

EVALUATING ROBUST DESIGN METHODS USING A MODEL OF INTERACTIONS IN COMPLEX SYSTEMS

Daniel D. Frey

Assistant Professor of Mechanical Engineering and Engineering Systems
Massachusetts Institute of Technology

Xiang Li

Graduate Research Assistant
Massachusetts Institute of Technology

Abstract. Uncertainty, robustness, and flexibility are three of the issues widely viewed as central to the field of complex engineered systems. A key challenge is to make the consideration of these issues more quantitative and theoretically sound. This paper presents a detailed technical analysis of one question inter-relating all three issues. How can we quantify the benefits of flexibility in the design processes used to seek robustness to uncertainty? To address this question, an approach is presented for evaluating robust design methods via probabilistic simulation. Different robust design methods are repeated on a large number of systems and the results are analyzed statistically. A key to the approach is appropriate modeling of interactions among the variables in complex engineering systems. The weak heredity model is used to generate simulated systems with the properties of effect sparsity, hierarchy, and inheritance. A case study is presented in which different combinations of factorial designs and adaptive plans are used in robust design experiments. It is shown that a combination of an adaptive inner array (a flexible plan) with a fractional factorial outer array (an efficient but inflexible plan) provides the best results. Thus, the case study provides specific operational advice for deploying flexibility in one particular systems engineering task. Because of this bit of progress, the new evaluation technique is proposed one example of rigorous quantitative treatment of foundational issues in complex engineering systems made possible by combining advances in computing power and statistical theory.

1. Introduction

Uncertainty is an increasingly important consideration in the design of large-scale engineering systems. As systems include more components, interconnections, and variables, the effects of uncertainties (unless carefully managed) tend to accumulate and may lead to unacceptable risks. One effective countermeasure is to reduce the sensitivity of the design to uncertain or randomly varying factors (a.k.a. noise factors). Robust design is a set of engineering methods widely successful in reducing sensitivity to such noise factors as customer use conditions, manufacturing variability, and degradation of a system over time. Therefore, robust design is an important discipline for complex engineering systems and research to improve robust design methods is a significant opportunity.

One widely known means to accomplish greater robustness is through parameter design. Within a given system architecture, the designer searches the space of controllable parameters seeking lower

sensitivity to noise factors. Effective parameter design requires coordination of two activities -- estimating sensitivity to noise factors and searching the space of design variables to seek reduced sensitivity. Many different robust design methods have been proposed including Taguchi methods (Taguchi, 1987) and response modeling (Welsch et. al., 1990; Shoemaker, Tsui, and Wu, 1991). Recently, Frey, Engelhardt, and Greitzer (2003) proposed that an adaptive one-factor-at-a-time technique will, under some conditions, provide large improvements with an economy of run size due to its flexibility and adaptability.

The controversies regarding choice of robust design methods have been divisive. Many practitioners strongly defend the use of Taguchi Methods on the basis of their effectiveness in practice. Many other practitioners and scholars offer persuasive arguments for alternative techniques based primarily on statistical theory. The purpose of this paper is to offer an approach to evaluating robust design methods so that the choice among alternatives can be tied directly to their effectiveness in applications while still maintaining consistency with statistical theory.

Sections 2 and 3 review selected technical background on Design of Experiments and Robust Design. Section 4 describes the proposed simulation-based approach for evaluating robust parameter design methods. Section 5 comprises a case study in which the approach is applied to evaluate combinations of factorial and adaptive plans. Section 6 presents conclusions and suggests future directions for research.

2. Background on Design of Experiments

Design of experiments (DOE) is a body of knowledge and techniques for planning a set of experiments, analyzing the resulting data, and drawing conclusions from the analysis (Wu and Hamada 2000). R. A. Fisher introduced the full factorial experiment in which every combination of factors and levels are tested. Table 1 depicts the 2^3 factorial experiment which is used for systems with three factors each having two levels (which are coded here as +1 and -1). If the effects of the experimental factors obey the superposition principle, a full factorial experiment provides the same precision of effect estimation “as if the whole experiment were dedicated to a single factor” (Fisher 1949). If the effects of the control factors are not simply additive, the full factorial design enables the experimenter to estimate all the interactions among the factors. A significant disadvantage of the full factorial design is that the number of experiments required rises geometrically with the number of factors.

TABLE 1. A full factorial design 2^3

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

In order to conduct experiments more efficiently, fractional factorial experiments may be employed which retain the desirable properties (e.g., balance and orthogonality) of the full factorial, but reduce the size of the experiment. For example, consider an experimental scenario for estimating the effects of seven factors each having two discrete levels using eight experiments. Given this scenario, the 2^{7-4} design (Table 2) is D-optimal for fitting a first-order model -- the design minimizes the volume of the ellipsoidal confidence region of the effect estimates.

The economy of run size afforded by the fractional factorial design comes at a cost. In the 2^{7-4} fractional factorial design, each main effect is clear of other main effects, but is aliased with two-factor interactions. Designs with this property are called resolution III. A design constructed so that main effects are clear of two-way interactions, is said to have resolution IV. Higher resolution requires more experimental runs or fewer experimental factors. For example, a 2^{4-1}_{IV} can be constructed by striking out columns C, E, and F from Table 2. Alternatively, a 2^{7-3}_{IV} can be formed by “folding over” which involves adding eight more runs to the 2^{7-4}_{III} that have the opposite settings as the first eight runs.

