

RAPPORTS D'ACTIVITE

AGROPEDOLOGIE

1989

Organisation et développement
des moyens informatiques du Laboratoire
d'Agropédologie de Nouméa

Philippe de LABORDE de MONPEZAT

Rapport d'activité de VAT

Avril 1988 - juillet 1989

INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE
POUR LE DÉVELOPPEMENT EN COOPÉRATION

Centre de Nouméa

ORSTOM

RAPPORTS D'ACTIVITE

AGROPEDOLOGIE

1989

Organisation et développement
des moyens informatiques du Laboratoire
d'Agropédologie de Nouméa

Philippe de LABORDE de MONPEZAT

Rapport d'activité de VAT

Avril 1988 - juillet 1989



INSTITUT FRANCAIS DE RECHERCHE SCIENTIFIQUE
POUR LE DEVELOPPEMENT EN COOPERATION

CENTRE DE NOUMEA

ORGANISATION ET DEVELOPPEMENT DES MOYENS INFORMATIQUES
DU LABORATOIRE D'AGROPEDOLOGIE DE NOUMEA

(U.R. 3D/A "INFLUENCE DES SYSTEMES D'EXPLOITATION SUR LE SOL")

RAPPORT D'ACTIVITE DE VAT
Avril 1988 - Juillet 1989

par

Philippe de LABORDE de MONPEZAT

REMERCIEMENTS

Je tiens à remercier tout particulièrement

M. Bernard BONZON pour l'aide qu'il m'a apportée tant sur le plan professionnel que sur le plan privé,

les techniciens Edouard OUCKEWEN, Léon TAPUTUARAI, William NIGOTE, sans qui toutes les expérimentations ne pourraient être réalisées,

Bruno BUISSON et tous les autres membres du laboratoire d'informatique de l'ORSTOM pour leurs précieux conseils,

et toutes les personnes qui ont oeuvré pour rendre mon séjour agréable.

TABLE DES MATIERES

<u>1. ACTIVITES DU LABORATOIRE</u>	1
<u>1.1 Objectifs et méthodologie</u>	1
<u>1.2 Analyses statistiques</u>	3
<u>2. MATERIEL INFORMATIQUE UTILISE</u>	3
<u>2.1 Historique</u>	3
<u>2.2 Description</u>	4
-micro-ordinateurs.....	4
-périphériques.....	4
-divers.....	4
<u>3. LOGICIELS</u>	5
<u>3.1 Logiciels commerciaux</u>	5
-logiciels de statistiques.....	5
-logiciel de gestion de base de données.....	5
-logiciel de traitement de textes.....	5
-logiciels intégrés.....	5
<u>3.2 Développement de programmes spécifiques</u>	6
<u>4. ORGANISATION</u>	7
<u>4.1 Organisation matérielle</u>	7
<u>4.2 Organisation logicielle</u>	8
-répartition géographique des fichiers.....	8
-utilisation de PONT.....	8
-utilisation de SYTOS.....	9
<u>5. PERSPECTIVES</u>	9
<u>5.1 Activités du laboratoire</u>	9
<u>5.2 Informatique</u>	10
<u>6. CONCLUSION</u>	10

<u>ANNEXE 1</u>	12
<u>ANNEXE 2</u>	15
A ORGANISATION DES UNITES DE PROGRAMME DU LOGICIEL AGROSTAT.	15
-Listing des principales unités de programme.....	16
-Organisation du logiciel par le choix des menus.....	23
B GESTION DES FICHIERS DE DONNEES PAR LE LOGICIEL AGROSTAT..	24

1. ACTIVITES DU LABORATOIRE

1.1 Objectifs et méthodologie

Le laboratoire d'agropédologie de l'ORSTOM, sous la direction de M. B. BONZON, et en étroite collaboration avec les Services Ruraux du Territoire (en particulier avec le Centre de Recherche et d'Expérimentation Agronomique de Nessadiou), s'intéresse depuis dix ans au problème de la fertilité des sols de Nouvelle Calédonie.

Les études réalisées s'appuient toujours sur des dispositifs expérimentaux. Leur objectif principal est double:

-adapter les fumures et les amendements aux sols et aux cultures,

-caractériser l'évolution des sols dans ces conditions.

Se fondant sur le travail antérieur des pédologues en ce qui concerne la classification des sols, les principaux types de sols intéressant l'agriculture ont été définis. Pour chacun des types étudiés, il a été choisi un champ de référence, correspondant au type "modal" du sol (c'est à dire le type regroupant les caractéristiques moyennes définies dans la nomenclature pédologique).

La méthode d'étude prévoit une ou plusieurs expérimentations en serre, sur vases de végétation, avec le sol provenant du champ de référence choisi. Ces essais préliminaires ont pour but de définir précisément les problèmes de fertilité (carences, déséquilibres...) et de présager des améliorations à apporter (amendements, apports d'engrais...). Ensuite, une expérimentation au champ est mise en place, sur la base des résultats obtenus en serre, et cela pour une durée de cinq ans.

A l'issue de ces 5 années, on étudie la variabilité du type de sol, c'est à dire que l'on procède à des expérimentations multilocales sur des types de sol légèrement différents du type modal (soit par le pH, soit par la profondeur du sol ...) qui ont pour but d'adapter la fertilisation optimale définie sur le type modal.

Ainsi, par exemple, l'étude sur les vertisols hypermagnésiens, caractérisés par un rapport Mg/Ca très élevé (supérieur à 10), ce qui entraîne de graves déséquilibres pour les cultures et notamment une carence en calcium.

Les premières expérimentations en serre ont mis en évidence outre cette carence en calcium, des carences en azote et phosphore. Elles ont aussi permis de définir les formes et doses d'amendements calciques à apporter.

Ces résultats en serre furent à la base de l'expérimentation au champ: un essai selon le modèle du carré latin en combinant deux formes d'amendements calciques (croute calcaire et gypse) et trois doses (4, 8, 12 t/ha de CaO), essai qui a été mis en place en Août 1984 et dont le cinquième et dernier cycle s'est terminé début avril 1989.

