TRENDS IN SCIENCES 2024; 21(11): 8193 RESEARCH ARTICLE
https://doi.org/10.48048/tis.2024.8193

Comparison of the Three Types of Central Composite Designs Over Subsets of Reduced Models by Design Optimality Criteria

Chawanee Suphirat[1] and Wasinee Pradubsri[2,*]

[1]*Department of Mathematics and Statistics, Faculty of Science and Technology, Rajamangala University of Technology Phra Nakhon, Bangkok 10800, Thailand*
[2]*Department of Science and Mathematics, Rajamangala University of Technology Isan, Surin Campus, Surin 32000, Thailand*

([*]Corresponding author's e-mail: wasinee.pr@rmuti.ac.th)

Received: 13 March 2024, Revised: 27 June 2024, Accepted: 4 July 2024, Published: 10 October 2024

Abstract

The purpose of this article is to compare the 3 important types of central composite designs (CCDs) consisting of central composite circumscribed design (CCCD), central composite inscribed design (CCID), and central composite face-centroid design (CCFD) in response surface methodology (RSM). The difference among these designs is the distance from the center design to the axial points. The comparison was performed across the full second-order response surface model and across a set of reduced models for 3, 4, and 5 design factors ($k = 3$, 4, and 5) including 1, 3, and 5 center runs ($n_c = 1$, 3, and 5). This study used D-, A-, and G- optimality criteria to evaluate the performance of CCDs by presenting the design optimality criteria comparison ranking throughout the reduced-model subsets of 43, 224, and 839 models for 3, 4, and 5 design variables, respectively. The results showed that CCCD was superior to CCID and CCFD according to A- and D- optimality criteria, while CCCD and CCID performed better than CCFD based on G- optimality criterion over a set of reduced models for 3, 4, and 5 design factors. It was observed that D-, A-, and G- optimality efficiencies were robust to changes in the linear and cross-product terms and sensitive to deviation in the square terms. The study will provide recommendations to assist the experimenters in the choice of the best design among the candidate designs for practice applications when some model effects may be insignificant.

Keywords: CCD, Design optimality criterion, Reduced models, RSM, Weak heredity

Introduction

Experiments are often conducted by practitioners in numerous fields including science, food, agriculture, chemical engineering, and industrial processes in order to study and model the effects of explanatory variables and the measurable response of interest. The original principles of response surface methodology (RSM) were presented by Box and Wilson [1]. RSM is a collection of mathematical and statistical techniques used to establish and analyze empirical model in which the response is a function of various independent variables so as to optimize this response [2-4]. An experimental design which applies such methodology is typically called the response surface design. In practice, the form of true relationships between design factors and the measurable response is unknown, however, it can be approximated by the second-order polynomial response surface model, as shown in Eq. (1).

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{j=i+1}^{k} \sum_{i=1}^{k} \beta_{ij} x_i x_j + \varepsilon \quad (1)$$

where, y represents the response or dependent variable, β_0 is the y - intercept, $\beta_1, \beta_2, \dots, \beta_k$ are the parameter coefficients, $x_1, x_2, \dots, x_k$ are the design or independent factors, k is the number of design variables, and ε is the random error.

Several response surface designs, particularly central composite designs (CCDs), have traditionally utilized the proposed model in Eq. (1). Furthermore, CCD remains the most popular response surface design for estimating a quadratic model. CCDs are a type of RSM that allows for the optimization of multiple factors and interactions in an experiment. They have been applied to various fields of engineering and science, such as chemical engineering, mechanical engineering, electrical engineering, biotechnology, agriculture, and medicine. For instance, CCD can be used to optimize the adsorption process for the removal of cadmium from contaminated water [5]. Because the specified model in Eq. (1) has various choices of response surface designs used in real-world situation, choosing an appropriate design is important. To help researchers achieve a suitable design prior to running the design, single-valued criteria are suggested. Single-valued criteria refer to alphabetic-optimality criteria (such as D-, A-, and G- optimality criteria) that are commonly used to evaluate competing designs. D- and A- optimality criteria are concerned with model parameter estimation, while G- optimality criterion is focused on prediction variance characteristics [6]. Numerous authors have used optimality criteria to compare response surface designs, see [7-9].

In addition, some publications considered the performance comparison CCDs across full second-order response surface design in Eq. (1). Zhang and Baixiaofeng [10] compared the 3 classes of CCDs consisting of central composite circumscribed design (CCCD), central composite inscribed design (CCID) and central composite face-centroid design (CCFD) with respect to the region of interest and robustness of design throughout simulated experimental data. The distinct designs are distance from the center design to the axial points. CCCD is the original form of CCD, while CCID is a scale down CCCD. CCFD is a special type of CCD which distance from the center design to star points equals 1 [11]. Kiwu-Lawrence *et al.* [11] evaluated the performances of the CCDs including CCCD, CCID, and CCFD in terms of the D-, A-, and G- optimality criteria for 3, 4, …, 10 design factors ($k =$ 3, 4, …, 10) with 0, 1, …, 5 center points ($n_c =$ 0, 1, …, 5). They recommended that the CCD is superior if the center points are replicated and added to the design. For additional information see [12,13].

Prior to running the experimental design, the data were collected, and then the model parameters tested. The response surface model may not always turn out to be the full second-order model stated in Eq. (1) because one or more model parameters may be insignificant. Therefore, the research may subsequently adapt the model, leading to the selection of model with fewer parameter terms as well as reduced models for approximating the response values in actual experiment. This implies that the experimenter is encountered the problem with choosing desirable optimal design in the analysis phase. It is important to note the fact that a design is the best not only over the full second-order model but also over a set of reduced models. Consequently, the experimental designs should have robustness properties across subsets of the reduced models as well as the design optimal criteria considered. The finding of this article has extensive applications in industrial processes, particularly chemical engineering processes. For example, optimization and kinetic modelling of water extracts from the radius of the particle size, temperature, microwave power and irradiation time, see [14]. Additional examples can be found in [11,15]. When the researchers find some model effects may be insignificant, they will determine select a reduced model in the analysis phases.

Borkowski and Valeroso [16] presented and evaluated the robustness characteristics in CCDs and other response surface designs, while Chomtee and Borkowski [17] and Chomtee [18] accessed the performance in the spherical region over a collection of reduced models for $k = 3$, 4, and 5 based on D-, A-, G- and IV- optimality criteria.

In this article, we compare the performance of CCD classifications with the replication of the center points by extending the researches of Kiwu-Lawrence *et al.* [11], Chomtee and Borkowski [17], and Chomtee [18]. The 3 types of CCDs (CCCD, CCID, and CCFD) were evaluated in terms of D-, A-, and G-optimality criteria for $k = 3$, 4, and 5 over classes of reduced models by using weak heredity [19]. The full factorial portion of the CCDs stated in this paper were examined for factors $k = 3$ and 4, whereas the fractional factorial portion of the CCDs were examined for only factor $k = 5$. The performances of these designs were considered for $n_c = 1$, 3, and 5.

