

CENTRE O.R.S.T.O.M. DE NOUMÉA
LABORATOIRE D'AGRONOMIE

R A P P O R T D ' A C T I V I T É .

Fonds Documentaire IRD
Cote : B* 25726 Ex :

H = 54/100

~~N° AG 1/10/1/1/1~~

→ ~~2010/1/1/1/1~~

Jean-Luc JICQUEL
V.A.T.
Décembre 83 à Mars 85.

Fonds Documentaire IRD



010025726

Centre O.R.S.T.O.M. de NOUMEA
Laboratoire d'Agronomie



R A P P O R T D ' A C T I V I T E .
=====

Fonds Documentaire IRD
Cote : ~~B~~* 25726 Ex : *unipe*

Jean-Luc JICQUEL
V.A.T.
Decembre 83 à
Mars 85.



Vue générale des essais en serre

Le laboratoire d'Agronomie du Centre O.R.S.T.O.M. de Nouméa étudie la fertilité des sols de Nouvelle-Calédonie. Ce programme de recherche est mené en étroite collaboration avec le laboratoire de Pédologie et les principales recherches entreprises, le sont dans le cadre de conventions passées avec la Direction du Développement de l'Economie Rurale (Service de la Recherche, de la Formation et de la Diffusion).

Ces 16 mois ont été pour moi l'occasion de participer activement à la vie d'un laboratoire de recherche agronomique outre-mer. Je n'insisterai que sur quelques tâches qui m'ont été confiées et dont j'ai eu la responsabilité :

- un micro-ordinateur ayant été acquis par le laboratoire peu après mon arrivée, j'ai eu à concevoir entièrement le traitement informatique de nos données expérimentales. Ce travail a abouti à la mise au point d'un ensemble de programmes permettant de réaliser des analyses statistiques plus fines et entièrement automatiques ;

- j'ai assuré le suivi des essais en plein champ sur vertisol hypermagnésien à la TAMOA et sur sol sodique acide à POUEMBOUT pendant les 4 mois d'absence de Monsieur BONZON.

- j'ai pu également concevoir et réaliser deux expérimentations en serre sur vertisol hypermagnésien concernant la première, l'influence de 3 doses de CaO de 3 amendements calcaïques sur une culture de maïs, et la seconde l'influence comparée de 2 engrais azotés sur culture de maïs ayant reçu 5 tonnes de CaO/ha sous 3 formes différentes.

I

LES ESSAIS

Les essais auxquels j'ai participé concernaient l'étude des effets des amendements calcaïques sur les sols de Nouvelle-Calédonie : vertisol hypermagnésien de la TAMOA, sol sodique acide de POUEMBOUT.

Il s'agit d'essais en serre et d'essais au champ.

Au cours de ces essais, la croissance des plantes est suivie pendant environ les deux premiers mois du cycle végétatif (mesures de hauteur, calculs de vitesse de croissance). Lors de la récolte, les différentes composantes du rendement sont mesurées et des analyses sont effectuées sur le matériel végétal récolté (teneurs en éléments minéraux) et sur des échantillons de sol (caractéristiques physiques et chimiques) en début et fin de cycle.

Mon rôle a consisté non seulement en une participation à ces observations mais aussi, et surtout en l'interprétation statistique des données.

I - LES ESSAIS EN POTS SOUS SERRE (SUR MAIS).

La technique mise au point à NOUMEA pour caractériser certaines aptitudes des sols, présente deux avantages :

- la rapidité des tests et leur faible coût ;
- la possibilité de fixer les conditions expérimentales de façon à ce que l'influence des paramètres secondaires ne gêne pas les essais et ne les rende pas difficilement interprétables.

La première étude à laquelle nous avons participé (décembre 1983 - janvier 1984) était constituée par :

- un essai soustractif dont le but était la mise en évidence de carences. Furent trouvées significatives des carences en azote, phosphore et molybdène ;
- un essai d'amendement calcaïque avec du gypse : qui a montré que l'on était effectivement devant une forte carence en calcium.

La deuxième (mars 1984) et la troisième études (juillet 84) étaient des essais d'amendements calcaïques. Ces essais ont permis de déterminer le "meilleur" mélange gypse-carbonate de calcium.

C'est cette dernière expérimentation que j'ai conçue et dirigée : 3 doses de CaO (5, 10 et 15t/ha) dont 10 %, 25 % ou 40 % étaient apportés sous forme de gypse (le reste étant du carbonate de calcium).

L'optimum semble avoir été approché par le mélange 10 tonnes de CaO avec 25 % du calcium apportés sous forme de gypse.

Ces expérimentations en serre ont permis de fixer les quantités qui seraient apportées sur l'essai au champ (juillet 84).

Voici les listings des analyses de variance pour la variable poids de matière sèche des tiges et feuilles pour les 2 expérimentations.



Influence du gypse sur la croissance en hauteur du maïs (de gauche à droite 0t/ha, 15t/ha, 30t/ha : étude expérimentale en serre de décembre 83 - janvier 84).

INFLUENCE DE TROIS DOSES DE CAO DE TROIS AMENDEMENTS CALCIFIQUES
SUR UNE CULTURE DE MAIS SUR VERTISOL HYPERMAGNESIEN

Analyse de variance

Date: 2-9-84

Parametre: PMSTF

Unite: g

	Bloc 1	Bloc 2	Bloc 3	Bloc 4	Bloc 5	Bloc 6
X 11k	13.070	14.120	11.700	16.520	13.570	13.450
X 12k	14.910	12.890	15.070	16.850	14.970	15.970
X 13k	7.480	7.890	9.320	9.640	6.920	11.320
X 21k	10.630	15.740	16.930	12.590	14.920	16.960
X 22k	18.870	17.040	16.700	16.830	15.720	17.970
X 23k	12.950	12.670	15.410	13.350	15.070	15.720
X 31k	16.570	15.860	15.390	12.130	18.100	15.700
X 32k	18.990	12.340	13.460	15.490	13.830	16.320
X 33k	18.270	12.420	17.140	14.820	16.460	19.930

Moyennes et ecartis relatifs :

X... = 14.536

X1.. = 12.537 (-13.76) X2.. = 15.337 (5.51) X3.. = 15.734 (8.24)

X..1 = 14.664 (0.88) X..2 = 15.790 (8.63) X..3 = 13.154 (-9.51)

