Measures of Lack of Fit for Response Surface Designs and Predictor Variable Transformations

G. E. P. Box and N. R. Draper
Department of Statistics
University of Wisconsin-Madison
Madison, WI 53706

Some first-order ($2^{k-p}$ two-level factorials and fractional factorials plus center points) and second-order (cube plus star plus center-point composite) response surface designs are discussed from the point of view of their ability to detect certain likely kinds of lack of fit of degree one higher than has been fitted. This leads to consideration of conditions for representational adequacy of first- and second-order models in transformed predictor variables. It is shown how to use the estimated regression coefficients from the higher degree model to check if power transformations of the predictor variables could eliminate the lack of fit, and also actually to estimate the transformations.

KEY WORDS: First-order designs; Lack of fit; Response surface designs; Second-order designs; Transformations on predictors.

1. INTRODUCTION

Whenever we design an experiment and tentatively entertain a model to fit to the resulting data, we face the possibility that a more complex model may be needed. We can try to resolve our doubts by employing a larger design that makes it possible to fit the more complex model, but then similar questions arise concerning the model, and so on. Clearly we cannot guard against all possibilities. A practical compromise is

1. to entertain initially at each stage of the experimental iteration, a model (containing, say, $p$ parameters) that we hope will be adequate;
2. to employ an associated design that with a number of runs only modestly larger than $p$ provides for checks sensitive to particularly feared discrepancies;
3. if such discrepancies occur, to consider first the possibility of their elimination by transformation (or retransformation) of the response $y$ and/or the predictor variables $\xi_1, \xi_2, \ldots, \xi_k$;
4. because there are situations in which a more complex model cannot be avoided, to employ design arrangements that can be conveniently augmented to form larger designs appropriate for fitting and checking the more complex model.

Our objective here is to develop methods that enable this type of procedure to be implemented in the case of first- and second-order polynomial models and certain associated, frequently used designs. We discuss ways of checking for specific types of lack of fit and how to eliminate these, if they do occur, by power transformations applied to the predictor variables. Transformation of the response $y$ (Bartlett 1947; Box and Cox 1964; Kruskal 1968; Draper and Hunter 1969) will not be considered here.

Consider a response surface study in which a polynomial of degree $d$ in $k$ predictor variables $\xi_1, \xi_2, \ldots, \xi_k$, is used to represent the expected response $r = \eta(\xi_1, \xi_2, \ldots, \xi_k)$. For example, for $d = 3$, we have a third-order model

$$\eta = \beta_0 + \sum_{i=1}^{k} \beta_i \xi_i + \sum_{i,j=1}^{k} \beta_{ij} \xi_i \xi_j + \sum_{i,j,k=1}^{k} \beta_{ijk} \xi_i \xi_j \xi_k.$$  (1.1)

Denote the observed response by $y$ and suppose that the errors $\epsilon_i = y_i - \eta_i$ are independently and identically normally distributed with variance $\sigma^2$. If, as is usual, we fit models of this kind with predictors coded...
in “design units”

\[ x_i = (\xi_i - \xi_{i0})/S_i, \]  

(1.2)

where \( \xi_{i0} \) is some convenient central value in the experimental range of \( \xi_i \) and \( S_i \) is a suitably chosen scale factor, then we can write the model of degree \( d = 3 \) in the form

\[ \eta = \beta_0 + \left\{ \sum_{i=1}^{k} \beta_i x_i \right\} + \left\{ \sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} x_i x_j \right\} + \left\{ \sum_{i=1}^{k} \sum_{j>i}^{k} \sum_{l>j} \beta_{ijk} x_i x_j x_l \right\}. \]  

(1.3)

In Section 2 we consider the case in which a first-degree (\( d = 1 \)) model is fitted when, at worst, a second-degree model (\( d = 2 \)) might be needed. The second-degree fit case when third-degree might be more appropriate is treated in Section 3. In both cases we show how power transformation of the \( \xi \)'s might make a lower-order model feasible and usable.

