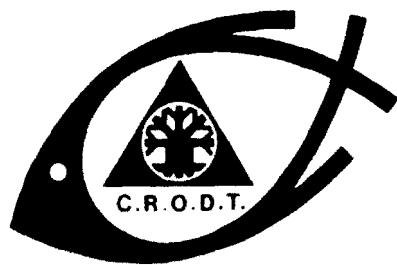


00000027

SEMINAIRE DE FORMATION A L'UTILISATION  
DU PROGICIEL "GENSTAT" DAKAR-CRODT  
28 JANVIER - 6 FÉVRIER 1985

L. LALOË



ARCHIVE

CENTRE DE RECHERCHES OCÉANOGRAPHIQUES DE DAKAR • TIAROYE

N° 139

\* INSTITUT SÉNÉGALAIS DE RECHERCHES AGRICOLES \*

MARS 1985

SÉMINAIRE DE FORMATION A L'UTILISATION  
DU PROGICIEL "GENSTAT" DAKAR-CRODT  
28 JANVIER - 6 FÉVRIER 1985

*par*

Franc i s LALOË \*

I N T R O D U C T I O N

Le "Bureau calcul" de l'ISRA (Institut Sénégalais de Recherches Agricoles), situé au CRODT (Centre de Recherches Océanographiques de Dakar-Thiaroye) est équipé d'un ordinateur IBM 4331. Il s'agit d'un ordinateur disposant d'un espace disque, d'une mémoire centrale et d'un processeur largement suffisants pour le stockage et le traitement des données collectées par l'ensemble des équipes de recherches de l'ISRA.

Pour satisfaire les besoins en calcul, les principaux langages sont présents FORTAN, COBOL... La grande quantité de calculs "statistiques" justifie et impose la disposition de programmes ou langages statistiques. Nous avons installé le progiciel GENSTAT (General Statistical program) écrit par les chercheurs de Rothamsted Agricultural Station en Angleterre. Il est actuellement utilisé pour un grand nombre de traitements.

Un des grands problèmes rencontrés, pour une utilisation optimale d'outils de ce type, est celui de la formation de "programmeurs". Une telle formation était l'objectif du séminaire organisé du 28 janvier au 6 février 1985. Les divers exposés et programmes réalisés pour ce séminaire sont décrits dans ce document. Le séminaire a été suivi par 15 chercheurs de l'ISRA.

L'appellation de "programmeur" utilisée plus haut est en fait fort mal choisie. En effet, la difficulté d'utilisation d'un langage statistique réside moins dans l'informatique ou la programmation, que dans l'acquisition de connaissances en statistiques. Pour illustrer ce fait, il suffit de signaler que les langages statistiques modernes (GENSTAT, SAS, BMDP, SPSS...) sont utilisés par certaines universités et écoles pour enseigner les statistiques à des étudiants "non informaticiens". Pour cette raison, il était demandé aux participants à ce séminaire de disposer de connaissance de base en statistique, et le lecteur observera également que les exposés sont généralement organisés autour d'un thème "statistique", le programme écrit en GENSTAT étant une illustration.

Le présent document ne doit pas être considéré comme un manuel d'utilisation. De tels manuels existent déjà. Il constitue un exemple de ce qui peut être fait dans un cours de formation d'une durée limitée (7 jours de cours) et donne une idée d'un nombre important d'applications possibles à partir d'exemples pour la plupart issus de travaux effectués au CRODT. Ces travaux ont

---

\* Biostatisticien de l'ORSTOM en service au CRODT (ISRA)

pu être réalisés avec un temps de programmation très réduit.

Il faut insister sur le fait que deux domaines des statistiques ont été **couverts** lors du séminaire : Modèles linéaires et analyses multivariées. Il est très fréquent que seul un de ces domaines soit couvert au cours d'un **seminaire**. C'est dommage car les deux approches se complètent et sont fondamentales. Il faut pouvoir les aborder toutes deux. Lors d'un séminaire un peu plus long, une partie supplémentaire pourrait être incorporée concernant les études de séries chronologiques. Les modèles **ARIMA** peuvent être traités d'une manière totalement satisfaisante avec GENSTAT, et dans la plus récente version déjà disponible dans la plupart des sites, de nouvelles directives permettent les calculs de transformées de Fourier. Une part plus importante pourrait également être accordée aux modèles non linéaires juste évoqués ici.

Dans la première partie de ce document le programme général du séminaire est exposé. La deuxième partie est consacrée à la description des exposés et programmes. L'annexe 1 contient les listings des programmes. Dans l'annexe 2 on trouvera quelques renseignements utiles; documentation, adresse des distributeurs.

## 1. P R O G R A M M E G E N E R A L

Le séminaire était organisé en trois parties :

1) Une initiation à l'utilisation de l'ordinateur. Cette partie étant liée à l'équipement local n'est pas décrite dans ce document (1 journée).

2) Formation à l'utilisation de GENSTAT. Elle a été faite par des exposés illustrés par des programmes GENSTAT (6 journées).

Les trois premiers exposés sont consacrés à des thèmes "informatiques" : lectures, écritures, déclarations de structures, calculs élémentaires (moyennes, **variances**, covariances, calcul matriciel, regroupement en catégories, fonctions mathématiques, boucles, sorties graphiques).

Les 7 exposés suivants ont été faits sur des thèmes "statistiques" :

- Tabulation (exemple de traitement d'un recensement).
- Modèle linéaire (3 exposés) :
  - Présentation générale (analyse par régression et analyse de **variance** d'un même jeu de données)
  - Analyses de **variance**
  - Regressions.
- Analyses multivariées (3 exposés) :
  - Analyses en composantes principales et analyse **discriminante**
  - Analyse en coordonnées principales - Etude de matrice de proximité.
  - Analyse de correspondances.

Un exposé supplémentaire a été consacré à un traitement intégrant la plupart des méthodes décrites ci-dessus (plus un ajustement d'un modèle non linéaire).

3) Discussion avec les participants à partir de leurs propres études, Plusieurs traitements ont été effectués au cours de ces "Travaux pratiques" :

- Analyses de **variances**
- Analyses en **composantes** principales
- Regressions multiples (step-wise, analyses combinant variables qualitatives et quantitatives)
- Tabulation
- Analyse de correspondances multiples

- Etude de séries chronologiques.  
Certains de ces traitements ont fait l'objet d'exposés supplémentaires.

## 2. DESCRIPTION DES EXPOSES ET PROGRAMMES

### 2.1. EXPOSES-PROGRAMME N° 1

Tous les programmes doivent commencer par la directive 'REFE' et finir par les directives 'CLOSE' et 'STOP'.

Ce programme comporte les déclarations de plusieurs types de structures :

- Les variables introduites par la directive 'VARIABLE' ou 'VARI' (seules les quatre premières lettres sont prises en compte) :

'VARI' LONGUEUR, POIDS \$ 10 (ligne 3 du programme) indique que les structures

LONGUEUR et POIDS contiennent chacune 10 valeurs,  $l_i, p_i, i = 1 \dots 10$  (analogue au fortran DIMENSION X (10) ).

- Les facteurs, ou variables qualitatives : 'FACT' SEXE \$ 2,10 (ligne 7) indique que la structure SEXE est une variable qualitative à 2 niveaux et 10 valeurs.

- Les scalaires

'SCAL' LM, PM, LV, LE, PE (ligne 23) indique que ces variables ne contiennent qu'une valeur.

D'autres structures existent, nous en verrons dans les autres programmes.

Les calculs s'effectuent aussi par directives, par exemple :

'CALC' PM, LP = MEAN (POIDS, LONGUEUR) (ligne 26) permet le calcul des moyennes des longueurs et des poids rangées dans LM et PM. Sur le listing n° 1, figurent d'autres exemples de calculs par la directive 'CALC'.

Les lectures s'effectuent par la directive READ :

'READ/PRIN = DEM' SEXE, POIDS, LONGUEUR (ligne 9) provoque la lecture en parallèle des trois structures. Dans une lecture en parallèle toutes les structures sont de même longueur (ici 10). Les premières valeurs de chaque structure sont lues, puis les secondes etc... A la fin du jeu de données l'instruction 'EOD' doit être rencontrée et il est vérifié que le nombre de valeurs lues est bien égal au nombre de structure multipliée par la longueur commune (ici  $3 \times 10 = 30$ ).

Dans la directive READ écrite ci-dessus, le signe/ suivi de PRIN = DEM signifie qu'on demande une option. Le D de DEM provoque l'impression des données (elles figurent sur le listing) le E signifie que les éventuelles erreurs de lectures seront signalées et le M provoque l'impression des moyennes, minimum, maximum, nombre de valeurs, nombre de données manquantes pour chaque variable. Si on ne fait pas appel à l'option PRIN, l'option par défaut est PRIN = EM. Si on ne veut pas d'impression on écrit :

'READ/PRIN = Z'. Beaucoup d'autres options sont disponibles pour 'READ'.

Enfin on notera qu'aucun format de lecture n'est donné. Ici toutes les données étaient séparées par un ou plusieurs blancs, considérés par la machine comme des séparateurs, aucun format n'est donc nécessaire. On peut également lire avec des formats lorsque c'est nécessaire.

Les impressions peuvent se faire par la directive PRINT. Quelques exemples figurent dans le listing. Nous verrons plus loin que certaines directives de calcul provoquent elles mêmes des impressions de résultats.

La directive 'RUN' provoque l'exécution des directives précédentes. Si dans ces directives figure un ordre 'READ' les données correspondantes doivent suivre l'instruction 'RUN' (sauf si ces données sont lues sur une autre unité). Plusieurs directives 'RUN' peuvent figurer dans un programme.

## 2.2 EXPOSE ET PROGRAMME N° 2

La première partie du programme concerne le calcul matriciel.

Les matrices sont déclarées par 'MATR' M64 § 6,4 (ligne 5) (déclaration de la matrice M64 à 6 lignes et 4 colonnes). Les calculs se font par la directive 'CALC' :

```
'CALC' M44 = PDT (M46 ; M64) (produit matriciel, ligne 22)
'CALC' M44 = INV (M44) (inversion, ligne 26).
```

Les fonctions de calcul de déterminants et traces sont indiquées. De nombreuses autres fonctions existent, en particulier la diagonalisation. La deuxième partie concerne le calcul des matrices de covariances et certains calculs sur les variables (sommées, différenciations) .

Enfin un exemple de boucle est montré ainsi que l'utilisation de la directive 'EQUA' pour transférer les valeurs de structures à structures, et de la fonction ELEM pour accéder à certains éléments d'une structure.

## 2.3 EXPOSE ET PROGRAMMES N° 3

La directive 'UNIT' § 15 (ligne 3) indique que toutes les structures par défaut ont une longueur 15. Les lectures se font en série (option FORM = S)

Diverses manières de regrouper des valeurs en facteur sont décrites. 'GROUP' QU (1, 2), SEXT = QUANT (X (1...3) ; 4, 4, 6) signifie que les valeurs de X (1), X (2), X (3) sont regroupées respectivement en quartiles, quartiles et sextiles dans les facteurs QU (1), QU (2) et SEXT. Le facteur SEXT a donc 6 niveaux. On regroupe ensuite dans le premier niveau du facteur GR 1 les niveaux 1, 2 et 3 de SEXT, dans le deuxième niveau de GR 1 les niveaux 4 et 6 de SEXT et le niveau 5 de SEXT dans le 3ème niveau de GR 1.

Le facteur GR 1 à 4 niveaux est ensuite calculé par la fonction LIMITS :

```
GR2 = 1      si      x (1) a 1
GR2 = 2      si      1 < x (1) ≤ 4
GR2 = 3      si      4 < X (1) ≤ 7
GR2 = 4      si      7 < X (1)
```

On peut noter que la faute consistant à déclarer 3 niveaux pour GR2 (ligne 9) a été corrigée. Ensuite on montre un exemple de calculs avec conditions logiques (ligne 21).

La fin du programme est consacrée à des sorties graphiques.

Au cours de l'exposé une description complète de la directive 'GRAPH' a été faite avec l'ensemble des options possibles. En même temps a été exposée la syntaxe générale des directives. Les possibilités de sorties sur table traçante, figurant dans la nouvelle version, non encore installée à Dakar, ont été rapidement évoquées.

## 2.4. EXPOSE ET PROGRAMME SUR LA TABULATION (TRAITEMENT D'UN RECENSEMENT)

Un jeu de données tirées d'un recensement du parc piroguier sénégalais est traité. On dispose pour 30 pirogues d'un certain nombre de renseignements : lieu d'enquête, lieu d'origine, activité, puissance du moteur, type de pêche et prise moyenne par sortie.

Le programme crée de nouvelles catégories : région d'enquête ou d'origine, type d'engin utilisé, migrants ou sédentaires. Diverses tables à 2 ou 3 entrées sont calculées et imprimées.

Les différents niveaux des facteurs sont indicés par des noms, déclarés par des structures 'NAME'.

