TRANSACTIONS OF NAMRI/SME
2009

Transactions of NAMRI/SME, Vol. 37, 2009, pp. 589-596
TP09PUB101

Multivariate Process Capability Analysis Using Non-Parametric Model and Bootstrap Sampling

authors

ANKIT PAHWA and WENZHEN HUANG
Mechanical Engineering Department
University of Massachusetts Dartmouth
North Dartmouth, MA

abstract

Current process capability analysis methods rely on the normality and univariate assumptions. This paper aims at developing a novel method general multivariate process capability analysis. Conformity of product or yield is proposed as an index for process capability evaluation. A nonparametric technique, i.e. kernel density (KD) estimation and a sampling technique are proposed for model fitting and yield estimation. The unique features of the new method include: 1) conceptually transparent to practitioners; 2) easy to be implemented; 3) no assumptions on the statistical model of the process data. Case study is presented for illustration.

terms

Manufacturing Process Capability, Quality, Kernel Density Estimation, Bootstrap

North American Manufacturing Research Institution of the
Society of Manufacturing Engineers
One SME Drive • PO Box 930
Dearborn, MI 48121 • (313) 425-3000 • www.sme.org

MULTIVARIATE PROCESS CAPABILITY ANALYSIS USING NONPARAMETRIC MODEL AND BOOTSTRAP SAMPLING

Ankit Pahwa and Wenzhen Huang
Mechanical Engineering Department
University of Massachusetts Dartmouth
North Dartmouth, MA

KEYWORDS

Manufacturing Process Capability, Quality, Kernel Density Estimation, Bootstrap

ABSTRACT

Current process capability analysis methods rely on the normality and univariate assumptions. This paper aims at developing a novel method general multivariate process capability analysis. Conformity of product or yield is proposed as an index for process capability evaluation. A nonparametric technique, i.e. kernel density (KD) estimation and a sampling technique are proposed for model fitting and yield estimation. The unique features of the new method include: 1) conceptually transparent to practitioners; 2) easy to be implemented; 3) no assumptions on the statistical model of the process data. Case study is presented for illustration.

1. INTRODUCTION

Process capability ratio (PCR) indices such as C_p, C_{pk} prevail in industry for manufacturing process evaluation and quality control. C_p, C_{pk} essentially use specification limits and estimated process mean and standard deviation to indicate potential and actual process capability of conforming to specifications. The initial introduction and interpretation of these indices were strongly tied to normality assumption and largely confined to univariate processes [Kotz et al. 2002]. More and more modified indices were introduced to accommodate non-normality, off-center, and different interpretation requirements (e.g., quality loss). The bewildering diversified versions of indices made the interpretation and comprehension per se an important issue and attracted intensions of academic and industrial community. Early comprehensive interpretations of these indices were given by Kane [1986], Sullivan et al. [1984], and Kotz et al. [1993]. More recent review and discussion include Tsui [1997], Palmer and Tsui [1999]. Kotz et al. [2002] gave an exhaustive summary and review in process capability indices definition, interpretation, and new development.

The non-normal population in manufacturing quality control is common as stated by Bisgaard et al. [1997]. To target non-normality and to align with 6σ dispersion of normal cases, Clements [1989] introduced an alternative of a percentile range, i.e., replacing 6σ with the range between upper and lower 0.135 percentage points of a non-normal quality characteristic in the index

definition. This led to similar capability indices. Along the similar line, variety of parametric models were developed for percentile range assessment, such as Pearson system [Clements 1989], Johnson system, Weibull, lognormal, generalized lambda [Pal 2005], *t*, gamma, etc. [Kotz et al. 2002]. These efforts broadened generality and applicability of initial indices. They also complicate assessment, comparison, and interpretation of the indices. For instance, with the same value of an index C_p (or C_{pk}), the answer may not be straightforward on whether the two processes have the same capability in terms of nonconformity. These parametric techniques assume a functional form for the density of quality characteristic. More involved model (distribution family) assumption, parameter estimation, and model adequacy checking are required.

