A new graphical approach for comparing response surface designs on the basis of the mean squared error of prediction criterion

S. Mukhopadhyay[1] and A.I. Khuri[2]

[1]*Department of Mathematics, Indian Institute of Technology Bombay, Powai, Mumbai, India.*

[2]*Department of Statistics, University of Florida, Gainesville, FL, USA.*

Abstract

The quality of prediction of a response surface model is measured by the size of its mean squared error within the region of experimentation. The so-called mean squared error of prediction (MSEP) consists of the prediction variance and a measure of bias caused by model misspecification. The purpose of this article is to present a new graphical technique for evaluating and comparing response surface designs using the minimization of the MSEP as a design criterion. Four MSEP-related criteria functions are introduced and plots of their quantile values are obtained on concentric spheres within a region of interest. These plots provide complete information concerning the distribution of each criterion function over the selected spheres. Such information readily depicts the performance of a given design under model misspecification. Furthermore, the proposed criteria are free of any unknown parameters that pertain to the unfitted true model and error variance. Several examples are presented to illustrate the application of the proposed graphical approach and its potential in design augmentation.

Key words: Bias; Model misspecification; Quantile plots; Prediction variance; Response surface model.

1 Introduction

One of the objectives of *response surface methodology* (RSM) is the choice of a design for fitting a hypothesized model. In a typical response surface investigation, such a model is represented by a low-degree polynomial, usually chosen to be of the first degree or the second degree. As a result, designs for fitting models having such a representation are of considerable interest. Given the fact that any

fitted model may not adequately represent the unknown functional relationship that depicts the true mean response in a given experimental situation, one should always be concerned about model bias and the possibility of fitting the wrong model (that is, model misspecification). For this reason, researchers in RSM have introduced several criteria for the choice of design in a manner that protects against a sizeable model bias, that is, when the fitted model differs markedly from what the experimenter fears as the "true" model. Box and Draper (1959, 1963) emphasized the importance of bias contribution in the choice of design. They introduced the so-called *integrated mean squared error* (IMSE) criterion which incorporates prediction variance and model bias. They advocated choosing designs that reduce the bias component of the IMSE unless the prediction variance contribution is considerably larger than that of the bias. This can be accomplished by imposing certain conditions on the moments of the chosen design [see, for example, Chapter 6 in Khuri and Cornell (1996)]. By contrast, Kiefer and several of his co-workers paid less attention to bias, promoting instead the reduction of the prediction variance through the use of the *D-optimality* criterion. The well-known *Equivalence Theorem* of Kiefer and Wolfowitz (1960) led to the development of a practical algorithm for the construction of a D-optimal design through the use of the *G-optimality* criterion. Other related variance criteria are *A-optimality* and *E-optimality*. All such criteria are usually referred to as *alphabetic optimality* criteria.

Whether the design is chosen on the basis of bias, by requiring it to satisfy certain conditions on its design moments, or on the basis of an alphabetic optimality criterion, design construction is traditionally done using single-valued criteria functions. This, however, does not give adequate information about the quality of prediction afforded by the design throughout the experimental region. More recently, several authors in RSM addressed such concerns by proposing certain graphical techniques for comparing response surface designs. Giovannitti-Jensen and Myers (1989) introduced the so-called *variance dispersion graphs* (VDGs), which consist of two-dimensional plots that display the maximum, minimum, and average of the prediction variance on concentric spheres, chosen within the experimental region, against their radii. Khuri, Kim, and Um (1996) proposed the use of *quantile plots* (QPs) of the prediction variance values on the surfaces of concentric spheres of varying radii within the experimental region.

The QPs provide complete information about the distribution of the prediction variance through its quantiles on a given sphere. It is obviously true that having knowledge of the entire distribution of the prediction variance can be more informative than knowing only its extremes (maximum and minimum), or its average, as is the case with the VDGs. More recently, Zahran, Anderson-Cook, and Myers (2003) developed the *fraction of design space* plots where the prediction variance is plotted against the fraction of design space that has prediction variance at or below a certain value.

The aforementioned graphical techniques were concerned only with the prediction variance. Vining and Myers (1991) proposed a graphical approach for evaluating and comparing response surface designs on the basis of the mean squared error of prediction (MSEP), which incorporates prediction variance and bias contributions. As in the VDGs approach, Vining and Myers considered plots of the maximum, minimum, and average of the MSEP over the surfaces of concentric spheres against their radii. Their approach, however, requires the specification of the value of an unknown index, denoted by w_B, which measures the relative contribution of the bias. Furthermore, having knowledge of only the maximum, minimum, or average of the MSEP values on a given sphere is not sufficient to give adequate information about the actual distribution of the MSEP on the sphere.

In this article, we present an alternative approach to the graphical technique by Vining and Myers (1991). The proposed approach introduces four criteria functions derived from the MSEP which do not require the specification of the value of the unknown index, w_B. Quantile plots of the values of each of these functions on concentric spheres within the experimental region are obtained. Such plots provide complete descriptions of the distributions of the criteria functions over the spheres. This makes it possible to evaluate and compare several response surface designs on the basis of the size of MSEP throughout the experimental region. Several examples are presented to illustrate the application of the proposed methodology. One of these examples demonstrates how this approach can be utilized to augment a given design with added design points that can lead to a reduction in the size of the MSEP.

2   The mean squared error of prediction

Suppose that the true underlying relationship, $y = g(\mathbf{x}) + \epsilon$, between the response y and the vector of k control variables $\mathbf{x} = (x_1, \ldots, x_k)'$ is unknown, where ϵ is a random error. The experimenter believes that the function g can be approximated by fitting a polynomial model of degree d_1 in the control variables over a region of interest, R, using an n-point design D. The fitted model is then

$$y(\mathbf{x}) = \boldsymbol{f}_1'(\mathbf{x})\hat{\boldsymbol{\beta}}_1, \tag{1}$$

where $\boldsymbol{f}_1'(\mathbf{x})\hat{\boldsymbol{\beta}}_1$ is a polynomial of degree d_1 in $x_1, \ldots, x_k$ and

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}.$$

Here, $\mathbf{X}_1$ is the $n \times p_1$ matrix whose uth row is of the form $\boldsymbol{f}_1'(\mathbf{x}_u)$, $\mathbf{x}_u$ is the value of $\mathbf{x}$ at the uth experimental run and $\mathbf{y}$ is the $n \times 1$ vector of observed responses.

The true relationship g over the entire region R is assumed to be a polynomial of degree $d_2(> d_1)$ in $\mathbf{x}$, that is,

$$y = \boldsymbol{f}_1'(\mathbf{x})\boldsymbol{\beta}_1 + \boldsymbol{f}_2'(\mathbf{x})\boldsymbol{\beta}_2 + \epsilon, \tag{2}$$

where $\boldsymbol{\beta}_2$ is a $p_2 \times 1$ vector of parameters. Corresponding to the n-point design D there exists an $n \times p_2$ matrix $\mathbf{X}_2$ whose uth row is of the form $\boldsymbol{f}_2'(\mathbf{x}_u)$.

The mean squared error of prediction (MSEP) at a point $\mathbf{x}$ in the region R, denoted by MSE($\mathbf{x}$), is then [see Vining and Myers (1991, p. 316)]

$$MSE[\hat{y}(\mathbf{x})] = \sigma^2 \boldsymbol{f}_1'(\mathbf{x})(\mathbf{X}_1'\mathbf{X}_1)^{-1}\boldsymbol{f}_1(\mathbf{x}) + \boldsymbol{\beta}_2'\mathbf{A}^*(\mathbf{x})\boldsymbol{\beta}_2, \tag{3}$$

where $\boldsymbol{A}^*(\mathbf{x}) = [\boldsymbol{f}_1'(\mathbf{x})\mathbf{A} - \boldsymbol{f}_2'(\mathbf{x})]'[\boldsymbol{f}_1'(\mathbf{x})\mathbf{A} - \boldsymbol{f}_2'(\mathbf{x})]$ and $\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$. Note that $\boldsymbol{A}^*(\mathbf{x})$ is a positive semidefinite matrix of rank 1.

