

Exemples simulés et réels (archéologie et économie) de la recherche d'une sériation

0 - FICHE TECHNIQUE.

A - Nombreux tableaux d'incidence simulés de dimensions très diverses : 25×50, 25×250, 100×1000 et 250×25000.

Tableau d'incidence de données en archéologie de dimension 26×59, tableau appelé MERO ci-dessous.

Tableau de mesures numériques en économie de dimension 20×26.

B - Mise en évidence d'une *sériation* sous-jacente au tableau des données. Cette mise en évidence est effectuée à partir de représentations euclidiennes de l'ensemble des lignes (resp. colonnes) du tableau. Ce type de représentation permet de déterminer une permutation des lignes (resp. colonnes) mettant en évidence, dans le cas des tableaux d'incidence, une forme "diagonale" significativement chargée de "1". Dans les figures suivantes la case comportant 1 est noircie.

C - L'indice de proximité entre lignes (resp. colonnes) du tableau d'incidence se met sous la forme

$$S(\ell, k) = [s(\ell, k) - n(\ell)n(k)/n] / \sqrt{(n(\ell)n(k)/n)}, \quad (1)$$

où $s(\ell, k)$ est le nombre d'associations positives entre les lignes (resp. colonnes) ℓ et k qui représentent des vecteurs logiques ; $n(\ell)$ (resp. $n(k)$) est le nombre de composantes égales à 1 de ℓ (resp. k) et n est le nombre total de composantes.

Dans le cas d'un tableau de mesures numériques ; variables x objets de dimension $m \times n$, tel que celui qui concernera l'exemple en économie. Si x_{ij} désigne la mesure brute de la j -ème variable sur le i -ème objet ; soit X_{ij} la mesure "centrée réduite" par rapport à la distribution de la j -ème variable :

$$X_{ij} = (x_{ij} - \mu_j) / \sqrt{\sigma_j^2 / n}, \quad (2)$$

où μ_j et σ_j^2 sont respectivement la moyenne et la variance de la distribu-

tion de la j -ème variable sur l'ensemble des objets défini par l'échantillon. Dans ces conditions, on a la même expression formelle de l'indice de proximité entre variables ou entre objets :

- entre variables :

$$Q(\ell, k) = \sqrt{n} \rho(\ell, k) = \sum_{1 \leq i \leq n} X_{i\ell} X_{ik} , \quad (3)$$

où ρ est le coefficient de corrélation ;

- entre objets :

$$P(r, s) = \sum_{1 \leq j \leq m} X_{rj} X_{sj} . \quad (4)$$

Ces indices sont sous-jacents à l'Analyse en Composantes Principales Normée.

D—Algorithmes de représentations euclidiennes de MPATS (Méthode des Pôles d'Attraction utilisant les Similarités, critère de la variance), de ACP (Analyse en Composantes Principales Normée) et de AFC (Analyse Factorielle des Correspondances). Algorithme de classification de MPAGD (Méthode des Pôles d'Agrégation utilisant les Distances) (cf. Chap. 6 et 8).

E—Tester l'aptitude de l'algorithme de représentation euclidienne de MPATS pour mettre en évidence une sériation sous-jacente à un tableau de données dans des cas simulés dans le cadre d'une hypothèse probabiliste, supposant une sériation plus ou moins nette et dans des cas réels. La validité des résultats de l'algorithme est notamment établie par rapport aux résultats de ceux classiques de l'ACP et de l'AFC.

F—H. LERREDDE.

G—P. PERIN pour l'exemple en Archéologie et J.-F. OUTREVILLE pour l'exemple en Economie.

La suite des titres des sous paragraphes est

A. SÉRIATIONS SIMULÉES.

- A.1. Introduction.
- A.2. Présentation des résultats.
- A.3. Conclusion.

B. EXEMPLE EN ARCHEOLOGIE MEROVINGIENNE.

- B.1. Introduction.
- B.2. Les traitements graphiques manuels.
- B.3. Le traitement par l'analyse des correspondances.
- B.4. L'analyse par MPATS.

C. EXEMPLE EN ECONOMIE.

- C.1. Introduction ; présentation des données.
- C.2. Résultats comparés de MPATS, ACP et MPAGD.

A. SERIATION SIMULEES.

A.1. Introduction.

Pour évaluer l'aptitude de la représentation euclidienne (cf. Chap. 8 § V.4), fournie dans le cadre du programme MPATS, à restituer autour des deux premiers pôles une sériation qui, sans être stricte, domine statistiquement ; nous avons commencé par travailler avec des tableaux d'incidence simulant une forme σ "floue" (cf. Chap. 8 § VII). Rappelons qu'un tel tableau d'incidence à m lignes et n colonnes est la réalisation d'un élément aléatoire de $\{0,1\}^{m \times n}$, muni d'une mesure de probabilité qui résulte du produit des mn aléas indépendants $\{\varepsilon'_{\ell j} / 1 \leq \ell \leq m, 1 \leq j \leq n\}$ sur $\{0,1\}$ où $\Pr\{\varepsilon'_{\ell j} = 1\} = p$ (resp. q) selon que la position (ℓ, j) tombe à l'intérieur du bloc diagonal (parallélogrammatique en général) (resp. à l'extérieur). On suppose évidemment $p \gg q$; la grandeur de p (resp. petitesse de q) indiquent la "netteté" statistique avec laquelle on peut espérer voir apparaître la forme σ après simulation.

Nous avons mené de nombreuses expériences en variant les paramètres du tableau d'incidence aléatoire :

Relativement à une taille $m \times n = 25 \times 50$ où l'analyse a été le plus poussée, on a considéré

- . une forme σ parallélogrammatique fortement enchaînée (cf. Chap. 8, § III) avec trois degrés de flou respectivement définis par $(p = 0,9 ; q = 0,1)$, $(p = 0,8 ; q = 0,2)$ et $(p = 0,6 ; q = 0,15)$;

- . une forme parallélogrammatique σ faiblement enchaînée (cf. Chap. 8, § III) avec également les trois degrés de flou qu'on vient de préciser : $(p = 0,9 ; q = 0,1)$, $(p = 0,8 ; q = 0,2)$ et $(p = 0,6 ; q = 0,15)$;

- . une forme σ trapézoïdale avec toujours les mêmes degrés de flou : $(p = 0,9 ; q = 0,1)$, $(p = 0,8 ; q = 0,2)$ et $(p = 0,6 ; q = 0,15)$. Ici, le nombre de composantes par ligne, égales à 1 avec la probabilité p , va en augmentant de la première à la dernière ligne.

Nous avons voulu ensuite étudier l'influence du nombre de colonnes et nous avons considéré, avec différents degrés de flou, une forme σ dans le cadre d'un tableau de taille 25×250 .

Nous avons enfin envisagé, toujours avec différents degrés de flou, une forme σ dans le cadre de "grands" tableaux de tailles respectives 50×500 , 100×1000 et 250×2500 ; ce cadre nous permettant notamment de tester les aspects informatiques des algorithmes proposés.

