Asian Journal of Probability and Statistics

5(2): 1-15, 2019; Article no.AJPAS.51644

ISSN: 2582-0230

Selection of Second Order Models' Design Using D-, A-, E-, T-Optimality Criteria

Wangui Patrick Mwangi[1*], Ayubu Anapapa[1] and Argwings Otieno[1]

[1]Department of Mathematics and Computer Science, University of Eldoret, Eldoret, Kenya.

Authors' contributions

This work was carried out in collaboration among all authors. Authors AO and AA developed the idea, designed the study and oversaw the whole process of the manuscript development. Author WPM managed the literature searches, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2019/v5i230134

Editor(s):

(1) Dr. Seemon Thomas, Associate Professor, Department of Statistics, St. Thomas College, Pala, Mahatma Gandhi University, India.

(2) Dr. Oguntunde, Pelumi Emmanuel, Department of Mathematics, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria.

Reviewers:

(1) Olutosin A. Otekunrin, Federal University of Agriculture, Abeokuta, Nigeria.

(2) Olumide Adesina, Olabisi Onabanjo University, Nigeria.

Complete Peer review History: http://www.sdiarticle4.com/review-history/51644

Received: 25 July 2019
Accepted: 28 September 2019
Published: 03 October 2019

Original Research Article

Abstract

There are numerous designs for fitting second order models that can be used in conjunction with the response surface methodology (RSM) technique in optimization processes, be it in agriculture, industries and so on. Some of the designs include the equiradial, Notz, San Cristobal, Koshal, Hoke, Central Composite and Factorial designs. However, RSM can only be applied in conjunction with a single design at a time. This research aimed at choosing a design out of the most widely employed designs for fitting 2nd order models involving 3 factors for optimization of French beans in conjunction with the RSM technique. The most commonly used designs for second order models were first identified as Box-Behnken designs, Hoke D2 and Hoke D6 designs, 3^k factorial designs, CCD face centred, CCD rotatable and CCD spherical. Design matrices for these 7 designs were formed and augmented with 5 centre points (chosen through lottery methods), and information and optimal design matrices were formed. Then, for each design, the analysis of D-, A- E-, T- optimality (D-Determinant, A-Average Variance, E-Eigen Value and T-Trace) was carried out according to Pukelsheim's definitions. The results were ranked for each criterion and the ranks corresponding to each design were averaged. The design chosen was Hoke D2 with the least average- 1.75. The Hoke D2 was found to be optimal in minimizing the variance of prediction and the most economical design among the seven. The findings are in agreement with other

**Corresponding author: E-mail: patrickmwangi72@yahoo.com, kheteroskedasticity@gmail.com;*

researchers and scientist that a design may be optimal in one criterion but fails in another criterion. Further, Hoke designs are in the class of the economical designs. It is recommended that more optimality criteria be applied and a wide range of designs be involved to see whether the results would still agree with these findings.

Keywords: Optimal; D-; A- E-; T- optimality; analysis; Hoke D2.

1 Introduction

Optimal designs are experimental designs that can be generated on the basis of a specific optimality criterion such as minimum variance, largest Eigen-value among other criteria. Optimal designs have some advantages over non-optimal/sub-optimal experimental designs such as reduced costs of experimentation due to use of fewer experimental runs, accommodation of multiple types of factors, such as process, mixture, and discrete factors and can be used when design-space is constrained (e.g. when factor settings point to infeasibility due to, say wellbeing worries). Optimality is the aspect of minimizing or maximizing something of interest. In the design of experiments, optimality has to do with minimization of variance and/or cost as well as maximizing the precision of estimates. In this case, the optimal design is selected based on D-, A-, E-, T- Optimality criteria and 3 factors of interest that are represented by $X_1, X_2\ and\ X_3$. This is because, given some factors and their levels, there are many designs that can be formed out of that and response surface methodology, RSM, technique can only be applied with a single design at any given time. Other criteria used in optimization include: the I-optimality and G-optimality among others. Therefore, out of the commonly used designs for fitting 2nd order models, one was to be selected based on the D-, A- E-, T- optimality criteria as the widely used optimality criteria in minimizing variances.

According to some sources, after the French beans had originated in Southern Mexico, Guatemala, Costa Rica and Honduras, it had spread to other parts of USA like Florida and Virginia by the year 1492 with the farmers starting to breed the crops by 1890 for cultivars of interest [1]. In developing nations, majority of French bean farmers do not treat the beans as high-input-demanding crops but they channel the limited resources to other crops resulting in very low yields compared to developed nations [2]. Also, competition from other producers has also dealt a heavy blow to the developing nations. However, these challenges have been overcome through various ways including the cooperation between the private and the public sectors to create links that ensure the small-scale producers are not left out or excluded from the networking chains. Such collaboration ensures that the produce meets the safety standards requirements in order to have a space in the international markets [3].

1.1 Literature review

1.1.1 Response Surface Methodology (RSM)

One of the effective ways to solve problems is to conduct experiments. Response surface methodology (RSM) is now used extensively in cases of optimization, designing of products, developing processes, and partly in modern framework for robust parameter design [4]. The method has been applied in a wide range of experiments in different fields and areas in life and has been proven reliable. Central composite design (CCD) is the most widely used design with RSM as was described by Montgomery and Myers although there are other second-order designs [5]. The model generated to optimize the response is $\Lambda = \mathrm{E(y)} = \mathrm{f}\,(X_1, X_2, X_3, X_4, \ldots)$ where y is the response of interest, $X_1, X_2, X_3, X_4, \ldots$ are the independent/explanatory variables or the treatments in the experiment. In general, the response is a function of the controllable variables where a second order model is adequate in achieving the objective of maximization or minimization. Therefore, the model is given by

$$\hat{y} = b_0 + \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} b_{ii} X_i^2 + \sum_i \sum_j b_{ij} X_i X_j \quad (1)$$

In this case, b's are the estimated regression coefficients.

