Efficiency of Repaired Resolution Central Composite Design to the Standard Central Composite Design

Anyanwu, Bridget. Chisa. [1,*]**; & Iwundu, Mary Pascal**[2]

[1,2]Department of Statistics, University of Port Harcourt, Rivers State

[1,2]Correspondence: anyanwubridget7@gmail.com

mary.iwundu@uniport.edu.ng

Abstract

Repaired resolution acts as an essential tool in reducing the size of experimental runs. In this paper, the efficiencies of Repaired Resolution Central Composite Design (RRCCD) compared to the Standard (Box-Wilson) Central Composite Design (CCD) in seven design variables were considered. A second-order response surface model was employed. Two cases of RRCCD, two of Box-Wilson central composite and one case of Sequential 3-Level Design (S3LD) are considered. D- and A-efficiencies are studied for these designs; run size efficiencies are also considered. For the RRCCD, the quarter fraction of resolution IV with a run size of 58 has D-efficiency of 82.24% and A-efficiency of 58.79%. The quarter fraction of resolution III with a run size of 54 has 81.89% D-efficiency and 59.54% highest values of A-efficiency when compared to other designs. For the Box-Wilson CCD, the quarter fraction of resolution V has a minimum run size of 50 with a D-efficiency value of 67.07% and an A-efficiency value of 37.72%. While the half fraction of resolution VII has a run size of 82, 85.52% D-efficiency and 57.23% A-efficiency. For the S3LD, the quarter fraction of resolution IV with a run size of 74 has a D-efficiency value of 39.33% and 30.34% A- efficiency. For the Box-Wilson CCD, considered the half-fraction of resolution VII, the D-efficiency value is slightly higher than that of the RRCD. Thus, RRCCD offers competitive efficiency with fewer runs, making it a viable CCD alternative.

Keywords: Central composite design, Repaired Resolution, Aliases, Factorial Designs, Efficiency.

1. Introduction

A central composite design is a statistical technique employed for the purpose of establishing the coefficients of a polynomial with quadratic terms in alternative to a complete three-level factorial experiment. A central composite design is an experimental design useful in the response surface method for building a second-order model for the response variable (Iwundu, 2017). Response surface methodology is defined by (Montgomery, 2013) as a collection of mathematical and statistical techniques useful for the modelling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimise this response. A response surface is a graphical representation of the response with the control factor. Assuming y is a response of interest and x is a vector of control variables on which the value of y is dependent on, then the relationship between x and y is expressed as:

$$y = f\left(x\right) + \varepsilon \tag{1}$$

such that ε is a random error and f(x) is a function of x.

A standard central composite design, also referred to as a Box-Wilson central composite design, is made up of three different parts: a full or fractional factorial design points, a set of axial points and centre points. For all important effects to be estimable, it is required of a standard CCD to have full factorial points or a half-fractional factorial of resolution V or more in its factorial portion. This is efficient if the number of design variables is few, but for factors greater than five (ie k > 5), the design becomes too large. As such, Block and Mee (2001) proposed an economical second-order design that was considered more efficient, especially in terms of run size, than the standard Box-Wilson central composite design. The design was named the repair resolution central composite design (RRCCD). The design is obtained by augmenting an initial half-fraction of a 2^k factorial design of resolution III or IV with a repaired portion, the axial points and the centre points, respectively. This design is recommended when the number of factors is greater than five (ie k > 5).

Design resolution as a means of selecting a 2^k fractional factorial experiment was introduced by (Box and Hunter 1961) and is defined as the length of the shortest word in a defining relation. That is to say, in a given design, if all the words in a defining relation are of length three, the resolution of such a design is III. In this study, design resolution acts as an essential tool in choosing the fractional factorial experiments. It shows the pattern of confounding or aliasing between main effects and interactions.

Another economic design proposed by (Block and Mee, 2001) is a three-level sequential design which is also obtained by augmenting an initial half-fractional factorial design of resolution III or IV.

