

CARTOGRAPHIE SYNTHETIQUE D'UN ENSEMBLE DE VARIABLES REGIONALISEES PAR UN LOGICIEL INTEGRE D'ANALYSE DES DONNEES

(Marie-Madeleine THOMASSIN)

RESUME - L'article décrit l'ensemble des ressources offert par le logiciel MacSAIF (*) pour le traitement de variables régionalisées, et en particulier l'aspect d'aide à l'interprétation qui utilise les méthodes statistiques de l'analyse des données (analyse factorielle de correspondances et classification ascendante hiérarchique).

En s'appuyant d'abord sur un cas simple, pour ensuite montrer le processus de traitement général dans le cas de données complexes réquerant un codage préalable, l'auteur met en évidence les trois étapes principales du travail : choix du découpage en modalités des variables à traiter, choix de la partition et choix des trames utilisées sur la carte.

L'utilisation des différents programmes qui composent le logiciel, et qui permettent de réaliser chacune de ces étapes, est décrit et les principales difficultés d'application sont relevées et illustrées par des exemples.

INTRODUCTION

Le but de cette communication est de montrer aux géographes, aux chercheurs des autres disciplines ainsi qu'aux spécialistes et responsables de la planification, la variété de ressources que présente ce logiciel Mac SAIF. Au chercheur isolé sur le terrain, il permet, à partir d'une base de données complexes, de créer une carte de synthèse intégrant les informations élémentaires apportées par des variables, dont le nombre et le format ne sont soumis à aucune contrainte.

Pour présenter ce logiciel, on a fait très largement appel aux notices des programmes, publiées par le Pr. J.-P. et F. Benzécri dans les Cahiers d'Analyse de Données.

Après avoir brièvement rappelé le traitement par l'analyse des données (c'est-à-dire par l'analyse factorielle des correspondances, AFC et la classification ascendante hiérarchique, CAH) du cas le plus simple : le tableau de contingence, on a abordé le cas des données complexes ou requérant un codage préalable.

(*) Logiciel MacSAIF, "Système d'analyse des informations", créé par le Professeur et F. BENZECRI. Laboratoire de statistique, Université de Paris VI.

Le rôle des programmes de création de tableaux a été illustré à partir de leur application aux données relatives à l'agriculture et à l'élevage de la Sierra équatorienne. La carte de synthèse construite par le programme 'carthage' traduit une typologie clairement explicitée, en délimitant des zonages dont le graphisme exprime les affinités entre classes.

1. ANALYSE DES DONNEES DANS LE CAS D'UN TABLEAU SIMPLE

Partons d'un exemple de géographie économique : le tableau de données croisant l'ensemble des départements de la France, avec un ensemble de variables (les consommations mensuelles du gazole, de 1972 à 1981).

Ce tableau comporte 95 lignes, correspondant aux 95 départements français (la Corse étant considérée comme constituant un seul département), rangés depuis l'Ain jusqu'au Val-d'Oise dans l'ordre de leurs numéros minéralogiques. Le tableau a 12 colonnes, qui sont les mois de Janvier à Décembre. A l'intersection, par exemple, de la ligne Ardèche et de la colonne Juillet, on lit le nombre 236 : ceci signifie que la consommation en gazole dans le département de l'Ardèche au cours du mois de Juillet des années 1972 à 1981 a été au total de 236 centaines de m³.

Ce tableau de données va être soumis tel quel à l'AFC.

L'analyse de ce tableau montrera d'abord une opposition entre les départements caractérisés par une importante consommation hivernale et très faible consommation estivale, et ceux où prédomine la consommation estivale, ou tout au moins les départements où elle s'avère moins réduite qu'il n'est le cas en moyenne. Apparaîtra ensuite le cas des départements betteraviers, à forte consommation automnale.

Afin de mettre en évidence avec précision cette structure, et ultérieurement de produire une carte représentant la régionalisation de ces consommations différenciées selon les saisons, nous devons utiliser l'ensemble des ressources de l'analyse des données :

- l'**analyse factorielle**, qui assigne, d'après le tableau initial, des coordonnées, ou facteurs, à chacun des départements ainsi qu'à chacun des mois. Autrement dit, qui traduit en termes de *proximité* sur un graphique les *ressemblances* existant entre *individus* et entre *variables*, ainsi que les *affinités* des *uns* avec les *autres* ; et ensuite,

- la **classification automatique**, qui, au sein de cette représentation géométrique, *détermine des classes* (dans notre cas, principalement des classes de départements à chacune desquelles sera affectée ultérieurement, lors du tracé de la carte, une trame unique, tandis que les classes de mois, déjà en évidence sur les graphiques-plans, nous sont connues a priori).

La première étape d'une analyse de données sera évidemment l'introduction de ce tableau dans l'ordinateur.

Le tableau étant saisi suivant un format convenable (et d'ailleurs très souple), on entre dans le programme '**qori**' d'analyse des correspondances. Outre un fichier de type texte qui contient sous forme de tableaux des résultats numériques dont une partie s'affiche à l'écran, à la fin du déroulement du programme, le programme '**qori**' crée des fichiers numériques, où sont gardés les mêmes résultats d'analyse des correspondances et qui seront utilisés dans toute la suite des programmes de la chaîne : dans les programmes '**planF**' et '**planX**', pour produire les graphiques-plans, puis dans la classification automatique, et enfin dans l'élaboration de la carte proprement dite.

