Statistics and Applications {ISSN 2454-7395 (online)}
Special Issue in Memory of Prof. C R Rao
Volume 22, No. 3, 2024 (New Series), pp 243–257
http://www.ssca.org.in/journal

Split-plot Designs with Main Plot Treatments in Incomplete Blocks

B. N. Mandal[1], Rajender Parsad[2] and Sukanta Dash[2]
[1]*ICAR-Indian Agricultural Research Institute, Jharkhand 825405, India*
[2]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110012, India*

Received: 14 January 2024; Revised: 16 August 2024; Accepted: 20 August 2024

Abstract

Split-plot designs are widely used in agricultural experiments because of its ability to allocate different factors to plots of different sizes. In standard split-plot designs, main plot treatments are allocated either in a completely randomized design or in a randomized complete block design and subplot treatments are allocated within each main plot. In this paper, we consider split-plot designs where main plot treatments are allocated in a connected incomplete block design. We propose a method of construction and present a catalogue of such designs. We also propose a method of analysis of such split-plot designs. We have implemented proposed construction and analysis methods using R language.

Key words: Split-plot design; Main plot; Whole plot; Subplot; Construction; Analysis.

AMS Subject Classifications: 62K10, 62K15

Prologue

Today, we are all united in our desire to pay our respect to Late Prof. Calyampudi Radhakrishna Rao. Prof. Rao, an Oracle in the field of Statistics, left an indelible mark on the fields of statistics, mathematics, and scientific research worldwide. His groundbreaking contributions have influenced diverse areas, including economics, genetics, anthropology, and medicine. Rao received numerous accolades, including the US National Medal of Science in 2002, and was awarded the International Prize in Statistics in 2023 - a distinction often likened to the 'statistics' equivalent of the Nobel Prize. His legacy continues to inspire generations, and he remains one of the most influential statisticians of all time. It gives us immense pleasure to know that the Society of Statistics, Computer and Applications has decided to bring out a Special Issue of the Statistics and Applications in memory of Late Prof. C R. Rao. This paper is a tribute in honour and loving memory of Late Prof. C R Rao who had a strong bondage with ICAR-Indian Agricultural Statistics Research Institute (ICAR-IASRI), New Delhi and the Indian Society of Agricultural Statistics. He visited the Institute during 2001 to receive Sankhyiki Bhusan Title conferred upon him by the Indian Society of Agricultural Statistics. His keynote address on 'Has Statistics a Future ? If So, in

Corresponding Author: Rajender Parsad
Email: rajender.parsad@icar.gov.in

What Form?' during the 60th Annual Conference of Indian Society of Agricultural Statistics and International Conference on Statistics and Informatics organized by ICAR-IASRI, New Delhi was published in the Journal of the Indian Society of Agricultural Statistics. The paper had set the tone for the requirement of transformation in Statistics in the era of Information and Communication Technology and Big Data. He has made monumental contributions to Design of Experiments. We have also prepared a Technical Bulletin entitled 'CR Rao's Life Sketch and its Influence on Designing of Experiments with a special reference to Agricultural Sciences' available at `http://krishi.icar.gov.in/jspui/handle/123456789/41295`.

By giving us an opportunity to contribute to the Special Issue, we have been given a chance to say thank you, Prof Rao, for paving the way and developing the playground of Statistics where all statisticians like us are working. We express our profound thankfulness to the Guest Editors of this Special Issue and the Chair Editor of Statistics and Applications for giving this opportunity to contribute in such an invaluable Special Issue.

1. Introduction

A split-plot design is a special kind of design in which two factors A and B with m and s levels, respectively, are allocated such that m levels of factor A (also called main plot treatments) are allocated in main plots using a suitable design and s levels of factor B (also called subplot treatments) are allocated to s smaller subplots within each main plot. These designs were originally developed by Fisher (1925). Popular choice of suitable design for levels of factor A is either a completely randomized design or a randomized complete block design. In a split-plot design, the main effect of B and interaction AB are estimated with higher precision and main effect of A are estimated with lesser precision. The main advantage of a split plot design is that the design can accommodate two different plot sizes for two different factors and is, thus, used in many agricultural and other experiments where one of the factor requires comparatively bigger plot size than the other factor. For example, consider an experiment involving irrigation methods (factor A) and fertilizer doses (factor B). It is possible to apply fertilizer doses in smaller plots but application of irrigation methods require bigger plots. So one can apply irrigation methods to bigger plots first and then each bigger plot is subdivided into smaller plots for application of different fertilizer doses. Other such experiments include study of tillage systems (factor A) and various management practices such as doses of fertilizer, pesticides *etc.* as factor B. Split plot designs are adopted in all such experiments where it is not practical to apply both the levels of factor A and B to plots of same size.

In certain situations it may not be possible to allocate all the m levels of factor A in a randomized complete block design and the number of main plots in each block may be restricted to k such that $k < m$. When m is moderately large, then it may not be possible to maintain homogeneity within the blocks with m main plots as these plots are bigger in size. Hence, it is advisable to use lesser number of main plots in such cases and as a result, an incomplete split-plot designs with blocks being incomplete with respect to main plot treatments is preferable. Robinson (1970) pioneered the idea of incomplete split-plot designs in which he arranged the levels of factor A and B in balanced incomplete block (BIB) designs. Bhargava and Shah (1975) considered incomplete split-plot design with main plot treatments in an incomplete block design where they considered unequal block sizes for main plot treatments and mainly studied tests for main effects of factor B and interaction AB.

Mathew and Sinha (1992) went a step further and presented various optimum and exact tests under fixed, random and mixed effects models in the case of unbalanced split-plot designs where main plot treatments are replicated unequal number of times. Mejza (1985) considered incomplete split-plot designs with main plot treatments in incomplete blocks and presented an analysis procedure with a different model than we study here.

There are some other works on incomplete split plot designs where particular classes of incomplete block designs were used either to allocate main plot and / or subplot treatments. Ozawa *et al.* (2004) obtained incomplete split-plot designs using Kronecker product of two component designs, one for levels of factor A and another for levels of factor B. Ozawa and Kuriki (2006) constructed incomplete split-plot designs using semi-Kronecker product of two types of α-resolvable designs. Kuriki and Nakajima (2007) constructed incomplete split-plot designs by semi-Kronecker product of two resolvable designs with second design being a square lattice design for factor B. Kristensen (2012) proposed four methods of constructing incomplete split-plot designs using α-designs. Works on incomplete split-plot designs considering subplot treatments in an incomplete block design are also available, see, for example, Robinson (1967); Mejza and Mejza (1984) and Mandal *et al.* (2020).

In this article, we consider incomplete split-plot designs where m levels of factor A are arranged in a connected incomplete block design with blocks of each of size k such that $k < m$ and s levels of factor B are allocated in s subplots within each main plot. We propose a methodology of analysis of data from experiments conducted using such designs following the standard fixed effects additive linear model approach. Since in agricultural experiments, generally factors and their levels are only a carefully chosen entities among which comparisons are desired, and blocks are also not a random sample from bigger population of blocks, random effects and mixed effects models for analysis of split-plot designs are not considered here and thus, we restrict ourselves to fixed effects model only.

