# ON SOME ASPECTS OF SECOND ORDER RESPONSE SURFACE METHODOLOGY.

A thesis presented in partial fulfilment of the requirements for the degree of M.Sc. in Mathematics at Massey University.

Vernon John Thomas

1972

# Abstract

A unified development of the theoretical basis of response surface methodology, particularly as it applies to second order response surfaces, is presented. A rigorous justification of the various tests of hypothesis usually used is given, as well as a convenient means of making tests on whole factors, rather than on terms of a given degree, as is customary at present. Finally, the super--imposition of some elementary classification designs on a response surface design is considered.

## Acknowledgement

I would like to take this opportunity of expressing my gratitude to Dr B.S.Weir for his comments and encouragement in the preparation of this thesis.

# Table of contents

# 1. Response Surface Methodology.

## Introduction

Response surface methodology seeks to estimate, by regression methods, that linear combination of previously specified graduating functions of a number of independent variables which provides, in some sense, the best fit to an observed response.

While the techniques of fitting are identical with, or closely related to, those of multiple linear regression, the emphasis is slightly different, in that considerable stress is laid on the design aspect of the problem. It is assumed that the levels of the independent variables may be pre-specified at will, within broad limits. The space defined on the independent variables, and within these limits, is termed the region of operability. The sub-space of this region, in which estimates of response are of interest to the experimenter, is termed the region of interest.

Typically, a number of experiments are carried out, according to some previously decided experimental plan. Each experiment consists of the measurement of an observed response at a point defined by some combination of the independent variables. In some cases, sequential designs are used - that is, the curve fitted to date is used as information to assist in the specification of the combination of independent variables to be used in the next experiment.

The basic variability of the observed response is measured by replication of experimental points, or by the residual error of the observed response from the fitted surface. This latter error can arise from a true observational error or from inadequate specification of the model, whereas the error based on point replication estimates true experimental error only. For this reason, when point replication is used, the residual error may be used to test model adequacy.

## The model

The model is developed by assuming p independent variables, given by

$$\underset{\sim}{\xi} = (\xi_1 \ldots \xi_p)^T$$

and k pre-specified vector graduating functions of these variables, given by

$$\underset{\sim}{x} = \underset{\sim}{x}(\underset{\sim}{\xi}) \text{ where } \underset{\sim}{x} \text{ is kx1}$$

The observational response is assumed (or known) to be

$$y = \eta + \epsilon$$

where $\epsilon$ is a random variate with zero mean, and the so-called "true response" $\eta$ is given by the exact relationship

$$\eta = \underset{\sim}{x}^T \underset{\sim}{\beta} \qquad (1.1)$$

where $\underset{\sim}{\beta}$ is a vector of unknown coefficients. The measurement of an observed y, for some known $\underset{\sim}{\xi}$, is termed an experiment. The values of $\epsilon$ arising from different experiments are assumed to be statistically independent, with constant, unknown, variance $\sigma^2$.

The aim of the sequence of experiments is to estimate $\boldsymbol{\beta}$ by $\underset{\sim}{b}$, and, from this, to estimate the response at any point of the region of interest by

$$\hat{y} = \underset{\sim}{x}^T \underset{\sim}{b}$$

To achieve this, n experiments are conducted, at the points $\underset{\sim}{\xi}_u$, u=1, ... n, yielding n observed responses

$$\underset{\sim}{y} = (y_1 \dots y_n)^T$$

Now let

$$\Xi = (\underset{\sim}{\xi}_1 \dots \underset{\sim}{\xi}_n)^T \text{ of dimension } n \times p$$

$$\underset{\sim}{x}_u = \underset{\sim}{x}(\underset{\sim}{\xi}_u)$$

and

$$X = (\underset{\sim}{x}_1 \dots \underset{\sim}{x}_n)^T \text{ of dimension } n \times k$$

so that $\underset{\sim}{y}$ is the observed value of the true response $X\boldsymbol{\beta}$.

Properly speaking, $\Xi$ is the design matrix, since $\Xi$ determines X. However, once $\Xi$ is chosen, according to some design criterion, it is convenient to refer to X as the design matrix, since all operations are in terms of X.

