

## Chapitre 7

# Les expériences en parcelles divisées (*split-plot*) et en bandes croisées (*split-block*)

### Sommaire

#### 7.1 Principes

- 7.1.1 Les expériences en blocs aléatoires complets et parcelles divisées
- 7.1.2 Autres dispositifs expérimentaux en parcelles divisées
- 7.1.3 Les expériences en bandes croisées

#### 7.2 Analyse des résultats

- 7.2.1 Les expériences en blocs aléatoires complets et parcelles divisées
- 7.2.2 Autres dispositifs expérimentaux en parcelles divisées
- 7.2.3 Les expériences en bandes croisées

#### 7.3 Discussion

#### 7.4 Exemple 1 : expérience en blocs aléatoires complets et parcelles divisées

- 7.4.1 Présentation et données
- 7.4.2 Analyse des résultats

#### 7.5 Exemple 2 : expérience en blocs aléatoires complets et bandes croisées

- 7.5.1 Présentation et données
- 7.5.2 Analyse des résultats

## 7.1 Principes

### 7.1.1 Les expériences en blocs aléatoires complets et parcelles divisées

#### 1° Deux facteurs

Nous introduisons le principe des *parcelles divisées*<sup>1</sup> en considérant tout d'abord le cas le plus classique des expériences factorielles organisées en blocs aléatoires complets. La figure 7.4.1 en donne un exemple, relatif à deux facteurs.

Dans cet exemple, que nous détaillerons ultérieurement (§ 7.4.1), le premier facteur possède six modalités et le deuxième facteur en possède trois. L'expérience comporte quatre blocs, correspondant aux quatre quadrants de la figure 7.4.1, et la répartition des 18 objets au sein des blocs a été réalisée en deux temps, correspondant aux deux facteurs.

Dans un premier temps, on a défini six parcelles au sein de chacun des quatre blocs, et on leur a attribué les différentes modalités du premier facteur (différentes variétés, dans l'exemple considéré), comme dans toute expérience en blocs aléatoires complets qui ne ferait intervenir qu'un seul facteur (§ 6.1.1.1°). Il s'agit des premiers chiffres (1 à 6) des nombres de deux chiffres qui identifient les différents objets. À ce stade, on ne tient donc aucun compte de l'existence du deuxième facteur.

Dans un deuxième temps, on a divisé chacune des 24 premières parcelles en trois parcelles plus petites, et on a réparti les trois modalités du deuxième facteur (différentes fumures, dans l'exemple considéré) au sein de chacun des 24 groupes de trois parcelles, comme s'il s'agissait de 24 blocs (deuxièmes chiffres, à savoir 1 à 3, des nombres qui désignent les différents objets). À ce stade de la randomisation, on néglige donc tout à fait le premier facteur.

D'une manière générale, pour deux facteurs présentant respectivement  $p$  et  $q$  modalités et pour  $r$  blocs, la première étape est une répartition classique des  $p$  modalités du premier facteur au sein des  $r$  blocs, dans  $pr$  parcelles de relativement grande taille. La deuxième étape consiste ensuite en une répartition aléatoire et indépendante des  $q$  modalités du deuxième facteur à l'intérieur de ces  $pr$  parcelles, en constituant un total de  $pqr$  parcelles de plus petite dimension.

Les premières parcelles, de plus grande taille, sont appelées *grandes parcelles* et parfois *sous-blocs*<sup>2</sup>, tandis que les secondes, de plus petite taille, sont qualifiées de *petites parcelles* ou *sous-parcelles*<sup>3</sup>. D'une manière plus générale, en dehors des

---

<sup>1</sup> En anglais : *split-plot*.

<sup>2</sup> En anglais : *whole plot, sub-block*.

<sup>3</sup> En anglais : *sub-plot*.

expériences en champ, en forêt, etc., on parle aussi d'*unités du premier degré*<sup>4</sup> et d'*unités du deuxième degré*<sup>5</sup>.

Comme au paragraphe 6.1.2, les notions de bloc, de grande parcelle et de petite parcelle peuvent en effet être appliquées à d'autres situations que les expériences en champ. Dans une expérience qui serait relative par exemple à la fabrication de papier, les quatre blocs pourraient consister en quatre journées de travail, les six « grandes parcelles » pourraient correspondre à six types de matière première, pour chacun desquels une fabrication serait réalisée au cours de chacune des quatre journées, et les « petites parcelles » pourraient correspondre à trois modalités de séchage du papier.

### ⊖ 2° Trois et plus de trois facteurs

Le principe des blocs aléatoires complets avec parcelles divisées peut être étendu de trois façons différentes au cas de trois facteurs.

Pour une expérience factorielle comportant trois facteurs, respectivement à  $p$ ,  $p'$  et  $p''$  modalités, et  $r$  blocs, on peut tout d'abord associer les différentes modalités d'un seul facteur aux grandes parcelles (ou aux unités du premier degré), puis toutes les combinaisons des deux autres facteurs aux petites parcelles (ou aux unités du deuxième degré), en définissant ainsi  $pr$  grandes parcelles et  $pp'p''r$  petites parcelles.

On peut aussi associer toutes les combinaisons des modalités de deux facteurs aux grandes parcelles, puis les différentes modalités du troisième facteur aux petites parcelles, en définissant  $pp'r$  grandes parcelles et, aussi,  $pp'p''r$  petites parcelles. On notera que les grandes parcelles de la première option sont en fait plus étendues que celles de la deuxième option, puisqu'elles réunissent chacune  $p'p''$  petites parcelles, au lieu de  $p''$  petites parcelles.

Mais, toujours pour une expérience qui comporterait trois facteurs, on peut envisager en outre d'effectuer la répartition des objets en trois temps, en définissant des *parcelles divisées à deux niveaux*<sup>6</sup>.

Les différentes modalités d'un premier facteur sont alors réparties tout d'abord à l'intérieur des  $r$  blocs, dans  $pr$  grandes parcelles (ou unités du premier degré). Les différentes modalités d'un deuxième facteur sont ensuite réparties à l'intérieur des  $pr$  grandes parcelles, en  $pp'r$  parcelles intermédiaires ou sous-parcelles (ou unités du deuxième degré). Et les différentes modalités du troisième facteur sont enfin réparties à l'intérieur des  $pp'r$  parcelles intermédiaires, en  $pp'p''r$  petites parcelles ou *sous-sous-parcelles* (ou *unités du troisième degré*)<sup>7</sup>, toutes les répartitions étant effectuées « au hasard » et indépendamment les unes des autres.

---

<sup>4</sup> En anglais : *whole unit*.

<sup>5</sup> En anglais : *sub-unit*.

<sup>6</sup> En anglais : *split-split-plot*.

<sup>7</sup> En anglais : *sub-sub-plot, sub-sub-unit*.

Les mêmes principes peuvent aussi être appliqués à un nombre quelconque de facteurs, avec évidemment une diversité de possibilités encore accrue.

### 7.1.2 Autres dispositifs expérimentaux en parcelles divisées

#### 1° Dispositif complètement aléatoire

Le principe des parcelles divisées peut être associé aussi au cas des expériences complètement aléatoires. Comme première illustration, considérons par exemple une expérience d'alimentation de truies et de porcelets.

Dans cette optique, nous supposons qu'on souhaite étudier la croissance de porcelets en fonction, à la fois, de l'alimentation qui est donnée à leurs mères durant leurs gestations ou leurs périodes d'allaitement, et de l'alimentation qui est donnée ensuite aux porcelets eux-mêmes, après sevrage.

Si on considère deux alimentations différentes des truies et quatre alimentations différentes des porcelets, et si on dispose de 10 truies, on peut répartir « au hasard » dans un premier temps les deux alimentations entre les 10 truies, puis choisir quatre porcelets dans chacune des portées, et répartir entre eux au hasard et indépendamment, dans chacune des portées, les quatre alimentations destinées aux porcelets. Il s'agirait ainsi d'une expérience factorielle  $2 \times 4$ , avec cinq répétitions et utilisant le principe des parcelles divisées.

Une deuxième illustration du même dispositif est donnée par la séquence :

33 31 32 13 12 11 32 31 33 21 22 23 11 13 12 21 22 23,

que nous avons envisagée au paragraphe 5.3.3°, et dans laquelle les trois modalités du deuxième facteur sont chaque fois associées successivement à chacune des modalités du premier facteur. Il s'agit ici d'une expérience factorielle  $3^2$  de même nature que la précédente, mais en deux répétitions.

#### ⊖ 2° Carré latin et blocs incomplets

Le principe des expériences en parcelles divisées peut être adapté à de nombreux autres dispositifs expérimentaux, dont les dispositifs en carré latin et en blocs incomplets (§ 8.1, 9.1 et 10.1).

Dans le cas du carré latin  $4 \times 4$  qui est présenté à la figure 8.1.1 et qui fait intervenir un facteur objets, un facteur lignes et un facteur colonnes, on pourrait subdiviser par exemple chacune des 16 parcelles en trois sous-parcelles, en vue d'introduire dans l'expérience un facteur supplémentaire à trois modalités. En dehors des facteurs lignes et colonnes, l'expérience serait alors une expérience factorielle  $4 \times 3$  en carré latin et parcelles divisées, avec quatre répétitions.

### ⊖ 3° Expériences factorielles fractionnaires, surfaces de réponse, etc.

Les différentes applications du principe des parcelles divisées que nous avons envisagées jusqu'à présent concernent toutes des expériences factorielles complètes (§ 2.3.2). Mais le même principe peut être appliqué aussi aux expériences non factorielles (au sens strict) : expériences factorielles fractionnaires (§ 2.3.3), étude des surfaces de réponse (§ 2.4.1 et 2.4.2), etc.

En matière d'expérience factorielle fractionnaire, dans le cas de l'exemple du paragraphe 2.3.3.6° (expérience de type  $3^{3-1}$ ), la répartition des objets pourrait être :

$$221 \quad 232 \quad 213 \quad 312 \quad 331 \quad 323 \quad 133 \quad 122 \quad 111,$$

une première répartition au hasard ayant été réalisée tout d'abord pour le premier facteur (premiers chiffres), et une deuxième répartition au hasard étant réalisée ensuite globalement pour les deux autres facteurs (deux derniers chiffres).

De même, en ce qui concerne l'étude des surfaces de réponse, dans le cas du plan de DOEHLERT à deux facteurs (partie gauche de la figure 2.4.3 et partie supérieure gauche du tableau 2.4.1), on pourrait envisager de procéder en un premier temps à une répartition aléatoire des trois niveaux du facteur  $x_2$ , puis à une répartition aléatoire, pour chacun de ces niveaux, des deux ou des trois niveaux du facteur  $x_1$ . Le résultat pourrait être par exemple :

$$(-1; 0) \quad (1; 0) \quad (0; 0) \quad (0,5; -0,866) \quad (0,5; 0,866) \quad (-0,5; 0,866) \quad (-0,5; -0,866).$$

De telles situations, essentiellement liées aux expériences organisées de façon consécutive, se présentent principalement dans le domaine industriel.

## 7.1.3 Les expériences en bandes croisées

### 1° Blocs aléatoires complets

Aux expériences en blocs aléatoires complets et parcelles divisées, on peut associer les expériences en blocs aléatoires complets et *bandes croisées* ou *blocs divisés*<sup>8</sup>. Il s'agit d'expériences factorielles qui font généralement intervenir deux facteurs et qui sont conçues de telle sorte qu'à l'intérieur de chacun des blocs, les différents traitements soient appliqués à des bandes de terrain allongées dans une première direction pour un des facteurs et dans une deuxième direction, perpendiculaire à la première, pour l'autre facteur.

La figure 7.5.1 donne une illustration de ce type d'expérience. Il s'agit de deux blocs, dans chacun desquels apparaissent, d'une part, huit bandes verticales relativement étroites correspondant aux huit modalités d'un premier facteur (premiers chiffres, 1 à 8, des nombres qui désignent les différents objets), et d'autre part, deux

<sup>8</sup> En anglais : *split-block*, *strip-block*, *strip-plot*, *criss-cross*.

bandes horizontales plus larges correspondant aux deux modalités d'un deuxième facteur (deuxièmes chiffres, 1 ou 2, des nombres qui désignent les différents objets). Dans une direction comme dans l'autre, la répartition des différentes modalités est réalisée « au hasard » et indépendamment dans chacun des blocs.

Cette façon de procéder permet de disposer de grandes parcelles, qui correspondent aux bandes verticales et horizontales, pour chacun des deux facteurs, et de petites parcelles, qui correspondent aux intersections des deux réseaux de bandes, en ce qui concerne l'interaction des deux facteurs.

### ⊖ 2° Carré latin, blocs incomplets, etc.

Le même principe peut être appliqué à d'autres dispositifs expérimentaux, dont le carré latin et les blocs incomplets (§ 8.1, 9.1 et 10.1). Il s'agit toutefois de situations peu courantes. On peut trouver des informations à ce sujet dans les publications de FEDERER [1955, 1975, 1977] et KUEHL [2000].

De nombreuses variantes des dispositifs expérimentaux en parcelles divisées et en bandes croisées sont également présentées par FEDERER et KING [2007].

## 7.2 Analyse des résultats

### 7.2.1 Les expériences en blocs aléatoires complets et parcelles divisées

#### 1° Analyse de la variance

Dans le cas de deux facteurs, auquel nous nous limitons, l'interprétation des résultats des expériences en blocs aléatoires complets et parcelles divisées consiste généralement en une analyse de la variance à trois critères de classification, le plus souvent avec une seule observation par petite parcelle et par variable étudiée [STAT2, § 11.2.4].

Les différentes sources de variation d'une telle analyse sont, classiquement, le premier facteur considéré au départ, le deuxième facteur considéré au départ, le facteur blocs, l'interaction des deux facteurs initiaux, et les interactions facteur 1 – blocs, facteur 2 – blocs et facteur 1 – facteur 2 – blocs. Ces trois dernières interactions ne peuvent toutefois pas être regroupées en une seule composante « résiduelle », comme dans le cas général des blocs aléatoires complets (§ 6.2.2°), car on doit s'attendre à ce que la première soit plus importante, et parfois même beaucoup plus importante, que les deux autres.

L'interaction facteur 1 – blocs se mesure en effet au sein des blocs, au niveau des grandes parcelles, tandis que les interactions facteur 2 – blocs et facteur 1 – facteur 2 – blocs résultent essentiellement des disparités qui peuvent exister entre les petites parcelles, au sein des grandes parcelles. Or, l'hétérogénéité observée au

premier niveau (grandes parcelles) est normalement supérieure à celle observée au deuxième niveau (petites parcelles).

En conséquence, l'interaction facteur 1 – blocs est généralement considérée individuellement et utilisée comme base de comparaison pour tester l'influence du premier facteur. Cette composante est souvent appelée aussi *variation résiduelle 1* ou *erreur 1* (ou *a*)<sup>9</sup>. Par contre, les deux autres interactions sont généralement regroupées en une composante unique, appelée *variation résiduelle 2* ou *erreur 2* (ou *b*)<sup>10</sup>, qui sert de base de comparaison à la fois pour le deuxième facteur et pour l'interaction des deux facteurs<sup>11</sup>.

En outre, on présente souvent séparément l'ensemble des éléments relatifs aux grandes parcelles et l'ensemble des éléments qui concernent les petites parcelles. Le tableau 7.2.1 donne le schéma d'une telle analyse, pour  $p$  modalités du premier facteur, associé aux grandes parcelles,  $q$  modalités du deuxième facteur, associé aux petites parcelles, et  $r$  blocs.

