Optimal Three-Level Designs for Response Surfaces in Spherical Experimental Regions

ROBERT W. MEE

University of Tennessee, Knoxville, TN 37996-0532

Since 1960, Box-Behnken designs have been very popular with experimenters wishing to estimate a second-order model in three or four factors. This popularity is due to these three-level designs' simplicity and high efficiency. However, as the number of factors increases, the run size of Box-Behnken designs increases rapidly, making them less attractive. The purpose of this article is to recommend the use of D-optimal and I-optimal three-level designs for spherical design regions involving three or more factors. Using an optimal design algorithm offers flexibility of run size. In addition, optimal design criteria permit construction of second-order designs for more complex response surface applications involving mixture or qualitative factors, or involving restrictions on randomization. The restriction to three-level designs provides great convenience and often little loss in efficiency, provided the design is not nearly saturated. Finally, when the number of factors is very large and the region is spherical, I argue that three-level designs are preferred to central composite designs.

KEY WORDS: Box-Behnken design; D-efficiency; I-efficiency; mixture design; second-order model; split-plot.

Dr. Mee is Professor in the Department of Statistics, Operations, and Management Science. He is a member of ASQ. His email address is rmee@utk.edu.

Introduction

Box and Behnken (1958, 1960) proposed the use of three-level designs for estimating second-order models in three or more factors. The hundreds of technical journal articles that have utilized these designs attest to their popularity. However, literature offers little guidance regarding adapting these designs for applications with additional mixture factors or qualitative variables. It is the purpose of this article to highlight the simplicity and flexibility offered by three-level designs for fitting second-order models in a wide variety of situations.

This article focuses on constructing three-level experimental designs for spherical regions. Lucas (1976, p. 411) summarizes the relevance of such regions as follows:

> "Spherical regions are considered because of the experimental fact that changes of a given maximum size can often be made in only a few variables, simultaneously, without running into operability problems… (E)xperimentation in the corners of the hypercube is impossible if the size of the hypercube is defined by the magnitude of the permissible change of one or a few variables at a time."

Instead of reducing the range of levels for each factor and exploring a smaller hypercube inscribed in a hypersphere, we assume a spherical design region which allows exploration safely with wider ranges for each factor.

The construction of these designs will be based on existing optimal design algorithms that utilize an initial candidate set provided by the user. *SAS*'s Proc Optex, *Design Expert*, and *Minitab* are all suitable for this task, as well as the specialized split plot optimal design algorithms by Goos and Vandebroek (2003, 2004). When constructing D-optimal designs for a spherical region, the candidate set should not be the full 3^t factorial, where t denotes the number of factors, but rather a particular subset (orbit) of points all the same distance from the center. A later section will show which candidate sets are best for 4 – 12 factors. First however, we review the optimal design literature and existing Box-Behnken designs. The article includes both standard and more complex design applications to illustrate the wide applicability of this approach.

The D-optimality criterion emphasizes precision of the estimated coefficients of the assumed model, and can seriously neglect other design characteristics of importance. I-optimality, the criterion that minimizes the average variance of prediction over a specified region of interest, provides (more) replication at the center. In addition to producing lower prediction variance over the center of the design region, this facilitates estimation of the error variance and testing for lack of fit. However, many three-level designs are rather ineffective for detecting lack of fit for second-order models. We address this concern briefly in a later section.

In spite of its title, Bisgaard (1997) is not a criticism of three-level designs per se. Rather, it criticizes the inefficiency of: i) using 3^t factorial and 3^{t-s} fractional factorial designs to estimate first-order and second-order models, and ii) conducting large experiments rather than a sequence of smaller experiments. For the purpose of screening, saturated first-order, three-level orthogonal designs exist for every t (see Mee 2002), whereas saturated, orthogonal two-level designs only exist for t = 3, 7, 11, … Thus, three-level designs provide fully-efficient designs to determine the direction of steepest ascent. Also, as this article shows, three-level designs for estimating a second-order model for a spherical region can be constructed for any number of factors t and run size n, provided n is greater than or equal to the number of parameters to be estimated.

This article explains how to effectively utilize optimal design algorithms to construct three-level, second-order designs for spherical experimental regions in a wide variety of applications. For readers less acquainted with optimal design, Carlyle, Montgomery, and Runger (2000, pp. 11-13) provide a concise summary of the formulation and numerical search for optimal designs, as well as weaknesses of these designs if the assumed model is not adequate. Snee (1975) gives practical advise about the use and modification of optimal designs to achieve a variety of different objectives. Atkinson (1996) documents the varied usefulness of optimal design criteria and algorithms for producing designs to fit prescribed models, especially when no standard designs exist. For a book-length treatment, see Atkinson and Donev (1992).

Insights from Optimal Design Literature

Kiefer (1960) showed that a D-optimal second-order design for t factors in a spherical region must reside entirely on the boundary of the hypersphere and at the center, with no other interior points. For additional details, including the further requirement of rotatability for D-optimality, see Appendix 1. Hardin and Sloane (1993, 2001) showed that designs of the same form were essentially I-optimal, with the only difference being the proportion of runs at the center. Let n_B and n_C denote the number of design points on the boundary and at the center, respectively, so that $n = n_B + n_C$, and let p denote the number of parameters for the second-order model, $p = (t+1)(t+2)/2$. Then D-optimality requires $n_C = n/p$, or equivalently $n_C = n_B/(p-1)$, whereas I-optimality requires $n_C \approx 2.06 n_B/(p-1)$. The close connection between D-optimality and

I-efficiency here suggests that I-efficient designs can be obtained by searching for a D-optimal design and then increasing the number of centerpoint replicates.

This article considers D-optimal and I-optimal designs for second-order designs in a spherical region with the additional condition that each factor assumes just three levels. In many applications, the added convenience of restricting the number of levels for each factor will far outweigh the small loss of variance efficiency that results from this restriction. Mitchell and Bayne (1978) investigated D-optimal three-level, second-order designs for hypercube regions. Results here differ – and are simpler - because of the shape of the experimental region. As in Mitchell and Bayne, it is helpful to partition the 3^t treatment combinations in the full factorial into orbits of various radii. An orbit contains all the points the same distance from the center. For example, for t = 2, there are two orbits, each with four points (see Figure 1). For t = 3, there are three orbits: the first orbit consists of six face-centered axial points; the second orbit consists of 12 edge-centered points; the third orbit consists of eight factorial corners. For general t, the binomial expansion $3^t = (1+2)^t = 1 + \sum_{k=1}^{t} \binom{t}{k} 2^k$ reveals that in addition to the one point at the center, there are t orbits, with $\binom{t}{k} 2^k$ points in the k^{th} orbit (k = 1, 2, …, t). For any k = 2,…, t-1, the points in the k^{th} orbit plus center point replicates form a second-order design that obviously satisfies the condition that all points reside on the boundary or the center, provided the factor levels are scaled so that the k^{th} orbit is on the design region boundary. As will be discussed in the next section, if $k = (t+2)/3$, such designs are also second-order rotatable. Except for small t, taking all the points in the kth orbit requires an excessively large n_B. The next section reviews how Box and Behnken (1960) proposed taking a subset of the $\binom{t}{k}2^k$ points in the kth orbit.

What about other three-level designs? Face-centered central composite designs are based on taking the first and t^{th} orbits (i.e., axial and factorial points), plus centerpoint replicates. These designs do not satisfy the conditions for D-optimality, and provide very poor precision for pure quadratic coefficients unless the face-centered axial points are replicated. Similarly, Morris (2000) proposed a series of three-level second-order designs that consist of an initial fraction of the 2^t, plus augmented pair points that reside on several interior orbits. While his designs are well-suited to sequential experimentation in a hypercube region, they are not efficient for estimation of parameters or prediction of the response for spherical regions. If the region of interest is spherical, three-level designs that include the factorial points of the t^{th} orbit are to be avoided.

Box-Behnken Designs

Box and Behnken (BB) (1958, 1960) proposed 17 three-level designs for estimating a second-order model in t factors. For 3 – 5 factors, the BB designs involve all $\binom{t}{2}2^2$ points from the second orbit. For six or more factors, each design contains just a subset of the k^{th} orbit, with $k > 2$:

- Box and Behnken used incomplete block designs (IBD) for blocks of size k to select a subset of the $\binom{t}{k}$ combinations. Whenever possible, BB used balanced IBD; in other cases, they used *regular graph* IBD - designs where all pairs of treatment levels occur together within a block for either λ or λ+1 blocks (John and Mitchell 1977).

- For BB designs with $k \geq 5$ (i.e., t = 11 and 16), only a half fraction of each set of 2^k runs was used.

For instance, for t = 11, the full 5th orbit contains $\binom{11}{5}2^5 = 462 \cdot 2^5 = 14,784$ points, while the corresponding BB design requires only $n_B = 11 \cdot 2^{5-1}$ = 176 points. Crossier (1991) added additional BB-type designs for t = 9 and 13 utilizing balanced IBD. Table 1 presents a summary of all BB and Crossier designs and their efficiencies. In every case, the D-optimal and I-optimal n_C's correspond to rounding $n_B/(p\text{-}1)$ and $2.06n_B/(p\text{-}1)$, respectively, to the nearest integer. Note the high efficiencies for these designs. This implies that for these design sizes, the restriction to three-level designs has not substantially impacted either D- or I-efficiency.

The BB designs based on regular graph IBD do have lower efficiencies than the designs based on balanced IBD because they provide uneven (and, on average, lower) precision for the second-order coefficient estimates. For BB designs based on balanced IBD, the design is rotatable if and only if $k = (t+2)/3$. (Appendix 2 contains a proof.) Thus, the BB designs based on balanced IBD with (t=4, k=2), (t=7, k=3), (t=10, k=4), and (t=16, k=6) are perfect designs (in a sense defined in Appendix 1), although the designs for t = 10 and 16 require too many runs to be very useful.

