Experimental designs for the selection of integrated farming system components

Sayantani Karmakar¹, Cini Varghese²*, Mohd Harun², L. N. Vinaykumar¹ and Eldho Varghese³

¹Post Graduate School, ICAR-Indian Agricultural Research Institute, Pusa, New Delhi 110 012, India
²Division of Design of Experiments, ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi 110 012, India
³Division of Fishery Resources Assessment, Economics and Extension, ICAR-Central Marine Fisheries Research Institute, Kochi 682 018, India

Integrated farming system (IFS) approach is a powerful tool for ensuring the livelihood security of small and marginal farmers. The precision of IFS experiments can be enhanced using statistical and computational tools. Two-part designs are helpful in selecting the best possible components in IFS. They involve two groups of treatment arranged in incomplete blocks with respect to both groups, and the concurrence of treatment pairs within and between groups is constant. The fusion of two incomplete block designs in a systematic manner can yield two-part designs. Further, for situations where certain experimental units are not available, two-part structurally incomplete designs are proposed.

Keywords: Block designs, integrated farming system, livelihood security, small and marginal farmers, two-part designs.

The Indian rural population mainly consists of small and marginal farmers dependent on agriculture and its allied sectors for their livelihood. Though the contribution of agriculture to the Gross Domestic Product (GDP) has decreased post-independence, it has made the country self sufficient in food and made it a net exporter of agriculture and allied products. The ever-growing population, as well as globalization effects in India, demand increased quantity and better quality of food along with an increase in income. Therefore, pressure on diminishing available cultivable land to produce more quantity and quality of food keeps increasing.

Integrated farming system (IFS) is considered a powerful tool and is key to ensuring income, employment, livelihood and nutritional security for small and marginal farmers. It meets the above goals through multiple uses of available resources, thus giving scope for year-round income from various enterprises of the system. Improving the farming system through experimentation to attain household-level self-sufficiency, land utilization efficiency, and sustainable livelihood security is the major need of the hour. A sustainable livelihood security index of improved IFS compared with benchmark farming is available for semiarid regions³.

Often, it may not be possible to adopt all the components of IFS available in a particular location due to management issues pertaining to resources and proper land utilization. Besides the compulsory components like location-specific dominant crops and dairy, various possible components, viz. poultry, piggery, goastry, fishery, apiculture, horticulture and sericulture, can be included in IFS to maximize the overall profit of the farmers. So, it is pivotal to identify and adopt appropriate statistical techniques for choosing the best combination of components (along with crops and dairy) based on the availability of resources in order to generate optimum income for the farmers. The supplementary material (Figure 1 of Appendix 1) explains the steps for selecting the best combination of components to ensure sustainable livelihood security for small and marginal farmers.

Designing an experiment is inevitable in almost every agricultural and other scientific research. A properly designed experiment enhances the efficient use of available resources. Implementing experimental designs in farming systems can improve their outcomes. Likewise, in IFS, designs can be used to find the best combination of components available to obtain maximum profit.

Incomplete block designs (IBDs) have been established as an important benchmark for obtaining useful information from planned scientific experiments with high precision utilizing fewer resources. Some important classes of these designs include balanced incomplete block (BIB) designs² and partially balanced incomplete block (PBIB) designs³. The concept of association schemes was introduced to study these designs, and the schemes were classified into two associate classes⁷. The two associate-class PBIB designs were tabulated in the range $r, k \leq 10$, where $r$ and $k$ are the number of replications and block size respectively⁵. If the experimenter is constrained by resources, PBIB designs with three associate classes serve as an alternative to BIB designs or two-associate-class PBIB designs. Several association schemes have been defined for designs with three associate classes⁸-¹¹. 

Amalgamation of two (or more) IBDs, viz. BIB designs, PBIB designs or $r$-designs may help obtain higher dimensional designs, thus facilitating dealing with more complex problems. Method for constructing row-column designs (RCDs) was introduced earlier¹², but the method involved complexity. This complex algorithm had been generalized¹³ and an R package¹⁴ was developed to ease the construction

*For correspondence. (e-mail: cini2204@gmail.com)
RESEARCH ARTICLES

procedure. Sometimes more than one treatment must be applied to an experimental unit in order to identify the best treatment combination.

There may be situations where certain experimental units are not available for the application of treatments. In such cases, where the use of complete RCDs may not be feasible, two-part structurally incomplete RCDs (SIRCDs) may prove advantageous. Saharay\textsuperscript{35} studied SIRCDs, where treatments were allotted only to some row–column interactions. Sometimes, due to unavoidable constraints, we may have to allot the same treatment to either rows or columns, thereby making one of the blocking factors ineffective. This gives rise to two-part SI block designs, which form a particular case of the two-part SIRCDs.

