Common experimental designs used in plant breeding (RCBD, IBD, Rowcol)

Luis Fernando Delgado Munoz
luis.delgado@cgiar.org
Faisalabad, November 2023

@BiovIntCIAT_eng
@BiovIntCIAT_esp

#Alliance4Science

Statistical analysis of multi-environment data

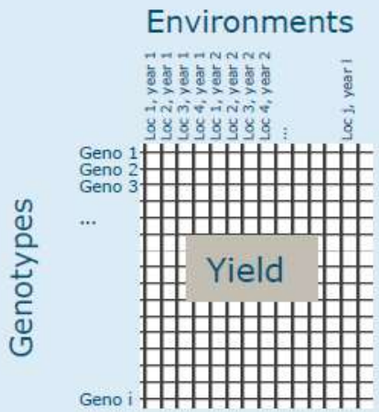
Single trials

- Experimental designs
- Models to separate genetic signal from experimental noise
- Calculation of adjusted means and heritability

Analysis per trial

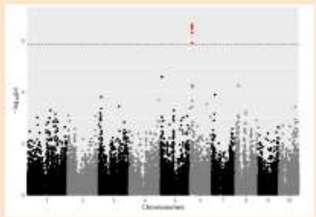
GxE analysis

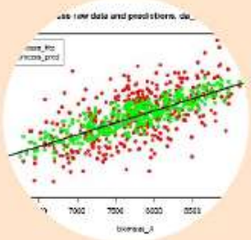
- Mixed models for GLY



Genotype to phenotype modelling

- QTL detection and GWAS
- Genomic prediction





WAGENINGEN UNIVERSITY & RESEARCH

Alliance Bioversity & CIAT

CGIAR

What is experimental design?

Experimental Design: the study of strategies for efficient data collection that lead to proper estimates of parameters relevant to the researcher's objective.(Lentner and Bishop, 1993).

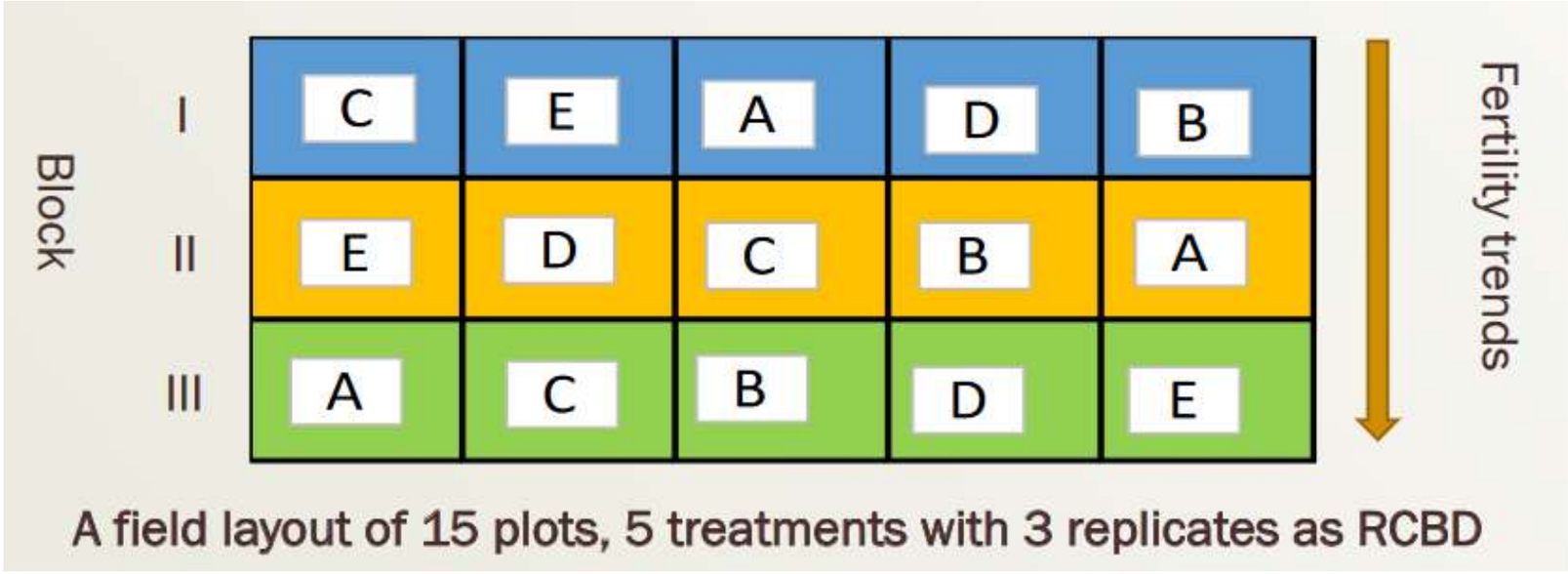
- Design structure – grouping of homogeneous plots into blocks
- Treatment structure - the set of treatments or treatment combination the researcher has selected to study and/or compare.