While no perfect three-level design exists when $(t+2)/3$ is not an integer, using a balanced IBD with $k \approx (t+2)/3$ to determine the subset of boundary points always produces a BB-type design with very high efficiency. The main drawback of these designs is the inflexibility of the design size. In the next section, we investigate D-optimal fractions of the 3^t and find empirically that the optimal design always consists of a subset of a single orbit.

D-Optimal Three-Level Designs

We now explore optimal three level designs in a sphere for situations where no perfect designs exist. The cases with 4, 5, and 6 factors are examined in detail. Lessons drawn from these cases are subsequently applied to larger t. Applications with $t = 10$ and 14 factors are discussed at the end of this section.

D-Optimal Designs for t = 4 Factors

The 3^4 full factorial consists of 81 points. Let x_1, x_2, x_3, x_4 denote the four factors. If the levels are centered about zero, a design point's distance from the center is given by $d = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$. Typically, the levels are coded (-1, 0, 1). Using this coding, the distance from the design center is shown for each of the 81 points of the full factorial design in Figure 2. We see that the points are distributed at the center and on four orbits at distances 1, √2, √3, and √4 (=2), with 8, 24, 32, and 16 points on each, respectively. To identify a D-optimal three-level design with a given number of trials, n, one starts with a (larger) set of candidate points from which the optimal subset is selected. For example, we could use all points from orbits 1 and 2 and the center point as the set of candidate points; this would give a candidate set of 33 points. This is one of three alternative candidate sets listed below. To compare "optimal" designs found from different candidate sets, it is necessary to rescale the levels so that the points lie inside or on the boundary of the assumed spherical design region. If we take the design region to be a hypersphere with radius 1, then the 33 point candidate set consisting of points from the center, orbit 1, and orbit 2 are scaled so that the low, medium, and high levels are $-\sqrt{1/2}$, 0, and $\sqrt{1/2}$, respectively. In each case, the scaling assures that points from the outermost orbit in each candidate set lie on the unit sphere and the remaining points all lie inside the unit sphere. The three candidate sets we consider for $t = 4$ factors are:

1. Orbits 2 and 1, plus the center (33 points). The levels are coded 0 and $\pm\sqrt{1/2}$.
2. Orbits 3, 2, and 1, plus the center (65 points). The levels are coded 0 and $\pm\sqrt{1/3}$.
3. The full 3^4 (81 points). The levels are coded 0 and $\pm\sqrt{1/4}$, so that the fourth orbit lies on the unit sphere and all other points are interior.

For a given candidate set and design run size n, one may use the Fedorov method as implemented in SAS's (2004) Proc Optex, or some similar algorithm, to find the best subset. For instance, with the 33 point candidate set and $n = 26$, output for the best design appears in Table 2. The last column indicates the frequency for each candidate set point in the optimal design of size 26 runs. Note that for the optimal design, 24 points (the entire second orbit) each appear once, and the center point has frequency 2.

In each case (here and later), for each candidate set and run size n, the most D-efficient subset was identified using Proc Optex with Fedorov's method and 100 or more iterations. Figure 3 shows the D-efficiency for $n = 15$ (the saturated design) up to $n = 34$ for each candidate set. These D-efficiencies were calculated relative to a theoretical upper bound for second-order model matrices which is not usually achievable for a specific n (for details, see Appendix 1).

For the 33-point candidate set and $n = 26$, the D-optimal design is the BB design with $n_C = 2$ and D-efficiency = 99.92% (refer to Table 2). For the same candidate set and

$n < 26$, the D-optimal design is a subset of the BB design, while for $n > 26$, the design contains duplicates of the BB points. That is, for candidate set 1, the optimal design only includes points from the second orbit and the center. Though not shown in Figure 3, from $n = 35$ to 51 the D-efficiency increases; at $n = 51$ the D-optimal design is a full replicate of the second orbit plus 3 center runs.

For the candidate set 2 with 65 points, the highest D-efficiency occurs at $n = 34$; this design consists of the entire 3rd orbit plus $n_C = 2$. Each smaller design selected from this candidate set is a subset of this design; that is, no design included points from interior orbits. For $n > 34$, the D-efficiencies initially decline as 3rd orbit points are duplicated.

For the full 3^4 candidate set, designs necessarily include points from multiple orbits, since the outer orbit contains only two levels for each factor. As evident from Figure 3, all designs obtained from this candidate set have inferior D-efficiency. In summary, the D-optimal three-level designs for $n < 26$ are subsets of the BB design, and for $n > 26$ involve duplication the same points.

D-Optimal Designs for $t = 5$ Factors

The 243 points in the full 3^5 consist of one at the center, plus 10, 40, 80, 80, and 32 points in orbits 1 – 5, respectively. The BB design utilizes the center point and the entire second orbit, and is D-optimal for a three-level design of size 42. Constructing a BB-type design using a balanced IBD with $k = 3$ or $k = 4$ would double the design size, requiring the entire orbit of 80 points in each case. Furthermore, since $k = 2$ is the closest to $(t+2)/3 = 2.33$, Box and Behnken's choice is closer to being rotatable and so has higher efficiency.

Figure 4 displays results from Proc Optex using the following three candidate sets:

1. Orbits 2 and 1, plus the center (51 points). Levels are 0 and $\pm\sqrt{1/2}$.
2. Orbits 3, 2, and 1, plus the center (131 points). Levels are 0 and $\pm\sqrt{1/3}$.
3. Orbits 4, 3, 1 and 1, plus the center (211 points). Levels are 0 and ± 0.5.

Taking the entire 2^5 as a candidate set is not reported here, since it fairs so poorly. For candidate sets 1-3, the only design to contain points from multiple orbits was the saturated design selected from candidate set 3. Every other design with efficiencies displayed in Figure 4 contains no interior points other than the center. Furthermore, for $31 < n < 42$, the D-optimal fraction from the 3^5 factorial is a subset of the Box-Behnken design. For $n \leq 31$, the flexibility offered by selecting from the 80 points in the third orbit results in a better design. However, for $n > 31$, the D-optimal design reverts to the $k = 2$ orbit.

D-Optimal Designs for $t = 6$ Factors

Box and Behnken (1960) propose two six-factor designs with $k = 3$, a large design based on a balanced IBD and a smaller design based on a regular graph IBD. While a second-order design could be based on the center and the k^{th} orbit for $k = 2, 3, 4$, or 5, $k = 3$ is closest to $(t+2)/3$ and best over the entire range of n explored. Figure 5 displays D-efficiency results for designs built from three candidate sets:

1. Orbits 2 and 1, plus the center (73 points). Levels are 0 and $\pm\sqrt{1/2}$.
2. Orbits 3, 2, and 1, plus the center (233 points). Levels are 0 and $\pm\sqrt{1/3}$.
3. Orbits 4, 3, 2 and 1, plus the center (473 points). Levels are 0 and ± 0.5.

Including points from the 5^{th} (or 6^{th}) orbit results in inferior designs which we do not bother to display. No design represented in Figure 5 included interior points other than the center.

The D-optimal three-level design is always a subset of the 3^{rd} orbit. For n between 63 and 83, the D-efficiencies generally increase, although not monotonically, until they reach the maximum of 99.5%, the D-efficiency for the BB design with 83 runs.

D-Optimal Designs for $t \geq 7$ Factors

Since, in every case considered thus far, the most D-efficient three-level design obtained by Proc Optex has included only boundary points and the center, subsequent searches will be confined to candidate sets restricted to the center point and a single orbit. This restriction will generally speed the search algorithm. In addition, provided the optimal design in fact includes no interior orbit points, reducing the candidate set increases the likelihood of actually finding the optimum subset within a specified time.

For t = 7, the best saturated design found had D-efficiency = 87.78%; this design took 35 points from the third orbit, plus the center point. (This result was far better than for any other orbit.) As run size is increased, D-efficiencies increase until at n = 58, the rotatable BB design is obtained (refer to Figure 6). Since this BB design does not contain the entire third orbit, but only 7 of the $\binom{7}{3} = 35$ combinations of subsets of three factors, the optimal designs for $n < 58$ are not generally a subset of the BB design with $n_B = 7(2^3)$; the optimal designs for other run sizes utilize points from more than seven of the 35 subsets of the factors.

For t = 8, Box and Behnken (1958) suggested two designs that are 4-5 times larger than the saturated design. The smaller of the two designs contains n_B = 192 points obtained using a regular graph IBD with 24 blocks of size k = 3, since no BIB design exists with fewer than $\binom{8}{3}$ blocks. A candidate set consisting of the third orbit and the center was found to be best over the range $45 \leq n \leq 90$. Figure 6 shows D-efficiencies for the designs Proc Optex found. While more extensive searching may find designs with slightly higher efficiencies than those displayed in the figure, it is apparent that the best candidate set consists of the third orbit plus the center. Finally, the BB design with n_B = 192 (D-efficiency = 97.9%) is not D-optimal, since Proc Optex found a design of that size with D-efficiency above 99.3%.

For t = 9, the best designs come from the third orbit. This is somewhat surprising, since k = 4 is nearest to $(t+2)/3$, but also a welcome outcome, since the 3rd orbit with 672 points is substantially smaller than the 4th orbit with 2016 points. Results in Figure 6 for n = 55, 60, …, 100 show an increase in D-efficiencies as design sizes approach the D-optimal n = 98 design from Crossier (1991), which is based on a balanced IBD with 12 blocks of size 3.

