

# EXPÉRIMENTATION AGRONOMIQUE PLANIFIÉE

Document obtenu sur le site Cirad du réseau <http://agroecologie.cirad.fr>

**(SUPPORT DE COURS)**

Philippe LETOURMY

**révision** Eric GOZÉ

**janvier 1999**

Centre de Coopération Internationale  
en Recherche Agronomique pour le Développement (CIRAD)

Document obtenu sur le site Cirad du réseau <http://agroecologie.cirad.fr>

**B.P. 5035 – 34032 Montpellier Cedex 1**

**philippe.letourmy@cirad.fr**

INTRODUCTION .....	3
CHAPITRE 1 .....	4
QUELQUES PRINCIPES DE L'EXPÉRIMENTATION AGRONOMIQUE .....	4
1) EXPÉRIMENTATION ET ENQUÊTE : .....	4
2) DÉMARCHE DE L'EXPÉRIMENTATION : .....	5
3) QUELQUES DÉFINITIONS : .....	7
3.1 Le traitement .....	7
3.2 L'unité expérimentale .....	7
3.3 Le champ d'expérience et le champ d'application .....	8
3.4 L'observation .....	9
4) LES TROIS PRINCIPES DE LA PLANIFICATION EXPÉRIMENTALE : .....	9
CHAPITRE 2 .....	11
DISPOSITIFS EXPÉRIMENTAUX CLASSIQUES .....	11
1) PLANS EN RANDOMISATION TOTALE : .....	11
1.1 Un facteur étudié : .....	11
1.2 Deux facteurs étudiés sans répétition : .....	20
1.3 Deux facteurs étudiés avec répétitions : .....	21
2) LES PLANS EN BLOCS COMPLETS : .....	23
2.1 Répartition des traitements totalement aléatoires dans les blocs .....	25
2.2 Répartition des traitements non totalement aléatoire dans les blocs .....	26
3) LES PLANS EN CARRÉ LATIN : .....	30
1. Carré latin arbitraire .....	31
2. Randomisation des lignes .....	31
3) Randomisation des colonnes .....	31
4) LES AUTRES PLANS D'EXPÉRIENCE : .....	33
CHAPITRE 3 .....	36
LES REGROUPEMENTS D'ESSAIS .....	36
CHAPITRE 4 .....	43
ANALYSE EN COMPOSANTES PRINCIPALES .....	43
BIBLIOGRAPHIE .....	48
1) Expérimentation et analyse de variance. ....	48
2/ Méthodes non-paramétriques .....	49
3) Enquêtes, échantillonnage. ....	49
4) Analyse des données .....	50
5) Logiciel statistique .....	50

## INTRODUCTION

Le présent document a été rédigé pour servir de support à des cours sur l'expérimentation agronomique. Ces cours comportaient des applications sur le logiciel STATITCF. Le document y fait fréquemment référence.

On ne fait pas état ici des exemples d'application qui sont vus en cours. Seuls sont exposés quelques principes généraux de l'expérimentation agronomique et de l'analyse de variance unidimensionnelle, ainsi que la construction, l'analyse, les avantages et les inconvénients des principaux dispositifs. Enfin, le problème des regroupements d'essais est abordé en se focalisant sur l'analyse exploratoire de l'interaction traitement-essai par des techniques d'analyse de données, d'où le chapitre de présentation de l'analyse en composantes principales (ACP) à la fin du document.

L'attention a été portée plus particulièrement sur les dispositifs expérimentaux analysables par STATITCF. Les dispositifs équilibrés non analysables par ce logiciel ne sont que survolés. Ceux non équilibrés ne sont pas abordés car leur analyse utilise des notions sur le modèle linéaire général qui ne sont pas développées ici.

## CHAPITRE 1

# QUELQUES PRINCIPES DE L'EXPÉRIMENTATION AGRONOMIQUE

Les agriculteurs sont confrontés à des choix dans leurs activités : choix de la culture, de la variété, de l'engrais, des techniques culturales, etc. Pour prendre ces décisions, ils peuvent se baser sur leur expérience personnelle, ou sur les traditions locales, ou encore sur des échanges d'informations avec des spécialistes agricoles. Ce sont les sources d'information de ces derniers qui vont être développées ici.

### 1) EXPÉRIMENTATION ET ENQUÊTE :

---

Tout recueil d'informations est conçu en fonction d'un objectif, et cet objectif conditionne la manière dont seront collectées les données et les questions auxquelles on tentera d'apporter une réponse.

On considère deux modes de recueil de l'information : l'expérimentation et l'enquête.

En expérimentation, on cherche à comparer les effets de différents traitements sur une population expérimentale. Les données doivent être "fabriquées" par l'expérimentateur de manière à respecter dans la mesure du possible le principe : "les traitements sont comparés, toutes choses étant égales par ailleurs".

Aussi, la caractéristique essentielle de l'expérimentation est de créer des conditions contrôlées. La population expérimentale est choisie dans ce but. La démarche de l'expérimentation est donc rigoureusement analytique. En effet, un résultat d'essai est relatif à une situation particulière : celle créée par les conditions de l'expérience, contrôlables ou non (par exemple, dans le domaine agronomique, la variété, les techniques culturales, le type de sol, les conditions climatiques, etc.).

Le dispositif expérimental doit assurer que les suppositions nécessaires à l'analyse statistique sont réalisées et permettre une bonne précision des résultats, en spécifiant la manière d'affecter les traitements aux unités expérimentales.

L'autre mode de recueil d'informations est l'enquête (appelée ainsi ou bien sondage lorsque les observations portent sur des personnes et échantillonnage lorsque les observations portent sur des animaux ou des objets).

Dans une enquête, on cherche à collecter des informations sur une population, en la représentant par une partie de celle-ci, l'échantillon. La notion de représentativité de la population par l'échantillon est donc essentielle dans une enquête.

On dit qu'un échantillon est représentatif si tout individu de la population peut figurer dans l'échantillon avec une probabilité non nulle et connue. De plus, on appelle base de sondage, la liste exhaustive des individus de la population.

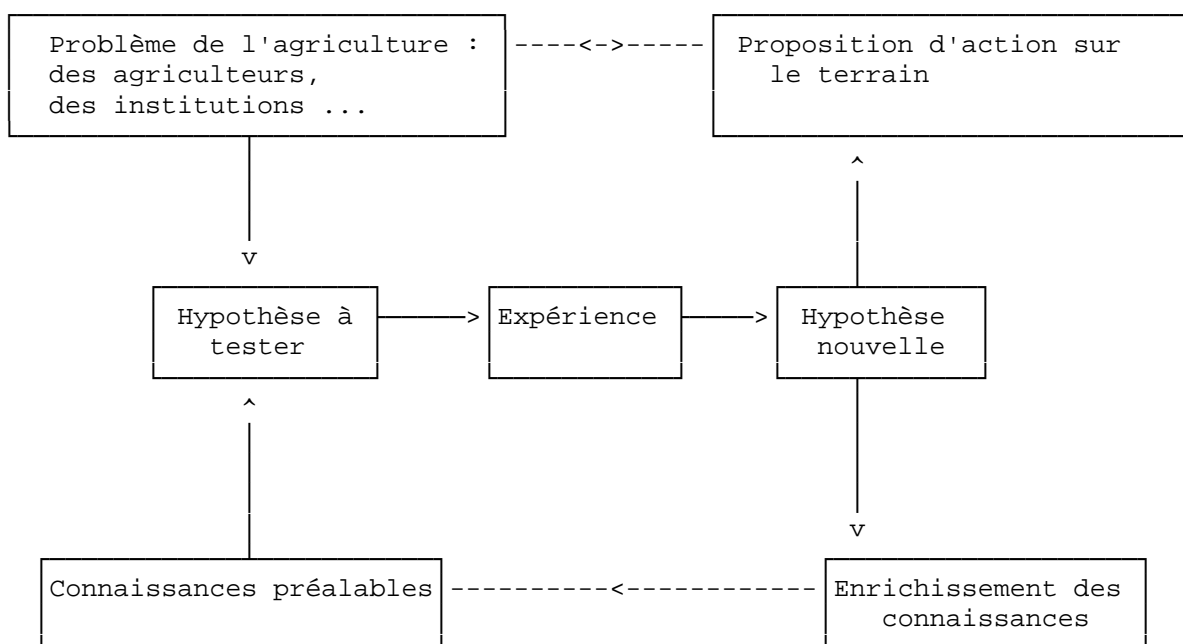
L'exposé des techniques d'enquête sort du cadre de ce document consacré à l'expérimentation. Mais les lecteurs intéressés peuvent se reporter aux trois documents sur ce sujet signalés en références bibliographiques.

## 2) DÉMARCHE DE L'EXPÉRIMENTATION :

---

L'expérimentation et l'inférence sont deux éléments essentiels de toute méthodologie scientifique. Le développement de la statistique a été en grande partie le résultat de la réflexion sur ces deux points.

La démarche générale de l'expérimentation agronomique peut se schématiser comme suit :



Les problèmes que rencontre l'agriculture d'une région et les connaissances que l'on peut avoir sur le sujet concourent à la définition d'un objectif de recherche. Celui-ci se traduit par une hypothèse à tester dans une expérience qui aboutit à une conclusion pouvant déboucher d'une part sur une proposition d'action sur le terrain et d'autre part sur un enrichissement des connaissances, car les résultats peuvent être utilisés pour résoudre des problèmes similaires.

### 3) QUELQUES DÉFINITIONS :

---

#### **3.1 Le traitement**

On appelle facteur toute série d'éléments de même nature conditionnant le phénomène étudié, en agriculture le comportement agronomique de la culture (par exemple la variété, la dose d'engrais, etc.).

Les facteurs qui sont l'objet même de l'expérience sont appelés facteurs étudiés. Ceux qui sont liés à la variabilité du milieu et introduits de façon à ce que leurs effets puissent être éliminés sont appelés les facteurs de contrôle (ou contrôlés, bien que ce dernier qualificatif puisse être à double sens).

Les différents éléments qui constituent un facteur sont appelés niveaux ou modalités. Le premier terme étant plutôt employé pour un facteur quantitatif et le deuxième pour un facteur qualitatif.

Enfin, on appelle objet ou traitement toute combinaison de niveaux ou de modalités de tous les facteurs étudiés. Si un seul facteur est étudié, il y a identité entre un traitement et un niveau.

#### **3.2 L'unité expérimentale**

L'unité expérimentale est l'unité élémentaire qui reçoit un traitement et sur laquelle est faite chaque mesure. Par exemple, dans les essais au champ, cette unité est la parcelle expérimentale. Elle est constituée d'une certaine étendue de terrain et d'un certain nombre de plants.

Un problème important est celui des effets de bordure : dans beaucoup d'essais au champ, il y a un risque d'interférences entre parcelles voisines (par exemple, il peut y avoir ombrage d'une variété haute sur une voisine plus basse, ou bien un engrais peut partiellement profiter aux parcelles voisines). Ceci peut affecter sérieusement la comparaison entre les traitements. La solution est de ne réaliser les mesures que sur la partie centrale de la parcelle expérimentale et pas sur les bordures. On distingue alors la parcelle utile sur laquelle sont effectuées les mesures et la parcelle expérimentale (= parcelle utile + bordures) qui reçoit le traitement.

Il n'existe pas de théorie générale pour décider de la dimension et de la forme de la parcelle expérimentale. On dispose toutefois de quelques principes pouvant guider ce choix.

Plus le matériel végétal est hétérogène, plus la parcelle sera de grande taille. Et, plus le contrôle de l'application des traitements sera difficile, plus la parcelle sera aussi de grande taille. Le problème est de trouver un compromis entre une parcelle pas trop petite pour pouvoir extrapoler les résultats à l'hectare et pas trop grande pour ne pas avoir un champ d'expérience (cf. définition ci-dessous) trop grand, c'est à dire prenant trop de place et risquant de rassembler des parcelles très hétérogènes. Il n'y a donc pas de règle absolue pour la dimension. Par exemple, l'usage est de prendre des parcelles de 10 à 100 m<sup>2</sup> pour les cultures annuelles.