X..1 = 14.638 (0.70) X..2 = 13.441 (-7.53) X..3 = 14.569 (0.23)

X..4 = 14.247 (-1.99) X..5 = 14.396 (-0.97) X..6 = 15.927 (9.57)

X11. = 13.738 (-5.49) X12. = 15.110 (3.95) X13. = 8.762 (-39.72)

X21. = 14.628 (0.63) X22. = 17.188 (18.25) X23. = 14.195 (-2.35)

X31. = 15.625 (7.49) X32. = 15.072 (3.68) X33. = 16.507 (13.56)

Variances et coefficients de variations

Se1 = 10.654 Se2 = 9.063 Se3 = 2.474

cv1 = 22.455 cv2 = 20.710 cv3 = 10.821

SA = 109.360 SD = 62.956 SB = 11.691

Teste F :

FA = 10.265 (0.996) FD = 6.946 (0.987) FAD = 21.952 (1.000)

Test de BARTLETT : K12 = 8.621

L'hypothese de l'egalite des variances residuelles est rejete
dans X_{ijk}

i est le pourcentage de Calcium apporté sous forme de gypse 10 %, 25 %, 40 %

j est la quantité de CaO en T/ha 5t, 10t, 15t.

INFLUENCE COMPARE DE 2 ENGRAIS AZOTE SUR CULTURE DE MAIS
 AYANT RECU 5 TONNES DE CaO/ha SOUS 3 FORMES DIFFERENTES

Analyse de variance

Date: 2-8-84

Parametre: PMSTF

Unite: g

	Bloc 1	Bloc 2	Bloc 3	Bloc 4	Bloc 5	Bloc 6
X 11k	13.070	14.120	11.700	16.520	13.570	13.450
X 12k	16.490	14.400	15.070	16.410	14.130	13.490
X 21k	10.630	15.740	16.930	12.590	14.920	16.960
X 22k	13.270	14.410	13.240	15.410	9.640	13.310
X 31k	16.570	15.860	15.390	12.130	18.100	15.700
X 32k	14.370	12.390	13.310	13.490	17.710	15.720

Moyennes et ecartes relatifs :

$$X_{...} = 14.456$$

$$X_{1..} = 14.368 (-0.61) \quad X_{2..} = 13.937 (-3.59) \quad X_{3..} = 15.062 (4.19)$$

$$X_{.1.} = 14.664 (1.44) \quad X_{.2.} = 14.249 (-1.44)$$

$$X_{..1} = 14.067 (-2.69) \quad X_{..2} = 14.467 (0.21) \quad X_{..3} = 14.273 (-1.26)$$

$$X_{..4} = 14.425 (-0.21) \quad X_{..5} = 14.712 (1.77) \quad X_{..6} = 14.772 (2.18)$$

$$X_{11.} = 13.738 (-4.96) \quad X_{12.} = 14.998 (3.75)$$

$$X_{21.} = 14.628 (1.19) \quad X_{22.} = 13.247 (-8.36)$$

$$X_{31.} = 15.625 (8.09) \quad X_{32.} = 14.498 (0.29)$$

Variances et coefficients de variations

$$Se1 = 12.446 \quad Se2 = 5.687 \quad Se3 = 5.907$$

$$cv1 = 24.404 \quad cv2 = 16.496 \quad cv3 = 16.813$$

$$SA = 7.720 \quad SD = 3.117 \quad SB = 0.844$$

Tests F :

$$FA = 0.620 (0.439) \quad FD = 0.548 (0.497) \quad FAD = 2.157 (0.834)$$

$$\text{Test de BARTLETT : } K_{12} = 1.672$$

on accepte l'hypothese de l'egalite des variances residuelles

$$Se = 8.479 \quad cv = 20.143$$

$$F_{1A} = 0.911 (0.418) \quad F_{1D} = 0.368 (0.444) \quad F_{1AD} = 1.503 (0.759)$$

dans X_{ijk}

i est le pourcentage de calcium apporté sous forme de gypse 10 %, 25 %, 40 %

j est la nature de l'engrais azoté, nitrate d'amoniaque ou sulfate d'amoniaque.

II - LES ESSAIS AU CHAMP.

Ces essais sont réalisés en collaboration avec le Centre de Recherche et d'Etudes Agronomiques des Services Ruraux Territoriaux (C.R.E.A.).

Les trois champs d'essais sont équipés d'un réseau d'irrigation d'appoint, indispensable compte tenu des périodes de sécheresse qui peuvent survenir pendant la croissance des plantes.

Les semis ont lieu généralement en juin, parfois début juillet et les récoltes s'échelonnent de novembre à décembre ; avant la période des cyclones. Les trois champs sont situés sur la côte Ouest de la Grande Terre, zone où l'on trouve les principales cultures céréalières de l'île.

a) - Sur vertisol (Pouembout) ; Plante-test : maïs.

Ce type de sol est souvent appelé "Argiles noires" en Nouvelle-Calédonie. Il couvre d'importantes surfaces sur le Côte Ouest de la Grande Terre.

Cet essai n'a pu être mis en place cette année en raison du prolongement de la saison des pluies.

b) - Essai Amendement Calcique sur Sol Sodique Acide (Pouembout) ; Plante-tests : Maïs grain, Maïs fourrage et haricot.

Cet essai fut mis en place en 1981 à Pouembout, sur un "pâturage naturel à repousses de niaoulis".

Sont testées 4 doses de CaO (0, 2, 4 et 6 t/ha). Cette année le traitement subsidiaire prévu au départ a été utilisé pour tester l'effet d'un amendement calcique en profondeur (2t/ha de CaO).

Ce traitement n'a pas influencé en apparence les rendements. Une forte attaque de rouille explique sans doute en partie l'égalisation des rendements.

c) Essai amendement calcique sur vertisol hypermagnésien ; Plante-test : Maïs.

Cet essai fut mis en place en 1984. A la suite des essais en serre, les facteurs suivants ont été choisis : doses de CaO: 4, 8 et 12 t/ha.

teneurs en gypse : 0 % et 10 %.

Une bande de maïs sans amendement mais avec fertilisation a été conservée. Les résultats sont spectaculaires (voir l'analyse de variance pour la variable poids de grain par pied de référence).