3   Design comparison criteria

An important criterion for the choice of design is the minimization of $MSE(\mathbf{x})$, given in formula (3), over the region R. Since $MSE(\mathbf{x})$ depends on $\boldsymbol{\beta}_2$ and σ^2, which are unknown, we propose instead three

other criteria functions, namely, Δ, L_1 and L_2 (defined below). We show that small values of Δ, L_1 and L_2 indicate small mean-squared error of prediction (MSEP) values.

Using the Cauchy-Schwartz inequality we can write

$$
\begin{aligned}
MSE[\hat{y}(\mathbf{x})] &= \sigma^2 \boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}) + \boldsymbol{\beta}'_2\mathbf{A}^*(\mathbf{x})\boldsymbol{\beta}_2 \\
&\leq \sigma^2 \boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}) + \boldsymbol{\beta}'_2\boldsymbol{\beta}_2 Tr[\mathbf{A}^*(\mathbf{x})] \\
&\leq [\sigma^4 + (\boldsymbol{\beta}'_2\boldsymbol{\beta}_2)^2]^{1/2}[\{\boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x})\}^2 + \{Tr(\mathbf{A}^*(\mathbf{x}))\}^2]^{1/2}.
\end{aligned}
\tag{4}
$$

Small values of $\Delta(\mathbf{x}, D) = \{\boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x})\}^2 + \{Tr(\mathbf{A}^*(\mathbf{x}))\}^2$ indicate small MSEP values for design D.

To define L_1 and L_2, we consider next the maximum of the MSEP over a sphere of radius r in the $\boldsymbol{\beta}_2$ space. Using formula (1.2) in Vining and Myers (1991), we have

$$
\max_{|\boldsymbol{\beta}_2|=r} MSE[\hat{y}(\mathbf{x})] = \sigma^2 \boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}) + r^2 Tr[\mathbf{A}^*(\mathbf{x})]
\tag{5}
$$

scaling the above equation by $\sigma^2 + r^2$ we get

$$
L(\mathbf{x}, D, w) = (1 - w)\boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}) + w Tr[\mathbf{A}^*(\mathbf{x})],
\tag{6}
$$

where $L(\mathbf{x}, D, w) = \frac{\max_{|\boldsymbol{\beta}_2|=r} MSE[\hat{y}(\mathbf{x})]}{\sigma^2 + r^2}$ and $w = \frac{r^2}{\sigma^2 + r^2}$.

Therefore, since $0 \leq w \leq 1$, we have

$$
L_1(\mathbf{x}, D) \leq L(\mathbf{x}, D, w) \leq L_2(\mathbf{x}, D),
\tag{7}
$$

where

$$
\begin{aligned}
L_1(\mathbf{x}, D) &= \min\{\boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}), Tr[\mathbf{A}^*(\mathbf{x})]\}, \text{ and} \\
L_2(\mathbf{x}, D) &= \max\{\boldsymbol{f}'_{\boldsymbol{1}}(\mathbf{x})(\mathbf{X}'_1\mathbf{X}_1)^{-1}\boldsymbol{f}_{\boldsymbol{1}}(\mathbf{x}), Tr[\mathbf{A}^*(\mathbf{x})]\}.
\end{aligned}
$$

Small values of the upper bound (L_2) indicate small MSEP values for design D. The closeness of the L_1 and L_2 values indicate robustness of $L(\mathbf{x}, D, w)$, as induced by the design D, to changes in the values of w. So, it will be desirable for both L_1 and L_2 to be small. Thus instead of studying $MSE[\hat{y}(\mathbf{x})]$ directly, we shall compare designs on the basis of their $\Delta(\mathbf{x}, D)$, $L_1(\mathbf{x}, D)$ and $L_2(\mathbf{x}, D)$ values over the entire region R. It is to be noted that all three of these criteria functions do not depend on w, which is unknown.

4   Quantile plots

In this section, we show how to obtain estimated quantiles of the chosen criterion on a sphere $S(\rho)$ of radius ρ centered at the origin inside R. Any point $\mathbf{x} = (x_1, \ldots, x_k)'$ on $S(\rho)$ can be represented by using independent $k-1$ spherical coordinates $\psi_1, \ldots, \psi_{k-1}$ [see Khuri, Kim and Um (1996)] such that

$$
\begin{aligned}
x_1 &= \rho \cos \psi_1 \\
x_2 &= \rho \sin \psi_1 \cos \psi_2 \\
&\vdots \\
x_{k-1} &= \rho \sin \psi_1 \ldots \sin \psi_{k-2} \cos \psi_{k-1} \\
x_k &= \rho \sin \psi_1 \ldots \sin \psi_{k-2} \sin \psi_{k-1},
\end{aligned}
$$

where $0 \leq \psi_1 \leq \pi, \ldots, 0 \leq \psi_{k-2} \leq \pi$ and $0 \leq \psi_{k-1} \leq 2\pi$. To generate points on $S(\rho)$, $\psi_1, \ldots, \psi_{k-1}$ are randomly chosen from independent uniform distributions, namely $\psi_i \sim U(0, \pi)$, $i = 1, \ldots, k-2$ and $\psi_{k-1} \sim U(0, 2\pi)$.

4.1   Quantile plots of $\Delta(\mathbf{x}, D)$ values

For a given design D and a set of generated values of $\mathbf{x}$ on $S(\rho)$, we obtain a sample, denoted by $\tau_{\rho,\Delta}(D)$, consisting of $\Delta(\mathbf{x}, D)$ values on $S(\rho)$. Let $Q_{\rho,\Delta}(D, p)$ denote the p-th quantile of $\tau_{\rho,\Delta}(D)$. These quantiles provide a description of the distribution of $\Delta(\mathbf{x}, D)$ for values of $\mathbf{x}$ on $S(\rho)$.

Plotting these values against p results in the quantile plots of $\Delta(\mathbf{x}, D)$ over the surface of $S(\rho)$. By repeating the same process for several selected values of ρ, we obtain plots that portray the prediction capability associated with the design D throughout the region R. Such plots can be constructed for each of several candidate designs for the model.

It should be noted that for a given ρ, a desirable feature of a design is to have small values of $Q_{\rho,\Delta}(D, p)$ over the range of p ($0 \leq p \leq 1$). The smallness of $Q_{\rho,\Delta}(D, p)$ indicates small MSEP values.

There are several advantages to this approach, namely, the performance of a design can be evaluated throughout the region R, and detailed information can be extracted about the distribution of $\Delta(\mathbf{x}, D)$ on $S(\rho)$, including, but not limited to, the median ($p = 0.50$), the first quartile (p = 0.25), and the third quartile ($p = 0.75$).

4.2   Quantile plots for $L_1(\mathbf{x}, D)$ and $L_2(\mathbf{x}, D)$

For a given design D and a set of generated values of $\mathbf{x}$ on $S(\rho)$, we obtain samples $\tau_{\rho, L_1}(D)$ and $\tau_{\rho, L_2}(D)$ consisting of $L_1(\mathbf{x}, D)$ and $L_2(\mathbf{x}, D)$ values on $S(\rho)$. Let $Q_{\rho, L_1}(D, p)$ and $Q_{\rho, L_2}(D, p)$ denote the p-th quantiles of $\tau_{\rho, L_1}(D)$ and $\tau_{\rho, L_2}(D)$, respectively. Plotting $Q_{\rho, L_1}(D, p)$ and $Q_{\rho, L_2}(D, p)$ against p results in the quantile plots of $L_1(\mathbf{x}, D)$ and $L_2(\mathbf{x}, D)$, respectively, over the region R.

