

Étude régionale de l'agriculture française

0 - FICHE TECHNIQUE

A - Tableau de données résultant de la concaténation, selon un seul côté, de tableaux de contingence. Une telle table de données se trouve indexée par un ensemble de la forme $I \times (J_1 \cup J_2 \cup \dots \cup J_\ell)$ où les différents tableaux $I \times J_1, I \times J_2, \dots, I \times J_\ell$ sont de contingence. I est ici défini par l'ensemble des 89 départements. $\ell=3$; J_1 est l'ensemble des 8 modalités de la variable "structure d'exploitation", J_2 est l'ensemble des 6 modalités de la variable "principales cultures" et J_3 l'ensemble des 3 modalités de la variable "importation du cheptel".

B - Classification hiérarchique de l'ensemble I et de l'ensemble $J=J_1 \cup J_2 \cup J_3$ où I est décrit à travers J et J à travers I .

C - Dans le cas d'un seul tableau de contingence $I \times H$, l'indice de proximité entre deux éléments j et h de H est défini par la formule (7) du paragraphe V du chapitre 2 dont nous rappelons ici l'expression

$$\rho(j, h) = \frac{\sum_{i \in I} \frac{1}{p_{i.}} (f_{ij} - p_{i.} p_{.j}) (f_{ih} - p_{i.} p_{.h})}{\{p_{.j} p_{.h} [\sum_{i \in I} \frac{f_{ij}^2}{p_{i.} p_{.j}} - p_{.j}] [\sum_{i \in I} \frac{f_{ih}^2}{p_{i.} p_{.h}} - p_{.h}]\}^{1/2}}$$

où $\{f_{ij}/(i,j) \in I \times H\}$ désigne la distribution des fréquences sur $I \times H$ et où $\{p_{i.}/i \in I\}$ et $\{p_{.j}/j \in H\}$ désignent, respectivement, les distributions marginales sur I et sur H .

A partir de l'expression de $\rho(j, h)$, on en déduit celle de l'indice $\alpha(i, i')$ entre les deux sommets i et i' de I .

Dans le cas d'un tableau résultant de la juxtaposition, ici horizontale, d'une suite de tables de contingence ; un tableau indexé donc par $I \times (J_1 \cup J_2 \cup \dots \cup J_\ell)$; désignons par $\sigma_k(i, i')$, $1 \leq k \leq \ell$, l'indice de proximité

entre i et i' calculé au niveau de la table de contingence $I \times J_k$. L'indice global qui est ici adopté au niveau de l'ensemble du tableau $I \times (J_1 \cup J_2 \cup \dots \cup J_\ell)$, est la moyenne des indices locaux :

$$\sigma(i, i') = \frac{1}{\ell} \sum_{1 \leq k \leq \ell} \sigma_k(i, i') \quad ;$$

pour une approche plus fine voir [6].

Pour la classification de l'ensemble des modalités de $J = J_1 \cup J_2 \cup \dots \cup J_\ell$; on tient compte de la représentation du nuage $\mathcal{N}(I)$ dans chacun des espaces $\mathbb{R}_{J_1}, \mathbb{R}_{J_2}, \dots, \mathbb{R}_{J_\ell}$ (cf. chap. 6 § II.4 et chap. 5 § IV.2) et on aboutit à l'expression précédente $\rho(j, h)$ de la proximité mais où $p_{i.}$ est remplacé par

$$\hat{p}_{i.} = \frac{1}{\ell} \sum_{1 \leq k \leq \ell} p_{i.}^k,$$

où $p_{i.}^k$ est la fréquence marginale de i , au niveau de la table de contingence $I \times J_k$; pour tout i de I .

La réduction globale des similarités a été effectuée en centrant et en réduisant par rapport à la distribution observée de l'indice sur l'ensemble des paires de l'ensemble à organiser.

D - Algorithme de la Vraisemblance des Liens (A.V.L.). Chaîne de programmes ORDON-POLON-ARBRE qui suppose un programme spécifique de calcul des proximités entre lignes ou colonnes du tableau $I \times (J_1 \cup J_2 \cup \dots \cup J_\ell)$ (ici $\ell=3$), conformément aux formules exprimées ci-dessus. Programme de construction de Classifications Ascendantes Hiérarchiques de M. Jambu ; critère de la part d'inertie expliquée associée à la métrique du χ^2 (cf. chap. 5, § IV. 2).

E - Tester dans le cadre de A.V.L. les nouveaux indices établis pour la comparaison des lignes (resp. colonnes) d'une juxtaposition de tables de contingence. Comparaison des résultats de cette classification hiérarchique avec celle fournie par l'algorithme que nous noterons C.A.H. (M_2, χ^2) où on optimise la part d'inertie expliquée, conformément à la métrique du χ^2 .

F - B. TALLUR.

G - P. RAINELLI.

La suite des titres des sous paragraphes est

- 1 - Introduction.
- 2 - Les données.
- 3 - La structure du tableau des données.
- 4 - Résultats.
 - 4.1 - Classification de l'ensemble des modalités.
 - 4.2 - Classification des départements selon la structure d'exploitation
 - 4.3 - Classification des départements selon les principales cultures.
 - 4.4 - Classification des départements selon les principales cultures et la structure d'exploitation.

- 4.5 - Classification des départements selon la suite des trois caractères : structure, cultures et cheptel.
- 5 - Comparaison des résultats issus de l'algorithme A.V.L. (Cor. χ^2) avec ceux de C.A.H. (M_2, χ^2).
- 6 - Autres exemples
- 7 - Conclusion.

1. INTRODUCTION

L'économiste étudiant l'agriculture française constate une grande diversité de situations. Chaque région est caractérisée par certaines productions. Il apparaît aussi une spécificité régionale des structures agraires. L'orientation et la taille des exploitations reflètent une spécialisation plus ou moins poussée. Les divers caractères ne sont pas immuables dans le temps. Il y a une évolution qui ne se fait pas de partout selon le même modèle.

Compte tenu des phénomènes observés et de la nature des données, les méthodes classiques basées sur les modèles linéaires ne fournissent pas un schéma explicatif satisfaisant des situations et de leur évolution.

D'où l'intérêt de proposer une autre démarche.

Dans un premier temps, nous avons cherché à dégager les principaux types d'agriculture à l'aide d'une classification hiérarchique. Celle-ci porte sur l'ensemble des départements français décrits selon le mode d'utilisation du sol, l'importance du cheptel et selon la structure des exploitations agricoles. Ceci permet d'obtenir une "photographie" de la situation économique à un moment donné. L'évolution de la structure économique pourra, ensuite, être étudiée par l'examen d'une série de clichés pris à des intervalles de temps réguliers. L'objet de la présente étude sera de mettre au point les moyens nécessaires à la réalisation des clichés et à l'examen comparatif de ceux-ci.

En premier temps, nous avons mis au point un indice de proximité permettant de classer les départements français, ainsi que des variables descriptives en utilisant la chaîne de programmes de classification hiérarchique mise au point par I.C. Lerman et ses collaborateurs. Nous avons ensuite analysé les données de l'année 1972, ce qui a permis de définir un certain nombre de "types" d'économie agricole en France, et les résultats sont comparés à ceux obtenus par la méthode de la Classification Ascendante Hiérarchique [C.A.H.], programmés par M. Jambu, [3], utilisant la distance du χ^2 et l'agrégation par la maximisation à chaque pas du moment centré d'ordre deux de la nouvelle classe formée par réunion de deux classes. Nous appellerons ce dernier, l'algorithme de C.A.H. (M_2, χ^2).

2. LES DONNEES

Les données ont été fournies par P. Rainelli (Station d'Economie Rurale INRA, Rennes) qui a participé à l'interprétation des résultats. Ces données relatives aux 89 départements français⁽¹⁾ sont de trois types

(1) Afin de poursuivre l'évolution dans le temps des variables retenues, l'ancien découpage de la région parisienne a été conservé avec la Seine et Oise. Par contre le département de la Seine a été exclu en raison de la très faible importance de son activité agricole.

a) Le mode d'Utilisation du Sol

Afin de décrire l'orientation des productions, on a retenu pour chaque département la répartition de la surface agricole utile selon les six principales modalités d'occupation du sol.

- Les surfaces toujours en herbe (STH). Avec les prairies et herbages, ce système à orientation animale est assez extensif.
- Les cultures fourragères (FOUR). Ce type de culture que l'on trouve aussi dans les orientations animales est plus intensif.
- Les céréales (CER) reflètent une utilisation assez spécifique du sol.
- Les cultures permanentes (C.P.) qui incluent les plantations fruitières et les vignes, ont une dominante intensive.
- Les cultures maraîchères et potagères (CMP) correspondent à une utilisation très poussée de la terre.
- Les plantes sarclées (SARC) avec essentiellement les pommes de terre, les betteraves et les choux fourragers, reflètent un système particulier.

b) L'importance du Cheptel

L'occupation du sol ne donnant qu'un aperçu de l'orientation générale, il a été nécessaire de compléter ces données par des renseignements sur l'importance du cheptel présent dans chaque département. On a retenu le nombre de têtes de bovins, le nombre de têtes de porcins et le nombre de têtes d'ovins.

c) La structure des Exploitations Agricoles

Huit classes de taille on été retenues :

- moins de 1 ha
- 1 à 2 ha
- 2 à 5 ha
- 5 à 10 ha
- 10 à 20 ha
- 20 à 50 ha
- 50 à 100 ha
- plus de 100 ha

Pour chaque département, on a la répartition en valeur absolue du nombre d'exploitations par classe de surface.

