

Chapitre 9

Les expériences en blocs aléatoires incomplets : expériences non factorielles

Sommaire

9.1 Principes

9.1.1 Généralités

⊖ 9.1.2 Les blocs incomplets équilibrés

⊖ 9.1.3 Les réseaux carrés équilibrés et les réseaux équilibrés

⊖ 9.1.4 Les carrés latins incomplets

⊖ 9.1.5 Les dispositifs partiellement équilibrés

⊖ 9.1.6 Le cas des surfaces de réponse et des plans optimaux

9.2 Analyse des résultats

⊖ 9.2.1 Les blocs incomplets équilibrés

⊖ 9.2.2 Les réseaux carrés équilibrés et les carrés latins incomplets

⊖ 9.3 Discussion

9.4 Exemple 1 : expérience en blocs incomplets équilibrés

⊖ 9.4.1 Présentation et données

⊖ 9.4.2 Analyse des résultats

9.5 Exemple 2 : expérience en réseau carré équilibré

⊖ 9.5.1 Présentation et données

⊖ 9.5.2 Analyse des résultats

9.1 Principes

9.1.1 Généralités

Nous avons signalé antérieurement que, bien souvent, le dispositif expérimental en blocs aléatoires complets ne convient pas pour l'étude de grands nombres d'objets (plus de 15 ou 20 objets par exemple) et, parfois aussi, pour l'étude de petits nombres d'objets (§ 6.3.6°). De telles restrictions existent également pour le dispositif en carré latin et, dans une moindre mesure, pour le dispositif avec permutation des objets (§ 8.3.1°).

On est donc fréquemment amené à avoir recours à des dispositifs en *blocs incomplets*¹, qui sont tels que les différents objets étudiés ne sont pas tous présents dans chacun des blocs. Ces dispositifs sont particulièrement nombreux et diversifiés.

Au cours de ce chapitre, nous présentons un certain nombre de dispositifs relatifs aux expériences non factorielles (expériences à un facteur ou expériences à deux ou plusieurs facteurs qui ne possèdent pas une structure factorielle, au sens du paragraphe 2.3.1.3°). Au cours du chapitre suivant, nous envisagerons le même problème dans le cas des expériences factorielles, complètes ou fractionnaires.

La bibliographie relative aux expériences en blocs incomplets est très abondante. On peut trouver des informations à leur sujet dans la plupart des livres que nous avons cités dans l'introduction générale, et en particulier dans le livre de KUEHL [2000]. L'ouvrage tout à fait classique de COCHRAN et COX [1957] reste aussi un outil de base en la matière, notamment du fait des nombreuses listes de dispositifs qu'il contient.

⊖ 9.1.2 Les blocs incomplets équilibrés

1° Présentation

Les *blocs incomplets équilibrés*² sont tels que :

- les blocs sont tous de même dimension k , quant au nombre de parcelles ou d'unités expérimentales qu'ils contiennent, ce nombre étant inférieur au nombre d'objets p ;
- les p objets ne sont jamais présents plus d'une fois dans un même bloc ;
- les p objets sont tous répétés un même nombre de fois r ;
- chacun des p objets se trouve associé un même nombre de fois λ , dans les différents blocs, à chacun des $p - 1$ autres objets³.

¹ En anglais : *incomplete blocks*.

² En anglais : *balanced incomplete blocks*, *BIB*.

³ Dans la littérature relative aux blocs incomplets, le terme *variétés* (en anglais : *varieties*) est souvent utilisé à la place de *traitements* ou *objets*. Ces dispositifs ont en effet été développés très largement en fonction des problèmes rencontrés dans l'étude de grands nombres de variétés de plantes cultivées.

Les deux dernières conditions sont celles qui assurent le caractère équilibré de ces dispositifs.

La figure 9.1.1 présente trois exemples de plans d'expériences qui respectent ces différentes conditions. Dans les trois cas, les blocs, qui correspondent aux lignes de la figure, sont constitués de quatre unités expérimentales.

4	2	5	7
2	1	7	6
7	6	3	5
3	1	2	5
6	5	4	1
6	2	3	4
1	3	4	7

8	6	3	1
7	4	5	2
7	3	5	1
6	4	2	8
5	4	6	3
7	2	8	1
7	6	1	4
2	5	8	3
5	4	1	8
6	3	2	7
6	5	7	8
3	4	2	1
3	7	4	8
2	5	6	1

1	10	6	8
7	3	8	1
9	1	4	10
2	7	10	4
5	4	8	3
7	5	6	4
7	2	8	9
4	2	1	3
8	6	4	9
7	9	5	1
3	10	9	5
6	5	2	1
10	3	6	7
2	6	3	9
10	8	5	2

Figure 9.1.1. Exemples de dispositifs en blocs incomplets équilibrés.

Si en plus des notations introduites ci-dessus, on désigne par b le nombre de blocs, les caractéristiques de ces trois dispositifs sont respectivement :

$$p = 7, \quad k = 4, \quad b = 7, \quad r = 4, \quad \lambda = 2,$$

$$p = 8, \quad k = 4, \quad b = 14, \quad r = 7, \quad \lambda = 3,$$

et
$$p = 10, \quad k = 4, \quad b = 15, \quad r = 6, \quad \lambda = 2.$$

Les valeurs du paramètre λ indiquent que les différents couples d'objets sont tous présents deux fois dans un même bloc pour le premier et le troisième exemple, et trois fois dans un même bloc pour le deuxième exemple.

On peut aussi remarquer que, dans le deuxième exemple, les blocs incomplets sont réunis par paires, de manière à constituer sept répétitions complètes. Un tel regroupement n'est par contre pas réalisable dans le premier et le troisième exemple, le nombre de blocs b n'étant pas un multiple du nombre de répétitions r .

Les dispositifs en blocs incomplets qui, comme le deuxième exemple, sont constitués de répétitions complètes sont parfois qualifiés de dispositifs *en répliques* ou

*résolubles*⁴. Une autre situation particulière, qui peut également se présenter, est l'existence de groupes de répétitions complètes.

Le football, comme d'autres sports, fournit des exemples supplémentaires de blocs incomplets équilibrés. Ainsi, une compétition qui met en jeu quatre équipes peut être considérée comme un ensemble de blocs incomplets équilibrés de paramètres :

$$p = 4, \quad k = 2, \quad b = 6, \quad r = 3 \quad \text{et} \quad \lambda = 1,$$

dès le moment où les quatre équipes sont opposées deux à deux au cours de six matchs, chacune d'entre elles jouant trois fois et rencontrant une fois chacune des autres équipes.

De même aussi, un championnat qui oppose 16 équipes en aller-retour se déroule selon un dispositif semblable, dont les paramètres sont :

$$p = 16, \quad k = 2, \quad b = 240, \quad r = 30 \quad \text{et} \quad \lambda = 2.$$

En outre, les 240 matchs sont en général réunis en 30 journées ou 30 week-ends de huit matchs, qui constituent chacune ou chacun une répétition complète.

2° Conditions

On peut facilement démontrer que les conditions suivantes sont nécessaires pour assurer l'existence de blocs incomplets équilibrés :

$$pr = bk \quad \text{et} \quad r(k-1)/(p-1) = \lambda,$$

tous les paramètres étant des nombres entiers positifs.

La première relation permet d'attribuer, pour un total de bk unités expérimentales, un même nombre r de répétitions à chacun des p objets, tandis que la deuxième relation garantit que tous les couples d'objets sont également représentés dans les différents blocs. La formulation de cette deuxième condition provient du fait que, si on considère un objet quelconque et les r blocs dans lesquels cet objet est présent, les $r(k-1)$ unités expérimentales qui, dans ces blocs, ne sont pas affectées à l'objet en question doivent pouvoir être réparties uniformément entre les $p-1$ autres objets.

On notera que ces relations ne sont pas, en général, des conditions suffisantes d'existence de dispositifs en blocs incomplets équilibrés⁵.

⁴ En anglais : *resolvable design*.

⁵ FISHER et YATES [1982] donnent diverses indications quant aux combinaisons des valeurs p , k , b , r et λ qui satisfont les conditions énoncées ci-dessus, mais pour lesquelles la non-existence d'un dispositif en blocs incomplets équilibrés a été démontrée, et quant à certains cas qui n'ont pas été résolus. Parmi les combinaisons pour lesquelles aucune solution n'existe, on peut citer notamment :

$$p = 15, \quad k = 5, \quad b = 21, \quad r = 7 \quad \text{et} \quad \lambda = 2.$$

3° Dispositifs symétriques et complémentaires

Quand, en plus des propriétés énoncées ci-dessus, on a :

$$p = b \quad \text{et} \quad r = k,$$

le dispositif est dit *symétrique*⁶. Tel est le cas notamment pour le premier exemple de la figure 9.1.1.

D'autre part, deux dispositifs en blocs incomplets équilibrés sont dits *complémentaires*⁷ quand chacun des blocs de l'un est constitué des objets qui sont absents dans un bloc de l'autre. Ainsi, l'ensemble complémentaire du premier dispositif de la figure 9.1.1 est formé des blocs suivants :

$$(1, 3, 6), (3, 4, 5), (1, 2, 4), (4, 6, 7), (2, 3, 7), (1, 5, 7) \text{ et } (2, 5, 6).$$

Et le dispositif complémentaire du deuxième exemple de la figure 9.1.1 est ce dispositif lui-même.

On peut démontrer que le complément d'un ensemble de blocs incomplets équilibrés de paramètres p, k, b, r et λ est lui-même un ensemble de blocs incomplets équilibrés, dont les paramètres sont :

$$p' = p, \quad k' = p - k, \quad b' = b, \quad r' = b - r \quad \text{et} \quad \lambda' = \lambda + b - 2r.$$

Sur cette base, les paramètres du dispositif complémentaire du troisième exemple de la figure 9.1.1 sont :

$$p' = 10, \quad k' = 6, \quad b' = 15, \quad r' = 9 \quad \text{et} \quad \lambda' = 5.$$

4° Randomisation

En dehors de la propriété qui vient d'être énoncée en ce qui concerne les dispositifs complémentaires, nous ne donnons pas d'informations particulières relatives à la construction des dispositifs en blocs incomplets équilibrés. Nous renvoyons simplement à l'emploi de certains logiciels ou à la consultation de livres tels que celui de COCHRAN et COX [1957].

Mais on doit bien sûr toujours procéder à la répartition aléatoire des objets. Celle-ci peut être réalisée en attribuant au hasard des numéros d'ordre aux différents objets, en permutant le cas échéant de façon aléatoire les répétitions complètes, en permutant de façon aléatoire les blocs au sein de l'ensemble du dispositif ou éventuellement au sein de chacune des répétitions complètes, et en permutant également de façon aléatoire et indépendamment les objets à l'intérieur de chacun des blocs.

⁶ En anglais : *symmetrical design*.

⁷ En anglais : *complementary design*.

5° Comparaisons avec un ou plusieurs témoins

Diverses extensions du principe des blocs incomplets équilibrés ont été proposées, notamment dans l'optique de la comparaison d'un ensemble d'objets avec un ou plusieurs témoins. La figure 9.1.2 donne trois exemples de dispositifs qui correspondent à cet objectif.

4	0	2	5	7
2	1	0	7	6
0	7	6	3	5
3	0	1	2	5
6	5	4	1	0
6	2	3	4	0
1	3	4	0	7

7	5	1	0
5	2	0	6
5	3	4	0
0	7	3	2
0	4	1	2
6	7	4	0
1	0	6	3

5	0	11	2
0	1	12	8
7	3	4	0
11	10	0	3
8	4	2	0
9	0	1	10
0	6	5	7
0	12	6	9

Figure 9.1.2. Exemples de dispositifs en blocs incomplets équilibrés relatifs aux comparaisons avec un témoin.

Le premier exemple est construit à partir du premier schéma de la figure 9.1.1, en ajoutant simplement un témoin (0) de façon aléatoire dans chacun des sept blocs. Les paramètres correspondants sont :

$$p = 8, \quad k = 5, \quad b = 7, \quad r = 3 \text{ ou } 7 \quad \text{et} \quad \lambda = 2 \text{ ou } 4,$$

le nombre de répétitions étant égal à 7 pour le témoin et à 3 pour les autres objets, et le nombre de présences simultanées de deux objets dans un même bloc étant égal à 4 pour les couples d'objets qui associent le témoin à chacun des autres objets et à 2 pour les autres couples d'objets.

Le deuxième exemple de la figure 9.1.2 permet également de comparer sept objets (1 à 7) à un témoin (0), mais cette fois dans un dispositif de dimensions identiques à celles du premier exemple de la figure 9.1.1. Il s'agit en réalité d'un ensemble initial de blocs incomplets équilibrés de paramètres :

$$p = 7, \quad k = 3, \quad b = 7, \quad r = 3 \quad \text{et} \quad \lambda = 1,$$

dans lequel un témoin a été inséré au hasard dans chacun des blocs. Les paramètres du dispositif résultant sont en conséquence :

$$p = 8, \quad k = 4, \quad b = 7, \quad r = 3 \text{ ou } 4 \quad \text{et} \quad \lambda = 1 \text{ ou } 3.$$

Par comparaison avec le premier exemple de la figure 9.1.1, chacun des objets à comparer au témoin s'est vu priver ici d'une répétition, pour faire place au témoin.

Enfin, le troisième exemple de la figure 9.1.2 a pour but de comparer 12 objets (1 à 12) à un témoin (0). Ses paramètres sont :

$$p = 13, \quad k = 4, \quad b = 8, \quad r = 2 \text{ ou } 8 \quad \text{et} \quad \lambda = 0, 1 \text{ ou } 2,$$

24 des 66 couples d'objets autres que le témoin se trouvant une et une seule fois dans un même bloc et les 42 autres couples de tels objets n'étant jamais présents.

On remarquera que les deux premiers dispositifs sont non seulement équilibrés en ce qui concerne les comparaisons avec le témoin, mais aussi équilibrés pour les comparaisons entre les objets autres que le témoin. Tel n'est par contre pas le cas pour le troisième exemple.