Signalons pour terminer qu'à titre technique, nous avons considéré la situation où on intervertit les rôles des densités de chargement p et q où alors le bloc parallélogrammatique σ apparaît "creux".

Pour chacun des cas étudiés nous avons simulé une réalisation du tableau d'incidence aléatoire où la forme σ apparaît. Nous introduisons alors une permutation des lignes et des colonnes du tableau rendant strictement insoupçonnable à l'oeil nu la forme σ sous-jacente. C'est ainsi que le tableau entre en machine pour être traité. On applique le programme de représentation euclidienne autour des deux premiers pôles de MPATS utilisant la variance à l'ensemble L des lignes.

A partir de la représentation graphique de L , on déduit un ordre ou permutation sur L . A partir de cette permutation, un programme REORD permet de ranger les éléments de C afin de dégager, au niveau du tableau d'incidence la forme diagonale σ .

Le tableau ainsi réordonné qui doit laisser réapparaître la forme σ , est comparé avec le tableau initial.

On a d'autre part considéré l'ordonnement des lignes et colonnes du tableau à partir des représentations de L et de C sur le premier plan factoriel de l'Analyse des Correspondances ; méthode apte à dégager clairement des sériations, mais qui met en oeuvre des algorithmes beaucoup plus élaborés que ceux que nous proposons, comprenant notamment la réduction d'un endomorphisme symétrique.

Nous nous limiterons ici à présenter un seul exemple de tableau simulé ; celui correspondant à une forme parallélogrammatique, faiblement enchaînée, où les paramètres de chargement sont respectivement $p = 0,9$ et $q = 0,1$.

Pour une analyse complète on se reportera à [1] et pour un aperçu général des résultats, à la conclusion de ce paragraphe.

A.2. Présentation des résultats.

La figure 1 reproduit la forme parallélogrammatique σ , faiblement enchaînée, obtenue par simulation avec les probabilités de chargement $p = 0,9$ et $q = 0,1$. Rappelons que l'hypothèse définissant le tableau d'incidence aléatoire, peut être considérée comme celle de la liaison la plus faible compatible, avec la configuration σ fixée et les densités de chargement la délimitant.

La figure 2 montre le même tableau après une permutation des lignes (resp. colonnes) perturbant complètement la forme diagonale significativement chargée de 1. C'est rappelons-le sous cette forme que le tableau est traité.

La figure 3 montre le tableau réordonné où la reconstitution de l'ordre sur les lignes est directement fournie par la représentation euclidienne autour des deux premiers pôles de L , obtenue par MPATS. Les points, représentant les éléments de L , se répartissent sur le graphique le long d'une courbe en forme d'"s" (lettre s manuscrite minuscule). Une telle répartition aurait été parfaite si la forme σ était nette ; c'est-à-dire, si $p = 1$ et $q = 0$; mais, compte tenu du flou, des "zig-zag" apparaissent le long de la deuxième direction entraînée par le second pôle d'attraction. L'expérience montre que pour retrouver l'ordre sur L , il y a lieu de parcourir la suite des points à partir de l'extrémité de la forme "s" dans le sens ↷.

La figure 4 montre l'ordre de parcours de la succession des points ayant permis de retrouver la forme σ de la figure 3 où on notera avec intérêt un résultat "meilleur" que le tableau simulé (figure 1) ; nous voulons dire, où la partie diagonale est davantage chargée.

La figure 5 montre la représentation de l'ensemble C des colonnes fournie par MPATS. On constate une répartition des points autour d'un cercle, sans doute révélatrice d'une sériation ; toutefois, il est difficile de savoir où est le point origine de la lecture de la répartition des points.

L'analyse des correspondances fait apparaître une répartition des points autour d'une parabole aussi bien pour la représentation de L que pour celle de C (cf. figures 6 et 7). La représentation simultanée de L et de C permet dans une certaine mesure de montrer la dépendance entre ligne et colonne dans la sériation à mettre en évidence.

A.3. Conclusion.

Relativement à l'ensemble des expériences que nous avons conduites sur les tableaux d'incidence simulés, il ressort clairement que

a) la représentation de L par MPATS est d'autant plus nette et plus étirée le long d'une forme \nearrow que la forme σ est plus nette (i.e. p grand et q petit). Il en résulte naturellement une plus grande aisance à préciser l'ordre sur L définissant la sériation. Dans le cas où le degré de flou augmente (i.e. p diminue et q augmente avec $p > q$), la représentation a tendance à se tasser autour de l'origine : point de rencontre des deux axes.

b) La sériation se dégage plus nettement lorsqu'il s'agit d'une forme σ fortement enchaînée ; d'autre part, le premier et le second pôle correspondent respectivement davantage à une ligne extrême et à une ligne centrale du tableau ordonné. Cette constatation est tout à fait conforme à l'analyse théorique (cf. Chap. 8, § IV et V).

c) Le nombre important de colonnes par rapport au nombre de lignes rend la répartition des points représentant L plus régulière et plus facile à suivre. Pour un nombre de colonnes devenant de plus en plus important la forme \nearrow , de représentation de L, devient \wedge .

d) L'analyse factorielle des correspondances étale davantage la représentation du nuage associé à L et permet de dégager la sériation en étant moins sensible aux lignes (resp. colonnes) faiblement chargées de uns. Il s'agit certes là d'une plus grande performance technique ; mais il ne faut pas oublier dans ce cas la complexité des algorithmes utilisés qui comprennent notamment une diagonalisation de matrices. Au contraire notre algorithme est étonnant dans sa simplicité et sa très grande rapidité : la partie essentielle du calcul réside dans la formation de la matrice des proximités entre éléments de L. Cette plus grande robustesse par rapport à l'importance de la charge peut être simplement due à la métrique du χ^2 de l'analyse des correspondances. Or, on a établi (cf. Chap. 2, § V) un indice de corrélation conforme à cette métrique que nous sommes en train de tester dans le cadre de notre problème. De toute façon, cette qualité de plus grande robustesse peut ne pas être sans inconvénient dans le cas de données réelles où ces lignes faiblement chargées peuvent correspondre à des attributs de description secondaires que le spécialiste veut voir écartés a posteriori. C'est ce qui se produira dans le cas de l'exemple réel suivant.

