Using Statistically Designed Experiments to Improve Reliability and Achieve Robust Reliability

M.S. Hamada
University of Waterloo

RR-94-06 (NT)
May 1994

USING STATISTICALLY DESIGNED EXPERIMENTS TO IMPROVE RELIABILITY AND ACHIEVE ROBUST RELIABILITY

Michael S. Hamada
The Institute for Improvement in Quality and Productivity
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

ABSTRACT

Statistically designed experiments provide a proactive means to improve reliability as has been advocated recently by Genichi Taguchi, the Japanese quality engineer. That is, by systematic experimentation, the important parameters (factors) affecting reliability can be identified along with parameter values that yield reliability gains. In addition to improving reliability, Taguchi's robust design can be used to achieve robust reliability; that is, to make a process' or product's reliability insensitive to factors which are hard or impossible to control. Robust design is also implemented using statistically designed experiments. In this paper, different classes of experimental plans for reliability improvement and robust reliability are presented. An important feature of the reliability data collected from such experiments is censoring which occurs when all experimental units have not failed by the end of the experiment. Consequently, the analysis methodology must account for these censored data which are likely to occur in light of the ever increasing reliability of today's products. Several appropriate methods are discussed briefly. These experimental plans and analysis methods are illustrated using three documented experiments which improved fluorescent lamp and industrial thermostat reliability and which achieved robust reliability for night vision goggles.

Key words: Bayesian, Box-Behnken design, Censored data, Central-Composite design, Control and noise factors, Estimability, Experimentation, Fractional factorial, Full factorial, Lognormal, Maximum likelihood estimation, Mixed-level design, Plackett-Burman design, Regression model, Robust design, Product array, Response surface design, Screening, Sequential approach, Weibull.

1. INTRODUCTION

Statistically designed experiments have been used extensively for estimating or demonstrating existing reliability (Nelson [14]). Until recently, they appear to have seldom been used to to improve reliability by identifying the important parameters (factors) affecting reliability out of many potentially important ones. For example, Genichi Taguchi (Taguchi and Wu [26], Taguchi [24], [25]) advocated their use as a proactive means for improving reliability and provides examples of experiments to improve clutch springs and fluorescent lamps. Taguchi is perhaps best known for robust design, whose aim is to make processes/products insensitive to *noise* factors which are hard or impossible to control. Such products/processes are said to be *robust* to the noise factors. Examples of noise factors include manufacturing variables that cannot easily be controlled and environmental conditions in which the product is used. This important paradigm for improving products/processes, which attracted the attention of industry in the 1980's (Kackar [11]), can also be applied to reliability. In order to ensure good stability and adequate reliability, Taguchi [24] (page 149) recommended that noise factors be considered in any experiment to improve reliability when it is practical to do so.

Since the early 1980's, several experiments for improving reliability have been documented. In the *Symposia on Taguchi Methods 1984-1993*, Specht [22] reported the improvement of heat-exchanger reliability in a commercial heating system; Montmarquet [13] discussed the improvement of drill bit reliability in a multilayer printed circuit board drilling operation and Reed [18] presented the improvement of night vision goggle reliability. From an early application of Taguchi's methodology at AT&T, Phadke [15] discussed the improvement of router-bit reliability in a printed circuit board cutting operation. In *Reliability Improvement with Design of Experiments*, Condra [4] gave several examples from the electronics industry. Taguchi's robust design philosophy figures prominently in Condra's book and is recommended reading. Recently, Bullington, Lovin, Miller and Woodall [3], reported on the improved reliability of industrial thermostats.

The purpose of this paper is to illustrate how statistically designed experiments can be used to improve reliability and to achieve robust reliability. First, separate classes of experimental plans for improving reliability and achieving robust reliability are discussed with examples in Sections 2 and 3, respectively. In Section 4, a brief discussion of analysis methodology for extracting the information from the experimental data is given. The methodology must account for an important feature of such experiments known as censoring which occurs when all the experimental units have not failed by the end of the experiment; this type of censoring produces what is referred to as Type I or right-censored data. There are also other types of censoring which typically arise in such experiments. In situations where units cannot be monitored continuously, units must be inspected periodically until failure. Periodic inspection produces left-censored data for units failing before the first inspection and interval-censored data, otherwise, i.e., from units failing between two consecutive inspections. Appropriate analysis techniques which handle censored data are illustrated in Sections 5 through 7 which present respective analyses of experiments to improve fluorescent lamp and thermostat reliability and to achieve robust reliability of night vision goggles.

Notation

2^{k-p}	two-level fractional factorial, k factors
3^{k-p}	three-level fractional factorial, k factors
Y_i	ith response
T_i	ith lifetime
x_i	ith vector of covariate values
β	regression parameters or effects
σ	scale parameter
ϵ_i	ith error associated with Y_i
log	natural logarithm
$x_{control}$	vector of control factors, also covariates in the regression model involving only the control factors
x_{noise}	vector of noise factors, also covariates in the regression model involving only the noise factors
$x_{control \times noise}$	covariates in the regression model corresponding to control by noise interactions
$l(\cdot)$	loss function
$L(x_{control})$	loss at $x_{control}$
$Y(x_{control}, x_{noise})$	response at $(x_{control}, x_{noise})$
$f(\cdot)$	joint pdf of noise factors
C	control factor main effect
N	noise factor main effect
$N \times N$	noise factor by noise factor interaction
$C \times N$	control factor by noise factor interaction
$C \times C$	control factor by control factor interaction
ML	maximum likelihood
MLE	ML estimate
$\{i \in CEN\}$	set of censored data
$\{i \in FAIL\}$	set of failure data
$\mathcal{L}(\cdot)$	likelihood function
$p(\cdot)$	prior density function
A_j	jth level of factor A
$gaud(w)$	$(1/\sqrt{2\pi})\exp(-w^2/2)$
$gaufc(w)$	$1 - \int_{-\infty}^{w} gaud(y)dy$

