JIRSS (2023)
Vol. 22, No. 02, pp 59-75
DOI: 10.22034/jirss.2024.1999151.1015

Robustness of Augmented Box-Behnken Designs to Two and Three Missing Observations

Oluchukwu Chukwuemeka Asogwa [1], Oluwaseun Aramide Otekunrin [2]

[1] Department of Mathematics and Statistics, Faculty of Physical Sciences, Alex Ekwueme Federal University Ndufu-Alike, Ebonyi State, Nigeria
[1,2] Statistical Design of Investigations Unit, Department of Statistics, Faculty of Science, University of Ibadan, Nigeria

Received: 28/03/2023, Accepted: 03/07/2024, Published online: 30/10/2024

Abstract. In real experimentation, second order response surface models may in some cases be inadequate and unrealistic due to lack of fit introduced by the presence of third-order terms in the response surface model. This willingly opens the door for augmentation of a second order response surface model to a third order model. The third–order design that takes care of the estimates of the model in this work is augmented Box Behnken Design (ABBD), which has always been confronted with missing observation scenario. This scenario that has affected the power of its test and desirable fundamental properties of this design, often surfaces when an observation is lost, ignored, miss-collected, etc. Therefore, this work constructed a minimaxloss design for ABBD that is robust to two and three missing observations under minimaxloss criterion.

Keywords. Box-Behnken Design, Lack of Fit, Response Surface Model, Second-order Designs, Third-order Design, Robustness of Design, Missing Observations.

MSC: 62K20, 62K25.

1 Introduction

Response surface methodology is an efficient modern statistical tool which emanated from Box and Wilson (1951). It has been widely used in building models and exploring

Corresponding Author : Oluchukwu Chukwuemeka Asogwa (qackasoo@yahoo.com)
Oluwaseun Aramide Otekunrin (seramide2003@yahoo.co.uk)

relationships between one or more response variables and several explanatory variables (Zhou and Xu, 2016; Ahmad and Gilmour, 2010). Its application has spanned many fields, including; manufacturing industries for process development (Lamidi et al., 2022), chemical and food industries (Yolmeh et al., 2017; Nwabueze, 2010; Akram 2002; Baş et al., 2007), biological and biomedical (Safari et al., 2018; Malekjani, 2020), bio-pharmaceutical (Rebollo-Hernanz et al., 2021), bioengineering and agriculture (Mead and Pike, 1975).

However, in response surface methodology, second order models, may be inadequate in some cases and unrealistic due to lack of fit caused by the presence of third-order or higher-order terms in the response surface model (Rashid et al., 2017). This situation creates a need for a third-order or higher-order model and this third-order or higher-order model can be obtained by augmenting the second order response surface designs in order to estimate the third-order or higher order terms in the response surface model. Although the augmented designs have large design points and are efficient, experimenters usually prefer small response surface designs because it saves costs.

Arshad et al. (2012) constructed augmented Box-Behnken designs which can handle the estimates of the third-order models and these are known as third-order designs. Rashid et al. (2019) investigated the robustness of Augmented Box-Behnken Designs (ABBDs) and Augmented Fractional Box-Behnken Designs (AFBBDs) to one missing observation using minimaxloss criteria. Rashid et al. (2022) investigated the impact of missing one observation on the estimation and predictive capabilities as well as on the relative A, D and G-efficiencies of Augmented Box-Behnken designs. These third-order designs in general have been used in many experimental situations for response surface modelling and however may have been faced with experiments that include one or more missing observations.

Akram (2002) developed minimaxloss3 designs robust to three missing observations by using minimaxloss criteria, and then compared these designs with cuboidal, orthogonal, rotatable, box and draper outlier robust designs, spherical, minimaxloss1 (loss due to one missing observation) and minimaxloss2 (loss due to two missing observations) designs for their robustness to a combination of three missing observations and examined the consequences of missing three observations on different kinds of CCDs which include cuboidal, spherical, orthogonal, rotatable, minimum variance, Box and Draper outlier robust designs.

Missing observation is a phenomenon that surfaces even in a carefully planned experiment, where some observations may be lost during the process of data collection, damaged or may be suspicious in some way (Patchanok, 2015). Missing observation affects the statistical power of a test, destroy some of the fundamental design properties like: orthogonal, optimality, balanced structure of a design, offer biased estimates of parameters and give invalid conclusions drawn from the data (Chen et al., 2017). This phenomenon which is missing at random (MAR) in this setting, can be resulted from many causes, for example, the loss of experimental units, cancellation of runs that take too long to achieve and miscoded data where their correct values are non-tractable (MacEachern et al., 1995). Many researchers have presented ways of dealing with missing observations which includes: Imputation method (Marina, 2013) and robustness-to-missing value criteria (Andrews & Herzberg, 1979; Ghosh, 1979; Akhtar & Prescott, 1986; Tanco et al., 2013).

Robustness of a design against missing observation is the ability of the design to work well (being able to estimate parameters without too much loss of precision) in the presence of uncontrollable conditions (missing observations). It could also be expressed as the experimental designs which provide protection in the presence of conditions such as outlying observations, missing observations, non-normality of error distribution, inadequate response surface model and auto-correlation in time (Akram, 2002). The robustness of designs against missing observation has been studied by several authors, including Hedayat and John (1974), John (1976), Kageyma (1980), Ghosh (1978, 1980, 1981, 1982), Ghosh et al. (1983), Akhtar and Prescott (1986), Ahmad and Gilmour (2010), Ahmad et al. (2012), Alrweili et al. (2019), and Georgiou et al. (2014). Therefore, the goal of this paper is to construct a minimaxloss design for augmented Box-Behnken design that is robust to two and three missing observations using a minimaxloss criterion.

This paper is organized in this way, materials and methods, which captured source of the design, third order model, robustness to missing observation criterion, minimaxloss criterion and construction of the proposed minimaxloss criterion are well discussed in section two. Section three deliberated on robust designs and its application, while section four and five presented results and conclusion and discussion respectively.

