Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.1232

Published online in Wiley Online Library

Comparison of Pearson Distribution System and Response Modeling Methodology (RMM) as Models for Process Capability Analysis of Skewed Data

Michal Shauly*† and Yisrael Parmet

Clements' approach to process capability analysis for skewed distributions, based on fitting the Pearson distribution system to data, is widely used in industry. In this paper we compare the accuracy of the Pearson system and the RMM (response modeling methodology) distribution, as distributional models for process capability analysis of non-normal data. The accuracy of the estimates of C_p and C_{PU} is measured by the relative mean square errors. Three factors that may affect the accuracy of RMM and Pearson are examined: the data-generating distribution (Weibull, log-normal, gamma), the skewness (0.5, 1.25, 2) and the sample size (50, 300, 2000). The results show that RMM consistently outperforms Pearson, even for samples from gamma, which is a special case of Pearson. This implies that when observations are visibly skewed yet their underlying distribution is unknown, RMM estimators for C_p and C_{PU} take account of the information stored in the data more precisely than the Pearson model, and may therefore constitute a preferred distributional model to pursue in process capability analysis. Copyright © 2011 John Wiley & Sons, Ltd.

Keywords: Clements' method; relative mean square error; Pearson distribution system; process capability analysis; response modeling methodology

1. Introduction

Process capability analysis is an essential part of the statistical process control. The capability of a process to meet its specifications is commonly measured by process capability indices (PCI). Two widely used PCIs are C_p and C_{pk}, defined by the following equations:

$$C_p = \frac{USL-LSL}{6\sigma}, \tag{1}$$

$$C_{pk} = \min\left(C_{PL} = \frac{\mu-LSL}{3\sigma}, \quad C_{PU} = \frac{USL-\mu}{3\sigma}\right), \tag{2}$$

where USL and LSL are the upper and lower specification limits, respectively, and μ and σ are the process mean and standard deviation, respectively.

Notwithstanding their wide prevalence in implementations of process capability analysis, these two indices suffer from some problems, not the least of which is the underlying assumption that the true process distribution is normal. This assumption is often unrealistic, with dire consequences on the validity of the analysis results. Somerville and Montgomery[1] analyzed the errors that can occur in calculating C_p or C_{pk} for a non-normal distribution, by examining the symmetric heavy tailed t-distribution and the three commonly used skewed distributions—gamma, Weibull and log-normal. They have found meaningful errors in inferencing on the process fallout or PPM in nearly all examined cases.

To deal with this issue, several approaches have been advanced. Some researchers have suggested using normalizing transformations, like the Box–Cox[2] power transformations. Somerville and Montgomery[1] suggested using the square-root transformation or some other non-linear transformation. Rivera *et al.*[3] presented a continuous monotonically increasing function for which a normal distribution is a satisfactory approximation. Niaki and Abbasi[4] presented the best root transformation to transform skewed

Ben-Gurion University of the Negev, Industrial Engineering and Management, Beer-Sheva, Israel

Correspondence to: Michal Shauly, Ben-Gurion University of the Negev, Industrial Engineering and Management, Beer-Sheva, Israel.

†*E-mail: shauly@bgu.ac.il*

discrete multivariate distributions to multivariate normal. However, practitioners often tend to regard the transformation approach unfavorably, as they lose orientation regarding the measures in their problems, due to the change of scale. Another approach, when the distribution is skewed, is to construct indices for skewed processes based on the weighted standard deviation method[5]. This method decomposes the standard deviation of a characteristic into upper and lower standard deviations and adjusts the value of the PCI in accordance with the skewness, estimated from sample data.