In the vast majority of applications, $\underset{\sim}{x}$ consists of all powers of the $\xi$, separately or together, up to some maximum degree d. The design is then referred to as a dth order design. Thus, for a second order design

$$\underset{\sim}{x}(\underset{\sim}{\xi}) = (1;\ \xi_1 \dots \xi_p;\ \xi_1^2 \dots \xi_p^2;\ \xi_1\xi_2 \dots \xi_{p-1}\xi_p)^T$$

For this type of design, it is convenient to use the subscripts occurring in the corresponding element of $\underset{\sim}{x}$ to identify the elements of $\underset{\sim}{\beta}$, thus, for second order designs,

$$\underset{\sim}{\beta} = (\beta_0; \beta_1 \cdots \beta_p; \beta_{11} \cdots \beta_{pp}; \beta_{12} \cdots \beta_{(p-1)p})^T$$

In general, for a dth order design, there will be $\binom{p+d}{d}$ coefficients.

Within this framework, $\underset{\sim}{x}^T\underset{\sim}{\beta}$ is a general dth order polynomial in p variables.

The exceptions to this kind of polynomial are of two types. In the first type, the elements of $\underset{\sim}{x}$ are not powers of the elements of $\underset{\sim}{\xi}$. For example, M.J.Box (1968) considerdd the functions given by

$$x_i = \exp(\xi_i)$$

as well as other non-polynomial functions.

The second type occurs when certain of the terms of the polynomial $\underset{\sim}{x}^T\underset{\sim}{\beta}$ cannot be estimated, and must, therefore, be omitted. For example, in the bivariate case (p=2), if $\Xi$ specified the points of a 3x5 factorial design, necessarily the polynomial elements of $\underset{\sim}{x}$ must be a subset of

$$1, \xi_1, \xi_2, \xi_1^2, \xi_2^2, \xi_1\xi_2, \xi_2^3, \xi_1^2\xi_2, \xi_1\xi_2^2, \xi_2^4, \xi_1^2\xi_2^2, \xi_1\xi_2^3, \xi_1\xi_2^4, \xi_1^2\xi_2^4$$

which omits the combinations $\xi_1^3$, $\xi_1^4$, and $\xi_1^3\xi_2$, whose coefficients cannot be estimated because an insufficient number of levels of $\xi_1$ was used. Similarly only two coefficients of degree five or higher may be estimated from this design. In practice it is unlikely that an attempt would be made to estimate the coefficients of $\xi_1\xi_2^4$ or $\xi_1^2\xi_2^4$. If it were, and if the factorial were unreplicated, an exact fit would be obtained.

Estimation

Methods, culminating in the estimates $\underset{\sim}{b}$ and $\hat{y}$, may be divided into design procedures and estimation procedures. Design procedures are those used to specify $\Xi$ and hence X. Discussion on methods of selecting the design is outside the scope of this thesis. Estimation procedures are those which, given X and $\underset{\sim}{y}$, seek to estimate $\underset{\sim}{b}$. In general, $\underset{\sim}{b}$ is assumed to be a linear combination of the observed responses, of the form

$$\underset{\sim}{b} = T\underset{\sim}{y}$$

where T depends only on X (not, for example, on $\underset{\sim}{\beta}$).

The commonest estimator arises from minimization of the sum of squares of the errors $y_u - \hat{y}_u$. This is known as the least squares estimator, and is, in fact, identical to that obtained when $\epsilon$ is assumed to have a normal distribution, and maximum likelihood estimation used.

The quantity to be minimized is

$$(\underset{\sim}{y} - X\underset{\sim}{b})^T(\underset{\sim}{y} - X\underset{\sim}{b}) \qquad (1.2)$$

Differentiation with respect to $\underset{\sim}{b}$ and equation to zero yields

$$-2X^T(\underset{\sim}{y} - X\underset{\sim}{b}) = \underset{\sim}{0}$$

from which

$$\underset{\sim}{b} = (X^TX)^{-1}X^T\underset{\sim}{y}$$

or

$$T = (X^TX)^{-1}X^T$$

provided that $X^TX$ is non-singular. If $X^TX$ is singular, a generalized inverse may be used, but is unnecessary in the present case.