2 Materials and Methods

This section is subdivided into source of designs, third order model, robust-to-missing observations criteria, and construction of minimaxloss augmented Box-Behnken Designs.

2.1 Source of Designs

The design presented in Table 1 was obtained from a work titled augmented Box-Behnken designs for fitting third order response surfaces by Arshad et al. (2012; table 5, page 4235).

In Table 1 above, $B[3] = n_b = b \times 2^t = 3 \times 2^2 = 12$, is the number of Box Behnken points, $F[a]^3 = n_f = 2^k = 2^3 = 8$, is the number of factorial points, $A[\alpha]^3 = n_a = 2k = 2(3) = 6$, is the number of axial points, n_c is the number of center points which is replicated three times and N is the total number of design points in augmented Box-Behnken designs for factor $k = 3$. However, the design matrix for $k = 3$ factors in Augmented

Box-Behnken Design (ABBD) is given as in Table 2 (note that here, values of factorial points was coded as -1 and +1):

Table 1: Augmented Box-Behnken third-order designs (ABBD) for k = 3.

Factor (k)	Original BBD + Added Points s	Design Points
3	$B[3]$	12
	$F[a]^3$	8
	$A[\alpha]^3$	6
		$\mathbf{N = 26 + n_c}$

2.2 Third Order Model

A mathematical model is an equation that represents or describes the relationship between a response variable and one or more explanatory (or predictor) variables. The form of the relationship between the response and the predictor variables is unknown in many response surface methodology cases and the curvature in the response function and the presence of model lack of fit will not be adequately modeled by first and second order model, respectively. Therefore, polynomial of higher-degree polynomials, such as the third-order model in optimization experiments, is employed to obtain an assessment of presence of lack of fit of the second order model. For k quantitative factors denoted by $x_1, x_2, \ldots, x_k$, a third-order response model is

$$\begin{aligned} y &= \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{i<j}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{iii} x_i^3 + \sum_{i=1}^{k-1} \sum_{i<j}^{k} \beta_{ijj} x_i x_j^2 \\ &+ \sum_{i=1}^{k-1} \sum_{i<j}^{k} \beta_{iij} x_i^2 x_j + \sum_{i=1}^{k-2} \sum_{j=i+l}^{k-1} \sum_{l=j+i}^{k} \beta_{ijl} x_i x_j x_l + \varepsilon, \end{aligned} \tag{2.1}$$

where $\beta_0, \beta_i, \beta_{ii}, \beta_{ij}, \beta_{iii}, \beta_{ijj}, \beta_{iij}$ *and* β_{ijl} represents the intercept, linear, quadratic, bilinear and cubic effects, respectively, and ϵ_i is a random error with mean zero, variance σ^2. The number of unknown parameters to be estimated is denoted as $p = \left((k+3)\ \complement_3 = \frac{(k+3)!}{3!\,k!}\right)$ or $p = \frac{(k+d)!}{k!\,d!}$ where k represents the number of factor under consideration and d is the order of the design under investigation. In other to have sufficient degrees of freedom to estimate the model coefficients, the number of runs n (number of observations) must be greater than or equal to p (the number of parameters in the model). In matrix notation, equation (1) above can be written as:

$$y = X\beta + \varepsilon, \tag{2.2}$$

where **X** is the $n \times p$ model matrix derived from the $n \times k$ design matrix with $x = (1,\ x_1, x_2, \ldots, x_k;\ \ x_1^2, x_2^2, \ldots, x_k^2;\ \ x_1x_2, x_1x_3, \ldots, x_{k-1}x_k;\ \ x_1^3, x_2^3, \ldots, x_k^3)$.

Table 2: The design matrix for k = 3 factor in Augmented Box-Behnken Design (ABBD) at n_c = 3

	S/N	X_0	X_1	X_2	X_3
	1	1	-1	-1	0
	2	1	1	-1	0
	3	1	-1	1	0
	4	1	1	1	0
	5	1	-1	0	-1
Box Behnken point: n_b = 12	6	1	1	0	-1
	7	1	-1	0	1
	8	1	1	0	1
	9	1	0	-1	-1
	10	1	0	1	-1
	11	1	0	-1	1
	12	**1**	**0**	**1**	**1**
	13	1	-1	-1	-1
	14	1	1	-1	-1
	15	1	-1	1	-1
Factorial point: n_f=8	16	1	1	1	-1
	17	1	-1	-1	1
	18	1	1	-1	1
	19	1	-1	1	1
	20	**1**	**1**	**1**	**1**
	21	1	α	0	0
	22	1	-α	0	0
	23	1	0	α	0
Axial point: n_α = 6	24	1	0	-α	0
	25	1	0	0	α
	26	**1**	**0**	**0**	**-α**
	27	1	0	0	0
Center point: n_c = 3	28	1	0	0	0
	29	1	0	0	0

2.3 Robust-to-Missing Observations Criterion

Robust-to-missing observation criterion is a criterion that is used to construct or evaluate the robustness of designs in the presence of missing observations. The following are robust to missing observation criteria: Estimability criterion by Ghosh (1979, 1982), Loss of D-efficiency by Ghosh (1979), and Minimaxloss criteria by Akhtar and Prescott (1986). This research made use of minimaxloss criterion proposed by Akhtar and Prescott (1986).

2.3.1 Minimaxloss Criterion

Akhtar and Prescott (1986) developed a minimaxloss criterion which is minimization of the maximum loss due to missing observations in the reduction of the determinant of the information matrix denoted by $\left|\mathbf{X}'\mathbf{X}\right|$. However, Andrews and Herzberg (1979) established a relationship between the reduced determinants of the information matrix, with m missing design points ($m = 1, 2, ...$), denoted by $\left|\mathbf{X_r}'\mathbf{X_r}\right|$ and the determinant of the information matrix without any missing observation known as full or complete information matrix $\left|\mathbf{X}'\mathbf{X}\right|$ which is expressed as

$$\frac{\left|\mathbf{X_r}'\mathbf{X_r}\right|}{|\mathbf{X}'\mathbf{X}|}=R_j, \tag{2.3}$$

where $R_j = 1 - h_{ij}$ is the j^{th} diagonal element of $(\mathbf{I}-\mathbf{H})$ and $\mathbf{H}$ is the hat matrix, $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. Now the loss for the j^{th} design point missing, is defined as

$$l_j=1-\frac{\left|\mathbf{X_r}'\mathbf{X_r}\right|}{|\mathbf{X}'\mathbf{X}|}. \tag{2.4}$$

Equation (2.2) can also be rewritten according to (Rashid et al. 2019) as

$$l_j = 1 - R_j = h_{jj}. \tag{2.5}$$

The loss l_j is a relative measure of efficiency with $0 \leq l_j \leq 1$. It has been observed that a small value of l_j indicates a low reduction in the determinant of the information matrix and in this sense, less loss of information. We will choose a design with a smallest value of the maximumloss over design points (Ahmad and Gilmour, 2010).