Succinctement, le programme '**planF**' crée, en mode conversationnel, des plans croisant deux axes factoriels qui s'affichent instantanément à l'écran ; tandis que le programme '**planX**', également conversationnel, crée les mêmes graphiques sur un fichier texte, qui peut être consulté grâce à un éditeur de texte et, ensuite, imprimé sur papier.

On entre ensuite dans le programme de '**CAH**' (**Classification Ascendante Hiérarchique**).

Ce programme a de multiples fonctions; notamment le tracé de l'arbre, la création de tableaux d'aides à l'interprétation, d'une grande utilité pour le praticien de la classification automatique, et tout particulièrement pour le géographe.

Le programme '**CAH**' consulte alors le fichier numérique des facteurs, qui a été créé par le programme '**qori**' d'analyse des correspondances sans lequel le programme de classification automatique ne pourrait s'exécuter ; il lit ces facteurs et, par des calculs de distances, il range les uns avec les autres les éléments qui sont le plus proches. Ceci fait, il nous reviendra d'extraire de l'arborescence ainsi construite une partition destinée à être traduite cartographiquement.

En bref, il s'agit de retenir de la classification ascendante hiérarchique, qui comporte dans l'exemple retenu, au-dessus des 95 éléments (départements), 94 noeuds, une sous-hiérarchie à la base de laquelle se trouvera une partition, par exemple en neuf classes, avec, au-dessus de ces neuf classes, des réunions successives en classes supérieures, jusqu'à parvenir au sommet. Pour effectuer un tel choix de façon tout à fait conséquente, il faut avoir une connaissance approfondie des résultats de la classification, et c'est pourquoi l'on prévoit, après une première consultation, de rentrer dans le programme '**CAH**' pour demander à bon escient une sous-hiérarchie.

Quant aux **aides à l'interprétation**, nous nous bornerons à souligner l'importance du listage **Vacor**, qui donne l'interprétation des classes par leurs profils, c'est-à-dire directement en termes des variables du tableau des données.

2. CAS DES DONNEES COMPLEXES OU REQUERANT UN CODAGE PREALABLE

2.1. Tableau de contingence avec un nombre élevé de variables.

Cet exemple de la consommation mensuelle du gazole est particulièrement simple parce que, d'une part, le *profil de chaque ligne* (c'est-à-dire la suite des quotients des éléments de la ligne par leur total) est *directement interprétable* : on dit qu'il s'agit d'un *tableau de contingence* ; et que, d'autre part, le *nombre des variables* (colonnes du tableau) n'est pas très élevé.

Si le *nombre des variables est élevé*, on peut, par le programme 'CAH', faire un *Vacor* expliquant la hiérarchie des départements i , non par les variables j elles-mêmes (les mois), mais par les *classes j_q de la partition de l'ensemble des variables* qu'il vient d'obtenir par classification des j mois. Dans le cas d'un ensemble de douze variables seulement, les douze mois de l'année, l'interprétation de la classification, et donc celle de la cartographie, directement en termes de ces variables est possible, et même facile. Mais prenons dans le même domaine un tableau à 48 colonnes : le *tableau donnant les consommations mensuelles de quatre produits pétroliers*. L'interprétation directe en termes de ces 48 variables devient difficile, sinon impossible. Il faut donc *procéder à l'agrégation de ces variables*, et, donc, considérer en quelque sorte un *tableau intermédiaire* donnant, par exemple, au lieu des consommations en gazole, fuel lourd, fuel domestique, carburant auto, mois par mois, les *consommations de chacun de ces produits suivant des périodes plus longues choisies selon le produit* : saison entière, moitié de l'année, même dans certains cas, trois quarts de l'année opposées à la période d'été etc.. De sorte que l'utilisateur, finalement placé en présence d'un tableau d'une dizaine de colonnes, et de pourcentages calculés sur ce tableau par le programme d'aide à l'interprétation, et disposant de graphiques-plans où s'affichent les sigles de ces agrégats de variables, pourra comprendre sans peine l'interprétation des classes de départements créées.

L'interprétation de la classification se fait donc en termes d'agrégats de variables et partant, la cartographie.

La légende de la carte distingue pour chaque produit, selon des périodes variables, les excédents ou les déficits par rapport à la moyenne, forts ou faibles des consommations mensuelles, de 1972 à 1981. De là, la nécessité d'une recherche au niveau des modes d'expression graphique, l'utilisation des symboles s'étant avérée peu satisfaisante et de lecture peu aisée, comme chacun sait, la création de nouvelles trames par le programme 'carthage' s'est révélée indispensable.

2.2. Tableau des données requérant un codage préalable

Un des cas les plus simples est celui des tableaux de réponses à un questionnaire clos.

Lorsque l'utilisateur non programmeur se trouve confronté, et c'est très souvent le cas du géographe, à un tableau de données *croisant un ensemble d'unités administratives avec un ensemble de variables numériques hétérogènes*, il se trouve dans la même situation que s'il avait à traiter un tableau de réponses à un questionnaire clos, où chaque réponse est initialement notée comme un numéro de modalité. En effet, en découpant en p sous-intervalles successifs l'intervalle de variation d'une variable continue, on peut coder celle-ci comme une variable discrète à p modalités. On peut donc assimiler la variable à une question, et la circonscription au sujet qui y répond. Un tel tableau doit être éclaté, sous forme binaire, en (zéro,un), en attribuant à chaque modalité une colonne.

