

J.-P. BENZÉCRI

F. BENZÉCRI

État des recherches sur la cartographie automatique des données statistiques

Les cahiers de l'analyse des données, tome 14, n° 4 (1989),
p. 449-466

http://www.numdam.org/item?id=CAD_1989__14_4_449_0

© Les cahiers de l'analyse des données, Dunod, 1989, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

ÉTAT DES RECHERCHES SUR LA CARTOGRAPHIE AUTOMATIQUE DES DONNÉES STATISTIQUES

[STAT. CART.]

J.-P. & F. BENZÉCRI

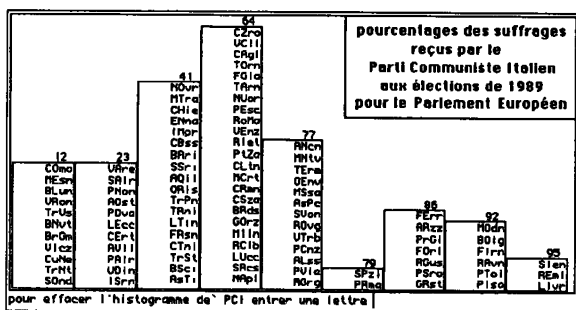
Nous considérerons successivement les thèmes traités jusqu'ici en cartographie automatique avec le logiciel MacSAIF, et les perfectionnements apportés aux programmes depuis la publication de la notice [NOT. PROG. CARTE]; nous terminerons en montrant sur un exemple comment peut être menée à bien une étude complexe.

1 Thèmes des études de cartographie automatique

Chaque fois que, dans un tableau de données, l'ensemble des individus est un ensemble d'unités territoriales, par exemple l'ensemble des départements français, toute partition de cet ensemble, construite par le statisticien, peut être utilement représentée sur une carte, en affectant à chaque classe de la partition une trame qui lui soit propre.

L'exemple le plus simple est sans doute celui la partition qu'on réalise, d'après une seule variable v , quand on construit un histogramme. Si, par exemple, l'histogramme a 9 créneaux, les unités territoriales sont réparties en 9 classes, dont chacune contient les unités pour lesquelles v est compris dans l'un des 9 sous-intervalles en lesquels on a divisé l'intervalle total de variation de v . Il est naturel d'affecter à ces classes des trames d'autant plus foncées que la valeur de v y est plus élevée: c'est ce qu'on peut faire avec le programme 'zrang', pour toute variable d'un tableau de données, lorsqu'on utilise l'option 'Z', "observer les données sans créer de tableau", à condition qu'ait été créé un fichier de fond de carte (cf. [NOT. CRÉ. TAB.], §2.1, in *CAD*, Vol XIV, n°1, 1989).

Sur une carte, apparaissent, aux yeux de tous, des cohérences spatiales que seul un géographe à l'imagination précise perçoit sur un histogramme ou une colonne de tableau. Ainsi, la carte ci-jointe, extraite d'un travail, en préparation, de M. Maravalle, montre la concentration spatiale des circonscriptions où est maxima l'influence du *Parti Communiste Italien*.



La confrontation de cartes analogues afférentes aux principaux partis donne une vue d'ensemble de l'équilibre des forces politiques dans le pays.

Toutefois, c'est par l'analyse multidimensionnelle qu'apparaît le mieux la structure d'un ensemble de variables. De ce point de vue, la cartographie automatique met en relief les résultats obtenus par l'analyse des données, et, réciproquement, l'analyse des données devient la méthode de base de la cartographie thématique. Pour rester dans le domaine de l'analyse des votes, nous citerons d'abord la cartographie des votes par départements aux premiers

tours des élections présidentielles françaises de 1981 et 1988; étude complétée par l'analyse des votes au référendum calédonien (in *CAD*, Vol XIII, n°3 et Vol XIV, n°2); puis, dans un tout autre domaine, l'analyse, par E. Dimara, des productions agricoles de la Grèce par régions (in *CAD*, vol XIV, n° 2).

L'analyse des correspondances, préparant la classification ascendante hiérarchique, permet de traiter non seulement un tableau de contingence homogène (tel qu'un tableau dénombrant des votes, comme ci-dessus; ou le tableau des consommations mensuelles, par département, d'un ou plusieurs produits pétroliers: cf. *CAD*, Vol XII, n°2, et [NOT. CRÉ. TAB.], déjà cité); mais un tableau hétérogène dont, par 'zrang', chaque colonne, considérée comme donnant la réponse à une question, est éclatée suivant plusieurs modalités: c'est ainsi qu'on a procédé pour représenter sur une carte un ensemble de 33 variables, collationnées par l'hebdomadaire *Le POINT*, afin de déterminer où vit-on le mieux en France (cf. [CARTE VIE FRANCE], in *CAD*, Vol XIII, n°3, 1988).

Il va sans dire que, tandis que la carte, par elle-même, permet d'apprécier la cohérence spatiale des classes d'une partition, seuls les instruments usuels d'aide à l'interprétation (notamment la représentation sur les plans issus de l'analyse factorielle du nuage des centres des classes, et les listages Facor et Vacor, cf. [NOT. CORR. CAH], in *CAD*, Vol XIV, n°1) donnent, en terme de variables, la signification de ces classes.

Dans une certaine mesure, le choix judicieux des trames peut offrir, de cette interprétation, un rappel mnémorique: c'est ce qu'aide à faire le programme 'carthage', en affichant sur un même plan (issu de l'analyse factorielle) les variables (ou les centres de classes de variables) et les centres de classes d'unités territoriales avec, dans un caisson rectangulaire la trame adoptée pour la classe, trame qui peut être ensuite modifiée.