2.3.2 Construction of the Proposed Minimaxloss Design.

The detailed procedure for constructing robust Augmented Box-Behnken Design (ABBD) for two and three missing observations are presented as follows:

Algorithm

Step 1. Choose an ABBD for factor $k = 3$ (Rashid et al., 2022), we represent Box-Behnken point by $'bb'$, factorial point by $'f'$, axial point by $'a'$ and centre point by $'c'$. None of the design points are replicated, except for the center point which is fixed at 3 according to Akhtar and Prescott (1986).

Step 2. Calculate the loss function l_{bb}, l_{ff}, l_{aa}, l_{cc}, l_{bf}, l_{ba}, l_{bc}, l_{fa}, l_{fc} and l_{ac} based on X for the third-order model (1), where l_{bb} is the loss for missing a pair of Box-Behnken point, l_{ff} is the loss for missing a pair of factorial points, l_{aa} is the loss for missing a pair of axial points, l_{cc} is the loss for missing a pair of center points, l_{bf} is the loss for missing one Box-Behnken point and one factorial point, l_{ba} is the loss for missing one Box-Behnken point and one axial point, l_{bc} is the loss for missing one Box-Behnken point and one centre point, l_{fa} is the loss for missing one factorial point and one axial point, l_{fc} is the loss for missing one factorial point and one center point and l_{ac} is the loss for missing one axial point and one center point.

Step 3. Numerically search for α value that minimizes the maximum losses of l_{bb}, l_{ff}, l_{aa}, l_{cc}, l_{bf}, l_{ba}, l_{bc}, l_{fa}, l_{fc} and l_{ac}

Step 4. Substitute the chosen α value in the chosen design.

Step 5. Repeat step 2 by calculating the loss function l_{bbb}, l_{fff}, l_{aaa}, l_{ccc}, l_{cca}, l_{ccf}, l_{ccb}, l_{faa}, l_{fac}, l_{ffa}, l_{ffc}, l_{baa}, l_{bba}, l_{bbf}, l_{bbc}, l_{bff}, l_{bac}, l_{bfc}, l_{bfa}, and l_{aac} based on X for the same model (1). l_{bbb} is the loss for missing three Box-Behnken point, l_{fff} is the loss of missing three factorial points, l_{aaa} is the loss for missing three obssrvations under axial point, l_{ccc} is the loss for missing three observations of center point, whereas, l_{cca} is the loss for missing two center points and one axial point, l_{ccf} is the loss for missing two observations under center part and one factorial point, l_{ccb} is the loss of missing two center points and one Box-Behnken part, l_{faa} is the loss of missing one factorial point and two of axial point, l_{fac} is the loss for missing one factorial, one axial and one center point respectively, l_{ffa} is the loss for missing two factorial points and one axial point, l_{ffc} is the loss for missing two factorial points and one center point, l_{baa} is the loss for missing one Box-Behnken point and two axial points, l_{bba} is the loss for missing two Box-Behnken points and one axial point, l_{bbf} is the loss for missing two Box-Behnken points and one factorial point, l_{bbc} is the loss for missing two Box-Behnken points and one center point, l_{bff} is the loss for missing one Box-Behnken point and two factorial points, l_{bac} is the loss for missing one Box-Behnken, one axial point and one center point, l_{bfc} is the loss for missing one Box-Behnken point, one factorial point and one center point, l_{bfa} is the loss for missing one observation in the Box-Behnken point, factorial point, and axial point respectively and l_{aac} is the loss for missing two axial points and one center point.

Step 6. Search for α value that minimizes the maximum losses of (l_{bbb}, l_{fff}, l_{aaa}, l_{ccc}, l_{cca}, l_{ccf}, l_{ccb}, l_{faa}, l_{fac}, l_{ffa}, l_{ffc}, l_{baa}, l_{bba}, l_{bbf}, l_{bbc}, l_{bff}, l_{bac}, l_{bfc}, l_{bfa}, and l_{aac})

Step 7. Substitute the new α in the chosen design.

2.4 Robust Designs

To find the minimaxloss designs robust to two and three missing observations, we have to explore the case design (ABBD) with number of factor $k = 3$, center point replications $n_c = 3$ and number of design points $N = 29$). In the structure of the augmented Box-Behnken designs, all losses are classified into four types for all k, where k is the number of factors; loss of Box-Behnken points, loss of factorial points, loss of axial points, and loss of center points. To search the α (axial distance) value at which the maximum loss is minimized for the case of missing two observations, loss of pairs of factorial points (l_{ff}) is equated to the loss of pairs of axial points (l_{aa}) and the minimum value of α at which (l_{ff}) is equal to (l_{aa}) is chosen as the value of α that makes the design robust. However, for three missing observations, the value of α at which the maximum loss is minimized, loss of missing a pair of factorial and one axial point (l_{ffa}) is equated to the loss of missing a pair of axial points and one factorial point (l_{faa}) and the minimum value of α at which l_{ffa} is equal to l_{faa} is chosen as the value of α at which robustness of the design occurred. Note that it is not always certain that the search for α (axial distance) value that will make a design with missing observations robust exists. At times some designs with missing observation is not robust because there existed no such axial distance value that will make two functions to be equal.