5   Examples

In this section, we present several examples to illustrate the application of the quantile plots approach. All the designs considered in this section are three-variable designs.

5.1   Example 1: Choice of center runs

The following first-degree model in k control variables $x_1, \ldots, x_k$ is fitted using a design D with n runs,

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \epsilon. \qquad (8)$$

Now suppose that the analyst wishes to protect against the full second-degree model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum\sum_{i<j} \beta_{ij} x_i x_j + \epsilon. \qquad (9)$$

For $k = 3$ control variables we consider: a 2^3-factorial design with no center runs (D_1), a 2^3-factorial design with one center run (D_2), a 2^3-factorial design with two center runs (D_3), and a 2^3-factorial design with five center runs (D_4). These designs were also considered in Vining and Myers (1991) and were scaled so that all the design points fall within a sphere of radius 1.

For each radius $\rho\,(= 0.3, 0.6, 0.9, 1)$, 10,000 points are generated at random on $S(\rho) = \{\mathbf{x} : x_1^2 + x_2^2 + x_3^2 = \rho^2\}$ by using the following

equations,

$$
\begin{aligned}
x_1 &= \rho \cos \psi_1 \\
x_2 &= \rho \sin \psi_1 \cos \psi_2 \\
x_3 &= \rho \sin \psi_1 \sin \psi_2,
\end{aligned}
\tag{10}
$$

where $\psi_1 \sim U(0, \pi)$ and $\psi_2 \sim U(0, 2\pi)$. The resulting value of $\mathbf{x} = (x_1, x_2, x_3)'$ is used to evaluate Δ, L_1 and L_2 on $S(\rho)$.

Quantile plots of $Q_{\rho,\Delta}(D, p)$ for designs D_1 and D_4 for selected values of $\rho \, (= 0.3, 0.6, 0.9, 1)$ and $p = 0(0.5)1$ are given in Figure 1. While, plots of $Q_{\rho,L_1}(D, p)$ and $Q_{\rho,L_2}(D, p)$ for designs D_1, D_2, D_3 and D_4 for selected values of $\rho \, (= 0.3, 0.6, 0.9, 1)$ and $p = 0(0.5)1$ are shown in Figures 2 and 3.

From Figure 1 we note that quantiles of Δ for design D_4 are smaller than those of design D_1 for values of $\rho \, (= 0.3, 0.6, 0.9)$ and all values of p. Increasing ρ from 0.3 to 0.9 decreases the distance between the quantiles for designs D_1 and D_4. For $\rho = 1$ and $p > 0.8$, the quantiles of design D_4 are slightly larger than those of design D_1.

From the quantile plot of $Q_{\rho,L_1}(D, p)$ in Figure 2, we note that for $\rho = 0.3$, $Q_{\rho,L_1}(D_1, p) > Q_{\rho,L_1}(D_i, p)$ for $i \, (= 2, 3, 4)$, while $Q_{\rho,L_1}(D_4, p) < Q_{\rho,L_1}(D_i, p)$ for $i \, (= 1, 2, 3)$. When ρ is 0.6, the quantiles of L_1 for design D_1 increase slightly and are higher than the other three designs, while D_4 has the lowest $Q_{\rho,L_1}(D, p)$ values among all four designs. Thus for values of $\rho \, (= 0.3, 0.6)$, we can say that adding center runs helps the design's performance. As we approach the perimeter of the experimental region (i.e. for $\rho = 0.9, 1$) the $Q_{\rho,L_1}(D, p)$ values for each of the four designs increase and almost overlap one another. When ρ is 1, the $Q_{\rho,L_1}(D_4, p)$ values are higher than all the other three designs for $p < 0.4$, and lower than all the other three designs for $p > 0.4$.

Figure 3 shows that when ρ is 0.3, $Q_{\rho,L_2}(D_1, p) > Q_{\rho,L_2}(D_i, p)$ for $i \, (= 2, 3, 4)$, while $Q_{\rho,L_2}(D_4, p) < Q_{\rho,L_2}(D_i, p)$ for $i \, (= 1, 2, 3)$. For $\rho = 0.6$, the quantiles of L_2 for designs D_2, D_3 and D_4 increase, and the $Q_{\rho,L_2}(D, p)$ values of all four designs come close to one another. As the value of ρ increases to 0.9, there is a general increase in the quantiles of L_2. Thus, we note from Figure 3 that for $\rho \, (= 0.3, 0.6, 0.9)$, D_1 is the worst design while D_4 is the best, but when $\rho = 1$ and $p > 0.5$, D_4 is the worst design.

From Figures 1-3 we can then conclude that adding center runs helps the design's performance near the center, but does not improve its performance near the perimeter. Comparing the above results

with Figure 5 of Vining and Myers' (1991) paper, we come to the same conclusion that adding center runs appear to offer some benefit, but this benefit is not uniform over the entire experimental region.

5.2   Example 2: 2^3-factorial and orthogonal simplex designs

Here, we again consider the same models as in (8) and (9). For $k = 3$ control variables, the designs used are a 2^3-factorial design with no center runs (D_1) and an orthogonal simplex design with replicated points (D_2). The designs (shown in Tables 1 and 2) were scaled so that all the design points fall within a sphere of radius 1.

For each radius ρ, 10,000 points are generated at random on $S(\rho)$ using formula (10). Quantile plots of $Q_{\rho,\Delta}(D,p)$, $Q_{\rho,L_1}(D,p)$ and $Q_{\rho,L_2}(D,p)$ for designs D_1 and D_2 and for selected values of $\rho\,(=0.3, 0.6, 0.9, 1)$ and $p = 0(0.5)1$ are shown in Figures 4, 5 and 6.

From the quantile plots of $Q_{\rho,\Delta}(D,p)$ in Figure 4, we note that for $\rho\,(= 0.3, 0.6)$ and almost all values of p, the $Q_{\rho,\Delta}(D,p)$ values for designs D_1 and D_2 overlap with each other. As ρ increases (i.e. for $\rho = 0.9, 1.0$), the quantiles of Δ for design D_2 increase sharply and are higher than those of design D_1. We can then conclude that the performance of the orthogonal simplex design (D_2) deteriorates as we approach the perimeter of the experimental region.

From Figure 5 we note that when ρ is 0.3, the $Q_{\rho,L_1}(D,p)$ values for both designs D_1 and D_2 overlap each other. For $\rho = 0.6$, the quantiles of L_1 for design D_2 increase and are higher than those for design D_1. As we increase ρ to 0.9, there is an increase in the $Q_{\rho,L_1}(D,p)$ values. For $\rho = 0.9$ and values of p $(0.2 < p < 0.8)$, $Q_{\rho,L_1}(D_2,p) > Q_{\rho,L_1}(D_1,p)$; for $p < 0.2$, $Q_{\rho,L_1}(D_2,p) < Q_{\rho,L_1}(D_1,p)$; and for $p > 0.8$, $Q_{\rho,L_1}(D_2,p)$ overlaps with $Q_{\rho,L_1}(D_1,p)$. As we approach the perimeter of the experimental region (i.e. $\rho = 1$) and for $p > 0.6$, the $Q_{\rho,L_1}(D,p)$ values for each of the two designs increase further and overlap each other; for $p < 0.2$, $Q_{\rho,L_1}(D_2,p) < Q_{\rho,L_1}(D_1,p)$; and for $0.2 < p < 0.6$, $Q_{\rho,L_1}(D_2,p) > Q_{\rho,L_1}(D_1,p)$. Thus we can say that the performance of both designs deteriorates as we approach the perimeter of the experimental region.

