S. Ghosh and C. R. Rao, eds., *Handbook of Statistics, Vol. 13*

Response Surface Designs

Norman R. Draper and Dennis K. J. Lin

1. Response surfaces and models

Suppose we have a set of data containing observations on a response variable y and k predictor variables $\xi_1, \xi_2, \ldots, \xi_k$. A response surface model is a mathematical model fitted to y as a function of the ξ's in order to provide a summary representation of the behaviour of the response, as the predictor variables are changed. This might be done in order to (a) optimize the response (minimize a cost, maximize a percentage yield, minimize an impurity, for example), (b) find what regions of the ξ-space lead to a desirable product (viscosity within stated bounds, transparency not worse than a standard, appropriate color maintained, for example), or (c) gain knowledge of the general form of the underlying relationship with a view to describing options such as (a) and (b) to customers.

When the mechanism that produced the data is either unknown or poorly understood, so that the mathematical form of the true response surface is unknown, an *empirical model* is often fitted to the data. An empirical model is usually linear in the parameters and often of polynomial form, either in the basic predictor variables or in transformed entities constructed from these basic predictors. The purpose of fitting empirical models is to provide a mathematical French curve that will summarize the data. This chapter will discuss only design of experiments for such empirical models.

There is another useful type of model, however, the mechanistic model. If knowledge of the underlying mechanism that produced the data *is* available, it is sometimes possible to construct a model that represents the mechanism reasonably well. An empirical model usually contains fewer parameters, fits the data better, and extrapolates more sensibly. (Polynomial models often extrapolate poorly.) However, mechanistic models are often nonlinear in the parameters, and more difficult to formulate, to fit, and to evaluate. For information on this topic, see Bates and Watts (1988) and Seber and Wilde (1989).

When little is known of the nature of the true underlying relationship, the model fitted will usually be a polynomial in the ξ's. The philosophy applied here is that we are approximating the true but unknown surface by low-order (equivalently: low degree) terms in its Taylor's series expansion. Most used in practice are polynomials of first and second order. The first-order model is

$$y_u = \beta_0' + \beta_1'\xi_{1u} + \beta_2'\xi_{2u} + \cdots + \beta_k'\xi_{ku} + \varepsilon_u, \tag{1}$$

where $(y_u, \xi_{1u}, \xi_{2u}, \ldots, \xi_{ku}), u = 1, 2, \ldots, n$, are the available data and where it is usually tentatively assumed that the errors $\varepsilon_u \sim N(0, \sigma^2)$ and are independent. Such assumptions are always carefully checked by examining the residuals (the differences between observed and predicted values of y) for possible contradictory patterns. The second-order model contains additional terms

$$\beta'_{11}\xi_{1u}^2 + \beta'_{22}\xi_{2u}^2 + \cdots + \beta'_{kk}\xi_{ku}^2 + \beta'_{12}\xi_{1u}\xi_{2u} + \cdots + \beta'_{k-1,k}\xi_{k-1,u}\xi_{ku}. \quad (2)$$

Polynomial models of order higher than 2 are rarely fitted, in practice. This is partially because of the difficulty of interpreting the form of the fitted surface, which, in any case, produces predictions whose standard errors are greater than those from the lower-order fit, and partly because the region of interest is usually chosen small enough for a first- or second-order model to be a reasonable choice. Exceptions occur only when two or three ξ's are used. When a second-order polynomial is not adequate, and often even when it is, the possibility of making a simplifying transformation in y or in one or more of the ξ's would usually be explored before reluctantly proceeding to higher order. A more parsimonious representation involving fewer terms is generally more desirable.

Coding

In actual applications, it is common practice to code the ξ's via $x_{iu} = (\xi_{iu} - \xi_{i0})/S_i$, $i = 1, 2, \ldots, k$, where ξ_{i0} is some selected central value of the ξ_i range to be explored, and S_i is a selected scale factor. For example, if a temperature (T) range of 150–170°C is to be covered using three levels 150, 160, 170°C, the coding $x = (T - 160)/10$ will code these levels to $x = -1, 0, 1$, respectively. The second-order model would then be recast as

$$\begin{aligned} y_u = {} & \beta_0 + \beta_1 x_{1u} + \cdots + \beta_k x_{ku} \\ & + \beta_{11} x_{1u}^2 + \cdots + \beta_{kk} x_{ku}^2 \\ & + \beta_{12} x_{1u} x_{2u} + \cdots + \beta_{k-1,k} x_{k-1,u} x_{ku} + \varepsilon_u \end{aligned} \quad (3)$$

or

$$y = X\beta + \varepsilon$$

in matrix form, and would usually be fitted by least squares in that form. Substitution of the coding formulas into (3) enables the β''s to be expressed in terms of the β's, if desired.

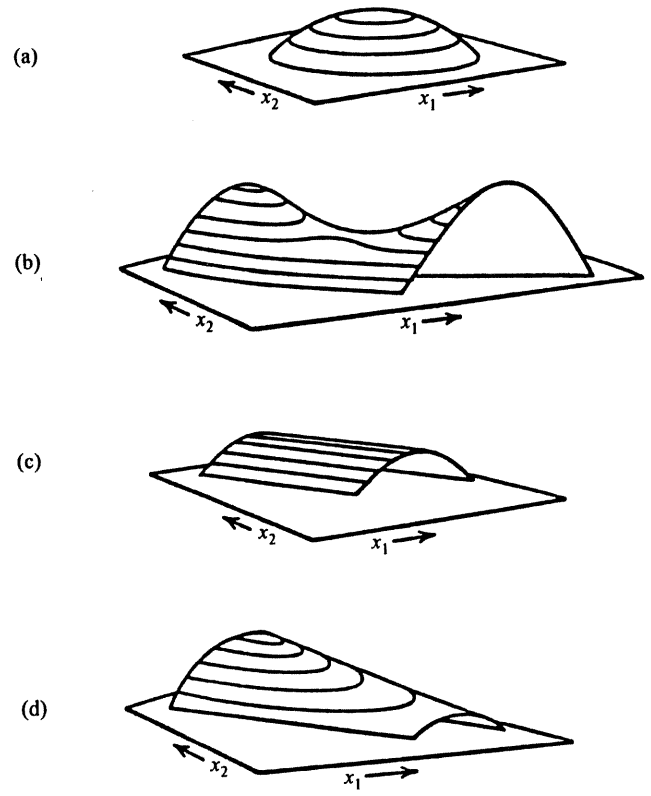


Fig. 1. Examples of surfaces representable by a second-degree equation: (a) simple maximum, (b) saddle (or col or minimax), (c) stationary ridge, and (d) rising ridge.

2. Second order surfaces

A model of the form (3) can represent a variety of surfaces, one of which will best fit a given set of data. Figure 1 shows examples of the four basic types that occur when $k = 2$. "Upside down" versions can also occur. For example, linked with the simple maximum of Figure 1(a) (a hill) is a simple minimum (a hollow). The "upside down" version of a rising ridge (Figure 1(d)) is a falling valley, and so on. In higher dimensions the drawings become more complicated, but the two dimensional sections of higher dimensional surfaces are always one of the basic types illustrated in Figure 1, together with their upside down versions. For additional details on surface types in various dimensions, including the important reduction of such surfaces to canonical form, see Davies (1978) or Box and Draper (1987).

3. Criteria for experimental designs

First and second order models have proved valuable in a variety of subject areas. Sometimes they are fitted to data that have been obtained by observing a running process. More typically, the data will result from a carefully planned series of experimental runs (individual experiments) which, taken as a whole, are called the *experimental design*, and denoted, in the coded x-space by the n sets of values $(x_{1u}, x_{2u}, \ldots, x_{ku})$, $u = 1, 2, \ldots, n$. These coordinates define a pattern of n points in a k-dimensional space. A response surface design is simply an experimental arrangement of points in x-space that permits the fitting of a response surface to the corresponding observations y_u. We thus speak of first-order designs (if a first-order surface can be fitted), second-order designs, and so on. Obviously, a design of a particular order is also necessarily a design of lower order.

The choice of a response surface design is thus one of selecting a set of suitable points in k-dimensional x-space according to some preselected criterion or criteria of goodness. The technical literature of experimental design contains many discussions of so-called "optimal designs". However, skepticism is called for in reading many books and papers, because their authors often concentrate on one criterion only (and sometimes one that by practical experimental standards is inappropriate) and then derive the best designs under that single criterion. While this often provides interesting mathematics, it does not necessarily constitute useful practical advice. There are many possible desirable characteristics for a "good" response surface design. Box and Draper (1987) gave 14 such characteristics. The design should:

1. Generate a satisfactory distribution of information about the behaviour of the response variable throughout a region of interest, R;
2. Ensure that the fitted value at $\boldsymbol{x}, \hat{y}(\boldsymbol{x})$, be as close as possible to the true value at $\boldsymbol{x}, \eta(\boldsymbol{x})$;
3. Give good detectability of lack of fit;
4. Allow transformations to be estimated;
5. Allow experiments to be performed in blocks;
6. Allow designs of increasing order to be built up sequentially;
7. Provide an internal estimate of error;
8. Be insensitive to wild observations and to violation of the usual normal theory assumptions;
9. Require a minimum number of experimental points;
10. Provide simple data patterns that allow ready visual appreciation;
11. Ensure simplicity of calculation;
12. Behave well when errors occur in the settings of the predictor variables, the x's;
13. Not require an impractically large number of predictor variable levels;
14. Provide a check on the "constancy of variance" assumption.

It is impossible for a design to satisfy all the characteristics simultaneously. Indeed some characteristics work against others, for example, (9) conflicts with the need to add extra points to attain (3). In any given experimental situation, certain characteristics will loom larger than others, depending on what the desired objectives are. If we wish to examine a large number of variables and pick the most effective few, for example, criteria (1)–(4) may not be of much interest temporarily. A good statistician will be able to size up the current situation and give emphasis to the various criteria accordingly. Moreover, certain types of designs satisfy many, if not all of the criteria. Such designs are especially valuable.

4. Sequential experimentation

Although each experimental design is an important step in itself, experimentation is rarely a one-step process. Typically one proceeds through a series of steps:

1. Identify all the variables currently of interest. If there are many, because current knowledge is sparse, consider a screening experiment that will enable us to eliminate unimportant predictor (x) variables and retain influential ones. (The WHICH? stage.) We would usually initially consider the possibility that a first-order model might be satisfactory, and perform a first-order design. A simple but good choice (see Box, 1952) would be a regular simplex design with one or more center points. The general regular simplex in k dimensions has $n = k+1$ points (runs) and can be oriented to have its coordinates given as in Table 1, where $a_i = \{cn/[i(i+1)]\}^{1/2}$, and c is a scaling constant to be selected. Alternatively, a two-level factorial or fractional factorial, or a Plackett and Burman design with added center point(s) would be excellent. In all cases, the center point(s) average response can be compared to the average response at the noncentral points to give a measure of nonplanarity. For additional details, see Box et al. (1978, p. 516) or Box and Draper (1987), as well as Section 6 below.