Pour énumérer tous les domaines qui offrent matière à la cartographie automatique, il faudrait citer tous les tableaux de données imaginables dont l'ensemble des individus soit un ensemble d'unités territoriales; tâche impossible, et d'ailleurs inutile, une fois admis que la cartographie est le complément obligé de toute classification effectuée sur un tel ensemble. Mais il vaut la peine de citer encore un exemple, celui des tableaux de flux, qui présente cette particularité que les variables elles-mêmes, comme les individus, sont de nature spatiale.

Considérons, avec A. Aït Hamlat (travail en cours), les flux interdépartementaux de prestations d'une entreprise dont les agents, domiciliés chacun dans un département déterminé, peuvent toutefois servir des clients d'autres départements. Pour caractériser, de ce point de vue, un département *d*, il faut le mettre en correspondance avec l'ensemble *D* des départements *d'*

considérés deux fois: d'une part, *id'*, comme importateur à partir de *d*, et, d'autre part, *xd'*, comme exportateur vers *d*; avec:

$$k(d, id') = \text{importations de } d' \text{ à partir de } d ;$$

$$k(d, xd') = \text{exportations de } d' \text{ vers } d .$$

L'analyse d'un tel tableau met en évidence des régions au sein desquelles les échanges sont actifs; et délimite les zones d'attraction des principales métropoles.

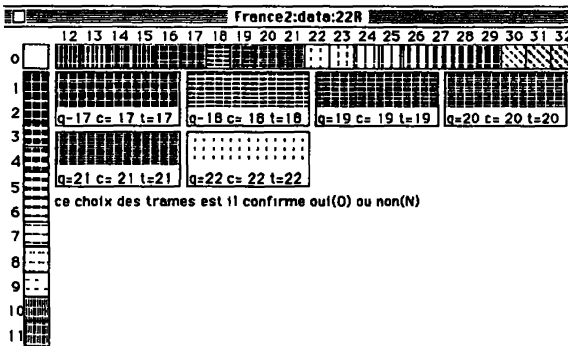
2 Perfectionnements apportés au logiciel de cartographie

Ces perfectionnements, suggérés par les études d'une année, concernent d'une part le programme 'carthage' (et , accessoirement 'zrang') qui crée des cartes à partir de données; et d'autre part le programme 'litexcarte', qui crée, pour les divers pays ou autres territoires subdivisés en unités, des fichiers numériques de "fond de carte". Comme il s'agit de modifications à des programmes dont l'état antérieur est déjà expliqué dans [NOT. PROG. CART.], nous supposons que le lecteur a accès à cette notice.

2.1 État du programme 'carthage'

2.1.1 Nombre maximum de classes de la partition

La version présentée par [NOT. PROG. CART.] limite à 16 le nombre des classes: ainsi peuvent s'afficher simultanément, sur l'écran du Macintosh, les caissons rectangulaires qui présentent les trames attribuées à chacune des classes en offrant de les modifier; et l'arbre de la hiérarchie des classes (présenté, lui aussi avec les trames) tient sur un seul écran. Ce nombre de 16 nous paraissait d'ailleurs suffire, dans la mesure où fractionner le territoire en un grand nombre de classes tend à supprimer toute image de cohérence spatiale.



Cependant, nous étant proposé de réaliser une carte de la partition en 22 Régions de programme de l'ensemble des départements français, nous avons dû dépasser ce nombre limite de 16. Il a suffi pour cela de permettre que le tableau

des caissons, ou l'arbre hiérarchique, s'affiche sur deux écrans, comme le fait la carte elle-même pour laquelle on passe alternativement du Nord au Sud. Ainsi, la limite a été portée de 16 à 32. Simultanément, il a fallu porter de 29 à 32 le nombre des trames sélectionnées pour constituer la palette qui s'affiche dans des petits caissons carrés, autour des grands caissons rectangulaires des classes, lors du dialogue d'affectation des trames aux classes (cf. [NOT. PROG. CART.], p. 63).

La figure présente la fin du dialogue de modification des trames pour la partition de la France en 22 régions: on y voit les caissons des régions 17 à 22, les caissons des 16 premières régions occupant l'écran précédent (non publié).

2.1.2 Extension Nord-Sud et extension Est-Ouest

Afficher le territoire complet sur deux écrans en passant alternativement du Nord au Sud, n'est possible que si le rapport de la dimension verticale à la dimension horizontale ne dépasse guère 2. Or ce rapport est nettement dépassé par la carte d'Italie présentée au §2, et plus encore par la carte de la *Sierra* équatorienne considérée au §3.

Il a donc fallu modifier 'carthage' afin qu'il puisse présenter une carte sur un nombre quelconque d'écrans se succédant du Nord au Sud et du Sud au Nord: par exemple, l'affichage de l'Italie, Sicile comprise, se fait sur 3 écrans. Le même perfectionnement a été introduit dans 'zrang' pour la présentation de carte associée à l'affichage conversationnel des histogrammes (cf. §1).

L'utilisateur dispose encore d'une autre voie qui est de cadrer la carte en laissant, à gauche ou à droite de l'écran, une bande non utilisée pour le tracé. Ceci est obtenu très simplement en introduisant dans le fichier 'stx' (qui fournit à 'litxcarte' les coordonnées des sommets pour le fond de carte; cf. §2.2) un point supplémentaire, écarté vers la gauche ou vers la droite, et n'intervenant dans le périmètre d'aucune unité territoriale (fichier 'utx'): en effet, 'litxcarte' prend en compte ce point dans le calcul des valeurs Max et min de x servant au cadrage.

2.1.3 Éléments supplémentaires surimpression et vérification de cohérence

Les statistiques françaises étaient, jusqu'à une date récente, ventilées en comptant l'*Île de beauté* pour un département unique numéroté 20. Mais il y a désormais deux départements distincts, *Hauts de Corse* et *Corse du Sud*, pour lesquels on publie des statistiques séparées. Il faut donc prévoir pour la France deux fonds de carte, selon que cette division nouvelle est ou non prise en compte.

