Process Capability Assessment for Univariate and Multivariate Non-normal Correlated Quality Characteristics

By

Shafiq Ahmad

Master of Engineering
Asian Institute of Technology, Bangkok, Thailand &
Technical University Hamburg Harburg (TUHH), Germany

This thesis is submitted in total fulfillment of the requirement for the degree of Doctor of Philosophy

School of Mathematics & Geospatial Sciences

RMIT University

Melbourne, AUSTRALIA

Thesis Supervisors:
Dr. Malihe Abdollahian and Professor Panlop Zeephongsekul

September, 2009

Abstract

In today's competitive business and industrial environment, it is becoming more crucial than ever to assess precisely process losses due to non-compliance to customer specifications. To assess these losses, industry is extensively using Process Capability Indices for performance evaluation of their processes. Determination of the performance capability of a stable process using the standard process capability indices such as C_p and C_{pk} requires that the underlying quality characteristics data follow a normal distribution. However it is an undisputed fact that real processes very often produce non-normal quality characteristics data and also these quality characteristics are very often correlated with each other. For such non-normal and correlated multivariate quality characteristics, application of standard capability measures using conventional methods can lead to erroneous results.

The research undertaken in this PhD thesis presents several capability assessment methods to estimate more precisely and accurately process performances based on univariate as well as multivariate quality characteristics. The proposed capability assessment methods also take into account the correlation, variance and covariance as well as non-normality issues of the quality characteristics data.

It is an established fact that the fundamental objective of all capability measures is to help process engineers and managers decide whether to

accept or reject the process outcomes based on conformance to customer (engineering) specifications. This research has therefore focused on assessing the efficacy of our proposed methods using the Proportion of Non-Conformance (PNC) criterion, which is frequently used in practice to assess the utility of PCI methods. This has been further supplemented by using our proposed methods to estimate the capability of processes from the real world through application to data obtained from the manufacturing industry.

A comprehensive review of the existing univariate and multivariate PCI estimations have been provided. We have proposed fitting Burr XII distributions to continuous positively skewed data. The proportion of nonconformance (PNC) for process measurements is then obtained by using Burr XII distribution, rather than through the traditional practice of fitting different distributions to real data. Maximum likelihood method is deployed to improve the accuracy of PCI based on Burr XII distribution. Different numerical methods such as Evolutionary and Simulated Annealing algorithms are deployed to estimated parameters of the fitted Burr XII distribution.

We have also introduced new transformation method called Best Root Transformation approach to transform non-normal data to normal data and then apply the traditional PCI method to estimate the proportion of non-conforming data. Another approach which has been introduced in this thesis is to deploy Burr XII cumulative density function for PCI estimation

using Cumulative Density Function (CDF) technique. The proposed approach is in contrast to the approach adopted in the research literature i.e. use of best-fitting density function from known distributions to non-normal data for PCI estimation. The proposed CDF technique has also been extended to estimate process capability for bivariate non-normal quality characteristics data.

A new multivariate capability index based on the Generalized Covariance Distance (GCD) is proposed in this research thesis. This novel approach reduces the dimension of multivariate data by transforming correlated variables into univariate ones through a metric function. This approach evaluates process capability for correlated non-normal multivariate quality characteristics. Unlike the Geometric Distance (GD) approach cited in the research literature, GCD approach takes into account the scaling effect of the variance–covariance matrix and produces a Covariance Distance (CD) variable that is based on the Mahanalobis distance. Another novelty introduced in this research is to approximate the distribution of these distances by a Burr XII distribution and then estimate its parameters using numerical search algorithm. It is demonstrates that the proportion of nonconformance (PNC) using proposed method is very close to the actual PNC value.

Statement of Authorship

The work of this thesis has not been submitted previously for a degree or diploma at any university or institution. Except where explicit reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or been awarded another degree or diploma. No other person's work has been relied upon or used without due acknowledgment in the main text and bibliography of the thesis. Some of the material, methods, figures and results from this dissertation have appeared in the refereed conference and journal publications during my Ph.D. study period.

Signed: --

Dated: --

Shafiq Ahmad

Candidate

Acknowledgement

I would like to start off my words here by giving my humble thanks to Almighty Allah for bestowing His blessings in the form of a supporting family, strength, knowledge and a very supportive group of friends and supervisors, which have enabled the completion of this research. I take this opportunity to express my appreciation to the special people who have shaped my career and a fulfilling research candidature at RMIT University.

Foremost, I wish to express my profound gratitude to my supervisors Dr. Malihe Abdollahian and Professor Panlop Zeephongsekul. Their invaluable support, advice and periodic reviews on day-to-day activities enhanced my confidence to accomplish this research. I am indebted for their encouragement and inspirational advices received throughout the association I had with them.

I would like to thank the Head of the School Professor John Hearne, Head of the Mathematics discipline Associate Professor John Shepherd, Research Coordinator Associate Professor David Fraser, Senior Lecturer Kaye Marion and many other staff of the School of Mathematics for providing me with assistance during my study period at RMIT. I am indebted also to my colleagues of the School for their valuable suggestions, and include them in my thanks.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1. Introduction

The evolution of market dynamics over the past century has virtually shifted the business paradigm from producer to customer. At the beginning of 19th century the market was known to be "producer oriented", namely, anything produced in bulk could be sold. Today the customer plays a driving force in design, development and the quality of the product. With frequent advances in technology, customers need products that meet their not only challenging requirements but also with competitive price (Ashraf 2009). This competition has forced manufacturers to continuously improve their products and services, which can address the stringent customer expectations with minimal operational losses.

Since the early eighties of the last century, process capability indices (PCIs) played vital role to improve the operational efficiency of manufacturing products and processes thus resulting in significant reduction of process losses that occur due to non-compliance to customer specifications. Although process capability indices such as C_p and C_{pk} are being extensively applied in industry to assess process performances but there is a lack of understanding among quality practitioners that these capability measures are essentially based on statistical theory of normality. If the basic

assumptions of statistical theory are violated the capability assessments can mislead to wrong conclusions (Deleryd, M. 1998).

The research undertaken in this thesis presents capability assessment methods to estimate more precisely and accurately process performances based on single as well as multiple quality characteristics. Proposed methods also take into account the correlation, variance and covariance as well as non-normality issues of the quality characteristics data. It is an established fact that the fundamental objective of all capability measures is to help process engineers and managers decide whether to accept or reject the process outcomes based on conformance to customer (engineering) specifications. This research has therefore focused on presenting the efficacy of our proposed methods using the Proportion of Non-Conformance (PNC) criterion, which is frequently used in practice to assess the utility of PCI methods.

The remainder of this chapter elaborates the basics of the capability indices and provides literature review of this research, the main research problems and questions, the proposed approaches to the research problem, the major contributions of this research and a detailed outline of this thesis. The list of publications resulting from this research is given in the last section.

1.2.   Background

Process capability analysis together with statistical process control and design of experiments are statistical methods that have been used for decades with the main purpose being to reduce the variability in industrial processes and products (Albing 2006). The need to understand and control

processes is getting more and more relevant due to the increasing complexity in technical systems in industry. Moreover, the use of statistical methods in industry is also increasing by the introduction of quality management concepts such as the Six Sigma programme, where statistical methods, including process capability indices, are important parts (Hahn et al. 1999).

Process capability analysis deals with how to assess the capability of a manufacturing process, where information about the process is used to improve the capability. With process capability analysis one can determine how well the process will perform relative to product requirements or specifications. However, before assessing the capability of a process it is important that the process is stable and repeatable. That is, only natural (common) causes of variation should be present. It should be noted that a process capability analysis could be preformed even if the process is unstable. However, such an analysis will give an indication of the capability at that very moment only and hence the results are of limited use (Deleryd, M & Vännman 1999).

To check if the process is stable, statistical process control is usually applied. The purpose of statistical process control is to detect and eliminate assignable causes of variation and control charts are usually used in order to determine if the process is in statistical control and reveal systematic patterns in process output. An introduction to statistical process control can be found in Montgomery (Montgomery, DC 2005b).

When the process is found stable, different techniques can be used within the concept of process capability analysis in order to analyse the capability (Montgomery, DC 2005b). For instance, a histogram along with sample statistics such as average and standard deviation gives some information about the process performance and the shape of the histogram gives an indication about the distribution of the studied quality characteristic. Another simple technique is to determine the shape, centre and spread of the distribution by using a normal probability plot.

The above-mentioned tools give some approximated information only about the process capability. The most frequently used tool when performing a capability analysis is called process capability index. A process capability index is a unit-less measure that quantifies the relation between the actual performances of the process and its specified requirements. In general, the higher the value of the index, the lower the amount of products outside the specification limits. If the process is not producing an acceptable level of conforming products, improvement efforts should be initiated. These efforts can be based on design of experiments. By using design of experiment one can, for instance, identify process variables that influence the studied characteristic and find directions for optimizing the process outcome. An introduction to design of experiments can also be found in, e.g.(Montgomery, DC 2005a).

Process capability indices (PCIs), as well as many other statistical methods, are based on fundamental assumptions. For instance, the most widely used process capability indices in industry today analyse the capability of a

process under the assumptions that the process is stable and that the studied quality characteristic is independent and normally distributed (Albing 2006). To understand these conventional capability indices, consider $X_1,.......,X_n$ to be the actual values of a certain quality characteristic which correspond to n randomly selected items from a production process and suppose that such characteristic should lie between lower specification limit (Lsl) and upper specification limit (Usl) to conform to engineering specifications. Items which lie outside (Usl, Lsl) specifications will be considered non-conforming. The special cases where only one specification limit is required are obtained by letting $Lsl \to -\infty$ or $Usl \to \infty$.

Process capability is designed to monitor the proportion of items which are expected to fall outside the engineering specifications to prevent an excessive production of non-conforming output. This is usually done at a specified rating periods, using the measurements $\{Y_1, Y_2..............,Y_n\}$ taken on, say, n produced items and assuming that

- there is no measurement error, so that $Y_i = X_i, i = 1,......,n$, i.e. the measurements are taken to be the actual values, and
- the X_i are identically distributed with, say, process mean μ and standard deviation σ.

Traditional capability analysis then proceeds to evaluate capability indices which relate the allowable spread of the process $Usl - Lsl$ to its natural spread, customarily taken to be 6σ (Bernardo J 1996). Under these

5

assumptions the most frequently used index "C_p" in industry was presented by (Juran 1974).

$$C_p = \frac{Usl - Lsl}{6\sigma} \qquad (1.1)$$

i.e. C_p is the ratio of the allowable spread to the natural spread. In particular, if the X_i is normally distributed $N(\mu, \sigma^2)$, and the process is centred at nominal mean defined by

$$m = \frac{Usl + Lsl}{2} \qquad (1.2)$$

i.e. $E(x) = \mu = m$, then a capable process, is a process for which $C_p \geq 1$, will result in , at most, 0.27 % of non-conforming items, i.e. 2700 non-conforming items per million items produced in a production process.

The traditional capability index in equation (1.1) only takes care of the process spread. Obviously it would be possible to have any proportions of items outside the specification limits by merely relocating the process mean, thus, C_p only quantifies the potential performance of the process , which will only be attained if the process is centred at the midpoint of the specification limits. If the process is not centred but skewed to upper or lower side of the

6

specification limits, then the upper and lower one-sided capability indices can be applied. These one-sided capability indices are given in equations (1.3) and (1.4):

$$C_{pu} \ (\text{Upper } C_p) \ = \ \frac{Usl - \mu}{3\,\sigma} \qquad (1.3)$$

$$C_{pl} \ (\text{Lower } C_p) \ = \ \frac{\mu - Lsl}{3\,\sigma} \qquad (1.4)$$

where μ and σ are the mean and standard deviation of the in-control process respectively.

If the process in not centred at the midpoint of engineering specifications (Usl, Lsl), then the actual performance of a production process is traditionally measured by C_{pk}, off-centred process capability index (equation 1.5). C_{pk} was defined by (Kane 1986) and is the minimum value of the upper or lower capability indices.

$$C_{pk} \ = \ \min\left\{\frac{Usl - \mu}{3}, \ \frac{\mu - Lsl}{3}\right\} \qquad (1.5)$$

which is a normalized distance between the process mean and its closest specification limit.

From above discussion it is evident that the capability index C_p measures the allowable range of measurements related to the actual range of

measurements and C_{pk} measures the distance between the expected value and the closest specification limit related to half of the actual range of measurements. If the quality characteristic is normally distributed and the process is well centred, i.e. the process mean is located at the midpoint of the two-sided specification interval i.e. $\mu = m = (Usl + Lsl)/2$, $C_p \geq 1$ implies that the number of values of the studied characteristic outside the specification limits will be small (Pearn & Kotz 1994) and obviously non-conformance to customer (engineering) specifications will be minimum. In fact non-conformity ratio (PNC) is the main interpretation of process capability C_p index.

In this research the efficacy of the proposed methods will be presented using PNC criterion. Therefore it is necessary to understand the basics of PNC and its relationship with C_p first. It is interesting to note that C_p in its computation has a direct link with the proportions of items falling outside the specifications limits (Telmoudi 2005). Consider a quality characteristic X, under the normality assumption, the proportion of non-conforming items (PNC) is expressed as:

$$P\big[X > Usl\big] + P\big[X < Lsl\big]$$

$$\Rightarrow \quad P\left[\frac{X-\mu}{\sigma} > \frac{Usl-\mu}{\sigma}\right] + P\left[\frac{X-\mu}{\sigma} < \frac{Lsl-\mu}{\sigma}\right]$$

$$\Rightarrow \quad PNC = 1 - \Phi\left[\frac{Usl - \mu}{\sigma}\right] + \Phi\left[\frac{Lsl - \mu}{\sigma}\right]$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du \qquad (1.6)$$

If μ is substituted by midpoint of specification spread ($(Usl + Lsl)/2$) as given in (1.2) i.e. then the probability of non-conformance (PNC) can be expressed as

$$PNC = 2\Phi(-3C_p) \qquad (1.7)$$

where $\Phi(x)$ is the cumulative distribution function of the unit Gaussian.

It is important to notice that the C_p index depends heavily on standard deviation (σ) of the in-control process. Table 1.1 shows the probability of non-conformance (PNC) for some given values of C_p (Kotz & Johnson 1993). It is established fact that if the C_p index exceeds 1.33, that means the items falling outside the required limits are small in numbers (63 parts only falls outside the specifications per million produced parts) and the process is deemed to be capable enough to fulfil customer requirements. Contrary to this, if the index is smaller than 1.33 but larger than 1, it is recommended to examine the process as the PNC number will rise. For index values less than 1, the process is considered incapable.

Table 1.1 Minimum expected Proportions of Non-Conformance (PNC) items

C_p	2.00	5/3=1.67	4/3=1.33	1.00	2/3=0.667	1/3=0.33
PNC	0.002ppm	0.57ppm	63ppm	2700ppm	45500ppm	317300ppm

[ppm = parts per million]

It is evident from above discussions that the intuitive basis of the capability indices C_p and C_{pk} heavily depends on an implicit assumption of normality of the process output.

1.3. Research problems

It is an established fact that quality has always been an integral part of virtually all businesses and services. In today's competitive business environment, quality improvement paradigm has shifted from being a departmental goal to being an umbrella objective for an entire organization. Companies which can enhance quality levels of their manufactured products faster by reducing non-conformities to customer specifications/ requirements can only lead and sustain their market position in this ever changing era.

Since the advent of statistical quality control, process capability measures C_p and C_{pk} has played a pivotal role for quality improvement endeavours in industry. However, as discussed earlier the traditional capability indices depend heavily on the assumption that the process under examination must be under control and stable and the quality characteristics of the process outcome must be independent and normally distributed. In the real world, process data do not always follow a normal distribution. Numerous statisticians and quality engineers have performed research on non-normal process capability indices; e.g. (Zimmer & Burr 1963), (Burr, IW 1967), (Clements 1989), (Pyzdek 1992), (Kotz & Johnson 1993), (Boyles 1994),

(Wright 1995), (Castagliola, P & Castellanos 2005) (Somerville & Montgomery 1996), (Chou & Cheng 1997), (Kotz & Lovelace 1998), (Deleryd, M. 1998),(Wu, Wang & Liu 1998), (Tang & Than 1999), (Kotz & Johnson 2002), (Liao, Chen & Li 2002), (Liu, P & Chen 2006) but the main focus of their research was concentrating on processes with a single quality characteristic. A new "Burr XII distribution based percentile" approach to estimate PCI for non-normal univariate data is proposed by Liu and Chen (Liu, P & Chen 2006) which uses Burr XII distribution for non-normal quality characteristics data. It is well known that Burr XII distribution is very versatile to fit any real data (Burr, IW 1973). This makes it more attractive to extend its application to the multivariate PCI's estimations.

To date there are a few articles which have discussed the capability measures for non-normal multivariate quality characteristics. In recent years, a number of research papers appeared on multivariate process capability analysis, for example, the work by (Boyles 1994; Kotz & Johnson 1993), (Somerville & Montgomery 1996), (Bernardo J 1996; Wang & Hubele 1999), (Noorossana 2002),(Pal 2005), (Wang 2006), (Chen, K, Hsu & Wu 2006), (Niaki & Abbasi 2007). However, most of the research focuses on multivariate normal process capability measures. In reality, very often these quality characteristics are non-normal and correlated. According to the literature (e.g. (Wang 2006)) multivariate PCIs that have been proposed by many researchers suffer from the following constraints and limitations:

▪ Normality assumption on multivariate data is usually required.

- Confidence intervals of the multivariate capability indices are difficult to derive.

- Higher-dimension (more than three quality variables) PCIs are not readily obtainable, except through projection of multivariate data into univariate variables such as the geometric distance approach proposed by Wang et al. ((Wang & Hubele 1999), (Wang & Du 2000)).

Due to the above limitations, it is evident that the application of conventional methods is somewhat limited. In order to deal with non-normal multivariate and correlated quality characteristics data, there is ample opportunity for researchers to develop more suitable PCIs that can address the complex situation of multivariate non-normal and correlated data. To investigate the possible solutions for the multivariate PCI with correlated non-normal quality characteristics, this research attempts to find the answers to the following questions:

- What current methods for calculating the process capability are widely available and commonly used for non normal process data with single quality characteristic? The need to perform a comprehensive comparison with respect to their applicability to the real world problems is also carried for.

- Can Burr XII distribution be utilized with real data and be able to provide the best estimate of non-normal process capability indices with univariate data?

- Can we extend Burr XII distribution function to more than one quality characteristics when the quality characteristics are dependent and correlated?

- Can we use multivariate Burr model to calculate the process capability indices for simulated and real data?

- Will application of Burr XII distribution provide an accurate, efficient and practical procedure for process capability estimation with multivariate quality characteristics?

Throughout the course of this thesis, these research questions were addressed and investigated. The proposed approaches to deal with the problems that arise and the outcomes of the research are described in the next section.

1.4. Proposed approaches to research problems

As mentioned earlier, when the data is non-normal, measuring process capability using conventional methods can lead to erroneous conclusions. Different PCI methods have been proposed to deal with the non-normality issue. Although these methods are practiced in industry, there is insufficient literature to assess the accuracy of these methods under mild and severe departures from normality. In this thesis, we will firstly review the performance of the existing capability estimation methods (e.g. Clements percentile method, Burr based percentile method and Box Cox method) for non-normal univariate quality characteristics data. A simulation study using known non-normal distributions (e.g. Weibull, Gamma, Beta and lognormal) will be conducted to compare the performance of some of the commonly

used methods. Accuracy of the Burr based percentile method will be further improved using numerical search algorithms (e.g. Simulated Annealing, Evolutionary and Compass Direct Search algorithms). We will further explore new capability estimation methods such as cumulative density function and Best root transformation methods by fitting Burr distribution to simulated and real data. In the later part of this research we will extend Cumulative Density Function method to bivariate process capability estimation by fitting bivariate Burr distribution to bivariate non-normal quality characteristics data.

To achieve our major objective later in this thesis, we will propose an approach for dealing with multiple correlated quality characteristics. In the proposed approach, we will first cluster correlated quality characteristics and then define a variable, referred to as the Covariance Distance (CD) variable which is the distance of individual quality characteristics from their respective targets scaled by their variance–covariance matrix. CD is well known in pattern recognition literature as the Mahanalobis distance (Devroye, Györfi & Lugosi 1996). The proposed approach will be similar to the geometric distance (GD) approach adopted by Wang (Wang 2006), but it differs in that the scaling factor of the variance–covariance matrix is absent in GD. Furthermore, unlike the approach in (Wang 2006), we will fit Burr XII distribution (Burr, IW 1942) to the CD data instead of fitting different distributions to GD data, as was done by Wang. The parameters of the fitted Burr XII distribution are obtained using different numerical search techniques. Application examples with real data from manufacturing industry

have been presented in this research study to show the relevance of theory developed to industry.

1.5. Contributions

The major contributions arising from this research are:

1.5.1.   Univariate process capability

- A comprehensive review of the existing PCI estimation methods.

- Instead of applying the traditional moment matching method cited in the research literature, Maximum Likelihood Estimation method has been deployed to improve the accuracy of PCI estimation based on Burr percentile method.

- Novel numerical methods have been proposed to estimate the parameter's of the fitted Burr distribution.

- A new approach called Best Root Transformation (BRT) is proposed which enables to transform non-normal data to normal by searching for the optimal root for data transformation.

- Another contribution to this thesis is to propose Burr cumulative density function for PCI estimation using Cumulative Density Function approach. The proposed approach is in contrast to the approach adopted in the research literature i.e. use of best-fitting density function from known distributions to non-normal data for PCI estimation.

1.5.2.　Multivariate process capability

- Cumulative Density Function method has been extended to estimate bivariate non-normal PCI where bivariate Burr distribution is fitted to bivariate non-normal quality characteristics.

- A novel approach to estimate multivariate non-normal PCI has been introduced. This proposed approach called "Generalized Covariance Distance (GCD)" approach, evaluates process capability for correlated non-normal multivariate quality characteristics. Proposed approach has the following novel features:

  i.   It is based on the idea of reducing the dimension of multivariate data by transforming correlated variables into univariate ones through a metric function.

  ii. Unlike the Geometric Distance (GD) approach cited in the research literature, our approach takes into account the scaling effect of the variance–covariance matrix and produces a CD variable that is based on the Mahanalobis distance.

  iii. It is demonstrated that the proposed GCD approach does not assume that the CD variables are mutually independent, which is implicitly assumed in the Geometric Distance approach.

  iv. In contrast to the GD approach, where different distributions are fitted to different GD variables, a single distribution, the Burr XII distribution is fitted to the CD data. Numerical search techniques are used to estimate the parameters of the Burr distribution.

- Application examples using real data with several non-normal quality characteristics from the manufacturing industry have been presented.

1.6. Organization of the thesis

The subsequent chapters have been organized as follows.

Chapter 2 covers a literature review and basic principles of process capability measures for non-normal quality characteristics. Current approaches to estimate PCIs for univariate and multivariate non-normal quality characteristics are also covered in this chapter.

Numerical techniques and procedures to estimate PCI's developed during this research thesis have been presented in Chapter 3.

In Chapter 4, a comprehensive comparison of existing PCI estimation methods have been discussed. This chapter also covers proposed Cumulative Density Function (CDF) and Best Root Transformation PCI estimation methods. A comparison of the results using simulated and real data have also been presented in this chapter

PCI estimation using bivariate non-normal data are discussed in Chapter 5.

The existing multivariate PCI approaches; particularly Geometric Distance (GD) approach and the proposed Generalized Covariance Distance approach are discussed in chapter 6. Several application examples based on real data are also included in this chapter.

The thesis concludes with Chapter 7, which also includes recommendations for further research in this area.

1.7. Publications

The publications, based on research carried out during the tenure of this research, are listed below:

1.7.1.   Refereed journal papers

1. S. Ahmad, M. Abdollahian, P. Zeephongsekul "Multivariate Non-normal Process Capability Analysis, International Journal of Advance Manufacturing Technology, (2009) 44:757-765.

2. S. Ahmad, M. Abdollahian, P. Zeephongsekul, "Process Capability Estimation for non-normal quality characteristics using Clement, Burr and Box-Cox methods", The Australian and New Zealand Industrial and Applied Mathematics Journal, pp. C642-C665, 2008. ISSN: 1446-8735.

3. S. Ahmad, M. Abdollahian, P. Zeephongsekul, B. Abbasi, "Performance Analysis of Skewed Data", Ubiquitous Computing and Communication Journal, UBICC Journal, Volume 3, January 2008, pp.8-12:ISSN:1992-8424.

4. B. Abbasi, S. Ahmad, M. Abdollahian and P. Zeephongsekul, Measuring Process Capability for Bivariate Non-Normal Process Using the Bivariate Burr Distribution, WSEAS Transaction on Business and Economics, Issue 5, Volume 4, 2007, ISSN: 1109-9526.

5. S.Z. Hosseinifard, B. Abbasi, S. Ahmad, M. Abdollahian, A Transformation Technique to Estimate Process Capability Index for

1.7.2. Refereed conference papers

6. Ahmad S., Huda S. Bakir S., Abdollahian M., Zeephongsekul P. "Constraint-Based Evolutionary Learning Approach to the Process Performance Evaluation", accepted for publication, 3rd International Conference on Informatics and Technology, Kuala Lumpur, Malaysia, 27th - 28th October 2009.

7. Ahmad S., Huda S. Bakir S., Abdollahian M., Zeephongsekul P. "Process Performance Evaluation Using Evolutionary Algorithm, accepted for publication, WORLDCOMP'09 - The 2009 World Congress in Computer Science, Computer Engineering, and Applied Computing, Las Vegas, USA, July 3- 16 2009.

8. Nazari, A., Ahmad, S., Abdollahian, M. and Zeephongsekul, P. "A Model to Estimate Proportion of Non-conformance for Multi-Characteristics Product" 6th International Management Conference, Dec 20-22, Tehran Iran, 2008, pp. 3-14.

9. Ahmad, S., Abdollahian, M. and Zeephongsekul, P. "Evaluating Process Capability by Fitting Burr Distribution to Multivariate Data" Proceedings of the 14th ISSAT International Conference on Reliability and Quality in Design, Seattle, Washington, USA, 2008, pp. 49 – 53.

10. Ahmad, S., Abdollahian, M. and Zeephongsekul, P. "Fitting Burr XII distribution to continuous positive data using Hybrid Search

11. Ahmad, S., Abdollahian, M. and   Zeephongsekul, P. "Process Capability for a Non-Normal Quality Characteristics Data", 4ᵗʰ International Conference on Information Technology- New Generations (ITNG), April 2-4, Las Vegas, Nevada, U.S.A, 2007, pp. 420-424.

12. Ahmad, S., Abdollahian, M. and   Zeephongsekul, P.   "Process Capability Analysis for Non-Normal Quality Characteristics Using Gamma Distribution", 4ᵗʰ International Conference on Information Technology- New Generations (ITNG), April 2-4, Las Vegas, Nevada, U.S.A, 2007, pp. 425- 430.

13. Ahmad, S., Abdollahian, M. and   Zeephongsekul, P. "Process Capability Analysis under a Weibull Shock Model" 13ᵗʰ ISSAT International Conference on Reliability and Quality in Design, August 2-4,Seattle, Washington, U.S.A, 2007, pp. 88-92.

14. Ahmad, S., Abdollahian, M. and   Zeephongsekul, P. "Process Capability Analysis for a Skewed Population Data", The 2007 World Congress in Computer Science, Computer Engineering, & Applied Computing, June 25-28, Las Vegas, Nevada, U.S.A, 2007, pp. 327 - 333.

15. Ahmad, S., Abdollahian, M. and   Zeephongsekul, P. "Non-Normal Process Capability Evaluation", IKE'07- The 2007 International

1.8. Summary

This chapter laid the foundations for the thesis. It introduced the research problem and research questions. The methodology was briefly described and justified, the organization of the thesis was outlined and the list of publications based on this research during this tenure has been provided.

In the proceeding Chapter 2, we will discuss PCIs when the quality characteristics data is not normal and traditional capability indices C_p and C_{pk} are unable to provide accurate estimates of process performances.

Chapter 2

PROCESS CAPABILITY FOR NON-NORMAL

QUALITY CHARACTERISTICS

2.1. Introduction

Process capability indices are random variables and it is well known that random variables are always associated with a probability distribution. This distribution also provides a description of the expected value and variance of the index (Kotz & Lovelace 1998).Keeping in view this notion, it is important to understand the theory of probability and statistical methods in order to fully comprehend the nature and behavior of the traditional process capability indices presented in the research literature.

After presenting basic theoretical background of traditional capability indices C_p and C_{pk} in Chapter 1, this chapter will focus on those methods and techniques commonly used to estimate process capability when quality characteristics are not normal.

2.2. Univariate process capability for non-normal data

The basis of traditional capability indices C_p and C_{pk} heavily depends on an implicit assumption that the underlying quality characteristic measurements

are independent and normally distributed. However, this basic assumption is not usually fulfilled in practice. Many physical processes produce non-normal quality characteristics data and quality practitioners need to verify the above basic assumptions before deploying any conventional PCI techniques to determine the capability of their processes.