On peut bien sûr envisager aussi d'introduire dans chacun des blocs deux ou plusieurs témoins ou deux ou plusieurs répétitions d'un même témoin, et éventuellement même de ne considérer qu'une seule répétition de chacun des objets autres que le ou les témoins. Dans le dernier dispositif de la figure 9.1.2, il pourrait s'agir par exemple d'une seule répétition de 24 objets, qui seraient associés trois par trois au témoin dans les différents blocs.

Ces diverses possibilités s'inscrivent dans la ligne de ce qui a été dit antérieurement quant à l'opportunité d'accroître le nombre de répétitions du ou des témoins, par rapport au nombre de répétitions des différents objets qui leur sont comparés (§ 2.2.2.2°).

De tels dispositifs sont parfois qualifiés d'*augmentés*⁸. On peut trouver des informations complémentaires à leur sujet dans les articles de BECHHOFFER et TAMHANE [1981], FEDERER et RAGHAVARAO [1975], et JACROUX [1989].

⊖ 9.1.3 Les réseaux carrés équilibrés et les réseaux équilibrés

1° Réseaux carrés équilibrés

Les *réseaux carrés équilibrés*, aussi appelés parfois *treillis carrés équilibrés*⁹, sont des dispositifs expérimentaux constitués à la fois de répétitions complètes et d'un double réseau de blocs incomplets, croisés les uns par rapport aux autres et qui correspondent aux lignes et aux colonnes des répétitions complètes.

La figure 9.1.3 en donne un exemple, pour 16 objets, cinq répétitions, 20 lignes et 20 colonnes de quatre unités expérimentales (quatre blocs-lignes et quatre blocs-colonnes de quatre unités expérimentales dans chacune des cinq répétitions). À l'intérieur des différentes répétitions complètes, tous les couples d'objets sont présents une et une seule fois dans une même ligne, et une et une seule fois dans une même colonne.

⁸ En anglais : *augmented incomplete blocks*.

⁹ En anglais : *balanced lattice square*.

9	1	13	5	7	16	9	2	10	16	5	3	8	9	3	14	2	1	4	3
16	8	12	4	1	10	15	8	7	1	12	14	5	12	2	15	10	9	12	11
2	10	6	14	14	5	4	11	4	6	15	9	6	11	1	16	14	13	16	15
7	15	3	11	12	3	6	13	13	11	2	8	7	10	4	13	6	5	8	7

Figure 9.1.3. Exemple de réseau carré équilibré.

2° Conditions

De tels dispositifs n'existent évidemment que quand le nombre d'objets est un carré ($p = 9, 16, 25, \dots$), le nombre d'unités expérimentales par bloc k et le nombre de répétitions r étant tels que :

$$k = \sqrt{p} \text{ ou } p = k^2 \text{ et } r = k + 1.$$

Dans le cas de la figure 9.1.3, les paramètres du dispositif sont :

$$p = 16, \quad k = 4, \quad b = 2 \times 20, \quad r = 5 \quad \text{et} \quad \lambda = 2 \times 1,$$

les mentions 2×20 et 2×1 indiquant qu'il y a deux systèmes de 20 blocs (20 lignes et 20 colonnes), et que les différents couples d'objets sont présents une et une seule fois dans chacun de ces deux systèmes de blocs.

On notera toutefois que ces dispositifs n'existent pas pour tout nombre d'objets égal à un carré. Il a été démontré par exemple que le cas $p = 36$ n'existe pas.

D'autre part, quand k est un nombre impair ($p = 9, 25, 49, \dots$), il est possible de diminuer de moitié le nombre de répétitions, en considérant que tous les couples d'objets doivent être présents une et une seule fois dans une même ligne *ou* dans une même colonne. Les différents paramètres sont alors tels que :

$$k = \sqrt{p} \text{ ou } p = k^2 \text{ et } r = (k + 1)/2.$$

La figure 9.1.4 présente le cas de 25 objets, trois répétitions, 15 blocs-lignes et 15 blocs-colonnes de cinq unités expérimentales, pour lequel on a :

$$p = 25, \quad k = 5, \quad b = 2 \times 15, \quad r = 3 \quad \text{et} \quad \lambda = 1.$$

3° Randomisation

Des réseaux carrés équilibrés peuvent être générés et randomisés par ordinateur à l'aide de divers logiciels, et des tables de tels dispositifs sont données notamment par COCHRAN et COX [1957].

Dans tous les cas, on doit attribuer au hasard des numéros d'ordre aux différents objets, et on doit procéder aussi à une permutation aléatoire des répétitions, et une permutation aléatoire des lignes d'une part, et des colonnes d'autre part, indépendamment au sein de chacune des répétitions.

18	24	5	4	6	17	1	5	16	19	7	14	8	5	23
21	1	9	14	25	13	23	25	12	4	20	6	17	12	9
19	12	11	15	7	7	2	10	18	9	4	10	11	21	16
23	10	3	17	22	21	6	15	3	8	1	13	18	22	15
20	8	13	2	16	24	11	20	14	22	3	19	25	2	24

Figure 9.1.4. Exemple de demi-réseau carré équilibré.

4° Réseaux équilibrés

Les *réseaux équilibrés* ou *treillis équilibrés*¹⁰ sont des dispositifs analogues aux réseaux carrés équilibrés, qui ne font intervenir, à l'intérieur des répétitions, qu'un seul système de blocs (lignes ou colonnes). Ils existent dans les mêmes conditions que les réseaux carrés équilibrés et peuvent être déduits de ces derniers.

Pour 16 objets par exemple, un dispositif en réseau équilibré peut être établi à partir du schéma de la figure 9.1.3, en permutant de façon aléatoire et indépendamment les quatre unités expérimentales de chacune des 20 colonnes (et non pas les quatre lignes de quatre unités de chacune des cinq répétitions complètes). Ce dispositif comporte alors 20 blocs incomplets de quatre unités, correspondant aux colonnes et groupés en cinq répétitions.

⊖ 9.1.4 Les carrés latins incomplets

1° Présentation

Les *carrés latins incomplets*, aussi appelés *carrés de YOUDEN*¹¹, sont des dispositifs expérimentaux équilibrés qui peuvent être obtenus en éliminant, dans certaines conditions, une ou plusieurs lignes ou une ou plusieurs colonnes des carrés latins classiques (§ 8.1.1). La figure 9.1.5 en donne un exemple, pour 13 objets et quatre répétitions, le nombre de lignes étant égal au nombre d'objets et le nombre de colonnes étant égal au nombre de répétitions.

Comme les carrés latins classiques et les dispositifs avec permutation des objets et comme les réseaux carrés équilibrés, les carrés latins incomplets assurent un double contrôle de la variabilité du matériel expérimental. Ils superposent en effet un ensemble de blocs complets (colonnes de la figure 9.1.5) et un ensemble de blocs incomplets équilibrés (lignes de la figure 9.1.5).

En permutant les lignes et les colonnes du dispositif de la figure 9.1.5, on obtient aussi un schéma tout à fait comparable aux dispositifs avec permutation des objets,

¹⁰ En anglais : *balanced lattice*.

¹¹ En anglais : *incomplete Latin square, YOUDEN's square*.

12	4	3	6
2	7	6	9
10	2	1	4
5	10	9	12
4	9	8	11
9	1	13	3
7	12	11	1
1	6	5	8
3	8	7	10
13	5	4	7
11	3	2	5
6	11	10	13
8	13	12	2

Figure 9.1.5. Exemple de carré latin incomplet.

tel que celui de la figure 8.1.3. Dans le cas présent, il s'agirait d'un dispositif équilibré qui permettrait de comparer 13 objets sur 13 individus, avec quatre répétitions, au cours de quatre périodes consécutives.

2° Conditions

Les conditions d'existence des carrés latins incomplets sont les mêmes que celles des blocs incomplets équilibrés, avec en outre les restrictions suivantes :

$$p = b \quad \text{et} \quad k = r.$$

Les carrés latins incomplets sont donc des dispositifs symétriques, au sens du paragraphe 9.1.2.3°.

En particulier, pour l'exemple de la figure 9.1.5, les paramètres sont :

$$p = b = 13, \quad k = r = 4 \quad \text{et} \quad \lambda = 1.$$

On notera que certains dispositifs en blocs incomplets équilibrés peuvent être construits à partir de carrés latins incomplets, par une permutation aléatoire des objets à l'intérieur des lignes. Ainsi, le premier dispositif de la figure 9.1.1 pourrait être obtenu à partir d'un carré latin incomplet constitué de sept lignes et quatre colonnes.

3° Randomisation

Comme pour les autres dispositifs, des carrés latins incomplets peuvent être générés par ordinateur et des listes figurent dans des livres tels que celui de COCHRAN et COX [1957].

La répartition aléatoire des objets se fait en attribuant au hasard des numéros d'ordre aux différents objets, et en permutant de façon aléatoire, d'une part les lignes, et d'autre part les colonnes du dispositif choisi.

Dans la construction des carrés latins incomplets, on peut aussi appliquer la propriété de complémentarité présentée au paragraphe 9.1.2.3°, en ce qui concerne les dispositifs en blocs incomplets équilibrés. Cette propriété permet notamment d'affirmer qu'au dispositif de la figure 9.1.5 correspond un carré latin incomplet complémentaire de 13 lignes et 9 colonnes, dont les paramètres sont :

$$p' = b' = 13, \quad k' = r' = 9 \quad \text{et} \quad \lambda' = 6.$$

⊖ 9.1.5 Les dispositifs partiellement équilibrés

1° Généralités

Parallèlement aux différents dispositifs équilibrés que nous avons présentés ci-dessus, existent des dispositifs *partiellement équilibrés*¹². D'une manière générale, ceux-ci sont tels que tous les objets y possèdent le même nombre de répétitions, mais tous les couples d'objets n'y sont pas présents un même nombre de fois. La dernière condition du début du paragraphe 9.1.2.1° n'est donc pas remplie, de telle sorte que toutes les comparaisons entre objets ne sont pas également précises ou également puissantes.

Dans les dispositifs les plus classiques, les différents couples d'objets sont divisés en deux catégories. Certains couples sont tous représentés un même nombre de fois, tandis que les autres couples sont tous caractérisés par un autre nombre de présences communes.

D'une manière générale, des informations relatives à ces dispositifs peuvent être trouvées dans différents livres, tels que ceux de COCHRAN et COX [1957], et KUEHL [2000].

2° Blocs incomplets partiellement équilibrés

La figure 9.1.6 donne un exemple de dispositif en *blocs incomplets partiellement équilibrés*¹³, qui présente les caractéristiques dont il vient d'être question. Les paramètres de ce dispositif sont :

$$p = 8, \quad k = 3, \quad b = 8, \quad r = 3 \quad \text{et} \quad \lambda = 0 \text{ ou } 1.$$

Un examen attentif de cette figure montre que les couples 1-3, 2-6, 4-7 et 5-8 sont absents de ce dispositif ($\lambda = 0$), et que les 24 autres couples d'objets y sont tous présents une fois ($\lambda = 1$).

¹² En anglais : *partially balanced design*.

¹³ En anglais : *partially balanced incomplete blocks, PBIB*.

8	3	7
6	3	5
2	4	3
8	2	1
4	1	5
4	6	8
7	2	5
1	6	7

Figure 9.1.6. Exemple de dispositif en blocs incomplets partiellement équilibrés.

Ce dispositif peut être comparé au deuxième exemple de la figure 9.1.1, dans la mesure où il permet aussi d'étudier huit objets, mais ici en blocs de trois unités expérimentales, au lieu de quatre.

Comme au paragraphe 9.1.2.5°, on pourrait bien sûr envisager aussi d'insérer dans un tel dispositif un ou plusieurs témoins répétés plus fréquemment que les autres objets.

Des tables particulières relatives aux blocs incomplets partiellement équilibrés ont été publiées par BOSE *et al.* [1954] et CLATHWORTHY [1973]. L'utilisation de tables cède toutefois largement la place au recours à des logiciels ou des algorithmes, dont traitent notamment NGUYEN [1994] et VENABLES et ECCLESTON [1993]. On peut mentionner aussi d'autres travaux, tels que ceux de BOFILL et TORRAS [2004], et RECK et MORGAN [2005].

3° Réseaux rectangulaires

Les *réseaux rectangulaires*¹⁴ sont des dispositifs comparables aux réseaux équilibrés du paragraphe 9.1.3.4°. Comme les réseaux carrés équilibrés et les réseaux équilibrés, ils sont constitués de répétitions complètes, mais ils sont tels que le nombre d'objets p est lié à la dimension des blocs k par la relation :

$$p = k(k + 1).$$

Avec des nombres d'objets égaux à 12, 20, 30, ..., ces dispositifs se situent donc presque à mi-chemin entre les nombres d'objets des réseaux carrés équilibrés et des réseaux équilibrés ($p = 9, 16, 25, \dots$).

La figure 9.1.7 présente un réseau rectangulaire de paramètres :

$$p = 20, \quad k = 4, \quad b = 15, \quad r = 3 \quad \text{et} \quad \lambda = 0 \text{ ou } 1.$$

Pour 20 objets, ce dispositif comporte 15 blocs de quatre unités expérimentales, constituant trois répétitions complètes.

¹⁴ En anglais : *rectangular lattice*.

1	5	16	20	11	18	19	16	10	3	20	6
10	12	11	9	12	20	13	2	14	1	18	12
8	3	18	4	7	10	14	8	19	5	13	8
2	6	19	7	9	3	17	5	9	7	16	15
15	13	14	17	6	4	15	1	2	11	4	17

Figure 9.1.7. Exemple de réseau rectangulaire.

Les 15 blocs de quatre parcelles permettent de faire apparaître 90 couples d'objets ($15 \times 4 \times 3/2 = 90$), alors que le nombre total de couples d'objets est égal à 190 ($20 \times 19/2 = 190$). Effectivement, 100 couples d'objets sont absents du dispositif ($\lambda = 0$), et 90 couples y sont présents chacun une fois ($\lambda = 1$).

Avec ses 20 objets, ce dispositif peut être considéré comme se situant à mi-chemin entre le réseau carré équilibré de la figure 9.1.3 (16 objets) et le demi-réseau carré équilibré de la figure 9.1.4 (25 objets). Par son nombre de répétitions, il est aussi plus proche de ce dernier.