For t = 10, designs from the third orbit were again superior to designs from the fourth orbit for n over the range of $p = 66 \leq n \leq 132 = 2p$, i.e., for designs up to twice the size of the saturated design. For n = 140 similar efficiencies were obtained for k = 3 and 4, so that k = 4 is expected to be optimal for most $n > 140$.

Even for t = 11, k = 3 is preferred for $n \leq 110$ rather than k = 4. For n between 120 and 150, there is little difference in D-efficiencies for optimal subsets from the third and fourth orbits. For $n > 150$, use $k = 4$, or even $k = 5$ for n near 178, the size of the BB design, although adequately searching such large candidate sets will require more time.

For $t = 12$, the fourth orbit produces more efficient designs than the third orbit over the range of sample sizes $91 \le n \le 160$. D-efficiencies using $k = 4$ ranged from 72.2% for the saturated design to 92.5% at $n = 160$.

I have not investigated constructing optimal subsets for larger t, although it seems reasonable to recommend using the fourth orbit for $t = 13$, 14, and 15 factors. Apart from the complication of an increasingly large candidate set, this simple procedure ensures that one can easily find highly D-efficient designs. When there are many factors, SAS (2004) recommends restricting the candidate set for Proc Optex, and illustrates this suggestion with the case of a 3^{15}. While their suggestion of using a 3^{15-11} orthogonal array as the candidate set may be suitable for a hypercube region, we know that choosing the $\binom{15}{3}2^3$ $= 3640$ points of the third orbit or $\binom{15}{4}2^4 = 21{,}840$ points of the fourth orbit is better for second-order designs in spherical regions. In the later case, it seems reasonable that the candidate set might be trimmed further without limiting the D-efficiencies of resulting designs. How best to do so has not been explored, since this only becomes necessary for $t > 11$.

The largest application personally discussed by the author involved 14 factors for engineers at Ford Motor Company investigating steering wheel "nibble"; see Thomas, Soderborg, and Borders (2005) for an earlier stage of this investigation. For 14 factors, $k = 4$ is recommended. While a saturated design would require only 120 computer runs, 150 – 180 runs would ensure better D-efficiency and also provide some ability to check for the adequacy of the second-order model, a topic addressed in the next section. The largest case seen in the literature involved 32 factors in a sensitivity analysis for simulation of upper percentiles of exposure to a particular toxin (Albert and Gauchi 2002). Those researchers used a resolution V 3^{32-24} (= 6561 run) fraction to estimate the p = 561 parameters of a second-order model. Given the model they intended to fit, a substantially smaller design with perhaps 700 – 800 runs would seem sufficient.

A 10-Factor Example Based on a Deterministic Model

While six factors may be a large number for a physical experiment, computer model experiments often involve 10 or more factors. When the number of factors is very large, a central composite design (CCD) is may be poorly suited for spherical regions, since either the location of the axial points require extreme levels for each factor, or the axial points are interior to the design region and so provide poorer efficiency. Srinivasaiah and Bhat (2005) discuss the results of a screening experiment in which they identified 10 of 21 process parameters as important to the performance of integrated circuits, as represented by a deterministic computer model. They mention as ongoing work a 1045-run second-order design in 10 parameters. Apparently, they have in mind a 2^{10} full factorial plus 20 axial points and a center point. Using a 2_V^{10-3} fraction in their CCD rather than the full factorial would reduce the design size to 149. Now what axial point levels ($\pm\alpha$) would be appropriate? For a hypercube region, α = 1. If this experimental design were being used to investigate a stochastic computer model, one could replicate the axial points to improve the precision of the pure quadratic coefficients.

However, replication is useless for deterministic computer models. For a spherical region, $\alpha = \sqrt{10} \doteq 3.16$ would place the axial points on the same sphere as the factorial points. For physical experiments such extreme spacing in each factor can result in failed runs or even dangerous operating conditions, unless the actual low and high levels in the factorial points were chosen with the much larger range for the axial points in mind. Furthermore, the loss of data at axial points has drastic consequences for the ability to estimate pure quadratic coefficients. For computer model experiments, there is no danger in collecting data at treatment combinations with extreme levels; it is simply a matter of whether the second-order model holds for extreme changes in the factors.

Now consider instead a three-level design for the same size experimental region. There exists a perfect design with n_B = 240 and k = 4 (refer to Table 1). However, for designs up to twice the size of a saturated design ($n \leq 132$), taking points from the 3rd orbit is preferred. In Figure 6, we see that a design with t = 10, k = 3 and n = 122 has very high efficiency (D-efficiency = 97.1%). Upon inspection, one finds that this design has recognizable structure, being half of a usual BB-type design. That is, balanced IBD exist with t = 10 and 30 blocks of size k = 3 (see, e.g., the design in Cochran and Cox 1957, p. 475). A BB design constructed from this balanced IBD would have $n_B = 30(2^3)$ = 240, while the smaller D-optimal design has $n_B = 30(2^{3-1}) = 120$. Again, one would use only one centerpoint run, since the model is deterministic.

Now we compare the spherical CCD with the D-optimal three-level design. Taking the design region as the unit sphere, the five levels of the CCD with α = 3.16 are rescaled to (-1, -.316, 0, .316, 1). Similarly, the three-level design taken from the third orbit must be coded with levels (-.577, 0, .577). With this coding, the CCD points and the D-optimal three-level design points (not at the center) are all a distance 1 from the center. The equal spacing for three-level designs with many runs at each level is surely more robust for estimating pure quadratic terms than is the CCD with only one run at each extreme level.

Bias from Third Order Terms and Checking Lack of Fit

If a design is not saturated, then the error degrees of freedom can be used to investigate the adequacy of the model, and even perhaps to estimate some additional terms. If we have replication, then a formal lack of fit test may be performed - although the D-optimal allocation of $n_C = n/p$ generally furnishes too little replication to be useful. As mentioned in the introduction, I-optimal designs are better in this regard. So what types of higher-order terms can be detected? Checking for lack of fit of a second-order model is usually based on considering the need for third order terms, i.e., x_i^3, $x_i^2 x_j$, and $x_i x_j x_l$ terms. Three-level designs have no ability to detect pure cubic terms; assuming the levels are equally spaced, pure cubic effects are fully aliased with linear effects. For example, with (-1,0,1) coding, $x_i^3 = x_i$. [We utilize (-1,0,1) coding in this section to make such correlations more obvious.] The situation is better for third order terms involving interactions. The column $x_i^2 x_j$ will be positively correlated with x_j, but the presence of runs for which $x_i^2 x_j \neq x_j$ provide some ability to detect such higher-order interactions. That is, if a quadratic-by-linear interaction term exists but is omitted from our model, it will bias the linear effect estimate for the j^{th} factor; however, by including the column $x_i^2 x_j$ for various (i,j), evidence for such terms may be considered. (In fact, the D-optimal designs presented here are better than face-centered CCD's for detecting the need for such terms, since for CCDs, $x_i^2 x_j = x_j$ except for x_j's axial points.] Finally, a three-level design constructed from the k^{th} orbit for $k \geq 3$ can detect the presence of some three factor interactions. However, for designs constructed only from the second orbit, $x_i x_j x_l = 0$ for all (*i,j,k*), and so this lack-of-fit cannot be detected. We now illustrate how one may check for lack of fit using a three-level design.

Srinivas, Rao, Théodore, and Panda (1995) report an experiment based on a six-factor BB design with $n_B = 48$ and $n_C = 6$. Six centerpoint replicates is more than either D-optimality or I-optimality require, but this is beneficial for the lack-of-fit test. The experiment involved optimizing a fermentation process for producing ethanol. The primary design involved six medium constituents studied in the 54 runs of a BB design. For each of these 54 runs, the response was the maximum observed ethanol production achieved in a small experiment in pH and temperature.

There appears to be errors in six treatment combinations – rows 3, 19, 23, 26, 33, and 43 of Srinivas et al.'s Table 1. Eliminating these six rows and the first centerpoint run (row 49), which is a distinct outlier, an unbalanced design with $n_B = 42$ and $n_C = 5$ remains. Fitting a second-order model with these 47 runs, we observe significant lack of fit (see Table 3).   Augmenting a second-order model using forward selection, considering eligible all third order terms that are estimable from this design, five terms were added: $x_1 x_4 x_5$, $x_4^2 x_6$, $x_1 x_5^2$, $x_1^2 x_5$, and $x_3^2 x_4$. This model has R^2 of 95.3% and lack-of-fit F = 3.48 (p-value > .05). Due to aliasing of the four included $x_i^2 x_j$ terms with similar terms that are not included, in addition to the inestimability of most other third-order terms, there is considerable ambiguity about the correct model. Even so, this design reveals that the second-order model is deficient and that several third-order terms of the form $x_i x_j^2$ would improve the model. Clearly, augmenting the design is necessary.

In cases such as this with serious lack-of-fit, one may need a third-order design. D-optimal third-order designs consist of points on two spheres, with the inner sphere having a radius of 50% - 60% of the radius of the boundary sphere (Farrell, Kiefer, and Walbran, 1967, pp. 134-6)). How best to augment a BB design or a CCD is an open question.

Loss of Efficiency Due to Restriction to Three-Level Designs

Although three level designs are popular, if one considers designs with more levels, higher variance efficiencies can often be achieved, especially for nearly saturated designs. To assess the loss of efficiency from this restriction, it is necessary to search for D-optimal and I-optimal designs within the sphere, and then compare efficiencies of three-level designs and many-level designs. JMP 6.0 software provides a very efficient search for D-optimal and I-optimal designs in a sphere without having to specify a candidate set. Table 4 summarizes the results.

