

CH 07. *Split plot designs*

Plan en parcelles partagées

Contexte

Dans le RCBD comme dans le carré latin, les unités expérimentales ont toutes la même taille.

Il arrive que les unités expérimentales aient des dimensions différentes.

Ex. 1: une expérience au champ impliquant une irrigation, qui est réalisé (en une fois) sur des surfaces importantes, et des variétés qui peuvent être semées sur des petites surfaces.

Ex. 2: une expérience en chambre de culture, où des régimes de température sont appliqués à la chambre de culture, et où les effets d'autres facteurs (nutrition, variété...) sont appliqués à l'intérieur des chambres.

Un cas général est celui du *Split Plot*, un facteur A est appliqué aux grandes UE (*whole plots*), et un facteur B est appliqué à des petites UE (*sub plots*, qui sont des parties des *whole plots*).

Plan

1. Présentation de l'exemple
2. Le plan en parcelles partagées
3. Modèle d'analyse
4. Inférence sur les traitements
5. Génération du plan

Exemple

Une expérience au champ investigant les effets compétitifs d'adventices, avec ou sans irrigation, sur le rendement du blé d'hiver.

Les traitements sont combinés de manière factorielle (2 x 2) avec deux types d'irrigation (pluvial - irrigué) et quatre types d'adventices (Alopecurus m., Galium a., Stellaria m. et contrôle sans adventice).

L'expérience comporte 4 blocs.

Le fait que l'irrigation ne puisse être appliquée que sur des larges bandes a mené les chercheurs à considérer un plan en parcelles partagées.

Plan de l'expérience

	Block 1		Block 2		Block 3		Block 4	
Whole plot 1	-	Am	-	Ga	Sm	Ga	Am	-
	7.92	3.62	8.02	5.72	4.91	2.20	3.71	7.16
Whole plot 2	Sm	Ga	Sm	Am	-	Am	Ga	Sm
	5.70	4.49	6.32	3.19	5.54	1.97	6.51	6.65
	-	Sm	Am	Ga	Am	Sm	Ga	Sm
	9.11	6.77	2.52	4.70	2.92	6.64	4.91	5.78
	Ga	Am	-	Sm	Ga	-	Am	-
	7.59	4.12	7.05	5.91	6.90	8.18	2.73	8.22

Bloc

Sous-bloc
Whole plot
UE facteur A

UE facteur B
Sub plot

Le nombre de réplifications pour un niveau du facteur A est moins élevé que s'il pouvait être appliqué aux petites UE. Ce plan sera donc moins puissant qu'un RCBD.

Le plan en parcelles partagées (*Split Plot*)

Caractéristiques du plan:

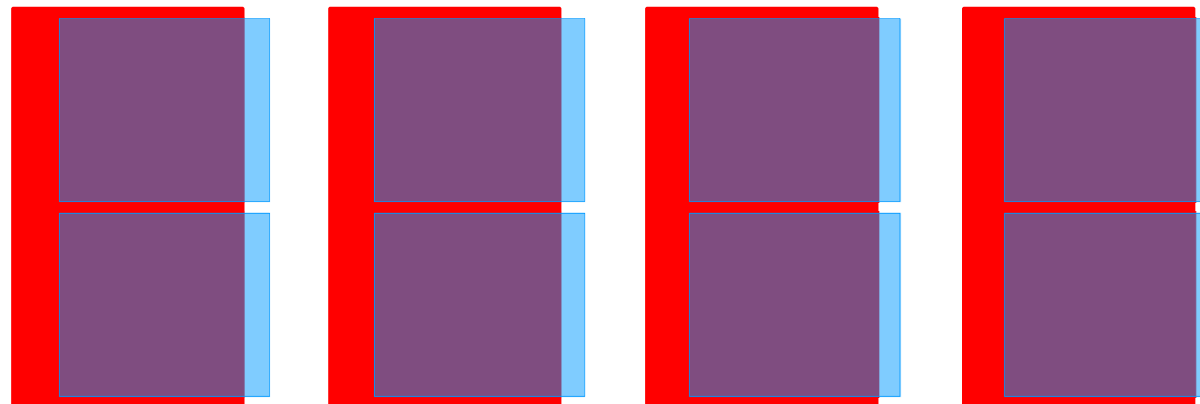
- Le plan comporte m blocs complets
- Chaque bloc comporte t_A sous blocs, auxquels sont appliqués les t_A niveaux du facteur A. Ces sous-blocs sont les grandes UE (*whole plots*)
- Chaque sous-bloc comporte t_B parcelles, auxquelles sont appliqués les t_B niveaux du facteur B. Ces parcelles sont les petites UE (*sub plots*)
- Il y a donc $t = t_A \times t_B$ traitements.