Toutes ces expérimentations sont réalisées avec le maïs comme plante-test. Ce choix repose sur les avantages suivants:

- rapidité du développement: elle permet de limiter dans le temps les opérations de culture, d'où une flexibilité plus grande quant au choix de la période de mise en place des essais et de leur nombre, et une diminution des charges (techniques et financières) dues à la conduite de l'essai;

- sensibilité aux déséquilibres minéraux du sol: elle permet la mise en évidence aisée des problèmes et une approche plus fine des solutions à mettre en oeuvre;

- extrapolation possible entre les résultats en serre et ceux du champ: en tant que plante annuelle, les premiers stades du développement ne dépendent que des conditions du moment (climat, état du sol) et non des conditions antérieures comme ce serait le cas pour une plante perenne (initiation florale, accumulation de réserves...).

Afin de déterminer les effets (positifs ou négatifs) des traitements utilisés, on mesure de nombreux paramètres;

- sur le végétal:

- croissance en hauteur des plants
- vitesse de croissance
- rendement en grains
- rendement en tiges et feuilles
- teneurs en éléments minéraux des différents organes (grains, tiges et feuilles, racines)

- sur le sol

- teneurs en éléments minéraux du sol en début de cycle (avant la plantation)
- teneurs en éléments minéraux du sol en fin de cycle (après la récolte)

L'interprétation rationnelle et efficace de ces résultats ne se conçoit que si l'on dispose de moyens permettant le traitement statistique de toutes ces données.

Devant la masse importante des résultats, seule l'utilisation de méthodes de calcul automatique au moyen d'ordinateurs permet une interprétation rapide et systématique de ces résultats.

1.2 Analyses statistiques

Les calculs statistiques utilisables dans le cas d'essais contrôlés sont l'analyse de variance et la comparaison de moyennes.

L'analyse de variance estime les effets de chaque facteur contrôlé de l'essai et teste la probabilité que ces effets soient nuls en calculant le rapport de la variance associée à un facteur contrôlé à la variance résiduelle (associée aux facteurs aléatoires).

La comparaison des moyennes d'un facteur consiste en le calcul des différences entre ces moyennes prises deux à deux et le calcul de la probabilité que ces différences soient nulles.

Ces calculs et tests statistiques indiquent des différences éventuelles de comportement entre les traitements et permettent à l'utilisateur d'interpréter les résultats et de tirer des conclusions quant à leur efficacité.

2. MATERIEL INFORMATIQUE UTILISE

2.1 Historique

Durant les premières années, le laboratoire ne disposait que de moyens réduits, à savoir quelques calculatrices Hewlett Packard programmables, permettant le calcul de moyennes, de variances, d'écart-types sur des séries limitées de données. En raison de la lenteur et de l'aspect incommode des opérations (saisie des données, calculs, transcription des résultats), les opérations effectuées étaient le plus souvent réduites à l'indispensable: effet sur le rendement, les teneurs et les immobilisations en éléments minéraux dans la plante.

L'acquisition en 1985 d'un micro-ordinateur Goupil G3, puis d'un Goupil G40 en 1987 a permis une meilleure utilisation des données expérimentales: réalisation de fichiers de données facilement accessibles, possibilité de calcul plus étendue permettant d'envisager d'autres traitements statistiques (comparaison de moyennes par le test de Student ou le test de Newman et Keuls, étude de corrélations résiduelles ...).

Durant mon séjour, l'acquisition d'un Goupil G5 a encore augmenté les capacités de stockage d'informations et la rapidité des calculs.

2.2 Description

Actuellement, le laboratoire dispose des équipements suivants:

-micro-ordinateurs

.GOUPIL G40 286

640 ko de mémoire RAM
 co-processeur 287
 disque dur 20 Mo
 lecteur disquettes 5"¼ de capacité 1,2 Mo
 lecteur disquettes 3,5" de capacité 1,44 Mo
 écran couleur avec carte EGA
 souris

.GOUPIL G5 386

2 Mo de mémoire RAM (640 Ko utilisable par DOS
 et 1024 Ko de mémoire étendue)
 co-processeur 387
 disque dur 40 Mo (2 partitions DOS de 20 Mo
 chacune)
 streamer (lecteur de cartouche magnétique de 60
 Mo pour la sauvegarde du disque dur)
 lecteur disquettes 5"¼ de capacité 1,2 Mo
 écran multimode couleur (émulation carte EGA)
 souris

Ces deux ordinateurs sont reliés par une liaison RS232 permettant le transfert d'informations entre les deux machines.

-périphériques

.imprimante matricielle STAR NB 15
 .imprimante matricielle FUJITSU DL 3400 (option
 ruban 4 couleurs)
 .table traçante ROLAND DG DXY 880A

-divers

.2 onduleurs Imunelec 600 VA qui assurent la
 régulation de l'alimentation électrique des ordinateurs.

Remarque: il est conseillé de solliciter régulièrement les batteries (une fois par mois) en débranchant pendant quelques minutes l'onduleur afin de maintenir celles-ci en charge.

3. LOGICIELS

3.1 Logiciels commerciaux

Le laboratoire dispose de plusieurs logiciels:

-logiciels de statistiques

STAT-ITCF:

il permet de réaliser la plupart des traitements statistiques (régressions, analyses de variance, analyse multidimensionnelle ...).

STAT-GRAPHIC:

il est plus performant que le précédent en ce qui concerne la présentation des résultats, surtout sur le plan des graphiques.

-logiciel de gestion de base de données

DBASE III

-logiciel de traitement de textes

WORD 4

-logiciels intégrés

FrameWork II:

tableur, graphique, traitement de texte.

Son utilisation est assez peu pratique, et la sortie des graphiques sur table traçante pose quelques problèmes (l'impression ne correspond pas à la visualisation écran).

EXCEL:

tableur, graphique, base de données.

Sa récente acquisition n'a pas permis pour l'instant de tester ni d'exploiter toutes ses possibilités. Une première approche permet néanmoins de se rendre compte de ses performances supérieures à FW II quant à sa pratique d'utilisation, et à ses capacités graphiques.

