abstract The process capability indices (PCIs) C_p and C_{pk} are commonly used in industry to measure the process performance. The implementation of these indices required that process should follow a normal distribution. However, in many cases the underlying processes are non-normal which influence the performance of these indices. In this paper, median absolute deviation (MAD) is used as a robust measure of variability in two PCIs, C_p and C_{pk}. Extensive simulation experiments were performed to evaluate the performance of MAD-based PCIs under low, moderate and high asymmetric condition of Weibull, Log-Normal and Gamma distributions. The point estimation of MAD-based estimator of C_p and C_{pk} is encouraging and showed a good result in case of Log-Normal and Gamma distributions, whereas these estimators perform very well in case of Weibull distribution. The comparison of quantile method and MAD method showed that the performance of MAD-based PCIs is better for Weibull and Log-Normal processes under low and moderate asymmetric conditions, whereas its performance for Gamma distribution remained unsatisfactory. Four bootstrap confidence intervals (BCIs) such as standard (SB), percentile (PB), bias-corrected percentile (BCPB) and percentile-t (PTB) were constructed using quantile and MAD methods under all asymmetric conditions of three distributions under study. The bias-corrected percentile bootstrap confidence interval (BCPB) is recommended for a quantile (PC)-based PCIs, whereas CIs were recommended for MAD-based PCIs under all asymmetric conditions of Weibull, Log-Normal and Gamma distributions. A real-life example is also given to describe and validate the application of proposed methodology.

The original version of this article was revised: The spelling of the fourth author's name has been corrected.

✉ Muhammad Aslam
aslam_ravian@hotmail.com

Muhammad Kashif
mkashif@uaf.edu.pk

Chi-Hyuck Jun
chjun@postech.ac.kr

Ali Hussein Al-Marshadi
aalmarshadi@kau.edu.sa

G. Srinivasa Rao
gaddesrao@gmail.com

1 Department of Statistics, Faculty of Sciences, King Abdulaziz University, Jeddah 21551, Saudi Arabia

2 Department of Industrial and Management Engineering, POSTECH, Pohang 37673, Republic of Korea

3 Department of Statistics, School of Mathematical Sciences, CNMS, The University of Dodoma, Dodoma PO. Box: 259, Tanzania

Keywords Nonparametric confidence intervals · Process capability index · Median absolute deviation (MAD) · Percentile-t bootstrap (PTB) method

1 Introduction:

One of the important applications of statistical tools in manufacturing industries is to determine the process performance numerically, whose measures are known as process capability indices (PCIs). Among numerous PCI's, the most applicable indices are C_p, C_{pk}, C_{pm} and C_{pmk} [1–3]. All other indices can be viewed as further modification of the above four indices [4]. The four basic PCIs are summarized by [5] in a single measure using two nonnegative parameters (η, κ). The resulting measure is known as superstructure for the four basic indices and can be defined as

$$C_p(\eta, \kappa) = \frac{d - \eta |\mu - m|}{3\sqrt{\sigma^2 + \kappa (\mu - T)^2}} \quad (1)$$

where $d = USL - LSL/2$, $m = USL + LSL/2$ and T is the target value. Here, μ and σ^2 represent the mean and variance of the process and the values of (η and κ) lie between 0 and 1. The correct interpretation of the four PCIs depends mainly on the assumption that process follows a normal distribution [6].

However, in practice many processes are non-normal due to the involvement of various factors [7]. Therefore, the conventional PCIs defined in Eq. (1) are not appropriate for handling non-normal processes because both mean and standard deviation are not efficient estimator to deal with non-normality. Moreover, in non-normal distributed processes, the magnitude of the error varies substantially which makes mean and standard deviation non-appropriate estimators [8,9]. Also Hosseinifard *et al.* [10] pointed out that the performance of conventional methods is ineffective when normality assumption fails. Therefore, dealing with non-normality in PCIs is a core issue and is getting more attention by the researchers [6,7,9,11,12].

The literature reveals that the non-normality issues in PCIs can be resolved by applying five major approaches, but these approaches have been criticized by the researchers because of their variable performance under different situations [11]. This is an evidence that no single method has been recommended yet that works accurately in all situations [7].

Among five approaches, one most popular direction is to use the quantiles estimator introduced by Clements [13]. Based on this idea, the generalization of $C_p(\eta,\kappa)$ was proposed by Pearn and Chen (PC) [14], which is defined as

$$C_{Np}(\eta,\kappa) = \frac{d - \eta\,|M - m|}{3\sqrt{\left[\frac{Q_{99.865} - Q_{0.135}}{6}\right]^2 + \kappa\,(M - T)^2}} \tag{2}$$

where M is the sample median and $Q(0.9985)$ and $Q(0.00135)$ are the 99.865th and 0.135th quantiles, respectively. The basic idea is to use the normal distribution property to yield only 0.27% of nonconforming products. Hence, the process variability (6σ) is replaced by $Q(0.9985) - Q(0.00135)$ and the process mean is replaced by the median of the distribution.

But it is well established that the use of these PCIs, for heavily skewed distributions, did not provide accurate results [15]. Instead of using quantiles, there are other measures that were more suitable when median is used as measure of central tendency [16]. These measures are known as robust measures and have ability to provide more accurate results than the traditional statistical measures especially under non-normality. In the literature, it has been argued that the median absolute deviation (MAD) is a good robust estimator of standard deviation (σ) in case of a non-normal distribution [17–21].

There are some studies on the efficiency of MAD in calculating the PCI. Shu *et al.* [18] investigated the performance of C_p and C_{pk} for non-normal data using four different estimators. They have used classical, smooth adaptive estimator, MAD and Clements estimators for normal, student t and Gamma distributions. On the basis of mean square error (MSE) and bias, they suggested that MAD-based and Clements-based indices do not provide reliable estimates of C_p and C_{pk} as compared to the other indices. Recently, Adekeye [21] measures four PCI using MAD and compares their performance with quantile-based estimators. He conducted a simulation study using Weibull and exponential distributions. He concluded that MAD-based estimators gave better performance under heavily skewed process data compared to the quantile-based estimators.

However, Shu *et al.* [18] ignored the other commonly used distributions like Weibull and Log-Normal distributions in their research. Therefore, a more comprehensive study on using MAD as a measure of variability for Weibull and Log-Normal distributions is still required.

Beside the point estimation, the construction of confidence intervals for non-normal PCIs is also among major interests of researchers [14]. Though a point estimation is very common in PCIs, a confidence interval provides more comprehensive picture of the characteristics of interest than a point estimation [2]. Confidence intervals are useful in the correct interpretation of PCIs [22]. The construction of confidence intervals for PCI has been first studied by [23]. Over the years researchers have developed quite many techniques to construct confidence intervals for the PCIs. At the beginning, most of the confidence intervals for the PCIs were constructed for a normally distributed process. But later on, there have been some efforts to develop estimation techniques free from the normal assumption. The PCIs for non-normal distributions have been studied, for example, by [2,11,22,24]. For this purpose, a nonparametric statistical method called a bootstrap method introduced by [25] is frequently used. The main attraction in the use of this method is that it does not require the assumption of normality for calculating the confidence intervals.

In this work, our objectives are (1) to propose a comprehensive methodology for the construction of bootstrap confidence intervals using MAD as measure of variability for estimating $C_p(\eta,\kappa)$, (2) to investigate the behavior of the proposed estimator for low, moderate and high asymmetric non-normal distributions and (3) to compare its performance with quantile-based estimator proposed by [14]. The article is organized as follows: In Sect. 2, the methodology and relevant terminologies are introduced to describe PCI $C_p(\eta,\kappa)$ and their bootstrap confidence intervals under three commonly used non-normal distributions. Section 3 presents the results of the Monto Carlo simulation study to demonstrate the above methodology. In Sect. 4, a real-world data set is used to illustrate the application of the proposed approach. We conclude in a Sect. 5 with a brief discussion of the results of the study.

2 Methodology:

2.1 PCIs Using Median Absolute Deviation:

Suppose that the sample median (MD) is computed from a random sample $(x_1, x_2, \ldots, x_n)$. Then, MAD from the sample median is defined as [19,26,27]

$$MAD = b * median\,\{|x_i - MD|\} \tag{3}$$

The value of constant b in (3) is used to make the parameter of interest as a consistent estimator. In the case of an unbiased estimator of σ, we need to set $b = 1.4826$ if a random sample is taken from a normal distribution. For a non-normal distribution, this value changes to $b = 1/\,Q\,(0.75)$, where Q (0.75) is the 0.75 quantile of the underlying distribution. In case of normality, 1/Q (0.75) = 1.4826 [28]. Thus, the unbiased estimator of σ is

$$\hat{\sigma} = 1.4826\,(MAD)$$

Using the above relationship, the superstructure defined in (1) can be modified by

$$C_{MAD}\,(\eta,\kappa) = \frac{d - \eta\,|M - m|}{3\sqrt{(1.4826\,(MAD))^2 + \kappa\,(M - T)^2}} \tag{4}$$

The PCIs of C_p and C_{pk} can be derived from Eq. (2) or Eq. (4) based on PC and MAD estimators by letting $(\eta,\kappa) = [(0,0)\,,(1,0)]$ and $T = 0$, which are given in Table 1.

