Model Assisted Statistics and Applications 6 (2011) 81–87
DOI 10.3233/MAS-2011-0176
IOS Press

Robustness of neighbour balanced complete block designs against missing observation(s)

N.R. Abeynayake[a], Seema Jaggi[b,*] and Cini Varghese[b]
[a]*Department of Agribusiness Managemant, Wayamba University, Gonawila* (*NWP*), *Sri Lanka*
[b]*Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi, India*

Abstract. A block design with neighbour effect(s) is said to be neighbour balanced if every treatment has every other treatment appearing constant number of times as neighbour(s). These designs are used when the treatment applied to one experimental plot may affect the response on neighbouring plot(s) besides the response to which it is applied. Neighbour-balanced designs ensure that no treatment is unduly disadvantaged by its neighbour(s). However, there is a possibility that some of the observations could become unavailable for analysis. In this paper, we examine the robustness of neighbour balanced complete block designs when specific observations are missing. The information matrix for direct treatment effects of the resultant design (one-sided neighbour effects) after missing of an observation from a block is derived and the efficiency of resulting design is investigated. Robustness of neighbour balanced complete block design has also been investigated against missing of more than one observation. The efficiencies are found to be quite high indicating the designs to be fairly robust against missing observations.

Keywords: Neighbour balanced block design, missing observations, information matrix, eigenvalues, direct and neighbour effects

1. Introduction

Experiments conducted in agriculture often show neighbour effects i.e., the response on a given plot is affected by the treatments on the neighbouring plots as well as by the treatment applied to that particular plot. Neighbour-balanced designs, wherein the allocation of treatments is such that every treatment occurs equally often with every other treatment as neighbours, are used under these situations. Neighbour-balanced designs ensure that no treatment is unduly disadvantaged by its neighbours. These designs permit the estimation of neighbour effects besides the direct effects of treatments. Understanding the structure of the neighbour effects helps in minimizing the bias in direct treatment effects to great extent.

Series of circular neighbour balanced block (NBB) designs have been developed in the literature. Azais et al. [5] have given a catalogue of complete and incomplete NBB designs. Tomar et al. [6] have developed series of incomplete NBB designs and Jaggi et al. [15] have obtained series of partial NBB designs. Jaggi et al. [14] have studied the optimal properties of complete block design with neighbouring competition effects.

In a well-planned experimental work, situation may arise where some observations are lost or destroyed or unavailable due to certain reasons that are beyond the control of the experimenter. Unavailability of the observations destroys the orthogonality and the balance of the design and also affects the inference.

In the literature, robustness of designs has been studied by many research workers with different angles. The common factors which generally disturb the structure of efficient/optimal designs are missing data [e.g. missing observation(s), missing treatment(s), missing block(s) etc.]; presence of outlier(s); presence of a common trend effect in one or more spatial dimension; inadequacy of assumed model (e.g. correlated error structure, inequality of error variances, incomplete model, etc.); exchange or interchange of treatments. Among all these disturbances, the commonest factor responsible for disturbances is the missing observation(s).

*Corresponding author. E-mail: seema@iasri.res.in.

ISSN 1574-1699/11/$27.50

The robustness of several kinds of block designs against the unavailability of data has been investigated in abundance, for example, see Hedayat and John [3], Ghosh [8–11], Ghosh et al. [13], Srivastava et al. [7], Bhaumik and Whittinghill [4], Ghosh et al. [12], Das and Kageyama [1] and Dey [2]. Gupta and Srivastava [16] investigated the robustness of block design against the unavailability of some disjoint blocks. As a special case, they also showed that resolvable balanced incomplete block (BIB) designs are fairly robust against the unavailability of one resolution set consisting of disjoint blocks.

Ghosh [8] introduced connectedness criterion of robustness, in connection with BIB designs. The criterion considers a connected design *d*, i.e., a design in which all the elementary treatment contrasts are estimable, and the residual design $(d*)$, the design which has actually remained after some disturbance, say η, has occurred. The design d is said to be robust against the disturbance η, if the design d* is connected.