Trial designs were introduced to streamline drug development in rare diseases\textsuperscript{16}. Two-part balanced incomplete block designs or two-part 2 designs for cancer trials with only a limited number of cancer types and a limited number of drugs have been discussed\textsuperscript{17}. The concept of basket trials\textsuperscript{18}, where several different drugs are tested on various diseases in a single protocol, had been considered in the above-mentioned studies.

Here, we develop two-part SI block designs and two-part SIRCDs for identifying the best components of IFS. We present some experimental situations and define two-part SIRCDs and two-part SI block designs.

Experimental conditions

Let us consider an IFS research trial where the aim is to find out the best farming system at a given location. Besides the compulsory components, field crops and dairy in all the systems, there are others like poultry, piggery, apiculture, goatery and fishery. The aim is to find the best combination that will yield maximum profit. \((\delta_1)\) Combinations were allotted to ten households in each of the locations. Every location may not have resources for every combination. Here, the comparison of household groups is not of interest, so locations can be treated as blocks and a design can be developed for such a situation.

Further, if we consider farmers (grouped based on their economic status) in place of household groups, it will add to another source of variation. For such a situation, an RCD can be developed.

Units occurring at certain row–column intersections may not be available for the experiment and hence do not receive any treatment; such a design is structurally incomplete.

Materials and methods

Preliminaries

We define two IBDs, viz. design 1 (\(\delta_1\)) and design 2 (\(\delta_2\)) with parameters \(v_1, b_1, r_1, k_1\) and \(v_2, b_2, r_2, k_2\) respectively, where \(v_1\) and \(v_2\) are the number of treatments in \(\delta_1\) and \(\delta_2\) respectively. Each treatment in \(\delta_1\) and \(\delta_2\) is replicated \(r_1\) and \(r_2\) times respectively. \(b_1\) is the number of blocks in \(\delta_1\), each of size \(k_1\) and \(b_2\) is the number of blocks in \(\delta_2\), each of size \(k_2\). Fusion of these IBDs may result in more complicated two-part SIRCDs or two-part SI block designs for testing a set of treatment combinations.

Some definitions

Two-part block designs: A two-part block design can be defined on \(v\) treatment combinations arranged in \(b\) blocks of size \(k\) (<\(v\) ) such that:

- Each treatment combination appears in each block at most once.
- All blocks contain the same number \((r_1)\) of treatment combinations.
- Every component in the combinations must appear in each block equally frequently.

Two-part RCDs: A two-part RCD (constructed using \(\delta_1\) and \(\delta_2\)) can be defined on \(v\) treatment combinations with \(p\) rows and \(q\) columns such that:

- Each treatment combination appears in each row at most once.
- All treatment combinations appear in each column equally frequently.
- All rows involve the same number \((k_1)\) of treatment combinations \((k_1 < v_1)\).
- All columns contain the same number \((r_1)\) of treatment combinations.
- Every component in the combinations of size \(k_2\) from \(v_2\) \((k_2 < v_2)\) treatments must appear in each column equally frequently.

Note: Two-part RCDs and block designs are considered SI, if at least one experimental unit in the design does not receive any treatment combination.

Model and experimental set-up

The model for a two-part block design for \(v\) treatments, \(b\) blocks each of size \(k\) and each treatment being replicated \(r\) times is as follows:

\[
y = \mu 1 + D_1 \tau + D_2 \beta + \epsilon.
\]

where \(y\) is the \(n \times 1\) vector of observations, \(\mu\) the general mean, \(1\) the \(n \times 1\) vector of ones, \(D_1\) the \(n \times v\) design matrix of observations versus treatments, \(\tau\) the \(v \times 1\) vector of treatment effects, \(D_2\) the \(n \times b\) design matrix of observations versus rows, \(\beta\) the \(b \times 1\) vector of column effects and \(\epsilon\) is the \(n \times 1\) vector of random errors with \(E(\epsilon) = 0\) and \(D(\epsilon) = \sigma^2 I_n\).
The information (C) matrix under two-part SI block design set-up is:

\[ C = X'_1X_1 - X'_2X_2(X'_2X_2)^{-1}X'_2X_1 = \nu v \frac{1}{r_1} N'_1N'_1, \]

upon simplification, \( r_1 \) where \( N'_1 \) is the incidence matrix of treatment versus blocks.

