FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 4, April, 2025, pp 295 – 299
DOI: https://doi.org/10.33003/fjs-2025-0904-3356

ON THE DEVELOPMENT OF MINIMAX LOSS AUGMENTED BOX-BEHNKEN DESIGN ROBUST TO TWO MISSING OBSERVATIONS

***[1,2]Oluchukwu C. Asogwa and [2]Oluwaseun A. Otekunrin**

[1]Department of Mathematics and Statistics, Alex Ekwueme Federal University, Ndufu-Alike, Nigeria
[2]Department of Statistics, University of Ibadan, Ibadan, Nigeria

*Corresponding authors' email: qackasoo@yahoo.com

ABSTRACT

Augmented Box-Behnken Designs (ABBDs) are efficient third-order designs for complete observations. Previous studies have focused on the development of minimax loss ABBDs for one missing observation. In practice, two or more observations may be missing, leading to design inefficiency and generation of invalid inference. This study was therefore aimed at developing minimax loss ABBD that is robust to two missing observations for more valid inferences. A complete ABBD (k, n_c, n_b, n_f, n_a, α) was adopted, where k, n_c, n_b, n_f, n_a and α are the number of factors, centre, Box-Behnken, factorial, axial points and axial distance, respectively. Missing observations were introduced into the ABBD by removing all possible combinations of two observations, from each of n_c, n_b, $n_f$and $n_a$points of the ABBD. Minimax loss criterion was used to compute the losses in precision for the ABBD when two observations are missing from each point. A greed search algorithm was used to determine the αvalue that minimised the maximum loss in precision. A minimax loss ABBD was obtained when the maximum loss in precision between any two different types of points was equal at a particular value of α. The loss in precision values when two observations were missing at α = 1.6021were 0.5259, 0.8833, 0.9664 and 0.9664 and the α value which minimised the maximum losses in precision was1.6021. The minimax loss ABBD obtained was minimax loss ABBD robust to two missing observations (k=3,$n_c = 3$, $n_b = 12$, $n_f = 8$, $n_a = 6$, α =1.6021).

Keywords: Robustness of designs, Missing observations, Third-order designs

INTRODUCTION

Response surface methodology is an efficient modern statistical techniques which emanated from Box and Wilson (1951). It has widely been used in building models and exploring relationships between one or more response variables and several explanatory variables (Zhou and Xu, 2016; Ahmad and Gilmour, 2010).Its application has cut across many fields of life, including; manufacturing industries, Lamidi *et al.* (2022); chemical and food industries, Yolmeh *et al.* (2017); biological and biomedical, Safari *et al.* (2018); biopharmaceutical, Rebollo-Hernanz *et al.* (2021); agriculture, Mead and Pike (1975); etc. However, in response surface methodology, second-order models, in some cases may be inadequate and unrealistic because of 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 presents a need for a third-order or higher-order model to be augmented from the second order response surface designs in order to estimate the third-order or higher-order terms in the response surface model. Arshad *et al.* (2012) constructed an augmented Box-Behnken designs which can handle the estimates of the third-order models and these are known as third-order designs. These third-order designs in general have been useful in many experimental situations for response surface modelling and however may have been confronted with experiments that include missing observations.

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 tests, destroys 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 were prolonged 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 and Herzberg, 1979; Ghosh, 1979; Akhtar and Prescott, 1986; Tanco 2013).

Robustness of a design against missing observation is the ability of a design being able to estimate parameters without too much loss of precision in the presence of unavoidable conditions (missing observations). The robustness of designs against missing observation has been studied by several authors, including Hedayat and John (1974), Ghosh *et al.* (1983), Akhtar and Prescott (1986), Ahmad and Gilmour (2010), Ahmad *et al*. (2012), Alrweili *et al*. (2019), and more. Therefore, the goal of this paper is to develop a minimax loss augmented Box-Behnken design robust to two missing observations using a minimax loss criterion.

Akhtar and Prescott (1986) discussed the effects of missing observations in response surface designs and provided a number of robustness criteria. Akram (2002) developed minimaxloss3 designs robust to three missing observations by using minimax loss criteria, and then compared these designs with cuboidal, orthogonal, rotatable, Box and Draper outlier robust designs, spherical, minimaxloss1 and minimaxloss2 designs. Alrweili *et al.* (2019) constructed minimax loss response surface designs which are robust to missing observations. The construction applied minimax loss criteria to the minimal central composite designs of Georgiou *et al.* (2014) and Angelopoulos *et al.* (2014). Rashid *et al.* (2019) investigated the robustness of augmented Box-Behnken designs and augmented fractional Box-Behnken designs to one missing observation using minimax loss criteria. Oladugba *et al.* (2019) looked at "robustness of Space-filling Orthogonal-Array Composite Design (SOACD)", when observations are declared missing. The authors developed a novel design called "*Space-filling Orthogonal-Array Composite Minimax loss design (SOACM)*". The formulated Second-Order Augmented Composite Model (SOACM) design was implemented using Second-Order Augmented Composite Designs (SOACDs), with the aim of determining an optimal axial distance (α) which minimizes maximum loss of efficiency in the cases of missing data. This was achieved through the application of a minimum loss of efficiency criterion. The study compared the performance of SOACDs with other established designs, such as orthogonal array composite designs, centre composite designs, and small composite designs, focusing on their robustness and efficiency under conditions where data points are missing. The results provided insights into how SOACMs and other designs handle such challenges, contributing to the broader understanding of design reliability in experimental settings. The precision of coefficient estimate of regression model and *D*- efficiency value were the basis for their comparisons. SOACMs showed a preferred performance in general. Full information maximum likelihood was also considered in judging the influence of missing observation (MCAR and MAR) on the performances of the designs and the judgement discovered no association betwixt the missing values and variable in the dataset. Rashid *et al.* (2019) demonstrated that Augmented Box-Behnken designs are generally more robust when one observation is lost for 3 to 6 factors, using minimax loss 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.