2.2 Comparison of PCIs for Three Non-Normal Distributions

The performance of the MAD- and PC-based PCIs C_p and C_{pk} was compared by using significantly different tail behavior distributions, i.e., Log-Normal, Weibull and Gamma [7,24,29–31]. The skewness is calculated by using different shape and scale parameters of each distribution which are categorized as low, moderate and high asymmetric levels as shown in Fig. (1) by plotting their PDF.

2.3 The method of bootstrap:

The complete bootstrap procedure is given for proposed study by following [30–32] in Sect. 2.3–2.6. The bootstrap procedure for independent and identically distributed random variables can be explained in the following way [32]. Let $z_1, z_2, z_3, \cdots z_n$ be a random sample of size n drawn from any distribution of interest say $\mathcal{F}$., i.e, $z_1, z_2, z_3, \cdots z_n \sim \mathcal{F}$. Let $\hat{\gamma}$ represent the estimator of PCI say C_p and C_{pk} based on PC or MAD method. Then,

1. Draw a bootstrap sample of size n, i.e., $z_1^*, z_2^*, z_3^* \cdots z_n^*$ from original sample by putting $1/n$ as mass at each point.
2. Let M_m^* where $1 \le m \le B$ be the mth bootstrap sample, then mth bootstrap estimator of γ is computed as

$$\widehat{\gamma_m}^* = \hat{\gamma}\left(z_1^*, z_2^*, z_3^* \cdots z_n^*\right) \tag{5}$$

3. Since there are total n^n resamples, there are total n^n values of $\widehat{\gamma_m}^*$. Each of these would be estimate of $\hat{\gamma}$. The ascending arrangement of the entire collection would constitute an empirical bootstrap distribution of $\hat{\gamma}$.

The construction of three bootstrap confidence intervals of the PCI $\hat{\gamma}\epsilon\left(C_p \text{ and } C_{pk}\right)$ based on $B = 1000$ bootstrap resamples is described as follows:

2.4 Standard Bootstrap (SB) Confidence Interval:

From $B = 1000$, bootstrap estimates of $\hat{\gamma}^*$, calculate the sample average and standard deviation as

$$\overline{\gamma^*} = (1000)^{-1}\sum\nolimits_{i-1}^{1000}\hat{\gamma}^*\,(i) \tag{6}$$

$$S_{\hat{\gamma}^*}^* = \sqrt{\left(\frac{1}{999}\right)\sum\nolimits_{i=1}^{1000}(\hat{\gamma}^*\,(i) - \overline{\gamma^*})^2} \tag{7}$$

The SB $(1-\alpha)\,100\%$ confidence interval is

$$CI_{SB} = \overline{\gamma^*} \pm Z_{1-\frac{\alpha}{2}}\,S_{\hat{\gamma}^*}^* \tag{8}$$

where $Z_{1-\frac{\alpha}{2}}$ is obtained by using $\left(1-\frac{\alpha}{2}\right)th$ quantile of the standard normal distribution.

2.5 Percentile Bootstrap (PB) Confidence Interval:

From the ordered collection of $\hat{\gamma}^*\,(i)$, choose $100\left(\frac{\alpha}{2}\right)\%$ and the $100\left(1-\frac{\alpha}{2}\right)\%$ points as the end points to calculate PB. Then, confidence interval would be

Table 1 Two PCIs using PC and MAD estimators for bilateral specifications

Estimator	C_p	C_{pk}
Pearn and Chen (PC)	$\widehat{C_p} = \frac{USL-LSL}{Q_{99.865}-Q_{0.135}}$	$\widehat{C_{pk}} = min\left\{\frac{USL-M}{\frac{Q_{99.865}-Q_{0.135}}{2}}, \frac{M-LSL}{\frac{Q_{99.865}-Q_{0.135}}{2}}\right\}$
MAD	$\widehat{C_p} = \frac{USL-LSL}{8.9MAD}$	$\widehat{C_{pk}} = min\left\{\frac{USL-M}{4.45MAD}, \frac{M-LSL}{4.45MAD}\right\}$

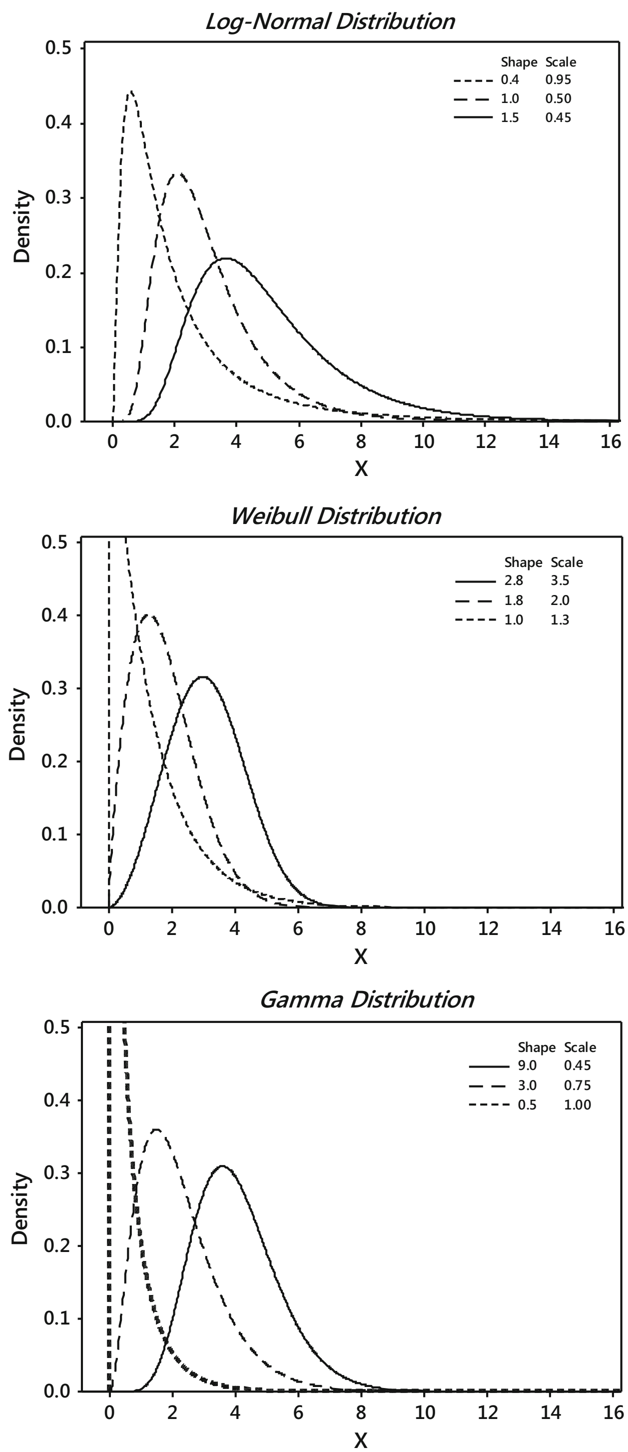

Fig. 1 PDFs of three distributions with selected shape and scale parameters

$$CI_{PB} = \left(\hat{\gamma}^*_{B\left(\frac{\alpha}{2}\right)}, \hat{\gamma}^*_{B\left(1-\frac{\alpha}{2}\right)}\right) \tag{9}$$

For a 95% confidence interval with $B = 1000$, it is:

$$CI_{PB} = \left(\hat{\gamma}^*_{(25)}, \hat{\gamma}^*_{(975)}\right) \tag{10}$$

2.6 Bias-Corrected Percentile Bootstrap (BCPB) Confidence Interval:

This method corrects the potential bias. Bias is generated because the bootstrap distribution is based on a sample from the complete bootstrap distribution and may be shifted higher or lower than what it would be expected. The calculation of this method is based on the following steps.