Figure 6 shows that for $\rho = 0.3$, $Q_{\rho,L_2}(D,p)$ values for both designs are small and overlap each other. As we increase ρ $(\rho = 0.6, 0.9, 1)$, $Q_{\rho,L_2}(D_1,p) < Q_{\rho,L_2}(D_2,p)$. Thus, the performance of D_1 is better

than D_2 near the perimeter of the experimental region.

5.3 Example 3: Central-composite and Box-Behnken designs

In this example, we fit the full second-degree model (9) in k control variables while protecting against a cubic model of the form

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \sum_{j=1}^{k} \beta_{ijj} x_i x_j^2 + \epsilon. \tag{11}$$

For $k = 3$ control variables the designs used are a central composite design (CCD) consisting of a 2^3 design plus 6 axial points and one center point (D_1), and a 15-point Box-Behnken design (D_2) [Khuri and Cornell (1996, p. 119)]. The designs (given in Tables 3 and 4) were scaled so that all the design points fall within a sphere of radius 1.

For each radius ρ, 10,000 points are generated at random on $S(\rho)$ using formula (10). Quantile plots of $Q_{\rho,\Delta}(D,p)$, $Q_{\rho,L_1}(D,p)$ and $Q_{\rho,L_2}(D,p)$ for designs D_1 and D_2 and selected values of $\rho\,(= 0.3, 0.6, 0.9, 1)$ and $p = 0(0.5)1$ are given in Figures 7, 8 and 9.

From Figure 7 we note that for $\rho\,(= 0.3, 0.6)$, the values of $Q_{\rho,\Delta}(D_1,p)$ are much higher than $Q_{\rho,\Delta}(D_2,p)$. Increasing ρ, $Q_{\rho,\Delta}(D_1,p)$ values remain stable, while $Q_{\rho,\Delta}(D_2,p)$ values increase. For $\rho\,(= 0.9, 1)$, $Q_{\rho,\Delta}(D_1,p)$ are lower than $Q_{\rho,\Delta}(D_2,p)$.

The quantile plots in Figure 8 show that for $\rho = 0.3$, the values of $Q_{\rho,L_1}(D,p)$ for both designs are small and are also very close to each other. As we approach the perimeter of the experimental region, the $Q_{\rho,L_1}(D,p)$ values for both designs increase and $Q_{\rho,L_1}(D_2,p) > Q_{\rho,L_1}(D_1,p)$ for almost all values of p.

From Figure 9 it is seen that for $\rho\,(= 0.3, 0.6)$, the values of $Q_{\rho,L_2}(D_1,p)$ are much higher than $Q_{\rho,L_2}(D_2,p)$. As we increase ρ, $Q_{\rho,L_2}(D_1,p)$ values remain stable, while $Q_{\rho,L_2}(D_2,p)$ values increase. For $\rho = 0.9, 1$, $Q_{\rho,L_2}(D_1,p)$ are lower than $Q_{\rho,L_2}(D_2,p)$.

Thus we can conclude from Figures 7-9 that the 15-point Box Behnken design (D_2) tends to perform worse than the CCD (D_1) near the perimeter of the experimental region, but D_2 is better than D_1 near the center.

5.4 Example 4: Design augmentation

In this example, we fit the first-degree model (8) in k control variables while protecting against a full second-degree model (9). The purpose of this example is to show how the MSEP is affected by the addition of one single point to the original design. By choosing several candidate points to augment the design, we can utilize the quantile plots for each augmented design to select the "best" point that gives the lowest quantile values.

For $k = 3$ control variables, we start with a 2^3-factorial design with no center runs (D_1). Design D_1 is our original design. Next, we augment D_1 with the center point $(0, 0, 0)$ and denote this new design by D_2 (Table 5). Design D_3 is the original design D_1 augmented with the point $(0.5, 0, 0)$, and D_4 is design D_1 augmented with the point $(0.75, 0, 0)$. Augmenting D_1 with the point $(1, 0, 0)$ gives D_5. The designs were scaled so that all the design points fall within a sphere of radius 1.

For $\rho = 1$, 10,000 points are generated at random on $S(\rho)$ using formula (10). Quantile plots of $Q_{\rho, L_2}(D, p)$ for designs D_1, D_2, D_3, D_4 and D_5 for $p = 0(0.5)1$ are given in Figure 10. We can also generate similar quantile plots for the five designs $D_1 - D_5$ for other values of ρ.

From Figure 10 we note that design D_2 has $Q_{\rho, L_2}(D, p)$ values very close to the original design D_1. For values of $p < 0.5$, the $Q_{\rho, L_2}(D_2, p)$ values are slightly lower than $Q_{\rho, L_2}(D_1, p)$.

Design D_3 has $Q_{\rho, L_2}(D, p)$ values lower than the original design D_1. For values of $p < 0.4$, there is a substantial gap between $Q_{\rho, L_2}(D_3, p)$ and $Q_{\rho, L_2}(D_1, p)$, but this gap decreases and the quantiles of L_2 for both designs D_1 and D_3 come closer together for $p > 0.6$.

Augmenting the original design by the point $(0.75, 0, 0)$ we note that values of $Q_{\rho, L_2}(D_4, p)$ are lower than $Q_{\rho, L_2}(D_1, p)$ for $p < 0.8$.

For values of $p < 0.5$, $Q_{\rho, L_2}(D_5, p)$ are lower than $Q_{\rho, L_2}(D_1, p)$. But as p increases, we see that D_1 is better than D_5.

Thus, from Figure 10 we can conclude that augmenting the original design D_1 by the extra point, $(0.5, 0, 0)$, causes a decrease in the quantiles of L_2 for almost all values of p. Augmenting the original design by $(0, 0, 0)$ fails to make a noticeable change in the quantiles of L_2. While, augmenting D_1 by $(0.75, 0, 0)$ or $(1, 0, 0)$ increases the values of the quantiles of L_2 for $p > 0.8$ and $p > 0.5$, respectively. Thus, the point $(0.5, 0, 0)$ gives the lowest quantile values among the

four chosen points for almost all values of p. The above process can be continued in a sequential manner by augmenting D_1 with additional points beyond the point $(0.5, 0, 0)$.

6   Comparison based on the quantiles of $L(\mathbf{x}, D, w)$

When comparing designs using the upper bound $L_2(\mathbf{x}, D)$, we have to be careful about forming our conclusions since the value of this upper bound may sometimes be large while the actual value of $L(\mathbf{x}, D, w)$ in formula (6) is small. To avoid such a situation, it is recommended to work directly with the actual expression of $L(\mathbf{x}, D, w)$ itself. The only problem here is that of the unknown w. To overcome this problem, we suggest generating points on the sphere $S(\rho)$ for a given ρ as before. Then we select several values of w from the interval $[0, 1]$. Let T denote the set consisting of such selected values. For each design D, we compute $L(\mathbf{x}, D, w)$ for all values of w in T and the generated points on $S(\rho)$. Quantiles of the resulting values of L can then be obtained and plotted against p. For illustration, let us now consider two examples. In both examples we fit a first-degree model (8) in three control variables while protecting against a full second-degree model (9). The designs are scaled so that all the points fall within a sphere of radius 1.

6.1   Example 5: Comparison of 2^3 factorial and orthogonal simplex designs based on $L(\mathbf{x}, D, w)$ values

We compare the same designs, a 2^3 factorial design with no center points (D_1) and an orthogonal simplex design (D_2), as in Example 2, using the quantile plots of $L(\mathbf{x}, D, w)$ for all $\mathbf{x}$ on $S(\rho)$ and w in T.

