

Case Study : Augmented Block Design. Analysis with Fixed Model

Prof L. Gentzbittel Skoltech, Project Center for Agro Technologies *

Prof C. Ben, Skoltech, Project Center for Agro Technologies †

Oct. 20th , 2021 - Skoltech

CASE STUDY PRESENTATION

A toy dataset : three blocks, 4 checks and 8 new varieties

PREPARATION OF THE WORKING INTERFACE IN R

```
### I. Set working directory ###
# On RStudio: tab 'Session' -> Set Working Directory -> Choose Directory.
# Choose the directory containing the Alpha_latticefile and the associated R script.

### II. Possibly, installation of new R packages needed for the analysis on RStudio:
# Click on the 'Packages' tab in the bottom-right window of R Studio interface -> 'Install Packages'
# Comment #1: R package installation requires a connection to internet
# Comment #2: Once packages have been installed,
# no need to re-install them again when you close-open again RStudio.

### III. Initialisation of the working space
# To erase all graphs
graphics.off()
# To erase objects from the working space - Clean up of the memory
rm(list = ls())

# this is a trick to detect which folder contains the R script and the Alpha_lattice
main_dir <- dirname(rstudioapi::getSourceEditorContext()$path)
setwd(main_dir)
```

LOADING REQUIRED METHODS FOR ANALYSIS

```
library(agricolae)
library(emmeans)
library(ggplot2) ; library(gridExtra)

# constraints on effects for ANOVA
options(contrasts=c("contr.sum", "contr.poly"))
```

*l.gentzbittel@skoltech.ru

†c.ben@skoltech.ru

ANALYSIS OF THE CASE STUDY

The case study will be analysed using a classical method based on fixed model (GLM)

```
## ### the data, with map of the field
```

```
Produc <- read.table("AugmentedBlockDesign1b.csv", sep = ";", dec = ".", header = TRUE)
str(Produc)
```

```
## 'data.frame':    20 obs. of  6 variables:
## $ Accession: chr  "C1" "C1" "C1" "C2" ...
## $ Type      : chr  "check" "check" "check" "check" ...
## $ Bloc      : chr  "B1" "B2" "B3" "B1" ...
## $ weight    : int  10 8 6 3 5 6 10 9 8 15 ...
## $ x         : int  2 5 8 4 5 9 1 7 8 4 ...
## $ y         : int  2 2 2 2 1 2 1 2 1 1 ...
```

```
### the trial ?
```

```
with(Produc,
      table(Accession,Bloc))
```

```
##           Bloc
## Accession B1 B2 B3
##          C1  1  1  1
##          C2  1  1  1
##          C3  1  1  1
##          C4  1  1  1
##          V1  1  0  0
##          V2  0  1  0
##          V3  0  1  0
##          V4  1  0  0
##          V5  0  0  1
##          V6  0  0  1
##          V7  0  0  1
##          V8  1  0  0
```

```
## visualize the field:
```

```
library(desplot)
x11(width = 7, height = 5)
desplot(Bloc ~ x + y, Produc,
        col = Type, text = Accession, cex = 1,
        out1 = Bloc, out2 = Accession,
        out2.gpar=list(col = "gray50", lwd = 1, lty = 1),
        main = "Augmented Block Design - Example 1")
```

	C1	V8	C2	C1	V2	C3	C1	C2	V5	V6
C3	V1	V4	C4	C2	V3	C4	C3	C4	V7	

```
##          Df Sum Sq Mean Sq F value    Pr(>F)
## Accession 11 1673.2   152.11   40.265 0.000102 ***
## Bloc      2     0.7     0.33    0.088 0.916711
## Residuals 6    22.7     3.78
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## the two ANOVA are different because some combinations are lacking
```

```
##### Analysis of the checks only #####
# this is a simple RCBD
```

```
## SS of accessions corrected for block effect
```

```
AnaChecks <- aov(weight ~ Bloc + Accession , data = subset(Produc, subset = Type == "check"))
summary(AnaChecks)
```

```
##          Df Sum Sq Mean Sq F value Pr(>F)
## Bloc      2   0.67    0.33   0.088 0.91671
## Accession  3 191.33   63.78  16.882 0.00249 **
## Residuals  6  22.67    3.78
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##### Phenotypic values of new entries, "corrected for the block effect"
```