Materials and methods

This section gives an overview of CCDs and presents design optimality criteria used in this study. After that, the full and reduced models are summarized.

Central composite design

This part provides the composition of central composite designs (CCDs) and describes structures of the 3 types of CCDs considered in this research: Central composite circumscribed design (CCCD), central composite inscribed design (CCID) and central composite face-centroid design (CCFD).

CCD, introduced by Box and Wilson [1], is the most popular response surface design for fitting the second-order polynomial response surface model as in Eq. (1). CCD comprises of 3 main parts as follows:

(i) f factorial portion with the form $(x_1, x_2, \dots, x_k)$ = (± 1, ± 1, …, ± 1) from a 2^k full or 2^{k-g} fractional factorial design of at least resolution V where k represents the number of design variables, and g represents an integer. $k \geq 2$ are assumed.

(ii) $2k$ axial portion or star points with the form (± α, 0, …, 0), (0, ± α, 0, …, 0), …, (0, 0, …, ± α), α is the axial distance from the center points.

(iii) n_c center points (0, 0, …, 0).

Therefore, the total number of design runs is expressed by $N = f + 2k + n_c$, where N is the number of design points.

For CCDs, the factorial portion provides to estimate the effect of linear and cross-product terms, the axial portion allows to approximate the effect of quadratic terms, and the center points provide to determinate the effect of error terms in the model (1) related with the square terms [2-4]. This article studied the impact of center point replications because the center points were used to investigate the distribution of scale prediction variance. Additionally, they also provided to give a better estimation of the pure error terms, see [12,15].

The size of these design points can increase dramatically as these designs have high number of design factors as well as factorial portion, especially $k = 5$. The fractional factorial portion is used to minimize the design size in the structure of the CCD. In this research, the CCD utilized the full factorial portion for 3 and 4 design factors, whereas the CCD used fractional factorial portion (2^{k-1}) for 5 factors with $n_c = 1$, 3, and 5. Additional center points are utilized to estimate pure error for the lack of fit test.

Choosing a suitable value of α in axial points depends on the region of interest and the types of the CCD. In this study, CCCD and CCID were applied to a rotatable design in which the variance of the predicted response was constant at points that are the same distance from the design center. The concept of rotatability for experimental design was provided by Box and Hunter [20]. The values of α that make rotatable or near-rotatable are $\alpha = (f)^{1/4}$, where f is the number of factorial points in the composition of the CCD. For CCFD, we considered the cuboidal region, $\alpha = \pm 1$.

Central composite circumscribed design

The central composite circumscribed design (CCCD) is the fundamental composition of CCD which composes of factorial, star and center points. To construct CCCD, either a 2-level full factorial design or fractional factorial design of resolution V is generated. A $2k$ axial portion is selected and located at a distant α from the center [2,11]. Axial points are coded with low and high factor levels (-1,1). Thus, the CCCD offers 5 levels for each design factor. To determine the values of α, for rotatable design, the CCCD is rotatable when $\alpha = (f)^{1/4}$. An example of CCCD design matrix for $k = 3$ and $n_c = 1$ is shown in **Table 1**.

Table 1 CCCD design matrix for $k = 3$ and $n_c = 1$.

Structure	Points	x_1	x_2	x_3
factorial portion (f)	1	−1	−1	−1
	2	−1	−1	1
	3	−1	1	−1
	4	−1	1	1
	5	1	−1	−1
	6	1	−1	1
	7	1	1	−1
	8	1	1	1
axial portion ($2k$)	9	$-\alpha$	0	0
	10	α	0	0
	11	0	$-\alpha$	0
	12	0	α	0
	13	0	0	$-\alpha$
	14	0	0	α
center points (n_c)	15	0	0	0

Central composite inscribed design

The central composite inscribed design (CCID) is one of the varieties of CCD in which the design space of operability coincides with the region of interest. The CCID has a similar structure to the CCD with the difference between them being the nature of factorial and star points. The star points are put on the interior of the factorial design, equidistant from the center [10,13]. The CCID is created by dividing each factor level of the CCCD by α. This means that the CCID is considered to be a scaled down CCCD. Moreover, the CCID also offers 5 levels for each design factor. An example of CCID design matrix for $k = 3$ and $n_c = 1$ is given in **Table 2**.

Table 2 CCID design matrix for $k = 3$ and $n_c = 1$.

Structure	Points	x_1	x_2	x_3
factorial portion (f)	1	$-1/\alpha$	$-1/\alpha$	$-1/\alpha$
	2	$-1/\alpha$	$-1/\alpha$	$1/\alpha$
	3	$-1/\alpha$	$1/\alpha$	$-1/\alpha$
	4	$-1/\alpha$	$1/\alpha$	$1/\alpha$
	5	$1/\alpha$	$-1/\alpha$	$-1/\alpha$
	6	$1/\alpha$	$-1/\alpha$	$1/\alpha$
	7	$1/\alpha$	$1/\alpha$	$-1/\alpha$
	8	$1/\alpha$	$1/\alpha$	$1/\alpha$
axial portion ($2k$)	9	−1	0	0
	10	1	0	0
	11	0	−1	0
	12	0	1	0
	13	0	0	−1
	14	0	0	1
center points (n_c)	15	0	0	0

Central composite face-centroid design

The central composite face-centroid design (CCFD) is a special case of the CCD with setting $\alpha = 1$. This is to mean that all factorial and axial portions of the CCFD are located on the surface of the cube. Consequently, the values of $\alpha = 1$ give the cuboidal CCFD in response surface design. The CCFD also requires 3 levels for each design factor as well as a 3-level design. An example of CCFD design matrix for $k = 3$ and $n_c = 1$ is shown in **Table 3**.

Table 3 CCFD design matrix for $k = 3$ and $n_c = 1$.

Structure	Points	x_1	x_2	x_3
factorial portion (f)	1	−1	−1	−1
	2	−1	−1	1
	3	−1	1	−1
	4	−1	1	1
	5	1	−1	−1
	6	1	−1	1
	7	1	1	−1
	8	1	1	1
axial portion ($2k$)	9	−1	0	0
	10	1	0	0
	11	0	−1	0
	12	0	1	0
	13	0	0	−1
	14	0	0	1
center points (n_c)	15	0	0	0

Design optimality criteria

We now demonstrate the design optimality criteria for evaluating the competing designs related to this work. An optimal design refers to an experimental design which is generated with respect to satisfying an optimality criterion. These optimal design procedures utilize single-valued criteria or the alphabetic-optimality criteria to summarize and measure precise parameter estimation and variance characteristics for each experimental design. D-, A-, and G- optimality criteria are commonly used for design comparison. Most optimality criteria focus on the properties of the information matrix or $X'X$, where X represents the design matrix. See [21,22] for a comprehensive detail of design optimality criteria.