Optimality of experimental designs was systematically studied by (Kiefer, 1958, 1959). Several optimality criteria were discussed. When several designs are proposed for comparison, then their optimality properties may be compared for the choice of design. A design that is superior to other designs in one optimality criterion may be inferior when compared with other optimality criteria. According to (Akram *et al.*, 2003), the performance of a design may be dependent on the choice of an evaluation criterion. Four commonly used criteria for design evaluation are A -, D-, E-, and G-optimality criteria.

Iwundu (2017) considered the loss in Relative A-, D-, and G-efficiency due to missing single or multiple observations using a central composite cuboidal design. Results showed that higher losses in Relative A- and D-efficiencies are attributed to missing vertex points. Relative A-, D- and G- is unaffected by missing one or two centre points, but there was a percentage loss in Relative A-, D- and G-efficiencies when missing both centre points and axial points. Relative G-efficiency is mildly affected by the missing vertex or axial point or both. The losses in Relative efficiency are less when compared with the losses in Relative A- and D-efficiencies.

Ranade *et al.* (2017) studied the selection of design for response surface; optimal design has the ability to fit into any type of model, either a first-order model, a second-order model or even a quadratic or cubic model. Optimal design could also be used for screening experiments and for generating response surfaces. The optimal design has the least number of experimental runs when compared to the classic CCD and BBD.

The main aim of this study is to compare the efficiency of Repaired Resolution Central Composite Design to the standard (Box-Wilson) Central Composite Design. To achieve this aim, we considered the following specific objectives: to obtain A and D optimality constraints for RRCCDs and Box-Wilson CCDs in seven design variables, involving design resolution III, IV, V and VII; to establish A- and D-efficiency for the RRCCDs and Box-Wilson CCDs; to establish run-size efficiency of the design; and to establish design preferences in terms of efficiency measures. The findings of this study contribute to the existing body of knowledge on the efficiencies of Repaired Resolution Central composite design compared to the standard Box-Wilson Central composite design by providing insight into the design with better statistical properties and more reliable parameter estimates while conserving experimental runs.

2. Materials and Methods

Where there is an indication of quadratic effects, a model that detects the presence of Curvature is usually preferred to approximate the optimum response. Therefore, the second-order model is employed for this study and is defined as follows:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2 + \sum_{i<j} \hat{\beta}_{ij} x_i x_j \quad (2)$$

It is expressed in matrix form as:

$$Y = X\beta + \varepsilon \quad (3).$$

Where, y is an N-vector of observations, β is a p×1 column vector of estimates of parameters, X is an N design model matrix, and $\boldsymbol{\varepsilon}$ is an N-vector.

The number of parameters for a second-order model for response surface design is therefore given as follows:

$$p = \binom{k+2}{2} = \frac{(k+1)(k+2)}{2}; \quad (4)$$

where k is the number of design variables. The second-order model for seven design variables is as follows:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_7 x_7 + \hat{\beta}_{11} x_1^2 + \hat{\beta}_{22} x_2^2 + \cdots + \hat{\beta}_{77} x_7^2 + \hat{\beta}_{12} x_1 x_2 + \hat{\beta}_{13} x_1 x_3 + \cdots + \hat{\beta}_{67} x_6 x_7 \quad (5).$$

The compressed form is given as

$$\hat{y} = \hat{\beta}_0 + \textstyle\sum_{i=1}^{7} \beta_i x_i + \sum_{i=1}^{7} \beta_{ii} x_i^2 + \sum_{i=1}^{6} \sum_{j=2}^{7} \beta_{ij} x_i x_j \; i,j = 1,2,\ldots,7 \quad (6)$$

The number of parameters (p) is given as:

$$p = \frac{(k+1)(k+2)}{2} = \frac{(7+1)(7+2)}{2} = 36. \quad (7)$$

Therefore, for k = 7, the number of parameters is 36.

2.1. Study Design

In this study, five cases of composite designs are considered. Two cases of Repaired resolution central composite design, two cases of Box-Wilson central composite design and one case of sequential three-level design.

2.1.1. The Structure of Box-Wilson Central Composite Design

Central composite design involves three components: the 2k factorial or fractional factorial points, the axial points and the centre points. The construction of CCD depends on the number of design variables; for a k =7 design variable, assuming a half fractional factorial experiment, we have 64 experimental runs, and an additional 14 axial points and 4 centre points. A sequential three-level design, on the other hand, is obtained by augmenting an initial design with additional experimental points sequentially.