the *block effect* for correction is estimated in the ANOVA of the checks

```
## What are the block effects ?
```

```
model.tables(AnaChecks)
```

```
## Tables of effects
```

```
##
```

```
## Bloc
```

```
## Bloc
```

```
##      B1      B2      B3
```

```
## 0.1667 -0.3333 0.1667
```

```
##
```

```
## Accession
```

```
## Accession
```

```
##      C1      C2      C3      C4
```

```
## -1.333 -4.667 -0.333 6.333
```

```
model.tables(AnaChecks)$tables$Bloc
```

```
## Bloc
```

```
##      B1      B2      B3
```

```
## 0.1666667 -0.3333333 0.1666667
```

```
### Syntax to correct the phenotypic values of the 'new entries'
```

```
Produc$EffetBloc <- sapply(Produc$Bloc, switch,
```

```
    B1=model.tables(AnaChecks)$tables$Bloc[1],
```

```
    B2=model.tables(AnaChecks)$tables$Bloc[2],
```

```
    B3=model.tables(AnaChecks)$tables$Bloc[3]
```

```
)
```

```
## use of the switch() fnction
```

```
## the Bi block effect is taken from
```

```
## the results of ANOVA table
```

```
## create the vector of all data "corrected for the block effect"
```

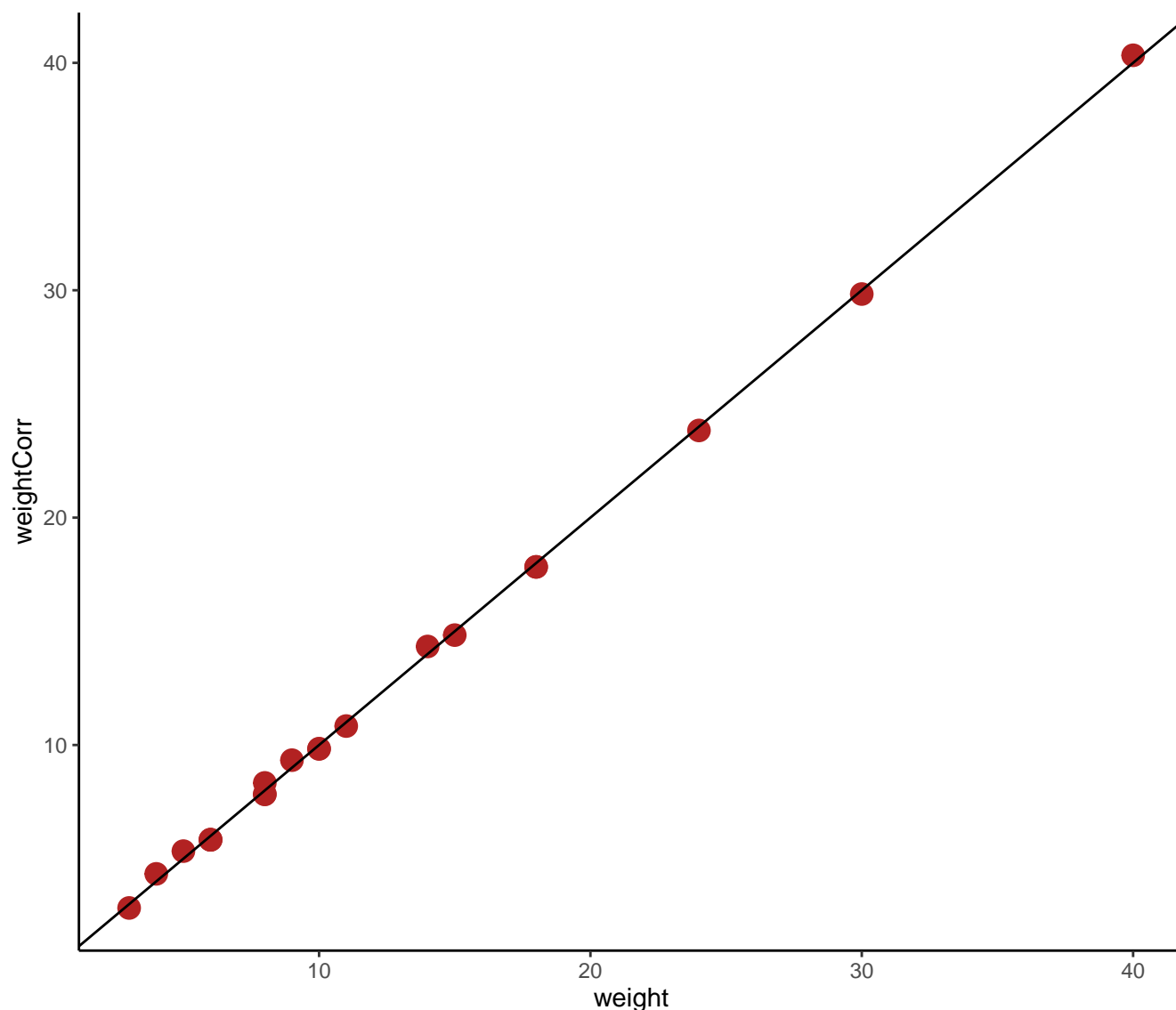
```
Produc$weightCorr <- Produc$weight - Produc$EffetBloc
```

```
Produc
```

```