Now, if we add another source of variation in eq. (1), we get the experimental model for two-part RCDs for \( v \) treatments, each replicated \( r \) times arranged in \( p \) rows and \( q \) columns. It can be represented in matrix notation as

\[ y = D'_1\alpha + D'_2\beta + \mu + \varepsilon. \]  

where \( D'_1 \) is the \( n \times p \) design matrix of observations versus rows, \( \alpha \) the \( p \times 1 \) vector of row effects, \( D'_2 \) the \( n \times q \) design matrix of observations versus columns, \( \beta \) the \( q \times 1 \) vector of column effects and the other terms are as defined in eq. (1).

The \( C \)-matrix can be derived and simplified as:

\[ C = q_1 - N_1C'_1 \left[ \left( X'_2X_2 \right)^{-1} \right] \nu \left[ \left( X'_2X_2 \right)^{-1} \right] \]

\[ \times \left( \left( X'_2X_2 \right)^{-1} \right) J_{qvy}. \]

where \( N'_1 \) and \( N'_2 \) are the incidence matrices of treatment versus rows and rows versus columns respectively. \( N'_2 \) is further partitioned into \( N''_2 \) of order \( p \times (q - 1) \), so that it will be conformable for multiplication and \( N''_2N''_2 \) is a full rank sub-matrix of \( N'_2N'_2 \).

**Results and discussion**

**Two-part SI block designs**

Consider two IBDs, \( \delta_1 \) and \( \delta_2 \), which satisfy the condition \( b_1 = b_2 \). Fill \( k_1 \) positions in each block of \( \delta_1 \) with contents of any one block of \( \delta_2 \). The block contents of \( \delta_2 \) may have to be rearranged so that the conditions of not a single treatment combination occurring more than once in any given row, each treatment combination occurring at most once in any given column, and each individual component of treatment combination occurring equally frequently in each column are satisfied. Thus, the resultant design will be a two-part SI block design with columns as blocks. The parameters of the two-part SI block design are: \( v \) treatment combinations, \( r = 2k_1 \) replications, and \( b = v_1 \) blocks, each of size \( k = r_1 \).

Example 4.1: Suppose we have four components, say poultry, piggy, goatery and fishery. The Supplementary material (Figure 2 of Appendix 1) shows all possible two-tuple combinations of the four components (Figure 1).

The experiment was conducted at nine locations. Our objective was to find the best combination of components for each location. For this, we need to ensure that each location receives all the combinations. Let us consider two IBDs, \( \delta_1 \) (lattice partially balanced IBD) with parameters \( v = 9, \ b_1 = 12, \ r_1 = 4, \ k_1 = 3 \), representing the location and \( \delta_2 \) (an unreplicated balanced IBD with repeated blocks) with parameters \( v_2 = 4, \ b_2 = 6, \ r_2 = 3, \ k_2 = 2 \), representing component combinations.

Adoption of every farming system may not be possible by each household due to a shortage of resources. Hence, one may apply the same treatment combination (a particular block content of \( \delta_2 \)) in all the \( k_1 = 3 \) positions (locations where the components are available), such that the component treatments occur equally frequently in each block. Thus, the resultant design will be a two-part SI block design with parameters \( v = 6, \ b = 9, \ r = 6 \) and \( k = 4 \), considering the locations as blocks (Figure 2).

The two-part SI block design so obtained is structurally incomplete and equi-replicate.

In continuation with Example 4.1, if we add another source of variation, e.g. if households are replaced by farmers (grouped on the basis of economic status), then a SIRCD can be developed in a manner similar to block designs. The resultant design will be a two-part SIRCD with \( v \)
treatment combinations, each replicated \( r = v_1 \) times, \( p = b_1 = b_2 \) rows and \( q = v_1 \) columns.

**Example 4.2:** We consider here two designs, \( \delta_1 \) (a triangular partially balanced IBD\(^{2}\)) and \( \delta_2 \) (an unreduced BIB design) with \( v_1 = 10, b_1 = 6, r_1 = 3, k_1 = 5 \) and \( v_2 = 6, b_2 = 6, r_2 = 3, k_2 = 2 \) respectively, with block contents of \( \delta_1 \) and \( \delta_2 \) as given in Figure 3.

Each block of \( \delta_2 \) represents a type of two-component combination. The first row of the design will receive any five out of the six types of combinations in positions 1, 3, 4, 5, and 9. The remaining five rows are filled similarly, manner corresponding to the treatment positions of \( \delta_1 \). Here, \( r_1 = 3 \); hence each age group can be given only three treatment combinations. To find the best combination for each age group, every combination needs to be attempted by each age group. Therefore, blocks of \( \delta_1 \) may be replicated again so that all six combinations appear in each column exactly once. Thus, the resultant design so obtained will be a two-part SIRCD with \( v = 6 \) treatment combinations, each replicated \( r = 10 \) times, \( p = 12 \) rows and \( q = 10 \) columns (Figure 4).

