Forecasting Manufacturing Variation Using Historical Process Capability Data: Applications for Random Assembly, Selective Assembly, and Serial Processing

by

Daniel C. Kern

M.S., Mechanical Engineering
Georgia Institute of Technology, 1998

B.S., Mechanical Engineering
Georgia Institute of Technology, 1996

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirement for the Degree of

Doctor of Philosophy in Mechanical Engineering

at the

Massachusetts Institute of Technology

June 2003

© 2003 Massachusetts Institute of Technology
All rights reserved

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
AUG 26 2003
LIBRARIES

Signature of Author ______________________________
Department of Mechanical Engineering
May 15, 2003

Certified by ______________________________
Anna Thornton
Thesis Supervisor

Accepted by ______________________________
Ain A. Sonin
Chairman, Department Committee on Graduate Students

BARKER

Forecasting Manufacturing Variation Using Historical Process Capability Data: Applications for Random Assembly, Selective Assembly, and Serial Processing

by

Daniel C. Kern

Submitted to the Department of Mechanical Engineering on May 15, 2003
in Partial Fulfillment of the Requirement for the Degree of
Doctor of Philosophy in Mechanical Engineering
at the
Massachusetts Institute of Technology

ABSTRACT

In today's competitive marketplace, companies are under increased pressure to produce products that have a low cost and high quality. Product cost and quality are influenced by many factors. One factor that strongly influences both is manufacturing variation. Manufacturing variation is the range of values that a product's dimensions assume. Variation exists because no production process is perfect. Often times, controlling this variation is attempted during production when substantial effort and resources, e.g., time, money, and manpower, are required. The effort and resources could be reduced if the manufacturing variation could be forecast and managed during the design of the product.

Traditionally, several barriers have been present that make forecasting and managing variation during the design process very challenging. The first barrier is the effort required of a design engineer to know the company's process capability, which makes it difficult to specify tolerances that can be manufactured reliably. The second barrier is the difficulty associated with understanding how a single manufacturing process or series of processes affects the variation of a product. This barrier impedes the analysis of tradeoffs among processes, the quantifying of the impact incoming stock variation has on final product variation, and the identification of sources of variation within the production system. The third barrier is understanding how selective assembly influences the final variation of a product, which results in selective assembly not being utilized efficiently.

In this thesis, tools and methods to overcome the aforementioned barriers are presented. A process capability database is developed to connect engineers to manufacturing data to assist with detailing a design. A theory is introduced that models a production process with two math functions, which are constructed using process capability data. These two math functions are used to build closed-form equations that calculate the mean and standard deviation of parts exiting a process. The equations are used to analyze tradeoffs among processes, to compute the impact incoming variation has on output, and to identify sources of variation. Finally, closed-form equations are created that compute the variation of a product resulting from a selective assembly operation. Using these tools, forecasting and managing manufacturing variation is possible for a wide variety of products and production systems.

Thesis Committee:
Dr. Anna Thornton (Thesis Supervisor)
Professor David Hardt (Thesis Committee Chairperson)
Professor Sanjay Sarma

Table of Contents

1 INTRODUCTION 15

1.1 Overview 17
1.2 Hypothesis 17
1.3 Challenges 19
1.4 Approach 20
1.5 Related Work 21
1.6 Thesis Structure and Advancements 27
1.7 Chapter Summary 28

2 PROCESS CAPABILITY DATABASE DESIGN 29

2.1 Introduction 31
2.1.1 Overview 31
2.1.2 Background 31
2.2 Indexing Scheme 35
2.2.1 Attribute Combination Descriptor (ACD) 36
2.2.2 Benefits of the Attribute Combination Descriptor 38
2.2.3 Method to Index Design Characteristics 41
2.3 User Interface 52
2.4 Record Set 59
2.5 Chapter Summary 62

3 MODELING OF MANUFACTURING PROCESSES 67

3.1 Introduction 69
3.1.1 Overview 69
3.1.2 Background 69
3.2 Theoretical Framework 72
3.2.1 Theory for Computing Output Mean 73
3.2.2 Theory for Computing Output Variation 76
3.3 Method to Compute μ_{output} and σ_{output} 78
3.3.1 Distribution Binning and Processing 78
3.3.2 Approximate Solution Transformation 82
3.3.3 Forms of DeltaP and SigmaP 83
3.3.4 Closed-form Solution for μ_{output} and σ_{output} 84
3.4 Coefficients of DeltaP and SigmaP 85
3.5 Applications of the Modeling Methodology 89
3.6 Validation of DeltaP and SigmaP Theory 91
3.6.1 Turning 94
3.6.2 Rolling 108
3.6.3 Sheet Bending 117
3.7 Guidelines to Equation Construction and Usage 127
3.8 Chapter Summary 130

4 MODELING OF SELECTIVE ASSEMBLY 135
4.1 Introduction 137
4.1.1 Overview 137
4.1.2 Background 138
4.2 Selective Assembly Techniques 143
4.2.1 Equal Width Binning to Minimize Variation 144
4.2.2 Equal Width Binning to Minimize Scrap 146
4.2.3 Equal Area Binning to Minimize Variation 147
4.2.4 Equal Area Binning to Minimize Scrap 149
4.2.5 Optimal Width Binning 151
4.2.6 Selective Machining 151
4.2.7 Queuing Method 152
4.3 Selective Assembly Considerations 152
4.3.1 Quality Considerations 153
4.3.2 Economic Considerations 153
4.4 Selective Assembly Calculations 156
4.4.1 Closed-form Equations 158
4.4.2 Sensitivity Analysis 164
4.5 Selective Assembly Equation Validation 165
4.6 Example Problems 170
4.6.1 Selective Assembly of Two Normal Distributions 170
4.6.2 Selective Assembly of Two Skewed Distributions 180
4.6.3 Selective Assembly of a Skewed and an Irregular Distribution 183
4.7 Chapter Summary 185
5 SYSTEM APPROACH: VRUM 189
5.1 Introduction 191
5.2 System Model Composition 192
5.2.1 Icons 192
5.2.2 System Diagram 193
5.2.3 Output Equations 194
5.3 Role of VRUM 196
5.4 Implementation 198
5.5 Example Problem 200
5.6 Chapter Summary 217
6 CONCLUSION 219
6.1 Introduction 221
6.2 Research Advancements 221
6.2.1 Process Capability Database 221
6.2.2 Modeling of Manufacturing Processes 223
6.2.3 Modeling of Assembly Operations 224
6.2.4 System Approach 225
6.3 Future Research 226
6.3.1 Process Capability Database Future Work 226
6.3.2 Modeling of Manufacturing Processes Future Work 226
6.3.3 Modeling of Assembly Operations Future Work 227

6.3.4 System Approach Future Work 227
6.4 Conclusion 228

APPENDICES 231

A. GENERAL NOMENCLATURE 233

B. SURVEY RESULTS 235

C. SELECTIVE ASSEMBLY VARIATION COEFFICIENTS 245

D. SELECTIVE ASSEMBLY SCRAP COEFFICIENTS 249

E. SELECTIVE ASSEMBLY SIMULATION DATA 253

F. ADDITIONAL PCDB USER INTERFACE INFORMATION 283

G. DETAILED DERIVATION FOR CHAPTER 3 287

H. HEURISTIC GUIDELINE DEVELOPMENT 317

I. NOMENCLATURE USED IN VALIDATION MODELS 325

J. DATA FOR EXAMPLE PROBLEM 327

K. DERIVATION OF MIXTURE DISTRIBUTION EQUATIONS 329

GLOSSARY 335

BIBLIOGRAPHY 337

Table of Figures

Figure 1-1: Graphical depiction of hypothesis 19
Figure 1-2: Schematic of approach to forecast manufacturing variation 20
Figure 2-1: General Attribute Combination Descriptor 36
Figure 2-2: Sample database hierarchy 37
Figure 2-3: Visual example of attributes 37
Figure 2-4: Example Attribute Combination Descriptor 38
Figure 2-5: Hierarchy tree with Feature only 39
Figure 2-6: Hierarchy trees of Feature and Geometry 39
Figure 2-7: Side by side comparison of database hierarchies 40
Figure 2-8: Query a single instance of geometry instead of many 40
Figure 2-9: Illustration of extensibility 41
Figure 2-10: Example F-matrix template 43
Figure 2-11: Sample key characteristic listing 44
Figure 2-12: Visual sample key characteristics 44
Figure 2-13: Partially populated F-matrix 45
Figure 2-14: F-matrix populated 45
Figure 2-15: Initial G-matrix after F-matrix is populated 46
Figure 2-16: G-matrix populated 47
Figure 2-17: Blank M-matrix 47
Figure 2-18: M-matrix populated 48
Figure 2-19: P-matrix1 partially populated 48
Figure 2-20: P-matrix1 populated 49
Figure 2-21: P-matrix2 populated 49
Figure 2-22: Introduction screen to the PCDB 54
Figure 2-23: Displayed security level once logged on 54
Figure 2-24: Screen capture of query page 56
Figure 2-25: Material_2 menu items before Material_1 is specified 57
Figure 2-26: Aluminum being specified for Material_1 57
Figure 2-27: Material_2 selections once aluminum is specified 57
Figure 2-28: Add Record Page 58
Figure 2-29: Traditional view of process capability data 60
Figure 2-30: Proposed mental model of process capability data 61
Figure 3-1: Process transforms a distribution 70
Figure 3-2: End result 72
Figure 3-3: Graphical definition of DeltaP 75
Figure 3-4: Graph of sample DeltaP 75
Figure 3-5: Graphical definition of SigmaP 77
Figure 3-6: Graph of sample SigmaP 77
Figure 3-7: Sample of a binned input distribution 79
Figure 3-8: Bin transformed into output sub-distribution 80
Figure 3-9: Method for finding coefficients of DeltaP 87
Figure 3-10: Method for finding coefficients of SigmaP 88
Figure 3-11: Illustration of modeling a manufacturing process 89
Figure 3-12: Creation of closed-form equations for theory validation 92
Figure 3-13: Validating calculated values vs. simulated values 94
Figure 3-14: Dimensions of interest in turning 96
Figure 3-15: Cutting forces during turning 96
Figure 3-16: End views of cutting tool and workpiece for ideal and modeled cases 99
Figure 3-17: Geometry considerations for calculating actual final diameter 99
Figure 3-18: Algorithm for Monte Carlo simulation method 101
Figure 3-19: DeltaP for turning operation used for validation 103
Figure 3-20: SigmaP for turning operation used for validation 104
Figure 3-21: Schematic of rolling operation 109

Figure 3-22: Force acting on rolls 110
Figure 3-23: Algorithm for Monte Carlo simulation method for rolling 113
Figure 3-24: DeltaP for rolling operation used for validation 115
Figure 3-25: SigmaP for rolling operation used for validation 116
Figure 3-26: Final desired shape of bent component 118
Figure 3-27: Workpiece in the die 119
Figure 3-28: Graphical representation of springback 119
Figure 3-29: Algorithm for Monte Carlo simulation method for sheet bending 122
Figure 3-30: DeltaP for sheet bending operation used for validation 125
Figure 3-31: SigmaP for sheet bending operation used for validation 125
Figure 3-32: Pictorial summary of mathematical derivation 134
Figure 4-1: Mental model of assembly operation 138
Figure 4-2: Graphical representation of selective assembly 141
Figure 4-3: Equal Width Binning to Minimize Variation 145
Figure 4-4: Equal Width Binning to Minimize Scrap 146
Figure 4-5: Equal Area Binning to Minimize Variation 148
Figure 4-6: Equal Area Binning to Minimize Scrap 151
Figure 4-7: Graphical illustration of the cost of selective assembly 155
Figure 4-8: Graphical depiction of the selective assembly algorithm 157
Figure 4-9: R^2 values closest to one for assembly standard deviation 166
Figure 4-10: Lowest R^2 for assembly standard deviation 167
Figure 4-11: Example of confidence intervals with high R^2 values 168
Figure 4-12: Confidence intervals with the lowest R^2 values 168
Figure 4-13: R^2 values closest to one for scrap 169
Figure 4-14: Typical case for a constant value best representing a scrap equation 170
Figure 4-15: Assembly standard deviation versus number of bins for multiple approaches 172
Figure 4-16: Scrap fraction versus the number of bins used for multiple approaches 173
Figure 4-17: Assembly standard deviation versus scrap fraction 174
Figure 4-18: Solution point to example problem 175
Figure 4-19: Assembly standard deviation with ±95% confidence interval 176
Figure 4-20: Scrap fraction with ±95% confidence interval 177
Figure 4-21: Comparing the assembly standard deviation for multiple approaches using four bins 178
Figure 4-22: Comparing the scrap fraction for multiple approaches using four bins 179
Figure 4-23: Two skewed distributions to be combined 181
Figure 4-24: Histograms of assembly dimensions for different numbers of bins 182
Figure 4-25: Assembly standard deviation versus number of bins 183
Figure 4-26: Skewed distribution and irregular distribution 184
Figure 4-27: Histograms of assembly dimensions for different numbers of bins 184
Figure 5-1: How the "pieces" fit 191
Figure 5-2: Icons used for VRUM system model diagrams 193
Figure 5-3: Example of a VRUM system model diagram 194
Figure 5-4: Cascading of information from one process to the next 195
Figure 5-5: Decisions enhanced with VRUM 197
Figure 5-6: Generic four phase product development process 198
Figure 5-7: Schematic of steps used in example problem 201
Figure 5-8: Illustration of widget being designed 202
Figure 5-9: Top view of widget showing the critical dimension 202
Figure 5-10: VRUM system diagram for example problem 203
Figure 5-11: DeltaP curves for the available roughing operations 205
Figure 5-12: SigmaP curves for the available roughing operations 206
Figure 5-13: DeltaP curves for the available finishing operations 207
Figure 5-14: SigmaP curves for the available finishing operations 207
Figure 5-15: Illustration of different inputs for roughing and finishing 210
Figure 5-16: Possible solutions 214
Figure 5-17: Possible solutions with Pareto Front drawn 215
Figure 5-18: 95% worst case scenario for Pareto Front feasible points 216

Figure F-1: Administration main page 283
Figure F-2: Manufacturing Plant Management Page.. 284
Figure F-3: Page to manage material attribute listings in the PCDB... 285
Figure F-4: Page used to manage user accounts 286

To Karen and Cyler

Acknowledgements

My sincerest thanks to God for blessing me with the ability and determination to complete the work presented here. I thank my wife, Karen, for her dedication and support over the years at MIT, especially for her understanding of the late nights of typing. A special thanks to my parents for being lifelong encouragers of my educational pursuits and never letting me doubt my abilities.

My appreciation is extended to Dr. Anna Thornton, Dr. David Hardt, and Dr. Sanjay Sarma for the guidance and suggestions they shared with me. A thanks to my past lab mates for their help and friendship. I also appreciate the assistance and insight from Ann O'Connell and Dr. Agus Sudjianto. My gratitude goes to The Ford Motor Company for their financial support that made this work possible.

1 Introduction

1.1 Overview

The goal of any manufacturing firm is to produce products that meet and exceed customer based requirements by producing high quality products. Producing quality products is challenging because they are comprised of one or more components, each having many features that have measurable dimensions, e.g., length, width, diameter, etc. Producing every dimension to an exact value is impossible because of variation. Manufacturing variation is the range of values that a dimension actually spans and directly affects the quality of a product. Companies are motivated to better understand and manage manufacturing variation to lower costs and improve quality. Managing manufacturing variation is the act of making design and production decisions so that a product can be produced with an acceptable level of variation at an acceptable cost.

In an ideal world, engineers could understand and manage manufacturing variation to create product designs that could be manufactured with a low cost and high level of quality. Design engineers would have a full appreciation and knowledge of the firm's capability to manufacture goods. They would be able to see into the future and know the level of quality a product will have before it is even produced. Decisions regarding large, complex manufacturing systems would be based on real data.

Unfortunately, this ideal world does not exist because of several major barriers. Design engineers are often unfamiliar with their firm's capability to manufacture parts (Tata and Thornton 1999). Data that could help them understand are usually geographically dispersed, difficult to retrieve, and hard to interpret. Even if they did find the data, design engineers would not know how to always use it to help design high quality parts. Finally, products are often manufactured by large, complex manufacturing systems. Understanding how these systems directly affect a product's quality is very challenging and making decisions as to the types of processes and assembly operations to use are often done in an ad hoc manner. In other words, designing high quality products is difficult because forecasting manufacturing variation is very challenging.

1.2 Hypothesis

The work in this dissertation addresses the hypothesis that a design tool can be developed that uses manufacturing process capability data to construct math models to forecast the quality of a dimension on a design prior to being manufactured. This design tool will assist engineers in understanding and managing variation so that products are designed that can be manufactured with an acceptable level of quality at the lowest possible cost despite the presence of manufacturing variation.

Figure 1-1 graphically illustrates the hypothesis using a map of a product development process. Elements of the picture represent different items in a product development process. The dark circles represent manufacturing data. The white circles represent artifacts of the product development process, i.e., general product requirements labeled "functional intent," detailed drawings and other specifications labeled "the design," and the plans to be used to produce the parts labeled "production plans." Actions are represented by the rectangles. Each arrow in the picture represents a flow of information.

The map starts with the functional intent, which drives the design and tolerancing of the product. The result of these two actions is the design, i.e., detailed drawings, material specifications, detailed operation specifications, etc. The design helps in the selection of suppliers and processes to use to manufacture the product. The process selection outputs production plans. The design and production plans directly influence the production operations used. The lower sequence of elements demonstrates that a measurable mean and standard deviation of parts enters the first operation of the production system. These values are known in the sense they can be measured before any operation takes place and, the company has some control over their values, e.g., the company can buy higher grade stock to have less variation present in the incoming stock. These values are affected by the series of operations that make up the production line, and an unknown mean and standard deviation exit the series of operations. These output values are unknown in the sense they cannot typically be determined until all operations are complete.

The hypothesis is that these output values, μ_{out} and σ_{out}, can be managed during the portion of the product development process that is contained within the gray area, i.e., decisions during the design, tolerancing, supplier selection, and process selection phases can positively influence of a product's final mean and standard deviation. To accomplish this, a tool must be developed that can forecast the mean and standard deviation of a product prior to its production. Using this tool, engineers can calculate the effects design, tolerancing, supplier selection, and process selection decisions have on the end quality; thus, the quality can be managed early in the product development process. The tool must take into account the effects random assembly, selective assembly, manufacturing processes, and serial processes have on the final quality of a product's dimension.

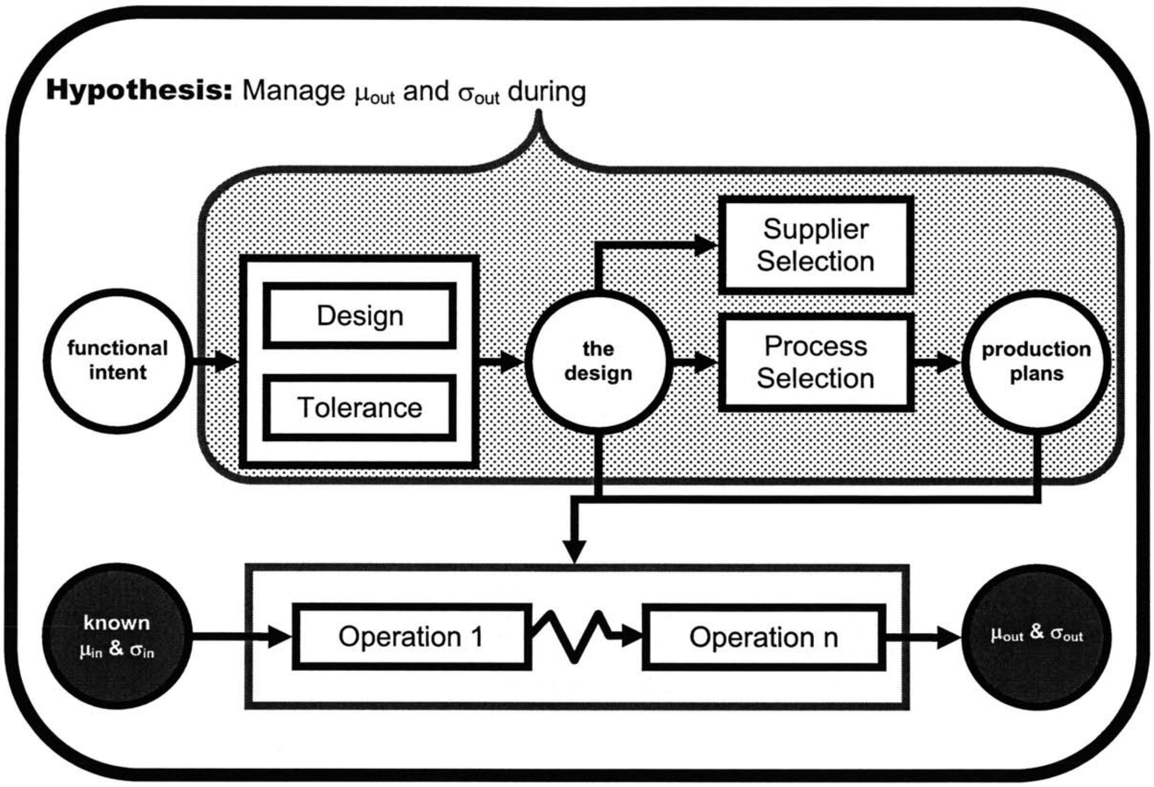

Figure 1-1: Graphical depiction of hypothesis

1.3 Challenges

Utilizing manufacturing data during the design process to manage variation is very difficult for a manufacturing organization. One challenge is that the appropriate data are not readily available to the appropriate engineers. Making manufacturing data easy to retrieve for design engineers is difficult because it is often dispersed throughout an organization and can be in numerous unique forms, making it difficult to interpret.

Manufacturing variation is challenging to understand and manage because it is influenced by many factors. Product dimensions can be affected by multiple processes in a series. This makes determining where the variation is introduced and how it changes as the part passes through the manufacturing system difficult. Product variation also depends on whether selective assembly is used. When selective assembly is implemented, the variation of a dimension can be reduced. Unfortunately, selective assembly is often implemented as a "last resort" and in an ad hoc manner using trial and error, which results in not managing variation effectively.

1.4 Approach

The approach taken in this thesis is to develop a design tool that assists an engineer in designing products with high levels of quality by forecasting manufacturing variation during the design of the products. This design tool is comprised of a database and math functions. The database is designed to store and make readily available manufacturing capability data to design engineers. A mathematical framework is created that allows design engineers to compute the variation a dimension has upon exiting a manufacturing process. Math models for selective assembly are created to help engineers decide when to implement selective assembly into a manufacturing system.

A schematic of the design tool being developed is shown below in Figure 1-2. The gray cylinder represents the Process Capability Database. Data from the database are used to construct the math functions representing manufacturing processes. These functions compute the mean and standard deviations of parts exiting processes, which feed into the equations for assembly operations. The mean and standard deviation, μ_{out} and σ_{out}, for the final dimension are then output. The chapters that discuss each portion of the design tool are listed in Figure 1-2.

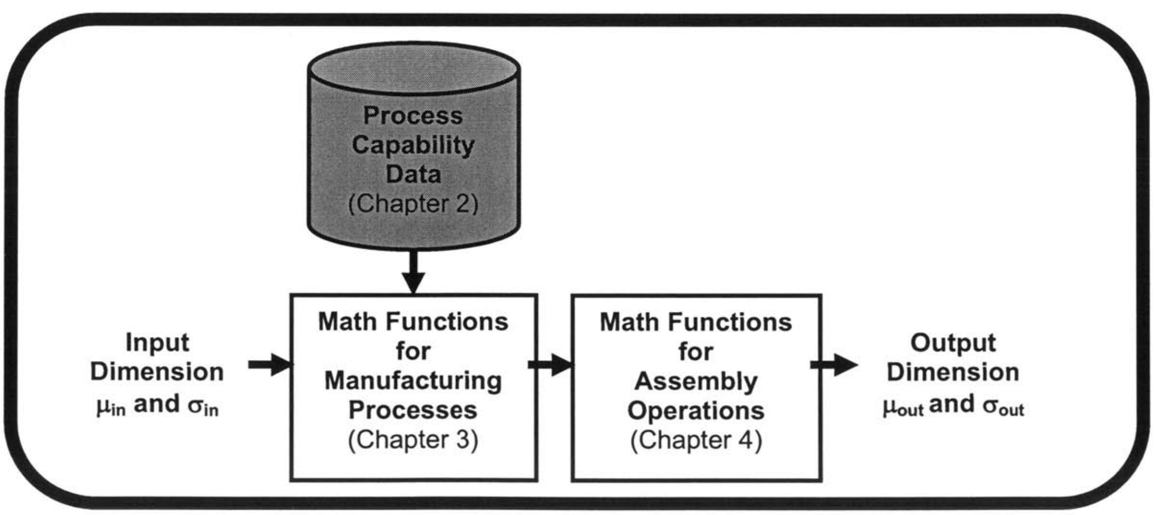

Figure 1-2: Schematic of approach to forecast manufacturing variation

The design tool is called *Variation **R**isk **U**nderstanding & **M**anagement* (*VRUM*, pronounced "vroom"). The term "risk" is included to represent the uncertain negative effects associated with neglected manufacturing variation. Therefore, VRUM signifies a proactive approach to understanding and managing, through design and manufacturing decisions, the risk associated with manufacturing variation. VRUM is discussed in depth in Chapter 5.

1.5 Related Work

Creating high quality parts with low costs is very challenging. A great amount of research has been performed to study methods and techniques to accomplish this task. The existing work spans many fields and includes a diverse body of literature. The research in this thesis results in a design tool (presented in Chapter 5) that relates to a large portion of the existing research. The bodies of work that are relevant to the content in this thesis are discussed in this section. Each body of work is listed followed by a brief description and a discussion of how that body of work relates to this thesis.

Methodologies to "Design for Quality"

Description: Two popular methodologies exist to provide a framework to design high quality products. Design for Six Sigma (DFSS) (Harrold 1999; Antony and Coronado 2002) emphasizes the creation of designs that can be manufactured so that there are only 3.4 defects per million parts produced. Quality Function Deployment (QFD) is another methodology (http://www.qfdi.org). QFD is concerned with capturing customer needs and translating these needs into requirements that can be met through designed products.

Relevance: Both DFSS and QFD are high level methodologies intended to guide the design and development of a product. Neither has specific design tools associated with it. The design tool created in this thesis (see Chapter 5) can be used with either methodology to assist in the detailed design stage of a product development process. Six Sigma initiatives often include the collection of process capability data. These data can be used to populate a Process Capability Database (see Chapter 2).

Manufacturing Variation Modeling

Description: Manufacturing variation is always present because no process is perfect. Modeling variation is a good step to understanding it and designing products that are unaffected by it. Research has been done to investigate managing the effect variation has on the final cost and quality of a product while it is being designed. A tool set has been developed to use during the design process to identify where variation will have significant impact (Thornton 1999). This procedure allows attention to be focused on the identified product characteristics to improve the overall quality of the product. Kazmer et al. use finite-element models of processes to analyze the variation within a single operation (Kazmer, Barkan et al. 1996). Suri and Otto (Suri and Otto 1999) use variational modeling to select manufacturing process input parameters that render the process outputs insensitive to variation. This technique identifies the product characteristics and process parameters that affect the quality of a manufactured product; however, it does not provide a method to forecast a product's quality while it is being designed. Variational modeling has also been used to design a product while optimizing its quality and the production cycle time (Wei 2001). A physics-based approach to modeling variation was introduced by Suri that predicts the nominal values and variation of each output quality characteristic in a manufacturing system (Suri 1999). This methodology was adopted by Ota who used it to model the variation present in cold rolled steel manufacturing (Ota 2000). The concept of a process capability matrix and bias vector to model a manufacturing system was introduced by Frey (Frey 1997). The model developed can be used to compute the yield from a manufacturing process.

Relevance: The above mentioned works help identify and manage product variation; however, each method often requires advanced modeling or expertise. The work in this thesis is concerned with understanding manufacturing variation while a product is being designed. A tool is developed that forecasts the variation that will be present when a product is produced. This tool is unique from the above approaches because it does not require any physical modeling, uses historical process capability data, is simple to use, and produces closed-form equations that are computationally inexpensive.

Manufacturing Variation Propagation

Description: Research in the field of variation propagation attempts to identify sources of variation in a manufacturing system, identify how that variation propagates through the system, and suggest adjustments to minimize the final product variation. A "Stream of Variation" theory has been developed by Hu to examine variation propagation in the assembly of flexible components (Hu 1997). A model of how variation is transmitted through a production system is presented by Lawless, Mackay, and Robinson (Lawless, MacKay et al. 1999). Their approach is to measure parts at each process in the manufacturing system and determine the amount of variation added by each process and carried through the next process. Other research has been conducted to model dimensional variation in multistage machining processes (Zhou, Huang et al. 2003). The approach developed uses differential motion vectors as state vectors to represent geometric deviations. Several bodies of work exist that use control theory to assist in the identification of root causes of manufacturing variation in multistage manufacturing systems (Ding, Shi et al. 2002; Huang, Zhou et al. 2002). Control theory is also used as the basis for research into modeling and controlling variation propagation in assemblies (Mantripragada and Whitney 1999). The concept is to use the models to reduce assembly variation by rearranging the assembly sequence and by making in-process adjustments. These approaches are predominantly used after the manufacturing system is operational, not during the design of a product.

Relevance: The work in this thesis is not specifically focused on variation propagation; however, the closed-form equations can be "linked" together to calculate how a product dimension's variation is propagated through a series of manufacturing operations, including assembly. The work presented here is unique in that it is easy to use, based on process capability data (does not require detailed knowledge of any process), and produces closed-form results. The design tool to accomplish this is presented in Chapter 5.

Robust Design

Description: Robust design is the idea of creating products with functionality that is insensitive to variation present in manufacturing and the environment in which the products are used. Robust design methods are based on the work of Taguchi

(Clausing 1988; Phadke 1989; Taguchi 1993). The approach is to set design and process parameters in regions where their response functions are "flat," i.e., small shifts in the parameters do not cause major effects to the product functionality. A detailed understanding of how design and process parameters relate to product performance is required and is often difficult to develop; however, Design of Experiments is one method often implemented to gain this understanding.

Relevance: Robust design is broadly concerned with creating products that are insensitive to many types of variation. The design tool created in this thesis helps design products that are insensitive to manufacturing variation.

Response Surface Methodology (RSM)

Description: Response Surface Methodology is a technique that uses a polynomial surface to approximate the relationship between multiple input variables and a single output variable (Montgomery 1984). The polynomial surface is explored to find the optimal value of the output variable. The concept of RSM has been extended to include Multiple Response Surfaces (MRS). The Multiple Response Surfaces approach uses less data to fit multiple low-order surfaces and can be more accurate than using a single high-order surface (Guo and Sachs 1993). A related methodology is the Robust Concept Exploration Method (RCEM) (Chen, Allen et al. 1996; Chen, Allen et al. 1997), which combines aspects from Response Surface Methodology, robust design techniques, Axiomatic design, and the Decision Support Problem. RCEM is used to evaluate design alternatives and develop top-level specifications. Another similar technique for finding input parameters that produce the desired output is known as Evolutionary Operation (EVOP) (Box and Draper 1969). EVOP uses a design of experiment to test different parameter settings and find a good operating point. The technique is repeated on a continuous basis, and the operating point is updated using the new information gathered.

Relevance: The above approaches are mostly interested in finding process parameters that produce a desired output. The design tool developed in this thesis is not concerned with specific process parameters. Instead, it assists in the design of products that can be manufactured with a process that has a known set of parameters.

Tolerance Allocation

Description: Tolerance allocation is the act of assigning tolerances to different components that comprise an assembly. Performing tolerance allocation in a thoughtful, intelligent manner improves product quality while reducing manufacturing costs because some components are easier and less expensive to manufacture with tighter tolerances than others. Therefore, an assembly's dimension that is created by the stacking of several components can be controlled to a tighter tolerance by assigning tight tolerances to the components that are easier to manufacture. A great deal of work has been performed to optimize the tolerance allocation of a product by considering Taguchi quality loss functions and manufacturing costs (Bloebaum, Mulubagal et al. 1994; Krishnaswami and Mayne 1994; Soderberg 1994); the goal being to minimize the cost of a product by properly assigning tolerances to the appropriate components.

Relevance: Tolerance allocation is an important aspect of the detailed design stage of developing a new product. The act of allocating tolerances can be enhanced using the design tool developed in this thesis (see Chapter 5). The tool forecasts the manufacturing variation a dimension on a component will have when produced. This information can be used when assigning tolerances, thus improving the reliability of the tolerance allocation.

Statistical Process Control (SPC)

Description: Statistical Process Control (SPC) is a technique used in a manufacturing environment to ensure quality parts are produced. SPC uses control charts to plot measurements of component dimensions being produced. These charts are used to alert the operator to shifts in the mean of the measurement. The measurements are also used to compute process capability indexes such as C_{pk} and C_p. The indexes indicate how "good" a process is at producing parts that meet specification. SPC is not a design tool. It is a quality tool implemented in the manufacturing plant.

Relevance: SPC requires a great deal of measurements to be taken on many parts. Traditionally, the gathered data are used only for short-term quality monitoring. One concept in this thesis is that the data gathered as part of an SPC plan can be a valuable asset to an engineer designing a new product. The data provide insight into the available manufacturing capability that would not exist otherwise. A method of making process capability available to design engineers is discussed in Chapter 2, and a technique to use process capability data to construct mathematical equations describing processes is presented in Chapter 0.

Using Manufacturing Process Capability Data during the Design of a Product

Description: Research has shown that the product development process can be improved through the use of historical manufacturing process capability data (Hix and Kittleson 1998). Other work has been done to incorporate historical supplier data into the design process to improve quality (Lucca, Berti et al. 1995). The goal of this concept is to design products that have a high level of manufacturability.

Relevance: These are good steps toward designing products while considering manufacturing variation; however, each method had very specific applications. The design tool in this thesis can be broadly applied to assist engineers with using process capability data effectively to improve a product's design.

Selective Assembly

Description: Selective assembly is the act of measuring components, grouping them, and then selecting mates from the appropriate groups. Research has been performed on several aspects of selective assembly including grouping techniques and methods. See Section 0 for a detailed literature review.

Relevance: The selective assembly work in this thesis extends existing research and has several unique aspects. New selective assembly techniques are defined to use when component distribution variances are unequal. Closed-form equations to calculate the standard deviation of an assembly that is created by a selective assembly operation are created. The robustness of selective assembly to changes in incoming component variances is also explored.

Each of the above bodies of work provides some advancement in the area of creating high quality parts while reducing manufacturing costs. Some provide a framework in which to design while others are used during the production of a product to ensure high levels of quality are maintained.

1.6 Thesis Structure and Advancements

This section gives a brief description of the topics covered in each chapter. Following the descriptions are bullet point lists of the major research advancements made in each chapter.

Chapter 2: The barriers to successfully implementing a database to store manufacturing capability data are discussed. Solutions are developed to overcome these barriers.

Research Advancements

- Introduces a new method of representing data
- Creates a systematic approach to transforming existing product information into a form suitable for database storage
- Designs a new graphical user interface for a database storing process capability data
- Discusses the required data fields to ensure data is useful

Chapter 3: A new theory is introduced that represents an arbitrary manufacturing process with two math functions. These functions are used to construct closed-form equations for the output mean and standard deviation from the process as functions of the incoming mean and standard deviation.

Research Advancements

- Introduces theory to represent a manufacturing process with two math functions
- Develops a method to construct closed-form equations for the mean and standard deviation of parts exiting a manufacturing process
- Validates theory with simulation model and actual data

Chapter 4: The effects that assembly operations have on product quality are discussed and modeled. A detailed discussion of selective assembly is presented including descriptions of different techniques and the advantages and disadvantages of each technique. Math models are created to help in the decision process of when to implement selective assembly and the specific technique to use.

Research Advancements

- Introduces two new selective assembly techniques
- Develops algorithm to simulate selective assembly operations

- Produces closed-form equation for the standard deviation of a dimension resulting from selective assembly
- Produces closed-form equation for selective assembly scrap
- Validates selective assembly closed-form equations using Monte Carlo simulation

Chapter 5: The process capability database, equations for manufacturing process variation developed in Chapter 0, and the equations for assembly operations developed in Chapter 4 are integrated to create a design tool that is used to forecast the end quality of a product dimension upon exiting a production system.

Research Advancements

- Integrates research from previous three chapters into one design tool to forecast the end quality of a product's dimension
- Demonstrates a method to graphically represent a manufacturing system
- Discusses the implementation of the design tool
- Presents an example of the design tool

1.7 Chapter Summary

The overarching theme of this thesis is that process capability data can be used to forecast and manage the manufacturing variation associated with a product during the design of the product. The opportunity and motivation exist to explore this theme because of the present difficulty manufacturing firms face when designing a new product to have a high level of quality. Variation Risk Understanding & Management (VRUM) is introduced as a design tool that assists with design and manufacturing decisions that affect the manufacturing variation associated with a product. VRUM relies on process capability data and math models to describe the effect manufacturing processes and assembly operations have on manufacturing variation.

2 Process Capability Database Design

Chapter Highlights

- Provide overview of process capability data
- Introduce new indexing scheme
- Show method to index data
- Develop database user interface
- Present fields in the database record set

2.1 Introduction

Process capability data are measurements and information related to the variation of dimensions on products a firm produces. Typically, process capability data include the target mean, the actual mean, the upper limit, the lower limit, and the standard deviation for critical component dimensions. A Process Capability Database (PCDB) stores manufacturing data that is recorded during the production process. A Process Capability Database centrally stores this information, making it easier to retrieve by engineers regardless of their geographical location, e.g., an engineer in Michigan is able to retrieve manufacturing data from a plant located in England. Implementing a PCDB that can be used effectively is challenging. This chapter discusses some of the challenges and offers solutions to them.

2.1.1 Overview

The main focus of this chapter is to address the barriers that prevent process capability databases from being used effectively.

Chapter Goals

- Develop a method of representing data so that it is easily retrievable
- Formalize the steps required to transform component data into a form suitable for database entry and storage
- Design a user interface for process capability databases
- Suggest a list of data that should be included in process capability databases

2.1.2 Background

Producing products that meet and exceed customer based requirements is the goal of any manufacturing firm. This goal is accomplished by designing quality into manufactured components. Process capability data can aid a design through ensuring the part tolerances are achievable with the current manufacturing capability (Hix and Kittleson 1998). Many companies want to store process capability data and make it available to all engineers. One way to achieve this is to create a database of process capability data (Tata and Thornton 1999; Thornton and Tata 2000). A database of this type is known as a Process Capability Database (PCDB). A PCDB stores manufacturing data that is recorded during the production process such as the target mean, actual mean, standard deviation, upper limit, and lower limit for important characteristics of manufactured components or assemblies. These measurements are used to calculate process capability indexes, e.g., C_p, C_{pk}, etc. Historically this information has been dispersed throughout the organization in files at the manufacturing site, at the desk of an engineer, or on the plant floor. A Process Capability Database centralizes the storage of this information.

Companies realize the importance and value of knowledge management because it helps propagate lessons learned, avoid repeat mistakes, and capture employee know-how. Even though knowledge management is regarded as important, the issues of poor application implementation, limited language development, knowledge base maintenance, and knowledge under utilization plague companies (Sainter, Oldham et al. 2000). Many opportunities exist for companies to use knowledge management tools to improve product design and quality. One example of using knowledge management to aid in the design of a product is suggested by Zhang and Xue (2002) who present a method of using a database of features for design purposes. However, their work is limited to the generation of designs and does not analyze if they can be produced effectively. Kwon *et al* (2000) propose a different method that creates designs based on stored information. Their approach is to integrate the design and manufacturing processes using a stored set of information divided into the categories of tasks, processes, and constraints. Different inputs from these categories are combined to create a flow diagram to represent a design and one or more process flow diagrams based on the stored knowledge. This approach is useful in creating alternative processing diagrams, but it does not take into account process capability data from manufacturing. Another approach to design with existing knowledge is presented by Xue and Dong (1997) who suggest a method of using combinations of pre-defined features to generate new designs to meet a set of functional requirements. They suggest a separate, yet extensive set of features is needed to integrate the design with manufacturing. The goal of their research is to create new designs, then generate a process/production plan; however, variation is not a consideration. While these knowledge management tools have important roles to play, none of them assess the manufacturability of a product, provide a method for engineers to understand manufacturing capability, or consider manufacturing variation during the design of a product.

Manufacturing process capability data are a source of knowledge that could be used to improve product quality by assessing the manufacturability of a product, providing insight into a company's manufacturing capability, and considering variation when designing a product, but it is not immune to the issues that plague knowledge management tools (Sainter, Oldham et al. 2000). A survey and informal interviews conducted with engineers from an American automobile manufacturer that does not have a PCDB revealed that process capability data are not being used effectively to improve product designs. The survey results showed that design engineers often do not know where to obtain accurate, reliable manufacturing process capability data, resulting in the engineers specifying product tolerances without considering manufacturing variation. The interviews revealed that any new knowledge management tool, e.g. a database, must be easy to use and understand, be easily accessible, and be easy to learn. (A detailed discussion about the survey and interview results, including a listing of the survey questions, is presented in Appendix B.) A study showed that even companies that have PCDBs are not effectively using process capability data to improve product designs (Tata and Thornton 1999). The major barriers preventing process capability data from being used to effectively improve designs are knowing which product characteristics to measure and store, representing information about the product characteristics in the database so that it is easy to retrieve, transforming information about the product characteristics into a form for database storage, and creating a PCDB user interface that is easy to use.

The barrier of knowing what product characteristics to store is important because not every dimension can be measured and tracked statistically due to the overwhelming number, even on the simplest part. To address this problem, the concept of key characteristics is used. A key characteristic is a physical aspect of a component or assembly that has a significant effect on fit, performance, or service life when it varies from its nominal value (Thornton 1999). Storing only the information related to product key characteristics makes the amount of information being stored in the database much more manageable.

Manufacturing data stored in a database is useless unless it is easily retrievable. To ensure data is retrievable, the product characteristics must be represented in a manner that is familiar to engineers. Several attempts have been made to create methods to effectively represent product characteristics. In some firms, the process capability data associated with a component is identified by the component's part number, which makes it difficult for a new engineer or an engineer unfamiliar with the part numbering to locate data. For example an engineer in one design group would not know the part numbers used by a different design group. This representation also causes data to be even more difficult to find when a component's part number changes since all the database is referenced by the old, obsolete part number. A numerical code to represent data in a database has been suggested that allows the data to be found quickly and interpreted (Hanson 2001). Another approach is to use information about a part family and process family to identify process capability data (Hix and Kittleson 1998). This is an interesting approach, but it is effective only when engineers have prior knowledge of the product families. Another approach to representing product characteristics is the use of a Feature Code to describe the attributes of the characteristic (Thornton and Tata 2000). The Feature Code includes information about a characteristic's material, the process that made it, and the feature itself. This representation is sufficient in describing a component aspect, but an improvement can be made to simplify the PCDB structure and make querying the database easier. This improvement comes in the form of the Attribute Combination Descriptor (ACD) being introduced in this thesis. The ACD uses the Feature Code as a foundation and builds upon it by adding a geometry attribute. An ACD is used to fully represent a product's characteristic by specifying a feature, a geometry, a material, and a process attribute. The details and advantages of the ACD are discussed in Section 2.2.

No formal approach exists specifically to overcome the barrier of transforming information about a product characteristic into a form for database storage. Traditionally, this has been done in an ad hoc manner. Without a formal approach, key characteristic representation is non-uniform, which leads to confusion when the PCDB is accessed. A related research area is feature recognition (Han, Pratt et al. 2000), which is the process of analyzing a CAD solid model and extracting feature information. However, these techniques require a 3D CAD model of the product, may not be accessible to the engineers concerned with PCDB maintenance, and do not generate sufficient information for PCDB entry. A simple, formal method is introduced in this chapter that uses a set of matrices to assist an engineer in systematically transforming information about product characteristics into a form for entry into a database. A positive aspect of this method is that the matrices themselves provide insight into the product characteristics that should be tracked and the manufacturing processes that create them. This new method is discussed in detail in Section 2.2.3.

The final barrier preventing the effective usage of process capability data is the creation of a PCDB user interface that is easy to use. The user interface must facilitate the retrieval of data. Engineers do not want to learn a new language or program to be able to use their company's database; therefore, a command prompt or text based user interface is undesirable. The user interface must also allow the database to be fully exploited without being complicated. A graphical user interface is proposed in this chapter that is easy to use without training. The design is intentionally simple looking to make engineers comfortable with it; however, the underlying structure is more complex. A menu system is driven by data in the database to dynamically update as selections are made. This dynamic behavior helps guide the engineer and prevents faulty or empty queries. The user interface is presented and discussed in Section 2.3.

2.2 Indexing Scheme

An indexing scheme is simply a means of using one thing to represent another. For example, a social security number is used to identify a person. In the case of a PCDB, the indexing scheme defines how process capability data is referenced in the database. For example, a component dimension that is measured and tracked when it is produced must be identified with an indexing scheme so that a person looking for the data at a later time can find it.

One of the most important aspects of creating a PCDB is the design and creation of the indexing scheme since it directly affects the ease with which an engineer can locate and retrieve data of interest. If the indexing scheme is poorly designed, it presents a barrier to engineers who want to access data. For example, several companies store process capability information based on part numbers, making it difficult for a design engineer to find relevant data without a familiarity with existing engineering documents (Tata and Thornton 1999).

The development of the ideal indexing structure is non-trivial. It must be detailed enough to fully describe a physical characteristic of a component. However, a too detailed structure will reduce the usability of the PCDB through increased search times and unnecessary complexities. The indexing scheme also needs to be understood by design engineers. Typically designers think in terms of "features" and "component attributes." "Features" describe the geometrical aspects of a component, e.g., rod of length 50 mm, and "component attributes" describe the overall characteristics, such as material, how it is manufactured, etc.

In addition, the indexing scheme must be designed so it is expandable and easily maintained. There are thousands of dimensions and processes that a manufacturing organization may track. To make any process capability database usable and efficient, the commonalities between these dimensions should be exploited. This is similar to the concept of Group Technology (Raffish 1985), which sorts components into groups to exploit their similarities to reduce manufacturing costs. For example, there may be many different cases where faces are ground. This process may be used to generate a variety of parts and or features. However, if a designer were able to understand the general capability of "face grinding," through analyzing a broad set of data, he/she is more likely to accurately specify the component's tolerances.

The Feature Code mentioned earlier in the background section contains three attributes to describe a key characteristic. The attributes are feature, material, and process (Thornton and Tata 2000). When planning the implementation of a system for a large scale manufacturing company, several limitations to this structure were identified. As a result, the indexing scheme is expanded to include a fourth attribute, geometry, to improve both the implementation and usability of the PCDB. This new combination of attributes is called the Attribute Combination Descriptor (ACD).

2.2.1 Attribute Combination Descriptor (ACD)

An Attribute Combination Descriptor (ACD) represents a key characteristic's feature, geometry, material, and the process that created it. Each attribute of the ACD can have several sub-attributes to provide more detail about a key characteristic. A specific ACD is not unique to a specific key characteristic. Multiple key characteristics can have all attributes in common; therefore, the same ACD describes each of them. Figure 2-1 below represents a generic ACD with three levels for each attribute. The number of levels in an ACD depends on the manufacturing company's needs, but three levels usually provide sufficient detail to describe a key characteristic.

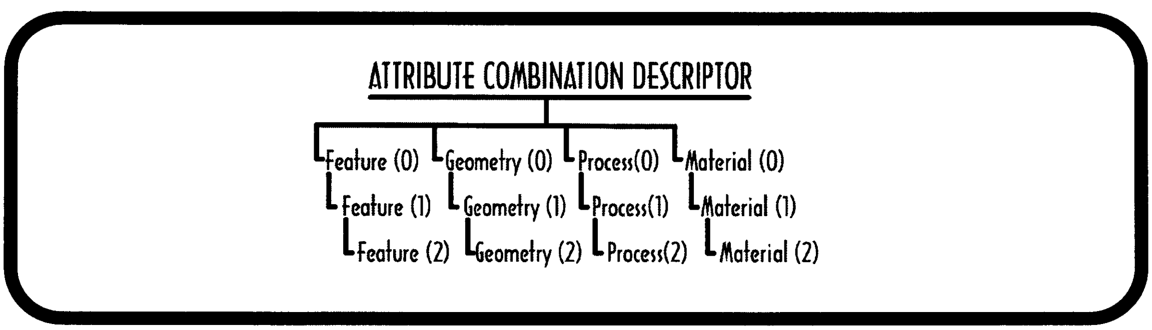

Figure 2-1: General Attribute Combination Descriptor

The database hierarchy is structured based on the ACD. A sample database hierarchy is shown in Figure 2-2. The graphic shows the top-levels of Feature, Geometry, Material, and Process. Each of these then has lower level branches extending out of them. For example, thrust is a type of face, which is a Feature.

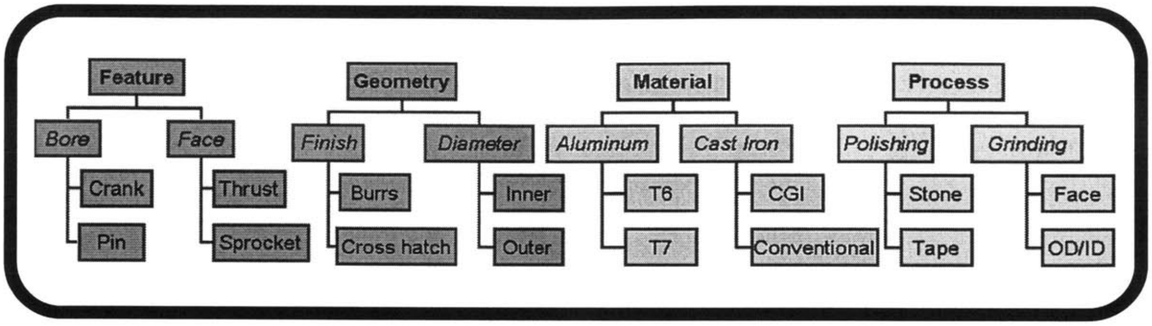

Figure 2-2: Sample database hierarchy

The term feature attribute is used to describe a physical entity or aspect of a component. This is usually something that is dimensioned or has other measurable properties. An example of a feature attribute is a FACE with possible sub-attributes of FRONT and REAR. The geometry attribute describes a property of a feature attribute, e.g., the face could have a geometry attribute of POSITION. The feature attribute differs from the geometry attribute in that it represents a three-dimensional entity that is measured while the geometry attribute represents the measurement itself. The process attribute depicts the manufacturing process used to create the key characteristic. The material of the component is revealed by the material attribute. For example, the component could be made of a type of ALUMINUM from a MILLING process. A visual example of attributes is displayed in Figure 2-3 while an example of the corresponding ACD is graphically represented in Figure 2-4. The picture in Figure 2-3 would appear on a detailed drawing of the part.

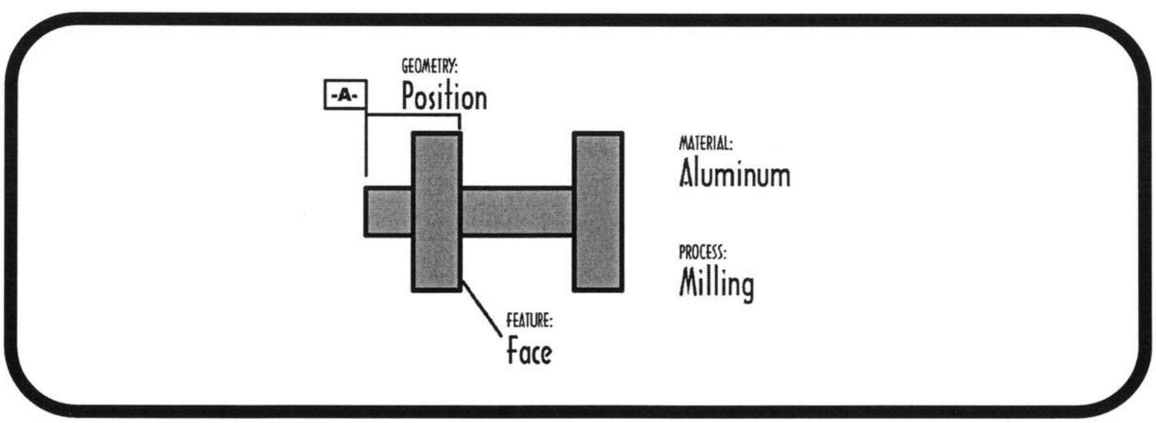

Figure 2-3: Visual example of attributes

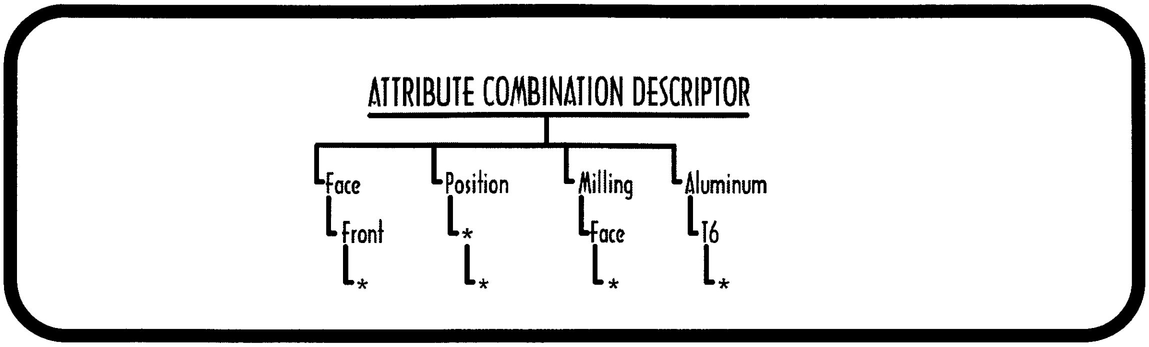

Figure 2-4: Example Attribute Combination Descriptor

Each record in a PCDB contains fields for every attribute in the ACD, i.e., a feature(0) attribute field, a feature(1) attribute field, etc. These fields allow queries to be made by specifying a partial or complete ACD. A complete ACD has at least one level specified for each attribute while a partial attribute has one or more attributes as wildcards. The use of a partial ACD allows for a broad search of a PCDB, while specifying multiple attributes and sub-attributes generates more precise query results. As a result of a query, the PCDB outputs records of data from components that match the query. The records contain process capability data as well as information such as part name, plant location, production line, collection date, and other pertinent data (See Section 2.4). Based on this descriptive data, it is up to the designer to determine if the query results are relevant.

2.2.2 Benefits of the Attribute Combination Descriptor

The three attributes of feature, material, and process are obviously required to identify a characteristic of a component. Although less obvious, the geometry attribute is a powerful inclusion. Adding geometry as an attribute is a natural extension to represent information about a feature and is loosely based on Geometric Dimensioning and Tolerancing (GD&T). Some of the geometry attribute terminology is borrowed from GD&T such as concentricity, parallelism, runout, etc.

An obvious advantage to using the ACD to index data instead of by part number is that part number changes have no effect on how the data is stored in the database. Three major benefits exist for using the ACD instead of just the feature, material, and process attributes where geometry is nested under the feature attribute. They include simplifying the PCDB structure, allowing for additional types of queries to be made, and simplifying expansion of the PCDB.

The inclusion of the geometry attribute in addition to the feature attribute allows for the hierarchical structure to be substantially flattened and significantly reduces redundancy within the database by moving branches from the feature tree to the geometry tree. This creates a clearer perspective for the unfamiliar engineer to perform queries. It also improves an algorithm's speed when searching the database because the branch depth is decreased, typically in half.

Assume a PCDB contains a hundred feature attributes and fifty geometry attributes. If geometry is included in the feature attribute tree and half the geometry attributes are applicable to half the feature attributes, there would be a hierarchy tree with 1,250 branches. Separating the feature and geometry attributes creates two hierarchy trees, one with 100 branches and one with 50 branches. This is a simplification of 88%. Figure 2-5 and Figure 2-6 below visually compare two simple sample hierarchal structures. Figure 2-5 shows what a hierarchy looks like when the geometry attribute is included as branches of the feature attribute while Figure 2-6 shows the same information but uses two, simpler hierarchy trees. A schematic showing a side by side comparison is showed in Figure 2-6.

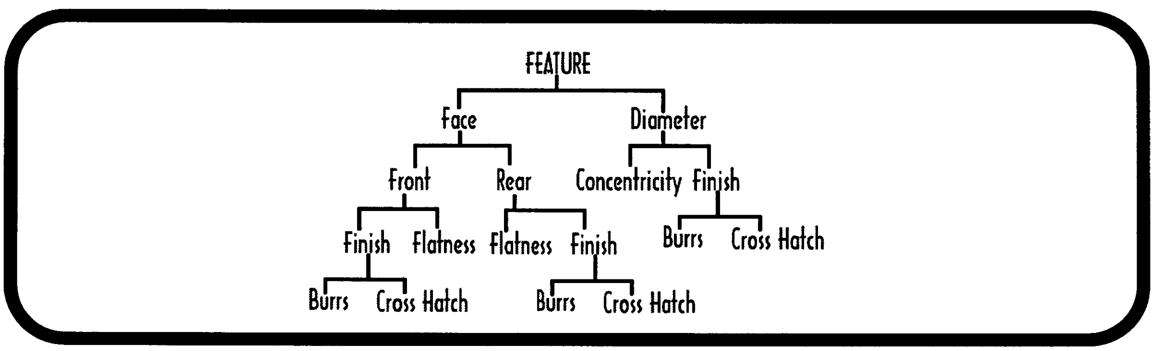

Figure 2-5: Hierarchy tree with Feature only

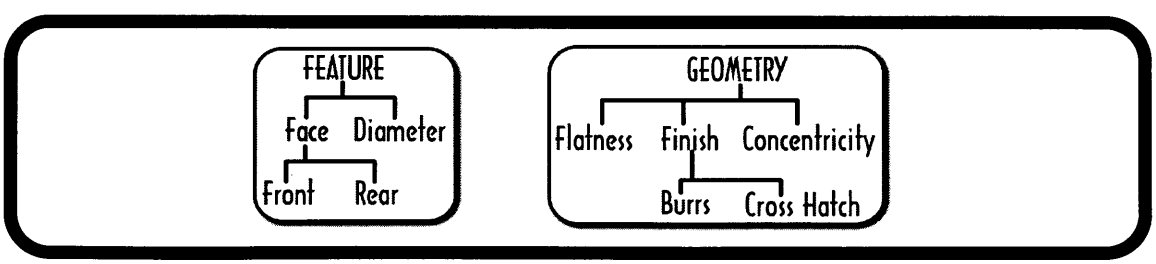

Figure 2-6: Hierarchy trees of Feature and Geometry

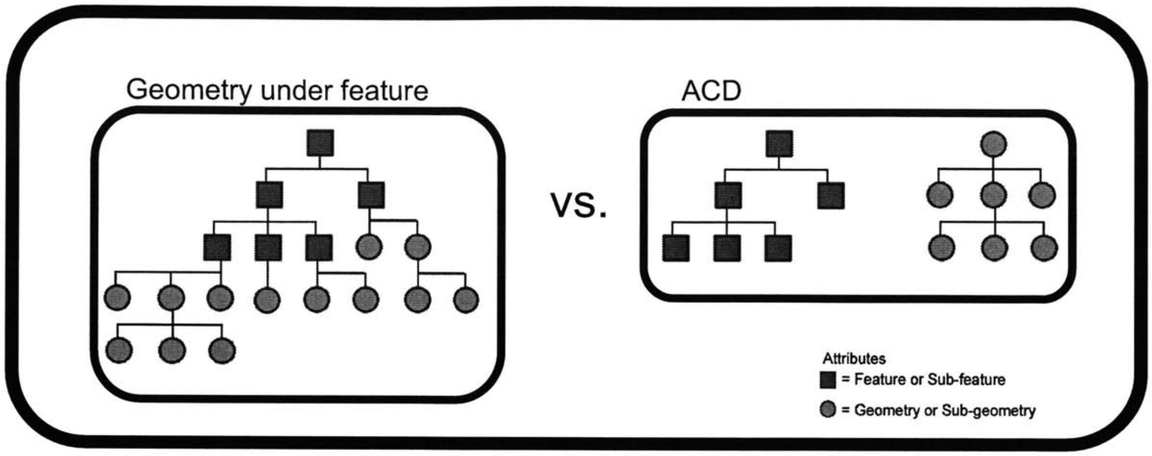

Figure 2-7: Side by side comparison of database hierarchies

The second advantage is that a geometry attribute allows a query of all instances of a specific geometry. For example, if an engineer wanted to compare the process capability data pertaining to surface finishes of aluminum parts, all records with the material attribute of aluminum and geometry attribute of surface finish could be retrieved. If the geometry attribute were nested under the feature attribute, it would be a tedious process to query all surface finish entries, as an engineer would have to locate surface finish entries deep in the branches of the feature tree and perform separate queries on each instance. This concept is graphically illustrated in Figure 2-8 where the black circle represents the geometry attribute of interest in the two data hierarchies. Clearly, the geometry attribute is shown to appear repeatedly in the left graphic while it only appears once in the right graphic.

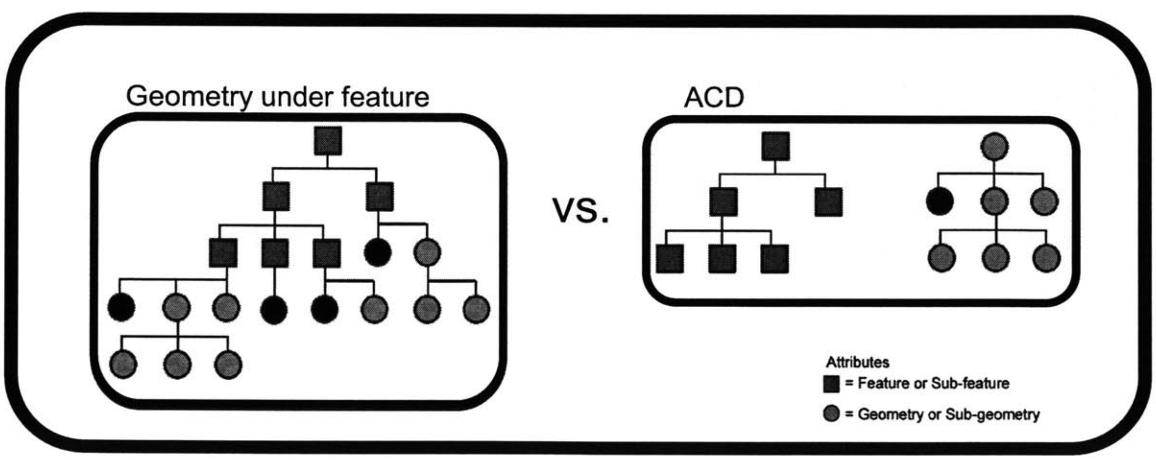

Figure 2-8: Query a single instance of geometry instead of many

The third advantage of having a separate geometry attribute is that it simplifies PCDB expansion. When a new key characteristic is labeled, its attributes may not exist in the PCDB listing. If this happens for a material or process, the expansion is as simple as adding a new branch or sub-branch to the appropriate tree. If the geometry attribute is included with the feature attribute, the new geometry attribute listing may apply to multiple feature branches; therefore, each applicable feature branch must be modified. Also in this case, if the feature attribute is a new listing, a new feature branch must be added and should include all applicable geometry attributes as lower level branches. The number of applicable geometry attributes can make this a difficult task. Expansion is much simpler if feature and geometry are separate attributes. Adding a single branch or sub-branch to the appropriate tree completes the task. This concept is illustrated in Figure 2-9 where the white circles are extensions of an existing hierarchy. The left graphic represents the structure when geometry is included as a sub-attribute to feature and multiple updates are required as represented by the three white circles. The hierarchy on the right uses the ACD and only one update is necessary, regardless of the number of branches already existing.

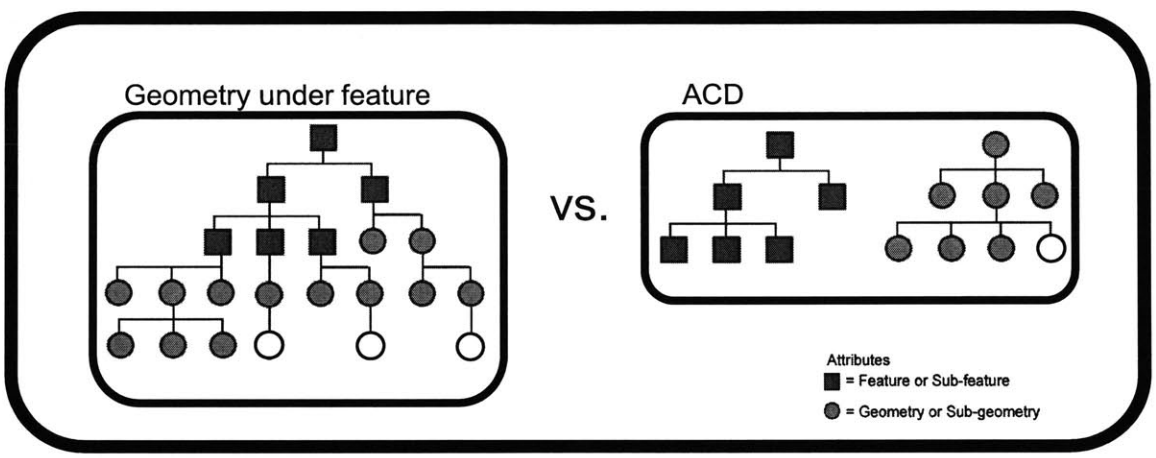

Figure 2-9: Illustration of extensibility

2.2.3 Method to Index Design Characteristics

Prior to being entered and tracked in a PCDB, a component's key characteristics must be indexed into ACDs. Indexing is the process of assigning attributes for feature, geometry, material, and process for each key characteristic. In most cases, sub-attributes will be required to fully represent a key characteristic. The required indexing is often done in an *ad hoc* manner. Using the indexing scheme to construct ACDs is very important. If the indexing does not match the PCDB, the ACDs will be incomplete and the data will be difficult, if not impossible to retrieve. If the terminology is ignored, retrieving the data will be confusing, as entries will appear to be redundant or nonexistent. In this section, a tool is introduced to assist engineers in the indexing of product data into Attribute Combination Descriptors. This tool is a set of matrices called an ACD Matrix Set and is comprised of five individual matrices. The ACD Matrix Set is unique to a company because it only needs to contain listings of features, geometries, materials, and processes that are specific to the company.

A difficulty in indexing key characteristics into ACDs is maintaining consistency to avoid redundant entries or misleading representations. This happens when two different terms are used to describe the same attribute e.g., SURFACE and FACE. In this example, the person indexing a "flat surface" into an ACD could rightfully use FACE, SURFACE, FLAT SURFACE, or many other possible terms to describe the feature attribute. This freedom leads to multiple terms used to describe the same attribute, which in turn makes querying the PCDB very difficult. If an engineer querying the PCDB selects SURFACE, the results would definitely be incomplete, unknowingly ignoring FACE and FLAT SURFACE entries.

There are several reasons the consistency within the PCDB is difficult to maintain. Numerous sources of information, e.g., engineering drawings, production data, material specifications, process plans, engineers' input, etc., are used to categorize key characteristics; however, each of these sources uses a slightly different set of terminology. Also, an engineer unfamiliar with a PCDB may generate ACDs that are incomplete and use unique terms. These issues account for two or more unique text descriptions being used to describe the same key characteristic. A simple example is "-A- to rear flange face" versus "face position relative to -A-."

2.2.3.1 Assistance Tool

A tool to assist with indexing is the ACD Matrix Set. An ACD Matrix Set is comprised of five individual matrices. There is one matrix for the feature attribute (F-matrix), one for the geometry attribute (G-matrix), one for the material attribute (M-matrix), and two for the process attribute (P-matrix1 and P-matrix2). The goal of the ACD Matrix Set is to guide an engineer through the indexing of a component's key characteristics and the creation of ACDs for them. The matrices allow key characteristic information to be mapped to an attribute for ACD generation. Templates need to be created for the F-matrix, G-matrix, and M-matrix before the ACD Matrix Set can be used. (The creation of the P-matrixes differs slightly and is discussed in a later section.)

A PCDB contains a listing of attributes used to describe the key characteristics. These attributes are related to each other in a hierarchal fashion as seen in Figure 2-6 so that each attribute has a "level" associated with it. The "level" identifies where the attribute or sub-attribute lies in the tree and corresponds to the attribute level within the ACD (see Figure 2-1). For example, FRONT or REAR, if applicable, would be placed into the feature(1) attribute of an ACD. The listing of attributes and their levels is used to form the top two rows of the matrix templates. A sample matrix template is shown in Figure 2-10.

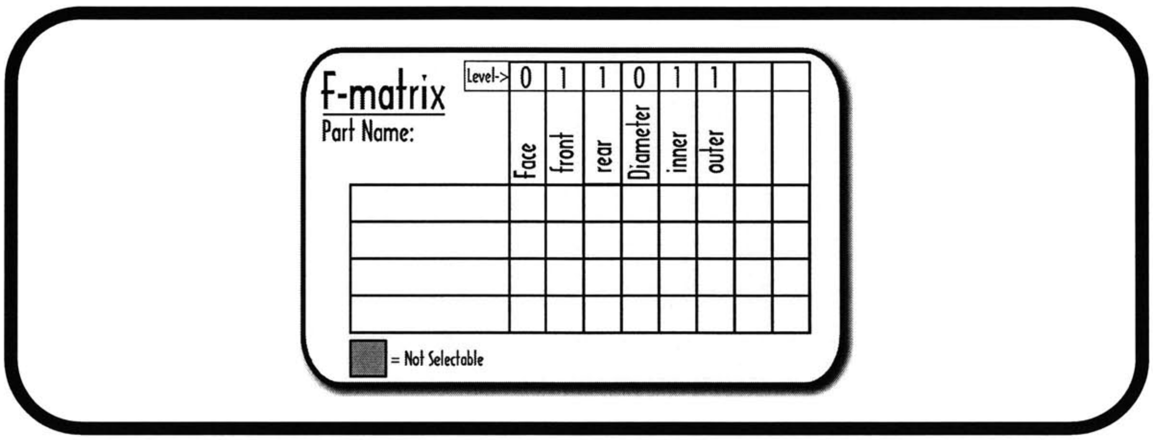

Figure 2-10: Example F-matrix template

Algorithms are built into the template to prevent generating infeasible ACDs. The algorithms are based on the hierarchal relationships of the PCDB and "shade out" sub-attributes or subsequent attributes that would not be relevant. For example, if FACE is the feature attribute selected, then lower level feature attributes FRONT and REAR would be selectable while INNER would not.

The matrix templates can be created by a PCDB administrator or generated automatically from the PCDB since they rely entirely on the hierarchal attribute listing. They can be implemented using standard spreadsheet software.

A component with key characteristics has an ACD Matrix Set associated with it. The ACD Matrix Set for a specific component is constructed by starting with the templates. The key characteristics of the component are then placed into the leftmost column of the F-matrix, G-matrix, and M-matrix. Therefore, a particular row contains all the information for a particular key characteristic while each column represents a particular attribute or sub-attribute.

Example ACD Matrixes are shown to demonstrate how the sample key characteristics listed in Figure 2-11 and shown in Figure 2-12 are indexed. The component can be thought of as an overly simplified crankshaft. The information in Figure 2-11 could be obtained from the product specification, and the picture in Figure 2-12 could be retrieved from the component's detailed drawing.

Part Name: Two Tier Rod

Key Characteristic List

Key Characteristic 1 (KC1): Diameter surface finish

Key Characteristic 2 (KC2): Surface to end

Key Characteristic 3 (KC3): Back face to end

Figure 2-11: Sample key characteristic listing

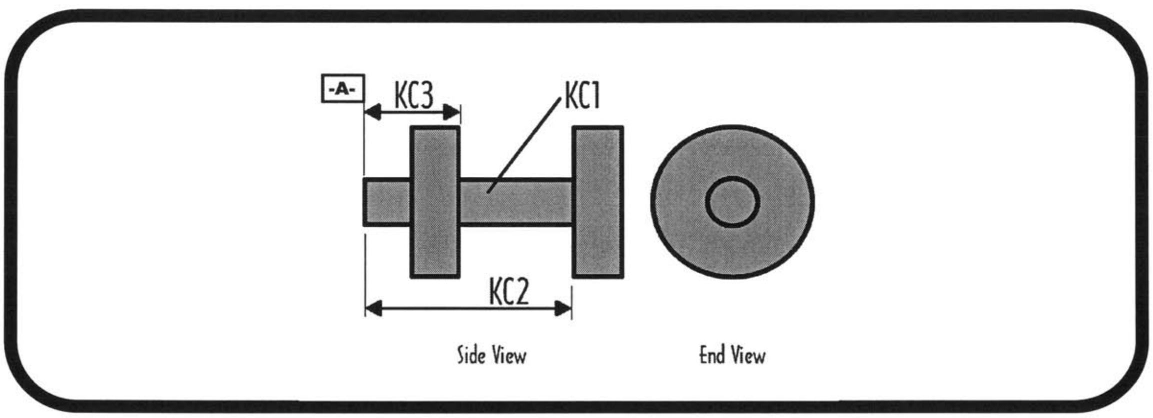

Figure 2-12: Visual sample key characteristics

The feature attribute is the first matrix to be populated. The engineer fills in the matrix row/column intersections (cells) with a number between 0 and 1 to represent the probability that the key characteristic (row) has the listed attribute (column). If there is no chance the attribute in a certain column describes the attribute of the key characteristic, the cell value is 0 or left blank. The cell value is 1 if the column attribute definitely describes the key characteristic. A number between 0 and 1 is used when an engineer is not absolutely confident of the column attribute of the key characteristic.

The F-matrix in Figure 2-13 below is partially populated. The feature attributes for KC1 and KC2 are determined while the level 1 attribute for KC3 has yet to be marked. Notice the infeasible feature attributes are “shaded out” once a selection has been made.

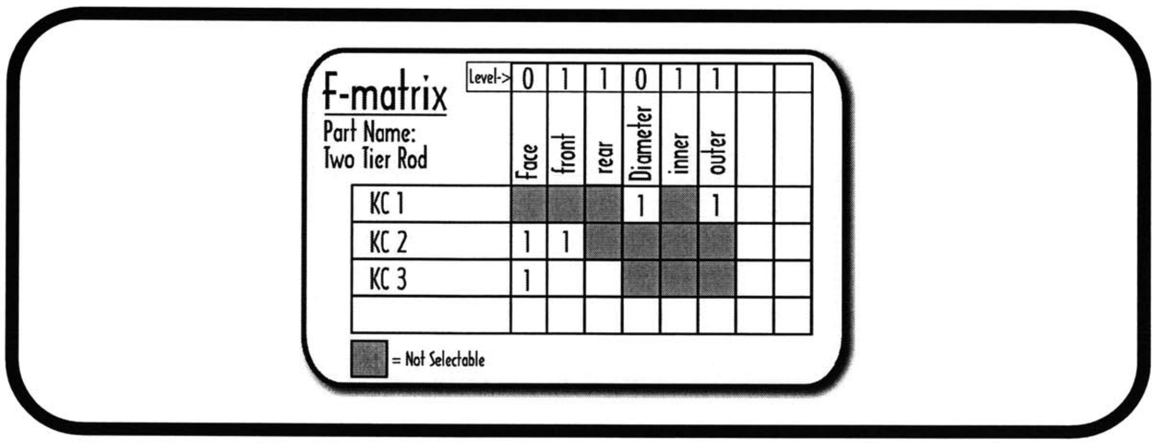

Figure 2-13: Partially populated F-matrix

Once the final selection is made for key characteristic 3, the F-matrix is fully populated and is shown in Figure 2-14.

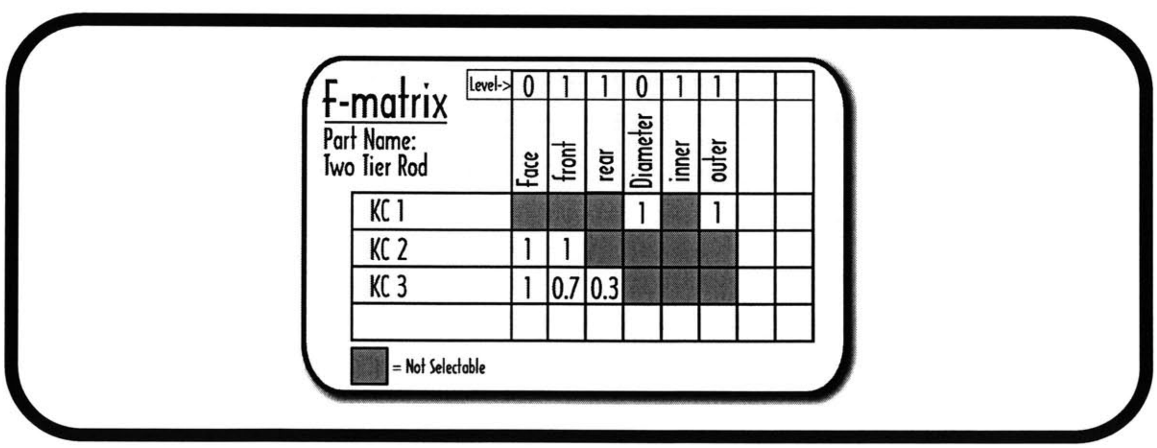

Figure 2-14: F-matrix populated

Notice the engineer populated the type of face for KC3 using "fuzzy" numbers. The engineer is unsure if a front face is one that points inward or toward datum A. In this case, the engineer is 70% confident that a front face is a face that points inward. The use of the "fuzzy" numbers will generate multiple ACD sets. It is then the engineer's responsibility to confirm the inward pointing face is defined to be a front face, at which time the appropriate ACD can be selected with complete confidence.

The ability to enter "fuzzy" numbers allows an engineer to continue the population of the ACD Matrices even when uncertainty arises. Since the engineer is not forced to make an absolute choice, an incorrect selection is less likely.

The geometry attribute is populated after the F-matrix since geometry may depend on feature, i.e., some geometry attributes are infeasible with some feature attributes, e.g., a FACE does not have a DEPTH. Shading out the infeasible geometry selections represents these relationships. The G-matrix in Figure 2-15 below demonstrates this concept by shading out the infeasible cells.

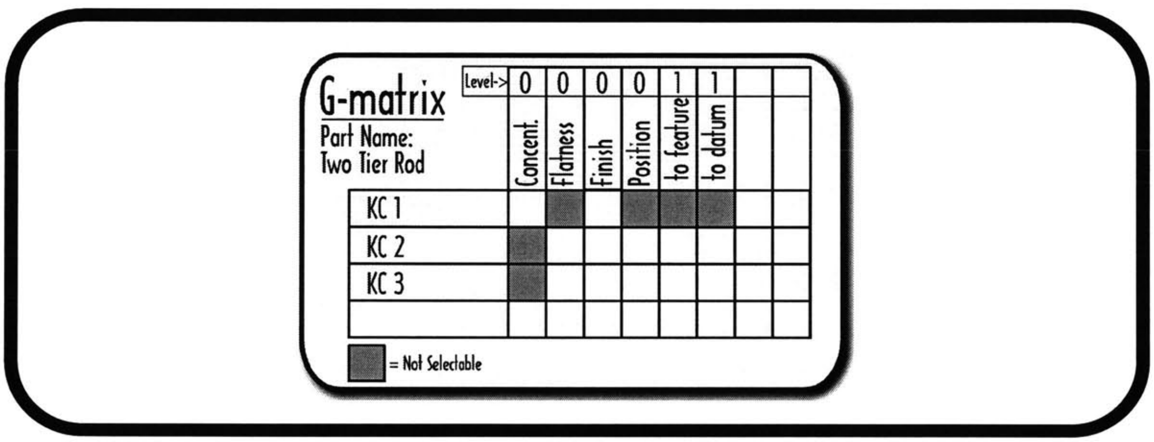

Figure 2-15: Initial G-matrix after F-matrix is populated

The same numerical value scheme used for the F-matrix is used for the G-matrix. The end result is shown in Figure 2-16 where the G-matrix has been populated.

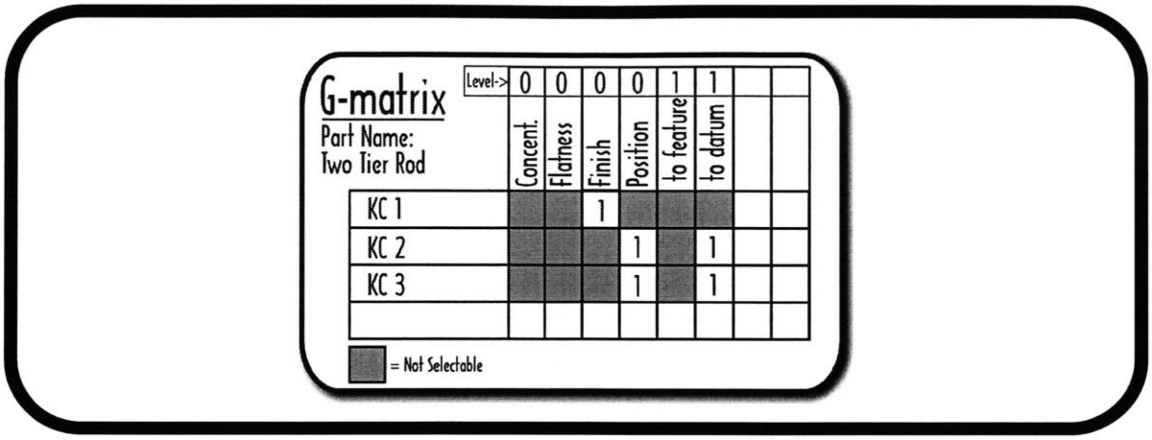

Figure 2-16: G-matrix populated

The material attribute is typically the easiest to index. The material information can usually be obtained from the engineering drawing. When not explicitly stated, the engineering drawing references a material specification. It may be necessary to review the specification to fully populate the M-matrix. For a component with uniform material properties, the M-matrix can be simplified to a single row. Figure 2-17 and Figure 2-18 show an unpopulated simplified M-matrix and the same matrix populated, respectively.

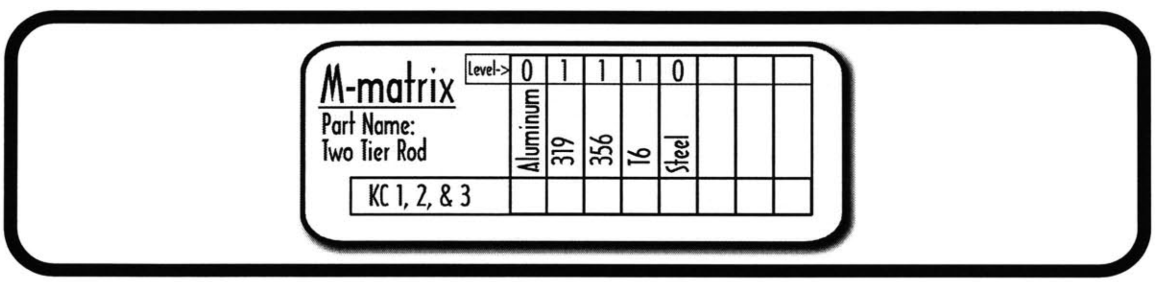

Figure 2-17: Blank M-matrix

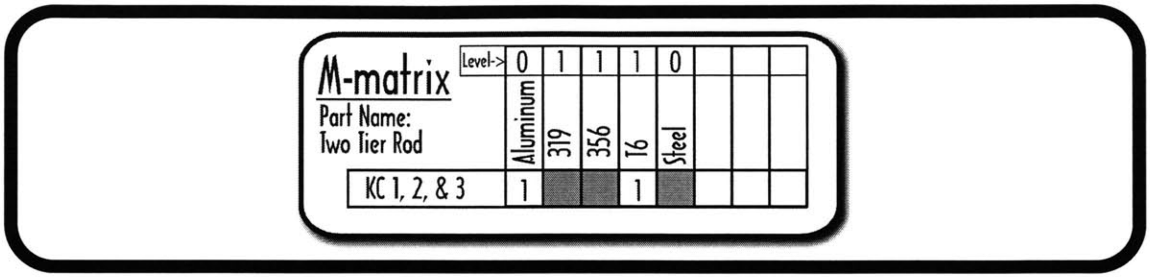

Figure 2-18: M-matrix populated

A firm that has truly integrated the use of its PCDB into its design practices may include the material index in the material specification once it has been determined. This allows instant access to the information required to populate the M-matrix during future indexing on components that reference the same material specification.

When indexing the process attribute, the processes that contribute to a key characteristic's creation must be identified. This requires two separate matrixes, the P-matrix1 and P-matrix2. The P-matrix1 decomposes the production operations into processes a PCDB recognizes. The top two rows contain the processes listed in the PCDB and their levels while the leftmost column contains a list of the operations used to produce the component. The matrix is populated in a manner similar to the previous attribute matrixes to map which processes are used for each operation. Figure 2-19 shows a partially populated P-matrix1. The fully populated P-matrix1 is shown in Figure 2-20.

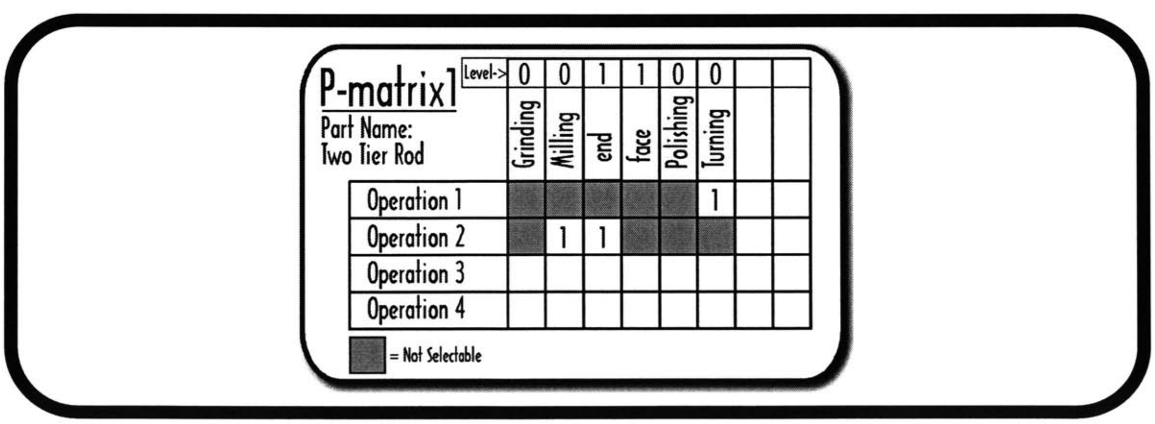

Figure 2-19: P-matrix1 partially populated

P-matrix1

Part Name:
Two Tier Rod

Level->	0	0	1	1	0	0		
	Grinding	Milling	end	face	Polishing	Turning		
Operation 1						1		
Operation 2		1	1					
Operation 3	1							
Operation 4					1			

= Not Selectable

Figure 2-20: P-matrix1 populated

The top row of the P-matrix2 contains the operations. This allows an engineer to quickly select the operations affecting the key characteristic using the production plan, since it is based on operations, not processes. Similar to the other attribute matrixes, a value of one is placed in the cells where an operation contributes to the creation of a key characteristic. Figure 2-21 is an example of a populated P-matrix2.

P-matrix2

Part Name:
Two Tier Rod

	Operation 1	Operation 2	Operation 3	Operation 4				
KC 1	1	0	1	1				
KC 2	1	1	0	0				
KC 3	1	1	0	0				

Figure 2-21: P-matrix2 populated

Now that the P-matrix2 is populated, a relationship is created that reveals the processes used to create each of the key characteristics.

There are four incentives for using two process attribute matrixes. The first is the fact each operation only has to be indexed once, regardless of how many key characteristics it affects. If there were a sole P-matrix with the key characteristics listed in the leftmost column and the processes along a top row, the indexing would be repeated for every row. Having two P-matrixes is very helpful when the number of key characteristics, e.g., two hundred, dwarfs the number of operations, e.g., ten.

The fact serial processing can be easily represented in the P-matrix2 is the second incentive. If there were only one P-matrix, the key characteristics affected by multiple processes would need to be identified beforehand and listed multiple times, one for each operation.

The third incentive for using two P-matrixes is that it eases the extraction of information from the production plan since it usually lists key characteristics by operation. This information allows the P-matrix2 to easily be populated. Also, the production plan contains details about each operation that makes populating the P-matrix1 simple.