## 2.5. EXPOSE ET PROGRAMME SUR LE MODELE LINEAIRE

Dans un premier temps on fabrique des données à l'aide de nombres générés par la fonction RANDU (valeurs suivant une loi uniforme sur (0, 1) 15 valeurs  $X_k$  et 15 valeurs  $E_k$  sont ainsi créées suivant des lois normales  $N(20, 10)$  et  $N(0, 5)$ . 15 valeurs  $Y_k$  sont ensuite calculées comme suit :

$$Y_k = 3 \cdot X_k + E_k \text{ si } k = 1, 2, 3, 4 \text{ ou } 5$$

$$Y_k = 10 + 3 \cdot W_k + E_k \text{ si } k = 6, 7, 8, 9 \text{ ou } 10$$

$$Y_k = 25 + 3 \cdot X_k + E_k \text{ si } k = 11, 12, 13, 14, \text{ ou } 15$$

On peut ainsi noter ces valeurs :

$$Y_{ij} = a_i + \beta X_{ij} + E_{ij} \quad \begin{array}{l} i = 1 \dots 3. \quad j = 1 \dots 5 \\ \beta = 3, \quad a_1 = 0, a_2 = 10, a_3 = 25 \end{array}$$

Il s'agit d'un modèle linéaire dont les paramètres sont  $a_1$ ,  $a_2$ ,  $a_3$ , et  $\beta$

$Y_{11}$	1	0	0	$X_{11}$	$E_{11}$
$Y_{12}$	10	0	0	$X_{12}$	$E_{12}$
$Y_{13}$	10	0	0	$X_{13}$	$E_{13}$
$Y_{14}$	1	0	0	$X_{14}$	$E_{14}$
$Y_{15}$	10	0	0	$X_{15}$	$E_{15}$
$Y_{21}$	0	1	0	$X_{21}$	$E_{21}$
$Y_{22}$	0	1	0	$X_{22}$	$E_{22}$
$Y_{23}$	0	1	0	$X_{23}$	$E_{23}$
$Y_{24}$	3	1	0	$X_{24}$	$E_{24}$
$Y_{25}$	0	1	0	$X_{25}$	$E_{25}$
$Y_{31}$	0	0	1	$X_{31}$	$E_{31}$
$Y_{32}$	0	0	1	$X_{32}$	$E_{32}$
$Y_{33}$	0	0	1	$X_{33}$	$E_{33}$
$Y_{34}$	0	0	1	$X_{34}$	$E_{34}$
$Y_{35}$	0	0	1	$X_{35}$	$E_{35}$

qui peut être écrit :

$$Y = X \cdot \theta + E$$

où  $Y$  et  $E$  sont des matrices (15,1)

où  $X$  est une matrice (15,4)

et où  $\theta$  est une matrice (4,1)

Supposer tous les  $a_i$  égaux revient à faire une régression  $Y = a + BX$  :

Si on désire séparer les  $a_i$  on doit ajouter le facteur  $F$  à 3 niveaux dans les "variables explicatives" cela revient à ajouter des variables explicatives valant 0 ou 1 pour ce faire on peut utiliser la directive 'REGRESSION' : 'REGR/DVSET = F'  $Y = X \cdot F$

'Y'Y

'FIT'X : F : F + X ; RES = R : F + X.F

'REGR' est suivi de la liste des structures (variables ou facteurs) qui pourront être utilisées en variables dépendantes (variables) ou indépendantes (variables et/ou facteurs).

'Y' est suivi de la variable dépendante

'FIT' est suivi du modèle à ajuster :

'FIT' X provoque l'ajustement :  $Y_{ij} = a + \beta X_{ij} + e_{ij}$

'FIT' F " " :  $Y_{ij} = a_i + e_{ij}$

'FIT' X + F " " :  $Y_{ij} = a_i + \beta X_{ij} + e_{ij}$

'FIT' F + X.F " " :  $Y_{ij} = a_i + \beta_i X_{ij} + e_{ij}$

(Pour indiquer que les pentes diffèrent pour chaque niveau de  $F$  on introduit  $X.F$ , l'interaction entre  $X$  et  $F$ . L'option DVSET = F de 'REGR' est nécessaire pour permettre l'établissement des contraintes d'identification lorsqu'une interaction (ici  $X.F$ ) est utilisée en l'absence d'un effet principal (ici  $X$  est absent dans  $F + X.F$ ).

Le symbole "?" signifie la répétition de la directive précédente :  
 'FIT'X : F équivaut à 'FIT'X 'FIT'F.  
 En écrivant : 'FIT' X + F ; RES = R, on range les résidus de l'analyse dans R. RES est un mot clé. Il est possible de conserver ainsi toutes les structures utiles pour des travaux ultérieurs.  
 Tous les ajustements ont été faits par les moindres carrés. On peut également par un jeu d'options demander un ajustement par maximum de vraisemblance **lorsque** les résidus suivent certaines lois non normales.

Dans le listing figurent les différents estimateurs et tableaux d'analyses de **variance**.

on peut aussi aborder le problème par analyse de **variance** :

Si on veut introduire la "**covariable**" X on a  $Y_{ij} = a_i + \beta X_{ij} + E'_{ij}$

on écrit alors 'TREAT'F 'COVA'X 'ANOVA'Y ; RES = R1

Dans les deux cas on a conservé les résidus (R et R1). On observe que les sommes de carrés résiduels sont identiques, ainsi que les différents estimateurs, de même tous les résidus R et R1 sont les mêmes.

## 2.6. EXPOSE ET PROGRAMME SUR LA REGRESSION

Il s'agit d'une étude sur la taille moyenne de crevettes pêchées en Casamance (données de L. LE RESTE).

Pour chaque opération de pêche on dispose de la taille moyenne des crevettes des mesures de courant en **surface**, en pleine eau et au fond, ainsi que de la salinité (avec quelques données manquantes).

Plusieurs modèles sont essayés. Les méthodes permettant de faire des tests sur les coefficients à partir des tableaux d'analyse de **variance** sont exposées. La signification des valeurs figurant dans les colonnes "T" est expliquée.

Le modèle final est une régression multiple :  $Y_i = a + bC_i + cS_i + dS_i^2 + e_i$   
 où  $C_i$  est le courant de surface,  $S_i$  la salinité est  $S_i^2$  la salinité au carré.  
 La répartition des résidus est commentée. Dans la dernière partie du programme, 3 tableaux sont édités : effectifs des opérations, moyennes ajustées et **observées** des tailles par niveau de salinité et de courant.

## 2.7. EXPOSE ET PROGRAMME SUR L'ANALYSE DE VARIANCE

Le jeu de données étudiées est **tiré** d'une étude sur les erreurs commises lors des estimations à vue du poids de tas de poissons par divers enquêteurs. 10 enquêteurs estiment le poids de 60 tas. On dispose donc de 600 estimations. La variable étudiée est :

$$Y_{ijk} = \frac{\hat{P}_{ijk} - P_{ijk}}{P_{ijk}} \text{ où } \hat{P}_{ijk} \text{ est}$$

l'estimation faite par l'enquêteur i du même tas de catégorie j. (3 catégories sont définies, la première constituée par les 20 tas les plus petits, la seconde par les 20 tas de poids intermédiaires et la troisième par les 20 plus gros).  $P_{ijk}$  est le vrai poids du tas,

2 modèles sont étudiés :

1)  $Y_{ijk} = u + e_i + c_i + ec_{ij} + ijk$

défini par 'TREA' ENQT \* CAT

2)  $Y_{ijk} = u + e_i + c_j + ct_{jk} + ec_{ij} + ijk$

(on ajoute un effet tas intra catégorie)

défini par 'TREAT' ENQT + CAT/TAS + ENQT.CAT

## 2.8. EXPOSE ET PROGRAMME SUR L'ANALYSE EN COMPOSANTES PRINCIPALES ET DISCRIMINANTE.

Le principe général des analyses factorielles est exposé. L'exemple traité est tiré du manuel GENSTAT édité par l'INRA, 23 vaches (Zébu et charolaises) avec chacune six observations (poids vif, carcasse, viande de première qualité, viande totale, gras, os) constituent un nuage de 23 points dans un espace à 6 dimensions. L'analyse se fait par la directive 'PCP' 'PCP/PRIN = L, CORR = Y, NLR = 4' X (1...6) ; SCORES = COMP (ligne IX) La lettre L de l'ophon PRIN signifie qu'on ne demande que les valeurs propres, % de variation et vecteurs propres. L'option CORR = Y signifie qu'on fait une A.C.P normée, sur les corrélations et non les covariances. NLR = 4 signifie qu'on ne demande que 4 axes (par défaut on aurait l'analyse sur les covariances et tous les axes). Les valeurs des points (vache) sur les 2 premiers axes sont rangées dans la matrice COMP pour faire un graphique. L'utilisation de la directive 'EQUA' avec un format est expliquée. Les points sur le graphe sont identifiés par le nom des vaches.

Pour effectuer l'analyse discriminante, on calcule la matrice de covariance "intra-race".

Elle est déclarée par :

```
'DSSP' INTRA $ X (1...6); RACE; M (1...6), N
```

et calculéepar :

```
'SSP/PRIN = SC' INTRA
```

La matrice INTRA est calculée en fonction du facteur RACE, les moyennes de chaque race son% dans les variables M (1 . ...6) et les effectifs dans N.

L'option PRIN = SC dans 'SSP' provoque les impressions des matrices de sommes de carrés (lettre S) et des corrélations (lettre C).

L'analyse est faite par la directive 'CVA' avec une partie des sorties possibles demandées par l'option PRIN.

## 2.9. EXPOSE ET PROGRAMME SUR L'ANALYSE EN COORDONNEES PRINCIPALES

Le problème est analogue à celui d'une analyse en composantes principales. On dispose d'une matrice de distances entre villes sénégalaises (déclarée par la directive 'SYMMAT', matrice symétrique). On cherche le plan qui rend le mieux compte de ces distances (voir sortie graphique). Des erreurs sur les mesures des distances faites sur une carte à l'aide d'un double décimètre rendent sans doute compte de l'ajustement non parfait sur un plan.

L'analyse est faite par la directive 'PCO'. Une étude est ensuite effectuée pour rechercher les 3 villes les plus proches de chaque ville ;

'NEIGHBOUR/3'DISTANCE, ou DISTANCE est une matrice de proximités déduite de la matrice des distances). Ensuite le "minimum spanning tree" est imprimé ('MST' DISTANCE ).

## 2.10. EXPOSE ET PROGRAMME SUR L'ANALYSE DES CORRESPONDANCES

Quelques particularités de l'analyse des correspondances sont décrites (distance du chi 2...).

Le programme a été fait à partir d'un tableau donnant les effectifs de certaines plaintes par classe d'âge de plus de 6000 consultants dans des dispensaires de Pikine (données ORSTOM, équipe urbanisation santé Pikine).

Il n'y a pas dans GENSTAT de directive provoquant directement une analyse



de. correspondances. Un programme a été écrit à cet effet par Mr ASTIER de l'Université Paris Sud. Il a été complété par un sous programme qui permet la suppression des lignes et colonnes dont la somme est inférieure à un certain seuil (25 dans notre programme). En GENSTAT un sous programme est une structure déclarée par la directive 'MACRO'. Ici le programme lit ces macros dans un fichier binaire (unité 11 dans notre exemple).

Le sous programme est appelé par la directive 'USE' (ligne 40).

La suite de l'exposé est consacrée à la description des différentes **sorties** (% de variance, coordonnées, contributions relatives et absolues, graphes),

Les **sous-programmes** CORRESP2 et PREPCORR ont été lus sur un fichier binaire permanent. La gestion de ces fichiers a été décrite.

#### 2.11. EXPOSE SUR UN PROGRAMME GENERAL FAISANT INTERVENIR PLUSIEURS TRAITEMENTS DIFFERENTS.

Cet exposé peut être fait en introduction. Il effectue rapidement la plupart des **traitements** exposés dans les autres programmes (avec en plus un ajustement non linéaire ; directives 'MODEL' et 'OPTIMISER'). Il permet de montrer, sur un exemple "sur mesure" un éventail de traitements disponibles sans effort de programmation.

#### 2.12. CONCLUSION

Toutes les possibilités de GENSTAT n'ont évidemment pas été présentées dans ces programmes. Des chapitres entiers ont même été ignorés ou juste abordés (ajustements non linéaires, séries chronologiques, analyses **multi-dimensionnelles**...). Les participants ayant suivi les "travaux pratiques" ont été néanmoins familiarisés avec la syntaxe générale du langage et sont maintenant susceptibles de rechercher eux-mêmes dans la documentation les méthodes non décrites dans le cours.

## A N N E X E 1

Listings des programmes exposés

Les pages sont numérotées par deux nombres :

$n - k$  signifie kième page du nième exposé.

