CSCanada
Progress in Applied Mathematics
Vol. 4, No. 2, 2012, pp. [110–126]
DOI: 10.3968/j.pam.1925252820120402.1517

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net
www.cscanada.org

Comparison of Optimality Criteria of Reduced Models for Response Surface Designs with Restricted Randomization

Angela U. Chukwu[a] and Yisa Yakubu[b],*

[a] Department of Statistics, University of Ibadan, Oyo State, Nigeria.
[b] Department of Mathematics and Statistics, Federal University of Technology, Minna, Niger State, Nigeria.

* Corresponding author.
Address: Department of Mathematics and Statistics, Federal University of Technology, Minna, Niger State, Nigeria; E-Mail: yisa_yakubu@yahoo.com

Received: March 12, 2012/ Accepted: July 2, 2012/ Published: October 31, 2012

Abstract: In this work, $D-$, $G-$, and $A-$ efficiencies and the scaled average prediction variance, IV criterion, are computed and compared for second-order split-plot central composite design. These design optimality criteria are evaluated across the set of reduced split-plot central composite design models for three design variables under various ratios of the variance components (or degrees of correlation d). It was observed that D, A, G, and IV for these models strongly depend on the values of d; they are robust to changes in the interaction terms and vary dramatically with the number of, and changes in the squared terms.

Key words: Response surface methodology; Split-plot central composite design; Reduced models; Design optimality criteria; Efficiency; Robustness

Chukwu, A.U., & Yakubu, Y. (2012). Comparison of Optimality Criteria of Reduced Models for Response Surface Designs with Restricted Randomization. *Progress in Applied Mathematics, 4*(2), 110–126. Available from http://www.cscanada.net/index.php/pam/article/view/j.pam.1925252820120402.1517 DOI: 10.3968/j.pam.1925252820120402.1517

1. INTRODUCTION

Experiments are performed by researchers in virtually all fields of inquiry so as to study and model the potential effects of several design variables on the responses of interest. The foundation for response surface methodology (RSM) was laid by Box

and Wilson (1951). Response surface methodology comprises a group of statistical and mathematical techniques for empirical model building and model exploitation. By careful design and analysis of experiments, it seeks to relate a *response*, or *output* variable, to the levels of a number of *predictors*, or *input* variables, that affect it. The form of such a relationship is usually unknown, but can be approximated by a low-order polynomial such as the second-order response surface model

$$y_u = \beta_0 + \sum_{i=1}^{k} \beta_i x_{iu} + \sum_{i=1}^{k} \beta_{ii} x_{iu}^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_{iu} x_{ju} + e_u \tag{1.1}$$

Where y is the measured response, the β's are parameter coefficients, x_i's are the input variables and e is an error term. Popular designs that utilize this model include—the central composite designs (CCD), introduced by Box and Wilson (1951), and Box-Behnken design (BBD), introduced by Box and Behnken (1960).

The most extensive applications of RSM are in the industrial world, particularly in situations where potential influence of several process variables on some quality characteristic of the process is being investigated. RSM is sequential in nature and this allows the experimenter to learn about the process or system under study as the investigation proceeds. See, for instance, Khuri and Cornell (1996), Box and Draper (2007), and Myers *et al.* (2009) for more details.

After data are generated from the experiment and a model is fit, many parameters in the fitted model are deemed insignificant. Therefore, a reduced model retaining only the significant terms is adopted for use. Design optimality criteria based on the adopted reduced model are equally if not more important than the optimality criteria for the proposed full model (Borkowski & Valereso, 2001). Therefore, a design should be *robust* over classes of reduced models; that is, the design should maintain high optimality criteria over a wide assortment of potential models.

Many authors (e.g., Box & Draper, 1959, 1963; Karson, Manson & Hader, 1969) have studied the design-selection problem when the proposed approximating model is an *underparameterized* approximation of the true response surface. In such cases, use is made of a low-order polynomial when a higher-order polynomial is a better approximating function. With regard to this design problem, some authors (e.g., Box & Draper, 1987; Myers & Montgomery, 1995; and, Khuri & Cornell, 1996) also have used the integrated mean squared error (IMSE), where $IMSE = V + B$, and

$$B = \frac{N\Omega}{\sigma^2} \int_R [E(\hat{y}(x) - \eta(x)]^2 dx \tag{1.2}$$

is the systematic (squared) bias resulting from underestimation of the true response surface with the fitted low-order model;

$$V = \frac{N\Omega}{\sigma^2} \int_R Var[(\hat{y}(x))] dx \tag{1.3}$$

is the prediction variance, and $\Omega^{-1} = \int_R dx$.

The research by Borkowski and Elsie (2001) addresses the problem in a different dimension. These authors provide an evaluation of the robustness properties of some standard response surface designs (CCD, SCD, NHD, and computer- generated algorithmic designs) over a collection of reduced models based on $D-$, $G-$, $A-$, and

$IV-$ optimality criteria. These reduced models are formed by removing terms one after the other from the proposed full model. They showed that design optimality criteria can be sensitive to deviations from the full second-order response surface model, and that the CCD is robust with respect to the set of reduced models as well as across the four optimality criteria considered.

Angela and Yisa (2012) studied the role of several model characteristics and center point replications on the properties of $A-$, $D-$, $G-$, and $IV-$optimal designs for a full second-order model, and also investigate the impacts of these same designs under various numbers of experimental runs. They showed that the pure linear model is the best model in terms of quality of estimation and prediction, and also that the $A-$ and $D-$ efficiencies for this model is insensitive to changes in number of experimental runs, as compared to other models under consideration.

The similarity with all of the above efforts is the assumed possibility of complete randomization of the experimental run order. That is, all the experiments considered are (assumed to be) conducted in a completely randomized (CRD) mode. This means that all factors are independently reset with each run (see, for instance, Ganju & Lucas 1997, 1998, 2005).

However, in most industrial experiments it becomes a challenge to adhere to the CRD principles at all times. This is due to the fact that a large fraction of these experiments contain two sets of factors: the hard-to-change (HTC) factors (such as pressure, humidity, process temperature, mechanical set-ups, etc.), which are not reset from run to run because of the time or cost involved in doing so, and the easy-to-change (ETC) factors (like variety, processing time, etc.), whose levels are easy to reset from run to run.