The efficiencies displayed in Table 4 were computed as follows. As in Appendix 1, let $\boldsymbol{D}$ denote an $n \times t$ design matrix for a second order design constrained to points on or inside the unit sphere, and let $\boldsymbol{X}$ denote the corresponding $n \times p$ second-order model matrix. The theoretical upper bound for the determinant of the matrix $(\boldsymbol{X'X})/n$ is denoted D_∞ in Appendix 1. For $t = 4$, this upper bound is 7.5043e-017. The best determinant for three-level designs with $n = 15$ runs had $|(\boldsymbol{X'X})/n|$ = 2.2829e-018, and, as displayed in Figure 3,

$$\text{D-efficiency} = [2.2829\text{e-}018 / 7.5043\text{e-}017]^{1/15} = .7923.$$

This low D-efficiency does not imply that there exists a 15-run design that is 100% efficient. In fact, the best 15-run spherical design JMP found has $|(\boldsymbol{X'X})/n|$ = 1.4933e-017, and efficiency $[4.8031\text{e-}017 / 7.5043\text{e-}017]^{1/15} = .9707$. Thus for this saturated design, the three-level restriction produces loss of D-efficiency of about 18%, since .7923/.9707 = .82. A saturated design with $t = 4$ is perhaps the worst case for a three-level design. Table 4 shows a pattern that the loss of efficiency diminishes as the number of factors increase, and as the extra runs n-p increase, For example, at $t = 7$ and $n = 41$ (= $p + 5$), the loss of efficiency is only 2%.

The I-efficiency of three-level designs relative to the best design from JMP is displayed in Table 5. The three-level designs upon which Table 5 is based were obtained by adding one run to each of the D-optimal, three-level designs represented in Table 4 [except for the case of ($t = 4$, $n = 21$), where ($n_B = 18$, $n_c = 3$) is I-optimal; there we added two runs to the 19-run D-optimal design]. This was done because we have no convenient manner to search for spherical I-optimal designs from a candidate set. These I-efficient designs are compared with spherical I-optimal second-order designs found using 1000 starts with JMP 6.0's custom design feature. Although the I-efficiencies in Table 5 are much smaller than the D-efficiencies in Table 4, the overall pattern is the same. The loss of efficiency is greatest (60%) for the smallest saturated design ($t = 4$, $n = 16$), but only 14% for the largest design ($t = 7$, $n = 42$) considered in Table 5. The efficiencies for truly I-optimal three-level designs might be better than Table 5 indicates, since we have relied on a D-optimal search procedure to obtain the boundary points for these three-level designs. Similarly, the efficiencies for the many-level designs might also be understated slightly, since further search might achieve better results than we found with JMP.

The conclusion from Tables 4 and 5 is that saturated designs suffer the most inefficiency due to imposing the restriction of using three-level factors. Further, three-level designs with six or seven factors fare better than those with five or less.

Examples of Three-Level Designs for More Complex Applications

Optimal design algorithms provide a practical means for design construction in non-standard settings. The restriction to three-level factors adds to the simplicity of implementing the design, often without seriously impacting the efficiency. In this section, three applications are presented. The first is taken from Vuchkov, Damgaliev and Yontchev (1981) and involves three quantitative factors plus three mixture factors. The second example involves both qualitative and quantitative factors. The final example from Trinca and Gilmour (2001) involves a restriction on randomization, for which the restriction to three-level factors facilitates construction of a split-plot design.

Example with Process and Mixture Factors

Vuchkov, Damgaliev and Yontchev's (1981) rubber compound example involved selecting a 30-run second-order design for three types of rubber (x_1, x_2, x_3) and three other ingredients (x_4 - *oil,* x_5 - resin, x_6 - accelerator). The first three are mixture factors, in that the experiment involved different rubber blends, while the other ingredients could be set independently of the other factor levels. The constraint $x_1 + x_2 + x_3 = 1$ produces numerous singularities for the usual second-order model. One means of avoiding these singularities is to write the model in terms of only five factors, say $x_2, \ldots, x_6$, since the level of $x_1 = 1 - x_2 - x_3$. Instead, Vuchkov et al. (1981) avoid these singularities by excluding the intercept, pure quadratic terms for the mixture factors and linear terms for the other factors. The resulting (Scheffe canonical) form for the second order model is:

$$E(y) = \sum_{i=1}^{3} \beta_i x_i + \sum_{i=4}^{6} \beta_{ii} x_i^2 + \sum_{i=1}^{5} \sum_{j=i+1}^{6} \beta_{ij} x_i x_j$$

[see equation (6) of Kowalski, Cornell, and Vining (2000, p. 2260)].

Vuchkov et al.'s (1981) candidate set contained the product of the six point second-order simplex lattice design {(1,0,0), (0,1,0), (0,0,1), (.5,.5,0), (.5,0,.5), (0,.5,.5)} for the rubber factors and a 3^3 factorial for the other ingredients. I revised the candidate set by excluding the third orbit of the 3^3 and re-scaling so that the second orbit matches the unit sphere. This candidate set has 6 x 19 = 114 treatment combinations. The D-optimal 30-run, three-level design with $x_4^2 + x_5^2 + x_6^2 \le 1$ appears in Table 6 and consists of:

- Mixture factor combinations: 19 pure blends and 11 50/50 blends, which is similar to the 20 pure and 10 binary blends for Vuchkov et al.'s sequentially determined design
- Process factor combinations: 28 points from the second orbit and two center point replicates. As expected, no points from the first orbit appeared.

Suppose instead that one uses the same ($6 \times 27 = 162$ run) candidate set as Vuchkov et al. (1981), with the process factors scaled to have levels $\pm\sqrt{1/3}$ and 0. The best design Proc Optex finds has D-efficiency:

- 3% higher than their sequentially determined design. As expected, it is slightly better to select a 30-run optimal design than it is to use the sequentially determined design of size 30 from the Vuchkov et al. (1981) tables.
- 21% lower than the D-optimal three-level design in Table 6.

Thus, for a spherical region, excluding the third orbit produces a much more efficient design. This D-optimal three-level design is more than 99% efficient compared to the D-optimal design found using *JMP 6.0* constrained by $x_4^2 + x_5^2 + x_6^2 \leq 1$ but not restricted to three levels.

Example with Process and Qualitative Factors

Several papers have been written on constructing response surface designs when some qualitative factors must be included; see, e.g., Draper and John (1988) and Wu and Ding (1998). Optimal design construction is particularly useful here, because the models to be estimated depend not only on the numbers of quantitative and qualitative factors, but also on the types of interaction terms included. Myers and Montgomery's (1995, p. 434) example 9.7 involves construction of a 24-run design with four quantitative factors (x_1, x_2, x_3, x_4) and one two-level qualitative factor (x_5). Their second-order model is

$$E(y) = \beta_0 + \sum_{i=1}^{5} \beta_i x_i + \sum_{i=1}^{4} \sum_{j=i+1}^{5} \beta_{ij} x_i x_j + \sum_{i=1}^{4} \beta_{ii} x_i^2$$

with 20 terms, and their candidate set has 50 points – the rotatable CCD at each level of z. There are several 24-run designs with equivalent D-efficiencies. Each involves omitting one point from the CCD, with 8 points at one level for x_5 and 16 points at the other level.

For an optimal three-level design, we exclude the third and fourth orbits, and obtain a design with the same efficiency as the fraction of the CCD found by Myers and Montgomery. This occurs because the 25-run CCD and BB design are equivalent for t = 4; with n = 24, any one point can be deleted. As with the CCD, this D-optimal design splits the points, eight to one level of x_5 and 16 to the other.

If higher-order terms are included, three-level designs may be a poor choice. Myers and Montgomery's (1995, p.437) example 9.8 with t = 3, plus two 2-level qualitative factors (z_1, z_2) involves a model that is not well suited for three-level factors. Due to the inclusion of six third-order terms ($x_i^2 z_j$), three-level designs are inefficient.

Split Plot Example

Trinca and Gilmour (2001) consider response surface designs with split plot unit structure. Their protein extraction example involves five three-level factors, with one factor (feed position) held constant each day and four factors varied within a day. Their proposed 42-run split plot design is a face-centered CCD with 16 factorial runs, 20 axial points, and six center point replicates, partitioned into 21 whole units of size two. Goos and Vandebroek (2003) assume a hypercube experimental region as well, but utilize a 3^5 candidate set to construct a D-optimal split-plot design under the same run size specification.

The D-optimal split plot design Goos and Vandebroek (2003, Table 1) find for the hypercube region has points from all five orbits. They proceed to state (p. 240) that the 3^5 candidate set is not sufficient for a spherical design region and recommend the addition of axial points to the candidate set. We now know that one can retain a three-level design and accommodate the spherical region by reducing the candidate set, excluding the outer orbits of the 3^5 from the candidate set. In this way we achieve a candidate set suited to the experimental region.

Peter Goos searched for the D-optimal three-level design with $n = 42$ and whole plots of size 2, using the same candidate sets as in the previous section for $t = 5$ factors. The design in Table 7 was found to be D-optimal for a wide range of variance ratios. In particular, Goos and Vandebroek's (2003) algorithm returns this design (or an equivalent arrangement) for whole-plot-to-split-plot variance ratios of 0.5, 1, and 2. The design in Table 7 is the Box-Behnken design for $t = 5$ factors arranged into 21 whole plots. Assuming a whole plot error variance component equal to the split plot component (so that the correlation within each whole plot is .5), this design is much more efficient that the CCD in Trinca and Gilmour (2001). For instance, Trinca and Gilmour's design, modified to be a spherical design by taking $\alpha = \sqrt{5}$, is only 65.6% efficient compared to the BB design in Table 7. If one chooses $\alpha < \sqrt{5}$ for the CCD, its relative efficiency would be even lower.