The multivariate PCR is a challenging topic. Current methods rely heavily on the normality assumption [Kotz et al. 2002], and the regularity of the specification regions are also required (e.g., rectangular cube). With these assumptions the derivation of similar PCRs can be carried out. Yeh et al. [1998, 2001] proposed the ratio of expected fallout rate to the observed one as an index for both univariate and multivariate processes. However, none of these methods prevails in application.

It is well known that C_p, C_{pk} generally do not provide us with the proportion of nonconforming products. Only for normal distribution this proportion can be easily derived from C_p, C_{pk}. Several authors preferred going back to the basic concept of what process capability really means, i.e., replacing C_p, C_{pk} with more transparent and coherent indices of nonconformity or fallout rate ($\hat{p}$). Carr [1991] was probably the first who proposed the estimated $\hat{p}$ as an index. Tsui [1997] also preferred the yield or conforming rate ($1-\hat{p}$) for capability evaluation, and he proposed alternatives, i.e., quality yield indices to accommodate conformity and quality loss evaluation. Parametric models were suggested for nonconformity assessment by almost all the authors who preferred nonconformity indices. If the population is believed to follow certain distribution family of distributions, the related parameters are estimated based on the sample data, and then the nonconforming probability $\hat{p}$ is assessed with the estimated distribution model. However, none of the available distribution family can accommodate all possible behaviors in actual data densities.

Nonparametric approach was also proposed to estimate density for process capability analysis, and Gaussian kernel was adopted for density estimation [Polansky 1998, 2000]. Nonparametric kernel density (KD) estimate is purely data driven, requiring no model training process and prior model knowledge/assumption on population density. Thus it can be a promising approach to accommodate any distribution for capability analysis. Gaussian kernel is one of the commonly used kernels which can simplify conformity assessment [Polansky 1998]. For irregular specification regions and alternative kernels there is no available method for yield assessment. The irregular specification region is common in engineering [Huang et al. 2009], in particular in GD&T evaluation problems. Exploring other kernel alternatives may increase flexibility and improve performance of density estimate of truncated distributions. One way to do this, instead of randomly sampling an actual process, is to generate random samples from the estimated KD, an idea of bootstrap. $\hat{p}$ can thus be simply estimated by the fallout ratio.

It has been recognized that process capability analysis for non-normal multivariate processes remains a great challenge for years to come [Kotz et al. 2002]. The goal of this paper is to contribute a novel approach to enable multivariate, non-normal process capability analysis. To this end, kernel density estimation and bootstrap sampling techniques are integrated to provide a new tool for assessing the nonconforming probability.

2. KERNEL DENSITY ESTIMATION

In process capability analysis the accuracy of results of C_p, C_{pk} are essential for interpretation and decision making. Model (normality) assumption is one of the most critical factors, detrimentally affecting the quality of the results if the distribution feature is incorrectly assumed. Non-normality also complicates the process capability analysis in multivariate processes [Kotz et al. 2002]. When normality is moderately even slightly violated but still accepted the conformity of a process represented by C_p, C_{pk} can be very misleading, as was shown in Polanski [1998]. These non-normality cases are common in practice. For example, as presented by Bisgaard

et al. [1997] the *process bias* in a machining process if a tolerance is subject to *maximum material condition* (MMC) the distribution of a radius of a hole tends to be skewed to its lower tolerance boundary side and a radius of a shaft tends to be skewed to its upper tolerance boundary side because one cannot add materials to a part in machining processes. Thus incurable mistakes of an out-of-spec larger hole or a smaller shaft can be avoided.

Parametric method is appealing because of simplicity. However, it lacks the flexibility to accommodate these non-normal distributions. Nonparametric density estimation is a well established method. The most appealing advantage of it is the flexibility of accommodating to any unknown distributions, i.e., "*letting the data speak for themselves*" or independent of model assumptions.