1. Using the (ordered) distribution of $\hat{\gamma}^*(i)$, compute the probability p_0 as $p_0 = pr\left(\hat{\gamma}^* \leq \hat{\gamma}\right)$
2. Let Ø and $Ø^{-1}$ represent the cumulative and inverse cumulative distribution functions of standard normal variable Z, then

$$Z_0 = Ø^{-1}(p_0)$$

3. The percentiles of the ordered distribution of $\hat{\gamma}^*$ are obtained as

$$P_L = Ø\left(2Z_0 + z_{\frac{\alpha}{2}}\right)$$
$$P_U = Ø\left(2Z_0 + z_{1-\frac{\alpha}{2}}\right)$$

Finally, the BCPB confidence interval is given as

$$CI_{BCPB} = \left(\hat{\gamma}^*_{(P_L B)}, \hat{\gamma}^*_{(P_U B)}\right) \tag{11}$$

The performance of the three confidence intervals: SB, PB and BCPB was compared using coverage probabilities and average widths. The coverage probability and average width of each BCI are calculated as

$$\text{Coverage Probability} = \frac{\left(Lw \leq \hat{C}_p \leq Up\right)}{B} \tag{12}$$

$$\text{Average Width} = \frac{\sum_{i=1}^{B}(Up_i - Lw_i)}{B} \tag{13}$$

3 Simulation Results:

In this section, the simulation results obtained by applying PC- and MAD-based estimators of two PCIs C_p and C_{pk} are reported. The bilateral limits for simulation were calculated, by fixing target value equal to 1.33, as

$$USL = \left(C_{pu} * 3\sigma_d\right) + \mu_d. \tag{14}$$
$$LSL = \mu_d - \left(C_{pu} * 3\sigma_d\right). \tag{15}$$

where μ_d and σ_d are the mean and standard deviation of each distribution calculated by using different shape and scale parameters [18]. The upper and lower specification limits

Table 2 Bilateral specification limits used for each distribution to calculate non-normal PCI

Parameter (shape, scale)	USL	LSL	σ
Log-Normal distribution			
(0.45, 1.50)	14.33	−4.42	2.35
(0.50, 1.00)	9.63	−3.47	1.64
(0.95, 0.40)	13.66	−8.97	2.84
Weibull distribution			
(2.8, 3.5)	7.924	−1.691	1.21
(1.8, 2.0)	5.858	−2.301	1.02
(1.0, 1.3)	6.487	−3.887	1.30
Gamma distribution			
(9.0, 0.45)	46.60	−6.60	6.67
(3.0, 0.75)	13.21	−5.21	2.31
(0.5, 1.0)	3.32	−2.32	0.71

for Log-Normal, Weibull and Gamma distribution are presented in Table 2. The negative limits are used for simulation purposes; however, it may be considered as zero because negative limits sometime have no implementation in real-world studies. For each distribution, 10,000 replications were performed with a sample size of n=25,50,75 and 100 to study the performance of PC- and MAD-based estimators of two PCIs, C_p and C_{pk}

The simulation results are presented in Tables (3, 4, 5). The complete simulation was run by using the R software.

These tables depict the mean, SD, simulated bias in standard units and mean square error (MSE) corresponding to the target value equal to 1.33 by using both indices. The results are arranged according to the low, moderate and high asymmetric behavior of each distribution.

3.1 Results for Log-Normal Distribution :

The mean value of C_p or C_{pk} using PC-based estimator is smaller than the target value in all cases. The trend demonstrates that the value of both indices decreases as the asymmetry and sample size increase. On the other hand, the values of these indices become higher and get closer to the target values as the sample size increases for the MAD-based estimator. Indeed, the values of C_p and C_{pk} greater than 1.33 allow the user to declare that a process is in highly capable condition. In general, for high asymmetry both indices of Cp and Cpk behave in different ways. In case of PC estimator, both indices of Cp and Cpk perform poorly. In particular, the index C_{pk} does not demonstrate a good pattern to deal with high asymmetry. On the other hand, both indices are always far above the values of the true C_p and C_{pk} in the case of the MAD estimator. So neither of the two estimators of Cp and Cpk can be classified as a good estimator. Regarding the simulated bias and mean square error, both decrease as the sample size increases for both estimators. The negative bias values indicate that PC-based estimator underestimates the indices. On the other hand, the direction of bias of MAD-based estimator is positive in all cases which indicates that it overestimates the indices. However, it is observed that under high asymmetry both estimators provide higher bias and MSE especially for the MAD. So we must be very careful with values of bias and MSE because these values increase from low to high asymmetry for both estimators.

3.2 Results for Weibull Distribution :

Unlike the Log-Normal distribution, both indices of Cp and Cpk showed a different pattern in the case of Weibull distribution. From Table 4, it is observed that the performance of MAD-based estimator is consistently better than that of PC-based estimator from low to high asymmetry. In the case of low and moderate asymmetry, PC-based estimator overestimates and underestimates in the case of high asymmetry. On the other hand, MAD-based estimator of Cp and Cpk yields values that are very close to the target values except in high asymmetry using large sample sizes. For high asymmetry, MAD-based estimator overestimates both indices. The bias and MSE using PC-based estimator are larger than MAD-based estimator in case of low and moderate asymmetry and less in high asymmetry for both indices.

3.3 Results for Gamma Distribution :

In the case of Gamma distribution, the PC-based estimator gives more accurate and precise estimates of C_p and C_{pk} under low and moderate asymmetry and showed poor performance in case of high asymmetry. The MAD-based estimator consistently overestimates both indices in all cases.

4 Comparison of Bootstrap Confidence Intervals of C_p and C_{pk} for Log-Normal, Weibull and Gamma Distributions

In this section, we compare the performance of four types of bootstrap confidence intervals using MAD- and PC-based estimators of C_p and C_{pk} for all distributions under studies. For each case $B = 1000$, bootstrap resamples were drawn to estimate the coverage probabilities and average width of above four confidence intervals. The simulated results using bilateral specifications and different combination of values for the shape and scale parameters (Table 2) are presented in Figs. 2, 3, 4 and 5.

Comparison of the average width of four type of confidence intervals for both estimators of Cp and Cpk using three distributions is presented in Figs. 2 and 3, whereas the cover-

Table 3 Comparison of statistical indicators for the estimators of C_p and C_{pk} under three Log-Normal distributions

C_p		LGN (1.5,0.45)		LGN (1.0,0.50)		LGN (0.40,0.95)	
Indicator	n	PC	MAD	PC	MAD	PC	MAD
Mean	25	1.2619	1.8450	1.2324	1.9332	1.0586	3.4966
	50	1.2103	1.7079	1.1861	1.7835	0.9685	3.2390
	75	1.1918	1.6659	1.1612	1.7344	0.9364	3.1342
	100	1.1865	1.6443	1.1567	1.7171	0.9217	3.0830
SD	25	0.3071	0.6040	0.3244	0.6478	0.4874	1.4060
	50	0.2049	0.3752	0.2169	0.4017	0.3142	0.9070
	75	0.1660	0.2922	0.1730	0.3089	0.2460	0.7051
	100	0.1407	0.2489	0.1478	0.2653	0.2102	0.5918
Bias/σ	25	−0.0290	0.2192	−0.0595	0.3678	−0.1655	1.3211
	50	−0.0510	0.1608	−0.0878	0.2765	−0.2204	1.1640
	75	−0.0588	0.1429	−0.1029	0.2466	−0.2400	1.1001
	100	−0.0611	0.1338	−0.1056	0.2361	−0.2490	1.0689
MSE	25	0.0047	0.2653	0.0096	0.3640	0.0737	4.6940
	50	0.0143	0.1428	0.0207	0.2057	0.1307	3.6443
	75	0.0191	0.1128	0.0285	0.1636	0.1549	3.2552
	100	0.0206	0.0988	0.0300	0.1499	0.1667	3.0729
C_{pk}							
Mean	25	1.1931	1.7510	1.1595	1.8122	0.9688	3.2279
	50	1.1480	1.6190	1.1175	1.6769	0.8917	2.9746
	75	1.1333	1.5825	1.0992	1.6432	0.8641	2.8871
	100	1.1269	1.5620	1.0925	1.6225	0.8539	2.8559
SD	25	0.2824	0.5648	0.2894	0.5923	0.4249	0.6683
	50	0.1831	0.3447	0.1941	0.3618	0.2825	0.5791
	75	0.1465	0.2719	0.1558	0.2868	0.2222	0.5483
	100	0.1261	0.2330	0.1323	0.2423	0.1853	0.5373
Bias/σ	25	−0.0583	0.1792	−0.1040	0.2940	−0.1272	0.6683
	50	−0.0775	0.1230	−0.1296	0.2115	−0.1543	0.5791
	75	−0.0837	0.1075	−0.1407	0.1910	−0.1641	0.5483
	100	−0.0864	0.0987	−0.1448	0.1783	−0.1676	0.5373
MSE	25	0.0188	0.1774	0.0291	0.2326	0.1306	3.6021
	50	0.0331	0.0835	0.0452	0.1204	0.1921	2.7048
	75	0.0387	0.0638	0.0533	0.0981	0.2171	2.4246
	100	0.0412	0.0538	0.0564	0.0855	0.2267	2.3283

age probabilities of these confidence intervals are presented in Figs. 4 and 5. It is observed that the average width of all confidence intervals reduces when the sample size increases in all cases under studies. The asymmetric level affects the average width. Average width increases as asymmetry level increases except Weibull distribution. In case of Weibull distribution, we observed a reverse pattern.