When comparing the performance for each design using D-, A-, and G- optimality criteria, it is crucial to remember that while an experimental design may perform best among all designs under 1 optimality criterion, it may perform worse under another or more criteria. As a result, choosing an optimality criterion is a primary consideration prior to running a design. Furthermore, a design is said to be efficient when it performs very well both over the full model and over a set of reduced models [16-18,23]. In this article, D-, A-, and G- optimality criteria were examined and their efficiencies calculated to evaluate the 3 types of CCDs (CCCD, CCID, and CCFD) over the full second-order response surface model and collections of reduced models.

D- optimality criterion is focused on minimizing the volume of the joint confidence region for the vector of model coefficients or maximizing the determinant of information design matrix ($|X'X|$). For this criterion, an experimental design with the largest $|X'X|$ is desirable in terms of parameter prediction.

The D- efficiency of an experimental design can be computed as:

$$\text{D-eff} = 100\frac{|X'X|^{1/p}}{N} \tag{2}$$

A- optimality criterion requires the minimizing the sum of variances of the model coefficients, or equivalently, minimizing the trace of inverse of $|X'X|$. An experimental design with the lowest sum of the diagonal elements of $(X'X)^{-1}$ is preferred with respect to parameter precision.

The A- efficiency of an experimental design can be given by:

$$\text{A-eff} = 100\frac{p}{\text{trace}\left[N(X'X)^{-1}\right]} \tag{3}$$

G- optimality criterion is based on prediction variance properties and focused on minimizing a design's maximum prediction variance. The G-optimality criterion requires values at a particular location in the experimental design region for the desirable prediction. The prediction variance function at a point x over the design space can be written as:

$$\text{var}\left[\hat{y}(x)\right] = \sigma^2 f'(x)(X'X)^{-1} f(x) \tag{4}$$

where $\text{var}\left[\hat{y}(x)\right]$

$f(x) = [f_1(x),..., f_p(x)]'$ represents the vector of design points in design space according to p parameter model terms, and X is the design matrix. For example, if there are 3 factors and the model is the full second-order response surface model, then $f'(x) = (1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1^2, x_2^2, x_3^2)$.

To compare the competing designs, it is often convenient to scale the prediction variance. From Eq. (4), the division by σ^2 makes a scale-free measure and multiplication by the design size (N) penalizes the prediction variance for larger experimental designs. The resulting scaled prediction variance (SPV) is as follows:

$$\text{SPV} = \frac{N\,\text{var}\left[\hat{y}(x)\right]}{\sigma^2} = Nf'(x)(\text{X}'\text{X})^{-1}f(x) \tag{5}$$

Consequently, the SPV is used to facilitate comparison among competing experimental designs of various sizes. That means, additional design sizes or N does not affect the prediction variance capabilities. The SPV is an appropriate measure to estimate response precision at any point in design space. Consequently, the G-optimal design searches a design by minimizing the maximum SPV over design space.

The G-efficiency of an experimental design is calculated by:

$$\text{G-eff} = \frac{100p}{N\sigma^2_{\text{max}}} \tag{6}$$

where p is the number of model parameters, and σ^2_{max} is the maximum of $f'(x)(\text{X}'\text{X})^{-1}f(x)$ approximated over the subsets of design points in the design space. According to this criterion, a large G-efficiency measure is desirable.

In this research article, the D-, A-, and G- optimality values and efficiencies were computed by using the MATLAB software version R2021b for the 3 classes of CCDs (CCCD, CCID, and CCFD) across the full model and a set of the reduced models.

Full and reduced models

In this section, the full second-order response surface and a set of reduced models related to this article were summarized. To compare each design, robustness was quantified by calculating D-, A-, and G-optimality efficiencies over the reduced-model subsets for the full second-order response surface model in Eq. (1). Unfortunately, the full model frequently failed to fit the response surface model. Consequently, reduced models could satisfactorily form the appropriate model. Weak heredity [19] was used to construct the subsets of reduced models, with the procedures for construction as following: (i) If a quadratic term (x_i^2) is contained in the model, x_i term must be contained in the model or (ii) If an interaction term is contained in the model, x_i or x_j must be contained in the model.

Suppose the full model is the second-order response surface model in Eq. (1). The subsets of weak heredity reduced models of the full response surface model were 43, 224, and 839 models considered for k = 3, 4, and 5 factors, respectively. According to numerous a possible set of reduced models, if the practitioners have limited resources due to time, design sizes, and cost constraints, design optimal criteria are commonly utilized to assess a proposed design over subset of reduced models before running it. Most industrial experiment follow CCDs to find the optimal production or the response values. Elhag *et al.* [14] performed the CCCD to identify the optimal empirical models to represent the kinetic characteristics of extraction process.

An example of a set of reduced models for $k = 3$ is presented in **Table 4**. Let 0 and 1 denote the absence and presence of that x_i term in each reduced model, respectively. p is the number of model parameters, dv is the number of design variables appearing in the model. $l, c,$ and q represent the number of linear, cross-product, and square terms presented in the model, respectively.

Table 4 Reduced models for k = 3.

Model	p	dv	L	C	Q	(l, c, q)
1	10	3	(1, 1, 1)	(1, 1, 1)	(1, 1, 1)	(3, 3, 3)
2	9	3	(1, 1, 1)	(1, 1, 1)	(0, 1, 1)	(3, 3, 2)
3	9	3	(1, 1, 1)	(0, 1, 1)	(1, 1, 1)	(3, 2, 3)
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
41	3	2	(0, 0, 1)	(0, 0, 1)	(0, 0, 0)	(1, 1, 0)
42	3	1	(0, 0, 1)	(0, 0, 0)	(0, 0, 1)	(1, 0, 1)
43	2	1	(0, 0, 1)	(0, 0, 0)	(0, 0, 0)	(1, 0, 0)

Source: Borkowski and Valeroso [16].
Note: $L = (x_1, x_2, x_3)$, $C = (x_1x_2, x_1x_3, x_2x_3)$, and $Q = (x_1^2, x_2^2, x_3^2)$.

