Generation and Use of a Statistical As-Built/As-Is Model
by
Daniel J. Keating

Submitted to the Department of Mechanical Engineering
in Partial Fulfillment of the Requirements for the Degrees of
Master of Science in Mechanical Engineering
and
Bachelor of Science in Mechanical Engineering

at the
Massachusetts Institute of Technology
June 1999

© 1999 Massachusetts Institute of Technology
All rights reserved

Signature of Author

Department of Mechanical Engineering
May 7, 1999

Certified by

Daniel Frey
Assistant Professor of Aeronautical and Astronautical Engineering
Thesis Supervisor

Certified by

Kevin N. Otto
Robert N. Noyce Assistant Professor of Mechanical Engineering
Thesis Reader

Accepted by

Ain A. Sonin
Chairman, Department Committee on Graduate Students
GENERATION AND USE OF A STATISTICAL AS-BUILT/AS-IS MODEL

By

Daniel J. Keating

Submitted to the Department of Mechanical Engineering on May 7, 1999 in Partial Fulfillment of the Requirements for the Degrees of Master of Science in Mechanical Engineering and Bachelor of Science in Mechanical Engineering

ABSTRACT

Using the principles of as-built/as-is engineering and decision analysis, it is possible to generate a highly accurate as-built model of a manufactured product, and an as-is model of an aged and/or used product from a relatively small amount of measurement information. Using manufacturing data from multiple product instantiations, one can determine probabilistic distributions for design parameters, establish influence relationships between variables, and identify the key characteristics which represent the entire group of instantiations in a single model. As-built/as-is engineering emphasizes the importance of knowing the actual state of a manufactured, aged, and/or used product, as opposed to the nominal design or the state of the product at an earlier time. Decision analysis is a commonly used procedure for determining the optimal course of action in the presence of uncertainty. Combining the principles of as-built/as-is engineering with decision analysis can lead to the creation of a highly accurate and relatively inexpensive numerical model, which can be used to enhance analysis, simulation, or decision making. The merits of as-built/as-is engineering are discussed, as well as the application of decision theory to this topic. The methodology for creating and verifying the as-built model provides the core of this research, and examples are given to display its effectiveness.

Thesis Supervisor: Daniel Frey
Title: Assistant Professor of Aeronautical and Astronautical Engineering

Thesis Reader: Kevin N. Otto
Robert N. Noyce Assistant Professor of Mechanical Engineering
Table Of Contents

1. Introduction ........................................................................................................... 7
  1.1. Description of Project ..................................................................................... 7
  1.2. The Need for the Project ............................................................................... 9
  1.3. Introduction to As-Built/As-Is Engineering .................................................. 10
  1.4. Introduction to Decision Analysis ................................................................ 11
  1.5. Introduction to Bayesian Inferencing ............................................................. 12
  1.6. Integration ...................................................................................................... 13
  1.7. References ..................................................................................................... 15

2. As-Built/As-Is Engineering ................................................................................. 17
  2.1. History of Product Realization ...................................................................... 17
  2.2. Imperfections of Nominal Design .................................................................. 20
    2.2.1. Tolerances ............................................................................................... 20
    2.2.2. Example of Deviations from Nominal Design ...................................... 22
  2.3. Principles of As-Built/As-Is Engineering ...................................................... 23
  2.4. Benefits of As-Built/As-Is Engineering ......................................................... 25
  2.5. Compensation Factors ................................................................................... 28
  2.6. Current Limitations of As-Built/As-Is Engineering ...................................... 30
  2.7. Summary ....................................................................................................... 32
  2.8. References ..................................................................................................... 33

3. Decision Analysis and Bayesian Inferencing ..................................................... 35
  3.1. Decision Theory ............................................................................................ 35
  3.2. Utility Theory ................................................................................................. 35
  3.3. Bayes' Theorem and Bayesian Inferencing .................................................... 38
    3.3.1. Bayes' Theorem ...................................................................................... 38
    3.3.2. Example: Converting Conditional Probabilities to Posterior Probabilities 39
  3.4. Decision Analysis .......................................................................................... 39
  3.5. The Decision Analysis Process ...................................................................... 41
    3.5.1. The Deterministic Phase ......................................................................... 42
    3.5.2. The Probabilistic Phase .......................................................................... 43
    3.5.3. The Informational Phase ......................................................................... 45
  3.6. Influence Diagrams ....................................................................................... 46
  3.7. Uses of Decision Analysis ............................................................................. 47
  3.8. Summary ....................................................................................................... 49
  3.9. References ..................................................................................................... 50

4. Generating a Statistical As-Built/As-Is Model .................................................. 53
  4.1. Overview of the Uses of a Statistical As-Built/As-Is Model ......................... 53
  4.2. Minimizing the Quantity of Data .................................................................. 54
  4.3. Key Characteristics ....................................................................................... 55
  4.4. Identifying Key Characteristics ..................................................................... 57
    4.5. Use of Decision Analysis to Identify Key Characteristics ......................... 58
      4.5.1. The Influence Diagram ......................................................................... 59
1. Introduction

1.1. Description of Project

Using the principles of as-built engineering and decision analysis, I will develop a method to generate and verify a highly accurate numerical model of the current state of a machined and/or aged mechanical product from a relatively small amount of measurement information. When the data for a large number of instantiations of the product design are considered, a single, statistical numerical model representing the entire set of products can be developed. The goal is to characterize, with as little information as possible, the entire inventory of a set of manufactured, used, and/or aged products.

The first step in this process is measuring the as-built/as-is characteristics of a sufficient number of instantiations of the product to generate results that the engineer feels are statistically significant. The number of measurements required will depend on the level of variation present and the confidence level required by the engineer/analyst. An as-built/as-is model is the term given for a numerical model that captures and characterizes the current state of a product. While this is only a one-time investment of resources, it is the most difficult and costly stage of the process. Because I am interested in achieving highly accurate results, a large sample size of the product must be measured to obtain statistically significant results. In addition, the amount of information that will need to be measured from each instance of a product can be significant. Accurate as-built/as-is measurements of products can be fairly costly. However, the measurements taken at this stage will, it is hoped, greatly reduce the number of measurements required
farther down the line. Finally, it will also produce a more accurate input for analyses and simulations, which could in turn reduce or eliminate the amount of testing required for product validation or certification.

After I have acquired the necessary data, it must then be analyzed. I will fit a probability distribution to the data for each of the variables considered. In addition, I will determine correlations and dependencies between variables at this stage. After I have gathered this information, I can create an influence diagram, showing the relationships between the various model parameters.

Finally, I will perform decision analyses to determine both the vital parameters necessary to characterize the product in an as-is model and the relationship between the measured product parameters and the important model performance parameters. As stated previously, measurement of as-built/as-is systems can be prohibitively time consuming and/or expensive. Therefore, the fewer measured product parameters necessary to accurately reflect the product's performance parameters, the more feasible and useful as-built engineering can be in the engineering environment.

I will explain in detail the methodologies for generating the probability distributions and influence diagrams, determining the vital model parameters, and relating the model parameters to product performance parameters, as well as the underlying theories of as-built engineering and decision analysis. It is believed that this process, when used correctly in a product realization environment, can greatly improve the quality and performance of a final product. At the same time, the numerical models used in analyses and simulations will be more realistic, greatly enhancing a decision-maker’s ability to rely on those numerical results.
1.2 The Need for the Project

In cases where extreme fidelity is required in knowing the dimensions and/or performance parameters of a manufactured, used, and/or aged product, it is possible that the nominal design parameters are not going to provide sufficient information to characterize the product. Deviations from the nominal design are inevitable and may affect a product's performance. In other situations, while nominal data will provide acceptable results (e.g., interchangability of parts), as-built data can lead to optimal results (e.g., vacuum fits, ideal clearances/interferences, etc.).

Unfortunately, taking large numbers of measurements on every single part in an inventory is often not a reasonable solution. Taking accurate and useful measurements can be expensive, time consuming, and inconvenient. Therefore, it is necessary to create a method whereby the benefits of as-built engineering can be feasibly achieved. If the number (and thus, overall cost) of the measurements required to characterize a part can be reduced, a large step will be taken towards improving the attractiveness of as-built engineering.

By expending the effort to measure and characterize the as-built properties of an initial sample of a product, it is hoped that the relationships between various model parameters, as well as the relationships between model parameters and product parameters, can be determined. In addition, probability distributions can be generated, and correlations and/or causalities can be identified. Finally, the key characteristics (i.e., the subset of the measured parameters that can sufficiently characterize a product) will be identified, reducing the number of measurements that must be taken in the long run. I will then utilize this information to create a statistical as-built model, which represents
the entire inventory of a product, and to infer the overall as-built state of a specific instance of a product from a relatively sparse set of measurement information.

1.3 Introduction to As-Built/As-Is Engineering

Since the beginning of the industrial revolution, an important goal during product realization has been mass production and, to achieve this, part interchangeability. The overall objective has been to produce acceptable piece parts and assemblies as quickly and as inexpensively as possible. As a result, engineers to this day strive (through methods such as tolerancing) to ensure that normal variations that occur during product realization do not make the product unusable or unacceptable [Hindle and Lubar, 1986]. However, while product realization capabilities have progressed tremendously during the industrial revolution, this underlying tenet of engineering design has remained constant.

Dolin and Hefele [1996] were the first to lay down the principles of as-built/as-is engineering, although aspects of the philosophy were already in use in the practice of selective assembly. However, even during selective assembly, the objective is merely to make acceptable assemblies in order to minimize waste and facilitate production. Using as-built engineering and compensation factors, one can create optimal products, not merely acceptable ones. As-built engineering acknowledges the fact that products are going to deviate from their nominal design, and states that only by capturing and characterizing these deviations can the actual state of a part or assembly be determined with a high degree of accuracy.

One of the primary applications of as-built engineering, the use of compensation factors during product realization, is a variation of selective assembly. However,
traditional selective assembly is an extension of tolerancing, in that it is generally used to combine parts into acceptable assemblies when the production methods are unable to meet the tolerance criteria [Gutierrez et al, 1995] or to widen tolerance bands to make part production easier and less expensive [Chen, 1995][Chen 1996]. Dolin and Hefele [1997] first showed the possibilities of using compensation factors during product realization to improve the quality of the product created. By determining the deviations from the nominal design at a certain point during the product realization process, one can alter the remaining steps in order to optimize the overall product performance. Keating et al [1997] formulated an algorithm to optimize these alterations.

1.4 Introduction to Decision Analysis

Engineers, managers, and public policy makers are just a handful of those who are often forced to make important decisions in the face of uncertainty. Many factors that cannot be known with total certainty can have major effects on the outcome of a decision, from the size of a potential market for a new product, to the strategies of the competition, to forthcoming government regulations, to weather conditions on a particular day. Under these conditions, it is impossible to guarantee what the consequences of a particular decision will be. However, with the use of decision analysis, a much more informed choice can be made which accounts for the decision-maker's beliefs about the uncertainties and his/her preferences for various outcomes, as well as for the principles of logic and probability.

Decision analysis can be subdivided into three phases. These are the deterministic phase, the probabilistic phase, and the informational (post-mortem) phase
Howard, 1968]. A decision-maker enters the deterministic phase with a certain amount of prior information and eventually leaves the informational phase with an optimal decision. However, decision analysis is not always a simple three step process. Several iterations through the various phases may be necessary before a well-informed decision can be made.

One of the keys to the successful application of decision analysis is realizing the difference between a good outcome and a good decision. A good outcome is one that is favorable and/or profitable to the decision-maker. A good decision is one that is made after careful and logical consideration of the uncertainties and potential outcomes involved. Because uncertainty is at the heart of all crucial decisions, it follows that a good decision may lead to a bad outcome, or vice versa. Overall, however, good decisions are more likely to produce good outcomes if the decision analysis was performed competently [Howard, 1964].

1.5 Introduction to Bayesian Inferencing

Accurately determining probability distributions for uncertain variables is essential to successful decision analysis. Often, the probabilities of interest cannot be ascertained directly, but must be inferred from other, known probabilities. An example of such a problem is determining the probabilities of the various possible causes of a known effect. In these cases, Bayesian inferencing may be used to determine the relevant probabilities.

Bayesian inferencing or analysis is based on Bayes' theorem, which states that the probability of state \( A_i \) (of \( n \) possible states) given outcome \( B \) is given by
\[ P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} \]  \hspace{1cm} (EQ 1.1)

[Drake, 1967].

Bayesian inferencing is often used to convert a priori (initial) and conditional probabilities into a posteriori (final, informed) probabilities. For example, the probabilities of various states may be known, as well as the conditional probabilities of various outcomes based on the existence of these states. Once the outcome has been observed, the probabilities for the underlying states may be inferred to more accurately reflect all of the knowledge obtained at this point.

1.6 Integration

The goal of this project is to use the principles of decision analysis and Bayesian inferencing to make as-built engineering a feasible product realization alternative. The potential advantages of as-built/as-is engineering are clear. Unfortunately, accurately capturing every single aspect of every single product created can be prohibitively expensive. It is hoped that through the use of decision analysis, the number of measurements required can be substantially reduced, thereby reducing the time and resources required to achieve as-built/as-is engineering.

I will use the principles of as-built/as-is engineering to develop a statistical as-built/as-is model. Only by characterizing a substantial number of products can one determine the as-built correlations. Reducing the cost of realizing the advantages of as-built/as-is engineering are the driving force behind this project, as it is hypothesized that
knowing the state in which the products are generated is more important and useful than knowing the state in which they are designed.

I will employ decision analysis to alleviate the cost of full-scale as-built/as-is engineering. That is, the primary goal will be to allow the benefits of as-built/as-is engineering to be realized without fully inspecting and characterizing every single instantiation of every single part produced. By determining the key characteristics of the as-built products, one can greatly reduce the overall number of measurements that need to be taken to characterize a part. I will also use decision analyses to generate the individual as-built models from the statistical as-built model using only sparse model parameter and product parameter information.

Because many of the probability distributions necessary for the decision analyses or for the generation of the as-built model may not be explicitly known, it will be necessary for us to use Bayesian inferencing to determine them. Probability distributions for feature dimensions given certain product dimensions, or vice versa, are examples of probabilities that may need to be inferred.

