QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2007; **23**:597–614
Published online 7 November 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/qre.836

Research

Measuring Process Capability Based on C_{pmk} with Gauge Measurement Errors

B. M. Hsu[1], M. H. Shu[2,*,†] and W. L. Pearn[3]

[1]*Department of Industrial Engineering and Management, Cheng Shiu University, 840 Cheng Cing Road, Niaosong, Kaohsiung 83347, Taiwan, Republic of China*
[2]*Department of Industrial Engineering and Management, National Kaohsiung University of Applied Sciences, 415 Chien Kung Road, Seng-Min District, Kaohsiung 80778, Taiwan, Republic of China*
[3]*Department of Industrial Engineering and Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, Republic of China*

Due to their effectiveness and simplicity of use, the process capability indices C_p, C_{pk}, and C_{pm} have been popularly accepted in the manufacturing industry as management tools for evaluating and improving process quality. Combining the merits of those indices, a more advanced index, C_{pmk}, is proposed that takes into account process variation, process centering, and the proximity to the target value, and has been shown to be a very useful index for manufacturing processes with two-sided specification limits. Most research works related to C_{pmk} assume no gauge measurement errors. However, such an assumption inadequately reflects real situations even when highly advanced measurement instruments are employed. Conclusions drawn regarding process capability are therefore unreliable and misleading. In this paper, we conduct a sensitivity investigation for the process capability index C_{pmk} in the presence of gauge measurement errors. We consider the use of capability testing of C_{pmk} as a method for obtaining lower confidence bounds and critical values for true process capability when gauge measurement errors are unavoidable. The results show that using the estimator with sample data contaminated by measurement errors severely underestimates the true capability, resulting in an imperceptibly smaller test power. To measure the true process capability, three methods for the adjusted confidence bounds are presented and their performances are compared using computer simulation. Copyright © 2006 John Wiley & Sons, Ltd.

Received 4 January 2006; Revised 2 March 2006

KEY WORDS: gauge measurement error; lower confidence bound; critical value; process capability indices; generalized confidence interval

*Correspondence to: M. H. Shu, Department of Industrial Engineering and Management, National Kaohsiung University of Applied Sciences, 415 Chien Kung Road, Seng-Min District, Kaohsiung 80778, Taiwan, Republic of China.
†E-mail: workman@cc.kuas.edu.tw

Contract/grant sponsor: National Science Council of Taiwan, R.O.C.; contract/grant number: NSC92-2213-E-251-005

Copyright © 2006 John Wiley & Sons, Ltd.

WILEY InterScience®
DISCOVER SOMETHING GREAT

1. INTRODUCTION

Process capability indices, including C_p, C_{pk}, and C_{pm} (Kane[1], Chan *et al.*[2], Zimmer *et al.*[3], Vännman and Hubele[4], Perakis and Xekalaki[5]), have been proposed in the manufacturing industry to provide numerical measures on whether a process is capable of reproducing items meeting the manufacturing quality requirement preset in the factory. Combining the merits of those indices, Pearn *et al.*[6] proposed a more advanced capability index called C_{pmk}, which has been shown to be a useful capability index for processes with two-sided specification limits. Those indices are defined as

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

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

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} \tag{3}$$

$$C_{pmk} = \min\left\{\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right\} \tag{4}$$

where USL is the upper specification limit, LSL is the lower specification limit, μ is the process mean, σ is the process standard deviation, and T is the target value predetermined by the product designer or the manufacturing engineer.

Process variation (product quality consistency), process departure (targeting), process yield, and process loss (relating to product reliability) are considered crucial benchmarks for measuring manufacturing quality. The index C_p measures only the distribution spread (process consistency/precision), which only reflects the consistency of the product quality characteristic. The yield-based index C_{pk} provides lower bounds on process yield by taking the process location into consideration, which offsets some of the weaknesses in C_p but can fail to distinguish between on-target and off-target processes (Hoffman[7]). The index C_{pm} takes the proximity of process mean from the target value into account, which is more sensitive to process departure than C_{pk}. Since the design of C_{pm} is based on the average process loss relative to the manufacturing tolerance, the index C_{pm} provides an upper bound on the average process loss.

The index C_{pmk} is constructed by combining the modifications to C_p that produced C_{pk} and C_{pm}, and therefore inherits the merits of both indices. We note that a manufacturing process satisfying the capability requirement '$C_{pk} \geq c$' may not satisfy the capability condition '$C_{pm} \geq c$'. On the other hand, a process satisfying the capability requirement '$C_{pm} \geq c$' may not satisfy the capability requirement '$C_{pk} \geq c$'. However, a manufacturing process does satisfy both capability requirements '$C_{pk} \geq c$' and '$C_{pm} \geq c$' if the process satisfies the capability requirement '$C_{pmk} \geq c$' since $C_{pmk} \leq C_{pk}$ and $C_{pmk} \leq C_{pm}$. Thus, the index C_{pmk} provides a greater level of quality assurance with respective to process yield and process loss to the customers than the other two indices. This is a desired property according to today's modern quality theory, as a reduction of process loss (variation from the target) is just as important as increasing process yield (meeting the specifications). While C_{pk} remains the more popular and widely used index, C_{pmk} is considered a very useful index for processes with two-sided manufacturing specifications. For semiconductor or microelectronics manufacturing in particular, C_{pmk} is an appropriate index for capability measurement due to the high standard and stringent requirements on product quality and reliability.

2. THE INDEX C_{pmk} AND THE GAUGE MEASUREMENT ERROR

Most research works related to C_{pmk} have assumed no gauge measurement errors. For example, Chen and Hsu[8] investigated the asymptotic sampling distribution of the estimated C_{pmk}. Wright[9] derived an explicit but rather complicated expression for the probability density function of the estimated C_{pmk}. Jessenberger and Weihs[10]

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

Table I. Guidelines for gauge capabilities

Gauge capability	Result
$\lambda < 0.1$	Gauge system OK
$0.1 < \lambda < 0.3$	May be acceptable based on importance of application, cost of gauge, cost of repair, and so on
$0.3 < \lambda$	Gauge system needs improvement; make every effort to identify the problems and have them corrected

studied the behavior of C_{pmk} for processes with asymmetric tolerances. Pearn *et al.*[11] obtained an alternative but simpler form of the probability density function of the estimated C_{pmk} and considered the capability testing based on C_{pmk}. Pearn and Lin[12] and Pearn and Shu[13] developed efficient Maple/Matlab computer programs to calculate the critical values, the p-value, and the lower confidence bounds for estimating and testing process capability based on C_{pmk}.

However, capability analysis with no gauge measurement errors cannot reflect real situations closely, even with highly sophisticated and advanced measurement instruments (Bordignon and Scagliarini[14,15]). Any measurement error has some impact on the determination of true measurement systems and process capabilities. Montgomery and Runger[16,17] and Burdick *et al.*[18] noted that the quality of the collected data relies heavily on the gauge accuracy. Clearly, conclusions drawn regarding process capability, based on the empirical index values calculated from data contaminated with gauge measurement errors, are highly unreliable. To analyze the effect of gauge measurement errors on the true capability measure, Mittag[19,20] and Levinson[21] quantified the percentage error on process capability evaluation in the presence of gauge measurement errors. Bordignon and Scagliarini[14] presented the statistical analysis on the estimation of confidence intervals for C_p with data contaminated with measurement errors.

Suppose that the measurement errors can be described as a random variable $M \sim N(0, \sigma_M^2)$. Montgomery and Runger[16,17] expressed the gauge capability as

$$\lambda = \frac{6\sigma_M}{USL - LSL} \tag{5}$$

For the measurement system to be deemed acceptable, the measurement variability due to the measurement system must be less than a predetermined percentage of the engineering tolerance. Montgomery[22] noted that the automotive industry action group recommended the guidelines for gauge acceptance given in Table I. In this paper, the gauge capability, λ, in Equation (5) provided by the gauge manufacturing factory is assumed to be known.

The organization of this paper is as follows. In Section 3, we consider the sensitivity of the C_{pmk} index with gauge measurement errors. In Sections 4 and 5, the sampling distribution and bias and mean squared error (MSE) are studied when using $\hat{C}_{pmk}^G$ as an estimator for C_{pmk}. In Sections 6 and 7, we show that a large measurement error will cause underestimation of the true process capability. In Section 8 sampling distribution (standard distribution (SD) and mean squared distribution (MSD)) approaches and generalized confidence intervals (GCIs) approach are proposed to establish more reliable lower confidence bounds. In Section 9, a simulation study is conducted for the performance comparison of these methods. Section 10 concludes the paper.

3. EMPIRICAL PROCESS CAPABILITY C_{pmk}^G

Suppose that $X \sim N(\mu, \sigma^2)$ represents the quality characteristic of the manufacturing process under investigation. In practice, the observed variable G (with gauge measurement errors) is measured rather than the true variable X. Assume that X and M are stochastically independent, then we have $G \sim N(\mu, \sigma_G^2 = \sigma^2 + \sigma_M^2)$ and the empirical process capability index C_{pmk}^G is obtained after substituting σ_G for σ. The relationship between

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

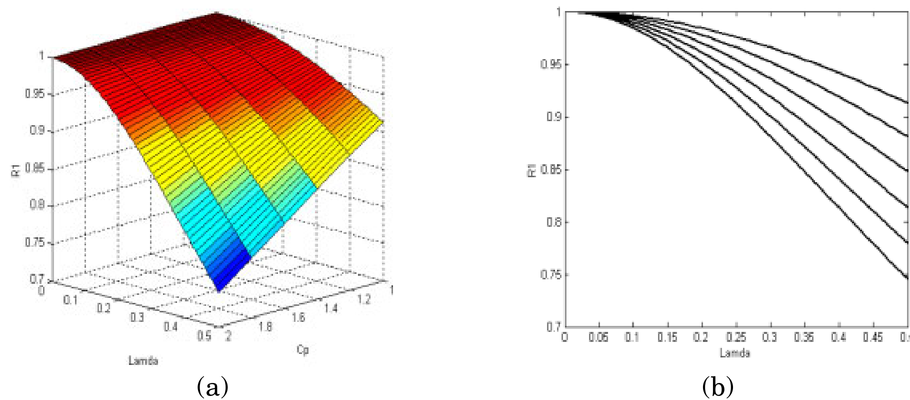

Figure 1. (a) Surface plot and (b) plots of R_1 versus λ in [0,0.5] for $C_p = 1.0(0.2)2.0$ with $\xi = 0.5$

the true process capability, C_{pmk}, and the empirical process capability, C_{pmk}^G, can be expressed as

$$\frac{C_{pmk}^G}{C_{pmk}} = \frac{\sqrt{1+\xi^2}}{\sqrt{1+\lambda^2 C_p^2 + \xi^2}} \tag{6}$$

where $\xi = (\mu - T)/\sigma$. Since the variation of the observed data is larger than the variation of the original data, the denominator of the index C_{pmk} becomes larger and the true capability of the process is understated if the empirical data G are used.