Kane (Kane 1986) has drawn the attention to such problems that may occur with non-normal data and (Gunter 1989), in Parts 2 and 3, highlighted this even more. To overcome these problems several approaches have been suggested. Here in section 2.3, we will discuss two common approaches, namely techniques of non-normal quantile estimation and transformations. Furthermore we will consider some more recent developments to handle the issue of non-normality by fitting known non-normal distributions to non-normal quality characteristics data in Chapter 3. For a thorough discussion of different methods to handle a non-normally distributed process outcome see, e.g. (Kotz & Johnson 1993), (Kotz & Lovelace 1998) and (Kotz & Johnson 2002).

One of the first indices for data that are non-normally distributed was suggested by Clements (Clements 1989). He used the technique of non-normal quantile estimation and proposed that 6σ and μ in C_p and C_{pk} be replaced with $q_{0.99875} - q_{0.000135}$ and $q_{0.5}$, respectively, where q_α for the specified α values, represents the quantiles for a distribution in the Pearson family. If the distribution of the quality characteristic is normally distributed then $q_{0.99875} - q_{0.000135} = 6\sigma$. We will discuss this method in some detail in section 2.3.1. (Pearn & Kotz 1994) has extended Clements' method to

develop some new indices which can be deployed for non-normal PCI estimations.

Another approach for dealing with non-normal data is to transform the original non-normal data to normal or at least close to normal (section 2.3.2 of this thesis will provide some details of data transformation techniques). Gunter (Gunter 1989) suggested application of data transformation approach to perform calculations of capability indices when the process data is non-normal. Calculation of C_{pk} for non-normal data was also discussed in (Rivera, Hubele & Lawrence 1995). Furthermore, (Polansky, Chou & Mason 1998, 1999) proposed a method for assessing the capability of a process using data from a truncated normal distribution, where Johnson transformations ((Johnson, NL. 1949)) were used to transform the non-normal process data into normal. However, one can not be sure that the capability of the transformed distribution will reflect the capability of the true distribution in a correct way, see, e.g. (Gunter 1989). Furthermore, Kotz & Lovelace (Kotz & Lovelace 1998) point out that practitioner may be uncomfortable working with transformed data due to the difficulties in translating the results of calculations back to the original scale. Another disadvantage from a practitioner's point of view is that transformations do not relate clearly enough to the original specifications according to Kotz & Johnson (Kotz & Johnson 2002).

In addition to percentile and transformation approaches, some other methods have also been proposed in the research literature. For the case with skew distributions and two-sided specification limits, (Wu et al. 1999)

introduced a new process capability index based on a weighted variance method. The main idea of this method is to divide a skewed distribution into two normal distributions from its mean to create two new distributions which have the same mean but different standard deviations. Chang et al. (Chang, Choi & Bai 2002) proposed a different method of constructing simple process capability indices for skewed populations, based on a weighted standard deviation method. Some properties for the proposed indices are also investigated by Wu et al. (Wu et al. 1999) and Chang et al. (Chang, Choi & Bai 2002) and the estimators are compared to other methods for non-normal data.

Several authors have made comparative studies between different methods to handle non-normal process data. Heuvel & Ion (Heuvel & Ion 2003) compared indices for skew distributions proposed by Munchechika (Munchechika 1986) and Bai & Choi (Bai & Choi 1997), for a number of distributions corresponding to C_{pk} . One conclusion from their study is that for many practical situations the true value of C_{pk} lies between the values of the indices presented by Munchechika and Bai & Choi. Using Monte Carlo simulations Wu et al. (Wu & Swain 2001) compared traditional indices for Clements' method, the Johnson-Kotz-Pearn method (Johnson, NL, Kotz & Pearn 1994) and the weighted variance method (Wu et al. 1999)for the Johnson family of distributions. They found that for skewed bounded cases none of these three methods performs well in estimating the nominal value. Furthermore, Clements' method was misleading for skewed unbounded cases. For log-normal cases, the weighted variance method underestimates

the nominal values while the Johnson-Kotz-Pearn method consistently overestimates the nominal values. Clements' method neither overestimates nor underestimates the results on a consistent basis.

2.2.1. Process capability for one-sided specification limit when data is non-normal

Process capability indices for one-sided specification limit and a non-normally distributed characteristic have not been discussed much in the literature. This is not an uncommon situation in industry, however, it should be noted that Clements (Clements 1989) treated the indices for one-sided specification limits similar to C_p, i.e. he replaced μ with median and denominator by the lengths of interval between the upper and lower 0.135 percentage points of the distribution of X (refer to equations 1.3 and 1.4). Sakar et al. (Sarkar & Pal 1997) considered an extreme value distribution for the C_{pu} case. Furthermore, Tang et al. (Tang & Than 1999) studied estimators of C_{pu} for a number of methods designed to handle non-normal process data, in particular when the underlying distribution are Weibull and Lognormal. This was done by Monte Carlo simulations. They found that methods involving transformations provide estimates of C_{pu} that is closer to the nominal value compared to non-transformation methods, e.g. the weighted variance method discussed by Choi & Bai (Choi & Bai 1996). However, even though a method performs well for a particular distribution, that method can give erroneous results for another distribution with different tail behavior. In fact, the effect of the tail area can be quite

dramatic (Vännman & Albing 2007). Ding (Ding 2004) introduced a process capability index based on the effective range by using the first four moments of non-normal process data. He also considered the situation with univariate positively skewed data and proposed an index for this situation. However, the proposed index contains no target value and furthermore, as far as we know, no decision procedures or tests have been presented.

2.3. Methods to estimate process capability for non-normal quality characteristics data

This section provides information about some methods commonly used in industry for non-normal process capability estimations. Clements percentile method, Box Cox power transformation method and root transformation method are among them. The general description of these methods and some historical review are presented here.

2.3.1.   Clements percentile method

Pearson (Pearson 1895) identified four types of distributions which include a rich class of populations with non-normal characteristics. Clements (Clements 1989) proposed a method for calculating process capability indices C_p and C_{pk} based on non-normal pearsonian distributions. This method provides an easy approach to handle the issue of non-normality for process capability estimations and uses non-normal percentiles to modify the traditional capability indices. The main advantage of this approach is that it requires no complicated distribution fitting and is simple to use by non-statisticians. Clements method is a popular method to use among

quality practitioners in industry today. This method used Pearson curves to provide more accurate estimates of percentile points $X_{0.00135}$, $X_{0.5}$ and $X_{0.99865}$ when the underlying process data follow non-normal distribution (Kotz & Lovelace 1998). The underlying concept of this method is again based on normal distribution. As mentioned in chapter 1, in the traditional capability indices, we are critically interested in three points within the process distribution i.e. the upper tail, the point of central tendency and the lower tail. In terms of quantiles, these points for the normal distribution correspond, respectively, to

- Upper Tail = $X_{0.99865} = \mu + 3\sigma$,

- Lower tail = $X_{0.00135} = \mu - 3\sigma$ and

- Mean = Central tendency = $X_{0.50} = \mu$.

When the data is normal, it is quite easy to estimate these three points. However, in case of non-normal data, it is not easy to estimate these quantiles when we don't know the distribution of the underlying variable. More importantly, these quantiles do not necessarily corresponds to $\mu + 3\sigma$, μ , $\mu - 3\sigma$ respectively. For instance, quantile $X_{0.50}$ corresponds to the mean μ for the normal case, but in the non-normal case, it corresponds to the median.

Clements used the same approach to estimate non-normal quantiles and replaced 6σ in equation (1.1) by the lengths of interval between the upper and lower 0.135 percentage points of the distribution of X . Therefore, the denominator in equation (1.1) can be replaced by $(U_p - L_p)$, i.e.

$$C_p = \frac{Usl - Lsl}{U_p - L_p} \qquad (2.1)$$

where U_p is the upper percentile i.e. 99.865 percentile of observations and L_p is the lower percentile i.e. 0.135 percentile of observations. Since the median "M" is the preferred central value for a skewed distribution, so Clements estimated C_{pu} and C_{pl} as follows:

$$C_{pu} = \frac{Usl - M}{U_p - M} \qquad (2.2)$$

$$C_{pl} = \frac{M - Lsl}{M - L_p} \qquad (2.3)$$

and

$$C_{pk} = \min\left[C_{pu}, C_{pl}\right] \qquad (2.4)$$

As mentioned earlier, Clements' approach does not require a mathematical transformation of the data. It is easy for non-statisticians to understand and no complicated distribution fitting are required (see (Kotz & Lovelace 1998)). However, Clements' method requires knowledge of the skewness and kurtosis that are based on 3rd and 4th moments respectively, and these may not be reliably estimated for very small sample sizes (Liu, P & Chen 2006). Wu et al. (Wu, Wang & Liu 1998) have conducted a research study indicating that the Clements method cannot accurately measure the capability indices, especially when the underlying data distribution is skewed.

2.3.2. Data transformation method

Data transformation refers to the application of a known deterministic mathematical function to each point in a quality characteristics data i.e. each data point X_i is replaced with the transformed value $Y_i = f(X_i)$, where the function $f(.)$ is an appropriate mathematical function. The main objective of data transform technique is to transform the non-normal data to normally distributed data so that it can closely meet the assumptions of a statistical inference procedure that need to be applied to improve the interpretability of the quality data.

Data transformation techniques are straightforward and easy to deploy and are popular among quality practitioners in industry. Johnson (Johnson, NL. 1949) proposed a system of distributions based on the moment method to transform the non-normal data to normally distributed data. It is called the Johnson transformation system. Box and Cox (Box & Cox 1964) presented a useful family of power transformation. Somerville et al. (Somerville & Montgomery 1996) proposed a square root transformation to transform the non-normal data to normal data. In recent years Niaki et al. (Niaki & Abbasi 2007) also presented root transformation to handle the issue of non-normal quality data and deployed this transformation technique to design multi attribute control charts.

In this research we will use the power transformation and root transformation techniques to estimate non-normal PCI and we will discuss

these two techniques in some detail in the proceeding sections of this chapter.

2.3.2.1 Box-Cox power transformation method

Power transformation is a family of transformations that map non-normal quality characteristics data from one space to another using power functions. This is a useful data transformation technique employed to reduce data variation and make the data normally distributed. The Box-Cox power transformation is the most commonly used technique in industry. This technique was proposed by George E. P. Box and David R. Cox in 1964 (Box & Cox 1964). The Box-Cox power transformation on necessarily positive response variable X is expressed by

$$X(\lambda) = \begin{cases} \dfrac{(X^{\lambda} - 1)}{\lambda} & \text{for} \quad \lambda \neq 0 \\ \text{Ln}(X) & \text{for} \quad \lambda = 0 \end{cases} \qquad (2.5)$$

$$\text{where} \quad -5 \leq \lambda \leq +5$$

This transformation depends upon a single parameter λ that can be estimated by Maximum Likelihood Estimation (MLE) method (Tang, 1999 #20). λ Can be chosen from the given range and for each chosen λ evaluate:

$$L_{max}(\lambda) = -\frac{1}{2}\ln(\hat{\sigma}^2) + (\lambda - 1)\sum_{i=1}^{n}\ln(X_i) \qquad (2.6)$$

31

The estimate of $\hat{\sigma}^2$ for fixed λ is obtained by

$$\hat{\sigma}^2 = \frac{S(\lambda)}{n} \qquad\qquad (2.7)$$

where $S(\lambda)$ is the residual sum of square in the analysis of variance of X^λ. After calculating $L_{max}(\lambda)$ for several values of λ within the given range one can plot $L_{max}(\lambda)$ against λ. The maximum likelihood estimator of λ is the value of λ that maximizes $L_{max}(\lambda)$. Using the optimal $\overset{*}{\lambda}$ value, data values for each individual X data are transformed to a normal variate using equation (2.5) (for details refer to (Box & Cox 1964)). Box-Cox transformation can be applied to non-zero, positively skewed data. The transformation method is available in most statistical software packages as a standard feature. Consequently, the users can deploy this technique directly and with ease to evaluate process capability indices for non-normal data first transforming these data to normal data using Box-Cox transformation.

2.3.2.2 Root transformation method

This is another data transformation technique to handle the issue of non-normal data. In this section we will briefly discuss two data transformation techniques. The first one, called Square Root Transformation, was proposed by Somerville et al. (Somerville & Montgomery 1996) and the second approach, called Best Root Transformation, was proposed by Niaki et al. (Niaki & Abbasi 2007) to solve the problem of non-normality in the data.

Somerville and Montgomery (Somerville & Montgomery 1996) proposed the square root transformation. According to a detailed survey of non-normal distributions, they have presented in their research that some of the non-normal distributions can be transformed to normal distribution by means of an appropriate transformation function; for example, a skewed distribution may respond well to the square root transformation. The square root of each data point is taken and then the transformed data is evaluated for normality test. If the normality test shows that the transformed data is normally distributed, then one can apply statistical procedures to the transformed data in order to obtain useful information such as capability index values. The advantage of this method is that it is easy to understand by non-statisticians and simple to deploy. However, due to it being computationally intensive it is hard to implement this method as a standard method in industry.

As we know the most serious problem with the non-normal data is its existing skewness. Niaki & Abbasi (Niaki & Abbasi 2007) proposed an approach that can transform the existing skewness of data to zero or close to zero to make it normal. They called it Best Root Transformation technique and used it to design multi attribute control charts. This approach enables user to transform non-normal data to normal by searching for the optimal root for data transformation. According to this approach, one can search for the best root of the non-normal data $(\mathbb{X})$. The obtained best root $(\mathbb{r})$ of the non-normal data $(\mathbb{X})$ is the root that if we raise the power of the data $(\mathbb{X})$ to

that root (r) (i.e. $\mathbb{X}^{\mathbb{r}}$), the skewness of the transformed data will become

zero. Bisection method is used to find the (r) value. This method is based on the fact that a function will change sign when it passes through zero. By evaluating the function at the middle of an interval and replacing whichever limit has the same sign, the bisection method can halve the size of the interval in each iteration and eventually find the optimal root. An application example using this method will be presented in Chapter 4. For more details refer to Niaki et al. (Niaki & Abbasi 2007).

2.4.   Multivariate non-normal process capability

It is an established fact that production processes not only produce non-normal quality characteristics data but also there is always more than one quality characteristics of interest in process outcomes, and very often, these quality characteristics are correlated with each other. The traditional capability indices C_p and C_{pk} consider only one quality characteristic at a time. However quality of many products is determined by more than one characteristic. Moreover, these quality characteristics that jointly determine the quality are often jointly inter-related (Wierda 1993). This situation increases the complexity of the problem of trying to produce a meaningful measure of capability indices for multivariate data. For example, in a detailed description of a connecting rod for a combustion engine (Taam, Subbaiah & Liddy 1993), the crank bore inner diameter, pin bore inner diameter, rod length, bore true location, bore-to-bore parallelism, and other features are specified. To represent how well this connecting rod is made, one may examine numerical summaries of each of these individual characteristics separately or consider all characteristics together to see how

they interact with each other. The latter is preferred if one treats the rod as one entity. In situations where the design intention of a product is prescribed by a number of related characteristics, the functionality of this product cannot be represented by individual quality characteristic separately. Many other such examples are scattered throughout the quality control literature, which points towards the need of developing accurate measures of process capability that can address the complex nature of multivariate non-normal quality characteristics data.

Multivariate capability indices usually produce one number jointly representing capability for two or more quality characteristics. Generally multivariate process capability indices can be obtained from a number of different methods such as:

- the ratio of a specification limit to process variation or modified process variation.

- the probability of nonconforming products over rectangular tolerance zone implementing loss functions and vector representation.

- theoretical proportion of non-conforming products over convex polygons and

- global approach of viewing multivariate quality control.

Taam et al. (Taam, Subbaiah & Liddy 1993) defined the first multivariate capability index based on the ratio of a specification limit to process variation or modified process variation. Chen (Chen, H 1994) also proposed

a method in order to estimate the multivariate C_p using a non-conforming proportion approach. Shahriari et al. (Shahriari, Hubele & Lawrence 1995) proposed a process capability multivariate vector in order to evaluate the process performance. Braun (Braun 2001) defined C_p and C_{pk} as EC_p and EC_{pk}, where both the multivariate process region and the multivariate tolerance region are of elliptical shape. Castagliola et al. (Castagliola, P & Castellanos 2005) defined two new capability indices BC_p and BC_{pk} dedicated to two quality characteristics, based on the computation of the theoretical proportion of non-conforming products over convex polygons. Bothe (1999) proposed a method in order to compute the multivariate C_{pk} index. Wang et al. (Wang & Du 2000) proposed multivariate equivalents for C_p and C_{pk} based on the PCA (Principal Component Analysis) decomposition. Other researchers who worked in the Multivariate PCIs are (Beck & Ester 1998), (Bernardo & Irony 1996), (Boyles 1994), (Davis, Kaminsky & Saboo 1992), (Wierda 1993), (Hellmich & Wolff 1996), (Li & Lin 1996), (Mukherjee & Singh 1994), (Yeh & Bhattacharya 1998), (Veevers 1998) and (Niverthi & Dey 2000). But the main limitation with the existing multivariate PCIs in the research literature is that they are all based on multivariate normal quality characteristics data. Wang and Du in 2000 proposed the same multivariate C_p and C_{pk} indices and extended their research work to the non-normal multivariate case. Wang (Wang & Hubele 1999) and(Wang 2006) proposed a "Geometric Distance Variable" approach to reduce the dimensionality of the multivariate (normal and non-normal) data to univariate data and use the established univariate PCI techniques for

process performance analysis. A detailed review of their Geometric Distance approach along with our newly proposed Generalized Covariance Distance (GCD) approach (Ahmad et al. 2009) will be presented in Chapter 6

2.5. Summary

The main objective of this chapter is to review the basic theory of non-normal PCIs and present a literature review of PCI estimations when the quality characteristics data don't follow normal distribution. Commonly used techniques have been discussed. In the proceeding Chapter 3, we will provide details of the numerical techniques used in this research study.

Chapter 3

METAHEURISTIC APPROACHES AND DISTRIBUTION FITTING TO NON-NORMAL QUALITY CHARACTERISTICS DATA

3.1. Introduction

In contrast to non-normal quantile estimation and data transformations discussed earlier in Chapter 2, this chapter describes another simple approach to solve the problem of non-normal PCI estimations. This approach deals with fitting known distribution to non-normal quality characteristics data and use proportions of non-conformance criterion to assess process performances. The chapter also describes metaheuristics approaches employed to estimate the parameters of the fitted distribution.

This thesis discusses metaheuristic approaches such as Simulated Annealing, Compass Direct search and their hybrids with local search approaches for solving the local convergence problems. Constraint-based approaches to the population based metaheuristic (Evolutionary Algorithm, EA) will also be discussed. This chapter

provides background knowledge of these heuristics and metaheuristics approaches which will be used in this thesis.

3.2. Fitting known distributions to non-normal quality characteristics data

Although percentile estimation and data transformation techniques are commonly used to solve the non-normality problems of quality characteristics data, there is another simple approach to handle the issue of non-normal quality characteristics data. According to this technique, a generic known distribution (e.g. Gamma, Weibull, Beta and Lognormal) is fitted to the actual sample quality characteristics data (Somerville & Montgomery 1996). Then, process capability indices can be simply evaluated using the percentage falling outside the specification limits of the fitted distribution. Hahn & Shapiro (Hahn & Shapiro 1967) suggested using the calculated sample skewness and kurtosis and plotting them on a distribution "mapping" which may suggest the proper distribution to use for the fit.

Fitting known distributions to quality characteristics data can easily be done by using available statistical software packages. The analyst can easily determine which distribution can fit best to the sample data. Consequently, the analyst can deploy this technique directly to evaluate process capability indices. Fitting a known distribution to

sample data has several advantages, e.g. it is a straightforward approach and the analyst deal with the quality sample data directly; instead of making adjustments and approximations. The process analyst may gain insight into the underlying process when fitting different distributions. This insight could lead the analyst to implement real process improvement endeavors. The shortcoming to this approach is that a relatively large sample must be obtained so that the distribution fit could be accomplished with some degree of confidence (Somerville & Montgomery 1996).

In the proceeding sections of this chapter we will discuss some metaheuristic numerical techniques used to fit Burr XII distribution in this research study.

3.3. Parameter estimation techniques

In dealing with the application of statistical theory to industrial problems, the analyst should take care with regard to sample selection and other experimental details and follow strict guidelines. Otherwise, results based on inferential techniques which include the method of parameter estimation can lead to erroneous estimates of process capability. The need for accurate parametric estimation has become increasingly important, as indicated by its wide application and theoretical literature which have appeared on this subject

(Gruska, Mirkhani & Lamberson 1989).There are several techniques used for parametric estimation which include Method of Moments and Maximum Likelihood Method. In the next section we will discuss parametric estimation methods used in this research study.

3.3.1. Method of moments

The method of moments equates sample moments to population moments. It has the advantage of simplicity; however, the disadvantage is that they are often not as accurate as other parametric estimation techniques such as Maximum Likelihood Method and Least Squares Method. In this research study we have applied this method to estimate the Burr distribution parameters (i.e. c and k) using Burr tables (Burr, IW 1973). An illustrative example is provided in chapter 4.

3.3.2. Maximum likelihood method

Maximum Likelihood estimation uses the mathematical expression known as a likelihood function of the sample data to estimate parameters. In more general form, we can say the likelihood of a set of data is the probability of obtaining that particular set of data given the chosen probability model. This expression contains the unknown parameters. Those values of the parameter that maximize the sample likelihood are known as the maximum likelihood estimates.

There are several advantages using this method, in particular this method provides a consistent approach to parameter estimation and can be applied widely to industrial problems such as reliability analysis of censored data under various censoring models. Being an established procedure, several popular statistical software packages provide excellent algorithms for maximum likelihood estimates for many of the commonly used distributions. This helps to mitigate the computational complexity of maximum likelihood estimation, although it does not totally eliminate the problems of finding the correct solutions.

However, the likelihood equations need to be specifically worked out for a given distribution and estimation problem. The mathematics is often non-trivial, particularly if confidence intervals for the parameters are desired. Except for a few cases where the maximum likelihood formulas are in fact simple and the numerical estimation is usually non-trivial, it is therefore generally best to rely on high quality statistical software to obtain maximum likelihood estimates. Fortunately, high quality maximum likelihood software is becoming increasingly common. Another shortcoming to this method is that it can be heavily biased for small samples. The optimality properties may not apply for small samples. Further parametric estimates obtained using this method can be sensitive to the choice of starting

values. In the proceeding section we will describe background information and historical review of some of the numerical (heuristic and metaheuristics) techniques used in this thesis for estimation of the parameters of the Burr distribution.

3.4. Heuristic numerical techniques for parameter estimation

Heuristic techniques help to solve many optimization problems. A heuristic is a rule of thumb which is used to solve a complex problem where no exact method is present for solving the problem. Usually a heuristic approach uses some knowledge about the domain of the problem under consideration and its structure to devise a technique for solving the problem. A typical heuristic is a best-first search which is used to search a decision tree and other tree-like data structure (Baum & Sell 1968). Another example of a heuristic is Kruskal's polynomial time algorithm for finding a minimum spanning tree (Pearl 1984), (Kruskal 1956). Heuristic techniques can be broadly classified into two main groups: constructive and improvement heuristics.

Constructive heuristics find a solution from the scratch incrementally. Usually a constructive heuristic starts from an empty solution and successively augments the solution component at each step and finds a final solution. It relies on knowledge of the problem to allow

the development of the solution. Constructive heuristics are highly problem dependent. Therefore, it is often true that if a problem provides sufficient knowledge about its domain and structure then good constructive heuristics can be developed to solve the problem. Some successful constructive heuristics are the greedy heuristic and the nearest neighbor heuristic (Rosenkrantz, Stearns & Lewis 1977), (Higgins, Kozan & Ferreira 1997).

Improvement heuristics take a feasible initial solution as input and try to find a better solution by searching through the neighbors of the current solution. The next solution of the search step is computed by finding the neighbors of the current solution and choosing the best from these neighbors. The initial solution is changed over a number of iterative steps so that the solution quality is gradually improved. The set of all possible changes that can be applied to a particular solution is referred to as the neighborhood of the solution. These search approaches are referred to as neighborhood search or local search heuristic ((Rosenkrantz, Stearns & Lewis 1977), (Johnson, D 1990), (Michalewicz & Fogel 2000), (Batti & Protasi 2001)), and for a detailed discussion refer to Huda (Huda, S 2009) and Huda (Huda, S , Yearwood & Togneri 2009). An iterative improvement heuristic is more general than a constructive heuristic and has wider application. One of the important characteristics of a neighborhood search heuristic (improvement

heuristic) is that its deterministic selection criteria to choose the next solution may result locally optimized solution and not give any global solution. However, by applying different selection criteria (stochastic selection criteria (Osman & Kelley 1996)), threshold selection criteria (Jiang & Yang 2002)) or applying intelligent control strategy (Ganesh & Punniyamoorthy 2005) on the search process, neighborhood search may be capable of overcoming the local optimization problem. This latter approach of heuristics is generally called Metaheuristic.

Metaheuristic approaches are techniques that can be generally applied to solve an optimization problem like a black-box optimization algorithm (Lin & Keringhan 1973). According to (Osman & Kelley 1996), *Metaheuristic is an iterative generation process which guides a subordinate heuristic* by combining intelligently different concepts for exploring and exploiting the search space.

In many real world problems, we do not have any strong insight into how a problem might be solved. Sometimes we could find a *constructive heuristic* for a problem which is too complex to implement. In these cases, it is best to use more general *heuristics* which can be referred to as *metaheuristics*. Metaheuristics are good candidates for solving many optimization problems. Metaheuristic approaches can be generally divided into two main categories:

- Neighborhood search based approaches

- Population based approaches

Neighborhood search based approaches are Simulated Annealing (Osman & Kelley 1996), Tabu Search (Ganesh & Punniyamoorthy 2005), and Threshold Accepting (Jiang & Yang 2002) which are known as single candidate model based approach. Population based approaches involve Evolutionary Algorithm (EA) (Dueck & Scheuer 1990), Ant Colony Systems (Applegate, Cook & Rohe 2003), Particle Swarm systems ((Stutzle 1998), (Kennedy, Eberhart & Shi 2001), (Liu, B et al. 2005)).

3.4.1.   Simulated Annealing (SA) approach

Simulated Annealing (SA) (Osman & Kelley 1996) is one of the important *naturally motivated metaheuristics* that combines a naturally motivated acceptance criterion with the general structure of an improvement heuristic. SA is used on a variety of large optimization problems and requires little problem-specific knowledge other than a fitness or energy information. The basic idea of SA comes from the physical annealing process. In a metallurgical annealing process, a metal body is heated to near its melting point and then slowly cooled back down to room temperature. At very high temperatures atoms of metal obtain very high energy. If the temperature of metal is slowly decreased, then atoms reach an

absolute minimum energy. If the temperature is decreased too quickly the metal ends up in a poly-crystalline or amorphous state which is not pure crystal with a higher energy than the minimum energy of the metal. The behavior of the metal with temperature and structure of the atoms inside the metal (which is called the state of the metal) can be explained by *statistical mechanics*.

Let the state *(ss)* of a metal be identified with the set of spatial position of the atoms. If the metal is in thermal equilibrium condition at temperature *(T')*, then the probability $Pr_{T'}$ *(ss)* that the metal is in a given state *(ss)* depends on the energy E *(ss)* and follows the Boltzmann distribution given in equation (3.1).

$$\mathrm{Pr}_{T'}(ss) = \frac{\exp\left[\dfrac{-E(ss)}{k_b T'}\right]}{\displaystyle\sum_{\omega \in SS}\left[\dfrac{-E(\omega')}{k_b T'}\right]} \tag{3.1}$$

where k_b is the Boltzmann constant and (SS) is the set of all states of the metal. Let us consider that at time *(t')* the metal is in a state *(q')*. A candidate state *(r')* at time *(t'+1)* can be generated randomly and accepted with the probability $(P_{T'})$ given by equation 3.2.

$$P_{T'} = \frac{Pr_{T'}(r')}{Pr_{T'}(q')} = \exp\left[\frac{E(r') - E(q')}{k_b \, T'}\right] \qquad (3.2)$$

If $(P_{T'} > 1)$, the energy of state (r') is strictly greater than energy of state (q'). It has been proved that as time (t') increases to infinity, the probability that the metal is in a given state S_∞ equals $P_{T'}(S_\infty)$ and converges to Boltzmann distribution (Osman & Kelley 1996). $Pr_{T'}(q')$ and $Pr_{T'}(r')$ can be determined using equation 3.1. However, it is not the case that lower temperature gives the lower energy state. We must adapt an annealing process where the temperature of the metal is raised to very high temperature at the beginning and then slowly decreased, spending sufficient time at each temperature to reach thermal equilibrium. This physical annealing phenomenon is used as a computational technique for optimization problem to avoid local optimum problem.

Before applying the metallurgical annealing techniques in an optimization problem we must find the analogous of the physics concept of annealing. Here, the energy function of metal, *E (ss),* corresponds to the objective function *F(x)* in an optimization problem. States of the metal *"ss"* becomes the values for parameters *(x)* of the optimization problem. We must also find a function

(neighborhood generation operator) to generate the neighbor/new solution from current solution and a cooling schedule as well.

Unlike other search algorithms such as Tabu Search which may generate a local optimum of objective function, the chance of being trapped at local maximum of the objective function using SA is avoided by the use of a method similar to physical annealing technique.