Combined, these techniques can be used to generate, affordably and accurately, a statistical as-built model that encompasses an entire inventory or stockpile of a particular product. In addition, one can use the methodology described to infer individual as-built/as-is models for specific instantiations of a product from a relatively small amount of measurement data. As a result, parts can be selectively assembled, repaired, or replaced in order to optimize product quality and performance. This represents a significant step forward from practicing blind part interchangeability, which leads to acceptable, but sub-optimal, products.
1.7 References


2. **As-Built/As-Is Engineering**

2.1 **History of Product Realization**

The industrial revolution has brought about drastic changes in nearly every aspect of the creation of products from their initial design. However, throughout the entire progression, the goal has remained constant: to create a product that performs as the customer expects it to at a cost that does not inhibit the sale of the product [Hindle and Lubar, 1986]. The industrial revolution was not one single sweeping change in production methods, but rather a series of major improvements that each attacked an existing problem with product variance and cost. Jaikumar [1988] traced the progress of the industrial revolution from the days of craft production of single units to numerically controlled, computer-integrated manufacturing systems. A brief synopsis of his work follows.

Early craftsmen made their products one at a time, usually from physical models (masters), but occasionally from dimensionless drawings. Products were compared to the model and refined in a continuous cycle until they were deemed acceptable. Only crude tools (e.g., hammers and chisels) and measuring devices (e.g., rulers and calipers) were available. As a result, products varied greatly (by modern standards) from one instance to the next. More accurate tools and measuring devices, as well as improved product designs, were vital to improving production economics.

These were precisely the improvements brought about by the English system of manufacture, which began in the late eighteenth and early nineteenth centuries. More accurate measuring devices were created, as well as highly accurate machine tools (such
as lathes). These advances greatly increased the ability to produce accurate products. In addition, the engineering drawing was created, and first described by Monge in his *La Géométrie Descriptive* in 1801, including dimensions and orthographic representation. However, a limitation with this system was that machinists skilled in the English system were in short supply.

The American system of manufacture, which began around the beginning of the nineteenth century, focused less on accuracy but more on the interchangeability of parts. Whereas the English system aimed to create ideal products with perfect fits by making each part to mate with other specific parts, the American system aimed to determine the maximum allowable variations that would still produce acceptable products in order to manufacture parts at a much faster rate [Hindle and Lubar, 1986]. With the introduction of tolerances, or allowable errors, the skill level required of the workers was reduced. At the time, it was not feasible to accomplish both goals, those of ideal products and mass production, simultaneously.

Worker productivity and machine efficiency were targeted by the Taylor System, which was introduced in the beginning of the twentieth century. Labor was further divided, such that workers could focus their efforts on the tasks vital to their job. For example, a machinist would not be responsible for repairing and/or replacing damaged tools, merely for using the machine to make a feature of a part. In addition, the discretion of individual workers was no longer allowed, as tasks were to be performed only in the pre-determined optimal method. Finally, management took control of work and compared the speed of each worker to a standard (also pre-determined). The Taylor System eliminated much variance caused by worker inconsistency.
Statistical process control was the next step in the industrial revolution, beginning in the 1950's. Having minimized the worker-induced variance in product realization, management now focused on machine-induced variance. Two types of machine variance, systematic and random, were identified. Systematic variances have discernible causes (e.g., tool wear) and results (e.g., consistently less material cut) and can be eliminated once identified. Random variances are considered to be intrinsic to a machine or process and cannot be controlled. The process capability index, a ratio relating tolerance to the random variance of the process, was created. Processes were monitored only by investigating variances beyond the normal random variance, improving the efficiency of manufacturing and reducing the number of rejected parts.

With the advent of numerically-controlled machines, which became economically viable in the 1970's, the quality and speed of production were drastically improved. Single machines were created that were able to perform the jobs of several older machines, thanks to their automatic tool-changing capabilities and NC programming. Electronic gauges were used to automate the statistical process control, whereby systematic errors would be automatically identified and corrected.

The final step examined by Jaikumar was computer integration, including flexible manufacturing systems and CAD/CAM integration, which were introduced by Beretta in 1987. The former optimized the rate of part production among several CNC machines, the latter drastically reduced the manpower required to design, test, and generate NC codes for products.

Each historical progression through these epochs improved the efficiency of, and decreased the variance in, the manufacturing process. However, whereas the previous
improvements aimed only to reduce variance, as-built/as-is engineering recognizes that variance is inevitable and seeks to characterize variations and use understanding of variation to improve product performance. With the availability of electronic gauges and other measuring devices, it is becoming more feasible to pursue the goal of perfection that characterized the English system of manufacture. Through the use of compensation factors, as-built/as-is engineering can allow the integration of the goal of producing optimal products into the modern manufacturing environment [Keating et al, 1997].

2.2 Imperfections of Nominal Design

Products are not created exactly as they are nominally designed. As discussed above, manufacturing engineers have been attempting for centuries to solve this problem. From the introduction of clearances and tolerances to the use of electronic gauges to locate systematic errors and trends in processes, manufacturers have aimed to produce as many acceptable parts as possible by maximizing the allowable process variances and minimizing the actual variances. No amount of automation or measurement can reduce the variances to zero, however. As a result, the practice of tolerancing still pervades all engineering product design.

2.2.1 Tolerances

An engineering tolerance is defined as "latitude given to the production shop in achieving a given dimension, i.e. the difference between upper and lower limit" [Conway, 1962]. This latitude is necessary due to the imperfections of various aspects of
manufacture. Despite the various advances in design, machining, and assembly, it remains impossible to create a product exactly to nominal specifications.

The use of tolerances began with the American system of manufacture, as stated previously. Interchangeability of parts was desired for various reasons, from cost of production to ease of repair and replacement [Hindle and Lubar, 1986]. To ensure that a part produced would be interchangeable with any other instance of that part, the maximum allowable variation that would still allow functionality was determined. As long as the dimensions were within this range, the part was deemed acceptable [Jaikumar, 1988].

Throughout the various developments brought about by the industrial revolution, the use of tolerances has remained constant. In nearly all cases, parts are inspected on a pass/fail basis. That is, the only matter of importance is whether its dimensions fall within the predetermined tolerances. Exactly where in the allowable range a part falls is not analyzed or recorded. The effects that any "allowable" deviations from nominal will have on product quality or performance are rarely investigated.

Occasionally, due to the nature of a product, the tolerances are too narrow for a process to reliably produce. In these cases, parts are subdivided into groups (e.g., small, nominal, and large) based on their dimensions and selectively assembled with mating parts in a corresponding subdivision [Gutierrez et al, 1995][Chen, 1996]. Even in these cases, however, the goal is simply to produce acceptable assemblies or products, and little attention is paid to optimizing the resulting assembly’s unique performance parameters.
2.2.2 Example of Deviations from Nominal Design

Figure 2.1 shows one view from a standard engineering drawing, including dimensions and tolerances. This particular part is combined with a bolt and washer (not shown), which are also to be manufactured, to create a single assembly. For various reasons, the mass of this part is of importance in the successful performance of its function. All dimensions are in millimeters, and the part is to be made of general purpose stainless steel.

![Engineering Drawing with Tolerances](image)

**Figure 2.1** Engineering Drawing with Tolerances (Dimensions in mm)

If the part is machined exactly to nominal specifications (or the mean of the lower and upper bounds if no nominal value is specified), the mass will be 8.240 grams. If, however, this part is machined such that all dimensions are at the limit that would maximize the volume of the part, the mass will be 8.712 grams, or 5.73% above the
nominal value. It is also possible to machine this part within the tolerances such that its mass is as low as 7.764 grams, or 5.77% below the nominal value.

The mass of this part, stated to be of functional importance, can range nearly a gram, or 11.50% of the nominal value. Most likely, any mass in this range will yield an acceptable part. That is, the part will function successfully, albeit sub-optimally. To reduce the variance of the mass, a design engineer would need to redesign the part with smaller tolerances, making manufacturing more difficult and more expensive.

2.3 Principles of As-Built/As-Is Engineering

The principles of as-built/as-is engineering were developed by Dolin and Hefele [1996] at the Center for Advanced Engineering Technology at the Los Alamos National Laboratory. The underlying tenet of this philosophy is that engineering analysis and decisions should be based not on the nominal, ideal design, but on the state of a part or product as it is actually produced. Somewhat similar in principle to the English system of manufacture discussed previously, as-built/as-is engineering can allow for optimization during product realization.

Before proceeding, the distinction between as-built and as-is must be clarified. Any model that depicts the actual, current state of a part or product is an as-is model. When this model is created immediately following fabrication (and prior to any other phase of the product life cycle), it is referred to as an as-built model. As manufacturing deviations were the catalyst for the birth of this philosophy, it was originally referred to as as-built engineering. However, as the benefits of analyzing the effects of storage and
use on product state have been realized, the philosophy has come to be known as as-
built/as-is engineering.

There are four primary principles of as-built/as-is engineering.

1. The various demands of mass production do not preclude customized product
   realization.

2. Nominal based product realization methods may no longer be necessary and the
   role of tolerances in engineering design and manufacturing should be reevaluated.

3. Products can and should be characterized using solid (three dimensional) models
   depicting their actual states.

4. Errors associated with manufacturing and application are inevitable and need to
   be captured as attributes of each part or assembly.

The feasibility of performing customized product realization at mass
manufacturing rates has been greatly enhanced by recent technological advances.
Improvements in measurement devices and computing capabilities allow for rapid part
inspection and analysis during the realization process. An algorithm created by Keating
et al [1997] facilitates the use of this analysis to customize the remaining production steps
to optimize product performance parameters.

Nominal based engineering was developed early in the industrial revolution to
facilitate the production of interchangeable parts [Hindle and Lubar, 1986]. Since that
time, engineers have been limited in their design possibilities to only those solutions that
are tolerant of small deviations [Edie et al, 1979][Krick, 1979]. With the incorporation of
as-built/as-is engineering and customized product realization, new design solutions will
become available to engineers.

The use of as-built and as-is solid models is a reliable way to keep track of the
states of products. A single, nominal engineering drawing depicts the ideal condition of a
product, which is most likely not the actual state of any instance of a product, and is
certainly not the actual state of all instances of a product. In addition, it is not a statistical
representation of these instantiations.

Deviations occur during the manufacturing of any product. This fact has been the
driving force behind nearly every advance in the industrial revolution [Jaikumar, 1988].
However, tolerances allow engineers and manufacturers, in most cases, to simply ignore
these deviations so long as they are within the predetermined acceptable boundaries.
Only by acknowledging and capturing these deviations in an as-built/as-is model can the
ture attributes of a part or assembly be known, allowing for performance optimization.

As-built/as-is engineering allows for customized product realization, which in
turn allows for optimization of a product's performance parameters. That is, by altering
the nominal manufacturing and/or assembly of other parts/features to compensate for the
deviations already present, a manufacturer can improve the quality and performance of
the final product. In addition, pre-existing inventories of parts can be selectively
assembled in such a manner as to yield the optimum assembly or batch of assemblies.
All of these capabilities were demonstrated by Doiin and Hefele [1997].

2.4 Benefits of As-Built/As-Is Engineering

This project is based on the principles and benefits of as-built/as-is engineering.
By minimizing the cost of generating viable as-built/as-is models, I hope to improve the
feasibility of implementing as-built/as-is engineering in a broader spectrum of
engineering environments. For this reason the benefits of using as-built/as-is engineering
will be discussed briefly.
The five primary benefits of as-built/as-is engineering were also delineated by Dolin and Hefele [1996].

1. As-built/as-is engineering can be implemented without affecting critical production activities.

2. It is robust and permits new methodologies and equipment to be incorporated into a product realization environment.

3. It permits product characterization, which in turn enables assembly optimization.

4. Through archiving and reinspection, it allows for evaluation of changes in a product.

5. It creates a highly accurate product model that can be used in analysis, simulations, and decision making.

For as-built/as-is engineering to be a feasible option, it must be possible to incorporate it without adversely affecting existing production capabilities. Due to recent technological advances, it is possible in many cases to perform customized product realization without impeding the rate or ease of manufacture. Therefore, the only tradeoffs that need to be considered are the increase in product performance and the initial overhead cost of implementing the system.

The robustness of as-built/as-is engineering allows new methodologies and equipment to be incorporated into the manufacturing infrastructure reliably and quantifiably. That is, the output from the manufacturing system can be analyzed before and after any changes are made to evaluate their effect(s). The ability to benchmark and document the effects that new processes or technologies have on production capabilities would be extremely useful in product realization.

The information stored in the as-built/as-is model enables assembly optimization. If a part has been fully characterized, and its deviations from the nominal design have
been captured, it can be mated with other parts such that the losses in performance attributes that the deviations would cause are mitigated or eliminated. In addition, customized part production allows for still greater levels of assembly optimization.

As-built/as-is engineering also allows for thorough analyses of the effects of time and/or use on parts and assemblies. A product that was inspected and characterized at the time it was created may then be reinspected at a later time. The changes and deformations that have occurred over time can thus be quantitatively measured. This information can be useful both for repairs and part replacement on aged products and for feedback into the design and manufacturing stages to prevent such deformations.

The key to as-built/as-is engineering is the model that is generated. To be useful, it must contain enough information to completely characterize the state of a part or assembly. This model can also be used for analyses (e.g. finite element analyses) and simulations, allowing for computer testing of a product before it is put into use. Again, if the ability to run these analyses and simulations are goals for a specific as-built/as-is engineering project, the models generated must contain enough information to generate accurate results.

If utilized correctly and consistently, as-built/as-is engineering offers many advantages over current, nominal based product realization processes. Assembly optimization, customized product realization, age and wear analysis, and computer analyses and simulations are all possible in an as-built/as-is engineering environment. All that remains is to reduce the theory to practice.
2.5 Compensation Factors

In mass production environments, the current manufacturing philosophy (i.e. the use of interchangeable parts) has proven to be an effective method for minimizing waste and costs [Hindle and Lubar, 1986]. However, in cases where product performance is of key importance, the approach will not always be ideal. While expense is minimized by the current process, all that is required of the product's performance parameters is acceptability. Once a product realization process has begun, parts are inspected simply on a pass/fail basis, using predetermined tolerances; no changes are made in the process between the final design and the final product. While the product will be acceptable, it will almost certainly not be optimal with respect to all of the design's performance parameters. With the use of compensation factors during the product realization process, deviations from the nominal can be measured, analyzed, and compensated for in the remaining steps of the process, thereby optimizing the performance parameters of the product.

Using the principles of as-built/as-is engineering, parts are analyzed after each step in the product realization process. After all of the deviations from nominal have been characterized, the remaining steps to be performed in the manufacturing process are analyzed and altered within the design specifications so as to optimally compensate for the deviations present in the already completed steps. In this way, the remaining parts or features are manufactured or selected based on what has actually been produced, not on what was originally designed.