Tous ces logiciels sont cependant actuellement peu, voire pas du tout utilisés, en raison soit de leur difficulté d'application (adaptation des logiciels statistiques aux conditions particulières des expérimentations réalisées), soit du manque de temps pour se familiariser avec les logiciels complexes et les utiliser de façon rationnelle.

3.2 Développement de programmes spécifiques

Les logiciels de statistiques ne sont pas toujours utilisables dans les conditions de l'essai réalisé: les répétitions des traitements à l'intérieur d'un même bloc ne sont pas prises en compte, on ne choisit pas toujours les tests effectués (pour la comparaison des moyennes par exemple), les modèles mathématiques utilisés ne sont pas apparents et peuvent entraîner des erreurs par suite d'un mauvais choix.

Jusqu'à présent, les traitements statistiques des données issues des différents essais ont été effectués à l'aide de programmes spécifiquement écrits pour l'essai considéré.

Par exemple;

-essai "Amendement calcique sur vertisol hyper-magnésien",
programme "ACVHM.COM"

-essai "Etude des carences en oligo-éléments des sols de Wallis et Futuna",
programme "OLIGO_WF.COM"

Ces programmes sont écrits en langage Pascal avec la version 3 du compilateur Turbo-Pascal.

Il s'est avéré que, bien que les essais soient différents les uns des autres, les traitements sont toujours assez semblables:

création des fichiers de données
analyse de variance (test de Fisher)
comparaison de moyennes (test de Student ou test de Newman et Keuls)

J'ai donc entrepris de réaliser un logiciel permettant le traitement des données quel que soit le type d'essai considéré. En se fondant sur les expériences précédentes et les projets envisagés par le laboratoire, il a été établie une liste des types d'essais réalisables et des modèles mathématiques correspondants (cf exemples en annexe 1).

La réalisation de ce logiciel baptisé AGROSTAT a été interrompue avant la mise en place de tous les cas considérés en raison de difficultés techniques rencontrées et provenant des limitations du compilateur Turbo-Pascal: la taille maximale octroyée au programme correspond à un segment soit 64 Ko. La généralisation des modèles statistiques entraîne des contraintes de programmation importantes pour gérer tous les paramètres du modèle (nombre de facteurs, niveaux subsidiaires...) qui conduisent à un dépassement de la capacité de mémoire.

Cela nous a amenés à utiliser la version n°4 du compilateur Turbo-Pascal. Dans cette version, il est possible de découper un programme en unités individuelles autonomes, lesquelles sont ensuite utilisées dans un programme principal exécutable. Si chaque unité reste limitée à 64 Ko, le code exécutable peut inclure plusieurs de ces unités.



Matériel informatique: disposition des machines



Matériel informatique: détail (Goupil G5, table traçante)

Face aux possibilités étendues du Turbo-Pascal 4, le logiciel AGROSTAT a été repensé et organisé de la façon suivante (cf détails en annexe 2.A):

- un programme principal (MAIN) qui, par une série de menus, détermine le travail à effectuer;

- un programme de gestion des fichiers de données (STATGEST);

- un programme de calculs statistiques (MOD_STAT) qui détermine le modèle mathématique et réalise les calculs.

Ce logiciel, dans sa version actuelle (2.10) est capable de traiter les types d'essais suivants:

- en blocs simples
- en blocs subdivisés
- en carré latin simple
- en carré latin subdivisé

Le cas d'un essai en blocs avec une subdivision par parcelles n'est pas utilisable dans l'état actuel du logiciel. Cependant il semble que les algorithmes de calcul utilisés dans le cas d'un carré latin subdivisé puissent être valables dans ce cas. Il conviendra alors, après vérification de cette hypothèse, d'adapter le programme de calcul (voir un informaticien).

Le logiciel permet en outre de réaliser des analyses statistiques sur un regroupement de plusieurs essais, en incluant un facteur "temps". Une extension au regroupement dans l'espace est prévue mais n'a pu être réalisée, faute de temps.

La construction du programme ainsi que l'organisation des fichiers de données sont détaillées dans l'annexe 2 (A et B).

4. ORGANISATION

4.1 Organisation matérielle

Chaque micro-ordinateur est branché sur un onduleur. L'ensemble des périphériques a été regroupé afin de faciliter les changements de branchement. En effet, les imprimantes étant de qualité différente, il peut être avantageux de choisir l'une ou l'autre pour les éditions, ou même la table traçante, sans être obligé de changer de poste de travail si la connexion voulue n'est pas la bonne.

Par défaut, le G5 est connecté à la FUJITSU, le G40 à la STAR, et la table traçante n'est pas connectée.

4.2 Organisation logicielle

-répartition géographique des fichiers

Tous les fichiers de données concernant les expérimentations sont regroupés sur le disque dur du G40 (C,label SERVEUR).

Les fichiers (données et programmes) réalisés par mes prédécesseurs se trouvent dans un répertoire nommé TURBO. L'ensemble de ces fichiers est accessible depuis le répertoire principal par le programme MENUSTAT.BAT.

Les fichiers de données des essais traités par AGROSTAT sont regroupés dans un répertoire nommé DATA, et accessible par le programme AGROSTAT.BAT.

Outre quelques utilitaires tels que SideKick, PCTools, se trouvent les logiciels suivants:

FrameWork II, Windows, Turbo-Pascal 3.

Le disque dur du G5 est partitionné en deux (C,label HD1 et D,label HD2).

Sur HD1 se trouvent les logiciels Word4, Excel, Turbo-Pascal 4, DBase III, les programmes-source de AGROSTAT, le logiciel de pilotage du streamer SYTOS, et le logiciel de liaison PONT.

Sur HD2 se trouvent les logiciels STAT-ITCF, STAT-Graphic.

-utilisation de PONT

Ce logiciel permet de définir une hiérarchie entre deux micro-ordinateurs reliés entre eux par un port série. Il est composé de deux parties, installées séparément sur chacun des micro-ordinateurs.