Nous avons choisi un tableau de données à 89 lignes (départements) et 17 colonnes (modalités ou items) relatif à l'année 1972.

3. LA STRUCTURE DU TABLEAU DES DONNEES

L'ensemble I à classifier est l'ensemble des 89 départements et l'ensemble J des variables descriptives est la réunion de trois caractères :

J_1 (structure d'exploitation à 8 modalités) ;

J_2 (principales cultures à 6 modalités).

J_3 (l'importance du cheptel à 3 modalités).

Le tableau des données est le tableau qui croise l'ensemble I indexant les lignes et l'ensemble $J=J_1 \cup J_2 \cup J_3$, indexant les colonnes. Etant donné que les caractères J_1, J_2 et J_3 ne sont pas de même nature, toutes les données du tableau ne se réfèrent pas à une même unité. En effet, l'élé-

ment $(i, j_1), i \in I, j_1 \in J_1$ est exprimé en nombre d'exploitations du département i et appartenant à la modalité j_1 de la "structure", alors que l'élément $(i, j_2), j_2 \in J_2$ est la surface en hectares du département i où se pratique la culture j_2 et enfin l'élément $(i, j_3), j_3 \in J_3$ indique le nombre d'animaux de la catégorie j_3 dans le département i .

Il s'agit donc de trois tableaux de contingence $I \times J_1, I \times J_2$ et $I \times J_3$ qui se trouvent concaténés. On considérera, pour commencer, le problème de classification des lignes (respectivement, colonnes) d'un tableau de contingence croisant l'ensemble I et l'ensemble J où ce dernier est composé des modalités d'un seul caractère, et ensuite on généralisera la méthode d'une façon naturelle au cas où J est formé de la réunion des modalités des trois caractères J_1, J_2 et J_3 .

Les indices de proximité utilisés sont ceux exprimés dans la partie C de la fiche technique (§ 0').

PRODUCTIONS ANIMALES	01 TOUJOURS EN HERBE	>-----*	
	08 BVINS	> 10----	*
	06 CULT.FOURRAGERES	>-----*	I
	09 PORCS	> 11----	I
	07 BOVINS	>-----*	I
	15 20 A 50 HECTARES	>-----*	I
SYST. EXTENSIF	03 CULT.MARAICHERES	>-----*	I
	05 PLANTES SARCLEES	> 8---- ----	I
	04 CEREALES	>-----*	I
	17 > 8U = A 100 H	> 1	I
	16 50 A 100 HECTARES	> *3	I
	02 CULT.PERMANENTES	>-----*	I
SYST. INTENSIF	10 INF.A 1 HECTARE	> 5	I I
	11 1 A 2 HECTARES	> 2	I
	12 2 A 5 HECTARES	>-----*	I
	13 5 A 10 HECTARES	> 4----	I
	14 10 A 20 HECTARES	>-----*	I

CLASSIFICATION DES VARIABLES 1972 (BOIS ET FORETS EXCLUS)

ALGORITHME: AVL (COR, χ^2)

Figure 1.

4. RESULTATS

En utilisant la chaîne de programmes de classification hiérarchique de I.C. Lerman, avec les indices définis au paragraphe 3 et par l'A.V.L. (Algorithme de Vraisemblance du Lien), on dira plus brièvement A.V.L.

(COR, χ^2), on a procédé aux analyses suivantes :

- 1- Analyse de l'ensemble de 17 modalités.
- 2- Classification des départements selon la structure d'exploitation.
- 3- Classification des départements selon les principales cultures.
- 4- Classification des départements selon les principales cultures et la structure d'exploitation.
- 5- Classification des départements selon tous les trois caractères réunis.

Enfin, on a classifié l'ensemble des départements par l'algorithme de C.A.H. (M_2, χ^2) [classification Ascendante Hiérarchique] programmé par M. Jambu, utilisant la distance du χ^2 et l'agrégation par la maximisation du moment centré d'ordre deux et les classes fournies par cette méthode sont comparées à celle obtenues plus haut.

4.1. Classification de l'ensemble de modalités

L'arbre de classification (figure 1) de l'ensemble de 17 modalités comporte au total 16 niveaux avec 3 noeuds significatifs à savoir 3, 6 et 12.

Classe A : Au niveau trois de l'arbre une classe est formée des modalités {"céréales", " \geq à 100 hectares", et "50 à 100 hectares"}. C'est le noyau d'une classe, à laquelle s'adjoint au niveau 12 une autre classe caractérisée par les variables "plantes sarclées", et "cultures maraîchères et potagères". Cette association, bien que choquante à première vue, s'explique parfaitement car, les plantes sarclées (notamment les pommes de terre, betteraves industrielles) sont cultivées sur des grandes surfaces et on trouve souvent des cultures potagères (par exemple petits-pois, haricots) sur de grandes surfaces.

Classe B : Parallèlement, on peut remarquer la formation d'une classe d'exploitations de type intensif. En effet, au niveau 6, la variable "cultures permanentes" se réunit au groupe déjà formé des modalités : {"inférieur à 1 hectare", "1 à 2 hectares", et "2 à 5 hectares"}. Le niveau 6 étant significatif, ces quatre modalités définissent très bien un système intensif basé sur de petites exploitations avec dominante viticulture ou arboriculture.

Classe C : Enfin apparaît une troisième classe comportant les modalités : "surface toujours en herbe", "ovins", "cultures fourragères", "porcs", "bovins" et "20 à 50 hectares" . On peut reconnaître dans cette classe un système orienté vers les productions animales ; les variables "surface toujours en herbe" et les "cultures fourragères" y figurent ; les exploitations sont de taille moyenne : 20 à 50 ha .

L'analyse générale met donc en évidence trois grands systèmes de production. Le premier est centré sur les grandes et très grandes exploitations associées aux céréales. Il correspond à l'agriculture la plus favorisée. Le second est basé sur de petites et très petites exploitations où

l'on pratique les cultures permanentes. C'est le système intensif par excellence. Enfin, dans le dernier se trouvent les productions animales associées aux cultures fourragères, aux surfaces toujours en herbe sur des exploitations moyennes.

4.2. Classification des départements selon la structure d'exploitation

L'arbre de classification (figure 2) donné par l'A.V.L. (COR, χ^2) est condensé aux niveaux 41, 43, 48, 50, 54, 56, 58, 62, 64, 68, 72, 75, 77, 80, 82 et 86 relatifs aux maximums locaux de la statistique Θ . La statistique globale atteint son maximum absolu au niveau 86, qui donne une partition de l'ensemble des départements en trois grandes classes.

a) Exploitations moyennes

Cette classe est constituée des départements où la grande majorité des exploitations agricoles sont de taille moyenne allant de 5 à 50 hectares. Cette classe comporte quatre sous-classes à l'intérieur desquelles on trouve regroupés les départements où les exploitations sont : plutôt grandes dans la première, plutôt petites dans la seconde, "très moyennes" dans la troisième et enfin moyennes et grandes dans la dernière.

1ère sous-classe : Cette classe comporte des départements tels que Nord, Pas-de-Calais, Mayenne, Tarn, Creuse, Haute-Garonne... etc. (voir le graphique). La taille d'exploitation dominante est de "20 à 50 hectares" suivie de "10 à 20 hectares".

2ème sous-classe : Elle regroupe les départements où la majorité des exploitations sont de "10 à 20 hectares" suivie de "5 à 10 hectares" ; par exemple Corrèze, Dordogne, Ariège, etc...

3ème sous-classe : Les départements tels que Manche, Isère, Landes, Hautes Pyrénées, Côtes du Nord, Rhône, Haute-Savoie... comportant le plus grand nombre d'exploitations de "10 à 20 hectares" suivies de "20 à 50 hectares", sont réunis.

4ème sous-classe : Dans cette dernière partie, petite classe des exploitations de "20 à 50 hectares" sont nettement dominantes suivies, de loin, par d'autres.

b) Grandes exploitations

C'est la classe des départements tels que {Aisne, Seine-et-Marne, Oise, Cher, Eure-et-Loire...}, où la grande majorité des exploitations agricoles ont "50 à 100 hectares", et celles ayant "plus de 100 hectares" ne sont pas minoritaires. A l'intérieur de cette classe, on peut distinguer trois petites sous-classes, qui se différencient les unes des autres par le pourcentage d'exploitations de très grande taille (≥ 100 hectares) ; ce pourcentage varie entre 6 à 12 dans la première, entre 2 à 5 dans la seconde et inférieur à 2 dans la dernière.