1.1.2 Previous research on French beans

Wondimu and Tana investigated the French bean yield and other parameters based on combinations of nitrogen and phosphorus fertilizers in Ethiopia [6]. The experiment had the aim of determining the optimum rates of nitrogen and phosphorus, fertilizers though the methodology used was not RSM, in order to achieve highest output in terms of yield of French bean. The design applied was the randomized complete block design (RCBD). There were 3 replicates and four bean varieties were tested: Awash Melka, Awash 1, Chercher and Red Wolayita as treatments. The best combination of the involved fertilizers was to be determined by determining the best response of interest. Kiptoo, Arunga and Kimno conducted an experiment involving ten varieties of the commercial French beans. This was done at the University of Eldoret in a glass house [7]. The aim was to help identify those types that are not adversely affected by anthracnose disease (the variety that is resistant to the disease). The experiment was in a controlled environment involving glass-house and involved 10 varieties of the bean; 2 control varieties and 8 others. Four replicates were used and randomized complete block design- RCBD- was the design applied. The results showed that three varieties could be classified as resistant including Julia, Mu_{H13} and Organdia.

Nyasani and others investigated thrips species composition as well as their population density, at KARI-Embu [8]. This was achieved by determining the composition and density on French beans when intercropped as well as when not intercropped. Four replicates were involved and the design employed was RCBD. The results indicated that, the thrips population was increasing over time and was at peak at flowering stage. In conclusion, more species of thrips are hosted when French bean is not intercropped while intercropping helps reduce the species and their effects. However, yield reduces with intercropping.

In conclusion, a lot has been done on French beans but no research has used the design chosen in this research. Most have employed the CCD, RCBD and CRD while very few have used the RSM technique. The researches have revealed that intercropping and pests and diseases as well as climate change can significantly affect the crop performance. The research done at KARI-Embu on French beans has investigated the effects of intercrops on yields and pests' concentration and since intercropping is inevitable in Kariua region, this research seeks to help farmers maximize output while still intercropping their French beans with various types of other crops. It focusses on what farmers deal with in each day activities in their farms in order to help them carry less burdens in terms of labour and resources for optimal gains.

1.1.3 Designs associated with RSM- second order RSM models case

Some of the designs applied in second order RSM include:

i. Central composite designs (CCD)- Its variations include rotatable CCD, spherical CCD, small composite design and the face-centred cube. There are other forms of CCD but cannot be exhausted. CCD for k = 2, each at 2 levels is given by Gunawan [9]:

Combinations	**Co-ordinate**
2^k factorial	(-1, -1), (-1, 1), (1, -1), (1, 1)
$2*k$ axial points	(±α, 0), (0, ±α), where $\alpha = (2^k)^{¼}$
1 centre point	(0.0)

ii. Box-Behnken designs (BBD)- These can be rotatable or nearly rotatable designs involving three level incomplete factorial designs [10]. However, they have limited capability for orthogonal blocking compared to CCD and require 3 levels of each factor.

iii. Hybrid family of designs- These designs are constructed without the aim to satisfy any optimality criteria. They are designed in such a way that the same degree of orthogonality in CCD and regular polyhedron designs is exhibited in them too. I.e. they are near rotatable as well as near minimum point in size as well as enable easy coding [11].

iv. Hoke designs- These are economical designs since they require fewer experiments compared to CCD and BBD. They require 3 or more factors. They are based on partially balanced designs of irregular fractions of 3^k factorial designs [12]. In case the region of interest is cuboidal, then these designs are appropriate.

v. The 3^k factorial designs- In these designs, there are k factors and each is at three levels while this research uses $k = 3$ factors.

All these are classified as standard designs. In case of non-standard situations, these designs are not applicable. Such situations include: unusual sample size requirements, non-standard blocking conditions, variations from standard models and non-normal distribution of the response, among others [4].

For the CCD design, there are the factorial design points, the axial/star points and the centre points. The axial points are at some α-value and -α-value on each axis. The α-values are determined in different ways to yield the various forms of CCD. E.g. If $\alpha = \sqrt{k}$, this yields the spherical CCD, k is the number of factors. The star points are equidistant from the centre as the corner points. If $\alpha = \sqrt[4]{n_f}$, this yields the rotatable CCD, where n_f is the number of factorial points. In this case, the variance of prediction is equal for all points a fixed from the centre, O. So, star points are shifted such that predicted values of the response have equal variance. A special case arises when α = 1, in which star points lie on the boundary. Allowing α = +1and -1 on the cube, the result is a face-centred cube design. It is good to note that, the cubic terms and higher order interactions can't be estimated when the CCD is employed. However, all the higher order interactions can be estimated if 3^k design is employed, but would be too costly because of the large number of runs required [13]. In CCD, factorial points are used in the estimation of first order terms as well as the interaction terms. The squared terms are estimated with the help of the star/axial points (axial terms are 2*k). The lack of fit is tested using pure error that is estimated using the centre points. The centre points also help in estimating the squared terms in the model.