Le tableau pris pour exemple d'application est extrait du tableau d'ensemble, ayant servi de base à une étude de la régionalisation agricole de la Sierra équatorienne (Cf bibliographie), croisant 238 paroisses avec 48 variables : les 26 premières relatives au milieu physique et à ses potentialités, les 9 suivantes à l'agriculture et à l'élevage, les 13 dernières aux critères socio-économique. Tableau d'ensemble découpé respectivement en 3 blocs (bloc 1, bloc 2, bloc 3). A chacun des blocs ainsi formés et au bloc d'ensemble ont été adjoints des éléments supplémentaires : des variables (en colonnes) et des individus, 42 paroisses (en lignes).

Le tableau de la présente application croise donc 280 paroisses, dont 42 en supplémentaires avec les 9 variables afférentes à l'agriculture et l'élevage.

3. LE ROLE DES PROGRAMMES DE CREATION DE TABLEAUX

Quelques programmes de création de tableaux suffisent à l'utilisateur non programmeur pour dépouiller par lui-même une enquête par questionnaire, et analyser un ensemble de variables hétérogènes recueillies sur le terrain. La présente notice est consacrée à un ensemble de tels programmes.

Le plus important de ces programmes, '**zrang**', sert à ranger, observer et découper les variables d'un tableau :

- *ranger*, c'est-à-dire ordonner par valeurs croissantes ;
- *observer*, en affichant des tableaux de valeurs, des histogrammes, des croisements, (et même des cartes analytiques, servant à établir celle qui rendra compte de la synthèse générale des données effectuée par l'analyse des correspondances et la classification automatique) ;
- *découper*, en fixant un ensemble de modalités et en créant un tableau de numéros de modalités ou un tableau binaire.

Un programme créant des histogrammes doit être conversationnel. Il suffit, en effet, d'un individu isolé très écarté pour contraindre tous les autres à s'enfermer dans un seul créneau. D'autre part, connaître les noms des individus rentrant dans chaque créneau s'avère intéressant. On s'est efforcé, dans le programme '**zrang**', d'utiliser au mieux la place disponible. Ainsi, il est facile de décider du choix des bornes des modalités en lesquelles sont découpées les variables (c'est-à-dire déterminer ce que l'on entend par : très fort, fort, moyen, faible et très faible pour chacune des variables).

Le listage du fichier 'Dcodx' (en annexe) résulte de l'application du programme 'Zrang' (option D) au tableau de données - ici agriculture - élevage. Il indique les bornes du découpage des 9 variables en 34 modalités. Ces bornes correspondent aux bornes initiales, le recodage de ces variables, à la différence de celles afférentes aux autres thèmes, ayant été jugé inutile.

Le tableau à analyser croise donc 280 paroisses avec les 34 modalités issues des 9 variables précitées. Sur ces 280 paroisses, 42 d'entre elles comportant des lacunes de l'information, ont été mises en supplémentaires par le programme 'qori'.

En analyse des correspondances, on appelle *élément supplémentaire* une ligne ou une colonne dont on ne tient pas compte pour déterminer les axes factoriels mais qui est projetée sur ceux-ci. 'Qori' permet de mettre en supplémentaire un sous-ensemble de lignes ou de colonnes du tableau soumis à l'analyse.

Il peut s'agir d'éléments apportant à un axe une contribution telle que l'analyse en est perturbée ; ou encore d'éléments pour lesquels certaines données sont absentes, comme c'est le cas ici, ou imprécises. Souvent, il est préférable de faire totalement disparaître ces éléments, comme permet de le faire le programme 'soustab'.

Parfois, on désire projeter en supplémentaires des éléments qui n'appartiennent pas au tableau de base analysé pour éclairer les résultats de l'analyse et vérifier les hypothèses qu'elle suggère. Si, comme il est d'usage, on place en lignes les individus et en colonnes les variables, on aura, dans le premier cas, un *tableau externe de lignes supplémentaires* ; et, dans le deuxième cas, un *tableau externe de colonnes supplémentaires*. Nous disons *tableau externe* parce qu'il ne s'agit pas d'un sous-tableau du tableau de base soumis à l'analyse. C'est souvent de l'adjonction de tableaux externes appropriés que dépendra l'interprétation des données principales que l'on se propose de représenter sur la carte de synthèse.

Le programme 'qorlsup' sert à créer, pour un ou plusieurs tableaux supplémentaires externes, (de lignes ou de colonnes), un fichier numérique de facteurs et un listage usuel.

Dans l'étude de la Sierra, l'analyse d'un bloc comportait systématiquement la mise en supplémentaires des deux autres, ainsi qu'un certain nombre de variables relatives aux coordonnées géographiques et à la superficie des paroisses (en colonnes) et les 42 paroisses déjà mentionnées (en lignes). L'exploitation des listages relatifs aux éléments supplémentaires à été déterminante pour l'interprétation du tableau principal et la définition des sept facteurs issus de l'analyse, et ce, d'autant plus que les coordonnées géographiques et la taille des paroisses, facilitaient l'association entre certains secteurs de la Sierra et/ou certains types de paroisses (bassins, versants) et les caractéristiques exprimées par les demi-axes considérés.

Enfin, le programme 'discr', partant de deux fichiers numériques de facteurs : le fichier des centres de classes et le fichier des individus permet d'affecter chaque individu au centre dont il est le plus proche. Dans la mesure où chaque individu est censé relever d'une classe, il s'agit de ce qu'on appelle classiquement une *procédure de discrimination* ; d'où le nom du programme. Un exemple

simple est fourni par l'application, où l'ensemble des individus (paroisses) est l'ensemble des usuels ; et les centres sont les i_q , centres de gravité des classes d'une partition des i , obtenues par **CAH**, soit les onze classes de la partition (Cf. figures 3 et 4).