Although the design d may be robust in the sense of connectedness, the residual design d* may not be efficient as compared to the original design. Hence, efficiency criterion is also of much importance. According to this criterion, a design is robust against the loss of observations, if the efficiency of the residual design is close to the efficiency of the original design. If $\mathbf{C}_d$ is the information matrix of the original connected design d and is $\mathbf{C}_{d^*}$ that of the residual design d*, then the efficiency E of the residual design relative to the original design is given by

$$E = \frac{\text{Harmonic mean of non-zero eigen values of } \mathbf{C}_{d*}}{\text{Harmonic mean of non-zero eigen values of } \mathbf{C}_d} \tag{1}$$

The purpose of this paper is to assess the consequences of missing observations from blocks of NBB design. In particular, we investigate the robustness of complete NBB designs under one-sided neighbour effects model against missing of an observation. Further, the efficiency of complete NBB designs has also been studied with more than one observation missing.

2. Model

Consider v number of treatments to be studied in b blocks with n experimental units under the following additive fixed effect model with one-sided (say, left side) neighbour effect:

$$\mathbf{Y} = \mu\mathbf{1} + \Delta'\tau + \Delta'_1\delta + \mathbf{D}'\beta + \mathbf{e},$$

where, **Y** is $n \times 1$ vector of observations, μ is grand mean, **1** is $n \times 1$ vector of unities, Δ' is n $\times$ v incidence matrix of observations versus direct treatments, τ is $v\times$ 1 vector of direct treatment effects, Δ'_1 is $n \times v$ incidence matrix of observations versus left treatments, δ is $v \times 1$ vector of left neighbour effects, **D'** is $n \times b$ incidence matrix of observations versus blocks, β is $b \times 1$ vector of block effects and **e** is $n \times 1$ vector of errors.

Further let,

$\Delta\Delta'_1 = \mathbf{M}$, $v \times v$ incidence matrix of direct versus left neighbour treatments

$\Delta\mathbf{D}' = \mathbf{N}_1$, $v \times b$ incidence matrix of direct treatments versus blocks

$\Delta_1\mathbf{D}' = \mathbf{N}_2$, $v \times b$ incidence matrix of left neighbour treatments versus blocks and $\mathbf{r} = (r_1, r_2, \ldots, r_v)$' be the $v\times$ 1 replication vector of direct treatments with r_s ($s =$ 1, 2, ..., v) being the number of times the s^{th} treatment appears in the design.

$\mathbf{r}_1 = (r_{11}, r_{12}, \ldots, r_{1v})$ be the v $\times$ 1 replication vector of the left neighbour treatments with r_{1s} being the number of times the treatments in the design has s^{th} treatment as left neighbour.

$\mathbf{R}_\tau =$ diag $(r_1, r_2, \ldots, r_v) =$ diagonal matrix of replication of direct treatments

$\mathbf{R}_\delta =$ diag $(r_{11}, r_{12}, \ldots, r_{1v})$

$\mathbf{K} =$ diag $(k_1, k_2, \ldots, k_b) =$ diagonal matrix of block sizes, where $k_1, k_2, \ldots, k_b$ are the sizes of the b blocks.

The joint information matrix (**C**) for estimating direct and left neighbour effects is as follows:

$$\mathbf{C} = \begin{bmatrix} \mathbf{R}_\tau - \mathbf{N}_1\mathbf{K}^{-1}\mathbf{N'}_1 & \mathbf{M} - \mathbf{N}_1\mathbf{K}^{-1}\mathbf{N'}_2 \\ \mathbf{M}' - \mathbf{N}_2\mathbf{K}^{-1}\mathbf{N'}_1 & \mathbf{R}_\delta - \mathbf{N}_2\mathbf{K}^{-1}\mathbf{N'}_2 \end{bmatrix} \tag{2}$$

From above, the information matrix for estimating the direct effect of treatments is obtained as given below:

$$\mathbf{C}_\tau = \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-}\mathbf{C}_{21} \tag{3}$$

with $\mathbf{C}_{11} = [\mathbf{R}_\tau - \mathbf{N}_1\mathbf{K}^{-1}\mathbf{N'}_1]$; $\mathbf{C}_{12} = [\mathbf{M} - \mathbf{N}_1\mathbf{K}^{-1}\mathbf{N'}_2]$ and $\mathbf{C}_{22} = [\mathbf{R}_\delta - \mathbf{N}_2\mathbf{K}^{-1}\mathbf{N'}_2]$.