Figure 1(a) displays the surface plot of the ratio $R_1 = C_{pmk}^G / C_{pmk}$ for λ in [0,0.5] with $C_p \in [1, 2]$. Figure 1(b) plots the ratio R_1 versus λ for $C_p = 1.0(0.2)2.0$. These figures show that the measurement errors result in a decrease in the estimated value. A small process variation has the same effect as the presence of a measurement error. Since R_1 is small if λ becomes large, the gauge becomes more important as the true capability improves. For instance, if $\lambda = 0.5$, $C_p = 2$, and $\xi = 0.5$ (the ratio $R_1 = 0.7454$), $C_{pmk}^G = 0.7454$ with $C_{pmk} = 1$ and $C_{pmk}^G = 1.8634$ with $C_{pmk} = 2.50$. The empirical process capability diverges from the true process capability as measurement errors increase.

4. SAMPLING DISTRIBUTION OF $\hat{C}_{pmk}^G$

In practice, sample data must be collected in order to estimate the empirical process capability C_{pmk}^G. For a stably normal process, the empirical data (observed data contaminated with errors) G_i, for $i = 1, 2, \ldots, n$, is collected. The maximum likelihood estimator (MLE) of C_{pmk}^G is defined as

$$\hat{C}_{pmk}^G = \min\left\{\frac{USL - \bar{G}}{3\sqrt{\tilde{S}_n^2 + (\bar{G} - T)^2}}, \frac{\bar{G} - LSL}{3\sqrt{\tilde{S}_n^2 + (\bar{G} - T)^2}}\right\} \tag{7}$$

where $\bar{G} = \sum_{i=1}^n G_i/n$ and $\tilde{S}_n^2 = \sum_{i=1}^n (G_i - \bar{G})^2/n$ are the MLEs of μ and σ_G^2. We note that the statistic $\tilde{S}_n^2 + (\bar{G} - T)^2 = \sum_{i=1}^n (G_i - T)^2/n$ in the denominator of $\hat{C}_{pmk}^G$ is the uniformly minimum variance unbiased estimator of $\sigma_G^2 + (\mu - T)^2 = E[(G - T)^2]$. For processes with a symmetric manufacturing tolerance ($T = m$), the estimator $\hat{C}_{pmk}^G$ can alternatively be expressed as follows

$$\hat{C}_{pmk}^G = \frac{d - |\bar{G} - m|}{3\sqrt{\tilde{S}_n^2 + (\tilde{G} - T)^2}} \tag{8}$$

where $d = (USL - LSL)/2$.

Copyright © 2006 John Wiley & Sons, Ltd.

Obviously, if $\sigma_M = 0$, then the empirical process capability, C^G_{pmk}, reduces to the basic index, C_{pmk}. Pearn *et al.*[6] considered the MLE of C_{pmk}, expressed as

$$\hat{C}_{pmk} = \min\left\{\frac{USL - \bar{X}}{3\sqrt{S_n^2 + (\bar{X} - T)^2}}, \frac{LSL - \bar{X}}{3\sqrt{S_n^2 + (\bar{X} - T)^2}}\right\} = \frac{d - |\bar{X} - T|}{3\sqrt{S_n^2 + (\bar{X} - T)^2}} \tag{9}$$

where $\bar{X} = \sum_{i=1}^{n} X_i/n$ and $S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/n$. Using the same technique as Vännman[22] and Shu and Chen[23], the cumulative distribution function (CDF) of $\hat{C}^G_{pmk}$ can be expressed in terms of a mixture of the Chi-square distribution and the normal distribution

$$F_{\hat{C}^G_{pmk}}(x) = 1 - \int_0^{b_G\sqrt{n}/(1+3x)} F_K\left(\frac{(b_G\sqrt{n} - t)^2}{9x^2} - t^2\right)[\phi(t + \xi_G\sqrt{n}) + \phi(t - \xi_G\sqrt{n})]\,\mathrm{d}t, \tag{10}$$

for $x > 0$, where $b_G = d/\sigma_G = 3C^G_p$, $F_K(\cdot)$ is the CDF of the ordinary central Chi-square distribution χ^2_{n-1} and $\phi(\cdot)$ is the probability density function (PDF) of the standard normal distribution $N(0, 1)$ where

$$\xi_G = \frac{\mu - m}{\sigma_G} = \frac{-C^G_p/3 + \sqrt{(C^G_p)^2/9 + [(C^G_{pmk})^2 - 1/9][(C^G_p)^2 - (C^G_{pmk})^2]}}{(C^G_{pmk})^2 - 1/9}$$

$$C^G_p = \frac{C_p}{\sqrt{1 + \lambda^2 C_p^2}}, \quad C^G_{pmk} = \frac{C_{pmk}\sqrt{1 + \xi^2}}{\sqrt{1 + \lambda^2 C_p^2 + \xi^2}}$$

The rth moment of the estimator $\hat{C}^G_{pmk}$ can be obtained (Vännman[23] and Shu and Chen[24]) as

$$E[(\hat{C}^G_{pmk})^r] = 3^{-r}\sum_{i=0}^{r}(-1)^i\binom{r}{i}\left(\frac{D}{\sqrt{2}}\right)^{r-i}\sum_{\ell=0}^{\infty}\frac{\beta^\ell e^{-\beta/2}}{2^\ell \ell!} \times \frac{\Gamma(c-a)\Gamma(b)}{\Gamma(b - i/2)\Gamma(c)} \tag{11}$$

where $D = \sqrt{n}d/\sigma_G$, $\beta = n(\mu - T)^2/\sigma_G^2$, $a = r/2$, $b = (1 + i + j)/2$, and $c = (n + i + j)/2$.

Obviously, if the $\sigma_M = 0$, then the CDF of $\hat{C}_{pmk}$ reduces to

$$F_{\hat{C}_{pmk}}(x) = 1 - \int_0^{b\sqrt{n}/(1+3x)} F_K\left(\frac{(b\sqrt{n} - t)^2}{9x^2} - t^2\right)[\phi(t + \xi\sqrt{n}) + \phi(t - \xi\sqrt{n})]\,\mathrm{d}t \tag{12}$$

for $x > 0$, where $b = d/\sigma = 3C_p$ and

$$\xi = \frac{\mu - m}{\sigma} = \frac{-C_p/3 + \sqrt{C_p^2/9 + [C_{pmk}^2 - 1/9][C_p^2 - C_{pmk}^2]}}{C^2_{pmk} - 1/9}$$

5. BIAS AND THE MSE ANALYSIS

To investigate how measurement errors may affect the sampling distribution, we conduct bias and MSE analyses. Noting that from the expression $C_{pmk} = (C_p - |\xi|/3)/\sqrt{1 + \xi^2}$, Pearn and Shu[13] and Pearn and Lin[12] show that the lower confidence bounds and critical values for C_{pmk} can be obtained by setting $\xi = 0.5$ (for test reliability purposes). We then set $C_{pmk} = (C_p - 1/6)/\sqrt{1.25}$ and consider cases of $(C_p, C_{pmk}) = (1.285, 1.00)$ and $(1.844, 1.50)$ as examples. Figures 2(a) and (b) plot the bias of $\hat{C}^G_{pmk}$ versus $n = 5(5)100$ with $\lambda = 0(0.1)0.5$ for $C_p = 1.285$ and $C_{pmk} = 1.00$ and $C_p = 1.844$ and $C_{pmk} = 1.5$, respectively. Note that when $\lambda = 0$, the bias of $\hat{C}^G_{pmk}$ is equal to the bias of $\hat{C}_{pmk}$, and the bias of $\hat{C}^G_{pmk}$ increases as λ increases or n decreases. Figures 3(a) and (b) are the surface plots of the ratio $R_2 = MSE(\hat{C}^G_{pmk})/MSE(\hat{C}_{pmk})$ with $n = 5(5)100$ and λ in [0,0.5] for $C_p = 1.285$ and $C_{pmk} = 1.00$ and $C_p = 1.844$ and $C_{pmk} = 1.50$, respectively. The maximum values of R_2 in Figures 3(a) and (b) occur at $(n, \lambda) = (100, 0.5)$ and the minimum values of R_2 occur at $(n, \lambda) = (5, 0.5)$.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