We assume a process is in a stable status (in-control) but the quality characteristic $\boldsymbol{x}$ is not necessarily normally distributed. The quality characteristics and its distribution are denoted by x and $f(x)$, respectively. If we collected n samples from the process as a training data set, denoted by $\{ \boldsymbol{X}_i = \{X_{i1},...,X_{id}\} \}$, i=1,…n, the density function can be expressed as the smooth *Parzen* estimate for a univariate process:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

or for a multivariate process:

$$\hat{f}(\boldsymbol{x}) = \frac{1}{nh_1...h_d}\sum_{i=1}^{n}\left\{\prod_{j=1}^{d} K\left(\frac{x_j - X_{ij}}{h_j}\right)\right\} \qquad (1)$$

where, d is the dimensions of the process. K(z) is a kernel function. Any smooth function, satisfying $K(z) \geq 0, \quad \int K(z)dz = 1, \quad \int zK(z)dz = 0 \quad \int z^2 K(z)dz \neq 0$ and < ∞, can be used as a kernel. h represents the window width of the kernel function, controlling smoothness of $\hat{f}(\boldsymbol{x})$. The kernel K(.) is a local weighting function, giving more weight to the point $\boldsymbol{X}$ that is closer to the training point $\boldsymbol{X}_i$ and the weight decreases as its distance from $\boldsymbol{X}_i$ increases.

Eq. (1) is essentially a regression, requiring very little model training. Another appealing feature is that the important distribution patterns (skewness, multimodal, thick tails etc.) can be reserved automatically. This is the so called "*letting data speak for themselves*".

There are several widely used kernels in literature. The choice of kernels is less important than the choice of h in terms of the behavior of $\hat{f}(\boldsymbol{x})$ [Marron et al. 1988]. In this paper, we propose using Gaussian kernel:

$$K\left(\frac{x\text{-}X_i}{h}\right) = \frac{1}{(2\pi)^{1/2}h}\exp\left(\frac{-(x\text{-}X_i)^2}{2h^2}\right) \qquad (2)$$

Bandwidth parameter h is a more important factor in density estimation, controlling the smoothness of $\hat{f}(\boldsymbol{x})$. Selection of h strongly relies on how the samples scattered, i.e., σ of x rather than the sample size n as shown below in Eq. (3). When h is too large important features of the underlying distribution (e.g. multimodal) are smoothed away. When h is too small the $\hat{f}(\boldsymbol{x})$ tends to be too wiggly, representing the sample randomness rather than true patterns. Various techniques have been developed for selection of optimal bandwidth. By optimum it means to achieve a trade-off between biase $\hat{f}(\boldsymbol{x})$ and variance of $\hat{f}(\boldsymbol{x})$, and achieving a minimum estimation error. A common measure of the estimation error is the mean integrated squared error (MISE).

$$MISE(h) = E\int(\hat{f}_h(x) - f(x))^2\,dx$$

A review on the optimal bandwidth selection was given in Jones et al. [1996]. These techniques are especially beneficial in automatic bandwidth determination for multiple estimates and dimensionality reduction where manual selection is impractical.

In this paper, we propose to use the *rule of thumb* method for bandwidth selection [Turlach 1993; Jones et al. 1996]. This can be interactively used with visual choice of bandwidth. In process capability analysis one usually does not have many processes to be estimated simultaneously. Another reason is the simplicity for implementation, thus, it is more appealing and desirable for practitioners.

If we take *K* as the Gaussian kernel, the *rule of thumb* optimal *h* for multivariate processes is adopted as presented below:

$$h_j = \left(\frac{4}{n(d+2)} \right)^{\frac{1}{d+4}} \hat{\sigma}_j \quad (3)$$

where $\hat{\sigma}_j$ is the estimated standard deviation of j[th] variable x_j in $\boldsymbol{x}$. h_j in Eq. (3) is sensitive to outliers in $\{\boldsymbol{x}_i\}$ which may cause a too-large estimate of $\hat{\sigma}_j$, and hence a too-large h_i, resulting in an oversmoothed *f*. A more robust alternative to estimate $\hat{\sigma}$ is to use the estimated quantile range $\hat{R} = x_{j,0.75n} - x_{j,0.25n}$. $x_{j,0.25n}$, $x_{j,0.75n}$ denote 25% and 75% quantile points of entry x_j in $\boldsymbol{x}_i$. The robust estimate of $\hat{\sigma}$ is

$$\hat{\sigma} = \frac{\hat{R}}{1.34}$$

With the selected kernel and bandwidth the yield Y_{pk} (will be defined shortly) can be estimated with the proposed sampling algorithm as presented below.