From the results of Log-Normal distribution, we conclude the followings.

1. For PC-based estimator of both indices, SB and BCBP methods showed the better coverage probabilities under all asymmetric levels. However, BCBP performs better because it provides a smaller average width compared to SB method.
2. In the case of MAD, PTB method is recommended based on coverage probabilities and average widths.

The results of average widths and coverage probabilities using Weibull distribution are presented in Figs. 2 (b), 3 (h), 4 (f) and 5 (i). From these results, the following recommendations can be made.

1. Using both estimators of Cp and Cpk, the average width decreases as sample size and asymmetry level increase. On the other hand, Log-Normal and Gamma distributions showed wider width as asymmetry level increases.

Table 4 Comparison of statistical indicators for the estimators of C_p and C_{pk} under three Weibull distributions

C_p		W (2.8,3.5)		W (1.8,2.0)		W (1.0,1.3)	
Indicator	n	PC	MAD	PC	MAD	PC	MAD
Mean	25	1.5454	1.4911	1.5327	1.5598	1.3554	2.2423
	50	1.5040	1.3768	1.4823	1.4355	1.2837	2.0343
	75	1.4946	1.3425	1.4695	1.4059	1.2507	1.9709
	100	1.4883	1.3293	1.4640	1.3871	1.2442	1.9399
SD	25	0.2109	0.4430	0.2788	0.4778	0.4309	0.8977
	50	0.1413	0.2695	0.1840	0.2872	0.2812	0.5331
	75	0.1127	0.2121	0.1482	0.2283	0.2220	0.4137
	100	0.0957	0.1805	0.1288	0.1938	0.1939	0.3516
Bias	25	0.1657	0.1240	0.1559	0.1768	0.0196	0.7018
	50	0.1338	0.0360	0.1171	0.0811	−0.0356	0.5418
	75	0.1266	0.0096	0.1073	0.0583	−0.0610	0.4930
	100	0.1218	−0.0005	0.1031	0.0439	−0.0660	0.4691
MSE	25	0.0464	0.0260	0.0411	0.0528	0.0007	0.8324
	50	0.0303	0.0022	0.0232	0.0111	0.0022	0.4960
	75	0.0271	0.0002	0.0195	0.0058	0.0063	0.4108
	100	0.0251	0.0000	0.0180	0.0033	0.0074	0.3720
C_{pk}							
Mean	25	1.4732	1.4250	1.4543	1.4771	1.2501	2.0634
	50	1.4582	1.3282	1.4266	1.3804	1.1893	1.8837
	75	1.4561	1.3082	1.4124	1.3485	1.1618	1.8238
	100	1.4537	1.2942	1.4104	1.3337	1.1494	1.8013
SD	25	0.2077	0.4278	0.2611	0.4473	0.3894	0.7615
	50	0.1403	0.2618	0.1758	0.2727	0.2571	0.4698
	75	0.1119	0.2074	0.1424	0.2159	0.2037	0.3636
	100	0.0965	0.1735	0.1240	0.1825	0.1734	0.3044
Bias	25	0.1101	0.0731	0.0956	0.1132	−0.0614	0.5641
	50	0.0986	−0.0014	0.0743	0.0388	−0.1082	0.4259
	75	0.0970	−0.0167	0.0634	0.0142	−0.1294	0.3798
	100	0.0951	−0.0275	0.0618	0.0029	−0.1390	0.3626
MSE	25	0.0205	0.0090	0.0155	0.0217	0.0064	0.5379
	50	0.0164	0.0000	0.0093	0.0025	0.0198	0.3066
	75	0.0159	0.0005	0.0068	0.0003	0.0283	0.2438
	100	0.0153	0.0013	0.0065	0.0000	0.0326	0.2221

2. Overall, MAD-based estimator provides higher coverage probabilities as compared to PC.
3. The BCPB method performed better in the case of PC-based estimator and PTB method considered better based on the smaller width and stable coverage probabilities.

The simulated average widths of both indices in case of Gamma distribution are presented in Figs. 2 (c) and 3 (j), and the coverage probabilities are shown in Figs. 4 (g) and 5 (m) which indicate the following.

1. For high asymmetry, MAD estimator showed very large average widths for small samples sizes especially in the case of C_p. However, it decreases rapidly when sample size increases for both C_p and C_{pk}. Overall PB and PTB methods provide almost similar probabilities and average widths, but PB considered superior.
2. The BCBP method showed better performance as compared to other three methods for both indices in case of PC-based estimators followed by PTB method.

In short, this comparative study showed that the MAD-based estimator of Cp and Cpk gives very consistent and accurate results for a Weibull distribution under low and moderate asymmetry levels. The PC-based estimator gives more stable results for Gamma distributions except the high asymmetry case. The performance of both estimators in case of a

Table 5 Comparison of statistical indicators for the estimators of C_p and C_{pk} under three Gamma distributions

C_p		GA(9.0,0.45)		GA(3.0,0.75)		GA(0.5,1.0)	
Indicator	*n*	PC	MAD	PC	MAD	PC	MAD
Mean	25	1.3956	1.5863	1.3852	1.7058	1.2547	4.0664
	50	1.3561	1.4683	1.3347	1.5701	1.1702	3.4832
	75	1.3452	1.4399	1.3220	1.5357	1.1493	3.3268
	100	1.3391	1.4228	1.3156	1.5187	1.1326	3.2276
SD	25	0.2311	0.4907	0.2706	0.5458	0.4463	2.9245
	50	0.1546	0.3025	0.1764	0.3367	0.2796	1.4899
	75	0.1234	0.2423	0.1419	0.2664	0.2228	1.1260
	100	0.1075	0.2021	0.1228	0.2246	0.1868	0.9197
Bias	25	0.0098	0.0384	0.0239	0.1627	−0.1061	3.8541
	50	0.0039	0.0207	0.0020	0.1039	−0.2250	3.0326
	75	0.0023	0.0165	−0.0035	0.0891	−0.2545	2.8124
	100	0.0014	0.0139	−0.0062	0.0817	−0.2780	2.6727
MSE	25	0.0043	0.0657	0.0031	0.1412	0.0057	7.4879
	50	0.0007	0.0191	0.0000	0.0577	0.0255	4.6364
	75	0.0002	0.0121	0.0001	0.0423	0.0327	3.9872
	100	0.0001	0.0086	0.0002	0.0356	0.0390	3.6010
Cpk							
Mean	25	1.3331	1.5238	1.3014	1.6036	0.8317	2.7100
	50	1.3079	1.4170	1.2733	1.5022	0.7797	2.3189
	75	1.3007	1.3870	1.2586	1.4636	0.7612	2.2047
	100	1.2972	1.3772	1.2528	1.4478	0.7549	2.1562
SD	25	0.2135	0.4747	0.2385	0.5072	0.2822	1.8494
	50	0.1437	0.2898	0.1594	0.3151	0.1743	0.9648
	75	0.1172	0.2309	0.1260	0.2481	0.1394	0.7050
	100	0.0997	0.1957	0.1080	0.2086	0.1186	0.5880
Bias	25	0.0005	0.0291	−0.0124	0.1185	−0.7019	1.9437
	50	−0.0033	0.0130	−0.0245	0.0745	−0.7751	1.3929
	75	−0.0044	0.0085	−0.0309	0.0579	−0.8012	1.2320
	100	−0.0049	0.0071	−0.0334	0.0510	−0.8099	1.1636
MSE	25	0.0000	0.0376	0.0008	0.0749	0.2483	1.9045
	50	0.0005	0.0076	0.0032	0.0297	0.3028	0.9783
	75	0.0009	0.0033	0.0051	0.0179	0.3236	0.7652
	100	0.0011	0.0022	0.0060	0.0139	0.3307	0.6826

Log-Normal distribution is not satisfactory as compared to Weibull and Gamma distributions.

The outcome of our study reveals deviation from the findings of Shu *et al* [18] which states that median-based PCIs are not good estimators of C_p and C_{pk} due to their large MSE and bias. Our findings are not surprising because every method depicts varying performance for a particular distribution with a different tail behavior [12,33]. To validate the results in a comprehensive way, performance comparison of current findings is made with classical estimator in [18]. The results of MAD-based estimator of both PCIs for Weibull distribution and PC-based estimator of two PCIs for Gamma distribution under low and moderate asymmetry were included in comparison with classical estimator only. Moreover, it was noted that there is little difference between classical and smooth adaptive index when there is no outlier in the data set [34]. Therefore, only classical estimator is taken for comparison in this study.