**Tableau 7.2.1.** Schéma de l'analyse de la variance relative aux expériences en blocs aléatoires complets et parcelles divisées, dans le cas général  $p \times q \times r$  et dans le cas particulier  $6 \times 3 \times 4$ .

Sources de variation	Degrés de liberté	
Facteur 1	$p - 1$	5
Blocs	$r - 1$	3
Variation résiduelle 1	$(p - 1)(r - 1)$	15
Facteur 2	$q - 1$	2
Facteur 1 – facteur 2	$(p - 1)(q - 1)$	10
Variation résiduelle 2	$p(q - 1)(r - 1)$	36
Totaux	$pqr - 1$	71

Ce tableau présente également, en particulier, les nombres de degrés de liberté relatifs à l'exemple du paragraphe 7.1.1.1° (comparaison de variétés et de fumures).

## 2° Comparaisons de moyennes

Dans les comparaisons de moyennes, au sens des comparaisons particulières ou multiples notamment [STAT2, § 12.2, 12.3 et 12.4], il faut être attentif au fait que la manière de calculer les erreurs standards des différences de moyennes dépend du ou des facteurs qui sont pris en considération.

<sup>9</sup> En anglais : *error 1*.

<sup>10</sup> En anglais : *error 2*.

<sup>11</sup> Comme au paragraphe 6.2.2° et comme dans de nombreux autres cas, l'utilisation des expressions « variation résiduelle », « carré moyen résiduel » ou « variance résiduelle » est quelque peu abusive, dans la mesure où il ne s'agit pas de variances résiduelles au sens propre, mais bien d'interactions ou de groupes d'interactions servant de bases de comparaison pour d'autres éléments de l'analyse.

Pour les différences  $\bar{x}_{i..} - \bar{x}_{i'..}$ , relatives au premier facteur, et pour les différences  $\bar{x}_{.j.} - \bar{x}_{.j'.$ , relatives au deuxième facteur, les erreurs standards sont respectivement :

$$\sqrt{2 \text{CM}_{(1)}/(qr)} \quad \text{et} \quad \sqrt{2 \text{CM}_{(2)}/(pr)},$$

si on désigne par  $\text{CM}_{(1)}$  et  $\text{CM}_{(2)}$  les carrés moyens qui concernent les deux sources de variation « résiduelle ». De même, pour les différences  $\bar{x}_{ij.} - \bar{x}_{i'j.}$ , relatives également au deuxième facteur, mais en se limitant à une modalité donnée  $i$  du premier facteur, l'erreur standard est :

$$\sqrt{2 \text{CM}_{(2)}/r}.$$

⌈ La situation se complique quelque peu quand on considère les différences  $\bar{x}_{ij.} - \bar{x}_{i'j.}$ , relatives au premier facteur en se limitant à une modalité donnée  $j$  du deuxième facteur. L'erreur standard est alors :

$$\sqrt{2 [\text{CM}_{(1)} + (q-1) \text{CM}_{(2)}]/(qr)}.$$

Le cas échéant, le nombre de degrés de liberté correspondant peut être déterminé comme suit, de façon approchée, selon la méthode de SATTERTHWAITTE [STAT2, § 11.2.3.10°] :

$$p(p-1)(r-1) [\text{CM}_{(1)} + (q-1) \text{CM}_{(2)}]^2 / [p \text{CM}_{(1)}^2 + (p-1)(q-1) \text{CM}_{(2)}^2].$$

En outre, l'analyse de la variance et les erreurs standards des différences de moyennes qui concernent le cas des parcelles divisées à deux niveaux (*split-split-plot*) sont présentées notamment par KUEHL [2000].

## 7.2.2 Autres dispositifs expérimentaux en parcelles divisées

### 1° Dispositif complètement aléatoire

Des principes semblables peuvent être appliqués au cas des expériences complètement aléatoires organisées en parcelles divisées.

En particulier, pour deux facteurs, l'analyse de la variance est une analyse à deux critères de classification. Le tableau 7.2.2 donne le schéma de cette analyse, pour  $p$  et  $q$  modalités des facteurs et pour  $r$  répétitions des  $pq$  objets, ainsi que les nombres de degrés de liberté relatifs au premier exemple du paragraphe 7.1.2.1° (alimentation de truies et de porcelets).

### ⊖ 2° Autres situations

La méthode qui vient d'être présentée ne s'applique évidemment pas quand on ne dispose que d'une seule répétition, les deux composantes résiduelles disparaissant alors des analyses de la variance. D'autres approches doivent donc être



**Tableau 7.2.2.** Schéma de l'analyse de la variance relative aux expériences complètement aléatoires organisées en parcelles divisées, dans le cas général  $p \times q \times r$  et dans le cas particulier  $2 \times 4 \times 5$ .

Sources de variation	Degrés de liberté	
Facteur 1	$p - 1$	1
Variation résiduelle 1	$p(r - 1)$	8
Facteur 2	$q - 1$	3
Facteur 1 – facteur 2	$(p - 1)(q - 1)$	3
Variation résiduelle 2	$p(q - 1)(r - 1)$	24
Totaux	$pqr - 1$	39

envisagées, et il en est de même pour des effectifs inégaux et pour les expériences non factorielles au sens strict, qui ont aussi été évoquées au paragraphe 7.1.2.3°.

Des informations peuvent être trouvées à ce sujet dans les articles de GOOS *et al.* [2006], NAES *et al.* [2007], et SMITH et JOHNSON [2007].

### 7.2.3 Les expériences en bandes croisées

#### 1° Analyse de la variance

Selon des principes semblables à ceux du paragraphe 7.2.1.1°, pour les expériences en blocs aléatoires complets et bandes croisées, l'interprétation des résultats consiste normalement en une analyse de la variance à trois critères de classification, dans laquelle n'intervient aucun regroupement a priori des interactions liées au facteur blocs [STAT2, § 11.2.4]. On doit s'attendre en effet, dans ce cas, à ce que les interactions facteur 1 – blocs et facteur 2 – blocs soient toutes deux plus importantes que l'interaction facteur 1 – facteur 2 – blocs, sans qu'elles ne soient cependant nécessairement du même ordre de grandeur.

Chacune des trois sources de variation auxquelles on s'intéresse principalement, à savoir les deux facteurs initiaux et leur interaction, doit alors être comparée à son interaction avec le facteur blocs. Certains regroupements peuvent néanmoins être envisagés au cas par cas, moyennant la réalisation de tests préliminaires.

Les interactions qui servent de bases de comparaison sont parfois appelées variation résiduelle 1, variation résiduelle 2 et variation résiduelle 3, ou erreur 1 (ou  $a$ ), erreur 2 (ou  $b$ ) et erreur 3 (ou  $ab$ ), comme le montre le tableau 7.2.3, pour  $pq$  objets et  $r$  blocs.

Ce tableau présente également les nombres de degrés de liberté relatifs à l'exemple qui a été évoqué au paragraphe 7.1.3.1°.

**Tableau 7.2.3.** Schéma de l'analyse de la variance relative aux expériences en blocs aléatoires complets et bandes croisées, dans le cas général  $p \times q \times r$  et dans le cas particulier  $8 \times 2 \times 2$ .

Sources de variation	Degrés de liberté	
Facteur 1	$p - 1$	7
Blocs	$r - 1$	1
Variation résiduelle 1	$(p - 1)(r - 1)$	7
Facteur 2	$q - 1$	1
Variation résiduelle 2	$(q - 1)(r - 1)$	1
Facteur 1 – facteur 2	$(p - 1)(q - 1)$	7
Variation résiduelle 3	$(p - 1)(q - 1)(r - 1)$	7
Totaux	$pqr - 1$	31

### ⊖ 2° Comparaisons de moyennes

Des informations relatives au calcul des erreurs standards des différences de moyennes sont données notamment par KUEHL [2000]. D'autre part, une large discussion relative aux expériences en parcelles divisées et en bandes croisées, comportant notamment de nombreux exemples et diverses mises en garde, est présentée par FEDERER [1975, 1977].

## 7.3 Discussion

### 1° Avantages et inconvénients

Les expériences en parcelles divisées et en bandes croisées ont l'avantage de permettre l'utilisation de parcelles ou, d'une manière plus générale, d'unités expérimentales de plus grandes dimensions, pour un facteur dans le premier cas et pour les deux facteurs dans le deuxième cas.

Mais il faut savoir que cet avantage est compensé par une perte, parfois importante, de précision ou de puissance pour les comparaisons qui concernent le ou, éventuellement, les facteurs associés aux grandes parcelles. Cette perte de précision ou de puissance résulte non seulement des différences d'ordre de grandeur entre les composantes « résiduelles » qui servent de bases de comparaison (§ 7.2.1.1° et 7.2.3.1°), mais aussi des différences qui existent entre les nombres de degrés de liberté de ces composantes. Ce dernier élément intervient de façon d'autant plus marquée que le ou les facteurs associés aux grandes parcelles comportent un petit nombre de modalités.

Par contre, on peut souvent observer un gain de précision ou de puissance pour le facteur qui est associé aux petites parcelles et pour l'interaction des deux facteurs.

## 2° Applications agronomiques

Dans le domaine agronomique, l'utilisation de parcelles divisées ou de bandes croisées peut donc se justifier en premier lieu quand, pour l'une ou l'autre raison (mécanisation, facilité de réalisation de l'expérience, etc.), un des deux ou les deux facteurs considérés nécessitent des parcelles ou des unités expérimentales de relativement grandes dimensions. Au moment de prendre une décision allant éventuellement dans ce sens, il ne faut cependant pas perdre de vue les réductions de précision ou de puissance auxquelles nous venons de faire allusion.

D'autre part, le recours au principe des parcelles divisées peut aussi se justifier quand on souhaite mettre l'accent plus sur un des facteurs que sur l'autre ou, à la limite, quand on s'intéresse uniquement à un des deux facteurs et à l'interaction des deux facteurs. Tel peut être le cas par exemple si on souhaite comparer différents produits phytosanitaires et étudier leur interaction éventuelle avec un certain nombre de variétés d'une culture donnée, sans s'intéresser particulièrement aux différences pouvant exister entre les variétés.

Enfin, un troisième cas d'application du principe des parcelles divisées concerne l'introduction, en cours d'expérience, d'un facteur supplémentaire non prévu initialement. Une telle situation peut se présenter en particulier dans les expériences de longue durée (expériences forestières notamment), où il est parfois utile ou même nécessaire de subdiviser les parcelles, si leur dimension le permet, de manière à pouvoir tenir compte d'un facteur supplémentaire qui n'avait pas été prévu ou jugé suffisamment important au départ (lutte contre une maladie ou un parasite qui a pris de l'extension en cours d'expérience, par exemple).

## 3° Applications industrielles

D'origine agronomique, le principe des parcelles divisées a été introduit et largement développé dans le domaine industriel et dans certains domaines connexes où les expériences sont organisées en séquence. On notera d'ailleurs que la dénomination « *split-plot* » reste régulièrement d'application dans de nombreuses situations où la notion de « parcelle » n'a aucune raison d'être.

Le but poursuivi dans ce cas est essentiellement de faire face autant que possible aux difficultés de modification des modalités de certains facteurs (§ 5.3.3°). Dans cette optique, le principe des parcelles divisées est assez souvent associé à l'un ou l'autre critère d'optimalité, appliqué au coût global de réalisation de l'expérience.

[ Des informations complémentaires sont données notamment par GOOS [2002],  
[ JONES et NACHTSHEIM [2009], et KOWALSKI *et al.* [2007].

## 4° Efficacité relative

On peut déterminer de façon approchée l'efficacité relative des dispositifs en parcelles divisées et en bandes croisées, par comparaison avec les dispositifs classiques correspondants, en calculant les rapports des carrés moyens qui servent de

bases de comparaison, d'une manière semblable à celle que nous avons présentée au paragraphe 6.3.2°.

Dans le cas des expériences en blocs aléatoires complets et parcelles divisées par exemple, il s'agit des quotients :

$$CM_r/CM_{(1)} \quad \text{et} \quad CM_r/CM_{(2)},$$

si on désigne par  $CM_{(1)}$  et  $CM_{(2)}$  les carrés moyens relatifs aux deux sources de variation « résiduelle » de l'expérience en parcelles divisées, et par  $CM_r$  le carré moyen « résiduel » qui est obtenu en supposant que la répartition des objets à l'intérieur des blocs a été réalisée de manière complètement aléatoire.

Toujours dans ce cas, on peut démontrer que  $CM_r$  est en fait une moyenne pondérée de  $CM_{(1)}$  et  $CM_{(2)}$ . Il en résulte que toute augmentation de l'un ou l'autre de ces deux derniers carrés moyens est nécessairement compensée par une diminution de l'autre, ce qui confirme le fait que tout gain de précision ou de puissance lié aux petites parcelles est compensé par une perte de précision ou de puissance pour les grandes parcelles.

Des formules plus précises, qui tiennent compte notamment de la correction relative aux nombres de degrés de liberté (§ 6.3.3°), peuvent être trouvées dans plusieurs ouvrages classiques, dont ceux de FEDERER [1955] et KUEHL [2000].

[ On peut facilement démontrer la propriété de moyenne pondérée en partant des définitions des différentes composantes. Si on désigne par  $SCE_{ac}$ ,  $SCE_{bc}$  et  $SCE_{abc}$  les sommes des carrés des écarts relatives aux trois interactions qui font intervenir le facteur blocs, affecté de l'indice  $c$  [STAT2, § 11.2.2], on a :

$$CM_{(1)} = SCE_{ac}/[(p-1)(r-1)], \quad CM_{(2)} = (SCE_{bc} + SCE_{abc})/[p(q-1)(r-1)]$$

$$\text{et} \quad CM_r = (SCE_{ac} + SCE_{bc} + SCE_{abc})/[(p-1)(r-1)].$$

On en déduit :

$$CM_r = \frac{(p-1)(r-1)CM_{(1)} + p(q-1)(r-1)CM_{(2)}}{(p-1)(r-1) + p(q-1)(r-1)},$$

ce qui est bien une relation de moyenne pondérée, les coefficients de pondération étant les nombres de degrés de liberté des carrés moyens.

### 5° Données manquantes

La question des données manquantes se complique quelque peu dans le cas des expériences en parcelles divisées et en bandes croisées, car elle doit être envisagée, le cas échéant, de manière séparée pour les grandes et pour les petites parcelles.

Cette question peut être traitée soit par l'utilisation de formules particulières semblables à celles qui concernent par exemple les blocs aléatoires complets classiques [STAT2, § 10.4.3], soit par l'analyse de la variance non orthogonale, l'analyse de la covariance ou le modèle linéaire mixte (§ 12.4.1, 12.4.2 et 12.4.4).

[ Des informations peuvent être trouvées notamment à ce sujet dans le livre de  
[ FEDERER et KING [2007].

### ⊖ 6° Parcelles divisées et observations successives

Les observations qui sont répétées à différents moments sur les mêmes individus sont parfois traitées comme s'il s'agissait de parcelles divisées, en assimilant les différents individus observés à de grandes parcelles et les différentes observations successives à de petites parcelles.

Dans cette optique, pour une expérience complètement aléatoire qui ferait intervenir  $p$  modalités d'un seul facteur et  $q$  observations successives (croissance d'animaux soumis à différentes alimentations et observés à différentes dates, par exemple), l'analyse serait du type de celle du tableau 7.2.2, le « facteur 2 » étant en fait le facteur temps ou dates. De même, dans le cas d'une expérience en blocs aléatoires complets, pour  $p$  modalités d'un seul facteur,  $q$  répétitions successives des mêmes mesures et  $r$  blocs, l'analyse de la variance serait celle du tableau 7.2.1.