The matrix $\mathbf{C}_\tau$ is symmetric, non-negative definite with zero row and column sums and Rank ($\mathbf{C}_\tau$) $\leqslant$ (v–1). Similarly, the information matrix for estimating left neighbour effects ($\mathbf{C}_\delta$) can be obtained.

A block design with one-sided neighbour effects is said to be **neighbour balanced** if every treatment has every other treatment appearing constant (say λ) number of times as a left neighbour. Similarly, under two-sided neighbour effects model, a block design is neighbour balanced if every treatment has every other treatment as left and right neighbour constant number of times. These designs are circular in the sense that treatment in the left border is the same as the treatment in the right end inner plot and the treatment in the right border is same as the treatment in the left end inner plot. It may be mentioned here that the observations are not recorded from the border plots; these plots are taken only to have the neighbour effects of treatments at the end plots of the blocks.

3. Robustness of one-sided neighbour balanced block designs

We consider here the class of complete NBB design with v treatments (v prime) in $b = v - 1$ blocks, $r_1 = r_2 = \ldots = r_v = v - 1$, $k_1 = k_2 = \ldots = k_b = v$ and $\lambda = 1$ (Azais et al. [5]) obtained by taking the j^{th} block ($j = 1, 2, \ldots, v - 1$) of the design as:

$v \;\; j \;\; 2j \ldots (v-1)j$ modulo v

The structure of various matrices for this class is as follows:

$$\mathbf{M} = \mathbf{J} - \mathbf{I}, \quad \mathbf{N}_1 = \mathbf{N}_2 = \mathbf{J}, \quad \mathbf{R}_\tau = \mathbf{R}_\delta = (v-1)\mathbf{I} \text{ and } \mathbf{K} = v\mathbf{I}$$

J is the $v \times v$ matrix of unities and **I** is an identity matrix of order v.

The joint information matrix **C** as given in Eq. (2) reduces here to

$$\mathbf{C} = \begin{bmatrix} (v-1)\left[\mathbf{I} - \frac{\mathbf{J}}{v}\right] & \frac{\mathbf{J}}{v} - \mathbf{I} \\ \frac{\mathbf{J}}{v} - \mathbf{I} & (v-1)\left[\mathbf{I} - \frac{\mathbf{J}}{v}\right] \end{bmatrix}$$

with,

$$\mathbf{C}_{11} = (v-1)\left[\mathbf{I} - \frac{\mathbf{J}}{\mathbf{v}}\right], \mathbf{C}_{12} = \left[\frac{\mathbf{J}}{v} - \mathbf{I}\right], \mathbf{C}_{22} = \left[(v-1)\left[\mathbf{I} - \frac{\mathbf{J}}{v}\right]\right] \text{ and } \mathbf{C}_{22}^{-} = \frac{1}{(v-1)}\mathbf{I}.$$

Therefore, the information matrix for estimating the direct effects and left neighbour effects of treatment is

$$\mathbf{C}_\tau = \mathbf{C}_\delta = \frac{v(v-2)}{(v-1)}\left[\mathbf{I} - \frac{\mathbf{J}}{v}\right]$$

Example 1: Following is a circular complete NBB design with parameters $v = 7$, $b = 6$, $r = 6$, $k = 7$ and $\lambda = 1$ balanced for left neighbour:

6	7	1	2	3	4	5	6
5	7	2	4	6	1	3	5
4	7	3	6	2	5	1	4
3	7	4	1	5	2	6	3
2	7	5	3	1	6	4	2
1	7	6	5	4	3	2	1

Now, let us assume that the observation pertaining to right most plot of a block containing the direct effect of any treatment and respective left neighbour effect is missing. This is feasible since in experimental layout in the field, edge of blocks is more vulnerable for physical damages.

For mathematical simplification, it is assumed that the observation pertaining to right most plot of the last block containing the direct effect of treatment number v and respective left neighbour effect is missing. Since the design is circular, without loss of generality the contents of the last block can be rearranged in such a way that right most plot of last block have treatment number v.