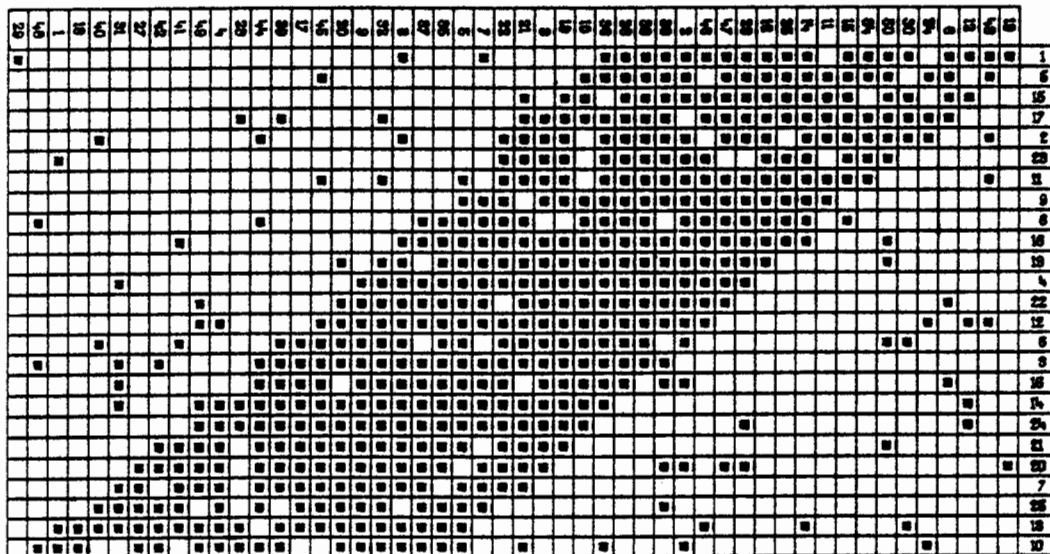


Fig. 1

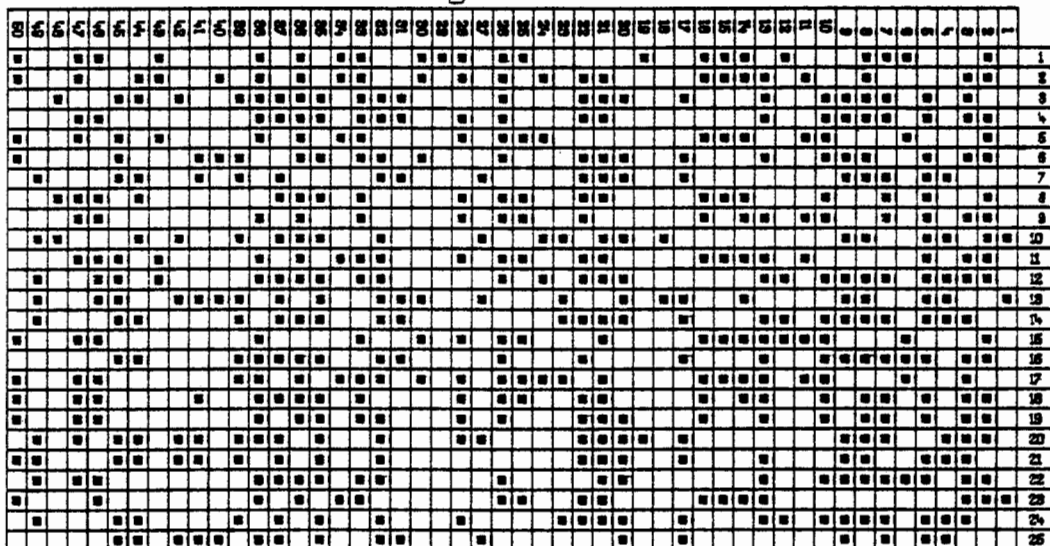


Fig. 2

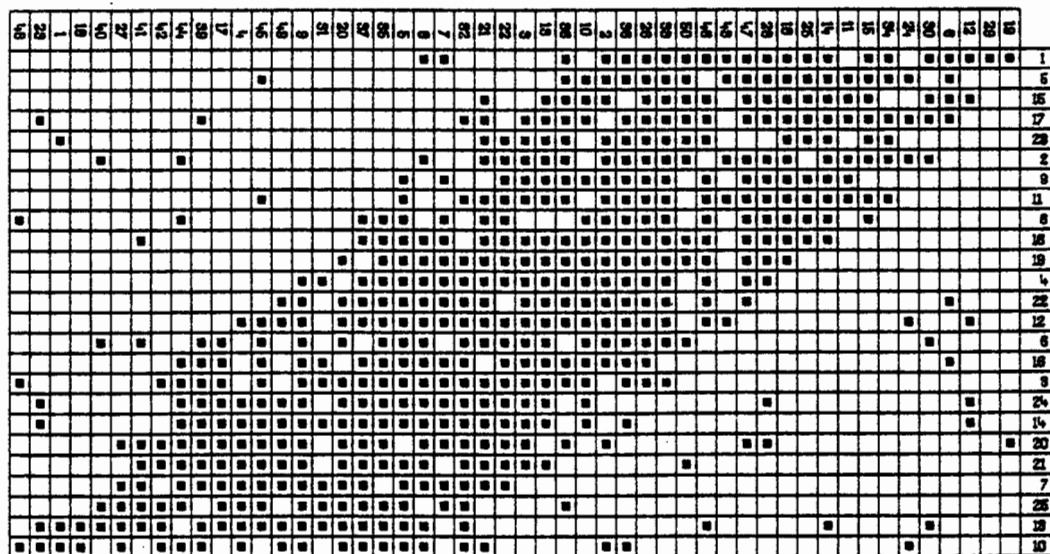


Fig. 3.