An important aspect of this process is to exactly determine the performance parameters of the part or assembly, and conversely to determine which design
characteristics are not intrinsically vital but merely means to these ends. For example, the manufacturing engineer must know whether the length of a part is vital (and thus must be kept as close to nominal as possible), or whether it is simply a means by which to achieve the required moment of inertia (in which case it can be altered for this same purpose). The ability to optimize the performance of the product increases with the freedom to modify the original nominal design.

Keating et al [1997] developed an optimization algorithm for this purpose. Manufacturing problems are posed as constrained, single objective function optimizations. The constrained objective function is converted to an unconstrained objective function using the interior penalty function method. The unconstrained problem is then solved using the Boyden-Fletcher-Goldfarb-Shanno (BFGS) variation of the Davidson-Fletcher-Powell method for unconstrained optimizations. For more information on these algorithms, see Rao [1984].

An objective function with j constraints, of the form

\[
\begin{align*}
\text{Min } f(x) \\
\text{s.t.} \\
g_j(x) &\leq 0
\end{align*}
\]  

(EQ 2.1)

is converted, by the interior penalty function method, to the unconstrained problem of

\[
\begin{align*}
\text{Min } \phi(x) = f(x) + rk \sum(-1/g_j(x)) \\
\end{align*}
\]  

(EQ 2.2)

where rk is a user defined coefficient. In this manner, the objective function will have extremely high values near the constraint limits, forcing the solution inside of the constraints. The BFGS algorithm then searches the function for a minimum using the steepest descent gradient method, with the gradient defined as

\[
\nabla \Phi(x) = \nabla f(x) + rk\Sigma(\nabla g_j(x)/[g_j(x)^2])
\]  

(EQ 2.3)
The user must then locate a local minimum in the direction of steepest descent. Once this is obtained, the rk value is decreased to allow the function to approach the constraint limits if necessary, and the objective function and its gradient are calculated at this new point. This process continues until the objective function value converges within a user specified limit, or until a maximum number of iterations (also user specified) are run. More work needs to be done to develop an algorithm to optimize a multiple objective function problem.

2.6 Current Limitations of As-Built/As-Is Engineering

While in principal as-built/as-is engineering offers the potential for significant improvement in product performance parameters, limitations exist which may prevent it from being fully realized in all types of mass manufacturing environments. The number of measurements required to fully characterize a product may be prohibitively large, the initial cost of setting up the inspection and analysis machines may be too expensive, or the benefits may be marginal enough in some cases that the effort to implement as-built/as-is engineering would not be worthwhile.

A particularly complex product may have many dimensions that need to be measured. Simple inspections are fast and relatively easy; however, as the number and accuracy of measurements required increases, the time and money required can become a burden on the product realization process. Unless the number of measurements or the speed and cost of inspection can be decreased, as-built/as-is engineering will not be a feasible solution for many complex products.
The initial cost of as-built/as-is engineering may also prohibit its adoption in some cases. Inspection and analysis machines must be purchased and incorporated into the product realization process. In addition to the initial capital investment, there are the issues of limited shop floor space and line interruption to incorporate the new machines. As with any other change in production methods or machines, as-built/as-is engineering must be viewed as an investment and its costs and benefits weighed against each other. As the cost of switching to as-built/as-is engineering drops, it will become a more attractive option.

Finally, in some cases, as-built/as-is engineering will not significantly enhance the vital product performance parameters. If the process variance present in production truly has no noticeable effect on the performance of the final product, then as-built/as-is engineering would not be a worthwhile improvement.

There is little that can be done about the final case, where as-built/as-is engineering simply does not offer a significant increase in product performance. In addition, assessing the cost of purchasing and incorporating inspection and analysis machines is beyond the scope of our project. However, if the number of measurements required to completely characterize a product can be significantly reduced, a large step will be taken towards bringing about the feasibility of as-built/as-is engineering. This is one of the primary goals of the project.

There do exist a number of cases in which as-built/as-is engineering is already a viable solution. When precision products are being manufactured that must adhere to specific performance characteristics, as-built models of manufactured parts can be combined with nominal designs of unmanufactured parts to suggest modifications for
improving the overall performance. As-built/as-is models are also useful when estimating the service interval, reliability, and expected life of a product, as they show the evolution of the product during the previous time periods. Finally, as-built/as-is engineering are useful in accident scenarios, when decisions must be made quickly and optimally in order to stabilize a situation and prevent further damage [Dolin et al, 1998].

2.7 Summary

Throughout the industrial revolution, many advances have been made towards reducing the variance of the product realization process. However, this variance can never be completely eliminated. In cases where even small variances can affect the performance of a product, it is worthwhile to measure and characterize the as-built state of that product rather than merely depending on the nominal design, which most likely does not accurately depict the actual state of the product.

The principles of as-built engineering are:

1. Customized product realization is feasible in a mass manufacturing environment

2. Nominal-based engineering methods may no longer be necessary, and the role of tolerances in engineering design should be reinvestigated

3. Products can and should be characterized by solid models depicting their actual states

4. Deviations caused by manufacturing and application are inevitable and need to be captured as specific product attributes.

The use of as-built/as-is engineering also allows for the mitigation of the effects of these deviations. Customized manufacturing, from simple selective assembly to the use of proactive compensation factors discussed previously, is feasible with the
information contained in a product's as-built/as-is model. In this manner it is possible to create not merely acceptable products, but optimal ones.

There exist some barriers to the full implementation of as-built/as-is engineering in a manufacturing environment. First, the number of measurements required to characterize a product may be very large. Second, the initial cost of purchasing and incorporating the necessary inspection and analysis machine may be prohibitively expensive in some cases. Finally, the gains in product performance may not be significant in all cases. The focus of our project is to attack the first barrier, the number of measurements required.

2.8 References


3. Decision Analysis and Bayesian Inferencing

3.1 Decision Theory

Nearly all important decisions must be made in the face of some degree of uncertainty. Factors that may have either small or significant effects on the outcome of a decision may not be known for certain in advance. The eventual outcome resulting from various options cannot be known, and thus, it is impossible to determine which decision set will lead to the best outcome. However, with the use of decision analysis, the information that is known is combined with the decision maker's beliefs about the various uncertainties as well as his/her utility function to determine an optimal course of action.

The two underlying principles of decision theory are probability theory and utility theory. It is assumed that the reader is familiar with general probability theory, although Bayes' theorem and Bayesian inferencing will be discussed. Utility theory is a method for assigning the value that a particular outcome or set of potential outcomes has to a decision-maker, based on his/her value system and risk preference.

3.2 Utility Theory

In order to determine an optimal course of action in a decision problem, the values that the decision-maker associates with the various possible outcomes must be determined. While in some decisions (particularly business decisions), the outcomes will be dollar amounts, this is not the case in every decision problem. In addition, the utility a decision-maker assigns to monetary amounts is not always directly proportional to the
dollar amount. First laid down by Von Neumann and Morgenstern [1944], utility theory describes the methodology for establishing these values.

Utility theory is based on the principle of preference. That is, that given a choice of two outcomes, the decision-maker is able to state with certainty which outcome he/she prefers or that he/she is indifferent between the two outcomes. In addition, if offered a choice of a certain outcome or an outcome lottery, he/she would again be able to select the more preferable outcome (or state indifference). One key point is that utility can vary greatly from person to person, as different people can have varying preference strengths or even preference orders [Von Neumann and Morgenstern, 1944].

For example, consider three possible events A, B, and C. The decision-maker, when presented with these three alternatives, states that event A is preferred to event B and that event B is preferred to event C. The utility of these functions can then be described by

\[ u(A) > u(B) > u(C), \]  (EQ 3.1)

in which case \( u(x) \) represents the utility to the decision maker of outcome \( x \).

Now suppose that the decision-maker is offered a choice of taking outcome B for certain or for entering a lottery in which outcome A will occur with some probability \( p \) and outcome C will occur with the complementary probability \( 1-p \). There exists some probability \( p \) where he/she will express indifference. That is,

\[ p[u(A)] + (1 - p)[u(C)] = u(B). \]  (EQ 3.2)

Equation 3.2 can be expanded and adapted to accommodate multiple or even continuous possibilities.
A decision-maker's utility function is very significant in determining his/her risk attitude. For example, consider a person who has ten dollars in his/her pocket. This person is offered a chance to flip a fair coin to double his/her money. That is, if the coin comes up heads, he/she will have twenty dollars, but if the coin comes up tails, he/she will have no money.

If this person believes that the utility difference between zero and ten dollars is greater than the utility difference between ten and twenty dollars, he/she will refuse the bet, and can be said to be risk-averse for the ten dollar bet. However, if this person believes that the utility difference between ten and twenty dollars is greater than the utility difference between zero and ten dollars, he/she will take the bet, and can be said to be risk-preferring for the ten dollar bet. Finally, if this person believes the utility of ten dollars is constant regardless of wealth (i.e., his/her utility is proportional to wealth), he/she will be indifferent to accepting the bet and can be said to be risk-neutral for the ten dollar bet.

Determining the utilities of the various possible outcomes is a vital step in solving a decision analysis problem. However, it is among the most difficult steps. Most people cannot explicitly state their utility functions. Many people are risk-preferring for losses and risk-averse for gains even without knowledge of the break-even point [Howard, 1980]. It is the job of the decision analyst to elicit this information in interviews with the decision-maker by offering alternatives and lottery outcomes, as explained above. In addition, utility functions can change over time, due to internal changes in a person or external influences. Also, short term utility and long term utility are often in opposition
to one another, and both must be accounted for [Kreile, 1977]. Only by knowing the utility of every possible outcome can a fully informed decision analysis be performed.

3.3 Bayes' Theorem and Bayesian Inferencing

Determining accurate probability distributions is also vital to the successful performance of decision analysis. Occasionally, the probabilities of interest are known or can be directly discerned by analyzing available data and fitting a histogram. In other cases, however, the important probabilities cannot be directly determined. In these cases, the probabilities of interest must be inferred from other, known probabilities. This is the purpose of Bayesian inferencing—to provide a method whereby unknown probability distributions can be calculated from other, known probabilities [Antelman, 1997].

3.3.1 Bayes' Theorem

Equation 1.1, shown again here, is the numerical representation of Bayes' Theorem.

\[
P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} \quad \text{(EQ 1.1)}
\]

[Drake, 1967].

Again, it is used to combine initial, or a priori, probabilities with conditional probabilities to generate the final, or a posteriori, probabilities, which more accurately reflect the possible states of a variable given all known information.
3.3.2 Example: Converting Conditional Probabilities to Posterior Probabilities

Suppose there are four possible (mutually exclusive) states that a car can be in. There is a 35% chance that the car has low quality gas in the tank, a 25% chance that the air pressure in the tires is low, a 15% chance that the car has a bad or dirty air filter, and a 25% chance that the car is fine. If the car has low quality gas, there is a 40% chance that it will get poor gas mileage. Low tire pressure has a 50% chance of reducing gas mileage, and the air filter 25%. If there is nothing wrong with the car, there is a 4% chance that the gas mileage will be poor.

Now suppose the result (gas mileage) is observed. With this new information, it is possible to infer more accurate probabilities for the state of the car. If the car is getting poor gas mileage, Bayes' Theorem can be used to show that there is now a 44.8% chance that the car has poor gas in the tank, a 40% chance that the air pressure in the tires is low, an 11.3% chance that the air filter is dirty or bad, and a 3.2% chance that there is nothing wrong with the car. However, if the gas mileage is fine, there is a 30.5% chance of poor gas, a 18.2% chance of low tire pressure, a 16.4% chance of bad air filter, and a 34.9% chance that nothing is wrong. The probability distribution for the possible states of the car change significantly in the face of this new information (gas mileage), suggesting that the initial probability distribution would not be the most accurate means by which to determine the state of the car.

3.4 Principles of Decision Analysis

The nine assertions of decision analysis laid down by Howard [1964] represent a very concise description of its basic tenets.
1. Decision-making is involved in most technical, business, and government problems.

2. Nearly all non-trivial decisions have a certain amount of inherent uncertainty.

3. Probability theory is the only formal way to quantify this uncertainty.

4. Most probabilities are subjective, based on an expert's opinion, and not based on physical objects.

5. All prior knowledge and experience is necessary in accurately assessing probabilities.

6. In addition to probabilities, a value system must be established.

7. The criteria for choosing alternatives must be established in advance (e.g., maximize expected utility, maximize utility of worst-case scenario, maximize probability of best-case scenario, etc.).

8. The long-term implications of the present decision cannot be neglected.

9. Good decisions must be distinguished from good outcomes.

   In addition, when practicing decision analysis, analyst bias must be carefully avoided. In most cases, the decision analyst is not an expert in the field being analyzed, only in the practice of decision analysis. In these cases, the decision analyst's purpose is to elicit information from the decision-maker, formulate the problem, and present a solution. Always, but specifically when the decision analyst is not an expert in the field, he/she must not deliberately or inadvertently inject his/her beliefs and/or values in to the decision analysis problem at any stage [Matheson, 1969].

   There have been various criticisms of decision analysis. Some contend that it is inferior to using intuition and experience when making a decision. It has also been stated that decision analysis leads to distorted framing of the problem [Howard, 1980].

   Due to human nature, many people make decisions simply by opting for blind adherence, blind change, or procrastination. Even those who carefully think out their
decisions, using intuition and experience, can improve their performance by incorporating their knowledge into a formal decision analysis. Shaping the problem and establishing of the end-value system are entirely up to the decision-maker. The decision analyst simply encodes the information given to him into the analysis [Howard, 1980].

3.5 The Decision Analysis Process

There are three main phases to solving a decision analysis problem. The first phase is the deterministic phase, in which the problem, options, variables, and outcome utilities are defined. The second phase is the probabilistic phase, in which the various uncertainties are encoded as probability distributions and an optimal decision is determined. The final phase is the informational phase, in which the decision analyst determines whether the decision should be carried out, or whether additional information should be obtained to reduce uncertainty. If more information is obtained, the problem reverts to the deterministic phase and is recalculated [Howard, 1968].