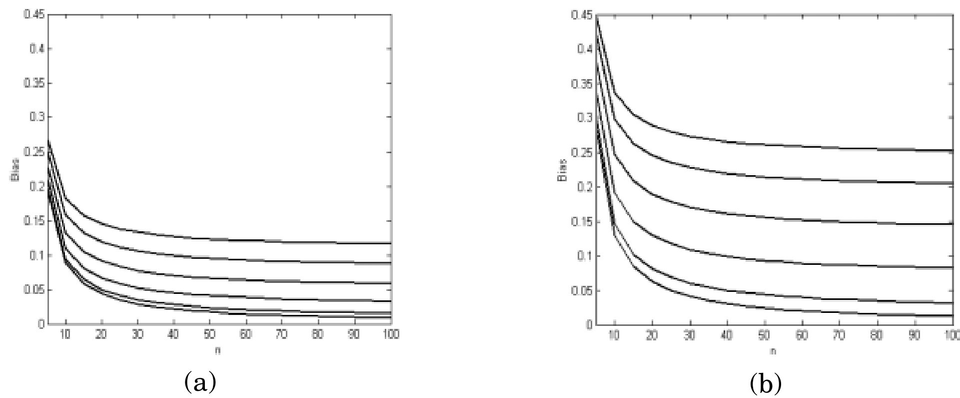

Figure 2. Plots of the bias of $\hat{C}_{pmk}^{G}$ for $n = 5(5)100$, $\lambda = 0(0.1)0.5$ (bottom to top): (a) $C_p = 1.285$ and $C_{pmk} = 1.00$; (b) $C_p = 1.844$ and $C_{pmk} = 1.5$

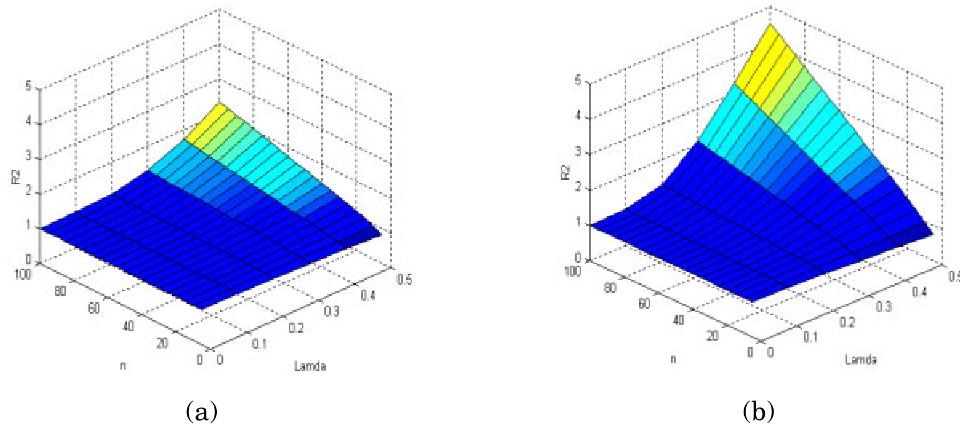

Figure 3. Surface plot of R_2 with $n = 5(5)100$ and λ in [0,0.5] for (a) $C_p = 1.285$ and $C_{pmk} = 1.00$; (b) $C_p = 1.844$ and $C_{pmk} = 1.50$

6. LOWER CONFIDENCE BOUND BASED ON $\hat{C}_{pmk}^{G}$

The lower confidence bounds estimate the minimum process capability based on sample data. To find reliable $100\gamma\%$ lower confidence bound L for C_{pmk} (where γ is the confidence level), Pearn and Shu[13] solved the following equation:

$$\int_0^{b\sqrt{n}/(1+3\hat{C}_{pmk})} F_K\left(\frac{(b\sqrt{n}-t)^2}{9\hat{C}_{pmk}^2} - t^2\right)[\phi(t+\xi\sqrt{n}) + \phi(t-\xi\sqrt{n})]\,\mathrm{d}t = 1-\gamma \qquad (13)$$

Note that the parameter b can be expressed as $b = 3C_p = 3L\sqrt{1+\xi^2} + |\xi|$. Since the process parameters μ and σ are unknown, then the distribution parameter $\xi = (\mu - m)/\sigma$ is also unknown. To eliminate the need for further estimation of the distribution characteristic parameter ξ, Pearn and Shu[13] investigated the behavior of the lower confidence bound, L, against ξ. They performed extensive calculations to obtain the lower confidence bound values for $\xi = 0(0.05)3.00$, $\hat{C}_{pmk} = 0.7(0.1)3.0$, and $n = 10(5)200$, and found that the lower confidence bound obtains its minimum at $\xi = 0.5$ in all cases. Thus, for practical purposes they recommended solving Equation (13) with $\xi = 0.5$ to obtain the required lower confidence bounds, without having to further estimate ξ.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

In practice, the observed sample data are contaminated with errors. Thus, $\hat{C}^G_{pmk}$ is substituted into Equation (13) with $\xi_G = 0.5$ to obtain the confidence bounds, which can be written as

$$\int_0^{b_G\sqrt{n}/(1+3\hat{C}^G_{pmk})} F_K\left(\frac{(b_G\sqrt{n}-t)^2}{9(\hat{C}^G_{pmk})^2} - t^2\right)[\phi(t+0.5\sqrt{n}) + \phi(t-0.5\sqrt{n})]\,\mathrm{d}t = 1-\gamma \tag{14}$$

We denote the bound originated from $\hat{C}^G_{pmk}$ as L_G where $b_G = 3C^G_p = 3L_G\sqrt{1.25} + 0.5$.

The confidence coefficient for the lower confidence bound L_G (denoted as γ_G) is

$$\gamma_G = 1 - \int_0^{b_G\sqrt{n}/(1+3\hat{C}^G_{pmk})} F_K\left(\frac{(b_G\sqrt{n}-t)^2}{9(\hat{C}^G_{pmk})^2} - t^2\right)[\phi(t+\hat{\xi}_G\sqrt{n}) + \phi(t-\hat{\xi}_G\sqrt{n})]\,\mathrm{d}t \tag{15}$$

where $b_G = 3L_G\sqrt{1+\hat{\xi}^2_G} + |\hat{\xi}_G|$ and γ_G is no less than γ.

Figures 4(a)–(d) plot L_G versus $\lambda \in [0, 0.5]$ with $n = 30, 50, 70, 100, 150$ for $\hat{C}_{pmk} = 1.00, 1.50$ and $\hat{C}_p = \hat{C}_{pmk} + R_3$, $R_3 = 0.285$ and 0.50 with 95% confidence level. It should be noted that for sufficiently large sample size, we have

$$\hat{C}^G_{pmk} = \frac{\hat{C}_{pmk}\sqrt{1.25}}{\sqrt{1.25 + \lambda^2\hat{C}^2_p}} \tag{16}$$

Therefore, we set $\hat{C}^G_{pmk} = \hat{C}_{pmk}\sqrt{1.25}\Big/\sqrt{1.25+\lambda^2\hat{C}^2_p}$ to obtain $\hat{C}^G_{pmk}$. We see that in Figures 4(a)–(d), L_G decreases in λ, especially for large $\hat{C}_p$ values, and the reduction of L_G is more significant for large $\hat{C}_{pmk}$. A large measurement error results in significant underestimation of true process capability.

In current practice, a process is called 'inadequate' if $C_{pmk} < 1.00$, 'marginally capable' if $1.00 \le C_{pmk} < 1.33$, 'satisfactory' if $1.33 \le C_{pmk} < 1.50$, 'excellent' if $1.50 \le C_{pmk} < 2.00$, and 'super' if $2.00 \le C_{pmk}$. If capability measures do not include the measurement errors, a significant underestimation of the true process capability may result in high production costs, reducing competitiveness. For instance, suppose that a process has a 95% lower confidence bound, 1.211 ($\hat{C}_{pmk} = 1.50$) with $n = 50$, which meets the threshold of an 'excellent' process. However, the bound may be calculated as 0.997 with measurement errors $\lambda = 0.42$ and the process is determined as 'inadequate'.

7. TESTING PROCESS CAPABILITY BASED ON $\hat{C}^G_{pmk}$

To determine whether a given process meets the preset capability requirement, we could consider statistical testing with the null hypothesis $H_0 : C_{pmk} \le c$ (the process is not capable) versus the alternative $H_1 : C_{pmk} > c$ (the process is capable), where c is the required process capability. If the calculated process capability is greater than the corresponding critical value, we reject the null hypothesis and conclude that the process is capable. The test rejects the null hypothesis $H_0 : C_{pmk} \le c$ if $\hat{C}_{pmk} \ge c_0$ with type I error α (α-risk), which is the chance of incorrectly concluding an incapable process (with $C_{pmk} \le c$) as capable (with $C_{pmk} > c$). That is, $P(\hat{C}_{pmk} \ge c_0 \mid C_{pmk} = c) = \alpha$. Given values of the capability requirement c, sample size n, and risk α, the critical value c_0 can be obtained by solving Equation (17) using available numerical methods:

$$\int_0^{b\sqrt{n}/(1+3c_0)} F_K\left(\frac{(b\sqrt{n}-t)^2}{9c_0^2} - t^2\right)[\phi(t+\xi\sqrt{n}) + \phi(t-\xi\sqrt{n})]\,\mathrm{d}t = \alpha \tag{17}$$