Aucune difficulté particulière ne se présente en ce qui concerne la première partie de telles analyses, c'est-à-dire pour le seul facteur considéré au départ. Il n'en est pas de même par contre pour la deuxième partie de l'analyse, c'est-à-dire pour le facteur temps ou dates et pour l'interaction des deux facteurs. Les différentes observations successives étant réalisées sur les mêmes individus, il faut en effet s'attendre à ce que les résidus relatifs à chacun des individus ne soient pas indépendants les uns des autres, mais au contraire plus ou moins fortement corrélés entre eux.

Nous donnerons quelques informations complémentaires à ce propos au paragraphe 11.2.2.3°.

## 7.4 Exemple 1 : expérience en blocs aléatoires complets et parcelles divisées

### 7.4.1 Présentation et données

#### 1° Présentation générale

Nous illustrons le principe des blocs aléatoires complets avec parcelles divisées en considérant le cas d'une expérience réalisée sur haricot en vue d'étudier, pour plusieurs variétés, l'influence d'une inoculation bactérienne par *Rhizobium*.

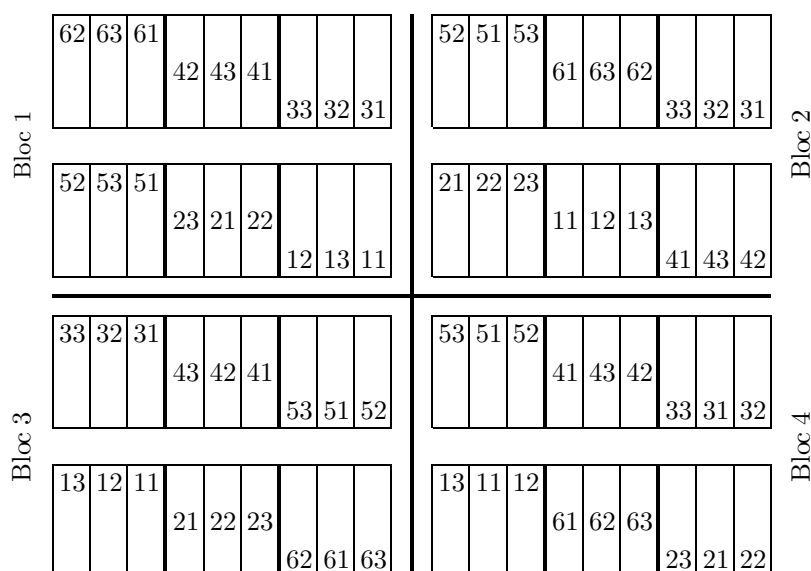
Les variétés considérées sont au nombre de six et les « traitements » appliqués à ces variétés au nombre de trois, l'ensemble constituant une expérience factorielle complète comportant 18 objets. Les « traitements » sont, dans chaque cas, un té-

moins, une inoculation et un apport d'azote minéral (50 unités d'azote par ha), cet apport étant destiné à chiffrer autant que possible l'équivalent en azote de l'inoculation bactérienne.

Le détail des modalités de culture et de traitement est exposé dans la publication originale de BRAKEL et MANIL [1965].

## 2° Dispositif expérimental

La figure 7.4.1 représente, à l'échelle de 1/400, le plan exact de l'expérience. Les parcelles unitaires ont 2 m de largeur sur 6 m de longueur et sont réunies en huit groupes de neuf parcelles, les différents groupes étant séparés et entourés de tous côtés par des chemins d'accès de 2 m de largeur<sup>12</sup>.



**Figure 7.4.1.** Expérience d'inoculation bactérienne sur haricot : plan de l'expérience, les numéros d'ordre des variétés et des traitements étant chaque fois juxtaposés.

Les 72 parcelles ont été divisées en quatre blocs, constitués chacun de deux groupes de neuf parcelles. En l'absence de toute information relative à l'hétérogénéité du terrain, ces blocs ont été définis de manière à être aussi compacts que possible. Ils correspondent aux quatre quadrants de la figure 7.4.1.

<sup>12</sup> Dans sa version électronique, la transmission du présent document en format PDF peut avoir pour conséquence de modifier quelque peu l'échelle de la figure 7.4.1.

Nous avons indiqué au paragraphe 7.1.1.1° comment la répartition des 18 objets a été effectuée, en deux temps, au sein de chacun des blocs. L'objectif poursuivi en ayant recours au principe des parcelles divisées était, à la fois, de simplifier la réalisation de l'expérience (semis de 24 grandes parcelles au lieu de 72 parcelles unitaires), et de concentrer l'attention sur les différences entre traitements et sur leur interaction avec le facteur variétés, éventuellement aux dépens de la précision des comparaisons entre variétés, considérées comme relativement secondaires.

### 3° Données

Chacune des 72 petites parcelles comportait cinq lignes de haricots, semées à 40 cm d'écartement. Seules les trois lignes centrales de chaque parcelle ont été l'objet de mesures, de telle sorte que la surface utile par parcelle est égale à 7,2 m<sup>2</sup>.

Les rendements obtenus de cette manière et exprimés en kg de gousses fraîches par parcelle, ainsi que les moyennes correspondantes, sont présentés dans le tableau 7.4.1.

**Tableau 7.4.1.** Expérience d'inoculation bactérienne sur haricot : rendements observés et moyennes, en kg de gousses fraîches par parcelle.

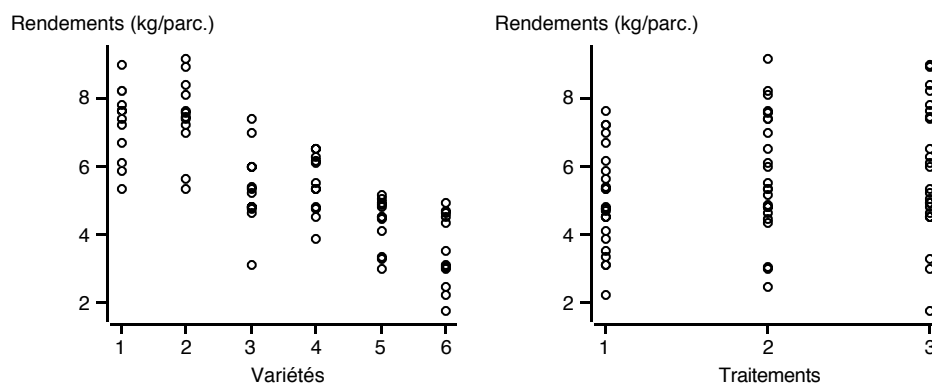
Variétés et traitements		Blocs				Moyennes
		1	2	3	4	
1	1	7,20	6,70	5,87	5,35	6,28
1	2	7,38	8,20	7,60	6,08	7,32
1	3	7,82	8,98	7,64	8,20	8,16
2	1	6,95	7,60	7,22	5,60	6,84
2	2	7,55	9,15	8,10	5,30	7,52
2	3	7,42	8,89	8,40	7,37	8,02
3	1	5,37	4,75	3,08	5,30	4,62
3	2	6,95	7,41	4,80	6,00	6,29
3	3	6,00	5,20	4,80	4,60	5,15
4	1	3,87	6,15	4,75	4,50	4,82
4	2	5,35	6,50	4,79	5,50	5,54
4	3	5,35	6,50	6,28	6,10	6,06
5	1	4,50	4,80	3,35	4,11	4,19
5	2	5,14	4,45	4,85	2,95	4,35
5	3	4,50	4,91	3,28	5,02	4,43
6	1	3,10	4,65	2,20	3,50	3,36
6	2	3,00	4,59	2,46	4,30	3,59
6	3	2,95	4,90	1,75	4,48	3,52

## 7.4.2 Analyse des résultats<sup>13</sup>

### 1° Examen préliminaire

La figure 7.4.2 présente les données sous forme graphique. Dans cette figure, comme dans le tableau 7.4.1, on observe tout d'abord que les données sont extrêmement variables (valeur minimum égale à 1,75 et valeur maximum égale à 9,15, soit un rapport de plus de 1 à 5, avec une moyenne générale égale à 5,56 kg par parcelle). On observe aussi que les différences concernent essentiellement les variétés et, à première vue, relativement peu les traitements.

En raison notamment de l'existence de ces différences importantes, les rendements ont été soumis à une transformation logarithmique (logarithmes décimaux), destinée à stabiliser les variances.



**Figure 7.4.2.** Expérience d'inoculation bactérienne sur haricot : représentation graphique de l'influence des deux facteurs considérés.

### 2° Analyse de la variance

Le tableau 7.4.2 donne les résultats de l'analyse de la variance, réalisée selon le schéma du tableau 7.2.1. L'analyse détaillée à trois critères de classification montrerait que la ligne « Variation résiduelle 1 » du tableau 7.4.2 correspond bien à l'interaction variétés-blocs, et que le regroupement des interactions traitements-blocs et variétés-traitements-blocs, qui constitue la « Variation résiduelle 2 », se justifie pleinement. Les carrés moyens de ces deux interactions sont en effet égaux respectivement à 0,003854 et 0,002951, avec 6 et 30 degrés de liberté.

Le tableau 7.4.2 montre qu'il existe des différences très hautement significatives entre variétés et entre traitements, et que l'interaction des deux facteurs n'est pas significative.

<sup>13</sup> Un traitement complet de cet exemple à l'aide du logiciel R est présenté à l'adresse <[www.dagnelie.be/docexr/exp074/exp074.htm](http://www.dagnelie.be/docexr/exp074/exp074.htm)>.



**Tableau 7.4.2.** Expérience d'inoculation bactérienne sur haricot : tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Variétés	5	1,055289	0,211058	15,0 ***	0,0000
Blocs	3	0,147833	0,049278		
Variation résiduelle 1	15	0,211102	0,014073		
Traitements	2	0,056757	0,028378	9,15 ***	0,0006
Variétés-traitements	10	0,044062	0,004406	1,42	0,21
Variation résiduelle 2	36	0,111646	0,003101		
Totaux	71	1,626689			

Ces conclusions peuvent paraître surprenantes, par comparaison avec ce que semblent indiquer les diagrammes de la figure 7.4.2. Elles résultent du fait que les tests relatifs au facteur traitements et à l'interaction variétés-traitements sont beaucoup plus puissants que le test relatif au facteur variétés, en raison, d'une part, de la très grande différence entre les carrés moyens des deux composantes résiduelles (rapport de 1 à 4,5 environ), et d'autre part, de la différence entre les nombres de degrés de liberté qui sont associés à ces deux composantes (respectivement 15 et 36).

### 3° Différences de moyennes

Comme il a été dit ci-dessus, les différences entre variétés ne constituent pas, dans le cas présent, un des éléments essentiels de l'expérimentation. Par contre, les différences entre traitements doivent être prises en considération de manière plus précise.

Les moyennes des 24 observations relatives à chacun des traitements sont, après transformation logarithmique :

témoin : 0,68144 ,  
 inoculation : 0,73875 ,  
 fumure azotée : 0,74301 .

On peut en déduire les intervalles de confiance suivants, d'une part pour l'influence de l'inoculation :

$$0,73875 - 0,68144 \pm 2,028 \sqrt{2(0,003101)/24} = 0,05731 \pm 0,03260 \\ = 0,02471 \text{ et } 0,08991 ,$$

et d'autre part pour l'influence de la fumure azotée :

$$0,74301 - 0,68144 \pm 2,028 \sqrt{2(0,003101)/24} = 0,06157 \pm 0,03260 \\ = 0,02897 \text{ et } 0,09417 .$$

La valeur 0,05731 est en fait une estimation de la différence des logarithmes des rendements moyens et, donc aussi, du logarithme du quotient des rendements moyens. Cette différence indique que les rendements obtenus après inoculation sont égaux, dans l'ensemble, à 114 % des rendements des parcelles témoins, puisque :

$$10^{0,05731} = 1,141 \quad \text{ou} \quad \log_{10} 1,141 = 0,0573.$$

De même, aux valeurs 0,02471 et 0,08991, correspondent des limites égales à 106 et 123 %.

Ces différents résultats montrent que l'accroissement de rendement lié à l'inoculation peut être estimé à 14 %, avec des limites de confiance égales à 6 et 23 %.

Quant à l'accroissement de rendement lié à la fumure azotée, il peut être estimé d'une manière identique à 15 %, avec des limites de confiance égales à 7 et 24 %.

L'accroissement de rendement dû à l'inoculation étant égal à environ 93 % de l'accroissement lié à l'apport de 50 unités d'azote par ha, on peut en conclure que l'inoculation bactérienne équivaut dans ce cas à l'apport de 46 ou 47 unités d'azote.

#### 4° Efficacité relative

Le tableau 7.4.2 montre aussi que, pour une répartition supposée complètement aléatoire des 18 objets au sein des quatre blocs, c'est-à-dire en négligeant la structure propre aux parcelles divisées, la variation « résiduelle » relative aux deux facteurs étudiés et à leur interaction aurait été :

$$CM_r = (0,211102 + 0,111646)/(15 + 36) = 0,006328.$$

On peut en déduire qu'en première approximation, par rapport à une expérience classique en blocs aléatoires complets, l'efficacité du test relatif aux variétés est, dans le cas des parcelles divisées :

$$CM_r/CM_{(1)} = 0,006328/0,014073 = 0,45 \text{ ou } 45 \text{ \%}.$$

On obtient de la même façon, pour les tests relatifs aux traitements et à l'interaction des deux facteurs :

$$CM_r/CM_{(2)} = 0,006328/0,003101 = 2,04 \text{ ou } 204 \text{ \%}.$$

On peut en conclure que le quasi-doublement de l'efficacité des tests relatifs aux traitements et à l'interaction, qui constituent les principaux objectifs de l'expérience, est contrebalancé par une perte de plus de 50 % de l'efficacité du test relatif aux variétés.

[ En raison de la réduction du nombre de degrés de liberté, la perte d'efficacité est sans doute sensiblement plus importante que 55 %, et elle serait plus importante ]  
[ encore si le nombre de variétés considérées avait été plus réduit.

### 5° À propos de la transformation logarithmique

L'emploi de la transformation logarithmique soulève quelques questions complémentaires, qui méritent d'être discutées ici et qui s'appliquent également à d'autres exemples présentés plus loin.

En ce qui concerne les moyennes, il faut se souvenir du fait que le retour à la variable initiale, à partir des moyennes des logarithmes, donne naissance à des moyennes géométriques, qui sont toujours inférieures ou égales aux moyennes arithmétiques correspondantes [STAT1, § 3.5.2]. À titre indicatif, les moyennes relatives aux trois traitements, obtenues par l'intermédiaire des logarithmes, sont :

$$10^{0,68144} = 4,80, \quad 10^{0,73875} = 5,48 \quad \text{et} \quad 10^{0,74301} = 5,53 \text{ kg/parcelle,}$$

alors que les moyennes arithmétiques, déduites directement des données du tableau 7.4.1, sont :

$$5,02, \quad 5,77 \quad \text{et} \quad 5,89 \text{ kg/parcelle.}$$

On peut noter aussi que, globalement, les conclusions auxquelles on aboutit dans le cas présent, avec ou sans transformation de variable, sont fort semblables. Mais l'emploi de la transformation logarithmique, qui complique quelque peu les calculs, présente cependant deux avantages.

D'une part, cette transformation permet d'exprimer toutes les différences entre objets en valeur relative, c'est-à-dire indépendamment notamment des unités de mesure utilisées, de l'étendue des parcelles, etc. Cet avantage n'est pas propre au cas présenté ici, mais est, au contraire, tout à fait général.