Several papers have appeared which consider BB designs run as split plot experiments. The comments by Letsinger, Myers and Lentner (1996) and Goos and Vandebroek (2001) regarding BB designs as split plot designs assume that all runs with the same whole plot factor treatment combination appear together in one whole plot. Vining, Kowalski and Montgomery (2005) modify BB designs with three and four factors by adding runs to make them more suitable as split plot designs; adding replication of runs and increasing the number of whole plots facilitates estimation of variance components. Draper and John (1998, p. 493) noted that the attractiveness of BB designs for split plot experiments diminishes if there are more whole plot factors. However, they concur that the restriction to three-levels for the whole plot factors is attractive. Thus, while some BB designs might not adapt well to some split-plot applications, optimal three-level designs obtained using the Goos and Vandebroek algorithms and a candidate set limited to the center point and a single orbit of the 3^t provides an excellent general solution.

Conclusion

We have learned that D-efficient three-level designs for estimating second-order models with run size less than twice the number of parameters are obtained from points on a single orbit:

- For three or four factors, use the second orbit
- For 6-10 factors, use the third orbit
- For 12-15 factors, use the fourth orbit

For the in-between case of five factors (11 factors), the choice between the 2nd and 3rd orbits (3rd and 4th orbits) depends on the sample size. These suggestions are generally applicable even for more complicated applications that involve additional qualitative factors or mixture variables. Furthermore, since the loss of efficiency can be considerable for nearly saturated designs, a run sizes of 20% or more than the saturated design (i.e., $n \geq 1.2p$) are recommended.

The choice of low and high level for each three-level factor should depend on which orbit is used to construct the design. For example, if the third orbit is used, then the low and high levels should be chosen with the understanding that only three of the factors are to be set to their high or low level at a time, with the remaining factors all set to their middle level. This is true whether the experiment involves five factors or ten factors. This is somewhat simpler than choosing levels for a spherical CCD, since there one must take account of both the levels for the factorial points and the corresponding axial point levels ($\alpha \approx \sqrt{t}$) required to put these points near the same boundary sphere.

Due to computational convenience, optimal designs obtained here were based on D-efficiency rather than I-efficiency. While perfect designs on the hypersphere can be made either D-optimal or I-optimal, depending on the number of centerpoint replicates, it is not proven how closely I-optimal and D-optimal three-level designs coincide when the best designs are not rotatable. This is a topic for further research. In the absence of a candidate set algorithm to compute I-optimal three-level designs in a sphere, compute a D-optimal design for a slightly smaller size, and then increase the number of centerpoint runs. For instance, for 6 factors and 50 runs, the D-optimal design will have two centerpoint replicates, while the I-optimal design will have four. Thus, to construct a 50-run I-efficient design, compute a 48-run D-efficient design and add two more centerpoint replicates.

Acknowledgments

This work was funded by UT's College of Business Summer Research Grant program. Adraine Jordan also contributed to the initial investigation as part of her M.S.

independent study project "Some New Three-Level Designs." I thank Peter Goos for sending me his optimal split plot design algorithm, and for running several cases of the five-factor split plot example for me. Both referees provided many helpful comments that greatly improved the presentation of these results for *JQT* readers.

Appendix 1: Pertinent Optimal Design Results

Kiefer (1960) showed that a second-order design for factors in a spherical region is 100% D-efficient if the following three conditions hold:

1. The design resides entirely on the boundary of the hypersphere or at the center
2. The design is second-order rotatable, which imposes requirements on design moments up to fourth order
3. The proportion of the design at the center equals $1/p$, where p is the number of coefficients to be estimated, i.e., $p = 1 + 2t + \binom{t}{2} = (t+1)(t+2)/2$.

Let $\boldsymbol{D}$ denote a design with n rows, and let $\boldsymbol{X}$ denote the second-order model matrix corresponding to $\boldsymbol{D}$. For designs constrained to the unit ball (i.e., the diagonal elements of $DD' \le 1$), $|X'X/n| \le D_\infty$, where

$$D_\infty = \quad 2^t \left(t+1\right)^{-p} \left(t+2\right)^{-t(t+2)} (t+3)^{p-1}$$

(Hardin and Sloane 1993). D-efficiency for a design is measured relative to this upper bound, i.e.,

$$D - efficiency = [|X'X/n|/D_\infty]^{1/p}.$$

Hardin and Sloane (1993) argue that, since a response surface is used for prediction, I-optimality – the criterion of minimizing the average prediction variance - should be primary. However, for second-order designs within a spherical region, a 100% D-optimal design can be made essentially 100% I-efficient by the addition of center point replicates. Suppose a design satisfies the first two conditions given above for D-optimality. Hardin and Sloane (2001) refer to such a design as *perfect*, and show that its average variance over the design region, multiplied by n, achieves the lower bound

$$\frac{n_B + n_C}{(t+2)(t+4)} \left\{ \frac{t^2(t^2+5t+10)}{2n_B} + \frac{8}{n_C} \right\} \tag{1}$$

among all designs satisfying condition 1, where n_B and n_C denote the number of boundary points and center points, respectively, for the design. With the correct assignment of n_C, perfect designs are fully efficient. Whereas D-optimality requires $n_C/(n_B+n_C) = 1/p$, which is equivalent to $n_C = n_B/(p-1)$, the average variance over the sphere (1) is minimized when

$$n_C = \frac{4t\sqrt{t^2+5t+10} - 16}{(t-1)(t+2)(t^2+4t+8)} n \approx 2.06 n_B/(p-1).$$

At this optimum allocation, (1) achieves the minimum value of

$$I_\infty = \frac{(t-1)^2(t+2)(t^2+4t+8)^2}{2(t+4)\{t\sqrt{t^2+5t+10} - 4\}^2}.$$

I-efficiency will be evaluated as

$$\text{I-efficiency} = I_\infty / trace[M(X'X)^{-1} n],$$

where $\boldsymbol{M}$ is the $p \times p$ moment matrix for the sphere (Hardin and Sloane 2001).

Appendix 2: Three-Level Second-Order Rotatable Designs

Let ***D*** denote a t-factor, three-level design with levels -1, 0 and 1, and let x_{ij} denote the level of factor *j* for the i^{th} run. Further we restrict attention to designs constrained to the center and the boundary of the hypersphere. Let *k* denote the orbit for the n_B boundary points. Finally, let r_j denote the number of rows for which $|x_{ij}| = 1$ and let $s_{jj'}$ denote the number of rows for which $|x_{ij}x_{ij'}| = 1$. The matrix ***D*** contains kn_B non-zero elements. With (-1, 0, 1) coding, rotatability implies that for all *j*, *j'*:

- $r_j = \sum_{i=1}^{n} x_{ij}^4 = kn_B / t$
- $s_{jj'} = \sum_{i=1}^{n} x_{ij}^2 x_{ij'}^2 = (kn_B / t)(k-1)/(t-1)$, since each of the r_j rows with $|x_{ij}| = 1$ contain (*k*-1) other non-zero elements that must be spread evenly among the other *t*-1 factors

Finally, rotatability requires $\sum_{i=1}^{n} x_{ij}^4 = 3\sum_{i=1}^{n} x_{ij}^2 x_{ij'}^2$ for all *j*, *j'*, which given the results above implies that (*t*-1)/(*k*-1) = 3. Thus, for three-level designs with points only from the *kth* orbit and the center, *k* = (*t*+2)/3 is a necessary condition for rotatability.

References

Albert, I. and Gauchi, J.P. (2002). "Sensitivity Analysis for High Quantiles of Ochratoxin A Exposure Distribution," *International Journal of Food Microbiology* 75, pp. 143-155.

Atkinson, A.C. (1996). "The Usefulness of Optimum Experimental Designs," *Journal of the Royal Statistical Society Ser. B* 58, pp. 59-76.

Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs*. New York, NY: Oxford University Press.

Bisgaard, S. (1997). "Why Three-Level Designs Are Not So Useful for Technological Applications", *Quality Engineering* 9, pp. 545-550.

Box, G.E.P., and Behnken, D.W. (1958). "Some New Three Level Second Order Designs for Surface Fitting," Statistical Techniques Research Group Technical Report No. 26. Princeton University.

Box, G.E.P., and Behnken, D.W. (1960). "Some New Three Level Designs for the Study of Quantitative Variables," *Technometrics* 2, pp. 455-475.

Carlyle, W.M., Montgomery, D.C., and Runger, G.C. (2000). "Optimization Problems and Methods in Quality Control and Improvement," *Journal of Quality Technology* 32, pp. 1-17.

Cochran, W.G., and Cox, G.M. (1957). *Experimental Designs.* New York, NY: Wiley & Sons.

Crossier, R.B. (1991). "Some New Three-Level Response Surface Designs," Technical Report CRDEC-TR-308, U.S. Army Chemical Research, Development and Engineering Center.

Draper, N.R., and John, J.A. (1988). "Response Surface Designs for Quantitative and Qualitative Variables," *Technometrics* 30, pp. 423-428.

Draper, N.R., and John, J.A. (1998). "Response Surface Designs where Levels of Some Factors are Difficult to Change," *Australian and New Zealand Journal of Statistics* 40, pp. 487-495.

Farrell, R.H., Kiefer, J., and Walbran, A. (1967). "Optimum Multivariate Designs." *Proceedings of the 5th Berkeley Symposium*, vol. 1, pp. 113-138 (U.C. Press).

Goos, P. and Vandebroek, M. (2001). "Optimal Split Plot Designs," *Journal of Quality Technology* 33, pp. 436-450.

Goos, P. and Vandebroek, M. (2003). "D-Optimal Split Plot Designs with Given Numbers of Sizes and Whole Plots," *Technometrics* 45, pp. 235-245.