The most general and practiced approach, which is also programmed in most current statistical software packages, is to select a multi-parameter distributional model for the (unknown) process distribution and use some fitting procedure to estimate the (true) endpoints of the process distribution. This is the widely practiced Clements' method[6], which replaces C_p (Equation (1)) and C_{pk} (Equation (2)) with

$$C_p = \frac{USL - LSL}{X_{0.99865} - X_{0.00135}} \tag{3}$$

and

$$C_{pk} = \min\left[C_{PL} = \frac{X_{0.5} - LSL}{X_{0.5} - X_{0.00135}}, C_{PU} = \frac{USL - X_{0.5}}{X_{0.99865} - X_{0.5}}\right], \tag{4}$$

where X_p is the $(100 \cdot p)$th percentile of the process quality characteristic, X. As a distributional model of this characteristic, Clements[6] proposed using the Pearson family of distributions. Pearn and Kotz[7] applied Clements' method to obtain estimators for two other PCIs, C_{pm} and C_{pmk}.

Although Clements' method is widely used and implemented in ISO standards (e.g. ISO-TR 2254-4[8]), some claim that this method can be inaccurate, especially when the underlying data distribution is extremely skewed (e.g. Wu *et al.*[9]). Therefore, several alternatives to the use of the Pearson family were suggested. Shore[10] introduced a new versatile family of distributions, Kotz and Lovelace[11] and Lin[12] used the folded normal distribution and Lin[13] used the generalized folded normal distribution as the fitting distribution. Liu and Chen[14] proposed the Burr XII family of distributions. Abbasi[15] used a multilayer perceptron neural network with one hidden layer to estimate PCI: skewness, kurtosis and upper specification limit are used as input variables, and PCIs are the output of the network.

While C_p and C_{pk} are appropriate measures for processes with two-sided specifications, C_{PU} and C_{PL} (Equation (4)) have been designed specifically for processes with one-sided specification limit. C_{PU} measures the capability of a smaller-the-better process with a *USL*, whereas C_{PL} measures the capability of a larger-the-better process with an *LSL*. Such processes, which have skewed distributions that are bounded on one side, occur frequently in industry (e.g. a shearing process and a chemical dip process[16], the microelectronics manufacturing[17]). Therefore, the C_{PU} and C_{PL} are widely used in industry. Kotz and Lovelace[11] studied this type of process using the log-normal distribution and proposed an index $C_{PU}(\ln)$. Kotz and Lovelace[11] and Lin[12] used the folded normal distribution on index C_{PU} for zero-bounded data. Albing and Vännman[18] proposed a new class of PCIs for quality characteristics with skewed zero-bounded distribution, and studied properties of two estimators of these PCIs. In the case of 'centered' right-skewed processes (i.e. in which $USL - X_{0.5} = X_{0.5} - LSL$), it is easy to show that C_{pk} equals C_{PU}. For this reason, we will focus on C_p and C_{PU} in this paper.

Recently, Shore[19] has introduced a new modeling methodology—response modeling methodology (RMM). The error distribution associated with the RMM model has been shown to deliver good representation to a large spectrum of skewed distributions[20–22]. Shauly *et al.*[23] compared the goodness-of-fit of the RMM distributional model and the Pearson distribution system by using simulated data, and found RMM to compete favorably with Pearson.

In this paper, we examine the accuracy of the RMM distributional model in estimating C_p and C_{PU}, relative to the Pearson model. Accuracy is measured by the relative *MSE* (mean square error divided by the squared value of the parameter). The latter has been found to be an appropriate measure for the comparison, as will be expounded in this paper.

The outline of the paper is as follows. In Section 2, we introduce the estimators of the PCIs investigated in this study, outline the Pearson and the RMM models and detail the estimation procedures for the studied indices. The simulation design is described in Section 3. The results are displayed in Section 4, and some conclusions are given in Section 5.

2. The PCIs and their estimation by Pearson and RMM

Suppose that the distribution of X is approximated by some parametric distributional model (in our simulation these are the RMM and the Pearson families of distributions). Estimators for C_p and C_{PU}, as these are given in Equations (3) and (4), are

$$\hat{C}_p = \frac{USL - LSL}{\hat{X}_{0.99865} - \hat{X}_{0.00135}}, \quad \hat{C}_{PU} = \frac{USL - \hat{X}_{0.5}}{\hat{X}_{0.99865} - \hat{X}_{0.5}}, \tag{5}$$

where $\hat{X}_p$ is a maximum likelihood estimator (MLE) for the $(100 \cdot p)$th quantile of X, derived from the fitted distribution (Pearson or RMM, in this paper). Namely, $\hat{X}_p = F^{-1}(p|\hat{\theta}_{\text{MLE}})$. We briefly introduce these two models in the following section.