Pour le résultat X_{ijk} :

i est le pourcentage de calcium sous forme de gypse
(i = 1 pour 0 % et i = 2 pour 10 %);

j est la quantité de CaO apportée à l'hectare
(j = 1 pour 4 tonnes, j = 2 pour 8 tonnes et j = 3 pour 12 tonnes) .

La liste des variables mesurées ou calculées au 1er avril 1985 figure également en annexe (il manque les résultats de l'analyse complète de sol de juin 1984).



Vue générale du champ expérimental sur vertisol hypermagnésien (vallée de la TAMOA) juste après l'enfouissement du pâturage naturel en avril 84.

ANALYSE DE LA VARIANCE

CRITERES RETENUS	DISPOSITIFS EXPERIMENTAUX CONSIDERES		
	Blocs pris sur les lignes	Blocs pris sur les colonnes	Carre latin
VARIANCES RESIDUELLES			
VR1 (cv1%)	1.43902E+02 (10.6)	2.48305E+02 (13.9)	1.62806E+02 (11.3)
VR2 (cv2%)	5.95279E+01 (6.8)	5.30659E+01 (6.4)	
VR3 (cv3%)	1.05149E+02 (9.1)	1.06442E+02 (9.1)	
HOMOGENEITE DES V.R.			
KI 2 DE BARTLETT	1.45274E+00	5.97256E+00	
PROBABILITE DU KI 2	1.00000	1.00000	
OBSERVES ET PROBABILITES			
- AMENDMENT	1.854 (0.817)	1.075 (0.689)	1.639 (0.787)
- DOSE	16.158 (1.000)	9.364 (0.999)	14.282 (1.000)
- AMENDMENT*DOSE	0.017 (0.016)	0.010 (0.009)	0.015 (0.014)
- SUBSIDIAIRE	1.138 (0.664)	1.276 (0.690)	
- AMENDMENT*SUBSIDIAIRE	0.731 (0.405)	0.722 (0.408)	
- DOSE*SUBSIDIAIRE	0.276 (0.236)	0.273 (0.233)	
- AMENDEM*DOSE*SUBSID	0.337 (0.279)	0.333 (0.276)	
OBSERVES ET PROBABILITES			
LES V.R. SONT HOMOGENES			
- AMENDMENT	2.299 (0.868)	1.691 (0.803)	
- DOSE	20.029 (1.000)	14.730 (1.000)	
- AMENDMENT*DOSE	0.021 (0.020)	0.016 (0.015)	
- SUBSIDIAIRE	0.583 (0.455)	0.429 (0.478)	
- AMENDMENT*SUBSIDIAIRE	0.662 (0.425)	0.487 (0.495)	
- DOSE*SUBSIDIAIRE	0.250 (0.217)	0.184 (0.166)	
- AMENDEM*DOSE*SUBSID	0.305 (0.258)	0.225 (0.198)	

VARIABLES AC/VHM TAMDA

NO	PARAMETRE	DATE	UNITE	FORMULE	DEFINITION
0	H 38	12/09/84	cm	-	Hauteur au 38 ^{eme} jour
1	H 45	19/09/84	cm	-	Hauteur au 45 ^{eme} jour
2	H 52	26/09/84	cm	-	Hauteur au 52 ^{eme} jour
3	H 59	03/10/84	cm	-	Hauteur au 59 ^{eme} jour
4	H 66	10/10/84	cm	-	Hauteur au 66 ^{eme} jour
5	H 72	16/10/84	cm	-	Hauteur au 72 ^{eme} jour
6	V 38-45	19/09/84	cm/j	(6) = ((1) - (0)) / 7	Vitesse de croissance entre le 38 ^{eme} et 45 ^{eme} jour
7	V 45-52	26/09/84	cm/j	(7) = ((2) - (1)) / 7	Vitesse de croissance entre le 45 ^{eme} et 52 ^{eme} jour
8	V 52-59	03/10/84	cm/j	(8) = ((3) - (2)) / 7	Vitesse de croissance entre le 52 ^{eme} et 59 ^{eme} jour
9	V 59-66	10/10/84	cm/j	(9) = ((4) - (3)) / 7	Vitesse de croissance entre le 59 ^{eme} et 66 ^{eme} jour
10	V 66-72	16/10/84	cm/j	(10) = ((5) - (4)) / 6	Vitesse de croissance entre le 66 ^{eme} et 72 ^{eme} jour
11	DL	19/12/84	nbre pied/m ²	-	Densite de peuplement
12	NOTE	03/09/84	note subj.	-	Note subjective sur l'aspect de la vegetation
13	NER	19/12/84	-	-	Nombre d'epis des pieds de reference
14	PGERP	19/12/84	g	-	Poids de grains secs des pieds de reference
15	PRERP	19/12/84	g	-	Poids des rachis des pieds de reference
16	PTFRF	19/12/84	g	-	Poids frais total des tiges et feuilles des pieds de reference
17	PTFRF	19/12/84	g	-	Poids frais des tiges et feuilles des pieds de reference (echantillon)
18	PTFRS	19/12/84	g	-	Poids sec des tiges et feuilles des pieds de reference (echantillon)
19	NEPR	19/12/84	nbre epis/plt	(19) = (13) / 18	Nombre d'epis par pied de reference
20	PSR	19/12/84	g/plt	(20) = (14) / 18	Poids de grain sec par pied de reference
21	PGER	19/12/84	g/epi	(21) = (14) / (13)	Poids de grain sec par epi des pieds de reference
22	PSR	19/12/84	g/plt	(22) = (15) / 18	Poids de rachis sec par pied de reference
23	PRER	19/12/84	g/epi	(23) = (15) / (13)	Poids de rachis sec par epi sur les pieds de reference
24	TGE	19/12/84	%	(24) = 100 * (14) / ((14) + (15))	Taux de remplissage des grains sur les pieds de reference
25	PTFR	19/02/84	g/plt	(25) = ((16) / 18) * 1 / (((17) - (18)) / (18) + 1)	Poids sec de tiges et feuilles sur les epis de references
26	NPUR	19/12/84	-	-	Nombre de pieds utiles restants
27	BSR	19/12/84	g/m ²	(27) = ((14) * (11)) / 18	Rendement en grain estime a partir des pieds de reference