Il existe plusieurs variations du split-plot...

Blocs et sous blocs hiérarchisés contrôlent l'hétérogénéité

	Block 1		Block 2		Block 3		Block 4	
Whole plot 1	-	Am	-	Ga	Sm	Ga	Am	-
	7.92	3.62	8.02	5.72	4.91	2.20	3.71	7.16
Whole plot 2	Sm	Ga	Sm	Am	-	Am	Ga	Sm
	5.70	4.49	6.32	3.19	5.54	1.97	6.51	6.65
Whole plot 2	-	Sm	Am	Ga	Am	Sm	Ga	Sm
	9.11	6.77	2.52	4.70	2.92	6.64	4.91	5.78
Whole plot 2	Ga	Am	-	Sm	Ga	-	Am	-
	7.59	4.12	7.05	5.91	6.90	8.18	2.73	8.22

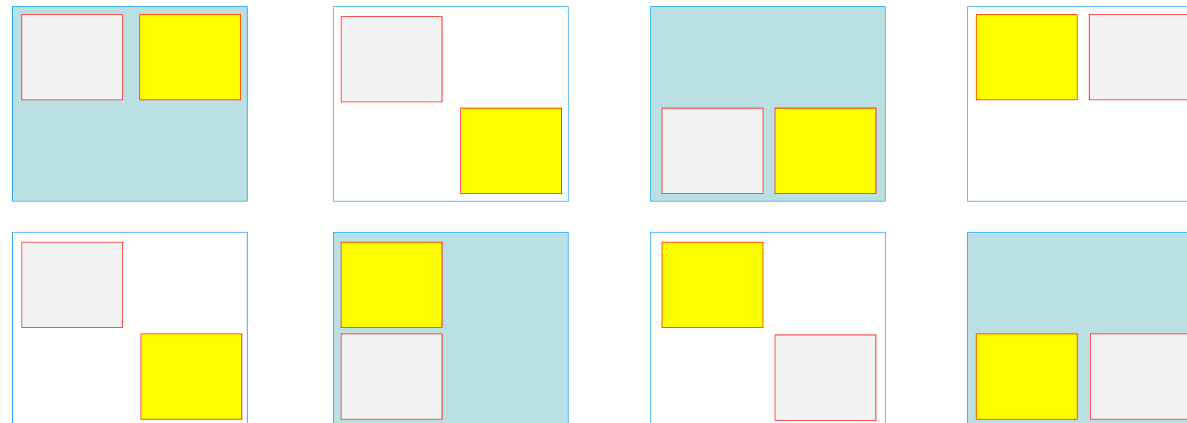
$$b_i + w_{j(i)}$$



Grandes et petites UE correspondent aux facteurs d'intérêt

	Block 1		Block 2		Block 3		Block 4	
Whole plot 1	-	Am	-	Ga	Sm	Ga	Am	-
	7.92	3.62	8.02	5.72	4.91	2.20	3.71	7.16
Whole plot 2	Sm	Ga	Sm	Am	-	Am	Ga	Sm
	5.70	4.49	6.32	3.19	5.54	1.97	6.51	6.65
Whole plot 2	-	Sm	Am	Ga	Am	Sm	Ga	Sm
	9.11	6.77	2.52	4.70	2.92	6.64	4.91	5.78
Whole plot 2	Ga	Am	-	Sm	Ga	-	Am	-
	7.59	4.12	7.05	5.91	6.90	8.18	2.73	8.22

$$\alpha_j + \beta_k + (\alpha\beta)_k$$



Equation du modèle

Le modèle utilisé pour analyser les données de l'expérience est le suivant:

$$y_{ijk} = \mu + b_i + w_{j(i)} + \alpha_j + \beta_k + (\alpha\beta)_k + e_{ijk}$$

avec :

y_{ijk} : le rendement de l'UE située dans le bloc i , dans la parcelle $j(i)$ et qui a reçu la combinaison des niveaux j de A et k de B