3. YIELD CALCULATION WITH SAMPLING

Widely used process capability indices C_p, C_{pk} usually require normality assumption. These indices indicate the ratios of specification range and the process dispersion. 6σ is used to represent process dispersion (equivalent to a 99.73% conformity range). Alternatives have also been proposed when normality is violated, using the range of distribution quantiles (0.00135 and 0.99865) as the measure of process dispersion.

These indices can be interpreted by the corresponding conformity, indicated by defective part per million or *ppm*. However, the conformity interpretation of C_p, C_{pk} will be significantly different and misleading when normality assumption is moderately even slightly violated.

In this paper, we assume the process under study is in-control which can be assessed by control chart techniques with subgroup samples. We further assume the process quality characteristic (individual sample) can be characterized any known or unknown statistical model. We propose directly using yield or conformity to characterize process capability below. It gives a coherent and unified index for any distribution. This prevents confusion and misunderstanding caused by vagueness in existing indices (C_p, C_{pk}).

The yield is defined as the probability of quality characteristic(s) falling in specification ranges or regions (for multivariate processes). Thus,

$$\begin{aligned} &\text{Yield} = \text{Prob}[\text{LSL} \le x \le \mathit{USL}] \quad \text{(univariate)} \\ &\text{or} \\ &\text{Yield} = \text{Prob}[\boldsymbol{X} \in \Omega_S] \quad \text{(multivariate)} \end{aligned} \quad (4)$$

where x, $\boldsymbol{X}$ are the individual quality characteristic and the multivariate quality characteristic vector, respectively. Ω_S denotes a specification region. Ω_S usually represents a hypercube but can be a complex irregular region. Examples of such irregular region can be found in semiconductor and automotive manufacturing [Huang et al. 2009]. For instance, a composite tolerance on two dimensional variables creates an irregular Ω_S. *USL* and *LSL* are upper and lower specification limits. In a univariate case, the conformity is expressed as:

$$Y_{pk} = \int_{LSL}^{USL} f(x)dx \quad (5)$$

where Y_{pk} can be interpreted as actual conformity (Yield). $f(x)$ is the density function of a quality characteristic x. In the symmetric distribution case this means well centered. *ppm* can thus be conveniently expressed by:

$$ppm_{actual} = (1 - Y_{pk})10^6$$

In multivariate cases, a general expression of conformity is

$$Y_{pk} = \int_{\Omega_S} f(\boldsymbol{X})d\boldsymbol{X} \quad (6)$$

The potential capability similar to C_p is

$$Y_p = \max_{\boldsymbol{\mu}}\{Y_{pk}\} = \max_{\boldsymbol{\mu}}\{\int_{\Omega_S} f(\boldsymbol{X})d\boldsymbol{X}\}$$

where $\boldsymbol{\mu}$ denotes the mean vector of $\boldsymbol{X}$.

Y_{pk} defines the actual capability of process for producing quality products and preferred in application. The interpretation is coherent and independent of model assumption. However, the applicability of Eq. (6) relies on how easy and accurate Y_{pk} can be obtained from process data.

The KD $\hat{f}(\boldsymbol{x})$ in Section 2 estimates a nonparametric statistical model to approximate the true unknown model of the process. With Gaussian kernel density estimate [Polansky 1996, 1999] gave closed-form expressions of process conformity. However, for a multivariate process it is not convenient if one uses different window widths h_j for each dimension or non-Gaussian kernels. Monte Carlo integration provides a generic and flexible method for evaluating the yield in Eq. (6). Since the pattern of the estimated density $\hat{f}(\boldsymbol{x})$ can be complex there is no straightforward approach to directly draw samples from $\hat{f}(\boldsymbol{x})$ (*direct methods)* as in Monte Carlo simulation with well-known parametric models such as normal, uniform, gamma, etc.