Dans la pratique, '**discr**' servira particulièrement à affecter à des centres, résultant de l'analyse d'un premier ensemble de données, un ensemble d'individus adjoints en supplémentaires à cette analyse (ici les 42 paroisses présentant des lacunes de l'information). D'où le lien de '**discr**' avec '**qorlsup**'.

4. LA CARTOGRAPHIE PROPREMENT DITE

Outre les données statistiques, résultats de l'analyse factorielle et de la C.A.H., le programme '**carthage**' utilise des données numériques représentant le fond de carte, pour construire la carte.

L'interprétation des classes de paroisses en termes de classes de variables (agrégats de modalités), grâce à un Vacor, a permis d'établir une typologie clairement définie (Cf. figures 1 et 2).

Pour cartographier cette typologie relative aux dominantes culturelles et à l'élevage de la Sierra, plusieurs étapes ont été nécessaires. Tout d'abord créer un fond de carte, c'est-à-dire *numériser la carte*. Cette numérisation a été effectuée à partir du 1:750 000 (échelle des cartes de l'étude), aucun document d'échelle plus grande n'étant disponible. Pour ce faire, créer deux fichiers :

- le premier, *fichier des coordonnées géographiques des points*, au nombre de 1217, choisis comme sommets des lignes polygonales délimitant les unités administratives (paroisses), et numérotés séquentiellement ;

- le second concernant *les contours des unités administratives elles-mêmes*, au total 280 (238 principales, 42 supplémentaires), assimilées chacune à un polygone et décrites par son numéro de code, le nombre de sommets placés sur ses contours et les numéros de ces sommets.

Précisons que les mesures de la latitude et de la longitude ont été effectuées en mm, à partir d'une origine située au nord-ouest de la feuille. Il n'a pas été nécessaire de se soucier de l'échelle du document servant de base à la numérisation, le programme de création des fichiers numériques mettant lui-même les nombres à l'échelle commode pour l'exécution du tracé de la carte.

Le programme '**litxcarte**' a permis de compiler les fichiers de coordonnées de points et de contours des paroisses, pour créer un *fond de carte* au format voulu par '**carthage**'. Il convient de souligner qu'il rend possible également la vérification, unité par unité, de la *validité des contours de chacune d'elles*. Plusieurs parcours exploratoires ont permis de corriger les erreurs commises lors de la saisie.

Pour imprimer le fond de carte on se sert de la trame n° 12 (Cf. figure 5)

En outre, '**carthage**' propose à l'utilisateur un choix de trames pratiquement illimité ; car, d'une part, lui sont directement accessibles 32 trames prises dans un tableau de 100, à l'entrée du programme (et il peut modifier le choix

Fig.1
 Consommation mensuelle du gazole en France par département
 (1972 - 1981) par A.Moussaoui

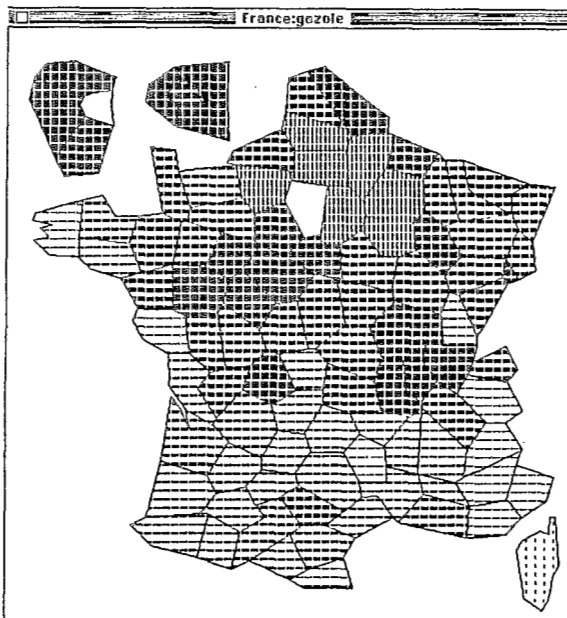
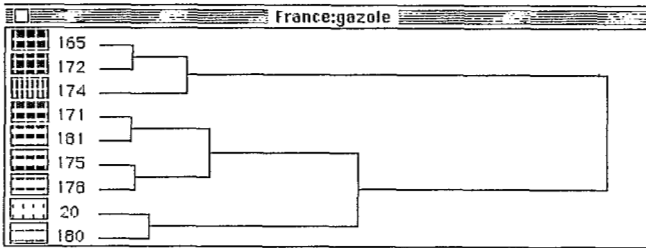
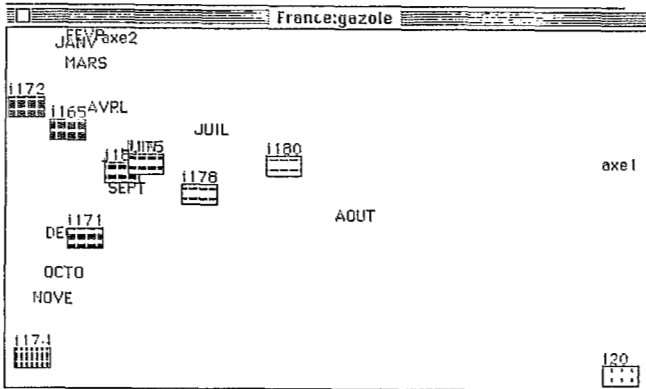