Performing CRD in the presence of HTC factors adds considerable time and expense to the experiment. When hard-to-change factors exist in an experiment, it is typically more cost-effective to randomize the treatment combinations in such a manner as to minimize the number of times the levels of these (HTC) factors are changed. In this sense, the experimenter typically will fix the level of the HTC factor (i.e., restrict the randomization) and then run all combinations or a fraction of all combinations of the ETC factor levels. This strategy leads to a split-plot design where the experimental unit for the hard-to-change factors is subdivided into experimental units for the easy-to-change factors.

For every split-plot experiment, there are two separate randomizations. The hard-to-change factor level combinations, often called whole-plot treatments, are randomly assigned to whole plots based on the whole plot design. Within each whole plot, the easy-to-change factor level combinations, often called subplot treatments, are randomly assigned to subplots with a separate randomization for each whole plot. This leads to two error terms, one for the whole-plot treatments and one for the subplot treatments.

Most industrial experiments follow the split-plot structure but in many cases are incorrectly analyzed as completely randomized experiments (Jones & Christopher, 2009). In fact, recent work, most notably by Lucas and his coworkers (Ambari & Lucas, 1994; Ganju & Lucas, 1997, 1999, 2005; Ju & Lucas, 2002). Web *et al.* (2004) has demonstrated that many experiments previously thought to be completely randomized experiments also exhibit split-plot structure.

The issue of randomization restriction in response surface methods (RSM) was first investigated by Letsinger *et al.* (1996). Based on simulation study, Letsinger *et al.* (1996) recommended restricted maximum likelihood method (REML) for pa-

rameter estimation as compared to ordinary least squares (OLS) method. Vining *et al.* (2005) modified the standard central composite design (CCD) to accommodate split-plot structure and developed equivalent Cestimation split-plot designs for response surfaces. Kowalski *et al.* (2006) modified a central composite design to model the process mean and variance when there are hard Cto-change factors.

1.1. Model and Notations

We define the general form of the split-plot model as

$$\mathbf{Y} = \mathbf{X}\beta + Z\gamma + \varepsilon, \tag{1.1.1}$$

where $\mathbf{Y}$ is the $n \times l$ vector of responses, $\mathbf{X}$ is the $n \times p$ model matrix, β is the $p \times 1$ vector of coefficients, Z is the $n \times b$ incidence matrix assigning the n observations to the b whole plots: the $(i,j)^{th}$ entry of Z is equal to one if the j^{th} observation belongs to the i^{th} whole plot, and zero otherwise. γ is the $b \times 1$ vector of random whole-plot effects, ε is the $n \times 1$ vector of random subplot effects. It is assumed that γ and ε each has a mean of 0 and variance σ_δ^2 and σ_ε^2 respectively, where σ_ε^2 is the subplot error variance, and σ_δ^2 is the whole-plot error variance. It is assumed that

$$\begin{aligned} &E(\varepsilon) = 0 \ \text{ and } \ Cov(\varepsilon) = \sigma_\varepsilon^2 I_n, \\ &E(\gamma) = 0 \ \text{ and } \ Cov(\gamma) = \sigma_\gamma^2 I_b, \\ &Cov(\gamma, \varepsilon) = 0. \end{aligned}$$

Under these assumptions, the variance-covariance matrix of the observations $\mathrm{Var}(y)$ can be written as

$$V = \sigma_\varepsilon^2 I + \sigma_\delta^2 ZZ' = \sigma_\varepsilon^2 \left[I + dZZ'\right] \tag{1.1.2}$$

Where $d = \dfrac{\sigma_\delta^2}{\sigma_\varepsilon^2}$ is a measure for the extent to which observations within the same whole plot are correlated, and ZZ' is a block diagonal matrix of the form

$$ZZ' = \begin{bmatrix} 1_{n_1} 1'_{n_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1_{n_w} 1'_{n_w} \end{bmatrix} \tag{1.1.3}$$

where 1_{ni} is a vector of ones of length n_i and n_i is the number of subplot runs in the i^{th} whole-plot. We can observe from the structure of ZZ' that observations in different whole-plots are independent, while observations within a whole-plot are correlated.

Expanding equation (1.2.1) illustrates the terms in the model matrix for second-order designs as

$$\begin{aligned} E(y) = \beta_0 + \sum_{i=1}^{w} \beta_i z_i + \sum_{i=1}^{w-1} \sum_{j=i+1}^{w} \beta_{ij} z_i z_j + \sum_{i=1}^{w} \beta_{ii} z_i^2 + \sum_{i=1}^{k} \theta_i x_i \\ + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \theta_{ij} x_i x_j + \sum_{i=1}^{w} \sum_{j=1}^{k} \gamma_{ij} z_i x_j + \sum_{i=1}^{k} \theta_{ii} x_i^2 \end{aligned} \tag{1.1.4}$$

From (1.1.4), z is the whole-plot factor, x is the subplot factor, the β's are the regression coefficients at the whole-plot level, θ's and γ's are the regression coefficients at the subplot levels.

The generalized least squares (GLS) estimate of the model coefficients is

$$\hat{\beta}_{GLS} = (\mathbf{X}'V^{-1}\mathbf{X})^{-1}\mathbf{X}'V^{-\mathbf{1}}\mathbf{y} \tag{1.1.5}$$

with the variance-covariance matrix as

$$\text{var}(\hat{\beta}_{GLS}) = (\mathbf{X}'V^{-1}\mathbf{X})^{-1} \tag{1.1.6}$$

Equation (1.1.5) is also the maximum likelihood estimate of β under normality.

Goos and Vandebroek (2004) investigated the role of several model characteristics on the properties of $D-$optimal split-plot designs. They considered three models (pure linear model, linear model with interaction, and a full quadratic model) and computed the $D-$efficiency gains for each of these models in the presence of one whole-plot and two whole-plot factors separately. The authors observed that the gains in efficiency in the presence of two whole-plot factors are smaller due to the more whole-plot factor levels produced by this design, which reduces the possibility to group observations and to benefit from the correlation. This means that the higher the correlation, the larger the efficiency gains of using a split-plot design instead of a CRD.