Cependant, même si les deux nouveaux départements (qui existaient en fait jadis sous les noms de *Golo* et de *Liamone*) ont reçu les numéros 201 et 202, le numérotage des autres départements métropolitains subsiste, de 1 à 95, avec

seulement une lacune à 20. Il est clair que cette incohérence dans la numérotation risque d'engendrer des erreurs dans la compilation automatique des statistiques. Il nous a donc paru préférable de laisser subsister dans le nouveau fond de carte, appelé "France2", l'ancien département de la Corse, avec son numéro 20, les nouveaux départements étant placés aux rangs 96 et 97.

Quand donc, on analyse un tableau de données distinguant entre Nord et Sud de l'île, il faut prévoir dans ce tableau, à la 20-ème ligne, une ligne de cumul, ou éventuellement une simple ligne de zéro, qui sera traitée en supplémentaire par le programme 'qori'.

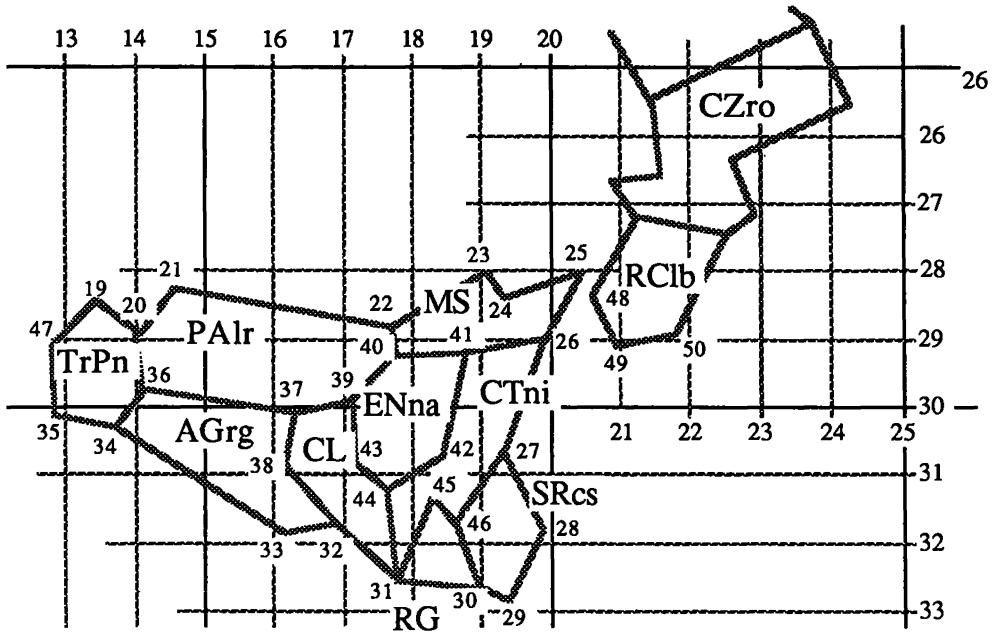
Que fait cependant 'carthage' avec ces départements Cors, Hcrs et SCrs, dont les deux derniers recouvrent le premier? Il les imprime tous les trois, selon les instructions qui lui sont données. En fait, si comme il est logique, soit le tout, soit les deux parties sont en supplémentaire, il n'y a pas de confusion à redouter de ces impressions superposées, car 'carthage' laisse en blanc les éléments supplémentaires. Il faut toutefois noter au passage que cette surimpression n'allait pas de soi; et qu'il a fallu choisir, selon les normes du Macintosh, le mode d'écriture approprié par l'instruction: penmode(pator).

Il a fallu encore prévoir que l'utilisateur pourrait se tromper de fond de carte, et répondre pour nom du pays, "France" et non "France2"! On a donc introduit en tête de 'carthage' une vérification de concordance du nombre des unités territoriales, entre fond de carte et tableau de données.

On signalera l'intérêt de la surimpression: si, dans le fond de carte, un groupe d'unités territoriales figure deux fois, on peut mettre chacune des unités de ce groupe dans deux classes différentes de la partition représentée: les unités apparaissant alors peintes de deux trames croisées. Ceci permet notamment de rendre compte d'une évolution diachronique (comme l'a fait - en créant des trames combinées - E. Dimara; cf. [AGRI. GREC.], in *CAD*, Vol XIV, n°2, p. 237). Un autre d'exemple d'utilisation, pour introduire en les signalant des éléments supplémentaires, sera donné au §3.

2.2 État du programme 'litxcarte' de création du fond de carte

Nous rappelons (cf. [NOT. PROG. CART.], §1) que 'litxcarte', utilise deux fichiers de texte, dont les suffixes sont 'stx' et 'utx' donnant respectivement les coordonnées des sommets relevés sur les contours des unités territoriales, et la suite des numéros des sommets du schéma polygonal attribué à chaque unité. Dans sa version initiale, 'litxcarte' procède en deux étapes: lecture et vérification des coordonnées (fichier 'stx'); puis lecture et vérification des suites de numéros ('utx') avec création du fichier numérique de fond de carte (suffixe 'reg'). La version présente ajoute à ces deux étapes nécessaires deux étapes facultatives: vérification de la forme des unités territoriales affichées une par une; et affichage de la carte dans sa totalité. Nous décrivons ces deux étapes, après avoir préalablement commenté la création des fichiers 'stx' et 'utx'.



2.2.1 Relevé des points et de leurs coordonnées sur la carte

Voici comment on a procédé pour créer le fond de carte pour 'Italia' (cf. §1). Une carte d'Italie, s'inscrivant dans un rectangle (30cm × 33cm), a été d'abord décalquée en schématisant sous forme polygonale, les côtes, frontières et limites internes des 95 unités territoriales (appelées '*Provincia*') qui sont l'analogue des départements de la France. Sur chaque unité, on a écrit son sigle en deux capitales, en usage pour l'immatriculation des automobiles, en le complétant, pour plus de clarté, de deux minuscules (dans la mesure de l'espace disponible): e.g. AGRg, pour *Agrigente*; SRcs, pour *Siracusa*; etc...

