FAKULTI PERTANIAN
FACULTY OF AGRICULTURE
فاكولتي ڤرتانين

UNIVERSITI PUTRA MALAYSIA

AGR5201

ADVANCED STATISTICAL METHODS

Split plot experiment

Analysis of variance of split plot experiments and mean comparison

Topic outline

1.0 SPLIT PLOT DESIGN
- 1.1 Experimental layout
- 1.2 Why do you need a split plot
- 1.3 Linear additive model
- 1.4 Split plot ANOVA
- 1.5 Numerical example
- 1.6 Factorial RCBD vs. split plot

2.0 ANALYSIS USING R
- ANOVA using agricolae package
- ANOVA using AgroR package

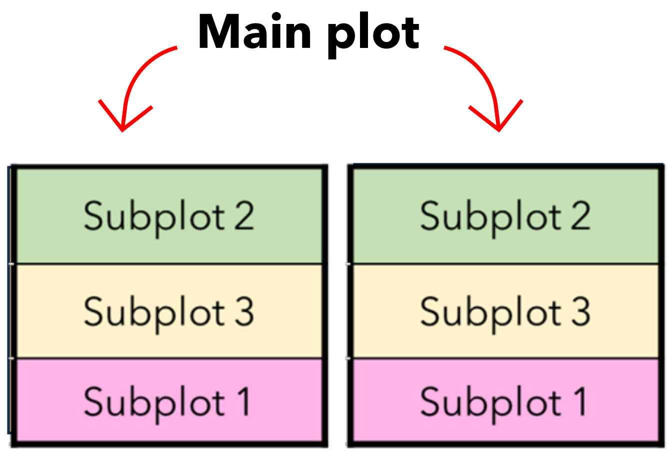
Reference book:
Gomez, A.G & Gomez, A.A. (1984). Statistical procedures for agricultural research. John Wiley & Sons. Page 97

Website:
https://online.stat.psu.edu/stat503/lesson/14/14.3

1.0 SPLIT PLOT DESIGN

Introduction

- A design that involves factorial experiments (2 factors)
  - Note: If 3 factors → split-split plot design
- Two types of experimental units:
  - Factor A and factor B
- One unit is nested within a bigger unit
  - The bigger unit → Main plot
  - The smaller unit → Subplot



1.0 SPLIT PLOT DESIGN

Introduction

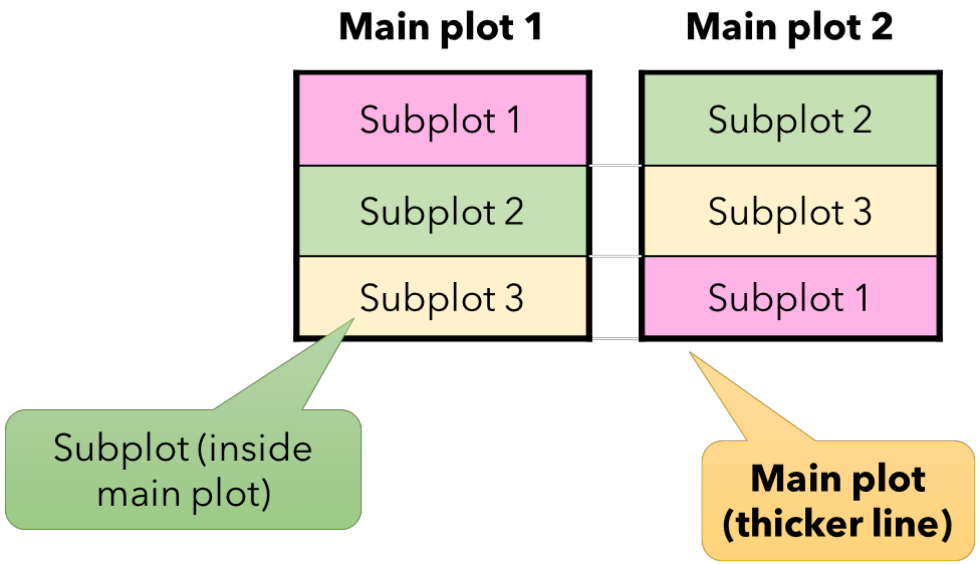
- Why use split plot design?
  1. Management practices
  2. Degree of precision
  3. Relative size of the main effects
- Uses:
  - In experiments where different factors require different size plots
  - To introduce new factors into an experiment that is already in progress

1.0 SPLIT PLOT DESIGN

Different size requirements

- The split plot is a design which allows the levels of one factor to be applied to large plots while the levels of another factor are applied to small plots
    - Large plots are whole plots or main plots
    - Smaller plots are split plots or subplots

General layout of a split plot RCBD



1.0 SPLIT PLOT DESIGN

Split plot RCBD example

- Treatments:
  - Factor A (Main plot): 4 nitrogen rates (N1, N2, N3, N4)
  - Factor B (Subplot): 3 varieties (V1, V2, V3)
  - Treatment combination:

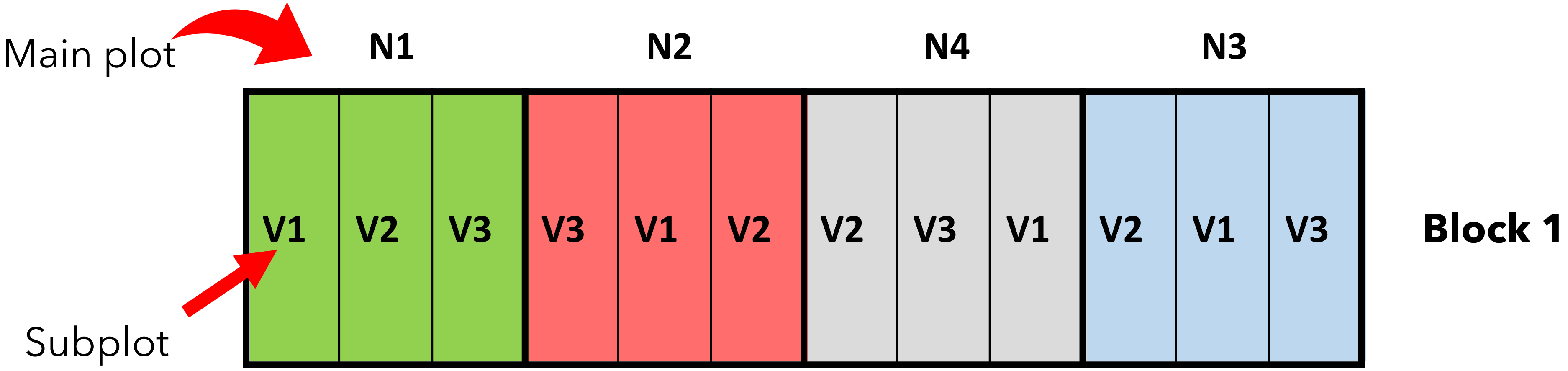
    = 4 (factor A) x 3 (factor B)

    = **12 treatment combinations**
- Rep/Block: 4
- Experimental unit: 4 x 3 x 4 = **48**

1.0 SPLIT PLOT DESIGN

Randomization

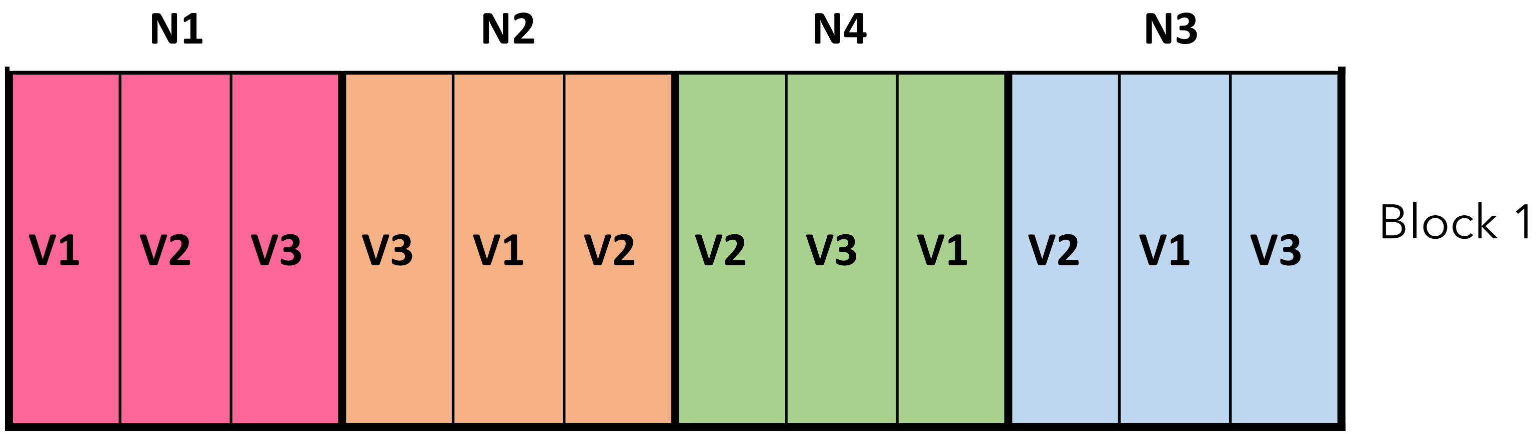
- Levels of the whole-plot factor are randomly assigned to the main plots, using a different randomization for each block (for RCBD)
- Levels of the subplots are randomly assigned within each main plot using a separate randomization for each main plot
- Example:
  - Factor A (Main plot): Nitrogen rates (N1, N2, N3, N4)
  - Factor B (Subplot): Variety (V1, V2, V3)