Sur la machine "maître" (celle qui va réaliser les opérations demandées par l'utilisateur, en pratique le G5) sont installés les programmes de pilotage dans un répertoire PONT (il faut de plus ajouter la commande d'installation du pont dans le fichier CONFIG.SYS). Sur la machine "esclave" (celle qui joue le rôle de serveur, en pratique le G40) est installé le programme PONT.COM.

Pour établir une première fois la liaison, il faut d'abord lancer le programme PONT sur la machine "esclave", puis relancer le système sur le G5 (Ctrl + Alt + Del). Si lors du chargement apparaît le message "ERREUR... PONT non installé", vérifier que le programme PONT a bien été lancé sur la machine "esclave", et que le cordon de raccord entre les deux micros est bien branché sur les ports série 1. Par la suite, l'accès au serveur depuis la machine "maître" se fera directement lorsque le programme PONT sera relancé sur la machine "esclave".

Attention! Si l'on relance le système sur l'une des deux machines, la liaison sera interrompue. Il faudra alors procéder pour la réactiver comme indiqué ci-dessus pour la première fois.

Lorsque les deux micros sont reliés par le pont, on peut accéder aux disques de la machine "esclave" depuis la machine "maître". En pratique, le G5 disposant de 5 disques; A,B (lecteur 5"¼), C,D (disques durs), E (disque virtuel), les 3 disques du G40 s'appelleront F (lecteur 5"¼), G (lecteur 3"½), H (disque dur).

-utilisation de SYTOS

Ce logiciel permet de gérer les opérations du streamer (lecteur de cartouche magnétique) et de sauvegarder les informations stockées sur le disque dur. L'utilisation du logiciel est précisée dans le manuel. On veillera toutefois lors de la sauvegarde de la totalité du disque à ce que les options de configuration soient:

"Sélectionner Inclure: *.*"

"Sélectionner Arbres : Oui"

Utiliser pour la sauvegarde l'option "Sauvegarder Disque".

Pour une restauration, utiliser l'option "Restaurer Disque Liste: *.*" (ou tout autre nom de fichier).

L'organisation des sauvegardes peut être réalisée comme suit:

le premier de chaque mois, sauvegarde sur la cartouche I,
le quinze de chaque mois, sauvegarde sur la cartouche II,
soit deux sauvegardes mensuelles.

Remarque: On peut sauvegarder le disque dur du G40 sur la cartouche par l'intermédiaire du pont.

5. PERSPECTIVES

5.1 Activités du laboratoire

Le CREA, qui vient tout récemment de s'équiper d'une serre, et qui bénéficie de l'expérience du laboratoire de l'ORSTOM quant à la technique de culture sur vases de végétation, peut dorénavant réaliser une partie des expérimentations sur la fertilité des sols jusqu'alors effectuée par l'ORSTOM. Leur première expérimentation sur les sols oxydiques du sud de la Calédonie est d'ailleurs en train de s'achever.

Ce transfert de technologie va d'autre part permettre au laboratoire d'agronomie de l'ORSTOM de se consacrer à d'autres problèmes; ainsi le lancement de l'opération "cocotier" dont le but est d'une part d'adapter la technique de culture sur vases de végétation au cocotier, technique qui permettrait sans doute d'accélérer le développement des jeunes plantules obtenues par culture in vitro et donc de limiter les délais avant la plantation, et d'autre part de déterminer les besoins réels de la plante et sa capacité à utiliser les réserves minérales d'un sol, en vue d'améliorer les conditions de culture du cocotier dans la majeure partie des petits archipels du Pacifique Sud, dont les sols sont souvent très pauvres.

5.2 Informatique

Grace à la réalisation du logiciel AGROSTAT, la mise en oeuvre des traitements statistiques pour des essais classiques sera dorénavant une opération de routine, ce qui va permettre d'utiliser à d'autres tâches le temps jusqu'alors consacré à la réalisation des programmes d'analyse de variances.

On pourra alors envisager une interprétation plus poussée des résultats, par exemple avec la recherche de corrélations entre différentes mesures expérimentales.

Cette recherche de corrélation peut tout d'abord faire l'objet d'une étude graphique, puis se définir mathématiquement par des calculs de corrélation sur les variances résiduelles.

L'utilisation de logiciels tels que EXCEL se justifie alors pleinement.

Toujours dans une optique d'optimisation, chaque nouvelle réalisation de programme devra autant que possible envisager une généralisation du problème à résoudre de façon à limiter par la suite la réalisation de programmes similaires, ce qui, outre l'aspect répétitif et peu valorisant pour le programmeur, pénalise l'efficacité du travail.

On peut aussi envisager de la part du VAT en poste une meilleure utilisation de ses compétences agronomiques (puisque celles-ci font partie des critères pour le choix des candidats et leur recrutement).

6. CONCLUSION

Les capacités de recherche du laboratoire sont actuellement limitées par le manque de chercheur. En effet, l'équipe de l'U.R.3D/A qui comprenait à mon arrivée deux chercheurs, (un agronome, B. Bonzon et un pédologue, C. Thomann) se réduit actuellement à un seul chercheur. De plus, l'U.R.3J ("Etudes de synthèse et cartographie), dirigée par le pédologue P. Podwojewski, lequel collaborait avec nous, a été fermée depuis son départ.

Si l'on veut pouvoir relancer le potentiel de recherche agropédologique de l'ORSTOM sur le territoire, il semble indispensable d'augmenter les effectifs de cette U.R.

D'autre part, sur le plan technique, la présence d'un VAT me semble insuffisante pour assurer le développement statistique et informatique des méthodes d'analyses des résultats d'expérimentations, d'autant plus que leur activité comporte aussi et devrait être surtout une participation aux autres travaux de l'U.R: recherche documentaire, organisation et suivi d'expérimentations en serre, etc. En effet, outre le fait que le niveau des VAT successifs n'est pas identique d'un appelé au suivant, la continuité du travail se heurte au problème du recrutement et des impératifs du calendrier des affectations militaires. En conséquence, il semblerait judicieux qu'il y ai parmi le personnel permanent du laboratoire un ingénieur informaticien ayant en outre des connaissances en statistique afin de pouvoir aller plus loin dans la réalisation de logiciels informatiques ainsi que dans la conception de méthodes statistiques plus appropriées aux besoins du laboratoire.