This straight-forward estimator has many desirable properties. In particular,

$$E(\underset{\sim}{b}) = (X^TX)^{-1}X^TX\underset{\sim}{\beta} = \underset{\sim}{\beta} \tag{1.3}$$

$$\begin{aligned} Var(\underset{\sim}{b}) &= (X^TX)^{-1}X^T \, Var(\underset{\sim}{y}) \, X(X^TX)^{-1} \\ &= \sigma^2 \quad (X^TX)^{-1} \end{aligned} \tag{1.4}$$

from the assumptions about $\epsilon$. Hence the estimator is unbiased from (1.3). It can also be shown that (1.4) gives the minimum variance arising from an unbiased linear estimator.

Finally

$$\begin{aligned} Var(\hat{y}) &= Var(\underset{\sim}{x}^T\underset{\sim}{b}) \\ &= \sigma^2 \, \underset{\sim}{x}^T(X^TX)^{-1}\underset{\sim}{x} \end{aligned}$$

for arbitrary $\underset{\sim}{x}$, not necessarily one of the $\underset{\sim}{x}_u$. Thus the various variances can easily be derived from $(X^TX)^{-1}$.

This leads naturally to the concept of a rotatable design, which is a polynomial design for which $Var(\hat{y})$ depends only on $\sigma^2$ and $\underset{\sim}{\xi}^T\underset{\sim}{\xi}$. That is, $Var(\hat{y})$ is invariant under orthogonal rotation of the $\xi$-axes.

It should be emphasized that the estimator defined above is not the only linear estimator possible. In particular, in conditions where the specified model (1.1) is inadequate, that is, where y contains other terms than those in a linear combination of the specified $\underset{\sim}{x}$, a different estimator may assist in compensating, in some degree, for this inadequacy, at the expense of greater variance.

Hypothesis testing

From this point hypothesis testing will be considered, and the additional assumption that the $\epsilon$ are normally distributed will be required.

Now let

$$M = I - X(X^T X)^{-1} X^T$$
$$N = X(X^T X)^{-1} X^T$$

Note that both M and N are idempotent matrices, nxn, and that MN=0, MX=0. Also

$$\text{tr } N = \text{tr}\left\{X(X^T X)^{-1} X^T\right\}_{nxn}$$
$$= \text{tr}\left\{(X^T X)^{-1} X^T X\right\}_{pxp}$$

since compatible matrices commute under the trace operator. Hence

$$\text{tr } N = \text{tr } I_{pxp} = p$$
$$\text{tr } M = \text{tr } I_{nxn} - \text{tr } N = n-p$$

The residual sum of squares (1.2) is equal, on expansion, to

$$\begin{aligned} SSE &= \underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T X \underset{\sim}{b} \\ &= \underset{\sim}{y}^T \underset{\sim}{y} - \underset{\sim}{y}^T X(X^T X)^{-1} X^T \underset{\sim}{y} \qquad (1.5) \\ &= \underset{\sim}{y}^T M \underset{\sim}{y} \end{aligned}$$

It is necessary to recall a theorem on the distribution of quadratic forms (see, for example, Graybill (1961)).

Theorem: If $\underset{\sim}{y} \sim N(\underset{\sim}{\mu}, \sigma^2 I)$, then $\underset{\sim}{y}^T A \underset{\sim}{y}/\sigma^2$ is distributed as $\chi'^2(k,\lambda)$, where $\chi'^2$ represents the non-central chi-squared distribution, and $\lambda = \frac{1}{2\sigma^2} \underset{\sim}{\mu}^T A \underset{\sim}{\mu}$, if, and only if, $A$ is an idempotent matrix and tr $A = k$.

In the present situation, $\underset{\sim}{y} \sim N(X\underset{\sim}{\beta}, \sigma^2 I)$ and hence

$$\frac{SSE}{\sigma^2} = \frac{\underset{\sim}{y}^T M \underset{\sim}{y}}{\sigma^2} \sim \chi'^2(n-p,\lambda)$$

where $\lambda = \frac{1}{2\sigma^2} \underset{\sim}{\beta}^T X^T M X \underset{\sim}{\beta} = 0$. Thus $SSE/\sigma^2$ has a central $\chi^2$ distribution with n-p degrees of freedom.