Like other neighborhood search, SA starts with an initial solution (x_{nn}). At each search step "nn", a new solution (x_{nn+1}) is generated by using neighborhood generation operator from the current solution (x_{nn}). SA accepts the new solution (x_{nn+1}) if $F(x_{nn+1}) > F(x_{nn})$. However this deterministic method may terminate at local maximum of $F(x)$. Applying an annealing process similar to physical annealing, SA allows the search process to change its state to a state with lower objective function value so that it gets a chance to jump out of the local maxima and seek better maximum from that point again. Here the lower value of objective function is accepted with a probability given by equation (3.3)

$$P_{T'}(x_{nn+1}) = \exp\left[\frac{F(x_{nn+1}) - F(x_{nn})}{k_b T'}\right] \qquad (3.3)$$

Uphill moves in SA are always accepted. However, downhill moves are accepted with an acceptance probability which is function of temperature given by equation (3.3). The performance of the SA is dependent on the cooling schedule (Osman & Kelley 1996) and (Huda, S 2009).

One of the important cooling schedules is the Lundy schedule (Liaw 2000) which is described below:

In Lundy schedule (Lundy & Teng 1986),two temperature values, T'_{nn} and T'_{nn+1}, which are in nn^{th} and $(nn+1)^{th}$ iterations are related by the following formula:

$$T'_{nn+1} = \frac{T'_{nn}}{1 + \beta\, T'_{nn}} \qquad (3.4)$$

where β is defined by equation (3.5)

$$\beta = \frac{T'_i - T'_f}{i_t\, T'_i\, T'_f} \qquad (3.5)$$

Here i_t = total number of iteration, T'_f = the final temperature, T'_i = the initial temperature. β is greater than 0.

SA starts with a high initial temperature T_i' and any random initial solution (x_{nn}). A neighborhood operator is applied to the current solution (x_{nn}) having objective function values $F(x_{nn})$ to produce a new solution (x_{nn+1}) having objective function values $F(x_{nn+1})$. The new solution (x_{nn+1}) is accepted if $F(x_{nn+1}) > F(x_{nn})$ and becomes the current solution, otherwise (x_{nn+1}) becomes the current solution with a probability $P_{T'}(x_{nn+1})$ from equation (3.3). If (x_{nn+1}) is not accepted, then (x_{nn}) remains as the current solution. The application of the neighborhood generation operator and the probabilistic acceptance of the newly generated solution are repeated either a fixed number of iterations or until a quasi-equilibrium is reached. The whole process is repeated each time starting from the current solution with a lower temperature. For any given temperature T', a sufficient number of iterations always lead to equilibrium. The cooling schedule is such that at high temperature any change is accepted. This means the SA visits a very large neighborhood of current solution. At lower temperatures, transition to lower values of the objective function becomes less frequent and the solution stabilizes. The complete algorithm for SA is described in the Table 3.1.

Table 3.1: Algorithm for Simulated Annealing (SA)

begin

Choose an initial solution x

Initialize temperature $T' \leftarrow T'_i$

trial $\leftarrow 1$

repeat

   for (trials $= 1$ to TOTAL TRIALS) do

    Generate a neighboring solution $x' \in V(x)$

    if $F(x') > F(x)$ then

     $x \leftarrow x'$

     $rn \in \{0,1\}$,   else if $\left(\min\left[1, \exp\left[\dfrac{F(x') - F(x)}{k_b\, T'}\right]\right] \geq rn\right)$ then

      $x \leftarrow x'$,   else

       Current solution x is unchanged

      end if

   end for

  $T' \leftarrow T'/(1 + \beta\, T')$

 until (termination reached)

RETURN x

end

SA is used in many optimization algorithms. The connection between this method and mathematical minimization was first observed by (Pincus 1970), but it was (Kirkpatrick, Gerlatt & Vecchi 1983) who proposed it as an optimization technique for combinatorial and other optimization problems. Ease of use and provision of good solutions to real-world problems makes this method one of the most powerful and popular meta-heuristics to solve many optimization problems (Niaki & Abbasi 2007).

3.4.2.   Compass Direct Search method

Direct search methods form a class of optimization methods that don't use any exact or approximate information of derivatives. The direct search methods were first used by Hooke and Jeeves (Hooke & Jeeves 1961) and in the Simplex algorithm of Nelder and Mead (Nelder & Mead 1965). At that time, the methods were considered heuristic without any mathematical convergence proof. Here we will consider one general frame of direct search methods called Compass Direct Search. Most of the direct search methods could be described as special cases of this method.

For a function with two variables the method can be summarized as follow: Try steps to the east, West, North and South. If one of these trials yields a better point in the function, the improved point is

taken as the new iterate. If none of the trials makes any improvement, try again with step length which is half the original step. The mathematical description of this method for general n-dimension is given below.

Let $x_k \in R^n$ denotes the K_{th} iteration, where x_0 is the chosen initial value. Also let D denotes the set of $2n$ coordinate directions which is positive and negative unit coordinate vectors,

$$D = \{e_1, e_2, \ldots, e_n, -e_1, -e_2, \ldots, -e_n\}.$$

Let Δ_k denotes the step length control parameter with the starting value of Δ_0.

Initialization:

Let $f : R^n \rightarrow R$ be given

Let $x_0 \in R^n$ be the initial value.

Let $\Delta_{tol} > 0$ be the tolerance

Let $\Delta_0 > \Delta_{tol}$ be the initial value of the step length.

Let D be the set of coordinate directions $\{\pm e_i\}$

Table 3.2 describes the generic algorithm of Direct Search Method.

Table 3.2: Direct search method algorithm

For each iteration $k = 1,2,...$

1. If there exists $d_k \in D$ such that $f(x_k + \Delta_k d_k) < f(x_k)$ then :

1.1. Set $x_{k+1} = x_k + \Delta_k d_k$ (change the iterate)

1.2. Set $\Delta_{k+1} = \Delta_k$ (no change to the step length)

2. If $f(x_k + \Delta_k d_k) \geq f(x_k)$ for all $d \in D$, then:

2.1. Set $x_{k+1} = x_k$ (no change to the iterate)

2.2. Set $\Delta_{k+1} = \frac{1}{2}\Delta_k$ (half the step length)

2.3. If $\Delta_{k+1} < \Delta_{tol}$ then stop.

Whenever there is a trial point that improves the objective function, we conclude that iterate is successful. Regardless of the procedure chosen for evaluation of trial points, the value of Δ_k is not reduced unless every trial point has been evaluated and is found unacceptable. In this case, none of the $2n$ trial points has lead to an improved solution. Such iteration is called unsuccessful.

Furthermore, following each unsuccessful iteration, Δ_k is compared to the preset stopping criteria Δ_{tol} to test the convergence. Once the step length falls below Δ_{tol}, the search terminates with $x_* = x_{k+1}$. The convergence of the compass search method has been proved for continuously differentiable functions.

The method is considered as a local method and is only guaranteed to find a stationary point. The direct search method is a reliable method to find the global optimizer of the function when the initial step length is chosen long enough and when it is combined with any heuristic global search method (Kolda, Lewis & Torczon 2003).

3.4.3. Hybrid search approach

One of the successful strategies in dealing with global optimization is to combine a local search method and global search method. With the global search method, we are trying to explore the whole search space to find a rough estimate of the global optimal point. The most important feature of the global search method is its ability to escape from local optimal points. After the Global search method finds some information about the global optimal point, we can use the local search method to locate it more precisely.

In the hybrid method used in this thesis, we apply simulated annealing as a global search method. SA has proved ability to explore the search space and it can escape from local minima by means of probability of acceptance (refer to section 3.4.1 for details). After some information about the global optima, we use the Secant Method or Direct Search method to find the global point more precisely.

Table 3.3: Hybrid search method

1. Use global search method from an initial value to find x_k. In this thesis, we used simulated annealing method.

2. Apply local search method from x_k to find a better point x_{k+1}. In this paper, we used compass direct search method

3. If the required accuracy has not been achieved , go to the step 1 and start global search by setting initial value to x_{k+1}

3.4.4. Evolutionary Algorithm (EA)

Stochastic global search such as Evolutionary Algorithm (EA) (Dueck & Scheuer 1990)are *population based metaheuristics*. Compared to *neighborhood based metaheuristics*, in the population based approach, the probability of choosing an inappropriate initial point is minimized due to the use of a large number of initial points of EA distributed over the whole search space. The search is then focused on promising regions of the search space by successively narrowing the regions until the search converges.

The *population based metaheuristic* Evolutionary Algorithm (EA) is an iterative and stochastic optimization techniques inspired by the concepts from Darwinian evolution theory (Goldberg 1989). Several authors have proposed different versions of Evolutionary Algorithms

(EA) including theory of evolution (Darwin 1859), Evolution Strategies (ES) (Fogel 1994) and Evolutionary Programming (EP) (Hwang & He 2006). In general, an EA performs an evolutionary process on a population of solutions with the purpose of evolving the best possible approximate solution to the optimization problem. It operates on a given initial population of potential solutions to the problem and applies the principle of survival of the fittest to produce better and better approximations to a solution of the given problem. At each iteration, a pair of solutions is selected from a pool of solutions according to their level of fitness in the problem domain, which are bred together to produce a new set of solutions using the reproduction operators. The process of creating new solutions by combining the selection process of the parent pool and breeding processes directs the evolution of a population of solutions that are better suited to the problem domain. The whole process is executed over several iterations (generations) until a candidate solution of the problem with sufficient quality is found (Huda, S 2009).

EA provides significant advantages over traditional optimization algorithms because of the simultaneous use of several search techniques and heuristics such as population based search, a continuous balance between exploitation (convergence), exploration (maintained diversity) and the principle of building-block in its search

process. Some of the important features of EA over traditional neighborhood based metaheuristics and global optimization algorithms are noted in the following:

- Neighborhood based metaheuristics are single candidate model based and affected by the choice of initial point. However in the EA, the probability of choosing an inappropriate initial point is minimized due to the use of a large number of initial points of EA distributed over the whole search space.

- Compared to other global optimization algorithms, EA does not require any derivative information of the objective function or other knowledge about the structure of the problem. Therefore, EA can be applied on wide varieties of optimization problems as a black-box optimization algorithm (Lin & Keringhan 1973).

- Simultaneous use of non deterministic transition operators for generating new solutions and the use of several solutions in the population, implements a good diversification strategy in the search process which gives EA higher global exploration capability than other metaheuristic approaches.

Structure of the EA

Evolutionary algorithm models natural evolution processes. Thus, a typical EA incorporates many of the sub-process logically similar to

the sub-process of natural evolution including selection, evolutionary operations (re-combination, mutation etc.), and fitness evaluation. Figure 3.1 describes the structure of a simple EA and Algorithm in Table 3.4 shows the basics steps of an EA.
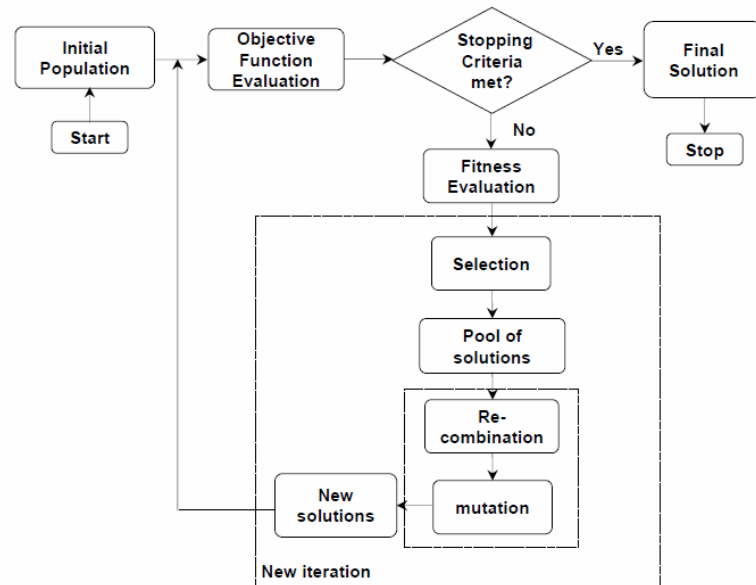
Figure 3.1: Structure of Evolutionary Algorithm (EA)

Table 3.4: Basic steps in an Evolutionary Algorithm (EA)

Step-1: Create an initial Population.

repeat

Step-2: Evaluate the initial population using the objective function.

Step-3: Compute the fitness of the population.

Step-4: Build the pool of solutions using selection operator.

Step-5: Apply re-combination operator to create the new pool of solution.

Step-6: Apply mutation operator to the new pool of solution.

Step-7: Apply replacement strategy to form the new population for next iteration (generation).

Until (*Termination criterion is reached*)

RETURN Best solution from the final population

End

Initial Population: Initial population of the EA comprises a number of solutions and specifies the starting point of the search. Initial population could be created using random initialization. The main goal of initialization process is to create a population with a good coverage of the search space. Any knowledge about the problem domain also may be used to create the initial population.

Objective function and fitness evaluation: The objective function measures the performance of a solution with respect to its parameters and is related to problem under consideration. The value of the objective function for one solution is independent of the values of the parameters of other solution in the population.

However, the fitness of a solution measures its reproductive ability and ability to survive. Unlike the objective function, the fitness of a solution is always defined with respect to other solutions of the

population being assessed. The fitness function transforms the value of objective function into a measure of reproductive ability.

Selection operator and pool of solutions: Selection operator is used to build a pool of solutions for reproduction of new solutions for next iteration from the current population. According to the Schema theorem (Holland 1975), a reproductive opportunity is allocated to each solution in the current population in proportion to their relative fitness. Therefore, solution with higher fitness gets higher probability of being selected for reproduction. Thus, the selection operator is implementing a survival-of-the-fittest strategy to build the pool of solutions. Many selection mechanisms have been proposed including Stochastic Universal Sampling, Roulette Wheel selection (Rechenberg 1978) and Tournament Selection (Baker 1987).

Re-combination operator and new solutions: The re-combination operator produces new solutions by exchanging some corresponding attribute value between the two solutions. Many re-combination operators have been proposed including one-point re-combination, two-point re-combination (Goldberg & Deb 1991), discrete re-combination and intermediate re-combination (Jong 1975). The simplest re-combination is the one-point re-combination where a position along the two solutions is randomly chosen. Then the one-point re-combination exchanges the sub-sequences before

and after that position between two solutions to create two new solutions.

Mutation Operator: After re-combination, the new solutions are passed through mutation. Mutation operators are stochastic operator which provides small amount of randomness to the variable of the new solutions and maintains a sufficient level of variety in the domain value. This in turn, re-introduces necessary solution features into populations that have been unintentionally lost after several iterations have passed.

Mutation prevents premature convergence:

At the beginning, the values of the new solutions in the population are randomly distributed providing a wide spread of individual fitness. As the iteration progresses, it is possible that selection operator will drive most of the solutions in the population to share the same value for some variables. Then the range of fitness level of the population reduces. As a result, EA looses the ability to continue to search for better solutions. If this happens without the EA converging to a satisfactory solution, then the search process has *prematurely converged*. This may particularly happen if the *population size is small*. In this situation, re-combination operator alone cannot prevent the premature *convergence*. By providing a small amount of randomness to the new solutions in the vicinity of the population, mutation operator maintains sufficient level of

diversity in the domain value and prevents any possible premature convergence of the search process.

Replacement Scheme for new iteration: Once the solutions are produced by evolutionary operators, then the current population is replaced with the new solutions. Different replacement strategy can be applied. An *elitist* replacement strategy replaces the worst solutions so that the significant features of the best solutions from the previous population can be transformed into next iteration. A *non- elitist* strategy replaces all solutions from the current population.

The fundamental steps of an EA have been mentioned here. Although, for many years, EA has been applied in many applications including optimization, design and creative systems (Dueck & Scheuer 1990), (Muhlenbein 1993), (Chambers 2001), it faces difficulties to find a high quality solution. Like *premature convergence*, another major problem of EA is the *slow convergence*. After many generations, the average values of the fitness of the solutions will be high and the range of fitness of the population becomes small which indicates a small gradient in the fitness function. Therefore, the selective pressure is also reduced. Due to this, population slowly advances towards a global maximum. The problem can partially be avoided by using a *fitness scaling* (Bentley &

Corne 2002), (Dueck & Scheuer 1990). Since EA explore the large solution space, to obtain high quality solution we also have to employ good evolutionary operators and tune the parameters as well.

3.5. Summary

This chapter discusses techniques of fitting known distribution to non-normal quality characteristics data and parameter estimation techniques of fitted distributions. In the later part of this chapter important metaheuristic approaches (including Simulated Annealing, Compass Direct Search and Evolutionary Algorithm) have been described with their basic structure. One important hybrid approach using single candidate model metaheuristic with local search has also been presented.

The proceeding chapters are our main contribution to this research study and application of methods and procedures described in Chapter 3 will be discussed in the subsequent chapters.

Chapter 4

UNIVARIATE PROCESS CAPABILITY

ANALYSIS

4.1. Introduction

This chapter describes the estimation of process capability indices C_p and C_{pk} for non-normal single quality characteristics data using methods and procedures presented in previous chapters. A newly proposed PCI estimation method (Liu, P & Chen 2006) bases on Burr XII distribution percentiles is discussed first in this chapter. Subsequently, a comprehensive review of the existing non-normal univariate PCI estimation methods using simulation study as well as real data examples has been presented. Maximum Likelihood Method is proposed to improve the accuracy of PCI estimation based on Burr percentile method and a metaheuristic technique (Simulated Annealing) has been deployed to estimate the parameter's of the fitted Burr distribution.

Later in the chapter, a new root transformation technique called Best Root Transformation (BRT) to estimate PCI for non-normal quality characteristics data is proposed. Lastly, Burr cumulative density

function for PCI estimation using Cumulative Density Function method has also been presented which is in contrast to the approach adopted in the research literature i.e. use of best-fitting density function from known distributions to non-normal data for PCI estimation.

4.2. Process capability estimation using Burr XII distribution

Although Clements method presented in section 2.3 is popular among quality practitioners, however, research studies (Wu, Wang & Liu 1998), (Ahmad, Abdollahian & Zeephongsekul 2007a) indicated that Clements method can not accurately measure the capability indices when the underlying data distribution is non-normal. Liu et al. (Liu, P & Chen 2006) and Ahmad et al. (Ahmad, Abdollahian & Zeephongsekul 2008) has conducted a detailed analysis of Clements method and introduced a new approach based on Burr distribution percentiles for evaluation of capability measures for non-normal quality characteristics data. In the proceeding section we will first present a shorter review of Burr XII distribution followed by a comprehensive review of Burr based method vs. existing conventional non-normal PCI estimation methods.

4.2.1. Review of Burr XII distribution

Burr (Burr, IW 1942) developed a number of useful cumulative frequency functions which can describe various non-normal distributions. One of them is the Burr XII distribution. This is widely used in reliability and quality literature. The probability density function of the Burr XII distribution is defined as follows:

$$f(y) = \begin{cases} \dfrac{cky^{c-1}}{(1+y^c)^{k+1}} & \text{if } y \geq 0; c \geq 1; k \geq 1 \\ 0 & \text{if } y < 0 \end{cases} \qquad (4.1)$$

Note that c and k represent the skewness and kurtosis coefficients of the Burr XII distribution respectively. Therefore, the cumulative distribution function of the Burr XII distribution is derived as:

$$F(y) = \begin{cases} 1 - \dfrac{1}{(1+y^c)^k} & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases} \qquad (4.2)$$

Burr (Burr, IW 1973) presented a wide range of skewness and kurtosis coefficients of various probability distributions that can be

approximated using different values of Burr distribution parameters c and k. For example, the normal density function can be estimated by a Burr distribution with c =4.85437 and k =6.22665 and a Gamma distribution with shape parameter 16 can be approximated by a Burr XII distribution with c = 3 and k = 6, and log-logistic distribution is also a special case of Burr XII distribution. Rodriguez (Rodriguez 1977) demonstrated that the Weibull distribution is a limiting distribution of the Burr XII distribution. In practice, it has been observed that majority of the quality characteristics follow Weibull distribution. Hence, the two-parameter Burr XII distribution can be used to describe the data in the real world.

Burr (Burr, IW 1973) has tabulated the means and standard deviations as well as skewness and kurtosis coefficients for the family of Burr distribution. These tables enable users to make a standardized transformation between a Burr variate (say Q) and another random variate (say X). The expression of the transformation is defined by

$$\frac{X - \bar{x}}{s} = \frac{Q - \mu}{\sigma} \qquad (4.3)$$

where $\bar{x}$ and s are the values of sample mean and standard deviation for the original sample data. μ and σ are the mean and

standard deviation respectively, for the family of Burr distribution relative to the original sample data.

The Burr XII distribution has been applied in areas of quality control, reliability analysis, and failure time modeling. Zimmer and Burr (Zimmer & Burr 1963) developed a method for sampling variables from non-normal populations using the Burr XII distribution. Burr (Burr, IW 1967) used his distribution to investigate the effect of non-normality on the limits of $\bar{X}$ and R control chart. Castagliola (Castagliola, P 1996) used Burr's approach to compute the proportion of nonconforming items.

4.2.2. Use of Burr distribution for non-normal PCI estimation

When the quality characteristics data is non-normally distributed, Burr XII distribution can be applied to estimate capability indices. It can provide better estimate of the process capability than the commonly used Clements's method. Liu and Chen (Liu, P & Chen 2006) (Ahmad, Abdollahian & Zeephongsekul 2007b) introduced a modification based on the Clements method, whereby instead of using Pearson curve percentiles, they replaced them with percentiles from an appropriate Burr distribution. Their proposed modified method involves the following steps:

- Estimate the sample mean, sample standard deviation, skewness and kurtosis of the original sample data.

- Calculate standardized moments of skewness (α_3)and kurtosis (α_4) for the given sample size n as follows:

$$\alpha_3 = \frac{n^2}{((n(n-1))^{3/2}} \, \Sigma \left\{ \frac{(x_j - \overline{x})}{s} \right\}^3 \qquad (4.4)$$

where $\overline{x}$ is mean of the observations and s is the standard deviation.

$$\alpha_4 = \frac{n}{(n-1)^2} \, \Sigma \left(\frac{x_j - \overline{x}}{s} \right)^4 - \frac{12(n-1)}{(n+1)(n-3)} \qquad (4.5)$$

Kurtosis used in the above equation is excess kurtosis; i.e. when calculating kurtosis, a result of +3.00 indicates the absence of kurtosis (distribution is mesokurtic). For simplicity in its interpretation, some statisticians adjust this result to zero (i.e. kurtosis minus 3 equals zero), and then any reading other than zero is referred to as excess kurtosis. Negative numbers indicate a platykurtic distribution; positive numbers indicate a leptokurtic distribution.

- Use the values of α_3 and α_4 to select the appropriate Burr parameters c and k (Burr, 1942). Then use the standardized

$Z = (X - \bar{X})/s = (Q - \mu)/\sigma$, where X is the random variate of the original data, Q is the selected Burr variate, μ and σ its corresponding mean and standard deviation respectively. The mean and standard deviations, as well as skewness and kurtosis coefficients, for a large collection of Burr distributions are found in the tables of Burr (Burr, IW 1973) and Liu and Chen (Liu, P & Chen 2006). From these tables, the standardized lower, median and upper percentiles are obtained. These tables enable users to make a standardized transformation between a Burr variate and another random variate using equation (4.3).

- Calculate estimated percentiles using Burr table for lower, median, and upper percentiles as follows:

$$L_p = \bar{X} + s\,Z_{0.00135} \qquad (4.6)$$

$$M = \bar{X} + s\,Z_{0.50} \qquad (4.7)$$

$$U_p = \bar{X} + s\,Z_{0.99865} \qquad (4.8)$$

- Calculate process capability indices using equations 2.1-2.4 presented in the previous chapter under section 2.3, i.e.

$$C_p = \frac{Usl - Lsl}{U_p - L_p} \quad , \quad C_{pu} = \frac{Usl - M}{U_p - M} \quad , \quad C_{pl} = \frac{M - Lsl}{M - L_p}$$

$$C_{pk} = \min\left[C_{pu} , C_{pl}\right]$$

A practical example following the above steps is presented in the following Table 4.1.

Table 4.1: Process capability calculation procedure using the Burr percentile method

Step	Procedure		Calculations
1	Enter specifications Upper tolerance limit Lower tolerance limit	 Usl Lsl	 32 4
2	Estimate sample statistics Sample size Mean Standard deviation Skewness Kurtosis	 n $\overline{X}$ s S_k K_u	 100 10.5 3.142 1.14 2.58
3	Estimate standardized moments of skewness (α_3) and kurtosis (α_4) using S_k and K_u values from step 2.	α_3 α_4	1.12 4.97
4	Based on α_3 and α_4 from step 3, select the parameters c and k values using the Burr XII distribution table Burr (Burr, IW 1973), Liu and Chen (Liu, P &	c k	2.347 4.429

	Chen 2006)		
5	With reference to parameters c and k obtained in step 4, use the table of standardized tails of the Burr XII distribution to determine standardized lower, median and upper percentiles Liu and Chen (2006). where Z_p = standardized Burr variate at percentile p.	$Z_{0.00135}$ $Z_{0.5}$ $Z_{0.99865}$	-1.808 -0.140 4.528
6	Calculate estimated 0.135 percentile using equation (4.6)	L_p	=10.5+(-1.808 x 3.142) =4.819
7	Calculate estimated 99.865 percentile using equation (4.8)	U_p	=10.5+ (4.528 x 3.142) =24.727
8	Calculate estimated median using equation (4.7)	M	=10.5+(-0.140 x 3.142) = 10.06
9	Calculate non-normal process capability indices (C_p, C_{pu}, C_{pl}, C_{pk}) using equations 2.1-2.4.	C_p C_{pu} C_{pl} C_{pk}	=(32-4)/(24.727-4.819) =1.40 =(32-10.06)/(24.727-10.06) = 1.49 = (10.06-4)/(10.06-4.819)= 1.15 =1.15

Instead of using moments of skewness and kurtosis to estimate parameters of a Burr distribution as presented here, other methods such as Maximum Likelihood, Method of Probability-Weighted Moments and Methods of L-Moments (O'Connell & Shao 2004) can also be used. However, the choice is determined by the fact that quality control practitioners with little background in theoretical statistics will find the estimation procedure adopted here, which is simply a moment matching process, much easier to comprehend and apply.

4.3. A simulation study

4.3.1. Comparison criteria

Different comparison yardsticks can lead to different conclusions. It is imperative to adopt such a criterion common among researchers and easy to understand and apply by industry professionals. In practice, capability indices are commonly used for tracking process performances and comparison between different processes. But such uses without examining the underlying distribution can lead to erroneous outcomes. A good surrogate capability index for non-normal data should be compatible with the process capability computed under normality when the corresponding fractions non-conforming are about the same. This approach has been adopted in

literature. Most of the researchers, for example, English et al.(English & Taylor 1993) used fixed value of capability indices C_p and C_{pk} =1.0 for all their simulation runs in investigating the robustness of these capability indices to non-normality. The basis for their comparison was the proportion of estimated C_p and C_{pk} from simulation data greater than 1.0 for the normal distribution case. This leads to a similar approach to Rivera et al. (Rivera, Hubele & Lawrence 1995). This approach is widely recognized yardstick for tackling the non-normality problem for process capability estimation. Rivera et al. (Rivera, Hubele & Lawrence 1995) used upper tolerance limits of the underlying distributions to calculate the actual number of non-conformance items and equivalent C_{pk} values. Estimated C_{pk} values calculated from the transformed data are then compared with the target C_{pk} values. Deleryd (Deleryd, M 1996) also used proportion of no-conforming items to derive the corresponding value of capability index. Then bias and dispersion of the estimated capability index have been compared with the target C_p values. A similar motivated scheme has been used as a comparison yardstick for one-sided C_{pu} by Tang et al. (Tang & Than 1999) and Liu and Chen (Liu, P & Chen 2006), (Albing 2006) and (Ahmad, Abdollahian & Zeephongsekul 2008) in their non-normal PCIs studies. For a

target C_{pu} value, the fraction of non-confirming units for a normal distribution can be determined using

$$\text{Fraction of non} - \text{confor min g items} = \Phi(-3C_{pu}) \qquad (4.9)$$

where $\Phi(x)$ refers to the cumulative distribution function of the standard normal random variable (Kotz & Lovelace 1998).

In this thesis, the criterion for comparing all methods is to determine the precision and accuracy of their estimated process capability indices. A most suitable method will have the mean of the estimated C_{pu} values closest to the target value i.e. greater accuracy and will have the smallest variability, measured by standard deviation of the estimated C_{pu} values i.e. greater precision (Tang & Than 1999).

4.3.2. Underlying distributions

Weibull, Gamma and Lognormal distributions are used to investigate the effect of non-normal data on the process capability index. These distributions are known to have the parameter values that can represent mild to severe departures from normality. These parameters are selected so that we can compare our simulation results with existing results using the same parameters in the literature Tang et al. (Tang & Than 1999), Liu & Chen (Liu, P & Chen 2006).