Prior to entering the deterministic phase, a decision-maker should have a certain amount of a priori information. This can include, but is not limited to, his/her options, the various possible outcomes, conditional probabilities, etc. However, he/she may not have all of the necessary information to make an informed decision. In addition, he/she may or may not be able to quantitatively express the information that is known. It is the job of the decision analyst to formulate a useful problem from the known and unknown information.
3.5.1 The Deterministic Phase

The decision problem is defined and bounded during the deterministic phase. The overall scope of the decision being considered must be known exactly so as to ensure both that no vital information is omitted and that no irrelevant information is included. With the problem bounded, all decision points and alternatives must be identified. Often, a decision-maker may have overlooked a decision point (by assuming it to be trivial, for example) or several alternatives that for various reasons were not realized as feasible options. During this stage, the utility function of the decision-maker must also be determined. The variables that may influence the outcome of the decision are identified, as well as any correlations and/or causalities that may be present. When the problem has been well defined, the alternatives are examined to locate (and remove) dominated alternatives (i.e., alternatives which always result in an outcome inferior to another alternative). Finally, sensitivity analyses are performed to determine which variables are crucial to the decision at hand and which are more or less irrelevant [Howard, 1968].

A popular example of decision analysis is the "party problem" [Howard, 1989]. The host has to decide whether to hold the party indoors, outdoors, or on a partially covered porch. However, he does not know what the weather will be like at the time of the party. During the deterministic phase, it is determined that there is one decision point, the location of the party, with three alternatives, indoor (I), outdoor (O), and porch (P). In addition, there is one uncertain variable which will influence the outcome of the decision, the weather, with two possibilities, sunny (S) and rainy (R). Table 3.1 shows the host's utility for each potential combination. There are no dominated alternatives, and the variable (weather) is clearly relevant.
<table>
<thead>
<tr>
<th></th>
<th>Sunny</th>
<th>Rainy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indoor</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Pr.ch</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Outdoor</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 3.1** Host Utility for Party Problem. Values in boxes represent utility for the host for each weather/location combination.

### 3.5.2 The Probabilistic Phase

Uncertainty is present in nearly all vital decisions. However, for the decision analysis to effectively account for these uncertainties, they must be encoded in a quantifiable manner. Most importantly, the crucial variables (identified in the deterministic phase) must be encoded as accurately as possible. Among the most common mistakes in decision analysis are neglecting crucial variables and underestimating the amount of uncertainty present in their values [Menke, 1979][Abt et al 1979]. Some probabilities will be subjective, in that they are based on an expert's beliefs, not on actual physical objects. In the context of this project, an example of a subjective probability would be the probability that a certain variance in a certain feature would result in product failure. In these cases, different experts may give different probability distributions for the same variables. It is up to the decision-maker in these cases to reconcile the differences based on his/her beliefs [Stäti von Holstein, 1984].

After the probability distributions for the underlying variables have been determined, the same methods can be used to generate profit lotteries for the outcomes of the various alternatives the decision-maker has. If a certain action will result in a definite
outcome, the utility of the alternative is simply the utility of the outcome. However, if a
certain action may result in multiple (or a continuous range of) outcomes, the decision-
maker's risk attitude must be taken into account in order to determine the utility of the
alternative. At this point, the optimal alternative (based on the pre-established criteria) is
chosen.

Returning to the party problem, suppose that the host hears a weather forecast,
which he believes to be credible, stating that there is a 60% chance of rain at the time of
the party. At this point, a decision tree, like the one shown in Figure 3.1, can be created.
The expected utility of the various options is: Indoor, 3.8; Porch, 4.4; and Outdoor, 4.0.
Therefore, if the host wishes to optimize his expected utility, under the current
conditions, the Porch would be the optimal choice of party locations.

![Decision Tree for Party Problem](image)

**Figure 3.1** Decision Tree for Party Problem
3.5.3 The Informational Phase

The final stage of a decision analysis problem is the informational phase, also called the post-mortem phase. During this stage, the value of obtaining additional information to reduce the uncertainty in one or more of the key variables is determined. The maximum that a decision-maker should be willing to pay for additional information is the increase in expected utility that would result from having perfect (completely certain) information. If information is available at a justifiable cost, the information should be obtained and the process is reverted back to the deterministic phase. If no information can be obtained, or obtaining it is too expensive to justify, the decision reached in the probabilistic phase is carried out.

For example, in the party problem, suppose a machine existed which could perfectly predict what the weather was going to be like at the time of the party. To determine how much the host should be willing to pay for this machine, we consider the possible results. Assuming that the original weather forecast was correct, there is a 60% chance that the machine will predict rain. In this case, the host would choose to have the party indoors, maximizing his utility at six. There is also a 40% chance that the machine will predict sun, in which case the host would choose to have the party outdoors, maximizing his utility at ten. The expected utility in this case is thus 7.6, and the host should be willing to pay the monetary equivalent of up to 3.2 units of utility for the use of this machine. In addition, if the machine was not perfect, but accurate 90% of the time, the analysis could be repeated to determine the expected utility using the machine and determining the new value to the host.
3.6 Influence Diagrams

One of the most powerful tools in performing a decision analysis is the influence diagram. By visualizing the variables of interest and their correlations, a decision-maker can be certain that the problem is structured correctly (based on all available knowledge) and that no vital variables and/or influences have been omitted. As stated previously, omissions can have devastating effects on the accuracy of a decision analysis [Menke, 1979].

Simply put, influence diagrams are graphical representations of decision problems. They consist of nodes representing the crucial variables and arcs that represent the interrelationships between them [Agogino et al, 1988]. There are three types of nodes possible in an influence diagram, as shown in Figure 3.2. Decision nodes, including variables which the decision maker has direct control over, are placed in rectangles. Variables that are uncertain (and not under the control of the decision-maker) are placed in circles. Finally, there is a single diamond shaped node which contains the final value or utility obtained from the problem.

![Diagram](image)

**Figure 3.2:** Influence Diagram Node Types

The presence (or absence) and direction of arcs between the various nodes of an influence diagram convey the modeler's state of knowledge about the problem. For example, an arc from a variable node to a control node indicates that the value of the variable is known at the time of that particular decision and that the value is relevant to the decision being made. An arc from a control node to a variable node indicates that the
decision may have some bearing on the value of the variable. A lack of an arc between these two nodes indicates definitively that the value is not known at the time of the decision (or that its value is irrelevant to the decision), and that the decision has absolutely no effect on the variable's value. In addition, a logical (static) influence diagram cannot contain any cycles, in that variables cannot (even indirectly) influence themselves [Agogino et al, 1988]. Influence diagram representations of the party problem from Section 3.3 are shown in Figure 3.3. Figure 3.3a is the original formulation of the problem, where the host has no knowledge of the weather at the time of the decision. The influence diagram shows that the value of the variable (weather) is neither known at the time of the decision (location) nor affected by the decision. Figure 3.3b shows the problem in the case in which the host knows what the weather will be.

![Influence Diagram Representation of the Party Problem](image)

(a) Location

(b) Location

**Figure 3.3:** Influence Diagram Representation of the Party Problem a) with weather unknown and b) with weather known

### 3.7 Uses of Decision Analysis

Decision analysis can be applied to any situation in which decisions must be made in the face of uncertainty. Public policy decisions, business strategies, and engineering problems can all benefit from the use of a formal decision analysis. By forcing a
decision-maker to consciously consider the variables involved as well as his/her utility function, an optimal course of action can be determined.

Business and economic problems were the original motivation for the formulation of formal game (utility) theory, although its applications had been realized in gambling for some time prior [Von Neumann and Morgenstern, 1944]. Since then, decision analysis has been applied to numerous business decisions. Egger and Menke [1984] detailed its use in determining investment strategies for a large company. Menke [1979] and Matheson [1983] went so far as to apply decision analysis to the entire strategic plan for a company. Business managers have found that decision analysis is a vital tool for both minute and immense problems.

Many engineering problems can be aided by decision analysis, as well. Agogino et al [1988] used decision analysis to monitor and control automated milling machines. It has also been employed in determining the optimal scheduling of outages in power plants for maintenance [Agogino, 1997]. Decision analysis can also be used to evaluate various manufacturing techniques (analyzing accuracy vs. cost), to determine allowable errors (tolerances) which will not result in product failure, to perform a risk/benefit analysis during a product development cycle prior to production, as well as to solve various other engineering problems.

Even public policy problems, which could have a great impact on an entire population, can be facilitated by a formal decision analysis. Howard et al [1972] put forth an example of the use of using decision analysis to determine an optimal course of action in the face of a potential natural disaster. Matheson [1969] applied decision analysis to the planning of a space program. While the utility functions in these cases
may be more complex and difficult to calculate, the benefits of decision analysis are still present.

We seek to apply decision analysis to determine an optimal set of key characteristics that can characterize a product in an as-built model. From this information we hope to generate a single statistical as-built model which represents an entire inventory of a product and from that, generate sufficient individual as-built models from relatively few measurements. By determining the vital measurements (and thus, eliminating the superfluous ones), we can greatly reduce the cost of implementing as-built engineering.

3.8 Summary

From engineering to business to public policy, crucial decisions often must be made in the face of uncertainty. While some rely on intuition and experience to solve these problems, decision analysis offers a truly analytical method whereby the relevant uncertainties, as well as the values of the decision-maker, are quantitatively incorporated into the decision process to determine an optimal course of action.

The nine assertions of decision analysis, as laid down by Howard [1964] are:

1. Decision-making is involved in most technical, business, and government problems
2. Nearly all non-trivial decisions contain some inherent uncertainty
3. Probability theory is the only way to quantify this uncertainty
4. Most probabilities are subjective (based on an expert's opinion), not objective
5. All prior knowledge and experience is necessary in assessing probabilities
6. A value system is needed in addition to a probability system
7. Decision criteria must be established in advance

8. Long-term implications of decisions cannot be neglected

9. Good decisions must be distinguished from good outcomes

There are three stages to decision analysis. During the deterministic phase, the scope of the problem is defined, all decision points and alternatives are identified, the utility function of the decision-maker is determined, and all important variables (as well as their correlations) are identified. During the probabilistic phase, the probability distributions for the variables and the outcomes are established. An optimal decision, based on the established probabilities, the decision-maker's utility function and risk attitude, and the pre-established criteria, is determined. Finally, in the informational phase, the value of reducing uncertainty by gathering additional information is determined. If the information is gathered, the problem is re-evaluated. Otherwise, the decision reached in the probabilistic phase is carried out.

We will now apply formal decision analysis to as-built engineering. By determining the vital parameters that characterize a product, the number of measurements necessary to perform as-built engineering will be greatly reduced. A single statistical model representing an entire inventory of a product will be generated. In addition, decision analysis will be used to generate individual as-built models from the statistical model using relatively few measured data points.

3.9 References


4. Generating a Statistical As-Built/As-Is Model

4.1 Overview of the Uses of a Statistical As-Built/As-Is Model

Before discussing the procedure by which a statistical as-built/as-is model may be generated, the uses of such a model should be explained. The general purpose of creating as-built/as-is models was discussed in detail in Chapter 2. The generation of a single statistical model that represents the current state of an entire inventory of parts or products has many of the same advantages. Trends during the manufacturing process can be identified (e.g., certain features being manufactured too large or small), the reliability of the manufacturing process can be investigated, and the creation of individual as-built/as-is models in the future is facilitated.

The statistical as-built/as-is model allows for visualization of variances in the product realization process. While current inspection techniques and statistical process control can provide information on features as they are created, the statistical as-built/as-is model can integrate all of this information to give an overall picture of what is happening during the product realization process. In addition, it can generate as-is information for the product performance parameters by assimilating all of the information into a complete model.

Another application of the statistical as-built/as-is model is assisting in the generation of individual as-built/as-is models for specific instantiations of a part or product. If the overall statistical information about an inventory is known, a relatively small number of measurements will be required to generate a model of a specific product. In this way the benefits of as-built engineering can be realized without the prohibitive
cost of taking every possible inspection on every product instantiation, which would otherwise be necessary in order to fully characterize a part or product in a model. The uses of the statistical as-built/as-is model will be discussed in more detail in Chapter 5.

4.2 Minimizing the Quantity of Data

To limit the cost of generating the statistical as-built/as-is model (and the later individual as-built/as-is models), the number of measurements required to capture and characterize the state of a part or assembly must be minimized. For any product, there are an infinite number of measurements that may be taken. Obviously, some of these will be vital in characterizing the product, while others will be redundant and/or superfluous. By determining the simplest (not necessarily smallest) set of measurements that can characterize the product, an engineer can greatly reduce the time and effort required to implement the as-built/as-is approach.

Another reason to minimize the amount of necessary information is that many methods that are used to control or measure variation can become bogged down and rendered useless if the amount of data they are required to process becomes too great. Variation analysis (a simulation determining how variations in part features affect product dimensions), for example, can generate too much information to be useful if is performed on a data set that is too broad. That is, if too many features are investigated using a variation analysis, the data that is output may be too complex and time consuming to analyze to be of any use to the decision maker [Thorton 3, 1996]. Cost-loss functions (measurments of the cost and/or loss associated with feature variation) can also be performed more efficiently on a limited data set.
In the following sections the use of key characteristics will be discussed, as well as methods for identifying these key characteristics. Key characteristics are those features that can have the greatest effect on overall product performance and/or customer satisfaction, and must therefore be closely monitored [Thornton, 1996]. Influence diagrams and decision analysis will be used to generate an optimal set of product key characteristics that will define the statistical as-built/as-is model.

4.3 Key Characteristics

Key characteristics are defined as "product features, manufacturing process parameters, and assembly process features that significantly affect a product's performance, function, and form" [Lee and Thornton 1, 1996]. Key characteristics (KCs) are used throughout industry; however, the specifics of their implementation are not consistent among the companies that use them [Lee and Thornton 1, 1996][Lee and Thornton 2, 1996]. KCs are generally subdivided into product, manufacturing, and assembly KCs, and the ones that pose the greatest potential threat to product performance are singled out as the vital StatKCs.

As previously stated, procedures for measuring, controlling, or reducing variation function much more efficiently if the data that must be processed is reduced. The use of KCs allows manufacturers and post-production analysts to focus their efforts on only those features that have the potential to significantly vary the overall form or function of the final product.

Although many companies have begun using KCs in their production facilities, the implementation has been inconsistent and often sub-optimal [Thornton, 1996]. Many
companies use KCs only in a reactive manner. That is, they wait until problems arise and only then identify the key characteristics that can be used to control or eliminate the problems identified. In addition, many companies assign KC status to more features than they can feasibly monitor, defeating the purpose of utilizing key characteristics. Finally, many companies lump all KCs together whether they refer to a part feature, a manufacturing process, or an assembly process, making it difficult to differentiate among the three.