Il n'y a pas non plus de règle absolue pour la forme des parcelles d'après DAGNELIE (1981) : lorsque le terrain est assez homogène, il y a intérêt à prendre des parcelles aussi carrées que possible, ce qui limite les effets de bordure. Mais l'homogénéité est rare et, dans le cas où il y a un gradient de fertilité, il y a plutôt intérêt à prendre des parcelles allongées dans le sens du gradient. De plus, le même auteur conseille de prendre des parcelles allongées en cas d'hétérogénéités "en taches", ce qui permettrait de recouper les zones de fertilités différentes.

La réalisation d'un essai à blanc est le moyen le plus utilisé pour avoir des informations sur la dimension et la forme des parcelles expérimentales à prendre dans un lieu donné. Un tel essai permet en outre de mesurer l'hétérogénéité des parcelles et de prévoir comment les regrouper en blocs homogènes, s'il y a lieu.

### **3.3 Le champ d'expérience et le champ d'application**

On appelle champ d'expérience l'ensemble des unités expérimentales, placées dans les conditions de l'expérience.

Et on appelle champ d'application ou champ de référence, le domaine sur lequel les résultats d'une expérience sont applicables. Il contient au moins le champ d'expérience.

Le champ d'application est en général mal connu en expérimentation du fait du caractère très analytique de la démarche. Mais c'est une notion importante dès lors que l'on analyse des regroupements d'essais.



### **3.4 L'observation**

L'observation se traduit par des mesures (ou variables) de caractères quantitatifs ou qualitatifs associés à chaque unité expérimentale.

On appelle erreur technique ou erreur de mesure, l'erreur que l'on fait lors de l'observation. Cette erreur doit être aussi petite que possible pour avoir des mesures fiables et précises, et par là avoir une expérimentation avec des résultats interprétables.

## **4) LES TROIS PRINCIPES DE LA PLANIFICATION EXPÉRIMENTALE :**

---

Un plan d'expérience définit : le champ d'expérience, les unités expérimentales et leur nombre, le mode de regroupement de celles-ci (s'il y a lieu) et l'affectation des traitements aux unités.

Pour assurer la validité de l'analyse statistique, trois principes doivent impérativement être appliqués :

1. **RANDOMISATION** : l'allocation des traitements aux unités doit être faite par un tirage aléatoire,
2. **RÉPÉTITIONS** : chaque traitement doit être affecté à plusieurs unités, afin de pouvoir estimer une erreur expérimentale,
3. **CONTRÔLE DE L'ERREUR** : il faut réduire la part non contrôlée de l'expérience, donc diminuer l'erreur expérimentale.

On a vu ce qu'est l'erreur technique. Mais il existe une autre source d'erreur. On appelle erreur unitaire l'erreur due à l'hétérogénéité des unités expérimentales, appelée aussi erreur d'hétérogénéité. On l'appelle aussi erreur de randomisation, car c'est la randomisation qui en fait une variable aléatoire dont on peut étudier la distribution. On appelle erreur expérimentale ou erreur résiduelle la somme de l'erreur unitaire et de l'erreur technique.

Pourquoi la randomisation ? tout d'abord elle permet d'éviter tout biais plus ou moins conscient. En termes mathématiques, elle permet d'assurer que l'erreur unitaire est d'espérance nulle.

Il est important de savoir que la randomisation n'annule pas, ni même ne diminue l'erreur. Elle permet d'éviter qu'elle soit systématique et de connaître suffisamment sa distribution pour assurer la validité du test des effets traitement.

La randomisation doit être faite indépendamment pour chaque essai. Aucun plan ne doit être réutilisé sans nouvelle randomisation.

A chaque plan d'expérience, correspond une procédure de randomisation spécifique. Ces procédures seront décrites pour chaque dispositif étudié ci-dessous.

#### Pourquoi des répétitions ?

Leur rôle est de permettre d'évaluer la part de l'erreur expérimentale dans la construction de l'observation. En effet, sans des répétitions de chaque traitement sur plusieurs unités, on ne peut pas distinguer l'effet dû au traitement de l'erreur expérimentale.

Enfin, il est nécessaire d'avoir une erreur expérimentale aussi petite que possible, c'est à dire de contrôler l'erreur.

On commence par choisir un champ expérimental aussi homogène que possible. Ceci suppose de ne pas se placer sur des emplacements d'anciennes souches etc. On améliore grandement l'homogénéité par la pratique de l'assolement expérimental. Il ne faut pas implanter un essai à l'emplacement d'un précédent essai qui a forcément introduit une hétérogénéité sur le terrain. C'est évident pour les essais de fumure, mais également tous les autres essais provoquent des exportations de minéraux différentes selon les parcelles ; les allées provoquent également une hétérogénéité supplémentaire. C'est pourquoi il est indispensable de faire une ou deux années de "repos" du terrain entre deux essais (jachère travaillée, multiplication d'une même variété, ou autres).

Lorsqu'une hétérogénéité est reconnue sur le terrain, on peut former des groupes homogènes de parcelles, ou blocs. Ceci permet de contrôler un facteur parasite non étudié (d'où le nom de facteur de contrôle) et d'améliorer la précision sur les facteurs étudiés.

## CHAPITRE 2

### DISPOSITIFS EXPÉRIMENTAUX CLASSIQUES

Ces dispositifs ont été classés d'après le mode de contrôle de l'hétérogénéité : pas de contrôle pour le plan en randomisation totale, un seul facteur de contrôle pour le plan en blocs et deux facteurs de contrôle pour le carré latin. Seuls les plans analysables par le logiciel STATITCF sont vus dans un premier temps mais une quatrième partie est consacrée aux autres plans d'expériences classiques non analysables par ce logiciel.

Enfin, une annexe est consacrée à divers sujets intéressant l'expérimentation : le calcul d'un nombre de répétitions, la gestion des données manquantes et le problème des écarts aux hypothèses de l'analyse de variance.

#### 1) PLANS EN RANDOMISATION TOTALE :

---

Ce type de dispositif, bien que peu utilisé en expérimentation agronomique, nous sert ici à introduire des notions fondamentales de l'analyse de variance.

##### **1.1 Un facteur étudié :**

On considère un champ expérimental composé de  $N$  unités et un facteur composé de  $t$  modalités (les traitements). On fixe  $n_1, n_2, \dots, n_t$  les nombres d'unités expérimentales auxquelles seront respectivement affectés les traitements 1, 2, ...,  $t$  (on a  $N = n_1 + n_2 + \dots + n_t$ ). Le dispositif est dit en randomisation totale si l'affectation des traitements se fait par un tirage aléatoire équiprobable parmi l'ensemble des unités expérimentales.

On peut réaliser ce tirage de la manière suivante : on numérote les unités de 1 à  $N$ . On place dans un sac  $n_1$  papiers portant le nom du traitement 1,  $n_2$  papiers portant le nom du traitement 2, ...,  $n_t$  papiers portant le nom du traitement  $t$ . Puis après avoir mélangé dans le sac, on tire les papiers un par

un et on affecte à l'Unité 1 le premier traitement tiré, puis à l'Unité 2 le deuxième traitement tiré, etc., jusqu'au dernier traitement tiré affecté à l'unité N.

Il existe d'autres méthodes de tirage aléatoire : l'utilisation d'un générateur de nombre au hasard d'un ordinateur, l'utilisation de tables de nombres au hasard ou de permutations aléatoires.

En général, et c'est le cas du logiciel STATITCF, on impose que le nombre de répétitions par traitement soit identique pour tous les traitements (égal à n). On a alors  $N = n t$ .

Après avoir réalisé l'expérience, on obtient des valeurs mesurées sur les unités expérimentales (par exemple des rendements s'il s'agit d'une expérience au champ). On note  $Y_{ij}$  la valeur obtenue sur la  $j^{\text{e}}$  répétition du traitement  $i$ .

L'objectif est de savoir si les traitements sont ou non identiques. On pose alors deux hypothèses. Une première hypothèse est appelée  $H_0$  : les traitements sont identiques. Dans ce cas, toutes les unités ont sensiblement la même valeur, que l'on note  $\mu$ .

Une deuxième hypothèse est appelée  $H_1$  : les traitements ne sont pas identiques. Dans ce cas, seules les unités portant un même traitement ont sensiblement la même valeur mesurée. On note  $\mu_1$  cette valeur pour le traitement 1,  $\mu_2$  pour le traitement 2, etc.

On pose souvent  $\mu_i = \mu + \alpha_i$  (avec  $\sum_i \alpha_i = 0$  si le nombre de répétitions est identique pour tous les traitements), où  $\alpha_i$  est appelé l'effet du traitement  $i$ .

Il faut décider si l'on retient  $H_0$  ou  $H_1$  à partir des données mesurées. Dans cette optique, les données sont analysées au moyen du modèle :

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

où  $\varepsilon_{ij}$  est l'erreur résiduelle faite sur la  $j^{\text{e}}$  répétition du traitement  $i$ .

On suppose que  $E(\varepsilon_{ij}) = 0$ ,  $\text{var}(\varepsilon_{ij}) = \sigma^2$  et les  $\varepsilon_{ij}$  sont indépendantes entre elles.

Du fait de ces erreurs, dont on a vu qu'elles étaient dues à l'hétérogénéité des unités expérimentales et à la mesure, on n'a pas accès aux valeurs exactes de  $\mu$  et des  $\alpha_i$ . On ne peut que les estimer à partir des  $Y_{ij}$ .

On estime  $\mu$  par  $M = \sum_{ij} Y_{ij}/N$  ;

On estime  $\alpha_i$  par  $A_i = \sum_j Y_{ij}/n - M$

(On suppose qu'on a  $n$  répétitions par traitement, donc que  $N = n t$ ).

$M$  est la moyenne de toutes les données mesurées (moyenne générale) et  $A_i$  est l'écart entre la moyenne des données du traitement  $i$  et la moyenne générale. On en tire une estimation de chaque erreur résiduelle (appelée résidu) qui est l'écart entre la donnée observée ( $Y_{ij}$ ) et la donnée estimée par le modèle ( $M + A_i$ ) ; ici la différence entre la donnée observée et la moyenne du traitement correspondant :

$$E_{ij} = Y_{ij} - M - A_i = Y_{ij} - \sum_j Y_{ij}/n$$

Pour décider de l'hypothèse à retenir ( $H_0$  : les  $\alpha_i$  sont tous nuls ou  $H_1$  : les  $\alpha_i$  ne sont pas tous nuls), on fait ce que l'on appelle de l'analyse de variance :

- on estime la dispersion des effets des traitements, elle sera d'autant plus grande que les  $\alpha_i$  seront différents ;
- on estime la dispersion due aux fluctuations aléatoires, qui sert de base de référence ; elle sera d'autant plus grande que les résidus seront grands ;
- et on compare ces deux dispersions.

La dispersion des effets traitement est estimée par ce qu'on appelle la somme des carrés traitement (notée SCT) et la dispersion due aux fluctuations aléatoires est estimée par ce qu'on appelle la somme des carrés résiduelle (notée SCR). On obtient ces estimations sachant que :

$$Y_{ij} - M = A_i + E_{ij}$$

et on montre que :

$$\sum_{ij} (Y_{ij} - M)^2 = \sum_i n_i A_i^2 + \sum_{ij} E_{ij}^2$$

$$SC \text{ totale} = SCT + SCR$$

Mais on ne peut pas comparer directement les deux sommes de carrés d'écarts. Il faut les diviser par leur nombre de degrés de liberté respectif, c'est à dire par le nombre de valeurs indépendantes qui les composent :  $t-1$  pour SCT et  $N-t$  pour SCR.

