

AD-A100 614

WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1  
MEASURES OF LACK OF FIT FOR RESPONSE SURFACE DESIGNS AND PREDIC--ETC(U)  
APR 81 8 E BOX, N R DRAPER DAAG29-80-C-0041  
MRC-TSR-2211 NL

UNCLASSIFIED

1 of 1  
AC  
AD06676



END  
DATE  
FILMED  
7 81  
DTIC

**LEVEL**

2

AD A100614

MRC Technical Summary Report #2211

MEASURES OF LACK OF FIT FOR RESPONSE  
SURFACE DESIGNS AND PREDICTOR VARIABLE  
TRANSFORMATIONS,

G. E. P. Box and N. R. Draper

Mathematics Research Center  
University of Wisconsin—Madison  
610 Walnut Street  
Madison, Wisconsin 53706

April 1981

DTIC  
EXCISE  
JUN 25 1981  
D  
C

Received March 25, 1981

DTIC FILE COPY

Approved for public release  
Distribution unlimited

Sponsored by  
U. S. Army Research Office  
P. O. Box 12211  
Research Triangle Park  
North Carolina 27709

81 0 23

UNIVERSITY OF WISCONSIN - MADISON  
MATHEMATICS RESEARCH CENTER

MEASURES OF LACK OF FIT FOR RESPONSE SURFACE DESIGNS  
AND PREDICTOR VARIABLE TRANSFORMATIONS<sup>†</sup>

G. E. P. Box and N. R. Draper

Technical Summary Report #2211  
April 1981



ABSTRACT

Some first and second order response surface designs are discussed from the point of view of their ability to detect certain likely kinds of lack of fit. This leads to consideration of conditions for representational adequacy of first and second order models in transformed predictor variables.

AMS(MOS) Subject Classification: 62J02, 62J05, 62K99

Key Words: First order designs, Lack of fit, Response surface designs, Second order designs, Transformations on predictors

Work Unit No. 4 - Statistics and Probability

---

<sup>†</sup>Department of Statistics Technical Report #639, University of Wisconsin, Madison.

---

Sponsored by the United States Army under Contract Nos. DAAG29-80-C-0041 and DAAG29-80-C-0113.

DISTRIBUTION STATEMENT  
Approved for public release  
Distribution is unlimited

MEASURES OF LACK OF FIT FOR RESPONSE SURFACE DESIGNS  
AND PREDICTOR VARIABLE TRANSFORMATIONS<sup>†</sup>

G. E. P. Box and N. R. Draper

1. INTRODUCTION

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

$$\begin{aligned} \eta = & \beta_0' + \left\{ \sum_{i=1}^k \beta_i' \xi_i \right\} + \left\{ \sum_{i=1}^k \sum_{j \geq 1}^k \beta_{ij}' \xi_i \xi_j \right\} \\ & + \left\{ \sum_{i=1}^k \sum_{j \geq i}^k \sum_{\ell \geq i}^k \beta_{ij\ell}' \xi_i \xi_j \xi_\ell \right\}. \end{aligned} \quad (1.1)$$

Denote the observed response by  $y$  and suppose that the errors  $\epsilon_u = y_u - \eta_u$  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"

$$x_i = (\xi_i - \xi_{i0})/s_i, \quad (1.2)$$

then we can write the model of degree  $d = 3$  in the form

$$\begin{aligned} \eta = & \beta_0 + \left\{ \sum_{i=1}^k \beta_i x_i \right\} + \left\{ \sum_{i=1}^k \sum_{j \geq i}^k \beta_{ij} x_i x_j \right\} \\ & + \left\{ \sum_{i=1}^k \sum_{j \geq i}^k \sum_{\ell \geq i}^k \beta_{ij\ell} x_i x_j x_\ell \right\}. \end{aligned} \quad (1.3)$$

<sup>†</sup>Department of Statistics Technical Report #639, University of Wisconsin, Madison.

Models of orders  $d = 1$  and  $d = 2$  are obtained by omitting the appropriate bracketed terms in (1.1) and (1.3). To obtain the most parsimonious representation, (usually corresponding to the lowest possible value of  $d$ ) it may be necessary to transform the response  $y$  and/or certain of the predictor variables  $\xi$ .

Now whenever we fit a model we face the possibility that some more complex model may be needed. We can try to resolve our doubts by employing a larger design which makes it possible to fit the more complex models, but then similar questions arise concerning that 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) which it is hoped will be adequate.
2. to employ an associated design which, 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  $y$  and/or  $\xi_1, \xi_2, \dots, \xi_k$ .
4. because there are situations where a more complex model cannot be avoided, to employ design arrangements which can be conveniently augmented to form larger designs appropriate for fitting and checking the more complex model.

Our object is to consider some appropriate checks and the possible elimination of the associated lack of fit by power transformations of the predictors  $\xi$ , for first and second order polynomial models and designs. Transformation of the response  $y$  (Bartlett, 1947; Box and Cox, 1964; Kruskal, 1968; Draper and Hunter, 1969) will not be considered here.

|                    |                                     |
|--------------------|-------------------------------------|
| Accession For      |                                     |
| ERIC CB&I          | <input checked="" type="checkbox"/> |
| ERIC T&B           | <input type="checkbox"/>            |
| Unannounced        | <input type="checkbox"/>            |
| Subscription       |                                     |
| Distribution/      |                                     |
| Availability Codes |                                     |
| Ann. no. of        |                                     |
| Dist. in           |                                     |
| <b>A</b>           |                                     |

## 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 which use (respectively) all or some of the  $2^k$  runs  $(\pm 1, \pm 1, \dots, \pm 1)$ . 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 contrast  $c_2$  between the average response  $\bar{y}_{c_0}$  at the center and the average response  $\bar{y}_c$  at the factorial points is

$$c_2 = \bar{y}_c - \bar{y}_{c_0}$$

with

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

(2.1)

Thus, in the common situation where the  $\beta_{ij}$  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 which 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  $\xi_1^{\lambda_1}, \xi_2^{\lambda_2}, \dots, \xi_k^{\lambda_k}$  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

$$\eta(\xi, \lambda) = \dot{\beta}_0 + \sum_{i=1}^k \dot{\beta}_i \xi_i^{\lambda_i} + \sum_{i=1}^k \sum_{j>i}^k \dot{\beta}_{ij} \xi_i^{\lambda_i} \xi_j^{\lambda_j}. \quad (2.2)$$

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  $j$ . In Appendix C, we show that this requires that

$$\eta_{ij} = 0, \quad i \neq j \quad (2.3)$$

$$\eta_{ii} + \delta_i(1-\lambda_i)\eta_i = 0, \quad i = 1, 2, \dots, k. \quad (2.4)$$

where

$$\eta_i = \left[ \frac{\partial \eta}{\partial x_i} \right]_{x_i=0}, \quad \eta_{ij} = \left[ \frac{\partial^2 \eta}{\partial x_i \partial x_j} \right]_{x_i=0} \quad (2.5)$$

and where

$$\delta_i = S_i / \xi_{i0}. \quad (2.6)$$



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

$$\hat{\eta}_i = b_i/S_i; \quad \hat{\eta}_{ij} = b_{ij}/(S_i S_j), \quad i \neq j; \quad \hat{\eta}_{ii} = 2b_{ii}/S_i^2, \quad (2.7)$$