Rotatability of a design depends on choice of α, and if the prediction variance depends only on distance of the design point from the centre of the design, then a design is said to be rotatable. Box-Behnken designs avoid all corner points and star points and end up eliminating the extreme treatment combinations in the design. So, they define the boundaries in the experiment. This design can be applied when extreme points of the experimental region are a problem. Both Box-Behnken and CCD can work well though they have different structures and a quadratic model can be fit with Box-Behnken [13].

1.1.4 Optimal designs and D-, A-, E-, T- optimality

An optimal design is selected based on some criteria. Some of the criteria used in optimization include: D-optimality, A-optimality, E-optimality, T-optimality, I-optimality and G-optimality, MV-optimality [14] among others. Let X be the model/design matrix (n*p model matrix constructed by expanding the design matrix to model form) and $X^T X$ be the information matrix. The matrix M = $(\frac{X^T X}{N})$ is called the moment matrix, N is the total number of runs. The N seeks to penalize larger designs. The moment matrix determines the estimated response surface statistical properties. D-optimal design is the one that has the maximum determinant in the information matrix among all the possible designs. This results in minimization of the generalized variance of the parameter estimates. A-optimal design is the one that has the minimum trace in the inverse of the information matrix among all the possible designs. It helps minimize the average variance of the estimated regression coefficients. It doesn't make use of covariates. Note that, replication of star points lowers the optimality of D and G in CCDs [15]. E-optimal design is the one that has the largest Eigen value in the information matrix among all the possible designs. T-optimal design is the one that has the maximum trace in the information matrix among all the possible designs. Note that, exact designs are actually the designs for a specified number of runs [16] and all designs are exact designs in practice [17].

According to Pukelsheim [18],

$$\text{D-optimal, } \phi_0(\text{C}) \text{ refers to determinant criterion} = (\det(C))^{1/p} \quad (2)$$

$$\text{A-optimal, } \phi_{-1}(\text{C}) \text{ refers to average-variance criterion} = (\frac{1}{p}*\text{trace}(C)^{-1})^{-1} \quad (3)$$

$$\text{E-optimal, } \phi_{-\infty}(\text{C}) \text{ refers to smallest-eigenvalue criterion} = \lambda_{min}(C) \quad (4)$$

$$\text{T-optimal, } \phi_1(\text{C}) \text{ refers to trace criterion} = \frac{1}{p}*\text{trace}\ (C) \quad (5)$$

In all these cases, C is the information matrix of the optimal design of interest defined as C = $(K^{'}M^{-}K)^{-1}$; M is the moment matrix, M^{-} is the generalized inverse of matrix M, K is the submatrix of parameters of interest and p is the number of parameters.

On the other hand, [19] define optimal designs in some varied form. First, we note that, some optimizing criteria aim at estimating good parameters of the model while some bring about good prediction in the region of the design. For moment matrix M = $(\frac{X^TX}{N})$, |M| = $(\frac{|X^TX|}{N^p})$, $M^{-1} = N(X^TX)^{-1}$ is the scaled dispersion matrix. The $|X^TX|$ and the square of the volume of the confidence region on the regression coefficients are inversely proportional to each other. A D-optimal design maximizes $|M| = (\frac{|X^TX|}{N^p})$. The A-optimal design minimizes the $tr(\frac{X^TX}{N})^{-1}$ = $\min\{tr(N(X^TX)^{-1})\}$. A-optimal designs are such that min[A] = min $[tr[(X^TX)^{-1}]]$ = min $[\Sigma\lambda_j]$. For the case of D-optimal designs, we have the Min[D] = min $[|X^TX|]$ = min$[\prod(\frac{1}{\lambda_j})]$, λ_j is the j[th] Eigen value of $(X^TX)^{-1}$. The E-optimal designs have min[E] = min $[\lambda_j]$ and this means that they minimize the largest Eigen value. One can see that, the design chosen becomes more suitable with increase in D but with decrease in A and E [20].

1.1.5 Multiple response optimization

In many cases in life and in practice, the researcher is usually interested in several responses and not just one [19]. For example, one may be fitting a model that is maximizing crop (say maize) output while interested in investigating the plant height and base diameter based on some factors. One way of optimizing multiple responses is to overlay the contour plots for each response when there are few process factors. Examining the overlaid contour plots can help determine appropriate operating conditions. Therefore, the researcher determines the necessary operating conditions that optimizes all the responses concurrently from the plot of overlaid contour plots. Note that, one needs a lot of guess-works in determining the factors to hold constant and the levels to select for the best view of the surface since there is no ‘best’ formal way of optimizing multiple responses [19].

2 METHODOLOGY

2.1 Materials

The material needed for the achievement of the objective was computer hardware and software. The Ms-Excel and R software were used in creating the necessary matrices and the analysis of those matrices.

2.2 The design and methods

The CCD (there are several variations of CCD), the Box-Behnken designs (BBD), the Hoke designs and the 3^k factorial designs are the most widely used designs for second order models in optimization. Specifically, the seven designs most widely used are face-centred CCD, spherical CCD and rotatable CCD, Hoke D2, Hoke D6, Box-Behnken designs and 3^k factorial designs, and one was chosen. The method used was the analysis of D-, A-, E- and T- optimality criteria (D- Determinant, A- Average Variance, E- Eigen Value and T- Trace) in coming up with only one design out of the seven designs based on 3 factors.