30	NEURA	03/01/85	nbre epi	-	Nombre d'epis attaques des pieds utiles restants
31	NEURO	03/01/85	nbre epi	$(31)=(30)+(29)$	Nombre d'epis total des pieds utiles restants
32	PEURS	03/01/85	kg	-	Poids d'epis sains des pieds utiles restants
33	PEURA	03/01/85	kg	-	Poids d'epis attaques des pieds utiles restants
34	PEURT	03/01/85	kg	$(34)=(33)+(32)$	Poids total d'epis des pieds utiles restants
35	PGUF	06/02/85	g	-	Poids des grains frais des pieds utiles restants
36	PGUR	6/02/85	g/plt	$(36)=(35)/(26)$	Poids de grain frais par pied utile restant
37	HG	6/02/85	%	-	Humidite des grains par pied utile restant
38	NEV	03/01/85	epi vole	-	Nombre d'epis voles sur les pieds utiles restants par parcelle
39	QG	06/01/85	g/m2	$(39)=(((14)+(41))*(26))/(((26)-(38))*20.25)$	Rendement estime sur les pieds utiles restants par parcelle
40	PG	06/01/85	g/plt	$(40)=(((14)+(41))*(26))/(((26)+18)*((26)-(38)))$	Rendement en grains par pied utile restant
41	PGUS	06/02/85	g	$(41)=(35)/(((37)/100)+1)$	Poids de grains secs des pieds utiles restants
42	CATF	01/01/85	%	-	pourcentage de calcium dans les tiges et feuilles
43	MGTF	01/01/85	%	-	pourcentage de magnesium des tiges et feuilles
44	NTF	01/01/85	%	-	pourcentage d'azote dans les tiges et feuilles
45	P205TF	01/01/85	%	-	pourcentage de phosphore dans les tiges et feuilles
46	KTF	01/01/85	%	-	pourcentage de potassium dans les tiges et feuilles
47	BCATFR	01/01/85	g/m2	$(47)=((25)*(11)*(42)/100)$	immobilisation de Ca/m2 dans les tiges et feuilles calculee a partir des pieds de references
48	BMGTFR	01/01/85	g/m2	$(48)=((25)*(11)*(43)/100)$	immobilisation de Mg/m2 dans les tiges et feuilles calculee sur les pieds de reference
49	BNTFR	01/01/85	g/m2	$(49)=((25)*(11)*(44)/100)$	immobilisation de N/m2 dans les tiges et feuilles calculee sur les pieds de reference
50	BP25FR	01/01/85	g/m2	$(50)=((25)*(11)*(45)/100)$	immobilisation de P205/m2 dans les tiges et feuilles calculee sur les pieds de reference
51	BKTFR	01/01/85	g/m2	$(51)=((25)*(11)*(46)/100)$	immobilisation de K/m2 dans les tiges et feuilles calculee sur les pieds de reference

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70

II

PROGRAMMATION DES CALCULS STATISTIQUES SUR LE
GOUPIIL 3.

Les programmes que j'ai réalisés sur GOUPIIL 3 concernent l'analyse statistique des résultats observés sur :

- . une expérimentation du type à parcelle subdivisée une fois
C'est le cas de l'essai d'amendement calcique sur solonetz solodisé. Les facteurs principaux sont le facteur dose (0,2,4,6 t/ha de CaO). et le facteur subsidiaire est un apport d'amendement calcique en profondeur (2 t/ha de CaO).
- . une expérimentation du type factoriel 3^3 à 2 répétitions.
C'est le cas de l'essai fertilisation nitro-phospho-potassique du maïs sur vertisol. Les facteurs principaux sont le facteur azote (niveau 1,2,3) le facteur phosphore (1,2,3) et le facteur potasse (1,2,3).
- . un carré latin à parcelles subdivisées.
C'est le cas de l'essai amendement calcique sur vertisol hyper-magnésien.
- . une expérimentation factorielle à 2 critères de classification (3x4 ou 3x3) avec 6 répétitions.
La plupart des essais en serre en 1984 étaient de ce type.

La programmation des essais en serre m'a permis de me familiariser avec le GOUPIIL et d'améliorer à chaque fois la technique de programmation notamment l'organisation et de déroulement des calculs.

Le dernier programme que j'ai mis au point est celui de l'analyse du carré latin à parcelles subdivisées. Il bénéficie de toutes les améliorations successives de ces 12 mois et propose la nomenclature indispensable des fichiers susceptible d'être adoptée définitivement. D'autre part ce programme, dont les estimations des variances sont les plus complexes, démontre que les techniques de calcul des moyennes et des variances adoptées sont bien au point. Pour toutes ces raisons, j'ai choisi de le prendre en exemple.

Je rappellerai d'abord les caractéristiques du GOUPIIL :

- mémoire : 64K (+ carte 256K)
- processeur : 6809 FLEX
- capacité de stockage : 1 Mo (2 disquettes de 8")
- mémoire graphique : (512 x 256)

D. FARNIER responsable informatique du Centre avait tout de suite attiré l'attention des laboratoires sur une spécificité du GOUPIIL : le tableau virtuel qui est en quelque sorte un tableau sur disque. Il s'est révélé que ce type de "fichier" est effectivement de loin le mieux adapté aux calculs statistiques et je l'ai rapidement adopté pour stocker l'essentiel des données.

La carte mémoire 256K a permis d'utiliser au mieux les tableaux virtuels grâce à des accès à l'information extrêmement courts. Tous les fichiers nécessaires au programme (.BAC et .DAT) résident sur cette carte. Un inconvénient cependant : la taille des fichiers doit tenir compte de la capacité de la carte.

PLAN

1. MENU GENERAL ET PROGRAMMES CORRESPONDANTS
2. ORGANISATION DES FICHIERS DE DONNEES
3. BIBLIOTHEQUE DE SOUS PROGRAMMES
4. MISE A JOUR DES FICHIERS DE DONNEES
5. ANALYSE DE LA VARIANCE
6. COMPARAISON DE MOYENNES
7. CORRELATION SIMPLE.

BILAN ET PERSPECTIVES.