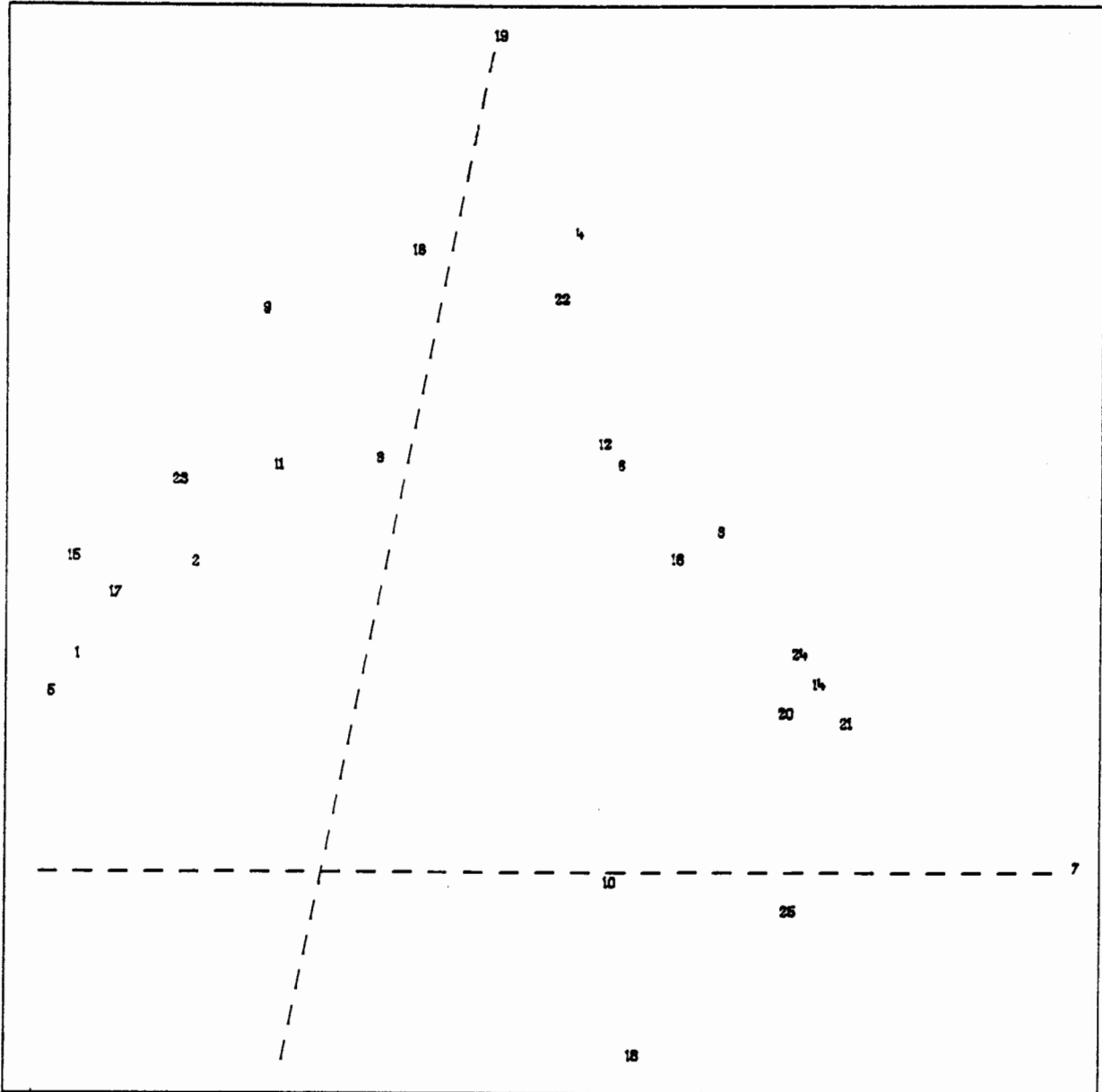


Figure 4

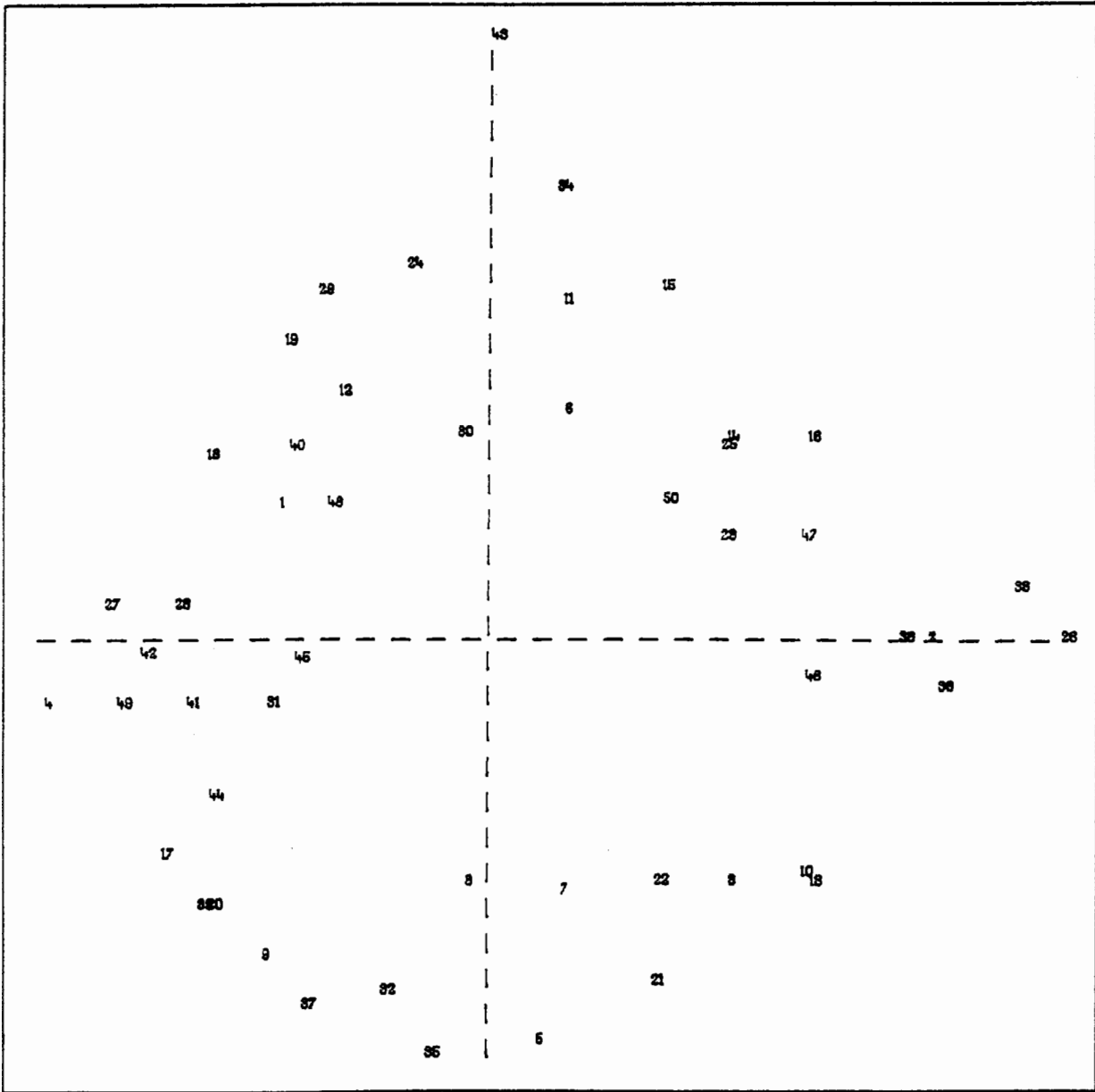


Figure 5

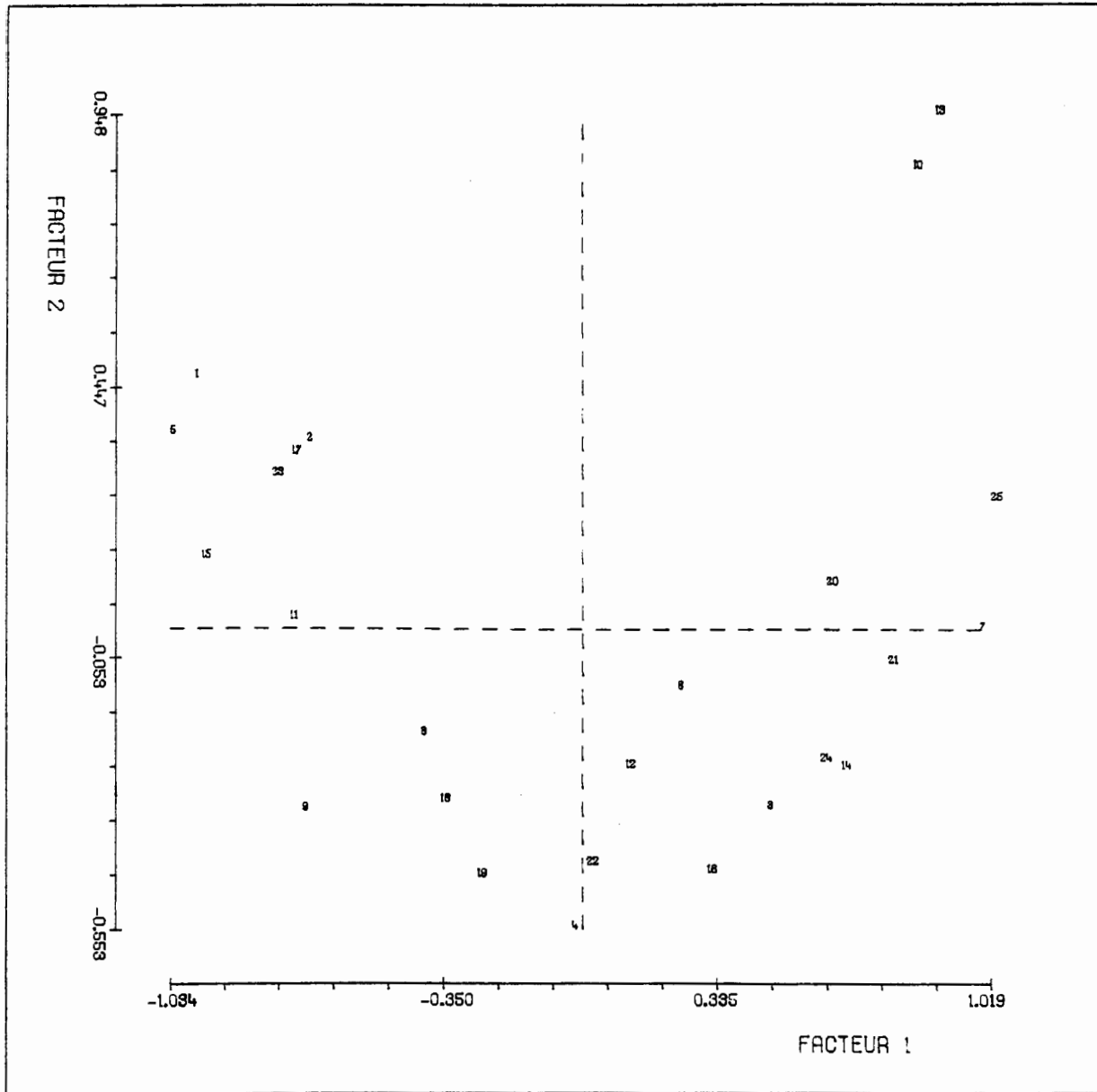


Figure 6

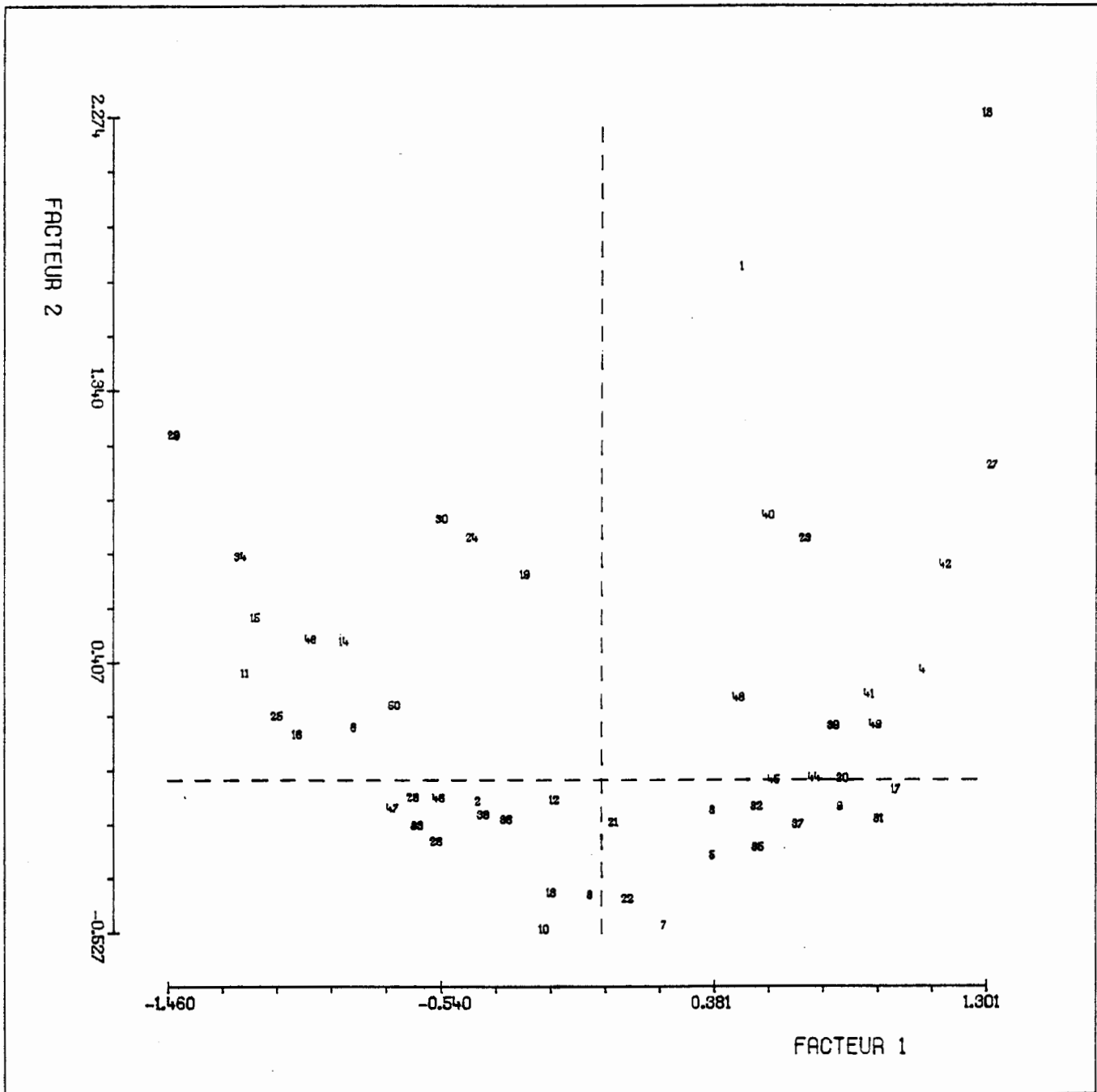


Fig. 7.

B. EXEMPLE EN ARCHEOLOGIE MEROVINGIENNE.

B.1. Introduction.

Les données concernant cet exemple nous ont été fournies par P. Périn qui dirige le séminaire d'archéologie Mérovingienne à la 4ème section de l'E.H.S.S. . Il s'agit d'étudier un ensemble de plaques-boucles de ceinture damasquinées du nord-est de la France s'échelonnant entre la fin du VIe siècle et le début du VIIIe siècle. Sur cet ensemble de 59 plaques, le spécialiste a noté la présence ou l'absence de 26 attributs de technique de fabrication, de forme et de décor notés de A à Z (cf. figure 8).