D'autre part, l'emploi de la transformation logarithmique rend non significative l'interaction variétés-traitements, qui sans transformation, serait juste significative ( $F_{obs} = 2,36$  et  $P = 0,029$ ). Cette absence d'interaction, en termes de logarithmes, signifie que les effets de l'inoculation et de la fumure azotée, exprimés en valeur relative, ne dépendent pas de la variété considérée. Les pourcentages d'augmentation de rendement qui ont été calculés ci-dessus globalement, pour l'ensemble des variétés, peuvent donc être appliqués indifféremment à chacune d'entre elles.

[ En outre, en ce qui concerne la variabilité « résiduelle », la relation existant entre le coefficient de variation d'une variable aléatoire et l'écart-type de son logarithme [STAT1, ex. 5.8.9] permet de calculer aisément des estimations des deux coefficients de variation résiduelle relatifs aux observations initiales. On obtient pour les grandes parcelles (variétés) :

$$\sqrt{0,014073}/0,4343 = 0,273 \text{ ou } 27,3 \%,$$

et pour les petites parcelles (traitements et interaction des deux facteurs) :

[ 
$$\sqrt{0,003101}/0,4343 = 0,128 \text{ ou } 12,8 \%.$$

### ⊖ 6° Codification du modèle d'analyse de la variance

Au cours des chapitres précédents, nous n'avons pas présenté les codes, en général relativement simples, qui peuvent être employés pour réaliser les analyses de la variance. En ce qui concerne le présent chapitre, nous donnons au contraire les codes relatifs aux deux exemples, de manière à bien faire apparaître la distinction entre le cas des parcelles divisées et le cas des bandes croisées.

Le contenu du tableau 7.4.2 peut être obtenu à l'aide du code suivant, associé aux commandes ou procédures « *anova* » ou « *glm* » :

$$L10R = V \ B \ V*B \ T \ V*T,$$

et cela tant avec le logiciel Minitab qu'avec le logiciel SAS. Dans ce modèle, L10R désigne les logarithmes décimaux des rendements, B, T et V les facteurs blocs, traitements et variétés, V\*B l'interaction variétés-blocs, qui correspond à la « variation résiduelle 1 », et V\*T l'interaction variétés-traitements.

On peut remarquer que la « variation résiduelle 2 » n'apparaît pas dans le code, cette source de variation étant en fait considérée comme un résidu obtenu par différence. En outre, il y a lieu de préciser chaque fois comment les tests doivent être réalisés, ce qui peut être fait à l'aide d'une instruction « *random* » pour Minitab et « *test* » pour SAS.

Quant au logiciel R, la formulation peut être :

$$L10R \sim V*T + \text{Error}(B + B:V),$$

en relation avec la fonction « *aov* ». L'écriture V\*T désigne ici à la fois les facteurs V et T et leur interaction, tandis que B:V représente la seule interaction de B et V, le terme « Error » étant relatif à la procédure de réalisation des tests.

Il faut noter aussi que, pour les logiciels R et SAS, les symboles B, T et V doivent être définis comme étant des facteurs, et non pas des variables (respectivement par des instructions « *factor* » et « *class* »).

De plus, pour les trois logiciels, d'autres commandes, procédures ou fonctions, liées notamment au modèle linéaire mixte (§ 12.4.4), pourraient aussi être utilisées.

Ces dernières remarques, de même que les remarques relatives à la réalisation des tests, s'appliquent tout autant à d'autres exemples présentés plus loin.

## 7.5 Exemple 2 : expérience en blocs aléatoires complets et bandes croisées

### 7.5.1 Présentation et données

#### 1° Présentation générale

L'exemple que nous envisageons ici est issu d'un projet de coopération entre la Tunisie et la Belgique, dont le but essentiel était l'amélioration de l'élevage

bovin [ANTOINE *et al.*, 1976]. Dans ce cadre, de nombreuses expériences relatives aux cultures fourragères devaient évidemment être réalisées, sans toutefois que ces expériences ne deviennent trop complexes et ne prennent une place trop importante dans l'ensemble des activités.

D'une façon plus particulière, l'objectif de l'expérience que nous prenons en considération est de comparer différents mélanges de plantes fourragères (avoine et vesce), associés à deux doses d'engrais azoté (30 et 60 unités d'azote par hectare).

Les différents mélanges étudiés sont définis dans le tableau 7.5.1, les lettres A, B et C désignant trois variétés différentes d'avoine. On notera que ces mélanges peuvent être considérés comme constituant un ensemble factoriel  $4 \times 2$ , qui fait intervenir d'une part les variétés d'avoine (variétés A, B et C, et mélange A-B), et d'autre part les proportions d'avoine et de vesce (50-50 et 25-75)

**Tableau 7.5.1.** Comparaison de différents mélanges fourragers associés à deux doses d'engrais azoté : définition des objets (pourcentages d'avoine et de vesce).

Mélanges	Avoine			Vesce
	A	B	C	
1	50	–	–	50
2	25	–	–	75
3	–	50	–	50
4	–	25	–	75
5	–	–	50	50
6	–	–	25	75
7	25	25	–	50
8	12,5	12,5	–	75

## 2° Dispositif expérimental et données

Les parcelles ont une étendue totale de 8 m sur 20 m, dont une partie seulement est pesée à la récolte.

Dans l'optique de l'organisation d'expériences très simples, deux répétitions sont prévues et les « traitements » (mélanges de semences et engrais) sont appliqués en bandes croisées. À cette fin, chaque répétition est divisée dans un sens en huit bandes, auxquelles sont affectés au hasard les huit mélanges de semences, et dans l'autre sens en deux bandes, auxquelles sont affectées au hasard les deux doses d'engrais.

La figure 7.5.1 donne une représentation schématique du dispositif expérimental qui a été adopté, ainsi que les résultats obtenus, exprimés en tonnes de matière sèche par hectare.

Il faut noter qu'en choisissant un tel dispositif expérimental, on doit savoir avant d'entamer l'expérience, sur la seule base de la répartition des nombres de

62	42	12	82	52	72	22	32	71	51	81	11	31	41	61	21
5,79	8,67	7,97	7,61	8,69	10,61	7,72	8,78	6,68	9,61	3,55	4,83	4,32	7,25	5,30	3,89
61	41	11	81	51	71	21	31	72	52	82	12	32	42	62	22
6,03	7,16	4,92	4,63	7,70	6,36	6,14	5,79	5,52	5,81	5,07	8,16	9,12	8,85	5,57	6,19
Bloc 1								Bloc 2							

**Figure 7.5.1.** Comparaison de différents mélanges fourragers associés à deux doses d’engrais azoté : plan de l’expérience (numéros des mélanges de semences et des doses d’engrais) et rendements observés, en tonnes par ha.

degrés de liberté, que le facteur engrais sera considéré avec une très faible précision (distribution  $F$  de FISHER-SNEDECOR à 1 et 1 degré de liberté), tandis que le facteur mélanges fourragers et son interaction avec le facteur engrais seront connus dans de meilleures conditions (distribution  $F$  à 7 et 7 degrés de liberté, dans les deux cas).

## 7.5.2 Analyse des résultats<sup>14</sup>

### 1° Examen préliminaire

Une représentation graphique des rendements observés est donnée par la figure 7.5.2. Comme dans l’exemple précédent, cette figure met en évidence la grande variabilité des observations, dont les valeurs extrêmes sont égales à 3,55 et 10,61 t/ha (soit un rapport de 1 à 3), la moyenne générale étant égale à 6,70 t/ha.

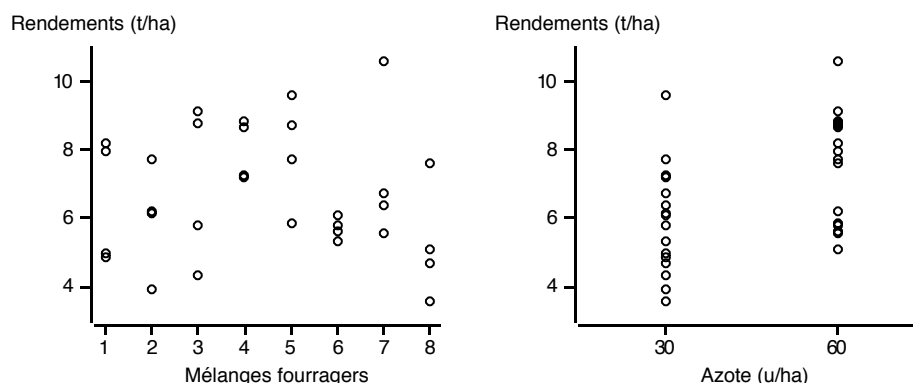
On observe en outre des différences, à première vue difficiles à interpréter, entre les mélanges fourragers, ainsi qu’une certaine influence du facteur fumures.

Comme dans l’exemple précédent également, en fonction notamment de l’hétérogénéité des rendements observés, une transformation logarithmique (logarithmes décimaux) a été appliquée aux données initiales.

### 2° Première analyse de la variance

Les résultats de l’analyse de la variance, présentés comme dans le tableau 7.2.3, sont donnés dans le tableau 7.5.2.

<sup>14</sup> Un traitement complet de cet exemple à l’aide du logiciel R est présenté à l’adresse <[www.dagnelie.be/docexr/exp075/exp075.htm](http://www.dagnelie.be/docexr/exp075/exp075.htm)>.



**Figure 7.5.2.** Comparaison de différents mélanges fourragers associés à deux doses d'engrais azoté : représentation graphique de l'influence des deux facteurs.

**Tableau 7.5.2.** Comparaison de différents mélanges fourragers associés à deux doses d'engrais azoté : tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	$F$	P
Mélanges fourragers	7	0,136850	0,019550	4,77 *	0,028
Blocs	1	0,037397	0,037397		
Mélanges-blocs	7	0,028719	0,004103		
Azote	1	0,096953	0,096953	28,4	0,12
Azote-blocs	1	0,003408	0,003408		
Mélanges-azote	7	0,090646	0,012949	1,91	0,21
Mélanges-azote-blocs	7	0,047413	0,006773		
Totaux	31	0,441386			

Le facteur mélanges fourragers apparaît comme significatif, au niveau de probabilité 0,05, tandis que le facteur azote apparaît comme non significatif, bien que la valeur de la variable  $F$  de FISHER-SNEDECOR soit très élevée, et l'interaction des deux facteurs n'est pas significative non plus.

### 3° Deuxième analyse de la variance

Comme les carrés moyens des trois interactions dans lesquelles intervient le facteur blocs sont du même ordre de grandeur (0,004103, 0,003408 et 0,006773), il peut se justifier de les regrouper en une seule composante de variation « résiduelle ». Cette façon de procéder revient en fait à considérer a posteriori l'expérience comme si elle avait été réalisée en blocs aléatoires complets de manière tout à fait classique. C'est aussi constater que, contrairement au principe général des expériences en

bandes croisées, l'inférence relative aux deux facteurs considérés n'est pas caractérisée ici par une précision ou une puissance moindre que l'interaction des deux facteurs.

D'autre part, il est utile de subdiviser la somme des carrés des écarts du facteur mélanges fourragers, de manière à individualiser ses deux composantes (proportions d'avoine et de vesce et variétés d'avoine) et leur interaction. Une décomposition semblable peut aussi être réalisée pour l'interaction mélanges-azote.

Le tableau 7.5.3 donne les résultats de l'analyse de la variance, tels qu'ils se présentent après ces diverses modifications.

**Tableau 7.5.3.** Comparaison de différents mélanges fourragers associés à deux doses d'engrais azoté : tableau d'analyse de la variance, après regroupement des composantes « résiduelles » et subdivision des composantes mélanges et mélanges-azote.

Sources de variation	Degrés de liberté	S. des carrés des écarts	Carrés moyens	F	P
Mélanges fourragers	7	0,136850	0,019550	3,69 *	0,016
Proportions d'avoine	1	0,030384	0,030384	5,73 *	0,030
Variétés d'avoine	3	0,040384	0,013461	2,54	0,096
Proportions-variétés	3	0,066082	0,022027	4,15 *	0,025
Azote	1	0,096953	0,096953	18,3 ***	0,0007
Mélanges-azote	7	0,090646	0,012949	2,44	0,069
Proportions-azote	1	0,000156	0,000156	0,03	0,87
Variétés-azote	3	0,063962	0,021321	4,02 *	0,028
Prop.-variétés-azote	3	0,026528	0,008843	1,67	0,22
Blocs	1	0,037397	0,037397		
Variation résiduelle	15	0,079540	0,005303		
Totaux	31	0,441386			

#### 4° Interprétation

L'élément le plus frappant, à l'examen du tableau 7.5.3, est le fait que le facteur azote, qui n'était pas significatif avec une base de comparaison égale à 0,003408 et un degré de liberté (tableau 7.5.2), devient très hautement significatif avec une base de comparaison légèrement supérieure (0,005303) et 15 degrés de liberté. Ceci souligne l'importance qu'il faut toujours accorder à l'existence, dans toute la mesure du possible, d'un nombre suffisant de degrés de liberté pour les carrés moyens qui servent de bases de comparaison.

D'autre part, il apparaît que les différences entre mélanges fourragers sont liées au facteur proportions d'avoine et de vesce et à l'interaction de ce facteur avec le facteur variétés d'avoine. Et en outre, une interaction significative variétés-azote apparaît également.



En ce qui concerne la fumure azotée, en procédant comme au paragraphe 7.4.2.3°, à partir des moyennes relatives aux deux niveaux, on obtient :

$$0,86542 - 0,75533 \pm 2,131 \sqrt{2(0,005303)/16} = 0,11009 \pm 0,05487 \\ = 0,05522 \text{ et } 0,16496.$$

En négligeant la légère interaction variétés-azote, l'accroissement de rendement lié aux 30 unités supplémentaires d'azote par hectare peut donc être estimé globalement à 29 %, avec des limites de confiance égales à 14 et 46 %.

Quant aux facteurs proportions d'avoine et de vesce et variétés, ils s'avèrent plus difficiles à interpréter.

Le tableau 7.5.4 présente les moyennes relatives aux huit mélanges fourragers, en distinguant les deux proportions d'avoine et de vesce et les quatre variétés ou mélange de variétés d'avoine, et cela tant pour les données transformées que pour les données initiales.

**Tableau 7.5.4.** Comparaison de différents mélanges fourragers associés à deux doses d'engrais azoté : moyennes relatives aux différents mélanges fourragers, pour les données transformées (logarithmes décimaux) et pour les données initiales (t/ha).

Proport.	Variétés			
	A	B	C	A-B
50-50	0,797	0,825	0,893	0,849
25-75	0,764	0,900	0,753	0,701
50-50	6,47	7,00	7,95	7,29
25-75	5,98	7,98	5,67	5,22

Ce tableau montre que l'interaction proportions-variétés résulte largement d'une différence de comportement de la variété B, par comparaison avec les deux autres variétés et avec le mélange A-B (rendement plus élevé de la proportion 25-75 pour la variété B, contrairement aux trois autres cas). Ce tableau montre aussi que le comportement et les rendements du mélange A-B sont loin d'être intermédiaires entre ceux des variétés A et B considérées séparément.

La poursuite éventuelle de l'interprétation de ces résultats devrait être fonction de l'objectif précis qui a été (ou qui aurait dû être) fixé au départ. On peut en effet, par exemple, soit essayer d'approfondir l'analyse des facteurs et de leur interaction, soit rechercher le « meilleur » mélange de semences, au sens notamment de la méthode de GUPTA [STAT2, § 12.3.3].

On pourrait aussi examiner de la même manière l'interaction variétés-azote, qui s'avère légèrement significative.