ANNEXES

ANNEXE 1EXEMPLES DE MODELES LINEAIRES
D'ANALYSE DE VARIANCE

Un modèle linéaire d'analyse de variance est la décomposition d'une valeur expérimentale en une somme de termes qui correspondent chacun à un effet d'un facteur du modèle expérimental.

Le modèle le plus simple avec répétitions est celui d'un essai en blocs complets équilibrés à un facteur contrôlé. Il s'écrit:

$$X_{ij} = x + F_i + B_j + e_{ij}$$

où

X_{ij} représente la valeur mesurée pour le $i^{\text{ème}}$ traitement du facteur F et le $j^{\text{ème}}$ bloc B,

x représente la moyenne générale,

F_i représente l'effet de la $i^{\text{ème}}$ modalité du facteur F,

B_j représente l'effet du $j^{\text{ème}}$ bloc B,

e_{ij} représente l'ensemble des effets aléatoires (facteurs non contrôlés).

On estime chaque terme par les formules suivantes:

$$F_i = x_i - x$$

$$B_j = x_j - x$$

$$e_{ij} = X_{ij} - F_i - B_j - x = X_{ij} - x_i - x_j + x$$

dans lesquelles

$$x_i = \sum_j (X_{ij} / n_B) \text{ avec } n_B = \text{nombre de blocs}$$

$$x_j = \sum_i (X_{ij} / n_F) \text{ avec } n_F = \text{nombre de modalités du facteur F}$$

La variance de chaque terme se calcule comme suit:

$$s^2_F = \sum_i \sum_j (F_i^2) / (n_F - 1) = n_B \sum_i (x_i - x)^2 / (n_F - 1)$$

$$s^2_B = \sum_i \sum_j (B_j^2) / (n_B - 1) = n_F \sum_j (x_j - x)^2 / (n_B - 1)$$

$$s^2_e = \sum_i \sum_j (e_{ij}^2) / (n_F - 1)(n_B - 1)$$

Lorsque le modèle comporte plusieurs facteurs, il intervient des interactions entre les facteurs. Si l'on prend comme exemple l'essai fertilisation NPK sur vertisol équilibré de Pouembout, le modèle est le suivant:

$$X_{ijk1} = x + N_i + P_j + K_k + B_1 + \\ (NP)_{ij} + (NK)_{ik} + (PK)_{jk} + (NPK)_{ijk} + \\ e_{ijk1}$$

termes du modèle	estimations des termes
N_i	$X_i - X$
P_j	$X_j - X$
K_k	$X_k - X$
B_1	$X_1 - X$
$(NP)_{ij}$	$X_{ij} - X_i - X_j + X$
$(NK)_{ik}$	$X_{ik} - X_i - X_k + X$
$(PK)_{jk}$	$X_{jk} - X_j - X_k + X$
$(NPK)_{ijk}$	$X_{ijk} - X_{ij} - X_{ik} - X_{jk} + X_i + X_j + X_k - X$
e_{ijk1}	$X_{ijk1} - X_{ijk} - X_1 + X$

Il existe des modèles plus complexes, lorsqu'on fait intervenir un ou plusieurs facteurs subsidiaires. On appelle facteur subsidiaire un facteur dont les n_s modalités sont, pour des raisons techniques, appliquées aléatoirement à autant de sous-parcelles découpées à l'intérieur de chacune des parcelles principales de chaque bloc. Le nombre de parcelles principales correspond, dans ce cas, au nombre des combinaisons des modalités des autres facteurs contrôlés (bloc excepté), appelés pour cette raison "facteurs principaux". Par exemple, un amendement calcique sur un essai au champ sera plus aisé à mettre en place sur un ensemble regroupé de parcelles que sur des parcelles dispersées. Lors de regroupement d'essais annuels, par exemple, le facteur année pris en compte peut être considéré aussi comme un facteur subsidiaire dans la mesure où chaque parcelle de l'essai porte plusieurs relevés (un par année).

Dans le cas des essais de fertilisation NPK sur le vertisol de Pouembout, le regroupement des essais par campagne correspond ainsi au modèle suivant:

$$\begin{aligned}
 X_{ijklm} = & X + \\
 & N_i + P_j + K_k + B_1 + \\
 & (NP)_{ij} + (NK)_{ik} + (PK)_{jk} + (NPK)_{ijk} + \\
 & e_{ijk1} + \\
 & T_m + \\
 & e_{1m} + \\
 & (NT)_{im} + (PT)_{jm} + (KT)_{km} + (NPT)_{ijm} + (NKT)_{ikm} + \\
 & (PKT)_{jkm} + (NPKT)_{ijkm} + \\
 & e_{ijk1m}
 \end{aligned}$$

Pour un essai en carré latin, le principe reste le même, mais on remplace l'effet bloc par un effet ligne et un effet colonne. Si de plus on a un facteur subsidiaire, on ne peut pas définir de résiduelle associée à ce facteur, et il faut faire apparaître les interactions avec les facteurs de répétition.

Ainsi par exemple le modèle de l'essai amendement calcique sur vertisol hyper-magnésien:

$$X_{ij1r} = x + A_i + D_j + (AD)_{ij} + L_1 + C_r + e_{ij(1r)} + T_m + (LT)_{1m} + (CT)_{rm} + (AT)_{im} + (DT)_{jm} + (ADT)_{ijm} e_{ij(1r)m}$$

où

A représente le facteur "forme de l'amendement calcique",
 D le facteur "dose de l'amendement calcique",
 L le facteur "lignes",
 C le facteur "colonnes",
 T le facteur subsidiaire "apport complémentaire de gypse".