2. Construction

In this section, we present construction of incomplete split plot designs where m levels of factor A are arranged in a connected proper binary incomplete block design with blocks of same size and s levels of factor B are arranged randomly within each main plot. To construct a design, take a binary connected proper incomplete block design D with number of treatments m, number of blocks b and block size $k < m$. Arrange the m levels of factor A using design D. Within each level of factor A, apply s subplot treatments at random. Obtained design is an incomplete split plot design where blocks are incomplete with respect to factor A and whole plots are complete with respect to factor B.

We illustrate the construction with an example.

Example 1: Let $m = 5, s = 5, b = 5, k = 3$. So a connected binary proper incomplete block design D for factor A is

$$\begin{matrix} (1 & 4 & 5) \\ (2 & 3 & 5) \\ (1 & 3 & 4) \\ (2 & 3 & 4) \\ (1 & 2 & 5) \end{matrix}$$

In D, there are 5 blocks and in each block, three main plot treatments are allocated. Now, randomly assign each of the s levels of factor B in each of the main plots. We get the following incomplete split-plot design.

Block 1	1 (5 4 3 1 2)	4 (5 4 3 1 2)	5 (2 3 4 5 1)
Block 2	2 (3 4 2 1 5)	3 (2 3 4 5 1)	5 (4 1 5 3 2)
Block 3	1 (3 4 1 5 2)	3 (3 4 1 5 2)	4 (3 4 2 1 5)
Block 4	2 (3 4 1 5 2)	3 (3 4 2 1 5)	4 (4 1 5 3 2)
Block 5	1 (4 1 5 3 2)	2 (2 3 4 5 1)	5 (5 4 3 1 2)

Remark 1: We recommend that the design D should be so chosen that it has high A- and D-efficiency. One can use the available efficient incomplete block designs in literature for this purpose. We utilized A-efficient incomplete block designs generated by the R package *ibd* (Mandal, 2019). If the design D is equireplicate with r replications for each of the m levels of factor A, then in the incomplete split-plot design, each AB combination appears r times. Had a complete split-plot design with b blocks been chosen, each AB treatment combination would have appeared b times. Since number of main plots in an incomplete split-plot design is k in each block, it is expected that blocks would be more homogeneous than a block containing m main plots. This will increase precision of comparisons among main effects of factor A. Further, whenever $r < b$, incomplete split-plot designs is expected to be more resource efficient because then they will require lesser number of main plots. For example, consider an experiment conducted by Pandey *et al.* (2000) who used $m = 5$ levels of irrigation regimes as factor A and $s = 5$ levels of Nitrogen doses as factor B and they used complete split-plot design with four blocks. This experiment required 20 main plots and 100 subplots in total. Had an incomplete split-plot design as given in Example 1 with 5 blocks with block size 3 been used, only 15 main plots and 75 subplots would have been required.

We have used the method to construct incomplete split-plot designs in the restricted parametric range of $m \leq 6, s \leq 6$ and $b \leq 10$. The list of parameters for which design has been generated is available, see Mandal *et al.* (2019c). However, the proposed method is general and works for any m, s, b, k provided a suitable connected incomplete block design D with parameters (m, b, k) exists and is available in literature.

3. Analysis

In this section, we present a methodology for analysis of data from experiments conducted using incomplete split-plot designs considered in this paper. We consider fixed effect additive linear model for this purpose:

$$y_{jil} = \mu + \rho_j + \alpha_i + \gamma_{ji} + \beta_l + \delta_{il} + \epsilon_{jil} \quad (1)$$

where y_{jil} denote the observation from the experimental unit in jth block receiving ith level of factor A and lth level of factor B, μ is the general mean, ρ_j is the effect of jth block, α_i is the main effect of ith level of factor A, γ_{ji} is the interaction terms between blocks and ith level of factor A, β_l is the main effect of lth level of factor B, δ_{il} is the interaction effect of ith level of factor A and lth level of factor B and ϵ_{jil} is the random subplot error with zero mean and constant variance $\sigma^2, j = 1, 2, ..., b; i = 1, 2, ..., m; l = 1, 2, ..., s$. Here all the effects are fixed effects except subplot error. Note here that data do not exist for all (j, i, l) combinations since all levels of factor A do not appear within each block. Here it may be mentioned that Mathew and Sinha (1992) also considered a model similar to (1). However, they considered unbalanced cases, *i.e.*, the blocks may be of unequal sizes and may contain different number of main plot treatments and they also considered cases of random and mixed effects scenarios. In our case, each block is of constant size k and contains $k < m$ main plot treatments and we do not consider random and mixed effect models.

In matrix notation, the model (1) may be represented as

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma} + \mathbf{X}_4\boldsymbol{\beta} + \mathbf{X}_5\boldsymbol{\delta} + \boldsymbol{\epsilon} \tag{2}$$

where $\mathbf{y}$ denotes the vector of n observations, $\mathbf{1}$ denotes the vector of ones, $\mathbf{X}_1$ denotes $n \times b$ observation versus block incidence matrix, $\boldsymbol{\rho}$ denotes $b \times 1$ vector of block effects, $\mathbf{X}_2$ denotes $n \times m$ observation versus factor A incidence matrix, $\boldsymbol{\alpha}$ denotes $m \times 1$ vector of main effects of factor A, $\mathbf{X}_3$ denotes $n \times bk$ observation versus block-A incidence matrix, $\boldsymbol{\gamma}$ denotes $bk \times 1$ vector of block versus factor A interactions, $\mathbf{X}_4$ denotes $n \times s$ observation versus factor B incidence matrix, $\boldsymbol{\beta}$ denotes $s \times 1$ vector of main effects of factor B, $\mathbf{X}_5$ denotes $n \times ms$ observation versus AB interaction incidence matrix, $\boldsymbol{\delta}$ denotes $ms \times 1$ vector of AB interaction effects and $\boldsymbol{\epsilon}$ denotes $n \times 1$ vector of errors. We assume that errors are i.i.d. normal with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}_n$. Under the given set-up of design construction, $n = bks$ and $\mathbf{X}_1 = \mathbf{I}_b \otimes \mathbf{1}_{ks}$. The model (2) can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{3}$$

where $\mathbf{X} = (\mathbf{1} : \mathbf{X}_1 : \mathbf{X}_2 : \mathbf{X}_3 : \mathbf{X}_4 : \mathbf{X}_5)$ and $\boldsymbol{\theta} = (\mu, \boldsymbol{\rho}', \boldsymbol{\alpha}', \boldsymbol{\gamma}', \boldsymbol{\beta}', \boldsymbol{\delta}')'$.