4° Réseaux généralisés, α -plans et carrés semi-latins

On qualifie parfois de *réseaux généralisés*¹⁵ l'ensemble des dispositifs en blocs incomplets pour lesquels le nombre d'objets p est un multiple du nombre d'unités expérimentales par bloc k , tandis que le nombre de blocs b est le même multiple du nombre de répétitions r :

$$p = ck \quad \text{et} \quad b = cr,$$

c étant une constante entière positive. Ces dispositifs sont tels que les blocs incomplets peuvent toujours être groupés en répétitions complètes.

On retrouve, parmi ces dispositifs, certains cas particuliers de blocs incomplets équilibrés, tel que le deuxième exemple de la figure 9.1.1, les réseaux carrés équilibrés et les réseaux équilibrés (§ 9.1.3), pour lesquels le facteur multiplicatif c est égal à k , et les réseaux rectangulaires que nous avons évoqués au paragraphe précédent, pour lesquels le facteur c est égal à $k + 1$.

Mais les réseaux généralisés comprennent également d'autres ensembles de dispositifs non équilibrés, tels que les α -plans (*alpha-plans*)¹⁶, les α -plans latinisés¹⁷, les carrés semi-latins¹⁸, et les carrés troyens¹⁹ [BAILEY, 1992; EDMONDSON, 1998; PATTERSON et WILLIAMS, 1976; PATTERSON *et al.*, 1978; PREECE et FREEMAN, 1983].

¹⁵ En anglais : *generalized lattice*.

¹⁶ En anglais : *α -design, alpha-design*.

¹⁷ En anglais : *Latinized α -design*.

¹⁸ En anglais : *semi-Latin square*.

¹⁹ En anglais : *Trojan square*.

Le plus souvent, dans ces différents dispositifs, le nombre λ de présences simultanées des différents couples d'objets est égal à 0 ou 1, comme dans le cas des blocs incomplets partiellement équilibrés et des réseaux rectangulaires. Parfois cependant, le paramètre λ peut prendre les valeurs 0, 1 et 2.

Dans certains cas également, on utilise des blocs de deux dimensions différentes, la différence étant d'une unité seulement et les nombres d'unités expérimentales par bloc pouvant donc être désignés par k et $k + 1$.

La figure 9.1.8 présente, à titre d'illustration de ces différentes possibilités, un carré semi-latin relatif à 12 objets et quatre répétitions, de paramètres [ITCF, 2001] :

$$p = 12, \quad k = 3, \quad b = 16, \quad r = 4 \quad \text{et} \quad \lambda = 0 \text{ ou } 1.$$

8	12	4	1	6	10	2	5	11	3	7	9
3	11	6	9	8	2	1	7	12	5	10	4
1	5	9	4	11	7	8	3	10	6	12	2
2	7	10	5	3	12	4	9	6	1	11	8

Figure 9.1.8. Exemple de carré semi-latin.

Dans cet exemple, les 12 objets sont présents une et une seule fois dans chacune des quatre lignes de 12 unités, et de même, une et une seule fois dans chacun des ensembles de quatre blocs incomplets qui constituent les répétitions complètes. Cette disposition assure, comme les carrés latins notamment (§ 8.1.1), un double contrôle de la variabilité expérimentale, ce qui justifie la dénomination de carré semi-latin. D'autre part, on peut constater aussi que les 16 blocs incomplets de trois unités expérimentales réunissent 48 des 66 couples d'objets ($\lambda = 1$), les 18 autres couples d'objets étant absents ($\lambda = 0$).

L'ensemble des dispositifs dont il est question ici a l'avantage de couvrir une très large gamme de valeurs possibles pour le nombre d'objets p , le nombre de répétitions r , et le nombre d'unités expérimentales par bloc k , ce qui peut s'avérer particulièrement utile en présence de grands nombres d'objets.

Des algorithmes de construction sont présentés notamment par NGUYEN et WILLIAMS [1993] et VENABLES et ECCLESTON [1993], et des généralisations ont été introduites entre autres par EDMONDSON [2002], JOHN *et al.* [2002], et SINHA *et al.* [2002].

5° Autres dispositifs

Au-delà des dispositifs équilibrés et partiellement équilibrés classiques, certains auteurs préconisent d'utiliser des blocs incomplets de dimensions inégales, tels qu'ils se présentent naturellement [JOHNSTONE, 2003; MEAD, 1988, 1990; PEARCE, 1964]. Il peut s'agir par exemple de portées d'animaux d'effectifs inégaux,

au sein desquelles on s'efforce de répartir un certain nombre d'objets de manière aussi équilibrée que possible, sans imposer aucune restriction quant à la dimension des blocs.

D'autre part, des dispositifs en lignes et colonnes relatifs aux comparaisons avec un ou plusieurs témoins ont aussi été décrits par FEDERER [2002], TURE [1994], et WILLIAMS et JOHN [2003], notamment dans le cas où seuls le ou les témoins sont l'objet de répétitions.

⊖ 9.1.6 Le cas des surfaces de réponse et des plans optimaux

1° Surfaces de réponse

L'utilisation de dispositifs expérimentaux en blocs incomplets, au sens large, intervient également dans l'étude des surfaces de réponse, y compris les problèmes de mélanges (§ 2.4.1 et 2.4.2).

Comme premier exemple, on peut envisager le cas du plan composite centré à deux facteurs, qui comporte neuf objets et qui est illustré dans la partie gauche de la figure 2.4.1, en supposant que les contraintes d'organisation de l'expérience imposent de ne faire qu'une seule répétition des huit points périphériques et un nombre très limité de répétitions du point central. L'expérience peut alors être organisée notamment en deux blocs de six unités expérimentales, constitués de la manière suivante :

$$\begin{aligned} & (-1, -1), \quad (-1, +1), \quad (+1, -1), \quad (+1, +1), \quad (0, 0), \quad (0, 0) \\ \text{et} \quad & (-\sqrt{2}, 0), \quad (+\sqrt{2}, 0), \quad (0, -\sqrt{2}), \quad (0, +\sqrt{2}), \quad (0, 0), \quad (0, 0). \end{aligned}$$

Le premier bloc réunit les quatre points factoriels du plan composite et deux répétitions du point central, tandis que le deuxième bloc regroupe les quatre points radiaux du plan composite et deux autres répétitions du point central.

De même, dans le cas du plan composite centré à trois facteurs, qui comporte 15 objets et qui correspond à la partie droite de la figure 2.4.1, on peut envisager de travailler en deux blocs de six unités et un bloc de huit unités expérimentales. Les blocs de six unités doivent alors réunir chacun quatre des huit points factoriels et deux répétitions du point central, tandis que le bloc de huit unités comprend les six points radiaux et deux répétitions supplémentaires du point central.

On dispose ainsi d'une seule répétition de l'ensemble des points périphériques et de six répétitions du point central. La répartition des huit points factoriels en deux groupes de quatre points doit être réalisée autant que possible comme dans le cas des blocs incomplets relatifs aux expériences factorielles (§ 10.1.2). En outre, pour des raisons d'orthogonalité, on peut montrer qu'il y a éventuellement intérêt à modifier quelque peu les valeurs $\pm 1,682$ qui figurent dans les coordonnées des points radiaux (§ 2.4.1.3°).

Des informations complémentaires peuvent être trouvées à ce sujet dans les ouvrages généraux à orientation industrielle que nous avons cités antérieurement, ainsi que dans les travaux de GOOS et DONEV [2006a], PARK et KIM [2002], et TRINCA et GILMOUR [2000, 2002] pour les surfaces de réponse classiques, et de GOOS et DONEV [2006b, 2007], et PRESCOTT et DRAPER [1998] pour les mélanges.

2° Plans optimaux

Enfin, des dispositifs expérimentaux en blocs incomplets peuvent être associés également aux plans optimaux qui ont été présentés au paragraphe 2.4.3. Mais les critères d'optimalité interviennent ici, non pas uniquement dans le choix des traitements ou des objets, mais aussi dans la recherche des dispositifs expérimentaux eux-mêmes.

De tels critères apparaissent en effet dans certaines des publications que nous avons citées au cours des paragraphes précédents, et on peut y ajouter plus globalement le livre de GOOS [2002], ainsi que la discussion de PAGÈS et PÉRINEL [2007].

9.2 Analyse des résultats

⊖ 9.2.1 Les blocs incomplets équilibrés

1° Analyse de la variance

Le tableau 9.2.1 présente le schéma de l'analyse de la variance relative aux expériences en blocs incomplets équilibrés, tant d'une manière générale que pour le premier exemple de la figure 9.1.1.

Tableau 9.2.1. Schéma de l'analyse de la variance relative aux expériences en blocs incomplets équilibrés, dans le cas général et dans le cas particulier $p = b = 7$ et $k = r = 4$.

Sources de variation	Degrés de liberté	
Objets (ajustés)	$p - 1$	6
Blocs	$b - 1$	6
Variation résiduelle	$pr - p - b + 1$	15
Totaux	$pr - 1$	27

Il s'agit d'une analyse à deux critères de classification (objets et blocs), sans distinction entre une éventuelle interaction et la réelle variation résiduelle. Mais il faut noter que cette analyse est non orthogonale, puisque tous les objets ne sont pas présents dans tous les blocs (§ 12.4.1).

La somme des carrés des écarts totale et la somme des carrés des écarts relative au facteur blocs peuvent être calculées de manière tout à fait classique, tandis que la somme des carrés des écarts relative au facteur objets doit être ajustée, de façon à tenir compte de la présence simultanée du facteur blocs dans l'analyse. Enfin, la somme des carrés des écarts « résiduelle » peut être obtenue par différence.

De telles analyses peuvent être réalisées sans problème à l'aide des différents logiciels qui permettent de traiter les cas non orthogonaux d'analyse de la variance. Nous en exposons néanmoins le principe de manière assez détaillée, de telle sorte que le lecteur qui le souhaite puisse comprendre les fondements de ces analyses.

2° Effets principaux et moyennes ajustées

Dans le cas des blocs aléatoires complets, les effets principaux relatifs aux différentes modalités du facteur objets peuvent être estimés de la manière suivante [STAT2, § 10.2.5 et 10.3.6] :

$$\hat{a}_i = \bar{x}_{i.} - \bar{x}_{..},$$

$\bar{x}_{i.}$ désignant les moyennes relatives aux différentes modalités ($i = 1, \dots, p$) et $\bar{x}_{..}$ étant la moyenne générale de toutes les observations. Les moyennes estimées relatives aux différentes modalités sont aussi :

$$\hat{m}_i = \bar{x}_{i.} \quad \text{ou} \quad \hat{m}_i = \bar{x}_{..} + \hat{a}_i.$$

Dans le cas des blocs incomplets par contre, les effets principaux ne peuvent pas être estimés en comparant les moyennes par modalité $\bar{x}_{i.}$ avec la moyenne générale de toutes les observations, mais bien en comparant chacune des moyennes $\bar{x}_{i.}$ avec une moyenne calculée uniquement à partir des observations des blocs dans lesquels la modalité considérée est présente.

Ainsi, pour le premier exemple de la figure 9.1.1, la moyenne $\bar{x}_{1.}$, qui est relative à la première modalité et qui est calculée à partir des quatre unités expérimentales « 1 », doit être comparée à une « moyenne générale » déterminée en fonction uniquement des observations du deuxième, du quatrième, du cinquième et du septième bloc.

Mais cette « moyenne générale », qui doit servir de base de comparaison, n'est pas simplement la moyenne arithmétique des 16 observations des quatre blocs en question, car cette dernière moyenne est en réalité une moyenne pondérée, dans laquelle la première modalité intervient quatre fois ($r = 4$), et chacune des six autres modalités deux fois seulement ($\lambda = 2$).

Une moyenne générale non pondérée doit donc être calculée pour chacune des modalités, en fonction chaque fois des seuls blocs concernés. On peut démontrer que ces différentes moyennes non pondérées sont :

$$[(r + \lambda p - \lambda) \bar{x}'_{i.} - (r - \lambda) \bar{x}_{i.}] / (\lambda p),$$

si on désigne par \bar{x}_i la moyenne des r observations relatives à la modalité i , et par \bar{x}'_i la moyenne des kr observations relatives aux r blocs dans lesquels la modalité i est présente.

On peut en déduire que les effets principaux estimés sont :

$$\hat{a}_i = \frac{k(p-1)}{(k-1)p} (\bar{x}_i - \bar{x}'_i),$$

et des *moyennes ajustées*²⁰ peuvent alors être calculées à l'aide de la même relation que ci-dessus, pour les différentes modalités :

$$\hat{m}_i = \bar{x}_{..} + \hat{a}_i.$$

On peut démontrer aussi, sur la même base, que la somme des carrés des écarts ajustée relative au facteur objets est :

$$\text{SCE}'_a = \frac{k(p-1)r}{(k-1)p} \sum_{i=1}^p (\bar{x}_i - \bar{x}'_i)^2 = \frac{(k-1)pr}{k(p-1)} \sum_{i=1}^p \hat{a}_i^2,$$

alors que la somme des carrés des écarts correspondante est, pour r blocs complets :

$$\text{SCE}_a = r \sum_{i=1}^p (\bar{x}_i - \bar{x}_{..})^2 = r \sum_{i=1}^p \hat{a}_i^2.$$

De même, l'erreur standard de toute différence de deux moyennes ajustées est :

$$\sqrt{\frac{2k(p-1)}{(k-1)pr}} \text{CM}_r,$$

au lieu de, classiquement pour r blocs complets :

$$\sqrt{2\text{CM}_r/r}.$$

3° Blocs incomplets en répétitions complètes

L'analyse de la variance est quelque peu différente quand les blocs incomplets sont groupés en répétitions complètes, comme l'illustre notamment le deuxième exemple de la figure 9.1.1. Le tableau 9.2.2 donne le schéma de l'analyse qui correspond à cette situation et, en particulier, à l'exemple en question.

La seule différence est le fait que la composante « blocs », à $b-1$ degrés de liberté (tableau 9.2.1), est divisée en une composante « répétitions », à $r-1$ degrés de liberté, et une composante « blocs (dans répétitions) », à $b-r$ degrés de liberté.