1.1 Experimental layout

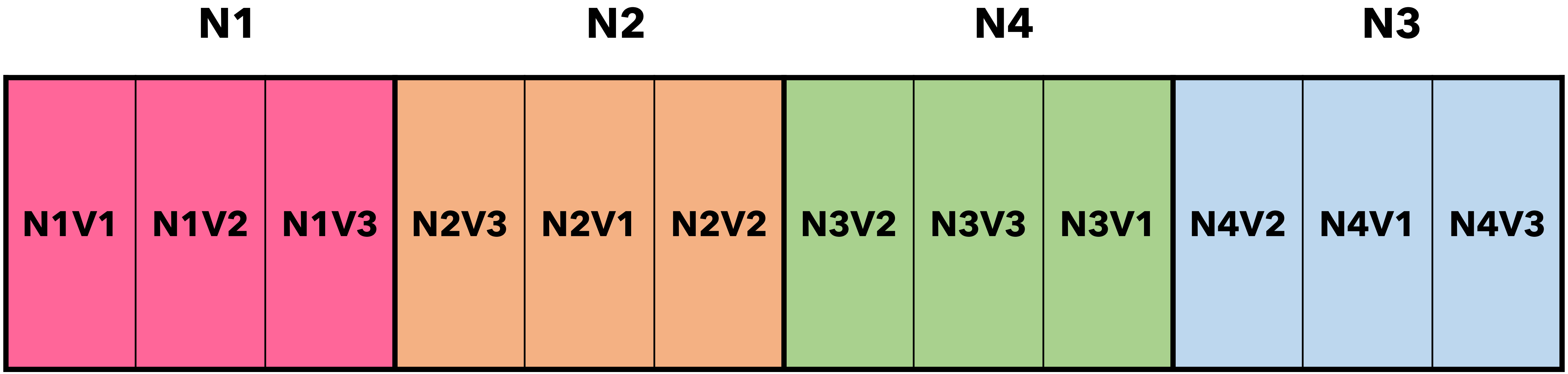
Split plot - RCBD

Treatments:
Factor A (Main plot): Nitrogen rate (N1, N2, N3, N4)
Factor B (Subplot): Variety (V1, V2, V3)



1.1 Experimental layout

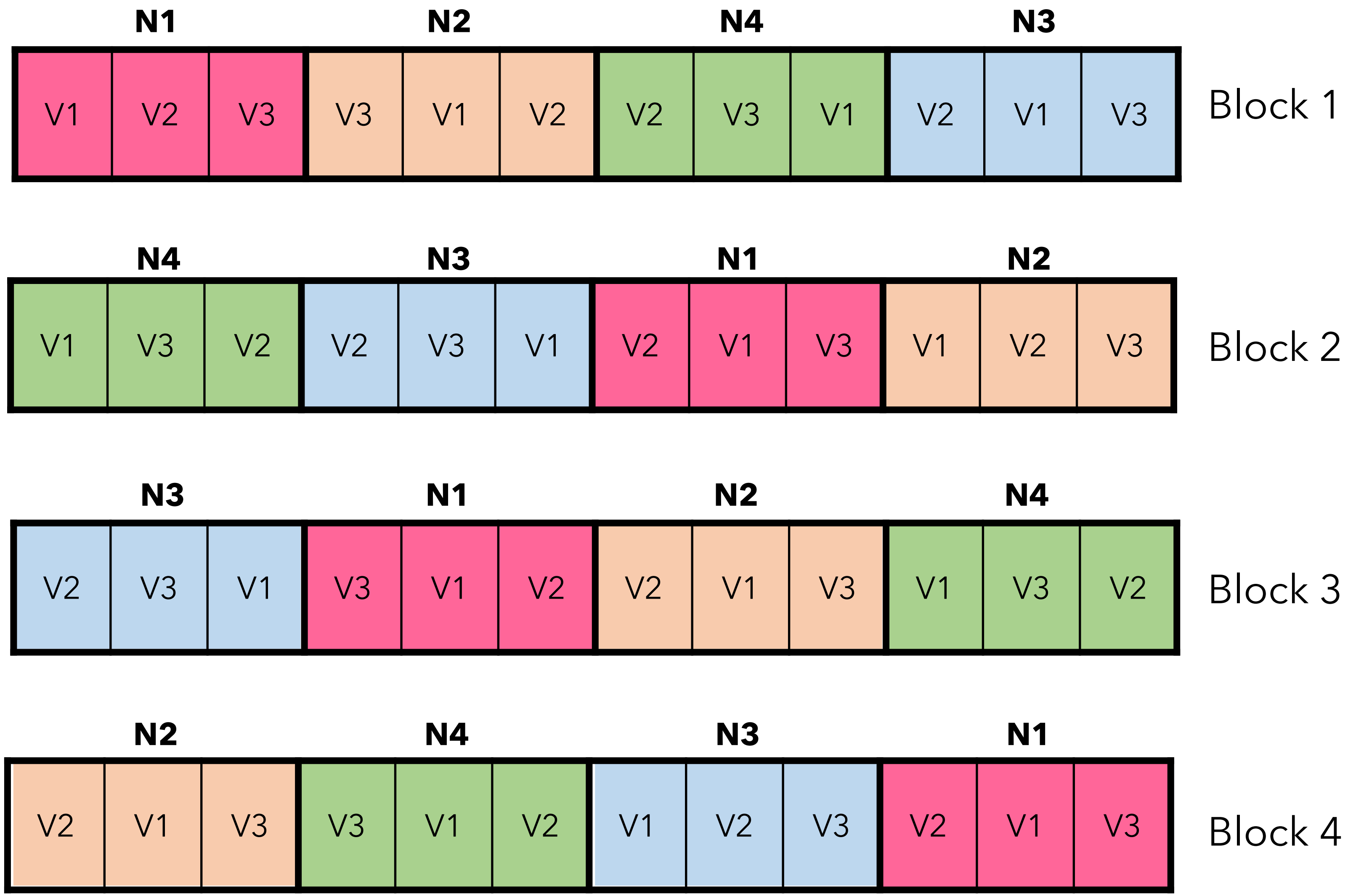
Treatment combination (one block)



Block 1

1.1 Experimental layout

The whole experiment



1.2 Why do you need a split plot?

1. Degree of precision

- For a greater precision for the measurement of one factor:
  - Precision: Factor B > Factor A.
  - Make factor B → subplot; factor A → main plot.
- The subplot has more precision than main plot.
  - Subplot has > rep compared to main plot = more precise.

    Main plot replication = # block = 4

    Subplot replication = # block x # main plot level

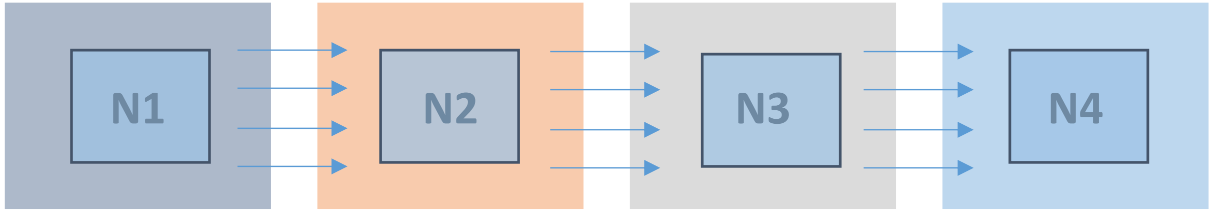
    → 4 blocks x 4 nitrogen rates = 16 reps for subplot

1.2 Why do you need a split plot?

2. Management practices

Examples:

a) Cultural practices required by a factor dictate the use of large plot:
   - Nitrogen fertilizer → Leaching of N to neighboring plots that allow you to harvest only a small part of your experimental unit.