The probability density functions of Weibull, Gamma and Lognormal distributions are given respectively by

Weibull (α, β):

$$f(x|\alpha, \beta) = \frac{\alpha}{\beta} x^{\alpha-1} e^{-x^{\alpha}/\beta}, \quad \alpha > 0, \ \beta > 0, \ x \geq 0 \qquad (4.10)$$

Gamma (α, β):

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \quad \alpha > 0, \ \beta > 0, \ x \geq 0 \qquad (4.11)$$

Lognormal (μ, σ^2):

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\frac{e^{-(\log x - \mu)^2/2\sigma^2}}{x}, \qquad -\infty < \mu < \infty, \ x \geq 0$$
$$(4.12)$$

Figures 4.1, 4.2 and 4.3 below show respectively the density functions of the Weibull, Gamma and Lognormal distributions used in our simulation study.
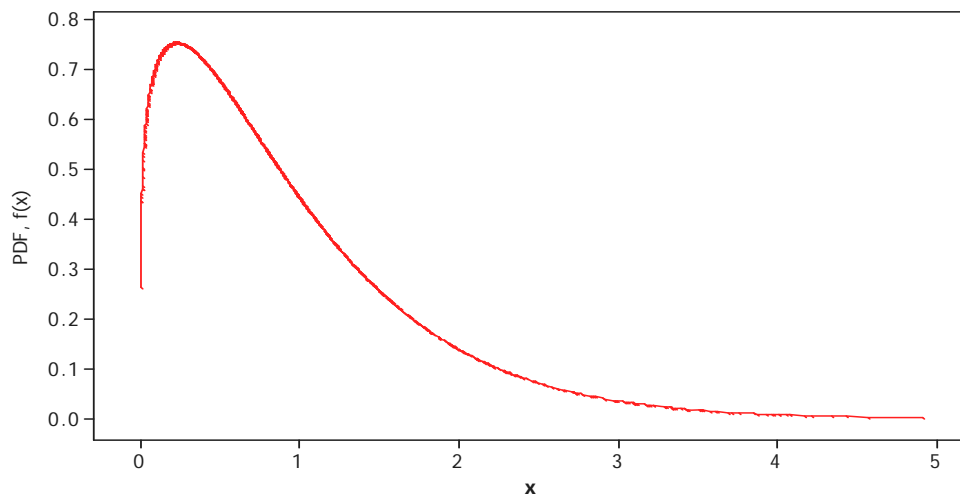
Figure 4.1: pdf of Weibull distribution with parameters (α = 1.2, β = 1.0)
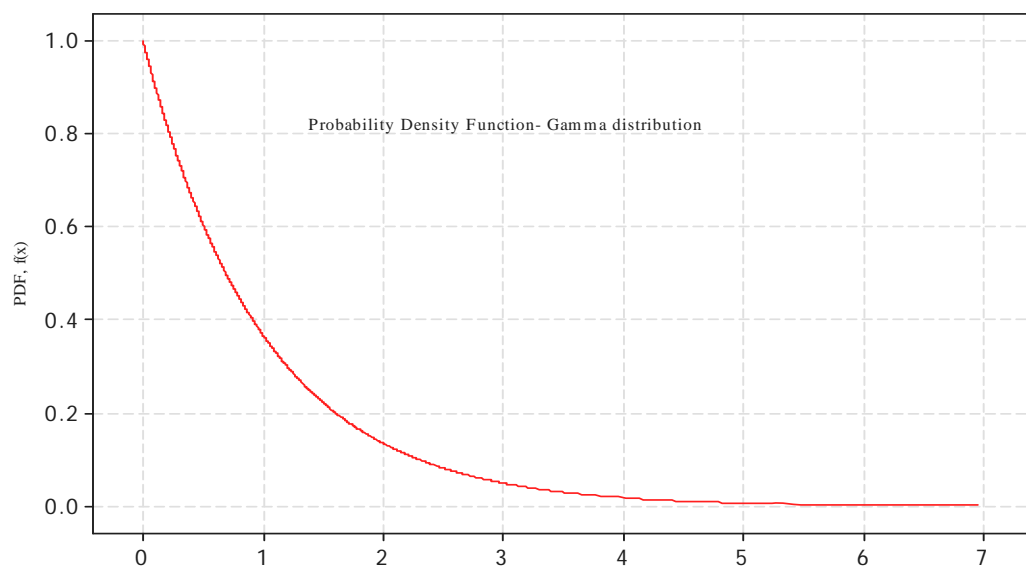
Figure 4.2: pdf of Gamma distribution with parameters (shape= 1.0, scale= 1.0)

Note that Gamma (1, 1) is in fact the Exponential distribution with mean 1.
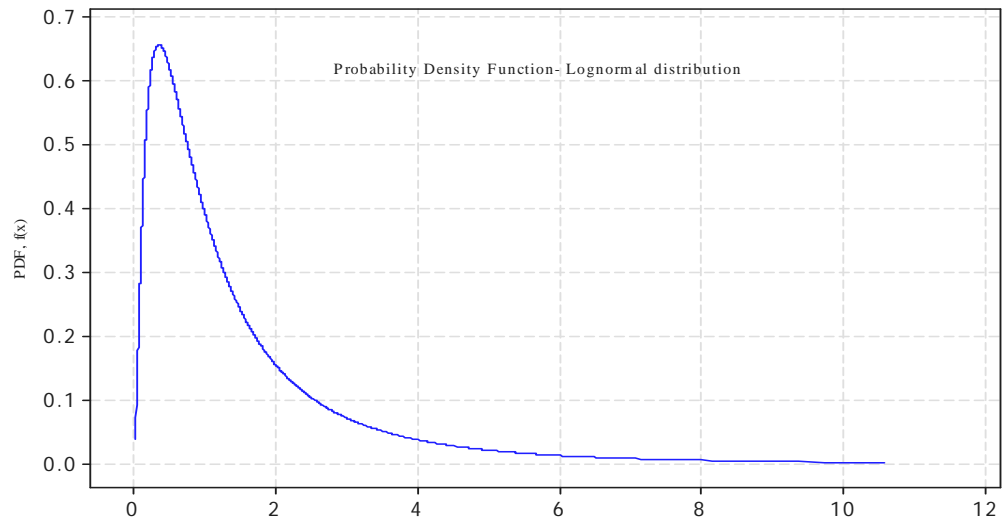
Figure 4.3:pdf of Lognormal distribution with parameters ($\mu = 0$, $\sigma^2 = 1.0$)

4.3.3.   Simulation methodology

In the simulation studies, targeted values of C_{pu} = 0.5, 1.0, 1.5 and

2.0 have been used. The corresponding value of upper specification

limit (*Usl*) for each distribution is obtained with same fraction of non-

conforming items as follows:

$$Usl = C_{pu}(X_{0.99865} - X_{0.5}) + X_{0.5} \qquad (4.13)$$

where $X_{0.99865}$ and $X_{0.5}$ are the designated percentiles of the

corresponding distribution. For example, if C_{pu} is equal to 1.5, and

we use say Weibull with parameter values 1.2 and 1, then using any

statistical package one can obtain $X_{0.99865}$ and $X_{0.5}$ for this

distribution which are 4.8236 and 0.7368 respectively. Then we use

equation (4.13) to find the corresponding *Usl* which is 6.867.

Table 4.2: Simulation methodology

<u>Step 1</u>	Choose a distribution with known parameters, e.g. Weibull α =1.2 and β =1.0
<u>Step 2</u>	Find $X_{0.99865}$ and $X_{0.5}$ for this distribution using any statistical package.
<u>Step 3</u>	Choose a target C_{pu} value, say C_{pu} = 1.5
<u>Step 4</u>	Use equation (4.13) to calculate Usl value, which equals Usl =6.867 for this example.
<u>Step 5</u>	Next we compare 3 methods, Clements (C), Box-Cox (B-C) and Burr (Bu) using following steps.
<u>Step 5-1</u>	Simulate values from underlying distribution
<u>Step 5-2</u>	Use each method to estimate $X_{0.99865}$ and $X_{0.5}$
<u>Step 5-3</u>	For fixed C_{pu} value, say 1.5 and corresponding *Usl* value, say 6.867, calculate the C_{pu} values using all three methods (similar to Table 4.1)
<u>Step 5-4</u>	Compare these calculated C_{pu} values using standard statistical measures and graphs to decide which method leads to the most accurate estimate of target C_{pu} value.

We simulate 30 samples of size 100 from each distribution and follow the steps outlined in Table 4.2 to calculate the corresponding C_{pu} for each sample. *Usl* value obtained using equation (4.14) and is used to estimate the C_{pu} index pertaining to the different methods from the simulated data. These estimated $C_{pu}(s)$ are then compared with the targeted C_{pu} values. A superior method is one with its sample mean of the estimated $C_{pu}(s)$ having the smallest deviation from the target value C_{pu} (accuracy) and with the smallest variability, measured by the spread or standard deviation of the estimated $C_{pu}(s)$ values (precision). A graphical representation that conveniently depicts these two characteristics is the simple Box-and Whisker plot.

4.3.4. Simulation runs

For our simulation study, we have generated 30 samples of size 100 from Weibull, Gamma and Lognormal distributions. After each simulation run, the necessary statistics, such as mean, standard deviation, median, skewness, kurtosis, upper and lower 0.135 percentiles were obtained. In this chapter, C_{pu} a process capability index with unilateral tolerance limit is used as comparison criterion. This representative capability index for the non-normal data should be compatible with that computed under normality assumption, given the same fraction of non-confirming parts (Tang & Than 1999).

The estimates for C_{pu} were determined using Burr, Clements and Box Cox methods following steps outlined in Table 4.2. The average value of all 30 estimated C_{pu} values and their standard deviations were calculated and presented in Tables 4.3-4.5 below.

Table 4.3: The mean and standard deviation of 30 C_{pu} values with n=100 (Weibull)

		Burr		Clements		Box Cox	
Cpu	*Usl*	*mean*	*Std*	*mean*	*Std*	*mean*	*Std*
0.5	2.780	0.596	0.090	0.590	0.099	0.621	0.100
1.0	4.824	1.152	0.159	1.159	0.175	0.956	0.194
1.5	6.867	1.708	0.228	1.727	0.252	1.204	0.283
2.0	8.910	2.264	0.297	2.296	0.328	1.407	0.367

Table 4.4 :The mean and standard deviation of 30 C_{pu} values with n=100 (Gamma)

		Burr		**Clements**		**Box Cox**	
Cpu	*Usl*	*mean*	*Std*	*mean*	*Std*	*mean*	*Std*
0.5	3.650	0.578	0.091	0.593	0.105	0.611	0.075
1.0	6.608	1.117	0.166	1.159	0.188	0.897	0.132
1.5	9.565	1.655	0.241	1.725	0.271	1.099	0.185
2.0	12.522	2.194	0.316	2.290	0.354	1.262	0.233

Table 4.5: The mean and standard deviation of 30 C_{pu} values with n=100 (Lognormal)

		Burr		Clements		Box Cox	
Cpu	*Usl*	*mean*	*Std*	*mean*	*Std*	*mean*	*Std*
0.5	4.482	0.499	0.084	0.496	0.097	0.503	0.048
1.0	8.339	1.024	0.166	1.031	0.183	0.710	0.061
1.5	12.009	1.523	0.243	1.541	0.265	0.832	0.070
2.0	15.679	2.022	0.320	2.050	0.348	0.921	0.076

To investigate the most suitable method for dealing with non-normality presented by Weibull, Gamma, and Lognormal distributions, we present box plots of estimated C_{pu} values using all three methods. The box plots are presented for different targeted C_{pu} values. Box plots (Figures 4.4-4.6) are able to graphically display important features of the simulated C_{pu} values, such as the median, variability and outlier. Figures 4.4-4.6 using Weibull, Gamma and Lognormal data indicate that the mean C_{pu} for Burr method is the closest to the targeted C_{pu} values and the spread of the C_{pu} values using the Burr method is smaller than that of Clements method, therefore, indicating a better approximation. Box Cox method indicates comparable results for smaller target C_{pu} values.
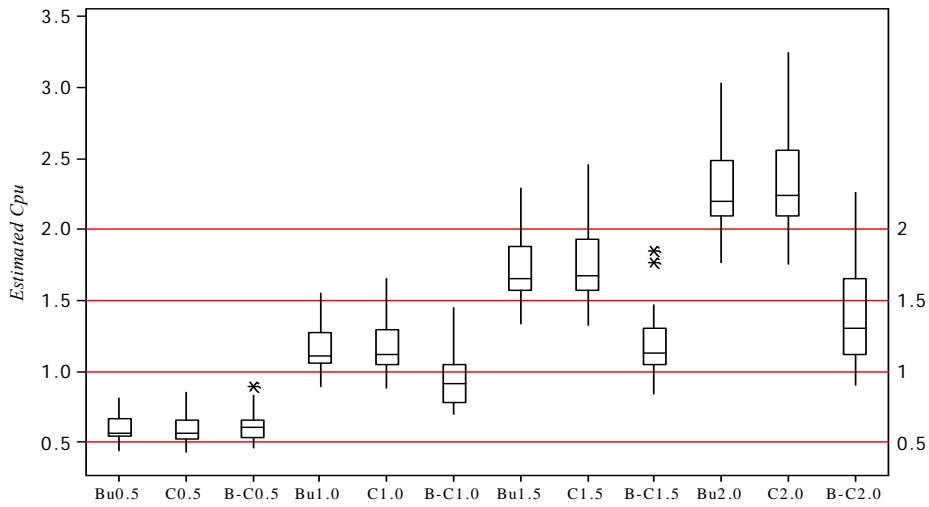
Figure 4.4: Box plot of estimated C_{pu} values with target C_{pu} =0.5, 1.0, 1.5, and 2.0 for Weibull
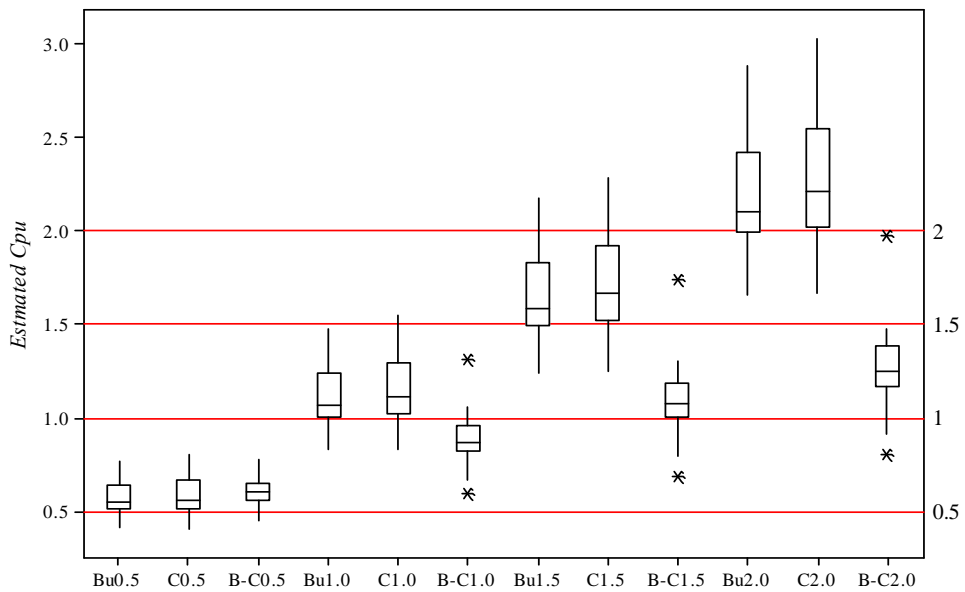
Figure 4.5: Box plot of estimated C_{pu} values with target C_{pu} =0.5, 1.0, 1.5, and 2.0 for Gamma
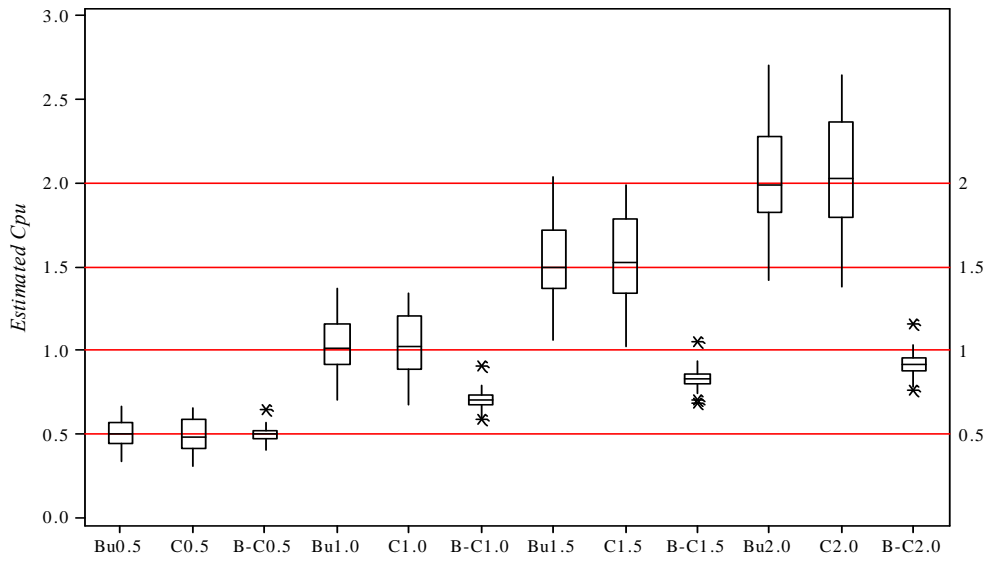
Figure 4.6: Box plot of estimated C_{pu} with target C_{pu} =0.5, 1.0, 1.5, and 2.0 for Lognormal

4.4. Discussion of results

Both Clements and Burr's methods included in this simulation study yield estimates which are close to the target C_{pu} values. However, Box Cox method performs well for smaller targeted C_{pu} values but underestimates for higher targeted C_{pu} values. As mentioned above, the performance yardstick is to determine the accuracy and precision with the given sample size. To determine accuracy, we have looked at the mean of the estimated C_{pu} values using all three methods. For precision, we have focused on the standard deviation of the estimated C_{pu} values for all three methods.

Looking at the results depicted in Tables 4.3 – 4.5, we conclude that the Burr method is the one for which the mean of the estimated C_{pu} value deviates least from the target C_{pu} value. Also for a given sample size, standard deviation of estimated C_{pu} values using the Burr method is smaller than Clements method. However, Box Cox method does not yield results close to any targeted C_{pu} values except for small targeted C_{pu} values.

During simulation investigation, we also observed that a larger sample size yields better estimates for all methods. Therefore sample size does have impact on process capability estimates. It was also observed that a larger target value of C_{pu} leads to slightly worse estimates using all methods.

4.5. Case studies

4.5.1. Real example 1

The data set used in this case study is obtained from a semiconductor manufacturing industry. The data are the measurements of bonding area between two surfaces with upper specification $Usl = 24.13$. $\overline{X} - R$ Chart was used to check whether the process is stable or not before analyzing experimental data. Figure 4.7 shows histogram of the data. The histogram indicates that the

underlying distribution is not normal and has a long right tail. Using a Goodness of Fit Test, the data is best fitted by a Gamma distribution with $\hat{\alpha}$ = 2543.8, $\hat{\beta}$ = 0.00921.
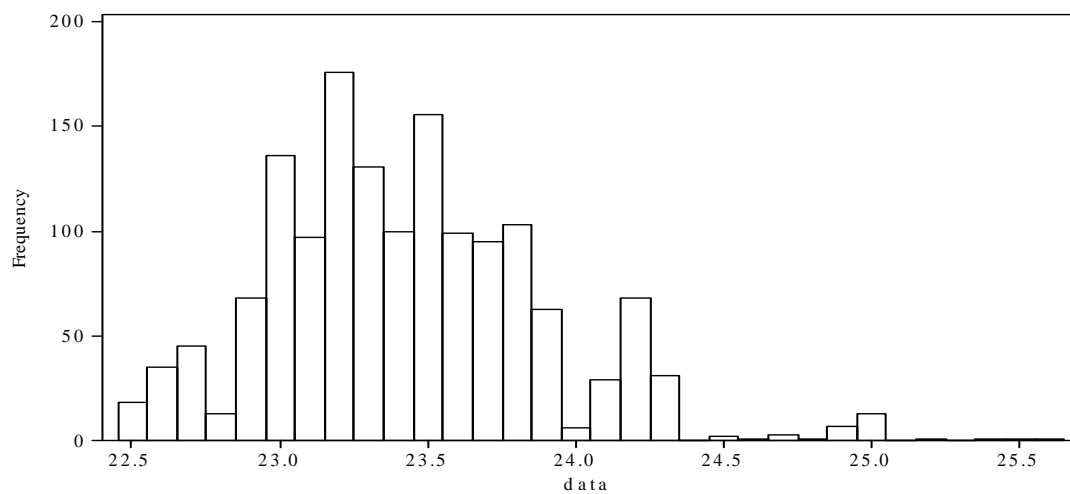
Figure 4.7: Histogram of measurement data from an attach assembly process

We have used all three methods to estimate C_{pu} values for 30 samples of size 50 from this process. The mean and standard deviation of the estimated C_{pu} values using each method is presented in Table 4.6.

Table 4.6: Result of the real example by using 30 Samples of size n=50.

method	upper tolerance (Usl)	C_{pu} mean	C_{pu} standard deviation
Clements	24.13	0.4368	0.0934
Burr	24.13	0.4207	0.0768
Box Cox	24.13	0.4999	0.0608

The actual C_{pu} value of this process (using equation (4.10)) based on 1500 data is 0.3775. This C_{pu} value is obtained by using $X_{0.99865}$ and $X_{0.5}$ of the actual data. The results in Table 4.6 show that mean C_{pu} value obtained using the Burr is closest to the actual C_{pu} value.

4.5.2. Real example 2

In this case study, we present the capability analysis of a real set of data called connector obtained from computer manufacturing industry in Taiwan (Wang 2006). The data set has one sided specification limit i.e. Usl = 0.2 mm. Figure 4.8 shows histogram of the data. The histogram indicates that the underlying distribution is not normal and has a long right tail. Using a Goodness of Fit Test has shown that data is not best fitted to any known distribution.
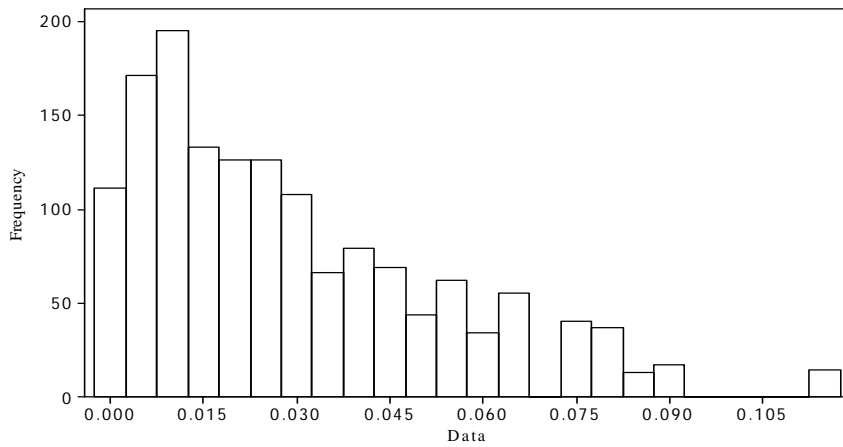
Figure 4.8: Histogram of measurement data from connector assembly process

All three methods have been applied to estimate process capability of this right skewed data. The estimated C_{pu} results are presented in Table 4.7.

Table 4.7: Process capability estimation results using n=100.

method	upper tolerance (*Usl*)	C_{pu} mean	C_{pu} standard deviation
Clements	0.2mm	2.0432	0.4354
Burr	0.2mm	1.9251	0.3490
Box Cox	0.2mm	1.2431	0.2266

The actual C_{pu} value of this process (using equation (4.10)) is 1.8954. This C_{pu} value is obtained by using $X_{0.99865}$ and $X_{0.5}$ of the actual data. The results in Table 4.7 show that mean C_{pu} value

obtained using the Burr and Clements are closer to the actual C_{pu} value.

4.6. Further application of Burr distribution to estimate process capability indices

This section is devoted to new research to investigate Burr distribution and its application to assess process capability for non-normal data. In the previous section, Burr distribution has been introduced in the percentile method where as the following section deals with its application using two new methods (Ahmad et al. 2008) and (Hosseinifard et al. 2009).

4.6.1.   Cumulative Density Function method using Burr distribution

In this section, we compare and contrast the Cumulative Distribution Function (CDF) method using Burr XII distribution. The results are then compared with the latest proposed process capability evaluation methods such as Burr percentile method and commonly used Clements percentile method when the underlying distribution is non-normal. Wierda (Wierda 1993) introduced a new approach to evaluate process capability for a non-normal data using Cumulative Distribution Function (CDF). Castagliola (Castagliola, P 1996) used

CDF approach to compute proportion of non-conforming items and then estimate the capability index using this proportion. Castagliola showed the relationship between process capability and proportion of non-conforming items and used CDF function to evaluate PCI for non normal data by fitting a Burr distribution to the process data. He used a polynomial approximation to replace empirical function in the Burr distribution, and then applied the equation given in (4.14). Using CDF method, C_p and C_{pk} are defined by

$$C_p = \frac{\Phi^{-1}(0.5 + 0.5 \int_{lsl}^{usl} f(x)dx)}{3} \tag{4.14}$$

where

$$C_{pl} = \frac{\Phi^{-1}(0.5 + \int_{lsl}^{T} f(x)dx)}{3} \tag{4.15}$$

$$C_{pu} = \frac{\Phi^{-1}(0.5 + \int_{T}^{usl} f(x)dx)}{3} \tag{4.16}$$

where f(x) represents the normal probability density function of the process and T represents the process mean for normal data and process median for non-normal data.

Here is a short proof of the above mentioned capability index equation (4.14).

Conventionally, if the process X is normally distributed with mean μ and standard deviation σ, i.e. $X \sim N (\mu,\sigma^2)$, then capability index is as defined as

$$C_p = \frac{1}{3}\Phi_z^{-1}\left(\frac{1}{2}+\frac{B}{2}\right)$$
(4.17)

where $B = P(Lsl < X < Usl)$.

Note that

$$P(X < Usl) = \frac{1}{2}+\frac{B}{2}$$

since

$$Z = \frac{X-\mu}{\sigma}$$

so $$P\left(Z < \frac{Usl-\mu}{\sigma}\right) = \frac{1}{2}+\frac{B}{2}$$

which is equivalent to

$$\Phi_z^{-1}\left(\frac{1}{2}+\frac{B}{2}\right)=\frac{Usl-\mu}{\sigma} \qquad (4.18)$$

As the *pdf* of *Z* is symmetric about the origin,

$$-\frac{(Lsl-\mu)}{\sigma}=\frac{(Usl-\mu)}{\sigma}=\Phi_z^{-1}\left(\frac{1}{2}+\frac{B}{2}\right) \qquad (4.19)$$

By equating (5.9), finally,

$$C_p=\frac{Usl-Lsl}{6\sigma}=\frac{1}{6}\left(\frac{Usl-\mu}{\sigma}-\frac{Lsl-\mu}{\sigma}\right)$$

$$=\frac{1}{6}\left[\Phi_z^{-1}\left(\frac{1}{2}+\frac{B}{2}\right)+\Phi_z^{-1}\left(\frac{1}{2}+\frac{B}{2}\right)\right] \qquad (4.20)$$

where we have used (4.19) and (4.20), which concludes the proof.

In this method $f(x)$ in Equation (4.15) is replaced by Burr density function (refer to chapter 3). We first fit Burr distribution function $f(x)$ to process data and then evaluate the PCI using CDF method.

To fit the data distribution with Bur distribution, we need to estimate c and k parameters. The likelihood function of univariate Burr is:

$$L(c,k;x_1,....,x_n) = \frac{c^n k^n \prod_{i=1}^{n} (x_i)^{c-1}}{\prod_{i=1}^{n} (1 + x_i^c)^{k+1}} \qquad (4.21)$$

In univariate Burr distribution there are two parameters c and k; and to estimate these parameters the maximum likelihood function with sample size n is:

$$\log L = n\log(c) + \log(k) - (1+k)\sum_{i=1}^{n}\log(1+x_i^c) + (c-1)\sum_{i=1}^{n}\log x_i \quad (4.22)$$

Differentiating with respect to parameters c and k give:

$$\frac{\partial l}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} \log x_i - (k+1)\sum_{i-1}^{n} \frac{\log x_i \log x_i^c}{1 + x_i^c} \qquad (4.23)$$

$$\frac{\partial l}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n}\log(1+x_i^c) \qquad (4.24)$$

Unknown Burr parameters c and k have been determined by maximizing equation (4.22) using a systematic random search algorithm called Simulated Annealing (refer to chapter 3 for details). The steps for SA method to obtain Burr parameters c and k are describes in Table 4.8.

Table 4.8: Simulated Annealing algorithm

1.	Obtain a covariance distance variable data.
2.	Decide control parameters of SA, i.e. T_o, T_f, C, I
3.	Generate random values c, k
4.	Compute the likelihood function, L, at this randomly generated values.
5.	If $T > T_o$ Then $T = CT$
6.	For I = 1 to I

      6. 1. Generate neighboring values, say c_1, k_1, for c, k

      6. 2. Compute the likelihood function (L_o) at the new solution obtained in step 6.1.