μ : la moyenne générale de la population des échantillons

b_i : l'effet du bloc i , exprimé en différence de la moyenne générale

$w_{j(i)}$: l'effet du sous-bloc j , exprimé en différence de la moyenne du bloc i

α_j : l'effet du niveau j de A exprimé en différence de la moyenne générale

β_k : l'effet du niveau k de B exprimé en différence de la moyenne générale

$\alpha\beta_{jk}$: l'effet d'interaction entre les niveaux j de A et k de B

e_{ijk} : la déviation (erreur résiduelle) entre la valeur observée y_{ijk} et sa valeur attendue $(\mu + b_i + w_{j(i)} + \alpha_j + \beta_k + (\alpha\beta)_k)$

$i=1 \dots m, j=1 \dots t_A, k=1 \dots t_B, N=m \cdot t_A \cdot t_B$

$e_{ijk} \sim iN(0, \sigma^2)$

Et, si on considère un modèle mixte, $b_i \sim iN(0, \sigma_b^2)$ $w_{j(i)} \sim iN(0, \sigma_w^2)$

Décomposition de la variance (approche GLM)

df	Sum of Squares	Mean Square	Ratio	
$m - 1$	BkSS	BkMS	BkMS/WPtMS	Effet Bloc
$t_A - 1$	SS(A)	MS(A)	MS(A)/WPtMS	Effet facteur A
$(t_A - 1) \times (m - 1)$	WPtSS	WPtMS	WPtMS/ResMS	Effet Whole plot
$t_B - 1$	SS(B)	MS(B)	MS(B)/ResMS	Effet facteur B
$(t_A - 1) \times (t_B - 1)$	SS(A.B)	MS(A.B)	MS(A.B)/ResMS	Interaction AxB
$t_A \times (t_B - 1) \times (m - 1)$	ResSS	ResMS		
$N - 1$	TotSS			

De ce tableau, on retiendra notamment que les degrés de liberté au dénominateur pour les tests sur A et B ne sont pas les mêmes.

Modélisation R (GLM)

```
data <- read.table(file="C:/Users/draye/Home/Cours/BRAI2222 - Mixtes et plans/R/EXAMPLES/COMPETITION.DAT",
header = TRUE)
data$Block <- as.factor(data$Block)
data$WholePlot <- as.factor(data$WholePlot)
data$SubPlot <- as.factor(data$Subplot)

## Aproche Modele GLM

mod1 <- lm(Grain ~ Block + WholePlot:Block + Irrigation*Species, data)
summary(mod1)

# test F sur les facteurs fixes: traitement et Block
anova(mod1)
```

Modélisation R (GLM)

```
> summary(mod1)
```

Call:

```
lm(formula = Grain ~ Block + WholePlot:Block + Irrigation * Species,  
    data = data)
```

Residuals:

```
      Min       1Q   Median       3Q      Max  
-0.7456 -0.3650  0.1169  0.3097  0.5656
```

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	7.9281	0.4947	16.026	4.24e-12	***
Block2	-0.2175	0.5109	-0.426	0.675366	
Block3	-1.7775	0.3613	-4.920	0.000110	***
Block4	-0.0225	0.3613	-0.062	0.951026	
Irrigationyes	-0.1988	0.4779	-0.416	0.682422	
SpeciesAm	-4.6325	0.3613	-12.823	1.72e-10	***
SpeciesGa	-1.4375	0.3613	-3.979	0.000880	***
SpeciesSm	-1.5225	0.3613	-4.214	0.000521	***
Block1:WholePlot2	0.8675	0.5109	1.698	0.106739	
Block2:WholePlot2	-0.1700	0.5109	-0.333	0.743176	
Block3:WholePlot2	1.9075	0.5109	3.734	0.001521	**
Block4:WholePlot2	NA	NA	NA	NA	
Irrigationyes:SpeciesAm	0.1600	0.5109	0.313	0.757755	
Irrigationyes:SpeciesGa	-1.6700	0.5109	-3.269	0.004266	**
Irrigationyes:SpeciesSm	-0.0850	0.5109	-0.166	0.869721	