For each radius ρ, we generated 10,000 points on $S(\rho)$ at random. For each of these 10,000 values of $\mathbf{x}$ and $w = 0(0.1)1$, values of $L(\mathbf{x}, D, w)$ were obtained for both designs D_1 and D_2. These were respectively denoted by $\tau_{\rho,L}(D_1)$ and $\tau_{\rho,L}(D_2)$. Let $Q_{\rho,L}(D, p)$ denote the quantiles of $\tau_{\rho,L}(D)$. Figure 11 shows the quantile plots of $Q_{\rho,L}(D, p)$ for designs D_1 and D_2 and selected values of ρ (= $0.3, 0.6, 0.9, 1$). From Figure 11 we note that for $\rho = 0.3$, the quantile

values for both designs are small, but close to each other. Near the perimeter of the experimental region, $Q_{\rho,L}(D_2, p) > Q_{\rho,L}(D_1, p)$ for almost all values of p. Thus, D_1 and D_2 are similar in performance near the center, but as we go towards the perimeter of the region, D_1 has a better prediction capability than D_2. From Figures 6 and 11 we note that the quantile plots of L and L_2 are very similar for both designs.

6.2 Example 6: Comparison of 2^3 factorial and a non-standard design based on $L(\mathbf{x}, D, w)$ values

Next, we compare a 2^3 factorial design with no center runs (D_1) with a non-standard design D_2 given in Table 6. Following the same steps as in Section 6.1, we obtain the quantile plots of L and L_2 (Figures 12 and 13, respectively) for selected values of ρ ($= 0.3, 0.6, 0.8, 1$) and $w = 0(0.1)1$. From Figure 12 we observe that both designs D_1 and D_2 have similar prediction capabilities near the center of the experimental region, i.e., $\rho = 0.3, 0.6$ and for almost all values of p. As we increase ρ to $0.8, 1$, D_2 has higher $Q_{\rho,L}(D, p)$ values than D_1 for $p \geq 0.4$. From Figure 13 we note that for $\rho = 0.6, 0.8, 1$ and almost all values of p, the $Q_{\rho,L_2}(D_2, p)$ values are higher than the $Q_{\rho,L_2}(D_1, p)$ values.

Figure 14 provides a comparison of $Q_{\rho,L}(D_2, p)$ with $Q_{\rho,L_2}(D_2, p)$. From Figure 14 we note that the gap between the quantiles of L and L_2 for design D_2 increases with an increase in ρ. This indicates that near the perimeter of the experimental region, the actual value of L for design D_2 is much smaller than the value of the upper bound L_2. Thus, conclusions drawn using quantile plots of L_2 suggest a bigger difference in the prediction capabilities of D_1 and D_2 than if we had used L.

7 Concluding remarks

In this article, we propose four criterion functions for comparing designs, Δ, L, L_1 and L_2. Small values of all four criteria indicate that the designs being compared have small MSE values. In situations where the values of L_2 are large, it is advisable to make a comparison between L and L_2. A big difference between L and L_2 suggest that

the actual value of L is much smaller than the upper bound L_2, and design comparisons should be based on the L values.

The examples presented in Sections 5 and 6 demonstrate that the proposed approach of quantile plots is an effective tool for evaluating and comparing several response surface designs over the entire experimental region. These plots provide more information concerning the distribution of the MSEP for a given design than the graphical approach of Vining and Myers (1991). The plots reveal which parts of the experimental region are more sensitive to model misspecification. Example 4 shows how the proposed approach can be utilized to augment a given design with a single point in order to select the "best" point that gives the lowest quantile values.

It should be noted that, thus far, the technique of using quantile plots has traditionally been used only to compare several given designs, but not to find a design prior to the experiment. However, using design augmentation we can start with a given design and augment it with points in order to obtain a new design with lower quantile values. This new design can then be used for experimentation as it has better prediction capabilities than the original design. Finally, It is important to point out that the design comparison criteria proposed in Section 3 do not require the specification of an unknown index, as was the case in Vining and Myers (1991).

References

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

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

Giovannitti-Jensen, A. and Myers, R. H. (1989). Graphical assessment of the prediction capability of response surface designs. *Technometrics* **31**, 159-171.

Khuri, A. I. and Cornell, J. A. (1996). *Response Surfaces*, 2nd Edition, Dekker, New York.

Khuri, A. I., Kim, H. J. and Um, Y. (1996). Quantile plots of the prediction variance for response surface designs. *Computational*

Statistics and Data Analysis **22**, 395-407.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12**, 363-366.

Vining, G. G. and Myers, R. H. (1991). A graphical approach for evaluating response surface designs in terms of the mean squared error of prediction. *Technometrics* **33**, 315-326.

Zahran, A., Anderson-Cook, C. M. and Myers, R. H. (2003). Fraction of design space to assess prediction capability of response surface designs. *J. Quality Technology* **35**, 377-386.

S. Mukhopadhyay
Department of Mathematics
Indian Institute of Technology, Bombay
Powai, Mumbai 400076, India.
Email: siuli@math.iitb.ac.in

A.I. Khuri
Department of Statistics
University of Florida
Gainesville, FL 32611, USA.
Email: ufakhuri@stat.ufl.edu

Table 1: Design D_1, 2^3-factorial design on a radius of 1 with no center runs. (Examples 1 and 2)

x_1	x_2	x_3
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$

Table 2: Design D_2, orthogonal simplex design with replicated points on a radius of 1 (Example 2).

x_1	x_2	x_3
$-\sqrt{\frac{2}{3}}$	$-\frac{\sqrt{2}}{3}$	$-\frac{1}{3}$
$-\sqrt{\frac{2}{3}}$	$-\frac{\sqrt{2}}{3}$	$-\frac{1}{3}$
$\sqrt{\frac{2}{3}}$	$-\frac{\sqrt{2}}{3}$	$-\frac{1}{3}$
$\sqrt{\frac{2}{3}}$	$-\frac{\sqrt{2}}{3}$	$-\frac{1}{3}$
0	$\frac{2^{3/2}}{3}$	$-\frac{1}{3}$
0	$\frac{2^{3/2}}{3}$	$-\frac{1}{3}$
0	0	1
0	0	1

Table 3: Design D_1, a CCD on a radius of 1 (Example 3).

x_1	x_2	x_3
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
1	0	0
-1	0	0
0	1	0
0	-1	0
0	0	1
0	0	-1
0	0	0

Table 4: Design D_2, a 15-point Box Behnken design on a radius of 1 (Example 3).

x_1	x_2	x_3
$-\frac{1}{\sqrt{2}}$	0	$-\frac{1}{\sqrt{2}}$
0	0	0
0	$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$
$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0
$-\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$
0	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$
$\frac{1}{\sqrt{2}}$	0	$-\frac{1}{\sqrt{2}}$
0	0	0
$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0
$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
0	$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
0	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
0	0	0
$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0
$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$

Table 5: Design D_2, original design D_1 augmented with one center run (Example 4).

x_1	x_2	x_3
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$
$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$
0	0	0

