Center for Quality and Productivity Improvement
UNIVERSITY OF WISCONSIN
610 Walnut Street
Madison, Wisconsin 53705

(608) 263–2520
(608) 263–1425 FAX

Report No. 62

Robust Product Designs, Part I: First-Order Models with Design × Environment Interactions

George Box and Stephen Jones

May 1990

CQPI
UNIVERSITY OF WISCONSIN
Report No. 62
Report Series
in Quality and Productivity

Robust Product Designs, Part I: First-Order Models with Design × Environment Interactions

George Box
Center for Quality and Productivity Improvement
University of Wisconsin
Madison, Wisconsin

Stephen Jones
Boeing Computer Services
The Boeing Company
Seattle, Washington

ABSTRACT

Professor Genichi Taguchi has emphasized the use of designed experiments in several novel and important applications. In this paper we focus on the use of statistical experimental designs in designing products to be robust to environmental conditions. The engineering concept of robust product design is very important since it is frequently impossible or prohibitively expensive to control or eliminate sources of variation due to environmental conditions. Robust product design enables the experimenter to discover how to modify the design of the product to minimize the effect due to variation from environmental sources.

In experiments of this kind, Professor Taguchi's total experimental arrangement consists of a cross-product of two experimental designs, an inner array containing the design factors and an outer array containing the environmental factors. Except in situations where both of these arrays are small, this arrangement may involve a prohibitively large amount of experimental work.

In the previous paper (Box and Jones, 1990) we showed how this amount of work could be reduced. In that paper we developed a robustness measure for this particular experimental situation. Application of this robustness measure led to classes of experimental designs that generally required significantly fewer runs than the designs proposed by Professor Taguchi.

In the present paper we apply the strategy developed in Box and Jones (1990) to a simple first-order model with interactions between the design and the environmental factors. For this model we derive the robustness measure and give a series of tables of designs that are appropriate for the possible objectives of the experimenter.

KEYWORDS: Parameter design; Robustness; Experimental design; Response surface methodology; Taguchi.

This work was sponsored by the National Science Foundation under grant number DDM-8808138 and by the Vilas Trust of the University of Wisconsin–Madison.

 This report is part of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics) at the University of Wisconsin–Madison. Thesis adviser: George Box.

Table of Contents

First-Order Models with Design × Environment Interactions

CQPI Report No. 62, May 1990

Robust Product Designs, Part I: First-Order Model with Design × Environment Interactions

George Box and Stephen Jones

Professor Genichi Taguchi has emphasized the use of designed experiments in several novel and important applications. In this paper we focus on the use of statistical experimental designs in designing products to be robust to environmental conditions. The engineering concept of robust product design is very important since it is frequently impossible or prohibitively expensive to control or eliminate sources of variation due to environmental conditions. Robust product design enables the experimenter to discover how to modify the design of the product to minimize the effect due to variation from environmental sources.

In experiments of this kind, Professor Taguchi's total experimental arrangement consists of a cross-product of two experimental designs, an inner array containing the design factors and an outer array containing the environmental factors. Except in situations where both of these arrays are small, this arrangement may involve a prohibitively large amount of experimental work.

In the previous paper (Box and Jones, 1990) we showed how this amount of work could be reduced. In that paper we developed a robustness measure for this particular experimental situation. Application of this robustness measure led to classes of experimental designs that generally required significantly fewer runs than the designs proposed by Professor Taguchi.

In the present paper we apply the strategy developed in Box and Jones (1990) to a simple first-order model with interactions between the design and the environmental factors. For this model we derive the robustness measure and give a series of tables of designs that are appropriate for the possible objectives of the experimenter.

1. INTRODUCTION

Much attention has been focused in recent years on the impact of the use of statistics and, in particular, experimental design on the quality of Japanese products and the consequent competitiveness of Japanese industry in the world market-place. A leading quality consultant in that country, who has advocated the widespread use of design of experiments, is Genichi Taguchi. Professor Taguchi has emphasized the use of designed experiments in several novel and important applications. One such application is that of designing a product that is robust to environmental variation.

As an example of such an experiment, consider the set of data given in Table 1, which is typical of tests that have been run in the United States food industry for many years. The manufacturer is seeking the best recipe for a cake mix composed of several ingredients to be sold in a box containing simple

Table 1.
Cake Mix Data

	DESIGN FACTORS			*ENVIRONMENTAL FACTORS*					
Recipe				*T:*	0	−	+	−	+
(*k*)	*F*	*S*	*E*	*t:*	0	−	−	+	+
(0)	0	0	0		6.7	3.4	5.4	4.1	3.8
(1)	−	−	−		3.1	1.1	5.7	6.4	1.3
(2)	+	−	−		3.2	3.8	4.6	4.3	2.1
(3)	−	+	−		5.3	3.7	5.1	6.7	2.9
(4)	+	+	−		4.1	4.5	6.4	5.8	5.2
(5)	−	−	+		5.9	4.2	6.8	6.5	3.5
(6)	+	−	+		6.9	5.0	6.0	5.9	5.7
(7)	−	+	+		3.0	3.1	6.3	6.4	3.0
(8)	+	+	+		4.5	3.9	5.5	5.0	5.4

simple instructions. The quality of the cake made by the purchaser of the cake mix will depend not only on the ingredients but also on the extent to which the baking instructions printed on the box are followed.