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

$$\hat{\lambda}_i = 1 + 2b_{ii}/(\delta_i b_i), \quad i = 1, 2, \dots, k. \quad (2.8)$$

---

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

### 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 Eq. (1.3) with  $d = 2$  consists of the central composite arrangements in which a "cube", consisting of a two-level factorial with coded points  $(\pm 1, \pm 1, \dots, \pm 1)$  or a fraction of resolution  $R \geq 5$  is augmented by an added "star", with axial points at coded distance  $\alpha$ , and by  $n_0$  added center points at  $(0, 0, \dots, 0)$ . More generally, both the cube and the star might be replicated. A simple example of a design of this type for  $k = 2$  is defined by the columns headed  $x_1$  and  $x_2$  in Table 1. (We use  $n_c$  and  $n_s$  for the number of cube and star points and  $n_{co}$  and  $n_{so}$  for the number of center points associated with them, respectively. Thus  $n_0 = n_{co} + n_{so}$ .) Also shown are some manufactured data whose mode of generation is discussed in Appendix A.

The fitted least squares second degree equation is

$$\hat{y} = 30.59 - 4.22x_1 - 5.91x_2 - 1.66x_1^2 - 1.44x_2^2 - 3.41x_1x_2 \quad (3.1)$$

$\pm 0.25 \quad \pm 0.17 \quad \pm 0.17 \quad \pm 0.14 \quad \pm 0.14 \quad \pm 0.24$

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

Table 1. A composite design for  $k = 2$  predictor variables and its associated estimator columns;  $n_c = 8$ ,  $n_{co} = 1$ ,  $n_s = 4$ ,  $n_{so} = 5$ ,  $\alpha = 2$ .

| I | $x_1$ | $x_2$ | $x_1^2$ | $x_2^2$ | $x_1x_2$ | $x_{111}/1.5$ | $x_{222}/1.5$ | 8CC | Blocks | y    |
|---|-------|-------|---------|---------|----------|---------------|---------------|-----|--------|------|
| 1 | -1    | -1    | 1       | 1       | 1        | 1             | 1             | 1   | -1     | 37.5 |
| 1 | -1    | -1    | 1       | 1       | 1        | 1             | 1             | 1   | -1     | 38.3 |
| 1 | 1     | -1    | 1       | 1       | -1       | -1            | 1             | 1   | -1     | 34.7 |
| 1 | 1     | -1    | 1       | 1       | -1       | -1            | 1             | 1   | -1     | 35.1 |
| 1 | -1    | 1     | 1       | 1       | -1       | 1             | -1            | 1   | -1     | 27.7 |
| 1 | -1    | 1     | 1       | 1       | -1       | 1             | -1            | 1   | -1     | 29.2 |
| 1 | 1     | 1     | 1       | 1       | 1        | -1            | -1            | 1   | -1     | 12.2 |
| 1 | 1     | 1     | 1       | 1       | 1        | -1            | -1            | 1   | -1     | 11.4 |
| 1 | .     | .     | .       | .       | .        | .             | .             | -8  | -1     | 30.1 |
| 1 | -2    | .     | 4       | .       | .        | -2            | .             | -1  | 1      | 30.2 |
| 1 | 2     | .     | 4       | .       | .        | 2             | .             | -1  | 1      | 16.1 |
| 1 | .     | -2    | .       | 4       | .        | .             | -2            | -1  | 1      | 31.4 |
| 1 | .     | 2     | .       | 4       | .        | .             | 2             | -1  | 1      | 16.7 |
| 1 | .     | .     | .       | .       | .        | .             | .             | 0.8 | 1      | 30.5 |
| 1 | .     | .     | .       | .       | .        | .             | .             | 0.8 | 1      | 29.9 |
| 1 | .     | .     | .       | .       | .        | .             | .             | 0.8 | 1      | 29.9 |
| 1 | .     | .     | .       | .       | .        | .             | .             | 0.8 | 1      | 29.8 |
| 1 | .     | .     | .       | .       | .        | .             | .             | 0.8 | 1      | 30.2 |

Table 2. Analysis of variance associated with the second order model and its checks, for the data of Table 1.

| Source             | df | SS         | MS      | F     |
|--------------------|----|------------|---------|-------|
| Mean               | 1  | 13,938.934 |         |       |
| Blocks             | 1  | 7.347      | 7.347   |       |
| First order extra  | 2  | 842.907    | 421.453 |       |
| Second order extra | 3  | 188.142    | 62.714  |       |
| Lack of fit        | 3  | 88.923     | 29.691  | 64.85 |
|                    |    |            |         |       |
|                    |    |            |         |       |
|                    |    |            |         |       |
| Pure error         | 8  | 3.657      | 0.457   |       |
| Total              | 18 | 15,069.910 |         |       |

Before accepting the utility of the fitted equation we would 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 to throw light 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

$$(x_1^3, x_1 x_2^2); (x_2^3, x_2 x_1^2). \quad (3.2)$$

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  $\underline{x}$ -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  $x_1 x_2^2$ ), and so on. Then

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

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.

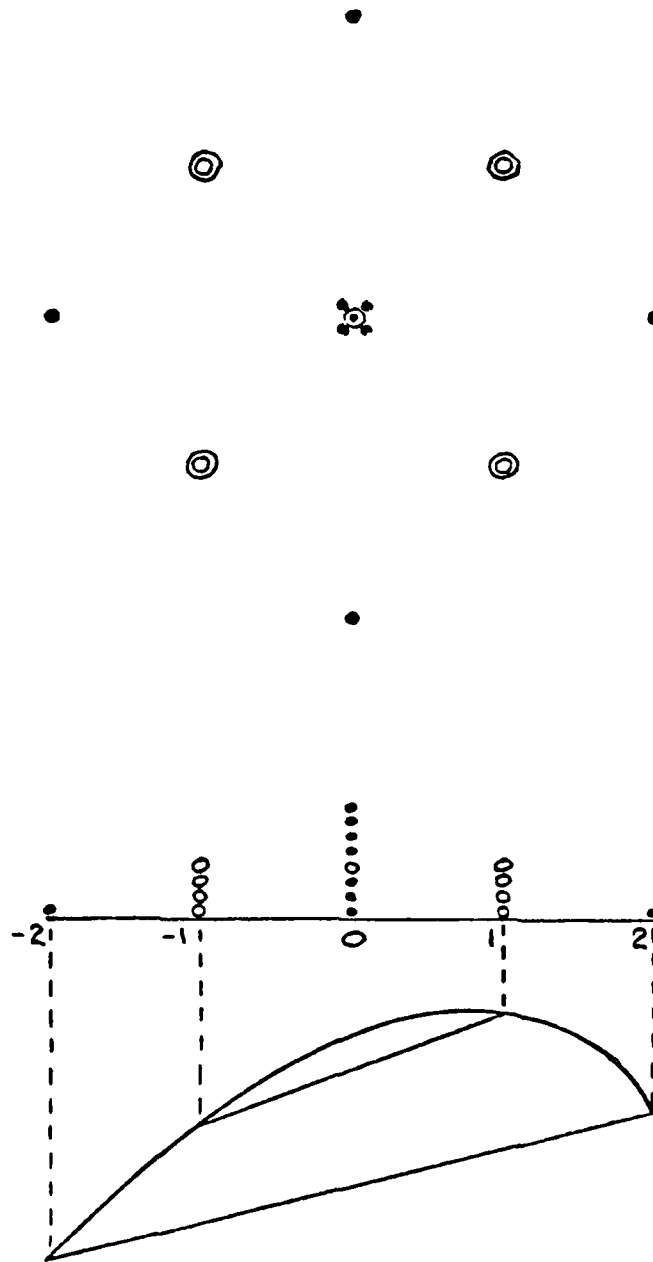
Consider now the column  $x_{111}$  in relation to Figure 1, which shows the projection of the points of the composite design onto the  $x_1$  axis. Denoting the average of the responses at  $x_1 = -\alpha, -1, 1, \alpha$  by  $\bar{y}_{-\alpha}, \bar{y}_{-1}, \bar{y}_1, \text{ and } \bar{y}_{\alpha}$ , respectively, we see that a contrast  $c_{31}$  associated with  $x_{111}$  is

