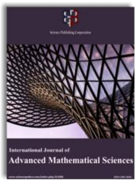
International Journal of Advanced Mathematical Sciences, 5 (2) (2017) 43-49

International Journal of Advanced Mathematical Sciences

Website: www.sciencepubco.com/index.php/IJAMS
doi: 10.14419/ijams.v5i2.7786
Research paper



Missing observations: The loss in relative A-, D- and G-efficiency

Iwundu, M. P. *

Department of Mathematics and Statistics, University of Port Harcourt, Nigeria
**Corresponding author E-mail: mary.iwundu@uniport.edu.ng*

Abstract

The loss in Relative A-, D- and G-efficiency due to missing single or multiple observations is studied using cuboidal designs associated with response models. Higher losses in Relative A- and D-efficiencies are attributed to missing vertex points. The absence of one or two center points does not affect any of Relative A-, D- and G-efficiency, but when its absence is in combination with either a vertex or axial point, there is some negative effect on the design efficiency resulting in some percentage loss in Relative efficiency. The loss in relative efficiency is higher when the missing center point is in combination with missing vertex point. Losses in Relative A- and D-efficiencies are generally higher than losses in Relative G-efficiency. In fact, Relative G-efficiency is mildly affected by the missing vertex or axial point or both.

Keywords: *Missing Observations; Cuboidal Design; Loss in Relative A-Efficiency; Loss in Relative D-Efficiency; Loss in Relative G-Efficiency.*

1. Introduction

Missing data points or observations are practical issues in many experiments as it is clearly not certain that all responses of experimental trials would be realized during experimentation. Encountering missing data points is not far-fetched as instrument malfunction could lead to data point that is not consistent with the other data points thus resulting in outlying observation that is discarded from further consideration in the analysis of the experiment. It could also arise when no value is obtained at one or more factor settings. Since missing data points could grossly affect the statistical power of a test, offer biased estimates of parameters and give invalid conclusions, handling missing data points raises many interests among researchers. Many methods have been proposed to handle missing data and include substitution, imputation, use of robust statistics, etc. It is obvious from literature that attention has been given to missing observations in many areas of statistics, including the area of optimal design of experiments as seen in Herzhberg and Andrews [9], Anbrews and Herzhberg [7], Akhtar and Prescott [4], Akhtar and Prescott [5], Akhtar [2], Whittinghill [14]. Concentrations are generally on the effect of one or more missing observations on certain aspects of optimality. For example, Akhtar [2] considered the effect of one or two missing observations on the determinant of information matrix using the five-factor Box and Behnken design and compared its robustness with the five-factor mimimax loss central composite design. Attention is still rising in the recent years among researchers in the area of optimal design of experiments on aspects of either missing model coefficients or missing observations with particular interests in design robustness to missing model coefficients and design robustness to missing observations. Akhtar [3] considered the case of two missing observations in three different configurations of factorial and axial parts of a five-factor central composite design. Akram [6] assessed the effects of all possible combinations of one to three missing observations of factorial, axial and center points on parameter estimation using the D-efficiency criterion. The setup for the experimental design was multiple replications of center points and single replication of factorial and axial parts of central composite designs. Results showed that the effect of missing a center point is less severe than the effect of missing a factorial point or an axial point. Ahmad [1] considered the construction of different types of second-order response surface designs that are robust to missing observations. Yakubu *et al.* [15] considered the impact of single missing observations of the various composite design points on estimation and predictive capability of central composite designs with four center points and defined for varying axial distances. Largest loss in precision of parameters was associated with a missing factorial point. An extensive study on robust response surface designs against missing observation was made by Srisuradetchai [13] who considered robustness of response surface designs to one missing observation. In particular, the behaviours of D-, A-, G- and IV-efficiencies were studied with a missing response in central composite designs for varying axial distances and center point. Smucker *et al.* [12] considered the robustness of classical and optimal designs to missing observations using D- and I-efficiencies. One to three missing observations were examined to see their impact on D- and I-efficiencies as well as on the leverage and THA criteria. Results showed that optimal designs fared relatively well in terms of robustness when compared to classical designs. Designs that are "more robust" to missing observations have been proposed by da Silva *et al.* [8]. The loss in Relative A-, D- and G-efficiency due to missing single or multiple observations using central composite cuboidal designs shall be the focus of this work. The design variables shall be for two and three cases. The unreplicated factorial and axial parts of the central composite designs together with replicated center points shall be used. Maximum and minimum losses in Relative efficiency shall be presented for the two and three control variable cases of cuboidal designs.