Goos, P. and Vandebroek, M. (2004). "Outperforming Completely Randomized Designs," *Journal of Quality Technology* 36, pp. 12-26.

Hardin, R.H., and Sloane, N.J.A. (1993). "A New Approach to Construction of Optimal Designs," *Journal of Statistical Planning and Inference* 37, pp. 339-369. (Updated version of this paper appears at http://www.research.att.com/~njas/doc/de.html.)

Hardin, R.H., and Sloane, N.J.A. (2001). "Computer-Generated Minimal (and Larger) Response-Surface Designs: (I) The Sphere," http://www.research.att.com/~njas/doc/doeh.pdf.

John, J.A. & Mitchell, T.J. (1977). "Optimal Incomplete Block Designs," *Journal of the Royal Statistical Society, ser. B* 39, pp. 39-43.

Kiefer, J. (1960). "Optimum Experimental Designs V, with Applications to Systematic and Rotatable Designs," Fourth Berkeley Symposium on Mathematical Statistics and Probability. J. Neyman ed.

Kowalski, S., Cornell, J.A., and Vining, G.G. (2000). "A New Model and Class of Designs for Mixture Experiments with Process Variables," *Communications in Statistics – Theory and Methods* 29, pp. 2255-2280.

Letsinger, J.D., Myers, R.H., and Lentner, M. (1996). "Response Surface Methods for Bi-Randomization Structures," *Journal of Quality Technology* 28, pp. 381-397.

Lucas, J.M. (1976). "Which Response Surface Design is Best," *Technometrics* 18, pp. 411-417.

Mee, R.W. (2002). "Three-Level Simplex Designs and Their Use in Sequential Experimentation," *Journal of Quality Technology*, 34, pp. 152-164.

Mitchell, T.J., and Bayne, C.K. (1978). "D-Optimal Fractions of Three-Level Factorial Designs," *Technometrics* 20, pp. 369-380.

Morris, M. D. (2000). "A Class of Three-Level Experimental Designs for Response Surface Modeling," *Technometrics* 42, pp. 111-121

Myers, R.H., and Montgomery, D.C. (1995). *Response Surface Methodology.* New York, NY: Wiley & Sons.

SAS (2004). SAS Version 9.1. Cary, NC: SAS Institute..

Snee, R.D. (1985). "Computer-Aided Design of Experiments – Some Practical Experiences," *Journal of Quality Technology* 17, pp. 222-236.

Srinivas, D., Rao, J.K., Théodore, K., and Panda, T. (1995). "Direct Conversion of Cellulosic Material to Ethanol by the Intergeneric Fusant Trichoderma Reesei QM 9414/Saccharomyces Cerevisiae NCIM 3288," *Enzyme and Microbial Technology* 17, pp. 418-423.

Srinivasaiah, H.C., and Bhat, N. (2005). "Characterization of Sub-100 nm CMOS Process Using Screening Experiment Technique," *Solid-State Electronics* 49, pp. 431-436.

Thomas, R.S., Soderborg, N.S., and Borders, S.G. (2005). "Using Computer-Aided Engineering to Find and Avoid Steering Wheel "Nibble" Failure Mode," SAE 2005 World Congress Presentation: Detroit, MI.

Trinca, L.A., and Gilmour, S.G. (2001). "Multistratum Response Surface Designs," *Technometrics* 43, pp. 25-33.

Vining, GG., Kowalski, S.M., and Montgomery, D.G. (2005). "Response Surface Designs Within a Split-Plot Structure," *Journal of Quality Technology* 37, pp. 115-129

Vuchkov, I.N., Damgaliev, D.L., and Yontchev, C.A. (1981). "Sequentially Generated Second Order Quasi D-Optimal Designs for Experiments with Mixture and Process Variables," *Technometrics* 23, pp. 233-238.

Wu, C.F.J., and Ding, Y. (1998). "Construction of Response Surface Designs for Qualitative and Quantitative Factors," *Journal of Statistical Planning and Inference* 71, pp. 331-348.

TABLE 1. Summary of Box and Behnken's (1958, 1960) and Crossier's (1991) Designs, with D-Optimal and I-Optimal Number of Center Run Replicates

	Design		Number of Runs		Efficiency (%)	
t	Structure	p	n_B	n_C	D-eff.	I-eff.
3	BIB: 3×2^2	10	12	1	97.00	82.64
				3	93.82	95.25
4	BIB: 6×2^2	15	24	2	99.92	96.39
				4	97.17	99.83
5	BIB: 10×2^2	21	40	2	98.83	93.83
				4	97.50	98.02
6	RG: 6×2^3	28	48	2	94.61	88.82
				4	93.25	89.83
	BIB: 10×2^3	28	80	3	99.48	95.71
				6	98.41	98.99
7	BIB: 7×2^3	36	56	2	99.93	98.67
				3	99.35	99.95

8	RG: 24×2^3	45	192	4	97.91	93.37
				9	97.21	95.92
	BIB: 14×2^4	45	224	5	99.06	95.82
				10	98.45	98.08
9	BIB: 12×2^3	55	96	2	98.93	96.91
				4	98.18	98.11
	RG: 15×2^3	55	120	2	95.06	89.62
				4	94.72	91.55
	RG: 9×2^4	55	144	3	92.96	84.71
				5	92.56	85.49
	BIB: 18×2^4	55	288	5	99.81	97.40
				11	99.22	99.62
10	RG: 10×2^4	66	160	2	93.56	85.46
				5	93.04	87.31
	BIB: 15×2^4	66	240	4	100.00	98.75
				8	99.42	99.99
11	BIB: $11 \times 2^{5-1}$	78	176	2	99.57	97.26
				5	99.08	99.12
12	RG: 12×2^4	91	192	2	97.12	94.01
				4	96.86	95.21
13	BIB: 13×2^4	105	208	2	99.11	97.46
				4	98.82	98.42
16	RG: 24×2^4	153	384	3	93.83	89.38
				5	93.66	89.64
	BIB: $16 \times 2^{6-1}$	153	512	3	100.00	99.05
				7	99.78	100.00

The first design for each t (except $t = 8$, 9 and 13) appear in Box and Behnken (1960). The first designs for $t = 9$ and $t = 13$ are from Crossier (1991). All other designs appear in Box and Behnken (1958). The first (second) n_C value listed here for each design is D-optimal (I-optimal). BIB = balanced incomplete block. RG = regular graph.

TABLE 2. D-Optimal Design with Run Size n = 26 Constructed from a Candidate Set of 33 Points

Candidate Point No.	x_1	x_2	x_3	x_4	Frequency
1	-0.70711	-0.70711	0	0	1
2	-0.70711	0	-0.70711	0	1
3	-0.70711	0	0	-0.70711	1
4	-0.70711	0	0	0	0
5	-0.70711	0	0	0.70711	1
6	-0.70711	0	0.70711	0	1
7	-0.70711	0.70711	0	0	1
8	0	-0.70711	-0.70711	0	1
9	0	-0.70711	0	-0.70711	1
10	0	-0.70711	0	0	0
11	0	-0.70711	0	0.70711	1
12	0	-0.70711	0.70711	0	1
13	0	0	-0.70711	-0.70711	1
14	0	0	-0.70711	0	0
15	0	0	-0.70711	0.70711	1
16	0	0	0	-0.70711	0
17	0	0	0	0	2
18	0	0	0	0.70711	0
19	0	0	0.70711	-0.70711	1
20	0	0	0.70711	0	0
21	0	0	0.70711	0.70711	1
22	0	0.70711	-0.70711	0	1
23	0	0.70711	0	-0.70711	1
24	0	0.70711	0	0	0
25	0	0.70711	0	0.70711	1
26	0	0.70711	0.70711	0	1
27	0.70711	-0.70711	0	0	1
28	0.70711	0	-0.70711	0	1
29	0.70711	0	0	-0.70711	1
30	0.70711	0	0	0	0
31	0.70711	0	0	0.70711	1
32	0.70711	0	0.70711	0	1
33	0.70711	0.70711	0	0	1

TABLE 3: Fitted Second-Order Model, using 47 Runs from Srinivas et al. (1995)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	27	0.02204393	0.000816	1.8212
Error	19	0.00851756	0.000448	Prob > F
C. Total	46	0.03056149		0.0897

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	15	0.00837036	0.000558	15.1637
Pure Error	4	0.00014720	0.000037	Prob > F
Total Error	19	0.00851756		0.0088
				Max RSq
				0.9952

TABLE 4: Loss of D-Efficiency from Restricting Factors to Three Levels

	$n = p$			$n = p + 5$		
t	D-eff for 3-level design	D-eff for many-level design	Efficiency Ratio	D-eff for 3-level design	D-eff for many-level design	Efficiency Ratio
4	0.7923	0.9707	0.82	0.9126	0.9972	0.92
5	0.8281	0.9747	0.85	0.8853	0.9979	0.89
6	0.8703	0.9994	0.87	0.9208	0.9968	0.92
7	0.8780	0.9737	0.90	0.9244	0.9444	0.98

TABLE 5: Loss of I-Efficiency from Restricting Factors to Three Levels

	$n = p + 1$			$n = p + 6$		
t	I-eff for 3-level design	I-eff for many-level design	Efficiency Ratio	I-eff for 3-level design	I-eff for many-level design	Efficiency Ratio
4	0.3727	0.9429	0.40	0.7628	0.9712	0.79
5	0.6658	0.9488	0.70	0.7823	0.9937	0.79
6	0.7443	0.9999	0.74	0.8408	0.9954	0.84
7	0.7471	0.9487	0.79	0.8510	0.9848	0.86