Les fichiers contenus sur le disque sont les suivants :

Catalogue du lecteur 1
Disque: ACVHM #0

Nom	Type	Taille	Date	Prt
FISH-SUB.	BAC	6	27-Fev-85	
EDIT-E	.BAC	9	5-Mar-85	
ACVHM-CS.	BAC	35	27-Mar-85	
TRAVAIL1.	BAC	37	5-Mar-85	
CHI2-SUB.	BAC	4	5-Mar-85	
STUD-SUB.	BAC	5	27-Fev-85	
BART-SUB.	BAC	6	27-Fev-85	
EDIT-I	.BAC	7	5-Mar-85	
TRAVAIL2.	BAC	41	5-Mar-85	
ACVHM-F	.BAC	7	5-Mar-85	
EDIT-F	.BAC	5	5-Mar-85	
ACVHM-SS.	BAC	17	5-Mar-85	
CALC-R	.BAC	23	27-Mar-85	
CALC-C	.BAC	48	2-Avr-85	
ACVHM-SD.	BAC	11	5-Mar-85	
ACVHM-M	.BAC	46	5-Mar-85	
MENU	.BAC	10	5-Mar-85	

Secteurs occupes=1189, libres=2763

1 - MENU GENERAL ET PROGRAMMES CORRESPONDANTS.

Après avoir introduit la disquette "ACVHM" et lancer le système, le menu suivant est affiché à l'écran.

EXPERIMENTATION AC/VHM TAMOA (MENU)	PROGRAMME CORRESPONDANT.
1. Liste des variables	
* Edition sur écran	2. EDIT-E . BAC
* Edition sur imprimante	2. EDIT-I . BAC
2. Liste du fichier de données	2. EDIT-F . BAC
3. Enregistrement de variables	
* Variable simple	2. ACVHM-SS . BAC
* Variable dérivée	2. ACVHM-SD . BAC
4. Calculs	
* Analyse de variance	2. ACVHM-P . BAC
* Corrélation simple	2. ACVHM-CS . BAC
* Comparaison de moyennes	2. ACVHM-M . BAC
5. Fin de travail.	

Après validation, le programme choisi par l'utilisateur se chaine sur 2.MENU . BAC, ex : 2.EDIT-E . BAC pour l'édition sur écran.

NB : Les programmes EDIT-E, EDIT-F, EDIT-I sont des programmes indépendants du modèle à quelques modifications près. Les terminologies -SS, -SP, -P, -CS, -M sont également susceptibles d'être adoptés définitivement.

2 - ORGANISATION DES FICHIERS DE DONNEES.

Les données sont stockées dans des tableaux virtuels dont l'avantage est l'accès immédiat aux éléments du tableau.

Ces fichiers et leurs contenus sont les suivants :

X-C.DAT : dates et unités des variables (tableau à 2 dimensions)

X-D.DAT : définitions des variables (tableau à 1 dimension)

X-F.DAT : formules de calcul des variables lorsqu'il s'agit de variables dérivées (tableau à 1 dimension)

X-P.DAT : données expérimentales mesurées ou calculées. C'est un tableau à 2 dimensions A(I,J) dans lequel j est le numéro de la variable et i la valeur mesurée ou calculée pour une des sous parcelles.

EXEMPLE : soit un essai comportant 3 facteurs

i = 1 à I, j = 1 à J, k = 1 à K.

les données seront rangées de telle sorte que

k + j x K + i x J x K soit croissant.

Cependant dans le cas d'un carré latin, cette technique n'est pas possible, on doit alors ranger les données suivant les lignes ou suivant les colonnes (j'ai choisi les lignes).

Pour retrouver le numéro de la colonne de la parcelle, on utilise un DATA qui contient dans l'ordre les numéros de colonne des parcelles.

La liste des fichiers de données stockées sur la disquette est la suivante :

Catalogue du lecteur 1
Disque: ACVHM #0

Nom	Type	Taille	Date	Prt
NEWMAN1	.DAT	47	17-Fev-84	D
NEWMAN5	.DAT	47	17-Fev-84	D
ACVHM-F	.DAT	235	28-Jun-84	D
ACVHM-C	.DAT	28	29-Dec-84	D
ACVHM-D	.DAT	56	14-Fev-85	D
ACVHM-F	.DAT	31	14-Fev-85	D
COR1	.DAT	43	2-Avr-85	

Secteurs occupes=1189, libres=2763

3 - BIBLIOTHEQUE DE SOUS-PROGRAMMES.

- Les principaux tests nécessaires en statistiques sont programmés.
Il s'agit :

- du test F : "2.FISH-SUB.BAC" qui permet le calcul de la probabilité d'un "F" observé ;
- du test de BARTLETT : "2.BART-SUB.BAC" qui permet de tester l'homogénéité de plusieurs variances ;
- du test du KHI2 : "KHI2-SUB.BAC" qui permet le calcul de la probabilité d'un KH12. ;
- du test de STUDENT : "2.STUD-SUB.BAC".

Ce sont des sous programmes que l'on charge en OVERLAY pour réduire la taille des programmes.

- Pour le test de NEWMAN et KEULS, test de comparaison de plusieurs moyennes, la programmation du calcul des probabilités est très complexe et j'ai préféré créer les deux tableaux virtuels des valeurs critiques du test :

2.NEWMAN-5.DAT pour le seuil $x = 0,05$

2.NEWMAN 1.DAT pour le seuil $x = 0,01$

Un programme d'interpolation permet de compléter les tables.

4 - MISE A JOUR DES FICHIERS DE DONNEES.

Enregistrement des variables simples : 2 ACVHM-SS.BAC

La saisie des informations se fait sur un plan de l'expérimentation qui se dessine à l'écran. L'utilisateur se déplace d'une sous parcelle à l'autre grâce aux flèches. Cette technique permet de vérifier et de corriger d'éventuelles erreurs. Elle permet d'autre part d'éviter une transcription : à la fin de la saisie les données sont rangées automatiquement dans l'ordre dans le fichier 2.ACVMH-P.DAT.

