Analytical Techniques of Quality and Cost: Robust Design, Design of Experiments, and the Prediction of Mean Shift

by

Justin Ruflin

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

Bachelor of Science

at the

Massachusetts Institute of Technology
[June 2004]
May 2004

Signature of Author..
Department of Mechanical Engineering
May 19, 2004

Certified by..
Daniel Frey
Assistant Professor of Mechanical Engineering
Thesis Supervisor

Accepted by..
Ernest G. Cravalho
Chairman, Undergraduate Thesis Committee

Analytical Techniques of Quality and Cost: Robust Design, Design of Experiments, and the Prediction of Mean Shift

by

Justin E. Ruflin

Submitted to the Department of Mechanical Engineering on May 19, 2004 in partial fulfillment of the requirements for the Degree of Bachelor of Science in Mechanical Engineering

ABSTRACT

The quality of a product to a large extent determines the success of that product in competitive markets. Measuring and improving quality is thus a primary objective of the designer. The aim of the following work is to provide an introduction to the methods of quality optimization and to illustrate these techniques through examples. Quality is first defined and quantified. The robust design method, which is a technique that focuses on improving quality without adding cost, is then described. Particular attention is paid to experiment design, which is a major factor in the effectiveness and efficiency of the robust design process. The effect of product variability on the mean performance of a product is also explained along with the various ways that can be used to predict a shift in the mean value of the performance.

Two examples are then developed. The first focuses on the application of the robust design method to illustrate the steps of the process. The second example primarily focuses on creating a comparison of the Monte Carlo, Latin Hypercube, and star pattern sampling methods on predicting mean shift. The benefits of the star pattern sampling method are apparent through the example. The error in the prediction of mean shift of the star pattern is less than 1%, and the execution time was less than one fifth the times of the Monte Carlo and Latin Hypercube methods.

Thesis Supervisor: Daniel Frey

Title: Assistant Professor of Mechanical Engineering

1.0 Introduction

The goal of a firm is to produce products that are useful, that people want to buy, that make money, and that are efficient (cause minimal loss). In many markets, the products that gain market share and make the most money are those that are the highest quality. The better a product can achieve its desired performance in spite of the variability in the conditions of its operating environment, the manufacturing process which produced it, and the conditions imposed on it by the consumer, the more robust and higher quality a product is considered to be. High quality products are also the most efficient in that they produce the least amount of loss (often in the form of a cost) to both the firm and to the consumer. As an engineer and as part of a company, the goal of the designer is to produce high quality products.

Many methods of improving quality exist, but the real goal is to improve the quality of a product without adding cost. However, system variability comes from component variability, and the general misconception of the engineer is that the only way to get rid of this variability is to add higher tolerance components. Adjusting tolerances can be a step in the design process, but it will result in higher costs, and there are other techniques that don't create higher costs. The objective in the next several sections is to describe these alternative methods, specifically focusing on robust design and the supporting techniques of design of experiments and computer simulated sampling techniques. The final sections contain two examples. The first focuses primarily on the robust design method, and the second pays more attention to the idea of mean shift and the ways that it can be quantified. The first several sections, however, define and quantify quality, and place the quality design techniques in the larger setting of the design of a product and the objectives of the firm.

2.0 Understanding Quality and its Benefits

Firms differentiate their products "by their characteristics or attributes" to gain market share (Church and Ware, 2000, p. 369). Products within a market are similar in that they perform similar functions and can thus be substituted for each other. The closer one product is to the next, the more substitutable the products, and the more they compete directly on price. In a market where all products are essentially identical and no differentiation occurs, consumers simply buy the product with the lowest price. In an undifferentiated market, if no agreement takes place between firms on what price to set (implied or illegally), then the price of the product is driven to the marginal costs of the firms, and the firms that actually sell the product because they have the lowest marginal costs make no money. By differentiating a product, a firm sets it apart from the other products in the market, making it less substitutable, which gives the firm market power and consequently the ability for increased profits and market share. It is thus in the interest of a firm to differentiate its product by any means possible.

In economic terms, products can be differentiated either horizontally or vertically. A product is in a horizontally differentiated market when consumers have heterogeneous preferences for products within that market; that is, consumers cannot agree on which products are best. The markets for breakfast cereals, toothpaste, and music are all horizontally differentiated. A product is in a vertically differentiated market when consumers have homogenous preferences. If all prices were the same in a vertically differentiated market, all consumers would purchase the same brand because consumers (all?) agree on the same quality index and thus the brand with the highest quality. Quality is defined as "the vertical attributes of a product" (Church and Ware, 2000, p. 184). "Automobile brands within a class – like subcompacts – are horizontally differentiated… but automobile classes are vertically differentiated – if the price of a Ford Escort was the same as a Saab Turbo 900, most (all?) consumers would purchase the Saab" (Church and Ware, 2000, p. 369). The noticeable quality difference between these two cars is due to the hardware that goes along with the Saab. A turbo, a larger engine, and a more aggressive suspension are all attributes that can improve a car's quality. To determine the quality between a Saab and a BMW however, to say that these two cars are only horizontally differentiated is not always true (the goal of the manufacturer is to make sure it is not true). In fact, quality that is obtained by attributes other than larger engines and moving into different car classes is a major goal of the manufacturer in the automotive market (as well as any other market).

Aside from quality obtained from the "stuff" added to a product, or the materials used, quality can also be improved by improving the product's ability to perform on target. Consider again the Saab 900 and the Ford Escort. Looking at the hardware and the class of these cars tells the consumer about price, and what he or she gets by buying the car. However, the hardware tells the consumer nothing about the durability of the car, how often it might break down, how long it will last, or if the performance will deteriorate over time. From Escort to Escort this may even be difficult. The target performance of a car is to have a life of maybe 100 thousand miles, to never break down, and to not deteriorate in performance. Any deviation from this target value results in a loss of quality. The vertical integration of these two products might completely disappear to a customer if he knew that the Saab he was buying was a "lemon" and the Ford would never break down. Customers want a car that is reliable, and reliability in the automotive industry is thus a huge issue in asserting quality.

The problem with automobiles in determining quality is that they are essentially experience goods. Referring again to economic terms, experience goods are those products that must be used after purchase in order to determine quality (Church and Ware, 2000). Companies can and do attempt to signal quality (which is why a buyer may not be worried about getting a Saab that is a lemon) by advertising, warranties, brand reputation, and any other way a firm can communicate to a consumer the quality of their product. Even with these signaling methods however, consumers never have complete information about an experience good until they have tried that product. Looking at the product from an engineering perspective, how exactly a firm conveys more information to consumers is not a concern, but how to make the product perform closer to the target value very much is the concern of the engineers. If the marketing team is trying to sell a fleet of lemons, the reputation and the brand will not last very long.

The quality of a product directly relates to how well the product performs on target. Deviation from the target results in a loss of quality which essentially results in a cost to the consumer and ultimately the manufacturer; the consumer has to replace or fix the product, and the manufacturer sells less of the product and loses money from warranties. From the viewpoint of the engineer, several methods exist to control quality. Quality is considered in both the manufacturing phase (online quality control) and the design phase (offline quality control) of a product.

3.0 Online vs. Offline Quality Control Methods

Quality control techniques applied to the manufacturing process are known as on-line quality control methods. Manufacturing control of quality is a major step in the creation of a product due to the variability involved in the manufacturing process and the quality loss that can result if the process is not in control. On-line control methods include control charts, reliability studies, cause and effect diagrams, process capability studies, and statistical process control (Logothetis and Wynn, 1989). The advantage of these techniques is that they can reduce waste in the manufacture process since low quality products often are discarded through the screening phase. They can also improve quality loss due to the manufacturing process itself; that is, variability introduced by poor manufacturing. However statistical process control and the other methods cannot compensate for poor quality of design (Logothetis and Wynn, 1989). These processes also cannot "guarantee a product robust to deterioration and variability due to uncontrollable environmental factors" that arise once the product is actually in use (Logothetis and Wynn, 1989, p. 241). Thus even with these processes optimized, a design can still produce waste and loss once it has left the factory. In order to get the quality to the desired level, sometimes the necessary on-line control strategies can also become prohibitively expensive. "Economically successful design is about ensuring high product quality while minimizing manufacturing cost" (Ulrich and Eppinger, 2000, p. 237). The alternative to online control methods is to attempt to design quality into the product before the product is introduced to environmental factors, manufacturing variability, and process control issues.

Offline quality control methods seek to improve quality before production. Doctor W.E. Deming, one of the major contributors to the development of the offline techniques, stressed that industry must "cease dependence on inspection to achieve quality… by building quality into the product in the first place" (Logothetis and Wynn, 1989, p. 3). Dr. Deming was an American who worked for 30 years in Japan on improving quality by using statistical techniques. Dr. Genichi Taguchi is one of Dr. Deming's major followers, and the application and ideas of the robust design method described in later sections is largely a result of his efforts. The techniques developed by these two individuals for improving quality while reducing costs have "been an important factor in the rapid industrial growth and subsequent domination of international markets in [the electronics, automotive products, photography] industries by Japan" (Phadke, 1989, p. 3). Japan is a good example of how reduced costs and improved quality results in increased market share and profits. Before considering the robust design technique, the next section reviews the general design process.

4.0 Overview of the Design Process

The design of a product from conception to product launch and entrance into the market is a detailed process. The process begins essentially with the creation of an idea, taken from research and technology development activities, and ends with the launch of a product. The steps of design are

0. **Planning**: The result of the planning stage is the product mission statement which specifies the target market, business goals, key assumptions, and constraints.
1. **Concept Development**: One or more concepts are chosen out of many ideas to proceed in the design process. A concept includes the "description of the form, function, and features of a product" but not the details (Ulrich and Eppinger, 2000, p. 17).
2. **System-level Design**: The product architecture is determined, and the product is broken into subsystems and components for further design.
3. **Detail Design**: The geometry, materials, and tolerances of all product features are specified. The suppliers for the individual parts are determined. The process plan is also created, and the necessary tooling identified. This stage also includes the Design for X (DOX) processes.
4. **Testing and Refinement**: The product performance is validated and improved.
5. **Production Ramp-up**: The work force is trained. The final problems are solved, and the product is launched.

(Ulrich and Eppinger, 2000, p. 17)

Quality is a topic that can be in the mind of the designers throughout the entire process. For instance, quality can be considered in concept development, which "can play an important role in reducing the sensitivity to noise factors as well as in reducing the manufacturing cost" (Phadke, 1989, p. 33). "Quality Function Deployment (QFD) and Pugh's concept selection method are two techniques that can improve the quality and productivity of the concept design step" (Phadke, 1989, p. 33). The most substantial quality improvement occurs in the detail design phase, in which the parts are defined and design for X takes place.

4.1 Design for X

By the end of system-level design, the design is no longer significantly driven by needs of the customers and product specifications. Customer needs and specifications are a huge part of the process, but by this stage, design is more focused on details of the actual parts. As a goal then in this phase, the team focuses on design for X, where "X may correspond to one of dozens of quality criteria such as reliability, robustness, serviceability, environmental impact, or manufacturability" (Ulrich and Eppinger, 2000, p. 237). As stated above, improvements in quality and cost result in vertical differentiation and cost reduction which are hugely beneficial to product and company success; thus, design for X is very beneficial.

Two major steps for quality improvement in the DOX process are design for manufacturing (DFM) and parameter design (robust design). Both have the same goal in mind, reducing cost while keeping or increasing quality, but the two methods don't always necessarily achieve the goal in the same manner. The Steps of DFM entail reducing the costs of components, the assembly, the manufacturing, and the supporting production. DFM tries to find more efficient, sometimes simply cleverer, ways of creating a product and choosing better components and manufacturing techniques. Robust design can be a step within this process. However, DFM and Robust design are not necessarily the same, since if robust design is not considered during the DFM phase, "in some cases actions to decrease manufacturing cost can have adverse effects on product quality (such as reliability or robustness)" (Ulrich and Eppinger, 2000, p. 257). For instance, reducing the cost of a component part often results in a lower quality component with increased performance variability. Adding this higher variability part to the system will increase the variability of the system performance (in relation to the target value) which will reduce the quality of the system. Thus a negative relationship between quality and cost is often observed and understood by designers. "Engineers and managers, unaware of the Robust Design method, tend to achieve higher quality by using more costly parts, components, and manufacturing processes" (Logothetis and Wynn, 1989, p. 5). This negative relationship is true in certain situations, but is a major misconception if applied to all levels of design.

4.2 Taguchi's Design Steps

Taguchi defines design at the DOX stage as consisting of two steps: parameter design and tolerance design. Both aim to decrease the output of the system's deviation from the target value, but their methods are very different.

4.2.1 Tolerance Design

Allowance or tolerance design, or what Taguchi also called tertiary design, is the classic quality improvement step (Logothetis and Wynn, 1989). If the variability of the output of the system is even considered, often the means to deal with improving it is by buying parts with higher tolerances. This step is where the understanding that increased quality results from increased cost, and in this step, this relationship is true. If the designers want an output with a higher tolerance, then they buy inputs that have higher tolerances.

4.2.2 Parameter Design: Taking Advantage of Non-linearity

In contrast to tolerance design, parameter design increases quality without increasing costs. Parameter design (secondary design) is the step that is often skipped, but can be highly beneficial. In parameter design, the designer takes advantage of system characteristics in order to be able to use low cost components and still obtain high quality system outputs. In parameter design, increasing quality often results in decreasing costs.

Parameter design entails choosing different values of system inputs to improve system output. The "product and process" are optimized "to make the performance minimally sensitive to various causes of variation" and is ultimately the key element of the technique known as robust design (Phadke, 1989, p. 6). Parameter design is an involved process, but the general concept is very straightforward. Figure 1 shows a simple function, **f(x)**, where the value of **f** depends on **x** in a non-linear way. This could be a function for a host of different processes and systems. The idea can be applied equally well to all of them. An input of the system, **x**, is assumed to have variability. **f(x)** could represent the output of a manufacturing process involving a stove, and **x** may be the temperature of the stove. The operator of the stove sets **x** to say 500 C. However, due to the inconsistencies in the system, the temperature could be anywhere from 480 C to 520 C. This variability is represented by the dotted lines in Figure 1. **x** may be set to $\mathbf{x_2}$, but it ranges in values between the vertical dotted lines of $\mathbf{x_2-\Delta_2}$ and $\mathbf{x_2+\Delta_2}$. $\mathbf{x_1}$ exhibits the same type of variability, but it changes over a smaller range. In the example of the manufacturing process, variability in the temperature may result in variability in the yield of the process; Figure 1 illustrates this fact, showing the tolerances seen in the output. Figure 1 also illustrates the effect of non-linearity. As is apparent from the figure, the range of **f(x)** for $\mathbf{x_2}$ is considerably smaller than the range of **f(x)** for $\mathbf{x_1}$, even though $\mathbf{x_2}$ has greater variability. By changing the value of **x** from $\mathbf{x_1}$ to $\mathbf{x_2}$, the variability of the process can thus be reduced. The idea of parameter design is to change the values of parameters to where the function is minimally sensitive to noise.

x cannot be the only parameter (variable) involved in the process, however. Changing $\mathbf{x_1}$ to $\mathbf{x_2}$ is going to change the output of the process (or system or product), so to compensate there must be an additional factor known as an adjustment factor. The adjustment factor, by definition, is a parameter which changes the output of the process, but does not contribute to variability. By the use of the adjustment factor and setting **x** and all the other factors involved in a process to their optimum levels (levels that produce the least variability in the output) the process can be designed robustly.