2.5 Application of the Study

Practical Example: considering a study carried out by (Khuri and Cornell 1996, Table 4.2, Page 119) to evaluate the effects of temperature, pressure, and ethanol on the extraction of oil and alkaloids from ground seeds, a three factor 12-points Box-Behnken design was adopted. The design settings in its original and coded variables are listed in Table 3. This example was used to obtain the Augmented Box-Behnken Design (ABBD) by adding Factorial point of levels $'a'$ denoted by $F[a]^k = n_f = 2^k = 8$, axial point denoted by $A[\alpha]^3 = n_a = 2k = 6$, of level $'\alpha'$ and center point, denoted by ($n_c = 3$) to the Box-Behnken design, where $k = 3$. This augmented Box-Behnken design was then used to obtain the minimaxloss design robust to two and three missing observations respectively. The loss after missing a (two and three) Box-Behnken point (n_b), a factorial point (n_f), an axial n_a and a center point (n_c) of augmented Box-Behnken design has been calculated for certain values of axial distances (α). The results of losses are tabulated (Table 4 and 5) and plotted against the values of α (figure 1 and 2, for two and three missing observations respectively) and the minimaxloss design is obtained by using plots and also by an iteration. It is observed that minimum loss for missing two observation is $l_{ff} = l_{ff} = 0.9664$ at $\alpha = 1.6021$, and that of missing three observation is $l_{faa} = l_{ffa}$ and that point is 0.9937 at α =1.5599. So, for three factors with $n_b = 12$, $n_f = 8$, $n_a = 6$, $n_c = 3$ for $\alpha = 1.6021$ *and* 1.5599, the design is a minimaxloss augmented Box-Behnken design robust to two and three missing observations respectively.

3 Results

3.1 Losses Due to Two Missing Observations in Augmented Box-Behnken Design.

When a pair of design points are missing in an augmented Box-Behnken design there are ten possible outcomes of the combination of their losses, which are: loss due to missing a pair of Box- Behnken points (l_{bb}), loss due to missing a pair of factorial points (l_{ff}), loss due to missing a pair of axial points (l_{aa}), loss due to missing a pair of center points (l_{cc}), loss due to missing one factorial and one axial point (l_{fa}), loss due to missing one Box-Behnken and one factorial point (l_{bf}), loss due to missing one Box-Behnken and one center point (l_{bc}), loss due to missing one factorial and one center point (l_{fc}), loss due to missing one Box-Behnken and one axial point (l_{ba}) and loss due to missing one axial and one center point (l_{ac}). The losses for all possible pairs of augmented

Box-Behnken design with three factors $k = 3$, $n_c = 3$, with no replication at any other points except the center point, $N = 29$, $N_R = 27$, with Box-Behnken points (n_{bb}=12), factorial points ($n_f = 8$), axial point ($n_a = 6$) and center points ($n_c = 3$), where N_R is the design point with a pair of missing observations and N is the total design points.

Table 3: Residual oil content after supercritical extraction

Run	Temperature()	Pressure(MPa)	Ethanol (%)	Coded design settings x_1	x_2	x_3
1	35	20.7	0	-1	0	-1
2	45	10.4	0	0	-1	-1
3	55	10.4	5	1	-1	0
4	35	20.7	10	-1	0	1
5	45	31.0	0	0	1	-1
6	55	20.7	0	1	0	-1
7	35	10.4	5	-1	-1	0
8	35	31.0	5	-1	1	0
9	45	10.4	10	0	-1	1
10	45	31.0	10	0	1	1
11	55	31.0	5	1	1	0
12	55	20.7	10	1	0	1

It is observed in Table 4 above that loss incurred by missing a pair of center points (l_{cc}) remain less than loss due to missing a pair of Box-Behnken points (l_{bb}) and every other pair of augmented Box-Behnken design for the whole ranges of axial point distance (α). This is because of the location dependency of h_{jj}. The minimaxloss design point for augmented Box-Behnken design with $k = 3$, $n_c = 3$ and a pair of observations missing is chosen from the values of alpha (α) ranging from 0.25 *to* 4.00, for which the loss due to missing a pair of factorial points (l_{ff}) is equal to the loss due to missing a pair of axial points (l_{aa}). The minimaxloss design point occurs at a point where $l_{ff} = l_{aa}$ and that point is 0.9664 at $\alpha = 1.6021$.

Thus, a three factor augmented Box-Behnken design with total design points $N = 29$, Box- Behnken points (n_{bb}=12), factorial points ($n_f = 8$), axial points ($n_a = 6$), center points ($n_c = 3$) and $\alpha = 1.6021$ is a minimaxloss design robust to a pair of missing observations (minimaxloss2). However, Figure 1 shows the plots of the loss due to missing a pair of Box-Behnken, a pair of factorial, a pair of axial and a pair of center points in an augmented Box-Behnken design for $k = 3$ and $n_c = 3$. The minimaxloss point is traced where the line plot of loss due to missing a pair of factorial points intercepted that of the line plot of loss due to missing a pair of an axial point.

Table 4: The maximum losses due to a pair of missing observations of augmented Box-Behnken design at Box-Behnken, factorial, axial and center points for different values of α with $k = 3$ and $n_c = 3$.

