JIRSS (2024)
Vol. 23, No. 01, pp 33-49
DOI: 10.22034/jirss.2024.2015205.1044

Robustness of Augmented Third-order Response Surfaces Designs to Missing Observation

Brenda Mbouamba Yankam [1], **Abimibola Victoria Oladugba** [1]

[1] Department of Statistics, Faculty of Physical Sciences, University of Nigeria, Nsukka, Nigeria.

Received: 06/11/2023, Accepted: 06/07/2024, Published online: 04/11/2024

Abstract. Missing observations are practical problems that occur frequently even in a well-planned experiment and can significantly impact the statistical accuracy of the experiment. This work introduces a new class of third-order designs called augmented orthogonal uniform composite minimax loss (AOUCM) designs, which are more robust to a single missing design point as a variation of the existing third-order augmented orthogonal uniform composite designs (AOUCDs). The AOUCM designs are constructed using the minimax loss criterion. The constructed AOUCM designs are evaluated and compared with AOUCDs based on the relative D- and G-efficiency criteria, generalized scaled deviation, and the fraction of design space plot. The AOUCM designs are shown to be robust and more efficient in estimating the parameters of the third-order model. Moreover, although the AOUCDs and AOUCM designs are stable and uniformly distributed throughout the design space, the AOUCM designs have the least scaled prediction variance.

Keywords. Third-order Models, Augmented Orthogonal Uniform Composite Designs, Minimax Loss Criterion, Missing Observation, Relative Efficiency, Fraction of Design Space Plot.

MSC: 62K15, 62D10.

1 Introduction

Response surface designs (RSDs) are designs used for fitting models of either first-order, second-order, or even higher-order models such as third-order models when

Corresponding Author : Brenda Mbouamba Yankam (brenda.yankam.pg03685@unn.edu.ng)
Abimibola Victoria Oladugba (abimibola.oladugba@unn.edu.ng)

significant curvature is observed in the second-order model (Xu et al. (2014); Yankam and Oladugba (2022)). Some examples of the second-order response surface designs are orthogonal uniform composite designs (OUCDs), orthogonal array composite designs (OACDs), central composite designs, etc. Augmented orthogonal uniform composite designs (AOUCDs) are often used in a situation in which the orthogonal uniform composite designs (OUCDs) for second-order models cannot estimate the response surface model due to the lack of fit caused by third-order terms in the response surface model (Zhang et al. (2020); Yankam and Oladugba (2022)). Although, missing observations can disrupt or mar the desirable properties of these designs.

Missing observations are observations that are not present for a variable in a given dataset, and its occurrences can hardly be avoided (Chen et al. (2018); Oladugba and Nwanonobi (2021); Yankam and Oladugba (2021a)). Missing observations can significantly affect the statistical accuracy of an experiment (a measure of how well the results obtained from a sample of data reflect the true characteristics or parameters of the population from which the sample was drawn), thereby resulting in biased conclusions (Arshad et al. (2012); Ahmad and Akhtar (2014); Chen et al. (2018); Oladugba et al. (2018)). In order to minimize missing observations impacts on response surface designs researchers have figured out ways of treating them either by method of estimation (Marina (2013); Ossai and Oladugba (2018)) or analyzing them using robust criteria (Akhtar and Prescott (1986); Alrweili et al. (2019)). A response surface design is said to be robust if when constructed under a certain criterion, it helps to reduce the impacts of missing values, outliers, and non-normal errors, etc on data analysis. Akhtar and Prescott (1986) first introduced the minimax loss criterion that is based on D-efficiency, to construct designs robust against missing observation. This criterion aids in minimizing the maximum loss of a missing experimental design point resulting. Many authors have adopted the use of minimax loss criterion to construct robust designs to missing observations, (see Ahmad and Gilmour (2010); Ahmad et al. (2012); Chen et al. (2018); Alrweili et al. (2019); Rashid et al. (2019); Oladugba and Nwanonobi (2021); Yankam and Oladugba (2021a); Hemavathi et al. (2022)) etc.

In this paper, a new class of third-order response surface designs called augmented orthogonal uniform composite minimax loss (AOUCM) designs are constructed for $3 \leq k \leq 8$ factors using $1 \leq n_o \leq 5$ center points. These designs are constructed based on the structure of augmented orthogonal uniform composite designs (AOUCDs) of Yankam and Oladugba (2022) using the minimax loss criterion of Akhtar and Prescott (1986) that minimizes the maximum loss of a missing design point. The motivation for constructing these designs is to find new α values for the nonzero coordinates of the AOUCD points that minimize the impact of a missing experimental observation. This impact is assessed using a loss function approach, specifically a minimax loss based on the relative change in the determinant of the information matrix. The relative D- and G-efficiency criteria, generalized scaled deviation (GSD), and fraction of design space (FDS) plots were used to evaluate and compare the constructed AOUCM designs with

AOUCDs.

The remaining part of this work is organized as follows. In Section 2, description of AOUCDs, minimax loss criterion, and procedure for constructing robust AOUCM designs and construction of the robust designs are discussed. Design comparisons based on the relative D- and G-efficiency criteria are discussed in Section 3. Section 4 discussed the comparison based on the generalized scaled deviation (GSD). In Section 5, the fraction of design space (FDS) plot is presented, and Section 6 presents the conclusion.