TABLE 2. A fractional factorial design 2^{7-4}

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

The effectiveness of the fractional factorial designs rests on three principles which have been formulated by the statistics community based on a large body of experience:

- 1) Sparsity of effects -- among several experimental effects examined in any experiment, a small fraction will usually prove to have a statistically significant effect.
- 2) Hierarchy -- main effects are generally larger than two-factor interactions, two-way interactions are larger than three-factor interactions, and so on.
- 3) Inheritance -- an interaction is more likely to be active when its “parent” factors are active.

In addition to the simple experimental plans discussed in this section, many other experimental plans have been developed and are used widely by industry to improve products and processes. The purpose of this paper is to articulate a method for evaluating different experiments for use in robust design. The paper therefore focuses on just a few simple experimental designs.

3. Background on Robust Design

Taguchi pioneered robust parameter design in which products and processes are made less sensitive to noise due to “noise factors” such as manufacturing variations, customer use, and degradation over time (Taguchi 1987; Phadke 1989). The decreased sensitivity to noise is accomplished by systematically searching a space of controllable design variables or “control factors” while deliberately inducing variations in the noise factors. In Taguchi Methods of robust design, the changes in both the control and noise factors are accomplished with “orthogonal arrays”, a tabulated set of fractional factorial designs including the 2_{III}^{7-4} which Taguchi called the L8. Taguchi advocates the use of “crossed arrays” in which one orthogonal array is selected for the control factors (the inner array), another orthogonal array is selected for the noise factors (the outer array). For each row of the inner array, the entire outer array of noise is carried out and is used to compute a signal to noise ratio. The effects of the control factors on the signal to noise ratio are analyzed and used in an effort to maximize robustness of the system. Taguchi methods emphasize the definition of an “ideal function” of the system and the use of dynamic signal to noise ratio wherever possible in order to measure robustness across a range of signal inputs to the system.

Classical DOE has also been adapted for robustness optimization (Logothetis and Wynn 1994; Wu and Hamada 2000). There are several key differences between Taguchi methods and the methods suggested by the statistics community. The use of signal to noise ratios is generally discouraged. Instead classical approaches emphasize modeling of mean and dispersion effects or direct estimation of control by noise interactions. The use of a single array including both control and noise factors is often recommended.

This paper undertakes an evaluation of a recently proposed concept which deploys greater adaptability (flexibility) in design of experiments. Frey, Engelhardt, and Greitzer (2003) showed that an adaptive one factor-at-a-time (OFAT) plan can, under some conditions, be more effective for optimizing a system’s response than a fractional factorial design. Specifically, an OFAT plan provides more improvement when the experimental error is less than 40% of the factor effects or interactions are more than 25% of all factor effects. This result was demonstrated on 66 response variables cataloged from journals across a wide range of engineering system. The results had not been applied to problems of robust design because the sort of full factorial data sets needed are not widely available. The development of the new technique presented in section 4 made possible the first rigorous evaluation of the adaptive OFAT approach as applied to robust design (presented in section 5).

4. An Approach for Evaluating Robust Design Methods

This section describes an approach employing computational simulation to compare the effectiveness of robust design methods. The approach described here relies on statistical characterization of the robust design methods by probabilistic simulation. Therefore, each robust design method must be repeated on a large enough number of individual systems to determine its average performance and the variations about the average. A key to the approach is appropriate modeling of interactions among the system input variables. The approach has four main steps:

- 1) Instantiate models of multiple engineering systems that will be subject to robust parameter design.
- 2) For each system, simulate every robust design method to determine the control factor settings preferred according to that method.
- 3) For each system/method pair, perform a confirmation experiment to determine the actual variance of the response at the chosen control factor settings.

- 4) For each method, analyze the data across all instantiated systems to determine the mean improvement and inter-quartile range for the improvement.

Each of these steps requires substantial elaboration as presented in the following subsections.

4.1 A MODEL OF INTERACTIONS IN ENGINEERING SYSTEMS

In order to carry out the approach described herein, it is necessary to have a large number of engineering systems which are subject to robust design experiments. One approach would be to create a database of thousands of simulations of actual engineering systems. However, most simulations used to design large-scale engineering systems require hours, days, or more to produce realistic estimates of system performance at a single design point. The method described here requires thousands of such performance estimates for each system. A much faster approach is needed for the present purpose.

Another approach is to assemble a collection of data from experiments on a large number of authentic engineering systems. This approach was used in previous evaluations of robust design methods (Frey, Engelhardt, and Greitzer, 2003). However, this approach did not allow enough flexibility for the present purposes or enable a study of systems with adequate scale. Systems with seven or more factors are rarely documented in adequate detail to enable the kind of approach presented here.

Motivated by the considerations described previously, we developed a way to create multiple instances of computationally efficient simulated systems which are realistic in certain critical regards. In particular, the interactions between control factors and noise factors must be modeled realistically because these provide opportunities for improved robustness. Other system interactions such as those among noise factors are also important because they create difficulties that may adversely affect the robust design methods being evaluated. Thus, for the present purpose, the structure of all interactions must be modeled in an appropriate way. As discussed in Section 2, actual engineering systems are known to exhibit sparsity of effects, hierarchy, and inheritance. These properties are therefore central to the model proposed here.