GENSTAT V MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (RUTHAMSTED EXPERIMENTAL STATION)  
 -- INSTALLE AU C.R.O.D.T.

```

1  *REFERENCE*EXER_1
2  **DECLARATION DE DEUX VARIABLES DE LONGUEUR 10**
3  *VARIABLE* POIDS, LONGUEUR $10
-4  **DECLARATION D'UN FACTEUR "SEXE" A DEUX NIVEAUX
-5  ET DE LONGUEUR 10 ET D'UN FACTEUR "ESPECE" A CINQ
  b NIVEAUX ET DE LONGUEUR 10**
7  *FACTEUR* SEXE $2,10
8  **LECTURE EN PARALLELE DES TROIS STRUCTURES**
9  *READ/PRIN=DEM* SEXE, POIDS, LONGUEUR
10 *RUN*
```

```

11 2 10 20
12 1 11 22
13 1 12 21
14 1 13 24
15 1 15 25
16 2 10 20
17 1 13 19
18 2 917
19 2 12 20
20 2 14 21
21 *EOD*
```

POIDS	MNMINMAX	11.6000	9.0000	15.0000	10	VALUES	0	MISSING
LONGUEUR	MNMINHAX	20.9000	17.0000	25.0000	10	VALUE 5	0	MISSING

```

22  **DECLARATION D E SCALAIRES**
23 *SCALAIRES* LM, PM, LV, PV, LE, PE
24 **CALCUL DE PARAMETRES DES DISTRIBUTIONS DES VARIABLES**
25 **VARIABLES -----> SCALAIRES**
26 *CALCULER* PM, LM =MEAN(POIDS, LONGUEUR)
27 *CALCULER* PV, LV =VAR (POIDS, LONGUEUR)
28 **SCALAIRES -----> SCALAIRES**
29 *CALCULER* LE, PE =SQRT(LV, PV)
30 **VARIABLES -----> VARIABLES**
31 *VARI* COND $10
32 *CALC* COND=2*POIDS/LONGUEUR
33 **IMPRESSION DE RESULTATS**
34 *CAPT* ** TITRE (RESULTAT DE LA DIRECTIVE (CAPTION))**
35 "IMPRESSIONS"
36 *PRINT/P* SEXE, POIDS, LONGUEUR, COND $10.3 ** EN PARALLELE AVEC FORMAT**
37 *PRINT/S* SEXE, LONGUEUR . * EN SERIE ET SANS FORMAT"
38 *PRINT* PM, LM $5.1
39 *PRINT* PV, LV, PE, LE
40 *RUN*
```

TITRE (RESULTAT DE LA DIRECTIVE CAPTION)

SEXE	POIDS	LONGUEUR	COND
2	10.000	20.000	1.000
1	11.000	22.000	1.000
1	12.000	21.000	1.143
1	13.000	24.000	1.083
1	15.000	25.000	1.200
2	10.000	20.000	1.000
1	10.000	19.000	1.053
2	9.000	17.000	1.059
2	12.000	20.000	1.200
2	14.000	21.000	1.333

SEXE  
2  
1  
1  
1  
1  
1  
2  
1  
2  
2  
2

LONGUEUR  
2.0000E 1  
2.2000E 1  
2.1000E 1  
2.4000E 1  
2.5000E 1  
2.0000E 1  
1.9000E 1  
1.7000E 1  
2.0000E 1  
2.1000E 1

PM 11.6  
LM 20.9  
PV 3.8222E 0  
LV 5.4333E 0  
PL 1.9550E 0  
LF 2.3309E 0

41 \*CLOSE\*  
\*\*\*\*\* END OF EXERC\_1 AT LINE 36 USED 465 LEFT 49535

1 2

GENSTAT v MARK 4.33  
 (C)1980 LAWES AGRICULTURAL TRUST (RUTHAMSTEAD EXPERIMENTAL STATION)  
 -- INSTALLE AU C.R.O.D.T. --

```

1 *REFERENCE*EXER_2
2 ** CALCUL: DIVERS ET VARIES AVEC MATRICES, SCALAIRES ET VARIABLES**
3 **DECLARATIONS**
4 *MATRICES*M46$4,6
5 *MATRICES*M64$6,4
6 *MATRICES*M44$4,4
7 *MATRICES*M66$6,6
8 *SCALAIRES*DETER,TR *VARI* V(1),V(2),V(3),V(4)$6
9 *READ/S,PRIN=D*M46,M64
10 *RUN*
```

```

11 1 2 3 4 5 6
12 2 3 4 5 7 9
13 0 1 0 0 0 0
14 1 0 0 0 0 0 *EOD*
15 1 2 3 5
16 1 3 5 7
17 2 4 1 3
18 1 0 0 0
19 1 2 0 0
20 1 2 1 0 *EOD*
```

```

21 ** PRODUIT MATRICIEL **
22 *CALC*M44=PDT(M46;M64)
23 ** IMPRESSIONS **
24 *PRIN*M46$4,0 *PRIN*M44$7,0
25 ** INVERSION **
26 *CALC*M44=INV(M44) *PRIN*M44$9,2
27 ** CALCULS DES DETERMINANTS ET TRACE **
28 *CALC*DETER=DET(M44)
29 *CALC*TR =TRACE(M44)
30 *PRIN*DETER,TR$10,2
31 *RUN*
```

M46

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	3	4	5	7	9
3	0	1	0	0	0	0
4	1	0	0	0	0	0

M44

	1	2	3	4
1	24	42	22	28
2	34	61	34	43
3	1	3	5	7
4	1	2	3	5

M44

	1	2	3	4
1	h.42	-4.54	4.71	-3.62
2	-4.54	3.23	-3.62	2.69
3	4.35	-3.08	4.54	-4.23
4	-2.08	1.46	-2.23	2.38

DETER 0.04

TR 16.58

```

32 *READ/PRIN=DEM*V(1...4)
33 *RUN*

34 1 2 3 4
35 0 1 0 2
36 1 3 2 4
37 1 3 5 7
38 0 0 # 0
39 2 3 1 4
40 *EOD*
    V(1) MNMINMAX      0.8333      0.0000      2.0000      6 VALUES      0 MISSING
    V(2) MNMINMAX      2.0000      0.0000      3.0000      6 VALUES      0 MISSING
    V(3) MNMINMAX      2.2000      0.0000      5.0000      6 VALUES      1 MISSING
    V(4) MNMINMAX      3.5000      0.0000      7.0000      6 VALUES      0 MISSING

41 **CALCUL DE MATRICES DE COVARIANCES**
42 **DECLARATION **
43
44 **CALCUL AVELOPTIONO'IMPRESSION**
45
46 **EXEMPLE DE BOUCLE, QUATRE PASSAGES SONT EFFECTUES LES DEUX LISTES
47 V EI S SONT BALAYEES EN PARALLELE **
48 *SCAL*          S(1...4)          * VARI*SV,OSV4
49 *FOR*          V=V(1...4);S=S(1...4)
50 *CALC*         S=SUM(V)
51 *REPETE*
52 * UTILISATION DE LA DIRECTIVE EQUATE POUR METTRE DANS LA VARIABLE
53 SV DE LONGUEUR 4 LES 4 SCALAIRES S(1...4)**
54 *EQUA*         SV=S(1...4)
55 **LA FONCTION ELEH (SV;N) PERMET DE TROUVER LE NIEME ELEMENT DE LA
56 VARIABLE SV **
57 *CAL*         S(2)=ELEM(SV;2)      *PRIN*S(2)
58 **CALCUL DE DIFFERENCES ENTRE OBSERVATIONS (POUR SERIES
59 CHRONOLOGIQUES **
60 *CALC*         DSV=DIFF(SV;2)      *PRIN/P*SV,DSV%6.1
61 *RUN*

```

45.....

\*\*\*\*SUMS OF SQUARES AND PRODUCTS MATRIX\*\*\*\*

DF = 4

		SM				
V(1)	1	2.0000E 0	0			
V(2)	2	2.0000E 0	3.2000E 0			
V(3)	3	1.0000E 0	3.6000E 0	1.4800E 1		
V(4)	4	2.0000E 0	4.6000E 0	1.2800E 1	1.2800E 1	
MEAN	5	10.0000E -1	2.4000E 0	2.2000E 0	4.2000E 0	5.0000E 0
			1	2	3	4
						5

\*\*\*\*CORRELATION MATRIX\*\*\*\*

DF = 3

		1	2	3	4
V(1)	1	1.0000			
V(2)	2	0.7906	1.0000		
V(3)	3	0.1838	0.5231	1.0000	
V(4)	4	1.3953	0.7187	0.9300	1.0000
S(2)	1	1.2000E 1			

SV	DSV
5.0	*
12.0	*
11.0	6.0
21.0	9.0

1 4

GENSTAT v MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 - - - - - INSTALLE AU C.R.O.O.T. - - - - -

```
1 'REFE' EXER_3
2 OPERATIONS ENTKE VARIABLES ET FACTEURS
3 'UNITS' $15
4 'READ/PRIN=DEM,FORM=S' X(1...3)
5 'RUN'
```

6	1	2	3	4	1	2	3	4	5	6	1	4	1	7	8	'EOD'					
	X(1)				MNMINMAX											1.0000	8.0000	15	VALUES	0	MISSING
7	1	3	4	5	14	11	14	13	15	12	11	12	13	13	14	1.0000	15.0000	15	VALUES	0	MISSING
	X(2)				MNMI																
8	1	2	5	3	8	9	0	0	0	1	1	2	3	1	2						
	X(3)				HMINYAX											0.0000	9.0000	15	VALUES	0	MISSING

```
9 'FACT' QU(1,2)$4 : SEXT$6 : GR1,GR2$3
10 PUANTILES
11 'GROUPE' QU(1,2),SEXT=QUANT(X(1...3));4,4,6)
12 KE GROUPEMENTS DE FACTEURS
13 'INTE' IN=1,2,-3,4,-6,-5
14 'GROUPE' GR1=GROUP(SEXT;IN)
15 REGROUPEMENTS PAR INTERVALLES
16 'VARI' LIMITE=1,4,7
17 'GROU' GR2=LIMITS(X(1);LIMITE)
18 'PRIN/P' QU(1,2),SEXT,GR1,GR2
19 'RUN'
```

QU(1)	QU(2)	SEXT	GR1	GR2
1	1	3	1	1
2	1	4	2	2
2	1	5	3	2
3	1	5	3	2
1	4	6	2	1
2	2	6	2	2
2	3	1	1	2
3	3	1	1	2
3	4	1	1	3
4	2	3	1	3
1	2	2	1	1
3	2	4	2	2
1	3	5	3	1
4	3	2	1	3
4	4	3	1	4

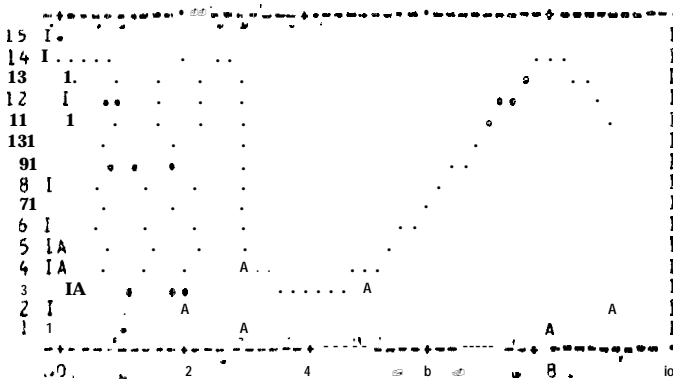
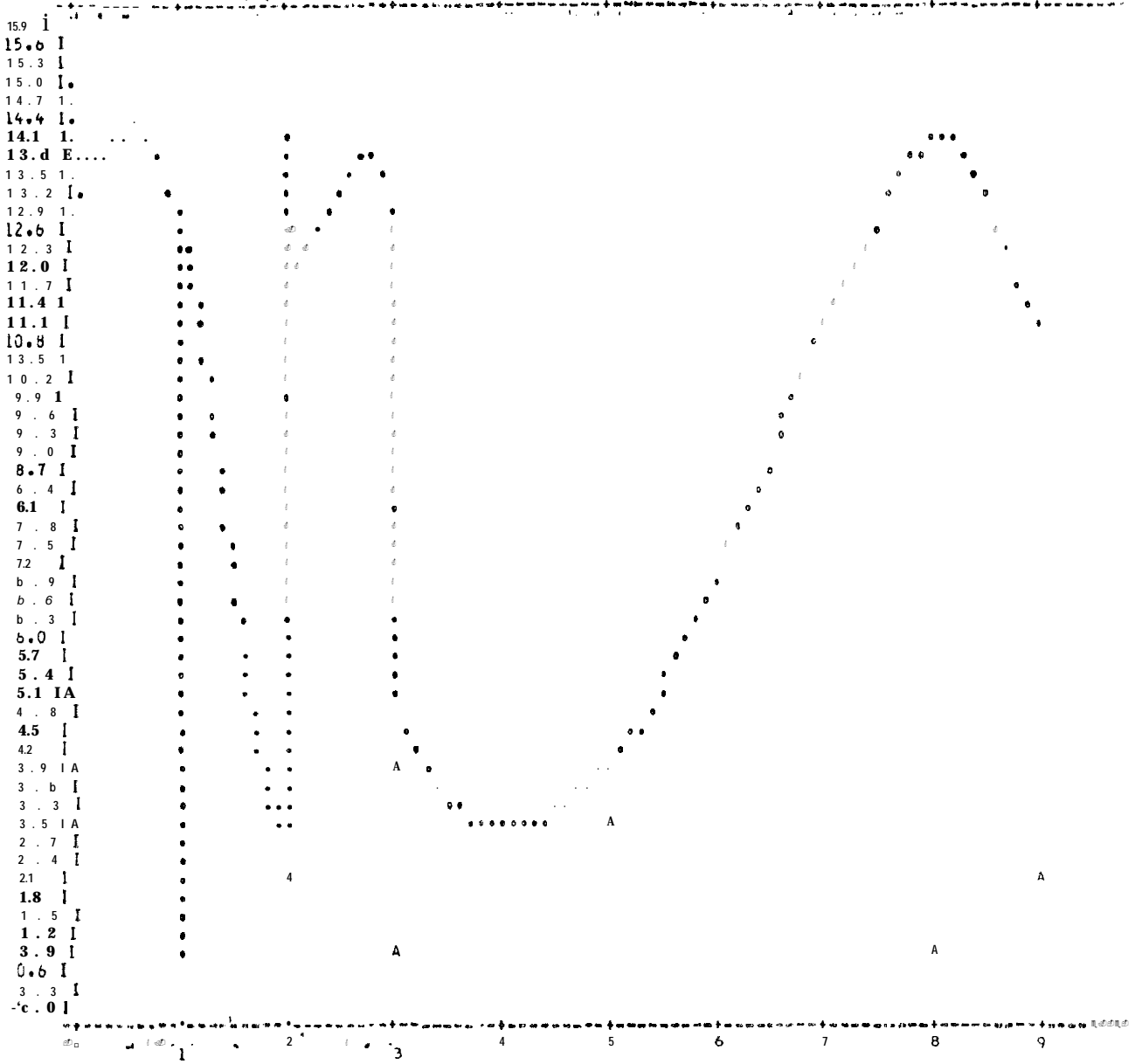
```
20 EXPRESSIONS AVEC VARIABLES "LOGIQUES"
21 'CAL:' Y=1*(X(1).GE.2.AND.X(1).LT.4)+3*(X(1).GE.6)
22 'PRIN/P' Y,X(1)$5.0
23 QUELQUES SORTIES GRAPHIQUES
24 'VARI' X(4)$45 *EQUA*X(4)=X(1...3)
25 HISTOGRAMMES
26 'HIST' X(4)
27 GRAPHS
28 'HEAD' H1='PS' : H2='A'
29 'GRAPH' X(1,2);X(3)$H1;H2
30 'GRAPH/NCF=50,NRF=15' X(1,2);X(3)$H1;H2
31 'RUN'
```