2.3 Procedures

For each design, the design matrix **X** involving all the factors and interactions as well as augmented with additional factor denoted as X_0 for estimating the intercept was created. The design matrix **X** for each of the seven designs was constructed using the standard way of listing the factor levels. Then, each design was augmented with 5 centre points based on the lottery method employed in simple random sampling. According to literature, number of centre points should be between 3 and 5, and hence through lottery, 5 was chosen. The X matrix was of the form $\mathbf{X} = (X_0, X_1, X_2, X_3, X_1^2, X_2^2, X_3^2, X_1X_2, X_1X_3, X_2X_3)$ where X_0 is a column of units while the rest of $X_i's$ and $X_iX_j's$ are factor levels. Note that, X_1X_2, X_1X_3, X_2X_3 are denoted as X_{12}, X_{13}, X_{23} respectively in the matrices. From this matrix, the information matrix of the design was constructed as X^TX, which makes it become a square matrix. The moment matrix of the design was obtained as $M = \frac{\mathrm{X^TX}}{N}$ where N is the number of runs of the design. The information matrix for the optimal design was computed as $C = (K^TM^-K)^{-1}$ where K is the identity matrix representing the sub-system matrix of the parameters of interest. This means that the matrix $C = M$ since K is identity.

Each of the seven designs mentioned here, in form of C, was subjected to all the criteria, one criterion at a time. That's to say, determinant, trace, average variance and Eigen-value were computed for each matrix C according to [18] definition. The value and score of every design in each criterion was noted- meaning that, the scores were ranked for all the designs. In other words, for the determinant criterion, the determinant of each matrix was computed. Then the smallest determinant was ranked 1 while the largest was ranked 7. This process was repeated for the rest of the criteria. Then, the ranks for each design were averaged. In the end, the design with the least average rank-score was the design employed. This ensured that the design chosen averaged the optimality of all criteria.

3 Results and Discussion

Design matrices were generated and data fed into Ms-Excel and R. In R, the necessary analysis like finding moment and information matrices, their determinants, Eigen values, traces, average-variances and inverses was done. Averaging and ranking of these results was done in Ms-Excel.

Seven designs for second order models were selected as the most widely used. In all these designs, there were only three factors namely; manure, water and spacing. These were coded as X_1, X_2 and X_3 for manure, water and spacing respectively for easier fitting of the models. The Hoke D2 designs are formed as a result of combinations of factor levels from the following sets: (-1, -1, -1), (1, 1, -1), (1, -1, -1) and (-1, 0, 0) for D2 with 3 factors and (-1, -1, -1, -1), 1, 1, 1, -1), (1, 1, -1, -1) and (-1, 0, 0, 0) for D2 with 4 factors. For Hoke D6, the sets (1, 0, 0) for 3 factors and (1, 0, 0, 0) for 4 factors are added.

3.1 Designs of interest

The factor levels corresponding to each of these seven designs without augmenting with the centre points are shown below.

Table 1. Factor level combinations for Box-Behnken, face centred, rotatable and spherical designs

Box-Behnken design			CCD face centred			CCD rotatable			CCD spherical		
X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
-1	-1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	0	1	-1	-1	1	-1	-1	1	-1	-1
1	-1	0	-1	1	-1	-1	1	-1	-1	1	-1
1	1	0	1	1	-1	1	1	-1	1	1	-1
-1	0	-1	-1	-1	1	-1	-1	1	-1	-1	1
-1	0	1	1	-1	1	1	-1	1	1	-1	1
1	0	-1	-1	1	1	-1	1	1	-1	1	1
1	0	1	1	1	1	1	1	1	1	1	1
0	-1	-1	-1	0	0	-1.68	0	0	-1.73	0	0
0	-1	1	1	0	0	1.68	0	0	1.73	0	0
0	1	-1	0	-1	0	0	-1.68	0	0	-1.73	0
0	1	1	0	1	0	0	1.68	0	0	1.73	0
			0	0	-1	0	0	-1.68	0	0	-1.73
			0	0	1	0	0	1.68	0	0	1.73

Source: Myers, Montgomery and Cook, Response Surface Methodology, 3rd Ed. 2009

Table 2. Factor level combinations for 3^k Factorial, Hoke D2 and D6 designs

3^k Factorial						Hoke D2			Hoke D6		
X_1	X_2	X_3				X_1	X_2	X_3	X_1	X_2	X_3
-1	-1	-1	1	0	0	-1	-1	-1	-1	-1	-1
0	-1	-1	-1	1	0	1	1	-1	1	1	-1
1	-1	-1	0	1	0	1	-1	1	1	-1	1
-1	0	-1	1	1	0	-1	1	1	-1	1	1
0	0	-1	-1	-1	1	1	-1	-1	1	-1	-1
1	0	-1	0	-1	1	-1	1	-1	-1	1	-1
-1	1	-1	1	-1	1	-1	-1	1	-1	-1	1
0	1	-1	-1	0	1	-1	0	0	-1	0	0
1	1	-1	0	0	1	0	-1	0	0	-1	0
-1	-1	0	1	0	1	0	0	-1	0	0	-1
0	-1	0	-1	1	1				1	1	0
1	-1	0	0	1	1				1	0	1
-1	0	0	1	1	1				0	1	1
0	0	0									

Source: Myers, Montgomery and Cook, Response Surface Methodology, 3rd Ed. 2009

3.2 Information matrices for the designs of interest

The following are the information matrices for each of these seven designs. Each design has been augmented with the centre points at this stage of comparison.