      6. 3. Evaluate parameters

          6. 3. 1. if $L_o > L$ then $c = c_1, k = k_1$, and $L = L_o$

          6. 3. 2. else

             generate a random value, u, from uniform distribution $Uni(0,1)$

          6. 3. 2. 1. if $u < e^{-\frac{(F_o - F)}{T}}$ then $c = c_1, k = k_1$

  7.   Print c, k and L

      ** c, k are the estimates of Burr distribution (c, k) parameters.

4.6.1.1 Simulation study

Three non-normal distributions; Gamma, Weibull and Beta have been used to generate random data in this simulation. These distributions are used to investigate the effects of non-normal data on the process capability index. Same parameters for these distributions have been selected as given in research literature (Liu, P & Chen 2006), (Tang

& Than 1999). We then compare our simulation results with existing results using the same parameters in the literature.

The probability density functions of Weibull as well as for Gamma distributions have been given in equation (4.11) and equation (4.12) respectively. The probability distribution function of Beta distribution with shape 1 (α) and shape 2 (β) is given by

$$f(x) = \left[\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1 \qquad (4.21)$$

The parameters used in this simulation are: $\alpha = 4.4$ and $\beta = 13.3$
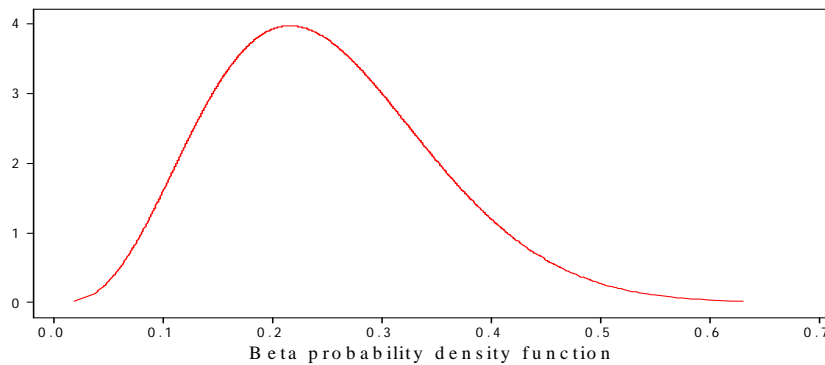
Figure 4.9: pdf of Beta distribution with parameters ($\alpha = 4.4$, $\beta = 13.3$)

Table 4.9: Comparison results

Distribution	Usl	C_{pu} *	C_{pu} Clements	C_{pu} Burr	C_{pu} CDF
Gamma(4,0.5)	6.3405	1.000	0.8698	0.9069	1.0000
Weibull(1,1.2)	5.0	1.043	0.9694	0.9738	1.0292
Beta(4.4,13.3)	0.5954	1.002	0.7434	0.7965	1.0028

4.6.1.2   Simulation results

The C_{pu} values in Table 4.9 are computed using equation (4.16) where $f(x)$ is replaced by the corresponding distributions (i.e. Gamma, Weibull and Beta). The C_{pu}* values in Table 4.9 are used to access the efficacy of the three methods in estimating process capability index for non-normal data. The simulations results show that C_{pu} values obtained using Clements method are not closer than those obtained using Burr and CDF methods. The C_{pu} values obtained using the CDF method are the closet to those C_{pu} values obtained using direct distribution percentiles in the conventional approach; thus, leading to better estimates of the PCIs compare with the Burr method.

Table 4.10 presents the probability of non-conformance using different methods. Probability of Non-Conforming (PNC) items is calculated using equation (4.22) as suggested by Castagliola (Castagliola, P 1996) for all three methods e.g. for Gamma distribution with C_{pu} value 0.8698, corresponding PNC value using equation (4.22) will be 0.0045351. The exact PNC value (p) in Table 4.10 is obtained using following equation.

$$PNC = 1 - \int_{0}^{usl} f(x)dx \qquad (4.22)$$

where $f(x)$ represents the corresponding distribution function of Gamma, Weibull and Beta distributions. Comparison criteria is that the method which yields expected proportion of non-nonconformities closest to that obtained using exact distribution would be the most superior method.

Table 4.10: Comparison of expected proportion of non-conformaning items with exact PNC

	Clements p_3	Burr p_2	CDF p_1	Exact p
Gamma	0.00454	0.00326	0.00135	0.0013
Weibull	0.00182	0.00170	0.00101	0.0010
Beta	0.01287	0.00844	0.00131	0.0013

Results in Table 4.10 show that PNC values obtained using Clements method are worse than the other two methods. In this table PNC

values using CDF method are close to the PNC values obtained using exact distribution (equation (4.22)). Thus the later method is giving better estimates of proportions of non-conformances as compared to the commonly used Clements and Burr methods (for details refer (Ahmad, Abdollahian & Zeephongsekul 2008)).

4.6.2.   Root Transformation Technique

As mentioned in chapter 2, the most serious issue with non-normal data is its existing skewness. Niaki et al. (Niaki & Abbasi 2007) used root transformation technique to design multi attribute control charts. In the proposed root transformation technique, we extend this technique to search the best root (r) of the non-normal data in such a way that if we raise the data to power r (i.e. X^r), then the transformed data will have zero skewness. We use the bisection method to find the r value. The bisection method is based on the fact that a function will change sign when it passes through zero. By evaluating the function at the middle of an interval and replacing whichever limit has the same sign, the bisection method can halve the size of the interval in each iteration and eventually find the root. Best root transformation search algorithm is given in Table 4.11 and an application example of this method using real data is presented in section 4.8.

Table 4.11: Best root transformation search algorithm

Let

$$k = 0$$

while

$$\left| f(x_{k+1}) \right| > \varepsilon$$

$$x_{k+1} = \frac{a_k + b_k}{2}$$

if

$$\left(f(x_{k+1}) f(a_k) < 0 \right)$$

then

$$a_{k+1} = a_k \text{ and } b_{k+1} = x_{k+1}$$

else

$$b_{k+1} = b_k \text{ and } a_{k+1} = x_k$$

end If

$$k = k + 1$$

end

while $x^* = x_k$

After finding the best root value, we calculate new specification limit by using $Usl = (Usl)^r$, and then estimate mean and standard deviation of the transformed data to use classical method to estimate PCI. We stop the iteration in the bisection procedure when skewness is less than 0.05 or after 200 replications.

4.6.2.1   Simulation study

In the simulation study we use three non-normal distributions: Gamma, Weibull and Beta. We need some target values for PCI to

compare the estimated PCIs of different methods. Hence to provide target values for PCI in each distribution we obtain appropriate upper specification limits. We use four target C_{pu} values of 0.5, 1, 1.5 and 2 for this simulation study.

English & Taylor (English & Taylor 1993) used fixed values of C_p and C_{pk} (equal to 1.0) for all their simulation runs in investigating the robustness of PCIs to non-normality. The basis for their comparison was the proportion of C_p and C_{pk} (estimated from simulation) greater than 1.0 for the normal distribution case (Tang & Than 1999)).

In this thesis a procedure similar to that of Rivera et al. (Rivera, Hubele & Lawrence 1995) and Tang et al. (Tang & Than 1999) mentioned earlier in this chapter is used to calculate the upper specification limit corresponding to the target PCI. If we have the *pdf* of the data then the exact C_{pu} can be obtain from equation (4.23) (see Castagliola (Castagliola, P 1996)). Therefore corresponding *Usl* for each target C_{pu} should be true in equation (4.23). Equation (4.24) is used to obtain the upper specification limit for the given target C_{pu}.

$$C_{pu} = \frac{\Phi^{-1}\left[\int_{-\alpha}^{Usl} f(x)d(x)\right]}{3}$$

$$= \frac{1}{3}\Phi^{-1}\left[F(usl)\right] \qquad (4.23)$$

$$usl = F^{-1}(\Phi(3C_{pu})) \qquad (4.24)$$

where $f(x)$ and $F(x)$ are probability density function and cumulative density function of x respectively.

In our simulation study, target values of C_{pu} = 0.5,1, 1.5 and 2 are used, the corresponding *Usl* values for Gamma, Beta and Weibull distributions are obtained from equation (4.23). These Usl values are then used to estimate the C_{pu} using root transformation, Box-Cox and percentiles methods. These estimated C_{pu} values are then compared with the targeted C_{pu} values. Again, a superior method is the one with sample mean of the estimated C_{pu} closest to the target C_{pu} value (accuracy) and with the smallest variability, measured by the spread or standard deviation of the estimated C_{pu} values (precision). Table 4.12 and Table 4.13 show the results of simulation study.

Table 4.12: Simulation study results (sample size n= 100)

Distribution	Usl	Cpu	Mean Cpu-RT	Std BRT	Mean Box-Cox	Std Box-Cox
Gamma(4,0.5)	3.657	0.5	0.5084	0.0540	0.5061	0.0525
	6.3405	1	1.0180	0.1405	0.9991	0.1270
	9.966	1.5	1.5289	0.2961	1.4775	0.2428
	14.584	2	2.0271	0.4991	1.9418	0.3853
Weibull(1,1.2)	2.2924	0.5	0.5058	0.0506	0.5014	0.0487
	4.8235	1	0.9612	0.1101	0.9394	0.1010
	8.256	1.5	1.3729	0.1936	1.3280	0.1649
	12.511	2	1.7583	0.2955	1.6729	0.2420
Beta(4.4,13.3)	0.41068	0.5	0.5035	0.0509	0.5066	0.0505
	0.59438	1	0.9547	0.1137	0.9424	0.1081
	0.7551	1.5	1.3004	0.1840	1.2755	0.1718
	0.87162	2	1.5347	0.2429	1.4921	0.2201

Table 4.13 : Simulation study results (sample size n= 1000)

Distribution	Usl	C_{pu}	Mean $C_{pu\text{-}RT}$	Std $C_{pu\text{-}RT}$	Mean $C_{pu\text{-}Box\text{-}Cox}$	Std $C_{pu\text{-}Box\text{-}Cox}$
Gamma(4,0.5)	3.657	0.5	0.5007	0.016	0.5006	0.0162
	6.3405	1	0.9890	0.040	0.9870	0.0375
	9.966	1.5	1.4614	0.078	1.4551	0.0714
	14.584	2	1.9147	0.123	1.9053	0.1144
Weibull(1,1.2)	2.2924	0.5	0.5011	0.015	0.4975	0.0151
	4.8235	1	0.9475	0.034	0.9320	0.0300
	8.256	1.5	1.3456	0.059	1.3169	0.0496
	12.511	2	1.7072	0.088	1.6627	0.0726
Beta(4.4,13.3)	0.4107	0.5	0.5038	0.015	0.5022	0.0154
	0.5944	1	0.9411	0.034	0.9345	0.0317
	0.7551	1.5	1.2752	0.056	1.2601	0.0508
	0.8716	2	1.4962	0.073	1.4774	0.0655

The results presented in Table 4.12 and Table 4.13 of this simulation study are based on samples of size *n=100* and *n=1000.* The results indicate that root transformation method provides more accurate results compare with the percentiles method even when the exact distribution percentiles for PCI calculations are used. The results also

indicate that in most cases root transformation method performs better than Box-Cox method.

4.6.2.2   Case Study

A case study is presented using real world data from a semiconductor manufacturing industry. Percentiles, Box-Cox and root transformation methods have been used to estimate PCI for the experimental data. The data are the measurements of the contact area between two surfaces with Usl = 24. Thirty samples of size 50 are selected from these data. The summary statistics for 1500 data are Mean = 23.4809, standard deviation = 0.5650, median = 23.3963, skewness = 1.1098 and kurtosis= 4.9740. For each sample; C_{pu} value is computed using Clements, Burr, Box-Cox and root transformation methods.

The results presented in Tables 4.14 indicate that root transformation method provides a better estimation of PCI compared to the other three PCI methods. The proportion of nonconforming (PNC) of the process data is 0.168 (that is 16.8% of data are above Usl = 24) which is very close to the expected proportion of non-conforming based on the proposed root transformation method.

Table 4.14: Result of the real example based on 30 samples of size n=50

methods		C_{pu}		
		Mean	Standard	Corresponding PNC
Transformation methods	BRT	0.328	0.1097	0.1626
	Box-Cox	0.358	0.0794	0.1416
Percentile methods	Burr	0.348	0.066	0.1483
	Clements	0.361	0.081	0.1396

The desired performance of root transformation method is depicted in Table 4.14. Both simulation and case study using experimental data from a semiconductor industry have indicated that root transformation technique provides better estimate of the process capability for non-normal processes compared with Box-Cox and percentiles methods; however Box-Cox method has smaller standard deviation.

4.7. Summary

The main purpose of this chapter is to compare and contrast among different methods of obtaining process capability indices and determine which method is more capable in achieving higher accuracy in estimating these indices for non-normal quality characteristics data. Simulation study as well as experimental data indicates that among percentile methods Burr percentile method

generally provides better estimate of the process capability for non-normal data. CDF and Root transformation methods have also indicated a potential to accurately estimate proportion of non-conforming products as well as process capability index, when underlying data is non-normal when compared to the percentile methods.

In the following chapter, application of Burr distribution for a bi-variate non-normal data will be discussed.

Chapter 5

BIVARIATE PROCESS CAPABILITY

ANALYSIS

5.1. Introduction

It is well known that process capability analysis for more than one quality variables is a complicated and sometimes contentious area with several quality measures vying for recognition. When these variables exhibit non-normal characteristics, the situation becomes even more complex. The aim of this chapter is to estimate Process Capability Indices (PCIs) for bivariate non-normal process using the bivariate Burr distribution. In the previous chapter, we have seen that by using Burr XII distribution, the accuracy of estimates of process capability for univariate non-normal distributions (see for example, (Castagliola et al. (Castagliola, P & Castellanos 2005)and Liu & Chen (Liu, P & Chen 2006), Ahmad et al. (Ahmad, Abdollahian & Zeephongsekul 2008)) is much improved. In this chapter, we will extend application of bivariate Burr distribution to estimate process capability for bivariate non-normal data. Cumulative Density Function (CDF) method will be deployed for non-normal bivariate data in contrast to its (CDF) method original application to bivariate

normal data (Castagliola et al. (Castagliola, P & Castellanos 2005)). The process of obtaining these PCIs will be accomplished in a series of steps involving estimating the unknown parameters of Burr distribution using maximum likelihood estimation coupled with simulated annealing. Finally, the Proportion of Non-Conformance (PNC) obtained using this method will be compared with those obtained from variables distributed under the bivariate Beta, Weibull, Gamma and Weibull-Gamma distributions.

5.2. Non-normal process capability for more than one quality characteristics

In the field of statistical quality control, it is generally assumed that the distributions of quality characteristics are normal. But, in most practical cases this assumption is not valid and the distribution of the quality characteristics may follow non-normal distributions such as Gamma, Beta, and Weibull distributions. In the past decade, several modifications of classical process capability indices have been proposed to resolve the issue of non-normality for quality characteristics data.

Although many researchers have proposed several methods to handle the issue of non-normality for univariate quality characteristics data, however, it is often observed that quality of a product does not depend only on a single quality characteristic. Thus

process capability estimation for more than one quality characteristic is vital. In the research literature few efforts have been presented for construction of bivariate as well as multivariate PCIs.

PCIs for more than one quality characteristics, in general, can be obtained from (a) the ratio of a specification region to a process region or modified process variation region, (b) the probability of nonconforming items over rectangular tolerance zone, and (c) implementing loss functions and vector representation, (d) theoretical proportion of non-conforming items over convex polygons, (e) global approach viewing multivariate quality control (Zahid & Sultana 2008). Taam et al. (Taam, Subbaiah & Liddy 1993) developed the first multivariate process capability index based on ratio of tolerance region to a process region approach. Chen (Chen, H 1994) also proposed a method for multivariate PCI using a non-conforming proportion approach. Shahriari et al. (Shahriari, Hubele & Lawrence 1995) proposed a process capability multivariate vector in order to evaluate process performance. Castagliola et al. (Castagliola, P & Castellanos 2005) proposed a new capability index dedicated to two quality characteristics. This approach is based on the computation of the theoretical proportion of non-conforming items (PNC). This approach has some interest for researchers as it is straightforward, logical and easy to deploy by non-statisticians (engineers and front line quality practitioners), for normal as well as

for non-normal data. Castagliola et al. (Castagliola, P & Castellanos 2005) have also extended univariate method as presented in Castagliola (Castagliola, P 1996) to bivariate distribution but again limited its application to bivariate normal data, and compared the results against existing methods for multivariate normal processes.

In this research work, Cumulative Density Function (CDF) approach has been extended to estimate bivariate non-normal PCI where bivariate Burr distribution is fitted to bivariate non-normal quality characteristics. Preliminary to this, we also use the bivariate Burr distribution with three parameters (Durling 1975) to fit bivariate non-normal data. Another contribution of this thesis is to estimate parameters of fitted Burr distribution to the bivariate non-normal data using heuristics technique called Simulated Annealing (SA), (refer to chapter 3 for details).

5.2.1. Cumulative Density Function (CDF) approach for bivariate data

Castagliola (Castagliola, P 1996) defined a relationship between process capability and proportion of non-conforming items and presented a new approach to evaluate process capability index for non-normal data. This approach is based on the generalized Burr distribution to assess the capability of the process data. Through the sample empirical distribution function, he used a polynomial function

to approximate a Burr distribution and from this obtained the Process Capability Indices.

As mentioned in Chapter 4, using Cumulative Density Function (CDF) method for the normal quality characteristics data, C_p and C_{pk} are defined by

$$C_p = \frac{\Phi^{-1}(0.5 + 0.5\int_{lsl}^{usl} f(x)dx)}{3} \qquad (5.1)$$

$$C_{pk} = \min(C_{pu}, C_{pl}) \qquad (5.2)$$

where

$$C_{pl} = \frac{\Phi^{-1}(0.5 + \int_{lsl}^{T} f(x)dx)}{3} \qquad (5.3)$$

and

$$C_p = \frac{\Phi^{-1}(0.5 + \int_{T}^{usl} f(x)dx)}{3} \qquad (5.4)$$

where $f(x)$ represents the probability density function of the process and T represents process target. For non-normal distribution, the above equations can still be used to obtain process capability; but T would represent the median instead of mean of the process data.

Castagliola et al. (Castagliola, P & Castellanos 2005) extended above mentioned univariate CDF method to multivariate normal distribution; by replacing the univariate probability density function $f(x)$ in equation (5.1) with the multivariate normal probability density function $f(x_1, x_2, ..., x_p)$ with a multivariate normal pdf.

$$C_p = \frac{\Phi^{-1}(0.5 + 0.5 \int_{lsl1}^{usl1} \int_{lsl2}^{usl2} \int_{lsl3}^{usl3} .... \int_{lslp}^{uslp} f(x_1, x_2, ..., x_p) dx_1 dx_2 .... dx_p)}{3} \qquad (5.6)$$

Keeping in view the above literature survey, there is an opportunity for researchers to explore a suitable capability evaluation method that can address the complex situation of multivariate non-normal data. In this research study, we replace probability density function $f(x_1, x_2)$ in equation (5.6) with the bivariate Burr distribution. The efficacy of the proposed method will be assessed by using the Proportion of Non-Conformance (PNC) criterion. In the proceeding section, a review of bivariate Burr distribution is presented first.

5.3. Bivariate Burr distribution

Burr (Burr, IW 1942) developed a number of useful cumulative frequency functions which can describe various non-normal distributions. One of them is the Burr XII distribution. This distribution has been reviewed thoroughly in chapter 4. In this section we will

review bivariate Burr distribution. Durling (Durling 1975) introduced the bivariate Burr distribution as follows:

$$f(x_1, x_2) = \frac{\Gamma(p+2)}{\Gamma(p)} b_1 b_2 x_1^{b_1-1} x_2^{b_2-1} (1 + x_1^{b_1} + x_2^{b_2})^{-(p+2)},$$

(5.7)

$$x_1, x_2 \geq 0, \quad b_1, b_2, p \geq 0$$

The cumulative distribution function has the form:

$$F(x_1, x_2) = 1 - (1 + x_1^{b_1})^{-p} - (1 + x_2^{b_2})^{-p} + (1 + x_1^{b_1} + x_2^{b_2})^{-p},$$

(5.8)

$$x_1, x_2 \geq 0, \quad b_1, b_2, p \geq 0$$

In the bivariate Burr distribution there are three parameters, b_1, b_2 and p to be estimated. These parameters can be estimated by maximizing the log-likelihood function based on a sample of size n given by L:

$$L(b_1, b_2, p : x_1, ..., x_n) = n \ln b_1 + n \ln b_2 + n \ln p + n \ln(1 + p) +$$

(5.9)

$$(b_1 - 1) \sum_{j=1}^{n} \ln x_{1j} + (b_2 - 1) \sum_{j=1}^{n} \ln x_{2j} - (2 + p) \sum_{j=1}^{n} \ln(1 + x_{1j}^{b_1} + x_{2j}^{b_2})$$

(x_{1j}, x_{2j}), $j = 1, 2, \ldots\ldots n$ is an observed bivariate sample. The first order condition for maximizing L with respect to b_1, b_2 and p lead to the following equations:

$$\frac{\partial L}{\partial b_1} = \frac{n}{b_1} + \sum_{j=1}^{n} \ln x_{1j} - (2 + p) \sum_{j=1}^{n} \frac{x_{1j}^{b_1} \ln x_{1j}}{1 + x_{1j}^{b_1} + x_{2j}^{b_2}} = 0 \quad (5.10)$$

$$\frac{\partial L}{\partial b_2} = \frac{n}{b_2} + \sum_{j=1}^{n} \ln x_{2j} - (2 + p) \sum_{j=1}^{n} \frac{x_{2j}^{b_2} \ln x_{2j}}{1 + x_{1j}^{b_1} + x_{2j}^{b_2}} = 0 \quad (5.11)$$

$$\frac{\partial L}{\partial p} = \frac{n}{p} + \frac{n}{1+p} - \sum_{j=1}^{n} \ln(1 + x_{1j}^{b_1} + x_{2j}^{b_2}) = 0 \quad (5.12)$$

Since the process of obtaining the solutions to (5.10) – (5.12) is numerically challenging, in this chapter we will use a systematic random search algorithm called "Simulated Annealing" (refer to Chapter 3 for details) to obtain the estimated parameters directly from equation (5.9).

5.4. Fitting bivariate Burr distribution to bivariate non-normal data

As mentioned earlier, the Cumulative Density Function (CDF) method will be used to evaluate process capability for non-normal bivariate characteristics. To use equation (5.6), one needs to calculate the probability of quality characteristics falling between specification limits. In order to calculate this probability we first need to know the distribution of the data.

As shown in the earlier chapter Burr distribution can easily be fitted to any real data, in this chapter we use bivariate Burr distribution to calculate the probability of non-conforming products in a bivariate non-normal process. Maximum likelihood estimation (MLE) method is used to estimate its unknown parameters b_1, b_2 and p. Since the maximum likelihood function (MLF) of bivariate Burr is complex and may have some local optima, and numerical methods used to solve equations may also give local optima, we will maximize likelihood function by using Simulated Annealing algorithm (SA). Abbasi et al. (Abbasi et al. 2006) used simulated annealing to estimate three parameters of Weibull distribution through MLE method and they observed that it was fast and the results were very accurate.

5.4.1. Process capability evaluation using bivariate Burr distribution

After obtaining bivariate Burr distribution and fitting to bivariate non-normal data, we will then use equation (5.6), replacing $f(x_1, x_2, \ldots x_p)$ in the numerator with the bivariate Burr distribution (equation (5.12)) to compute process capability. Table 5.1 outlines the procedure of the proposed procedure.

Table 5.1: C_p computation procedure

Step 1	Select a sample from the process.
Step 2	Write down the maximum likelihood function (MLF) for sample based on bivariate Burr distribution.
Step 3	Maximize MLF by using Simulated Annealing and estimate b_1, b_2 and p.
Step 4	From Eq (5.6) compute the difference between cumulative densities function at the upper specification limits ((Usl_1, Usl_2) and the lower specification limits (Lsl_1, Lsl_2), i.e. $$B = F(Usl_1 \,,\, Usl_2) - F(Lsl_1 \,,\, Lsl_2) \,.$$
Step 5	$$C_p = \frac{\Phi^{-1}(0.5 + 0.5B)}{3}$$

Step 6	Compute the corresponding $PNC = \Phi(-3C_p)$ and compare it with the PNC obtained from the exact distribution for example Gamma.

5.5. Simulation studies

The purpose of this section is to show the ability of the proposed method for estimating the C_p value of non normal bivariate processes. Simulation studies have been conducted for bivariate non-normal processes. As discussed in the previous chapter different comparison yardsticks can lead to different conclusions. It is imperative to adopt such a criterion common among researchers and easy to understand and applicable by industry professionals. In this simulation study we will use the same comparison criterion as mentioned under Section 4.3, i.e. we use upper tolerance limits of the underlying distributions to calculate the actual number of non-conformance items and their corresponding C_p values. Estimated C_p values calculated from the fitted burr distribution are then compared with the target C_p values.

For this simulation study, underlying bivariate non normal distributions such as Gamma, Beta and Weibull and Weibull- Gamma

are used. Table 5.2 depicts the simulation methodology for this research study. NORmal-To-Anything (NORTA) method is used to generate simulation bivariate data. Refer to Cario & Nelson (Cairo & Nelson 1997) and Niaki et al. (Niaki & Abbasi 2007) for discussion of the procedures used to generate this simulation data. Table 5.2 illustrates the simulation methodology procedure for this research study.

Table 5.2: The flowchart for simulation methodology

Step 1	Generate 100 vectors from bivariate non-normal using one of above distributions. Compute expected proportion of non-conformance (p*) by using 1,000,000 data from the corresponding distribution e.g., Gamma and calculate the proportion of data falling outside the given USL.
Step 2-1	Fit Maximum likelihood function of Bivariate Burr distribution to data.
Step 2-2	Estimate parameters of the fitted bivariate Burr distribution using SA.
Step 3	Use Castagliola method to compute C_p for Bivariate Burr distribution Eq (5.6)
Step 4	Compute proportion of non-conforming for Cpu $PNC = \Phi(-3C_p)$ say p**
Step 5	Compare p* and p** to evaluate the accuracy of the proposed method

Table 5.3 presents the parameters of the bivariate non-normal distributions in the simulation study. The C_p value computed using the exact bivariate distributions, for example Gamma, is presented under exact distribution, then we have generated m=30 samples of size n=100 and fitted a bivariate Burr distribution to each sample. The parameters b_1, b_2 and p for the fitted Burr distribution are estimated using simulated annealing (SA) algorithm.

Table 5.3: Simulation methodology for bivariate non-normal distribution

Distribution	Parameters[1]	Specification Limits[2]		C_p	Burr (n=100) (m=30)	
		Lsl	Usl	Exact Distribution	Mean	Std
Gamma	a=[3,4] β=[2.5,3]	[0,0]	[20,45]	0.7761	0.7124	0.1117
Gamma	a =[3,6] β=[8,10]	[0,0]	[100,160]	1.0405	1.0109	0.2371
Beta	a =[2,4] β=[5,4]	[0.0025, 0.005]	[0.9,0.92]	0.8645	0.8100	0.0403
Weibull	a =[3,4] β =[1,2]	[0,0]	[15,12]	0.8943	0.9089	0.3830
Gamma, Weibull	a =[5,2] β=[3,5]	[0,0]	[45,6.8]	0.8541	0.8260	0.3381

[1]Note that each value in the pair represents the corresponding marginal distribution.
[2]Specification limits are selected to represent natural Specification limits.

The C_p(s) of these 30 samples are calculated using equation (5.6). The mean and standard deviation for 30 computed C_p(s) are presented in the last column of Table 5.3. The results in Table 5.3

show that the mean C_p for different bivariate non-normal distributions are very close to the exact C_p value calculated using exact distribution (for example C_p value for using exact Beta distribution data in Table 5.3 is 0.8645 vs. fitted Burr distribution to the same data which C_p value of 0.8100). From the results in the last column of Table 5.3 one can imply that the proposed CDF method enables user to estimate C_p value reasonably accurate by fitting bivariate non-normal data with the bivariate Burr distribution.

Table 5.4: Proportion of non-conformance

Distribution	Burr PNC (p^{**})	Expected PNC (p^*)
Gamma	0.033	0.020
Gamma	0.002	0.002
Beta	0.015	0.010
Weibull	0.006	0.007
Gamma, Weibull	0.013	0.010

To further assess the efficacy of the proposed method, we have also calculated the Proportion of Non-Conformance (PNC) data using $\text{PNC} = \Phi(-3C_p)$. Table 5.4 shows that the proportions of non-conforming items using proposed method is close to the expected proportion of non-conforming items obtained using the true underlying distributions.

5.6. Conclusion

In this chapter, Cumulative Density Function approach for bivariate data has been discussed and applied to estimate the process capability index for bivariate non-normal quality characteristic data. We used the bivariate Burr distribution to fit the probability density function to the bivariate non-normal data. We have used simulated annealing algorithm which will maximize the log likelihood function based on Burr distribution. We have presented the results using simulated data from non-normal bivariate distribution such as Gamma, Beta and Weibull. The results revealed that the proposed method provides close estimates of process capability when compared with the values of process capability obtained using exact distribution. Using the expected non-conformance proportion criterion, the results indicate that proportions of non-conformance obtained using the proposed method is close to those obtained under the exact distributions.

In the following chapter, application of Burr distribution to estimate process capability for a multivariate non-normal correlated quality characteristics data will be discussed.

