P.ISSN: 0971-8184 II E ISSN: 0976-1926
DOI: https://doi.org/10.56093/ijpgr.v39i3.185048

Indian Journal of Plant Genetic Resources

RESEARCH ARTICLE

Resource Efficient pRep Block Designs for Comparing Test Lines vs. Control Line(s) in Early Generation Varietal Trials

LN Vinaykumar[1], Cini Varghese[1]*, Mohd Harun[1], Sayantani Karmakar[2], Eldho Varghese[3] and Seema Jaggi[4]

Abstract

Efficient resource use is vital in early generation varietal trials (EGVTs), where replicating test lines across environments is often constrained. Partially replicated (pRep) designs address this by allowing a selective number of replications of test lines. This study introduces two series of pRep designs for EGVTs, facilitating comparisons between test and control line(s). The general form of the information matrix for these designs is derived to evaluate their properties, and SAS code was developed to calculate efficiency and variance factors. The proposed designs offer a flexible and effective framework for conducting EGVTs with fewer replications, enhancing their applicability across various environments.

Keywords: Early generation varietal trials, Information matrix, Multi-environmental trials, Partially replicated designs

[1]ICAR-Indian Agricultural Statistics Research Institute, PUSA, New Delhi, 110012, India

[2]ICAR-National Institute of Abiotic Stress Management, Baramati, Pune, Maharashtra, 413115, India

[3]Fishery Resources Assessment, Economics and Extension Division, ICAR-Central Marine Fisheries Research Institute, Kochi, Kerala, 682018, India

[4]Krishi Anusandhan Bhavan, ICAR, New Delhi, 110012, India

Received: 04/08/2025

Revised: 14/10/2026

Accepted: 15/10/2026

***Author for correspondence:**
cini2204@gmail.com

How to cite this article: LN Vinaykumar, Cini Varghese, Mohd Harun, Sayantani Karmakar, Eldho Varghese and Seema Jaggi (2026) Resource Efficient pRep Block Designs for Comparing Test Lines vs. Control Line(s) in Early Generation Varietal Trials. *Indian J. Plant Genet. Resour.* 39(3): 61-69.

DOI: https://doi.org/10.56093/ijpgr.v39i3.185048

Introduction

Most plant breeding programmes follow a method of breeding referred to as a pedigree selection method. This method implies that programmes are structured around the grouping of breeding lines which have been derived as progeny of a fixed number of crosses between elite parents. Different crosses are made each year, and the cohort of breeding lines then undergoes selection through preliminary and advanced stages of testing. Traits of interest for selection in the preliminary stage include disease and herbicide tolerance, phenology type and functional quality of yield. Typically, selection in the advanced stages occurs in a sequential manner. The key selection trait in the advanced stages is yield. For example, correlation and path analysis in mothbean (*Vigna aconitifolia*) identified yield-associated traits, including pods per plant and harvest index, as useful selection criteria for improving seed yield (Meena *et al*., 2026). Yield data for each stage are generated from a series of field trials across several environments. Trials typically include the set of breeding lines of interest and some control lines, which will collectively be termed entries.

In general, varietal trials are classified as early generation varietal trials (EGVTs) and late generation varietal trials (LGVTs). The number of entries that are evaluated at the EGVTs often exceeds 1,000, while at LGVTs the number of entries tested is generally less than 100. In EGVTs, resource and seed limitations reduce the number of environments and plots that can be used for each breeding line. Replication within an environment is

© IJPGR, 2025. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit https://creativecommons.org/licenses/by-nc-sa/4.0/.

often less than two, and the number of environments is usually less than four. For LGVTs entries are evaluated at up to ten environments, with between two and three replicates per environment. The different stages of breeding trials and the key selection criteria of plant breeding programmes are summarized in Fig. 1.

When complete replication of every line in every environment is unfeasible, researchers advocate for grid-plot designs (Kempton, 1984). These designs involve replicating well-known commercial varieties, check varieties, and un-replicated test lines in row-column arrangements. As an alternative to grid-plot designs, resolved row-column designs, known as partially replicated (pRep) designs, were introduced to enhance the accuracy of selecting breeding lines compared to traditional designs (Cullis *et al.,* 2006). The pRep designs are more efficient in addressing the challenges associated with EGVTs conducted across multiple environments, particularly when the number of plots per breeding line is fewer than two. A construction method for pRep designs has been outlined, which is based on eliminating entries from α-arrays (Williams *et al.*, 2011; Williams *et al.*, 2014). Moreover, the adoption of pRep designs in plant breeding experiments has been widely discussed (Clarke and Stefanova, 2011; Moehring *et al.*, 2014; Sermarini *et al.*, 2020).

