

CR

REPUBLIQUE DE HAUTE-VOLTA



MINISTERE DU DEVELOPPEMENT RURAL

DIRECTION DES SERVICES AGRICOLES

INSTITUT DE RECHERCHES AGRONOMIQUES
TROPICALES ET DES CULTURES VIVRIERES

I. R. A. T.

CONCEPTION, ORGANISATION ET CONDUITE
DES EXPERIMENTATIONS AGRONOMIQUES
REMARQUES SUR L'INTERPRETATION STATISTIQUE ET AGRONOMIQUE

Par

J. - F. POULAIN

Ingénieur Agronome, Ingénieur de Recherches à l'I.R.A.T.

Directeur de l'I.R.A.T. en Haute-Volta

OUAGADOUGOU - Janvier 1977

Fonds Documentaire

N° : 02680

Cote B

Date : 10 JANV 1977

CONCEPTION, ORGANISATION ET CONDUITE
DES EXPERIMENTATIONS AGRONOMIQUES
REMARQUES SUR L'INTERPRETATION STATISTIQUE ET AGRONOMIQUE

"Elaborate statistics are not substitute for meticulous experimentation"

G.W. SNEDECOR

LIMINAIRE

Le chercheur dispose de trois méthodes de travail :

- l'observation
- l'analyse
- l'expérimentation.

Chacune a ses limites et ses servitudes et l'utilisation simultanée de ces trois méthodes constitue l'idéal pour mener à bien tout travail de recherche. En effet, le chercheur ne s'intéresse pas seulement aux résultats bruts d'un essai, il doit se préoccuper aussi et surtout des mécanismes qui expliquent ces résultats pour mieux en autoriser la généralisation.

C'est dans cette perspective que les deux premières méthodes sont essentielles, mais l'expérimentation agricole qui intègre dans un résultat brut l'ensemble des effets reste toutefois indispensable pour établir les liaisons entre le comportement des cultures et les données visuelles et analytiques. Elle permet également de chiffrer, pour un sol donné qui n'a pas atteint son niveau optimum de fertilité, les variations de "marge brute" sous l'effet de la fumure minérale.

Pour le vulgarisateur, l'expérimentation est le point de départ d'une action de développement et seuls les résultats l'intéressent dans la mesure où ils constituent des bases techniques et économiques indispensables.

La littérature sur l'expérimentation agricole est abondante et notre propos est seulement d'apporter une modeste contribution à la lumière d'une expérience de plus de quinze ans en Afrique de l'Ouest.

Nous nous limiterons à quelques aspects essentiels :

- Conception de l'expérimentation
- Organisation de l'expérimentation
- Conduite de l'expérimentation.

Nous compléterons par quelques remarques sur l'interprétation statistique et agricole.

1.- CONCEPTION DE L'EXPERIMENTATION

1.1.- Analyse du problème

Il est essentiel avant tout de rassembler l'information existante sur le sujet et d'analyser soigneusement les expériences antérieures en isolant les faits valables obtenus par d'autres chercheurs dans des conditions écologiques identiques. Il est en effet regrettable de voir des expérimentations se répéter dans le temps et l'espace sans que les responsables prennent connaissance de la masse des résultats déjà acquis sur le même sujet. Doit-on par exemple reconduire des expérimentations précises sur les phosphates tricalciques dans tel ou tel pays à la suite de la découverte d'un nouveau gisement, alors que celles-ci ont déjà été abondamment réalisées au niveau de l'Afrique de l'Ouest dans des conditions pédoclimatiques identiques ? Certes, une expérimentation simple de contrôle reste indispensable, mais il est parfaitement inutile de reprendre le problème comme si rien n'avait déjà été fait.

Par contre, si le problème est vraiment nouveau, on doit prévoir une expérience d'exploration préliminaire avant d'installer un dispositif complet et coûteux.

1.2.- Définition claire des questions à résoudre par l'expérimentation

Après cet inventaire le plus soigné possible des travaux anciens (et récents) sur le sujet, il importe de planifier correctement la nouvelle expérimentation. Il est indispensable de choisir les idées à vérifier et de bien préciser les problèmes à résoudre. Dans tous les cas les questions devront être réduites ; il est en effet préférable de poser une ou deux questions précises auxquelles on apportera une réponse sûre, que plusieurs (probablement justifiées) qui resteront sans réponse valable.

On établira ensuite la liste des traitements à mettre en comparaison en veillant à ce que seulement les facteurs objet de l'étude varient. Un certain nombre de précautions doivent être prises. Nous en citerons quelques unes.:

- Le témoin absolu reste à notre avis indispensable au stade actuel de nos travaux.

- On évitera d'associer deux facteurs dans le même traitement. Une courbe de réponse à l'azote utilisant du sulfate NH_4 est également une courbe de réponse au soufre.

- On ne se souciera pas dans un essai conduit en station d'utiliser des niveaux de fertilisants élevés, même si à priori ceux-ci sont aberrants du point de vue de la rentabilité (indispensable pour les essais courbe de réponse).

- Dans une comparaison entre phosphate tricalciques et phosphates solubles, ces derniers ont probablement une action supérieure à un niveau faible, mais les deux engrais pourront être équivalents à dose élevée.