Normal equations are given by

$$\mathbf{X}'\mathbf{X}\boldsymbol{\theta} = \mathbf{X}'\mathbf{y}$$

where

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{X}_1 & \mathbf{1}'\mathbf{X}_2 & \mathbf{1}'\mathbf{X}_3 & \mathbf{1}'\mathbf{X}_4 & \mathbf{1}'\mathbf{X}_5 \\ \mathbf{X}_1'\mathbf{1} & \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{X}_1'\mathbf{X}_3 & \mathbf{X}_1'\mathbf{X}_4 & \mathbf{X}_1'\mathbf{X}_5 \\ \mathbf{X}_2'\mathbf{1} & \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 & \mathbf{X}_2'\mathbf{X}_3 & \mathbf{X}_2'\mathbf{X}_4 & \mathbf{X}_2'\mathbf{X}_5 \\ \mathbf{X}_3'\mathbf{1} & \mathbf{X}_3'\mathbf{X}_1 & \mathbf{X}_3'\mathbf{X}_2 & \mathbf{X}_3'\mathbf{X}_3 & \mathbf{X}_3'\mathbf{X}_4 & \mathbf{X}_3'\mathbf{X}_5 \\ \mathbf{X}_4'\mathbf{1} & \mathbf{X}_4'\mathbf{X}_1 & \mathbf{X}_4'\mathbf{X}_2 & \mathbf{X}_4'\mathbf{X}_3 & \mathbf{X}_4'\mathbf{X}_4 & \mathbf{X}_4'\mathbf{X}_5 \\ \mathbf{X}_5'\mathbf{1} & \mathbf{X}_5'\mathbf{X}_1 & \mathbf{X}_5'\mathbf{X}_2 & \mathbf{X}_5'\mathbf{X}_3 & \mathbf{X}_5'\mathbf{X}_4 & \mathbf{X}_5'\mathbf{X}_5 \end{pmatrix}. \tag{4}$$

Now, following relations can be verified:

$$\begin{aligned} &\mathbf{X}_1'\mathbf{1} = ks\mathbf{1}_b && \mathbf{X}_2'\mathbf{1} = s\mathbf{r} \\ &\mathbf{X}_3'\mathbf{1} = s\mathbf{1}_{bk} && \mathbf{X}_4'\mathbf{1} = bk\mathbf{1}_s \\ &\mathbf{X}_5'\mathbf{1} = \mathbf{r} \otimes \mathbf{1}_s && \mathbf{X}_1'\mathbf{X}_1 = ks\mathbf{I}_b \\ &\mathbf{X}_2'\mathbf{X}_1 = s\mathbf{N}_1, \text{ say} && \mathbf{X}_3'\mathbf{X}_1 = s\mathbf{1}_k \otimes \mathbf{I}_b' \\ &\mathbf{X}_4'\mathbf{X}_1 = k\mathbf{1}_s\mathbf{1}_b' && \mathbf{X}_5'\mathbf{X}_1 = \mathbf{N}_3, \text{ say} \\ &\mathbf{X}_2'\mathbf{X}_2 = s\mathbf{R} && \mathbf{X}_3'\mathbf{X}_2 = \mathbf{N}_2, \text{ say} \\ &\mathbf{X}_4'\mathbf{X}_2 = \mathbf{r}' \otimes \mathbf{1}_s && \mathbf{X}_5'\mathbf{X}_2 = \mathbf{R} \otimes \mathbf{1}_s \\ &\mathbf{X}_3'\mathbf{X}_3 = s\mathbf{I}_{bk} && \mathbf{X}_4'\mathbf{X}_3 = \mathbf{1}_s\mathbf{1}_{bk}', \text{ say} \\ &\mathbf{X}_5'\mathbf{X}_3 = \mathbf{N}_4, \text{ say} && \mathbf{X}_4'\mathbf{X}_4 = bk\mathbf{I}_s \\ &\mathbf{X}_5'\mathbf{X}_4 = \mathbf{r} \otimes \mathbf{I}_s, \text{ say} && \mathbf{X}_5'\mathbf{X}_5 = \mathbf{R} \otimes \mathbf{I}_s \end{aligned}$$

with $\mathbf{r}$ being the vector of replications of levels of factor A and $\mathbf{R}$ being diagonal matrix with elements of $\mathbf{r}$. Therefore, equation (4) can be written as

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} bks & ks\mathbf{1}'_b & s\mathbf{r}' & s\mathbf{1}'_{bk} & bk\mathbf{1}'_s & \mathbf{r}'\otimes\mathbf{1}'_s \\ ks\mathbf{1}_b & ks\mathbf{I}_b & s\mathbf{N}'_1 & s\mathbf{1}'_k\otimes\mathbf{I}_b & k\mathbf{1}_b\mathbf{1}'_s & \mathbf{N}'_3 \\ s\mathbf{r} & s\mathbf{N}_1 & s\mathbf{R} & \mathbf{N}'_2 & \mathbf{r}\otimes\mathbf{1}'_s & \mathbf{R}\otimes\mathbf{1}'_s \\ s\mathbf{1}_{bk} & s\mathbf{1}_k\otimes\mathbf{I}_b & \mathbf{N}_2 & s\mathbf{I}_{bk} & \mathbf{1}_{bk}\mathbf{1}'_s & \mathbf{N}'_4 \\ bk\mathbf{1}_s & k\mathbf{1}_s\mathbf{1}'_b & \mathbf{r}'\otimes\mathbf{1}_s & \mathbf{1}_s\mathbf{1}'_{bk} & bk\mathbf{I}_s & \mathbf{r}'\otimes\mathbf{I}_s \\ \mathbf{r}\otimes\mathbf{1}_s & \mathbf{N}_3 & \mathbf{R}\otimes\mathbf{1}_s & \mathbf{N}_4 & \mathbf{r}\otimes\mathbf{I}_s & \mathbf{R}\otimes\mathbf{I}_s \end{pmatrix}. \tag{5}$$

It may be seen that

$$\mathbf{X}'\mathbf{y} = (y_{...} : \mathbf{y}'_{B..} : \mathbf{y}'_{.M.} : \mathbf{y}'_{BM.} : \mathbf{y}'_{.S.} : \mathbf{y}'_{.MS})'$$

where $y_{...}$ denote the gross total of all observations, $\mathbf{y}_{B..}$ is the vector of block totals, $\mathbf{y}_{.M.}$ is the vector of totals for m levels of factor A, $\mathbf{y}_{BM.}$ is the vector of totals corresponding to block-factor A combinations, $\mathbf{y}_{..S}$ is the vector of totals for s levels of factor B and $\mathbf{y}_{.MS}$ is the vector of totals corresponding to AB combinations.

One can verify that the number of rows in $\mathbf{X}'\mathbf{X}$ is $1+b+m+bk+s+ms$, but there are total $1+1+(m+b-1)+1+(m+s-1) = 1+b+2m+s$ linearly dependent rows and they are as follows: sum of 2nd to $(b+1)$th row is equal to the first row, sum of $(b+2)$th row to $(b+m+2)$th row is equal to the first row, summing rows for each level of γ_{ji} over i keeping j fixed gives row corresponding to jth $(j = 1, 2, ..., b)$ block and similarly summing rows for each level of γ_{ji} over j keeping i fixed gives row corresponding to row of ith $(i = 1, 2, ..., m)$ level of factor A, summing of rows corresponding to s levels of factor B gives the first row, summing rows for each level of δ_{il} over i keeping l fixed gives row corresponding to lth $(l = 1, 2, ..., s)$ level of factor B, summing rows for each level of δ_{il} over l keeping i fixed gives row corresponding to ith level of factor A. Therefore, to get a solution to the normal equations (3), one can set $(1+b+2m+s)$ parameter estimates to zero. We set $\hat{\mu} = 0, \hat{\rho}_j = 0, \hat{\alpha}_i = 0, \hat{\beta}_l = 0 \forall j, i, l$ and we also set every sth component of $\hat{\boldsymbol{\delta}}$ as zero, *i.e.*, $\hat{\delta}_s = 0, \hat{\delta}_{2s} = 0, ..., \hat{\delta}_{ms} = 0$. As a result, we get,