Voici un listing de la saisie de la variable pH obtenu par la commande HARDCOPY.CMD

EXPERIMENTATION AC/VHM TAMOA (SAISIE DE VARIABLES SIMPLES).					
6.35	7.5	7.1	6.35	6.7	7.35
5.95	7.4	7.2	6.25	6.9	7.65
7.05	6.5	6.4	7.5	6.7	7.2
7.0	6.6	6.7	7.3	7.1	7.55
7.25	6.6	6.8	7.35	7.45	7.4
7.05	6.7	6.9	7.0	7.7	7.3
7.1	7.2	7.5	7.65	6.2	6.1
6.9	7.2	7.6	7.2	6.5	6.5
7.0	7.5	7.7	6.5	6.1	7.4
7.35	7.2	7.4	6.1	6.55	7.5
6.65	7.85	7.1	7.4	7.4	6.55
6.5	7.5	6.6	7.1	7.65	6.8
FIN DE TRAVAIL : tapez F					

EXPERIMENTATION AC/VHM TAMOA (SAISIE DE VARIABLES SIMPLES)

NUMERO DE LA VARIABLE : ? 99
PARAMETRE : ? PH
DATE..... : ? 14/09/84
UNITE ... : ?-
DEFINITION : ? PH

Enregistrement des variables dérivés : 2 ACVHM-SD.BAC.

Ce programme permet de calculer de nouvelles variables fonctions linéaires ou non linéaires de variables déjà entrées.

ex (15) = (14)/(1)

3 - ANALYSE DE VARIANCE : ESTIMATIONS DES TERMES DU MODELE ; VARIANCE DES EFFETS DES FACTEURS CONTROLES ; TESTS F.

Modèle linéaire : $x_{ijl} = \bar{x} + a_i + v_j + (av)_{ij} + b_l + e_{ijl}$

Termes du modèle	Significations	Estimations des termes du modèle	Degrés de liberté	Variances des effets	Tests F	Observations
x_{ijl}	Valeur de X observée sur la parcelle du l ^{ème} bloc soumise aux effets des variantes i et j des facteurs contrôlés A et V.	-	-	-	-	-
\bar{x}	Moyenne générale de X	-	-	-	-	-
a_i	Effet sur X de la i ^{ème} variante de A	$a_i = \bar{x}_i - \bar{x}$	$d_A = n_A - 1$	$s_A^2 = \{n_V \cdot n_B \cdot \sum_i a_i^2\} / d_A$	$F_A = s_A^2 / s_E^2$	-
v_j	Effet sur X de la j ^{ème} variante de V	$v_j = \bar{x}_j - \bar{x}$	$d_V = n_V - 1$	$s_V^2 = \{n_A \cdot n_B \cdot \sum_j v_j^2\} / d_V$	$F_V = s_V^2 / s_E^2$	-
$(av)_{ij}$	Interaction sur X des i ^{ème} et j ^{ème} variantes des facteurs contrôlés A et V	$(av)_{ij} = \bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}$	$d_{AV} = (n_A - 1)(n_V - 1)$	$s_{AV}^2 = \{n_B \cdot \sum_{ij} (av)_{ij}^2\} / d_{AV}$	$F_{AV} = s_{AV}^2 / s_E^2$	-
b_l	Effet sur X de la l ^{ème} répétition	$b_l = \bar{x}_l - \bar{x}$	$d_B = (n_B - 1)$	$s_B^2 = \{n_A \cdot n_V \cdot \sum_l b_l^2\} / d_B$	$F_B = s_B^2 / s_E^2$	Illicite en principe
e_{ijl}	Effet sur X des facteurs aléatoires attachés à la ij ^{ème} parcelle	$e_{ijl} = \bar{x}_{ijl} - \bar{x}_{ij} - \bar{x}_l + \bar{x}$	$d_E = (n_A n_V - 1)(n_B - 1)$	$s_E^2 = \{\sum_{ijl} e_{ijl}^2\} / d_E$	-	-
g_{ijl}	Variation générale de X	$g_{ijl} = x_{ijl} - \bar{x}$	$d_G = n_A n_V n_B - 1$	$s_G^2 = \{\sum_{ijl} g_{ijl}^2\} / d_G$	$F_G = s_G^2 / s_E^2$	Illicite en principe

5 - L'ANALYSE DE LA VARIANCE. 2.ACVMH-P.BAC.

C'est un programme qui exécute les analyses de variance en série. 2.TRAVAIL1.BAC (calculs des moyennes et des variances) et 2.TRAVAIL2.BAC (tests statistiques et impression des résultats) s'exécutent alternativement.

La liste des analyses de variance est dans un tableau virtuel constitué par l'utilisateur.

Après de multiples tâtonnements, j'ai mis au point une méthode originale de calcul des moyennes et des variances.

Les calculs se déroulent en 2 temps : le calcul des moyennes puis le calcul des variances. Un système de boucles imbriquées (il y a autant de boucle que de facteurs) permet de calculer toutes les moyennes, puis toutes les variances.

Le principe est le suivant. Pour une variable donnée, chaque élément du tableau défile successivement tel que $k+Kx_j+Kx_jx_i$ (pour 3 facteurs) est croissant. Grâce aux boucles qui fonctionnent comme des compteurs, on connaît les niveaux de chacun des facteurs (c'est à dire i, j, k, \dots). Il suffit alors d'incrémenter les moyennes correspondantes avec cet élément du tableau.

La méthode est simple et l'écriture est automatique et sans risque d'erreur lorsqu'on dispose des estimations des termes du modèle.

Exemple : Analyse de la variance d'une expérimentation factorielle, à 2 critères de classification (IxJ) avec K répétitions.

Ce modèle simple permet de comprendre le système des boucles imbriquées pour le calcul des variances.

Calcul des moyennes.

Les moyennes nécessaires au calcul des variances sont les suivantes :

$$\bar{x}_i = \sum_{jk} (x_{ijk}/I)$$

$$\bar{x}_j = \sum_{ik} (x_{ijk}/J)$$

$$\bar{x}_k = \sum_{ij} (x_{ijk}/K)$$

$$\bar{x}_{ij} = \sum_k (x_{ijk}/IxJ)$$

Toutes les moyennes peuvent être calculées dans le même système de boucles imbriquées.

Calcul des variances.

Le même système de boucles imbriquées permet de calculer toutes les variances. On peut mettre les variances des effets sous la forme suivante :

$$S_A^2 = \sum_{ijl} (a_i^2/da)$$

$$S_V^2 = \sum_{ijl} (v_j^2/dv)$$

$$S_{AV}^2 = \sum_{ijl} ((av)_{ij}^2 / dav)$$

$$S_B^2 = \sum_{ijl} (b_1^2 / dl)$$

$$S_E^2 = \sum_{ijl} (e_{ijl}^2 / de)$$

$$S_G^2 = \sum_{ijl} (g_{ijl}^2 / dg)$$

Voici le listing de calcul des moyennes et des variances dans le cas d'un modèle à 2 critères de classification (3x6) à 4 répétitions.

