Randomized Block Designs II

Sean Hellingman ©

Design for Data Science (ADSC2030)

shellingman@tru.ca

Winter 2025

THOMPSON RIVERS UNIVERSITY

Topics

Introduction

- Blocking can be used to keep experimental units as homogeneous as possible
  - Reduces the variance (σ^2) of the experimental error.
  - Allows for better detection of treatment factor effects.

- Blocking can be included to help designs that we have already covered.

Factorial Designs in Blocks

- Blocking is even more effective when combined with a factorial design.
- A **randomized complete block factorial** (RCBF):

  When more than one treatment factor is studied, the number of experimental units in each block must be equal to the product of levels of all the factors.

Road Map

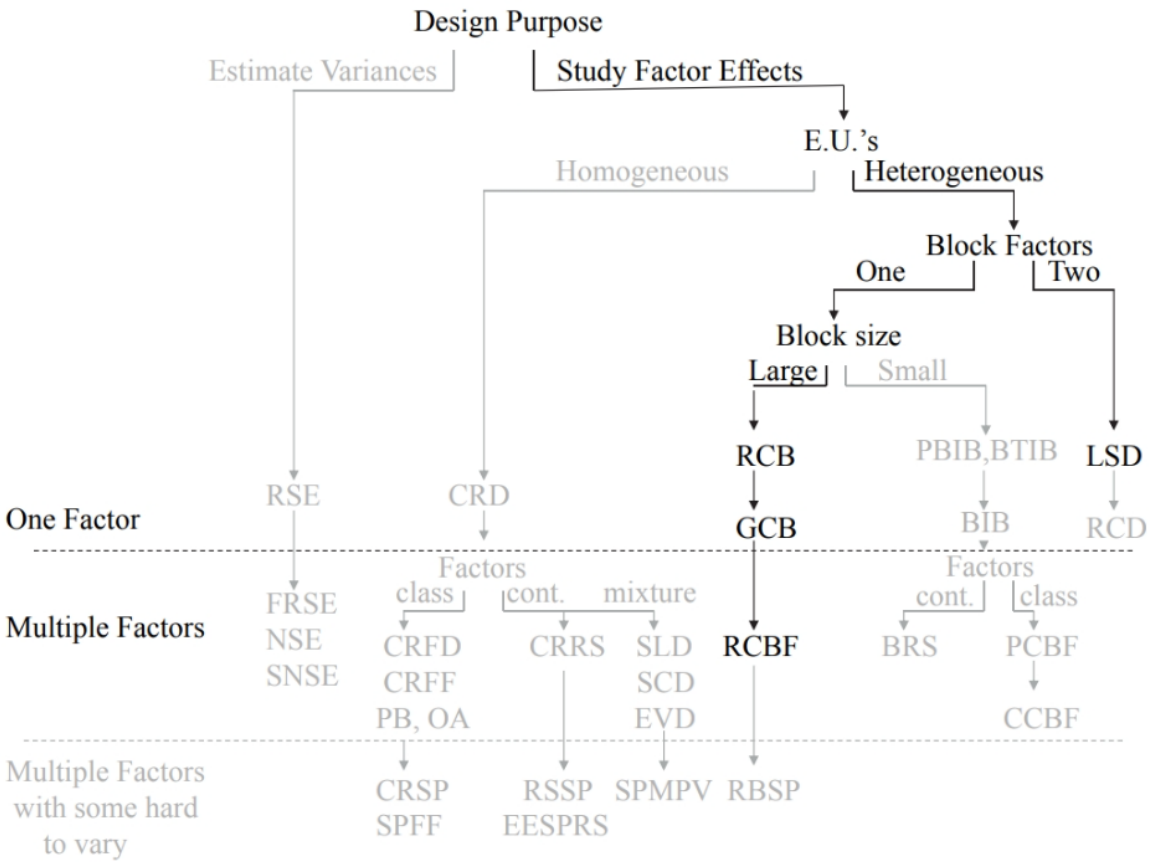

Figure: Source: (1)

Factorial Designs in Blocks Model

- The effects model:

$$Y_{ijk...m} = \mu + b_i + \alpha_j + \beta_k + ... + \gamma_m + \alpha\beta_{jk} + ... + \alpha\beta \cdots \gamma_{jk...m} + \epsilon_{ijk..m}$$

- b_i represents the block effect.
- α_j, β_k, ... γ_m are the main effects.

- Notice: Interactions are included, but not with the blocks (b_i).

Factorial Designs in Blocks Model in R

- In R:
  - `model <- aov(response ~ block + factor.1*factor.2*...*factor.m, data = data)`
  - `summary(model)`
- *Examine the ANOVA table for significant effects*

Example 1

- Import the *bha* data into R.
- Take some time to understand the data.
- Construct an appropriate ANOVA table for the data obtained from the RCBF design.
  - What does the resulting table tell us?
- Visualize the interactions.