- Experimental designs are divided into two categories:
  - Complete block design (blocks are complete) CRD, RCBD, LD, SPD
  - Incomplete block design (blocks are incomplete) Augmented design, lattice design

Principle of experimental design

3 principles are:

- **Replication**
- **Blocking**
- **Randomization**



Replication – Repeating a treatment more than once in an experiment.

The reasons for replication are:

- Understand the variability in the experiment
- Provide an estimate of experimental error
- Estimate the difference between mean effects of different treatments.
- To improve the precision of the experiment by reducing standard erro

Principle of experimental design

Blocking: process of partitioning experimental units into homogeneous groups such that.

- Intrablock plot variability minimized.
- Intrbock plot variability maximized.

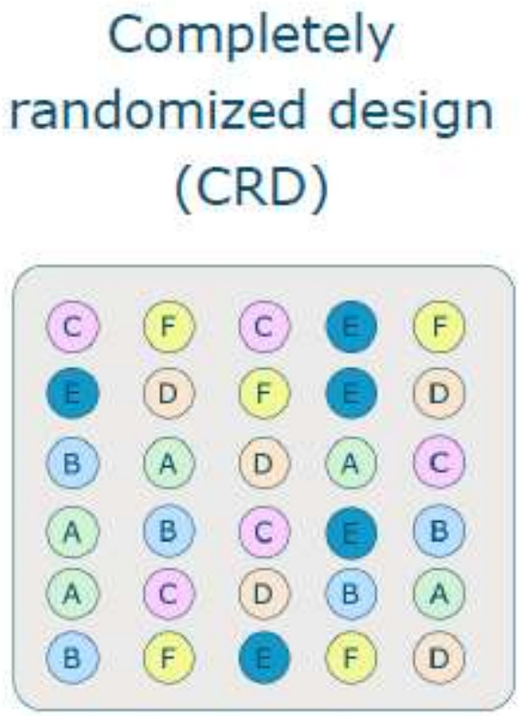
Randomization: assignment of treatments to experimental units so that all units have an equal chance of receiving a treatment

Why randomization?

- It helps to assure unbiased estimates of treatment means and experimental error
- Statistical estimation and tests of hypotheses on effects are theoretically valid

Completely Randomized Design (CRD)

- CRD is best suited for experiments with a small number of treatments.
- It is used when experimental units are essentially homogeneous.
- CRD is generally applicable to the lab experimental conditions.
- Because of homogeneity requirement, it may be difficult to use this design for field experiments.



A typical layout of 30 experimental units, 6 treatments with 5 replicates as CRD

Statistical Model CRD

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

$\mu = overall\ mean$
$\tau_i = treatment\ effect$
$\varepsilon_{ij} = random\ error$
$\varepsilon \sim N(0, \sigma^2)$

Skeleton of ANOVA Table

Source	DF	SS	MS	F	P-value
Treatment	t-1	SS_t	$SS_t/(t-1)$	MS_t/MS_e	
Error	(n-1)–(t-1)	SS_e	$SS_e/[(n-1)-(t-1)]$		
Total	n-1	TotalSS			

Randomized Complete Block Design (RCBD)

- Probably the most used experimental design in agricultural field trial.
- Each treatment replicated once in each block (balanced and complete).
- The block of experimental units should be as uniform as possible.
- The randomization process for RCBD is applied separately to each block (Gomez and Gomez, 1984).
- The design is based on three principles of experimental design.