For example, for a 3-factor CCCD, if the reduced model for k = 3 is Model 2 (from **Table 4**), then $f'(x) = (1,\ x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_2^2, x_3^2)$ and $p = 9$. The structure of the expanded design matrix X for k = 3, n_c = 1, $\alpha = (8)^{1/4}$ = 1.6818, and $N = 17$ is shown below:

$$
X = \begin{array}{c} \begin{array}{ccccccccc} \boldsymbol{x}_0 & \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_1\boldsymbol{x}_2 & \boldsymbol{x}_1\boldsymbol{x}_3 & \boldsymbol{x}_2\boldsymbol{x}_3 & \boldsymbol{x}_2^2 & \boldsymbol{x}_3^2 \end{array} \\ \begin{pmatrix}
1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \\
1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1.6818 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1.6818 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1.6818 & 0 & 0 & 0 & 0 & 2.8285 & 0 \\
1 & 0 & 1.6818 & 0 & 0 & 0 & 0 & 2.8285 & 0 \\
1 & 0 & 0 & -1.6818 & 0 & 0 & 0 & 0 & 2.8285 \\
1 & 0 & 0 & 1.6818 & 0 & 0 & 0 & 0 & 2.8285 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \end{array}.
$$

The corresponding $X'X$ matrix is:

$$X'X = \begin{pmatrix} 17 & 0 & 0 & 0 & 0 & 0 & 0 & 13.6569 & 13.6569 \\ 0 & 13.6569 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 13.6569 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 13.6569 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 13.6569 & 0 & 0 & 0 & 0 & 0 & 0 & 24.0003 & 8 \\ 13.6569 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 24.0003 \end{pmatrix}.$$

The D-, A- and G- efficiencies of a 3-factor CCCD are the following:

$$\text{D-eff} = 100\frac{|X'X|^{1/p}}{N} = \frac{100(3.5678\times10^{10})^{1/9}}{15} = 67.7532.$$

$$\text{A-eff} = 100\frac{p}{\text{trace}\left[N(X'X)^{-1}\right]} = \frac{100(9)}{16.0437} = 56.0967.$$

For calculating G- efficiencies, if maximum of $f'(x)(X'X)^{-1}f(x)$ approximated over the subsets of design points in the design space occurred at $[x_1\ x_2\ x_3] = [1.6473\ -0.5090\ 0.1654]$, $\sigma^2_{\max} = f'(x)(X'X)^{-1}f(x) = 0.6699$, then $\text{G-eff} = \frac{100p}{N\sigma^2_{\max}} = \frac{100(9)}{17(0.6699)} = 79.4123$.

Results and discussion

For each design comparison, robustness was quantified by computing D-, A-, and G- optimality measures for the 3 types of CCDs (CCCD, CCID, and CCFD) across the full second-order response surface model and the collections of reduced models including k = 3, 4, and 5 with n_c = 1, 3, and 5. The 3 competing CCDs were assessed on a per point basis. Note that larger D-, A-, and G- efficiency values are desired.

Full second-order response surface model

The results of the D-, A-, and G- efficiency values for each design over the full second-order model when k = 3, 4, and 5 with n_c = 1, 3, and 5 are summarized in **Tables 5 - 7**.

Table 5 Summary of D-, A-, and G- efficiencies for CCCD, CCID, and CCFD over the full second-order model when k = 3.

Designs	N	n_c	D – eff	A – eff	G – eff
CCCD	15	1	68.7039	32.0718	67.4514
	17	3	67.6079	49.3151	87.8147
	19	5	63.6517	51.9063	78.5802
CCID	15	1	14.4430	7.7746	67.4513

Designs	N	n_c	D – eff	A – eff	G – eff
	17	3	14.2126	9.6712	87.8124
	19	5	13.3809	9.4336	78.5782
CCFD	15	1	44.7163	31.2907	83.6237
	17	3	41.2965	29.7416	74.0181
	19	5	38.1206	27.5997	66.3237

Table 6 Summary of D-, A-, and G- efficiencies for CCCD, CCID, and CCFD over the full second-order model when $k = 4$.

Designs	N	n_c	D – eff	A – eff	G – eff
CCCD	25	1	76.7266	31.6484	60.0000
	27	3	76.4417	55.2876	95.2381
	29	5	73.6352	57.7386	88.6700
CCID	25	1	8.3493	4.5000	60.0000
	27	3	8.3183	5.5556	95.2381
	29	5	8.0129	5.5419	88.6700
CCFD	25	1	44.5233	25.4855	90.9961
	27	3	42.1055	24.2751	84.3091
	29	5	39.8347	22.9789	78.5233

Table 7 Summary of D-, A-, and G- efficiencies for CCCD, CCID, and CCFD over the full second-order model when $k = 5$.

Designs	N	n_c	D – eff	A – eff	G – eff
CCCD	27	1	72.4641	40.4332	88.1890
	29	3	70.5493	53.9803	82.3325
	31	5	67.5089	56.4569	77.0722
CCID	27	1	7.1893	4.6823	88.1890
	29	3	6.9994	5.0292	82.3325
	31	5	6.6977	4.8875	77.0722
CCFD	27	1	42.6899	25.2011	80.5931
	29	3	40.2099	23.7483	75.0571
	31	5	37.9682	22.3915	70.2280

The results in **Tables 5 - 7** corresponding to Kiwu-Lawrence *et al.* [11] demonstrated that:

(1) Increasing the center points tended to reduce values of D-efficiency for the CCCD and CCFD, while there was a slight decrease for the CCID in all cases [11,12].

(2) Replicating the center points increased the A-efficiency for the CCCD but decreased that of the CCFD for all cases; however, the CCID had quite similar efficiency values.

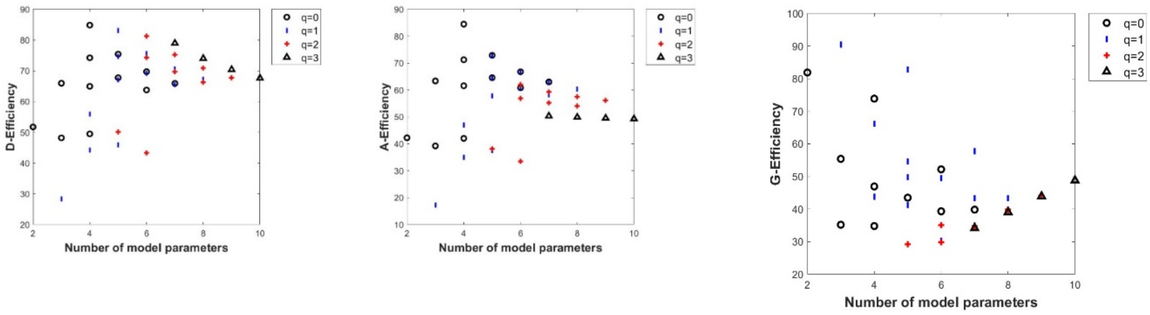
(3) It was clearly observed that the 3 CCD designs had very high G-efficiency values for all cases. Moreover, the replicated center points tended to decrease G-efficiency values for CCFD only. Nevertheless, G-efficiency values for the CCCD and CCFD quite fluctuated.

