

CN0100911

(m. I. D / d. D)

DOCUMENT N. 22 / 83

MARS 83

*LA RIGUEUR DU DISPOSITIF EXPERIMENTAL
DANS L'EXPLOITATION STATISTIQUE DES RESULTATS*

par

m. I. DIEDHIU

biometriste ISRA - CNRA BAMBEY

DEUXIEME REUNION DES COMITES SCIENTIFIQUES NATIONAUX
DU PROJET REGIONAL D'AMELIORATION
DES MIL SORGHO MAIS ET NIEBE

BANJUL - 7-10 MARS 1983

I - INTRODUCTION :

Le but du présent article est de présenter l'importance du dispositif expérimental dans la conception, la réalisation et l'interprétation des résultats.

Tout n'a pu être abordé. Nous ne parlerons pas de l'analyse de la variance qui est la base de la planification expérimentale, Nous rappellerons au passage les conditions de validité de celle-ci, Notre attention se portera sur la façon de faire une planification, ,

II - HISTORIQUE DE L'EXPERIMENTATION AGRICOLE :

L'analyse de la variance et la planification des expériences ne sont pas de nouvelles techniques,

L'outil statistique a été introduit dans la planification expérimentale au 20^e siècle par les travaux de Ronald Aylmer Fisher (1890-1962) dans le cadre de la station de Recherches de Rothamsted (Rothamsted Experimental Station, Harpenden, Hertfordshire en Grande Bretagne). Ses travaux remontent à 1930.

Déjà dès le 17^e siècle, les principes de base furent énoncés par Francis BACON (1561-1626) et René Descartes (1596-1650).

Aujourd'hui, la statistique connaît un développement considérable grâce à l'évènement de l'informatique.

III - CRITERES DE LEGITIMITE D'UNE ANALYSE DE LA VARIANCE (ANOVA) :

L'analyse de la variance ne fournit des conclusions valables que si :

- la caractéristique numérique étudiée a une distribution normale ;
- les variances des diverses populations sont égales (homoscédasticité) ;
- les échantillons sont prélevés de façon aléatoire et indépendante dans chacune des populations,

NE pas vérifier ces conditions, risque de conduire à de graves erreurs d'interprétation,

IV - COMMENT MENER UN PLAN?

Pour mener un plan, il faut définir très précisément le but, les diverses questions auxquelles l'essai doit répondre, car l'organisation de l'expérience et la collecte des données peuvent être bien différentes selon les problèmes posés.

L'élaboration d'un dispositif expérimental est conditionnée par les qualités que l'on exige d'une "bonne" expérience ; ces qualités sont :

- l'absence d'erreur,
- la généralisation des résultats,
- la simplicité dans la réalisation,
- la précision des comparaisons à effectuer,
- la possibilité d'analyser les résultats.

Pour éliminer tout risque d'erreur systématique on fait recours à la randomisation ou affectation au hasard. Elle consiste à déterminer par des tirages au sort quelles pièces seront affectées à un traitement donné. La randomisation donne à n'importe quel couple de traitements une chance égale, d'être appliquée à deux unités adjacrites, si bien que les effets de corrélation ont tendance à se compenser lorsque le nombre de répétitions est assez grand. L'analyse des résultats supposant l'indépendance des erreurs, peut conduire à des évaluations biaisées si les erreurs sont corrélées. C'est pour réaliser cette hypothèse d'indépendance que l'on fait appel à la randomisation.

Pour ce qui est de la généralisation des résultats, il est bon de savoir si les facteurs sont fixes ou aléatoires ; si les facteurs sont considérés comme fixes, les conclusions ne concernent que le cas étudié ; dans le cas aléatoire, la conclusion est souvent générale.

La précision peut être atteinte en augmentant le nombre d'unités expérimentales. Cette augmentation n'est pas sans inconvénient. On peut aussi utiliser des unités expérimentales aussi homogènes que possible.

L'utilisation des méthodes statistiques ne sera possible que s'il existe un nombre suffisant de données pour caractériser les populations étudiées.

Les répétitions, l'affectation au hasard ou randomisation, rendent un test de signification valable. Le contrôle local (par exemple la réalisation des blocs, le groupement ou l'équilibrage des unités expérimentales), rend le test plus sensible ou plus puissant. La puissance d'un test est mesurée par son aptitude à détecter une différence quand il en existe réellement une.

L'affectation au hasard, peut être effectuée au moyen de dés, de cartes numérotées ou en utilisant des tables de nombres au hasard.

4.1. - étapes de la planification :

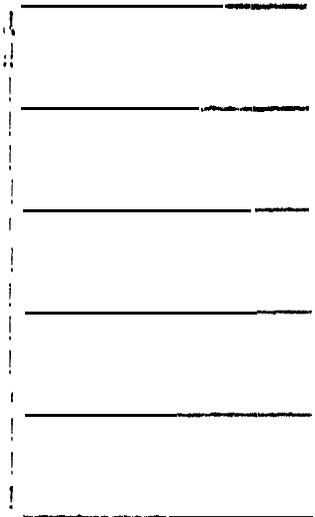
La première étape dans la planification consiste à faire différents choix : choix des objets, choix des unités expérimentales et du nombre de répétitions, choix des observations et de leur mode de collecte et choix du dispositif expérimental, choix des facteurs et de leurs niveaux et détermination de la façon dont ils se présentent l'un par rapport à l'autre (facteurs croisés ou hiérarchisés).