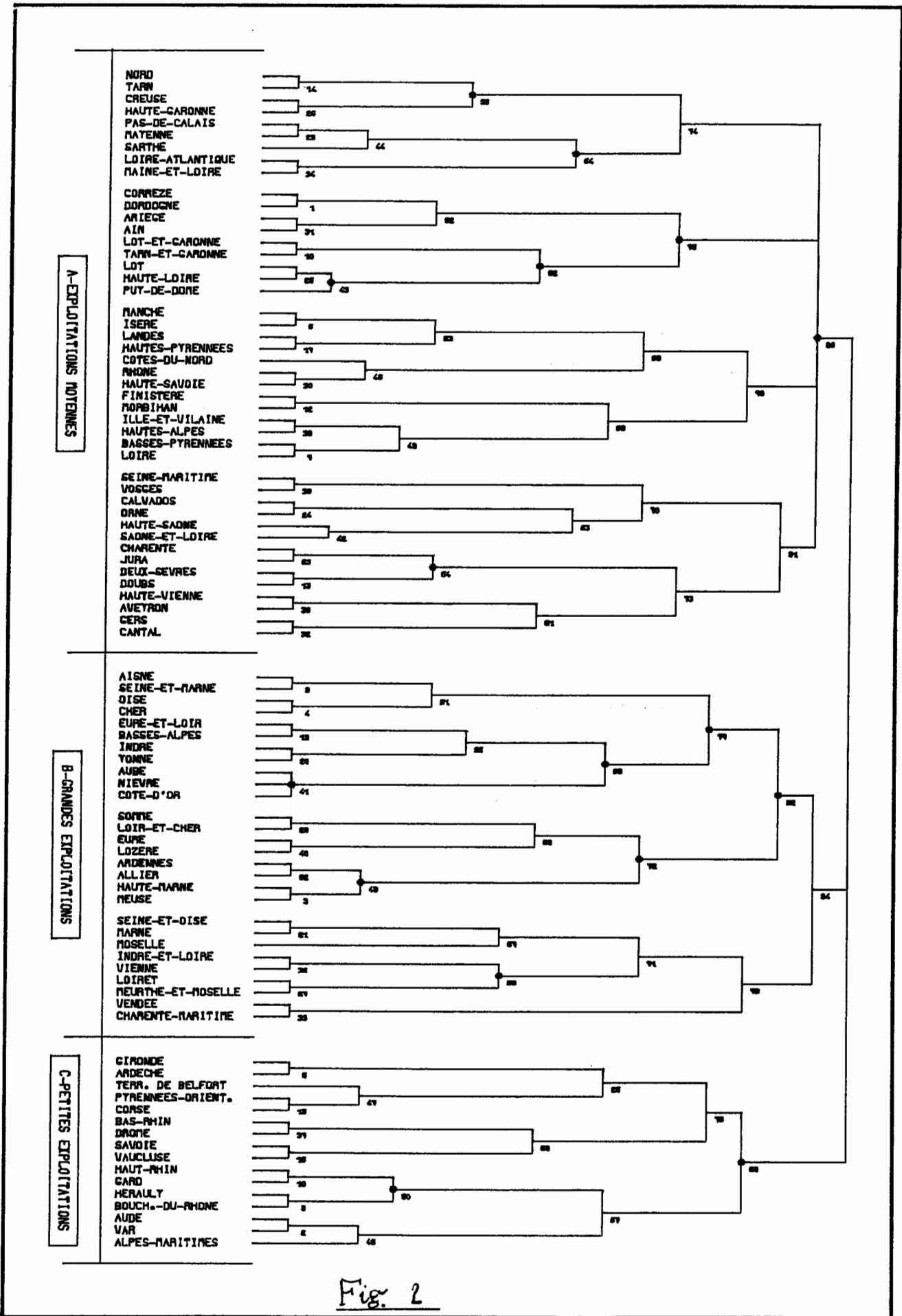


Fig. 2

c) Petites exploitations

Enfin, une classe très "compacte" des départements de très petites exploitations se dessine dès le niveau 2 pour s'achever au niveau 80 par un noeud significatif. La plupart des exploitations agricoles de ces départements ont moins de 2 hectares avec un pourcentage assez élevé de "2 à 5 hectares". Cette classe comprend les départements suivants :

{Gironde, Ardèche, Pyrénées Orientales, Savoie, Vaucluse, Gard, Hérault, Aude, Var, Corse, Bas Rhin, Haut Rhin, Bouches-du-Rhône, Alpes Maritimes, Drôme et Territoire de Belfort}.

Conclusion : On retrouve par cette analyse une classification très naturelle qui satisfait pleinement aussi bien l'économiste que le statisticien.

4.3. Classification des départements selon les principales cultures

L'arbre de classification (figure 3) issu de l'A.V.L. (COR, χ^2), comporte 88 niveaux avec 15 noeuds significatifs. La partition en quatre grandes classes est obtenue au niveau 85.

a) Céréales - Fourrages

Cette grande classe est composée de 22 départements qui se caractérisent, tous, par la culture céréalière (50% de surface, en moyenne), S.T.H. (de 10 à 30%), et cultures fourragères (12% en moyenne). On peut néanmoins y discerner trois sous-groupes :

Le premier groupe : possède, en plus des céréales et des fourrages, la culture des plantes sarclées (15% en moyenne). On peut y citer les départements : {Nord, Pas-de-Calais, Bas-Rhin, Haut-Rhin, Aisne, Oise et Somme}.

Le deuxième groupe : {Seine-et-Marne, Loiret, Eure-et-Loire, Aube, Seine-et-Oise et Marne} est essentiellement la zone céréalière avec plus de 70% de surface en céréales.

Le troisième groupe : {Cher, Côte d'Or, Eure, Indre, Yonne, ...} représente le groupe moyen avec 50% de céréales, 20% de S.T.H., et 12% de fourrages.

b) Cultures permanentes

Les 17 départements de la classe B se divisent en deux sous classes d'effectifs 8 et 9, respectivement. La caractéristique commune de tous les départements est la présence non négligeable des cultures permanentes.

1ère sous-classe : {Charente, Charente-Maritime, Tarn-et-Garonne, Drôme, Lot-et-Garonne, Gers, Dordogne et Tarn}. Ces départements ont 10 à 16% de surface en cultures permanentes, à l'exception de Tarn, bien que la culture dominante soit celle des céréales.

2ème sous-classe : {Gironde, Pyrénées Orientales, Rhône, Var, Vaucluse, Aude, Bouches du Rhône, Gard et Hérault}. Ici les cultures permanentes sont largement dominantes.

c) S.T.H. - Céréales - Fourrages

Celle-ci a un effectif de 20 départements, qui se divisent en deux groupes.

1er groupe : {Seine-Maritime, Ardennes, Sarthe, Hautes Pyrénées, Haute Saône, Nièvre, Allier, Mayenne, Haute-Vienne, Aveyron et Creuse}. Ces 17 départements ont, chacun, plus de 55% de surface toujours en herbe ; viennent ensuite les cultures céréalières et fourragères.

2ème groupe : {Côtes du Nord, Finistère, Ille-et-Vilaine, Morbihan, Vienne, Deux Sèvres, Vendée et Loire Atlantique et Maine-et-Loire}. Ces 9 départements sont dominés par les cultures fourragères ; les surfaces en céréales et S.T.H. suivent.

d) S.T.H.

Cette classe est la plus grande, ayant à son effectif 30 départements. Il s'agit des départements où la S.T.H. est de loin la plus importante. Certains d'entre eux consacrent presque totalité de la surface à l'herbe : par exemple : Savoie (89%), Alpes Maritimes (92%) et Cantal (80%). Cependant, un petit groupe : {Basses Pyrénées, Meuse, Haute-Marne, Ain, Meurthe-et-Moselle, Territoire de Belfort et Isère} se détache du reste.

Dans cette zone, on consacre, en moyenne, presque 30% de surface à la culture céréalière et 10% aux cultures fourragères après avoir laissé 50 à 60% en herbe.

Conclusion : Cette analyse a permis de retrouver, à partir des variables caractérisant l'utilisation du sol, les principaux systèmes agricoles en France.

4.4. Classification des départements selon les principales cultures et la structure d'exploitation

L'arbre de classification (figure 4) issu de l'algorithme A.V.L. (COR, χ^2) est condensé à ses noeuds les plus significatifs. Au niveau 84, on a une partition en cinq classes :

a) Très grandes exploitations associées aux céréales et fourrage

Cette première classe est, à quelques exceptions près, la première classe de l'analyse 4.3. Elle est composée des départements : {Nord, Pas de Calais, Aisne, Oise, Somme, Seine-et-Marne, Cher, Côte d'Or, ...}. Ce sont de grandes exploitations avec les cultures céréalières et fourragères. On y retrouve le système "extensif".

b) Moyennes exploitations associées aux spéculations fourragères.

{Côtes du Nord, Morbihan, Finistère, Ille-et-Vilaine, ... Charente, Tarn, ..., Tarn et Garonne}. Ce sont des exploitations moyennes avec les cultures fourragères, S.T.H. et un peu de céréales. Le petit groupe des 6 premiers départements se sépare du reste ; cette zone comporte plus de surface en cultures fourragères et moins en céréales que les autres départements de la classe.

c) Moyennes et grandes exploitations associées à des S.T.H.