In this example, experiments are to be run with three design factors – flour (F), shortening (S), and egg powder (E) – in a 2^3 factorial design with a center point, indicated by (0, 0, 0). It is known the temperature indicator for an oven in a typical domestic stove may be considerably biased up or down. Furthermore, in practice, people frequently overcook or undercook a cake. The manufacturer wants a robust recipe so that the cake will taste reasonably good even when the time (t) and temperature (T) of baking differ somewhat from the recommended levels given in the instructions on the box, and indicated here by (0, 0). To supply data on this, for each recipe the temperature and time of baking are varied about the standard conditions in a 2^2 factorial. The cakes are then baked corresponding to all combinations of these conditions. In this case the ingredients are design variables and the baking conditions given in the instructions printed on the box represent environmental variables. Table 1 gives a typical set of average scores of a taste panel rating the quality of these on a scale of 1 through 7.

Taguchi, in conducting experiments to investigate the robustness of a product to variation from environmental factors, advocates the use of the cross-product of two orthogonal arrays that he calls inner and outer arrays (see Taguchi, 1986; Taguchi and Phadke, 1984; Taguchi and Wu, 1980; Kackar, 1985). The inner array is composed of the factor levels of the variables associated with the design of the product. For each experimental point of the inner array there is an outer array composed of the different levels of the environmental variables. We, however, will use the expressions *design array* instead of inner array and *environmental array* instead of outer array, and refer to the total arrangement as a *cross-product array* since it is a cross-product of the design array and the environmental array. In general, if there are n_1 runs in the design array and n_2 runs in the environmental array, and the runs are made independently, then the cross-product array will require $n_1 \times n_2$ runs for the total experiment. Thus, except where n_1 and n_2 are both small, this would involve a large amount of experimental work.

An alternative approach to robust product experiments that sought to reduce the amount of experimental work required was described in an earlier paper (Box and Jones, 1990). In that paper we first considered the possible objectives for an experiment of this kind. From these objectives we developed performance measures that the investigator might seek to optimize. One of the performance measures that we developed is an overall measure of environmental robustness. The approach is to then entertain a class of models that we suppose, tentatively at least, approximates the underlying system over the (product design × environmental) factor space of interest. We then determine the coefficients of the model that we need to estimate so as to be able to optimize the various performance measures of interest. Having determined the coefficients that it is necessary to estimate, we are then able to choose experimental designs that will yield data that, when the model is true, enable these coefficients to be estimated and so make it possible to optimize the desired performance measures.

In this paper we will apply this strategy described to a simple first-order model with certain interactions. Let us suppose that over the region of interest the behavior of the product design and environmental variables ($\underline{x}$, $\underline{z}$) can be represented by a first-order model in the product design and environmental factors with, in addition, cross-product terms between the product design and the environmental factors.

Therefore, if there are n product design variables, $x_1, x_2, \ldots, x_n$, and m environmental variables $z_1, z_2, \ldots, z_m$, then the response y_{xz} can be represented by

$$y_{xz} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{j=1}^{m} \gamma_j z_j + \sum_{i=1}^{n} \sum_{j=1}^{m} \delta_{ji} x_i z_j \quad (1)$$

with $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

In matrix notation we have

$$y_{xz} = \beta_0 + \underline{x}'\underline{\beta} + \underline{z}'\underline{\gamma} + \underline{z}'\mathbf{D}\underline{x} \quad (2)$$

where $\underline{\beta}$ is $n \times 1$, $\underline{\gamma}$ is $m \times 1$ and $\mathbf{D}$ is $m \times n$.

In Section 2 we derive the Estimation Table appropriate for the first-order model given in Equation (1). The Estimation Table indicates the coefficients in the model of Equation (1) that need to be estimated to meet the possible objectives of interest of the investigator. The potential objectives that we advocated in Box and Jones (1990) were: (a) to minimize $M(x)$, a measure of how far the mean response over the environmental variables is the ideal value; (b) to minimize $V(x)$, a measure of the variability of the response over the environmental variables; (c) to derive the locus of minimum $R(x)$ for different values of λ where $R(x) = \lambda V(x) + (1 - \lambda)M(x)$ is a weighted average of $M(x)$ and $V(x)$; and (d) to minimize $R(x)$ for a specific value of λ_0. To achieve these objectives we need to be able to estimate different constants in the model. In Section 2, expressions for $M(x)$, $V(x)$, and $R(x)$ are obtained and the required coefficients are identified. In particular, we show that to minimize $R(x)$ over $\underline{x}$ we need unbiased estimates of β_0, $\underline{\beta}$, $\underline{\gamma}$, and **D**.

In Section 3 we propose a class of designs that would provide unbiased estimates of the parameters that are needed to minimize $R(x)$. This objective is referred to in the previous paragraph as objective (d). In Section 3.1 we describe a method for constructing a suitable class of designs, which we illustrate with an example. In Section 3.2 we give a series of tables of these designs. An example that illustrates the use of the designs in Section 3.2 to find the levels of the design factors to minimize $R(x)$ is given in Section 4. It will be clear from Section 2 and Table 1 that the designs that are appropriate when the objective is to minimize $R(x)$, that is objective (d), will also be appropriate for describing the locus of minimum $R(x)$, that is objective (c) in the previous paragraph.

In Section 5 we consider designs that would be appropriate if the objective of interest was to minimize $V(x)$. This objective is referred to above as objective (b). We show that the designs given in Section 3 can be used for this situation. In Section 6 we propose a class of designs that would provide unbiased estimates of the parameters that we need to minimize $M(x)$. This objective is referred to above as Objective (a). In Section 6.1 we describe a method for constructing a suitable class of designs, that we illustrate with an example. In Section 6.2 we give a series of tables of these designs.