La seconde étape consiste au choix de la forme des parcelles et des bordures. Pour éviter que les interférences, les "contagions" entre parcelles voisines, n'influencent les résultats, on peut limiter l'observation à la partie centrale de chacune des parcelles, en considérant le pourtour comme une simple bordure (cas épandage d'engrais).

4.2 - Forme des unités expérimentales :

Lorsque le terrain est relativement homogène, on a intérêt à adopter une forme de parcelle aussi carrée que possible.

Si le terrain est hétérogène dans une direction donnée, appelée gradient de fertilité, il faut adopter des parcelles rectangulaires, allongées parallèlement à la direction générale de cette hétérogénéité (voir figure).



-----) Direction du gradient de fertilité

V - PRINCIPAUX PLANS LES PLUS UTILISES EN AGRONOMIE -
AVANTAGES ET INCONVENIENTS :

5.1 - Les plans complètement aléatoires : (Completely randomized design)

Dans ce plan, les niveaux des facteurs ou traitements sont répartis de façon aléatoire sur les unités expérimentales. C'est le dispositif le plus simple à réaliser. Le nombre de facteurs et de niveaux n'est pas limité.

Les résultats peuvent être analysés par l'analyse de la variance à un, deux ou trois critères de classification suivant le nombre de facteurs étudiés.

On pourra recourir avant de faire l'ANOVA, aux différents tests de normalité, d'égalité des variances et de transformation des variables suivant les besoins.

Du fait de l'hétérogénéité du matériel expérimental, ces plans sont souvent peu efficaces, peu sensibles et peu puissants.

5.2 - Les plans en blocs aléatoires complets : (Randomized complete block design)

Contrairement au plan complètement aléatoire, il y a contrôle d'un facteur d'hétérogénéité. Le bloc est constitué de parcelles semblables entre elles. Ces blocs sont dits complets lorsque tous les traitements mis en expérience sont présents dans chacun d'eux. Tous les blocs

ont la même taille. Lorsque le terrain est relativement homogène, les blocs comme les parcelles auront une forme carrée. Si le terrain présente un gradient de fertilité marqué dans une direction, les blocs auront une forme allongée perpendiculaire à la direction du gradient.

Le facteur "bloc" est implicitement considéré comme aléatoire. Il y a possibilité de supprimer certains blocs ou même certains traitements en cours de réalisation.

Ce plan a une efficacité supérieure à celle du plan complètement aléatoire. La puissance ou la sensibilité des comparaisons est plus élevée.

Lorsque le nombre de traitements est grand (supérieure à 20 ou 30) on utilise les blocs incomplets.

5.3 - Le carré latin :

Dans ce plan, il y a contrôle de deux facteurs d'hétérogénéité. Ces facteurs sont appelés "Lignes" et "Colonnes". Il est plus efficace que les deux précédents plans. On suppose les interactions négligeables.

5.4 - Le Split-plot :

C'est un plan en parcelles divisées. Le principe du split-plot consiste en une répartition des objets en deux ou plus de deux étapes.

Illustrons le principe du split-plot par ^{un} exemple. Soit l'étude de l'influence de deux modes d'irrigation, (facteur A) et de trois fumures différentes (facteur B) sur le rendement d'une culture ; on dispose de trois blocs.

La disposition expérimentale se fait comme suit :

1-chaque bloc est subdivisé en deux parcelles : l'une recevant l'irrigation 1 et l'autre l'irrigation 2 (la répartition est faite aléatoirement) ;

2-chaque parcelle définie ci-dessus est partagée en trois sous-parcelles et les trois fumures (niveaux du facteur B) sont réparties de façon aléatoire sur les sous-parcelles.

L'analyse des résultats se fait par l'analyse de la variance à trois critères de classification. Les comparaisons relatives au facteur lié aux grandes parcelles seront toujours moins précises que les comparaisons relatives

Il y a perte de précision pour le facteur appliqué aux grandes parcelles et gain de précision pour le deuxième facteur et pour l'interaction des deux facteurs. La perte de précision relative au premier facteur est généralement d'autant plus importante que le nombre de niveaux de ce facteur est petit.

On utilise le split-plot, si on s'intéresse plus particulièrement à un des facteurs qu'à l'autre ou, à la limite, lorsqu'on s'intéresse uniquement à l'un des deux facteurs et à leur interaction; on peut aussi utiliser le split-plot pour introduire en cours d'expérience un facteur supplémentaire non prévu initialement.

On envisagera d'autres plans suivant le problème à étudier.