2 Methodology

2.1 Augmented Orthogonal Uniform Composite Designs

Augmented orthogonal uniform composite designs (AOUCDs) introduced by Yankam and Oladugba et al. (2022) are class of third-order designs used to fit the third-order model. The AOUCDs exist for cuboidal and spherical regions for $k \geq 2$, where k is the number of factors. The AOUCDs have three components; the factorial portion (n_f), the axial portion (n_α), and the center portion (n_o). The factorial portion used are either a 2^k full factorial with levels at ± 1 or a $2^{\,k-l}$ resolution V fractional factorial, in which neither a main effect nor a two-factor interaction is aliased with another main effect or two-factor interaction. This portion is used mainly for the estimation of the linear terms and two-factor interactions. The axial portion, also called the star points or additional points, is located at a distance of α from the design center, and are four-level uniform designs ($-\alpha$,$-\frac{\alpha}{3}$,$\frac{\alpha}{3}$ and α), (α is the distance between the center and the non-zero coordinate in the axial design point). The axial portion mainly contributes to the estimation of quadratic and cubic terms. The center portion is located in the center of the design space and helps estimate the pure error. The AOUCDs are good space-filling and efficient designs and can perform multiple analyses with different portions of the design for cross validation. A summary of the choice of the factorial and axial portions in the construction of AOUCDs for $3 \leq k \leq 8$ factors are as follows:

- Case 1: For $k = 3$ and 4, the AOUCDs are constructed by selecting two-level full factorial designs 2^3 and 2^4 made of 8 and 16 runs respectively combined with four-level uniform designs made of UD (16, 4^3), where UD represents the uniform design constructed from the R package named uniform design of experiment (UniDOE).

- Case 2: For $k = 5$, the AOUCD is constructed by selecting two-level fractional factorial designs 2_V^{5-1} of resolution V with 16 runs and four-level uniform designs UD (16, 4^5) of 16 runs.

- Case 3: For $k = 6$, the AOUCD is constructed by selecting two-level fractional factorial designs 2_{VI}^{6-1} of resolution VI with 32 runs and four-level uniform designs UD (32, 4^6) of 32 runs.

- Case 4: For $k = 7$ and 8, the AOUCD is constructed by selecting two-level fractional factorial designs, 2^{7-1}_{VII} of resolution VII and 2^{8-2}_{V} of resolution V with 64 runs and four-level uniform designs UD (32, 4^7) and UD (32, 4^8) respectively of 32 runs.

For full factorial design, the geometry and design matrix of the AOUCD for k = 3 are shown in Figure 1 and Table 1, respectively. Also, for fractional factorial design, the design matrix of the AOUCD for k = 5 is shown in Table 2.

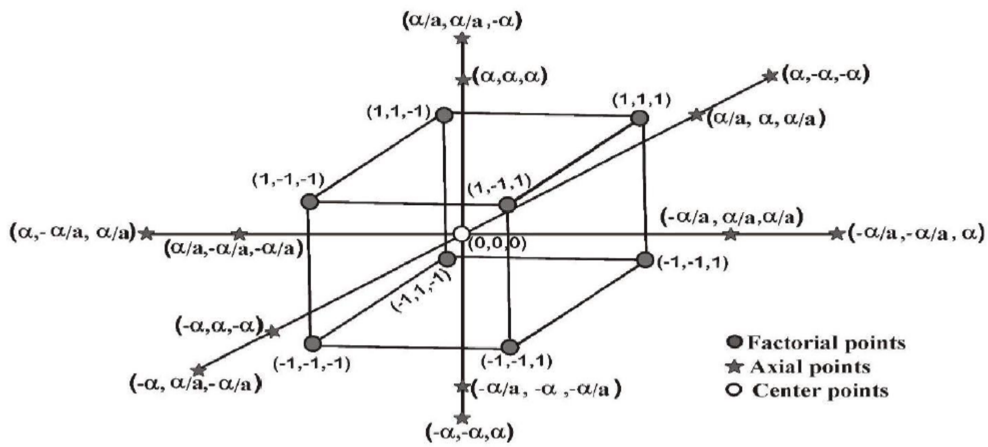

Figure 1: The geometry of a three-factor AOUCD.

Table 1: A typical layout of 26-runs AOUCD for k = 3 factors with n_o = 2 center point.

Factorial	portion	(n_f = 8)	Axial	portion	(n_α = 16)	Center	point	(n_o = 2)
-1	1	1	α	-α	α	0	0	0
1	-1	1	α	-$\frac{\alpha}{3}$	-α	0	0	0
1	-1	-1	α	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
-1	-1	-1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-α			
1	1	1	$\frac{\alpha}{3}$	-$\alpha(\alpha)$	$\frac{\alpha}{3}$			
-1	-1	1	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
1	1	-1	-α	α	α			
-1	1	-1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$			
			-α	α	α			
			-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	α			
			-α	-α	-$\frac{\alpha}{3}$			
			-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
			$\frac{\alpha}{3}$	-α	$\frac{\alpha}{3}$			
			-$\frac{\alpha}{3}$	α	-α			
			-α	$\frac{\alpha}{3}$	-α			
			α	α	$\frac{\alpha}{3}$			
where	$\alpha = 1$							

Table 2: A typical layout of 34-runs AOUCD for k = 5 factors with n_o = 2 center point.

	Factorial	portion	(n_f = 16)			Axial	portion	(n_α = 16)			Center	point	(n_o = 2)	
1	-1	-1	-1	-1	α	-α	α	α	α	0	0	0	0	0
-1	1	-1	-1	-1	α	-$\frac{\alpha}{3}$	-α	-α	-α	0	0	0	0	0
-1	-1	1	-1	-1	α	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$					
-1	-1	-1	1	-1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-α	-α	-α					
-1	-1	-1	-1	1	$\frac{\alpha}{3}$	-α α	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$					
1	1	1	-1	-1	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$					
1	1	-1	1	-1	-α	α	α	α	α					
1	1	-1	-1	1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$					
1	-1	1	1	-1	-α	α	α	α	α					
1	-1	1	-1	1	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	α	α	α					
1	-1	-1	1	1	-α	-α	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$					
-1	1	1	1	-1	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$					
-1	1	1	-1	1	$\frac{\alpha}{3}$	-α	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$					
-1	1	-1	1	1	-$\frac{\alpha}{3}$	α	-α	-α	-α					
-1	-1	1	1	1	-α	$\frac{\alpha}{3}$	-α	-α	-α					
1	1	1	1	1	α	α	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	$\frac{\alpha}{3}$					
where	α = 1													

2.2 Model Description

The third-order model given k factors is defined as:

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

where y denotes the response of interest, $\beta's$ the unknown parameters, x's is the coded units of k factors, and ε the random experimental error assumed to follow a normal distribution with a mean of zero (0) and variance, σ^2.