$$c_{31} = \frac{1}{36} x'_{111} y = \frac{1}{3} \left[ \frac{\bar{y}_{\alpha} - \bar{y}_{-\alpha}}{2\alpha} - \frac{\bar{y}_1 - \bar{y}_{-1}}{2} \right], \quad (3.4)$$

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 non-quadraticity in the  $x_1$  direction. A corresponding measure in the  $x_2$  direction is, of course, given by  $c_{32} = x'_{222} y / 36$ .

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 return to discuss these data later.

Figure 1. Projection of design points on the  $x_1$  axis for the composite design in two factors given in Table 1. The contrast  $c_{3j}$  is an estimate of the difference between the slopes of the two chords.



### 3.2 General formulas.

In general, a composite design contains

(a) A "cube", consisting of a  $2^k$  factorial, or a  $2^{k-p}$  fractional factorial, made up of points of the type  $(\pm, \pm 1, \dots, \pm 1)$ , of resolution  $R \geq 5$  (Box and Hunter, 1961) replicated  $f(\geq 1)$  times. There are thus  $n_c = f2^{k-p}$  such points (where  $p$  may be zero).

(b) A "star", that is,  $2k$  points  $(\pm\alpha, 0, 0, \dots, 0)$ ,  $(0, \pm\alpha, 0, \dots, 0)$ ,  $\dots$ ,  $(0, 0, 0, \dots, \pm\alpha)$  on the predictor variable axes, replicated  $r$  times, so that there are  $n_s = 2kr$  points in all.

(c) Center points  $(0, 0, \dots, 0)$ ,  $n_o$  in number, of which  $n_{co}$  are in cube blocks and  $n_{so}$  in star blocks.

It is shown in Appendix B that, for any such design,  $k$  sets of columns can be isolated with the  $i$ th set containing the  $k$  columns  $x_i x_j^2$ ,  $j = 1, 2, \dots, k$ . This  $i$ th set is associated with a single vector  $x_{iii}$  which is orthogonal to the  $(k+1)(k+2)/2$  columns required for fitting the second degree equation and is also orthogonal to the  $(k-1)$  similarly constructed vectors  $x_{jjj}$ ,  $j \neq i$ .

The elements of these vectors are such that:

for the cube points,  $x_{iii} = \phi x_i$ , with  $\phi = 2r\alpha^2(1-\alpha^2)/(n_c + 2r\alpha^2)$ ;

for the star points,  $x_{iii} = \gamma x_i$ , with  $\gamma = n_c(\alpha^2 - 1)/(n_c + 2r\alpha^2)$ ;

for the center points  $x_{iii} = 0 = x_i$ .

Thus, the  $k$  estimates of third order lack of fit,  $c_{31}, c_{32}, \dots, c_{3k}$  are



$$c_{3i} = \frac{\sum_{iii} x_{iii}^3 y_{iii}}{\sum_{iii} x_{iii}^3} = \frac{1}{\alpha^2 - 1} \left\{ \frac{\bar{y}_{\alpha i} - \bar{y}_{-\alpha i}}{2\alpha} - \frac{\bar{y}_{1i} - \bar{y}_{-1i}}{2} \right\} \quad (3.5)$$

with standard deviation

$$\sigma_{c_3} = \frac{1}{\alpha^2 - 1} \left\{ \frac{1}{n_c} + \frac{1}{2r\alpha^2} \right\}^{1/2} \sigma. \quad (3.6)$$

Also

$$E(c_{3i}) = \beta_{iii} + (1 - \alpha^2)^{-1} \sum_{j \neq i}^k \beta_{ijj}. \quad (3.7)$$

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

$$SS(c_{3i}) = (\alpha^2 - 1)^2 c_{3i}^2 / \left\{ \frac{1}{n_c} + \frac{1}{2r\alpha^2} \right\}. \quad (3.8)$$

Note that even if  $E(c_{3i}) = 0$ , this does not necessarily mean there are no cubic coefficients. A combination of non-zero  $\beta_{iii}$  and  $\beta_{ijj}$  could occur for which  $\beta_{iii} + (1 - \alpha^2)^{-1} \sum \beta_{ijj} = 0$ . 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. In order to determine conditions that must be satisfied to make it possible to remove lack of fit in this way, and 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. In Appendix C we show that the conditions that must then apply if all third order coefficients of the transformed  $\xi_i^{\lambda_i}$  are to be zero are

$$\eta_{ijl} = 0, \quad \text{all } i \neq j \neq l = 1, 2, \dots, k; \quad (3.9)$$

$$\eta_{ijj} + \delta_i(1-\lambda_j)\eta_{ij} = 0, \quad i \neq j = 1, 2, \dots, k; \quad (3.10)$$

$$\eta_{iii} + 3\delta_i(1-\lambda_i)\eta_{ii} + \delta_i^2(1-\lambda_i)(1-2\lambda_i)\eta_i = 0, \quad i = 1, 2, \dots, k. \quad (3.11)$$

An important conclusion is thus the following. The possibility of second order representation in the transformed variables is contra-indicated if one (or more) interaction estimates  $b_{ijl}$  is (are) non-zero.

In practice, the estimation of the transformation (when not contra-indicated) is best done using nonlinear least squares directly on the model of Eq. (2.2); however some interesting light on how the curvature measures  $c_{3i}$  relate to these transformations is obtained by considering how Eqs. (3.10) and (3.11) could be used to obtain estimates of the  $\lambda_i$ .