The model of engineering systems used in this paper is an adaptation (and extension) of a hierarchical prior probability model called the weak heredity model or WH model developed by Chipman, Wu, and Hamada (1997). The model used in the proposed method expressed in Equations 1 through 10 below.

$$y(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \sum_{\substack{j=1 \\ j>i}}^n \beta_{ij} x_i x_j + \sum_{i=1}^n \sum_{\substack{j=1 \\ j>i}}^n \sum_{\substack{k=1 \\ k>j}}^n \beta_{ijk} x_i x_j x_k + \varepsilon \quad (1)$$

$$x_i \sim NID(0, w_1) \quad i \in 1 \dots m \quad (2)$$

$$x_i \in \{+1, -1\} \quad i \in m+1 \dots n \quad (3)$$

$$\varepsilon \sim NID(0, w_2) \quad (4)$$

$$f(\beta_i | \delta_i) = \begin{cases} N(0, 1) & \text{if } \delta_i = 0 \\ N(0, c) & \text{if } \delta_i = 1 \end{cases} \quad (5)$$

$$f(\beta_{ij} | \delta_{ij}) = \begin{cases} N(0, s_1) & \text{if } \delta_{ij} = 0 \\ N(0, c \cdot s_1) & \text{if } \delta_{ij} = 1 \end{cases} \quad (6)$$

$$f(\beta_{ijk} | \delta_{ijk}) = \begin{cases} N(0, s_2) & \text{if } \delta_{ijk} = 0 \\ N(0, c \cdot s_2) & \text{if } \delta_{ijk} = 1 \end{cases} \quad (7)$$

$$\Pr(\delta_i = 1) = p \quad (8)$$

$$\Pr(\delta_{ij} = 1 | \delta_i, \delta_j) = \begin{cases} p_{00} & \text{if } \delta_i + \delta_j = 0 \\ p_{01} & \text{if } \delta_i + \delta_j = 1 \\ p_{11} & \text{if } \delta_i + \delta_j = 2 \end{cases} \quad (9)$$

$$\Pr(\delta_{ijk} = 1 | \delta_i, \delta_j, \delta_k) = \begin{cases} p_{000} & \text{if } \delta_i + \delta_j + \delta_k = 0 \\ p_{001} & \text{if } \delta_i + \delta_j + \delta_k = 1 \\ p_{011} & \text{if } \delta_i + \delta_j + \delta_k = 2 \\ p_{111} & \text{if } \delta_i + \delta_j + \delta_k = 3 \end{cases} \quad (10)$$

Equation 1 represents the measured response of the engineering system y whose mean and variance might be optimized in robust design. The independent variables x_i are both control factors and noise variables; no distinction is made in this notation except via the indices. Equation 2 shows that the first set of x variables (x_1, x_2, \dots, x_m) are regarded as “noise factors” and are assumed to be normally distributed. Equation 3 shows that the other x variables ($x_{m+1}, x_{m+2}, \dots, x_n$) are the “control factors” which are assumed to be two level factors. The variable ε represents the pure experimental error in the observation of the response which was assumed to be normally distributed. Since control factors are usually explored over a wide range compared to the noise factors, the parameter w_1 is included to set the ratio of the control factors range to the standard deviation of the noise factors. The parameter w_2 is included to set the ratio of the standard deviation of the pure experimental error to the standard deviation of the noise factors.

The response is a third order polynomial in the independent variables x_i . The coefficients β_i are the main effects. The coefficients β_{ij} model two-way interactions including control by noise and noise by noise interactions. Similarly, the coefficients β_{ijk} model three-way interactions including control by control by noise and control by noise by noise interactions. The weak heredity model originally proposed by Chipman, Wu and Hamada (1997) did not include three-way interaction effects, but their addition was useful for the present investigation.

The values of the coefficients are determined by a random process that models the properties of effect sparsity, hierarchy, and inheritance. Equation 5 determines the probability density function for the first order coefficients. Factors can be either active or inactive depending on the value (0 or 1 respectively) of their corresponding parameters δ_i . The parameter strength of active effects is assumed to be c times that of inactive effects. Similarly, Equations 6 and 7 determine the probability density function for the second order and third order coefficients respectively. In Equations 3 and 4, the hierarchy principle is reflected in the fact that second order effects are only s_1 times as strong (on average) as first order effects ($s_1 < 1$) and third order effects are only s_2 times as strong as second order effects.

Equation 8 enforces the sparsity of effects principle. There is a probability p of any main effect being active. Equations 9 and 10 enforce inheritance. The likelihood of any second order effect being active is low if no participating factor has an active main effect and is highest if all participating factors have active main effects. Thus generally one sets $p_{11} > p_{01} > p_{00}$ and so on.

4.2 SIMULATE EVERY ROBUST DESIGN METHOD ON EACH SYSTEM

Every robust design method to be evaluated must be encoded as an automatic procedure for sending inputs to the engineering system. Each set of inputs sent to a system represents a simulated experiment

within a robust design method. Every robust design method to be evaluated must also include a description of how data are analyzed and how the “optimized” control factor settings are determined. The exact procedure used in this step depends very strongly on the particular method being simulated. Section 5.2 provides several examples.