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

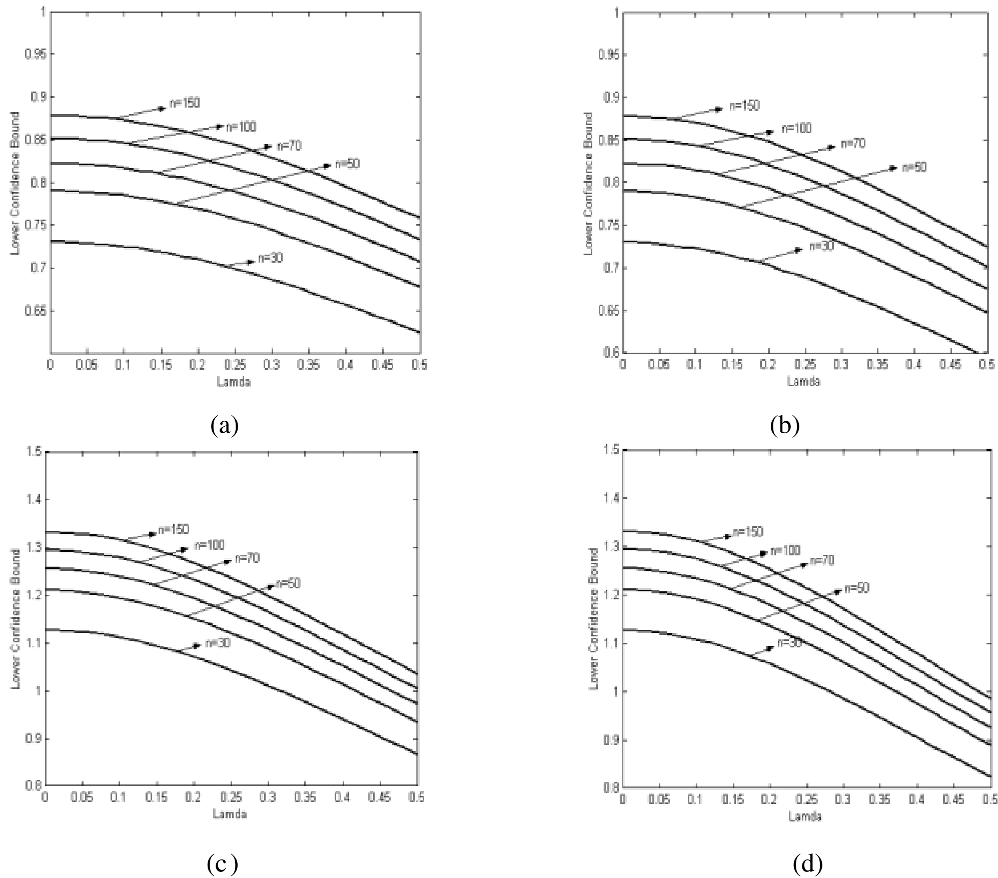

Figure 4. Plot of L_G versus λ with $n = 30, 50, 70, 100, 150$ with 95% confidence level: (a) $\hat{C}_p = 1.285$, $\hat{C}_{pmk} = 1.00$; (b) $\hat{C}_p = 1.50$, $\hat{C}_{pmk} = 1.00$; (c) $\hat{C}_p = 1.785$, $\hat{C}_{pmk} = 1.50$; (d) $\hat{C}_p = 2.00$; $\hat{C}_{pmk} = 1.50$

where $b = 3c\sqrt{1+\xi^2} + |\xi|$. The test power (where $b = 3C_{pmk}\sqrt{1+\xi^2} + |\xi|$) is

$$\pi(C_{pmk}) = P(\hat{C}_{pmk} \geq c_0 \mid C_{pmk} > c)$$
$$= \int_0^{b\sqrt{n}/(1+3c_0)} F_K\left(\frac{(b\sqrt{n}-t)^2}{9c_0^2} - t^2\right)[\phi(t+\xi\sqrt{n}) + \phi(t-\xi\sqrt{n})]\,\mathrm{d}t \tag{18}$$

To eliminate the need for estimating the characteristic parameter ξ, we apply the technique used by Pearn and Lin[12] to examine the behavior of the critical values, c_0, against the parameter ξ. We perform extensive calculations to obtain the critical values for $\xi = 0(0.01)3$, $c = 1.00$, 1.33, 1.50, 1.67, 2.00, 2.5, and 3.0, $n = 10(50)300$, and $\alpha = 0.05$. The results show that the critical value obtains its maximum, uniformly, at $\xi = 0.5$ in all cases. For practical purposes, we could greatly simplify the solution procedure by solving Equation (17) with $\xi = 0.5$ to obtain the required critical values, without having to further estimate ξ. In practice, sample data are collected, contaminated with measurement errors, to estimate the empirical process capability. Thus, the α-risk corresponding to the test using the sample estimate $\hat{C}_{pmk}^G$ (denoted by α_G) becomes

$$P(\hat{C}_{pmk}^G \geq c_0 \mid C_{pmk} \leq c) = \alpha_G$$
$$\int_0^{b_G\sqrt{n}/(1+3c_0)} F_K\left(\frac{(b_G\sqrt{n}-t)^2}{9c_0^2} - t^2\right)[\phi(t+0.5\sqrt{n}) + \phi(t-0.5\sqrt{n})]\,\mathrm{d}t = \alpha_G \tag{19}$$

where $b_G = 3.75c\Big/\sqrt{1.25+\lambda^2C_p^2} + 0.5$ and $C_p = \sqrt{1.25}c + \frac{1}{6}$.

Copyright © 2006 John Wiley & Sons, Ltd.

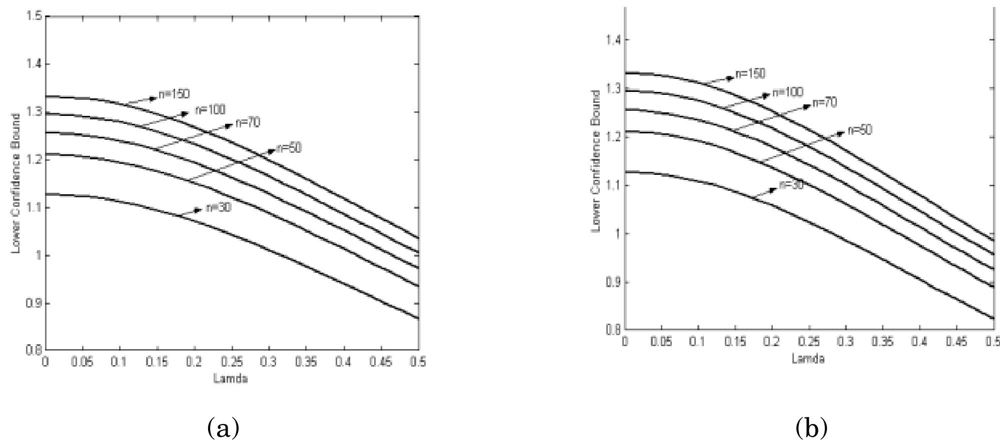

Figure 5. Plots of α_G with $n = 30, 50, 70, 100, 150$ and $\lambda \in [0, 0.5]$ for $\alpha = 0.05$: (a) $c = 1.00$; (b) $c = 1.50$

The test power (denoted by π_G) is

$$\pi_G(C_{pmk}) = P(\hat{C}^G_{pmk} \geq c_0 \mid C_{pmk} > c)$$
$$\pi_G(C_{pmk}) = \int_0^{b_G\sqrt{n}/(1+3c_0)} F_K\left(\frac{(b_G\sqrt{n}-t)^2}{9c_0^2} - t^2\right)[\phi(t+0.5\sqrt{n}) + \phi(t-0.5\sqrt{n})]\,\mathrm{d}t \tag{20}$$

where $b_G = 3.75C_{pmk}\Big/\sqrt{1.25+\lambda^2C_p^2} + 0.5$ and $C_p = \sqrt{1.25}C_{pmk} + \frac{1}{6}$.

Earlier discussions indicate that the true process capability would be severely underestimated if $\hat{C}^G_{pmk}$ is used. The probability that $\hat{C}^G_{pmk}$ is greater than c_0 would be less than that of using $\hat{C}_{pmk}$. Thus, the α-risk using $\hat{C}^G_{pmk}$ is less than the α-risk using $\hat{C}_{pmk}$ (α) for C_{pmk}. The test power using $\hat{C}^G_{pmk}$ is also smaller than the test power using $\hat{C}_{pmk}$. That is, $\pi_G < \pi$. Figures 5(a) and (b) are the plots of α_G with $n = 30, 50, 70, 100, 150$, $\lambda \in [0, 0.5]$, and $\alpha = 0.05$ for $c = 1.00$ and 1.50, respectively. Figures 6(a) and (b) plot π_G versus λ with $n = 100$, $\alpha = 0.05$, and $C_{pmk} = c_1(0.02)c_2$ (where $c_1 = c + 0.2$ and $c_2 = c + 1$) for $c = 1.00$ and 1.50, respectively. Note that if $\lambda = 0$, then $\alpha_G = \alpha$ and $\pi_G = \pi$. In Figures 5(a) and (b), α_G decreases as λ or n increases, and the decreasing rate is more significant with large c. In fact, for large λ, α_G is smaller than 10^{-2}. In Figures 6(a) and (b), π_G decreases as λ increases, but increases as n increases. The reduction of π_G in λ is more significant for large c. In the presence of measurement errors, the test power π_G decreases. For instance, in Figure 6(b) the π_G value ($c = 1.50$, $n = 100$) for $C_{pmk} = 2.30$ is $\pi_G = 0.9957$ if there is no measurement error ($\lambda = 0$). However, when $\lambda = 0.5$, π_G significantly decreases to 0.0834, a reduction of 0.912.