Lee and Thornton [1, 1996][2, 1996] suggested that KCs be subdivided into three classes—product, manufacturing, and assembly key characteristics. Product key characteristics (PKCs) consist of the vital geometric features and material properties of a product. Manufacturing process key characteristics (MKCs) consist of machine settings, tool states, etc. Assembly process key characteristics (AKCs) consist of assembly methods that affect the product either directly or through later stages of assembly. For our research, PKCs will be the features of importance, as we are more concerned with the actual state of the product than the method by which this state was achieved.

In addition, those KCs that, due to either their level of variation or the severity of their effect on product performance parameters, add significant risk to proper product performance are labeled StatKCs. StatKCs are considered to be the highest priority. By focusing efforts on these StatKCs, a company can greatly reduce the cost and effort required to perform analysis before, during, and after production. The important tasks, therefore, are to identify those features to classify as KCs and, more importantly, StatKCs.
4.4 Identifying Key Characteristics

When done reactively, KCs identification entails finding those features which have the most significant (in terms of affecting overall product performance) variations which could be causing problems in the performance of the product. This process can be very tedious, and can lead to a very large number of KCs. Lee and Thornton [1, 1996] described a more efficient iterative process to identify KCs, which can be performed either proactively or reactively.

Beginning with the highest level (normally customer requirements and/or product functionality) the KCs are identified (e.g., total mass, moment of inertia, power output, etc.). The product is then decomposed one level (to the assembly, sub-assembly, or part level, whichever is relevant) and the KCs at this level which affect the previously determined KCs are identified. This process is continued until the KCs at the individual feature level have been identified. In this way, the high level KCs are translated to the low level ones which can be monitored during production.

During this process, a fairly large number of key characteristics may be identified. In some cases, it may be necessary to monitor all of them to accurately characterize the as-built/as-is state of a part or product. However, in other cases some of the features that were initially identified as KCs may not have a significant effect on product performance. These features are then stripped of their KC status. By identifying those KCs that are StatKCs, resources can be further focused on those features that are absolutely vital to the product performance parameters [Lee and Thornton 2, 1996]. The goal is, through analysis, to identify those KCs that can vary significantly or that can greatly affect product performance with small variations.
For the purposes of this research, the definitions of KCs and StatKCs will be altered slightly, but they will retain the same general meaning. As the goal is to generate an accurate numerical model that represents an entire inventory of parts at minimum cost, we must determine which features must be measured for the purposes of generating this model. We will first start by identifying all features that could have an effect, either directly or indirectly, on the overall performance of the product. Any feature that influences product performance will be labeled as a KC. However, to minimize the number of measurements required to generate the model, the StatKCs must be identified. For the purposes of this research, StatKCs will be defined as those characteristics that full knowledge and measurement of is absolutely vital to the generation of an accurate model. The methodology for identifying these KCs and StatKCs will be described in a later section. This slight variation from Lee and Thornton's definitions reflect the differences in our philosophies: while Lee and Thornton use KCs to minimize and control variance, this research seeks to use them to capture and characterize variance in a numerical model.

4.5 Use of Decision Analysis to Identify Key Characteristics

While the use of key characteristics has the potential to greatly improve the efficiency of the development of a statistical as-built/as-is model, the procedure for identifying and prioritizing these KCs has only been addressed in a general fashion. In this section a procedure will be outlined, using influence diagrams and other decision analysis tools, for determining the KCs that will comprise the as-built/as-is model.
4.5.1 The Influence Diagram

The first step in identifying the key characteristics is to build an influence diagram. As described in Section 3.6, influence diagrams show the relationships between the various decisions and uncertainties in a decision analysis problem. To determine the key characteristics, an influence diagram will be built in reverse, starting with the utility node, and will initially contain no decision nodes. The uncertainty nodes will represent the values of the key characteristics.

The utility will be defined to be the overall product performance. Normally, there will be several product performance parameters, which when combined, will give the product performance. In this way we will begin to capture and define design intent in the KCs. These parameters will all be incorporated as variable nodes with arrows to the utility node. Figure 4.1 shows an influence diagram (generated in Hugin) at this early stage for a baseball bat with five performance parameters deemed to be of importance.

![Influence Diagram](attachment:image.png)

**Figure 4.1** Product Performance Influenced by Many Parameters (Baseball Bat)
In this case, the parameters of interest are the mass, length, strength, and moments of inertia (the bat is assumed to be radially symmetric, thus there are only two moments of inertia considered). The mass is important because it affects both the speed with which a hitter can swing the bat and the momentum the bat will carry upon impact with the ball. The length is important as it determines both the range of pitch locations that the hitter can reach and affects the location of the “sweet spot” (place where maximum power is transferred from bat to ball). The strength of the bat is important, as it should not break when the hitter contacts a fast pitch. The moments of inertia will affect the mechanics of the hitter’s swing.

After the various product performance parameters have been selected, the factors that directly influence those parameters must be identified. For simplicity, one parameter should be considered at a time. Some features may influence more than one performance parameter. This does not pose a problem, assuming that no cycles are created in the influence diagram.

In the baseball bat example, the type of wood selected will affect both the mass and strength of the bat. Some of the previously identified parameters (e.g., mass, length) will influence others (e.g., moments of inertia). Other features (e.g., barrel width, handle width, handle length) may be added to the diagram as the analyst sees fit.

This process is repeated for each of the product performance parameters and then for the next level of decomposition. The identified features may in turn be affected by other features (or possibly by features already identified as KCs). The initial setup of the influence diagram is only complete when every independent variable (a node with no inputs) has been identified.
At this stage the influence diagram is most likely very complex and/or jumbled, and thus not very useful. The number of independent nodes is probably too great for each one to be treated as a KC. It is at this time that the decision analysis begins, with the goal of reducing the influence diagram to create a manageable tool for identifying the StatKCs.

4.5.2 Reducing the Influence Diagram

The influence diagram that was initially created contains every performance parameter, as well as every feature or dimension that may directly or indirectly affect those parameters. Assuming that the influence diagram was generated correctly, knowledge of the independent variables identified should provide enough information to completely characterize the overall product performance. However, the number of measurements, inspections, and tests that are required to obtain this knowledge may be prohibitively expensive or time consuming. At this point the analyst has defined one end of the spectrum—a model with minimum uncertainty but requiring maximum cost.

At this point a related but separate influence diagram must be created. This influence diagram will be used to determine both the number of measurements to take on each inspected part or product and the number of parts/products to inspect. The utility node in this diagram will represent the utility gained from the model generation process (over doing nothing). The two inputs to this utility node will be the cost of generating the model and the value of the model. While the cost of generating the model is a clearly definable function of the number and complexity of measurements per product and the number of products inspected, the value of the model created is a much more subjective
number and thus more difficult to quantify. It is possible that generating an accurate model will save money on defects and/or recalls, by allowing sub-optimal parts to be compensated for in later manufacturing/assembly stages. Further, it may alleviate the need for more accurate (and more expensive) machinery, software, and/or workers. The value of the model generated will also be a function of the measurements taken per product instantiation and the number of product instantiations inspected.

At this point the decisions must be made regarding what measurements to take and how many product instantiations to inspect. Each feature identified as a KC is placed into the second (decision) influence diagram, with additional influence arcs (arrows) going from each measurement to both the model cost and model value nodes. At this point a decision node must be placed in parallel with each KC variable node, and a new variable node representing the knowledge of the KC is created as the output of this decision node and the KC variable node. That is, the knowledge that the engineer/analyst/decision-maker has of the KC is dependent upon both the actual state of the KC and the decision on whether or not to measure the KC. If the measurement is taken, the distribution of the dimension will be known.

The remaining steps in the completion of the final influence diagram consist of analyzing trade-offs between accuracy and cost. This analysis again begins with the highest level of KCs; that is, the product performance parameters. If the level of variance present in a parameter is known for certain (in advance) to be small enough that its measurement is not worth the expense, that parameter can be deleted from the influence diagram. However, if the variance is unknown or is known to be significant enough to justify the cost of measurement, the analysis proceeds to the direct inputs to the
parameter in question. Only after the independent variables have been subjected to the same scrutiny can the StatKCs be identified. These analyses and trade-offs will often be subjective, and must be investigated by the engineer, analyst, or other relevant expert to determine what a "significant variance" is. Depending on the design and use of a part, the amount of variation that can be ignored as "insignificant" may have vastly different interpretations. The cost of obtaining the measurement information must be weighed against the cost (loss of model value or cost of potential failures/recalls) of ignoring the potential variation when generating the model.

When a feature or characteristic is determined to have a variance that is not significant enough to justify its measurement, that node is deleted from the influence diagram. The procedure for deleting a node is fairly simple. First, all arrows that originate from the node are removed, so that the node no longer has any outputs. Second, the node itself is removed. Next, any arrows that went into the now deleted node (inputs to this node) are deleted. Finally, any nodes that (due to the previous step) no longer have any output arrows are deleted in the same manner. When there are no nodes without outputs (excepting of course the utility node), the deletion is complete.

4.5.3 Dealing with Unknown Variances

The previous section dealt with those features whose variation was known to impact product performance too insignificantly to justify the cost of their measurement. In some cases, however, the inherent variation in a feature (or the ramifications these variations will have on higher level product characteristics) is unknown. In these cases, it will be necessary to measure these variances and/or their effects until the manufacturer or
analyst is confident that enough information has been obtained to make an informed decision about the relevance of the variance to product performance.

While feature variances are relatively easy to measure, their effects on other features or overall product performance may be difficult to quantify. Design of experiments and variation analysis are used in many such cases to measure the effect of feature variance on product performance [Thornton, 1996]. When generating an as-built/as-is model, any features or characteristics whose variances (or the effects thereof) are unknown must be treated as significant until further information is obtained.

4.5.4 Inferring Information about Key Characteristics

In some cases, it will not be feasible to inspect certain features or characteristics of interest. Reasons can include the cost of accurate measurement or the availability of necessary inspection equipment, the presence of features within sealed assemblies, and features for which no measurement tools exist, among other things. Regardless of the reason, if information about these characteristics is deemed to be important, it will be necessary to devise other methods to determine their values. Bayesian inferencing offers the capability to generate a probability distribution for these features if information about related features is available.

For example, suppose there is a feature that, for one reason or another, cannot feasibly be accurately measured. However, it is known that this feature has an effect on a (measurable) product performance parameter. The product design engineers (the relevant experts in this case) have determined that if the feature in question is manufactured too large, there are certain probabilities (conditional on the dimensions of other features
which affect the parameter) that the parameter will be large, nominal, or small. The probabilities have also been calculated for the cases where the feature is manufactured too small or to nominal specifications. By measuring the product performance parameter (and the other features which affect it), information can be obtained about the feature, using Bayes' theorem as described in Section 3.3.2. If the feature influences other parameters (or features), the analysis can be repeated to gain further information about the feature.

4.6 Generating the Statistical As-Built/As-Is Model

Through the use of measurements and, the feature variances and their effects are monitored during the generation of the statistical as-built/as-is model. Those characteristics whose variance was deemed to be significant in the overall product performance are tracked and recorded, as well as those characteristics whose variance or variance effects were unknown.

The number of measurements that must be taken (i.e., the number of part or product instantiations that must be inspected) must be determined on a case by case basis. In some cases, only a relatively small number of measurements may be needed to produce an acceptable confidence level in the statistics to characterize product performance. In other cases, the number of measurements required may be much greater. The number of features being tracked, the complexity of the final influence diagram, the level of inherent variance, and the accuracy required may all contribute to the final number of inspections required. Chi-squared analyses can be performed to determine the level of uncertainty in the statistical representation as a function of the number of
inspections taken [Drake, 1967]. The decision-maker also knows the marginal cost of inspecting a part or product, as well as the confidence level required of the model. By considering each of these factors, an optimal number of inspections can be determined.

During the generation of the statistical as-built model, each part/product measurement must be recorded for two reasons. First, because each measurement is a data point in the statistical as-built model. That is, each measurement point contributes to the distribution that will eventually characterize the product inventory. Second, in the case of uncertain influences, the data will be needed at a later time for analysis. For example, if the effect that variations in feature A have on product performance parameter B is unknown, the data will need to be analyzed after the fact to determine any influence or correlations. In addition, because many features may influence parameter B, each of these features must be monitored to ensure that any synergetic and/or compensating effects may be accurately accounted for.

The statistical model must be updated and rebuilt with each new batch of measurement information. When the manufacturer is satisfied that the confidence level obtained in the distributions is sufficient, the statistical model is complete. Again, this determination can be made with the use of decision analysis. The confidence level of the model as a function of the number of products inspected is objective, as is the marginal cost of product inspections. However, the value of a model as a function of its confidence level is subjective. There is most likely some point where the model ceases to have any value (the uncertainty is too great for the model to serve any useful purpose). There is also most likely some point where the value to the model of inspecting additional products is negligible (the model's value has been maximized). In between,
the value of the model increases as a function of the number of products inspected, as does the overall cost of the inspections. A relatively simple decision analysis will result in the optimal number of products to be inspected.

At that point, the model contains probability distributions for each StatKC as identified in the influence diagram. In addition, data should have been generated for the purposes of investigating uncertain influences. At this stage the data is analyzed to determine if any StatKCs may be eliminated.

If any KCs whose variance effects were unknown are determined to be insignificant, their nodes may be deleted from the influence diagram in the previously described manner. Likewise, any characteristics whose variance was determined to be large or have a significant impact on product performance are retained as StatKCs. In this manner the number of measurements required to completely characterize the as-built/as-is state of a product (or inventory of products) in the future has been minimized, which in turn minimizes the marginal cost of determining the as-built/as-is state of products.

4.7 Example: Surface Interpolation

To demonstrate the process of generating a statistical as-built/as-is model, a relatively simple example will be employed. Suppose a part is designed to be a perfect hemisphere with a radius of 100 mm, as shown in Figure 4.2. A number of these parts were turned on a lathe and placed in storage for a period of time. Information regarding the topography of these parts is required for manufacturing analysis. A statistical model will be generated representing the entire inventory of this part. This model must contain
sufficient information to evaluate the effects of manufacturing and storage in general and facilitate the future generation of individual as-built/as-is models.

As these parts are fictional and have never been produced, all data contained in this section is also fictional. However, the goal of this example is to illustrate the capabilities of the procedures outlined; therefore, the specifics of the data are of little import. The data will be selected to accurately reflect potential product states and in some cases for mathematical purposes.