Table 6: Design D_2, a non-standard design on a radius of 1 with no center runs. (Example 6)

x_1	x_2	x_3
$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$
$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$
$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$
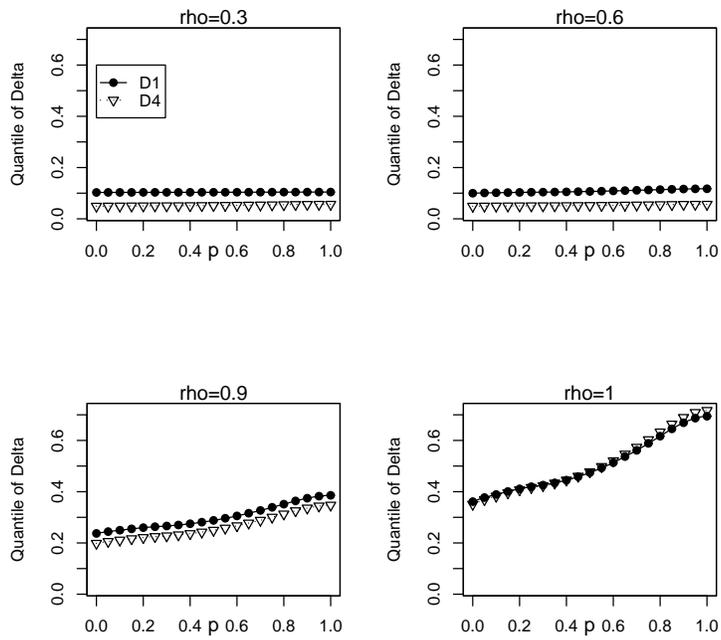
Figure 1: Quantile plots of Δ for designs D_1 (a 2^3-factorial design with no center runs) and D_4 (a 2^3-factorial design with five center runs) (Example 1).
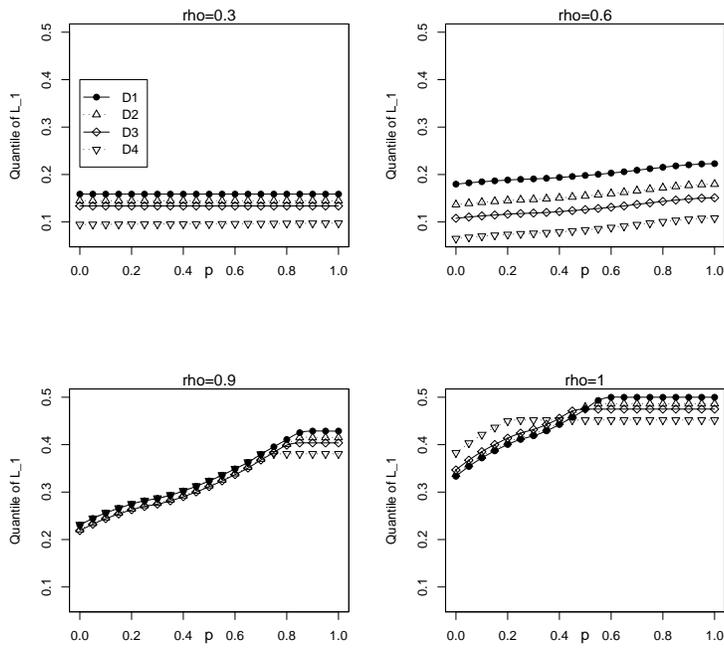
Figure 2: Quantile plots of L_1 for designs D_1 (a 2^3-factorial design with no center runs), D_2 (a 2^3-factorial design with one center run), D_3 (a 2^3-factorial design with two center runs) and D_4 (a 2^3-factorial design with five center runs) (Example 1).
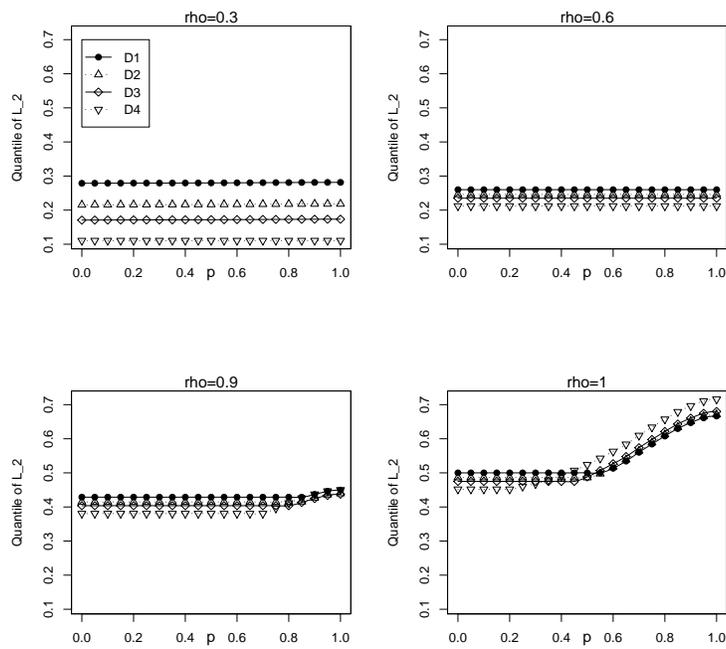
Figure 3: Quantile plots of L_2 for designs D_1 (a 2^3-factorial design with no center runs), D_2 (a 2^3-factorial design with one center run), D_3 (a 2^3-factorial design with two center runs) and D_4 (a 2^3-factorial design with five center runs) (Example 1).