2. Methodology

2.1. Linear model setting

For the purpose of this study, it is assumed that the typical linear regression model in (1) will be fitted:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum\sum_{i<j} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \varepsilon \qquad (1)$$

Written in matrix form as

$$y = X\beta + \varepsilon \qquad (2)$$

where

y is an $N \times 1$ vector of observations.

ε is an $N \times 1$ vectors of random experimental error assumed to have a zero mean and constant variance σ^2.

X is an $N \times p$ expanded design matrix otherwise called model matrix.

β is a $p \times 1$ vectors of unknown parameters estimated as $\hat{\beta} = (X'X)^{-1}X'y$ using the least square approach.

The variance of the estimate is

$$\mathrm{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

$$E(y) = X\hat{\beta}$$

$$V(\hat{y}(\underline{x})) = \sigma^2 \underline{x}(X'X)^{-1}\underline{x}'$$

Letting $f^T(x)$ be a row of the model matrix, X, so that

$$f^T(x)\beta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1}\sum_{j}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 \qquad (3)$$

The procedures employed in studying the effects of missing observations are outlined as follows.

2.2. The central composite cuboidal design

Factorial, axial and center parts make up the design points for a central composite design and each central composite design is defined for a specified axial distance α. According to Myers *et al.* [10], the axial distance lies in the range $1.0 \le \alpha \le \sqrt{k}$, where k is the number of controllable variables. For a two-variable central composite design with $n_c = 4$ replicated center points, the design measure is made up of $N = 12$ design points and listed as

$$\xi_{12} = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The design matrix associated with the 12-point design measure is

$$D = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The corresponding model matrix is

$$X = \begin{pmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The moment matrix is defined as

$$M = \frac{X^T X}{N}$$

$$= \begin{pmatrix} 1.0000 & 0 & 0 & 0 & 0.5000 & 0.5000 \\ 0 & 0.5000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.3333 & 0 & 0 \\ 0.5000 & 0 & 0 & 0 & 0.5000 & 0.3333 \\ 0.5000 & 0 & 0 & 0 & 0.3333 & 0.5000 \end{pmatrix}$$

Its inverse is

$$M^{-1} = \begin{pmatrix} 2.5000 & 0 & 0 & 0 & -1.5000 & -1.5000 \\ 0 & 2.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.0000 & 0 & 0 \\ -1.5000 & 0 & 0 & 0 & 4.5000 & -1.5000 \\ -1.5000 & 0 & 0 & 0 & -1.5000 & 4.5000 \end{pmatrix}$$

For a three-variable central composite design with $n_c = 4$ replicated center points, the design measure is made up of $N = 18$ design points and listed as

$$\xi_{18} = \begin{pmatrix} -1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The design matrix associated with the 18-point design measure is

$$D = \begin{pmatrix} -1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The corresponding model matrix is

$$X = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The moment matrix is defined as

$$M = \frac{X^T X}{N}$$

$$= \begin{pmatrix} 1.0000 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5556 & 0.5556 & 0.5556 \\ 0 & 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4444 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4444 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.4444 & 0 & 0 & 0 \\ 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5556 & 0.4444 & 0.4444 \\ 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4444 & 0.5556 & 0.4444 \\ 0.5556 & 0 & 0 & 0 & 0 & 0 & 0 & 0.4444 & 0.4444 & 0.5556 \end{pmatrix}$$