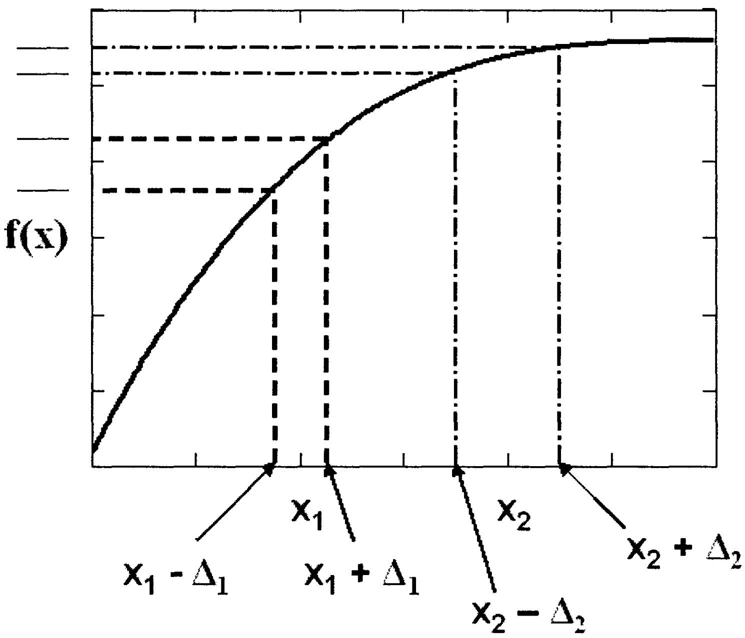


Figure 1: The effect of non-linearity on the variability of process output due to a process input.

5.0 Deviation from Target Performance

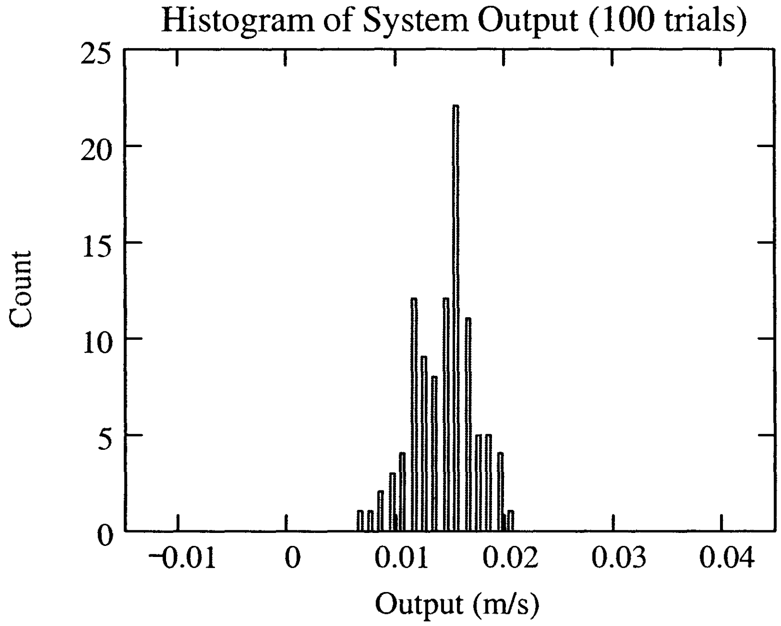
The deviation from the ideal system performance, the tolerances in Figure 1, can be quantified by means of the standard deviation of the data set and the mean of the data set. If a system is run 100 times, a graph like in Figure 2 might be developed. Figure 2 is a histogram which shows the number of data points that fall within certain ranges of the data set. Two parameters are used that describe a set of data. The mean, **μ**, is defined as

$$\mu = \frac{\sum x}{N}, \tag{1}$$

where **N** is the number of elements in the data set, and **x** represents each individual element. The standard deviation, **σ**, can then be defined as

$$\sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}}. \tag{2}$$

The mean is the average of all the points of the system. The standard deviation is a measure of the "width" of the distribution (Del Vecchio, 1997, p. 11). "The majority of naturally occurring things have distributions of [the normal distribution] form" (also called a Gaussian distribution and recognized by a bell-shaped curve) (Del Vecchio, 1997, p. 11). The normal distribution, of which Figure 2 is an example, is 6 standard deviations wide.



$$\text{mean}(N) = 0.01433 \qquad \text{stdev}(N) = 0.00278 \qquad \frac{\text{stdev}(N)}{\text{mean}(N)} = 0.19402$$

Figure 2: Histogram of a process output containing 100 data points.

6.0 Noise factors

The better the designer's understanding of the noise factors that the system or process will encounter, the greater the number of factors that he can consider in the design process, the more complete the design will be, and the more likely that the optimum conditions will be discovered. The major groups of noise factors are described in the following sections.

6.1 Outer Noise

External noise factors (or what Taguchi referred to as outer noises) include "variations in environmental conditions such as temperature, humidity, vibrations, dust, or human interactions in operating a product" (Logothetis and Wynn, 1989, p. 243). As an example consider the life of a car in San Diego versus Philadelphia. The hot and sunny environment characteristic of San Diego is going to affect the life of a car in a very different way compared to the snowy and cold weather characteristic of Philadelphia in winter. The life of the car in either environment is also going to depend on how hard it is run (the conditions at which it is driven). The more often the car spends at wide-open throttle (the accelerator pedal fully depressed) the harder the engine has to work and the more deterioration that will take place.

6.2 Inner Noise

Taguchi's inner noises include anything that results in "deviations of the actual characteristics of a product from the corresponding nominal settings" (Logothetis and Wynn, 1989, p. 243). Inner noise results from several different sources, of which, manufacturing and deterioration are two of the most significant. The MIG welding process is an example of manufacturing noise. MIG welding is often used as the joining process for a huge number of products. If the MIG welding process is not in control, if the speed, wire feed-rate, or current vary, then the quality of the weld will vary. Variability in the quality of a weld results in variability in the strength and durability of the weld, which will result in variability in the quality of the product that has been welded.

Deterioration also is a source of inner noise. Initially, "the functional characteristics of the product may be on target" (Phadke, 1989, p. 23). As the product is used however, the characteristics change because of wear. Changes in the characteristics of a product result in a reduction in product performance due to deviation from the ideal performance. One example of deterioration is in the brakes of a car. Over time the brake pads wear and the brake fluid leaks, producing weaker more sluggish breaks (Phadke, 1989, p. 23).

7.0 Quantification of Quality

If quality is to be optimized, then it must be quantified. One method of quantifying quality is fraction defective where the quality can be related by a step function, as shown in Figure 3. Quality is acceptable as long as it is within the range of $\mathbf{m}-\Delta$ and $\mathbf{m}+\Delta$. The use of fraction defective has classically been applied to manufacturing where the tolerance limits of Figure 3 describe which parts coming out of the process are acceptable enough to enter the market and which must be discarded. The major shortcoming of fraction defective is that it assumes all parts within the tolerances are equally good.

The other major method of quantifying quality is quality loss. Taguchi described quality as "the loss a product causes to society after being shipped, other than any losses caused to society by its intrinsic function" (Logothetis and Wynn, 1989, p. 5). Ideal quality is where the performance of the system, the output, is the ideal (target) value. Thus any deviation, however small, results in a loss of quality. These losses include direct losses due to warranty and increased service costs and the dissatisfaction experienced by the customers as well as indirect losses due to lost market share and the need for increased marketing to sell the product (Logothetis and Wynn, 1989). All these losses can ultimately be expressed as a cost.

In equation form, quality loss can be stated as

$$L(x) = k\cdot(x-\mu)^2, \tag{3}$$

where **k** is the quality loss coefficient described as

$$k = \frac{A_0}{\Delta_0^2}, \tag{4}$$

where $\mathbf{A_0}$ is the cost of quality loss and $\mathbf{\Delta_0}$ are the functional limits in Figure 3. As Equation 3 shows, quality immediately begins to decrease to a squared power as the performance moves off target.

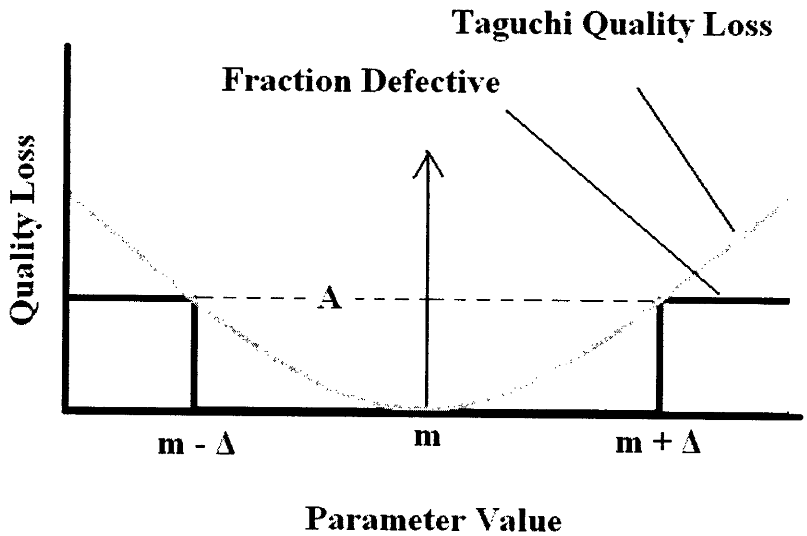


Figure 3: Taguchi Quality Loss versus fraction defective (Phadke, 18).

Equation 3 describes the quality loss for each element in a data set. An average quality loss can be assigned to the entire data set by manipulation of Equation 3. The result is

$$Q = k\cdot\left[(\mu - m)^2 + \sigma^2\right], \tag{5}$$

where **m** is the desired mean of the value. Note that the average quality loss has two components. The first is the standard deviation around the mean of the process, and the second is a result of the entire process shifting from the desired value. The greater the standard deviation, the further the data points are from the target value, and the lower the level of the average quality (the greater the quality loss). As would be expected, a shift in the mean of the data results in an immediate loss of quality due to the offset from performance. Both the deviation and the mean shift must be minimized to achieve optimum quality.

8.0 The Robust Design Method

The idea of the robust design method is to minimize the quality loss (Equation 5) of a system by the use of parameter design (and eventually tolerance design as economically justifiable). "Robust design is a methodology for finding the optimum settings of the control factors to make the product or process insensitive to noise factors" (Phadke, 1989, p. 67). Graphing the system with relation to each parameter as shown in Figure 1 is not practical, and in many cases is impossible when there is no equation to describe the system. Also an equation like in Figure 1 does not take into account the effect of interactions, which can be important. Section 8.1 lists the steps of the robust design method. Sections 8.2 and 8.3 then discuss the ideas of the experiments used in the robust design method and the choice of the variable that should be optimized.

8.1 The Taguchi Procedure (the Steps of Robust Design)

The robust design process follows essentially eight steps (Logothetis and Wynn, 1989):

1. **Define the problem**: Establish what the experiments and the application of the method is intended to achieve.
2. **Determine the objective:** The output of the system which is to be optimized needs to be determined, as well as the appropriate method of measuring the chosen output.
3. **The Brainstorming Session**: Once the problem is understood and the objective set, a meeting must take place between statisticians and engineers to determine the controllable and uncontrollable factors, the experimental range, and the appropriate factor levels.
4. **Design of the Experiment**: The appropriate efficient experiment is chosen. The result of this step are reduced experiments that seek to achieve the desired results in the shortest amount of time and lowest cost commitment possible (refer to the next section).
5. **Set up the experiment and collect data**
6. **Analyze the data:** Determine the optimum levels of the control parameters. Run analysis of variance techniques, or alternative methods.
7. **Interpret the results**
8. **Run a confirmatory experiment and plan for future actions**

Step 4 is paid substantial attention to and is described in the next section. Steps 5 through 8 are explained in detail in the examples. More information can be found on the steps of robust design in the references.

8.2 Design of Experiments

If the problem is not clearly defined, the objective determined, and the factor levels correctly chosen, then design of experiments has little use. However, if care is taken in the early steps of Robust Design, a well-designed experiment can result in a fast and efficient means to find out a lot about a system. The information from the experiment is used to set the control factors at the desired levels in order to achieve a more robust system design.

8.2.1 Basic Construction of a Simple Design

Objective is to "extract a large fraction of the information in a matrix from a smaller fraction of the numbers in that matrix" (Del Vecchio, 1997, p. 1). If it is desired to test 4 factors (parameters/ inputs to a system) at 3 levels each, the required number of experiments to test every possible combination is 81 (3^4). However, with efficient experimentation, "doing as few as 16... can reveal a great deal about the process; doing 25 experiments... can furnish almost as much information as doing 81 might" (Del Vecchio, 1997, p. 2). Thus in order to save time as well as avoid costly trial and error, design of experiments can be very beneficial.

The goal is to avoid 1-FAT (one factor at a time) experiments in which an operator simply changes one control factor at a time to see how the output of the process or system changes. If not organized, such as when an experienced technician just trys to "feel out" the best settings, 1-FAT experimentation can take a very long time and will most likely result in suboptimal conditions. "1-FAT experiments by their nature are not capable of finding" interactions, which can also be very important in finding the optimum of a process (Del Vecchio, 1997, p. 3).

Consider an experiment of three factors at two levels each. The full experiment, where all combinations are considered is shown in Figure 4, where the + and – signs represent the higher or lower level of factor A, B, or C. This pattern is the same as a cube in space on x, y, and z axes, as shown in Figure 5. Statistical theory states that from 8 pieces of data, 7 comparisons can be made. We can thus expand are initial experiment by either considering interactions or considering other levels.

	A	B	C
1	+	+	+
2	+	+	-
3	+	-	+
4	+	-	-
5	-	+	+
6	-	+	-
7	-	-	+
8	-	-	-

Figure 4: Full design involving 3 factors at two levels each.