Generalized Complete Block Design

- When experimental units represent trials rather than physical entities larger block sizes may be okay.
  - In general, block sizes should be small (more homogeneous).

- **Generalized Complete Block Design** (GCB): A design *with replicates* of each treatment level within a block.
  - *Did not have the required replicates to estimate an interaction with the block before.*

Road Map

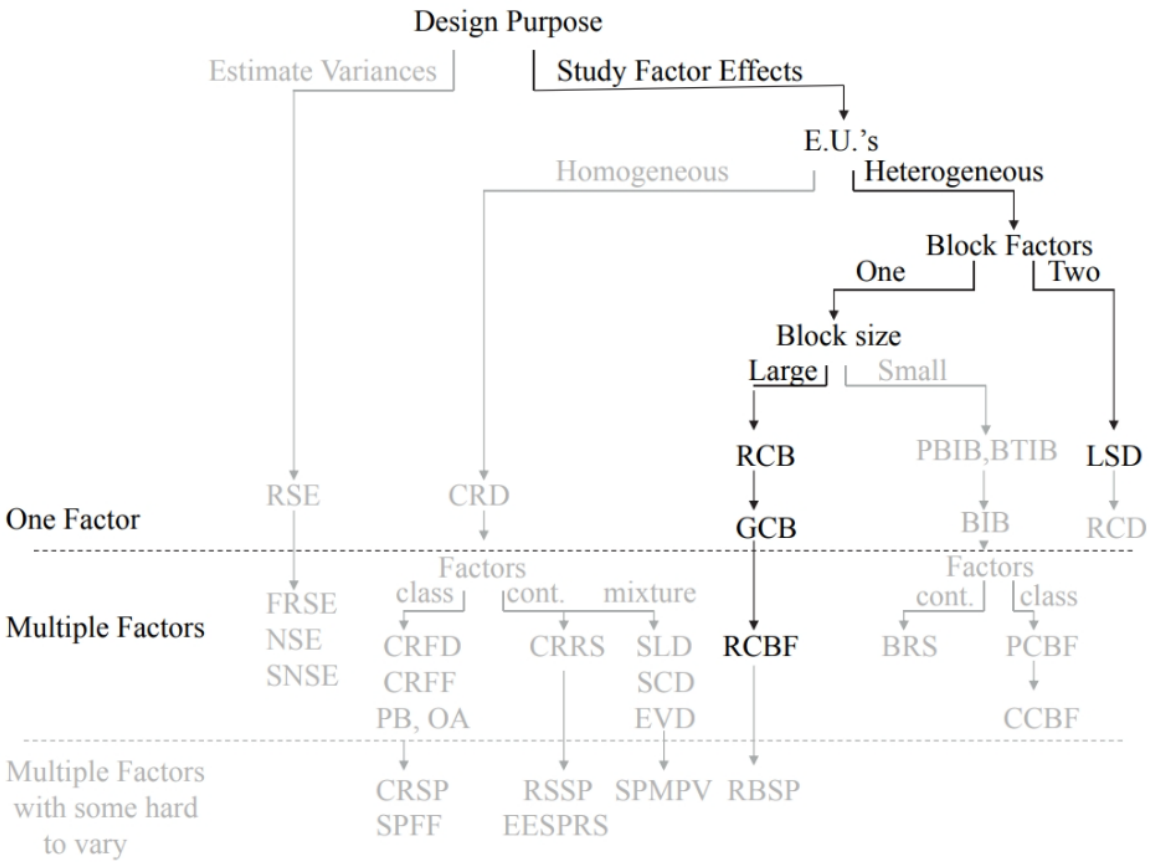

Figure: Source: (1)

Generalized Complete Block Design Model

- The mathematical model for a GCB (one factor):

$$y_{ijk} = \mu + b_i + \tau_j + b\tau_{ij} + \epsilon_{ijk}$$

- b_i is the block effect and τ_j is the treatment effect.

- **This leaves a problem with generalizing the results:**
  - If the interaction is significant, it means that the effects of the treatment are different in each block.

Generalizing the Results of GCB

- *To make a general recommendation, the treatment factor should be tested using the block by treatment interaction mean square as the denominator of the F-test.*

- *If the mean square for treatment factor is significantly larger than the mean square for the interaction between the block and treatment factor, generalizations can be made.*

Generalizing the Results of GCB in R I

- Testing this difference is not automatic in the `aov()` function in R.
- It can be specified:
  - `model <- aov(response ~ factor + Error(block/factor), data = data)`
  - `summary(model)`
- Results:
  - First section: sum of squares and mean square of the block
  - Second section: test for the treatment effect
  - Third section: residual and mean sum of squares (not used in testing)

Example 2

- Import the *rcb* data into R.
- Take some time to understand the data.
- Construct an ANOVA table to determine the significance of an interactions.
  - Can we generalize these results?
- Construct an appropriate ANOVA table that will allow you to generalize the results.
  - What do these results imply?