On dresse ainsi le tableau d'analyse de variance suivant :

Source de variation	SCE	DI	CM	F
Traitement	$SCT = \sum iniAi^2$	$t-1$	$CMT = SCT/(t-1)$	$F = CMT/CMR$
Résiduelle	$SCR = \sum ijEij^2$	$N-t$	$CMR = SCR/(N-t)$	
Totale	$SC_{totale} = \sum ij(Y_{ij}-M)^2$	$N-1$		

**Légende :**

**SCE** : Somme des carrés des écarts

**dl** : Degrés de liberté

**CM** : Carré moyen

Le carré moyen résiduel (CMR) est appelé aussi variance résiduelle ; c'est la base de référence qui sert à tester les écarts entre traitements. En effet, si l'hypothèse  $H_0$  est vraie, le rapport des deux carrés moyens suit une loi de Fisher-Snedecor avec  $t-1$  degrés de liberté au numérateur et  $N-t$  degrés de liberté au dénominateur.

On s'intéresse alors à la probabilité qu'un  $F(t-1 ; N-t)$  dépasse la valeur effectivement calculée. Si cette probabilité est inférieure à un seuil préalablement fixé, égal à  $\alpha$  (par exemple 5 % ou 1 %), on dit que les traitements sont significativement différents au niveau  $\alpha$ . Ceci veut dire qu'on rejette  $H_0$  et qu'on retient  $H_1$ .

Au contraire, si cette probabilité est supérieure à  $\alpha$ , on dit que les traitements ne sont pas significativement différents au niveau  $\alpha$ . Ceci veut dire qu'on accepte  $H_0$  et qu'on ne retient pas  $H_1$ .

On voit que la conclusion du test (et la décision prise) dépend du niveau  $\alpha$  auquel il est fait. Ce niveau  $\alpha$  est la probabilité consentie de rejeter  $H_0$  alors qu'en réalité elle est vraie, c'est la probabilité de faire une erreur de 1re espèce. On l'appelle le risque de 1re espèce.

Mais un autre type d'erreur peut se produire ; on peut accepter  $H_0$  alors qu'en réalité  $H_1$  est vraie (les traitements sont en fait différents). C'est l'erreur de 2e espèce. Sa probabilité est notée  $\beta$ , le risque de 2e espèce.

On s'intéresse aussi à la grandeur  $1 - \beta$ , qui est la probabilité de retenir  $H_1$  alors qu' $H_1$  est vraie (la décision est bonne).  $1 - \beta$  est appelée la puissance du test.

Ces probabilités sont récapitulées dans le tableau ci-dessous :

		Décision	
		$H_0$	$H_1$
Situation Réelle	$H_0$ vraie	$1 - \alpha$	$\alpha$ risque de 1re espèce
	$H_1$ vraie	$\beta$ risque de 2e espèce	$1 - \beta$ Puissance

Mais les deux risques ne sont pas symétriques : le risque  $\alpha$  est une valeur préalablement fixée, alors que le risque  $\beta$  est une fonction des différences entre traitements dans le cadre d'une expérience. En fait  $\beta$  (et donc aussi la puissance  $1 - \beta$ ) est une fonction :

- du risque  $\alpha$ ,
- des différences vraies entre traitements,
- de la variabilité résiduelle,
- du nombre de répétitions.

On a donc cinq paramètres liés. Il suffit d'en fixer quatre pour déterminer le cinquième. C'est ce qui permet en particulier de déterminer un nombre de répétitions dans une expérience.

Le test F qui vient d'être présenté est un test global, il permet de dire si les traitements sont différents ou non. Mais on ne sait pas quels traitements se différencient. Pour le savoir il faut faire des

comparaisons multiples de moyennes. Il existe de nombreux tests, et le choix de l'un d'entre eux est fonction de l'objectif : quelques uns vont être détaillés ici.

Toute comparaison de moyennes s'écrit sous la forme d'un contraste. Par définition, un contraste est une combinaison linéaire des moyennes des traitements dont la somme des coefficients est nulle.

$$C = \sum_i a_i \mu_i \text{ est un contraste si } \sum_i a_i = 0$$

Par exemple,

$$\mu_1 - \mu_2 \text{ est un contraste estimé par } M_1 - M_2,$$

$$\mu_1 - 2 \mu_2 + \mu_3 \text{ est un contraste estimé par } M_1 - 2 M_2 + M_3,$$

(on note  $M_i$  la moyenne des observations du traitement  $i$ , d'après nos précédentes notations,  $M_i = M + A_i$ ).

Le contraste du type  $\mu_1 - \mu_2$  (simple comparaison entre deux traitements) est appelé contraste élémentaire. Beaucoup de méthodes de comparaisons multiples ne s'intéressent qu'aux comparaisons 2 à 2 entre traitements, donc aux contrastes élémentaires.

Deux contrastes sont dits orthogonaux si la somme des produits de leurs coefficients est nulle, c'est à dire si leur produit scalaire est nul.

$$C_1 = \sum_i a_i \mu_i \text{ et } C_2 = \sum_i b_i \mu_i ,$$

$$C_1 \text{ et } C_2 \text{ orthogonaux si } \sum_i a_i b_i = 0$$

$$\text{Par exemple, } C_1 = \mu_1 - \mu_3 \text{ et } C_2 = \mu_1 - 2 \mu_2 + \mu_3$$

sont des contrastes orthogonaux. De plus, avec  $t$  traitements, on ne peut avoir que  $t-1$  contrastes orthogonaux en tout : ils correspondent à une décomposition des  $t - 1$  degrés de liberté de la variation traitement en  $t - 1$  sommes de carrés dont la somme vaut SCT.

Maintenant que la notion de contraste a été vue, il faut envisager la notion de risque d'erreur, très importante dans les comparaisons multiples. On distingue :



- le taux d'erreur au niveau de la comparaison (EC), égal au nombre de comparaisons faussement déclarées significatives divisé par le nombre total de comparaisons dont le contraste est nul ( $H_0$  vraie),

- et le taux d'erreur au niveau de l'expérience, qui peut se définir de deux manières, et qui est relatif aux comparaisons que l'on considère simultanément ; il peut se définir tout d'abord comme le nombre d'expériences, dans lesquelles les traitements sont équivalents, avec au moins une comparaison déclarée faussement significative divisé par le nombre total d'expériences dans lesquelles tous les traitements sont équivalents ( $H_0$  complète vraie) ; on l'appelle taux d'erreur au niveau de l'expérience pour l'hypothèse nulle complète (EEC). Et il peut se définir sinon comme le nombre d'expériences avec au moins une comparaison déclarée faussement significative divisé par le nombre d'expériences avec au moins une comparaison dont le contraste est nul ( $H_0$  complète ou partielle vraie) , on l'appelle taux d'erreur au niveau de l'expérience pour l'hypothèse nulle complète ou partielle (EECP).

Il existe un grand nombre de méthodes de comparaisons multiples, elles ne diffèrent pas seulement par les contrastes qu'elles considèrent mais aussi par le risque qu'elles contrôlent. Par exemple, le logiciel STATITCF propose quatre méthodes :

- La méthode de NEWMAN-KEULS permet de tester toutes les comparaisons deux à deux (contrastes élémentaires) en contrôlant le risque que l'on a noté EEC. Elle ne s'applique que lorsqu'on veut faire des comparaisons entre des traitements indifférenciés a priori. On commence par classer les moyennes en ordre décroissant.

Première étape : on considère l'ensemble des  $t$  moyennes, soient  $v$  le nombre de degrés de liberté pour estimer l'erreur résiduelle,  $ETR$  l'écart-type résiduel ( $ETR = \sqrt{CMR}$ ) et  $n$  le nombre de répétitions par traitement. On détermine la plus petite amplitude significative ( $ppas$ ), c'est à dire la

$$ppas = q_{\alpha}(t,v) ETR \sqrt{1/n}$$

plus petite différence significative entre la plus grande et la plus petite des moyennes :

$q_{\alpha}(t,v)$ : valeur lue dans la table de NEWMAN-KEULS au niveau  $\alpha$   
(on prend  $\alpha$  égal à 5 % ou 1% ici)

Si l'amplitude observée est plus grande que la  $ppas$  on rejette l'hypothèse d'égalité des  $t$  traitements et on passe à la deuxième étape, sinon on s'arrête (homogénéité des  $t$  moyennes).

Deuxième étape : on considère les deux groupes de  $t - 1$  moyennes. La ppas est alors donnée par:

$$ppas = q_{\alpha}(t-1, \nu) ETR \sqrt{1/n}$$

Si l'amplitude considérée est inférieure à la ppas, on s'arrête et on accepte l'hypothèse d'homogénéité du groupe. Sinon on continue la procédure sur les groupes plus petits jusqu'à arriver à des groupes homogènes.

Le logiciel STATITCF, en plus du résultat final, fournit les ppas calculées pour les groupes entre 2 et  $t$  moyennes.

- La méthode de BONFERRONI permet également de tester toutes les comparaisons deux à deux, mais en contrôlant le risque EECP. Cette méthode contrôle un risque plus étendu que le risque EEC contrôlé par la méthode de NEWMAN-KEULS.

Cette méthode consiste à comparer toutes les différences entre deux moyennes de traitement à la valeur :

$$t_x(\nu) \sqrt{2CMR/n}$$

où, pour avoir un risque EECP égal à  $\alpha$ , on lit la valeur  $t_x(\nu)$  dans une table de Student ( $\nu$  degrés de liberté pour la résiduelle) au niveau  $\alpha = 2\alpha/(t - 1)$ .

- La méthode de DUNNETT contrôle, comme celle de BONFERRONI, le risque EECP, mais ne s'intéresse qu'à quelques comparaisons deux à deux : les  $t - 1$  comparaisons à un témoin. Pour cela, on détermine un plus petit écart significatif au témoin (ppes) :

$$ppes = d_{\alpha}(t-1, \nu) \sqrt{2 CMR/n}$$

$d_{\alpha}(t-1, \nu)$  valeur lue dans la table de DUNNETT au niveau  $\alpha$   
(on prend  $\alpha$  égal à 5 % ou 1 % ici)

On définit aussi trois ensembles :

- l'ensemble des traitements significativement supérieurs au témoin,
- l'ensemble des traitements non significativement différents du témoin,
- l'ensemble des traitements significativement inférieurs au témoin.

- La méthode des contrastes proposée par STATITCF permet de tester des contrastes orthogonaux préalablement définis. Cette méthode contrôle le risque qui a été noté EC (au niveau de la comparaison).

Sur le contraste  $C = \sum_i a_i \mu_i$ , on teste sa nullité au niveau  $\alpha$  en comparant  $\sum_i a_i M_i$  à la valeur :

$$t_\alpha(v) \sqrt{CMR \sum_i a_i^2/n}$$

En fait, le logiciel STATITCF calcule la probabilité  $x$  telle que :

$$t_x(v) = \sum_i a_i M_i \sqrt{n/(CMR \sum_i a_i^2)}$$

Deux petits exemples vont aider à faire comprendre l'utilité de cette méthode des contrastes.

Exemple 1 : on considère cinq traitements, un témoin et des applications de deux produits (A et B) à deux doses (1 et 2). On les note : T, A1, A2, B1 et B2.

On définit alors quatre contrastes orthogonaux testant les effets suivants :

	$\mu_T$	$\mu_{A1}$	$\mu_{A2}$	$\mu_{B1}$	$\mu_{B2}$
Témoin vs application	4	- 1	- 1	- 1	- 1
Effet produit (A vs B)	0	1	1	- 1	- 1
Effet dose	0	1	- 1	1	- 1
Interaction produit x dose	0	1	- 1	- 1	1

Exemple 2 : on considère un produit appliqué à quatre doses équidistantes (notées de 1 à 4). La courbe de réponse peut être étudiée alors par les trois contrastes orthogonaux suivants :

$$\mu_1 \quad \mu_2 \quad \mu_3 \quad \mu_4$$

Effet linéaire	- 3	- 1	1	3
Effet quadratique	1	- 1	- 1	1
Effet cubique	1	- 3	3	- 1

## **1.2 Deux facteurs étudiés sans répétition :**

Ce dispositif n'est, à juste titre, pratiquement pas utilisé en expérimentation agronomique, mais il va servir ici à introduire une notion importante : la notion de "répétitions cachées".