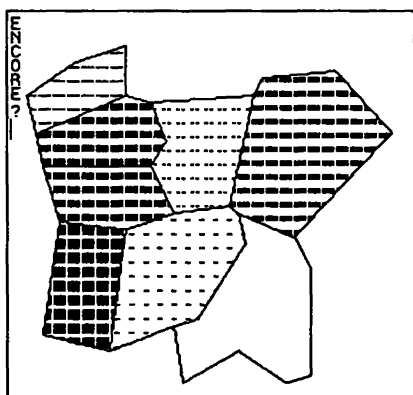
On a ensuite photocopié le calque, et tracé sur la copie un quadrillage centimétrique, couvrant seulement les zones utiles, et où ne sont appuyées que les lignes de 5 en 5 cm. Suivant les normes des procédures de dessin du Macintosh (Quickdraw), l'abscisse est comptée à partir du haut vers le bas, et l'ordonnée de la gauche vers la droite.

Avant de commencer à relever les coordonnées des sommets, on a, grâce à l'éditeur de texte 'EDIT', ouvert les fichiers de texte Italiautx et Italiastx. Sur Italiautx, on a, après la ligne de titre usuelle (cf. [NOT. PROG. CART.] §1.2) préparé 95 lignes commençant chacune par l'un des numéros de 01 à 95 suivi du sigle de la *provincia* correspondante, de AGRgente à VTrho: on sait en effet, que

dans ce listage, tout ce qui n'est pas chiffre est considéré par 'litxcarte' comme un simple séparateur. Sur Italiastx, on a seulement écrit le titre.

Le numérotage et le relevé des sommets s'est fait ensuite par étapes. D'abord, 18 sommets pour la Sardaigne; puis les sommets 19 à 47 pour la Sicile, et ainsi de suite. Le schéma publié se place au début de la troisième étape: Italie méridionale. Compte tenu de la précision désirée ici, les coordonnées des sommets peuvent être lues à vue, en millimètres sur le quadrillage centimétrique. Ce qu'il faut craindre, ce n'est pas un décalage de 1mm, mais des erreurs majeures qui, comme on verra, déforment la carte!

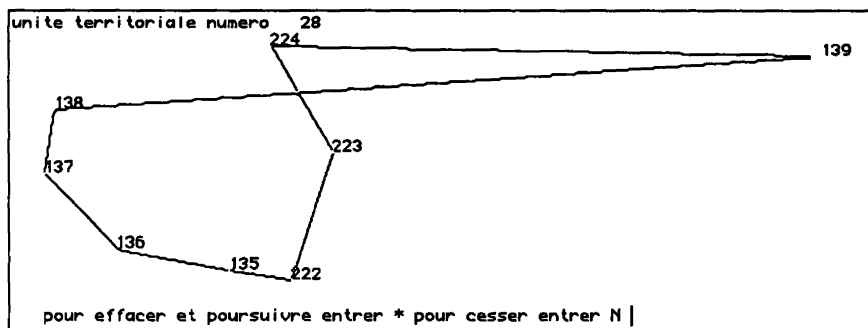
À chaque étape, on écrit effectivement sur le listage 'stx' les coordonnées des points, sur des lignes numérotées consécutives; puis l'on porte la suite des sommets du périmètre des *province* déjà complètes sur les lignes du listage 'utx' préparées à cet effet. En travaillant par étape, on se garde de donner à deux sommets différents le même numéro; erreur qu'on ne peut être absolument sûr d'éviter qu'en marquant les numéros des sommets, non directement sur la carte, mais sur des pastilles numérotées au préalable et que l'on colle au fur et à mesure.



Le cartographe impatient peut réaliser graphiquement un état partiel de la carte. Il suffit d'écrire sur les lignes du listage 'utx' afférentes aux unités dont les sommets n'ont pas encore été mesurés:

3s 1 1 1 ,

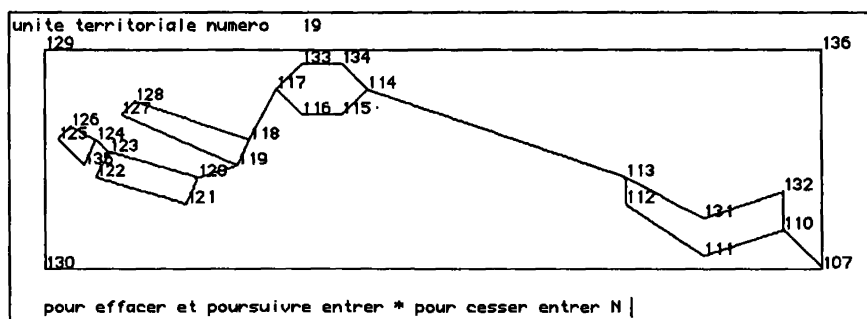
comme si chacune de ces unités était réduite à un triangle dont les trois sommets sont le même point numéroté '1'. Le programme 'litxcarte' accepte en leur état provisoire les fichiers 'stx' et 'utx' et construit une carte partielle telle que celle, reproduite ici, qui donne le Nord du Portugal, avec 29 sommets mesurés et 7 unités déjà complétées (sur les 20 que compte le pays).



2.2.2 Vérification de la forme des unités territoriales

Le programme 'litxcarte' ouvre d'abord le fichier 'stx' des coordonnées des sommets et en vérifie la cohérence; dans la présente version, l'interprétation d'éventuels messages d'erreur est facilitée par l'affichage continu, à l'écran, des numéros des sommets avec les coordonnées de ceux-ci.