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

Residual standard error: 0.5109 on 18 degrees of freedom

Multiple R-squared: 0.961, Adjusted R-squared: 0.9328

F-statistic: 34.09 on 13 and 18 DF, p-value: 5.537e-10

Modélisation R (GLM)

```
> anova(mod1)
Analysis of Variance Table

Response: Grain

      Df Sum Sq Mean Sq  F value    Pr(>F)
Block    3  6.647   2.2158    8.4885 0.0009994 ***
Irrigation 1 14.231  14.2311   54.5188 7.528e-07 ***
Species   3 85.926  28.6419  109.7259 9.314e-12 ***
Block:WholePlot 3  4.504   1.5012    5.7509 0.0061059 **
Irrigation:Species 3  4.371   1.4571    5.5823 0.0069108 **
Residuals 18  4.699   0.2610

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

Si on regarde bien, on voit que `lm` utilise le même carré moyen résiduel au dénominateur de tous les tests...

Modélisation R (en mixte)

```
require(lme4)
require(emmeans)
require(lmerTest)

## Approche Modele Mixte

mod2 <- lmer(Grain ~ Irrigation*Species + (1 | Block) + (1 | WholePlot:Block), data)
summary(mod2)

# test F sur les facteurs fixes
anova(mod2)

# test LRT sur les facteurs aleatoire block
ranova(mod2)

# Estimation des moyennes par (ici SE est l'erreur std sur une moyenne (SEM))
emm2 <- emmeans(mod2, ~ Irrigation*Species)
emm2
```

Modélisation R

```
> summary(mod2)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Grain ~ Irrigation * Species + (1 | Block) + (1 | WholePlot:Block)
Data: data
```

```
REML criterion at convergence: 58.6
```

```
Scaled residuals:
```

```
      Min       1Q   Median       3Q      Max
-1.6093 -0.6387  0.1940  0.6135  1.1657
```

```
Random effects:
```

```
Groups          Name          Variance Std.Dev.
WholePlot:Block (Intercept) 0.31003  0.5568
Block           (Intercept) 0.08932  0.2989
Residual                               0.26103  0.5109
```

```
Number of obs: 32, groups: WholePlot:Block, 8; Block, 4
```

```
Fixed effects:
```

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	8.1175	0.4063	11.0585	19.978	4.99e-10 ***
Irrigationyes	-0.9350	0.5343	6.6449	-1.750	0.125915
SpeciesAm	-4.6325	0.3613	18.0000	-12.823	1.72e-10 ***
SpeciesGa	-1.4375	0.3613	18.0000	-3.979	0.000880 ***
SpeciesSm	-1.5225	0.3613	18.0000	-4.214	0.000521 ***
Irrigationyes:SpeciesAm	0.1600	0.5109	18.0000	0.313	0.757755
Irrigationyes:SpeciesGa	-1.6700	0.5109	18.0000	-3.269	0.004266 **
Irrigationyes:SpeciesSm	-0.0850	0.5109	18.0000	-0.166	0.869721

```
---
```

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

SEM

SED (pour A)

SED (pour B)

Modélisation R

```
> anova(mod1)
```

```
Type III Analysis of Variance Table with Satterthwaite's method
```

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Irrigation	2.475	2.4746	1	3	9.4800	0.054180 .
Species	85.926	28.6419	3	18	109.7259	9.314e-12 ***
Irrigation:Species	4.371	1.4571	3	18	5.5823	0.006911 **

Le nombre de dl est correct

```
---
```

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

```
> emm2 <- emmeans(mod2, ~ Irrigation*Species)
```

```
> emm2
```

Irrigation	Species	emmean	SE	df	lower.CL	upper.CL
no	-	8.12	0.406	11.1	7.22	9.01
yes	-	7.18	0.406	11.1	6.29	8.08
no	Am	3.48	0.406	11.1	2.59	4.38
yes	Am	2.71	0.406	11.1	1.82	3.60
no	Ga	6.68	0.406	11.1	5.79	7.57
yes	Ga	4.08	0.406	11.1	3.18	4.97
no	Sm	6.59	0.406	11.1	5.70	7.49
yes	Sm	5.58	0.406	11.1	4.68	6.47

SEM

```
Degrees-of-freedom method: kenward-roger
```

```
Confidence level used: 0.95
```


Création d'un Split Plot avec DiGGer

```
require(DiGGer)
.pardefault <- par() # Keep copy of graphical parameters

# Création du design
```

Création d'un Split Plot avec DiGGer

Représentation graphique