4° Information intra-blocs et inter-blocs

Nous n'avons envisagé jusqu'à présent que l'étude des comparaisons entre objets qui sont réalisées à l'intérieur des blocs, c'est-à-dire l'étude de l'*information*

²⁰ En anglais : *adjusted mean*.

Tableau 9.2.2. Schéma de l'analyse de la variance relative aux expériences en blocs incomplets équilibrés avec répétitions complètes, dans le cas général et dans le cas particulier $p = 8$, $k = 4$, $b = 14$ et $r = 7$.

Sources de variation	Degrés de liberté	
Objets (ajustés)	$p - 1$	7
Répétitions	$r - 1$	6
Blocs (dans répétitions)	$b - r$	7
Variation résiduelle	$pr - p - b + 1$	35
Totaux	$pr - 1$	55

*intra-blocs*²¹. Quand le nombre de blocs est important (au moins 10 à 15 blocs), et que le nombre d'objets et le nombre d'unités expérimentales par bloc sont relativement faibles, une information complémentaire utile relative aux différences entre objets peut être déduite des différences entre blocs. On peut alors procéder éventuellement à une certaine *recupération*²² de cette information, dite *inter-blocs*²³.

Dans le dispositif qui sera présenté au paragraphe 9.4.1.1° par exemple (tableau 9.4.1), on peut remarquer que la différence entre les deux premiers blocs (feuilles 1 et 2) se confond en réalité avec la différence entre les objets 4 et 5, et il en est de même pour les différences entre les blocs 3 et 8 d'une part, et 5 et 9 d'autre part. Il en est de même aussi, de manière équilibrée, pour les différences entre d'autres couples de blocs et les différences entre les autres couples d'objets.

Il apparaît donc bien que la comparaison des blocs peut fournir une certaine information relative aux différences entre objets.

⊖ 9.2.2 Les réseaux carrés équilibrés et les carrés latins incomplets

1° Analyse de la variance

Nous ne présentons pas de façon détaillée les modalités d'analyse des résultats pour tous les types de dispositifs que nous avons introduits au paragraphe 9.1. Mais nous donnons toutefois, dans les tableaux 9.2.3 et 9.2.4, les schémas des analyses de la variance qui concernent les réseaux carrés équilibrés et les carrés latins incomplets (§ 9.1.3 et 9.1.4), et cela chaque fois dans le cas général et pour les exemples que nous avons envisagés (figures 9.1.3 et 9.1.5).

Dans un cas comme dans l'autre, il s'agit d'analyses de la variance non orthogonales, dans lesquelles le facteur objets doit être ajusté en fonction des autres

²¹ En anglais : *intra-block information*.

²² En anglais : *recovery*.

²³ En anglais : *inter-block information*.

Tableau 9.2.3. Schéma de l'analyse de la variance relative aux expériences en réseaux carrés équilibrés, dans le cas général et dans le cas particulier $p = 16$, $k = 4$, $b = 2 \times 20$ et $r = 5$.

Sources de variation	Degrés de liberté	
Objets (ajustés)	$p - 1$	15
Répétitions	$r - 1$	4
Lignes (dans répétitions)	$r(k - 1)$	15
Colonnes (dans répétitions)	$r(k - 1)$	15
Variation résiduelle	$pr - p - 2rk + r + 1$	30
Totaux	$pr - 1$	79

Tableau 9.2.4. Schéma de l'analyse de la variance relative aux expériences en carrés latins incomplets, dans le cas général et dans le cas particulier $p = b = 13$ et $k = r = 4$.

Sources de variation	Degrés de liberté	
Objets (ajustés)	$p - 1$	12
Lignes	$p - 1$	12
Colonnes	$r - 1$	3
Variation résiduelle	$(p - 1)(r - 2)$	24
Totaux	$pr - 1$	51

facteurs. Il en est de même d'ailleurs, d'une manière générale, pour les autres dispositifs que nous avons présentés antérieurement et que nous ne reprenons pas en considération ici.

On notera aussi que, pour les réseaux carrés équilibrés (tableau 9.2.3), les nombres de degrés de liberté peuvent être exprimés comme suit en fonction du seul paramètre k :

$$k^2 - 1, \quad k, \quad k^2 - 1, \quad k^2 - 1, \quad (k^2 - 1)(k - 2) \quad \text{et} \quad k^3 + k^2 - 1.$$

2° Effets principaux et moyennes ajustées

Des ajustements semblables à ceux que nous avons présentés au paragraphe 9.2.1 doivent également intervenir dans l'estimation des effets principaux, dans l'estimation des moyennes et dans la détermination des erreurs standards relatives aux différences de moyennes. Ces ajustements doivent évidemment être adaptés dans chaque cas au type de dispositif considéré.

Nous envisageons le cas des réseaux carrés équilibrés au paragraphe 9.5.2.3° et d'autres informations peuvent être trouvées à ce sujet dans certains des ouvrages généraux que nous avons cités, dont celui de KUEHL [2000].

⊖ 9.3 Discussion

1° Avantages et inconvénients

Les dispositifs en blocs incomplets, au sens large, permettent tout d'abord de faire face aux situations dans lesquelles l'emploi de blocs de très petites dimensions s'impose (blocs de deux unités expérimentales par exemple). D'autre part, le recours à ces dispositifs permet souvent d'augmenter la précision des résultats quand le nombre d'objets à étudier est élevé ou très élevé (plus de 15 ou 20 objets par exemple).

L'utilisation des dispositifs en blocs incomplets implique toutefois quelques difficultés supplémentaires, de calcul et d'interprétation, parmi lesquelles on peut citer la réalisation d'analyses de la variance non orthogonales et l'emploi de moyennes ajustées.

2° Utilisation

En raison sans doute de ces difficultés, les dispositifs en blocs incomplets étaient traditionnellement peu utilisés. En matière agronomique par exemple, différentes sources d'informations relatives à un large éventail de pays et de conditions montrent qu'environ 60 % des expériences réalisées dans les stations de recherche faisaient appel uniquement au principe des blocs aléatoires complets (chapitre 6), environ 30 % des expériences étaient constituées de blocs aléatoires complets avec parcelles divisées, de carrés latins et de dispositifs apparentés (chapitres 7 et 8), et environ 10 % des expériences faisaient intervenir des blocs incomplets, avec confusion d'effets dans le cas des expériences factorielles (chapitres 9 et 10) [CLAUSTRIAUX, 1977 ; COX, 1950 ; KHOSLA *et al.*, 1973 ; OGER, 1980].

Sans qu'on puisse se référer à des informations très générales à ce propos, on peut penser que la situation a sensiblement évolué au cours du temps, du fait notamment du développement des moyens de calcul mis à la disposition des chercheurs. En particulier, les réseaux carrés et rectangulaires et les α -plans ou, d'une manière plus large, les réseaux généralisés (§ 9.1.5) seraient très fréquemment utilisés dans les expériences de comparaison de variétés de plantes cultivées [ITCF, 2001 ; OGER et ROISIN, 1992 ; PATERSON, 1988].

3° Choix

Dans le choix d'un dispositif en blocs incomplets, il y a lieu de privilégier dans la mesure du possible deux propriétés essentielles, qui ne sont pas toujours conciliables, à savoir le caractère équilibré du dispositif et le fait que les blocs puissent être groupés en répétitions complètes.

Le caractère équilibré du dispositif se justifie bien sûr pleinement quand tous les éléments pris en considération doivent être comparés sur pied d'égalité. Quant au groupement des blocs incomplets en répétitions complètes, il permet éventuel-

lement de traiter les résultats, de façon quelque peu abusive sans doute, comme s'il s'agissait de blocs complets, notamment quand un nombre important de données sont manquantes ou quand un certain nombre d'objets ou un certain nombre de blocs doivent être écartés de l'analyse.

Le tableau 9.3.1 permet d'orienter le choix en fonction de ces principes. Ce tableau donne en effet une liste de dispositifs en blocs incomplets équilibrés, au sens large, en faisant la distinction entre les blocs incomplets équilibrés proprement dits, avec répétitions complètes (B1) et sans répétitions complètes (B2), les carrés latins incomplets (C) et les réseaux carrés équilibrés (R). On se rappellera à ce propos, d'une part, que tous les carrés latins incomplets (C) peuvent être considérés aussi comme des blocs incomplets équilibrés sans répétitions complètes (B2), et d'autre part, que tous les dispositifs affectés d'une des mentions B1, C ou R comprennent, sous une forme ou sous une autre, des répétitions complètes.

Il est bon de noter également que l'existence de dispositifs équilibrés pour certains nombres d'objets seulement n'est pas un inconvénient aussi contraignant qu'on pourrait le croire à première vue, dans la mesure où il est souvent possible, et même parfois souhaitable, de répéter un plus grand nombre de fois certains objets de référence (§ 2.2.2.2°). Le dispositif en réseau carré équilibré 5×5 ($p = 25$ et $k = 5$), avec trois, six ou neuf répétitions complètes, peut ainsi être utilisé par exemple dans le cas d'expériences destinées à comparer 20, 21 ou 22 nouvelles variétés avec une variété témoin, pour autant que le nombre de répétitions de celle-ci soit multiplié par 5, par 4 ou par 3.

Dans la mesure du possible, il est souhaitable enfin de donner la préférence aux dispositifs qui assurent un double contrôle de la variabilité expérimentale (dispositifs « en lignes et colonnes »), à savoir les carrés latins incomplets, les réseaux carrés équilibrés, et parmi les dispositifs non équilibrés, les carrés semi-latins et les α -plans latinisés notamment.

4° Efficacité relative

Comme pour les autres dispositifs (§ 6.3.2°, 6.3.3°, etc.), l'efficacité relative des expériences en blocs incomplets peut être calculée a posteriori dans tous les cas, par rapport à l'expérience complètement aléatoire correspondante. L'efficacité relative peut aussi être calculée par rapport à une expérience en blocs aléatoires complets, quand les blocs incomplets peuvent être groupés en répétitions complètes.

D'autre part, le quotient $[k(p-1)]/[(k-1)p]$, qui intervient notamment dans l'erreur standard de la différence de deux moyennes ajustées (§ 9.2.1.2°), peut être considéré comme une efficacité relative a priori.

Ainsi que nous l'avons signalé au paragraphe 6.3.6°, LIN et BINNS [1984] et PATTERSON et HUNTER [1983] donnent diverses indications quant à l'efficacité relative des expériences en blocs incomplets, par rapport aux expériences en blocs complets, et quant aux dimensions à donner aux blocs.

Tableau 9.3.1. Dispositifs en blocs incomplets équilibrés, en carrés latins incomplets et en réseaux équilibrés : nombres d'objets p (< 100), nombres d'unités expérimentales par bloc k , nombres de répétitions r (≤ 10) et types de dispositifs (B1 = blocs incomplets équilibrés avec répétitions complètes ; B2 = blocs incomplets équilibrés sans répétitions complètes ; C = carrés latins incomplets ; R = réseaux carrés équilibrés).

p	k	r	Type	p	k	r	Type	p	k	r	Type
3	2	4, 6, 8, 10	C	10	2	9	B1	21	3	10	B1
4	2	3, 6, 9	R		3	9	C		5	5, 10	C
	3	3, 6, 9	C		4	6	B2		7	10	B2
5	2	4, 8	C		5	9	B2	25	4	8	C
	3	6	C		6	9	B2		5	3, 6, 9	R
	4	4, 8	C		9	9	C		9	9	C
6	2	5, 10	B1	11	2	10	C	28	4	9	B1
	3	5, 10	*		5	5, 10	C		7	9	B2
	4	10	B2		6	6	C	31	6	6	C
	5	5, 10	C		10	10	C		10	10	C
7	2	6	C	13	3	6	C	37	9	9	C
	3	3, 6, 9	C		4	4, 8	C	41	5	10	C
	4	4, 8	C		9	9	C	49	7	4, 8	R
	6	6	C	15	3	7	B1	57	8	8	C
8	2	7	B1		7	7	C	64	8	9	R
	4	7	B1		8	8	C	73	9	9	C
	7	7	C	16	4	5, 10	R	81	9	5, 10	R
9	2	8	C		6	6	C	91	10	10	C
	3	4, 6, 8, 10	R		6	9	B2				
	4	8	C		10	10	C				
	5	10	C	19	3	9	C				
	6	8	B2		9	9	C				
	8	8	C		10	10	C				

* B2 pour 5 répétitions et B1 pour 10 répétitions.

9.4 Exemple 1 : expérience en blocs incomplets équilibrés

⊖ 9.4.1 Présentation et données

1° Présentation

Nous illustrons le cas des expériences en blocs incomplets équilibrés en considérant une partie des données de l'exemple initial de YATES [1936]. L'étude en

question était destinée à comparer cinq traitements de lutte contre la mosaïque du tabac (maladie virale du tabac).

Le schéma de l'expérience est présenté dans la deuxième et la quatrième colonne du tableau 9.4.1, par les chiffres 1 à 5 qui y désignent les différents traitements.

Tableau 9.4.1. Étude de la mosaïque du tabac : numéros des traitements et nombres de lésions par demi-feuille.

Fe.	Moitié gauche		Moitié droite	
1	5	26	2	40
2	4	16	2	26
3	3	21	5	14
4	2	11	3	16
5	5	12	1	12
6	2	34	1	49
7	3	69	1	68
8	3	42	4	35
9	1	22	4	31
10	5	19	4	25

Il s'agit en fait de 10 feuilles de tabac dont la moitié gauche a été l'objet dans chaque cas d'un premier traitement et la moitié droite d'un autre traitement. La première feuille, par exemple, s'est vu affecter les traitements 5 et 2, la deuxième feuille les traitements 4 et 2, etc.

Chacune des feuilles constitue ainsi un bloc de deux unités expérimentales, de telle sorte que les paramètres du dispositif considéré sont :

$$p = 5, \quad k = 2, \quad b = 10, \quad r = 4 \quad \text{et} \quad \lambda = 1,$$

les 10 couples de traitements (1-2, 1-3, ...) apparaissant bien chacun une fois dans le dispositif²⁴.

2° Données

Les données récoltées sont relatives aux nombres de lésions de mosaïque observés sur chaque demi-feuille. Ces données figurent dans la troisième et la cinquième colonne du tableau 9.4.1, en regard des numéros des traitements.