Its inverse is

$$M^{-1} = \begin{pmatrix} 2.7857 & 0 & 0 & 0 & 0 & 0 & 0 & -1.0714 & -1.0714 & -1.0714 \\ 0 & 1.8000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.8000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.8000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2.2500 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.2500 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.2500 & 0 & 0 & 0 \\ -1.0714 & 0 & 0 & 0 & 0 & 0 & 0 & 6.6429 & -2.3571 & -2.3571 \\ -1.0714 & 0 & 0 & 0 & 0 & 0 & 0 & -2.3571 & 6.6429 & -2.3571 \\ -1.0714 & 0 & 0 & 0 & 0 & 0 & 0 & -2.3571 & -2.3571 & 6.6429 \end{pmatrix}$$

In studying the loss in efficiency due to missing observations, some rows of the design matrix and the model matrix shall be deleted. The deleted rows correspond to the missing observations. If there are m missing observations, exactly m rows in the design matrix *D* and exactly m rows in the model matrix M will be missing. Hence, for m missing observations, the design matrix *D* shall be a matrix of size (N-m) x k and the model matrix shall be a matrix of size (N-m) x p. However, the moment or information matrix shall maintain its dimension irrespective of the missing observations. The design measure associated with the missing observations shall be regarded as the reduced design measure while the design measure associated with the complete observations shall be regarded as the complete design measure. Notationally, D shall describe the design matrix for a design with complete observations. D_r shall describe the design matrix for a design with missing observations. X shall describe the model matrix for a design with complete observations. X_r shall describe the model matrix for a design with missing observations. M shall describe the information matrix for a design with complete observations and M_r shall describe the information matrix for a design with missing observations. M^{-1} and M_r^{-1} shall describe the respective matrix inverses. The relationship between the complete and reduced information matrices may be written as

$$M = M_m + M_r$$

where M_m is associated with the missing observations and equals $X_m{}'X_m$.

Normalization of the information matrix allows comparison of designs with varying design sizes.

2.3. The relative loss in A-efficiency

The concept of Loss in A-efficiency takes its bearing from the criterion of A-optimality, which focuses on good model parameter estimation. It considers making the variances of the parameter estimates small and does not take into account the covariances among model parameters. As in Rady *et al.* [11], A-optimality criterion is one in which the sum of the variances of the model coefficients is minimized. It is defined by the criterion function

$$\Phi\ (M\ (\xi)) = \text{Min tr}\ (M^{-1})$$

where Min implies that minimization is over all designs and tr represents trace.

The A-efficiency of a design ξ is defined as

$$A(\xi) = \frac{trace\ \{(M^{-1}(\xi^*)\}}{trace\ \{(M^{-1}(\xi)\}}$$

where ξ^* is A-optimal.

When comparing two designs ξ_N and ξ_{N-m} for a p-parameter model, the A-efficiency of ξ_{N-m} relative to ξ_N is the ratio of the separate A-efficiencies and is given as

$$A_{Rel}(\xi_{N-m}/\xi_N) = \frac{trace\ \{(M^{-1}(\xi^*)\}}{trace\ \{(M^{-1}(\xi_{N-m})\}} \Big/ \frac{trace\ \{(M^{-1}(\xi^*)\}}{trace\ \{(M^{-1}(\xi_N)\}}$$

$$= \frac{trace\ \{(M^{-1}(\xi_N)\}}{trace\ \{(M^{-1}(\xi_{N-m})\}}$$

$$= \frac{trace\ M^{-1}}{trace\ M_r^{-1}}$$

where, as earlier stated, M is the information matrix for a design with complete observations and M_r is the information matrix for a design with missing observations. In comparing designs, the best design is one with the largest A-efficiency value. The relative A-efficiency of a design is such that

$$0 \leq A_{Rel}(\xi_{N-m}/\xi_N) \leq 1,$$

However, $A_{Rel}(\xi_{N-m}/\xi_N)$ may be negative if the design ξ_{N-m} is better than the design ξ_N.