##      Accession  Type Bloc weight  x y  EffetBloc weightCorr
## 1          C1  check  B1     10  2  2  0.1666667   9.833333
## 2          C1  check  B2      8  5  2 -0.3333333   8.333333
## 3          C1  check  B3      6  8  2  0.1666667   5.833333
## 4          C2  check  B1      3  4  2  0.1666667   2.833333
## 5          C2  check  B2      5  5  1 -0.3333333   5.333333
## 6          C2  check  B3      6  9  2  0.1666667   5.833333
```

```
## 7      C3  check  B1    10  1  1  0.1666667  9.833333
## 8      C3  check  B2     9  7  2 -0.3333333  9.333333
## 9      C3  check  B3     8  8  1  0.1666667  7.833333
## 10     C4  check  B1    15  4  1  0.1666667 14.833333
## 11     C4  check  B2    14  7  1 -0.3333333 14.333333
## 12     C4  check  B3    18  9  1  0.1666667 17.833333
## 13     V1 variety  B1     6  2  1  0.1666667  5.833333
## 14     V2 variety  B2    40  6  2 -0.3333333 40.333333
## 15     V3 variety  B2     4  6  1 -0.3333333  4.333333
## 16     V4 variety  B1    24  3  1  0.1666667 23.833333
## 17     V5 variety  B3    11 10  2  0.1666667 10.833333
## 18     V6 variety  B3    18 11  2  0.1666667 17.833333
## 19     V7 variety  B3     8 10  1  0.1666667  7.833333
## 20     V8 variety  B1    30  3  2  0.1666667 29.833333
```

```
x11()
(gr1 <- ggplot(Produc) + aes(x = weight, y = weightCorr) +
  geom_point(color = 'firebrick', cex = 4) +
  geom_abline(slope = 1, intercept = 0) +
  theme_classic()
)
```



```
##### to verify / validate :
```