Y	X(1)
0	1
1	2
1	3
0	4
0	1
1	2
1	3
0	4
0	5
3	6
0	1
0	4
0	1
3	7
3	8

2.5	17	*****
5.0	12	*****
7.5	2	**
10.0	3	***
12.5	4	****
15.0	7	*****
	0	

MISSING VALUES 0

SCALE: 1.30 ASTERISK(S) REPRESENT UNIT





GENSTAT V MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 -- INSTALLE AU C.R.O.D.T. --

```

1 *REFE* EXER_6
2 'UNIT' $30
3 *NAME* NREG=NORD,CENTRE,SUD
4 *NAME* NLIEU=ST_LOUIS,KAYAR,YOFF,SOUMB.,HANN,MBOUR,JOAL,DJIFERE,
5 RUFISQUE
6 *FACT* LIEU,ORIGINE$NLIEU
7 *FACT* REGION,REGORI$NREG
8 *READ/PRIN=DEM,FLEV=F* LIEU,ORIGINE,ACT,MOT,P,PRISE
9 *RUN*
```

```

10 1 1 3 25 10000 100
11 1 1 3 40 01000 100
12 1 1 1 8 0 0 0 0 1 3 4
13 1 1 1 8 0 0 0 0 1 5 0
14 1 1 0 0 0 0 0 0 1 7 5
15 2 1 3 2 5 1 0 0 0 0 6 3
16 2 1 0 40 01000 70
17 2 2 1 8 00001 22
18 221 8 00001 80
19 2 2 3 0 00001 56
20 2 2 1 25 00010 73
21 3 1 3 25 01000 111
22 3 3 1 8 03001 12
23 3 3 0 8 00010 57
24 4 4 1 0 0 0 5 0 1 50
25 4 4 1 25 10000 85
26 4 4 1 40 01000 86
27 551 8 00001 43
28 511 8 00001 41
29 5 1 1 8 00001 65
30 5 1 3 25 00001 100
31 9 9 1 00 00001 22
32 6 3 3 01 00001 23
33 771 8 00001 34
34 7 1 1 8 03001 34
35 7 1 1 8 00100 32
36 7 1 1 25 00100 88
37 7 5 1 25 10000 107
38 7 5 1 25 01000 56
39 8 8 1 00 00001 12
```

```

4 0 *EOD*
ACT MNMINMAX 0.6667 0.0000 1.0000 30 VALUES 0 MISSING
MOT MNMINMAX 14.7333 0.0000 40.0000 30 VALUES 0 MISSING
P MNMINMAX 1507.8999 1.0000 10000.0000 30 VALUES 0 MISSING **
P R I S E MNMINMAX 59.3667 12.0000 111.0000 30 VALUES 0 MISSING
```

```

41 *VARI* MLIMITE=.5,1.5,10,26
42 *NAME* NMOT=VOILE,P_INCON,8,25,40
43 *FACT* PUISSANCESNMOT
44 *GROUP* PUISSANCE=LIMITS(MOT;MLIMITE)
45 *VARI* ELIMITE=2,20,200,2000
46 *NAME* NENGIN=LFDC,SP,FME,STPP,STPF
47 *FACT* ENGIN$NENGIN
48 *GROUP* ENGIN=LIMITS(P;ELIMITE)
49 *INTEGER* REGROUP=1,-2,3,4,5,-9,6,7,-8
50 *GROUP* REGION=GROUP(LIEU;REGROUP)
51 *GROUP* REGORI=GROUP(ORIGINE;REGROUP)
52 *TABLE/M* TX,MIGRE,TP$LIEU,ORIGINE
53 *TABU* ACT;TX;MIGRE
54 *CALC* TX=100*TX/MIGRE
55 *PRIN* MIGRE, TX 19.0
56 *PERC* TP;MIGRE$LIEU
57 *PRIN* TP$9.0
58 *LINE* 3
59 *CAPT* .. TAUX D'ACTIVITE **
60 *TABLE/M* T$REGION,PUISSANCE,ENGIN
61 *TABJ* ACT;T
62 *PRIN* T$8.0
63 *RUN*
```

ORIGINE	MIGRE										
ST_LOUIS	ST_LOUIS	KAYAR	YOFF	SOUMB.	HANN	MBOUR	JOAL	OJIFERE	RUFISQUE	MARGIN	
LIEU											
ST_LOUIS	5	0	0	0	0	0	0	0	0	5	
KAYAR	2	4	0	0	0	0	0	0	0	6	
YOFF	1	0	2	0	0	0	0	0	0	3	
SOUMB.	0	0	0	3	0	0	0	0	0	3	
HANN	3	0	0	0	1	0	0	0	0	4	
MBOUR	0	0	1	0	0	0	0	0	0	1	
JOAL	3	0	0	0	0	2	1	0	0	6	
DJIFERE	0	0	0	0	0	0	0	1	0	1	
RUFISQUE	0	0	0	0	0	0	0	0	1	1	
MARGIN	14	4	3	3	1	2	1	1	1	30	

ORIGINE	TX									
ST_LOUIS	ST-LOUIS	KAYAR	YOFF	SOUMB.	HANN	MBOUR	JOAL	OJIFERE	RUFISQUE	MARGIN
LIEU										
ST_LOUIS	40	*	*	*	*	*	*	*	*	40
KAYAR	0	75	*	*	*	4	*	*	*	50
YOFF	0	*	50	*	*	*	*	*	*	33
SOUMB.	*	*	*	100	*	*	*	*	*	100
HANN	67	*	*	*	100	*	*	*	*	75
MBOUR	*	*	3	*	*	*	*	*	*	0
JOAL	100	*	*	*	*	100	100	*	*	100
DJIFERE	*	*	*	*	*	*	*	100	*	100
RUFISQUE	*	*	*	*	*	*	*	*	100	100
MARGIN	50	75	33	100	100	100	100	100	100	67

ORIGINE	TP									
ST_LOUIS	ST-LOUIS	KAYAR	YOFF	SOUMB.	HANN	MBOUR	JOAL	DJIFERE	RUFISQUE	MARGIN
LIEU										
ST_LOUIS	36	0	0	0	0	0	0	0	0	17
KAYAR	14	100	67	0	0	0	0	0	0	20
YOFF	7	0	0	0	0	0	0	0	0	10
SOUMB.	0	0	0	100	0	0	0	0	0	10
HANN	21	0	0	0	100	0	0	0	0	13
MBOUR	0	0	33	0	0	0	0	0	0	3
JOAL	21	0	0	0	0	100	100	0	0	20
DJIFERE	0	0	0	0	0	0	0	100	0	3
RUFISQUE	0	0	0	0	0	0	0	0	100	3
MARGIN	100	100	100	100	100	100	100	100	100	100

UNKNOWN TP \*

TAUX D'ACTIVITE

REGION	ENGIN	LFDC	SP	FME	STPP	STPF	MARGIN
NORD	PUISSANC						
	VOILE	2	3	0	0	0	2
	P-INCON	0	0	0	0	0	0
	8	4	0	0	0	0	4
	25	0	1	0	0	2	3
	40	0	0	0	2	0	2
	YARGIN	6	1	0	2	2	11
CENTRE	VOILE	2	0	0	0	0	2
	P-INCON	0	0	0	0	0	0
	8	4	1	0	0	0	5
	25	1	0	0	1	1	3
	40	0	0	0	1	0	1
	MARGIN	7	1	0	2	1	11
	SUD	VOILE	1	0	0	0	0
P-INCON		1	0	0	0	0	1
8		2	0	1	0	0	3
25		0	0	1	1	1	3
40		0	0	0	0	0	0
HARGIN		4	0	2	1	1	8
MARGIN		VOILE	5	0	0	0	0
	P-INCON	1	0	0	0	0	1
	8	10	1	1	0	0	12
	25	1	1	1	2	4	9
	40	0	0	0	3	0	3
	YARGIN	17	2	2	5	4	30

```

64 *CALC*          L,0=FLOAT(LIEU,ORIGINE)
65 *CALC*          X=L.NE.0
66 *VARI*          LIM=0.5
67 *FACT*          MIGRANT$2
68 *GROU*          MIGRANT=LIMITS(X;LIM)
69 *REST*          REGION,REGORI,ACT$MIGRANT=2
70 *TABLE/M*       TM$REGION,REGORI
71 *TABJ*          ACT; ;TM
72 *PRI Y*         TM$9.0
73 *RUN*
    
```

	TM			
REGORI	NORD	CENTRE	SUD	HARGIN
REGION				
NORD	2	0	0	2
CENTRE	4	0	0	4
SUD	3	1	2	6
MARGIN	9	1	2	12

```

74 *FACT*          QUART$4
75 *TABLE/M*       PUIS_PEC$REGION,ENGIN
76 *TABJ*          PRISE; ;PUIS_PEC
77 *PRIN/M*        ARG=MEAN*PUIS_PEC$8.0
79 *RUN*
    
```

ENGIN	REGION	LFDC	SP	FME	STPP	STPF	MEAN
	NORD	0	0	0	70	63	67
	CENTRE	69	0	0	111	0	79
	SUD	29	0	60	56	107	57
MEAN		53	0	60	79	85	66

GENSTAT V MARK 4.03  
 (C) 1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 -- INSTALL EA U C.R.O.D.T.

```

1 *REFE/NID=1000*
2 *VARI*EE$900 *CALC*EE=RANDU(40)
3 *UNIT*$15
4 *VARI*A(1...15),B(1...15)$30
5 *SCAL*MX(1...15),ME(1...30),SE,SX *CALC*SE=SQRT(1800):SX=SQRT(36001
6 *EQUA*A(1...15),B(1...15)=EE
7 *CALC*MX(1...15),ME(1...15)=MEAN(A(1...15),B(1...15))
8 *EQUA*X,E=MX(1...15),ME(1...15)
9 *CALC*E=(E-.5)*SE :X=20+(X-.5)*SX :SE,SX=VAR(E,X) *PRIN*SE,SX$10.2
10 *VART*FV=5(1,2,3)
11 *CALC*Y=(E+3*X)+10*(FV.EQ.2)+25*(FV.EQ.3)
12 *PRIN/P*X,E,FV,Y$10.2
13 *FACT*F$3=5(1...3)
14 *RUN*
```

SE 5.14  
 SX 9.92

X	E	FV	Y
19.68	-2.19	1.00	56.85
24.91	-2.18	1.00	72.54
21.09	0.43	1.00	63.69
19.04	2.97	1.30	60.08
19.78	-1.24	1.00	58.09
23.28	2.25	2.00	82.09
22.98	2.73	2.00	01.67
22.19	-4.77	2.00	71.79
16.91	-2.13	2.00	58.61
19.64	0.89	2.00	69.81
20.92	-2.19	3.00	85.56
22.93	-2.11	3.00	91.70
19.08	0.55	3.00	82.78
13.67	1.73	3.00	67.73
14.70	-1.27	3.00	67.82

```

15 *UNIT*$15
16 *REGR/OVSET=F*Y+X*F *Y*Y *FIT*X :F : F*X;RES=R :F*X.F
17 *TREAT* F *COVA*X *ANOVA*Y;RES=R1
18 *PRIN/P* R.R1RES10.2
19 *RUN*
```

16.....

\*\*\*\* REGRESSION ANALYSIS \*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VAR I ATE:Y

	ESTIMATE	S.E.	T
CONSTANT	38.6720	17.6184	2.19
I	1.6315	0.8687	1.88

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	1	370	369.6
RESIDUAL	13	1302	104.8
TOTAL	14	1732	123.7
CHANGE	-1	-370	369.6

PERCENTAGE VARIANCE ACCOUNTED FOR 15.3

16.....

2 0

\*\*\*\*\* REGRESSION ANALYSIS \*\*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: Y

	ESTIMATE	S.E.	T
CONSTANT	79.1169	4.0943	19.32
F 1	-16.8675	5.7903	-2.91
F 2	-6.3241	5.7903	-1.09
F 3	0.0000	0.0000	*

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	2	726	363.06
RESIDUAL	12	1006	83.82
TOTAL	14	1732	123.71
CHANGE	*	-356	*

PERCENTAGE VARIANCE ACCOUNTED FOR 31.1

16.....