Axial point	Loss in precision due to missing									
distance (α)	bb	ff	aa	cc	bf	ba	bc	fa	fc	ac
0.25	0.9691	0.9710	0.8505	0.2105	0.9804	0.9156	0.8098	0.9323	0.8480	0.6532
0.50	0.9657	0.9715	0.8508	0.2314	0.9796	0.9154	0.8075	0.9334	0.8514	0.6603
0.75	0.9534	0.9722	0.8646	0.2722	0.9767	0.9162	0.7961	0.9381	0.8573	0.6888
1.10	0.9207	0.9729	0.9078	0.3782	0.9680	0.9213	0.7691	0.9500	0.8695	0.7617
1.25	0.9064	0.9724	0.9273	0.4394	0.9631	0.9244	0.7614	0.9547	0.8741	0.7915
1.50	0.8883	0.9689	0.9561	0.5198	0.9536	0.9321	0.7548	0.9617	0.8727	0.8334
1.60	0.8834	0.9665	0.9661	0.5260	0.9495	0.9371	0.7511	0.9650	0.8672	0.8501
1.6021*	0.8833	0.9664	0.9664	0.5259	0.9494	0.9372	0.7510	0.9651	0.8670	0.8505
1.6025	0.8833	0.9664	0.9665	0.5259	0.9494	0.9371	0.7510	0.9651	0.8670	0.8503
1.6050	0.8832	0.9655	0.9666	0.5257	0.9493	0.9374	0.7509	0.9652	0.8668	0.8510
1.75	0.8781	0.9629	0.9785	0.4992	0.9439	0.9468	0.7416	0.9709	0.8549	0.8767
2.00	0.8728	0.9583	0.9912	0.4000	0.9379	0.9643	0.7185	0.9807	0.8333	0.9185
2.25	0.8697	0.9564	0.9965	0.3047	0.9356	0.9774	0.6973	0.9879	0.8190	0.9484
2.50	0.8676	0.9559	0.9986	0.2389	0.9351	0.9856	0.6822	0.9923	0.8113	0.9668
2.75	0.8662	0.9560	0.9994	0.1966	0.9352	0.9905	0.6722	0.9949	0.8074	0.9780
3.00	0.8652	0.9563	0.9997	0.1690	0.9355	0.9935	0.6655	0.9965	0.8054	0.9814
3.25	0.8645	0.9566	0.9998	0.1505	0.9359	0.9954	0.6609	0.9976	0.8044	0.9892
3.50	0.8640	0.9569	0.9999	0.1376	0.9363	0.9967	0.6576	0.9982	0.8038	0.9921
3.75	0.8637	0.9571	1.0000	0.1283	0.9366	0.9975	0.6553	0.9987	0.8036	0.9941
4.00	0.8634	0.9573	1.0000	0.1214	0.9368	0.9981	0.6535	0.9990	0.8034	0.9955

At that intercept, the axial distance value (α) is 1.6021 and the loss value is 0.9664. In Figure 1 also, other loss line plots due to missing a pair of Box-Behnken points and a pair of center points were sighted. The line plot of center point is always lower than every other plots displayed and that was justified by its loss values as seen in Table 4. This is because of the closeness of the central points to the explanatory points. The shape of the line plot due to missing a pair of center point is bell shaped.

3.2 Losses Due to Three Missing Observations in Augmented Box-Behnken Design.

When three of design points are missing in an augmented Box-Behnken design with $k = 3$, $n_c = 3$, there are about twenty plus possible outcomes of the combination of their losses. This researchworked with those losses that are of interest to the researcher. The losses that are of interest are those losses that do not give a singular matrix. This work considered about twenty different losses at three possible combinations of the factor. Those losses considered were: loss due to missing three Box-Behnken points (l_{bbb}), loss due to missing three factorial points (l_{fff}), loss due to missing three axial points (l_{aaa}), loss due to missing three center points (l_{ccc}), loss due to missing two Box-Behnken and one factorial points $\left(l_{bbf}\right)$, others were loss due to missing two Box- Behnken and one axial points (l_{bba}), loss due to missing one Box-Behnken and two factorial points (l_{bff}), loss due to missing one Box-Behnken, one factorial and one axial points (l_{bfa}).

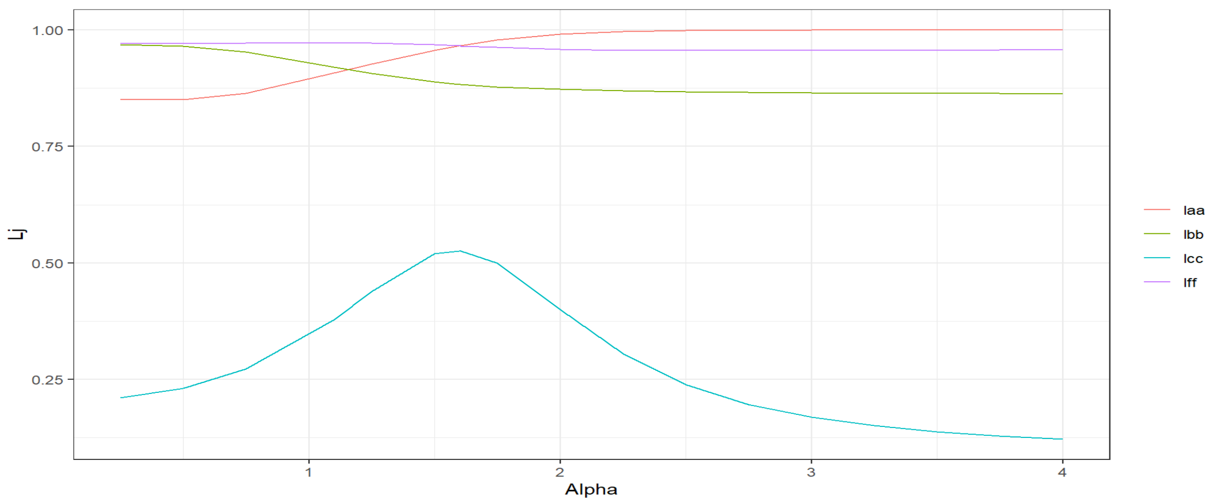

Figure 1: Loss due to two Box-Behnken, two factorial, two axial and two center points missing for $k = 3$ and $n_c = 3$.

Among the losses were loss due to missing one Box-Behnken and two axial points (l_{baa}), loss due to missing one Box-Behnken, one factorial and one center points (l_{bfc}), loss due to missing two Box- Behnken and one center points (l_{bbc}) was also included. Others losses were: loss due to missing a Box-Behnken and two center points (l_{bcc}), loss due to missing two Box-Behnken and one center points (l_{bbc}), loss due to missing two factorial and one axial points (l_{ffa}), loss due to missing a factorial and two axial points (l_{faa}), loss due to missing one factorial, one axial and one center points (l_{fac}), loss due to missing two factorial and one center points (l_{ffc}), loss due to missing one factorial and two center points (l_{fcc}), loss due to missing two axial and one center points (l_{aac}), loss due to missing one axial and two center points (l_{acc}).