²⁴ Les données considérées par YATES concernaient en réalité, non pas 10 feuilles, mais bien 30 feuilles (ou 30 blocs), les paramètres du dispositif étant :

$$p = 5, \quad k = 2, \quad b = 30, \quad r = 12 \quad \text{et} \quad \lambda = 3.$$

Le fait de n'étudier ici que 10 feuilles doit permettre au lecteur qui le souhaite de suivre plus facilement la procédure de calcul, à partir d'un nombre très limité d'observations (20 observations).

⊖ 9.4.2 Analyse des résultats

1° Examen préliminaire

Un examen même rapide du tableau 9.4.1 met clairement en évidence la grande variabilité des résultats obtenus (valeurs extrêmes égales à 11 et 69, avec une moyenne égale à 29,4 lésions par demi-feuille). En raison notamment de cette variabilité, la représentation graphique des observations en fonction des différents traitements ne fournit guère d'indications précises (figure 9.4.1).

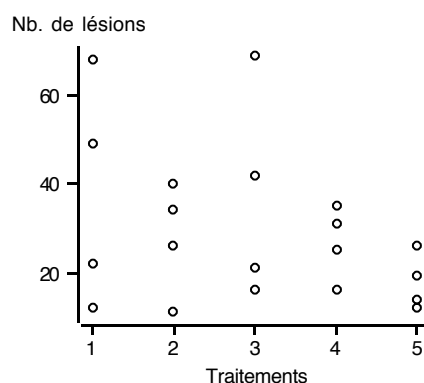


Figure 9.4.1. Étude de la mosaïque du tabac : représentation graphique de l'influence des différents traitements.

La variabilité des résultats et le type d'observations considéré (nombre de lésions) pourraient suggérer l'emploi d'une transformation racine carrée [STAT2, § 4.3.2]. Dans l'optique d'une présentation permettant de suivre éventuellement de façon détaillée la marche des calculs, nous préférons toutefois nous abstenir de procéder à un tel changement de variable.

2° Analyse de la variance

Le tableau 9.4.2 donne les résultats de l'analyse de la variance non orthogonale, la somme des carrés des écarts relative aux traitements étant ajustée en fonction des blocs, comme nous l'avons signalé au paragraphe 9.2.1.1°.

Ce tableau montre que les différences entre traitements ne sont pas significatives, et aussi que les différences entre blocs, c'est-à-dire entre feuilles, sont très importantes, ce qui apparaissait déjà à l'examen du tableau 9.4.1.

L'analyse de la variance à un critère de classification qui serait réalisée en tenant compte uniquement du facteur traitements est caractérisée par un carré moyen résiduel égal à 299,53, de telle sorte que l'efficacité relative du dispositif

Tableau 9.4.2. Étude de la mosaïque du tabac : tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Traitements (ajust.).	4	117,40	29,35	0,67	0,64
Blocs	9	5.203,80	578,20		
Variation résiduelle	6	263,60	43,93		
Totaux	19	5.584,80			

en blocs incomplets, par rapport à une expérience complètement aléatoire, est en première approximation :

$$299,53/43,93 = 6,82 \text{ ou } 682 \text{ \%}.$$

Pour atteindre la même précision que l'expérience en blocs incomplets basée sur 20 observations, soit quatre répétitions, il aurait donc été nécessaire de traiter et d'examiner environ 140 feuilles, soit 25 à 30 répétitions, dans le cadre d'une expérience complètement aléatoire.

3° Moyennes ajustées

Bien qu'on n'observe pas de différences significatives, et qu'aucune comparaison particulière de moyennes ne s'impose donc, nous présentons les valeurs des moyennes ajustées relatives aux différents traitements et leur procédure de calcul dans le tableau 9.4.3.

Tableau 9.4.3. Étude de la mosaïque du tabac : calcul des moyennes ajustées.

i	\bar{x}_i	\bar{x}'_i	\hat{a}_i	\hat{m}_i
1	37,75	37,125	1,0	30,4
2	27,75	27,250	0,8	30,2
3	37,00	34,500	4,0	33,4
4	26,75	27,000	-0,4	29,0
5	17,75	21,125	-5,4	24,0

Les valeurs \bar{x}_i sont les moyennes classiques des ensembles de quatre observations (12, 49, 68 et 22 pour \bar{x}_1 , par exemple), tandis que les quantités \bar{x}'_i sont des moyennes de huit observations (12, 12, 34, 49, 69, 68, 22 et 31 pour \bar{x}'_1). Quant aux effets principaux estimés \hat{a}_i et aux moyennes ajustées \hat{m}_i , leurs valeurs sont obtenues en appliquant les relations du paragraphe 9.2.1.2°.

On peut constater que certaines des moyennes ajustées sont très différentes des moyennes observées correspondantes. On peut vérifier aussi que la somme des

carrés des écarts ajustée relative aux traitements est liée de la manière suivante aux effets principaux estimés :

$$SCE'_a = 2,5 (1,0^2 + 0,8^2 + 4,0^2 + 0,4^2 + 5,4^2) = 117,40.$$

Enfin, l'erreur standard des différences de moyennes est :

$$\sqrt{0,8 (43,93)} = 5,93.$$

4° Codification du modèle d'analyse de la variance

En vue d'obtenir le tableau 9.4.2, le modèle d'analyse de la variance peut être présenté de la manière suivante :

$$N = B \ T \quad \text{et} \quad N \sim B + T,$$

respectivement pour les logiciels Minitab et SAS d'une part (commande ou procédure « *glm* »), et R d'autre part (fonction « *aov* »), N, B et T désignant les nombres de lésions et les facteurs blocs (feuilles) et traitements.

On notera que l'ordre « B suivi de T » doit être respecté, pour que les résultats relatifs au facteur traitements soient ajustés en fonction de l'existence du facteur blocs.

9.5 Exemple 2 : expérience en réseau carré équilibré

⊖ 9.5.1 Présentation et données

1° Présentation générale

Une expérience destinée à comparer différentes variétés de betterave sucrière, réalisée dans le cadre de l'Institut belge pour l'Amélioration de la Betterave, à Tirlemont, constitue une deuxième illustration des dispositifs en blocs incomplets, au sens large. Les variétés étudiées sont au nombre de 13, y compris une variété témoin.

Les parcelles ont une surface utile de 6,75 m², abstraction faite des bordures, et comportent normalement 50 à 60 betteraves (trois lignes de 5 m de longueur, espacées de 45 cm).

Les observations réalisées concernent le nombre de betteraves par parcelle, le rendement en betteraves par parcelle et la teneur en sucre (ou richesse saccharine) des betteraves, cette dernière caractéristique étant déterminée à partir d'échantillons de betteraves prélevés dans les productions des différentes parcelles. Les trois variables ainsi observées permettent de calculer d'autres données, telles que le poids moyen des betteraves, le rendement en betteraves par hectare, le rendement en sucre par parcelle et par hectare, et la valeur financière de la récolte.

2° Dispositif expérimental

En vue d'augmenter au maximum la précision des comparaisons entre les nouvelles variétés et la variété témoin, il a été décidé de multiplier par trois ou par quatre le nombre de répétitions du témoin (§ 2.2.2.2°). D'autre part, afin d'assurer le meilleur contrôle possible de l'hétérogénéité du champ d'expérience, un dispositif en réseau carré équilibré a été adopté.

En considérant les 12 nouvelles variétés et quatre fois la variété témoin, soit au total 16 « objets », et avec cinq répétitions, un réseau carré équilibré permet de disposer à la fois de cinq répétitions de chacune des 12 nouvelles variétés et de 20 répétitions de la variété témoin.

La figure 9.1.3, qui nous a servi à introduire la notion d'expérience en réseau carré équilibré (§ 9.1.3), présente en fait le dispositif qui a été choisi pour l'expérience dont il est question ici, sans respecter toutefois la forme exacte des parcelles et sans faire état des sentiers et des chemins d'accès.

Les numéros d'ordre des « objets » ont été tirés au hasard, la variété témoin se voyant affecter les numéros 3, 5, 9 et 16, et les autres numéros étant attribués aux 12 nouvelles variétés.

3° Données

Nous considérons à titre d'exemple les teneurs en sucre, exprimées en pourcentages. Les valeurs observées pour cette variable figurent dans le tableau 9.5.1.

⊖ 9.5.2 Analyse des résultats

1° Examen préliminaire

La figure 9.5.1 donne une représentation graphique des teneurs en sucre relatives aux différentes « variétés », celles-ci étant rangées dans l'ordre croissant de leurs moyennes (tableau 9.5.1). On remarquera que les numéros 3, 5, 9 et 16 ont été remplacés par la lettre T, en vue de rappeler qu'il s'agit de quatre répétitions d'un même témoin.

Aucune tendance très nette n'apparaît dans la figure en question, les moyennes par « variété » étant peu différentes les unes des autres, avec des valeurs extrêmes égales à 16,28 et 17,30 %, respectivement pour un des témoins (n° 16) et pour la variété n° 8. On peut toutefois observer que les quatre témoins se situent dans la partie inférieure du graphique, leur moyenne générale étant égale à 16,495 %.

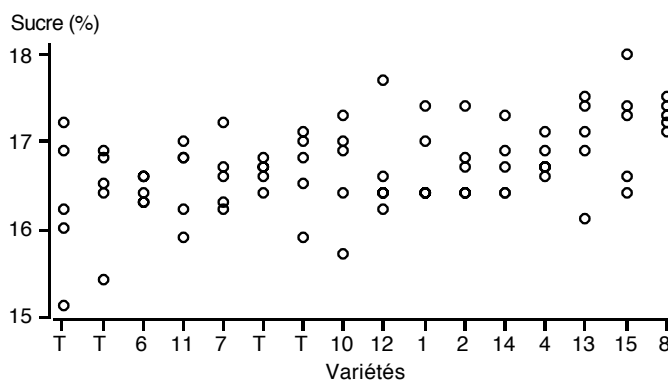
En outre, la figure 9.5.1 ne met en évidence aucune anomalie particulière.

2° Analyse de la variance

Le tableau 9.5.2 présente les résultats de l'analyse de la variance, réalisée sans transformation de variable, les valeurs observées étant des pourcentages allant de

Tableau 9.5.1. Comparaison de variétés de betterave sucrière : teneurs en sucre, moyennes observées et différentes valeurs dérivées, en pourcentages.

Var.	Répétitions					Moy. obs.	Moyennes		Effets princ.	Moy. ajust.
	1	2	3	4	5		\bar{x}'_i	\bar{x}''_i		
1	17,4	16,4	16,4	17,0	16,4	16,72	16,645	16,790	-0,012	16,695
2	16,7	16,4	17,4	16,8	16,4	16,74	16,730	16,795	-0,129	16,578
3	17,1	16,8	15,9	17,0	16,5	16,66	16,660	16,615	0,154	16,862
4	16,7	16,7	16,6	16,9	17,1	16,80	16,765	16,740	0,004	16,712
5	16,5	16,9	15,4	16,4	16,8	16,40	16,450	16,640	0,029	16,737
6	16,6	16,6	16,3	16,3	16,4	16,44	16,850	16,505	-0,346	16,362
7	17,2	16,7	16,3	16,6	16,2	16,60	16,690	16,720	-0,171	16,537
8	17,1	17,2	17,5	17,4	17,3	17,30	16,870	16,875	0,438	17,145
9	16,7	16,7	16,4	16,6	16,8	16,64	16,695	16,765	-0,187	16,520
10	17,3	16,4	15,7	16,9	17,0	16,66	16,590	16,765	0,021	16,728
11	16,2	16,8	15,9	17,0	16,8	16,54	16,915	16,570	-0,396	16,312
12	16,4	17,7	16,2	16,6	16,4	16,66	16,720	16,720	-0,121	16,587
13	16,1	17,4	17,5	17,1	16,9	17,00	16,855	16,705	0,246	16,953
14	17,3	16,9	16,4	16,7	16,4	16,74	16,715	16,695	0,062	16,770
15	18,0	17,3	17,4	16,4	16,6	17,14	16,740	16,835	0,454	17,162
16	17,2	16,0	15,1	16,9	16,2	16,28	16,430	16,585	-0,046	16,662

**Figure 9.5.1.** Comparaison de variétés de betterave sucrière : représentation graphique des observations relatives aux différentes variétés²⁵.

15,1 à 18,0 %. Cette analyse montre que les différences entre les variétés ne sont pas significatives.

²⁵ Certains points correspondent en fait à deux ou trois valeurs identiques, de telle sorte que certaines variétés ne sont représentées que par quatre ou même trois points seulement.

Tableau 9.5.2. Comparaison de variétés de betterave sucrière :
tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Variétés (ajust.)	15	2,584	0,1723	1,22	0,31
Répétitions	4	2,482	0,6205		
Lignes (dans rép.)	15	7,009	0,4673		
Colonnes (dans rép.)	15	3,874	0,2583		
Variation résiduelle	30	4,227	0,1409		
Totaux	79	20,176			

Le tableau 9.5.2 indique aussi que le facteur répétitions joue un rôle relativement prépondérant, et que les différences entre les lignes et entre les colonnes (dans les répétitions) ne sont pas négligeables non plus.

Plus concrètement, l'analyse de la variance à deux critères de classification qui ne ferait intervenir que les facteurs variétés et répétitions, à l'exclusion des lignes et des colonnes, est caractérisée par un carré moyen résiduel égal à 0,2087, de telle sorte que l'efficacité relative du dispositif en réseau carré équilibré, par rapport à une expérience en blocs aléatoires complets, est en première approximation :

$$0,2087/0,1409 = 1,48 \text{ ou } 148 \text{ \%}.$$

De même, on peut montrer que l'efficacité relative de ce dernier dispositif par rapport à une expérience complètement aléatoire est :

$$0,2344/0,2087 = 1,12 \text{ ou } 112 \text{ \%}.$$

L'introduction des facteurs lignes et colonnes s'avère donc particulièrement justifiée dans le cas envisagé.

3° Moyennes ajustées

Comme au paragraphe 9.4.2.3°, bien que les différences observées ne soient pas significatives, nous effectuons à titre indicatif le calcul des moyennes ajustées.