2. If only a few (of many) x-variables were effective, the results could be projected into those fewer x-dimensions, and a first order surface could be refitted. Then, if

Table 1
The rows are the coordinates of the $(k+1)$ points of a simplex design in k dimensions

x_1	x_2	x_3	$\cdots$	x_i	$\cdots$	x_k
$-a_1$	$-a_2$	$-a_3$	$\cdots$	$-a_i$	$\cdots$	$-a_k$
a_1	$-a_2$	$-a_3$	$\cdots$	$-a_i$	$\cdots$	$-a_k$
0	$2a_2$	$-a_3$	$\cdots$	$-a_i$	$\cdots$	$-a_k$
0	0	$3a_3$	$\cdots$	$-a_i$	$\cdots$	$-a_k$
.	.	0		.		.
.	.	.		.		.
.	.	.		.		.
				ia_i		
				0		
				.		
				.		
				.		
0	0	0		0		ka_k

the reduced first-order surface fitted well, one could either interpret its nature if the local relationship were being sought, or else move out along a path of steepest ascent (or descent) if improved conditions were sought; see Box and Draper (1987). If the first-order surface were an inadequate representation of the local data, either initially or after one or more steepest ascent(s) (or descent(s)), it would be sensible to consider transformations of the response and/or predictor variables that *would* allow a first-order representation. When the possibilities of using first-order surfaces had been exhausted, one would then consider the need for a second-order surface. Second order designs will be discussed in Sections 7, 8 and 12.

3. If a second order surface were deemed inadequate, it would again be sensible to seek suitable variable transformations. Proceeding to models of order higher than second would be a last resort in most applications.

5. "Value for money" in designs

When we have a limited budget for experimentation (typically the case in practice), we wish to choose our experimental design to get full value for what we spend. We can assess value in a design by considering what the degrees of freedom (the "money") buy for us, that is, what benefits they provide. Consider the first order simplex design for k factors with an additional n_0 center points, a total of $n = k + 1 + n_0$ degrees of freedom (df) available. Of these, $(k+1)$ are used to estimate the coefficients of the first order model, $(n_0 - 1)$ provide a pure error estimate of σ^2, and 1 df provides a test for non-planarity. So we have "good value for money" here, in the sense that the design performs well on criteria 1, 3, 6, 7, 9 and 11. Of course, the design can be criticised, in various degrees, with respect to other criteria. It makes excellent sense, in considering any specific design to evaluate exactly what the available degrees of freedom will provide in terms of model estimation, pure error, and lack of fit, particularly when a choice between competing designs needs to be made.

When considering the next experimental design we most often choose the location of its center as the point representing the current "best" (whatever that is defined to mean) conditions. Three common general objectives of response surface methodology are:

1. To find the local nature of the relationship between the response and the predictors and so "explain" the response's behavior. It may, for example, be desired to keep the response within specifications requested by a customer, and/or to check whether the predictor variable settings are critical and sensitive.
2. To move from the current "best" conditions to better conditions (lower cost, higher yield, improved tear resistance, and so on).
3. To use the fitted surface as an intermediate step to mechanistic understanding of the underlying process.

For other possible objectives, see Herzberg (1982).

6. Screening designs and projection properties

Practitioners projection properties are constantly faced with distinguishing between the factors that have an actual effect and those factors whose effects are due to random error. Typically, many possible factors are suggested for investigation, but it is often anticipated that only a "small" subset of these (k, say) will be effective, the so called "effect sparsity" situation. Thus, it is believed that perhaps even a smaller subset of the specific terms in (1) and (2) are actually needed to describe the behaviour of the response. A design suitable for screening out the k relevant factors from the q total factors is called a screening design. See Box at al. (1978, pp. 545–546).

Screening designs are typically used in the initial stages of an experimental investigation. (Sometimes, several responses are measured in each experiment.) Because of their relative simplicity of use, two-level screening designs are very popular in practice. For example, the 2^{q-p} fractional factorial designs (that is, a 2^{-p} fraction of a 2^q two-level factorial design; see Box et al., 1978) and the Plackett and Burman (1946) designs are widely used. When such an n-run screening design is employed, it is not expected that every factor will show up as important, merely a subset. This permits the use of fractionated designs with complicated alias structures. After the initial analysis, the whole design is then projected into a lower dimensional space which contains only the k apparently important factors.

We employ throughout the standard notation introduced by Box and Hunter (1961) in which I represents an n-run column of plus signs and, for example, 123 represents a column of signs determined by taking the product of the signs $\pm$ in columns 1, 2, and 3 of the two-level factorial design, where $-$ and $+$ denote the two levels of the factor allocated to any column.

Consider, for a simple example, the eight run, four factor 2^{4-1}_{IV} design $I = 1234$ consisting of the runs $(x_1, x_2, x_3, x_4) = (----), (+--+), (-+-+), (++--), (--++), (+-+-), (-++-), (+++)$. Suppose one of the four factors is inactive; we do not know which factor it might be. No matter *which* factor it is, if we drop that variable from the design, the remaining three variables are represented by a full 2^3 factorial. For example, let us drop variable x_1 by removing the first sign from the parentheses above. We are left with $(---), (--+), (+-+), (+--), (-++), (-+-), (++-), (+++)$, a 2^3 factorial design in (x_2, x_3, x_4). Such a design will remain no matter which variable is dropped from the design.

Details of the projection properties of two-level designs will be discussed next. This knowledge also allows us to see what *additional* runs can be of value, after the results of the initial screening are available. In addition, it provides insight into how to allocate the design variables to the factors that are thought, a priori, to be important.

Projections of 2^{q-p} designs

When a 2^{q-p} screening design is used, all projections are either standard two-level full factorials or fractional factorials. For $k = 2$ and $n \geqslant 4$, the projection is always a 2^2 design, with multiplicity $n/4$. For $k = 3$ and $n \geqslant 8$, there are two types of

projections: a 2^3 design, with multiplicity $n/8$, or a 2^{3-1}_{III} design ($I = \pm 123$), with multiplicity $n/4$. For $k = 4$ and $n \geqslant 16$, there are three possibilities: a 2^4 design, with multiplicity $n/16$, or a 2^{4-1}_{IV} design ($I = \pm 1234$), with multiplicity $n/8$, or a 2^{4-1}_{III} design ($I = \pm 123$), with multiplicity $n/8$.

For $k = 5$, the possibilities for the projected designs are given below:

Type	Generators	Multiplicity
2^5	–	$n/32$
2^{5-1}_{V}	$I = \pm 12345$	$n/16$
2^{5-1}_{IV}	$I = \pm 1234$	$n/16$
2^{5-1}_{III}	$I = \pm 123$	$n/16$
2^{5-2}_{III}	$I = \pm 124 = \pm 1235$	$n/8$

The extension to $k \geqslant 6$ is similar and straightforward. However, if the projected design remains resolution III, estimated main effects are confounded with two-factor interactions. The usual advice given in such circumstances to eliminate such blurring is to "fold over" the design (Box et al., 1978, pp. 340, 399), that is, repeat the projected design with all signs reversed. Foldover always converts a resolution III design into a resolution IV design (see Box et al., 1978, p. 398). It also doubles the size of the experiment, however, which can be disadvantageous.

Plackett and Burman screening designs

Table 2 shows a 12-run Plackett and Burman design, obtained as follows.

(a) Write down the set of signs $++-+++---+-$, provided by Plackett and Burman (1946).

Table 2
A 12-run Plackett and Burman design

Run No.	Factors										
	1	2	3	4	5	6	7	8	9	10	11
1	+	+	–	+	+	+	–	–	–	+	–
2	–	+	+	–	+	+	+	–	–	–	+
3	+	–	+	+	–	+	+	+	–	–	–
4	–	+	–	+	+	–	+	+	+	–	–
5	–	–	+	–	+	+	–	+	+	+	–
6	–	–	–	+	–	+	+	–	+	+	+
7	+	–	–	–	+	–	+	+	–	+	+
8	+	+	–	–	–	+	–	+	+	–	+
9	+	+	+	–	–	–	+	–	+	+	–
10	–	+	+	+	–	–	–	+	–	+	+
11	+	–	+	+	+	–	–	–	+	–	+
12	–	–	–	–	–	–	–	–	–	–	–

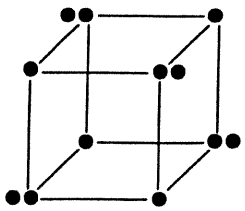


Fig. 2. A 12-run, 11 factor Plackett and Burman design projected into any three dimensions.

(b) Permute the signs in 11 rows total, by taking the sign from the right hand side and moving it to the left hand side.

(c) Add a 12th row of all minus signs.

For $n \leqslant 24$, all of the Plackett and Burman designs can be obtained by such a cyclic permutation. The signs for the first rows are:

$n = 8$: $+++-+--$,
$n = 12$: $++-+++---+-$,
$n = 16$: $++++-+-++--+---$,
$n = 20$: $++--+++-+-+-----++-$,
$n = 24$: $++++++-+-++--++--+-+----$.

For $n = 8$ and 16, we obtain a standard 2^{q-p} design, so these cases are covered by the previous section. For the projection of the 12-run design in any k of the 11 factor dimensions, we select k columns and examine the design that results by ignoring the other $11 - k$ columns. For example, suppose $k = 3$ and we select the 1, 2, 3 columns. The reduced 12-point design consists of a 2^3 design plus a 2^{3-1} design with $I = 123$, shown in Figure 2. This very desirable arrangement provides complete coverage of all the factorial effects plus additional pure error information obtained at four different locations well spread out over the experimental region. Moreover, **no matter which three factors are designated as the survivor columns**, a similar design is always obtained, that is, a 2^3 plus a 2^{3-1} with $I = \pm ABC$ where A, B, C represent any three of the eleven factors. See, also, Lin and Draper (1992) and Box and Bisgaard (1993).

For $k = 3$, and $n = 20$, two types of projections can occur:

1. A 2 : 3 type. (This means two full 2^3 factorials and an additional 2^{3-1}. At the corners of the cube there are either two or three points.)

2. A 1 : 4 type. (This means a 2^3 factorial and three identical 2^{3-1} designs. At each corner of the cube there is either one point or there are four points.)

The notation "$(r : s)$" means r points lie at four of the "2^{3-1} locations" and s points lie at the other four. 2^{3-1} locations are always defined by $I = \pm$ the relevant three factor interaction. Another point to note is that, in some cases, we can proceed from a three column n runs projection to a three columns $(n + 4)$ runs projection by simply adding a 2^{3-1} design. For $n = 20$, a (2 : 3) can be converted into either a

(3 : 3) or a (2 : 4) with $n + 4 = 24$, depending on which 2^{3-1} is added. Similarly, a (1 : 4) can become a (2 : 4) or a (1 : 5); however, the latter is not a three-column projection of a 24-run Plackett and Burman design. Other possibilities for $k = 3$ are given in Lin and Draper (1991).