![Diagram of a hemispherical part](image)

**Figure 4.2** Hemispherical part, as designed. Dimensions in mm.

### 4.7.1 Identification of Key Characteristics

The first step, as previously described, is to determine the parameters of importance and begin the creation of the influence diagram. These parameters will be represented by the nodes with direct input into the overall product performance, which in turn will be represented by the utility node. As shown in Figure 4.3, there are a number of mass and geometric properties that, for various reasons, are vital to product performance. These are the surface area, volume, center of gravity, and pitch and roll moments of inertia.
Figure 4.3  Vital product performance parameters for hemispherical part

Some of these parameters can be measured directly (e.g., center of gravity, moments of inertia), while others are not as simple to measure and must be inferred from the part's geometry. The part's geometry, in turn, must be inferred from a number of topography points. Determining the number and location of these topographic points is among the most vital steps in generating the statistical as-built/as-is model. An infinite number of points exist on the surface of the part. The ideal subset to be inspected will be small (to minimize cost) but contain sufficient information to generate a useful model. Figure 4.4 shows the intermediate (complex) influence diagram for this part. In this case, the lack of influence arcs (arrows) from the "geometry" node to some of the performance parameters does not imply that the part's geometry does not affect these values. It merely implies that these (independent) parameters can be measured directly and do not need to be inferred from the part's geometry.
Figure 4.4  Complex influence diagram for hemispherical part

4.7.2  Generating the Final As-Built/As-Is Model

Having generated the complex form of the influence diagram for the (nominally) hemispherical part, the final task in generating the statistical as-built/as-is model is reducing the influence diagram and identifying the StatKCs (the dimensions/properties that must be inspected to capture and characterize the state of the part). In some cases, assumptions will need to be made to simplify the inspection process. In others, decisions will need to be made in determining which measurements are absolutely vital.

In this case, we will first make the assumption that the parts are radially symmetric about the pole. As the parts were turned on a lathe and stored face down (resting on their flat side), this is probably a valid assumption to make. It also reduces the problem of generating a surface for the model from a three-dimensional problem to a
two-dimensional one. All that remains is to determine the radius as a function of the angle between the pole and the equator.

Even in this reduced form, there still exist an infinite number of potential inspection points. It is now necessary to determine how many measurements will be required to generate an accurate surface model. Three topography points are probably insufficient, while measuring 180 would probably be overly expensive and unnecessary. It is at this point that the decision analysis must occur. Quantifying the cost of generating this model as a function of the number of measurements is a fairly straightforward task. Quantifying the value of the model generated as a function of these measurements is a more difficult (and often subjective) task. The uncertainty that will remain in the model and the effects it will have on the usefulness of the model must be considered. In the end, a number $x$ will be identified which is the optimum number of surface points to be inspected. The final influence diagram is shown in Figure 4.5.

**Figure 4.5** Final influence diagram for hemispherical part
Each topographic point is identified as a StatXC, and measurements are taken for each instantiation of the part. Distributions are generated, as well as inter-variable influences. Of particular interest in this case is the relationship between the radii at adjacent inspection points on the parts, as this information may allow for better surface interpolation procedures requiring fewer data points in the future. Various pieces of information could be contained in the model. For example, it is possible that the parts are being consistently manufactured too large (radius too large all around), or that the method of storage (on the flat side) is leading to sagging, resulting in the radius being slightly decreased at the pole and increased at the equator. This, or any other information contained in the statistical as-built model, can be used to analyze the product realization process and help to generate individual models in the future.

Data will be necessary for later stages of model generation. Suppose measurement data has shown that although the parts are being manufactured as designed, the storage period is indeed causing the parts to sag. In addition, the surface is for the most part continuous, with no extreme changes in radius. That is, the radius follows a fairly smooth (albeit not necessarily linear) function of angle from base to equator. This information will be used in the generation of individual models in Chapter 5.

4.8 Summary

The procedure for generating a statistical model, which represents the as-built/as-is state of an entire inventory of a part or product, has been developed. Using key characteristics and the tools of decision analysis, the number of inspections required to fully characterize the actual state of a part or product (or an inventory thereof) can be
minimized. Data is then gathered about these key characteristics to generate the probability distributions and inter-variable relationships that define the statistical as-built/as-is model.

Key characteristics are used in manufacturing to identify the features that can significantly affect overall product performance. Typically, these features are monitored to ensure that they do not vary enough to adversely affect the product performance. If the variance can not be controlled within the required limits, the feature is labeled as a StatKC, and some aspect of the feature must be changed (e.g., its dimensions, the manufacturing process, etc.). For this research, the definitions of KCs and StatKCs have been altered slightly. Any feature that influences product performance will be labeled as a KC. However, to minimize the number of measurements required to generate the model, the StatKCs must be identified. For the purposes of this research, StatKCs will be defined as those characteristics that full knowledge and measurement of is absolutely vital to the generation of an accurate model. Decision analyses must be applied to determine which KCs need to be identified as StatKCs.

By building an influence diagram relating individual features to overall product performance parameters, the initial, complex problem is identified. A cost-benefit analysis is applied to each feature to determine whether measuring its variance would be worthwhile. Those features whose variances are too insignificant to justify the cost of their inspection are removed from KC status. The features remaining are the StatKCs. However, in the context of generating the statistical model, the StatKCs are not to be altered in any way (as they would be in the traditional KC usage), but merely monitored and recorded. KCs whose variance effects are unknown are measured in the generation
of the statistical as-built/as-is model to determine if they should retain their KC status. The following chapter will detail the uses of this statistical as-built/as-is model.

4.9 References


5. Utilizing the Statistical As-Built/As-Is Model

5.1 Introduction

The uses of a single statistical as-built/as-is model, which were discussed briefly in Section 4.1, will be expanded upon in this chapter. Realizing the advantages of as-built engineering (listed in Section 2.4) is the general objective when creating an as-built/as-is model, although the goals are slightly altered in the case of a single statistical model. It is still important that the process can be implemented unobtrusively (i.e., without disrupting the existing manufacturing process), and that the process be robust and allow for the incorporation of new processes and/or equipment. Product characterization is still the ultimate goal, although assembly optimization is not an immediate objective in this case. The evaluation of changes in the inventory of products over time is among the primary uses of the statistical model. The product model can also be used in decision making when the question being investigated pertains to the general state of an entire inventory, as opposed to a deterministic question regarding a specific product instantiation.

The statistical as-built/as-is model is very useful in post-production analysis. It allows for consistent trends in the manufacturing, use, and/or aging of a part or product to be identified and investigated. If parts are consistently manufactured sub-optimally, or if aging and/or use consistently cause deviations that adversely affect the overall product performance, changes may need to be made in either the nominal design or the manufacturing process. Further, the statistical model can locate those areas in which
precision (in the case of manufacturing) and/or consistency (in the case of storage or usage) needs to be improved.

The final application of the statistical as-built/as-is model that will be discussed is facilitating the creation of individual models, which represent specific instantiations of the part or product. The identification of the StatKCs using the process described in Chapter 4 will have minimized the number of inspections required to characterize a part or product, and the knowledge contained in the statistical model can be used to further improve the fidelity of the individual models. At the same time, this represents a closed loop process, as each individual as-built/as-is model provides additional statistical information that may be used to refine the statistical as-built/as-is model.

5.2 Post-Production Analysis

One of the primary uses of the statistical as-built/as-is model is post-production analysis of the parts or products. The as-built model contains a large amount of information about the production process and its imperfections. Likewise, the as-is models contain information about the various effects that storage and/or use of the parts or products has on their overall performance. Manufacture, storage, and use are the three events that can potentially cause a part or product's state to deviate from the nominal design. The more recent the latest statistical model is, the better chance an engineer will have to identify the underlying cause(s) of observed variances. For this reason, it is important to ensure that the statistical model is kept up to date with each possible change in the products' states.
The first model to be generated should be a statistical as-built model, which represents the state of products as they are being manufactured and assembled. At this point any observed variances or errors can be attributed solely to the manufacturing process. It then remains a matter of using the influence diagram (as described in Chapter 4) to identify the underlying cause(s) of these observations. In some cases it will be possible to rectify the situation (e.g., by adjusting machine settings or replacing worn/broken tools), and in others the problem may need to be analyzed to determine if fixing it is worthwhile (e.g., buying a better machine to reduce variance). However, if the information contained in the statistical as-built model prompts any changes in the product realization process, a new model must then be generated in order to accurately reflect the updated manufacturing environment.

Merely storing a part or product for a certain period of time is not likely to cause significant variations in its performance parameters; however, it is possible. Depending on the conditions (e.g., temperature, humidity, stacking, etc.) under which the products are stored and the material properties (e.g., yield stress, hardness, etc.), an extended storage period could result in alterations in vital product performance parameters. Generating a new statistical as-is model at this point (after storage but before usage) and comparing it to the statistical as-built model will provide information about the overall effects of long term storage on the product. It is possible that storage procedures may be altered to reduce these effects, or that the effects will be insignificant enough to ignore. Again, any change in product handling requires the generation of a new statistical as-is model for future reference.
The use of a product, whether it is under normal or abnormal conditions, can cause variations in the product's form and/or function. For example, driving on a tire for an extended period of time will cause wear and tear to the tire, even if it is driven under normal conditions. Certain abnormal conditions (racing, poor alignments, repeated skidding, etc.) may further alter the performance (traction, speed, etc.) of the tires. If all instantiations of a part or product are to be used in a similar fashion, a single statistical as-is model representing the state of the products after a period of use may be appropriate to evaluate the effects of use. If a product has multiple uses, or may be used under extremely different conditions by different users, some grouping of the products may be necessary before generating a set of statistical models. For example, one would most likely not want to lump tires from family sedans and police cruisers in the same group. By comparing the post-usage as-is model to the as-built model (if the product went directly into use from production) or the post-storage as-is model (if the product was stored for a time between production and use), the effects of use on the product's performance parameters can be quantified. This information can be used to alter the product design to improve its longevity or to design rework to restore the product to full functionality.

By continually updating the statistical model, it is possible to identify the causes of all variances and/or alterations in product performance. However, as a number of inspections must be taken for the generation of each statistical as-is model, it may not be cost efficient to attempt to generate the models too frequently. As with any decision, trade-offs must be considered. In this case, the benefits of having a more up-to-date statistical as-is model must be weighed against the cost of generating that model. When
it is generated, however, the statistical as-is model can be a very powerful tool in post-production analysis of a part or product.

These models show what is happening to an entire inventory of products over the course of the product's life cycle. This information can be used in several ways to improve overall product performance. Even at their most basic level, they are all-encompassing statistical process control results. Random and/or systematic errors in manufacturing can be identified, not only by their own deviations, but by their effect on higher level product performance parameters. However, these models not only determine if the products are being realized as designed, but if they perform as they were designed to perform. In addition, these models can point out areas for improvement, either in manufacturing processes, storage procedures, or even product design. If the model represents a part or sub-assembly that is to be mated with other parts or sub-assemblies, compensation factors can be employed in the production or selection of the mating pieces. By knowing the actual state of the products, the possibilities for optimizing their performance are greatly enhanced.

5.3 Generation of Individual As-Built/As-Is Models

In situations in which individual as-built/as-is models representing specific instantiations of a part or product are necessary or could be advantageous, the statistical model can serve as a useful starting point. The influence diagram, which has already been generated and simplified for the statistical model, will greatly reduce the amount of inspections needed to generate each individual model. In addition, the knowledge obtained during the generation of the statistical model may further simplify the task of
generating individual models. Finally, the knowledge contained in the statistical model may increase the fidelity of the individual models when incorporated with the inspection data for a specific part or product instantiation.

When generating an as-built/as-is model, the number and complexity of inspections required can have a profound impact on the overall cost. It is important to take enough measurements to fully capture and characterize the current state of the part or product; however, the cost of doing so for every instantiation of a part or product may be prohibitively high. The existence of a statistical as-built/as-is model can help minimize the marginal cost of generating individual models, as much of the work has already been done. The influence diagrams created using the procedure outlined in Chapter 4 has already identified those features whose measurements provide sufficient information to characterize the as-is state of a part or product. All of the necessary trade-offs between accuracy and cost have already been evaluated and decided upon. Only those features identified as StatKCs need to be inspected for the generation of the individual as-built/as-is models.

It is possible that the knowledge gained during the generation of the statistical as-built/as-is model can be used to further simplify the influence diagram, which in turn will simplify the generation of individual as-built/as-is models. For example, the variance in a certain feature may have been overestimated. If the statistical as-built model shows that a feature does not vary as significantly as was originally believed, it may not need to be included in the individual as-built/as-is models. Unanticipated correlations may be discovered during the creation of the statistical model. Upon investigation, the underlying causes of these correlations may be determined, and they may be included in
later statistical and individual models. However, even if underlying mechanisms to these correlations can not be adequately explained, they may still prove to be of use in later models.

The statistical model may also identify shortcomings in the influence diagram as it was initially created. If the selected StatKCs fail to fully capture and characterize the state of the part or product (i.e., there are unexplainable variations in higher level product features and parameters), additional StatKCs will need to be identified to complete the model. Features' variances may have been underestimated or their effects improperly neglected, or influences may have been overlooked. These situations must be rectified before additional models can be generated.

Finally, combining the knowledge contained in the statistical model with the inspection data for a specific part or product instantiation may further increase the fidelity of the individual as-built/as-is model. An example is surface interpolation. If the statistical model contains information about consistent trends or characteristics of the surface of a part or product, the surface generated for the individual model may be created more accurately and/or less expensively. For example, an individual part may be inspected at points other than those used in the generation of the statistical model. In this case, the points contained in the statistical model, if they are of high fidelity and low variance, can offer additional insight into the topography of the part or product, improving the accuracy of the model.

A single statistical as-built/as-is model offers a powerful starting point from which to generate individual models for specific product instantiations. Much of the background work to simplify the inspections required to create a model has already been
performed, and additional knowledge has been gained that could offer insight to the likely states of the part or product.

5.4 Information Contained in an As-Built/As-Is Model

The key question an engineer must ask him/herself when generating an as-built/as-is model is what purpose the model is going to be required to serve. This is true whether the model being generated is a statistical model representing an entire inventory of products or an individual model representing one specific product instantiation. "Capture and characterize the state of the product" is a very broad term, and may entail vastly different information for different products. What information, exactly, is the model supposed to convey? Mode shapes, feature geometry, part topography, and mass properties are only a few of the potential outputs that may be required. The model may need to contain all of the information necessary to perform an accurate finite element analysis or computer simulation of the dynamic response of the product.

Usually, the individual models will need to convey much different information than the statistical model. The statistical model will be used to answer questions and make decisions regarding the design, manufacturing, storage, and use of a product in general, while the individual model will more often be used to determine the exact current state of a specific part or product.