Additionally, Partially Balanced Incomplete Block (PBIB) designs and their association schemes (Bose and Nair, 1939) are commonly used in constructing designs suitable for breeding experiments. Harun *et al.*, (2016) developed three-way cross designs for comparing test lines against control lines using PBIB designs and their association scheme. Recently, PBIB designs have also been explored for constructing mating environmental designs and enhancing pRep designs (Vinaykumar *et al.*, 2023; Vinaykumar *et al.*, 2026). Furthermore, a new R package called "pRepDesigns" has been developed by

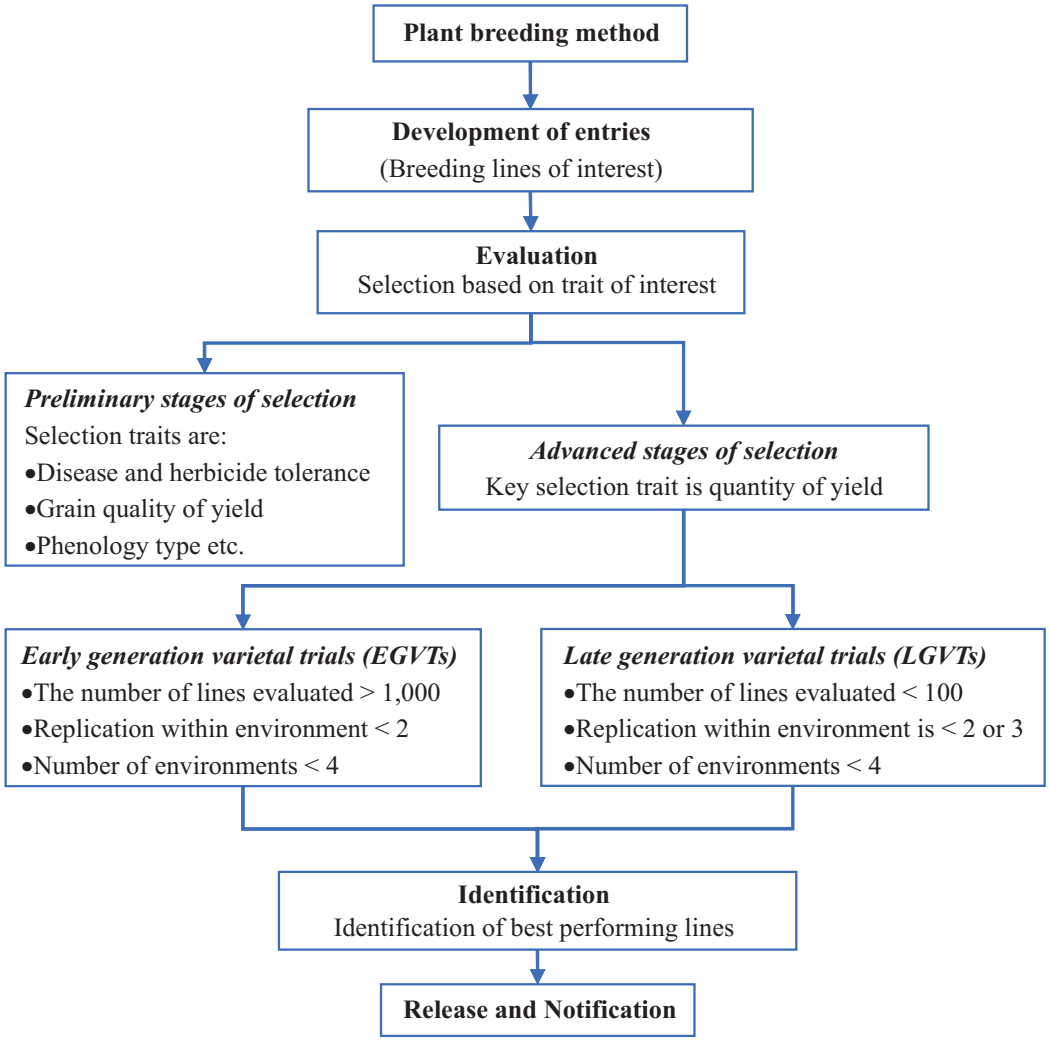


Fig. 1: Major stages of plant breeding programme

Vinaykumar *et al.*, (2024) to facilitate the easy generation of pRep designs for various lines and environmental structures. These designs empower breeders to comprehend environmental variations in breeding lines during data analysis. Recently, Santos *et al.*, (2024) conducted a simulation study and concluded that pRep designs were more efficient than grid-plot designs, showing higher relative realized genetic gain, accuracy, and selection probability.

In comparative breeding experiments, a control line is often established, and the primary objective is to compare new lines (or test lines) against control line(s) to identify that worth further study, rather than comparing new lines to select the best genotype. Efficient experimental designs are therefore essential for maximizing the use of limited seed material and experimental resources. The relative efficiency of different experimental designs for evaluating plant genetic resources has been investigated by Gangopadhyay *et al.*, (2010). The efficiency of such experiments can be improved by replicating control lines more frequently than test lines, as in partially replicated (pRep) designs. Availability of pRep designs for comparing test lines with control line(s) in EGVTs is limited. Such designs are henceforth referred to as pRepTC designs. This paper presents two new series of pRep designs tailored for such comparisons. The information matrix is derived to examine the properties of these designs, and SAS code has also been developed to calculate their efficiency and variance factors.