According to Draper (1962), Cornell and Montgomery (1996), Arshad et al. (2012), Arshad et al. (2018) and Zhang et al. (2018), if the number of parameters p of the third-order model is too large, the third-order interaction parameters such as $x_1x_2x_3$,$x_1^2x_2$,. . . are frequently ignored by some researchers in order to obtain economical designs as is the case in this work. The reduced model used in this work is defined as:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{iii} x_i^3 + \varepsilon. \tag{2.2}$$

Equation (2.2) can be expressed in matrix form as:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{2.3}$$

where $y = (y_1, \ldots, y_n)'$ represents the observed response's vector, X model matrix $n \times p$, β the $p \times 1$ vector of parameters, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^{'}$ the vector of random error associated with y. The error, ε, is a random vector assumed to follow a multivariate normal distribution with E(ε) = 0 and Var (ε) = $\sigma^2 I_n$.

The model matrix X is an $n \times p$ corresponding to the parameters of the third-order model is partitioned as:

$$X = (1, X_l, X_q, X_b, X_c), \tag{2.4}$$

where 1 represents the model's constant term which denotes an $n \times 1$ vector, X_l, X_q, X_b, and X_c represent the linear, quadratic, bilinear, and cubic coefficients that denote the $n \times k$, $n \times k$, $n \times (k - 1)k/2$, and $n \times k$ vector-matrix respectively.

2.3 Minimax Loss Criterion

The minimax loss criterion introduced by Akhtar and Prescott (1986), is a criterion (that is, based on D-efficiency) for reducing the losses from missing observations by minimizing the maximum loss due to missing observations. In other words, the minimax loss criterion is a criterion used in finding new values of α (that makes a response surface design robust to missing observations) such that the maximum loss of missing design points is minimized.

In order to use the minimax loss criterion, we suppose t observations (design points) are missing in a design, this implies that the model matrix X is reduced by t rows. Let X_t be the $t \times p$ matrix of the t missing rows corresponding to the t missing (incomplete) observations and X_R be a matrix of the $n-t$ remaining rows of the model matrix X for the reduced design (i.e., not containing the missing observations). The information matrix for the full, incomplete and reduced designs will be $X'X$, $X_t'X_t$ and $X_R'X_R$ (Ahmad et al., 2012; Chen et al., 2018; Yankam and Oladugba, 2021a). Also, the model matrix is partitioned as $X = \begin{pmatrix} X_t' \\ X_R \end{pmatrix}$, and the information matrix is rewritten as:

$$X'X = X_t'X_t + X_R'X_R. \tag{2.5}$$

Let g denote the determinant of the information matrix of the full design ($g = |X'X|$) and g_R denote the determinant of the information matrix of the reduced design ($g_R = \left|(X_R'X_R)\right|$). Akhtar and Prescott (1986) defined loss as the relative reduction in the determinant of the information matrix due to missing observation. In other words, the loss due to missing one or more observation(s) is the relative reduction in the determinant of the information matrix of the designs.

The loss (L_j) due to missing an observation relative to the complete observation is defined as:

$$L_j = \frac{\left|X'X\right| - \left|X_R'X_R\right|}{|X'X|} = 1 - \frac{g_R}{g}. \tag{2.6}$$

The loss, L_j, takes the range $0 \leq L_j \leq 1$ and it is a relative measure of efficiency. When $L_j = 0$, it implies that there is no reduction in the determinant of the information matrix of the design. When $L_j = 1$, it implies that the design has broken down owing to the unavailability of the j^{th} design point which prevents the estimation of the model coefficients. Note that L_j represents the loss of a factorial point L_f, the loss of an axial point L_{α}, and the loss for a center point L_o.

2.4 Procedure for Construction of AOUCM Designs

The AOUCM designs are constructed by:

Step 1: Selecting an initial design matrix, $X = \begin{pmatrix} U_1 \\ U_2 \\ U_o \end{pmatrix}$, without missing observation, where U_1 is the factorial portion with n_f runs, U_2 the axial portion with n_α runs, and U_o the center points with n_o runs. In this paper, n_o is $1 \leq n_o \leq 5$ center points.

Let $X^* = \begin{pmatrix} U_1 \\ \alpha U_2 \\ U_o \end{pmatrix}$ be the robust design, where α in αU_2 is the new value of α used to construct the robust design to a missing observation. The missing observation occurs at the factorial, axial, and center portions one at a time.

Step 2: Calculate the loss functions, L_f, $L_{\alpha,}$ and L_o, using equation (2.6), based on X^* for a missing observation, where L_f, $L_{\alpha,}$ and L_o are the loss of the factorial, loss of axial and loss of center point, respectively.

Step 3: Finding the new α that minimizes the maximum loss L_f, $L_{\alpha,}$ and L_o.

Step 4: Substituting the new α in X^*, an augmented orthogonal uniform composite minimax loss (AOUCM) design with $n = n_f + n_\alpha + n_o$ experimental runs and k factors is constructed.

Remark 1. The bisection method may be used to search for new alpha values (α) in Step 3 (Ahmad and Gilmour (2010); Ahmad et al. (2012); Chen et al. (2018)). However, in this paper, the R code in the Appendix was used in plotting the graph of the loss function (L_f, $L_{\alpha,}$ and L_o) in order to obtain the new alpha values. The new alpha values are obtained from the plot at the point where the L_f and L_α equals.