\*\*\*\*\* REGRESSION ANALYSIS \*\*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: Y

	ESTIMATE	S.E.	T
CONSTANT	26.9446	4.4181	6.10
F 1	-24.4062	1.7015	-14.34
F 2	14.1562	1.7104	8.28
F 3	0.0000	0.0000	*
X	2.8574	0.2341	12.21

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	3	1662.81	554.270
RESIDUAL	11	69.13	6.285
TOTAL	14	1731.94	123.713
CHANGE	*	-936.69	9

PERCENTAGE VARIANCE ACCOUNTED FOR 94.9

16.....

\*\*\*\*\* REGRESSION ANALYSIS \*\*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: Y

	ESTIMATE	S.E.	T
CONSTANT	29.7081	5.8372	5.09
F 1	-20.6701	12.5449	-1.65
F 2	-28.8563	11.3979	-2.53
F 3	0.0000	0.0000	*
X.F 1	2.5464	0.5287	4.82
X.F 2	3.4258	0.4631	7.40
X.F 3	2.7061	0.3138	8.62

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	5	1675.81	335.163
RESIDUAL	9	56.13	6.236
TOTAL	14	1131.94	123.110
CHANGE	*	-13.00	*

PERCENTAGE VARIANCE ACCOUNTED FOR 95.0

\*\*\*\*\* ANALYSIS OF VARIANCE \*\*\*\*\*

VARIATE: X

SOURCE OF VARIATION	DF	SS	SS%	MS	VU
*UNITS* STRATUM					
F	2	24.140	17.38	12.070	1.263
RESIDUAL	12	114.722	32.62	9.560	
TOTAL	14	138.862	100.00	9.919	
GRAND TOTAL	14	138.862	100.00		
GRAND MEAN		20.05			
TOTAL NUMBER OF OBSERVATIONS		15			

\*\*\*\*\* ANALYSIS OF VARIANCE \*\*\*\*\*

VARIATE: Y

SOURCE OF VARIATION	DF	SS	SS%	MS	VR
*UNITS* STRATUM					
F	2	726.12	41.93	363.06	4.331
RESIDUAL	12	1005.82	58.07	83.82	
TOTAL	14	1731.94	100.00	123.71	
GRAND TOTAL	14	1731.94	100.00		
GRAND MEAN		71.4			
TOTAL NUMBER OF OBSERVATIONS		15			

\*\*\*\*\* ANALYSIS OF VARIANCE \*\*\*\*\*  
(ADJUSTED FOR COVARIATES)

VARIATE: Y

SOURCE OF VARIATION	DF	SS	SS%	MS	VR	COV EF
*UNITS* STRATUM						
F	2	1293.182	14.67	646.591	102.884	0.905
COVARIATES	1	936.690	54.08	936.690	149.044	
RESIDUAL	11	69.131	3.99	6.285		13.337
TOTAL	14	2299.004	132.74	164.215		
GRAND TOTAL	14	2299.004	132.14			
GRAND MEAN		71.4				
TOTAL NUMBER OF OBSERVATIONS		15				

\*\*\*\*\* COVARIANCE REGRESSIONS \*\*\*\*\*

COVARIATE	COEFFICIENT	SE	R	RI	c
*UNITS* STRATUM					
X	2.86	0.234			
			-1.92	-1.92	-2.19
			-1.16	-1.16	-2.18
			0.90	0.90	0.43
			3.15	3.15	2.97
			-0.96	-0.96	-1.24
			2.78	2.78	2.25
			3.22	3.22	2.73
			34.39	-4.39	4.77
			-2.51	-2.51	-2.13
			0.95	0.90	0.89
			-1.16	-1.16	-2.19
			-0.78	-0.78	-2.11
			1.33	1.33	0.55
			1.73	1.73	1.73
			-1.12	-1.12	-1.27

\*\*\*\*\* STANDARD ERRORS OF DIFFERENCES OF MEANS \*\*\*\*\*

TABLE	F
REP	5
SED	1.67

\*\*\*\*\* STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION \*\*\*\*\*

STRATUM	OF	SE	CV%
*UNITS*	11	2.51	3.5

GENSTAT v MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 AUTAC.R.O.D.T.

```

1 *REFE*
2 *UNIT*$50
3 *READ/PRIN=DEM*TAILLE,CRTM,CRTS,CRTF,SAL
4 *REGR/PRIN=C*TAILLE,CRTM,CRTS,CRTF,SAL
5 *Y* TAILLE
6 *FIT*CRTM,CRTS,CRTF,SAL
7 ● FIT*CRTS,SAL
8 *REGR/PRIN=C*TAILLE,CRTM,CRTS,CRTF
9 *Y* TAILLE
10 *FIT*CRTM,CRTS,CRTF
11 *CALC*SAL2=SAL*SAL
12 *REGR/PRIN=C*CRTS,SAL,SAL2,TAILLE
13 *Y* TAILLE
14 *FIT*SAL,CRTS;FVAL=F1
15 *ADD*SAL2;FVAL=F;RES=R *HIST*R
1b *GROU*FCRTS=INTPT(CRTS) *GROU*FSAL=QUANT(SAL;5)
17 *RUN*
```

```

18 18.8 74 1014 1 41
19 17.7 74 101 41 *
20 19.2 74 101 41 24
21 17.8 74 101 41 31
22 18.2 74 101 41 4 6
23 18.5 74 101 41 33
24 18.3 74 101 41 42
25 19 74 101 41 43
26 20.6 62 84 22 3 8
27 18.4 62 84 22 6
28 21.6 62 84 22 3 9
29 22.9 62 84 22 4 7
30 23.1 52 84 22 2 3
31 23.1 62 84 22 *
32 21.8 62 84 22 3 3
33 21.8 62 84 22 5 0
34 20.3 62 84 22 3 3
35 22.4 62 84 22 4 8
36 20.8 62 84 22 4 4
37 19.7 62 84 22 5 2
38 20.9 34 65 18 *
39 21.2 34 65 18 *
40 21.7 34 65 18 *
41 26.9 34 65 18 60
42 29.9 34 65 18 23
43 20.4 34 65 18 3
44 26.2 34 65 18 35
45 20.6 34 65 18 54
46 25.3 34 65 18 30
47 21.6 34 65 18 56
48 21 34 65 18 49
49 16.5 34 6 5 1861
50 19.6 34 36 20 *
51 26.5 34 36 20 3 3
52 22.6 34 36 20 5 5
53 29.1 34 36 20 3 3
54 23.3 34 36 20 5 9
55 25.7 34 36 20 5 2
56 16.2 34 36 20 6 9
57 28.5 28 30 14 30.5
58 32.7 23 30 14 2 9
59 26.28 30 14 5 1
60 20.6 28 33 14 6 0
61 28.0 23 30 14 5 4
62 13.1 28 30 14 7 7
63 19.1 * 101 * 34.5
64 20.9 * 84 * 38
65 4 * 65 * 43
66 22.9 * 36 * 47
67 23.9 * 30 * 50
68 * EOD*
```

TAILLE	MNMINMAX	21.9550	13.1000	32.7000	50	VALUES	1 MISSING
CRTM	MNMINMAX	47.7778	28.0000	74.0000	50	VALUES	5 MISSING
CRTS	MNMINMAX	66.8800	30.0000	101.0000	50	VALUES	0 MISSING
CRTF	MNMINMAX	22.9333	14.0000	41.0000	50	VALUES	5 MISSING
SAL	MNMINMAX	43.1628	4.0000	77.0000	50	VALUES	7 MISSING

4.....

\*\*\*\*\* CORRELATION MATRIX \*\*\*\*\*

DF = 36

TAILLE	1	2	3	4	5
CRTM	1.0000				
CRTS	-0.5009	1.0000			
CRTF	-0.4799	0.9220	1.0000		
SAL	-0.4752	0.8321	0.7707	1.0000	
	-0.3019	-0.31363	-0.3871	-0.2735	1.0000

6.....

\*\*\*\*\* REGRESSION ANALYSIS \*\*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: TAILLE

	ESTIMATE	S.E.	T
CON47 ANT	38.224518	2.617660	14.60
CRTM	-0.100739	0.082384	-1.22
CRTS	-0.039043	0.049295	-0.79
CRTF	-0.045299	0.099664	-0.45
SAL	-0.172801	0.037640	-4.59

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	4	369.9	92.473
RESIDUAL	33	301.6	9.139
TOTAL	37	671.5	18.148
CHANGE	-4	-369.9	92.473

PERCENTAGE VARIANCE ACCOUNTED FOR 49.6

7.....

\*\*\*\*\* REGRESSION ANALYSIS \*\*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: TAILLE

	ESTIMATE	S.E.	T
CONSTANT	37.18913	1.58304	14.40
CRTS	-0.11440	3.02091	-5.47
SAL	-0.16928	7.03787	-4.47

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	2	342.5	171.241
RESIDUAL	35	329.0	9.400
TOTAL	37	671.5	18.148
CHANGE	*	17.4	*

PERCENTAGE VARIANCE ACCOUNTED FOR 50.2

8.....

\*\*\*\*\* CORRELATION MATRIX \*\*\*\*\*

DF = 43

TAILLE	1	2	3	4
CRTM	1.0000			
CRTS	0.4575	1.0000		
CRTF	-0.4583	0.9049	1.0000	
SAL	-0.4676	0.8362	0.7630	1.0000

10.....



\*\*\*\* REGRESSION ANALYSIS \*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: TAILLE	ESTIMATE	S.E.	T
CONSTANT	27.4878998	1.6213398	16.95
CHTH	-0.0028236	0.0846922	-0.03
CRTS	1.0374253	0.0508618	0.74
CRTF	-0.1253507	0.1118116	-1.12

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	3	171.0	57.01
RESIDUAL	41	531.9	12.97
TOTAL	44	702.9	15.97
CHANGE	3	-171.0	57.01

PERCENTAGE VARIANCE ACCOUNTED FOR 18.8

12.....

\*\*\*\* CORRELATION MATRIX \*\*\*\*

DF = 40

CRTS	1	1.0000			
SAL	2	-0.4072	1.0000		
SAL2	3	0.4410	0.9729	1.0000	
TAILLE	4	-0.4859	-0.2788	-0.3398	1.0000
		1	2	3	4

14.....

\*\*\*\* REGRESSION ANALYSIS \*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: TAILLE	ESTIMATE	S.E.	T
CONSTANT	36.76437	2.41057	15.25
SAL	-0.16816	0.03617	-4.65
CRTS	0.11009	0.01883	5.85

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	2	346.7	173.357
RESIDUAL	39	335.1	8.592
TOTAL	41	681.8	16.629
CHANGE	-2	-346.7	173.357

PERCENTAGE VARIANCE ACCOUNTED FOR 48.3

15.....

\*\*\*\* REGRESSION ANALYSIS \*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: TAILLE	ESTIMATE	S.E.	T
CONSTANT	27.5467529	2.61541142	10.53
SAL	0.3843000	0.1123288	3.42
CRTS	-0.1263380	0.0150620	-8.39
SAL2	-0.0066060	0.0012999	-5.08

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	3	482.3	160.766
RESIDUAL	38	199.5	5.250
TOTAL	41	681.8	16.629
CHANGE	-1	-135.6	135.584

PERCENTAGE VARIANCE ACCOUNTED FOR 68.4



GENSTAT v MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 INSTALLE AU C.R.O.D.T.

```

1 *REFE/NUNN=300,NID=300* ESTIME
2 *UNITS* $60
3 *NAME* TAILLE=PETIT,MOYEN,GROS
4 *FACT* CAT$TAILLE,600
5 *NAME* NE=ENQT-1,ENQT-2,ENQT-3,ENQT-4,ENQT-5,
6 ENQT-6,ENQT-7,ENQT-8,ENQT-9,ENQT-10
7 *FACT* ENQT$NE=60(1...10)
8 *FACT* TAS$60=(1...60)10
9 *READ* Y1,X(1...10)$S ,1X,1,1X,2,1X,8,1X
10 *GPOLIPER* C1=QUANT(Y1;3)
11 *VARI* R,Y,X,XT$600
12 *EQUA* Y=10(Y1) :X=X(1...10) :CAT=10(C1)
13 *CALC* XT=(X-Y)/Y
14 *HEAD* N=0 :ERREURS RELATIVES LORS D'ESTIMATIONS A VUE"
15 *DESC* XT$3;N
16 *TREAT* ENQT*CAT
17 *ANOVA* XT
18 *TREAT* ENQT+CAT/TAS+ENQT.CAT
19 *ANOVA/PR=10* XT;RES=R
20 *HIST* R
21 *RUN*
```

Variable	MNMINMAX	47.6000	1.0000	265.0000	60	VALUE S	0	WISSING	**
Y1	MNMINMAX	47.6000	1.0000	265.0000	60	VALUE S	0	WISSING	**
X(1)	MNMINMAX	41.9138	1.0000	200.0000	60	VALUES	2	MISSING	**
X(2)	MNMINMAX	49.7750	1.5000	165.0000	60	VALUES	0	MISSING	**
X(3)	MNMINMAX	41.8700	0.6000	150.0000	60	VALUE S	0	HISSING	**
X(4)	MNMINMAX	78.2500	1.5000	450.0000	60	VALUES	0	MISSING	**
X(5)	MNMINMAX	32.3627	0.4000	160.0000	60	VALUES	1	MISSING	**
X(6)	MNMINMAX	36.5017	0.4000	160.0000	60	VALUES	0	MISSING	**
X(7)	MNMINMAX	41.7917	1.0000	200.0000	60	VALUES	0	WISSING	**
X(8)	MNMINMAX	30.1817	1.0000	125.0000	60	VALUES	0	MISSING	**
X(9)	MNMINMAX	39.6661	0.8000	180.0000	60	VALUES	1	MISSING	**
X(10)	MNMINMAX	69.8772	0.5000	300.0000	60	VALUE S	3	MISSING	**