In this paper, an alternative indirect method, i.e., *acceptance-rejection method* [Robert et al. 2006] is adopted. The idea is to draw large number (*M*) of random samples that follow the distribution $\hat{f}(\boldsymbol{x})$, and then use them to estimate conformity probability. This KD based nonparametric bootstrap sampling is different from the direct bootstrap because it can predict the tails beyond the range of the observed data. And the information in the tails is of fundamental importance in yield estimation. In comparison, this method is more flexible, easy to implement, and independent of kernel function and the types of specification region. The training data collected from a manufacturing process are usually costly and limited, cannot be used for direct yield assessment. Instead, large sample set can be generated economically from KD for the evaluation. The computation cost of the sampling for density evaluation at *M* target points is linear, i.e., *O*(*M*). The algorithm involves: i) generating random samples from $\hat{f}(\boldsymbol{x})$; ii) calculating yield or the ratio of the number of conforming samples to the total number of the samples, i.e.

$$Y_{pk}=1-\frac{\#\ nonconforming\ samples}{total\ \#\ of\ samples} \qquad (7)$$

The one-side fall-outs can also be easily counted, which provides off-center information and is useful in estimating the optimal position of the process (Y_{pk}). Y_p can also be obtained by searching the maximum Y_{pk} when the mean $\boldsymbol{\mu}$ is running through Ω_s. In addition, sampling method can accommodate both regular and complex irregular specification regions in multivariate processes. This property is highly desirable in geometrical and dimensional tolerancing (GD&T) related capability study. The interrelated tolerance zones in GD&T often result in complicated specification regions, rendering the capability analysis intractable with current PCR based methods.

The *acceptance-rejection* method involves specifying a function *t* that *majorizes* the KD $\hat{f}(\boldsymbol{x})$, i.e., $t(\boldsymbol{x}) \geq \hat{f}(\boldsymbol{x})$ for all $\boldsymbol{x}$. To facilitate sampling, we need to transform $t(\boldsymbol{x})$ into a density function, i.e., using a density function $r(\boldsymbol{x})$:

$$r(\boldsymbol{x})= t(\boldsymbol{x})/\mathrm{c},\ \text{and}\ \mathrm{c}=\int_{\infty}^{\infty} t(x)dx$$

in the following procedure:

1. Generate *Z* having density $r(z)$
2. Generate *U*~U(0,1), independent of *Z*
3. If $U \leq \hat{f}(\boldsymbol{Z})\ /t(Z)$, return *X=Z* (*Accept Z*). Otherwise, go back to step 1 (*Reject Z*) and try again.

All of the accepted *Z*'s give the sample set drawn from $\hat{f}(\boldsymbol{x})$ and will be used in yield evaluation in Eq. (7). The $t(\boldsymbol{x})$ must be selected so that it is easy to generate samples from $r(\boldsymbol{x})$. Such a $r(\boldsymbol{x})$ can be determined as below:

1. Let $t = \underset{x}{Max}\{\hat{f}(\boldsymbol{x})\}$
2. Let *V* be the volume of the domain in which $\hat{f}(\boldsymbol{x})$ is defined. *V* can be determined from the process knowledge
3. $r(\boldsymbol{x})=t/V$.

The $r(\boldsymbol{x})$ is a constant defined on *V*. Hence it is easy to generate samples from by using uniform distributions. The $\underset{x}{Max}\{\hat{f}(\boldsymbol{x})\} = \hat{f}(\boldsymbol{x}^*)$ and $\boldsymbol{x}^*$ can be derived by letting the partial derivatives equate zero, i.e.,

$$\frac{\partial \hat{f}}{\partial x_j} = 0,\ j = 1,\ldots,d$$

Thus $$x_j^* = \frac{\sum_{i=1}^{n} exp\left(-\frac{1}{2}\sum_{j=1}^{d}\left\{\frac{x_j^* - X_{ij}}{h_j}\right\}^2\right) X_{ij}}{\sum_{i=1}^{n} exp\left(-\frac{1}{2}\sum_{j=1}^{d}\left\{\frac{x_j^* - X_{ij}}{h_j}\right\}^2\right)}.$$

This is a univariate equation that can be conveniently solved numerically.