(4) The columns in **Tables 5 - 7** show that the values of D- and A- efficiencies of the CCCD are larger than those of the CCFD and the CCID, respectively, for all cases. Therefore, the CCCD was superior to the other designs based on D- and A-optimality criteria [11,12].

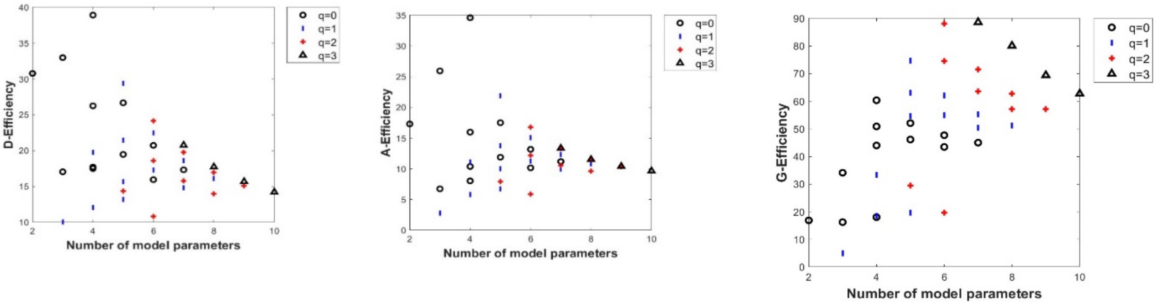
(5) For G-efficiencies, the CCCD was better than the other 2 designs for almost all cases. The exceptions occurred for k = 3 and 4 with n_c = 1, and k = 5 with n_c = 3, where the CCFD dominated the CCCD. However, the CCCD and the CCID were nearly identical resulting in G-efficiency values that were very close when k = 3, 4, and 5, see [7,10,12]. This means the CCCD and the CCID were more efficient than the CCFD when G-optimality criterion was considered.

A set of reduced models

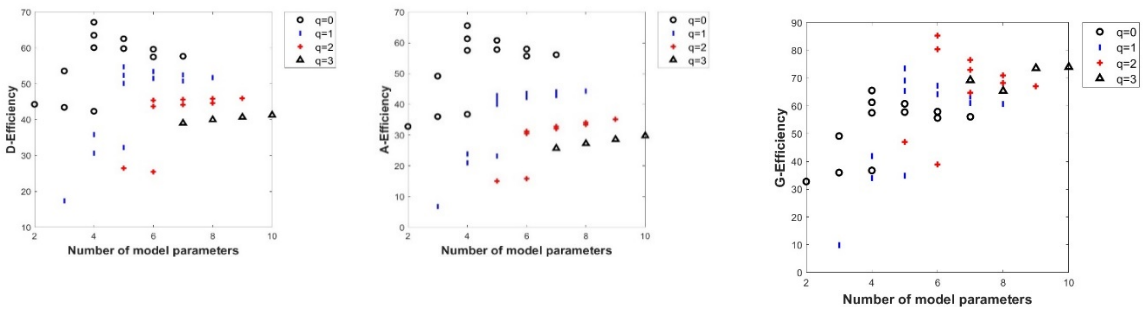
We first considered D-, A-, and G- efficiencies for a collection of reduced models for 3-factor, 4-factor, and 5-factor CCD's.



(a) **(b)** **(c)**



(d) **(e)** **(f)**



(g) **(h)** **(i)**

Figure 1 Plots of reduced model efficiencies for the 3-factor (3 center points) CCCD CCID, and CCFD. Plots (a), (b), and (c) contain the D-, A-, and G- efficiencies for the 3-factor CCCD; (d), (e), and (f) contain the D-, A-, and G- efficiencies for the 3-factor CCID; (g), (h), and (i) contain the D-, A-, and G- efficiencies for the 3-factor CCFD.