Chapter 6

MULTIVARIATE PROCESS CAPABILITY ANALYSIS

6.1. Introduction

This chapter describes process capability measures for multiple quality characteristics. Unfortunately, multivariate capability measures that are currently employed, except for a handful of cases, depend intrinsically on the underlying multivariate data being normally distributed. In reality, the quality characteristics data is not only multivariate but also non-normal and most often quality characteristics are interrelated with each other. This non-normality and correlated characteristics of multivariate data poses a challenge to researchers to investigate accurate and effective process performance yardstick in the area of quality control.

In this chapter, we will present different methods to investigate a suitable multivariate performance measure. In the first section we will deploy geometric distance introduced by Wang (Wang 2006) to reduce the dimensionality of the correlated non-normal multivariate data and then fit Burr distribution to the geometric distance variable. The optimal parameters of the fitted Burr distribution will be

estimated using different numerical techniques. The proportion of non-conformance (PNC) will be used as a criterion for process performance measurements.

In the later part of this chapter we will introduce an innovative approach for a multivariate capability index based on the Generalized Covariance Distance (GCD). This proposed approach (Ahmad et al. 2009) is easy to use by frontline managers and quality practitioners. Another novelty introduced in this methodology is to approximate the distribution of these distances by a Burr XII distribution and then estimate its parameters using different numerical techniques. Examples based on real manufacturing process data are also presented which demonstrate that the proportion of nonconformance using proposed GCD method is very close to the actual proportion of nonconformance value.

6.2. Background of multivariate PCIs

It is an established fact that production processes very often produce non-normal data, and there is always more than one quality characteristics of interest in process outcomes and very often these characteristics are correlated with each other. For example in Taam et al. (Taam, Subbaiah & Liddy 1993), an engineering drawing of a connecting rod for a combustion engine specifies the dimensions of

crank bore inner diameter, pin bore inner diameter, rod length, bore true location, bore to bore parallelism and other features. To represent how well this connecting rod is made, one may examine numerical summaries of individual characteristics or a comprehensive summary for all characteristics. If one treats the rod as one entity, the latter is preferred. In situations where the design intent of a product is prescribed by a number of related characteristics, the functionality of this product cannot be represented by individual characteristics separately. Many other such examples are scattered in quality control literature. This poses the need of multivariate process capability analysis.

As mentioned in Chapter 5, in general, multivariate capability indices can be obtained from (a) the ratio of a specification region to a process region or modified process variation region, (b) the probability of nonconforming items over rectangular tolerance zone, and (c) implementing loss functions and vector representation, (d) theoretical proportion of non-conforming items over convex polygons, (e) global approach viewing multivariate quality control (Zahid & Sultana 2008). Taam et al. (Taam, Subbaiah & Liddy 1993) defined the first multivariate process capability index based on ration of volume of the modified tolerance region (R_1) to the volume of 99.73% process region (R_2) approach.

$$MC_{pm} = \frac{\text{Volume of } R_1}{\text{Volume of } R_2} \qquad\qquad 6.1$$

If the process data are multivariate normal, the R_2 is an elliptical region. The modified tolerance region is the largest ellipsoid completely within the engineering tolerance region and centered at the target.

Chen (Chen, H 1994) also proposed a method for multivariate PCI using a non-conforming proportion approach over a rectangular tolerance zone. In this method a general tolerance zone is defied by

$$V = \left\{ x \in R^V : h(x - T) \leq r_0 \right\} \qquad\qquad (6.2)$$

where $h(x)$ is a specific positive function with the same scale as x, T is a target value and $T \in R^V$ is a constant vector and r_0 is a positive number. Then a rectangular solid tolerance zone is defined by

$$V = \left\{ x \in R^V : |x_i - T_i| \leq n, i = 1, \ldots, v \right\} \qquad\qquad (6.3)$$

The process is capable if $P(x \in V) \geq 1 - \alpha$.

Let $r = \min\{c : P(h(x - T) \le c) \ge 1 - \alpha\}$. If the cumulative distribution function of $h(x - T)$ is increasing in a neighborhood of r, then r is simply the unique root of equation $P(h(x - T) \le c) = 1 - \alpha$. The process is deemed capable if $r \le r_0$. Here. r_0 is the half-width of the tolerance interval centered at the target value μ_0 and r is the half width of an interval centered on the target value such that the probability of a process realization falling within this interval is $1 - \alpha$.

Shahriari et al. (Shahriari, Hubele & Lawrence 1995) proposed a process capability multivariate vector in order to evaluate process performance. Hubele et al. (Hubele, Shahriari & Cheng 1991), using multivariate normal distribution, defined PCI as the ratio of the rectangular tolerance region to modified process region which is the smallest rectangle around the ellipse with type I error $\alpha = 0.0027$. The number of quality characteristics in the process is taken into account by taking the θth root of the ratio where θ represents the number of quality characteristics. This leads to the following index:

$$C_{PM} = \left[\frac{\text{Vol. of the engineering tolerance region}}{\text{Vol. of the engineering process region}} \right]^{\frac{1}{\theta}} \quad (6.4)$$

.

Here the modified tolerance region is the largest ellipsoid centered at the target which falls completely within the original tolerance region.

Wang (Wang et al. 2000) compared the above two multivariate process capability indices and presented some graphical examples to illustrate them. Chen et al. (Chen, K, Hsu & Wu 2006) extend Boyles' work (Boyles 1994) for multivariate normal distribution. They have also extended Huang et al.'s (Huang, Chen & Hung 2002) work for multivariate data but they have not considered the correlation between the variables. They computed process capability for multivariate data (without correlation) and for each individual variable.

Further from literature review, Kotz and Johnson reviewed the multivariate process capability indices thoroughly for assessing multivariate processes (Kotz & Johnson 1993), (Kotz & Johnson 2002)). According to (Wang 2006), multivariate capability indices proposed by many researchers in recent years suffer from the following restrictions:

- Normality assumption on multivariate data is usually required.
- Confidence intervals of the multivariate capability indices are difficult to derive.

- Higher dimension (more than three quality variables) capability indices are not readily obtainable except by the geometric distance approach and the principal component analysis method proposed by Wang et al (Wang & Hubele 1999)and Wang and Du (Wang & Du 2000).

Due to the above restrictions, it is evident that application of the conventional methods is limited. In order to deal with non-normal multivariate and correlated quality characteristics data, there is an opportunity for researchers to develop a more suitable PCI that can address the complex situation of multivariate non-normal and correlated data.

Geometric distance approach is used to reduce higher dimensionality of multivariate data in the first part of this chapter. Here we propose to fit just one distribution Burr XII distribution to the geometric distance variable instead of the traditional practice adopted and cited in statistical literature, i.e. fitting different distributions to the geometric data (Wang et al. (Wang & NF 1999), Wang (Wang 2006)). Furthermore, different numerical techniques such as Simulated Annealing, Compass direct search and Evolutionary Algorithm will be deployed to estimate parameters of the fitted Burr distribution. Also a comprehensive analysis based on conformance or nonconformance to customer specifications will be conducted. The

efficacy of the proposed methods will be assessed by using the proportion of nonconformance (PNC) criterion. The performance of geometric distance variables using Best-fit method and Burr fit methods is presented in the first part of this chapter.

6.3. Analysis methodology

Although several methods have been proposed to deal with non-normal univariate quality characteristics data, there has not been much research work devoted to process capability studies for multivariate non-normal quality characteristics. This field is still wide open for researchers. In the proceeding section; we will briefly review research methodology in regards to the subject mater and discuss fitting Burr distribution to geometric distance data.

6.3.1. Geometric distance approach

Geometric Distance (GD) approach was proposed by Wang and Hubele (Wang & Hubele 1999). It reduces the dimension of the multivariate process data and renders them more tractable for statistical analysis. The GD approach utilizes the Euclidean distance (or $L2$ norm) which is defined as follows:

Let $X = (x_1, x_2, \ldots \ldots, x_n)$ represent a point from a sample space and let $T = (t_1, t_2, \ldots \ldots, t_n)$ be the corresponding target value. Then the Geometric Distance (GD) variable is defined by

$$GD = \sqrt{(X - T)'(X - T)}$$

$$GD = \sqrt{(x_1 - t_1)^2 + (x_2 - t_2)^2 + \ldots + (x_n - t_n)^2} \qquad (6.5)$$

A comprehensive study of the distribution of GD when the underlying variables have a multivariate normal distribution was undertaken in Wang et al. (Wang & Hubele 1999). When the underlying distribution is non-normal, Wang (Wang 2006) combined correlated quality characteristics to form GD and determined the distribution that best fits GD by using Best-Fit statistical software. In this chapter, instead of using different distribution as practiced in the Best-fit approach (Wang 2006), we will fit just one distribution, the Burr XII distribution to the geometric distance data. Burr XII distribution has been applied extensively in the area of quality control, reliability analysis, and failure time modeling ((Ahmad, Abdollahian & Zeephongsekul 2007b), (Liu, P & Chen 2006), (Castagliola, P 1996)) and also cited in the research literature for its versatility to fit any

132

real data. Another hurdle to measure multivariate process performance is that every critical characteristic has its own specification limit. In the next section we will discuss how these multiple specification limits can be converted to single specification limit.

6.3.1.1 Conversion of multiple specification limits to single specification limit using maximum radial distance approach

The Maximum Radial Distance (MRD) {Wang, 1999 #18} is used as the upper specification of the geometric distance variable. MRD is the distance between the target and the perimeter of the tolerance region. One sided specification as proposed by many researchers ((Ahmad et al. 2008), (Liu, P & Chen 2006), (Tang & Than 1999), (Singpurwalla 1998)) is used here as a performance yardstick when the quality characteristics data do not follow normal distribution. In this case median = 0 and the upper specification limit (Usl) is defined by MRD:

$$\text{MRD} = \sqrt{(\text{Tol}_{X_1})^2 + (\text{Tol}_{X_2})^2 + (\text{Tol}_{X_3})^2, \dots (\text{Tol}_{X_k})^2} \qquad (6.6)$$

where Tol_{Xi} = Tolerance perimeter (s) of the quality characteristic X_i
i=1,2,......,k.

133

6.3.1.2 Estimation of the proportion of nonconforming products (PNC)

The criterion which is used to assess the efficacy of the proposed method is to determine the proportion of non-conformance as proposed by many researchers in the quality literature ((Ahmad, Abdollahian & Zeephongsekul 2008; Liu, P & Chen 2006; Singpurwalla 1998; Tang & Than 1999). Hence, using MRD as upper bound (Wang 2006) the estimated proportion of non-conformance for each geometric distance variable is given by:

$$PNC \ = 1 - F(MRD\ ) \ = 1 - \int_{0}^{MRD} f(x)dx \qquad (6.7)$$

and

Probability of the product conforming for a single variable

$$= \int_{0}^{MRD} f(x)dx \qquad (6.8)$$

where $f(x)$ is the density function of the Burr XII distribution.

It is straightforward to generalize this to higher dimension; the estimated proportion of nonconforming for a manufactured product with multiple quality characteristics is given by (Wang 2006)

$$PNC_{(Total)} = 1 - \prod_{i=1}^{k} \int_{0}^{MRD} f(x_i)dx_i \qquad (6.9)$$

and

Probability of the product conforming for multiple variables

$$= \prod_{i=1}^{k} \int_{0}^{MRD} f(x_i)dx_i \qquad (6.10)$$

where $f(x_i)$ is the density function of the i_{th} GD variable, $i = 1, 2, \ldots, k$.

Table 6.1: A flowchart of the proposed methodology using Geometric Distance Approach

1. Identify the desired multiple quality characteristics along with their respective engineering specifications.

2. Collect measurements of these quality characteristics data from a manufacturing process.

3. Determine the correlated and uncorrelated quality characteristics using statistical software

4. Compute geometric distance variable *(GD)* for correlated and uncorrelated quality characteristics using equation (6.5).

5. Compute maximum radial distances *(MRD(s))* from the target

6. Determine geometric distance variables that do not have any significant correlation with each other using statistical software.

7. Use random search method to estimate the parameters of the fitted univariate Burr distribution to each geometric distance variable.

8. Compute the proportion of conforming items for each geometric distance variable (using equation (6.8)).

9. Compute the proportion of nonconformance (PNC) for each geometric distance variable using equation (6.7).

10. Compute total proportion of nonconformance *(PNC)* value using equation (6.9).

6.3.1.3 Distribution Fitting to Geometric Data and Parameter Estimation

The parameters of the fitted Burr distribution to Geometric data are estimated using different numerical techniques (refer Chapter 3 and 4 for details). Our approach contrasts with that adopted by Wang (Wang 2006), where different distributions are fitted to different sets of geometric distance data using best-fit technique.

6.4. Comparisons with results using Geometric Distance (GD) approach

In order to demonstrate the application of the proposed methodology, a real data set is used from Wang' paper (Wang 2006). Wang discussed a manufacturing product (called connector) from a computer industry having multivariate (seven) quality characteristics. These seven characteristics are X_1 (contact gap X), X_2 (contact loop T_p), X_3 (LLCR), X_4 (contact x T_p), X_5(contact loop diameter), X_6 (LTGAPY) and X_7 (RTGAPY), respectively. The specification limits for these characteristics can be two-sided or one-sided, and they are 0.10 ± 0.04 mm, $0 + 0.50$ mm, 11 ± 5 mΩ, $0 + 0.2$ mm, 0.55 ± 0.06 mm, 0.07 ± 0.05 mm and 0.07 ± 0.05 mm, respectively. The full data is given in Appendix A1.

6.4.1. Example 1

For the first example, we selected a sample of 100 parts with three quality characteristics. The specification limits for these three quality characteristics are 0.10 ± 0.04 mm, $0 + 0.50$ mm, 11 ± 5 mΩ respectively. Histogram of selected three quality characteristics is given in Figure 5.1.

137
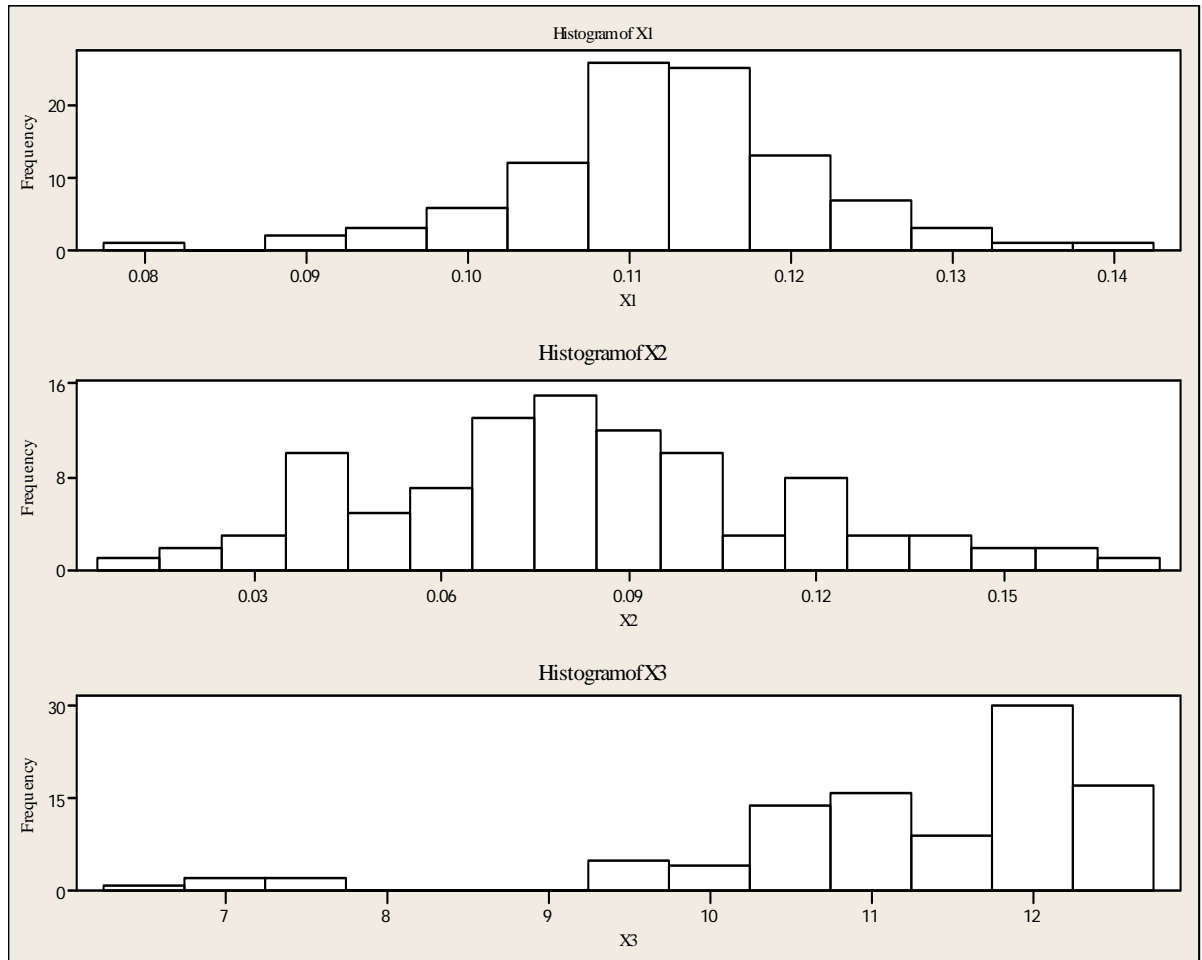
Figure 6.1: Histogram of X_1, X_2, X_3 characteristics.

Using a statistical package we found that variables $\{X_1, X_2, X_3\}$ are correlated. Correlation and Covariance matrix are given in Table 6.2 below:

Table 6.2: Correlation and Covariance Matrix

Correlations:X1, X2, X3	Covariances:X1, X2, X3

	X1	X2		X1	X2
X2	0.168		X1	0.0000879	
	0.095		X2	0.0000526	0.0011161
X3	-0.044	-0.017	X3	-0.0005381	-0.0007216
	0.66	0.869			

Using the geometric distance approach (Wang 2006), new univariate variable (GD) is given by

$$GD = \sqrt{(X_1 - 0.1)^2 + (X_2 - 0)^2 + (X_3 - 11)^2}$$

The numerical results of the geometric variable are given in Table 6.3. The data is reduced from multivariate dimension to univariate dimension.

Table 6.3: Geometric distance variable data for X_1, X_2 and X_3

GD	GD	GD	GD
0.22670	3.74730	1.28260	0.93040
0.16490	0.44080	1.44450	1.06170
0.09300	4.48960	1.63410	1.01460
0.42890	4.11990	0.48420	1.37030
0.29490	0.22090	0.91660	1.21480
0.29310	0.57590	1.41780	1.24580
0.51990	0.34790	0.83700	0.33240
0.22270	0.43450	0.81690	0.67060
0.16070	0.22740	1.02110	0.35920
0.23290	0.62170	0.65070	0.75850
0.34510	0.96020	0.89690	1.00190
0.40060	1.27150	0.78950	0.22600
0.12760	1.38680	1.70420	0.71950

0.27120	1.36900	1.53540	0.73560
0.39440	1.45620	0.88310	0.83630
0.37460	1.28410	1.18320	1.66870
0.33520	0.84090	0.91810	1.25220
0.33730	1.22940	1.35020	1.15580
0.10550	1.51320	1.12100	1.68570
0.75520	1.08030	0.55750	0.71970
0.22610	1.02610	1.03290	0.82650
0.09850	0.82730	1.19960	0.78920
0.19680	1.38710	1.02670	1.63370
3.60420	1.34210	0.94180	1.61810
4.21820	1.25510	0.96150	1.55800

The maximum radial distance (MRDs) using equation (6.6) is

$$\text{MRD} = \sqrt{(\text{Tol}_{X_1})^2 + (\text{Tol}_{X_2})^2 + (\text{Tol}_{X_3})^2}$$

$$= \sqrt{(0.04)^2 + (0.50)^2 + (5.0)^2} = 5.025$$

Burr XII distribution is fitted to the geometric distance variable (GD) using Secant and Compass numerical search algorithms. The estimated parameters of fitted Burr distribution to GD data are displayed in Table 6.4. In addition, the probability of the product nonconforming for GD variable is calculated using equation (6.7) and also summarized in Table 6.5

Table 6.4: Burr distribution parameter (c, k) estimation

Parameter Search Algorithm	Burr distribution parameter "c"	Burr distribution parameter "k"
Simulated Annealing	1.887	1.487
Compass Direct	1.95934	1.47977

Table 6.5: PNC for Geometric Distance Data

Parameter Search Method	MRD using Eq.(6.6)	Probability of the product conforming using Eq.(6.8)	Probability of the product nonconforming using Eq.(6.7)	Actual Proportion of nonconforming
SA	5.0251	0.999	0.001	0.01
Compass	5.0251	0.991	0.009	

Using the PNC criterion, Table 6.5 shows that the probability of the product nonconforming obtained by using both search methods (Compass and Simulated Annealing search methods) is close to the actual proportion of nonconforming. The actual PNC in Table 6.5 represents the actual proportion of data that fall outside their respective specification limits given by the computer manufacturer. Results in Table 6.5 indicate that estimation of fitted Burr distribution parameters using both search methods yield comparable results. This approach contrasts with that adopted by Wang (Wang 2006), where different distributions are fitted to different sets of geometric distance variables. It is shown that Burr distribution parameter's estimation using search methods have leaded to PNC value that closer to the exact value. We therefore recommend that the Burr

distribution be fitted to other non-normal multivariate data to analyze their performance.

6.4.2. Example 2: Using five quality characteristics and different numerical method to estimate Burr parameters

In the second example we have selected five (5) multivariate quality characteristics (X_1, X_2, X_3, X_6, X_7) [refer to Appendix A1 for complete data set] and used different numerical (simulated annealing and hybrid search) algorithms to estimate parameters of Burr distribution. The specification limits for these quality characteristics are 0.10 ± 0.04 mm, 0 + 0.50 mm, 11 ±5 mΩ, 0.07 ± 0.05 mm and 0.07 ± 0.05 mm, respectively. Using a statistical package we found that first three variables (X_1, X_2, X_3) are correlated and the other two (X_6, X_7) are correlated. Based on their correlation, we reduce multivariate data from five quality characteristics to two variables called geometric variables (GD_1 and GD_2). Table 6.6 and Table 6.7 shows correlation and covariance matrix respectively.

Table 6.6: Correlation matrix

Correlation	X_1	X_2	X_3	X_6
X_2	0.168			
p-value	0.095			
X_3	-0.044	-0.017		
p-value	0.66	0.869		
X_6	0.024	-0.046	0.184	

p-value	0.81	0.65	0.067	
X_7	0.089	-0.047	0.002	0.613
p-value	0.378	0.639	0.985	0

Table 6.7: Covariance matrix

	X1	X2	X3	X6	X7
X1	8.79E-05				
X2	5.26E-05	0.001116			
X3	-0.00054	-0.00072	1.665644		
X6	3.4E-06	-2.3E-05	0.003567	0.0002265	
X7	1.45E-05	-2.8E-05	4.24E-05	0.0001604	0.0003025

The numerical results of the geometric variables GD_1, GD_2 are given in the Table 6.8

Table 6.8: Geometric distance variable data

GD1	GD2	GD1	GD2	GD1	GD2	GD1	GD2	GD1	GD2
0.22670	0.01290	0.22610	0.00840	1.28410	0.04130	0.89690	0.00910	1.24580	0.00920
0.16490	0.03500	0.09850	0.03380	0.84090	0.05080	0.78950	0.02950	0.33240	0.02760
0.09300	0.03920	0.19680	0.04490	1.22940	0.04110	1.70420	0.05190	0.67060	0.03020
0.42890	0.02080	3.60420	0.02710	1.51320	0.05340	1.53540	0.02220	0.35920	0.02730
0.29490	0.01990	4.21820	0.00860	1.08030	0.02460	0.88310	0.04620	0.75850	0.02280
0.29310	0.01130	3.74730	0.03960	1.02610	0.02110	1.18320	0.04920	1.00190	0.01910
0.51990	0.01860	0.44080	0.03370	0.82730	0.03940	0.91810	0.02020	0.22600	0.05260
0.22270	0.01500	4.48960	0.05070	1.38710	0.04660	1.35020	0.03430	0.71950	0.01600
0.16070	0.01840	4.11990	0.02260	1.34210	0.03090	1.12100	0.03880	0.73560	0.03130
0.23290	0.03100	0.22090	0.01810	1.25510	0.02830	0.55750	0.04880	0.83630	0.05040
0.34510	0.01370	0.57590	0.02700	1.28260	0.03460	1.03290	0.02550	1.66870	0.01280
0.40060	0.04010	0.34790	0.05050	1.44450	0.03120	1.19960	0.02300	1.25220	0.05760
0.12760	0.03130	0.43450	0.00530	1.63410	0.02500	1.02670	0.01820	1.15580	0.02430
0.27120	0.03120	0.22740	0.02340	0.48420	0.00460	0.94180	0.02230	1.68570	0.04190
0.39440	0.01200	0.62170	0.01220	0.91660	0.00260	0.96150	0.02950	0.71970	0.01710
0.37460	0.02550	0.96020	0.04770	1.41780	0.00470	0.93040	0.03490	0.82650	0.05420
0.33520	0.00780	1.27150	0.01500	0.83700	0.03250	1.06170	0.00960	0.78920	0.03710
0.33730	0.01650	1.38680	0.01400	0.81690	0.03220	1.01460	0.03770	1.63370	0.04080

| 0.10550 | 0.01180 | 1.36900 | 0.02170 | 1.02110 | 0.03130 | 1.37030 | 0.00500 | 1.61810 | 0.03040 |
| 0.75520 | 0.01960 | 1.45620 | 0.03940 | 0.65070 | 0.03070 | 1.21480 | 0.02310 | 1.55800 | 0.00540 |

The maximum radial distance (MRD) using equation (6.5) is given by

$MRD_1 = 5.025$, $MRD_2 = 0.0707$.

The estimated parameters of the fitted Burr distributions are obtained using simulated annealing and hybrid (simulated annealing and direct search) methods and are displayed in Table 6.9.

Table 6.9: Burr distribution parameter (c, k) estimation

GD	Parameter estimation using Hybrid		Parameter estimation using SA	
	c	k	c	k
GD_1	1.9593	1.47978	1.7794	46.9234
GD_2	2.1155	1530.476	1.5819	236.5164

The comparison of the PNC values obtained using numerical method vs. exact PNC value presented in Table 6.10.

Table 6.10: PNC for Geometric Distance Data

GD	Prob. of the product conforming using Eq.(6.8)		Total Prob. of the product conforming using Eq.(6.10)		Total PNC using Eq.(6.9)		Actual PNC Value
	Hybrid	SA	Hybrid	SA	Hybrid	SA	0.010

GD$_1$	0.9913	0.9999	0.9877	0.9713	0.0123	0.0287
GD$_2$	0.9964	0.9714				

Using the PNC criterion, results in Table 6.10 shows that the PNC obtained by using hybrid search approach (which is the combination of Simulated Annealing and Compass Direct Search) is relatively closer to the actual PNC in comparison with PNC obtained using the simulated annealing approach. This is again a proof that using a numerical search algorithm and fitting Burr distribution to geometric variables we can achieve comparable results instead of a traditional approach i.e. fitting different distributions to different sets of geometric distance variables (Wang 2006).

6.4.3. Example 3: Using seven quality characteristics

Here we consider seven (7) quality characteristics $\{X_1, X_2, ..., X_7\}$ from the same multivariate connector data (Wang 2006) and use numerical search (Simulated Annealing (SA) and Evolutionary Algorithms (EA)) methods to estimate fitted Burr distribution parameters to all geometric variables instead of fitting different distributions to different geometric variables as proposed by Wang (Wang 2006). The full data is given in Appendix A1.

The specification limits for all seven quality characteristics are as 0.10 ± 0.04 mm, 0 + 0.50 mm, 11 ±5 mΩ, 0 + 0.2 mm, 0.55 ± 0.06 mm, 0.07 ± 0.05 mm and 0.07 ± 0.05 mm, respectively. Using a statistical package we found that variables $\{X_1, X_2, X_3\}$ are correlated; variables X_4 and X_5 are uncorrelated with other variables, and variables $\{X_6, X_7\}$ are correlated.