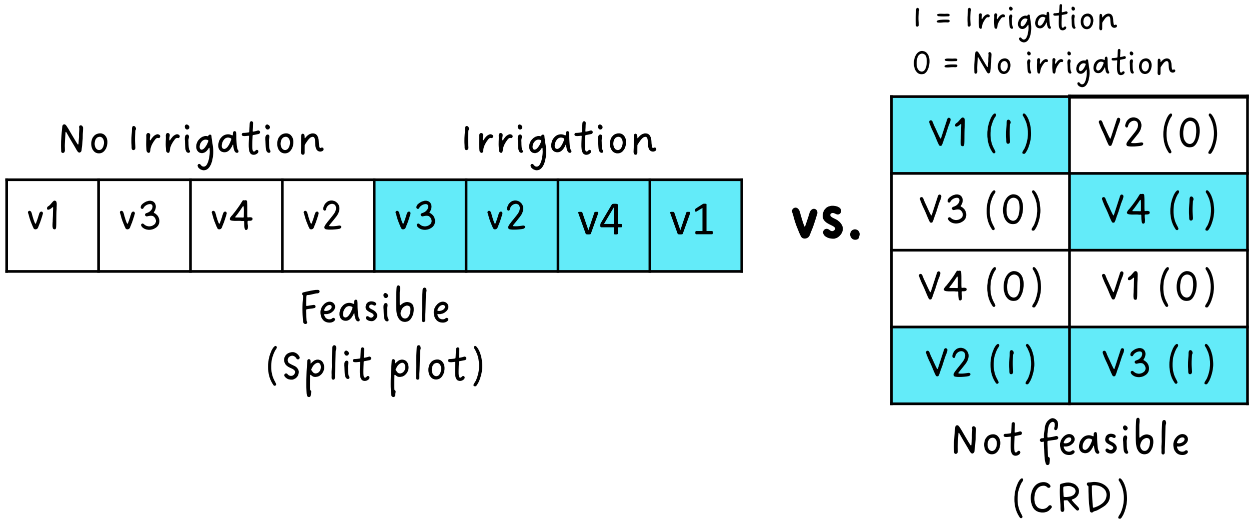


1.2 Why do you need a split plot?

2. Management practice (cont'd)

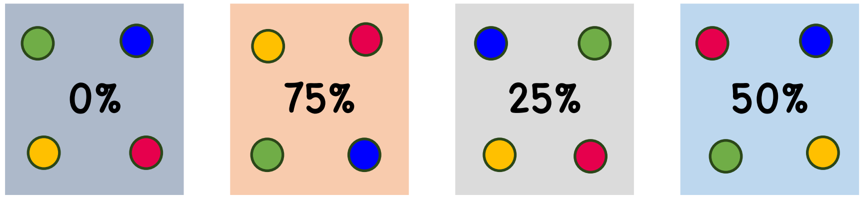
b) Water management (irrigation system)

- Difficult to set up irrigation system (F1) to each experimental plot, so apply irrigation to one big area. Then the 2nd factor (e.g. variety) is randomized within each Irrigation level.



c) Shade

- Require different set up / shade hut for different shade levels.



1.2 Why do you need a split plot?

3. Relative size of the main effects

- If the main effect of one factor (factor B) is expected to be larger and easier to detect than the other factor (factor A):
  - assign factor B to the main plot, and factor A to the subplot
  - Example:
    - Factor A: variety
    - Factor B: fertilizer

    → Assign fertilizer to the main plot as fertilizer effect is expected to be larger than the varietal effect

1.2 Why do you need a split plot?

Split-plot: Pros and Cons

Advantages

- Permits the efficient use of some factors that require different sizes of plot for their application
- Permits the introduction of new treatments into an experiment that is already in progress

Disadvantages

- Main plot factor is estimated with less precision so larger differences are required for significance – may be difficult to obtain adequate degrees of freedom for the main plot error
- Statistical analysis is more complex because different standard errors are required for different comparisons

1.3 Linear additive model

Split plot RCBD

Observation → $$Y_{ijk} = \mu + \gamma_k + \alpha_i + (\alpha\gamma)_{ik} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$ ← Error (b)

Overall mean

Block effect

Factor A (Whole plot) effect

Error (a)

Factor B (Subplot) effect

A*B interaction effect

Error a = block x main plot (factor A)
Error b = random error (residuals)

1.4 Split plot ANOVA

Skeleton of ANOVA table - RCBD

Linear model RCBD

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + \varepsilon_{ijk}$$

Source	Degree of freedom	Sum of square	Mean square	F test
Block	r-1	SSR	MSR	MSR/MSE_A
Factor A (Main plot)	a-1	SSA	MSA	MSA/MSE_A
Error a (Block*Factor A)	**(r-1)*(a-1)**	SSE_A	MSE_A	
Factor B	b-1	SSB	MSB	MSB/MSE_B
Factor A*Factor B	(a-1)*(b-1)	SSAB	MSAB	MSAB/MSE_B
Error b	**a(r-1)(b-1)**	SSE_B	MSE_B	
Total	(rab) -1	SST		

1.3 Linear additive model

Split plot CRD

Observation

$$Y_{ijk} = \mu + \alpha_i + (\alpha\gamma)_{ik} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Error (b)

Overall mean

Factor A (Whole plot) effect

Error (a)

Factor B (Subplot) effect

A*B interaction effect

Error a = replication x factor A
Error b = random error (residuals)

1.4 Split plot ANOVA

Skeleton of ANOVA table - CRD

Linear model CRD

$$Y_{ijk} = \mu + \tau_i + \rho_{k(i)} + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk}$$

Source	Degree of freedom	Sum of square	Mean square	F test
Factor A (Main plot)	a-1	SSA	MSA	MSA/MSE_A
Error a (Rep (Factor A))	**a(r-1)**	**SSE_A**	**MSE_A**	
Factor B	b-1	SSB	MSB	MSB/MSE_B
Factor A*Factor B	(a-1)*(b-1)	SSAB	MSAB	MSAB/MSE_B
Error b	**a(r-1)(b-1)**	**SSE_B**	**MSE_B**	
Total	(rab) -1	SST		

1.4 Split plot ANOVA

Computations – sum of squares (SS)

- Only the **error terms** are different from the usual two- factor analysis

SSTot	$\sum_i \sum_j \sum_k \left(Y_{ijk} - \bar{\bar{Y}}\right)^2$
SSR	$ab\sum_k \left(\bar{Y}_{..k} - \bar{\bar{Y}}\right)^2$
SSA	$rb\sum_i \left(\bar{Y}_{i..} - \bar{\bar{Y}}\right)^2$
SSE$_A$	$b\sum_i \sum_k \left(\bar{Y}_{i.k} - \bar{\bar{Y}}\right)^2 - SSA - SSR$
SSB	$ra\sum_j \left(\bar{Y}_{.j.} - \bar{\bar{Y}}\right)^2$
SSAB	$r\sum_i \sum_j \left(\bar{Y}_{ij.} - \bar{\bar{Y}}\right)^2 - SSA - SSB$
SSE$_B$	**SSTot - SSR - SSA - SSE$_A$ - SSB - SSAB**

1.4 Split plot ANOVA

F Ratios

F ratios are computed somewhat differently because there are two errors

- $F_R = MSR/MSE_A$ tests the effectiveness of blocking
- $F_A = MSA/MSE_A$ tests the sig. of the A main effect

- $F_B = MSB/MSE_B$ tests the sig. of the B main effect
- $F_{AB} = MSAB/MSE_B$ tests the sig. of the AB interaction

1.4 Split plot ANOVA

Standard Errors (SE) of Treatment Means

- Factor A Means $\sqrt{\dfrac{MSE_A}{rb}}$
- Factor B Means $\sqrt{\dfrac{MSE_B}{ra}}$
- Treatment AB Means $\sqrt{\dfrac{MSE_B}{r}}$

r = number of replication or block
a = the number of level for factor a
b = the number of level for factor b

The calculation of SE is calculated based on r*b, r*a and r experimental units for Factor A, B and AB interaction, respectively.

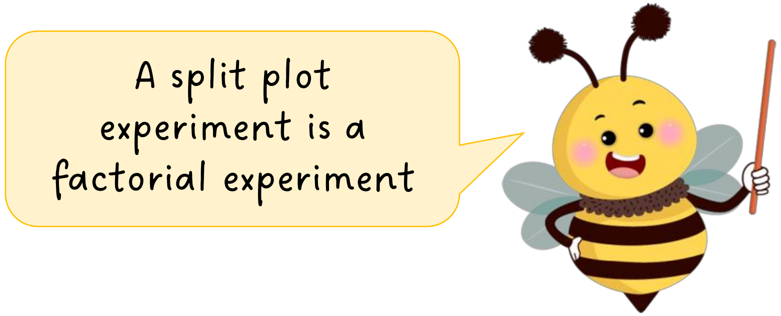
Example:

- **r = 4 (R1, R2, R3, R4)**
- **a = 3 (A1, A2, A3),**
- **b = 2 (B1, B2)**

Thus,

- For each main effect A1- A3, there are r*b = 4*2 = 8 experimental units.
- For each main effect B1 & B2, there are r*a = 4*3 = 12 experimental units.
- For each A1B1 – A4B2 treatment combination, there are r = 4 experimental units.