Projections into $k = 4$ dimensions

For $n = 12$, one could complete a 2_{IV}^{4-1} design by adding one run; there are five additional runs as well. For example, suppose we use columns 1–4 of Table 3. The eight runs 1, 3, 5, 6, 7, 9, 10, and 11 all have a negative product of signs, but runs 3 and 11 are identical, $(+ - - +)$. The new run $(- + - -)$ completes the 2_{IV}^{4-1} with $I = -1234$. This additional point is always uniquely determined as the foldover complement of the duplicate point. An alternative to this one run addition would be to fit a "main effects plus two factor interactions" model to all the initially available data.

For $n = 20$, only three types of projections exist (apart from sign changes in the columns, permutations of the columns, and rearrangements of the rows). For $n = 24$, there are four types of projections. Two types of projections provide a full 2^4 design plus a 2_R^{4-1}; for one R = IV, and for the other R = III. Another two types of projections require two additional runs to complete a full 2^4 design. Which projection is actually attained depends on the specific four factors retained after analysis.

Projections into $k = 5$ dimensions

For $n = 12$, two types of designs are possible, one with a repeat run pair ("type 5.1", say) and one with a mirror image pair ("type 5.2"); see Draper (1985, Table 2). A number of possibilities exist for supplementing these designs. To determine which of the two design types has been obtained via projection, one must check to see if the specific design has a repeat run pair, or a mirror image run pair, an easy thing to do.

For example, if we choose columns 1–5 of Table 2, we see that runs 7, $(+ - - - +)$, and 10, $(- + + + -)$, are mirror image runs. Thus we have a design of type 5.2 which we could convert to a standard form, in which the mirror image runs are $(- - - - -)$ and $(+ + + + +)$, by changing the signs in either columns 1 and 5 or in 2, 3, and 4 and perhaps rearranging the columns appropriately. Even without making those changes, it is clear that the product of signs in the columns 1, 2, 3, 4, and 5 is "−" for runs 1, 5, 7, 8, 11, and 12, and "+" for the remaining six runs. A 2_V^{5-1} can thus be produced in two alternative ways, by adding 10 runs with the same signed products in each case.

In examining the possibilities for Design 5.2, we discover that the 32 runs of a 2^5 design can be divided as follows:

(a) Into a 12-run portion and a 20-run portion so that the two portions are the projections into five dimensions of (respectively) 12-run and 20-run Plackett and Burman designs.

(b) Into 8-run and 24-run portions which are projections into five dimensions of 8 and 24 run Plackett and Burman designs.

In all such cases, the model appropriate to the completed 2_R^{q-p} design could be fitted by least squares, and the runs already made in addition to the 2_R^{q-p} runs will provide some residual degrees of freedom in an analysis of variance table.

All the nine possible different projected designs for $n = 20$ as well as the nine projections that occur for $n = 24$ are given in detail in Lin and Draper (1991, 1994).

Non-equivalent Hadamard matrices for $n = 16, 20$

By adding a column of 1's to a Plackett and Burman design, we obtain a Hadamard matrix H which satisfies $H'H = nI$. For $n = 12$, H is unique, but for higher n this is not true. Non-equivalent Hadamard matrices have different projection properties. We illustrate using the cases $n = 16$ and $n = 20$.

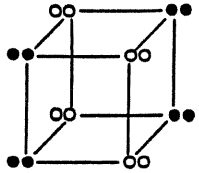
There are five non-equivalent Hadamard matrices for $n = 16$; see Hall (1961). Only one of these corresponds to a Plackett and Burman design, that is, only one (called H16-1 here, and I by Hall) provides a 2^{q-p} 16-run design of the type whose projections were studied in the previous section. We now briefly discuss the projection patterns of the other four types, which we designate as H16-2, H16-3, H16-4, and H16-5. These designs are, respectively designs II, III, IV, and V in Hall (1961, pp. 23–24).

For $k = 3$, there are three different possible projections. Two of these arise from *all* the five Hadamard matrices. In addition, H16-5 produces projected designs of type 1 : 3.

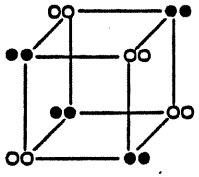
For $k = 4$, five projections occur, as shown in Figure 3. Two of these, (a) and (c), arise from *all* of the Hadamard matrices. In all parts of Figure 3, the cube represents the space of three of the four factors and the fourth is represented by open dots for the lower level, and solid dots for the upper level. Note the "unbalanced" structure of designs (d) and (e). Whereas the other three designs (a), (b), and (c) are replicated 2_{III}^{4-1}, replicated 2_{IV}^{4-1} and a full 2^4 respectively, designs (d) and (e) are not of this form.

For $k = 5$, there are eight different projections; see Figure 4. Two of these, (a) and (b) arise from *all* of the Hadamard matrices. In all parts of Figure 4, the cube represents the space of three of the five factors. Each circle represents a run of the projected design and in each circle, the left portion is for the fourth factor and the right portion is for the fifth factor. An open half-circle represents the lower level of a factor and a solid (black) half-circle represents the upper level of a factor. Note the "unbalanced" structure of projections (e), (f), (g), and (h). Designs (a), (b), (c), and (d) are, respectively, a replicated 2_{III}^{5-2}, a 2_{III}^{5-1}, a 2_{IV}^{5-1}, and a 2_V^{5-1}. There are three non-equivalent Hadamard matrices for $n = 20$ (Hall, 1965), only one of which is equivalent to a Plackett and Burman design. The other two give exactly similar projections for $k = 3$ and 4. For $k = 5$, however, there is one projection additional to the nine listed previously. For additional details, see Lin and Draper (1991, 1994) and Wang and Wu (1995).

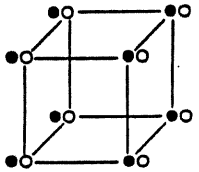
(a) 2^{4-1}_{III} (I = ± 124). Twice over



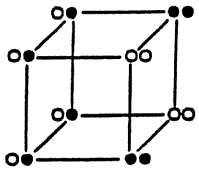
(b) 2^{4-1}_{IV} (I = ± 1234). Twice over



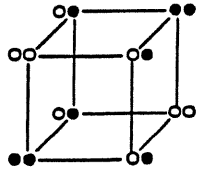
(c) 2^4

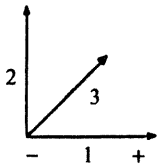


(d)



(e)



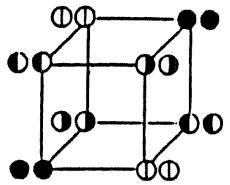


o x_4 at low level

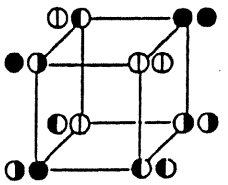
● x_4 at high level

Fig. 3. Projections of 16-Run Hadamard matrix type designs into four dimensions.

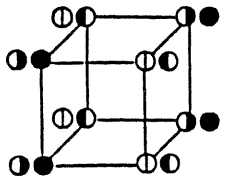
(a) 2^{5-2}_{III} (I = ± 124 = ± 135)



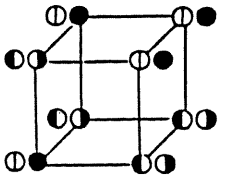
(e)



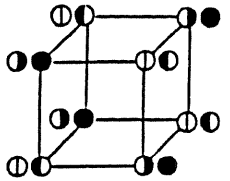
(b) 2^{5-1}_{III} (I = ± 125)



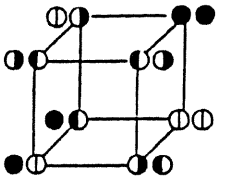
(f)



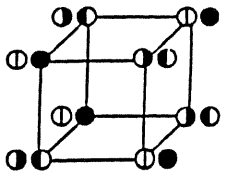
(c) 2^{5-1}_{IV} (I = ± 1235)



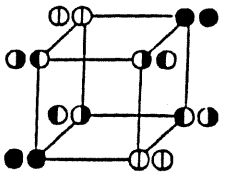
(g)



(d) 2^{5-1}_{V} (I = ± 12345)



(h)



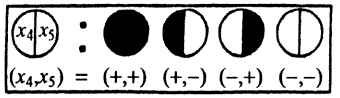


Fig. 4. Projections of 16-run Hadamard matrix type designs into five dimensions.

7. The central composite design

The *central composite design* was one of the early design suggestions (see Box and Hunter, 1957) for obtaining data for fitting a second order surface. It turned out that this design (often called just the *composite design*) satisfied many of the desirable criteria previous listed, and so it has become a cornerstone of response surface methodology. It is constructed from three sets of points which can be described in the coded x-space as follows.

(a) the 2^k vertices $(\pm 1, \pm 1, \ldots, \pm 1)$ of a k-dimensional "cube" ($k \leqslant 4$), or a fraction of it ($k \geqslant 5$);

(b) the $2k$ vertices $(\pm\alpha, 0, \ldots, 0), (0, \pm\alpha, \ldots, 0), \ldots, (0, 0, \ldots, \pm\alpha)$ of a k-dimensional cross-polytope or "star";

(c) a number, n_0, of "center points", $(0, 0, \ldots, 0)$.

Set (a) is simply a full 2^k factorial design or a 2^{k-p} fractional factorial if $k \geqslant 5$. The notation $(\pm 1, \pm 1, \ldots, \pm 1)$ means that 2^k points obtained by taking all possible combinations of signs are used for full factorial cases. (In response surface applications, these points are often referred to as a "cube", whatever the number of factors.)

Set (b) consists of pairs of points on the coordinate axes all at a distance α from the origin. (The quantity α has yet to be specified; according to its value the points may lie inside or outside the cube.) In three dimensions the points are the vertices of an octahedron and this word is sometimes used for other values of $k \neq 3$. However, a more convenient name for such a set of points in k dimensions is "star" or, more formally, cross-polytope.

These sets and the complete design (the n_0 center points represented by a single center point) are shown diagrammatically in Figures 5 and 6 for the cases $k = 2$ and 3.

Fractionation of the cube is possible whenever the resulting design will permit individual estimation of all the coefficients in equation (3). This is guaranteed for fractions of resolution $\leqslant 5$. The smallest usable fraction is then a 2^{k-1} design (a half-fraction) for $k = 5, 6, 7$, a 2^{k-2} design (a quarter-fraction) for $k = 8, 9$, a 2^{k-3} for $k = 10$, and so on. (See Box et al., 1978, p. 408.) Table 3, adapted from Box and Hunter (1957, p. 227) shows the number of parameters in equation (3) and the number of noncentral design points in the corresponding composite design for $k = 2, \ldots, 9$. The values to be substituted for p are $p = 0$ for $k = 2, 3$, and 4; $p = 1$ for $k = 5, 6$, and 7; and $p = 2$ for $k = 8$ and 9; they correspond to the fraction, $1/2^p$, of the cube used for the design.