The task of indexing a newly identified key characteristic is simplified when two P-matrixes are used. The new key characteristic only requires the operations that make it be known, which is available in the production plan. Since the operations are already indexed in the P-matrix1, the most difficult work has been completed and documented.

2.2.3.2 Creation of Indexed Data

Once the ACD Matrix Set is populated, the ACDs are generated. In the case where only zeros and ones are used, ACDs are generated with a high confidence level, i.e., the engineer indexing the key characteristics was 100% when selecting every attribute. In the case when an engineer is not 100% confident of the attribute to assign a key characteristic, fuzzy numbers between zero and one are issued, and multiple sets of ACDs are created for the key characteristic in question. The ACDs are rank ordered by a confidence level that is proportional to the fuzzy value. The engineer then performs a manual check to determine and select the proper ACD from those generated. The ACDs generated from the example ACD Matrix Set are listed in Table 2-1. These are the ACDs for the sample Key Characteristics listed in Figure 2-11 and shown in Figure 2-12.

Table 2-1: Generated ACDs for component in Figure 2-12

	ACD	Feature	Geometry	Material	Process
KC1 100% Set1	1	Diameter outer	Finish	Aluminum T6	Turning
	2	Diameter outer	Finish	Aluminum T6	Grinding
	3	Diameter outer	Finish	Aluminum T6	Polishing
KC2 100% Set1	1	Face front	Position to datum	Aluminum T6	Turning
	2	Face front	Position to datum	Aluminum T6	Milling end
KC3 70% Set1	1	Face front	Position to datum	Aluminum T6	Turning
	2	Face front	Position to datum	Aluminum T6	Milling end
KC3 30% Set2	1	Face rear	Position to datum	Aluminum T6	Turning
	2	Face rear	Position to datum	Aluminum T6	Milling end

2.2.3.3 Interpretation of Indexed Data

The ACDs listed in Table 2-1 provide a complete set of information for KC1 and KC2 so that measured data for these key characteristics can be easily input and retrieved from a process capability database. The two ACDs for KC3 require follow-up by the engineer to confirm the correct type classification of the face associated with KC3. Once the definition for the type of face is fully understood, the engineer selects the appropriate ACD for KC3 and deletes the other. Now each key characteristic has the appropriate ACDs defined.

Notice KC1 has three corresponding ACDs while KC2 and KC3 have two ACDs. This is because an ACD is generated for each combination of attributes. KC1 has three combinations since three processes are required to create it. The multiple ACDs for a key characteristic allow data to be stored and retrieved for each process that affects it.

There are several distinct advantages to using an ACD Matrix Set. The first is the reliability of the method. Using an ACD Matrix Set ensures the correct ACDs are generated. Robustness is another advantage. The method allows for a wide range of input sources to be used effectively without causing a disruption in the ACD generation process.

The ACD Matrix Set is also an effective knowledge management tool since it encapsulates more than just a list of key characteristics. It captures and documents the relationships attributes have among each other, i.e., it is easy to discern what processes create a geometry attribute via a review of the ACD Matrix Set. This is especially useful when more than one process affects the creation of a key characteristic.

Detailed information is also revealed about which attributes have variation that requires management, e.g., a quick glance at a G-matrix reveals the geometry attributes that are in the majority of key characteristics. This information is helpful in identifying overlooked key characteristics or specifying key characteristics on new designs, i.e., examine designs for attributes similar to the ones commonly listed. This same information can be used to leverage design and manufacturing resources to manage variation more effectively by tackling the "big hitters," e.g., if flatness appears frequently in the G-matrix, resources can be directed to improve it.

2.3 User Interface

The user interface of a Process Capability Database is important because it is the interactive portion of the database an engineer sees and determines the usability of the database (Powell 2000). The Process Capability Database developed for this thesis is accessible via the internet through a standard web browser. This section describes the various aspects of the user interface and shows screen captures from the beta prototype version.

Survey results (see Appendix B) revealed that any database developed needs to be easy to learn and simple to use. This revelation led to the development of a Graphical User Interface (GUI) for the database. A Graphical User Interface does not require the learning or memorization of text command lines. One common method for querying a database is to use Structured Query Language (SQL), which is a standard interactive and programming language for getting information from and updating a database. The use of a GUI eliminates the need for engineers to know SQL, which is beneficial since most are unfamiliar with the programming language.

The Graphical User Interface went through several revisions. First, a paper prototype was created to explore the necessary items and their placement relative to each other. Next, an alpha prototype was created in Microsoft® Excel. This prototype was run locally on a desktop computer and was interactive. It was used to solicit feedback from engineers as to their likes and dislikes. The feedback was mostly positive. One suggestion was to provide the ability to query data from specific manufacturing plants. Another suggestion was to allow the engineer to select the query results displayed. Both suggestions were implemented in the beta prototype through the use of check boxes. The beta prototype was created using Hypertext Markup Language (HTML) and Active Server Pages (ASP) for the user interface and a Microsoft® Access database as the backend. Using these technologies allows the database to be stored on a web server and accessed via the internet using a standard web browser. The beta prototype was populated with "made up" data for testing purposes. It was never used to store actual measurements from a manufacturing plant.

An important concern raised by engineers was the security aspect of the database. Therefore, the database is password protected. The introduction screen is displayed in Figure 2-22. The username and password have a hidden security level associated with them. The security level allows different abilities for different users. For example, a security level of five allows a user to query the database but not to add new manufacturing plants to the list; a task reserved for someone with a higher security level. Once logged in, the person's security level is displayed as seen in Figure 2-23.

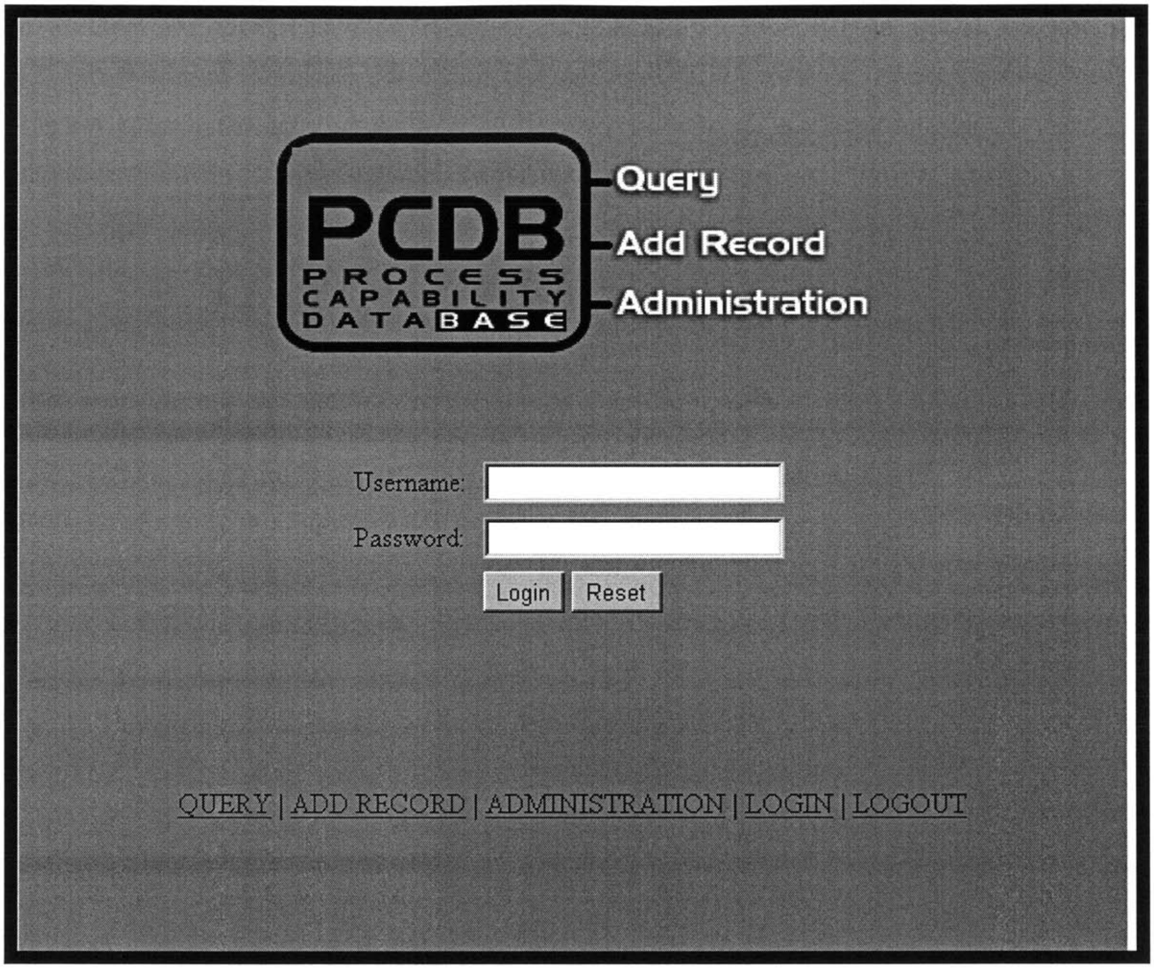

Figure 2-22: Introduction screen to the PCDB

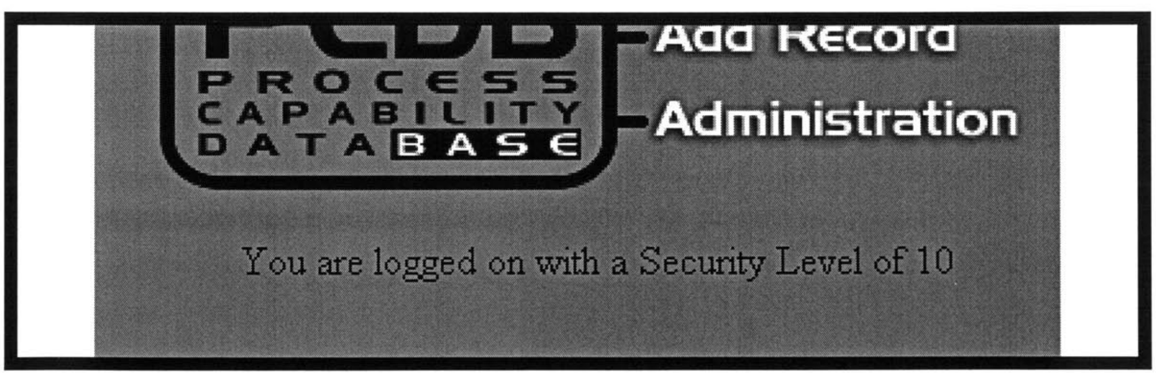

Figure 2-23: Displayed security level once logged on

The query page is the most frequently used. A screen capture of the query page is displayed as Figure 2-24. The GUI uses drop down menus to allow selection of the feature, geometry, material, and process attributes. Checkboxes allow for any number of manufacturing plants to be included. Clicking the **All Plants** checkbox places checks in all the checkboxes associated with the various plants. Subsequently clicking a checkbox associated with any specific plant unchecks the **All Plants** checkbox and the specific plant's checkbox. The next section of the query page is for inputting information related to the dimension of interest. The mean, tolerance limits, and process capability indexes can be used to limit the query results. The last section is for specifying what data to return. Checkboxes are again used to allow for the selections to be made. Any number of checkboxes can be selected. Once the selections are made, the submit query button is clicked to process the query.

The user interface is flexible. A user can leave as many of the fields unpopulated and drop down boxes set to "all" to run a query that returns a very general set of data. The user can then return to the query page and narrow the search by specifying values for more fields.

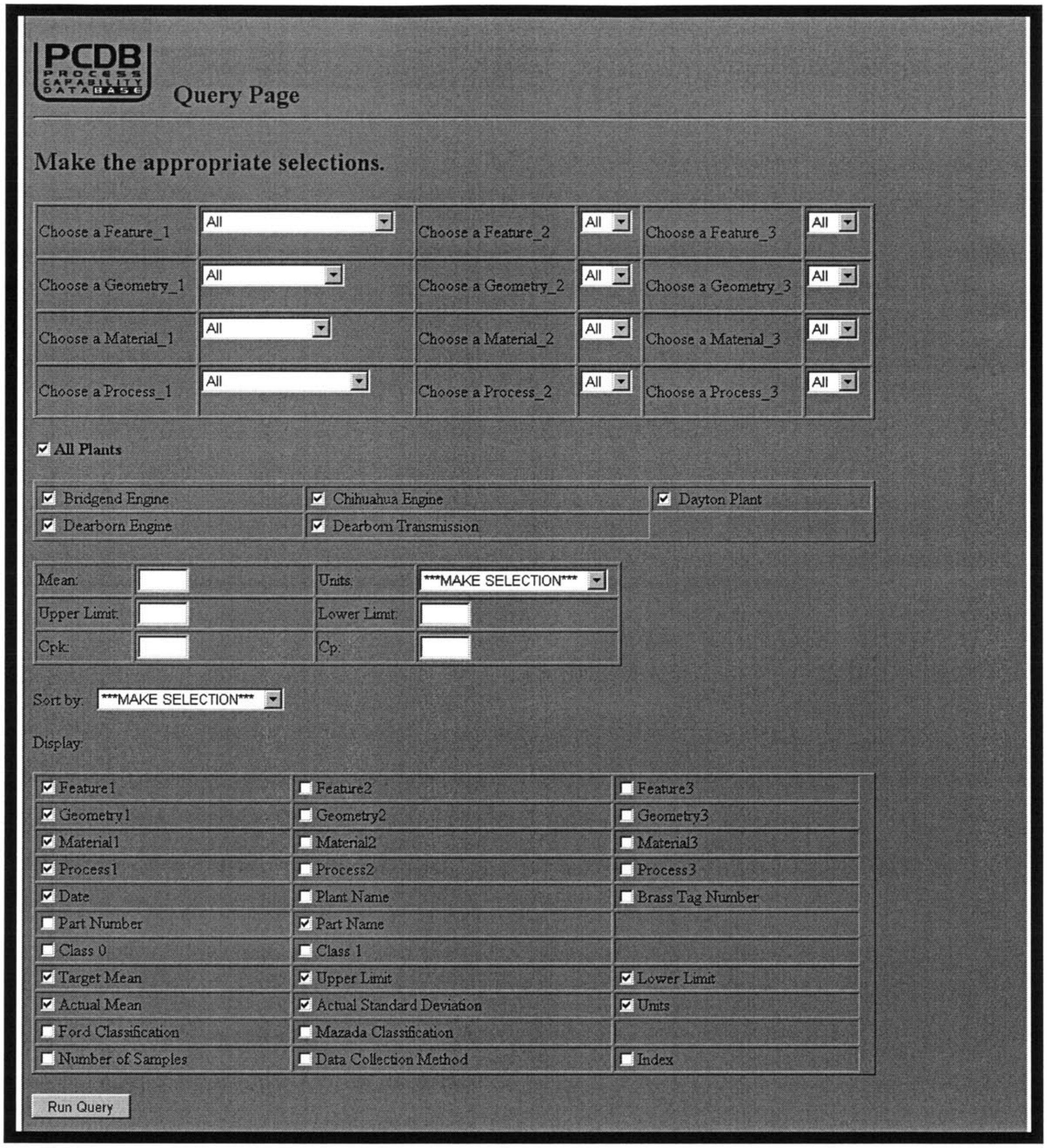

Figure 2-24: Screen capture of query page

One aspect of the user interface that makes it easier to use is the real-time updating of subsequent attribute menus. For example, before a Material_1 attribute is specified, the Material_2 drop down menu is not populated as seen in Figure 2-25 below. Once aluminum is selected as shown in Figure 2-26, the second level material attribute menu is updated to show possible types of aluminum as seen in Figure 2-27. If the Material_1 selection is changed, the Material_2 menu reverts back to "all," and its contents are updated to reflect possible sub-attributes for the newly selected material. The ability for the menu contents to change in real-time makes it easier for engineers to make lower level attribute selections. Once the highest level attribute is selected, the engineer does not have to scroll through every possible second level attribute, just those that are possibly applicable.

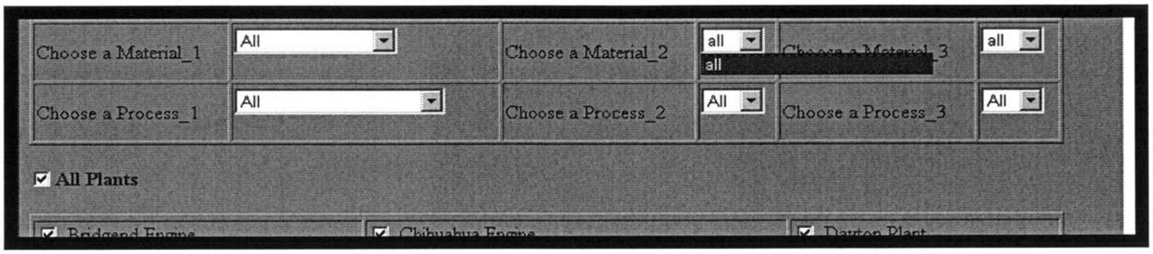

Figure 2-25: Material_2 menu items before Material_1 is specified

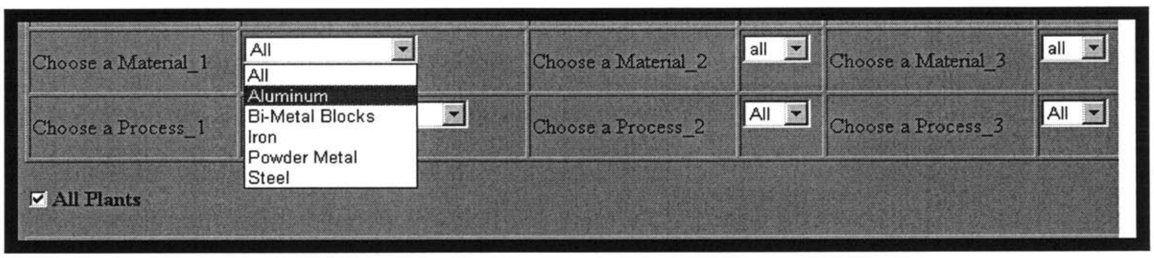

Figure 2-26: Aluminum being specified for Material_1

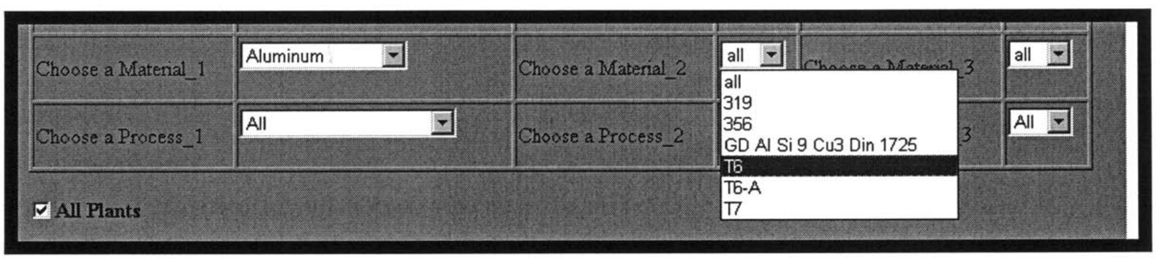

Figure 2-27: Material_2 selections once aluminum is specified

The results of the query are displayed in tabular form. The engineer can select what field to use to sort the results. The final version of the PCDB will also display the query results graphically based on the work done by Thornton and Tata (Thornton and Tata 2000). Displaying the query results graphically gives the engineer the ability to make quick comparisons and decisions regarding feasibility. The beta prototype developed for this thesis did not have the capability to present graphical results.

Another important aspect of the user interface is the page used to add a record to the database. A screen capture of the add record page is shown as Figure 2-28.

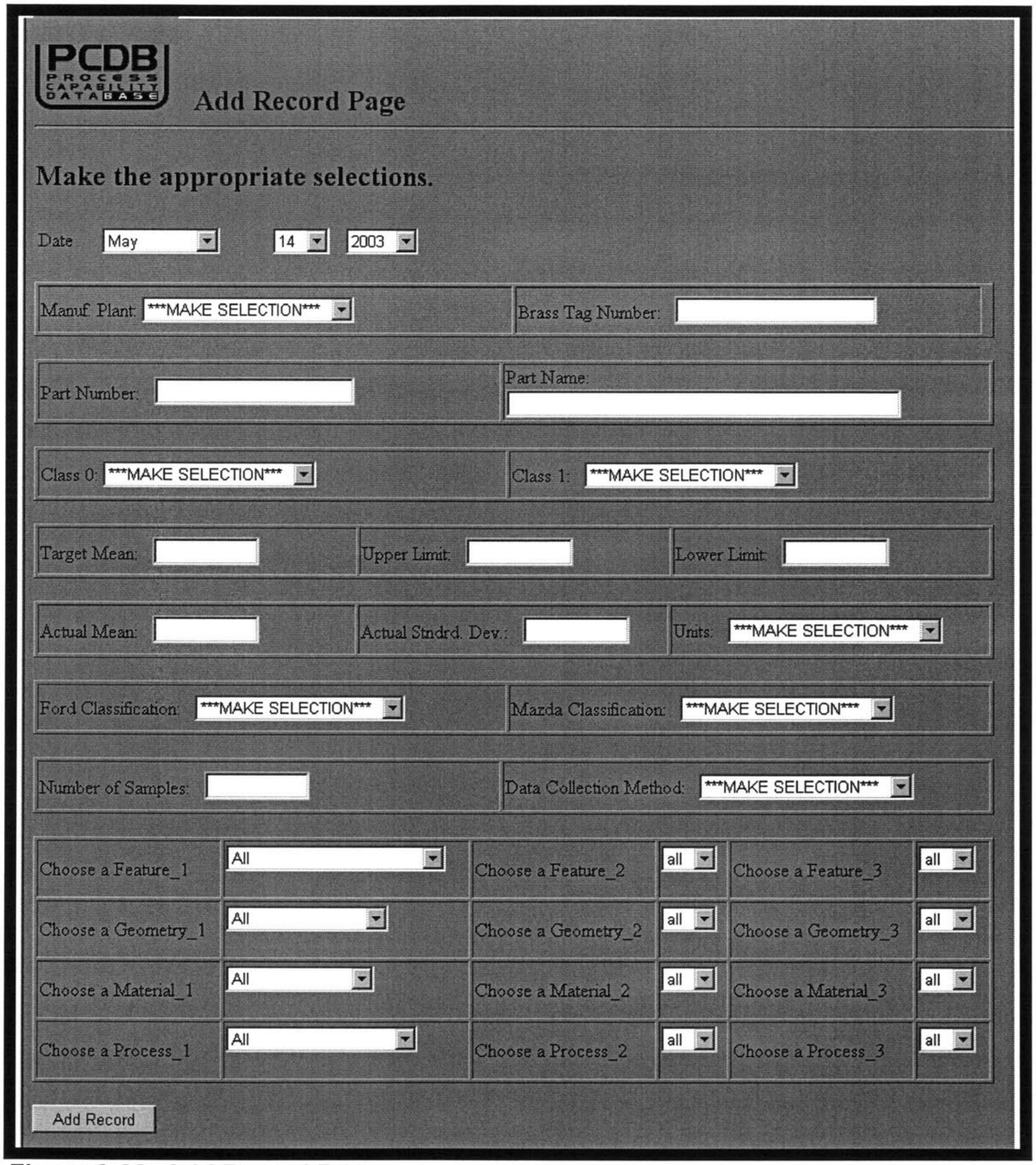

Figure 2-28: Add Record Page

The Add Record Page is organized slightly differently than the query page. Information concerning how the part is manufactured is input in the first section. The next section is where the measurement data are input. Finally, the last section is for inputting the indexing scheme to identify the characteristic. The process capability database also contains a section for database administration. This portion of the database is discussed in Appendix F.

2.4 Record Set

The Process Capability Database contains records of data stored in tables. The record set is the group of fields or pieces of information that each record contains. Of course, the record set contains fields for process capability data. Each record set also contains fields for descriptive data. These data include information such as part name, plant location, production line, collection date, and other pertinent data. The descriptive data reveals the origin of the specific set of data. This descriptive data helps an engineer determine if the query results are relevant. The remaining fields are for the indexing scheme and are used to relate the data to the product characteristics the data represent. The specific fields are listed in groups in Table 2-2, Table 2-3, and Table 2-4 below. The structure of the database is not explicitly discussed in this thesis. A good introduction and reference to general relational database design is *Database Design for Mere Mortals* (Hernandez 1997).

Table 2-2: Process capability data fields in record set

Record Set Field	Description of Field
Target Mean	Specified mean for the characteristic
Upper Limit	Specified upper limit for the characteristic
Lower Limit	Specified lower limit for the characteristic
Actual Output Mean	Mean calculated from measurements
Actual Output Standard Deviation	Standard deviation calculated from measurements
Input Mean*	Mean of incoming distribution
Input Standard Deviation*	Standard deviation of incoming distribution
Measurement Units	The measurement units used for capability data

* Required for the variation modeling technique presented in this thesis

Notice the inclusion of the mean and standard deviation of the incoming distribution to the particular process in the record set. These data are required to construct the equations that are used to compute the effects a process has on the variation of an incoming distribution. The technique that uses this type of data to construct these equations is presented in Chapter 3.

Traditionally process capability data only include data related to parts as they exit a process of interest as illustrated by Figure 2-29. Including data about the input distribution to the process requires a shift in the perception of process capability data and its uses. In the case of this thesis, the incoming and outgoing distribution parameters are needed to successfully build the models discussed in Chapter 3. Sometimes including data describing the incoming distribution requires more effort in the form of measurement or inspection of the incoming parts. In other instances, the data is known because the components were measured upon exiting the previous manufacturing operation or a supplier provides the data with the parts. Capturing and storing this information in the database record set is beneficial. In the case when the data is not "free," the costs of acquisition has to be weighed against the value of the data. In addition to including data about the incoming and outgoing distributions, process capability data can also include other pieces of data. These data could include process parameters and setting used to produce a population of parts, known disturbances to the process, physical properties of the workpieces, and a wealth of other information. The information included is highly dependent on the manufacturing firm's needs and expected uses of the data. An illustration of the proposed mental model of process capability data is shown below in Figure 2-30, which shows the inclusion of information about parts that enter a process, information about the process parameters, and information about parts exiting the process.

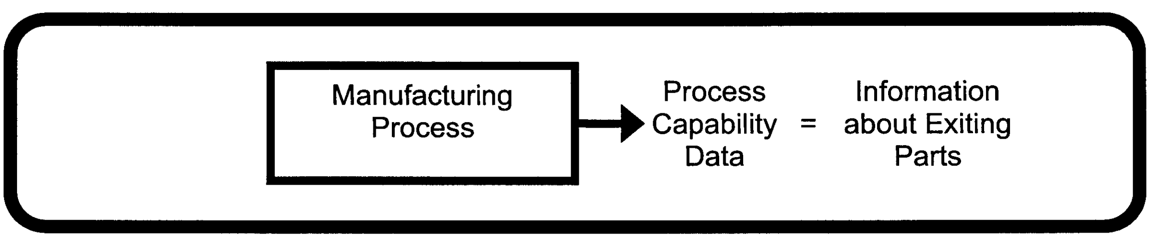

Figure 2-29: Traditional view of process capability data

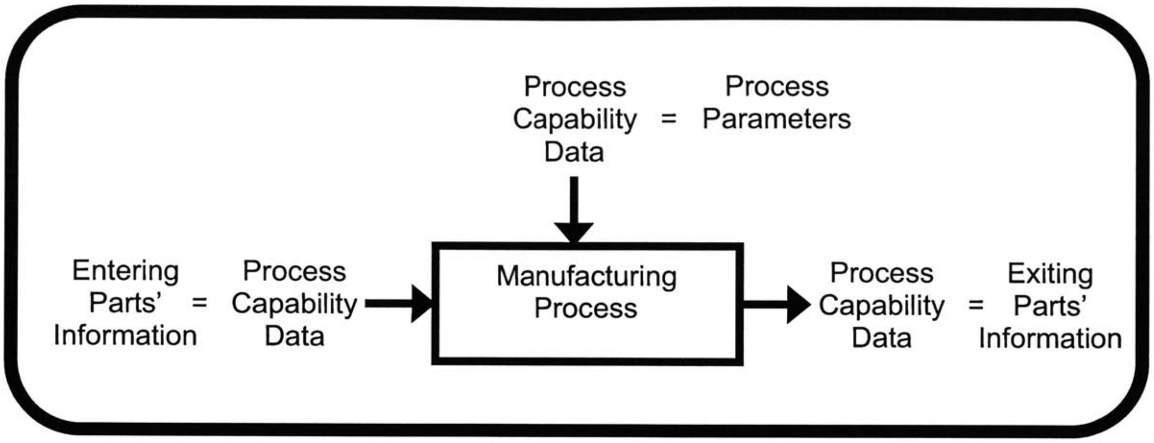

Figure 2-30: Proposed mental model of process capability data

Table 2-3: Descriptive data fields in record set

Record Set Field	Description of Field
Date	Date the samples were manufactured
Brass Tag Number	Unique number for each manufacturing machine used.
Plant Location	Where the part was manufactured
Part Name	Descriptive name of part
Part Number	Unique identifying number
Class0	Engine, Automatic Transmission, Manual Transmission
Class1	Type of engine, AT, or Manual Transmission
Data Collection Method	Automatic collection vs. manual collection vs. etc.
Last Time Record was Updated	When was "Meta Database" updated?
Company Significance Level	The significance of the characteristic.
Number of Samples	number of measurements used to compute mean, std dev.
Process Approvement Level	Red = Risky, Yellow = Caution, Green = OK
Process Production Volume	Low, Medium, or High

Table 2-4: Indexing scheme fields in record set

Record Set Field	Description of Field
Feature 0	Feature index
Feature 1	Feature index, Level 1
Feature 2	Feature index, Level 2
Feature 3	Feature index, Level 3
Geometry 0	Geometry index
Geometry 1	Geometry index, Level 1
Geometry 2	Geometry index, Level 2
Geometry 3	Geometry index, Level 3
Material 0	Material index
Material 1	Material index, Level 1
Material 2	Material index, Level 2
Material 3	Material index, Level 3
Process 0	Process index
Process 1	Process index, Level 1
Process 2	Process index, Level 2
Process 3	Process index, Level 3
Feature Notes	Specific Notes about the Feature
Geometry Notes	Specific Notes about the Geometry
Process Notes	Specific Notes about the Process
Material Notes	Specific Notes about the Material

The fields listed in the tables above were determined based on the needs of a specific manufacturing firm to ensure the data would have meaning to the engineers within the firm. These fields were decided upon through discussions and surveys. All fields deemed important were included. This is not a comprehensive, universal list. The exact fields a firm includes in its database depend on the firm and the needs of its engineers.

2.5 Chapter Summary

The focus of this chapter was addressing the barriers that prevent process capability databases from being used effectively. The challenges associated with accomplishing this, along with possible solutions, are summarized in this section. The solutions that are original ideas and concepts introduced for the first time in this thesis are identified by the ❋**NEW!**❋ symbol. The thesis section where each solution is discussed is listed to link to more detail. Each solution has advantages and disadvantages listed to provide a quick comparison.

Challenge: **Representing manufacturing data in a database so that it is easy to retrieve**

Solution	Section	Advantages (+) & Disadvantages (-)
Use part numbers to identify data	2.1.2	+ Easy to use because part numbers are well defined – Information is lost when part number changes – Engineers unfamiliar with product line cannot find data
Identify data with Feature, Process, and Material attributes	2.1.2	+ Engineers can find data easily + Familiar terminology can be used – Feature attribute not easily extensible – Some types of queries are cumbersome Reference: (Thornton and Tata 2000)
Use Group Technology (GT) to organize parts	2.2	+ Sorts parts into similar groups to reduce manufacturing costs + Familiar terminology can be used – Relies on a large number of parts to be indexed to be beneficial, i.e., to take advantage of similarities Reference: (Raffish 1985)
❋**NEW!**❋ Identify data with Feature, Geometry, Process, and Material attributes (Attribute Combination Descriptor)	2.2.1	+ Better defined than Group Technology, i.e., four attributes are identified and populated to fully characterize a physical aspect of a component + Engineers can find data easily + Familiar terminology can be used + Attributes are easily extensible + New types of queries are possible

Challenge: **Transforming part data into a form for easy database entry and storage**

Solution	Section	Advantages (+) & Disadvantages (-)
Ad hoc	n/a	– Difficult to ensure standard terminology is used – Results may not be correct
Computer Aided Design (CAD) feature recognition	2.1.2	+ Automated – Requires the existence of a 3-D CAD model – Resulting information insufficient for database entry, e.g., process information not available – Engineers interested in indexing data may not have access to CAD models Reference: (Han, Pratt et al. 2000)
❋NEW!❋ ACD Matrix Set	2.2.3	+ Well structured + Mistakes are avoided + Serves as a knowledge management tool + Helps identify important part characteristics that may need to be tracked – Templates must be created and linked to data structure

Challenge: **Providing a user interface to a process capability database**

Solution	Section	Advantages (+) & Disadvantages (-)
Command prompt	n/a	+ Many types of queries possible – Not intuitive – Engineers reluctant to learn new "language"
❋NEW!❋ Graphical User Interface (GUI)	2.3	+ Intuitive fields guide user + Dynamic menus eliminate infeasible combinations + Accessible via a standard web browser

Challenge: **Knowing the data to include in the process capability database**

Solution	Section	Advantages (+) & Disadvantages (-)
Statistical Process Capability (SPC) Data	1.5	+ Allows for Statistical Process Control (SPC) – Does not contain enough information to be effective during the design of new products
❋NEW!❋ Proposed record set	2.4	+ Includes the data needed to compute process capability indexes, e.g., Cpk, Cp + Data fields to identify when and where the parts were produced + Contains data needed to construct math functions that are introduced in Chapter 0 – Additional inputs must be made when data is entered

A PCDB is a valuable tool in designing components that can be manufactured with a high level of quality. The challenge of creating a PCDB that is easy to query has plagued their usefulness. A new data indexing scheme, the Attribute Combination Descriptor, was introduced that overcomes this challenge by using information related to a characteristic's feature, geometry, material, and process attributes. The usefulness of this indexing scheme is enhanced through the use of the ACD Matrix Set, a tool to assist an engineer in the systematic indexing of a component's key characteristics. The PCDB user interface and record set were also discussed.

Design engineers primarily search or query the PCDB to determine a company's manufacturing process capability to gain insight into the amount of variation a dimension on a product will have when produced. Querying the PCDB is very helpful when making qualitative comparisons, e.g., comparing the C_{pk} values of a few processes to find the one that produces the highest quality parts. However, using the data to determine the effects a manufacturing process has on an incoming distribution of parts is still difficult. This difficulty is compounded when several processes in a series affect the same dimension. Understanding the sources of the final product's variation is almost impossible using only the PCDB. To overcome these limitations, a mathematical framework is developed in the next chapter.

3 Modeling of Manufacturing Processes

Chapter Highlights

- Introduce and define ***Desired Dimensional Change***
- Introduce and define ***Process Imparted Dimensional Change*** and ***DeltaP***
- Introduce and define ***Process Imparted Variation*** and ***SigmaP***
- Validate DeltaP & SigmaP using physics-based math models of three processes
- Compare theoretical results to actual measurements

3.1 Introduction

Manufacturing processes are the building blocks of a production line. Most products go through more than one process, for example turning, grinding, etc. Process capability analysis usually focuses on the impact of a single process; however, in reality, the combined impact of many processes ultimately determines the final quality of a product. This chapter presents the development of the mathematical equations used to forecast the output mean and variation of a distribution produced by a manufacturing process as a function of the incoming distribution's mean and standard deviation. In Chapter 5, the manufacturing process equations are combined with the math models for assembly operations created in Chapter 4 to allow for the mean and standard deviation of a product's dimension to be computed when it is affected by multiple processes and assembly operations.

3.1.1 Overview

The goals of this chapter are to create and validate a general method for constructing closed-form equations for the mean and standard deviation of a dimension exiting a manufacturing process, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{target})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{target})$.

Chapter Goals

- Produce a simple method to model and understand variation propagation through a manufacturing process, i.e., how does the level of variance of a dimension entering a process affect the variance of it exiting the process?
- Develop a general method for constructing closed-form equations for the mean and standard deviation of a dimension exiting a manufacturing process as a function of the input dimension's mean and standard deviation, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$
- Validate constructed equations

3.1.2 Background

The dimensions of produced parts are represented statistically by probability density functions. Each probability density function has a mean and a standard deviation. A manufacturing process receives an input distribution and transforms it into an output distribution. Figure 3-1 displays a graphical representation of this transformation. In this thesis, manufacturing process variation is represented by the standard deviation of the output distribution.

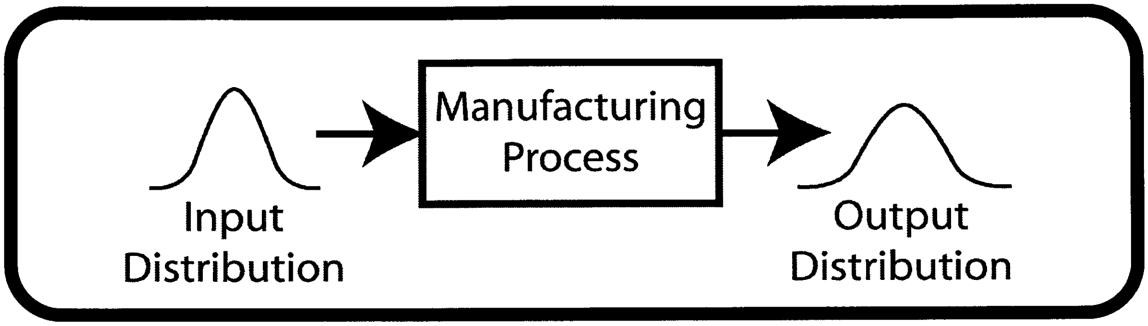

Figure 3-1: Process transforms a distribution

Engineers are very interested in understanding the relationship between the input and output distributions. Equations to compute the output mean and standard deviation as a function of the input mean and standard deviation allows a design engineer to specify tolerances that can be produced with an acceptable level of quality. The same equations assist a manufacturing engineer in selecting a process that ensures components are manufactured within specification. A project engineer also uses the equations to select suppliers by using their process capability data as input to the equations and calculating the resulting quality of the output. For example, a Design for Six Sigma (Antony and Coronado 2002) approach would greatly benefit from equations to calculate manufacturing output variation because the quality of a product could be forecast prior to its production, allowing the design to be created so that it meets the strict quality guidelines set forth by the Design for Six Sigma methodology.

There has been a wide variety of research in the area of product and manufacturing variation. Managing the effect variation has on the final cost and quality of a product while it is being designed is explored by Thornton (Thornton 1999). She developed a tool set to use during the design process to identify where variation will have significant impact. Suri and Otto (Suri and Otto 1999) use variational modeling to select manufacturing process input parameters that render the process outputs insensitive to variation. Their technique uses a designed experiment (Phadke 1989) to generate sensitivity functions for a manufacturing process so that its robustness can be improved. These techniques are good for identifying the product characteristics and process parameters that affect the quality of a manufactured product; however, they do not provide a method to forecast a product's quality while it is being designed.

A physics-based approach to modeling variation was introduced by Suri (Suri 1999). The approach creates an Integrated System Model that predicts the nominal values and variation of each output quality characteristic in a manufacturing system. Frey (Frey 1997) introduced the concept of a process capability matrix and bias vector to model a manufacturing system. The model developed can be used to compute the yield from the manufacturing process. A model of how variation is transmitted through a manufacturing process is presented by Lawless, Mackay, and Robinson (Lawless, MacKay et al. 1999). Their approach is to measure parts at each process in the manufacturing system and determine the amount of variation added by each process and carried through the next process. These approaches are only useful after the manufacturing system is in operation and do not assist in designing quality into a product.

The above approaches are useful when studying different aspects of product and manufacturing variation; however, none of them provide a simple method to compute a manufacturing process' output mean and standard deviation as a function of the mean and standard deviation of the distribution entering the process, i.e., none of the approaches produce $\mu_{out} = f(\mu_{in},\sigma_{in})$ and $\sigma_{out} = f(\mu_{in},\sigma_{in})$. One approach used to generate equations for the output mean and standard deviation of a distribution exiting a manufacturing process as a function of the incoming distribution's mean and standard deviation is simply to fit curves to measured data. The major downside to this approach is the inability to know what terms to include, i.e., combinations of the input mean and standard deviation and higher order terms, since the true curve is unknown. Also, the math functions have no real meaning or significance to an engineer, thus using them can be misleading and testing their validity is difficult.

Developing a method to create equations of the form $\mu_{out} = f(\mu_{in},\sigma_{in})$ and $\sigma_{out} = f(\mu_{in},\sigma_{in})$ for a manufacturing process is very challenging. Knowing how the process affects the variation is often times a mystery because there are an infinite number of combinations of the input mean and standard deviation. Running a design of experiment varying each parameter independently is very time consuming and costly since each experiment would require the processing of an entire distribution with a fixed mean and standard deviation. Adding to the challenge, this experiment would have to be run for every process, i.e., the results from one process could not be generalized and used to create equations for other processes.

In this chapter, a new method to create equations that model the transformation a distribution goes through when acted upon by a manufacturing process is presented. This new method is simple and quick to use. The method results in closed-form equations for the output mean and output standard deviation as a function of input mean and standard deviation. The method also produces mathematical graphs for a process that can be used to quickly compare two processes qualitatively. The foundation is the concept of characterizing a manufacturing process with two math functions called *DeltaP* and *SigmaP*. DeltaP and SigmaP represent Process Imparted Dimensional Change and Process Imparted Variation, which conceptually represent the effects a manufacturing process has on a dimension. Using these functions, closed-form solutions for the mean and standard deviation of a distribution exiting a manufacturing process can be constructed. The background of the theory is presented as well as an overview of the derivation of the closed-form equations for the output mean and standard deviation for a generic manufacturing process.

The end result of the above derivation is a set of equations for a given process that accept the mean and standard deviation of an incoming distribution as inputs to calculate the mean and standard deviation of the output distribution. This is shown conceptually in Figure 3-2.

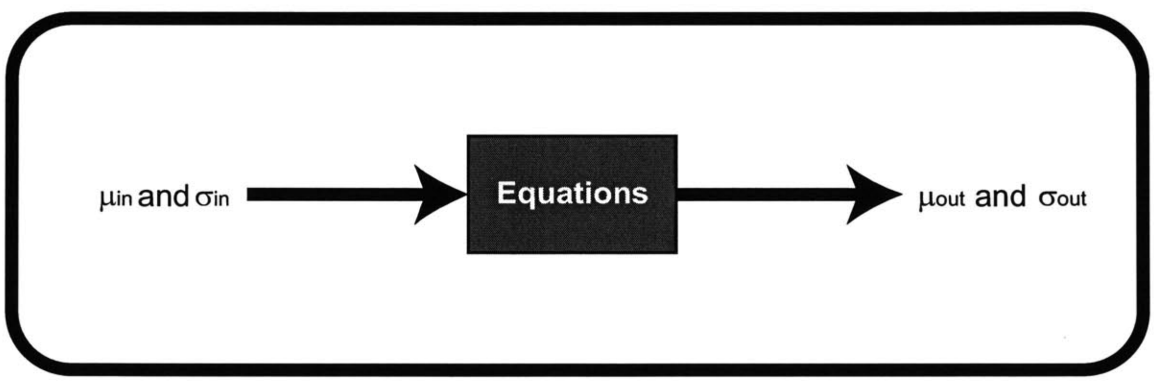

Figure 3-2: End result

3.2 Theoretical Framework

The derivation of the closed-form equations, $\mu_{out} = f(\mu_{in},\sigma_{in})$ and $\sigma_{out} = f(\mu_{in},\sigma_{in})$, is founded on the theory that two functions, DeltaP and SigmaP, can be used to represent the manufacturing process of interest. This section introduces, describes, and defines DeltaP and SigmaP.

3.2.1 Theory for Computing Output Mean

Manufacturing processes are the building blocks of a production line. A manufacturing process transforms the dimensions of a component. The process acts on an incoming dimension and attempts to transform it to a desired target value.

Desired Dimensional Change, M, is the change a process must make to a part's dimension to result in the dimension being equal to the desired (target) value.

The *desired dimensional change* will be referred to as M. An equation for M can be written as

$$M = \mu_{target} - x \qquad \text{3-1}$$

where μ_{target} is the desired target value for the output dimension and x is a component's incoming dimension.

A process is setup to produce a distribution with a desired target mean, μ_{target}; however, the output distribution does not always have a mean equal to the target mean. The difference is often referred to as the mean shift or the bias, b. The equation for bias is seen below in Equation 3-2.

$$b = \mu_{output} - \mu_{target} \qquad \text{3-2}$$

A process transforms an input dimension into an output dimension, e.g., a lathe reduces an incoming diameter to a smaller diameter. *Process Imparted Dimensional Change* is the actual change a dimension undergoes when operated on by a process when there is no process variation present. For example, an incoming distribution of rod diameters has a mean of 25mm. The output target mean for a turning process is 20mm. A particular rod with a diameter of 26mm is operated on and results in a rod with a diameter of 20.1mm. In this case, the Process Imparted Dimensional Change is -5.9mm for a desired dimensional change, M, of -6.0mm.

Process Imparted Dimensional Change is the actual change a dimension undergoes when operated on by a process when there is no process variation present.

Process Imparted Dimensional Change is different than bias. The bias is the actual difference between the target mean and the mean of the output distribution. Process Imparted Dimensional Change is the change a single dimension undergoes when a process operates on the part in the absence of variation.

Process Imparted Dimensional Change is a function of process parameters. Typical parameters to consider are the desired dimensional change, work piece material, cutter speed, feed, type of machine, and type of tool, etc. Most other parameters that affect Process Imparted Dimensional Change are specific to the process being modeled.

A special case of Process Imparted Dimensional Change is when all process parameters except M, desired dimensional change, are fixed. This special case is named DeltaP. A graphical definition is shown below in Figure 3-3 where the distribution on the left has zero variance (no spread) and is transformed by the process to produce a distribution that has no variance but is shifted by an amount DeltaP.

DeltaP is a mathematical representation of the actual change a dimension undergoes when operated on by a process with a given set of parameters when no process variation is present. It is a function of the desired dimensional change, M.

$$DeltaP = f(M) \qquad 3\text{-}3$$

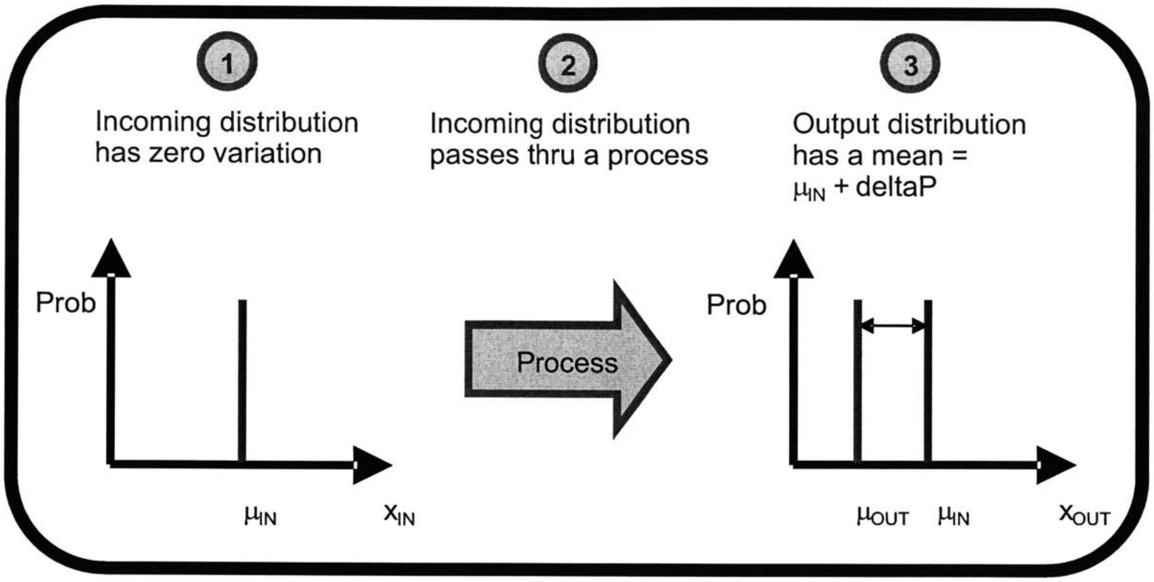

Figure 3-3: Graphical definition of DeltaP

A method of generating the function DeltaP is discussed later. A sample graph of DeltaP versus desired dimensional change is shown in Figure 3-4. In the sample shown, the DeltaP function represents a material removal process since it is valid over negative values of M. The graph shows that the process will produce oversized parts since it does not remove as much material as it should, i.e., the process removes less material than the ideal since DeltaP lies above the ideal line.

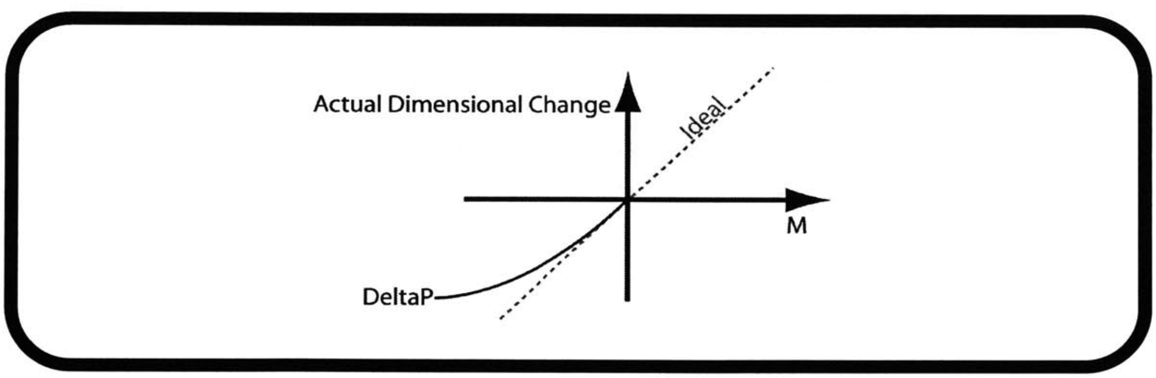

Figure 3-4: Graph of sample DeltaP

As seen in Figure 3-4, the ideal manufacturing process would have an associated Delta P = M. In this case, a manufacturing process produces all dimensions exactly equal to the output target dimension by making the actual dimensional change to a part equal to the desired dimensional change.

A DeltaP exists for a manufacturing process with a fixed set of operating parameters. Therefore, a particular process can have an infinite number of DeltaPs, one for each set of operating parameters. For example, a lathe has one DeltaP when it is operating on aluminum at 1500 RPM, and the same lathe operation has another DeltaP when operating at 2000 RPM.

3.2.2 Theory for Computing Output Variation

Process Imparted Variation is the variation an under control process inherently imparts on an incoming distribution that has zero variance. For example, a lathe is set to machine rods to a desired output diameter. A hundred rods with identical diameters are machined by this process. The result is a distribution of rod diameters with a mean and a standard deviation. This standard deviation is the Process Imparted Variation.

Process Imparted Variation is the standard deviation an under control process inherently imparts on an incoming distribution that has zero variance.

Similar to Process Imparted Dimensional Change, Process Imparted Variation is a function of process parameters. Typical parameters to consider are amount of material removed, work piece material, cutter speed, feed, type of machine, and type of tool. Most other parameters that affect Process Induced Variation are specific to the process.

A special case of Process Imparted Variation is when all process parameters except M, desired dimensional change, are fixed. This special case is named SigmaP. A graphical definition is given below in Figure 3-5 where a distribution with zero variance (no spread) is transformed by a process into a distribution with a standard deviation equal to SigmaP.

SigmaP is a mathematical representation of the standard deviation an under control process with a set of given parameters inherently imparts on an incoming distribution that has zero variance. It is a function of the desired dimensional change, M.

$$SigmaP = f(M) \qquad 3\text{-}4$$

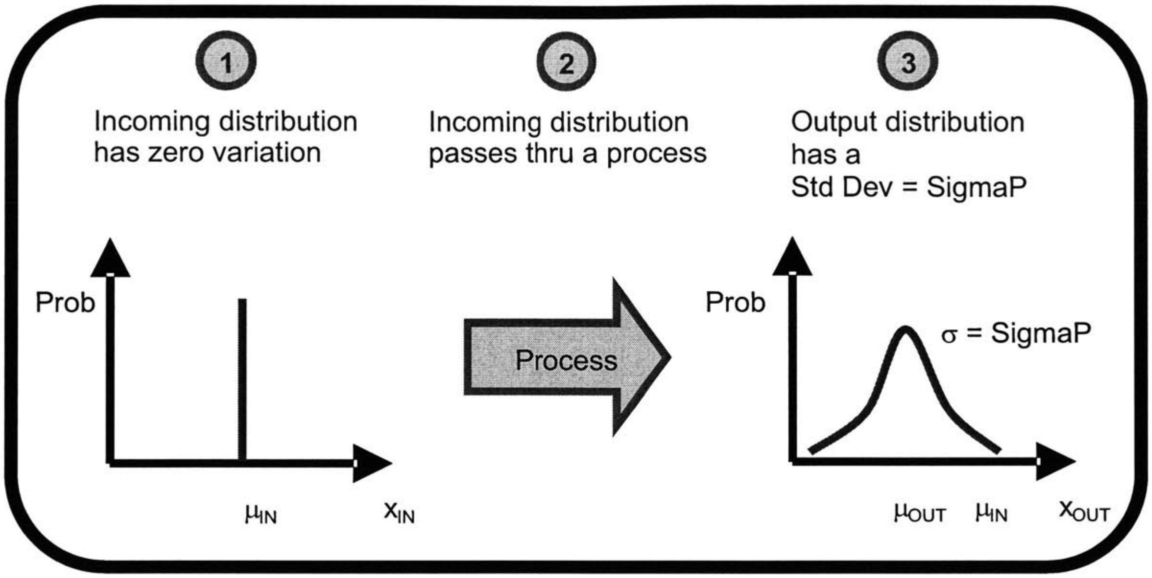

Figure 3-5: Graphical definition of SigmaP

The exact method of generating the function SigmaP is discussed in a later section. A sample graph of SigmaP versus Desired Dimensional Change is shown in Figure 3-6. In the sample shown, the SigmaP function represents a material removal process since it is valid over negative values of M. The graph shows that the process will produce parts with increasing variance as M increases in magnitude. SigmaP

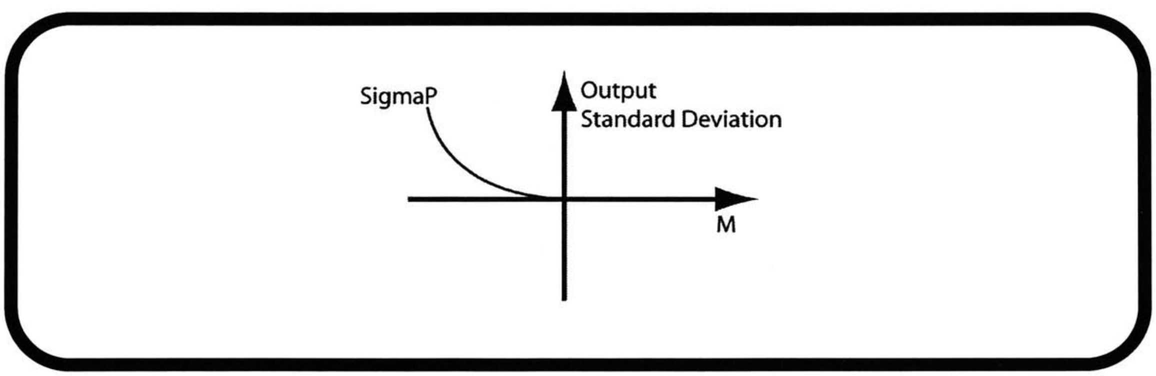

Figure 3-6: Graph of sample SigmaP

The theory of a manufacturing process being represented by the two mathematical functions of DeltaP and SigmaP can be used as a powerful qualitative tool. The DeltaPs for a process at different operating parameters or for different processes can be reviewed to select a process that more accurately produces dimensions at the desired value. For example, a process with a DeltaP that is almost a 45° line would result in dimensions closer to the desired value than a process with a DeltaP having a smaller slope. DeltaP can also be used to evaluate how sensitive a process is to the incoming dimension. For example, a highly non-linear DeltaP may reveal the need to set the process up to remove less material each pass to ensure output dimensions close to the target value.

The SigmaPs can also be used to compare how much variation different processes impart to a distribution. For example, a process with a low, flat SigmaP would be preferred to one with a steep slope. A SigmaP curve can also reveal a process' "sweet spot" where the variation imparted could be minimized.

3.3 Method to Compute μ_{output} and σ_{output}

The qualitative uses of Delta P and SigmaP can be beneficial to an engineer; however, the main goal is to develop equations to calculate the mean and variance of a process' output distribution given the input mean and variance. This goal requires the following derivation.

3.3.1 Distribution Binning and Processing

The first step in deriving the equations for the output mean and output variance is to bin the incoming distribution. Binning is the procedure of dividing a distribution into discrete divisions, or bins. Once the width of a bin is fixed, the height is set so that the area of the bin is equal to that under the probability distribution function over that range. Therefore, the areas in all the bins add to unity. The area in a bin can be calculated regardless of the type of probability density function according to Equation 3-5. A picture of a sample binned input distribution is shown in Figure 3-7.

$$Area_i = \int_{x_{i-1}}^{x_i} \left(pdf_{input}\right) dx \qquad \text{3-5}$$

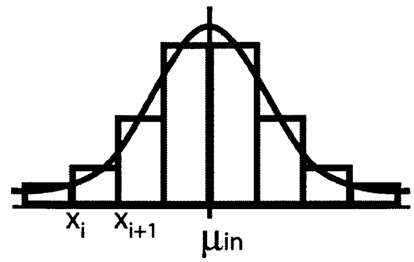

Figure 3-7: Sample of a binned input distribution

The bin width and the number of bins, N, are arbitrary at this point. The smaller the bin width, the greater number of bins required to traverse the original distribution. Also, as the bin width decreases, the discrete representation approaches that of the original distribution.

The next step is to treat each bin as an individual distribution. Each bin will be "processed" using the theory of DeltaP and SigmaP, resulting in an output sub-distribution. As an approximation, each bin is considered to contain all equal dimensions. For example, the i[th] bin from x_{i-1} to x_i contains components that all have the same dimension, equal to $(x_{i-1} + x_i)/2$. (This approximation becomes mute at a later point.) Using this approximation, the Desired Dimensional Change for the i[th] bin is calculated to be

$$M_i = \mu_{target} - \frac{(x_{i-1} + x_i)}{2} \qquad \text{3-6}$$

Each bin is processed to create an output sub-distribution as shown in Figure 3-8 which displays a bin entering a process and exiting as a probability density function. Using the definitions and equations for DeltaP and SigmaP, the mean, μ_i, and standard deviation, σ_i of the output sub-distribution can be approximated as

$$\mu_i \cong \frac{x_{i-i} + x_i}{2} + DeltaP(M_i) \qquad \text{3-7}$$

$$\sigma_i \cong SigmaP(M_i) \qquad \text{3-8}$$

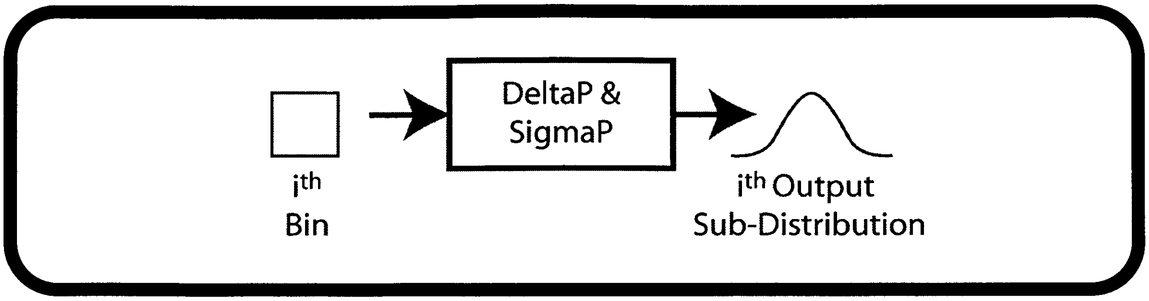

Figure 3-8: Bin transformed into output sub-distribution

After every bin has been "processed," the result is N sub-distributions. Each sub-distribution has a mean and standard deviation that has been calculated using Equations 3-7 and 3-8.

The main goal is still to derive equations to calculate the mean and standard deviation of a manufacturing process' output distribution. At this point, the means and standard deviations are known for the N distributions created by "processing" the N bins of the input distribution. The manufacturing process output distribution is simply the combination of these N sub-distributions. The distribution that results from mixing multiple distributions into one is often referred to as a mixture distribution. The equations for calculating the mean and variance of a mixture distribution are shown below.

$$\mu_{mixed} = p_1\mu_1 + p_2\mu_2 + \ldots + p_N\mu_N \tag{3-9}$$

$$\begin{aligned} \sigma^2_{mixed} &= \left(p_1\sigma_1^2 + p_2\sigma_2^2 + \ldots + p_N\sigma_N^2\right) \\ &+\left(p_1\mu_1^2 + p_2\mu_2^2 + \ldots + p_N\mu_N^2\right) \\ &-(p_1\mu_1 + p_2\mu_2 + \ldots + p_N\mu_N)^2 \end{aligned} \tag{3-10}$$

The μ_i's are the means of the distributions being mixed, and the σ_i's are the standard deviations of the distributions being mixed. The p_i term is equal to the number of measurements in the i^{th} distribution divided by the total number of measurements.

Equations 3-9 and 3-10 can be used to generate equations for the approximate values of an output distribution's mean and standard deviation. The p_i terms are replaced with Equation 3-5 and Equations 3-7 and 3-8 replace the means and standard deviations. These substitutions result in the following equations.

$$\mu_{output} \cong \sum_{i=1}^{N} \left\{ \int_{x_{i-1}}^{x_i} \left(pdf_{input}\right) dx \left[\frac{x_{i-1} + x_i}{2} + DeltaP\left(M_i\right) \right] \right\} \qquad \text{3-11}$$

$$\begin{aligned} \sigma^2_{output} \cong & \sum_{i=1}^{N} \left(\int_{x_{i-1}}^{x_i} \left(pdf_{input}\right) dx \left[SigmaP\left(M_i\right)\right]^2 \right) \\ & + \sum_{i=1}^{N} \left(\int_{x_{i-1}}^{x_i} \left(pdf_{input}\right) dx \left[\frac{x_{i-1} + x_i}{2} + DeltaP\left(M_i\right) \right]^2 \right) \\ & - (\mu_{output})^2 \end{aligned} \qquad \text{3-12}$$

Equations 3-11 and 3-12 are approximations because they were derived from a finite number of bins. (Remember the bin approximation made earlier?)

Manufacturing processes often generate dimensions that are normally distributed (Kalpakjian and Schmid 2001). With this assumption, the pdf_{input} terms in Equations 3-11 and 3-12 can be replaced with the equation for a normal distribution. At this time, it is also advantageous to standardize the equations above to use the reduced variable z in place of x. This allows for the standard form of the normal distribution, mean of zero and one standard deviation is at z equal to 1, equation to be used. The main advantage to this substitution is the limits on the integrals can be generalized. The equation for a normal distribution is

$$pdf_{Normal} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\} \qquad \text{3-13}$$

The reduced variable z is related to x according to

$$z = \frac{(x-\mu)}{\sigma} \qquad \text{3-14}$$

The standard normal distribution is

$$pdf = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \quad \text{3-15}$$

Equations 3-6 and 3-14 are combined to result in the equation below for the Desired Dimensional Change as a function of the reduced variable z.

$$M_i = \mu_{target} - \left\{ \frac{\{z_{i-1} + z_i\}\sigma_{input}}{2} + \mu_{input} \right\} \quad \text{3-16}$$

Now the appropriate substitutions can be made into Equations 3-11 and 3-12. Equation 3-15 is substituted in for the pdf_{input} terms. Equation 3-16 replaces the M_i terms. Finally, the x's are translated into z's using Equation 3-14. The results of these changes are the following equations.

$$\mu_{output} \cong \sum_{i=1}^{N} \left\{ \begin{array}{l} \frac{1}{\sqrt{2\pi}} \int_{z_{i-1}}^{z_i} \left\{ \exp\left(-\frac{z^2}{2}\right) \right\} dz \\ \left[\left[\frac{\{z_i + z_{i+1}\}\sigma_{input}}{2} + \mu_{input} + DeltaP\left(\mu_{target} - \left\{ \frac{\{z_{i-1} + z_i\}\sigma_{input}}{2} + \mu_{input} \right\} \right) \right] \right] \end{array} \right\} \quad \text{3-17}$$

$$\begin{aligned} \sigma^2_{output} \cong & \sum_{i=1}^{N} \left(\frac{1}{\sqrt{2\pi}} \int_{z_{i-1}}^{z_i} \left(\exp\left\{-\frac{z^2}{2}\right\} \right) dz \left[SigmaP\left(\mu_{target} - \left\{ \frac{\{z_{i-1} + z_i\}\sigma_{input}}{2} + \mu_{input} \right\} \right) \right]^2 \right) \\ & + \sum_{i=1}^{N} \left(\frac{1}{\sqrt{2\pi}} \int_{z_{i-1}}^{z_i} \left(\exp\left\{-\frac{z^2}{2}\right\} \right) dz \left[\frac{(z_{i-1} + z_i)\sigma_{input}}{2} + \mu_{input} + DeltaP\left(\mu_{target} - \left\{ \frac{\{z_{i-1} + z_i\}\sigma_{input}}{2} + \mu_{input} \right\} \right) \right]^2 \right) \\ & -(\mu_{output})^2 \end{aligned} \quad \text{3-18}$$

3.3.2 Approximate Solution Transformation

Equations 3-17 and 3-18 are one step closer to the goal; however, they are still only approximate solutions because of the original assumption of a finite number of bins. The solutions are made exact by taking the limit as the bin width goes to zero.

Several changes occur when the bin width goes to zero. This limit converts a discrete function into a continuous one. The discrete variable z_i becomes the continuous variable z. The integral over the normal distribution becomes the value of the integrand at that value of z. The summations become integrals with limits of negative infinity to positive infinity so that it spans the entire distribution. Making these modifications and simplifying results in the following equations.

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[DeltaP\left(\mu_{target} - \{ z\sigma_{input} + \mu_{input} \} \right) \right] \right\} dz \quad \text{3-19}$$

$$\begin{aligned}
\sigma^2_{output} = & \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[SigmaP\left(\mu_{target} - \{ z\sigma_{input} + \mu_{input} \} \right) \right]^2 \right) dz \\
& + \sigma^2_{input} + \mu^2_{input} \\
& + \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[2z\sigma_{input} DeltaP\left(\mu_{target} - \{ z\sigma_{input} + \mu_{input} \} \right) \right] \right) dz \\
& + 2\mu_{input} \{ \mu_{output} - \mu_{input} \} \\
& + \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[\left[DeltaP\left(\mu_{target} - \{ z\sigma_{input} + \mu_{input} \} \right) \right]^2 \right] \right) dz \\
& - (\mu_{output})^2
\end{aligned} \quad \text{3-20}$$

The above equations are exact solutions for the mean and variance of an output distribution. The equations require knowledge of DeltaP and SigmaP to be of practical use. Constructing DeltaP and SigmaP is the topic of the next section.

3.3.3 Forms of DeltaP and SigmaP

Until this point in the derivation, the functions DeltaP and SigmaP have remained in the equations for μ_{output} and σ_{output}. General forms of DeltaP and Sigma P are presented below in Equations 3-21 and 3-22.

$$DeltaP = a_0 + a_1 M + a_2 M^2 + \ldots + a_i M^i + \ldots + a_\alpha M^\alpha \quad \text{3-21}$$

$$SigmaP = b_0 + b_1 M + b_2 M^2 + \ldots + b_k M^k + \ldots + b_\beta M^\beta \quad \text{3-22}$$

The functions DeltaP and SigmaP can be approximated as quadratic equations. This approximation matches closely with simulation results. Using quadratic equations also reduces the chance of "over fitting" the curves, i.e., having the curve pass close to all the data points but not represent the trend of the data. Quadratic forms of DeltaP and SigmaP are used in Section 3.6 to validate the theory with very good results. The simplified forms of DeltaP and SigmaP are below.

$$DeltaP = a_0 + a_1 M + a_2 M^2 \qquad \text{3-23}$$

$$SigmaP = b_0 + b_1 M + b_2 M^2 \qquad \text{3-24}$$

3.3.4 Closed-form Solution for μ_{output} and σ_{output}

Finally, closed form equations for μ_{output} and σ_{output} can be generated. Equations 3-23 and 3-24 are substituted into Equations 3-19 and 3-20, expanded, and then simplified to produce the following.

$$\mu_{output} = \mu_{input} + a_0 + a_1 \left\{ \mu_{target} - \mu_{input} \right\} + a_2 \left\{ \mu_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2 + \sigma_{input}^2 \right\} \qquad \text{3-25}$$

$$
\begin{aligned}
\sigma_{output}^2 &= b_0^2 + b_1^2\left\{\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right\} \\
&-2b_1 b_2\left[\mu_{input} - \mu_{target}\right]\left[\mu_{input}^2 + 3\sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right] \\
&+2b_0\left\{b_1\left[\mu_{target} - \mu_{input}\right] + b_2\left[\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right]\right\} \\
&+b_2^2\left[\mu_{input}^4 + 3\sigma_{input}^4 - 4\mu_{input}^3\mu_{target} + 6\sigma_{input}^2\mu_{target}^2 + \mu_{target}^4 + 6\mu_{input}^2\left\{\sigma_{input}^2 + \mu_{target}^2\right\}\right] \\
&+b_2^2\left[-4\mu_{input}\left\{3\sigma_{input}^2\mu_{target} + \mu_{target}^3\right\}\right] \\
&\sigma_{input}^2 + \mu_{input}^2 + 2\mu_{input}\left\{\mu_{output} - \mu_{input}\right\} \\
&-(\mu_{output})^2 + a_0^2 + a_1^2\left\{\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right\} \\
&+2a_0\left\{a_1\left[\mu_{target} - \mu_{input}\right] + a_2\left[\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right]\right\} \\
&-2a_1\left\{\sigma_{input}^2 + a_2\left[\mu_{input} - \mu_{target}\right]\left[\mu_{input}^2 + 3\sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right]\right\} \\
&+4a_2\sigma_{input}^2\left[\mu_{input} - \mu_{target}\right] \\
&+a_2^2\left[\mu_{input}^4 + 3\sigma_{input}^4 - 4\mu_{input}^3\mu_{target} + 6\sigma_{input}^2\mu_{target}^2 + \mu_{target}^4 + 6\mu_{input}^2\left\{\sigma_{input}^2 + \mu_{target}^2\right\}\right] \\
&a_2^2\left[-4\mu_{input}\left\{3\sigma_{input}^2\mu_{target} + \mu_{target}^3\right\}\right]
\end{aligned}
\qquad 3\text{-}26
$$

Although Equation 3-26 appears to be complicated, it is merely a function of μ_{input}, σ_{input}, μ_{target}, DeltaP coefficients, and SigmaP coefficients. The above equations can easily be input into a spreadsheet to compute the mean and standard deviation of a process' output distribution for different values of input distributions' means and standard deviations if DeltaP and SigmaP are known for the given process.

3.4 Coefficients of DeltaP and SigmaP

DeltaP and SigmaP are constructed using manufacturing data related to the process of interest. Using the second order polynomial forms of DeltaP and SigmaP allows Equations 3-25 and 3-26 to be used to find coefficients by fitting manufacturing data to them.

The first step is to locate values for μ_{input}, σ_{input}, μ_{target}, μ_{output}, and σ_{output} for the process of interest. This is a "set" of process capability data. At least three sets of data are required to perform a fit. As the number of sets increases, the reliability of the fit improves. The number of sets is represented by n with an individual set identified by the subscript j. Next, guesses are made for the coefficients. The guesses are used with the sample μ_{input}, σ_{input}, and μ_{target} to calculate values for μ_{output} and σ_{output}. Non-negative error terms are computed by subtracting these calculated values from the actual values of the sample data and squaring the result. The coefficients are then adjusted, and the squared error is recalculated. Repeating this until the sum of the squared errors is minimized identifies the best set of coefficients. These steps are easily implemented using the solver feature in popular spreadsheet applications, e.g., Microsoft® Excel. Graphical representations of the procedures for computing coefficients for DeltaP and SigmaP are shown in Figure 3-9 and Figure 3-10 respectively.

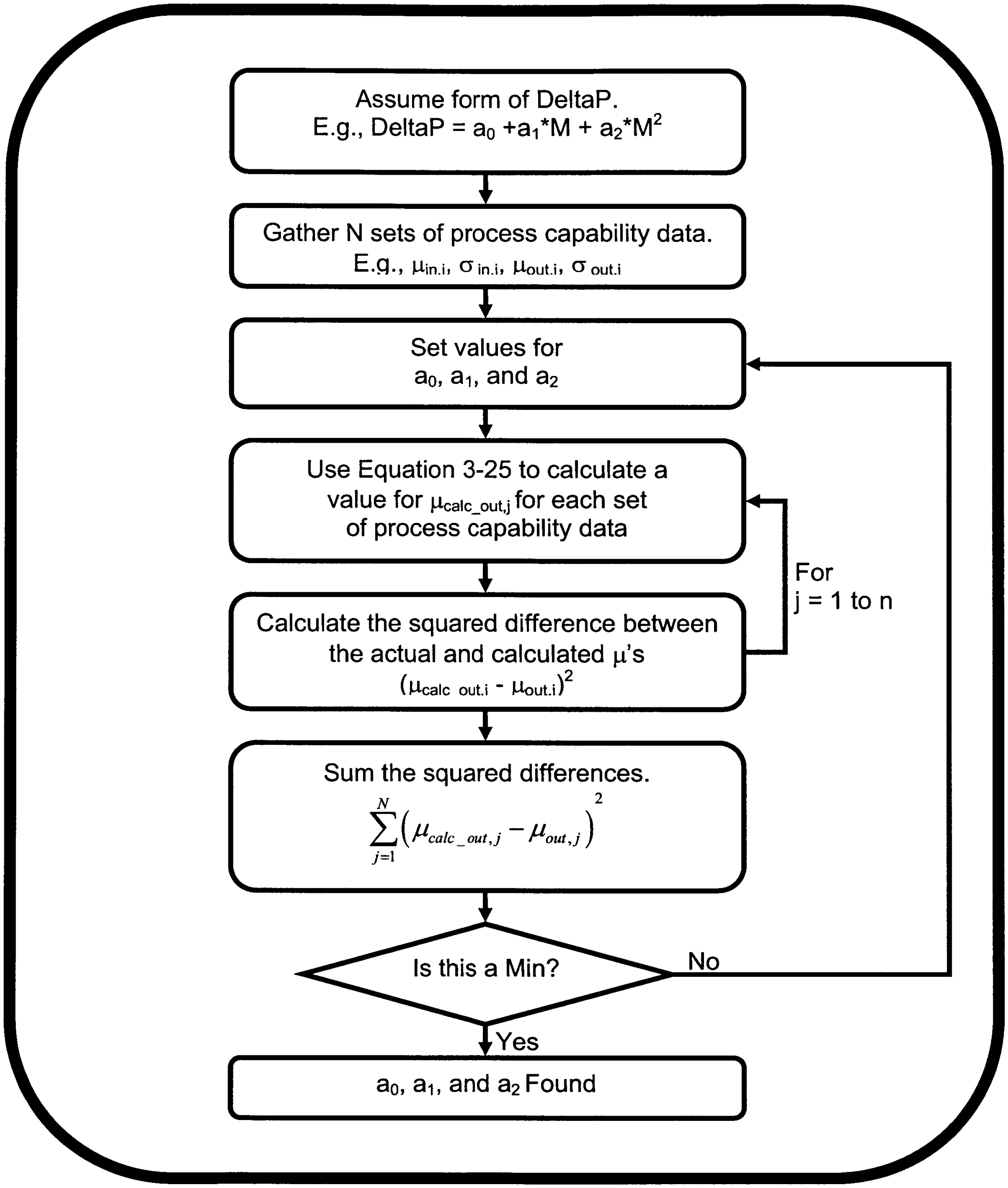

Figure 3-9: Method for finding coefficients of DeltaP

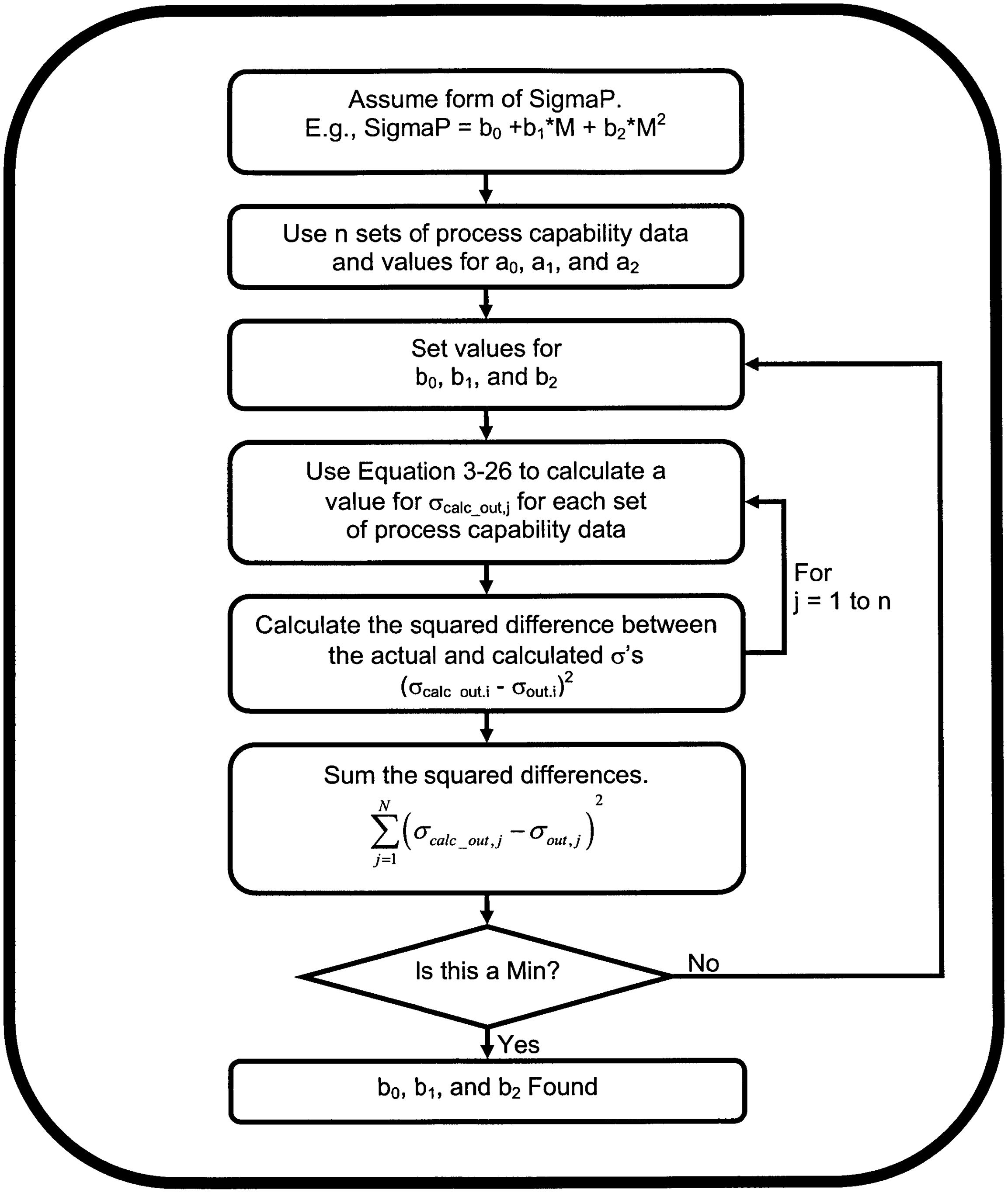

Figure 3-10: Method for finding coefficients of SigmaP

3.5 Applications of the Modeling Methodology

The methodology presented in this chapter is applicable to model many different manufacturing processes. The modeling technique produces closed-form equations that accept the mean and standard deviation of parts entering a process and the target value of the process output to calculate forecasted values for the mean and standard deviation of parts exiting the process. The basic requirement for applying this modeling technique is that the process must transform a dimension. This requirement exists because DeltaP and SigmaP are functions of the Desired Dimensional Change, M. The modeling methodology is demonstrated in the example presented in Section 5.4.

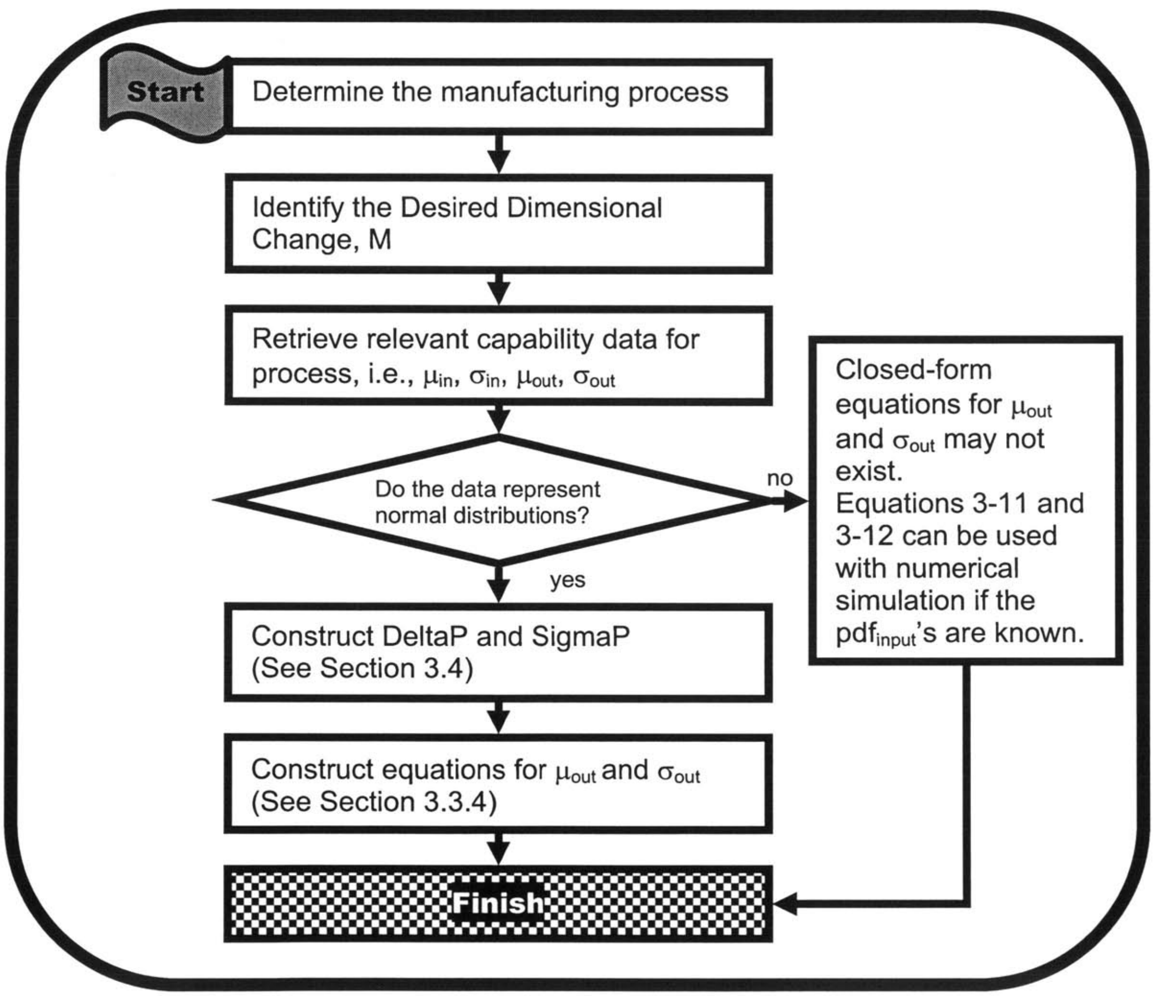

Figure 3-11: Illustration of modeling a manufacturing process

The graphical illustration in Figure 3-11 shows an overview of applying the technique to create closed-form equations to calculate forecasted values for the mean and standard deviation of parts upon exiting a manufacturing process. The starting point is determining the process of interest. The process is dependent on the type of part being produced. The process selection may depend on other non-functional considerations such as what processes are available, which processes are under utilized, etc.

Once the process has been determined, the Desired Dimensional Change must be identified. Identifying the Desired Dimensional Change for material removal processes is often very straightforward. For example, the depth of cut is the Desired Dimensional Change in a turning operation (see Section 3.6.1). Another example is end milling a block where the depth of cut is again the Desired Dimensional Change. Drilling is another material removal process that can be modeled using DeltaP and SigmaP. Modeling drilling is interesting because the incoming hole diameter distribution has a zero mean and zero variance, i.e., no hole exists. Using either the definitions of DeltaP and SigmaP or simplifying Equations 3-25 and 3-26 using $\mu_{input} = 0$ and $\sigma_{input} = 0$, it is shown that $\mu_{output} = \text{DeltaP(M)}$ and $\sigma_{output} = \text{SigmaP(M)}$ where M, the Desired Dimensional Change, is equal to the target mean of the drilled hole.

Forming operations can be modeled too. A rolling operation is modeled using the desired change in thickness as M (see Section 3.6.2). A forging operation could use the change in height of the workpiece as the Desired Dimensional Change. A forming operation that does not have an obvious Desired Dimensional Change is sheet bending. Choosing the bend angle as the Desired Dimensional Change is not appropriate because it cannot be used as the independent variable for DeltaP and SigmaP, i.e. it cannot be used as M. In this case, springback is considered, and the thickness of incoming sheets is factored into M along with a desired final dimension. Sheet bending is one of the operations used to validate the theory. To see specifics on how it is modeled, please refer to Section 3.6.3.

Net-shape manufacturing operations can also be modeled. Similar to drilling, the incoming mean and distribution of the dimension are both zero, i.e., non-existent. The output mean once again reduces to DeltaP(M), and the output standard deviation reduces to SigmaP(M) where M is the target dimension. A specific example is the injection molding of a wall where the thickness of the wall is the dimension of interest. The input mean is zero with zero variance since the wall does not initially exist.

After M has been identified, process capability data for the process and parameters of interest are retrieved. These data must be analyzed to determine whether they represent normal distributions or not. If they do not represent normal distributions, then closed-form solutions for μ_{out} and σ_{out} probably do not exist. Values for μ_{out} and σ_{out} can still be computed using Equations 3-11 and 3-12 with numerical simulation, if the probability density functions for the retrieved data are known. If the retrieved data do represent normal distributions (often a safe assumption, at least as a very good approximation), then DeltaP and SigmaP functions can be generated for the manufacturing process of interest per the steps in Section 3.4. Finally, the equations for μ_{out} and σ_{out} are constructed as described in Section 3.3.4. These equations are then used to compute the mean and standard deviation of dimensions leaving the process. The equations can also be linked with equations for other processes to compute the effects of serial processing (several sequential processes acting on the dimension) or linked with equations for assembly operations to compute the final quality of a dimension. These ideas are fully explored in Chapter 5.

The major limitation to applying this methodology is that it requires access to process capability data. The process capability data needed are values for the mean and standard deviation of parts entering and exiting the process being modeled. This is the reason the concepts related to the Process Capability Database in Chapter 2 are so important.

3.6 Validation of DeltaP and SigmaP Theory

Validating the theory of DeltaP and SigmaP using several manufacturing processes helps ensure it is general and broadly applicable. The first step toward validation is to create adequate physics-based math models of several processes. These math models are then used to simulate manufacturing process capability data through Monte Carlo simulations. Once the sample process capability data are generated, the math models representing the processes are "forgotten." The simulated process capability data from the models are used to compute DeltaP and SigmaP for each process. Once DeltaP and SigmaP are computed (see Section 3.4), the closed-form solutions for the output mean and output standard deviation can be constructed using Equations 3-25 and 3-26. These steps are outlined below in Figure 3-12.