$$\begin{pmatrix} s\mathbf{I}_{bk} & \tilde{\mathbf{N}}'_4 \\ \tilde{\mathbf{N}}_4 & \mathbf{R}\otimes\mathbf{I}_{s-1} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\boldsymbol{\delta}}_{(-m)} \end{pmatrix} = \begin{pmatrix} \mathbf{y}_{BM.} \\ \tilde{\mathbf{y}}_{.MS} \end{pmatrix} \tag{6}$$

where $\tilde{\mathbf{N}}_4$ is the matrix obtained after removing every sth row of $\mathbf{N}_4$, $\hat{\boldsymbol{\delta}}_{(-m)}$ is the vector after removing every sth element of $\hat{\boldsymbol{\delta}}$ and $\tilde{\mathbf{y}}_{.MS}$ is the vector obtained after removing every sth element of $\mathbf{y}_{.MS}$. From (6), we get,

$$\hat{\boldsymbol{\gamma}} = \frac{1}{s}\left(\mathbf{y}_{BM.} - \tilde{\mathbf{N}}'_4\hat{\boldsymbol{\delta}}_{(-m)}\right).$$

After a little algebra, it may be seen that

$$\hat{\boldsymbol{\delta}}_{(-m)} = \mathbf{C}_{MS}^{-1}\mathbf{Q}_{MS}$$

where $\mathbf{C}_{MS} = \left(\mathbf{R}\otimes\mathbf{I}_{s-1} - \frac{1}{s}\tilde{\mathbf{N}}_4\tilde{\mathbf{N}}'_4\right)$ and $\mathbf{Q}_{MS} = \left(\tilde{\mathbf{y}}_{.MS} - \frac{1}{s}\tilde{\mathbf{N}}_4\mathbf{y}_{BM.}\right)$.

Denoting the model sum of squares due to fitting parameters $\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\delta}$ with $R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\delta})$, we get

$$R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\delta}) = \hat{\boldsymbol{\gamma}}'\mathbf{X}_3'\mathbf{y} + \hat{\boldsymbol{\delta}}'\mathbf{X}_5'\mathbf{y} = \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} + \mathbf{Q}_{MS}'\mathbf{C}_{MS}^{-1}\mathbf{Q}_{MS}.$$

Similarly, it may be verified that

$$R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \hat{\boldsymbol{\gamma}}'\mathbf{X}_3'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}_4'\mathbf{y} = \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} + \mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S$$

where $\mathbf{C}_S = bk\mathbf{I}_{s-1} - \frac{bk}{s}\mathbf{1}_{s-1}\mathbf{1}_{s-1}'$ and $\mathbf{Q}_S = \tilde{\mathbf{y}}_{..S} - \frac{y_{...}}{s}\mathbf{1}_{s-1}$

$$R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) = \hat{\boldsymbol{\rho}}'\mathbf{X}_1'\mathbf{y} + \hat{\boldsymbol{\alpha}}'\mathbf{X}_2'\mathbf{y} + \hat{\boldsymbol{\gamma}}'\mathbf{X}_3'\mathbf{y} = \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.}$$

$$R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}) = \hat{\boldsymbol{\rho}}'\mathbf{X}_1'\mathbf{y} + \hat{\boldsymbol{\alpha}}'\mathbf{X}_2'\mathbf{y} = \frac{1}{ks}\mathbf{y}_{B..}'\mathbf{y}_{B..} + \mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M$$

where $\mathbf{C}_M = s\mathbf{R}_{m-1} - \frac{s}{k}\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1'$ and $\mathbf{Q}_M = \tilde{\mathbf{y}}_{.M.} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{y}_{B..}$

$$R(\mu, \boldsymbol{\rho}) = \hat{\boldsymbol{\rho}}'\mathbf{X}_1'\mathbf{y} = \frac{1}{ks}\mathbf{y}_{B..}'\mathbf{y}_{B..}$$

and

$$R(\mu) = \frac{1}{bks}y_{...}{}^2$$

Residual sum of squares after fitting the model (3) is given by

$$SSE = \mathbf{y}'\mathbf{y} - \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} - \mathbf{Q}_{MS}'\mathbf{C}_{MS}^{-1}\mathbf{Q}_{MS}. \tag{7}$$

Theorem 1: Under model (3), $SSE/\sigma^2 \sim \chi^2_{bks-bk-ms+m}$.

Proof: It is well known that in a fixed effects linear model (3), $SSE/\sigma^2 \sim \chi^2_{n-rank(\mathbf{X})}$. Here, $rank(\mathbf{X}) = rank(\mathbf{X}'\mathbf{X}) = bk + ms - m$. So the result follows. □

3.1. Testing significance of interactions between A and B

Consider the null hypothesis $H_0 : \delta_{i1} = \delta_{i2} = \cdots = \delta_{is} \forall i = 1, 2, ..., m$ versus H_1 : At least two of them are different. Under the null hypothesis, the reduced model is

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma} + \mathbf{X}_4\boldsymbol{\beta} + \boldsymbol{\epsilon}.$$

The residual sum of squares under reduced model is

$$SSE_1 = \mathbf{y}'\mathbf{y} - R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \mathbf{y}'\mathbf{y} - \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} - \mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S.$$

Theorem 2: $SSE_1/\sigma^2 \sim \chi^2_{bks-bk-s+1}$.

Proof: The rank of the model matrix $\mathbf{X}_{r1} = (\mathbf{1} : \mathbf{X}_1 : \mathbf{X}_2 : \mathbf{X}_3 : \mathbf{X}_4)$ is $bk + s - 1$ since out of $(1 + b + m + bk + s)$ rows of $\mathbf{X}'_{r1}\mathbf{X}_{r1}$, there are $b + m + 2$ dependencies. Hence, the result follows. □ Now, $SSE_1 - SSE = \mathbf{Q}'_{MS}\mathbf{C}^{-1}_{MS}\mathbf{Q}_{MS} - \mathbf{Q}'_S\mathbf{C}^{-1}_S\mathbf{Q}_S$. Therefore, the test statistic for testing $H_0 : \delta_{i1} = \delta_{i2} = \cdots = \delta_{is} \forall i = 1, 2, ..., m$ is

$$\begin{aligned} F_1 &= \frac{(SSE_1 - SSE)/(m-1)(s-1)}{SSE/(bks - bk - ms + m)} \\ &= \frac{(\mathbf{Q}'_{MS}\mathbf{C}^{-1}_{MS}\mathbf{Q}_{MS} - \mathbf{Q}'_S\mathbf{C}^{-1}_S\mathbf{Q}_S)/(m-1)(s-1)}{SSE/(bks - bk - ms + m)} \sim F_{(m-1)(s-1),(bks-bk-ms+m)} \end{aligned}$$

under null hypothesis. Null hypothesis is rejected whenever calculated value of $F_1 > F_{\alpha,(m-1)(s-1),(bks-bk-ms+m)}$ where $F_{\alpha,(m-1)(s-1),(bks-bk-ms+m)}$ denotes the upper α percent point of an F-distribution with $(m-1)(s-1)$ and $(bks - bk - ms + m)$ degrees of freedom.