Points 1, 4, 5, 6, 7, 9, 10, 11 and 12 in Section 3 can all be satisfied by the composite design. Satisfaction of some requires suitable choices of α, n_0, and shrinking or expanding all the design points relative to the region R (see Box and Draper, 1959, 1963; see also, Welch, 1984). Overall the composite design is an excellent choice for many investigations.

Table 3
Features of certain composite designs

No. of variables	k	2	3	4	5	6	7	8	9
No. of parameters	$(k+1)(k+2)/2$	6	10	15	21	28	36	45	55
Cube + star	$2^k + 2k$	8	14	24	–	–	–	–	–
$\frac{1}{2}$(cube) + star	$2^{k-1} + 2k$	–	–	–	26	44	78	–	–
$\frac{1}{4}$(cube) + star	$2^{k-2} + 2k$	–	–	–	–	–	–	80	130
α (rotatable)	$2^{(k-p)/4}$	1.414	1.682	2	2	2.378	2.828	2.828	3.364
Suggested n_0		2–4	2–4	2–4	0–4	0–4	2–4	2–4	2–4

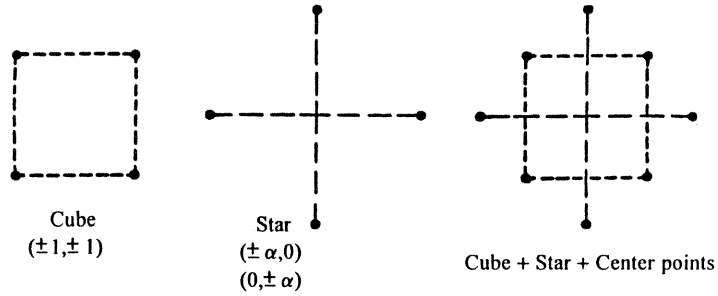


Fig. 5. Composite design for $k = 2$ variables.

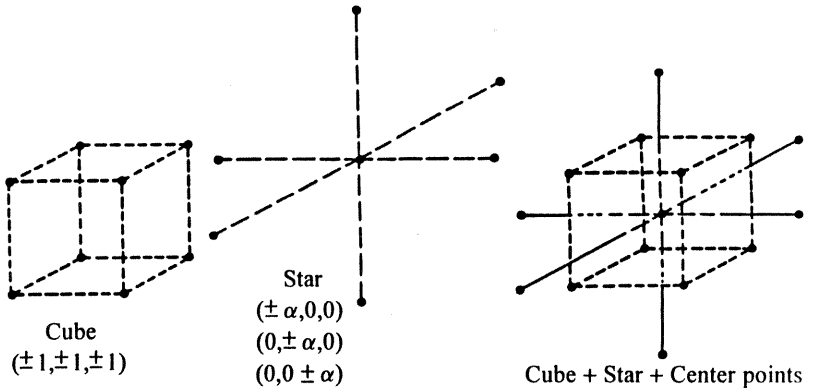


Fig. 6. Composite design for $k = 3$ variables.

Choices of α and n_0

What values should be chosen for α and n_0? The value of α determines if the star points fall inside the cube ($\alpha < 1$), outside the cube ($\alpha > 1$), or on the faces of the cube ($\alpha = 1$). Note that when $\alpha = 1$ only three experimental levels $(-1, 0, 1)$ are

required, which may be an advantage or necessity in some experimental situations. For additional comments and specific designs see De Baun (1956) and Box and Behnken (1960), also discussed in Section 12.

If three levels are not essential, what value of α should be selected? One criterion that can be applied to decide this is that of *rotatability*. A design (of any order) is rotatable when the contours of the variance function $V\{\hat{y}(x)\}$ are spheres about the origin in the k-dimensional factor space defined by variables $x_1, x_2, \ldots, x_k$. Box and Hunter (1957) showed that the required values (given in Table 3) are $\alpha = 2^{(k-p)/4}$, where $p = 0, 1$, or 2 according to the fraction of the cube used in the design.

Note that the rotatability property is specifically related to the codings chosen for the x's. It is usually assumed that these codings have been chosen in a manner that anticipates (roughly speaking) that one unit of change in any x will have about the same effect on the response variable. In such a case, obtaining equal information at the same radial distance in any direction (which is what rotatability implies) is clearly sensible. Codings are rarely perfect; the codings are adjusted in future designs as a result of information gained in current and past experiments. Exact rotatability is not a primary consideration. However, knowledge of the tabulated values provides a target to aim at, while attempting to satisfy other desirable design features.

How large a value should be selected for n_0? There are many possible criteria to apply; these are summarized by Draper (1982, 1984). The suggested values in the table appear to be sensible with respect to many criteria, the overall message being that only a few center points are usually needed. (Whenever α is chosen so that all the design points lie on a sphere, at least one center point is needed so that all of the coefficients can be individually estimated.) A few additional center points will do no harm. Nevertheless, additional runs are probably better used to duplicate selected noncentral design points, unless special considerations apply, as below. Repeated points spread over the design provide a check of the usual "homogeneous variance" assumption; see Box (1959) and Dykstra (1959, 1960).

For a wide variety of examples, see Box and Draper (1987).

8. Small composite designs

When experimentation is expensive, difficult or time-consuming, small designs might be appropriate, especially when an independent estimate of experimental error is available. Hartley (1959) pointed out that, for estimation of the quadratic surface, the cube portion of the composite design need not be of resolution V. It could be of resolution as low as III, provided that two-factor interactions were not aliased with other two-factor interactions. Hartley employed a smaller fraction of the 2^k factorial than is used in the original Box–Wilson designs and so reduced the total number of design points. Hartley's cubes may be designated *resolution* III*, meaning a design of resolution III but with no words of length four in the defining relation; see Draper and Lin (1990b). Hartley thus obtained minimal- or near-minimal-point second-order designs for $k = 2, 3, 4$ and 6. For $k = 5, 7, 9$, and higher numbers, there was then the possibility that a worthwhile improvement could be made.

Table 4
Numbers of cube points in some small composite designs

	Factors, k							
	3	4	5	6	7	8	9	10
Coefficients $p = (k+1)(k+2)/2$	10	15	21	28	36	45	55	66
Star points $2k$	6	8	10	12	14	16	18	20
Minimal points in cube	4	7	11	16	22	29	37	46
Box and Hunter (1957)	8	16	16	32	64	64	128	128
	(2^3)	(2^4)	(2_V^{5-1})	(2_V^{6-1})	(2_V^{7-1})	(2_V^{8-2})	(2_V^{9-2})	(2_V^{10-3})
Hartley (1959)	4	8	–	16	32	–	64	–
	(2_{III}^{3-1})	(2_{III}^{4-1})	–	(2_{III*}^{8-2})	(2_{III*}^{7-2})	–	(2_{III*}^{9-3})	–
Westlake (1965)	–	–	12	–	26	–	44	–
			$(3/8 \times 2^5)$	–	$(13/64 \times 2^7)$	–	$(11/128 \times 2^9)$	–
Draper (1985)	–	–	12	–	28	–	44	–
Minimal runs via Plackett and Burman	4	8	12	16	24	36	40	48
After elimination of repeat	4	8	11	16	22	30	38	46

Westlake (1965) provided a method for generating composite designs based on irregular fractions of the 2^k factorial system rather than using the complete factorials or regular fractions of factorials employed by Box and Wilson (1951) and Hartley (1959). Westlake gave designs for the following:

1. $k = 5$, based on a 3/8 fraction of the 2^5 factorial;
2. $k = 7$, based on a 13/64 fraction of the 2^7 factorial;
3. $k = 9$, based on a 11/128 fraction of the 2^9 factorial.

An alternative approach to obtaining small composite designs was used by Draper (1985), who employed columns of the Plackett and Burman designs rather than regular or irregular fractions. An advantage of this Plackett and Burman type of approach is that the designs are easy to construct. Specifically, (a) we can use, for the cube portion of the design, k columns of a Plackett and Burman (1946) design, and (b) where repeat runs exist, we can remove one of each set of duplicates if we wish to reduce the number of runs required.

Applying this method, Draper (1985) used 12-run, 28-run, and 44-run Plackett and Burman designs and obtained second-order response-surface designs with 22, 42, and 62 total runs (i.e., cube plus star points) for $k = 5, 7$, and 9, respectively. Deleting one of each duplicate pair gave 21 runs for $k = 5$ (a minimal-point design, beating Westlake's design by one run), 39 runs for $k = 7$ (again, one run fewer than Westlake's), and 60 runs for $k = 9$ (two runs fewer than Westlake's designs). A subsequent paper, Draper and Lin (1990a), provided new designs for $k = 7, 8, 9$, and 10, with improvements for cases $k = 7$ and $k = 9$.

Table 4 summarizes the major results related to fitting a second-order model, listing the cube points needed for the composite designs discussed previously. In all cases, center point and star points have been omitted from the table.

Comments on Table 4

Case $k = 3$. As discussed previously, the four-run Plackett and Burman design is a minimal-point design. It is equivalent to Hartley's design and is a 2^{3-1}_{III} design.

Case $k = 4$. The minimum possible number of cube points required is 7, so the eight-run Plackett and Burman design is considered. Columns (1, 2, 3, 6) give the highest D value, where $D = |X'X|^{1/p}/n$, known as the "information per point", is a popular measure of goodness for experimental designs. This 2^{4-1}_{III} design is equivalent to Hartley's design. There is one run more than the minimum number required in the cube.

Case $k = 5$. Five columns of the 12-run Plackett and Burman design are used, because 11 is the minimum possible number of cube points required. As Draper (1985, p. 174) showed, there are two basic types of designs, one with a repeat pair and one with a mirror-image pair. All other choices are equivalent to one of these. The columns (1, 2, 3, 4, 5) produce a mirror-image pair and the higher D value. The columns (1, 2, 3, 9, 11) produce a repeat pair, leading to a minimal-point design with 11 runs in the cube portion after removal of a duplicate run.

Case $k = 6$. Again, a minimal-point design is automatically obtained when six appropriate columns are chosen from a 16-run Plackett and Burman design. Based on the D criterion, the choice of columns (1, 2, 3, 4, 5, 14) is recommended. This is equivalent to Hartley's 2^{6-2}_{III} design.

Case $k = 7$. There are 36 coefficients to estimate and 14 star points. Thus a minimum of 22 cube points is required. The smallest Plackett and Burman design that can be used is thus the one with 24 runs. We wish to pick seven columns. There are 12 possible projection patterns, 5 of which produce nonsingular second-order $X'X$ matrices (see Draper and Lin, 1990a). The choice of columns (1, 2, 3, 5, 6, 7, 9) will give the highest D value. The choice of columns (1, 2, 5, 6, 7, 9, 10), however, will produce two repeat pairs, permitting the elimination of two runs, one from each pair. This minimal-point 22-run design is not only smaller than Hartley's 32-run design, but it is also smaller than Westlake's 26-run design.