TABLE 6: D-Optimal Three-Level Second-Order Design for 3 Mixture Factors and 3 Other Factors

Point	x1	x2	x3	x4	x5	x6
1	0.0	0.0	1.0	0.000	-0.707	-0.707
2	0.0	0.0	1.0	0.000	-0.707	0.707
3	0.0	0.0	1.0	-0.707	0.000	0.707
4	0.0	0.0	1.0	0.707	0.000	-0.707
5	0.0	0.0	1.0	-0.707	0.707	0.000
6	0.0	0.0	1.0	0.707	0.707	0.000
7	0.0	0.5	0.5	0.000	0.000	0.000
8	0.0	0.5	0.5	0.000	0.707	-0.707
9	0.0	0.5	0.5	0.707	0.000	0.707
10	0.0	0.5	0.5	-0.707	-0.707	0.000
11	0.0	1.0	0.0	0.000	-0.707	-0.707
12	0.0	1.0	0.0	0.000	-0.707	0.707
13	0.0	1.0	0.0	0.000	0.707	0.707
14	0.0	1.0	0.0	-0.707	0.000	-0.707
15	0.0	1.0	0.0	0.707	0.000	-0.707
16	0.0	1.0	0.0	-0.707	0.707	0.000
17	0.0	1.0	0.0	0.707	0.707	0.000
18	0.5	0.0	0.5	0.000	0.000	0.000
19	0.5	0.0	0.5	0.000	0.707	0.707
20	0.5	0.0	0.5	-0.707	0.000	-0.707
21	0.5	0.0	0.5	0.707	-0.707	0.000
22	0.5	0.5	0.0	0.000	0.707	-0.707
23	0.5	0.5	0.0	-0.707	0.000	0.707
24	0.5	0.5	0.0	0.707	-0.707	0.000
25	1.0	0.0	0.0	0.000	-0.707	-0.707
26	1.0	0.0	0.0	0.000	-0.707	0.707
27	1.0	0.0	0.0	0.707	0.000	-0.707
28	1.0	0.0	0.0	0.707	0.000	0.707
29	1.0	0.0	0.0	-0.707	0.707	0.000
30	1.0	0.0	0.0	0.707	0.707	0.000

TABLE 7: D-Optimal Three-Level Split-Plot Design for 5 Factors in 21 Whole Plots

Whole Plot	x_1	x_2	x_3	x_4	x_5
1	-1	-1	0	0	0
1	-1	1	0	0	0
2	-1	0	-1	0	0
2	-1	0	1	0	0
3	-1	0	0	-1	0
3	-1	0	0	1	0
4	-1	0	0	0	-1
4	-1	0	0	0	1
5	0	-1	-1	0	0
5	0	-1	1	0	0
6	0	-1	0	-1	0
6	0	-1	0	1	0
7	0	0	1	0	1
7	0	0	1	0	-1
8	0	1	0	-1	0
8	0	0	-1	0	1
9	0	0	1	1	0
9	0	0	1	-1	0
10	0	0	0	-1	-1
10	0	0	0	-1	1
11	0	1	1	0	0
11	0	0	0	1	1
12	0	0	0	1	-1
12	0	1	-1	0	0
13	0	0	-1	1	0
13	0	-1	0	0	1
14	0	1	0	1	0
14	0	0	-1	0	-1
15	0	1	0	0	1
15	0	0	0	0	0
16	0	0	0	0	0
16	0	0	-1	-1	0
17	0	1	0	0	-1
17	0	-1	0	0	-1
18	1	-1	0	0	0
18	1	1	0	0	0
19	1	0	-1	0	0
19	1	0	1	0	0
20	1	0	0	-1	0
20	1	0	0	1	0
21	1	0	0	0	-1
21	1	0	0	0	1

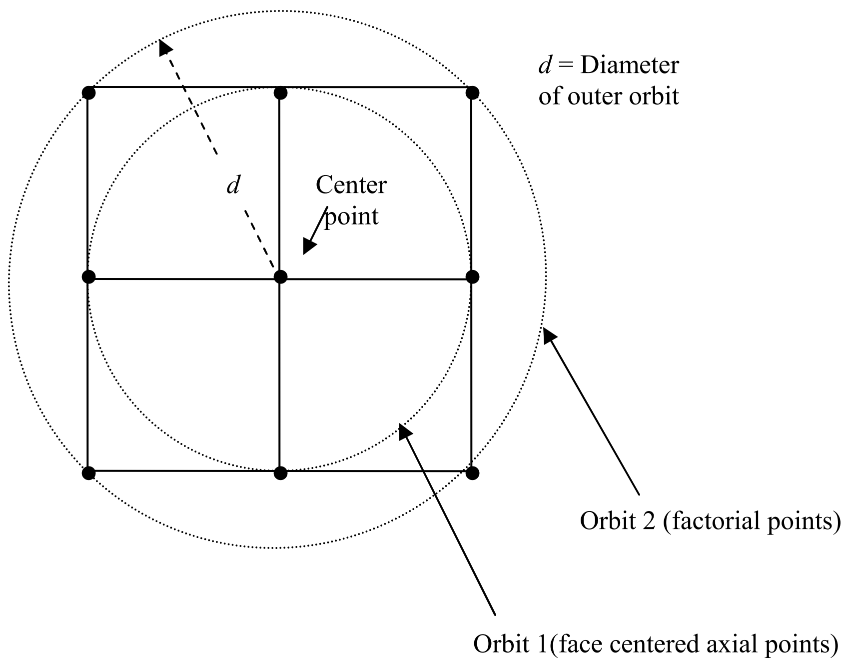

FIGURE 1. A Simple 3^2 Design Showing the Two Orbits of Points

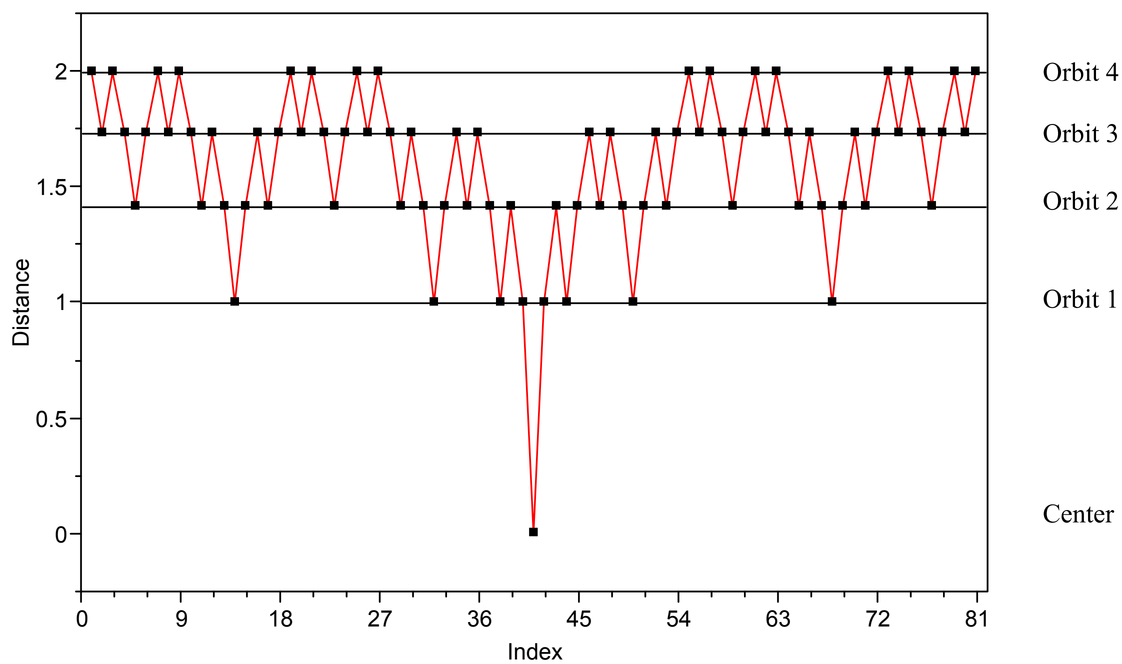

FIGURE 2. Orbits for a Full 3^4 Factorial Design with Levels (-1, 0, 1)

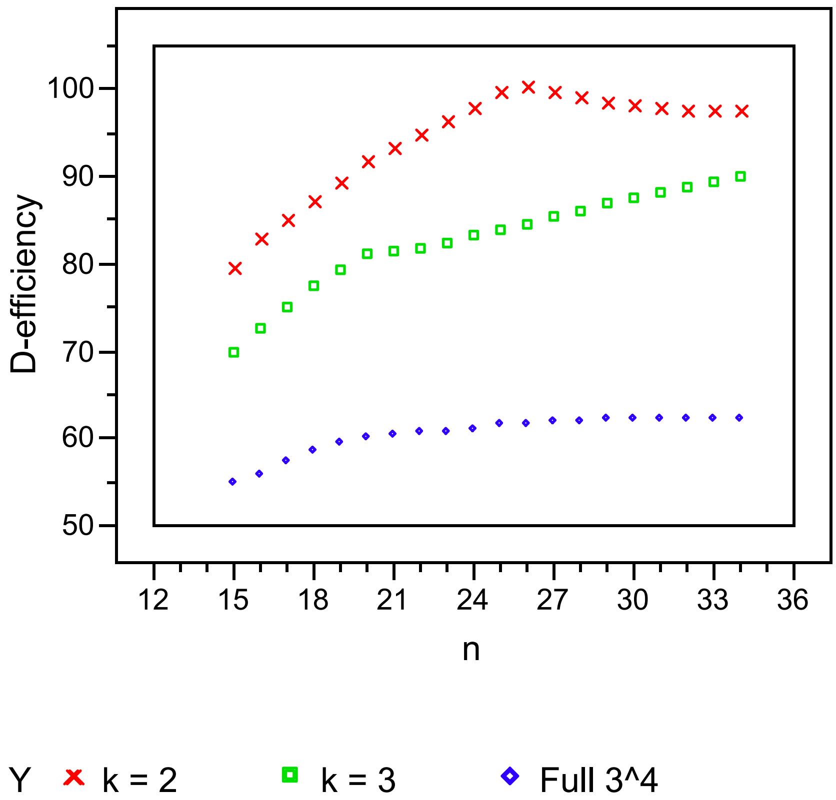

FIGURE 3. D-Optimal Three-Level Designs for Four Factors: Three Different Candidate Sets