Fig. 2
Classification des variables

c Partition en 17 classes : Sigles des individus de la classe numero c		
361	BV0 CR2	SUPERFICIES CEREALIERES TRES DEVELOPPEES (40 A 70%/sup.tot.utilisee) PEU D ELEVAGE BOVIN(0 A 1 ^o TETES/100 ha d une U.P.A.)
511	MZ1 LFO OV1	PAS DE CULTURES MARAICHERES, FRUITIERES ;PEU DE MAIS(0 A 40 %/sup. tot.utilisee) ELEVAGE OVIN MODERE (20 A 39 TETES/100 ha d une U.P.A.)
421	PT1 TR0	PAS DE CULT. TROPICALES , CULT. SPECULATIVE DE LA POMME DE TERRE (0 A 40%/sup.tot. utilisee)
41	BV3	ELEVAGE BOVIN INTENSIF (60 A 99 TETES/100 ha d une U.P.A.)
351	OV2 BV2	ELEVAGES OVIN ET BOVIN DEVELOPPEES(40 A 59 TETES/100 ha d une U.P.A.)
431	CR3 MZ0	PLUS DE 70% en CEREALES/sup.tot.utilisee;pas de MAIS
271	LF2	40 A 70% en CULTURES MARAICHERES ET FRUITIERES
501	OV3 PA2 IR2	ELEVAGE OVIN INTENSIF(60 A99 TETES/100 ha d une U.P.A.) PATURAGES ET SUPERFICIES IRRIGUEES TRES DEVELOPPEES (40 A 70 %/sup.tot. utilisee)
381	IR3 PA3	PLUS DE 70% EN PRAIRIES IRRIGUEES
301	TR2	CULTURES TROPICALES TRES DEVELOPPEES (40 A 70% /sup. tot. utilisee)
471	BV1 OV0 TR1	ELEVAGE BOVIN MODERE(20 A 39 TETES /100 ha d une U. P. A.); PAS D OVIN;PEU DE CULTURES TROPICALES(0 A 40%/sup.tot. utilisee)
491	LF1 CR1 MZ2	FAIBLE % EN CULTURES MARAICHERES ET FRUITIERES(0 A 40%/sup. tot. utilisee);40 A 70% EN MAIS)
401	IR1 PA1	FAIBLE % EN SUPERFICIES IRRIGUEES ET EN PATURAGES(0 A 40%/sup. tot. utilisee)
441	OV4 BV4	ELEVAGES BOVIN ET OVIN TRES INTENSIFS(>100 TETES/100 ha d une U.P.A.)
481	MZ3 PA0	CULTURE DU MAIS TRES DEVELOPPEE(>70%/sup.tot. utilisee)
311	IR0	PAS DE CULTURES IRRIGUEES
451	CR0 PT0	PAS DE CEREALES NI DE POMMES DE TERRE

de ces 32 trames sans sortir de 'carthage'), d'autre part il peut écrire des fichiers numériques décrivant des trames nouvelles qui se placent automatiquement à la fin du tableau des 100. De là, une totale liberté pour arrêter les modalités de l'expression cartographique de son choix.

Il importe de signaler que le choix des trames peut être guidé conversationnellement par l'affichage des résultats de l'analyse factorielle et de la classification automatique ; résultats qui ne sont autres que l'expression la plus simple des similitudes existant entre les réalités géographiques correspondant à chaque classe.

On peut en effet demander que s'affichent, sur un plan croisant deux axes factoriels, non seulement les numéros des classes de paroisses, précédés de la lettre i, par exemple i 460 (classe n° de paroisses, dont la composition peut être recherchée sur les listages de classification automatique et apparaîtra sur la carte avec une trame particulière), mais aussi l'affichage de ce n° avec, en-dessous dans un caisson rectangulaire, la trame actuellement choisie pour représenter les paroisses de cette classe (Cf. figure 3).

Il importe de noter qu'avant l'affichage d'une première version de la carte, sont proposées, dans des caissons carrés alignés aux marges gauche et supérieure de l'écran, des trames qu'on peut choisir pour remplacer celles proposées dans de grands caissons rectangulaires associés aux classes (Cf. annexe). La disposition des trames dans les caissons carrés apparaît, dès l'entrée dans le programme, avec un texte offrant de modifier cette palette de 32 trames à partir du tableau de 100.