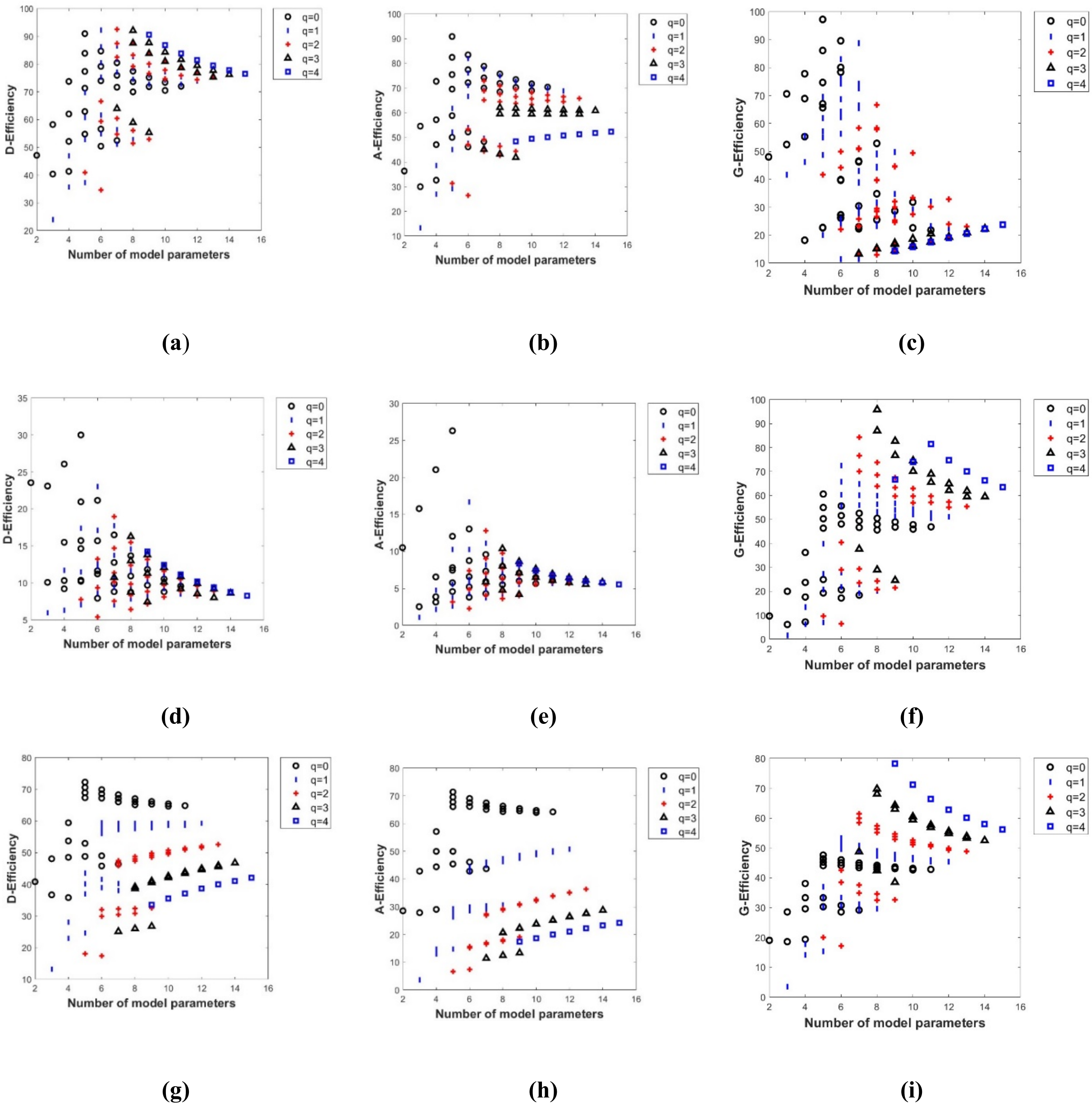


Figure 2 Plots of reduced model efficiencies for the 4-factor (3 center points) CCCD CCID and CCFD. Plots (a), (b) and (c) contain the D-, A-, and G-efficiencies for the 4 -factor CCCD; (d), (e) and (f) contain the D-, A-, and G-efficiencies for the 4-factor CCID; (g), (h) and (i) contain the D-, A-, and G-efficiencies for the 4-factor CCFD.

Figures 1(a) - 1(i), **Figures 2(a) - 2(i)** and **Figures 3(a) - 3(i)** show, respectively, plots of corresponding D-, A-, and G- efficiencies for the 3-factor (3 center-point), the 4-factor (3 center-point) and the 5-factor (3 center-point) CCCD, CCID and CCFD against the number of model parameters (p). The plotting symbol q represents the number of x_i^2 terms in each reduced model. The plots of D-, A-, and G-efficiencies for the CCCD, CCID, and CCFD with 3 center points are similar to those plots of 1 and 5 center points, thus we present only the plots with 3 center points for the 3, 4 and 5 factors.

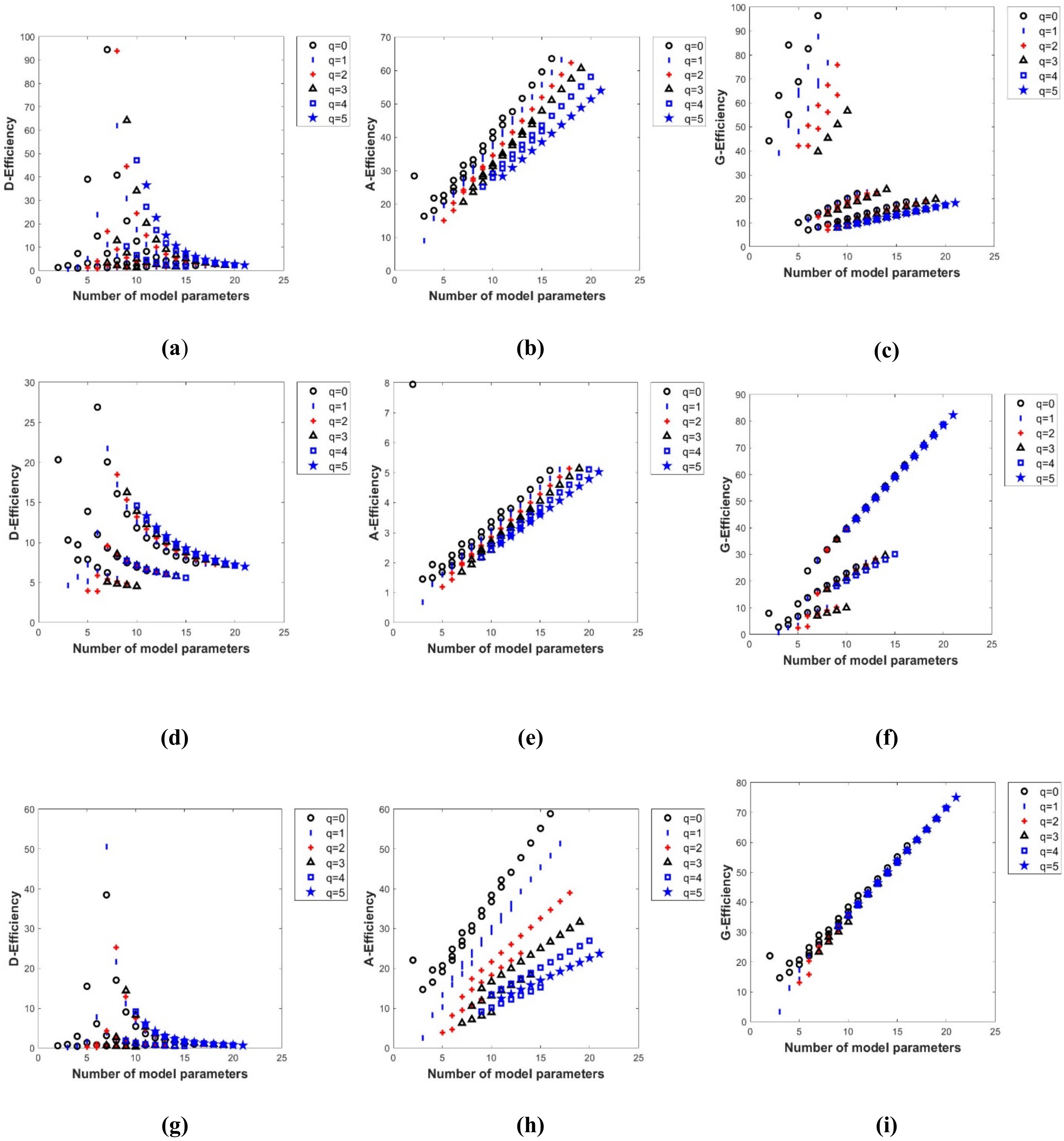


Figure 3 Plots of reduced model efficiencies for the 5-factor (3 center points) CCCD CCID and CCFD. Plots (a), (b) and (c) contain the D-, A-, and G-efficiencies for the 4-factor CCCD; (d), (e) and (f) contain the D-, A-, and G-efficiencies for the 5-factor CCID; (g), (h) and (i) contain the D-, A-, and G-efficiencies for the 4-factor CCFD.

Common patterns were observed for the 3-factor, 4-factor, and 5-factor CCCD, CCID, and CCFD with varying n_c. From these plots, we can conclude that removing an x_i^2 term from a model has varying effect on each design optimality efficiencies. The **Figures 1 - 3** can be observed that:

1) There were similar patterns for all competing designs for D- and A- optimality criteria. Different groups were formed by the values of q, with D- and A-efficiencies increasing considerably with a decrease

in q for the 3-factor, 4-factor, and 5-factor CCCD, CCID, and CCFD with n_c = 1, 3, and 5. In addition, if all square terms are removed ($q = 0$) from a model, D- and A- efficiencies are the largest.