On considère par exemple, une expérience avec deux facteurs croisés : la dose d'engrais (3 modalités) et le travail du sol (2 modalités). Ceci fait six traitements que l'on ne répète pas.

Valeurs Observées		Doses		
		1	2	3
Travail du sol	1	$Y_1$	$Y_2$	$Y_3$
	2	$Y_4$	$Y_5$	$Y_6$

Les traitements ne sont pas répétés, mais  $Y_1 - Y_4$ ,  $Y_2 - Y_5$  et  $Y_3 - Y_6$  sont trois répétitions d'un même écart entre les deux modes de travail du sol.

De même,  $Y_1 - Y_3$  et  $Y_4 - Y_6$  sont deux répétitions d'un même écart entre deux doses d'engrais.

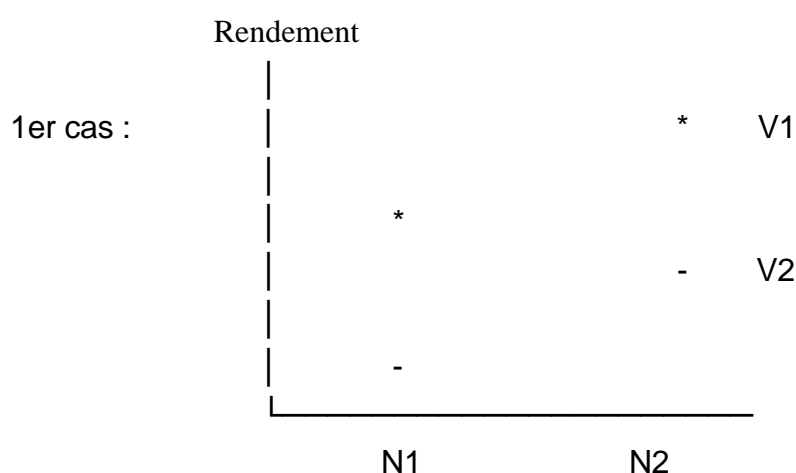
Aussi, l'analyse statistique de ce type de dispositif est éventuellement possible, pour peu que l'on suppose que les fluctuations entre les observations d'un même écart soient aléatoires. Ce qui revient à dire que l'interaction entre les deux facteurs est aléatoire. Et on se sert de cette interaction comme base de référence (variance résiduelle) pour tester les effets des deux facteurs.

Ceci amène à la notion d'interaction.

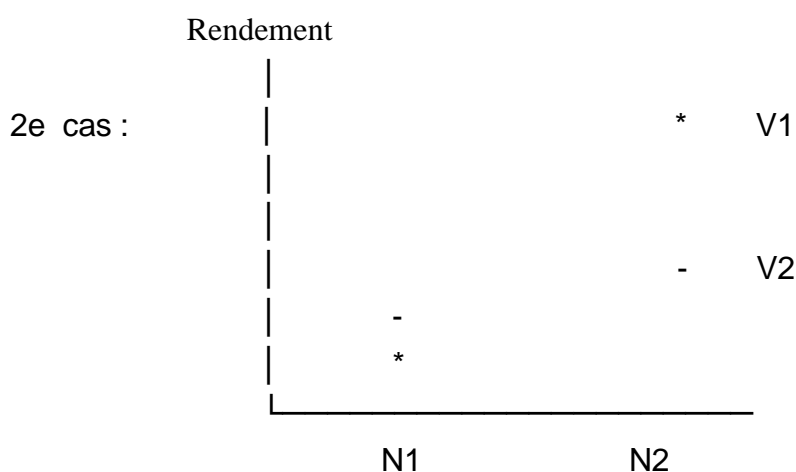
### 1.3 Deux facteurs étudiés avec répétitions :

On considère ici que les traitements sont issus de la combinaison de deux facteurs : le facteur A à I modalités et le facteur B à J modalités. Cela fait  $IJ$  traitements en tout, répétés chacun  $n$  fois. Donc  $N = IJn$ .

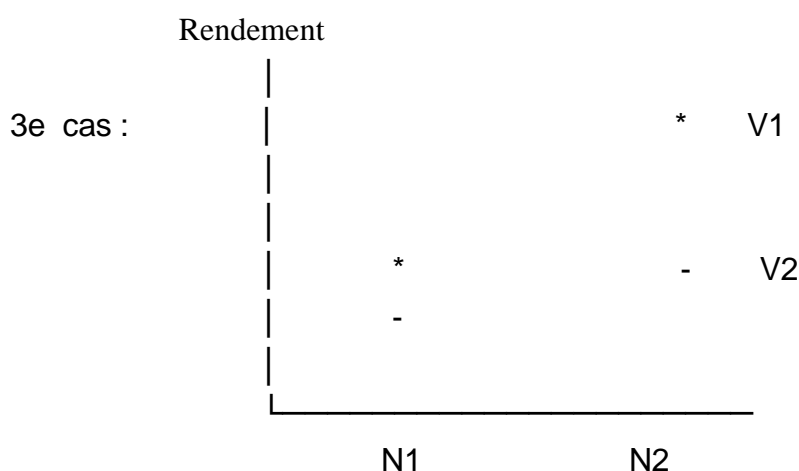
Une notion très importante pour étudier la combinaison de ces facteurs est la notion d'interaction. Un exemple va permettre de la mettre en lumière. On considère deux facteurs croisés, la variété avec deux modalités (V1 et V2) et la dose d'azote avec deux modalités aussi (N1 et N2), et on observe le rendement d'une culture.



L'écart entre les deux variétés ne dépend pas de la dose d'azote. On dit qu'il n'y a pas d'interaction entre les deux facteurs.



L'écart entre les deux variétés dépend de la dose d'azote. On dit qu'il y a une interaction entre les deux facteurs. De plus, celle-ci ne conserve pas l'ordre des traitements.



On a ici une interaction qui conserve l'ordre des traitements. Alors que les notions d'effet variété et d'effet dose d'azote peuvent conserver un sens contrairement au 2e cas.

Ce qui vient d'être vu a une conséquence importante : la première chose à tester, dans un essai étudiant deux facteurs A et B, est l'interaction entre A et B, car en présence d'interaction les effets principaux de A et de B peuvent ne pas avoir de sens.

L'analyse statistique est faite au moyen du modèle suivant :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

où  $Y_{ijk}$  est la donnée observée sur la  $k^e$  répétition du traitement (i,j), c'est à dire le traitement constitué de la modalité i du facteur A et de la modalité j du facteur B,

$\mu$  est la moyenne générale,

$\alpha_i$  est l'effet principal de la modalité i de A ( $\sum_i \alpha_i = 0$ ),

$\beta_j$  est l'effet principal de la modalité j de B ( $\sum_j \beta_j = 0$ ),

$(\alpha\beta)_{ij}$  est l'interaction du traitement (i,j), c'est à dire l'écart au modèle additif  
 $(\sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0)$ ,

$\varepsilon_{ij}$  est l'erreur résiduelle (avec les mêmes hypothèses qu'au 1.1).

**Notations :**

$Y_{...}$  = moyenne de toutes les observations, elle estime  $\mu$

$Y_{i..}$  = moyenne des observations portant la modalité i de A, on estime  $\alpha_i$  par  $Y_{i..} - Y_{...}$

$Y_{.j.}$  = moyenne des observations portant la modalité j de B, on estime  $\beta_j$  par  $Y_{.j.} - Y_{...}$

$Y_{ij.}$  = moyenne des observations portant le traitement (i,j),

on estime  $(\alpha\beta)_{ij}$  par  $Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}$

Le tableau d'analyse de variance s'écrit :

Variation	SCE	DI	CM	F
Facteur A	$\sum_i nJ (Y_{i..} - Y_{...})^2$	I - 1	SCA/(I-1)	CMA/CMR
Facteur B	$\sum_j nI (Y_{.j.} - Y_{...})^2$	J - 1	SCB/(J-1)	CMB/CMR
Interaction A x B	$\sum_{ij} n (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...})^2$	(I-1)(J-1)	SCI/(I-1)(J-1)	CMI/CMR
Résiduelle	$\sum_{ijk} (Y_{ijk} - Y_{ij.})^2$	N - IJ	SCR/(N-IJ)	

## 2) LES PLANS EN BLOCS COMPLETS :

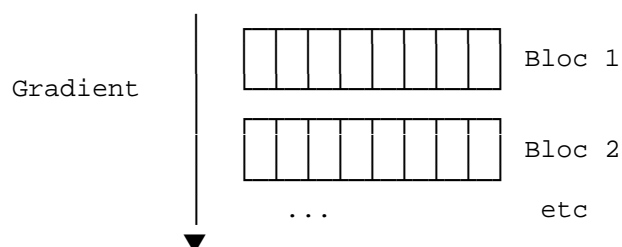
---

Le chapitre 1 a permis d'insister sur la nécessité de contrôler l'erreur expérimentale. La manière la plus courante de contrôler une part de celle-ci, l'erreur d'hétérogénéité, est de constituer des blocs.

Par définition, un bloc est un groupe homogène d'unités expérimentales. C'est à dire que la variabilité du phénomène étudié doit être plus faible entre unités appartenant à un même bloc qu'entre unités appartenant à des blocs différents.

Le regroupement des unités expérimentales en blocs résulte d'une information a priori dont on dispose sur le champ d'expérience. En particulier, dans le domaine agronomique, il est nécessaire de construire le dispositif à partir de relevés précis dans le champ d'expérience.

Le cas le plus classique se trouve lorsqu'il y a un gradient, induit par une pente, ou par une lisière brise-vent ou apportant de l'ombrage, ou bien par la proximité d'un cours d'eau, ou encore par d'autres causes.



Les blocs sont alors allongés perpendiculairement au gradient, les parcelles étant allongées dans le sens de ce gradient.

Le plus souvent les blocs sont compacts, c'est à dire d'un seul tenant. Il peut toutefois être nécessaire de les "éclater" pour éviter de fortes sources d'hétérogénéité (anciennes termitières ou souches) ou pour tenir comptes d'informations cartographiées (essais à blanc, analyses de sols, etc.).

La mise en place des blocs doit pouvoir être justifiée. Il arrive souvent, lorsque l'expérimentateur n'a pas de critère intéressant, que les unités expérimentales voisines soient regroupées pour constituer des blocs les plus compacts possible. En agissant ainsi, il subodore de moindres différences entre des parcelles géographiquement proches qu'entre parcelles éloignées. Si cette attitude relève d'un "certain bon sens", elle n'en constitue pas moins, dans de nombreux cas, une solution mal justifiée.

Enfin les blocs doivent être différents entre eux mais pas trop, pour éviter une éventuelle interaction bloc-traitement.



## **2.1 Répartition des traitements totalement aléatoires dans les blocs**

La procédure de randomisation consiste en un tirage aléatoire par bloc pour affecter les traitements aux unités expérimentales. On se limite ici au cas le plus simple où chaque traitement figure une fois et une seule dans chaque bloc. Donc un bloc comporte  $t$  traitements, si  $n$  est le nombre de blocs, le nombre total de parcelles est  $N = n t$ .

La randomisation peut être réalisée de la manière suivante : on numérote les parcelles d'un bloc de 1 à  $t$ , on place dans un sac  $t$  papiers portant chacun le nom d'un traitement différent, et on effectue un tirage sans remise : on affecte le premier traitement tiré à la première parcelle, le deuxième traitement tiré à la deuxième parcelle, etc. Et on renouvelle cette opération pour tous les blocs.

D'autres méthodes sont possibles pour faire la randomisation. On peut utiliser des tables de permutations aléatoires, ou bien utiliser une fonction génératrice de nombres aléatoires d'un micro-ordinateur. Il faut signaler de plus que le logiciel STATITCF a une option qui permet de générer directement les plans d'expérience qu'il peut analyser, en particulier le plan en blocs complets.