2. FIRST-ORDER MODELS AND DESIGNS

2.1 Some First-Order Designs With Discrepancy Checks

Useful first-order response surface designs are the two-level factorials and fractional factorials of resolution three or more that use, respectively, all or some of the \( 2^k \) runs (+1, ±1, . . . , ±1). As discussed, for example, by Box and Wilson (1951), fractions can be chosen so that checks are associated with the residual degrees of freedom containing feared interactions. Alternatively, or in addition (see De Baun 1956), by adding \( n_0 \) center points to any such design, a test for second-order curvature can be made. The contrast \( c_2 \) between the average response \( \bar{y}_{i0} \) at the center and the average response \( \bar{y}_c \) at the factorial points is

\[ c_2 = \bar{y}_c - \bar{y}_{i0} \]

with expected value

\[ E(c_2) = \sum_{i=1}^{k} \beta_{ii}. \]  

(2.1)

Thus, in the common situation in which the \( \beta_{ii} \) are either of the same sign or are near to zero, \( c_2 \) provides an overall check for curvature of second order.

2.2 Can We Use a First-Order Model in Transformed Predictor Variables?

When a curvilinear response relationship exists that is monotonic in the predictor variables over the current region of interest, it may be possible to use a first-order model in which power transformations \( \xi_1^{p_{11}}, \xi_2^{p_{22}}, \ldots, \xi_k^{p_{kk}} \) are applied to the \( \xi \)'s.

Assume that, at worst, the response may be represented by a second-degree polynomial in transformed variables, namely by

\[ \eta(\xi, \lambda) = \beta_0 + \sum_{i=1}^{k} \beta_i \xi_i^{p_{i1}} + \sum_{i=1}^{k} \sum_{j>i}^{k} \beta_{ij} \xi_i^{p_{i1}} \xi_j^{p_{j1}}. \]  

(2.2)

Then a first-degree polynomial model will be appropriate if the \( \lambda_i \) may be chosen so that \( \beta_{ij} = 0 \) for all \( i \) and \( j \). By requiring that all second derivatives of (2.2) with respect to the \( \xi_i^{p_{i1}} \) vanish, we obtain the conditions

\[ \eta_{ij} = 0, \quad i \neq j \]  

(2.3)

\[ \eta_{ii} + \delta(1 - \lambda_i) \eta_i = 0, \quad i = 1, 2, \ldots, k, \]  

(2.4)

where

\[ \eta_i = \left[ \frac{\partial^2 \eta}{\partial x_i \partial x_j} \right]_{x_0 = 0}, \quad \eta_{ij} = \left[ \frac{\partial^3 \eta}{\partial x_i \partial x_j \partial x_l} \right]_{x_0 = 0}. \]  

(2.5)

(the third piece of (2.5) is needed in Sec. 3, but not here), and where

\[ \delta_i = S_i/\xi_{i0}. \]  

(2.6)

Now suppose that a second-order model of the form of (1.3) with \( d = 2 \) has been fitted to the data from an appropriate design. Then we could approximate the derivatives of (2.5) by

\[ \tilde{\eta}_i = b_i; \quad \tilde{\eta}_{ij} = b_{ij}, \quad i \neq j; \quad \tilde{\eta}_{ii} = 2b_{ii}. \]  

(2.7)

Thus (a) the possibility of a first-order representation in transformed variables \( \xi_i^{p_{i1}} \) is contraindicated (see (2.3)) if one or more interaction estimates \( b_{ij}, i \neq j, \) are significantly different from zero, and (b) supposing such a transformation to be possible, the appropriate transformation parameters are roughly estimated by

\[ \hat{\lambda}_i = 1 + 2b_{ii}/(\delta_i b_{ii}), \quad i = 1, 2, \ldots, k. \]  

(2.8)

More precise estimates can be found by application of standard nonlinear least squares, fitting the model of (2.2) with \( \beta_{ij} = 0 \) directly to the data.