2.1. The Pearson distribution system

In 1895, Pearson suggested a family of distributions, represented by the following differential equation:

$$\frac{\mathrm{d}f}{\mathrm{d}y}=\frac{a-y}{c_0+c_1y+c_2y^2}f, \tag{6}$$

where f is the density function. Solution of Equation (6) yields the general form of Pearson's density function

$$f_Y(y)=C(c_0+c_1y+c_2y^2)^{-1/2c_2}\exp\left[\frac{(c_1+2ac_2)\tan^{-1}\left(\frac{c_1+2c_2y}{\sqrt{4c_0c_2-c_1^2}}\right)}{c_2\sqrt{4c_0c_2-c_1^2}}\right], \tag{7}$$

where C is a normalizing parameter and $\{a,c_0,c_1,c_2\}$ are parameters that need to be estimated in the distribution fitting process. The form of Equation (7) depends on the roots of

$$c_0+c_1y+c_2y^2=0. \tag{8}$$

Denoting by $\{a_1,a_2\}$ the roots of Equation (8), various distributions, such as gamma, beta and t-distribution, may be obtained in terms of these roots. Further details are given in Weisstein[24] and in Johnson *et al.*[25].

2.2. RMM distribution

The RMM distributional model (detailed in Shore[18] and in Shauly *et al.*[23]) expresses the quantile function of a non-negative random variable, Y, in terms of the corresponding quantile of the standard normal variable, Z

$$Y_p=\exp\left\{\mu+\frac{A}{B}[\exp(BZ_p)-1]+CZ_p\right\}, \tag{9}$$

where $\{\mu,A,B,C\}$ are parameters that need to be determined, and $\{Z_p,Y_p\}$ are the $(100\cdot p)$th quantiles of Z and Y, respectively, namely

$$F_Y(Y_p)=\Phi(Z_p)=p \tag{10}$$

with F and Φ being the corresponding cumulative distribution functions (CDF). Note that for $Z_p=0{:}p=0.5$, and in this case $Y_{0.5}=\exp(\mu)$. Denoting $Y_{0.5}=M$ (the median of Y), we may rewrite Equation (9) as

$$Y_p=M\cdot\exp\left\{\frac{A}{B}[\exp(BZ_p)-1]+CZ_p\right\}. \tag{11}$$

Y_p is a proper quantile function and the PDF of Y exists if and only if $A>0$ and $C>0$ (see proof in[23]).

3. The simulation study—design

3.1. The experimental design

The aim of this study is to assess and compare the accuracy of the estimators $\hat{C}_p$ and $\hat{C}_{PU}$ when the underlying distributional model is either RMM or Pearson. Estimator's accuracy is measured by the relative *MSE*, MSE_R, defined for $\hat{C}_p$ by

$$MSE_R(\hat{C}_p)=\frac{MSE(\hat{C}_p)}{C_p^2}=\frac{E((\hat{C}_p-C_p)^2)}{C_p^2}=E\left(\left(\frac{\hat{C}_p}{C_p}-1\right)^2\right)=E\left(\left(\frac{X_{0.99865}-X_{0.00135}}{\hat{X}_{0.99865}-\hat{X}_{0.00135}}-1\right)^2\right). \tag{12}$$

The justification for this definition is that MSE_R does not depend on the specification limits. Thus, it allows a general comparison between various approximating models, irrespective of the specification limits . By contrast, the MSE_R of C_{PU} is not independent of *USL*:

$$MSE_R(\hat{C}_{PU})=\frac{MSE(\hat{C}_{PU})}{(C_{PU})^2}=E\left(\left(\frac{\hat{C}_{PU}}{C_{PU}}-1\right)^2\right)=E\left(\left(\frac{(USL-\hat{X}_{0.5})(X_{0.99865}-X_{0.5})}{(USL-X_{0.5})(\hat{X}_{0.99865}-\hat{X}_{0.5})}-1\right)^2\right). \tag{13}$$