VI - EFFICACITE RELATIVE DU PLAN EN BLOCS ALEATOIRES COMPLETS PAR RAPPORT AU PLAN COMPLETEMENT ALEATOIRE :

Pour évaluer le gain de précision qu'apporte le plan en blocs aléatoires complets (Randomized complete block design) par rapport au plan complètement aléatoire (completely randomized design), on calcule l'efficacité relative (relative efficiency)

$$R.E = \frac{(r-1)E_b + \frac{\{(t-1) + (r-1)(t-1)\}}{(rt-1)} E_e}{(rt-1)E_e}$$

où r = nombre de blocs

t = nombre de traitements

E_b et E_e sont respectivement le carré moyen du bloc et de l'erreur. Si le nombre de degré de liberté de l'erreur est inférieur à 20, le R.E est ajusté par un coefficient k :

$$k = \frac{\left\{ \frac{(r-1)(t-1) + 1}{t(r-1) + 1} \right\} \left\{ \frac{(r-1)(t-1) + 3}{(r-1)(t-1) + 3} \right\}}{\left\{ \frac{(r-1)(t-1) + 1}{t(r-1) + 1} \right\} \left\{ \frac{(r-1)(t-1) + 3}{(r-1)(t-1) + 3} \right\}}$$

(k) \times (R.E), Efficacité Relative Ajustée

Considérons une expérience dans laquelle on désire comparer sept variétés de sorgho. On dispose de données sur le rendement en grains à l'hectare. Une ANOVA donne le tableau suivant :

Source de variation	Degré de liberté	Somme des Carrés	Carré moyen	F observée	F tabulée	
					5%	1%
Bloc	4	36467857.14	9116964.29	6.52**	2.78	4.22
Traitement	6	120121000	20020166.67	1431**	2.51	3.67
Erreur	24	33581142.86	1399214.29			

** = Hautement significatif

CV (a) = 13.52%

Le degré de liberté ^{de l'erreur} étant de 24, on va calculer l'efficacité relative avec $r = 5$, $t = 7$:

$$R.E = \frac{4 \times 9116964.29 + (6 + 24) \times 1399214.29}{(35 \cdot 1) \times 1399214.29} = 1,65$$

En termes de pourcentage, on obtient 165%, ce qui signifie que le plan en blocs aléatoires complets augmente la précision de 65% par rapport au plan complètement aléatoire.

VII - CONCLUSION :

L'association de la planification et de l'analyse de la variance permettra :

- de donner des réponses non ambiguës lors de l'interprétation des résultats ;
- de s'assurer que les buts fixés à l'expérience ont des chances d'être atteints ;
- de faire une meilleure utilisation des ressources disponibles, en faisant une répartition convenable du nombre d'échantillons dont on dispose entre les différents essais constituant l'expérience, en vue de tirer le maximum d'informations possibles.

En dehors du dispositif expérimental à mettre en oeuvre pour étudier un problème, pour que les objectifs fixés soient atteints avec le plus de précision, il faut une collaboration étroite entre Biométricien et Agronome, entre Statisticien et Expérimentateur. Le Biométricien et l'Agronome doivent avoir des points de vue complémentaires et non contradictoires. C'est pourquoi FERAUD a écrit : "la méthode statistique ne doit pas constituer un domaine à part dans l'expérimentation ; on ne saurait sans de graves inconvénients séparer l'expérimentateur du statisticien" (5).

L'information que l'on peut tirer de l'utilisation de l'outil statistique n'est exploitable que si le dispositif expérimental retenu se conforme à celui établi par le statisticien et l'expérimentateur. Ce dispositif doit être établi en fonction du but poursuivi par l'essai, le choix de l'un d'entre eux sera dicté par les questions que l'on se pose avant de lancer l'expérience.

La collaboration doit être permanente. Elle commence au moment où une idée est retenue, un problème clairement exprimé, ce n'est que dans ce cadre que l'analyse statistique des résultats sera facile et les conclusions sans ambiguïté.

Le biométricien ou le statisticien pourra éviter de commettre "l'erreur de troisième espèce", celle commise en donnant une réponse exacte à un problème faux, c'est-à-dire, à un problème posé qui n'est pas le problème qu'il conviendrait de traiter".

BIBLIOGRAPHIE

- 1 - CHAPOUILLE P. (1973) - Planification et analyse des expériences, MASSON et CIE, 190 pages.
- 2 - DAGNELIE O. (1980) - Principes d'expérimentation. Gembloux, Presses agronomiques, 182 pages.
- 3 - GERARD G. (1980) - Biométrie
syllabus Agro 2110 UCL
- 4 - GOMEZ K. A. and GOMEZ A. A. (1976) - Statistical procedures for agricultural research with emphasis on rice, I.R.R.I, 294 pages.
- 5 - PHILIPPEAU G. (1973) - Théorie des plans d'expérience application à l'agronomie, I.T.C.F, 205 pages.
- 6 - Cahiers de nutrition et de diététique (1977), Vol 12, fascicule 1, pp 37-43.