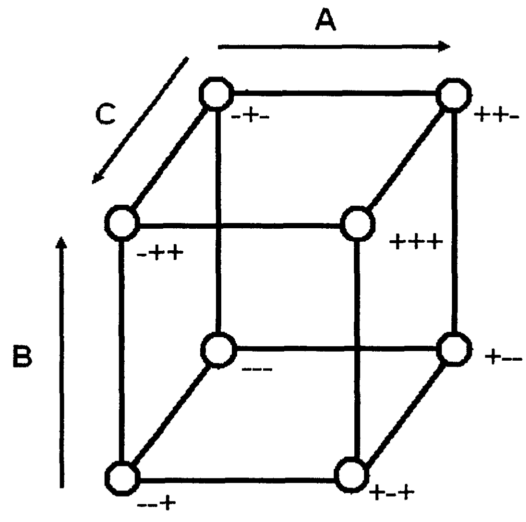


Figure 5: Cube representing a 3 factor (A, B, C) design where each factor has two levels (+, -).

The first option is to add 4 additional columns to the initial 3 (for a total of 7 comparisons) in order to determine interactions, as shown in Figure 6. The numbers 1 through 8 on the left of Figure 6 denote the experiment number. The matrix shown in Figure 6 is a fully expanded design, every combination of the factors occurs and all the interactions have been considered.

The way Figure 6 is created is by multiplying combinations of the first three columns. As an example, consider column AB. The first element in column A (the higher value of A represented by a plus sign) is multiplied by the first element in column B (B's higher value) to give the first element in column AB (+). Likewise, the third element of AB (in the third row) is created by multiplying the positive in A with the negative in B to give a negative value in AB. Going down column AB, one will note that it is simply the multiplication of A and B. The other columns have been created in the same manner. If the average of the results of all experiments marked with a positive sign in column AB is compared to the average of all the results marked with a negative sign, the difference between these two data sets gives a solid understanding of the effect of the AB interaction on the system output value. Columns AC through ABC can be analyzed in the same manner. The difference between the positive and negative values of each factor can then be plotted as a quick means to determine the magnitude of the effect of that factor. This type of plot, known as a Scree plot, will be used in the example section and is an effective means of separating the important effects from the unimportant or those simply caused by background noise.

	A	B	C	AB	AC	BC	ABC
1	+	+	+	+	+	+	+
2	+	+	-	+	-	-	-
3	+	-	+	-	+	-	-
4	+	-	-	-	-	+	+
5	-	+	+	-	-	+	-
6	-	+	-	-	+	-	+
7	-	-	+	+	-	-	+
8	-	-	-	+	+	+	-

Figure 6: Full L8 design where every possible combination of the three factors is considered (all interactions are accounted for).

If the pluses and minuses are added in each column, the total is zero. If the settings in any one column are multiplied by settings in any other column, the total sum is again zero. This property of the table makes the design orthogonal, and thus called an orthogonal array. "Orthogonal arrays are typically most efficient" (Del Vecchio, 1997, p. 29).

The exact same design in Figure 6, instead of being used to determine the effects of interactions can be used to determine the effects of other factors. Columns AB, AC, BC, and ABC can be replaced with new factors, as shown in Figure 7. The array shown in Figure 7 can no longer compare interactions; however if interactions are not present or insignificant, this experiment reduces a 2^7 experiment, which would require 128 experiments, to a mere 8. A design such as that shown in Figure 7, known as a saturated design, is both useful as a screening test to determine which factors are important to the process and which are not, as well as to run a heavily reduced experiment where the designer is certain that no interactions are present.

	A	B	C	D	E	F	G
1	+	+	+	+	+	+	+
2	+	+	-	+	-	-	-
3	+	-	+	-	+	-	-
4	+	-	-	-	-	+	+
5	-	+	+	-	-	+	-
6	-	+	-	-	+	-	+
7	-	-	+	+	-	-	+
8	-	-	-	+	+	+	-

Figure 7: Saturated L8 design; seven factors at two levels each.

As an analogy to saturated versus full designs, consider again the cube in Figure 5. The purpose of a saturated design is to get information about the full data set from only a sampling of that data set. Likewise, the shape of the cube shown in Figure 5 (the full data set) can be described from only the shadow of the cube cast by selected points, as shown in Figure 8. The information contained in the cube in Figure 5 can be fully described by the shadow in Figure 8 if no interactions are present. The cube in Figure 5 is the full design (an L8 with three factors), and Figure 8 is the saturated design represented in Figure 9.

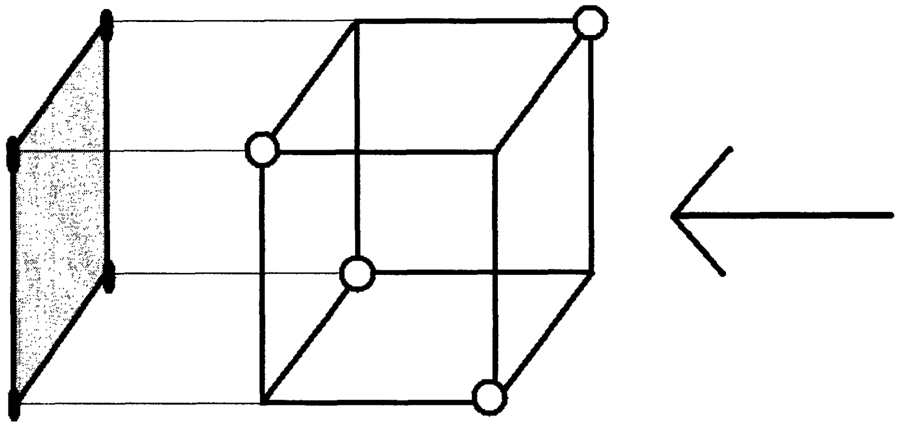


Figure 8: A cube can be described just by its shadow, just as a full design can be described just by the equivalent saturated design.

	A	B	C		A	B	AB
1	+	+	+	or	+	+	+
2	+	-	-		+	-	-
3	-	+	-		-	+	-
4	-	-	+		-	-	+

Figure 9: The design on the left of Figure 9 is the saturated design for the cube shown in Figure 8. The design on the right is the equivalent two factor design with an interaction.

8.2.2 False Data due to Interaction

"Taguchi believes that it is generally preferable to consider as many factors (rather than many interactions as economically feasible for the initial screening" (Logothetis and Wynn, 1989, p. 249). Ignoring interactions is often okay: "In many processes, factor interactions are not common, and the use of saturated designs is very appropriate and highly efficient" (Del Vecchio, 1997, p. 30). However, in "many mechanical/ chemical systems, interactions are regularly encountered, and the use of saturated designs can result in false or misleading data" (Del Vecchio, 1997, p. 30). Determining interaction levels and accounting for them if they are present is thus very important.

If interactions do exist their effects become mixed in with the rows already in the experiment. For the design in Figure 9, if the columns are multiplied together to create interaction columns, the resulting interaction columns are identical to the single factor columns, as shown in Figure 10. Examination of Figure 10 reveals that A is the same as BC, B is the same as AC, and C is the same as AB. The interactions are now part of a particular column; the isolation of the effect of the interaction as seen in Figure 6 has been lost. Thus if an interaction between A and C is occurring, this will show up falsely as an effect due to factor B.

	A	B	C	AB	AC	BC	ABC
1	+	+	+	+	+	+	+
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	-	-	+	+	-	-	+

Figure 10: The L4 design compared to columns representing the interactions between columns in the L4. The interaction columns are the same as the single factor columns; the interactions cannot be separated.

If there are interactions, then the more lengthy experiments must be performed. The opposite of the L4 matrix in Figure 10 can be run to improve data and help the design consider interactions. The opposite matrix along with the original L4 is shown in Figure 11. Combined, these two matrices happen to be the full L8.

	A	B	C	A	B	C
1	+	+	+	-	-	-
2	+	-	-	-	+	+
3	-	+	-	+	-	+
4	-	-	+	+	+	-

Figure 11: An L4 design and the complimentary matrix. The combination of these two designs produces the L8 design.

Saturated designs and full designs are not the only two types available to experimenters. Some columns in a design can be left empty for the purpose of interactions, while others are used for additional factors. A linear graph is a tool used to quickly describe how a design accounts for interactions. Figure 12 is the same design as in Figure 6, but the columns have been rearranged slightly and numbered instead of lettered. The individual factors do not have to be placed in the first 3 columns; they and more factors can be placed in any column. The two possible linear graphs for the L8 design in Figure 12 are shown in Figure 13. Referring to option 1, if the designer wants to estimate the interaction between columns 1 and 2 then he leaves column 3 empty. The second option depicts that if the only interactions that occur are between factor 1 and the other factors, then these interactions can be estimated by leaving columns 3, 5, and 6 empty and putting the other factors in 2, 4, and 7. By leaving a column empty, just that is meant. Nothing is placed within the column, and the extra experiments that the saturated design would have required to estimate the effects of the factor in that column are now instead used as extra data to estimate factor interactions between columns. Linear graphs are very effective for quickly determining what experiments would be best for a particular system.

	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

Figure 12: Rearranged L8 (2^7) design.

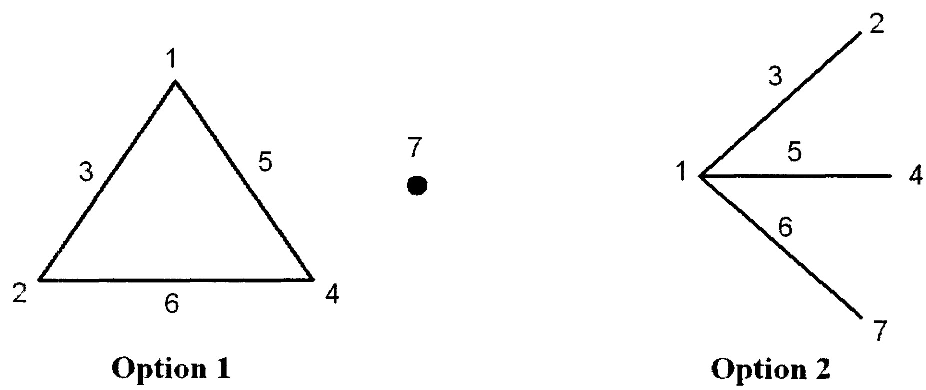


Figure 13: Linear graphs for the L8 shown in Figure 12. Each number represents a column in Figure 12.

8.2.3 More Complex Designs

Designs with only two levels for each factor are the most straightforward because the way they are created and analyzed is readily apparent. However more detailed designs involving more levels as well as different types of designs are shown in the references and used in the examples. The L9 design is listed below in Figure 14 to provide an example. The linear graph for Figure 14 is shown in Figure 15. As is apparent from Figure 15, to estimate interactions between columns 1 and 2, all other columns must be left empty. Designs can become more complicated and more detailed in their methods and include other types of designs other than the straightforward orthogonal arrays described here (such as Plackett-Burman, Box-Behnken, etc). For a full discussion, refer to the references. Regardless, the designs described here can be extremely useful in introducing efficiency to the robust design method. The next step is to consider the parameter that will actually be measured in the experiments.

	1	**2**	**3**	**4**
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Figure 14: Taguchi L9: Saturated design with four factors at three levels each.

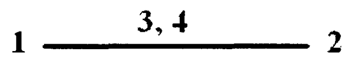


Figure 15: Linear graph for Taguchi L9.

8.3 The Signal to Noise Ratio

The signal to noise ratio (SN ratio) is the main parameter used by the robust design method to obtain the optimal settings of the control factors. The signal to noise ratio is the output used for each of the experiments shown in the last section when they are applied to robust design techniques. The choice of the signal to noise ratio depends on the nature of the system.

8.3.1 Signal to Noise Ratio for Nominal-is-Best Type Problems

The idea of robust design is to decrease the sensitivity of a system to variability in the inputs of that system. As the quality loss equation (Equation 5) describes, both the shift of the mean from the target value and the standard deviation need to be considered when attempting to optimize a system to a particular value. Equation 5 can however be misleading. As was discussed, there is an adjustment factor that is used to shift the mean back onto target, so the sensitivity to noise of Equation 5 needs to be isolated from the corresponding mean shift that goes along with each new combination of control factors since the adjustment factor will eliminate this mean shift. If a parameter was used that didn't include adjustment, the mean shift would most likely dominate, and the system would not be properly optimized. If the standard deviation without adjustment is σ, then the predicted standard deviation after adjusting the mean on target is

$$\sigma_a = \frac{\mu_0}{\mu} \cdot \sigma \quad , \qquad (6)$$

where σ_a is the standard deviation after adjustment, μ_0 is the mean before adjustment, and μ is the mean after adjustment. If σ_a is then placed into Equation 5, the result is

$$Q_a = k \cdot \mu_0^2 \cdot \frac{\sigma^2}{\mu^2} . \qquad (7)$$

The constant in Equation 7, $k\mu_0^2$, does not contribute to the optimization of the system, so it can be ignored. Equation 7 rearranged is given as

$$\eta = 10 \cdot \log\left(\frac{\mu^2}{\sigma^2}\right), \qquad (8)$$

where η is now the common form of the signal to noise ratio for nominal-is-best type problems. In contrast to Equation 7, which must be minimized to achieve the highest quality, Equation 8 should be maximized. Maximizing Equation 8 becomes the primary objective of the robust design method.

All problems are not nominal-is-best type problems, so multiple variations for signal-to-noise ratios exist. The designer of a system may be attempting to maximize or minimize the system. The system may also not be static, where the value of the output is dependent on the value of a signal factor, such as with a temperature controller (a continuous dynamic system). Appropriate signal to noise ratios for these different situations can be found in the references.

8.3.2 SN Ratios for Larger-is-Better or Smaller-is-Better Type Problems

In a maximum or minimum type problem, the concern is not about the standard deviation since the standard deviation, in itself, is part of what produces the mean shift and does not allow the function to reach the ideal level. In problems where larger/smaller is better, there also exists no adjustment factor. The signal to noise ratios for these type of problems are thus only functions of essentially the mean of the system. The SN ratio for a smaller-the-better type function for a static problem is given as

$$\eta = -10 \cdot \log\left[\frac{1}{n} \cdot \sum_{i=1}^{n} (y_i)^2\right], \qquad (9)$$

where **y** is each individual data element and **n** is the number of elements in the sample. For a larger the better type problem, the SN ratio is

$$\eta = -10 \cdot \log\left[\frac{1}{n} \cdot \sum_{i=1}^{n} \frac{1}{(y_i)^2}\right]. \qquad (10)$$

In both Equations 9 and 10, the designer is just looking for essentially an average. In contrast to the thinking of many designers, this average can often be lower than expected due to a mean shift.

9.0 Mean Shift

The occurrence of a mean shift in a system's output is often a completely unexpected event. Whether or not a system is being designed to minimize variation, the designers of the system know that variability in the inputs will result in variability in the output. The designers expect to see variability in the performance of the product. When it comes to the mean performance however, designers rarely consider that the variability will have an effect. The misconception is that plugging the optimum input values (the nominal values) into either a prototype or a simulation and reading the output will give the mean performance with variability even though variability has not been considered; the assumption is that variation will just occur around the nominal mean. For some systems using only the intended inputs without considering variations may predict the output even under variability. However in systems where a maximization or minimization occurs, the ideal values will predict a mean value closer to the optimum than what will be apparent in the actual system; a mean shift will occur in the performance of the real hardware.