L'analyse statistique de ce type de dispositif est faite au moyen du modèle d'analyse de variance :

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

avec les mêmes notations que précédemment pour  $Y_{ij}$ ,  $\mu$  et  $\varepsilon_{ij}$ ,

et où  $\alpha_i$  = effet du traitement  $i$  ( $\sum_i \alpha_i = 0$ )

et  $\beta_j$  = effet du bloc  $j$  ( $\sum_j \beta_j = 0$ )

Il s'agit d'un modèle additif à deux facteurs (on suppose l'absence d'interaction bloc-traitement) : un facteur étudié (traitement) et un facteur de contrôle (bloc).

$\mu$  est estimé par  $Y_{..}$  ,

$\alpha_i$  est estimé par  $Y_{i.} - Y_{..}$  ,

$\beta_j$  est estimé par  $Y_{.j} - Y_{..}$  ,

et les erreurs résiduelles  $\varepsilon_{ij}$  sont estimées par les résidus  $Y_{ij} - Y_{i.} - Y_{.j} + Y_{..}$ .

D'où le tableau d'analyse de variance :

Variation	SCE	dl	CM	F
Traitement	$\sum_i n(Y_{i.} - Y_{..})^2$	t-1	SCT/(t-1)	CMT/CM R
Bloc	$\sum_j t(Y_{.j} - Y_{..})^2$	n-1	SCB/(n-1)	CMB/C MR
Résiduelle	$\sum_{ij} (Y_{ij} - Y_{i.} - Y_{.j} + Y_{..})^2$	(t-1)(n-1)	SCR/(t-1)(n-1)	

Ce qui vient d'être présenté correspond à l'analyse d'un plan en blocs complets avec un seul facteur étudié. Le cas où il y a plusieurs facteurs se traite de la même manière qu'en randomisation totale, on tient compte en plus du facteur bloc, comme ci-dessus.

## **2.2 Répartition des traitements non totalement aléatoire dans les blocs**

On considère des dispositifs avec au moins deux facteurs étudiés (factoriels). Ces dispositifs qui vont être vus ici, split-plot (à parcelles divisées) et criss-cross (en bandes croisées), ne sont pas des plans en blocs complets mais ont un rapport étroit avec eux, au moins en ce qui concerne ceux analysables par le logiciel STATITCF. En effet, ceux-ci peuvent être considérés comme des plans en blocs complets dont la répartition des traitements n'est pas totalement aléatoire à l'intérieur de chaque bloc.

Le split-plot (dispositif à parcelles subdivisées) se réalise de la manière suivante : on constitue un plan en blocs complets avec des unités expérimentales qu'on appelle les grandes

parcelles (ce terme provient de l'expérimentation au champ, mais recouvre une notion plus générale). Les modalités d'un premier facteur sont randomisées dans ces grandes parcelles. Puis on subdivise des grandes parcelles en autant de petites parcelles (les unités expérimentales définitives du dispositif) qu'il y a de modalités du deuxième facteur. Ces modalités sont affectées aléatoirement aux petites parcelles, et ce indépendamment pour chacune des grandes parcelles (ou sous-blocs).

Ce dispositif à deux facteurs est le split-plot le plus simple, mais il peut être un peu plus compliqué lorsque le premier facteur (ou le deuxième) est lui-même la combinaison de deux facteurs. On a alors un split-plot à trois facteurs étudiés. Il peut aussi y avoir trois niveaux de parcelles expérimentales dans le split-plot trois étages : de grandes parcelles, des parcelles intermédiaires et des petites parcelles. Un premier facteur est affecté aléatoirement aux grandes parcelles, un deuxième facteur aux parcelles intermédiaires et un troisième facteur aux petites parcelles.

Les split-plots analysables par le logiciel STATITCF sont des plans dont le dispositif de base (des grandes parcelles) est un plan en blocs complets. Mais il faut savoir que le split-plot peut éventuellement être un plan dont le dispositif de base est différent : plan en randomisation totale, carré latin, etc. Ils sont toutefois beaucoup moins utilisés.

L'analyse statistique du split-plot à deux facteurs (A à I modalités et B à J modalités) le plus simple est réalisé au moyen du modèle :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \delta_k + B_{ik} + \varepsilon_{ijk}$$

où  $Y_{ijk}$  est la donnée observée dans le bloc k du traitement (i,j), c'est à dire la modalité i du facteur A et j du facteur B,

$\mu$  est la moyenne générale estimée par  $Y_{...}$ ,

$\alpha_i$  est l'effet de la modalité i du facteur A, estimé par  $Y_{i..} - Y_{...}$ ,

$\beta_j$  est l'effet de la modalité j du facteur B, estimé par  $Y_{.j.} - Y_{...}$ ,

$(\alpha\beta)_{ij}$  est l'interaction du traitement (i,j), estimée par  $Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}$ ,

$\delta_k$  est l'effet du bloc k, estimé par  $Y_{..k} - Y_{...}$ ,

$B_{ik}$  est l'effet aléatoire de la grande parcelle du bloc k qui porte la modalité i du facteur A,  
 et  $\varepsilon_{ijk}$  est l'erreur résiduelle sur la petite parcelle du bloc k qui porte le traitement (i,j).

D'où le tableau d'analyse de variance à deux étages, d'une part l'analyse du facteur A (grandes parcelles), d'autre part l'analyse de l'interaction A x B et du facteur B (petites parcelles) :

Variation	SCE	dl	CM	F
Facteur A	$\sum_i nJ (Y_{i..} - Y_{...})^2$	I - 1	SCA/(I-1)	CMA/CMR1
Bloc	$\sum_k IJ (Y_{..k} - Y_{...})^2$	n - 1	SCb/(n-1)	CMb/CMR1
Résiduelle 1	$\sum_{ik} J (Y_{i.k} - Y_{i..} - Y_{..k} + Y_{...})^2$	(I-1)(n-1)	SCR1/ (I-1)(n-1)	
Facteur B	$\sum_j nI (Y_{.j.} - Y_{...})^2$	J - 1	SCB/ (J-1)	CMB/CMR2
Interaction A x B	$\sum_{ij} n (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...})^2$	(I-1) (J-1)	SCI/(I-1)(J-1)	CMI/CMR2
Résiduelle 2	$\sum_{ijk} (Y_{ijk} - Y_{i.k} - Y_{ij.} + Y_{i..} + Y_{.j.} - Y_{...})^2$	I (J-1)(n-1)	SCR2/I(J-1)(n-1)	

Pour comparer les moyennes en l'absence d'interaction entre A et B, on compare les modalités du facteur A en utilisant la variance résiduelle 1 et les modalités du facteur B en utilisant la variance résiduelle 2.

On estime  $\text{var}(Y_{1..} - Y_{2..})$  par  $2 \text{CMR1}/nJ$ , et  $\text{var}(Y_{.1.} - Y_{.2.})$  par  $2 \text{CMR2}/nI$

Lorsqu'il y a une interaction A x B, c'est plus compliqué, mais on peut faire l'approximation que toutes les comparaisons entre deux traitements sont de variance  $2 \text{CMR2}/n$ . Cette approximation est d'autant plus valable que le nombre de modalités du deuxième facteur (J) est grand.

Ce type de dispositif comporte quelques avantages pratiques et des inconvénients d'ordre statistique. Il permet d'utiliser des parcelles de tailles plus importantes pour des facteurs le

nécessitant (travail du sol, irrigation, etc.) sans augmenter trop fortement la surface du champ d'expérience. De plus, la disposition sur le terrain est assez claire, ce qui est intéressant lorsque l'essai a aussi une fonction de démonstration.

Mais il introduit une perte de précision importante sur le facteur A par rapport à un essai factoriel en blocs complets. Le facteur B et l'interaction sont testés avec une variance résiduelle dont le nombre de degrés de liberté est moindre que dans l'essai factoriel en blocs complets. Cela peut introduire aussi une petite perte de précision s'il n'y a pas d'effet grandes parcelles, mais au contraire un peu de gain de précision s'il y a des effets grandes parcelles non négligeables.

Le criss-cross (dispositif en bandes croisées) se réalise de la manière suivante : on constitue des blocs de parcelles homogènes. On divise ces blocs en autant de colonnes qu'il y a de niveaux de facteur A, en autant de lignes qu'il y a de niveaux du facteur B. Une unité expérimentale est donc l'intersection d'une ligne et d'une colonne. Puis on affecte aléatoirement les niveaux de A aux colonnes, et les niveaux de B aux lignes. On effectue ces opérations indépendamment pour chaque bloc.

Ce type de dispositif peut se justifier lorsque les deux facteurs testés nécessitent des parcelles de taille importante. Toutefois, il introduit une grande perte de précision sur les effets principaux de A et de B. Seule l'interaction n'est pas sacrifiée s'il y a des effets ligne et colonne non négligeables. L'analyse statistique doit se faire à trois niveaux : les colonnes pour tester le facteur A, les lignes pour tester le facteur B et les unités expérimentales pour tester l'interaction.

Cette analyse est réalisée au moyen du modèle :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \delta_k + C_{ik} + L_{jk} + \varepsilon_{ijk}$$

où  $Y_{ijk}$  est la donnée observée dans le bloc k du traitement (i,j), avec les mêmes notations que dans le split-plot pour  $\mu$ ,  $\alpha_i$ ,  $\beta_j$ ,  $(\alpha\beta)_{ij}$ ,  $\delta_k$  et  $\varepsilon_{ijk}$ ,

$C_{ik}$  est l'effet aléatoire de la colonne qui porte la modalité i de A dans le bloc k,

et  $L_{jk}$  est l'effet aléatoire de la ligne qui porte la modalité j de B dans le bloc k.

D'où le tableau d'analyse de variance à trois étages :

Variation	SCE	DI	CM	F
Facteur A	$\sum_i nJ(Y_{i..}-Y_{...})^2$	I-1	SCA/(I-1)	CMA/ CMR1
Résiduelle 1	$\sum_{ik} J(Y_{i.k}-Y_{i..}-Y_{..k} + Y_{...})^2$	(I-1)(n-1)	SCR1 / (I-1)(n-1)	-
Facteur B	$\sum_j nI(Y_{.j.}-Y_{...})^2$	J-1	SCRB / (J-1)	CMB/ CMR2
Résiduelle 2	$\sum_{jk} I(Y_{.jk}-Y_{.j.}-Y_{..k} + Y_{...})^2$	(J-1)(n-1)	SCR2 / (J-1)(n-1)	-
Interaction A x B	$\sum_{ij} n(Y_{ij.}-Y_{i..}-Y_{.j.} + Y_{...})^2$	(I-1)(J-1)	SCI / (I-1)(J-1)	CMI/C MR3
Résiduelle 3	$\sum_{ijk} (Y_{ijk}-Y_{ij.}-Y_{i.k} - Y_{.jk} + Y_{i..} + Y_{.j.} + Y_{..k} - Y_{...})^2$	(I-1)(J-1)(n-1)	SCR3 / (I-1)(J-1)(n-1)	-

En conclusion, il est préférable de n'utiliser ces dispositifs, avec des restrictions dans la randomisation des traitements dans les blocs, que lorsqu'on a des contraintes fortes de réalisation du fait de pertes de précision qu'ils induisent sur les facteurs en grandes parcelles ou en bandes.

### 3) LES PLANS EN CARRÉ LATIN :

Ce type de dispositif permet un double contrôle de l'hétérogénéité, par opposition au plan en randomisation totale qui ne contrôle pas l'hétérogénéité et au plan en bloc qui permet un simple contrôle de celle-ci. On considère simultanément des blocs dans le sens des lignes et des blocs dans le sens des colonnes.

Dans un carré latin on impose, d'une part, que le nombre de répétitions soit égal au nombre de traitements :

$$n = t \text{ et } N = t^2,$$

d'autre part, que chaque traitement soit placé une fois et une seule dans chaque ligne et dans chaque colonne.

La procédure de randomisation est la suivante : on commence par construire un carré latin arbitraire, puis on randomise les lignes, enfin on randomise les colonnes pour obtenir le dispositif définitif. A titre d'exemple, on considère un carré latin à cinq traitements (notés A, B, C, D et E) :

### **1. Carré latin arbitraire**

	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

On peut constater que les cinq traitements sont présents dans chaque ligne et dans chaque colonne.