Case $k = 8$. There are 45 coefficients to estimate, and 16 star points, so a minimum of 29 cube points is required. The 32-run Plackett and Burman design thus suggests itself. The choice of eight columns from this design constitutes a 2^{8-3} design. There is no 2^{8-3} design of resolution III*, however. (The table by Westlake (1965, p. 325) incorrectly suggested that there is.) Thus we use the 36-run Plackett and Burman design instead. Columns (1, 3, 4, 5, 6, 7, 8, 9) will give the highest D value. Columns (1, 3, 4, 6, 8, 10, 16, 17) will produce six repeat pairs, of which one run each can be eliminated to obtain only 30 runs in the cube portion, one run more than the minimum number required.

Case $k = 9$. There are 55 coefficients to estimate and 18 star points. Thus a minimum of 37 cube points is required. This suggests use of nine columns of the 40-run Plackett and Burman design. There are at least 50 different projection patterns (see Draper and Lin, 1990a). The highest D value found is obtained by choosing columns

Table 5
Columns that provide the highest relative D values found

k	p	n_{pb}	Columns chosen	Total points N
3	10	4	(1, 2, 3)	10
4	15	8	(1, 2, 3, 6)	16
5	21	12	(1, 2, 3, 4, 5)	22
6	28	16	(1, 2, 3, 4, 5, 14)	28
7	36	24	(1, 2, 3, 5, 6, 7, 9)	38
8	45	36	(1, 3, 4, 5, 6, 7, 8, 9)	52
9	55	40	(1, 2, 5, 6, 8, 21, 22, 23, 26)	58
10	66	48	(1, 2, 3, 4, 5, 6, 7, 11, 12, 25)	68

(1, 2, 5, 6, 8, 21, 22, 23, 26). Columns (1, 2, 3, 33, 34, 35, 36, 37, 38) provide two repeat pairs, however, in each of which one run could be eliminated to give a two-level design of 38 points. This compares with 128 runs for Box and Hunter (1957, p. 233), 64 runs for Hartley (1959), 44 runs for Westlake (1965, p. 331) and 44 runs for Draper (1985, p. 179).

Case $k = 10$. For $k = 10$ factors, the smallest 2^{k-p}_{V} design requires 128 runs. There are 66 coefficients to estimate and 20 star points, so a minimum of 46 cube points is required. The obvious choice is to try 10 columns of the 48-run Plackett and Burman design. At least 32 types of projected designs exist, and the highest D value among them is obtained by choosing the columns (1, 2, 3, 4, 5, 6, 8, 9, 17, 18). Choice of the columns (1, 4, 5, 7, 10, 11, 14, 16, 17, 20), however, produces two repeat pairs, permitting elimination of one run from each pair to obtain a minimal-point design.

Table 5 summarizes, for $3 \leqslant k \leqslant 10$, those column choices already described that provide the highest relative D values.

Repeat runs provide information on pure error. Some repeat runs can be eliminated, however, if reduction in the total number of runs is critical. For more details, see, Draper and Lin (1990a). Note that when repeats runs are eliminated, the orthogonality is lost, causing correlations among the estimates.

9. Orthogonal blocking of second order designs

When experiments are spread out over space or time or material or equipment, it is possible for extraneous changes to occur which affect the response values, over and above the effects induced by changes in the predictor variables. For example, test bread ovens may hold only a few loaves, so that several baking sessions are needed. Or, a run might take nearly two hours, so that only four runs are possible on an eight hour shift. Or, the raw material for the experiment may come from two manufacturers. Or the response values at given conditions might drift over time periods, or be affected by changing weather conditions. In all these cases, it is usually desirable to divide the whole experiment up into blocks of runs in such a way that the responses within a

block are "consistent", apart, of course, from the effects of the predictor variables. In some circumstances, second order designs can be *orthogonally blocked*, that is, divided into two or more sections or blocks in such a manner that this split does not affect the estimates (of the second order model parameters) that are obtained from an ordinary least squares regression analysis. The necessary and sufficient conditions were given by Box and Hunter (1957); see also De Baun (1956) and Box (1959). They are:

1. Each block must, by itself, be a first-order orthogonal design. Thus for $i \neq j = 1, 2, \ldots, k$, $\sum_u x_{iu}x_{ju} = 0$, for each block.
2. The fraction of total sum of squares of each variable x_i contributed by every block must be equal to the fraction of the total observations allotted to the block. Thus, for each block,

$$\frac{\sum_u x_{iu}^2}{\sum_{u=1}^n x_{iu}^2} = \frac{n_b}{n}, \tag{4}$$

where n_b denotes the number of runs in the block under consideration, $\sum_u$ denotes summation *only in that block*, and the denominators of (4) refer to the entire design.

The simplest orthogonal block division of the composite design is into the orthogonal design pieces:

Block 1. Cube portion (2^{k-p} points) plus c_0 center points.
Block 2. Star portion ($2k$ points) plus s_0 center points.

Application of (4) then implies that

$$\alpha = \{2^{k-p-1}(2k + s_0)/(2^{k-p} + c_0)\}^{1/2}. \tag{5}$$

For example, if $k = 4$ and $p = 0$, so that the first block is a 2^4 factorial plus c_0 center points and the second block is an eight-point octahedron plus s_0 center points, then

$$\alpha = \{8(8 + s_0)/(16 + c_0)\}^{1/2} \tag{6}$$

achieves orthogonal blocking. Rotatability is achieved when $\alpha = 2^{(k-p)/4} = 2$. Substituting this in (6) shows that this design is both rotatable *and* orthogonally blocked whenever $c_0 = 2s_o$. The satisfaction of *both* criteria is not possible in general. Consider the case $k = 3$ and $p = 0$, so that the first block is a 2^3 factorial plus c_0 center points and the second block is a six-point octahedron plus s_0 center points. Then for orthogonal blocking, we need

$$\alpha = \{4(6 + s_0)/(8 + c_0)\}^{1/2}. \tag{7}$$

If $c_0 = 4$ center points are added to the cube and no center points are added to the star ($s_0 = 0$), then $\alpha = 2^{1/2} = 1.414$. This design is orthogonally blocked but is *not* rotatable. However, values of α closer to the rotatable value 1.682 are possible. For

example, if $c_0 = 0$, $s_0 = 0$, and $\alpha = (24/8)^{1/2} = 1.732$ or if $c_0 = 4$, $s_0 = 2$, and $\alpha = (32/12)^{1/2} = 1.633$. The choices are, of course, limited by the fact that c_0 and s_0 must be integers. Generally, orthogonal blocking (α from (5)) takes precedence over rotatability, for which $\alpha = 2^{(k-p)/4}$ is needed. The general condition for both to be achieved simultaneously is

$$2^{k-p} + c_0 = 2^{0.5(k-p)-1}(2k + s_0) \tag{8}$$

for integer (k, p, c_0, s_0). Some possibilities are $(2, 0, s_0, s_0), s_0 \geqslant 1; (4, 0, 2s_0, s_0)$, already discussed below (6); $(5, 1, (4 + 2s_0), s_0); (7, 1, 4(s_0 - 2), s_0), s_0 \geqslant 2$; and $(8, 2, 4s_0, s_0)$, where $s_0 = 0, 1, 2, \ldots$, unless otherwise specified. (Note that some of these arrangements call for more center points than recommended in the table, an example of how applications of different criteria can produce conflicting conclusions.)

Further division of the star will not lead to an orthogonally blocked design. However, it is possible to divide the cube portion into smaller blocks and still maintain orthogonal blocking if $k > 2$. As long as the pieces that result are fractional factorials of resolution III or more (see Box et al., 1978, p. 385), each piece will be an orthogonal design. All fractional factorial pieces *must* contain the same number of center points or else (4) cannot be satisfied. Thus c_0 must be divisible by the number of blocks.

Replication of point sets

In a composite design, replication of either the cube portion or the star portion, or both can be chosen if desired. An attractive example of such possibilities is given by Box and Draper (1987, p. 362). This is a 24-run second-order design for three factors that is both rotatable *and* orthogonally blocked into four blocks of equal size. It consists of a cube (fractionated via $x_1x_2x_3 = \pm 1$) plus replicated (doubled) star plus four center points, two in each 2^{3-1} block. This particular design also provides an interesting example of estimating σ^2 in the situation where center points in *different* blocks of the design are no longer directly comparable due to possible block effects.

Obtaining the block sum of squares

When a second-order design is orthogonally blocked, one can

1. Estimate the β coefficients of the second-order model in the usual way, ignoring blocking.
2. Calculate pure error from repeated points *within* the same block only, and then combine these contributions in the usual way. Runs in different blocks cannot be considered as repeats.
3. Place an extra term

$$SS\ (\text{blocks}) = \sum_{w=1}^{m} \frac{B_w^2}{n_w} - \frac{G^2}{n},$$

with $(m-1)$ degrees of freedom in the analysis of variance table, where B_w is the total of the n_w observations in the wth block and G is the grand total of all the observations in all the m blocks.

If a design is not orthogonally blocked, the sum of squares for blocks is conditional on terms taken out before it. An "extra" sum of squares calculation is needed; see Draper and Smith (1981).

10. Rotatability

Rotatability is a useful property of an experimental design. Any given design produces an X matrix whose columns are generated by the x-terms in the model to be fitted (e.g., (3)) and whose rows correspond to values from the n given design points. If z' is a vector of the form of a row of X but generated by a selected point at which a predicted response is required after estimation of the model's coefficients, then it can be shown that the variance of that prediction is $\mathrm{V}(\hat{y}(x)) = z'(X'X)^{-1}z\sigma^2$ where σ^2 is the variance of an observed response value, assumed to be constant. For any given design, contours of $\mathrm{V}(\hat{y}(x))$ = constant can be plotted in the k-dimensional x-space. If those contours are *spherical*, the design is said to be *rotatable*. In practice, *exact* rotatability is not important, but it is a plus if the design is at least "close to being rotatable" in the sense that $\mathrm{V}\{\hat{y}(x)\}$ changes little for points that are a constant distance from the origin in the region covered by the design points. For more on rotatability, see Box and Draper (1987).