1.4 Split plot ANOVA



Interpretation

Much the same as a two-factor factorial:

- First, test the AB interaction. If it is **significant:**
  - The main effects have no meaning even if they test significant
  - Summarize in a two-way table of AB means
- If AB interaction is **not significant**
  - Look at the significance of the main effects
  - Summarize in one-way tables of means for factors with significant main effects

1.4 Split plot ANOVA

Example

- **Objective:**
  The effect of shade intensity on yield of herb of three herb species.
- **Factor A - Shades:**
  - 4 levels (0%, 25%, 50%, 75%)
- **Factor B - Herb species:**
  - A, B, C
- **Block:** 4 blocks

Hypotheses:

Interaction (Shade x species)

Ho: There is no significant interaction between shade and species on yield of herb

Ha: There is significant interaction between shade and species on yield of herb

Main effects – Shades

Ho: There is no significant effects of shade on yield of herb

Ha: There is significant effects of shade on yield of herb

Main effects – Species

Ho: There is no significant effects of species on yield of herb

Ha: There is significant effects of species on yield of herb

1.4 Split plot ANOVA

Split plot RCBD

Block = 4; Shade = 4;
Species = 3 ; N = 4x4x3 = 48

Source	Degree of freedom	Sum of square	Mean square	F
Block	r-1 = 4-1 = 3			MS block/MS Err(A)
Shade (Factor A)	a-1 = 4-1= 3			MS shade/MS Err(A)
Block x Shade (Error A)	(r-1) x (a-1) 3 x 3 = 9			
Herb (Factor B)	b-1 = 3-1 = 2			MS herb/MS Err(B)
Shade x Herb (SxH)	(a-1) x (b-1) 3 x 2 = 6			MS SxH/MS Err(B)
Error B	a(r-1)(b-1) = 4*(4-1)*(3-1) = 24			
Total	rab-1 = 48-1 = 47			

1.4 Split plot ANOVA

Two-factors (RCBD)

Block = 4; Shade = 4;
Species = 3 ; N = 4x4x3 = 48

Source	Degree of freedom	Sum of square	Mean square	F
Block	r-1 = 4-1 = 3			MS block/MS Error
Shade (Factor A)	a-1 = 4-1= 3			MS shade/MS Error
Herb (Factor B)	b-1 = 3-1 = 2			MS herb/MS Error
Shade x Herb (SxH)	(a-1) x (b-1) 3 x 2 = 6			MS SxH/MS Error
Error	(r-1)(ab-1) (4-1)(4x3-1) = 33			
Total	rab-1 = 48-1 = 47			

1.4 Split plot ANOVA | Comparison

In **split plot**, the ***precision*** **for main plot factor is lower** (low dfe_a)

Subplot has greater precision than main plot (greater dfe_b)

Split plot in RCBD

Source	Degree of freedom
Block	r-1 = 4-1 = 3
Shade (*Factor A*)	a-1 = 4-1= 3
Block x Shade (*Error A*)	(r-1) x (a-1) **3 x 3 = 9**
Herb (*Factor B*)	b-1 = 3-1 = 2
Shade x Herb (SxH)	(a-1) x (b-1) 3 x 2 = 6
Error B	a(r-1)(b-1) = 4*(4-1)*(3-1) = **24**
Total	rab-1 = 48-1 = 47

Two-factorial RCBD

Source	Degree of freedom
Block	r-1 = 4-1 = 3
Shade (Factor A)	a-1 = 4-1= 3
Herb (Factor B)	b-1 = 3-1 = 2
Shade x Herb (SxH)	(a-1) x (b-1) 3 x 2 = 6
Error	(r-1)(ab-1) (4-1)(4x3-1) = **33**
Total	rab-1 = 48-1 = 47

1.5 Numerical example

- A wheat breeder wanted to determine the effect of planting date on the yield of four varieties of winter wheat
- Two factors:
    1. Planting date (15 Oct, 1 Nov, 15 Nov)
    2. Variety (V1, V2, V3, V4)
- Because of the machinery involved, planting dates were assigned to the main plots
- Used a randomized complete block design with 3 blocks

1.5 Numerical example

Comparison split plot with regular factorial RCBD

- With a split-plot, there is better precision for sub-plots than for main plots, but neither has as many error df as with a conventional factorial
- There may be some gain in precision for subplots and interactions from having all levels of the subplots in close proximity to each other

Split plot RCBD

Source	df
Block	2
Date	2
Error (a)	4
Variety	3
Var x Date	6
Error (b)	18
Total	35

Factorial in RCBD

Source	df
Block	2
Date	2
Variety	3
Var x Date	6
Error	22
Total	35

1.5 Numerical example

Raw data

Block	R1			R2			R3		
Date	**15 Oct**	**1 Nov**	**15 Nov**	**15 Oct**	**1 Nov**	**15 Nov**	**15 Oct**	**1 Nov**	**15 Nov**
Variety 1	25	30	17	31	32	20	28	28	19
Variety 2	19	24	20	14	20	16	16	24	20
Variety 3	22	19	12	20	18	17	17	16	15
Variety 4	11	15	8	14	13	13	14	19	8

Grand mean = 18.722

1.5 Numerical example

Hypotheses

Interaction (Date x variety)

- Ho: There is no significant interaction between date and variety on yield of wheat
- Ha: There is significant interaction between date and variety on yield of wheat

Main effects – Date

- Ho: There is no significant effects of date on yield of wheat
- Ha: There is significant effects of date on yield of wheat

Main effects – Variety

- Ho: There is no significant effects of variety on yield of wheat
- Ha: There is significant effects of variety on yield of wheat

1.5 Numerical example

ANOVA table

Source	df	SS	MS	F
Block	2	1.55	0.78	0.22
Date	2	227.05	113.53	32.16**
Error (a)	4	14.12	3.53	
Variety	3	757.89	252.63	37.82**
Var x Date	6	146.28	24.38	**3.65***
Error (b)	18	120.33	6.68	
Total	**35**	**1267.22**		

*, ** Significant at $p < 0.05$ and $p < 0.01$, respectively.

Interpretation:
The interaction term is significant, thus mean comparison should be done separately for each date or variety

1.5 Numerical example

Coefficient of variation, cv

- The cv for main plot, cv(a) and sub plot. cv(b) :

$$cv\ (a) = \frac{\sqrt{MSEa}}{Grand\ mean} \times\ 100 = \frac{\sqrt{3.53}}{18.7222} = 10.03\ \%$$

cv(a)
Degree of precision attached to the **main plot factor**

$$cv\ (b) = \frac{\sqrt{MSEb}}{Grand\ mean} \times\ 100 = \frac{\sqrt{6.68}}{18.7222} = 13.80\ \%$$

cv(b)
Degree of precision attached to the **sub plot factor** and **AB interaction**

1.5 Numerical example

Report and summarization – mean values

	Variety				
Date	1	2	3	4	Mean
Oct 15	28.00	16.33	19.67	13.00	19.25
Nov 1	30.00	22.67	17.67	15.67	21.50
Nov 15	18.67	18.67	14.67	9.67	15.42
Mean	25.55	19.22	17.33	12.78	18.72