The loss in relative A-efficiency is thus

$$A_{R.l.e} = 1 - A_{Rel}(\xi_{N-m}/\xi_N)$$

$= 1 - \frac{trace\ M^{-1}}{trace\ M_r^{-1}}$

3.4. The relative loss in D-efficiency

The concept of Loss in D-efficiency takes its bearing from the criterion of D-optimality, which also focuses on good model parameter estimation. As in Rady *et al.* [11], D-optimality criterion considers making both the variances of the parameter estimates and the covariances among model parameter estimates small. It is one in which the determinant of the moment matrix

$$M = \frac{X'X}{N}$$

is maximized over all designs, where X represents the design matrix associated with the design and X' represents its transpose. Equivalently, a design is D-optimum if it minimizes the generalized variance of the parameter estimates. It is defined by the criterion function

$$\Phi\ (M\ (\xi)) = \max\ \{\det M\ (\xi)\}$$

$$= \min\ \{\det\ (M^{-1}(\xi))\}$$

where det (.) represents determinant.
The D-efficiency of a design ξ is given as

$$D(\xi) = (det\ M(\xi))^{\frac{1}{p}}$$

When comparing two designs ξ_N and ξ_{N-m} for a p-parameter model, the D-efficiency of ξ_{N-m} relative to ξ_N is the ratio of the separate D-efficiencies and is given as

$$D_{Rel}(\xi_{N-m}/\xi_N) = \left(\frac{det\ M(\xi_{N-m})}{det\ M(\xi_N)}\right)^{\frac{1}{p}}$$

$$= \left(\frac{det\ M_r}{det\ M}\right)^{\frac{1}{p}}$$

where, as earlier stated, M is the information matrix for a design with complete observations and M_r is the information matrix for a design with missing observations. In comparing designs, the best design is one with the largest D-efficiency value. The relative D-efficiency of a design is such that

$$0 \le D_{Rel}(\xi_{N-m}/\xi_N) \le 1$$

However, $D_{Rel}(\xi_{N-m}/\xi_N)$ may be negative if the design ξ_{N-m} is better than the design ξ_N.
The loss in relative D-efficiency is thus

$$D_{R.l.e} = 1 - D_{Rel}(\xi_{N-m}/\xi_N)$$

$$= 1 - \left(\frac{det\ M_r}{det\ M}\right)^{\frac{1}{p}}$$

3.5. The relative loss in G-efficiency

The concept of Loss in G-efficiency takes its bearing from the criterion of G-optimality. As in Rady *et al.* [11], G-optimality criterion considers designs whose maximum scaled prediction variance, $v(\underline{x})$, in the region of the design is not too large. Hence, a G-optimal design minimizes the maximum scaled prediction variance and is defined by the criterion function

$$\Phi\ (M\ (\xi)) = \text{Min}\ \{\max_{\underline{x}\in R} v(\underline{x})\}$$

The G-efficiency of a design is defined as

$$\frac{p}{V(\underline{x})_{max}}$$

where p is the number of parameters in the model and $V(\underline{x})_{max}$ is the maximum scaled variance of prediction. G-efficiency compares the maximum value of scaled variance of prediction within the design region with respect to its theoretical minimum variance p.
When comparing two designs ξ_N and ξ_{N-m} for a p-parameter model, the G-efficiency of ξ_{N-m} relative to ξ_N is the ratio of the separate G-efficiencies and is given as

$$G_{Rel}(\xi_{N-m}/\xi_N) = \frac{p}{V(\underline{x})_{max(\xi_{N-m})}} \Big/ \frac{p}{V(\underline{x})_{max(\xi_N)}}$$

$$= \frac{V(\underline{x})_{max(\xi_N)}}{V(\underline{x})_{max(\xi_{N-m})}}$$

where $V(\underline{x})_{max(\xi_N)}$ is associated with the design with complete observations and $V(\underline{x})_{max(\xi_{N-m})}$ is associated with the design with missing observations. In comparing designs, the best design is one with the largest G-efficiency value. The relative G-efficiency of a design is such that