Generalizing the Results of GCB in R II

- Another way to do this is to average the responses in each block by treatment combination and then fit a RCB model.

- In R:
  - `cellmeans <-tapply( data$response, list(data$block, data$factor), mean)`
  - `dim(cellmeans) <-NULL`
  - `factor <-factor(rep(c(1,2,...,t), each = b) )`
  - `block<-factor(rep(c(1,2,...,b), t) )`

  - `model <-aov(cellmeans ~ block + factor)`

  - `summary(model)`

Using the *RCB with averages* Model

- The model estimated in the previous slide can be used to compare the effects of the factor levels.
- The `model.tables(model, type = "means")` function in R allows you to examine the means.
- The `TukeyHSD(model,"factor")` function in R allows you to compare differences in the means.

Example 3

- Using the *rcb* data in R complete the following tasks:
  1. Construct the alternative (RCB with averages) ANOVA table and interpret the results.
  2. Use the `model.tables(model, type = "means")` function to examine the mean distances of golf shots taken at the three different factor levels.
  3. Use the `TukeyHSD(model,"factor")` function to determine if these differences are significant.

Comments on Generalized Complete Block Designs

- If the interaction is not significant, the additive model can be used to fit the data: $y_{ijk} = \mu + b_i + \tau_j + \epsilon_{ijk}$
- Normally, the preliminary F-test of the interaction is conducted with a higher α ($\alpha = 0.25$).
- If the interaction is *significant*, use the interaction mean squares as the denominator for the F-test (Generalizing the Results of GCB in R I).
- Sometimes called, *Pool or not to Pool.*

Two Block Factors

- Originally proposed for agriculture, to contend with like soil types or conditions.
- A RCB design can be used for agricultural plots if there is a clear fertility gradient in the soil.
- If adjacent plots tend to be more alike than plots in the same block a different approach is needed.

Latin-square Block Designs

- **Latin-square design** (LSD) is blocked both horizontally and vertically.
- Each treatment level is assigned only once to each row and each column.
- The number of row blocks **must** equal the number of column blocks.