The current work investigates the impact of split-plot structure on the robustness properties of three-factor split-plot central composite designs consisting of one whole-plot factor and two subplot factors over a collection of reduced models based on $D-$, $A-$, $G-$, and $IV-$optimality criteria, and for three different degrees of correlation ($d = 0.5,\ 1,\ 2$).

2. DESIGN OPTIMALITY CRITERIA FOR REDUCED SPLIT-PLOT CCD MODELS

The desire of every experimenter is to minimize the number of experimental trials while still being able to estimate adequately the underlying model. After considering practical constraints (e.g., design size, time and money), *design optimality criteria* are often used to evaluate a proposed experimental design prior to running it. If several alternative designs are proposed, their optimality properties can be compared to aid in the choice of design. Design optimality criteria are primarily concerned with "optimal properties" of the information matrix for the model matrix X.

The four commonly-used optimality criteria are $A-$, $D-$, $G-$, and $IV-$ optimality criteria. Let

$$M(\xi) = (X'V^{-1}X) \tag{2.1}$$

be the information matrix for a split-plot design, then

$D-$criterion goal $\rightarrow$ maximize $|M(\xi)|$, or equivalently, minimize $|M^{-1}(\xi)|$,
$A-$criterion goal $\rightarrow$ minimize trace $[M^{-1}(\xi)]$,
$G-$criterion goal $\rightarrow$ minimize maxx?R $[Nf'(z,x)(M^{-1}(\xi))f(z,x)]$,
$IV-$criterion goal $\rightarrow$ minimize average $[Nf'(z,x)(M^{-1}(\xi))f(z,x)]$ over $x \in R$,

(2.2)

Where X is the model matrix of a split-plot CCD associated with the response surface model given in (1.1.4) above, x is any point in the design region R, N is the design size and $f(z,x) = [f_1(z,x), ..., f_p(z,x)]$ is a vector of p real- valued functions

based on the p model terms. A and D criteria examine the design's estimation quality while G and IV criteria are based on the scaled prediction variance $V(z,x)$, which is a function of the variance for the above fitted response model (1.2.4).

For split-plot response surface designs, the predicted value of the mean response at any point $(z,x)' = (z_1, z_2, ..., z_w,\ x_1, x_2, ..., x_s)'$ in the experimental region R, is

$$\hat{Y}_{(z,x)} = (z,x)'\hat{\beta} \tag{2.3}$$

where the vector $(z,x)'$ reflects a location in the design space at which an experimenter wants to predict the response and also reflects the nature of the model; $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$ is the generalized least squares estimator of β, with X as the model matrix and y as the vector of observations.

Then we define the prediction variance at this point, $(z,x)'$, as

$$\mathrm{Var}(\hat{Y}_{(z,x)}) = f(z,x)'(X'V^{-1}X)^{-1}f(z,x) \tag{2.4}$$

Here, $f(z,x)$ is the general form of the $1 \times p$ model vector for second-order split-plot response surface designs and is given as

$$\begin{aligned} f(z,x)' = [1|z_1, ..., z_w|x_1, ..., x_s|z_1z_2, ..., z_{w-1}z_w|z_1x_1, ..., z_wx_s| \\ x_1x_2, ..., x_{s-1}x_s|z_1^2, ..., z_w^2|x_1^2, ..., x_s^2|] \end{aligned} \tag{2.5}$$

where z and x are the whole plot and subplot factors respectively. It is important to note that the value of the prediction variance as given by the function (2.4) above depends on the location at which one is predicting.

Now, we denote the "Hat" matrix of the split-plot response surface design by

$$H = X(X'V^{-1}X)^{-1}X'V^{-1}. \tag{2.6}$$

By post multiplying both sides of this matrix by the error matrix V of the entire data observations, we have

$$HV = X(X'V^{-1}X)^{-1}X' \tag{2.7}$$

Then our computations have shown that the prediction variance at a point $x_j = (z,x)_j$, in the design space as given in (2.4) above, is

$$\mathrm{Var}\left(\hat{Y}_{(x_j)}\right) = x_j'(X'V^{-1}X)^{-1}x_j = (hv)_{jj} \tag{2.8}$$

where $(hv)_{jj}$ is the corresponding j^{th} diagonal element of the matrix (2.7).

Now, in split-plot response surface designs, the variance of an individual observation is the sum of the subplot and whole plot error variances, $\sigma_\gamma^2 + \sigma_e^2$.

By scaling Equation (2.4) by $\sigma_\gamma^2 + \sigma_e^2/N$ (i.e., observation error variance divided by the design size), we obtain the scaled prediction variance (SPV) function for split-plot response surface designs as

$$v_{(z,x)} = \frac{\mathrm{Var}\left(\hat{Y}_{(z,x)}\right)}{\sigma_\gamma^2 + \sigma_e^2/N} = Nf(z,x)'(X'R^{-1}X)^{-1}f(z,x) \tag{2.9}$$

where $R = \dfrac{V}{\sigma_\gamma^2 + \sigma_e^2}$ is the correlation matrix of the responses in a split-plot design. Where the structure of the R matrix is

$$R = \begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & R_a \end{bmatrix} \quad \text{where } R_j = \begin{bmatrix} 1 & \frac{d}{1+d} & \cdots & \frac{d}{1+d} \\ \frac{d}{1+d} & \ddots & \cdots & \vdots \\ \vdots & & \ddots & \frac{d}{1+d} \\ \frac{d}{1+d} & \cdots & \frac{d}{1+d} & 1 \end{bmatrix} \tag{2.10}$$

where $d = \dfrac{\sigma_\gamma^2}{\sigma_e^2}$, and R_j denotes the correlation matrix of responses for the j^{th} whole plot.