ANNEXE 2A ORGANISATION DES UNITES DE PROGRAMME DU LOGICIEL AGROSTAT

Le logiciel est constitué de plusieurs unités de programmes dont chacune est un ensemble de procédures regroupées par thème;

unité	thème
AIDE	aide, titres
UTILAFF	routines d'affichage (couleur, cadres...)
UTILIMP	routines d'impression à l'imprimante
UTILCHOI	routines de choix dans des menus
DIRLISTE	affichage d'une liste de fichiers lus sur disque
KI2_F	calcul des probabilités du Ki2 et de Fisher
STUDENT	calcul des probabilités de Student
DECLAR_G	déclarations de types et constantes
DECLAR	déclarations des variables globales d'AGROSTAT
UTIL_G	utilitaires divers (tri, sup, inf...)
UTIL	utilitaires d'affichage des résultats
UTIL_MOD	utilitaires utilisés par MOD_STAT
PARESSAI	saisie des caractéristiques de l'essai
INIT	initialisation des variables globales
MENUS	menus proposés à l'utilisateur
UTILGES1	affichage et déplacement dans la grille de saisie
UTILGES2	routines pour les fichiers grilles de saisie
GESTION	lecture et écriture sur fichiers données
LIST	affichage de la liste des paramètres
SAIS_DER	gestion des paramètres dérivés (calculs)
PROC_GEN	procédures générales (appel menus, aiguillages)
STATGEST	gestion des données (procédure principale)
MOD_STAT	création modèle statistique et calculs

-Listing des principales unités de programme:

PROGRAMME PRINCIPAL: il définit le type de traitement à effectuer.

```
program AGROSTAT;
```

```
uses crt,declar_g,declar,utilaff,utilchoi,dirliste,aide,
    menus,init,statgest,mod_stat;
```

```
var
```

```
    choix,nb,n    : integer;
    fin, finl     : boolean;
    ess           : vmenu;
    tab_choix     : ptab_bool;
    nom_sup       : tsup;
    mod_sup       : tmod;
    current_dir   : string;
```

```
procedure en_tete;
```

```
(...)
```

```
function calc_annee_fich( nom : chain80; num : integer;
                          complet : boolean ) : integer;
```

```
(...)
```

```
procedure select_choix( nom : chain80; var choisi : boolean );
```

```
(...)
```

```
procedure deselect_choix( nom : chain80; var choisi : boolean );
```

```
(...)
```

```
procedure select_annee( complet : boolean ) : integer;
```

```
(...)
```

```
begin
```

```
    en_tete;
```

```
    disk := lecteur;
```

```
    fin := FALSE;
```

```
    repeat
```

```
        clrscr;
```

```
        case type_traitement of
```

```
            0 : begin (* traitement d'un essai *)
```

```
                n_fich := 1;
```

```
                titre_expe := '';
```

```
                getdir( 0,current_dir );
```

```
                gest_main;
```

```
                chdir( current_dir );
```

```
            end;
```

```
            1 : begin (* regroupement de plusieurs essais *)
```

```
                finl := FALSE;
```

```
                repert := affiche_liste( disk + DATAPATH, '*',
                                          Directory, FALSE );
```

```
                repeat
```

```
                    clrscr;
```

```
                    case menu_regroup of
```

```
                        0 : begin (* regroupement par CAMPAGNE *)
```

```
                            clrscr;
```

```
                            n_fich := select_annee( TRUE );
```

```
                            init_modele( 0,1 );
```

```
                            init_car_adv( TRUE );
```

```

        titre_expe := 'Modèle Campagnes';
        calcul{};
    end;
1 : begin (* regroupement par ANNEE/SAISON *)
    n_fich := select_annee( FALSE );
    init_modele( 1,2 );
    init_car_adv( TRUE );
    titre_expe := 'Modèle Années/Saisons';
    calcul{};
end;
2 : begin (* regroupement par LOCALITE *)
    clrscr;
    writeln('Regroupement d''essais par
            localité');
    writeln;
    writeln('Cette option n''est pas
            disponible pour l''instant. ');
    writeln('Le programme doit être modifié
            (voir MAIN.PAS)');
    arret_defil{};
end;
3 : begin (* regroupement par un autre modèle à
            déterminer *)
    clrscr;
    writeln('Regroupement d''essais par modèle
            à définir');
    writeln;
    writeln('Cette option n''est pas
            disponible pour l''instant. ');
    writeln('Le programme doit être modifié
            (voir MAIN.PAS)');
    arret_defil{};
end;
4 : fin1 := TRUE;
end;
until ( fin1 = TRUE );
end;
2 : fin := TRUE;
end;
until ( fin = TRUE );
clrscr;
end.

```

PROGRAMME PRINCIPAL DE GESTION DES DONNEES: il définit les fichiers de données à utiliser et les opérations à effectuer (gestion des données, liste des paramètres, calculs).

```
unit statgest;
```

```
interface
```

```
uses dos,crt,declar_g,utilaff,utilchoi,dirliste,aide,
    declar,init,paressai,menus,proc_gen,mod_stat;
```

```
procedure gest_main;
```

```
implementation
```

```

procedure gest_main;
var info      : SearchRec;
    ch        : chain255;
    nouv      : boolean;
    erreur,
    i,
    j          : integer;
begin
  clrscr;
  nouv := FALSE;
  window(1,1,80,25);
  clrscr;
  gotoxy(17,1);
  textcolor(13);
  write('LISTE DES ETUDES EN COURS OU DEJA EFFECTUEES');
  aide10;
  repert := affiche_liste( disk + DATAPATH, '*', Directory, TRUE );
                (* affiche la liste des répertoires *)

  clrscr;
  if repert = 'NOUVE' then
  begin
    repeat
      repeat
        gotoxy(5,12);
        write('nom générique du sous-répertoire correspondant à
              l''essai');
        gotoxy(6,14);
        clreol;
        cadre(5,13,11,15);
        lire_chaine(6,14,5,ch);
        repert:=ch;
      until repert <> '';
      {$I-}
      mkdir(disk+DATAPATH + '\\' + repert);
      (* crée un sous-répertoire qui contiendra les fichiers de
      l'application*)
      {$I+}
      erreur:=IOResult;
      if erreur<>0 then
      begin
        beep;
        gotoxy(10,20);
        write('ce nom a déjà été utilisé...');
        delay(2000);
        gotoxy(10,20);
        clreol;
      end;
    until erreur=0;
  end;
  clrscr;
  gotoxy(17,1);
  textcolor(13);
  write('LISTE DES CAMPAGNES EFFECTUEES SUR ', repert);
  aide10;
  nomfich := affiche_liste( disk + DATAPATH + '\\' +
                            repert, '*.DON', Archive, TRUE );
  (* affiche la liste des applications *)