A composite design does not permit all third order terms in Eq. (1.3) to be separately estimated. Suppose, however, that a second order model augmented with only cubic terms

$$\beta_{111}x_1^3 + \beta_{222}x_2^3 + \dots + \beta_{kkk}x_k^3, \quad (3.12)$$

was fitted. If the response  $\eta$  could be represented by the third order model of Eq. (1.3), the estimates  $b_i$  and  $b_{iii}$  obtained from the composite design would have expectations

$$E(b_i) = \eta_i - \frac{1}{2}\alpha^2(1-\alpha^2)^{-1} \sum_{j \neq i}^k \eta_{ijj} \quad (3.13)$$

$$E(b_{iii}) = \frac{1}{6}\eta_{iii} + \frac{1}{2}(1-\alpha^2)^{-1} \sum_{j \neq i}^k \eta_{ijj} \quad (3.14)$$

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 Eqs. (3.9)-(3.11) we obtain the following  $k$  equations for the  $\lambda_i$ . (In these equations,  $b_{iii} = c_{3i}$ .)

$$\begin{aligned}
 & b_{iii} + \delta_i(1-\lambda_i)b_{ij} + \frac{1}{6}\delta_i^2(1-\lambda_i)(1-2\lambda_i)b_i \\
 & + \frac{1}{2}(1-\alpha^2)^{-1} \left\{ 1 - \frac{1}{6}\alpha^2\delta_i^2(1-2\lambda_i)(1-\lambda_i) \sum_{j \neq i}^k \delta_j(1-\lambda_j)b_{ij} \right\} = 0, \\
 & i = 1, 2, \dots, k. \qquad (3.15)
 \end{aligned}$$

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

For the example data, this procedure converges to the values  $\hat{\lambda}_1 = -0.23$ ,  $\hat{\lambda}_2 = -0.93$ . These may be compared with the values  $\hat{\lambda}_1 = 0.09$ ,  $\hat{\lambda}_2 = -0.82$  provided by nonlinear least squares (these are maximum likelihood estimates under the standard normal error assumptions) and with  $\lambda_1 = 0$ ,  $\lambda_2 = -1$ , 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 Section 3.4 for details.)

Table 3. Analysis of variance for second order model in predictor variables  $\ln \xi_1$  and  $\xi_2^{-1}$ .

| Source  | df | SS         | MS      |      |        |
|---|----|------------|---------|------|--------|
| Mean  | 1  | 13,938.934 |         |      |        |
| Blocks  | 1  | 7.347      | 7.347   |      |        |
| First order extra   | 2  | 552.713    | 276.357 |      | 535.57 |
| Second order extra  | 3  | 565.238    | 188.413 |      | 365.14 |
| Lack of fit   | 3  | 2.021      | 0.674   | 1.47 |        |
| <div style="display: flex; align-items: center;"> <div style="margin-right: 10px;"> <math>\left\{ \begin{array}{l} \text{CC} \\ \text{Third order} \end{array} \right.</math> </div> <div style="margin-right: 10px;"> <math>\left\{ \begin{array}{l} 1 \\ 2 \end{array} \right.</math> </div> <div style="margin-right: 10px;"> <math>\left\{ \begin{array}{l} 1.396 \\ 0.625 \end{array} \right.</math> </div> <div style="margin-right: 10px;"> <math>\left\{ \begin{array}{l} 1.396 \\ 0.313 \end{array} \right.</math> </div> <div style="margin-right: 10px;"> <math>\left\{ \begin{array}{l} 3.05 \\ 0.68 \end{array} \right.</math> </div> </div> |    |            |         |      |        |
| Pure error  | 8  | 3.657      | 0.457   |      |        |
| Total   | 18 | 15,069.910 |         |      |        |

### 3.4. A curvature contrast

Consider the overall curvature measure  $c_2$  of Eq. (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  $\bar{y}_{s_0}$  and the average over all the star points is  $\bar{y}_s$ , then the contrast

$$c_2' = \frac{k}{\alpha} (\bar{y}_s - \bar{y}_{s_0})$$

has expectation

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

Thus the statistic

$$c_2 - c_2' = \bar{y}_c - \bar{y}_{c_0} - \frac{k}{\alpha^2} (\bar{y}_s - \bar{y}_{s_0}),$$

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  $c_2'$  associated with the star (block 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  $c_2 - c_2'$  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  $c_2 - c_2'$  is given by

$$\sigma_{c_2 - c_2'} = \left\{ \frac{1}{n_c} + \frac{1}{n_{c_0}} + \frac{k^2}{\alpha^4} \left[ \frac{1}{2kr} + \frac{1}{n_{s_0}} \right] \right\}^{1/2} \sigma$$

and the associated sum of squares for the analysis of variance table entry of Table 2 is obtained from

$$SS(c_2 - c'_2) = (c_2 - c'_2)^2 / \left\{ \frac{1}{n_c} + \frac{1}{n_{co}} + \frac{k}{2r\alpha^4} + \frac{k^2}{\alpha^4 n_{so}} \right\} .$$

For our example we find

$$c_2 - c'_2 = 1.3925 \pm 0.7520$$

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 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.

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 if it is possible to eliminate it by a power transformation in the predictors, but also actually to estimate the powers needed to effect the transformation. The theory is derived in Appendix C, and a worked two-factor example shows how to carry out the calculations.

It is also shown 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 are derived as a special case



REFERENCES

- Bartlett, M.S. (1947). The use of transformations. Biometrics, 3, 39-52.
- Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. J. Roy. Statist. Soc., B-26, 211-243, discussion 244-252.
- Box, G.E.P. and Draper, N.R. (1975). Robust designs. Biometrika, 62, 347-352.
- Box, G.E.P. and Hunter, J.S. (1961). The  $2^{k-p}$  fractional factorial designs, Part I. Technometrics, 3, 311-351; Part II, 449-458.
- Box, G.E.P., Hunter, W.G. and Hunter, J.S. (1978). Statistics for Experimenters. New York: Wiley.
- Box, G.E.P. and Wetz, J. (1973). Criteria for judging adequacy of estimation by an approximating response function. University of Wisconsin-Madison Statistics Department Technical Report No. 9.
- Box, G.E.P. and Wilson, K.B. (1951). On the experimental attainment of optimum conditions. J. Roy. Statist. Soc. B, 13, 1-45.
- DeBaun, R.M. (1956). Block effects in the determination of optimum conditions. Biometrics, 12, 20-22.
- Draper, N.R. and Herzberg, A.M. (1971). Lack of fit. Technometrics, 13, 231-241.
- Draper, N.R. and Hunter, W.G. (1969). Transformations: some examples revisited. Technometrics, 11, 23-40.
- Draper, N.R. and Smith, H. (1981). Applied Regression Analysis, 2nd edition. New York: Wiley.
- Kruskal, J.B. (1968). Statistical analysis: transformations of data. International Encyclopedia of the Social Sciences. New York: Macmillan (Free Press Div.).

ACKNOWLEDGEMENTS

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 A1 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 A1(b) shows a quadratic response function with a simple maximum in variables  $(\ln \xi_1, 100 \xi_2^{-1})$ . This figure is redrawn in the metrics  $(\xi_1, \xi_2)$  in Figure A1(a). In the  $(\xi_1, \xi_2)$  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,  $\xi_{10} = 2.5$ ,  $\xi_{20} = 12.5$ ,  $S_1 = 0.75$ , and  $S_2 = 3.75$ .