### ⊖ 5° Codification des modèles d'analyse de la variance

En ce qui concerne la première analyse de la variance (tableau 7.5.2), le code relatif aux logiciels Minitab et SAS (commandes ou procédures « *anova* » ou « *gjm* ») est :

$$L10R = M \ B \ M*B \ A \ A*B \ M*A ,$$

L10R désignant les logarithmes décimaux des rendements, et A, B et M les facteurs azote, blocs et mélanges fourragers.

La distinction par rapport au cas des parcelles divisées (§ 7.4.2.6°) est l'introduction de l'interaction du deuxième facteur (A) et du facteur blocs (B), c'est-à-dire la « variation résiduelle 2 », l'interaction des trois facteurs (« variation résiduelle 3 »), qui doit être calculée par différence, n'apparaissant pas dans le code.

Quant à la deuxième analyse de la variance (tableau 7.5.3), elle peut être réalisée en adoptant l'écriture :

$$L10R = P \ V \ P*V \ A \ P*A \ V*A \ P*V*A \ B ,$$

P et V désignant en outre les proportions et les variétés d'avoine.

En ce qui concerne la fonction « *aov* » du logiciel R, les codes correspondants peuvent être, d'une façon moins explicite, pour la première analyse :

$$L10R \sim M*A + \text{Error}(M*A*B - M*A),$$

et de façon très condensée, pour la deuxième analyse :

$$L10R \sim P*V*A + B .$$

Rappelons en outre que diverses remarques formulées au paragraphe 7.4.2.6° s'appliquent également à cet exemple et à d'autres exemples ultérieurs.

## Chapitre 8

# Les expériences en carré latin et avec permutation des objets (*cross-over*)

### Sommaire

- 8.1 Principes**
  - 8.1.1 Le carré latin
  - 8.1.2 Le dispositif avec permutation des objets (*cross-over*)
- 8.2 Analyse des résultats**
  - 8.2.1 Le carré latin
  - 8.2.2 Le dispositif avec permutation des objets (*cross-over*)
  - ⊖ 8.2.3 Notions complémentaires
- 8.3 Discussion**
- ⊖ **8.4 Quelques cas particuliers et extensions du carré latin**
- 8.5 Exemple 1 : expérience en carré latin**
  - 8.5.1 Présentation et données
  - 8.5.2 Analyse des résultats
- 8.6 Exemple 2 : expérience avec permutation des objets (*cross-over*)**
  - 8.6.1 Présentation et données
  - 8.6.2 Analyse des résultats

## 8.1 Principes

### 8.1.1 Le carré latin

#### 1° Expériences en champ

En ce qui concerne les expériences en champ, toujours au sens large, on appelle *carré latin*<sup>1</sup> un dispositif qui est constitué d'un nombre de parcelles égal à un carré (9, 16, 25, ...), qui comporte autant de lignes de parcelles que de colonnes de parcelles, et au sein duquel chaque objet est présent une et une seule fois dans chaque ligne et dans chaque colonne.

La figure 8.1.1 en donne un exemple. Au vu de cette figure, on observe que ni les parcelles, ni l'ensemble du champ d'expérience ne doivent cependant être des carrés.

3	1	2	4
1	4	3	2
4	2	1	3
2	3	4	1

**Figure 8.1.1.** Exemple de carré latin  $4 \times 4$ .

Les *lignes*<sup>2</sup> et les *colonnes*<sup>3</sup> d'un carré latin constituent en quelque sorte un double réseau de blocs complets, perpendiculaires les uns aux autres et conduisant à un double contrôle de l'hétérogénéité du matériel expérimental<sup>4</sup>.

#### 2° Autres situations

Comme dans le cas des blocs aléatoires complets (§ 6.1.2), le contrôle de la variabilité qui concerne le carré latin peut être étendu à de nombreuses situations autres que les expériences en champ.

Ainsi, en expérimentation humaine comme en expérimentation animale, les lignes peuvent correspondre à des périodes différentes et les colonnes à des patients ou à des animaux différents, le dispositif étant tel que chacun des objets (traitements thérapeutiques ou alimentations, par exemple) soit considéré une et une seule fois au cours de chacune des périodes et une et une seule fois pour chacun des patients ou des animaux.

<sup>1</sup> En anglais : *Latin square*.

<sup>2</sup> En anglais : *row*.

<sup>3</sup> En anglais : *column*.

<sup>4</sup> L'utilisation française du mot « contrôle », qui est faite ici et qui est courante en expérimentation, est relativement impropre, dans la mesure où il s'agit bien plus de réduire ou de « maîtriser » que de « contrôler » la variabilité ou l'hétérogénéité du matériel expérimental.

De même, dans une expérience industrielle ou de laboratoire, les lignes et les colonnes peuvent correspondre par exemple à des opérateurs ou des ouvriers, à des appareils de mesure ou des machines, à des périodes de travail, etc., l'expérience étant toujours organisée de telle sorte que chacun des objets soit associé une et une seule fois à chacun des éléments qui correspondent aux lignes et aux colonnes.

D'une manière générale, le carré latin permet ainsi de prendre en considération simultanément trois facteurs. Dans l'exemple fort ancien de CRETTE DE PALLUEL, qui a été évoqué dans l'introduction générale, au facteur essentiel alimentations, étaient associés un facteur races et un facteur dates d'abattage. Les 16 moutons considérés appartenaient en effet à quatre races différentes et, à chacune des quatre dates, étaient abbatu un mouton de chacune des quatre races et un mouton correspondant à chacune des quatre alimentations. La figure 8.1.1 illustre parfaitement cette situation, si on considère par exemple que les lignes sont relatives aux dates d'abattage, les colonnes aux races, et les chiffres 1 à 4 aux alimentations.

On notera toutefois que, dans de telles conditions, le carré latin ne permet pas d'étudier les interactions qui peuvent exister entre les différents facteurs.

En outre, dans un champ d'expérience qui présenterait un gradient de fertilité important dans une seule direction, ou dans un laboratoire où on pourrait craindre que les conditions d'expérience évoluent de façon sensible au cours du temps, ainsi que dans d'autres situations semblables, toutes les unités expérimentales (parcelles, périodes de travail, etc.) peuvent être juxtaposées selon cette seule tendance générale, tout en constituant néanmoins un carré latin. La figure 8.1.2 illustre cette possibilité dans le cas d'une expérience en champ, en reprenant côte à côte les différentes lignes de la figure 8.1.1 et, aussi, en modifiant la forme des parcelles.

3	1	2	4	1	4	3	2	4	2	1	3	2	3	4	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

**Figure 8.1.2.** Exemple de carré latin  $4 \times 4$  dont les unités expérimentales sont disposées de façon linéaire.

Ce dispositif pourrait correspondre également à quatre journées de travail, au cours de chacune desquelles quatre essais ou manipulations sont effectués, de telle sorte que les quatre objets considérés interviennent chacun à leur tour en début de journée, chacun à leur tour en deuxième position, etc. Une telle approche permettrait de tenir compte à la fois des différences entre les journées et des différences qui pourraient exister au cours des différentes journées.

### 3° Randomisation

En particulier, on appelle *carré latin de base*<sup>5</sup> un carré latin dans lequel les objets sont rangés par ordre numérique croissant (ou par ordre alphabétique) dans la première ligne et dans la première colonne.

L'étude des carrés latins sous l'angle de l'analyse combinatoire permet de montrer qu'il n'existe qu'un seul carré latin de base de type  $3 \times 3$ , quatre carrés latins de base distincts de type  $4 \times 4$ , 56 carrés latins de base de type  $5 \times 5$ , etc. Des tables relativement complètes en sont données notamment par FISHER et YATES [1982]. Et en annexe, la table 2 en fournit aussi quelques exemples.

La répartition aléatoire des objets peut se faire ici en choisissant « au hasard » un carré latin de base de la dimension souhaitée, et en permutant ensuite de façon aléatoire, d'une part les différentes lignes, et d'autre part les différentes colonnes.

### 8.1.2 Le dispositif avec permutation des objets (*cross-over*)

Le dispositif expérimental avec *permutation des objets*<sup>6</sup> peut être considéré comme résultant de la juxtaposition de deux ou plusieurs carrés latins, choisis indépendamment l'un de l'autre, et de la permutation aléatoire de l'ensemble des lignes et de l'ensemble des colonnes. La figure 8.1.3 en donne un exemple.

4	3	3	2	1	1	2	4
2	1	4	3	4	3	1	2
3	2	1	4	3	2	4	1
1	4	2	1	2	4	3	3

**Figure 8.1.3.** Exemple de dispositif avec permutation des objets.

Dans un tel dispositif, on retrouve chacun des objets une et une seule fois dans chacune des colonnes et un même nombre de fois, supérieur à un, dans chacune des lignes.

Il est évident qu'ici également, les lignes et les colonnes peuvent être tout autre chose que des rangées de parcelles contiguës dans un champ. En particulier, le dispositif avec permutation des objets est largement utilisé en expérimentation humaine et en expérimentation animale, où le nombre de patients ou d'animaux susceptibles d'être pris en considération dépasse en général largement le nombre de traitements qu'on peut appliquer en séquence à chacun d'eux, dans des conditions suffisantes d'homogénéité. Comme ci-dessus, les patients ou les animaux sont alors généralement associés aux colonnes et les périodes aux lignes du dispositif.

<sup>5</sup> En anglais : *standard Latin square*.

<sup>6</sup> En anglais : *cross-over, change-over, switch-over*.

De manière plus particulière encore, le dispositif qui fait intervenir *deux objets et deux périodes*<sup>7</sup> figure parmi les plus employés en expérimentation humaine, pour comparer un traitement à un placebo ou deux traitements entre eux.

La randomisation des expériences avec permutation des objets peut être réalisée comme celle des expériences en carré latin.

## 8.2 Analyse des résultats

### 8.2.1 Le carré latin

#### 1° Analyse de la variance

L'analyse de la variance relative aux expériences en carré latin qui comportent au départ un seul facteur, auquel s'ajoutent un facteur lignes et un facteur colonnes, est une analyse à trois critères de classification d'un type particulier. Ainsi que nous l'avons déjà signalé (§ 8.1.1.2°), le dispositif en question a en effet comme caractéristique d'exclure toute possibilité d'estimation des interactions qui pourraient exister entre les différents facteurs.

Quand on ne dispose que d'une seule observation par parcelle ou par unité expérimentale, l'analyse de la variance ne fait donc intervenir que les trois facteurs et une composante résiduelle. Le tableau 8.2.1 présente la répartition des nombres de degrés de liberté dans le cas général d'un carré latin  $p \times p$  et dans le cas particulier de la figure 8.1.1.

**Tableau 8.2.1.** Schéma de l'analyse de la variance relative aux expériences en carré latin, dans le cas général  $p \times p$  et dans le cas particulier  $4 \times 4$ .

Sources de variation	Degrés de liberté	
Objets	$p - 1$	3
Lignes	$p - 1$	3
Colonnes	$p - 1$	3
Variation résiduelle	$(p - 1)(p - 2)$	6
Totaux	$p^2 - 1$	15

#### 2° Modèle mathématique et test d'additivité

Le modèle mathématique sous-jacent peut s'écrire [STAT2, § 11.2.3] :

$$X_{ijk} - m_{...} = a_i + b_j + c_k + D_{ijk},$$

<sup>7</sup> En anglais : *two-treatment two-period cross-over*.

$a_i$ ,  $b_j$  et  $c_k$  désignant les effets des trois facteurs, et  $D_{ijk}$  la composante résiduelle ( $i, j$  et  $k = 1, \dots, p$ ). Ce modèle implique l'additivité des effets des trois facteurs.

Par analogie notamment avec le cas des blocs aléatoires complets, quand les facteurs lignes et colonnes sont aléatoires, l'hypothèse d'additivité n'est pas restrictive et le facteur objets peut être testé sans problème par rapport à la variation résiduelle.

Par contre, quand les facteurs lignes et colonnes (ou l'un d'entre eux) sont fixes, l'hypothèse d'additivité est plus contraignante. Cette hypothèse doit alors être vérifiée par la réalisation d'un *test d'additivité* ou *test de TUKEY*<sup>8</sup>, semblable à celui qui peut intervenir aussi dans le cas des expériences en blocs aléatoires complets [STAT2, § 10.3.6].

[ Ce test peut être réalisé en calculant la composante de non-additivité [TUKEY, 1955] :

$$\text{SCE}_{add} = \left( \sum_{i=1}^p \sum_{j=1}^p [d_{ijk} (x_{ijk} - d_{ijk})^2] \right)^2 / \text{SCE}_r (x_{ijk} - d_{ijk})^2,$$

dans laquelle  $d_{ijk}$  désigne les résidus de l'analyse de la variance et la quantité  $\text{SCE}_r (x_{ijk} - d_{ijk})^2$  est la somme des carrés des écarts résiduelle de l'analyse de la variance qui peut être réalisée sur les valeurs  $(x_{ijk} - d_{ijk})^2$ . On peut remarquer que la sommation ne porte que sur deux des trois indices, puisque deux indices suffisent pour prendre en considération l'ensemble des  $p^2$  observations.

La somme des carrés des écarts ainsi obtenue, à un degré de liberté, doit être soustraite de la somme des carrés des écarts résiduelle et comparée au carré moyen [ résiduel qui peut alors être recalculé avec  $(p - 1)(p - 2) - 1$  degrés de liberté.

### ⊖ 3° Carré latin et expériences factorielles fractionnaires

Principalement quand les trois facteurs sont fixes, les expériences en carré latin peuvent être considérées comme étant des expériences factorielles fractionnaires (§ 2.3.3). Il s'agit en effet de fractions  $1/p$  d'expériences  $p^3$ , c'est-à-dire aussi d'expériences de type  $p^{3-1}$ .

Ainsi, dans le cas de la figure 8.1.1, l'expérience qui comporte trois facteurs, chacun à quatre modalités, et qui ne fait intervenir que 16 unités expérimentales, correspond au quart d'une expérience  $4^3$  et est donc de type  $4^{3-1}$ . De telles expériences sont de résolution III (§ 2.3.3.7°), ce qui corrobore le fait que les interactions ne peuvent pas être identifiées.

En supposant que chacun des trois facteurs de la figure 8.1.1 résulte lui-même de la combinaison factorielle de deux facteurs à deux modalités, l'expérience peut aussi être considérée comme étant de type  $2^{6-2}$ , c'est-à-dire comme le quart d'une expérience factorielle  $2^6$ . Cette situation est envisagée notamment par COPELAND et NELSON [2000].

<sup>8</sup> En anglais : *test of additivity*, TUKEY's *test*.



## 8.2.2 Le dispositif avec permutation des objets (*cross-over*)

### 1° Analyse de la variance

Pour le dispositif avec permutation des objets, l'analyse de la variance est une extension de celle qui concerne le carré latin. Le tableau 8.2.2 en donne le schéma dans le cas général de  $p$  lignes et  $k p$  colonnes ou, ce qui est équivalent, de  $p$  objets et  $r$  répétitions ( $r = k p$ ), et dans le cas particulier de la figure 8.1.3.

**Tableau 8.2.2.** Schéma de l'analyse de la variance relative aux expériences avec permutation des objets, dans le cas général  $p \times k p$  (ou  $r$ ) et dans le cas particulier  $4 \times 8$ .

Sources de variation	Degrés de liberté	
Objets	$p - 1$	3
Lignes	$p - 1$	3
Colonnes	$k p - 1$ ou $r - 1$	7
Variation résiduelle	$(p - 1)(k p - 2)$ ou $(p - 1)(r - 2)$	18
Totaux	$k p^2 - 1$ ou $p r - 1$	31

Les mêmes remarques que ci-dessus peuvent être formulées quant aux tests d'hypothèses et aux problèmes d'additivité.