A comparison of both indices with classical estimator in terms of relative bias and MSE is presented in Figs. (6) and (7). In case of Weibull distribution, it is observed that for large sample size, under low and moderate asymmetry MAD-based PCIs perform better than classical PCIs and showed less bias and MSE. On the other hand, PC-based indices dominate classical indices in case of Gamma distribution

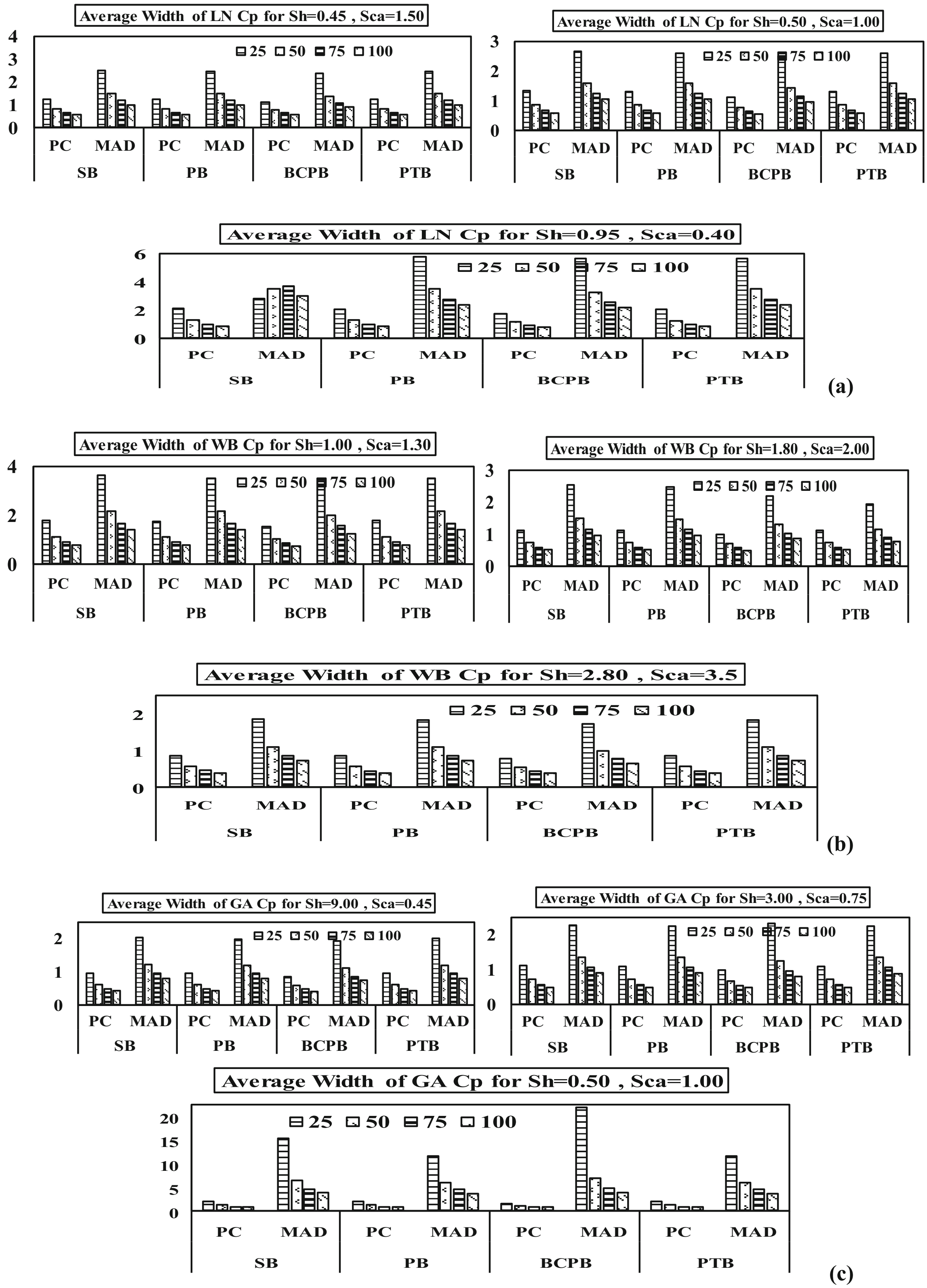

Fig. 2 Comparison of average widths of three distributions for index C_p using both methods

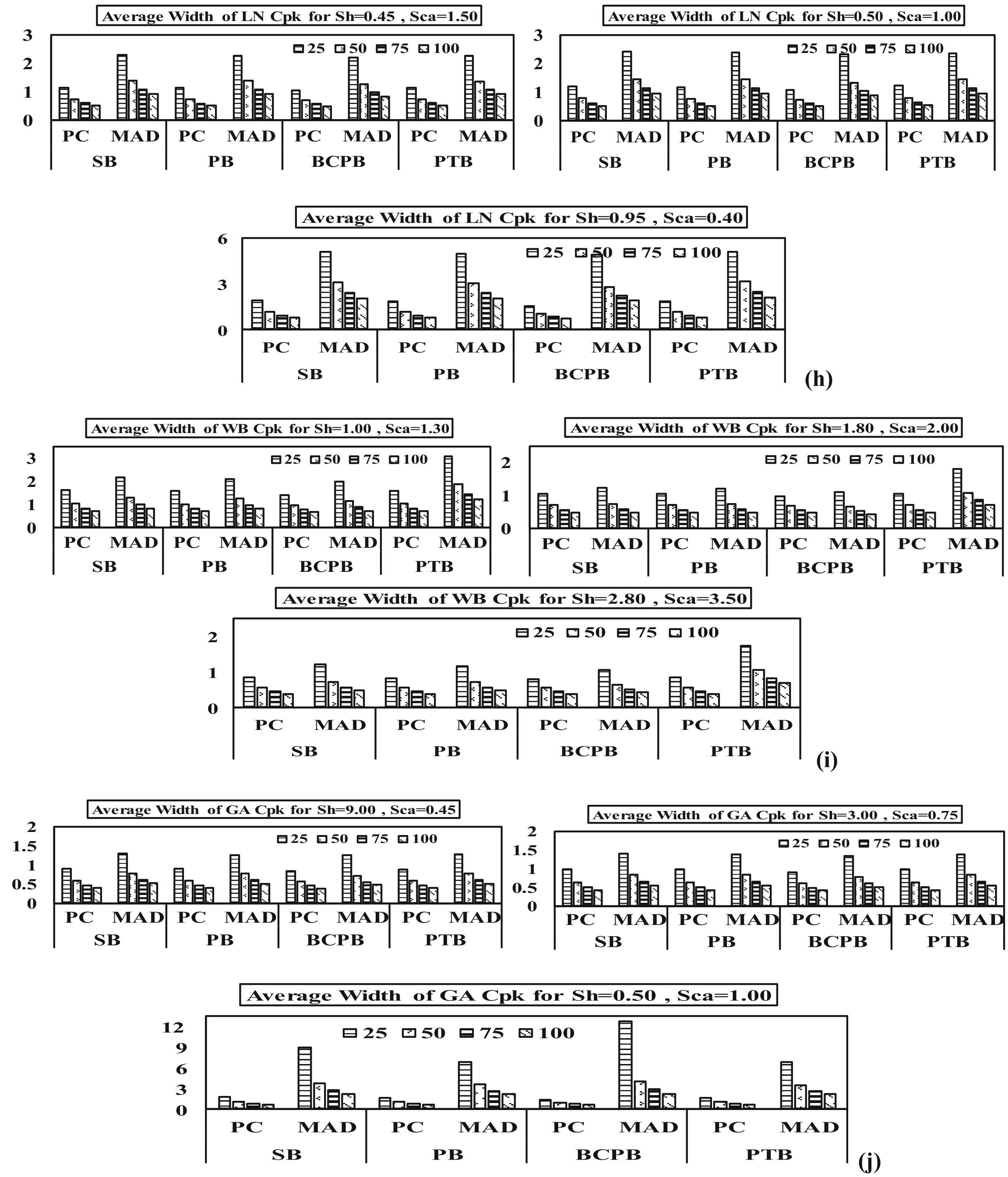

Fig. 3 Comparison of average widths of three distributions for index C_{pk} using both methods

especially under low asymmetry. For moderate asymmetry, it gives good results for small sample sizes.