The scaling by the observation error variance, $(\sigma_\gamma^2 + \sigma_e^2)$, causes the resulting information matrix to be without units. A unitless information matrix,

$$M = (X'R^{-1}X) \tag{2.11}$$

is desirable especially for design comparison purposes. While scaling by $1/N$ causes the information matrix M to be reflective of the design size. Now, the 'hat' matrix in terms of this observational correlation matrix R is

$$H = X(X'R^{-1}X)^{-1}X'R^{-1} \tag{2.12}$$

Post multiplying both sides of this matrix by R, we obtain

$$HR = X(X'R^{-1}X)^{-1}X' \tag{2.13}$$

Then, our computations have shown that the scaled prediction variance at a point $x_j = (z, x)_j$ in the design space is

$$v_{x_j} = Nx_j'(X'R^{-1}X)^{-1}x_j = N(hr)_{jj} \tag{2.14}$$

where $(hr)_{jj}$ is the j_{th} corresponding diagonal element of the product (HR) for the given split-plot response surface design. The prediction variance v_{x_j} will be equal for all observations in the same subset. The scaled prediction variance (SPV), as given by (2.14) above, allows the practitioner to measure the variance of the predicted response on a per observation basis and it penalizes larger designs over small designs. Scaling facilitates comparisons among designs of various sizes. On the other hand, the unscaled prediction variance (UPV) in equation (2.4) is useful to compare designs of different sizes so as to determine if the additional runs in a larger design are of value in substantially reducing the variance of the predicted response. The UPV also allow for an estimate of the quality of prediction in absolute terms.

Directly associated with the prediction variance $v_{(z,x)}$ are the G and IV optimality criteria. The $G-$ criterion seeks to minimize the maximum prediction variance over the experimental region R. That is, a $G-$optimal design(ζ) is one which min $[\max_{x \in R} v(z, x)]$. $G-$optimal designs give the researcher an upper bound for the prediction variance for a proposed design.

The integrated prediction variance (IV) criterion gives a single average measure of the design's prediction capability throughout the experimental region of interest. This is obtained by integrating the prediction variance function $v_{(z,x)}$ over the region R. For a split-plot central composite design, the standardized $IV-$ criterion can be expressed as

$$IV = \left(\min_{\zeta} \frac{1}{K} \int_R v(z,x) dz dx \right) = \left(\min_{\zeta} \frac{1}{K} \int_R f(z,x)'(X'R^{-1}X)^{-1} f(z,x) dz dx \right) \tag{2.15}$$

where K is the volume of the experimental region R given as $K = \int_R dz dx$.

If the experimenter's interest is in finding a design with precise estimates of the mean, the $G-$ and $IV-$ efficiency of the design are popular choices.

For each of the designs considered, robustness was quantified by calculating $D-$, $A-$, $G-$, and $IV-$optimality measures over reduced models of the second-order model in (1.1.4) above. These measures are

$$\begin{aligned} D- \text{ efficiency} &= 100 \frac{|M(\xi)|^{1/p}}{N}, \\ A- \text{ efficiency} &= 100 \frac{p}{trace[N(M^{-1}(\xi))]}, \\ G- \text{ efficiency} &= 100 \frac{p}{N\hat{\sigma}^2_{max}}, \\ IV- \text{ efficiency} &= N\sigma^2_{ave}, \end{aligned} \tag{2.16}$$

where N is the design size, p is the number of model parameters, σ^2_{ave} is the average of $f'(x)(M^{-1}(\xi))f(x)$ over the design region, and σ^2_{max} is the maximum of $f'(x)(M^{-1}(\xi))f(x)$ approximated over the set of points from a $5k$ factorial designs (with factor levels 0, ± 1.73205, ± 1), and $M(\xi)$ is as defined in (2.1).

These design optimality measures are used to compare designs across the set of reduced models. For the designs considered in this article, the optimality criteria values were computed using *Maple13* package.

3. MATERIALS AND METHODS

Sixty-four (64) split-plot central composite design models were considered in this work for $k =$ three, one whole plot factor and two subplot factors (i.e., $w = 1$, $s = 2$) under various degrees of correlation d. These comprise one full quadratic model and sixty-three (63) reduced models each consisting of pure linear part in combination with second-order part. Adopting the format of Borkowski and Valeroso (2001) for standard CCD, the 64 split-plot design models considered are given in Table 1 below. From this table, p is the number of model parameters; l, q, and c represent, respectively, number of linear terms, quadratic terms and interaction terms appearing in the model. The 1s and 0s in the L, Q and C columns indicate, respectively, the presence or absence of that term in the reduced model.

For each of the designs considered, robustness was quantified by calculating optimality criteria measures over reduced models of the second-order model in (1.1.4) above for three different degrees of correlation (i.e., $d = 0.5$, 1, 2).

Table 1
Reduced Models ($k = 3,\ w = 1,\ s = 2$)