```

```

clrscr;
if nomfich = 'NOUVEAU ' then
begin
(* saisit les paramètres d'une nouvelle application *)
nouv := TRUE;
i:=1;
repeat
  clrscr;
  gotoxy(5,8);
  writeln('titre de l''étude (ligne ',i,')');
  gotoxy(1,23);
  writeln('la ligne 1 doit contenir le titre principal, les 4
    suivantes les compléments');
  write('et sous-titres éventuels. Tapez FIN après la
    dernière ligne entrée');
  cadre(5,9,75,11);
  lire_chaine(6,10,69,ch);
  titre[i]:=ch;
  i:=i+1;
until (ch='FIN') or (ch='fin') or (i=6);
for j:=i to 5 do titre[j]:=' ';
repeat
  repeat
    gotoxy(5,12);
    write('nom à donner aux fichiers de données');
    gotoxy(6,14);
    clreol;
    cadre(5,13,14,15);
    gotoxy(6,14);
    write(repert);
    lire_chaine(6+length(repert),14,8-length(repert),ch);
    nomfich:=repert + ch;
  until nomfich<>repert;
  FindFirst(disk+DATAPATH + '\\' +repert + '\\' +
    nomfich+'.PAR',Archive,info);
  if DosError = 0 then
  begin
    beep;
    gotoxy(10,20);
    write('ce nom a déjà été utilisé...');
    delay(2000);
    gotoxy(10,20);
    clreol;
  end;
until ( DosError = 2 ) or ( DosError = 18 );
type_essai; (* charge les paramètres dans un fichier *)
grille_essai( TRUE,TRUE ); (* création du plan de l'essai *)
(* initialise les fichiers .DON et .PAR par un rewrite *)
assign(fich1,disk+DATAPATH + '\\' +repert + '\\' +nomfich+'.DON');
assign(fich3,disk+DATAPATH + '\\' +repert + '\\' +nomfich+'.PAR');
rewrite(fich1);
rewrite(fich3);
close(fich1);
close(fich3);
end;
init_car_adv( nouv );
FindFirst( disk + DATAPATH + '\\' + repert + '\\' + repert +
  '.GR1',Archive,info );

```

```

if ( DosError <> 0 ) and ( not nouv )
then grille_essai( TRUE,not nouv )
else begin
  grille_essai( FALSE,not nouv );
  nomgrille1:=repert +'.GR1';
  nomgrille2:=repert +'.GR2';
end;
nomfich3:=nomfich+'.PAR';
assign( fich3,disk + DATAPATH + '\ ' + repert + '\'+ nomfich3 );
nomfich1:=nomfich+'.DON';
assign( fich1,disk + DATAPATH + '\ ' + repert + '\'+ nomfich1 );
while( TRUE ) do begin
  normal;
  case Menu_Principal{} of
    0: menu_gestion{};
    1: liste_var{};
    2: begin
        init_modele( 10,0 );
        calcul{};
      end;
    3: begin
        clrscr;
        exit;
      end;
  end;
end;
end;
end.

```

PROGRAMME DE GESTION DES DONNEES: il définit les opérations à effectuer sur les données (saisie, modification, consultation, paramètres dérivés).

```

unit proc_gen;

interface

uses crt,declar_g,utilaff,utilimp,ki2_f,aide,util_g,
  declar,menus,utilges1,gestion,list,sais_der;

procedure liste_var;
procedure menu_gestion;

implementation

type trecalcul = ( TOTAL,PARTIEL );

var car : char;

procedure saisie;
(...);
procedure consultation;
(...);
procedure modification;
(...);
procedure liste_var;

```

```

(...)
procedure sais_derive( recalcul : trecalcul );
(...)
procedure menu_gestion;
begin
  while (TRUE) do begin
    titre1{};
    case menu_sais{} of
      0 : saisie{};
      1 : consultation{};
      2 : modification{};
      3 : sais_derive( TOTAL );
      4 : sais_derive( PARTIEL );
      5 : exit;
    end;
  end;
end;

end.