Ce tableau a d'abord été traité manuellement par permutations de lignes et colonnes au Laboratoire Graphique de J. Bertin. C'est après ces manipulations graphiques que nous nous sommes engagés dans le traitement automatique de ces données.

Le problème posé était d'étudier si on peut mettre en évidence, au delà d'une simple typologie, une évolution dans les techniques de fabrication de ces plaques et savoir si une telle évolution est conforme à la chronologie.

Pour tenter de répondre au problème posé, nous allons analyser les résultats des différents traitements, manuels ou automatiques, qui ont été opérés sur ces données.

B.2. Les traitements graphiques manuels.

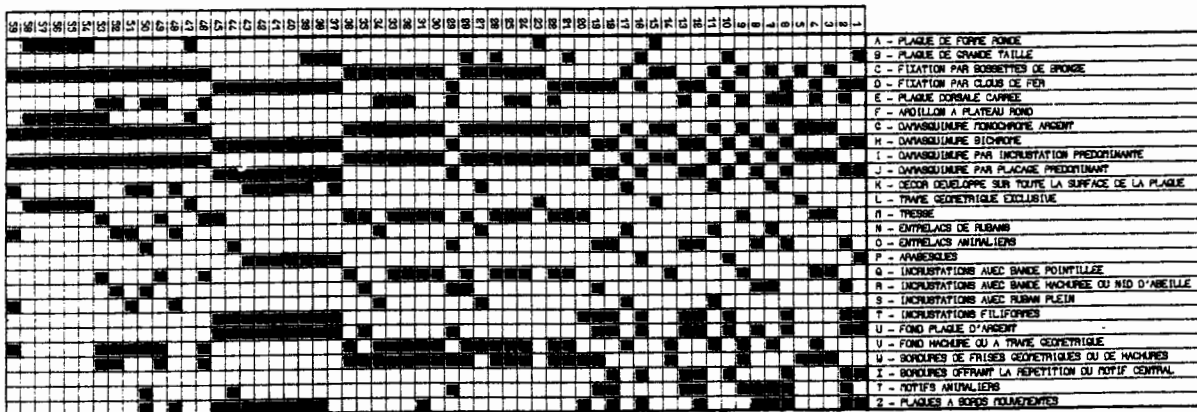


Figure 8

Le résultat du traitement tableau [MERO] met nettement en évidence un phénomène sériel (cf. figure 9) où apparaissent cependant très clairement deux grands sous ensembles de critères et de plaques.

La typologie délimitée par le chercheur (cf. figure 10) comprend les groupes de plaques suivants :

Groupe A : damasquinure monochrome (incrustation prédominante, absence presque totale de motifs animaliers, bossettes de fixation en laiton prédominantes.

Ce groupe ou classe se compose de :

A_1 : plaque ronde, trame géométrique exclusive imitant le cloisonné.

A_2 : Plaque triangulaire étroite, trame géométrique exclusive (comme pour A_1). La classe contient exactement un élément mais il s'agit d'un type bien représenté ailleurs.

A_3 : Tresse centrale (le plus souvent faite de rubans pointillés) avec encadrement de hachures, plaques triangulaires ou trapézoïdales,...

A_4 : Entrelacs de rubans pleins (le plus souvent sur toute la surface).

Groupe B : damasquinure bichrome (placage argent prédominant), fixation par clous de fer prédominants, bords mouvementés très fréquents.

Cette classe se compose de

B_1 : plaque longue et étroite avec tasse simple.

B_2 : Motif central d'entrelacs animaliers

B_{2-1} : bordures géométriques

B_{2-2} : bordures reprenant le décor animalier.

B_3 : Motif d'arabesques filiforme

B_{3-1} : sur plaque étroite (toute la surface)

B_{3-2} : sur de très grandes plaques

B_{3-2-1} : sur toute la plaque

B_{3-2-2} : par compartiments.

Cette typologie, fonction du type de décor, de l'organisation de ce décor et éventuellement de la morphologie du support a été retenue pour l'étude et la publication de ces plaques (cf. [3]). Elle met en évidence une évolution technique qui s'accompagne d'une évolution chronologique tout à fait nette. En effet, les plaques rondes damasquinées monochromes à décor géométrique exclusif imitant le cloisonné (classe A_1) apparaissent vers la fin du Ve siècle et jusqu'à vers 600. Ensuite ce sont les plaques relevant de la technique monochrome de forme trapézoïdale large et plus ou moins rectangulaire (classe A_3 et A_4) qui s'échelonnent entre 600 et 650 avec des prolongements possibles. Enfin domine dans la deuxième moitié du VIIe siècle la mode des plaques à décor damasquiné bichrome avec

placage prédominant et incrustations en réserve (groupe B) se prolongeant jusqu'au début du VIII^e siècle.

B.3. Le traitement par l'analyse des correspondances.

La projection des points-variables ainsi d'ailleurs que des points-objets sur le premier plan factoriel qui explique 53,5 % d'inertie, met clairement en lumière un phénomène sériel (cf. figure 11). De plus, il est possible de distinguer dans ce phénomène évolutif deux groupes assez nets aussi bien sur l'ensemble des attributs que sur l'ensemble des objets.

La figure 12 montre le croisement de deux ordres : celui sur les attributs et celui sur les objets ; où on a spécifié les groupes d'objets.

Les commentaires de l'archéologue à propos de ce résultat sont les suivants :

"Groupe A. Les sous-classes A_1 et A_2 sont bien placés dans les zones 1 et 2. Les types A_3 et A_4 se retrouvent bien en zones 3 et 4 ; seule la plaque 34 est mal placée. La zone 5 est un mélange des types A_3 et A_4 regroupés ici en raison de critères "secondaires" qui peuvent les attirer dans cette zone ; pour bien faire, il faudrait ventiler la zone 5 entre les zones 3 et 4 comme suit : les plaques 26, 31, 9, 28, 4 et 22 dans A_3 et les plaques 48, 35, 21, 7, 50 et 20 dans A_4 . Soulignons cependant que ces deux types A_3 et A_4 sont contemporains et ont en commun certains attributs associés au groupe B. De plus, le passage du type A au type B vers les années 650 ne constitue pas une rupture brutale, mais une évolution progressive.

Groupe B. Nous retrouvons en zone 6 les plaques à motifs animaliers et entrelacs de rubans hachurés. La zone 7 met en évidence les plaques à arabesques développées sur toute la surface avec deux subdivisions : les plaques de taille normale (7a) et les plaques de grande taille (7b). Les plaques de la zone 8 s'apparentent avec celles de la zone 6, mais constituent un groupe à part. La zone 9 est constituée de plaques avec décor d'arabesques. Les plaques de très grande taille, pouvant comporter également un compartimentage du décor, devraient en réalité former un sous-groupe (7c) de la zone 7".