9.1 The Importance of Mean Shift

Even a small shift in the mean can be important. For example, consider a large shipping company that intends to buy a large volume of engines for a fleet of new trucks. The performance of the engines, measured in terms of the efficiency of the engines, is of considerable interest to the shipping company. The more efficient the engines, the better the gas-mileage of the vehicles, and the greater the savings the company gets on gasoline. Assume that the shipping company was quoted a fuel consumption rate for the engines (a parameter that directly relates to the engine efficiency). In considering which engines from which company to buy, the shipping company takes into account fuel efficiency. Also, having then bought the engines, the company budgets money for fuel. If the engines have a lower efficiency than expected, this is going to result in an unexpected cost for the company.

Engine performance can often be described by detailed computer simulations. A prototype of the engine could also be built to give an idea of how the engine performs. The computer program, with the nominal values of the engine parameters, as well as the prototype may give a specific description of engine performance in terms of the nominal values of the input parameters. However, most likely variability will not be considered by the engine manufacturer.

To get an idea of how this shift in engine performance takes place, consider the plot of fuel conversion efficiency versus the fuel air ratio in Figure 16, where the line plotted is for a particular engine compression ratio. This graph is for an ideal cycle, a simple prediction of engine performance, but the trends it conveys are the same seen in real engine performance. The vertical dashed lines on the graph represent the limits of the data variability. A sudden change in the slope of the graph takes place at a fuel-air equivalence ratio of 1. If variability occurs in the fuel-air ratio, which is common in an engine, that variability will have an effect on the efficiency of the engine. If the fuel-air ratio is set at 1, which is often the case, the difference in slope on the left versus the right of this point can result in a mean shift from the expected value of the efficiency, as shown in the figure. The mean shift, even if small, over a large number of vehicles can result in costs to the company that would not have been considered before the purchase of the fleet. If the variability of the process is not accounted for, unexpected loss in quality can result.

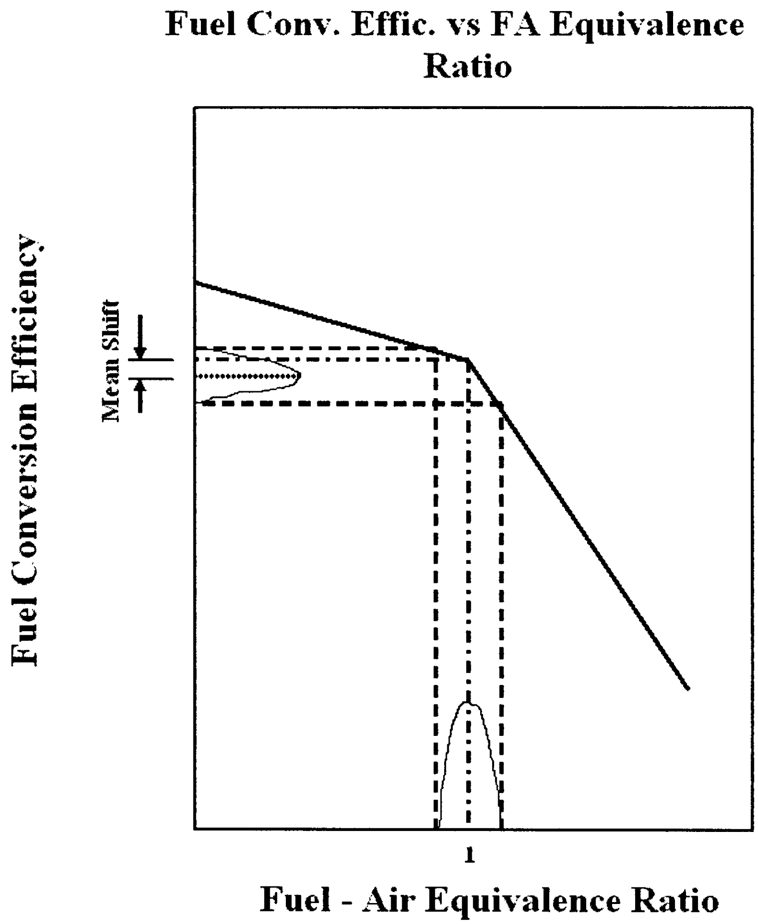


Figure 16: Engine fuel conversion efficiency vs. the fuel air equivalence ratio for a particular compression ratio (Adapted from Heywood, 1988, p. 182). Variability in a nominal fuel air ratio of 1 can result in a mean shift in the output.

9.2 Sampling Methods

Up to this point, the discussion has been general in the sense that all the ideas discussed can be applied to essentially any system. The system can be described by an equation, and the variability in the inputs can be simulated by giving the data a mean and a standard deviation. The system or process could also be a real system where the only way to test variability and run experiments is to actually run the hardware of the system or process. For the real processes, the sampling methods described in the following sections are not very applicable. However for computer simulations, sampling methods can be very helpful.

9.2.1 Monte Carlo

Monte Carlo is the most straightforward method of simulating variability in a system process. The inputs to the simulation are chosen from a probability distribution. In the case of the normal distribution, the input values would be chosen randomly from a distribution with a mean and standard deviation. Since there is no method to choosing the samples from the distribution, the Monte Carlo sampling method is described as a "random walk" type of method. After a certain number of trials, the output of the computer program settles down to having its own mean and standard deviation due to the variability in the inputs. The issue with the Monte Carlo method is that the number of trials before the final value settles down can be large. The fact that a computer program is being used certainly helps, but in functions of many variables the time commitment of the computer can still be extensive when running experiments.

9.2.2 Latin Hypercube

Latin Hypercube sampling is not a random walk, but attempts to apply a method to the sampling to decrease the ultimate number of trials needed before the output values settle. Latin hypercube sampling breaks the distribution of the variable into sections of equal probability. Values of the parameter are then taken from each "bin." The result is that fewer samples are needed to get a set of input values that portrays the desired characteristics (mean and standard deviation) assigned to the input. Histograms comparing Latin Hypercube sampling to Monte Carlo sampling are shown in Figures 17–19. The histograms are each broken into 25 sections. As is apparent from the figures, the Latin Hypercube sampling quickly converges to the normal distribution, while even after 600 trials (not shown), the Monte Carlo sampling hasn't completely settled down. The faster the inputs exhibit their intended behavior (in this case, with a mean and a standard deviation) the faster the output will settle to its respective mean and standard deviation. Statistically, the Latin Hypercube sample set will become valid faster than the Monte Carlo sample set.

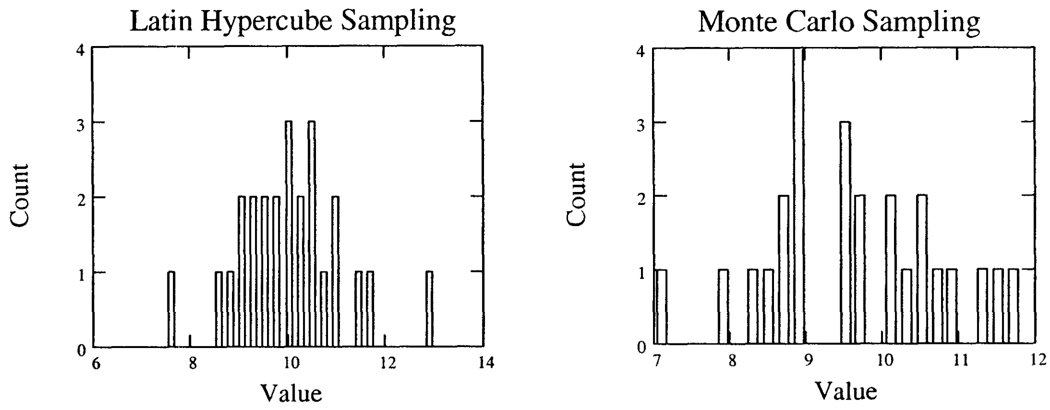


Figure 17: 25 bins, 25 samples

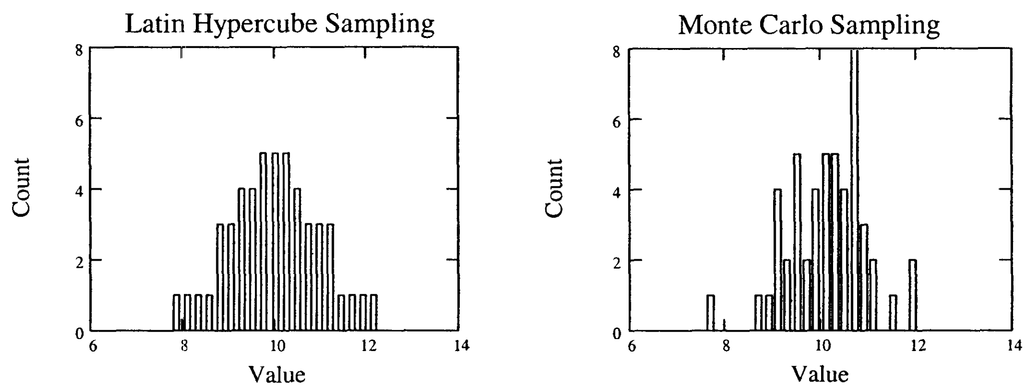


Figure 18: 25 bins, 50 Samples

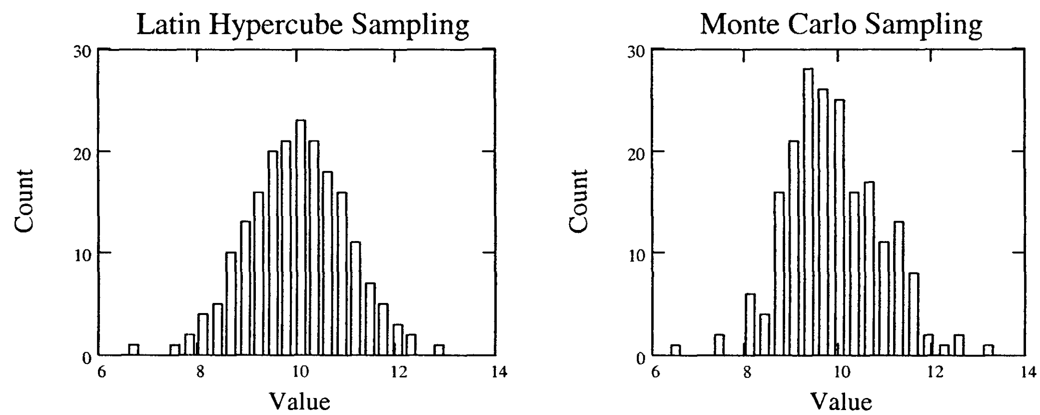


Figure 19: 25 bins, 200 Samples

9.2.3 Star Patterns, 2n+1 and 4n+1

The third sampling method is very different from the Monte Carlo and Latin Hypercube techniques because it no longer uses random sampling. The star patterns instead attempt to create a curve fit at the point of interest. Consider the function **f(x)** as shown in Figure 20. Variation in the value of **x** causes variability in the output value of **f(x)**. Since the output is at a maximum, any value other than the ideal value of **x** causes the output to be less than the maximum value of **f** (similar situation to that shown in Figure 16). The result is that all the data is bunched against the peak of the graph, and a mean shift results, as shown in Figure 20.

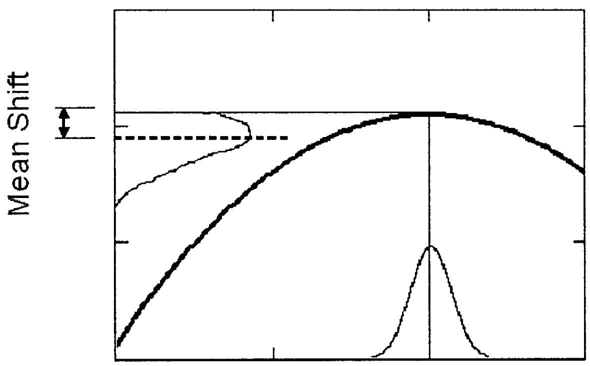


Figure 20: Mean shift in the output due to variability in the input at a maximum.

The star pattern creates a curve fit by first reassigning the origin. The value of **f(x)** at the nominal value of the input, **x**, is redefined as the value at x = 0, as shown in Figure 21. After the origin has been redefined, the curve fit for the star pattern is created by solving a system of equations. The equations are developed by defining an equation from the peak and then solving for different deviations from the new x=0 point. Figure 22 is a depiction of the function input values and the sampling that takes place. The center of the figure is not the original origin, but instead is the redefined origin as shown in Figure 21 for each input factor. For the 2n+1 pattern, each factor is considered at three different points.

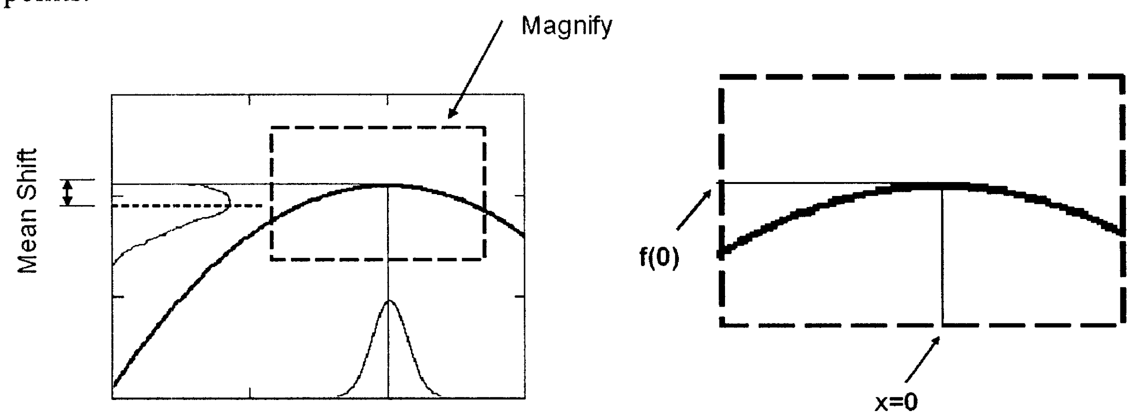


Figure 21: To create a star pattern, the origin of a function is first reassigned to the point of interest for an input value.

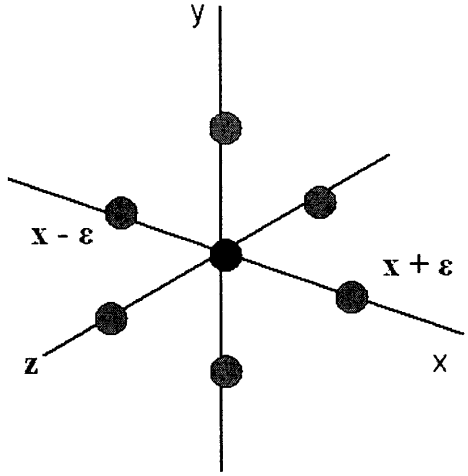


Figure 22: Star pattern for 2n+1 sampling method.