Material and Methods

In this section, we describe the experimental conditions in which pRepTC designs can be implemented. pRep designs are defined along with the experimental setup and model. The model for pRepTC designs remains the same as that of pRep designs.

Experimental situations

Situation 1: There may be experimental situations where several new lines are developed in the initial stage of an experiment. Only a few of the new lines are expected to be investigated further. The new lines may first be compared with a control line already being used by the experimenter to screen out the best number of new lines for further investigation. In such cases, the experimenter would like to compare test lines with control lines with as much precision as possible, whereas comparisons within test lines may be of lesser interest.

Situation 2: In some experiments, an experimenter may have two groups of parental lines, one group containing lines that are of primary importance and the other group containing lines of secondary interest. It may always be desirable to have superior or well-adapted lines represented in a high proportion of crosses in breeding experiments.

Situation 3: The simplest case is a nested structure of entries, where entries fall into groups or families such that entries within a group are more closely related than between groups. In such cases, the experimenter would like to make a comparison between groups of close relatives with greater precision, whereas comparison within groups of close relatives may not be of much interest.

pRep designs for comparing test lines with control line(s): A pRepTC design is an arrangement of $v = v_t + v_c$ (where v_t and v_c represent the number of test lines and control line(s), respectively) lines in b blocks, each of size k_w (<v), where $1 \leq w \leq b$ distributed over e environments such that:

- each environment consists of $\frac{b}{e}$ blocks
- test lines are replicated (e + 1) times whereas control lines are replicated more [> (e + 1)]
- e distinct sets of $\frac{v_t}{e}$ test lines are formed from v_t lines and in each environment, a distinct set of $\frac{v_t}{e}$ test lines are replicated twice and the rest $(1-\frac{v}{e})$ v_t lines occur once.

For example, in a maize breeding trial with 602 hybrid lines, consisting of 600 newly developed test lines and 2 existing control lines, these lines are evaluated across three distinct environments. Using pRepTC designs, a proportion (p %) of the test lines were replicated. If p = 30%, then 180 test lines had two replicates across the environments, while the remaining 400 test lines, along with the 2 control lines, each had a single replication.

Model and experimental setup: To analyse the data obtained using a pRep design, the following nested model can be considered:

$$y_{ij(l)} = \mu + \tau_i + \eta_l + \beta_{j(l)} + \varepsilon_{ij(l)}$$

where, $y_{ij(l)}$ is the response of i^{th} line (i=1, 2, ..., v) belonging to j^{th} block (j=1, 2, ..., b), nested within 1^{th} environment (1=1, 2, ..., e), μ is the general mean, τ_i is the effect of i^{th} line, η_l is the effect of 1^{th} environment, $\beta_{j(l)}$ is the effect of j^{th} block in l^{th} environment and $\varepsilon_{ij(l)}$ is the random error component which follows iid $N(0_2\sigma^2)$.

Results and Discussion

The construction procedure for obtaining pRepTC designs is outlined in the following methods.

Method I: The steps involved in this method are as follows:

Step 1: Let v = p(p - 1), where p ⩾ 3, and L = {1, 2, 3, …, v} is the set of lines. Consider two subsets of L, viz., R = {r_h| h = 1, 2, ..., (p- 1)}, where r_h = (h - 1)p + 1, (h - 1)p + 2, …, hp, and S = {s_g|g=1,2,…, p}, where s_g = g, g + p, g + 2p, ..., g + p(p-2). Clearly, R∩S = L, n(R) =n(S) =p(p-1), n(r_h) =p, and n(s_h) =p - 1, where n(.) denotes cardinality of set. For a line θ∈L, θ∈r_h, or s_g.

Step 2: Form the first replication (R1) with b_1 = p - 1 blocks, each of size k_1 = p, by considering r_h as the h^{th} block. Similarly, form the second replication (R2) with b_2 = p blocks, each of size k_2 = p - 1, by considering s_g as the g^{th} block. This results in a series of three associate resolvable PBIB designs.

Step 3: Divide the b_1 blocks of R1 into e sets (e≤p) of blocks and associate them with the b_2 blocks of the R2, or vice versa. Here, e sets are considered as environments.

Step 4: Reinforce the control lines $(v+1)^{th}$, $(v+2)^{th}$, ..., $(v+p)^{th}$ into the smaller blocks of 1^{st}, 2^{nd}, …, p^{th} environments respectively.

Here, the design obtained through Steps 1-2 is a resolvable PBIB(3) design. However, there is a constant number of test lines shared between any two blocks from different sets, defining these designs as affine resolvable; hence they are optimal (Bailey *et al.*, 1995). Therefore, these designs emerge as the superior choice among competing classes of resolvable designs.