2. EXPERIMENTS FOR IMPROVING RELIABILITY

While there may be potentially many factors (parameters) that affect reliability, some factors will tend to be more important, i.e., have a bigger impact on reliability as the values of these factors are changed. These important factors can be identified empirically through experimentation which involves making deliberate changes in the factor values and observing the resulting reliability. Besides identifying the important factors, values for these factors that yield reliability gains can be recommended. Statistically designed experiments provides a systematic and efficient plan of experimentation to achieve thse goals. Several factors can be studied simultaneously using as few resources as possible. Designed experiments have been used successfully to improve other quality characteristics (see the *Symposia on Taguchi Methods*, 1984-1993) and can be employed to improve reliability.

Some terminology will be helpful in describing various plans below. The plan of experimentation is referred to as the *experimental design* or *design*. The experimental design consists of a list of *runs*, where a run is a combination of values (*levels*) at which the factors in the experiment are set. The number of runs in the experimental design is called the *run size*. The experiment then involves making units according to the conditions specified by the runs in the experimental design and life testing these units to failure. Table 1 gives an experimental design for three factors (denoted by A-C) with each factor being studied at two levels (denoted by 1 and 2). Run 1 indicates that units are made with all the factors set at their respective first levels. This particular design is called a two-level *full factorial* since it consists of all possible combinations of the two levels for the three factors.

The run size for a two-level full factorial design in k factors is 2^k, which quickly becomes prohibitive for more than five factors. Designs with more than two levels, say three, allow curvature effects to be assessed but require even more runs. Even for three factors, the run size of the full factorial design is already 27 (i.e., 3^3). Because of their large run sizes, these designs tend not to be used in an initial experiment unless there are only a few potentially important factors to be studied.

For the typical industrial situation, a large number of factors needs to be studied in a relatively small number of runs. A sequential approach to experimentation provides one such strategy. An initial experiment using only a few levels (often two) for each factor is

Table 1: Two-Level Full Factorial Design for Three Factors

	Factor		
Run	A	B	C
1	1	1	1
2	1	1	2
3	1	2	1
4	1	2	2
5	2	1	1
6	2	1	2
7	2	2	1
8	2	2	2

used to screen out the unimportant factors. A follow-up experiment involving much fewer factors but at more levels can then be performed to explore the response-factor relationship in more detail. For the initial experiment, a subset or fraction of the full factorial design design (called a *fractional factorial* design) can be used. For two-level designs, there are two types, the geometric or 2^{k-p} designs (Box, Hunter and Hunter [2]) and the non-geometric Plackett-Burman [16] designs. The notation for the geometric designs indicate the degree of fractionation, i.e., a 2^{-p} fraction of a full factorial with run size 2^{k-p}.

Taguchi [25] (page 930) provides an example of a 2^{k-p} design which was used in an experiment to improve the reliability of fluorescent lamps. The experiment studied five two-level factors (denoted by A-E) in eight runs using a 2^{5-2} design or a quarter fraction of a full factorial as given in Table 2. No further details on the factor names and levels were provided, presumbably for reasons of confidentiality. Two lamps were made at each run and life testing was conducted over 20 days with inspections for failure being performed every two days. The lifetime data also appear in Table 2, with $(14, 16)$ meaning the lamp failed between days 14 and 16 and $(20, \infty)$ indicating that the lamp was still working at the 20 day inspection. Note that seven of the 16 lamps had not failed by the 20 day inspection which yielded right-censored data.

Bullington et al. [3] provides an example of a 12-run Plackett-Burman design which was used in an experiment to improve the reliability of industrial thermostats. Eleven factors (denoted by A-K) were studied using the design given in Table 3 in which ten thermostats

Table 2: Design and Lifetime Data for the Fluorescent Lamp Experiment

Factor								
A	B	C	D	E			Lifetime	
1	1	1	1	1	1	1	(14,16)	(20,∞)
1	1	2	2	2	1	2	(18,20)	(20,∞)
1	2	1	1	2	2	2	(08,10)	(10,12)
1	2	2	2	1	2	1	(18,20)	(20,∞)
2	1	1	2	1	2	1	(20,∞)	(20,∞)
2	1	2	1	2	2	2	(12,14)	(20,∞)
2	2	1	2	2	1	2	(16,18)	(20,∞)
2	2	2	1	1	1	1	(12,14)	(14,16)

Table 3: Design and Lifetime Data for the Thermostat Experiment

(with censoring time of 7342)