2.5 Construction of Robust Designs

In order to illustrate the construction of augmented orthogonal uniform composite minimax loss (AOUCM) designs robust to a single missing observation, we will consider an example based on the AOUCD structure in Table 1. For instance, an AOUCD for $k = 3$ factors has a total of $n = 26$ design points, comprising $n_f = 8$ factorial points, $n_\alpha = \text{UD}\,(16, 43) = 16$ axial points, and two center points ($n_o = 2$). In order to assess the robustness of AOUCM designs, we compute the determinant of the information matrix for the full design ($g = |X'X|$) and the loss for factorial points (L_f), axial points (L_α), and center points (L_o) at different α values, ranging from 0.50 to 2.25 using equation (2.6). The results are summarized in Table 3.

Table 3 reveals distinct trends: L_f decreases as α increases, L_α increases with increasing α, and L_0 initially increases, then decreases as α increases. For the purpose of visualizing these trends, the loss functions L_f, L_α, and L_o are plotted as shown in Figure 2, employing the R code provided in the Appendix. Figure 2 shows that as α increases from 0.50 to 2.25, L_f decreases steadily while L_α increases continuously with increasing α, and L_o

shows a bell-shaped curve, remaining relatively constant in the mid-range of α values. Notably, $1.2 \leq \alpha \leq 1.3$ is the range where L_f and L_α become equal for $k = 3$ with $n_o = 2$ center points. Using the R code, we determine that the maximum loss of missing a design point is minimized at $\alpha = 1.2333$ for $k = 3$ with $n_o = 2$ center points. A typical design layout of the AOUCM design for $k = 3$ with $n_o = 2$ center points is presented in Table 5. This construction process ensures that the AOUCM designs are robust to the impact of a single missing observation, enhancing their reliability in response surface experiments.

Table 3: A typical layout of 34-runs AOUCD for $k = 5$ factors with $n_o = 2$ center point.

α	g	L_f	L_α	L_0
0.50	4.2278×10^{6}	0.8563	0.3347	0.0741
0.55	1.0871×10^{7}	0.8484	0.3330	0.0793
0.60	2.3651×10^{7}	0.8385	0.3314	0.0856
0.65	4.4116×10^{7}	0.8267	0.3298	0.0930
0.70	7.1137×10^{7}	0.8132	0.3281	0.1018
0.75	9.9795×10^{8}	0.7984	0.3260	0.1119
0.80	1.2307×10^{8}	0.7826	0.3230	0.1232
0.85	1.3758×10^{8}	0.7663	0.3180	0.1356
0.90	1.5309×10^{8}	0.7479	0.3102	0.1485
0.95	2.1302×10^{8}	0.7184	0.3034	0.1610
1.00	5.1211×10^{8}	0.6601	0.3161	0.1723
1.05	2.3630×10^{9}	0.5820	0.3572	0.1840
1.10	1.6192×10^{10}	0.5202	0.3972	0.1953
1.15	1.2274×10^{11}	0.4809	0.4218	0.2036
1.20	8.9537×10^{11}	0.4556	0.4359	0.2080
1.25	6.0265×10^{12}	0.4386	0.4451	0.2089
1.30	3.7187×10^{13}	0.4268	0.4520	0.2071
1.35	2.1130×10^{14}	0.4187	0.4575	0.2032
1.40	1.1132×10^{15}	0.4132	0.4622	0.1979
1.45	5.4759×10^{15}	0.4095	0.4662	0.1919
1.50	2.5309×10^{16}	0.4070	0.4697	0.1854
1.55	1.1051×10^{17}	0.4054	0.4728	0.1789
1.60	4.5802×10^{17}	0.4043	0.4755	0.1725
1.65	1.8094×10^{18}	0.4036	0.4778	0.1664
1.70	6.8377×10^{19}	0.4031	0.4799	0.1607
1.75	2.4796×10^{19}	0.4027	0.4817	0.1553
1.80	8.6537×10^{19}	0.4024	0.4833	0.1503
1.85	2.9137×10^{20}	0.4022	0.4847	0.1457
1.90	9.4859×10^{20}	0.4019	0.4859	0.1414
1.95	2.9922×10^{21}	0.4017	0.4870	0.1376
2.00	9.1617×10^{21}	0.4014	0.4880	0.1340
2.05	2.7274×10^{22}	0.4012	0.4889	0.1307
2.10	7.9061×10^{22}	0.4009	0.4897	0.1277
2.15	2.2347×10^{23}	0.4006	0.4904	0.1249
2.20	6.1667×10^{23}	0.4003	0.4910	0.1224
2.25	1.6633×10^{24}	0.3999	0.4915	0.1201

Remark 2. Similarly, based on the preceding procedures, $g = |X'X|$, L_f, $L_{\alpha,}$ and L_o for other cases were computed at different α values $0.50 \leq \alpha \leq 2.25$, to obtain the new α values. These values are presented in Table 4.

From Table 4 it is observed that all the new α values are higher than the initial $\alpha = 1$ in

AOUCDs. Furthermore, as the number of factors (k) and center points (n_o) increase the new α values either increase or decrease without a definite pattern except for k = 3 and 4 where the value decreases from 1.2335 to 1.2312 and 1.1158 to 1.1111 respectively.

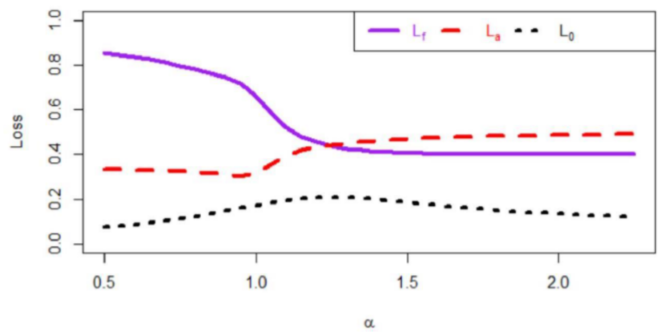

Figure 2: Choice of α for the AOUCD for 3 factors AOUCD with n_o = 2 center points.