Dans cette optique, les dernières colonnes du tableau 9.5.1 contiennent successivement les moyennes observées $\bar{x}_{i..}$, les moyennes $\bar{x}'_{i.}$ des différentes lignes dans lesquelles se trouve chacune des variétés, les moyennes $\bar{x}''_{i.}$ des différentes colonnes dans lesquelles se trouve chacune des variétés, les effets principaux estimés \hat{a}_i et les moyennes ajustées \hat{m}_i . Ces derniers résultats peuvent être obtenus à l'aide des relations :

$$\hat{a}_i = (5/3) (\bar{x}_{i.} - \bar{x}'_{i.} - \bar{x}''_{i.} + \bar{x}_{..}) \quad \text{et} \quad \hat{m}_i = \bar{x}_{..} + \hat{a}_i,$$

le facteur 5/3 correspondant à $(k+1)/(k-1)$ et le symbole $\bar{x}_{..}$ désignant la moyenne générale de toutes les observations, égale à 16,7075 %.

On peut vérifier en outre que la quantité :

$$SCE'_a = (k - 1) \sum_{i=1}^p \hat{a}_i^2,$$

correspond bien à la somme des carrés des écarts ajustée relative aux variétés.

Enfin, si les différences entre variétés avaient été significatives, on aurait pu comparer les 12 nouvelles variétés à la variété témoin à l'aide du test de DUNNETT [STAT2, § 12.3.2].

À cet égard, on peut noter que l'erreur standard des différences entre la moyenne de l'ensemble des résultats ajustés relatifs à la variété témoin, égale à 16,495 %, et les moyennes ajustées relatives aux autres variétés est :

$$\sqrt{\frac{5}{3} \left(\frac{1}{5} + \frac{1}{20} \right)} 0,1409 = 0,242 \text{ \%}.$$

De même, l'erreur standard des différences entre les moyennes ajustées relatives à d'autres variétés que la variété témoin est :

$$\sqrt{\frac{5}{3} \frac{2}{5}} 0,1409 = 0,306 \text{ \%}.$$

4° Codification du modèle d'analyse de la variance

Le contenu du tableau 9.5.2 peut être obtenu à l'aide des expressions :

$$S = R \ L(R) \ C(R) \ V \quad \text{et} \quad S = R \ L|R \ C|R \ V,$$

respectivement pour les logiciels Minitab et SAS (commande ou procédure « *glm* »), S désignant les teneurs en sucre et R, L, C et V les facteurs répétitions, lignes, colonnes et variétés.

Quant au logiciel R (fonction « *ao*v »), la forme suivante peut être adoptée :

$$S \sim R + L + C + V,$$

à condition de numéroter les lignes et les colonnes (L et C) de 1 à 20, globalement pour l'ensemble des cinq répétitions.

Comme dans les autres cas d'analyse de la variance non orthogonale (§ 9.4.2.4° notamment), l'ordre de présentation des variables doit être respecté.

Chapitre 10

Les expériences en blocs aléatoires incomplets : expériences factorielles

Sommaire

- 10.1 Principes**
 - 10.1.1 Généralités
 - ⊖ 10.1.2 La confusion complète
 - ⊖ 10.1.3 La confusion partielle
 - ⊖ 10.1.4 Les expériences factorielles fractionnaires et à répétition unique
 - ⊖ 10.1.5 Quelques principes de construction
- ⊖ **10.2 Analyse des résultats**
- ⊖ **10.3 Discussion**
- 10.4 Exemple : expérience avec confusion complète**
 - ⊖ 10.4.1 Présentation et données
 - ⊖ 10.4.2 Analyse des résultats

10.1 Principes

10.1.1 Généralités

Les problèmes que pose la prise en considération d'objets en nombre plus élevé que le nombre d'unités expérimentales par bloc se présentent également pour les expériences factorielles, complètes ou fractionnaires. Les nombres d'objets augmentent en effet très rapidement avec le nombre de facteurs étudiés, dans le cas de ces expériences (§ 2.3.1.3°).

Le principe de la *confusion d'effets*¹ permet de remédier à ces difficultés. Ce principe conduit à constituer des blocs incomplets, parfois réunis en répétitions complètes, de telle sorte que les différences entre blocs soient confondues avec certaines interactions, auxquelles on ne s'intéresse pas particulièrement ou qu'on peut négliger lors de l'interprétation des résultats.

On notera que les dispositifs présentés au chapitre 9, en faisant allusion uniquement aux expériences non factorielles, peuvent en réalité être utilisés également pour les expériences factorielles. Dans ce dernier cas toutefois, le recours à la confusion d'effets, présentée ici, est en général plus adéquat.

Enfin, la notion de confusion d'effets permet aussi d'explicitier certains concepts relatifs à la structure des objets, dans le cas des expériences factorielles fractionnaires (§ 2.3.3).

⊖ 10.1.2 La confusion complète

1° Exemple

La figure 10.1.1 indique ce que peut être, par confusion d'effets, la disposition des huit objets d'une expérience 2^3 organisée en blocs de quatre unités expérimentales, avec trois répétitions. Chaque ligne de cette figure constitue un bloc de quatre unités, et chaque groupe de deux lignes constitue une répétition complète des huit objets.

On peut remarquer qu'indépendamment de la répartition aléatoire des blocs au sein des répétitions et des objets au sein des blocs, les huit combinaisons des lettres a, b et c sont toujours groupées de la même manière : (1), ab, ac et bc d'une part, a, b, c et abc d'autre part. On peut démontrer que l'interaction des trois facteurs est ainsi confondue avec les éventuelles différences aléatoires qui peuvent exister entre les blocs, dans les répétitions, du fait de l'hétérogénéité du matériel expérimental.

On notera aussi l'analogie qui existe entre ce dispositif et le deuxième exemple de la figure 9.1.1, relatif à une expérience non factorielle qui comportait également huit objets réunis en blocs de quatre unités expérimentales.

¹ En anglais : *confounding*.

ac	bc	ab	(1)
b	a	abc	c
a	abc	c	b
bc	(1)	ac	ab
c	a	b	abc
bc	ab	(1)	ac

Figure 10.1.1. Exemple de confusion complète.

2° Démonstration

La démonstration du fait que le dispositif présenté ci-dessus « confond » les différences entre blocs avec l'interaction des trois facteurs peut être réalisée assez simplement en se basant, comme au paragraphe 2.3.3.3°, sur le modèle mathématique de l'analyse de la variance à trois critères de classification [STAT2, § 11.2.3] :

$$X_{ijkl} = m_{...} + a_i + b_j + c_k + (ab)_{ij} + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + D_{ijkl}.$$

Dans ce modèle, le terme d'interaction des trois facteurs est défini de la manière suivante, en fonction des différentes moyennes théoriques :

$$(abc)_{ijk} = m_{ijk} - m_{ij.} - m_{i.k} - m_{.jk} + m_{i..} + m_{.j.} + m_{..k} - m_{...}.$$

Pour des facteurs possédant chacun deux modalités seulement, ce terme est, en valeur absolue, indépendant des valeurs des indices i , j et k , et est donc toujours égal à $\pm (abc)_{111}$. De plus, ce dernier terme est alors :

$$\begin{aligned} (abc)_{111} &= m_{111} - (m_{111} + m_{112})/2 - (m_{111} + m_{121})/2 - (m_{111} + m_{211})/2 \\ &\quad + (m_{111} + m_{112} + m_{121} + m_{122})/4 + (m_{111} + m_{112} + m_{211} + m_{212})/4 \\ &\quad + (m_{111} + m_{121} + m_{211} + m_{221})/4 \\ &\quad - (m_{111} + m_{112} + m_{121} + m_{122} + m_{211} + m_{212} + m_{221} + m_{222})/8, \end{aligned}$$

ou encore, après réduction au même dénominateur et simplification :

$$(abc)_{111} = [(m_{111} + m_{122} + m_{212} + m_{221}) - (m_{112} + m_{121} + m_{211} + m_{222})]/8.$$

En tenant compte de l'équivalence entre les différentes notations relatives aux expériences factorielles (§ 2.3.2), cette relation peut aussi s'écrire :

$$(abc)_{111} = [(m_{(1)} + m_{bc} + m_{ac} + m_{ab}) - (m_c + m_b + m_a + m_{abc})]/8.$$

Cette formulation met bien en évidence la distinction entre les deux groupes d'objets (1), ab, ac et bc d'une part, a, b, c et abc d'autre part. Les différences

de moyennes qui existeraient éventuellement entre ces deux groupes d'objets sont donc effectivement intégrées aux termes d'interaction des trois facteurs.

3° Randomisation

Des listes de dispositifs relatifs aux expériences factorielles et basés sur le principe de la confusion d'effets figurent dans différents ouvrages classiques, tels que celui de COCHRAN et COX [1957]. Et de tels dispositifs peuvent être générés aussi à l'aide de certains des logiciels que nous avons mentionnés dans l'introduction générale.

Dans tous les cas, la répartition au hasard doit se faire par une permutation aléatoire éventuelle des répétitions et une permutation aléatoire des blocs au sein des répétitions et des objets au sein des blocs.

⊖ 10.1.3 La confusion partielle

1° Confusion complète et partielle

Dans l'exemple qui vient d'être présenté de manière détaillée, la même interaction (interaction des trois facteurs) est confondue dans toutes les répétitions avec les différences entre blocs. Une telle confusion est dite *complète*².

Dans d'autres cas, plusieurs composantes (interactions ou même facteurs principaux) sont confondues avec les différences entre blocs, ces composantes variant d'une répétition à l'autre. La confusion est alors dite *partielle*³.

2° Cas 2²

Pour une expérience 2² en blocs de deux unités expérimentales, on peut adopter par exemple le dispositif de la figure 10.1.2, qui comprend six répétitions complètes.

b (1)	b ab	ab (1)	a (1)	(1) ab	a ab
a ab	(1) a	a b	ab b	a b	(1) b

Figure 10.1.2. Exemple de confusion partielle.

Dans la première et la dernière répétition, les objets (1) et b forment un premier bloc et les objets a et ab un deuxième bloc : les différences entre blocs sont ainsi confondues avec l'influence du premier facteur (A). De même, dans la deuxième et la quatrième répétition, les objets (1) et a sont réunis dans un bloc et les objets b et ab dans l'autre, les différences entre blocs étant confondues avec l'influence du deuxième facteur (B). Enfin, dans la troisième et la cinquième répétition, on

² En anglais : *complete confounding*.

³ En anglais : *partial confounding*.

retrouve d'une part les objets (1) et ab et d'autre part les objets a et b : les différences entre blocs sont alors confondues avec l'interaction des deux facteurs.

Au total, les trois sources de variation (deux facteurs principaux et interaction des deux facteurs) sont donc confondues un même nombre de fois avec les différences entre blocs.

3° Cas 3^3

La figure 10.1.3 présente aussi le schéma d'une expérience 3^3 , comportant 27 objets et organisée en blocs de neuf unités expérimentales, avec deux répétitions. Ce cas est dans une certaine mesure intermédiaire entre la confusion complète et la confusion partielle.

321	313	233	222	131	112	211	332	123
331	111	323	232	122	221	213	133	312
311	132	121	113	333	322	212	231	223
132	321	312	213	123	222	231	111	333
311	212	131	233	122	323	332	113	221
223	112	232	331	121	133	211	313	322

Figure 10.1.3. Exemple de confusion d'effets dans le cas 3^3 .

L'interaction des trois facteurs, avec huit degrés de liberté, peut en effet être décomposée en huit contrastes orthogonaux. Et on peut démontrer que deux de ces huit contrastes sont confondus avec les différences entre blocs dans la première répétition (trois premières lignes de la figure 10.1.3), tandis que deux autres composantes de cette interaction sont confondues avec les différences entre blocs dans la deuxième répétition (trois dernières lignes de la figure 10.1.3).

Des éléments distincts sont donc confondus avec les différences entre blocs dans les différentes répétitions, mais ces éléments appartiennent tous à la même interaction.

4° Confusion d'effets et expériences en parcelles divisées

Il est intéressant de signaler que les expériences en blocs aléatoires complets avec parcelles divisées (*split-plot*) peuvent être considérées aussi, dans une certaine mesure, comme des cas particuliers de confusion complète.

Dans le cas le plus courant de deux facteurs, si on admet que les grandes parcelles du paragraphe 7.1.1.1° correspondent aux blocs dont il est question ici, l'élément confondu avec le facteur blocs est le facteur principal associé à ces grandes parcelles, et cette confusion d'effets intervient de la même manière dans les différentes répétitions. Il s'agit donc bien d'une confusion complète, relative à un des deux facteurs principaux.

⊖ 10.1.4 Les expériences factorielles fractionnaires et à répétition unique

1° Confusion d'effets et expériences factorielles fractionnaires

Les expériences factorielles fractionnaires, auxquelles est consacré le paragraphe 2.3.3, sont étroitement liées à la notion de confusion d'effets. Les sous-ensembles d'objets qui sont pris en considération dans ces expériences correspondent en fait aux blocs incomplets qui sont définis ici.

Ainsi, les sous-ensembles (1), ab, ac et bc, et a, b, c et abc de la figure 10.1.1 sont identiquement ceux qui ont été envisagés au paragraphe 2.3.3.2°.

Mais les expériences factorielles fractionnaires peuvent aussi être organisées elles-mêmes en blocs incomplets. Le principe de la confusion d'effets intervient alors à la fois dans la définition du sous-ensemble d'objets qui est considéré, et dans la constitution des blocs.

2° Exemple

Pour une demi-répétition d'une expérience 2^6 par exemple, c'est-à-dire pour une expérience de type 2^{6-1} , on peut répartir les 32 objets étudiés en deux blocs de 16 unités expérimentales, sans perdre aucune information en ce qui concerne les six facteurs principaux et les 15 interactions de deux facteurs. La répartition peut être notamment celle qui est présentée à la figure 10.1.4.

bc	acde	de	ab	bcd	(1)	acdf	abde
bcef	ef	acef	abef	abdf	bcde	df	ac
abcdef	ae	abcd	ad	af	abce	cf	bf
be	ce	cdef	adef	cd	bdef	bd	abcf

Figure 10.1.4. Exemple d'expérience factorielle fractionnaire en blocs incomplets avec confusion d'effets.