Note that since all higher-order (than two) derivatives of (2.2) should be identically zero, the presence of higher-order interactions would indicate the inadequacy of (2.2).

3. SECOND-ORDER MODELS AND DESIGNS

3.1 Some Second-Order Designs With Discrepancy Checks

A useful class of second-order designs (see Box and Wilson 1951) appropriate for fitting (1.3) with \( d = 2 \) consists of the central composite arrangements in which a “cube,” consisting of a two-level factorial with coded points (+1, ±1, . . . , ±1) or a fraction of reso-
Table 1. A Composite Design for \( k = 2 \) Predictor Variables and Its Associated Estimator Columns; \( n_e = 8, n_{co} = 1, n_s = 4, n_{so} = 5, \alpha = 2 \)

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_1^2 )</th>
<th>( x_2^2 )</th>
<th>( \sum_{i=1}^{2} x_{1i} )</th>
<th>( \sum_{i=1}^{1.5} x_{2i} )</th>
<th>( \text{CCC Blocks} )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>-8</td>
<td>-1</td>
<td>30.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>1</td>
<td>30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>4</td>
<td>-2</td>
<td>-1</td>
<td>31.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>-2</td>
<td>-2</td>
<td>-1</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>-1</td>
<td>30.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>1</td>
<td>30.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>1</td>
<td>29.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>1</td>
<td>29.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>1</td>
<td>29.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>1</td>
<td>30.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
(a) The \( x_{1ii} \) column is the portion of an \( x_i \) column that is orthogonal to all the columns of lower order. A divisor of 1.5 is attached for convenience.
(b) The CC column is a curvature contrast column which compares the average response at the non-central points with that at the center points, within each block.

The fitted least squares second-degree equation is

\[
y = 30.59 - 4.22x_1 - 5.91x_2 \\
\pm 0.25 \pm 0.17 \pm 0.17 \\
- 1.66x_1^2 - 1.44x_2^2 - 3.41x_1x_2, \\
\pm 0.14 \pm 0.14 \pm 0.24 (3.1)
\]

where \( \pm \) limits beneath each estimated coefficient indicate estimated standard errors, using the pure error estimate \( s^2 = 0.457 \) to estimate \( \sigma^2 \). An associated analysis of variance table is shown as Table 2.

Before accepting the utility of the fitted equation we need to be reassured on two questions:

1. Is there evidence from the data of serious lack of fit? If not,

2. Is the change in \( \hat{y} \) over the experimental region explored by the design, large enough compared with the standard error of \( \hat{y} \) to indicate that the response surface is adequately estimated?

The analysis of variance of Table 2 sheds light on both these questions. Its use on the second was studied by Box and Wetz (1973); see also Box, Hunter, and Hunter (1978, p. 524) and Draper and Smith (1981, pp. 129–133).

Clearly, for this example, it is the marked lack of fit of the second-order model that immediately concerns us. In particular, it is natural to be concerned with the possible effects of third-order terms. Associated with the design of Table 1 are four possible third-order columns, namely, those formed by creating entries of the form

\[
(x_1^3, x_1x_2^2, x_1^2x_2, x_1x_2^3) (3.2)
\]

These form two sets of two items, as indicated by the parentheses.

Now suppose these third-order columns are orthogonalized with respect to the lower-order \( X \) vectors. This may be accomplished by regressing them against the first six columns and taking residuals to yield columns \( x_{111} \) (from \( x_1^3 \), \( x_{122} \) (from \( x_1^2x_2 \)), and so on. Then

\[
x_{iii} = -3x_{ijj}, \quad i \neq j (3.3)
\]

and the residual vectors are confounded in two sets of two. Furthermore, the columns \( x_{111} \) and \( x_{222} \) are orthogonal to each other. These vectors, reduced by a convenient factor of 1.5 to show their somewhat remarkable basic form, are given in Table 1.