The second term in (1.5) is the sum of squares accounted for by the regression, and is

$$SSR = \underset{\sim}{y}^T N \underset{\sim}{y}$$

By a process of reasoning similar to that for SSE, it is an easy matter to establish that

$$\frac{SSR}{\sigma^2} \sim \chi'^2(p,\lambda)$$

where $\lambda = \frac{1}{2\sigma^2} \underset{\sim}{\beta}^T X^T N X \underset{\sim}{\beta} = \frac{1}{2\sigma^2} \underset{\sim}{\beta}^T X^T X \underset{\sim}{\beta}$

Again from the theory of quadratic forms, a necessary and sufficient condition for $\underset{\sim}{y}^T A \underset{\sim}{y}$ and $\underset{\sim}{y}^T B \underset{\sim}{y}$ to be independent is that $AB=0$.

Hence, since $MN=0$, SSE and SSR are independent and

$$F = \frac{SSR}{p} \Big/ \frac{SSE}{n-p}$$

has a non-central F-distribution with p and n-p degrees of freedom and non-centrality parameter $\frac{1}{2\sigma^2} \underset{\sim}{\beta}^T X^T X \underset{\sim}{\beta}$. Thus F may be used to test the hypothesis that $\underset{\sim}{\beta} = \underset{\sim}{0}$.

In response surface design it is usual to further subdivide SSE by taking advantage of point replication.

As a preliminary, suppose that the model specification (1.1) is incorrect and that while the model

$$\eta = \underset{\sim}{x}_1^T \underset{\sim}{\beta}_1$$

has been assumed, the true model is

$$\eta = \underset{\sim}{x}_1^T \underset{\sim}{\beta}_1 + \underset{\sim}{x}_2^T \underset{\sim}{\beta}_2$$

Using $X_1$ and $X_2$ in an obvious way,

$$\underset{\sim}{b} = (X_1^T X_1)^{-1} X_1^T \underset{\sim}{y}$$

In these circumstances

$$E(\underset{\sim}{b}) = (X_1^T X_1)^{-1} X_1^T (X_1 \underset{\sim}{\beta}_1 + X_2 \underset{\sim}{\beta}_2)$$
$$= \underset{\sim}{\beta}_1 + (X_1^T X_1)^{-1} X_1^T X_2 \underset{\sim}{\beta}_2$$

and $\underset{\sim}{b}$ is a biased estimator of $\underset{\sim}{\beta}_1$. The matrix $A = (X_1^T X_1)^{-1} X_1^T X_2$ is known as the alias matrix (Box and Wilson (1951)) and measures the extent of the bias.

Putting

$$M_1 = I - X_1 (X_1^T X_1)^{-1} X_1^T$$

and using the same expansion as before,

$$SSE = \underset{\sim}{y}^T M_1 \underset{\sim}{y}$$

and $\frac{SSE}{\sigma^2} \sim \chi'^2(n-p, \lambda)$.

However, in the present case

$$\underset{\sim}{y} \sim N(X_1 \underset{\sim}{\beta}_1 + X_2 \underset{\sim}{\beta}_2,\ \sigma^2 I)$$

and thus

$$\lambda = \frac{1}{2\sigma^2} (\underset{\sim}{\beta}_2^T X_2^T + \underset{\sim}{\beta}_1^T X_1^T) M_1 (X_1 \underset{\sim}{\beta}_1 + X_2 \underset{\sim}{\beta}_2)$$
$$= \frac{1}{2\sigma^2} \underset{\sim}{\beta}_2^T X_2^T M_1 X_2 \underset{\sim}{\beta}_2 \neq 0$$

in general, and the F-test described above is no longer available.

Suppose, however, that point replication has been used, and that r distinct points have been included in the design, the sth of them $n_s$ times, $s=1, \dots r$, and $\Sigma_{s=1}^{r} n_s = n$. Also, let $\bar{y}_s$ be the group mean of the y-values measured at the sth distinct point.

Without loss of generality, the points may be arranged in such a way that the $n_s$ points in the sth group are together in the $\underset{\sim}{y}$ and X matrices.