2. DERIVATION OF THE ESTIMATION TABLE

Consider the first-order model with design × environment interactions given in Equation (1). Now let

$$g_j(\underline{x}) = \left[\frac{\partial y_{xz}}{\partial z_j}\right]_{\underline{z}=\underline{0}} = \gamma_j + \sum_{i=1}^{n} \delta_{ji} x_i \text{ for } j = 1, \ldots, m. \quad (3)$$

Then

$$g'(\underline{x}) = \left\{\left[\frac{\partial y_{xz}}{\partial z_1}\right]_{\underline{z}=\underline{0}}, \left[\frac{\partial y_{xz}}{\partial z_2}\right]_{\underline{z}=\underline{0}}, \ldots, \left[\frac{\partial y_{xz}}{\partial z_m}\right]_{\underline{z}=\underline{0}}\right\} \quad (4)$$

and $g(\underline{x}) = \underline{\gamma} + \mathbf{D}\underline{x}$ is a measure of the change in the response in the direction of $\underline{z}$ at $\underline{z} = \underline{0}$.

Therefore, we have

$$y_{xz} = \beta_0 + \underline{x}'\underline{\beta} + \underline{z}'g(\underline{x}) = y(\underline{x}) + \underline{z}'g(\underline{x}) \quad (5)$$

where $y(\underline{x}) = y(\underline{x}, \underline{0})$.

Now the mean response over the environmental variables, $\bar{y}_x$, is

$$\bar{y}_x = k \int_{R_z} y_{xz} d\underline{z} = y(\underline{x}) \quad (6)$$

where $k^{-1} = \int_{R_z} d\underline{z}$ is an integrating constant.

Let us suppose that the ideal, but possibly unattainable value for the response is some value τ. One possible objective is to choose $\underline{x}$ so that the mean response over the environmental variables is as close as possible to the ideal value τ. A measure of the closeness of the mean response to τ is

$$M(x) = k \int_{R_z} (\tau - \bar{y}_x)^2 d\underline{z} = \left[\tau - \left(\beta_0 + \underline{x}'\underline{\beta}\right)\right]^2. \quad (7)$$

It would also be desirable to have the variation of the response about the mean to be as small as possible. This variation could be due to both random variation and to trends that occur in the response as different values of the environmental variables are considered. This variation could be measured by

$$\begin{aligned} V(x) &= k \int_{R_z} (y_{xz} - \bar{y}_x)^2 d\underline{z} = k \int_{R_z} \left(\underline{z}'g(\underline{x})\right)^2 d\underline{z} \\ &= \frac{1}{3} g'(x) g(x) \end{aligned}, \quad (8)$$

where $g(\underline{x}) = \underline{\gamma} + \mathbf{D}\underline{x}$.

Both $M(x)$ and $V(x)$ are quadratic forms. Now the relative importance of $V(x)$ and $M(x)$ to the experimenter will vary according to the different experimental context. An overall measure of robustness will be a weighted linear combination of $V(x)$ and $M(x)$ and can be written as

$$R(x) = \lambda V(x) + (1 - \lambda) M(x), \quad (9)$$

where $0 \le l \le 1$.

Therefore, the appropriate value for $\underline{x}$ that minimizes the robustness measure will depend on the relative importance that the experimenter gives to the two objectives of having the mean close to the ideal value τ and of having low variation about the mean due to the environmental variation. This trade-off between the two objectives can be represented by λ in Equation (9).

Now if you believe that the quadratic loss is an appropriate representation of the experimental situation then choosing $\underline{x}$ to minimize the quadratic loss,

$$L(x) = k \int_{R_z} (\tau - y_{xz})^2 d\underline{z}, \quad (10)$$

is equivalent to choosing $\underline{x}$ to minimize $R(x)$ with $\lambda = 1/2$. However, it does not appear that this particular weighting of $V(x)$ and $M(x)$ is the only one that should be considered.

Now the shape of the contours of $M(x)$ depend only on $\underline{\beta}$, and the value of the contours of $M(x)$ also depend on β_0 and τ. The shape and value of the contours of $V(x)$ depend only on $\underline{\gamma}$ and **D**. Thus to determine the locus of points $R^*(\overline{x})$ that minimizes $R(x)$ for different values of λ we need to estimate $\underline{\beta}$, $\underline{\gamma}$, and **D**. If we are only concerned with minimizing $\overline{M}(x)$ we only need to estimate $\underline{\beta}$. If we only concerned with minimizing $V(x)$ we only need to estimate $\underline{\gamma}$ and **D**. If we want to minimize $R(x)$ for a particular value of λ then we need to estimate β_0, $\underline{\beta}$, $\underline{\gamma}$, **D**, and we need to know τ. Since minimizing the quadratic loss function $L(x)$ is equivalent to minimizing $R(x)$ with $\lambda = 1/2$, to minimize $L(x)$ we need to estimate the same coefficients as for minimizing $R(x)$, that is we need to estimate the same coefficients as for minimizing $R(x)$, that is we need to estimate β_0, $\underline{\beta}$, $\underline{\gamma}$, **D**, and we need to know τ.

Therefore to achieve the various objectives outlined above we need to be able to estimate the constants indicated in the Estimation Table shown in Table 2. From this table we see that if we regard the first objective as the primary one, that is computation and minimization of $R(x)$, then we are looking for experimental arrangements that permit the estimation of the constant term, β_0, all of the linear effects $\underline{\beta}$ and $\underline{\gamma}$, and all of the two-factor interactions, $\underline{\delta_{ji}}$, between design and environmental factors. Of course, if some of these interactions can realistically be assumed to be negligible then alternative experimental designs that require fewer runs can be derived. Furthermore, if we are interested in minimizing $M(x)$ or $V(x)$, or in tracing the locus of minimum $R(x)$, then smaller experimental designs may be possible since we do not need to estimate as many constants.