C'est la classe associant les surfaces toujours en herbe et d'assez grandes exploitations. On peut remarquer une sous-classe nettement dominée par la S.T.H. et qui comporte deux petits groupes d'exploitations de taille moyenne et grande respectivement. 1er groupe : {Seine-Maritime,, Basses Alpes} et le second : {Haute Marne, Meuse, Nièvre, Allier ...} Une 2ème sous-classe de départements avec un pourcentage important de

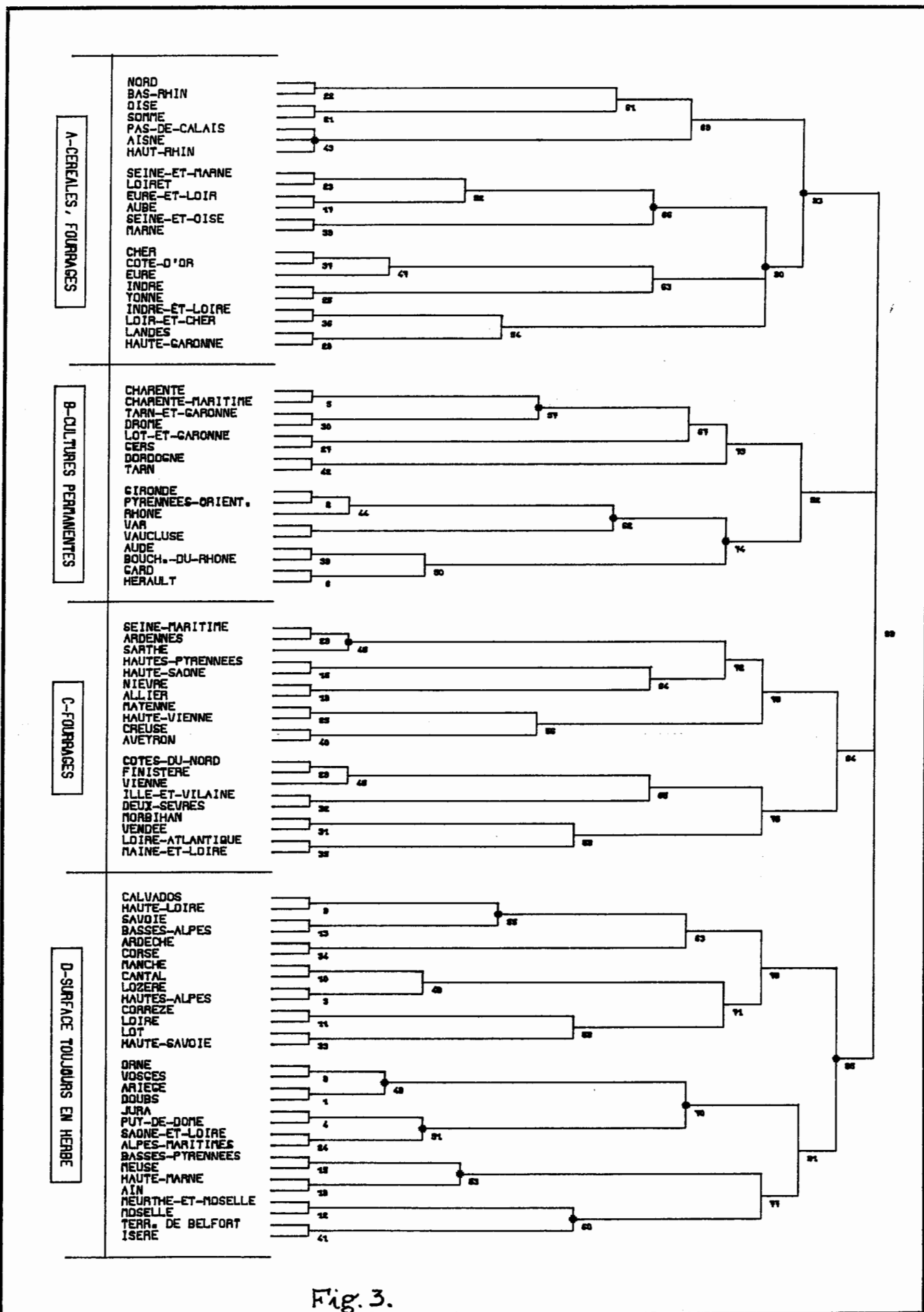


Fig. 3.

S.T.H., et qui possède aussi des surfaces en céréales et fourrages est divisée en petits groupes en fonction de la taille des exploitations (taille moyenne ou plutôt petite).

d) Petites exploitations dominées par des S.T.H.

{Manche, Haute-Savoie, Loire, ... Ain, Hautes Pyrénées, Isère}. C'est la classe des petites exploitations où des S.T.H. est dominante et est suivie des cultures céréalières et fourragères.

e) Très petites exploitations associées à des cultures permanentes

C'est, enfin, la fameuse classe d'exploitations de très petite taille et des cultures permanentes. On peut percevoir deux grands groupes de départements qui diffèrent l'un par rapport à l'autre par leur degré d'intensité en cultures permanentes : le premier : {Gironde, Gard, ..., Bouches du Rhône, ..., Vaucluse} ayant le pourcentage de C.P. beaucoup plus fort que le second : {T. Belfort, Savoie, Ardèche, Corse et Alpes Maritimes}.

Conclusion : Cette analyse est tout à fait conforme aux deux précédentes, et permet de définir une typologie, plus fine des systèmes de production, en ce sens que les départements sont classés dans des catégories distinctes non seulement selon la taille de leurs exploitations agricoles, mais aussi selon le type de cultures qui y sont pratiquées.

4.5. Classification des départements selon les trois caractères : Structure + Cultures + Cheptels.

En utilisant, toujours, l'algorithme A.V.L. (COR, χ^2) avec l'indice équipondéré de proximité, on obtient une classification en quatre grandes classes au niveau 85 de l'arbre (figure 5). Celle-ci est tout à fait en accord avec les trois analyses précédentes. Mais en plus elle tient compte du caractère "importance du cheptel".

a) Système extensif

Il s'agit de grandes et très grandes exploitations, avec des cultures céréalières et fourragères et des bovins.

b) Productions animales et spéculations fourragères

Moyennes exploitations, avec des surfaces toujours en herbe, fourrages, et plantes sarclées, où l'élevage de bovins est assez important. A noter que les quatre départements bretons qui figurent dans cette classe sont réunis par un noeud significatif. Ils forment un petit groupe caractérisé par l'élevage de porcs et la culture de plantes sarclées. Ces deux variables marquent le caractère intensif de cette région. A noter aussi le transfert des départements Rhône et Gironde de la classe des petites exploitations associées aux cultures permanentes à cette classe. Ceci est parfaitement justifié car, relativement aux caractères "Cheptel" et "Structure d'exploitation" ces deux départements ressemblent plus aux Lot-et-Garonne et Gers qu'à Var ou Vaucluse.

c) Productions Animales (bovins) associées à des S.T.H.

C'est la grande classe de S.T.H., céréales et fourrages associées à l'élevage de bovins ; elle se divise en trois groupes qui correspondent respectivement aux petites, grandes et très grandes exploitations.