There might be a scenario where one encounters more than one loss value in a particular combination, for instance when considering l_{bbb}, picking serial number $l_{1,2,3}$ to represent the l_{bbb} and another serial number $l_{1,5,12}$, to represent the same l_{bbb} at the same alpha value, remembering that in augmented Box-Behnken designs, Box- Behnken point is represented as ($b \times 2^t = 3 \times 2^2 = 12$). In this situation, some combination of the possible serial numbers of the considered loss will give the same loss value. For example, $l_{1,2,3}$ and $l_{1,2,5}$ gave the same loss value for missing three Box-Behnken points in an augmented Box-Behnken design but some combinations will offer either a singular matrix (a design matrix with zero determinant) or a different loss value. In this work, we tried to be consistent across all the combinations considered.

If $l_{bbb} = l_{1,2,3}$ was considered at $\alpha = 0.25$, then the same $l_{bbb} = l_{1,2,3}$ will be considered when looking at $0.50 \leq \alpha \leq 3.00$, etc. This will create a consistency in their loss values. The losses for all possible combination of three design points missing in augmented Box-Behnken design with three factors $k = 3$, $n_c = 3$, with no replications of any other design points, $N = 29$, $N_R = 26$, Box-Behnken points (n_{bb}=12), factorial point ($n_f = 8$), axial point ($n_a = 6$) and center point ($n_c = 3$), where N_R is the total design points with three missing observations and N is the total number of design points, is presented.

Table 5 above, shows the loss values for all different three combination of points when three observations are declared missing in augmented Box-Behnken design. From the Table 5 above, it was observed that the loss value due to missing triple center points l_{ccc} remains lower than loss value due to missing any other triple possible combination of points for the whole ranges of axial point distance(α). Finding the minimax loss (minimaxloss3) design for ABBD with $k = 3$ and $n_c = 3$, is searching for that value of alpha (α) from the whole ranges of alpha ($0.25 \leq \alpha \leq 3.00$) values for which loss due to missing one factorial point and two axial points l_{faa} is equal to the loss due to missing two factorial points and one axial point l_{ffa}. That is $l_{faa} = l_{ffa}$. The minimaxloss occurs at a point where $l_{faa} = l_{ffa}$ and that point is 0.9937 at α =1.5599. Thus a three factor augmented Box-Behnken design with total design points $N = 29$ *and* $N_R = 26$, Box- Behnken points (n_{bb}=12), factorial points ($n_f = 8$), axial points ($n_a = 6$), center points ($n_c = 3$) and $\alpha = 1.5599$ is a minimaxloss design robust to a three missing observations (minimaxloss3).

Figure 2 shows the plots for the loss due to missing three Box-Behnken points, three factorial points, three axial points, two factorial and one axial points, one factorial and two axial points and three center design points in an augmented Box-Behnken designs for $k = 3$ and $n_c = 3$. The minimaxloss point is tracked where the line plot of loss due to missing two factorial points and one axial point intercepted that of the line plot of loss due to missing one factorial point and two axial points. At that intercept, α =1.5599 and the loss value is 0.9937. Figure 2 also contained other loss line plots due to missing three Box-Behnken points, three factorial points, three axial points and three center points respectively. The line plot of center point is always lower than every other plots displayed in the Figure 2 above and that was justified by its loss values as seen in Table 5. The shape of the line plot due to missing three center points is bell shaped.

Table 5: The maximum losses due to three missing observations of augmented Box-Behnken designs at Box-Behnken, factorial, axial and center points for different values of α with $k = 3$ and $n_c = 3$.

Axial point distance (α)	Loss in precision due to missing																			
	faa	*ffa*	*bbb*	*baa*	*bac*	*fff*	*aaa*	*ccc*	*bba*	*bbc*	*bbf*	*bcc*	*bfa*	*bfc*	*bff*	*fac*	*fcc*	*ffc*	*aac*	*acc*
0.25	0.9749	0.9885	0.9915	0.9685	0.9270	0.9951	0.9502	0.3157	0.9881	0.9726	0.9974	0.8326	0.9922	0.9825	0.9969	0.9415	0.8664	0.9742	0.8767	0.7058
0.50	0.9751	0.9888	0.9907	0.9682	0.9279	0.9952	0.9485	0.3470	0.9876	0.9699	0.9970	0.8331	0.9921	0.9690	0.9968	0.9433	0.8716	0.9750	0.8785	0.7170
0.75	0.9776	0.9897	0.9876	0.9703	0.9306	0.9954	0.9526	0.4083	0.9855	0.9600	0.9954	0.8287	0.9918	0.9800	0.9964	0.9489	0.8813	0.9764	0.8923	0.7503
1.10	0.9848	0.9918	0.9769	0.9785	0.9392	0.9955	0.9721	0.5673	0.9813	0.9362	0.9913	0.8239	0.9916	0.9746	0.9950	0.9617	0.9036	0.9790	0.9300	0.8275
1.25	0.9880	0.9925	0.9710	0.9827	0.9434	0.9954	0.9811	0.6591	0.9801	0.9276	0.9892	0.8299	0.9917	0.9721	0.9942	0.9664	0.9143	0.9797	0.9457	0.8571
1.50	0.9927	0.9935	0.9628	0.9893	0.9505	0.9945	0.9921	0.7797	0.9798	0.9188	0.9855	0.8440	0.9922	0.9666	0.9925	0.9724	0.9221	0.9781	0.9676	0.8925
1.5599*	0.9937	0.9937	0.9613	0.9907	0.9526	0.9942	0.9939	0.7888	0.9802	0.9171	0.9845	0.8446	0.9924	0.9648	0.9920	0.9738	0.9202	0.9770	0.9722	0.8992
1.60	0.9940	0.9943	0.9604	0.9917	0.9542	0.9940	0.9949	0.7889	0.9807	0.9160	0.9839	0.8439	0.9926	0.9636	0.9916	0.9748	0.9181	0.9762	0.9750	0.9036
1.6250	0.9941	0.9947	0.9599	0.9922	0.9552	0.9938	0.9954	0.7866	0.9810	0.9152	0.9836	0.8429	0.9928	0.9628	0.9914	0.9754	0.9164	0.9756	0.9767	0.9063
1.75	0.9964	0.9949	0.9579	0.9947	0.9609	0.9930	0.9975	0.7487	0.9830	0.9111	0.9818	0.8326	0.9937	0.9609	0.9903	0.9788	0.9046	0.9725	0.9841	0.9197
2.00	0.9985	0.9966	0.9555	0.9978	0.9730	0.9919	0.9993	0.6000	0.9881	0.9012	0.9793	0.7944	0.9956	0.9512	0.9888	0.9855	0.8750	0.9667	0.9934	0.9444
2.25	0.9994	0.9978	0.9540	0.9991	0.9825	0.9914	0.9998	0.4571	0.9923	0.8924	0.9782	0.7567	0.9972	0.9468	0.9882	0.9906	0.8521	0.9632	0.9974	0.9632
2.50	0.9998	0.9986	0.9531	0.9996	0.9886	0.9913	1.0000	0.3584	0.9950	0.8860	0.9778	0.7296	0.9982	0.9445	0.9880	0.9939	0.8382	0.9616	0.9989	0.9755
2.75	0.9999	0.9991	0.9524	0.9998	0.9923	0.9913	1.0000	0.2948	0.9967	0.8816	0.9776	0.7117	0.9988	0.9434	0.9879	0.9959	0.8301	0.9609	0.9995	0.9833
3.00	0.9999	0.9994	0.9519	0.9999	0.9947	0.9914	1.0000	0.2536	0.9977	0.8787	0.9777	0.6998	0.9992	0.9429	0.9880	0.9972	0.8253	0.9606	0.9998	0.9883