2) With removal of the linear or the interaction terms from a model, D- and A- efficiencies showed little variation. Consequently, D- and A- optimality criteria were more insensitive or robust to deviations from linear and interaction terms than to deviations from a square term. This paper focused only on the square terms for each model when evaluating competing designs for a set of reduced models because there were less factor variables for all models including the identical number of the quadratic terms or the same q when $q > 0$ [16].

3) D- and A-efficiencies for the full model tended to be smaller for CCID and CCFD, whereas D- and A-efficiencies for the pure lineal model tended to be lower for CCCD, compared to the sets of reduced models.

4) For G- optimality criterion, the values of q formed diagonal G-bands for the 3-factor, 4-factor, and 5-factor CCCD, CCID, and CCFD including n_c = 1, 3, and 5. The full model had the higher G-efficiency for CCID and CCFD, indicating better prediction variance capabilities. On the other hand, the full model had the lower G- efficiency for CCCD, compared to the candidate sets of reduced models. The pure linear model had the worst G- efficiency for CCID and CCFD.

Tables 8 - 10 show the results of the comparison ranking for a set of reduced models when k = 3, 4, and 5 with n_c = 1, 3, and 5. To compare each design, the percentage of models for which one design was better than another design was considered for each design optimality criterion (D-, A-, and G- optimality criteria). These percentages were then ranked.

Table 8 Ranking of design performance comparison based on D-, A-, and G- efficiencies for the set of reduced models when k = 3.

Design criteria	N	n_c	CCCD	CCID	CCFD
D	15	1	1, 1, 1	3, 3, 3	2, 2, 2
	17	3	1, 1, 1	3, 3, 3	2, 2, 2
	19	5	1, 1, 1	3, 3, 3	2, 2, 2
A	15	1	1, 1, 1	3, 3, 3	2, 2, 2
	17	3	1, 1, 1	3, 3, 3	2, 2, 2
	19	5	1, 1, 1	3, 3, 3	2, 2, 2
G	15	1	1.5, 2.5, 3	1.5, 1, 1	3, 2.5, 2
	17	3	1.5, 2, 2.5	1.5, 1, 1	3, 3, 2.5
	19	5	1.5, 2, 2.5	1.5, 1, 1	3, 3, 2.5

Note: r_0, r_1, r_2 represent design rank over 43, 31, 15 models for $q \geq 0$, $q \geq 1$, and $q \geq 2$, respectively. Decimal represents the 2 designs with nearly identical rank.

Table 9 Ranking of design performance comparison based on D-, A-, and G- efficiencies for the set of reduced models when k = 4.

Design criteria	N	n_c	CCCD	CCID	CCFD
D	25	1	1, 1, 1	3, 3, 3	2, 2, 2
	27	3	1, 1, 1	3, 3, 3	2, 2, 2
	29	5	1, 1, 1	3, 3, 3	2, 2, 2
A	25	1	1, 1, 1	3, 3, 3	2, 2, 2
	27	3	1, 1, 1	3, 3, 3	2, 2, 2
	29	5	1, 1, 1	3, 3, 3	2, 2, 2
G	25	1	1, 1.5, 2.5	2, 1.5, 1	3, 3, 2.5
	27	3	1.5, 1.5, 2.5	1.5, 1.5, 1	3, 3, 2.5
	29	5	1.5, 1.5, 2	1.5, 1.5, 1	3, 3, 3

Note: r_0, r_1, r_2 represent design rank over 224, 181, 109 models for $q \geq 0$, $q \geq 1$, and $q \geq 2$, respectively. Decimal represents the 2 designs with nearly identical rank.

Three comparison were determined for k = 3 and 4: (i) over the full set of reduced models ($q \geq 0$), (ii) over the set of reduced models with at least 1 square term ($q \geq 1$), and (iii) over the set of reduced models including at least 2 quadratic terms ($q \geq 2$), as seen in **Tables 8** and **9**. For k = 5, 4 comparisons were observed: (i) - (iii) are similar to k = 3 and 4, and (iv) across the set of reduced models with at least 3 square terms ($q \geq 3$), as shown in **Table 10**.

Table 10 Ranking of design performance comparison based on D-, A-, and G- efficiencies for the set of reduced models when k = 5.

Design criteria	N	n_c	CCCD	CCID	CCFD
D	27	1	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
	29	3	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
	31	5	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
A	27	1	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
	29	3	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
	31	5	1, 1, 1, 1	3, 3, 3, 3	2, 2, 2, 2
G	27	1	1, 1.5, 1.5, 2	2, 1.5, 1.5, 1	3, 3, 3, 3
	29	3	1, 1, 1, 1	2, 2, 2, 2	3, 3, 3, 3
	31	5	1, 1, 1, 1	2, 2, 2, 2	3, 3, 3, 3

Note: r_0, r_1, r_2, r_3 represent design rank over 839, 736, 525, 289 models for $q \geq 0, q \geq 1, q \geq 2$ and $q \geq 3$, respectively. Decimal represents the 2 designs with nearly identical rank.

According to the results for $k = 3$ and 4 in **Tables 8** and **9**, each row/column Table entry comprised of 3 ranks (r_0, r_1, r_2). Each rank began with 1 (best) to 3 (worst). Ranks r_0, r_1, and r_2 represent a design's rank relative to the other 2 designs across a collection of reduced models including at least a square term $(q \geq a)$ when $a = 0$, 1, and 2. For any ties, the average ranks were applied. Regarding $k = 5$ in **Table 10**, each row/column Table also contained 4 ranks (r_0, r_1, r_2, r_3). Each rank started with 1 (best) to 4 (worst).

Similar to $k = 3$ and 4, average ranks were applied for tied ranks. It should be noted that when $k = 5$, some CCDs for some reduced model designs appeared as singular design matrices. As a result, the designs with singular matrices could not be computed and D-, A-, and G- optimality criteria could not be examined [17].

As a result, the CCCD was the superior design in terms of D- and A- optimality criteria, followed by the CCFD and the CCID, while the CCID ($k = 3$ and 4), and CCCD ($k = 5$) performed very well with respect to G- optimality criterion over subsets of reduced models for almost all cases. Conversely, the CCFD performed poorly for all cases. In addition, the CCCD had quite a similar performance to the CCID based on G- optimality criterion for $k = 3$ and 4. Generally, G- optimality criterion was preferred to D- and A-optimality criteria because the scaled prediction variance properties have been investigated. These proposed results are consistent with the work in [16,24]. Consequently, we recommend using for identifying the optimal condition in various applications [14,15] because the CCCD has been proven to be a better design over the other 2 designs across the reduced-subset models. Moreover, replicating center runs could be added to take advantage of the lack of model fit.