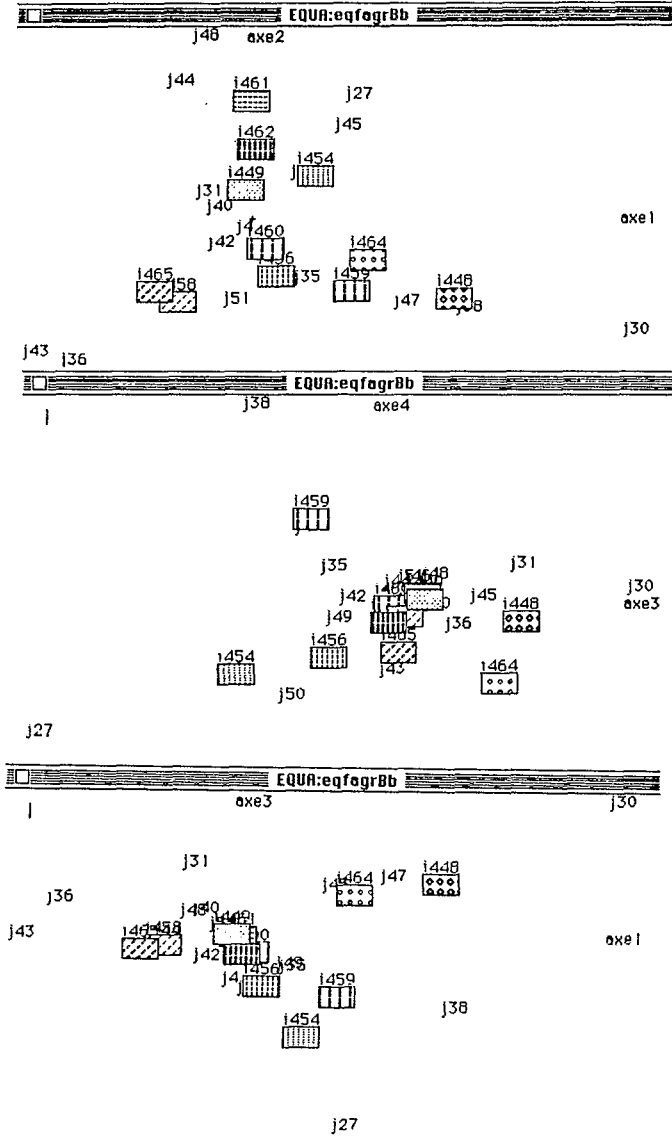
Le choix des trames est effectué en conformité avec l'interprétation des classes de paroisses en termes de classes de variables, en vue de rendre compte de leurs caractéristiques et traduire les degrés divers de spécialisation de l'agriculture et de l'élevage. Ce choix exprime aussi les affinités et les oppositions entre classes, délimitées sur la carte par des zonages (Cf. figure 4). Pour ce faire, on a joué sur l'orientation, le pas et la texture des trames, la valeur des grisés, la forme des éléments graphiques. La taille de la paroisse a influé grandement sur le choix du graphisme, et il a fallu, en effet, faire plusieurs parcours avant de l'arrêter définitivement.

L'expression cartographique des paroisses principales se différencie des supplémentaires. Sur les cartes de l'étude, imprimées en couleurs, les aplats représentaient les premières, les lignés et tiretés représentaient les secondes, selon l'importance des lacunes de l'information. Dans le cas présent, on a traduit le caractère incertain de l'appartenance à une classe en superposant des lignes obliques lâches (trame n° 13) à la trame individualisant la classe d'appartenance supposée.

Le programme 'carthage' permet en effet de combiner des trames. Il suffit pour cela, dans le fichier de fond de carte, de faire figurer deux fois la même paroisse, la première comme appartenant à une classe donnée, la seconde comme appartenant à la classe des supplémentaires.

Lorsque la carte s'avère pleinement satisfaisante, on peut immédiatement imprimer ce qui s'affiche à l'écran, et également le saisir comme un fichier-image, que l'ordinateur permet de modifier par un logiciel graphique, pour ajouter des

Fig 3
Dominantes culturelles et élevage
Plans graphiques issus de l'AFC



titres et d'abord pour assembler les différentes parties qui constituent la Sierra du nord au sud, au nombre de quatre.

Lorsqu'on désire revenir sur le choix des trames, il n'est que d'utiliser la possibilité qu'offre le programme de recommencer depuis le début toutes les opérations que nous avons sommairement décrites :

- choix d'une palette appropriée de 32 trames parmi 100 ;
- consultation des résultats de l'analyse des données ;
- affectation de nouvelles trames à tout ou partie des classes des unités territoriales considérées.

CONCLUSION

Du fait de son caractère conversationnel, le logiciel Mac SAIF permet de remodeler à volonté les trois étapes principales du travail, d'abord choix et découpage en modalités des variables retenues grâce au programme "**zrang**", ensuite choix de la partition adoptée en conjuguant les critères statistiques et critères proprement cartographiques, enfin choix des trames utilisées sur la carte, créées et présentées à l'écran par le programme '**carthage**'.

REFERENCES BIBLIOGRAPHIQUES

- J.-P. & F. BENZECRI : *Notices d'utilisation des programmes du logiciel Mac-SAIF* : [NOT. CORR. CAH], [NOT. CRE TAB.], [NOT. PROG. CART.], in C.A.D. Vol XIV n°1, pp.7-72,, 1989.
- J.-P. BENZECRI : *D'Aristote à l'analyse multidimensionnelle : logique et ontologie dans l'induction statistique* [LOG.STAT.]; in C.A.D. Vol XIII N°3, pp.325-338, 1988 ; (dans cet article est publiée la carte de la consommation du gazole en France, due à A.Moussaoui).
- J.-P. BENZECRI : *Où vit-on le mieux en France ? représentation cartographique de l'analyse de 33 variables* [CARTE VIE FRANCE]; in C.A.D. Vol XIII n°3, pp.349-366, 1988.
- M.-M. THOMASSIN : *Cartographie automatique fondée sur l'Analyse des Données ; une application sur micro-ordinateur du programme 'carthage' à la régionalisation de la consommation des produits pétroliers en France (1972-1981)*; in *Bulletin du Comité Français de Cartographie*, fascicule n°115, (Mars 1988). 8e Assemblée Générale de l'A.C.I., 13e Conférence internationale Morelia (Mexique)-12-21 Octobre 87.
- M.-M. THOMASSIN : *La Sierra équatorienne. Régionalisation agricole et Analyse des Données*. Vol. 1, Texte, 626 p ; Vol. 2, Cartographie : 28 cartes et plans, en couleurs et en noir et blanc. Collection Travaux et Documents. Editions de l'ORSTOM, Paris. 1988.