### ⊖ 2° Cas particulier de deux objets

L'analyse des résultats peut être réalisée de façon beaucoup plus simple dans le cas particulier de deux objets et deux lignes (le plus souvent deux périodes). On peut en effet montrer que, dans ces conditions, le test  $F$  de l'analyse de la variance à trois critères de classification qui concerne le facteur objets peut être remplacé par un test  $t$  d'égalité de deux moyennes ou par un test  $F$  d'analyse de la variance à un critère de classification, comparant les deux groupes de colonnes et portant sur les différences entre les lignes (première ligne – deuxième ligne ou deuxième ligne – première ligne).

## ⊖ 8.2.3 Notions complémentaires

### 1° Effets résiduels et interactions entre les objets

Quand différents traitements sont appliqués consécutivement aux mêmes individus, on peut craindre que les résultats de l'expérience soient affectés par l'existence d'éventuels *effets résiduels*<sup>9</sup> des traitements au cours des périodes successives ou par l'existence d'éventuelles interactions entre les traitements successifs.

<sup>9</sup> En anglais : *residual effect, carry-over effect*.

Des difficultés de ce type se présentent principalement dans le cas des expériences avec permutation des objets, mais peuvent aussi intervenir en ce qui concerne les expériences en carré latin.

Au niveau de la planification des expériences, une première précaution consiste à choisir des dispositifs *équilibrés pour les effets résiduels*<sup>10</sup>. On qualifie ainsi les dispositifs dans lesquels chacun des objets est suivi un même nombre de fois par chacun des autres objets.

Tel est le cas pour le plan de la figure 8.1.3, mais il n'en est pas ainsi pour le carré latin de la figure 8.1.1. Ce carré latin pourrait toutefois être rendu équilibré pour les effets résiduels en en permutant les deux dernières lignes.

Les dispositifs équilibrés permettent de séparer les effets directs ou immédiats des différents objets au cours des différentes périodes, des effets résiduels relatifs aux périodes ultérieures. L'analyse de la variance qui fait intervenir à la fois les effets directs et les effets résiduels est toutefois non orthogonale (§ 12.4.1). Le paragraphe 8.6.2 en donne un exemple.

Des dispositifs *presque équilibrés*<sup>11</sup> ont également été proposés pour faire face à certains situations particulières, de même que l'existence d'éventuels effets résiduels multiplicatifs ou proportionnels aux effets directs a aussi été envisagée.

Des informations complémentaires relatives à ces diverses possibilités sont données notamment par BAILEY et KUNERT [2006], BATE et JONES [2008], BOSE et STUFKEN [2007], et PARK *et al.* [2011].

## 2° Autres méthodes d'analyse

D'autres méthodes que l'analyse de la variance classique ont été envisagées pour prendre en compte l'existence d'effets résiduels, d'éventuelles interactions entre les effets directs et les effets résiduels, et les possibles non-indépendances des résidus de l'analyse. Ces méthodes font appel notamment à des concepts tels que l'étude des observations successives et l'analyse de la variance à plusieurs variables (§ 11.2.2 et 12.4.5) [CHEN *et al.*, 2002 ; CHEN et WEI, 2003 ; GOAD et JOHNSON, 2000].

D'autre part, diverses méthodes non paramétriques et robustes ont été proposées pour faire face à des situations qui ne permettraient pas l'utilisation de l'analyse de la variance [ÖHRVIK, 1998 ; PUTT et CHINCHILLI, 2004 ; TARDIF *et al.*, 2005]. En particulier, dans le cas du dispositif avec permutation des objets qui ne fait intervenir que deux modalités, le test *t* et l'analyse de la variance à un critère de classification peuvent être remplacés par les méthodes de MANN et WHITNEY ou de KRUSKAL et WALLIS [STAT2, § 8.4.3 et 9.3.6].

On notera enfin que, d'une manière générale, la littérature relative aux expériences avec permutation des objets est particulièrement abondante. On peut citer

<sup>10</sup> En anglais : *balanced for carry-over effects, uniform-balanced design.*

<sup>11</sup> En anglais : *nearly balanced design.*

notamment certains livres spécialisés, tels ceux de JONES et KENWARD [2003] et SENN [2002b].

## 8.3 Discussion

### 1° Avantages et inconvénients

Du fait de leur double contrôle de la variabilité du matériel expérimental, les expériences en carré latin et avec permutation des objets sont généralement plus efficaces que les expériences complètement aléatoires et en blocs aléatoires complets.

Mais le carré latin souffre de deux limitations relativement importantes. D'une part, pour trois objets, le nombre de degrés de liberté de la variation résiduelle est égal à 2 seulement, ce qui est extrêmement faible, et pour quatre objets, il est égal à 6, ce qui est encore fort peu. D'autre part, le nombre d'objets est nécessairement égal au nombre de répétitions (et au nombre de lignes et de colonnes) et ce nombre dépasse rarement 8 ou 10 dans la majorité des expériences. En conséquence, la gamme d'utilisations possibles du carré latin se limite le plus souvent à 5, 6, 7 ou 8 objets, éventuellement 4 et parfois plus de 8.

Il en résulte que le carré latin est en réalité relativement peu utilisé, alors que le dispositif avec permutation des objets est au contraire largement employé, principalement, comme nous l'avons signalé, dans le domaine de la recherche médicale et relative aux animaux.

### 2° Efficacité relative

Diverses efficacités relatives peuvent être calculées dans le cas du carré latin et du dispositif avec permutation des objets, comme pour les blocs aléatoires complets et les parcelles divisées (§ 6.3.2°, 6.3.3° et 7.3.4°).

Il s'agit ici de deux efficacités relatives par rapport à des expériences en blocs aléatoires complets, si on néglige le facteur lignes *ou* le facteur colonnes, et d'une efficacité relative par rapport au dispositif complètement aléatoire, si on néglige à la fois le facteur lignes *et* le facteur colonnes.

Ces efficacités relatives sont souvent fort élevées, ce qui souligne l'intérêt des dispositifs considérés ici, quand ils sont applicables. GARCIA *et al.* [2004] signalent par exemple des efficacités relatives de l'ordre de 400 à 1.000 %, dans le domaine médical. Ces ordres de grandeur correspondent à ceux qui apparaissent dans les exemples des paragraphes 8.5 et 8.6.

### 3° Données manquantes

Les éventuelles données manquantes peuvent être estimées selon une procédure analogue à celle qui concerne les blocs aléatoires complets [STAT2, § 10.4.3]. Par

extension, on a en effet, pour le carré latin :

$$\hat{x}_{ijk} = [p(X'_{i..} + X'_{.j.} + X'_{..k}) - 2X'_{...}] / [(p-1)(p-2)],$$

et pour le dispositif avec permutation des objets :

$$\hat{x}_{ijk} = [p(X'_{i..} + X'_{.j.}) + kpX'_{..k} - 2X'_{...}] / [(p-1)(kp-2)],$$

$X'_{i..}$ ,  $X'_{.j.}$  et  $X'_{..k}$  étant dans les deux cas les sommes relatives respectivement aux différents objets, aux différentes lignes et aux différentes colonnes, et  $X'_{...}$  étant la somme générale.

Il faut noter aussi que, contrairement aux expériences en blocs aléatoires complets, les expériences en carré latin et avec permutation des objets ne permettent pas d'éliminer un ou plusieurs objets ou une ou plusieurs lignes ou colonnes (ou répétitions) sans modifier la nature même du dispositif expérimental.

## ⊖ 8.4 Quelques cas particuliers et extensions du carré latin

### 1° Carré gréco-latin

Divers cas particuliers et diverses extensions du concept de carré latin, autres que le dispositif avec permutation des objets, ont été définis. On peut citer en premier lieu le *carré gréco-latin*<sup>12</sup>.

La figure 8.4.1 en illustre le principe. Les lettres latines y définissent un carré latin en tout point semblable à celui de la figure 8.1.1. De même, les lettres grecques y présentent un autre carré latin, à première vue indépendant du premier, et toujours semblable à celui de la figure 8.1.1. Mais on peut remarquer en outre que, dans l'ensemble du dispositif, chacune des lettres latines est associée une et une seule fois à chacune des lettres grecques.

D $\gamma$	A $\delta$	B $\alpha$	C $\beta$
C $\alpha$	B $\beta$	A $\gamma$	D $\delta$
A $\beta$	D $\alpha$	C $\delta$	B $\gamma$
B $\delta$	C $\gamma$	D $\beta$	A $\alpha$

**Figure 8.4.1.** Exemple de carré gréco-latin  $4 \times 4$ .

D'une manière générale, un carré gréco-latin est un dispositif carré, quant au nombre d'unités expérimentales, qui fait intervenir deux facteurs autres que les

<sup>12</sup> En anglais : *Graeco-Latin square*.

lignes et les colonnes, dans lequel chacune des modalités de chacun des deux facteurs est présente une et une seule fois dans chaque ligne et dans chaque colonne, et chacune des modalités d'un des facteurs est associée une et une seule fois à chacune des modalités de l'autre facteur.

Les carrés latins qui sont à la base de la constitution d'un carré gréco-latin sont dits *orthogonaux*<sup>13</sup> l'un à l'autre. On a pu démontrer qu'il n'existe pas de carrés latins orthogonaux, et donc pas de carré gréco-latin, de dimension  $6 \times 6$ , cette restriction concernant uniquement cette dimension<sup>14</sup>. On peut consulter notamment, à ce propos, l'article de COLBOURN et DINITZ [2001].

Les possibilités classiques d'utilisation du carré gréco-latin, dans le domaine agronomique, sont plus limitées encore que celles du carré latin, car l'introduction d'un facteur supplémentaire provoque une réduction plus importante du nombre de degrés de liberté de la variation résiduelle (trois degrés de liberté seulement pour un carré gréco-latin  $4 \times 4$  et huit degrés de liberté seulement pour un carré gréco-latin  $5 \times 5$ , par exemple).

Mais au même titre que le carré latin permet d'étudier simultanément trois facteurs quelconques, autres que des lignes et des colonnes de parcelles dans un champ, le carré gréco-latin peut intervenir dans l'étude de quatre facteurs quelconques, pour autant que ceux-ci présentent tous le même nombre de modalités. Dans cette optique, et par extension de ce qui a été dit au paragraphe 8.2.1.3°, un carré gréco-latin de dimension  $4 \times 4$  par exemple, tel que celui de la figure 8.4.1, peut être considéré comme un ensemble factoriel fractionnaire  $4^{4-2}$  ou  $2^{8-4}$ .

## 2° Quelques autres dispositifs carrés

Des carrés latins de divers types particuliers ont été étudiés et parfois utilisés dans certaines circonstances. La figure 8.4.2 en réunit quatre exemples, présentés sans randomisation.

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

1	2	3	4
3	4	1	2
2	1	4	3
4	3	2	1

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

**Figure 8.4.2.** Quelques cas particuliers de carré latin : carré latin complet, carré latin auto-orthogonal, carré latin magique et carré de VIK.

<sup>13</sup> En anglais : *orthogonal Latin squares*.

<sup>14</sup> L'absence de carré gréco-latin  $6 \times 6$  est un élément d'une conjecture due à EULER et datant de 1782, mais qui n'a été démontrée qu'en 1960 [BOSE *et al.*, 1960; ULLRICH, 2002].

Le premier exemple s'inscrit dans la ligne des dispositifs équilibrés pour les résidus (§ 8.2.3.1°). Il est tel que chacun des objets est suivi une et une seule fois par chacun des autres objets, non seulement dans les colonnes, mais aussi dans les lignes. Ces carrés latins sont qualifiés de *complets*<sup>15</sup> [FREEMAN, 1979].

Dans l'optique des carrés latins orthogonaux, le deuxième dispositif de la figure 8.4.2 est un carré latin *auto-orthogonal*<sup>16</sup>, c'est-à-dire orthogonal à lui-même, après transposition des lignes et des colonnes [HEDAYAT, 1973, 1975].

Le troisième exemple de la figure 8.4.2 superpose, aux lignes et aux colonnes d'un carré latin, un troisième réseau de blocs, de forme carrée dans le cas considéré. Ces carrés latins sont dits *magiques*<sup>17</sup> [DAGNELIE 1959; FEDERER, 1955].

Enfin, le dernier dispositif de la figure 8.4.2 est tel que chacun des objets est présent une et une seule fois, non seulement dans chaque ligne et dans chaque colonne, mais aussi dans chacune des deux diagonales ou parallèlement à chacune des deux diagonales. Ces dispositifs sont connus sous le nom de *carré de VIK* ou de Knuth VIK<sup>18</sup> [ATKIN *et al.*, 1977]<sup>19 20</sup>.

En plus de ces différents cas particuliers, on peut mentionner diverses extensions du carré latin, dont les *carrés F*<sup>21</sup>, qui sont tels que les différents objets sont présents un même nombre de fois, supérieur à un, dans chacune des lignes et chacune des colonnes du dispositif [HEDAYAT et SEIDEN, 1970], les *cubes latins*<sup>22</sup>, qui peuvent être utiles pour des expériences à réaliser dans des espaces physiquement à trois dimensions (fours, incubateurs, frigos, installations de séchage, etc.) [PREECE *et al.*, 1973], et les *hypercubes latins*<sup>23</sup>.

Le concept d'hypercube latin intervient en particulier dans les expériences qui comportent un nombre important de facteurs, dont les expériences numériques organisées sur ordinateur (§ 2.4.5) [BUTLER, 2001, 2005; STEINBERG et LIN, 2006].

Nous présenterons en outre plus loin les *carrés semi-latins* et différents types de *réseaux carrés* (§ 9.1.3 et 9.1.5).

---

<sup>15</sup> En anglais : *complete Latin square*.

<sup>16</sup> En anglais : *self-orthogonal Latin square*.

<sup>17</sup> En anglais : *magic Latin square*.

<sup>18</sup> En anglais : *VIK's square*.

<sup>19</sup> Chacun de ces types de dispositifs possède bien sûr ses propres propriétés. Ainsi par exemple, les carrés latins magiques n'existent pas pour des nombres d'objets qui sont des nombres premiers, tandis que les carrés de VIK n'existent pas pour des nombres d'objets qui sont des multiples de 2 ou de 3.

<sup>20</sup> Il est amusant de constater que les carrés latins magiques 9×9 correspondent exactement aux grilles les plus courantes du jeu de sudoku [DAGNELIE, 2008c], et que les carrés de VIK présentent une disposition des objets qui s'identifie aux déplacements des cavaliers du jeu d'échecs.

<sup>21</sup> En anglais : *frequency square*, *F-square*.

<sup>22</sup> En anglais : *Latin cube*.

<sup>23</sup> En anglais : *Latin hypercube*.

### 3° Quelques dispositifs rectangulaires

Nous envisageons ici différents dispositifs expérimentaux qui, comme le dispositif avec permutation des objets (§ 8.1.2), ne possèdent pas le même nombre de lignes et de colonnes, mais qui se distinguent de telles expériences par le fait que les différents objets ne sont pas tous présents un même nombre de fois dans chacune des lignes et dans chacune des colonnes.

D'une manière générale, ces dispositifs peuvent être qualifiés de *dispositifs en lignes et colonnes*<sup>24</sup>. La figure 8.4.3 en présente trois exemples. Nous y ajouterons ultérieurement divers autres dispositifs, dont les *carrés latins incomplets* ou *carrés de YOUNG*, qui combinent un réseau de blocs complets et un réseau de blocs incomplets (§ 9.1.4).