```

PROGRAMME DE CALCUL STATISTIQUE: il crée le modèle mathématique et exécute les calculs et tests statistiques.

```

unit mod_stat;

interface

uses crt,declar_g,declar,aide,utilaff,utilimp,util,util_g,
    util_mod,student,ki2_f,init;

procedure calcul;

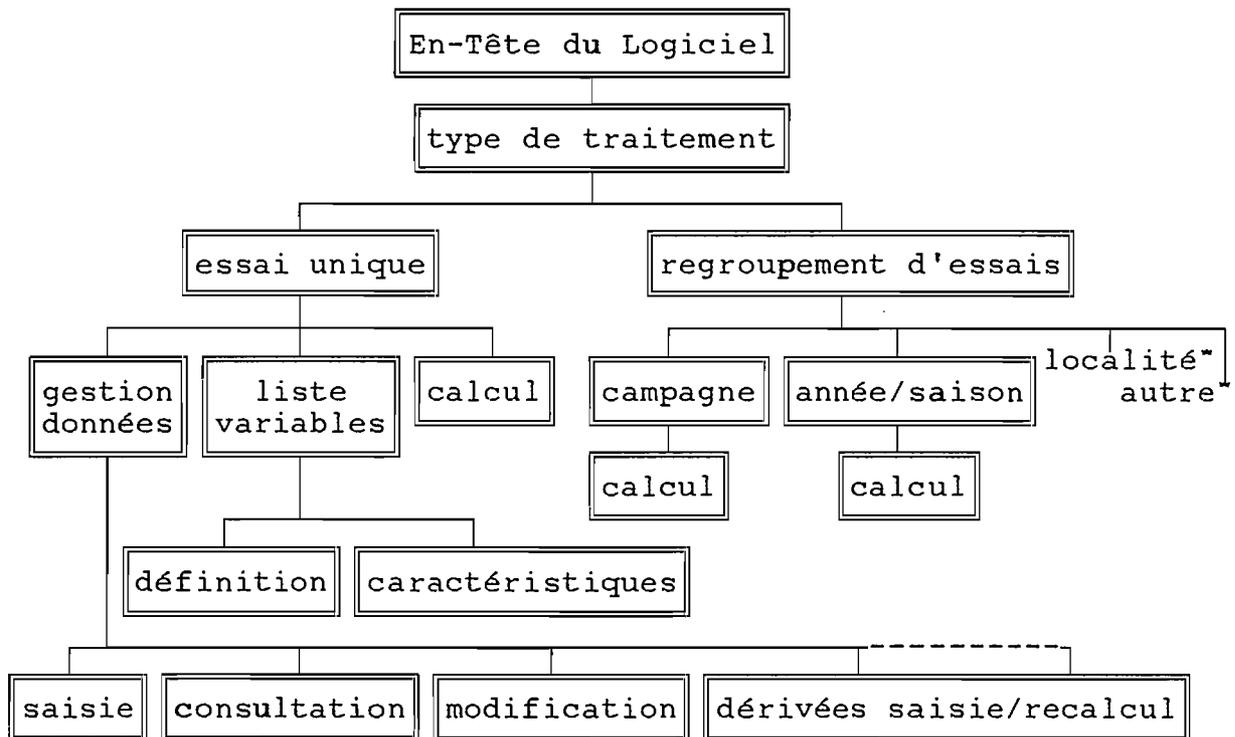
implementation
(...)
procedure calcul;
begin
  clrscr;
  assign( fich_ind,disk+DATAPATH +'\' +repert+'\' +repert+'.GR1' );
  cons_modele;
  choix_var( var_debut,var_fin );
  imprime := impression;
  for numvar := var_debut to var_fin do begin
    read_param( n_fich,numvar );
    clrscr;
    titre9;
    gotoxy( 18,15 );
    write( ' VARIABLE EN COURS DE TRAITEMENT ' );
    normal;
    ecran_param( numvar );
    if lit_numero( n_fich,numvar ) then begin end;
    read_data( n_fich,num_dat );
    reset( fich_ind );
    calc_moyennes;
    reset( fich_ind );
    calc_variances( varresH,ddlrH );
    titre_experimentation( n_fich,titre_expe );
    imp_adv;
  end;
end;

```

```
imp_param( numvar );
if n_fich = 1 then aff_donnees( numvar );
aff_moy;
titre_experimentation( n_fich,titre_expe );
imp_adv;
imp_param( numvar );
homogene := homogene_var;
calcul_fisher( varresH,homogene );
aff_fisher( ddlrH,homogene );
clrscr;
titrell;
centre( ' VARIABLE EN COURS DE TRAITEMENT ',15 );
normal;
ecran_param( numvar );
compare( numvar,varresH,ddlrH );
close( fich_ind );
free_data{};
end;
free_modele{};
end;

end.
```

-Organisation du logiciel par le choix des menus:



* ces possibilités ne sont pas opérationnelles dans l'état actuel du logiciel.

B GESTION DES FICHIERS DE DONNEES PAR LE LOGICIEL AGROSTAT

Le logiciel AGROSTAT crée et utilise des fichiers contenant les valeurs des mesures effectuées lors d'un essai expérimental. Ces fichiers sont placés dans un sous-répertoire du disque choisi nommé DATA.

Le disque usuel sera un disque dur (C en général), mais si l'on veut utiliser des fichiers stockés sur disquette, il faudra s'assurer de l'existence du sous-répertoire DATA (le créer au besoin) avant toute utilisation du logiciel.

Lors de la mise en place d'un nouvel essai, le logiciel demande le nom générique de l'essai (sur 5 caractères), et crée un sous-répertoire à ce nom, qui contiendra les fichiers relatifs à cet essai.

exemple:

essai amendement calcique sur vertisol hyper-magnésien
nom: ACVHM

Après la saisie des paramètres de l'essai (modèle, nombre de facteurs...), le programme crée 3 fichiers (format ASCII) de nom le nom générique choisi et d'extension ".CAR", ".GR1", ".GR2", qui contiennent respectivement les caractéristiques du modèle, la grille de saisie selon le plan de l'essai, et la grille de saisie selon l'ordre des facteurs.

Pour chaque répétition de l'essai, dans le temps et/ou dans l'espace, le programme va créer 2 fichiers structurés;

l'un contenant les caractéristiques de la mesure effectuée, (extension ".PAR"):

nom de la variable,
date,
unité,
définition,
éventuellement formule de calcul,

l'autre contenant les valeurs de cette mesure (extension ".DON").

Le nom de ces fichiers sera constitué du nom générique de l'essai (5 caractères), plus l'année de réalisation du fichier (2 caractères), plus éventuellement le lieu de l'expérimentation (1 caractère).

exemple:

ACVHM87.PAR
ACVHM87.DON

Il est recommandé de ne pas mettre dans le répertoire DATA et dans les répertoires de données d'autres répertoires ou fichiers inaccessibles par AGROSTAT. En effet, lors du lancement de l'application, le programme recherche sur le disque les répertoires contenant les fichiers des différents essais et propose la liste à l'utilisateur pour y faire son choix. Celle ci ne doit bien entendu contenir que les noms des fichiers créés par le logiciel.

Imprimé par le Centre ORSTOM
de Nouméa
Juillet 1989

ORSTOM ORSTOM
ORSTOM ORSTOM

Centre ORSTOM de Nouméa
B.P. A5 Nouméa Cédex Nouvelle-Calédonie

© 1989