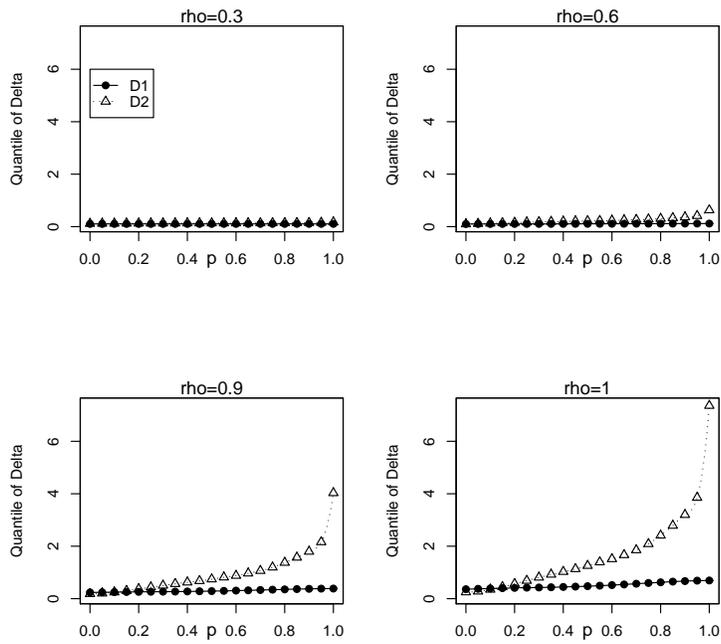
Figure 4: Quantile plots of Δ for designs D_1 (2^3-factorial design on a radius of 1 with no center runs) and D_2 (orthogonal simplex design with replicated points on a radius of 1) (Example 2).
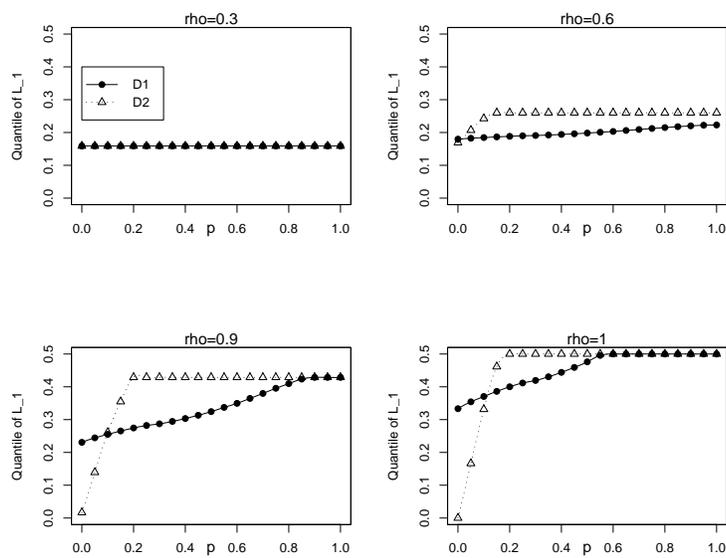
Figure 5: Quantile plots of L_1 for designs D_1 (2^3-factorial design on a radius of 1 with no center runs) and D_2 (orthogonal simplex design with replicated points on a radius of 1) (Example 2).
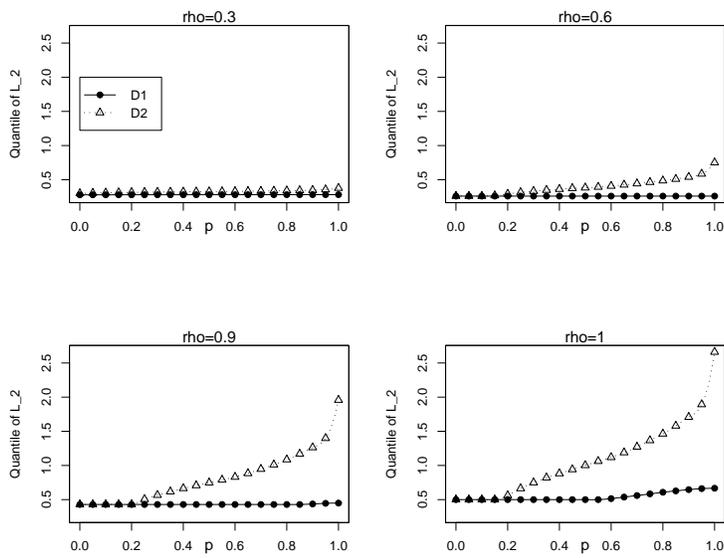
Figure 6: Quantile plots of L_2 for designs D_1 (2^3-factorial design on a radius of 1 with no center runs) and D_2 (orthogonal simplex design with replicated points on a radius of 1) (Example 2).
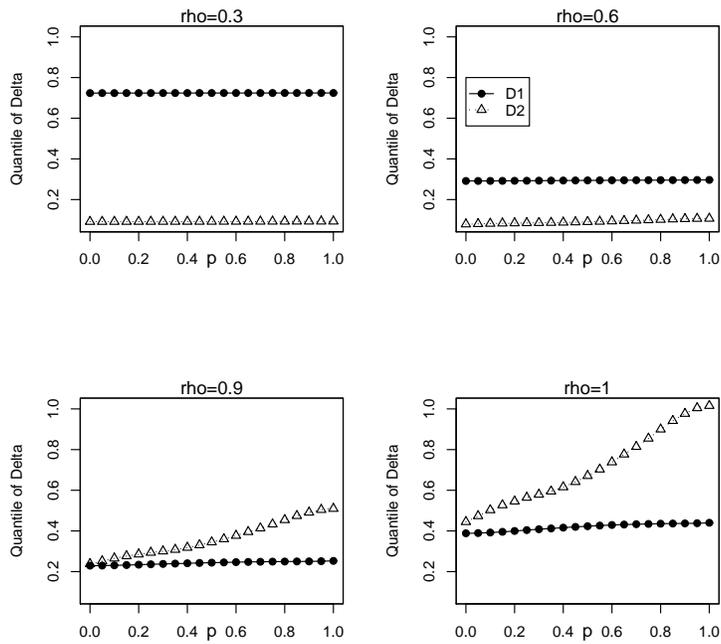
Figure 7: Quantile plots of Δ for designs D_1 (CCD) and D_2 (Box-Behnken) (Example 3).
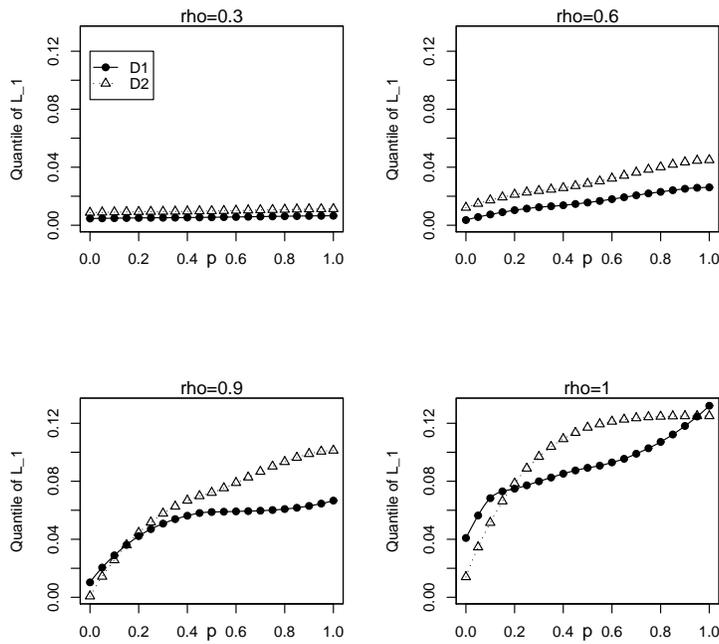
Figure 8: Quantile plots of L_1 for designs D_1 (CCD) and D_2 (Box-Behnken) (Example 3).
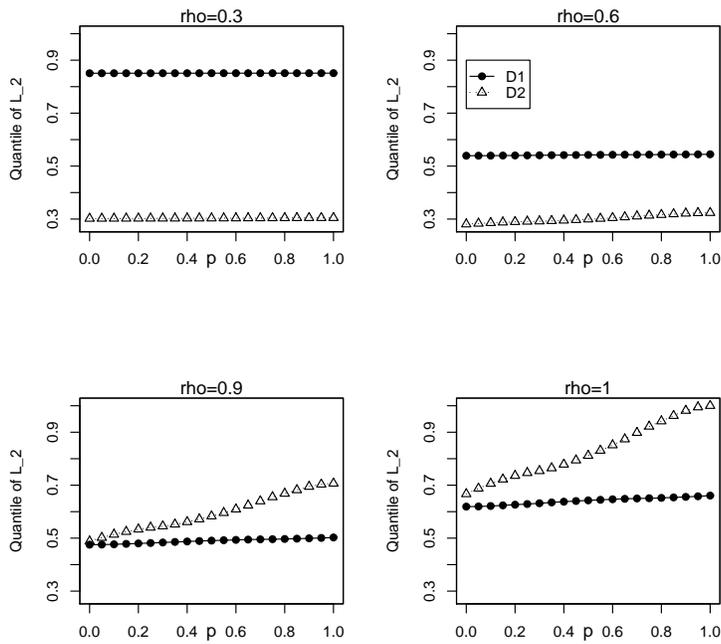