To assess how close a design is to being rotatable, we can use a criterion of Draper and Pukelsheim (1990). We describe this in the context of second order designs, although the concept is completely general for any order. For easy generalization a special expanded notation is needed. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)'$. We shall denote the terms in the second-order model by a vector with elements

$$1;\ \boldsymbol{x}';\ \boldsymbol{x}' \otimes \boldsymbol{x}',$$

where the symbol $\otimes$ denotes the Kronecker product. Thus there are $(1+k+k^2)$ terms,

$$1;\ x_1, x_2, \ldots, x_k;\ x_1^2, x_1x_2, \ldots, x_1x_k;$$

$$x_2x_1, x_2^2, \ldots, x_2x_k; \ldots; x_kx_1, x_kx_2, \ldots, x_k^2.$$

(An obvious disadvantage of this notation is that all cross-product terms occur twice, so the corresponding $\boldsymbol{X'X}$ matrix is singular. A suitable generalized inverse is obvious, however, and this notation is very easily extended to higher orders. For example, third order is added via $\boldsymbol{x}' \otimes \boldsymbol{x}' \otimes \boldsymbol{x}'$, and so on.)

Consider any second-order rotatable design with second-order moments $\lambda_2 = N^{-1}\sum_u x_{iu}^2$ and $\lambda_4 = N^{-1}\sum_u x_{iu}^2 x_{ju}^2$, for $i, j = 1, 2, \ldots, k$ and $i \neq j$. We can write its moment matrix $\boldsymbol{V}$ of order $(1+k+k^2) \times (1+k+k^2)$, in the form

$$\boldsymbol{V} = \boldsymbol{V}_0 + \lambda_2(3k)^{1/2}\boldsymbol{V}_2 + \lambda_4[3k(k+2)]^{1/2}\boldsymbol{V}_4, \tag{9}$$

where $\boldsymbol{V}_0$ consists of a one in the $(1,1)$ position and zeros elsewhere, where $\boldsymbol{V}_2$ consists of $(3k)^{-1/2}$ in each of the $3k$ positions corresponding to pure second-order moments in $\boldsymbol{V}$ and zeros elsewhere, and $\boldsymbol{V}_4$ consists of $3[3k(k+2)]^{-1/2}$ in the k positions corresponding to pure fourth-order moments, $[3k(k+2)]^{-1/2}$ in the $3k(k-1)$ positions corresponding to mixed even fourth-order moments in $\boldsymbol{V}$, and zeros elsewhere. Note that $\boldsymbol{V}_0, \boldsymbol{V}_2$, and $\boldsymbol{V}_4$ are symmetric and orthogonal so that $\boldsymbol{V}_i\boldsymbol{V}_j = \boldsymbol{0}$, and also the $\boldsymbol{V}_i$ have norms $\| \boldsymbol{V}_i \| = [\mathrm{tr}(\boldsymbol{V}_i\boldsymbol{V}_i)]^{1/2} = 1$.

Suppose we now take an arbitrary design with moment matrix $\boldsymbol{A}$, say. Draper et al. (1991) showed that, by averaging $\boldsymbol{A}$ over all possible rotations in the x space, we obtain

$$\overline{\boldsymbol{A}} = \boldsymbol{V}_0 + \boldsymbol{V}_2\mathrm{tr}(\boldsymbol{A}\boldsymbol{V}_2) + \boldsymbol{V}_4\mathrm{tr}(\boldsymbol{A}\boldsymbol{V}_4). \tag{10}$$

We call $\overline{\boldsymbol{A}}$ the *rotatable component of* $\boldsymbol{A}$. The measure of rotatability is

$$\begin{aligned} Q^* &= \| \overline{\boldsymbol{A}} - \boldsymbol{V}_0 \|^2 / \| \boldsymbol{A} - \boldsymbol{V}_0 \|^2 \\ &= \{\mathrm{tr}(\overline{\boldsymbol{A}} - \boldsymbol{V}_0)^2\}/\{\mathrm{tr}(\boldsymbol{A} - \boldsymbol{V}_0)^2\}. \end{aligned} \tag{11}$$

The rotatability measure Q^* is essentially an R^2 statistic for the regression of the design moments of second and fourth order in $\boldsymbol{A}$ onto the "ideal" design moments represented by $\boldsymbol{V}$. Such a criterion is easy to compute and is invariant under design rotation. It enables us to say how rotatable a design is, and to improve the design's Q^* value by adding new design points. For examples, see Draper and Pukelsheim (1990).

11. Variance, bias and lack of fit

Suppose that $\mathrm{E}(y) = \eta(\boldsymbol{\xi})$ where $\boldsymbol{\xi}$ is a vector of predictor variables and let $\boldsymbol{f}(\boldsymbol{\xi})$ be the vector with polynomial elements used to approximate y. We choose the form of f in the hope that it will provide a good approximation to η over some limited region of interest, R say. Two type of errors then need to be considered:

1. Systematic, or bias, errors $\boldsymbol{\delta}(\boldsymbol{\xi}) = \eta(\boldsymbol{\xi}) - \boldsymbol{f}(\boldsymbol{\xi})$, the difference between the expected value of the response, $\mathrm{E}(y) = \eta(\boldsymbol{\xi})$ and the approximating function $\boldsymbol{f}(\boldsymbol{\xi})$.
2. Random errors $\boldsymbol{\varepsilon}$.

Although the above implies that systematic errors $\boldsymbol{\delta}(\boldsymbol{\xi})$ are always to be expected, they are often wrongly ignored. Yet it is only rarely true that bias can be totally ignored. Suppose that $\hat{y}(\boldsymbol{\xi})$ is the fitted value obtained at a general point $\boldsymbol{\xi}$ in the experimental space, when the function $f(\boldsymbol{\xi})$ is fitted to available data on y and $\boldsymbol{\xi}$, then the associated mean square error, standardized for N, the number of observations and σ^2, the error variance is

$$\begin{aligned} &(N/\sigma^2)\mathrm{E}\{\hat{y}(\boldsymbol{\xi}) - \eta(\boldsymbol{\xi})\}'\{\hat{y}(\boldsymbol{\xi}) - \eta(\boldsymbol{\xi})\} \\ &\quad = (N/\sigma^2)\mathrm{E}\{\hat{y}(\boldsymbol{\xi}) - \mathrm{E}\hat{y}(\boldsymbol{\xi}) + \mathrm{E}\hat{y}(\boldsymbol{\xi}) - \eta(\boldsymbol{\xi})\}^2 \\ &\quad = (N/\sigma^2)\mathrm{V}\{\hat{y}(\boldsymbol{\xi})\} + (N/\sigma^2)\{\mathrm{E}\hat{y}(\boldsymbol{\xi}) - \eta(\boldsymbol{\xi})\}^2 \end{aligned}$$

after some reduction. We can write this as

$$\mathrm{M}(\boldsymbol{\xi}) = \mathrm{V}(\boldsymbol{\xi}) + \mathrm{B}(\boldsymbol{\xi})$$

and describe it as "the standardized mean square error at a point $\boldsymbol{\xi}$ is equal to the variance $\mathrm{V}(\boldsymbol{\xi})$ of prediction plus the squared bias $\mathrm{B}(\boldsymbol{\xi})$". We can also make an assessment of variance and bias over any given region of interest R by averaging (and normalizing) $\mathrm{V}(\boldsymbol{\xi})$ and $\mathrm{B}(\boldsymbol{\xi})$ over R. More generally, if $\omega(\boldsymbol{\xi})$ is a weight function, we can write

$$V = \int \omega(\boldsymbol{\xi})\mathrm{V}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \Big/ \int \omega(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \quad \text{and} \quad B = \int \mathrm{B}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \Big/ \int \omega(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}$$

and integrate it over the entire $\boldsymbol{\xi}$-space. Most often in practice we would have

$$\omega(\boldsymbol{\xi}) = \begin{cases} 1 & \text{within } R, \\ 0 & \text{outside } R, \end{cases}$$

whereupon V and B would represent integrals taken over R. If we denote the integrated mean squared error by M, we can write

$$M = V + B.$$

In practice, of course, the true relationship $\eta(\boldsymbol{\xi})$ would be unknown. To make further progress, we can proceed as follows:

1. Given that we are going to fit a polynomial of degree d_1 (say) to represent the function over some interval R, we can suppose that the true function $\eta(\boldsymbol{\xi})$ is a polynomial of degree d_2, greater than d_1.
2. We need also to say something about the *relative* magnitudes of systematic (bias) and random (variance) errors that we could expect to meet in practical cases. An investigator might typically employ a fitted approximating function such as a straight line, if he believed that the average departure from the truth induced by the approximating function were no worse than that induced by the process of fitting. We shall suppose this to be so, and will assume, therefore, that the experimenter will tend to choose the weight function $\omega(\boldsymbol{\xi})$, the size of his region R, and the degree of his approximating function in such a way that the integrated random error and the integrated systematic error are about equal. Thus we shall suppose that the situation of typical interest is that where B is roughly equal to V.

All-bias designs

If the problem of choosing a suitable experiment design is considered in the context described above, a major result can be deduced. An appropriate experimental design for an "average situation" when V and B are roughly equal has size roughly 10% greater than the *all bias design*, appropriate when $V = 0$. This result is important because the moments of the all-bias design are easily determined. Suppose that we now work in terms of variables $\boldsymbol{x}$, where the x's are coded forms of the ξ's, and centered around the origin, a conventional step. Suppose further that a polynomial model of degree d_1

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{x}_1'\boldsymbol{b}_1$$

is fitted to the data, while the true model is a polynomial of degree d_2,

$$\eta(\boldsymbol{x}) = \boldsymbol{x}_1'\boldsymbol{\beta}_1 + \boldsymbol{x}_2'\boldsymbol{\beta}_2.$$

Thus, for the complete set of N data points

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{X}_1\boldsymbol{b}_1,$$

$$\eta(\boldsymbol{x}) = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2.$$

Quite often, it would reasonable to choose $d_2 = d_1 + 1$. Let us now write

$$\boldsymbol{M}_{11} = N^{-1}\boldsymbol{X}_1'\boldsymbol{X}_1, \qquad \boldsymbol{M}_{12} = N^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2,$$

$$\boldsymbol{\mu}_{11} = \int_O w(\boldsymbol{x})\boldsymbol{x}_1\boldsymbol{x}_1'\,\mathrm{d}\boldsymbol{x}, \qquad \boldsymbol{\mu}_{12} = \int_O w(\boldsymbol{x})\boldsymbol{x}_1\boldsymbol{x}_2'\,\mathrm{d}\boldsymbol{x}.$$

It can now be shown that, whatever the values of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, a necessary and sufficient condition for the squared bias B to be minimized is that

$$\boldsymbol{M}_{11}^{-1}\boldsymbol{M}_{12} = \boldsymbol{\mu}_{11}^{-1}\boldsymbol{\mu}_{12}.$$

A sufficient (but not necessary) condition for B to be minimised is that

$$\boldsymbol{M}_{11} = \boldsymbol{\mu}_{11} \quad \text{and} \quad \boldsymbol{M}_{12} = \boldsymbol{\mu}_{12}.$$

Now the elements of $\boldsymbol{\mu}_{11}$ and $\boldsymbol{\mu}_{12}$ are of the form

$$\int_O w(\boldsymbol{x})x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_k^{\alpha_k}\,\mathrm{d}\boldsymbol{x}$$

and the elements of $\boldsymbol{M}_{11}$ and $\boldsymbol{M}_{12}$ are of the form

$$N^{-1}\sum_{u=1}^{N} x_{1u}^{\alpha_1}x_{2u}^{\alpha_2}\cdots x_{ku}^{\alpha_k}.$$

These typical elements are, respectively, moments of the weight function and moments of the design points of order

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_k.$$

Thus, the sufficient condition above states that, up to and including order $d_1 + d_2$, all the moments of the design are equal to all the moments of the weight function.