L'analyse des correspondances permet donc de mettre clairement en évidence le phénomène sériel sous-jacent au tableau [MERO] révélateur d'une double évolution technique et chronologique. On peut d'autre part mentionner ici l'intérêt de la représentation simultanée qui permet d'associer plaques et critères dans leur évolution. On peut enfin souligner que l'analyse factorielle a bien mieux que les manipulations graphiques manuelles, opéré implicitement une mise en forme diagonale de la matrice des données, tentant de prendre en compte tous les critères, même ceux "secondaires" (c'est ce qui apparaît sur le figure 12). Mais ceci est fait au détriment d'une distinction plus nette des groupes bien établis pour lesquels l'apparition de certains critères "secondaires" n'avait pas été prise en compte par le spécialiste (cf. figure 10).

B.4. L'analyse par MPATS.

Nous avons expliqué à ces données l'algorithme de représentation euclidienne autour des deux premiers pôles de MPATS, avec le critère de la va-

riance, dans le but de mettre en évidence la sériation et de comparer le résultat obtenu avec ceux des traitements précédents (manipulations graphiques manuelles et analyse factorielle des correspondances).

La figure 13 montre la projection des points-variables à partir de laquelle nous avons obtenu un premier ordre sur l'ensemble des attributs descriptifs par lecture directe de la projection de ces points sur le premier axe, parallèlement au second axe. A partir de cet ordre sur les attributs, nous avons déduit un ordre sur les plaques moyennant l'algorithme REORD construit à cette fin. C'est le résultat du croisement de ces deux permutations qui est visualisé dans la figure 14. Il nous est aussitôt apparu qu'il était souhaitable de déplacer les trois attributs L, A et F, sans que cela ait d'incidence sur l'organisation générale du tableau (cf. figure 15). La position de ces trois attributs, primitivement ainsi placés, s'explique par les effets de taille. Précisons tout de suite que les deux premières plaques 5 et 20 sont mal positionnées, mais c'est seulement dû à l'algorithme REORD.

C'est ce dernier tableau de la figure 15 que nous avons soumis au spécialiste, sans strictement rien lui dire des méthodes utilisées pour le former. Son avis est le suivant :

"Ce tableau est le plus satisfaisant au niveau de la lecture archéologique" ; on distingue les groupes et sous-groupes suivants (cf. figure 16) :

Groupe A : composé de A_1 et A_2 qui sont corrects ; A_3 et A_4 sont bien définis, mais les plaques 7 et 50 devraient se fondre dans A_4 et la plaque 20 dans A_3 .

Groupe B : alternance des critères "décors animaliers" et "arabesques". Tous les types définis à partir du tableau de la figure 10 se retrouvent ; à l'exception du groupe B_1 constitué des plaques 8 et 45 qui a éclaté. En explicitant :

- zone 1 - cas particulier emprunté à l'ex. B_1
- zone 2 - $B_{3.2.1}$.
- zone 3 - $B_{2.1}$.
- zone 4 - $B_{3.2.2}$.
- zone 5 - $B_{2.2}$.
- zone 6 - $B_{3.1}$.
- zone 7 - cas particulier emprunté à l'ex. B_1 .

Cette classification a permis de plus, de montrer un lien étroit entre le critère R "rubans hachurés" avec le type $B_{2.1}$. (zone 3) que le spécialiste n'avait pas vu sur le tableau de la figure 10 qu'il avait exploité.

Pour tenir compte des remarques de l'archéologue ; nous avons réorganisé ce dernier tableau en replaçant correctement les plaques 5 et 20 et en déplaçant le groupe A_4 pour le mettre derrière A_3 (cf. figure 17) ; ce qui est licite puisque les groupes A_3 et A_4 sont contemporains.

C'est ce dernier tableau qui constitue l'aboutissement des différents traitements automatiques et qui est le "meilleur" aux yeux de l'archéologue.

Si le résultat brut de l'algorithme de sériation fourni dans le cadre de MPATS n'a pas permis de tomber directement sur l'organisation finale, du moins en étions-nous particulièrement proches. Remarquons aussi que si le phénomène sériel sous-jacent au tableau des données est parfaitement détecté par l'Analyse des Correspondances ; le résultat du croisement des ordres obtenus est moins satisfaisant pour l'archéologue bien que la mise sous forme diagonale soit plus nette. Enfin, bien que le spécialiste ait longuement travaillé à la main ces données et les connaissait parfaitement, nous avons pu l'éclairer sur certains liens très fins entre critères et plaques.