Design											Lifetime Data									
A	B	C	D	E	F	G	H	I	J	K										
1	1	1	1	1	1	1	1	1	1	1	957	2846	7342	7342	7342	7342	7342	7342	7342	7342
1	1	1	1	1	2	2	2	2	2	2	206	284	296	305	313	343	364	420	422	543
1	1	2	2	2	1	1	1	2	2	2	63	113	129	138	149	153	217	272	311	402
1	2	1	2	2	1	2	2	1	1	2	76	104	113	234	270	364	398	481	517	611
1	2	2	1	2	2	1	2	1	2	1	92	126	245	250	390	390	479	487	533	573
1	2	2	2	1	2	2	1	2	1	1	490	971	1615	6768	7342	7342	7342	7342	7342	7342
2	1	2	2	1	1	2	2	1	2	1	232	326	326	351	372	446	459	590	597	732
2	1	2	1	2	2	2	1	1	1	2	56	71	92	104	126	156	161	167	216	263
2	1	1	2	2	2	1	2	2	1	1	142	142	238	247	310	318	420	482	663	672
2	2	2	1	1	1	1	2	2	1	2	259	266	306	337	347	368	372	426	451	510
2	2	1	2	1	2	1	1	1	2	2	381	420	7342	7342	7342	7342	7342	7342	7342	7342
2	2	1	1	2	1	2	1	2	2	1	56	62	92	104	113	121	164	232	258	731

were manufactured at each of the 12 run settings. These factors were chosen from many across a 14 stage manufacturing process and include the Beryllium copper grain size (factor E), the heat treatment (factor H), the power element electroclean (factor J) and power element plating rinse (factor K). Each factor was studied at two levels such as factor E with grain sizes of 0.008 and 0.018 inches or factor H at 45 minutes and 240 minutes at 600 degrees Fahrenheit. Table 3 also presents the lifetime data; the experiment was stopped at 7432 (×1000) cycles resulting in 22 right-censored observations at runs 1, 6 and 11. See Bullington et al. [3] for a detailed account of the experiment.

While highly fractionated 2^{k-p} and Plackett-Burman designs are ideally used as screening designs, in practice, the initial experiment may be the only one performed. Consequently, a properly chosen 2^{k-p} design can allow some potential interactions to be studied. For example in the fluorescent lamp experiment, besides the factors A-E main effects, the experimenter also thought that the $A \times B$ interaction might be potentially important. The design given in Table 2 allows the $A \times B$ interaction to be estimated. By factor *main effects*, it is meant the additive effects of the factors on reliability. The *interaction* between two factors indicates the degree of non-additivity of the factor effects; that is, if interaction is present, the effect of changing the levels of one factor on reliability depends on the level of the other factor. That is, the presence of an interaction can impact the recommendations made for setting the important factors. See Box et al. [2] for more discussion. Also, for Plackett-Burman designs, Hamada and Wu [9] has shown that some information on interactions may be obtained.

Taguchi [25] often initially uses designs with more than two levels. These include the 3^{k-p} designs (i.e., three-level fractional factorial designs) and *mixed-level* designs such as the 18 run design which can be used to study one two-level factor and up to seven three-level factors. For example, the clutch spring experiment in Taguchi [24] (chapter 9) used a 3^{7-4} design to study seven factors in 27 runs, a quarter replicate of a three-level full factorial design. Phadke [15] also used a mixed-level 32-run design to study two four-level factors and seven two-level factors. Dey [5] and Wang and Wu [27] catalogue other mixed-level designs.

In contrast with the initial use of multi-level factor designs, the sequential approach to experimentation uses such designs in a follow-up experiment. Box and Draper [1] give different designs referred to as *response surface designs* which allow the response-factor relationship to explored in more detail. For example, an internal document from a North American automobile manufacturer reports the use of an eight run experiment to screen seven factors. (The same design given in Table 2 consisting of the first seven columns was employed.) Four factors were identified and studied further using a 27-run Box-Behnken design given in Table 4. $(-1, 0, +1)$ denotes the three levels for each factor; each of the rows 1-2, 4-5 and 7-8 specify four runs since the $\pm$ notation means all combinations of the first and third levels are used. Other designs such as the central-composite designs (with five factor levels) can be employed. See Box and Draper [1] for more details.

The lifetime data from these experimental designs can be analyzed using a parametric model such as the lognormal or Weibull regression models. These models have the form (Lawless [12]):

Table 4: Box-Behnken Response Surface Design for Four Factors

Factor			
A	B	C	D
±1	±1	0	0
0	0	±1	±1
0	0	0	0
±1	0	0	±1
0	±1	±1	0
0	0	0	0
±1	0	±1	0
0	±1	0	±1
0	0	0	0

$$Y_i = \log(T_i) = x_i{}^T\beta + \sigma\epsilon_i,\ i = 1,\ \ldots,\ n, \tag{1}$$

where the $\{T_i\}$ are the lifetimes, the $\{x_i\}$ are the corresponding vectors of covariates values, β is the vector of location parameters and σ is the scale parameter. For the Weibull model, the errors $\{\epsilon_i\}$ are independent and identically distributed (i.i.d.) standard extreme-value random variables (r.v.s), whose probability density function (pdf) and survival function (Sf) are $exp(w - exp(w))$ and $exp(exp(-w))$, respectively. For the lognormal model, the errors $\{\epsilon_i\}$ are i.i.d. standard normal r.v.s, whose pdf and Sf are $gaud(w)$ and $gaufc(w)$, respectively.