La généralisation à un modèle à 2 critères de classification (nxn') à n'' répétitions ne comporterait pas plus de lignes.

```

280 REM
290 REM      calcul des moyennes : I=amend., L=bloc, J=variété
300 REM
310 FOR I%=0 TO 2
320 FOR L%=0 TO 3
330 FOR J%=0 TO 5
360 X=IJL(J%+6*I%+18*L%)
370 I(I%)=I(I%)+X/24
380 J(J%)=J(J%)+X/12
400 L(L%)=L(L%)+X/18
410 IJ(I%,J%)=IJ(I%,J%)+X/4
450 S=S+X/72

470 NEXT J%
480 NEXT L%
490 NEXT I%
510 REM
520 REM      calcul des variances et des résidus RES
530 REM
540 FOR I%=0 TO 2
550 FOR L%=0 TO 3
560 FOR J%=0 TO 5
580   SEA=SEA+((I(I%)-S).^2)/2
590   SEB=SEB+((J(J%)-S)^2)/5
610   SED=SED+((L(L%)-S)^2)/3
630   SEAB=SEAB+((IJ(I%,J%)-I(I%)-J(J%)+S)^2)/10
660   SRABC=SRABC+((IJL(J%+6*I%+18*L%)-IJ(I%,J%)-L(L%)+S)^2)/51
665   SG=SG+((IJL(J%+6*I%+18*L%)-S)^2)/71
670   RES(J%+6*I%+18*L%)=IJL(J%+6*I%+18*L%)-IJ(I%,J%)-L(L%)+S
690 NEXT J%
700 NEXT L%
710 NEXT I%

```

6 - COMPARAISON DE MOYENNES PAR LE TEST DE NEWMAN ET KEULS : 2.ACVM-M.BAC.

Le menu proposé est le suivant :

EXPERIMENTATION AC/VHM TAMOA (COMPARAISON DE MOYENNES)

COMPARAISON DE MOYENNES PAR LE TEST DE NEWMAN ET KEULS

Quelle moyenne voulez-vous comparer (premier niveau)?

- dose(1)
- interaction ... MOYENNE BRUTE.(2)
- ...ou INTERACTION (3)
- ligne (4)
- colonne (5)

pour une autre variable (F)
 RETOUR AU MENU..... (M)

En 1984, aucun traitement subsidiaire n'a été apporté, c'est la raison pour laquelle le menu ne propose que la comparaison de moyennes du premier niveau pour ne pas alourdir ce programme qui nécessite toute la mémoire disponible.

Le résultat du test est fourni au choix à l'écran ou sur imprimante. Il donne dans l'ordre le nombre de degrés de liberté, le nombre d'individus qui permet de calculer une moyenne, la variance, la valeur de chacune des moyennes, enfin le tableau à l'entrée dans lequel les moyennes sont rangées par ordre croissant :

- un "+" signifie que les moyennes considérées sont différentes au seuil 5 %
- deux "+" signifie que les moyennes considérées sont différentes au seuil 1 %.

Voici un exemple de sortie imprimante :

```

-----
                COMPARAISON DE MOYENNES PAR LE TEST
                  DE NEWMAN ET KEULS

    date :19/12/84      parametre :PGR      unite : g/plt
DONNEES :

    DDL = 20      NI = 12      SE = 162.806
    4 T = 101.840
    8 T = 117.773
    12 T = 119.819

RESULTATS :

           1      2      3
           4 T    8 T    12 T
1  4 T
2  8 T
3  12 T
  
```

7 - PROGRAMME DE CORRELATION SIMPLE : 2.ACVMH-CS.BAC.

Le calcul des résidus se fait dans le programme 2.CALC-R.BAC et les tests de normalité, de linéarité dans 2.CALC-C.BAC.

a) calcul des résidus : 2 CALC-R.BAC.

Le calcul des résidus se fait suivant la même technique que l'analyse de la variance et aboutit à la création de 6 tableaux virtuels à une dimension U et X (3 résidus).

b) le programme de corrélation : 2.CALC-C.BAC.

Voici un exemple de sortie imprimante

ACVMH/TAMOA-CORRELATIONS SIMPLES

SUR RESIDUS:1ER,2EME,3EME NIVEAU ET RESIDU MOYEN

DM VARIABLE	NORMU	NDRMX	LINUX	LINUXU	RPCORUX	RPCORXU	COR	REGUX	REGXU	COR(>0)
-H 66	:	:	:	:	:	:	:	:	:	:
: R 1	: -	: -	: F=2.89E+00:	F=2.12E-01:	:	:	: -.77E-02:	-.70E-02:	-.85E-02:	3.70E-02
:	:	:	: (0.924)	: (0.188)	:	:	:	:	:	: (0.515)
: R 2	: IMP	: IMP	:	:	:	:	:	:	:	:
: R 3	: -	: -	: F=2.46E-01:	F=3.32E-01:	:	:	: 1.14E-02:	1.41E-02:	9.25E-03:	5.48E-02
:	:	:	: (0.214)	: (0.274)	:	:	:	:	:	: (0.522)
: R 4	: -	: -	: F=1.61E-01:	F=2.70E-01:	:	:	: 6.36E-03:	6.31E-03:	6.41E-03:	4.63E-02
:	:	:	: (0.026)	: (0.074)	:	:	:	:	:	: (0.518)
-H 72	:	:	:	:	:	:	:	:	:	:
: R 1	: -	: -	: F=2.27E+00:	F=1.48E-01:	:	:	: 7.68E-02:	5.53E-02:	1.07E-01:	3.69E-01
:	:	:	: (0.875)	: (0.136)	:	:	:	:	:	: (0.642)
: R 2	: IMP	: IMP	:	:	:	:	:	:	:	:
: R 3	: -	: -	: F=2.26E-01:	F=2.30E-01:	:	:	: 5.62E-03:	5.44E-03:	5.80E-03:	2.69E-02
:	:	:	: (0.198)	: (0.202)	:	:	:	:	:	: (0.511)
: R 4	: -	: -	: F=9.61E-02:	F=1.46E-01:	:	:	: 5.01E-02:	3.94E-02:	6.37E-02:	3.65E-01
:	:	:	: (0.010)	: (0.021)	:	:	:	:	:	: (0.642)

L'organigramme du déroulement des calculs et des tests est le suivant (page suivante)

n couples de valeurs
 observées (x_i, u_i)
 ou
 transgénérées (x'_i, u'_i)
 $n \geq 1$

$n \geq 5$?