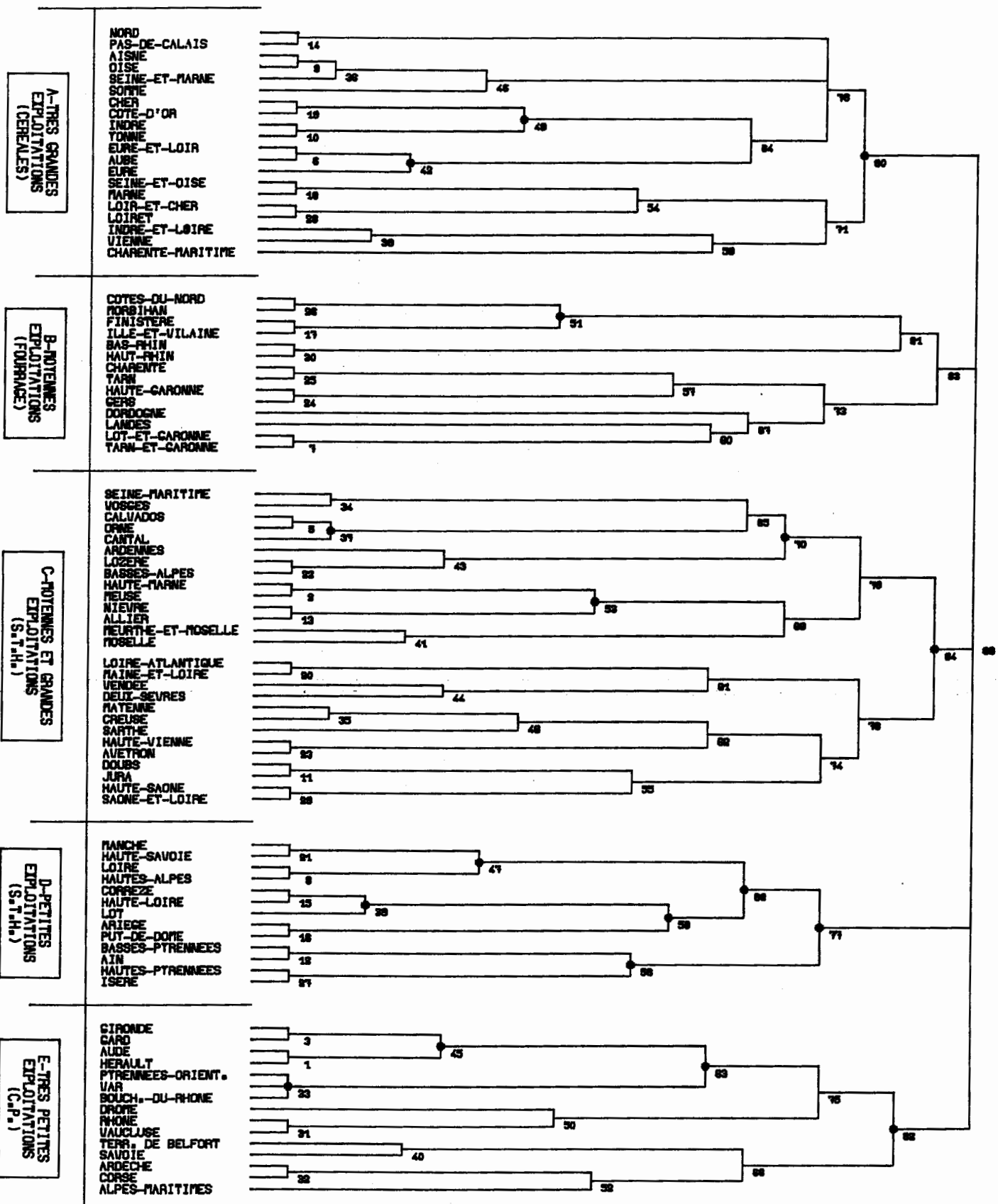


Fig. 4

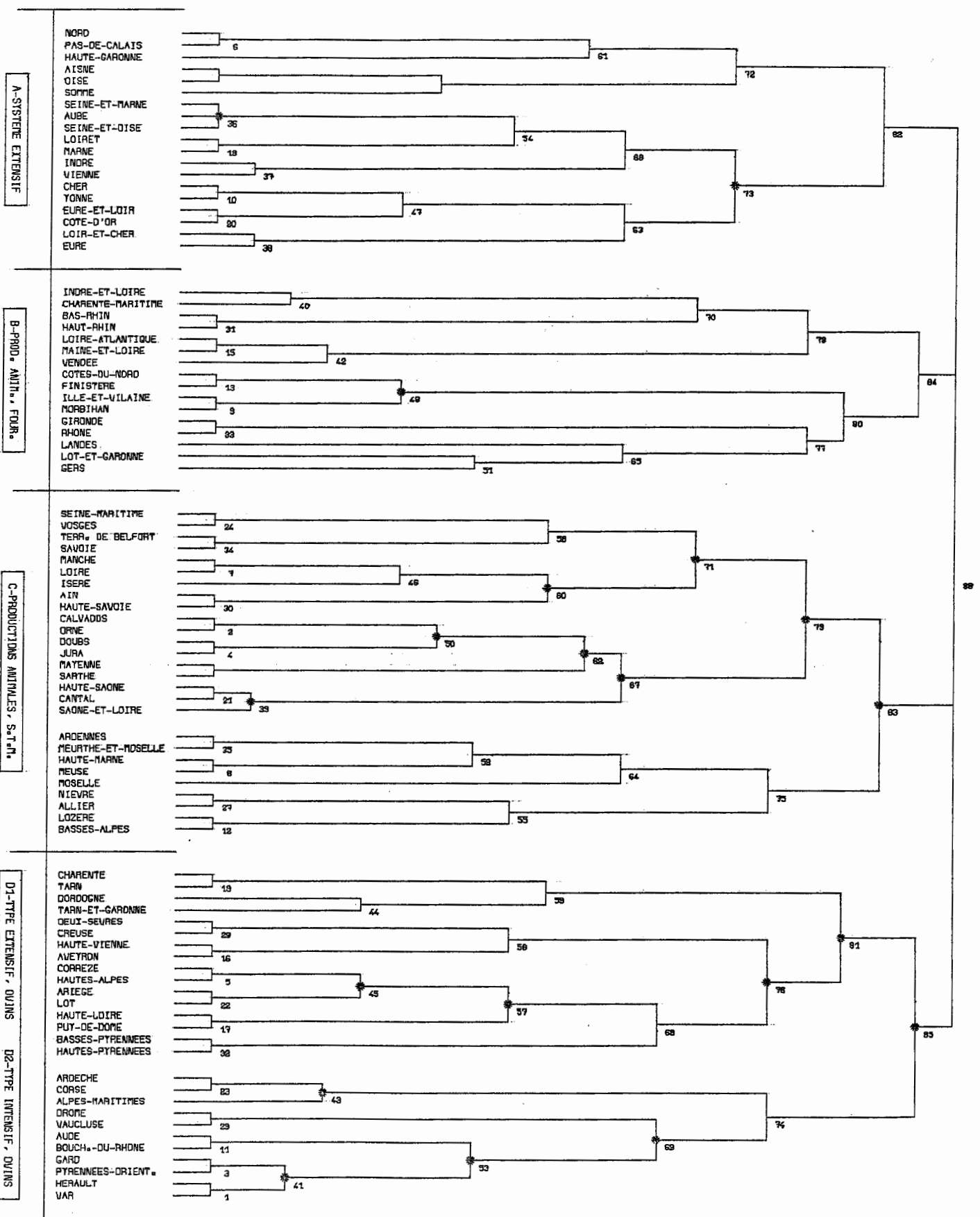
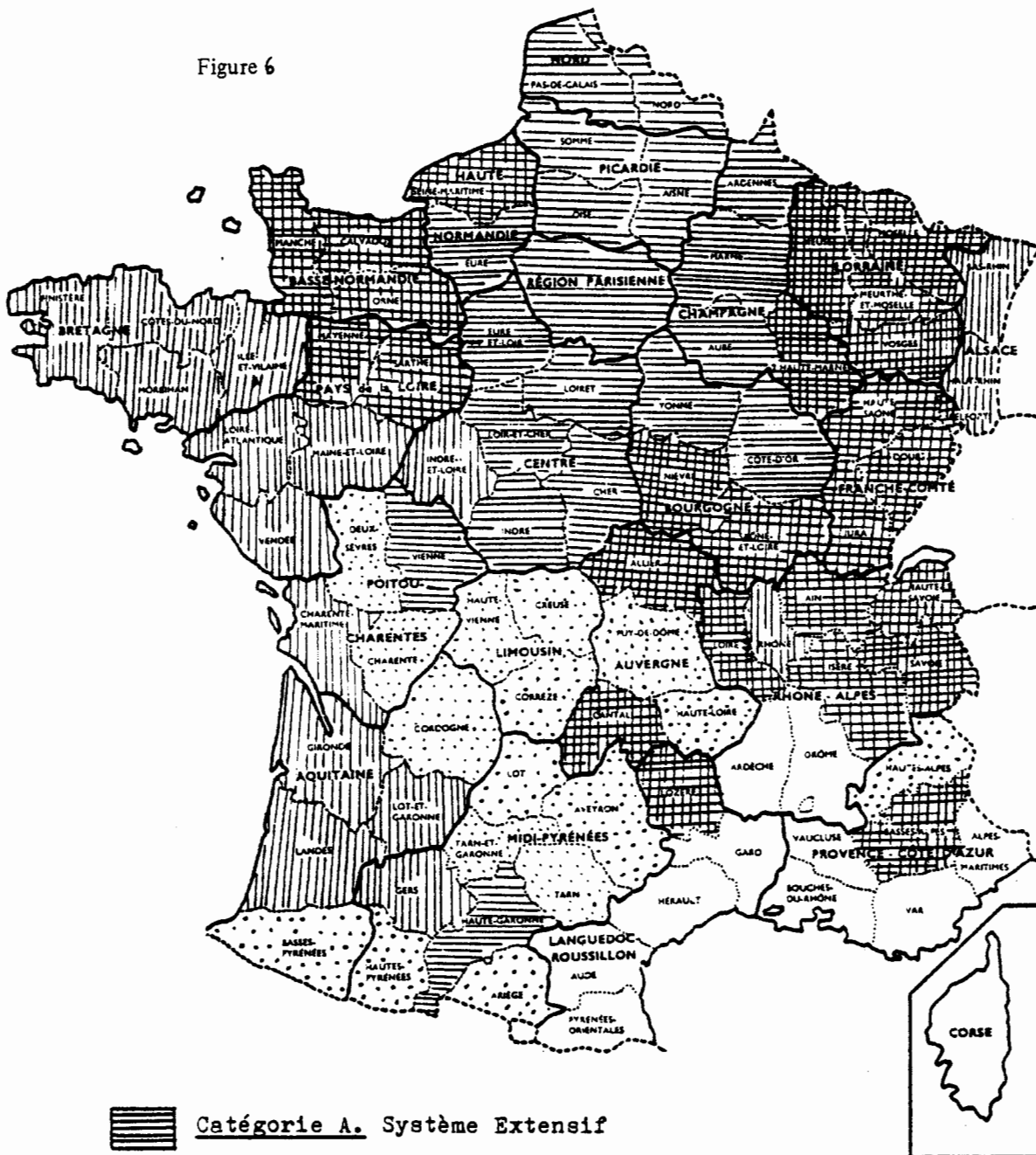







Fig. 5.