Step 1: Create Math Models

- Formulate physics-based math model for a manufacturing process, e.g., turning, rolling, etc.
- Math model accepts input dimension and computes the output dimension

Step 2: Simulate Capability Data

1. Pick values for the mean and standard deviation of an incoming distribution
2. Use random number generator to create 1,000 normally distributed values that represent incoming dimensions
3. Input values to math models to calculate outgoing dimensions
4. Compute mean and standard deviation of outgoing dimensions, use this as process capability data

Step 3: Construct closed-form equations

1. Use simulated process capability data from Step 2 to compute DeltaP and SigmaP curves
2. Construct $\mu_{out} = f(\mu_{in}, \sigma_{in}, \text{DeltaP})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \text{DeltaP}, \text{SigmaP})$

Figure 3-12: Creation of closed-form equations for theory validation

Different values for an "incoming" mean and standard deviation can be input into the closed-form equations to compute output means and standard deviations. The same "incoming distribution" values can be used to generate random variables with probability density functions that have the appropriate mean and standard deviation. These random numbers are fed into the math models and a Monte Carlo simulation run, resulting in simulated values for the output mean and output standard deviation. These values are taken to be the "true" values the process would generate. The output values from the closed-form solution are then compared to the simulated "true" values to see how well they match. A close match indicates the closed-form solution created using the theory of DeltaP and SigmaP works well in forecasting the mean and standard deviation of a distribution as it exits a manufacturing process. These validation steps are outlined in Figure 3-13 below. The theory is validated using three different processes. The processes are turning, rolling, and sheet bending.

At first glance, the validation plan shown in Figure 3-12 and Figure 3-13 may seem to suffer from circular logic, i.e., isn't the same physics-based math model that created the DeltaP and SigmaP functions being used to generate the simulated data used for the final comparison? This, however, is not the case. The DeltaP and SigmaP functions are not created directly from knowledge of the physics that describe the manufacturing process. They are constructed from data that comes from Monte Carlo simulations that use the physics-based math models. This is an important distinction. The math model can be thought of as a "black box" with no specific information known about how it works. For example, a thousand numbers are put into the "black box" and a thousand numbers come out just like a thousand pieces go into a real manufacturing process and a thousand pieces come out. The mean and standard deviation are computed for the thousand numbers going in and the thousand numbers coming out and are treated as process capability data. This is the process capability data utilized in the construction of the DeltaP and SigmaP functions. As for the simulated "true" data used for comparison, the physics-based math model is used to "process" a set of random numbers that have a known mean and standard deviation, each random number represents one part. This is again analogous to a real manufacturing operation processing parts from a known distribution. The mean and standard deviation of the numbers "leaving" the math model represent the "true" values the real-world manufacturing process would generate. These are the values to which the closed-form calculated results are compared. Simply put, the physics-based math models are only used in Monte Carlo simulations to generate numbers whose mean and standard deviations are treated as process capability data. The models are never used to directly influence the construction or creation of DeltaP or SigmaP functions.

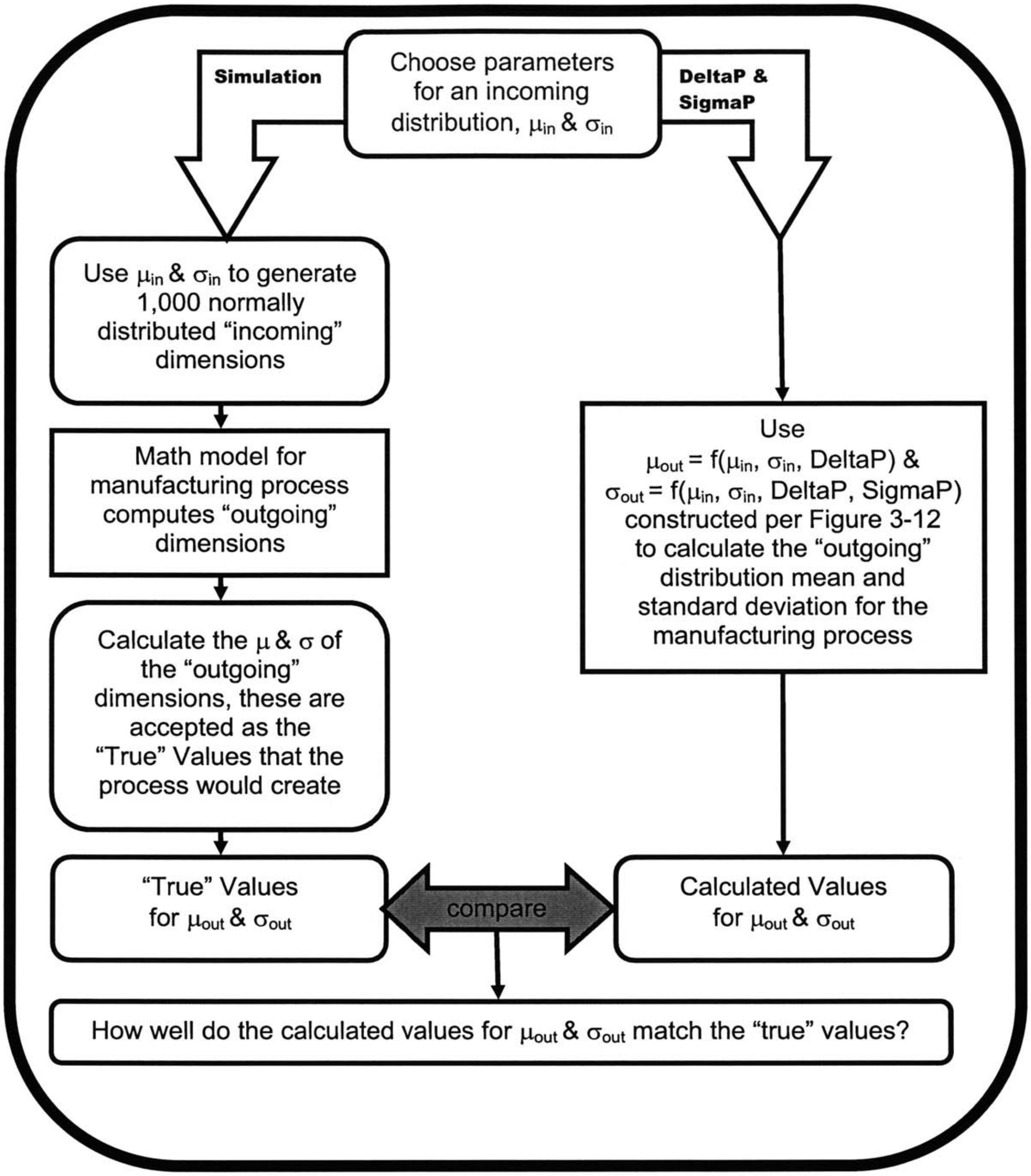

Figure 3-13: Validating calculated values vs. simulated values

3.6.1 Turning

Turning is a cutting operation that typically produces parts that are round in shape. In a turning operation, the workpiece is rotating while being machined. A cutting tool moves parallel to the rotating workpiece to remove a layer of material. The thickness of the layer of material being removed is known as the depth of cut.

3.6.1.1 Physics-based Math Model

This section develops a math model for a turning process to calculate the actual depth of cut as a function of turning parameters. There are three forces acting on the cutting tool and workpiece during turning. The actual depth of cut is partially determined by the deflections of the cutting tool and workpiece due to these forces. The math models presented in this thesis are based on those presented by Kalpakjian (Kalpakjian and Schmid 2001). Figure 3-14 shows a graphic of a turning operation with dimensions of interest labeled including depth of cut (d), initial workpiece diameter (D_0), final workpiece diameter (D_f), cutting tool feed (f), and the angular velocity of the workpiece (N).

The three forces involved in a turning operation are the thrust force, the radial force, and the cutting force. The thrust force, F_t, acts in the feed direction and resists the motion of the cutting tool. The radial force, F_r, acts in the radial direction and pushes the cutting tool from the workpiece. The cutting force, F_c, acts downward on the cutting tool and upward on the workpiece. These forces are represented graphically in Figure 3-15.

The math model for turning used in this thesis calculates the actual depth of cut at the tip of the workpiece for a turning operation with given parameters. The force that is most important to consider when calculating the actual depth of cut is the cutting force. The cutting force dominates the model because it causes the cutting tool to deflect away from the workpiece in a radial direction. The cutting force also deflects the workpiece in the radial direction. Both deflections influence the final diameter of the workpiece. Neglecting the thrust force is justified because it deflects the cutting tool in the workpiece's axial direction. The radial force can be neglected because it acts in the cutting tool's stiffest direction.

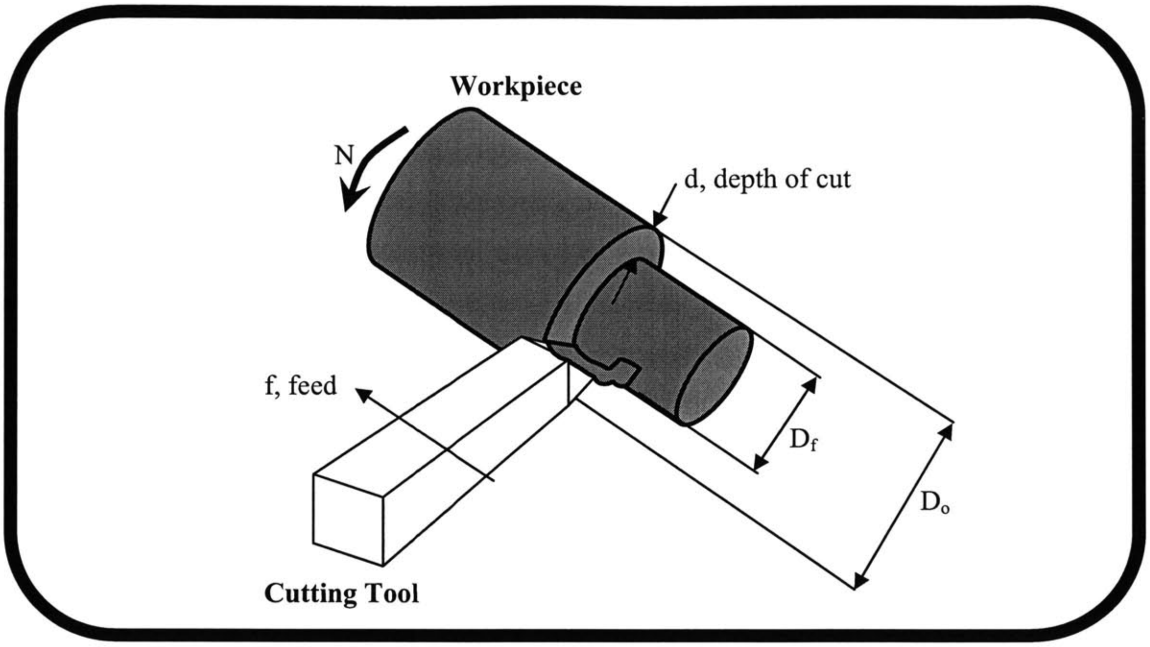

Figure 3-14: Dimensions of interest in turning

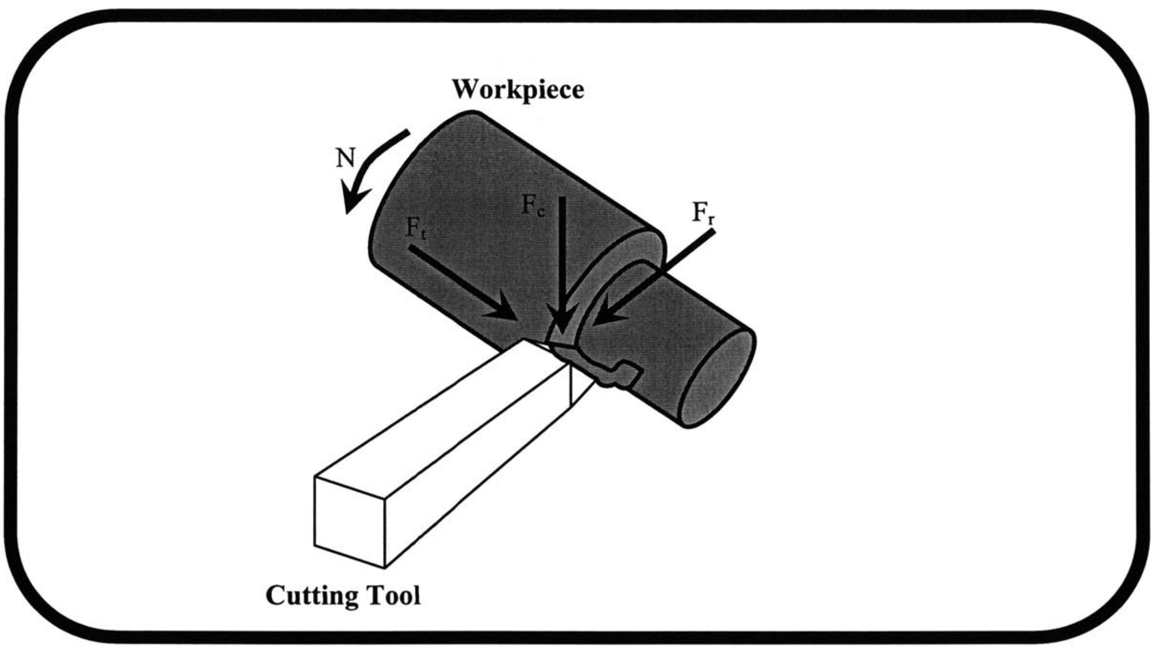

Figure 3-15: Cutting forces during turning

The equation for cutting force is derived from calculating the power required for the turning operation. The power required for a turning operation is related to the Material Removal Rate and the total specific energy of the workpiece material according to the following equation.

$$Power = MRR * u_t \tag{3-27}$$

The equation for Material Removal Rate is shown below.

$$MRR = (\pi)(D_{avg})(d_d)(f)(N) \tag{3-28}$$

Power is related to the cutting force according to the following equation.

$$Power = (Torque)(\omega) = \frac{1}{1000}(F_c)\left(\frac{D_{avg}}{2}\right)\omega \tag{3-29}$$

The factor of 1000 is required to maintain unit consistency. Now, Equations 3-27, 3-28, and 3-29 can be combined to get the following equation.

$$\frac{1}{1000}(F_c)\left(\frac{D_{avg}}{2}\right)\omega = (\pi)(D_{avg})(d_d)(f)(N)(u_t) \tag{3-30}$$

Rearranging the above equation and noting that $\omega = 2\pi N$ results in an equation for cutting force.

$$F_c = (1000)(d_d)(f)(u_t) \tag{3-31}$$

The workpiece and cutting tools can be modeled as beams. The deflection at the end of a beam can be calculated with the following equation.

$$\delta = \frac{F_c L^3}{3EI} \tag{3-32}$$

The workpiece has a circular cross section and the cutting tool has a rectangular cross section. Using the subscript w for the workpiece and the subscript t for the cutting tool, the deflections of each can be written as follows.

$$\delta_w = \frac{F_c L_w^3}{3E_w I_w} = \frac{F_c L_w^3}{3E_w\left(\frac{\pi D_o^4}{64}\right)} \qquad \text{3-33}$$

$$\delta_t = \frac{F_c L_t^3}{3E_t I_t} = \frac{F_c L_t^3}{3E_t\left(\frac{wh^3}{12}\right)} \qquad \text{3-34}$$

The total deflection is simply the sum of the workpiece deflection and the cutting tool deflection.

$$\delta_{total} = \delta_w + \delta_t \qquad \text{3-35}$$

The total deflection causes the actual final diameter of the workpiece to be larger than the desired final diameter. Now that the total deflection is known, the actual final diameter of the workpiece, D_a, at the tip can be computed. The equation for the actual final diameter is derived from geometrical relationships. End views of the cutting tool and workpiece for the case where both are perfectly rigid and for the case with the total deflection represented are shown below in Figure 3-16. Figure 3-17 shows the important aspects of the geometry used in deriving the equation for the actual final diameter of the workpiece.

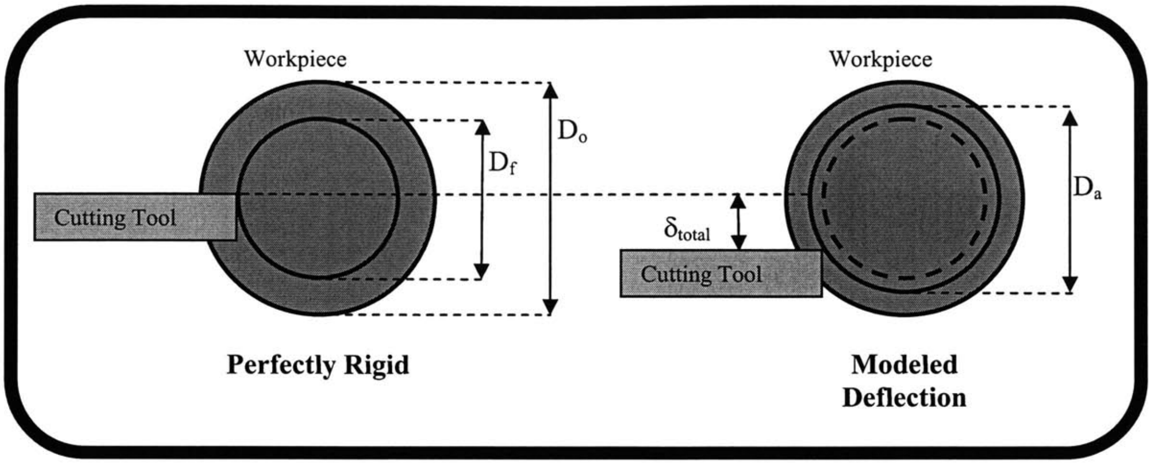

Figure 3-16: End views of cutting tool and workpiece for ideal and modeled cases

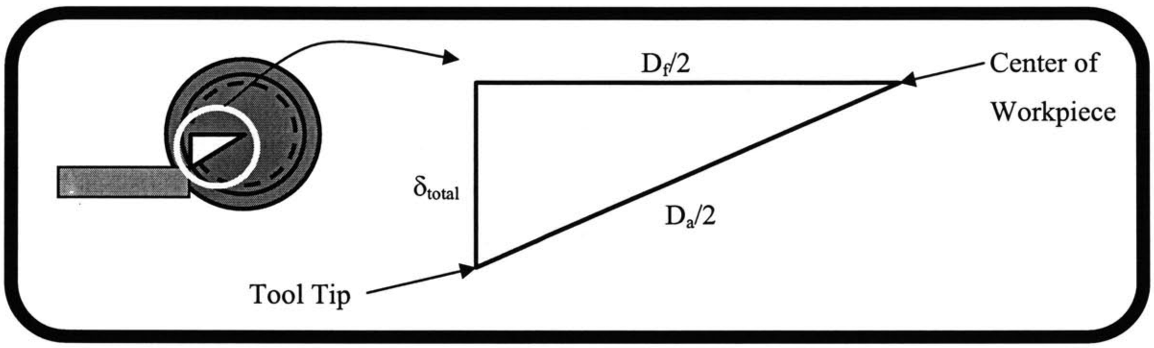

Figure 3-17: Geometry considerations for calculating actual final diameter

Using the geometry drawn in Figure 3-17, the following relation can be written.

$$\left(\frac{D_a}{2}\right)^2 = \left(\frac{D_f}{2}\right)^2 + \left(\delta_{total}\right)^2 \qquad \text{3-36}$$

This equation can be solved for D_a.

$$D_a = 2\sqrt{\left(\frac{D_f}{2}\right)^2 + \left(\delta_{total}\right)^2} \qquad \text{3-37}$$

Equations 3-31, 3-33, 3-34, and 3-37 are combined to result in the following equation for the actual final diameter.

$$D_a = 2\sqrt{\left(\frac{D_f}{2}\right)^2 + \left(\frac{1000(d_d)(f)(u_t)(L_w^3)}{3E_w\left(\frac{\pi D_o^4}{64}\right)} + \frac{1000(d_d)(f)(u_t)(L_t^3)}{3E_t\left(\frac{wh^3}{12}\right)}\right)^2} \qquad 3\text{-}38$$

3.6.1.2 Simulation of Process Capability Data

A Monte Carlo simulation is run to generate sample process capability data. Equation 3-38 is used to calculate the actual final diameter of the workpiece for the simulation. The total specific energy, u_t, is treated as a stochastic variable.

The model is used to generate ten sets of process capability data. Each set contains a value for an input mean, input standard deviation, output mean, and output standard deviation. These ten sets of data represent the type of information that would be retrieved from a Process Capability Database in a manufacturing firm and used to construct DeltaP and SigmaP functions for a particular process. To be completely unbiased, the ten values for the input means and the ten values of the standard deviations are randomly generated and paired. Each simulation run comprises one thousand calculations.

All parameters except for the total specific energy, u_t, and the original diameter of the workpiece, D_o are set. The average value and standard deviation for u_t are set. One thousand normally distributed random numbers are generated to represent values of u_t. Next, one thousand normally distributed numbers are generated that have a mean equal to the first value of input mean and a standard deviation equal to the value of the first input standard deviation. These values are generated with a random number generator in Microsoft® Excel. Using the generated numbers, the fixed parameter values, and Equation 3-38, one thousand D_o's are calculated. The mean and standard deviation are calculated for these one thousand values. The calculated mean is the output mean for the given input mean and standard deviation, and the calculated standard deviation is the output standard deviation for the given input mean and standard deviation. Repeat these steps ten times to generate the remaining simulated process capability data. This method is presented graphically in Figure 3-18.

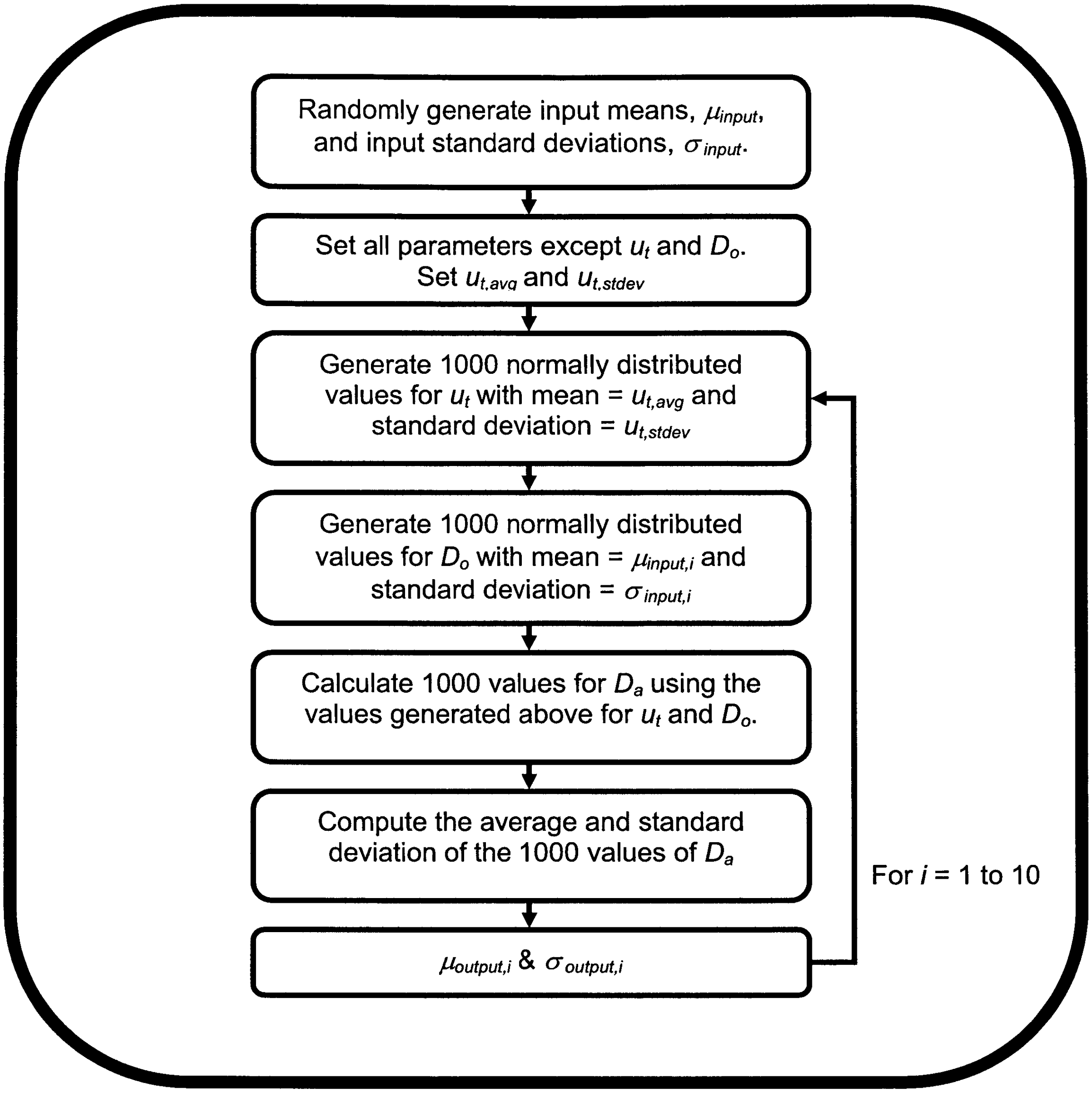

Figure 3-18: Algorithm for Monte Carlo simulation method

The workpiece material is Aluminum. The cutting tool is made of steel. The parameters of the workpiece and cutting tool are summarized in Table 3-1 and Table 3-2 respectively.

Table 3-1: Workpiece parameters for turning simulation

Length of workpiece, L_w	50.0 mm
Mean of Total Specific Energy, $u_{t,avg}$	1.0 W*s/mm^3
Standard Deviation of Total Specific Energy, $u_{t,stdev}$	0.05 W*s/mm^3
Modulus of Elasticity of Workpiece, E_w	70,000 MPa

Desired face diameter, D_f	30.0 mm

Table 3-2: Cutting tool parameters for turning simulation

Length of cutting tool, L_w	40.0 mm
Height of cutting tool, h	8.0 mm
Width of cutting tool, w	8.0 mm
Modulus of Elasticity of Cutting Tool, E_t	200,000 MPa
Desired face diameter, D_f	30.0 mm
Feed, f	0.35 mm/rev

Since the desired final diameter of the workpiece is 30.0 mm, the incoming mean should be around 35.0 mm. Values for input means will be randomly generated to fall close to 35.0 mm. Values for standard deviations will be randomly generated to fall close to 0.5 mm.

Table 3-3: Simulated process capability data for turning operation

Input Mean, μ_{in} [mm] Randomly Generated	Input Standard Deviation, σ_{in} [mm] Randomly Generated	Output Mean, μ_{out} [mm] Simulated	Output Standard Deviation, σ_{out} [mm] Simulated
39.860	0.4702	30.080	0.0106
36.490	0.5829	30.036	0.0072
34.719	0.3784	30.019	0.0036
37.606	0.2113	30.048	0.0054
37.265	0.4868	30.044	0.0074
36.408	0.6690	30.035	0.0077
37.264	0.1906	30.044	0.0049
37.593	0.6031	30.048	0.0087
34.625	0.4665	30.018	0.0040

3.6.1.3 Creation of DeltaP and SigmaP

The DeltaP and SigmaP functions can be created for the turning operation following the technique outlined in Section 3.4. Remember that DeltaP and SigmaP are functions of M, the desired dimensional change. In the case of turning, M is equal to the desired depth of cut, so

DeltaP = f(d_d) and SigmaP = f(d_d). The process capability data in Table 3-3 are used with the Microsoft® Excel solver to result in the following DeltaP and SigmaP.

$$M = \text{desired depth of cut, } d_d \quad \text{3-39}$$

$$DeltaP(d_d) = 0.99970 d_d + 0.00079 d_d^2 \quad \text{3-40}$$

$$SigmaP(d_d) = -0.00011 d_d + 0.00007 d_d^2 \quad \text{3-41}$$

These equations are plotted in Figure 3-19 and Figure 3-20 below. Notice the curve of DeltaP is almost a straight line with a forty five degree slope, i.e., it is close to the ideal curve. This implies the turning operation will actually make a depth of cut very close to the desired depth of cut. The SigmaP curve in Figure 3-20 is shaped as one would expect in that it is positive everywhere, since a negative variation cannot be imparted to a dimension.

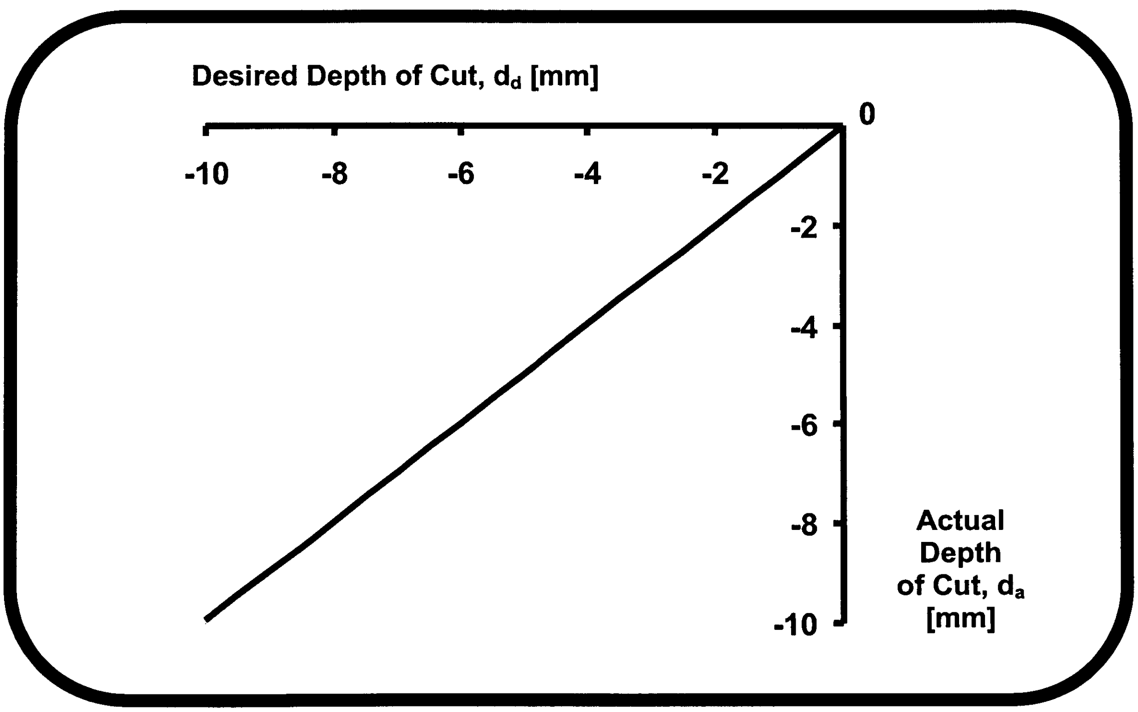

Figure 3-19: DeltaP for turning operation used for validation

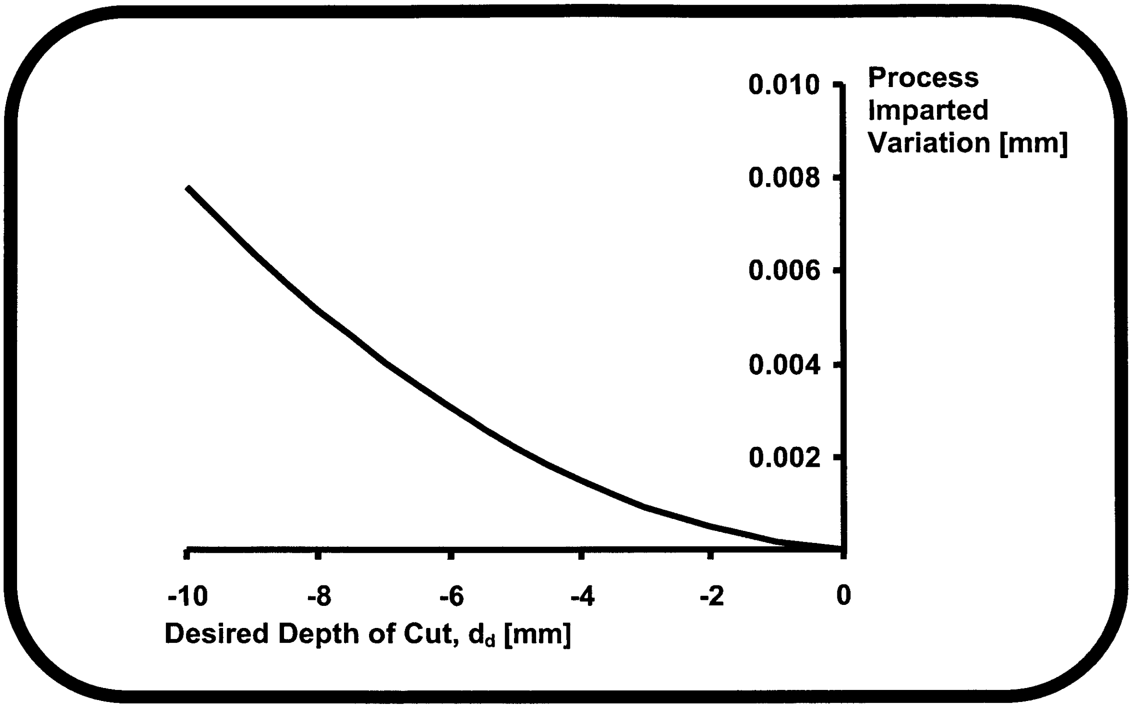

Figure 3-20: SigmaP for turning operation used for validation

Now that DeltaP and SigmaP are known, Equations 3-25 and 3-26 are used to created closed form equations for the output mean and standard deviation of diameters produced by the turning operation with the specified parameters.

3.6.1.4 Closed-form Results versus Simulation Results

The final part of validating the theory is to compare calculated values of the output mean and standard deviation to "true" values. The math model used to simulate the sample process capability data is used to simulate output data that is accepted as the true value the turning operation would produce. Three values for the incoming mean, from 32.0 to 39.0 mm, and three values for the incoming standard deviation, from 0.2 to 0.6 mm, are used for a total of nine combinations. The values for the incoming mean and standard deviation are used with a random number generator to produce a thousand values that are normally distributed. The mean and standard deviation are computed for these thousand values and are used as inputs to the closed-form equations. The incoming distribution values are listed in the first two columns of Table 3-4 along with the computed values for the thousand values in parentheses. The simulated and calculated output values are listed in the third through sixth columns with the percent differences listed in the last two columns of Table 3-4.

Table 3-4: Closed-form results versus simulated results

Incoming Distribution of Diameters		Simulated "True" Values for Output Distribution of Diameters		Closed-Form Results for Output Distribution of Diameters		% Error	
Mean μ_{in} (Calculated μ of 1,000 Generated Random Numbers)	Standard Deviation σ_{in} (Calculated σ of 1,000 Generated Random Numbers)	Output Mean $\mu_{sim,out}$	Standard Deviation $\sigma_{sim,out}$	Mean $\mu_{calc,out}$	Standard Deviation $\sigma_{calc,out}$	Mean $\frac{\text{(calc-sim)}}{\text{sim}}$	Standard Deviation $\frac{\text{(calc-sim)}}{\text{sim}}$
39.0 (38.978)	0.6 (0.5910)	30.067	0.0107	30.067	0.0107	0.00%	-0.14%
39.0 (39.005)	0.4 (0.4074)	30.067	0.0087	30.067	0.0088	0.00%	0.43%
39.0 (38.995)	0.2 (0.2057)	30.067	0.0073	30.067	0.0071	0.00%	-3.34%
35.0 (34.986)	0.6 (0.5913)	30.021	0.0053	30.021	0.0054	0.00%	1.63%
35.0 (34.998)	0.4 (0.3978)	30.021	0.0038	30.021	0.0040	0.00%	3.91%
35.0 (35.013)	0.2 (0.2117)	30.021	0.0028	30.021	0.0028	0.00%	2.14%
32.0 (31.973)	0.6 (0.6034)	30.004	0.0021	30.004	0.0022	0.00%	3.00%
32.0 (32.009)	0.4 (0.3943)	30.004	0.0014	30.004	0.0015	0.00%	4.47%
32.0 (32.004)	0.2 (0.2121)	30.004	0.0008	30.004	0.0009	0.00%	10.97%

All values in millimeters.

The comparison in Table 3-4 demonstrates the closed-form equations' ability to calculate the output mean and output standard deviation for any given input mean and input standard deviation. The closed-form solution for the mean matched the simulated results perfectly. The values for the output standard deviations were very close to the simulated result. All closed-form calculated standard deviation values were within 11% of the corresponding simulated value and represent the "true" values very well on an absolute scale.

3.6.1.5 Validation Using Actual Measurements

The above validation compares results calculated with the closed-form equations with results from simulation. The closed-form equations were constructed using data generated by Monte Carlo simulation. This section uses actual measurements from turned parts to construct closed-form equations. The equations are then used to calculate the expected mean and standard deviation for a known input mean and standard deviation. The calculated values are then compared to the mean and standard deviation of the actual output measurements. Three separate cases are presented.

The first case consists of process capability data measured at an American engine manufacturing plant. The mean and standard deviations are based on measurements of diameters made before and after a turning operation. There are a total of five sets of data. Three sets are used to construct the closed-form equations. These data are presented in Table 3-5. The two other sets are used as test sets.

Table 3-5: Case 1 – measurements used to construct closed-form equations

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Mean [mm]	Measured Outgoing Std Dev [mm]	Outgoing Target [mm]
50.998	0.0108	49.988	0.0011	49.99
50.995	0.0073	49.988	0.0008	49.99
50.992	0.0097	49.987	0.0008	49.99

Table 3-6: Case 1 – comparison of measured and calculated means

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Mean [mm]	Calculated Outgoing Mean [mm]	% Error
51.015	0.0069	49.984	49.988	0%
51.016	0.0084	49.988	49.988	0%

Table 3-7: Case 1 – comparison of measured and calculated standard deviations

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Std Dev [mm]	Calculated Outgoing Std Dev [mm]	% Error
51.015	0.0069	0.00122	0.0013	-6.3%
51.016	0.0084	0.00127	0.0014	-9.6%

The Case 1 calculated values for the output mean shown in Table 3-6 show that they match very well with the actual measurements. The standard deviation values calculated also matched well with the actual measured values. The percent error was less than ten percent in both sets. The values are summarized in Table 3-7.

The second case consists of process capability data measured at the same engine manufacturing plant. The mean and standard deviations are based on measurements of diameters made before and after a turning operation. There are a total of four sets of data. Three sets are used to construct the closed-form equations. One set is used as a test set. The data sets used to construct the DeltaP and SigmaP functions are listed in Table 3-8.

Table 3-8: Case 2 – measurements used to construct closed-form equations

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Mean [mm]	Measured Outgoing Std Dev [mm]	Outgoing Target
63.8190	0.0103	62.9936	0.0010	62.99
63.8017	0.0099	62.9903	0.0012	62.99
63.8187	0.0080	62.9877	0.0012	62.99

Table 3-9: Case 2 – comparison of measured and calculated means

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Mean [mm]	Calculated Outgoing Mean [mm]	% Error
63.8186	0.0083	62.9926	63.1392934	0.2%

Table 3-10: Case 2 – comparison of measured and calculated standard deviations

Measured Incoming Mean [mm]	Measured Incoming Std Dev [mm]	Measured Outgoing Std Dev [mm]	Calculated Outgoing Std Dev [mm]	% Error
63.8186	0.0083	0.0012	0.000915118	-23.7%

The Case 2 calculated value for the output mean shown in Table 3-9 matches well with the actual measurement. The standard deviation value is off by about twenty four percent from the actual measured value as seen in Table 3-10. This larger error is explained by the fact the process capability data used to construct the equations were tightly clustered with overlap and few in number. Some spread in the mean and standard deviations allows a better fit for DeltaP and SigmaP. A better fit for these functions would have produced results closer to the measured value.

The third case consists of measurements of parts turned in the Edgerton Machine Shop at MIT specifically for this thesis. The mean and standard deviations are based on measurements of diameters made before and after a turning operation on a manual lathe. Four "runs" were made. Each "run" consisted of measuring thirty rods with a hand-held micrometer, turning them to remove 0.1 mm, and re-measuring them. Each "run" is used as a set of data. Three sets are used to construct the closed-form equations. One set is used as a test set.

Table 3-11: Case 3 – measurements used to construct closed-form equations

Measured Incoming Mean [in]	Measured Incoming Std Dev [in]	Measured Outgoing Mean [in]	Measured Outgoing Std Dev [in]	Outgoing Target [in]
0.498	0.00027	0.398	0.00040	0.4
0.500	0.00016	0.391	0.00049	0.4
0.398	0.00040	0.296	0.00050	0.3

Table 3-12: Case 3 – comparison of measured and calculated means

Measured Incoming Mean [in]	Measured Incoming Std Dev [in]	Measured Outgoing Mean [in]	Calculated Outgoing Mean [in]	% Error
0.398	0.00168	0.29776	0.295	-1.1%

Table 3-13: Case 3 – comparison of measured and calculated standard deviations

Measured Incoming Mean [in]	Measured Incoming Std Dev [in]	Measured Outgoing Std Dev [in]	Calculated Outgoing Std Dev [in]	% Error
0.398	0.00168	0.00273	0.00217	-20.4%

The Case 3 calculated value for the output mean shown in Table 3-12 matches well with the actual measurement. The standard deviation value is off by about twenty percent from the standard deviation of the actual measured values as seen in Table 3-13. This error is due to the fact there were not many sets of data used when fitting DeltaP and SigmaP. Also, measurement error more than likely plagues these values since they were done with a hand-held micrometer that probably has low repeatability.

3.6.2 Rolling

Rolling is the process of changing the thickness of a workpiece by applying compressive forces through a set of rolls. The rolls rotate and pull the workpiece into the gap between the rollers. The workpiece thickness is reduces and exits with the modified thickness.

3.6.2.1 Physics-based Math Model

This section develops a math model for a rolling operation to calculate the actual change in thickness a workpiece undergoes as a function of rolling parameters. The compressive forces that act to reduce the thickness of the workpiece also act to deflect the rolls. Figure 3-21 shows a schematic of a typical rolling operation. Two rollers are shown operating on a piece of sheet metal. The equations presented in this thesis are based on those presented by Kalpakjian (Kalpakjian and Schmid 2001).

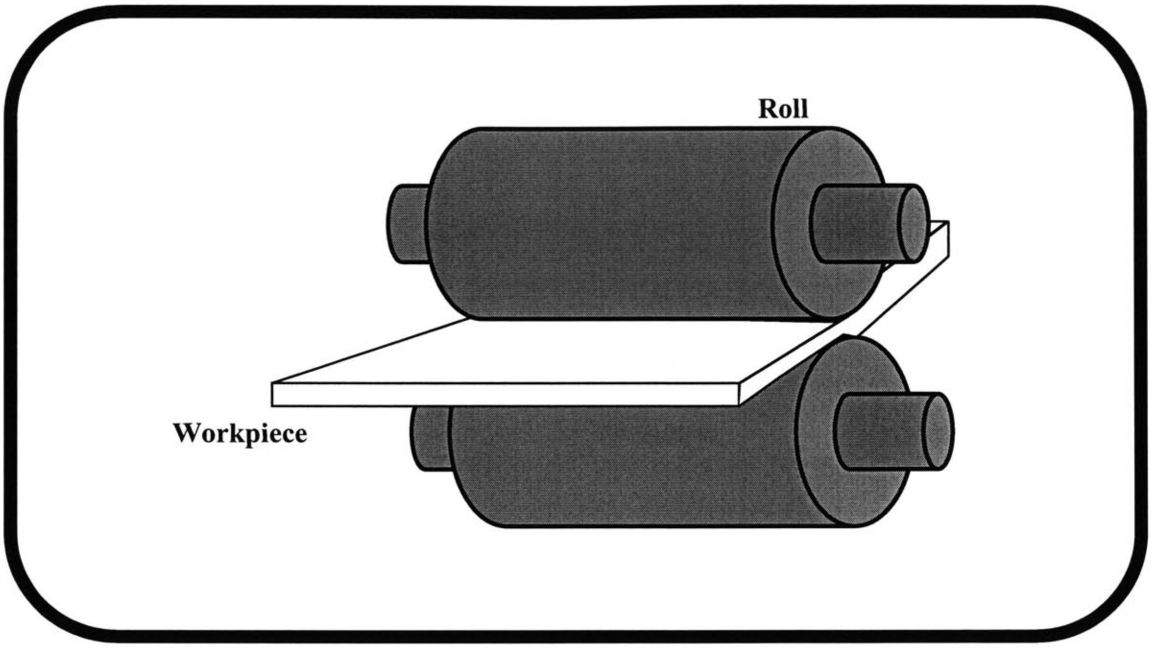

Figure 3-21: Schematic of rolling operation

The math model for rolling in this thesis calculates the change in thickness of the workpiece at its center. This calculation is made by modeling the rolls as two beams with a distributed force applied to each. The distributed force is the roll force per unit length. The applied force causes the rolls to deflect. This deflection results in a final workpiece thickness greater than the roll gap, the distance between the rolls in their non-deflected state. A diagram of the force acting on the rolls is shown below in Figure 3-22.

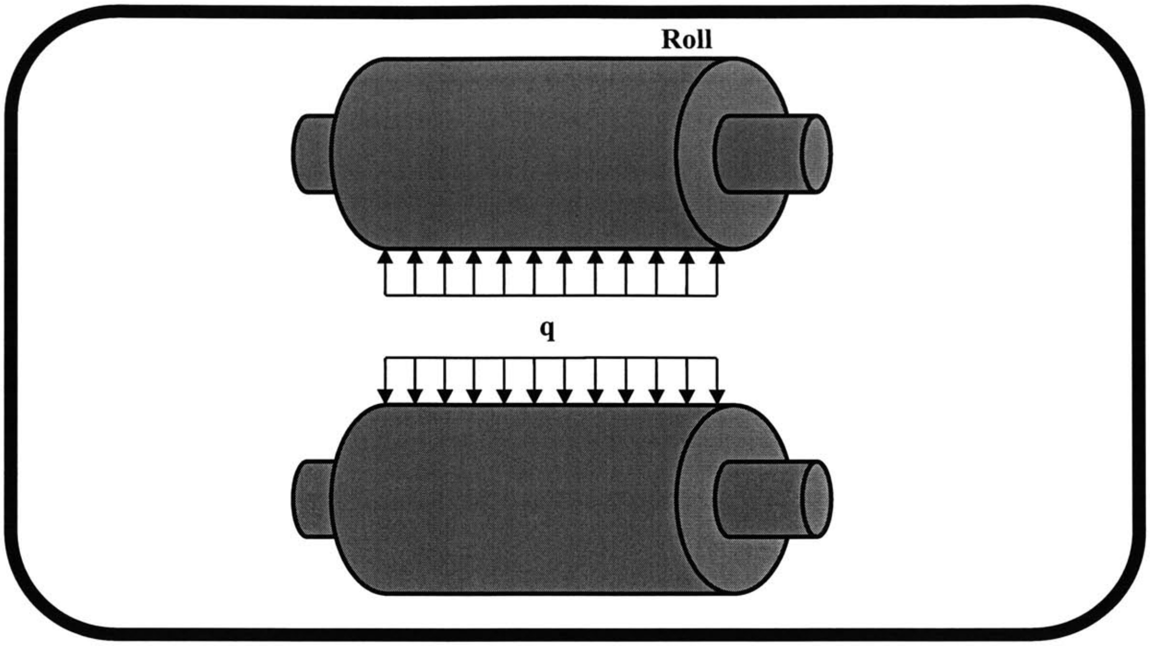

Figure 3-22: Force acting on rolls

The model assumes the workpiece to have a width equal to the rolls. Also, the model assumes the rolling operation acts on discrete workpieces (as compared to continuous sheet rolling) so that distributions of parts can be input into the operation.

The equation for the distributed force acting on a roll (roll force per unit width) depends on the roll-strip contact length (L), the width of the workpiece (w), and the average true stress of the workpiece in the roll gap (Y_{avg}). The relationship between these values is

$$q = \frac{F}{w} = LY_{avg} \tag{3-42}$$

The roll-strip contact length (L) is given by the following equation.

$$L = \sqrt{R\left(h_o - h_{f,desired}\right)} \tag{3-43}$$

where R is the roll radius, h_o is the initial thickness of the workpiece, and h_f is the final thickness of the workpiece.

The average true stress of the workpiece in the roll gap (Y_{avg}) can be calculated using the following equation

$$Y_{avg} = K\varepsilon^{n} \qquad \text{3-44}$$

where K is the strength coefficient for the workpiece material, epsilon is the true strain of the material in the roll gap, and n is the strain-hardening (or work-hardening) exponent. The true strain of the material in the roll gap can be computed with the following equation

$$\varepsilon = \ln\left(\frac{h_o}{h_f}\right) \qquad \text{3-45}$$

The deflection the rollers undergo when subjected to the distributed load is

$$\delta_{roll} = \frac{5qw^4}{384E_{roll}I_{roll}} \qquad \text{3-46}$$

where E_{roll} is the modulus of elasticity of the roll material, and I_{roll} is the moment of inertia of a roll and is described by the following equation.

$$I_{roll} = \frac{\pi R^4}{4} \qquad \text{3-47}$$

The final thickness, $h_{f,actual}$, of the workpiece is the desired thickness (roll gap) plus the deflection of both rollers. The equation for the final thickness is

$$h_{f,actual} = h_{f,desired} + 2\delta_{roll} \qquad \text{3-48}$$

Combining Equation 3-42 through Equation 3-48 and simplifying results in the following equation for the final workpiece thickness.

$$h_{f,actual} = h_{f,desired} + \frac{5Kw^4\sqrt{R\left(h_o - h_{f,desired}\right)}\left[\ln\left(\frac{h_o}{h_f}\right)\right]^n}{48E_{roll}\pi R^4} \quad \text{3-49}$$

The above equation represents a simple physics-based math model for the thickness of the center of a workpiece upon exiting a rolling operation and is a sufficient model for the purpose of this thesis. A more complete treatment of modeling a sheet rolling operation including the variation in sheet thickness is presented in work by Ota (Ota 2000).

3.6.2.2 Simulation of Process Capability Data

A set of Monte Carlo simulations is run to generate sample process capability data. Equation 3-49 is used to calculate the thickness of the center of a workpiece upon exiting a rolling operation. The strain-hardening coefficient, n, is treated as a stochastic variable. Treating n as a stochastic variable is simply a method "introducing" variation into the process model.

The model is used to generate ten sets of process capability data. Each set contains a value for an input mean thickness, input standard deviation of the thickness, output mean thickness, and output standard deviation of the thickness. These ten sets of data represent the type of information that would be retrieved from a Process Capability Database in a manufacturing firm and used to construct DeltaP and SigmaP functions for a rolling operation. To be completely unbiased, the ten values for the input means and the ten values of the standard deviations are randomly generated and paired. Each simulation run comprises one thousand calculations.

All operation parameters except for the strain-hardening coefficient, n, and the original thickness of the workpiece, h_o, are set. The average value and standard deviation for the strain-hardening coefficient are set. One thousand normally distributed random numbers are generated to represent values of n. Next, one thousand normally distributed numbers are generated that have a mean equal to the first value of input mean and a standard deviation equal to the value of the first input standard deviation. These values are generated with a random number generator in Microsoft® Excel. Using the generated numbers, the fixed parameter values, and Equation 3-49, one thousand $h_{f,actual}$'s are calculated. The mean and standard deviation are calculated for these one thousand values. The calculated mean is the output mean for the given input mean and standard deviation, and the calculated standard deviation is the output standard deviation for the given input mean and standard deviation. These steps are repeated ten times to generate the remaining simulated process capability data. This method is presented graphically in Figure 3-23.

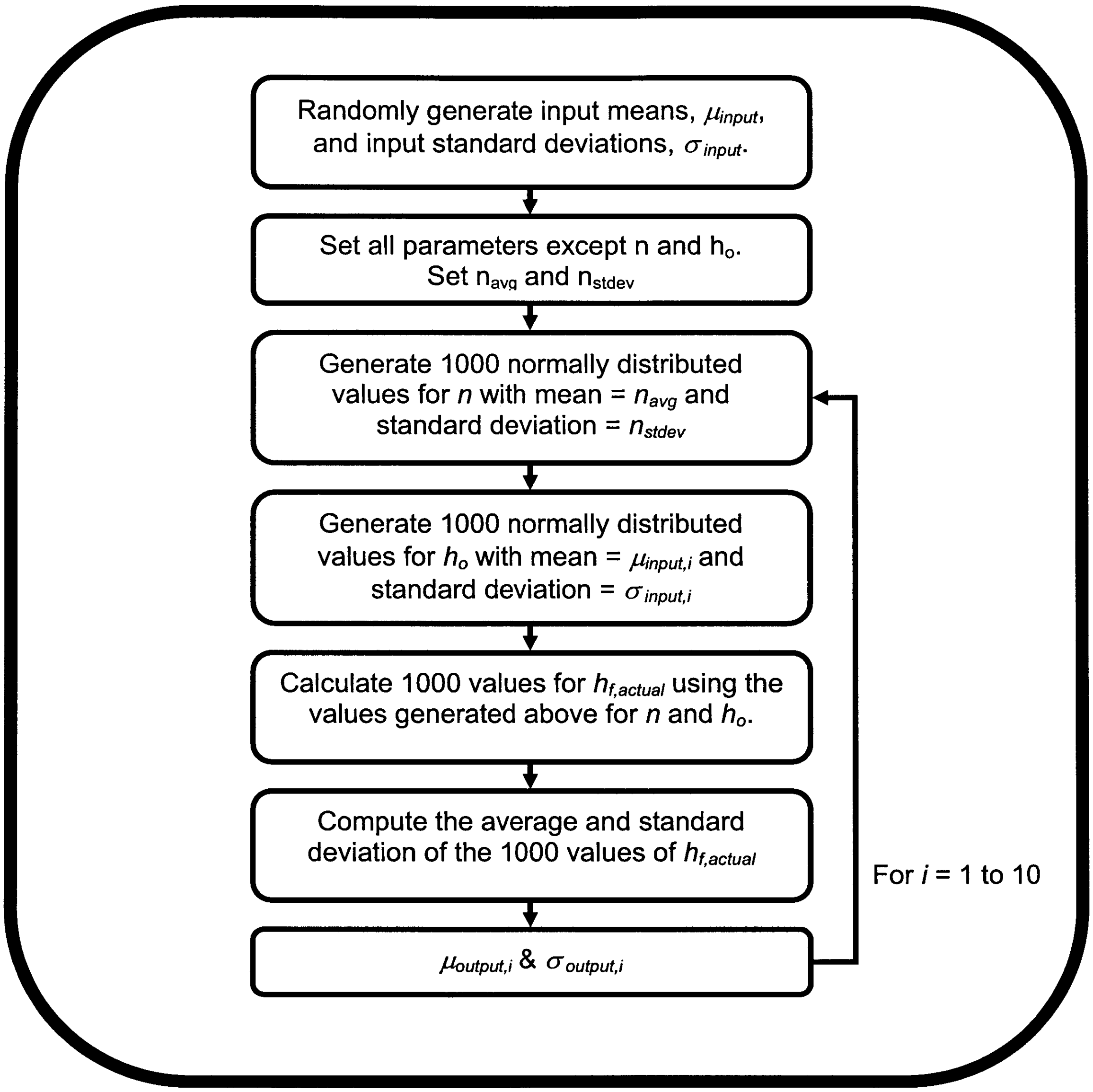

Figure 3-23: Algorithm for Monte Carlo simulation method for rolling

The workpiece material is copper. The rolls are made of steel. The parameters of the workpiece and rolls are summarized in Table 3-14 and Table 3-15 respectively.

Table 3-14: Workpiece parameters for rolling simulation

Width of workpiece, w	200.0 mm
Mean of strain-hardening coefficient, n_{avg}	0.54
Standard Deviation of strain-hardening coefficient, n_{stdev}	0.05
Strength coefficient, K	315 MPa
Desired thickness, $h_{f,desired}$	20.0 mm

Table 3-15: Roll parameters for rolling simulation

Width of roll, w	200.0 mm
Radius of roll, R	100.0 mm
Modulus of Elasticity of Roll, E_{roll}	200,000 MPa

Since the desired final thickness of the workpiece is 20.0 mm, the incoming mean should be around 25.0 mm. Values for input means will be randomly generated to fall close to 25.0 mm. Values for standard deviations will be randomly generated to fall close to 0.5 mm.

Table 3-16: Simulated process capability data for rolling operation

Input Mean, μ_{in} [mm] Randomly Generated	Input Standard Deviation, σ_{in} [mm] Randomly Generated	Output Mean, μ_{out} [mm] Simulated	Output Standard Deviation, σ_{out} [mm] Simulated
22.816	0.7852	20.002	0.0007
27.198	0.3309	20.006	0.0004
24.522	0.8948	20.004	0.0008
29.237	0.6397	20.008	0.0006
28.466	0.0776	20.007	0.0004
27.399	0.3291	20.006	0.0004
27.557	0.7207	20.006	0.0007
25.493	0.1918	20.005	0.0004
22.449	0.2332	20.002	0.0003
24.394	0.5911	20.004	0.0006

3.6.2.3 Creation of DeltaP and SigmaP

The DeltaP and SigmaP functions can be created for the rolling operation following the technique outlined in Section 3.4. Remember that DeltaP and SigmaP are functions of M, the desired dimensional change. In the case of rolling, M is equal to the desired change in thickness, so DeltaP = f($\Delta_{thickness}$) and SigmaP = f($\Delta_{thickness}$). The process capability data in Table 3-3 is used with the Microsoft® Excel solver to result in the following DeltaP and SigmaP.

$$M = \text{desired change in thickness, } \Delta_{thickness} \tag{3-50}$$

$$DeltaP = (0.9991)(\Delta_{thickness}) + (0.0000045)(\Delta_{thickness})^2 \tag{3-51}$$

$$SigmaP = (-0.000093)(\Delta_{thickness}) + (0.000006)(\Delta_{thickness})^2 \tag{3-52}$$

These equations are plotted in Figure 3-24 and Figure 3-25 below. Notice the curve of DeltaP is almost a straight line but diverges from the ideal case represented by the dotted line. Unfortunately, this implies the rolling operation will decrease the thickness of a sheet by an amount less than desired. The SigmaP curve in Figure 3-25 is shaped as one would expect in that it is positive everywhere, since a negative variation cannot be imparted to a dimension.

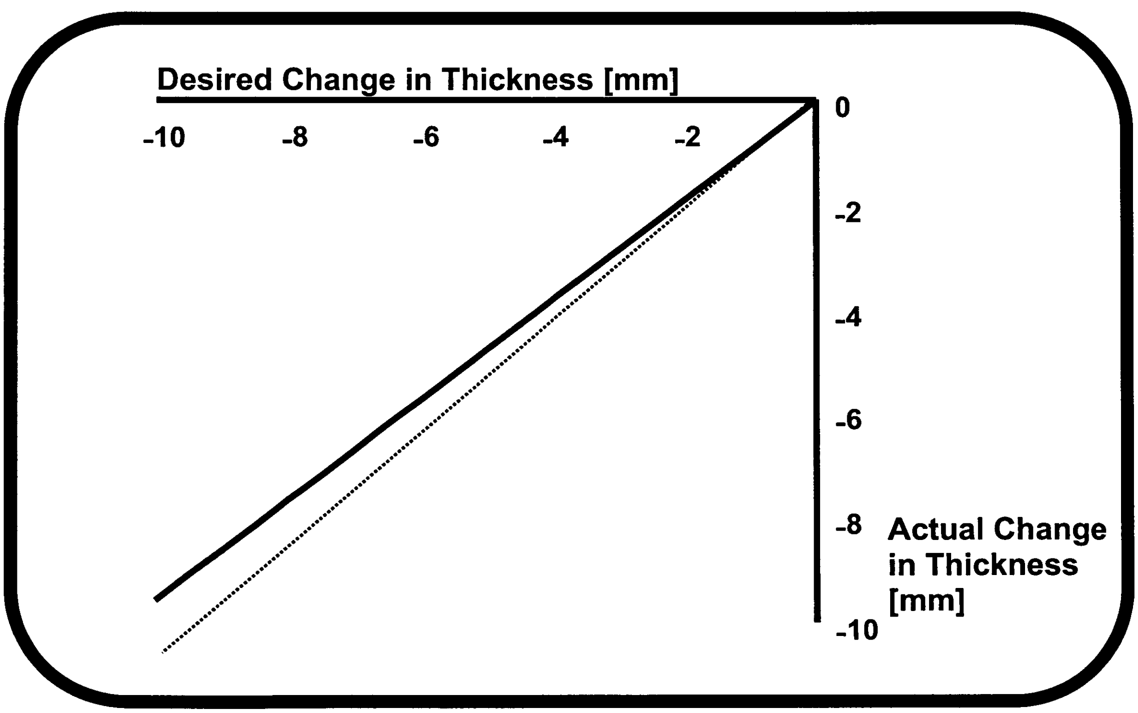

Figure 3-24: DeltaP for rolling operation used for validation

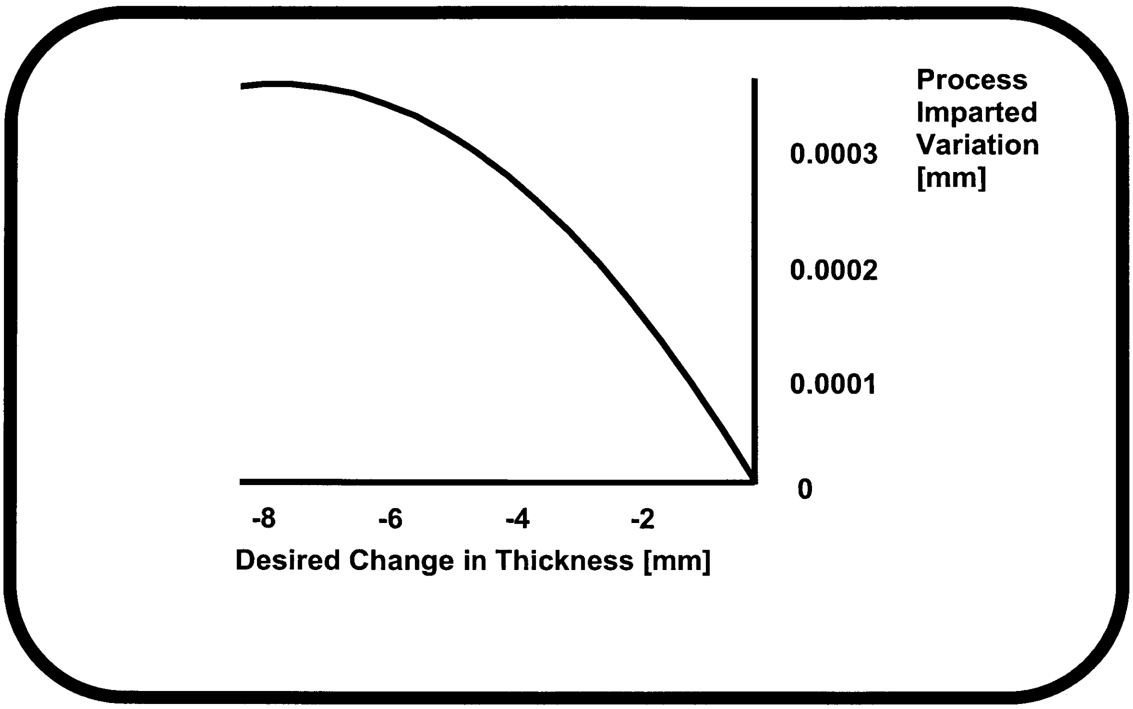

Figure 3-25: SigmaP for rolling operation used for validation

Now that DeltaP and SigmaP are known, Equations 3-25 and 3-26 are used to created closed form equations for the output mean and standard deviation of thicknesses produced by the rolling operation with the specified parameters.

3.6.2.4 Closed-form Results versus Simulation Results

The final part of validating the theory is to compare calculated values of the output mean and standard deviation to "true" values. The math model used to simulate the sample process capability data is used to simulate output data that is accepted as the true value the turning operation would produce. Three values for the incoming mean, from 22.0 to 27.0 mm, and three values for the incoming standard deviation, from 0.2 to 0.6 mm, are used for a total of nine combinations. The values for the incoming mean and standard deviation are used with a random number generator to produce a thousand values that are normally distributed. The mean and standard deviation are computed for these thousand values and are used as inputs to the closed-form equations. The incoming distribution values are listed in the first two columns of Table 3-17 along with the computed values for the thousand values in parentheses. The simulated and calculated output values are listed in the third through sixth columns with the percent differences listed in the last two columns of Table 3-17.

Table 3-17: Closed-form results versus simulated results

Incoming Distribution of Sheet Thickness		Simulated "True" Values for Output Distribution of Sheet Thickness		Closed-Form Results for Output Distribution of Sheet Thickness		% Error	
Mean μ_{in} (Calculated μ of 1,000 Generated Random Numbers)	Standard Deviation σ_{in} (Calculated σ of 1,000 Generated Random Numbers)	Output Mean μ_{simout}	Standard Deviation σ_{simout}	Mean $\mu_{calcout}$	Standard Deviation $\sigma_{calcout}$	Mean (calc-sim) / sim	Standard Deviation (calc-sim) / sim
27.0 (26.998)	0.6 (0.6261)	20.006	0.00059	20.006	0.00061	0.00%	2.60%
27.0 (26.988)	0.4 (0.4108)	20.006	0.00047	20.006	0.00048	0.00%	1.60%
27.0 (26.990)	0.2 (0.2041)	20.006	0.00038	20.006	0.00039	0.00%	3.60%
25.0 (25.004)	0.6 (0.6089)	20.004	0.00060	20.004	0.00059	0.00%	-3.00%
25.0 (25.000)	0.4 (0.4018)	20.004	0.00046	20.004	0.00045	0.00%	-1.20%
25.0 (25.000)	0.2 (0.2054)	20.004	0.00036	20.004	0.00036	0.00%	-0.30%
22.0 (21.987)	0.6 (0.5985)	20.002	0.00054	20.002	0.00053	0.00%	-3.00%
22.0 (21.992)	0.4 (0.4032)	20.002	0.00039	20.002	0.00038	0.00%	-4.10%
22.0 (22.001)	0.2 (0.2001)	20.002	0.00025	20.002	0.00023	0.00%	-6.90%

All values in millimeters.

The comparison in Table 3-17 demonstrates the closed-form equations' ability to calculate the output mean and output standard deviation for a variety of given input means and input standard deviations. The closed-form solution for the mean matched the simulated results perfectly. The values for the output standard deviations were very close to the simulated results. All standard deviation values were within 7% of the corresponding simulated value.

3.6.3 Sheet Bending

Sheet bending is a sheet-metal forming process. In sheet bending, a force is applied to a metal sheet to deform it around a die. The force is then removed, and the sheet maintains some of the imposed deformation.

3.6.3.1 Physics-based Math Model

Sheet bending is different from the two previous operations presented, turning and rolling, in the sense that it is does not change a dimension of the incoming workpiece; it affects the shape of the incoming workpiece. For example, in turning, the operation removes material from the workpiece to produce the desired diameter. In sheet bending, the workpiece is a sheet of metal that is bent to produce a desired shape. In order to apply the theory of DeltaP and SigmaP, the actual dimensional change must be a function of the desired dimensional change, *M*. To accomplish this for sheet bending, the design of the component being made will be considered, instead of focusing on the workpiece. The characteristic of interest will be the distance from the -A- datum to the outer surface of the part. This characteristic is labeled *X* in Figure 3-26.

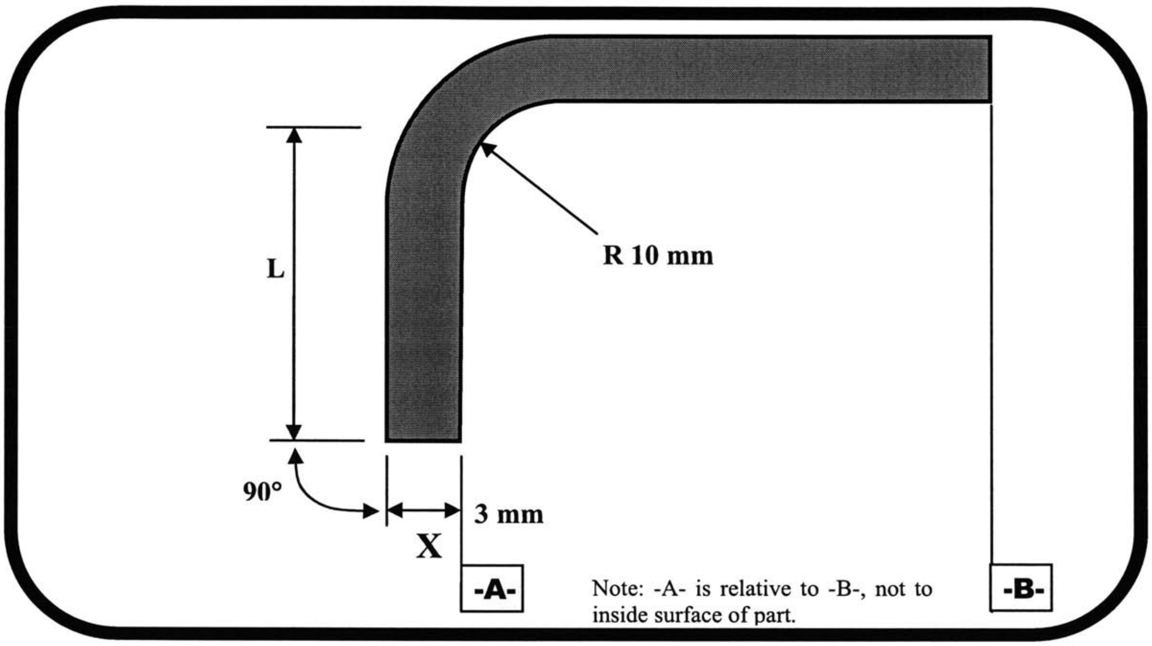

Figure 3-26: Final desired shape of bent component

The deformed shape is held by the metal due to plastic deformation; however, elastic recovery is present. This elastic recovery is known as springback. To produce the desired shape, over bending is required to compensate for springback. For example, to produce the desired shape in Figure 3-26, the part must be initially bent greater than 90°. The die is made so that an incoming workpiece with the nominal thickness and nominal material properties will be over bent and springback to achieve the final desired shape. Figure 3-27 shows a picture of a workpiece in the die after it has been release and allowed to springback to the desired shape. Figure 3-28 shows a graphical representation of springback by representing the part in the fully deflected state with dotted lines and the part after springback with solid lines.

In reality, the workpiece thickness varies from piece to piece; therefore, the final dimension of the characteristic of interest, *X*, also varies. If no springback existed, then the dimension *X* would be equal to the thickness of the workpiece. The existence of springback complicates the understanding of how the dimension X will vary as the variation of the workpiece thickness varies. The theory developed in this thesis assists in calculating the mean and variation of X as a function of the workpiece thickness distribution.

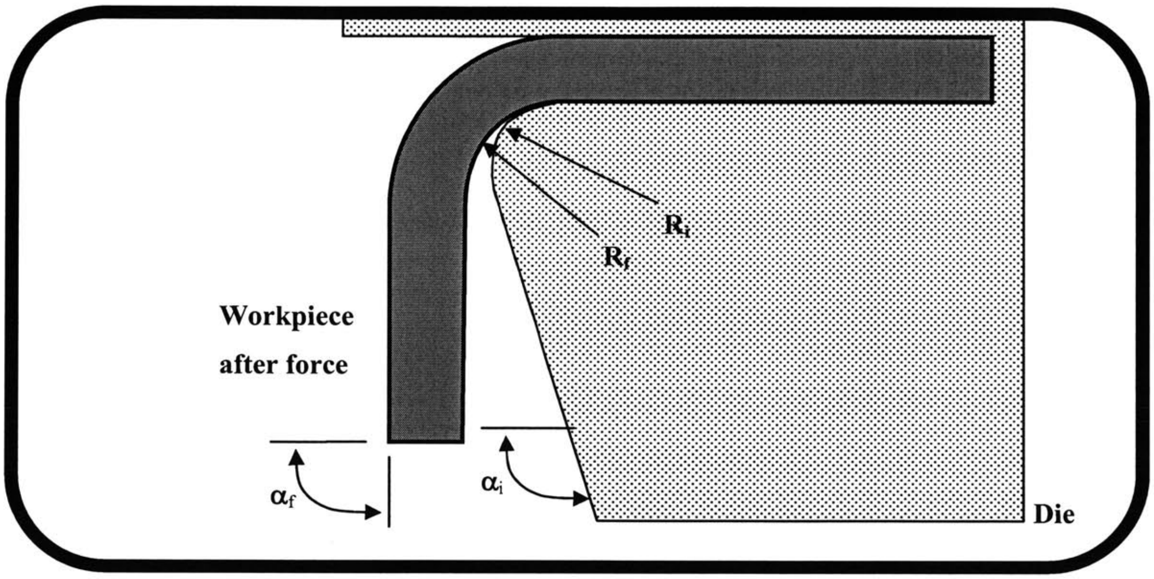

Figure 3-27: Workpiece in the die

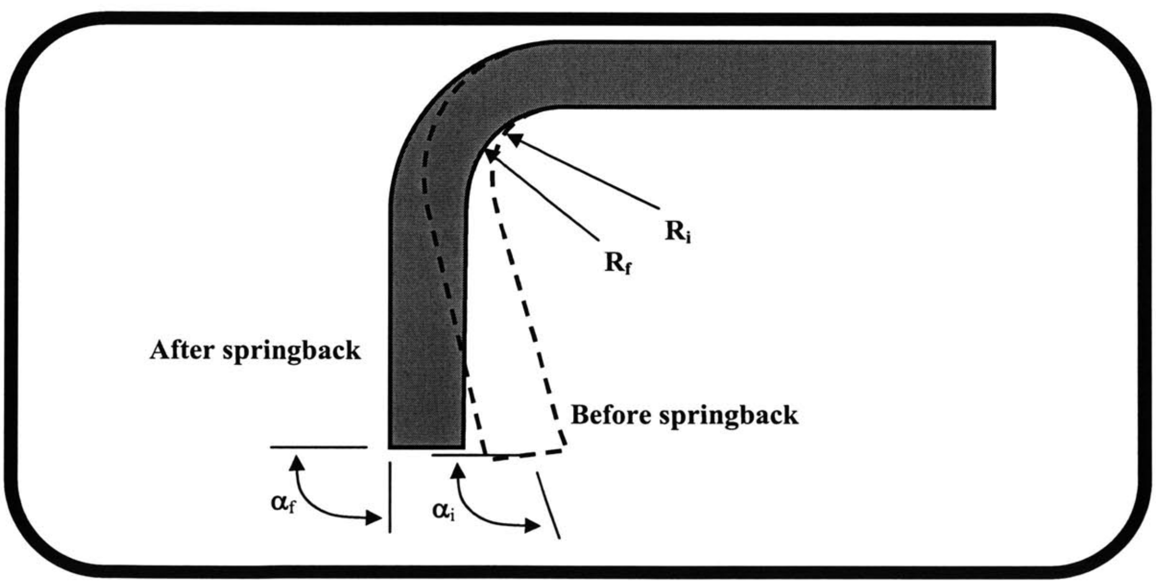

Figure 3-28: Graphical representation of springback

This math model for a sheet bending operation developed in this section is based on equations presented in Kalpakjian (Kalpakjian and Schmid 2001). The final radius of a workpiece after springback occurs is given by R_f. The equation for springback is a function of the die radius (R_i), the thickness of the workpiece (T), the Modulus of Elasticity (E) of the workpiece, and the Yield Stress (Y) of the workpiece and is shown below.

$$\frac{R_i}{R_f} = 4\left(\frac{R_i Y}{ET}\right)^3 - 3\left(\frac{R_i Y}{ET}\right) + 1 \qquad \text{3-53}$$

The initial and final angles are related to the radii per the following equation.

$$R_i \alpha_i = R_f \alpha_f \qquad \text{3-54}$$

Now, the characteristic of interest can be related to the operation parameters.

$$X = \frac{L}{\tan \alpha_f} + T \qquad \text{3-55}$$

where L is the length from the center of the die radius to the end of the bent workpiece. Combining the above three equations produces an equation for X as a function of die and workpiece parameters.

$$X = \frac{L}{\tan\left[\alpha_i\left(4\left(\frac{R_i Y}{ET}\right)^3 - 3\left(\frac{R_i Y}{ET}\right) + 1\right)\right]} + T \qquad \text{3-56}$$

Equation 3-56 allows for the characteristic of interest to be calculated for any combination of process and workpiece parameters. This equation will be used in the next section to run Monte Carlo simulations.

3.6.3.2 Simulation of Process Capability Data

A set of Monte Carlo simulations is run to generate sample process capability data. Equation 3-56 is used to calculate X (see Figure 3-26) upon exiting the sheet bending operation. The yield stress of the workpiece material, Y, is treated as a stochastic variable. Treating Y as a stochastic variable is simply a method "introducing" variation into the process model.

The model is used to generate ten sets of process capability data. Each set contains a value for an input mean thickness, input standard deviation of the thickness, output mean for X, and output standard deviation of X. These ten sets of data represent the type of information that would be retrieved from a Process Capability Database in a manufacturing firm and used to construct DeltaP and SigmaP functions for a sheet bending operation. To be completely unbiased, the ten values for the input means and the ten values of the standard deviations are randomly generated and paired. Each simulation run comprises one thousand calculations.

All operation parameters except for the yield stress of the workpiece material, Y, and the original thickness of the workpiece, T, are set. The average value and standard deviation for the yield stress are set. One thousand normally distributed random numbers are generated to represent values of Y. Next, one thousand normally distributed numbers are generated that have a mean equal to the first value of input mean and a standard deviation equal to the value of the first input standard deviation. These values are generated with a random number generator in Microsoft® Excel. Using the generated numbers, the fixed parameter values, and Equation 3-56, one thousand X's are calculated. The mean and standard deviation are calculated for these one thousand values. The calculated mean is the output mean for the given input mean and standard deviation, and the calculated standard deviation is the output standard deviation for the given input mean and standard deviation. These steps are repeated ten times to generate the remaining simulated process capability data. This method is presented graphically in Figure 3-29.

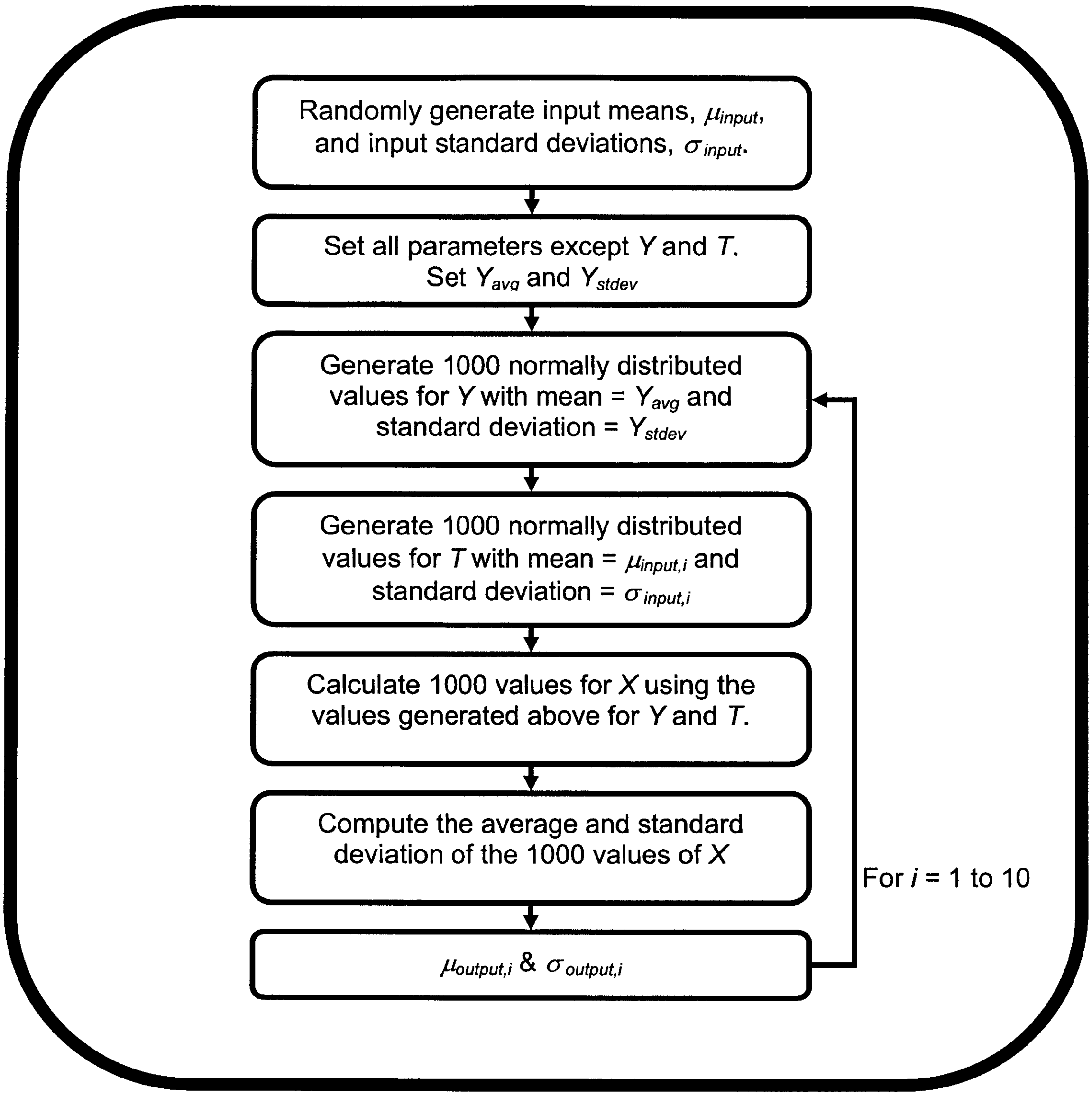

Figure 3-29: Algorithm for Monte Carlo simulation method for sheet bending

The workpiece material is steel. The parameters of the workpiece and die are summarized in Table 3-18 and Table 3-19 respectively.

Table 3-18: Workpiece parameters for sheet bending simulation

Desired mean of thickness	3.0 mm
Mean of yield stress, Y_{avg}	500 MPa
Standard Deviation of yield stress, Y_{stdev}	25 MPa
Modulus of elasticity	200,000 MPa

Table 3-19: Die parameters for sheet bending simulation

Radius of die, R_i	10.0 mm
Die angle, α_i	92.3°
Length from center of die radius to end of piece, L	20.0 mm

The desired final dimension of X is 3.0 mm. The incoming mean of the thickness should be around 3.0 mm, since the thickness is a major determining factor to the final value of X and the die in the example is made so that the springback associated with a 3.0 mm thick part would result in the final desired shape, an angle of ninety degrees. Allowing input dimensions less than the output target dimension is unique to the sheet bending operation. For turning and rolling, the input dimension had to be greater than the target dimension because both processes are reduction operations, whereas sheet bending is simply transforming the shape. The nature of sheet bending and the design example warrant the use of input material close to 3.0 mm; therefore, generating inputs far from this value does not make realistic sense. So, values for input means for thickness are randomly generated to fall close to 3.0 mm. Values for standard deviations are randomly generated to fall close to 0.05 mm.

Table 3-20: Simulated process capability data for sheet bending operation

Thickness, T Input Mean, μ_{in} [mm] Randomly Generated	Thickness, T Input Standard Deviation, σ_{in} [mm] Randomly Generated	X Output Mean, μ_{out} [mm] Simulated	X Output Standard Deviation, σ_{out} [mm] Simulated
2.894	0.064	2.924	0.062
3.108	0.026	3.082	0.044
2.978	0.072	2.986	0.067
2.782	0.031	2.847	0.049
3.174	0.057	3.130	0.058
3.118	0.048	3.088	0.052
3.123	0.061	3.089	0.062
3.023	0.036	3.018	0.048
2.874	0.045	2.909	0.054
2.972	0.080	2.980	0.069

3.6.3.3 *Creation of DeltaP and SigmaP*

The DeltaP and SigmaP functions can be created for the sheet bending operation following the technique outlined in Section 3.4. Remember that DeltaP and SigmaP are functions of *M*, the desired dimensional change. In the case of sheet bending, M is equal to the target value of *X* minus the input thickness.

$$M = X_{target} - T \tag{3-57}$$

Notice that M can be either positive or negative depending on the incoming workpiece thickness. The process capability data in Table 3-20 is used with the Microsoft® Excel solver to result in the following DeltaP and SigmaP.

$$DeltaP = -0.00033 + 0.26669(M) + 0.12428(M)^2 \tag{3-58}$$

$$SigmaP(M) = 0.04098 + 0.01850*(M) - 0.01289*(M)^2 \tag{3-59}$$

These equations are plotted in Figure 3-30 and Figure 3-31 below. Notice the curve of DeltaP is almost flat and diverges from the ideal case represented by the dotted line. This implies the sheet bending operation will produce parts that are not close to the desired output as the sheet thickness varies from the desired value, which makes sense because the dimension under consideration is heavily dependent on the incoming sheet thickness. The SigmaP curve in Figure 3-31 has an interesting shape. It has a positive slope as it crosses the y-axis. This means that slightly thicker pieces of sheet metal produce parts with less variation, which makes sense because the springback is less severe for thicker stock. The SigmaP curve does appear to be approaching the x-axis to the left of -1. These functions are only valid close to the y-axis due to the nature of sheet metal bending, i.e., the die was designed to handle sheets with a nominal thickness of 3mm, varying too much from this invalidates the equations generated with the retrieved manufacturing data.

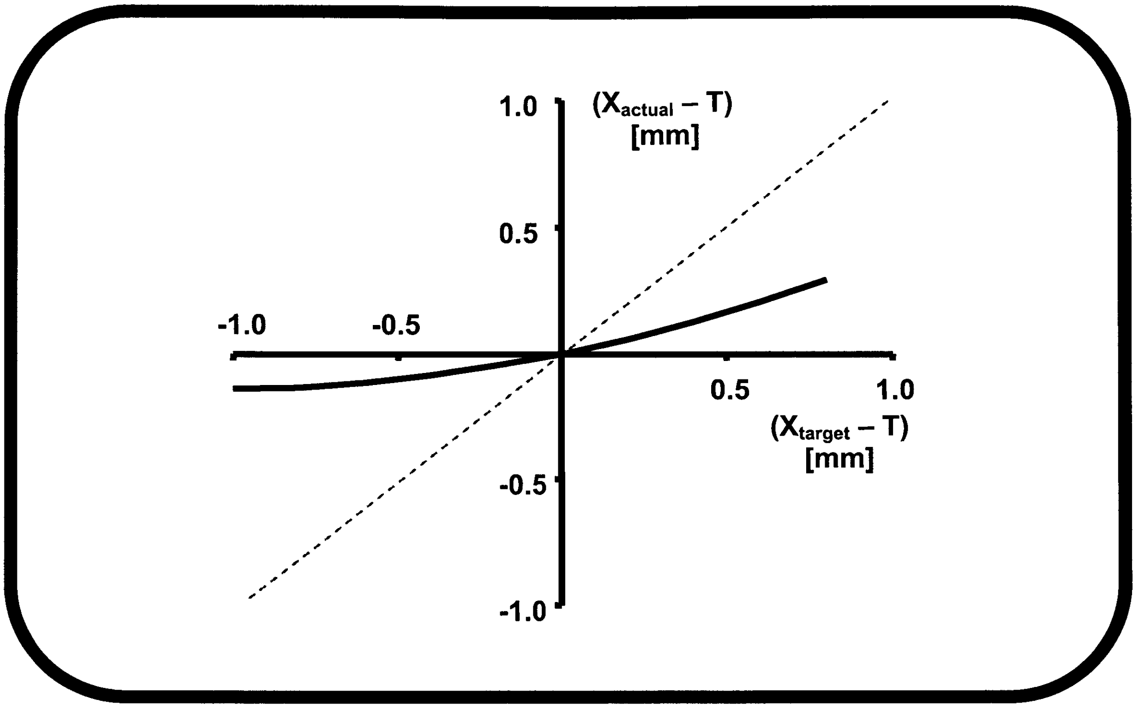

Figure 3-30: DeltaP for sheet bending operation used for validation

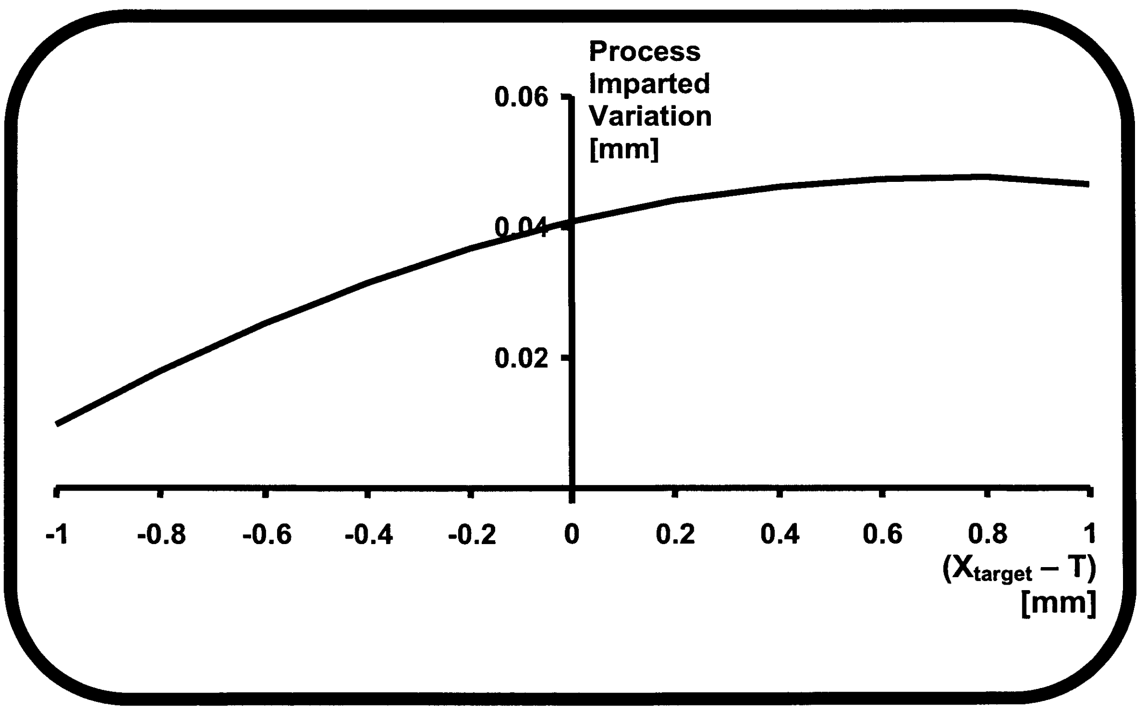

Figure 3-31: SigmaP for sheet bending operation used for validation

Now that DeltaP and SigmaP are known, Equations 3-25 and 3-26 are used to created closed form equations for the output mean and standard deviation of thicknesses produced by the sheet bending operation with the specified parameters.