In previous sections we pointed out the problems associated with using C^G_{pmk} as an estimator for C_{pmk}. In this situation the lower confidence bound is underestimated, and the α-risk and the test power decrease with measurement errors. The probability of passing non-conforming product units decreases, but the probability of correctly judging a capable process as capable also decreases. Since the lower confidence bound is severely underestimated and the test power becomes low, the producers cannot firmly state that their processes meet the capability requirement even if their processes are sufficiently capable. Good product units would be incorrectly rejected in this case. These incorrect decisions may lead to unnecessary costs for the producers. Improving the gauge capability and providing sufficient training for the operators are both essential to reduce measurement error. Nevertheless, measurement errors are inevitable in most industry applications. In order to provide a better capability assessment, three methods for adjusted confidence bounds are proposed. A simulation study is also conducted for the performance comparison of three methods.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

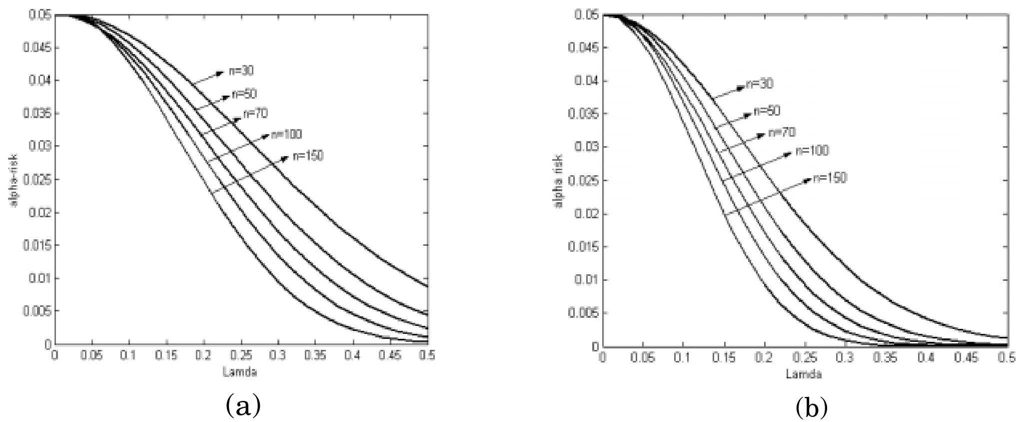

Figure 6. Plots of π_G versus λ with $n = 100$, $\alpha = 0.05$: (a) $c = 1.00$, $C_{pmk} = 1.2(0.20)2.00$; (b) $c = 1.50$, $C_{pmk} = 1.70(0.20)2.50$

8. METHODS FOR ADJUSTED CONFIDENCE BOUNDS

8.1. Sampling distribution approaches

From Equation (10), suppose that the desired confidence coefficient is γ and the adjusted confidence interval of $\hat{C}^G_{pmk}$ with the adjusted lower confidence bound L_{A} can be established as

$$P(\hat{C}^G_{pmk} > L_{\mathrm{A}}) = \gamma$$

then

$$1 - \int_0^{b_G\sqrt{n}/(1+3\hat{C}^G_{pmk})} F_K\left(\frac{(b_G\sqrt{n} - t)^2}{9(\hat{C}^G_{pmk})^2}\right)[\phi(t + \xi_G\sqrt{n}) + \phi(t - \xi_G\sqrt{n})]\,\mathrm{d}x = \gamma \tag{21}$$

where $b_G = 3L_{\mathrm{A}}(1 + \xi^2)\Big/\sqrt{1 + \xi^2 + \lambda^2 C_p^2} + |\xi_G|$

Equation (21) involves the unknown parameter ξ and C_p. As recommended by Bolyes[25] we directly replaced these variables by their MLEs. Thus, an adjusted lower confidence bound based on sampling distribution of $\hat{C}^G_{pmk}$ (denoted as $L^{\mathrm{SD}}_{\mathrm{A}}$) can be obtained by solving Equation (21) with $^*b_G = 3L^{\mathrm{SD}}_{\mathrm{A}}(1 + \hat{\xi}^2_G)(1 + \hat{\xi}^2)/\sqrt{1 + \hat{\xi}^2 + \lambda^2 \hat{C}_p^2} + |\hat{\xi}_G|$ for given sample data, *LSL*, *USL*, γ, and λ:

$$\int_0^{^*b_G\sqrt{n}/(1+3\hat{C}^G_{pmk})} F_K\left(\frac{(^*b_G\sqrt{n} - t)^2}{9(\hat{C}^G_{pmk})^2}\right)[\phi(t + \hat{\xi}_G\sqrt{n}) + \phi(t - \hat{\xi}_G\sqrt{n})]\,\mathrm{d}x = 1 - \gamma \tag{22}$$

In finding the lower confidence bounds, Pearn and Shu[13] recommended placing $\hat{\xi} = 0.5$ for reasons of quality assurance purposes to eliminate additional sampling errors from estimating ξ. Therefore, we may solve Equation (23) to obtain the adjusted lower confidence bounds (denoted as $L^{\mathrm{MSD}}_{\mathrm{A}}$) with $^{**}b_G = 3.75L^{\mathrm{MSD}}_{\mathrm{A}}\Big/\sqrt{1.25 + \lambda^2\hat{C}_p^2} + 0.5$ for given sample data, *LSL*, *USL*, γ, and λ:

$$\int_0^{^{**}b_G\sqrt{n}/(1+3\hat{C}^G_{pmk})} F_K\left(\frac{(^{**}b_G\sqrt{n} - t)^2}{9(\hat{C}^G_{pmk})^2}\right)[\phi(t + 0.5\sqrt{n}) + \phi(t - 0.5\sqrt{n})]\,\mathrm{d}x = 1 - \gamma \tag{23}$$

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

8.2. GCI approach

Tsui and Weerahandi[26] introduced the concept of generalized inference for testing hypotheses and constructing GCIs in situations such as the SD approach (see Equations (22) and (23)) where the exact confidence intervals are not exact. GCIs have been used in recent articles for a variety of problems including the construction of tolerance intervals by Liao *et al.*[27] and the development of tests for variance components by Mathew and Webb[28]. Hamada and Weerahandi[29] and Adamec and Burdick[30] used GCIs to handle measurement error problems. Burdick *et al.*[31,32] presented GCIs for misclassification rates in a gauge R&R study. Daniels *et al.*[33] proposed using GCIs for comparing capability measures when there is no measurement error. To compute a GCI for monitoring process capability based on C_{pmk} with gauge measurement errors, one must define generalized pivotal quantities (GPQs) for μ and σ^2. Using the previous notation of $G \sim N(\mu, \sigma_G^2 = \sigma^2 + \sigma_M^2)$, using the method of Iyer and Patterson[34] as described in Appendix B2 of Burdick *et al.*[32], the following GPQs can be defined for μ and σ_G^2

$$\mu(GPQ) = \bar{G} - Z\sqrt{\frac{\sigma_G^2(GPQ)}{n}} \tag{24}$$

$$\sigma_G^2(GPQ) = \frac{n\tilde{S}_n^2}{W} \tag{25}$$

where Z is a standard normal variable and W is a χ^2 random variable with $n-1$ degrees of freedom. Since $\sigma_M^2 = [(USL - LSL)\lambda/6]^2$ is known, then the GPQ for σ^2 is

$$\sigma^2(GPQ) = \max(\varepsilon, \sigma_G^2(GPQ) - \sigma_M^2) \tag{26}$$

where ε is a small positive quantity to maintain non-negative variance components, for example 0.0001.

The following procedure can be used to construct a $100(1-\alpha)\%$ lower confidence bound on C_{pmk}:

(1) compute $\bar{G}$ and $\tilde{S}_n^2$ for the collected data and denote the realized value as $\bar{g}$ and $\tilde{s}_n^2$;
(2) simulate $N = 2000$ values of $\mu(GPQ)$ and $\sigma^2(GPQ)$ using Equations (24)–(26) by simulating N independent values of Z and W;
(3) for each simulated pair of $\mu(GPQ)$ and $\sigma^2(GPQ)$, compute $C_{pmk}(GPQ)$ where

$$C_{pmk}(GPQ) = \min\left\{\frac{USL - \mu(GPQ)}{3\sqrt{\sigma^2(GPQ) + (\mu(GPQ) - T)^2}}, \frac{\mu(GPQ) - LSL}{3\sqrt{\sigma^2(GPQ) + (\mu(GPQ) - T)^2}}\right\} \tag{27}$$

(4) order the $N = 2000$ values of $C_{pmk}(GPQ)$ from the least to greatest;
(5) the $100\gamma\%$ lower bounds on $\hat{C}_{pmk}^G$, denoted as $L_{\mathrm{A}}^{\mathrm{GCI}}$, is the value in position $N \times \alpha$ of the ordered step in (4).

9. ADJUSTED LOWER CONFIDENCE BOUND COMPARISONS: A SIMULATION STUDY

The important consideration in choosing methods for determining the adjusted lower confidence bound is the performance of each method. In order to ascertain the performance of the adjusted lower confidence bound methods (Equation (22) for $L_{\mathrm{A}}^{\mathrm{SD}}$, Equation (23) for $L_{\mathrm{A}}^{\mathrm{MSD}}$, and Equations (24)–(27) for $L_{\mathrm{A}}^{\mathrm{GCI}}$) a simulation study is conducted (the Matlab program for performing this calculation is available upon request). Random samples of $n = 20(10)150$ different sample sizes are drawn 2000 times from processes with different gauge capabilities $\lambda = 0, 0.1, 0.2, 0.25$ and alternative values of μ and σ^2 for which $\xi = 0(0.05)1$, so as to detect any dependence on the coverage rate for the three methods. Tables II–VI present the mean value (ME) of $\hat{C}_{pmk}^G$, the percentage coverage rate (CR) observed in the simulation for a nominal confidence level of 95%, and mean

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

Table II. The simulated results for 95% lower confidence bounds of SD, MSD, and GCI methods with $C_{pmk} = 1$ ($LSL = -3$, $USL = 3$, $\mu = 0$, and $\sigma = 1$) at $N = 2000$