Using the geometric distance approach equation (6.4), four new univariate variables are given in Table 6.11 below:

Table 6.11: Geometric distance variable data

GD1	GD2	GD3	GD4	GD1	GD2	GD3	GD4
0.226700	0.009000	0.003600	0.012900	1.282600	0.007700	0.029200	0.034600
0.164900	0.009100	0.021000	0.035000	1.444500	0.009300	0.009800	0.031200
0.093000	0.015300	0.005600	0.039200	1.634100	0.021900	0.022900	0.025000
0.428900	0.021600	0.024700	0.020800	0.484200	0.008100	0.032200	0.004600
0.294900	0.008800	0.012100	0.019900	0.916600	0.005600	0.032700	0.002600
0.293100	0.022600	0.030400	0.011300	1.417800	0.035100	0.002700	0.004700
0.519900	0.010400	0.032500	0.018600	0.837000	0.055300	0.022200	0.032500
0.222700	0.006400	0.015300	0.015000	0.816900	0.004800	0.029600	0.032200
0.160700	0.078200	0.025500	0.018400	1.021100	0.028700	0.013900	0.031300
0.232900	0.031000	0.025800	0.031000	0.650700	0.002400	0.013900	0.030700
0.345100	0.026500	0.017900	0.013700	0.896900	0.062200	0.006500	0.009100
0.400600	0.056400	0.028600	0.040100	0.789500	0.049500	0.000100	0.029500
0.127600	0.006900	0.004900	0.031300	1.704200	0.021700	0.025300	0.051900
0.271200	0.116200	0.019800	0.031200	1.535400	0.032100	0.021900	0.022200
0.394400	0.034100	0.017400	0.012000	0.883100	0.026800	0.033400	0.046200
0.374600	0.067100	0.019000	0.025500	1.183200	0.016000	0.008300	0.049200
0.335200	0.020800	0.012200	0.007800	0.918100	0.025900	0.030800	0.020200
0.337300	0.001900	0.036900	0.016500	1.350200	0.028100	0.006700	0.034300

0.105500	0.034300	0.021900	0.011800	1.121000	0.055100	0.001400	0.038800
0.755200	0.063100	0.003200	0.019600	0.557500	0.017300	0.005900	0.048800
0.226100	0.075000	0.004000	0.008400	1.032900	0.005500	0.031100	0.025500
0.098500	0.013800	0.027500	0.033800	1.199600	0.009700	0.010500	0.023000
0.196800	0.002100	0.002800	0.044900	1.026700	0.037200	0.003100	0.018200
3.604200	0.040600	0.028800	0.027100	0.941800	0.020500	0.006200	0.022300
4.218200	0.011100	0.017800	0.008600	0.961500	0.025800	0.032300	0.029500
3.747300	0.018000	0.016700	0.039600	0.930400	0.038300	0.011600	0.034900
0.440800	0.019500	0.023200	0.033700	1.061700	0.017200	0.009000	0.009600
4.489600	0.007300	0.054700	0.050700	1.014600	0.011700	0.013600	0.037700
4.119900	0.029300	0.019900	0.022600	1.370300	0.002100	0.001700	0.005000
0.220900	0.003400	0.023100	0.018100	1.214800	0.005500	0.000700	0.023100
0.575900	0.040800	0.030200	0.027000	1.245800	0.072800	0.008100	0.009200
0.347900	0.021100	0.039300	0.050500	0.332400	0.014900	0.008300	0.027600
0.434500	0.045100	0.024000	0.005300	0.670600	0.000300	0.015300	0.030200
0.227400	0.003200	0.014700	0.023400	0.359200	0.011600	0.016400	0.027300
0.621700	0.013700	0.016900	0.012200	0.758500	0.010200	0.008100	0.022800
0.960200	0.074500	0.010700	0.047700	1.001900	0.059200	0.001400	0.019100
1.271500	0.025300	0.020600	0.015000	0.226000	0.005600	0.015900	0.052600
1.386800	0.067400	0.024700	0.014000	0.719500	0.022800	0.009900	0.016000
1.369000	0.010400	0.025500	0.021700	0.735600	0.011000	0.020000	0.031300
1.456200	0.002400	0.028800	0.039400	0.836300	0.040300	0.026100	0.050400
1.284100	0.000300	0.019500	0.041300	1.668700	0.044400	0.009300	0.012800
0.840900	0.047600	0.016600	0.050800	1.252200	0.042900	0.002700	0.057600
1.229400	0.021500	0.029700	0.041100	1.155800	0.077800	0.021900	0.024300
1.513200	0.030900	0.027600	0.053400	1.685700	0.006900	0.013200	0.041900
1.080300	0.088300	0.019600	0.024600	0.719700	0.024700	0.015400	0.017100
1.026100	0.012900	0.023800	0.021100	0.826500	0.043100	0.004500	0.054200
0.827300	0.038200	0.014800	0.039400	0.789200	0.057400	0.010600	0.037100
1.387100	0.034900	0.041100	0.046600	1.633700	0.051600	0.013400	0.040800
1.342100	0.016600	0.006100	0.030900	1.618100	0.025500	0.005400	0.030400
1.255100	0.030000	0.037000	0.028300	1.558000	0.084500	0.010200	0.005400

The maximum radial distances (MRDs) using equation (6.6) are

$\text{MRD}_1 = 5.025$, $\text{MRD}_2 = 0.2$, $\text{MRD}_3 = 0.06$, $\text{MRD}_4 = 0.0707$. Also, from

the correlation matrix of these four geometric distance variables, we

found that these variables do not have any significant correlation

with each other. Burr distribution is fitted to the geometric data with

the estimated parameters displayed in Table 6.12. The comparison of

the PNC values obtained using numerical method vs. exact PNC value

presented in Table 6.13.

Table 6.12: Burr distribution parameter (c, k) estimation

GD Variables	Burr XII distribution parameter estimation			
	Simulated Annealing (SA)		Evolutionary Algorithm (EA)	
	c	k	c	k
GD_1	1.8869	1.4868	1.0003492	121.34626
GD_2	1.203	67.5758	1.007604	127.91319
GD_3	1.2009	122.114	1.0023824	80.130296
GD_4	1.5819	236.5164	1.0001383	106.44024

Table 6.13: PNC for geometric distance data

GD Variables	Prob. of the product conforming using Equation (6.7)			Total Prob. of the product conforming using Equation (6.10)			Total Prob. of the product non-conforming (PNC) using Equation (6.9)		
	Best-fit	SA	EA	Best-fit	SA	EA	Best-fit	SA	EA
GD_1	1.0000	0.9999	1.0000						
GD_2	0.6123	0.9999	1.0000	0.3628	0.9454	0.9896	0.6372	0.0546	0.0104
GD_3	0.9999	0.9833	0.9903						
GD_4	0.5932	0.9714	0.9993						

Results in Table 6.13 shows that the PNC value obtained using the

proposed Burr-fit using SA & EA method is relatively closer to the

actual PNC (i.e. 0.01). As we know, here the actual PNC presents the

actual proportion of nonconforming data that fall outside their

respective specifications. Results further show that GD_3 (representing

X_5) and GD_4 (representing X_6 & X_7) have significantly larger PNC

value, and thus indicating that quality characteristics X_5, X_6 and X_7 are candidates for improvement first rather than other 2 quality characteristics.

All three examples mentioned in this section clearly indicate that instead of fitting different distributions to geometric variable data (a conventional approach cited in the research literature), it is sufficient to deploy only one distribution, the Burr distribution. The results also indicate that one can increase the accuracy of determining the probability of non-conforming items by using different numerical techniques to estimate the fitted Burr distribution parameters.

In the proceeding section, we propose a multivariate capability index based on the Generalized Covariance Distance (GCD) which is easy to use. Another novelty introduced in this section is to approximate the distribution of these distances by Burr distribution (as mentioned in section 6.4) where its parameters are estimated using a Simulated Annealing (SA) and Evolutionary Algorithms. An example, based on real manufacturing process data, is also given which demonstrates that the proportion of nonconformance (PNC) using proposed method is very close to the actual PNC value.

6.5. A new proposed methodology to evaluate multivariate process capability

A novel approach called "generalized covariance distance (GCD)" is presented in this section (Ahmad et al. 2009). According to this approach we first cluster correlated quality characteristics. Then we define a new variable referred to as the "covariance distance (CD)" variable that takes into account the distance of individual quality characteristics from their respective specifications scaled by their variance covariance matrix. The proposed approach is similar to the geometric distance approach adopted by Wang (Wang 2006) but differs insofar as there, the scaling effect of the variance–covariance matrix is absent. Furthermore, unlike the approach adopted in (Wang 2006), we fit Burr distribution using simulated annealing algorithm to the covariance distance data instead of fitting different distributions to the geometric data.

Another novelty to our proposed method is that it is not restricted to normal multivariate data as is commonly assumed by many traditional multivariate PCI methods. The proposed method resembles the linear discriminant classification approach, commonly adopted in multivariate analysis, which group correlated or uncorrelated quality characteristics into homogeneous subgroups based on a linear discriminant function. In practice, the fundamental

objective of process capability analysis is to help engineers and managers decide whether to accept or reject the process outcomes based on conformance to engineering specifications ((Bernardo & Irony 1996), (Singpurwalla 1992)) Keeping in mind this objective, the efficacy of the proposed method is assessed by using the proportion of nonconformance (PNC) criterion to evaluate the performance of each covariance distance variable using the proposed methodology.

6.5.1. Generalized Covariance Distance (GCD) Approach

Although several methods have been proposed to handle the issues of non-normality for univariate quality characteristic data in PCI studies, there has not been much research on this aspect devoted to multivariate non-normal data. In this section, we will discuss our proposed methodology based on generalized covariance distance (GCD) variable.

Our approach is closely related to the Geometric Distance (GD) approach (refer to section 6.3) proposed by Wang and Hubele (Wang & Hubele 1999), which reduces the dimension of the multivariate process data and render them more tractable for a statistical

analysis. The GD approach utilizes the Euclidean distance (or *L2* norm) which is defined as follows:

let $X = (x_1, x_2, \ldots \ldots, x_n)$ represents a point from a sample space and let $T = (t_1, t_2, \ldots \ldots, t_n)$ be the corresponding target value. Then the Geometric Distance (GD) variable Equation (6.5) is defined by

$$GD = \sqrt{(X - T)'(X - T)}$$

$$= \sqrt{(x_1 - t_1)^2 + (x_1 - t_1)^2 + \ldots + (x_n - t_n)^2}$$

A comprehensive study of the distribution of GD when the underlying variables have a multivariate normal distribution was undertaken in (Wang & Hubele 1999). When the underlying distribution is non-normal, Wang (Wang 2006) combined correlated quality characteristics to form GD and determined the distribution that best fit GD by using Best-Fit statistical software.

Instead of *GD*, we propose using the Generalized Covariance Distance (GCD) defined by

$$CD = \sqrt{(X - T)' \Sigma^{-1} (X - T)} \qquad (6.11)$$

where Σ refers to the variance-covariance matrix of X. The choice of *CD* is motivated by the following result whose proof easily follows from the Maximization Lemma (Johnson, R & Wichern 2007).

$$\max_{a \neq 0} \frac{((X - T)' a)^2}{a' \Sigma\, a} = (X - T)' \Sigma^{-1} (X - T) \quad (6.12)$$

where the maximum is achieved when $a = c\Sigma^{-1}(X - T)$ for any $c \neq 0$.

Therefore, CD^2 provides the maximum of the ratio of the squared weighted distance between the point X and its target T and its variance. The Geometric Distance corresponds to the case when $\Sigma = I$, the identity matrix.

We present the Maximization Lemma (c.f. p.80 of Johnson and Wichern, 2007) and its proof below:

Let B be positive definite matrix of order (pxp) and d be a given vector of dimension p . Then, for an arbitrary nonzero vector X,

$$\max_{a \neq 0} \frac{(x'd)^2}{x' B\, x} = (d)' B^{-1} d \qquad (6.13)$$

With the maximum attained when $X = CB^{-1}d$ for any constant $c \neq 0$.

By the extended Cauchy=Schwarz inequality, $(x'd)^2 \leq (x'Bx)(d'B^{-1}d)$. Because $x \neq 0$ and B is positive definite, $(x'Bx) > 0$. Dividing both sides of the inequality by the positive scalar $(x'Bx)$ yields the upper

bound $\quad \dfrac{(x'd)^2}{x'Bx} \leq d'B^{-1}d$

Taking the maximum over x gives equation (6.11) because the bound is attained for $x = c\ \mathrm{B}^{-1}d$

To illustrate the calculation of Generalized Covariance Distance variable "CD", consider the values of X_1, X_2 and X_3 to be 0.1165, 0.0614 and 10.7824 respectively. Assume that the specification limits for these quality characteristics are $0.10 \pm 0.04\ mm$., $0 + 0.50 mm$. and $11 \pm 5\ mm$ respectively. Here 0.04, 0.5 and 5 are referred to as tolerances of X_1, X_2 and X_3 respectively.

Therefore

$$X - T = (0.1165 - 0.10,\ 0.0614 - 0,\ 10.7824 - 11)$$

and

$$\Sigma^{-1} = \begin{pmatrix} 118.502 & -5.565 & 0.036 \\ -5.565 & 9.314 & 0.002 \\ 0.036 & 0.002 & 0.006 \end{pmatrix}$$

resulting in $CD = 0.2368$.

The maximum radial distance (MRD) (refer equation (6.6)) used as upper specification of the generalized covariance distance variables is the distance from the target to a perimeter of the tolerance region.

Using MRD as upper bound, the estimated proportion of non-conformance for each CD variable is estimated using equation (6.7)

The flowchart of the proposed methodology is presented in Table 6.14.

Table 6.14: A flowchart of the proposed methodology

1. Identify the desired quality characteristics (multiple) along with their respective engineering specifications.

2. Collect measurements of these quality characteristics data from a manufacturing process.

3. Determine the correlated and uncorrelated quality characteristics using statistical software.

4. Compute generalized covariance distance variables CD's for correlated and uncorrelated quality characteristics using equation (6.11).

5. Compute maximum radial distances *(MRD(s))* from the target value using equation (6.6).

6.   Determine covariance distance variables that do not have any significant correlation with each other (using any statistical software).

7. Use SA and EA search algorithms to find out parameters of the fitted univariate Burr distribution to each covariance distance variable (refer to section 3 of this Chapter).

8. Compute the proportion of conforming items for each covariance distance variable (using equation (6.8)).

9. Compute the proportion of nonconformance *(PNC)* for each covariance distance variable (using equation (6.7)).

10. Compute total proportion of nonconformance *(PNC)* value using equation (6.9).

As mentioned earlier in this section that we fit Burr distribution function f(x) to CD variables. To fit the appropriate Burr distribution, we need to estimate the parameters c and k. We will use the method of Maximum Likelihood Estimation (MLE) to estimate these parameters. Simulated Annealing and Evolutionary Algorithms are used to estimate Burr distribution parameters (refer to Chapter 3 for details).

6.6. A manufacturing example

In this section, we demonstrate the proposed methodology using real data from Wang (Wang 2006). The data set is from a manufacturing process with multivariate quality characteristics. It contains a sample of 100 parts that were tested on seven quality characteristics $\{X_1, X_2, \ldots, X_7\}$ of interest to the manufacturer. The full data set is given in Appendix A1. The specification limits for these seven quality characteristics can be two-sided or one-sided, and they are 0.10 ±

0.04 mm, 0 + 0.50 mm, 11 ±5 mm, 0 + 0.2 mm, 0.55 ± 0.06 mm, 0.07 ± 0.05 mm and 0.07 ± 0.05 mm, for X_1 to X_7 respectively. Based on quality characteristics and the manufacturing processes, it was found that the variables $\{X_1, X_2, X_3\}$ are correlated; variables X_4 and X_5 are uncorrelated with other variables, and variables $\{X_6, X_7\}$ are correlated.

Using the GCD approach, the following four new univariate variables are defined by

$CD_1 = (X - T) \, \Sigma^{-1} \, (X - T)'$, where T is the target distance and Σ^{-1} is inverse of variance-covariance matrix and for

$CD_1 : (X - T) = \{(X_1 - 0.1), (X_2 - 0), (X_3 - 11)\}$

$CD_2 = (X_4)$,

$CD_3 = (X_5 - 0.55)$ and $CD_4 = (X - T) \, \Sigma^{-1} \, (X - T)'$ where   for

$CD_4 : (X - T) = \{(X_6 - 0.07), (X_7 - 0.07)\}$ respectively.

The values of the covariance variables CD_1, CD_2, CD_3 and CD_4 are tabulated in Table 6.15.

Table 6.15 : Covariance distance variable data

CD_1	CD_2	CD_3	CD_4	CD_1	CD_2	CD_3	CD_4
0.476130	0.378290	0.329800	0.113580	1.132520	0.323650	2.675030	0.186010
0.406080	0.382500	1.923830	0.187080	1.201870	0.390900	0.897790	0.176640

0.304960	0.643100	0.513020	0.197990	1.278320	0.920510	2.097890	0.158110
0.654900	0.907900	2.262790	0.144220	0.695840	0.340460	2.949870	0.067820
0.543050	0.369890	1.108490	0.141070	0.957390	0.235380	2.995670	0.050990
0.541390	0.949940	2.784970	0.106300	1.190710	1.475350	0.247350	0.068560
0.721040	0.437140	2.977350	0.136380	0.914880	2.324400	2.033760	0.180280
0.471910	0.269010	1.401640	0.122470	0.903830	0.201760	2.711680	0.179440
0.400870	3.286950	2.336070	0.135650	1.010490	1.206340	1.273390	0.176920
0.482600	1.303010	2.363560	0.176070	0.806660	0.100880	1.273390	0.175210
0.587450	1.113860	1.639830	0.117050	0.947050	2.614430	0.595470	0.095390
0.632930	2.370640	2.620070	0.200250	0.888540	2.080620	0.009160	0.171760
0.357210	0.290030	0.448890	0.176920	1.305450	0.912110	2.317750	0.227820
0.520770	4.884190	1.813890	0.176640	1.239110	1.349250	2.006280	0.149000
0.628010	1.433310	1.594030	0.109540	0.939730	1.126470	3.059800	0.214940
0.612050	2.820390	1.740600	0.159690	1.087750	0.672520	0.760370	0.221810
0.578960	0.874280	1.117650	0.088320	0.958180	1.088650	2.821610	0.142130
0.580780	0.079860	3.380440	0.128450	1.161980	1.181120	0.613790	0.185200
0.324810	1.441720	2.006280	0.108630	1.058770	2.316000	0.128260	0.196980
0.869020	2.652260	0.293150	0.140000	0.746660	0.727160	0.540500	0.220910
0.475500	3.152450	0.366440	0.091650	1.016320	0.231180	2.849090	0.159690
0.313850	0.580050	2.519300	0.183850	1.095260	0.407720	0.961910	0.151660
0.443620	0.088270	0.256510	0.211900	1.013260	1.563610	0.283990	0.134910
1.898470	1.706520	2.638390	0.164620	0.970460	0.861670	0.567990	0.149330
2.053830	0.466560	1.630670	0.092740	0.980560	1.084440	2.959030	0.171760
1.935790	0.756590	1.529900	0.199000	0.964570	1.609850	1.062680	0.186820
0.663930	0.819640	2.125370	0.183580	1.030390	0.722960	0.824500	0.097980
2.118870	0.306840	5.011110	0.225170	1.007270	0.491780	1.245910	0.194160
2.029750	1.231560	1.823050	0.150330	1.170600	0.088270	0.155740	0.070710
0.470000	0.142910	2.116210	0.134540	1.102180	0.231180	0.064130	0.151990
0.758880	1.714930	2.766650	0.164320	1.116150	3.059980	0.742050	0.095920
0.589830	0.886890	3.600300	0.224720	0.576540	0.626290	0.760370	0.166130
0.659170	1.895670	2.198660	0.072800	0.818900	0.012610	1.401640	0.173780
0.476860	0.134500	1.346680	0.152970	0.599330	0.487580	1.502420	0.165230
0.788480	0.575850	1.548220	0.110450	0.870920	0.428730	0.742050	0.151000
0.979900	3.131430	0.980240	0.218400	1.000950	2.488330	0.128260	0.138200
1.127610	1.063430	1.887180	0.122470	0.475390	0.235380	1.456610	0.229350
1.177620	2.833000	2.262790	0.118320	0.848230	0.958340	0.906950	0.126490
1.170040	0.437140	2.336070	0.147310	0.857670	0.462360	1.832220	0.176920
1.206730	0.100880	2.638390	0.198490	0.914490	1.693910	2.391040	0.224500
1.133180	0.012610	1.786410	0.203220	1.291780	1.866250	0.851980	0.113140
0.917010	2.000750	1.520740	0.225390	1.119020	1.803200	0.247350	0.240000
1.108780	0.903700	2.720840	0.202730	1.075080	3.270140	2.006280	0.155880
1.230120	1.298810	2.528460	0.231080	1.298350	0.290030	1.209260	0.204690
1.039370	3.711480	1.795570	0.156840	0.848350	1.038210	1.410810	0.130770
1.012970	0.542220	2.180340	0.145260	0.909120	1.811610	0.412250	0.232810
0.909560	1.605650	1.355840	0.198490	0.888370	2.412670	0.971070	0.192610
1.177750	1.466940	3.765200	0.215870	1.278160	2.168880	1.227580	0.201990
1.158490	0.697740	0.558830	0.175780	1.272050	1.071830	0.494700	0.174360

1.120310   1.260980   3.389600   0.168230   1.248200   3.551760   0.934430   0.073480

The maximum radial distances are $MRD_1 = 5.025$, $MRD_2 = 0.2$, $MRD_3 = 0.06$, and $MRD_4 = 0.0707$ respectively. Also, from the correlation matrix of these four covariance distance variables, we found that these variables do not have any significant correlation with each other. Burr distribution is fitted to the covariance distance data and the probability of the product nonconforming items for each CD variable is calculated using equation (6.7). The estimated parameters of the fitted Burr distribution are displayed in Table 6.16.

Table 6.16 : Burr distribution parameter (c, k) estimation

CD Variables	Burr XII distribution parameter estimation			
	Simulated Annealing (SA)		Evolutionary Algorithm (EA)	
	c	k	c	k
CD_1	3.4041	41.1290	1.0027	317.6904
CD_2	1.2030	67.5758	1.0032	371.5740
CD_3	1.2009	122.1140	1.6162	398.4719
CD_4	1.1003	202.8647	1.0074	307.9073

Table 6.17: PNC for covariance distance data

CD Variables	Prob. of the product conforming using Equation (6.8)			Total Prob. of the product conforming using Equation (6.10)			Total Prob. of the product non-conforming (PNC) using Equation (6.9)		
	GD (Best-fit)	GCD (SA)	GCD (EA)	GD (Best-fit)	GCD (SA)	GCD (EA)	GD (Best-fit)	GCD (SA)	GCD (EA)
CD_1	1.0000	1.0000	1.0000						
CD_2	0.6123	0.9999	1.0000	0.3628	0.9830	0.9850	0.6372	0.0170	0.0150
CD_3	0.9999	0.9833	0.9850						
CD_4	0.5932	1.0000	1.0000						

The results in Table 6.17 shows that the proposed GCD method yields a much more accurate probability of the product non-conforming (PNC) (compared with the true proportion of nonconforming items falling outside their respective specifications in the real data i.e. 0.01 than the GD method (0.68%, 0.50% and 62.72% relative percentage difference respectively) which points to the superiority of the GCD method over the more conventional GD approach. Results further show that CD_3 has significantly larger PNC value (i.e. 1-0.9850=0.015); consequently quality characteristics X_5, is the first candidate for improvement in comparison with the other four quality characteristics.

6.7. Conclusion

This chapter discusses evaluation of process capability for correlated multivariate non-normal quality characteristics using Geometric Distance (GD) and proposed Generalized Covariance Distance (GCD)

approaches. A detailed discussion with real data examples has been devoted to fitting Burr distribution to GD and GCD data. Different numerical techniques have been deployed to estimate fitted Burr distribution parameters. Fitting Burr distribution to GD and GCD data has yielded comparable probability of the product non-conforming (PNC) results as compared to those achieved using Best-fit method cited in the research literature (Wang 2006).

Unlike the Geometric Distance (GD) approach, our proposed GCD approach takes into account the scaling effect of variance-covariance matrix while reducing the dimension of multivariate data, which then enables univariate statistical analysis to be performed on the generalized covariance distance variables. It is also demonstrated that the proposed GCD approach does not assume that the CD variables are mutually independent, which is implicitly assumed in the Geometric Distance approach. The parameters of the fitted Burr distribution are estimated using different numerical techniques. This approach contrasts with that adopted in Wang (Wang 2006), where different distributions are fitted to different sets of geometric distance data. It also resulted in PNC values which are much closer to the exact values.

Chapter 7

CONCLUSIONS AND RECOMMENDATIONS

7.1. Conclusions

This thesis has investigated several aspects of process capability indices C_p and C_{pk} when the underlying distribution of quality characteristics is non-normal. Initially, we reviewed the performance of the existing capability estimation methods (e.g. Clements percentile method, Burr based percentile method and Box Cox method) for non-normal univariate quality characteristics data. A simulation study using known non-normal distributions along with experimental data were then conducted to compare and contrast the performance of some of the commonly used methods. Burr percentile method yielded better results when compared to other methods (e.g. Clements and Box-Cox methods). Accuracy of the Burr based percentile method was further improved using several state of the art numerical techniques. A detailed discussion of these numerical techniques has been presented in Chapter 3.

Application of Burr distribution was then further explored to estimate PCI(s) of non-normal univariate quality characteristic data. New capability estimation methods, such as Cumulative Density Function

(CDF) and Best Root Transformation (BRT) methods, were evaluated by fitting Burr distribution to simulated and experimental data in Chapter 4. In Chapter 5, we've further extended Cumulative Density Function method to bivariate process capability estimation by fitting bivariate Burr distribution to bivariate non-normal quality characteristics data.

One of the major objectives of research in this thesis is presented in Chapter 6. Here, a novel approach to estimate multivariate non-normal PCI was introduced and implemented. This approach, called "Generalized Covariance Distance (GCD)" approach, evaluates process capability for correlated non-normal multivariate quality characteristics data. GCD approach has several novel features such as: 1) It is based on the idea of reducing the dimension of multivariate data by transforming correlated variables into univariate ones through a metric function). Unlike the Geometric Distance (GD) approach cited in the research literature, our proposed approach takes into account the scaling effect of the variance–covariance matrix and produces a CD variable that is based on the Mahanalobis distance. 3) It is demonstrated that the proposed GCD approach does not assume that the CD variables are mutually independent, which is implicitly assumed in the Geometric Distance approach. 4) In contrast to the GD approach, where different distributions are fitted to different GD variables, a single distribution, the Burr XII

distribution is fitted to the CD data. Several numerical search techniques are then applied to estimate the parameters of the Burr distribution. Finally, several application examples using real data with several non-normal quality characteristics from the manufacturing industry are presented in Chapter 6 which serve to illustrate the theory presented.

7.2. Recommendations and future work

Quantitative measure of process performance for multivariate quality characteristics is of great interest to quality control practitioners and has a huge potential of expanding its application to other multivariate industrial quality research areas. Hence, there is much scope in extending the present work. For example, instead of using univariate Burr, a multivariate Burr distribution (Takahasi 1965) could be directly employed to fit multivariate CD data. However, it is anticipated that the numerical work involved in estimating the parameters of the distribution could prove to be extremely laborious. This leads to the consideration of other search metaheuristics algorithms for their potential in estimating parameters of multivariate distributions involving the Maximum Likelihood Estimation (MLE) procedure.

Metaheuristics Algorithms (e.g. Simulated Annealing, Evolutionary Algorithm) have been successfully applied in this research and also

to many other optimization problems. However combination of these two or more metaheuristics can be made that may be better suited for estimating the parameters of the multivariate distributions. Usually, using combination of metaheuristics, component of one metaheuristic is added to another metaheuristic to enhance the performance. This type of combination strategies may compensate the disadvantages of each other.

Another possible improvement can be made using hybrid of Evolutionary Algorithm (EA) and MLE process (local search). Standard EA can find globally competitive solutions, but the EA often suffers from lack of accuracy and faces slow convergence. The complementary properties of EA and local search / deterministic search may give several advantages over either of the methods when applied alone such as improvement of the performance of the EA regarding convergence speed as well as improvement of the quality of the solutions obtained due to the incorporation of domain-specific knowledge from local search.

Even though in this thesis, we have discussed some of the distributional properties of correlated non-normal multivariate quality characteristics, however there are still essential difficulties in trying to assess the value of multivariate systems in terms of a single index. Clearly, further investigations are needed bearing in mind that

any new technique should be mapped to information that are useful to the frontline quality practitioners and engineers who are directly involves in assessing process performance.

Finally, since the approaches we have introduced in this thesis have led to significant improvement over existing methods, we recommend that the proposed methods be applied to other non-normal multivariate PCI studies for further comparisons.