Appropriate analysis methodology for fitting these models using censored data are discussed in Section 4. Analyses of the fluorescent lamp and thermostat experiments will be presented in Sections 5 and 6, respectively.

3. EXPERIMENTS FOR ACHIEVING ROBUST RELIABILITY

Taguchi's robust design is also referred to as parameter design because its objective is to find levels of engineering parameters (called *control factors*) that yield a robust product/process, i.e., that make the product/process insensitive to the variation of hard or impossible to control *noise factors*. Robust design is therefore strikingly different from the traditional approach of handling sources of variation by control which can be costly. For example, the Ina Tile Company was faced with reducing an unacceptable amount of variation in their tiles' size caused by an uneven temperature distribution in the kiln (Kackar [11]). Rather than purchasing an expensive kiln which would have controlled the temperature distribution better, it was found through designed experiments that increasing the lime content in the tile formulation decreased the tile size variation by a factor of ten. In other words, a tile formulation was found that was insensitive to the existing oven's uneven temperature distribution.

Taguchi's tactics for carrying out robust design are to specify a criterion for assessing the effect of the noise factors and to estimate it by experimentation. Note that while noise factors are difficult or impractical to control in production or in use, for purposes of the experiment (i.e., to learn about the effect of the noise factors), the noise factors need to be controlled during the experiment. The criterion for assessing the effect of the noise factors (termed the loss and denoted by $L(\cdot)$) at a particular combination of control factor levels $x_{control}$ can be defined for a general loss function $l(\cdot)$ (Welch, Yu, Kang and Sacks [28]) as:

$$L(x_{control}) = \int l(Y(x_{control}, x_{noise}))f(x_{noise})dx_{noise} , \tag{2}$$

where $Y(x_{control}, x_{noise})$ is the observed lifetime at a particular combination of control and noise factor levels $(x_{control}, x_{noise})$ and $f(\cdot)$ is the joint probability density function of the noise factors. The objective of robust design then is to find a product/process design $x_{control}$ with minimum loss. Some appropriate loss functions for reliability will be discussed below.

Taguchi [25] proposed estimating the loss (1) via experimentation and then modelling the estimated losses in terms of the control factors. Taguchi uses specialized experimental plans referred to as product arrays. A product array consists of two plans or arrays, one for the control factors called the *control array* and the other for the noise factors called the *noise array*. The product array design is so named because all the noise factor combinations specified by the noise array are run with every combination of the control factors specified by the control array.

As an example, consider an experiment that was performed to improve the reliability of a night vision goggle tube sub-assembly (Reed [18]). The tube sub-assembly is insulated by a combination of protective coatings which degrades over time and exposure to humidity and temperature. One goal of the experiment was to make the tube insulation reliability robust to the handling of tube when the coatings are applied. The noise factor (denoted by N) had two levels, whether the tube is handled or not. Ten control factors (denoted by A-J) were chosen from many across a 17 step manufacturing process and include factors related to the tube packaging such as type of coating, primer, and electrical connection configuration. Each control factor was also studied at two levels according to a 12-run Plackett-Burman design; e.g., two types of primer coating (factor B) and two types of lead coating material (factor H) were used. Therefore, the product array consisted of a a 12-run Plackett-Burman design for the control factor array and a simple noise factor array (a single factor at two levels) as displayed in Table 5. At each of the 24 control and noise factor combinations, one tube was manufactured and then life tested under a high temperature cycling and humidity regimen. The tubes were inspected for failure every two days for 20 days. The lifetime data presented in Table 5 makes some assumptions because it is unclear from Reed [18] whether 16 of the tubes failed between days 18-20 or whether they were still functioning at 20 days. For purposes of illustration, they are treated here as still functioning, resulting in right censored data. Note also that six of the tubes failed before the first inspection at day two yielding left-censored data.

For analyzing the product array data, Taguchi [25] proposed estimating the loss $L(x_{control})$ (2) for each $x_{control}$ specified by the control array from the data obtained by varying the noise factors according to the noise array and then modeling the estimated losses in terms of the control factors. That is, he proposed constructing responses from the noise array data and analyzing them by standard methods for designed experiments such as analysis of variance. Alternatively, Welch et al. [28] proposed modeling the response Y directly in terms of both the control and noise factors and then evaluating the loss using the estimated response model. Their rationale for the latter approach, called the response-model approach by Shoemaker, Tsui and Wu [20], was that it would be more likely to find a simple model for the response than one for the much more complicated estimated loss. Examples in Welch et al. [28] and Shoemaker et al. [20] give evidence for preferring the response-model approach because it also provides additional insight.