Example 1: Using Steps 1 and 2, a layout of a resolvable three-associate PBIB design for 20 (i.e., p = 5) test lines in 9 blocks is obtained as follows:

Replication 1				Replication 2				
B1	B2	B3	B4	B5	B6	B7	B8	B9
1	6	11	16	1	2	3	4	5
2	7	12	17	6	7	8	9	10
3	8	13	18	11	12	13	14	15
4	9	14	19	16	17	18	19	20
5	10	15	20					

Following Step 3, the layout of the pRep design for two environments (e = 2) is obtained as:

Environment 1							Environment 2						
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14
1	6	1	2	3	4	5	11	16	1	2	3	4	5
2	7	6	7	8	9	10	12	17	6	7	8	9	10
3	8	11	12	13	14	15	13	18	11	12	13	14	15
4	9	16	17	18	19	20	14	19	16	17	18	19	20
5	10						15	20					

The control lines, 21 and 22, are then augmented to the smaller blocks of Environment 1 and Environment 2, respectively, to obtain the layout of the pRepTC design, as shown below:

Environment 1							Environment 2						
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14
1	6	1	2	3	4	5	11	16	1	2	3	4	5
2	7	6	7	8	9	10	12	17	6	7	8	9	10
3	8	11	12	13	14	15	13	18	11	12	13	14	15
4	9	16	17	18	19	20	14	19	16	17	18	19	20
5	10	**21**	**21**	**21**	**21**	**21**	15	20	**22**	**22**	**22**	**22**	**22**

The parameters of resulting design are: v_t = 20, v_c = 2, b = 14, r_t = 3, r_c = 5, k = 5, where v_t and v_c are the number of test and control lines with replications r_t and r_c, respectively. The canonical efficiency factor (Dey, 1995) and average variance factor of this design have been computed as 0.7923 and 0.6678, respectively, using SAS code (given in APPENDIX II).

Example 2: For p = 6, the layout of the pRepTC design for 33 lines (v_t = 30 test lines and v_c = 3 control lines) in 21 blocks is presented below:

Environment 1

B1	B2	B3	B4	B5	B6	B7
1	7	13	19	25	1	2
2	8	14	20	26	7	8
3	9	15	21	27	13	14
4	10	16	22	28	19	20
5	11	17	23	29	25	26
6	12	18	24	30	**31**	**31**

Environment 2

B8	B9	B10	B11	B12	B13	B14
1	7	13	19	25	3	4
2	8	14	20	26	9	10
3	9	15	21	27	15	16
4	10	16	22	28	21	22
5	11	17	23	29	27	28
6	12	18	24	30	**32**	**32**

Environment 3

B15	B16	B17	B18	B19	B20	B21
1	7	13	19	25	5	6
2	8	14	20	26	11	12
3	9	15	21	27	17	18
4	10	16	22	28	23	24
5	11	17	23	29	29	30
6	12	18	24	30	**33**	**33**

Method II: The steps involved in this method are as follows:

Step1: Consider $v = s^2$ lines, where $s \geq 5$ is an odd prime or prime power. Construct s - 1 arrays of size s×s as follows: Start with the 1st row in natural order. Keep the 1st entry unchanged, and add increments of 1, 2, ..., s - 1 to the subsequent entries (mod s). Develop these rows cyclically (mod s) to form s - 1 square arrays.

Step 2: Now, arrange the s^2 lines in an s×s array and superimpose it on s - 1 arrays of Step 1. Then, form s - 1 new arrays by considering the line numbers common to a particular s^{th} symbol as the entries in the respective s^{th} column.

Step 3: From the arrays of Step 2, delete the entries of s^{th} row, then delete the 1^{st}, 2^{nd}, ..., $(s-1)^{th}$ row from 1^{st}, 2^{nd}, .., $(s-1)^{th}$ array, respectively. Assign the first two arrays to Environment 1, the next two arrays to Environment 2, and so on. These columns form the blocks of a pRep design with parameters: v = s(s-1), b = s(s-1), r = 3, k = s-2.

Step 4: Replace the higher-numbered test lines with control line(s).

Example 2: For s = 5, the four arrays of order 5×5 are constructed by cyclically developing the initial rows: (1, 2, 3, 4, 5), (1, 3, 5, 2, 4), (1, 4, 2, 5, 3) and (1, 5, 4, 3, 2). Following Step 2, these arrays are then superimposed onto a 5×5 arrangement of 25 lines to generate four new arrays. The resulting arrays are given below:

Array I					Array II				
1	2	3	4	5	1	4	2	5	3
10	6	7	8	9	8	6	9	7	10
14	15	11	12	13	15	13	11	14	12
18	19	20	16	17	17	20	18	16	19
22	23	24	25	21	24	22	25	23	21

Array III					Array IV				
1	3	5	2	4	1	5	4	3	2
9	6	8	10	7	7	6	10	9	8
12	14	11	13	15	13	12	11	15	14
20	17	19	16	18	19	18	17	16	20
23	25	22	24	21	25	24	23	22	21