.../...

- Il est inutile de vouloir mettre en évidence des différences qui seront probablement inférieures au seuil de signification. Il est de peu d'intérêt de comparer deux engrais phosphatés très solubles à même dose et contenant les mêmes éléments. Leur efficacité est probablement identique et seul l'aspect économique doit rentrer en ligne de compte.

- De même, il est illusoire de comparer deux doses relativement faibles ou à l'opposé très fortes d'engrais de même nature. La précision en matière de recommandation des fumures, est parfaitement illusoire en raison du nombre de facteurs entrant en jeu pour obtenir leurs effets (importance des interactions dans le milieu).

2.- ORGANISATION DE L'EXPERIMENTATION

2.1. - Choix du dispositif

Choix les plantes annuelles, les variations non dépendantes des traitements sont surtout imputables à des facteurs "positionnels", c'est-à-dire liés à la situation de la parcelle par rapport à l'hétérogénéité du terrain. Les erreurs d'implantation d'un dispositif devront donc être facilement rectifiables.

L'expérience acquise en matière d'expérimentation nous permet d'affirmer que les dispositifs simples à randomisation totale et estimation unique de l'erreur sont toujours préférables aux dispositifs compliqués. Ces derniers n'apportent souvent rien de plus qu'une analyse difficile. Sauf implication technique (travaux du sol en particulier), les plans d'expérience doivent s'dresser aux simples blocs avec répartition des parcelles au hasard (6 répétitions minimum), la nature factorielle de l'expérimentation permet, chaque fois qu'on le peut, d'utiliser la méthode des comparaisons orthogonales qui introduit une grande clarté dans les conclusions agronomiques (détermination des effets). La décomposition en fonctions orthogonales est du plus grand intérêt dans les essais conduits en série du fait de l'hétérogénéité fréquente de la variance des interactions. Elle permet le test de chacun des contrastes par sa composante dans l'interaction. Le dispositif en carré latin peut être utilisé également avec profit si on soupçonne une variabilité appréciable liée aux deux directions du terrain. Cependant il est recommandé de n'utiliser ce dispositif que si le nombre de traitements est compris entre 5 et 8.

Dans les recherches de développement, conduites dans une perspective de diffusion, le nombre de questions dans les expérimentations est souvent très réduit. Les plans d'expérience font appel à des dispositifs simples du type blocs à répartition des parcelles au hasard. Dans les études préalables on peut toutefois s'adresser à des dispositifs multifactoriels pour inclure plusieurs facteurs et déterminer ceux qui présentent des interactions ou ceux qui sont indépendants.

Il y aura lieu tous les cas d'adapter les dispositifs aux "conditions locales", mais il ne faut pas pour autant rejeter pour des raisons dites d'adaptation ou de simplification des dispositifs clairs et précis parfaitement utilisables dans les conditions de station et recommander des schémas simplistes qui n'apportent souvent que des réponses partielles (voire erronées) et retrécissent ainsi le champ des recherches sans raison valable. Il faut selon nous avoir une politique raisonnée de l'expérimentation et l'adapter avant tout aux moyens disponibles.

2.2. - Dispositif en série - Binomes de culture

Les variations de la pluviométrie annuelle entraînent de grandes différences chaque année dans la réponse aux fertilisants. Le dispositif en série permet de tester plusieurs années successives le même traitement sur la même culture et de diminuer ainsi l'importance de l'effet annuel. En prenant comme base de comparaison pour l'étude des différences observées entre traitements l'interaction traitement x année, il est alors possible de tester si ces différences ont des chances de rester valables pour une suite d'années (4 années successives semblent un minimum : le coefficient de variation de la pluviométrie passe de 25 % pour une année à 15 % pour 4 années consécutives).

Cependant, la grande variabilité annuelle entraîne souvent une hétérogénéité importante des variances de l'erreur expérimentale. Ce fait rend souvent difficilement interprétable l'analyse en série et il est toujours souhaitable que l'expérimentation annuelle reste exploitable.

Pour ce faire, il importe que le nombre de répétitions par série soit au minimum de quatre (six étant préférable). Il y aura autant de séries que de phases (ou cultures) de la rotation, celle-ci pouvant être conduite sur plusieurs cycles. Le plus souvent et afin de limiter le nombre des séries à deux, les études se feront sur des binomes de culture à priori les plus significatifs de la région où est conduit l'expérimentation. Ces systèmes coïncident le plus souvent avec des systèmes cibles pour une zone homogène donnée. Exemple : Céréales - Culture de rente : Sorgho-cotonnier, mil-arachide ou Maïs - Soja. Ce binome peut d'ailleurs être modifié dans le temps après résultats complets de quelques cycles. Les recherches en matière de fertilisation seront toujours conduites selon ce schéma et permettront une meilleure approche de la fumure minérale au niveau du binome ou de la succession des cultures.