EXAMPLE 1. Suppose we wish to fit a straight line $y = \beta_0 + \beta_1 x + \varepsilon$ to data to be taken over the region R, $-1 \leqslant x \leqslant 1$, where the weight function $\omega(x)$ is uniform within R and zero outside R. Suppose quadratic bias is slightly feared. Then the all-bias design is obtained when the design moments m_1, m_2, m_3, where

$$m_i = N^{-1} \sum_{u=1}^{N} x_u^i$$

are chosen to be $m_1 = m_3 = 0$, because $\mu_1 = \mu_3 = 0$, and

$$m_2 = \mu_2 = \int_{-1}^{1} x^2 \, dx \Big/ \int_{-1}^{1} dx = \frac{1}{3}.$$

It follows that, if we use a three-site, three-point design at positions $x = -a, 0, a$, we must choose $2a^2/3 = 1/3$ or $a = 2^{-1/2} = 0.707$. For a typical case where $V = B$ roughly, we could increase a slightly to (say) 0.75 or 0.80, about 10% or so.

EXAMPLE 2. In k dimensions, fitting a plane and fearing a quadratic, with R the unit sphere, an all bias design is a 2^{k-p}_{III} design with points $(\pm a, \pm a, \ldots, \pm a)$ such that

$$2^{k-p} a^2 / n = k/(k+2)$$

which implies, if n_0 center points are used that

$$a = \left\{ \frac{k(2^{k-p} + n_0)}{(k+2)2^{k-p}} \right\}^{1/2}.$$

Note that the special case $k = 1, p = 0, n_0 = 1$ is Example 1. For $k = 4, p = 0, n_0 = 2$, we have $a = 0.866$ for the all bias case. Note that this places the factorial points at distances $r = (4a^2)^{1/2} = 3^{1/2}$ from the origin, that is, outside R.

Detecting lack of fit

Consider the mechanics of making a test of goodness of fit using the analysis of variance. Suppose we are estimating p parameters, observations are made at $p+f$ distinct points, and repeated observations are made at certain of these points to provide e pure error degrees of freedom, so that the total number of observations is $N = p + f + e$.

The expectation of the unbiased pure error mean square is σ^2, the experimental error variance, and the expected value of the lack of fit mean square equals $\sigma^2 + \Lambda^2/f$ where Λ^2 is a noncentrality parameter. The test for goodness of fit is now made by comparing the mean square for lack of fit against the mean square for pure error, via an $F(f, e)$ test.

In general, the noncentrality parameter takes the form

$$\Lambda^2 = \sum_{u=1}^{N} \{\mathrm{E}(\hat{y}_u) - \eta_u\}^2 = \mathrm{E}(S_L) - f\sigma^2, \tag{12}$$

where S_L is the lack of fit sum of squares. Thus, good detectability of general lack of fit can be obtained by choosing a design that makes Λ^2 large. It turns out that this requirement of good detection of model inadequacy can, like the earlier requirement of good estimation, be achieved by certain conditions on the design moments. Thus, under certain sensible assumptions, it can be shown that a $(d_1 =)$ dth order design would provide high detectability for terms of order $(d_2 =)(d+1)$ if (1) all odd design moments of order $(2d+1)$ or less are zero, and (2) the ratio

$$\frac{N^d \sum_{u=1}^{N} r_u^{2(d+1)}}{\left\{\sum_{u=1}^{N} r_u^2\right\}^{d+1}} \tag{13}$$

is large, where

$$r_u^2 = x_{1u}^2 + x_{2u}^2 + \cdots + x_{ku}^2.$$

In particular, this would require that, for a first-order design $(d = 1)$ the ratio $N \sum r_u^4 / \{\sum r_u^2\}^2$ should be large to provide high detectability of quadratic lack of fit; for a second-order design $(d = 2)$, the ratio $N^2 \sum r_{iu}^6 / \{\sum r_{iu}^2\}^3$ should be large to provide high detectability of cubic lack of fit.

Note that, for the 2^{k-p} design of Example 2 above, the detectability criterion is independent of the size of a, which cancels out. Increasing the number of center points slightly increases detectability, however, since this is determined by contrasting the factorial point average response minus the center point average response.

Further reading

For additional commentary, see Chapter 13 of Box and Draper (1987) and Chapter 6 of Khuri and Cornell (1987). Related work includes Draper and Sanders (1988), DuMouchel and Jones (1994), and Wiens (1993).

12. Some other second order designs

The central composite design is an excellent design for fitting a second-order response surface, but there are also other useful designs available. We now mention some of these briefly.

The 3^k factorial designs

The 3^k factorial design series consists of all possible combinations of three levels of k input variables. These levels are usually coded to $-1, 0$, and 1. For the $k = 2$ case, the design matrix is

$$D = \begin{array}{c} \begin{matrix} x_1 & x_2 \end{matrix} \\ \begin{bmatrix} -1 & -1 \\ 0 & -1 \\ 1 & -1 \\ -1 & 0 \\ 0 & 0 \\ 1 & 0 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \end{array}$$

Such a design can actually be used to fit a model of form $\mathrm{E}(y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2 + \beta_{122} X_1 X_2^2 + \beta_{112} X_1^2 X_2 + \beta_{1122} X_1^2 X_2^2 + \varepsilon$, although the cubic and quartic terms would usually be associated with "error degrees of freedom".

To reduce the total number of experimental design points when k is large, fractional replications 3^{k-p} would often be employed if the number of runs were enough to fit the full second-order model. An extended table of fractional 3^{k-p} designs is given by Connor and Zelen (1959).

The Box–Behnken designs

The Box–Behnken (1960) designs were constructed for situations in which it was desired to fit a second-order model (3), but only three levels of each predictor variable $x_1, x_2, \ldots, x_k$, coded to $-1, 0$, and 1, could be used. The design points are carefully chosen subsets of the points of 3^k factorial designs, and are generated through balanced incomplete block designs (BIBD) or partially balanced incomplete block designs (PBIBD). They are available for k = 3–7, 9–12, and 16. (See Table 6 for k = 3–7.) They are either rotatable (for $k = 4$ and 7) or close to rotatable. Except for the designs for which $k = 3$ and 11, all can be blocked orthogonally. The designs

Table 6
The Box–Behnken (1960) designs, $3 \leqslant k \leqslant 7$

Number of factors k	Design matrix	No. of points	Blocking and association schemes
3	$\begin{bmatrix} \pm1 & \pm1 & 0 \\ \pm1 & 0 & \pm1 \\ 0 & \pm1 & \pm1 \\ 0 & 0 & 0 \end{bmatrix}$	12 3 $N = 15$	No orthogonal blocking BIB (one associate class)
4	$\begin{bmatrix} \pm1 & \pm1 & 0 & 0 \\ 0 & 0 & \pm1 & \pm1 \\ 0 & 0 & 0 & 0 \\ \hline \pm1 & 0 & 0 & \pm1 \\ 0 & \pm1 & \pm1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline \pm1 & 0 & \pm1 & 0 \\ 0 & \pm1 & 0 & \pm1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	8 1 8 1 8 1 $N = 27$	3 blocks of 9 BIB (one associate class)
5	$\begin{bmatrix} \pm1 & \pm1 & 0 & 0 & 0 \\ 0 & 0 & \pm1 & \pm1 & 0 \\ 0 & \pm1 & 0 & 0 & \pm1 \\ \pm1 & 0 & \pm1 & 0 & 0 \\ 0 & 0 & 0 & \pm1 & \pm1 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \pm1 & \pm1 & 0 & 0 \\ \pm1 & 0 & 0 & \pm1 & 0 \\ 0 & 0 & \pm1 & 0 & \pm1 \\ \pm1 & 0 & 0 & 0 & \pm1 \\ 0 & \pm1 & 0 & \pm1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	20 3 20 3 $N = 46$	2 blocks of 23 BIB (one associate class)
6	$\begin{bmatrix} \pm1 & \pm1 & 0 & \pm1 & 0 & 0 \\ 0 & \pm1 & \pm1 & 0 & \pm1 & 0 \\ 0 & 0 & \pm1 & \pm1 & 0 & \pm1 \\ \pm1 & 0 & 0 & \pm1 & \pm1 & 0 \\ 0 & \pm1 & 0 & 0 & \pm1 & \pm1 \\ \pm1 & 0 & \pm1 & 0 & 0 & \pm1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	48 6 $N = 54$	2 blocks of 27 First associates: (1, 4); (2, 5); (3, 6)
7	$\begin{bmatrix} 0 & 0 & 0 & \pm1 & \pm1 & \pm1 & 0 \\ \pm1 & 0 & 0 & 0 & 0 & \pm1 & \pm1 \\ 0 & \pm1 & 0 & 0 & \pm1 & 0 & \pm1 \\ \pm1 & \pm1 & 0 & \pm1 & 0 & 0 & 0 \\ 0 & 0 & \pm1 & \pm1 & 0 & 0 & \pm1 \\ \pm1 & 0 & \pm1 & 0 & \pm1 & 0 & 0 \\ 0 & \pm1 & \pm1 & 0 & 0 & \pm1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	56 6 $N = 62$	2 blocks of 31 BIB (one associate class)

Table 7
Rechtschaffner's (1967) point sets

Number	Points	Design generator (point set)	Typical point
I	1	$(+1,+1,\ldots,+1)$ or $(-1,-1,\ldots,-1)$	$(+1,+1,\ldots,+1)$
II	k	One $+1$ and all other -1	$(+1,-1,\ldots,-1)$
III	$k(k-1)/2$	Two $+1$ and all other -1	$(+1,+1,-1,\ldots,-1)$
IV	k	One $+1$ and all other 0	$(+1,0,\ldots,0)$

Table 8
Point sets of Box and Draper (1972, 1974)

Number	Points	Design generator (point set)	Typical point
I	1	$(+1,+1,\ldots,+1)$ or $(-1,-1,\ldots,-1)$	$(-1,-1,\ldots,-1)$
II	k	One $+1$ and all other -1	$(+1,-1,\ldots,-1,\ldots,-1)$
III	$k(k-1)/2$	Two λ and all other -1	$(\lambda,\lambda,-1,\ldots,-1)$
IV	k	One μ and all other 1	$(\mu,1,\ldots,1)$

have a relatively modest number of runs compared to the number of parameters in the corresponding second-order models. For additional appreciation of the usefulness of these designs, see Draper, Davis, Pozueta and Grove (1994).