Road Map

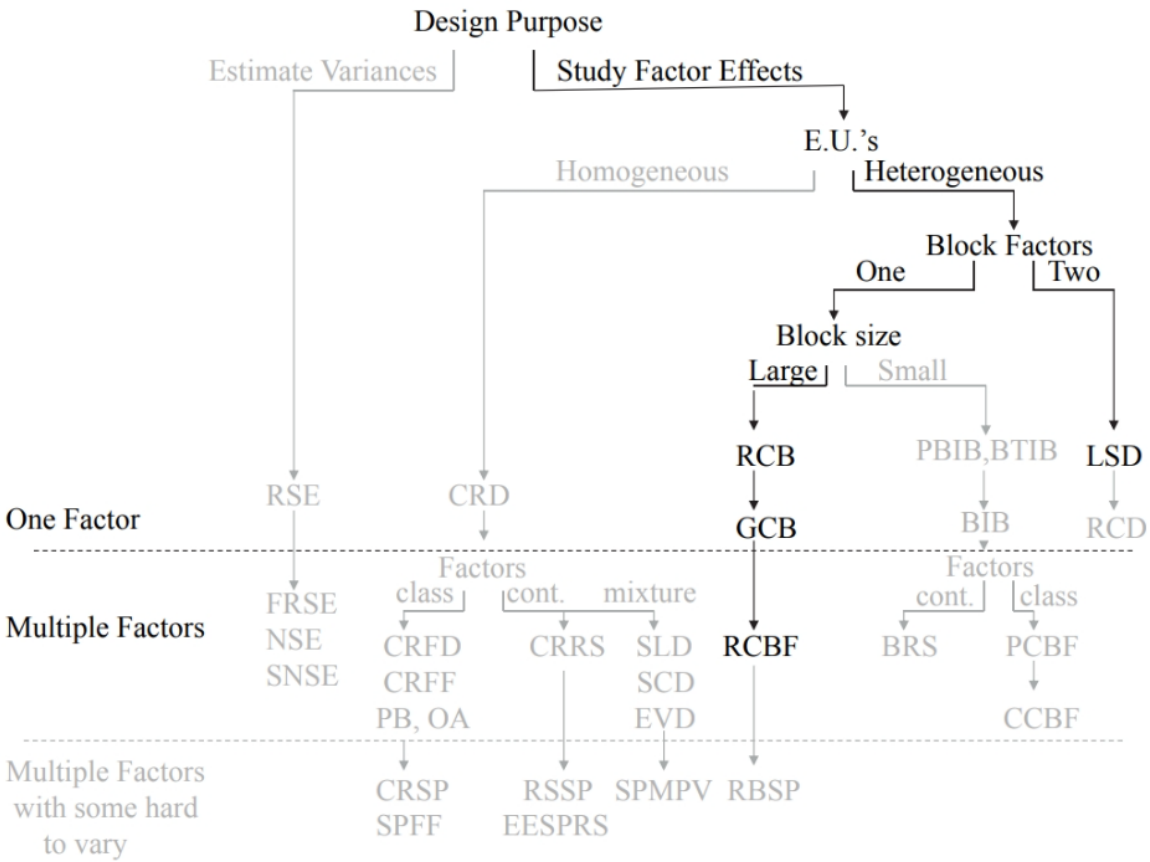

Figure: Source: (1)

Latin-square Block Designs Illustrative Example

RCB

Block				
1	B	C	D	A
2	A	D	C	B
3	B	D	A	C
4	C	B	A	D

Latin Square

B	D	C	A
D	C	A	B
C	A	B	D
A	B	D	C

Figure: Source: (1)

LSD Model

- The model for an LSD:

$$y_{ijk} = \mu + r_i + c_j + \tau_k + \epsilon_{ijk} \tag{1}$$

- r_i represents the row of the blocking factor.
- c_j represents the column of the blocking factor.
- τ_k represents the treatment factor.

- **No interaction terms are included to generalize the results**

LSD Comments

- Latin-square designs may be used when there are two independent blocking factors.
- Illustrative Example:
  - Experiment to determine the effect of tread design on the wear life of automobile tires.
    - EU: A wheel on the car.
    - Treatment: The tread design of the mounted tire.
    - Block 1: Type of automobile (weights may impact wear)
    - Block 2: Wheel position on car (front differs from back)
- The number of row blocks **must** equal the number of column blocks.

Creating a LSD in R

- Randomisation is still an important step in creating this kind of design.

- Creating a LSD in R:
  - `library(agricolae)`
  - `treatments <- c(1,2,...,t)`
  - `outdesign <- design.lsd(treatments, seed = 2030)`
  - `lsd <- outdesign$book`
  - `levels(lsd$row) <- c("Block1.A",...,"Block1.R")`
  - `levels(lsd$col) <- c("Block2.A",...,"Block2.C")`

- The number of row blocks **must** equal the number of column blocks.

Example 4

- Use R to create a LSD for the tire tread illustrative example.
- Assume the following:
  - Block 1 (Automobile): Truck, Van, SUV, Car
  - Block 2 (Wheel position): FR, FL, BR, BL
  - Treatments: Winter, All-season, Summer, Wet

Analysis of LSD in R

- We can analyse the data from a LSD in R:
  - `model <- aov(response ~ Block.row + Block.column + factor, data = data)`
  - `summary(model)`
- We can also examine the means and test the differences in factor levels:
  - `model.tables(model, type = "means")`
  - `TukeyHSD(model,"factor")`

Example 5

- Import the *bioeqv* data into R.
- Take some time to get to know the data.
- Construct an ANOVA table to examine the effects of the different treatment levels.
- Use the `model.tables(model, type = "means")` and `TukeyHSD(model,"factor")` functions to verify your results.

Number of Replicates

- The number of replicates for each treatment factor level in an LSD with t rows and t columns must equal t.
- In Example 5, we could increase the power of the test by increasing the number of subjects (replicate the entire square).
- Replicated Latin square has $r = nt$ rows, t columns, and t factor levels.

Exercise 1

- Using the *GCBD.csv* data conduct the following analysis:
  1. Construct an ANOVA table to determine the significance of an interactions.
     - Can we generalize these results?
  2. Construct an appropriate ANOVA table that will allow you to generalize the results.
     - What do these results imply?

Exercise 2

- Using the *GCBD.csv* data and the **RCB with averages Model** conduct the following analysis:
  1. Construct the alternative (RCB with averages) ANOVA table and interpret the results.
  2. Use the `model.tables(model, type = "means")` function to examine the mean distances of golf shots taken at the three different factor levels.
  3. Use the `TukeyHSD(model,"factor")` function to determine if these differences are significant.

References & Resources

1. Lawson, J. (2014). *Design and Analysis of Experiments with R (Vol. 115)*. CRC press.

- `design.rcbd()`
- `contr.poly()`
- `Fpower()`