4.3 PERFORM CONFIRMATION EXPERIMENTS

In robust design, any particular method will result in a set of control factor setting which are predicted to be optimal. In our notation these, “predicted optimal” control factor settings are a vector of values at coded levels $x_i \in \{+1, -1\}$ where $i \in m+1 \dots n$. Most methods provide a prediction of a robustness measure at the optimized condition. The predicted optimum is not necessarily the actual optimum, therefore good engineering practice requires a confirmation experiment in which the robustness of predicted optimal design is evaluated experimentally. Within the computational approach proposed here, the confirmation experiment can be simulated by computing the exact solution for the mean response μ and the standard deviation of the response σ . The variance of the response in the WH model is due to the variance in noise factors which are a subset of the x_i and also the pure experimental error ε . Given the response is a polynomial (Equation 1), and given the assumption that the input variables 1 through k are independent standard normal variables, it can be shown that

$$\mu(x_{k+1}, x_{k+2}, \dots, x_n) = \sum_{i=m+1}^n \beta_i x_i + \sum_{i=m+1}^n \sum_{\substack{j=m+1 \\ j>i}}^n \beta_{ij} x_i x_j + \sum_{i=m+1}^n \sum_{\substack{j=m+1 \\ j>i}}^n \sum_{\substack{k=m+1 \\ k>j}}^n \beta_{ijk} x_i x_j x_k \quad (11)$$

and

$$\sigma^2(x_{m+1}, x_{m+2}, \dots, x_n) = w_1^2 + \sum_{i=1}^m \left[\beta_i + \sum_{\substack{j=m+1 \\ j>i}}^n \beta_{ij} \cdot x_j + \sum_{\substack{j=m+1 \\ j>i}}^n \sum_{\substack{k=m+1 \\ k>j}}^n \beta_{ijk} \cdot x_j \cdot x_k \right]^2 + \sum_{i=1}^m \sum_{\substack{j=1 \\ j>i}}^m \left[\beta_{ij} + \sum_{\substack{k=m+1 \\ k>j}}^n \beta_{ijk} \cdot x_k \right]^2 + \sum_{i=1}^m \sum_{\substack{j=1 \\ j>i}}^m \sum_{\substack{k=1 \\ k>j}}^m \beta_{ijk}^2 \quad (12)$$

There are many ways to use the computed mean and standard deviations to form a measure of robustness. In Taguchi Methods, robustness is often measured using the Type I signal to noise ratio $20\log(\mu/\sigma)$ used when the standard deviation is a function of the mean. An alternative is the Type II signal to noise ratio $20\log(1/\sigma)$ used when the standard deviation is a independent of the mean. Other researchers in robust design suggest different measures of robustness including the simplest option, the standard deviation σ . The standard deviation is the most widely understood measure among practicing engineers, so it will be used for the remainder of the paper.

4.4 ANALYZE THE DATA AND COMMUNICATE

The data generated in step 3 of the approach can be analyzed and displayed in many different ways. To choose an approach, it is useful to consider how the data might be used. The designer must choose

among the alternative methods based on an *a priori* estimate of the system structure and behavior. Let us assume that this *a priori* estimate is captured by the weak heredity model (Equations 1-10). The designer seeks to balance the total cost of carrying out the method against the expected increase in value of the product due to reduced variability. The value of the product rises monotonically as variance in the response is reduced. This is consistent with the concept of a quadratic loss function (Phadke 1989; Taguchi 1987). In comparing one system to another, it is assumed that a given percentage decrease in variance from the average starting point design is equally valued across all simulated systems. The designer also seeks to assess the risks attendant in any given method. If this description characterizes the design scenario adequately, then a tabular presentation of the alternatives methods with the expected percent reductions in variance and the number of experiments should be the primary basis for choice of method. The need to assess risks suggests that the variability of the outcomes of the method across a population of systems should also be presented. An inter-quartile range is useful for communicating such variability. Examples of such a presentation of data are provided in Section 5.4.

5. Example – Assessing the Effectiveness of Factorial and Adaptive Experimental Plans

5.1 SELECTING PARAMETERS OF THE WEAK HEREDITY MODEL

The weak heredity (WH) model described in section 4.1 was used to create the simulated engineering systems for this case study. The systems in this case study all have three noise factors and seven control factors, therefore $m=3$ and $n=10$. These values were chosen because they represent a reasonable number of factors that might be considered in industrial practice of robust design. They also allow several variants of single array plans to be evaluated. This choice results in a potentially very complex model including 175 coefficients. Of these, 120 coefficients represent three-way interactions, but the vast majority of these will be small due to the assumption of hierarchy and inheritance.

The WH model has several real valued parameters which may have a significant effect on the inferences drawn from its use. To provide a balanced view, six different sets of parameter settings were used as shown in Tables 3 and 4.