Standard errors: Date=0.542; Variety=0.862; Variety x Date=1.492

1.5 Numerical example

Visualizing interactions | Method 1

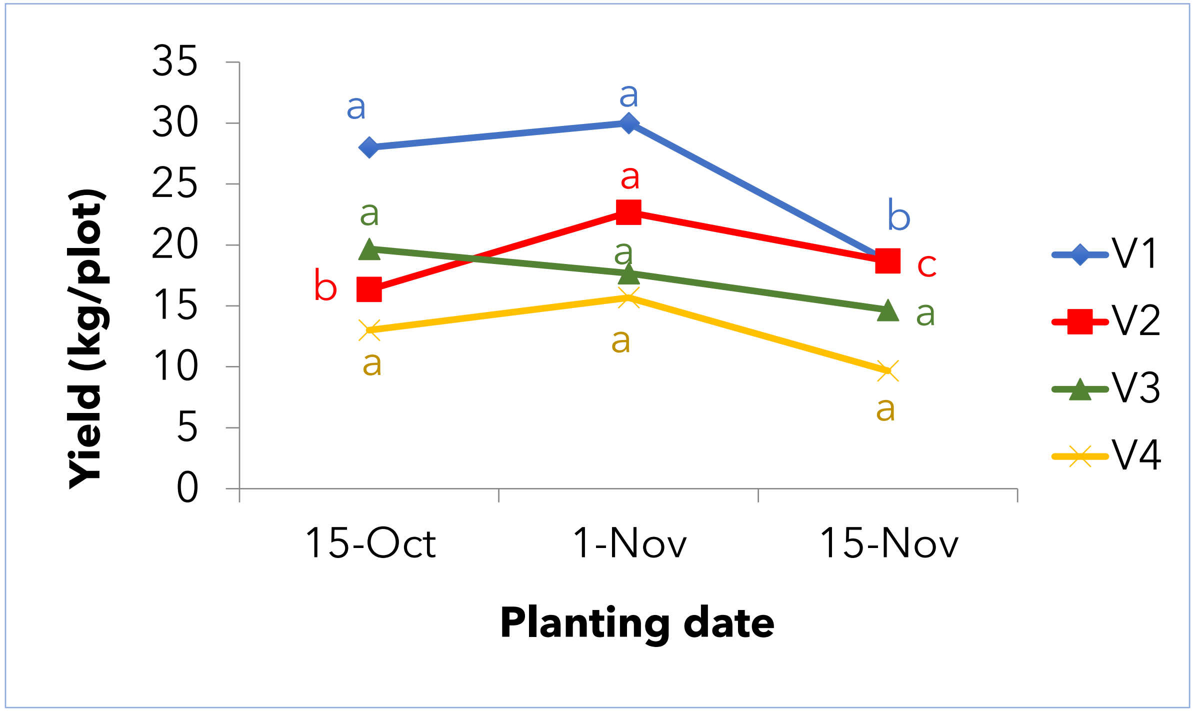


Figure 1. The interaction between planting date and variety on herb yield (kg/plot). **Within variety**, means with different letters are significantly different at $p<0.05$ using LSD.

Interpretations

- Differences among varieties depended on planting date
- Even so, variety differences and date differences were highly significant
- Except for variety 3, each variety produced its maximum yield when planted on 1 November
- For each variety, the highest yield was obtained when grown on 1 Nov. Except for variety 3, all varieties show no significant yield when grown earlier on 15 Oct.
- In fact, V3 and V4 show no significant yield on all planting dates, but has a tendency to reduce yield on 15 Nov.

1.5 Numerical example

Visualizing interactions | Method 2

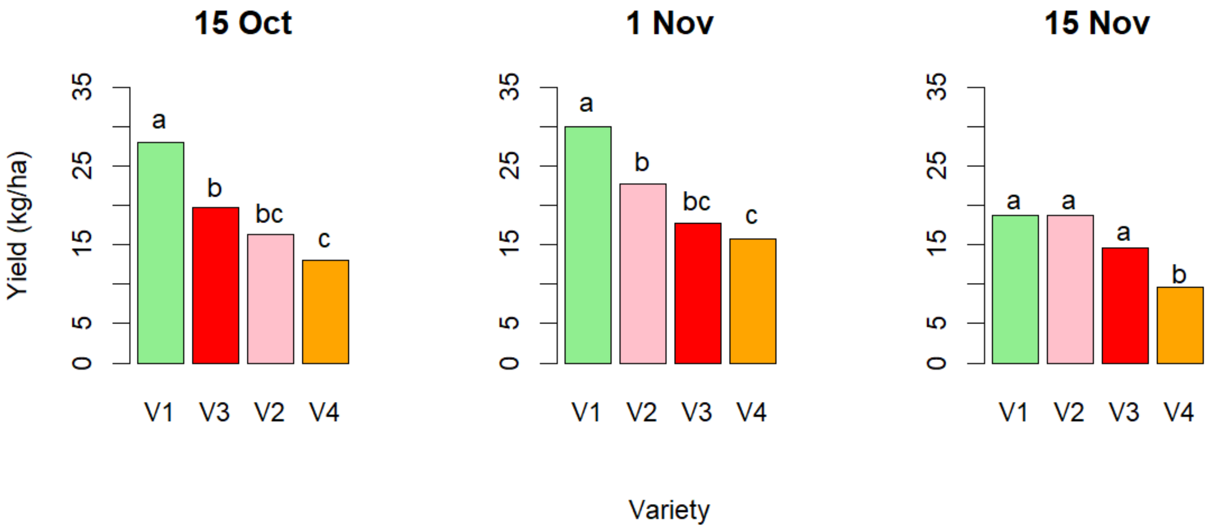


Figure 1. The interaction between planting date and variety on herb yield (kg/ha). **Within planting date**, means with different letters are significant different at $p<0.05$ using LSD.

Interpretations

- Differences among varieties depended on planting date
- Even so, variety differences and date differences were highly significant.
- On Oct 15, the highest yield was observed in V1 and the lowest were V2 and V4
- On 1 Nov, The highest yield was V1 and the lowest were V3 and V4.
- Var 4 produced the lowest yield at each planting date. No significant different between V4 and V2 on Oct 15 and between V4 and V3 on 1 Nov

1.6 Factorial RCBD vs. Split Plot RCBD

FACTORIAL RCBD

V2N4	V1N4	V3N1	V1N2	V3N3	V2N1	V1N3	V1N1	V3N2	V2N3	V2N2	V3N4	**BLOCK 1**
V2N3	V3N1	V3N2	V1N4	V1N2	V3N4	V2N4	V1N1	V1N3	V2N1	V3N3	V2N2	**BLOCK 2**
V2N4	V2N3	V2N1	V2N2	V1N1	V3N1	V3N4	V3N3	V1N4	V3N2	V1N3	V1N2	**BLOCK 3**
V2N2	V1N1	V2N3	V2N1	V1N2	V1N3	V1N4	V3N4	V3N2	V3N3	V2N4	V3N1	**BLOCK 4**

FACTOR 1 = NITROGEN (4)
FACTOR 2 = VARIETY (3)

SPLIT PLOT IN RCBD

N1			N2			N4			N3			
V1	V2	V3	V3	V1	V2	V2	V3	V1	V2	V1	V3	Block 1

N4			N3			N1			N2			
V1	V3	V2	V2	V3	V1	V2	V1	V3	V1	V2	V3	Block 2

N3			N1			N2			N4			
V2	V3	V1	V3	V1	V2	V2	V1	V3	V1	V3	V1	Block 3

N2			N4			N3			N1			
V2	V1	V3	V3	V1	V2	V1	V2	V3	V2	V1	V3	Block 4

Statistical analysis using RStudio

ANOVA
Mean comparison

RStudio

2.0 ANALYSIS USING R

Change data format – wide to long + split plot ANOVA

#Read data into R:

```
split_plot<-read.csv ("sp.plot.data.csv", sep = ",",
header = T)
```

 # Change data format (wide to long data format)

```
library (reshape2)
sp.data<- melt(split_plot, id.vars=c("Date", "Block"))
names(sp.data)<- c("Var", "Rep", "Date", "Yield")
```

#Set variable to factor

```
sp.data$Var<- as.factor(sp.data$Var)
sp.data$Rep<- as.factor(sp.data$Rep)
sp.data$Date<- as.factor(sp.data$Date)
str(sp.data)
```

Wide format

	A	B	C	D	E
1	**Variety**	**Block**	**D1**	**D2**	**D3**
2	**V1**	**R1**	25	30	17
3	**V2**	**R1**	19	24	20
4	**V3**	**R1**	22	19	12
5	**V4**	**R1**	11	15	8
6	**V1**	**R2**	31	32	20
7	**V2**	**R2**	14	20	16
8	**V3**	**R2**	20	18	17
9	**V4**	**R2**	14	13	13
10	**V1**	**R3**	28	28	19
11	**V2**	**R3**	16	24	20
12	**V3**	**R3**	17	16	15
13	**V4**	**R3**	14	19	8

Long format

	A	B	C	D
1	Var	Rep	Date	Yield
2	V1	R1	D1	25
3	V2	R1	D1	19
4	V3	R1	D1	22
5	V4	R1	D1	11
6	V1	R2	D1	31
7	V2	R2	D1	14
8	V3	R2	D1	20
9	V4	R2	D1	14
10	V1	R3	D1	28
11	V2	R3	D1	16
12	V3	R3	D1	17
13	V4	R3	D1	14
14	V1	R1	D2	30
15	V2	R1	D2	24
16	V3	R1	D2	19
17	V4	R1	D2	15
18	V1	R2	D2	32
19	V2	R2	D2	20
20	V3	R2	D2	18
21	V4	R2	D2	13
22	V1	R3	D2	28
23	V2	R3	D2	24
24	V3	R3	D2	16
25	V4	R3	D2	19
26	V1	R1	D3	17
27	V2	R1	D3	20
28	V3	R1	D3	12
29	V4	R1	D3	8
30	V1	R2	D3	20
31	V2	R2	D3	16
32	V3	R2	D3	17
33	V4	R2	D3	13
34	V1	R3	D3	19
35	V2	R3	D3	20
36	V3	R3	D3	15
37	V4	R3	D3	8