3.6.3.4 Closed-form Results versus Simulation Results

The final part of validating the theory is to compare calculated values of the output mean and standard deviation to "true" values. The math model used to simulate the sample process capability data is used to simulate output data that is accepted as the true value the turning operation would produce. Three values for the incoming mean, from 2.9 to 3.1 mm, and three values for the incoming standard deviation, from 0.01 to 0.10 mm, are used for a total of nine combinations. The values for the incoming mean and standard deviation are used with a random number generator to produce a thousand values that are normally distributed. The mean and standard deviation are computed for these thousand values and are used as inputs to the closed-form equations. The incoming distribution values are listed in the first two columns of Table 3-21 along with the computed values for the thousand values in parentheses. The simulated and calculated output values are listed in the third through sixth columns with the percent differences listed in the last two columns of Table 3-21.

Table 3-21: Closed-form results versus simulated results for sheet bending

Incoming Distribution for Sheet Thickness		Simulated "True" Values for Output Distribution of *X* Dimension		Closed-Form Results for Output Distribution of *X* Dimension		% Error	
Mean μ_{in} (Calculated μ of 1,000 Generated Random Numbers)	Standard Deviation σ_{in} (Calculated σ of 1,000 Generated Random Numbers)	Output Mean $\mu_{sim,out}$	Standard Deviation $\sigma_{sim,out}$	Mean $\mu_{calc,out}$	Standard Deviation $\sigma_{calc,out}$	Mean (calc-sim)/sim	Standard Deviation (calc-sim)/sim
3.1 (3.102)	0.10 (0.105)	3.077	0.087	3.077	0.089	-0.01%	2.35%
3.1 (3.101)	0.05 (0.049)	3.076	0.054	3.075	0.054	-0.01%	0.20%
3.1 (3.101)	0.01 (0.010)	3.075	0.039	3.075	0.040	0.00%	0.61%
3.0 (3.003)	0.10 (0.095)	3.005	0.078	3.003	0.081	-0.07%	3.46%
3.0 (3.003)	0.05 (0.049)	3.000	0.054	3.002	0.054	0.06%	0.65%
3.0 (3.000)	0.01 (0.010)	3.000	0.040	3.000	0.042	-0.03%	3.62%
2.9 (2.901)	0.10 (0.101)	2.930	0.082	2.929	0.083	-0.01%	1.74%
2.9 (2.899)	0.05 (0.053)	2.926	0.056	2.927	0.057	0.05%	0.43%
2.9 (2.900)	0.01 (0.010)	2.928	0.041	2.928	0.043	-0.03%	5.05%

All values in millimeters.

The comparison in Table 3-21 demonstrates the closed-form equations' ability to calculate the output mean and output standard deviation for a variety of given input means and input standard deviations. The closed-form solution for the mean matched the simulated results very closely with the maximum percent error less than 0.1%. The values for the output standard deviations were very close to the simulated results. All standard deviation values were within 6% of the corresponding simulated value.

3.7 Guidelines to Equation Construction and Usage

Several considerations are important when constructing and using closed-form equations for the mean and standard deviation exiting a process. In this section, several heuristic guidelines are presented that are helpful when deciding on the data to use to construct the equations, the amount of data required to construct accurate equations, and the range of possible inputs for which the equations are valid. Quantitatively deriving universal rules is impossible because of the infinite combination of possible processes, parameters, and input values; thus, these guidelines are based on heuristics. The detailed development of these guidelines is presented in Appendix H.

Guideline #1: Using more data to construct the closed-form equations produces more accurate results

Using as much data as possible to construct the equations eliminates the chance that a few outlier data points would cause misleading results. Also, using more data implies the range of data being used would be larger and thus produce better results. See Guideline #3.

Guideline #2: A minimum of n+1 sets of unique data are required to construct the equations, where n is the highest power appearing in the assumed form of DeltaP and SigmaP

This requirement exists to ensure the bare minimum of data are used to capture the shape of DeltaP and SigmaP. The uniqueness requirement is in place to guarantee that no two sets of data are identical, i.e., same values for the means and standard deviations. If two sets of data are identical, the second set of data does not add any new information that can be used to construct the equations. The minimum number of data sets required is needed because that is the number of coefficients that are being computed. For example, if DeltaP and SigmaP are assumed to be second order polynomials (n=2), then a minimum of three unique sets of data are required to find a_0, a_1, and a_2.

Guideline #3: The more spread present in the data used to construct the closed-form solutions, the better.

The technique presented in this thesis of modeling the mean and standard deviation exiting a manufacturing process is very good at "interpolating" values. In other words, the closed-form equations produce accurate results when the inputs to them fall within the range of values used to construct the equations. For example, if values for means spanning from 50 to 60 were used to construct the equations, then the equations would produce accurate results for an input mean of 55. Using even a few data sets that have a lot of spread to construct the closed-form equations produces equations that are more accurate over a wider range of input values than using a lot of data sets that are clustered around a single set of values.

Guideline #4: "Extrapolation" is possible with the closed-form equations.

The closed-form equations for the mean and standard deviation exiting a manufacturing process can accept input values that lie outside the range of values used to construct them and still compute accurate results. Estimating how far outside this range is difficult. The wider the spread of data used during construction, the further outside this range input values can be used, i.e., the use of highly spread data to construct the equations does a better job of capturing the true shape of the curves, thus produces better results when inputs are used that fall outside the original range. A rule of thumb is that the equations are valid for inputs that fall within the range of "construction data" and for one time the range of "construction data." For example, if values for means spanning from 50 to 60 were used to construct the equations, the equations would be valid for inputs ranging from 40 to 70. Of course, physical limitations do exist and should always be considered, e.g., if the above example were for a turning operation that creates an output diameter of 49, then it makes no physical sense to use any number less than 49 as an input. An upper physical limit also exists, e.g., maybe the lathe in the example physically cannot accept an incoming rod larger than 65, thus using inputs greater than 65 for the equations is fruitless.

Guideline #4: New data sets can be used to verify and "train" the equations.

The purpose of using the closed-form equations is to forecast the mean and standard deviation that will be produced with a particular process. This implies that a calculation will more than likely be followed by the making of parts. The parts can be measured and the data used to verify the equations forecasted accurate values. The data can also be used to "train" the equations in the sense that the new data can be combined with the original "construction data," and the closed-form equations can be re-constructed incorporating the newly measure data. The new equations will most likely produce more accurate results over a wider range of input values (see previous Guidelines).

3.8 Chapter Summary

The goals of this chapter were to create and validate a general method for constructing closed-form equations for the mean and standard deviation of a dimension exiting a manufacturing process, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{target})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{target})$. The challenges associated with accomplishing these goals along with possible solutions are summarized in this section. The solutions that are original and were introduced for the first time in this thesis are identified by the ❋**NEW!**❋ symbol. The section of the thesis where each solution is mentioned or discussed is listed in the summary to link to more detail. Each solution has its advantages and disadvantages listed to provide a quick comparison.

Challenge: **Understanding variation propagation through a manufacturing process**

Solution	Section	Advantages (+) & Disadvantages (-)
Process specific, physics-based modeling	n/a	+ Can provide insight into a process + Can be used to quantify the effects of changing operating parameters – Requires great expertise – Time consuming to build – Often done after process is in operation
Curve fit output data to input data	3.1.2	+ Simple approach + Requires no prior knowledge of the process – No physical significance to justify answers, i.e., data is blindly curve fit – Requires trustworthy process capability data – Difficult to know what terms to include in curve fit
❋NEW!❋ Use equations developed in this thesis, i.e., Equations 3-25 and 3-26	3.2 to 3.5	+ Simple to use + Backed by physical significance, i.e., the foundation is the theory of a process imparting a dimensional change and variation to components, which are both physical concepts (see Section 3.2) + Requires no prior knowledge of the process + Quickly construct equations for many processes to make comparisons + Generally applicable, i.e., this approach can be used to develop equations for many different types of manufacturing processes + Computationally inexpensive + Can be created early in the design process + Link equations together to represent serial processing (See Section 5.2.3) – Requires trustworthy process capability data – Cannot be used to optimize operating parameters – Closed-form solutions possible only if "incoming" dimensions are normally distributed

Challenge: **Defining a framework to represent the effects a manufacturing process has on component dimensions**

Solution	Section	Advantages (+) & Disadvantages (-)
❋NEW!❋ Theory of Desired Dimensional Change, Process Imparted Dimensional Change, Process Imparted Variation	3.2 3.2.2	+ Can be represented with equations or graphically + Easy to use to qualitatively compare processes' abilities to create quality parts + Building block for constructing closed-form solutions, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$ – Requires trustworthy process capability data to be effective

Challenge: **Generating closed-form equations**

Solution	Section	Advantages (+) & Disadvantages (-)
❋NEW!❋ Use the theory of Process Imparted Dimensional Change and Process Imparted Variation to generate closed form equations	3.5	+ Straightforward steps + Easy to implement in a wide variety of computer applications – Process capability data must be available

Challenge: **Calculating coefficients of functions that represent a manufacturing process**

Solution	Section	Advantages (+) & Disadvantages (-)
Find coefficients that minimize squared error	3.4	+ Simple to solve – Relies on trustworthy process capability data

Challenge: **Validating theory and resulting closed-form equations**

Solution	Section	Advantages (+) & Disadvantages (-)
Compare closed-form equation results to simulated data	3.6	+ Simulation models not difficult to construct + Multiple simulations can be run with little effort – Simulation model does not perfectly represent actual manufacturing process
Use actual measurements to construct and test closed-form equations	3.6.1s	+ Demonstrates theory works with real data + Very easy to perform required steps – Limited availability of actual measurements – Measurement error could be present

A production line is comprised of manufacturing processes. Understanding how these processes affect the quality of the parts they produce is very important. This chapter presented the development of the mathematical equations to forecast the output mean and variation of a distribution produced by a particular process as a function of the incoming distribution's mean and standard deviation. As a part of this development, a theory was introduced stating that a manufacturing process with a specified set of operating parameters can be represented by two math functions, DeltaP and SigmaP. These functions are used to construct the desired closed-form equations used to calculate the output mean and standard deviation. Three physics based math models were used to validate the theory. The three processes used were turning, rolling, and sheet bending. The theory of DeltaP and SigmaP performed very well in forecasting the quality of a component or assembly dimension upon exiting the appropriate manufacturing operation.

A pictorial overview of the μ_{out} and σ_{out} closed-form equation derivation is shown in Figure 3-32. The first step of the derivation is the binning or dividing of the incoming distribution. The bins are then treated as individual distributions, and each bin is "processed" to compute the mean and standard deviation it has upon exiting the process. The individual exiting distributions are then combined to produce the final distribution of all the parts. An overview of the derivation was presented in this chapter, a complete derivation is presented in Appendix G.

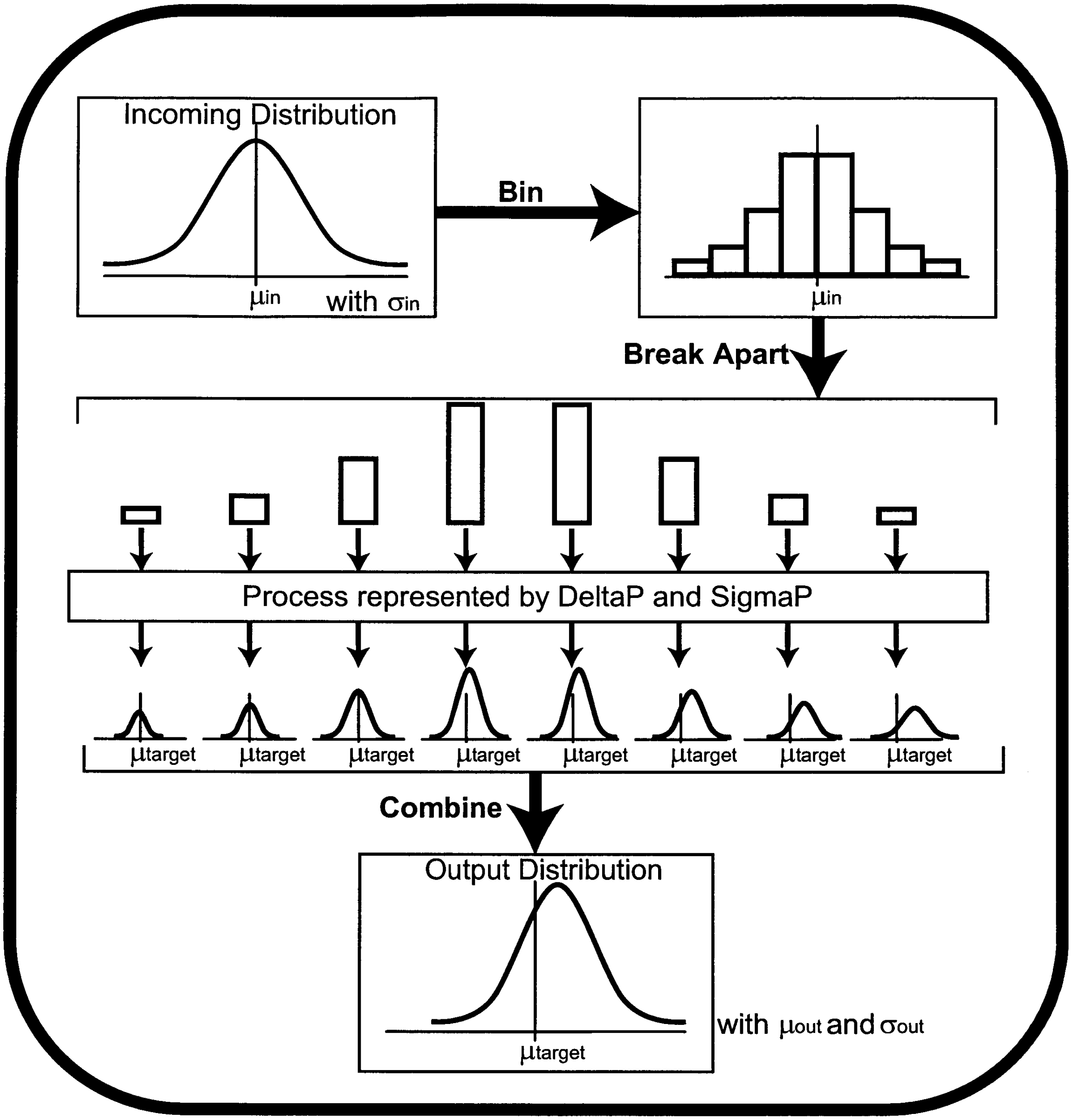

Figure 3-32: Pictorial summary of mathematical derivation

4 Modeling of Selective Assembly

Chapter Highlights

- Define and describe multiple selective assembly techniques
- Introduce two new techniques
- Develop algorithm to simulate selective assembly operations
- Produce closed-form equation for selective assembly standard deviation
- Produce closed-form equation for selective assembly scrap
- Present example problems of non-normal distributions
- Validate selective assembly closed-form equations using Monte Carlo simulation

4.1 Introduction

Most products are comprised of more than one component. These products require some type of assembly during production. Assembly is defined as the putting together of manufactured parts to make a completed product. An assembly process can be the joining of two or more parts, the fitting of one part to another, the fastening of two mating parts, or a number of other scenarios. The purpose of this chapter is to explore how assembly operations, both random assembly and selective assembly, affect the final mean and variation of a product's dimensions and functional characteristics such as weight, resistance, etc. Math models are created to be combined with the math functions for manufacturing processes developed in Chapter 0 to forecast the mean and standard deviation of a dimension as it exits a production system.

Selective assembly is the act of measuring components and selecting two mating components based on those measurements to create an assembly. A review of existing popular techniques for implementing selective assembly is presented in this chapter. Two new approaches that minimize scrap are introduced and discussed. A very general algorithm is introduced that quantifies the effect selective assembly has on the cost and quality of an assembly. This algorithm is used to construct closed-form equations for the variation and scrap produced by selective assembly when two normal distributions are combined. Engineers can use the algorithm and resulting closed-form equations to help decide when and what type of selective assembly to implement. For example, the equations can be used to compare the final variation a dimension will have if the product is assembled using selective assembly so that an engineer can decide if the reduced variation is worth the added cost.

4.1.1 Overview

The main focus of this chapter is to demystify selective assembly so that it can be effectively and intelligently implemented to manage product variation.

Chapter Goals

- Discuss the advantages and disadvantages of implementing selective assembly
- Define new selective assembly techniques for use when component distribution variances are unequal
- Develop criteria and analysis approach to use when choosing a selective assembly technique
- Create closed-form equations to calculate the standard deviation of an assembly that is created by a selective assembly operation
- Explore the robustness of selective assembly to changes in incoming component variances

4.1.2 Background

A general case of assembly is the joining of any number of parts. For the discussion in this thesis, the example used will be the joining of an "A" part with a "B" part to form a "C" assembly. This simplification does not cause any loss in generality because it could be repeated multiple times to represent the joining of many parts. A picture of this mental model is shown below in Figure 4-1 where a Part "A" is joined with a Part "B" to create an Assembly "C."

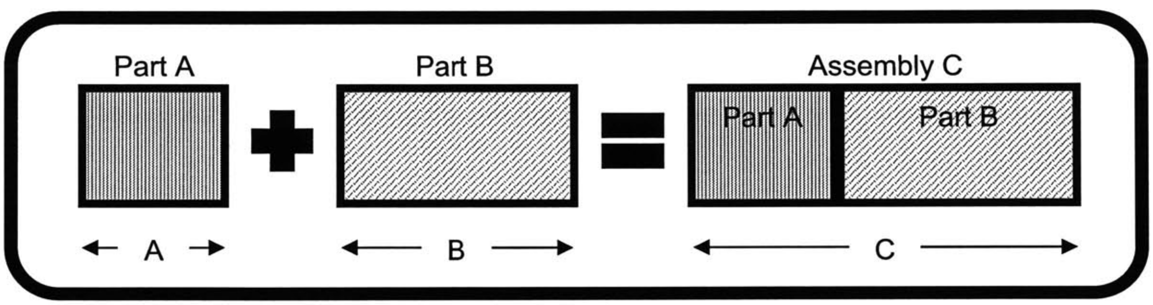

Figure 4-1: Mental model of assembly operation

Random assembly is the most common type of assembly operation. In random assembly, an "A" part is randomly chosen from a population of "A" parts, and a "B" part is randomly chosen from a population of "B" parts. These two chosen parts are then assembled together to form a "C" assembly.

The idea of random assembly is extended to statistical populations. The equations for the mean and standard deviation of Assembly "C" are generated very easily based on simple statistical equations. The mean of Assembly "C," μ_C, is simply the sum of the means of Part "A" and Part "B." The standard deviation of Assembly "C," σ_C, can also be computed easily by recognizing the variances of Part "A" and Part "B" can be summed to produce the variance of Assembly "C." The equations for the mean and standard deviation of Assembly "C" are shown below.

$$\mu_C = \mu_A + \mu_B \tag{4-1}$$

$$\sigma_C = \sqrt{\sigma_A^2 + \sigma_B^2} \tag{4-2}$$

Two common methods for setting tolerances exist for assembly dimensions created by a random assembly operation. The first method is usually referred to as the worst case stack-up method. With this method, the tolerances for Part "A" and Part "B" are set in such a way that a Part "A" at its upper tolerance limit mated with a Part "B" at its upper tolerance limit will produce an Assembly "C" that still falls within its tolerance limit, usually equal to its upper tolerance limit. This method guarantees that as long as Part "A" and Part "B" fall within their respective tolerance bands, then Assembly "C" will fall within its tolerance band. Setting tolerances using the worst case stack-up method is sometimes overly cautious and can lead to higher manufacturing costs to achieve the tight tolerance for the components. The other common method for setting tolerances is based on statistical principles and is known as the root sum squared method (RSS) (Liggett 1993). This method assumes the component distributions are normal and uses the idea that it is unlikely for a very large Part "A" to be randomly mated to a very large Part "B." The tolerance range of the assembly is set equal to the square root of the sum of the squares of the component tolerances. This relationship is based on the sum of the variances as seen in Equation 4-2.

A company's ability to manufacture quality products determines its success. Manufacturing products that have a high level of quality can be costly and difficult. The use of selective assembly may be a valuable addition to a manufacturing system to improve the quality of a product in a cost effective manner. Selective assembly is the process of measuring parts in each distribution and then purposefully selecting a Part "A" and a Part "B" that when mated together will produce an Assembly "C" that is close to the desired assembly dimension. There are several specific techniques for implementing selective assembly. Each approach has specific advantages and disadvantages associate with it. Selective assembly is often implemented when the cost of part precision is greater than the cost of implementing selective assembly. For example, low precision parts that are inexpensive to produce can be combined using selective assembly to produce high precision assemblies instead of spending large amounts of money on making high precision parts to assemble with random assembly.

The general steps for performing a selective assembly operation are as follows. First, each Part "A" is measured. Depending on the measured value, the part is placed in a bin with other parts that have similar measured values. A bin can mean a figurative grouping or a literal bin or bucket that contains parts with close dimensions. Secondly, each Part "B" is measured and placed in an appropriate bin. Finally, a Part "A" is randomly selected from any bin and a random Part "B" is selected from the corresponding bin. These two parts are then assembled together.

A graphical representation of selective assembly is shown below in Figure 4-2 where a slightly oversized Part "A" is picked and mated with a slightly undersized Part "B" to produce an overall assembly dimension close to the desired value. The distributions are numbered so that corresponding bins have the same identification number. Notice the bin numbers for Part "A" increase from left to right while the bin numbers decrease for Part "B." This is true when the desired assembly dimension is an additive dimension, e.g., an overall length. In this case, combining a small part from one distribution with a large part from the other distribution produces an assembly that will be close to the target value. If the dimension of interest on the assembly were a clearance, both distributions would be binned with increasing identification numbers from left to right, i.e., a small rod would be mated with a small bore to produce a clearance close to the desired value.

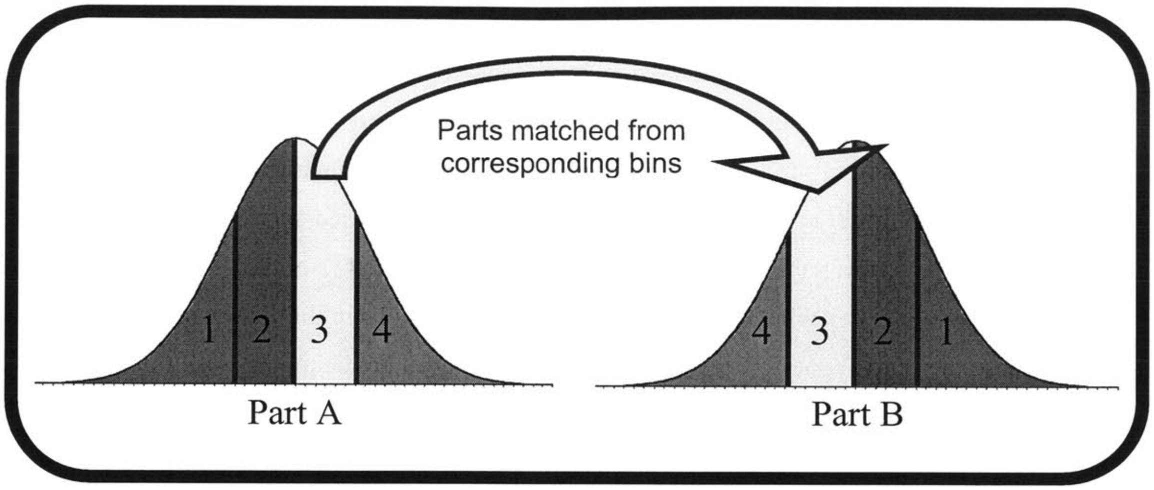

Figure 4-2: Graphical representation of selective assembly

Selective shimming is a subset of selective assembly. Selective shimming, or just shimming, is the process of using a very inexpensive part to mate with a more expensive part or sub-assembly to result in an assembly that is close to the desired value. Shimming is different from selective assembly because one of the two mating parts or sub-assemblies is much less expensive than the other. This situation gives rise to some unique decision making opportunities and is discussed later.

Selective assembly has several advantages over random assembly. The primary advantage is the assembly variation is reduced and can be controlled to a tighter tolerance. Another advantage is the fact the assembly scrap rate can be lowered substantially. The higher quality, lower variation, also reduces the rework required to get assemblies to meet specification. Another advantage of selective assembly is the fact the production line can implement statistical process control for almost free since selective assembly requires every part to be measured.

Along with the advantages, selective assembly also introduces several disadvantages. The obvious disadvantage is the requirement to measure every part. These measurements more than likely add cost and can increase production time. Inventory costs may also increase when selective assembly is implemented since more inventory of one part may be required to ensure a match is available. Part scrap cost can also rise when selective assembly is used to mate two parts with very dissimilar variances because a large number of the parts with the larger variance may have to be scrapped.

Another disadvantage to using selective assembly is a loss in serviceability of the product. This occurs when a part that was originally assembled into the product using selective assembly needs to be replaced in the field. Not just any service part will fit. A part from the appropriate range or bin is required. To address this issue, multiple service parts may have to be made and stocked, each from a different bin or range. The downside to this is increased service part inventory and additional unique part numbers are required and must be tracked. Another way to address the serviceability issue is to replace the entire assembly when failure in the field occurs. Again, this may increase the service cost since some parts that did not fail are being replaced (along with the failed part). In the case of selective shimming, a service kit may be made available containing multiple inexpensive shim parts, one form each bin. This kit allows the appropriate part to be selected and used in the repair while the remaining parts in the kit are discarded. If the shim part is inexpensive enough, this may be the most cost effective solution to the serviceability issue.

Research in the area of selective assembly has been focused on several key aspects of the problem. Pugh (Pugh 1986) discusses the use of equal width binning and introduces a computer program that computes the output mean and standard deviation of an assembly of two parts that are each normally distributed. The queuing method was introduced by Boyer and Nazemetz (Boyer and Nazemetz 1985) to measure a small number of parts and queue them, then measure a mating part and match it with the best fit from the queue. A strategy for optimally binning two distributions was introduced by Mease and Sudjianto (Mease, Nair et al. 2002). They presented the statistical framework for calculating the bin sizes that will minimize several loss functions. Their work also showed that optimal results from selective assembly can come from combining two distributions that have unequal variances, an idea that was previously thought to be untrue.

A limited amount of literature does exist on the subject of implementing selective assembly when the two distributions being combined have unequal variances. Pugh (Pugh 1992) establishes the concept of truncating the wider distribution to make the variance match more closely with the tight distribution. His analysis only looks at the number of assemblies that meet the specification without regard to the level of variation present. Chann and Linn (Chan and Linn 1998) propose a grouping method based on the cumulative distribution functions of the mating parts. Their goal was to ensure mated parts meet the assembly specification while minimizing the number of scrap. This was accomplished by using bins that have equal probability. From a quality perspective, both of these approaches were concerned only with meeting specification and did not consider the effect on the assembly's variation. Some selective assembly research deals with the availability of mating parts in the two distributions and are concerned with making sure a match exists for every part (Fang and Zhang 1996; Zhang and Fang 1999). There is limited literature on making selective assembly decisions based on economic considerations (Kwon, Kim et al. 1999). The existing literature assumes the distributions to be normal and have equal variance. These two assumptions limit the effective application of selective assembly.

Each of the above approaches focuses on one aspect of the selective assembly problem. None of the work takes a general approach to selective assembly that is applicable when the distributions' variations are different. Also, the existing research usually focuses on "goal post" analysis, i.e., they are only concerned if the assembly meets its specification or not. This focus can cause possible quality improvements to be overlooked. Another lacking item in the literature are closed-form equations describing the variation of an assembly created with selective assembly.

Historically, selective assembly has been used and implemented in an ad hoc manner. One goal of this chapter is to develop quantitative models that can assist in the decision to implement selective assembly into a manufacturing system. These quantitative models focus on the quality and cost impact selective assembly can have on a product. The models are created to be as general as possible so that they can be applied to a variety of situations, e.g., the assumption of distributions with equal variances is not made.

4.2 Selective Assembly Techniques

There are several techniques to performing selective assembly. Each technique has its own subset of advantages and disadvantages. The most common techniques for selective assembly are discussed in this section.

Since the discussion has been limited to the assembly of two parts, Part "A" and Part "B," there are two distributions under consideration. Often, in selective assembly work, the assumption is made that the two distributions have equal variances. That assumption is not made in this thesis because a general approach to selecting and implementing a selective assembly technique is the goal. The distribution that has the larger variance will be referred to as the "wide" distribution while the distribution with the smaller variance will be referred to as the "tight" distribution. For mathematical reasons, the distributions are truncated at plus or minus three standard deviations to give them finite widths. Also, as in previous chapters, the Δ symbol represents the width of a bin and N represents the number of bins by which a distribution is divided.

4.2.1 Equal Width Binning to Minimize Variation

The *Equal Width Binning to Minimize Variation* technique relies on both distributions being separated into bins that have equal widths. This technique is historically referred to as simply the Equal Width Binning technique. The new title for this technique is being introduced to differentiate it from a new technique, Equal Width Binning to Minimize Scrap, which is presented later. The Equal Width Binning to Minimize Variation and the Equal Width Binning to Minimize Scrap are the same technique when the two distributions being combined are normally distributed with equal variances.

In this technique, the bin width is the same for both distributions. This is expressed mathematically in Equation 4-3 below and shown below in Figure 4-3 where the width of every bin is identical and the areas of parts scrapped from the wide distribution are shown in black.

$$\Delta = \Delta_{wide} = \Delta_{tight} \tag{4-3}$$

The width of the bin is determined by the tight distribution. The tight distribution is truncated at plus or minus three standard deviations. The range, equal to six standard deviations, is then divided by the desired number of bins, *N*, to produce the bin width.

$$\Delta = \frac{6\sigma_{tight}}{N} \tag{4-4}$$

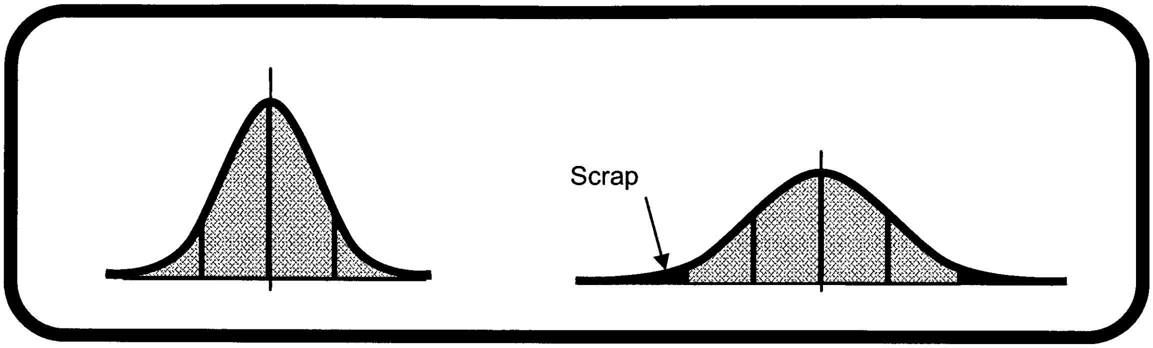

Figure 4-3: Equal Width Binning to Minimize Variation

Several advantages are associated with this technique. The resulting assembly variance is greatly reduced. Also, the tolerance limit on the assembly can be clearly set and meeting this specification can be guaranteed for combinations from any set of corresponding bins. For example, Part "A" has a mean of 100 mm and a standard deviation of 3 mm, and Part "B" has a mean of 50 mm and a standard deviation of 6 mm. If the number of bins is six, then the bin width is 1 mm. This creates a bin for Part "A" from 100 mm to 101 mm, and a mating bin for Part "B" from 49 mm to 50 mm. Any assembly generated by combining parts from these two bins, or any two corresponding bins, is guaranteed to fall between 149 mm and 151 mm. This fact is independent of the variance of Part "B." If the bin width for Part "B" was influenced by the variance of Part "B," the guarantee of a tolerance could not be made. Another advantage to this technique is its simplicity. The bin width calculations are very straightforward. This simplicity adds to the flexibility of this technique. For example, if the characteristics of an incoming distribution change, the bin widths can be quickly and easily adjusted.

There are also disadvantages associated with this technique. The biggest disadvantage is the amount of scrap that can result from truncating the wide distribution. This becomes worse when the variances of the two distributions are very different. The other major disadvantage is the improvement in quality may not be as great as competing techniques. Later sections quantify the improvements for each technique.

This technique lends itself to the selective shimming type of operation mentioned earlier. This is due to the fact the wider distribution parts would have to be very inexpensive to make this a feasible technique. Again, the models presented later quantify these concerns.

4.2.2 Equal Width Binning to Minimize Scrap

The *Equal Width Binning to Minimize Scrap* technique relies on a distribution being separated into bins that have equal widths. This is a new selective assembly technique and is being presented for the first time in this thesis. In the past, selective assembly has been mostly reserved for cases where the two distributions being combined have equal variance. This assumption is not being made in this thesis; therefore, introducing this new technique is necessary.

The bin width for this technique is unique to each distribution, i.e., the bin width for Part "A" is different than the bin width for Part "B" when the variances are unequal. The bin widths can be computed using the following equations assuming each distribution is truncated at its respective plus and minus three standard deviations. A sample of two binned distributions with the number of bins, N, equal to 4 is shown in Figure 4-4.

$$\Delta_{tight} = \frac{6\sigma_{tight}}{N} \tag{4-5}$$

$$\Delta_{wide} = \frac{6\sigma_{wide}}{N} \tag{4-6}$$

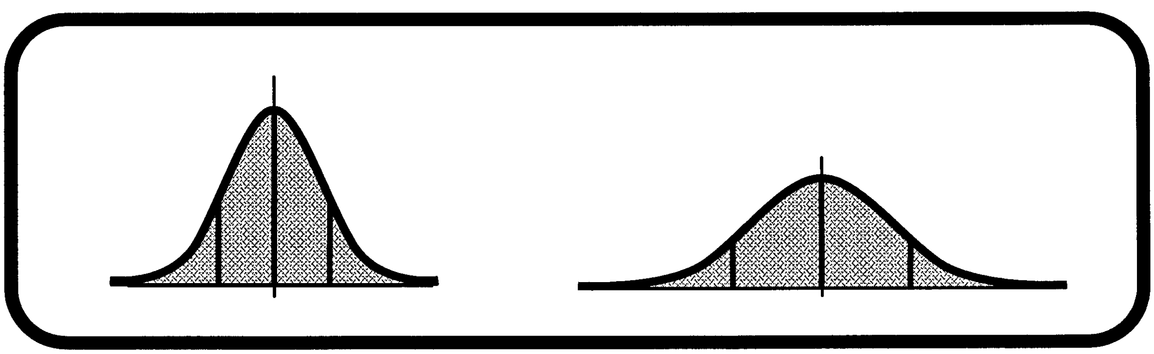

Figure 4-4: Equal Width Binning to Minimize Scrap

The biggest advantage to this technique is that more of the parts from the wider distribution can be used since it is not truncated to match the tighter distribution. The other advantage is its simplicity. The bin widths for the two distributions are independent and can be easily recalculated when the variance of one distribution changes.

The biggest disadvantage to implementing this technique is the range of assembly dimensions is dependent on which set of corresponding bins is used. For example, Part "A" has a mean of 100 mm and a standard deviation of 3 mm, and Part "B" has a mean of 50 mm and a standard deviation of 6 mm. If the number of bins is six, then the bin width for Part "A" is 1 mm and is 2 mm for Part "B." This leads to bins for Part "A" from 100 mm to 101 mm and from 101 mm to 102mm. The Part "B" corresponding bins range from 48 mm to 50 mm and from 46 mm to 48 mm respectively. Therefore, the first set of corresponding bins can produce assemblies from 148 mm to 151 mm, and the second set of bins can produce assemblies from 147 mm to 150 mm. Therefore, the entire range of possible assembly values is determined by bins at the extremities. The range of possible values is also larger than for the previous technique of Equal Width Binning to Minimize Variation.

This technique may be viable when the overall dimension of the assembly is not as critical as the variation of the assemblies. This is true because this technique does not limit the values the assembly can have as much as other techniques, but this technique does reduce the variation of the assembly dimension.

4.2.3 Equal Area Binning to Minimize Variation

The *Equal Area Binning to Minimize Variation* technique relies on both distributions being separated into bins that contain equal numbers of parts. This technique is historically referred to as simply the Equal Area Binning technique. The new title for this technique is being introduced to differentiate it from a new technique, Equal Area Binning to Minimize Scrap, which is presented later.

For this technique, the width of each bin is set so that a certain percentage of area under the probability density function is contained within each bin. The bins that are closer to the mean are narrower than those at the tails because the distribution is denser in the middle.

The tight distribution determines the limits of the bins in this technique. The wide distribution is truncated and then divided into bins according to the widths specified by dividing the tight distribution. The bin widths are the same for both distributions, i.e., the i^{th} bin of Part "A" is the same width as the corresponding bin for Part "B." A picture of two distributions binned using this technique is shown below in Figure 4-5. The areas that represent scrapped parts are shown in black. Each bin width is labeled to make it easier in identifying the bins that have the same width. This is expressed mathematically in Equation 4-7 below.

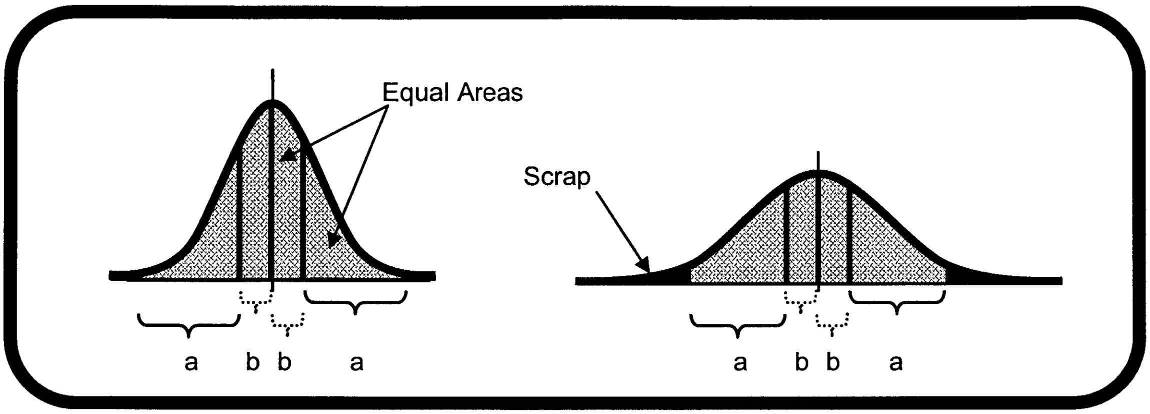

Figure 4-5: Equal Area Binning to Minimize Variation

$$\Delta_i = \Delta_{wide,i} = \Delta_{tight,i} \qquad \text{4-7}$$

The widths of the bins are determined by the tight distribution. The tight distribution is truncated at plus or minus three standard deviations. The area under the remaining truncated curve is then divided by the desired number of bins, *N*, according to the following equation where *x* is a measurement within the tight distribution.

$$\frac{1}{N} = \int_{x_{i-1}}^{x_i} \left(pdf_{tight} \right) dx \qquad \text{4-8}$$

For a normal distribution, Equation 4-8 can be rewritten as

$$\frac{1}{N} = \int_{x_{i-1}}^{x_i} \left(\frac{1}{\sigma_{tight}\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu_{tight}}{\sigma_{tight}} \right)^2 \right\} \right) dx \qquad \text{4-9}$$

The integral limits are the variables in the above equation and are set in a manner to produce symmetrical bins that contain the appropriate amount of area, which is fixed by the number of bins. The bin widths are then computed for each set of integral limits per the following equation.

$$\Delta_i = x_i - x_{i-1} \quad \text{4-10}$$

The bin limits are then set for the wide distribution to produce bins that have the same widths as their corresponding bins in the tight distribution. In the case of distributions with unequal variances, the part with the wider distribution does not have the same area in a bin as the corresponding bin of the tight distribution because of heavier truncation. If the more expensive part comprises the tight distribution, then the wide distribution may be overproduced to ensure a match. This would be the case for a shimming operation.

The main advantage associated with this technique is the resulting assembly variance is greatly reduced when compared to random assembly; however, there are disadvantages associated with this technique. The biggest disadvantage is the amount of scrap that can result from truncating the wide distribution. This becomes worse when the variances of the two distributions are very different. Another disadvantage is that the possible range of assembly dimensions is greater than the Equal Width Binning methods. This is due to the fact the bins near the tails are much wider than those near the middle. When parts are combined from two corresponding bins near the tails, the range of possible values of the assembled dimension can be large.

This technique lends itself to the selective shimming type of operation mentioned earlier. This is due to the fact the wider distribution parts would have to be very inexpensive to make this a feasible technique. The models presented later quantify the effect this technique has on assembly variance and part scrap.

4.2.4 Equal Area Binning to Minimize Scrap

The *Equal Area Binning to Minimize Scrap* technique relies on both distributions being separated into bins that contain equal numbers of parts. This is a new selective assembly technique and is being presented for the first time in this thesis. It is being introduced to allow engineers to have the option to keep scrap low while reducing the standard deviation of an assembly.

In this technique, the width of each bin is set so that a certain percentage of area under the probability density function is contained within the bin. In the case of a normal distribution, the bins closer to the mean are narrower than those at the tails because the distribution is denser in the middle.

The two distributions are binned independently. The bins span the entire range of the distributions; therefore, the wider distribution is not as truncated as in the *Equal Area Binning to Minimize Variation* technique. This implies that the wide distribution part does not need to be overproduced and the scrap rate is not as severe.

The widths of the bins are determined by each distribution. The distributions are truncated at plus or minus three standard deviations. The areas under the remaining truncated curves are then divided by the desired number of bins, N, according to the following equations where x is a measurement within the tight distribution and y is a measurement within the wide distribution. For normal distributions, the equations are shown below.

$$\frac{1}{N} = \int_{x_{i-1}}^{x_i} \left(\frac{1}{\sigma_{tight}\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu_{tight}}{\sigma_{tight}} \right)^2 \right\} \right) dx \qquad \textbf{4-11}$$

$$\frac{1}{N} = \int_{y_{i-1}}^{y_i} \left(\frac{1}{\sigma_{wide}\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{y-\mu_{wide}}{\sigma_{wide}} \right)^2 \right\} \right) dx \qquad \textbf{4-12}$$

The integral limits are the variables in the above equations and are set in a manner to produce symmetrical bins that contain the appropriate amount of area, which is fixed by the number of bins. The bin widths are then computed for each set of integral limits per the following equations. Figure 4-6 shows two sample distributions binned using this technique. The bins with equal area are labeled in the picture. The bins identified as "Not Necessarily Equal Areas" are only equal when the two distributions have the same variance.

$$\Delta_{tight,i} = x_i - x_{i-1} \qquad \textbf{4-13}$$

$$\Delta_{wide,i} = y_i - y_{i-1} \qquad \textbf{4-14}$$

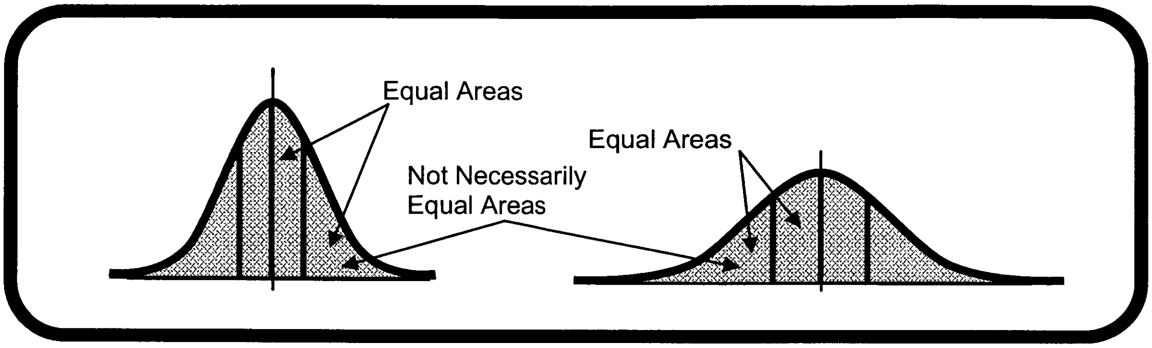

Figure 4-6: Equal Area Binning to Minimize Scrap

The main advantage associated with this technique is the wide distribution is not heavily truncated; making the scrap rate lower.

There are disadvantages associated with this technique. The biggest disadvantage is that the possible range of assembly dimensions is greater than most of the other techniques. This is due to the fact the bins near the tails are much wider than those near the middle, especially in the wide distribution. When parts are combined from two corresponding bins near the tails, the range of possible values of the assembled dimension can be large.

4.2.5 Optimal Width Binning

Optimal width binning is the concept of determining the bin widths of the two distributions so that the assembly of them will minimize some objective function. The optimal width binning technique used in this thesis is based on the work of Mease and Sudjianto (Mease, Nair et al. 2002) where the objective function being minimized by the selection of the bin widths is the squared error loss function.

4.2.6 Selective Machining

Selective machining is the act of measuring a component and then machining it or its mate to produce the overall desired assembly dimension. This technique has the advantage of producing assemblies with very low variance and little to no scrap. However, there are several disadvantages. This technique can add cost to the process by requiring high precision machining to achieve the desired level of variation. This negates the major positive effect of selective assembly, i.e., obtaining high precision assemblies with low precision machining processes. The other major drawback to this technique is that it may add time to the assembly process because of the measure and machining loop. The selective machining is usually only employed in low volume production operations such as a job shop or an aircraft assembly line.

4.2.7 Queuing Method

The queuing technique for selective assembly involves measuring a certain number of Part "A" and a certain number of Part "B," placing the measured parts into queues, and then selecting the parts from the queues that would create an assembly with a dimension closest to the desired value. This technique to selective assembly was introduced by Boyer and Nazametz (Boyer and Nazemetz 1985).

The main advantage of this technique is that it does not require the binning of parts. This technique can be implemented in-line as part of production flow. The other advantage is that there is no scrap as long as components meet their respective specifications.

There are a few disadvantages to this technique. The first is the fact the assembly variation is highly dependent on the ordering of the components. For example, if there is low part to part variation for both assemblies and the components are not from "corresponding tails," this technique may produce assembly dimensions that are consistent but not close to the desired value. This is related to another disadvantage; that is the inability to take advantage of the entire component distributions simultaneously. For example, in the binning techniques, when a small Part "A" is selected, a large Part "B" can be selected. Because of the limited, and usually low, number of parts in the queue, this technique can not achieve the same flexibility. Another downside to this technique is the need for special equipment to keep track of the parts in the queues. This equipment can add cost that the binning techniques do not have.

The queuing method could be successfully implemented in assembly operations that are automated. Pick-and-place machines or other types of robotic assembly lend themselves to this type of assembly because the computer controller could also keep track of the queued location of the parts.

4.3 Selective Assembly Considerations

Two major factors to consider when deciding whether or not to use selective assembly are quality and cost. This section covers details about both considerations.

4.3.1 Quality Considerations

The main reason for implementing selective assembly in a manufacturing system is to produce high precision assemblies from low precision parts. Ironically, a "goal post" mentality is often taken when selective assembly is utilized. A "goal post" mentality is one that only considers whether an assembly is within its specified tolerance limits. If the assembly dimension is within the limits, it is good; otherwise it is rejected. This type of characterization is binary, and the quality of the assembly is measured strictly as a percentage of assemblies that fall within the tolerance limits.

A better quality metric to use is the amount of variation present in the assembly dimension. This variation is quantified as the standard deviation of the assembly dimension's distribution. Using variation as a measure for quality allows selective assembly to be chosen based on reducing variation and the quality loss associated with that variation (Phadke 1989), instead of a means to just improve the yield of a production system.

Considering the variation of an assembly has several positive implications on a design. The impact of design changes can be easily quantified, thus leading to better designs. Tolerance allocation can be understood better when selective assembly's impact is known, leading to the opening of more expensive tolerances to reduce machining costs. The quality loss associated with varying from the desired value can be easily computed and used to make design or manufacturing modifications.

The models and algorithms presented later in this chapter quantify the variation an assembly will have based on the components' distributions. The binning techniques described earlier in the chapter are used. The result is the ability to calculate the variation of an assembly as a function of the specific technique, the number of bins utilized, and the part distributions. These results assist an engineer in making decisions regarding when to implement selective assembly and what technique to use in a given circumstance.

4.3.2 Economic Considerations

Along with the impact on quality, implementing selective assembly may have economic consequences. The decision to implement selective assembly is usually a result of not being capable of meeting an assembly specification with low precision components. Choosing to use selective assembly comes naturally when the cost of improving the components' precision is deemed too high compared to the costs associated with selective assembly. The difficulty arises when the true costs of selective are elusive, making it challenging to justify the use of selective assembly.

Selective assembly does have an advantageous effect on assembly costs. With selective assembly, the assembly yield should increase, thus decreasing the number of assembly rejects. This improvement also alleviates some, if not all, of the rework costs associated with fixing assemblies that do not meet spec.

Selective assembly can have a negative impact on the economics of a manufacturing system. The most obvious is the added cost of measuring every component. This cost is independent of the technique chosen and is a per piece cost. Selective assembly can increase the amount of component scrap. This can especially be severe when combining two distributions of dissimilar variance to minimize the assembly variation. The scrap cost is dependent on the selective assembly technique used, the components' distributions, and the number of bins used in the selective assembly operation. Another cost related to the scrap is the cost of inventory. This cost accumulates as a result of having to have more components from the wide distribution on hand to ensure a mating part exists at all times. The cost of selective assembly is given by Equation 4-15 and illustrated in Figure 4-7 where the part costs are identified as increasing the total cost and the assembly costs (savings) are shown as reducing the total cost.

$$C_{Total} = -C_{AsmScrap} - C_{Asm\,\mathrm{Re}work} - C_{Asm\,\mathrm{Re}pair} + C_{ComponentInspection} + C_{Inventory} + C_{PartScrap} \qquad \text{4-15}$$

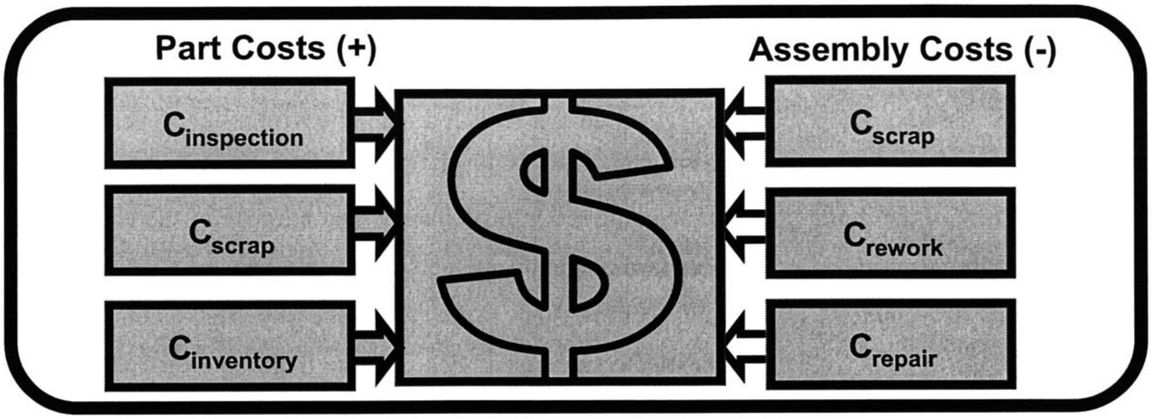

Figure 4-7: Graphical illustration of the cost of selective assembly

The costs associated with the assembly in Equation 4-15 are proportional to the assembly variation expected when random assembly is used and the tolerance range of the assembly dimension. This is expressed mathematically as

$$C_{AsmScrap} + C_{AsmRework} + C_{AsmRepair} \propto \sqrt{\sigma^2_{PartA} + \sigma^2_{PartB}} \qquad 4\text{-}16$$

The inspection cost of the components, $C_{ComponentInspection}$, is simply the fixed, per piece cost to make the required measurement. The inventory cost, $C_{Inventory}$, is the cost of having to keep extra inventory on hand to ensure a part match, e.g., in the case of selective shimming. The inventory cost is affected by a company's policy for internal charges and holding costs.

The cost in Equation 4-15 that is most affected by the choice of a particular selective assembly technique and number of bins is the part scrap cost, $C_{PartScrap}$. The models and algorithms presented in the next section quantify the effect a selective assembly operation has on the fraction of parts scrapped. The binning techniques described earlier in the chapter are used. The models calculate the scrap associated with using selective assembly as a function of the specific technique used, the number of bins utilized, and the part distributions. These results assist an engineer in making decisions regarding when to implement selective assembly and what technique to use in a given circumstance.

4.4 Selective Assembly Calculations

A selective assembly algorithm is presented here to compute the variation of an assembly dimension when two component distributions are combined. The algorithm is general and can be applied to any of the binning techniques outlined in this chapter. The purpose of the algorithm is to compute the resulting assembly standard deviation and percent scrap when any two component distributions are entered. This allows comparisons to be made and the best technique for a particular case to be selected. The best selective assembly technique is the one that produces an acceptable level of variation at the lowest cost.

The algorithm is implemented using numerical methods and Monte Carlo simulation because no analytically derived closed-form solutions exist. The first step is to generate a large number of random variables for Part A and Part B. In the case of the work presented here, one thousand random variables were generated for each part. The random variables are then binned according to a specific selective assembly technique for a given number of bins, ranging from one to the specified maximum number of bins allowed. Next, a value from a Part "A" bin is randomly chosen and combined with a Part "B" value randomly chosen from the corresponding bin to create a value for Assembly "C." A value left in a bin that has no corresponding match is added to the scrap count. Finally, the mean and standard deviation are calculated for Assembly "C." This algorithm is depicted graphically below in Figure 4-8.

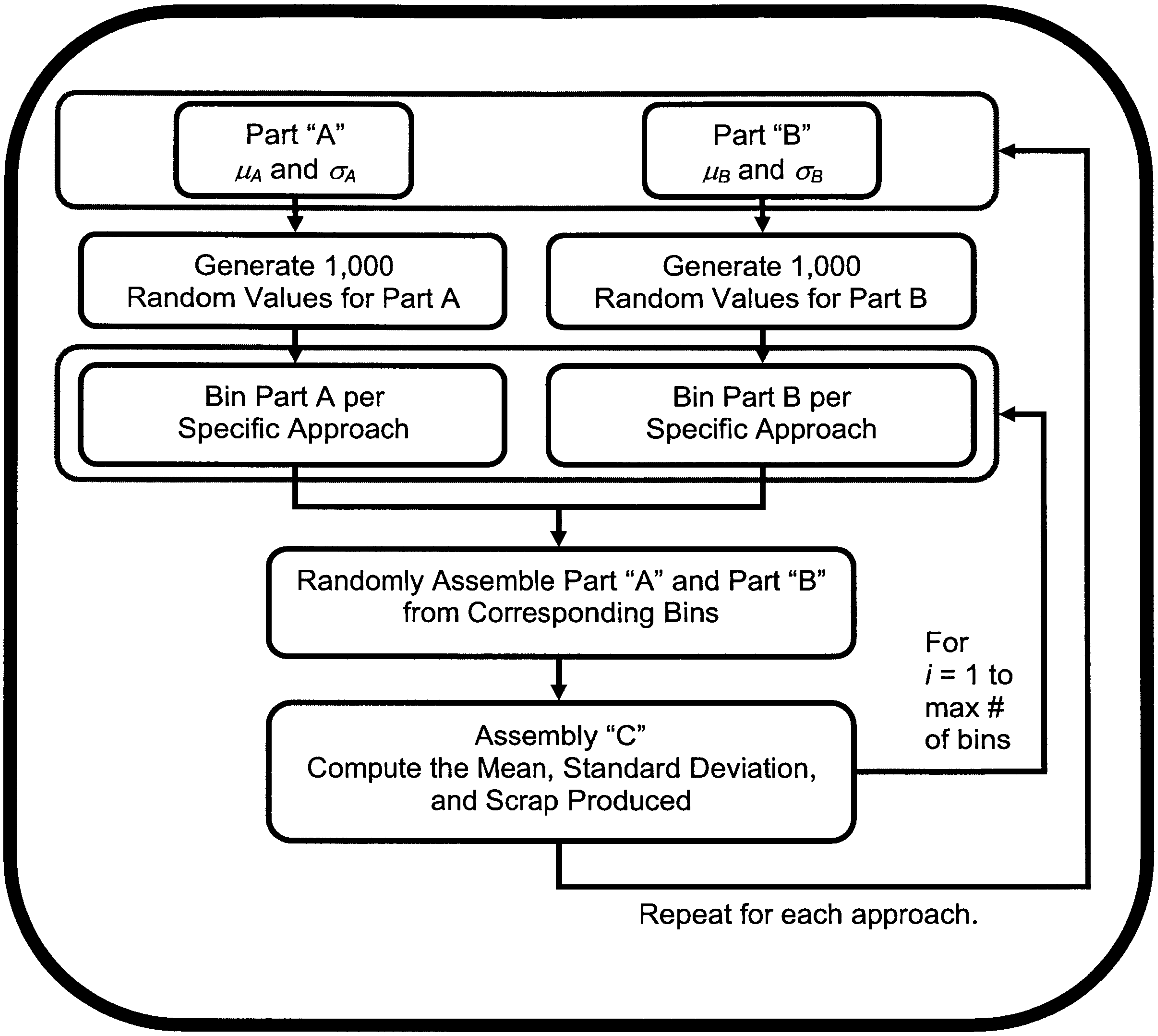

Figure 4-8: Graphical depiction of the selective assembly algorithm

Since the algorithm relies only on numerical Monte Carlo simulations, any type of distribution or combination of distributions can be evaluated; however, there are practical considerations to be made before blindly applying the selective assembly techniques discussed above to any pair of general distributions. This is the topic of discussion in Sections 4.6.2 and 4.6.3. In the next section, the algorithm is used to simulate data for the creation of closed-form equations to describe the impact selective assembly has on an assembly's variation and the amount of part scrap produced.

4.4.1 Closed-form Equations

Practical closed-form equations for computing an assembly dimension's variance when it is created using selective assembly and the percent of scrapped parts would alleviate the need to run computationally intensive Monte Carlo simulations. The desired closed-form equations would accept the standard deviations of the parts being combined and compute the standard deviation of the resulting assembly and the percent scrap. General forms of the desired equations are seen below in Equations 4-17 and 4-22 for the assembly's standard deviation and percent scrap respectively.

$$\sigma_{assembly} = f\left(\sigma_{PartA}, \sigma_{PartB}\right) \tag{4-17}$$

Deriving useful forms of these equations, so that they can be used when two normal distributions are being combined, from statistical principles is not possible because the integrals involved cannot be solved analytically. Therefore, numerical integration would be required. This approach is no better than using Monte Carlo simulation.

Another approach to creating these equations would be to run Monte Carlo simulation data and curve fit math functions to the data. The major drawback to this is the infinite number of combinations of the part distributions' standard deviations. The simulation would be required of every possible combination; thus making this approach infeasible.

A feasible approach to deriving a form of Equation 4-17 is to create an equation for a normalized assembly standard deviation. The normalized standard deviation is simply the selective assembly created standard deviation divided by the standard deviation that would be expected if random assembly were used. This normalized assembly standard deviation is then treated as a function of the ratio of the two part distributions' standard deviations. Taking this approach and using the notation introduced in Figure 4-1 the dimensionless relationship shown below in Equation 4-18 can be created where N is the number of bins used.

$$\frac{\sigma_{C,SelAsm(N)}}{\sigma_{C,RandomAsm}} = f\left(\frac{\sigma_A}{\sigma_B}\right) \tag{4-18}$$

Using Equation 4-2, the above equation can be rewritten as

$$\frac{\sigma_{C,SelAsm(N)}}{\sqrt{\sigma_A^2+\sigma_B^2}} = f\left(\frac{\sigma_A}{\sigma_B}\right) \qquad \text{4-19}$$

A function in the form of Equation 4-19 must be created for each combination of selective assembly technique and number of bins, e.g., a function exists for the Equal Width Binning to Minimize Scrap approach with five bins and another unique function exists for the Equal Area Binning to Minimize Variation approach with seven bins. Equations for the five approaches discussed in Sections 4.2.1 to 4.2.5 and for one to ten bins are created for a total of fifty unique equations.

The Monte Carlo based simulation model presented earlier is used to generate data that can be normalized and used to curve fit functions for each of the fifty scenarios. The part distributions are created so that they are normal and are truncated at plus and minus three standard deviations so that a finite number of bins can span the range. Values for the left hand side of Equation 4-19 are generated for values of σ_A/σ_B ranging from one to two. Running simulations for values outside of this range is not necessary because the validity of the desired equation is over the same range. The justification for using only this range is that combining distributions with standard deviations that vary by a factor of more than two is rarely done with selective assembly. Designating one distribution as Part "A" and the other as Part "B" is arbitrary. Therefore, if our function is good over the range of one to two for σ_A/σ_B, it can be used for any combination of standard deviations as long as one is not more than a factor of two larger. The Part "A" and Part "B" designations are made so that the ratio σ_A/σ_B falls between one and two, i.e., Part "A" will be the wide distribution and Part "B" will be the tight distribution.

The model (see Figure 4-8) is run approximately fifty times for each input σ_A/σ_B ranging from one to two with a step size of 0.2. Multiple runs were executed for each point to "smooth" the randomness associated with the Monte Carlo simulation and so that confidence intervals could be generated for each function, i.e., a ±95% probability band for each function. The data resulting from the Mote Carlo simulations are listed in tabular form in Appendix E.

Polynomial functions are fit to the simulated data. A third order term is the highest order term required for a good fit of the data, i.e., a third order polynomial does a good job of representing the shape of the data points for each case and results in a goodness of fit, R^2, close to one. The data could be fit to higher order terms, but the fit does not improve much with terms higher than cubic. Knowing that terms higher than cubic are not necessary, Equation 4-19 can be written in the general form below.

$$\frac{\sigma_{C,SelAsm(N)}}{\sqrt{\sigma_A^2+\sigma_B^2}} = (C_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (C_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (C_1)\left(\frac{\sigma_A}{\sigma_B}\right) + C_0 \qquad \text{4-20}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

The C_i's in the above equation are simply coefficients calculated from the curve fit. The coefficients for all of the techniques and numbers of bins are listed in Appendix C. For example, the C_i's for a Equal Width Binning to Minimize Scrap approach using four bins are C_3 = -0.1726, C_2 = 0.8266, C_1 = -1.1071, and C_0 = 0.8529.

Three cases exist where the normalized assembly standard deviation is better represented by a constant than a third order polynomial, i.e., $C_3 = 0$, $C_2 = 0$, and $C_1 = 0$ in the above equation. These cases are for the selective assembly techniques that bin the two distributions independently using one bin, i.e. Equal Width Binning to Minimize Scrap using one bin, Equal Area Binning to Minimize Scrap using one bin, and Optimal Width Binning using one bin. The reason the normalized assembly standard deviation is a constant in these three cases is because each case is simply truncating the part distributions and then combining them regardless of the distributions' standard deviations. This results in a value that is independent of the ratio of the standard deviations and is slightly less than the assembly standard deviation resulting from random assembly, i.e., the square root of sum of the variances. Aside from these three cases, the normalized assembly standard deviation is very well represented by a third order polynomial. More information about the goodness of fit of the curves is presented in Section 4.5.

The real goal is to have a closed-form equation for the standard deviation resulting from a selective assembly operation. This equation is easily derived from the normalized equation above. For a given number of bins, the standard deviation of the assembly is written as

$$\sigma_{C,SelAsm(N)} = \left((C_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (C_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (C_1)\left(\frac{\sigma_A}{\sigma_B}\right) + C_0 \right)\sqrt{\sigma_A^2 + \sigma_B^2} \qquad \text{4-21}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

Again, the C_i's are listed in Appendix C.

Because a selective assembly operation is a stochastic process, the standard deviation calculated using Equation 4-21 is actually the expected value of the assembly's standard deviation when selective assembly is used. Because of this, understanding the possible range of values the standard deviation can have is very important. To gain this understanding, confidence intervals can be calculated. A 95% confidence interval is calculated for the assembly standard deviation. The equations for the upper and lower bands of the confidence interval are of the same form as Equation 4-21 with different coefficients. Appendix C contains the values for the coefficients for the upper and lower 95% confidence interval bands.

The goal is to use these equations to help forecast quality when a selective assembly operation is being considered. The first step is to determine which selective assembly technique is being used and the number of bins to use. If these are unknown, results for each technique and each number of bins can be computed. So, for the first approach and number of bins, the coefficients for σ_C are looked up in the tables in Appendix C. These coefficients are used with Equation 4-21 to compute the expected value of the assembly standard deviation. The coefficients for the 95% confidence upper and lower bands are looked up and used with Equation 4-21 to calculate the confidence interval of the assembly standard deviation. For the given set of two part distributions, the true value of the assembly standard deviation has a 95% probability of being in this interval.

Now that a closed-form equation for the standard deviation resulting from selective assembly has been created, focus can be turned to creating a closed-form equation for the fraction of parts that do not have a corresponding match. These unmatched parts are known as scrap. Here the term scrap fraction is used to imply a fraction of the total parts that do not have a match.

Scrap is generated because of distribution truncation and when corresponding bins have an unequal number of parts in them. The equations developed in this section assume that the mating parts are produced in equal quantities, i.e., one distribution is not overproduced to ensure a match. This assumption implies that the scrap is equally divided among the two mating part distributions. For example, twenty unmatched parts from the "A" distribution means there are twenty unmatched parts from the "B" distribution. The goal of this section is to generate a closed-form equation for scrap fraction of the form seen below.

$$Scrap = f\left(\sigma_{PartA}, \sigma_{PartB}\right) \qquad 4\text{-}22$$

The scrap is calculated as a fraction of parts that are not used in an assembly. The desired equation is the fraction of parts not used and is already a normalized quantity. The simulation data used above to generate the equations for an assembly's standard deviation are also used with curve fitting to produce equations for scrap. The highest power term required to accurately fit any of the data, but not overfit, is a third order term, i.e., all the data could be well represented with equations that contain terms no higher than cubic. The resulting equation for the scrap, as a fraction of total parts, is shown below as Equation 4-23.

$$Scrap_{(N)} = \left((D_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (D_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (D_1)\left(\frac{\sigma_A}{\sigma_B}\right) + D_0 \right) \qquad 4\text{-}23$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

The D_i's in the above equation are simply coefficients calculated from the curve fit. The coefficients for all of the techniques and numbers of bins are listed in Appendix D. For example, the D_i's for a Equal Width Binning to Minimize Variation technique using six bins are D_3 = -0.0993, D_2 = 0.3686, D_1 = -0.1379, and D_0 = -0.1044 .

The selective assembly techniques that bin the part distributions independently, Equal Width Binning to Minimize Scrap, Equal Area Binning to Minimize Scrap, and Optimal Width Binning, have scrap equations that are better represented by a constant value than a polynomial, i.e., $Scrap_{(N)} = D_0$. This representation is due to the fact that the distributions are binned from minus to plus three sigma; therefore, approximately the same number of parts are being scrapped regardless of the distributions variances. The techniques that bin the wide distribution according to the tight distribution's variance, Equal Width Binning to Minimize Variation and Equal Area

Binning to Minimize Variation, truncate more of the wide distribution, especially as its variance increases relative to the tight distribution. For these techniques, the third order polynomial is an excellent fit with goodness of fit values close to one. More information about the goodness of fit of the curves is presented in Section 4.5.

Due to the stochastic nature of a selective assembly operation, the scrap fraction calculated using Equation 4-23 is actually the expected value of the fraction of scrap generated when selective assembly is used. Because of this, understanding the possible range of scrap is very important. To gain this understanding, a 95% confidence interval is calculated. The equations for the upper and lower bands of the confidence interval are of the same form as Equation 4-23 with different coefficients. Appendix D contains values for coefficients for the upper and lower 95% confidence interval bands.

The goal is to use these equations to help forecast the scrap fraction when a selective assembly operation is being considered. The first step is to determine which selective assembly technique is being used and the number of bins to use. If these are unknown, results for each technique and each number of bins can be computed. So, for the first technique and number of bins, the coefficients for Equation 4-23 are looked up in the tables in Appendix D. These coefficients are used with Equation 4-23 to compute the expected value of the scrap fraction. The coefficients for the 95% confidence upper and lower bands are looked up and used with Equation 4-23 to calculate the confidence interval of the scrap fraction. For the given set of two part distributions, the true amount of scrap has a 95% probability of being in this interval.

The closed-form equations do have two limitations worth noting. First, the equations are valid only over the range of one to two for the ratio of component standard deviations, i.e., $1 \leq \sigma_A/\sigma_B \leq 2$. This limitation is by design since this was the range considered when the simulations were run to produce the data for curve fitting. The second limitation is the equations are valid only when two normal distributions are being assembled. Again, this limitation is by design and is a result of the choice to run simulations with normal distributions, since they are the most common. If the need arose for closed-form equations to combine other types of distributions, the development used in this section could be repeated with the different types of distributions to create another set of closed-form equations, with no guarantee of existence.

4.4.2 Sensitivity Analysis

Sensitivity analysis is performed to determine how small changes of input values affect the output value. An output value that does not change much when the inputs are perturbed is considered to be insensitive and is known as a robust design. Considering the sensitivity of a selective assembly operation is important when choosing the technique to use.

The sensitivity around a specific point is calculated by taking the first derivative of the output with respect to the input variables. For the standard deviation resulting from selective assembly, this can be written as follows.

$$\Delta\sigma_{C,SelAsm(N)} = \left(\left. \frac{\partial \sigma_{C,SelAsm(n)}}{\partial \sigma_A} \right|_{(\sigma_A^*,\sigma_B^*)} \right) \Delta\sigma_A + \left(\left. \frac{\partial \sigma_{C,SelAsm(n)}}{\partial \sigma_B} \right|_{(\sigma_A^*,\sigma_B^*)} \right) \Delta\sigma_B \qquad \text{4-24}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

where σ_A^*, σ_B^* is the point about which the analysis is being performed. Substituting Equation 4-21 into the above equation yields the following equation.

$$\begin{aligned} \Delta\sigma_{C,SelAsm(N)} = & \left[\left(3C_3 \frac{\sigma_A^2}{\sigma_B^3} + 2C_2 \frac{\sigma_A}{\sigma_B^2} + C_1 \frac{1}{\sigma_B} \right) \left(\sqrt{\sigma_A^2 + \sigma_B^2} \right) + \frac{\sigma_A \left(C_3 \left(\frac{\sigma_A}{\sigma_B} \right)^3 + C_2 \left(\frac{\sigma_A}{\sigma_B} \right)^2 + C_1 \left(\frac{\sigma_A}{\sigma_B} \right) + C_0 \right)}{\sqrt{\sigma_A^2 + \sigma_B^2}} \right]_{(\sigma_A^*,\sigma_B^*)} \Delta\sigma_A \\ & + \left[\left(-3C_3 \frac{\sigma_A^3}{\sigma_B^4} - 2C_2 \frac{\sigma_A^2}{\sigma_B^3} - C_1 \frac{\sigma_A}{\sigma_B^2} \right) \left(\sqrt{\sigma_A^2 + \sigma_B^2} \right) + \frac{\sigma_B \left(C_3 \left(\frac{\sigma_A}{\sigma_B} \right)^3 + C_2 \left(\frac{\sigma_A}{\sigma_B} \right)^2 + C_1 \left(\frac{\sigma_A}{\sigma_B} \right) + C_0 \right)}{\sqrt{\sigma_A^2 + \sigma_B^2}} \right]_{(\sigma_A^*,\sigma_B^*)} \Delta\sigma_B \end{aligned} \qquad \text{4-25}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

A similar analysis can be performed to calculate the scrap fraction sensitivity. The basic equation is as follows.

$$\Delta Scrap_{(N)} = \left(\left. \frac{\partial Scrap_{(N)}}{\partial \sigma_A} \right|_{(\sigma_A^*, \sigma_B^*)} \right) \Delta\sigma_A + \left(\left. \frac{\partial Scrap_{(N)}}{\partial \sigma_B} \right|_{(\sigma_A^*, \sigma_B^*)} \right) \Delta\sigma_B \qquad \text{4-26}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

As before, σ_A^*, σ_B^* is the point about which the analysis is being performed. Substituting Equation 4-23 into the above equation yields the following equation.

$$\begin{aligned} \Delta Scrap_{(N)} = & \left[\left. \left(3D_3 \frac{\sigma_A^2}{\sigma_B^3} + 2D_2 \frac{\sigma_A}{\sigma_B^2} + D_1 \frac{1}{\sigma_B} \right) \right] \right|_{(\sigma_A^*, \sigma_B^*)} \Delta\sigma_A \\ & + \left[\left. \left(-3D_3 \frac{\sigma_A^3}{\sigma_B^4} - 2D_2 \frac{\sigma_A^2}{\sigma_B^3} - D_1 \frac{\sigma_A}{\sigma_B^2} \right) \right] \right|_{(\sigma_A^*, \sigma_B^*)} \Delta\sigma_B \end{aligned} \qquad \text{4-27}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

Equations 4-25 and 4-27 are used to calculate the effect small changes in the part distributions have on the assembly distribution and scrap fraction respectively. However, it is more interesting to use the values of the partial derivative terms to indicate how sensitive a selective assembly technique is to changes in the part distributions' standard deviations. These sensitivities reveal where batch to batch variation is most crucial. For example, if the standard deviation of Part "A" changes from batch to batch, choosing a selective assembly operation that is less sensitive to changes in σ_A would be beneficial to creating more consistent assemblies. The sensitivities also aid in choosing a selective assembly approach when multiple machining lines feed into the selective assembly operation. For example, if multiple machining lines are used to create Part "B," the standard deviations may be different for each of the distributions coming from the multiple lines. In this case, picking a selective assembly approach that is insensitive to the variance of Part "B" would be beneficial.

4.5 Selective Assembly Equation Validation

The closed-form equations for the assembly standard deviation and scrap fraction developed in Section 4.4.1 can be validated by quantifying the goodness of fit for the third order polynomial equations to the simulated data and checking to see if the equations accurately represent the data. The goodness of fit that a curve has to a group of data is represented by R^2. R^2 can have a value between zero and one. A value of zero indicates that a horizontal line passing through the average values of Y fits the data as well as the "best-fit" line. A value of one indicates all the data points lie exactly on the line. A value close to one is preferred but should not be the only test as to how well the fit curve represents the data. A simple visual test provides insight into the accuracy of the fit curve (1999).

The best fit normalized assembly standard deviation curves to the data (Equation 4-20), i.e., where the R^2 values for each technique were the closest to one, is when five bins are used. When five bins are used, the R^2 values range from 0.993 to 1. The points representing the simulated data along with the best-fit curves are shown below in Figure 4-9. Each point represents the average of thirty Monte Carlo simulations, e.g., averaging the results from thirty simulations of combining one thousand variables using the Equal Width Binning to Minimize Scrap technique produced a normalized assembly standard deviation of 0.4 for a σ_A/σ_B ratio of 1.4.

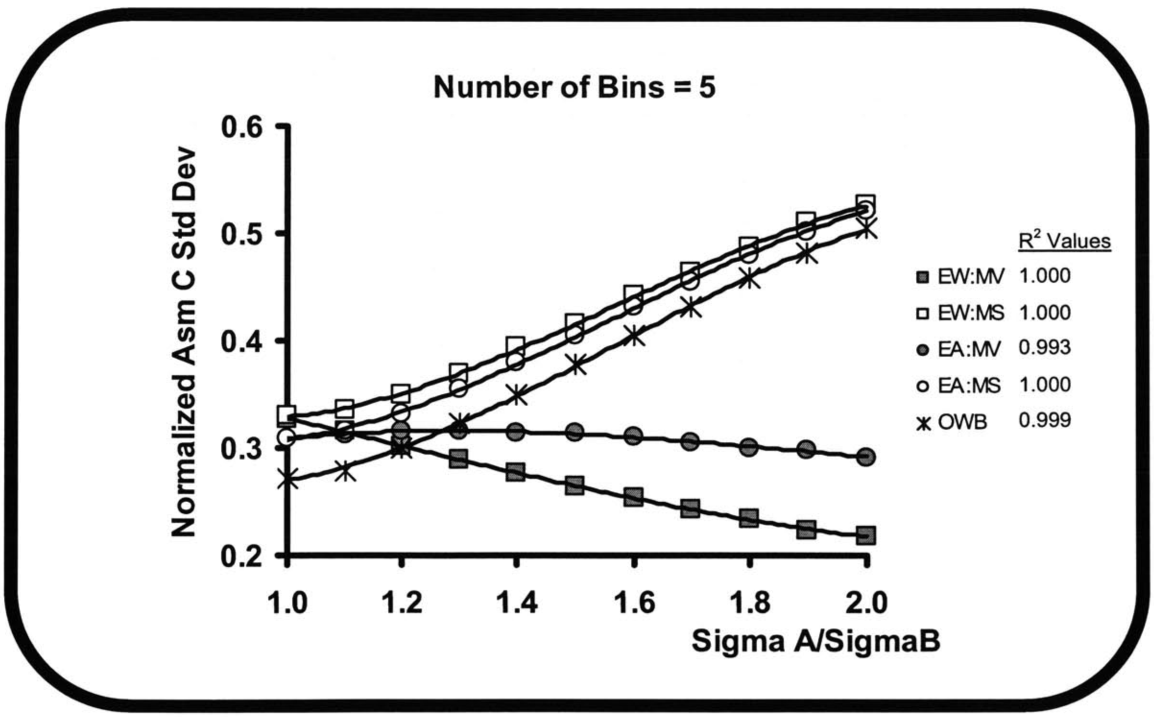

Figure 4-9: R^2 values closest to one for assembly standard deviation

For the case of five bins, the curves appear to represent the data very accurately capturing the trends of the data. The lowest value for R^2 occurs when fitting normalized assembly standard deviation curves to the data for the case of using seven bins. The curve for Equal Area Binning to Minimize Scrap has an R^2 value of 0.974, which is still very close to the desired value of one. Looking at the curves and the simulated data points in Figure 4-10 shows that the curves do fit the data extremely well.

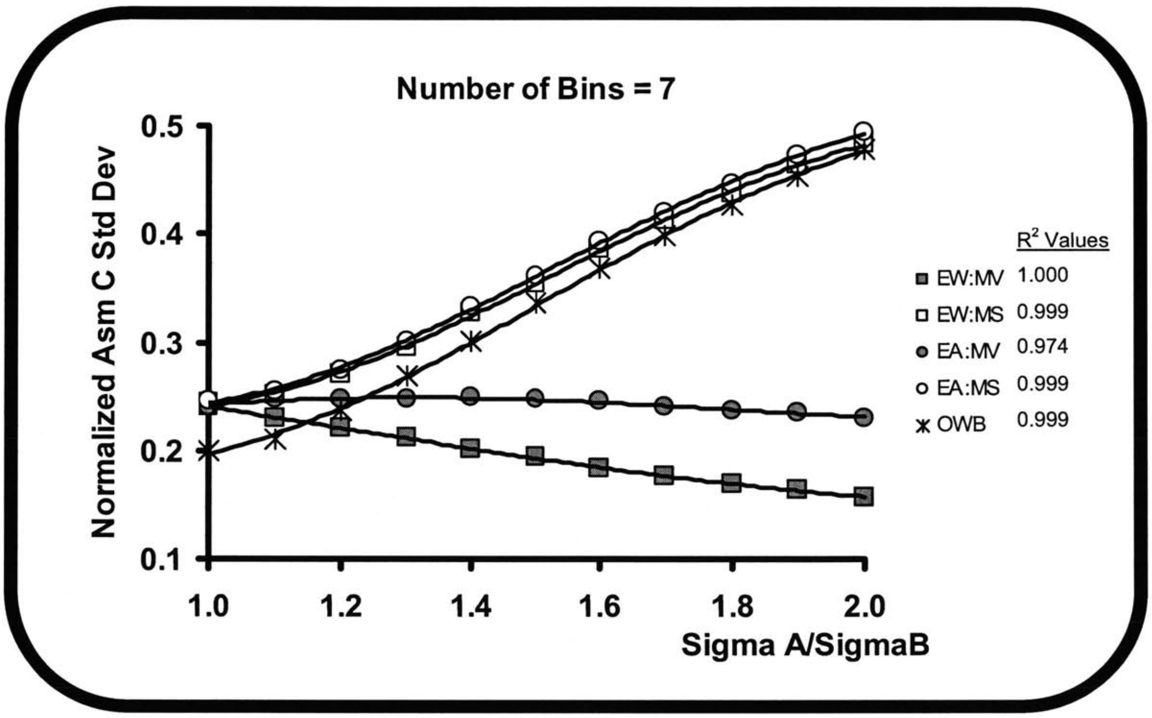

Figure 4-10: Lowest R^2 for assembly standard deviation

Investigating the curve fit for the ±95% confidence interval reveals similar results. One case where the curves for expected value, -95% confidence interval band, and +95% confidence interval band fit the data almost perfectly is Equal Width Binning to Minimize Variation using four bins. For this case, the R^2 values are 1, 0.999, and 1 respectively. The data points and curves fit to them are shown below in Figure 4-11. The curves fit the data very accurately. The worst fit of confidence intervals occurs for the Equal Area Binning to Minimize Variation technique using ten bins. For this case, the lower band has an R^2 value of 0.654 and the upper band has an R^2 value of 0.799. Looking at the plots of the data points and curves shown in Figure 4-12 reveals that even with these R^2 values the curves still accurately reflect the nature of the data. Note that the scale of the vertical axis is set in each of the graphs to show as much detail as possible.

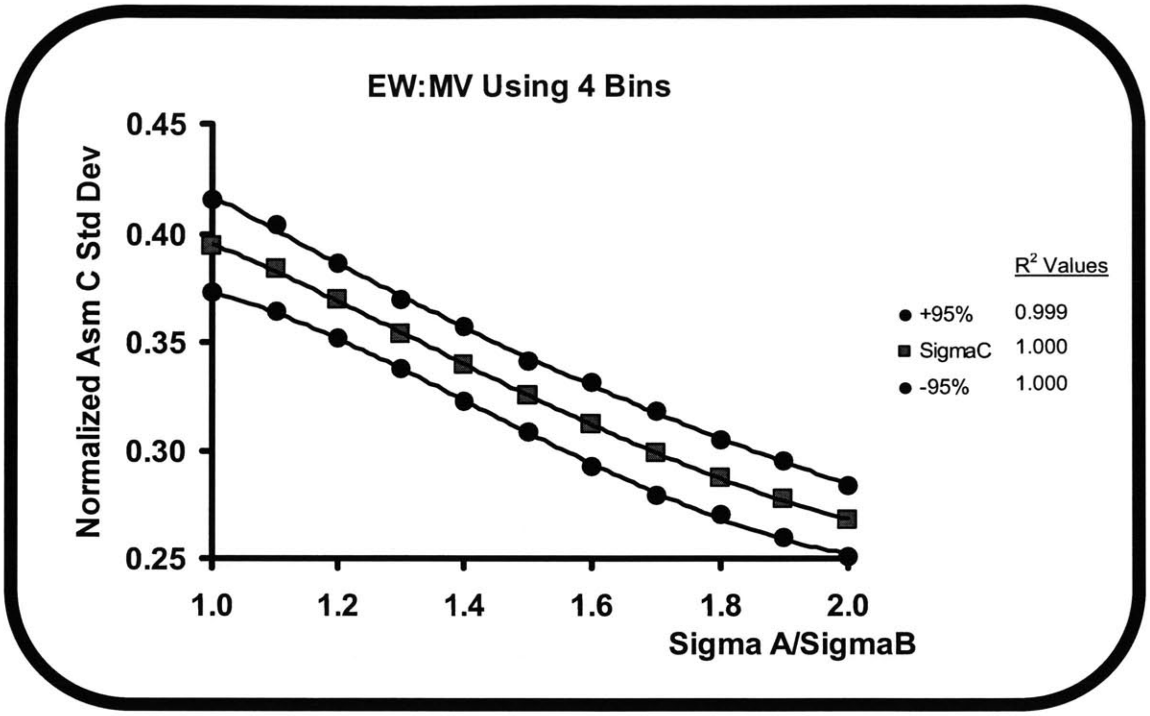

Figure 4-11: Example of confidence intervals with high R^2 values

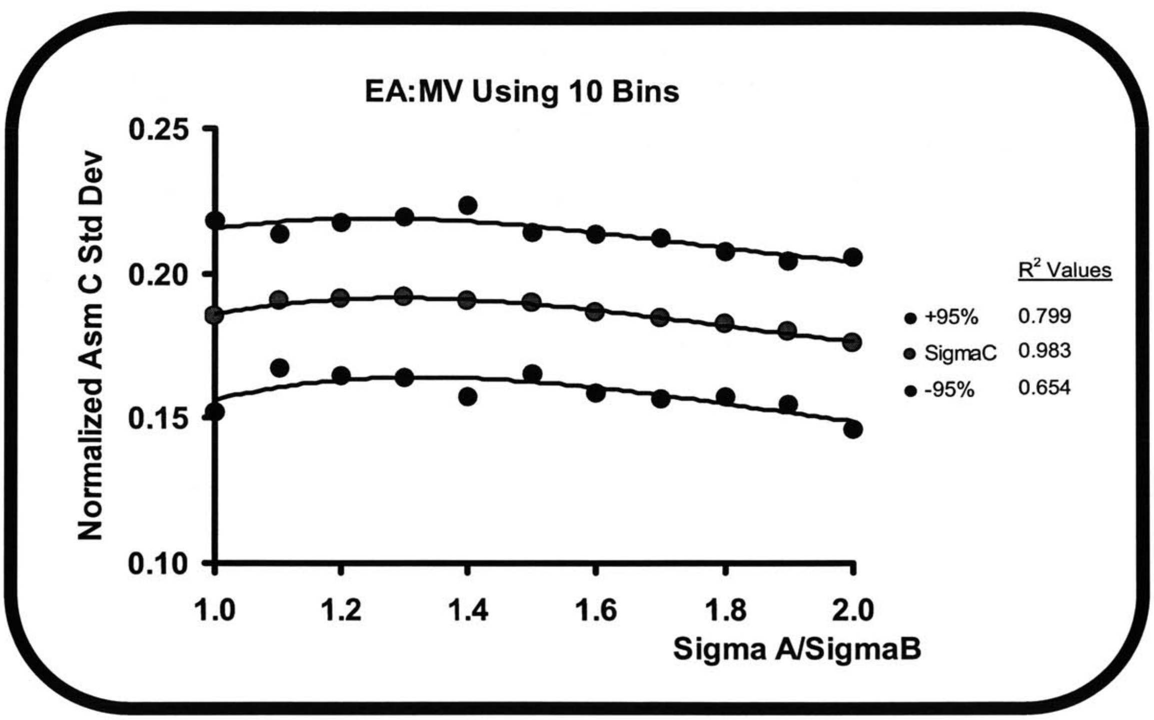

Figure 4-12: Confidence intervals with the lowest R^2 values

The scrap equations created also need to be checked for accuracy. As discussed earlier, the selective assembly techniques where the wide distribution is binned relative to the tight distribution, Equal Width Binning to Minimize Variation and Equal Area Binning to Minimize Variation, have scrap equations that are well represented by third order polynomials. The case with the highest R^2 values for the scrap equations is when three bins are used. The fit curves and the data are graphed in Figure 4-13 below. A worst case for the scrap curve fit for Equal Width Binning to Minimize Variation or Equal Area Binning to Minimize Variation does not exist because all the R^2 values are above 0.997, and all the equations fit the data very closely.