Consider now the column \( x_{111} \) in relation to Figure 1, which shows the projection of the points of the composite design onto the \( x_1 \) axis. Denoting the average of the responses at \( x_1 = -\alpha, -1, 1, \alpha \) by \( y_{-\alpha}, y_{-1}, y_1, y_{\alpha} \), respectively, we see that a contrast \( c_{31} \) associated with \( x_{111} \) is

\[
c_{31} = \frac{1}{2\alpha} x_{111}^2 y = \frac{1}{2\alpha} \left( \frac{y_{-\alpha} - y_{-1}}{2} - \frac{y_{1} - y_{\alpha}}{2} \right). (3.4)
\]
where, for our example, \( \alpha = 2 \). The expression in the parentheses is an estimate of the difference in slope of the two chords joining points equidistant from the design center. For a quadratic response curve this difference is zero. Thus \( c_{31} \) is a natural measure of overall nonquadracity in the \( x_1 \) direction. A corresponding measure in the \( x_2 \) direction is, of course, given by \( c_{32} = x_{12} \bar{y} / 36 \).

The corresponding sums of squares for these contrasts, given in Table 2, indicate a highly significant lack of fit. Corresponding plots of the residuals against \( x_1 \) and against \( x_2 \) show a characteristic pattern. A line joining residuals for observations at \( x_i = \alpha \) and \( x_i = -\alpha \) slopes up, while the tendency of the remaining residuals is down as \( x_i \) is increased. We discuss these data later. First, however, we provide formulas for general \( k \), and then show how it may be possible to use information provided by contrasts like \( c_{31} \) to estimate simplifying transformations in the predictor variables.

3.2 General Formulas

In general a composite design contains

1. A “cube,” consisting of a \( 2^k \) factorial or a \( 2^k - p \) fractional factorial, made up of points of the type \(( \pm 1, \pm 1, \ldots, \pm 1)\), of resolution \( R \geq 5 \) (Box and Hunter 1961) replicated \( f \) (\( f \geq 1 \)) times. There are thus \( n_j = 2^{k-1} f \) such points (where \( p \) may be zero).
2. A “star,” that is, \( 2^k \) points \(( \pm \alpha, 0, 0, \ldots, 0), (0, \pm \alpha, 0, 0, \ldots, 0), \ldots, (0, 0, 0, \ldots, \pm \alpha)\) on the predictor variable axes, replicated \( r \) times, so that there are \( n_j = 2^{kr} \) points in all. The interesting case is where \( \alpha > 1 \), and we assume that this is so here. (The case \( 0 < \alpha < 1 \) is mathematically identical. When \( \alpha = 1 \), there are only three levels of \( x_i \), which causes singularities that we have investigated but do not discuss here.)
3. Center points \((0, 0, \ldots, 0)\), \( n_0 \) in number, of which \( n_{00} \) are in cube blocks and \( n_{00} \) in star blocks.

Appendix B shows that for any such design, \( k \) sets of columns can be isolated with the \( i \)th set containing the \( k \) columns \( x_{1j}, x_{2j}, \ldots, x_{kj}, j = 1, 2, \ldots, k \). This \( i \)th set is associated with a single vector \( x_{iii} \) that is orthogonal both to the \( (k+1)(k+2)/2 \) columns required for fitting the second-degree equation and to the \( (k-1) \) similarly constructed vectors \( x_{iii}, j \neq i \).

The elements of these vectors are such that for the cube points, \( x_{iii} = \phi x_i \), with \( \phi = 2 x_i^2 (1 - x_i^2)/(n_i + 2 x_i^2) \); for the star points, \( x_{iii} = \gamma x_i \), with \( \gamma = n_i(x_i^2 - 1)/(n_i + 2 x_i^2) \); for the center points \( x_{iii} = 0 = x_i \). Thus the \( k \) estimates of third-order lack of fit, \( c_{31}, c_{32}, \ldots, c_{3k} \) are

\[
c_{3i} = \frac{x_{iii} \bar{y}}{x_{iii}^2} = \frac{1}{\sigma^2 - 1} \left\{ \frac{\bar{y}_{ii} - \bar{y}_{i-1}}{2\alpha} - \frac{\bar{y}_{ii} - \bar{y}_{i-1}}{2} \right\}
\]

with standard deviation

\[
\sigma_{c3} = \frac{1}{\alpha^3 - 1} \left( \frac{1}{n_i} + \frac{1}{2\alpha^2} \right)^{1/2} \sigma.
\]