2.0 ANALYSIS USING R

R codes – split plot RCBD (AgroR package)

Read data into R:

```
sp.data<-read.csv ("sp.plot.data.csv", sep = ",", header = T)
```

Split plot ANOVA using 'AgroR' package:

```
library (AgroR)
with(sp.data, PSUBDBC(Date, Var, Rep, Yield,
                      ylab="Yield (kg/ha)",
                      xlab = "Planting date",
                      names.fat = c("date", "variety")))
```

Dataset name

Codes for split plot RCBD in AgroR package

Variable names in data.
(Main, sub, block, response)

Label for factor in the output, F1 = date, F2 = variety

Data format

	A	B	C	D
1	Var	Rep	Date	Yield
2	V1	R1	D1	25
3	V2	R1	D1	19
4	V3	R1	D1	22
5	V4	R1	D1	11
6	V1	R2	D1	31
7	V2	R2	D1	14
8	V3	R2	D1	20
9	V4	R2	D1	14
10	V1	R3	D1	28
11	V2	R3	D1	16
12	V3	R3	D1	17
13	V4	R3	D1	14
14	V1	R1	D2	30
15	V2	R1	D2	24
16	V3	R1	D2	19
17	V4	R1	D2	15
18	V1	R2	D2	32
19	V2	R2	D2	20
20	V3	R2	D2	18
21	V4	R2	D2	13
22	V1	R3	D2	28
23	V2	R3	D2	24
24	V3	R3	D2	16
25	V4	R3	D2	19
26	V1	R1	D3	17
27	V2	R1	D3	20
28	V3	R1	D3	12
29	V4	R1	D3	8
30	V1	R2	D3	20
31	V2	R2	D3	16
32	V3	R2	D3	17
33	V4	R2	D3	13
34	V1	R3	D3	19
35	V2	R3	D3	20
36	V3	R3	D3	15
37	V4	R3	D3	8

2.0 ANALYSIS USING R

R output - split plot RCBD (AgroR package)

```
-----------------------------------------------------------
Analysis of Variance
-----------------------------------------------------------
               Df Sum Sq     Mean Sq     F value    Pr(>F)
date            2 227.055556 113.5277778 32.1811024 0.003
Block           2   1.555556   0.7777778  0.2204724 0.811
Error A         4  14.111111   3.5277778
variety         3 757.888889 252.6296296 37.7894737 p<0.001
date : variety  6 146.277778  24.3796296  3.6468144 0.015
Error B        18 120.333333   6.6851852
```

```
-----------------------------------------------------------
Significant interaction: analyzing the interaction
-----------------------------------------------------------
Analyzing  date  inside of each level of  variety
-----------------------------------------------------------
                      GL        SQ         QM        Fc  p.value
date:variety V1  2.00000 219.55556 109.777778 18.619552  1.9e-05
date:variety V2  2.00000  61.55556  30.777778  5.220259 0.014025
date:variety V3  2.00000  38.00000  19.000000  3.222615 0.059369
date:variety V4  2.00000  54.22222  27.111111  4.598351 0.021541
Combined error  21.84744 128.80885   5.895833

-----------------------------------------------------------
Analyzing  variety  inside of the level of  date
-----------------------------------------------------------
               GL       SQ         QM        Fc  p.value
variety:date D1  3 372.9167 124.305556 18.594183    1e-05
variety:date D2  3 367.0000 122.333333 18.299169  1.1e-05
variety:date D3  3 164.2500  54.750000  8.189751 0.001198
Error b         18 120.3333   6.685185
```

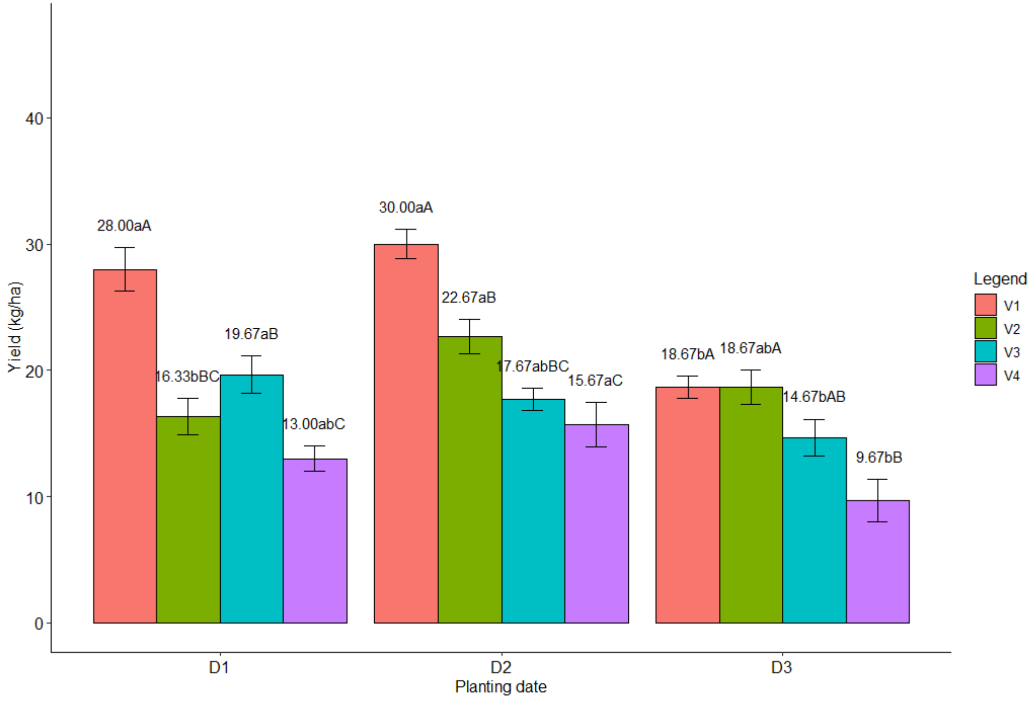
2.0 ANALYSIS USING R

R output - split plot RCBD (AgroR package)

```
-----------------------------------------------
Final table
-----------------------------------------------
         V1        V2         V3        V4
D1 28.00 aA 16.33 bBC   19.67 aB 13.00 abC
D2 30.00 aA  22.67 aB 17.67 abBC  15.67 aC
D3 18.67 bA 18.67 abA  14.67 bAB   9.67 bB
```

```
Averages followed by the same lowercase letter
in the column and uppercase in the row do not
differ by the tukey (p< 0.05 )
```

The graph below is produced based on significant interaction in ANOVA



AgroR always provide the bar graph (default setting) shown on the right. If a different types of graph or different method of mean comparison is needed, then do a separate analysis.

2.0 ANALYSIS USING R

AgroR output interpretation

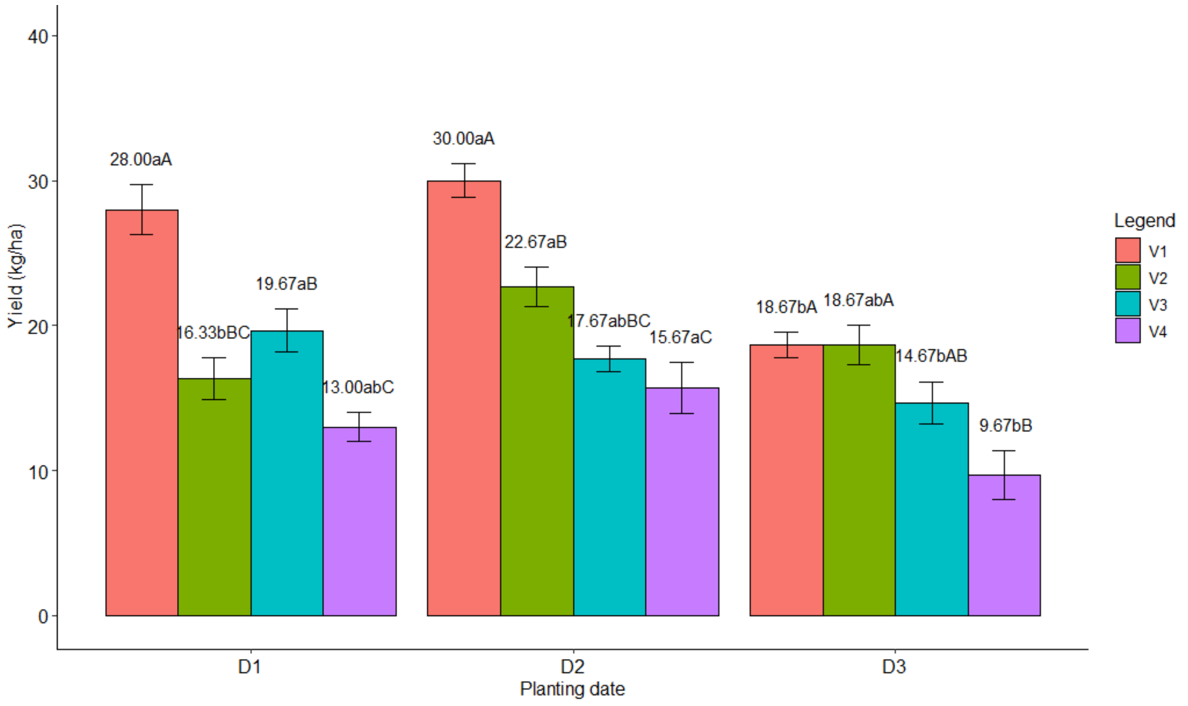


Figure 1. The interaction effects between date and variety. ***Within date, means with different upper letters*** are significantly different at p< 0.05 using Tukey

How to interpret the output?
The output from the AgroR package show that the mean comparison is done both ways (within variety and within date)
To interpret the result, we need to choose one method.