Design	P	L	Q	C	$(l,\ q,\ c)$
1	10	(1,1,1)	(1,1,1)	(1,1,1)	(3,3,3)
2	9	(1,1,1)	(1,1,1)	(1,1,0)	(3,3,2)
3	9	(1,1,1)	(1,1,1)	(1,0,1)	(3,3,2)
4	9	(1,1,1)	(1,1,1)	(0,1,1)	(3,3,2)
5	9	(1,1,1)	(1,1,0)	(1,1,1)	(3,2,3)
6	9	(1,1,1)	(1,0,1)	(1,1,1)	(3,2,3)
7	9	(1,1,1)	(0,1,1)	(1,1,1)	(3,2,3)
8	8	(1,1,1)	(1,1,1)	(1,0,0)	(3,3,1)
9	8	(1,1,1)	(1,1,1)	(0,1,0)	(3,3,1)
10	8	(1,1,1)	(1,1,1)	(0,0,1)	(3,3,1)
11	8	(1,1,1)	(1,1,0)	(1,1,0)	(3,2,2)
12	8	(1,1,1)	(1,1,0)	(1,0,1)	(3,2,2)
13	8	(1,1,1)	(1,1,0)	(0,1,1)	(3,2,2)
14	8	(1,1,1)	(1,0,1)	(1,1,0)	(3,2,2)
15	8	(1,1,1)	(1,0,1)	(1,0,1)	(3,2,2)
16	8	(1,1,1)	(1,0,1)	(0,1,1)	(3,2,2)
17	8	(1,1,1)	(0,1,1)	(1,1,0)	(3,2,2)
18	8	(1,1,1)	(0,1,1)	(1,0,1)	(3,2,2)
19	8	(1,1,1)	(0,1,1)	(0,1,1)	(3,2,2)
20	8	(1,1,1)	(1,0,0)	(1,1,1)	(3,1,3)
21	8	(1,1,1)	(0,1,0)	(1,1,1)	(3,1,3)
22	8	(1,1,1)	(0,0,1)	(1,1,1)	(3,1,3)
23	7	(1,1,1)	(1,1,1)	(0,0,0)	(3,3,0)
24	7	(1,1,1)	(1,1,0)	(1,0,0)	(3,2,1)
25	7	(1,1,1)	(1,1,0)	(0,1,0)	(3,2,1)
26	7	(1,1,1)	(1,1,0)	(0,0,1)	(3,2,1)
27	7	(1,1,1)	(1,0,1)	(1,0,0)	(3,2,1)
28	7	(1,1,1)	(1,0,1)	(0,1,0)	(3,2,1)
29	7	(1,1,1)	(1,0,1)	(0,0,1)	(3,2,1)
30	7	(1,1,1)	(0,1,1)	(1,0,0)	(3,2,1)
31	7	(1,1,1)	(0,1,1)	(0,1,0)	(3,2,1)
32	7	(1,1,1)	(0,1,1)	(0,0,1)	(3,2,1)
33	7	(1,1,1)	(1,0,0)	(1,1,0)	(3,1,2)
34	7	(1,1,1)	(1,0,0)	(1,0,1)	(3,1,2)
35	7	(1,1,1)	(1,0,0)	(0,1,1)	(3,1,2)
36	7	(1,1,1)	(0,1,0)	(1,1,0)	(3,1,2)
37	7	(1,1,1)	(0,1,0)	(1,0,1)	(3,1,2)
38	7	(1,1,1)	(0,1,0)	(0,1,1)	(3,1,2)
39	7	(1,1,1)	(0,0,1)	(1,1,0)	(3,1,2)
40	7	(1,1,1)	(0,0,1)	(1,0,1)	(3,1,2)
41	7	(1,1,1)	(0,0,1)	(0,1,1)	(3,1,2)
42	7	(1,1,1)	(0,0,0)	(1,1,1)	(3,0,3)

To be continued

Continued

Design	P	L	Q	C	$(l,\ q,\ c)$
43	6	(1,1,1)	(1,1,0)	(0,0,0)	(3,2,0)
44	6	(1,1,1)	(1,0,1)	(0,0,0)	(3,2,0)
45	6	(1,1,1)	(0,1,1)	(0,0,0)	(3,2,0)
46	6	(1,1,1)	(1,0,0)	(1,0,0)	(3,1,1)
47	6	(1,1,1)	(1,0,0)	(0,1,0)	(3,1,1)
48	6	(1,1,1)	(1,0,0)	(0,0,0)	(3,1,1)
49	6	(1,1,1)	(0,1,0)	(1,0,0)	(3,1,1)
50	6	(1,1,1)	(0,1,0)	(0,1,0)	(3,1,1)
51	6	(1,1,1)	(0,1,0)	(0,0,1)	(3,1,1)
52	6	(1,1,1)	(0,0,1)	(1,0,0)	(3,1,1)
53	6	(1,1,1)	(0,0,1)	(0,1,0)	(3,1,1)
54	6	(1,1,1)	(0,0,1)	(0,0,1)	(3,1,1)
55	6	(1,1,1)	(0,0,0)	(1,1,0)	(3,0,2)
56	6	(1,1,1)	(0,0,0)	(1,0,1)	(3,0,2)
57	6	(1,1,1)	(0,0,0)	(0,1,1)	(3,0,2)
58	5	(1,1,1)	(1,0,0)	(0,0,0)	(3,1,0)
59	5	(1,1,1)	(0,1,0)	(0,0,0)	(3,1,0)
60	5	(1,1,1)	(0,0,1)	(0,0,0)	(3,1,0)
61	5	(1,1,1)	(0,0,0)	(1,0,0)	(3,0,1)
62	5	(1,1,1)	(0,0,0)	(0,1,0)	(3,0,1)
63	5	(1,1,1)	(0,0,0)	(0,0,1)	(3,0,1)
64	4	(1,1,1)	(0,0,0)	(0,0,0)	(3,0,0)

Note: $L = (z_1, x_1, x_2)$, $Q = (z_1^2, x_1^2, x_2^2)$, $C = (z_1x_1, z_1x_2, x_1x_2)$.

4. DESIGN COMPARISONS FOR D, A, G, AND IV

In this section, D, A, G, and IV are compared for sets of reduced models for various degrees of correlation (i.e., $d = 0.5, 1, 2$). It is important to note that large $D-$, $A-$, and $G-$efficiency and small $IV-$criterion measures are desirable. Figure 1, (a), (b), (c), and (d); Figure 2, (a), (b), (c), and (d), and Figure 3, (a), (b), (c), and (d) shows, respectively, plots of D, A, G, and IV for the three-factor 24 point, one whole plot and two subplot (four center-point) split-plot CCD against the number of model parameters, for $d = 0.5$, $d = 1$, and $d = 2$. The plotting symbol is q, the number of x_i^2 terms in the reduced model.

The following patterns exist and the first five are common for the three scenarios (i.e., for $d = 0.5$, $d = 1$, and $d = 2$).

1. For $p-$parameter models containing any, or a combination of any of the pure quadratic terms (z_i^2 and x_i^2), D, A, G, and IV are invariant to removal of any of the interaction terms (z_ix_i and x_ix_j). That is, D, A, G, and IV are robust to changes in the interaction terms.

2. For $p-$parameter models containing combination of pure whole plot and subplot quadratics, D, A, G, and IV vary dramatically with removal of pure whole plot quadratic term, while the effect on the optimality measures due to removal of a pure subplot quadratics is constant for models containing the same whole plot quadratic term. That is, D, A, G, and IV are robust to changes in the pure subplot quadratic terms.

3. For $p-$parameter models containing only the pure subplot quadratic terms, D, A, G, and IV are invariant to changes in these terms.

4. D, A, G, and IV are unaffected by the presence or absence of cross-product terms involving the variable whose squared term is removed from the model.