Figure 6



-  **Catégorie A.** Système Extensif
-  **Catégorie B.** Productions Animales, Spéculations Fourragères
-  **Catégorie C.** Productions Animales associées à des S.T.H. et aux Céréales
-  **Catégorie D1.** Type Extensif avec des ovins; Petites ou moyennes Exploitations
-  **Catégorie D2.** Type intensif (Cultures Permanentes) associé avec les ovins

d) On peut remarquer deux sous-classes

1°) Type extensif avec des ovins

{Charente, Tarn, Dordogne, ..., Basses Pyrénées, et Hautes Pyrénées}. Il s'agit des départements caractérisés par S.T.H., taille moyenne ou petite et le nombre important des têtes d'ovins.

2°) Type intensif avec des ovins

{Ardèche, Corse, Alpes Maritimes, ..., Hérault, Var}. On retrouve les cultures permanentes, petites et très petites exploitations et élevage d'ovins.

Le lien entre ces deux sous-classes est motivé par les ovins et les structures d'exploitation. C'est un lien sensible car la réunion des deux sous-classes est accompagnée d'un noeud significatif avec un saut important de la statistique locale.

Conclusion : Par cette analyse, nous avons mis en évidence comment à un moment donné les variables caractérisant l'utilisation du sol, l'importance du cheptel et les structures de production, permettent de définir quelques grands systèmes agricoles.

Pour la répartition géographique des différentes catégories d'économie agricole, voir la figure 6.

5. COMPARAISON DES RESULTATS ISSUS DE L'ALGORITHME A.V.L. (COR, χ^2) AVEC CEUX DE C.A.H. (M2) χ^2 .

Pour la comparaison des deux méthodes au niveau de nos données sur l'agriculture française, nous avons retenu deux classifications de l'ensemble des départements par chaque algorithme :

- (a) Classification basée sur un seul caractère (la structure d'exploitation) : cas d'un seul tableau de contingence. Il était, en effet, plus facile pour ce caractère de contrôler la qualité relative des deux résultats.
- (b) Celle basée sur les trois caractères : cas de trois tableaux de contingence juxtaposés.

a) Un seul caractère : la structure d'exploitation

Nous avons déjà vu que l'A.V.L. (COR, χ^2) définit trois grandes classes d'après la structure d'exploitation :

- A. Exploitation de taille moyenne ou grande,
- B. Très grandes exploitations, et
- C. Petites ou très petites exploitations.

L'algorithme C.A.H. (M2, χ^2) produit une partition en sept classes (voir la liste, page 31) qui se réunissent selon le schéma ci-dessous (voire figure 7).

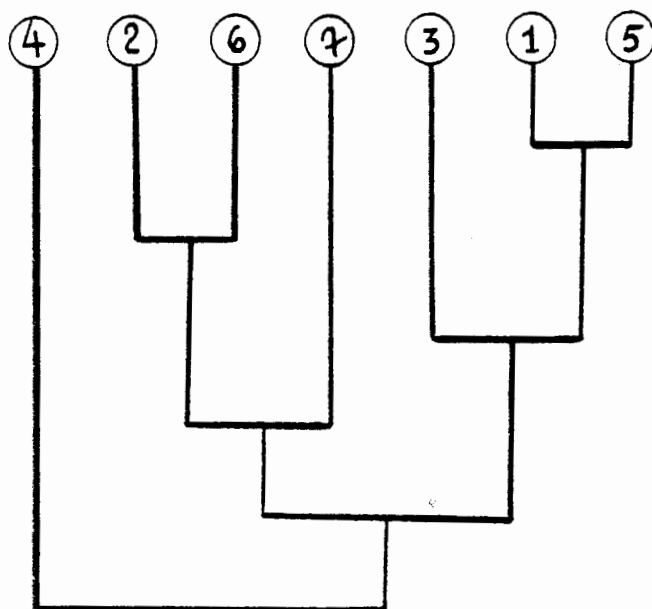


Figure 7 : Schéma d'agrégation de sept classes issues de l'algorithme C.A.H. (M_2, χ^2)

Interprétation : Les sept petites classes s'interpètent comme suit :

La classe [4] est composée des départements dominés par de très grandes exploitations ;

La classe [2] est caractérisée par des exploitations de taille grande ou très grande.

La classe [6] est celle des exploitations de taille moyenne ou grande ;

La classe [7] est caractérisée par les exploitations de taille moyenne (à l'exception de Savoie et de Vaucluse) et

Les classes [3], [1] et [5] sont formées des départements où dominent les petites ou très petites exploitations.

Si on veut retenir une classification en trois grandes classes d'après l'arbre précédent, on sera conduit à la partition {I, II, III}, où

Classe I = la classe [4],

Classe II = la réunion des classes [2], [6] et [7],

Classe III = la réunion des classes [3], [1] et [5].

En croisant les classifications en trois classes dûes aux deux algorithmes A.V.L. (COR, χ^2) et C.A.H. (M_2, χ^2) on obtient le tableau des effectifs suivant :

A.V.L. (COR, χ^2) \ C.A.H. (M_2, χ^2)	I	II	III	Somme
A	0	45	0	45
B	15	12	1 (Moselle)	28
C	0	2 Savoie Vaucluse	14	16
Somme	15	59	15	89

où on a noté, à l'intersection d'une classe A.V.L. et celle C.A.H. le nombre de départements qui y appartiennent simultanément.

Il est clair que les classes A et C correspondent respectivement avec les classes II et III, alors que les éléments de la classe B sont réparties entre les classes I et II. On remarque que les départements appartenant à $B \cap II$ sont dominés par de grandes et très grandes exploitations ; ce sont, en effet, les éléments de la petite classe [2].

Les deux départements à l'intersection des classes C et II, à savoir la Savoie et le Vaucluse, sont caractérisés par de très petites exploitations, et il est plus naturel de les trouver en classe C, plutôt qu'en II.

Le seul département appartenant à $B \cap III$, Moselle, pourrait aussi bien être classé en B qu'en III, car on y trouve à la fois des exploitations de grande et de petite taille.

Le tableau de croisement ci-dessous des deux classifications en trois et sept classes données par A.V.L. (COR, χ^2) et C.A.H. (M_2, χ^2) respectivement, permet d'établir la correspondance entre elles.

A.V.L. (COR, χ^2) \ C.A.H. (M_2, χ^2)	4	2	6	7	3	1	5	Somme
A	0	0	<u>23</u>	<u>22</u>	0	0	0	45
B	<u>15</u>	<u>10</u>	2	0	1	0	0	28
C	0	0	0	2	<u>8</u>	<u>1</u>	<u>5</u>	16
Somme	15	10	25	24	9	1	5	89
	I		II		III			

Les cases comportant les parts élevées d'effectifs de la ligne ou de la colonne ont leur contenu porté en *italic*.

On voit que la classe A correspond à $[6] \cup [7]$, la classe B à $[4] \cup [2]$ et la classe C se réfère à $[3] \cup [1] \cup [5]$.

Les 45 départements de $A \cap ([6] \cup [7])$ appartiennent à la catégorie des "exploitations de taille moyenne", les départements de $B \cap ([4] \cup [2])$ à la catégorie des "grandes ou très grandes exploitations", et ceux de $C \cap ([3] \cup [1] \cup [5])$ à celle des "petites ou très petites exploitations".

D'où on peut conclure que dans la classification des départements en tenant compte de la structure d'exploitation, l'algorithme A.V.L. (COR, χ^2) est plus fidèle aux associations de haut niveau que celui C.A.H. (M_2, χ^2).

b) Cas de tous les trois caractères

Rappelons que les quatre grandes classes issues de l'A.V.L. (COR, χ^2) appliqué aux départements décrits cette fois-ci par l'ensemble des modalités de tous les caractères, définissent les quatre catégories suivantes :

- A. Système extensif avec des céréales .
- B. Productions animales et spéculations fourragères.
- C. Productions animales associées à des S.T.H.
- D { 1. Type extensif avec des ovins.
2. Type intensif avec des ovins et cultures permanentes.

L'application de l'algorithme C.A.H. (M_2, χ^2) a donné quant à elle une partition de l'ensemble des départements en 9 classes (voir la liste, page 32), notées [1], [2], ..., [9] ; le schéma ci-dessous (figure 8) indique la suite des partitions successives avec de moins en moins de classes.