Fig.5
Fond de carte de la Sierra équatorienne

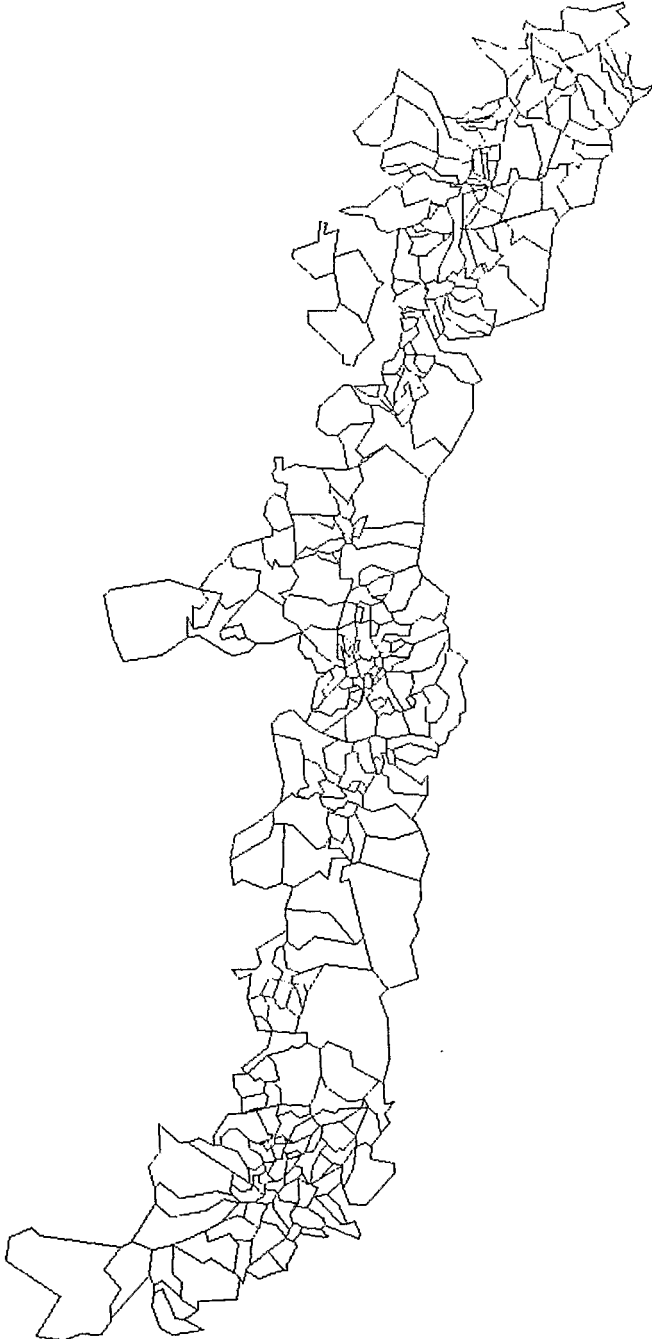


Fig. 6
Légende de la carte "Dominantes culturales-Elevage"
Typologie des paroisses expliquées par la typologie des modalités, d'après le Vacor

Classe n°	Caractéristique (s) *	Description
460	(O72 B72) +++ (IR1 FA1) ++	: Zones d'élevage intensif bovin et ovin. prairies irriguées modérément développées
456	(O7) FA2 IR2) +++	: Zones d'élevage ovin intensif avec prairies et cultures irriguées
458	(B70 CR2)+++ (MZ1 LFO OVI1)+(LFI CR1 MZ2)+	: Zones céréalières spécialisées avec élevage bovin ; culture du maïs modérée
465	(CR) MZ0) ++++++	: Zones céréalières très spécialisées. absence de maïs.
459	(IR) FA3) ++++++	: Zones d'élevage laitier intensif avec prairies irriguées très développées.
454	LF2) ++++++	: Zones spécialisées dans les cultures maraichères et fruitières.
461	(MZ) FA0) ++++++	: Zones maizicoles spécialisées. pas de pâturages
449	LFO +++ (CR0 FT0) ++	: Très peu de superficies irriguées. de céréales et de pommes de terre
462	(O74 B74) ++++ (MZ1 LFO OVI1)	: Zones d'agriculture maizicole avec élevage bovin et ovin intensif
448	TR2) ++++++	: Zones de cultures fruitières ou maraichères tropicales très spécialisées
464	(B71 G70 TRI) ++++++	: Zones de cultures fruitières ou maraichères moins développées que ci-dessus (448). avec élevage laitier modéré. peu d'ovins

* de la classe de variables d'après le Vacor, plus le nombre de + est élevé. plus les caractéristiques sont marquées.