2.3. - Observations sur les techniques culturales

- Chaque fois que possible nous donnerons la préférence à la disposition des traitements contigus sans allées entre les parcelles. Les effets de bordure sont en effet très importants et il semble plus difficile de "réussir" des cultures en petites parcelles non juxtaposées. Les lignes de bordure sont éliminées à l'intérieur des parcelles pour définir une parcelle utile bien représentative du traitement et ne subissant aucune influence du traitement voisin.

- Dans les essais de longue durée la définition des traitements doit rester rigoureusement la même mais il est impossible de définir une fois pour toutes les techniques culturales au moment de l'élaboration du protocole. Les techniques évoluent et les améliorations acquises se doivent d'être adoptées dans les expérimentations. Il n'est pas souhaitable de conserver des techniques de qui n'étant plus valables agronomiquement ôteraient toute signification à l'expérimentation au bout d'un certain nombre d'années.

Les techniques retenues doivent concilier deux impératifs :

- conserver à l'expérimentation toute sa rigueur, sa précision et sa sensibilité afin d'éliminer le plus grand nombre de facteurs de variation autres que ceux dont on veut mesurer l'effet.

- ne jamais créer de conditions artificielles qui risqueraient de rendre les conclusions inapplicables dans la pratique (proscrire le jardinage et se rapprocher le plus possible des conditions de culture en plein champ).

2.4. - Remarques sur les essais de longue durée

Dans un essai "Courbe de réponse" à la fumure azotée, la dose optimum et rentable sur maïs a été déterminée pendant 4 années consécutives. Cette dose voisine de 90 kg/ha N, a permis d'obtenir pendant ce cycle 4 t/ha maïs grains par an. La cinquième année, sous des conditions climatiques semblables les rendements ont chuté à 1,5 t/ha et la baisse de rendement a été d'autant plus nette que les doses d'azote appliquées étaient élevées. Une toxicité importante avec forte acidification (PH 3,6) est apparue brutalement compromettant gravement la croissance et le développement de la plante. Si l'essai avait été arrêté après 4 ans, le chercheur aurait pu en toute bonne foi conseiller pour un système intensif la fumure de 90 kg/ha N/an sur maïs. A court terme, on aurait abouti à une catastrophe et le sérieux des travaux de la recherche en aurait été fortement compromis.

Les essais de longue durée sont indispensables pour l'élaboration des doctrines de fertilisation. Malgré les erreurs qui ont été relevées dans leur conduite au cours des temps (erreur inévitable en raison même de la succession des responsables), les essais de longue durée constituent toujours des références précieuses quelqu'en soient les résultats.

Dans l'expérimentation citée précédemment, deux problèmes sont maintenant clairement posés.

- Qu'aurait-il fallu faire pour maintenir la fertilité du sol, éviter l'évolution défavorable observée et obtenir régulièrement de bons rendements ?
- Que faut-il faire maintenant pour récupérer le sol. L'évolution défavorable observée est-elle réversible et peut-on évaluer en terme de coût les opérations nécessaires pour ramener le système à son état initial ?

On ne saurait trop conseiller aux chercheurs qui ont la chance de suivre des essais de longue durée de réfléchir profondément avant de les abandonner ou de les modifier (une modification qui ne perturbe pas la nature même des traitements initiaux demeure cependant possible).

Un des intérêts majeurs des essais de longue durée est la création de situations variées avec des différences sensibles en réserves et proportions d'éléments nutritifs. Ce type d'essai se prête particulièrement aux études d'évolution du sol et à l'exploitation statistique des résultats. Il ne faut pas perdre de vue également que certains facteurs actuellement jugés secondaires (oligoéléments etc...) peuvent devenir importants à moyen terme. Si l'équipe pluridisciplinaire, ayant en charge les problèmes qui seront alors posés, dispose d'essais de longue durée avec des parcelles suffisamment grandes autorisant la subdivision, les solutions seront plus rapidement disponibles.

2.5. - Les essais hors station

Pour les recherches de développement, il apparaît indispensable de conduire des essais simples, élaborés à partir des résultats obtenus en station, dans des situations différentes (variation dans l'espace) et de les répéter plusieurs années successives (variation dans le temps). Il est en effet très fréquent d'observer dans une série d'essais des résultats très différents d'un essai à l'autre, bien que chacun soit réalisé dans des conditions qui paraissent semblables.

Pour cette raison on ne saurait trop conseiller d'implanter les points dans des zones homogènes (conditions pédoclimatiques similaires), afin que ceux-ci fournissent des éléments de répétition à l'intérieur du regroupement choisi.

3.- CONDUITE DES EXPERIMENTATIONS

"Dans l'investigation scientifique, les moindres procédés sont de la plus haute importance".

C1. BERNARD

Un résultat agronomique est le plus souvent exprimable par un chiffre. Ce chiffre doit être obtenu avec le maximum de sécurité car il engage le responsable de l'expérimentation vis-à-vis de l'équipe agronomique et des services de vulgarisation et de développement. L'expérimentation implique donc un certain nombre de servitudes indispensables pour ne pas rendre le travail vain ou malhonnête.

Il est toutefois matériellement impossible pour le responsable de participer activement à toutes les tâches ou de superviser toutes les opérations.

C'est pourquoi nous considérons qu'en dehors de quelques contrôles occasionnels (au semis et à l'épandage des engrais en particulier) le chercheur doit être entièrement libéré de la surveillance à tous les stades de réalisation de l'expérimentation.