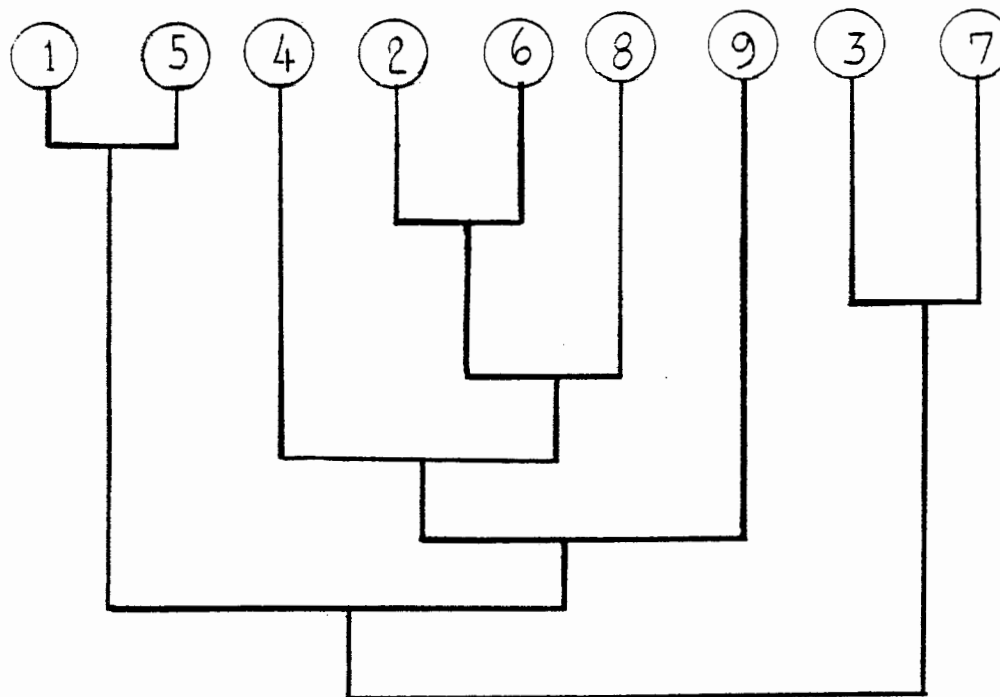


Figure. 8

Interprétation de la partition en 9 petites classes

Les caractéristiques communes des départements composant chacune des petites classes sont comme suit :

- [1] Système extensif : très grandes exploitations avec des céréales
- [5] Très grandes exploitations. Productions animales associées à des S.T.H. pour la moitié d'entre eux, et la culture des céréales pour les autres.
- [4] Exploitations de tailles diverses et de cultures très variées ; le seul point commun de ces départements est la présence des ovins.
- [2] Exploitations de taille moyenne associées à des S.T.H. et productions animales.
- [6] Exploitations de taille moyenne avec des cultures fourragères et des S.T.H., à l'exception de Loiret et de Loir-et-Cher, qui auraient dû se trouver en [1].
- [8] Exploitations de taille petite ou moyenne avec des S.T.H. et productions animales : avec des ovins dans certains départements et des bovins dans d'autres.
- [9] Exploitations de taille moyenne, productions porcines et plantes sarclées (système très intensif).
- [3] Petites et très petites exploitations avec des ovins.
- [7] Très petites exploitations avec des ovins et cultures permanentes.

Considérons dans ces conditions le tableau de croisement des deux classifications :

		I		II	III			IV	V		Somme
C.A.H.	A.V.L.	1	5	4	2	6	8	9	3	7	
A		<u>6</u>	<u>7</u>	1 Vienne	0	3	0	Nord 2 P.de.C	0	0	19
B		0	0	0	0	<u>7</u>	Rhône 2 Landes	<u>4</u>	<u>3</u>	0	16
C		0	<u>6</u>	Lozère 2 Basses Alpes	<u>13</u>	0	<u>5</u>	0	1 Savoie	0	27
D	D ₁	0	0	<u>3</u>	0	<u>5</u>	<u>8</u>	0	0	0	16
	D ₂	0	0	1 Corse	0	0	1 Drôme	0	Ardèche 2 Vaucluse	<u>7</u>	10
Somme		6	13	7	13	15	16	6	6	7	89

Comme précédemment les cases comportant des parts importantes d'effectifs des lignes ou des colonnes ont leur contenu porté en italique ; et les accolades indiquent les réunions des petites classes à un niveau plus élevé.

Si on veut retenir une classification en cinq classes d'après l'arbre (figure 8), on sera conduit à la partition {I, II, III, IV, V}, où $I = \{1\} \cup \{5\}$, $II = \{4\}$, $III = \{2\} \cup \{6\} \cup \{8\}$, $IV = \{9\}$, et $V = \{3\} \cup \{7\}$.

La classe I emprunte ses éléments principalement à la classe A et une partie de la classe C. Parmi les 13 départements de [5] on voit en effet que les 7 départements, à savoir Eure, Somme, Oise, Aisne, Côte d'Or, Cher et Indre, appartenant à A, se différencient des 6 autres appartenant à C quant au caractère "principales cultures" ;

Il s'agit des cultures céréalières qui dominent les premiers tandis que les derniers sont partagés entre les céréales et les S.T.H. On peut conclure en ce qui concerne cette comparaison que la classe I est "emportée" par les grandes exploitations, et les cultures ne sont pratiquement pas prises en compte.

La classe II a littéralement éclaté entre les classes A, C, D₁ et D₂. Considérons la répartition de sept départements de cette classe :

- Le département de Vienne est classé, par AVL (COR, χ^2) en A. En effet, il s'agit d'exploitations de taille plutôt grande avec des cultures céréalières.
- Lozère et Basses Alpes (Alpes de Haute Provence) appartiennent à C ; ils se caractérisent par de grandes exploitations avec des landes et des terres arables.
- Hautes Alpes, Aveyron et Haute Vienne se trouvent en D₁. On y voit des exploitations de taille moyenne associées à des S.T.H. et des ovins.
- Corse (D₂) possède de petites exploitations avec des cultures permanentes et des ovins.

Il semble ici que la seule caractéristique qui puisse réunir ces départements si différents tant par la structure d'exploitation que par les cultures, soit la présence des ovins : très forte dans les quatre derniers et plus faible dans les trois premiers.

C'est-à-dire que la classe II est "emportée" par les ovins alors que l'A.V.L. respecte tous les caractères.

La classe III est composée des éléments provenant notamment des classes B, C et D₁. Les 44 départements de cette classe sont répartis comme suit :

- Loiret, Loir-et-Cher et Haute-Garonne (1). Les deux premiers ont de grandes exploitations et tous les trois sont caractérisés par des céréales.
- Les 9 départements (B) sont caractérisés par des exploitations de taille moyenne (à l'exception, toutefois de Landes où la plupart des exploitations sont petites).
- Les 5 départements (C) ont des exploitations de taille moyenne associées à des S.T.H. et des productions animales.
- Les 13 départements (D₁) se caractérisent par des exploitations de taille petite ou moyenne avec des S.T.H. et des ovins.

- Enfin la Drôme (D_2) a de petites exploitations, des ovins et des cultures permanentes.

La classe IV comprend les quatre départements bretons (B) ; Nord et Pas de Calais (1).

- Les exploitations des départements du Nord et du Pas-de-Calais sont plus grandes que celles des quatre départements de Bretagne.

- Finistère, Côtes-du-Nord, Morbihan et Ille et Vilaine sont très typiques du système intensif avec des exploitations de taille moyenne, associées aux plantes sarclées, S.T.H. et des productions animales (porcs).

Enfin la classe V comportant 14 départements est la classe des petites exploitations ; leur répartition est la suivante :

- Les trois départements de $B \cap V$ sont Gironde, Bas Rhin et Haut Rhin. A l'exception de Gironde où abondent les cultures permanentes, les deux autres sont représentatifs des productions animales et spéculations fourragères.

- Le département de Savoie (en C) est associé à des S.T.H.

- Les 9 départements de $D_2 \cap V$ sont caractérisés par des ovins et des cultures permanentes.

Une fois encore, la "structure d'exploitation" est le seul caractère responsable dans la formation de la classe V par l'algorithme C.A.H. (M_2, χ^2).

CLASSIFICATION DES DEPARTEMENTS SELON LA STRUCTURE D'EXPLOITATION ALG :

C.A.H. (M_2, χ^2).

Nombre de terminaux demandés = 7

Niveau = 092

Ecart avec la partition suivante = 006

Ecart avec la partition précédente = 0033

Classe 1 : ALPM.

Classe 2 : VIEN, LRET, INLO, VNDE, ALIE, LOZR, INDR, SOMM, BALP, EURE.

Classe 3 : HTRH, GARD, ARDC, GDRE, TBLF, CORS? MOSL, DRME, BSRH.

Classe 4 : CHER, OISE, AISN, SETM, NVRE, EULO, MEUS, AUBE, HTMR, ARDN, YONN, MTML, CTOR, MARN, SETO.

Classe 5 : AUDE, BRHN, HRLT, PYOR, VAR.

Classe 6 : ORNE, CLVD, VOSG, SMRT, AVRN, HTVN, CHMM, LOCH, SNLO, JURA, CHRT, 2SVR, HTGR, PDEC, SRTH, MNLO, LRATp TARN, NORD, MYNN, HTSN, GERS, DOUB, CRES, CNTL.

Classe 7 : MRBH, FNST, ILVL, TAGA, LTGR, LOT, HALP, LOREq AIN, AREG, PDME, BSPY, DRDN, CREZ, HTLO, VAUC, SVIE, CIND, ISER, MNCH, HTSV, RHNE, HTPY, LNDE.

CLASSIFICATION DES DEPARTEMENTS TOUS CARACTERES ALG : C.A.H. (M_2, χ^2)

Nombre de terminaux demandés = 9

Niveau = 098

Ecart avec la partition suivante = 028

Ecart avec la partition précédente = 018

Classe 1 : YOMN, AUBE, EULO, SETM, MARN, SETO.