4. CASE STUDY

4.1. Case I: Simulation Based Validation

Beta functions are selected to simulate a non-normal multivariate process, i.e., $\boldsymbol{x} = \{x_1, x_2, x_3\}$, x_1 ~ Beta(2,3), x_2 ~ Beta(2,6), and x_3 ~ Beta(3,1.5). The baseline yield is obtained by generating a very large sample set (250,000) and calculating the conformity against a given specification limit. To validate the proposed method, another 50 samples are generated independently from the same distributions and are used for $\hat{f}(\boldsymbol{x})$ estimation. The scatter plot and histogram of the samples is shown in Figures 1 and 2. Then $t(x) = \underset{x}{Max}\{\hat{f}(\boldsymbol{x})\}$ can be obtained (here actually $t(x)$ is a constant). The volume of the variable space is known from the three beta distributions. Thus, $r(z)$ is a multivariate uniform distribution. Using $r(z)$ and following the procedure in Section 3, bootstrap sampling is then carried out to generate samples for yield calculation.

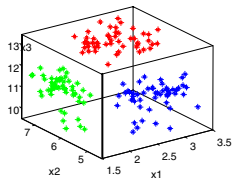

FIGURE 1. SCATTER PLOT OF x_1,x_2,x_3.

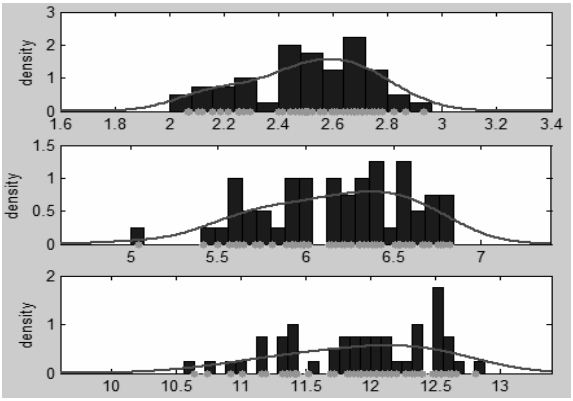

FIGURE 2. HISTOGRAMS OF 50 SAMPLES (UPPER, MIDDLE, AND LOWER ARE x_1, x_2, and x_3, RESPECTIVELY).

The baseline yield=0.9972, and the result estimated from the KD estimation and *acceptance-rejection* sampling (from 50,000 samples) gives yield=0.9771. The relative error is about 2%. This is a very good estimate considering only a small (50) training sample set is used. The computation elapses about 1832 *sec.* on a Pentium IV (1.8 GHz).

In the calculation process, we tried different window widths. Small window width generally gives better yield estimation results. This is because the window width controls the smoothness of the KD. Increasing the window width will over-smooth the KD at the tails (thickening the tails), and thus produce more nonconforming samples.

4.2. Case II: Impurity Profiles of Drug Substances

A case from pharmaceutical industry [Gonzalez et al. 2003] is picked to demonstrate the method. It involves the study of the consistency of the impurity profiles of drug substances. The impurity profile is usually multivariate in nature and is complicate, prohibiting examining each impurity by univariate statistical procedures.

A domestic company manufactures a crystalline drug substance, which has an impurity profile consisting of several major organic impurities chemically related to the drug substance. The names of the major impurities are confidential and denoted by capital letters A to G.

The Multivariate Statistical Process Control (MVSPC) chart (Hotelling's Multivariate T^2) and Mason–Tracy–Young (MYT) Decomposition Procedure was used to ensure the process in-control.

In the case study, 79 in-control points are selected. And three variables (A, C, and G) are used to represent the multivariate process. Logarithmic transformations are used to change the scale of different variables. The transformations do not affect the yield calculation because both the data and the specification limits are transformed simultaneously. It is common in process industries that the individual measurement is not normal as shown in Figure 3.

The kernel densities (KD) are estimated from these 79 samples as shown in Figure 3.