Si 'stx' est satisfaisant, on passe au fichier 'utx' qui donne les numéros des sommets sur les contours de chaque unité. Dans la présente version, 'litxcarte' propose de vérifier la forme des unités territoriales, avant de passer à la création du fichier numérique 'reg' du fond de carte. Si cette option est choisie, l'utilisateur voit les unités s'afficher à l'écran une par une, avec les numéros de leurs sommets: il peut ainsi déceler des pointes, ou parfois des croisements dus à une erreur grossière affectant une coordonnée. D'une telle erreur, la *provincia* de CuNeo, unité n° 28 de l'Italie, offre un exemple: il est manifeste que le point 139 a été fortement décalé vers la droite.



```

19 Açores      42s 107 110 111 112 113 114 115 116 117 118
                119 120 121 122 123 124 135 125 126 124
                123 120 119 127 128 118 117 133 134 114
                113 131 132 110 107 130 129 136 107 136
                129 130

```

Nous profiterons de cette option pour montrer comment on peut, dans le dessin du fond carte, introduire un cadre séparant un archipel du reste du pays.

L'exemple choisi est celui d'un schéma de l'archipel des Açores, joint à la carte du Portugal. En suivant les sommets dans l'ordre où les énumère l'extrait du fichier 'Portugalutx' publié ici, on part du coin '107' du cadre pour entourer toutes les îles; (une telle méthode avait déjà été utilisée par E. Dimara pour les îles de la Grèce); puis, une fois revenu à '107', on fait une boucle parcourant deux fois le cadre en sens opposés.

2.2.3 Affichage de l'ensemble de la carte

Si les fichiers 'stx' et 'utx' sont tous deux corrects quant au format (quelles que puissent être, d'autre part, les erreurs du tracé) 'litxcarte' crée le fichier numérique 'reg' du fond de carte. Il propose ensuite à l'utilisateur d'afficher l'ensemble de la carte soit en blanc (contours seuls) soit avec des hachures; avec la possibilité d'alterner les deux formes d'affichage plusieurs fois de suite.

Une telle vérification ne pouvait précédemment se faire qu'en passant par le programme 'carthage'; soit en utilisant les résultats de l'analyse d'un tableau de données relatif au pays considéré; soit en créant par 'litxicqi' (cf. [NOT. PROG. CART.] §3) le fichier d'une partition arbitraire.

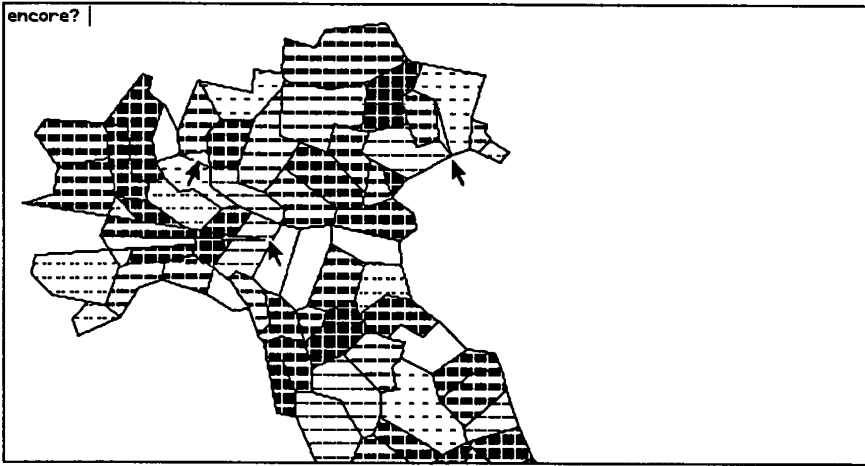
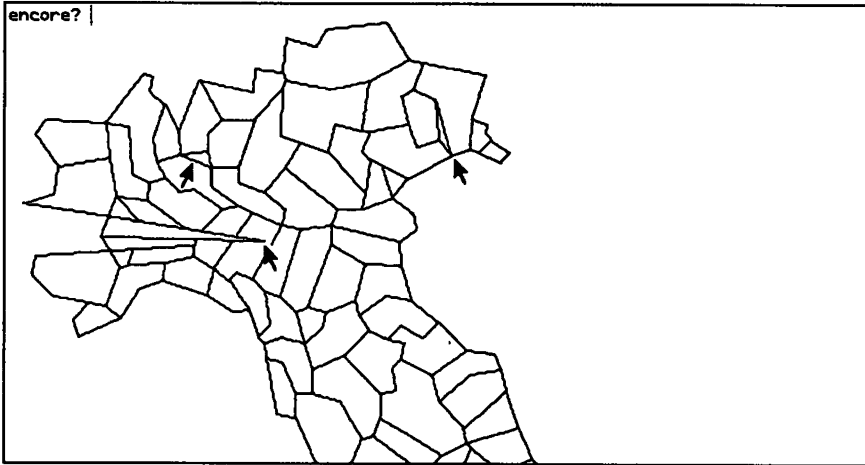
On peut ainsi déceler des erreurs qui n'apparaissent pas sur les unités affichées isolément. Mais les superpositions de traits, et l'absence des numéros des sommets rendent parfois plus difficile l'interprétation des erreurs; en sorte que les deux options de vérification ont des rôles complémentaires.

Prenons l'exemple des erreurs commises en créant le fond de carte pour l'Italie. On a déjà vu que le décalage du point 139 (en principe situé sur la frontière du Piémont avec la France) affectait la province de CuNeo; est également affectée la province limitrophe de ToRino (Turin). Il s'agit d'une erreur sur le fichier 'stx'.

Mais deux autres erreurs apparaissent, non décelables sur les unités territoriales isolées, à moins de comparer minutieusement l'image de l'écran à celle de la carte sur papier. En notant sur le listage 'utx' les sommets consécutifs d'UDine, le sommet 265 a été oublié. De même, 283 manque au contour de la *provincia* de Milano.

De telles omissions se voient le mieux sur la carte couverte de hachures où elles créent des triangles blancs. Sur la carte en blanc des contours seuls, ces triangles se remarquent aussi parfois comme des épaississements du trait (cas d'UDine). Mais la carte en blanc est surtout utile pour discerner les pointes intempêtes créées par les erreurs grossières de coordonnées.