A typical field layout of 36 plots, 6 treatments with 6 replicates as RCBD

Statistical Model RCRD

$$\boldsymbol{Y_{ij} = \mu + \beta_i \tau_i + \varepsilon_{ij}}$$

$\mu = overall\ mean$
$\beta_i = block\ effect$
$\tau_i = treatment\ effect$
$\varepsilon_{ij} = random\ error$
$\varepsilon \sim N(0, \sigma^2)$

Skeleton of ANOVA Table

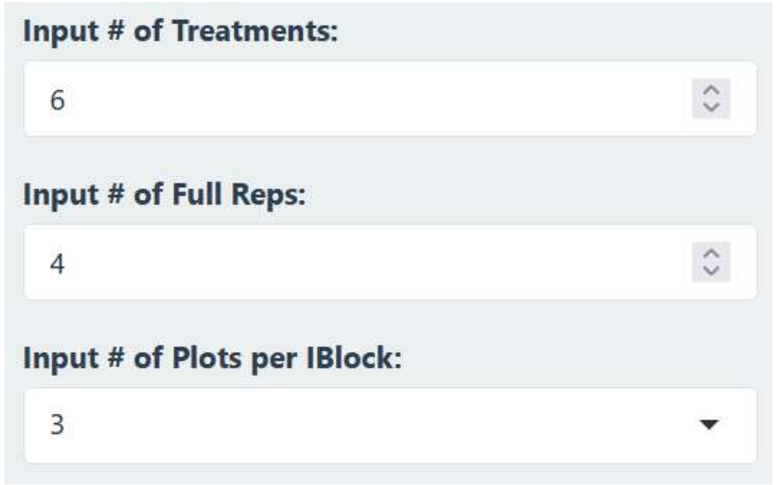
Source	DF	SS	MS	F	P-value
Rep	r-1	SS_r	$SS_r/(r-1)$		
Treatments	t-1	SS_t	$SS_t/(t-1)$	MS_t/MS_e	
Error	(r-1) (t-1)	SS_e	$SS_e/(r-1)(t-1)$		
Total	n-1	TSS			

Incomplete block design

- An appropriate design for evaluating large number of treatment.
- Useful when heterogeneity is great within block.
- It gives greater precision than RCBD.

- In an incomplete-block design, not every treatment is applied in every block. Instead, each block contains only a subset of the treatments.

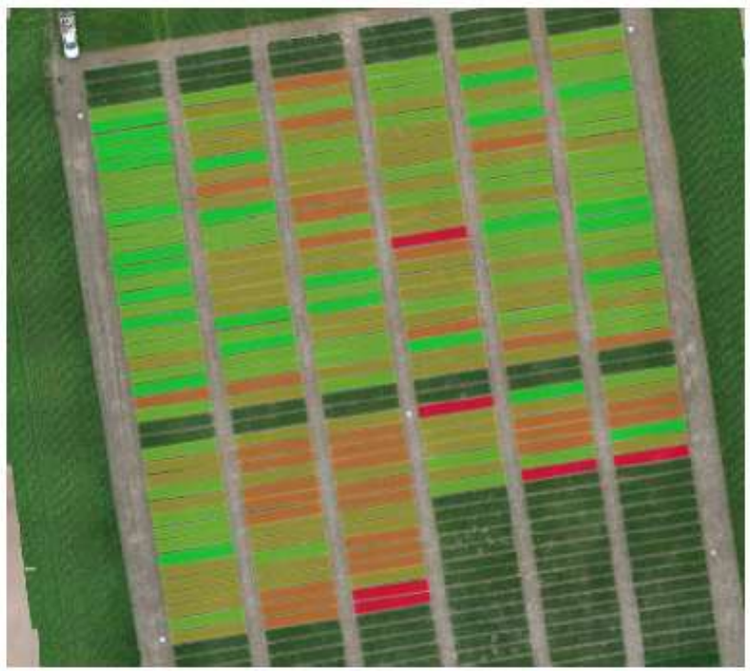
- Each superblock (the larger block) can be seen as a complete-block design because every treatment appears in the superblock. However, within the individual smaller blocks (which form the superblock), the design is still incomplete because not all treatments appear in each of these smaller blocks.



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

Complications in breeding trials

- The treatment structure in breeding trials typically is a single factor with many levels (100 to 1000 genotypes).
- Because of the many treatments, replication is costly.
- Fields can be heterogeneous in more than one direction.