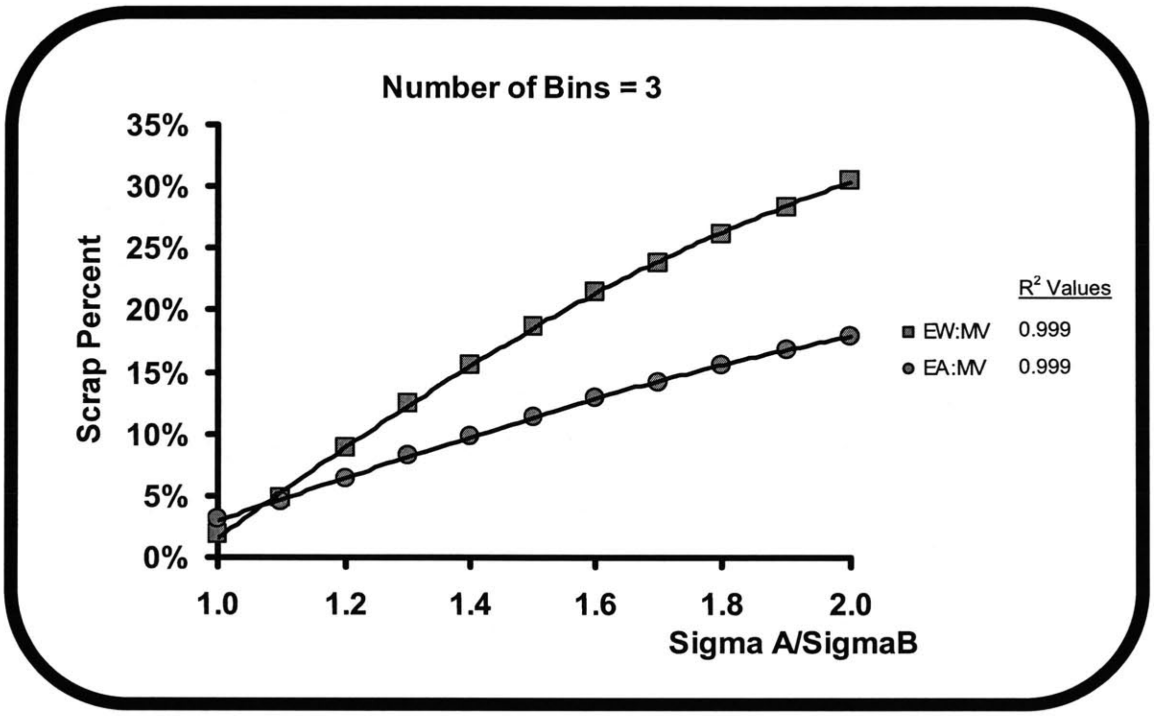

Figure 4-13: R^2 values closest to one for scrap

The selective assembly techniques where the part distributions are binned independently, Equal Width Binning to Minimize Scrap, Equal Area Binning to Minimize Scrap, and Optimal Width Binning, have scrap equations that are well represented by a constant as explained in Section 4.4.1. A typical example of how well a constant value represents the data is shown below in Figure 4-14.

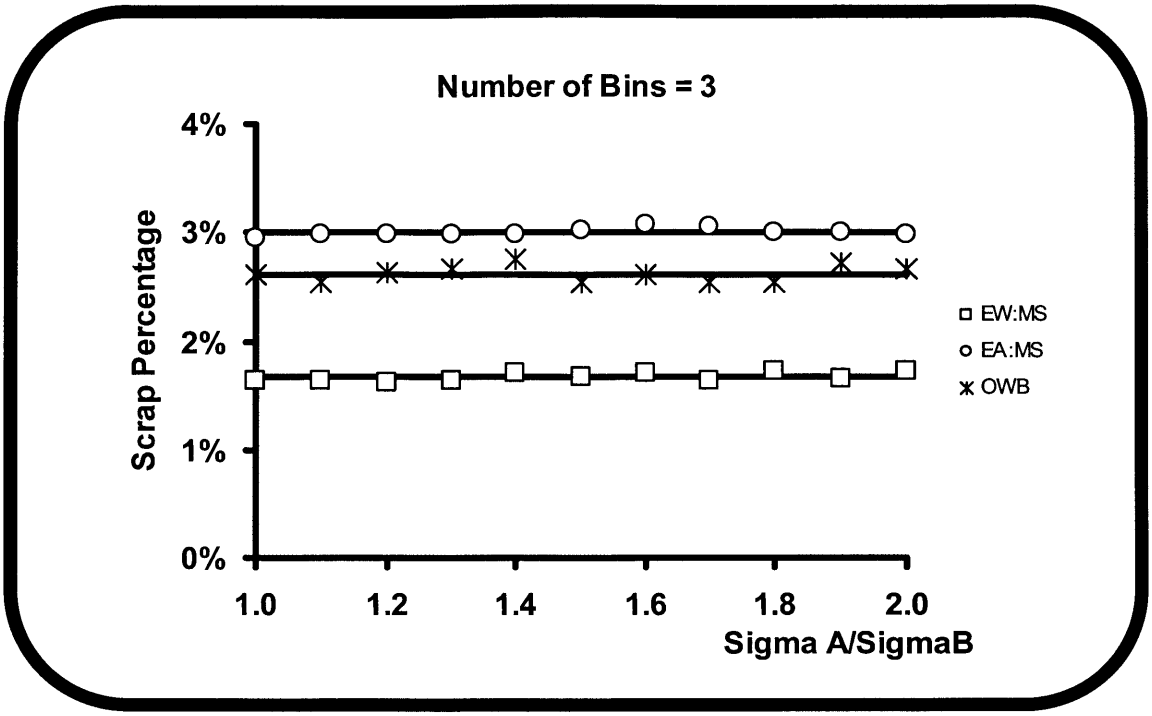

Figure 4-14: Typical case for a constant value best representing a scrap equation

The selective assembly simulation data used to curve fit the equations are listed in tabular format in Appendix E along with the R^2 values for all the third order polynomial fit curves.

4.6 Example Problems

Three examples are presented in this section to demonstrate the models presented above. One example is combining two normal distributions and takes advantage of the closed-form solutions derived earlier. The other two examples are combining non-normal distributions and use numerical simulation to do so.

4.6.1 Selective Assembly of Two Normal Distributions

An example problem is presented here to demonstrate how the methods and equations introduced in the previous few sections are used to choose an appropriate selective assembly approach.

Problem Statement

Two parts are joined together to form an assembly (see Figure 4-1). The length of Part "A" is normally distributed and has a standard deviation of 2.4. The length of Part "B" is also normally distributed and has a standard deviation of 2.0. Due to functional constraints, the overall length of the assembly, Assembly "C," must have a standard deviation less than 1.5. (Note that the numbers are kept dimensionless since the units would be arbitrary and not affect any calculations.) What selective assembly technique results in an assembly length distribution that will have a standard deviation less than 1.5 and minimize the expected amount of scrap?

Table 4-1: Nomenclature for example problem

Abbreviation	**What it means...**
Part A Std Dev, σ_A	Standard deviation of the length of Part "A"
Part B Std Dev, σ_B	Standard deviation of the length of Part "B"
Asm C Std Dev, σ_C	Standard deviation of the length of Assembly "C"
EW:MV	Equal Width Binning to Minimize Variation
EW:MS	Equal Width Binning to Minimize Scrap
EA:MV	Equal Area Binning to Minimize Variation
EA:MS	Equal Area Binning to Minimize Scrap
OWB	Optimal Width Binning
Scrap Fraction	(Number of unmatched parts)/(Total number of parts)
N	Number of bins used for a distribution in selective assembly

The value for the Asm C Std Dev if random assembly were used is used as the baseline. This value is calculated using Equation 4-2. For the given values for the part distributions' standard deviations, the Asm C Std Dev is 3.1.

Plotting the Asm C Std Dev versus the number of bins used for each selective assembly approach allows for a determination to be made as to how many bins are required to produce an assembly with a standard deviation below the desired level. Equation 4-21 and the coefficients in Appendix C are used to generate the necessary values. These values are plotted in Figure 4-15 below.

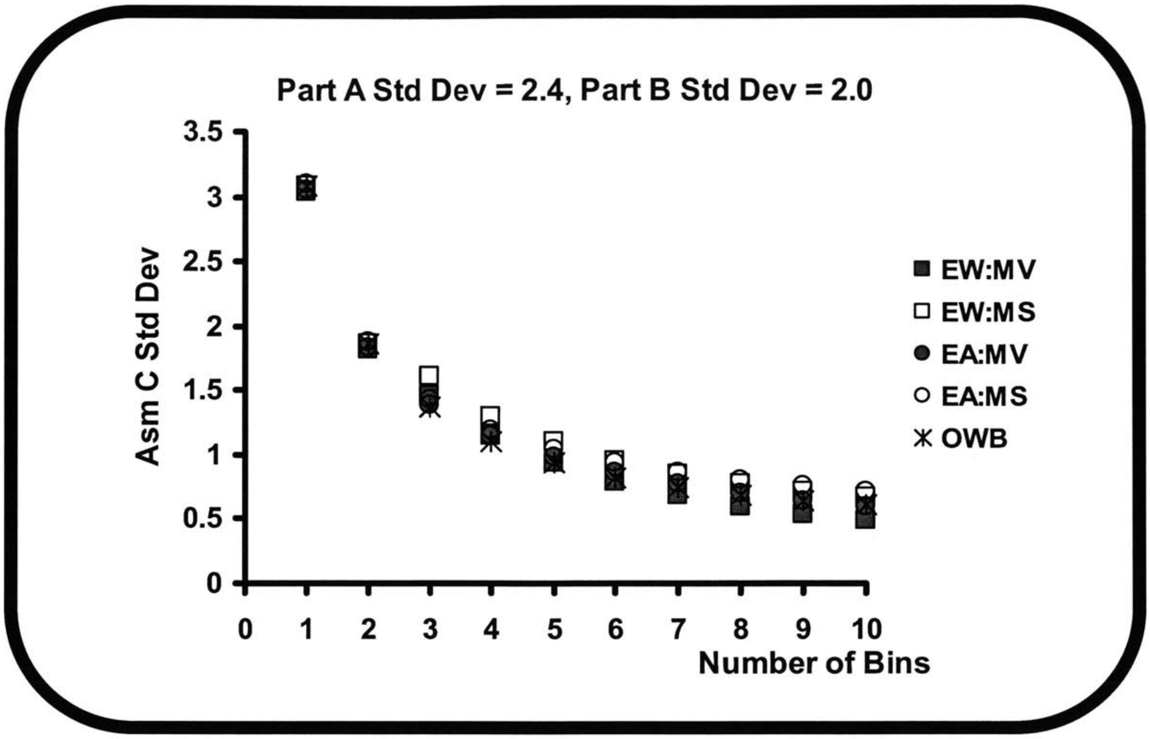

Figure 4-15: Assembly standard deviation versus number of bins for multiple approaches

To ensure the assembly has a standard deviation less than 1.5, the number of bins required is four. From Figure 4-15, determining exactly which selective assembly approach to use is difficult because all the points for each technique are so tightly grouped. Regardless of this tight grouping, Figure 4-15 does not provide any insight into the amount of scrap that is generated with each approach and number of bins. Therefore, Equation 4-23 is used to compute the expected amount of scrap, as a fraction, for each selective assembly approach and number of bins. These values are plotted in the graph shown below in Figure 4-16.

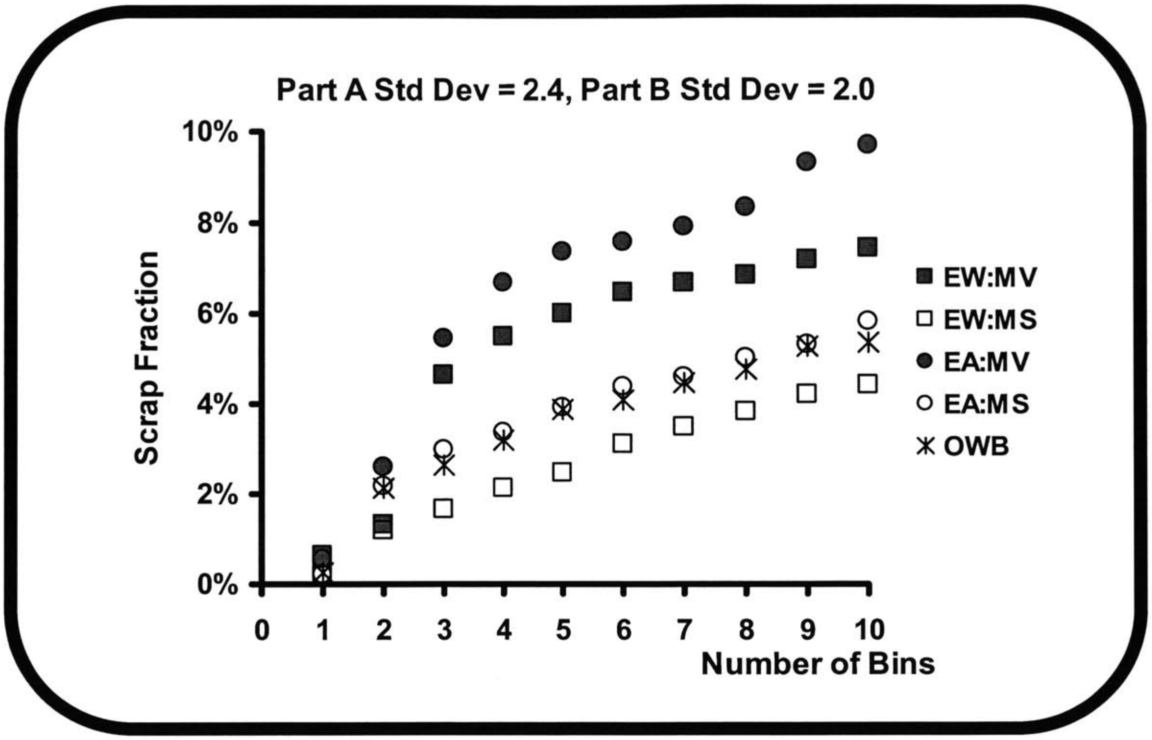

Figure 4-16: Scrap fraction versus the number of bins used for multiple approaches

Figure 4-16 reveals that for four bins, the Equal Width Binning to Minimize Scrap (EW:MS) approach generates the least amount of scrap of the five approaches studied. This implies that the Equal Width Binning to Minimize Scrap approach is the best to use for the given information and requirements since it generates the least amount of scrap and results in an assembly length that has a standard deviation less than the specified limit of 1.5.

This conclusion makes sense based on the definition of the Equal Width Binning to Minimize Scrap approach. Remember that this approach utilizes almost the entire range of both part distributions, thus leaving very few parts unmatched and not producing much scrap from distribution tail truncation. Knowing that this approach produces a desirable level of assembly length variation, choosing it is natural due to its given ability to not waste parts.

The same conclusion could have been reached more quickly through the use of a single graph. The graph plots Asm C Std Dev versus Scrap Fraction. These values are computed using Equations 4-21 and 4-23 respectively for each selective assembly technique and number of bins. The graph is shown below in Figure 4-17. Using this graph, a line representing the specification limit is drawn horizontally across from the appropriate value (1.5 in this case). All points above this line are unacceptable, while all points below this line are acceptable. The point that falls below the spec limit and is furthest to the left is selected as the best point since it meets the specification and produces the least amount of scrap. Figure 4-18 shows the spec limit and the chosen point. The chosen point represents the Equal Width Binning to Minimize Scrap approach with four bins. Notice this is the same conclusion derived before but in a quicker manner.

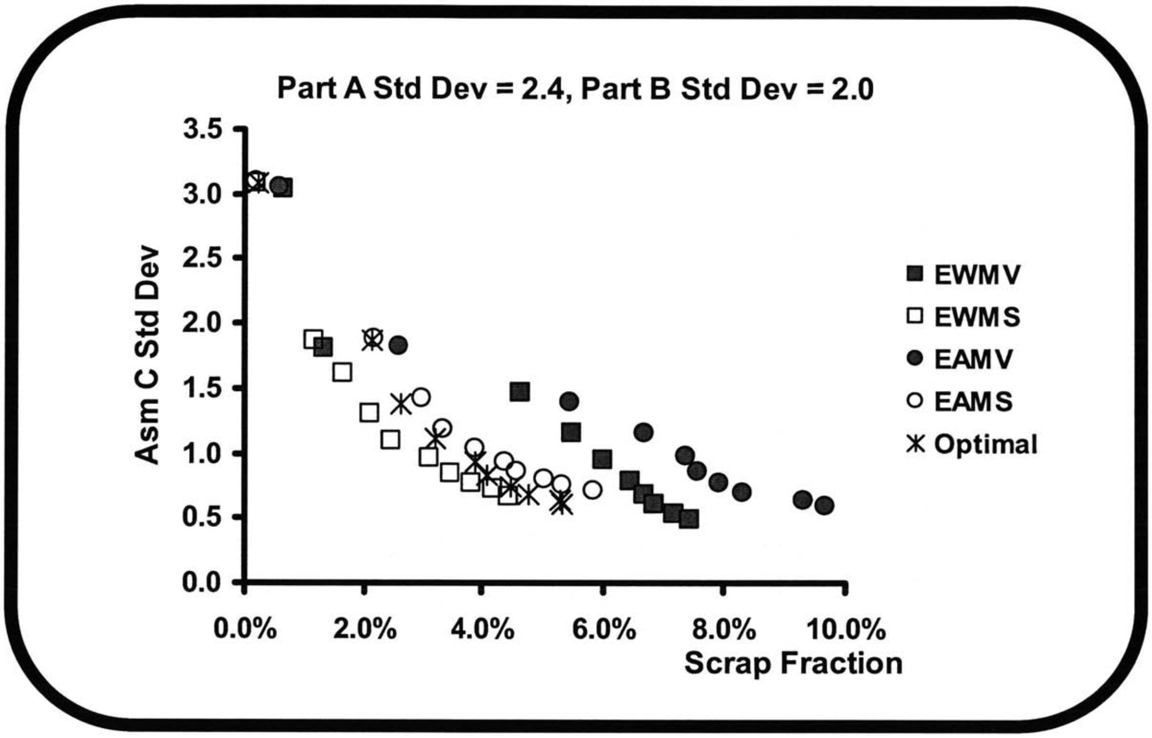

Figure 4-17: Assembly standard deviation versus scrap fraction

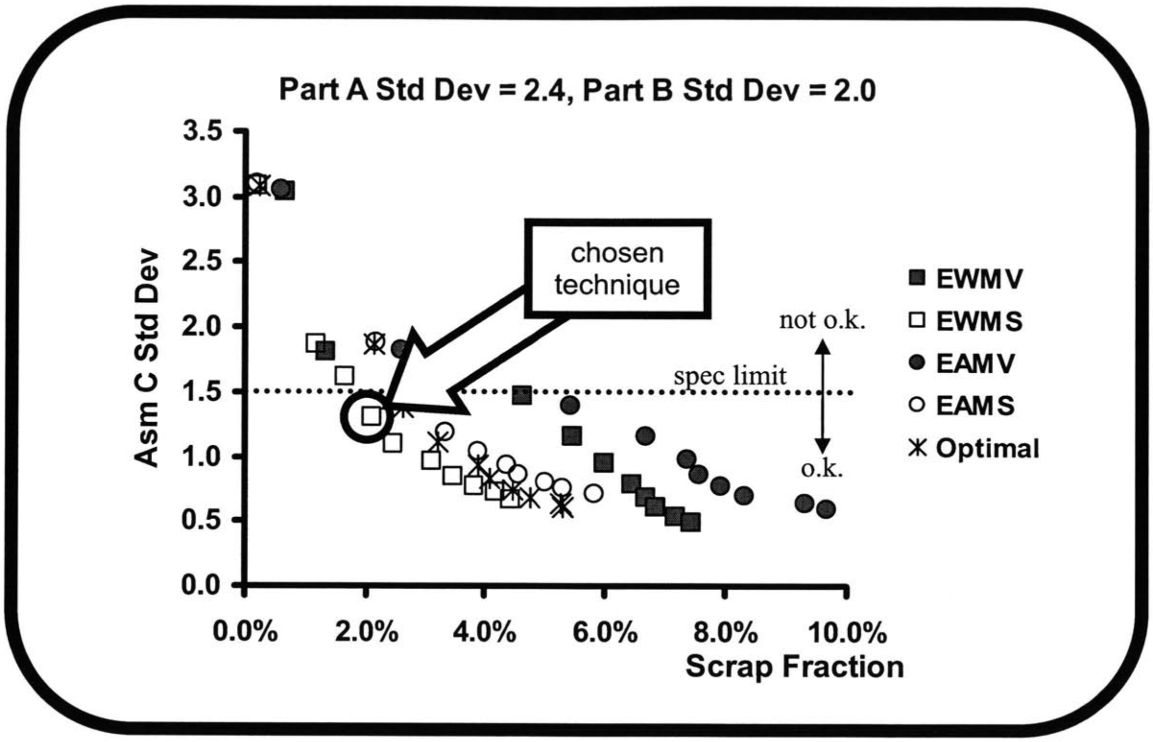

Figure 4-18: Solution point to example problem

Once the technique and number of bins have been chosen, it is important to consider the ±95% confidence interval for the σ_C and Scrap Fraction. Remember, due to the stochastic nature of selective assembly the values calculated by Equations 4-21 and 4-23 are the expected values of the assembly standard deviation and scrap fraction. To ensure the assembly standard deviation is below the specification limit of 1.5 with the given part distributions, the ±95% confidence interval for the σ_C is plotted using Equation 4-21 and values from Appendix C. This plot is seen below in Figure 4-19. The plot reveals that there is at least a 95% probability that with the given input distributions ($\sigma_A = 2.4$ and $\sigma_B = 2.0$), the assembly standard deviation will be less than 1.5.

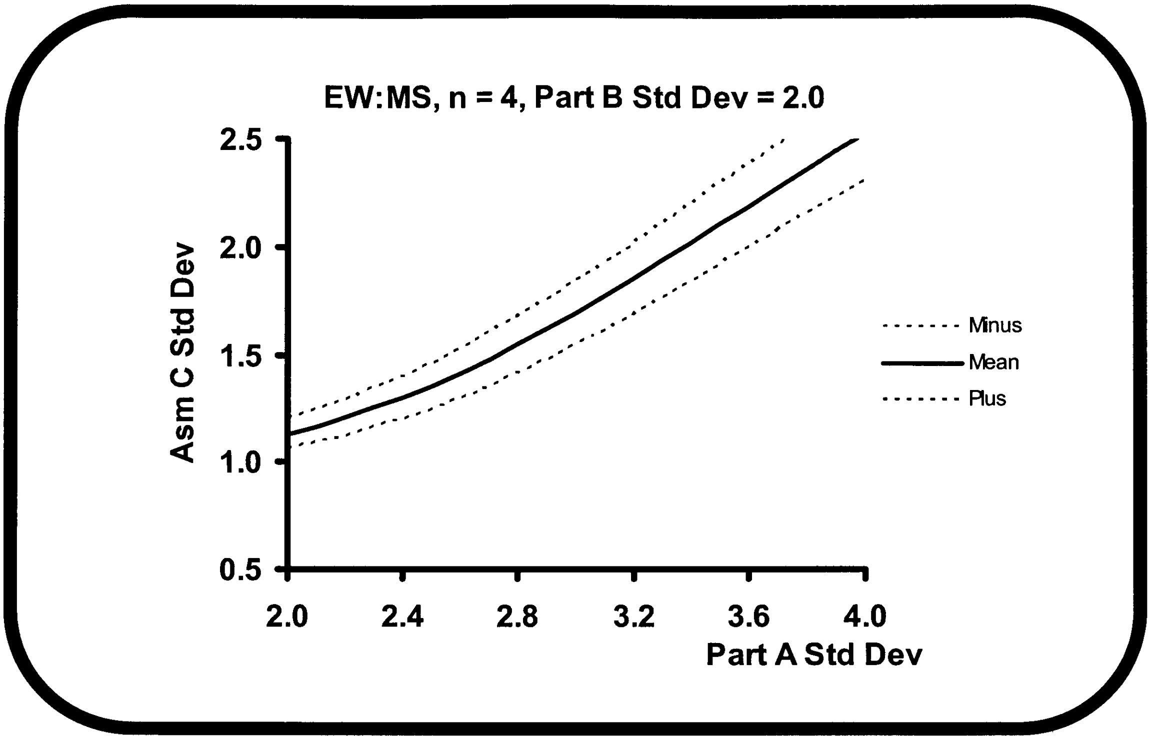

Figure 4-19: Assembly standard deviation with ±95% confidence interval

Figure 4-19 may also be used for sensitivity analysis. First, the sensitivities with respect to the input variables, σ_A and σ_B, are computed. From the results in Table 4-2, the value for the assembly standard deviation is much more sensitive to changes in the value of σ_A than to changes in the value of σ_B.

Table 4-2: Sensitivities for assembly standard deviation

$\left.\frac{\partial \sigma_{C,SelAsm(N)}}{\partial \sigma_A}\right\rvert_{(\sigma_A=2.4,\sigma_B=2.0)}$	$\left.\frac{\partial \sigma_{C,SelAsm(N)}}{\partial \sigma_B}\right\rvert_{(\sigma_A=2.4,\sigma_B=2.0)}$
0.525	0.021

With this conclusion, Figure 4-12 is used for sensitivity analysis, and the sensitivity to σ_B is safely neglected. From Figure 4-12, it is shown that σ_A can vary to a value as high as approximately 2.5 while maintaining a 95% probability that the assembly standard deviation remains below the specified limit of 1.5.

The scrap fraction is also a stochastic value; therefore, investigating its confidence interval is important. The ±95% confidence interval for the scrap fraction is plotted below in Figure 4-20. Note that the function and the confidence intervals are straight horizontal lines. This shape makes sense knowing that the approach is one that minimizes scrap. Therefore, regardless of the shape of the two part distributions, a similar number of scrap parts is generated. This intuition is proven mathematically using Figure 4-20 for sensitivity analysis, which shows the scrap fraction is insensitive to the value of the standard deviations of the part distributions.

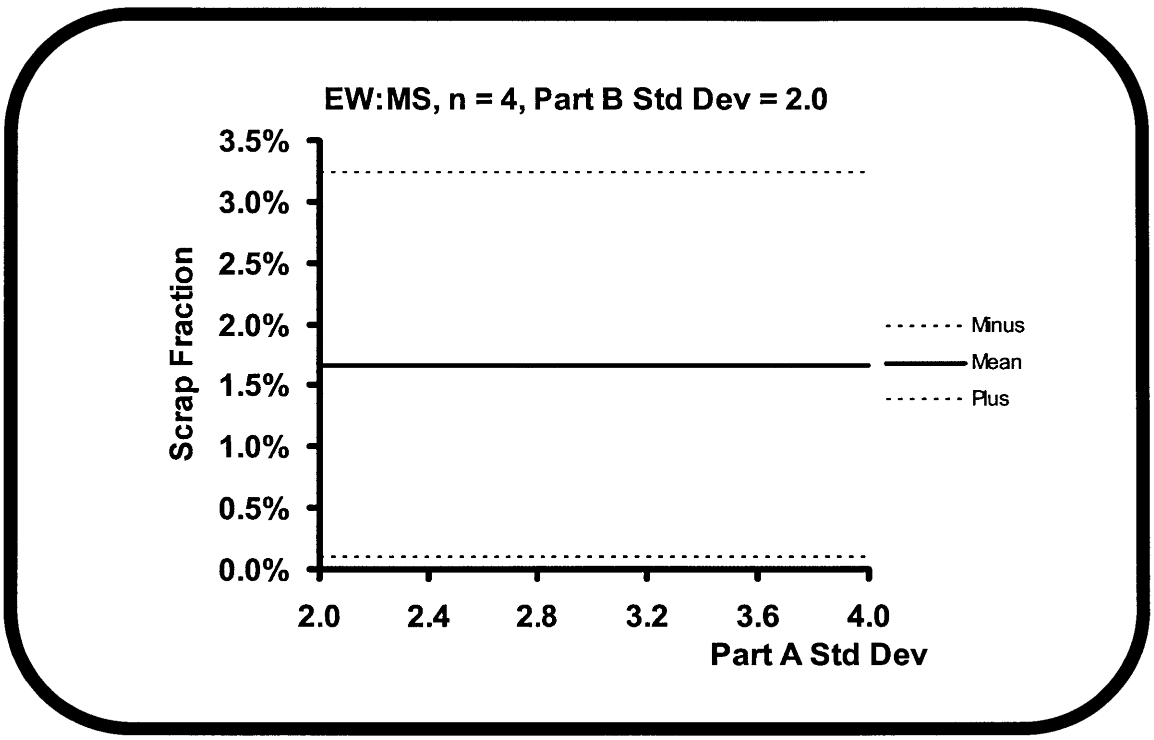

Figure 4-20: Scrap fraction with ±95% confidence interval

Even though the technique and number of bins that result in an acceptable level of assembly variation and minimizes the scrap has been identified, investigating the sensitivities of the other approaches for the same number of bins may provide insight into the strengths and weaknesses of each approach for the given problem. The assembly standard deviation versus σ_A is plotted below in Figure 4-21 for the case when σ_B is equal to 2.0 and four bins are used. The range of σ_A values represents a "shift" away from the case of two distributions with equal variance being combined (at $\sigma_A = 2.0$) to a case where the standard deviation of one distribution is twice that of the other (at $\sigma_A = 4.0$).

In previous selective assembly work, assuming the two distributions being combined have equal variance is common. Knowing that this equality is not always true, an engineer should make intelligent comparisons among the techniques for a particular case. For example, if the overall goal of an engineer is to minimize the variance regardless of the cost, then the Optimal Width Binning would be the proper choice when the standard deviation of Part "A" falls between 2.0 and 2.4. The proper choice would be Equal Width Binning to Minimize Variance for values of σ_A greater than 2.4. These conclusions are made based on examination of the plot in Figure 4-21 below. This choice matches intuition because of the nature of the approach. The EW:MV approach heavily truncates the wider distribution (in this case the distribution of Part "A") to maintain a "width" that is close to that of the mating part's distribution (Part "B"). This truncation keeps the assembly variation low and very consistent.

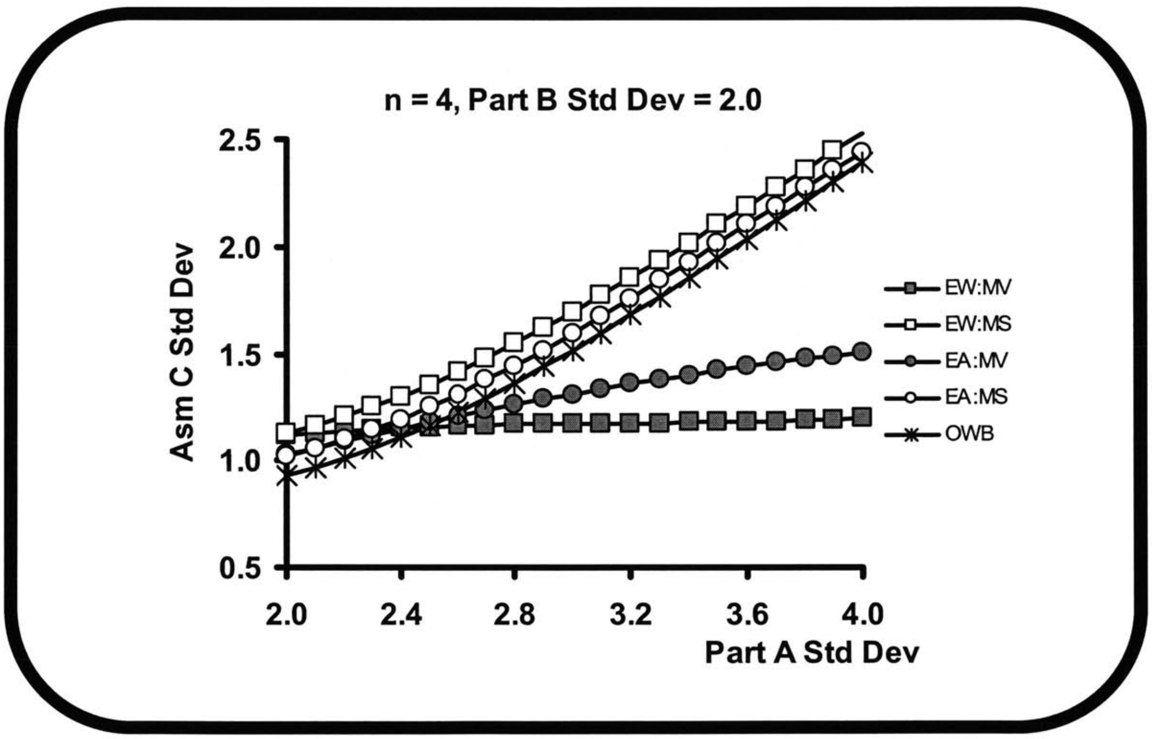

Figure 4-21: Comparing the assembly standard deviation for multiple approaches using four bins

As emphasized throughout this chapter, the variance of the assembly is only one consideration; the amount of scrap generated should not be ignored. The scrap fraction versus σ_A is plotted below in Figure 4-22 for the case when σ_B is equal to 2.0 and four bins are used. Again, the range of σ_A values represents a "shift" away from the case of two distributions with equal variance being combined (at $\sigma_A = 2.0$) to a case where the standard deviation of one distribution is twice that of the other (at $\sigma_A = 4.0$). Just as with the assembly's standard deviation, an engineer should make intelligent comparisons among the techniques for a particular case, e.g. $\sigma_B = 2.0$ and $N = 4$. For example, if the overall goal of an engineer is to minimize the scrap generated and gain whatever reduction in assembly standard deviation comes naturally, then the Equal Width Binning to Minimize Scrap would be the proper choice for all values of the standard deviation of Part "A." This conclusion is based on examination of the plot in Figure 4-22 below.

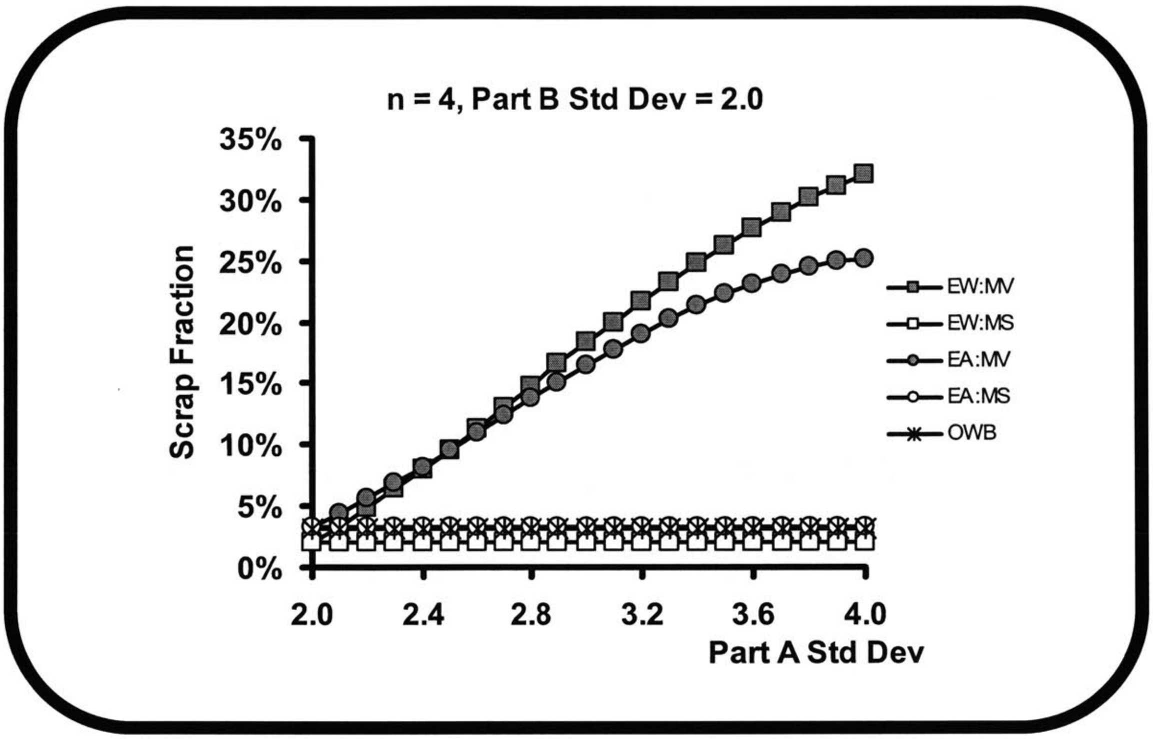

Figure 4-22: Comparing the scrap fraction for multiple approaches using four bins

Figure 4-21 and Figure 4-22 are also used to compare the robustness of the various selective assembly approaches. For σ_C, the selective assembly techniques that try minimizing variation, EW:MV and EA:MV, are both fairly insensitive to changes in the value of σ_A. The other three techniques are much more sensitive to changes in σ_A. This higher level of sensitivity is due to the fact these techniques use the entire range of the wide distribution; therefore, as its width increases, the standard deviation of the resulting assembly also increases.

The robustness of the various techniques to the amount of scrap generated is just the opposite. The techniques that use the majority of both distributions by binning the entire range of values, EW:MS, EA:MS, and OWB, are extremely robust to changes in σ_A. This robustness is due to the fact that these techniques use most of the parts in both distributions the best that they can. The scrap fraction generated by the two techniques that attempt to minimize variation are highly sensitive to the value of σ_A since these techniques truncate the wide distribution to maintain a low assembly standard deviation.

The example presented in this section demonstrates the considerations that need to be made when choosing to use selective assembly and making choices related to the technique and number of bins to implement. The same procedure can be used to analyze any situation where two normal distributions are being combined.

4.6.2 Selective Assembly of Two Skewed Distributions

Sometimes manufactured parts are not normally distributed. In reality, the parts that require assembly could form an infinite number of general, irregular distributions. Up until this point, the focus has been on parts that are normally distributed. The symmetry of normal distributions allows for any of the binning techniques in Section 4.2 to be used with some success; however, most of the techniques result in unacceptable levels of scrap when the distributions being combined are not normal.

There are logical reasons why most of the techniques do not produce positive results when applied to any two general distributions. The Equal Width techniques bin the distributions regardless of their shapes. Imagine an irregular distribution with several "humps" of dense regions. These regions could be split in such a way that an enormous amount of scrap would result. The Equal Area to Minimize Variation approach could result in the same failure mode if the tight distribution was symmetric and the wide distribution was highly irregular. The optimal binning technique would be possible if the distributions were known in advance and did not vary much over time; however, considerable computation is required to compute the optimal bin limits for any two general distributions.

The one technique that does not suffer from extremely high scrap rates and is easy to implement for any two general distributions is the Equal Area to Minimize Scrap approach. This technique, as described earlier, bins each distribution independently and in a manner that puts an equal amount of area under the probability density function in each bin. When applying this approach to two general distributions, the requirement to have an equal number of parts is forced upon the binning, whereas, in the case of two normal distributions being binned, the equation for the normal probability density function was used to set the bin limits. The forcing of an exact, equal number of parts in each bin results in zero scrap. Since the distributions being discussed here are not normal, using the standard deviation as a parameter of the spread may be misleading; therefore, in the following examples, histograms are shown to represent the distributions being combined and the resulting combination.

When the component distributions are non-normal, the successful reduction of an assembly's variance is highly dependent on the exact shapes of the two distributions. Also, the number of bins to use is a factor of the exact distribution shapes. The two examples below demonstrate how the simulation of a selective assembly process can help determine how many bins to use and the effects on the assembly's distribution when selective assembly is implemented on parts with arbitrary distributions.

An example is presented here to demonstrate the effect selective assembly has when combining parts that come from two non-normal distributions. Part "A" has a distribution that is skewed to the high side. The distribution of Part "B" is skewed to the low side. The histograms for these two distributions are seen below in Figure 4-23. Each histogram represents one thousand simulated values.

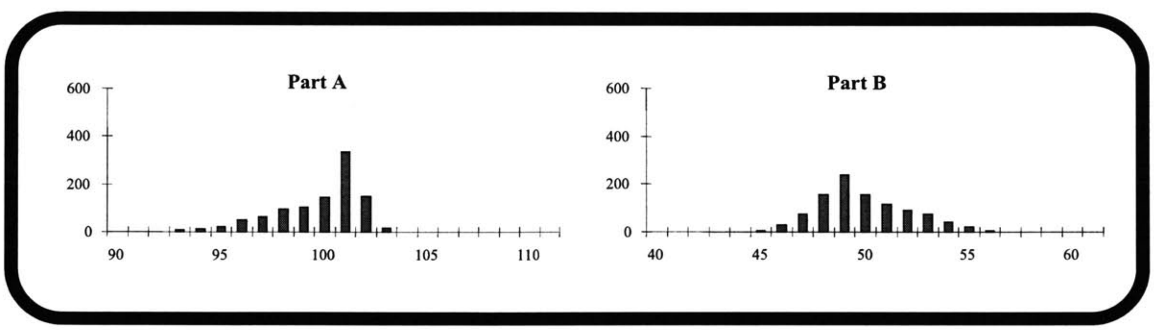

Figure 4-23: Two skewed distributions to be combined

The two distributions shown above in Figure 4-23 are combined using selective assembly via the Equal Area for Minimizing Scrap approach discussed above. The resulting assembly dimension histograms for the different numbers of bins are shown below in Figure 4-24.

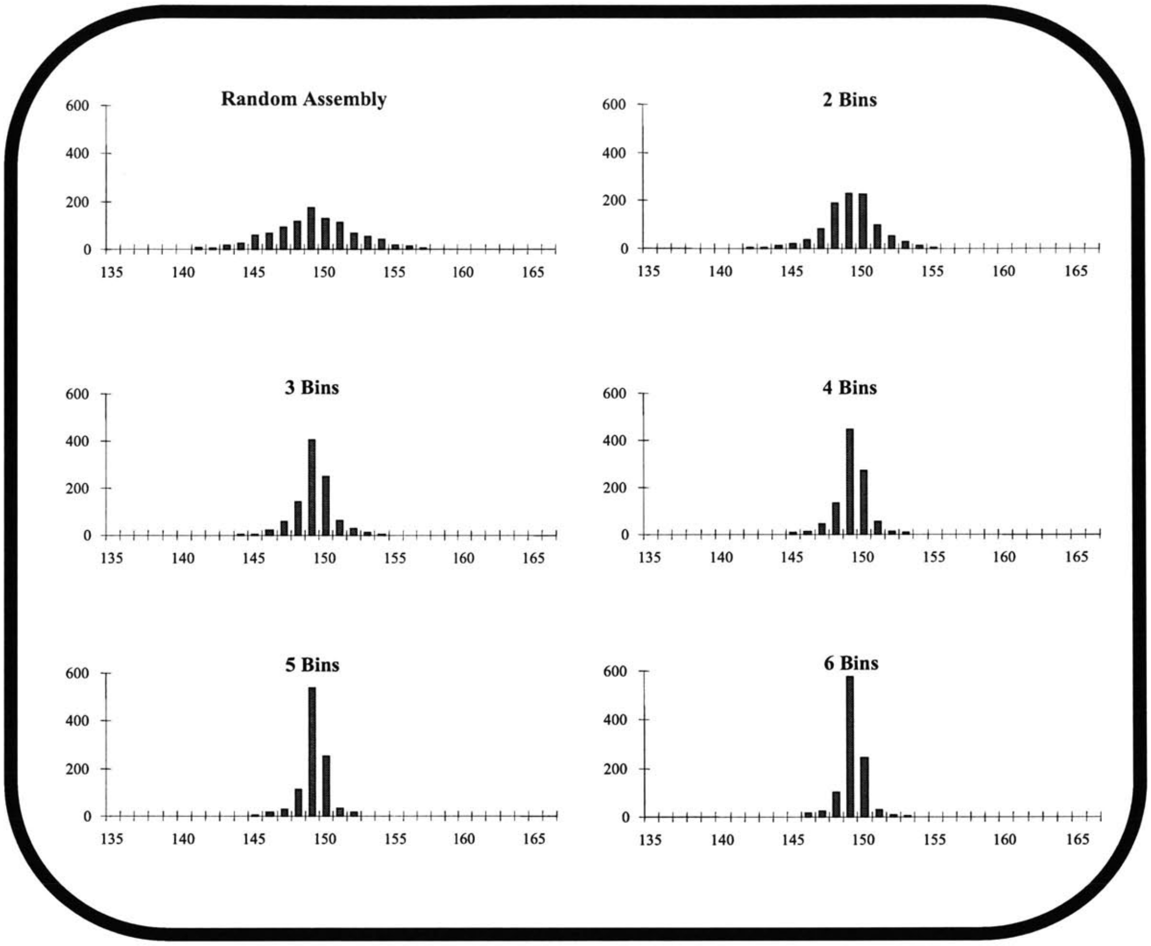

Figure 4-24: Histograms of assembly dimensions for different numbers of bins

Selective assembly does help reduce the spread and variation present in the assembly dimensions. An interesting observation is the fact the distribution of the assembly dimension does resemble a normal distribution regardless of the fact the two component distributions are non-normal. The reason this is true is because the two component distributions are skewed in opposite directions, and the components are added (see Figure 4-1) together to create the assembled dimension, i.e., components from the long left tail of the Part "A" distribution are added to components from the long right tail of the Part "B" distribution.

The increase in the number of bins does reduce the variation of the assembly dimension's distribution. Since the assembly dimension does form normal distributions, using the standard deviation is a valid parameter of the spread of the resulting assembly dimension. A plot of the assembly standard deviation versus the number of bins used is shown below in Figure 4-25.

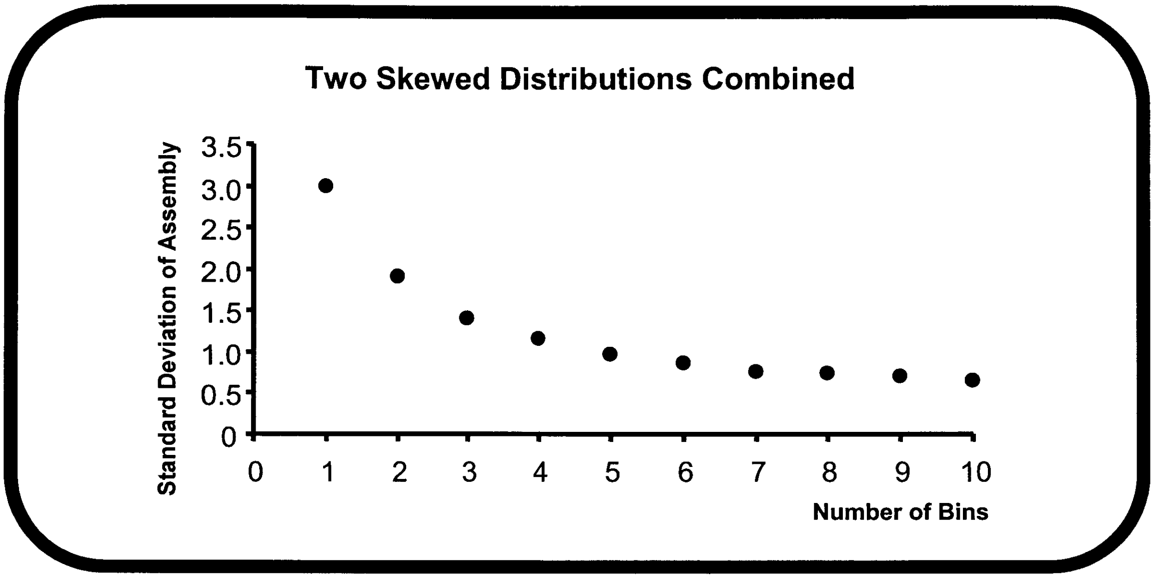

Figure 4-25: Assembly standard deviation versus number of bins

The plot in Figure 4-25 shows that the standard deviation for the assembly in this example can be greatly reduced through the use of selective assembly with an increased number of bins. In practice, if the component distributions resembled the ones in this example, the tradeoff between the number of bins (reduced variation) would have to be made against the costs of scrap, inventory, etc.

4.6.3 Selective Assembly of a Skewed and an Irregular Distribution

An example is presented here to demonstrate the effect selective assembly has when combining parts that come from two non-normal distributions. Part "A" has a distribution that is skewed to the high side. The distribution of Part "B" is non-symmetrical and has two regions of higher density. The histograms for these two distributions are seen below in Figure 4-26. Each histogram represents one thousand simulated values.

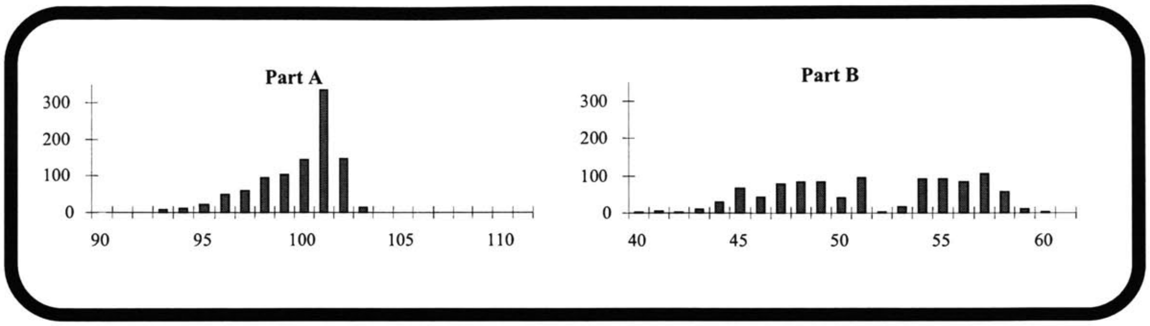

Figure 4-26: Skewed distribution and irregular distribution

The two distributions shown above in Figure 4-26 were combined using selective assembly via the Equal Area for Minimizing Scrap approach mentioned above. The resulting assembly dimension histograms for the different numbers of bins are shown below in Figure 4-27.

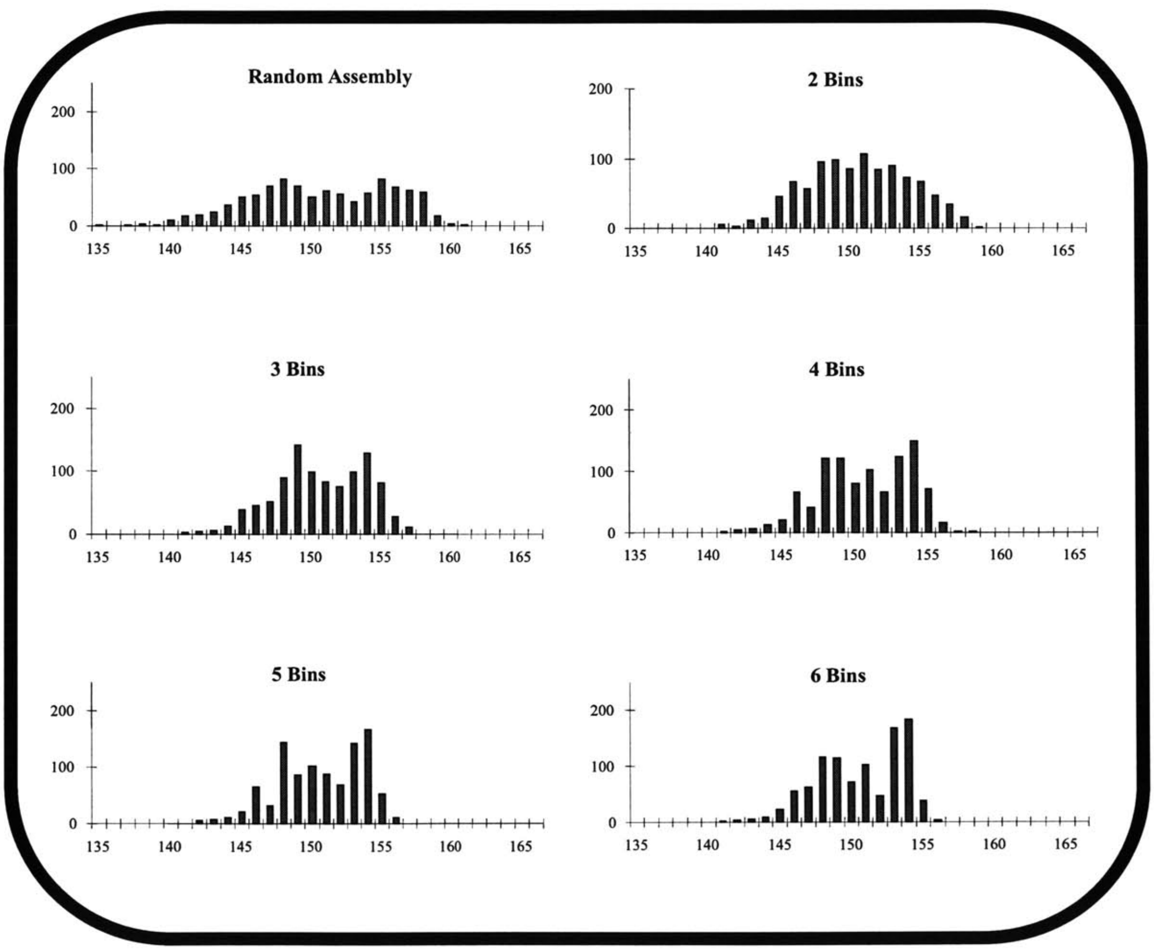

Figure 4-27: Histograms of assembly dimensions for different numbers of bins

Selective assembly does help reduce the spread and variation present in the assembly dimensions. The range of the assembly dimension when random assembly is used is 132 mm to 161 mm. When selective assembly is used with four bins, this range is reduced to 141 mm to 158 mm. This is a 41% reduction in the spread. The histograms also show a tighter, more centrally located distribution for the assembly dimension when selective assembly is used. The benefit of using more than four bins seems to be minimal. This is a result of the specific shapes of the parts' distributions.

4.7 Chapter Summary

The main focus of this chapter was to demystify selective assembly so that it can be effectively and intelligently implemented to manage product variation. The challenges associated with accomplishing this, along with possible solutions, are summarized below. The solutions that are original and introduced for the first time in this thesis are identified by the ❋**NEW!**❋ symbol. The section of the thesis where each solution is mentioned or discussed is listed in the summary to link to more detail. Each solution has its advantages and disadvantages listed to provide a quick summary.

Challenge: Understanding different selective assembly techniques

Solution	Section	Advantages (+) & Disadvantages (-)
Traditional selective assembly techniques	4.2.1 4.2.3 4.2.5 4.2.6 4.2.7	+ Each technique has its own set of advantages (See the corresponding section) – Each technique has its own set of disadvantages (See the corresponding section)
❋**NEW!**❋ Selective assembly techniques to minimize scrap	4.2.2 4.2.4	+ Provides a reduction in assembly variation while minimizing the cost impact of using selective assembly – May not provide the same reduction in assembly variation as other traditional methods

Challenge: **Analyzing tradeoffs between different selective assembly techniques**

Solution	Section	Advantages (+) & Disadvantages (-)
Traditional "goal post" analysis, i.e., parts are good if they meet the tolerance specified	4.1.2 4.3.1	+ Simple because bin limits can guarantee specified tolerances are met – Does not consider quality loss as variation increases
❋**NEW!**❋ Quality vs. Costs considerations	4.3	+ Provides a method to make tradeoffs between cost and quality – Used to make best choice for the selective assembly technique to implement

Challenge: **Calculating the standard deviation of a dimension resulting from a selective assembly operation**

Solution	Section	Advantages (+) & Disadvantages (-)
Numerical simulation	4.4	+ Can be used to compute the standard deviation of the assembly dimension regardless of the component distributions, e.g., the two component distributions do not necessarily need to be normally distributed – Computationally expensive – Time consuming to perform tradeoff analysis
❋**NEW!**❋ Closed-form equations	4.4.1	+ Easy to use + Can be implemented in a wide variety of computer applications, e.g., spreadsheets, math programs, etc. – Only applicable to a certain type of component distribution, e.g., in this thesis, the distributions being combined must be normal

Challenge: Developing closed-form equations for calculating the standard deviation of a dimension as it exits a selective assembly operation

Solution	Section	Advantages (+) & Disadvantages (-)
Analytical derivation	4.4	+ Produces exact solutions – Not possible to analytically derive useful form of equations, e.g., integrals appear that cannot be solved without numerical methods
❋NEW!❋ Use numerical simulation to generate results, then curve fit results to produce closed-form equations	4.4.1	+ Straightforward approach + Curve fit equations match simulated data very closely (See Section 4.5) – Equations only valid over specified range ± Initially seems like an infinite number of possible combinations exist for the two distributions' variances being combined, i.e., an infinite number of equations would have to be created so that there would be one for each combination. However, this is overcome by normalizing the equations (See 4.4.1)

Challenge: Performing sensitivity analysis on selective assembly operations

Solution	Section	Advantages (+) & Disadvantages (-)
Use numerical simulation	n/a	+ Can be used to compute the standard deviation of the assembly dimension regardless of the component distributions, e.g., the two component distributions do not necessarily need to be normally distributed – Computationally expensive – Time consuming to perform tradeoff analysis
❋NEW!❋ Use closed-form equations	4.4.2	+ Directly compute sensitivities by taking partial derivatives – Only applicable to a certain type of component distribution, e.g., in this thesis, the distributions being combined must be normal

Random assembly and selective assembly affect the variation of a product's dimensions. This chapter contained a brief review of random assembly followed by a detailed discussion of different selective assembly techniques. Two new techniques were introduced, Equal Width Binning to Minimize Scrap and Equal Area Binning to Minimize Scrap. An algorithm was presented to calculate the standard deviation and scrap generated as a result of a selective assembly operation. This algorithm was used to develop closed-form solutions for these two quantities. Several examples were shown to demonstrate the steps taken to make decisions regarding implementing a selective assembly operation.

5 System Approach: VRUM

Chapter Highlights

- Introduce *Variation Risk Management & Understanding* (*VRUM*) system model
- Define icons for system model diagrams
- Describe system model composition
- Formulate system model equations
- Discuss implementation of VRUM
- Present an example to demonstrate technique

5.1 Introduction

Variation Risk Understanding and Management (*VRUM*) is a design tool that uses process capability data to create mathematical equations that can calculate the expected mean and variance of a dimension on a manufactured part or assembly before being produced. The groundwork for VRUM has been created in the previous three chapters. Chapter 2 discussed the design of an efficient and effective Process Capability Database. The database is used to store and retrieve manufacturing data. The construction of the equations developed in Chapter 0 relies on these data. Once constructed, these equations calculate the effects a manufacturing process has on an incoming distribution. They accept an incoming distribution's mean and standard deviation as inputs to compute the expected mean and standard deviation of the outgoing distribution. Chapter 4 reviewed the formulation for computing how part variances add when random assembly is used to combine them. A closed-form equation for selective assembly was also developed that calculates the expected variance of an assembly's dimension when selective assembly is used to create it.

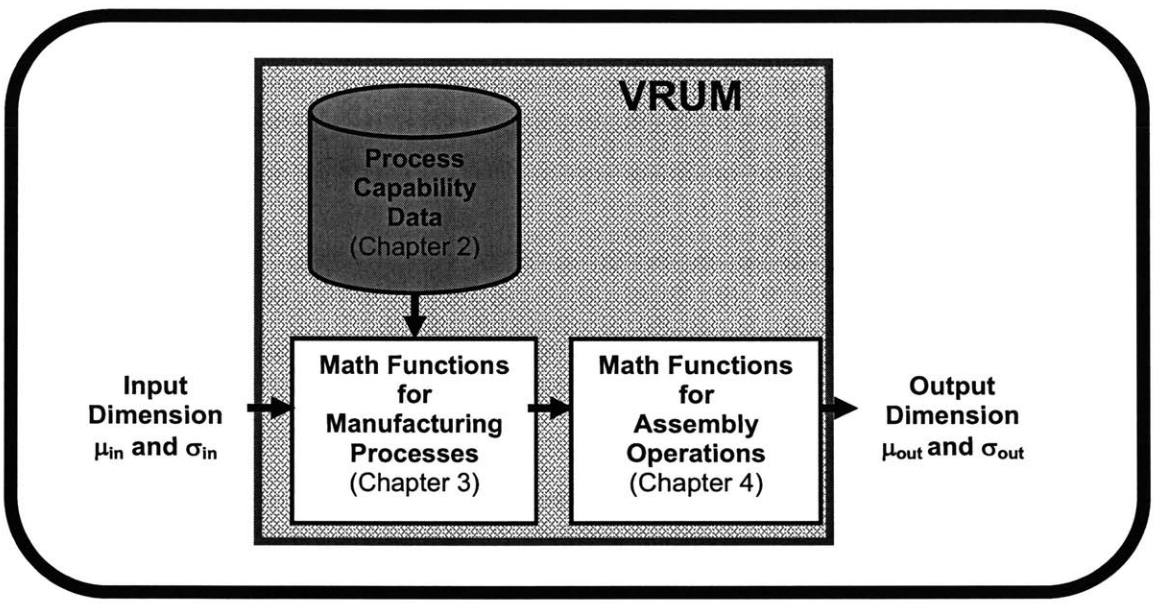

Figure 5-1: How the "pieces" fit

These developments are combined as shown schematically in Figure 5-1 to create a system level approach to forecast the mean and standard deviation of an assembly characteristic. The schematic shows how the previous developments fit together. The process capability data is used to construct math functions for manufacturing processes. These functions accept incoming dimensions' means and standard deviations and calculate the dimensions' means and standard deviations exiting the process. These values are then fed into the math functions representing assembly operations to compute the assembly dimension's mean and standard deviation. The large gray rectangle represents VRUM, i.e., the integration of the process capability database, math functions representing manufacturing processes, and math functions describing assembly operations. The schematic shows that VRUM accepts dimensions' μ_{in} and σ_{in} to compute μ_{out} and σ_{out}, thus, manufacturing variation can be forecast.

In this chapter, the method of how to compose a VRUM system level model and construct closed-form equations to compute output variation is described. Then, the role of VRUM in a product development process and the implementation of VRUM are discussed. Finally, an example is presented to demonstrate how VRUM can assist with design decisions.

5.2 System Model Composition

A VRUM system model is created to forecast the final mean and standard deviation of a characteristic on a part or assembly. The system model is comprised of manufacturing processes, assembly operations, process capability data, and part or assembly dimension distributions. A method for graphically and mathematically representing a VRUM system model is presented in this section.

5.2.1 Icons

Icons are used to graphically represent the elements in a VRUM system model. The icons are designed to be simple and easy to recognize. A small number of icons is used to keep the diagramming simple. A cylinder icon represents a Process Capability Database of the data stored within it. Different square icons denote different manufacturing operations. Manufacturing processes are symbolized with a square with an "X" in it since a process can act as a multiplier to an incoming distribution's variance. A square with a "+" sign signifies a random assembly operation since the variances of the parts being combined add. Selective assembly operations are designated by a square with a "-" because it works to reduce the assembly variance when compared to random assembly. A solid arrow represents a material flow while a dotted line represents information flow. These icons are shown below in Figure 5-2.

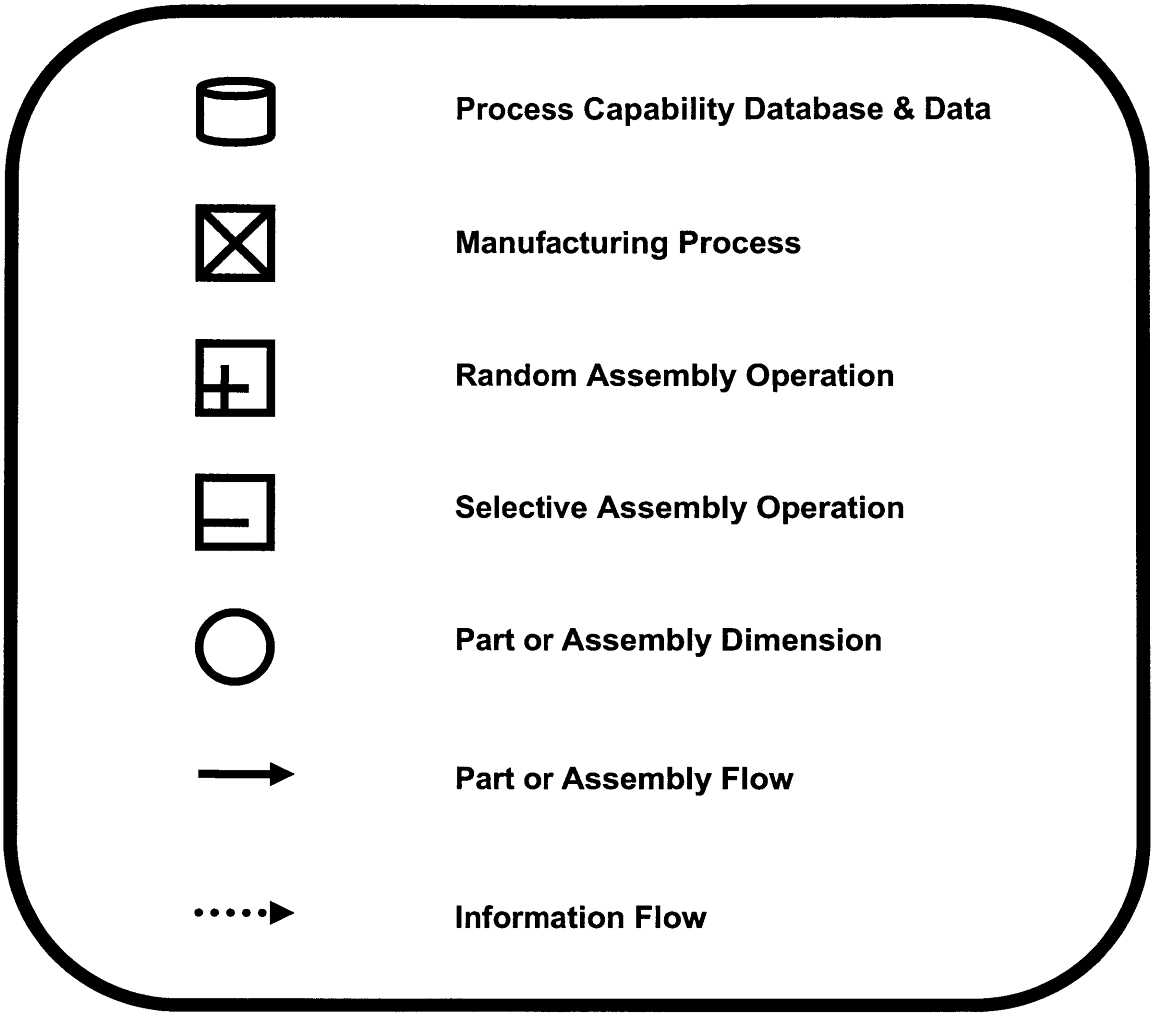

Figure 5-2: Icons used for VRUM system model diagrams

Individual processes do not have a corresponding unique icon because the number of existing processes is so large that remembering and identifying all the icons would become cumbersome. Instead the generic icons shown above are used and are simply labeled in the diagram. This approach keeps the diagram simple and effective for quick review.

5.2.2 System Diagram

A system diagram graphically represents a VRUM system model. The icons presented earlier are used to construct the system diagram. The diagram allows an engineer to quickly see the manufacturing processes and assembly operations that affect the characteristic of interest.

Constructing a diagram is simple. Circles are used as the starting point and ending point since they represent the part and assembly dimensions. Process and assembly icons are then placed into position. The Process Capability Database icon is then dropped into place. Finally, the solid arrows designating the part or assembly flow paths are used to connect the circles and squares, and the dotted lines are placed connecting the database to the processes for which data is used to build the corresponding equations.

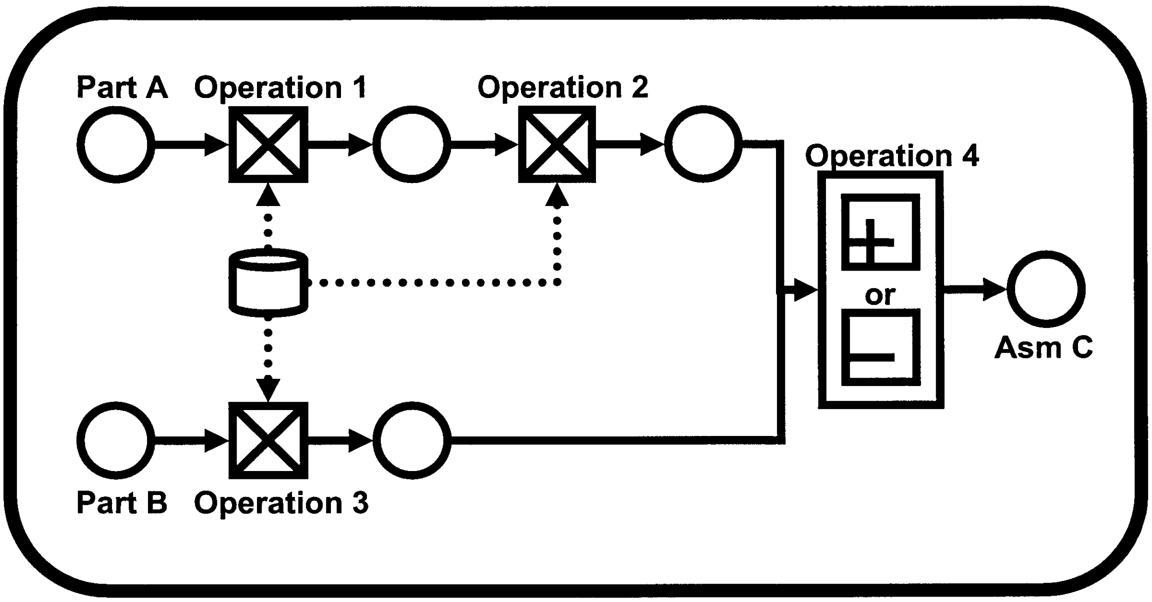

Figure 5-3: Example of a VRUM system model diagram

Figure 5-3 displays an example VRUM system model diagram used to forecast the mean and standard deviation of a characteristic of Assembly C. The assembly is made up of two components, Part A and Part B. The characteristic of Part A that affects Assembly C is operated on by two manufacturing processes. Only one manufacturing process affects Part B's characteristic of interest. Operation 4 is the operation that assembles Part A and Part B. At this point, the type of assembly operation, random or selective, has not been chosen so the diagram shows both.

5.2.3 Output Equations

The main advantage of VRUM is the ability to construct closed-form solutions for the mean and standard deviation of a dimension as it exits a manufacturing system. This is accomplished by associating equations with the operation icons in the VRUM system diagram. The equations used are those developed in Chapter 0, specifically Equation 3-25 for the mean and Equation 3-26 for the standard deviation of parts exiting a manufacturing process. The system does not have a single equation to represent it. Actually, each process has two associated equations, one for the output mean and one for the output standard deviation. The output from one process' equations or assembly operation is fed as an input to the equations for the next process downstream as illustrated in Figure 5-4. This cascading of information allows the final mean and standard deviation of a dimension as it exits a manufacturing system regardless of the number of processes or assembly operations in the production system. The example presented in Section 5.4 demonstrates this methodology.

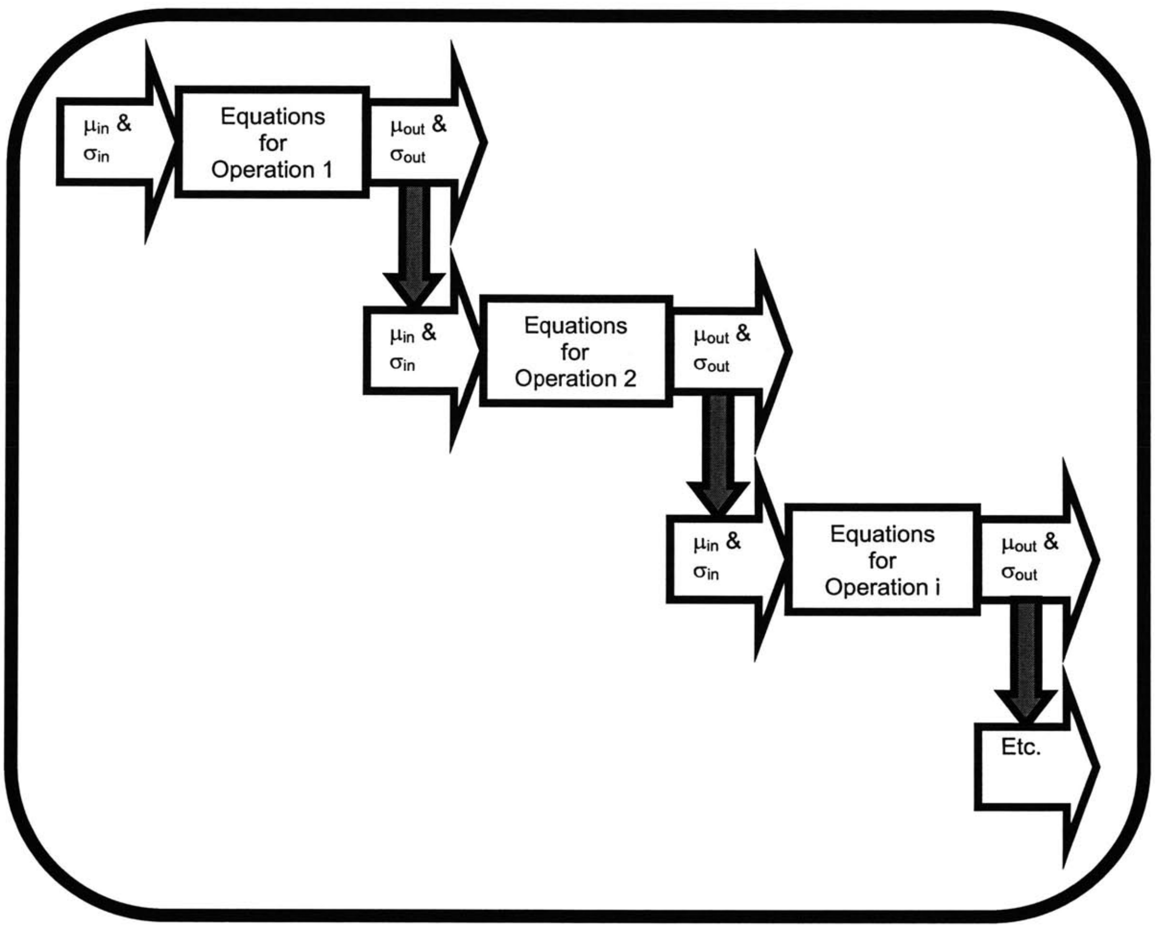

Figure 5-4: Cascading of information from one process to the next

The inputs and outputs for the individual operations can be examined to determine how variation propagates through the system. Once the operations that affect variation the most are identified, they can be modified or replaced to improve the system output. Examining the individual processes can also reveal if a downstream process "removes" variation from the parts going through it. This insight could allow the justification of using a less expensive operation with lower capability to be utilized upstream since the variance is removed by a later operation anyway.

5.3 Role of VRUM

Manufacturing companies are motivated to better understand and manage manufacturing variation to lower costs and improve quality. Managing manufacturing variation is the act of making design and manufacturing decisions so that a product can be manufactured with an acceptable level of variation at an acceptable cost. Managing variation does not always imply minimizing it. Sometimes the cost to minimize variation is not justified because there is no gain in functionality with the lower variation. Managing variation is finding a proper balance between manufacturing cost and variation in an intelligent manner.

VRUM assists in the management of variation. Variation can be managed through several decision making aspects of product development, which can be enhanced using VRUM. Supplier selection is enhanced using VRUM by comparing the quality of the products that different suppliers can produce. Process selection is enhanced because capable processes can be found and implemented using the process capability database and closed-form equations. The selective assembly closed-form equations are helpful in making tradeoffs between using more expensive manufacturing processes compared to using less expensive processes with selective assembly. VRUM helps during the design of a product by reassuring the design engineer that the product can be manufactured with an acceptable level of quality. Specifying tolerances on a product is assisted with VRUM by revealing how adjusting the tolerances affects the product quality. These decisions are further enhanced using VRUM in a feedback loop. The decisions are set and the output mean and standard deviation are computed. If these values do not meet specification, the decisions are re-evaluated, and the mean and standard deviation are re-computed. These steps are repeated until an acceptable mean and standard deviation are found. A schematic of these steps is shown in Figure 5-5 where the decisions are listed on the left and the large gray rectangle represents the VRUM system. The information flows are represented by solid black lines with arrow heads indicating the directions. Costs can also be considered to find the group of decisions that produce a product that meets specifications for the lowest cost.

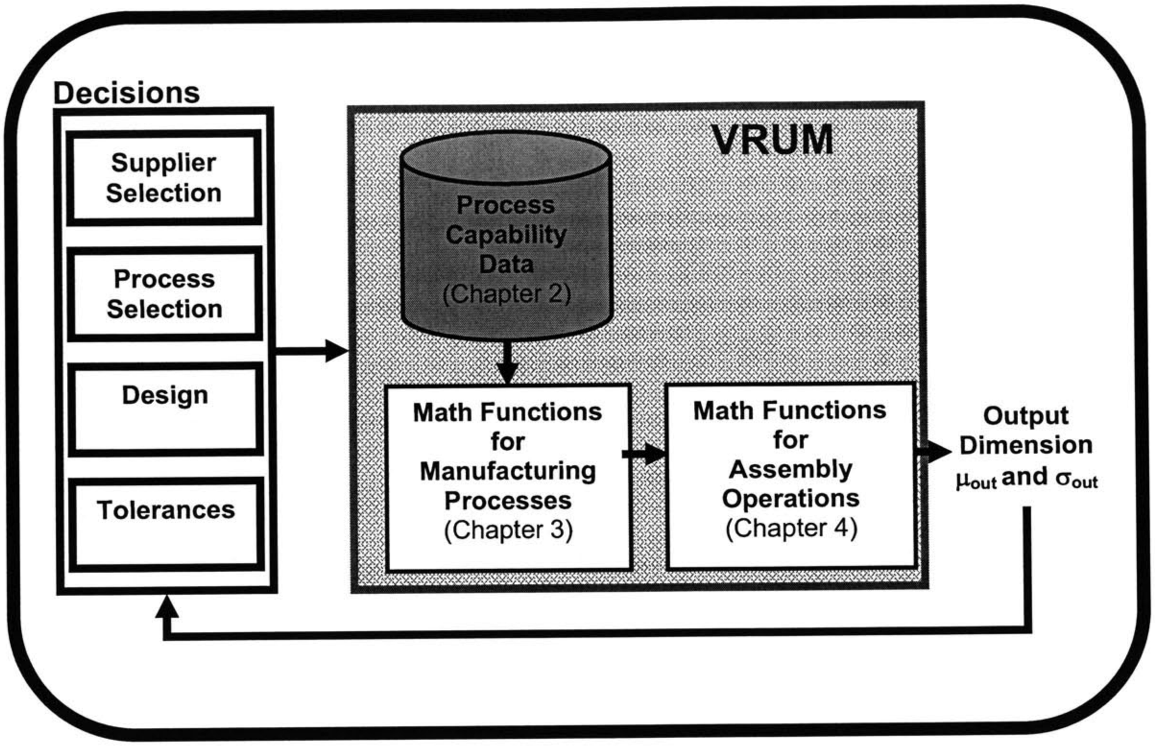

Figure 5-5: Decisions enhanced with VRUM

The exact role of VRUM depends on the product development phase when it is used. Figure 5-6 shows a generic four phase product development process consisting of planning, concept generation, detailed design, and production phases. (Details regarding product development processes can be found in *Product Design and Development* (Ulrich and Eppinger 2000).) The uses of VRUM are labeled along the timeline. During the early stages, VRUM is used to provide a general insight into the company's process capability. For example, an engineer could use VRUM to review capability data to get a sense as to the features and geometries that have been produced on previous products, the processes often used, and the materials that are commonly used. The engineer can then use this information to eliminate infeasible concepts. During the detailed design stage, VRUM guides an engineer in the design of a product that can be produced with an acceptable level of quality, e.g., using VRUM in a Design for Six Sigma environment. An example is when an engineer uses VRUM to forecast the final mean and standard deviation of a dimension and then makes decisions such as changing the overall magnitude of the dimension to make it easier to produce. VRUM also helps with tolerance allocation since it can be used to quantify the effect changing a tolerance has on the end quality. Selecting the best supplier is made easier using VRUM by forecasting the end product quality as a function of the supplier's process capability. For example, two suppliers can be quantitatively compared by using their capability data as inputs to VRUM equations. Process selection can be enhanced with VRUM by using the DeltaP and SigmaP functions to qualitatively compare processes and using the resulting closed-form equations to make quantitative comparisons. Finally, during production, VRUM helps identify where in the production system variation is introduced by examining intermediate values of the mean and standard deviation as it propagates through the system.

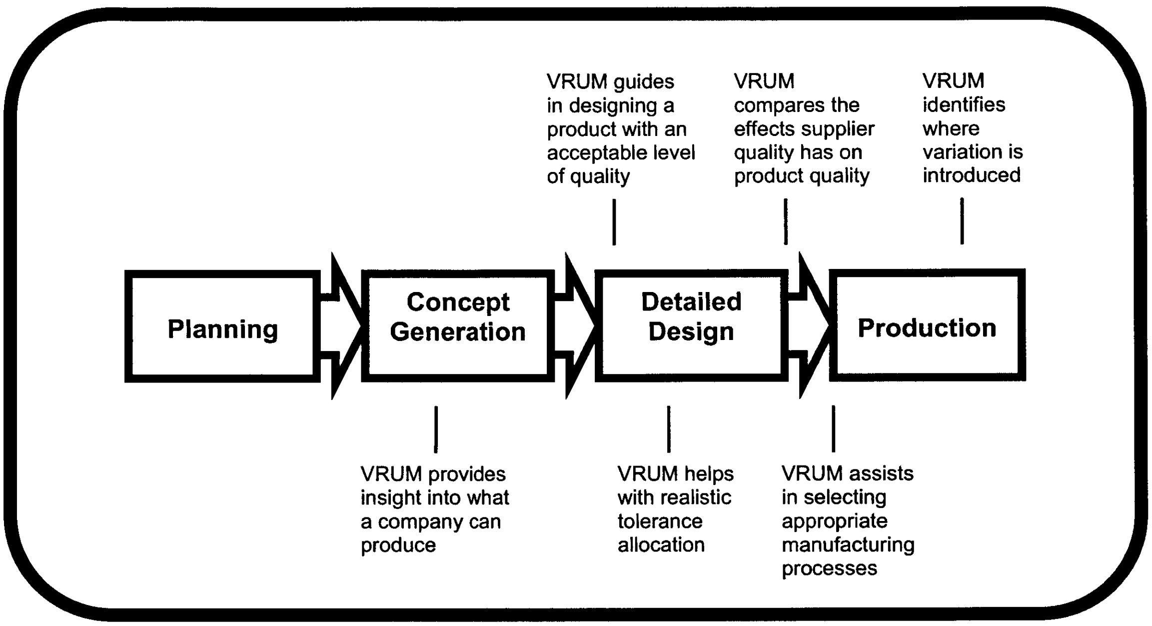

Figure 5-6: Generic four phase product development process

VRUM is solely focused on dimensional concerns and is not a process parameter optimizer. VRUM is not intended to be used to find the operating parameters of a particular process that produce high quality parts. VRUM is also unable to determine the sensitivity of a part's output quality to small deviations in an operating parameter, e.g., feed, rotational speed, etc. Instead, VRUM is used to compare different processes at different operating points, i.e., a fixed set of operating parameters. The main reason for this is that VRUM does not accept process parameters as inputs and none of the equations are functions of anything other than dimensional quantities. For example, VRUM does not help determine if a turning process will produce better parts if the angular velocity of the workpiece is changed slightly.

5.4 Implementation

Variation Risk Understanding and Management (VRUM) is a design tool that can be implemented in a wide variety of design and manufacturing environments ranging from small machine shops to international corporations. The foundation of VRUM is comprised of the three major components of a process capability database (Chapter 2), closed-form equations to compute the mean and standard deviation of a product dimension exiting a manufacturing process (Chapter 0), and closed-form equations to compute the variation of a product dimension upon exiting an assembly operation (Chapter 4). Each component could be implemented independently of the others; however, the usefulness is diminished. A process capability database implemented using the concepts presented in Chapter 2 allows for engineers to gain insight into a company's manufacturing capability and thus make better design decisions. The database also serves as a source for the data that is used to construct the closed-form equations for μ_{out} and σ_{out} exiting a manufacturing process. Once these equations are constructed, they can be used in series or as standalone equations. The closed-form equations for the variation produced by selective assembly can also be used alone. They are beneficial if a company is simply concerned with improving its product quality without changing any component production processes. The selective assembly equations are discussed and presented in Chapter 4.

The situation may arise where a company wants to use the closed-form equations for μ_{out} and σ_{out} exiting a manufacturing process without going through the effort of constructing, populating, and maintaining a process capability database. For example, a company purchases a few new pieces of equipment and wants to construct closed-form equations for how the new equipment affects the variation of parts they produce. In a situation like this, the equations can be produced by producing parts and using the measurements as process capability data to construct the equations per the instructions in Chapter 0. No hard-and-fast rules exist for how much data are required. Some guidelines are presented in Section 3.7 as to how much and what kind of data should be used to efficiently build equations that are accurate in forecasting quality. For example, using these guidelines, three groups of oversized stock could be processed followed by three groups of undersized stock to produce six sets of capability data with a reasonable amount of dispersion. These six groups of data are then used to construct the closed-form equations. Each group should contain enough parts to be statistically significant, e.g., thirty parts or more. Once the new equipment is used in production, parts produced on it can be measure and the data used to verify the accuracy of the equations.

Once VRUM is implemented, it is easy and quick to use. VRUM results in closed-form equations that can be solved using a simple spreadsheet. An engineer would follow the first few steps shown in Figure 5-7 through the step of constructing the DeltaP and SigmaP curves. At that point, the engineer would input the DeltaP and SigmaP coefficients into the spreadsheet and use the closed-form equations for the output mean and output standard deviations from a manufacturing process, Equations 3-25 and 3-26, along with the closed-form equation for selective assembly, Equation 4-21. Solutions are then calculated almost instantly for as many different cases as the engineer inputs. Finally, the engineer performs the last step of selecting and analyzing the best solution. Once the engineer has a spreadsheet setup, the entire cycle of computing new solutions only takes a matter of a few minutes.

5.5 Example Problem

A simple example problem is presented here to demonstrate how VRUM can assist with design decisions. The example problem is told through the eyes of a fictional engineer. The data used in the example problem were created using the physics-based math models introduced in Section 3.6.1. An outline of the steps used in the example problem is displayed in Figure 5-7.