For example, the statistical model can show if there is a shortcoming in a certain aspect of the product design or production process that is causing features to be created either incorrectly or inconsistently. It can also be used to perform certain analyses on an inventory as a whole, determining the product performance parameters for the
statistically "average" product or the worst-case scenario given the information gathered. In general, the statistical model can determine whether the product manufacturing process is generating parts or products as expected, and whether these parts or products are going to perform as expected.

Individual as-built/as-is models will be used when the general information contained in the statistical model is insufficient to answer questions or make decisions about a specific instantiation of a part or product. For example, if it is believed that a product has deviated significantly from the state indicated by the statistical as-built/as-is model, or if greater precision than the statistical model can provide is necessary to evaluate a product, an individual as-built/as-is model will need to be generated. Individual models can be used in finite element analyses, simulations, or other numerical applications in order to test the product. Topography, dynamic response, mode shapes, and thermal properties, among others, may not be able to be determined precisely enough in a statistical model.

As-built/as-is models must be generated with their applications in mind. That is, if a model is going to be used in a FEA, it must contain sufficient information to provide input for the analysis. If it is intended to analyze the process capability for a few manufactured features, enough products must be inspected to form a valid evaluation, and resources should not be wasted taking unnecessary measurements. A model should contain all of the information necessary to perform its intended task with a minimum (preferably an absence) of superfluous data.
5.5 Example: Use of Compensation Factors

To illustrate one of the uses of individual as-built/as-is models, an example of the use of compensation factors will be employed. Consider again the part from Section 2.2.2, shown again in Figure 5.1. Suppose that it has been manufactured in such a manner that the mass is at its minimum (7.764 g). That is, the part is acceptable (because it never violates its tolerance constraints), but every possible feature is at the tolerance limit which leads to a minimum volume and therefore, mass. (Refer back to Section 2.2.2 for clarification of nominal, minimum, and maximum mass for this part). Further, suppose that the parameter of interest is the total mass of the assembly. If the previously mentioned bolt and washer are manufactured as designed, the total mass of the assembly will be 11.526 grams, still 0.476 grams below the nominal value of 12.002 grams, for a deviation of 3.97%.

![Figure 5.1 Engineering Drawing with Tolerances (Dimensions in mm)](image-url)

84
However, if the bolt and washer are manufactured to have the maximum material allowed within the prespecified tolerances, the total assembly will have a mass of 11.732 grams, which is only 0.270 grams (2.25%) below the nominal value. If the manufacturing engineer were given latitude beyond the prescribed tolerances, the error could be reduced further or possibly eliminated.

By applying the principles of as-built/as-is engineering to customize the realization of the remaining parts, the error in the key product performance parameter has been reduced by over 40%, without forcing any dimensions outside of their tolerances. In addition, if the part had been inspected at various points during its production, the deviations would have been identified as they occurred, and the remaining features of the part could have been created in such a way as to compensate for these deviations.

5.6 Example: Surface Interpolation

To illustrate the use of a statistical as-built/as-is model in the generation of an individual model, we will return to the hemispherical part from Section 4.7. Suppose that these parts are going to be used in a new product. In order to determine if the parts will function in this new capacity, a finite element analysis is necessary for each instantiation. For these analyses to be accurate enough to be worthwhile, fairly high-fidelity models of the surfaces of these parts must be generated. Using the information contained in the previously generated statistical as-is model, as well as inspection points from the specific instantiations, individual as-is models can be created. However, cost is of vital importance in this case, and the number of inspections per part must be kept to a minimum.
5.6.1 Using Knowledge from the Statistical Model

The first step in creating the individual model is determining what, if any, useful information is contained in the statistical model. In this case, it has been shown that after manufacture and storage, the hemispherical parts have deformed in a consistent manner such that the height of the pole has decreased and the radius of the base has expanded (the part has sagged). While the deformation is not perfectly uniform (i.e., it varies some from part to part), the general trend is consistent. Therefore we know that the radii will be lower near the pole and higher near the base.

In addition, it has been discovered that the radius does not jump or oscillate as a function of theta. That is, the radius tends to follow a relatively smooth (albeit non-linear) pattern from the base to the pole. This knowledge will reduce the number of inspections necessary to capture and characterize the state of each part by increasing the ability of an analyst to infer surface points between the inspected data points.

5.6.2 Surface Interpolation Code

Software has been designed by Dolin and Keating for the purpose of generating geometric models from sparse data sets [Dolin et al, 1998]. In its current form, the software generates the radius of a part as a function of the angle from the pole to the equator (base). It then generates a radially symmetric part based on this radius. The generation of the radius as a function of the angle is the vital step in creating an accurate surface model.

The software accepts two data files as input. The first input file is the previously established canonical model of the product. When creating an as-built model, for
example, the nominal design can be used as the canonical model. This as-built model can in turn be used as the canonical model when generating an as-is model at a later point in time. A statistical as-built/as-is model of a part or product inventory can be used as the canonical model when generating an individual as-built/as-is model of a specific instantiation of that part or product. This canonical data is input as a series of Cartesian coordinates, one set for each known point. In the case in which the radius of the canonical model can be described by a known function of the angle, the number of points used to define the surface can be decided by the analyst. The second input file is the known (measured) deviations from the previously defined canonical model observed in the part or product. When creating an as-built model, for example, the deviations will be the difference between the nominal design and the actual product. Likewise, the deviations in later as-is models will represent the difference between the as-built model (or the most recent as-is model) and the current state of the product. This deviation data is input as the angle (from pole to equator) and the radial deviation observed at that angle.

In the absence of additional information or knowledge, the software’s functionality is limited to fitting Wilson-Fowler splines through the known (inspected) surface points. One spline represents the canonical surface model, the other represents the individual surface model. The more points that the spline is based on, the more accurate the curve will be. However, the fewer the number of measured points that are required, the easier and less expensive the inspection process will be. In order to create an accurate, useful, and inexpensive model, the number and location of the inspected points must be optimized.
As stated in Section 5.5.1, the statistical as-is model of the inventory of the hemispherical parts contains some very useful information. The sagging of the part was shown to be consistent throughout the inventory of the parts. In addition, the relative smoothness of the radius as a function of theta can be very helpful for the purpose of reducing the required number of measurement points. At this point the decision analysis must be performed. In light of the information contained in the statistical as-is model, how many inspection points are required to create an individual model accurate enough to run the finite element analysis? The cost of inspections and the fidelity required of the individual model must both be considered when making this determination.

As a side note, the statistical as-is model can also be used to perform an initial analysis on the entire inventory of the product. By generating models of the statistically average product, as well as the extremes of the possible product states (as determined by the probability distributions contained in the statistical model), an analyst can determine if the statistically average product can be expected to function as intended, and can possibly determine which types of variance will prevent the part from functioning in its new capacity.

Due to user demands and resource availability, I have developed a graphical user interface (GUI) for use in the Microsoft® Windows 95™ environment, using FORTRAN 77 and 90. The process is explained in clear language to the user (although it is assumed that the data input files have been previously created in the proper format), as well as the meaning of the results. The output from the software is twofold: text and graphics. A text output file is created which contains information for both the canonical and as-built/as-is models. Radial (R,θ) coordinates (at a user-specified interval) are given for
each surface, as well as various mass properties (surface area, volume, mass, center of gravity, and moments of inertia) for each model. This same information is presented graphically. Plots of radius vs. angle and radial deviation vs. angle are created. A band representing the user specified tolerance is included on the latter graph to show the "acceptability" of the part. Finally, a dialog box is created which displays the canonical and individual mass property information listed above, as well as the relative error between the two.

5.6.3 Minimizing the Required Number of Measurements

Among the primary goals of this research is minimizing the number of measurements that are required to generate an accurate as-built/as-is model of a part or product. Using the surface interpolation code, the accuracy of models generated using varying numbers of measurements per part will be evaluated. By removing some of the available information from the surface interpolation input files, I will force the program to infer more of the surface.

The manufacturing data used for this analysis is generated from actual surface deviation files generated for the Los Alamos National Laboratory. Six parts were inspected at two degree intervals from the base to the pole, for a total of 46 measurements per part. The original (nominal) shape could not be used in this research, so the deviations were applied to a hemisphere of radius 100 mm, which was determined to be a useful substitute. An example of the deviation data is given in Appendix A.

For each of the six parts, a surface model was created using all 46 available measurement points. Intermediate points were then taken away, and the surfaces were
regenerated using 24, 16, 13, and 7 measurement points. For each surface created, the accuracy was measured by calculated the sum of the squared errors at each of the original 46 measurement points. The results are presented in Table 5.1.

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Part 1</th>
<th>Part 2</th>
<th>Part 3</th>
<th>Part 4</th>
<th>Part 5</th>
<th>Part 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>2.279E-03</td>
<td>2.160E-03</td>
<td>1.151E-03</td>
<td>1.110E-03</td>
<td>1.541E-03</td>
<td>9.822E-04</td>
</tr>
<tr>
<td>16</td>
<td>4.200E-04</td>
<td>2.021E-04</td>
<td>2.188E-04</td>
<td>1.264E-04</td>
<td>5.402E-04</td>
<td>1.498E-04</td>
</tr>
<tr>
<td>24</td>
<td>2.870E-04</td>
<td>8.683E-05</td>
<td>1.630E-04</td>
<td>6.431E-05</td>
<td>1.242E-04</td>
<td>6.496E-05</td>
</tr>
<tr>
<td>46</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table 5.1**  Sum of Squared Errors for the Six Parts

Figure 5.2 presents this information graphically. Figure 5.2(a) shows the error as a function of the number of measurement points (in a log-log plot) for all six parts. Figure 5.2(b) averages the errors for all six parts and presents a least-squared solution to the data. In both graphs, the surface generated from all 46 points is omitted because the sum of the squared errors for these surfaces is zero, which cannot be plotted on a log-log graph.

According to the data, for these parts the sum of the squared errors can be expressed as a function of the number of measurement points by

\[ \ln(\text{SSE}) = -2.6225 - 2.0010 \ln(M) \quad (\text{EQ 5.1}) \]

or

\[ \text{SSE} = 0.0726M^{-2.001}, \quad (\text{EQ 5.2}) \]

where SSE is the sum of the squared errors at the original 46 inspection points and M is the number of measurements taken. By generating similar equations for a part that is to be inspected, an engineer can make a quantitative decision regarding the acceptable level of error and select the minimum number of measurement points to obtain this level of accuracy.
(a) Sum of Squared Error vs. Number of Measurements for Each Part

(b) Average Sum of Squared Errors for all six parts with least squares linear fit

Figure 5.2  Sum of Squared Error as a Function of Number of Measurements
To further reduce the required number of measurements per part, the engineer can investigate ways to reduce the coefficient and/or the exponent in Equation 5.2. The coefficient corresponds to the y-intercept of the graphs in Figure 5.2, and represents the information known from a single inspection. This is improved by learning more about the product and the manufacturing process used to create it, improving the knowledge base that an engineer has prior to taking any measurements. The exponent corresponds to the slope in the graphs, and represents the marginal increase in model accuracy provided by an additional measurement. By investigating trends across production runs and/or incorporating knowledge of the manufacturing process and its possible variances, the surface interpolation can be improved, reducing the number of measurement required to acceptably limit the sum of the squared errors.

5.6.4 Future Improvements to the Software

Currently, the code is limited to fitting the radial spline through the measured points; however, work is progressing on possible inferencing procedures. That is, methods whereby additional surface points can be inferred from the measured points and the engineer’s knowledge of the product. More complex than developing the various inference engines is incorporating knowledge to determine which method(s) to use in different cases. The statistical as-built/as-is model, as previously mentioned, contains a certain amount of information that may be useful in the generation of an individual model. Additional knowledge of the product in question, whether theoretical or empirical, can also be incorporated. Finally, information about the individual part instantiation may be used to assist in the model creation.
Measured part data, beyond the surface inspection points, can be of use in the
generation of individual as-built/as-is models. Product state parameters, including mass
properties, can be used to guide the inference of the surface model. It must be noted that
correlation between calculated model parameters and measured product parameters does
not guarantee that a model is valid. Combinations of errors may serendipitously result in
correct model parameters (e.g., creating a model which is too short and too wide may
result in correct model volume). However, while the existence of these correlations does
not guarantee model validity, their absence guarantees that there is some imperfection in
the model. If the calculated model volume is different than the measured product
volume, there is obviously some inaccuracy in the model. By including product
parameter knowledge in the surface interpolation procedure, the inferencing procedure
can be selected or fine tuned to result in accurate calculated model parameters. For
example, if the initial model generated results in a low model volume, methods may be
incorporated to result in higher inferred radii. Similarly, the model can be fine tuned to
align other mass properties.

Information contained in the statistical as-built/as-is model may also be of
significant use in the generation of the individual models. Relationships between various
parameters that were not realized at the time of the initial formulation of the influence
diagram may have since been identified, providing new information to the analyst for use
in the model generation. For example, it may have been determined that no inspection
point radius differed by more than a certain amount δ from an adjacent inspected radii. If
the radius of inspection point i (R_i) is measured, the radii at adjacent locations can be
constrained to R_i +/- δ. If deviations follow a definable pattern that can be constrained by
a small number of inspection points, the cost and/or difficulty of generating the model will be greatly reduced.

Theory may also provide a certain amount of information that will be of use in the generation of the individual as-built/as-is model. For example, materials properties, in conjunction with knowledge of the conditions of manufacture and/or storage may provide insight which can help guide the selection of model parameters.

The greatest challenge remaining in the improvement of the surface interpolation software is incorporating this knowledge and using it to fine tune the generation of the model. Currently, this is achieved manually. That is, the user must observe the results of the model generation, consider the knowledge available, make adjustments to the interpolation algorithms, and repeat the process until he/she is satisfied with the results. Automating this process, such that the software would run multiple iterations until an optimal model is generated, is the most significant improvement necessary in the code.

5.7 Verifying Results

Very few methods exist for verifying and validating a numerical model used to represent a mechanical part or product. Model verification is a long learning process during which time the model predictions are compared to observations and the model is updated accordingly [Knox, 1984]. Correlation of mass properties is a step in the right direction, but in and of itself does not guarantee that a model is correct. Combinations of errors, caused by unforeseen influences, inaccurate inspections, or simple luck can result in inadvertent parameter correlation. However, this correlation is necessary for model
validity to even be a possibility. A model that does not provide the correct mass property information can not possibly be a valid model.