The last row from each array is deleted, followed by the deletion of the 1^{st}, 2^{nd}, 3^{rd}, and 4^{th} rows from each respective array. Then considering columns as blocks, first two arrays to E1 and other two arrays to E2 are associated to obtain the test lines layout of pRep design with parameters v = 20, b = 20, r = 3, k = 3 as shown below:

Environment 1

B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
10	6	7	8	9	1	4	2	5	3
14	15	11	12	13	15	13	11	14	12
18	19	20	16	17	17	20	18	16	19

Environment 2

B11	B12	B13	B14	B15	B16	B17	B18	B19	B20
1	3	5	2	4	1	5	4	3	2
9	6	8	10	7	7	6	10	9	8
20	17	19	16	18	13	12	11	15	14

Finally, the layout of pRepTC design in two environments is obtained by replacing test lines 18, 19 and 20 with control line 18 as shown below:

Environment 1

B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
10	6	7	8	9	1	4	2	5	3
14	15	11	12	13	15	13	11	14	12
18	**18**	**18**	16	17	17	**18**	**18**	16	**18**

Environment 2

B11	B12	B13	B14	B15	B16	B17	B18	B19	B20
1	3	5	2	4	1	5	4	3	2
9	6	8	10	7	7	6	10	9	8
18	17	**18**	16	**18**	13	12	11	15	14

The parameters of resulting designs are: $v_t = 17$, $v_c = 1$, b = s(s-1), $r_t = 3$, $r_c = 9$, k = s-2. Further, in the following section, the information matrix required for analysing the characterization properties of the proposed designs has been computed.

Information matrix: The information matrix (**C**) for a pRepTC design framework (simplification given in APPENDIX I) is obtained as:

$$\mathbf{C} = \begin{bmatrix} r_t\mathbf{I}_t - \frac{r_t^2}{n}\mathbf{J}_t - \mathbf{N}_t\mathbf{M}\mathbf{N}_t' & -\frac{r_t r_c}{n}\mathbf{J}_{t\times c} - \mathbf{N}_t\mathbf{M}\mathbf{N}_c' \\ -\frac{r_t r_c}{n}\mathbf{J}_{c\times t} - \mathbf{N}_c\mathbf{M}\mathbf{N}_t' & r_c\mathbf{I}_c - \frac{r_c^2}{n}\mathbf{J}_c - \mathbf{N}_c\mathbf{M}\mathbf{N}_c' \end{bmatrix}$$

where $\mathbf{M} = \left[\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right]$, N_t and N_c are incidence matrices of test and control lines, respectively. Here, r_t and r_c represent the replication of the test and control lines, respectively. **I** and **J** denote the identity and unit matrices of appropriate orders. The reduced normal equations pertaining to line effects under the above model is $\mathbf{C}\hat{\tau} = \mathbf{Q}$ with $\hat{\tau} = 0$, where **Q** is the vector of adjusted line totals. The normal equations can be solved

to get an estimate of line effect $\hat{\tau} = \mathbf{C}^{-}\mathbf{Q}$, where $\mathbf{C}^{-}$ is the generalized inverse of **C** and $\mathbf{p}'\hat{\tau}$ be an estimable contrast of line effects such that $\mathbf{p}' = 0$. The best linear unbiased estimator of $\mathbf{p}'\tau$ s $\mathbf{p}'\hat{\tau} = \mathbf{p}'\mathbf{C}$–Q with variance $V(\mathbf{p}'\hat{\tau}) = \sigma^2\mathbf{p}'\mathbf{C}^{-}\mathbf{p}$.

Conclusion

This study introduces two innovative series of partially replicated (pRep) designs for EGVTs, focusing on comparisons between test lines and control line(s). Method I designs are constructed by reinforcing control lines into three-associate-class affine-resolvable PBIB designs, which are optimal within the considered class of resolvable PBIB designs. Method II designs are developed by replacing selected test lines with control line(s) in the constructed pRep designs using a systematic array-based procedure. SAS code was implemented to compute canonical efficiency factors and average variance factors for evaluating the statistical properties of the proposed designs.

The two series offer alternative construction strategies to accommodate different experimental requirements. Method I allows the allocation of blocks of different sizes across environments, whereas Method II provides a systematic construction with a prescribed block structure. Therefore, Method I designs are suitable for locations with uneven terrain, where maintaining equal block sizes may not be feasible, whereas Method II designs are preferable when sufficient resources are available to accommodate equal block sizes. Furthermore, the choice between the methods depends on the required number of test and control lines, block size, environmental allocation, and the resulting statistical efficiency and variance factors. Overall, the proposed designs provide resource-efficient alternatives for EGVTs by reducing the need for extensive replication of test lines while facilitating comparisons with control line(s) across environments. Their systematic construction and SAS-based evaluation support their practical application in early-stage varietal evaluation under resource constraints.