However, it can be seen from Equation (4) that C_{PU} is a linear transformation of *USL*. Therefore, in a simulation study one should determine either *USL* or C_{PU} and calculate the other. For a given C_{PU} and a sampled distribution, the *USL* is calculated as follows:

$$USL=X_{0.5}+C_{PU}\cdot(X_{0.99865}-X_{0.5}). \tag{14}$$

The MSE_R of both estimators (Equations (12) and (13)) cannot be calculated analytically. Hence, a simulation study was carried out, in which three factors that may affect the MSE_R of RMM and Pearson were examined, each of them in three fixed values

Table I. The experimental design of the simulation study

	Skewness	Sample size	Data-generating model	
			Distribution	Parameters
1	0.5	50	Gamma	
2	0.5	300	Gamma	$k=16,\ \theta=10$
3	0.5	2000	Gamma	
4	0.5	50	Weibull	
5	0.5	300	Weibull	$k=2.2,\ \lambda=4$
6	0.5	2000	Weibull	
7	0.5	50	Log-normal	
8	0.5	300	Log-normal	$\mu=3,\ \sigma=0.165$
9	0.5	2000	Log-normal	
10	1.25	50	Log-normal	
11	1.25	300	Log-normal	$\mu=0,\ \sigma=0.382$
12	1.25	2000	Log-normal	
13	1.25	50	Gamma	
14	1.25	300	Gamma	$k=2.56,\ \theta=0.25$
15	1.25	2000	Gamma	
16	1.25	50	Weibull	
17	1.25	300	Weibull	$k=1.36,\ \lambda=10$
18	1.25	2000	Weibull	
19	2	50	Weibull	
20	2	300	Weibull	$k=1,\ \lambda=1$
21	2	2000	Weibull	
22	2	50	Log-normal	
23	2	300	Log-normal	$\mu=1,\ \sigma=0.55$
24	2	2000	Log-normal	
25	2	50	Gamma	
26	2	300	Gamma	$k=1,\ \theta=20$
27	2	2000	Gamma	

(in brackets); the data-generating distribution (Weibull, gamma, log-normal), the skewness of the distribution (0.5, 1.25, 2) and the sample size (50, 300, 2000). Altogether $3^3=27$ experimental conditions. Hundred samples were simulated for each condition, and the two PCIs were estimated for each sample, by using Pearson and RMM as distributional models (the fitting methods are detailed in Section 3.2). The Weibull, gamma and log-normal distributions have been selected due to their extensive use in practice. However, a major consideration was that gamma is a special case of Pearson, the log-normal is a special case of RMM and Weibull is not a special case of either. Thus, it may be constructive to find out how RMM fares relative to Pearson when the generating distribution is from the Pearson family, and conversely, how Pearson fares when this distribution belongs to the RMM family. The parameters of the data-generating distributions are displayed in Table II.

Based on Equation (12), the MSE_R of $\hat{C}_p$ was estimated from the simulation study samples by

$$\hat{MSE}_R(\hat{C}_p)=\left[\sum_{i=1}^{n}\left(\frac{\hat{C}_p}{C_p}-1\right)^2\right]\Big/n=\left[\sum_{i=1}^{n}\left(\frac{X_{0.99865}-X_{0.00135}}{\hat{x}_{0.99865,i}-\hat{x}_{0.00135,i}}-1\right)^2\right]\Big/n, \tag{15}$$

where $n=100$ (the number of simulated samples for each experimental condition), X_p is the $(100\cdot p)$th percentile of the sampled distribution and $\hat{x}_{p,i}$ is the ith sample estimate of that percentile.