### **2. Randomisation des lignes**

	1	2	3	4	5
3	C	D	E	A	B
2	B	C	D	E	A
1	A	B	C	D	E
5	E	A	B	C	D
4	D	E	A	B	C

### **3) Randomisation des colonnes**

	5	2	1	4	3
3	B	D	C	A	E
2	A	C	B	E	D
1	E	B	A	D	C
5	D	A	E	C	B
4	C	E	D	B	A

D'où le dispositif à placer sur le terrain. Il faut ajouter qu'en raison de la nécessité d'estimer la variance résiduelle avec suffisamment de degrés de liberté et de ne pas avoir un dispositif trop lourd sur le terrain (avec trop de parcelles expérimentales), on ne pratique guère le carré latin en dehors d'un nombre de traitements compris entre cinq et huit. Toutefois, on peut envisager de faire des carrés latins avec moins de traitements si l'on répète ceux-ci ; et lorsqu'on a plus de traitements, la seule alternative est de faire des carrés latins incomplets, appelés plans de Youden.

L'analyse statistique d'un carré latin est réalisée au moyen du modèle :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_k + \varepsilon_{ijk}$$

où  $Y_{ijk}$  est la donnée observée dans la ligne  $j$  et la colonne  $k$ , portant le traitement  $i$ ,

$\mu$  est la moyenne générale, estimée par  $Y_{...}$ ,

$\alpha_i$  est l'effet du traitement  $i$ , estimé par  $Y_{i..} - Y_{...}$ ,

$\beta_j$  est l'effet de la ligne  $j$ , estimé par  $Y_{.j.} - Y_{...}$ ,

$\delta_k$  est l'effet de la colonne  $k$ , estimé par  $Y_{..k} - Y_{...}$ ,

et  $\varepsilon_{ijk}$  est l'erreur résiduelle.

D'où le tableau d'analyse de variance :



Variation	SCE	dl	CM	F
Traitement	$\sum_i t(Y_{i..} - Y_{...})^2$	t-1	SCT/(t-1)	CMT/ CMR
Ligne	$\sum_j t(Y_{.j.} - Y_{...})^2$	t-1	SCL/(t-1)	CML/ CMR
Colonne	$\sum_k t(Y_{..k} - Y_{...})^2$	t-1	SCC/(t-1)	CMC/ CMR
Résiduelle	$\sum_{ijk}(Y_{ijk} - Y_{i..} - Y_{.j.} - Y_{..k} + 2Y_{...})^2$	(t-1)(t-2)	SCR/(t-1)(t-2)	

Pour conclure, il est important de noter que ce type de dispositif (lignes et colonnes) n'est pas seulement utilisable dans le cas rare où il y a deux gradients d'hétérogénéité perpendiculaires, mais aussi dans le cas où la direction du gradient n'est pas connue avec précision et qu'elle a toute chance de traverser le champ d'expérience en diagonale. De plus, on constate que ce type de dispositif est souvent plus efficace que les blocs ou la randomisation totale pour contrôler une hétérogénéité en taches.

Enfin, il faut remarquer que les facteurs ligne et colonne ne sont pas forcément des positions sur un champ d'expérience. Ils peuvent correspondre à n'importe quel facteur de variation parasite des unités expérimentales.

#### **4) LES AUTRES PLANS D'EXPÉRIENCE :**

Seuls des dispositifs analysables par le logiciel STATITCF ont été vus jusqu'ici. Il en existe bien d'autres, et les présenter sort du cadre de cet exposé. D'autant que beaucoup d'ouvrages ont été consacrés aux plans d'expérience.

Toutefois, un panorama succinct des dispositifs pour lesquels le logiciel STATITCF ne propose pas d'analyse statistique s'impose, alors l'utilisation de logiciels plus performants est nécessaire.

On peut distinguer : les plans en blocs incomplets, les plans en lignes et colonnes incomplètes, les plans factoriels avec "confounding" (confusion d'effets ou d'interactions avec des facteurs de contrôle), etc.

Les plans en blocs incomplets sont des dispositifs en blocs où tous les traitements ne sont pas présents dans chaque bloc. En expérimentation, on s'intéresse plus particulièrement aux plans en blocs Binaires Incomplets Equilibrés (BIE) et aux plans en blocs Binaires Incomplets Partiellement Equilibrés (BIPE).

Les BIE et BIPE sont des plans connexes (on ne développera pas cette notion un peu complexe) où les traitements sont ou bien absents ou bien présents une seule fois dans chaque bloc (Binaires et Incomplets). Le nombre de parcelles par bloc est constant et égal à  $k$  et le nombre de répétitions par traitement dans l'essai est constant et égal à  $r$ , donc si on appelle  $b$  le nombre de blocs :

$$k < t \text{ et } N = bk = tr$$

Un BIE est un plan dans lequel chaque paire de traitements apparaît un même nombre de fois égal à  $\lambda$  (plan équilibré)

$$\lambda = r(k-1)/(t-1)$$

Un BIPE est un plan dans lequel les nombres d'occurrences des paires de traitements ne sont pas constants (plan partiellement équilibré).

exemple de BIE,  $t = b = 7$   
et  $r = k = 3$

Bloc 1	1	2	4
" 2	2	3	5
" 3	3	4	6
" 4	4	5	7
" 5	1	5	6
" 6	2	6	7
" 7	1	3	7

exemple de BIPE,  $t = b = 6$   
et  $r = k = 3$  :

Bloc 1	1	2	4
2	2	3	5
3	3	4	6
4	1	4	5
5	2	5	6
6	1	3	6

Dans le BIE (7 traitements), chaque paire de traitements apparaît dans un seul bloc. Dans le BIPE (6 traitements), la paire (1 ; 2) apparaît dans un bloc et la paire (1 ; 4) apparaît dans deux blocs, par exemple.

Pour la mise en place de tels plans, il faut d'une part randomiser les blocs, d'autre part randomiser les traitements à l'intérieur des blocs et enfin attribuer un numéro au hasard à chaque traitement.

Le plus souvent, on utilise des dispositifs dont les blocs peuvent être regroupés en répétitions complètes où chaque traitement figure une fois et une seule. On parle de lattices lorsque le nombre de traitements  $t = k^2$ , on parle d' $\alpha$ -plans lorsque  $t = kp$  ( $p$  étant un entier quelconque supérieur à 1). Les  $\alpha$ -plans sont donc des lattices généralisés aux cas où  $t$  n'est pas un carré. La randomisation des blocs pour ces dispositifs se fait par répétition complète.

Les plans en lignes et colonnes incomplètes sont des dispositifs avec un double contrôle de l'hétérogénéité. Les plus utilisés sont les carrés latins incomplets (ou plans de YOUDEN), qui sont des carrés latins auxquels on ôte des lignes ou des colonnes, et les lattices carrés, qui sont des plans en lignes et colonnes avec des répétitions complètes et avec  $t = k^2$ .

Enfin, les plans avec "confounding" sont des dispositifs dans lesquels certains effets, ou plus souvent certaines interactions, sont confondus avec les blocs. Il en existe une grande variété, et ils sont décrits dans les ouvrages sur les plans d'expérience. Un des plus utilisés comme essai d'orientation de fertilisation NPK est le  $3^3$  dont deux degrés de liberté de l'interaction entre les 3 facteurs sont confondus avec les blocs (3 niveaux de N, 3 niveaux de P et 3 niveaux de K : en tout 27 traitements) : les traitements sont disposés en 3 groupes de 9 de telle manière que ce soit l'interaction d'ordre 3 qui soit en partie confondue.

A part ce qui vient d'être vu, il existe encore les plans factoriels à répétitions fractionnaires, les plans associés à des surfaces de réponses, etc. Il n'entre pas dans notre objet de les décrire ici.

## CHAPITRE 3

### LES REGROUPEMENTS D'ESSAIS

La caractéristique essentielle d'un essai est de créer des conditions contrôlées. En effet, pour comparer des traitements, il faut qu'ils soient comparables, donc il faut se rapprocher de l'idéal : "toutes choses sont égales par ailleurs". La démarche de l'expérimentation est donc rigoureusement analytique.

La conséquence est qu'un résultat d'essai est relatif à une situation particulière : celle créée par les conditions de l'expérience (par exemple, le précédent, les techniques culturales, les conditions climatiques, etc.).

Mais le problème est de passer d'un résultat de recherche (analytique) à une innovation vulgarisable en milieu paysan, dans une région qui peut être étendue. Il est assez naturel de multiplier le même essai dans des conditions fort diverses, et de voir si l'on a ou non le même résultat. On pratique ce que l'on appelle des essais multilocaux ou pluriannuels. Par essais pluriannuels, on entend des essais de même protocole expérimental, réalisés lors d'années différentes, à des emplacements différents et avec des randomisations indépendantes. Ceci exclut les essais pérennisés, c'est à dire conduits pendant plusieurs années sur les mêmes parcelles avec les mêmes traitements. L'analyse statistique des différentes unités d'un essai pérennisé relève de la méthodologie des mesures répétées et non de celle des regroupements d'essais.

A un essai correspond l'ensemble de ses conditions de réalisation. Cet ensemble de conditions est appelé la situation de l'essai. Dans un regroupement d'essais, on cherche à obtenir une réponse des écarts entre traitements à des situations très diverses. Ceci revient à étudier l'interaction traitement x essai. Mais deux options sont possibles.

On peut considérer que les essais regroupés forment un échantillon représentatif d'un vaste champ d'application (par exemple une région). Ils sont alors le résultat d'un tirage aléatoire dans un ensemble de situations, et on considère l'effet essai et l'interaction traitement x essai comme des effets aléatoires. L'objectif est ici de tester les écarts moyens entre traitements sur l'ensemble des situations par rapport à une base de référence qui est l'interaction traitement x essai. On parle de régionalisation de la réponse.

L'hypothèse qui facilite les calculs dans ce cas est de considérer que les moyennes des traitements dans chaque essai sont estimées avec la même précision. Donc, on suppose l'égalité des variances résiduelles divisées par le nombre de répétitions par traitement. C'est l'option qui est prise par le logiciel STATITCF pour faire l'analyse de variance du regroupement.

L'autre option est de considérer que les essais regroupés forment un ensemble de situations ayant certaines caractéristiques (géographiques, pédologiques, culturelles, etc.). On suppose alors que les résultats peuvent être généralisés aux situations ayant des caractéristiques proches. L'effet essai et l'interaction traitement-essai sont pris comme des effets fixes. L'objectif est d'étudier l'interaction en tant que telle, en essayant de trouver les solutions adaptées à chaque situation. Tous les tests doivent alors être faits en prenant comme base de référence la variance résiduelle des essais. On parle de structuration de l'interaction (au sens large).

L'hypothèse habituellement faite dans ce cas (mais pas par STATITCF) est de considérer que les données parcellaires sont obtenues avec la même précision. Donc on suppose l'égalité des variances résiduelles. Cette hypothèse n'est identique à la précédente que si le nombre de répétitions par traitement est constant.

On peut remarquer que le test de l'interaction traitement x essai au sens le plus général nécessite de faire des répétitions dans chaque essai.

Les séries d'essais peuvent être réalisées en milieu contrôlé : les essais se déroulent en station ou sur points d'observation, leur mise en place et leur suivi sont assurés ou contrôlés par les chercheurs. Ou bien elles peuvent être réalisées en milieu paysan (ou milieu réel, ou milieu semi-contrôlé) : en général le chercheur assure ou contrôle la mise en place, l'application des traitements et les observations (dont la pesée de la récolte) ; le paysan assure le reste, à savoir la conduite de la culture selon ses propres techniques.

La justification essentielle des essais en milieu paysan est que la station, et plus généralement le milieu contrôlé, ne peut simuler ni toutes les contraintes ni tous les desiderata des paysans. Or ils peuvent s'exprimer dans des essais en milieu paysan.

Les essais que l'on prévoit de regrouper doivent avoir le même protocole, avec entre autres :

- les mêmes traitements (au minimum la plupart des traitements en commun),

- le même type de dispositif (randomisation totale, blocs complets, ou autres; tout en sachant que les regroupements de split-plots ou de criss-cross posent des problèmes avec STATITCF),

- la même surface parcellaire,
- si possible le même nombre de répétitions,
- mais des randomisations indépendantes.