3.2. Testing significance of main effects of factor B

Assuming that interactions between A and B is absent, we consider the null hypothesis $H_0 : \beta_1 = \beta_2 = ... = \beta_s = \beta$, say versus the alternative H_1 : At least two of them are different. Consider the following test statistic

$$F_2 = \frac{\mathbf{Q}'_S\mathbf{C}^{-1}_S\mathbf{Q}_S/(s-1)}{SSE/(bks - bk - ms + m)}$$

which follows $F_{(s-1),(bks-bk-ms+m)}$, See Appendix for proof. One can reject the null hypothesis when calculated value of $F_2 > F_{\alpha,(s-1),(bks-bk-ms+m)}$.

3.3. Testing significance of main effects of factor A

Since main effects of A can be tested when interactions of A with B and with blocks is absent, we assume that interactions between A and B is absent and then we consider the null hypothesis $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_m = \alpha$, say versus the alternative H_1 : At least two of them are different. One can see that

$$F_3 = \frac{\mathbf{Q}'_M\mathbf{C}^{-1}_M\mathbf{Q}_M/(m-1)}{SSW/(bk - b - m + 1)} \sim F_{(m-1),(bk-b-m+1)}.$$

$$SSW = R(\boldsymbol{\gamma}|\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}) = R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\gamma}) - R(\mu, \boldsymbol{\rho}, \boldsymbol{\alpha}) = \tfrac{1}{s}\mathbf{y}'_{BM.}\mathbf{y}_{BM.} - \tfrac{1}{ks}\mathbf{y}'_{B..}\mathbf{y}_{B..} - \mathbf{Q}'_M\mathbf{C}^{-1}_M\mathbf{Q}_M.$$

Above results can be summarized in the form of analysis of variance (ANOVA) table as given in Table 1 where,

$$\begin{array}{ll} SSR = \frac{1}{ks}\mathbf{y}'_{B..}\mathbf{y}_{B..} - \frac{1}{bks}y_{...}^2 & SSA = \mathbf{Q}'_M\mathbf{C}^{-1}_M\mathbf{Q}_M \\ SSW = \frac{1}{s}\mathbf{y}'_{BM.}\mathbf{y}_{BM.} - \frac{1}{ks}\mathbf{y}'_{B..}\mathbf{y}_{B..} - \mathbf{Q}'_M\mathbf{C}^{-1}_M\mathbf{Q}_M. & SSB = \mathbf{Q}'_S\mathbf{C}^{-1}_S\mathbf{Q}_S \\ SSAB = \mathbf{Q}'_{MS}\mathbf{C}^{-1}_{MS}\mathbf{Q}_{MS} - \mathbf{Q}'_S\mathbf{C}^{-1}_S\mathbf{Q}_S & SST = \mathbf{y}'\mathbf{y} - \frac{1}{bks}y_{...}^2 \end{array}$$

Remark 2: The model formulation under a split-plot design often involves a whole-plot error and a split-plot error, both of which are assumed to be random, satisfying the usual normality assumptions. This formulation leads to two ANOVA tables: a whole-plot ANOVA and a split-plot ANOVA. The present work considers a model that includes only one random

Table 1: ANOVA table depicting analysis of incomplete split plot designs

Source	Degrees of freedom	Sum of squares	Mean squares	F
Blocks	$b-1$	SSR	-	-
A	$m-1$	SSA	$MSA = SSA/(m-1)$	$F_3 = MSA/MSW$
Block × A	$bk-b-m+1$	SSW	$MSW = SSW/(bk-b-m+1)$	
B	$s-1$	SSB	$MSB = SSB/(s-1)$	$F_2 = MSB/MSE$
AB	$(m-1)(s-1)$	$SSAB$	$MSAB = SSAB/(m-1)(s-1)$	$F_1 = MSAB/MSE$
Error	$bks-bk-ms+m$	SSE	$MSE = SSE/(bks-bk-ms+m)$	-
Total	$bks-1$	SST	-	-

error term. For testing the significance of the main effects due to whole plot factor, we assume that A × B interaction and block × A interactions are absent and then use the mean square due to block × A interaction in the denominator of the F ratio and this F ratio coincides with the corresponding F ratio in the whole-plot ANOVA.

3.4. Estimation of treatment contrasts

First we consider estimation of treatment contrasts of factor B. It may be seen that

$$\hat{\boldsymbol{\beta}}_{s-1} = \mathbf{C}_S^{-1}\mathbf{Q}_S.$$

Thus, we can write

$$\hat{\boldsymbol{\beta}} = \mathbf{C}_S^{*-}\mathbf{Q}_S^{*} \tag{8}$$

where $\mathbf{C}_S^* = bk\mathbf{I}_s - \frac{bk}{s}\mathbf{1}_s\mathbf{1}_s'$ and $\mathbf{Q}_S^* = \mathbf{y}_{..S} - \frac{y_{...}}{s}\mathbf{1}_s$.

Theorem 3: Let $\mathbf{p}'\boldsymbol{\beta}$ be a linear parametric function such that $\mathbf{p}'\mathbf{1} = 0$. Then $\mathbf{p}'\boldsymbol{\beta}$ is estimable.

Proof: Consider the estimator $\mathbf{p}'\hat{\boldsymbol{\beta}}$ where $\hat{\boldsymbol{\beta}}$ is given by equation (8). Then,

$$\begin{aligned}
E(\mathbf{p}'\hat{\boldsymbol{\beta}}) &= E(\mathbf{p}'\mathbf{C}_S^{*-}\mathbf{Q}_S^{*}) \\
&= \mathbf{p}'\mathbf{C}_S^{*-}E(\mathbf{y}_{..S} - \frac{y_{...}}{s}\mathbf{1}_s) \\
&= \mathbf{p}'\mathbf{C}_S^{*-}(\mathbf{X}_4' - \frac{1}{s}\mathbf{X}_4'\mathbf{X}_3\mathbf{X}_3')E(\mathbf{y}) \\
&= \mathbf{p}'\mathbf{C}_S^{*-}(\mathbf{X}_4' - \frac{1}{s}\mathbf{X}_4'\mathbf{X}_3\mathbf{X}_3')(\mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma} + \mathbf{X}_4\boldsymbol{\beta}) \\
&= \mathbf{p}'\mathbf{C}_S^{*-}\mathbf{C}_S^{*}\boldsymbol{\beta} \text{ (after simplification)} \\
&= \mathbf{p}'\boldsymbol{\beta}
\end{aligned}$$

since $\mathbf{p}'\mathbf{1} = 0$. This completes the proof. □

It is easy to see that $v(\mathbf{p}'\hat{\boldsymbol{\beta}}) = \mathbf{p}'\mathbf{C}_S^{*-}\mathbf{p}\sigma^2$ where $v(.)$ denotes variance. So under normality of errors in model (1), $\mathbf{p}'\hat{\boldsymbol{\beta}} \sim N(\mathbf{p}'\boldsymbol{\beta}, \mathbf{p}'\mathbf{C}_S^{*-}\mathbf{p}\sigma^2)$. Thus, testing of hypothesis

$H_0 : \mathbf{p}'\boldsymbol{\beta} = b$ can be performed using the test statistic

$$F_s = \frac{(\mathbf{p}'\hat{\boldsymbol{\beta}} - b)^2/(\mathbf{p}'\mathbf{C}_S^{*-}\mathbf{p})}{SSE/(bks - bk - ms + m)}.$$

Under null hypothesis $F_s \sim F_{1,bks-bk-ms+m}$.