Classe 2 : SRTH, MYNN, ORNE, CLVD, SMRT, CNTL, HTSN, JURA, DOUB, MOSL, SNLO, TBLF, VOSG.

Classe 3 : GRDE, AROC, BSRH, HTRH, SVIE, VAUC.

Classe 4 : AVRN, HTVN, VIEN, CORS, HALP, BALP, LOZR.

Classe 5 : EURE, SOMM, OISE, AISN, CTOR, CHER, INDR, MTML, NVRE, ALIE, MEUS, HTMR, ARDN.

Classe 6 : CHMM, INLO, LRET, LOCH, VNDE, MNLO, LRAT, TARN, HTGR, CHRT, TAG, LIGR, GERS, CRES, 2SVR.

Classe 7 : ALPM, VAR, PYOR, HRLI, AUDE, GARD, BRHN.

Classe 8 : DRON, HTLO, CREZ, HTPY, AREG, LOT, BSPY, LNDE, ISER, RHNE, DRME, MNCH, HTSV, LORE, AIN, PDME.

Classe 9 : FNST, CTND, MRBH, ILVL, NORD, PDEC.

6. AUTRES EXEMPLES ETUDIES

6.1. Evolution de la construction des logements entre 1968 et 1975 sur les communes des Bouches du Rhône.

Etude à caractère démographique à partir de données portant sur le mode d'habitat et de financement de ce dernier en liaison avec les catégories socio-professionnelles et la structure de la commune.

Les données nous ont été communiquées par M. Roncayolo (Groupe de Géographie Sociale (E.H.E.S.S.)) qui a permis l'évaluation des résultats.

Dans l'exemple précédent (cf. fig. 6), la classification de l'ensemble des unités géographiques à partir de la seule proximité statistique a naturellement entraîné pour les classes obtenues un très bon respect de la condition de connexité spatiale.

Cette circonstance ne s'est pas produite pour ces données et les suivantes ; de sorte que A. Prod'homme, qui a travaillé cet exemple ainsi que le suivant dans le cadre d'une thèse de 3ème cycle (cf. [9]), a réalisé une adaptation de l'A.V.L. qui tient compte de la contrainte de contiguïté pour ne former que les "meilleures" classes connexes,

6.2. Répartition géographique de la viticulture girondine

Les données qui concernent l'année 1975, se rapportent par conséquent à une culture pérenne. Les éléments de base sont collectés à partir des déclarations de récoltes, totalisées commune par commune sur le département et relatives aux surfaces, aux volumes, au statut des exploitants et

à la taille des exploitations.

Les données ont été fournies par la "Chambre d'Agriculture de la Gironde" et nous ont été soumises par L. Breton (Maître-Assistant en M.I.A.G.E., Université Rennes I) qui a participé à l'interprétation et évaluation des résultats.

6.3. Répartition géographique des décès dans les communes de deux départements Bretons, relativement aux causes déclarées.

Les deux départements sont l'Ille et Vilaine et le Finistère. Un traitement séparé a été effectué pour chacun des deux départements, On a ainsi eu à prendre en compte deux tableaux de contingence où l'ensemble des lignes du premier (resp. du second) est indexé par l'ensemble des communes d'Ille et Vilaine (resp. du Finistère) et où l'ensemble des colonnes est indexé par les principales causes de décès déclarés. 19 ont été retenues (cf.[5]).

Ce type d'analyse a bien été effectué, mais à un niveau beaucoup plus global où l'unité géographique est le département. L'étude dont il est question ici est donc sensiblement plus difficile.

Les données nous ont été soumises par le Dr L. Massé (Epidémiologue, Ecole de la Santé de Rennes) qui a permis l'interprétation et l'évaluation des résultats.

Le traitement et l'élaboration d'une technique d'aide à l'interprétation ont été réalisés par A. Lelièvre dans le cadre de son travail de D.E.A. (cf.[5]) qui présente différents aspects dans l'appréhension et le codage de ces données.

6.4. Transmission du statut social

Courant 1977 de nombreuses familles de Loire Atlantique ont rempli un questionnaire sur leurs aspirations pour l'avenir professionnel de leurs enfants scolarisés de 14 à 16 ans en formulant d'une part la profession *souhaitée* (en absence de tout obstacle) et d'autre part, la profession *conseillée* (compte tenu de la réalité vécue) pour leur rejeton.

Plusieurs tableaux de données ont été analysés par la classification. Un même tableau a la structure d'une juxtaposition horizontale de tables de contingence. La variable dont les modalités indexent les lignes d'un tel tableau peut être définie par la catégorie socio-professionnelle d'un seul membre de la famille (e.g. C.S.P. du père) ou bien par le couple de catégories socio-professionnelle d'un couple de membres de la famille (e.g. (C.S.P. du père, C.S.P. de la mère) ou (C.S.P. du grand-père, C.S.P. du père)), ou même par un triplet de catégories socio-professionnelles de trois membres de la famille (e.g. (C.S.P. du grand-père, C.S.P. du père, C.S.P. de la mère)). Une même table de contingence de la juxtaposition croise la variable qu'on vient de définir, avec la variable "profession souhaitée" ou (exclusif) "profession conseillée" par le père ou (exclusif par la mère.

La classification des lignes a montré qu'il était significatif de remonter jusqu'à la deuxième génération et a de la sorte permis de mettre en évidence des "lignées" (cf.[8]).

Le traitement informatique et statistique a été effectué par G. Lecrenais dans le cadre d'un stage à l'I.N.S.E.E. de Nantes qui a largement débordé la stricte préparation de son mémoire de D.E.A. (cf. [4]).

L'interprétation des résultats a été conduite par Cl. Thélot (chargé de mission à l'I.N.S.E.E. de Nantes).

6.5. Reconnaissance de formes linguistiques utilisées

Ce sont des données provenant de conversations avec 120 individus à Montréal ; conversation sur la vie, les coutumes passées et présentes dans la ville. Pour chaque locuteur, on dispose d'une distribution des fréquences des formes linguistiques sur un échantillon de 8000 mots. On s'intéresse à la répartition des catégories socio-professionnelles à travers les classes d'une partition significative sur l'ensemble des 120 locuteurs ; partition obtenue par la classification des lignes du tableau de contingence indexé par Locuteurs \times Mots utilisés. Les données ont été mises à disposition par D. Sankoff (Centre de Recherches Mathématiques, Université de Montréal) ; elles sont travaillées sous différents aspects par A. Morin dans le cadre de sa future thèse de 3ème cycle.

7. CONCLUSION

L'indice de proximité que nous avons défini (cf. § V, chap. 2, partie I et étendu dans [6] ouvre pour notre méthode très générale de classification hiérarchique, la voie à une classe de structures de tableaux de données qui est le domaine privilégié de l'Analyse Factorielle des Correspondances et de la Classification des moindres carrés conformément à la métrique du χ^2 . Les résultats que nous avons pu obtenir jusqu'à présent sont d'une grande finesse et d'une grande cohérence et la comparaison avec les résultats de méthodes maintenant plus classiques, est pleine d'intérêt.

BIBLIOGRAPHIE

- [1] J.P. BENZECRI et Collaborateurs, "L'Analyse des Données" - Tome 1. "La Taxinomie", Tome 2. "Analyse des Correspondances". Dunod, 1973.
- [2] M. JAMBU et M.O. LEBEAUX, "Classification Automatique pour l'Analyse des Données" - Tome 1. "Méthodes et Algorithmes", tome 2. "Logiciels". Dunod 1978.
- [3] L. LEBART, A. MORINEAU et N. TABARD, "Techniques de la Description Statistique". Dunod, 1977.
- [4] G. LECRENAIS, "Classification de Catégories Socio-Professionnelles de Parents relativement aux Projets d'Avenir qu'ils formulent pour leurs enfants", rapport de D.E.A., Université de Rennes 1, Laboratoire de Statistiques, Septembre 1978.
- [5] A. LELIEVRE, "Méthodes de Classification Hiérarchiques Appliquées à la Répartition des Décès attribuables à certaines causes dans les communes de deux départements bretons", rapport de D.E.A., Université de Rennes 1, Laboratoire de Statistiques, Septembre 1979.
- [6] I.C. LERMAN et B. TALLUR, "Classification des éléments constitutifs d'une juxtaposition de tableaux de contingence", rapport I.R.I.S.A. n° 127, Université de Rennes I (1980) et Rev. Stat. Appl. (1980) n°28,3.
- [7] B. TALLUR, "Etude Régionale de l'Agriculture Française", rapport I.R.I.S.A. n° 103, Université de Rennes, I, 1978.

- [8] Cl. THELOT et Th. LEMAITRE, *"La transmission du statut social : l'exemple des Pays de la Loire"*, rapport interne, Direction régionale de l'I.N.S.E.E. de Nantes, 044 n° 3/NR, 2 Février 1979.
- [9] A. PROD'HOMME, *"Indice d'explication des classes obtenues par une méthode de classification hiérarchique respectant la contrainte de contiguïté spaciale. Application à la viticulture Griondine et à la construction de logements dans les Bouches du Rhône"*. Thèse de 3ème cycle, Université de Rennes I, Décembre (1980).