Also,

\[
E(c_{3i}) = \beta_{iii} + (1 - \alpha^2)^{-1} \sum_{j=1}^{k} \beta_{ijj},
\]

and the contribution to the lack-of-fit sum of squares is

\[
SS(c_{3i}) = (\alpha^2 - 1)^2 c_{3i}^2 \left( \frac{1}{n_i} + \frac{1}{2\alpha^2} \right)
\]

Note that even if \( E(c_{3i}) = 0 \), this does not necessarily mean there are no cubic coefficients. A combination of nonzero \( \beta_{iii} \) and \( \beta_{ijj} \) could occur for which \( \beta_{iii} + (1 - \alpha^2)^{-1} \sum_{j=1}^{k} \beta_{ijj} = 0 \). It is, of course, impossible to guard against every such possibility unless the full cubic model is fitted.

3.3 Can We Use a Second-Order Model in Transformed Predictor Variables?

In some instances lack of fit of a second-order model, revealed by significant curvature contrasts of the kind just described, might be removed by transformations of some or all of the predictors. A model of this kind contains many fewer parameters than a full third-order model and is much easier to analyze and interpret. To determine conditions that must be satisfied to make it possible to remove lack of fit in this way, and to determine the part that the curvature contrasts play in this, we suppose that, at worst, the response function may be represented by a third-order model in the transformed predictor variables. If a second-order representation were satisfactory, all third derivatives of the third-order model with respect to the \( \xi_i^h \) would be zero. Requiring this leads to the conditions

\[
\eta_{ij} = 0, \quad \text{all } i \neq j = 1, 2, \ldots, k;
\]

\[
\eta_{ij} + \delta(1 - \lambda)\eta_{ij} = 0, \quad \text{all } i \neq j = 1, 2, \ldots, k;
\]

\[
\eta_{ii} + 3\delta(1 - \lambda)\eta_{ii} + \delta^2(1 - \lambda)(1 - 2\lambda)\eta_{ii} = 0,
\]

\[
i = 1, 2, \ldots, k.
\]

Equation (3.9) provides the (otherwise obvious) conclusion that the possibility of second-order representation in the transformed variables is contraindicated if one or more interaction estimates \( h_{ij} \) are nonzero.

In practice the estimation of the transformation (when not contraindicated) is best done using nonlinear least squares directly on the model of (2.2); however we obtain some interesting information on how the curvature measures \( c_{3i} \) relate to these transformations by considering how (3.10) and (3.11) could be used to obtain estimates of the \( \lambda_i \).
however, that a second-order model augmented with only cubic terms
\[ \beta_{111}x_1^3 + \beta_{222}x_2^3 + \cdots + \beta_{k \cdot \cdot \cdot}x_k^3. \]  
was fitted. If the response \( \eta \) could be represented by the third-order model of (1.3), the estimates \( b_i \) and \( b_{iii} \) obtained from the composite design would have expectations
\[ E(b_i) = \eta_i - \frac{1}{\hat{\sigma}^2}(1 - \hat{\sigma})^{-1} \sum_{j \neq i} \eta_{ij}. \]  
(3.13)

If now \( b_i \) and \( b_{iii} \) are used as estimates of the quantities shown as their expectations, then, after appropriate substitutions have been made in (3.9) through (3.11), we obtain the following \( k \) equations for the \( \eta_i \). (In these equations \( b_{iii} = c_{3i} \).)
\[ \begin{align*}
  h_{iii} - \eta_{iii} &= \frac{1}{\hat{\sigma}} \eta_{ii} + \frac{1}{\hat{\sigma}^2} \sum_{j \neq i} \eta_{ij} + \frac{1}{\hat{\sigma}^3}(1 - \hat{\sigma})(1 - 2\hat{\sigma}) b_i \\
  \times \sum_{j \neq i} \delta(1 - \hat{\sigma})h_{ij} = 0, \quad i = 1, 2, \ldots, k.
\end{align*} \]
(3.15)