Le problème de la détermination du nombre d'essais et de répétitions par essai est abordé par GOZÉ (1992), dans le cadre de la régionalisation de la réponse, où l'effet essai est aléatoire.

La procédure d'analyse se fait en trois étapes, avec des variantes selon l'option prise.

Etape 1 : analyse des essais individuels

Etape 2 : sélection des essais de même variance résiduelle ou bien de même rapport variance résiduelle sur nombre de répétitions selon ce qui est choisi. Si les nombres de degrés de liberté pour estimer les variances résiduelles sont identiques, on peut appliquer le test de HARTLEY (facile à faire):

$$\text{test de } s^2 \text{ max} / s^2 \text{ min}$$

Sinon le test de BARTLETT (plus puissant, mais plus lourd à calculer) peut être appliqué. Ces deux tests ne sont pas programmés sur STATITCF et doivent être faits à la main, avec un tableur, ou programmés spécifiquement pour cela.

Si les variances résiduelles (resp. les rapports variance sur nombre de répétitions) ne sont pas égales, il est fréquent est de constituer des groupes d'essais de même précision en écartant les essais à trop forte variance résiduelle (resp. rapport variance sur nombre de répétitions). Mais ce procédé n'est pas très satisfaisant à cause de son manque de sélectivité. Il est préférable de pondérer le résultat de chaque essai en fonction de sa variance résiduelle, mais il faut alors que celle-ci soit estimée avec 15 degrés de liberté au moins par essai. Cette analyse, non disponible dans STATITCF, sort du cadre de cet exposé. Elle est décrite dans Cochran & Cox (1957).

Etape 3 : analyse de variance du regroupement.

L'analyse différera selon l'option choisie. Si l'effet essai et l'interaction traitement x essai sont supposés aléatoires et si les essais regroupés sont des plans en blocs complets, l'analyse statistique est réalisée au moyen du modèle :

$$Y_{ijk} = \mu + A_i + \beta_j + C_{ij} + D_{ik} + E_{ijk}$$

où  $Y_{ijk}$  est la donnée observée dans l'essai  $i$ , sur le bloc  $k$  et pour le traitement  $j$ ,  
 $\mu$  est la moyenne générale,

$A_i$  est l'effet aléatoire de l'essai  $i$ ,

$\beta_j$  est l'effet du traitement  $j$ ,

$C_{ij}$  est l'interaction essai  $i$ , traitement  $j$  (aléatoire),

$D_{ik}$  est l'effet bloc  $k$  de l'essai  $i$  (aléatoire),

et  $E_{ijk}$  est l'erreur résiduelle dans l'essai  $i$ , sur le bloc  $k$ , pour le traitement  $j$  (de variance  $n_i \sigma^2$ ).

Soit  $n_i$  le nombre de répétitions dans l'essai  $i$ ,  $I$  le nombre d'essais,  $t$  le nombre de traitements. On définit :

$$n_0 = \sum_i n_i$$

$$Y_{i..} = \sum_j Y_{ij.}/t$$

$$Y_{0j.} = \sum_i Y_{ij.}/I \text{ (estime } \mu + \beta_j)$$

$$Y_{0..} = \sum_{ij} Y_{ij.}/tI \text{ (estime } \mu)$$

Il faut noter que les  $Y_{0j.}$  et  $Y_{0..}$  ne sont pas égaux aux  $Y_{.j.}$  et  $Y_{...}$  si les nombres de répétitions par traitement diffèrent d'un essai à l'autre.

Alors le tableau d'analyse de variance s'écrit :

Variation	SCE	dl	CM	F
Traitement	$\sum_j I(Y_{0j\cdot} - Y_{0\cdot\cdot})^2$	t-1	SCT/(t-1)	CMT/C MI
Essai	$\sum_i t (Y_{i\cdot\cdot} - Y_{0\cdot\cdot})^2$	I-1	SCe/(I-1)	CMe/C MI
Interaction T x e	$\sum_{ij} (Y_{ij\cdot} - Y_{i\cdot\cdot} - Y_{0j\cdot} + Y_{0\cdot\cdot})^2$	(t-1) (I-1)	SCI/(t-1)(I-1)	CMI/C MR
Résiduelle	$\sum_{ijk} (Y_{ijk} - Y_{i\cdot k} - Y_{ij\cdot} + Y_{i\cdot\cdot})^2 / n_i$	(t-1)(n <sub>0</sub> -I)	SCR/(t-1)(n <sub>0</sub> -I)	-

La résiduelle est appelée aussi résiduelle pondérée. On teste l'interaction par rapport à cette résiduelle pondérée et l'effet traitement par rapport à l'interaction traitement-essai. D'où un classement global des traitements.

Si l'effet essai et l'interaction traitement-essai sont supposés fixes (étude de l'interaction en tant que telle), et si les essais regroupés sont des plans en blocs complets, l'analyse statistique est réalisée au moyen du modèle :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \delta_{ik} + \varepsilon_{ijk}$$

où  $Y_{ijk}$  est la donnée observée dans l'essai  $i$ , sur le bloc  $k$  et pour le traitement  $j$ ,

$\mu$  est la moyenne générale,

$\alpha_i$  est l'effet de l'essai  $i$ ,

$\beta_j$  est l'effet du traitement  $j$ ,

$(\alpha\beta)_{ij}$  est l'interaction essai  $i$ , traitement  $j$ ,

$\delta_{ik}$  est l'effet du bloc  $k$  dans l'essai  $i$ ,

et  $\varepsilon_{ijk}$  est l'erreur résiduelle (de variance  $\sigma^2$ ).



Le tableau d'analyse de variance s'écrit :

Variation	SCE	dl	CM	F
Traitement	$\sum_j n_0 (Y_{.j} - Y_{...})^2$	t-1	SCT/(t-1)	CMT/CMR
Essai	$\sum_i t n_i (Y_{i.} - Y_{...})^2$	I-1	SCe/(I-1)	CMe/CMR
Interaction t x e	$\sum_{ij} n_i (Y_{ij.} - Y_{i.} - Y_{.j} + Y_{...})^2$	(t-1)(I-1)	SCI/(t-1)(I-1)	CMI/CMR
Bloc	$\sum_{ik} t (Y_{i.k} - Y_{i.})^2$	$n_0 - I$	SCB/ ( $n_0 - I$ )	CMB/CMR
Résiduelle	$\sum_{ijk} (Y_{ijk} - Y_{i.k} - Y_{ij.} + Y_{i.})^2$	(t-1)( $n_0 - I$ )	SCR/(t-1)( $n_0 - I$ )	

Tout doit alors être testé par rapport à la variance résiduelle, car tous les effets sont fixes.

La première chose à voir dans ce tableau est le test de l'interaction traitement x essai. Si cette interaction est significative, l'analyse se poursuit pour essayer de la décrire ou de l'expliquer (structuration de l'interaction : au sens large). Le logiciel STATITCF propose diverses méthodes d'études de l'interaction, mais en supposant que les erreurs résiduelles  $\varepsilon_{ijk}$  sont de variance  $n_i \sigma^2$  (au lieu de  $\sigma^2$ ). Ceci revient à supposer que les moyennes par traitement et par essai ont la même précision dans tous les essais. Cette hypothèse est identique à celle du modèle ci-dessus si le nombre de répétitions est le même dans tous les essais.

Il existe un grand nombre de méthodes d'étude de l'interaction. On va passer en revue un certain nombre d'entre elles en considérant trois types de problèmes :

- On cherche à constituer des groupes d'essais homogènes. On peut utiliser:
  - des méthodes graphiques simples si le nombre d'essais est faible (cf. représentation de l'interaction au chapitre 2),
  - une analyse en composantes principales (ACP) non normée sur le tableau des interactions (essais en individus, traitements en variables, et à la croisée d'un individu i et d'une variable j la valeur de l'interaction  $Y_{ij.} - Y_{i.} - Y_{.j} + Y_{...}$ ),
  - ou une classification automatique sur le même tableau.

2. On cherche à constituer simultanément des groupes d'essais et des groupes de traitements. On peut utiliser :

- la structuration visuelle (matrice de BERTIN),
- la méthode de CARAUX, qui est une manière automatique de faire la structuration ci-dessus,
- la classification automatique sur les lignes (essais) et sur les colonnes (traitements).

Il faut noter que ces trois options sont proposées par STATITCF dans le module "structuration de l'interaction" (au sens strict).

3. On cherche à modéliser l'interaction. On peut utiliser :

- la régression factorielle qui consiste à prendre en compte des covariables explicatives des essais et/ou des traitements,
- la régression conjointe ou modèle d'étude de la stabilité du rendement dans les essais variétaux. Ceci consiste à prendre comme covariable la moyenne par essai des variétés communes à tous les essais.

La régression factorielle est disponible sur STATITCF, contrairement à la régression conjointe.

## CHAPITRE 4

### ANALYSE EN COMPOSANTES PRINCIPALES

L'analyse en composantes principales (ACP) est une des méthodes de l'analyse des données. Celles-ci ont été jusqu'à présent peu utilisées dans le domaine de l'expérimentation.

Pourtant ces méthodes, et plus spécifiquement l'ACP, sont intéressantes ici car elles sont d'une grande utilité pour analyser l'interaction traitement x essai dans des regroupements d'essais.

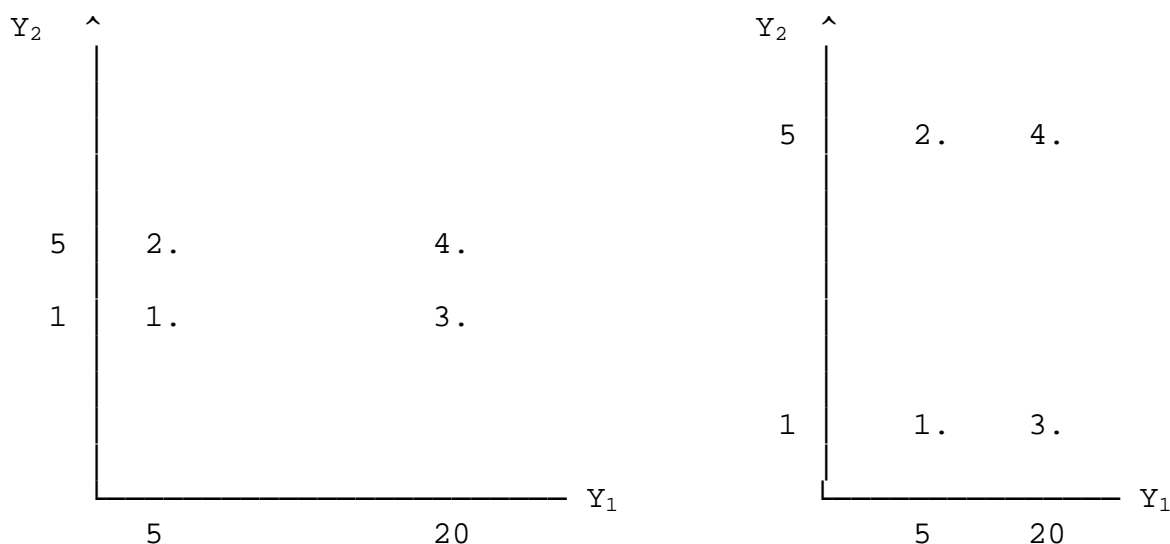
Les buts de l'analyse des données sont de représenter sous forme de graphiques ou de classer des données contenues dans des tableaux individus x variables, de manière à en résumer l'information.

L'ACP sert à représenter sous forme de graphiques des données provenant de tableaux de variables quantitatives, c'est-à-dire continues, entières ou polytomiques ordonnées. Les variables continues sont des variables de mensurations comme la hauteur, le rendement, etc. Les variables entières (ou discrètes) sont le plus souvent des variables de comptage : le nombre de plants par parcelle, le nombre de plants versés, etc. Les variables polytomiques ordonnées sont souvent des notations : l'intensité de l'attaque d'un champignon sur une culture peut être notée de 1 (culture saine) à 9 (très attaquée, voire détruite).