Information Matrix for Box-Behnken Design

$$\begin{bmatrix} & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\ X_0 & 17 & 0 & 0 & 0 & 8 & 8 & 8 & 0 & 0 & 0 \\ X_1 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_2 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_3 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1^2 & 8 & 0 & 0 & 0 & 8 & 4 & 4 & 0 & 0 & 0 \\ X_2^2 & 8 & 0 & 0 & 0 & 4 & 8 & 4 & 0 & 0 & 0 \\ X_3^2 & 8 & 0 & 0 & 0 & 4 & 4 & 8 & 0 & 0 & 0 \\ X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Information Matrix for CCD- Face Centred Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	19	0	0	0	10	10	10	0	0	0
X_1	0	10	0	0	0	0	0	0	0	0
X_2	0	0	10	0	0	0	0	0	0	0
X_3	0	0	0	10	0	0	0	0	0	0
X_1^2	10	0	0	0	10	8	8	0	0	0
X_2^2	10	0	0	0	8	10	8	0	0	0
X_3^2	10	0	0	0	8	8	10	0	0	0
X_{12}	0	0	0	0	0	0	0	8	0	0
X_{13}	0	0	0	0	0	0	0	0	8	0
X_{23}	0	0	0	0	0	0	0	0	0	8

Information Matrix for CCD- Rotatable Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	19	0	0	0	13.6448	13.6448	13.6448	0	0	0
X_1	0	13.6448	0	0	0	0	0	0	0	0
X_2	0	0	13.6448	0	0	0	0	0	0	0
X_3	0	0	0	13.64480	0	0	0	0	0	0
X_1^2	13.6448	0	0	0	23.9319	8	8	0	0	0
X_2^2	13.6448	0	0	0	8	23.9319	8	0	0	0
X_3^2	13.6448	0	0	0	8	8	23.9319	0	0	0
X_{12}	0	0	0	0	0	0	0	8	0	0
X_{13}	0	0	0	0	0	0	0	0	8	0
X_{23}	0	0	0	0	0	0	0	0	0	8

Information Matrix for CCD- Spherical Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	19	0	0	0	13.9858	13.9858	13.9858	0	0	0
X_1	0	13.9858	0	0	0	0	0	0	0	0
X_2	0	0	13.9858	0	0	0	0	0	0	0
X_3	0	0	0	13.9858	0	0	0	0	0	0
X_1^2	13.9858	0	0	0	25.9149	8	8	0	0	0
X_2^2	13.9858	0	0	0	8	25.9149	8	0	0	0
X_3^2	13.9858	0	0	0	8	8	25.9149	0	0	0
X_{12}	0	0	0	0	0	0	0	8	0	0
X_{13}	0	0	0	0	0	0	0	0	8	0
X_{23}	0	0	0	0	0	0	0	0	0	8

Information Matrix for 3^K Factorial Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	32	0	0	0	18	18	18	0	0	0
X_1	0	18	0	0	0	0	0	0	0	0
X_2	0	0	18	0	0	0	0	0	0	0
X_3	0	0	0	18	0	0	0	0	0	0
X_1^2	18	0	0	0	18	12	12	0	0	0
X_2^2	18	0	0	0	12	18	12	0	0	0
X_3^2	18	0	0	0	12	12	18	0	0	0
X_{12}	0	0	0	0	0	0	0	12	0	0
X_{13}	0	0	0	0	0	0	0	0	12	0
X_{23}	0	0	0	0	0	0	0	0	0	12

Information Matrix for Hoke D2 Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	15	−2	−2	−2	8	8	8	−1	−1	−1
X_1	−2	8	−1	−1	−2	−1	−1	−1	−1	−1
X_2	−2	−1	8	−1	−1	−2	−1	−1	−1	−1
X_3	−2	−1	−1	8	−1	−1	−2	−1	−1	−1
X_1^2	8	−2	−1	−1	8	7	7	−1	−1	−1
X_2^2	8	−1	−2	−1	7	8	7	−1	−1	−1
X_3^2	8	−1	−1	−2	7	7	8	−1	−1	−1
X_{12}	−1	−1	−1	−1	−1	−1	−1	7	−1	−1
X_{13}	−1	−1	−1	−1	−1	−1	−1	−1	7	−1
X_{23}	−1	−1	−1	−1	−1	−1	−1	−1	−1	7

Information Matrix for Hoke D6 Design

	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_{12}	X_{13}	X_{23}
X_0	18	0	0	0	10	10	10	0	0	0
X_1	0	10	0	0	0	0	0	0	0	−1
X_2	0	0	10	0	0	0	0	0	−1	0
X_3	0	0	0	10	0	0	0	−1	0	0
X_1^2	10	0	0	0	10	8	8	0	0	−1
X_2^2	10	0	0	0	8	10	8	0	−1	0
X_3^2	10	0	0	0	8	8	10	−1	0	0
X_{12}	0	0	0	−1	0	0	−1	8	−1	−1
X_{13}	0	0	−1	0	0	−1	0	−1	8	−1
X_{23}	0	−1	0	0	−1	0	0	−1	−1	8

According to Pukelsheim, [18], the information matrix for the optimal design would be defined as C = $(K'M^{-}K)^{-1}$, M = $(\frac{X'X}{N})$ and **X** is the design matrix while N is the number of runs. M^{-} is the generalized inverse of moment matrix M. In this research, all of the 10 parameters in the second order model are of interest. Therefore, $K = Ip$, which means K becomes identity matrix if and only if one is interested in all the parameters. In this case,

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Substituting K and M in the formula above yields the C = M hence our information matrix C of the optimal design is the same as the moment matrix M. Alternatively, C can be computed as follows: $C = LM^{-}L'$ where L is the left inverse of matrix K. Since K are identity matrices for this case, L are identity matrices too. The two cases yield C = M.