3. MINIMIZATION OF $R(x)$

Table 2 gives the constants that need to be estimated to achieve certain objectives. If we want to minimize $R(x)$ for a particular value of λ then we need to be able to estimate β_0, $\underline{\beta}$, $\underline{\gamma}$, **D**, and we also need to know τ. In this section we will find experimental designs that will enable us to satisfy this objective. To determine the locus of points that minimizes $R(x)$, or different values of $\lambda(0 \le \lambda \le 1)$ we need to be able to estimate $\underline{\beta}$, $\underline{\gamma}$, and **D**. From this it is clear that designs that are appropriate if our objective is to minimize $R(x)$ for particular value of λ will also be appropriate if our objective is to determine the locus of minimum $R(x)$.

Table 2.
Estimation Table for the First-Order Model with Design × Environment Interactions

OBJECTIVE	*CONSTANTS*				
Compute/Minimize	β_0	$\underline{\beta}$	$\underline{\gamma}$	**D**	Need to know τ
$R(x) = \lambda_0 V(x) + (1-\lambda_0) M(x)$	X	X	X	X	X
$L(x) = V(x) + M(x)$	X	X	X	X	X
Locus of minimum $R(x)$ for different λ		X	X	X	
$V(x)$			X	X	
$M(x)$		X			

3.1 CONSTRUCTION OF DESIGNS FOR THE MINIMIZATION OF $R(x)$

Addelman (1962) considered the problem of constructing experimental designs that permitted the unbiased estimation of all main effects and a specified number of two-factor interactions. He called his designs *compromise* plans because they offered a compromise (in terms of the number of runs required) between a resolution *V* design, which enables all main effects and all two-factor interactions to be estimated without bias (assuming all higher-order interactions are negligible), and a resolution *III* design that permits the estimation of main effects only, under the assumption that all interactions are negligible.

Addelman's Class 1 Compromise Plans can be used to provide unconfounded estimates of all main effects and all two-factor interactions among k specific factors. His Class 2 Compromise Plans enable the deviation of unconfounded estimates of all main effects, all the two-factor interactions among a subset of k factors and all the two-factor interactions among the l remaining factors. His Class 3 Compromise Plans can be used to provide unconfounded estimates of all main effects and all two-factor interactions that contain any one of k specific factors.

Many authors (Finney, 1945; Plackett and Burman, 1946; Davies and Hay, 1950; Box and Hunter, 1961a) have shown that the columns representing the interactions of a 2^n full factorial arrangement can be used to accommodate additional factors to construct orthogonal first-order designs for a factorial experiment involving up to $(2^n - 1)$ factors. In a similar way (see for example Box and Hunter, 1961b) resolution V designs, which enable the orthogonal estimation of all main effects and all two-factor interactions, can be constructed by taking a full factorial design and assigning additional factors to certain higher-order interactions that are assumed to be negligible.

Addelman constructed his experimental plans by taking full factorial designs and assigning additional factors to columns that represented higher-order interactions. The assignment was done in such a way that these higher-order interactions were not confounded with any main effects or with any of the required interactions.

For the first-order model with design × environment interactions, given in Equations (1) and (2), Table 2 indicates that to minimize $R(x)$ we need to be able to estimate all main effects and all two-factor interactions between the design and the environmental factors. This is an experimental situation that Addelman did not consider. In what follows we will be adapting his approach. We will condition on the number of one type of factor (design or environmental) and then determine the maximum number of the other type of factor that can be accommodated in a specific number of runs.

We observe that the model given in Equation (1) assumes that there are no interactions among the design variables and among the environmental variables. Main effects and the interactions of interest (those between design and environmental factors) will be confounded with these interactions that are assumed to be negligible.

In order to obtain unconfounded estimates of all main effects and all two-factor interactions between design and environmental variables it is necessary that none of the design factors are aliased with the interaction of any two of the environmental factors and none of the environmental factors are aliased with the interaction of any two of the design factors. However, the design factors may be confounded with two-factor interactions of other design factors and the environmental factors may be confounded with two-factor interactions of other environmental factors.

This suggests a method of construction. Let us assume that we have a full factorial design in $k + l$ factors so that we have $2^{(k+l)}$ treatment combinations. Then we can allocate n product design factors and m environmental factors provided $(n + 1)(m + 1) - 1 \leq 2^{(k+l)}$. The method of construction is as follows:

(*i*) choose as the product design (or environmental) factor representations k factors from the full factorial design and as many interactions among the k factors as required. The value k is chosen to be the smallest integer so that $2^n - 1$ is at least as large as the number of product design (or environmental) factors required; and

(*ii*) choose as the environmental (or product design) factor representations the remaining l factors from the factorial design and as many interactions among the l factors as required.

To illustrate the above method we will determine the maximum number of environmental factors that can be accommodated in a 32-run design that has six design factors, such that all main effects of design and environmental factors and all two-factor interactions between design and environmental factors can be estimated without confounding. Consider a full factorial design in the five factors A, B, C, D and E. Now if we take $k = 3$ then the design variables can be represented by A, B, C, AB, AC, and ABC, and the environmental variables can be represented by D, E, and DE. Thus we can accommodate a maximum of three environmental variables in a 32-run design that has six design variables. In fact we can accommodate seven design and three environmental factors by allocating the design variables to A B, C, AB, AC, BC, and ABC, and the environmental variables to D, E, and DE.

3.2 TABLES OF DESIGNS FOR THE MINIMIZATION OF $R(x)$

The following series of tables (Table 3) give the factor representations for all two-level designs using 8, 16, 32, 64, and 128 runs, such that all main effects and all interactions between the product design and the environmental variables can be estimated without confounding. In these tables, the second column gives the number of product design factors, and the third column gives the number of environmental factors that can be accommodated in the experimental design with the stated number of runs and number of product design factors. Similar tables could be produced that gave the number of product design factors that could be accommodated given the number of runs and the number of environmental factors. These tables would only be different to Table 3 in that the second and third column titles and the design and environmental factor representations would be switched.