Similarly, based on Equation (13) $\hat{MSE}_R(\hat{C}_{PU})$ is calculated from the simulation samples by

$$\hat{MSE}_R(\hat{C}_{PU})=\left[\sum_{i=1}^{n}\left(\frac{\hat{C}_{PU}}{C_{PU}}-1\right)^2\right]\Big/n=\left[\sum_{i=1}^{n}\left(\frac{(USL-\hat{x}_{0.5,i})(X_{0.99865}-X_{0.5})}{(USL-X_{0.5})(\hat{x}_{0.99865,i}-\hat{x}_{0.5,i})}-1\right)^2\right]\Big/n, \tag{16}$$

where USL for each sample is calculated by using Equation (14), for three values of C_{PU} (1.25, 1.6 and 2.0). These values were selected based on recommended minimum process capability for one-sided specification[26]. Correspondingly, the $\hat{MSE}_R$ of each selected value of C_{PU} was calculated, for each of the 27 combinations.

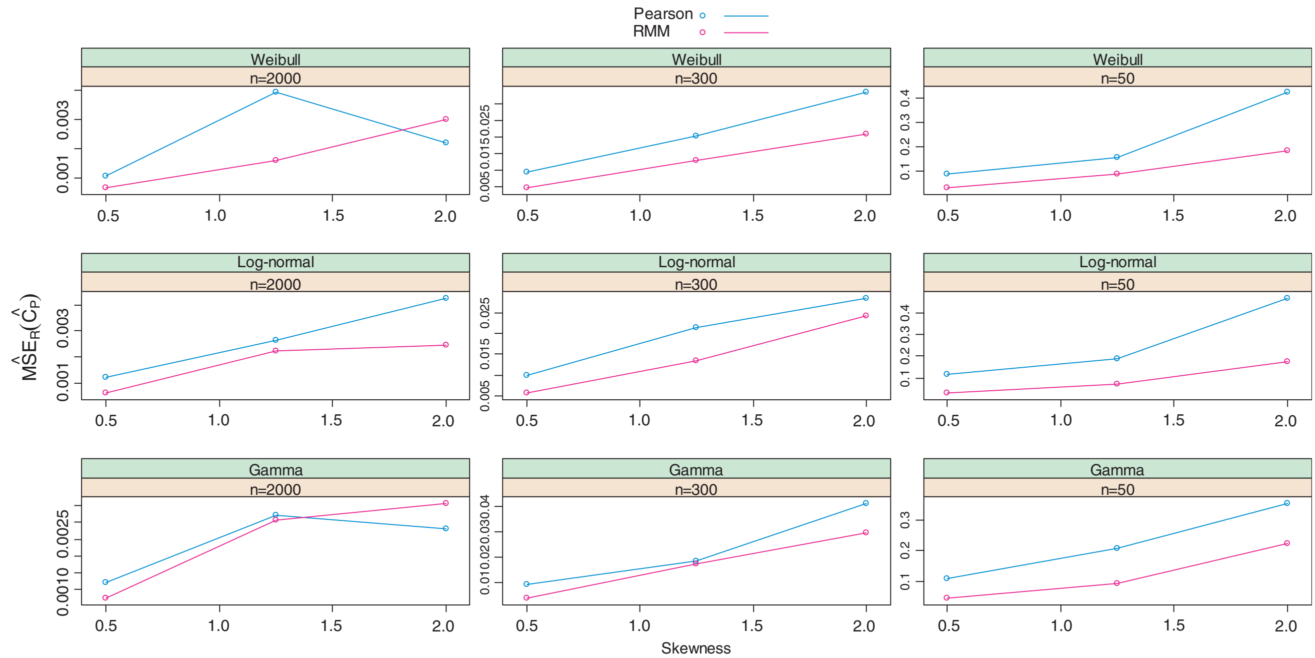

Figure 1. Profile plots of $\hat{MSE}_R(\hat{C}_p)$ by skewness and fitting model

3.2. Distribution fitting of Pearson and RMM

Unlike Clements[6] who adjusted the Pearson model to fit the first four moments of observed data, we fitted the Pearson distribution system to data by MLE, in order to avoid the lack of accuracy (high standard errors) associated with sample estimates of skewness and kurtosis. The Pearsonian system was fitted to data by the MLE *pearsonFitML* function in the *PearsonDS* package in the statistical software *R*[27]. This function finds MLEs for all sub-classes of the distribution system via numerical optimization, and returns the sub-class with the maximal likelihood value, along with the corresponding estimates. RMM was fitted to data by applying an ML procedure (presented in[23]) using *R*.