Advance experimental design

Design	Blocks are formed in	Blocks can be grouped in complete replicates (resolvable)
Row-column	Two directions (rows AND columns)	no
Resolvable row-column	Two directions (rows AND columns)	yes

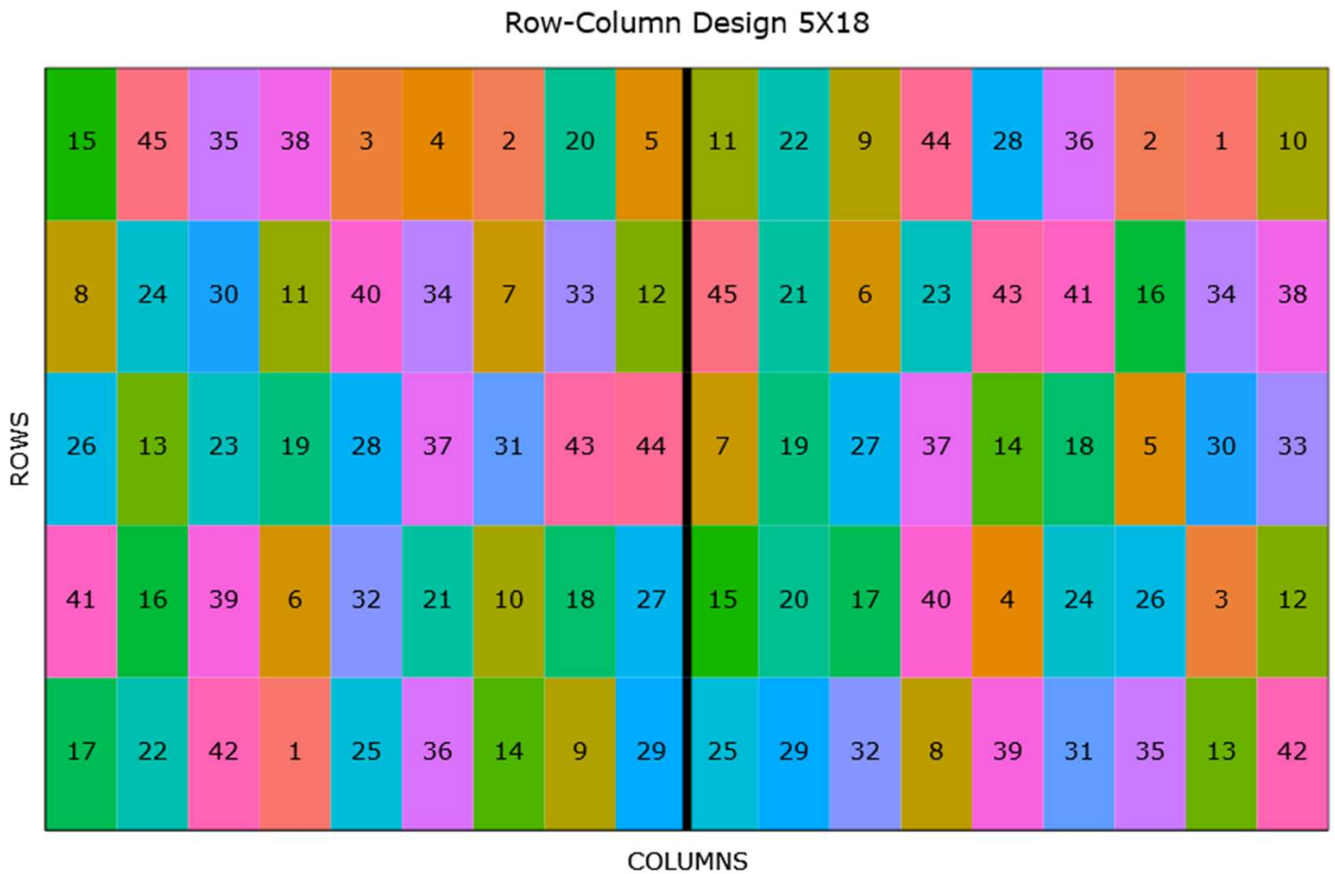
Row – col design

- Two types of blocks: rows/ columns
- Rows and/or columns can be incomplete.
- Randomization not simple: try to achieve balance in 2 directions!
- 12 treatments; 4 replicatesin 4 complete columns; 12 incomplete rows.
- Ideally: each treatment pair occurs once or twice in the same row.

	1	2	3	4
1	2	6	5	12
2	9	11	4	6
3	3	5	10	11
4	11	8	1	2
5	5	1	8	4
6	6	3	7	8
7	1	7	6	10
8	10	4	2	3
9	12	9	3	1
10	4	12	11	7
11	8	10	12	9
12	7	2	9	5

Resolvable row – col design

- Define two full replicates: make it resolvable.
- Rows and columns are incomplete.
- Genotypes within same row (column) in rep 1 will not be in the same row (column) in rep 2.



Thank you.