2.1.2. The Structure of Repaired Resolution Central Composite Design

According to (Ahmad et al., 2015), an RRCC design has four components: the first component is an initial two-level fractional factorial of resolution III or IV; the second is repaired fractional factorial points, 2^{w-1} of resolution w, for each word of Length w for which repairing is to be done; the third is the Axial points; and lastly is the Center points.

Illustratively, the first case of composite design considered is the quarter fraction of RRCCD with K=7 design variables of resolution IV; With Generators 6=123; 7=1245, the repair fraction is T=-1236, with rescaling factor ± (√k/√s) = ± 1.3229, where k is the number of design variables, and S is the design Resolution.

The repaired portion is as follows:

A	B	C	D	E	F	G	AB	AC	AD	AE
-1.3229	-1.3229	-1.3229	0	0	1.3229	0	1.75006441	1.75006441	0	0
1.3229	-1.3229	-1.3229	0	0	-1.3229	0	-1.75006441	-1.75006441	0	0
-1.3229	1.3229	-1.3229	0	0	-1.3229	0	-1.75006441	1.75006441	0	0
1.3229	1.3229	-1.3229	0	0	1.3229	0	1.75006441	-1.75006441	0	0
-1.3229	-1.3229	1.3229	0	0	-1.3229	0	1.75006441	-1.75006441	0	0
1.3229	-1.3229	1.3229	0	0	1.3229	0	-1.75006441	1.75006441	0	0
-1.3229	1.3229	1.3229	0	0	1.3229	0	-1.75006441	-1.75006441	0	0
1.3229	1.3229	1.3229	0	0	-1.3229	0	1.75006441	1.75006441	0	0

AF	AG	BC	BD	BE	BF	BG	CD	CE	CF	CG
-1.75006441	0	1.75006441	0	0	-1.75006441	0	0	0	-1.75006441	0
1.75006441	0	1.75006441	0	0	1.75006441	0	0	0	1.75006441	0
-1.75006441	0	-1.75006441	0	0	-1.75006441	0	0	0	1.75006441	0
1.75006441	0	-1.75006441	0	0	1.75006441	0	0	0	-1.75006441	0
1.75006441	0	-1.75006441	0	0	1.75006441	0	0	0	-1.75006441	0
-1.75006441	0	-1.75006441	0	0	-1.75006441	0	0	0	1.75006441	0
1.75006441	0	1.75006441	0	0	1.75006441	0	0	0	1.75006441	0
-1.75006441	0	1.75006441	0	0	-1.75006441	0	0	0	-1.75006441	0

DE	DF	DG	EF	EG	FG	A^2	B^2	C^2	D^2	E^2	F^2	G^2
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0
0	0	0	0	0	0	1.75006441	1.75006441	1.75006441	0	0	1.75006441	0

The four cases are also constructed and studied; the results are as stated in Table 1.

2.2. *Methods of Data Analysis*

The goodness or efficiency of an experimental design can be quantified. Common measures of the efficiency of an experimental design are based on the information matrix of the design (Iwundu, 2017). The D- and A-efficiency are studied for the various designs. Computations are done using R statistical software (4.6.0).

2.2.1. D-efficiency

The D-optimality criterion was introduced by Wald (1943), and it is expressed as:

$$\varphi(M(\xi)) = \text{Max } \{\det M(\xi)\} = \text{Min } \{\det (M^{-1}(\xi))\} \quad (8).$$

Maximising the determinant of the information matrix or minimising the determinant of the variance-covariance matrix. D-efficiency focuses on a good parameter estimate. The D-efficiency of a design $\xi^{(1)}$is represented as:

$$D_{eff} = 100\times(\det M(\ \xi^{(1)}))^{\ 1/p} \tag{9}$$

such that P is the number of parameters, and M is the normalised information matrix

2.2.1. A-Efficiency

Chernoff (1953) introduced the concept of the A-optimality criterion. A-efficiency of design ξ is defined as:

$$A_{eff}=100\ (p/\mathrm{trace}(N^*(X'X)^{-1})) \tag{10}$$

where P is the number of parameters.