It can be seen that in some of the designs given in Table 3, some of the interactions among the design factors are unconfounded with the main effects and interactions that we need to estimate. By estimating these interactions among the design factors we can test the validity of the model assumption implicit in equation (1) that these interactions are negligible.

For example, suppose we have five design variables that we want to investigate in an experiment

Table 3.
Factor Representations for Design for the Minimization of R(x)

NUMBER OF RUNS	NUMBER OF DESIGN FACTORS	NUMBER OF ENVIRONMENTAL FACTOR	TOTAL NUMBER OF FACTORS	FACTOR REPRESENTATIONS DESIGN FACTORS; ENVIRONMENTAL FACTORS
2^3	1	3	4	*A; B, C, BC*
	2	1	3	*A, B; C*
	3	1	4	*A, B, AB; C*
2^4	1	7	8	*A; B, C, D, BC, BD, CD, BCD*
	2	3	5	*A, B; C, D, CD*
	3	3	6	*A, B, AB; C, D, CD*
	4	1	5	*A, B, C, ABC; D*
	5	1	6	*A, B, C, AB, ABC; D*
	6	1	7	*A, B, C, AB, AC, BC; D*
	7	1	8	*A, B, C, AB, AC, BC, ABC; D*
2^5	1	15	16	*A; B, C, D, E,* all interactions excluding *A*
	2	7	9	*A, B; C, D, E*
	3	7	10	*A, B, AB; C, D, E*
	4	3	7	*A, B, C, ABC; D, E, DE*
	5	3	8	*A, B, C, AB, ABC; D, E, DE*
	6	3	9	*A, B, C, AB, AC, BC; D, E, DE*
	7	3	10	*A, B, C, AB, AC, BC, ABC; D, E, DE*
	8-15	1	9-16	*A, B, C, D,* and interactions in *A, B, C, D; E*
2^6	2	15	17	*A, B; C, D, E, F,* all interactions without *A, B*
	3	15	18	*A, B, AB; C, D, E, F,* all interactions without *A, B*
	4	7	11	*A, B, C, ABC; D, E, F, DE, DF, EF, DEF*
	5	7	12	*A, B, C, AB, ABC; D, E, F, DE, DF, EF, DEF*
	6	7	13	*A, B, C, AB, AC, BC; D, E, F, DE, DF, EF, DEF*
	7	7	14	*A, B, C, AB, AC, BC, ABC; D, E, F, DE, DF, EF, DEF*
	8-15	3	11-18	*A, B, C, D,* and interactions in *A, B, C, D; E, F, EF*
	16-31	1	17-32	*A, B, C, D, E,* and interactions in *A, B, C, D, E; F*
2^7	2	31	33	*A, B; C, D, E, F, G,* all interactions without *A, B*
	3	31	34	*A, B, AB; C, D, E, F, G,* all interactions without *A, B*
	4	15	19	*A, B, C, ABC; D, E, F, G,* all interactions without *A, B, C*
	5	15	20	*A, B, C, AB, AC; D, E, F, G* all interactions without *A, B, C*
	6	15	21	*A, B, C, AB, AC, BC; D, E, F, G,* all interactions without *A, B, C*
	7	15	22	*A, B, C, AB, AC, BC, ABC; D, E, F, G,* all interactions without *A, B, C*
	8-15	7	15-22	*A, B, C, D,* and interactions in *A, B, C, D; E, F, G, EF, EG, FG, EFG*
	16-31	3	19-34	*A, B, C, D, E,* and interactions in *A, B, C, D, E; F, G, FG*

with 32 runs. Then the table gives as the factor representations for the design factors, *A, B, C, AB, ABC*. With this design we can also estimate the interactions *AC* and *BC* and can conduct a test to determine whether these interactions are negligible. If they are not negligible then the model assumptions of Equation (1) do not hold. Similarly, if, for the same example, we only have two environmental variables of interest, then the interaction between the two variables can be estimated and can be used to test whether the assumption of negligible interactions between environmental variables is invalidated.

4. EXAMPLE

To illustrate how a design can be chosen that would enable us to find the levels at which the design factors to minimize $R(x)$, let us suppose that we require an arrangement with two product design variables, x_1, $x_{,2}$ and three environmental variables, z_1, z_2, z_3, with the response, y_{xz}, represented by the first-order model with interactions given in Equation (1). Therefore, we have

$$y_{xz} = \beta_0 + \sum_{i=1}^{2}\beta_i x_i + \sum_{j=1}^{3}\gamma_j z_j + \sum_{i=1}^{2}\sum_{j=1}^{3}\delta_{ji}x_i z_j \,. \quad (11)$$

From Table 3 we note that an appropriate experimental design can be obtained from a 16-run design by assigning *A*, *B* as design factors and *C, D*, and the interactions *CD* as the environmental factors. Table 4 gives the factor levels for this design and a set of hypothetical data. Table 5 shows the estimates of the effects that can be calculated from the data given in Table 4.