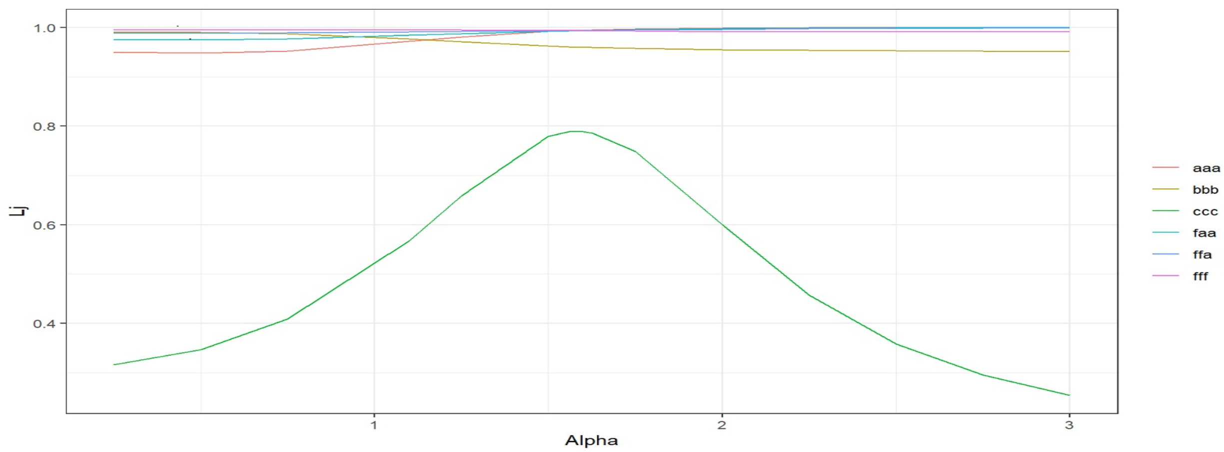

Figure 2: Loss due to three Box-Behnken, three factorial, three axial, one factorial and two axial, two factorial and one axial and three center points missing for $k = 3$ and $n_c = 3$.

4 Discussion and Conclusion

In response surface modeling, robustness to missing observations is a desirable property good response surface designs should hold, which is crucial to experimenters and therefore, designs that are minimally affected by the external sources of variability (such as missing observations) are desirous. Designs that are robust to missing observations are required so as to reduce the effects of the missing observations and also in such robust designs, the parameters of the assumed model can be estimated without much loss of efficiency. However, the augmented Box-Behnken design is one of the augmented third-order response surface designs for response surface exploration. The unique structure of the augmented Box-Behnken design with four components, the Box-Behnken, the factorial, the axial and the center design points, makes the design very flexible to use in industrial experiments. In this work, an augmented Box-Behnken design that is robust to two and three missing observations were constructed by using the minimaxloss criteria. It was shown that for $k = 3$, $n_b = 12$, $n_f = 8$, $n_a = 6$, and $n_c = 3$, with $\alpha = 1.6021$ and $\alpha = 1.5599$ are minimaxloss augmented Box-Behnken design robust to two and three missing observations respectively (minimaxloss2 and minimaxloss3). That is, the axial point distance (α) at the point at which $l_{ff} = l_{aa}$ and $l_{ffa} = l_{faa}$, obtained from the plot of loss against the axial point distance α, where l_{ff} is the loss incurred from losing a pair of factorial design points, l_{aa} is the loss acquired from losing a pair of axial design points, l_{ffa} is the loss encountered from missing three observations of a pair of factorial and axial points and l_{faa} is the loss experienced from missing three observations of a pair of axial and factorial points.

Acknowledgements

The authors are grateful to Dr. Abimibola Victoria Oladugba, Editor of this noble scholarly journal, Professor Nematollahi, the anonymous Associate Editor and the referees for their exceptionally insightful and constructive comments.

References

Ahmad, T., Akhtar, M. and Gilmour, S. G. (2012), Multilevel augmented pairs second-order response surface designs and their robustness to missing data. *Communications in Statistics - Theory and Methods*, **41**(3), 437–52.

Ahmad, T., and Gilmour, S. G. (2010), Robustness of subset response surface designs to missing observations. *Journal of Statistical Planning and Inference,* **140**(1), 92–103.