Acknowledgements

The first author gratefully acknowledges the Graduate School, ICAR-Indian Agricultural Research Institute, and ICAR-Indian Agricultural Statistics Research Institute for providing the necessary support and facilities to carry out this research.

References

Bose RC and KR Nair (1939) Partially balanced incomplete block designs. *Sankhya* 4(3): 337-372.

Bailey RA, H Monod and J Morgan (1995) Construction and optimality of affine-resolvable designs. *Biometrika* 82(1): 187-200. https://doi.org/10.2307/2337638

Clarke GPY and KT Stefanova (2011) Optimal design for early-generation plant-breeding trials with unreplicated or partially replicated test lines. *Aust. N. Z. J. Stat.* 53(4): 461-480. https://doi.org/10.1111/j.1467-842X.2011.00642.x

Cullis BR, AB Smith and NE Coombes (2006) On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11(4): 381-393.

Dey A (2008) Canonical efficiency factors and related issues revisited. *Indian J. Agric. Sci.* 62(2): 169-173.

Dos Santos DP, RA Sermarini, A dos Santos and CGB Demétrio (2024) Optimal designs in plant breeding experiments; a simulation study comparing grid-plot and partially replicated (p-rep) design. *Sugar Tech* 26(2): 387-395. https://doi.org/10.1007/s12355-024-01375-3

Gangopadhyay KK, RK Mahajan, R Prasad, G Kumar, BL Meena, R Kar and SK Mishra (2010) Relative efficiency of experimental designs in evaluation of plant genetic resources. *Indian J. Plant Genet. Resour.* 23(2): 164-167.

Harun M, C Varghese, E Varghese and S Jaggi (2016) Three-way cross designs for test lines vs. control line comparisons. *Electron. J. Plant Breed.* 7(1): 42-47

Kempton RA (1984) The design and analysis of unreplicated field trials. *Vortraege fuer Pflanzenzuechtung* (Germany) 7: 219-42.

Meena S, NK Sharma, P Yadav, R Chawla and AK Meena (2026) Correlation and path analysis for seed yield and components traits in mothbean [*Vigna aconitifolia* (Jacq) Marechal] under arid environment. *Indian J. Plant Genet. Resour.* 39(2): 105-113. https://doi.org/10.56093/ijpgr.v39i2.175071

Moehring J, ER Williams and HP Piepho (2014) Efficiency of augmented p-rep designs in multi-environmental trials. *Theor. Appl. Genet.* 127(5):1049-1060. https://doi.org/10.1007/ s00122-014-2278-y

Searle SR (1982) *Matrix Algebra Useful for Statistics.* John Wiley & Sons, New York, USA.

Sermarini RA, C Brien, CGB Demétrio and A dos Santos (2020) Impact on genetic gain from using misspecified statistical models in generating p-rep designs for early generation plant-breeding experiments. *Crop Sci.* 60(6): 3083-3095. https://doi.org/10.1002/csc2.20257

Vinaykumar LN, C Varghese, M Harun and S Karmakar (2023) Minimally replicated PBIB designs for multi-environmental trials. *Commun. Stat. Theory Methods* 53(13): 4696-4716. https://doi.org/10.1080/03610926.2023.2185753

Vinaykumar LN, C Varghese, M Harun, A Dalal, S Karmakar and Vinayaka (2024) *pRepDesigns: Partially replicated (p-Rep) designs*. Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=pRepDesigns

Vinaykumar LN, C Varghese, M Harun, S Karmakar, E Varghese, Vinayaka and S Jaggi (2026) Efficient partially replicated designs for multi-environment early-generation breeding trials. *Sugar Tech* 28(3): 874-884. https://doi.org/10.1007/ s12355-025-01684-1

Williams E, HP Piepho and D Whitaker (2011) Augmented p-rep designs. *Biometr. J.*53(1): 19-27. https://doi.org/10.1002/ bimj.201000102

Williams ER, JA John and D Whitaker (2014) Construction of more flexible and efficient p-rep designs. *Aust. N. Z. J. Stat.* 56(1): 89-96. https://doi.org/10.1111/anzs.12068

APPENDIX - I

Derivation of Information Matrix (C)

The simple block design model in matrix notation is:

$$\mathbf{y} = \mathbf{X}\theta + \varepsilon$$

where, X = [Δ' 1' D'], $\theta = \begin{bmatrix} \tau \\ \beta \\ \mu \end{bmatrix}$, **y** is the (n x 1) vector of observations, μ is the general mean, τ is the (v x 1) vector of line effects, β is the (b x 1) vector of block effects, **1** is the (n x 1) vector of unities, Δ_1 and Δ_2 are the (n x v) observations vs. lines and (n x b) observations vs. blocks design matrices, respectively and ε is the (n x 1) vector of errors following iid N(0, $\sigma^2\mathbf{I}$), where, I is the identity matrix of order n.