Among four bootstrap confidence intervals, the results showed that percentile-t bootstrap (PTB) method has the upper hand in the average width and coverage probability comparison over its competitors for MAD-based estimator of both indices. The BCBP method outperforms in all aspects when considering the results of PC-based estimators. Finally, among two indices C_{pk} is more sensitive to departure from normality than C_p in the case of high asymmetry.

4.1 Example

In this section, we present a real-life example to demonstrate the application of the proposed methodology for two process capability indices of C_p and C_{pk}. We considered

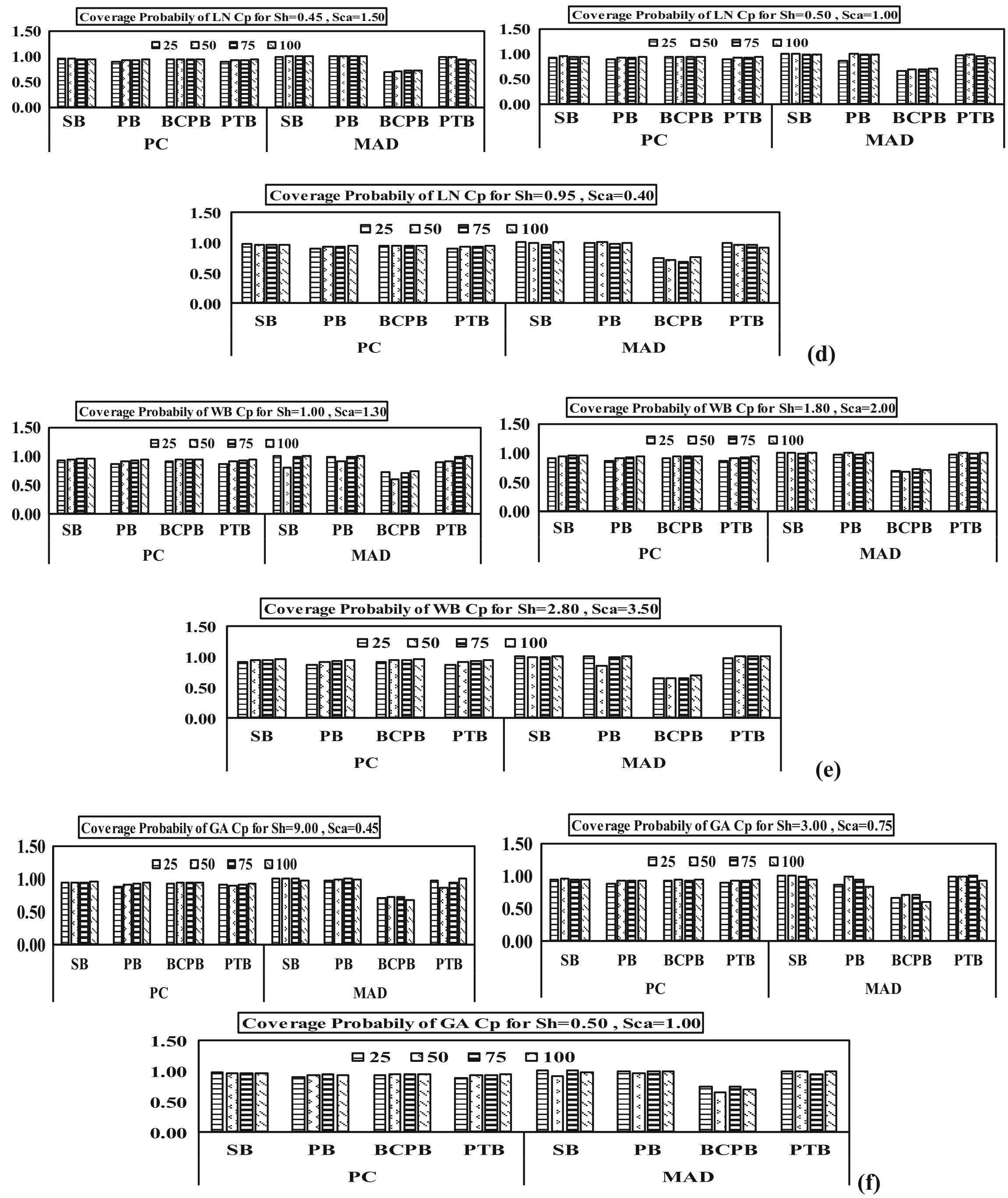

Fig. 4 Comparison of coverage probabilities of three distributions for index C_p using both methods

the empirical data given by [35–37]. To select the appropriate distribution for the selected data, the different goodness of fit statistics [38] is used and reported in Table 6. Based on AIC and BIC values, it is confirmed that two-parameter Weibull distribution is suitable for these data compared to other distributions. The basic descriptive statistics of the data is reported in Table 7. By fitting two-parameter Weibull distribution, the maximum likelihood estimator for shape and scale parameters is $\hat{\gamma} = 5.504809$, $\hat{\beta} = 2.650830$, respectively. To evaluate the adequacy of the data, Kolmogorov–Smirnov (K–S) goodness of fit test is used. The K–S distance value for these data is 0.056 with p-value 0.9816, which also shows in favor of Weibull distribution. The lower and upper specification limits used for the calculations of two process capability indices are (0.3989, 4.4960). The estimates of both indices using MAD and PC estimators and their corresponding boot-

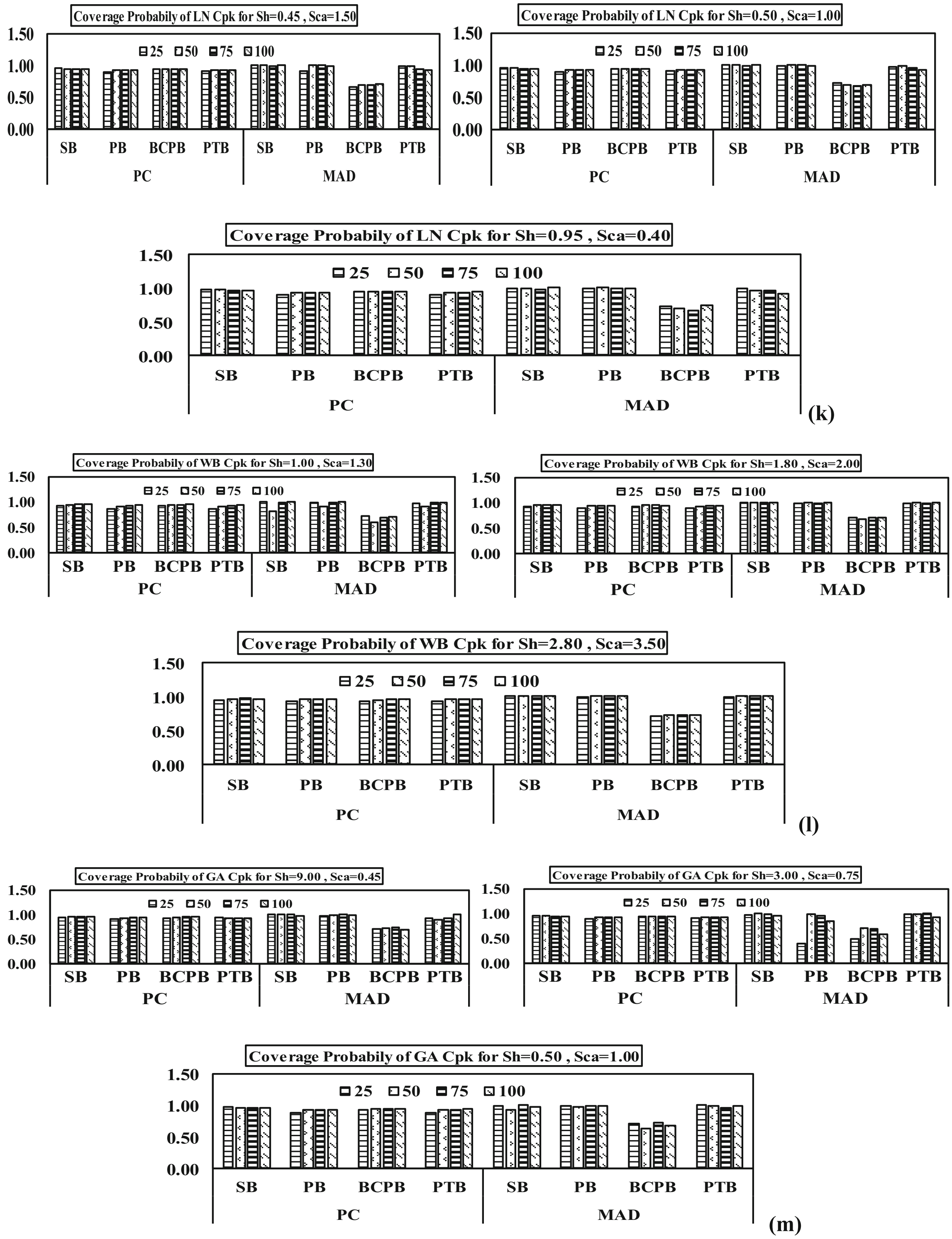

Fig. 5 Comparison of coverage probabilities of three distributions for index C_{pk} using both methods

strap CIs are reported in Table 8. The estimated values of both indices are close to target values using the MAD estimator. This showed good performance of MAD-based estimator of both indices but with wider width of CI. The comparison of four bootstrap confidence intervals for C_p and C_{pk} using both estimators showed that BCPB is good for Cp and PTB for Cp_k in both cases. For each method, confidence intervals are presented in brackets.