3.3 Information matrices for the optimal designs of interest

The information matrices, C's, for the seven designs (optimal designs) are as below, after rounding each matrix off to 4 decimal places where there's need.

Information Matrix for Optimal Box-Behnken Design

$$\frac{1}{17}\begin{bmatrix} & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\ X_0 & 17 & 0 & 0 & 0 & 8 & 8 & 8 & 0 & 0 & 0 \\ X_1 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_2 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_3 & 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1^2 & 8 & 0 & 0 & 0 & 8 & 4 & 4 & 0 & 0 & 0 \\ X_2^2 & 8 & 0 & 0 & 0 & 4 & 8 & 4 & 0 & 0 & 0 \\ X_3^2 & 8 & 0 & 0 & 0 & 4 & 4 & 8 & 0 & 0 & 0 \\ X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Information Matrix for Optimal CCD- Face Centred Design

$$\frac{1}{19}\begin{bmatrix} & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\ X_0 & 19 & 0 & 0 & 0 & 10 & 10 & 10 & 0 & 0 & 0 \\ X_1 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_2 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_3 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1^2 & 10 & 0 & 0 & 0 & 10 & 8 & 8 & 0 & 0 & 0 \\ X_2^2 & 10 & 0 & 0 & 0 & 8 & 10 & 8 & 0 & 0 & 0 \\ X_3^2 & 10 & 0 & 0 & 0 & 8 & 8 & 10 & 0 & 0 & 0 \\ X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$$

Information Matrix for Optimal CCD- Rotatable Design

$$\frac{1}{19}\left[\begin{array}{ccccccccccc}
 & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\
X_0 & 19 & 0 & 0 & 0 & 13.6448 & 13.6448 & 13.6448 & 0 & 0 & 0 \\
X_1 & 0 & 13.6448 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_2 & 0 & 0 & 13.6448 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_3 & 0 & 0 & 0 & 13.64480 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_1^2 & 13.6448 & 0 & 0 & 0 & 23.9319 & 8 & 8 & 0 & 0 & 0 \\
X_2^2 & 13.6448 & 0 & 0 & 0 & 8 & 23.9319 & 8 & 0 & 0 & 0 \\
X_3^2 & 13.6448 & 0 & 0 & 0 & 8 & 8 & 23.9319 & 0 & 0 & 0 \\
X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\
X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\
X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8
\end{array}\right]$$

Information Matrix for Optimal CCD- Spherical Design

$$\frac{1}{19}\left[\begin{array}{ccccccccccc}
 & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\
X_0 & 19 & 0 & 0 & 0 & 13.9858 & 13.9858 & 13.9858 & 0 & 0 & 0 \\
X_1 & 0 & 13.9858 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_2 & 0 & 0 & 13.9858 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_3 & 0 & 0 & 0 & 13.9858 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_1^2 & 13.9858 & 0 & 0 & 0 & 25.9149 & 8 & 8 & 0 & 0 & 0 \\
X_2^2 & 13.9858 & 0 & 0 & 0 & 8 & 25.9149 & 8 & 0 & 0 & 0 \\
X_3^2 & 13.9858 & 0 & 0 & 0 & 8 & 8 & 25.9149 & 0 & 0 & 0 \\
X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\
X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\
X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8
\end{array}\right]$$

Information Matrix for Optimal 3^K Factorial Design

$$\frac{1}{32}\left[\begin{array}{ccccccccccc}
 & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\
X_0 & 32 & 0 & 0 & 0 & 18 & 18 & 18 & 0 & 0 & 0 \\
X_1 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_2 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_3 & 0 & 0 & 0 & 18 & 0 & 0 & 0 & 0 & 0 & 0 \\
X_1^2 & 18 & 0 & 0 & 0 & 18 & 12 & 12 & 0 & 0 & 0 \\
X_2^2 & 18 & 0 & 0 & 0 & 12 & 18 & 12 & 0 & 0 & 0 \\
X_3^2 & 18 & 0 & 0 & 0 & 12 & 12 & 18 & 0 & 0 & 0 \\
X_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 & 0 & 0 \\
X_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 & 0 \\
X_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12
\end{array}\right]$$

Information Matrix for Optimal Hoke D2 Design

$$\frac{1}{15}\begin{bmatrix} & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\ X_0 & 15 & -2 & -2 & -2 & 8 & 8 & 8 & -1 & -1 & -1 \\ X_1 & -2 & 8 & -1 & -1 & -2 & -1 & -1 & -1 & -1 & -1 \\ X_2 & -2 & -1 & 8 & -1 & -1 & -2 & -1 & -1 & -1 & -1 \\ X_3 & -2 & -1 & -1 & 8 & -1 & -1 & -2 & -1 & -1 & -1 \\ X_1^2 & 8 & -2 & -1 & -1 & 8 & 7 & 7 & -1 & -1 & -1 \\ X_2^2 & 8 & -1 & -2 & -1 & 7 & 8 & 7 & -1 & -1 & -1 \\ X_3^2 & 8 & -1 & -1 & -2 & 7 & 7 & 8 & -1 & -1 & -1 \\ X_{12} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 7 & -1 & -1 \\ X_{13} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 7 & -1 \\ X_{23} & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 7 \end{bmatrix}$$