3	1	2	4
1	4	3	2
2	3	4	1
4	2	1	3
4	2	1	3

1	2	1	2	3	3
3	3	2	1	2	1
2	1	3	3	1	2
2	1	3	3	1	2

0	0	0	0	0	0
1	2	1	2	3	3
3	3	2	1	2	1
2	1	3	3	1	2

**Figure 8.4.3.** Quelques exemples d'extensions « rectangulaires » du carré latin et du dispositif avec permutation des objets.

Les exemples de cette figure concernent tous les trois le cas où les lignes correspondent à différentes périodes successives et les colonnes à des individus auxquels différents traitements sont appliqués en séquence (§ 8.1.1.2° et 8.1.2). Le but poursuivi est, dans les trois cas, d'améliorer la connaissance qu'on peut avoir des effets résiduels des traitements. On peut effectivement démontrer que la simple utilisation de dispositifs équilibrés pour les effets résiduels, dont nous avons parlé au paragraphe 8.2.3.1°, conduit à une forte disparité entre la précision obtenue pour les effets directs et la précision obtenue pour les effets résiduels des objets.

Le premier exemple est un *carré latin augmenté*<sup>25</sup>, qui prévoit simplement une répétition supplémentaire de la dernière ligne d'un carré latin classique, ce carré latin étant, dans le cas présent, du type équilibré pour les effets résiduels. Le deuxième exemple est de la même nature, mais dans le cas d'un dispositif avec permutation des objets, et cela pour trois objets seulement.

Enfin, le dernier exemple de la figure 8.4.3 suppose que, pour un dispositif avec permutation des objets, on possède des observations initiales (notées « 0 ») faites avant toute application des traitements, ces observations initiales intervenant ensuite dans l'analyse des résultats.

<sup>24</sup> En anglais : *row-and-column design*, *row-column design*.

<sup>25</sup> En anglais : *augmented Latin square*.

## 8.5 Exemple 1 : expérience en carré latin

### 8.5.1 Présentation et données

#### 1° Présentation générale

Nous considérons comme illustration du concept de carré latin le cas d'une expérience de chauffage du sol réalisée en serre, en Belgique en période hivernale, sur une variété décorative de *Ficus elastica* ROXB.

Les températures du sol prises en considération sont 15, 20, 25 et 30 degrés centigrades, à 12 centimètres de profondeur, toutes les autres conditions de culture étant en principe uniformes [GÉRARD, 1977].

Les unités expérimentales sont des ensembles de neuf plantes, cultivées en pots, ceux-ci étant enfoncés en terre. L'écartement entre les tiges est de 33 cm environ en tout sens, de telle sorte que les « parcelles » de neuf plantes ont une étendue de 1 m<sup>2</sup>.

#### 2° Dispositif expérimental

L'ensemble de l'expérience devant être réalisé dans deux petites serres, de 3 m de largeur sur 6 m de longueur, et des passages devant être prévus, à la fois pour pouvoir effectuer toutes les mensurations nécessaires et pour assurer des zones tampons entre les différentes « parcelles », il a été décidé de considérer quatre répétitions, disposées en carré latin. La figure 8.5.1 donne le plan de l'expérience, à l'échelle de 1/100.

Les serres étant allongées dans le sens ouest-est, le facteur lignes du dispositif en carré latin doit permettre d'éliminer, le cas échéant, les différences entre serres et les différences liées aux expositions nord et sud. De même, le facteur colonnes doit permettre d'éliminer, éventuellement, les différences liées aux expositions ouest (colonnes de gauche) et est (colonnes de droite), ainsi que les différences entre « parcelles » situées aux extrémités des serres (colonnes latérales) et au centre des serres (colonnes centrales)<sup>26</sup>.

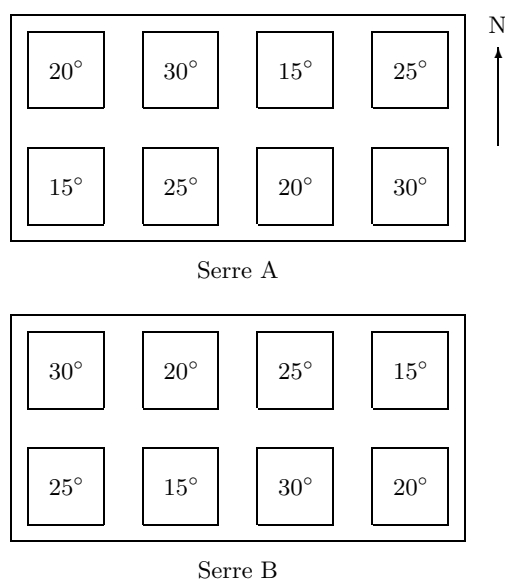
#### 3° Données

Les observations, effectuées de deux en deux semaines durant une période de huit mois (de début septembre à fin avril), ont eu trait à la hauteur des plantes, définie comme étant la distance séparant les points d'insertion de la première et

---

<sup>26</sup> On notera qu'en réalité, les deux serres n'étaient pas disposées côte à côte, comme semble l'indiquer la figure 8.5.1, mais bien dans le prolongement l'une de l'autre, la serre A étant située à l'ouest de la serre B. Cette disposition particulière ne modifie en rien la structure générale du carré latin.





**Figure 8.5.1.** Expérience de chauffage du sol sur *Ficus* : plan de l'expérience (températures du sol, en degrés centigrades).

de la dernière feuille, ainsi qu'au nombre de feuilles, aux dimensions (longueur et largeur) des feuilles et, subsidiairement, à la température du sol et de l'air. Ces observations ont permis d'étudier différentes variables dérivées, telles que l'accroissement total en hauteur, la longueur moyenne des entre-nœuds, etc.

Nous ne considérons ici que les accroissements totaux en hauteur, pour l'ensemble de la période de huit mois. Les moyennes de ces accroissements, calculées pour les différentes « parcelles » de neuf plantes, sont données dans le tableau 8.5.1. En vue de faire apparaître clairement la concordance entre ce tableau et la figure 8.5.1, les observations sont classées selon l'ordre croissant des numéros des lignes et des colonnes<sup>27</sup>.

On notera, à titre d'indication, que la hauteur totale moyenne générale était égale à 131 mm en début d'expérience, avec un écart-type résiduel (dans les « parcelles ») égal à 19 mm et un coefficient de variation égal à 15 %, les valeurs correspondantes étant respectivement de 333 mm, 37 mm et 11 % en fin d'expérience.

<sup>27</sup> Nous aurions bien sûr pu présenter les 144 valeurs individuelles et l'analyse relative à ces valeurs, au lieu des 16 moyennes, les résultats obtenus par les deux approches n'étant guère différents.

**Tableau 8.5.1.** Expérience de chauffage du sol sur *Ficus* : accroissements moyens en hauteur par « parcelle », en mm, en fonction des températures du sol, en degrés centigrades, des lignes et des colonnes.

Temp.	Lignes	Colon.	Haut.
20	1	1	185
30	1	2	242
15	1	3	177
25	1	4	214
15	2	1	117
25	2	2	229
20	2	3	209
30	2	4	238
30	3	1	200
20	3	2	200
25	3	3	222
15	3	4	154
25	4	1	218
15	4	2	174
30	4	3	247
20	4	4	205

## 8.5.2 Analyse des résultats<sup>28</sup>

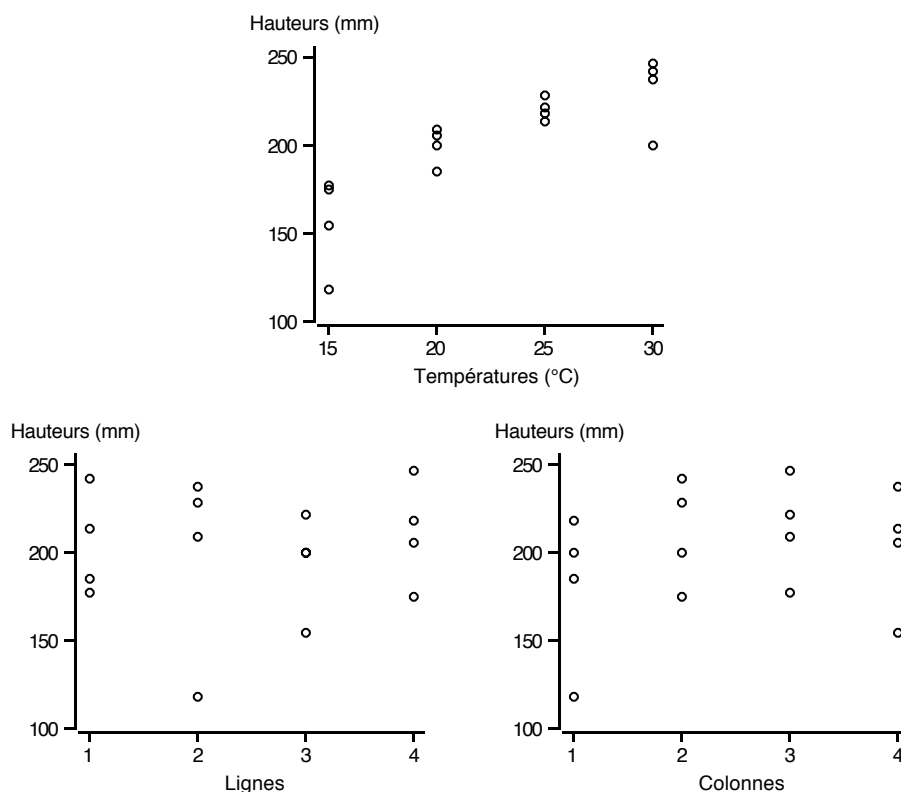
### 1° Examen préliminaire

La figure 8.5.2 donne une représentation graphique de l'influence des trois facteurs considérés. On peut constater que le facteur température du sol joue un rôle prépondérant, que le facteur lignes est à première vue négligeable, et que le facteur colonnes est également assez marqué.

En ce qui concerne le premier facteur, la croissance en hauteur est d'autant plus rapide que la température du sol est plus élevée, et cela selon une relation apparemment curvilinéaire. Quant au facteur colonnes, c'est la différence entre les positions centrales et latérales, dans les deux serres, qui apparaît.

L'examen des différents graphiques qui constituent la figure 8.5.2 met aussi en évidence le fait que la valeur 117 pourrait être considérée comme douteuse. L'application de la méthode de GRUBBS [STAT2, § 3.5.3] aux résidus de l'analyse de la variance dont il est question ci-dessous n'indique cependant pas que cette valeur doit être rejetée, et rien dans les données détaillées, relatives aux neuf plantes considérées, ne semble devoir conduire à une telle conclusion.

<sup>28</sup> Un traitement complet de cet exemple à l'aide du logiciel R est présenté à l'adresse <[www.dagnelie.be/docexr/exp085/exp085.htm](http://www.dagnelie.be/docexr/exp085/exp085.htm)>.



**Figure 8.5.2.** Expérience de chauffage du sol sur *Ficus* : représentation graphique de l'influence des trois facteurs considérés.

D'autre part, un examen de l'ensemble des 144 valeurs initiales montre que les conditions de normalité et d'égalité des variances, sous-jacentes à l'analyse de la variance, sont satisfaites.

## 2° Analyse de la variance et test d'additivité

Le tableau 8.5.2 donne les résultats de l'analyse de la variance, les trois facteurs étant comparés à la variation résiduelle.

Les facteurs lignes et colonnes, liés aux différences entre serres, entre expositions, etc., ne peuvent évidemment pas être considérés comme aléatoires. Pour que les différents tests soient valables, il s'impose donc de vérifier la condition d'additivité du modèle mathématique.

**Tableau 8.5.2.** Expérience de chauffage du sol sur *Ficus* : tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	$F$	P
Températures	3	13.616	4.539	43,0 ***	0,0002
Lignes	3	661	220	2,09	0,20
Colonnes	3	2.833	944	8,95 *	0,012
Variation résiduelle	6	633	105,5		
Totaux	15	17.743			

En appliquant la relation du paragraphe 8.2.1.2°, on obtient, pour le test de TUKEY, la somme des carrés des écarts :

$$SCE_{add} = (-56.449)^2 / 14.053.271 = 227,$$

et un carré moyen résiduel recalculé égal à :

$$(633 - 227) / 5 = 81,2.$$

La valeur  $F$  de FISHER-SNEDECOR est en conséquence égale à 2,80, avec 1 et 5 degrés de liberté, et la probabilité correspondante est égale à 0,16.

L'hypothèse d'additivité ne doit donc pas être rejetée et, sur base du tableau 8.5.2, on peut en conclure qu'il y a bien une influence très hautement significative du facteur températures et une influence significative du facteur colonnes.

### 3° Efficacité relative

Le tableau 8.5.2 montre aussi que, par rapport à une expérience complètement aléatoire, l'efficacité relative de l'expérience en carré latin est, en première approximation, de l'ordre de 330 % :

$$[(661 + 2.833 + 633) / (3 + 3 + 6)] / 105,5 = 343,9 / 105,5 = 3,26 \text{ ou } 326 \text{ \%}.$$

On peut s'interroger en outre quant à l'efficacité relative de l'expérience en carré latin par rapport à une éventuelle expérience en blocs aléatoires complets.

A priori, en tenant compte de l'existence de deux serres et des différences qu'on pouvait attendre entre les expositions nord et sud, on aurait pu imaginer une expérience organisée en quatre blocs aléatoires complets correspondant aux quatre lignes de la figure 8.5.1. Par rapport à une telle expérience, qui n'aurait pas fait intervenir le facteur colonnes, l'efficacité relative du carré latin serait, en première approximation :

$$[(2.833 + 633) / (3 + 6)] / 105,5 = 385,1 / 105,5 = 3,65 \text{ ou } 365 \text{ \%}.$$

L'expérience en carré latin s'avère donc nettement supérieure dans ce cas, à la fois, à l'expérience complètement aléatoire et à l'expérience en blocs aléatoires complets qui aurait pu être envisagée.

#### 4° Interprétation

Pour compléter l'interprétation des résultats, il est utile de poursuivre l'étude du facteur températures et, dans une moindre mesure, des facteurs lignes et colonnes. Cette étude complémentaire peut se faire par une subdivision complète des différentes sommes des carrés des écarts factorielles.

En ce qui concerne le premier facteur, on peut tester le caractère linéaire ou quadratique de la relation entre la croissance en hauteur et la température du sol, par la méthode des polynômes orthogonaux [STAT2, § 12.2.3] ou par régression multiple [STAT2, § 16.2]. Quant au facteur lignes, il peut être décomposé en un facteur serres, un facteur orientations (nord-sud) et une interaction serres-orientations. De même, la somme des carrés des écarts entre colonnes peut donner naissance à un facteur orientations (ouest-est), un facteur positions (centrales ou latérales) et une interaction orientations-positions. Le tableau 8.5.3 présente les résultats de ces différentes décompositions.

**Tableau 8.5.3.** Expérience de chauffage du sol sur *Ficus* : tableau d'analyse de la variance, après subdivision des sommes des carrés des écarts factorielles.

Sources de variation	Degrés de liberté	S. des carrés des écarts	Carrés moyens	<i>F</i>	P
Températures	3	13.616	4.539	43,0 ***	0,0002
Régression linéaire	1	12.475	12.475	118 ***	0,0000
Régression quadratique	1	1.106	1.106	10,5 *	0,018
Régression cubique	1	35	35	0,33	0,59
Lignes	3	661	220	2,09	0,20
Serres	1	5	5	0,05	0,83
Orientations (N-S)	1	116	116	1,10	0,33
Serres-orientations	1	540	540	5,12	0,064
Colonnes	3	2.833	944	8,95 *	0,012
Orientations (O-E)	1	638	638	6,05 *	0,049
Positions	1	1.785	1.785	16,9 **	0,0063
Orientations-positions	1	410	410	3,89	0,096
Variation résiduelle	6	633	105,5		
Totaux	15	17.743			

Ce tableau confirme très largement les premières indications que la figure 8.5.2 avait données, à savoir l'importance des facteurs températures et, dans une moindre mesure, positions (centrales ou latérales) dans les serres.