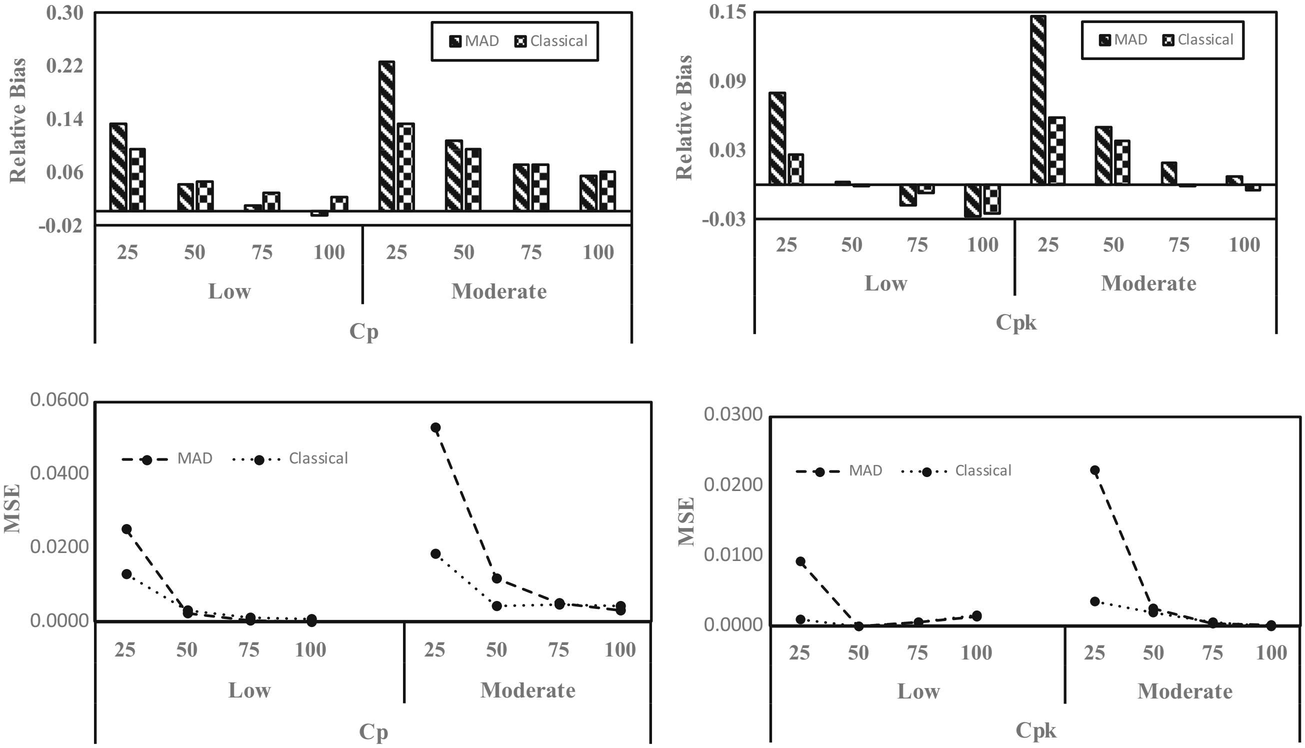

Fig. 6 Comparison of Bias and MSE between MAD and classical estimators for C_p and C_{pk} under low and moderate asymmetry for Weibull distribution

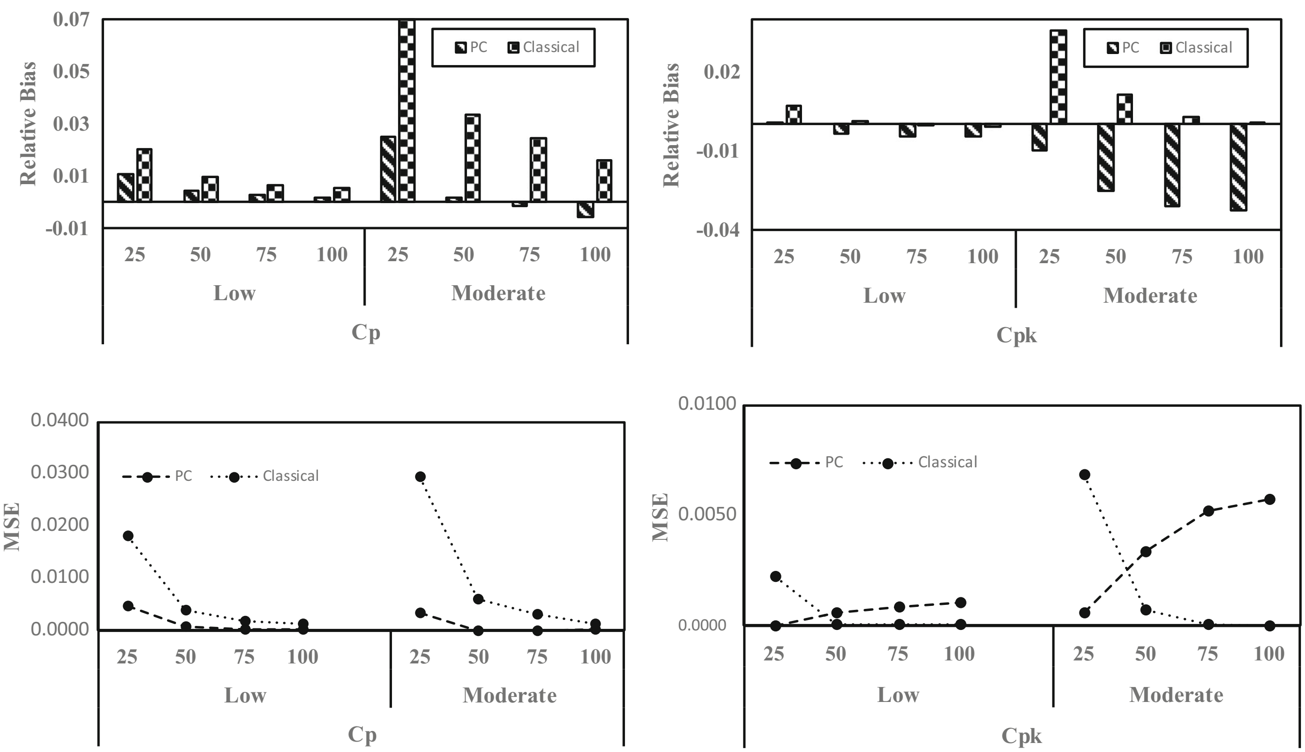

Fig. 7 Comparison of Bias and MSE between PC and classical estimators for C_p and C_{pk} under low and moderate asymmetry for Gamma distribution

Table 6 Goodness of fit statistics of three distributions

	LGN	W	GA
Kolmogorov–Smirnov (KS) Statistics	0.072	0.056	0.058
Cramer–von Mises (CvM) Statistics	0.076	0.034	0.045
Anderson – Darling (AD) statistics	0.544	0.274	0.334
AIC	106.76	103.19	104.07
BIC	111.24	107.66	108.54

Table 7 Summary statistics of example data

n	Min.	Max.	$Q_{0.00135}$	$Q_{0.50}$	$Q_{0.9986}$	Mean	SD	MAD	Sk	Ku
69	1.31	3.58	1.31	2.48	3.58	2.45	0.49	0.33	−0.03	3.03

Table 8 Bootstrap confidence interval widths for example data

	MAD	Quantile
C_p	1.3987	1.4150
SB	0.9331	0.3898
	(0.9264–1.8595)	(1.2219–1.6118)
PB	0.9241	0.3876
	(1.0056–1.9298)	(1.2427–1.6304)
BCPB	0.8215	0.3707
	(1.0527–1.8742)	(1.2201–1.5909)
PTB	0.9083	0.3889
	(0.9760–1.8843)	(1.2414–1.6303)
C_{pk}	1.3601	1.3745
SB	0.9040	0.3813
	(0.9013–1.8053)	(1.1860–1.5673)
PB	0.8957	0.3794
	(0.9771–1.8729)	(1.2042–1.5837)
BCPB	0.9118	0.3812
	(0.7583–1.6701)	(1.2077–1.5890)
PTB	0.8056	0.3807
	(1.0576–1.8632)	(1.2029–1.5836)