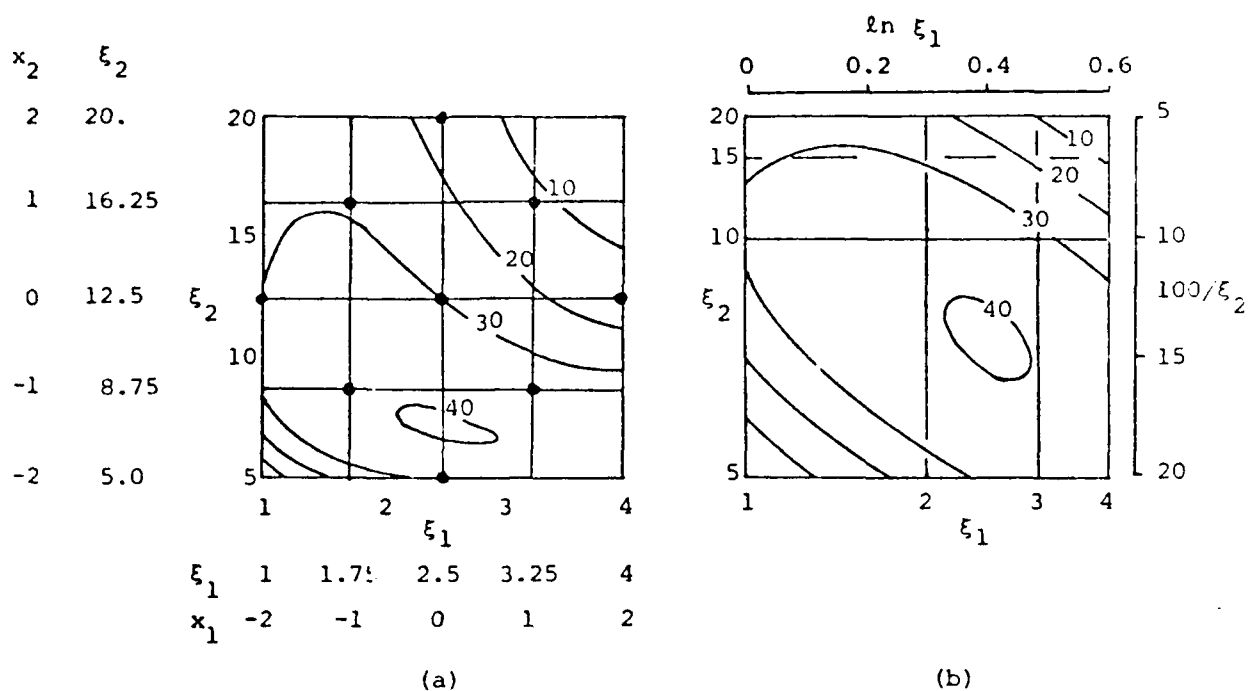


Figure A1. Generation of example data. Surface (a) becomes quadratic, as in (b), when variables  $\ln \xi_1$  and  $\xi_2^{-1}$  are employed. Data were taken from (a) with added errors.

Appendix B. The third order columns of the  $\underline{X}$ -matrix for a composite design.

We here prove the results summarized in Section 3.2. The full cubic model in  $k$  variables  $x_1, x_2, \dots, x_k$  is given by Eq. (3.2). The form of the  $\underline{X}$  matrix in the regression model  $\underline{y} = \underline{X}\underline{\beta} + \underline{\epsilon}$  when the design consists of  $f$  "cubes" plus  $r$  "stars" plus  $n_0$  center points is as shown in Table B1. We can denote columns by placing square brackets around the column head; for example  $[x_1]$  will denote the  $x_1$  column, and so on. We write  $n_c = f2^{k-p}$  for the number of cube points.

All of the cubic columns are orthogonal to all of the other columns with the following exceptions:  $[x_i^3]$  is not orthogonal to  $[x_i]$ , nor to  $[x_i x_j^2]$ ;  $[x_i x_j^2]$  is not orthogonal to  $[x_i]$ , nor to  $[x_i^3]$ , nor to  $[x_i x_j^2]$ . The first step is to regress the  $[x_i^3]$  and  $[x_i x_j^2]$  vectors on the  $[x_i]$  and take residuals. Because the columns involved are orthogonal to  $[x_0]$ , no adjustment for means is needed. We denote the "cube portion" of the  $[x_i]$  and  $[x_i x_j^2]$  vectors by  $c_i$ , as indicated in the table. These two sets of residuals are, where the prime denotes transpose,

$$[x_{ijj}] = [x_i^3] - \{[x_i]'[x_i^3]/[x_i]'[x_i]\}[x_i] \quad (A1)$$

and

$$[x_{ijj}] = [x_i x_j^2] - \{[x_i]'[x_i x_j^2]/[x_i]'[x_i]\}[x_i] \quad (A2)$$

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

Table B1. The  $X$  Matrix for a Cubic Model in  $k$  Predictors  $x_1, x_2, \dots, x_k$ .

| $x_0$    | $x_1$      | $x_2$      | $\dots$ | $x_k$     | $x_1^2$    | $x_2^2$    | $\dots$ | $x_k^2$    | $x_1x_2$   | $\dots$              | $x_{k-1}x_k$ | $x_1^3$     | $x_1x_2^2$ | $\dots$                                     | $x_1^2x_3$ | $\dots$    |
|----------|------------|------------|---------|-----------|------------|------------|---------|------------|------------|----------------------|--------------|-------------|------------|---|------------|------------|
| 1        | $\zeta_1$  | $\zeta_2$  | $\dots$ | $\zeta_k$ |            |            |         |            |            |                      |              |             |            |   |            |            |
|          | $\pm 1$ 's |            |         |           | 1's        |            |         |            | $\pm 1$ 's | Interaction patterns | $\pm 1$ 's   |             | $\pm 1$ 's | each column same pattern as in $x_1$ column |            | $\pm 1$ 's |
| 1        | $-\alpha$  | $\alpha^2$ |         | 0         | $\alpha^2$ | 0          |         | 0          |            |                      |              | $-\alpha^3$ |            |   |            |            |
| 1        | $\alpha$   | 0          |         | 0         | $\alpha^2$ | 0          |         | 0          |            |                      |              | $\alpha^3$  |            |   |            |            |
| 1        | $-\alpha$  | 0          |         | 0         | $\alpha^2$ | 0          |         | 0          |            |                      |              | $-\alpha^3$ |            |   |            |            |
| 1        | $\alpha$   | 0          |         | 0         | $\alpha^2$ | 0          |         | 0          |            |                      |              | $\alpha^3$  |            |   |            |            |
| $\vdots$ | $\vdots$   | $\vdots$   |         | $\vdots$  | $\vdots$   | $\vdots$   |         | $\vdots$   |            |                      |              | $\vdots$    |            |   |            |            |
| $\vdots$ | 0          | 0          |         | 0         | 0          | 0          |         | 0          |            |                      |              | 0           |            |   |            | 0          |
| $\vdots$ | 0          | 0          |         | 0         | 0          | 0          |         | 0          |            |                      |              | 0           |            |   |            | 0          |
| 1        | $-\alpha$  | $\alpha^2$ |         | 0         | $\alpha^2$ | $\alpha^2$ |         | $\alpha^2$ |            |                      |              | 0           |            |   |            | 0          |
| 1        | $\alpha$   | 0          |         | 0         | $\alpha^2$ | $\alpha^2$ |         | $\alpha^2$ |            |                      |              | 0           |            |   |            | 0          |
| $\vdots$ | $-\alpha$  | 0          |         | $\alpha$  | $\alpha^2$ | $\alpha^2$ |         | $\alpha^2$ |            |                      |              | 0           |            |   |            | 0          |
| 1        | 0          | 0          |         | 0         | 0          | 0          |         | 0          |            |                      |              | 0           |            |   |            | 0          |
| 1        | 0          | 0          |         | 0         | 0          | 0          |         | 0          |            |                      |              | 0           |            |   |            | 0          |
| $\vdots$ | $\vdots$   | $\vdots$   |         | $\vdots$  | $\vdots$   | $\vdots$   |         | $\vdots$   |            |                      |              | $\vdots$    |            |   |            | $\vdots$   |
| 1        | 0          | 0          |         | 0         | 0          | 0          |         | 0          |            |                      |              | 0           |            |   |            | 0          |