Table 4: New α values for $3 \leq k \leq 8$ using $1 \leq n_o \leq 5$ center points.

	Number	of	factors	(k)		
n_o	3	4	5	6	7	8
1	1.2335	1.1158	1.4991	1.3580	1.4056	1.0826
2	1.2333	1.1105	1.4693	1.3581	1.4123	1.0858
3	1.2321	1.1103	1.4676	1.3592	1.4187	1.0859
4	1.2317	1.1100	1.4688	1.3612	1.4211	1.0739
5	1.2312	1.1111	1.4632	1.3610	1.4200	1.1000

Table 5: A typical design layout of 26-runs AOUCM design for k = 3 factors with n_o = 2 center point.

Ffactors								
Factorial	portion	(n_f = 8)	Axial	portion	(n_α = 16)	Center	point	(n_o = 2)
-1	1	1	α	-α	α	0	0	0
1	-1	1	α	-$\frac{\alpha}{3}$	-α	0	0	0
1	-1	-1	α	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
-1	-1	-1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	-α			
1	1	1	$\frac{\alpha}{3}$	-α α	$\frac{\alpha}{3}$			
-1	-1	1	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
1	1	-1	-α	α	α			
-1	1	-1	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$			
			-α	α	α			
			-$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$	α			
			-α	-α	-$\frac{\alpha}{3}$			
			-$\frac{\alpha}{3}$	$\frac{\alpha}{3}$	-$\frac{\alpha}{3}$			
			$\frac{\alpha}{3}$	-α	$\frac{\alpha}{3}$			
			-$\frac{\alpha}{3}$	α	-α			
			-α	$\frac{\alpha}{3}$	-α			
			α	α	$\frac{\alpha}{3}$			
where α = 1.2333								

3 Comparison Based on the Relative D- and G-efficiencies

The relative D- and G-efficiencies of AOUCM design to AOUCD for $3 \le k \le 8$ using $1 \le n_o \le 5$ center points are obtained using the following definitions. The relative D- and G-efficiency values are presented in Table 6.

3.1 Relative D-efficiency

The relative D-efficiency (RD) of a design with missing observation ξ_B to a design without missing observation ξ_C is the ratio of the separate D-efficiencies and is defined as:

$$RD_{B/C} = \left[\frac{|M(\xi_B)|}{|M(\xi_D^*)|}\right]^{\frac{1}{p}} \div \left[\frac{|M(\xi_C)|}{|M(\xi_D^*)|}\right]^{\frac{1}{p}} = \left[\frac{|M(\xi_B)|}{|M(\xi_C)|}\right]^{\frac{1}{p}}, \tag{3.1}$$

where p represents the parameter of the model under study and ξ_D^* is D-optimal design.

3.2 Relative G-efficiency

The relative G-efficiency (RG) of a design with missing observation ξ_B to a design without missing observation ξ_C is the ratio of the separate G-efficiencies and is defined as:

$$RG_{B/C} = \frac{p}{v(\underline{x})_{\max}(\xi_B)} \Big/ \frac{p}{v(\underline{x})_{\max}(\xi_C)} = \frac{v(\underline{x})_{\max}(\xi_C)}{v(\underline{x})_{\max}(\xi_B)}, \tag{3.2}$$

where $v(\underline{x})_{max} = \max\left(n.var\frac{[\hat{y}(x)]}{\sigma^2}\right) = \left(n.x'\left(X'X\right)^{-1}x'\right)$ is the scale prediction variance, with x being the vector of design points in the design space, $x = (1, x_1, x_2, \ldots, x_k; x_1^2, x_2^2, \ldots, x_k^2; x_1, x_2, \ldots, x_{k-1}, x_k; x_1^3, x_2^3, \ldots, x_k^3)$, and also B is AOUCM design and C is AOUCD.

Remark 3. When two response surface designs are compared, the $RD_{B/C}$, and $RG_{B/C}$ values greater than one is often desirable implying that AOUCM designs are better than AOUCDs in terms of design efficiency.

From Table 6, it was observed that the AOUCM designs perform better with the highest D-efficiency values in all cases for $3 \le k \le 8$ with $1 \le n_o \le 5$ center points. Furthermore, AOUCM designs showed superior performance in terms of the relative G-efficiency values for $k = 6, 7$, and 8, but exhibited lesser performance for $k = 3, 4$, and 5. This implies that AOUCM designs are more efficient than AOUCDs in most cases considered in this paper. The superiority of the AOUCM designs could be attributed to the utilization of new α values (i.e., $\alpha > 1$) compared to $\alpha = 1$ used in constructing AOUCDs. Hence, the AOUCM designs can be seen as special cases of the AOUCDs by carefully choosing the new α values.

Table 6: A typical design layout of 26-runs AOUCM design for k = 3 factors with n_o = 2 center point.