	$\lambda = 0.0$							$\lambda = 0.1$						
	$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}		$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	0.9799	**0.9650**	**0.7099**	0.9795	0.6566	0.9900	0.6193	0.9748	**0.9635**	**0.7100**	0.9790	0.6560	0.9900	0.6175
50	0.9782	**0.9690**	**0.8117**	0.9895	0.7718	0.9935	0.7599	0.9732	**0.9675**	**0.8117**	0.9890	0.7709	0.9935	0.7585
70	0.9812	**0.9595**	**0.8417**	0.9835	0.8063	0.9890	0.7992	0.9762	**0.9590**	**0.8416**	0.9835	0.80530	0.9885	0.7979
100	0.9761	**0.9755**	**0.8605**	0.9910	0.8302	0.9930	0.8268	0.9767	**0.9590**	**0.8655**	0.9845	0.8342	0.9870	0.8307
150	0.9821	**0.9635**	**0.8881**	0.9875	0.8622	0.9875	0.8635	0.9781	**0.9665**	**0.8888**	0.9890	0.8623	0.9900	0.8618
	$\lambda = 0.2$							$\lambda = 0.25$						
	$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}		$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	0.9598	**0.9575**	**0.7105**	0.9780	0.6546	0.9900	0.6121	0.9578	**0.9615**	**0.7091**	0.9840	0.6528	0.9945	0.6106
50	0.9585	**0.9625**	**0.8116**	0.9885	0.7683	0.9930	0.7542	0.9576	**0.9635**	**0.8117**	0.9890	0.7675	0.9950	0.7530
70	0.9616	**0.9560**	**0.8153**	0.9830	0.8026	0.9888	0.7942	0.9572	**0.9635**	**0.8370**	0.9845	0.7986	0.9905	0.7904
100	0.9566	**0.9690**	**0.8600**	0.9910	0.8262	0.9930	0.8223	0.9625	**0.9600**	**0.8655**	0.9850	0.8317	0.9885	0.8280
150	0.9644	**0.9610**	**0.8893**	0.9890	0.8601	0.9905	0.8599	0.9645	**0.9660**	**0.8897**	0.9885	0.8601	0.9890	0.8598

Table III. The simulated results for 95% lower confidence bounds of SD, MSD, and GCI methods with $C_{pmk} = 1.6007$ ($LSL = -5.2$, $USL = 5.2$, $\mu = 0.25$, $\sigma = 1.0$) at $N = 2000$

	$\lambda = 0.0$							$\lambda = 0.1$						
	$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}		$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.6632	**0.9400**	**1.2014**	**0.9530**	**1.1667**	**0.9745**	**1.1101**	1.6275	**0.9380**	**1. 1927**	**0.9575**	**1.1571**	**0.9780**	**1.0949**
50	1.6288	**0.9380**	**1.3359**	**0.9545**	**1.3188**	**0.9690**	**1.2991**	1.5962	**0.9380**	**1.3272**	**0.9575**	**1.3083**	**0.9730**	**1.2852**
70	1.6194	**0.9435**	**1.3705**	**0.9550**	**1.3587**	**0.9655**	**1.3457**	1.5944	**0.9375**	**1.3693**	**0.9515**	**1.3541**	**0.9665**	**1.3386**
100	1.6043	**0.9555**	**1.3961**	**0.9665**	**1.3880**	**0.9705**	**1.3799**	1.5927	**0.9370**	**1.4061**	**0.9515**	**1.3949**	**0.9590**	**1.3852**
150	1.6108	**0.9400**	**1.4390**	**0.9480**	**1.4335**	**0.9515**	**1.4291**	1.5876	**0.9405**	**1.4383**	**0.9575**	**1.4299**	**0.9625**	**1.4248**
	$\lambda = 0.2$							$\lambda = 0.25$						
	$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}		$\hat{C}^{G}_{pmk}$	L^{SD}_{A}		L^{MSD}_{A}		L^{GCI}_{A}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.5785	0.9170	1.2155	**0.9360**	**1.1720**	**0.9725**	**1.0833**	1.5238	0.9045	1.2136	**0.9315**	**1.1659**	**0.9750**	**1.0593**
50	1.5346	0.9270	1.3306	**0.9465**	**1.3035**	**0.9665**	**1.2683**	1.4956	0.9020	1.3370	**0.9415**	**1.3037**	**0.9675**	**1.2589**
70	1.5273	0.9265	1.3654	**0.9525**	**1.3430**	**0.9710**	**1.3188**	1.4931	0.9095	1.3760	**0.9425**	**1.3465**	**0.9665**	**1.3154**
100	1.5277	0.9280	1.4042	**0.9490**	**1.3851**	**0.9580**	**1.3694**	1.4766	0.9190	1.3951	**0.9515**	**1.3705**	**0.9710**	**1.3502**
150	1.5255	0.9250	1.4386	**0.9535**	**1.4218**	**0.9600**	**1.4136**	1.4860	0.9060	1.4418	**0.9520**	**1.4191**	**0.9630**	**1.4080**

value of the lower confidence bounds (MLCB) for true process capability C_{pmk}. The CR and MLCB entries are used as a basis for evaluating the performance of various methods. Bold numbers denote acceptable performance measures.

Obviously, the ratio of CR and MLCB is negative. The lower the value of CR, the closer it is to the actual value of the MLCB. This is true, because the lower CR means that the much lower confidence bounds do not cover the actual value, and the MLCB is much closer to the actual value of C_{pmk}.

An advantage of obtaining the adjusted lower confidence intervals of the SD approach of Equation (22) and MSE approach of Equation (23), L^{SD}_{A} and L^{MSD}_{A}, respectively, are that both can be written in closed form and, unlike the GCI method, do not require Monte Carlo simulation.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

Table IV. The simulated results for 95% lower confidence bounds of SD, MSD, and GCI methods with $C_{pmk} = 1.3416$ ($LSL = -5$, $USL = 5$, $\mu = 0.5$, $\sigma = 1$) at $N = 2000$

	$\lambda = 0.0$							$\lambda = 0.1$						
	$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}		$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.4043	**0.9435**	**0.9887**	**0.9560**	**0.9737**	**0.9660**	**0.9497**	1.3817	**0.9350**	**0.9856**	**0.9465**	**0.9701**	**0.9665**	**0.9389**
50	1.3595	**0.9480**	**1.0959**	**0.9525**	**1.0926**	**0.9595**	**1.0836**	1.3487	**0.9430**	**1.0999**	**0.9470**	**1.0965**	**0.9600**	**1.0832**
70	1.3502	**0.9585**	**1.1276**	**0.9620**	**1.1259**	**0.9645**	**1.1195**	1.3455	**0.9385**	**1.1368**	**0.9410**	**1.1351**	**0.9530**	**1.1248**
100	1.3507	**0.9515**	**1.1640**	**0.9525**	**1.1630**	**0.9570**	**1.1583**	1.3395	**0.9485**	**1.1672**	**0.9500**	**1.1663**	**0.9550**	**1.1582**
150	1.3495	**0.9505**	**1.1968**	**0.9525**	**1.1963**	**0.9555**	**1.1931**	1.3336	**0.9510**	**1.1957**	**0.9525**	**1.1953**	**0.9610**	**1.1894**

	$\lambda = 0.2$							$\lambda = 0.25$						
	$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}		$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.3325	0.9195	0.9864	0.9305	0.9714	**0.9655**	**0.9190**	1.3028	0.9085	0.9917	0.9195	0.9772	**0.9660**	**0.9077**
50	1.3051	0.9275	1.0999	0.9345	1.0967	**0.9450**	**1.0706**	1.2761	0.9205	1.1012	0.9255	1.0983	**0.9550**	**1.0622**
70	1.3004	0.9335	1.1345	0.9370	1.1332	**0.9580**	**1.1117**	1.2716	0.9245	1.1353	0.9290	1.1344	**0.9575**	**1.1044**
100	1.2976	0.9415	1.1673	0.9445	1.1666	**0.9610**	**1.1492**	1.2688	0.9285	1.1676	0.9350	1.1671	**0.9610**	**1.1428**
150	1.2932	0.9375	1.2010	0.9398	1.2001	**0.9585**	**1.1818**	1.2622	0.9245	1.2014	0.9342	1.2011	**0.9585**	**1.1763**

Table V. The simulated results for 95% lower confidence bounds of SD, MSD, and GCI methods with $C_{pmk} = 1.800$ ($LSL = -7.5$, $USL = 7.5$, $\mu = 0.75$, $\sigma = 1$) at $N = 2000$

	$\lambda = 0.0$							$\lambda = 0.1$						
	$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}		$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.8636	**0.9475**	**1.3445**	**0.9520**	**1.3158**	**0.9615**	**1.3196**	1.8281	0.9420	1.3497	0.9440	1.3289	**0.9630**	**1.3056**
50	1.8263	**0.9455**	**1.4996**	**0.9490**	**1.4846**	**0.9520**	**1.4921**	1.7912	0.9370	1.5014	0.9375	1.4946	**0.9520**	**1.4120**
70	1.8235	**0.9535**	**1.5460**	**0.9555**	**1.5352**	**0.9570**	**1.5412**	1.7884	0.9420	1.5473	0.9415	1.5443	**0.9565**	**1.5322**
100	1.8160	**0.9500**	**1.5842**	**0.9545**	**1.5756**	**0.9520**	**1.5810**	1.7833	0.9480	1.5871	0.9460	1.5867	**0.9580**	**1.5757**
150	1.8076	**0.9595**	**1.6187**	**0.9630**	**1.6122**	**0.9620**	**1.6168**	1.7782	0.9395	1.6244	0.9380	1.6260	**0.9470**	**1.6157**

	$\lambda = 0.2$							$\lambda = 0.25$						
	$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}		$\hat{C}^G_{pmk}$	L_A^{SD}		L_A^{MSD}		L_A^{GCI}	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.7295	0.9035	1.3674	0.8955	1.3714	**0.9595**	**1.2652**	1.6669	0.8820	1.3834	0.8625	1.4107	**0.9615**	**1.2384**
50	1.6905	0.9170	1.5044	0.9020	1.5198	**0.9525**	**1.4507**	1.6332	0.9000	1.5096	0.8715	1.5403	**0.9530**	**1.4305**
70	1.6885	0.9200	1.5461	0.9040	1.5649	**0.9575**	**1.5019**	1.6269	0.9035	1.5497	0.8685	1.5832	**0.9580**	**1.4840**
100	1.6857	0.9230	1.5860	0.8960	1.6069	**0.9610**	**1.5499**	1.6241	0.8985	1.5881	0.8580	1.6226	**0.9600**	**1.5341**
150	1.6791	0.9300	1.6204	0.8975	1.6432	**0.9560**	**1.5914**	1.6176	0.9130	1.6213	0.8640	1.6578	**0.9575**	**1.5778**