*## CAUTION : This verification works ONLY for 'checks' because the design is balanced for checks
so the SS of accessions is independent from the SS of block: we and do the ANOVA in both direction
and it shall produce the SAME results*

```
Valid1 <- aov(weightCorr ~ Bloc + Accession, data = subset(Produc, subset = Type == "check"))
summary(Valid1)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Bloc      2  0.00    0.00    0.00 1.00000
## Accession 3 191.33   63.78   16.88 0.00249 **
## Residuals 6  22.67    3.78
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Valid2 <- aov(weightCorr ~ Accession + Bloc, data = subset(Produc, subset = Type == "check"))
summary(Valid2)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Accession 3 191.33   63.78   16.88 0.00249 **
## Bloc      2  0.00    0.00    0.00 1.00000
## Residuals 6  22.67    3.78
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Are the results expected ?

*##### the above procedure can be simply done by computing "marginal means"
also called 'lsmeans'*

```
emmeans(Ana1a, ~ Accession)
```

```
## Accession emmean SE df lower.CL upper.CL
## C1        8.00 1.12 6    5.254    10.75
## C2        4.67 1.12 6    1.921     7.41
## C3        9.00 1.12 6    6.254    11.75
## C4       15.67 1.12 6   12.921    18.41
## V1        5.83 2.10 6    0.696    10.97
## V2       40.33 2.10 6   35.196    45.47
## V3        4.33 2.10 6   -0.804     9.47
## V4       23.83 2.10 6   18.696    28.97
## V5       10.83 2.10 6    5.696    15.97
## V6       17.83 2.10 6   12.696    22.97
## V7        7.83 2.10 6    2.696    12.97
## V8       29.83 2.10 6   24.696    34.97
##
## Results are averaged over the levels of: Bloc
## Confidence level used: 0.95
```

*##### what accessions differ from checks ?
débobinage : quelles sont les accessions différentes des checks ?*

just use the formulas of the slides:

```
c <- 4
r <- 3
ddl <- (r-1)*(c-1)
MSE <- summary(Ana1a)[[1]]$'Mean Sq'[3] # 3rd term of 'Mean Sq' column of the summary table of ANOVA
```

SE of the difference between 2 adj means of selections in different blocks (cf dias)

```
Sv <- sqrt(2*(c+1)*MSE/c)
```

#vSE of the difference between adjusted selection mean and check. c'est la quantité utile à connaître !

```

Svc <- sqrt(((r+1)*(c+1)*MSE)/(r*c))

# We are , for example, interested to those corrected vaues that outperform the checks.
## the CI is given following a t distribution

# critical value of t:
qt(0.95,ddl)

## [1] 1.94318

## numerical value of Least Significant Interval:
LSI <- qt(0.95,ddl) * Svc
LSI

## [1] 4.87591

# so any new entry that outperforms the 'best check' by this quantity is significantly different from it.
Produc

##      Accession      Type Bloc weight  x y  EffetBloc weightCorr
## 1          C1    check   B1     10  2 2  0.1666667   9.833333
## 2          C1    check   B2      8  5 2 -0.3333333   8.333333
## 3          C1    check   B3      6  8 2  0.1666667   5.833333
## 4          C2    check   B1      3  4 2  0.1666667   2.833333
## 5          C2    check   B2      5  5 1 -0.3333333   5.333333
## 6          C2    check   B3      6  9 2  0.1666667   5.833333
## 7          C3    check   B1     10  1 1  0.1666667   9.833333
## 8          C3    check   B2      9  7 2 -0.3333333   9.333333
## 9          C3    check   B3      8  8 1  0.1666667   7.833333
## 10         C4    check   B1     15  4 1  0.1666667  14.833333
## 11         C4    check   B2     14  7 1 -0.3333333  14.333333
## 12         C4    check   B3     18  9 1  0.1666667  17.833333
## 13         V1  variety   B1      6  2 1  0.1666667   5.833333
## 14         V2  variety   B2     40  6 2 -0.3333333  40.333333
## 15         V3  variety   B2      4  6 1 -0.3333333   4.333333
## 16         V4  variety   B1     24  3 1  0.1666667  23.833333
## 17         V5  variety   B3     11 10 2  0.1666667  10.833333
## 18         V6  variety   B3     18 11 2  0.1666667  17.833333
## 19         V7  variety   B3      8 10 1  0.1666667   7.833333
## 20         V8  variety   B1     30  3 2  0.1666667  29.833333

So, what is the list of the new entreis that outperfrom the best check ?