$$[x_{111}]' = [c_1^1, d, -d, d, -d, \dots, d, -d, 0, 0, \dots, 0](1-\alpha^2)^m, \quad (A3)$$

where

$$d = n_c / (2r\alpha), \quad (A4)$$

and where there are  $r$  sets of  $(d, -d)$ 's in the vector. In general, for  $[x_{iii}]'$ ,  $c_1^1$  will be replaced by  $c_i^i$  and the position of the  $\pm d$ 's will correspond to those of the  $\mp \alpha$ 's in the corresponding  $[x_i]$ ' vector. Note that, because  $c_i^i c_j^j = 0$ ,  $i \neq j$ , it is obvious that  $[x_{iii}]$  and  $[x_{jjj}]$  are orthogonal.

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

$$\begin{aligned} \ell_{iii} &= [x_{iii}]' \underline{y} / [x_{iii}]' [x_{iii}] \quad (A5) \\ &= \{c_i^i \underline{y}_1 + d(-r\bar{y}_{\alpha i} + r\bar{y}_{-\alpha i})\} / \{n_c(1-\alpha^2)\} \end{aligned}$$

where  $\underline{y}_1$  is the portion of  $\underline{y}$  corresponding to the cube part of the design, and  $\bar{y}_{\alpha i}$ ,  $\bar{y}_{-\alpha i}$  are, respectively the averages of observations taken at the  $\alpha$  and  $-\alpha$  axial points on the  $x_i$  axis. If we similarly denote by  $\bar{y}_{1i}$  and  $\bar{y}_{-1i}$  the averages of the  $n_c/2$  observations in  $\underline{y}_1$  corresponding to 1 and -1 in  $c_i^i$ , respectively, it follows quickly that  $\ell_{iii} = c_{3i}$  where  $c_{3i}$  is given in Eq. (3.5). The expected value is

$$E(c_{3i}) = [x_{iii}]' \underline{X}' \underline{\beta} / [x_{iii}]' [x_{iii}] \quad (A6)$$

where  $\underline{X}$  is as in Table B1 and the coefficients of  $\underline{\beta}$  correspond to the columns in the obvious manner. Because  $[x_{iii}]$  is orthogonal to all columns of  $\underline{X}$  except the  $[x_i^3]$  and  $[x_i x_j^2]$  columns, Eq. (3.7) emerges almost immediately.

Appendix C. Conditions for efficacy of transformations on the predictor variables.

Suppose that an experimental design is run in the  $k$  coded variables

$$x_i = (\xi_i - \xi_{i0})/S_i, \quad i = 1, 2, \dots, k, \quad (C1)$$

in a situation where the underlying response function can be approximated by a second degree polynomial  $F\{(\xi_i^{\lambda_i})\}$  in the transformed original variables  $\xi_i^{\lambda_i}$ . Thus

$$F(\xi, \lambda) = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i \xi_i^{\lambda_i} + \sum_{i=1}^k \sum_{j \geq i}^k \hat{\beta}_{ij} \xi_i^{\lambda_i} \xi_j^{\lambda_j}. \quad (C2)$$

Assume all  $\lambda_i \neq 0$ . (The case when any  $\lambda_i = 0$  can be handled as a limiting case using the fact (see, for example, Box and Cox, 1964) that  $(\xi^\lambda - 1)/\lambda$  tends to  $\ln \xi$  as  $\lambda$  tends to zero.)

If (C2) is to be a suitable representation, then all third derivatives with respect to the  $\xi_i^{\lambda_i}$  must vanish identically. Note that

$$\frac{\partial F}{\partial \xi_i} \frac{\partial \xi_i}{\partial \lambda_i} = F_i(\xi_i^{1-\lambda_i}/\lambda_i) \quad (C3)$$

say, where  $F_i = \partial F / \partial \xi_i$ . Moreover, because of Eq. (C1),



$$\eta_i = \frac{\partial F}{\partial x_i} = \frac{\partial F}{\partial \xi_i} \frac{\partial \xi_i}{\partial x_i} = F_i S_i \quad (C4)$$

so that

$$\frac{\partial F}{\lambda_i \partial \xi_i} = \eta_i S_i^{-1} (\xi_i^{1-\lambda_i} / \lambda_i). \quad (C5)$$

The obvious extensions of these results also follow for the higher derivatives which involve terms of the form

$$\eta_{ij} = \frac{\partial^2 F}{\partial x_i \partial x_j} \quad \text{and} \quad \eta_{ijl} = \frac{\partial^3 F}{\partial x_i \partial x_j \partial x_l}. \quad (C6)$$

If we now carry out the appropriate differentiations, we obtain

$$\frac{\partial F}{\lambda_i \partial \xi_i} = \frac{\xi_i^{1-\lambda_i}}{S_i \lambda_i} \eta_i \quad (C7)$$

$$\frac{\partial^2 F}{\lambda_i \lambda_j \partial \xi_i \partial \xi_j} = \frac{\xi_i^{1-\lambda_i} \xi_j^{1-\lambda_j}}{S_i S_j \lambda_i \lambda_j} \eta_{ij} \quad (C8)$$

$$\frac{\partial^2 F}{\partial^2(\epsilon_i)} = \frac{\epsilon_i^{1-2\lambda_i}}{S_i \lambda_i^2} \left[ (1-\lambda_i) \eta_i + \frac{\epsilon_i}{S_i} \eta_{ii} \right] \quad (C9)$$

$$\frac{\partial^2 F}{\partial \epsilon_i^{\lambda_i} \partial^2(\epsilon_j)} = \frac{\epsilon_i^{1-\lambda_i} \epsilon_j^{1-2\lambda_j}}{S_i S_j \lambda_i \lambda_j^2} \left[ \frac{\epsilon_j}{S_j} \eta_{ijj} + (1-\lambda_j) \eta_{ij} \right] = 0 \quad (C10)$$

$$\frac{\partial^3 F}{\partial^3(\epsilon_i)} = \frac{\epsilon_i^{1-3\lambda_i}}{S_i \lambda_i^3} \left[ \left(\frac{\epsilon_i}{S_i}\right)^2 \eta_{iii} + \frac{3\epsilon_i}{S_i} (1-\lambda_i) \eta_{ii} + (1-\lambda_i)(1-2\lambda_i) \eta_i \right] = 0 \quad (C11)$$

$$\frac{\partial^3 F}{\partial \epsilon_i^{\lambda_i} \partial \epsilon_j^{\lambda_j} \partial \epsilon_\ell^{\lambda_\ell}} = \frac{\epsilon_i^{1-\lambda_i} \epsilon_j^{1-\lambda_j} \epsilon_\ell^{1-\lambda_\ell}}{S_i S_j S_\ell \lambda_i \lambda_j \lambda_\ell} \eta_{ij\ell} = 0. \quad (C12)$$