The equations for a **2n+1** pattern are given as

$$f(x) = f_0 + c\cdot x + k\cdot x^2 , \tag{14}$$

$$f(x+\varepsilon) = f_0 + c\cdot(x+\varepsilon) + k\cdot(x+\varepsilon)^2 , \tag{15}$$

$$f(x-\varepsilon) = f_0 + c\cdot(x-\varepsilon) + k\cdot(x-\varepsilon)^2 , \tag{16}$$

where **ε** is a step size to be determined, and $\mathbf{f_0}$, **c**, and **k** are constants. Equation 14 is the actual curve-fitting function being solved for. The other two equations are this same function, but at different **x** values just a small step away from the original value. The result of solving for $\mathbf{f_0}$, **c**, and **k** is

$$f_0 = f(0) , \tag{17}$$

$$c = \frac{1}{2}\cdot\frac{f(\varepsilon) - f(-\varepsilon)}{\varepsilon} , \tag{18}$$

$$k = \frac{1}{2}\cdot\frac{f(-\varepsilon) - 2\cdot f(0) + f(\varepsilon)}{\varepsilon^2} . \tag{19}$$

Thus, $\mathbf{f_0}$ is the value of the function at **x** = 0 (the reassigned origin). To deal with the fact that **f(x = 0)** in Equations 17 through 19 does not refer to the original origin is a trivial matter. The equations are modified to account for this discrepancy:

$$c = \frac{1}{2} \cdot \frac{f(x_0 + \varepsilon) - f(x_0 - \varepsilon)}{\varepsilon} , \qquad (20)$$

$$k = \frac{1}{2} \cdot \frac{f(x_0 - \varepsilon) - 2 \cdot f(x_0) + f(x_0 + \varepsilon)}{(\varepsilon)^2} , \qquad (21)$$

where $\mathbf{x_0}$ is the nominal value of the input parameter **x**. If the noise in the input variable is Gaussian then only the even parts of the curve fitting function are necessary to give an accurate prediction of the system output. Pure quadratic terms as well as fourth order terms have an influence, but 1st and 3rd order terms as well as any interactions between different factors (for instance x and y interactions) are not considered since they are odd. The prediction value of f(x) due to the shift for the 2n+1 pattern is then

$$f(x) = f_0 + k \cdot \sigma^2 , \qquad (22)$$

where $\mathbf{f_0}$ and **k** are the constants defined in Equation 17 and Equation19 (which is the same as Equation 21).

The prediction for the 4n +1 pattern is determined by the same procedure as for the 2n+1. The star pattern is shown in Figure 23.

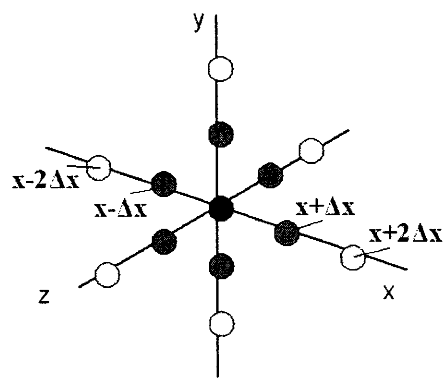


Figure 23: Star Pattern for 4n+1

The results of the 4n+1 pattern, listing only the quadratic and fourth order terms, is

$$g_0 = g(0) , \qquad (23)$$

$$k = \frac{-1}{24} \cdot \frac{-16 \cdot g(\varepsilon) + 30 \cdot g(0) - 16 \cdot g(-\varepsilon) + g(2 \cdot \varepsilon) + g(-2 \cdot \varepsilon)}{\varepsilon^2} , \qquad (24)$$

$$q = \frac{1}{24} \cdot \frac{-4 \cdot g(\varepsilon) + 6 \cdot g(0) - 4 \cdot g(-\varepsilon) + g(2 \cdot \varepsilon) + g(-2 \cdot \varepsilon)}{\varepsilon^4} , \qquad (25)$$

where the function is now **g(x)** and the step size ε in Equations 23-25 is different from the ε in Equations 14-22. Equations 23 – 25 must again be shifted back to the original x value at the point where mean shift is being estimated (as was done between Equations 18 and 19 and Equations 20 and 21). The mean for the 4n+1 pattern can then be predicted by

$$g(x) = g_0 + k \cdot \sigma^2 + 3 \cdot q \cdot \sigma^4 \quad . \tag{26}$$

The only term in the star pattern equations still needed is ε. For the 2n+1 pattern

$$\varepsilon = \sigma \cdot \sqrt{\frac{3}{2}} , \tag{27}$$

where **σ** is the standard deviation of the variable in question. For the 4n+1 pattern,

$$\varepsilon = \sigma \cdot \left(\frac{15}{11} \right)^{\frac{1}{4}} , \tag{28}$$

where again **σ** is the standard deviation of the input variable. Note that the constants have been considered for **f(x)** in terms of only **x**, but since only even functions have an influence and no interactions need to be considered, the technique is applied to many variables by simply adding up the terms and combining the prediction equations for each variable. Thus the 4n+1 patter for 3 variables can be written as

$$g(x) = g_0 + k_1 \cdot \sigma_1^2 + k_2 \cdot \sigma_2^2 + k_3 \cdot \sigma_3^2 + 3 \cdot \left(q_1 \cdot \sigma_1^4 + q_2 \cdot \sigma_2^4 + q_3 \cdot \sigma_3^4 \right) \quad , \tag{29}$$

where the subscript of each term refers to the variable.

10.0 Illustrative Model: Development of a Novel Fluid Heating Device

To illustrate the ideas of the robust design method, particularly the analysis methods, two examples are developed and then worked through in the following sections. The first model involves the optimization of a fluid heating device by the application of the robust design method. The goal of this model is to provide an illustrative example of the design method as well as to describe how the data obtained from the experiments is analyzed.

10.1 Description of the Device and Problem

A fluid flowing through a tube can be heated by means of a finned surface, as shown in Figure 24. The plate (surface) of the finned heating device is kept at a constant temperature higher than the fluid temperature. The goal of the device is to provide a constant heat transfer to the fluid flowing past the fins. However due to variability in the size of the fins, the material properties, the fluid properties, and the water velocity, the rate of heat transfer is also variable. Thus the objective is to apply the robust design method in an attempt to reduce the variability in the output of the process due to the variability of the input parameters. Figure 25 shows a 3-D view of the heating device imbedded in the tube shown in Figure 24. The fins used are pin fins.

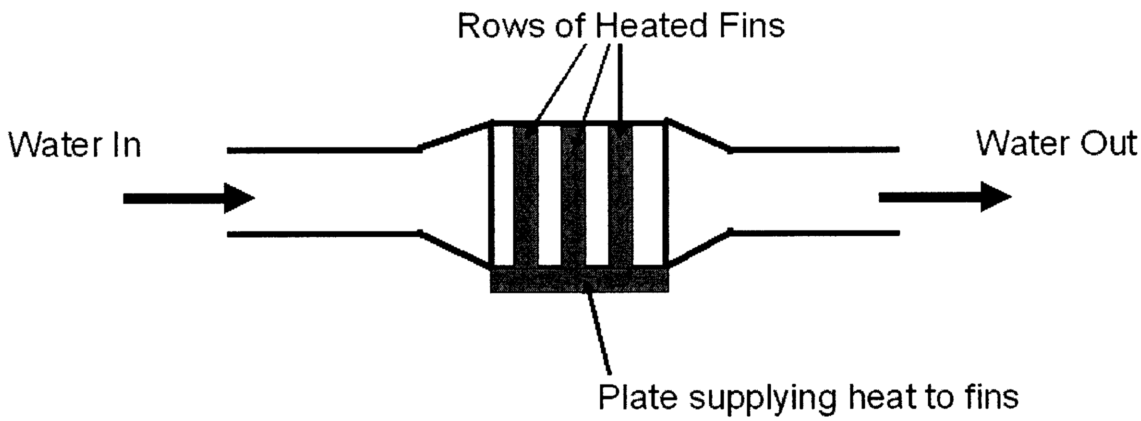


Figure 24: Finned heating device.

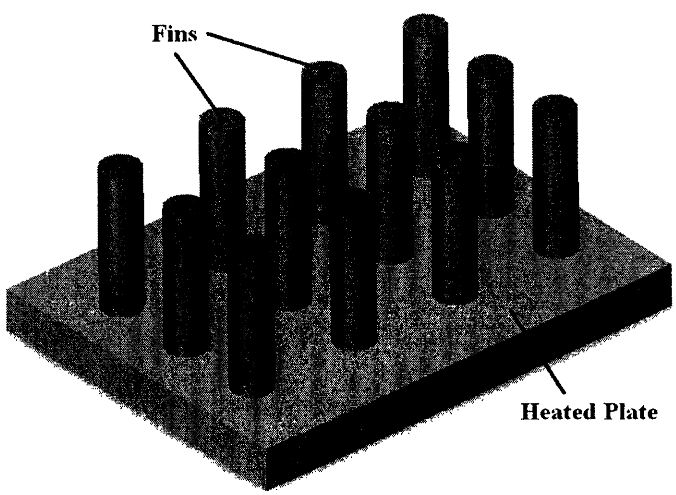


Figure 25: 3-D view of the actual heating device. A plate kept at a constant surface temperature heats aligned rows of pin fins which then heat the fluid flowing past the fins.

10.2 Model Derivation

The model is straightforward. The system is in steady state, so the heat transfer from the fins can be modeled according to the equation

$$\text{Qdot\_fin} = \sqrt{h_c \cdot P \cdot k \cdot A_c} \cdot (T_s - T_{inf}) \cdot \tanh(m \cdot L) \quad , \tag{30}$$

where $\mathbf{h_c}$ is the convective heat transfer coefficient, $\mathbf{T_s}$ is the temperature of the plate, $\mathbf{T_{inf}}$ is the bulk temperature of the fluid, **L** is the length of each fin, **P** is the parameter of a fin, $\mathbf{A_c}$ is the cross-sectional area of a fin, **k** is the thermal conductivity of the metal of the heating plate and fins, and **m** is the fin parameter. The fin parameter is defined as

$$m = \sqrt{\frac{h_c \cdot P}{k \cdot A_c}} . \quad (31)$$

The convective heat transfer coefficient in Equation 31 is determined by a correlation where the Nusselt number is related to the heat transfer coefficient by

$$h_c = \frac{Nu \cdot k_w}{D} , \quad (32)$$

where **Nu** is the Nusselt number, $\mathbf{k_w}$ is the thermal conductivity of the fluid, and **D** is the diameter of the fins shown in Figure 25. The appropriate Nusselt correlation for a bank of fins with aligned rows is

$$Nu = Pr^{0.36} \cdot 0.27 \cdot Re^{0.63} , \quad (33)$$

where **Re** is the Reynolds number and **Pr** is the Prandtl number. This correlation holds for $10^3 < \mathbf{Re} < 2 \cdot 10^5$ and **St/Sl** > 0.7 where **St** and **Sl** are the distances between fins in each row and each column respectively, as shown in Figure 26. Since Equation 23 is truly valid only for the rows in the center of the tube bundle, a correction factor is used as the number of rows is reduced. The correction factor can be found in the references (Lienhard, 2003, p. 383). The **Pr** number can be further defined as

$$Pr = \frac{\mu \cdot c_p}{k_w} = \frac{v}{\alpha} , \quad (34)$$

where $\boldsymbol{\mu}$ is the dynamic viscosity of the fluid, **v** is the kinematic viscosity, $\boldsymbol{\alpha}$ is the thermal diffusivity, $\mathbf{k_w}$ is the thermal conductivity of the water, and $\mathbf{c_p}$ is the heat capacity of the water. Finally, the Reynolds number is defined as

$$Re = \frac{\rho \cdot v \cdot D}{\mu} , \quad (35)$$

where the velocity, **v**, is the flow rate taking into account the fact that the area has been reduced due to the presence of the fins, and $\boldsymbol{\rho}$ is the density of the fluid.

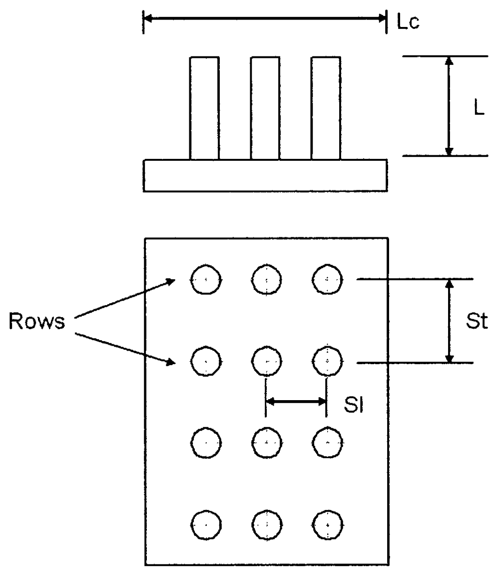


Figure 26: Schematic representation of the heating element.

The flow rate increases as the fluid enters the rows of tubes since the cross sectional area available to flow is reduced. The volumetric flow rate into the heater is set at a constant 1.93 kg /sec. The flow velocity into the heater is allowed to vary, so to compensate for the constant mass flow rate, the area of the inlet to the box varies. As the velocity increases, the area must decrease to keep the mass flow-rate constant. For an initial velocity of 0.75 m/s, the cross-sectional area of the "box" containing the fins which the flow enters is set at 4 square inches regardless of the length or diameter of the fins. Thus as the fins become longer the available row length, **Lc**, decreases. The number of fins per row is chosen so that the distance between 2 fins is 2.2 times the diameter of each fin. The number of fins per row is however many fins can fit given this requirement. The heat transfer obtained by all the fins is then given as

$$\mathrm{Qdot} = M \cdot N \cdot \sqrt{h_c \cdot P \cdot k \cdot A_c} \cdot \left(T_s - T_{inf}\right) \cdot \tanh(m \cdot L) \quad , \tag{36}$$

where **M** is the number of rows and **N** is the number of fins per row. The heat transfer to each successive row actually decreases since the temperature difference driving the heat transfer decreases as the fluid heats up; this effect however is ignored in the model.

10.3 Development of the Experiment

The control factors for the device are chosen to be the velocity (**v**) of the incoming fluid before it reaches the tube bundles, the diameter (**D**) of the fins, the length (**L**) of the fins, and the thermal conductivity (**k**) of the fin material. The aim is to design an experiment that incorporates three test levels for each factor. The nominal value originally assigned to the control factor is set as the intermediate level, and then a positive and negative deviation from this nominal value are set as the higher and lower values respectively for each factor. The desired levels for each factor are listed in Table

1. The range in each of the parameters was chosen to be large, so as to be able to consider many different options in the design of the heater.