5 Conclusions:

This study demonstrates the application of the median absolute deviation (MAD) as a measure of variability to commonly used process capability indices Cp and Cpk. The effectiveness of these estimators of PCIs was measured under low, mild and severe asymmetric level for three different distributions, and their performance was compared with Pearn and Chen (PC) estimator under similar conditions. The results show that the use of MAD approach is very attractive and it provides better results than PC estimator under different asymmetric levels of different distributions. In particular, MAD-based estimator of both indices showed a great potential for dealing with high asymmetry. It is observed that both estimators performed variably in different distributions. Under low and moderate asymmetry, in Weibull distribution, the MAD-based estimators are more close to the target values, whereas PC-based estimators performed better in Gamma distribution. Moreover, the MAD estimators showed more appropriate results than PC estimators in all distributions under high asymmetry. After this discussion, the median absolute deviation is recommended as a useful measure in dealing with asymmetry where the target value should not less than 2. So it is recommended that MAD can be used as an alternative measure of variability more properly in PCI especially in low and moderate asymmetry when process follows Weibull distribution. Along with point estimation, we also constructed four types of bootstrap confidence intervals using extensive simulation studies. Among four types of confidence intervals, BCBP and PTB methods provide reliable confidence intervals for PC and MAD estimators under all asymmetry levels and sample sizes. Finally, PTB method should be used for the construction of bootstrap confidence intervals when MAD is used as an alternative measure of variability.

Acknowledgements The authors are deeply thankful to editor and reviewers for their valuable suggestions to improve the quality of this manuscript. This article was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah. The author, Muhammad Aslam, therefore, acknowledges with thanks DSR technical and financial support.

References

1. Kane, V.E.: Process capability indices. J Qual Technol **18**(1), 41–52 (1986)
2. Panichkitkosolkul, W.: Confidence interval for the process capability index Cp based on the Bootstrap-t confidence interval for the standard deviation. Metodološki zvezki **11**(2), 79–92 (2014)
3. Z Zhang, J.: Conditional confidence intervals of process capability indices following rejection of preliminary tests (Order No. 3408970). Available from ProQuest Dissertations & Theses Global. (609568394). Retrieved from https://search.proquest.com/docview/609568394?accountid=135034. (2010)
4. Pearn, W.; Chen, K.: Estimating process capability indices for non-normal Pearsonian populations. Qual Reliab Eng Int **11**(5), 386–388 (1995)
5. Vännman, K.: A uniformed approach to capability indices. Statistica. **5**(2), 805–820 (1995)
6. Sennaroglu, B.; Senvar, O.: Performance comparison of box-cox transformation and weighted variance methods with weibull distribution. J. Aeronauti. Space Technol. **8**(2), 49–55 (2015)
7. Piña-Monarrez, M.R.; Ortiz-Yañez, J.F.; Rodríguez-Borbón, M.I.: Non-normal capability indices for the Weibull and Lognormal distributions. Qual. Reliab. Eng. Int. **32**(4), 1321–1329 (2016)
8. Senvar, O.; Kahraman, C.: Type-2 fuzzy process capability indices for non-normal processes. J. Intell. Fuzzy Syst. **27**(2), 769–781 (2014)

9. Senvar, O., Kahraman, C.: Fuzzy process capability indices using clements' method for non-normal processes. J. Mult.-Valued Logic & Soft Comput., **22** (2014)
10. Hosseinifard, Z.; Abbasi, B.; Niaki, S.: Process capability estimation for leukocyte filtering process in blood service: a comparison study. IIE Trans. Healthc. Syst. Eng **4**(4), 167–177 (2014)
11. Leiva, V.; et al.: Capability indices for Birnbaum–Saunders processes applied to electronic and food industries. J. Appl. Statis. **41**(9), 1881–1902 (2014)
12. Senvar, O.; Sennaroglue, B.: Comparing performance of clements, Box–Cox, Johnson methods with Weibull Distributions for assessing process capability. J. Ind. Eng Manag. **9**(3), 634–656 (2016)
13. Clements, J.A.: Process capability calculations for non-normal distributions. Qual. Prog. **22**(9), 95–97 (1989)
14. Pearn, W.; Chen, K.: Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing. Microelectr. Reliab. **37**(12), 1853–1858 (1997)
15. Chien-Wei Wu, W.L.P.; Chang, C.S.; Chen, H.C.: Accuracy analysis of the percentile method for estimating non normal manufacturing quality. Commun. Statis. Simul. Comput **36**(3), 657–696 (2007)
16. Nanthakumar, D.C.; Vijayalakshmi, M.S.V.: Construction of Interquartile range (IQR) control chart using process capability for mean. Int. J. Mod. Sci. Eng. Technol **2**(10), 8 (2015)
17. Rodriguez, R.N.: Recent developments in process capability analysis. J. Qual. Technol **24**(4), 176–187 (1992)
18. Shu, M.-H., Hsu, B.-M., Han, C.-P.: Estimation of industrial process capabilities: some estimators to overcome the obstacle of non-normal distributions. In: proceedings IIE annual conference Citeseer. (2002)
19. Abu-Shawiesh, M.O.: A simple robust control chart based on MAD. J. Math. Statis. **4**(2), 102 (2008)
20. Shahriari, H.; Maddahi, A.; Shokouhi, A.H.: A robust dispersion control chart based on M-estimate. J. Ind. Syst. Eng. **2**(4), 297–307 (2009)
21. Adekeye, K.S.: Process capability indices based on median absolute deviation. Int. J. Appl., **3**(4), (2013)
22. Balamurali, S.: Bootstrap confidence limits for the process capability index Cpmk. Int. J. Qual. Eng. Technol. **3**(1), 79–90 (2012)
23. Chou, Y.-M.; Owen, D.; BORREGO, S.: A Lower confidence limits on process capability indices. J. Qual. Technol. **22**(3), 223–229 (1990)
24. Leiva, V.; et al.: New control charts based on the Birnbaum–Saunders distribution and their implementation. Revista Colombiana de Estadística **34**(1), 147–176 (2012)
25. Efron, B.: Bootstrap methods: another look at the jackknife. The Ann. Statis., p. 1–26, (1979)
26. Rousseeuw, P.J.; Croux, C.: Alternatives to the median absolute deviation. J. Am. Statis. Assoc. **88**(424), 1273–1283 (1993)
27. Haque, M.E.; Khan, J.A.: Globally robust confidence intervals for location. Dhaka Univ. J. Sci. **60**(1), 109–113 (2012)
28. Leys, C.; et al.: Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J. Exp. Soc. Psychol. **49**(4), 764–766 (2013)
29. Liu, P.-H.; Chen, F.-L.: Process capability analysis of non-normal process data using the Burr XII distribution. The Int. J. Adv. Manuf. Technol. **27**(9–10), 975–984 (2006)
30. Kashif, M.; et al.: Capability indices for non-normal distribution using Gini's mean difference as measure of variability. IEEE Access **4**, 7322–7330 (2016)
31. Kashif, M., et al.: Bootstrap confidence intervals of the modified process capability index for Weibull distribution. Arab. J. Sci. Engi., p. 1–9, (2017)
32. Tosasukul, J.; Budsaba, K.; Volodin, A.: Dependent bootstrap confidence intervals for a population mean. Thail. Statis. **7**(1), 43–51 (2009)
33. TANG, L.C.; THAN, S.E.: Computing process capability indices for non-normal data: a review and comparative study. Qual. Reliab. Engng. Int. **15**, 339–353 (1999)
34. Hsu, B.-M.: Robust estimators of process capability indices using smooth adaptive estimator. The university of Texas at Arlington, Arlington (2002)
35. Badar, M.G.P., A.M., Statistical aspects of fiber and bundlestrength in hybrid composites, In: progress in science andengineering composites. Hayashi, T., Kawata, K. and Umekawa,S., Eds. In: Tokyo. p. 1129–1136, (1982)
36. Raqab, M.Z., Kundu, D.: Comparison of different estimators of P [Y $\leq$ X] for a scaled burr type X distribution. Commun. Statist.–Simulation and Computation®, 34(2):465–483 (2005).
37. Kundu, D.; Raqab, M.Z.: Estimation of R= P (Y $\leq$ X) for three-parameter Weibull distribution. Statis. Probab. Lett. **79**(17), 1839–1846 (2009)
38. Delignette–Muller, M.L., Dutang, C.: fitdistrplus: an R package for fitting distributions. J. Stat. Softw. **64**(4), 1–34 (2015)