5. D, A, G, and IV are less variable for all models having the same number, q, of squared terms.

6. For our first scenario ($d = 0.5$), removing a squared term from a model has varying effects on each of these optimality measures as can be seen directly from Figure (1) a-d. However, for D, distinct groups are formed by the values of q with D increasing dramatically as q increases, while A increases as q decreases. The values of q also formed diagonal D bands as shown in Figure 1 (a). For this scenario, D for the full model is the smallest, A for the pure linear model is the largest, and IV for the full model is the largest (note that smaller IV is better), relative to the set of other reduced models. There is a large decrease in D when a square term is removed from a model; a large increase in A is also observed when a squared term is removed from a model. Here, the pure linear model is the most $A-$, and $IV-$efficient.

7. For our second scenario ($d = 1$), removing a squared term from a model has varying effects on each of these optimality measures except A, where distinct groups are formed by the values of q. The values of q formed diagonal D and G bands as can be seen directly from Figure 2 (a) and (c) below. There is a large increase in A when a squared term is removed, as can be seen directly from Figure 2 (b). Also, the values of q formed diagonal G bands as shown in Figure 2 (c). For this scenario, the pure linear model is the most $A-$ and $IV-$efficient while the full model is the most $G-$efficient and also the worst $IV-$efficient.

8. For our third scenario ($d = 2$), removing a squared term from a model has varying effects on each of these optimality measures as can be seen directly from Figure 3 (a)-(d). However, distinct groups are formed for D by the values of q while these values of q formed diagonal G bands as shown in Figure 3 (c). There is a large change in each of these optimality measures when a squared term is removed from a model. Here, the pure linear model is the most $IV-$efficient.

9. For all the three scenarios, IV for the full model tends to be larger relative to the set of reduced models. While that of the pure linear model ($p = 4$) is smaller relative to the set of other models and is therefore the most preferred.

Next we compare D, A, G, and IV for sets of reduced models for the three scenarios combined. Figure 4 (a), (b), (c), and (d) shows plots of D, A, G, and IV for the three-factor 24 point, one whole plot and two subplot (four center-point) split-plot CCD against the number of model parameters, for the combined three scenarios (i.e., $d = 0.5$, 1, 2). From these plots, we observe that

1. D, A, G, and IV measures for these designs depend strongly on the ratio of the two variance components (i.e., on the degree of correlation d).

2. The three-factor 24 point, one whole plot and two subplot split-plot CCD with four replicates of center point and equal whole plot and subplot error variances (i.e., $d = 1$) is the most $D-$ and $A-$efficient.

3. The pure linear model is the most $A-$efficient model, while the full model is the worst $IV-$efficient model.

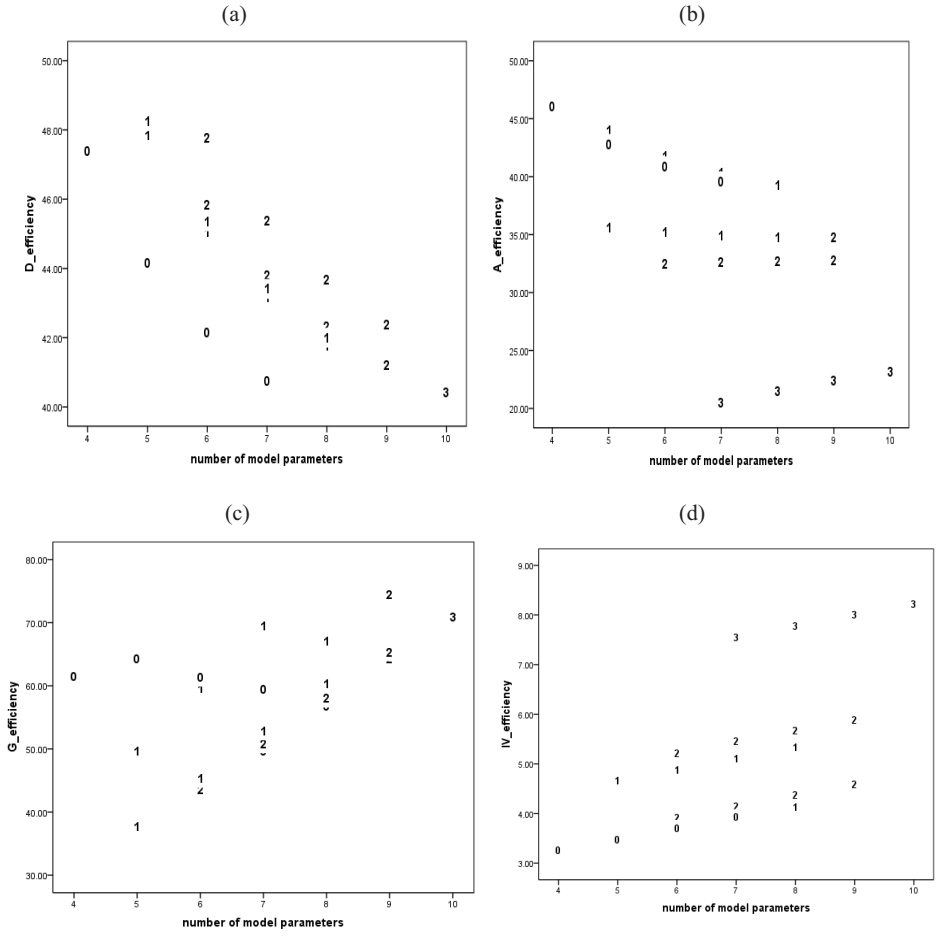

Figure 1
Plots of Reduced Model Efficiencies for the Three-Factor 24 Points Split-Plot CCD. The CCD Consists of 4 Center-Points. Plots (a), (b), (c), and (d) Contain the D, A, G, and IV Efficiencies for the CCD When $d = 0.5$