It can be noted that the lower bound of the SD method, L_A^{SD}, is a better performance measure in the absence of gauge measurement errors, i.e. $\lambda = 0.0$. When measurement errors are unavoidable, for small ξ (say $0 \le \xi < 0.15$), the SD method performs well providing the most accurate CRs of the methods studied here. The MSD and GCI methods, L_A^{MSD} and L_A^{GCI}, keep type I error (α-risk) from exceeding a predetermined value (such as 0.05 or 0.01) to provide necessary protection to the customers. However, the conservative lower confidence bounds for the true value of the same index can lead to a higher level of type II error. For $0.15 \le \xi < 0.40$, the MSD and GCI methods have the same acceptable performance measures with accurate CRs for all of the cases studied. However, the SD method ensures that type I error (α-risk) is greater than a predetermined value (such as 0.05 or 0.01) providing an overoptimistic process capability. For large values of ξ,

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

Table VI. The simulated results for 95% lower confidence bounds of SD, MSD, and GCI methods with $C_{pmk} = 2.1213$ ($LSL = -10$, $USL = 10$, $\mu = 1$, $\sigma = 1$) at $N = 2000$

	$\lambda = 0.0$							$\lambda = 0.1$						
	$\hat{C}^G_{pmk}$	L^{SD}_A		L^{MSD}_A		L^{GCI}_A		$\hat{C}^G_{pmk}$	L^{SD}_A		L^{MSD}_A		L^{GCI}_A	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	2.1852	**0.9475**	**1.6284**	**0.9595**	**1.5548**	**0.9565**	**1.6094**	2.1292	0.9355	1.6382	0.9405	1.5951	**0.9575**	**1.5837**
50	2.1466	**0.9445**	**1.8011**	**0.9615**	**1.7532**	**0.9500**	**1.7923**	2.0904	0.9280	1.8046	0.9325	1.7878	**0.9525**	**1.7739**
70	2.1451	**0.9500**	**1.8529**	**0.9660**	**1.8130**	**0.9555**	**1.8451**	2.0887	0.9360	1.8552	0.9370	1.8460	**0.9550**	**1.8290**
100	2.1341	**0.9490**	**1.8912**	**0.9640**	**1.8576**	**0.9525**	**1.8846**	2.0776	0.9355	1.8925	0.9310	1.8895	**0.9525**	**1.8704**
150	2.1315	**0.9475**	**1.9341**	**0.9640**	**1.9061**	**0.9470**	**1.9318**	2.0750	0.9295	1.9350	0.9215	1.9380	**0.9495**	**1.9166**
	$\lambda = 0.2$							$\lambda = 0.25$						
	$\hat{C}^G_{pmk}$	L^{SD}_A		L^{MSD}_A		L^{GCI}_A		$\hat{C}^G_{pmk}$	L^{SD}_A		L^{MSD}_A		L^{GCI}_A	
n	ME	CR	MLCB	CR	MLCB	CR	MLCB	ME	CR	MLCB	CR	MLCB	CR	MLCB
20	1.9804	0.8780	1.6658	0.8430	1.7228	**0.9560**	**1.5139**	1.8799	0.8580	1.7208	0.7640	1.8373	**0.9580**	**1.5805**
50	1.9441	0.8950	1.8132	0.8290	1.8775	**0.9495**	**1.7216**	1.8525	0.8675	1.8234	0.7575	1.9418	**0.9515**	**1.6878**
70	1.9366	0.8905	1.8568	0.8150	1.9277	**0.9575**	**1.7789**	1.8453	0.8640	1.8639	0.7215	1.9870	**0.9580**	**1.7488**
100	1.9341	0.8885	1.8992	0.7940	1.9736	**0.9610**	**1.8337**	1.8427	0.8590	1.9037	0.6875	2.0260	**0.9615**	**1.8071**
150	1.9270	0.8995	1.9354	0.7630	2.0147	**0.9560**	**1.8885**	1.8351	0.8740	1.9378	0.6460	2.0635	**0.9505**	**1.8584**

Table VII. Summary of the most effective methods for adjusted lower confidence bounds

	$\lambda = 0.0$	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.25$
$0.0 \le \xi < 0.15$	SD	SD	SD	SD
$0.15 \le \xi < 0.4$	SD, MSD, GCI	SD, MSD, GCI	MSD, GCI	MSD, GCI
$\xi \ge 0.4$	SD, MSD, GCI	GCI	GCI	GCI

$0.40 \le \xi$, the GCI method performs very well and the CRs achieved by the lower confidence bounds are quite robust and close to the nominal values for all of the methods studied. On the other hand, the CRs of the SD and MSD methods are significantly lower than the state level for all of the methods studied. Table VII shows the most effective methods for adjusted lower confidence bounds with $\lambda = 0.0, 0.1, 0.20, 0.25$ and $0.0 \le \xi < 0.15$, $0.15 \le \xi < 0.4$, and $\xi \ge 0.4$.

As a result of this discussion we present the following generalizations for practitioners of real-world factory applications:

(a) if no measurement errors exist, the adjusted lower confidence bound is in support of the use of the SD approach, L^{SD}_A;
(b) when $0 \le \xi < 0.15$, the adjusted lower confidence bound is in support of the use of the SD approach, L^{SD}_A;
(c) when $0.15 \le \xi < 0.40$, the adjusted lower confidence bound is in support of the use of the MSD and GCI approaches, L^{MSD}_A and L^{GCI}_A;
(d) when $0.4 \le \xi$, the adjusted lower confidence bound is in support of the use of the GCI approach L^{GCI}_A.

10. APPLICATION EXAMPLE: PRECISION VOLTAGE REFERENCE

When a data conversion system is designed, the system accuracy greatly depends on the accuracy of the voltage established by the internal or external DC voltage reference. The voltage reference is used to produce a precise

Copyright © 2006 John Wiley & Sons, Ltd.

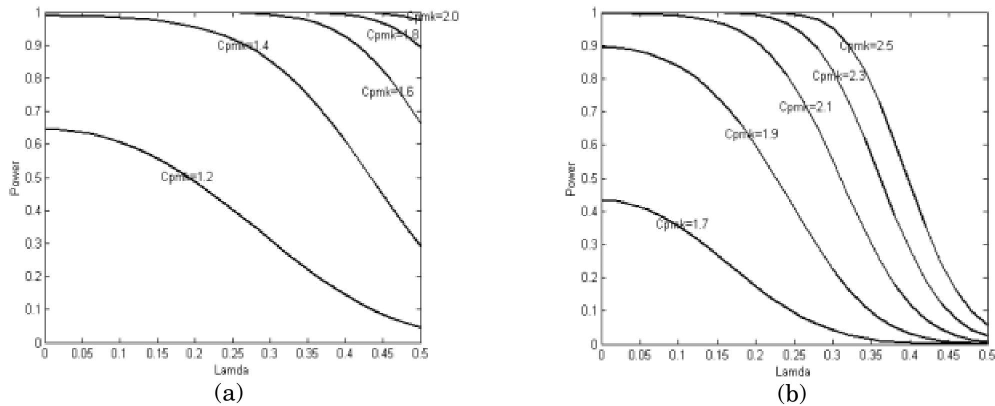

Figure 7. The eight-pin SOIC-30 and eight-pin TO-99 packages for PVR

value of the output voltage for setting the full-scale input of the data conversion system. In an analog-to-digital converter (ADC), the DC voltage reference together with the analog input signal are used to generate the digitized output signal. In a digital-to-analog converter (DAC), the DAC selects and produces an analog output from the DC reference voltage according to the digital input signal presented at the input of the DAC. Any errors in the reference voltage over the operating temperature range will adversely affect the linearity and spurious free dynamic range (SFDR) of the ADC/DAC. With the emergence of the portable battery-operated environment, low voltage and low power are key goals of the industry. The voltage reference can also be used in constructing a precision regulated supply that could have better characteristics than some regulator chips, which can occasionally dissipate too much power. In addition, voltage references are needed in the design of products which must be accurate, such as voltmeters, ohmmeters, and ammeters.

Consider the following case taken from a manufacturing factory in Taiwan making one type of low-power, fast-warm-up, and highly stable 15 V precision voltage reference (PVR). The output voltage is extremely insensitive to both line and load variations and can be externally adjusted with minimal effect on drift and stability. This PVR is offered in eight-pin SOIC-30 and eight-pin TO-99 packages, as depicted in Figure 7. They are ideal for communications equipment, data acquisition systems, instrumentation and process control, high-precision power supplies, and battery powered equipment. They may also be used in portable battery powered equipment (such as notebook computers, PDAs, DVMs, GPS, etc.). Initial accuracy is one critical quality characteristic of this PVR which has a significant impact on the PVR quality/reliability. This characteristic is usually only valid at room temperature, and it provides a starting point for most of the other specifications. The output voltage tolerance of a reference measured without a load applied after the device is turned on and warmed up. Manufacturers specify a reference with a small initial error so they do not have to perform room-temperature system calibration after assembly.

For this particular model of PVR product, the specification limits are $T = 15$ V, $USL = 15.025$, and $LSL = 14.975$. A total of 70 observations are collected and displayed in Table VIII. A histogram and a normal probability plot show that the collected data follow the normal distribution. A Shapiro–Wilk test is applied to further justify the assumption. To determine whether the process is 'excellent' ($C_{pmk} > 1.33$) with unavoidable measurement errors $\lambda = 0.24$, we first determine that $c = 1.33$ and $\alpha = 0.05$. Then, based on the sample data of 70 observations, we obtain the sample mean $\bar{g} = 15.0014$, the sample standard deviation $\tilde{s}_n = 0.0049$ ($\bar{g}$ and $\tilde{s}_n$ are the realized sample values for $\bar{G}$ and $\tilde{S}_n$), and the point estimator $\hat{C}^G_{pmk} = 1.5526$. Since $0.15 \leq \hat{\xi} < 0.4$, the GCI method is suggested for this PVR process capability assessment. The Matlab computer program (available upon request) reads $T = 15$ V, $USL = 15.025$, $LSL = 14.975$, 70 observations, and $\lambda = 0.24$ (provided by the gauge manufacturing factory), and the desired confidence coefficient $\gamma = 0.95$, so the 95% lower confidence bound of the true process capability can be obtained as $L^{\mathrm{GCI}}_{\mathrm{A}} = 1.3812$. We thus can assure that the production yield is 99.99658%, and the number of the non-conformities is less than 34.21 ppm (parts per million).