On considère donc un tableau de  $p$  variables quantitatives et  $n$  individus. On note  $Y_{ij}$  la valeur de la variable  $j$  pour l'individu  $i$ .

Cas de deux variables ( $p = 2$ ) : la représentation de ces données est très simple. On dresse un graphique (de nuage de points). Il sert d'une part à évaluer la liaison entre les deux variables, d'autre part à visualiser la répartition des individus (pour peu qu'on les identifie).

Se pose alors le problème de l'échelle des deux variables. Si l'on choisit de représenter l'étendue de  $Y_1$  (ou l'écart-type) sur une plus grande largeur que l'étendue de  $Y_2$  (ou l'écart-type), le nuage n'aura pas la même forme que si l'on choisit l'inverse : les distances entre individus changent.



Les proximités entre individus peuvent ainsi être fortement modifiées par le choix de l'échelle de représentation. On retrouve ce problème dans les notions d'ACP normée et non normée.

Cas de plus de deux variables (p>2) : la représentation directe n'est alors plus possible (sauf en dimension 3, en perspective). Mais on cherche la même chose que dans un graphique : d'une part à évaluer les liaisons entre variables, d'autre part à visualiser la répartition des individus.

L'idée de l'ACP est de chercher des projections du nuage de points sur des axes, de telle sorte que ces projections s'accompagnent d'un minimum de perte d'information. C'est-à-dire que les distances entre individus (facilement calculables pour toute paire d'individus à partir du tableau de données) doivent être restituées le mieux possible.

Mais il faut avant tout choisir la définition d'une distance entre les individus. On offre souvent deux options :

1. Le centrage et la réduction (ACP normée) qui donne le même poids à chaque variable, la distance entre deux individus est donnée par :

$$d^2(i,i') = \sum_{j=1}^p \frac{(Y_{ij} - Y_{i'j})^2}{\sigma_j^2}$$

où  $d(i,i')$  est la distance entre  $i$  et  $i'$  et  $\sigma_j^2 =$  variance de  $j$ .

2. Le centrage (ACP non normée) qui donne à chaque variable un poids égal à sa variance, la distance entre deux individus est donnée par :

$$d^2(i, i') = \sum_{j=1}^p (Y_{ij} - Y_{i'j})^2$$

Une fois une distance définie, ce qui correspond à une forme de représentation des individus dans un espace de dimension  $p$ , on cherche à les représenter dans un espace plus petit et visualisable sur un ou plusieurs plans. On commence par chercher à calculer la combinaison linéaire des variables de départ formant un axe qui permet de retenir le plus d'information sur les individus. On prend l'axe pour lequel les coordonnées des individus sont de variance maximale. Cet axe est appelé le 1er axe principal.

Une certaine proportion de la variation totale des individus (inertie totale du nuage de points) est expliquée par cette première composante principale. La variance des coordonnées des individus sur cet axe (inertie portée par l'axe) est la part de variance expliquée.

Puis on calcule la 2e composante principale en cherchant parmi l'ensemble des combinaisons linéaires orthogonales à la première celle qui a la plus grande variance. Ainsi le graphique (de nuage de points) formé par le 1er et le 2e axe principal forment le plan qui représente le mieux la dispersion des individus.

Le processus se déroule ainsi jusqu'à l'obtention de la  $p^e$  et dernière composante principale qui porte la plus faible variance.

Ainsi l'ACP consiste à transformer les  $p$  variables quantitatives initiales, plus ou moins corrélées, en  $p$  nouvelles variables quantitatives, non corrélées entre elles, appelées composantes principales.

Les résultats de l'ACP se présentent de la manière suivante :

1. calcul de statistiques élémentaires : moyenne et écart-type de chaque variable, ainsi que la matrice des corrélations,

2. calcul des composantes principales : valeurs propres et vecteurs propres (ces notions sont vues plus loin),

3. étude des variables au moyen des cercles de corrélation,

4. étude des individus à partir des coordonnées sur les axes principaux.

L'interprétation des statistiques élémentaires est un élément important de l'ACP. Les moyennes et les écarts-type sont à comparer aux références que l'on peut connaître. Et il ne faut pas oublier que, sur le plan des variables, l'ACP est avant tout l'étude de la matrice des corrélations (ou des variances-covariances).

Le calcul des composantes principales est fait par une technique mathématique appelée la diagonalisation. Il n'entre pas dans notre objet ici de détailler la justification de cette technique et les opérations à effectuer.

Il faut retenir que lors d'une ACP normée (données centrées et réduites), on diagonalise la matrice des corrélations, et lors d'une ACP non normée (données simplement centrées) on diagonalise la matrice des variances-covariances.

La diagonalisation fournit une valeur propre par axe principal et un vecteur propre associé. Les valeurs propres sont les variances des coordonnées des individus sur les axes principaux. Chaque valeur propre représente la part de l'inertie du nuage portée par l'axe correspondant.

En ACP normée la somme des valeurs propres est égale à  $p$  (le nombre de variables), et en ACP non normée cette somme est égale à la somme des variances des  $p$  variables.

Chaque vecteur propre est constitué des coefficients à affecter aux variables initiales pour calculer l'axe principal correspondant. Mais il ne faut pas oublier que ces variables initiales sont les variables centrées et réduites en ACP normée et les variables centrées en ACP non normée.

L'étude des variables est faite en dressant les cercles de corrélation. Ils servent à analyser les corrélations entre variables de départ et à interpréter les axes principaux.

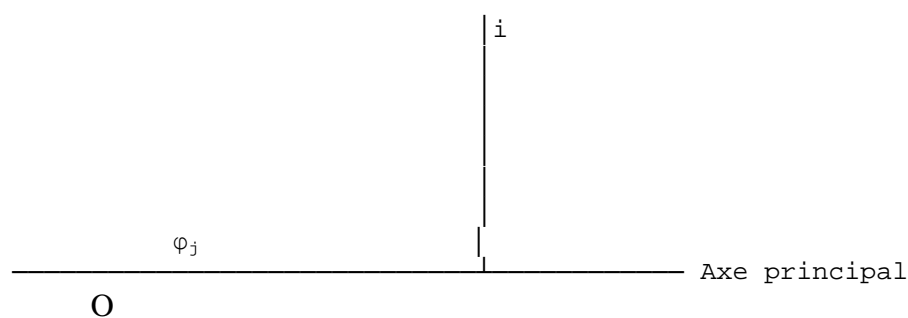
On attribue à chaque variable des coordonnées égales à ses corrélations avec les axes principaux. Les variables se placent sur une sphère de dimension  $p$  et de rayon 1. Les cercles de corrélation sont donc les projections des variables sur les plans considérés.

On peut retenir les règles d'interprétation suivantes :

1. On ne peut interpréter que les variables proches du cercle, donc bien représentées sur le plan considéré,
2. Ce sont les directions des variables par rapport aux axes qui sont à interpréter,
3. Deux variables dont les directions forment un angle faible sont corrélées positivement,
4. Deux variables qui sont diamétralement opposées sont corrélées négativement,
5. Deux variables dont les directions sont perpendiculaires sont non corrélées linéairement.

L'étude des individus est faite par les représentations graphiques dans les plans principaux choisis. On analyse la répartition des individus, en se demandant quels sont ceux qui se ressemblent, et quels sont ceux qui diffèrent.

On se fait une idée de la qualité de la représentation d'un individu en analysant les cosinus carrés des angles que font les axes principaux avec la direction de l'individu par rapport à l'origine :



Les  $\text{Cos}^2 \varphi_j$  permettent d'évaluer la qualité de la représentation de l'individu  $i$ .

\* Remarque :

L'ACP permet d'introduire également des variables et des individus supplémentaires qui ne participent pas à l'analyse, mais qui sont représentés en projection sur les axes formés par les variables et les individus actifs.

## BIBLIOGRAPHIE

### 1) Expérimentation et analyse de variance.

---

- \*\* COCHRAN W. G., COX G. M. Experimental designs, 2nd edition. Wiley, 1957.
- \*\*\* COTTRELL M., COURSOL J. La planification des expériences. Economica, 1989.
- \*\*\* COURSOL J. Technique statistique des modèles linéaires, 1. Aspects théoriques. Les cours du CIMPA, 1980.
- \*\* DAGNELIE P. Théorie et méthodes statistiques (vol. 1 et 2). Presses agronomiques de Gembloux, 1970.
- \* DAGNELIE P. Principes d'expérimentation. Presses agronomiques de Gembloux, 1981.
- \* DYKE G. V. Comparative experiments with field crops, 2nd edition. Charles Griffin, 1988.
- \*\*\* FEDERER W. T. Experimental design, Theory and Application. The Macmillan Company, 1955.
- \*\* GOZÉ E. Détermination de la dimension des réseaux d'essais. Coton et Fibres Tropicales Vol 47 (1992) n° 2: pp 81-94.
- \* GOUET J.-P. , PHILIPPEAU G. Comment interpréter les résultats d'une analyse de variance ? ITCF, 1986.
- \*\*\* KEMPTHORNE O. The design and analysis of experiments. Réédition corrigée, Wiley, 1975.
- \* LECOMPT M. L'expérimentation et les engrais. SPIEA, 1965.



- \*\* MEAD R., CURNOW R. N. Statistical methods in agriculture and experimental biology. Chapman and Hall, 1983.
- \*\* PEARCE S.C., CLARKE G.M., DYKE G.V., KEMPSON R.E. Manual of crop experimentation. Griffin, 1988.
- \*\* PHILIPPEAU G. Théorie des plans d'expérience. ITCF, 1977.
- \*\*\* SCHEFFE H. The analysis of variance. Wiley, 1959.
- \*\* SNEDECOR G.W., COCHRAN W. G. Méthodes Statistiques ACTA, 1974
- \*\* TOMASSONE R., AUDRAIN S., LESQUOY-DE TURCKHEIM E., MILLIER C. (1992). La régression, nouveaux regards sur une ancienne méthode statistique , 2e édition, Masson.

## 2/ Méthodes non-paramétriques

---

- \*\*\* CONOVER W.J. Practical nonparametric statistics, 2nd edition. Wiley, 1980.
- \* SIEGEL S. Non parametric statistics for behavioural sciences. Mc GrawHill, 1956.

## 3) Enquêtes, échantillonnage.

---

- \*\*\* COCHRAN W. G. Sampling techniques, 3rd edition. Wiley, 1977.
- \*\* DEROO M. , DUSSAIX A.-M. Pratique et analyse des enquêtes par sondage. PUF, 1980.
- \* GOUET J.-P. L'élaboration d'un protocole d'enquête. Proposition d'un plan-type détaillé et quelques commentaires. ITCF, 1978.

## 4) Analyse des données.

---

- \*\* BOUROCHE, J. M. and SAPORTA, G. L'analyse des données, P.U.F., 1980.
- \*\*\* CAILLIEZ F., PAGES J.-P. Introduction à l'analyse des données. SMASH, 1976.
- \* CIBOIS, P. L'analyse factorielle. P.U.F., 1987
- \* DERVIN C. Comment interpréter les résultats d'une analyse factorielle des correspondances ? ITCF, 1988.
- \* FENELON J.-P. Qu'est-ce que l'analyse des données ? Lefonen, 1981.
- \* PHILIPPEAU G. Comment interpréter les résultats d'une analyse en composantes principales ? ITCF, 1986.
- \*\* SAPORTA, G. Probabilités analyse des données et statistique. Technip, 1990

## 5) Logiciel statistique

---

- \*\* SAS Institute . SAS/STAT user's guide, version 6, 4<sup>th</sup> edition (2 volumes). SAS, 1989
- \* STATITCF. Manuel d'utilisation, 5e version. ITCF, 1991

### Légende:

- \* lecture facile, présentation intuitive,
- \*\* ouvrage nécessitant de bonnes bases mathématiques,
- \*\*\* ouvrage de base difficile, très mathématique.