4. The simulation study—results

4.1. Accuracy of $\hat{C}_p$

Figure 1 shows nine profile plots, in which the values of $\hat{MSE}_R(\hat{C}_p)$ are presented by skewness (the *X* axis of each plot) and the fitting model (different lines). In each row of plots the generating distribution of the data is the same, and in each column the sample size is equal.

In samples from the Weibull distribution, RMM is notably superior than Pearson, for almost all sample sizes and skewness levels (only one exceptional case: $n=2000$, skewness$=2$). In samples from the log-normal distribution, RMM consistently outperforms Pearson. This evidence is not surprising, since the log-normal distribution is a special case of RMM. By contrast, the results for samples from gamma distribution are unexpected: RMM is more accurate than Pearson in almost all sample sizes and skewness levels (only one exceptional case: $n=2000$, skewness$=2$), although this distribution is a special case of Pearson (type 3). Two unsurprising results are that $\hat{MSE}_R(\hat{C}_p)$ of both RMM and Pearson increases with respect to skewness, and decreases with respect to sample size.

4.2. Accuracy of $\hat{C}_{PU}$

Figure 2 displays profile plots of $\hat{MSE}_R(\hat{C}_{PU})$, for the case of $C_{PU}=2.0$. The profile plots are assigned in the same manner as in Figure 1.

The results are very much compatible with the results for $\hat{MSE}_R(\hat{C}_p)$. In 25 scenarios (out of 27), RMM performs better than Pearson; in samples from the Weibull and gamma distributions, RMM is superior than Pearson in small and medium samples (for all skewness levels), and in the case of large samples, it is superior in lower skewness levels. In samples from the log-normal distribution, RMM remarkably outperforms Pearson in all cases. Similar results were obtained in the other two examined values of C_{PU} $(1.25, 1.6)$.