Exactly on similar lines, it can be proved that for testing $H_0 : \mathbf{q}'\boldsymbol{\delta} = d$, one can use the test statistic

$$F_{ms} = \frac{(\mathbf{q}'\hat{\boldsymbol{\delta}} - d)^2/(\mathbf{q}'\mathbf{C}_{MS}^{*-}\mathbf{q})}{SSE/(bks - bk - ms + m)}$$

which follows F distribution with 1 and $(bks - bk - ms + m)$ degrees of freedom under null hypothesis. Here, $\hat{\boldsymbol{\delta}} = \mathbf{C}_{MS}^{*-}\mathbf{Q}_{MS}^*$ with $\mathbf{C}_{MS}^* = \mathbf{R} \otimes \mathbf{I}_{s-1} - \frac{1}{s}\mathbf{N}_4\mathbf{N}_4'$ and $\mathbf{Q}_{MS}^* = \mathbf{y}_{.MS} - \frac{1}{s}\mathbf{N}_4\mathbf{y}_{BM.}$.

Now consider a treatment contrast $\mathbf{w}'\boldsymbol{\alpha}$ of main effects of factor A. An estimator of this treatment contrast is given by $\mathbf{w}'\hat{\boldsymbol{\alpha}} = \mathbf{w}'\mathbf{C}_M^{*-}\mathbf{Q}_M^*$ where $\mathbf{C}_M^* = s\mathbf{R}_m - \frac{s}{k}\mathbf{N}_1\mathbf{N}_1'$ and $\mathbf{Q}_M^* = \mathbf{y}_{.M.} - \frac{1}{k}\mathbf{N}_1\mathbf{y}_{B..}$. To test $H_0 : \mathbf{w}'\hat{\boldsymbol{\alpha}} = a$, the following test statistic can be used:

$$F_m = \frac{(\mathbf{w}'\hat{\boldsymbol{\alpha}} - a)^2/(\mathbf{w}'\mathbf{C}_M^{*-}\mathbf{w})}{SSW/(bk - b - m + 1)}.$$

Under null hypothesis, $F_m \sim F_{1,(bk-b-m+1)}$ and inferences can be made accordingly.

4. Concluding remarks

In this paper, we have proposed a method of construction of incomplete split-plot designs where main plot treatments are allocated using a connected proper incomplete block design. We have also presented an analysis methodology for the proposed designs. We have implemented the proposed methods of construction and analysis using R language and the same is available as part of an R package 'ispd' which can be accessed on https://cran.r-project.org/web/packages/ispd/index.html, see (Mandal *et al.*, 2019a). Further, we have also implemented the construction and analysis methodology as part of an web application which is available on http://drsr.icar.gov.in/ISPD/Home.jsp, see (Mandal *et al.*, 2019b). The will enable the experimenters and statisticians to use these designs with ease.

Acknowledgements

The authors sincerely acknowledges the financial support received in the form of Extra Mural Research Grant by Science and Engineering Research Board (SERB), Department of Science and Technology, India. We are thankful to the Chair Editor and General Editors for their valuable guidance in improving the manuscript. We are also thankful to the anonymous reviewers for their enlightening remarks which led to considerable improvement in the presentation of the article.

References

Bhargava, R. and Shah, K. (1975). Analysis of some mixed-models for block and split-plot designs. *Annals of the Institute of Statistical Mathematics*, **27**, 365–375.

Fisher, R. A. (1925). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

Kristensen, K. (2012). Incomplete split-plot designs based on α-designs: a compromise between traditional split-plot designs and randomised complete block design. *Euphytica*, **183**, 401–413.

Kuriki, S. and Nakajima, K. (2007). Square lattice designs in incomplete split-plot designs. *Journal of Statistical Theory and Practice*, **1**, 417–426.

Mandal, B. N. (2019). *ibd: Incomplete Block Designs.* R package version 1.5.

Mandal, B. N., Dash, S., and Parsad, R. (2019a). *ispd: Incomplete Split-Plot Designs.* R package version 0.1.

Mandal, B. N., Parsad, R., and Dash, S. (2019b). Incomplete split plot designs : Construction and analysis. http://drsr.icar.gov.in/ISPD/how_to.jsp.

Mandal, B. N., Parsad, R., and Dash, S. (2019c). Incomplete split-plot designs: Blocks are incomplete, main plots are complete : Design resources server. http://drs.icar.gov.in.

Mandal, B. N., Parsad, R., and Dash, S. (2020). Incomplete split-plot designs: Construction and analysis. *Statistics and Probability Letters*, **166**, 108869.

Mathew, T. and Sinha, B. K. (1992). Exact and optimum tests in unbalanced split-plot designs under mixed and random models. *Journal of the American Statistical Association*, **87**, 192–200.

Mejza, I. and Mejza, S. (1984). Incomplete split plot designs. *Statistics and Probability Letters*, **2**, 327–332.

Mejza, S. (1985). A split-plot design with wholeplot treatments in an incomplete block design. In *Linear Statistical Inference*, pages 211–222. Springer.

Ozawa, K. and Kuriki, S. (2006). Incomplete split-plot designs generated from α-resolvable designs. *Statistics and Probability Letters*, **76**, 1245–1254.

Ozawa, K., Mejza, S., Jimbo, M., Mejza, I., and Kuriki, S. (2004). Incomplete split-plot designs generated by some resolvable balanced designs. *Statistics and probability letters*, **68**, 9–15.

Pandey, R., Maranville, J., and Admou, A. (2000). Deficit irrigation and nitrogen effects on maize in a sahelian environment: I. grain yield and yield components. *Agricultural Water Management*, **46**, 1–13.

Robinson, J. (1967). Incomplete split plot designs. *Biometrics*, **23**, 793–802.

Robinson, J. (1970). Blocking in incomplete split plot designs. *Biometrika*, **57**, 347–350.

Appendix

Proof of F_2 following $F_{(s-1),(bks-bk-ms+m)}$:
First we prove that $\mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S/\sigma^2 \sim \chi^{2'}_{s-1}$ with non-centrality parameter $\boldsymbol{\theta}'\mathbf{X}'(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}')$ $\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')\mathbf{X}\boldsymbol{\theta}/2\sigma^2$.