Eqs. (C10) - (C12) are exact conditions on the  $\lambda$ 's. The  $\eta_i$ ,  $\eta_{ij}$  and  $\eta_{ijk}$  also involve the  $\lambda$ 's, and cannot be specifically evaluated if the  $\lambda$ 's are unknown. If estimates of the  $\eta_i$ ,  $\eta_{ij}$ , and  $\eta_{ij\ell}$  are substituted, however, Eqs. (C10) - (C12) can be solved to provide estimates of the  $\lambda$ 's.

We now assume that the response surface can be approximately represented by a cubic polynomial in the coded predictor variables, namely, by

$$\begin{aligned}
 \hat{\eta} = & b_0 + b_1 x_1 + \dots + b_k x_k \\
 & + b_{11} x_1^2 + \dots + b_{kk} x_k^2 \\
 & + b_{12} x_1 x_2 + \dots + b_{k-1,k} x_{k-1} x_k \\
 & + b_{111} x_1^3 + b_{122} x_1 x_2^2 + \dots \\
 & + b_{222} x_2^3 + b_{112} x_1^2 x_2 + \dots \\
 & + b_{kkk} x_k^3 + b_{11k} x_1^2 x_k + \dots \\
 & + b_{123} x_1 x_2 x_3 + \dots
 \end{aligned}
 \tag{C13}$$

We can estimate the  $\eta_i$ ,  $\eta_{ij}$ , and  $\eta_{ijl}$  by the corresponding derivatives of  $\hat{\eta}$  evaluated at the center of the design, that is, at  $x_i = 0$ . In general then,

$$\begin{aligned}
 \hat{\eta}_i = b_i, & \quad \hat{\eta}_{ii} = 2b_{ii}, & \quad \hat{\eta}_{iii} = 6b_{iii}, \\
 \hat{\eta}_{ijl} = b_{ijl}, & \quad \hat{\eta}_{ijjj} = 2b_{ijjj}, & \quad \hat{\eta}_{ijjll} = b_{ijjll},
 \end{aligned}
 \tag{C14}$$

and we would substitute these in (C10) - (C12), at the same time setting  $\xi_i = \xi_{i0}$ , its value when  $x_i = 0$ . Note that, from Eq. (C12), we require that  $\eta_{ijl} = 0$ , which implies that the assumed transformations  $\xi_i^{\lambda_i}$  are not suitable representations unless all three-factor interactions are zero. In practice, then, we would want  $b_{ijl}$  to be small and nonsignificant when  $i \neq j \neq l$  or else we have a clear indication that the  $\xi_i^{\lambda_i}$  will not provide a satisfactory second order representation in (C2). If this aspect is satisfied, however, we now solve the  $\frac{1}{2}k(k+1) + k = \frac{1}{2}k(k+3)$  simultaneous equations:

$$\hat{n}_{ijj} + \delta_j(1-\lambda_j)\hat{n}_{ij} = 0, \quad i \neq j = 1, 2, \dots, k,$$

$$\hat{n}_{iii} + 3\delta_i(1-\lambda_i)\hat{n}_{ii} + \delta_i^2(1-\lambda_i)(1-2\lambda_i)\hat{n}_i = 0, \quad i = 1, 2, \dots, k \quad (C15)$$

where  $\delta_i = S_i/\epsilon_{i0}$ , and we have divided through Eqs. (C10) and (C11) by factors assumed to be non-zero, namely  $S_i$ ,  $\epsilon_{i0}$ ,  $\lambda_i$ , using the  $\hat{n}$  values from Eq. (C14).

#### Composite Designs

An additional complication arises with composite designs. For such designs, we cannot estimate all the third order coefficients individually, as we have explained in Appendix B. This means that, while Eqs. (C15) are still valid, the values in (C14) cannot be used. For a composite design, the column vectors  $x_i$ ,  $x_i^3$ , and  $x_i x_j^2$  are linearly dependent via

$$x_i x_j^2 = (-\alpha^2 x_i + x_i^3)/(1-\alpha^2). \quad (C16)$$

Thus in the general cubic model, we cannot estimate all  $k+k+k(k-1) = k(k+1)$  coefficients in the terms

$$\beta_i x_i + \beta_{iii} x_i^3 + \sum_{j \neq i}^k \beta_{ijj} x_i x_j^2 \quad (C17)$$

but only the  $2k$  coefficients of

$$\left\{ \beta_i - \frac{\alpha^2}{1-\alpha^2} \sum_{j \neq i}^k \beta_{ijj} \right\} x_i + \left\{ \beta_{iii} + \frac{1}{1-\alpha^2} \sum_{j \neq i}^k \beta_{ijj} \right\} x_i^3. \quad (C18)$$

It follows that, in the model so reduced,

$$b_j \text{ estimates } \eta_j - \frac{\alpha^2}{1-\alpha^2} \frac{1}{2} \sum_{j \neq i}^k \eta_{ijj} \quad (C19)$$

while

$$b_{iii} \text{ estimates } \frac{1}{6} \eta_{iii} + \frac{1}{1-\alpha^2} \frac{1}{2} \sum_{j \neq i}^k \eta_{ijj}. \quad (C20)$$

(By examining Eqs. (3.7), (C14), and the second portion of Eq. (C18), we infer that  $b_{iii} = c_{3i}$ .) Alternatively, if the model is fitted using  $x_i$  and the "orthogonalized  $x_i^3$ ", namely  $x_{iii} = x_i^3 - \psi x_i$ , the terms of the model are

$$\left\{ \beta_i + \psi \beta_{iii} + \theta \sum_{j \neq i}^k \beta_{ijj} \right\} x_i + \left\{ \beta_{iii} + \frac{1}{1-\alpha^2} \sum_{j \neq i}^k \beta_{ijj} \right\} (x_i^3 - \psi x_i) \quad (C21)$$

where

$$\theta = n_c / (n_c + 2r\alpha^2), \quad \psi = (n_c + 2r\alpha^4) / (n_c + 2r\alpha^2). \quad (C22)$$