$\tau \geq 20$?

X et U normales ?

On applique à chaque variable non-normale une fonction de transgénération et on recommence l'analyse.

X et U normales ?

η^2_{xu} significatif ?

$r_{xu} \neq 0$?

$r_{xu} \neq 0$?

$r_{xu} \neq 0$?

η^2_{ux} significatif ?

η^2_{ux} significatif ?

$U = g(X)$ linéaire ?

$X = f(U)$ linéaire ?

$U = g(X)$ linéaire ?

X et U ne sont pas liés significativement.

X et U sont liés significativement mais non linéairement. Il faut transgénérer X et, ou U et recommencer l'analyse si l'on désire estimer X en fonction de U et U en fonction de X. La transgénération portera :

- plutôt sur U si seule $X = f(U)$ n'est pas linéaire, ou si seul η^2_{xu} est significatif,
- plutôt sur X dans les cas contraires,
- sur X et U si les deux régressions ne sont pas linéaires

X et U sont liés significativement et linéairement :

$$x(u)_i = x + b_{xu} (u_i - \bar{u})$$

$$u(x)_i = u + b_{ux} (x_i - \bar{x})$$

X et U sont liés significativement mais la linéarité, des régressions ne peut être testée.

X et U ne sont pas liés significativement de façon linéaire. Aucune autre étude n'est possible.

X et U sont liés significativement mais la linéarité de leurs régressions et la normalité de leur distributions n'ont pu être testées

BILAN et PERSPECTIVES.

L'acquisition d'un microordinateur aura modifié profondément le traitement des données :

- les analyses statistiques sont beaucoup plus fines grâce aux calculs des probabilités des F, KHI2 etc...

- l'automatisation des calculs (analyse de variance, comparaison des moyennes, corrélation simple) est totale ;

- la technique de gestion des fichiers de données est définitivement au point. Elle évite les longues transcriptions avec tous les risques d'erreurs qu'elles représentent ;

- les éditions sont agréables à lire et directement publiables.

Les gains de temps pour l'ensemble de l'équipe sont considérables, notamment pour les 2 techniciens et la secrétaire.

Cependant le traitement des données peut encore beaucoup évoluer dans le domaine des logiciels et du matériel.

1. LE MATERIEL.

Les capacités du GOUPIL 3 (avec la carte 256K) ont été suffisantes pour le traitement des données en 1984. Les contraintes (mémoire et temps d'exécution) m'ont obligé à rechercher des solutions nouvelles qui se sont souvent révélées plus simples et plus rationnelles en ce qui concerne le déroulement des calculs et l'organisation des fichiers. Cependant dans le cas du carré latin, on a du tout de même limiter la taille du fichier à 100 variables alors que pour les 5 années, il en aurait fallu 500. Plusieurs solutions sont désormais possibles. :

- de nouvelles cartes 256K augmenteraient la capacité "utile" de stockage et aurait l'avantage de faciliter le travail de programmation. C'est la solution la moins coûteuse ;

- une autre solution est un disque dur 5 Mo (ou 10Mo). Sa capacité permet à l'utilisateur de ne plus se préoccuper de la taille de ses fichiers et l'accès aux données est très satisfaisant. Le lecteur 8" ne serait plus alors utilisé que pour la sauvegarde et la copie des fichiers. Les pannes du lecteur 8" seraient moins graves puisque l'essentiel du travail se ferait sur le disque dur. Cette solution plus coûteuse peut se justifier par la fragilité du lecteur 8" : en effet, les pannes sont fréquentes, et un des lecteurs a même été immobilisé 4 mois.

Ces extensions faciliteraient le travail de programmation mais aucune n'est absolument indispensable pour le moment et il est peut être préférable d'attendre un an ou deux avant d'acquérir un disque dur.

2. LOGICIELS.

Il n'existe pas encore de logiciel sur le marché vraiment adapté à des analyses statistiques aussi fines des expérimentations agronomiques. Chaque expérimentation est un cas particulier ne serait-ce qu'en raison du tirage aléatoire de la position des parcelles. Enfin, il est difficile de paramétrer certaines phases des calculs et surtout l'impression si on veut des sorties imprimantes agréables à lire et directement publiables. Mais les travaux que j'ai réalisés depuis un an ont contribué peu à peu à constituer une bibliothèque de programmes relatifs à des modèles fréquemment rencontrés. Ainsi en quelques heures, il est maintenant possible en modifiant le modèle le plus "proche" d'écrire l'ensemble des programmes pour un nouveau modèle : gestion des fichiers, analyse de variance, comparaison de moyenne et corrélation simple.

Cependant on peut encore améliorer le paramétrage pour les programmes de corrélation simple et de comparaison de moyennes. En ce qui concerne l'analyse de variance, on peut envisager 2 bibliothèques de programme : l'une pour les calculs, l'autre pour les éditions.

ex : on peut généraliser facilement les calculs d'une expérience factorielle à 2 critères de classification ($n \times n'$) avec n'' répétitions, mais le programme d'édition sera fonction de n , n' , n'' .

3. LOGICIELS D'ACCOMPAGNEMENT.

- AUTOGRAPHE judicieusement utilisé avec le SBASIC rend la programmation en BASIC très rapide. Cependant ce logiciel pourrait être mieux exploité : la rédaction des protocoles expérimentaux et un grand nombre de paragraphes des rapports pourrait être réalisé sur le GOUPIL.

- DYNAPLOT est un logiciel graphique. Il pourrait permettre en disposant d'une table traçante de présenter les résultats des expérimentations sous une forme plus simple (graphe, camembert,...)

- Un logiciel d'Analyse Factorielle des Correspondances pourrait compléter utilement l'étude statistique actuelle. Cette méthode pourrait permettre en fin d'expérimentation de mettre en évidence des relations linéaires ou non linéaires.