TABLE 3. Sets of model parameters considered in the case study

	c	s_1	s_2	w_1	w_2
Basic WH	10	1	1	1	1
Basic low w	10	1	1	0.1	0.1
Basic 2 nd order	10	1	0	1	1
Fitted WH	15	1/3	2/3	1	1
Fitted low w	15	1/3	2/3	0.1	0.1
Fitted 2 nd order	15	1/3	0	1	1

TABLE 4. Additional model parameters for each set considered in the case study

	p	p_{11}	p_{01}	p_{00}	p_{111}	p_{011}	p_{001}	p_{000}
Basic WH	0.25	0.25	0.1	0	0.25	0.1	0	0
Basic low w	0.25	0.25	0.1	0	0.25	0.1	0	0
Basic 2 nd order	0.25	0.25	0.1	0	N/A	N/A	N/A	N/A
Fitted WH	0.43	0.31	0.04	0	0.17	0.08	0.02	0
Fitted low w	0.43	0.31	0.04	0	0.17	0.08	0.02	0
Fitted 2 nd order	0.43	0.31	0.04	0	N/A	N/A	N/A	N/A

The basic weak heredity model (basic WH) is based on the parameters used in Bayesian model selection (Chipman, Wu, and Hamada, 1997). Two variants were developed from that basic model. The low w variant accounts for the fact that, in robust design, the control factors are generally explored over a wider range than the noise factors. The 2nd order variant zeros out the coefficients of all the three way interactions.

The fitted weak heredity model (fitted WH), was developed specifically for use in this case study. The parameters of the model were selected based on their fit to a set of empirical data. A group of 62 full factorial experiments were collected from technical journals (including 2^3 , 2^4 , 2^5 , and 2^6 designs). In this data set, 43% of main effects were active according to the step-down Lenth method (Lenth 1989). The other probabilities in Equations 9 and 10 were estimated by the same means.

The estimation of the other parameters (c , s_1 , etc.) was based exclusively on the four factor experiments 2^4 among the 62 experiments (there were 23 of these). First, the data were standardized so that we could meaningfully compare values across systems. The responses from each experiment were transformed so that the minimum observation was zero and the maximum observation was 100. The factor effects were then computed from the normalized data. The value of s_1 was estimated by computing the ratio of the standard deviation of all main effects and the standard deviation of all two-way interactions. The value of s_2 was estimated the same way but by comparison with three-way interactions. Histograms of factor effects were formed for the collection of 23 real systems and for a set of 1000 systems instantiated from the WH model. The shape of the distribution was used to adjust the parameter c . If the setting of c is too low, the histogram of factor effects has overly thick tails. This led to adjustment of c from 10 to 15 which provided a more reasonable fit of the model to the data (as shown in Figures 1 and 2).

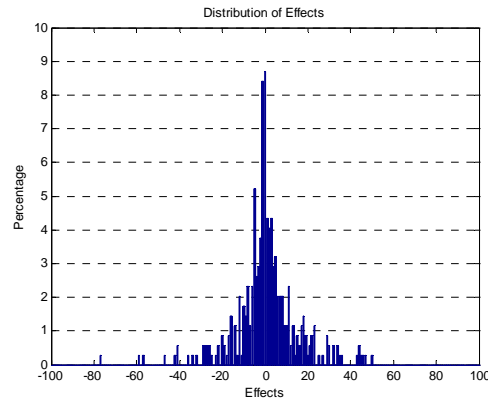


Figure 1. The distribution of factor effects from 23 real systems

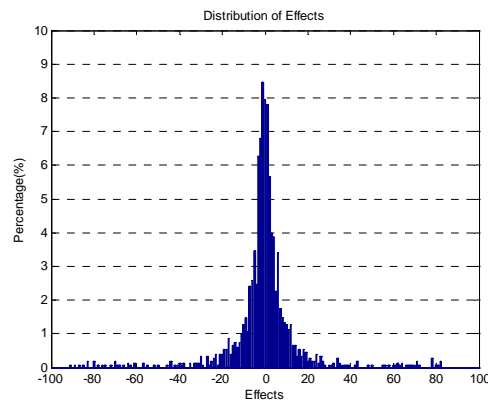


Figure 2. The distribution of factor effects from 1000 simulated systems sampled from the fitted weak heredity model with $c=15$

5.2 ROBUST DESIGN METHODS TO BE EVALUATED

Eight different approaches to robust design were evaluated. Five of the techniques are crossed array strategies of different sorts. Three are single array approaches. The details of each approach are given below.

$2^7 \times 2^3$ -- A full factorial 2^7 inner array of control factors was crossed with a full factorial 2^3 outer array of noise. This approach is expensive, requiring 1,024 experiments, but provides resolution to estimate all 175 coefficients in the model (Equation 1). Based on these parameters, the standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates. This design provides a baseline value since the standard deviation optimized by this approach will be the lowest possible value within the given discrete space of control factor settings.

$2^7 \times 2_{III}^{3-1}$ -- A full factorial 2^7 inner array of control factors was crossed with a fractional factorial 2_{III}^{3-1} outer array of noise. The data from the design were used to calculate all noise main effects, control by noise interactions, and control by control by noise interactions. Based on these parameters, the

standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates.

2^{10-4} -- A 64 run single array approach was employed as described by Wu and Hamada (2000) with design generators A=1, B=2, C=3, D=4, E=5, F=123, G=124, a=6, b=1345, c=23456. The single array was executed and the resulting data were used to calculate the main effect of the noise factors and control by noise interactions. Based on these parameters, the standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates.