Le contrôle d'un programme de travail très large conduit ainsi implique quatre conditions si on ne veut s'exposer qu'à quelques erreurs acceptables liées principalement aux facteurs du milieu.

- La présence d'un chef de culture rompu aux travaux de routine de l'expérimentation à tous les stades de réalisation. La formation de ce technicien nécessite souvent plusieurs années de travail côte à côte en augmentant progressivement les responsabilités et la difficulté des tâches. Il est nécessaire de faire constater au futur responsable ses propres erreurs pour lui faire accepter l'impératif du double ou triple contrôle de certaines opérations. Ce travail de formation est très important et explique que certains chercheurs préfèrent "supposer" la parfaite compétence du réalisateur pour se consacrer personnellement à des travaux plus "scientifiques" et que d'autres à l'opposé tente de contrôler eux-mêmes toutes les opérations.
- La préparation d'un document très complet, recueil des protocoles d'essais avec instructions détaillées permettant de faire face à toutes les éventualités. Certaines précautions ou remèdes peuvent paraître puérils, l'expérience prouve qu'il est souvent essentiel d'en faire état.
- L'établissement d'un planning précis pour chaque campagne.
- L'enregistrement complet des résultats, des notes et observations prises au champ tout au long de la campagne. Ces dernières ne seront pas publiées mais elles doivent être disponibles à tout moment et singulièrement au stade de l'interprétation agronomique finale.

Pour les essais hors-station principal (stations secondaires et points d'appui d'expérimentations multilocales), la présence d'un technicien qualifié par station est indispensable. La formation de ce dernier est en général assurée par un séjour plus ou moins long à la station principale au début de sa carrière et par un contrôle permanent et étroit pendant les premières années de son travail en station. Un recyclage périodique permettant de tester les capacités réelles de l'agent est cependant souhaitable.

Si des expérimentations sont conduites directement chez le cultivateur, il importe d'assurer la formation complète d'un personnel d'encadrement spécialisé dans ces travaux et dégagé de toute autre tâche. Les nombreux échecs enregistrés dans ce domaine ont conduit certains chercheurs à constituer des équipes volantes très spécialisées et entraînées aux différentes opérations (piquetage, semis, épandage, récolte, pesées...). Cette méthode est coûteuse mais permet d'obtenir généralement des résultats fiables.

Malgré toutes les précautions signalées, des contrôles et des observations fréquentes par les chercheurs qualifiés restent indispensables. On ne peut mieux conclure ce paragraphe qu'en citant le Docteur PEARCE de la station de recherche d'East Maling (arbres fruitiers) en Grande-Bretagne.

"On doit insister sur le fait qu'aucun système de relevés sur les arbres, si bien conçu et mis en pratique soit-il, ne peut remplacer le lien intime et presque personnel qui doit être entretenu entre l'expérimentation et ses arbres. Les plantes ne sont pas mesurées principalement pour fournir des données à des analyses statistiques, mais pour enregistrer des observations qui décriront aux autres ce qui est apparu à l'expérimentateur sur le moment. L'autre objet, celui qui est de confirmer ce que l'on suspecte rend possible un test de signification, qui bien qu'important, est cependant secondaire".

4.-INTERPRETATION STATISTIQUE ET AGRONOMIQUE

" The actual and physical conduct of an experiment must govern the statistical procedure of its interpretation"

R.A. FISCHER

4.1. Vérification des hypothèses de base

Le calcul statistique se propose de nous indiquer si les variations provenant des erreurs sont suffisamment petites pour que nous puissions admettre avec une grande probabilité que les différences enregistrées ont bien pour cause l'application des traitements. L'emploi du carré moyen de l'erreur à la base des tests de signification (Test F) est sous la dépendance de plusieurs postulats :

- la distribution de fréquence des données doit pouvoir être assimilée à la loi normale (normalité de l'erreur).
- les données doivent être obtenues de manière indépendante.
- la variance des données doit être homogène (homogénéité ou constance de la variance de l'erreur) ;
- les équations sur lesquelles repose l'analyse de variance doivent avoir un sens (additivité des effets, modèle linéaire).

4.1.1. Distribution des variables analysées

Ces études sont nécessaires quand les variables analysées sont des teneurs ou des variables dérivées des variables primitives : poids de grain par épi, rapport parties utiles sur parties végétatives... Il est nécessaire de déterminer les résidus par bloc ou traitement pour établir leur fonction de répartition. L'utilisation de transformations adéquates permet de "normaliser" certaines distributions dissymétriques évidentes ou confirmées par un test de symétrie ou de normalité.

4.1.2. - L'indépendance des données

Cette condition est presque toujours vérifiée si l'essai a été correctement pensé et réalisé, mais il est bon de rappeler la nécessité absolue de l'indépendance totale de chaque résultat par rapport aux autres (cas en particulier des traitements phytosanitaires).