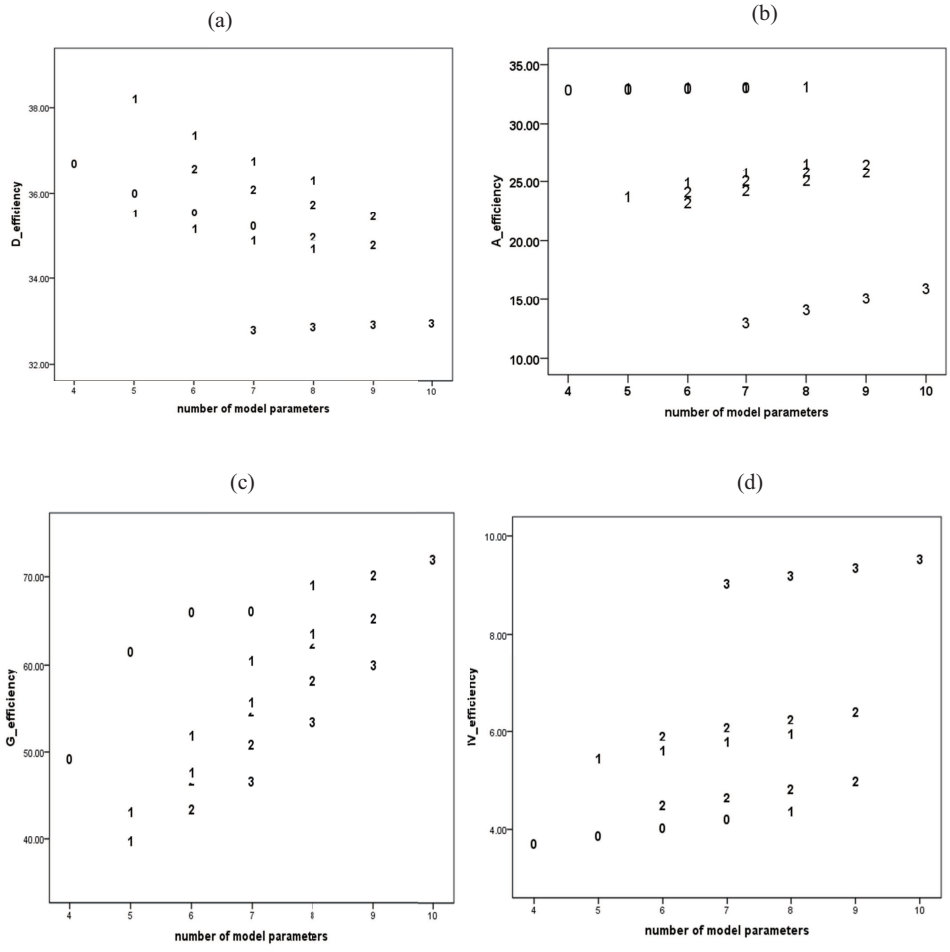

Figure 2
Plots of Reduced Model Efficiencies for the Three-Factor 24 Points Split-Plot CCD. The CCD Consists of 4 Center-Points. Plots (a), (b), (c), and (d) Contain the D, A, G, and IV Efficiencies for the CCD When $d = 1$

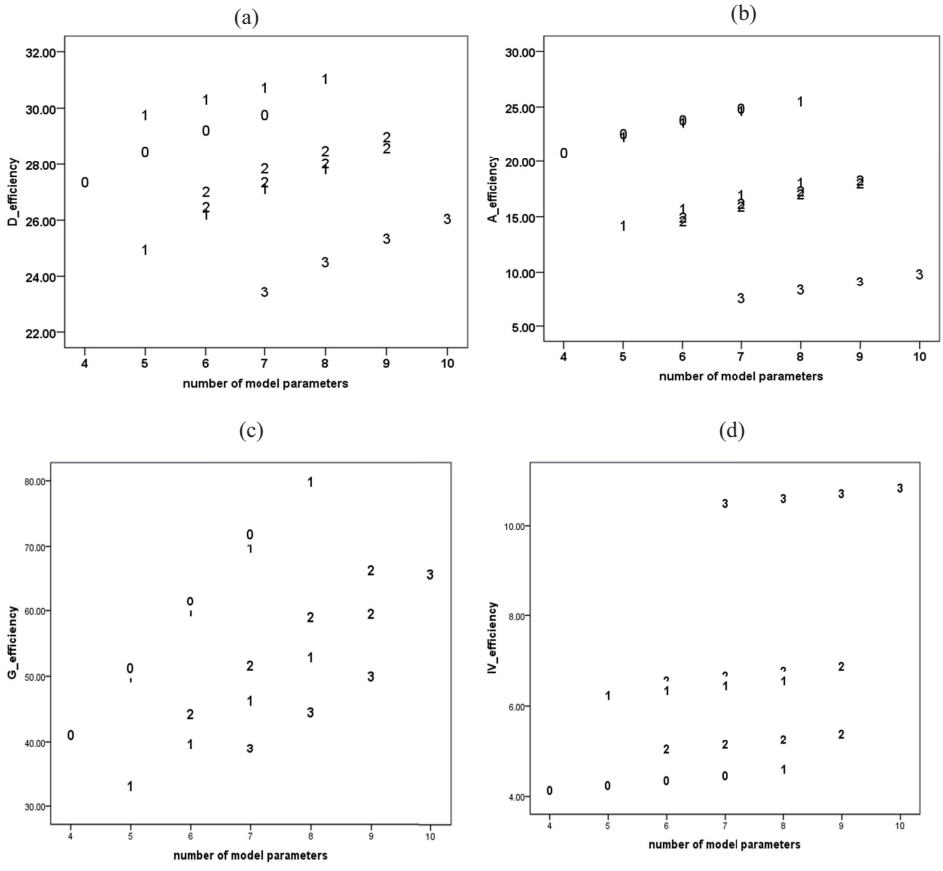

Figure 3
Plots of Reduced Model Efficiencies for the Three-Factor 24 Points Split-Plot CCD. The CCD Consists of 4 Center-Points. Plots (a), (b), (c), and (d) Contain the D, A, G, and IV Efficiencies for the CCD When $d = 2$