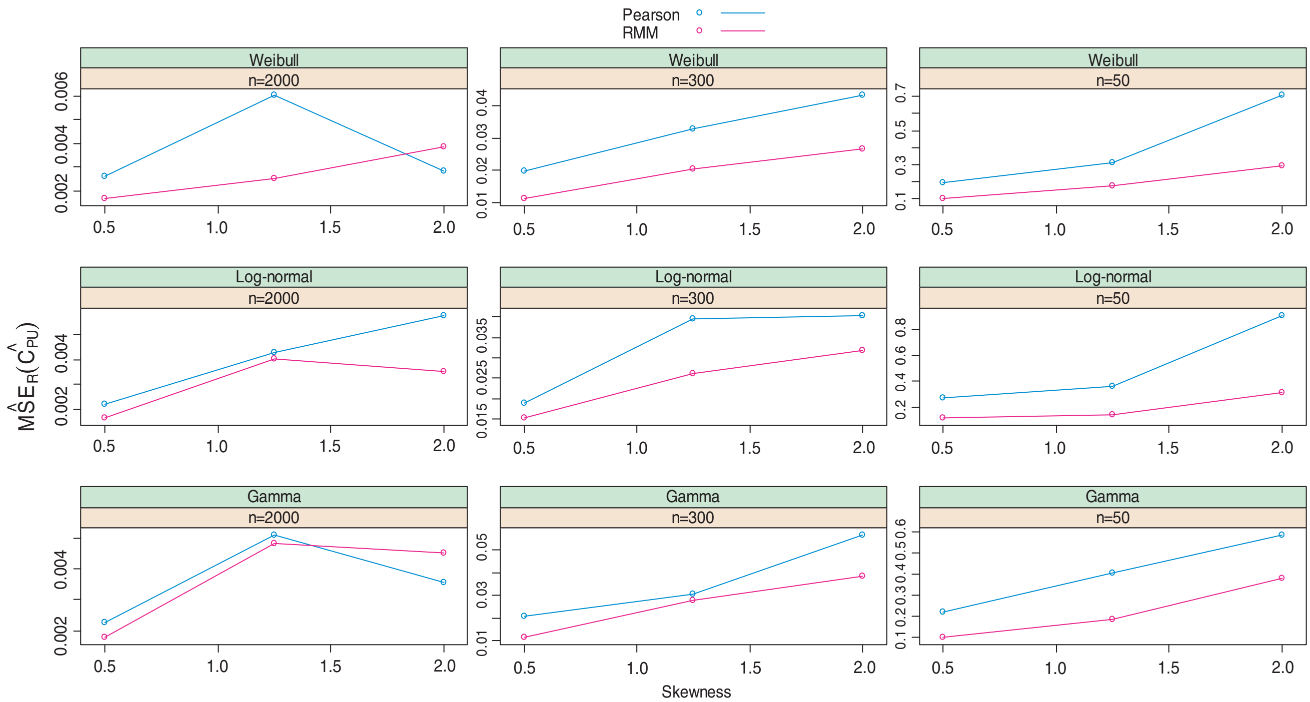

Figure 2. Profile plots of $\hat{MSE}_R(\hat{C}_{PU})$ for the case of $C_{PU}=2.0$ by skewness and fitting model

5. Conclusions

In this paper, we studied the accuracy of estimators for C_p and C_{PU}, in the case of skewed simulated samples, using two flexible parametric families of distributions, Pearson and RMM. The approach pursued was to first estimate (distribution fitting) the parameters of RMM and Pearson, and then use these estimates to evaluate the quantiles that appear in C_p and C_{PU}. The average of the square of the relative error, MSE_R, was compared for three generating distributions, three values of skewness, and three sample sizes. For both examined PCIs, RMM consistently produced more accurate estimates than those obtained by Pearson, even when the sampled distribution was gamma, which is a special case of Pearson. The main conclusion that one can derive from this study is that when the observations are visibly skewed yet their underlying distribution is unknown, RMM estimators for C_p and C_{PU} take account of the information stored in the data in a more precise fashion than the Pearson model. These findings corroborate the conclusions of Shauly *et al.*[23] that RMM better models skewed data, and may therefore constitute a preferred option to pursue in process capability analysis for skewed distributions.

Acknowledgements

The authors gratefully acknowledge the consultation of Professor Haim Shore from the Department of Industrial Engineering and Management at Ben-Gurion University.

References

1. Somerville SE, Montgomery DC. Process capability indices and non-normal distributions. *Quality Engineering* 1996; **9**:305–316.
2. Box GEP, Cox DR. An analysis of transformation. *Journal of the Royal Statistical Society, Series B* 1964; **26**:211–243.
3. Rivera LAR, Hubele NF, Lawrence FP. C_{pk} index estimation using data transformation. *Computers and Industrial Engineering* 1995; **29**:55–58.
4. Niaki STA, Abbasi B. Skewness reduction approach in multi-attribute process monitoring. *Communications in Statistics—Theory and Methods* 2007; **36**:2313–2325.
5. Chang YS, Choi IS, Bai DS. Process capability indices for skewed populations. *Quality and Reliability Engineering International* 2002; **18**:383–393.
6. Clements JA. Process capability calculations for non-normal distributions. *Quality Progress* 1989; **22**:95–100.
7. Pearn WL, Kotz S. Applications of Clements' method for calculating second and third generation process capability indices for non-normal Pearsonian populations. *Qualilty Engineering* 1994; **7**(1):139–145.
8. ISO/TR 22514-4: 2007(E). *Statistical Methods in Process Management—Capability and Performance—Part 4: Process Capability Estimates and Performance Measures* (1st edn). International Organization for Standardization (ISO).
9. Wu HH, Wang JS, Liu TL. Discussions of the Clements-based process capability indices. *Proceedings of the 1998 CIIE National Conference*, Chang-Hua, Taiwan, 1998; 561–566.
10. Shore H. A new approach to analyzing non-normal quality data with application to process capability analysis. *International Journal of Production Research* 1998; **36**(7):1917–1933.
11. Kotz S, Lovelace CR. *Introduction to Process Capability Indices in Theory and Practice*. Arnold: London, 1998.