Now define

$$K = I - \begin{pmatrix} \frac{1}{n_1} J_{n_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \frac{1}{n_r} J_{n_r} \end{pmatrix} = I - J$$

where $J_{n_s}$ is the $n_s x n_s$ matrix with all unit elements, so that, without point replication, $n_s=1$, and $K=0$.

Now $K^2=K$, hence K is idempotent, and tr K = n-r. Also, since $\Xi$, and hence $X_1$ and $X_2$, consist of r groups of $n_1, \dots n_r$ identical rows, $JX_1=X_1$ and $JX_2=X_2$, from which $KX_1=KX_2=0$. Hence $KM_1=K$ and $KN_1=0$.

If the y-values are standardized by $z_u=y_u-\bar{y}_s$, where $\bar{y}_s$ is the group mean containing $y_u$, then

$$\begin{aligned} \underset{\sim}{z} &= K\underset{\sim}{y} \\ &= K(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2+\underset{\sim}{\epsilon}) \\ &= K\underset{\sim}{\epsilon} \end{aligned}$$

Now SSW, the sum of squares within groups of observations at the same point, is given by

$$SSW = \underset{\sim}{z}^T\underset{\sim}{z} = \underset{\sim}{y}^T K\underset{\sim}{y} = \underset{\sim}{\epsilon}^T K\underset{\sim}{\epsilon}$$

and since $\underset{\sim}{\epsilon} \sim N(\underset{\sim}{0},\sigma^2 I)$, and tr K = n-r,

$$\frac{SSW}{\sigma^2} \sim \chi^2(n-r)$$

by the theorem quoted for SSE. Also, SSW and SSR are independent, since

$$KN_1 = KX_1(X_1^TX_1)^{-1}X_1^T = 0$$

Now consider SSF (for sum of squares due to lack of fit), defined by

$$SSF = SSE - SSW$$
$$= \underset{\sim}{y}^T(M_1-K)\underset{\sim}{y}$$

Now $$(M_1-K)^2 = M_1^2 - M_1K - KM_1 + K^2$$
$$= M_1-K$$
$$\text{tr }(M_1-K) = \text{tr } M_1 - \text{tr } K$$
$$= r - p$$

Hence, from the theorem,

$$\frac{SSF}{\sigma^2} \sim \chi'^2(r-p,\lambda)$$

where $$\lambda = \frac{1}{2\sigma^2}(\underset{\sim}{\beta}_2^TX_2^T+\underset{\sim}{\beta}_1^TX_1^T)(M_1-K)(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2)$$
$$= \frac{1}{2\sigma^2}\underset{\sim}{\beta}_2^TX_2^TM_1X_2\underset{\sim}{\beta}_2$$

This requires, reasonably enough, $r>p$.

Finally, SSR= $\underset{\sim}{y}^TN_1\underset{\sim}{y}$ where $N_1=X_1(X_1^TX_1)^{-1}X_1$, and SSR has a $\chi'^2(p,\lambda)$ distribution where

$$\lambda = \frac{1}{2\sigma^2}(\underset{\sim}{\beta}_2^TX_2^T+\underset{\sim}{\beta}_1^TX_1^T)N_1(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2)$$

$$= \frac{1}{2\sigma^2}[\underset{\sim}{\beta}_1^TX_1^TX_1\underset{\sim}{\beta}_1+2\underset{\sim}{\beta}_1^TX_1^TX_2\underset{\sim}{\beta}_2+\underset{\sim}{\beta}_2^TX_2^TX_1(X_1^TX_1)^{-1}X_1^TX_2\underset{\sim}{\beta}_2]$$

Now note that, where L is an arbitrary matrix,

$$E(\underset{\sim}{y}^TL\underset{\sim}{y}) = E(\text{tr }\underset{\sim}{y}^TL\underset{\sim}{y}) = E(\text{tr }L\underset{\sim}{y}\underset{\sim}{y}^T)$$
$$= \text{tr }[LE(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2+\underset{\sim}{\epsilon})(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2+\underset{\sim}{\epsilon})^T]$$
$$= \text{tr }[L(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2)(X_1\underset{\sim}{\beta}_1+X_2\underset{\sim}{\beta}_2)^T+\sigma^2L]$$
$$= \text{tr }\underset{\sim}{\beta}^TX^TLX\underset{\sim}{\beta} + \sigma^2\text{tr }L$$

where

$$\underset{\sim}{\beta} = \begin{pmatrix} \underset{\sim}{\beta}_1 \\ \underset{\sim}{\beta}_2 \end{pmatrix} \qquad X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

From this, the expected values of the various sums of squares are readily derived. The analysis is given in table 1.