Table 5: Product Array Design and Lifetime Data for the Goggle Experiment

Control Array										Noise Array N	
A	B	C	D	E	F	G	H	I	J	1	2
1	1	1	1	1	1	1	1	1	1	$(0,2)$	$(20,\infty)$
1	1	1	1	2	2	2	2	2	2	$(20,\infty)$	$(20,\infty)$
1	1	2	2	1	1	1	2	2	2	$(0,2)$	$(0,2)$
1	2	1	2	1	2	2	1	1	2	$(7,9)$	$(20,\infty)$
1	2	2	1	2	1	2	1	2	1	$(20,\infty)$	$(20,\infty)$
1	2	2	2	2	2	1	2	1	1	$(20,\infty)$	$(7,9)$
2	1	2	2	1	2	2	1	2	1	$(20,\infty)$	$(20,\infty)$
2	1	2	1	2	2	1	1	1	2	$(20,\infty)$	$(0,2)$
2	1	1	2	2	1	2	2	1	1	$(0,2)$	$(0,2)$
2	2	2	1	1	1	2	2	1	2	$(20,\infty)$	$(20,\infty)$
2	2	1	2	2	1	1	1	2	2	$(20,\infty)$	$(20,\infty)$
2	2	1	1	1	2	1	2	2	1	$(20,\infty)$	$(20,\infty)$

For reliability applications, the response-model approach is a natural one because the same parametric regression models given in Section 2 can be used. The product array data allows a model to be fit consisting of all C main effects (with possibly some $C \times C$ interactions), all $C \times N$ interactions and all N main effects (with possibly some $N \times N$ interactions), where C and N denote control and noise factors, respectively. The $C \times N$ interactions play an important role because the fact that the loss (2) changes for different control factor combinations means that these interactions must exist. Figure 1 displays a simplified relationship between a response Y and one control factor (at two levels) and one noise factor (over an interval) and shows that the effect of the noise factor is substantially smaller at control factor level 1 (C1). Therefore, robust design exploits the existence of interactions between control and noise factors.

Once estimates for the response model effects have been obtained, recommendations for the important control factors settings need to be made. For a simple model with few noise factors, they may be apparent from inspection of the model directly; i.e., by observing what the significant effects are with their signs and magnitudes. Shoemaker et al. [20] gave an example, but for complicated models, this approach may be tedious.

An alternative is to specify some meaningful criterion or loss (2) and use the identified model to evaluate them. The loss can then be evaluated using the estimated response model (1) for some distribution of the noise factors. In practice, because it may be difficult to specify such a distribution, the criterion can be evaluated over the noise combinations given by a noise array. The same noise array from the experiment need not be used, however. For example, instead of a fractional factorial design, the loss could be evaluated using a full factorial design. The noise combinations can also be weighted appropriately to reflect their probabilities of occurrence. Similarly, the loss can be evaluated for all possible settings of the control factors.

For achieving robust reliability, as little dependence as possible on the noise factors is desired. Also high reliability on average is required. Hamada [6] considered criteria based on the linear part ($x^T\beta$) of model (1), i.e., the mean log lifetime for the lognormal regression model. In this paper, reliability will be assessed in terms of the probability of exceeding a certain time T, such as a warranty period. Using model (1), this survival probability can be defined as:

$$Sf\{(\log(T) - x^T\beta)/\sigma\}, \tag{3}$$

where Sf is the appropriate survivor function and $x = (x_{control}, x_{control\times noise}, x_{noise})$. For a given control array combination, these probabilities can be evaluated over all the the noise array combinations with the evaluations representing a sample of probabilities. The sample can be summarized by various quantities such as its mean and standard deviation. Taking a worst case approach, the minimum probability can be used. Based on these criteria, control array combinations with large mean, large minimum probability and small standard deviation are desirable. Analysis of the goggle experiment in Section 7 will illustrate the use of these criteria.

4. ANALYSIS METHODS FOR CENSORED DATA

For model (1), when all the lifetimes are observed, i.e., complete data, the analysis is straightforward using maximum likelihood (ML) estimation (Lawless [12]). Problems arise with analyzing censored data, however, and will be discussed below. Next, a brief overview of some methods for analyzing censored data is given.

One method which continues to be used in practice treats the right-censoring times as actual failure times and then analyzes them by standard methods for complete data. (For interval-censored data, an interval endpoint or midpoint might be used.) Although simple, ignoring the censoring can lead to wrong decisions because the unobserved failure times and right-censoring times may differ greatly depending on the particular factor level combination. A simulation study in Hamada and Wu [8] showed that this method can perform quite poorly by missing some important effects and mis-identifying spurious effects.

The ML estimation methodology can easily handle both failure and censored data. The MLE's for (β, σ) are found by maximizing the following likelihoods: for the Weibull regression model,

$$\mathcal{L}(\beta,\sigma) = \prod_{i \in FAIL} (1/\sigma) exp\{[(y_i - x_i{}^T\beta)/\sigma] - exp[(y_i - x_i{}^T\beta)/\sigma]\} \times \prod_{i \in CEN} exp\{exp[-(y_i - x_i{}^T\beta)/\sigma]\} \quad (4)$$

and for the lognormal regression model,

$$\mathcal{L}(\beta,\sigma) = \prod_{i \in FAIL} (1/\sigma) gaud\{(y_i - x_i{}^T\beta)/\sigma\} \prod_{i \in CEN} gaufc\{(y_i - x_i{}^T\beta)/\sigma\}, \quad (5)$$

where $\{i \in FAIL\}$ denotes those observations which are failures and $\{i \in CEN\}$ denotes those observations which are censored. Standard errors for the MLE's can also be obtained (Lawless [12]). Various commercially available software perform these computations such as SURVIVAL, the SYSTAT survival analysis module (Steinberg and Colla [23]), or the LIFEREG procedure in SAS [19]. Note that for a censored datum, its contribution to the likelihood is simply the probability of being censored. Similarly for an interval-censored datum (a,b), its contribution to the likelihood is the probability of failing between times a and b.