Number	of	factors	(k)			
n_o	3	4	5	6	7	8
1	1.9540 [0.9002]	1.2700 [0.9346]	2.8421 [0.8865]	2.1714 [1.1629]	2.2256 [1.0851]	1.1129 [1.0338]
2	1.9597 [0.9005]	1.2574 [0.9098]	2.7094 [0.9192]	2.1785 [1.1624]	2.2629 [1.0838]	1.1198 [1.0341]
3	1.9574 [0.9074]	1.2610 [0.9366]	2.7152 [0.9215]	2.1884 [1.5029]	2.2959 [1.0816]	1.1208 [1.0345]
4	1.9585 [0.9011]	1.2627 [0.9373]	2.7307 [0.9210]	2.2014 [1.1599]	2.3104 [1.0811]	1.0987 [1.0347]
5	1.9582 [0.9013]	1.2689 [0.9362]	2.7076 [0.9269]	2.2032 [1.1595]	2.3092 [1.0808]	1.1505 [1.0353]
G-efficiency	values are	presented	in the	square brackets		

4 Comparison Based on the Generalized Scaled Deviation

In this section, the generalized scaled deviations (*GSDs*) for estimating the full model, linear, quadratic, bilinear, and cubic coefficients for the AOUCDs and AOUCM designs were obtained for $3 \leq k \leq 8$ using $n_o = 5$ and the results are presented in Table 7. The values of α obtained by the minimax loss criterion for the AOUCM designs are listed in the second column

The generalized scaled deviation (*GSD*) is a criterion that aids in estimating a subset of the model parameters in terms of precision. Response surface designs that minimize *GSD*, are D-optimal for a constant value of n number of runs, and those that minimize GSD_l, GSD_q, GSD_b, and GSD_c are D_s-optimal for estimating the corresponding subsets of coefficients, where *GSD*, GSD_l, GSD_q, GSD_b, and GSD_c are, the generalized scaled deviation for the full, linear, quadratic, bilinear and cubic coefficients, respectively, for the third-order model. These measures are defined respectively as:

$$GSD = \sqrt{n|X'X|^{-1/p}}, \tag{4.1}$$

$$GSD_l = \sqrt{n|X_l'X_l - X_l'\bar{X}_l(\bar{X}_l\bar{X}_l)^{-1}\bar{X}_l'X_l|^{-1/k}}, \tag{4.2}$$

$$GSD_q = \sqrt{n|X_q'X_q - X_q'\bar{X}_q(\bar{X}_q'\bar{X}_q)^{-1}\bar{X}_q'X_q|^{-1/k}}, \tag{4.3}$$

$$GSD_b = \sqrt{n|X_b'X_b - X_b'\bar{X}_b(\bar{X}_b'\bar{X}_b)^{-1}\bar{X}_b'X_b|^{-2/k(k-1)}}, \tag{4.4}$$

and

$$GSD_c = \sqrt{n|X_c'X_c - X_c'\bar{X}_c(\bar{X}_c'\bar{X}_c)^{-1}\bar{X}_c'X_c|^{-1/k}}, \tag{4.5}$$

where $\overline{X}$ is the complement of X. The design with the least *GSD* is often desired as it indicates good performance compared to other designs. From Table 7, it was observed that the AOUCM designs performed better in estimating the full model, linear, quadratic, bilinear, and cubic coefficients for $3 \leq k \leq 8$ factors.

Table 7: Generalized scaled deviations of AOUCDs and AOUCM designs.

k	α	Design	n	p	GSD	GSD_l	GSD_q	GSD_b	GSD_c
3	1.2312	AOUCD	29	13	2.4505	8.8292	3.0413	1.6397	9.0566
		AOUCM	29	13	1.7512	4.8743	2.1298	1.4860	3.8535
4	1.1111	AOUCD	37	19	2.358	10.4530	3.3666	1.4147	10.6229
		AOUCM	37	19	2.0939	8.2443	2.8216	1.3941	7.7637
5	1.4632	AOUCD	37	26	2.2763	9.8254	3.7763	1.4514	9.9550
		AOUCM	37	26	1.3833	3.2720	2.1403	1.2981	2.2793
6	1.3610	AOUCD	69	34	2.1031	8.8605	3.6828	1.3496	8.9885
		AOUCM	69	34	1.4168	3.5907	2.2887	1.1758	2.6971
7	1.4200	AOUCD	101	43	2.0245	11.3249	4.4290	1.2063	6.3527
		AOUCM	101	43	1.3319	3.1805	2.3440	1.0998	2.4270
8	1.1000	AOUCD	101	53	1.9923	11.6181	4.8950	1.2137	11.6921
	1.1000	AOUCM	101	53	1.8574	9.6025	4.2464	1.1984	9.3455

5 Comparison Based on Fraction of Design Space Plot

The variance dispersion graph (VDG) is an effective technique for comparing and assessing the experimental design's prediction variance. However, it does not account for differences in proportions of the full design area among concentric spheres for every radius (Fang et al., 2019; Oladugba & Yankam, 2021; Yankam & Oladugba, 2021b). Zahran et al. (2003) introduced the fraction of design space (FDS) plot to complement the VDG. The FDS plot displays the characteristics of the scaled prediction variance (SPV) in a fraction of design space on a two-dimensional graph with a single curve, allowing easy comparison of multiple designs on a single plot (Zahran et al., 2003). It enables the researchers to evaluate a design's prediction capability in showing the distribution of the variance of all predicted values.

The FDS is calculated by sampling points throughout the design spaces; thereafter, the SPV is computed from each design point. The cumulative fraction is then calculated for the SPV values and plotted on the FDS plot. The y-axis of the FDS plot displays the SPV values ranging from minimum to maximum SPV and the x-axis depicts the quantiles of the design space (which range from 0 to 1). According to Zahran et al. (2003), a design with a stable and consistent SPV is often desired over those that are not. Moreover, the smaller the fraction of design space at any given value, the better the design. In this work, the FDS plots for $3 \leq k \leq 8$ factors using $n_o = 2$ center points are presented in Figures 3 to 8.

From Figures 3 to 8, it was observed that the AOUCM designs for $3 \leq k \leq 8$ performed better with the least SPV and are stable throughout the design space. Although AOUCDs and AOUCM designs are uniformly distributed and stable throughout the design space, the AOUCM designs which are variations of the AOUCDs perform better with the least SPV. Moreover, it was observed that at 50% of the total design space, the SPV value of the AOUCM designs for $k = 3$, 6, 7, and 8 is below 100% G-efficiency as shown in Figures 3, 6, 7, and 8. Meanwhile, at 50% of the total design space, the SPV value of the AOUCM designs for $k = 4$, and 5 is at 100% G-efficiency as shown in Figures 4, and 5.