Table 1: Test levels of the four chosen control factors

	Levels		
Parameter	**1**	**2**	**3**
k (W/m-K)	30	60	90
D (mm)	0.769	5.083	9.398
L (mm)	12.7	25.4	38.1
v (m/s)	0.225	0.75	1.275

The output of the system that is chosen to be optimized is the heat transfer rate of the fins (Equation 36). The desired output is 3.85 kW, even though the greater concern is to attempt to reduce the variability. The signal to noise ratio for a nominal-the best type problem is chosen as the appropriate optimization parameter (Equation 8). The adjustment factor chosen is the number of fin rows, since the total heat output is only a multiplication of the number of rows.

10.4 Testing for Factor Interaction

To design the correct matrix experiment, the parameter values were first tested in an L16 design to determine what interactions occurred if any. The L16 design is shown in Figure 27, where the control factors, **k**, **D**, **L**, and **v** are assigned to columns A, B, C, and D respectively. As explained in earlier sections, a very convenient way to determine factor effects is to create a Scree chart. The output of each experiment (row) is a value for the heat transfer rate as a function of the parameter levels specified in columns 1 through 4. The experiment is run using the lowest level in Table 1 as the lowest level in Figure 27 and the highest level in Table 1 as the highest level in Figure 27. The output values of the experiments are placed into a single column. To then determine factor levels for say column 5 (the AB interactions) the average of all the output-values in the output column corresponding to the rows with a 2 in column 5 are subtracted from the average of the output values corresponding to a row with a 1 in column 5. The difference between these two values results in the factor effect. If this is done for each column, and then plotted, the result is the Scree chart shown in Figure 28. Figure 28 shows significant factor effects. Fortunately, Equation 36 can be linearized by taking the natural log of the equation. Using the natural log of the heat transfer as the output, the Scree chart for the interactions is shown in Figure 29, which shows virtually no interaction between factors, so a saturated design can be used.

	1 A	2 B	3 C	4 D	5 AB	6 AC	7 AD	8 BC	9 BD	10 CD	11 ABC	12 ABD	13 ACD	14 BCD	15 ABCD
1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	2
2	1	1	1	2	2	2	1	2	1	1	1	2	2	2	1
3	1	1	2	1	2	1	2	1	2	1	2	1	2	2	1
4	1	1	2	2	2	1	1	1	1	2	2	2	1	1	2
5	1	2	1	1	1	2	2	1	1	2	2	2	1	2	1
6	1	2	1	2	1	2	1	1	2	1	2	1	2	1	2
7	1	2	2	1	1	1	2	2	1	1	1	2	2	1	2
8	1	2	2	2	1	1	1	2	2	2	1	1	1	2	1
9	2	1	1	1	1	1	1	2	2	2	2	2	2	1	1
10	2	1	1	2	1	1	2	2	1	1	2	1	1	2	2
11	2	1	2	1	1	2	1	1	2	1	1	2	1	2	2
12	2	1	2	2	1	2	2	1	1	2	1	1	2	1	1
13	2	2	1	1	2	1	1	1	1	2	1	1	2	2	2
14	2	2	1	2	2	1	2	1	2	1	1	2	1	1	1
15	2	2	2	1	2	2	1	2	1	1	2	1	1	1	1
16	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Figure 27: L16 experiment; the full design includes four factors at two levels each.

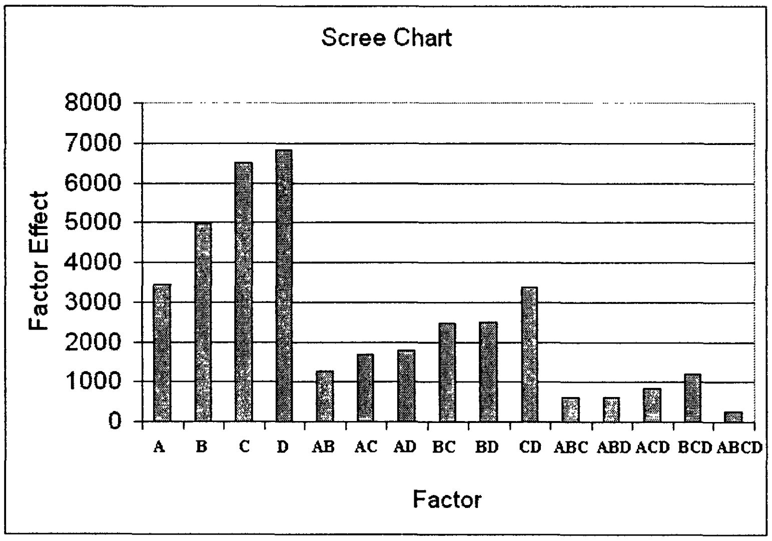


Figure 28: Scree chart for the L16 experiment where the heat transfer (Equation 36) is the measured output. The scree chart indicates that significant factor effects exist.

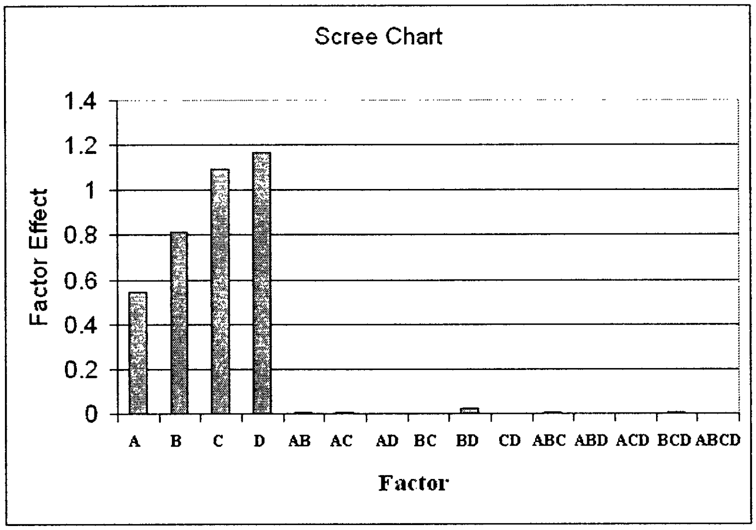


Figure 29: Scree chart for the L16 experiment where the natural log of the heat transfer (Equation 36) is the measured output. The Scree chart indicates that factor effects are negligible compared to single factor effects.

10.5 Choosing the Appropriate Design and Setting up the Experiment

Since for the natural log of the heat transfer interactions between factors are very small compared to the effects of the factors, interactions are ignored. As stated earlier, the preferred number of levels for each factor is three, so the resulting design used is the L9 shown in Figure 14. Since a computer simulation is used to estimate factor effects, each of the parameters that are to be varied is assigned a mean and standard deviation, as shown in Table 2. In order to ensure that the inputs have enough time to accurately portray their given characteristics, 300000 trials were run to obtain the value of the SN ratio. The results of the simulation are given in the next section.

Table 2: Mean and standard deviation of both the control and Noise factors.

	Mean	**Standard Deviation**
k	Levels in Table 1	10% of k
D	Levels in Table 1	10% of D
L	Levels in Table 1	2% of L
v	Levels in Table 1	15% of v
Tinf (K)	280	5.6
Ts (K)	353.15	3.53
μ (kg/m-s)	0.000867	0.0000867
kw (W/m-K)	0.6103	0.061
cp (J/kg-K)	4181	418.1

10.6 Results and Analysis of Data

The L9 experiment is depicted in Figure 30. The SN ratio, mean, and standard deviation for each experiment are listed to the right of the design. Table 3 compares the average values of the SN ratio for each level of each factor. The level for each parameter that produces the highest SN ratio is given in bold; this level for each factor is the optimum level as determined by the design.

	k	**D**	**L**	**v**	**SN**	**Mean**	**St Dev**
1	1	1	1	1	32.312	8.8196	0.21373
2	1	2	2	2	31.403	7.8866	0.2122
3	1	3	3	3	30.654	7.3718	0.2162
4	2	1	2	3	30.616	7.2978	0.21499
5	2	2	3	1	32.171	8.6518	0.21308
6	2	3	1	2	32.575	9.1538	0.2152
7	3	1	3	2	30.854	7.432	0.21301
8	3	2	1	3	32.233	8.8142	0.21554
9	3	3	2	1	32.977	9.4372	0.21183

Figure 30: L9 Experiment and Results

Table 3: Summary of results of the L9 experiment in Figure 30

	1	2	3	Average	SS
K	31.45629	31.78734	**32.0213**	31.754976	0.483558918
D	31.26039	31.93575	**32.06878**	31.754976	1.127302482
L	**32.37323**	31.66523	31.22647	31.754976	2.008819274
V	**32.48657**	31.61074	31.16762	31.754976	2.703075774

Analysis of variance and other statistical techniques are usually run after the experiments have been performed in order to extract more information from the data. One of the most important measurements taken in data analysis is sum of squares which is similar in idea to the Scree chart from before in that it gives a measurement of the importance of the factor and its levels on the outcome of the experiment. The sum of squares for factor A can be computed as

$$SS_A = \sum_i E_i \cdot (m_{Ai} - m)^2 \quad , \tag{37}$$

where $\mathbf{m_{Ai}}$ is the average of the SN ratios for the experiments in which factor A occurs at level **i** (where level **i** can be either 1, 2, or 3 for the L9 in Figure 30), **m** is the average of all the SN ratios as shown in the fifth column in Table 3, and **E** is the number of times the factor level **i** appears in the matrix experiment. In the L9 experiment, each level for each factor appears 3 times. As is apparent from the sum of squares values listed in Table 3, the velocity has the greatest effect on the heat transfer rate of the fins, followed next by fin length and then fin diameter.

10.7 Verification Experiment and Discussion

The final results of the experiment, along with the optimum settings with and without adjustment are shown in Table 4. In comparing the original and the adjusted values, one immediately notes that the standard deviation has not changed (and may have even increased slightly). On first glance it would appear as if the design did not effectively optimize the system. An L16 was subsequently run using the lowest and highest values in Table 1 to see how the optimum settings of the control factors might change if interactions are considered. Since all possible interactions are included in this type of design, heat transfer alone was optimized (no natural log function). The result of the L16 (not shown) was that the optimum values for the control factors were the minimum values tested in the L9 for every factor. This result is in contradiction to the results of the L9 which concluded that the diameter and the thermal conductivity should be increased, not decreased. When comparing the values for the mean and standard deviation after adjustment for both designs, one realizes that no significant improvement has taken place in the variability of the system with either optimization. The designs were adjusted to fractional row numbers of fins, **M**, (an impossible situation) to try to obtain a discrepancy even though none was found.

The conclusion of the experiments is that even for the wide range of input values listed, the non-linearities in the system are too small to make any significant difference through parameter design. The reason that the results of the L9 and L16 are different is simply due to the fact that not enough trials were run to obtain the level of accuracy desired by these results. The mean and standard deviation simply scale. More trials could be run to improve accuracy, but since the standard deviation changes from before to after the optimization by only about 1 %, the improvements are too small to justify the designers changing the values of the fins simply to make them more robust to variation. Within the ranges tested, the control factors cannot improve system robustness, and thus the heater control factors should be set taking other information into account. For instance, an **mL** greater than 4 results in a fin that is too long; the rate of convective heat transfer is high enough that heat is swept away before the length of the fin can be used. Thus the smaller diameter suggested by the L16 design should not be used. The Fin length should be set to the smallest value and the fins should be made with the larger diameter. If lower output variability is required, tolerance design should be used.

Table 4: Results of the L9 experiment for the finned heater

	Original	Optimized	Adjusted
mean	8.232842	10.144549	8.231271
st dev	0.213105	0.214751	0.214629
qout	3848.655	26044.84	3844.039
M	10	10	1.475
SN	31.73911	33.485957	31.67556
k	60	90	90
D(mm)	5.08	9.398	9.398
L(mm)	25.4	12.7	12.7
v	0.75	0.225	0.225
N	9	33	33
m*L	12.252691	2.257547	2.25757
st dev act	826.35827	5651.188	836.7254

11.0 Illustrative Model: Development of a Throttle-by-Wire System

A throttle-by-wire system is considered in the second example. The transient response of the throttle is treated as a smaller-the-better type problem. The main objective of this section is to compare the various sampling methods that can be used to determine mean shift.

11.1 Description of the System

Engines in most automobiles today are of the spark ignition type. The goal of the intake system of a spark ignition engine, as shown in Figure 31, is to control the flow of air into the engine and evenly distribute it to the engine cylinders. The mixing of fuel with the air entering the engine to create a combustible mixture used by the engine to produce power also takes place within the intake system. After passing through the air filter, incoming air moves across a throttle plate before entering the manifold and ultimately the engine cylinders. In modern engines, fuel is introduced into the air-stream right before it enters the cylinders by injectors incorporated into the manifold runners. Spark-ignition engines operate with the ratio of the mixture of fuel to air always at a roughly constant value, so the electronic control unit of the engine varies the amount of fuel injected into the air stream depending on the rate of air flow. The amount of fuel is always proportional to the amount of air, so the engine is controlled is by varying the amount of total mixture that enters the cylinders (in contrast, the diesel engines used in large trucks always have roughly the same amount of air entering the cylinders and vary instead the amount of fuel injected). The amount of mixture is controlled by a butterfly valve (the throttle) which varies the amount of air-flow into the engine depending on the opening angle of the throttle plate. The throttle creates a pressure drop between the outside atmosphere and the cylinders which reduces air flow into the engine. Figure 32 shows a picture of one of these throttle bodies.

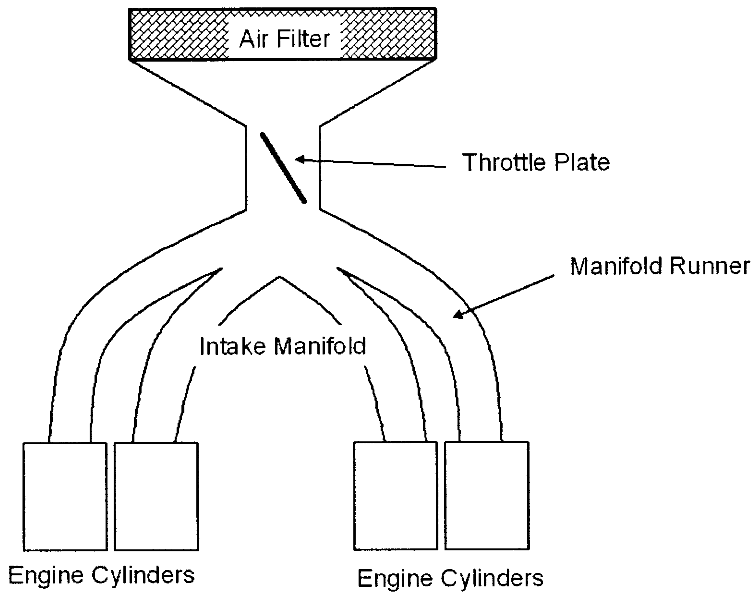


Figure 31: Intake system of a spark ignition engine. Spark Ignition engines are controlled by a throttle which varies the amount of air that enters the engine.