Akhtar, M., and Prescott, P. (1986), Response surface designs robust to missing observations. *Communications in Statistics-Simulation and Computation*, **15**(2), 345-363.

Akram, M. (2002), Central Composite Designs Robust to three missing observations, A Ph. D Thesis in Statistics, Islamia University, Baha-walpur, Pakistan.

Alrweili, H., Georgiou, S., and Stylianou, S. (2019), Robustness of response surface designs to missing data. *Quality and Reliability Engineering International*, **35**(5), 1288–1296.

Andrews, D. F., and Herzberg, A. M. (1979), The robustness and optimality of response surface designs. *Journal of Statistical Planning and Inference*, **3,** 249-257.

Arshad, H. M., Akhtar, M., and Gilmour, S. G. (2012), Augmented Box-Behnken designs for fitting third-order response surfaces. *Communications in Statistics-Theory and Methods*, **41**(23), 4225-4239.

Baş, D., Ismail H., and Boyacyı, I. H. (2007), Modeling and Optimization I: Usability of response surface methodology. *Journal of Food Engineering,* **78**, 836-845.

Box, G. E. P., and Wilson, K. B. (1951), On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B (Methodological)*, **13**(1), 1- 45.

Chen, X. P., Guo, B., Liu, M. Q., and Wang, X. L. (2017), Robustness of orthogonal-array based composite designs to missing data. *Journal of Statistical Planning and Inference*, **194**, 15-24.

Georgiou, S. D., Stylianou, S., and Aggarwal, M. (2014), A class of composite designs for response surface methodology. *Computational Statistics & Data Analysis*, **71**, 1124-1133.

Ghosh, S. (1978), On robustness *of* designs against incomplete data. *Sankhya,* B **40,** 204-2 08.

Ghosh, S. (1979), On robustness of designs against incomplete data. SankhyaÅ, Ser. B **40,** 204 -208.

Ghosh, S. (1980), On robustness of optimal balanced resolution V plans. ***J.*** *Stat.* ***Ph.*** *Inf.,* **4,** 313- 319.

Ghosh, S. (1981), Robustness of three dimensional designs. *Sankhya,* B **43,** 222-229.

Ghosh, S. (1982), Robustness of BIB designs against non-availability of data. *J Stat Plan Infer*., **6,** 29-32.

Ghosh, S., Rao, S. B., and Singhi, N. M. (1983), On a robust property of PBIB designs, *J. Statist. Plann. Inf.,* **8**, 355-364.

Hedayat, A., and John, P. (1974), Resistant and susceptible BIB designs. *Ann Stat*., **1,** 148-158.

John, P. M. W. (1976), Robustness of balanced incomplete block designs, *Ann. Statist.,* **4**, 960- 962.

Kageyama, S. (1980), Robustness of Connected Balanced Block Designs. *Ann. Inst. Statist. Math*., **32 A,** 255-261.

Lamidi, S., Olaleye, N., Bankole, Y., Obalola, A., Aribike, E., and Adigun, I. (2022), Applications of Response Surface Methodology (RSM) in Product Design, Development, and Process Optimization. *Research Advances and Applications, DOI:http://dx.doi.org/10.5772/intechopen.106763*

MacEachern, S. N., Notz W. I., Whittinghill, D. C., and Zhu, Y. (1995), Robustness to the unavailability of data in the linear model, with applications. *J. Statist. Plann. Inf*., **48,** 207- 213.

Malekjani, N., and Jafari, S. M. (2020), Food Process Modeling and Optimization by Response Surface Methodology (RSM). In: Sevda S., Singh A. (Editors.). *Mathematical and Statistical Applications in Food Engineering*, Florida: CRC Press, 181-203

Marina, S. (2013), Dealing with missing data: Key assumptions and methods for applied analysis, Technical Report 4, Boston University School of Public Health.

Mead, R., and Pike, D. J. (1975), A review of response surface methodology from a biometric viewpoint. *Biometrics*, **31**, 803–851.

Nwabueze, T. U. (2010), Basic steps in adapting response surface methodology as mathematical modelling for bioprocess optimization in the food systems. *International Journal of Food Science and Technology*, 45**,** 1768-1776.

Patchanok, S. (2015), Robust response surface designs against missing observations. A Ph. D dissertation, Montana State University, Bozeman, Montana.

Rashid, F., Akbar, A., and Arshad, H. M. (2022), Effects of missing observations on predictive capability of augmented Box-Behnken designs. *Communications in Statistics - Theory and Methods*, **51**(20), 7225-7242.

Rashid, F., Akbar, A., and Zafar, Z. (2019), Some new third order designs robust to one missing observation. *Communications in Statistics - Theory and Methods*, **48**(24)**,** 6054-6062.

Rashid, F., Akram, M., Akbar, A., and Javed, A. (2017), Some New Augmented Fractional Box-Behnken Designs. *Communications in Statistics-Theory and Methods*, **46**(4), 2007-2012.

Rebollo-Hernanz, M., Cañas, S., Taladrid, D., Segovia, A., Bartolomé, B., Aguilera, Y., and Martín-Cabrejas, M. A. (2021), Extraction of phenolic compounds from cocoa shell: Modeling using response surface methodology and artificial neural networks. *Sep. Purif. Technol.*, **270,** 118779.

Safari, M., Abdi, R., Adl, M., and Kafashan, J. (2018), Optimization of biogas productivity in lab-scale by response surface methodology. *Renewable Energy*, **118,** 368–375.

Tanco, M., del Castillo, E., and Viles, E. (2013), Robustness of three-level response surface designs against missing data. *IIE Transactions*, **45**(5), 544 - 553.

Yolmeh M, Jafari S. M. (2017), Applications of response surface methodology in the food Industry processes. *Food Bioprocess Technology*, **10,** 413-433.

Zhou, Y. D., and Xu, H. (2016), Composite designs based on orthogonal arrays and definitive screening designs. *Journal of the American Statistical Association*, ***112***(520), 1675-1683.