Information Matrix for Optimal Hoke D6 Design

$$\frac{1}{18}\begin{bmatrix} & X_0 & X_1 & X_2 & X_3 & X_1^2 & X_2^2 & X_3^2 & X_{12} & X_{13} & X_{23} \\ X_0 & 18 & 0 & 0 & 0 & 10 & 10 & 10 & 0 & 0 & 0 \\ X_1 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ X_2 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ X_3 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & -1 & 0 & 0 \\ X_1^2 & 10 & 0 & 0 & 0 & 10 & 8 & 8 & 0 & 0 & -1 \\ X_2^2 & 10 & 0 & 0 & 0 & 8 & 10 & 8 & 0 & -1 & 0 \\ X_3^2 & 10 & 0 & 0 & 0 & 8 & 8 & 10 & -1 & 0 & 0 \\ X_{12} & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 8 & -1 & -1 \\ X_{13} & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 8 & -1 \\ X_{23} & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & -1 & 8 \end{bmatrix}$$

3.4 Analysis of D-, A-, E-, and T- optimality criteria

From these information matrices, C's, the optimality criteria D-, A-, E-, and T- can be applied.

Determinant = $(\det(C))^{1/p}$, Average-variance = $(\frac{1}{p}*\text{trace } (C)^{-1})^{-1}$, Eigen-value = $\lambda_{min}(C)$ and Trace = $\frac{1}{p}*\text{trace } (C)$. The table below shows these values for all the seven designs.

The values in red in brackets are the ranks. The last column shows the averages of the ranks for each design.

Table 3. The optimality values, ranks & averages for the seven optimal designs

Design	D-Opt.	A-Opt.	E-Opt.	T-Opt.	Average
Box-Behnken	0.3403 (2)	0.2887 (4)	0.1550 (4)	0.4529 (1)	2.75
CCD Face Centred	0.3812 (3)	0.2760 (3)	0.1053 (3)	0.5421 (2)	2.75
CCD Rotatable	0.6357 (6)	0.5186 (6)	0.1905 (6)	0.8196 (6)	6.00
CCD Spherical	0.6587 (7)	0.5307 (7)	0.1927 (7)	0.8563 (7)	7.00
3^K Factorial	0.4054 (5)	0.3293 (5)	0.1695 (5)	0.5500 (3)	4.50
Hoke D2	***0.3190(1)***	***0.1700 (1)***	***0.0585 (1)***	***0.5600 (4)***	***1.75***
Hoke D6	0.3878 (4)	0.2723 (2)	0.1032 (2)	0.5667 (5)	3.25

Source: Authors

The design with the minimum average is the best design compared to the rest. In this case, Hoke D2 with an average of 1.75 was chosen as the optimal design. This design could be applied in the field experiments throughout the entire period of the research. It is the design with the minimum variance and it is very economical compared to the rest since it has the least number of runs among the seven designs.

3.5 Discussion

According to other researchers, a design may be optimal in one or more criteria but can fail to be optimal in another criterion/criteria. This can be seen in Hoke D2 in which it is optimal in the first three criteria (scoring 1 in D-, A- and E- criteria) but is not optimal in the 4th criteria (scoring 4th in T- criterion). Again, the Box-Behnken design is emerging as optimal (scoring 1 in T- criterion) but fails to be optimal in the rest of the criteria. Therefore, this research and all analysis involving computations are in agreement with other researchers' findings. Not only so, but Hoke D2 is classified by scientists under the economical class of designs and the findings of this research are in agreement too as can be seen in the number of runs (10 runs only) involved. These runs are the fewest among all the seven designs chosen. The runs being the fewest means that the cost is lowest too hence the phrase "economical class of designs".

4 Conclusion and Recommendation

4.1 Summary and Conclusion

There are a wide range of second order models' designs that can be used together with the response surface methodology (RSM) technique. Seven designs for fitting second order models were found to be common in practice. However, although the equiradial designs, Notz designs, San Cristobal Designs and Koshal designs among others- including hybrid/ Roquemore designs- are also for fitting 2nd order models, they are not widely employed in practice for some reasons. Out of the seven most commonly employed designs for second order models, one was chosen through the criteria of analysing the D-, A-, E- and T- optimality. Apart from these criteria in optimality, there are other criteria that can be used such as the I-optimality, G-optimality and so on. The D-, A-, E- and T- optimality was used since they are the most commonly used optimality criteria by most researchers, scientists and experimenters because they are easy to compute and reliable in minimizing variances. The design with the smallest value on average was chosen as the design to be used together with the RSM technique. In this case, RSM has been identified to be employed because of the optimization required as well as identification of factor levels that can bring about the optimal results. It is now the standard tool used in the analysis of data obtained from experiments meant to optimize responses of interest. Further, it was the only practical technique found that could help the researcher arrive at the desired results and it is now used extensively in cases of optimization. The analysis of D-, A-, E- and T-optimality showed that Hoke D2 was ranked first in D-, A- and E- criteria while Box-Behnken was first in T- criterion. The CCD Spherical and CCD Rotatable designs emerged last in all the criteria because they were placed 7th and 6th in all criteria respectively. Hoke D2 was chosen as the best design in minimizing the variance of prediction since it had the least average. This method of choosing a design ensures that the design so chosen averages all the criteria involved. The design was also found to have the least number of runs (10 runs) compared to the rest of the designs which had more than 10 runs. This means it is the most economical design among the considered designs. All the results concur with other researches. In general, Hoke D2 is the best design in minimizing variance of prediction and is the most economical design among the seven widely applied designs based on D-, A-, E-, T- optimality criteria and 3 independent variables. Based on all these, Hoke D2 is the design of choice and should be applied practically in the field by researchers.