Within date:
The comparison between varieties is denoted by the big letters (A,B, …)

Within variety:
The comparison between dates is denoted by small letter (a, b, …)

You need to mention the meaning of letter groupings displayed on the graph in the footnotes to inform the reader for correct interpretations.

2.0 ANALYSIS USING R

R Studio

R codes: Split plot RCBD (agricolae package)

Read data into R:

```
sp.data<-read.csv ("sp.plot.data.csv", sep = ",", header = T)
```

Split plot ANOVA using 'agricolae' package:

```
library (agricolae)
model<-with(sp.data,sp.plot(Rep, Date, Var, Yield))
```

Data format

	A	B	C	D
1	Var	Rep	Date	Yield
2	V1	R1	D1	25
3	V2	R1	D1	19
4	V3	R1	D1	22
5	V4	R1	D1	11
6	V1	R2	D1	31
7	V2	R2	D1	14
8	V3	R2	D1	20
9	V4	R2	D1	14
10	V1	R3	D1	28
11	V2	R3	D1	16
12	V3	R3	D1	17
13	V4	R3	D1	14
14	V1	R1	D2	30
15	V2	R1	D2	24
16	V3	R1	D2	19
17	V4	R1	D2	15
18	V1	R2	D2	32
19	V2	R2	D2	20
20	V3	R2	D2	18
21	V4	R2	D2	13
22	V1	R3	D2	28
23	V2	R3	D2	24
24	V3	R3	D2	16
25	V4	R3	D2	19
26	V1	R1	D3	17
27	V2	R1	D3	20
28	V3	R1	D3	12
29	V4	R1	D3	8
30	V1	R2	D3	20
31	V2	R2	D3	16
32	V3	R2	D3	17
33	V4	R2	D3	13
34	V1	R3	D3	19
35	V2	R3	D3	20
36	V3	R3	D3	15
37	V4	R3	D3	8

2.0 ANALYSIS USING R

R output

```
ANALYSIS SPLIT PLOT:  Yield
Class level information

Date      :  D1 D2 D3
Var       :  V1 V2 V3 V4
Rep       :  R1 R2 R3

Number of observations:  36
```

Note that the calculation of F value for rep is using $\frac{MSR}{MSEb}$, not $\frac{MSR}{MSEa}$


```
Analysis of Variance Table

Response: yield
         Df Sum Sq Mean Sq F value    Pr(>F)
Rep       2   1.56   0.778  0.1163  0.890833
Date      2 227.06 113.528 32.1811  0.003424 **
Ea        4  14.11   3.528
Var       3 757.89 252.630 37.7895 5.617e-08 ***
Date:Var  6 146.28  24.380  3.6468  0.015144 *
Eb       18 120.33   6.685
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

cv(a) = 10 %, cv(b) = 13.8 %, Mean = 18.72222
```

2.0 ANALYSIS USING R

R codes: Mean comparison using agricolae package → interaction

Method 1 – mean comparison between levels of factor 2 for each level of factor 1

```
#subset the data (based on one of the factor such as 'Date')
date1 <- subset(sp.data, Date == "D1")
date2 <- subset (sp.data, Date == "D2")
date3 <- subset (sp.data, Date == "D3")

#anova and LSD for date1
fitd1<-lm(Yield ~ Rep + Variety , date1); anova (fitd1)
(lsd_d1<-LSD.test(fitd1, "Variety"))

#anova and LSD for date2
fitd2<-lm(Yield ~ Rep + Variety , date2); anova (fitd2)
(lsd_d2<-LSD.test(fitd2, "Variety"))

#anova and LSD for date3
fitd3<-lm(Yield ~ Rep + Variety , date3); anova (fitd3)
(lsd_d3<-LSD.test(fitd3, "Variety"))
```

2.0 ANALYSIS USING R

R output: Mean comparison using agricolae package → interaction

Method 1 – mean comparison between levels of factor 2 for each level of factor 1

```
#Show output for the groups only

lsd_d1$groups
lsd_d2$groups
lsd_d3$groups
```

```
> lsd_d1$groups
      Yield groups
V1 28.00000      a
V3 19.66667      b
V2 16.33333     bc
V4 13.00000      c
```

```
> lsd_d2$groups
      Yield groups
V1 30.00000      a
V2 22.66667      b
V3 17.66667     bc
V4 15.66667      c
```

```
> lsd_d3$groups
       Yield groups
V1 18.666667      a
V2 18.666667      a
V3 14.666667      a
V4  9.666667      b
```

2.0 ANALYSIS USING R

Bar graph for interaction effect using agricolae package

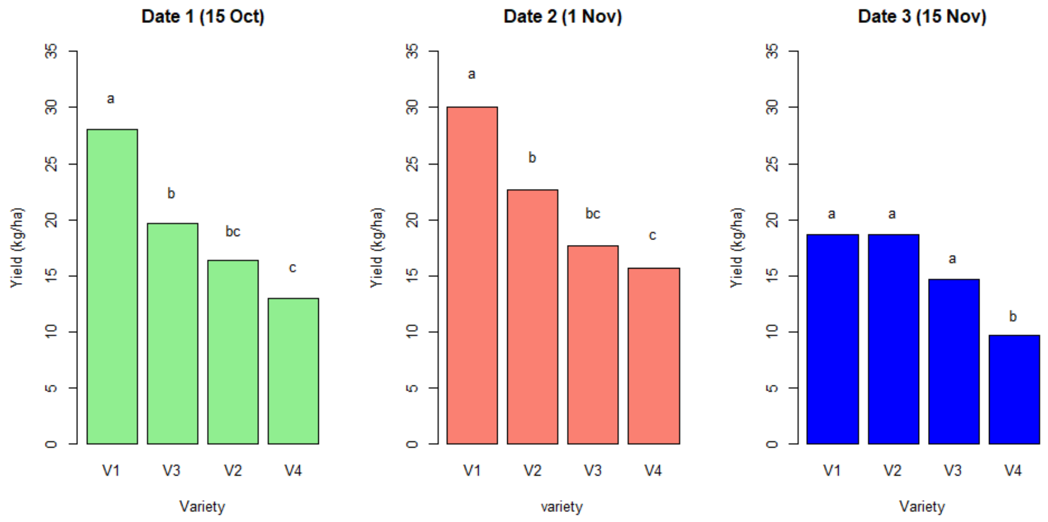
Method 1 – mean comparison between levels of factor 2 for each level of factor 1

```
#====================================#
create bar graph (3 graphs in a page)
#====================================#

par(mfrow=c(1,3))
bar.group(lsd_d1$groups, main = "Date 1 (15 Oct)",
ylim= c(0,35), ylab = "Yield (kg/ha)",
col= "lightgreen", xlab = "Variety")

bar.group(lsd_d2$groups, main = "Date 2 (1 Nov)",
ylim= c(0,35),ylab = "Yield (kg/ha)", col=
"salmon", xlab = "variety")

bar.group(lsd_d3$groups, main = "Date 3 (15 Nov)",
ylim= c(0,35), ylab = "Yield (kg/ha)",
col= "blue", xlab = "Variety")
```



2.0 ANALYSIS USING R

R codes: Mean comparison using agricolae package → interaction

Method 2 – mean comparison between treatment combinations

```
#  MEAN COMPARISON BTWN TRT COMBINATION #
add a new column in the dataset for treatment combination mean comparison
sp.data$trt_comb <- paste(sp.data$var, sp.data$date)
#paste will combine the content in the two columns
sp.data$trt_comb <- as.factor (sp.data$trt_comb)
str(sp.data)

#--- mean comparison of trt combination ---#
# Use MSEb value in the calculation of LSD (or Tukey or dmrt).
MSerror_b <- 6.685 # This is MS Error B from ANOVA
df <- 18 #df residuals
(LSD_trt <- with(sp.data, LSD.test(yield, trt_comb, df, MSerror_b)))
LSD_trt$groups #display only the LSD grouping
```