One problem with the ML estimation approach for censored data is that the MLEs may not exist, i.e., at least one parameter estimate is infinite, so that testing cannot be done by comparing the MLEs with their standard errors. Silvapulle and Burridge [21] gave necessary and sufficient conditions for the existence of MLEs for model (1). In the reliability context, Hamada and Tse [7] concluded that estimability problems will tend to occur for the designs discussed in Sections 2 and 3 where the fitted model has nearly the same number of parameters as number of observations.

The estimability problem of the ML approach motivated the Bayesian approach proposed in Hamada and Wu [10]. The Bayesian approach is a natural one because important factor effects might be expected to be large but not infinite. By using proper prior distributions, posterior distributions with finite modes result and can be used to obtain finite estimates. Also, posterior distributions allow the importance of factorial effects to be assessed without using the asymptotic approximations employed by the ML method. Hamada and Wu [10] considered the lognormal regression model (1) and used the natural conjugate prior (Raiffa and Schlaifer [17]):

$$p(\beta,\sigma) = \sigma^{-k} exp\{-(\beta-\beta_0)^T A_0(\beta-\beta_0)/2\sigma^2\} \times \sigma^{-(\nu_0+1)} exp(-\nu_0 s_0^2/2\sigma^2)\ . \qquad (6)$$

The posterior is proportional to the product of the likelihood (5) and the prior (6) and is relatively simple to obtain numerically using recent advances in Bayesian computing. An appropriate choice of the prior parameters $(\nu_0, s_0^2, \beta_0, A_0)$ give a very diffuse or essentially non-informative prior. In fact, the sensitivity of the results based on the current data can be assessed by trying more informative priors. See Hamada and Wu [10] for more details. The Bayesian approach is illustrated in Section 7 in the analysis of the goggle experiment.

5. ANALYSIS OF THE FLUORESCENT LAMP EXPERIMENT

Consider the fluorescent lamp experiment presented in Section 2. Recall that the experiment studied five factors (A-E) using the design in Table 2. Besides the five main effects (A-E), the experimenters thought that the $A \times B$ interaction might be important. Taking the ML approach, a lognormal regression model was fit using the lifetime data in Table 2. Table 6 gives the MLEs and significance levels (p values) for the five main effects (A-E) and the $A \times B$ interaction (with the intercept denoted by Int). Based on these results, the main effects D, B, E and A are important in the order given. Therefore, only four of the five factors are important, with A being only marginally important. The sign of these effects

Table 6: MLEs and P Values of Lognormal Regression Model for the Fluorescent Lamp Experiment

Effect	MLE	P Value
Int	2.939	0.000
A	-0.117	0.059
B	0.201	0.001
AB	-0.049	0.430
C	0.051	0.408
D	-0.273	0.000
E	0.153	0.015
σ	1.590	0.000

suggests that reliability gains can be achieved at recommended setting $A_1B_2D_1E_2$, where the subscript indicates the recommended level.

6. ANALYSIS OF THE THERMOSTAT EXPERIMENT

Next consider the thermostat experiment presented in Section 2 which studied eleven factors (A-K). Taking the ML approach, a lognormal regression model with eleven factor main effects (A-K) was fit using the Table 3 lifetime data whose results are given in Table 7 (Model 1). (At most 12 effects can be fit simultaneously because of the design run size of 12, so that only main effects could be considered in an initial analysis.) Based on Table 7, nine of the factors appear to be important. One potential reason for there being so many was pointed out by Bullington et al. [3]: each group of ten units was produced at the same time so that the variability among the ten units tends to be smaller that if they had been produced at different times. This reduced variability which is used in the statistical testing is one possible explanation for the large number of significant effects. Some additional analysis using Hamada and Wu [8] and [9] which account for the properties of the 12-run Plackett-Burman design used in this experiment suggests the presence of an $E \times H$ interaction, however. Further evidence of an interaction is seen in Table 7 by noting that the MLEs for all factors except E and H have nearly the same magnitude. Consequently, a model (Model 2) was fit in which the factor B main effect (the least significant from Model 1) was dropped and replaced by the $E \times H$ interaction. The results in Table 7 indicate that only

Table 7: MLEs and P Values of Lognormal Regression Models
for the Thermostat Experiment

Model 1			Model 2		
Effect	MLE	P Value	Effect	MLE	P Value
A	-0.312	.0001	A	-0.091	.3890
B	0.221	.0024	EH	0.663	.0024
C	-0.319	.0001	C	-0.098	.3474
D	0.285	.0001	D	0.064	.5174
E	-1.023	.0001	E	-1.023	.0001
F	0.231	.0016	F	0.010	.9219
G	-0.390	.0001	G	-0.169	.1075
H	-0.557	.0001	H	-0.557	.0001
I	-0.332	.0001	I	-0.112	.2872
J	-0.277	.0001	J	-0.056	.5958
K	-0.352	.0001	K	-0.131	.2149

E, H and $E \times EH$ effects are important. An alternate analysis in Bullington et al. [3] found E and H important and recommended $E_1 H_1$. Using the signs of the important effects in Table 7 (Model 2), the same recommendation is obtained. While the original analysis did not account for the possibility of an interaction, the same recommendations result because the two factors have a synergistic effect.