Arbre de classification des paroisses

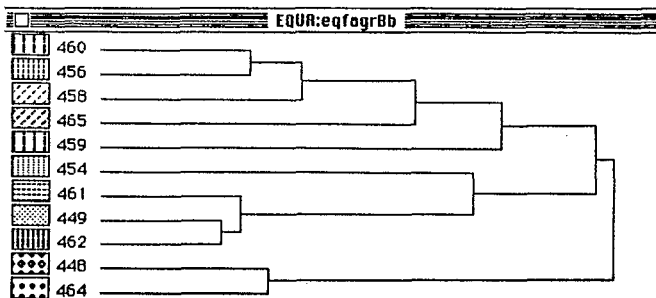
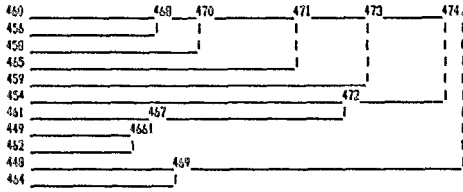
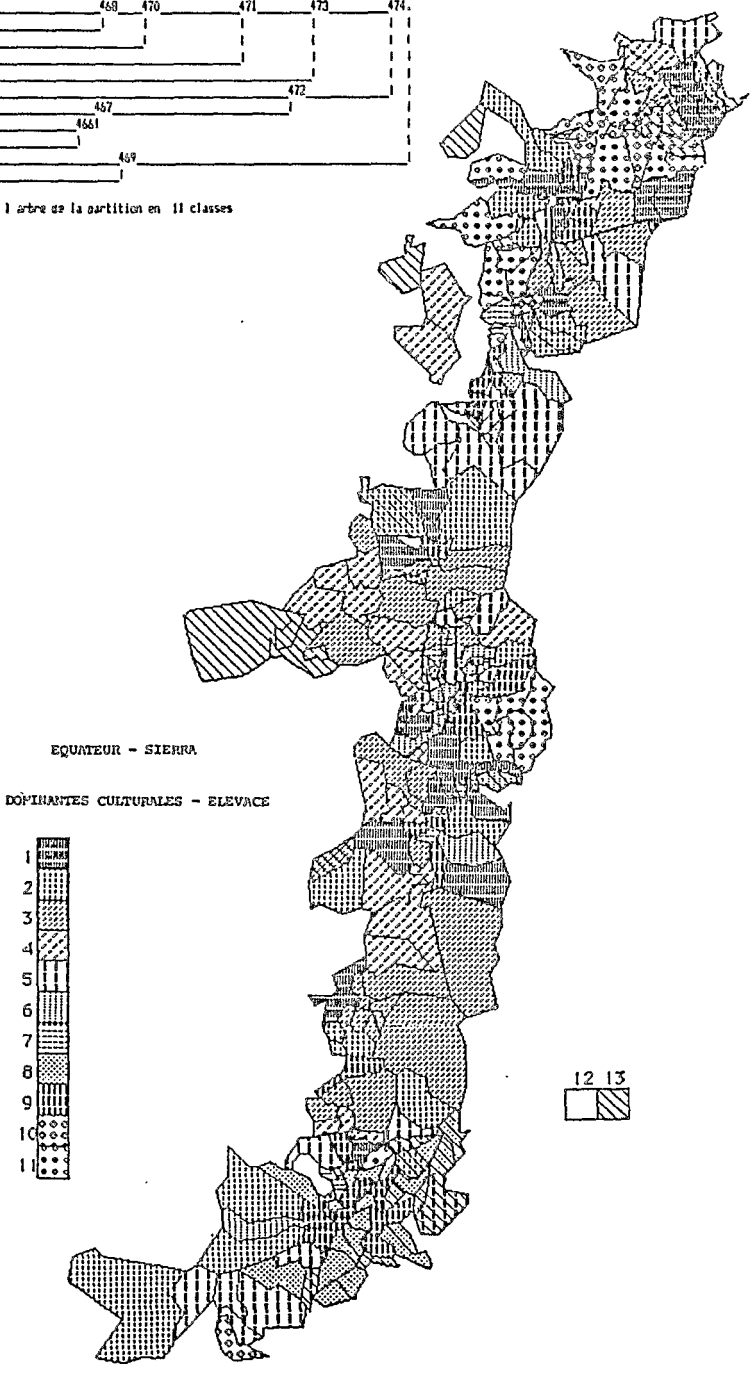


Fig.7

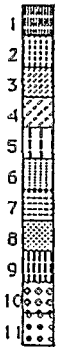


ci dessus l arbre de la partition en 11 classes



EQUINEUR - SIERRA

DOMINANTES CULTURALES - ELEVACE



ANNEXE

Variables principales codées par 'Dcodx'

EQUATEUR agriculture élevage

77:EQagrDcodx: bornes pour le découpage des variables

le nombre des variables est 9

27 a 5 modalités dont les signes et bornes sont

BV0 BV1 BV2 BV3 BV4

0 2 5 8 10

28 a 5 modalités dont les signes et bornes sont

OU0 OU1 OU2 OU3 OU4

0 2 5 8 10

29 a 4 modalités dont les signes et bornes sont

PA0 PA1 PA2 PA3

0 3 7 10

30 a 4 modalités dont les signes et bornes sont

MZ0 MZ1 MZ2 MZ3

0 3 7 10

31 a 4 modalités dont les signes et bornes sont

CR0 CR1 CR2 CR3

0 3 7 10

32 a 2 modalités dont les signes et bornes sont

PT0 PT1

0 3

33 a 3 modalités dont les signes et bornes sont

LF0 LF1 LF2

0 3 7

34 a 3 modalités dont les signes et bornes sont

TR0 TR1 TR2

0 3 7

35 a 4 modalités dont les signes et bornes sont

IR0 IR1 IR2 IR3

0 3 7 10