Les deux autres erreurs commises en créant la carte d'Italie ne perturbaient pas la topologie du maillage des unités territoriales. Affectant les coordonnées de



deux points, elles déplaçaient seulement des frontières intérieures; et déformaient suffisamment certaines unités pour ne pas rester inaperçues. Ces erreurs concernant le centre de l'Italie, elles ne sont pas visibles sur le présent graphique.

Nous signalerons pour conclure que si l'utilisateur demande plusieurs fois de suite l'affichage de la carte couverte de hachures, celles-ci sont modifiées, ce qui peut aider à distinguer les erreurs affectant des zones d'abord couvertes d'une trame trop sombre pour permettre de distinguer les replis erronés des contours.

3 Reprise d'une étude antérieure: l'agriculture de la Sierra équatorienne

On trouve dans [ANA. CART.] (in *CAD*, Vol X, n°1, pp. 25-51, 1985) un extrait des travaux de cartographie multidimensionnelle, consacrés par M.-M. Thomassin à la Sierra équatorienne. L'article est accompagné d'une carte hors-texte, réalisée par l'Institut Géographique National. Sans prétendre rivaliser avec des dessinateurs et des imprimeurs maîtres de leur art, nous voulons montrer que de tels travaux ne sont pas inabordables pour notre logiciel.

3.1 Structure des données et enchaînement des analyses

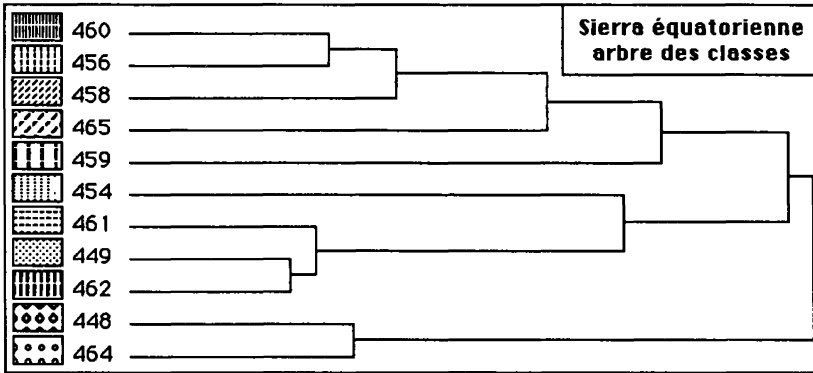
La Sierra équatorienne est divisée en quelques centaines d'unités territoriales, de très inégale étendue, appelées "paroisses". Une enquête de l'ORSTOM a permis de recueillir pour ces paroisses un ensemble de 48 variables réparties en trois blocs: milieu physique, utilisation du sol, niveau socio-économique. En fait, les données ne sont complètes que pour 238 paroisses; tandis que, pour 42 autres, on a des données lacunaires mais intéressantes.

Dans le présent §, nous nous bornerons à considérer les 9 variables du bloc 2: utilisation du sol. Celles-ci constituent un tableau de base 280×9 , noté 'eqagr'; dans ce tableau, les données manquantes sont simplement remplacées par un nombre convenu très élevé: 10000. Le programme 'manq' (dont l'usage est expliqué dans [ÉQ. PERS.], §2, in *CAD*, Vol XIV, n°3) fournit un listage, par lignes et par colonnes, des données manquantes.

Comme il est désormais classique, les variables ont été découpées en classes. On sait que le programme 'zrang' permet à l'utilisateur de choisir les bornes des classes au vu d'histogrammes, et crée un tableau éclaté (en 0,1): noté 'eqagrBb'; tableau où, à chaque variable, correspond un bloc comptant autant de colonnes que la variable a de modalités. Il faut seulement prendre garde aux données manquantes: on peut leur réserver une modalité dans chaque variable, ou (ce qui a été fait ici) tolérer qu'elles se rangent dans la modalité supérieure, quitte à s'en souvenir dans la suite.

Afin de faciliter la comparaison avec les analyses portant sur les autres blocs du questionnaire, on a gardé exclusivement comme lignes principales les 238 paroisses ne présentant pas de lacune sur l'ensemble de l'enquête. Parmi les 42 autres, 36 n'ont pas de lacunes dans le bloc 2, analysé ici; 6 seulement ne pourront être considérées.

On soumet au programme 'qori' le tableau 'eqagrBb', en spécifiant 42 lignes supplémentaires; 'qori' produit, outre un listage d'analyse des correspondances, des fichiers numériques de facteurs pour les trois ensembles i , is et j : 'eqagrBbiFacww', 'eqagrBbisFacww' et 'eqagrBbjFacww'. Ces fichiers



servent au programme 'CAH' pour effectuer des classifications hiérarchiques sur les ensembles i et j .

Cependant, en même temps que le listage Facor d'aide à l'interprétation, le programme 'CAH' crée un fichier de facteurs pour les centres des classes: 'eqagrBbiqFacww'. Grâce à ce fichier, on peut, par le programme 'discr', rattacher chaque individu supplémentaire au centre de la classe dont il est le plus proche; les résultats étant consignés dans le listage 'eqagrBbaffis->iq' (dans le titre duquel les lettres 'aff' rappellent qu'il s'agit d'affectation: cf. [NOT. CRÉ. TAB.], §6).

3.2 Cartographie pour l'ensemble des paroisses

Mise à part la réelle difficulté que présente la création d'un fichier de fond de carte pour 280 paroisses très inégales de taille et aux contours parfois imbriqués, la cartographie des éléments principaux peut se faire sans qu'il soit besoin d'innover (sinon pour le défilement du Nord au Sud, cf. *supra* §2.1.2): une partition des 238 paroisses principales en 11 classes ayant été retenue, chaque classe a reçu une trame propre; tandis que restaient en blanc les 42 autres paroisses (mises en éléments supplémentaires dans l'analyse du tableau 'eqagrBb'). L'arbre des classes, avec les trames choisies, est publié ici.