Following the same procedure we estimated $\hat{f}(\boldsymbol{x})$ and *t(x).* The volume *V* of the variable space is estimated based on the upper and lower boundaries of these variables which can be obtained from the process. A conservative (i.e., a larger one) will not affect the result but may reduce the computation efficiency. Figure 4 illustrates the samples drawn from the KDs. And the estimated yield is 0.9254. The similar computation time as in Case I is used in the calculation. It is largely consumed on kernel functions evaluation in the sample generation process. A large proportion of sample rejection is observed, showing a poor efficiency in sample generation.

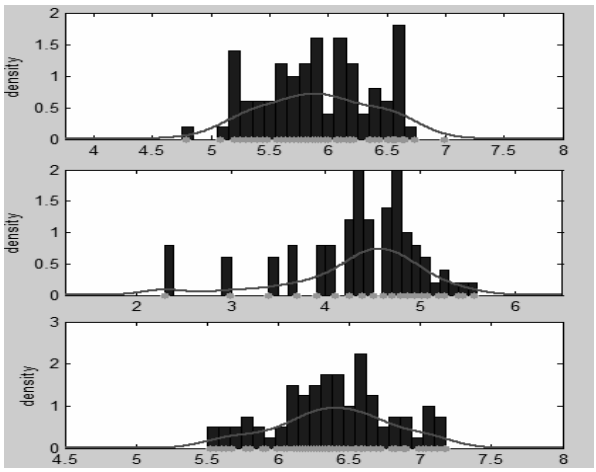

FIGURE 3. HISTOGRAMS AND KDs OF THE SAMPLES (UPPER, MIDDLE, AND LOWER ARE A, C, AND G, RESPECTIVELY).

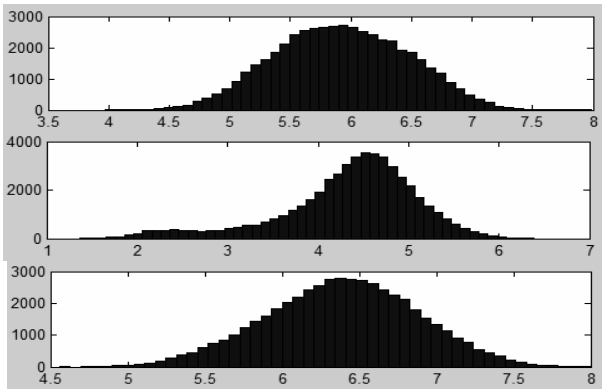

FIGURE 4. GENERATED SAMPLES FROM KERNEL DENSITIES FOR YIELD CALCULATION.

5. SUMMARY

Process capability study prevails in manufacturing for quality assurance. Current techniques involve the process capability ratio (PCR) analyses (e.g., C_p, C_{pk}) that require normality assumption and are largely confined to univariate processes. It has been very difficult to analyze process capability for non-normal multivariate processes in aerospace, automotive, semiconductor, chemical, and pharmaceutical industries. For instance, the interrelated tolerance zones (e.g., size, orientation, and form) and multivariate nature in GD&T evaluation render it extremely difficult, if not impossible, to evaluate the process capability using conventional PCR methods.

A new method is proposed in this paper. It integrates a nonparametric (kernel) method and a sampling technique to enable the capability analysis of multivariate and non-normal processes. Conformity or yield is introduced as an index for the capability evaluation. Yield provides direct quality information in terms of conforming rate (ppm) which has an unambiguous interpretation and is conceptually transparent to practitioners. Kernel density is estimated from a small sample set from a process to characterize the statistical model behind the process. The advantage is the model assumption free, i.e., "*letting the data speak for themselves*". The subjectivity and the reliance on the skill and expertise in model selection and fitting for non-normal processes can thus be avoided. The KD requires no model training and can capture the patterns in data automatically. The *acceptance-rejection* sample generation algorithm is adopted to enable Monte Carlo calculation for yield analysis.

The new method is easy to implement. The codes developed for KD estimation and sample generation in MATLAB are extremely simple. To enhance the applicability, the intensive mathematical and computational involvements are put behind the screen, i.e., *letting the computer sweat behind the screen*.