$$0 \le G_{Rel}(\xi_{N-m}/\xi_N) \le 1,$$

However, $G_{Rel}(\xi_{N-m}/\xi_N)$ may be negative if the design ξ_{N-m} is better than the design ξ_N.
The loss in relative G-efficiency is thus

$$G_{R.l.e} = 1 - G_{Rel}(\xi_{N-m}/\xi_N)$$

$$= 1 - \frac{V(\underline{x})_{max(\xi_N)}}{V(\underline{x})_{max(\xi_{N-m})}}$$

3. Results and discussion

Two categories of missing observations, namely a single missing observation and a pair of missing observations, are studied throughout this work. Single missing observation includes

i) Missing a vertex point (V_1).
ii) Missing an axial or star point (S_1).
iii) Missing a center point (C_1).

On the other hand, missing a pair of observations includes

i) Missing two vertex points (V_2).
ii) Missing two axial or star points (S_2).
iii) Missing two center points (C_2).
iv) Missing a vertex point and a center point (V_1C_1).
v) Missing an axial or star point and a center point (S_1C_1).
vi) Missing a vertex point and an axial or star point (V_1S_1).

3.1. One missing observation in vertex, axial and center parts of cuboidal design in two control variables

In studying the robustness of A-, D- and G-efficiency to missing points of vertex, axial and center parts of the cuboidal design in two control variables with four replicated center points, no loss is incurred by a missing center point even though losses are incurred by missing vertex or axial point. Higher losses in Relative A- and D-efficiencies are attributed to missing vertex point. Relative G-efficiency is mildly affected by missing vertex or axial point with maximum loss of 5.78% attributed to missing an axial point. Details of the results for each subcategory are as in Table 1.

3.2. One missing observation in vertex, axial and center parts of cuboidal design in three control variables

As with the two control variable case in section 3.1, there is no loss in Relative A-, D- and G-efficiencies to missing a center point

Table 1: Relative Efficiencies and Losses to A Single Missing Observation at Factorial, Axial and Center Points of Cuboidal Design with k = 2

Criterion	One missing vertex point V_1	One missing axial point S_1	One missing center point C_1	No missing point
Design size N	11	11	11	12
Determinant of normalized information matrix	0.0016	0.0039	0.0062	0.0046
Trace of the inverse of information matrix	24.9333	20.4722	17.9956	18.50
Maximum scaled predictive variance	9.5333	10.0833	8.7325	9.5
Minimum scaled predictive variance	2.3833	2.4444	2.8947	2.5
Relative D-efficiency	0.8386	0.9729	1.0510	-
Relative G-efficiency	0.9965	0.9422	1.0879	-
Relative A-efficiency	0.7420	0.9037	1.0280	-
Loss in Relative D-efficiency	0.1614	0.0271	-	-
Loss in Relative G-efficiency	0.0035	0.0578	-	-
Loss in Relative A-efficiency	0.2580	0.0963	-	-

of the central composite design in three control variables with four replicated center points. However, losses are incurred by missing a vertex or an axial point with higher losses in A- , D- and G-efficiencies being attributed to missing a vertex point. Details of the results for each subcategory are as in Table 2.

3.3. Two missing observation in vertex, axial and center parts of cuboidal design in two control variables

In studying the robustness of A-, D- and G-efficiency to a pair of missing observations at vertex point, axial point, center point as well as a combinations of two categories of points, associated with the cuboidal design in two control variables with four replicated center points, no loss is incurred by missing two center points. Three categories, namely V_2, $S_2$and V_1S_1, incurred multiple losses in efficiency. Maximum losses in A- and D-efficiency are attributed to missing two vertex points (V_2) followed by missing one vertex point and one axial point (V_1S_1). G-efficiency is mildly affected by missing points with a maximum of 5% loss in efficiency associated with missing two vertex points (V_2) and missing one vertex point and one axial point (V_1S_1). Details of the results for each subcategory are as in Table 3.