#####
##### Analysis using agricolae #####
#####

## another elegant way to see the design:
with(Produc,
by(Accession, Bloc, as.character)
)

## Bloc: B1
## [1] "C1" "C2" "C3" "C4" "V1" "V4" "V8"
## -----
## Bloc: B2
## [1] "C1" "C2" "C3" "C4" "V2" "V3"
## -----
## Bloc: B3
## [1] "C1" "C2" "C3" "C4" "V5" "V6" "V7"

```

```
## the agricolae::DAU() fits augmented design

modelDAU1 <- DAU.test(Produc$Bloc, Produc$Accession, Produc$weight,
                      method = "lsd", group = TRUE,
                      console = TRUE)

##
## ANALYSIS DAU: Produc$weight
## Class level information
##
## Block:  B1 B2 B3
## Trt   :  C1 C2 C3 C4 V1 V2 V3 V4 V5 V6 V7 V8
##
## Number of observations:  20
##
## ANOVA, Treatment Adjusted
## Analysis of Variance Table
##
## Response: Produc$weight
##
##           Df Sum Sq Mean Sq F value    Pr(>F)
## block.unadj      2   41.79   20.894
## trt.adj          11 1632.10  148.372   39.275 0.0001096 ***
## Control           3   191.33   63.778   16.882 0.0024938 **
## Control + control.VS.aug.  8 1440.76  180.095   47.672 7.176e-05 ***
## Residuals         6    22.67    3.778
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ANOVA, Block Adjusted
## Analysis of Variance Table
##
## Response: Produc$weight
##
##           Df Sum Sq Mean Sq F value    Pr(>F)
## trt.unadj      11 1673.22  152.11
## block.adj        2    0.67    0.33  0.0882 0.9167114
## Control          3   191.33   63.78  16.8824 0.0024938 **
## Augmented         7 1151.88  164.55  43.5583 0.0001001 ***
## Control vs augmented  1  330.01  330.01  87.3551 8.504e-05 ***
## Residuals         6    22.67    3.78
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## coefficient of variation: 15.4 %
## Produc$weight Means: 12.65
##
## Critical Differences (Between)
##
##                               Std Error Diff.
## Two Control Treatments          1.586984
## Two Augmented Treatments (Same Block)  2.748737
## Two Augmented Treatments(Different Blocks)  3.073181
## A Augmented Treatment and A Control Treatment  2.380476
##
##
## Treatments with the same letter are not significantly different.
##
##   Produc$weight groups
## V2      40.333333      a
## V8      29.833333      b
## V4      23.833333     bc
```



```
## V6      17.833333      cd
## C4      15.666667      de
## V5      10.833333      ef
## C3       9.000000       f
## C1       8.000000      fg
## V7       7.833333      fg
## V1       5.833333      fg
## C2       4.666667       g
## V3       4.333333       g
##
## Comparison between treatments means
##
## <<< to see the objects: comparison and means >>>

## The table "ANOVA, Treatment Adjusted" is the same as Ana1a ( Bloc + Accessions)
## the table "ANOVA, Block Adjusted" is the same as Ana1b (Accessions + Bloc)