Based on ranking comparison between the full second-order model and a possible set of reduced models for $k = 3$, 4, and 5, there was some variability, see **Tables 5 - 10**. Even though the CCCD was the best for the full model, it significantly differed from the rank of G-optimality criterion for a collection of reduced models. Similarly, the CCID had the lowest rank for the full model, but the first and second highest rank for reduced-subset models in terms of the G-optimality criterion. This should be of concern to researchers who select an appropriate design based on the second-order model proposed in Eq. (1) and ignore the final model gained by a reduction of model. It is important to note that ranking comparison does not provide the same results when using different design optimality criteria. This is because, while the experimental design may be better under 1 optimality criterion, it may be worse under another 1 or more. Consequently, the practitioners should select an appropriate design which performs very well over the full model and over reduced-subset models, and through the 3 optimality criteria before running the experimental designs.

Conclusions

This study compared the 3 classes of CCDs (CCCD, CCID, and CCFD) over the full second-order response surface model and the subsets of reduced models when $k = 3$, 4, and 5 with n_c = 1, 3, and 5. The results indicated that D-, A-, and G-optimality measures were more robust or sensitive to changes in the number of quadratic terms than to changes in linear and cross-product terms. Therefore, these optimality criteria are robust to variability transmitted into linear and cross-product terms.

The summarized results of the D-, A-, and G- optimality criteria were used in assessing the design performances. The results showed the CCCD was evidently a more efficient design with respect to the D- and A- optimality criteria when the replicated center points were considered over subsets of reduced models for all cases, this implied better precision of parameter estimate. It can also be concluded that the CCCD and the CCID were superior to the CCFD for almost cases over a collection of reduced models based on G-optimality criterion.

According to design comparison ranking, full-model ranks were identical to reduced-model ranks with respect to D-and A- optimality criteria, but the ranks for full model were very different from the ranks for a set of reduced models according to G-optimality criterion. D-and A- optimality criteria focus on parameter estimation, while G-optimality criterion is focused on the variance capabilities. When the experimenter must choose an appropriate design based on 1 or more design optimality criteria, it is critical to determine the best design among optimality criteria designs and across a subset of possible reduced models.

Acknowledgements

The authors thank Rajamangala University of Technology Phra Nakhon (RMUTP), Thailand and Rajamangala University of Technology Isan Surin Campus, Thailand for support this research.

References

[1] GEP Box and KB Wilson. *On the experimental attainment of optimum conditions*. *In*: S Kotz and NL Johnson. (Eds.). Breakthroughs in Statistics. Springer Series in Statistics. Springer, New York, 1992, p. 270-310.

[2] DC Montgomery. *Design and analysis of experiment*. 7th ed. John Wiley & Sons, Ontario, Cannada, 2008.

[3] CM Anderson-Cook, CM Borror and DC Montgomery. Response surface design evaluation and comparison. *J. Stat. Plann. Inference* 2009; **139**, 629-41.

[4] RH Myers, DC Montgomery and CM Anderson-Cook. *Response surface methodology: Process and product optimization using designed experiments*. 4th ed. John Wiley & Sons, New Jersey, 2016.

[5] S Hosamane1, S Bandekar, C Patil and K Joshi. Optimization of removal of cadmium (II) using response surface methodology. *Trends Sci.* 2022; **19**, 2997.

[6] AC Atkinson, AN Donev and RD Tobias. *Optimal experimental design with SAS*. Oxford University Press, Oxford, 2007.

[7] CU Onwuamaeze. Optimal prediction variance properties of some central composite designs in the hypercube. *Commun. Stat. Theory Methods* 2019; **50**, 1911-24.

[8] A Verma, R Kumar, Vinayaka, K Gavhane and B Ahmed. Assessment of prediction capability of response surface designs using graphical methods. *Pharma Innovation J.* 2022; **SP-11**, 113-20.

[9] H Alrweili, S Georgiou and S Stylianou. A new class of second-order response surface designs. *IEEE Access* 2020; **8**, 115123-32.

[10] Z Zhang and B Xiaofeng. Comparison about the three central composite designs with simulation. *In*: Proceedings of the international conference on advanced computer control, Singapore. 2009, p.163-67.

[11] FC Kiwu-Lawrence, LC Kiwu, DC Bartholomew, CP Obite, and AF Chikereuba. Evaluation and comparison of three classes of central composite designs. *Asian J. Probab. Stat.* 2021; **13**, 31-47.

[12] BA Oyejola and JC Nwanya. Selecting the right central composite design. *Int. J. Stat. Appl.* 2015; **5**, 21-30.

[13] JC Nwanya and KCN Dozie. Optimal prediction variance capabilities of inscribed central composite designs. *Asian J. Stat. Probab.* 2020; **8**, 1-8.

[14] HEEA Elhag, A Ajit, and AZ Sulaiman. Optimization and kinetic modelling of total water extracts and water soluble proteins in root extracts of Eurycoma apiculata by microwave assisted extraction. *Mater. Today Proc.* 2020; **31**, 1-8.

[15] O Eguasa and ME Eguasa. Multi-Response optimization of metformin -hydrochloride drug via experimental design techniques applied on the adaptive local linear regression model for response surface methodology. *FUDMA J. Sci.* 2022; **6**, 262-70.

[16] JJ Borkowski and ES Valeroso. Comparison of design optimality criteria of reduced models for response surface designs in the hypercube. *Technometrics* 2001; **43**, 468-77.

[17] B Chomtee and JJ Borkowski. Comparison of response surface designs in spherical region. *In*: Proceedings of the International Conference on Mathematical, World academy of science, engineering and technology, Chicago. 2012, p. 545-8.

[18] B Chomtee. Comparison of five design variables of response surface designs in a spherical region over a set of reduced models. *Agr. Nat. Res.* 2015; **49**, 305-12.
[19] H Chipman. Baysian variable selection with related predictors. *Can. J. Stat.* 1996; **24**, 17-36.
[20] GEP Box and JS Hunter. Multi-factor experimental designs for exploring response surfaces. *Ann. Math. Stat.* 1957; **28**, 195-241.
[21] RH Myers, AI Khuri and Jr Carter. Response surface methodology: 1966-1988. *Technometrics* 1989; **31**, 137-57.
[22] J Kiefer. Optimum experimental design. *J. R. Stat. Soc. Ser. B* 1959; **21**, 272-304.
[23] AU Chukwu and Y Yakubu. Comparison of optimality criteria of reduced models for response surface designs with restricted randomization. *Prog. Appl. Math.* 2012; **4**, 110-26.
[24] AU Chukwu and Y Yakubu. Sensitivity of design optimality criteria to second-order response surface model restrictions and center point replications. *Stud. Math. Sci.* 2012; **4**, 22-9.