Figure 9: Quantile plots of L_2 for designs D_1 (CCD) and D_2 (Box-Behnken) (Example 3).
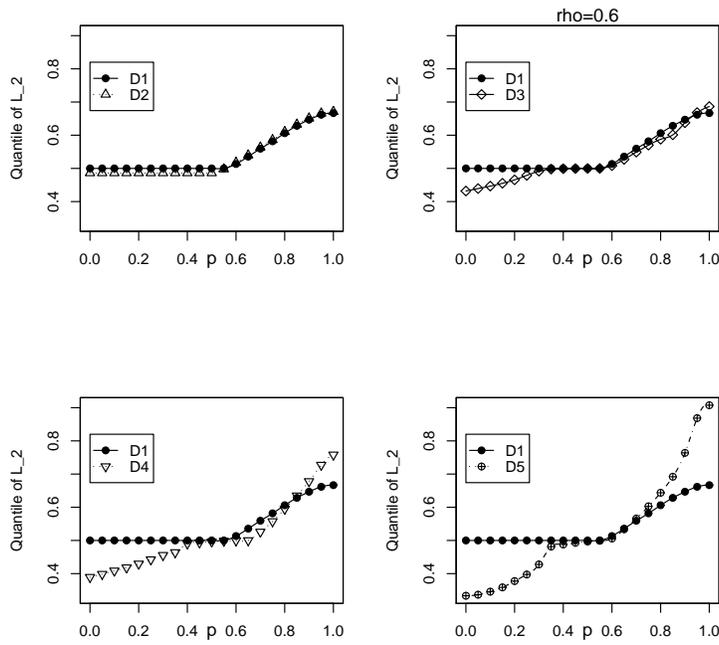
Figure 10: Quantile plots of L_2 for designs D_1 (2^3-factorial design with no center runs), D_2 (D_1 augmented with $(0, 0, 0)$), D_3 (D_1 augmented with $(0.5, 0, 0)$), D_4 (D_1 augmented with $(0.75, 0, 0)$) and D_5 (D_1 augmented with $(1, 0, 0)$) for $\rho = 1$ (Example 4).
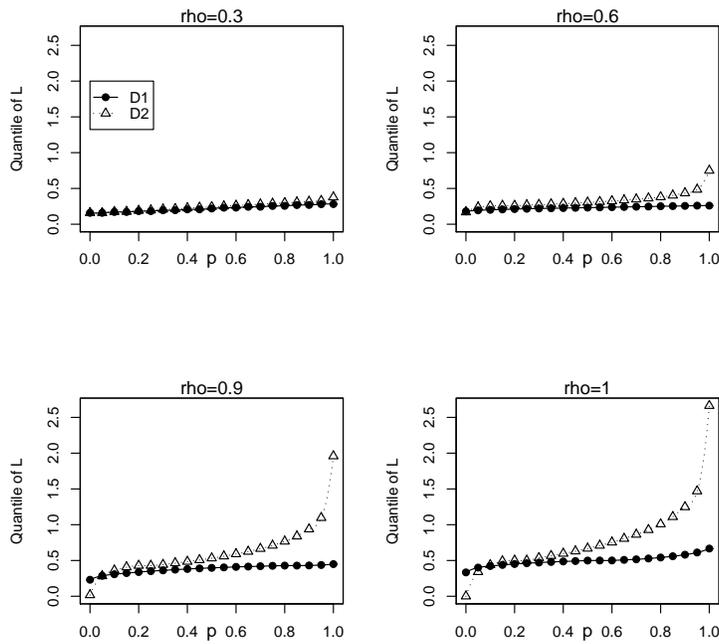
Figure 11: Quantile plots of L for designs D_1 (2^3-factorial design with no center runs) and D_2 (orthogonal simplex design) (Example 5).
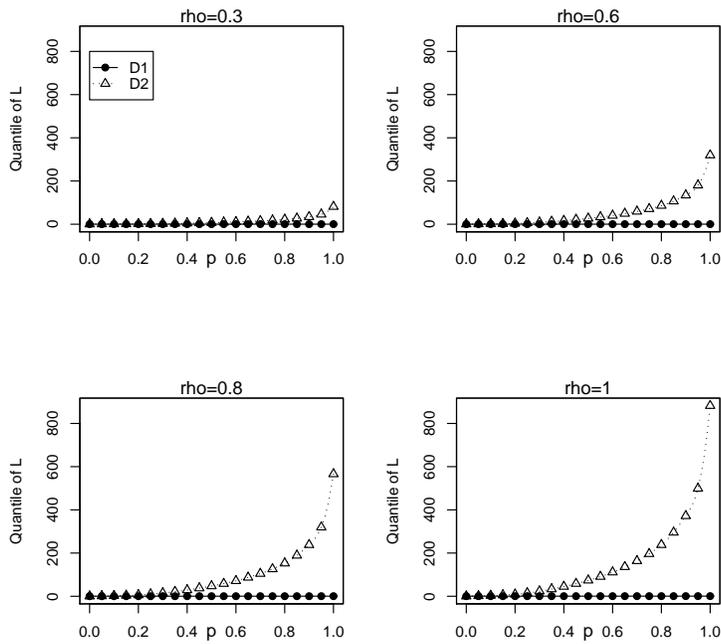
Figure 12: Quantile plots of L for D_1 (2^3 factorial design with no center runs) and the non-standard design D_2 (Example 6)
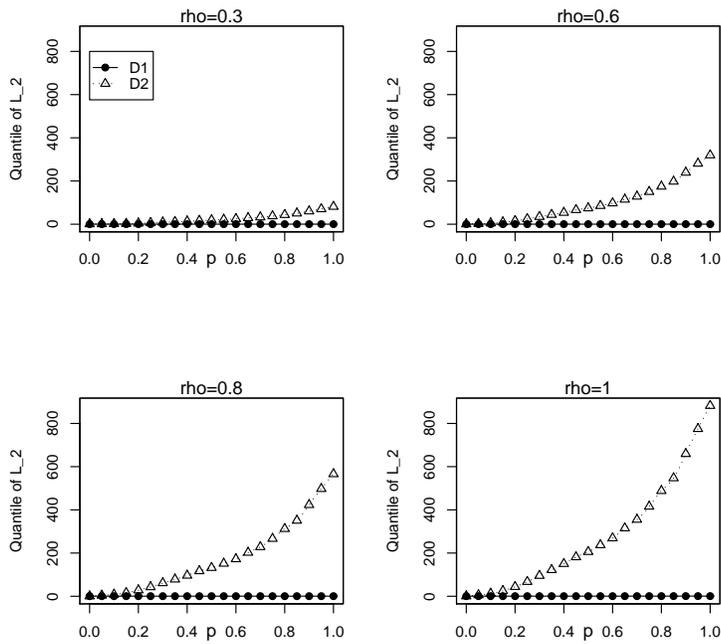
Figure 13: Quantile plots of L_2 for D_1 (2^3 factorial design with no center runs) and the non-standard design D_2 (Example 6)
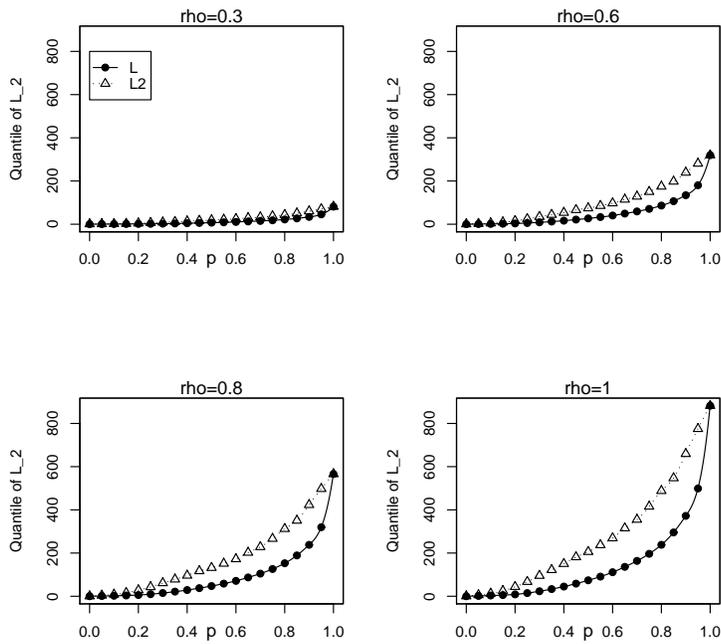
Figure 14: Quantile plots of L versus L_2 for D_2 (non-standard design) (Example 6)