Table 4.
Experimental Design and Response

RUN	*y*	x_1	x_2	z_1	z_2	z_3
1	76	−1	−1	−1	−1	1
2	73	1	−1	−1	−1	1
3	71	−1	1	−1	−1	1
4	68	1	1	−1	−1	1
5	94	−1	−1	1	−1	−1
6	105	1	−1	1	−1	−1
7	41	−1	1	1	−1	−1
8	52	1	1	1	−1	−1
9	58	−1	−1	−1	1	−1
10	55	1	−1	−1	1	−1
11	41	−1	1	−1	1	−1
12	38	1	1	−1	1	−1
13	60	−1	−1	1	1	1
14	75	1	−1	1	1	1
15	59	−1	1	1	1	1
16	74	1	1	1	1	1

Table 5.
Coefficients for Equation (11)

COEFFICIENT	*VALUE*	*COEFFICIENT*	*VALUE*
β_0	65.0	δ_{11}	4.0
β_1	2.5	δ_{12}	−4.0
β_2	−9.5	γ_{21}	0.5
γ_1	5.0	γ_{22}	5.0
γ_2	−7.5	γ_{31}	0.5
γ_3	4.5	γ_{32}	8.0

If we take $\tau = 80$ then the contours for $M(x)$ and $V(x)$, as given by Equations (7) and (8), respectively, are as shown in Figure 1. Also in this figure we indicate the locus of points of minimum $R(x)$ for different values of λ, where $R(x)$ is given by Equation (9).

Let us take $\lambda = 1/2$ so that minimizing $R(x)$ is equivalent to minimizing the loss function, $L(x)$, given in Equation (10). A FORTRAN 77 computer program using the IMSL subroutines DU4INF and DUMINF was written to conduct the minimization by a quasi-Newton method. Using this computer program we conclude that the point $\underline{x}$ that minimizes $L(x)$ is

$$\underline{x} = \begin{bmatrix} 0.38 \\ -1.00 \end{bmatrix}.$$

Now if we take the same model (Equation (11)) and the same coefficients (given in Table 5) but with $\tau = 75$, then the shape of the contours in Figure 1 are unchanged. However, in this case the point $\underline{x}$ that minimizes $L(x)$ is

$$\underline{x} = \begin{bmatrix} 0.00 \\ -0.70 \end{bmatrix}.$$

Thus we see that the minimum point has changed as we have changed τ. For a given value of λ, the value of $\underline{x}$ that minimizes $R(x)$ depends on the value of τ as well as the parameters indicated in Table 2. Thus, if there is a difference of opinion on the magnitude of the ideal value then this will result in a difference in the value of $\underline{x}$ that minimizes $R(x)$. This is an indication of some of the difficulties associated with these measures of environmental robustness.

5. MINIMIZATION OF $V(x)$

Table 2 gives the constants that need to be estimated to achieve certain objectives. If we want to minimize $V(x)$ then we only need to be able to estimate $\underline{\gamma}$ and $\mathbf{D}$. When we compare this list of coefficients with those that we need to estimate if the objective is the minimization of $R(x)$ consider in Section 3, it might be conjectured that since we do not need to be able to estimate the main effects of the design factors, the β_i, to minimize $V(x)$, we will be able to construct smaller designs that will satisfy this objective. However, in this section we will show that this is not the case.

For the designs developed in Section 3.1 we, in effect, constructed separate resolution *III* designs for he design factors and for the environmental factors. Now for minimizing $V(x)$ we might suppose that since we do not need to be able to estimate the main effects of the design factors that we can confound them in alias strings with one another. However, it is clear that if we do this then interactions of an environmental factor with these design factors will also be aliased with one another. Thus we will be unable to estimate the coefficients o $\mathbf{D}$.

To illustrate this, consider a design where two of the design factors, A and B, were aliased with one another. Then the identity relationship would be

$$I = AB = \ldots$$

Now let C be an environmental factor. Then the identity relationship gives

$$AC = BC = \ldots$$

showing that the two design $\times$ environment interactions, AC and BC, are confounded.