These equations can be solved iteratively. Guessed values for the \( \hat{\sigma} \) are first substituted in the grouping \( (1 - \hat{\sigma})(1 - 2\hat{\sigma}) \) wherever it occurs and the resulting linear equations are solved to provide improved estimates for a second iteration, and so on.

For the example data this procedure converges to the values \( \hat{\sigma}_1 = -.23, \hat{\sigma}_2 = -.93 \). These may be compared with the values \( \hat{\sigma}_1 = .09, \hat{\sigma}_2 = -.82 \) provided by nonlinear least squares (these are maximum likelihood estimates under the standard normal error assumptions) and with \( \hat{\sigma}_1 = 0, \hat{\sigma}_2 = -1 \), the values used to generate the data; see Appendix A.

An analysis of variance for the transformed data is shown in Table 3 where, as anticipated, no lack of fit appears. (See also Sec. 3.4 for details.) No reduction in the degrees of freedom is made for the estimates of \( \hat{\sigma} \) and \( \hat{\sigma}^2 \) (Box and Cox 1964, p. 240) because these degrees of freedom are identical to those attributed to third order in Table 3.

### 3.4 A Curvature Contrast

Consider the overall curvature measure \( c_2 \) of (2.1) used to check the first-order model. When, as in the design of Table 1, center points are available both in the factorial block(s) and in the star block(s), several (two for our example) such measures are available. Consider, specifically, the two-block case for a moment. If the average response at the center of the star is \( \bar{y}_{00} \) and the average over all the star points is \( \bar{y}_s \), then the contrast
\[ c_2 = \left( k/\alpha^2 \right) (\bar{y}_s - \bar{y}_{00}) \]
has expectation
\[ E(c_2) = \sum_{i=1}^{k} \beta_i. \]
Thus the statistic
\[ c_2 - \tilde{c}_2 = \bar{y}_s - \bar{y}_{00} - \left( k/\alpha^2 \right) (\bar{y}_s - \bar{y}_{00}), \]
which is the difference of the two measures of overall curvature, should be zero if the assumptions made about the model being quadratic are true.

From Figure 1 we see that the curvature measure \( c_2 \) associated with the cube (open circles) is contrasted with \( \tilde{c}_2 \) associated with the star (black dots). In general the distance from the center of the design to the cube points is \( k^{1/2} \) and that for the star points is \( \alpha \). When, as in our example, \( k^{1/2} \) and \( \alpha \) are different, a significant value of \( c_2 - \tilde{c}_2 \) could indicate (for example) a symmetric departure from quadratic fall-off on each side of the maximum, such as we see for example in a normal distribution curve.

In general, for two blocks, the standard deviation for \( c_2 - \tilde{c}_2 \) is given by
\[ \sigma_{c_2 - \tilde{c}_2} = \left[ \frac{1}{n_k} + \frac{1}{n_{00}} + \frac{k^2}{2\alpha^4} \left( \frac{1}{n_k} + \frac{1}{n_{00}} \right) \right]^{1/2}, \]
and the associated sum of squares for the analysis of variance table entry of Table 2 is obtained from
\[ SS(c_2 - \tilde{c}_2) = (c_2 - \tilde{c}_2)^2 \left[ \frac{1}{n_k} + \frac{1}{n_{00}} + \frac{k}{2\alpha^4} + \frac{k^2}{\alpha^8 n_{00}} \right]. \]
For our example we find
\[ c_2 - \tilde{c}_2 = 1.3925 \pm .7520 \]
with associated sum of squares 1.567 as shown in Table 2. There is clearly no evidence of this sort of lack of fit. When transformed predictors are used, again no lack of fit of this kind is evident, as we see from Table 3.