2^{10-5} -- A 32 run single array approach was employed as described by Wu and Hamada (2000) with design generators A=1, B=2, C=3, D=4, E=234, F=134, G=123, a=5, b=124, c=1245. The single array was executed and the resulting data were used to calculate the main effect of the noise factors and control by noise interactions. Based on these parameters, the standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates.

$2_{III}^{7-4} \times 2_{III}^{3-1}$ -- A fractional factorial 2_{III}^{7-4} inner array of control factors was crossed with a fractional factorial 2_{III}^{3-1} outer array of noise. This requires a total of 32 experiments. The data from the design were used to calculate all noise main effects and control by noise interactions. Based on these parameters, the standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates.

$OFAT \times 2_{III}^{3-1}$ -- An adaptive one-factor-at-a-time plan was used to modify the control factors. At a randomly selected baseline configuration of the control factors, a fractional factorial 2_{III}^{3-1} outer array of noise was executed. The data from the design were used to calculate all noise conditional main effects. Based on these parameters, the standard deviation was estimated based on Equation 12. Following this baseline assessment, a randomly selected control factor was varied and the outer array of noise was executed again. The standard deviation at the current control factor setting was compared to the baseline standard deviation. If the current settings resulted in the lowest observed standard deviation, then the change was adopted and the next control factor was toggled. This adaptive process repeated until all seven control factors were changed. In each case, the change in control factor settings was adopted only if it resulted in the lowest standard deviation observed so far in the experiment. The complete process requires 32 simulated experiments.

$OFAT \times OFAT$ -- An adaptive one-factor-at-a-time plan was used exactly as in the $OFAT \times 2_{III}^{3-1}$ method except that a one-factor-at-a-time outer array of noise was executed instead of the 2_{III}^{3-1} design. For each control factor setting explored, an experiment was conducted at a randomly selected baseline noise level. Each noise factor was changed individually to estimate its conditional main effect. Then these conditional main effects were assumed to be estimates of the first order effects of noise and the standard deviation of the response was calculated using equation 12.

2^{10-6} -- A 16 run single array approach was employed as described by Wu and Hamada (2000) with design generators A=1, B=2, C=3, D=4, E=1234, F=24, G=123, a=134, b=14, c=34. The single array was executed and the resulting data were used to calculate the main effect of the noise factors and control by noise interactions. Based on these parameters, the standard deviation was estimated based on Equation 12 and the control factor settings were optimized based on these estimates.

5.3 RESULTS OF THE APPROACH AS APPLIED IN THE CASE STUDY

The four step method described in section 4 was applied to 48 pairings of six system types and eight robust design methods. The results of the study are presented in Tables 5 and 6. The percentages depicted in Table 5 are the expected value (averaged across the systems) of the percentage of confirmed improvement in the standard deviation of the response. The confirmed improvement is the confirmed standard deviation from Equation 12 evaluated at the predicted optimum control factor settings. The percentage of confirmed improvement was defined so that 100% improvement implies the confirmed standard deviation is zero and 0% improvement implies that the confirmed standard deviation is no better than selecting the control factor settings at random from among the available discrete settings. The expectation is estimated by averaging the percentages across all 100 systems instantiated from that parameter set. The inter-quartile ranges are computed to provide an indication of the variability of the outcomes across different systems of the same type.

TABLE 5. Expected values of percent reduction in standard deviation for various robust design methods and system parameter sets

Method	Experiments	Basic			Fitted		
		WH	low w	2 nd order	WH	low w	2 nd order
$2^7 \times 2^3$	1,024	60%	81%	58%	50%	58%	40%
$2^7 \times 2_{III}^{3-1}$	512	44%	80%	52%	45%	58%	40%
2^{10-4}	64	8%	8%	56%	18%	9%	38%
2^{10-5}	32	9%	3%	33%	16%	9%	17%
$2_{III}^{7-4} \times 2_{III}^{3-1}$	32	12%	8%	51%	16%	25%	38%
<i>OFAT</i> $\times 2_{III}^{3-1}$	32	39%	56%	43%	36%	42%	35%
<i>OFAT</i> \times <i>OFAT</i>	32	31%	37%	41%	33%	31%	27%
2^{10-6}	16	4%	4%	8%	4%	2%	0%

TABLE 6. Inter-quartile ranges of percent reduction in standard deviation for various robust design methods and system parameter sets

Method	Experiments	Basic			Fitted		
		WH	low w	2 nd order	WH	low w	2 nd order
$2^7 \times 2^3$	1,024	51 to 69%	77 to 85%	41 to 71%	31 to 61%	38 to 77%	15 to 60%
$2^7 \times 2_{III}^{3-1}$	512	31 to 61%	76 to 85%	37 to 65%	27 to 58%	37 to 75%	15 to 62%
2^{10-4}	64	-12 to 28%	-15 to 40%	39 to 70%	2 to 35%	-7 to 33%	14 to 55%
2^{10-5}	32	-10 to 35%	-20 to 30%	18 to 52%	-2 to 34%	-14 to 31%	-8 to 34%
$2_{III}^{7-4} \times 2_{III}^{3-1}$	32	-17 to 38%	-24 to 35%	36 to 65%	-3 to 38%	9 to 49%	13 to 57%
<i>OFAT</i> $\times 2_{III}^{3-1}$	32	26 to 54%	45 to 68%	29 to 59%	19 to 49%	23 to 62%	9 to 49%
<i>OFAT</i> \times <i>OFAT</i>	32	15 to 46%	18 to 54%	26 to 55%	17 to 50%	12 to 51%	7 to 41%
2^{10-6}	16	-21 to 27%	-20 to 33%	-11 to 27%	-16 to 22%	-17 to 19%	-13 to 13%