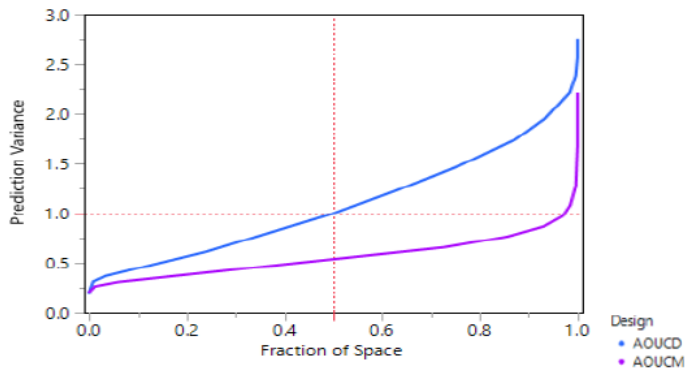

Figure 3: FDS plot of AOUCD and AOUCM design for $k = 3$.

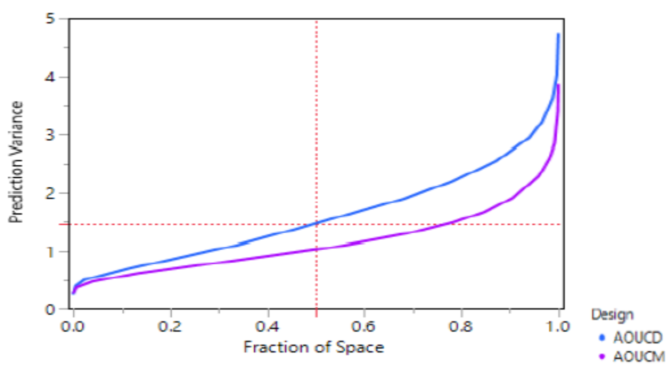

Figure 4: FDS plot of AOUCD and AOUCM design for $k = 4$.

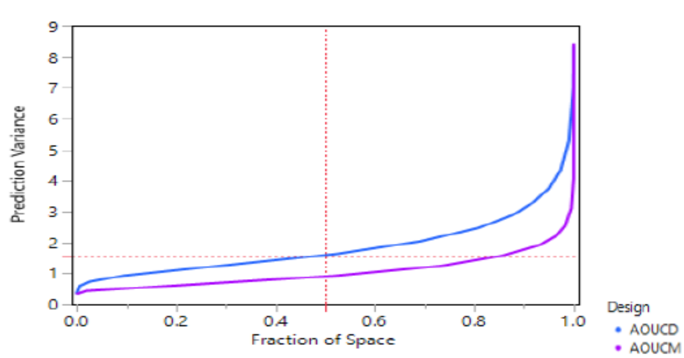

Figure 5: FDS plot of AOUCD and AOUCM design for $k = 5$.

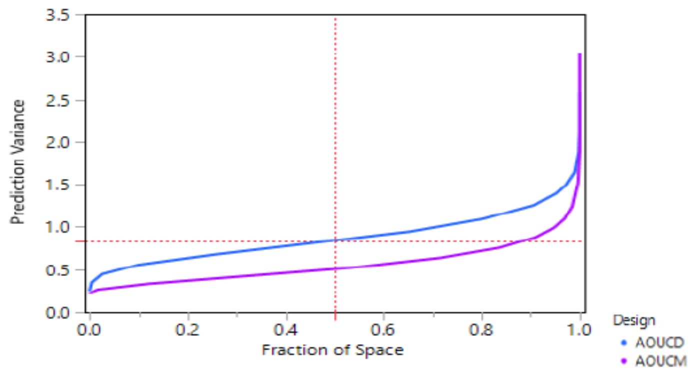

Figure 6: FDS plot of AOUCD and AOUCM design for $k = 6$.

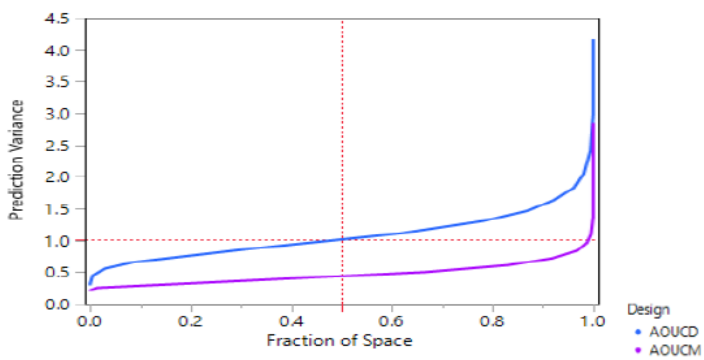

Figure 7: FDS plot of AOUCD and AOUCM design for $k = 7$.

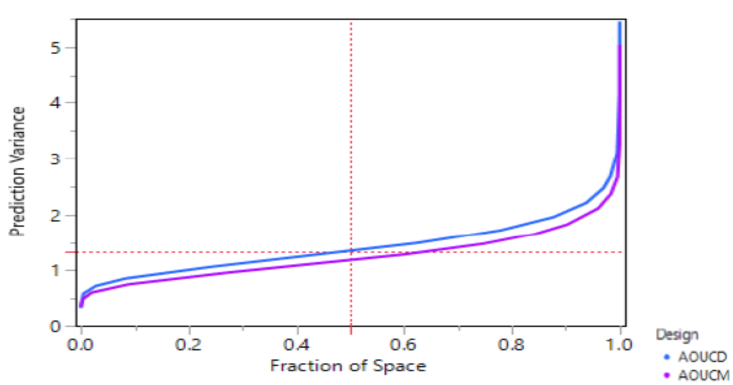

Figure 8: FDS plot of AOUCD and AOUCM design for $k = 8$.