2.0 ANALYSIS USING R

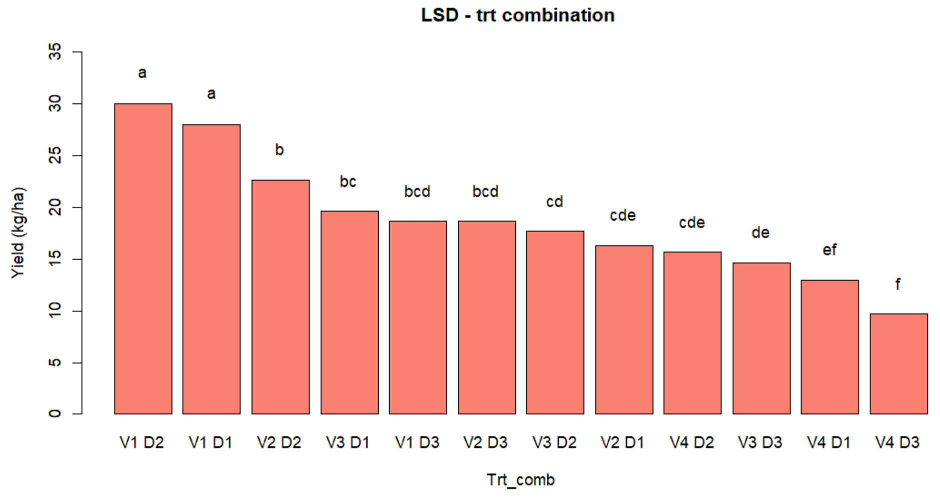
R output: Mean comparison using agricolae package → interaction

Method 2 – mean comparison between treatment combinations

```
> LSD_trt$groups #display only the LSD grouping

          yield groups
V1 D2 30.000000      a
V1 D1 28.000000      a
V2 D2 22.666667      b
V3 D1 19.666667     bc
V1 D3 18.666667    bcd
V2 D3 18.666667    bcd
V3 D2 17.666667     cd
V2 D1 16.333333    cde
V4 D2 15.666667    cde
V3 D3 14.666667     de
V4 D1 13.000000     ef
V4 D3  9.666667      f
```

```
bar.group(LSD_trt$groups,
main = "LSD - trt combination ",
ylim= c(0,35), ylab = "Yield (kg/ha)",
col= "salmon", xlab = "Trt_comb")
```



2.0 ANALYSIS USING R

Mean comparison using agricolae package → no interaction

If the interaction is not significant, the comparison (LSD) is done between levels of main effects only

In split plot, the F test for main effects of main plot uses MSEa (MS Error a) as denominator, thus the error term to calculate the LSD for main plot should use MSEa (default is MSEb).

For example, in the least significant difference (LSD) formula, the MSE should be replaced with MSEa for main plot mean comparison.

```
Analysis of Variance Table

Response: y
            Df  Sum Sq Mean Sq F value    Pr(>F)
rep          3  17.804   5.935  1.9039 0.2300990
fac_A        2 128.765  64.383 20.6547 0.0020399 **
Ea           6  18.703   3.117
fac_B        2 141.296  70.648 13.2088 0.0002947 ***
fac_A:fac_B  4  26.007   6.502  1.2156 0.3387297
Eb          18  96.273   5.349
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

cv(a) = 7.1 %, cv(b) = 9.3 %, Mean = 24.94972
```

$$LSD = t_{\alpha}\sqrt{\frac{2 * MSE}{r}}$$

2.0 ANALYSIS USING R

```
Analysis of Variance Table

Response: y
             Df  Sum Sq Mean Sq F value    Pr(>F)
rep           3  17.804   5.935  1.9039 0.2300990
fac_A         2 128.765  64.383 20.6547 0.0020399 **
Ea            6  18.703   3.117
fac_B         2 141.296  70.648 13.2088 0.0002947 ***
fac_A:fac_B   4  26.007   6.502  1.2156 0.3387297
Eb           18  96.273   5.349
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

cv(a) = 7.1 %, cv(b) = 9.3 %, Mean = 24.94972
```

R codes: Mean comparison using agricolae package

No interaction

LSD for main plot factor (Factor A)

$$LSD_A = t_{\alpha/2}\sqrt{\frac{2 * MSE}{rb}}$$

$$t_{\alpha/2} = t_{0.025,6} = 2.447$$

DF error a

$$LSD_A = 2.447\sqrt{\frac{2 * 3.117}{4 * 3}}$$

$$\boldsymbol{LSD_A = 1.764}$$

LSD for main plot factor (Factor B)

$$LSD_B = t_{\alpha/2}\sqrt{\frac{2 * MSE}{ra}}$$

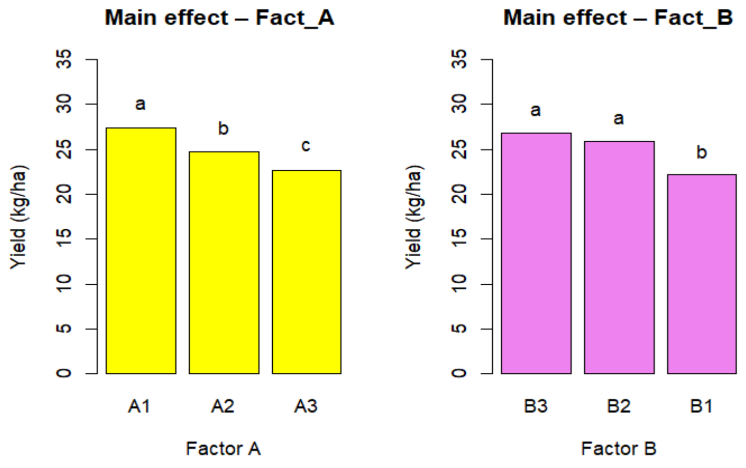
$$t_{\alpha/2} = t_{0.025,18} = 2.101$$

DF error b

$$LSD_B = 2.101\sqrt{\frac{2 * 5.349}{4 * 3}}$$

$$\boldsymbol{LSD_B = 1.984}$$

2.0 ANALYSIS USING R



R codes: Mean comparison using agricolae package

No interaction

```
#1.LSD mean comparison for mainplot (fact_A) main effect

MSerror_a <- 3.117 # This is Error A from anova
dfea <- 6 # df error a
LSD_A <- with(sp_long,LSD.test(yield, Fact_A, dfea,
MSerror_a))
LSD_A

#2.LSD mean comparison for subplot (fact_B) main effect

MSerror_b <- 5.349 # This is Error B from ANOVA
dfeb <- 18 #df error b
LSD_B <- with(sp_long, LSD.test(yield, Fact_B, dfeb,
MSerror_b))
LSD_B
```

```
#bar graph for main effects

par (mfrow = c(1,2))

bar.group(LSD_A$groups, main = "Main
effect - Fact_A ", ylim= c(0,30), ylab
= "Yield (kg/ha)",col= "yellow", xlab =
"Factor A")

bar.group(LSD_B$groups, main = "Main
effect - Fact_B", ylim= c(0,30), ylab =
"Yield (kg/ha)",col= "violet", xlab =
"Factor B")
```

2.0 ANALYSIS USING R

RStudio

LSD calculation in LSD.test of agricolae package - main effects

No interaction

```
> MSerror_a <- 3.117 # This is Error A from anova
> dfea <- 6 # df error a
> LSD_A <- with(sp_long,
+               LSD.test(yield, Fact_A, dfea, MSerror_a))
> LSD_A
$statistics
  MSerror Df     Mean       CV  t.value      LSD
    3.117  6 24.94972 7.076242 2.446912 1.763645
```

LSD for main plot factor (Factor A)

$$LSD_A = t_{\alpha/2}\sqrt{\frac{2*MSE}{rb}}$$

$t_{\alpha/2} = t_{0.025,6} = 2.447$

DF error a

$$LSD_A = 2.447\sqrt{\frac{2*3.117}{4*3}}$$

$$\boldsymbol{LSD_A = 1.764}$$

```
> MSerror_b <- 5.349 # This is Error B from ANOVA
> dfeb <- 18 #df error b
> LSD_B <- with(sp_long,
+               LSD.test(yield, Fact_B, dfeb, MSerror_b))
> LSD_B
$statistics
  MSerror Df     Mean       CV  t.value      LSD
    5.349 18 24.94972 9.269805 2.100922 1.983675
```

LSD for main plot factor (Factor B)

$$LSD_B = t_{\alpha/2}\sqrt{\frac{2*MSE}{ra}}$$

$t_{\alpha/2} = t_{0.025,18} = 2.101$

DF error b

$$LSD_B = 2.101\sqrt{\frac{2*5.349}{4*3}}$$

$$\boldsymbol{LSD_B = 1.984}$$