Table 2: Relative Efficiencies and Losses to A Single Missing Observation at Factorial, Axial and Center Points of Cuboidal Design with k = 3

Criterion	One missing vertex point V_1	One missing axial point S_1	One missing center point C_1	No missing point
Design size N	17	17	17	18
Determinant of normalized information matrix	3.5147x10^{-5}	8.4521x10^{-5}	1.4425x10^{-4}	9.6364x10^{-5}
Trace of the inverse of information matrix	39.8812	39.3615	33.6229	34.8543
Maximum scaled predictive variance	16.6069	14.0250	13.5102	14.2929
Minimum scaled predictive variance	2.6777	2.9423	3.1127	2.7857
Relative D-efficiency	0.9041	0.9870	1.0412	-
Relative G-efficiency	0.8607	1.0191	1.0579	-
Relative A-efficiency	0.8742	0.8857	1.0369	-
Loss in Relative D-efficiency	0.0959	0.0130	-	-
Loss in Relative G-efficiency	0.1393	-	-	-
Loss in Relative A-efficiency	0.1258	0.1143	-	-

Table 3: Relative Efficiencies and Losses to A Pair of Missing Observations at Factorial, Axial, Center and the Point Combinations of Cuboidal Design with k = 2

Criterion	Two missing vertex points V_2	Two missing axial points S_2	Two missing center points C_2	One missing vertex point and one missing axial point V_1S_1	One missing vertex point and one missing center point V_1C_1	One missing axial point and one missing center point S_1C_1	No missing point
Design size N	10	10	10	10	10	10	12
Determinant of normalized information matrix	5.76x10^{-4} 3.84 x10^{-4}	0.0031 0.0031	0.0081	0.0013 5.76 x10^{-4}	0.0023	0.0054	0.0046
Trace of the inverse of information matrix	31.1111 41.6667	24.1667 24.3750	17.9762	27.1429 44.4444	23.9007	19.7024	18.50
Maximum scaled predictive variance	10.0 8.3333	9.1667 9.7917	7.9762	9.5238 10.0000	8.7234	9.1667	9.5
Minimum scaled predictive variance	2.2222 2.5000	2.50 2.2917	3.5714	2.3810 2.2222	2.7660	2.8571	2.5
Relative D-efficiency	0.8386 0.6611	0.9363 0.9363	1.0989	0.8101 0.7073	0.8909	1.0271	-
Relative G-efficiency	0.95 1.1400	1.0364 0.9702	1.1910	0.9975 0.9500	1.0890	1.0364	-
Relative A-efficiency	0.5946 0.4440	0.7655 0.7590	1.0291	0.6816 0.9500	0.7740	0.9390	-
Loss in Relative D-efficiency	0.1614 0.3389	0.0637 0.0637	-	0.1899 0.2927	0.1091	-	-
Loss in Relative G-efficiency	0.05 -	- 0.0298	-	0.0025 0.0500	-	-	-
Loss in Relative A-efficiency	0.4054 0.5560	0.2345 0.2410	-	0.3184 0.5837	0.2260	0.0610	-

3.4. Two missing observation in vertex, axial and center parts of cuboidal design in three control variables

As with the two control variable case in section 3.3, there is no loss in A-, D- and G-efficiencies to missing of a pair of center points of the cuboidal design in three control variables with four replicated center points. Three categories, namely V_2, $S_2$and V_1S_1, incurred multiple losses in efficiency. Maximum loss in D-efficiency is attributed to missing two vertex points (V_2) followed by missing one vertex point and one axial point (V_1S_1). Maximum loss in A-efficiency is attributed to missing two vertex points (V_2) followed by missing two axial points (S_2) and thirdly, missing one vertex point and one axial point (V_1S_1). G-efficiency is again mildly affected by missing points with a maximum of approximately 9% loss in efficiency across the subcategories V_2, V_1S_1 and V_1C_1. Details of the results for each subcategory are as in Table 4.