The two-part SIRCD obtained is structurally incomplete, equi-replicate, and both row and column components are balanced.

**Example 4.3:** Sometimes, it may not be possible to obtain 12 rows and 10 columns as in Example 4.2 for the same number of treatment combinations. For such situations, one must obtain designs with a lesser number of rows and columns. Let us consider two designs, viz. \( \delta_1 \) (\( v_1 = 6, b_1 = 3, r_1 = 2, k_1 = 4 \)) and \( \delta_2 \) (\( v_2 = 6, b_2 = 6, r_2 = 3, k_2 = 2 \)), where \( \delta_1 \) is a singular group divisible design and \( \delta_2 \) an unreduced BIB design (Figure 5).

Following the procedure detailed above, a two-part SIRCD with \( v = 6 \) treatment combinations, \( r = 6, p = 9 \) and \( q = 6 \) is obtained (Figure 6).

Simplified general forms of the \( C \)-matrix pertaining to Examples 4.1, 4.2 and 4.3 are given in the Supplementary material (Appendix 2), along with association schemes followed by the treatment combinations.

**Randomization**

Due to the structurally incomplete nature of the designs discussed above, the randomization procedure is somewhat

![Figure 2](image-url) Two-part SI block design with parameters \( v = 6, b = 9, r = 6 \) and \( k = 4 \).

![Figure 3](image-url) Block contents of two designs, \( \delta_1 \) (\( v_1 = 10, b_1 = 6, r_1 = 3, k_1 = 5 \)) and \( \delta_2 \) (\( v_2 = 6, b_2 = 6, r_2 = 3, k_2 = 2 \)).

![Figure 4](image-url) Two-part SIRCD with \( v = 6, r = 10, p = 12, q = 10 \).

![Figure 5](image-url) Block contents of two designs, \( \delta_1 \) (\( v_1 = 6, b_1 = 3, r_1 = 2, k_1 = 4 \)) and \( \delta_2 \) (\( v_2 = 6, b_2 = 6, r_2 = 3, k_2 = 2 \)).

![Figure 6](image-url) Two-part SIRCD with \( v = 6, r = 6, p = 9 \) and \( q = 6 \).
restricted compared to block designs or RCDs. The steps are of randomization given below:

- Randomize \( v \) treatment combinations, i.e. random allocation of treatment combinations to 1, 2, ..., \( v \) random numbers.
- Randomize \( q \) columns, keeping rows undisturbed.
- Randomize \( p \) rows, keeping columns undisturbed and considering the available resources in each location.
- Randomization within rows and columns is also restricted.

**Efficiency factor**

The information (\( C \)) matrix pertaining to the estimation of treatment contrasts after adjusting the effects of rows and/or columns/blocks, according to the nature of the design considered, can be easily estimated using the derived expressions given in eq. (2) and/or eq. (4) by substituting \( N_1 \) and \( N_2 \) matrices as required. A program has been written in Statistical Analysis System (Supplementary Appendix 3) to facilitate the easy computation of the \( C \)-matrix and canonical efficiency factors of the proposed designs. In comparison to an equi-replicate and proper orthogonal design with the same number of treatments, the efficiency factors of these designs can be computed as \( \sqrt{\frac{2}{v}} \) times the harmonic mean of non-zero eigenvalues of the information matrix pertaining to the proposed designs\(^{20}\).

The canonical efficiency factors of the designs in Examples 4.1, 4.2 and 4.3 are 0.8824, 0.9559 and 0.8333 respectively.

**Summary**

The transition from a mono-cropping system to an IFS-based resilient system is pivotal for small and marginal farmers. Two-part RCDs and block designs obtained by combining two or more IBDs give a powerful solution in identifying the best suitable location-specific IFS combinations. However, every selected household/group of farmers may not be able to adopt/afford all IFS combinations. Hence, two-part SI designs for comparing a set of treatment combinations are suggested. These designs are easy to construct and have reasonably high-efficiency factors, which will encourage researchers to adopt them.

**Conflict of interest:** The authors declare that there is no conflict of interest.


ACKNOWLEDGEMENTS. We thank the Editor and reviewer for their valuable suggestions that helped improve manuscript. Sayantani Karmakar thanks the Post Graduate School, ICAR-Indian Agricultural Research Institute, New Delhi for providing the necessary facilities, and University Grants Commission, New Delhi for financial support. We thank the scientists of ICAR-Indian Institute of Farming Systems Research, Modipuram for interactive discussions that mooted the idea of constructing tailor-made designs for integrated farming systems research.

Received 3 November 2022; revised accepted 13 February 2023

doi: 10.18520/cs/v124/i9/1053-1057