### 3.5 Interaction With Blocks?

When composite designs are run in blocks, and if we allow the possibility that effects from the predictor variables could interact with blocks, then the various

Table 3. Analysis of Variance for Second-Order Model in Predictor Variables in \( x_1 \) and \( x_2 \)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1</td>
<td>13,938.934</td>
<td></td>
</tr>
<tr>
<td>Blanks</td>
<td>1</td>
<td>7.347</td>
<td></td>
</tr>
<tr>
<td>First order extra</td>
<td>2</td>
<td>552.713</td>
<td>276.357</td>
</tr>
<tr>
<td>Second order extra</td>
<td>3</td>
<td>555.238</td>
<td>188.413</td>
</tr>
<tr>
<td>Lack of fit</td>
<td>2</td>
<td>1.396</td>
<td>0.698</td>
</tr>
<tr>
<td>Pure error</td>
<td>1</td>
<td>0.457</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>15,069.910</td>
<td></td>
</tr>
</tbody>
</table>
measures of lack of fit would be confounded with block-effect interactions. Although such contingencies must always be borne in mind, it should be remembered that these particular block-effect interactions are no more likely than any others.

4. SUMMARY

In the traditional analysis of second-order response surface designs, a number of degrees of freedom are usually consigned to "lack of fit" and the corresponding sum of squares is used to test for overall lack of fit. Some split-up of lack of fit has been previously discussed (Draper and Herzberg 1971), but a much more ambitious and detailed division is described here. We show that it is possible to check for cubic lack of fit in the $k$ axial directions and if it exists, not only to check whether it can be eliminated by a power transformation in the predictors, but also actually to estimate the powers needed to effect the transformation. A worked two-factor example shows how to carry out the calculations.

We also show how a certain curvature contrast can be used to check overall quadratic fall-off away from the maximum of a response surface.

Simpler but similar considerations, one degree down, apply to the first-order model, and appropriate formulas for these arise as a special case.

5. ACKNOWLEDGMENTS

Both authors were partially sponsored at the Mathematics Research Center, University of Wisconsin by the United States Army under Contract No. DAAG29-80-C-0041. N. R. Draper was also sponsored at the Statistics Department, University of Wisconsin by the United States Army through the Office of Army Research under Contract No. DAAG29-80-C-0113.

APPENDIX A. GENERATION OF EXAMPLE DATA

Figure 2 is taken from the manuscript of a book on response surfaces by G. E. P. Box, N. R. Draper, and J. S. Hunter, in preparation. Figure 2(b) shows a quadratic response function with a simple maximum in variables $(\ln \xi, 100\zeta_2^2)$. This figure is redrawn in the metrics $(\xi, \zeta_2)$ in Figure 2(a). In the $(\xi, \zeta_2)$ representation, the doubled cube plus star plus center-points design of Table 1 is indicated by the positions of the dots. Response values were calculated at these points, and random error added to give the $y$ values in Table 1. For these generated data $\xi_{10} = 2.5$, $\zeta_{20} = 12.5, S_1 = 0.75$, and $S_2 = 3.75$.

APPENDIX B. THE THIRD-ORDER COLUMNS OF THE $X$ MATRIX FOR A COMPOSITE DESIGN

We here prove the results summarized in Section 3.2. The full cubic model in $k$ variables $x_1, x_2, \ldots, x_k$ is given by (1.3). The form of the $X$ matrix in the regression model $y = X\beta + \varepsilon$ when the design consists of $f$ "cubes" plus $r$ "stars" plus $n_0$ center points is as shown in Table 4. We denote columns by placing square brackets around the column head; for example, $[x_i]$ will denote the $x_i$ column, and so on. We write $n_i = 2^{k-p}$ for the number of cube points.