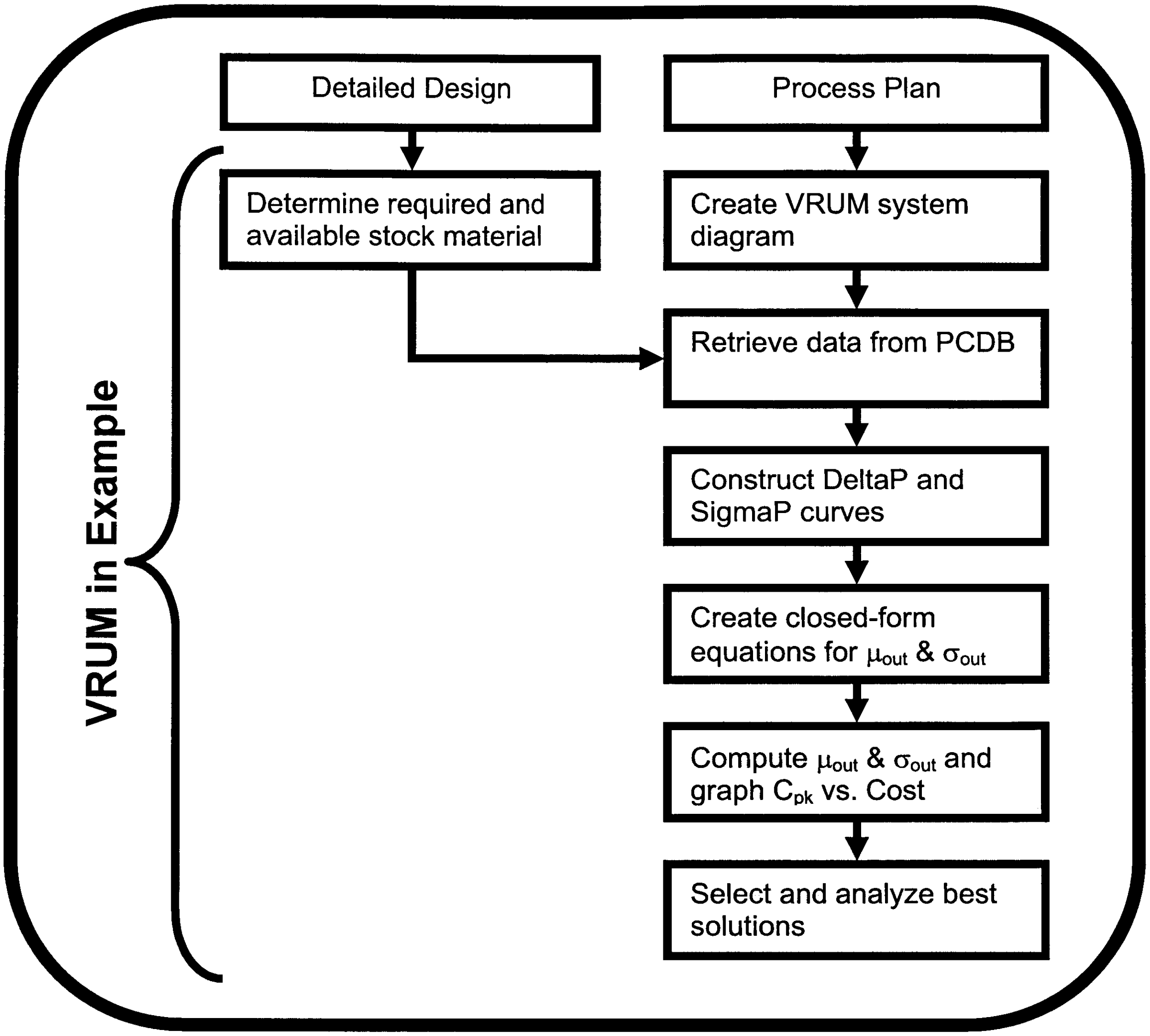

Figure 5-7: Schematic of steps used in example problem

Cyler is a design engineer working for a company that manufactures widgets. One type of widget is comprised of an aluminum cylindrical shaft mating with an aluminum plate that has a round bore. Figure 5-8 shows an illustration of the widget being designed. The gap between the two components is used for a cooling channel. Having the appropriate cooling fluid path requires a diametrical clearance of 1.00mm. Variation of the clearance between the mating parts can cause too little or too much fluid to pass by the parts, which causes early failure or degraded performance of the widget; therefore, the clearance is determined to be a critical dimension and must be held to tight tolerances, see Figure 5-9. Based on functional requirements, the tolerance range for the diametrical clearance is set to 0.99mm to 1.01mm. The target diameter of the shaft is 20.0mm, and the target hole diameter is 21.0mm. Cyler's company is striving toward a high level of quality. The company uses the process capability index of C_{pk} to measure the level of quality of a dimension. Cyler would like to produce widgets that have a clearance with a C_{pk} greater than two. Cyler assumes that the parts' dimensions are normally distributed and that the processes he has available to him produce normally distributed parts.

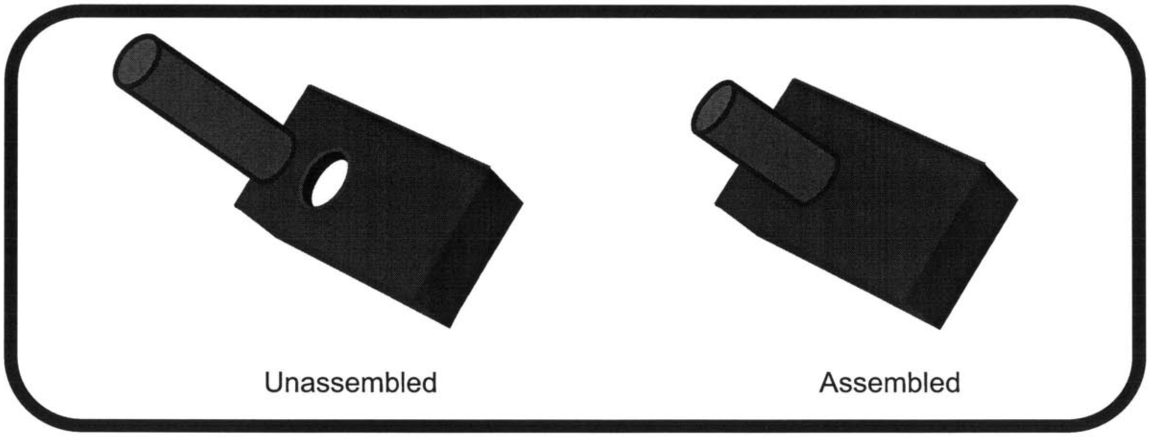

Figure 5-8: Illustration of widget being designed

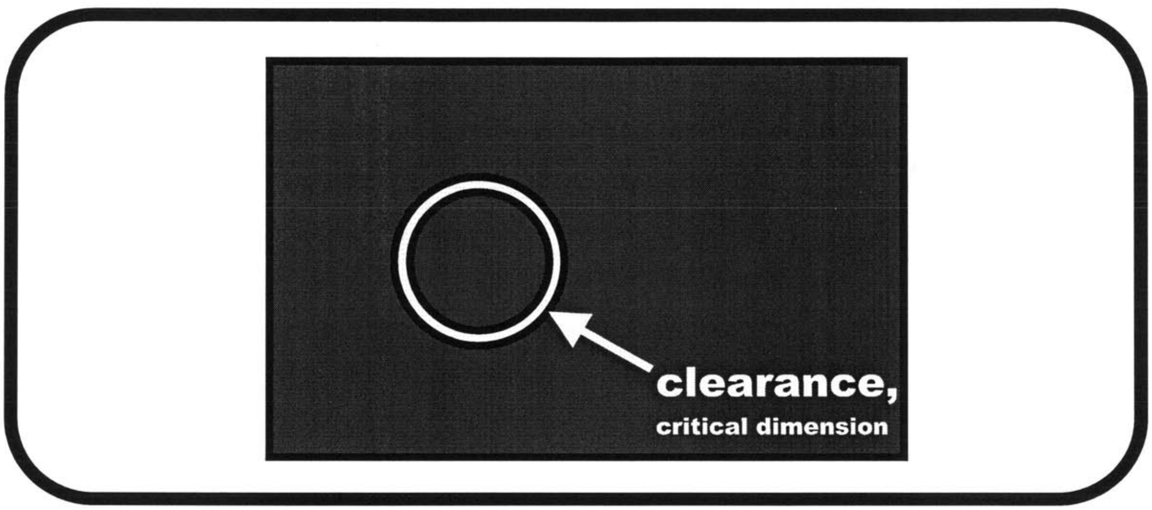

Figure 5-9: Top view of widget showing the critical dimension

Cyler is going to use aluminum bar stock that is cut to the correct length. He has three different potential suppliers. The suppliers' bar stock is specified by a mean and standard deviation. The summary of the supplier capability data is seen in Table 5-1. Cyler uses aluminum blocks machined to the proper dimension for the plate component of the widget. The aluminum blocks to be used for the plate component can be procured for $0.20 per piece.

Table 5-1: Supplier data for bar stock

Supplier	Diameter Mean [mm]	Standard Deviation [mm]	Piece Cost [$]
A	23.0	0.30	0.19
B	24.0	0.12	0.24
C	25.4	0.70	0.14

The bar stock obviously needs to be machined to the proper dimension. Cyler has three lathes at his disposal for the task. Cyler knows that to even get close to the required tolerance and level of quality, a roughing cut followed by a finishing cut is required. His plan is to use a roughing operation to reduce the bar stock to a diameter of approximately 20.5mm. The finishing operation then removes material to produce diameters equal to the final desired diameter of 20.0mm. He also has three drill presses available. He hopes to simply drill the hole with one pass. Cyler takes this information and constructs the VRUM system diagram shown in Figure 5-10. The VRUM diagram shows the bar stock is sent into a rough turning operation that is followed by a finish turning operation. The output from the finishing operation meets the drilled plate stock at the assembly operation where the components are mated together to create the clearance.

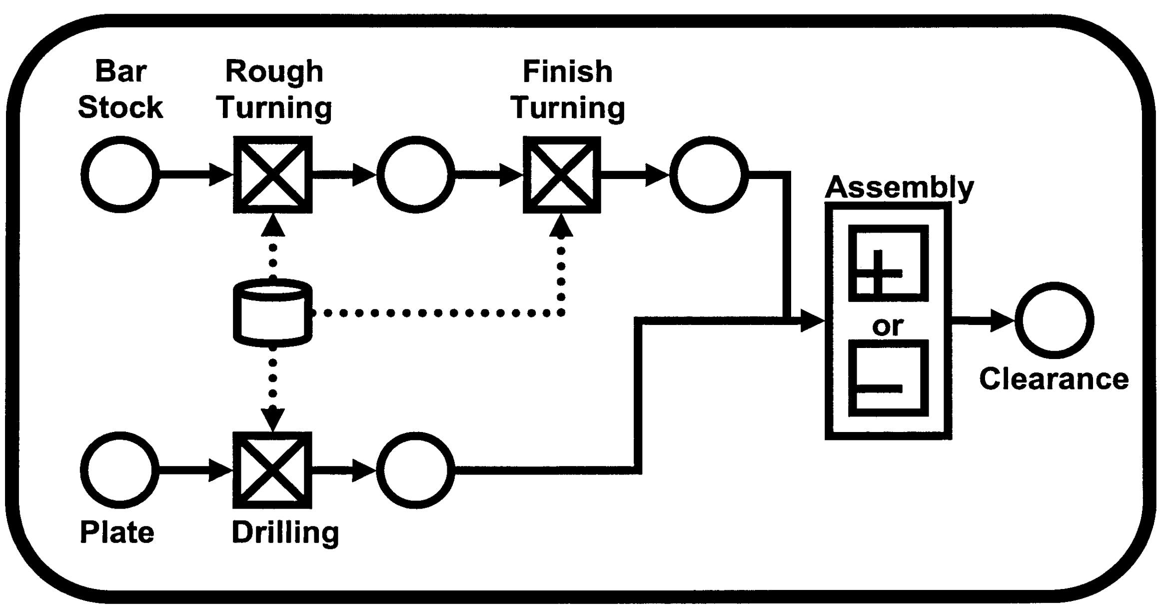

Figure 5-10: VRUM system diagram for example problem

Cyler queries his company's process capability database to retrieve data for the roughing, finishing, and drilling operations when the workpiece material is aluminum. The retrieved data are listed in tabular form in Appendix J. Cyler knows that these data represent normal distributions. The retrieved data also list the cost associated with each process. The costs are summarized below in Table 5-2. These costs are used to calculate the total cost of each assembly and help in determining the lowest cost combination of raw stock and processes that result in assemblies that meet the quality specification.

Table 5-2: Cost for each available processes

Process	Cost per Piece [$]
Roughing 1	0.21
Roughing 2	0.15
Roughing 3	0.11
Finishing 1	0.45
Finishing 2	0.27
Finishing 3	0.20
Drilling 1	0.47
Drilling 2	0.13
Drilling 3	0.33

The retrieved process capability data are used to calculate the coefficients for the DeltaP and SigmaP functions for each of the roughing and finishing operations. The DeltaP and SigmaP functions are assumed to be second order polynomials as expressed in Equations 5-1 and 5-2. The technique used to calculate the coefficients is outlined in Section 3.4. The coefficients are listed below in Table 5-3.

$$DeltaP = a_0 + a_1 M + a_2 M^2 \tag{5-1}$$

$$SigmaP = b_0 + b_1 M + b_2 M^2 \tag{5-2}$$

where the coefficients are listed below for the different available processes.

Table 5-3: Coefficients for DeltaP and SigmaP functions

	DeltaP Coefficients			SigmaP Coefficients		
Process	a_0	a_1	a_2	b_0	b_1	b_2
Roughing 1	0	0.988121	0.004126	0	-0.004485	0.002819
Roughing 2	0	0.981478	0.007011	0	-0.011743	0.003757
Roughing 3	0	0.985607	0.004973	0	-0.008009	0.002910
Finishing 1	0	0.998914	0.010450	0	-0.000468	0.006219
Finishing 2	0	0.998508	0.015554	0	-0.000247	0.009706
Finishing 3	0	0.998890	0.012274	0	-0.000454	0.007306

The coefficients listed in Table 5-3 are used to plot the DeltaP and SigmaP functions for the roughing and finishing processes. These graphs are shown in Figure 5-11 through Figure 5-14. The graphs allow for a quick qualitative comparison of the different processes. The graph in Figure 5-11 shows that the three roughing processes are close to the ideal case of the actual depth of cut being equal to the desired depth of cut, which would be represented by a line with a forty five degree slope. The graphs do reveal that the Roughing 2 operation produces parts that are not as close to the desired output value since it diverges from a forty five degree line more than the other two curves.

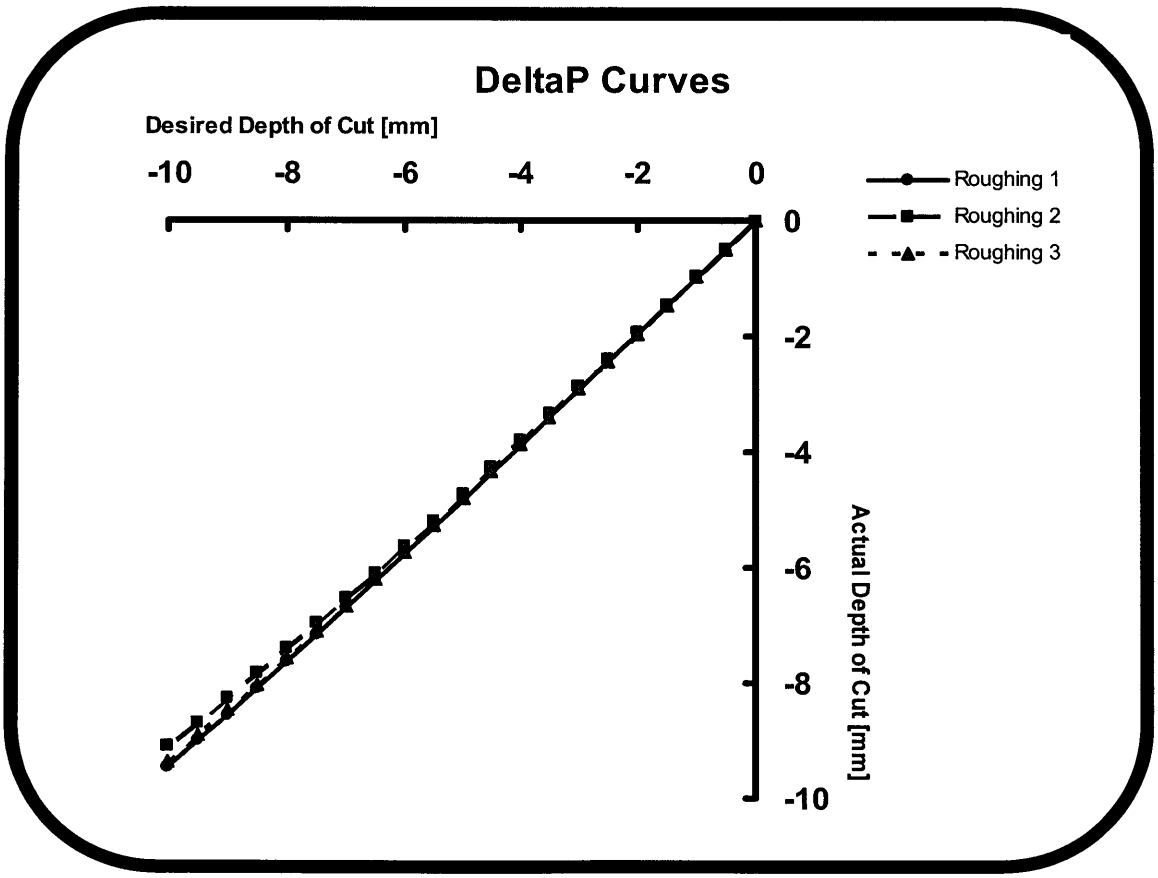

Figure 5-11: DeltaP curves for the available roughing operations

The SigmaP curves shown in Figure 5-12 are used to qualitatively compare the expected variation upon exiting the roughing process. The Roughing 2 curve lies above the other two curves; therefore, the Roughing 2 operation produces output parts that have higher variation than the other two processes. The roughing process that produces parts with the smallest amount of variation is the Roughing 1 operations since it lies below the other two curves.

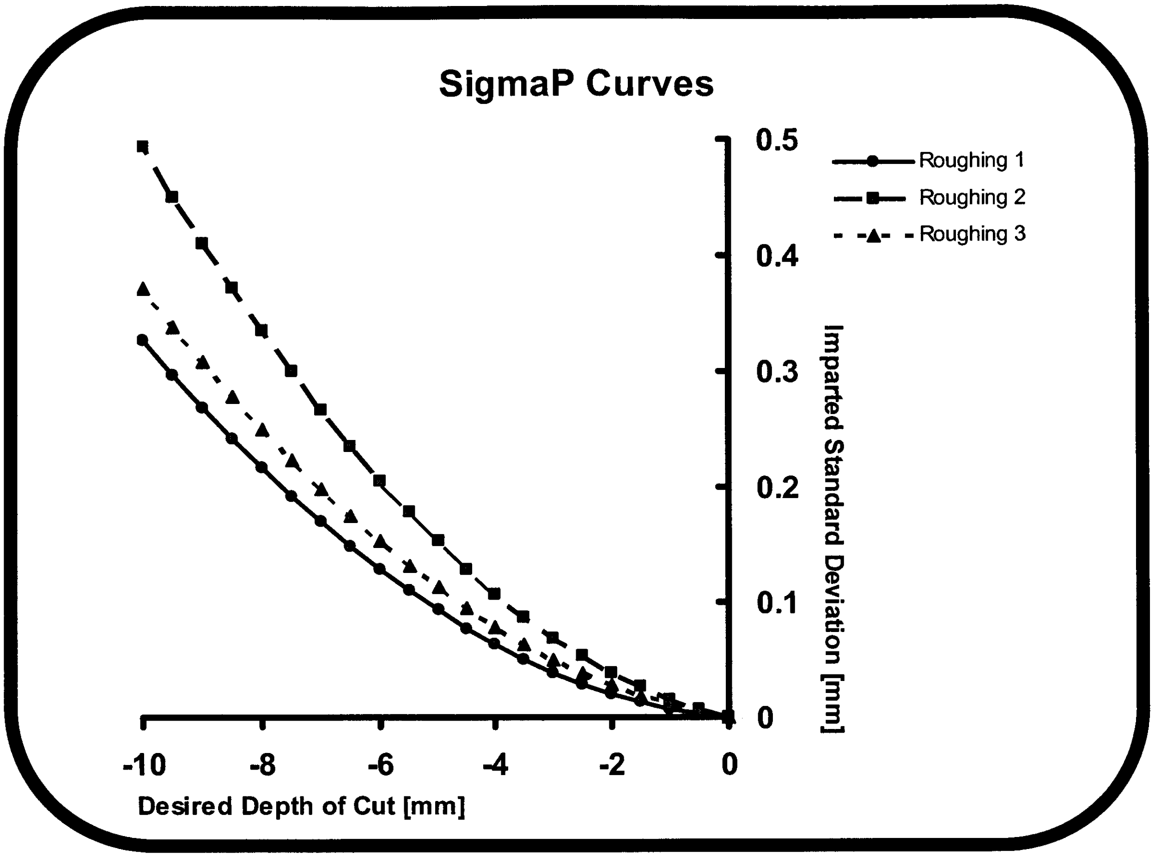

Figure 5-12: SigmaP curves for the available roughing operations

The DeltaP curves for the finishing operations are plotted in Figure 5-13 over the range of interest. The plots reveal that the three processes are almost identical in their abilities to remove a certain amount of material as a function of the desired amount of material removal. The SigmaP curves are plotted in Figure 5-14. These plots show that Finishing 1 produces the parts with the lowest amount of variation since it is below the other two curves. The parts produced with Finishing 2 have a higher level of variance than parts produced with the other two processes.

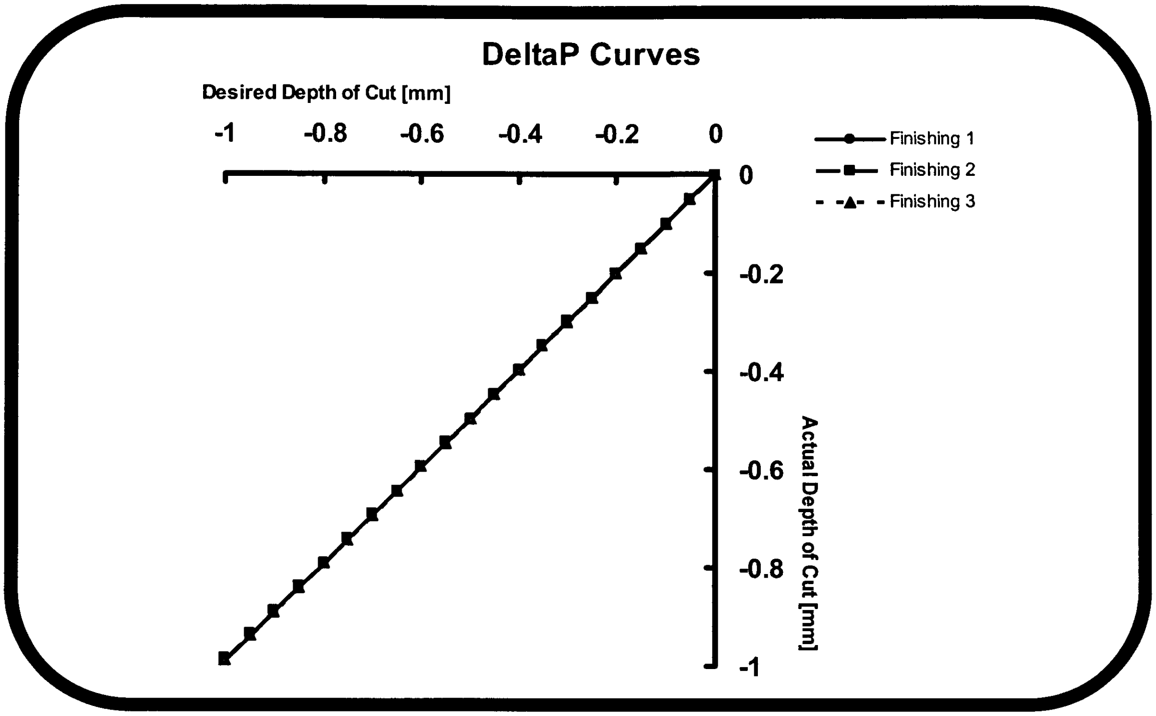

Figure 5-13: DeltaP curves for the available finishing operations

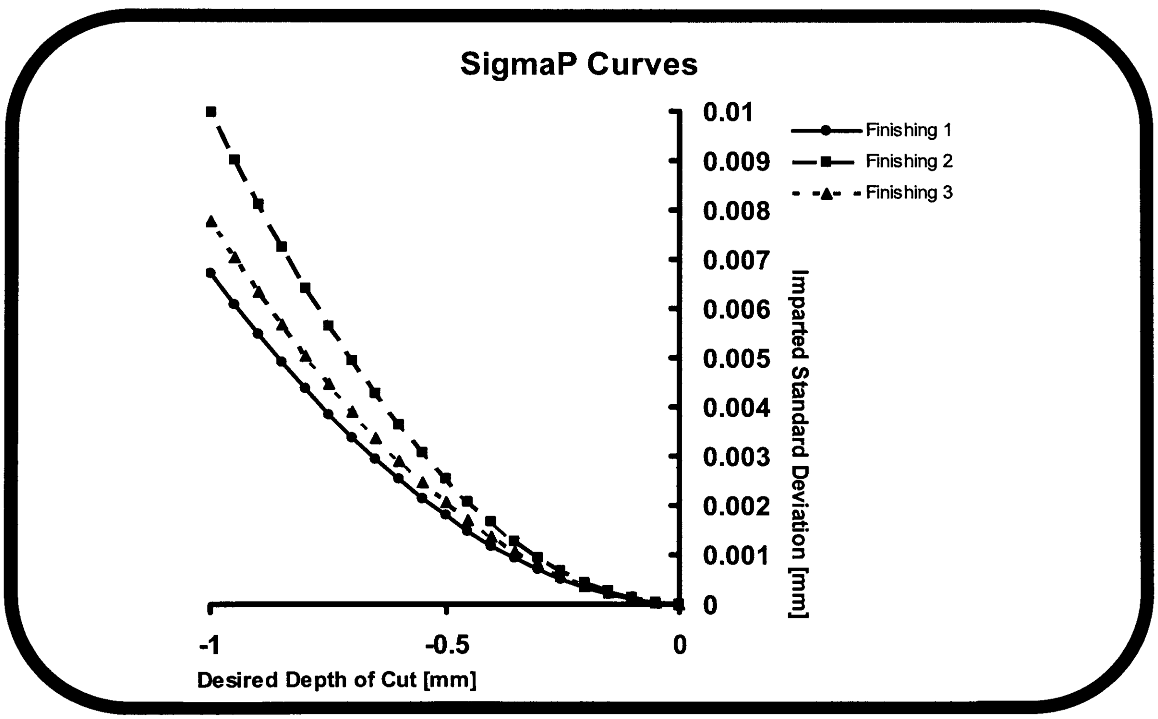

Figure 5-14: SigmaP curves for the available finishing operations

The plots of the DeltaP and SigmaP functions are a valuable resource for qualitative analysis. If all other aspects of the processes were equal, e.g., costs, cycle times, up time, etc., then the curves would be enough to make the decisions as to the best processes to use. However, other factors are different for the different processes. For this example, the only process aspect other than quality considered is the cost of producing one part. The values for the associated costs are given in Table 5-2.

Cyler must now select a supplier (see Table 5-1), a particular roughing operation, a particular finishing operation, a particular drilling operation, and an assembly operation to successfully produce the widgets with an acceptable level of quality. The roughing, finishing, and drilling operations he has available to him are listed in Table 5-2. For the assembly operation, he can choose to use random assembly or any of the five selective assembly binning techniques with any number of bins from one to ten. Cyler realizes a finite number of possible combinations of supplier, roughing operation, finishing operation, drilling, and assembly operations exist. The exact number can be computed using combinatorial math by multiplying the number of possibilities of each choice. Three possible suppliers are available. Three choices for the roughing operation, three choices for the finishing operation, and three choices for the drilling operation exist. Five selective assembly techniques can be used. Each technique can utilize between one and ten bins. Random assembly is also an option. Therefore, fifty one possibilities exist for the assembly operation. These numbers multiplied results in 4,131 possible combinations. Determining the best combination of choices is the responsibility of Cyler. A few sample combinations are listed below in Table 5-4.

Table 5-4: Samples of possible combinations Cyler can choose

Stock	Rough	Finish	Drill	Assembly
A	1	3	1	Random
A	2	2	1	EW:MV, N = 2
A	3	1	1	EW:MS, N = 7
B	1	3	2	OWB, N = 3
B	2	2	2	EA:MV, N = 6
B	3	1	2	EA:MS, N = 3
C	1	3	3	EW:MV, N = 3
. . .	. . .	. . .	. . .	. . .

The metric Cyler is using to represent the quality of the clearance is the process capability index known as C_{pk}. The equation below is used to compute C_{pk}.

$$C_{pk} = Min\left(\frac{\mu - LL}{3\sigma}, \frac{UL - \mu}{3\sigma}\right) \qquad 5\text{-}3$$

The symbol LL is used to represent the lower limit of the dimension of interest, and UL represents the upper limit of the dimension of interest. The mean and standard deviation of the dimension are represented by μ and σ respectively. The detailed design of the widget fixes the lower limit, LL, at 0.99mm and the upper limit, UL, at 1.01mm.

The mean and standard deviation of the clearance must be calculated to compute the C_{pk} values. To calculate these values for the clearance, these same values must be determined for the components. The mean and standard deviation for the hole diameter are determined by the drilling operation selected and are read from Table 5-5. The equation for the mean of the rod diameters exiting the roughing or finishing process is constructed using Equation 5-4, and the standard deviation of the rod diameters exiting the roughing or finishing process is constructed using Equation 5-5. The a_0, a_1, a_2, b_0, b_1, and b_2 coefficients are read from Table 5-3 for the specific process.

Table 5-5: Quality produced by available drilling operations

Drilling Operation	Mean of Holes Produced [mm]	Standard Deviation of Holes Produced [mm]
1	21.0	0.005
2	21.0	0.006
3	21.0	0.007

Notice that the same general equations, Equations 5-4 and 5-5, are used to calculate the mean and standard deviation for the diameters as they exit each process. Not only do the values of the a_i's and b_i's depend on the particular process (see Table 5-3), the values used for the μ_{input}, μ_{target}, and σ_{input} also depend on the process the equation is representing, i.e., the roughing operation accepts the bar stock as input while the finishing operation accepts the roughing operation's output. These differences are summarized in Table 5-6 and are demonstrated graphically in Figure 5-15 where the arrows represent the information that is being passed to and from operations and the rectangles represent the equations for each operation.

Table 5-6: Different inputs used for roughing and finishing equations

	Equation for Roughing Operation Outputs	Equation for Finishing Operation Outputs
μ_{input}	Raw bar stock mean diameter	Mean of diameters exiting the roughing operation
μ_{target}	20.5 mm	20.0 mm
σ_{input}	Standard deviation of raw bar stock	Standard deviation of diameters exiting the roughing operation

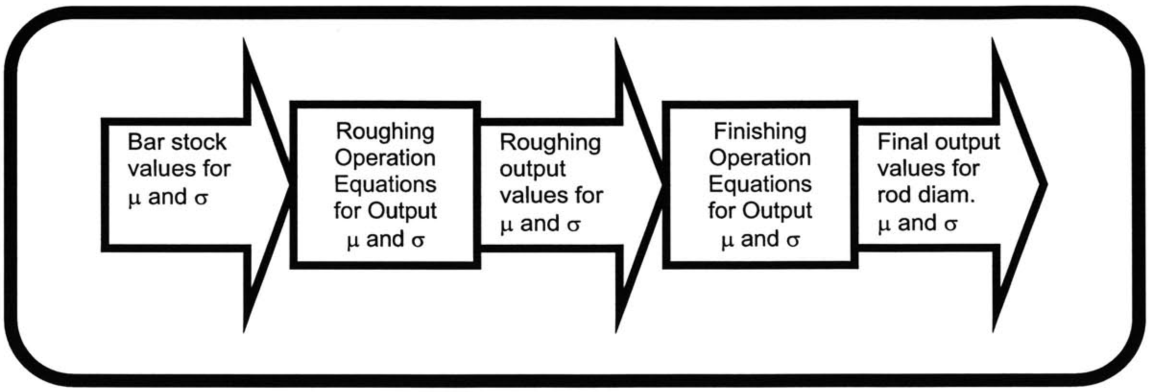

Figure 5-15: Illustration of different inputs for roughing and finishing

$$\mu_{output} = \mu_{input} + a_0 + a_1\left\{\mu_{target} - \mu_{input}\right\} + a_2\left\{\mu_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2 + \sigma_{input}^2\right\} \qquad 5\text{-}4$$

$$
\begin{aligned}
\sigma^2_{output} &= b_0^2 + b_1^2\left\{\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right\} \\
&-2b_1b_2\left[\mu_{input} - \mu_{t\arg et}\right]\left[\mu^2_{input} + 3\sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right] \\
&+2b_0\left\{b_1\left[\mu_{t\arg et} - \mu_{input}\right] + b_2\left[\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right]\right\} \\
&+b_2^2\left[\mu^4_{input} + 3\sigma^4_{input} - 4\mu^3_{input}\mu_{t\arg et} + 6\sigma^2_{input}\mu^2_{t\arg et} + \mu^4_{t\arg et} + 6\mu^2_{input}\left\{\sigma^2_{input} + \mu^2_{t\arg et}\right\}\right] \\
&+b_2^2\left[-4\mu_{input}\left\{3\sigma^2_{input}\mu_{t\arg et} + \mu^3_{t\arg et}\right\}\right] \\
&\sigma^2_{input} + \mu^2_{input} + 2\mu_{input}\left\{\mu_{output} - \mu_{input}\right\} \\
&-(\mu_{output})^2 + a_0^2 + a_1^2\left\{\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right\} \\
&+2a_0\left\{a_1\left[\mu_{t\arg et} - \mu_{input}\right] + a_2\left[\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right]\right\} \\
&-2a_1\left\{\sigma^2_{input} + a_2\left[\mu_{input} - \mu_{t\arg et}\right]\left[\mu^2_{input} + 3\sigma^2_{input} - 2\mu_{input}\mu_{t\arg et} + \mu^2_{t\arg et}\right]\right\} \\
&+4a_2\sigma^2_{input}\left[\mu_{input} - \mu_{t\arg et}\right] \\
&+a_2^2\left[\mu^4_{input} + 3\sigma^4_{input} - 4\mu^3_{input}\mu_{t\arg et} + 6\sigma^2_{input}\mu^2_{t\arg et} + \mu^4_{t\arg et} + 6\mu^2_{input}\left\{\sigma^2_{input} + \mu^2_{t\arg et}\right\}\right] \\
&a_2^2\left[-4\mu_{input}\left\{3\sigma^2_{input}\mu_{t\arg et} + \mu^3_{t\arg et}\right\}\right]
\end{aligned}
\tag{5-5}
$$

Cyler now has the closed-form equations for the mean and standard deviation of parts exiting the roughing and finishing operations. The equations are functions different variables as listed in

Table 5-6 and are written below for clarification.

$$\mu_{RoughingOutput} = f\left(\mu_{BarStock}, \mu_{RoughingT\arg et}, \sigma_{BarStock}\right) \tag{5-6}$$

$$\mu_{FinishingOutput} = f\left(\mu_{RoughingOutput}, \mu_{FinishingT\arg et}, \sigma_{RoughingOutput}\right) \tag{5-7}$$

$$\sigma^2_{RoughingOutput} = f\left(\mu_{BarStock}, \mu_{RoughingT\arg et}, \sigma_{BarStock}\right) \tag{5-8}$$

$$\sigma^2_{FinishingOutput} = f\left(\mu_{RoughingOutput}, \mu_{FinishingT\arg et}, \sigma_{RoughingOutput}\right) \tag{5-9}$$

This calculation demonstrates the capability of VRUM to quantify the effects serial processes have on a dimension. For example, the final mean and standard deviation of the rod diameter are easily computed even though two processes act on the bar stock. The capability arises from the fact that each process is represented independently by its own set of equations for its output mean and standard deviation. These output values are then cascaded as inputs into the next downstream process.

Now that Cyler can determine the components' means and standard deviations, he needs to be able to calculate the mean and standard deviation of the clearance. He uses Equation 5-10 to calculate the mean of the clearance where $\mu_{FinalRodDiameter}$ is the mean of the diameters exiting the Finishing operation and $\mu_{HoleDiameter}$ is the mean of the hole diameters resulting from the drilling operation. If he chooses to use random assembly, the equation for the clearance's variation level is shown as Equation 5-11. If Cyler opts to use selective assembly for the assembly operation, he uses Equation 5-12 to compute the standard deviation of the clearance. The values for the C_i coefficients are listed in Appendix C. The subscripts A and B are arbitrarily assigned to the hole diameter and rod diameter to produce a ratio of $1 \le \sigma_A/\sigma_B \le 2$. The subscript assignment can be swapped accordingly to cover the range of $0.5 \le \sigma_{HoleDiameter}/\sigma_{FinalRodDiameter} \le 2$. The three equations below were introduced and discussed in detail in Chapter 4.

$$\mu_{Clearance} = \mu_{HoleDiameter} - \mu_{FinalRodDiameter} \qquad \text{5-10}$$

$$\sigma_{Clearance,RandomAsm} = \sqrt{\sigma^2_{HoleDiameter} + \sigma^2_{FinalRodDiameter}} \qquad \text{5-11}$$

$$\sigma_{Clearance,SelAsm(N)} = \left((C_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (C_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (C_1)\left(\frac{\sigma_A}{\sigma_B}\right) + C_0 \right)\sqrt{\sigma_A^2 + \sigma_B^2} \qquad \text{5-12}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

Cyler now combines Equations 5-3 through 5-12 to calculate the C_{pk} for each combination of raw stock, roughing operation, finishing operation, drilling operation, and assembly operation. He still needs to compute the associated piece cost for each combination. The per piece cost is computed using Equation 5-13. The values for $C_{BarStock}$ are given in Table 5-1. The plate stock cost \$0.20 per piece. The cost for each operation is provided in Table 5-2. The inspection cost is zero if random assembly is used and is \$0.10 per piece if selective assembly is used. The part scrap cost is also zero if random assembly is used, i.e., all the parts will be used. The part scrap cost may be nonzero if selective assembly is used. The value of the part scrap cost is computed by computing the costs of the unmatched parts. This can be done as a percentage of the total number of parts. The percentage of scrap is given by Equation 5-14. This equation is used to construct the equation for the part scrap cost given by Equation 5-15. Again, the subscripts A and B are arbitrarily assigned to the hole diameter and rod diameter to produce a ratio of $1 \leq \sigma_A/\sigma_B \leq 2$.

$$\begin{aligned} C_{PerPiece} &= C_{BarStock} + C_{Plate} + C_{Roughing} \\ &\quad + C_{Finishing} + C_{Drilling} + C_{inspection} + C_{PartScrap} \end{aligned} \tag{5-13}$$

$$Scrap_{(N)} = \left((D_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (D_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (D_1)\left(\frac{\sigma_A}{\sigma_B}\right) + D_0 \right) \quad \text{for } 1 \leq \frac{\sigma_A}{\sigma_B} \leq 2 \tag{5-14}$$

$$C_{PartScrap} = \left((D_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (D_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (D_1)\left(\frac{\sigma_A}{\sigma_B}\right) + D_0 \right)\left(C_{BarStock} + C_{Plate}\right) \quad \text{for } 1 \leq \frac{\sigma_A}{\sigma_B} \leq 2 \tag{5-15}$$

Cyler now has all the equations needed to calculate the quality as measured by C_{pk} and cost of each combination. He plots the Cpk values versus the per piece cost as seen below in Figure 5-16. Each point on the graph represents a particular combination of raw stock, roughing operation, finishing operation, drilling operation, and assembly operation. The combinations that produced part distributions whose standard deviations were off by more than a factor of two are omitted from the graph. The squares on the graph represent combinations that use random assembly while the circles on the graph represent combinations that use selective assembly.

A Pareto Front can be created by connecting the best points on the graph with lines. The Pareto Front represents the best possible solutions, i.e., selecting a point below the Pareto Front would be sacrificing quality by reducing the C_{pk} while not reducing the cost. The points on the Pareto Front are labeled to identify them for further analysis. The Pareto Front is drawn and the points are labeled on the graph shown in Figure 5-17.

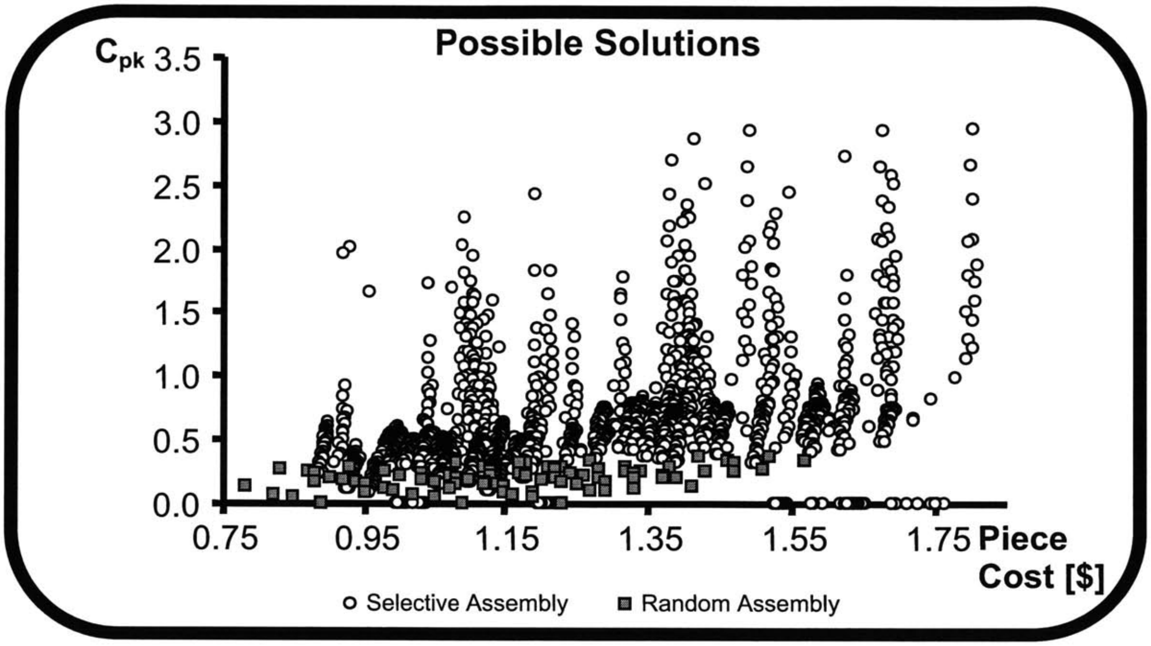

Figure 5-16: Possible solutions

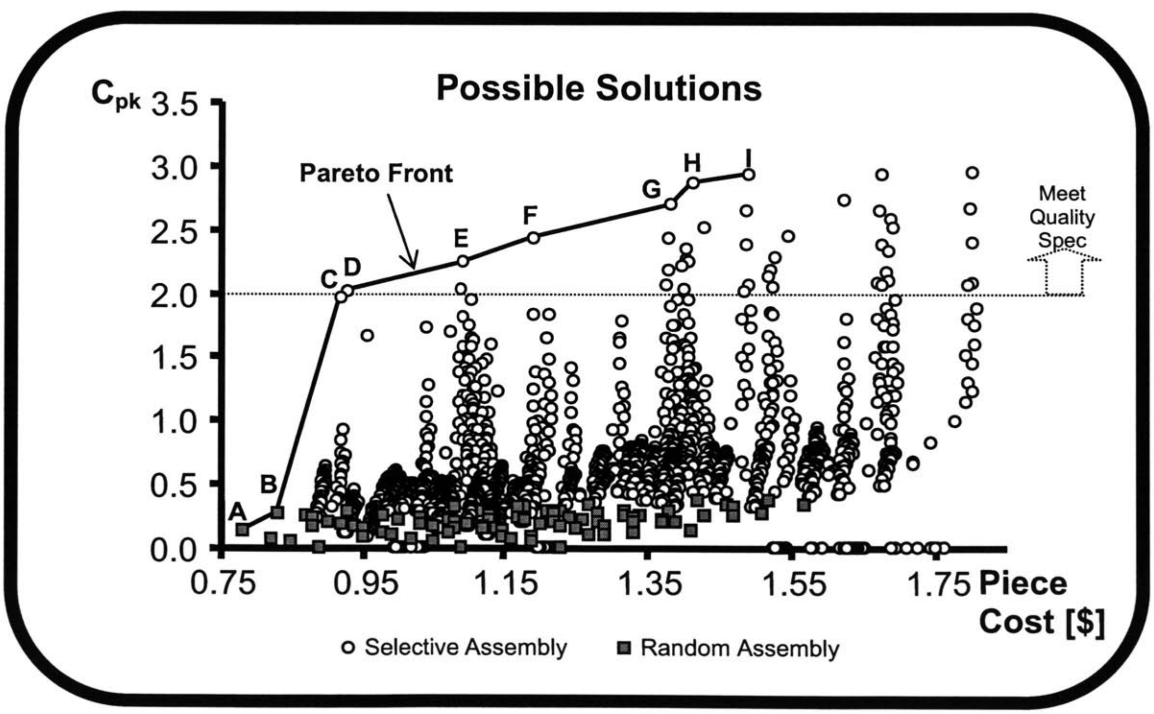

Figure 5-17: Possible solutions with Pareto Front drawn

The dotted horizontal line on the graph in Figure 5-17 represents the quality specification limit of a C_{pk} value of two. The points above the spec limit meet the quality requirement while those below do not. Points A and B clearly do not meet the required C_{pk}; therefore, they are eliminated from consideration. Points C through I are possible contenders for the combination that meets the quality specification. The combinations and numerical values associated with these points are listed below in Table 5-7.

Table 5-7: Feasible solution combinations

Point	Bar Stock	Roughing Operation	Finishing Operation	Drilling Operation	Assembly Operation	Cost Per Piece [$]	C_{pk}
C	C	3	3	2	EW:MV N = 3	0.92	1.96
D	C	2	3	2	EW:MV N = 3	0.93	2.01
E	A	2	3	2	EW:MV N = 10	1.09	2.24
F	C	2	1	2	EW:MV N = 3	1.19	2.42
G	A	3	3	1	EW:MV N = 10	1.38	2.69
H	B	2	3	1	EW:MV N = 3	1.41	2.86
I	A	1	3	1	EW:MV N = 10	1.49	2.93

Based on these results, Cyler probably chooses Point D, that is, he decides to use Supplier C, Roughing 2, Finishing 3, Drilling 2, and Equal Width Binning to Minimize Variation with three bins. This combination produces a clearance with a C_{pk} greater than two for the lowest cost.

Cyler realizes making this choice may be premature without investigating the worst case values this combination could produce. The numbers plotted in Figure 5-17 and listed in Table 5-7 are based on the expected values of the standard deviation and scrap percentage leaving the selective assembly operation. As stated in the development of the closed-form equations for the selective assembly techniques in Chapter 4, selective assembly is a stochastic process. The probabilistic nature of selective assembly operations allowed for the development of 95% confidence intervals. These confidence intervals are now used to compute a 95% worst case scenario for each of the Pareto points identified above. These worst case scenarios represent the highest cost and lowest C_{pk} values expected for the given combinations with a 95% confidence; i.e., for a given combination, there is a 95% probability that the value of C_{pk} is greater and the cost is less than the worst case values. The worst case values for the Pareto Front points are plotted in Figure 5-18 and are listed in Table 5-8.

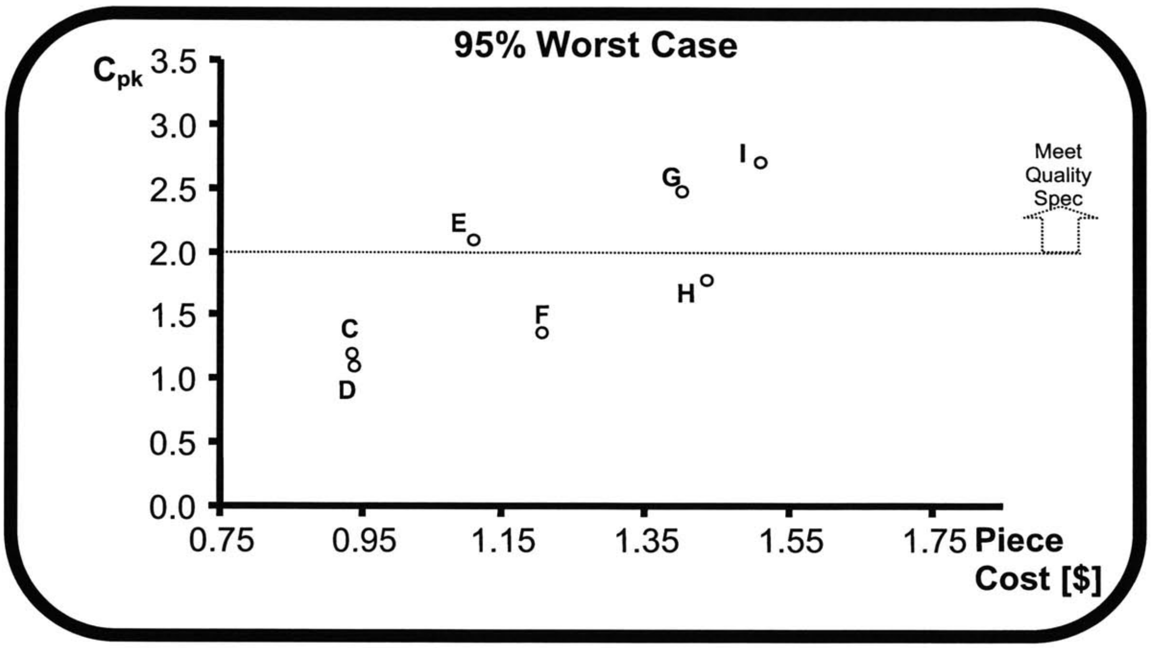

Figure 5-18: 95% worst case scenario for Pareto Front feasible points

Table 5-8: Values for 95% worst case scenario

Point or Combination	Cost Per Piece [$]	C_{pk}
C	0.93	1.21
D	0.94	1.10
E	1.11	2.11
F	1.20	1.37
G	1.40	2.49
H	1.43	1.79
I	1.51	2.74

Reviewing the worst case values, Cyler realizes that combination E is the least expensive combination that is most likely to produce assemblies with the desired quality, i.e, Point E has a 95% probability of producing parts with a C_{pk} greater than two.

The final choice Cyler has to make between combination C and combination E depends on how sensitive the widget is to cost and how sensitive Cyler is with risk. If the product is cost sensitive, then combination C is probably the best choice with the risk that sometimes the parts will not have a C_{pk} of two. If Cyler is not willing to take the risk that the quality level will dip below a C_{pk}

value of two, his company should probably pay the higher price associated with combination E and rest assured that the quality will be better than the desired C_{pk} value of two.

5.6 Chapter Summary

In this chapter, the developments in the areas of process capability databases, manufacturing process modeling, and assembly operation modeling were combined to form Variation Risk Understanding and Management (VRUM). VRUM is a design tool that is used to forecast the quality of a product's dimension before it is produced. The method of how to compose a VRUM system level model and equations was described. The role of VRUM in the product development process was discussed. Finally, an example was presented to demonstrate how VRUM can assist with design decisions.

6 Conclusion

6.1 Introduction

In this thesis, a design tool has been developed that uses process capability data to construct closed-form equations to forecast the mean and variation of a product's dimension upon exiting a manufacturing system. The design tool is comprised of a process capability database, math models of manufacturing processes, and math models of assembly operations as discussed in Chapter 2, 3, and 4 respectively. The research advancements in each of these areas are re-stated in the next section to summarize the overall impact of this thesis. Possible areas of future work that could further enhance this body of research are discussed in Section 6.3.

6.2 Research Advancements

Many challenges were solved during the development of the design tool presented in this thesis. The goals for each thesis topic are listed below. Each major, specific challenge is then listed followed by the solution to it. Each solution is followed by a bulleted list summarizing the advantages it provides. The goals, challenges, and solutions are grouped into the categories of process capability databases, the modeling of manufacturing processes, the modeling of assembly operations, and the system approach to correspond with the structure of the thesis, i.e., each group corresponds to a specific chapter.

6.2.1 Process Capability Database

Topic Goals

- Develop a method of representing data so that it is easily retrievable
- Formalize the steps required to transform component data into a form suitable for database entry and storage
- Design a user interface for process capability databases
- Suggest a list of data that should be included in process capability databases

Challenge: **Representing manufacturing data in a database so that it is easy to retrieve**

Solution: **Created the Attribute Combination Descriptor (ACD) (Section 2.2.1)**

+ Four attributes are identified and populated to fully characterize a physical aspect of a component
+ Engineers can find data easily
+ Familiar terminology can be used
+ Attributes are easily extensible

Challenge: **Transforming part data into a form for easy database entry and storage**

Solution: **Created the ACD Matrix Set to assist in the transforming of part data (Section 2.2.3.1)**

+ Well structured
+ Mistakes are avoided
+ Serves as a knowledge management tool
+ Helps identify important part characteristics that may need to be tracked

Challenge: **Providing a user interface to a process capability database**

Solution: **Designed Graphical User Interface (Section 2.3)**

+ Intuitive fields guide user
+ Dynamic menus eliminate infeasible combinations
+ Accessible via a standard web browser

Challenge: **Knowing the data to include in the process capability database**

Solution: **Devised a record set for process capability data (Section 2.4)**

+ Includes the data needed to compute process capability indexes, e.g., Cpk, Cp
+ Data fields to identify when and where the parts were produced
+ Contains data needed to construct math functions that are introduced in Chapter 0

6.2.2 Modeling of Manufacturing Processes

Topic Goals

- Produce a simple method to model and understand variation propagation through a manufacturing process, i.e., how does the level of variance of a dimension entering a process affect the variance of it exiting the process?
- Develop a general method for constructing closed-form equations for the mean and standard deviation of a dimension exiting a manufacturing process as a function of the input dimension's mean and standard deviation, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$
- Validate constructed equations

Challenge: **Understanding variation propagation through a manufacturing process**

Solution: **Developed closed-form equations (Section 3.2 to 3.5)**

+ Simple to use
+ Backed by physical significance, i.e., the foundation is the theory of a process imparting a dimensional change and variation to components, which are both physical concepts (see Section 3.2)
+ Requires no prior knowledge of the process
+ Quickly construct equations for many processes to make comparisons
+ Generally applicable, i.e., this approach can be used to develop equations for many different types of manufacturing processes
+ Computationally inexpensive
+ Can be created early in the design process
+ Link equations together to represent serial processing (See Section 5.2.3)

Challenge: **Defining a framework to represent the effects a manufacturing process has on component dimensions**

Solution: **Developed Theory of Desired Dimensional Change, Process Imparted Dimensional Change, Process Imparted Variation (Section 3.2)**

+ Can be represented with equations or graphically
+ Easy to use to qualitatively compare processes' abilities to create quality parts
+ Building block for constructing closed-form solutions, i.e., $\mu_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$ and $\sigma_{out} = f(\mu_{in}, \sigma_{in}, \mu_{in})$

Challenge: **Validating theory and resulting closed-form equations**

Solution: **Compare closed-form equation results to simulated data and to actual data (Section 3.2)**

+ Validated theory works well to forecast the mean and standard deviation of a dimension exiting a manufacturing process
+ Demonstrated theory works with real data

6.2.3 Modeling of Assembly Operations

Topic Goals

- Define new selective assembly techniques for use when component distribution variances are unequal
- Develop criteria and analysis approach to use when choosing a selective assembly technique
- Create closed-form equations to calculate the standard deviation of an assembly that is created by a selective assembly operation
- Explore the robustness of selective assembly to changes in incoming component variances

Challenge: **Understanding different selective assembly techniques**

Solution: **Created new selective assembly techniques to minimize scrap (Section 4.2.2 and 4.2.4)**

+ Provides a reduction in assembly variation while minimizing the cost impact of using selective assembly
+ Easy to implement

Challenge: **Calculating the standard deviation of a dimension resulting from a selective assembly operation**

Solution: **Developed closed-form equations for selective assembly techniques (Section 4.4.1)**

+ Easy to use
+ Used to determine the best selective assembly technique to implement
+ Can be implemented in a wide variety of computer applications, e.g., spreadsheets, math programs, etc.

Challenge: **Performing sensitivity analysis on selective assembly operations**

Solution: **Used closed-form equations for selective assembly to perform sensitivity analysis (Section 4.4.2)**

+ Directly compute sensitivities by taking partial derivatives
+ Computationally inexpensive

6.2.4 System Approach

The main goal of the system approach was to integrate the previous topics into a single design tool know as Variation Risk Management and Understanding (VRUM). Icons were created so that a graphical representation of a system could be easily constructed. The system diagram was then shown to help formulate the closed-form equations that describe the manufacturing system. The role of VRUM in a product development process was discussed followed by guidelines to help implement the tool. Finally, an example problem was presented to demonstrate how VRUM can assist with complex design decisions.

6.3 Future Research

The work presented in this thesis developed a complete and functional design tool; however, additional work could further enhance the usefulness of the tool. Topics for possible future work are discussed here. The topics are sorted into the categories of process capability databases, the modeling of manufacturing processes, the modeling of assembly operations, and the system approach.

6.3.1 Process Capability Database Future Work

Automate the Indexing Process

The implementation and maintenance of a process capability database would be enhanced if the creation of Attribute Combination Descriptor were automated. Algorithms consisting of possible attributes could be combined with advanced feature recognition software to accomplish this automation.

Use Process Math Models During Indexing

The ACD Matrix Set population process may be further improved by using the process models created in Chapter 0 to create mathematical relationships that could forecast the final product variation during the indexing process and offer alternative design options to improve the end capability. This is similar to the possible future work of integrating the elements of VRUM more tightly as stated below.

6.3.2 Modeling of Manufacturing Processes Future Work

Use Physics-based Models for Data Generation

The usefulness of the math models in Chapter 3 are highly dependent on the availability of process capability data or the ability to "run" sample parts to generate process capability data, e.g., the machinery does not exist yet. The idea of using physics-based models to generate capability data, and then use the generated data to construct closed form-equations is an interesting one that could extend the applicability of the math models. This idea is partially explored by Kern, Du, and Sudjianto (Kern, Du et al. 2003).

Extend Concept to Parameter Optimization

The concept in Chapter 0 represents a manufacturing process by two math functions. One related to mean shift, and one related to variation. The functions are then used to construct closed-form equations for these quantities. The equations accept dimensions as inputs and produce dimensions as outputs. It would be interesting to attempt to apply this concept to non-dimensional quantities. For example, develop a series of math functions to represent a particular process. Each math function would accept a different parameter as an input and produce a contribution to the output.

6.3.3 Modeling of Assembly Operations Future Work

Use Selective Assembly Closed-form Equations as Real Time In-line Tool

The closed-form equations developed for selective assembly were intended to help design a product or choose what type of selective assembly operation to implement. It would be interesting to use these equations to program a system that would continuously re-calculate the best technique and number of bins to use in an active manufacturing environment. For example, in a production system using selective assembly, this concept could be added so that shifts in distribution means or variances would be detected and the best selective assembly technique defined or the number of bins being used could be modified to produce better output at little to no extra cost. The new technique or number of bins would be implemented, and the steps repeated.

Include Assembly Adjustments

Presently, only part-defined assemblies are modeled, i.e., assemblies that are created by mating parts together. The other type of assembly that was not modeled in this thesis is the type that requires adjustments during the actual assembly operation. Examples of this type of assembly are automobile body panels, aircraft wings, etc. Including models for this type of assembly would broaden the usefulness of the design tool.

6.3.4 System Approach Future Work

Improve Design Tool Integration

Presently, the VRUM design tool spans several computer applications. The prototype process capability database was accessible via a web browser, but the other modeling was accomplished in desktop computer applications such as Mathematica® and Microsoft® Excel. A goal of future work is to integrate all of the math modeling into one, seamless web application or software package.

Create Interactive Graphical User Interface

Other than the process capability database graphical user interface, no other portion of the research is "user-friendly." Creating an interactive graphical user interface (GUI) would make the VRUM tool much easier to use. The GUI could have icons that could be "dragged and dropped" to construct VRUM system diagrams.

Automate the Design Tool

In the future, the aforementioned integration and GUI could be combined with algorithms to automate the construction of the closed-form equations. Just click a button and the equations are generated automatically.

Include Considerations for Inventory, Scheduling, and Capacity Issues

The design tool developed in this thesis was mainly concerned with dimensional variation and costs. Including considerations for inventory, scheduling, and capacity issues would greatly enhance its functionality. Then, not only could a product be designed that could be manufactured with a high level of quality, but a product could be designed so that it could be produced in a timely manner with no inventory shortfall or surplus.

Combine Design Tool with Optimization Techniques

The design tool could be combined with optimization techniques to broadly explore design spaces to find the design that could be manufactured with the highest level of quality.

Apply Concept to Other Semi-structured Data

The basic concept of this thesis is to take existing data (process capability data) and use it to create models that assist in the design of other products. This concept could be applied to other bodies of data that exist and are underutilized. For example, warranty data is collected and tracked but is rarely leveraged to improve the design of other future products.

6.4 Conclusion

A design tool that forecasts the mean and standard deviation of a product's dimension upon exiting a manufacturing system has been developed and presented. The tool is comprised of a process capability database, math models of manufacturing processes, and math models of assembly operations. The process capability database was created to connect engineers to manufacturing data to assist with detailing a design and to make the needed data available for model construction. The theory developed models a production process with two math functions, which are constructed using the process capability data. These two math functions make it possible to build closed-form equations that calculate the mean and standard deviation of parts exiting a process. These equations are used to analyze tradeoffs among processes, to compute the impact incoming variation has on output, and to identify sources of variation. Finally, closed-form equations were created that compute the variation of a product resulting from a selective assembly operation. Using these tools, forecasting and managing manufacturing variation is possible for a wide variety of products and production systems.

Appendices

A. General Nomenclature

The symbols and acronyms used throughout the thesis are listed here for easy reference. The definitions to most terms listed here are included in the Glossary at the end of the thesis. The nomenclature used specifically in the simulation models in Section 3.6 is listed in Appendix H.

Symbol or Acronym	**Meaning**
C_p	Process capability index, see the Glossary for a definition
C_{pk}	Process capability index, see the Glossary for a definition
Δ	Bin width for selective assembly operation
KC	Key Characteristic
μ	Mean
N	Number of bins used for a selective assembly operation
PCDB	Process Capability Database
PDP	Product Development Process
σ	Standard deviation

B. Survey Results

Identifying user needs is essential to developing a successful Process Capability Database. A survey and interviews were conducted to gather insight into the methods engineers use to retrieve manufacturing capability data and the tools design and manufacturing engineers use when specifying dimensions and tolerances on detailed drawings. The survey was administered to engineers working at an American automobile manufacturer to gain insight into the present practices of how manufacturing variation is considered while a new component is being designed. The survey focused on the information utilized during the design phase. The respondents categorized themselves into the categories of design engineers, manufacturing engineers, and quality engineers.

The term design engineer refers to a person responsible for the creation of detailed drawings of a component or an assembly. Detailed drawings are also known as engineering prints, blueprints, or engineering drawings. A detailed drawing usually contains different views of a component with dimensions specified. Each dimension has an associated tolerance range. During the creation of detailed drawings, a design engineer relies on physical laws, empirical data, experience, and other sources of input to dimension and tolerance a component in such a way that it will perform its intended function.

The term manufacturing engineer identifies the people responsible for aspects of the manufacturing process. These aspects include the design, creation, and supervision of a manufacturing process. A manufacturing process can have a few or many individual operations; therefore, multiple manufacturing engineers could be working together on a single production line. Since the variation of a key characteristic significantly affects the functionality of a component, key characteristics are often measured and monitored during manufacturing. This monitoring can take the form of Statistical Process Control and is usually overseen by a specific type of manufacturing engineer, typically known as a quality engineer. This manufacturing engineer makes adjustments to the manufacturing process if the monitoring indicates a shift in the component's dimension is taking place. A manufacturing engineer would like to be able to easily retrieve manufacturing capability information. The retrieved information could be used to assist in a new production layout, with the purchase of a new piece of equipment, quality control, and for numerous other reasons.

An overwhelming majority of respondents (95%) said that they would talk to a manufacturing engineer when they wanted data about production capability. This was the preferred method over locating the data in a local file at the manufacturing plant or locating the data in a centralized file system. This strongly implies an easily and readily available source from which manufacturing data can be retrieved does not exist. One engineer commented that the data would have to be generated (new measurements taken), because if any data that was stored could be found, it would not be reliable.

Most engineers (75%) responded that manufacturing process capability data is presently difficult to retrieve. A few (10%) responded that the data did not exist. A few comments focused on the fact engineers must rely on a personal contact at a manufacturing plant in order to be able to retrieve the data. These responses and comments reinforce the notion that any existing manufacturing capability data is difficult to access.

A majority (60%) of engineers said that the feature being dimensioned (e.g., hole diameter, plate thickness, etc.) is the first consideration for specifying a tolerance on a new component drawing. A minority of 15% stated they primarily consider what method is going to be used to manufacture the component when specifying tolerances (66% of these respondents had categorized themselves as manufacturing engineers). The remaining 25% listed their own individual considerations that mostly focused on the functional requirements of the component. Only one respondent mentioned manufacturing variation as a consideration. These responses indicate that variation in manufacturing is not a top concern for engineers.

Half the engineers surveyed stated they initially use an existing detailed drawing of a component similar to the new one being created when specifying a tolerance. The next largest percentage of respondents (15%) stated they reference process capability data as the first step in specifying tolerances on a new detailed drawing (100% of these engineers classified themselves as manufacturing engineers). Smaller percentages said they would consult a manufacturing engineer (10%), reference company standards (10%), reference an industry handbook (5%), review the proposed dimension with a group of design engineers (5%), or base the tolerance on the function of the feature being dimensioned (5%). These percentages reflect the dependence an engineer has on prior knowledge of an existing detailed drawing or a close contact with an appropriate manufacturing engineer to retrieve the process capability data when the engineer is creating tolerances for a new detailed drawing. This implies that without a familiarity with existing designs or a manufacturing contact person, an engineer may not be able to properly assign tolerances for a new drawing. This can be especially true for an engineer new to a company or product group.

When the above question was expanded to ask the engineers to check all the sources they use when detailing a new drawing the results were similar. The majority of respondents (75%) said they refer to existing detailed drawings of similar components and consult with a manufacturing engineer. A smaller majority (55%) stated they refer to a set of company standards. An industry standards handbook was listed and came in a distant fourth with only 20% of the people listing it as a source of information.

The survey included a question asking how often the engineer would use an easily accessible process capability database when specifying detailed drawing tolerances. The response could be always, frequently, sometimes, rarely, or never. A large majority (72%) would either always or frequently make use of an easily accessible process capability database. A smaller percent (17%) said they would sometimes use it. Only 11% said they would rarely or never make use of such a resource. It is interesting to note that these respondents were the only manufacturing engineers to respond. The design engineers who responded comprised the 89% that said they would at least use the database sometimes. This indicates design engineers are the primary target users for a process capability database while the manufacturing engineers are the secondary user group.

A question was included to solicit feedback on what an engineer would like to use as an input when utilizing manufacturing process capability data during the detailing of a new drawing. One option was to specify a C_{pk} value and use the capability data to calculate the minimum tolerance range. The other option was to specify a tolerance and calculate the corresponding C_{pk}. The first option was selected by 35%, and the second option was chosen by 55%. A few respondents (10%) listed "other" as there choice. This demonstrates that any methodology that is created needs to be flexible to be attractive to engineers with varying preferences.

Another interesting observation the survey made clear is the fact that material selection for a new component is almost solely based on the functional requirement of the component. An overwhelming 85% selected functional requirements as the material selection criterion. This may be expected, but as stated earlier, a large number of engineers use information from sources other than functional requirements (product specifications) to create tolerances, which may be counter-intuitive. This revelation implies that the impact the selected material has on manufacturing variation does not need to be heavily considered because it would not factor into the design process, i.e., the material is selected based on the component's functional requirements regardless of the impact it has on manufacturing variation.

Interviews were conducted with engineers to gain additional insight into their needs. The overarching theme of the discussions was to make the PCDB simple to use. The thought was that engineers are very busy and do not have time to learn a complicated system, no matter how useful. Another common theme was that design engineers do not understand manufacturing engineers and vice versa. Manufacturing engineers were quick to point out how some tolerances are specified that cannot be produced realistically. Design engineers shared their belief that manufacturing engineers do not have a grasp of tolerances being required to meet functional requirements. One example was that a manufacturing engineer could not understand why a tolerance that was difficult to maintain in production could not be loosened. Engineers wanted the PCDB to be easily accessible. They were not interested in specialized software or having to go to a computer other than the one on their desktops. This was mostly due to convenience and the amount of time required to learn how to use the PCDB and the time needed to use it in practice.

The interviews revealed a few concerns regarding the usage of a PCDB. The largest concern is that design engineers would use the data to specify tolerances that could not be met with an acceptable capability index. This situation could occur when a design engineer retrieves data from the PCDB regarding a very specialized or low volume production line, and then use this data to justify a tight tolerance on a part produced in high volume. A strategy to minimize this concern is to include information in each dataset that indicated the volume of production associated with the capability indexes. The production volume information enlightens design engineers using the PCDB as to the reality of achieving similar capability on a new design. Another concern is that upper management would use the PCDB as a means of monitoring the manufacturing engineering department's performance. This concern could be realized in the event a process went out of control and the data was recorded into the PCDB. This out of control data would not reflect the true, in control capability of the production line and could be misconstrued by management. A system of filtering the data prior to input is implemented to eliminate this concern. A person at each site is responsible for manually filtering the data prior to inputting it into the database. The filtering process eliminates any production measurements that came from parts manufactured during a time when the process is not in control. Finally, managers and engineers are concerned the data could be used in litigation against the company if it is intercepted and misinterpreted by outside parties. The real threat would be in the misrepresentation or skewing of the information contained in the database. In addition to the filtering mentioned above, security is implemented to minimize the risk of this concern. The security protocol includes limiting access to only those employees who require it.

The information gathered from the survey and interviews clearly reinforce the need for an easy to use Process Capability Database. The survey results show that design engineers often do not know where to obtain accurate, reliable manufacturing process capability data. This results in the engineers specifying tolerances without considering the impact of manufacturing variation. The interviews reveal that the database developed must be easy to use and understand, be easily accessible, and be easy to learn.

Table B-1: Respondents by department

Department	Number of Respondents
Design Engineering	10
Manufacturing Engineering	9
Quality Engineering	2

Table B-2: Respondents by powertrain system

Powertrain System	Number of Respondents
Engine	16
Transmission	3
Quality Engineering	2

Table B-3: Survey questions and responses

1) Typically, the first consideration for specifying a tolerance on a new component drawing is ______________ .	**Total**	Design Engineers	Manufacturing Engineers	Quality Engineers
A. what method will be used to manufacture the component	**3**	1	2	0
B. the feature being dimensioned (e.g., hole diameter, plate thickness, etc.)	**12**	6	4	2
C. the material which will be used to manufacture the component	**0**	0	0	0
D. other (please list)	**5**	3	2	0
Component durability requirements Function of feature, as well as material properties are considered initially. The mfg. method is next considered, and design/tol. modified if req'd. Product engineering are typically the first to specify tolerances To best suit the application of that component, the tolerance is to match the design intent, especially to accommodate manufacturing variations.				

2) The first step in specifying a tolerance on a new component is to ______________ .				
A. look at an existing detailed drawing of a similar component	**10**	6	2	2
B. consult a manufacturing engineer	**2**	1	1	0
C. review the proposed dimension with a group of design engineers	**1**	0	1	0
D. reference company standards	**2**	1	1	0
E. reference company process capability data	**3**	0	3	0
F. reference an industry handbook	**1**	0	1	0
G. other (please list)	**1**	0	0	0
the function of the feature being dimensioned				

3)I consult __________ before specifying tolerances on a new engine component detailed drawing. *[Check all that apply]*				
A. a manufacturing engineer	**15**	8	5	2
B. existing detailed drawings of similar components	**15**	9	4	2
C. handbooks	**3**	1	1	1
D. company standards	**11**	7	3	1
E. industry standards	**4**	2	2	0

F. other (please list)	**3**	3	0	0
Guide file drawing Full service suppliers usually create the detail drawings for cam drive and VCT system components.				

4) A material is selected for a new engine component based on ____________.				
A. the material of a similar component	**1**	0	0	1
B. cost	**1**	0	0	1
C. the functional requirements of the component	**17**	9	8	0
D. how it will be manufactured	**0**	0	0	0
E. the component's features	**0**	0	0	0
F. other (please list)	**1**	1	0	0
all of the above Cost, rough part and process requirements for ease of manufacture "Function", with qualification that cost is part of "function".				

5) I ____________________ when I want data about my company's production capability.				
A. talk to a manufacturing engineer	**18**	10	7	1
B. reference a local file at a particular plant	**0**	0	0	0
C. reference a central file which contains information from several plants	**0**	0	0	0
D. other (please list)5	**2**	0	1	1
find it out myself I am the manufacturing person The best match comparison is made which may mean one or more accessed Generate it myself – Measurement systems too unreliable for believing file data.				

6)Process capability data (i.e., Cpk, Cp, Ppk, Pp) is presently __________________.				
A. non-existent	**2**	1	1	0
B. difficult to retrieve	**15**	8	6	1
C. easy to retrieve	**1**	0	1	0
D. other (please list)6	**2**	1	0	1
Unless asked departments do not readily calculate this item Can be difficult at times Depends, easy when you know who to contact, at a plant that keeps a DCP on-line. Progressively more difficult as plant information system, and ability to make a correct personal contact get stretched.				

7) I would __________ use an easily accessible process capability database when specifying detailed drawing tolerances.				
A. always	**7**	3	2	2
B. frequently	**6**	3	3	0
C. sometimes	**3**	3	0	0
D. rarely	**1**	0	1	0
E. never	**1**	0	1	0

8) When I am dimensioning a new component, I would prefer _______________ for a Significant Characteristic based on process capability data.				
A. specifying a Cpk and deriving the corresponding tolerance	**7**	4	3	0
B. specifying a tolerance and deriving the corresponding Cpk	**11**	5	5	1
C. other (please list)	**2**	1	0	1
I would tend to specify a Cpk depending on the type of in process operation. Basing tolerance of the SC on DoE, regression, tolerance parallelograms, optimization,etc.				

Table B-4: Additional comments submitted with survey responses

Short term data from machine capability runoffs would probably be accessible if extracted shortly after runoff. Long term data is normally not available unless the particular dimension is deemed a critical characteristic.
The crankshaft community is somewhat unique in that our guide file drawing has been extensively deveolped based on manufacturing, product and gaugeing requirements.
Don't forget Casting in any database. Also, discussions of capability without strong knowledge of the measurement systems used to generate the data is dangerous to the point of invalidating your data base.
In most cases several things are considered simultaneously in the design process. The component function is obviously the driver for the design, but existing practice, cost, material and manufacturing process and capability all are considered. This proposed database would be a big enabler for engineers to bring robustness practices into their designs, rather than to rely on conversations with their various manufacturing engineers.
Please rewrite survey to be generic ... the engine is not the only part of the powertrain. Make sure any capability database includes capability for processes that are unique to transmission (valve body grinding, valve body boring,
Will the database include supplier information?...or can it be made flexible to include suppliers, or links to suppliers in the future.

The goal should be to maitain this database with some integrity. We should first walk before we try to start running. Current correlation of design intent to the production capability is absolutely essential to meet the standards of good product design, let alone design for six sigma. It would be better to gather data based on application than any general machining process, because application reference will be more specific than being generic. This database will be of value if the information it contains can be trusted. Confidence of design Engineers is based on their knowledge of real life process- not from any book. This communication between manufacturing and design is what we should strive for before even we start maintaining a database. So may be we should first start walking....Suggestion: Create a database of component characteristics that are critical (prioritize) and get the current data. Expand this database with more authentic information from global facilities to compare or benchmark....Use it to design a component for reliability.

C. Selective Assembly Variation Coefficients

The following equation is a restatement of Equation 4-20. It is the closed-form equation for the standard deviation of a dimension, $\sigma_{C,SelAsm(N)}$, created as a result of using selective assembly to combine two normal distributions of parts, a Part A having a standard deviation of σ_A with a Part B having a standard deviation of σ_B. The coefficients, C_0, C_1, C_2, and C_3, are retrieved from the tables below for the specific technique and number of bins of interest.

$$\frac{\sigma_{C,SelAsm(N)}}{\sqrt{\sigma_A^2+\sigma_B^2}} = (C_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (C_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (C_1)\left(\frac{\sigma_A}{\sigma_B}\right) + C_0 \qquad \text{C-1}$$

$$\text{for } 1 \le \frac{\sigma_A}{\sigma_B} \le 2$$

Table C-1: Equal Width Binning to Minimize Variation

	$\sigma_{C,SelAsm}$				Upper 95% Band				Lower 95% Band			
N	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0
1	0.1321	-0.6752	0.9126	0.6207	0.1626	-0.8188	1.1282	0.5691	0.1016	-0.5316	0.6970	0.6795
2	0.0498	-0.2868	0.4174	0.4070	0.1302	-0.6692	1.0040	0.1479	-0.0307	0.0957	-0.1693	0.666
3	0.0418	-0.1627	0.0431	0.5799	0.008	-0.0131	-0.1748	0.7074	0.0755	-0.3123	0.2610	0.4523
4	0.0551	-0.2235	0.1574	0.4062	0.0065	0.0019	-0.184	0.5928	0.1038	-0.4489	0.4988	0.2196
5	0.0309	-0.1083	-0.0012	0.4060	0.0298	-0.1059	-0.0059	0.427	0.0319	-0.1106	0.0034	0.3851
6	0.0014	0.0192	-0.1626	0.4185	-0.0104	0.0789	-0.2651	0.4898	0.0132	-0.0405	-0.0601	0.3472
7	0.0172	-0.0537	-0.0422	0.3188	0.0020	0.0230	-0.1727	0.4053	0.0324	-0.1304	0.0883	0.2324
8	0.0139	-0.0447	-0.0365	0.2772	-0.0295	0.1578	-0.3460	0.4418	0.0572	-0.2472	0.2729	0.1127
9	-0.0010	0.0249	-0.1348	0.3004	-0.0402	0.214	-0.4346	0.4659	0.0383	-0.1641	0.1651	0.1350
10	0.0094	-0.0230	-0.0567	0.2411	-0.0612	0.3066	-0.5610	0.5023	0.0800	-0.3526	0.4476	-0.0201

Table C-2: Equal Width Binning to Minimize Scrap

	$\sigma_{C,SelAsm}$				Upper 95% Band				Lower 95% Band			
N	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0
1	0	0	0	0.9889	0	0	0	1.0363	0	0	0	0.9414
2	-0.0954	0.4623	-0.6213	0.8428	-0.1157	0.5484	-0.7240	0.9103	-0.0751	0.3762	-0.5186	0.7753
3	-0.1266	0.6141	-0.8310	0.8488	-0.1259	0.5983	-0.7638	0.8208	-0.1274	0.6298	-0.8981	0.8768
4	-0.1726	0.8266	-1.1071	0.8529	-0.1678	0.7852	-0.9905	0.7962	-0.1774	0.8681	-1.2238	0.9097
5	-0.2036	0.9626	-1.2656	0.8358	-0.2087	0.9647	-1.2074	0.7980	-0.1986	0.9605	-1.3237	0.8737
6	-0.2259	1.0563	-1.3652	0.8143	-0.2264	1.0304	-1.2511	0.7427	-0.2254	1.0822	-1.4794	0.8859
7	-0.2442	1.1335	-1.4505	0.8029	-0.2155	0.9635	-1.1058	0.6143	-0.2728	1.3036	-1.7953	0.9915
8	-0.2591	1.1885	-1.4901	0.7723	-0.2122	0.9335	-1.0189	0.5262	-0.3060	1.4434	-1.9614	1.0184
9	-0.2655	1.2080	-1.4900	0.7377	-0.2198	0.9564	-1.0207	0.4904	-0.3113	1.4596	-1.9593	0.9849
10	-0.2757	1.2447	-1.5170	0.7185	-0.2162	0.9278	-0.9468	0.4217	-0.3352	1.5616	-2.0872	1.0153

Table C-3: Equal Area Binning to Minimize Variation

	$\sigma_{C,SelAsm}$				Upper 95% Band				Lower 95% Band			
N	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0
1	0.1351	-0.7054	0.9930	0.5679	0.2019	-1.0249	1.4735	0.3812	0.0683	-0.3860	0.5125	0.7545
2	0.0491	-0.2768	0.4031	0.4152	0.1167	-0.6103	0.9343	0.1708	-0.0185	0.0566	-0.1281	0.6597
3	0.0204	-0.1518	0.2688	0.3062	0.0667	-0.3734	0.6091	0.1634	-0.0259	0.0697	-0.0714	0.4490
4	0.0331	-0.2029	0.3509	0.1820	0.0451	-0.2643	0.4481	0.1565	0.0210	-0.1415	0.2536	0.2075
5	0.0417	-0.2388	0.4078	0.0981	0.0524	-0.2850	0.4721	0.0894	0.0309	-0.1925	0.3435	0.1069
6	0.0311	-0.1880	0.3337	0.0924	-0.0010	-0.0308	0.0814	0.2511	0.0632	-0.3452	0.5859	-0.0662
7	0.0389	-0.2165	0.3664	0.0529	-0.0017	-0.0159	0.0431	0.2490	0.0795	-0.4170	0.6898	-0.1433
8	0.0485	-0.2564	0.4211	0.0053	0.0245	-0.1358	0.2239	0.1370	0.0726	-0.3771	0.6183	-0.1263
9	0.0313	-0.1755	0.2963	0.0493	0.0013	-0.0331	0.0747	0.1898	0.0614	-0.3179	0.5179	-0.0913
10	0.0362	-0.1962	0.3259	0.0200	0.0346	-0.1821	0.2931	0.0698	0.0379	-0.2104	0.3587	-0.0298

Table C-4: Equal Area Binning to Minimize Scrap

	$\sigma_{C,SelAsm}$				Upper 95% Band				Lower 95% Band			
N	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0
1	0	0	0	0.9916	0	0	0	1.0393	0	0	0	0.9439
2	-0.0938	0.4579	-0.6213	0.8496	-0.0952	0.4546	-0.5884	0.8488	-0.0923	0.4611	-0.6542	0.8504
3	-0.1457	0.7026	-0.9395	0.8255	-0.1488	0.7087	-0.9169	0.8218	-0.1426	0.6965	-0.9622	0.8292
4	-0.1878	0.8917	-1.1765	0.8344	-0.1915	0.9024	-1.1625	0.8344	-0.1841	0.8811	-1.1905	0.8343
5	-0.2024	0.9528	-1.2290	0.7865	-0.2129	0.9935	-1.2616	0.8072	-0.1918	0.9120	-1.1964	0.7659
6	-0.2277	1.0581	-1.3468	0.7863	-0.2139	0.9803	-1.1927	0.7176	-0.2414	1.1358	-1.5010	0.8549
7	-0.2396	1.1081	-1.3983	0.7732	-0.2145	0.9814	-1.1782	0.6743	-0.2646	1.2349	-1.6183	0.8421
8	-0.2455	1.1283	-1.4034	0.7410	-0.2309	1.0490	-1.2518	0.6726	-0.2601	1.2075	-1.5551	0.8094
9	-0.2542	1.1596	-1.4251	0.7215	-0.2197	0.9861	-1.1299	0.5807	-0.2888	1.3331	-1.7203	0.8622
10	-0.2583	1.1721	-1.4249	0.6982	-0.2392	1.0734	-1.2500	0.6200	-0.2774	1.2707	-1.5997	0.7763

Table C-5: Optimal Width Binning

	$\sigma_{C,SelAsm}$				Upper 95% Band				Lower 95% Band			
N	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0	C_3	C_2	C_1	C_0
1	0	0	0	0.9883	0	0	0	1.0303	0	0	0	0.9462
2	-0.1153	0.5572	-0.7654	0.9127	-0.0952	0.4641	-0.6134	0.8627	-0.1355	0.6503	-0.9174	0.9626
3	-0.1242	0.6098	-0.7993	0.7358	-0.1734	0.8327	-1.1105	0.8942	-0.0749	0.3869	-0.4880	0.5774
4	-0.1805	0.8592	-1.1100	0.7616	-0.1972	0.9234	-1.1702	0.7942	-0.1637	0.7950	-1.0498	0.7289
5	-0.1968	0.9264	-1.1673	0.7074	-0.2361	1.0950	-1.3862	0.8158	-0.1576	0.7579	-0.9485	0.5991
6	-0.2304	1.0646	-1.3216	0.7145	-0.2547	1.1710	-1.4598	0.7907	-0.2061	0.9582	-1.1834	0.6382
7	-0.2381	1.0864	-1.3133	0.6617	-0.2666	1.2141	-1.4883	0.7574	-0.2095	0.9588	-1.1383	0.5661
8	-0.2404	1.0855	-1.2796	0.6081	-0.2770	1.2538	-1.5264	0.7451	-0.2038	0.9171	-1.0328	0.4712
9	-0.2538	1.1334	-1.3168	0.5922	-0.2915	1.3115	-1.5870	0.7448	-0.2162	0.9552	-1.0467	0.4396
10	-0.2509	1.1098	-1.2557	0.5362	-0.2873	1.2900	-1.5427	0.7043	-0.2145	0.9296	-0.9687	0.3681

D. Selective Assembly Scrap Coefficients

The following equation is a restatement of Equation 4-23. It is the closed-form equation for the percentage of scrap generated when using selective assembly to combine two normal distributions of parts, a Part A having a standard deviation of σ_A with a Part B having a standard deviation of σ_B. The coefficients, D_0, D_1, D_2, and D_3, are retrieved from the tables below for the specific technique and number of bins of interest.

$$Scrap_{(N)} = \left((D_3)\left(\frac{\sigma_A}{\sigma_B}\right)^3 + (D_2)\left(\frac{\sigma_A}{\sigma_B}\right)^2 + (D_1)\left(\frac{\sigma_A}{\sigma_B}\right) + D_0 \right) \quad \text{D-1}$$

for $1 \le \frac{\sigma_A}{\sigma_B} \le 2$

Table D-1: Equal Width Binning to Minimize Variation

	$Scrap_{SelAsm}$				Upper 95% Band				Lower 95% Band			
N	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0
1	-0.0614	0.3671	-0.5440	0.2404	-0.0832	0.4548	-0.6329	0.2661	-0.0396	0.2795	-0.4552	0.0215
2	-0.0635	0.3942	-0.6207	0.3025	-0.0719	0.4420	-0.6969	0.3557	-0.0551	0.3465	-0.5445	0.2492
3	-0.0181	-0.0206	0.4766	-0.4223	0.1217	-0.6963	1.5396	-0.9261	-0.1579	0.6551	-0.5864	0.0815
4	-0.1729	0.7261	-0.6678	0.1344	-0.1228	0.4712	-0.2387	-0.0663	-0.2230	0.9810	-1.0969	0.3352
5	0	0	0.2764	-0.2502	0	0	0.2879	-0.2341	0	0	0.2648	-0.2663
6	-0.0993	0.3686	-0.1379	-0.1044	0.0066	-0.1488	0.6902	-0.4982	-0.2051	0.8859	-0.9661	0.2894
7	-0.1622	0.6715	-0.5919	0.1148	-0.1449	0.5758	-0.4118	0.0422	-0.1795	0.7672	-0.7721	0.1873
8	-0.1322	0.5601	-0.4745	0.0826	-0.1109	0.4451	-0.2604	-0.0123	-0.1536	0.6750	-0.6886	0.1774
9	-0.1136	0.4591	-0.3075	-0.0006	-0.0532	0.1708	0.1530	-0.2096	-0.1739	0.7474	-0.7681	0.2085
10	-0.1494	0.6163	-0.5264	0.1002	-0.1080	0.4078	-0.1764	-0.0551	-0.1908	0.8248	-0.8764	0.2554

Table D-2: Equal Width Binning to Minimize Scrap

	Scrap,SelAsm				Upper 95% Band				Lower 95% Band			
N	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0
1	0	0	0	0.0021	0	0	0	0.0042	0	0	0	0
2	0	0	0	0.0118	0	0	0	0.0269	0	0	0	0
3	0	0	0	0.0167	0	0	0	0.0324	0	0	0	0.0010
4	0	0	0	0.0212	0	0	0	0.0437	0	0	0	0
5	0	0	0	0.0248	0	0	0	0.0497	0	0	0	0
6	0	0	0	0.0312	0	0	0	0.0527	0	0	0	0.0098
7	0	0	0	0.0349	0	0	0	0.0603	0	0	0	0.0095
8	0	0	0	0.0382	0	0	0	0.0632	0	0	0	0.0132
9	0	0	0	0.0420	0	0	0	0.0662	0	0	0	0.0179
10	0	0	0	0.0444	0	0	0	0.0708	0	0	0	0.0180

Table D-3: Equal Area Binning to Minimize Variation

	Scrap,SelAsm				Upper 95% Band				Lower 95% Band			
N	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0
1	-0.0578	0.3532	-0.5322	0.2389	-0.0750	0.4279	-0.6169	0.2687	-0.0407	0.2786	-0.4475	0.2092
2	-0.0061	0.1277	-0.2384	0.1407	0.0721	-0.2252	0.2604	-0.0497	-0.0843	0.4805	-0.7372	0.3312
3	-0.0234	0.0677	0.1093	-0.1239	-0.0004	-0.0521	0.3109	-0.1992	-0.0463	0.1876	-0.0924	-0.0486
4	-0.1901	0.7645	-0.7443	0.2028	-0.5964	2.5836	-3.3184	1.3958	0.2163	-1.0546	1.8299	-0.9903
5	-0.2304	0.9511	-0.9896	0.3064	-0.5201	2.1751	-2.6014	1.0099	0.0593	-0.2729	0.6222	-0.3971
6	-0.1691	0.7159	-0.6998	0.1927	-0.2545	1.1044	-1.2305	0.4519	-0.0837	0.3274	-0.1692	-0.0665
7	-0.1582	0.6639	-0.6108	0.1464	-0.2311	0.9922	-1.0503	0.3628	-0.0853	0.3356	-0.1712	-0.0700
8	-0.1840	0.7616	-0.7202	0.1862	-0.2442	0.9430	-0.8068	0.1744	-0.1238	0.5801	-0.6336	0.1979
9	-0.2113	0.9275	-1.0393	0.3848	-0.2193	0.9917	-1.1808	0.5410	-0.2032	0.8633	-0.8978	0.2287
10	-0.1831	0.7994	-0.8530	0.3009	-0.1678	0.7253	-0.7441	0.3145	-0.1984	0.8735	-0.9619	0.2873

Table D-4: Equal Area Binning to Minimize Scrap

	Scrap,SelAsm				Upper 95% Band				Lower 95% Band			
N	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0
1	0	0	0	0.0021	0	0	0	0.0044	0	0	0	0
2	0	0	0	0.0219	0	0	0	0.0559	0	0	0	0
3	0	0	0	0.0300	0	0	0	0.0594	0	0	0	0.0007
4	0	0	0	0.0335	0	0	0	0.0636	0	0	0	0.0035
5	0	0	0	0.0391	0	0	0	0.0689	0	0	0	0.0092
6	0	0	0	0.0438	0	0	0	0.0712	0	0	0	0.0165
7	0	0	0	0.0458	0	0	0	0.0732	0	0	0	0.0184
8	0	0	0	0.0504	0	0	0	0.0763	0	0	0	0.0244
9	0	0	0	0.0533	0	0	0	0.0803	0	0	0	0.0264
10	0	0	0	0.0583	0	0	0	0.0881	0	0	0	0.0286

Table D-5: Optimal Width Binning

	Scrap,SelAsm				Upper 95% Band				Lower 95% Band			
N	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0	D_3	D_2	D_1	D_0
1	0	0	0	0.0024	0	0	0	0.0055	0	0	0	-0.0007
2	0	0	0	0.0214	0	0	0	0.0551	0	0	0	-0.0123
3	0	0	0	0.0262	0	0	0	0.0549	0	0	0	-0.0025
4	0	0	0	0.0321	0	0	0	0.0603	0	0	0	0.0038
5	0	0	0	0.0388	0	0	0	0.0705	0	0	0	0.0071
6	0	0	0	0.0409	0	0	0	0.0702	0	0	0	0.0116
7	0	0	0	0.0448	0	0	0	0.0710	0	0	0	0.0187
8	0	0	0	0.0476	0	0	0	0.0725	0	0	0	0.0227
9	0	0	0	0.0528	0	0	0	0.0796	0	0	0	0.0260
10	0	0	0	0.0534	0	0	0	0.0816	0	0	0	0.0252

E. Selective Assembly Simulation Data

The selective assembly simulation data used to curve fit the third order polynomial equations are listed in this appendix. The data used to fit the curves for the standard deviation of an assembly created using selective assembly are listed in Table E-1 through Table E-5, and the data used to fit the curves for the percentage of scrap are listed in Table E-11 through Table E-15. Also, listed are the R^2 values representing the equations' goodness of fit in Table E-6 through Table E-10, Table E-16, and Table E-17.