## il y a une petite diff numérique sur Svc entre agricolae et calculs manuels

options(digits = 2)
modelDAU1$means

##      Produc.weight std r Min Max  Q25 Q50  Q75 mean.adj  SE block
## C1          8.0 2.0 3   6 10  7.0  8  9.0      8.0 1.1
## C2          4.7 1.5 3   3  6  4.0  5  5.5      4.7 1.1
## C3          9.0 1.0 3   8 10  8.5  9  9.5      9.0 1.1
## C4         15.7 2.1 3  14 18 14.5 15 16.5     15.7 1.1
## V1          6.0 NA 1   6  6  6.0  6  6.0      5.8 1.9    B1
## V2         40.0 NA 1  40 40 40.0 40 40.0     40.3 1.9    B2
## V3          4.0 NA 1   4  4  4.0  4  4.0      4.3 1.9    B2
## V4         24.0 NA 1  24 24 24.0 24 24.0     23.8 1.9    B1
## V5         11.0 NA 1  11 11 11.0 11 11.0     10.8 1.9    B3
## V6         18.0 NA 1  18 18 18.0 18 18.0     17.8 1.9    B3
## V7          8.0 NA 1   8  8  8.0  8  8.0      7.8 1.9    B3
## V8         30.0 NA 1  30 30 30.0 30 30.0     29.8 1.9    B1

### see lsmeans and weightCorr to compare

## If one wants pairwise comparisons détail : si on veut les comparaisons deux à deux, et non pas les gro
modelDAU2 <- DAU.test(Produc$Bloc, Produc$Accession, Produc$weight,
                      method="lsd", group = FALSE,
                      console = TRUE)

##
## ANALYSIS DAU: Produc$weight
## Class level information
##
## Block:  B1 B2 B3
## Trt   :  C1 C2 C3 C4 V1 V2 V3 V4 V5 V6 V7 V8
##
## Number of observations:  20
##
## ANOVA, Treatment Adjusted
## Analysis of Variance Table
##
## Response: Produc$weight
##
##      Df Sum Sq Mean Sq F value    Pr(>F)
## block.unadj      2      42     20.9
## trt.adj         11    1632    148.4     39.3 0.00011 ***
```

```
## Control          3    191    63.8    16.9 0.00249 **
## Control + control.VS.aug. 8   1441   180.1    47.7 7.2e-05 ***
## Residuals        6     23     3.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ANOVA, Block Adjusted
## Analysis of Variance Table
##
## Response: Produc$weight
##              Df Sum Sq Mean Sq F value    Pr(>F)
## trt.unadj      11   1673      152
## block.adj       2      1        0    0.09  0.9167
## Control         3    191        64   16.88  0.0025 **
## Augmented       7   1152       165   43.56  0.0001 ***
## Control vs augmented 1    330       330   87.36  8.5e-05 ***
## Residuals      6     23        4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## coefficient of variation: 15 %
## Produc$weight Means: 13
##
## Critical Differences (Between)
##                                     Std Error Diff.
## Two Control Treatments              1.6
## Two Augmented Treatments (Same Block) 2.7
## Two Augmented Treatments(Different Blocks) 3.1
## A Augmented Treatment and A Control Treatment 2.4
##
##
## Comparison between treatments means
##
## <<< to see the objects: comparison and means >>>
```

```
head(modelDAU2$comparison, 12)
```

```
##          Difference pvalue sig.
## C1 - C2      3.33 0.0804      .
## C1 - C3     -1.00 0.5518
## C1 - C4     -7.67 0.0030    **
## C1 - V1      2.17 0.3978
## C1 - V2    -32.33 0.0000   ***
## C1 - V3      3.67 0.1744
## C1 - V4    -15.83 0.0006   ***
## C1 - V5     -2.83 0.2790
## C1 - V6     -9.83 0.0062    **
## C1 - V7      0.17 0.9464
## C1 - V8    -21.83 0.0000   ***
## C2 - C3     -4.33 0.0342     *
```

```
## en conclusion : agricolae est utile et juste pour l'analyse des plans en blocs augmentés.
## Il utilise un modèle fixe d'ANOVA.
## on peut raffiner un peu mieux que agricolae.
```

```
#####
##### Analysis using plantbreeding
#####
```

```

install.packages("plantbreeding", repos="http://R-Forge.R-project.org") ## unclear if maintained
## requires reshape

library(plantbreeding)
Ana3 <- aug.rcb(dataframe = Produc, genotypes = "Accession", block = "Bloc", yvar = "weight")