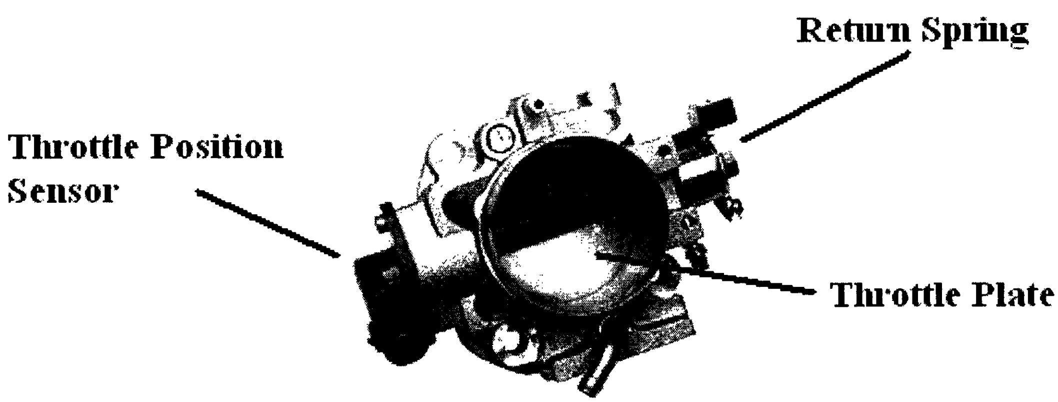


Figure 32: A single-barrel throttle body.

The throttle shown in Figure 32 is mechanical; the opening and closing of the throttle is by means of hardware attached directly to the accelerator pedal of the automobile. Recently, some companies have begun using throttle-by-wire systems which instead use a signal from the accelerator pedal to open the throttle via a motor.

11.2 Development of the Throttle Model

The hardware of the throttle-by-wire system consists of the actual throttle plate, a return spring, a throttle position sensor, a motor, a controller usually incorporated into the engine electronic control unit, and a possible damping mechanism. To improve the accuracy of the system, the controller generally uses a closed loop feedback system (Eriksson and Nielsen, 2000). The characteristic of the controller that makes it closed-loop is that it actively monitors the output of the system and then actively adjusts the system input accordingly. Closed-loop control results in both a faster response as well as the ability to eliminate any offset that would otherwise be seen in the system.

The block diagram for the closed loop controller is shown in Figure 33. The output of the system (the throttle angle) is read by the potentiometer, converted to a signal, and then subtracted from the input signal (that also represents an angle). The difference between these two signals is the error signal, which is then subsequently fed to the controller and ultimately the hardware of the system, which then produces the new output value. The system incorporates a proportional-integral (PI) controller. The proportional control allows for a fast response, and the integral control eliminates the steady-state offset of the throttle.

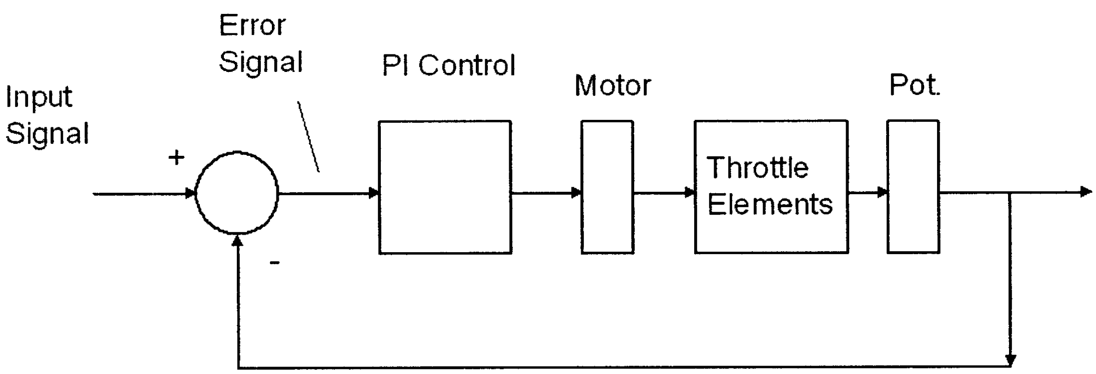


Figure 33: Block diagram of the closed- loop system

The hardware of the mass, spring, and damper system of the throttle can be described by the equation

$$J\cdot\frac{d^2}{dt^2}\theta = F - k_1\cdot(\theta + \theta_0) - k_2\cdot\left(\frac{d}{dt}\theta\right) - T_s\cdot \text{sgn}\left(\frac{d}{dt}\theta\right) \quad , \qquad (38)$$

where **F** is the applied force of the motor, $\mathbf{k_1}$ is the spring constant, $\mathbf{k_2}$ is the coefficient of the viscosity force, $\mathbf{T_s}$ is the static friction, $\boldsymbol{\theta_0}$ is the preset in the spring, **J** is the moment of inertia of the throttle mass, and $\boldsymbol{\theta}$ is the angular output of the system.

The value of **F** is essentially determined by a value related to the input signal to the system. In the closed loop system, **F** is determined by the magnitude of the error signal and the characteristics of the control plant. Since the controller will incorporate proportional control, **F** is simply replaced by the input signal to the motor, **u(t)**. Replacing **F**, and rewriting Equation 38 in the Laplace domain, the resulting transfer function for the throttle system is

$$\frac{\theta(s)}{u(s)} = \frac{1}{J \cdot s^2 + k_2 \cdot s + k_1}, \tag{39}$$

The transfer function of a PI controller can be written in terms of two parameters and is represented also in the Laplace domain as

$$k_3 \cdot \left(1 - \frac{1}{\tau \cdot s}\right), \tag{40}$$

where $\mathbf{k_3}$ and $\boldsymbol{\tau}$ are constants. The schematic of the closed-loop system incorporating the transfer functions is shown in Figure 34. The final transfer function for the system in is

$$\frac{\theta(s)}{\Theta(s)} = \frac{k_3 \cdot \left(1 + \frac{1}{\tau \cdot s}\right)}{k_3 \cdot \left(1 + \frac{1}{\tau \cdot s}\right) + J \cdot s^2 + k_2 \cdot s + k_1}, \tag{41}$$

where $\mathbf{\Theta(s)}$ is the desired throttle angle (the system input) in the Laplace domain. Note that Equation 41 has simplified Equation 38. Equation 41 ignores the effect of $\mathbf{T_s}$. To account for $\mathbf{T_s}$ and the sign function, the equations are not ultimately solved in the Laplace domain, however this domain does provide a clean description of the system, which is why it has been included.

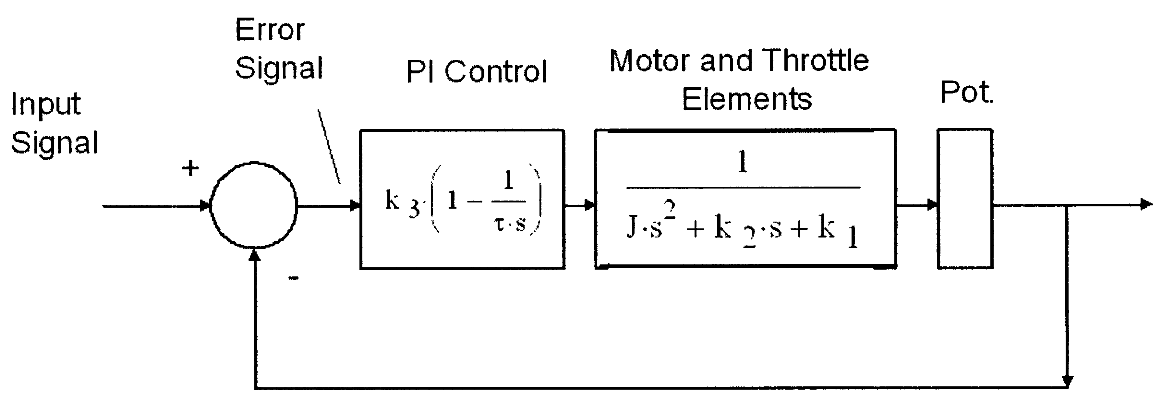


Figure 34: Closed-loop block diagram for the throttle-by-wire system. The transfer functions for the system components are included in the diagram.

The aspects of the system not shown in Equation 41 but incorporated into the final model deal with the static friction and the modeling of the input to the system. If the motor force and the spring force together are less than the static friction, and if the velocity equals zero, then the sum of the forces on the system equals zero; essentially the throttle stops because static friction is now greater than the forces driving the system. This term can be represented by the extra term in Equation 38. The throttle input signal is also not modeled as instantaneous, which is in contrast to a model developed in the references and on which this model is based (Eriksson and Nielsen, 2000). When a driver depresses the throttle pedal with the desire to reach nearly wide open throttle, that movement does not happen instantaneously, so the signal to the electronic throttle is not instantaneous. For fast throttle opening times, opening might occur within about 0.1 seconds. Thus the throttle opening is modeled as

$$\Theta(t) = \left(1 - \exp\left(\frac{-t}{T_{constant}}\right)\right) \cdot \Theta_{set} , \tag{42}$$

where $\mathbf{\Theta(t)}$ is the desired throttle angle now described in the time domain, $\mathbf{\Theta_{set}}$ is the steady-state angle of the throttle, $\mathbf{t}$ is time, and $\mathbf{T_{constant}}$ is a time constant.

11.3 Optimization Parameter

A particular throttle response may look like that shown in Figure 35. The broken curve is the angular input of the accelerator pedal, and the solid curve represents the throttle angle as a function of time. The ideal throttle angle is equal to the accelerator angle. Thus, ideally the solid and broken curves would follow each other. The loss to the system is thus the difference between these two curves. If we add up that distance at ever point, or calculate the total area between these two curves, as shown in Figure 36, this area is a good representation of the deviation of the throttle behavior from the ideal operation. Thus, the shaded area in Figure 36 is used as the optimization parameter in the design of experiments, where the goal is to make the shaded area as small as possible.

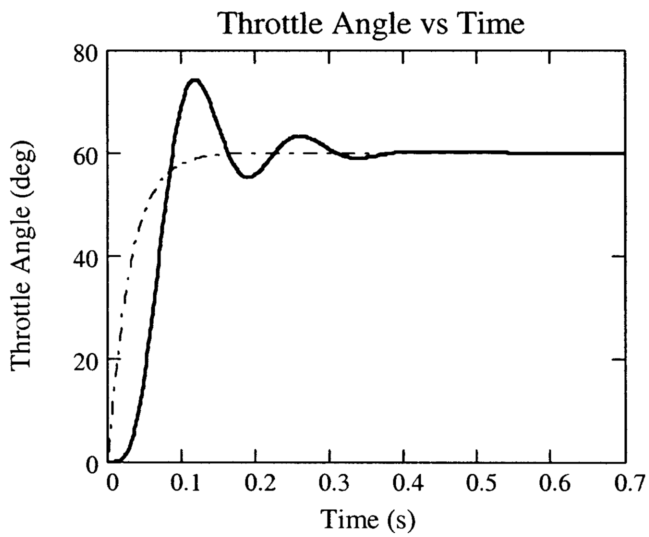


Figure 35: Typical response of the throttle valve. The solid curve is the response of the throttle. The broken line is the input produced by the acceleration pedal.

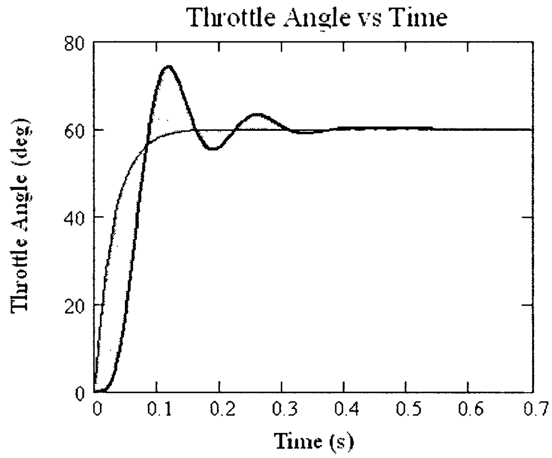


Figure 36: The area between the throttle response and the input curve must be made as small as possible in order to achieve optimum throttle conditions.

11.4 Test for Factor Interactions

The system is tested for 4 factors at 2 levels each, so again the L16 matrix is incorporated. The 4 factors chosen to be considered initially are $\mathbf{k_1}$, $\mathbf{k_2}$, $\mathbf{k_3}$, and $\boldsymbol{\tau}$. The interaction of these factors is shown in Figure 37. $\boldsymbol{\theta_0}$ is the other possible control factor, but for now, it is set to a value of 5 degrees, and only the interaction between the other control factors is considered. The factors other than those just mentioned are not changeable, so they are not included in the control factors (but will be considered as noise factors later). As is apparent from the Scree chart, the interactions cannot be ignored. Unlike with the heater example, the system model for the throttle cannot be linearized by taking the natural log of the output. Thus a full design must be used.

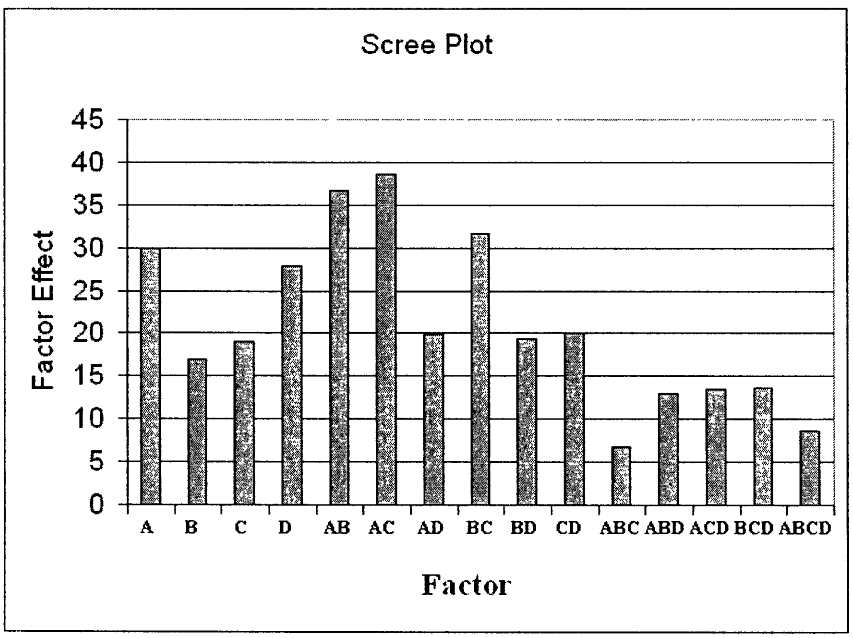


Figure 37: Interactions between $\mathbf{k_1}$, $\mathbf{k_2}$, $\mathbf{k_3}$, and $\boldsymbol{\tau}$.