6 Conclusion

It can be challenging to choose the right response surface model to fit specific polynomial regression models. A design that performs well in the absence of an experimental observation is always desirable.

In this paper, the issue of augmented orthogonal uniform composite designs (AOUCDs) with missing observation was considered. Augmented orthogonal uniform composite minimax loss (AOUCM) designs are constructed, which are more robust to missing an observation than AOUCDs. The construction used the minimax loss criterion to select new α values for a whole set of AOUCD points in the experimental region. In general, the AOUCM designs were shown to be efficient in estimating the parameters of third-order models. Also, the AOUCM designs while retaining the properties of the AOUCDs can be used with and without missing observation. Moreover, although the AOUCDs and AOUCM designs were stable and uniformly distributed throughout the experimental region, the AOUCM designs had the least SPV throughout the design space.

Based on this work, we recommend further exploration of the AOUCDs to model misspecification, two or more missing observations, and the percentage of missingness. Also, the construction of AOUCDs and AOUCM designs with lesser number of runs that provide the needed number of degrees of freedom for estimating lack of fit.

Acknowledgment
This work was supported by the German Academic Exchange Service (DAAD), Bonn, Germany (Funding number: 57423580 ()-0002).

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), 437–452.

Ahmad, T., and Akhtar, M. (2014). Efficient response surface designs for the second-order multivariate polynomial model robust to missing observation. *Journal of Statistical Theory and Practice*, **9**(2), 361–375.

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.

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.

Arshad, H. M., Ahmad, T., and Akhtar, M. (2018). Some sequential third order response surface designs. *Communications in Statistics: Simulation and Computation*, **49**(7), 1872–1885.

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.

Chen, X. P., Guo, B., Liu, M. Q., and Wang, X. L. (2018). Robustness of orthogonal-array based composite designs to missing data. *Journal of Statistical Planning and Inference*, **194**, 15–24.

Cornell, J. A., and Montgomery, D. C. (1996). Interaction models as alternatives to low-order polynomials. *Journal of Quality Technology*, **28**(2), 163–176.

Draper, N. R. (1962). Third order rotatable designs in three factors: Analysis'. *Technometrics*, **4**(2), 219–234.

Fang, J., He, Z., He, S., and Wang, G. (2019). The robustness of response surface designs with errors in factor levels. *Communications in Statistics - Theory and Methods*, **49**(10), 2365–2383.

Hemavathi, M., Varghese, E., Shekhar, S., Athulya, C. K., Ebeneezar, S., Gills, R., and Jaggi, S. (2022). Robustness of sequential third-order response surface design to missing observations. *Journal of Taibah University for Science*, **16**(1), 270-279.

Marina, S. (2013). Dealing with missing data: Key assumptions and methods for applied analysis. School of Public Health Department, Boston University. Technical Report No. 4.

Oladugba, A. V., and Nwanonobi, O. C. (2021). Robustness of definitive screening composite designs to missing observations. *Communications in Statistics-Theory and Methods*, 1-15.

Oladugba, A. V., and Yankam, B. M. (2021). Prediction variance properties of third-order response surface designs in the hypersphere. *Mathematics and Statistics*, **10**(1): 61-73.

Ossai, E. O., and Oladugba, A. V. (2018). The estimation of missing values in rectangular lattice designs. *Journal of Modern Applied Statistical Methods*, **17**(2), eP2670.

Oladugba, A. V., Ossai E. O., and Ugah T. E.(2018). On the Comparison of methods of estimating missing values in rectangular lattice designs. *Journal of Mathematics and Statistics*, **14**(1), 201–208.

Rashid, F., Akbar, A., and Zafar, Z. (2019). Some new third order designs robust to one missing observation. *Communications in Statistics—Theory and Methods*, 1–9.

Xu, H., Jaynes, J., and Ding, X.(2014). Combining two-level and three-level orthogonal arrays for factor screening and response surface exploration. *Statistica Sinica*, **24**(1), 269–289.

Yankam, B. M., and Oladugba, A. V. (2021a). Robustness of orthogonal uniform composite designs against missing data. *Communications in Statistics - Theory and Methods*, 1–16.

Yankam,B. M., and Oladugba, A. V.(2021b). Prediction variance capabilities of third-order response surface designs for cuboidal regions. *Mathematics and Statistics*, **9**(5), 760-772.

Yankam,B.M., and Oladugba,A.V.(2022). Augmented orthogonal uniform composite designs for fitting third-order model.*Journal of Statistical Theory and Practice*,**16**(2),1–14.

Zahran, A., Anderson-Cook, C. M., and Myers, R. H. (2003). Fraction of design space to assess prediction capability of response surface designs. *Journal of Quality Technology*, **35**(4), 377–386.

Zhang, X. R., Liu, M., and Zhou, Y. (2020). Orthogonal uniform composite designs. *Journal of Statistical Planning and Inference*, **206**, 100–110.

Zhang, X. R., Qi, Z. F., Zhou, Y. D., and Yang, F. (2018). Orthogonal-array composite design for the third-order models. *Communications in Statistics - Theory and Methods*, **47**(14), 3488–3507.

Appendix: R code to plot the graph and obtain Minimax loss design and new alpha (α) value

```
Data = read.csv ("k3.csv", header=T)

attach (Data)

plot (a, Lf, type="l", lty=1, ylim=c(0,1), ylab="Loss", xlab=expression(alpha), col="purple",
lwd=2)

lines (a, La, type="l", lty=2, lwd=2, col="red")

lines (a, L0, type="l", lty=3, lwd=2, col="black")

legend.txt= c(expression(L[f]), expression(L[a]), expression(L[0]))

legend ("topleft", legend =legend.txt, lty=c(1,2,3), col= c("purple", "red", "black"),
ncol=2, lwd=2, text.col=c("purple", "red", "black"))

abline (v=1.2333, h=0.5321, col="blue")
```