7. ANALYSIS OF THE GOGGLE EXPERIMENT

In the robust reliability experiment to improve night vision goggles presented in Section 3, there were ten control factors (A-J) and a single noise factor (N). Taking the response-model approach, a lognormal regression model (1) can be fit which consists of an intercept, ten C main effects, one N main effect and ten $C \times N$ interactions, where C and N denote control and noise factors, respectively. Using the product array data given in Table 5, the MLEs for this model do not exist. Consequently, the Bayesian approach (Hamada and Wu [10]) was taken. Using a relatively diffuse prior, Table 8 gives the central 0.95 and 0.99 intervals of the marginal posteriors for each effect. The important effects appearing in bold face are those whose central 0.95 intervals do not contain zero; in fact, the 0.99 intervals for all of these except the $I \times N$ interaction do not contain zero. Based on these results main effects, B and D-H, and interactions, $A \times N$, $C \times N$, $E \times N$, $I \times N$, $J \times N$, are important.

Table 8: Posterior Quantiles Using Diffuse Prior
for the Goggle Experiment

	Quantile					Quantile			
Effect	.005	.025	.975	.995	Effect	.005	.025	.975	.995
INT	2.60	2.64	2.99	3.07	N	-0.11	-0.05	0.29	0.35
A	-0.15	-0.08	0.32	0.40	**AN**	-0.68	-0.62	-0.21	-0.16
B	0.74	0.80	1.15	1.24	BN	-0.12	-0.05	0.29	0.35
C	-0.23	-0.18	0.23	0.28	**CN**	-0.82	-0.72	-0.34	-0.27
D	-0.79	-0.71	-0.32	-0.24	DN	-0.22	-0.17	0.24	0.31
E	0.31	0.37	0.77	0.84	**EN**	-0.58	-0.50	-0.07	-0.01
F	0.04	0.10	0.52	0.60	FN	-0.09	-0.03	0.39	0.45
G	-0.60	-0.52	-0.12	-0.07	GN	-0.46	-0.37	0.03	0.09
H	0.27	0.32	0.70	0.82	HN	-0.35	-0.24	0.16	0.23
I	-0.35	-0.27	0.12	0.20	**IN**	-0.53	-0.44	-0.04	0.03
J	-0.25	-0.18	0.23	0.30	**JN**	-0.79	-0.71	-0.33	-0.27
					σ	0.02	0.02	0.07	0.11

For this experiment, the relationship between the response and the control and noise factors is too complicated to make control factor level recommendations simply by inspecting the model. Consequently, the criteria discussed in Section 3 based on the survival probability (3) distribution (mean, standard deviation and minimum probability) can be evaluated over the two levels of the single noise factor for each of the possible combinations of control factors ($1024 = 2^{10}$) and then ranked appropriately (out of 1024, with 1 being the best). In calculating the survival probability (3), the posterior maximizer was used to estimate (β, σ) and the time T was taken to be 100. Table 9 presents the 25 best control factor combinations according to the mean criterion. Based on the first eight rows of Table 9, a good choice of factor levels would be $B_2D_1E_2F_2G_1H_2$ with factors A, C, I and J chosen according to the eight rows. The standard deviation criterion is also small so that these combinations are robust to the noise factor. Also, note that any choice of A, C, I and J will not do as can be seen by the last row of Table 9. Consequently, there are a number of possible factor settings at which high reliability and robust reliability can be achieved.

ACKNOWLEDGEMENTS

This research was supported by General Motors of Canada Limited, the Manufactur-

Table 9: Best Factor Settings for the Goggle Experiment

Settings										Criterion					
										Mean		Std Dev		Min Prob	
A	B	C	D	E	F	G	H	I	J	Value	Rank	Value	Rank	Value	Rank
1	2	1	1	2	2	1	2	2	2	1.000	1	0.000	478	1.000	1
1	2	2	1	2	2	1	2	2	1	1.000	2	0.000	479	1.000	2
2	2	2	1	2	2	1	2	1	1	1.000	3	0.000	524	1.000	3
2	2	1	1	2	2	1	2	1	2	1.000	4	0.000	525	1.000	4
1	2	1	1	2	2	1	2	1	2	1.000	5	0.000	588	1.000	5
1	2	2	1	2	2	1	2	1	1	1.000	6	0.000	590	1.000	6
2	2	2	1	2	2	1	2	2	1	1.000	7	0.000	621	1.000	7
2	2	1	1	2	2	1	2	2	2	1.000	8	0.000	622	1.000	8
1	2	1	1	2	2	2	2	2	2	0.999	9	0.000	639	0.999	9
1	2	2	1	2	2	2	2	2	1	0.999	10	0.000	641	0.999	10
1	2	1	1	2	1	1	2	2	2	0.999	11	0.000	640	0.999	11
1	2	2	1	2	1	1	2	2	1	0.999	12	0.000	643	0.999	12
2	2	1	1	2	2	1	2	2	1	0.999	13	0.001	669	0.998	13
2	2	2	1	2	2	2	2	1	1	0.995	14	0.007	701	0.990	14
2	2	1	1	2	2	2	2	1	2	0.995	15	0.007	702	0.990	15
2	2	2	1	2	1	1	2	1	1	0.995	16	0.007	704	0.990	16
2	2	1	1	2	1	1	2	1	2	0.995	17	0.007	705	0.990	17
1	2	2	1	2	2	1	2	1	2	0.992	18	0.011	714	0.984	18
1	2	1	1	2	2	2	2	1	2	0.947	19	0.076	759	0.893	21
1	2	2	1	2	2	2	2	1	1	0.946	20	0.076	761	0.893	22
1	2	1	1	2	1	1	2	1	2	0.946	21	0.076	762	0.892	23
1	2	2	1	2	1	1	2	1	1	0.946	22	0.077	763	0.891	24
2	2	2	1	1	2	1	2	1	1	0.927	23	0.010	712	0.919	19
2	2	1	1	1	2	1	2	1	2	0.927	24	0.011	713	0.919	20
2	2	1	1	2	2	1	2	1	1	0.926	25	0.104	774	0.853	25