After missing of last observation from last block, the various incidence matrices of the residual design (denoted with *) are changed as follows:

$$\mathbf{M}^* = \mathbf{M} - \begin{bmatrix} \mathbf{0}_{v-1} & \mathbf{0}_{v-1} \\ 1 & \mathbf{0'}_{v-1} \end{bmatrix},$$

$$\mathbf{N}_1^* = \mathbf{N_1} - \begin{bmatrix} \mathbf{0}_{v-1,v-2} & \mathbf{0}_{v-1} \\ \mathbf{0'}_{v-2} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{J}_{v-1,v-2} & \mathbf{1}_{v-1} \\ \mathbf{1'}_{v-2} & 0 \end{bmatrix},$$

$$\mathbf{N}_2^* = \mathbf{N}_2 - \begin{bmatrix} \mathbf{0'}_{v-2} & 1 \\ \mathbf{0}_{v-1,v-2} & \mathbf{0}_{v-1} \end{bmatrix} = \begin{bmatrix} \mathbf{1'}_{v-2} & 0 \\ \mathbf{J}_{v-1,v-2} & \mathbf{1}_{v-1} \end{bmatrix},$$

$$\mathbf{R}_\tau^* = \mathbf{R}_\tau - \begin{bmatrix} \mathbf{0}_{v-1} & \mathbf{0}_{v-1} \\ \mathbf{0'}_{v-1} & 1 \end{bmatrix} = \begin{bmatrix} (v-1)\mathbf{I}_{v-1} & \mathbf{0}_{v-1} \\ \mathbf{0'}_{v-1} & v-2 \end{bmatrix},$$

$$\mathbf{K}^* = \mathbf{K} - \begin{bmatrix} \mathbf{0}_{v-2} & \mathbf{0}_{v-2} \\ \mathbf{0'}_{v-2} & 1 \end{bmatrix} = \begin{bmatrix} v\mathbf{I}_{v-2} & \mathbf{0}_{v-2} \\ \mathbf{0'}_{v-2} & v-1 \end{bmatrix}.$$

The joint information matrix of the residual design can be expressed as follows:

$$\mathbf{C}^* = \begin{bmatrix} \mathbf{R}_\tau^* - \mathbf{N}_1^*\mathbf{K}^{*-1}\mathbf{N}_1^{*'} & \mathbf{M}^* - \mathbf{N}_1^*\mathbf{K}^{*-1}\mathbf{N}_2^{*'} \\ \mathbf{M}^* - \mathbf{N}_2^*\mathbf{K}^{*-1}\mathbf{N}_1^{*'} & \mathbf{R}_\delta^* - \mathbf{N}_2^*\mathbf{K}^{*-1}\mathbf{N}_2^{*'} \end{bmatrix}$$

with $\mathbf{C}_\tau^* = \mathbf{C}_{11}^* - \mathbf{C}_{12}^*\mathbf{C}_{22}^{*-}\mathbf{C}_{21}^*$.

Here,

$$\mathbf{C}_{11}^* = \mathbf{R}_\tau^* - \mathbf{N}_1^*\mathbf{K}^{*-1}\mathbf{N}_1^{*'} = \begin{bmatrix} (v-1)\mathbf{I}_{v-1} - \frac{(v^2-2v+2)}{v(v-1)}\mathbf{J}_{v-1} & -\frac{(v-2)}{v}\mathbf{1}_{v-1} \\ -\frac{(v-2)}{v}\mathbf{1'}_{v-1} & \frac{(v-1)(v-2)}{v} \end{bmatrix}$$

$$\mathbf{C}_{12}^* = \mathbf{M}^* - \mathbf{N}_1^*\mathbf{K}^{*-1}\mathbf{N}_2^{*'} = \begin{bmatrix} -\frac{(v-2)}{v} & \frac{(v-2)}{v(v-1)}\mathbf{1'}_{v-2} & \frac{(v-2)}{v(v-1)} \\ \frac{2}{v}\mathbf{1}_{v-2} & \frac{(v-2)}{v(v-1)}\mathbf{J}_{v-2} - \mathbf{I}_{v-2} & \frac{(v-2)}{v(v-1)}\mathbf{1}_{v-2} \\ -\frac{(v-2)}{v} & \frac{2}{v}\mathbf{1'}_{v-2} & -\frac{(v-2)}{v} \end{bmatrix}$$