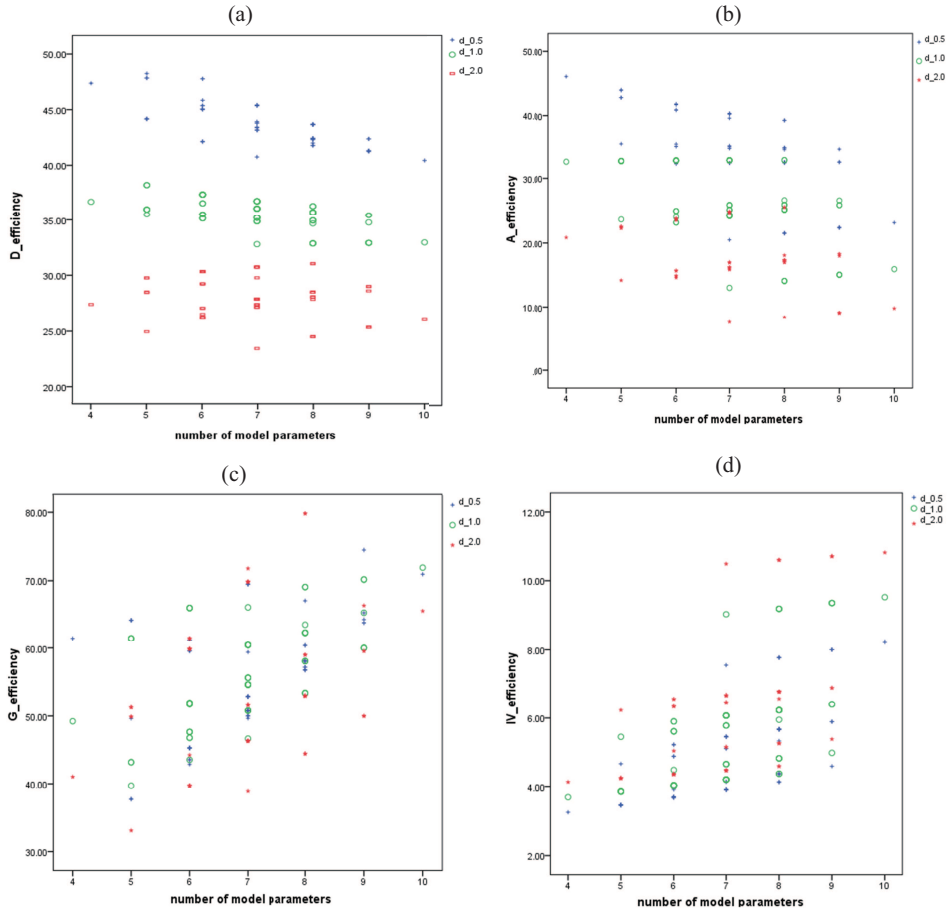

Figure 4
Plots of Reduced Model Efficiencies for the Three-Factor 24 Points Split-Plot CCD. The CCD Consists of 4 Center-Points. Plots (a), (b), (c), and (d) Contain the D, A, G, and IV Efficiencies for the CCD for the Three Scenarios Combined (i.e., $d = 0.5, 1,$ and 2)

5. CONCLUSION

From the second-order reduced split-plot CCD models considered in this work, we have shown that D, A, G, and IV are more sensitive to changes in the pure whole plot squared terms and also to changes in the number of squared terms (q) in the model than to changes in interaction terms. These criteria are robust to changes in the interaction terms and vary dramatically with changes in the pure whole plot squared terms.

D, A, G, and IV are unaffected by the presence or absence of cross-product terms involving the variable whose squared term is removed from the model. The IV for the full model tends to be larger relative to the set of reduced models, while that of the pure linear model ($p = 4$) is smaller relative to the set of other models and therefore the most preferred. In each of the three scenarios, the pure linear model is the most $IV-$efficient while the full quadratic model is the worst $IV-$efficient.

In the combined plots in Figure 4, the most $D-$ and $A-$efficient design is the one with equal whole plot and subplot error variances (i.e., $d = 1$).

In general we see that D, A, G, and IV are not robust across reduced models. They vary dramatically when a whole plot squared term is removed, and are insensitive to changes in the interaction terms. Most interestingly, these criteria measures depend strongly on the ratio of the two variance components d as can be directly seen in Figure 4.

Therefore, when a researcher is faced with a decision of which response surface designs to choose, when some variables are hard to change and some are easy to change, and based on one or more optimality criteria, it is important that these criteria be first determined over a subset of restricted models.

Each of the sixty-four models considered in this work contains full pure linear part and so the effects of linear terms are not investigated.

REFERENCES

[1] Atkinson, A.C., Donev, A.N., & Tobias, R.D. (2007). *Optimum experimental designs, with SAS* (1st ed.). New York: Oxford University Press.

[2] Box, G.E.P., & Wilson, K.B. (1951). On the experimental attainment of optimum conditions (with discussions). *Journal of the Royal Statistical Society (Series B), 13*(1), 1–45.

[3] Box, E.P., Hunter, G., Stuart, J., & Hunter, G.W. (2005). *Statistics for experimenters*. New Jersey: John Willey and Sons, Inc.

[4] Box, G.E.P., & Draper, N.R. (1975). Robust designs. *Biometrika, 62*(2), 347–352.

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

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

[7] Box, G.E.P., & Hunter, J.S. (1957). Multi-factor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics, 28*, 195–241.

[8] Daniel, C. (1976). *Applications of statistics to industrial experiments*. New York: Wiley.

[9] John J.B., & Valeroso, E.S. (2001). Comparison of design optimality criteria of reduced models for response surface designs in the hypercube. *Technometrics, 43*(4), 468–477.

[10] Khuri, A.I., & Cornell, J.A. (1996). *Response surfaces: design and analysis.* New York: Marcel Dekker.

[11] Montgomery, D.C. (2001). *Design and analysis of statistical experiments* (5th ed.). New York: John Wiley & Sons.

[12] Martijn, P.F., Berger, W., & Wong, K. (2009). *An introduction to optimal designs for social and biomedical research* (1st ed.). U.K.: John Wiley & Sons Ltd.

[13] Scott, M., Kowalski, G., Vining, G., Montgomery, G.C., & Borror, C.M. (2006). Modifying central composite design to model the process mean and variance when there are hard-to-change factors. *J. R. S. S. (Series C), 55*(5), 615–630.

[14] Vining, G.G., Kowalski, S.M., & Montgomery, D.C. (2005). Response surface designs within a split-plot structure. *Journal of Quality Technology, 37*, 115–129.