7 1

\*\*\*\*\*ANALYSIS OF VARIANCE\*\*\*\*\*

VARIATE: XT :ERREURS RELATIVES LORS D'ESTIMATIONS 3 VUE

SOURCE OF VARIATION	DF (MV)	SS	SS%	MS	VR
*UNITS* STRATUM					
ENQT	9	38.31779	45.29	4.25753	71.323
CAT	2	2.00456	2.37	1.00228	16.790
ENQT.CAT	18	11.60786	13.72	0.64488	10.803
RESIDUAL	563( 7)	33.60716	39.72	0.05969	
TOTAL	592	05.53796	101.10	0.14449	
GRAND TOTAL	592	05.53796	101.10		

ESTIMATED GRAND MEAN -0.026  
 TOTAL NUMBER OF OBSERVATIONS 600  
 NUMBER OF MISSING VALUES 7  
 MAXIMUM NUMBER OF ITERATIONS 2

UNIT NUMBER	ESTIMATED VALUE
19	-0.163
26	20.094
296	-0.575
534	0.161
568	0.507
574	0.507
593	-0.156

\*\*\*\*\*TABLES OF MEANS\*\*\*\*\*

VARIATE: XT :ERREURS RELATIVES LORS D'ESTIMATIONS 3 VUE

GRAND MEAN	-0.026									
ENQT	ENQT-1	ENQT-2	ENQT-3	ENQT-4	ENQT-5	ENQT-6	ENQT-7	ENQT-8	ENQT-9	ENQT-10
	-0.126	0.231	-0.140	0.450	-0.391	-0.039	-0.068	-0.341	-0.100	-0.261
CAT	PETIT	MOYEN	GROS							
	-0.097	0.044	-0.026							
CAT	PETIT	MOYEN	GROS							
ENQT-1	-0.094	-0.163	-0.121							
ENQT-2	3.351	5.313	3.028							
ENQT-3	-0.281	-0.042	-0.098							
ENQT-4	3.304	0.305	0.742							
ENQT-5	10.575	-0.302	-0.297							
ENQT-6	0.076	0.109	-0.303							
ENQT-7	-0.144	0.070	0.131							
ENQT-8	10.345	-0.327	-0.352							
ENQT-9	-0.111	-0.029	-0.161							
ENQT-10	10.156	0.500	0.431							

\*\*\*\*\*STANDARD ERRORS OF DIFFERENCES OF MEANS\*\*\*\*\*

TABLE	ENQT	CAT	ENQT.CAT
REP	60	200	20
SED	0.0446	0.0244	0.0773

(NOT ADJUSTED FOR MISSING VALUES)

\*\*\*\*\*STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION\*\*\*\*\*

STRATUM	DF	SE	CVY
*UNITS*	563	5.2443	922.8

44444 ANALYSIS OF VARIANCE 44444

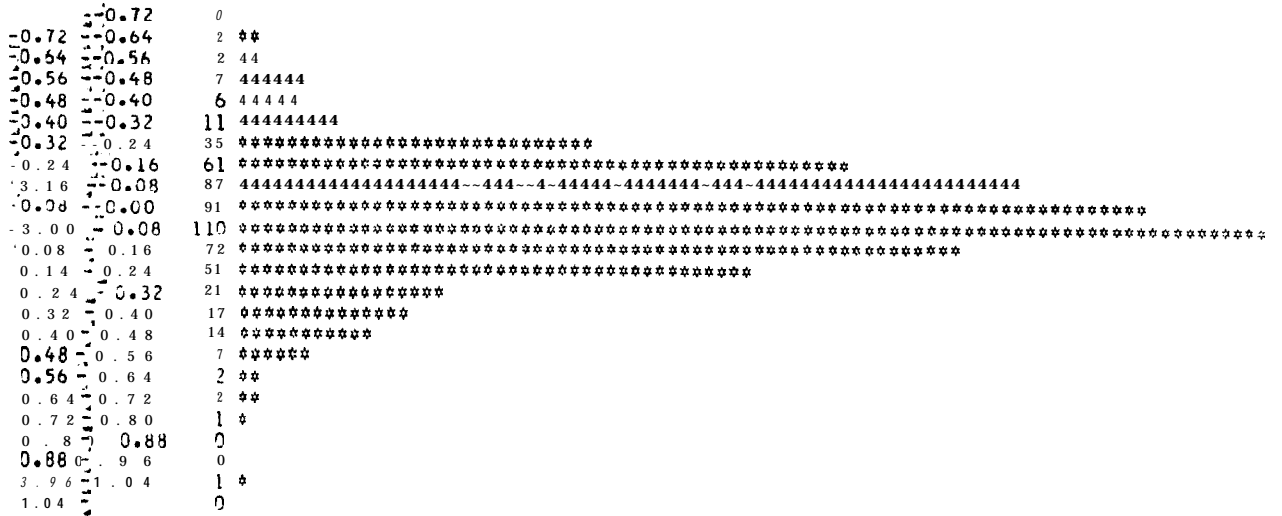
VARIATE: XT ERREURS RELATIVES LORS D'ESTIMATIONS.VUVE

SOURCE OF VARIATION	DF(MV)	SS	SS%	MS	VR
#UNITS# STRATUM					
ENQT	9	38.3721	45.35	4.26357	18.632
CAT	2	1.97694	2.34	0.98947	18.230
CAT.TAS	51	0.19143	7.32	0.10862	2.003
ENQT.CAT	18	11.61766	13.73	0.64543	11.903
RESIDUAL	5061 71	27.43622	32.43	0.05422	
TOTAL	592	85.59436	101.16	0.14459	
GRAND TOTAL	592	05.59436	101.16		

ESTIMATED GRAND MEAN -0.027  
TOTAL NUMBER OF OBSERVATIONS 600  
NUMBER OF MISSING VALUES 1  
MAXIMUM NUMBER OF ITERATIONS 2

UNIT NUMBER	ESTIMATED VALUE
19	-0.288
26	-0.059
296	-0.597
534	-0.200
568	0.509
574	0.488
593	-0.114

HISTOGRAM OF R



MISSING VALUES 0

SCALE: 3.82 ASTERISK(S) REPRESENT 1 UNIT

83 °CLOSE\*  
4446444 END OF ESTIME AT LINE 19 USED 3962 LEFT 46038

GENSTAT V MARK 4.03  
(C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
INSTALLÉ AU C.R.O.D.T.

1 \*REFE\*VACHES  
2 \*\* ETUDE PAR ANALYSE EN COMPOSANTES PRINCIPALES DE 23  
3 VACHES DE RACE ZÉBU ET CHAROLAISES LES SIX VARIABLES  
4 ETUDIÉES SONT :  
5 - LE POIDS VIF  
6 - LE POIDS DE LA CARCASSE  
7 - LE POIDS DE VIANDE DE PREMIÈRE QUALITÉ  
8 - LE POIDS DE VIANDE TOTAL  
9 - LE POIDS DE GRAS  
10 - LE POIDS D'OS  
11 (SOURCE : MANUEL GENSTAT - INRA - HAI 1982 )\*\*

12 \*UNIT\$23  
13 \*NAME\*NV  
14 \*READ/PRIN=DEM\*NV,X(1...6)  
15 \*FACT\*NUM\$NV=1...23  
16 \*MATR\*COMP\$23,2  
17 \*PCP/PRIN=L,CORR=Y,NLR=4\*X(1...6);SCORES=COMP  
18 \*REGR/PRIN=C\*X(1...6)  
19 \*EQUA\*Y(1,2)=COMP\$(1,1X)23,1X  
20 \*GRAP/NCF=50,NRF=20\*Y(1);Y(1)\$;NOM  
21 \*ACTRACES2=12111,11(21  
22 \*DSSP\*INTRA\$X(1...6);RACE;M(1...6),N  
23 \*SSP/PRIN=SC\*INTRA \*PRIN/P\*M(1...6),NS7.1  
24 \*CVA/PRIN=LMD\*INTRA  
25 \*RUN\*

Table with 6 columns of numerical data for various units (IA, IB, IC, etc.)

Summary statistics table for variables X(1) to X(6), including MN, MIN, MAX, and missing values.

\*\*\*\*\* PRINCIPAL COMPONENTS ANALYSIS \*\*\*\*\*

\*\*\* LATENT ROOTS \*\*\*

Table showing latent roots for components 2, 3, and 4.

\*PERCENTAGE VARIANCE\*

Table showing percentage variance for components 1, 2, 3, and 4.

\*\*\* LATENT VECTORS (LOADINGS) \*\*\*

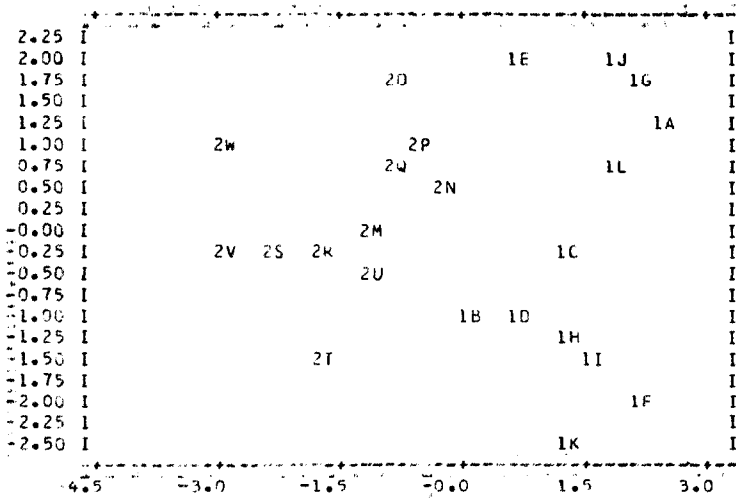
Table showing loadings for variables X(1) to X(6) across components 1, 2, 3, and 4.

3 0

\*\*\*\*\* CORRELATION MATRIX \*\*\*\*\*

DF = 21

X(1)	1						
X(2)	2	0.6914	1.0000	1.0000			
X(3)	3	-0.0344	0.2830	1.0330			
X(4)	4	-0.0585	0.3903	0.8950	1.0000		
X(5)	5	0.1011	3.3331	-0.8584	-0.9062	1.0000	
X(6)	6	0.1112	-0.1908	-0.0643	-0.1295	-0.2705	1.0000
		1	2	3	4	5	6



23.....

\*\*\*\*\* SUMS OF SQUARES AND PRODUCTS MATRIX \*\*\*\*\*

DF = 21

INTRA

X(1)	1	1.5948E	?											
X(2)	2	8.9582E	2	1.2122E	3									
X(3)	3	-1.0548E	2	-7.8191E	1	3.7935E	1							
X(4)	4	-1.0961E	2	-1.7891E	1	2.5215E	1	4.1645E	1					
X(5)	5	1.1849E	2	2.7045E	1	-2.7901E	1	-3.6003E	1	5.4073E				
X(6)	6	-1.7697E	1	-9.8365E	0	1.7790E	0	-2.4918E	0	-2.1171E	1	2.7816E	1	
MEAN	7	4.0161E	?	2.2883E	?	2.9922E	1	7.4670E	1	8.9391E	0	1.6422E	1	2.3000E
		1		2		3		4		5		6		

\*\*\*\*\* CORRELATION MATRIX \*\*\*\*\*

DF = 20

X(1)	1	1.00%					
X(2)	2	0.6889	1.3000				
X(3)	3	-0.4585	-0.3046	1.0000			
X(4)	4	-0.4548	-0.0796	0.6344	1.0000		
X(5)	5	3.4314	0.1056	-0.6160	-0.7587	1.0000	
X(6)	6	-0.0898	0.0536	0.0548	-0.0732	-0.5459	1.0000
		1	2	3	4	5	6

M(1)	M(2)	M(3)	M(4)	M(5)	M(6)	N
403.3	233.0	32.0	76.6	7.2	16.3	12.0
399.7	224.3	27.7	72.4	10.8	16.5	11.0

\*\*\* LATENT ROOTS\*\*\*

8 - 3

	<sup>1</sup>	<sup>2</sup>	<sup>3</sup>	<sup>4</sup>	<sup>5</sup>	<sup>6</sup>
	5.1347	0.0000	0.0000	-0.0000	-0.0000	-0.0000

\*PERCENTAGE VARIANCE\*

	<sup>1</sup>	<sup>2</sup>	<sup>3</sup>	<sup>4</sup>	<sup>5</sup>	<sup>6</sup>
	100.000	0.0000	0.0000	-0.0000	-0.0000	-0.0000

\*\*\* LATENT VECTORS(LOADINGS) \*\*\*

	<sup>1</sup>
X(1)	0.0520
X(2)	0.0430
X(3)	0.6698
X(4)	0.6238
X(5)	0.4565
X(6)	0.3662

VECTORS FOR ROOTS LESS THAN 0.0001 ARE NOT PRINTED.