$$\mathbf{C}_{22}^* = \mathbf{R}_\delta^* - \mathbf{N}_2^*\mathbf{K}^{*-1}\mathbf{N}_2^{*'} = \begin{bmatrix} \frac{(v-1)(v-2)}{v} & -\frac{(v-2)}{v}\mathbf{1'}_{v-1} \\ -\frac{(v-2)}{v}\mathbf{1}_{v-1} & (v-1)\mathbf{I}_{v-1} - \frac{(v^2-2v+2)}{v(v-1)}\mathbf{J}_{v-1} \end{bmatrix}$$

and

$$\mathbf{C}_{22}^{*-} = \begin{bmatrix} 0 & \mathbf{0'}_{v-1} \\ \mathbf{0}_{v-1} & \frac{1}{(v-1)}\mathbf{I}_{v-1} + \frac{(v^2-2v+2)}{(v-1)^2(v-2)}\mathbf{J}_{v-1} \end{bmatrix}.$$

Therefore, the information matrix for estimating direct effect of treatments after missing of one observation is

$$\mathbf{C}_\tau^* = \begin{bmatrix} (v-2) & -\frac{(v-2)}{(v-1)}\mathbf{1'}_{v-2} & -\frac{(v-2)}{(v-1)} \\ -\frac{(v-2)}{(v-1)}\mathbf{1}_{v-2} & \frac{(v^2-2v)}{(v-1)}\mathbf{I}_{v-2} - \frac{(v^4-6v^3+14v^2-13v+4)}{(v-1)^3(v-2)}\mathbf{J}_{v-2} & -\frac{(v^2-4v+2)}{(v-1)^2}\mathbf{1}_{v-2} \\ \frac{(v-2)}{(v-1)} & -\frac{(v^2-4v+2)}{(v-1)^2}\mathbf{1'}_{v-2} & \frac{(v^3-5v^2+7v-2)}{(v-1)^2} \end{bmatrix}.$$

The non-zero eigenvalues of $\mathbf{C}_\tau^*$ are obtained as $v(v-2)/(v-1)$with multiplicity $(v-2)$ and $v(v-3)/(v-1)$ with multiplicity one. The efficiency of residual design as per (1) is worked out as:

$$\mathrm{E} = (v-1)(v-3)/(v-2)^2.$$

Table 1
Efficiencies of complete NBB designs with missing observation(s)

Design parameters		Number of missing observation(s)	Efficiency of residual design for direct effects relative to original
Number of treatments	Number of blocks		
5	4	1	0.89
		2	0.79
7	6	1	0.96
		2	0.92
		3	0.88
		Right most plots of each block	0.79
11	10	1	0.99
		2	0.98
		3	0.96
		4	0.95
		5	0.94
		6	0.93
		7	0.92
		8	0.91
		Right most plots of each block	0.90
13	12	1	0.99
		2	0.98
		3	0.98
		4	0.97
		5	0.96
		6	0.95
		7	0.94
		Right most plots of each block	0.91
17	16	1	1.00
		2	0.99
		3	0.99
		4	0.98
		7	0.97
		9	0.96
		11	0.95
		13	0.95
		14	0.94
		Right most plots of each block	0.93
19	18	1	1.00
		2	0.99
		3	0.99
		5	0.98
		8	0.97
		11	0.96
		14	0.95
		17	0.94
		Right most plots of each block	0.94

We consider design to be robust if the loss in efficiency of the residual design is not more than 5% and fairly robust if the loss in efficiency is between 5% and 10%.

Table 1 gives the efficiency of residual design for $v < 20$ (only for prime numbers) obtained by missing one observation from last plot of blocks of the class of designs described in the beginning of this section. Since the design is totally balanced in the sense that variance of any estimated elementary contrast among the direct effects and among left neighbour effects of treatments is constant, therefore the efficiencies of only direct effects have been reported. It is seen that as v increases efficiency increases. Thus, the class of complete NBB designs considered are robust as per the efficiency criteria given in Eq. (1) for the number of treatments exceeding five when one observation is missing.

Obtaining theoretical expression of $\mathbf{C}_{\tau}^{*}$ for missing of more than one observation is complicated. Hence, it is difficult to find out the eigenvalues of this matrix in explicit form. The information matrix and eigenvalues of the information matrix have been thus obtained by developing a SAS code using PROC IML (SAS software package).