Table 4: Relative Efficiencies and Losses to A Pair of Missing Observations at Factorial, Axial, Center and the Point Combinations of Cuboidal Design with k = 3

Criterion	Two missing vertex points V_2	Two missing axial points S_2	Two missing center points C_2	One missing vertex point and one missing axial point V_1S_1	One missing vertex point and one missing center point V_1C_1	One missing axial point and one missing center point S_1C_1	No missing point
Design size N	16	16	16	16	16	16	18
Determinant of normalized information matrix	$1.0505x10^{-5}$ $9.7603x10^{-6}$ $1.4909x10^{-6}$	$4.7684x10^{-5}$ $7.3910x10^{-5}$	$2.1607x10^{-4}$	$3.01x10^{-5}$ $2.712x10^{-5}$	$5.4296x10^{-5}$	$1.2815x10^{-4}$	$9.6364x10^{-5}$
Trace of the inverse of information matrix	47.6851 50.7756 179.20	61.20 44.5032	32.5931	44.7485 46.0176	38.2669	37.9907	34.8643
Maximum scaled predictive variance	15.6595 15.6336 13.60	13.20 13.5355	12.7310	15.7228 15.6923	15.6707	13.20	14.2929
Minimum scaled predictive variance	2.6383 2.5344 4.00	4.0 2.9677	3.5862	2.8911 2.7692	2.9914	3.3488	2.7857
Relative D-efficiency	0.8012 0.7953 0.6591	0.9321 0.9738	1.0841	0.8902 0.8809	0.9442	1.0289	-
Relative G-efficiency	0.9127 0.9142 1.0509	1.0828 1.0560	1.1227	0.9091 0.9108	0.9121	1.0828	-
Relative A-efficiency	0.7311 0.6866 0.1946	0.5697 0.7834	1.0697	0.7791 0.7576	0.9111	0.9177	-
Loss in Relative D-efficiency	0.1988 0.2047 0.3409	0.0679 0.0262	-	0.1098 0.1191	0.0558	-	-
Loss in Relative G-efficiency	0.0873 0.0858 -	- -	-	0.0909 0.0892	0.0879	-	-
Loss in Relative A-efficiency	0.2689 0.3134 0.8054	0.4303 0.2166	-	0.2209 0.2424	0.0889	0.0823	-

4. Conclusion

This work has assessed the effect of one or two missing observations of factorial, axial and center points on three optimality measures. It has particularly examined losses incurred in A-, D- and G-efficiency by missing a single observation of the factorial, axial or center part of cuboidal designs as well as losses incurred in A-, D- and G-efficiency by a pair of missing observations. Although the absence of one or two center points was not seen to affect any of A-, D- and G-efficiency, but when its absence is in combination with either a vertex or axial point, there is some negative effect on the design efficiency resulting in some percentage loss in efficiency. The loss in efficiency is higher when the missing center point is in combination with missing vertex point. Relative efficiencies that exceed 1.0 simply imply that there is no loss in efficiency using the specified criterion but rather some percentage gain in efficiency has been incurred. By interpretation, such situation implies that using a design with the missing observation(s) is better than using the specified "complete" design. For instance, the relative D-efficiency of 1.0841 associated with a pair of missing observations at center part of the cuboidal design with three control variables implies that by the two missing points there is 8.41% gain in D-efficiency when compared with the 18 point cuboidal design having four center points. This result agrees with the result of Srisuradetchai [13] that in general, increasing the number of center runs does not much improve the robustness of central composite design with α = 1.0 but rather results in lower D-efficiency value. Unlike missing vertex points which seems to grossly affect design efficiency, missing one or two center runs do not incur losses in design efficiency for both two- and three-control variable cases and missing one center point plus one axial point incurs minimal or no loss in efficiency. For missing single observation associated with two-control variable case, higher losses in A- and D-efficiency are attributed to missing vertex point, the category V_1. On the other hand, higher losses in G-efficiency are attributed to missing axial point, the category S_1. For missing two observations, higher losses in D-efficiency are attributed to missing points in the category V_1S_1. On the other hand, higher losses in A- and G-efficiency are attributed to missing points in the category V_2. For missing single observation associated with three control variable case, higher losses in A- D- and G-efficiency are attributed to missing vertex point, the category V_1. For missing two observations, higher losses in A- and D-efficiency are attributed to missing vertex point, the category V_2 while higher losses in G-efficiency are attributed to missing points in the category V_1S_1 even though percentage losses are approximately identical across the subcategories V_2, V_1S_1 and S_1C_1. Losses in A- and D-efficiencies are generally higher than losses in G-efficiency. Summarily, G-efficiency is mildly affected by the missing observations. Unlike what is seen in most references cited, all information matrices used in the analysis have been normalized to remove the effect of changing design sizes.