Appendix A1

A real data set used in this research study (Chapter 6) is taken from Wang' paper (Wang 2006). Wang discussed a manufacturing product (called connector) from a computer industry having multivariate (seven) quality characteristics. These seven characteristics are X1 (contact gap X), X2 (contact loop Tp), X3 (LLCR), X4 (contact x Tp), X5(contact loop diameter), X6 (LTGAPY) and X7 (RTGAPY), respectively. The specification limits for these characteristics can be two-sided or one-sided, and they are 0.10 ± 0.04 mm, 0 + 0.50 mm, 11 ±5 mΩ ,0 + 0.2 mm, 0.55 ± 0.06 mm, 0.07 ± 0.05 mm and 0.07 ± 0.05 mm, respectively. The full data is given below:

X1	X2	X3	X4	X5	X6	X7
0.116500	0.061400	10.782400	0.009000	0.553600	0.064200	0.058500
0.125900	0.027700	10.839500	0.009100	0.529000	0.099400	0.088900
0.126500	0.076200	10.953800	0.015300	0.544400	0.098000	0.097400
0.118500	0.095700	10.582400	0.021600	0.574700	0.085600	0.083700
0.141400	0.131900	10.739500	0.008800	0.562100	0.080300	0.087000
0.092000	0.047600	10.710900	0.022600	0.519600	0.080400	0.074300
0.080400	0.044000	10.482400	0.010400	0.517500	0.061600	0.053400
0.110300	0.090100	10.796600	0.006400	0.565300	0.057300	0.062000
0.102200	0.091800	10.868100	0.078200	0.524500	0.065000	0.052300
0.110300	0.082300	10.782400	0.031000	0.524200	0.099200	0.080400
0.106900	0.094300	10.668100	0.026500	0.532100	0.056400	0.071800
0.107500	0.095000	10.610900	0.056400	0.521400	0.103500	0.092000
0.117000	0.117700	10.953800	0.006900	0.545100	0.096500	0.086700
0.127600	0.137800	10.768100	0.116200	0.530200	0.099700	0.079700
0.103900	0.064500	10.610900	0.034100	0.532600	0.078500	0.078500
0.125100	0.098800	10.639500	0.067100	0.531000	0.094500	0.077000
0.115300	0.155300	11.296600	0.020800	0.562200	0.072400	0.062600
0.096100	0.014200	10.965700	0.001900	0.513100	0.081000	0.057700
0.117000	0.080800	11.065700	0.034300	0.528100	0.078200	0.078500

0.115800	0.078000	10.249000	0.063100	0.553200	0.085400	0.082100
0.113400	0.151100	10.832400	0.075000	0.546000	0.067200	0.062100
0.135600	0.040700	11.082400	0.013800	0.522500	0.037700	0.060100
0.122200	0.065400	10.815700	0.002100	0.547200	0.109100	0.092100
0.112400	0.151600	7.399000	0.040600	0.578800	0.047300	0.055200
0.115500	0.069500	6.782400	0.011100	0.532200	0.076000	0.063900
0.122300	0.085300	7.253800	0.018000	0.533300	0.096700	0.099200
0.113000	0.086900	10.568100	0.019500	0.526800	0.094500	0.093200
0.113700	0.065200	6.510900	0.007300	0.495300	0.046300	0.114800
0.111200	0.135700	6.882400	0.029300	0.530100	0.075700	0.091900
0.111400	0.090900	10.799000	0.003400	0.526900	0.087900	0.072700
0.097800	0.097000	10.432400	0.040800	0.519800	0.093500	0.083200
0.120900	0.093900	10.665700	0.021100	0.510700	0.097100	0.112600
0.109200	0.041900	11.432400	0.045100	0.526000	0.067900	0.065100
0.116100	0.108800	11.199000	0.003200	0.535300	0.085600	0.087400
0.113100	0.069800	10.382400	0.013700	0.533100	0.074800	0.081200
0.119300	0.144500	11.949000	0.074500	0.539300	0.106400	0.100800
0.123300	0.090000	12.268100	0.025300	0.529400	0.068800	0.085000
0.089000	0.110200	12.382400	0.067400	0.525300	0.076300	0.082500
0.107400	0.050700	12.368100	0.010400	0.524500	0.087900	0.082200
0.094700	0.084500	12.453800	0.002400	0.521200	0.073300	0.030700
0.104800	0.066100	12.282400	0.000300	0.530500	0.100400	0.098000
0.119200	0.044800	11.839500	0.047600	0.533400	0.106600	0.105300
0.114300	0.100900	12.225200	0.021500	0.520300	0.096800	0.101200
0.110200	0.081500	12.510900	0.030900	0.522400	0.103800	0.111300
0.108400	0.161700	12.068100	0.088300	0.530400	0.085500	0.089100
0.107100	0.042300	12.025200	0.012900	0.526200	0.086100	0.083700
0.110600	0.058500	11.825200	0.038200	0.535200	0.086000	0.106000
0.105100	0.115100	12.382400	0.034900	0.508900	0.096300	0.108500
0.116700	0.082700	12.339500	0.016600	0.543900	0.093800	0.089700
0.106200	0.056600	12.253800	0.030000	0.513000	0.087500	0.092300
0.109100	0.022100	12.282400	0.007700	0.520800	0.092300	0.096500
0.113800	0.119000	12.439500	0.009300	0.540200	0.090000	0.093900

0.130900	0.069900	12.632400	0.021900	0.527100	0.083900	0.090800
0.107500	0.041700	11.482400	0.008100	0.517800	0.068300	0.074300
0.115700	0.037700	11.915700	0.005600	0.517300	0.071800	0.068100
0.119300	0.075200	12.415700	0.035100	0.552700	0.071400	0.065500
0.123900	0.084900	11.832400	0.055300	0.527800	0.055300	0.041000
0.113300	0.041800	11.815700	0.004800	0.520400	0.085800	0.098100
0.120600	0.081800	9.982400	0.028700	0.536100	0.086800	0.096400
0.122000	0.073300	10.353800	0.002400	0.536100	0.094700	0.088200
0.116100	0.117000	10.110900	0.062200	0.556500	0.066300	0.061700
0.118600	0.104400	11.782400	0.049500	0.549900	0.095600	0.084600
0.102000	0.052900	9.296600	0.021700	0.524700	0.094800	0.115600
0.111600	0.056300	9.465700	0.032100	0.528100	0.085000	0.086300
0.111100	0.035200	11.882400	0.026800	0.516600	0.104000	0.101300
0.098600	0.043700	12.182400	0.016000	0.541700	0.092600	0.113700
0.116600	0.065100	11.915700	0.025900	0.519200	0.076700	0.089100
0.118700	0.053500	12.349000	0.028100	0.543300	0.089000	0.098500
0.120600	0.107300	12.115700	0.055100	0.551400	0.092500	0.101600
0.113900	0.095600	11.549000	0.017300	0.544100	0.102700	0.106200
0.110300	0.033200	12.032400	0.005500	0.518900	0.068700	0.044500
0.108900	0.084300	12.196600	0.009700	0.539500	0.092500	0.075000
0.104300	0.055800	12.025200	0.037200	0.546900	0.071100	0.088200
0.105600	0.065100	11.939500	0.020500	0.543800	0.087300	0.084000
0.114500	0.121000	11.953800	0.025800	0.582300	0.091000	0.090700
0.127700	0.094500	11.925200	0.038300	0.538400	0.094900	0.094400
0.112000	0.128500	12.053800	0.017200	0.541000	0.072100	0.079400
0.113400	0.085400	12.010900	0.011700	0.563600	0.100100	0.092700
0.123700	0.075200	12.368100	0.002100	0.548300	0.066100	0.066900
0.127000	0.093300	12.210900	0.005500	0.549300	0.078000	0.091700
0.115400	0.124700	12.239500	0.072800	0.558100	0.073300	0.078600
0.096700	0.068500	11.325200	0.014900	0.541700	0.089100	0.089900
0.105600	0.057500	11.668100	0.000300	0.565300	0.096100	0.085200
0.111600	0.116800	11.339500	0.011600	0.566400	0.090000	0.088600
0.111300	0.083900	11.753800	0.010200	0.558100	0.082400	0.089100

0.114900	0.101900	11.996600	0.059200	0.548600	0.080300	0.086100
0.109100	0.017000	11.225200	0.005600	0.534100	0.104500	0.109700
0.112000	0.073400	11.715700	0.022800	0.540100	0.082600	0.079900
0.107400	0.068500	11.732400	0.011000	0.530000	0.078100	0.100200
0.111200	0.080700	11.832400	0.040300	0.523900	0.098200	0.111800
0.109700	0.099300	12.665700	0.044400	0.540700	0.078200	0.079800
0.108500	0.088200	12.249000	0.042900	0.547300	0.117200	0.103100
0.100900	0.125200	12.149000	0.077800	0.528100	0.072100	0.094200
0.107100	0.068600	9.315700	0.006900	0.563200	0.103000	0.095800
0.115700	0.051800	10.282400	0.024700	0.534600	0.086500	0.074600
0.112700	0.120300	10.182400	0.043100	0.554500	0.101600	0.114000
0.112400	0.102900	11.782400	0.057400	0.539400	0.067100	0.107000
0.101700	0.075300	9.368100	0.051600	0.563400	0.087300	0.106900
0.121000	0.029900	9.382400	0.025500	0.555400	0.064700	0.099900
0.110700	0.167100	12.549000	0.084500	0.539800	0.065600	0.066900

Bibliography

Abbasi, B, Eshragh, J, Arkat, J & Hosseinkouchack, M 2006, 'Estimating the parameters of Weibull distribution using simulated annealing algorithm', *Applied Mathematics and Computation*, vol. 183, pp. 85-93.

Ahmad, S, Abdollahian, M & Zeephongsekul, P 2007a, 'Process capability analysis for non-quality characteristics using Gamma distribution', paper presented to 4th international conference on information technology - new generations, Las Vegas, USA.

Ahmad, S, Abdollahian, M & Zeephongsekul, P 2007b, 'Process Capability Analysis under a Weibull Shock Model', paper presented to 13th ISSAT International Conference on Reliability and Quality in Design, Seattle, Washington.

Ahmad, S, Abdollahian, M & Zeephongsekul, P 2008, 'Process Capability Estimation for non-normal quality characteristics using Clement, Burr and Box-Cox methods', *The Australian and New Zealand Industrial and Applied Mathematics Journal*, pp. 642-665.

Ahmad, S, Abdollahian, M, Zeephongsekul, P & Abbasi, B 2008, 'Measuring Process Performance for Non-Normal Quality Characteristics Data', *Ubiquitous Computing and Communication Journal*, vol. 3, pp. 8-12.

Ahmad, S, Abdollahian, M, Zeephongsekul, P & Abbasi, B 2009, 'Multivariate Non-normal Process Capability Analysis', *International Journal of Advanced Manufacturing Technology*, vol. 44, pp. 757-765.

Albing, M 2006, 'Process capability analysis with Focus on indices for one-sided specification limits', Lulea University of Technology.

Applegate, D, Cook, W & Rohe, A 2003, 'Chained linkernighan for large traveling salesman problems', *Informs Journal On Computing*, vol. 15, no. 1, pp. 82-92.

Ashraf, M 2009, 'Prediction of remaining useful life of a polymer-polymer sliding contact through simulation of contact wear', University of South Australia.

Bai, D & Choi, S 1997, 'Process capability indices for skewed populations', Advanced Institute of Science and Technology, Taejon, South Korea.

Baker, E 1987, 'Reducing bias and inefficiency in the selection algorithm', paper presented to Proceedings of the Second International Conference on Genetic Algorithms.

Batti, R & Protasi, M 2001, 'Reactive local search for the maximum clique problem', *Algorithmica*, vol. 29, no. 4, pp. 610-637.

Baum, L & Sell, G 1968, 'Growth function for transformations on manifold', *Pacific Journal of Mathematics*, vol. 27, no. 2, pp. 211-227.

Beck, C & Ester, S 1998, 'Multicriteria capability index', *Spektrum*, vol. 24, pp. 179-187.

Bentley, P & Corne, D 2002, *Creative Evolutionary Systems*, Morgan Kaufmann Academic Press.

Bernardo, J & Irony, T 1996, 'A general multivariate process capability index', *The Statistician*, vol. 45, no. 4, pp. 487-502.

Box, G & Cox, D 1964, 'An analysis of transformation', *Journal of royal statistics society*, vol. 26, pp. 211-243.

Boyles, R 1994, 'Process capability with asymmetric tolerances', *Communications in Statistics – Simulation and Computation*, vol. 23, no. 3, pp. 615-643.

Braun, L 2001, 'New Methods in Multivariate Statistical Process Control (MSPC)', *Diskussionsbeiträge des Fachgebietes Unternehmensforschung*, pp. 1-12.

Burr, I 1942, 'Cumulative frequency functions', *The Annals of Mathematical Statistics*, vol. 13, no. 2, pp. 215-232.

Burr, I 1967, 'The effect of non-normality on constants for and R charts', *Industrial quality control*, pp. 563-569.

Burr, I 1973, 'Parameters for a general system of distributions to match a grid of á3 and á4', *Communications in Statistics*, vol. 2, no. 1, pp. 1-21.

Cairo, M & Nelson, B 1997, *Modeling and Generating Random Vectors with Arbitrary Marginal Distributions and Correlation Matrix*, Department of Industrial Engineering and Management Science, Northwestern University, Evanston.

Castagliola, P 1996, 'Evaluation of non-normal process capability indices using Burr's distributions', *Quality Engineering*, vol. 8, no. 4, pp. 587-593.

Castagliola, P & Castellanos, J 2005, 'Capability Indices Dedicated to the two quality characteristics case', *Quality Technology and Quantitative Management*, vol. 2, no. 2, pp. 201-220.

Chambers, L 2001, *The Practical Handbook of Genetic Algorithms Applications*, Chapman and Hall.

Chang, Y, Choi, I & Bai, D 2002, 'Process Capability Indices for Skewed Populations', *Quality and reliability engineering International*, vol. 17, pp. 397-406.

Chen, H 1994, 'A Multivariate Process Capability Index Over a Rectangular Solid Tolerance Zone', *Statistica Sinica*, vol. 4, pp. 749-758.

Chen, K, Hsu, C & Wu, C 2006, 'Process Capability Analysis for A Multi-process Product', *International Journal of Advanced Manufacturing Technology*, vol. 27, pp. 1235-1241.

Choi, I & Bai, D 1996, 'Process capability indices for skewed populations', paper presented to 20th International Conference on Computer and Industrial Engineering.

Chou, C & Cheng, P 1997, 'Ranges control chart for non-normal data', *J Chinese Inst Ind Eng*, vol. 14, no. 4, pp. 401-409.

Clements, J 1989, 'Process capability calculations for non-normal distributions', *Quality progress*, vol. 22, pp. 95-100.

Darwin, C 1859, *On the origin of Species by means of natural selection or the preservation of favoured races in the Struggle for life*, J. Murray, London.

Davis, R, Kaminsky, F & Saboo, S 1992, 'Process capability analysis for processes with either a circular or a spherical tolerance zone', *Quality Engineering*, vol. 5, pp. 41-54.

Deleryd, M 1996, 'Process capability studies in Theory and Practice ', Lulea University of Technology.

Deleryd, M 1998, 'On the Gap Between Theory and Practice of Process Capability Studies', *Journal of Quality & Reliability Management*, vol. 15, pp. 178-191.

Deleryd, M & Vännman, K 1999, 'Process Capability Plots—A Quality Improvement Tool', *Quality and reliability engineering International*, vol. 15, pp. 1-15.

Devroye, L, Györfi, L & Lugosi, G 1996, *A Probabilistic Theory of Pattern Recognition (Stochastic Modeling and Applied Probability)*, Springer, New York.

Ding, J 2004, 'A Method of Estimating the Process Capability Index from the First Four Moments of Non-normal Data', *Quality and reliability engineering International*, vol. 20, pp. 787-805.

Dueck, G & Scheuer, T 1990, 'Threshold accepting: A general purpose optimization algorithm superior to simulated annealing', *Journal of computational physics*, vol. 90, pp. 161-175.

Durling, F 1975, 'The bivariate Burr distribution', *Statistical Distributions in Scientific Work*, vol. 1, pp. 329-335.

English, J & Taylor, G 1993, 'Process capability analysis—a robustness study', *International Journal of Production Research*, vol. 31, pp. 1621-1635.

Fogel, D 1994, 'An introduction to simulated evolutionary optimization', *IEEE Transactions on Neural Networks*, vol. 5, no. 1, pp. 3-14.

Ganesh, K & Punniyamoorthy, M 2005, 'Optimization of continuous-time production planning using hybrid genetic algorithms-simulated

annealing', *International Journal of Advanced Manufacturing Technology*, vol. 26, pp. 148-154.

Goldberg, D 1989, *Genetic algorithm in search, optimization and machine learning.*, Addison Wesley Longman.

Goldberg, D & Deb, K 1991, 'A comparative analysis of selection schemes used in genetic algorithms', in *Foundations of Genetic Algorithms*, Morgan Kaufmann Publishers,, pp. 69-93.

Gruska, G, Mirkhani, K & Lamberson, L 1989, *Non-normal data analysis*, Applied Computer Solutions, Inc, St.Clair Shores, Michigan.

Gunter, B 1989, 'The use and abuse of Cpk, Parts 1-4', *Quality progress*, vol. 22, no. 1, 3, 5, 7, pp. 72-, 108-9, 79-80, 86-87.

Hahn, G, Hill, W, Hoerl, R & Zinkgraf, S 1999, 'The Impact of Six Sigma Improvement—A Glimpse into the Future of Statistics', *The American Statistician*, vol. 53, pp. 208-215.

Hahn, G & Shapiro, S 1967, *Statistical models in engineering By Gerald J. Hahn, Samuel S. Shapiro*, Wiley, University of Michigan.

Hellmich, M & Wolff, H 1996, 'A new approach for describing and controlling process capability for a multivariate process', paper presented to Proceedings of the ASA section on quality and productivity.

Heuvel, E & Ion, R 2003, 'Capability Indices and the Proportion of Nonconforming Items', *Quality Engineering*, vol. 15, pp. 427-439.

Higgins, A, Kozan, E & Ferreira, L 1997, 'Heuristic techniques for single line train scheduling', *Journal of heuristics*, vol. 3, pp. 42-62.

Holland, J 1975, *Adaptation in Natural and Artificial Systems*, University of Michigan Press.

Hooke, R & Jeeves, T 1961, 'Direct search solution of numerical and statistical problems', *Journal of the Association of Computing Machinery*, vol. 8, no. 2, pp. 212-229.

Hosseinifard, S, Abbasi, B, Ahmad, S & Abdollahian, M 2009, 'A Transformation Technique to Estimate Process Capability Index for Non-normal Processes', *International Journal of Advanced Manufacturing Technology*, pp. 512-517.

Huang, M, Chen, K & Hung, Y 2002, 'Integrated process capability. Analysis with an application in backlight module', *Microelectronic Reliability*, vol. 42, pp. 209–214.

Hubele, N, Shahriari, H & Cheng, C 1991, 'A Bivariate Capability Vector" in Statistics and Design in Process Control: Keeping Pace with Automated Manufacturing', in Marcel Dekker, Keats and Montgomery, pp. 299-310.

Huda, S 2009, 'Hybrid Training Approaches to Hidden Markov Model-based Acoustic Models for Automatic Speech Recognition', University of Ballarat.

Huda, S, Yearwood, J & Togneri, R 2009, 'A Constraint-based Evolutionary Learning approach to the Expectation Maximization for Optimal Estimation of the Hidden Markov Model for Speech Signal Modeling, 2009', *IEEE Transactions on Systems, Man and Cybernetics,* vol. 39, no. 1.

Hwang, S & He, R 2006, 'Improving real-parameter genetic algorithm with simulated annealing for engineering problems', *Advances in Engineering Software*, vol. 37, pp. 406-418.

Jiang, T & Yang, F 2002, 'An evolutionary Tabu search for cell image segmentation', *IEEE transactions on Systems, Man and Cybernetics: Part-B,* vol. 32, no. 5, pp. 576-578.

Johnson, D 1990, 'Local optimization and the travelling salesman problem', paper presented to Proceedings of 17th International colloquium on Automata, Language and programming.

Johnson, N 1949, 'System of frequency curves generated by methods of translation', *Biometrika*, vol. 36, pp. 149-176.

Johnson, N, Kotz, S & Pearn, W 1994, 'Flexible Process Capability Indices.' *Pakistan Journal of Statistics*, vol. 10, pp. 23-31.

Johnson, R & Wichern, D 2007, *Applied Multivariate Statistical Analysis*, 6th edn, Pearson Prentice Hall, New York.

Jong, K 1975, 'An analysis of the behavior of a class of genetic adaptive systems', University of Michigan.

Juran, J 1974, *Juran's Quality Control Handbook*, 3rd edn, McGraw-Hill, New York.

Kane, V 1986, 'Process capability indices', *Journal of Quality Technology*, vol. 18, pp. 41-52

Kennedy, J, Eberhart, R & Shi, Y 2001, *Swarm Intelligence*, Morgan Kaufmann Publishers, San Francisco, CA.

Kirkpatrick, S, Gerlatt, C & Vecchi, M 1983, 'Optimization by Simulated Annealing', *Science*, vol. 220, pp. 671-680.

Kolda, T, Lewis, R & Torczon, V 2003, 'Optimization by direct search: new perspectives on some classical and modern methods ', *Society for industrial and applied mathematics, SIAM review*, vol. 45, no. 3, pp. 385-482.

Kotz, S & Johnson, N 1993, *Process capability indices*, Chapman & Hall, New York.

Kotz, S & Johnson, N 2002, ' Process capability indices—a review, 1992–2000', *Journal of Quality Technology*, vol. 34, pp. 1-19.

Kotz, S & Lovelace, C 1998, *Process capability indices in theory and practice*, Arnold, London.

Kruskal, J 1956, 'On the shortest spanning subtree of a graph and the travelling sales man problem', paper presented to Proceedings of the American mathematical societies.

Li, Y & Lin, C 1996, 'Multivariate C_p value', *Chinese Journal of applied Probability and Statistics*, vol. 12, pp. 132-138.

Liao, S, Chen, K & Li, R 2002, 'Capability evaluation for process of the larger-the-better type for non-normal populations', *Advance Applied Statistics*, vol. 2, no. 2, pp. 189–198.

Liaw, C 2000, 'A hybrid genetic algorithm for the open shop scheduling problem', *European Journal of Operational Research*, vol. 124, pp. 28-42.

Lin, S & Keringhan, B 1973, 'An effective heuristic techniques for the travelling salesman problem', *Operation Research*, vol. 21, pp. 498-516.

Liu, B, Wang, L, Jin, H & Huang, D 2005, 'Advances in particle swarm optimization algorithm', *Control Instruments Chemical Industry*, vol. 32, no. 3, pp. 1-6.

Liu, P & Chen, F 2006, 'Process capability analysis of non-normal process data using the Burr XII distribution', *International Journal of Advanced Manufacturing Technologies*, vol. 27, pp. 975-984.

Lundy, M & Teng, H 1986, 'Convergence of an annealing algorithm', *Mathematical Programming*, vol. 34, pp. 111-124.

Michalewicz, Z & Fogel, D 2000, *How to solve it: Modern Heuristic*, Springer.

Montgomery, D 2005a, *Design of Experiments*, Sixth edition edn, John Wiley & Sons, Hoboken, NJ.

Montgomery, D 2005b, *Introduction to statistical quality control*, 5th edition edn, Wiley, New York.

Muhlenbein, H 1993, 'Predictive models for the breeder genetic algorithm: continuous parameter optimization', *Evolutionary Computation*, vol. 1, no. 1, pp. 25-49.

Mukherjee, S & Singh, N 1994, 'Sampling properties of an estimate of the bivariate process capability index', *Economic quality control*, vol. 9, pp. 73-78.

Munchechika, M 1986, 'Evaluation of Process Capability for Skew Distributions', paper presented to Proceedings 30th EOQC Conference, Stockholm.

Nelder, J & Mead, R 1965, 'A simplex method for function minimization', *Computer Journal*, vol. 7, no. 4, pp. 308-313.

Niaki, S & Abbasi, B 2007, 'Skewness Reduction Approach in Multi-Attribute Process Monitoring', *Journal of Communications in Statistics - Theory and Methods*, vol. 36, no. 12, pp. 2313 - 2325.

Niverthi, M & Dey, D 2000, 'Multivariate process capability: A Bayesian perspective', *Communications in statistics- Simulations and computation*, vol. 29, pp. 667-687.

Noorossana, R 2002, 'Process Capability Analysis in the Presence of Autocorrelation', *Journal of Quality and Reliability Engineering International*, vol. 18, pp. 75-77.

O'Connell, J & Shao, Q 2004, *Further investigation on a new approach in analyzing extreme events*, CSIRO Mathematical and Information Sciences.

Osman, I & Kelley, J 1996, *Metaheuristics: Theory and applications*, Kluwer academic publishers.

Pal, S 2005, 'Evaluation of non normal process capability indices using generalized lambda distribution', *Quality Engineering*, vol. 17, pp. 77-85.

Pearl, J 1984, *Heuristics: Intelligent search strategies for computer problem solving*, Addison Wesley.

Pearn, W & Kotz, S 1994, 'Application of Clements' method for calculating second and third generation process capability indices for non-normal Pearsonian populations', *Quality Engineering*, vol. 7, no. 1, pp. 139-145.

Pearson, K 1895, 'Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material', *Philosophical Transactions of the Royal Society of London ARRAY*, vol. 186, pp. 343–414.

Pincus, M 1970, 'A Monte Carlo method for the approximate solution of certain types of constrained optimization problems', *Operations research*, vol. 18, pp. 1225-1228.

Polansky, A, Chou, Y & Mason, R 1998, 1999, 'Estimating process capability indices for a truncated distribution', *Quality Engineering*, vol. 11, pp. 257-265.

Pyzdek, T 1992, 'Process capability analysis using personal computers', *Quality Engineering*, vol. 4, pp. 419-440.

Rechenberg, I 1978, 'Evolution strategy," ', *Simulations methoden in der Medizin and Biologie*, pp. 83-114.

Rivera, L, Hubele, N & Lawrence, F 1995, 'Cpk Index Estimating Using Data Transformation', *Computers & Industrial Engineering*, vol. 29, pp. 55-58.

Rodriguez, R 1977, 'A guide to the Burr type XII distributions', *Biometricka*, vol. 64, no. 1, pp. 129-134.

Rosenkrantz, D, Stearns, R & Lewis, P 1977, 'An analysis of several heuristics for the travelling salesman problem', *SIAM journal of computing,*, vol. 6, pp. 563-581.

Sarkar, A & Pal, S 1997, 'Estimation of Process Capability Index for Concentricity', *Quality Engineering*, vol. 9, pp. 665-671.

Shahriari, H, Hubele, N & Lawrence, F 1995, 'A multivariate process capability vector', paper presented to Proceedings of the Fourth Industrial Engineering Research Conference (Institute of Industrial Engineers.

Singpurwalla, N 1992, 'A Bayesian perspective on Taguchi's approach to quality engineering and tolerance design', paper presented to IIE Trans.

Singpurwalla, N 1998, 'The stochastic control of process capability indices', *Sociedad de Estadistica e Investigacion Operativa Test*, vol. 7, no. 1, pp. 1-74.

Somerville, S & Montgomery, D 1996, 'Process capability indices and non-normal distributions', *Quality Engineering*, vol. 19, no. 2, pp. 305-316.

Stutzle, T 1998, 'An ant approach to the flow shop problem ', paper presented to Proceedings of 6th European Congress Intelligent Technology and Soft Computing.

Taam, W, Subbaiah, P & Liddy, J 1993, 'A note on multivariate capability indices', *Journal of Applied Statistics*, vol. 20, no. 3, pp. 339–351.

Takahasi, K 1965, 'Note on the multivariate Burr's distribution', paper presented to Annual Proceedings of Institute of Statistics and Mathematics Tokyo.

Tang, L & Than, S 1999, 'Computing process capability indices for non-normal data : a review and comparative study', *Quality and reliability engineering International*, vol. 15, pp. 339-353.

Telmoudi, R 2005, 'A multi-stream process capability assessment using a nonconformity ratio based desirability function', University of Dortmund.

Vännman, K & Albing, M 2007, 'Process capability indices for one-sided specification intervals and skewed distributions', *Quality and reliability engineering International*, vol. 23, pp. 755-765.

Veevers, A 1998, 'Viability and capability indices for multiresponse processes', *Journal of Applied Statistics*, vol. 25, pp. 545-558.

Wang, F 2006, 'Quality evaluation of a manufactured product with multiple characteristics', *Quality and reliability engineering International*, vol. 22, pp. 225-236.

Wang, F & Du, T 2000, 'Using principal component analysis in process performance for multivariate data', *Omega*, vol. 28, pp. 185–194.

Wang, F & Hubele, N 1999, ' Quality evaluation using geometric distance approach', *International Journal of Reliability, Quality and Safety Engineering*, vol. 6, pp. 139-153.

Wang, F, Hubele, N, Lawrence, F, Miskulin, J & Shahriari, H 2000, 'Comparison of Three Multivariate Process Capability Indices', *Journal of Quality Technology*, vol. 32, no. 3, pp. 263-275.

Wang, F & NF, H 1999, 'Quality evaluation using geometric distance approach', *International Journal of Reliability, Quality and Safety Engineering*, vol. 6, pp. 139–153.

Wierda, S 1993, 'A multivariate process capability index', paper presented to Quality Congress.

Wright, P 1995, 'A process capability index seneitive to skewness', *Journal of Statistical Computation and Simulation*, vol. 52, pp. 195-203.

Wu, H & Swain, J 2001, 'A Monte Carlo Comparison of Capability Indices when Processes are Non-normally Distributed', *Quality and reliability engineering International*, vol. 17, pp. 219-231.

Wu, H, Swain, J, Farrington, P & Messimer, S 1999, 'A Weighted Variance Capability Index for General Non-normal Processes', *Quality and reliability engineering International*, vol. 15, pp. 397-402.

Wu, H, Wang, J & Liu, T 1998, 'Discussions of the Clements-based process capability indices', paper presented to Proceedings of the 1998 CIIE National Conference, China.

Yeh, A & Bhattacharya, S 1998, 'A robust capability index', *Communications in statistics- Simulations and computation*, vol. 27, pp. 565-589.

Zahid, A & Sultana, A 2008, 'Assessment and comparison of multivariate process capability indices in ceramic industry', *Journal of Mechanical Engineering*, vol. 39, no. 1, pp. 18-25.

Zimmer, W & Burr, I 1963, 'Variables sampling plans based on non normal populations', *Industrial quality control*, vol. July, pp. 18-26.