Let:

$$\mathbf{X}_1 = [\Delta'], \mathbf{X}_2 = [D\ 1], \theta_1 = (\tau), \theta_2 = [\beta\ \mu]'$$

Here, θ_1 represents the vector of the parameter of interest, and θ_2 represents the vector of nuisance parameters.

The matrix **X' X** is: $\mathbf{X}'\mathbf{X} = \begin{bmatrix} \Delta\Delta' & \Delta\mathbf{1} & \Delta\mathbf{D}' \\ \mathbf{1}'\Delta' & \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{D}' \\ \mathbf{D}\Delta' & \mathbf{D}\mathbf{1} & \mathbf{D}\mathbf{D}' \end{bmatrix}$

where, $\Delta\Delta' = \begin{bmatrix} r_t\mathbf{I}_t & 0_{t\times c} \\ 0_{c\times t} & r_c\mathbf{I}_c \end{bmatrix}_{v\times v}$, $\Delta\mathbf{1} = \begin{bmatrix} r_t\mathbf{1}_t \\ r_t\mathbf{1}_c \end{bmatrix}_{(t+c)\times 1}$,

$\Delta\mathbf{D}' = \mathbf{N}_{1(v\times b)}$, $\mathbf{1}\mathbf{1}' = n$, $\mathbf{1}'\mathbf{D}' = k\mathbf{1}_b'$, $\mathbf{D}\mathbf{1} = k\mathbf{1}_b$ and

$\mathbf{D}\mathbf{D}' = k\mathbf{I}_b$

The general form of the information matrix (**C**) for the line effects is:

$$\mathbf{C} = \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1$$

For proposed designs: $(\mathbf{X}_2'\mathbf{X}_2) = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{D}' \\ \mathbf{D}\mathbf{1} & \mathbf{D}\mathbf{D}' \end{bmatrix}$ is not full rank, i.e., r($\mathbf{X}_2'\mathbf{X}_2$) = b. Let, $(\mathbf{X}_2'\mathbf{X}_2) = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, then,

$\mathbf{A}_{21}\mathbf{A}_{12} = \mathbf{D}\mathbf{1}\mathbf{1}'\mathbf{D}' = k^2\mathbf{J}$, $\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} = \frac{k^2\mathbf{J}}{n}$, $\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} = (b-1)$, $r(A_{11}) = 1$ and $r(A_{11}) + r(A_{22} - A_{21}A_{11}^{-1}A_{12}) = r(X'_2X_2) = b$. As the rank condition $r(X'_2X_2) = r(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})$, is satisfied, by following Searle (1982), the generalized inverse of (X'_2X_2) obtained through the method of partitioning. Let

$$(\mathbf{X}_2'\mathbf{X}_2)^- = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}$$

where, $\mathbf{B}_{11} = \mathbf{A}_{11}^{-1} + (\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\xi^-(\mathbf{A}_{21}\mathbf{A}_{11}^{-1}) = \frac{1}{n}$, $\mathbf{B}_{12} = -(\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\xi^- = \mathbf{0}_{1\times b}$, $\mathbf{B}_{21} = -\xi^-(\mathbf{A}_{21}\mathbf{A}_{11}^{-1}) = \mathbf{0}_b'$, and $\mathbf{B}_{22} = \xi^+$ is the Moore-Penrose inverse of $\xi = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ obtained using full rank factorization (Searle, 1982) as:

$\xi^+ = \begin{bmatrix} \left(\frac{1}{k}\right)\mathbf{I}_{b-1} - \left(\frac{1}{kb}\right)\mathbf{J}_{b-1} & -\left(\frac{1}{kb}\right)\mathbf{1}_{b-1} \\ -\left(\frac{1}{k}\right)\mathbf{1}_{b-1}' + \left(\frac{1}{kb}\right)(b-1)\mathbf{1}_{b-1}' & \left(\frac{1}{kb}\right)\ b\text{-}1 \end{bmatrix}$, where $\xi = \begin{bmatrix} k\mathbf{I}_{b-1} - \left(\frac{k}{b}\right)\mathbf{J}_{b-1} & -\left(\frac{k}{b}\right)\mathbf{1}_{b-1} \\ -\left(\frac{k}{b}\right)\mathbf{1}_{b-1}' & \frac{k(b-1)}{b} \end{bmatrix}$

Now, on simplification: $(\mathbf{X}_2'\mathbf{X}_2)^- = \begin{bmatrix} \frac{1}{n} & \mathbf{0}_b' \\ \mathbf{0}_b & \xi^+ \end{bmatrix}$, and

$$\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1 =$$

$$\begin{bmatrix} \frac{r_t^2}{n}\mathbf{J}_t + \mathbf{N}_t\left(\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right)\mathbf{N}_t' & \frac{r_tr_c}{n}\mathbf{J}_{t\times c} + \mathbf{N}_t\left(\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right)\mathbf{N}_c' \\ \frac{r_tr_c}{n}\mathbf{J}_{c\times t} + \mathbf{N}_c\left(\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right)\mathbf{N}_t' & \frac{r_c^2}{n}\mathbf{J}_c + \mathbf{N}_c\left(\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right)\mathbf{N}_c' \end{bmatrix}$$