All of the cubic columns are orthogonal to all of the other columns with the following exceptions: $[x_i^2]$ is not orthogonal to $[x_i]$, nor to $[x_i x_j^2]$; $[x_i x_j]$ is not
MEASURES OF LACK OF FIT

![Figure 2. Generation of Example Data. Surface (a) Becomes Quadratic, as in (b), When Variables in $\xi_1$ and $\xi_2^{-1}$ are Employed. Data Were Taken From (a) With Added Errors]

Table 4. The $X$ Matrix for a Cubic Model in $k$ Predictors $x_1, x_2, \ldots, x_k$. The Notation $c_i$ is Used for the First $2^{k-p}$ Elements in the $x_i$ Column

TECHNOMETRICS ©, VOL. 24, NO. 1, FEBRUARY 1982
orthogonal to $[x_i]$, nor to $[x_i^2]$, nor to $[x_i x_j^2]$. The first step is to regress the $[x_i^2]$ and $[x_i x_j^2]$ vectors on the $[x_i]$ and take residuals. Because the columns involved are orthogonal to $[x_i]$, no adjustment for means is needed. We denote the “cube portion” of the $[x_i]$ and $[x_i x_j]$ vectors by $c_i$, as indicated in the table. These two sets of residuals are, where the prime denotes transpose,

$$
[x_{iii}] = [x_i^2] - \{[x_i]'[x_i^2]/[x_i]'[x_i]\}[x_i] \quad \text{(B.1)}
$$

and

$$
[x_{ijj}] = [x_i x_j^2] - \{[x_i]'[x_i x_j^2]/[x_i]'[x_i]\}[x_i] \quad \text{(B.2)}
$$

both of which reduce to multiples $(1 - \alpha^2)m$ and $m$, respectively, where $m = 2p x^2/(n_x + 2p x^2)$, of the same vector. For example,

$$
[x_{iii}]' = [c_i]' - d, 0, 0, \ldots, 0 \{1 - \alpha^2\}m, \quad \text{(B.3)}
$$

where

$$
d = n_x/(2\pi), \quad \text{(B.4)}
$$

where $c_i'$ is of length $2^{k-2}$, and there are $r$ sets of $(d, -d)$'s and $n_0$ zeros in the vector part of (B.3). For other $[x_{iii}]'$, $c_i'$ will be replaced by $c_i'$ and the position of the $-d$'s will correspond to those of the $\pm d$'s in the corresponding $[x_i]$ vector. Note that, because $c_i c_j = 0$, $i \neq j$, it is obvious that $[x_{iii}]$ and $[x_{ijj}]$ are orthogonal.

It follows that the $k$ cubic coefficients $\beta_{iii}, \beta_{ijj}$ ($i \neq l$, $j = 1, 2, \ldots, k$, otherwise) cannot be estimated individually but only in linear combination, and that an appropriate normalized estimating contrast for this is

$$
l_{iii} = [x_{iii}]'y/[x_{iii}]'x_{iii}\]

$$

= \{c_i' y + d(-r y_{+i} + r y_{-i})\}/(n_x(1 - \alpha^2)), \quad \text{(B.5)}
$$

where $y_1$ is the portion of $y$ corresponding to the cubic part of the design, and $\bar{y}_{+i}, \bar{y}_{-i}$ are, respectively, the averages of observations taken at the $z$ and $-z$ axial points on the $x_i$ axis. If we similarly denote by $y_{ii}$ and $y_{-i}$ the averages of the $n_x/2$ observations in $y_1$ corresponding to 1 and $-1$ in $c_i$, respectively, it follows quickly that $l_{iii} = c_{3i}$, where $c_{3i}$ is given in (3.5). The expected value is

$$
E(c_{3i}) = [x_{iii}]'x[\beta]/[x_{iii}]'x_{iii}, \quad \text{(B.6)}
$$

where $X$ is as in Table B.1 and the coefficients of $\beta$ correspond to the columns in the obvious manner. Because $[x_{iii}]$ is orthogonal to all columns of $X$ except the $[x_i^2]$ and $[x_i x_j^2]$ columns, (3.7) emerges almost immediately.

[Received February 1980. Revised October 1981.]

REFERENCES