## Phenotypes and adjusted values :
##
##   Accession   Type Bloc weight  x y EffetBloc weightCorr yvar.adj
## 13         V1 variety  B1      6  2 1      0.17      5.8      5.8
## 14         V2 variety  B2     40  6 2     -0.33     40.3     40.3
## 15         V3 variety  B2      4  6 1     -0.33      4.3      4.3
## 16         V4 variety  B1     24  3 1      0.17     23.8     23.8
## 17         V5 variety  B3     11 10 2      0.17     10.8     10.8
## 18         V6 variety  B3     18 11 2      0.17     17.8     17.8
## 19         V7 variety  B3      8 10 1      0.17      7.8      7.8
## 20         V8 variety  B1     30  3 2      0.17     29.8     29.8

## Standard error of different comparisons
##
## Difference between check means: 1.6
##
## Difference adjusted yield of two varieties in same block : 2.7
##
## Difference between two varieties in different blocks: 3.1
##
## Difference between two varieties and a check mean: 2.5

# analysis of variance
Ana3$anova

##           Df Sum Sq Mean Sq F value Pr(>F)
## Genotypes  3  191.3    63.8   16.88 0.0025 **
## Block      2    0.7     0.3    0.09 0.9167
## Residual   6   22.7     3.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## compare with Ana2 and modelDAU1

Ana3$adjusted_values # yield observed and expected value table

##   Accession   Type Bloc weight  x y EffetBloc weightCorr yvar.adj
## 13         V1 variety  B1      6  2 1      0.17      5.8      5.8
## 14         V2 variety  B2     40  6 2     -0.33     40.3     40.3
## 15         V3 variety  B2      4  6 1     -0.33      4.3      4.3
## 16         V4 variety  B1     24  3 1      0.17     23.8     23.8
## 17         V5 variety  B3     11 10 2      0.17     10.8     10.8
## 18         V6 variety  B3     18 11 2      0.17     17.8     17.8
## 19         V7 variety  B3      8 10 1      0.17      7.8      7.8
## 20         V8 variety  B1     30  3 2      0.17     29.8     29.8

## same as agricolae. Does not provide for checks.

str(Ana3)

## List of 6
## $ anova      :Classes 'anova' and 'data.frame': 3 obs. of 5 variables:
## ..$ Df       : int [1:3] 3 2 6
## ..$ Sum Sq   : num [1:3] 191.333 0.667 22.667

```

```
## ..$ Mean Sq: num [1:3] 63.778 0.333 3.778
## ..$ F value: num [1:3] 16.8824 0.0882 NA
## ..$ Pr(>F) : num [1:3] 0.00249 0.91671 NA
## $ adjusted_values:'data.frame': 8 obs. of 9 variables:
## ..$ Accession : Factor w/ 12 levels "C1","C2","C3",...: 5 6 7 8 9 10 11 12
## ..$ Type      : chr [1:8] "variety" "variety" "variety" "variety" ...
## ..$ Bloc      : Factor w/ 3 levels "B1","B2","B3": 1 2 2 1 3 3 3 1
## ..$ weight    : int [1:8] 6 40 4 24 11 18 8 30
## ..$ x         : int [1:8] 2 6 6 3 10 11 10 3
## ..$ y         : int [1:8] 1 2 1 1 2 2 1 2
## ..$ EffetBloc : num [1:8] 0.167 -0.333 -0.333 0.167 0.167 ...
## ..$ weightCorr: num [1:8] 5.83 40.33 4.33 23.83 10.83 ...
## ..$ yvar.adj  : num [1:8] 5.83 40.33 4.33 23.83 10.83 ...
## $ se_check    : num 1.59
## $ se_within   : num 2.75
## $ se_diff     : num 3.07
## $ se_genotype : num 2.51
## Ana3$se_genotype (Difference between two varieties /entries and a check mean) equates Svc , as expected
## approximation in agricolae ?
```