In this form, we have that

$$b_i^* \text{ estimates } \eta_i + \frac{1}{6}\psi\eta_{iii} + \frac{1}{2}\theta \sum_{j \neq i}^k \eta_{ijj} \quad (C23)$$

and

$$b_{iii}^* \text{ estimates } \frac{1}{6}\eta_{iii} + \frac{1}{1-\alpha^2} \frac{1}{2} \sum_{j \neq i}^k \eta_{ijj} \quad (C24)$$

where  $b_i^*$  is the estimated coefficient of  $x_i$  and  $b_{iii}^*$  is that for  $(x_i^3 - \psi x_i)$ . (Note that  $b_{iii}^* = b_{iii} = c_{3i}$ , but that  $b_i^* = b_i + \psi b_{iii}$ .) We now describe how this affects the estimation of the  $\lambda_i$ . From Eqs. (C10) and (C11), and setting  $\xi_i = \xi_{i0}$  we have

$$\eta_{ijj} + \delta_j(1-\lambda_j)\eta_{ij} = 0 \quad (C25)$$

and

$$\eta_{iii} + 3\delta_i(1-\lambda_i)\eta_{ii} + \delta_i^2(1-\lambda_i)(1-2\lambda_i)\eta_i = 0 \quad (C26)$$

We now combine  $\frac{1}{6}$  times (C26) with  $\left\{ \frac{1}{2(1-\alpha^2)} - \frac{\alpha^2}{12(1-\alpha^2)} \delta_i^2(1-2\lambda_i)(1-\lambda_i) \right\}$  times (C25) summed over  $j \neq i$  to give, for  $i = 1, 2, \dots, k$ ,

$$\begin{aligned}
 & \left\{ \frac{1}{6} \eta_{iii} + \frac{1}{2} \frac{1}{1-\alpha^2} \sum_{j \neq i}^k \eta_{ijj} \right\} \\
 & + \frac{1}{2} \delta_i (1-\lambda_i) \eta_{ii} \\
 & + \frac{1}{6} \delta_i^2 (1-\lambda_i) (1-2\lambda_i) \left\{ \eta_i - \frac{\alpha^2}{1-\alpha^2} \frac{1}{2} \sum_{j \neq i}^k \eta_{ijj} \right\} \\
 & + \frac{1}{2} \frac{1}{1-\alpha^2} \left\{ 1 - \frac{1}{6} \alpha^2 \delta_i^2 (1-2\lambda_i) (1-\lambda_i) \right\} \sum_{j \neq i}^k \delta_j (1-\lambda_j) \eta_{ij} = 0.
 \end{aligned} \tag{C27}$$

Thus, if the third order model is fitted in terms of  $x_i$  and  $x_i^3$  (rather than  $x_i$  and  $(x_i^3 - \psi x_i)$  as described above) we can substitute appropriate estimates to give the  $k$  simultaneous equations for  $i = 1, 2, \dots, k$ .

$$\begin{aligned}
 & b_{iii} + \delta_i (1-\lambda_i) b_{ii} + \frac{1}{6} \delta_i^2 (1-\lambda_i) (1-2\lambda_i) b_i \\
 & + \frac{1}{2} \frac{1}{1-\alpha^2} \left\{ 1 - \frac{1}{6} \alpha^2 \delta_i^2 (1-2\lambda_i) (1-\lambda_i) \right\} \sum_{j \neq i}^k \delta_j (1-\lambda_j) b_{ij} = 0
 \end{aligned} \tag{C28}$$

which can be solved for  $\lambda_1, \lambda_2, \dots, \lambda_k$ . These awkward equations can be difficult to solve unless some care is applied. We suggest an iterative procedure in which rough estimates of  $\lambda_i$  are used in the grouping  $Q_i \equiv (1-\lambda_i)(1-2\lambda_i)$  which occurs in two positions in Eq. (C28). The resulting linear equations in  $\theta_i = 1-\lambda_i$  are straightforward to solve, and the results are used in the grouping  $(1-\lambda_i)(1-2\lambda_i)$  for a second iteration and so on, until convergence is

achieved. To aid convergence, each new iteration can be started from the midpoint of the old and new values, if desired.

#### Alternative Cubic Case

If the cubic is fitted with estimated terms  $b_i^*x_i + b_{iii}^*(x_i^3 - \psi x_i)$  as described above, we obtain  $b_i$  and  $b_{iii}$  from

$$b_i = b_i^* - \psi b_{iii}^* \quad \text{and} \quad b_{iii} = b_{iii}^* \quad (\text{C29})$$

#### Related Work

The expression in (C20) can be alternatively written in the form

$$\beta_{iii} + (1-\alpha^2)^{-1} \sum \beta_{i,ij}$$

Draper and Herzberg (1971, p. 236) show that the sum of squares of these quantities for  $i = 1, 2, \dots, k$  occurs in the expected value of a general measure of lack of fit  $L_1$ . Thus, the  $c_{3i}$  contrasts essentially provide a split-up of  $L_1$  which permits a more detailed and sensitive analysis. The remaining degrees of freedom pertaining to  $L_1$  can be attributed to other contrasts as already described.



### The First Degree Case

We return to the beginning of this Appendix to examine the simpler case when the underlying response surface can be approximated by a first degree polynomial in the transformed variables  $\xi_i^{\lambda_i}$ . In such a case

1. All  $\beta'_{ij} = 0$ , in (C2).
2. All second derivatives with respect to  $\xi_i^{\lambda_i}$  must vanish identically.

This condition provides, from Eqs. (C8) and (C9), the equations (2.3) and (2.4).

| REPORT DOCUMENTATION PAGE  |                                      | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM  |
|--|--------------------------------------|--|
| 1. REPORT NUMBER<br>2211   | 2. GOVT ACCESSION NO.<br>AD-A100 614 | 3. RECIPIENT'S CATALOG NUMBER  |
| 4. TITLE (and Subtitle)<br>MEASURES OF LACK OF FIT FOR RESPONSE SURFACE<br>DESIGNS AND PREDICTOR VARIABLE TRANSFORMATIONS  |                                      | 5. TYPE OF REPORT & PERIOD COVERED<br>Summary Report - no specific<br>reporting period               |
|  |                                      | 6. PERFORMING ORG. REPORT NUMBER   |
| 7. AUTHOR(s)<br>G. E. P. Box and H. R. Draper  |                                      | 8. CONTRACT OR GRANT NUMBER(s)<br>DAAG29-80-C-0041<br>DAAG29-80-C-0113                               |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Mathematics Research Center, University of<br>610 Walnut Street Wisconsin<br>Madison, Wisconsin 53706   |                                      | 10. PROGRAM ELEMENT, PROJECT, TASK,<br>AREA & WORK UNIT NUMBERS<br>4 - Statistics and<br>Probability |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>U. S. Army Research Office<br>P. O. Box 12211<br>Research Triangle Park, North Carolina 27709   |                                      | 12. REPORT DATE<br>April 1981  |
|  |                                      | 13. NUMBER OF PAGES<br>37  |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)  |                                      | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED   |
|  |                                      | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE  |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited.  |                                      |  |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)   |                                      |  |
| 18. SUPPLEMENTARY NOTES  |                                      |  |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>First order designs, Lack of fit, Response surface designs, Second order<br>designs, Transformations on predictors   |                                      |  |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>Some first and second order response surface designs are discussed from the<br>point of view of their ability to detect certain likely kinds of lack of fit.<br>This leads to consideration of conditions for representational adequacy of first<br>and second order models in transformed predictor variables. |                                      |  |

**DAT  
FILM**