Hence, the information matrix of a pRepTC design framework is:

$$\mathbf{C} = \begin{bmatrix} r_t\mathbf{I}_t - \frac{r_t^2}{n}\mathbf{J}_t - \mathbf{N}_t\mathbf{M}\mathbf{N}_t' & -\frac{r_tr_c}{n}\mathbf{J}_{t\times c} - \mathbf{N}_t\mathbf{M}\mathbf{N}_c' \\ -\frac{r_tr_c}{n}\mathbf{J}_{c\times t} - \mathbf{N}_c\mathbf{M}\mathbf{N}_t' & r_c\mathbf{I}_c - \frac{r_c^2}{n}\mathbf{J}_c - \mathbf{N}_c\mathbf{M}\mathbf{N}_c' \end{bmatrix}$$

where, $\mathbf{M} = \left[\frac{1}{k}\mathbf{I}_b - \frac{1}{kb}\mathbf{J}_b\right]$ N_t and N_c are incidence matrices of test lines and control lines, respectively.

APPENDIX - II

SAS programme to generate information matrix, compute canonical efficiency factor and average variance factor of a pRepTC design

```
%let t = 20;/* Enter number of test lines*/
proc iml;
a={
1 10 14 22,
2 6 15 23,
3 7 11 24,
4 8 12 25,
5 9 13 21
};
print a;
block_rep={5,5};
a1=countn(loc(a));
m=j(a1,1,1);/*mean vector*/
trt=j(a1,max(a),0);/*design matrix -obs VS direct treatment*/
k=1;
do i=1 to nrow(a);
do j=1 to ncol(a);
if a[i,j]>0 then
                do;
                trt[k,a[i,j]]=1;
                k=k+1;
                end;
end;
end;
*print trt;
rep=j(a1,nrow(block_rep),0);/*design matrix -obs VS rep*/
k=1;
kk=0;
do i=1 to nrow(block_rep);
do j=1 to block_rep[i,];
do l=1 to ncol(a);
if a[kk+j,l]>0 then do;
                rep[k,i]=1;
                k=k+1;
                end;
                end;
                end;
kk=kk+block_rep[i,];
end;
Block_withinrep=j(a1,max(block_rep),0);/*design matrix - obs VS Blocks with in rep*/
k=1;
kk=0;
do i=1 to nrow(block_rep);
do j=1 to block_rep[i,];
do l=1 to ncol(a);
if a[kk+j,l]>0 then do;
                Block_withinrep[k,j]=1;
                k=k+1;
                end;
                end;
                end;
kk=kk+block_rep[i,];
end;
Block=j(a1,nrow(a),0);/*design matrix - obs VS Blocks*/
k=1;
do i=1 to nrow(a);
do j=1 to ncol(a);
if a[i,j]>0 then do;
                Block[k,i]=1;
                k=k+1;
                end;
                end;
end;
```

```
x=m||trt||rep||Block;/*design matrix*/
x1=trt;
x2=m||Block;
c_mat=(x1`*x1)-(x1`*x2*(ginv(x2`*x2))*x2`*x1)/*C matrix*/;
print c_mat; /*Information matrix*/
t1=comb(&t,2);/*total no of test vs test treatments*/
cot=j(t1,Max(a),0);/*for var_test*/
k=1;
do i=1 to &t-1;
do j=i+1 to &t;
cot[k,i]=1;
cot[k,j]=-1;
k=k+1;
end;
end;
*print cot;

covt=cot*(ginv(c_mat))*cot`;
vart1=diag(covt);
onet=j(t1,1,1);
var_test=vart1*onet;
print var_test;

tot=comb(max(a),2);/*total pairs of treatments*/
if (max(a)-&t)>1 then do;
                cc=comb((max(a)-&t),2);
                end;
                else do;
                cc=0;
                end;
tc=tot-(t1+cc);
cotc=j(tc,max(a),0);
k=1;
do i=1 to &t;
do j=&t+1 to max(a);
cotc[k,i]=1;
cotc[k,j]=-1;
k=k+1;
end;
end;
covtc=cotc*ginv(c_mat)*cotc`;
vartc=diag(covtc);
ones_tc=j(tc,1,1);
var_tc=vartc*ones_tc;
print var_tc;
av_var_test=var_test[+,]/nrow(var_test);
av_var_tc=var_tc[+,]/nrow(var_tc);
print av_var_test av_var_tc;
eig=eigval(c_mat);
eig1=eig[loc(eig>0.0000001),];/*positive eigen values*/
rep1=trt`*trt;
eig2=eig1/(rep1[1,1]);
eig3=1/eig2;
CanEffFactor=nrow(eig3)/sum(eig3);
print CanEffFactor;
quit;
```