\*\*\* CONSTANT TERMS(ADJUSTMENT FOR MEANS) \*\*\*

	<sup>1</sup>
	108.0169

\*\*\*\*\* CANONICAL VARIATE MEANS \*\*\*\*\*

<sup>1</sup>	2.5730
<sup>2</sup>	-2.2615

MEANS FOR ROOTS LESS THAN 0.0001 ARE NOT PRINTED

\*\*\*\*\* MAHALANOBIS'S DISTANCES \*\*\*\*\*

<sup>1</sup>	0.0000	
<sup>2</sup>	4.3373	0.0000
	<sup>1</sup>	<sup>2</sup>

50 \*CLOSE\*  
 \*\*\*\*\* END OF VACHS AT LINE 24 USED 986 LEFT 49014



GENSTAT v MARK 4.03
(C) 1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)
-- INSTALLE AU C.R.O.O.T. --

1 \*REFE\*EXER\_9
-2 \*\* RECONSTITUTION DE LA CARTE DU SENEGAL ...
3 (A PEU DE CHOSSES PRESIAVECL'ANALYSE EN COORONNEES PRINCIPALES \*\*
4 \*UNIT/N\* VILLES\$1=DAKAR, MBOUR, ZIGUINCH, KAOLACK, ST\_LOUIS, MATAM, BAKEL
5 , KEDOUGOU, LINGUERE, TAMBA, THIES
6 \*SYM\* DIST, DISTANCE\$VILLES
7 \*\* LECTURE DE LA MATRICE DES DISTANCES ENTRE 10 VILLES \*\*
8 \*READ\* DISTANCE
9 \*PRIN\* DISTANCE\$9.1
10 \*MATR\* COORD\$VILLES, 2 \*DIAG\*H\$2 \*SCAL\*1
-11 \*\* TRANSFORMATION DES DISTANCES, REALISATION DE L'ANALYSE (PCO...)
12 CONSERVATION DES COORDONNEES, VALEUR - PROPRES ET TRACE \*\*
13 \*CALCULER\* DIST=-0.5\*DISTANCE\*DISTANCE
14 \*PCO/PRIN=LT, NLR=3\*DIST; COORD, H, T
15 \*EQUA\*V(1,2)=COORD\$(1,1X)11,1X
16 \*NAME\*NVILLE=DK, MB, ZG, KA, SL, MA, HA, KD, LG, TB, TH
17 \*FACT\*FV\$NVILLE=1...11
18 \*GRAP/NRF=30, NCF=60\*V(2); V(1)\$; FV
19 \*SCAL\*M\* CALC\*M=MAX(DISTANCE) \*CALC\*DISTANCE=1 (DISTANCE/M)
20 \*PRIN\*DISTANCE\$10.2
21 \*NEIGHBOUR/3\*DISTANCE
22 \*MST\*DISTANCE
23 \*RUN\*

Table with 11 columns: DISTANCE, MN, MIN, MAX, and 8 city names (DAKAR, MBOUR, ZIGUINCH, KAOLACK, ST\_LOUIS, MATAM, BAKEL, KEDOUGOU, LINGUERE, TAMBA, THIES). Values represent distances between cities.

14.....

\*\*\*\* PRINCIPAL COORDINATE ANALYSIS \*\*\*\*

\*\*\* LATENT ROOTS \*\*\*

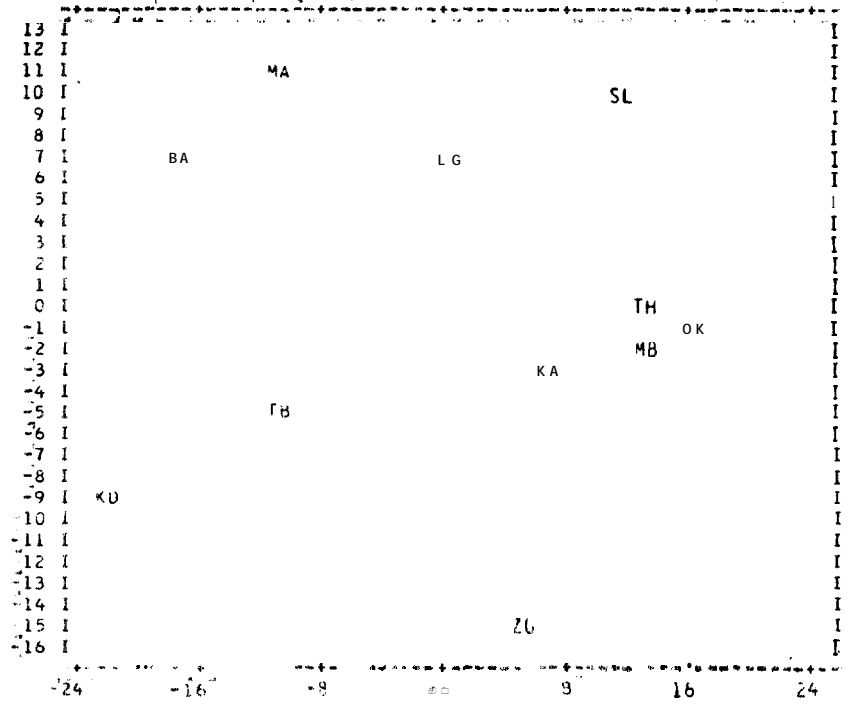
Table with 3 columns: H (1, 2, 3) and values (1907.0090, 660.9382, 49.1549).

\*PERCENTAGE VARIANCE\*

Table with 3 columns: H (1, 2, 3) and values (74.9172, 25.9359, 1.9326).

\*\*\* LATENT VECTORS (COORDINATES) \*\*\*

Table with 4 columns: COORD (1, 2, 3) and 10 city names. Values represent latent vectors for each city.



DISTANCE

DAKAR	1.00											
MBOUR	0.90	1.00										
ZIGUINCH	0.57	0.65	1.00									
KAOLACK	3.74	0.84	0.71	1.00								
ST_LOUIS	6.71	0.70	0.38	0.66	1.110							
MATAM	0.25	0.32	0.23	0.44	0.43	1.00						
BAKEL	0.13	0.20	0.21	0.35	0.27	0.81	1.00					
KEDOUGOU	0.00	0.10	0.20	0.26	0.02	0.41	0.57	1.00				
LINGUERE	0.58	0.63	3.45	0.72	0.74	0.67	0.53	0.28	1.00			
TAMBA	0.32	0.41	0.61	0.57	0.36	0.65	0.70	0.66	0.61	1.00		
THIES	0.91	0.93	0.59	0.81	0.17	0.34	0.22	0.08	0.06	0.06	1.00	

DAKAR MBOUR ZIGUINCH KAOLACK ST\_LOUIS MATAM BAKEL KEDOUGOU LINGUERE TAMBA THIES

\*\*\*\*\* NEIGHBOURS TABLE \*\*\*\*\*

DAKAR	1	11	90.9	2	90.1	4	74.3
MBOUR	2	11	93.3	1	90.1	4	83.7
ZIGUINCH	3	4	71.4	2	64.9	10	61.5
KAOLACK	4	2	03.7	11	81.2	1	74.3
ST_LOUIS	5	11	76.5	9	13.6	1	70.6
MATAM	6	1	80.7	9	67.4	10	65.4
BAKEL	7	6	80.7	10	70.4	8	57.3
KEDOUGOU	8	10	65.9	7	57.3	6	41.2
LINGUERE	9	5	73.6	4	71.9	6	07.4
TAMBA	10	7	70.4	8	65.9	6	65.4
THIES	11	2	93.3	1	90.9	4	91.2

\*\*\* MINIMUM SPANNING TREE \*\*\*

DAKAR	THIES	MBOUR	KAOLACK	ZIGUINCH			
1.....	11.....	2.....	4.....	3			
90.9	{	93.3	03.7	71.4			
	{						
	{	ST-LOUIS	LINGUERE	MATAM	BAKEL	TAMBA	KEDOUGOU
	{	5.....	9.....	6.....	7.....	10.....	8
	{	76.5	73.6	67.4	80.7	70.4	65.9

TOTAL LENGTH 773.6

36 \*CLOSE\*  
 \*\*\*\*\* END OF \*\*\*\*\* XER\_9 AT LINE 14 USED 568 LEFT 49432

3 4

GENSTAT V MARK 4.33  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 INSTALLÉ AU I.R.O.O.T.

```

1 *REFE/NID=1000*
2 *FILE*UF=11
3 *FETCH*UF$CORRESP2,PREPCORR
4 *GET*CORRESP2$CORRESP2 :PREPCORR$PREPCORR
5 *NAME*NCAT=<1,1-4,5-14,15-24,25-34,35-44,45-54,55-64,>65
6 *FACT*CAT_AGE$NCAT
7 *NAME*NPL=FEVRE,PALU,DIARRHEE,A.RESP.A,VERS_INT,CONVULS,MAL_YEUX,
8 STOMATO,OREILLE,PARALYS.,BRUL.,LES.PEAU,ROUGEOLE,COQ.,
9 URETRITE,JAUN.,HTA,TOUX_CH.,DIVERS
10 *FACT*PLAINTES$NPL
11 *TABLE*TSPLAINTES,CAT_AGE
12 *TABLE*TCALCULEE PARTIR DE 0464 PLAINTES**
13 *READ*IT
14 *PRINT$6
15 *VARI*GRAPHE=1,2 *SCAL*NFAC=4
16 *NAME*NVARI=1,4,14,24,34,44,54,64,99 *SCAL*NC=9
17 *NAME*NOBSI=A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S *SCAL*NL=19
18 *PRINT/P*NOBSI,NPL$10
19 *SCAL*PSUP=0:NSUP=0:SEUIL=25
20 *RUN*
```

T MNMINMAX 33.6316 0.0000 457.0000 171 VALUE S 0 MISSING \*\*

CAT_AGE	<1	1-4	5-14	15-24	25-34	35-44	45-54	55-64	>65
PLAINTES									
FEVRE	39	116	146	97	52	15	9	a	1
PALU	57	257	369	179	74	35	22	10	3
DIARRHEE	103	225	77	44	18	7	1	4	5
A.RESP.A	89	189	137	56	37	9	3	5	5
VERS-INT	1b	149	135	29	15	3	3	1	1
CONVULS	0	2	0	0	1	0	0	0	0
MAL-YEUX	15	43	36	22	12	6	3	6	2
STOMATO	19	46	32	18	9	a	2	4	1
OREILLE	15	72	58	37	16	6	5	4	1
PARALYS.	0	1	1	0	0	0	0	0	0
BRUL.	6	22	29	13	7	2	1	0	2
LES.PEAU	102	413	457	177	80	26	12	11	9
ROUGEOLE	14	56	20	a	0	1	0	0	0
COQ.	2	0	2	1	0	0	0	0	0
URETRITE	0	1	0	7	5	3	0	1	1
JAUN.	0	0	0	0	0	0	0	0	0
HTA	0	0	0	3	4	1	7	5	1
TOUX_CH.	0	5	11	22	12	4	4	3	2
DIVERS	37	131	179	201	134	57	32	21	15

UNKNOWN T \*

```

NOBSI NPL
A FEVRE
B PALU
C DIARRHEE
D A.RESP.A
E VERS-INT
F CONVULS
G MAL_YEUX
H STOMATO
I OREILLE
J PARALYS.
K BRUL.
L LES.PEAU
M ROUGEOLE
N COQ.
O URETRITE
P JAUN.
Q HTA
R TOUX_CH.
S DIVERS
```

41 \*USE/P\*PREPCORR\$

FC	1	4	14	24	34	44	54	64	99	MARGIN
FL										
A	39	116	146	97	52	15	9	a	1	483
B	57	257	369	179	74	35	22	10	3	1007
C	103	225	77	44	18	7	1	4	5	484
D	89	189	137	56	37	9	3	5	5	530
E	1b	149	135	29	15	3	3	1	1	352
F	15	43	36	22	12	6	3	6	2	145
G	19	46	32	18	9	a	2	4	1	139
H	15	72	58	37	16	6	5	4	1	214
I	0	1	1	0	0	0	0	0	0	82
J	6	22	29	13	7	2	1	0	2	82
K	102	413	457	177	80	26	12	11	9	1287
L	14	56	20	a	0	0	0	0	0	109
M	0	5	11	22	12	4	4	3	2	63
N	37	131	179	201	134	57	32	21	15	807
O										
P										
Q										
R										
S										
MARGIN	512	1734	1686	903	466	180	97	77	47	5702

UNKNOWN TRAVAIL \*

VALEURS PROPRES      %VARIANCE

1.0000000	*
3.1032724	611.619
3.0349625	23.231
3.0054073	3.593
0.0024987	1.660
0.0019745	1.312
0.0012539	0.833
0.0008843	0.588
0.0002481	0.165

COORDONNEES ET CONTRIBUTIONS RELATIVES ET ABSOLUES DES VARIABLES

FACTEUR	VARIABLES											
	1			2			3			4		
NOMVAR	COO	CTREL	CTABS	COO	CTREL	CTABS	COO	CTREL	CTABS	COO	CTREL	CTABS
1	422	539	155	360	391	333	150	68	374	-20	1	14
4	307	927	277	id	3	3	-83	69	392	5	0	3
14	0	0	0	-240	951	486	50	41	136	3	0	1
24	-339	946	177	35	10	5	7	0	1	-25	5	38
34	-472	a74	177	144	a1	48	19	1	6	72	20	167
44	-543	800	9c	189	97	32	-53	8	17	-70	13	62
54	-708	923	83	94	1b	4	-128	30	51	-41	3	12
54	-488	498	31	320	225	42	-88	1b	19	-195	79	205
39	-377	252	11	445	351	47	-56	6	5	389	268	500