12. Lin PC. The measurement of a process capability for folded normal process data. *Journal of Advanced Manufacturing Technology* 2004; **24**:223–228.
13. Lin PC. Application of the generalized folded-normal distribution to the process capability measures. *Journal of Advanced Manufacturing Technology* 2005; **26**:825–830.
14. Liu PH, Chen FL. Process capability analysis of non-normal process data using the Burr XII distribution. *International Journal of Advanced Manufacturing Technology* 2006; **27**:975–984.
15. Abbasi B. A neural network applied to estimate process capability of non-normal processes. *Expert Systems with Applications* 2009; **36**: 3093–3100.
16. Pyzdek T. Process capability analysis using personal computers. *Quality Engineering* 1992; **4**(3):419–440.
17. Lin PC, Pearn WL. Testing process capability for one-sided specification limit with application to the voltage level translator. *Microelectronics Reliability* 2002; **42**:1975–1983.
18. Albing M, Vännman K. Skewed zero-bound distributions and process capability indices for upper specifications. *Journal of Applied Statistics* 2009; **36**(2):205–221.
19. Shore H. *Response Modeling Methodology—Empirical Modeling for Engineering and Science*. World Scientific Publishing: Singapore, 2005.
20. Shore H. Comparison of Generalized Lambda Distribution (GLD) and Response Modeling Methodology (RMM) as general platforms for distribution fitting. *Communications in Statistics—Theory and Methods* 2007; **36**(15):2805–2819.
21. Shore H. Distribution fitting with the quantile function of Response Modeling Methodology. *Handbook of Fitting Statistical Distributions with R*, Karian ZA, Dudewicz EJ (eds.). Taylor and Francis: London, 2010; 537–556.
22. Shore H, A'wad F. Statistical comparison of the goodness-of-fit delivered by five families of distributions used in distribution fitting. *Communication in Statistics—Theory and Methods* 2010; **39**(10):1707–1728.
23. Shauly M, Shore H, Parmet Y. A semi-MLE algorithm for RMM distribution and a comparison of the goodness-of-fit of RMM distribution and the Pearson distribution system. *Proceedings of the 2010 Joint Statistical Meetings*, Vancouver, Canada, 2010; 3574–3587.
24. Weisstein EW. Pearson System, From MathWorld—A Wolfram Web Resource. http://www.mathworld.wolfram.com/PearsonSystem.html (2011).
25. Johnson NL, Kotz S, Balakrishnan N. *Distributions in Statistics II*: *Continuous Univariate Distributions*. Wiley: New York, 1995.
26. Montgomery DC. *Introduction to Statistical Quality Control* (5th edn). Wiley: New York, 2005.
27. R Development Core Team 2009. *R*: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org (2011).

Authors' biographies

Michal Shauly —Ph.D. student at the Department of Industrial Engineering and Management in the Ben-Gurion University of the Negev, Israel. She holds B.Sc. and M.Sc. degrees in Industrial Engineering & Management from the Ben-Gurion University of the Negev. She specializes in statistical modeling and biostatistics.

Yisrael Parmet —Senior Lecturer at the Department of Industrial Engineering and Management in Ben-Gurion University of the Negev, Israel. He holds a B.A. degree in Economics and Statistics and M.Sc. and Ph.D. degrees in Statistics from the Tel-Aviv University. He specializes in design of experiments and statistical modeling. During his studies he served as a research assistant at the statistical laboratory at the Department of Statistics and OR, Tel Aviv University, which granted him knowledgeableness in practical data analysis. In 2007–2008 Dr. Parmet was a visiting professor at the Department of Dermatology and Cutaneous Surgery in the UM Miller School of Medicine.