4.1.3. - L'homogénéité des variances résiduelles

L'hypothèse d'invariance de la variance n'est en général qu'approximativement vérifiée. Il arrive fréquemment que sous l'effet d'un traitement la valeur de la variance augmente avec celle de la variable (liaison moyenne-variance). L'hétérogénéité des variances résiduelles a pour conséquence de biaiser l'erreur qui peut être sous-estimée ou au contraire sur-estimée (augmentation de l'interaction traitement x répétition confondue dans l'erreur et par conséquent non identifiable). L'homogénéité peut être facilement vérifiée au moyen des critères de BARTLET *et de HARTLEY et dans le cas d'hétérogénéité, l'utilisation d'une variable transformée à variance constante permet de valider l'analyse. Il y a lieu de souligner que l'hétérogénéité de la variance peut être due à la présence d'un témoin absolu très inférieur à tous les autres traitements (cas assez fréquent dans les essais de fumure). Des rendements de l'ordre de 400 kg/ha ne peuvent pas avoir la même variance que des rendements de l'ordre de 5 000 kg/ha. La solution consiste à éliminer le témoin de l'interprétation (sa différence étant évidente avec les autres traitements). L'analyse se fait de la manière habituelle ; la suppression d'un traitement en totalité n'affectant pas l'orthogonalité de l'essai (essai bloc en particulier).

4.1.4. - L'additivité des effets

La non additivité peut avoir plusieurs causes. En premier lieu le modèle additif peut ne pas confondre au comportement biologique de la variable mesurée. On citera à ce propos l'exemple des céréales où un facteur peut agir sur le nombre d'épis, un autre facteur sur le poids des épis, et ceci d'une manière indépendante. Dans ce cas les effets des deux facteurs se multiplient au lieu de s'additionner. Une autre forme de non additivité se présente quand certaines données sont très nettement en dehors de la population. Si les causes de ces écarts sont connues (hétérogénéité manifeste de la parcelle) ou que le résultat est invraisemblable (erreur accidentielle très probable à la récolte, aux pesées, à la transcription des résultats...) les données aberrantes doivent être éliminées, puis calculées pour rétablir l'orthogonalité du dispositif. Le nombre de degrés de liberté du carré moyen de l'erreur doit être diminué d'une unité par parcelle estimée. Les essais comportant un trop grand nombre de parcelles calculées doivent être considérés avec beaucoup de prudence en particulier si le nombre de degrés de liberté de la variance de l'erreur est faible. On admet en pratique un maximum de 2 parcelles calculées (à moins de 5% d'observations aberrantes, l'estimation demeure possible). Si les parcelles aberrantes sont situées dans la même répétition, il est toujours préférable de supprimer cette répétition, ce qui n'affecte pas l'orthogonalité de l'essai (au moins pour un essai bloc). Le type de non additivité le plus fréquent est obtenu quand les traitements d'un premier facteur modifient le classement des traitements d'un deuxième facteur les uns par rapport aux autres.

* Le test de BARTLETT est sensible à la non normalité et n'est applicable que si le nombre de degrés de liberté des variances comparées est supérieur à 4. Le test de HARTLEY ne peut être appliqué que pour un nombre de variance inférieur à 12. Pour un petit nombre de degrés de liberté et un grand nombre de variances comparées on peut utiliser la méthode présentée par HEALEY et TAYLOR qui comporte en particulier l'étude de la relation Variance. Moyenne (hétérogénéité évidente mesurée par un coefficient de corrélation).

se traduisent par une interaction, les dispositifs expérimentaux qui mettent en évidence les interactions sont les seuls à pouvoir contrôler l'additivité. On retrouve là un des avantages du dispositif factoriel qui grâce au contrôle des interactions procure une meilleure estimation de l'erreur. Pour les essais blocs, le critère de TUKEY permet d'isoler la part due à la non additivité et d'indiquer les transformations pour la réduire.

Dans la plupart des essais, les postulats à la base de l'analyse de la variance sont remplis d'une manière acceptable. Toutefois il faut se méfier d'une application routinière de l'analyse de la variance suivant un schéma type et l'examen attentif des données brutes en liaison avec le plan d'expérimentation est toujours profitable avant d'aborder tout calcul. Pour les variables courantes la détermination de la loi de distribution, le contrôle et la stabilité de la variance et de l'additivité est toujours recommandable. Dans la plupart des cas, les transformations appropriées sont connues. D'autres cas, plus complexes, nécessitent par contre un travail supplémentaire important d'autant que le non-additivité et l'hétérogénéité des variances peuvent ne pas être obligatoirement contrôlées par la même transformation. Ces calculs sont alors du ressort d'un service de statistique spécialisé. La tâche de ce dernier sera grandement facilitée par l'utilisation de dispositifs simples (blocs ou blocs-factoriels) par les expérimentateurs, dispositifs pour lesquels la vérification des hypothèses de base est relativement aisée.

4.2. - Autres précautions et critères

La théorie de la régression permet de préciser l'analyse de la variance et de tenir compte de facteurs de variation reconnus et mesurés non compris dans l'expérimentation principale. Une autre précaution consiste à s'assurer de l'homogénéité de l'expérimentation. Pour clore ce chapitre, nous ferons état d'un indice très utilisé pour juger de la précision d'un essai : le coefficient de variation.