Each value for the assembly standard deviation in the tables below is the average from fifty Monte Carlo simulation runs. A thousand numbers were used for each simulation run. The output from one simulation run was a value for the normalized assembly standard deviation. For example, fifty Monte Carlo simulation runs were conducted for the Equal Width Binning to Minimize Variation technique where two bins are used and the ratio σ_A/σ_B is 1.4 to produce fifty values for the normalized assembly standard deviation. The resulting normalized assembly standard deviations from the fifty runs were averaged to produce a value of 0.56604. The standard deviation of the fifty runs was also calculated and used with the average value to compute the values used to curve fit the ±95% Confidence Interval at a level of plus and minus two standard deviations from the average value, i.e., for the case mentioned above, the normalized assembly standard deviation has a 95% probability of falling between 0.59953 and 0.53255.

Table E-1: Data for curve fitting σ_C (Equal Width Binning to Minimize Variation)

Number of Bins	σ_A/σ_B	Normalized Assembly Standard Deviation σ_C	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.98974	1.03497	0.94451
	1.1	0.98427	1.02714	0.94140
	1.2	0.96986	1.01375	0.92596
	1.3	0.95875	1.00600	0.91150
	1.4	0.93599	0.98517	0.88681
	1.5	0.91612	0.96319	0.86905
	1.6	0.89354	0.93232	0.85475
	1.7	0.86743	0.90624	0.82863
	1.8	0.84740	0.89093	0.80386
	1.9	0.82416	0.87118	0.77714

	2.0	0.80076	0.84047	0.76104
2	1.0	0.58747	0.61255	0.56239
	1.1	0.58487	0.61560	0.55414
	1.2	0.58154	0.61390	0.54918
	1.3	0.57430	0.61163	0.53698
	1.4	0.56604	0.59953	0.53255
	1.5	0.55378	0.58245	0.52510
	1.6	0.54662	0.57536	0.51789
	1.7	0.53253	0.56120	0.50387
	1.8	0.51849	0.54633	0.49065
	1.9	0.50638	0.53509	0.47766
	2.0	0.49299	0.51931	0.46668
3	1.0	0.50008	0.52534	0.47483
	1.1	0.48871	0.51097	0.46645
	1.2	0.47046	0.49740	0.44352
	1.3	0.45231	0.47474	0.42987
	1.4	0.43432	0.45717	0.41146
	1.5	0.41828	0.44130	0.39526
	1.6	0.40332	0.42457	0.38207
	1.7	0.38865	0.41109	0.36620
	1.8	0.37455	0.40044	0.34865
	1.9	0.36186	0.38569	0.33803
	2.0	0.34827	0.36653	0.33001
4	1.0	0.39415	0.41539	0.37291
	1.1	0.38403	0.40432	0.36375
	1.2	0.36907	0.38659	0.35156
	1.3	0.35366	0.36937	0.33796
	1.4	0.33950	0.35666	0.32234
	1.5	0.32498	0.34115	0.30882
	1.6	0.31183	0.33109	0.29257
	1.7	0.29874	0.31824	0.27923
	1.8	0.28751	0.30505	0.26997
	1.9	0.27760	0.29536	0.25984
	2.0	0.26748	0.28379	0.25118
5	1.0	0.32693	0.34476	0.30909
	1.1	0.31538	0.33250	0.29826
	1.2	0.30230	0.31843	0.28617
	1.3	0.28951	0.30645	0.27256
	1.4	0.27641	0.29320	0.25962
	1.5	0.26386	0.27937	0.24836
	1.6	0.25354	0.26851	0.23856
	1.7	0.24321	0.25737	0.22905
	1.8	0.23389	0.24866	0.21911
	1.9	0.22395	0.23658	0.21132
	2.0	0.21760	0.23039	0.20482
6	1.0	0.27548	0.29234	0.25862
	1.1	0.26659	0.28105	0.25214
	1.2	0.25372	0.26852	0.23891

	1.3	0.24222	0.25480	0.22964
	1.4	0.23194	0.24407	0.21981
	1.5	0.22229	0.23363	0.21095
	1.6	0.21343	0.22606	0.20081
	1.7	0.20465	0.21731	0.19198
	1.8	0.19705	0.20771	0.18639
	1.9	0.18894	0.19921	0.17866
	2.0	0.18119	0.19245	0.16994
7	1.0	0.24000	0.25816	0.22183
	1.1	0.23046	0.24402	0.21690
	1.2	0.22066	0.23581	0.20552
	1.3	0.21135	0.22483	0.19786
	1.4	0.20094	0.21224	0.18963
	1.5	0.19338	0.20645	0.18031
	1.6	0.18414	0.19572	0.17257
	1.7	0.17626	0.18693	0.16558
	1.8	0.16945	0.18148	0.15742
	1.9	0.16315	0.17410	0.15221
	2.0	0.15720	0.16792	0.14648
8	1.0	0.20916	0.22499	0.19333
	1.1	0.20252	0.21097	0.19406
	1.2	0.19337	0.20290	0.18383
	1.3	0.18447	0.19475	0.17419
	1.4	0.17597	0.18624	0.16570
	1.5	0.16854	0.17828	0.15880
	1.6	0.16063	0.16975	0.15151
	1.7	0.15479	0.16579	0.14378
	1.8	0.14803	0.15817	0.13789
	1.9	0.14143	0.15108	0.13178
	2.0	0.13616	0.14525	0.12707
9	1.0	0.18919	0.20590	0.17249
	1.1	0.18153	0.19153	0.17154
	1.2	0.17315	0.18274	0.16356
	1.3	0.16531	0.17551	0.15510
	1.4	0.15795	0.16645	0.14946
	1.5	0.15051	0.15944	0.14158
	1.6	0.14446	0.15369	0.13522
	1.7	0.13807	0.14706	0.12909
	1.8	0.13380	0.14334	0.12426
	1.9	0.12809	0.13580	0.12038
	2.0	0.12230	0.13119	0.11341
10	1.0	0.17019	0.18744	0.15293
	1.1	0.16385	0.17296	0.15474
	1.2	0.15684	0.16577	0.14791
	1.3	0.14933	0.15671	0.14195
	1.4	0.14174	0.15032	0.13316
	1.5	0.13530	0.14426	0.12634
	1.6	0.12959	0.13806	0.12111

	1.7	0.12460	0.13400	0.11520
	1.8	0.11968	0.12940	0.10996
	1.9	0.11523	0.12296	0.10751
	2.0	0.11019	0.11714	0.10325

Table E-2: Data for curve fitting σ_C (Equal Width Binning to Minimize Scrap)

Number of Bins	σ_A/σ_B	Normalized Assembly Standard Deviation σ_C	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.99107	1.03789	0.94426
	1.1	0.98576	1.03171	0.93982
	1.2	0.99110	1.03647	0.94573
	1.3	0.98690	1.03392	0.93988
	1.4	0.99092	1.03797	0.94386
	1.5	0.98733	1.03532	0.93933
	1.6	0.98939	1.03834	0.94043
	1.7	0.98906	1.03691	0.94122
	1.8	0.98841	1.03668	0.94013
	1.9	0.99125	1.03884	0.94366
	2.0	0.98651	1.03555	0.93747
2	1.0	0.58921	0.62026	0.55816
	1.1	0.59051	0.62131	0.55972
	1.2	0.59808	0.63155	0.56462
	1.3	0.60646	0.64060	0.57232
	1.4	0.61948	0.65723	0.58172
	1.5	0.62784	0.66662	0.58906
	1.6	0.64270	0.68291	0.60248
	1.7	0.65324	0.69441	0.61208
	1.8	0.66533	0.70873	0.62193
	1.9	0.67792	0.72165	0.63419
	2.0	0.68634	0.73031	0.64237
3	1.0	0.50605	0.53060	0.48151
	1.1	0.50785	0.53517	0.48052
	1.2	0.51693	0.54771	0.48616
	1.3	0.52683	0.56035	0.49332
	1.4	0.54352	0.58211	0.50494
	1.5	0.55619	0.59533	0.51705
	1.6	0.57308	0.61582	0.53034
	1.7	0.58816	0.63165	0.54466
	1.8	0.60227	0.64870	0.55584
	1.9	0.61881	0.66532	0.57230
	2.0	0.62990	0.67948	0.58032
4	1.0	0.40139	0.42525	0.37754
	1.1	0.40300	0.42958	0.37642
	1.2	0.41710	0.44973	0.38448

	1.3	0.43017	0.46413	0.39620
	1.4	0.45240	0.49252	0.41227
	1.5	0.46924	0.51005	0.42844
	1.6	0.49156	0.53549	0.44763
	1.7	0.51190	0.55714	0.46666
	1.8	0.52901	0.57518	0.48284
	1.9	0.55154	0.59998	0.50311
	2.0	0.56448	0.61389	0.51508
5	1.0	0.33065	0.34884	0.31246
	1.1	0.33535	0.35525	0.31545
	1.2	0.35072	0.37806	0.32337
	1.3	0.36897	0.39831	0.33963
	1.4	0.39452	0.43001	0.35904
	1.5	0.41579	0.45200	0.37958
	1.6	0.44200	0.48159	0.40242
	1.7	0.46503	0.50717	0.42290
	1.8	0.48683	0.53053	0.44312
	1.9	0.50985	0.55559	0.46411
	2.0	0.52622	0.57244	0.48001
6	1.0	0.28141	0.29873	0.26410
	1.1	0.28741	0.30702	0.26780
	1.2	0.30559	0.33360	0.27758
	1.3	0.32747	0.35883	0.29610
	1.4	0.35667	0.39417	0.31916
	1.5	0.38032	0.41930	0.34133
	1.6	0.41040	0.45332	0.36748
	1.7	0.43525	0.48013	0.39036
	1.8	0.45968	0.50603	0.41334
	1.9	0.48392	0.53175	0.43608
	2.0	0.50266	0.55186	0.45346
7	1.0	0.24410	0.26010	0.22811
	1.1	0.25085	0.27143	0.23027
	1.2	0.27132	0.30157	0.24107
	1.3	0.29541	0.33104	0.25978
	1.4	0.32755	0.36795	0.28716
	1.5	0.35368	0.39635	0.31101
	1.6	0.38539	0.42983	0.34096
	1.7	0.41265	0.45984	0.36546
	1.8	0.43749	0.48430	0.39068
	1.9	0.46451	0.51476	0.41425
	2.0	0.48368	0.53341	0.43394
8	1.0	0.21440	0.23257	0.19624
	1.1	0.22276	0.24636	0.19917
	1.2	0.24621	0.28146	0.21095
	1.3	0.27358	0.31148	0.23567
	1.4	0.30873	0.35298	0.26448
	1.5	0.33742	0.38181	0.29304
	1.6	0.37065	0.41845	0.32286

	1.7	0.40043	0.44866	0.35220
	1.8	0.42621	0.47571	0.37670
	1.9	0.45485	0.50603	0.40367
	2.0	0.47451	0.52626	0.42277
9	1.0	0.19321	0.21052	0.17590
	1.1	0.20291	0.22560	0.18022
	1.2	0.22819	0.26253	0.19385
	1.3	0.25812	0.29529	0.22095
	1.4	0.29491	0.33873	0.25109
	1.5	0.32499	0.36865	0.28132
	1.6	0.36005	0.40775	0.31236
	1.7	0.39019	0.43784	0.34255
	1.8	0.41758	0.46692	0.36824
	1.9	0.44582	0.49507	0.39658
	2.0	0.46677	0.51807	0.41546
10	1.0	0.17401	0.19088	0.15715
	1.1	0.18413	0.20774	0.16053
	1.2	0.21184	0.24805	0.17564
	1.3	0.24328	0.28256	0.20399
	1.4	0.28267	0.32692	0.23843
	1.5	0.31339	0.35885	0.26793
	1.6	0.35010	0.39798	0.30222
	1.7	0.38079	0.42977	0.33181
	1.8	0.40916	0.45897	0.35934
	1.9	0.43785	0.48906	0.38663
	2.0	0.45903	0.51074	0.40732

Table E-3: Data for curve fitting σ_C (Equal Area Binning to Minimize Variation)

Number of Bins	σ_A/σ_B	Normalized Assembly Standard Deviation σ_C	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.99095	1.03049	0.95140
	1.1	0.98582	1.03242	0.93922
	1.2	0.97527	1.02077	0.92976
	1.3	0.96569	1.01088	0.92050
	1.4	0.94662	0.99385	0.89939
	1.5	0.92454	0.95683	0.89225
	1.6	0.90384	0.94302	0.86465
	1.7	0.88069	0.91836	0.84302
	1.8	0.85745	0.89050	0.82440
	1.9	0.83469	0.86766	0.80172
	2.0	0.81236	0.84199	0.78274
2	1.0	0.59263	0.61496	0.57031
	1.1	0.58714	0.61312	0.56116
	1.2	0.58147	0.60629	0.55664

	1.3	0.58062	0.61376	0.54748
	1.4	0.57455	0.60761	0.54148
	1.5	0.56382	0.59613	0.53150
	1.6	0.55200	0.58031	0.52368
	1.7	0.54087	0.56540	0.51633
	1.8	0.52886	0.55397	0.50375
	1.9	0.51814	0.54342	0.49285
	2.0	0.50767	0.53285	0.48249
3	1.0	0.44446	0.46717	0.42176
	1.1	0.44378	0.46795	0.41961
	1.2	0.44562	0.47307	0.41818
	1.3	0.44454	0.46961	0.41947
	1.4	0.44095	0.46656	0.41533
	1.5	0.43664	0.46622	0.40706
	1.6	0.43224	0.45468	0.40981
	1.7	0.42356	0.44562	0.40150
	1.8	0.41698	0.43840	0.39557
	1.9	0.40896	0.43081	0.38710
	2.0	0.39985	0.42182	0.37788
4	1.0	0.36314	0.38685	0.33942
	1.1	0.36599	0.38745	0.34452
	1.2	0.36895	0.39189	0.34600
	1.3	0.36690	0.39198	0.34182
	1.4	0.36618	0.39103	0.34133
	1.5	0.36324	0.38396	0.34251
	1.6	0.35919	0.38349	0.33489
	1.7	0.35630	0.37922	0.33339
	1.8	0.34774	0.36723	0.32825
	1.9	0.34211	0.36207	0.32216
	2.0	0.33709	0.35813	0.31605
5	1.0	0.30908	0.32917	0.28898
	1.1	0.31274	0.33312	0.29235
	1.2	0.31603	0.33602	0.29605
	1.3	0.31639	0.33688	0.29590
	1.4	0.31465	0.33372	0.29558
	1.5	0.31425	0.33470	0.29381
	1.6	0.31059	0.33086	0.29031
	1.7	0.30560	0.32499	0.28621
	1.8	0.30075	0.32150	0.28000
	1.9	0.29803	0.31589	0.28018
	2.0	0.29192	0.31363	0.27020
6	1.0	0.26837	0.30253	0.23420
	1.1	0.27495	0.29950	0.25039
	1.2	0.27592	0.30267	0.24917
	1.3	0.27612	0.30001	0.25223
	1.4	0.27588	0.30640	0.24536
	1.5	0.27570	0.29952	0.25189
	1.6	0.27167	0.29908	0.24427

	1.7	0.27012	0.29764	0.24261
	1.8	0.26483	0.28839	0.24127
	1.9	0.26179	0.28688	0.23670
	2.0	0.25638	0.28413	0.22863
7	1.0	0.24116	0.27570	0.20662
	1.1	0.24755	0.27469	0.22041
	1.2	0.24696	0.27261	0.22132
	1.3	0.24775	0.27085	0.22465
	1.4	0.24877	0.27709	0.22045
	1.5	0.24716	0.27437	0.21995
	1.6	0.24497	0.27123	0.21871
	1.7	0.24088	0.26637	0.21539
	1.8	0.23645	0.26239	0.21050
	1.9	0.23527	0.26079	0.20974
	2.0	0.23062	0.25896	0.20227
8	1.0	0.21772	0.24998	0.18547
	1.1	0.22436	0.25128	0.19744
	1.2	0.22451	0.25180	0.19723
	1.3	0.22605	0.25082	0.20128
	1.4	0.22529	0.25457	0.19601
	1.5	0.22359	0.24853	0.19865
	1.6	0.22121	0.24800	0.19442
	1.7	0.21786	0.24603	0.18968
	1.8	0.21560	0.24209	0.18911
	1.9	0.21358	0.23950	0.18766
	2.0	0.20917	0.23796	0.18038
9	1.0	0.20153	0.23519	0.16787
	1.1	0.20435	0.22958	0.17913
	1.2	0.20553	0.23235	0.17871
	1.3	0.20720	0.23556	0.17883
	1.4	0.20645	0.23626	0.17664
	1.5	0.20482	0.22975	0.17990
	1.6	0.20125	0.22995	0.17255
	1.7	0.19937	0.22754	0.17120
	1.8	0.19691	0.22331	0.17051
	1.9	0.19381	0.22012	0.16749
	2.0	0.19006	0.21780	0.16231
10	1.0	0.18522	0.21835	0.15210
	1.1	0.19030	0.21344	0.16717
	1.2	0.19109	0.21757	0.16461
	1.3	0.19171	0.21948	0.16395
	1.4	0.19044	0.22334	0.15754
	1.5	0.19003	0.21458	0.16548
	1.6	0.18629	0.21358	0.15901
	1.7	0.18485	0.21262	0.15707
	1.8	0.18276	0.20787	0.15764
	1.9	0.17972	0.20471	0.15473
	2.0	0.17601	0.20561	0.14641

Table E-4: Data for curve fitting σ_C (Equal Area Binning to Minimize Scrap)

Number of Bins	σ_A/σ_B	Normalized Assembly Standard Deviation σ_C	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.99330	1.04126	0.94534
	1.1	0.99058	1.03574	0.94542
	1.2	0.99224	1.03781	0.94668
	1.3	0.99115	1.03819	0.94410
	1.4	0.99298	1.04043	0.94552
	1.5	0.99094	1.04007	0.94181
	1.6	0.99207	1.03904	0.94511
	1.7	0.99132	1.04195	0.94069
	1.8	0.99073	1.03742	0.94405
	1.9	0.99297	1.04384	0.94210
	2.0	0.98899	1.03612	0.94187
2	1.0	0.59324	0.62097	0.56550
	1.1	0.59399	0.62314	0.56485
	1.2	0.60103	0.63296	0.56910
	1.3	0.60890	0.64122	0.57659
	1.4	0.62186	0.65880	0.58493
	1.5	0.63096	0.66640	0.59552
	1.6	0.64384	0.68200	0.60568
	1.7	0.65543	0.69393	0.61693
	1.8	0.66635	0.70640	0.62630
	1.9	0.67963	0.71957	0.63970
	2.0	0.68810	0.72901	0.64719
3	1.0	0.44368	0.46644	0.42093
	1.1	0.44733	0.47029	0.42436
	1.2	0.45731	0.48445	0.43016
	1.3	0.47027	0.49890	0.44163
	1.4	0.48952	0.52279	0.45625
	1.5	0.50509	0.53795	0.47224
	1.6	0.52525	0.56106	0.48945
	1.7	0.54186	0.57807	0.50564
	1.8	0.55946	0.59802	0.52090
	1.9	0.57782	0.61817	0.53747
	2.0	0.59125	0.63232	0.55018
4	1.0	0.36363	0.38512	0.34213
	1.1	0.36663	0.38990	0.34335
	1.2	0.38187	0.40647	0.35726
	1.3	0.39811	0.42531	0.37092
	1.4	0.42252	0.45439	0.39065
	1.5	0.44259	0.47496	0.41022
	1.6	0.46676	0.50306	0.43045

	1.7	0.48773	0.52173	0.45374
	1.8	0.50885	0.54811	0.46959
	1.9	0.53033	0.56820	0.49246
	2.0	0.54662	0.58842	0.50481
5	1.0	0.30971	0.32802	0.29141
	1.1	0.31623	0.33666	0.29581
	1.2	0.33230	0.35339	0.31121
	1.3	0.35457	0.37728	0.33186
	1.4	0.37976	0.40745	0.35207
	1.5	0.40485	0.43189	0.37781
	1.6	0.43154	0.46203	0.40105
	1.7	0.45586	0.48615	0.42557
	1.8	0.47957	0.51243	0.44671
	1.9	0.50223	0.53428	0.47019
	2.0	0.52201	0.55657	0.48745
6	1.0	0.27228	0.29438	0.25017
	1.1	0.27906	0.30374	0.25439
	1.2	0.29812	0.32526	0.27098
	1.3	0.32341	0.35404	0.29278
	1.4	0.35211	0.38528	0.31894
	1.5	0.37994	0.41465	0.34524
	1.6	0.40824	0.44317	0.37330
	1.7	0.43536	0.47202	0.39871
	1.8	0.45962	0.49608	0.42316
	1.9	0.48494	0.52202	0.44786
	2.0	0.50484	0.54407	0.46562
7	1.0	0.24597	0.26599	0.22595
	1.1	0.25429	0.27688	0.23170
	1.2	0.27448	0.29987	0.24908
	1.3	0.30172	0.32928	0.27416
	1.4	0.33278	0.36362	0.30194
	1.5	0.36198	0.39235	0.33161
	1.6	0.39286	0.42486	0.36087
	1.7	0.42047	0.45212	0.38882
	1.8	0.44671	0.47895	0.41447
	1.9	0.47287	0.50605	0.43970
	2.0	0.49396	0.52910	0.45882
8	1.0	0.22325	0.24228	0.20422
	1.1	0.23211	0.25401	0.21022
	1.2	0.25468	0.27787	0.23149
	1.3	0.28401	0.31152	0.25650
	1.4	0.31688	0.34539	0.28836
	1.5	0.34776	0.37865	0.31687
	1.6	0.37914	0.41054	0.34773
	1.7	0.40852	0.44022	0.37682
	1.8	0.43566	0.46853	0.40279
	1.9	0.46312	0.49579	0.43046
	2.0	0.48455	0.51989	0.44920

9	1.0	0.20499	0.22058	0.18940
	1.1	0.21462	0.23520	0.19404
	1.2	0.23889	0.26085	0.21694
	1.3	0.27019	0.29641	0.24396
	1.4	0.30471	0.33176	0.27766
	1.5	0.33665	0.36575	0.30755
	1.6	0.36975	0.39871	0.34079
	1.7	0.39980	0.42918	0.37042
	1.8	0.42751	0.45771	0.39731
	1.9	0.45515	0.48491	0.42538
	2.0	0.47751	0.51023	0.44479
10	1.0	0.19061	0.20808	0.17314
	1.1	0.20091	0.22137	0.18044
	1.2	0.22691	0.24833	0.20549
	1.3	0.25959	0.28409	0.23509
	1.4	0.29478	0.32095	0.26861
	1.5	0.32881	0.35584	0.30178
	1.6	0.36187	0.38967	0.33407
	1.7	0.39336	0.42109	0.36563
	1.8	0.42136	0.45007	0.39266
	1.9	0.44938	0.47691	0.42185
	2.0	0.47245	0.50320	0.44171

Table E-5: Data for curve fitting σ_C (Optimal Width Binning)

Number of Bins	σ_A/σ_B	Normalized Assembly Standard Deviation σ_C	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.98737	1.02634	0.94840
	1.1	0.98759	1.02999	0.94520
	1.2	0.98857	1.02725	0.94989
	1.3	0.98740	1.02523	0.94956
	1.4	0.98815	1.03041	0.94589
	1.5	0.98839	1.03322	0.94356
	1.6	0.98774	1.02848	0.94700
	1.7	0.98890	1.03456	0.94324
	1.8	0.98916	1.03517	0.94315
	1.9	0.98919	1.03423	0.94416
	2.0	0.98857	1.02879	0.94835
2	1.0	0.58851	0.61841	0.55862
	1.1	0.59284	0.62223	0.56346
	1.2	0.59735	0.63138	0.56333
	1.3	0.60485	0.63980	0.56990
	1.4	0.61610	0.65314	0.57906
	1.5	0.63117	0.66618	0.59616
	1.6	0.64051	0.67811	0.60291

	1.7	0.65586	0.69353	0.61820
	1.8	0.66809	0.70690	0.62928
	1.9	0.67849	0.72116	0.63581
	2.0	0.68822	0.73008	0.64637
3	1.0	0.42352	0.44363	0.40342
	1.1	0.42696	0.44859	0.40533
	1.2	0.44013	0.46000	0.42026
	1.3	0.45486	0.47730	0.43242
	1.4	0.47157	0.49491	0.44824
	1.5	0.49022	0.51940	0.46104
	1.6	0.51053	0.53653	0.48453
	1.7	0.52901	0.56146	0.49656
	1.8	0.54773	0.58151	0.51395
	1.9	0.56598	0.59944	0.53252
	2.0	0.58396	0.61710	0.55082
4	1.0	0.33156	0.35132	0.31179
	1.1	0.33849	0.36038	0.31660
	1.2	0.35399	0.37793	0.33006
	1.3	0.37415	0.40118	0.34711
	1.4	0.39628	0.42331	0.36924
	1.5	0.42270	0.45395	0.39144
	1.6	0.44584	0.47641	0.41527
	1.7	0.47108	0.50525	0.43691
	1.8	0.49389	0.52915	0.45863
	1.9	0.51446	0.54943	0.47949
	2.0	0.53581	0.57074	0.50089
5	1.0	0.27207	0.28979	0.25436
	1.1	0.27917	0.30033	0.25800
	1.2	0.29936	0.31912	0.27960
	1.3	0.32316	0.34649	0.29982
	1.4	0.35036	0.37184	0.32889
	1.5	0.37813	0.40699	0.34926
	1.6	0.40604	0.43320	0.37887
	1.7	0.43285	0.46403	0.40166
	1.8	0.45825	0.49032	0.42618
	1.9	0.48216	0.51274	0.45158
	2.0	0.50549	0.53605	0.47493
6	1.0	0.22952	0.24837	0.21067
	1.1	0.23938	0.26182	0.21695
	1.2	0.26147	0.28340	0.23954
	1.3	0.28952	0.31376	0.26528
	1.4	0.31975	0.34013	0.29936
	1.5	0.35240	0.38071	0.32410
	1.6	0.38246	0.40935	0.35558
	1.7	0.41195	0.44208	0.38182
	1.8	0.43919	0.47006	0.40833
	1.9	0.46430	0.49541	0.43319
	2.0	0.48842	0.51894	0.45791

7	1.0	0.20001	0.21837	0.18164
	1.1	0.21102	0.23232	0.18973
	1.2	0.23617	0.25730	0.21504
	1.3	0.26833	0.29110	0.24555
	1.4	0.30071	0.31752	0.28389
	1.5	0.33559	0.36152	0.30966
	1.6	0.36811	0.39352	0.34271
	1.7	0.39844	0.42597	0.37092
	1.8	0.42692	0.45511	0.39873
	1.9	0.45355	0.48126	0.42585
	2.0	0.47875	0.50633	0.45118
8	1.0	0.17729	0.19763	0.15694
	1.1	0.18935	0.21221	0.16650
	1.2	0.21774	0.23832	0.19716
	1.3	0.25167	0.27249	0.23085
	1.4	0.28632	0.30342	0.26923
	1.5	0.32263	0.34619	0.29907
	1.6	0.35649	0.37930	0.33367
	1.7	0.38781	0.41251	0.36310
	1.8	0.41712	0.44237	0.39187
	1.9	0.44414	0.46907	0.41921
	2.0	0.47032	0.49387	0.44676
9	1.0	0.15880	0.17974	0.13786
	1.1	0.17246	0.19611	0.14881
	1.2	0.20231	0.22323	0.18138
	1.3	0.23882	0.25938	0.21825
	1.4	0.27565	0.29031	0.26099
	1.5	0.31379	0.33699	0.29059
	1.6	0.34813	0.36948	0.32679
	1.7	0.38091	0.40510	0.35672
	1.8	0.41091	0.43562	0.38621
	1.9	0.43813	0.46163	0.41464
	2.0	0.46411	0.48723	0.44100
10	1.0	0.14363	0.16642	0.12083
	1.1	0.15851	0.18329	0.13374
	1.2	0.19062	0.21217	0.16908
	1.3	0.22888	0.24973	0.20803
	1.4	0.26744	0.28036	0.25453
	1.5	0.30626	0.32849	0.28404
	1.6	0.34165	0.36275	0.32056
	1.7	0.37479	0.39791	0.35168
	1.8	0.40529	0.42894	0.38164
	1.9	0.43323	0.45618	0.41028
	2.0	0.45943	0.48270	0.43616

The closed-form equations for the assembly standard deviation and percentage of scrap developed in Section 4.4.1 can be tested by quantifying the goodness of fit for the third order polynomial equations to the simulated data and checking to see if the equations accurately represent the data. The goodness of fit that a curve has to a group of data is represented by R^2. R^2 can have a value between zero and one. A value of zero indicates that a horizontal line passing through the average values of Y fits the data as well as the "best-fit" line. A value of one indicates all the data points lie exactly on the line. A value close to one is preferred but should not be the only test as to how well the fit curve represents the data. A simple visual test provides insight into the accuracy of the fit curve (1999). The generated curves do accurately reflect the shape and trends of the data. The R^2 values for the normalized assembly standard deviation curves are all very close to the desired value of one and are listed in Table E-6 through Table E-10. The R^2 values for the percentage of scrap curves that were fit using third order polynomials are also very close to the desired value of one and are listed in Table E-16 and Table E-17.

Table E-6: Goodness of fit for σ_C (Equal Width Binning to Minimize Variation)

Number of Bins	R^2 for Normalized σ_C	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	0.999	0.996	0.998
2	0.999	0.996	0.996
3	0.999	0.997	0.998
4	1.000	0.999	1.000
5	1.000	1.000	1.000
6	0.999	0.999	0.998
7	1.000	0.998	0.999
8	0.999	0.998	0.995
9	0.999	0.999	0.996
10	0.999	0.999	0.994

Table E-7: Goodness of fit for σ_C (Equal Width Binning to Minimize Scrap)

Number of Bins	R^2 for Normalized σ_C	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	n/a	n/a	n/a
2	0.999	0.999	0.999
3	0.999	0.999	0.999
4	0.999	0.999	1.000
5	1.000	0.999	1.000
6	0.999	0.999	1.000
7	0.999	0.999	0.999
8	0.999	0.999	0.999
9	0.999	0.999	0.999
10	0.999	0.999	0.999

Table E-8: Goodness of fit for σ_C (Equal Area Binning to Minimize Variation)

Number of Bins	R^2 for Normalized σ_C	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	1.000	0.997	0.997
2	0.996	0.984	0.999
3	0.998	0.990	0.984
4	0.993	0.974	0.975
5	0.993	0.985	0.975
6	0.985	0.888	0.824
7	0.974	0.879	0.918
8	0.981	0.944	0.887
9	0.991	0.882	0.870
10	0.983	0.799	0.654

Table E-9: Goodness of fit for σ_C (Equal Area Binning to Minimize Scrap)

Number of Bins	R^2 for Normalized σ_C	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	n/a	n/a	n/a
2	0.999	0.998	0.999
3	0.999	0.999	1.000
4	0.999	0.999	0.999
5	1.000	0.999	1.000
6	0.999	0.999	1.000
7	0.999	0.999	1.000
8	0.999	0.999	0.999
9	0.999	0.999	0.999
10	0.999	0.999	0.999

Table E-10: Goodness of fit for σ_C (Optimal Width Binning)

Number of Bins	R^2 for Normalized σ_C	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	n/a	n/a	n/a
2	0.999	0.999	0.996
3	1.000	0.999	0.998
4	1.000	1.000	1.000
5	0.999	0.999	0.999
6	0.999	0.999	0.999
7	0.999	0.999	0.998
8	0.999	0.999	0.998
9	0.999	0.999	0.998
10	0.999	0.999	0.998

Each value for the percentage of scrap in the tables below is the average from fifty Monte Carlo simulation runs. A thousand numbers were used for each simulation run. The output from one simulation run was a value for the percentage of scrap. For example, fifty Monte Carlo simulation runs were conducted for the Equal Width Binning to Minimize Variation technique where two bins are used and the ratio σ_A/σ_B is 1.4 to produce fifty values for the percentage of scrap. The resulting values for the percentage of scrap from the fifty runs were averaged to produce a value of 0.03202. The standard deviation of the fifty runs was also calculated and used with the average value to compute the values used to curve fit the ±95% Confidence Interval at a level of plus and minus two standard deviations from the average value, i.e., for the case mentioned above, the percentage of scrap has a 95% probability of falling between 0.04906 and 0.01498.

Table E-11: Data for curve fitting Scrap (Equal Width Binning to Minimize Variation)

Number of Bins	σ_A/σ_B	Percentage of Scrap	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.00272	0.00552	-0.00008
	1.1	0.00402	0.00817	-0.00013
	1.2	0.00964	0.01820	0.00108
	1.3	0.01840	0.02880	0.00800
	1.4	0.03104	0.04409	0.01799
	1.5	0.04374	0.05938	0.02810
	1.6	0.05818	0.07600	0.04036
	1.7	0.07560	0.09718	0.05402
	1.8	0.09158	0.11517	0.06799
	1.9	0.11124	0.13355	0.08893
	2.0	0.13024	0.15475	0.10573
2	1.0	0.01236	0.02790	-0.00318
	1.1	0.01260	0.02980	-0.00460
	1.2	0.01572	0.03215	-0.00071
	1.3	0.02172	0.03890	0.00454
	1.4	0.03202	0.04906	0.01498
	1.5	0.04440	0.06098	0.02782
	1.6	0.05838	0.07640	0.04036
	1.7	0.07558	0.09721	0.05395
	1.8	0.09158	0.11512	0.06804
	1.9	0.11074	0.13382	0.08766
	2.0	0.13032	0.15468	0.10596
3	1.0	0.01868	0.04017	-0.00281
	1.1	0.04766	0.08403	0.01129
	1.2	0.08838	0.12959	0.04717
	1.3	0.12400	0.16716	0.08084
	1.4	0.15500	0.19741	0.11259

	1.5	0.18714	0.22925	0.14503
	1.6	0.21430	0.25169	0.17691
	1.7	0.23802	0.27600	0.20004
	1.8	0.26158	0.29958	0.22358
	1.9	0.28360	0.31936	0.24784
	2.0	0.30486	0.34155	0.26817
4	1.0	0.02388	0.04644	0.00132
	1.1	0.04212	0.07213	0.01211
	1.2	0.07856	0.11439	0.04273
	1.3	0.11522	0.15021	0.08023
	1.4	0.14976	0.18800	0.11152
	1.5	0.18422	0.22021	0.14823
	1.6	0.21818	0.25654	0.17982
	1.7	0.24802	0.28717	0.20887
	1.8	0.27392	0.31231	0.23553
	1.9	0.29854	0.33660	0.26048
	2.0	0.32250	0.36039	0.28461
5	1.0	0.02688	0.04995	0.00381
	1.1	0.04864	0.07917	0.01811
	1.2	0.08032	0.11550	0.04514
	1.3	0.11014	0.14186	0.07842
	1.4	0.13664	0.16784	0.10544
	1.5	0.16706	0.20270	0.13142
	1.6	0.19678	0.22911	0.16445
	1.7	0.22328	0.25768	0.18888
	1.8	0.24780	0.28084	0.21476
	1.9	0.27210	0.31068	0.23352
	2.0	0.29834	0.34030	0.25638
6	1.0	0.03220	0.05439	0.01001
	1.1	0.04971	0.08151	0.01792
	1.2	0.08764	0.12793	0.04734
	1.3	0.12169	0.16257	0.08080
	1.4	0.15570	0.19697	0.11443
	1.5	0.18727	0.22791	0.14663
	1.6	0.21043	0.25036	0.17050
	1.7	0.23864	0.27689	0.20039
	1.8	0.25848	0.29762	0.21934
	1.9	0.28204	0.32182	0.24226
	2.0	0.30212	0.34058	0.26366
7	1.0	0.03643	0.06575	0.00711
	1.1	0.05491	0.08752	0.02231
	1.2	0.08757	0.12452	0.05063
	1.3	0.12458	0.15978	0.08939
	1.4	0.16077	0.20259	0.11894
	1.5	0.19440	0.23585	0.15294
	1.6	0.21989	0.26116	0.17862
	1.7	0.25434	0.29580	0.21288
	1.8	0.27474	0.31656	0.23292

	1.9	0.29998	0.34502	0.25494
	2.0	0.32194	0.36484	0.27904
8	1.0	0.04027	0.06537	0.01517
	1.1	0.05589	0.08645	0.02534
	1.2	0.08940	0.12292	0.05589
	1.3	0.12188	0.15870	0.08505
	1.4	0.15662	0.19504	0.11820
	1.5	0.18762	0.22519	0.15005
	1.6	0.21456	0.25539	0.17373
	1.7	0.24564	0.28846	0.20282
	1.8	0.26776	0.31045	0.22507
	1.9	0.29470	0.33860	0.25080
	2.0	0.31806	0.36194	0.27418
9	1.0	0.04232	0.06716	0.01748
	1.1	0.05912	0.08557	0.03267
	1.2	0.09238	0.12555	0.05921
	1.3	0.12638	0.16290	0.08986
	1.4	0.16070	0.19694	0.12446
	1.5	0.19167	0.22922	0.15411
	1.6	0.21732	0.25420	0.18044
	1.7	0.24646	0.28226	0.21066
	1.8	0.26640	0.30446	0.22834
	1.9	0.29146	0.33219	0.25073
	2.0	0.31540	0.35679	0.27401
10	1.0	0.04540	0.07248	0.01832
	1.1	0.06152	0.09501	0.02803
	1.2	0.09572	0.13222	0.05922
	1.3	0.12884	0.16546	0.09222
	1.4	0.16476	0.20504	0.12448
	1.5	0.19700	0.23760	0.15640
	1.6	0.22322	0.26506	0.18138
	1.7	0.25352	0.29136	0.21568
	1.8	0.27328	0.31483	0.23173
	1.9	0.29850	0.34070	0.25630
	2.0	0.32024	0.36154	0.27894

Table E-12: Data for curve fitting Scrap (Equal Width Binning to Minimize Scrap)

Number of Bins	σ_A/σ_B	Percentage of Scrap	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.00216	0.00435	-0.00002
	1.1	0.00209	0.00433	-0.00016
	1.2	0.00211	0.00414	0.00009
	1.3	0.00213	0.00439	-0.00014
	1.4	0.00206	0.00416	-0.00003
	1.5	0.00210	0.00423	-0.00003

	1.6	0.00201	0.00416	-0.00014
	1.7	0.00205	0.00406	0.00004
	1.8	0.00205	0.00421	-0.00011
	1.9	0.00215	0.00427	0.00004
	2.0	0.00209	0.00437	-0.00019
2	1.0	0.01214	0.02773	-0.00345
	1.1	0.01154	0.02597	-0.00288
	1.2	0.01182	0.02686	-0.00321
	1.3	0.01108	0.02481	-0.00266
	1.4	0.01186	0.02709	-0.00337
	1.5	0.01236	0.02764	-0.00292
	1.6	0.01158	0.02676	-0.00360
	1.7	0.01223	0.02763	-0.00317
	1.8	0.01178	0.02706	-0.00350
	1.9	0.01201	0.02753	-0.00351
	2.0	0.01174	0.02674	-0.00326
3	1.0	0.01629	0.03141	0.00117
	1.1	0.01629	0.03235	0.00023
	1.2	0.01624	0.03054	0.00194
	1.3	0.01644	0.03249	0.00038
	1.4	0.01712	0.03207	0.00216
	1.5	0.01674	0.03268	0.00081
	1.6	0.01714	0.03298	0.00131
	1.7	0.01629	0.03249	0.00010
	1.8	0.01732	0.03364	0.00100
	1.9	0.01654	0.03245	0.00062
	2.0	0.01732	0.03343	0.00121
4	1.0	0.02111	0.04324	-0.00102
	1.1	0.02080	0.04279	-0.00120
	1.2	0.02110	0.04317	-0.00097
	1.3	0.02041	0.04302	-0.00220
	1.4	0.02088	0.04298	-0.00122
	1.5	0.02174	0.04591	-0.00243
	1.6	0.02065	0.04240	-0.00110
	1.7	0.02247	0.04692	-0.00197
	1.8	0.02137	0.04306	-0.00031
	1.9	0.02137	0.04476	-0.00202
	2.0	0.02076	0.04218	-0.00067
5	1.0	0.02420	0.04837	0.00003
	1.1	0.02541	0.05030	0.00051
	1.2	0.02351	0.04776	-0.00075
	1.3	0.02453	0.04840	0.00065
	1.4	0.02394	0.05034	-0.00247
	1.5	0.02410	0.04728	0.00093
	1.6	0.02614	0.05311	-0.00083
	1.7	0.02441	0.04822	0.00060
	1.8	0.02568	0.05301	-0.00165
	1.9	0.02404	0.04743	0.00064

	2.0	0.02635	0.05288	-0.00019
6	1.0	0.03101	0.05211	0.00991
	1.1	0.03075	0.05253	0.00897
	1.2	0.03066	0.05073	0.01059
	1.3	0.03065	0.05106	0.01024
	1.4	0.03164	0.05209	0.01120
	1.5	0.03223	0.05428	0.01018
	1.6	0.03128	0.05348	0.00909
	1.7	0.03140	0.05381	0.00899
	1.8	0.03132	0.05321	0.00943
	1.9	0.03094	0.05270	0.00917
	2.0	0.03158	0.05346	0.00970
7	1.0	0.03480	0.05940	0.01020
	1.1	0.03475	0.06156	0.00794
	1.2	0.03414	0.05844	0.00983
	1.3	0.03447	0.06034	0.00861
	1.4	0.03471	0.05916	0.01025
	1.5	0.03415	0.05794	0.01037
	1.6	0.03613	0.06208	0.01018
	1.7	0.03549	0.06017	0.01080
	1.8	0.03588	0.06243	0.00934
	1.9	0.03459	0.05968	0.00950
	2.0	0.03474	0.06159	0.00790
8	1.0	0.03824	0.06224	0.01424
	1.1	0.03837	0.06495	0.01178
	1.2	0.03810	0.06244	0.01376
	1.3	0.03783	0.06411	0.01155
	1.4	0.03788	0.06190	0.01387
	1.5	0.03842	0.06441	0.01243
	1.6	0.03863	0.06355	0.01371
	1.7	0.03858	0.06373	0.01342
	1.8	0.03878	0.06309	0.01447
	1.9	0.03731	0.06040	0.01421
	2.0	0.03847	0.06474	0.01221
9	1.0	0.04200	0.06747	0.01653
	1.1	0.04205	0.06615	0.01795
	1.2	0.04115	0.06621	0.01609
	1.3	0.04174	0.06573	0.01776
	1.4	0.04223	0.06551	0.01895
	1.5	0.04210	0.06593	0.01828
	1.6	0.04263	0.06635	0.01890
	1.7	0.04210	0.06611	0.01810
	1.8	0.04199	0.06554	0.01844
	1.9	0.04204	0.06641	0.01767
	2.0	0.04236	0.06651	0.01821
10	1.0	0.04382	0.07033	0.01732
	1.1	0.04425	0.07152	0.01699
	1.2	0.04339	0.06894	0.01785

	1.3	0.04336	0.06960	0.01712
	1.4	0.04379	0.06954	0.01805
	1.5	0.04450	0.07079	0.01821
	1.6	0.04558	0.07198	0.01917
	1.7	0.04492	0.07084	0.01901
	1.8	0.04522	0.07259	0.01785
	1.9	0.04394	0.06966	0.01822
	2.0	0.04523	0.07262	0.01784

Table E-13: Data for curve fitting Scrap (Equal Area Binning to Minimize Variation)

Number of Bins	σ_A/σ_B	Percentage of Scrap	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.00261	0.00524	-0.00002
	1.1	0.00349	0.00679	0.00019
	1.2	0.00855	0.01544	0.00166
	1.3	0.01660	0.02495	0.00825
	1.4	0.02816	0.03914	0.01718
	1.5	0.04058	0.05223	0.02893
	1.6	0.05471	0.06891	0.04051
	1.7	0.07173	0.08965	0.05381
	1.8	0.08723	0.10736	0.06710
	1.9	0.10558	0.12601	0.08515
	2.0	0.12546	0.14713	0.10379
2	1.0	0.02360	0.05620	-0.00900
	1.1	0.02484	0.06126	-0.01158
	1.2	0.02893	0.06589	-0.00803
	1.3	0.03339	0.06688	-0.00010
	1.4	0.04007	0.06911	0.01103
	1.5	0.04906	0.07634	0.02178
	1.6	0.06032	0.08419	0.03645
	1.7	0.07593	0.09914	0.05272
	1.8	0.09007	0.11149	0.06865
	1.9	0.10656	0.12611	0.08701
	2.0	0.12580	0.14675	0.10485
3	1.0	0.03113	0.05963	0.00263
	1.1	0.04512	0.07716	0.01308
	1.2	0.06319	0.09970	0.02668
	1.3	0.08249	0.11852	0.04646
	1.4	0.09849	0.13053	0.06645
	1.5	0.11393	0.14798	0.07988
	1.6	0.12912	0.16244	0.09580
	1.7	0.14175	0.17826	0.10524
	1.8	0.15506	0.18834	0.12178
	1.9	0.16829	0.20196	0.13462
	2.0	0.17885	0.21022	0.14748

4	1.0	0.03622	0.06359	0.00885
	1.1	0.05120	0.07716	0.02524
	1.2	0.08088	0.10941	0.05235
	1.3	0.11216	0.14249	0.08183
	1.4	0.13885	0.16804	0.10966
	1.5	0.16209	0.19615	0.12803
	1.6	0.19494	0.28062	0.10926
	1.7	0.21536	0.30530	0.12542
	1.8	0.22859	0.30848	0.14870
	1.9	0.24060	0.30992	0.17128
	2.0	0.25510	0.33113	0.17907
5	1.0	0.04117	0.06778	0.01456
	1.1	0.05757	0.08593	0.02921
	1.2	0.08841	0.11966	0.05716
	1.3	0.12114	0.15604	0.08624
	1.4	0.15104	0.18677	0.11531
	1.5	0.19065	0.27270	0.10860
	1.6	0.21665	0.29382	0.13948
	1.7	0.24141	0.32035	0.16247
	1.8	0.25861	0.32973	0.18749
	1.9	0.27493	0.33781	0.21205
	2.0	0.29275	0.35889	0.22661
6	1.0	0.04345	0.07251	0.01440
	1.1	0.05911	0.09228	0.02594
	1.2	0.08918	0.12888	0.04948
	1.3	0.12146	0.16012	0.08280
	1.4	0.15613	0.19998	0.11227
	1.5	0.18626	0.22838	0.14415
	1.6	0.20909	0.25380	0.16437
	1.7	0.24461	0.32109	0.16813
	1.8	0.26315	0.33050	0.19580
	1.9	0.28500	0.34689	0.22311
	2.0	0.30632	0.37615	0.23649
7	1.0	0.04541	0.07541	0.01541
	1.1	0.06139	0.09489	0.02790
	1.2	0.09503	0.13604	0.05403
	1.3	0.12683	0.16727	0.08639
	1.4	0.16170	0.20653	0.11688
	1.5	0.19326	0.23517	0.15135
	1.6	0.21734	0.26435	0.17033
	1.7	0.25474	0.32727	0.18221
	1.8	0.27123	0.33751	0.20495
	1.9	0.29453	0.35640	0.23266
	2.0	0.31813	0.38545	0.25081
8	1.0	0.04886	0.07626	0.02147
	1.1	0.06454	0.09426	0.03483
	1.2	0.09631	0.13377	0.05885
	1.3	0.13015	0.16954	0.09075

	1.4	0.17317	0.24486	0.10148
	1.5	0.20148	0.26511	0.13785
	1.6	0.23039	0.30416	0.15662
	1.7	0.26020	0.32943	0.19097
	1.8	0.27766	0.33986	0.21546
	1.9	0.30174	0.36213	0.24135
	2.0	0.32413	0.38931	0.25895
9	1.0	0.06613	0.13642	-0.00416
	1.1	0.07669	0.14326	0.01012
	1.2	0.10557	0.17349	0.03765
	1.3	0.13715	0.19802	0.07628
	1.4	0.17378	0.24379	0.10377
	1.5	0.19836	0.24123	0.15549
	1.6	0.23216	0.30193	0.16239
	1.7	0.26365	0.32917	0.19813
	1.8	0.28156	0.34283	0.22029
	1.9	0.30676	0.36748	0.24604
	2.0	0.32918	0.39549	0.26287
10	1.0	0.06930	0.13453	0.00407
	1.1	0.07942	0.14229	0.01655
	1.2	0.10948	0.17366	0.04530
	1.3	0.13961	0.19898	0.08024
	1.4	0.17591	0.24082	0.11100
	1.5	0.20492	0.26597	0.14387
	1.6	0.23231	0.29768	0.16694
	1.7	0.26492	0.32682	0.20302
	1.8	0.28276	0.34156	0.22396
	1.9	0.30502	0.35285	0.25719
	2.0	0.33247	0.39552	0.26942

Table E-14: Data for curve fitting Scrap (Equal Area Binning to Minimize Scrap)

Number of Bins	σ_A/σ_B	Percentage of Scrap	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.00210	0.00430	-0.00010
	1.1	0.00215	0.00449	-0.00020
	1.2	0.00205	0.00427	-0.00017
	1.3	0.00219	0.00449	-0.00011
	1.4	0.00199	0.00425	-0.00027
	1.5	0.00208	0.00436	-0.00021
	1.6	0.00199	0.00421	-0.00022
	1.7	0.00199	0.00418	-0.00020
	1.8	0.00201	0.00428	-0.00026
	1.9	0.00203	0.00425	-0.00020
	2.0	0.00208	0.00450	-0.00033
	1.0	0.02176	0.05477	-0.01125

2	1.0	0.02176	0.05477	-0.01125
	1.1	0.02160	0.05635	-0.01314
	1.2	0.02214	0.05511	-0.01083
	1.3	0.02054	0.05452	-0.01345
	1.4	0.02213	0.05654	-0.01227
	1.5	0.02249	0.05782	-0.01284
	1.6	0.02224	0.05629	-0.01182
	1.7	0.02265	0.05653	-0.01123
	1.8	0.02149	0.05547	-0.01248
	1.9	0.02181	0.05497	-0.01135
	2.0	0.02161	0.05607	-0.01285
3	1.0	0.02951	0.05670	0.00231
	1.1	0.02983	0.05917	0.00049
	1.2	0.02987	0.05964	0.00011
	1.3	0.02994	0.05765	0.00222
	1.4	0.02990	0.06056	-0.00075
	1.5	0.03023	0.05832	0.00214
	1.6	0.03070	0.06189	-0.00049
	1.7	0.03053	0.05851	0.00254
	1.8	0.02998	0.06142	-0.00146
	1.9	0.03001	0.05780	0.00223
	2.0	0.02981	0.06159	-0.00196
4	1.0	0.03294	0.06237	0.00351
	1.1	0.03365	0.06339	0.00392
	1.2	0.03363	0.06355	0.00372
	1.3	0.03304	0.06266	0.00342
	1.4	0.03370	0.06451	0.00289
	1.5	0.03429	0.06475	0.00384
	1.6	0.03433	0.06431	0.00434
	1.7	0.03291	0.06349	0.00233
	1.8	0.03404	0.06402	0.00407
	1.9	0.03247	0.06237	0.00258
	2.0	0.03401	0.06437	0.00365
5	1.0	0.03894	0.06907	0.00880
	1.1	0.03917	0.06785	0.01049
	1.2	0.04009	0.06944	0.01073
	1.3	0.03872	0.06754	0.00989
	1.4	0.03917	0.07013	0.00822
	1.5	0.03942	0.06945	0.00940
	1.6	0.03871	0.06935	0.00807
	1.7	0.03926	0.06921	0.00930
	1.8	0.03849	0.06847	0.00852
	1.9	0.03906	0.06877	0.00935
	2.0	0.03867	0.06872	0.00862
6	1.0	0.04366	0.07176	0.01556
	1.1	0.04347	0.07001	0.01694
	1.2	0.04384	0.07234	0.01533
	1.3	0.04340	0.06952	0.01727

	1.4	0.04403	0.07153	0.01653
	1.5	0.04413	0.07085	0.01740
	1.6	0.04402	0.07263	0.01541
	1.7	0.04376	0.07090	0.01662
	1.8	0.04417	0.07165	0.01670
	1.9	0.04365	0.07040	0.01691
	2.0	0.04402	0.07122	0.01681
7	1.0	0.04448	0.07210	0.01686
	1.1	0.04574	0.07318	0.01830
	1.2	0.04562	0.07409	0.01715
	1.3	0.04526	0.07118	0.01933
	1.4	0.04610	0.07439	0.01780
	1.5	0.04596	0.07147	0.02045
	1.6	0.04693	0.07506	0.01880
	1.7	0.04564	0.07177	0.01951
	1.8	0.04656	0.07518	0.01793
	1.9	0.04508	0.07198	0.01818
	2.0	0.04662	0.07507	0.01817
8	1.0	0.04965	0.07512	0.02417
	1.1	0.05012	0.07665	0.02359
	1.2	0.05009	0.07517	0.02501
	1.3	0.05012	0.07672	0.02351
	1.4	0.05010	0.07563	0.02456
	1.5	0.05040	0.07636	0.02444
	1.6	0.05079	0.07707	0.02451
	1.7	0.05087	0.07648	0.02527
	1.8	0.05106	0.07680	0.02531
	1.9	0.05060	0.07565	0.02555
	2.0	0.05034	0.07802	0.02266
9	1.0	0.05337	0.07937	0.02736
	1.1	0.05337	0.08040	0.02635
	1.2	0.05344	0.08024	0.02665
	1.3	0.05278	0.07874	0.02683
	1.4	0.05308	0.08054	0.02562
	1.5	0.05276	0.07757	0.02794
	1.6	0.05438	0.08386	0.02490
	1.7	0.05313	0.07866	0.02760
	1.8	0.05385	0.08308	0.02463
	1.9	0.05299	0.07871	0.02727
	2.0	0.05358	0.08234	0.02483
10	1.0	0.05757	0.08742	0.02772
	1.1	0.05874	0.08815	0.02933
	1.2	0.05880	0.08870	0.02891
	1.3	0.05759	0.08746	0.02772
	1.4	0.05880	0.08892	0.02868
	1.5	0.05741	0.08538	0.02944
	1.6	0.05899	0.08942	0.02856
	1.7	0.05794	0.08733	0.02854

	1.8	0.05937	0.09031	0.02842
	1.9	0.05753	0.08712	0.02793
	2.0	0.05902	0.08918	0.02886

Table E-15: Data for curve fitting Scrap (Optimal Width Binning)

Number of Bins	σ_A/σ_B	Percentage of Scrap	+95% Confidence Interval Band	-95% Confidence Interval Band
1	1.0	0.00237	0.00549	-0.00075
	1.1	0.00228	0.00544	-0.00088
	1.2	0.00240	0.00542	-0.00062
	1.3	0.00246	0.00557	-0.00065
	1.4	0.00234	0.00546	-0.00078
	1.5	0.00228	0.00544	-0.00088
	1.6	0.00242	0.00534	-0.00050
	1.7	0.00228	0.00544	-0.00088
	1.8	0.00228	0.00544	-0.00088
	1.9	0.00230	0.00536	-0.00076
	2.0	0.00246	0.00557	-0.00065
2	1.0	0.02109	0.05398	-0.01180
	1.1	0.02090	0.05524	-0.01344
	1.2	0.02088	0.05384	-0.01208
	1.3	0.02128	0.05300	-0.01044
	1.4	0.02230	0.05583	-0.01123
	1.5	0.02090	0.05524	-0.01344
	1.6	0.02270	0.05791	-0.01251
	1.7	0.02090	0.05524	-0.01344
	1.8	0.02090	0.05524	-0.01344
	1.9	0.02266	0.05793	-0.01261
	2.0	0.02128	0.05300	-0.01044
3	1.0	0.02607	0.05444	-0.00230
	1.1	0.02542	0.05350	-0.00266
	1.2	0.02632	0.05619	-0.00355
	1.3	0.02672	0.05560	-0.00216
	1.4	0.02756	0.05676	-0.00164
	1.5	0.02542	0.05350	-0.00266
	1.6	0.02608	0.05545	-0.00329
	1.7	0.02542	0.05350	-0.00266
	1.8	0.02542	0.05350	-0.00266
	1.9	0.02714	0.05586	-0.00158
	2.0	0.02672	0.05560	-0.00216
4	1.0	0.03173	0.05972	0.00374
	1.1	0.03078	0.05937	0.00219
	1.2	0.03286	0.06114	0.00458
	1.3	0.03268	0.06021	0.00515
	1.4	0.03306	0.06124	0.00488

	1.5	0.03078	0.05937	0.00219
	1.6	0.03376	0.06274	0.00478
	1.7	0.03078	0.05937	0.00219
	1.8	0.03078	0.05937	0.00219
	1.9	0.03296	0.06093	0.00499
	2.0	0.03268	0.06021	0.00515
5	1.0	0.03856	0.06946	0.00766
	1.1	0.03944	0.07088	0.00800
	1.2	0.03834	0.07166	0.00502
	1.3	0.03768	0.06824	0.00712
	1.4	0.03914	0.07257	0.00571
	1.5	0.03944	0.07088	0.00800
	1.6	0.03862	0.07173	0.00551
	1.7	0.03944	0.07088	0.00800
	1.8	0.03944	0.07088	0.00800
	1.9	0.03932	0.07054	0.00810
	2.0	0.03768	0.06824	0.00712
6	1.0	0.04058	0.06984	0.01132
	1.1	0.03964	0.06894	0.01034
	1.2	0.04106	0.06979	0.01233
	1.3	0.04152	0.07091	0.01213
	1.4	0.04222	0.07133	0.01311
	1.5	0.03964	0.06894	0.01034
	1.6	0.04252	0.07184	0.01320
	1.7	0.03964	0.06894	0.01034
	1.8	0.03964	0.06894	0.01034
	1.9	0.04184	0.07174	0.01194
	2.0	0.04152	0.07091	0.01213
7	1.0	0.04433	0.07030	0.01836
	1.1	0.04490	0.07265	0.01715
	1.2	0.04526	0.07065	0.01987
	1.3	0.04376	0.06805	0.01947
	1.4	0.04522	0.06946	0.02098
	1.5	0.04490	0.07265	0.01715
	1.6	0.04564	0.07181	0.01947
	1.7	0.04490	0.07265	0.01715
	1.8	0.04490	0.07265	0.01715
	1.9	0.04576	0.07236	0.01916
	2.0	0.04376	0.06805	0.01947
8	1.0	0.04764	0.07231	0.02297
	1.1	0.04768	0.07292	0.02244
	1.2	0.04742	0.07280	0.02204
	1.3	0.04760	0.07194	0.02326
	1.4	0.04790	0.07225	0.02355
	1.5	0.04768	0.07292	0.02244
	1.6	0.04744	0.07194	0.02294
	1.7	0.04768	0.07292	0.02244
	1.8	0.04768	0.07292	0.02244

	1.9	0.04756	0.07319	0.02193
	2.0	0.04760	0.07194	0.02326
9	1.0	0.05254	0.07945	0.02563
	1.1	0.05288	0.08020	0.02556
	1.2	0.05172	0.07890	0.02454
	1.3	0.05220	0.07896	0.02544
	1.4	0.05316	0.07786	0.02846
	1.5	0.05288	0.08020	0.02556
	1.6	0.05308	0.08088	0.02528
	1.7	0.05288	0.08020	0.02556
	1.8	0.05288	0.08020	0.02556
	1.9	0.05450	0.07959	0.02941
	2.0	0.05220	0.07896	0.02544
10	1.0	0.05302	0.08099	0.02505
	1.1	0.05236	0.07960	0.02512
	1.2	0.05426	0.08395	0.02457
	1.3	0.05368	0.08259	0.02477
	1.4	0.05342	0.08184	0.02500
	1.5	0.05236	0.07960	0.02512
	1.6	0.05488	0.08342	0.02634
	1.7	0.05236	0.07960	0.02512
	1.8	0.05236	0.07960	0.02512
	1.9	0.05510	0.08377	0.02643
	2.0	0.05368	0.08259	0.02477

The goodness of fit values, R^2, are presented for the curves that were fit for the Equal Width Binning to Minimize Variation and the Equal Area Binning to Minimize Variation selective assembly techniques in Table E-16 and Table E-17 respectively. The data representing the percentage of scrap for the selective assembly techniques that bin the two part distributions independently, Equal Width Binning to Minimize Scrap, the Equal Area Binning to Minimize Scrap, and Optimal Width Binning, are best represented by a constant value, i.e., a straight horizontal line passing through the average. For this type of curve fit, the R^2 value has no meaning, thus tables for the goodness of fit for these techniques do not exist.

Table E-16: Goodness of fit for Scrap (Equal Width Binning to Minimize Variation)

Number of Bins	R^2 for Scrap Curve	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	1.000	1.000	0.999
2	1.000	0.999	0.999
3	0.999	1.000	0.998
4	0.999	0.999	0.999
5	1.000	0.999	0.999
6	0.998	0.999	0.997
7	0.999	0.998	0.998

8	0.999	0.999	0.998
9	0.998	0.998	0.998
10	0.998	0.999	0.998

Table E-17: Goodness of fit for Scrap (Equal Area Binning to Minimize Variation)

Number of Bins	R^2 for Scrap Curve	R^2 for +95% Confidence Interval Band	R^2 for -95% Confidence Interval Band
1	1.000	1.000	0.999
2	1.000	0.997	1.000
3	0.999	0.999	0.998
4	0.998	0.983	0.969
5	0.998	0.986	0.990
6	0.999	0.994	0.992
7	0.998	0.995	0.994
8	0.998	0.989	0.997
9	0.998	0.989	0.995
10	0.998	0.992	0.998

F. Additional PCDB User Interface Information

The discussion regarding the process capability database graphical user interface from Section 2.3 is continued here. The user interface also included an administration section to manage the database. The administration section contains pages to add manufacturing plants to the database so they can be selected when running a query, attribute management pages to update the selectable attributes, and a user management page to create new users and control security levels. The main administration page is shown in Figure F-1.

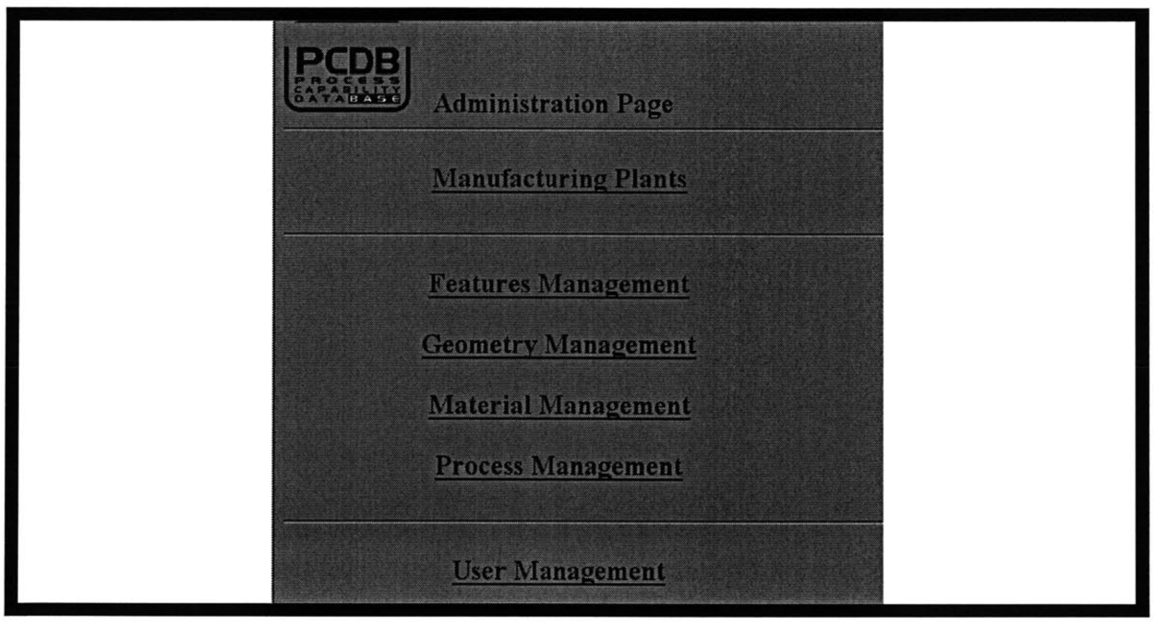

Figure F-1: Administration main page

The Manufacturing Plants page is used to manage what plants are stored in the database. From this page, a list of plants presently in the database can be displayed. Plant information is also added from this page. A screen capture of this page is shown as Figure F-2.

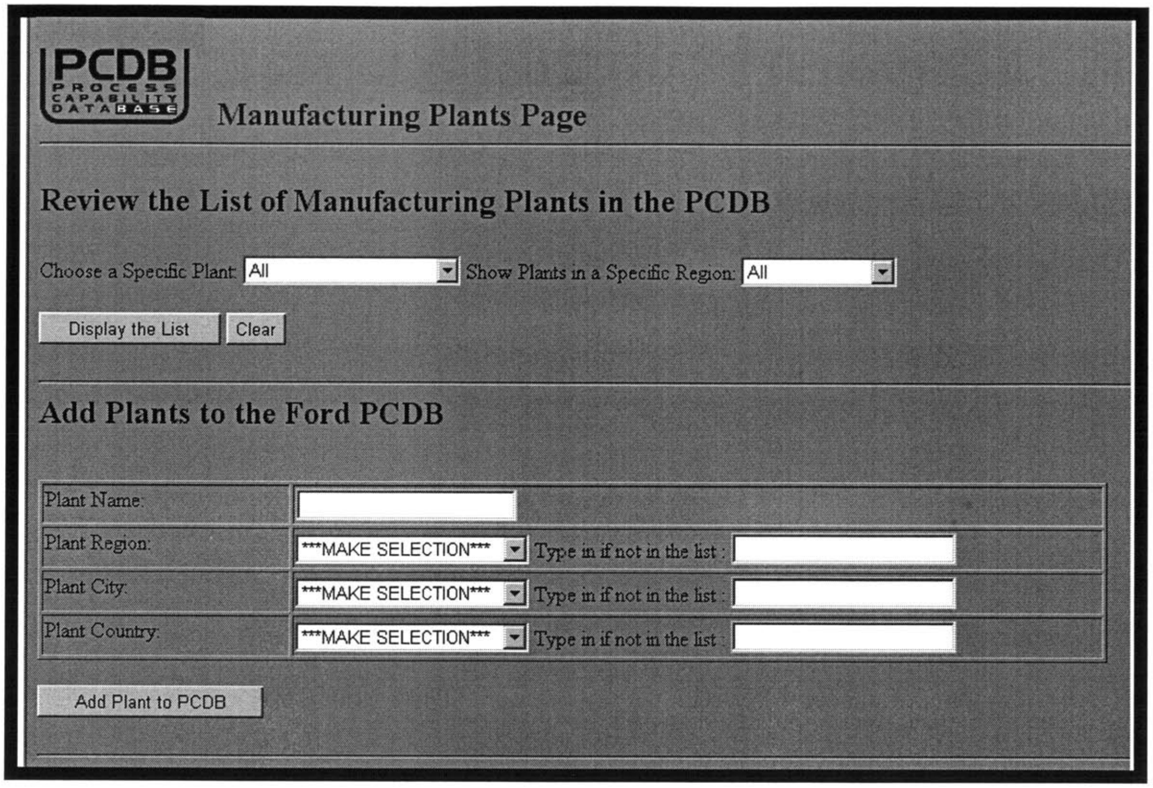

Figure F-2: Manufacturing Plant Management Page

Each attribute, feature, geometry, material, and process, has an associated management page. These pages allow the administrator to review the existing lists of attributes in the database and define new attributes at any level and define relationships among them. A screen capture of the Material Management Page is shown below as Figure F-3. Similar pages exist for each of the other attributes.

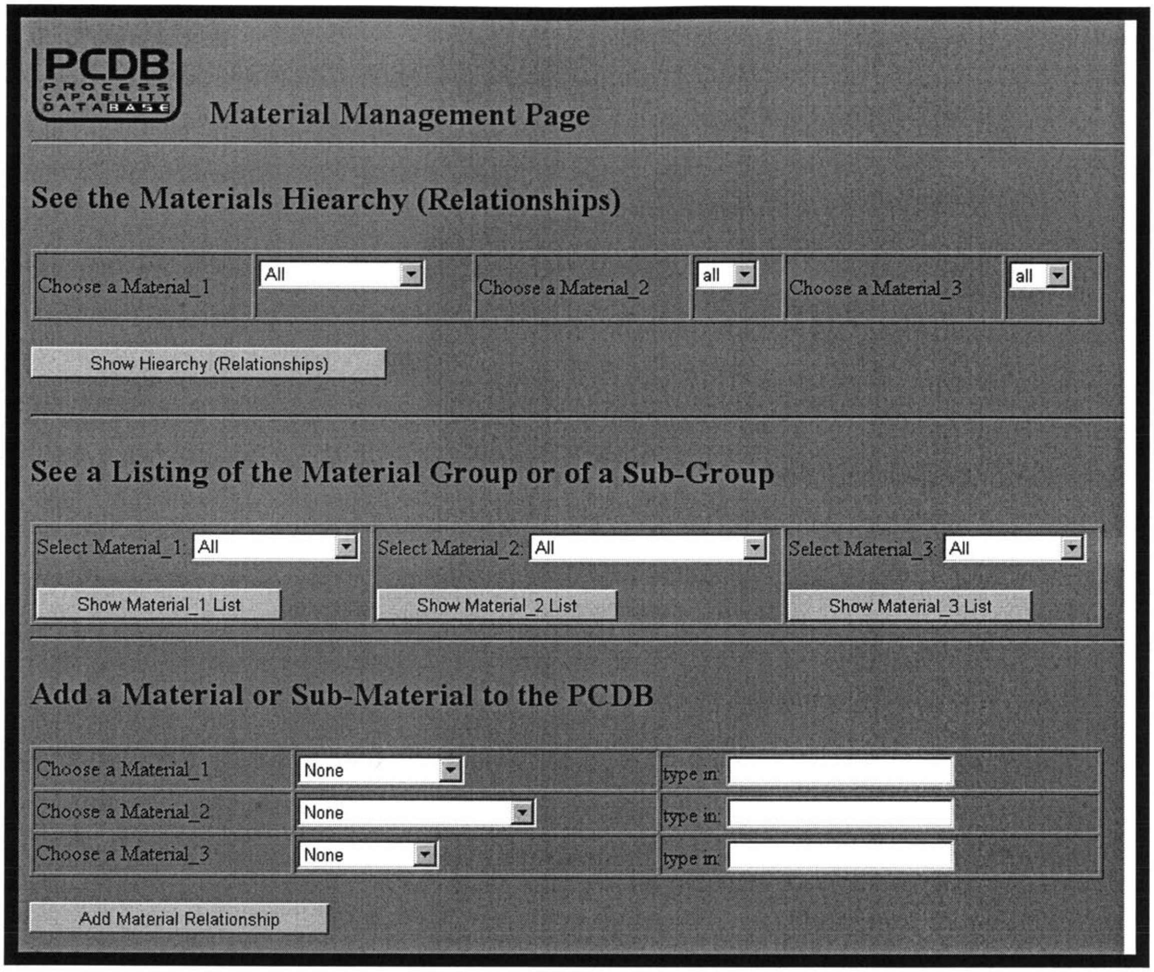

Figure F-3: Page to manage material attribute listings in the PCDB

The last part of the administration section is the user management page. This page is used by the administrator to add new users, display current users, and to adjust the security level of users. A screen capture of the User Management Page is shown below as Figure F-4.

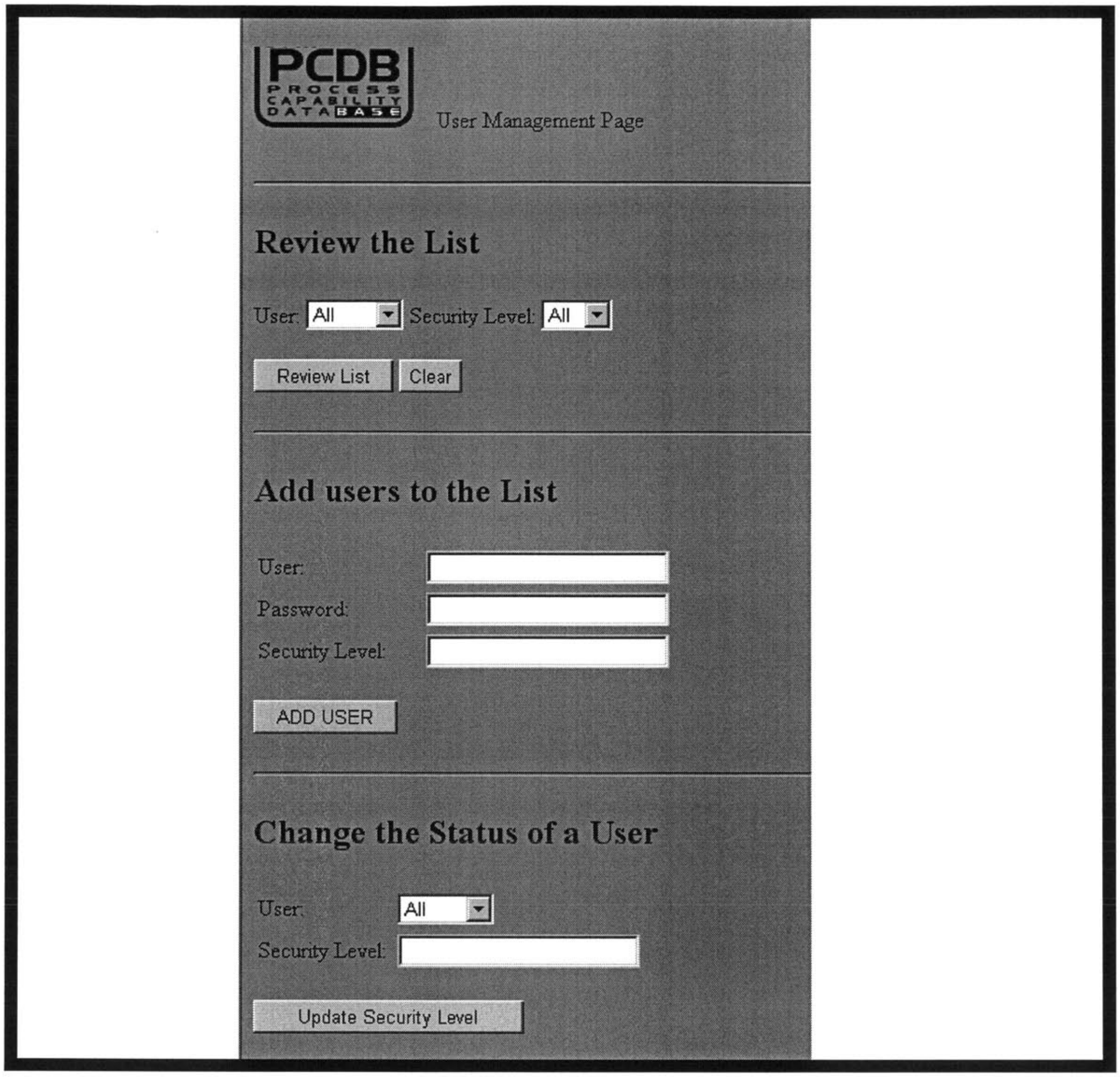

Figure F-4: Page used to manage user accounts

G. Detailed Derivation for Chapter 3

The methodology developed in this thesis requires a substantial amount of mathematical background to be implemented. A great deal of this framework is rooted in statistics and calculus. Engineers commonly think deterministically; yet, considering the stochastic nature of manufacturing a product is extremely important for engineers because it is impossible for any manufacturing process to make every component exactly the same.

A product is typically an assembly composed of many sub-assemblies and components. Even a product as simple in appearance as a VHS tape is often composed of more than thirty individual components. Each component has to be manufactured by some single process or series of processes. These processes create the values of the dimensions of the component's features. The dimensions have target values and tolerance ranges. The distribution of the actual values of these dimensions is what interests the engineer since these values ultimately determine how well the product will meet its functional requirements.

Another reason to use statistics when dealing with manufacturing a product is that often a large number of the same component is created by the same process. This lends itself very well to a statistical approach since powerful design decisions can be made based on properties of the distribution of a measurable quantity of a component, e.g. feature sizes can be set with confidence that they can be manufactured with a high level of repeatability. In this thesis, the concern is designing components that have features that can be produced with an acceptable level of quality. The design tool being developed is not intended to help design a unique one-of-a-kind assembly; a task for simple tolerance stack-up.

The equations developed in this thesis are not valid for predicting a measured quantity on the next part produced; the equations estimate the properties of the distribution of the next large batch of parts produced. What is large? Usually sample sizes of more than thirty parts begin to achieve statistical significance. The large batch requirement is rarely an issue because the design decisions made by engineers are typically long term and span many batches. Engineers want to design long term manufacturability into the product, not the manufacturability of one batch of components.

Data from manufactured components are often normally distributed (Kalpakjian and Schmid 2001). Working with normal distributions allows for closed-form equations to be derived. All the closed-form equations developed in this thesis are valid when working with normal distributions. A subset of these equations is valid for the more general case of any type of distribution.

A basic understanding of probability and statistics is assumed. To gain a better understanding of the fundamentals of probability and statistics, the reader should refer to a text on the subject matter. Two such texts are by Drake (Drake 1967) and by DeGroot (DeGroot 1986). Calculus is also heavily used as part of this derivation. Many texts exist covering the fundamentals of calculus. One text used for reference during the following derivations is by Grossman (Grossman 1992).

A probability distribution function is a mathematical description of how likely a variable is to take on a particular value. A normal distribution is a special type of distribution that is symmetrical and bell-shaped and is described by a mean, μ, and a standard deviation, σ. The standard deviation squared is equal to the variance of a distribution. Both terms are used in this thesis to refer to manufacturing variation. The probability density function (pdf) for a normal distribution is

$$pdf_{Normal} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \qquad \text{G-1}$$

The variable x can be converted to what is known as the reduced variable z. The equation for z is given by

$$z = \frac{(x-\mu)}{\sigma} \qquad \text{G-2}$$

Using z, any normal distribution can be re-written so that its mean is at zero and it has a standard deviation of one. This version of a normal distribution is known as a standard normal distribution and is given by the following equation.

$$pdf = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \quad \text{G-3}$$

A mixture distribution is formed by combining measurements from multiple distributions into one distribution. The equation for the probability density function of a mixture distribution comprised of any type of distributions is

$$pdf_{mixed} = (p_1)(pdf_1) + (p_2)(pdf_2) + \ldots + (p_i)(pdf_i) + \ldots (p_N)(pdf_N) \quad \text{G-4}$$

The mean and the variance of the mixture distribution can be computed if the means and standard deviations of the original distributions are known.

$$\mu_{mixed} = p_1\mu_1 + p_2\mu_2 + \ldots + p_N\mu_N \quad \text{G-5}$$

$$\begin{aligned} \sigma^2_{mixed} = & \left(p_1\sigma_1^2 + p_2\sigma_2^2 + \ldots + p_N\sigma_N^2\right) \\ & + \left(p_1\mu_1^2 + p_2\mu_2^2 + \ldots + p_N\mu_N^2\right) \\ & - (p_1\mu_1 + p_2\mu_2 + \ldots + p_N\mu_N)^2 \end{aligned} \quad \text{G-6}$$

where

N = Number of original distributions

σ = standard deviation of the distribution

μ = mean of the distribution

p_i = number of measurements in i^{th} distribution divided by the total number of measurements

The next step in laying the mathematical framework is to bin the incoming distribution. Binning is the procedure of dividing a distribution into discrete divisions, or bins. Once the width of a bin is fixed, the height is set so that the area of the bin is equal to that under the probability distribution function over that range. Therefore, the area in all the bins adds to unity. The area in a bin can be calculated regardless of the type of probability density function according to the following equation

$$Area_i = \int_{x_{i-1}}^{x_i} \left(pdf_{input} \right) dx \qquad \text{G-7}$$

$$x_i - x_{i-1} = \Delta \qquad \text{G-8}$$

where Δ is the bin width and x is the variable of interest. The bin width and the number of bins are arbitrary at this point. The smaller the width, the greater number of bins required to traverse the original distribution. Also, as the bin width decreases, the discrete representation approaches that of the original distribution.

Binning allows attention to be easily focused on a specific region of the distribution. As a mathematical approximation, each bin can be considered as an individual uniform distribution. This approximation allows the calculations presented in the next few sections to be simplified for example purposes. Once the methodology is developed, the approximation becomes mute as the bin width is driven to zero, resulting in an exact solution as will be seen in a later section.
Consider the original input distribution that has a mean of μ_{input}. The desired output dimension value is μ_{target}. Now assume the i[th] bin from x_{i-1} to x_i contains components that all have the same dimension, equal to $\frac{(x_{i-1} + x_i)}{2}$. The desired dimensional change for components in the i[th] bin to get them to the target value is

$$M_i = \mu_{target} - \frac{(x_{i-1} + x_i)}{2} \qquad \text{G-9}$$

The bins of the input distribution are now treated as individual input sub-distributions. The parts from the i[th] bin are transformed by the process to create an output sub-distribution with a probability density function of pdf_i.

Using the definitions of DeltaP and SigmaP, formulations for the mean and variance of the sub-distributions can be formulated. The mean of the i[th] sub-distribution can be written as

$$\mu_i \cong \frac{x_{i-i} + x_i}{2} + DeltaP(M_i) \qquad \text{G-10}$$

The standard deviation of the ith sub-distribution can be computed using the definition of SigmaP. The standard deviation of the ith output sub-distribution is

$$\sigma_i \cong SigmaP(M_i) \qquad \text{G-11}$$

These are approximations because they are based on finite bin widths.

The individual output sub-distributions, as well as information about those distributions, are of little use to an engineer. The engineer is concerned with the entire output distribution and its properties; therefore, the output distribution needs to be generated by combining the pdf_i's of the individual output sub-distributions using Equation G-4. The resulting equation for the combined pdf_{output} is

$$\begin{aligned} pdf_{output} &= (Area_1)(pdf_1) + (Area_2)(pdf_2) + \ldots \\ &\quad + (Area_i)(pdf_i) + \ldots + (Area_N)(pdf_N) \end{aligned} \qquad \text{G-12}$$

where $Area_i$ is the area in the ith Bin. Combining the above equation with Equation G-7 results in

$$\begin{aligned} pdf_{output} &= \int_{x_0}^{x_1} (pdf_{Input})\,dx\,(pdf_1) + \ldots + \\ &\int_{x_i}^{x_{i+1}} (pdf_{Input})\,dx\,(pdf_{i+1}) + \ldots + \int_{x_{N-1}}^{x_N} (pdf_{Input})\,dx\,(pdf_N) \end{aligned} \qquad \text{G-13}$$

These equations would be exact solutions except that the original distribution was binned.

Equation G-13 above is of little practical use because of the difficulty in knowing the actual equations for the pdf_i's. It is much more useful to the engineer to have the ability to calculate the mean and variance of the output distribution.