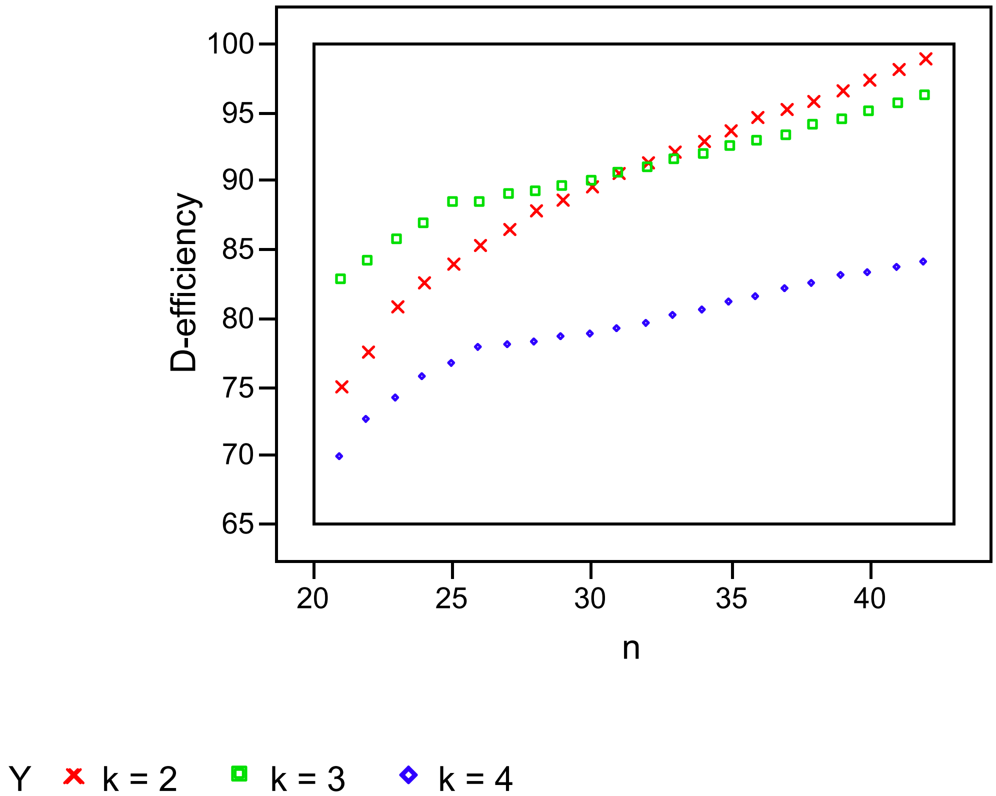

FIGURE 4. D-Optimal Three-Level Designs for Five Factors: Three Different Candidate Sets

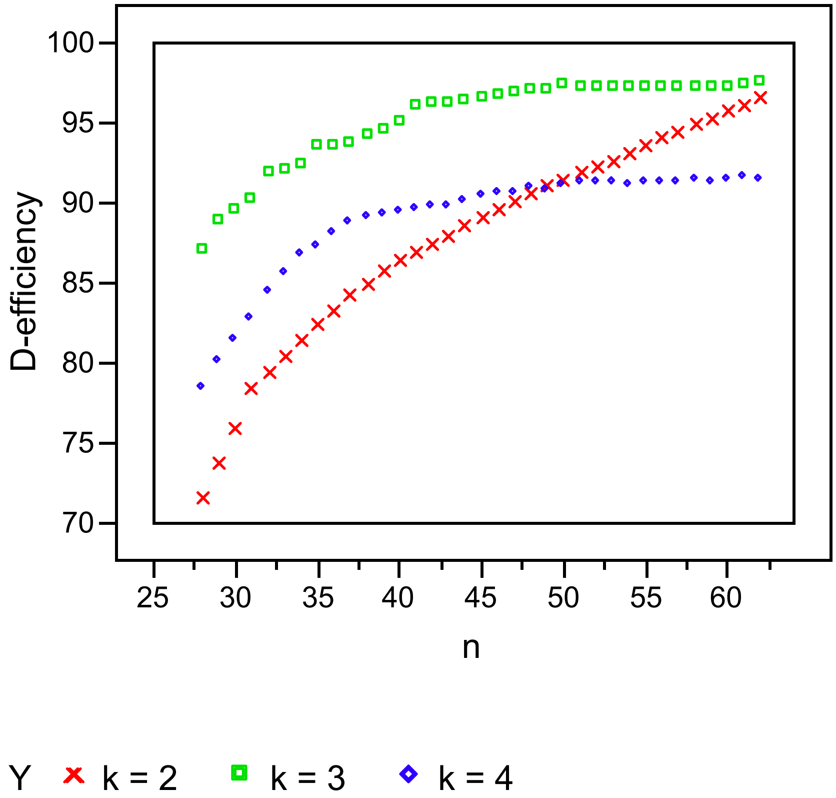

FIGURE 5. D-Efficient Three-Level Designs for Six Factors: Three Different Candidate Sets

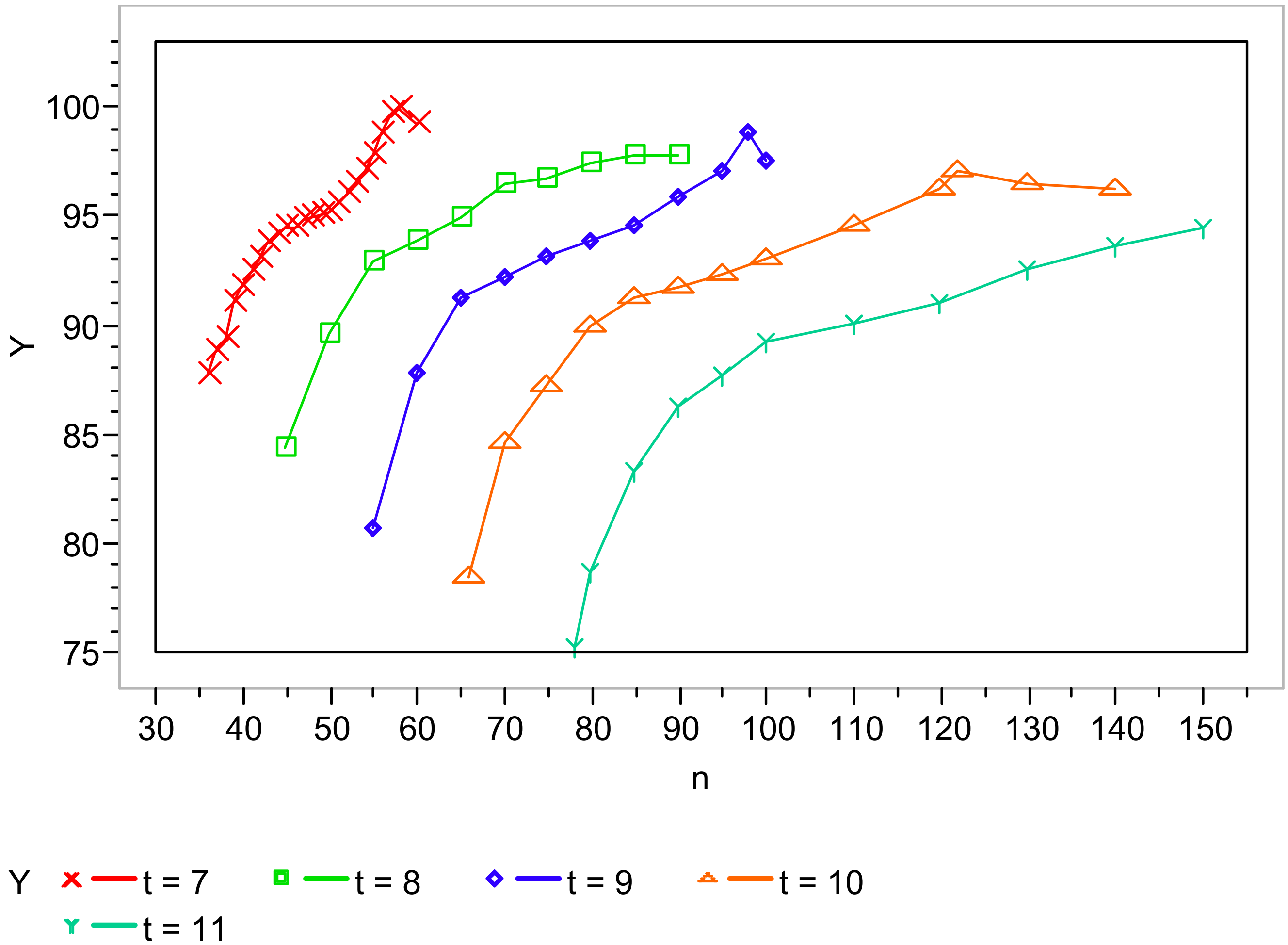

FIGURE 6. D-Efficient Three-Level Designs for t = 7 - 11 Factors, Each Using k = 3rd Orbit