5.4 DISCUSSION OF THE CASE STUDY

To begin the discussion of the data from this case study (as presented in Tables 5 and 6), it is important to note that the most effective robust design methods provided moderately large reductions in standard deviation. A typical application of robust parameter design using the best available method reduced standard deviation by a factor of about 2 to 4. Those are practically significant improvements and seem to be in-line with figures from industrial practice. However, the inter-quartile ranges show that even the best method sometimes accomplished only a modest improvement in robustness. Robust parameter design relies on the existence of control by noise interactions that can counter the main effects of noise. Some systems simply lack interactions of the right size or structure to effect much improvement.

In comparing the results of different methods, the most salient feature is the excellent performance of the $OFAT \times 2_{III}^{3-1}$ method for robust design. For all six versions of the WH model, the $OFAT \times 2_{III}^{3-1}$ provided the best results or quite competitive results compared to all the alternatives requiring less than 100 experiments. In comparison the best possible improvement (which was always provided by the $2^7 \times 2^3$ approach) the $OFAT \times 2_{III}^{3-1}$ always provided a large fraction of the benefit on average (around $\frac{3}{4}$) while requiring only three percent of the experimental cost.

It is instructive to make a direct comparison of the $OFAT \times 2_{III}^{3-1}$ with the $2_{III}^{7-4} \times 2_{III}^{3-1}$. The $OFAT \times 2_{III}^{3-1}$ provides superior results in almost every scenario. When third order terms are included in the model, the results of the $OFAT \times 2_{III}^{3-1}$ are better by a factor of about two or substantially more. For second order models, the two techniques provide comparable results. An explanation for these trends is that the adaptive one-factor-at-a-time approach can exploit some control by control by noise interactions. Despite the assumptions of hierarchy and inheritance embedded in the WH model, these three-way interactions seem to play an important role in robust design.

It is also instructive to make a direct comparison of the $OFAT \times 2_{III}^{3-1}$ with the $OFAT \times OFAT$ approach. The $OFAT \times 2_{III}^{3-1}$ provides slightly better results for all six versions of the WH model. A reasonable conclusion is that the greater statistical efficiency of the fractional factorial design is valuable for the outer (noise) array.

Another significant result of this study is that the single array approaches fare very poorly except when the 2^{10-4} is used on a second order system. The 2^{10-4} single array is arranged so that all 21 control by noise interactions and all noise main effects are clear of main effects and other two-factor interactions. In a second order system, all the opportunities for robustness arise from cancellation of noise main effects using control by noise interactions. As a result, the 2^{10-4} approach works exceptionally well for the second order systems. However, the results presented here show that the 2^{10-4} is affected very strongly by the presence of three-way interactions. In some cases, the lower quartile was less than zero indicating that confirmation experiments would, about 25% of the time, result in a standard deviation higher than the average system within the parameter space. This is a strong indication of an unreliable approach to robust design.

The other single array design performed even worse. The 2^{10-5} design does not fare as well as the well 2^{10-4} on second order systems. In the 2^{10-5} design only 14 of the 21 control by noise interactions are clear. This would seem to account for the fact that only about 2/3 of the available robustness improvement was attained. Like the 2^{10-4} , the 2^{10-5} fared very poorly on systems including three-way interactions. The 2^{10-6} fared very poorly on all the versions of the WH model. In the 2^{10-6} design, none of the control by noise interactions are clear. The data from this analysis suggest the 2^{10-5} is not an effective design for robustness improvement of systems with seven control factors and three noise factors.

To summarize, a principal result of the study is that an extremely effective combination is an adaptive one-factor-at-a-time inner array crossed with a fractional factorial outer array ($OFAT \times 2_{III}^{3-1}$). Across many different variants of the WH model, this technique provided most of the benefits to be gained by parameter design while exploring only 3% of the parameter space. No technique using a fractional factorial inner array or a single fractional factorial array was competitive with the $OFAT \times 2_{III}^{3-1}$ approach if there was allowed some possibility of three-way interactions. This is significant since no current text or recent research literature has suggested that OFAT should be used for robust design and most of the modern textbooks specifically argue against it.

6. A Broader Discussion in Relation to Complex Engineering Systems

The evaluation technique and case study presented here may have some interesting implications for the emerging field of complex engineering systems. This broader discussion seems to fall naturally into three clusters which will be addressed in the following paragraphs: 1) the role of robust design in engineering systems, 2) flexibility and the design process, and 3) possible directions for research methods.