La difficulté de la présente étude vient de ce que l'on a voulu figurer sur la carte, couvertes de trames appropriées, les paroisses supplémentaires; ou, du moins, les 36 paroisses pour lesquelles on a, dans le tableau 'eqagr' des données complètes. Chacune de ces paroisses ayant été, par 'discr', rattachée à une classe de paroisses principales, il était naturel de lui donner la trame de cette classe; mais, d'autre part, on désirait signaler, par des hachures parallèles superposées à la trame, qu'il s'agissait d'éléments supplémentaires.

Voici donc comment on a procédé. On a créé un fichier de fond de carte, 'eq2reg', comprenant 322 unités territoriales: les 238 principales, puis le bloc de 42 autres répété deux fois: il a suffi pour cela de modifier sur traitement de texte

	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
0	[Pattern]																				
1	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
2	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
3	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
4	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
5	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
6	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
7	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
8	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
9	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
10	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
11	[Pattern]					[Pattern]					[Pattern]					[Pattern]					
	ce choix des trames est il confirme oui(O) ou non(N)																				Sierra équatorienne choix des trames

le fichier 'utx' créé pour les 280 paroisses, sans rien changer au fichier 'stx' des coordonnées des sommets. On a ensuite créé un fichier 'eqagr2Bbicqkitx' spécifiant pour les 322 unités la partition suivante:

11 classes comprenant les 238 paroisses principales, avec, de plus, dans chaque classe, les paroisses supplémentaires (de 239 à 280) qui lui sont affectées; sous réserve que celles-ci comportent des données complètes dans le bloc 2;

une douzième classe, formée des 6 paroisses supplémentaires (de 239 à 280) pour lesquelles les données manquent;

une treizième classe, formée des unités 281 à 322, ou paroisses supplémentaires répétées une deuxième fois.

Comme on le voit sur le tableau du choix des trames, les 11 premières classes ont reçu les trames déjà choisies pour elles sur la carte restreinte aux 238 unités principales. Aux autres classes, on a affecté respectivement le blanc et des hachures parallèles continues: ainsi, sur la carte, toute paroisse supplémentaire est couverte de hachures parallèles; si, de plus, elle est du nombre des 36 pour lesquelles les données sont complètes dans le bloc 2, elle reçoit aussi la trame de la classe au centre de laquelle 'discri' l'a rattachée. Il est clair que la trame de hachures continues a été introduite dans le tableau des trames puis dans la palette en créant un fichier '\$tr', comme on l'explique dans [NOT. PROG. CART.], en *addendum* au §2.

Reste à dire comment a été créé commodément le fichier 'eqagr2Bbicqkitx' décrivant la partition des 322 unités territoriales en 13 classes. On est parti pour cela du fichier 'eqagrBbicqki' décrivant la partition des 238 paroisses principales, ce fichier étant relu sous forme de listage grâce au programme

```
eq: eqagr2Bb$tr
trame pour blanc $0 $0
trame pour barrer les éléments suppl
$8040201 $80402010;
```

```
eq: eqagrBbicqkitx
 11 20460 10026 1 111 110 126 125 124 123
 121 120 109 122 119 129 127 128 99 50
 116 117 118 204 203 199 202 200 201 20456
10022 2 4 3 229 230 234 236 235 237
 238 233 231 232 222 220 221 226 228 227
 223 224 225 20458 10034 .....
```

'relifent'. Puis on a inséré dans chaque classe les éléments qui lui revenaient selon le listage 'affis->iq', dont nous publions un extrait (annoté en donnant en plus des sigles des paroisses leurs numéros): par exemple, dans la première classe, i460, on a inséré les éléments {242, 247, 261, 262}.

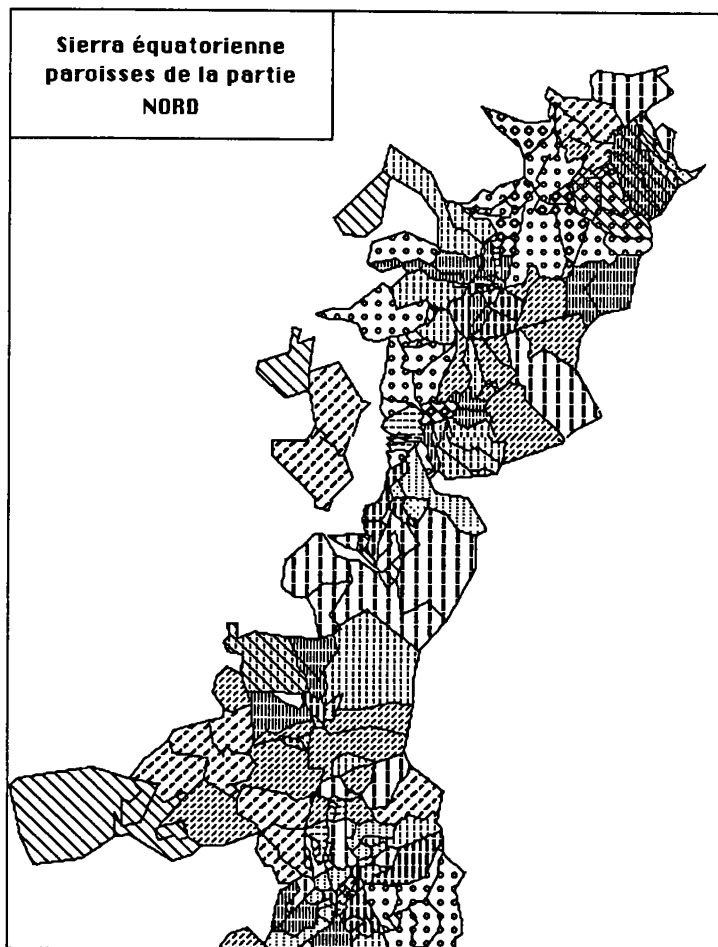
```
eq: eqagrBbaffis->iq ; nombre de facteurs utilisés = 8
affectation des is aux iq
.....
liste des is affectés à i460
obh* occ* baa* bae*
242 247 261 262
liste des is affectés à i456
oco* adb* beh*
244 255 271
.....
```