As the number of parameters compared between the product and the model increases, the likelihood of the model being valid also increases. However, for each comparison, a physical test or inspection must be conducted to compare to the numerical calculation or simulation results. By showing consistent accuracy for a number of parameters for a number of product instantiations, the methodology by which the models are generated may be deemed valid, thus allowing for the physical inspections and tests to be foregone in the future. With current technology, it is impossible to completely verify and validate a numerical model without a certain amount of physical testing. The amount and type of testing required depends on the information that the model will be required to provide and the precision demanded by the analyst.

5.8 Summary

The uses of statistical and individual as-built/as-is models have been explained, with examples given when necessary. In addition to the overall objectives of as-built engineering, the statistical model can be used for post-production analysis and to facilitate in the generation of individual as-built/as-is models. Software created for generating surface models of mechanical parts was described, and an example given for determining the number of measurements required to generate a surface with a specific level of accuracy. The software’s limitations were explained and recommendations for future improvements were put forth. Finally, the difficulties associated with verifying and validating numerical models were touched upon.
The statistical as-built/as-is model contains a great deal of information regarding the product realization process. As-built models highlight the design and manufacturing processes, and later as-is models focus on the effects of storage and/or use of the product. The KCs identified in the formulation of the model are analyzed to identify the imperfections that result from various phases of the product life cycle. Improvements can be made, or at the minimum deviations can be anticipated and accounted for, improving the accuracy of predictions regarding product performance. The statistical model should not be viewed solely as a collection of production statistics. While it can be used to perform statistical process control, its functionality is much greater. While SPC can determine if parts or products are being made as they are intended to be made, the statistical as-built/as-is model can determine if the product, as realized, performs in the expected or optimal manner.

The information contained in, and thus the uses of, statistical and individual as-built/as-is models is somewhat different. While the statistical model contains information regarding an entire inventory or production run, the individual model contains information only for one specific instantiation of a part or product. The statistical model should be used to evaluate the processes by which the parts/products are realized. The individual model is used to evaluate the performance of a specific part or product. For example, an analyst may use the statistical model to determine if the parts are being consistently manufactured large enough, or use the individual model to perform a finite element analysis on a specific part. However, finite element or other complex analyses can be performed on the statistical model, assuming the model was created with that
intent. This may be useful in locating previously unnoticed failure states resulting from synergistic combinations of acceptable but sub-optimal feature dimensions.

The statistical as-built/as-is model is also useful in creating these individual models. During the process of generating the statistical model, a good deal is learned about the product, its features, and the correlations and influences among these features. This knowledge can be used to improve both the fidelity and the efficiency of the individual model creation process. The software developed for this project lays the groundwork for incorporating knowledge into individual as-built/as-is models. While a fair amount of improvement needs to be made before the software reaches its full capabilities, it has already demonstrated the ability to incorporate prior knowledge into model generation.

5.9 References


6. Summary and Conclusions

6.1 Review of Underlying Theories

The theories of as-built/as-is engineering, decision analysis, and key characteristics provided the basis for this research. As-built/as-is engineering, and some of the difficulties associated with its implementation, provided the impetus for this research. Decision analysis and key characteristics helped to alleviate the problems with as-built/as-is engineering by allowing efforts to be focused on vital product features and parameters. The combination of the three results in a coherent procedure for generating accurate models reflecting the actual state of a part or product at minimum cost.

As-built/as-is engineering is the next logical step in the evolution of the product realization process. Rather than attempting to simply control and minimize variance and deviations from nominal designs, as-built/as-is engineering seeks to capture these deviations in order to characterize the actual state of a part or product. By basing analyses, simulations, and other numerical testing on the actual state of a product rather than the ideal state (which can never be exactly obtained), the value of the results obtained from those numerical tests will be much higher.

Decision analysis provides a quantifiable method for optimizing the expected output when decisions must be made in the face of uncertainty. By encoding the uncertainties in the form of probability distributions and the decision-maker's priorities in the form of utility functions, decisions can be made analytically rather than intuitively. While decision analysis does not always lead to an optimal outcome, it can help to maximize the chances of obtaining an optimal outcome.
Key characteristics are used to identify those product realization details that require extra attention to ensure that products are being manufactured correctly. Although key characteristics can be product features or parameters, manufacturing settings, or assembly processes, only product key characteristics were considered in this research. Because there may be too many features or parameters for an analyst to accurately track at a reasonable cost, it is often necessary to allocate resources where they are needed the most. By correctly identifying the key characteristics, analysts and manufacturers can focus their efforts on the areas in which there is the most room for improvement.

The combination of these three theories has resulted in the formulation of a process whereby the benefits of as-built/as-is engineering may be realized at a lower cost than previously believed. As a result, implementing as-built/as-is engineering may become feasible for a much wider segment of the product realization industry. By identifying the key characteristics of a part or product and using decision analysis to determine which of these are absolutely vital in order to capture and characterize the state of that part or product, a low-cost, high-fidelity model can be generated that represents an entire inventory of that part or product. This model may then be used for post-production analysis or to facilitate in the generation of individual as-built/as-is models.

6.2 Summary of Methods Outlined

Procedures were described for creating and using the statistical as-built/as-is model. Influence diagrams, the primary graphical tool of decision analysis, were used to identify the key characteristics necessary to formulate the statistical as-built/as-is model.
Uses of this statistical model were explained, including methods for creating individual models from the statistical model. Finally, software for generating geometric models of radially symmetric parts was created and described. Although no physical experimentation was undertaken, the outlined methods should be verified in this fashion in the future.

Influence diagrams have traditionally been used to visualize the formulation of a decision analysis problem. With nodes representing state variables, decision points, and the utility derived from the result, influence diagrams show the relationships between the various aspects of a decision to be made under uncertain conditions. By representing the overall performance of a part/product in the utility node, an influence diagram was created to identify the key characteristics of that part/product. Parameters and features that directly or indirectly affected performance were identified as key characteristics. After the initial influence diagram was completed, decisions were made to determine which nodes to retain in the final influence diagram and which nodes to delete based on their cost of inspection and the value each would add to the model generated. The final influence diagram contains all of the variables that will be used to define the statistical as-built/as-is model, and nothing more.

Once it has been created, the statistical as-built/as-is model can be a very useful tool in post-production analysis of both the product realization process and the products themselves. The statistical model contains information on every variable that was deemed to be of importance in determining the overall product performance. Further, the influence diagram has already been created, showing exactly how these variables affect product performance. Therefore, the statistical model can be used to determine not only
whether the products are being created as they were designed, but whether they can be expected to perform as intended. Imperfections in overall product performance can be traced to their root causes at the feature level, allowing for designs and/or processes to be fine tuned to yield better products. In addition, the statistical model can facilitate in the production of individual as-built/as-is models when they are deemed necessary.

In some cases, the information contained in the statistical as-built/as-is model may not be sufficient to perform the analyses or simulations required. When additional information is needed regarding a specific instantiation of a part or product, an individual model must be created. A previously generated statistical as-built/as-is model can facilitate the generation of this model in two ways. First, the statistical model contains a great deal of information about the state of all of the instantiations of the part/product. This information can be used to guide the formulation of the individual model. Second, during the creation of the statistical model, it is possible that additional information regarding the product, which was not originally reflected in the influence diagram, was learned. Possibilities include unforeseen correlations between variables, as well as the irrelevance of one or more variables that were assumed to have a significant impact on product performance. This additional knowledge can reduce the marginal cost of generating individual as-built/as-is models when they are deemed necessary. In addition, if the statistical model is deemed to be sufficient, the entire cost of generating individual models can be eliminated.

Software was developed for the purpose of generating geometric as-built/as-is models of parts from relatively sparse data sets. By minimizing the number of inspections required to capture and characterize the surface of a part, the cost of
generating the models may be reduced. Currently the software fits splines through inspected surface points. The software is still in the early stages of development, and it is currently limited to radially symmetric parts. Work is continuing with the goal of encoding methods for inferring intermediate points; however, the next major step in code revision will be incorporating knowledge of additional product parameters (e.g., mass properties) into the inference engine to guide the surface interpolation. In this way, the software will be able to choose an optimal surface model based on all of the information available to the analyst.

Through the use of these methods, the feasibility of implementing as-built/as-is engineering may be improved. While there is still a cost (monetary and/or time) associated with acquiring the inspection equipment and taking the measurements necessary to generate the statistical and, possibly, individual models, these methods will help to minimize those costs. This is a significant improvement over taking every feasible measurement on every single instantiation to generate models that only represent one instantiation each.

6.3 Recommendations for Future Work

Because this is a new field of study, this research has only laid the groundwork for the work that may be performed in the future. There are a number of improvements which need to be made to the software before it will reach full functionality. Physical tests need to be completed to further demonstrate the capabilities of these processes. Finally, methods for verifying numerical models must also be improved, as the entire
theory of as-built/as-is engineering rests on the ability to capture and characterize the state of a part or product in a numerical model.

The necessary improvements to the surface interpolation software have been mentioned previously. The incorporation of product knowledge remains as the most significant upgrade to be made to the software. Automating the inference procedure selection would provide a great increase in the utility provided by the software, but this can only be accomplished if the software has access to all of the knowledge at the analyst's disposal. The coding of additional inference engines would allow for greater flexibility in the surface interpolation process. Finally, removing the limitation to radially symmetric parts would allow the software to be applied to a much wider product base. Radial symmetry should remain as an option, but not a requirement, for surfaces generated by the software. These upgrades are currently in progress, but are not functional at the time of this report. Their completion will greatly improve the modeling capabilities available to the analyst.

While this research has described the methodology for creating and using a statistical as-built/as-is model, additional physical verification of the principles is necessary. A batch of parts should be designed and manufactured, and the outlined procedures should be performed to generate a statistical as-built model representing the entire batch. The nominal design, the as-built model, and the actual state of the products should be compared to show the improvements offered by as-built/as-is engineering.

The basic tenets of as-built/as-is engineering rest on the principle that product definition must be entrusted to a numerical model. There is no absolute method for model validation. Currently, if a model is considered valid, it simply has not been proven
to be invalid. That is, while model failure can be demonstrated, complete model success can not. As model validation procedures improve, the reliability of as-built/as-is engineering will improve with them.

There are a number of areas for further research in the field of as-built/as-is engineering. This research has built upon the work of Dolin, Hefele, and Keating at the Los Alamos National Laboratory and provided the basis for new directions in this area. Advances have been made in this relatively new field, but there is still much to be done.

6.4 Acknowledgements

This research would not have been possible without the support of many people. The author would like to recognize, first and foremost, Ronald M. Dolin, Ph.D., one of the pioneers in as-built engineering. Ron has guided this research and served as a very helpful mentor. Professor Daniel Frey of the Department of Aeronautical and Astronautical Engineering at the Massachusetts Institute of Technology served as the academic advisor for this research and offered invaluable insight and guidance. Jill Hefele, another of the founders of the theory of as-built engineering, has also been of great service. Denise DesRochers provided valuable proofreading and editing services. Finally, the author would like to thank the Engineering Analysis group of the Engineering Sciences and Applications division (ESA-EA) at the Los Alamos National Laboratory for funding this research and providing the necessary resources for its completion.
Appendix A: Example of Data Used in Surface Interpolation

Table A.1 contains the raw data used for the surface interpolations in Chapter 5. Every two (2) degrees from pole (zero degrees) to base (90 degrees), a measurement was taken. The deviation between the designed radius and the measured radius at that point was recorded. For confidentiality reasons, the designed radii could not be included, so the deviations were applied to a hemisphere of radius 100 mm, which was determined to be an acceptable substitute.

<table>
<thead>
<tr>
<th>Degree</th>
<th>Deviation (mm)</th>
<th>Degree</th>
<th>Deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0695</td>
<td>46</td>
<td>-0.0203</td>
</tr>
<tr>
<td>2</td>
<td>0.0608</td>
<td>48</td>
<td>0.0014</td>
</tr>
<tr>
<td>4</td>
<td>0.0512</td>
<td>50</td>
<td>0.0234</td>
</tr>
<tr>
<td>6</td>
<td>0.0336</td>
<td>52</td>
<td>0.0446</td>
</tr>
<tr>
<td>8</td>
<td>0.0192</td>
<td>54</td>
<td>0.0590</td>
</tr>
<tr>
<td>10</td>
<td>0.0016</td>
<td>56</td>
<td>0.0736</td>
</tr>
<tr>
<td>12</td>
<td>-0.0167</td>
<td>58</td>
<td>0.0863</td>
</tr>
<tr>
<td>14</td>
<td>-0.0340</td>
<td>60</td>
<td>0.0947</td>
</tr>
<tr>
<td>16</td>
<td>-0.0397</td>
<td>62</td>
<td>0.1016</td>
</tr>
<tr>
<td>18</td>
<td>-0.0633</td>
<td>64</td>
<td>0.1051</td>
</tr>
<tr>
<td>20</td>
<td>-0.0793</td>
<td>66</td>
<td>0.1118</td>
</tr>
<tr>
<td>22</td>
<td>-0.0879</td>
<td>68</td>
<td>0.1128</td>
</tr>
<tr>
<td>24</td>
<td>-0.0919</td>
<td>70</td>
<td>0.1133</td>
</tr>
<tr>
<td>26</td>
<td>-0.0980</td>
<td>72</td>
<td>0.1147</td>
</tr>
<tr>
<td>28</td>
<td>-0.0970</td>
<td>74</td>
<td>0.1089</td>
</tr>
<tr>
<td>30</td>
<td>-0.0852</td>
<td>76</td>
<td>0.1086</td>
</tr>
<tr>
<td>32</td>
<td>-0.0924</td>
<td>78</td>
<td>0.1020</td>
</tr>
<tr>
<td>34</td>
<td>-0.0860</td>
<td>80</td>
<td>0.0968</td>
</tr>
<tr>
<td>36</td>
<td>-0.0775</td>
<td>82</td>
<td>0.0902</td>
</tr>
<tr>
<td>38</td>
<td>-0.0696</td>
<td>84</td>
<td>0.0845</td>
</tr>
<tr>
<td>40</td>
<td>-0.0577</td>
<td>86</td>
<td>0.0755</td>
</tr>
<tr>
<td>42</td>
<td>-0.0468</td>
<td>88</td>
<td>0.0709</td>
</tr>
<tr>
<td>44</td>
<td>-0.0283</td>
<td>90</td>
<td>0.0655</td>
</tr>
</tbody>
</table>

Table A.1: Deviations from nominal for Part 1 from Chapter 5