ing Research Corporation of Ontario, and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. G.E.P. Box and N.R. Draper, *Empirical Model-Building and Response Surfaces*, 1987; John Wiley & Sons.

2. G.E.P. Box, W.G. Hunter and J.S. Hunter, *Statistics for Experimenters*, 1978; John Wiley & Sons.

3. R.G. Bullington, S.G Lovin, D.M. Miller and W.H. Woodall, "Improvement of an Industrial Thermostat Using Designed Experiments," *Journal of Quality Technology*, vol 25, 1993, pp 262-270.

4. L.W. Condra, *Reliability Improvement with Design of Experiments*, 1993; Marcel Dekker.

5. A. Dey, *Orthogonal Fractional Factorial Designs*, 1985; Halsted Press.

6. M. Hamada, "Reliability Improvement Via Taguchi's Robust Design," *Quality and Reliability Engineering International*, vol 9, 1993, pp 7-13.

7. M. Hamada and S.K. Tse, "On Estimability Problems in Industrial Experiments with Censored Data," *Statistica Sinica*, vol 2, 1992, pp 381-391.

8. M. Hamada and C.F.J. Wu, "Analysis of Censored Data from Highly Fractionated Experiments," *Technometrics*, vol 33, 1991, pp 25-38.

9. M. Hamada and C.F.J. Wu, "Analysis of Designed Experiments with Complex Aliasing," *Journal of Quality Technology*, vol 24, 1992, 130-137.

10. M. Hamada and C.F.J. Wu, "Analysis of Censored Data from Fractionated Experiments: a Bayesian Approach," University of Waterloo Institute for Improvement in Quality and Productivity Research Report RR-92-11, 1992.

11. R.N. Kackar, "Off-Line Quality Control, Parameter Design, and the Taguchi Method (with discussion)," *Journal of Quality Technology*, vol 17, 1985, pp 176-209.

12. J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, 1982; John Wiley & Sons.

13. F. Montmarquet, "Printed Circuit Drill Bit Design Optimization Using Taguchi's Methods – .013″ Diameter Bits," *Sixth Symposium on Taguchi Methods*, 1988; American Supplier Institute, pp 70-77.

14. W. Nelson, *Applied Life Data Analysis*, 1982; John Wiley & Sons.

15. M.S. Phadke, "Design Optimization Case Studies," *AT&T Technical Journal*, vol 65, 1986, pp 51-68.

16. Plackett, R.L. and Burman, J.P. "The Design of Optimum Multifactorial Experiments," *Biometrika*, vol 33, 1946, pp 305-325.

17. H. Raiffa and R. Schlaifer, *Introduction to Statistical Decision Theory*, 1961; McGraw-Hill.

18. T.L. Reed, "Reduction of Gain Change and Improved Humidity Reliability in Gen III Night Vision Goggle Tube Sub-Assembly Using an L_{12} Taguchi Orthogonal Array," *Sixth Symposium on Taguchi Methods*, 1988; American Supplier Institute, pp 247-268.

19. SAS Institute Inc., *SAS/STAT User's Guide*, Version 6, Fourth Edition, Volume 2, 1989; SAS Institute, Inc.

20. A.C. Shoemaker, K.L. Tsui and C.F.J. Wu, "Economical Experimentation Methods for Robust Design," *Technometrics*, vol 33, 1991, pp 415-427.

21. M.J. Silvapulle and J. Burridge, "Existence of Maximum Likelihood Estimates in Regression Models for Grouped and Ungrouped Data," *Journal of the Royal Statistical Society*, Ser. B, vol 48, 1986, pp 100-106.

22. N. Specht, "Heat Exchanger Product Design via Taguchi Methods," in *Third Symposium on Taguchi Methods*, 1985; American Supplier Institute, pp 302-318.

23. D. Steinberg and P. Colla, *SURVIVAL: a Supplementary Module for SYSTAT,* 1988; SYSTAT Inc.

24. G. Taguchi, *Introduction to Quality Engineering*, 1986; Asian Productivity Organisation.

25. G. Taguchi, *System of Experimental Design,* 1987; Unipub/Kraus International Publications.

26. G. Taguchi and Y.I. Wu, *Introduction to Off-Line Quality Control*, 1980; Japan Quality Control Association.

27. J.C. Wang and C.F.J. Wu, "An Approach to the Construction of Asymmetrical Orthogonal Arrays," *Journal of the American Statistical Association*, vol 86, 1991, pp 450-456.

28. W.J. Welch, T.K. Yu, S.M. Kang and J. Sacks, "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology*, vol 22, 1990, pp 15-22.