4.2.1. - L'analyse de covariance

Cette technique permet d'améliorer la précision des expériences grâce à la connaissance d'une variable concomitante. Il est indispensable que les facteurs mis à l'étude dans l'expérimentation principale n'aient aucune action sur la variable concomitante d'où la nécessité d'une première analyse sur cette variable. Il est nécessaire de s'assurer de la relation entre la variable mesurée et la variable concomitante par une étude sur les résidus et de tester le coefficient de corrélation résiduel. L'analyse de la covariance est particulièrement recommandée si on dispose d'un essai d'uniformité préliminaire à l'essai principal : essai sur plantes perennes (fourrages), essai dans lequel les traitements n'interviennent qu'après une récolte ou mesure (essais sur repousses de canne par exemple)...

4.2.2. - Homogénéité des réponses aux traitements dans les répétitions

Cette vérification est facile dans les expériences factorielles du type $2n$ ou $3n$ plusieurs répétitions où l'on doit avant de procéder à l'analyse globale effectuer les analyses répétition par répétition pour s'assurer de l'homogénéité des réponses. J. DEJARDIN propose pour les expérimentations non factorielles de faire l'analyse sur les répétitions groupées deux à deux. Dans le cas d'hétérogénéité des réponses, il y a lieu d'en rechercher la cause. Si celle-ci n'est pas manifeste, les conclusions de l'interprétation devront être émises avec beaucoup de prudence.

4.2.3. - Coefficient de variation ($Cv = \frac{s}{\bar{x}}$)

Le coefficient de variation est un indice très utilisé qui se définit par l'écart type de l'erreur parcelle en % de la moyenne générale de l'essai. Il permet de juger globalement la précision de l'essai. Il est utile de souligner que cet indice est seulement indicatif et doit être interprété en relation avec les dimensions de la moyenne (\bar{x}) et l'écart type de l'erreur (s). Il est extrêmement délicat de fixer la limite acceptable pour le coefficient de variation. En effet, il n'y a aucune raison statistique de ne pas tenir compte d'un effet significatif obtenu dans un essai à Cv élevé. D'autre part le Cv varie non seulement avec les plantes et les données observées sur ces plantes, mais il est également très fluctuant dans le temps (sécheresse ou humidité excessive accentuant les hétérogénéités naturelles). Un coefficient de variation élevé doit toutefois induire à la prudence et entraîner de la part de l'expérimentateur un (nouvel) examen attentif des données de base.

4.3.- Signification et Tests

Avant de commencer "l'analyse de la variance", l'examen des moyennes des traitements peut montrer des différences arithmétiques très faibles. A l'opposé, une moyenne peut être très nettement différente (inférieure ou supérieure) à la moyenne des autres traitements. Il est bien évident que dans ces deux cas, l'analyse statistique sera d'un faible intérêt, si ce n'est pour renseigner sur les dimensions de la variance résiduelle et la précision de l'expérimentation.

Une règle essentielle de l'expérimentation agronomique est celle de la désignation. Le principe consiste à fixer clairement les questions auxquelles devra répondre l'expérimentation, puis adopter un schéma expérimental utilisant des tests statistiques désignés répondant par oui ou par non aux questions posées.

Cette méthode implique donc d'avoir choisi le mode d'interprétation et par conséquent réalisé schématiquement celle-ci avant la mise en place de l'essai. Il est essentiel de souligner que l'examen des résultats ne doit jamais conduire à une modification des tests désignés.

Deux cas peuvent se présenter lors de la planification :

- Individualisation des degrés de liberté des traitements impossibles

Essai avec désignation d'effets ou contrastes impossibles par sans intérêt.

- Individualisation des degrés de liberté des traitements possibles

Cas des essais factoriels en particulier.

4.3.1. - Individualisation des degrés de liberté des traitements impossibles

(s_t^2/s_e^2) Deux cas sont à environ selon la valeur observée du rapport du carré moyen entre traitements au carré moyen de l'erreur.

La loi de distribution de ce rapport est celle de F dans l'hypothèse nulle d'égalité. Si le rapport est inférieur à la valeur F de la table au seuil P, on accepte l'hypothèse nulle d'égalité ; si au contraire la valeur trouvée est supérieure à la valeur critique de P, on rejette l'hypothèse nulle d'égalité.

4.3.1.1. - Carré moyen entre traitements non significative-ment plus élevé que la variance de l'erreur (effets des traitements du même ordre que les variations aléatoires ; acceptation de l'hypothèse nulle d'égalité).

La non signification (F calculé inférieur à la table) peut avoir 4 causes :

- Traitements identiques ou voisins
- Nombre de répétitions trop faible
- Variance de l'erreur expérimentale importante (expérience pas assez précise)
- Grand nombre de traitements.

Dans ce dernier cas, quelques traitements peuvent être significativement différents mais ils sont noyés parmi les traitements identiques et le test global demeure non significatif.