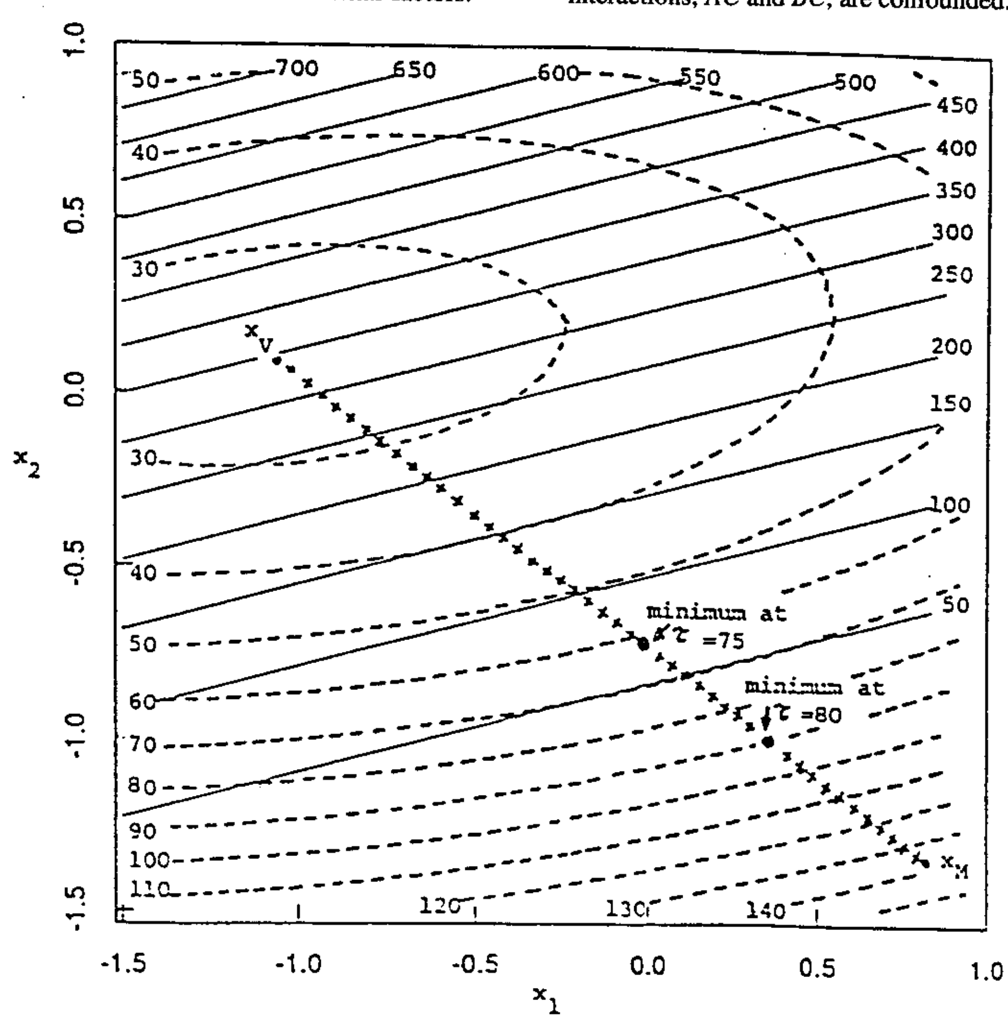

Figure 1. *Contours of M(x) (———), V(x) (– – –), and the locus of minimum R(x) (+ + + + +). (• indicates the minimum of L(x) for two different values of τ.*

Therefore, we conclude that we cannot obtain smaller designs for the objective of minimizing $V(x)$ by confounding the design factor main effects with one another. Thus we recommend that for minimizing $V(x)$ the experimenter should use the designs given in Section 3.

6. MINIMIZATION OF $M(x)$

Table 2 gives the constants that need to be estimated to achieve certain objectives. If we want to minimize $M(x)$ then we only need to be able to estimate $\underline{\beta}$. In this section we will propose a class of designs that can be used to estimate these coefficients and that are smaller than the designs given in Section 3. It should be noted that although we do not need to be able to estimate the environmental factor main effects and the two-factor interactions between the design factors and the environmental factors, we do not assume that these terms are negligible. Thus, the designs cannot confound these factors with the terms that we do need to be able to estimate, namely the main effects of the design factors.

6.1 CONSTRUCTION OF DESIGNS FOR THE MINIMIZATION OF $M(x)$

In order to obtain unbiased estimates of the main effects of the design factors it is necessary that none of the environmental factors are aliased with the interaction of any two of the design factors. This will ensure that none of the design factors will be aliased with a two-factor interaction between a design and an environmental factor. However, the environmental factors can be aliased with interactions of other environmental factors and also with interactions between design factors and environmental factors. The latter is satisfactory in this context but would not have been so in Section 3. The design factors can be aliased with two-factor interactions of other design factors and with two-factor interactions of environmental factors.

This suggests a method of construction. Let us assume that we have a full factorial design in $k + l$ factors so that we have $2^{(k+l)}$ treatment combinations. The method of construction is as follows:

- (*i*) choose as the product design factor representations k factors from the full factorial design and as many interactions among the k factors as required. The value k is chosen to be the smallest integer so that $2^k - 1$ is at least as large as the number of product design factors required. This is, in effect, a resolution *III* design in the k factors;
- (*ii*) choose as the environmental factor representations the remaining l factors from the factorial design and all of the interactions involving any one of the l factors; that is all $2^l - 1$ interactions among the l factors and the $(2^l - 1)(2^k - 1)$ interactions between one of the $2^l - 1$ combinations of the l factors and one of the $2^k - 1$ combinations of the k factors. From this method of construction we see that if we choose k factors from the full factorial design at step (*i*), then the maximum number of environmental factors that we can accommodate in 2^{k+l} runs is $2^k(2^l - 1)$.

To illustrate this method of construction we will determine the maximum number of environmental factors that can be accommodated in a 16-run design that has five design factors. Consider a full factorial design in the four factors *A*, *B*, *C*, and *D*. Now if we take $k = 3$ then the design factors can be represented by *A*, *B*, *C*, *AB*, and *ABC*. Then the environmental factors can be represented by *D*, *AD*, *BD*, *CD*, *ABD*, *ABCD*, *ACD*, and *BCD*. Therefore we can accommodate a maximum of eight environmental factors in a 16-run design that has five design factors. In fact we can accommodate seven design factors and eight environmental factors by allocating the design factors to *A*, *B*, *C*, *AB*, *AC*, *BC*, and *ABC*, and the environmental factors to *D*, *AD*, *BD*, *CD*, *ABD*, *ACD*, *BCD*, and *ABCD*.

6.2 TABLES OF DESIGNS FOR THE MINIMIZATION OF $M(x)$

The following series of tables (Table 6) give the factor representations for all two-level designs using 8, 16, 32, and 64 runs, that will provide unbiased estimates of the main effects of the design factors that can be used to minimize $M(x)$.

7. CONCLUSION

In this paper we have considered the problem of designing products that are robust to variation that is due to environmental factors. This is an important issue since it is frequently impossible or inappropriate for the product manufacturer to control the variation that comes from the environment. Thus, it is advantageous for the manufacturer to design a product that is insensitive to environmental variation. An issue of concern for the experimenter is the amount of experimental work that is usually required in the cross-product designs that Taguchi and others have advocated for robust product design.