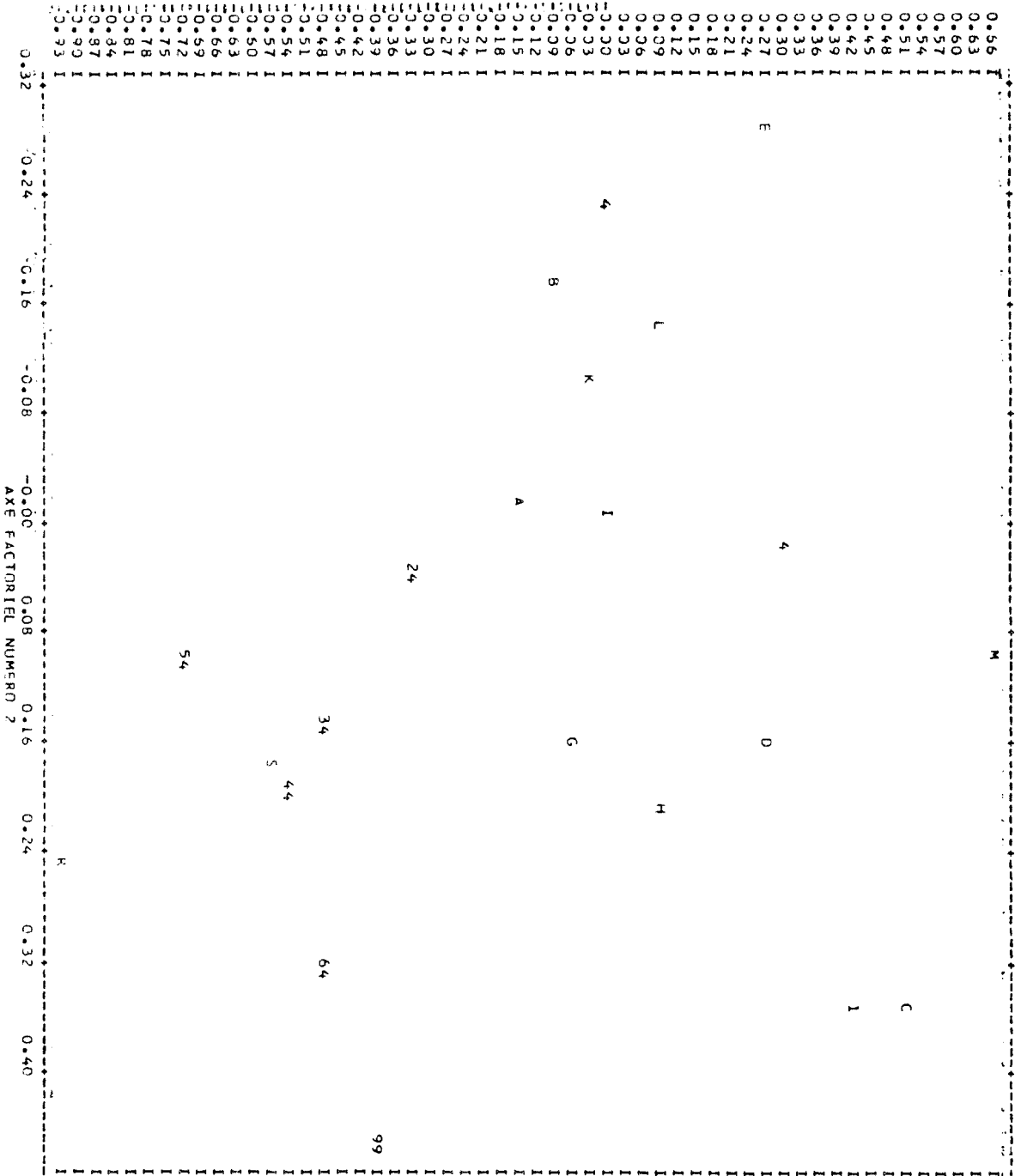
UNKNOWN      VARIABLES      \*

COORDONNEES ET CONTRIBUTIONS RELATIVES ET ABSOLUES OBSERVATIONS

FACTEUR	OBSERVAT											
	1			2			3			4		
NOMOBS	COO	CTREL	CTABS	COO	CTREL	CTABS	COO	CTREL	CTABS	COO	CTREL	CTABS
A	-156	b14	20	-13	5	0	68	117	72	-46	54	73
B	-105	236	19	-174	653	153	32	22	34	-46	46	150
C	519	682	221	352	314	301		0	0	2	0	0
D	280	673	71	-159	218	68	101	87	174	36	11	48
E	273	409	45	-290	458	148	-135	100	209	53	15	69
G	-57	44	1	150	339	19	-47	29	10	-86	98	75
H	83	76	2	208	483	30	-18	4	1	-131	193	168
I	-12	8	9	-6	2	0	-106	623	78	-56	173	47
K	37	21	0	-102	155	4	64	62	11	204	619	239
L	99	310	21	-141	637	128	21	14	18	23	17	47
Y	654	791	79	95	11	5	-300	167	319	-35	2	9
Z	-930	871	92	251	63	20	-126	16	33	38	1	6
S	-560	900	429	175	88	124	-39	4	39	35	4	71

UNKNOWN      OBSERVAT      \*

42 \*RUN\*



A X E F A C T I O R I O D 1

GENSTAT V MARK 4.03  
 (C)1980 LAWES AGRICULTURAL TRUST (ROTHAMSTED EXPERIMENTAL STATION)  
 U. INSTALLEE U (C.R.O.D.T.)

```

1 *REFE*EXER_G
2 *NAME* NS=MALE,FEMELLE
3 *VARIABLES* R,AGE,LONGUEUR,POIDS,COND$12
4 *FACTEUR* SEXE$NS,12 *FACTEUR*SS_ESP$3,12
5 *READ/PRIN=DM,FLFV=F* SEXE,SS_ESP,AGE,LONGUEUR,POIDS,COND
6 *RUN*
```

```

7 1 1 1 1 1 2 5
8 1 2 2 4 10 b
9 1 2 3 6 20 4
10 1 1 6 10 50 5
11 1 3 10 12 75 7
12 1 3 10 12 80 3
13 21 3 6 15 7
14 23 4 8 20 2
15 23 5 930 5
16 2 2 b 10 45 8
17 2 2 8 11 60 2
18 2 1 91165 3
19 *EOD*
```

AGE	MNHNMAX	5.5633	1.0000	10.0000	12	VALUES	0	HISSING
LONGUEUR	MNMINMAX	8.3333	1.0000	12.0000	12	VALUES	0	HISSING
POIDS	MNMINMAX	39.3333	2.0000	80.0000	12	VALUES	0	HISSING
CONO	MNMINMAX	4.7500	2.0000	8.0000	12	VALUES	0	HISSING

```

20 *REGRESSION/PRIN=C* AGE, LONGUEUR, POIDS, COND
21 *Y* POIDS
22 *FIT* AGE, LONGUEUR, COND
23 *TREAT* SEXE*SS_ESP
24 *ANOVA* POIDS; RES=R
25 *HIST* R
26 *PCP/CORR=Y, PRIN=LTS* AGE, LONGUEUR, POIDS, COND
27 *MODELE* CROISSANCE$LT=LINF*(1-EXP(-K*(AGE-T0)))
28 *SCALAIRE* LINF=12 :K=0.1 :T0=0
29 *OPTIMISER/NPAR=3 * MODELE=CROISSANCE; PARAM=LINF,K,T0; Y=LONGUEUR; Z=LT
30 *PRIN/P* LONGUEUR,LT$10,2
31 *TABLE/M* EFF,MOY$SEXE,SS_ESP
32 *TABU* LONGUEUR; ASSCT=EFF; MEANS=MOY
33 *PRIN* EFF$8.0
34 *PRIN* W1Y18.2
35 *RUN*
```

20.....

\*\*\*\* CORRELATION MATRIX \*\*\*\*

DF = 1 0

AGE	1	1.0000			
LONGUEUR	2	0.9363	1.0000		
POIDS	3	0.9904	3.9225	1.0000	
CONO	4	-0.2072	0.1841	-0.1817	1.0000
		1	2	3	4

22.....

\*\*\*\* REGRESSION ANALYSIS \*\*\*\*

\*\*\* REGRESSION COEFFICIENTS \*\*\*

Y-VARIATE: POIDS

	ESTIMATE	S.E.	T
CONSTANT	-9.04001	5.12922	-1.76
AGE	8.84870	1.17388	7.54
LONGUEUR	-0.31329	1.05688	-0.30
CONO	0.33238	0.65201	0.51

\*\*\* ANALYSIS OF VARIANCE \*\*\*

	DF	SS	MS
REGRESSN	3	7694.7	2564.92
RESIDUAL	8	143.9	17.99
TOTAL	11	7838.6	712.61

\* CHANGE -3 -7694.7 2564.92

PERCENTAGE VARIANCE ACCOUNTED FOR 97.5

3 8

\*\*\*\*\* ANALYSIS OF VARIANCE \*\*\*\*\*

VARIATE: PDIOS

SOURCE OF VARIATION	DF	SS	SS%	MS	VU
*UNITS* STRATUM					
SEX	1	0.3	0.00	0.3	0.001
SS-ESP	2	853.2	10.88	426.6	5.974
SEX*SS-ESP	2	4358.2	55.60	2179.1	4.977
RESIDUAL	6	2627.0	33.51	437.0	
TOTAL	11	7830.7	100.00	712.6	
GRAND TOTAL	11	7038.7	100.00		
GRAND MEAN		39.3			
TOTAL NUMBER OF OBSERVATIONS		12			

\*\*\*\*\* TABLES OF MEANS \*\*\*\*\*

VARIATE: PDIOS

GRAND MEAN	39.3		
SEX	MALE	FEMELLE	
	39.5	39.2	
SS-ESP	1	1	3
	33.0	33.7	51.2
SS-ESP	1	2	3
SEX			
MALE	26.0	15.0	11.5
FEMELLE	40.0	52.5	25.0

\*\*\*\*\* STANDARD ERRORS OF DIFFERENCES OF MEANS \*\*\*\*\*

TABLE	SEX	SS-ESP	SEX
			SS-ESP
REP	6	4	2
SED	12.08	14.80	20.92

\*\*\*\*\* STRATUM STANDARD ERRORS AND COEFFICIENTS OF VARIATION \*\*\*\*\*

STRATUM	DF	SE	CV%
*UNITS*	6	23.92	53.2

HISTOGRAM OF R

0 - 25 6 \*\*\*\*\*  
25 \* 6 \*\*\*\*\*  
0

MISSING VALUES 0

SCALE: 1.33 ASTERISK(S) REPRESENT 1 UNIT

26.....

\*\*\*\*\* PRINCIPAL COMPONENTS ANALYSIS \*\*\*\*\*

\*\*\* LATENT ROOTS \*\*\*

	1	2	3	4
	2.9558	0.9444	0.0914	0.0085

\*PERCENTAGE VARIANCE\*

	1	2	3	4
	73.0943	23.6089	2.2844	0.2123

\*\*\* LATENT VECTORS (LOADINGS) \*\*\*

	1	2	3	4
AGE	-0.5755	-0.0820	-0.3314	0.7431
LONGUEUR	-0.5608	-0.0996	0.8180	-0.0805
POIDS	-0.5714	-0.1075	-0.4701	-0.6640
COND	0.1668	-0.9858	-0.0038	0.0187

\*\*\* TRACE = 4.0000

\*\*\* SIGNIFICANCE TESTS FOR EQUALITY OF REMAINING ROOTS \*\*\*

NUMBERS OF UNITS AND VARIATES DIFFER BY LESS THAN 50 SO CHI-SQUARED APPROXIMATIONS ARE POOR

NO. OF ROOTS EXCLUDED	CHI SQ	DF
0	67.4836	9
1	44.5788	5
2	12.8403	2

29.....

\*\*\* PARAMETER ESTIMATES \*\*\*

	ESTIMATE	S.E.	CORRELATIONS		
LINE	12.62630	0.24142	1.0000		
K	0.29161	0.01741	-0.8964	1.0000	
TO	0.71864	0.06813	-0.4738	0.7080	1.0000
DF		SS	MS		
RESIDUAL	Y	0.458269	0.050919		

LONGUEUR	LT
1.00	0.99
4.00	3.94
6.00	6.13
10.00	9.92
12.00	11.78
12.00	11.78
6.00	6.13
8.00	7.78
9.00	9.00
10.30	9.92
11.00	11.12
11.00	11.50

SS_ESP	EFF			MARGIN
	1	2	3	
SEXE				
HALE	2	2	2	6
FEMELLE	2	2	2	6
MARGIN	4	4	4	12

SS_ESP	MOY			MARGIN
	1	2	3	
SEXE				
HALE	5.50	5.00	12.00	7.50
FEMELLE	8.50	10.50	8.50	9.17
MARGIN	7.00	7.75	10.25	8.33



## ANNEXE 2

GENSTAT est un programme écrit en FORTAN et disponible sur La plupart des gros ordinateurs. Une version a été écrite pour microordinateurs équipés du système MS DOS et disposant d'au moins 256 k de mémoire centrale. Les conditions de commercialisation de cette version ne sont pas encore connues.

Il existe un manuel d'utilisation livré avec le progiciel (en langue anglaise).

Un manuel en français a été écrit par des membres de l'INRA et de l'Université Paris Sud. Il est conçu également comme un guide sur les méthodes statistiques et comprend de très nombreux exemples traités.

- Pour obtenir le document de l'INRA s'adresser à :

SERVICE DES PUBLICATIONS DE L'INRA  
C.N.R.A Route de Saint Cyr  
78000 -- Versailles (Tél. 950 7522)

- Pour des renseignements concernant la location etc . . . :

The programs secretary  
Statistics department  
Rothamsted experimental station  
Harpenden  
Hertfordshire  
AL 5 2 J Q  
GRANDE BRETAGNE