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

Table VIII. 70 observations for output voltage (in volts)

15.0023	15.0056	15.0062	15.0030	15.0046	15.0098	14.9952	14.9968	15.0026	14.9959
14.9977	15.0021	15.0053	15.0069	15.0066	15.0001	14.9987	15.0014	15.0013	15.0048
14.9994	15.0003	15.0046	15.0080	14.9989	15.0059	14.9942	14.9895	14.9884	15.0030
14.9946	15.0022	15.0069	14.9990	14.9942	14.9978	15.0046	14.9983	15.0019	15.0000
14.9999	15.0057	15.0082	15.0065	15.0015	14.9954	15.0004	14.9947	14.9973	15.0063
14.9994	15.0132	15.0020	15.0009	15.0046	15.0052	14.9952	15.0035	15.0004	15.0098
15.0107	15.0035	15.0002	15.0001	15.0022	14.9947	14.9973	15.0049	15.0001	14.9961

Moreover, similar to the adjusted lower confidence bound, we obtain the adjusted critical value 1.498 for the MSD method based on $\alpha = 0.05$, $\lambda = 0.24$, and $n = 70$. Since $\hat{C}^G_{pmk} > 1.498$, we therefore conclude that the process is 'excellent'. However, we also see that if we ignore the measurement errors and evaluate the critical value without any correction, the critical value may be calculated as $c_0 = 1.585$. In this case we would say that the process is not 'excellent' since $\hat{C}^G_{pmk}$ is no greater than the uncorrected critical value 1.585.

11. CONCLUSIONS

Measurement errors are unavoidable in most industry applications. In this paper, we conducted a sensitivity study for process capability, C_{pmk}, in the presence of gauge measurement errors. The statistical properties and capability testing of estimating C_{pmk} were investigated to obtain lower confidence bounds and critical values for true process capability testing. In estimating and testing the capability, the estimator $\hat{C}^G_{pmk}$, using the sample data contaminated with the measurement error, severely underestimates the true capability and decreases testing power. The lower confidence bounds must be adjusted to improve the accuracy of capability assessment. The SD approaches and the GCI approach were presented to obtain the adjusted lower confidence bounds. An intensive simulation study was used to compare performance of the attained confidence levels and the average interval lengths of both approaches. The result recommends the appropriate method to practitioners for the real-world factory applications.

Acknowledgements

Work on this paper was partially funded by National Science Council of Taiwan, under grant NSC92-2213-E-251-005. The authors wish to thank the chief editor and the anonymous reviewers for encouraging and helpful comments that resulted in an improved presentation of our research.

REFERENCES

1. Kane VE. Process capability indices. *Journal of Quality Technology* 1986; **18**:41–52.
2. Chan LK, Cheng SW, Spiring FA. A new measure of process capability: C_{pm}. *Journal of Quality Technology* 1988; **20**:162–175.
3. Zimmer LS, Hubele NF, Zimmer WJ. Confidence intervals and sample size determination for C_{pm}. *Quality and Reliability Engineering International* 2001; **17**:51–68.
4. Vännman K, Hubele NF. Distributional properties of estimated capability indices based on subsamples. *Quality and Reliability Engineering International* 2003; **19**:111–128.
5. Perakis M, Xekalaki E. A new method for constructing confidence intervals for the index C_{pm}. *Quality and Reliability Engineering International* 2004; **20**:651–665.
6. Pearn WL, Kotz S, Johnson NL. Distributional and inferential properties of process capability indices. *Journal of Quality Technology* 1992; **24**:216–231.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre

7. Hoffman LL. Obtaining confidence intervals for C_{pk} using percentiles of the distribution of $\hat{C}_{pk}$. *Quality and Reliability Engineering International* 2001; **17**:113–118.
8. Chen SM, Hsu NF. The asymptotic distribution of the process capability index: C_{pmk}. *Communications in Statistics—Theory and Methods* 1995; **24**:1279–1291.
9. Wright PA. The probability density function of process capability index C_{pmk}. *Communications in Statistics—Theory and Methods* 1998; **27**:1781–1789.
10. Jessenberger J, Weihs C. A note on the behavior of C_{pmk} with asymmetric specification limits. *Journal of Quality Technology* 2000; **32**:440–443.
11. Pearn WL, Yang SL, Chen KS, Lin PC. Testing process capability using the index C_{pmk} with an application. *International Journal of Reliability, Quality and Safety Engineering* 2001; **8**:15–34.
12. Pearn WL, Lin PC. Computer program for calculating the p-value in testing process capability index C_{pmk}. *Quality and Reliability Engineering International* 2002; **18**:333–342.
13. Pearn WL, Shu MH. Measuring manufacturing capability based on lower confidence bounds of C_{pmk} applied to current transmitter process. *International Journal of Advanced Manufacturing Technology* 2004; **23**:116–125.
14. Bordignon S, Scagliarini M. Statistical analysis of process capability indices with measurement errors. *Quality and Reliability Engineering International* 2002; **18**:321–332.
15. Bordignon S, Scagliarini M. Estimation of C_{pm} when measurement error is present. *Quality and Reliability Engineering International* 2006; **22**(7):787–801. DOI: 10.1002/qre.728.
16. Montgomery DC, Runger GC. Gauge capability and designed experiments. Part I: Basic methods. *Quality Engineering* 1993; **6**:115–135.
17. Montgomery DC, Runger GC. Gauge capability analysis and designed experiments. Part II: Experimental design models and variance component estimation. *Quality Engineering* 1993; **6**:289–305.
18. Burdick RK, Borror CM, Montgomery DC. A review of methods for measurement systems capability analysis. *Journal of Quality Technology* 2003; **35**:342–354.
19. Mittag HJ. Measurement error effects on the performance of process capability indices. *Proceedings of the 5th International Workshop on Intelligent Statistical Quality Control*, Osaka, Japan, 1994.
20. Mittag HJ. Measurement error effects on the performance of process capability indices. *Frontiers in Statistical Quality Control* 1997; **5**:195–206.
21. Levinson WA. How good is your gage? *Semiconductor International* 1995; 165–168.
22. Montgomery DC. *Introduction to Statistical Quality Control* (4th edn). Wiley: New York, 2001.
23. Vannman K. A unified approach to capability indices. *Statistica Sinica* 1995; **5**:805–820.
24. Shu MH, Chen KS. Estimating process capability indices based on subsamples for asymmetric tolerances. *Communications in Statistics—Theory and Methods* 2005; **34**:485–506.
25. Bolyes RA. The Taguchi capability index. *Journal of Quality Technology* 1991; **23**:17–26.
26. Tsui K, Weerahandi S. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association* 1989; **84**:602–607.
27. Liao CT, Lin TY, Iyer HK. One and two-sided tolerance intervals for general balanced mixed models and unbalanced one-way random model. *Technometrics* 2005; **47**:323–335.
28. Mathew T, Webb DW. Generalized p values and confidence intervals for variance components: Applications to army testing and evaluation. *Technometrics* 2005; **47**:312–322.
29. Hamada M, Weerahandi S. Measurement system assessment via generalized inference. *Journal of Quality Technology* 2000; **32**:241–253.
30. Adamec E, Burdick RK. Confidence intervals for a discrimination ratio in a gauge R&R study with three random factors. *Quality Engineering* 2003; **15**:383–389.
31. Burdick RK, Borror CM, Montgomery DC. Design and analysis of gauge R&R studies: Making decisions with confidence intervals in random and mixed ANOVA models. *ASA–SIAM Series on Statistics and Applied Probability*, SIAM: Philadelphia, PA/ASA: Alexandria, VA, 2005. Available at: http://www.siam.org/books/sa17.
32. Burdick RK, Park YJ, Montgomery DC, Borror CM. Confidence intervals for misclassification rates in a gauge R&R study. *Journal of Quality Technology* 2005; **37**:294–303.
33. Daniels L, Edgar B, Burdick RK, Hubele N. Using confidence intervals to compare process capability indices. *Quality Engineering* 2005; **17**:23–32.
34. Iyer HK, Patterson P. A recipe for constructing generalized pivotal quantities and generalized confidence intervals. *Technical Report 2002-10*, Department of Statistics, Colorado State University, Fort Collins, CO, 2002. Available at: http://www.stat.colostate.edu/research/2002_10.pdf.

Copyright © 2006 John Wiley & Sons, Ltd.

Authors' biographies

M. H. Shu, received his PhD from the University of Texas, Arlington, U.S.A. and is an associate professor of Industrial Engineering and Management in National Kaohsiung University of Applied Sciences, Taiwan. His research interests include statistical/engineering process control, process capability analysis, and decision-making analysis.

W. L. Pearn received his MS degree in Statistics and PhD degree in Operations Research from the University of Maryland at College Park, U.S.A. He worked for Bell Laboratories as a quality engineer before he joined the National Chiao Tung University. Currently, he is a professor in the Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan.

B. M. Hsu received her PhD degree from the University of Texas, Arlington, U.S.A. She is currently an assistant professor of Industrial Engineering and Management in Cheng Shiu University, Taiwan.

Copyright © 2006 John Wiley & Sons, Ltd.

DOI: 10.1002/qre