3. Results and Discussions

Design efficiency gives insight into the goodness of an experimental design. A- and D-efficiency measures information on different areas of the performance of an experimental design. A-efficiency shows the average precision of parameter estimates, whereas D-efficiency shows the overall information available for estimating the model parameters jointly. Table 1 presents the results of the comparative study.

Table 1. Result summary of comparative study

Design type	Run size	Det of $\frac{X^1X}{N}$	Det of $\left(\frac{X^IX}{N}\right)^{-1}$	Trace of $\frac{X^1X}{N}$	Trace of $\left(\frac{X^IX}{N}\right)^{-1}$	D-efficiency	A-efficiency
RRCCD 2_{IV}^{7-2}	58	8.836×10^{-4}	1131.725	39.01852	61.2306	82.24%	58.79%
RRCCD 2_{III}^{7-2}	54	7.529×10^{-4}	1328.119	39.1984	60.4629	81.89%	59.54%
S3LD2_{IV}^{7-2}	74	2.567×10^{-15}	3.89616×10^{14}	19.05405	118.6241	39.33%	30.34%
BOX-WILSON CCD2_{V}^{7-2}	50	5.689×10^{-07}	175771.8	39.08108	95.4217	67.07%	37.72%
BOX-WILSON CCD2_{VII}^{7-1}	82	3.583×10^{-3}	279.0875	37.8665	62.9008	85.52%	57.23%

For the five composite designs considered in this study, the RRCD demonstrated superior performance, having the higher A-efficiency values of 59.54%, indicating a better precision of parameter estimation than the Box-Wilson CCD and the S3LD. The Box-Wilson CCD has a higher value of D efficiency, offering better overall information for joint parameter estimation. The RRCD offered a competitive efficiency with fewer run sizes, making it a viable CCD alternative.

3. Conclusion

This study provides valuable insight into the efficiencies of the Repaired Resolution Central composite design compared to the standard Box-Wilson Central composite design. It provides evidence in support of the proposed RRCD (Block and Mee, 2001). The RRCD repaired important confounded effects with fewer runs while maintaining a high level of competing efficiency. RRCD is therefore recommended for k>5.

Reference

Ahmad, T., & Akhtar, M. (2015). Efficient response surface designs for the second-order multivariate polynomial model robust to missing observations. *Journal of Statistical Theory and Practice, 9*(2), 361–375.

Akram, M., Akhtar, M., & Tahir, H. M. (2013). Comparison of different central composite designs. *International Journal of Agriculture & Biology, 5*(4), 571–575.

Block, R. M., & Mee, R. W. (2001). *Some new second-order designs*. Department of Statistics, University of Tennessee, Knoxville.

Box, G. E. P., & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B (Methodological), 13*(1), 1–45. https://doi.org/10.1111/j.2517-6161.1951.tb00067.x

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics, 24*(4), 586–602. https://doi.org/10.1214/aoms/1177728915

Iwundu, M. P. (2017). Missing observations: The loss in relative A-, D-, and G-efficiency. *International Journal of Advanced Mathematical Sciences, 5*(2), 43–49.

Kiefer, J. (1958). On the nonrandomized optimality and randomised nonoptimality of symmetrical designs. *Annals of Mathematical Statistics, 29*(3), 675–699. https://doi.org/10.1214/aoms/1177706536

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society: Series B (Methodological), 21*(2), 272–319.

Montgomery, D. C. (2013). *Design and analysis of experiments* (8th ed.). John Wiley & Sons.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2009). *Response surface methodology: Process and product optimisation using designed experiments* (3rd ed.). John Wiley & Sons.

Ranade, S. S., & Thiagaraj, P. (2017). Selection of a design for response surface. *IOP Conference Series: Materials Science and Engineering, 263*(2), Article 022043. https://doi.org/10.1088/1757-899X/263/2/022043

Wald, A. (1943). On the efficient design of statistical investigations. *Annals of Mathematical Statistics, 14*(2), 134–140. https://doi.org/10.1214/aoms/1177731498