The output distribution's mean can now be approximated by combining information about the sub-distributions with the equation for the mean of a mixture distribution. This approximation for the output distribution's mean is

$$\mu_{output} \cong p_1\mu_1 + p_2\mu_2 + \ldots p_i\mu_i + \ldots + p_N\mu_N \qquad \text{G-14}$$

Note p_i is equivalent to the area in the i^{th} bin; therefore, Equation G-14 can be combined with Equations G-7 and G-10 to produce

$$\begin{aligned}\mu_{output} \cong & \int_{x_0}^{x_1}\left(pdf_{input}\right)dx * \left[\frac{x_0 + x_1}{2} + DeltaP\left(M_1\right)\right] \\ & + \int_{x_1}^{x_2}\left(pdf_{input}\right)dx * \left[\frac{x_1 + x_2}{2} + DeltaP\left(M_2\right)\right] \\ & + \ldots + \int_{x_{i-1}}^{x_i}\left(pdf_{input}\right)dx * \left[\frac{x_{i-1} + x_i}{2} + DeltaP\left(M_i\right)\right] + \ldots \\ & + \int_{x_{N-1}}^{x_N}\left(pdf_{input}\right)dx * \left[\frac{x_{N-1} + x_N}{2} + DeltaP\left(M_N\right)\right]\end{aligned} \qquad \text{G-15}$$

Using a summation, the equation above can be simply written as

$$\mu_{output} \cong \sum_{i=1}^{N}\left\{\int_{x_{i-1}}^{x_i}\left(pdf_{input}\right)dx * \left[\frac{x_{i-1} + x_i}{2} + DeltaP\left(M_i\right)\right]\right\} \qquad \text{G-16}$$

In a similar manner to calculating the equation for approximating the output distribution's mean, an equation for the approximate output distribution's variation can be derived. Using Equation G-6 the approximate output distribution's variation is

$$\begin{aligned}\sigma^2_{output} \cong & \left(p_1\sigma_1^2 + p_2\sigma_2^2 + \ldots + p_i\sigma_i^2 + \ldots + p_N\sigma_N^2\right) \\ & + \left(p_1\mu_1^2 + p_2\mu_2^2 + \ldots + p_i\mu_i^2 + \ldots p_N\mu_N^2\right) \\ & - (p_1\mu_1 + p_2\mu_2 + \ldots + p_i\mu_i + \ldots p_N\mu_N)^2\end{aligned} \qquad \text{G-17}$$

The above equation is an approximation because it is based on finite bin widths.

Note the last term in the above equation in parentheses is equal to μ_{output}, thus the equation can be simplified to

$$\begin{aligned}\sigma_{output}^2 \cong &\left(p_1\sigma_1^2 + p_2\sigma_2^2 + \ldots + p_i\sigma_i^2 + \ldots + p_N\sigma_N^2\right)\\ &+\left(p_1\mu_1^2 + p_2\mu_2^2 + \ldots + p_i\mu_i^2 + \ldots p_N\mu_N^2\right)\\ &-(\mu_{output})^2\end{aligned} \qquad \text{G-18}$$

Now the μ_i's can be replaced using Equation G-10 to obtain

$$\begin{aligned}\sigma_{output}^2 \cong &\left(p_1\sigma_1^2 + p_2\sigma_2^2 + \ldots + p_i\sigma_i^2 + \ldots + p_N\sigma_N^2\right)\\ &+\left(\begin{array}{l} p_1\left[\frac{x_0+x_1}{2} + DeltaP(M_1)\right]^2 + p_2\left[\frac{x_1+x_2}{2} + DeltaP(M_2)\right]^2 \\ + \ldots + p_i\left[\frac{x_{i-1}+x_i}{2} + DeltaP(M_i)\right]^2 + \ldots + p_N\left[\frac{x_{N-1}+x_N}{2} + DeltaP(M_N)\right]^2 \end{array}\right)\\ &-(\mu_{output})^2\end{aligned} \qquad \text{G-19}$$

The σi's terms can now be replaced using Equation G-11 to obtain

$$\begin{aligned}\sigma_{output}^2 \cong &\left(p_1\left[SigmaP(M_1)\right]^2 + p_2\left[SigmaP(M_2)\right]^2 + \ldots + p_i\left[SigmaP(M_i)\right]^2 + \ldots + p_N\left[SigmaP(M_N)\right]^2\right)\\ &+\left(\begin{array}{l} p_1\left[\frac{x_0+x_1}{2} + DeltaP(M_1)\right]^2 + p_2\left[\frac{x_1+x_2}{2} + DeltaP(M_2)\right]^2 \\ + \ldots + p_i\left[\frac{x_{i-1}+x_i}{2} + DeltaP(M_i)\right]^2 + \ldots + p_N\left[\frac{x_{N-1}+x_N}{2} + DeltaP(M_N)\right]^2 \end{array}\right)\\ &-(\mu_{output})^2\end{aligned} \qquad \text{G-20}$$

Note p_i is equivalent to the area in the i^{th} Bin; therefore, the equation above can be combined with Equation G-7 to produce

$$\sigma_{output}^2 \cong \begin{pmatrix} \int_{x_0}^{x_1}(pdf_{input})dx\left[SigmaP(M_1)\right]^2 + \int_{x_1}^{x_2}(pdf_{input})dx\left[SigmaP(M_2)\right]^2 \\ +\ldots+\int_{x_{i-1}}^{x_i}(pdf_{input})dx\left[SigmaP(M_i)\right]^2 +\ldots+\int_{x_{N-1}}^{x_N}(pdf_{input})dx\left[SigmaP(M_N)\right]^2 \end{pmatrix}$$

$$+\begin{pmatrix} \int_{x_0}^{x_1}(pdf_{input})dx\left[\frac{x_0+x_1}{2}+DeltaP(M_1)\right]^2 + \int_{x_1}^{x_2}(pdf_{input})dx\left[\frac{x_1+x_2}{2}+DeltaP(M_2)\right]^2 \\ +\ldots+\int_{x_{i-1}}^{x_i}(pdf_{input})dx\left[\frac{x_{i-1}+x_i}{2}+DeltaP(M_i)\right]^2 +\ldots+\int_{x_{N-1}}^{x_N}(pdf_{input})dx\left[\frac{x_{N-1}+x_N}{2}+DeltaP(M_N)\right) \end{pmatrix}$$

$$-(\mu_{output})^2 \qquad \text{G-21}$$

Using summations, the above equation can be written in a shortened version such as

$$\sigma_{output}^2 \cong \sum_{i=1}^{N}\left(\int_{x_{i-1}}^{x_i}(pdf_{input})dx\left[SigmaP(M_i)\right]^2\right)$$

$$+\sum_{i=1}^{N}\left(\int_{x_{i-1}}^{x_i}(pdf_{input})dx\left[\frac{x_{i-1}+x_i}{2}+DeltaP(M_i)\right]^2\right) \qquad \text{G-22}$$

$$-(\mu_{output})^2$$

The equation above is an approximation for σ_{output}^2 because of the use of N discrete bins. The larger the number of bins, the better the approximation is to the true value. The exact solution is presented later.

Processes often produce normal distribution outputs (Kalpakjian and Schmid 2001). Working with distributions that are normal allows closed-form solutions to be derived.

Consider the original input distribution to be a normal distribution with a mean equal to μ_{input} and standard deviation equal to σ_{input}. Also assume the process generates a normal output sub-distributions with means μ_i.

For the above equation and all equations below for calculating the output standard deviation, it is assumed the input distribution does not overlap the target mean. The physical reasoning for this assumption is that if overlap exists, the overlapping components in the input distribution would not be processed as they would already be equal to or less than the target mean. Not processing some components would invalidate the equations below. Equations could be derived to handle the situation of overlap; however, they are not derived here because in a manufacturing situation that is considered to be in statistical control, the case of this type of overlap will not arise.

Standardizing the preceding equations eases calculations and allows for integration limits to be general for any case, versus being set uniquely for each case. Standardizing the equations involves transforming the probability density functions to standard normal density functions and x values to z values. A standard normal density function is a bell shaped probability density function with a mean of zero and a standard deviation of one. Values of z represent the number of standard deviations away from zero a particular value is.

Using Equation G-2 for standardization results in

$$x_i = z_i \sigma_{input} + \mu_{input} \qquad \text{G-23}$$

With the normal distribution assumption, the integral, $\int_{x_{i-1}}^{x_i} \left(pdf_{Input} \right) dx$, becomes

$$\frac{1}{\sqrt{2\pi}} \int_{z_{i-1}}^{z_i} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) dz \qquad \text{G-24}$$

The standardized desired dimensional change, M_i, becomes

$$M_i = \mu_{target} - \left\{ \frac{\{ z_{i-1} + z_i \} \sigma_{input}}{2} + \mu_{input} \right\} \qquad \text{G-25}$$

Now, the equations for the approximate mean and standard deviation of the output distribution can be transformed into standardized forms. The approximate output distribution's mean, Equation G-16, can be written as a function of z to obtain

$$\mu_{output} \cong \sum_{i=1}^{N}\left\{\frac{1}{\sqrt{2\pi}}\int_{z_{i-1}}^{z_i}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}dz\left[\frac{\{z_i+z_{i+1}\}\sigma_{input}}{2}+\mu_{input}+DeltaP\left(\mu_{target}-\left\{\frac{\{z_{i-1}+z_i\}\sigma_{input}}{2}+\mu_{input}\right\}\right)\right]\right\} \qquad \text{G-26}$$

The approximate output distribution's standard deviation, Equation G-22, can be written as a function of z to obtain

$$\begin{aligned}\sigma_{output}^2 \cong & \sum_{i=1}^{N}\left(\frac{1}{\sqrt{2\pi}}\int_{z_{i-1}}^{z_i}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)dz\left[SigmaP\left(\mu_{target}-\left\{\frac{\{z_{i-1}+z_i\}\sigma_{input}}{2}+\mu_{input}\right\}\right)\right]^2\right)\\ & +\sum_{i=1}^{N}\left(\frac{1}{\sqrt{2\pi}}\int_{z_{i-1}}^{z_i}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)dz\left[\frac{(z_{i-1}+z_i)\sigma_{input}}{2}+\mu_{input}+DeltaP\left(\mu_{target}-\left\{\frac{\{z_{i-1}+z_i\}\sigma_{input}}{2}+\mu_{input}\right\}\right)\right]^2\right)\\ & -(\mu_{output})^2\end{aligned} \qquad \text{G-27}$$

The above is an approximation. The exact solution is about to be presented.

An engineer is ultimately interested in the exact solution of the output distribution's mean, μ_{output}, and standard deviation,σ_{output}. The exact solutions are obtained by taking the limits of Equations G-26 and G-27 as the bin width goes to zero.

Mathematically speaking, the limits of the portions of the two equations of interest are

$$\lim_{\Delta\to 0}\left[z_{i-1}+z_i\right]=2z \qquad \text{G-28}$$

$$\lim_{\Delta\to 0}\left[\frac{1}{\sqrt{2\pi}}\int_{z_{i-1}}^{z_i}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)dz\right]=\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right) \qquad \text{G-29}$$

Assuming the bins were selected to span the entire original input distribution, as they should have been, the following is true.

$$\lim_{\Delta\to 0}\left[\sum_{i=0}^{N}\quad\right] \text{ becomes } \int_{-\infty}^{\infty}[\quad]dz \qquad \text{G-30}$$

Equations G-28, G-29, and G-30 are combined with Equations G-26 and G-27 to obtain exact solutions for the output distribution's mean, μ_{output}, and standard deviation, σ_{output}. These exact solutions are presented next.

As the bin width goes to zero, the desired dimensional change becomes

$$\lim_{\Delta \to 0}\left[M_i\right] = M = \mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\} \qquad \text{G-31}$$

All the pieces are in place to formulate the equation for the exact output distribution's mean. Equations G-28, G-29, and G-30 are used in taking the limit of Equation G-26 to produce

$$\mu_{output} = \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[z\sigma_{input} + \mu_{input} + DeltaP\left(\mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\} \right) \right] \right\} dz \qquad \text{G-32}$$

Expanding this equation results in

$$\begin{aligned} \mu_{output} = & \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[z\sigma_{input} \right] \right\} dz \\ & + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[\mu_{input} \right] \right\} dz \\ & + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[DeltaP\left(\mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\} \right) \right] \right\} dz \end{aligned} \qquad \text{G-33}$$

The first term in the above equation has a form that appears repeatedly in the derivations of μ_{output} and σ_{output}. It is beneficial to realize a term of this form can be simplified according to the equation below.

1 for k = 0 G-34

0 for odd k's

$$\int_{-\infty}^{+\infty}\left\{\frac{z^k}{\sqrt{2\pi}}\left[\exp\left(-\frac{z^2}{2}\right)\right]\right\}dz = \begin{cases} \\ \prod_{j=1}^{\frac{k}{2}}\left[k-\{2j-1\}\right] & \text{for even k's} \\ \\ \end{cases}$$

The physical reason why $\int_{-\infty}^{+\infty}\left\{\frac{z^k}{\sqrt{2\pi}}\left[\exp\left(-\frac{z^2}{2}\right)\right]\right\}dz$ equals zero for all odd i's is simple. Integration is the summation of the integrand, $\frac{z^k}{\sqrt{2\pi}}\left[\exp\left(-\frac{z^2}{2}\right)\right]$. Note the integrand is composed of the z^k and the $\frac{1}{\sqrt{2\pi}}\left[\exp\left(-\frac{z^2}{2}\right)\right]$ terms. The first term is simply z raised to a power, and the second term is the standardized normal distribution, which is symmetric about zero. The symmetry of this term combined with the fact z^i takes the sign of z when i is odd leads to a cancellation effect when the integral is computed. For example, when z takes on any value, K, and i is odd, there exists a value of –K that produces the canceling effect. The two canceling terms in this case are $\left[\frac{-K^3}{\sqrt{2\pi}}\left(\exp\left\{-\frac{K^2}{2}\right\}\right)\right]$ and $\left[\frac{K^3}{\sqrt{2\pi}}\left(\exp\left\{-\frac{K^2}{2}\right\}\right)\right]$.

The integral in the second term of Equation G-33 equals unity. With these two simplifications, Equation G-33 can be written as

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]\right\}dz \qquad \text{G-35}$$

Updating Equation G-27with Equations G-28 through G-30 generates the exact solution of the output standard deviation squared,

$$\sigma^2_{output} = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[z\sigma_{input}+\mu_{input}+DeltaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right)dz \quad \text{G-36}$$
$$-(\mu_{output})^2$$

Expanding this equation results in

$$\sigma^2_{output} = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\begin{matrix}z^2\sigma^2_{input}+2z\sigma_{input}\mu_{input}+\mu^2_{input}+2z\sigma_{input}DeltaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\\+2\mu_{input}DeltaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)+\left[DeltaP\left(\mu_{target}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\end{matrix}\right]\right)dz \quad \text{G-37}$$
$$-(\mu_{output})^2$$

Expanding even further produces

$$\sigma^2_{output} = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]^2\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[z^2\sigma^2_{input}\right]\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}\mu_{input}\right]\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\mu^2_{input}\right]\right)dz \qquad \text{G-38}$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2\mu_{input}DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left[DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]^2\right]\right)dz$$
$$-(\mu_{output})^2$$

Using Equation G-34 to simplify the above equation results in

$$\sigma^2_{output} = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]^2\right)dz$$
$$+1\sigma^2_{input}$$
$$+0$$
$$+1\mu^2_{input}$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]\right)dz \qquad \text{G-39}$$
$$+2\mu_{input}\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]\right)dz$$
$$+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left[DeltaP\left(\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}\right)\right]^2\right]\right)dz$$
$$-(\mu_{output})^2$$

The integral in the sixth term of the above equation appears in Equation G-35. Using this fact, Equation G-39 can be further simplified to

$$
\begin{aligned}
\sigma_{output}^2 = & \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right)dz \\
& +1\sigma_{input}^2 \\
& +0 \\
& +1\mu_{input}^2 \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}DeltaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]\right)dz \\
& +2\mu_{input}\left\{\mu_{output}-\mu_{input}\right\} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left[DeltaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right]\right)dz \\
& -(\mu_{output})^2
\end{aligned}
\qquad \text{G-40}
$$

The above equation can be rewritten in a more compact form to result in

$$
\begin{aligned}
\sigma_{output}^2 = & \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[SigmaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right)dz \\
& +\sigma_{input}^2+\mu_{input}^2 \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}DeltaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]\right)dz \\
& +2\mu_{input}\left\{\mu_{output}-\mu_{input}\right\} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left[DeltaP\left(\mu_{t\arg et}-\left\{z\sigma_{input}+\mu_{input}\right\}\right)\right]^2\right]\right)dz-(\mu_{output})^2
\end{aligned}
\qquad \text{G-41}
$$

The mean of the output distribution can be calculated using Equation G-35 once the equation for DeltaP is known. At this point, the exact form of DeltaP is not known; therefore, a general form is presented so that a closed-form solution for μ_{output} can be derived.

A general form for DeltaP is

$$DeltaP = a_0 + a_1 M + a_2 M^2 + \ldots + a_i M^i + \ldots + a_\alpha M^\alpha \qquad \text{G-42}$$

where the a_i's are numerical coefficients and M is the desired dimensional change and is equal to $\mu_{target} - \{z\sigma_{input} + \mu_{input}\}$. A method for finding DeltaP for a particular process is given in Section 3.4.

Combining Equations G-35 and G-42 results in

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[a_0 + a_1 M + a_2 M^2 + \ldots + a_k M^k + \ldots + a_\alpha M^\alpha \right] \right\} dz \qquad \text{G-43}$$

Expanding this equation and remembering that M is equal to $\mu_{target} - \{z\sigma_{input} + \mu_{input}\}$ results in

$$\begin{aligned} \mu_{output} &= \mu_{input} \\ &+ A_0 \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz + A_1 \int_{-\infty}^{+\infty} \left\{ \frac{z}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz \\ &+ A_2 \int_{-\infty}^{+\infty} \left\{ \frac{z^2}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz + \ldots + A_k \int_{-\infty}^{+\infty} \left\{ \frac{z^k}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz \\ &+ \ldots + A_\alpha \int_{-\infty}^{+\infty} \left\{ \frac{z^\alpha}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz \end{aligned} \qquad \text{G-44}$$

The coefficients, A_i's, are the result of the DeltaP coefficients being multiplied with the terms of M. They are simply combinations of the coefficients of the DeltaP function (a0, a1,…,aα), the standard deviation of the original input distribution, the mean of the original input distribution, and the target mean.

For simplicity, Equation G-34 is being restated below as Equation G-45.

$$\int_{-\infty}^{+\infty}\left\{\frac{z^k}{\sqrt{2\pi}}\left[\exp\left(-\frac{z^2}{2}\right)\right]\right\}dz=\begin{cases}1 & \text{for k = 0}\\ 0 & \text{for odd k's}\\ \prod_{j=1}^{\frac{k}{2}}\left[k-\{2j-1\}\right] & \text{for even k's}\end{cases} \qquad \text{G-45}$$

Using the above equation to simplify Equation G-44 results in the compact form of

$$\begin{aligned}\mu_{output} &= \mu_{input} + A_0 + A_2\prod_{j=1}^{1}\left[2-\{2j-1\}\right] + A_4\prod_{j=1}^{2}\left[4-\{2j-1\}\right]\\ &+\ldots+A_k\prod_{j=1}^{\frac{k}{2}}\left[k-\{2j-1\}\right]+\ldots+A_\alpha\prod_{j=1}^{\frac{\alpha}{2}}\left[\alpha-\{2j-1\}\right]\end{aligned} \qquad \text{G-46}$$

which can be further reduced to

$$\begin{aligned}\mu_{output} &= \mu_{input} + A_0 + A_2 + 3A_4 + \ldots\\ &+A_k\prod_{j=1}^{\frac{k}{2}}\left[k-\{2j-1\}\right]+\ldots+A_\alpha\prod_{j=1}^{\frac{\alpha}{2}}\left[\alpha-\{2j-1\}\right]\end{aligned} \qquad \text{G-47}$$

for all even k's. Remember $A_i = f\left(\mu_{target}, \mu_{input}, \sigma_{input}, a_0, a_1, \ldots, a_k, \ldots, a_\alpha\right)$.

The above equation is very powerful in that it is a closed-form solution for the output distribution's mean.

The variance of the output distribution can be calculated using Equation G-41 once the equation for SigmaP is known. At this point, the exact form of SigmaP is not known; therefore, a general form is presented so that a closed-form solution for σoutput can be derived.

A general form for SigmaP is

$$SigmaP = b_0 + b_1 M + b_2 M^2 + \ldots + b_k M^k + \ldots + b_\beta M^\beta \qquad \text{G-48}$$

where the bi's are numerical coefficients and M is the desired dimensional change and is equal to $\mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\}$.

Combining Equations G-41, G-42, and G-48 results in

$$\begin{aligned}
\sigma^2_{output} = & \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[b_0 + b_1 M + b_2 M^2 + \ldots + b_k M^k + \ldots + b_\beta M^\beta \right]^2 \right) dz \\
& + \sigma^2_{input} + \mu^2_{input} \\
& + \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[2z\sigma_{input} \left\{ a_0 + a_1 M + a_2 M^2 + \ldots + a_k M^k + \ldots + a_\alpha M^\alpha \right\} \right] \right) dz \\
& + 2\mu_{input} \left\{ \mu_{output} - \mu_{input} \right\} \\
& + \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[\left\{ a_0 + a_1 M + a_2 M^2 + \ldots + a_k M^k + \ldots + a_\alpha M^\alpha \right\}^2 \right] \right) dz - (\mu_{output})^2
\end{aligned} \qquad \text{G-49}$$

where M equals $\mu_{target} - \left\{ z * \sigma_{input} + \mu_{input} \right\}$.

The length and complexity of the above equation lends itself to being reduced term by term. Term 1, Term 4, and Term 6 will be discussed individually in detail.

Term 1 is

Term 1 of

$$\sigma^2_{output} = \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[b_0 + b_1 M + b_2 M^2 + \ldots + b_k M^k + \ldots + b_\beta M^\beta \right]^2 \right) dz \qquad \text{G-50}$$

which can be expanded to

$$\text{Term 1} = B_0 \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz + B_1 \int_{-\infty}^{+\infty} \left\{ \frac{z}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz$$
$$+ B_2 \int_{-\infty}^{+\infty} \left\{ \frac{z^2}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz + \ldots + B_k \int_{-\infty}^{+\infty} \left\{ \frac{z^k}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz \qquad \text{G-51}$$
$$+ \ldots + B_{2\beta} \int_{-\infty}^{+\infty} \left\{ \frac{z^{2\beta}}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz$$

The coefficients, B_i's, are the result of the SigmaP term being squared and multiplied with the terms of M. They are simply combinations of the coefficients of the SigmaP function (b0, b1,…,bβ), the standard deviation of the original input distribution, the mean of the original input distribution, and the target mean.

Using Equation G-45, the equation above can be rewritten as

$$\text{Term 1} = B_0 + B_2 \prod_{j=1}^{1} \left[2 - \{2j-1\} \right] + B_4 \prod_{j=1}^{2} \left[4 - \{2j-1\} \right]$$
$$+ \ldots + B_k \prod_{j=1}^{\frac{k}{2}} \left[k - \{2j-1\} \right] + \ldots + B_{2\beta} \prod_{j=1}^{\frac{2\beta}{2}} \left[2\beta - \{2j-1\} \right] \qquad \text{G-52}$$

which can be further reduced to

$$\text{Term 1} = B_0 + B_2 + 3B_4 + \ldots + B_k \prod_{j=1}^{\frac{k}{2}} \left[k - \{2j-1\} \right] + \ldots + B_{2\beta} \prod_{j=1}^{\beta} \left[2\beta - \{2j-1\} \right] \qquad \text{G-53}$$

Now, attention is turned to Term 4, which is

Term 4 of

$$\sigma^2_{output} = \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[2z\sigma_{input} \left\{ a_0 + a_1 M + a_2 M^2 + \ldots + a_k M^k + \ldots + a_\alpha M^\alpha \right\} \right] \right) dz \qquad \text{G-54}$$

The equation above can be expanded to produce

$$\text{Term } 4 = 2\sigma_{input} A_0 \int_{-\infty}^{+\infty} \left\{ \frac{z}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz +$$
$$2\sigma_{input} A_1 \int_{-\infty}^{+\infty} \left\{ \frac{z^2}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz$$
$$+2\sigma_{input} A_2 \int_{-\infty}^{+\infty} \left\{ \frac{z^3}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz \qquad \text{G-55}$$
$$+\ldots + 2\sigma_{input} A_k \int_{-\infty}^{+\infty} \left\{ \frac{z^{k+1}}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz$$
$$+\ldots + 2\sigma_{input} A_\alpha \int_{-\infty}^{+\infty} \left\{ \frac{z^{\alpha+1}}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \right\} dz$$

The above equation can be combined with Equation G-45to result in

$$\text{Term } 4 = 2\sigma_{input} A_1 \prod_{j=1}^{1} \left[2 - \{2j-1\} \right] + 2\sigma_{input} A_3 \prod_{j=1}^{2} \left[4 - \{2j-1\} \right]$$
$$+\ldots + 2\sigma_{input} A_k \prod_{j=1}^{\frac{k+1}{2}} \left[k - \{2j-1\} \right] + \ldots + 2\sigma_{input} A_\alpha \prod_{j=1}^{\frac{\alpha+1}{2}} \left[\alpha - \{2j-1\} \right] \qquad \text{G-56}$$

for all odd k's. The above equation simplifies to

$$\text{Term } 4 = 2\sigma_{input} A_1 + 6\sigma_{input} A_3 + \ldots + 2\sigma_{input} A_k \prod_{j=1}^{\frac{k+1}{2}} \left[k - \{2j-1\} \right]$$
$$+\ldots + 2\sigma_{input} A_\alpha \prod_{j=1}^{\frac{\alpha+1}{2}} \left[\alpha - \{2j-1\} \right] \qquad \text{G-57}$$

Now, Term 6 is discussed.

Term 6 of

$$\sigma^2_{output} = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left\{a_0 + a_1M + a_2M^2 + ... + a_kM^k + ... + a_\alpha M^\alpha\right\}^2\right]\right)dz \qquad \text{G-58}$$

Term 6 can be expanded to

$$\begin{aligned} \text{Term 6} &= C_0\int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\right\}dz + C_1\int_{-\infty}^{+\infty}\left\{\frac{z}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\right\}dz \\ &+ C_2\int_{-\infty}^{+\infty}\left\{\frac{z^2}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\right\}dz + ... + C_k\int_{-\infty}^{+\infty}\left\{\frac{z^k}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\right\}dz \\ &+ ... + C_{2\alpha}\int_{-\infty}^{+\infty}\left\{\frac{z^{2\alpha}}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\right\}dz \end{aligned} \qquad \text{G-59}$$

The coefficients, C_i's, are the result of the DeltaP term being squared and multiplied with the terms of M. They are simply combinations of the coefficients of the DeltaP function (a0, a1,…,aα), the standard deviation of the original input distribution, the mean of the original input distribution, and the target mean.

Using Equation G-45, Equation G-51 can be rewritten as

$$\begin{aligned} \text{Term 6} &= C_0 + C_2\prod_{j=1}^{1}\left[2-\{2j-1\}\right] + C_4\prod_{j=1}^{2}\left[4-\{2j-1\}\right] \\ &+ ... + C_k\prod_{j=1}^{\frac{k}{2}}\left[k-\{2j-1\}\right] + ... + C_{2\alpha}\prod_{j=1}^{\frac{2\alpha}{2}}\left[2\alpha-\{2j-1\}\right] \end{aligned} \qquad \text{G-60}$$

which can be further reduced to

$$\text{Term } 6 = C_0 + C_2 + 3C_4 + \ldots + C_k \prod_{j=1}^{\frac{k}{2}} \left[k - \{2j-1\} \right] + \ldots + C_{2\alpha} \prod_{j=1}^{\alpha} \left[2\alpha - \{2j-1\} \right] \qquad \text{G-61}$$

Recombining all the terms results in the following closed form solution for σoutput

$$\begin{aligned}
\sigma_{output}^2 = & \left\{ B_0 + B_2 + 3B_4 + \ldots + B_k \prod_{j=1}^{\frac{k}{2}} \left[k - \{2j-1\} \right] + \ldots + B_{2\beta} \prod_{j=1}^{\beta} \left[2\beta - \{2j-1\} \right] \right\} \\
& + \sigma_{input}^2 + \mu_{input}^2 \\
& + \left\{ 2\sigma_{input} A_1 + 6\sigma_{input} A_3 + \ldots + 2\sigma_{input} A_k \prod_{j=1}^{\frac{k+1}{2}} \left[k - \{2j-1\} \right] + \ldots + 2\sigma_{input} A_\alpha \prod_{j=1}^{\frac{\alpha+1}{2}} \left[\alpha - \{2j-1\} \right] \right\} \\
& + 2\mu_{input} \left\{ \mu_{output} - \mu_{input} \right\} \\
& + \left\{ C_0 + C_2 + 3C_4 + \ldots + C_k \prod_{j=1}^{\frac{k}{2}} \left[k - \{2j-1\} \right] + \ldots + C_{2\alpha} \prod_{j=1}^{\alpha} \left[2\alpha - \{2j-1\} \right] \right\} - (\mu_{output})^2
\end{aligned} \qquad \text{G-62}$$

where

$$A_i = f\left(\mu_{target}, \mu_{input}, \sigma_{input}, a_0, a_1, \ldots, a_k, \ldots, a_\alpha \right)$$

$$B_i = f\left(\mu_{target}, \mu_{input}, \sigma_{input}, b_0, b_1, \ldots, b_k, \ldots, b_\alpha \right)$$

$$C_i = f\left(\mu_{target}, \mu_{input}, \sigma_{input}, a_0, a_1, \ldots, a_k, \ldots, a_\alpha \right).$$

The equation for the output mean can be greatly simplified for special cases of DeltaP. This section presents three such simple cases.

The simplest case is when DeltaP is equal to a constant, a_0. Substituting this constant into Equation G-35 results in

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[a_0 \right] \right\} dz \qquad \text{G-63}$$

which reduces to

$$\mu_{output} = \mu_{input} + a_0 \quad \text{G-64}$$

The equation above makes physical sense because a constant DeltaP implies that every piece in the incoming distribution undergoes the exact same change in dimension, a_0.

The next simplest case is when DeltaP is a linear function with a y-intercept at zero. This results in

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[a_0 + a_1 M \right] \right\} dz \quad \text{G-65}$$

where M equals $\mu_{target} - \{ z * \sigma_{input} + \mu_{input} \}$. Making this substitution results in

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[a_0 + a_1 \left\{ \mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\} \right\} \right] \right\} dz \quad \text{G-66}$$

which reduces to

$$\begin{aligned} \mu_{output} = \mu_{input} &+ \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[a_0 + a_1 \mu_{target} - a_1 \mu_{input} \right] \right\} dz \\ &+ \int_{-\infty}^{+\infty} \left\{ \frac{1}{\sqrt{2\pi}} \left\{ \exp\left(-\frac{z^2}{2} \right) \right\} \left[-a_1 \sigma_{input} z \right] \right\} dz \end{aligned} \quad \text{G-67}$$

The above equation can be simplified using Equation G-45 to produce

$$\mu_{output} = \mu_{input} + a_0 + a_1 \mu_{target} - a_1 \mu_{input} \quad \text{G-68}$$

The equation above is an equation for the output distribution's mean when the process has a DeltaP that is linear with a y-intercept of a_0 and a slope of a_1.

Another special case of interest is when DeltaP is a quadratic function with a y-intercept at zero, i.e., DeltaP(M) = $a_0 + a_1*M + a_2*M^2$.

$$\mu_{output} = \mu_{input} + \int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[a_0 + a_1 M + a_2 M^2\right]\right\}dz \qquad \text{G-69}$$

where M equals $\mu_{target} - \{z\sigma_{input} + \mu_{input}\}$.

Expanding the above equation results in

$$\begin{aligned}\mu_{output} &= \mu_{input} + a_0 \\ &+ a_1\int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[\mu_{target} - \{z*\sigma_{input} + \mu_{input}\}\right]\right\}dz \\ &+ a_2\int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[\mu_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2 + 2\mu_{input}\sigma_{input}z - 2\mu_{target}\sigma_{input}z + \sigma_{input}^2 z^2\right]\right\}dz\end{aligned} \qquad \text{G-70}$$

Grouping terms in z produces

$$\begin{aligned}\mu_{output} &= \mu_{input} + a_0 \\ &+ \int_{-\infty}^{+\infty}\left\{\frac{1}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[a_1\mu_{target} - a_1\mu_{input} + a_2\mu_{input}^2 - 2a_2\mu_{input}\mu_{target} + a_2\mu_{target}^2\right]\right\}dz \\ &+ \int_{-\infty}^{+\infty}\left\{\frac{z}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[-a_1\sigma_{input} + 2a_2\mu_{input}\sigma_{input} - 2a_2\mu_{target}\sigma_{input}\right]\right\}dz \\ &+ \int_{-\infty}^{+\infty}\left\{\frac{z^2}{\sqrt{2\pi}}\left\{\exp\left(-\frac{z^2}{2}\right)\right\}\left[a_2\sigma_{input}^2\right]\right\}dz\end{aligned} \qquad \text{G-71}$$

Simplify this equation using Equation G-45 produces

$$\mu_{output} = \mu_{input} + a_0 + a_1\mu_{target} - a_1\mu_{input} \\ +a_2\mu_{input}^2 - 2a_2\mu_{input}\mu_{target} + a_2\mu_{target}^2 + a_2\sigma_{input}^2 \quad \text{G-72}$$

The like terms of this equation can be grouped together to result in

$$\mu_{output} = \mu_{input} + a_0 + a_1\left\{\mu_{target} - \mu_{input}\right\} \\ +a_2\left\{\mu_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2 + \sigma_{input}^2\right\} \quad \text{G-73}$$

The equation for the output standard deviation can be greatly simplified for special cases of SigmaP. This section presents three such simple cases. For each special case of SigmaP, DeltaP will be represented by a quadratic equation.

The most simple equation for the output standard deviation is when SigmaP is equal to a constant regardless of the desired dimensional change, i.e., SigmaP(M) = b_0. DeltaP will be represented by $a_0 + a_1M + a_2M^2$. Focusing on this simple form of SigmaP allows the development of the terms containing DeltaP. The next two sections will use these developments and expand the complexity of SigmaP.

The formulae for SigmaP and DeltaP are inserted into Equation G-41 to produce

$$\sigma_{output}^2 = \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)[b_0]^2\right)dz + \sigma_{input}^2 + \mu_{input}^2 \\ + \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}\left\{a_0 + a_1M + a_2M^2\right\}\right]\right)dz \\ +2\mu_{input}\left\{\mu_{output} - \mu_{input}\right\} \\ + \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left\{a_0 + a_1M + a_2M^2\right\}^2\right]\right)dz - (\mu_{output})^2 \quad \text{G-74}$$

Replacing M in Equation G-74 with $\mu_{target} - \left\{z\sigma_{input} + \mu_{input}\right\}$ results in

$$
\begin{aligned}
\sigma_{output}^{2} = & \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)[b_{0}]^{2}\right)dz+\sigma_{input}^{2}+\mu_{input}^{2} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)\left[2z\sigma_{input}\left\{a_{0}+a_{1}\left\{\mu_{target}-z\sigma_{input}-\mu_{input}\right\}+a_{2}\left\{\mu_{target}-z\sigma_{input}-\mu_{input}\right\}^{2}\right\}\right]\right)dz \\
& +2\mu_{input}\left\{\mu_{output}-\mu_{input}\right\} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)\left[\left\{a_{0}+a_{1}\left\{\mu_{target}-z\sigma_{input}-\mu_{input}\right\}+a_{2}\left\{\mu_{target}-z\sigma_{input}-\mu_{input}\right\}^{2}\right\}^{2}\right]\right)dz \\
& -(\mu_{output})^{2}
\end{aligned}
\qquad \text{G-75}
$$

Expanding this equation, combining terms, and eliminating odd powered terms in z per Equation G-45 results in

$$
\begin{aligned}
\sigma_{output}^{2} = & b_{0}^{2}+\sigma_{input}^{2}+\mu_{input}^{2}+2\mu_{input}\left\{\mu_{output}-\mu_{input}\right\}-(\mu_{output})^{2} \\
& +a_{0}^{2}+a_{1}^{2}\left\{\mu_{input}^{2}+\sigma_{input}^{2}-2\mu_{input}\mu_{target}+\mu_{target}^{2}\right\} \\
& +2a_{0}\left\{a_{1}\left[\mu_{target}-\mu_{input}\right]+a_{2}\left[\mu_{input}^{2}+\sigma_{input}^{2}-2\mu_{input}\mu_{target}+\mu_{target}^{2}\right]\right\} \\
& -2a_{1}\left\{\sigma_{input}^{2}+a_{2}\left[\mu_{input}-\mu_{target}\right]\left[\mu_{input}^{2}+3\sigma_{input}^{2}-2\mu_{input}\mu_{target}+\mu_{target}^{2}\right]\right\} \\
& +4a_{2}\sigma_{input}^{2}\left[\mu_{input}-\mu_{target}\right] \\
& +a_{2}^{2}\left[\mu_{input}^{4}+3\sigma_{input}^{4}-4\mu_{input}^{3}\mu_{target}+6\sigma_{input}^{2}\mu_{target}^{2}+\mu_{target}^{4}+6\mu_{input}^{2}\left\{\sigma_{input}^{2}+\mu_{target}^{2}\right\}-4\mu_{input}\left\{3\sigma_{input}^{2}\mu_{target}+\mu_{target}^{3}\right\}\right]
\end{aligned}
\qquad \text{G-76}
$$

A very similar case to the one above is when SigmaP is a linear function with a slope equal to b a_1 and a y-intercept equal to b_0, i.e., SigmaP(M) = b_0 + b_1*M. The formulae for SigmaP and DeltaP are inserted into Equation G-41 to produce

$$
\begin{aligned}
\sigma_{output}^{2} = & \int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)[b_{0}+b_{1}M]^{2}\right)dz+\sigma_{input}^{2}+\mu_{input}^{2} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)\left[2z\sigma_{input}\left\{a_{0}+a_{1}M+a_{2}M^{2}\right\}\right]\right)dz \\
& +2\mu_{input}\left\{\mu_{output}-\mu_{input}\right\} \\
& +\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^{2}}{2}\right\}\right)\left[\left\{a_{0}+a_{1}M+a_{2}M^{2}\right\}^{2}\right]\right)dz-(\mu_{output})^{2}
\end{aligned}
\qquad \text{G-77}
$$

Replacing M in this equation with $\mu_{target} - \left\{ z\sigma_{input} + \mu_{input} \right\}$ and using the simplified results for the terms not containing SigmaP derived earlier produces

$$
\begin{aligned}
\sigma^2_{output} = & \int_{-\infty}^{+\infty} \left(\frac{1}{\sqrt{2\pi}} \left(\exp\left\{ -\frac{z^2}{2} \right\} \right) \left[b_0 + b_1 \left\{ \mu_{target} - z\sigma_{input} - \mu_{input} \right\} \right]^2 \right) dz \\
& + \sigma^2_{input} + \mu^2_{input} + 2\mu_{input} \left\{ \mu_{output} - \mu_{input} \right\} - (\mu_{output})^2 \\
& + a_0^2 + a_1^2 \left\{ \mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right\} \\
& + 2a_0 \left\{ a_1 \left[\mu_{target} - \mu_{input} \right] + a_2 \left[\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right] \right\} \\
& - 2a_1 \left\{ \sigma^2_{input} + a_2 \left[\mu_{input} - \mu_{target} \right] \left[\mu^2_{input} + 3\sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right] \right\} \\
& + 4a_2\sigma^2_{input} \left[\mu_{input} - \mu_{target} \right] \\
& + a_2^2 \left[\mu^4_{input} + 3\sigma^4_{input} - 4\mu^3_{input}\mu_{target} + 6\sigma^2_{input}\mu^2_{target} + \mu^4_{target} + 6\mu^2_{input} \left\{ \sigma^2_{input} + \mu^2_{target} \right\} \right] \\
& - 4\mu_{input} a_2^2 \left\{ 3\sigma^2_{input}\mu_{target} + \mu^3_{target} \right\}
\end{aligned}
\qquad \text{G-78}
$$

Expanding the above equation, combining terms, and eliminating odd powered terms in z per Equation G-45 results in

$$
\begin{aligned}
\sigma^2_{output} = & b_0^2 + b_1^2 \left\{ \mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right\} \\
& + 2b_0 \left\{ b_1 \left[\mu_{target} - \mu_{input} \right] \right\} \\
& + \sigma^2_{input} + \mu^2_{input} + 2\mu_{input} \left\{ \mu_{output} - \mu_{input} \right\} - (\mu_{output})^2 \\
& + a_0^2 + a_1^2 \left\{ \mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right\} \\
& + 2a_0 \left\{ a_1 \left[\mu_{target} - \mu_{input} \right] + a_2 \left[\mu^2_{input} + \sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right] \right\} \\
& - 2a_1 \left\{ \sigma^2_{input} + a_2 \left[\mu_{input} - \mu_{target} \right] \left[\mu^2_{input} + 3\sigma^2_{input} - 2\mu_{input}\mu_{target} + \mu^2_{target} \right] \right\} \\
& + 4a_2\sigma^2_{input} \left[\mu_{input} - \mu_{target} \right] \\
& + a_2^2 \left[\mu^4_{input} + 3\sigma^4_{input} - 4\mu^3_{input}\mu_{target} + 6\sigma^2_{input}\mu^2_{target} + \mu^4_{target} \right] \\
& + a_2^2 \left[6\mu^2_{input} \left\{ \sigma^2_{input} + \mu^2_{target} \right\} - 4\mu_{input} \left\{ 3\sigma^2_{input}\mu_{target} + \mu^3_{target} \right\} \right]
\end{aligned}
\qquad \text{G-79}
$$

A very similar case to the one above is when SigmaP is a quadratic function with a general y-intercept equal to a_0, i.e., SigmaP(M) = $a_0 + a_1 * M + a_2 * M^2$. The formulae for SigmaP and DeltaP are inserted into Equation G-41 to produce

$$
\begin{aligned}
\sigma_{output}^2 = &\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[b_0 + b_1 M + b_2 M^2\right]^2\right)dz + \sigma_{input}^2 + \mu_{input}^2 \\
&+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[2z\sigma_{input}\left\{a_0 + a_1 M + a_2 M^2\right\}\right]\right)dz \\
&+2\mu_{input}\left\{\mu_{output} - \mu_{input}\right\} \\
&+\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[\left\{a_0 + a_1 M + a_2 M^2\right\}^2\right]\right)dz - (\mu_{output})^2
\end{aligned}
\qquad \text{G-80}
$$

Replacing M in Equation the above equation with $\mu_{target} - \{z * \sigma_{input} + \mu_{input}\}$ and using the simplified results for the terms not containing SigmaP derived earlier produces

$$
\begin{aligned}
\sigma_{output}^2 = &\int_{-\infty}^{+\infty}\left(\frac{1}{\sqrt{2\pi}}\left(\exp\left\{-\frac{z^2}{2}\right\}\right)\left[b_0 + b_1\left\{\mu_{target} - z\sigma_{input} - \mu_{input}\right\} + b_2\left\{\mu_{target} - z\sigma_{input} - \mu_{input}\right\}^2\right]^2\right)dz \\
&+\sigma_{input}^2 + \mu_{input}^2 + 2\mu_{input}\left\{\mu_{output} - \mu_{input}\right\} - (\mu_{output})^2 \\
&+a_0^2 + a_1^2\left\{\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right\} \\
&+2a_0\left\{a_1\left[\mu_{target} - \mu_{input}\right] + a_2\left[\mu_{input}^2 + \sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right]\right\} \\
&-2a_1\left\{\sigma_{input}^2 + a_2\left[\mu_{input} - \mu_{target}\right]\left[\mu_{input}^2 + 3\sigma_{input}^2 - 2\mu_{input}\mu_{target} + \mu_{target}^2\right]\right\} \\
&+4a_2\sigma_{input}^2\left[\mu_{input} - \mu_{target}\right] \\
&+a_2^2\left[\mu_{input}^4 + 3\sigma_{input}^4 - 4\mu_{input}^3\mu_{target} + 6\sigma_{input}^2\mu_{target}^2 + \mu_{target}^4 + 6\mu_{input}^2\left\{\sigma_{input}^2 + \mu_{target}^2\right\}\right] \\
&-4\mu_{input}a_2^2\left\{3\sigma_{input}^2\mu_{target} + \mu_{target}^3\right\}
\end{aligned}
\qquad \text{G-81}
$$

Expanding this equation, combining terms, and eliminating odd powered terms in z per Equation G-45 results in

$$
\begin{aligned}
\sigma_{output}^{2} = {} & b_0^2 + b_1^2 \left\{ \mu_{input}^{2} + \sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right\} \\
& -2b_1 b_2 \left[\mu_{input} - \mu_{target} \right]\left[\mu_{input}^{2} + 3\sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right] \\
& +2b_0 \left\{ b_1 \left[\mu_{target} - \mu_{input} \right] + b_2 \left[\mu_{input}^{2} + \sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right] \right\} \\
& +b_2^2 \left[\mu_{input}^{4} + 3\sigma_{input}^{4} - 4\mu_{input}^{3}\mu_{target} + 6\sigma_{input}^{2}\mu_{target}^{2} + \mu_{target}^{4} + 6\mu_{input}^{2} \left\{ \sigma_{input}^{2} + \mu_{target}^{2} \right\} - 4\mu_{input} \left\{ 3\sigma_{input}^{2}\mu_{target} + \mu_{target}^{3} \right\} \right] \\
& +\sigma_{input}^{2} + \mu_{input}^{2} + 2\mu_{input} * \left\{ \mu_{output} - \mu_{input} \right\} - (\mu_{output})^2 + a_0^2 + a_1^2 \left\{ \mu_{input}^{2} + \sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right\} \\
& +2a_0 \left\{ a_1 \left[\mu_{target} - \mu_{input} \right] + a_2 \left[\mu_{input}^{2} + \sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right] \right\} \\
& -2a_1 \left\{ \sigma_{input}^{2} + a_2 \left[\mu_{input} - \mu_{target} \right] * \left[\mu_{input}^{2} + 3\sigma_{input}^{2} - 2\mu_{input}\mu_{target} + \mu_{target}^{2} \right] \right\} + 4a_2\sigma_{input}^{2} \left[\mu_{input} - \mu_{target} \right] \\
& +a_2^2 \left[\mu_{input}^{4} + 3\sigma_{input}^{4} - 4\mu_{input}^{3}\mu_{target} + 6\sigma_{input}^{2}\mu_{target}^{2} + \mu_{target}^{4} + 6\mu_{input}^{2} \left\{ \sigma_{input}^{2} + \mu_{target}^{2} \right\} - 4\mu_{input} \left\{ 3\sigma_{input}^{2}\mu_{target} + \mu_{target}^{3} \right\} \right]
\end{aligned}
\qquad \text{G-82}
$$

H. Heuristic Guideline Development

The development of the heuristic guidelines presented in Section 3.7 is shown in this appendix. The amount of process capability data required to construct accurate DeltaP and SigmaP functions is discussed. Also, the range of possible input values the closed-form equations are valid over is investigated. In other words, this appendix attempts to answer the following two questions:

- Over what range are the constructed equations valid?
- How much process capability data are required to construct accurate, useful closed-form equations?

The validation in Section 3.6 demonstrated the ability of the closed-form equations to calculate the output mean and standard deviation of a diameter leaving a turning process. The DeltaP and SigmaP coefficients were computed using the ten sets (a set is comprised of μ_{in}, σ_{in}, μ_{out} and σ_{out}) of simulated process capability data shown in Table 3-3, which has values for the input mean spanning the range from 34.6 to 39.9 and values for the input standard deviation spanning from 0.19 to 0.60. Notice that these values encompass the entire range of values in Table 3-4 used to validate the equations. This implies that the closed-form equations are at least valid for input values, μ_{in} and σ_{in}, that fall inside the range of values originally used to construct DeltaP and SigmaP, i.e., the closed-form equations do a good job of interpolation.

Using the closed-form equations with input values that fall outside the range of process capability data would often times be very beneficial. For example, if a larger diameter bar stock than any that had been previously used becomes available for a cheaper price, understanding the effect on the output mean and standard deviation would be helpful in determining whether this new bar stock would be acceptable. Therefore, an important question to ask is how valid are the closed-form equations outside of the range of values used to construct them, e.g., can the equations be used with an input of 39 if the process capability data used to construct them only spans input values from 32 to 36? In other words, how good are the equations when extrapolation is necessary? In addition to this question, the amount of process capability data required to construct accurate, useful closed-form equations is an important consideration, i.e., how much data is enough to build "good" equations?

Quantitatively deriving universal answers to these questions is impossible because of the infinite number of combinations of variables; however, qualitative answers can be developed. The turning physics model is used in this section to explore answers to these questions. The goal is to determine heuristic rules that can be applied to the construction of closed-form equations for any process.

The steps used to answer these questions are similar to the steps used to validate the method in Section 3.6. Step one is to simulate process capability data that can be used to construct DeltaP and SigmaP functions. The data are generated in a very structured manner. The input mean values range from 32.0 to 40.0, and the standard deviation values are 0.15, 0.30, and 0.45. This produces the twenty seven sets of process capability data shown below in Table H-1. The first column identifies the set of data with a number. The second and third columns are the values of the input mean and standard deviation used to generate one thousand normally distributed random numbers. The mean and standard deviation of the thousand random numbers are listed in columns four and five. The last two columns list the simulated output values.

Table H-1: Simulated capability data sets used to construct DeltaP & SigmaP

	Incoming Distribution Values					
Sample Set #	Value for Mean Used to Generate Random Numbers	Value for Standard Deviation Used to Generate Random Numbers	Calculated μ of 1,000 Generated Random Numbers	Calculated σ of 1,000 Generated Random Numbers	Simulated Output Mean, μ_{out} [mm]	Simulated Output Standard Deviation, σ_{out} [mm]
1	32.00	0.15	32.0007	0.1538	30.0035	0.0006
2	32.00	0.30	31.9953	0.3016	30.0035	0.0011
3	32.00	0.45	32.0219	0.4581	30.0037	0.0016
4	33.00	0.15	33.0009	0.1528	30.0077	0.0011
5	33.00	0.30	33.0042	0.2972	30.0078	0.0017
6	33.00	0.45	33.0028	0.4553	30.0079	0.0025
7	34.00	0.15	33.9967	0.1489	30.0135	0.0017
8	34.00	0.30	33.9929	0.3047	30.0135	0.0024
9	34.00	0.45	34.0073	0.4636	30.0137	0.0034
10	35.00	0.15	34.992	0.1526	30.0209	0.0025
11	35.00	0.30	35.0038	0.3067	30.0211	0.0033
12	35.00	0.45	34.9984	0.4311	30.0211	0.0042
13	36.00	0.15	36.0034	0.1495	30.0301	0.0034
14	36.00	0.30	36.0033	0.3072	30.0301	0.0042
15	36.00	0.45	35.9693	0.4555	30.0299	0.0055
16	37.00	0.15	36.9962	0.1502	30.0407	0.0044

17	37.00	0.30	36.9987	0.3158	30.0408	0.0055
18	37.00	0.45	36.9744	0.437	30.0406	0.0067
19	38.00	0.15	37.9977	0.1559	30.0529	0.0059
20	38.00	0.30	38.0017	0.3064	30.0531	0.0069
21	38.00	0.45	37.9705	0.4458	30.0527	0.0078
22	39.00	0.15	39.0035	0.1545	30.0668	0.0072
23	39.00	0.30	39.0053	0.2978	30.0669	0.008
24	39.00	0.45	39.0121	0.4496	30.0671	0.0094
25	40.00	0.15	39.9971	0.1485	30.0821	0.0088
26	40.00	0.30	40.0052	0.3081	30.0823	0.0098
27	40.00	0.45	40.0118	0.4293	30.0825	0.0112

All values in millimeters.

The second step is to generate test cases of simulated data. These data are generated with the same model that produced the values in Table H-1. These data are taken to be the "true values" that the manufacturing process would produce, i.e., the turning process math model is used to "operate" on a thousand random numbers to produce a thousand values of output diameters. The test case input values were picked to span a similar range as the process capability data in Table H-1, but not to have an identical set of data, i.e., no input mean and standard deviation combination that is exactly the same. The test cases of data are listed below in Table H-2.

Table H-2: Simulated capability data sets used for test cases

	Incoming Distribution Values					
Test Case #	Value for Mean Used to Generate Random Numbers	Value for Standard Deviation Used to Generate Random Numbers	Calculated μ of 1,000 Generated Random Numbers	Calculated σ of 1,000 Generated Random Numbers	Simulated Output Mean, μ_{out} [mm]	Simulated Output Standard Deviation, σ_{out} [mm]
1	32.50	0.10	32.4972	0.0986	30.0053	0.0007
2	32.50	0.29	32.4880	0.2857	30.0054	0.0013
3	32.50	0.50	32.4632	0.4992	30.0054	0.0022
4	33.50	0.10	33.5006	0.1015	30.0104	0.0012
5	33.50	0.31	33.4919	0.3056	30.0104	0.0021
6	33.50	0.52	33.5027	0.5236	30.0106	0.0033
7	34.50	0.10	34.4990	0.1008	30.0171	0.0019
8	34.50	0.29	34.5061	0.2948	30.0172	0.0027
9	34.50	0.50	34.5114	0.4971	30.0173	0.0041
10	35.50	0.10	35.5006	0.0994	30.0253	0.0027
11	35.50	0.30	35.5013	0.3034	30.0254	0.0037
12	35.50	0.51	35.5052	0.5135	30.0256	0.0054
13	36.50	0.10	36.5018	0.0996	30.0352	0.0037
14	36.50	0.30	36.5041	0.2991	30.0353	0.0047
15	36.50	0.50	36.5052	0.5043	30.0354	0.0065

16	37.50	0.10	37.4989	0.1021	30.0466	0.0050
17	37.50	0.29	37.5139	0.2911	30.0469	0.0061
18	37.50	0.48	37.4957	0.4848	30.0468	0.0076
19	38.50	0.10	38.4970	0.1023	30.0596	0.0062
20	38.50	0.31	38.5020	0.3099	30.0597	0.0074
21	38.50	0.51	38.5336	0.5139	30.0603	0.0095
22	39.50	0.10	39.4894	0.0977	30.0741	0.0077
23	39.50	0.30	39.5045	0.3045	30.0744	0.0090
24	39.50	0.48	39.4942	0.4840	30.0744	0.0108
25	40.50	0.11	40.4986	0.1055	30.0904	0.0095
26	40.50	0.30	40.5023	0.3018	30.0905	0.0107
27	40.50	0.49	40.5054	0.4914	30.0907	0.0127

All values in millimeters.

The third step is to use various sets of data listed in Table H-1 to construct DeltaP and SigmaP functions. These functions are used to construct the closed form equations for μ_{out} and σ_{out}. Then, the values for μ_{in} and σ_{in} in Table H-2 are input into these closed-form equations to compute μ_{out} and σ_{out}. Finally, the calculated values of μ_{out} and σ_{out} are compared to the "true values" of the output from Table H-2. The percentage differences are listed in the tables below. The percent differences show how well the closed-form equations computed the values of μ_{out} and σ_{out}.

Table H-3: Percent difference of output means

	Sample Data Sets Used to Construct DeltaP and SigmaP					
Test Case #	1 - 3	1 - 6	1 - 9	1 - 12	1 - 15	1 - 18
1	-0.0004%	0.0002%	0.0001%	0.0000%	0.0000%	0.0001%
2	-0.0004%	0.0002%	0.0001%	0.0000%	0.0000%	0.0002%
3	-0.0004%	0.0002%	0.0001%	0.0001%	0.0000%	0.0002%
4	-0.0018%	-0.0006%	0.0000%	-0.0001%	-0.0001%	0.0000%
5	-0.0018%	-0.0007%	0.0000%	-0.0001%	-0.0001%	0.0000%
6	-0.0019%	-0.0007%	0.0000%	0.0000%	-0.0001%	0.0000%
7	-0.0038%	-0.0019%	0.0000%	0.0000%	-0.0001%	-0.0001%
8	-0.0038%	-0.0019%	0.0000%	0.0000%	-0.0001%	-0.0001%
9	-0.0038%	-0.0019%	0.0000%	0.0000%	-0.0001%	-0.0001%
10	-0.0062%	-0.0034%	0.0002%	0.0001%	0.0000%	-0.0001%
11	-0.0062%	-0.0034%	0.0002%	0.0001%	0.0000%	-0.0001%
12	-0.0063%	-0.0035%	0.0002%	0.0001%	0.0000%	-0.0001%
13	-0.0091%	-0.0053%	0.0005%	0.0004%	0.0003%	0.0000%
14	-0.0091%	-0.0053%	0.0006%	0.0005%	0.0003%	0.0000%
15	-0.0092%	-0.0054%	0.0005%	0.0004%	0.0003%	-0.0001%
16	-0.0125%	-0.0075%	0.0011%	0.0009%	0.0007%	0.0002%
17	-0.0126%	-0.0075%	0.0011%	0.0009%	0.0007%	0.0002%
18	-0.0125%	-0.0075%	0.0011%	0.0010%	0.0008%	0.0002%
19	-0.0162%	-0.0098%	0.0019%	0.0018%	0.0015%	0.0007%

20	-0.0162%	-0.0098%	0.0020%	0.0018%	0.0015%	0.0007%
21	-0.0164%	-0.0099%	0.0020%	0.0018%	0.0015%	0.0006%
22	-0.0203%	-0.0124%	0.0030%	0.0027%	0.0024%	0.0012%
23	-0.0204%	-0.0125%	0.0030%	0.0027%	0.0024%	0.0012%
24	-0.0204%	-0.0125%	0.0029%	0.0027%	0.0023%	0.0012%
25	-0.0248%	-0.0152%	0.0043%	0.0040%	0.0036%	0.0021%
26	-0.0249%	-0.0152%	0.0043%	0.0040%	0.0036%	0.0021%
27	-0.0250%	-0.0153%	0.0043%	0.0040%	0.0036%	0.0020%

Table H-4: Percent difference of output standard deviations

	Sample Data Sets Used to Construct DeltaP and SigmaP					
Test Case #	1 - 3	1 - 6	1 - 9	1 - 12	1 - 15	1 - 18
1	47.60%	7.97%	3.77%	3.61%	3.99%	3.62%
2	13.93%	0.36%	1.73%	1.65%	1.64%	1.50%
3	3.89%	-2.22%	0.50%	0.43%	0.33%	0.24%
4	108.34%	9.03%	-1.65%	-0.82%	-0.86%	-0.72%
5	45.44%	-0.18%	-0.93%	-0.68%	-0.79%	-0.93%
6	21.51%	-2.15%	0.34%	0.42%	0.28%	0.05%
7	155.96%	13.97%	-1.77%	-0.22%	-0.56%	0.00%
8	92.87%	6.94%	2.26%	2.95%	2.71%	2.74%
9	47.55%	-0.73%	-0.02%	0.28%	0.08%	-0.13%
10	186.98%	16.52%	-2.75%	-0.67%	-1.23%	-0.35%
11	122.38%	8.61%	0.78%	1.86%	1.48%	1.70%
12	69.54%	0.64%	-0.35%	0.15%	-0.11%	-0.26%
13	209.06%	18.41%	-3.36%	-0.89%	-1.63%	-0.51%
14	151.56%	11.15%	-0.17%	1.31%	0.80%	1.25%
15	95.57%	3.07%	-0.42%	0.36%	0.01%	0.00%
16	218.35%	17.30%	-5.77%	-3.07%	-3.92%	-2.65%
17	168.76%	10.19%	-3.99%	-2.18%	-2.81%	-2.15%
18	122.22%	5.74%	-0.70%	0.38%	-0.06%	0.10%
19	239.67%	21.53%	-3.60%	-0.59%	-1.57%	-0.13%
20	190.52%	14.73%	-1.42%	0.65%	-0.08%	0.73%
21	134.25%	5.90%	-1.95%	-0.69%	-1.20%	-0.94%
22	244.46%	20.42%	-5.54%	-2.39%	-3.43%	-1.89%
23	200.47%	13.73%	-4.34%	-2.04%	-2.86%	-1.89%
24	156.30%	7.75%	-3.12%	-1.54%	-2.15%	-1.69%
25	250.64%	20.61%	-6.01%	-2.74%	-3.84%	-2.23%
26	214.48%	15.34%	-4.62%	-2.08%	-2.98%	-1.87%
27	170.84%	9.39%	-3.28%	-1.50%	-2.18%	-1.59%

The percent differences in the output means in Table H-3 show that the closed-form equations do an adequate job of computing the output mean, even when only three sets of data are used to construct the equations. The percent differences do decrease as the number of sets used to construct the equations increase. The percent differences in the output standard deviations in Table H-4 tell a slightly different story. The second column shows the percent differences between the "true values" and the calculated values from equations constructed using only three sets of data. The closed-form equations are not adequate and produce values very different from the "true values," i.e., the percent differences are very large. When nine sets of data are used (column four), the closed-form equations do a better job and produce acceptable results over the entire range of inputs, i.e., the input values outside of the range of values used to construct the equation still produce acceptable results. For example, column four represents equations constructed with process capability data spanning values from 32 to 34, and yet these equations can compute the output standard deviation when the input mean is 40 within six percent (test cases #25, #26, and #27). This implies that if enough data are used to construct the equations, then values outside the original range of values can still be used as inputs to the closed-form equations. Table H-4 does show that using more data sets to construct the equations does produce equations that produce better results, i.e., the columns toward the right in Table H-4 have smaller percent differences. However, the benefit of using more and more data to construct the equations does diminish. For example the difference between using three sets and nine sets (column 2 vs. column 4) is great; the improvement when using eighteen sets instead of fifteen sets is small.

Another question is "Can the equations be improved with fewer sets of data that are selected to encompass a broad range of possible inputs?," i.e., what happens if sets 1 – 3 and 25 – 27 in Table H-1 are used to construct the closed-form equations? The results are surprisingly good and are summarized in Table H-5, which shows the percentage differences between mean and standard deviation values computed using the closed-form equations versus the "true values" in Table H-2. Selecting data sets that span a larger range to use in the construction of the closed-form equations is beneficial. The equations based on data sets 1 – 3 and 25 – 27 produce excellent results even though only six data sets are used (see columns four and five in Table H-5).

Table H-5: Percent difference of output values

	Sample Data Sets Used to Construct DeltaP and SigmaP					
	1 – 3 & 13 - 15		1 – 3 & 25 -27		13 – 15 & 25 - 27	
Test Case #	Mean % Difference	St Dev % Difference	Mean % Difference	St Dev % Difference	Mean % Difference	St Dev % Difference
1	0.0004%	1.36%	0.0001%	2.78%	0.0006%	-5.58%
2	0.0004%	1.06%	0.0002%	1.02%	0.0006%	-0.57%
3	0.0004%	0.22%	0.0002%	-0.15%	0.0006%	-0.22%
4	0.0002%	-2.56%	-0.0001%	-0.51%	0.0005%	-6.79%
5	0.0002%	-1.62%	-0.0001%	-1.15%	0.0005%	-3.05%

6	0.0002%	-0.31%	-0.0001%	-0.31%	0.0005%	-0.91%
7	0.0001%	-1.53%	-0.0003%	1.00%	0.0003%	-3.72%
8	0.0001%	1.87%	-0.0003%	2.91%	0.0003%	0.77%
9	0.0001%	-0.68%	-0.0003%	-0.32%	0.0003%	-1.26%
10	0.0000%	-1.59%	-0.0005%	1.22%	0.0001%	-2.13%
11	0.0000%	0.85%	-0.0005%	2.27%	0.0001%	0.44%
12	0.0000%	-0.87%	-0.0005%	-0.25%	0.0001%	-1.20%
13	-0.0001%	-1.51%	-0.0006%	1.49%	-0.0001%	-0.76%
14	0.0000%	0.44%	-0.0006%	2.23%	0.0000%	0.77%
15	-0.0001%	-0.65%	-0.0006%	0.27%	-0.0001%	-0.62%
16	-0.0001%	-3.44%	-0.0006%	-0.38%	-0.0002%	-1.74%
17	-0.0001%	-2.85%	-0.0006%	-0.79%	-0.0002%	-1.80%
18	0.0000%	-0.55%	-0.0006%	0.70%	-0.0002%	-0.06%
19	0.0002%	-0.79%	-0.0004%	2.46%	-0.0002%	1.76%
20	0.0002%	0.09%	-0.0004%	2.35%	-0.0002%	1.75%
21	0.0002%	-1.55%	-0.0005%	-0.16%	-0.0002%	-0.66%
22	0.0005%	-2.41%	-0.0002%	0.86%	-0.0002%	0.74%
23	0.0004%	-2.43%	-0.0002%	0.00%	-0.0002%	-0.19%
24	0.0004%	-2.25%	-0.0003%	-0.56%	-0.0002%	-0.83%
25	0.0009%	-2.64%	0.0003%	0.67%	0.0000%	1.01%
26	0.0009%	-2.33%	0.0003%	0.28%	0.0000%	0.45%
27	0.0009%	-2.10%	0.0002%	-0.24%	0.0000%	-0.24%

I. Nomenclature Used in Validation Models

The nomenclature used in the physics-based math models in Section 3.6 to validate the theory of DeltaP and SigmaP is presented in this Appendix.

Table I-1: Nomenclature used in the turning operation equation derivation.

$\text{Power} = \text{Power [W]}$

$\text{MRR} = \text{Material Removal Rate } [\text{mm}^3/\text{min}]$

u_t = Total Specific Energy of Workpiece $[\text{W} * \text{s/mm}^3]$

$\text{Torque} = F_c * D_{avg}/2$, $[\text{N} * \text{m}]$

D_{avg} = Average Diameter of Workpiece [mm]

d_d = Desired Depth of Cut [mm]

f = Feed [mm/rev]

N = Rotational speed of the workpiece [rpm]

ω = Angular Velocity [radians/min]

F_c = Cutting Force [N]

I = Moment of Inertia $[\text{mm}^4]$

E = Modulus of Elasticity [MPa]

δ_w = Deflection of Workpiece Tip [mm]

δ_t = Deflection of Cutting Tool Tip [mm]

D_o = Original Outer Diameter of Workpiece [mm]

L_w = Length of Workpiece [mm]

L_t = Length of Cutting Tool [mm]

w = Width of Cutting Tool [mm]

h = Height of Cutting Tool [mm]

D_f = Desired Final Outer Diameter of Workpiece [mm]

D_a = Actual Final Outer Diameter of Workpiece [mm]

d_a = Actual Depth of Cut [mm]

Table I-2: Nomenclature used in the rolling operation equation derivation.

q = Distribute force on rolls [N/mm]

w = Width of the workpiece [mm]

L = Roll-strip contact length [mm]

Y_{avg} = Average true stress of the workpiece in the roll gap [MPa]

R = Roll radius [mm]

h_0 = Initial thickness of the workpiece [mm]

h_f = Final thickness of the workpiece [mm]

K = Strength coefficient of the workpiece material [MPa]

ε = True strain of the workpiece material in the roll gap

δ_{roll} = Deflection of a roll [mm]

E = Modulus of elasticity [MPa]

I = Moment of inertia [mm^4]

Table I-3: Nomenclature used in the rolling operation equation derivation.

R_i = Die radius [mm]

R_f = Radius of workpiece bend after springback [mm]

α = Angle of bend

T = Thickness of the workpiece [mm]

Y = Yield stress of the workpiece [MPa]

E = Modulus of elasticity of the workpiece [MPa]

L = Length from center of die to end of workpiece [mm]

X = Dimension of interest on bent sheet [mm]

J. Data for example problem

This appendix list the sample process capability data used to construct the DeltaP and SigmaP functions for the example problem presented in Chapter 5. These data were generated using Monte Carlo simulation and the physics-based equations presented in Section 3.6.1.

Table J-1: Process capability data used in example problem

Process	Incoming Mean [mm]	Incoming Std Dev [mm]	Outgoing Mean [mm]	Outgoing Std Dev [mm]	Outgoing Target [mm]	Cost per Piece [$]
Roughing 1	24.995	0.349	20.640	0.084	20.5	0.21
Roughing 1	22.003	0.139	20.522	0.013	20.5	0.21
Roughing 1	26.995	0.665	20.752	0.152	20.5	0.21
Roughing 1	23.989	0.226	20.593	0.054	20.5	0.21
Roughing 1	26.512	0.439	20.721	0.129	20.5	0.21
Roughing 2	24.996	0.251	20.726	0.131	20.5	0.15
Roughing 2	24.992	0.386	20.726	0.135	20.5	0.15
Roughing 2	27.016	0.721	20.921	0.252	20.5	0.15
Roughing 2	24.295	0.364	20.671	0.103	20.5	0.15
Roughing 2	26.488	0.460	20.865	0.215	20.5	0.15
Roughing 3	24.952	0.213	20.665	0.096	20.5	0.11
Roughing 3	25.003	0.133	20.668	0.097	20.5	0.11
Roughing 3	26.994	0.695	20.804	0.183	20.5	0.11
Roughing 3	23.010	0.407	20.564	0.042	20.5	0.11
Roughing 3	23.995	0.404	20.612	0.069	20.5	0.11
Finishing 1	20.481	0.044	20.003	0.002	20.0	0.45
Finishing 1	20.599	0.077	20.005	0.003	20.0	0.45
Finishing 1	20.299	0.040	20.001	0.001	20.0	0.45
Finishing 1	21.003	0.110	20.012	0.007	20.0	0.45
Finishing 1	20.497	0.051	20.003	0.002	20.0	0.45
Finishing 2	20.201	0.052	20.001	0.001	20.0	0.27
Finishing 2	20.699	0.094	20.009	0.006	20.0	0.27
Finishing 2	20.799	0.050	20.011	0.007	20.0	0.27
Finishing 2	21.003	0.120	20.017	0.011	20.0	0.27
Finishing 2	20.496	0.081	20.005	0.003	20.0	0.27
Finishing 3	20.451	0.061	20.003	0.002	20.0	0.20
Finishing 3	20.698	0.090	20.007	0.004	20.0	0.20
Finishing 3	20.301	0.043	20.001	0.001	20.0	0.20
Finishing 3	20.499	0.050	20.004	0.002	20.0	0.20
Finishing 3	20.998	0.119	20.013	0.008	20.0	0.20
Drilling 1	0.000	0.000	21.000	0.005	21.0	0.47
Drilling 2	0.000	0.000	21.000	0.006	21.0	0.13
Drilling 3	0.000	0.000	21.000	0.007	21.0	0.33

K. Derivation of Mixture Distribution Equations

The equations for the mean and standard deviation of a mixture distribution presented in Chapter 0 are derived here.

The derivation begins with three groups of numbers.

$$\begin{aligned} &x_1, x_2, x_3, \ldots, x_{N_x} \\ &y_1, y_2, y_3, \ldots, y_{N_y} \\ &z_1, z_2, z_3, \ldots, z_{N_z} \end{aligned} \qquad \text{K-1}$$

The mixture distribution will be denoted by W.

$$w_i = x_1, x_2, x_3, \ldots, x_{N_x}, y_1, y_2, y_3, \ldots, y_{N_y}, z_1, z_2, z_3, \ldots, z_{N_z} \qquad \text{K-2}$$

The means for the original three groups are

$$\begin{aligned} \bar{x} &= \frac{1}{N_x}\sum_{i=1}^{N_x} x_i \\ \bar{y} &= \frac{1}{N_y}\sum_{i=1}^{Ny} y_i \\ \bar{z} &= \frac{1}{N_z}\sum_{i=1}^{N_z} z_i \end{aligned} \qquad \text{K-3}$$

Now, the mean of the w's is calculated.

$$\bar{w} = \frac{x_1 + x_2 + \ldots + x_{N_x} + y_1 + y_2 + \ldots + y_{N_y} + z_1 + z_2 + \ldots + z_{N_z}}{N_x + N_y + N_z} \qquad \text{K-4}$$

$$\overline{w} = \frac{x_1 + x_2 + \ldots + x_{N_x}}{N_x + N_y + N_z} + \frac{y_1 + y_2 + \ldots + y_{N_y}}{N_x + N_y + N_z} + \frac{z_1 + z_2 + \ldots + z_{N_z}}{N_x + N_y + N_z} \qquad \text{K-5}$$

$$\begin{aligned} \overline{w} = & \frac{N_x}{N_x + N_y + N_z}\left[\frac{x_1 + x_2 + \ldots + x_{N_x}}{N_x}\right] \\ & + \frac{N_y}{N_x + N_y + N_z}\left[\frac{y_1 + y_2 + \ldots + y_{N_y}}{N_y}\right] \\ & + \frac{N_z}{N_x + N_y + N_z}\left[\frac{z_1 + z_2 + \ldots + z_{N_z}}{N_z}\right] \end{aligned} \qquad \text{K-6}$$

Recognizing the terms in parentheses are equal to the means of the individual groups, the equation above can be simplified to

$$\overline{w} = \frac{N_x}{N_x + N_y + N_z}\overline{x} + \frac{N_y}{N_x + N_y + N_z}\overline{y} + \frac{N_z}{N_x + N_y + N_z}\overline{z} \qquad \text{K-7}$$

The three initial groups make up the w population. The percent contribution is represented by the letter p and a subscript.

$$\begin{aligned} p_x &= \frac{N_x}{N_x + N_y + N_z} \\ p_y &= \frac{N_y}{N_x + N_y + N_z} \\ p_z &= \frac{N_z}{N_x + N_y + N_z} \end{aligned} \qquad \text{K-8}$$

The above two equations are combined to result in the equation for a mixture distribution.

$$\overline{w} = p_x\overline{x} + p_y\overline{y} + p_z\overline{z} \qquad \text{K-9}$$

Equation K-9 can be generalized to a combination of any number of individual distributions.

Now, the equation to calculate the variance of a mixture distribution is derived. The equations for the variances for the original three groups are shown below.

$$\sigma_x^2 = \left[\frac{1}{N_x}\sum_{i=1}^{N_x} x_i^2\right] - \bar{x}^2$$
$$\sigma_y^2 = \left[\frac{1}{N_y}\sum_{i=1}^{N_y} y_i^2\right] - \bar{y}^2 \qquad \text{K-10}$$
$$\sigma_z^2 = \left[\frac{1}{N_z}\sum_{i=1}^{N_z} z_i^2\right] - \bar{z}^2$$

The variance for the mixture distribution, w, can be written as

$$\sigma_w^2 = \frac{1}{N_x + N_y + N_z}\begin{Bmatrix} x_1^2 + x_2^2 + \ldots + x_{N_x}^2 \\ +y_1^2 + y_2^2 + \ldots + y_{N_y}^2 \\ +z_1^2 + z_2^2 + \ldots + z_{N_z}^2 \end{Bmatrix} - \bar{w}^2 \qquad \text{K-11}$$

Terms can be grouped to simplify the equation.

$$\sigma_w^2 = \frac{1}{N_x + N_y + N_z}\left\{x_1^2 + x_2^2 + \ldots + x_{N_x}^2\right\} + \frac{1}{N_x + N_y + N_z}\left\{y_1^2 + y_2^2 + \ldots + y_{N_y}^2\right\}$$
$$+ \frac{1}{N_x + N_y + N_z}\left\{z_1^2 + z_2^2 + \ldots + z_{N_z}^2\right\} - \bar{w}^2 \qquad \text{K-12}$$

$$\sigma_w^2 = \frac{N_x}{N_x + N_y + N_z}\left\{\frac{1}{N_x}\right\}\left\{x_1^2 + x_2^2 + \ldots + x_{N_x}^2\right\}$$
$$+ \frac{N_y}{N_x + N_y + N_z}\left\{\frac{1}{N_y}\right\}\left\{y_1^2 + y_2^2 + \ldots + y_{N_y}^2\right\} \qquad \text{K-13}$$
$$+ \frac{N_z}{N_x + N_y + N_z}\left\{\frac{1}{N_z}\right\}\left\{z_1^2 + z_2^2 + \ldots + z_{N_z}^2\right\} - \bar{w}^2$$

Equation K-8 is substituted into Equation K-13 and summations are used to result in a simplified form.

$$\sigma_w^2 = p_x\left\{\frac{1}{N_x}\right\}\left\{\sum_{i=1}^{N_x} x_i^2\right\} + p_y\left\{\frac{1}{N_y}\right\}\left\{\sum_{i=1}^{N_y} y_i^2\right\} + p_z\left\{\frac{1}{N_z}\right\}\left\{\sum_{i=1}^{N_z} z_i^2\right\} - \overline{w}^2 \qquad \text{K-14}$$

Now, a term is added and subtracted for each group to allow for simplification using Equation K-10.

$$\begin{aligned}\sigma_w^2 = {} & p_x\left\{\frac{1}{N_x}\right\}\left\{\sum_{i=1}^{N_x} x_i^2\right\} + \left\{p_x\overline{x}^2 - p_x\overline{x}^2\right\} \\ & + p_y\left\{\frac{1}{N_y}\right\}\left\{\sum_{i=1}^{N_y} y_i^2\right\} + \left\{p_y\overline{y}^2 - p_y\overline{y}^2\right\} \\ & + p_z\left\{\frac{1}{N_z}\right\}\left\{\sum_{i=1}^{N_z} z_i^2\right\} + \left\{p_z\overline{z}^2 - p_z\overline{z}^2\right\} \\ & - \overline{w}^2\end{aligned} \qquad \text{K-15}$$

Simply regrouping terms results in

$$\begin{aligned}\sigma_w^2 = {} & p_x\left\{\frac{1}{N_x}\sum_{i=1}^{N_x} x_i^2 - \overline{x}^2\right\} + p_x\overline{x}^2 \\ & + p_y\left\{\frac{1}{N_y}\sum_{i=1}^{N_y} y_i^2 - \overline{y}^2\right\} + p_y\overline{y}^2 \\ & + p_z\left\{\frac{1}{N_z}\sum_{i=1}^{N_z} z_i^2 - \overline{z}^2\right\} + p_z\overline{z}^2 \\ & - \overline{w}^2\end{aligned} \qquad \text{K-16}$$

The above equation can be simplified by making substitutions using Equation K-10.

$$\sigma_w^2 = p_x\sigma_x^2 + p_x\overline{x}^2 + p_y\sigma_z^2 + p_y\overline{y}^2 + p_z\sigma_z^2 + p_z\overline{z}^2 - \overline{w}^2 \qquad \text{K-17}$$

The desired form is obtained by substituting Equation K-9 in for the mean of w and regrouping.

$$\sigma_w^2 = \left(p_x\sigma_x^2 + p_y\overline{y}^2 + p_z\sigma_z^2\right) + \left(p_x\overline{x}^2 + p_y\sigma_z^2 + p_z\overline{z}^2\right) - \left(p_x\overline{x} + p_y\overline{y} + p_z\overline{z}\right)^2 \qquad \text{K-18}$$

Equation K-18 can be generalized to a combination of any number of individual distributions.

Glossary

An **Attribute Combination Descriptor** (ACD) represents a key characteristic's feature, geometry, material, and the process that created it.

ACD Matrix Set is comprised of four matrices that assist in the transformation of product information into Attribute Combination Descriptors for entry into a process capability database.

$\mathbf{C_p}$ and $\mathbf{C_{pk}}$ are process capability indexes that are commonly used to compare process variation and engineering specifications (tolerances). The equations for C_p and C_{pk} are given below where UL is the upper tolerance limit, LL is the lower tolerance limit, σ is the standard deviation, and μ is the mean.

$$C_p = \frac{UL - LL}{6\sigma}$$

$$C_{pk} = Min\left(\frac{\mu - LL}{3\sigma}, \frac{UL - \mu}{3\sigma}\right)$$

DeltaP is a mathematical representation of the actual change a dimension undergoes when operated on by a process with a given set of parameters when no process variation is present. It is a function of the desired dimensional change, M. DeltaP is a special case of Process Imparted Dimensional Change.

Desired Dimensional Change, M, is the change a process must make to a part's dimension to result in the dimension being equal to the desired (target) value.

A **key characteristic** is a physical aspect of a component or assembly that has a significant effect on fit, performance, or service life when it varies from its nominal value (Thornton 1999)

A **Process Capability Database** (PCDB) stores manufacturing data that is recorded during the production process.

Process Imparted Dimensional Change is the actual change a dimension undergoes when operated on by a process when there is no process variation present.

Process Imparted Variation is the standard deviation an under control process inherently imparts on an incoming distribution that has zero variance.

SigmaP is a mathematical representation of the standard deviation an under control process with a set of given parameters inherently imparts on an incoming distribution that has zero variance. It is a function of the desired dimensional change, M.

Bibliography

(1999). How good is the fit?, http://www.curvefit.com. **2003**.

Antony, J. and R. Coronado (2002). Design for six sigma. Manufacturing Engineer: 24 - 26.

Bloebaum, C. L., G. S. Mulubagal, et al. (1994). "Multidisciplinary design with tolerance allocation for maximum quality". *35th AIAA/ASME/ASCE/AHS/ASC Structures, Structural, Dynamics, and Materials Conference*, Hilton Head, SC.

Box, G. and N. R. Draper (1969). *Evolutionary operation: a statistical method for process improvement*. New York, John Wiley and Sons.

Boyer, D. E. and J. W. Nazemetz (1985). "Introducing statistical selective assembly -- a means of producing high precision assemblies from low precision components". *Annual International Industrial Engineering Conference*.

Chan, K. C. and R. J. Linn (1998). "A grouping method for selective assembly of parts of dissimilar distributions." *Quality Engineering* **11**(2): 221-234.

Chen, W., J. K. Allen, et al. (1997). "The Robust Concept Exploration Method for Enhancing Concurrent Systems Design." *Concurrent Engineering: Research and Applications* **5**(3): 203 - 217.

Chen, W., J. K. Allen, et al. (1996). "A Procedure for Robust Design." *ASME Journal of Mechanical Design* **118**: 478 - 485.

Clausing, D. (1988). "Taguchi Methods to Improve the Development Process". *IEEE International Conference on Communications*, Philadelphia, PA.

DeGroot, M. H. (1986). *Probability and Statistics*. Reading, Massachusetts, Addison-Wesley Publishing Company.

Ding, Y., J. Shi, et al. (2002). "Diagnosability Analysis of Multi-Station Manufacturing Processes." *Journal of Dynamic Systems, Measurement, and Control* **124**: 1 - 13.

Drake, A. W. (1967). *Fundamentals of Applied Probability Theory*. New York, NY, McGraw-Hill Book Company.

Fang, X. D. and Y. Zhang (1996). "Assuring the matchable degree in selective assembly via a predictive model based on set theory and probability method." *Journal of Manufacturing Science and Engineering* **118**: 252-258.

Frey, D., 1997, "Using product tolerances to drive manufacturing system design," Mechanical Engineering, MIT, Cambridge, MA

Grossman, S. I. (1992). *Calculus*. Fort Worth, TX, Saunders College Publishing.

Guo, R.-S. and E. Sachs (1993). "Modeling, Optimization and Control of Spatial Uniformity in Manufacturing Processes." *IEEE Transactions on Semiconductor Manufacturing* **6**(1): 41 - 57.

Han, J., M. Pratt, et al. (2000). "Manufacturing feature recognition from solid models: A status report." *IEEE Transactions on Robotics and Automation* **16**(6): 782-796.

Hanson, J. A., 2001, "Structure and Strategy for Improving Design's Access to Process Capability Databases," Master of Science, Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Harrold, D. (1999). Designing for six sigma capability, Control Engineering Online, http://www.manufacturing.net/ctl/.

Hernandez, M. J. (1997). *Database Design for Mere Mortals: A Hands-On Guide to Relational Database Design*. Boston, Addison-Wesley.

Hix, C. E. and E. B. Kittleson, 1998, "Leveraging manufacturing process capability in integrated product development," MS, Management and Mechanical Engineering, MA, Cambridge

http://www.qfdi.org QFD Institute Website, QFD Institute. **2003**.

Hu, S. J. (1997). "Stream-of-Variation Theory for Automotive Body Assembly." *CIRP Annals 1997* **46**(1): 1 - 6.

Huang, Q., S. Zhou, et al. (2002). "Diagnosis of multi-operational machining processes through variation propagation analysis." *Robotics and Computer Integrated Manufacturing* **18**: 233 - 239.

Kalpakjian, S. and S. Schmid (2001). *Manufacturing Engineering and Technology*.

Kazmer, D. O., P. Barkan, et al. (1996). "Quantifying Design and Manufacturing Robustness through Stochastic Optimization Techniques". *22nd Annual ASME Design Engineering Conference*, Irvine, CA.

Kern, D., X. Du, et al. (2003). "Forecasting Manufacturing Quality During Design Using Process Capability Data". *Submitted to the 2003 ASME International Mechanical Engineering Congress and RD&D Expo*, Washington D.C.

Krishnaswami, M. and R. W. Mayne (1994). "Optimizing tolerance allocation based on manufacturing cost and quality loss". *ASME 20th Design Automation Conference*, Minneapolis, MN.

Kwon, H.-M., K.-J. Kim, et al. (1999). "An economic selective assembly procedure for two mating components with equal variance." *Naval Research Logistics* **46**: 809-821.

Kwon, P., M. J. Chung, et al. (2000). "A formal approach to integrate design and manufacturing processes". *Design Automation Conference*, Baltimore, MD.

Lawless, J. F., R. J. MacKay, et al. (1999). "Analysis of variation transmission in manufacturing processes." *Journal of Quality Technology* **31**(2): 131 -154.

Liggett, J. V. (1993). *Dimensional Variation Management Handbook*. Englewood Cliffs, NJ, Prentice Hall.

Lucca, A. M., K. N. Berti, et al. (1995). "Statistical tolerance allocation in design utilizing historical supplier process data". *International Intersociety Electronic Packaging Conference - INTERpack'95*, New York, NY.

Mantripragada, R. and D. E. Whitney (1999). "Modeling and Controlling Variation Propagation in Mechanical Assemblies Using State Transition Models." *IEEE Transactions on Robotics and Automation* **15**(1): 124 - 140.

Mease, D., V. N. Nair, et al. (2002). "Selective assembly in manufacturing: Statistical issues and optimal binning strategies."

Montgomery, D. C. (1984). *Design and Analysis of Experiments*. New York, John Wiley and Sons.

Ota, K., 2000, "Variation of cold rolled steel manufacturing," Ph.D. thesis, Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Phadke, M. (1989). *Quality Engineering Using Robust Design*. Englewood Cliffs, New Jersey, Prentice Hall.

Powell, T. A. (2000). *Web Design: The Complete Reference*. Berkeley, CA, Osborne/McGraw-Hill.

Pugh, G. A. (1986). "Partitioning for selective assembly." *Computers and Industrial Engineering* **11**: 175-179.

Pugh, G. A. (1992). "Selective assembly with components of dissimilar variance." *Computers and Industrial Engineering* **23**(1-4): 487-491.

Raffish, N. (1985). "Group technology: an overview". *28th Annual American Production and Inventory Control Society International Conference*, Toronto, On.

Sainter, P., K. Oldham, et al. (2000). "Product knowledge management within knowledge-based engineering systems". *ASME Design Automation Conference*, Baltimore, MD.

Soderberg, R. (1994). "Robust design by tolerance allocation considering quality and manufacturing cost". *ASME 20th Design Automation Conference*, Minneapolis, MN.

Suri, R., 1999, "An integrated system model for variation reduction in manufacturing systems," Mechanical Engineering, MIT, Cambridge, MA

Suri, R. and K. Otto (1999). "System-level robustness through integrated modeling". *ASME International Conference on Design Theory and Methodology*, Las Vegas, Nevada.

Taguchi, G. (1993). *Taguchi on Robust Technology Development: Bringing Quality Engineering Upstream*. New York, ASME Press.

Tata, M. and A. C. Thornton (1999). "Process capability database usage in industry: Myth vs. reality". *ASME Design Engineering Technical Conferences*, Las Vegas, Nevada.

Thornton, A. C. (1999). "Mathematical framework for the key characteristic process." *Research in Engineering Design Theory, Applications, and Concurrent Engineering* **11**(3): 145-157.

Thornton, A. C. (1999). "Variation risk management using modeling and simulation." *Journal of Mechanical Design* **121**(2): 297-304.

Thornton, A. C. and M. Tata (2000). "Use of graphic displays of process capability data to facilitate producibility analyses." *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM* **14**(3): 181-192.

Thornton, A. C. and M. Tata (2000). "Use of graphic displays of process capability data to facilitate producibility analyses." *Ai Edam-Artificial Intelligence for Engineering Design Analysis and Manufacturing* **14**(3): 181-192.

Ulrich, K. T. and S. D. Eppinger (2000). *Product design and development*. Boston, Irwin/McGraw-Hill.

Wei, Y. F., 2001, "Concurrent Design for Optimal Quality and Cycle Time," Ph.D., Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Xue, D. and Z. Dong (1997). "Coding and clustering of design and manufacturing features for concurrent design." *Computers in Industry* **34**(1): 139-153.

Zhang, F. and D. Xue (2002). "Distributed database and knowledge base modeling for concurrent design." *CAD Computer Aided Design* **34**(1): 27-40.

Zhang, Y. and X. D. Fang (1999). "Predict and assure the matchable degree in selective assembly via PCI-based tolerance." *Journal of Manufacturing Science and Engineering* **121**: 494-500.

Zhou, S., Q. Huang, et al. (2003). "State Space Modeling of Dimensional Variation Propagation in Multistage Machining Process Using Differential Motion Vectors." *IEEE Transactions on Robotics and Automation* **19**(2): 296 - 309.

5270-72