Ici, il convient de rappeler (cf. [NOT. PROG. CART.], §3) que le nombre en 10000 (qui suit le nombre en 20000 donnant le numéro attribué à chaque

```
eq: eqagr2Bbicqkitx (eqagrBb modifié avec elsup et hachures)
 13 20460 10030 1 111 110 126 125 124 123
 121 120 109 122 119 129 127 128 99 50
 116 117 118 204 203 199 202 200 201
 242 247 261 262 20456
10025 2 4 3 229 230 234 236 235 237
 238 233 231 232 220 221 226 228 227
 223 224 225 244 255 271 .....
.....
20555 10006 250 275 276 277 278 279
20666 10042
 281 282 283 284 285 286 287 288 289
 290 291 292 293 294 295 296 297 298 299
 300 301 302 303 304 305 306 307 308 309
 310 311 312 313 314 315 316 317 318 319
 320 321 322
```

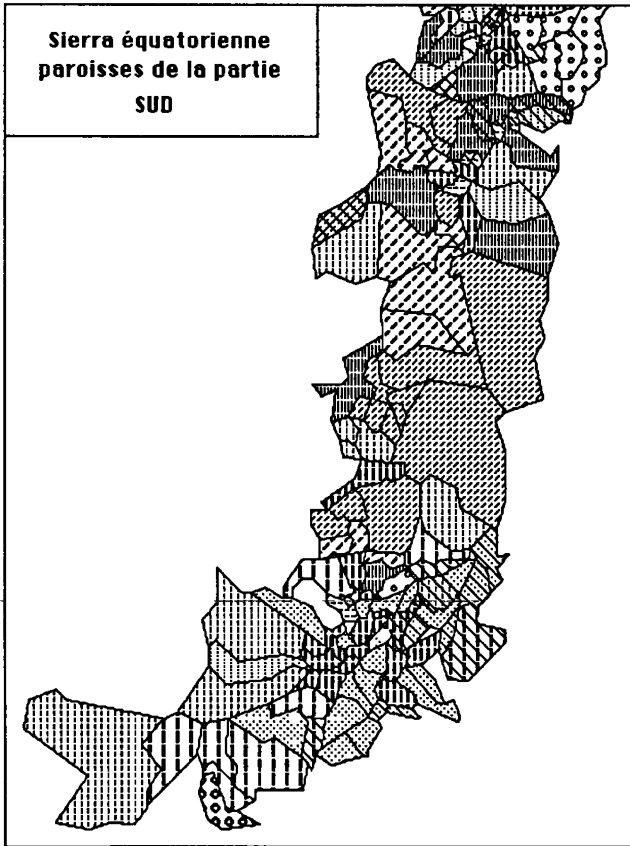
classe) sert à définir le cardinal de la classe (il vaut 10000 + cardq) et doit donc être modifié dans 'eqagr2Bbicqkitx' relativement à 'eqagrBbicqkitx': ainsi, le cardinal de la classe i460 passant de 26 à 30, on a, dans le listage modifié: 20460 10030 ..., au lieu de: 20460 10026...

Quant aux deux dernières classes (douzième et treizième, auxquelles on a attribué arbitrairement les numéros '555' et '666'), il a été aisé d'en énumérer



les éléments, particulièrement pour la dernière qui contient simplement tous les numéros de 281 à 322.

Malgré l'échelle réduite qui ne peut convenir aux plus petites paroisses, la carte donne une vue d'ensemble de la distribution spatiale des divers types d'agricultures propres à chaque classe; caractéristiques que donne succinctement un arbre étiqueté par les modalités (de 0 ou 1 = nul ou faible, à 3 ou 4 = fort) des variables: BoVins, OVins, PAturages, MaiZ, CéRéales, Pommes deTerre, Légumes-Fruits, cultures TRopicales, cultures IRriguées. Une mention telle que OV0++++ doit se lire: "absence caractéristique d'ovins"; tandis que



{IR1 PA1}++									
460	—{OV2 BV2}+++	—468	—470	—471	—473	—474			
456	—{OV3 PA2 IR2}+++ —								
458	—{BV0 CR2}++++								
465	—{CR3 MZ0}++++++								
459	—BV3+ {IR3 PA3}++++++								
454	—{F2}++++++						472		
461	—{MZ3 PA0}++++++	—467							
449	IR0+++ {CR0 PT0}++ —466								
462	—{OV4 BV4}++++ —								
448	—TR2++++++	—469							
464	—{BV1 OVO TR1}+++++ —								

l'association de IR3++++++, “très fort pourcentage (>70%) de cultures irriguées par rapport à la superficie utilisée”, avec PA3++++++, “très fort pourcentage de pâturage”, implique “prairies irriguées”, et, avec BV3+, caractérise la classe i459, des zones d'élevage laitier.

Bibliographie

Outre les nombreux articles, parus dans *CAD*, cités dans le cours de l'article nous mentionnerons:

M.-M. Thomassin: Cartographie automatique fondée sur l'Analyse des Données; in *Bulletin du Comité Français de Cartographie*, fascicule n°115, Mars 1988.

ejusdem: La Sierra équatorienne. Régionalisation agricole et Analyse des Données. Collection *Travaux et Documents*, Éditions de l'ORSTOM, Paris, 1988.

ejusdem: Communications aux Conférences Internationales de Cartographie de Morelia, Mexique (1987) et Budapest, Hongrie (1989).