Note that $\mathbf{Q}_S = \left(\tilde{\mathbf{y}}_{..S} - \frac{y_{...}}{s}\mathbf{1}_{s-1}\right) = \tilde{\mathbf{X}}_4'\mathbf{y} - \frac{1}{s}\tilde{\mathbf{X}}_4'\mathbf{X}_3\mathbf{X}_3'\mathbf{y}$ and hence $\mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S = \mathbf{y}'(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'\mathbf{y}$. Now, $(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'$ is idempotent because

$$\begin{aligned}
&(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)' \\
&= (\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}')\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}')\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3') \\
&= (\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}')\mathbf{C}_s^{-1}\mathbf{C}_s\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')
\end{aligned}$$

since

$$\begin{aligned}
&(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}') \\
&= \tilde{\mathbf{X}}_4'\tilde{\mathbf{X}}_4 - \frac{1}{s}\tilde{\mathbf{X}}_4'\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\tilde{\mathbf{X}}_4 + \frac{1}{s^2}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\mathbf{X}_3\mathbf{1}_{bk}\mathbf{1}_{s-1}' \\
&= bk(\mathbf{I}_{s-1} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{s-1}' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{s-1}' + \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{s-1}') \\
&= bk((\mathbf{I}_{s-1} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{s-1}') \\
&= \mathbf{C}_s.
\end{aligned}$$

Now, under H_0,

$$\begin{aligned}
\mathbf{X}\boldsymbol{\theta} &= \mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma} + \beta\mathbf{X}_4\mathbf{1}_s \\
&= \mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma} + \beta\mathbf{1} \\
&= (\mu+\beta)\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&(\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')\mathbf{X}\boldsymbol{\theta} \\
&= (\tilde{\mathbf{X}}_4' - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3')((\mu+\beta)\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha} + \mathbf{X}_3\boldsymbol{\gamma}) \\
&= (\mu+\beta)\tilde{\mathbf{X}}_4'\mathbf{1} + \tilde{\mathbf{X}}_4'\mathbf{X}_1\boldsymbol{\rho} + \tilde{\mathbf{X}}_4'\mathbf{X}_2\boldsymbol{\alpha} + \tilde{\mathbf{X}}_4'\mathbf{X}_3\boldsymbol{\gamma} - \frac{\mu+\beta}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\mathbf{1} \\
&\qquad - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\mathbf{X}_1\boldsymbol{\rho} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\mathbf{X}_2\boldsymbol{\alpha} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{X}_3'\mathbf{X}_3\boldsymbol{\gamma} \\
&= (\mu+\beta)bk\mathbf{1}_{s-1} + k\mathbf{1}_{s-1}\mathbf{1}_b'\boldsymbol{\rho} + \mathbf{r}'\otimes\mathbf{1}_{s-1}\boldsymbol{\alpha} + \mathbf{1}_{s-1}\mathbf{1}_{bk}'\boldsymbol{\gamma} - \\
&\qquad \frac{\mu+\beta}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'s\mathbf{1}_{bk} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'s(\mathbf{1}_k\otimes\mathbf{I}_b)\boldsymbol{\rho} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'\mathbf{N}_2\boldsymbol{\alpha} - \frac{1}{s}\mathbf{1}_{s-1}\mathbf{1}_{bk}'s\mathbf{I}_{bk}\boldsymbol{\gamma}
\end{aligned}$$

$$= \mathbf{r}' \otimes \mathbf{1}_{s-1}\boldsymbol{\alpha} - \frac{1}{s}\mathbf{1}_{s-1}s\mathbf{r}'\boldsymbol{\alpha}$$
$$= \mathbf{0}$$

As a result, the non-centrality parameter is zero. Thus, $\mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S/\sigma^2 \sim \chi^2_{s-1}$ under H_0. Here, the degrees of freedom is equal to the rank of the matrix of the quadratic form $(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'$ and the rank of this matrix is clearly $s-1$.

To check independence of $\mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S$ and SSE, we know that $SSE = \mathbf{y}'(\mathbf{I}-\mathbf{X}\mathbf{G}\mathbf{X}')\mathbf{y}$ where $\mathbf{G}$ is a generalized inverse of $\mathbf{X}'\mathbf{X}$. Now,

$$(\mathbf{I}-\mathbf{X}\mathbf{G}\mathbf{X}')(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'$$
$$= (\tilde{\mathbf{X}}_4 - \frac{1}{s}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4 - \mathbf{X}\mathbf{G}\mathbf{X}'\tilde{\mathbf{X}}_4 + \frac{1}{s}\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)\mathbf{C}_s^{-1}(\tilde{\mathbf{X}}_4 - \frac{1}{k}\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4)'$$
$$= \mathbf{0}$$

because $\mathbf{X}\mathbf{G}\mathbf{X}'\tilde{\mathbf{X}}_4 = \tilde{\mathbf{X}}_4$ and $\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4 = \mathbf{X}_3\mathbf{X}_3'\tilde{\mathbf{X}}_4$ due to properties of generalized inverse matrix $\mathbf{G}$. Hence, two quadratic forms $\mathbf{Q}_S'\mathbf{C}_S^{-1}\mathbf{Q}_S$ and SSE are independent. Hence, F_2 under null hypothesis follows F-distribution with $(s-1)$ and $(bks-bk-ms+m)$ degrees of freedom. □ Proof of $F_3 \sim F_{(m-1),(bk-b-m+1)}$: where

$$SSW = R(\boldsymbol{\gamma}|\mu,\boldsymbol{\rho},\boldsymbol{\alpha}) = R(\mu,\boldsymbol{\rho},\boldsymbol{\alpha},\boldsymbol{\gamma}) - R(\mu,\boldsymbol{\rho},\boldsymbol{\alpha}) = \tfrac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} - \tfrac{1}{ks}\mathbf{y}_{B..}'\mathbf{y}_{B..} - \mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M.$$

First we prove that $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M/\sigma^2 \sim \chi^2_{m-1}$ under null hypothesis.

$$\begin{aligned}\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M &= (\tilde{\mathbf{y}}_{.M.} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{y}_{B..})'\mathbf{C}_M^{-1}(\tilde{\mathbf{y}}_{.M.} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{y}_{B..})\\ &= (\tilde{\mathbf{X}}_2'\mathbf{y} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{y})'\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2'\mathbf{y} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{y})\\ &= \mathbf{y}'(\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')\mathbf{y}\\ &= \mathbf{y}'\mathbf{A}\mathbf{y}\end{aligned}$$

where $\mathbf{A} = (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')$. Now,

$$\mathbf{A}\mathbf{A} = (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1').$$

It may be seen that

$$\begin{aligned}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1') &= \tilde{\mathbf{X}}_2'\tilde{\mathbf{X}}_2 - \frac{1}{k}\tilde{\mathbf{X}}_2'\mathbf{X}_1\tilde{\mathbf{N}}_1' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\tilde{\mathbf{X}}_2 + \frac{1}{k^2}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{X}_1\tilde{\mathbf{N}}_1'\\ &= \tilde{\mathbf{X}}_2'\tilde{\mathbf{X}}_2 - \frac{s}{k}\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1' - \frac{s}{k}\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1' + \frac{ks}{k^2}\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1'\\ &= s\mathbf{R}_{m-1} - \frac{s}{k}\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1'\\ &= \mathbf{C}_M.\end{aligned}$$

As a result,

$$\begin{aligned}\mathbf{AA} &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}\mathbf{C}_M\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')\\ &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')\\ &= \mathbf{A}\end{aligned}$$