Some minimal-point second-order designs

Lucas (1974) gave minimal-point designs not of composite type that he called "smallest symmetric composite designs," which consist of one center point, $2k$ star points, and $\binom{k}{2}$ "edge points." An edge point is a $k \times 1$ vector having ones in the ith and jth location and zeros elsewhere. Note that the edge point designs do not contain any two-level factorial points.

Rechtschaffner (1967) used four different so-called *design generators* (actually point sets) to construct minimal-point designs for estimating a second-order surface (see Table 7). The signs of design generators I, II, and III can be varied (e.g., we may have one -1 and all other $+1$ in design generator II, say). Rechtschaffner's designs are available for $k = 2,3,4,\ldots$, but, as pointed out out by Notz (1982), they have an asymptotical D efficiency of 0 as $k \to \infty$ with respect to the class of saturated designs.

Box and Draper (1971, 1974) provided other minimal-point designs for $k = 2,3,4$, and 5, made up from the design generators (point sets) shown in Table 8. Values for λ and μ were tabulated in the 1974 article. Kiefer, in unpublished correspondence, established, via an existence result, that this type of design cannot be optimal for $k \geqslant 7$, however. Box and Draper's designs were given for $k \leqslant 5$, though they can be generated for any k.

Mitchell and Bayne (1976) used a computer algorithm called DETMAX that Mitchell (1974) developed earlier to find an n-run design that maximizes $|\boldsymbol{X}'\boldsymbol{X}|$, given n, a specified model, and a set of "candidate" design points. For each value of $k = 2,3,4$, and 5, they ran the algorithm 10 times, each time starting with a different randomly selected initial n-run design. The algorithm then improved the starting design by adding or removing points according to a so-called "excursion" scheme until no further improvement was possible.

Notz (1982) studied designs for which $p = n$. He partitioned $\boldsymbol{X}$ so that

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{12} \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22} \end{bmatrix},$$

where $\boldsymbol{Z}_1$ is $(p-k) \times p$ and $\boldsymbol{Z}_2$ is $(p-k) \times k$. Note that $\boldsymbol{Y}_{11}$ is $(p-k) \times (p-k)$, $\boldsymbol{Y}_{12}$ is $(p-k) \times k$, $\boldsymbol{Y}_{21}$ is $k \times (p-k)$, and $\boldsymbol{Y}_{22}$ is $k \times k$, and we can think of $\boldsymbol{Z}_1$ as representing the cube points and $\boldsymbol{Z}_2$ the star points; $\boldsymbol{Y}_{12}$ over $\boldsymbol{Y}_{22}$ consists of the columns $(x_1^2, x_2^2, \ldots, x_k^2)$. Thus (a) all elements in $\boldsymbol{Y}_{11}$ are either $+1$ or -1, (b) all elements in $\boldsymbol{Y}_{22}$ are either 1 or 0, and, more important, (c) all elements in $\boldsymbol{Y}_{12}$ are $+1$. It follows that $|\boldsymbol{X}| = |\boldsymbol{X}'\boldsymbol{X}|^{1/2} = |\boldsymbol{Y}_{11}| \cdot |\boldsymbol{Y}_{22} - \boldsymbol{J}_{k,k}|$, where $\boldsymbol{J}_{k,k}$ is a $k \times k$ matrix with all of its elements equal to 1. Maximization of $|\boldsymbol{X}'\boldsymbol{X}|$ is now equivalent to maximization of $|\boldsymbol{Y}_{11}|$ and $|\boldsymbol{Y}_{22} - \boldsymbol{J}_{k,k}|$ separately. Notz found new saturated designs for $k \leqslant 5$ and extended his result to the $k = 6$ case.

Most of the minimal-point designs available for $k \geqslant 7$ comprise the extensions of Lucas's (1974) or Rechtschaffner's (1967) or Box and Draper's (1971, 1974) designs. Minimal-point designs can also be obtained by using the methods given in Draper (1985) and Draper and Lin (1990a), employing projections of Plackett and Burman designs for $k = 3,5,6,7$, and 10. See Section 8. Their main virtues are that they are easy to construct and of composite form, providing orthogonal or near orthogonal designs and including other previously known small composite designs as special cases. For other designs and related considerations, see Khuri and Cornell (1987). A comparison of all the designs we have discussed in this section is made in Draper and Lin (1990a).

References

Bates, D. M. and D. G. Watts (1988). *Nonlinear Regression Analysis and Its Applications*. Wiley, New York.

Box, G. E. P. (1952). Multifactor designs of first order. *Biometrika* **39**, 40–57. (See also p. 189, note by Tocher.)

Box, G. E. P. (1959). Answer to query: Replication of non-center points in the rotatable and near-rotatable central composite design. *Biometrics* **15**, 133–135.

Box, G. E. P. and D. W. Behnken (1960). Some new three level designs for the study of quantitative variables. *Technometrics* **2**, 455–475.

Box, G. E. P. and S. Bisgaard (1993). What can you find out from 12 experimental runs. *Quality Eng.* 4

Box, G. E. P. and N. R. Draper (1959). A basis for the selection of a response surface design. *J. Amer. Statist. Assoc.* **54**, 622–654.

Box, G. E. P. and N. R. Draper (1963). The choice of a second order rotatable design. *Biometrika* 50, 335–352.

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

Box, G. E. P. and J. S. Hunter (1957). Multifactor experimental designs for exploring response surfaces. *Ann. Math. Statist.* **28**, 195–241.

Box, G. E. P. and J. S. Hunter (1961). The 2^{k-p} fractional factorial designs, Parts I and II. *Technometrics* **3**, 311–351 and 449–458.

Box, G. E. P., W. G. Hunter and J. S. Hunter (1978). *Statistics for Experimenters*. Wiley, New York.

Box, G. E. P. and K. B. Wilson (1951). On the experimental attainment of optimum conditions. *J. Roy. Statist. Soc. Ser. B* **13**, 1–38, discussion 38–45.

Box, M. J. and N. R. Draper (1971). Factorial designs, the $|X'X|$ criterion and some related matters. *Technometrics* **13**, 731–742. Corrections **14** (1972), 511 and **15** (1973), 430.

Box, M. J. and N. R. Draper (1974). Some minimum point designs for second order response surfaces. *Technometrics* **16**, 613–616.

Connor, W. S. and M. Zelen (1959). Fractional factorial experimental designs for factors at three levels. U.S. Dept. of Commerce, National Bureau of Standards, Applied Math. Series No. 54.

Davies, O. L., ed. (1978). *Design and Analysis of Industrial Experiments*. Longman Group, New York.

De Baun, R. M. (1956). Block effects in the determination of optimum conditions. *Biometrics* **12**, 20–22.

De Baun, R. M. (1959). Response surface designs for three factors at three levels. *Technometrics* **1**, 1–8.

Draper, N. R. (1982). Center points in response surface designs. *Technometrics* **24**, 127–133.

Draper, N. R. (1984). Schläflian rotatability. *J. Roy. Statist. Soc. Ser. B* **46**, 406–411.

Draper, N. R. (1985). Small composite designs. *Technometrics* **27**, 173–180.

Draper, N. R., T. P. Davis, L. Pozueta and D. M. Grove (1994). Isolation of degrees of freedom for Box–Behnken designs. *Technometrics* **36**, 283–291.

Draper, N. R. and H. Smith (1981). *Applied Regression Analysis*. Wiley, New York.

Draper, N. R. and D. K. J. Lin (1990a). Small response surface designs. *Technometrics* **32**, 187–194.

Draper, N. R. and D. K. J. Lin (1990b). Connections between two-level designs of resolution III^* and V. *Technometrics* **32**, 283–288.

Draper, N. R., N. Gaffke and F. Pukelsheim (1991). First and second-order rotatability of experimental designs, moment matrices, and information functions. *Metrika* **38**, 129–161.

Draper, N. R. and F. Pukelsheim (1990). Another look at rotatability. *Technometrics* **32**, 195–202.

Draper, N. R. and E. R. Sanders (1988). Designs for minimum bias estimation. *Technometrics* **30**, 319–325.

DuMouchel, W. and B. Jones (1994). A simple Bayesian modification of D-optimal designs to reduce dependence on an assumed model. *Technometrics* **36**, 37–47.

Dykstra, O. (1959). Partial duplication of factorial experiments. *Technometrics* **1**, 63–75.

Dykstra, O. (1960). Partial duplication of response surface designs. *Technometrics* **2**, 185–195.

Hall, M. J. (1961). Hadamard matrix of Order 16. Jet Propulsion Laboratory Summary **1**, 21–26.

Hall, M. J. (1965). Hadamard matrix of Order 20. Jet Propulsion Laboratory Technical Report **1**, 32–76.

Hartley, H. O. (1959). Small composite designs for quadratic response surfaces. *Biometrics* **15**, 611–624.

Herzberg, A. M. (1982). The robust design of experiments: A review. *Serdica Bulgaricae Math. Publ.* **8**, 223–228.

Khuri, A. I. and J. A. Cornell (1987). *Response Surfaces, Designs and Analyses*. Marcel Dekker/ASQC Quality Press, New York/Milwaukee.

Lin, D. K. J. and N. R. Draper (1991). Projection properties of Plackett and Burman designs. Tech. Report 885, Department of Statistics, University of Wisconsin.

Lin, D. K. J. and N. R. Draper (1992). Projection properties of Plackett and Burman designs. *Technometrics* **34**, 423–428.

Lin, D. K. J. and N. R. Draper (1995). Screening properties of certain two level designs. *Metrika* **42**, 99–118.

Lucas, J. M. (1974). Optimum composite designs. *Techometrics* **16**, 561–567.

Mitchell, T. J. (1974). An algorithm for the construction of D-optimal experimental designs. *Technometrics* **16**, 203–210.

Mitchell, T. J. and C. K. Bayne (1978). D-optimal fractions of three-level fractional designs. *Technometrics* **20**, 369–380, discussion 381–383.

Notz, W. (1982). Minimal point second order designs. *J. Statist. Plann. Inference* **6**, 47–58.

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

Rechtschaffner, R. L. (1967). Saturated fractions of 2^n and 3^n fractional designs. *Technometrics* **9**, 569–575.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. Wiley, New York.

Wang, J. C. and C. F. J. Wu (1995). A hidden projection property of Plackett–Burman and related designs. *Statist. Sinica* **5**, 235–250.

Welch, W. J. (1984). Computer-aided design of experiments for response estimation. *Technometrics* **26**, 217–224.

Westlake, W. J. (1965). Composite designs based on irregular fractions of factorials. *Biometrics* **21**, 324–336.

Wiens, D. P. (1993). Designs for approximately linear regression: Maximizing the minimum coverage probability of confidence ellipsoids. *Canad. J. Statist.* **21**, 59–70.