4.2 Recommendation

A wide range of optimality criteria to be employed in choosing design. This is because there are so many criteria that can be used in optimality of designs and hence a way of not limiting the criteria to just D-, A-, E- and T- optimality is required. Even more designs can be included apart from the seven designs considered in this research. This means to include the designs classified as 'not widely' employed such as hybrid, Cristobal etc.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Agricultural Alternatives. Snap Bean Production. PennState- College of Agricultural Sciences, Agrucultural Research and Cooperative Extension; 2002.
Available:http://agalternatives.aers.psu.edu
(Accessed on 26th April, 2019)

[2] OECD. Common bean (*Phaseolus vulgaris*). In Safety Assessment of Transgenic Organisms in the Environment, Volume 6: OECD Consensus Documents, OECD Publishing, Paris. 2016;187-198.
DOI: https://doi.org/10.1787/9789264253421-7-en
(Accessed in April 26, 2019)

[3] Okello JJ, Narrod C, Roy D. Food safety requirements in African green bean exports and their impact on small farmers. International Food Policy Research Institute, Washington DC, USA. IFPRI Discussion Paper 00737. 2007;1-6.

[4] Johnson RT, Montgomery DC. Choice of second-order response surface designs for logistic and Poisson regression models. Int. J. Experimental Design and Process Optimisation. 2009;1(1).

[5] Myers RH, Montgomery DC. Response surface methodology: Process and product optimization using designed experiments. New York: John Wiley & Sons; 1995.

[6] Wondimu W, Tana T. Yield response of common bean (*Phaseolus vulgaris* L.) varieties to combined application of nitrogen and phosphorus fertilizers at Mechara, Eastern Ethiopia. J Plant Biol Soil Health- Avens Publishing Group. 2017;4(2):7.

[7] Kiptoo GJ, Arunga EE, Kimno SK. Evaluation of French bean (*Phaseolus vulgaris* L.) varieties for resistance to anthracnose. Journal of Experimental Agriculture International. 2018;27(4):1-7.

[8] Nyasani JO, Meyhöfer R, Subramaniana S, Poehling HM. Effect of intercrops on thrips species composition and population abundance on French beans in Kenya. Entomologia Experimentalis et Applicata. 2012;142:236–246.
DOI: 10.1111/j.1570-7458.2011.01217.x

[9] Gunawan A, Chuin LH. Second order-response surface model for the automated parameter tuning problem. 2014 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM): 9-12 December 2014, Selangor. 2014;652- 656. Research Collection School of Information Systems.
Available:https://ink.library.smu.edu.sg/sis_research/2672

[10] Ferreira SLC, Bruns RE, Ferreira HS, Matos GD, David JM, Brandao GC, Santos WNL. Box-Behnken design: An alternative for the optimization of analytical methods. Analytica Chimica Acta. 2007;597:179–186.

[11] Roquemore KG. Hybrid designs for quadratic response surfaces. Technometrics. 1976;18(4):419-423.
(Retrieved from JSTOR Databases, 9th October, 2018)

[12] Neifar M, Kamoun A, Jaouani A, Ghorbel RE, Chaabouni SE. Application of Asymetrical and Hoke designs for optimization of laccase production by the White-Rot fungus fomes fomentarius in solid state fermentation. Enzyme Research. 2011;12. Article ID: 368525, (SAGE Hindawi Access to Research).
DOI: 10.4061/2011/368525

[13] Penn State Eberly College of Science. STA 503- Design of experiments. Department of Statistics Online Programs; 2018.
(Accessed in October, 2018)

[14] Das A. An introduction to optimality criteria and some results on optimal block design. Design Workshop Lecture Notes ISI, Kolkata. 2002;1-21.

[15] Oyejola BA, Nwanya JC. Selecting the right central composite design. International Journal of Statistics and Applications. 2015;5(1):21-30.

[16] Jacob E, Boon JE. Generating exact D-optimal designs for polynomial models. SpringSim '07. 2007;2.

[17] Xiang Feng W. Optimal designs for segmented polynomial regression models and web-based implementation of optimal design software. The Graduate School at Stony Brook University; 2007.

[18] Pukelsheim F. Optimal design of experiments. John Wiley & Sons, Inc., New York. 1993;Chapter 6:135.

[19] Myers HR, Montgomery DC, Cook CMA. Response surface methodology- process and product optimization using designed experiments. John Wiley & Sons, Inc., Hoboken, New Jersey. 3rd Ed; 2009.

[20] Frank IE, Todeschint R. The data analysis handbook. Data Handling in Science and Technology. 1994;14.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/51644