Table 2
Efficiencies of complete NBB design when observation(s) missing from last block

Design parameters		Number of missing observation(s)	Efficiency of residual design for direct effects relative to original
Number of treatments	Number of blocks		
5	4	1	0.89
		2	0.79
7	6	1	0.96
		2	0.92
		3	0.88
		Last block	0.80
11	10	1	0.99
		2	0.98
		3	0.96
		4	0.95
		5	0.94
		6	0.93
		7	0.92
		8	0.91
		9	0.90
		Last block	0.89
13	12	1	0.99
		2	0.98
		3	0.98
		4	0.97
		5	0.96
		6	0.95
		Last block	0.91
17	16	1	1.00
		2	0.99
		3	0.99
		4	0.98
		7	0.97
		9	0.96
		11	0.95
		13	0.95
		Last block	0.93
19	18	1	1.00
		2	0.99
		3	0.99
		5	0.98
		8	0.97
		11	0.96
		14	0.95
		Last block	0.94

The efficiency of residual design for the loss of more than one observation is also reported in Table 1. Here also, the efficiency is quite high for the loss of few observations except for the case $v = 5$. But there is a decreasing trend in efficiency with increase in number of missing observations.

The non availability of observations may also happen from blocks in many situations. Suppose in an agricultural experiment, there is a patch of pest damage in the field and as a consequence the experimenter was unable to get observations from those blocks which happened to be in that patch. It is therefore logical to study the robustness of the designs against loss of observations from blocks. Table 2 present the efficiency of residual design after missing of observations from last block of complete NBB design. Here also the efficiencies are quite high and the designs are fairly robust against missing observations.

Acknowledgement

We are grateful to the editor and a referee for the comments that helped to improve the paper.

References

[1] A. Das and S. Kageyama, Robustness of BIB and extended BIB designs, *Computational Statistic & Data Analysis* **14** (1992), 343–358.
[2] A. Dey, Robustness of block design against missing data, *Statistica Sinica* **3** (1993), 219–231.
[3] A. Hedayat and P.W.M. John, Resistant and susceptible BIB designs, *Ann Statist* **2** (1974), 148–158.
[4] D.K. Bhaumik and D.C. Whittinghill, Optimality and robustness to the unavailability of blocks in block designs, *J Roy Statist Soc B* **53**(2) (1991), 399–407.
[5] J.M. Azais, R.A. Bailey and H. Monod, A catalogue of efficient neighbour design with border plots, *Biometrics* **49** (1993), 1252–1261.
[6] J.S. Tomar, S. Jaggi and C. Varghese, On totally balanced block designs for competition effects, *J Appl Statist* **32**(1) (2005), 87–97.
[7] R. Srivastava, V.K. Gupta and A. Dey, Robustness of some designs against missing data, *J Appl Statist* **18** (1990), 313–318.
[8] S. Ghosh, On robustness of designs against incomplete data, *Sankhya B* **40** (1978), 204–208.
[9] S. Ghosh, Robustness of BIB designs against the unavailability of data, *J Statist Plann Inf* **6** (1982), 29–32.
[10] S. Ghosh, Robustness of designs against non-availability of data, *Sankhya, B* **44** (1982), 50–62.
[11] S. Ghosh, Information in an observation in robust designs, *Commun Statist – Theory Meth* **11** (1982), 1173–1184.
[12] S. Ghosh, S. Kageyama and R. Mukerjee, Efficiency of connected binary block designs when a single observation is unavailable, *Ann Inst Statist Math* **44** (1992), 593–603.
[13] S. Ghosh, S.B. Rao and N.M. Singh, On a robust property of PBIB designs, *J Statist Plann Inf* **8** (1983), 355–363.
[14] S. Jaggi, C. Varghese and V.K. Gupta, Optimal circular block designs for neighbouring competition effects, *J Appl Statist* **34**(5) (2007), 577–584.
[15] S. Jaggi, V.K. Gupta and J. Ashraf, On block designs partially balanced for neighbouring competition effects, *J Ind Statist Assoc* **44** (2006), 27–41.
[16] V.K. Gupta and R. Srivastava, Investigation of robustness of block designs against missing observations, *Sankhya, B* **54** (1992), 100–106.