Robust design is an important strategy for dealing with the uncertainties attendant in the design of complex engineered systems. If components and subsystems can be made robust to uncertain interface variables, then system integration is likely to proceed more smoothly. However, it is not at all obvious how best to deploy robust design in a Systems Engineering effort. Some robust design techniques provide good results reliably but at high cost. If these are used, then they can only be applied to the most critical components. Other robust design methods are simpler and less demanding of resources but occasionally result in disappointing outcomes. Such techniques can be taught to a large number of engineers and implemented (to some degree) on almost every component and subsystem. It falls upon the Systems Engineer to devise an overall strategy for deploying all available methods across the system. In order for Systems Engineers to perform that function, they need information about the pros and cons of every available technique. The literature on robust design includes arguments for and against different methods, but most of the arguments are made on the basis mathematical properties of the experimental designs. This paper presents an approach to evaluating robust design methods directly on the basis of their outcomes. I propose that information in this form is more valuable to the Systems Engineer than the kind of information previously available. Further, I believe that when Systems Engineers see that an adaptive OFAT approach will provide a 42% reduction in variance and that a much more expensive full factorial approach can only provide 58% (as shown in Table 5), there is likely to be a significant change in their behavior with more “quick and dirty” approaches deployed more broadly.

Flexibility is regarded as an important foundational issue in complex engineering systems. A central issue in designing any system is the tension between optimality for a fixed purpose and flexibility in the face of change. Most often, this tension is discussed vis-à-vis the flexibility/optimality of the *product*, but

the tension is equally relevant to the design *process*. Should an enterprise plan a design process that will lead to the best outcomes assuming current knowledge of the design scenario, or employ a flexible design process that may not be ideally suited to any one condition but can adapt well in the face of new information that comes to light? This paper has shown that, in robust design, the value of adaptability and flexibility is so significant that, when accounted for, it overturns the conventional wisdom. No modern text on the subject of design or experiments or robust design suggests any valid role for an OFAT approach. Nevertheless, the data from these new simulations strongly suggests OFAT is a preferred design for an inner array. It appears that flexibility of the design process is far more valuable than previously acknowledged by the technical literature on design of experiments or robust design. Similar misconceptions may exist in other areas of Systems Engineering.

If the results of this paper are borne out, how can it be that the benefits of an OFAT approach to robust design have gone unnoticed given two decades of strong research efforts in this area? My view is that there has been an over-dependence on closed form analysis. Closed form proofs are a genuine and widely recognized sign of scholarly accomplishment. It is very difficult to write a proof showing that a complex adaptive process will have desirable properties. Therefore, adaptive processes for robust design have garnered little attention from academic communities. On the other hand, it is often easy to demonstrate the behavior of a complex process by computational simulation once a reasonable set of assumptions can be articulated and defended. One key set of assumptions for this study concerned the structure of interactions among variables in engineering systems. The development of the relaxed weak heredity model by Wu and Hamada was a critical enabler for the research presented here. It enabled the creation of multiple instances of reasonably realistic simulated engineering systems upon which robust design methods could operate. It may be that other aspects of the Systems Engineering process could also be simulated if the other reasonable assumptions can be identified and codified. If so, design process simulations like the ones discussed here could represent a promising research methodology for a wide range of Systems Engineering topics.

Acknowledgements

The support of the Ford/MIT strategic alliance is gratefully acknowledged.

References

- Borror, CM, and Montgomery, DC: 2000, "Mixed Resolution Designs as Alternatives to Taguchi Inner/Outer Array Designs for Robust Design Problems," *Quality and Reliability International*, **16**:117-127.
- Box, GEP, Hunter, WG, and Hunter, JS: 1978, *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, John Wiley & Sons, Inc., NY.
- Chipman, HM, Hamada, M, and Wu, CFJ: 1997, "Bayesian variable-selection approach for analyzing designed experiments with complex aliasing," *Technometrics* **39**(4): 372-381.
- Fisher, RA: 1949, *The Design of Experiments*, 5th Ed., Oliver and Boyd, London and Edinburgh.
- Frey, DD: 2003, "How One-Factor-at-a-Time Experimentation Can Lead to Greater Improvements than Orthogonal Arrays," DETC 2003/DTM48646, ASME Design Engineering Technical Conference, September 2-6, Chicago, IL.

- Frey, DD, Engelhardt, F, and Greitzer, EM: 2002, "A Role for One Factor at a Time Experimentation in Parameter Design", *Research in Engineering Design* **14**(2): 65-74.
- Hou, XS: 2002, "On the use of compound noise factor in parameter design experiments", *Applied Stochastic Models in Business and Industry* **18**:225-243.
- Lenth, RV: 1989, "Quick and Easy Analysis of Unreplicated Factorials," *Technometrics* **31**(4): 469-473.
- Phadke, MS: 1989, *Quality Engineering Using Robust Design*, PTR Prentice-Hall, Inc.: A Simon & Schuster Company, Englewood Cliffs, NJ.
- Taguchi, G: 1987, *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Costs*, translated by Tung, L.W., QualityResources: A division of the Kraus Organization Limited, White Plains, NY; and American Supplier Institute, Inc. Dearborn, MI, Vol. 1, pp. 1-531.
- Wu, CFJ., and Hamada, M: 2000, *Experiments: Planning, Analysis, and Parameter Design Optimization*, Wiley & Sons, Inc., NY.