Dans le cas de non signification et quelle que soit la cause, il est inutile de chercher à comparer deux moyennes particulières. La différence entre certaines moyennes (la plus petite et la plus grande par exemple) pourrait être significative si l'on ne disposait que des deux traitements correspondants ; mais dans un groupe de plusieurs moyennes, il est normal que certaines d'entre elles diffèrent assez nettement sous la seule influence du hasard d'échantillonnage. Pour chaque test utilisé, il ne faut pas perdre de vue qu'il y a 5 chances sur 100 d'obtenir par hasard un résultat significatif (ce résultat est même possible avec des données prises au hasard).

Exceptionnellement, il est possible de calculer des différences significatives avant d'avoir calculé le F.

Les deux exceptions principales sont les suivantes :

- Parmi un certain nombre de traitements, un couple de traitements a été désigné au moment de la planification, mais d'autres traitements d'un intérêt secondaire figurent dans l'expérimentation et rendent l'ensemble non significatif ; quel que soit le résultat du test F général, les deux traitements désignés seront considérés comme significatifs après confirmation par le test.
- Chaque fois qu'intervient une comparaison d'un groupe de traitements avec un traitement ou témoin (essais variétaux, essais soustractifs...), on peut comparer un traitement au choix avec le témoin (la comparaison a été désignée dès la planification) même si le test F n'est pas significatif.

On utilise généralement pour ces comparaisons le test de DUNNETT. Les valeurs critiques des différences entre traitements et témoin sont calculées à partir d'une loi de t généralisée dans laquelle toutes les différences sont considérées ensemble et non plus séparément.

4.3.1.2. - Carré moyen entre traitements significativement plus élevé que la variance de l'erreur (rejet de l'hypothèse nulle d'égalité)

L'hypothèse d'homogénéité ayant été rejetée, on se heurte au problème des comparaisons multiples pour classer les traitements. Selon l'opinion des auteurs sur l'emploi et le sens des tests, différentes méthodes sont proposées.

L'emploi des différences significatives se justifie dans tous les cas où il y a plus de deux traitements.

Le test t peut seulement être utilisé dans le cas de deux traitements. On peut également l'utiliser pour la comparaison de deux moyennes désignées ou deux moyennes adjacentes dans un groupe de moyennes.

Par contre dans le cas de plusieurs moyennes à comparer, l'utilisation du test t multiple est incorrecte. Le test t suppose que les moyennes sont prises au hasard et sont indépendantes, ce qui n'est jamais réalisé dans un groupe de moyennes où le choix est décidé au vue des résultats et où les comparaisons ne peuvent pas être indépendantes. Les avantages et inconvénients des tests de comparaison deux à deux des moyennes de traitements ont été clairement exposés par P. DEJARDIN et N'GUYEN N'GOG QUOI.

Nous avons adopté depuis plusieurs années le test de KEULS qui est un test de comparaisons multiples au moyen de l'étendue. Cette méthode d'après les auteurs cités ci-dessus est actuellement la meilleure approximation pour les comparaisons multiples deux à deux (les tests proposés ne sont toujours qu'approximatifs).

Avec le test de KEULS, le risque de première espèce (rejet de l'hypothèse nulle d'égalité alors qu'elle est vraie : action du traitement qui en réalité est sans effet) est bien caractérisé par un seuil de probabilité. On court par contre un risque de deuxième espèce (acceptation de l'hypothèse nulle d'égalité alors qu'elle est fautive : non action du traitement qui en réalité a une action). Il est préférable dans les conditions souvent ingrates de l'expérimentation (coefficient de variation élevé) de courir ce risque, que l'on peut d'ailleurs réduire en augmentant le nombre de répétitions.

4.3.2. - Individualisation des degrés de liberté des traitements possibles

Nous avons déjà signalé l'intérêt des dispositifs factoriels qui sont les mieux adaptés à la recherche causale, et permettant de déterminer les facteurs à mettre en cause. Ces dispositifs sont seuls à pouvoir tester des hypothèses précises puisque la décomposition en degrés de liberté orthogonaux (synonyme du terme indépendant dans le langage des statisticiens) associée à une hypothèse un seul degré de liberté, par conséquent un test. A chacun de ces degrés de liberté correspond une fonction linéaire des résultats dénommée contraste, telle que la somme des coefficients soit nulle. Deux contrastes sont orthogonaux si la somme des produits des coefficients correspondants est nulle. Cela traduit une indépendance en probabilité, c'est-à-dire le test d'une hypothèse indépendement autres tests. Si tous les contrastes choisis sont orthogonaux, la décomposition est additive et la somme des carrés d'écart liés à chaque hypothèse avec un seul degré de liberté correspond à la somme des carrés des écarts calculés globalement pour l'essai.

Certains auteurs, mais en toute connaissance de cause mentionnent cependant dans les résultats, la dimension de la plus petite différence significative (PPDS ou LSD), même pour des essais mettant en comparaison plusieurs traitements.

4.3.2.1. - Dans le cas où le carré moyen des traitements n'est pas significativement plus élevé que la variance de l'erreur (F calculé) il est inutile de poursuivre l'analyse sauf si un contraste particulier a été désigné au moment de la planification.