En ce qui concerne le facteur températures, on peut aussi rechercher une équation du second degré, du type :

$$H = a + bT + cT^2,$$

si on désigne la croissance en hauteur par  $H$  et la température par  $T$ . L'ajustement d'une telle équation par la méthode des moindres carrés conduit au résultat suivant [STAT2, § 16.2.2] :

$$H = -68,39 + 19,96T - 0,3325T^2.$$

Le tableau 8.5.4 donne, pour les quatre températures considérées, les moyennes observées et les moyennes « théoriques » déduites de cette équation. On peut constater la très bonne adéquation de la régression du deuxième degré, dans le domaine étudié.

**Tableau 8.5.4.** Expérience de chauffage du sol sur *Ficus* : moyennes observées et théoriques relatives aux différentes températures du sol.

Températures	Moyennes	
	observées	théoriques
15°	156	156
20°	200	198
25°	221	223
30°	232	231

Enfin, comme au paragraphe 6.5.3 notamment, on pourrait s'efforcer de déterminer une température optimale de culture.

### ⊖ 5° Codification des modèles d'analyse de la variance

Le contenu du tableau 8.5.2 peut être obtenu à l'aide des modèles :

$$H = T L C \quad \text{et} \quad H \sim T + L + C,$$

respectivement pour les logiciels Minitab et SAS d'une part (commandes ou procédures « *anova* » ou « *glm* »), et R d'autre part (fonction « *aov* »),  $H$ ,  $T$ ,  $L$  et  $C$  désignant les accroissements en hauteur et les facteurs températures, lignes et colonnes.

Quant au tableau 8.5.3, une solution consiste à définir une série de variables auxiliaires, relatives aux composantes linéaire, quadratique et cubique du facteur températures, et aux facteurs subsidiaires serres, orientations nord ou sud, positions centrales ou latérales, et orientations est ou ouest.

Les variables auxiliaires relatives au facteur températures peuvent être trois polynômes orthogonaux [STAT2, § 12.2.3]. Il peut s'agir par exemple d'une varia-

ble T1 prenant les valeurs  $-3$ ,  $-1$ ,  $+1$  et  $+3$  respectivement pour  $15$ ,  $20$ ,  $25$  et  $30^\circ$ , en ce qui concerne la composante linéaire, et de même T2 prenant les valeurs  $+1$ ,  $-1$ ,  $-1$  et  $+1$  pour la composante quadratique, et T3 prenant les valeurs  $-1$ ,  $+3$ ,  $-3$  et  $+1$  pour la composante cubique. Les variables auxiliaires relatives aux quatre autres facteurs sont par contre des variables binaires pouvant prendre les valeurs 0 et 1 (ou 1 et 2), et pouvant être désignées respectivement par S (serres), N (nord-sud), P (positions) et E (est-ouest).

Dans ces conditions, le problème peut être traité par une régression multiple classique exprimant les accroissements en hauteur en fonction de ces sept variables et des produits SN et PE, et cela avec les trois logiciels considérés.

Mais, comme dans les différents cas envisagés précédemment, d'autres solutions sont possibles. Ainsi, le modèle :

$$H \sim T1 + T2 + T3 + S*N + P*E,$$

pourrait être utilisé avec la fonction « *avv* » du logiciel R.

## 8.6 Exemple 2 : expérience avec permutation des objets (*cross-over*)

### 8.6.1 Présentation et données

#### 1° Présentation

Nous considérons également à titre d'exemple le cas d'une expérience avec permutation des objets destinée à comparer, sur quelques bovins seulement, trois alimentations différentes. Il s'agit de six vaches laitières, qui ont été suivies individuellement, chacune d'entre elles ayant reçu consécutivement chacune des trois alimentations (A, B et C), selon le dispositif présenté dans le tableau 8.6.1.

**Tableau 8.6.1.** Comparaison de trois alimentations (A, B et C) sur vaches laitières : dispositif expérimental.

Périodes	Vaches					
	1	2	3	4	5	6
1	B	C	A	B	A	C
2	A	A	B	C	C	B
3	C	B	C	A	B	A

Plus concrètement, l'alimentation A est une alimentation témoin, à laquelle les alimentations B et C doivent être comparées, et les périodes d'observation sont

des périodes de quatre semaines, précédées chacune d'une période d'adaptation de deux semaines. Les principales observations concernent la production laitière.

[ À l'examen du tableau 8.6.1, on peut constater que toutes les permutations des lettres A, B et C, au sens de l'analyse combinatoire, sont prises en considération, [ de telle sorte que le dispositif est équilibré pour les effets résiduels.

### 2° Données

Le tableau 8.6.2 présente les productions journalières moyennes de lait, en kilogrammes, en fonction des alimentations, des périodes et des vaches. Ici également, les observations sont classées dans l'ordre des lignes et des colonnes du dispositif expérimental, et nous définirons ultérieurement le contenu de la colonne « Résid. ».

**Tableau 8.6.2.** Comparaison de trois alimentations sur vaches laitières : productions journalières moyennes de lait, en kg.

Aliment.	Périodes	Vaches	Product.	Résid.
B	1	1	21,7	0
C	1	2	20,4	0
A	1	3	25,9	0
B	1	4	23,5	0
A	1	5	19,1	0
C	1	6	19,0	0
A	2	1	19,1	B
A	2	2	19,9	C
B	2	3	24,0	A
C	2	4	20,5	B
C	2	5	17,0	A
B	2	6	19,1	C
C	3	1	16,4	A
B	3	2	19,5	A
C	3	3	22,2	B
A	3	4	21,9	C
B	3	5	20,0	C
A	3	6	17,6	B

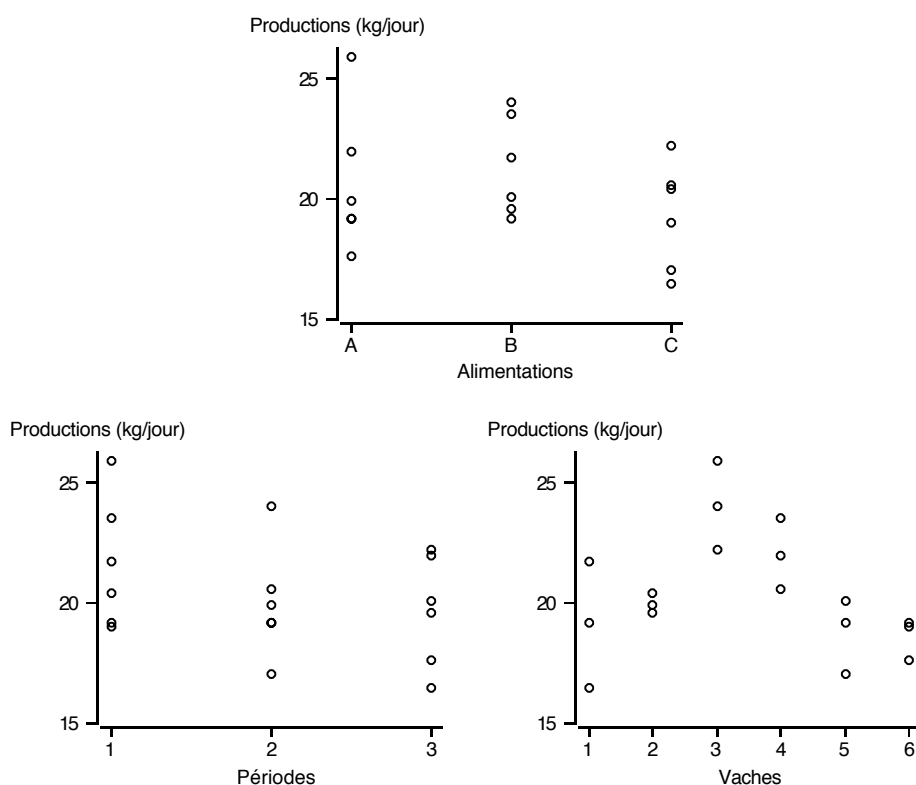
## 8.6.2 Analyse des résultats<sup>29</sup>

### 1° Examen préliminaire

La figure 8.6.1 donne une représentation graphique des résultats en fonction des trois facteurs considérés. On observe surtout l'influence du facteur vaches et, dans

<sup>29</sup> Un traitement complet de cet exemple à l'aide du logiciel R est présenté à l'adresse <[www.dagnelie.be/docexr/exp086/exp086.htm](http://www.dagnelie.be/docexr/exp086/exp086.htm)>.





**Figure 8.6.1.** Comparaison de trois alimentations sur vaches laitières : représentation graphique de l'influence des trois facteurs considérés.

une mesure nettement moindre, l'influence des deux autres facteurs (alimentations et périodes).

La moyenne des productions de lait est égale à 20,38 kg/jour, avec des valeurs extrêmes égales à 16,4 et 25,9 kg/jour.

### 2° Analyse de la variance et test d'additivité

Le tableau 8.6.3 contient les résultats de l'analyse de la variance, les trois facteurs étant comparés à la variation résiduelle.

Le facteur périodes ne peut cependant pas être considéré comme aléatoire, dans la mesure où les périodes se présentent évidemment dans un ordre systématique, et aussi, dans la mesure où, toutes choses étant égales par ailleurs, la production laitière est normalement décroissante d'une période à l'autre. Il importe donc de vérifier la condition d'additivité du modèle mathématique.

**Tableau 8.6.3.** Comparaison de trois alimentations sur vaches laitières :  
tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	$F$	P
Alimentations	2	12,9878	6,4939	10,0 **	0,0066
Périodes	2	13,7778	6,8889	10,6 **	0,0056
Vaches	5	71,6978	14,3396	22,1 ***	0,0002
Variation résiduelle	8	5,1878	0,6485		
Totaux	17	103,6512			

Ce contrôle peut être réalisé selon une procédure semblable à celle que nous avons présentée et appliquée dans le cas du carré latin (§ 8.2.1.2° et 8.5.2.2°). On obtient ainsi une somme des carrés des écarts relative à la non-additivité égale à 0,3016, une valeur  $F$  de FISHER-SNEDECOR égale à 0,43, avec 1 et 7 degrés de liberté, et une probabilité égale à 0,53. La conclusion de ce test est donc tout à fait favorable.

### 3° Efficacité relative

Le tableau 8.6.3 confirme tout d'abord l'importance du facteur vaches, auquel correspond la contribution la plus élevée, en termes de sommes des carrés des écarts et de carrés moyens. Mais ce tableau montre aussi que les facteurs alimentations et périodes sont nettement plus marqués que ne semblait l'indiquer la figure 8.6.1.

L'importance des facteurs périodes et vaches conduit d'ailleurs à une valeur particulièrement élevée de l'efficacité relative. Par rapport à une expérience complètement aléatoire, cette efficacité serait en effet, en première approximation :

$$[(13,7778 + 71,6978 + 5,1878)/(2 + 5 + 8)]/0,6485 = 6,0442/0,6485 = 9,32 \text{ ou } 932 \text{ \%}.$$

Dans les conditions considérées, pour atteindre une précision comparable à celle que donnent six vaches observées au cours de trois périodes successives, il faudrait disposer, en une seule période, de trois lots d'au moins 50 bêtes ( $9,32 \times 6 = 56$ ), soit un total d'au moins 150 têtes de bétail.

### 4° Interprétation

Quant au facteur alimentations, qui constitue la raison d'être de l'expérience, l'analyse de la variance montre qu'il doit être considéré comme hautement significatif.

En outre, en vue de répondre à la question posée au départ, les alimentations B et C doivent être comparées à l'alimentation A, considérée comme témoin. Dans l'optique d'identifier une ou des alimentations qui permettent d'obtenir une production supérieure à celle de l'alimentation de référence, cette comparaison peut être réalisée par la méthode de DUNNETT [STAT2, § 12.3.2].

Les moyennes observées sont :

$$\bar{x}_A = 20,58, \quad \bar{x}_B = 21,30 \quad \text{et} \quad \bar{x}_C = 19,25 \text{ kg/jour};$$

et la plus petite différence significative, au sens de DUNNETT, est égale à :

$$2,22 \sqrt{2(0,6485)/6} = 1,03 \text{ kg/jour.}$$

Aucune des deux alimentations B et C ne s'avère donc significativement supérieure à l'alimentation de référence A.

### ⊖ 5° Effets résiduels

Ainsi que nous l'avons signalé, le dispositif expérimental qui a été adopté est équilibré pour les effets résiduels et permet donc de vérifier l'existence éventuelle de tels effets.

En vue d'atteindre cet objectif, il y a lieu d'introduire le facteur supplémentaire à quatre modalités qui figure dans la dernière colonne du tableau 8.6.2. Il s'agit, d'une part, de 0 pour la première période, et d'autre part, des lettres A, B ou C relatives chaque fois à la période antérieure, dans le cas de la deuxième et de la troisième période.

Le tableau 8.6.4 présente les résultats de l'analyse de la variance, non orthogonale, qui fait intervenir ce facteur supplémentaire.

**Tableau 8.6.4.** Comparaison de trois alimentations sur vaches laitières : tableau d'analyse de la variance incluant les éventuels effets résiduels.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Alimentations	2	12,9878	6,4939		
Périodes	2	13,7778	6,8889		
Vaches	5	71,6978	14,3396		
Effets résid. (ajust.)	2	2,8536	1,4268	3,67	0,091
Variation résiduelle	6	2,3342	0,3890		
Totaux	17	103,6512			

Dans ce tableau, le facteur effets résiduels est un facteur ajusté, qui tient compte de la présence simultanée, ou préalable, des facteurs alimentations, périodes et vaches. La somme des carrés des écarts qui est associée à ce facteur correspond en effet à la réduction de la somme des carrés des écarts résiduelle qui résulte de l'introduction de ce facteur supplémentaire dans le modèle d'analyse de la variance qui contient déjà les trois autres facteurs [STAT2, § 10.4.4 et 16.4.4].

On peut remarquer aussi que, du fait des relations qui existent entre les différents facteurs, le nombre de degrés de liberté relatif aux effets résiduels est égal à 2, alors que ce facteur possède quatre modalités.

Les effets résiduels apparaissent ainsi comme non significatifs, ce qui ne met donc pas en cause l'analyse de la variance du tableau 8.6.3 et l'interprétation qui en a été donnée. Il faut toutefois noter que cette conclusion est sans doute liée dans une large mesure au faible nombre de degrés de liberté de la variation résiduelle, qui est fonction lui-même du petit nombre de bêtes mises en expérience.

### ⊖ 6° Codification des modèles d'analyse de la variance

La codification de l'analyse de la variance initiale (tableau 8.6.3) est particulièrement simple. Elle se présente sous la forme :

$$L = A P V \quad \text{ou} \quad L \sim A + P + V,$$

respectivement pour les logiciels Minitab et SAS d'une part (commandes ou procédures « *anova* » ou « *glm* »), et R d'autre part (fonction « *aov* »), L, A, P et V désignant les productions laitières et les facteurs alimentations, périodes et vaches.

Quant au tableau 8.6.4, on a de même :

$$L = A P V R \quad \text{ou} \quad L \sim A + P + V + R,$$

R désignant en outre le facteur effets résiduels. On notera qu'en raison du caractère non orthogonal du modèle envisagé ici, le facteur R doit nécessairement être placé en dernière position.