Table 6.
Factor Representations for Design for the Minimization of M(x)

NUMBER OF RUNS	NUMBER OF DESIGN FACTORS	NUMBER OF ENVIRONMENTAL FACTOR	TOTAL NUMBER OF FACTORS	FACTOR REPRESENTATIONS DESIGN FACTORS; ENVIRONMENTAL FACTORS
2^3	1	6	7	A; B, C, BC, AC, ABC
	2	4	6	A, B; C, AC, BC, ABC
	3	4	7	A, B, AB; C, AC, BC, ABC
2^4	1	14	15	A; B, C, D, and all interactions
	2	12	14	A, B; C, D, CD and all interactions except AB
	3	12	15	A, B, AB; C, D, CD, and all interactions except AB
	4	8	12	A, B, C, ABC; D, AD, BD, CD, ACD, ABD, BCD, ABCD
	5	8	13	A, B, C, AB, ABC; D, AD, BD, CD, ACD, ABD, BCD, ABCD
	6	8	14	A, B, C, AB, AC, BC; D, AD, BD, CD, ACD, ABD, BCD, ABCD
	7	8	15	A, B, C, AB, AC, BC, ABC; D, AD, BD, CD, ACD, ABD, BCD, ABCD
2^5	1	30	31	A; B, C, D, E, and all interactions
	2	28	30	A, B; C, D, E, and all interactions except AB
	3	28	31	A, B, AB; C, D, E, and all interactions except AB
	4	24	28	A, B, C, ABC; D, E, and all interactions involving D, E
	5	24	29	A, B, C, AB, ABC; D, E, and all interactions involving D, E
	6	24	30	A, B, C, AB, AC, BC; D, E, and all interactions involving D, E
	7	24	31	A, B, C, AB, AC, BC, ABC; D, E, and all interactions involving D, E
	8-15	16	24-31	A, B, C, D, and their interactions; E, and all interactions involving D; E
2^6	1	62	63	A; B, C, D, E, F, and all interactions
	2	60	62	A, B; C, D, E, F, and all interactions excluding AB
	3	60	63	A; B, AB; C, D, E, F, and all interactions excluding AB
	4	56	60	A, B, C, ABC; D, E, F, and all interactions involving D, E, F
	5	56	61	A, B, C, AB, ABC; D, E, F, and all interactions involving D, E, F
	6	56	62	A, B, C, AB, AC, BC; D, E, F, and all interactions involving D, E, F
	7	56	63	A, B, C, AB, AC, BC, ABC; D, E, F, and all interactions involving D, E
	8-15	48	56-63	A, B, C, D, and their interactions; E, F, and all interactions involving E, F
	16-31	32	48-63	A, B, C, D, E, and their interactions; F, and all interactions involving F

The choice of an appropriate design depends on the experimental circumstances. Box and Draper (1987, p. 502-503) list a series of experimental circumstances that should be considered by the investigator when selecting a response surface design. Many of the same considerations apply to the conducting of experiments to design products that are robust to environmental variation. In this paper we have illustrated the fact that the choice of experimental design should also depend on the objective of the experiment. For example, the investigator who is only interested in obtaining a mean performance close to the ideal value will choose a different design to the investigator whose objective is to achieve low variability about the mean. The designs that we have given in this paper are appropriate when the objective of the experimenter is to minimize a weighted average of two performance measures; the one measures how good is the product design on the average when it is exposed to different environmental conditions, the other measures how much variation there is in the performance at different environmental conditions. Experimental designs that require fewer runs may be appropriate if the objective of interest is different to that of minimizing a weighted average of these two performance measures.

In this paper we have supposed, tentatively at lest, that the model that approximates the underlying system over the (design × environmental) factor space is linear in the design and environmental factors with cross-product terms between those factors. We have determined the coefficients of the model that it is necessary to be able to estimate in order to optimize performance measures of interest. We have given experimental designs that will yield data that, when the model is true, enable those coefficients to be estimated and so make it possible to optimize the desired performance measures. These ideas have been illustrated with an example.

ACKNOWLEDGMENT

This work was sponsored by the National Science Foundation under grant DDM-8808138, and by the Vilas Trust of the University of Wisconsin-Madison.

REFERENCES

Addelman, S. (1962), "Symmetrical and asymmetrical fractional factorial plans," *Technometrics*, Vol. 4. p. 47-58.

Box, G.E.P. and N.R. Draper (1987), *Empirical Model-Building and Response Surfaces.* John Wiley; New York.

Box, G.E.P. and J.S. Hunter (1961a), "The 2^{k-p} fractional factorial designs, I", *Technometrics*, Vol. 3. p. 311-351.

Box, G.E.P. and J.S. Hunter (1961b), "The 2^{k-p} fractional factorial designs, II" *Technometrics*, Vol. 3. p. 449-458.

Box, G.E.P. and S.P. Jones (1990), "Designing products that are robust to the environment," (in preparation).

Davies, O.L. and W.A. Hay (1950), "The construction and uses of fractional factorial designs in industrial research," *Biometrics*, Vol. 6. p. 233-249.

Finney, D.J. (1945), "The fractional replication of factorial arrangements," *Annals of Eugenics*, Vol. 12. p. 291-301.

Kackar, R.N. (1985), "Off-line quality control, parameter design and the Taguchi method," *Journal of Quality Technology*, Vol. 17. p. 176-188, discussion p. 189-209.

Plackett, R.L. and J.P. Burman (1946), "The design of optimum multifactorial experiments," *Biometrika*, Vol. 33. p. 305-325.

Taguchi, G. (1986), *Introduction to Quality Engineering: Designing Quality into Products and Processes.* Kraus International Publications; White Plains, New York.

Taguchi, G. and M.S. Phadke (1984), "Quality engineering through design optimization," *Conference Record*, Vol. 3. IEEE GLOBECOM 1984 Conference. p. 1106-1113.

Taguchi, G. and Y. Wu (1980), *Introduction to Off-Line Quality Control.* Central Japan Quality Control Association; Nagaya, Japan.