Une répartition en blocs de quatre ou de huit unités expérimentales pourrait aussi être envisagée, mais on peut démontrer que certaines interactions de deux facteurs doivent alors être nécessairement confondues avec les différences entre blocs.

3° Répétition unique

Le principe de la confusion d'effets peut également être appliqué aux expériences factorielles qui ne comportent qu'une seule répétition (§ 2.3.2.8°). La figure 10.1.5 illustre le cas de cinq facteurs qui seraient étudiés en quatre blocs de huit unités expérimentales, sans perte d'information au sujet des facteurs principaux et des interactions de deux facteurs.

bc	abd	abe	(1)	acd	de	bcde	ace
acde	bce	ab	ac	bcd	e	abde	d
bde	ae	cde	abce	ad	b	c	abcd
be	ce	abcde	a	abc	cd	bd	ade

Figure 10.1.5. Exemple d'expérience factorielle à répétition unique en blocs incomplets avec confusion d'effets.

Un autre exemple est fourni par la première moitié ou la deuxième moitié de la figure 10.1.3, si on envisage d'étudier trois facteurs présentant chacun trois modalités en une seule répétition, constituée de trois blocs de neuf unités expérimentales.

⊖ 10.1.5 Quelques principes de construction

1° Expériences 2^k en deux blocs

Les différents exemples qui ont été présentés ci-dessus permettent d'introduire de façon intuitive quelques principes relatifs aux modalités de construction des dispositifs en blocs incomplets avec confusion d'effets, dans le cas des expériences factorielles 2^k .

Quand une ou plusieurs répétitions complètes doivent être divisées chacune en deux blocs, on confond généralement avec les différences entre blocs l'interaction qui fait intervenir tous les facteurs. Tel est le cas notamment pour l'exemple de la figure 10.1.1.

Tous les objets définis par un nombre impair de lettres (a, b, ..., abc, etc.) sont alors réunis dans un bloc, tandis que tous les objets définis par un nombre pair de lettres (ab, ac, etc.), y compris l'objet (1), constituent l'autre bloc. Nous avons vu qu'il en est ainsi également dans la constitution de demi-répétitions d'expériences factorielles 2^k (§ 2.3.3).

2° Expériences 2^k en quatre blocs

Quand une ou plusieurs répétitions complètes doivent être divisées chacune en quatre blocs, on doit choisir deux interactions à confondre avec les différences entre blocs. Ce choix implique nécessairement le choix d'une troisième interaction qui est également confondue avec les différences entre blocs.

Ainsi, dans l'exemple de la figure 10.1.5, les interactions qui ont été choisies pour constituer les blocs sont ABC et ADE, et l'interaction BCDE en découle. Cette dernière peut être obtenue par une multiplication symbolique des deux premières, en considérant que A^2 (ou, d'une façon plus générale, A^2, B^2, C^2 , etc.) est égal à 1 :

$$(ABC)(ADE) = A^2BCDE = BCDE.$$

Les trois interactions en question constituent les *contrastes* ou *mots de définition*⁴ des blocs, auxquels correspondent les trois degrés de liberté qui sont associés aux différences existant entre les quatre blocs.

On notera qu'il est indispensable de partir ici d'interactions de trois facteurs, au risque sinon de voir figurer, parmi les contrastes de définition, une interaction de deux facteurs, telle que par exemple AE en partant de ABCD et BCDE, puisque :

$$(ABCD)(BCDE) = AB^2C^2D^2E = AE.$$

3° Constitution des blocs

Les contrastes de définition des blocs étant choisis, on peut procéder à la constitution des blocs eux-mêmes selon les principes suivants, illustrés encore à l'aide de l'exemple de la figure 10.1.5.

Si seule l'interaction ABC devait être confondue avec les différences entre blocs, le premier bloc devrait contenir les objets :

(1), ab, ac, bc, d, abd, acd, bcd, e, abe, ace, bce, de, abde, acde, bcde,

et le deuxième bloc, les objets :

a, b, c, abc, ad, bd, cd, abcd, ae, be, ce, abce, ade, bde, cde, abcde.

Ce résultat peut être obtenu en partant des objets qui constituent les deux blocs associés à l'interaction ABC dans le cas 2^3 (figure 10.1.1), et en ajoutant à ces objets ceux dont le symbole contient en outre les deux lettres restantes, soit séparément (d ou e), soit conjointement (d et e).

De la même façon, on peut établir que, si seule l'interaction ADE devait être confondue avec les différences entre blocs, les deux blocs seraient, d'une part :

(1), ad, ae, de, b, abd, abe, bde, c, acd, ace, cde, bc, abcd, abce, bcde,

et d'autre part :

a, d, e, ade, ab, bd, be, abde, ac, cd, ce, acde, abc, bcd, bce, abcde.

Pour que les deux interactions soient simultanément confondues avec les différences entre blocs, dans une répétition alors divisée en quatre blocs, il faut par exemple que le premier des quatre blocs réunisse les objets communs au premier bloc relatif à ABC et au premier bloc relatif à ADE :

(1), bc, de, abd, abe, acd, ace, bcde;

⁴ En anglais : *defining contrast*.

que le deuxième bloc réunisse les objets communs au premier bloc relatif à ABC et au deuxième bloc relatif à ADE :

$$d, e, ab, ac, bcd, bce, abde, acde;$$

que le troisième bloc réunisse les objets communs au deuxième bloc relatif à ABC et au premier bloc relatif à ADE :

$$b, c, ad, ae, bde, cde, abcd, abce;$$

et que le quatrième bloc réunisse les objets communs au deuxième bloc relatif à ABC et au deuxième bloc relatif à ADE :

$$a, bd, be, cd, ce, abc, ade, abcde.$$

Tel est bien le contenu de la figure 10.1.5.

4° Relations entre les blocs

On peut remarquer que les quatre blocs ainsi définis sont aussi liés les uns aux autres par des relations de multiplication. Pour établir de telles relations, il faut adopter la convention :

$$a^2 = b^2 = c^2 = d^2 = e^2 = 1,$$

qui est analogue à celle utilisée ci-dessus pour les lettres majuscules.

Les objets du deuxième bloc peuvent alors être obtenus par une multiplication symbolique des objets du premier bloc par d ou par e :

$$(bc)d = bcd, (abd)d = ab, \text{ etc.} \quad \text{et} \quad (bc)e = bce, (abd)e = abde, \text{ etc.},$$

et vice-versa :

$$(acde)d = ace, (bce)d = bcde, \text{ etc.} \quad \text{et} \quad (acde)e = acd, (bce)e = bc, \text{ etc.}$$

De même, le premier et le troisième bloc sont liés par une multiplication symbolique par b ou par c :

$$(bc)b = c, (abd)b = ad, \text{ etc.} \quad \text{et} \quad (bc)c = b, (abd)c = abcd, \text{ etc.},$$

$$(bde)b = de, (ae)b = abe, \text{ etc.} \quad \text{et} \quad (bde)c = bcde, (ae)c = ace, \text{ etc.},$$

le premier et le quatrième bloc par une multiplication par a, le deuxième et le troisième bloc par une multiplication par a également, le deuxième et le quatrième bloc par une multiplication par b ou par c, et enfin, le troisième et le quatrième bloc par une multiplication par d ou par e.

Ces propriétés peuvent simplifier la construction des blocs et permettent d'effectuer dans certains cas d'utiles vérifications.

5° Compléments

Les principes qui viennent d'être présentés s'appliquent également au cas des expériences factorielles fractionnaires, tant en ce qui concerne le choix des sous-ensembles d'objets que la constitution des blocs.

Dans le cas 2^6 qui est envisagé à la figure 10.1.4 par exemple, les contrastes de définition qui sont utilisés sont les interactions ABC, DEF et ABCDEF, et seuls deux des quatre blocs qui peuvent être définis à l'aide de ces contrastes ont été retenus. De plus, dans les deux sens, le passage d'un bloc à l'autre peut être réalisé par une multiplication symbolique des objets par ad.

Tous ces principes peuvent être facilement étendus à n'importe quelle expérience factorielle du type 2^k , et peuvent être adaptés à d'autres types d'expériences, telle que la série 3^k . On trouvera des informations complémentaires à ce sujet notamment dans les livres de COCHRAN et COX [1957], FINNEY [1960] et KUEHL [2000], et dans les articles de COLLINGS [1989], FRANKLIN [1985], et GHOSH et BAGUI [1998].

On peut mentionner aussi d'autres publications, telles que celles de CHEN *et al.* [2006], CHENG et WU [2002], KERR [2006], et XU et LAU [2006]. Ces travaux associent parfois, aux blocs incomplets, certains critères d'optimalité (§ 2.4.3), de même que des notions de parcelles divisées (§ 7.1 et 7.3), en raison du fait que les modalités des différents facteurs pris en considération peuvent être plus ou moins difficiles à modifier (§ 5.3.3°).

⊖ 10.2 Analyse des résultats

1° Analyse de la variance

Une fois encore, l'analyse de la variance est ici très souvent à la base de l'interprétation des résultats des expériences.

Dans le cas de la confusion complète, le calcul des effets principaux et des interactions non confondues avec le facteur blocs doit être réalisé selon les principes habituels. Il en est de même pour les différences entre répétitions et entre blocs dans les répétitions.

Dans le cas de la confusion partielle par contre, des techniques particulières doivent être utilisées pour « récupérer » les éléments d'information qui restent disponibles au sujet des interactions et, éventuellement, des facteurs principaux qui sont confondus avec les différences entre blocs.

D'autres méthodes que l'analyse de la variance peuvent cependant aussi être utilisées. On peut penser notamment à l'étude des diagrammes ou des demi-diagrammes de probabilité, et aux méthodes qui sont parfois associées à ces diagrammes, dont la méthode LENTH (§ 5.2.5°) [LOEPPKY et SITTE, 2002].

2° Exemples

À titre d'illustrations des quelques principes qui viennent d'être exposés, les tableaux 10.2.1, 10.2.2 et 10.2.3 donnent les schémas des analyses de la variance qui correspondent aux différentes situations envisagées aux paragraphes 10.1.2 et 10.1.3, et cela tant d'une manière générale que pour les différents exemples considérés.

Tableau 10.2.1. Schéma de l'analyse de la variance relative à une expérience 2^3 en blocs de quatre unités expérimentales, avec confusion complète, dans le cas général et dans le cas particulier de trois répétitions.

Sources de variation	Degrés de liberté	
Facteur A	1	1
Facteur B	1	1
Facteur C	1	1
Interaction AB	1	1
Interaction AC	1	1
Interaction BC	1	1
Répétitions	$r - 1$	2
Blocs (dans répétitions) et interaction ABC	r	3
Variation résiduelle	$6(r - 1)$	12
Totaux	$8r - 1$	23

Tableau 10.2.2. Schéma de l'analyse de la variance relative à une expérience 2^2 en blocs de deux unités expérimentales, avec confusion partielle, dans le cas général ($r \geq 3$) et dans le cas particulier de six répétitions.

Sources de variation	Degrés de liberté	
Facteur A	1	1
Facteur B	1	1
Interaction AB	1	1
Répétitions	$r - 1$	5
Blocs (dans répétitions)	r	6
Variation résiduelle	$2r - 3$	9
Totaux	$4r - 1$	23

Aucun problème particulier ne se pose dans le premier cas (tableau 10.2.1). Dans le deuxième cas au contraire (tableau 10.2.2), il ne faut pas perdre de vue le fait que les deux facteurs principaux et leur interaction sont partiellement confondus avec le facteur blocs.

Tableau 10.2.3. Schéma de l'analyse de la variance relative à une expérience 3^3 en blocs de neuf unités expérimentales, avec confusion partielle, dans le cas général ($1 \leq r \leq 4$) et dans le cas particulier de deux répétitions.

Sources de variation	Degrés de liberté	
Facteur A	2	2
Facteur B	2	2
Facteur C	2	2
Interaction AB	4	4
Interaction AC	4	4
Interaction BC	4	4
Interaction ABC	$2(4 - r)$	4
Répétitions	$r - 1$	1
Blocs (dans répétitions) et interaction ABC	$2r$	4
Variation résiduelle	$26(r - 1)$	26
Totaux	$27r - 1$	53

Dans le troisième cas enfin (tableau 10.2.3), l'analyse qui est présentée suppose que les différences entre blocs sont confondues, dans les différentes répétitions, avec des composantes différentes de l'interaction des trois facteurs. La ligne « Interaction ABC » est alors relative aux composantes de l'interaction des trois facteurs qui ne sont pas confondues avec le facteur blocs, tandis que la ligne « Blocs (dans répétitions) et interaction ABC » concerne les composantes de l'interaction qui sont confondues avec le facteur blocs.

Dans ce dernier cas, pour une répétition unique, la ligne « Variation résiduelle » disparaîtrait, et on pourrait envisager d'effectuer les tests relatifs aux trois facteurs et à leurs interactions simples par rapport à la composante « Interaction ABC », avec six degrés de liberté.

⊖ 10.3 Discussion

Des commentaires analogues à ceux qui ont été formulés au paragraphe 9.3 peuvent être émis au sujet de la confusion d'effets, en ce qui concerne tant les avantages et les inconvénients des dispositifs envisagés, que leurs possibilités d'utilisation et le calcul d'efficacités relatives.

On peut ajouter que, comme dans le cas des expériences non factorielles (§ 9.1.3 et 9.1.4), des dispositifs en réseaux qui font intervenir à la fois des lignes et des colonnes ont été proposés pour les expériences factorielles [CHENG et MUKERJEE,

2003; JOHN et LEWIS, 1983; WILLIAMS et JOHN, 1996]. Parmi ceux-ci, figurent notamment des dispositifs appelés *carrés quasi latins*⁵.

10.4 Exemple : expérience avec confusion complète

⊖ 10.4.1 Présentation et données

1° Présentation générale

En ce qui concerne la confusion d'effets, nous considérons comme exemple le cas d'une expérience factorielle 2^4 réalisée sur cotonnier au Congo (RDC) [VAN RENTERGHEM et DAGNELIE, 1963]. Le but de cette expérience est de comparer l'effet de quatre engrais (nitrate d'ammonium, phosphate bicalcique, sulfate de potassium et sulfate de magnésium) sur le rendement en coton-graine. Chacun des quatre engrais est soit non appliqué, soit appliqué à une certaine dose, l'ensemble constituant un schéma factoriel complet comportant 16 objets.