MATERIALS AND METHODS

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).

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

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

where: $B[3] = n_b = b \times 2^t = 12$and it is the number of Box Behnken points, $F[a]^3 = n_f = 2^k = 8$,is the number of factorial points, $A[\alpha]^3 = n_a = 2(3) = 6$,is the number of axial points,$n_c$is the number of center points and it is replicated three times.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 -1and +1).

Table 2: The design matrix for $k = 3$ factor in Augmented Box-Behnken Design 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
Axial point: $n_\alpha = 6$	**23**	1	0	α	0
	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

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 (1982), Loss of D-efficiency by Ghosh (1979), Minimax loss criteria by Akhtar and Prescott (1986). This research made use of minimax loss criterion proposed by Akhtar and Prescott (1986)

Minimax loss Criterion

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

$$\frac{|X_r{}'X_r|}{|X'X|} = R_j \quad (1)$$

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

$$l_j = 1 - \frac{|X_r{}'X_r|}{|X'X|} \quad (2)$$

Equation (2) can also be rewritten according to (Rashid *et al.* 2019) as

$$l_j = 1 - R_j = h_{jj} \quad (3)$$

The loss $l_j$is a relative measure of efficiency with $0 \le l_j \le 1$. 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 maximum loss over design points (Ahmad and Gilmour, 2010).

Construction of the Proposed Minimax Loss design

The detailed procedure for constructing robust augmented Box-Behnken design for two 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}, 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, etc.

Step 3. Using greedy search, trace theα 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.

Robust Designs

To find the minimax loss designs robust to two 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 $N = 29$. However, 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 centre 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. 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 (Akram, 2002).

RESULTS AND DISCUSSION

Loss in precision 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: l_{bb}, l_{ff}, l_{aa}, …, l_{ac}.

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

Axial point distance (α)	Loss in precision due to missing									
	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
1.75	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

It is observed in Table 3 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 minimax loss 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 minimax loss 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 minimax loss design robust to a pair of missing observations (minimax loss2).

The minimax loss 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. 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 3. 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.

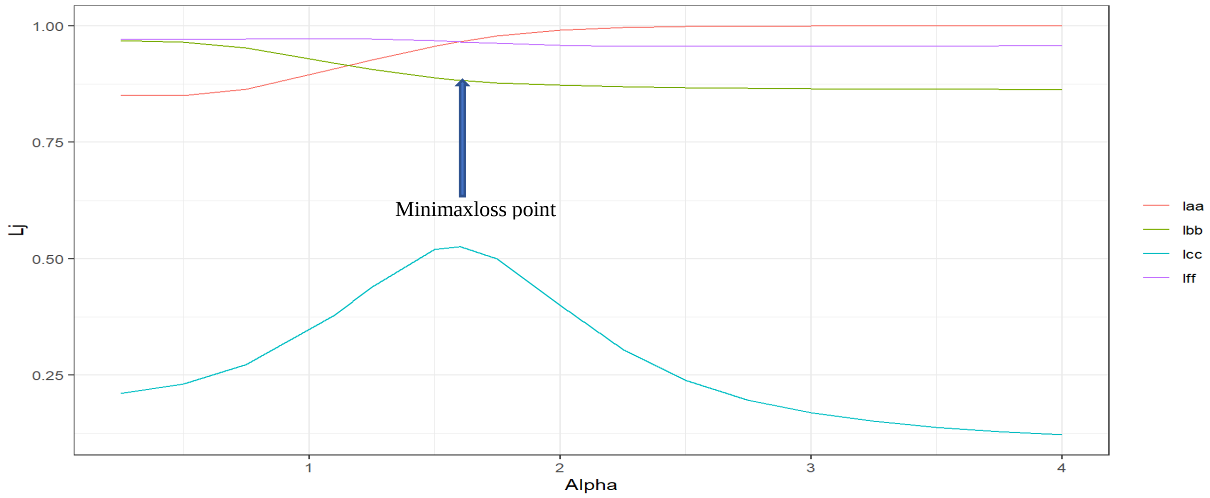

Figure 1: Loss due to two Box-Behnken, two factorial, two axial and two center points missing for $k = 3$ and $n_c = 3$.

CONCLUSION

In response surface modeling, robustness to missing observations is a property that 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 minimax loss criteria. It was shown that for $k = 3$, $n_b = 12$, $n_f = 8$, $n_a = 6$, and $n_c = 3$, with $\alpha = 1.6021$ is minimax loss augmented Box-Behnken design robust to two missing observations (minimax loss2. That is, the axial point distance (α) at the point at which $l_{ff} = l_{aa}$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.

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.

Angelopoulos P, Evangelaras H, and Koukouvinos, C. (2014). Small, balanced, efficient and near-rotatable central composite designs,*Journalof Statistical Planning and Inference*. 139, 2010-2013.

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.

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

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.

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*, 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.

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.

Oladugba, A. and Okeke, E. 2019. Robustness of Space-Filling Orthogonal Array Based Composite Design to Missing. Observation. Abstract Booklet –Royal statistical Society, Belfast, United Kingdom, Sept. 2019. Pp324

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 -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.