Table 1

Basic response surface ANOV

| Source | Sum of squares | DF | E(MS) |
|---|---|---|---|
| Regression | $SSR = \underset{\sim}{y}^T X_1 \underset{\sim}{b}$ | $p$ | $\sigma^2 + \frac{1}{p}\underset{\sim}{\beta}^T X^T M_1 X \underset{\sim}{\beta}$ |
| Lack of fit | SSF by subtraction | $r-p$ | $\sigma^2 + \frac{1}{r-p}\underset{\sim}{\beta}_2^T X_2^T M_1 X_2 \underset{\sim}{\beta}_2$ |
| Error within replicated points | $SSW = \underset{\sim}{y}^T K \underset{\sim}{y}$ | $n-r$ | $\sigma^2$ |
| Total | $SST = \underset{\sim}{y}^T \underset{\sim}{y}$ | $n$ | |

Note that without replication $n=r$, and if $\underset{\sim}{\beta}_2=0$, this table reduces to the simpler form derived earlier.

While the above argument establishes the theoretical justification for the use of the F-tests, the test of the whole regression is, in practice, of little use. However, it is perfectly general, and not dependent on a polynomial specification of $\underset{\sim}{x}$. In the event that a polynomial is used, the SSR is ordinarily broken down into the classification shown in table 2.

Table 2

Conventional ANOV for regression coefficients in polynomial model

| Source | DF |
|---|---|
| Mean, $\beta_0$ | 1 |
| Linear terms | p |
| Second order terms | $\frac{1}{2}p(p+1)$ |
| Third order terms | $\frac{1}{6}p(p+1)(p+2)$ |
| ... | |
| dth order terms | $\sum_{i=1}^{p} i^{d-1}$ |

While this is suitable for establishing the true degree of the polynomial, it is inadequate for establishing the importance, in the final response, of a particular $\xi$. Section 3 of this thesis considers the structure of $X^T X$, for the second order polynomial model, in some detail, in order to facilitate tests aimed at establishing the importance of particular elements of $\underset{\sim}{\xi}$.

## Further topics

In field experiments, each experiment usually consists of a plot of ground. In most circumstances, the number of such plots which can be assumed to represent essentially the same external conditions is quite limited.

In order to control this type of environmental variation, a block structure may be superimposed on the response surface design, yielding a model of the form

$$\eta = \alpha_w + \underset{\sim}{x}^T \underset{\sim}{\beta} \qquad (1.6)$$

where $\alpha_w$ is the block effect associated with the wth block, with $\sum_w \alpha_w = 0$.

Designs including such block structures were introduced by DeBaun (1956) and elaborated by Box and Hunter (1957) in the case of rotatable designs. These designs allow adequate control of environmental variation.

A natural extension of this type of design is to consider the possibility of superimposing a further treatment effect, which, in practice, could represent something like a species effect. The model would be

$$\eta = \alpha_w + \tau_v + \underset{\sim}{x}^T \underset{\sim}{\beta}$$

where now $\tau_v$ is the vth treatment effect. As far as treatments are concerned, such a model is identical to the analysis of covariance model, which uses the regression variables $\underset{\sim}{x}$ to reduce variation in the response, major interest being focussed on the superimposed treatment effects. A response surface approach would have equal interest in both parts of the fitted model.

Pursuing this line of enquiry further, section 4 of this thesis considers the implications of combining various classification designs with a response surface design.

One obvious extension of the model described by (1.6) is to allow $\beta$ to vary with the block, giving a model of the form

$$\eta = \alpha_w + x^T \beta_w$$

In many applications the question of the degree of correspondence between the individual regressions $\beta_w$ and the overall regression $\beta$ is of considerable importance. Section 4 also considers, briefly, this aspect of response surface methodology.