Hence, the matrix of the quadratic form of $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M/\sigma^2$ is idempotent. Thus, $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M/\sigma^2 \sim \chi^2_{m-1}$ with non-centrality parameter $\frac{1}{2\sigma^2}\boldsymbol{\theta}\mathbf{X}'\mathbf{A}\mathbf{X}\boldsymbol{\theta}$ where $\mathbf{X}\boldsymbol{\theta} = \mu\mathbf{1} + \mathbf{X}\boldsymbol{\rho} + \mathbf{X}_2\boldsymbol{\alpha}$ since the rank of the matrix $\mathbf{A}$ is $m-1$. Under null-hypothesis, $\boldsymbol{\alpha} = \alpha\mathbf{1}_m$. So the non-centrality parameter can be shown to be zero as

$$\begin{aligned}\mathbf{AX}\boldsymbol{\theta} &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\mu\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho} + \alpha\mathbf{X}_2\mathbf{1}_m)\\ &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')\{(\mu+\alpha)\mathbf{1} + \mathbf{X}_1\boldsymbol{\rho}\}\\ &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}\left\{(\mu+\alpha)\tilde{\mathbf{X}}_2'\mathbf{1} + \tilde{\mathbf{X}}_2\mathbf{X}_1\boldsymbol{\rho} - \frac{\mu+\alpha}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{1} - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{X}_1\boldsymbol{\rho}\right\}\\ &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}\left\{(\mu+\alpha)s\mathbf{r}_{m-1} + s\tilde{\mathbf{N}}_1\boldsymbol{\rho} - \frac{\mu+\alpha}{k}\tilde{\mathbf{N}}_1ks\mathbf{1}_b - \frac{1}{k}\tilde{\mathbf{N}}_1ks\mathbf{I}_b\boldsymbol{\rho}\right\}\\ &= (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}\left\{(\mu+\alpha)s\mathbf{r}_{m-1} + s\tilde{\mathbf{N}}_1\boldsymbol{\rho} - (\mu+\alpha)s\mathbf{r}_{m-1} - s\tilde{\mathbf{N}}_1\boldsymbol{\rho}\right\}\\ &= \mathbf{0}.\end{aligned}$$

Thus, $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M/\sigma^2 \sim \chi^2_{m-1}$.

Now note that

$$\begin{aligned}SSW &= \frac{1}{s}\mathbf{y}_{BM.}'\mathbf{y}_{BM.} - \frac{1}{ks}\mathbf{y}_{B..}'\mathbf{y}_{B..} - \mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M\\ &= \frac{1}{s}\mathbf{y}'\mathbf{X}_3\mathbf{X}_3'\mathbf{y} - \frac{1}{ks}\mathbf{y}'\mathbf{X}_1\mathbf{X}_1'\mathbf{y} - \mathbf{y}'\mathbf{A}\mathbf{y}\\ &= \mathbf{y}'\mathbf{B}\mathbf{y}\end{aligned}$$

where $\mathbf{B} = \frac{1}{s}\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}\mathbf{X}_1\mathbf{X}_1' - \mathbf{A}$.

To check independence of $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M$ and SSW, we need to prove that $\mathbf{AVB} = \mathbf{0}$ where $\mathbf{A}, \mathbf{B}$ are as defined above and here $\mathbf{V} = \sigma^2\mathbf{I}$. So it suffices to show that $\mathbf{AB} = \mathbf{0}$. Now,

$$\mathbf{AB} = (\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\frac{1}{s}\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}\mathbf{X}_1\mathbf{X}_1' - \mathbf{A}).$$

The last two terms of $\mathbf{AB}$ may be simplified as

$$\begin{aligned}&(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\frac{1}{s}\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}\mathbf{X}_1\mathbf{X}_1' - \mathbf{A})\\ &= \frac{1}{s}\tilde{\mathbf{X}}_2'\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}\tilde{\mathbf{X}}_2'\mathbf{X}_1\mathbf{X}_1' - \tilde{\mathbf{X}}_2'\mathbf{A} - \frac{1}{ks}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{X}_3\mathbf{X}_3' + \frac{1}{k^2s}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{X}_1\mathbf{X}_1' + \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{A} \qquad (9)\\ &= \frac{1}{s}\tilde{\mathbf{X}}_2'\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}s\tilde{\mathbf{N}}_1\mathbf{X}_1' - \tilde{\mathbf{X}}_2'\mathbf{A} - \frac{1}{ks}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{X}_3\mathbf{X}_3' + \frac{1}{k^2s}ks\tilde{\mathbf{N}}_1\mathbf{X}_1' + \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{A}.\end{aligned}$$

It may be checked that

$$\begin{aligned}\tilde{\mathbf{X}}_2'\mathbf{A} &= (\tilde{\mathbf{X}}_2'\tilde{\mathbf{X}}_2 - \frac{1}{k}\tilde{\mathbf{X}}_2'\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1') \\ &= (s\mathbf{R}_{m-1} - \frac{1}{k}s\tilde{\mathbf{N}}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1') \\ &= \tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1'.\end{aligned}$$

Also,

$$\begin{aligned}\frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1'\mathbf{A} &= \frac{1}{k}\tilde{\mathbf{N}}_1(\mathbf{X}_1'\tilde{\mathbf{X}}_2 - \frac{1}{k}\mathbf{X}_1'\mathbf{X}_1\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1') \\ &= \frac{1}{k}\tilde{\mathbf{N}}_1(s\tilde{\mathbf{N}}_1' - \frac{1}{k}ks\tilde{\mathbf{N}}_1')\mathbf{C}_M^{-1}(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1') \\ &= \mathbf{0}.\end{aligned}$$

Hence, equation (9) can be simplified as

$$\begin{aligned}&(\tilde{\mathbf{X}}_2' - \frac{1}{k}\tilde{\mathbf{N}}_1\mathbf{X}_1')(\frac{1}{s}\mathbf{X}_3\mathbf{X}_3' - \frac{1}{ks}\mathbf{X}_1\mathbf{X}_1' - \mathbf{A}) \\ &= \frac{1}{s}\tilde{\mathbf{X}}_2'\mathbf{X}_3\mathbf{X}_3' - \tilde{\mathbf{X}}_2' + \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1' - \frac{1}{ks}\tilde{\mathbf{N}}_1'\mathbf{X}_1'\mathbf{X}_3\mathbf{X}_3' \\ &= \tilde{\mathbf{X}}_2' - \tilde{\mathbf{X}}_2' + \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1' - \frac{1}{k}\tilde{\mathbf{N}}_1'\mathbf{X}_1' \\ &= \mathbf{0}.\end{aligned}$$

Thus, two quadratic forms $\mathbf{Q}_M'\mathbf{C}_M^{-1}\mathbf{Q}_M$ and SSW are independent and hence, under null hypothesis, the test statistic F_3 follows F-distribution with $(m-1)$ and $(bk-b-m+1)$ degrees of freedom. Null hypothesis should be rejected whenever calculated value of F_3 exceeds $F_{\alpha,(m-1),(bk-b-m+1)}$ where $F_{\alpha,(m-1),(bk-b-m+1)}$ denotes the upper α percent point of an F-distribution with $(m-1)$ and $(bk-b-m+1)$ degrees of freedom.