References

[1] T. Ahmad, Some new Second-order Response Surface Designs robust to Missing Observations,Thesis in Statistics, TheIslamia University of Bahawalpur,Thesis Abstractat Pakistan Journal of Statistics and Operations Research, (2011).

[2] M. Akhtar, One or two missing observations in five factor Box and Behnken Design,Journal of Engg.& App. Scs.,6(1) (1987) 87-89.

[3] M. Akhtar, Five-factor Central Composite Designs robust to a pair of missing Observations, Journal of Research (Science), Bahauddin Zakariya University, Multan, Pakistan, 12 (2) (2001) 105-115.

[4] M. Akhtar, P. Prescott, Response Surface Designs Robust to missing observations, Communications in Statistics- Simulation, 15 (2) (1986) 245-363.

[5] M. Akhtar, P. Prescott, A review of Response Surface Designs, Pak. Journal of Statistics, 3 (3) B (1987) 11-26.

[6] M. Akram, Central Composite Designs Robust to three missing observations, A Ph. D Thesis in Statistics, Islamia University, Bahawalpur, (2002).

[7] D. F. Andrews, A. M. Herzhberg, The Robustness and optimality of Response Surface Designs, Journal of Statistical Planning and Inference, 3 (1979) 249-257. https://doi.org/10.1016/0378-3758(79)90016-8.

[8] M. A. da Silva, S. G. Gilmour, L. A.Trinca, Factorial and Response Surface Designs robust to missing observation, Computational Statistics and Data Analysis,Available online on 7 June 2016.

[9] A. M. Herzhberg, D. F. Andrews, Some considerations in the optimal design of experiments in Non-optimal situation, Journal Roy. Statist. Soc. B38 (1976) 284-289.

[10] R. H., Myers, D. C. Montgomery, C. M. Anderson-Cook, Response Surface Methodology: Process and Product Optimization using designed experiments, 3rd Edition, John Wiley & Sons, Inc. New Jersey, (2009).

[11] E. A. Rady, M. M. E. Abd El-Monsef, M. M. Seyam, Relationship among several optimality criteria, interstat.statjournals.net> YEAR> articles (2009).

[12] B. J. Smucker, W. Jensen, Z. Wu, B. Wang, Robustness of Classical andOptimal Designs to Missing Observations, Computational Statistics & Data Analysis, Available online January 2017, 17 pages, http://doi.org/10.1016/j.csda.2016.12.001.

[13] P. Srisuradetchai, Robust response surface designs against missing observations. A Ph. D dissertation, Montana State University, Bozeman, Montana, (2015).

[14] D. C. Whittinghill, A note on the Robustness of Box-Behnkens Designs to unavailability of Data, Metrika, 48 (1998) 49-52. https://doi.org/10.1007/PL00003971.

[15] Y.Yakubu, A. U.Chukwu, B. T. Adebayo, A. G. Nwanzo, Effects of missing observations on predictive capability of central composite designs, International Journal on Computational Sciences & Applications, 4 (6) (2014) 1-18. https://doi.org/10.5121/ijcsa.2014.4601.