Future efforts are expected to alleviate the computation intensity, in particular for higher dimensions (*d*) and larger training samples. The *acceptance-rejection* algorithm usually has a low acceptance rate. Thus, to obtain sufficiently large samples for yield evaluation it requires very large number of kernel function evaluations. The

integration based alternatives, such as direct integration or Monte Carlo, and sampling algorithms in Markov Chain Monte Carlo (MCMC) techniques such as Metropolis-Hastings algorithm can be more appealing candidates in efficiency and applicability, motivating the authors' current research in this vein. Nevertheless, compared with the costly data from real manufacturing processes for capability study the computation cost would not be a serious concern.

REFERENCES

Bisgaard, S. and S. Graves (1997). A negative process capability index from assembling good components? A problem in statistical tolerancing. *Quality Engineering*, **10**(2):409-414.

Carr, W.E. (1991). A new process capability index: Parts per million. *Quality Progress*, **24**(2):152-154.

Clements, J.A. (1989). Process capability calculation for non-normal distributions. *Quality Progress*, **22**:95-100.

Flaig, J.J. (2002). Process capability optimization. *Quality Engineering*, **15**(2):233-242.

Flaig, J.J. (1999). Process capability sensitivity analysis. *Quality Engineering*, **11**:587-592.

Gonzalez-de la Parra, M.G. and P. Rodriguez (2003). Application of multivariate T^2 control chart and the Mason-Tracy-Young decomposition procedure to the study of the consistency of the impurity profiles of drug substances. *Quality Engineering*, **16**(1):127-142.

Huang, W., T. Phoomboplab, and D. Ceglarek (2009). Process capability surrogate model-based tolerance synthesis of multi-station manufacturing systems. *IIE Transactions*, **41**(4):309-322.

Jones, M.C., J.S. Marron, and S.J. Sheather (1996). Progress in data-based bandwidth selection for kernel density estimation. *Computational Statistics*, **11**:337-381.

Kane, V.E. (1986). Process capability indices. *Journal of Quality Technology*, **18**:41-52.

Kotz, S. and N.L. Johnson (1993). *Process capability indices*. London: Chapman and Hall.

Kotz, S. and N.L. Johnson (2002). Process capability indices—A review, 1992-2000. *Journal of Quality Technology*, **34**(1):2-19.

Marron, J.S. and D. Nolan (1988). Canonical kernels for density estimation. *Statistics & Probability Letters*, **7**(3):195-199.

Pal, S. (2005). Evaluation of nonnormal process capability indices using generalized lambda distribution. *Quality Engineering*, **17**:77-85.

Palmer, K. and K. Tsui (1999). A review and interpretations of process capability indices. *Annals of Operational Research*, **87**(1999):31-47.

Polansky, A.M. (1998). A smooth nonparametric approach to process capability." *Quality Reliability Engineering International*, **14**:43-48.

Polansky, A.M. (2000). An algorithm for computing a smooth non-parametric capability estimate. *Journal of Quality Technology*, **32**:284-289.

Robert, C.P. and G. Casella (2006). *Monte Carlo statistical methods*. New York: Springer-Verlag.

Silverman, B.W. (1986). *Density estimation for statistics and data analysis*, Vol. 26 of Monographs on Statistics and Applied Probability. London: Chapman and Hall.

Sullivan, L.P. (1984). Reducing variability: A new approach to quality. *Quality Progress*, **17**(7):15-21.

Tsui, K. (1997). Interpretation of process capability indices and some alternatives. *Quality Engineering*, **9**(4):587-596.

Turlach, B.A. (1993). Bandwidth selection in kernel density estimation: A review. *Discussion Paper 9307*, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

Wand, M.P. and M.C. Jones (1995). *Kernel smoothing*, Vol. 60 of Monographs on Statistics and Applied Probability. London: Chapman and Hall.

Yeh, A.B. and S. Bhattacharya (1998). A robust capability index. *Communication in Statistics—Simulation and Computation*, **27**:565-589.

Yeh, A.B. and H. Chen (2001). A nonparametric multivariate process capability index. *International Journal of Modeling & Simulation*, **21**(8):218-223.