4.3.2.2. - Si par contre le F calculé des traitements est significatif, nous disposons d'un jugement global sur les différences entre les traitements et il est alors possible de poursuivre l'analyse en effectuant un test rigoureux sur chacun des contrastes isolés avec seulement un degré de liberté (on doit à nouveau insister sur le fait qu'une façon possible de diviser les degrés de liberté pour les traitements n'est admissible que si les groupes ont été désignés au moment de la planification, le choix est relativement simple dans le cas d'essai à doses croissantes et équidistantes d'engrais). On répond alors à des questions précises qui ont le plus souvent l'avantage d'être hiérarchisées.

5. - CONCLUSION

Ce document pourra apparaître à la fois inutile et sans doute pour une part prétentieux, car présenté par un non spécialiste des méthodes statistiques. Il faut regretter que si la plupart des chercheurs sont appelés à utiliser les méthodes d'expérimentation aux champs, certains ignorent (ou feignent d'ignorer) les règles les plus élémentaires des techniques expérimentales.

Et puis, il nous a semblé opportun de rassembler et préciser quelques unes des observations réalisées au cours de plus de 15 ans de travaux en Afrique de l'Ouest.

- La conception d'une expérimentation et son organisation sont deux aspects essentiels. Si la question posée est simple et précise, le plan choisi sera simple. Encore faut-il définir la simplicité en matière de plan d'expérience ! Il nous apparaît essentiel pour planifier d'acquiescer "le raisonnement factoriel". Tous les problèmes doivent être posés en termes d'influence de facteur. On évitera donc les plans mettant en jeu un certain nombre de traitements à partir, desquels on ne pourra tirer aucune conclusion garantie de causalité. Il importe pour élaborer un bon protocole d'avoir prévu dès la conception de l'essai les modalités d'interprétation en utilisant en particulier les contrastes orthogonaux. L'expérience prouve que cette démarche, relativement aisée si les questions posées sont précises, devient vite automatique. La préférence devra toujours être accordée aux dispositifs à randomisation totale et estimation unique de l'erreur.

- Nous avons insisté sur les essais de longue durée qui sont indispensables pour définir les doctrines de la fertilisation. Les dispositifs doivent être soigneusement étudiés en raison de la perte de temps et d'argent qui pourrait résulter d'un mauvais choix. Ils constituent des références précieuses de "situations" pour les études d'évolution du sol. On a tout intérêt à multiplier le nombre d'observations qu'ils autorisent.

- Les quelques considérations sur la conduite des expérimentations sont le fruit de l'expérience. Nous avons souligné l'importance de l'aspect formation des responsables d'essai et l'intérêt d'établir des instructions soignées permettant de faire face à toute éventualité.

- Les remarques concernant l'analyse statistique peuvent sans doute prêter à critiques. Certains adoptent une attitude détachée vis-à-vis de l'utilisation de la statistique. Les méthodes statistiques se ramènent en fait à des simples calculs d'erreurs (d'où les notions de randomisation et répétitions et il ne faut pas vouloir leur attribuer une portée supplémentaire.

En fait la situation est claire : ou on utilise la statistique avec bon sens et discernement, et il n'y aucune raison d'adopter une telle attitude. Ou bien on ne s'en sert pas..... Or on constate souvent que des chercheurs font appel aux méthodes statistiques sans y croire vraiment.... Ce sont d'ailleurs souvent les mêmes qui exécutent seuls les calculs en consacrant ainsi à des travaux de routine un temps précieux qui pourrait probablement être mieux utilisé !

Nous avons rappelé un certain nombre d'hypothèses de bases que l'on suppose vraies quand on utilise l'analyse de la variance. L'intérêt des dispositifs factoriels (homogénéité des réponses) est à nouveau démontré. Le problème des comparaisons de moyenne a été exposé à la lumière des enseignements acquis lors de notre mission à la station expérimentale de Rothamsted*. Nous avons insisté sur l'intérêt de la décomposition en degrés de liberté orthogonaux pouvant être utilisés facilement dans les plans factoriels, mais également dans d'autres dispositifs sous réserve d'un choix préalable des traitements.

Il faut être prudent dans une interprétation des niveaux de signification. Les niveaux 5 % et 1 % ont été choisis arbitrairement par FISHER. On peut observer au cours d'une série d'expériences qu'une différence se produit toujours dans le même sens, - bien que jamais significative. Il ne faut pas oublier qu'un effet même observé à 10 % à 9 chances sur 10 d'être réel quoique non démontré d'une manière satisfaisante ! A l'opposé un résultat significatif mais portant sur une différence faible peut n'avoir aucune "signification" du point de vue économique.

En fait après l'analyse statistique rigoureuse, l'agronome doit reprendre ses droits et se livrer à nouveau à un examen minutieux des données brutes pour utiliser les résultats des calculs avec bon sens et impartialité.

L'auteur a eu la faveur de suivre au début de sa carrière un stage à la station expérimentale de Rothamsted en Grande Bretagne. La compétence et la notoriété de l'équipe d'agronomes - statisticiens qui s'est constitué sous l'impulsion de F. YATES est reconnu mondialement. Nous avons retiré de ce séjour un très grand profit. Les enseignements reçus et l'expérience acquise en Afrique par la suite ont largement contribué à alimenter les réflexions contenues dans le présent document.