Sans spécifier exactement les doses utilisées, qui figurent dans le travail original cité ci-dessus, nous représentons comme suit les différents objets :

$$(1), n, p, k, m_g, np, nk, nm_g, pk, pm_g, km_g, \\ npk, npm_g, nkm_g, pkm_g, npkm_g,$$

les symboles n, p, k et m_g étant simplement, en minuscules, les symboles chimiques classiques relatifs à l'azote, au phosphore, au potassium et au magnésium.

Les unités expérimentales sont des parcelles d'un are et l'expérience comporte deux répétitions, soit un total de 32 parcelles.

Les observations principales sont, comme il a été dit ci-dessus, le rendement en coton-graine.

2° Dispositif expérimental

Le dispositif expérimental est constitué de répétitions complètes avec, dans chacune des deux répétitions, confusion de l'interaction des quatre facteurs (NPKMg) et du facteur blocs. Ce dispositif permet de réduire la dimension des blocs à huit unités expérimentales, au lieu de 16.

Le schéma adopté est esquissé à la figure 10.4.1, ici également sans en donner tous les détails (bordures, sentiers, etc.).

3° Données

Les rendements observés, exprimés en quintaux par hectare, sont présentés dans le tableau 10.4.1, en fonction des différents éléments fertilisants (codés en 1 et 2), des répétitions et des blocs.

⁵ En anglais : *quasi-Latin square*.

pk _g	k	km _g	np	np	km _g	npm _g	n
n	npk	npkm _g	(1)	nm _g	pk	pkm _g	nkm _g
m _g	p	pm _g	nk	(1)	nk	p	k
npm _g	nkm _g	nm _g	pk	npkm _g	pm _g	mg	npk
Bloc 1		Bloc 2		Bloc 1		Bloc 2	
Répétition 1				Répétition 2			

Figure 10.4.1. Comparaison de fumures sur cotonnier : plan schématique de l'expérience.

Tableau 10.4.1. Comparaison de fumures sur cotonnier : rendements observés, en quintaux par hectare.

N	P	K	Mg	Rép.	Bl.	Rdt	N	P	K	Mg	Rép.	Bl.	Rdt
1	1	1	1	1	2	8,43	2	1	1	1	1	1	9,84
1	1	1	1	2	1	9,95	2	1	1	1	2	2	14,94
1	1	1	2	1	1	6,57	2	1	1	2	1	2	12,46
1	1	1	2	2	2	9,16	2	1	1	2	2	1	13,30
1	1	2	1	1	1	7,93	2	1	2	1	1	2	9,29
1	1	2	1	2	2	11,53	2	1	2	1	2	1	10,24
1	1	2	2	1	2	9,83	2	1	2	2	1	1	10,74
1	1	2	2	2	1	12,64	2	1	2	2	2	2	11,70
1	2	1	1	1	1	7,52	2	2	1	1	1	2	10,35
1	2	1	1	2	2	10,41	2	2	1	1	2	1	11,63
1	2	1	2	1	2	8,22	2	2	1	2	1	1	10,12
1	2	1	2	2	1	7,10	2	2	1	2	2	2	11,06
1	2	2	1	1	2	10,49	2	2	2	1	1	1	9,21
1	2	2	1	2	1	9,52	2	2	2	1	2	2	22,88
1	2	2	2	1	1	7,90	2	2	2	2	1	2	9,69
1	2	2	2	2	2	16,77	2	2	2	2	2	1	14,77

⊖ 10.4.2 Analyse des résultats

1° Examen préliminaire

La représentation graphique des données en fonction des différents facteurs étudiés (figure 10.4.2) met tout d'abord en évidence la présence d'une valeur particulièrement élevée (22,88). Cette représentation graphique suggère également l'existence d'un effet possible de la fumure azotée et de la fumure potassique.

La grande variabilité des observations, allant de 6,57 à 22,88 (avec une moyenne générale égale à 10,82 q/ha), pourrait justifier à elle seule l'application d'une trans-

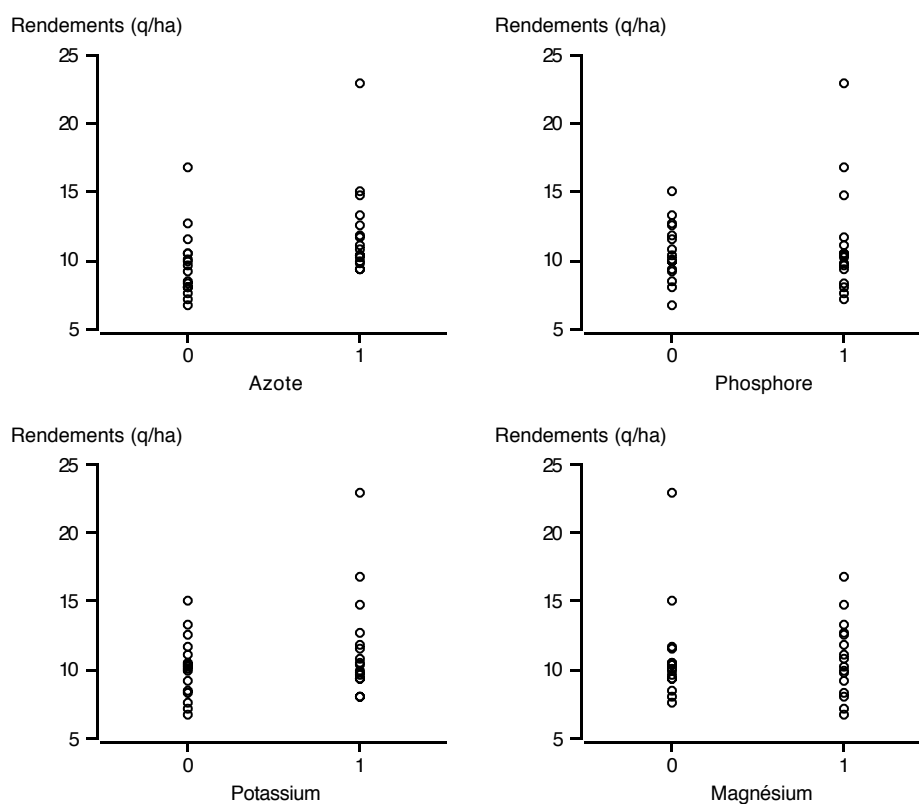


Figure 10.4.2. Comparaison de fumures sur cotonnier : représentation graphique de l'influence des différents facteurs considérés⁶.

formation logarithmique. Mais cette transformation s'impose surtout en raison du fait que l'expérience qui est considérée ici s'insère dans un ensemble plus vaste d'expériences établies en plusieurs endroits et au cours de plusieurs années, qu'on souhaite étudier globalement, alors que les moyennes relatives à ces expériences varient de 1,04 à 12,18 q/ha, avec des écarts-types pratiquement proportionnels aux moyennes. L'étude de l'ensemble des expériences en question est d'ailleurs l'objet de l'exemple du paragraphe 11.4.

2° Analyse de la variance

Les logarithmes décimaux des rendements ont été soumis à une analyse de la variance à six critères de classification, y compris les facteurs répétitions et blocs.

⁶ Les valeurs 0 et 1, qui remplacent les indices 1 et 2 du tableau 10.4.1, désignent respectivement l'absence et la présence de chacun des éléments fertilisants.

Les résultats de cette analyse sont donnés dans le tableau 10.4.2. On peut remarquer que 14 degrés de liberté seulement, au lieu de 15 normalement, sont affectés aux 16 objets et que la décomposition en effets principaux et en interactions ne fait pas apparaître l'interaction NPKMg, qui est confondue avec le facteur blocs.

Tableau 10.4.2. Comparaison de fumures sur cotonnier : tableau d'analyse de la variance.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Facteur N	1	0,073460	0,073460	13,3 **	0,0027
Facteur P	1	0,000877	0,000877	0,16	0,70
Facteur K	1	0,022835	0,022835	4,13	0,062
Facteur Mg	1	0,000056	0,000056	0,01	0,92
Interaction NP	1	0,000464	0,000464	0,08	0,78
Interaction NK	1	0,018688	0,018688	3,38	0,087
Interaction NMg	1	0,000004	0,000004	0,00	0,98
Interaction PK	1	0,018463	0,018463	3,34	0,089
Interaction PMg	1	0,004193	0,004193	0,76	0,40
Interaction KMg	1	0,006615	0,006615	1,20	0,29
Interaction NPK	1	0,009443	0,009443	1,71	0,21
Interaction NPMg	1	0,004482	0,004482	0,81	0,38
Interaction NKMg	1	0,011073	0,011073	2,00	0,18
Interaction PKMg	1	0,001687	0,001687	0,31	0,59
Répétitions	1	0,102048	0,102048		
Blocs (dans rép.)	2	0,035287	0,017643		
Variation résiduelle	14	0,077459	0,005533		
Totaux	31	0,387134			

L'analyse ainsi réalisée montre que seul le facteur azote présente une contribution significative et qu'aucune des interactions n'est significative. L'examen des résidus de cette analyse révèle également que la valeur initiale 22,88 ne doit pas être considérée comme anormale.

D'autre part, le tableau 10.4.2 indique que l'efficacité relative de l'expérience avec confusion d'effets par rapport à une expérience en blocs aléatoires complets classique est, en première approximation :

$$[(0,035387 + 0,077459)/(2 + 14)]/0,005533 = 0,007047/0,005533 = 1,27 \text{ ou } 127 \%,$$

tandis que l'efficacité de ce dernier dispositif par rapport à une expérience complètement aléatoire peut être estimée à :

$$\begin{aligned} [(0,102048 + 0,035387 + 0,077459)/(1 + 2 + 14)]/0,007047 &= 0,012635/0,007047 \\ &= 1,79 \text{ ou } 179 \%. \end{aligned}$$

Comme le montrent aussi les carrés moyens correspondants, la composante répétitions joue donc un rôle beaucoup plus important que la composante « blocs (dans répétitions) ».

Dans la mesure où aucune interaction n'est significative, on peut envisager de regrouper les sommes des carrés des écarts qui les concernent avec la somme des carrés des écarts de la variation résiduelle, de manière à augmenter le nombre de degrés de liberté de cette dernière (tableau 10.4.3). Ce regroupement ne modifie cependant pas les conclusions de l'analyse.

Tableau 10.4.3. Comparaison de fumures sur cotonnier : tableau d'analyse de la variance, après regroupement de différentes composantes.

Sources de variation	Degrés de liberté	Sommes des carrés des écarts	Carrés moyens	F	P
Facteur N	1	0,073460	0,073460	11,6 **	0,0024
Facteur P	1	0,000877	0,000877	0,14	0,71
Facteur K	1	0,022835	0,022835	3,59	0,070
Facteur Mg	1	0,000056	0,000056	0,01	0,93
Répétitions	1	0,102048	0,102048		
Blocs (dans rép.)	2	0,035287	0,017643		
Variation résiduelle	24	0,152571	0,006357		
Totaux	31	0,387134			

3° Interprétation

Comme au paragraphe 7.4.2.3°, on peut chiffrer l'importance du facteur azote par l'intermédiaire des logarithmes décimaux.

Les moyennes des logarithmes des rendements des 16 parcelles sans azote et des 16 parcelles avec azote étant respectivement 0,97124 et 1,06707, la valeur estimée et l'intervalle de confiance relatifs à l'influence de la fumure azotée peuvent être déterminés comme suit :

$$1,06707 - 0,97124 \pm 2,145 \sqrt{2(0,005533)/16} = 0,09583 \pm 0,05641 \\ = 0,03942 \text{ et } 0,15224.$$

Le retour aux valeurs initiales met donc en évidence un accroissement moyen de rendement estimé à 25 %, avec des limites de confiance égales à 10 et 42 %.

Il faut signaler enfin que l'étude des quatre facteurs et de leurs interactions aurait pu être réalisée également en utilisant d'autres méthodes que l'analyse de la variance, comme dans le cas de l'exemple du paragraphe 5.5 (demi-diagramme de probabilité et régression multiple).

4° Codification des modèles d'analyse de la variance

En ce qui concerne le logiciel SAS (procédure « *anova* » ou « *glm* »), le contenu des tableaux 10.4.2 et 10.4.3 peut être obtenu à l'aide des modèles :

$$\begin{aligned} L10R = N \ P \ K \ Mg \ N^*P \ N^*K \ N^*Mg \ P^*K \ P^*Mg \ K^*Mg \\ N^*P^*K \ N^*P^*Mg \ N^*K^*Mg \ P^*K^*Mg \ R \ B|R, \end{aligned}$$

et
$$L10R = N \ P \ K \ Mg \ R \ B|R,$$

si on désigne par L10R les logarithmes décimaux des rendements, par N, P, K et Mg les quatre éléments fertilisants, et par R et B les répétitions et les blocs.

En ce qui concerne Minitab (commande « *glm* »), les mentions « B|R » doivent être remplacées par « B(R) », et le symbole K doit éventuellement être remplacé par exemple par Ka (Kalium), la lettre K ne pouvant pas être utilisée seule comme symbole de variable, au moins pour certaines versions de ce logiciel.

Quant au logiciel R (fonction « *aov* »), on peut utiliser, dans le premier cas (tableau 10.4.2), le modèle :

$$\begin{aligned} L10R \sim N + P + K + Mg + N:P + N:K + N:Mg + P:K + P:Mg + K:Mg \\ + N:P:K + N:P:Mg + N:K:Mg + P:K:Mg + R + B, \end{aligned}$$

ou la forme simplifiée :

$$L10R \sim N^*P^*K^*Mg - N:P:K:Mg + R + B,$$

et dans le deuxième cas (tableau 10.4.3), le modèle :

$$L10R \sim N + P + K + Mg + R + B.$$

Ces trois formulations supposent toutefois que les blocs ont été numérotés de 1 à 4, pour l'ensemble des deux répétitions.