11.5 Optimization

In addition to $\mathbf{k_1}$, $\mathbf{k_2}$, $\mathbf{k_3}$, and $\boldsymbol{\tau}$, $\boldsymbol{\theta_0}$ is also considered as a factor to optimize. Thus a matrix experiment incorporating an L32 matrix, with 5 factors at 2 levels each where all interactions can be considered is used. The SN ratio is essentially the same as shown in Equation 9, except that the square is ignored. The average of the optimizing parameter for each level of each factor is calculated (just as in Table 3). A sum of squares was also determined for each level. At the end of each experiment, the optimum factor levels replaced the previous levels in the experiment, and the tests were run again. The optimum level of both the spring, $\mathbf{k_2}$, and the controller parameter, $\mathbf{k_3}$, (whose value is largely characteristic of the size of the motor) were capped at the values in Table 5. The larger the motor and spring, the faster the response and the less the inertia of the mass affects the response. However the motor cannot be too large for cost, weight, and size issues. The optimum values of the constants after about 20 runs are shown in Table 5. Figure 38 is a trace of the optimum response.

Table 5: Optimum settings for the throttle-by-wire system.

Factors	Setting at Optimum
k1 (N-m)	1.8
k2 (kg-m^2/s)	0.0435
k3	0.7
T	0.0114
θ zero (degrees)	1.3525
J (kg-m^2)	7.32E-04
Θset (degrees)	60
Tconstant (s)	0.03
Ts (N-m)	0.0075

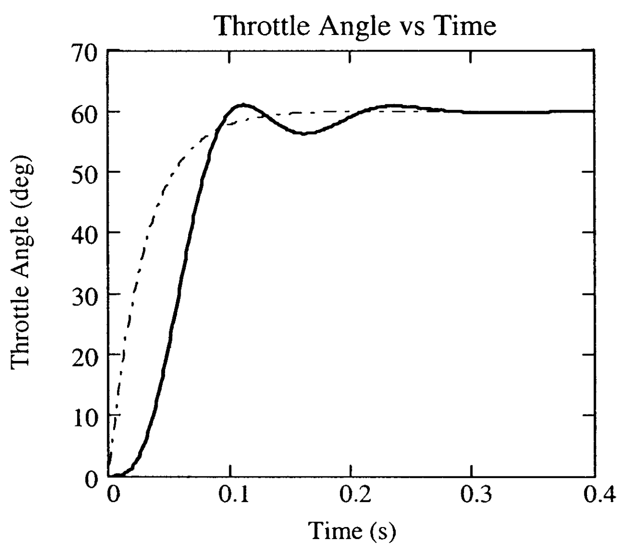


Figure 38: Optimum response for the throttle-by-wire system where the motor and spring sizes have been capped.

11.6 Calculation of Mean Shift

Each factor of the throttle system was assigned a standard deviation listed in Table 6. The Monte Carlo, Latin Hypercube, 2n+1 pattern, and 4n+1 pattern were all then used to predict the mean shift of the throttle from its ideal response. For the Monte Carlo method, a total of 40000 trials were tested. For the Latin Hypercube method, 60000 trials were performed. The trials were done in batches, or runs, of 100, 200, 400, 1000 and 2000 trials each to try to obtain a sense of the quickness with which the mean of the data output converges.

Table 6: The settings of the control and noise factors and the standard deviation for each factor.

Control Factors	Setting at Optimum	Std. Dev.
k1 (N-m)	1.8	0.18
k2 (kg-m^2/s)	0.0435	0.00435
k3	0.7	0.07
Tau	0.0114	0.00114
Theta zero (degrees)	1.3525	0.13525
Noise Factors	**Settings at Optimum**	**Std. Dev.**
J (kg-m^2)	7.32E-04	7.32E-05
Theta set (degrees)	60	6.00E+00
Tconstant (s)	0.03	0.0015
Ts (N-m)	0.0075	0.00075

11.7 Results and Discussion

Figure 39 compares the results of the Monte Carlo to the Latin Hypercube method. The graph is a plot of the mean output value of the system for a specific run versus the number of trials sampled within that run. The Monte Carlo runs, which are represented by the dots in the figure, exhibit more scatter than the Latin Hypercube results, especially for the runs that included low numbers of trials. As the run sizes become larger, the two appear to converge, even though the Latin Hypercube results still appear to have slightly less scatter.

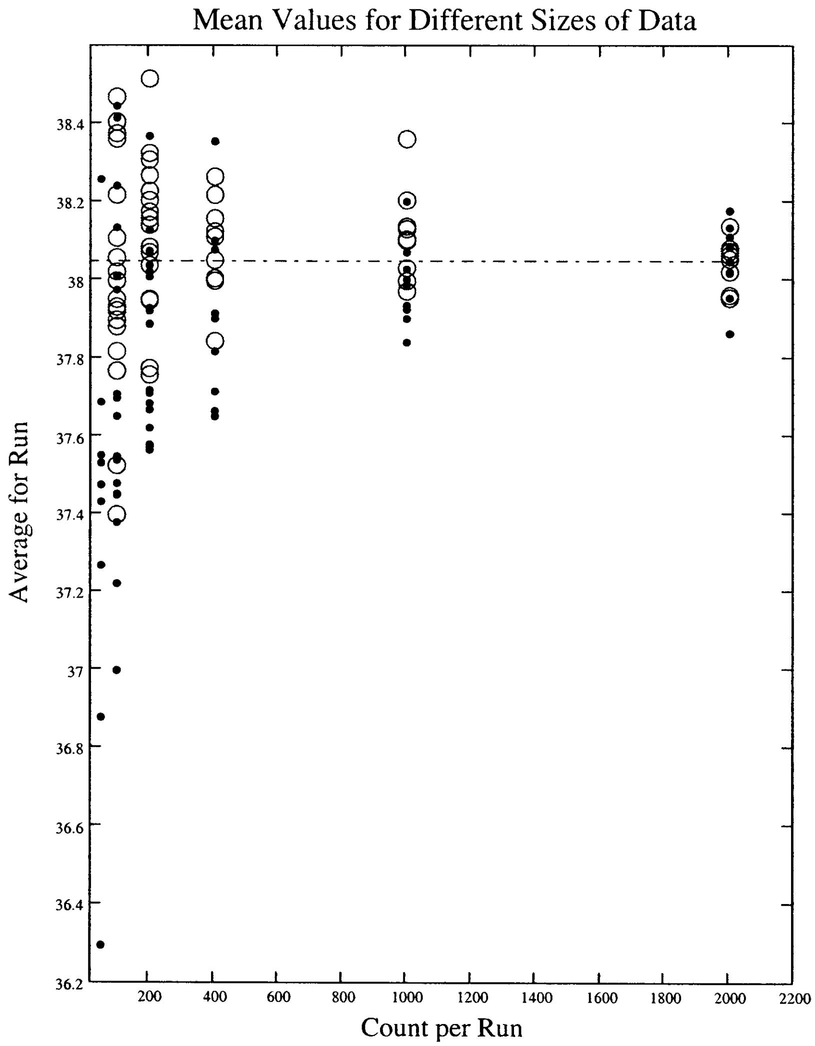


Figure 39: Average output for each run versus the number of trials in the run. The Latin Hypercube data is represented by the circles. The Monte Carlo data is shown as dots.

Figures 40 is a plot of the standard deviations for each of the runs shown in Figure 39. The run numbers are low, around 10 to 20 each, so the standard deviations are by no means exact. However Figure 40 does noticeably illustrate that the data sets of the Latin Hypercube method are less scattered, especially for the lower trial counts. As a final comparison between these two methods, Figures 41 and 42 are histograms for the 2000 trial size runs for both the Monte Carlo and Latin Hypercube methods. The histogram for the Latin Hypercube appears slightly more ordered than the Monte Carlo results. From Figures 39 through 41, it is reasonable to conclude that the Latin Hypercube offers an advantage over the Monte Carlo method in predicting the output of the system, especially for the smallest trial sizes.

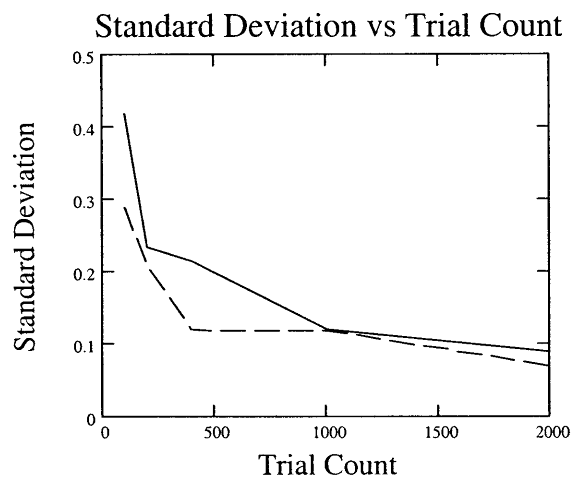


Figure 40: Average standard deviation for all runs at the same trial count. The Monte Carlo standard deviations are represented by the solid line. The Latin Hypercube data is represented by the dashed line.

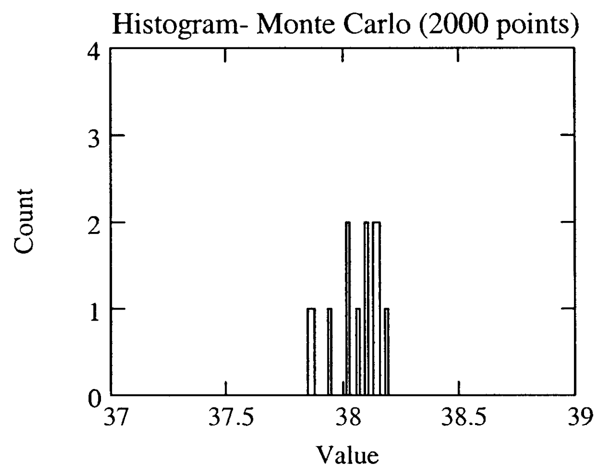


Figure 41: Histogram of the runs for the Monte Carlo sampling 2000 trials per run.

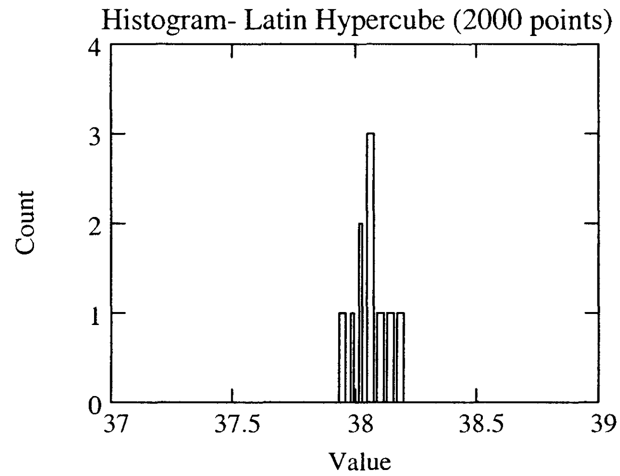


Figure 42: Histogram of the runs for the Latin Hypercube sampling 2000 trials per run.

Finally, Table 7 compares the results of the four sampling methods. The value of the mean shift from the original value is listed for each sampling method as well as the percent difference between each sampling method and the Monte Carlo method. The 2n+1 pattern consequently under-predicts the actual final value by about 0.29%, and the 4n+1 pattern under-predicts the actual value by about 0.05%.

As mentioned previously, one of the major uses of these sampling methods is in the robust design of a maximizing or minimizing system. This experiment was performed for essentially one set of values for the system. However, if the variability was to actually be determined for the robust design method, many experiments would have been run. The 4n+1 pattern took about 20 minutes to set up and run and the 2n+1 pattern took about 10 minutes. In contrast, the Hypercube and Monte Carlo methods took about an hour and a half each to run all of the trials. Thus for a system described by a complex differential equation, if a slight inaccuracy in the data can be afforded, the star patterns have a huge advantage over the other techniques.

Table 7: Comparison of the final values predicted by the various sampling methods.

	Prediction of Method	Value of Mean Shift	Diff. from Monte Carlo Prediction
Monte Carlo	37.9977	2.9367	0
Latin Hypercube	38.0687	3.0077	0.187 %
2n+1	37.8868	2.8258	-0.292 %
4n+1	37.98	2.919	-0.047 %
Value Before Shift	**35.061**		

12.0 Conclusion

The benefits of the techniques described in the last several sections are readily apparent from the experiments and the discussions in these sections. The robust design method when applied correctly can increase quality substantially while still reducing costs. Quality distinguishes the vertical integration of the product, so applying this method and increasing quality can have significant benefits to a company's success in the market place. The efficiency of design of experiments is also apparent, both when applied to robust design as well as when used on its own to increase productivity and results when working on complicated systems with many inputs. The point is that the effect of the variability on system performance can be quantified, and needs to be in order to improve quality as well as accurately determine system performance (such as when a mean shift occurs).

The designs involved in robust design can become complicated, and some engineers focus their careers on the robust design techniques and the statistics behind them. However, the main point to get across from the ideas presented is that accounting for and manipulating variability by using simple statistical techniques is a relatively easy and straightforward process that results in significant benefits to a company even when applied at the more elementary level. For the design of a system or process that is either known to have variability or is exhibiting variability in the early or prototype stages of the design, robust design is a necessary tool that the engineers should be aware of. An engineer should never be improving the quality of the design by tightening the tolerances without having first tried to robust design the system.

References

1. Church, Jeffrey and Roger Ware. Industrial Organization. Boston, MA: McGraw-Hill, 2000.
2. Logothetis, N. and H.P. Wynn. Quality Through Design: Experimental Design, Off-line Quality Control and Taguchi's Contributions. Oxford, England: Clarendon Press, 1989.
3. Ulrich, Karl T. and Steven D. Eppinger. Product Design and Development. Boston, Massachusetts: McGraw-Hill, 2000.
4. Phadke, Madhav S. Quality Engineering Using Robust Design. Englewood Cliffs, New Jersey: Prentice Hall, 1989.
5. Del Vecchio, R.J. Understanding Design of Experiments. Cincinnati, Ohio: Hanser/ Gardner Publications, 1997.
6. Heywood, John B. Internal Combustion Engine Fundamentals. New York, New York: McGraw-Hill, 1988.
7. Lienhard, John H. V and John H. Lienhard IV. A Heat Transfer Textbook. Cambridge, Massachusetts: Phlogiston Press, 2003.
8. Eriksson, Lars and Lars Nielsen. Non-Linear Model-Based Throttle Control. SAE Paper No. 2000-01-0261, 2000.
9. Palm, William J. Modeling, Analysis, and Control of Dynamic Systems. New York, New York: John Wiley and Sons, 2000.