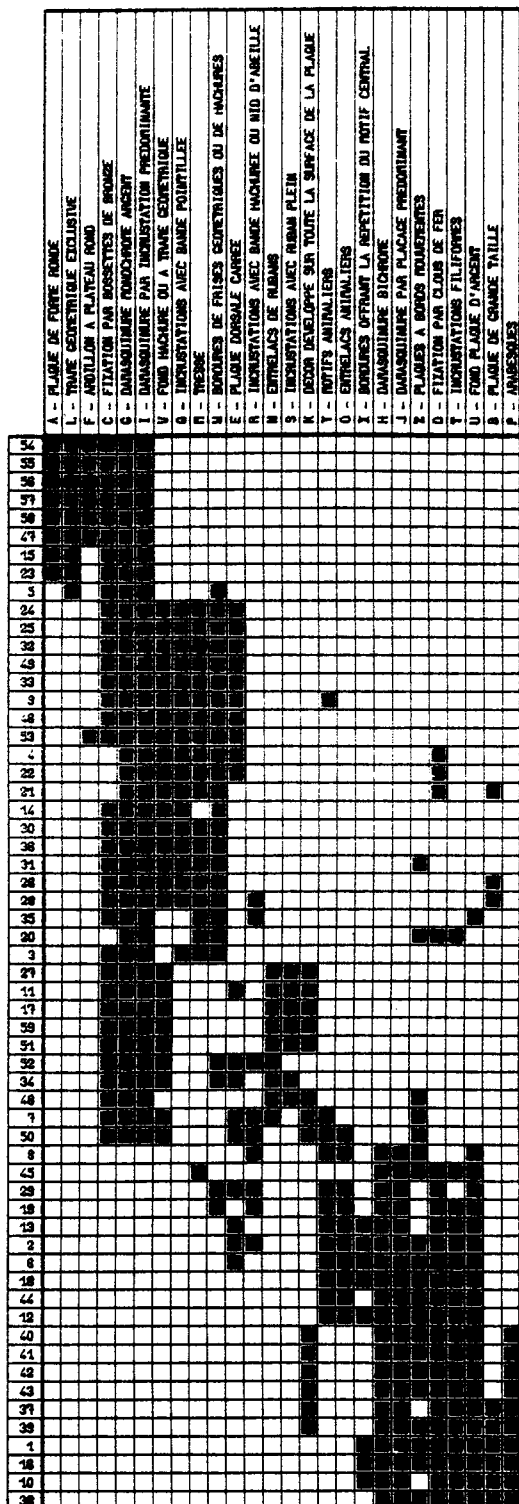


Figure 9

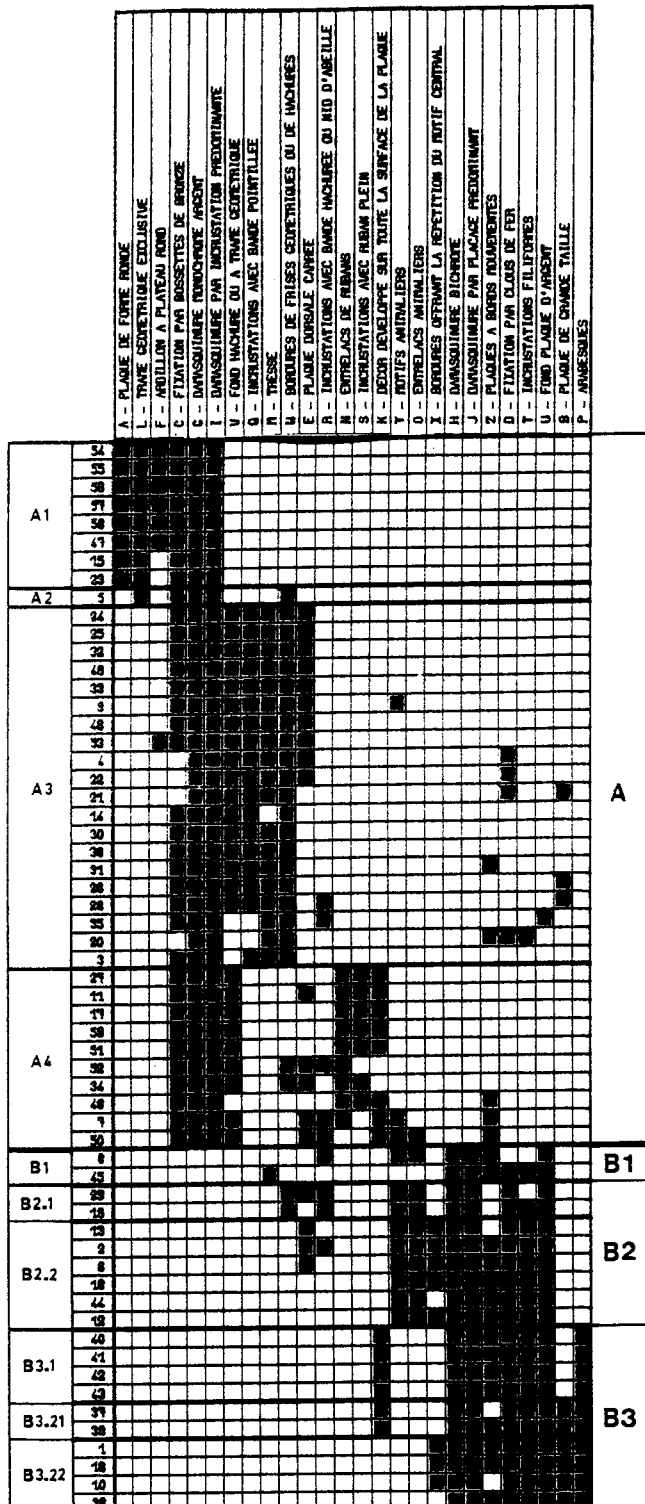


Figure 10

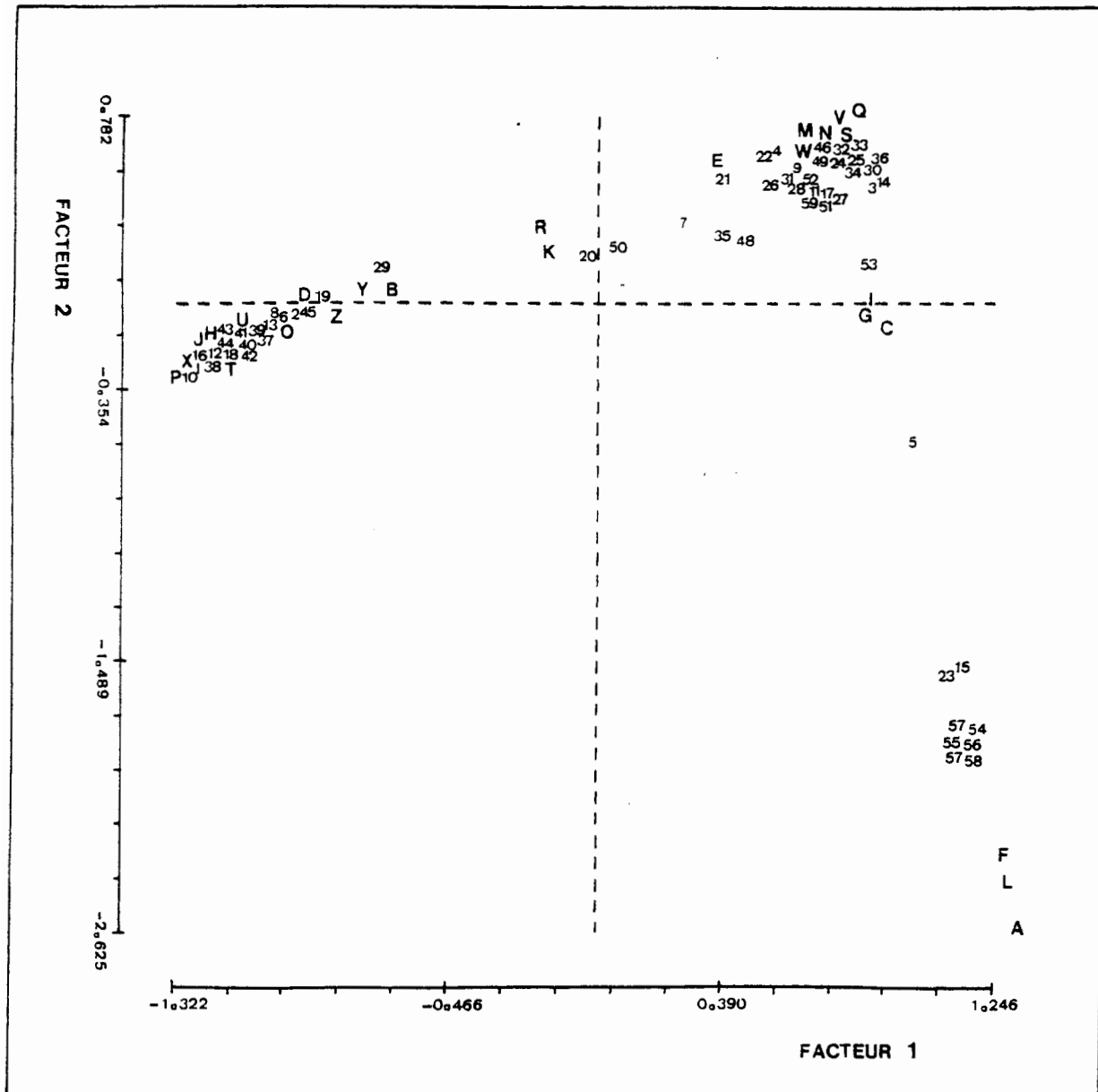


Figure 11

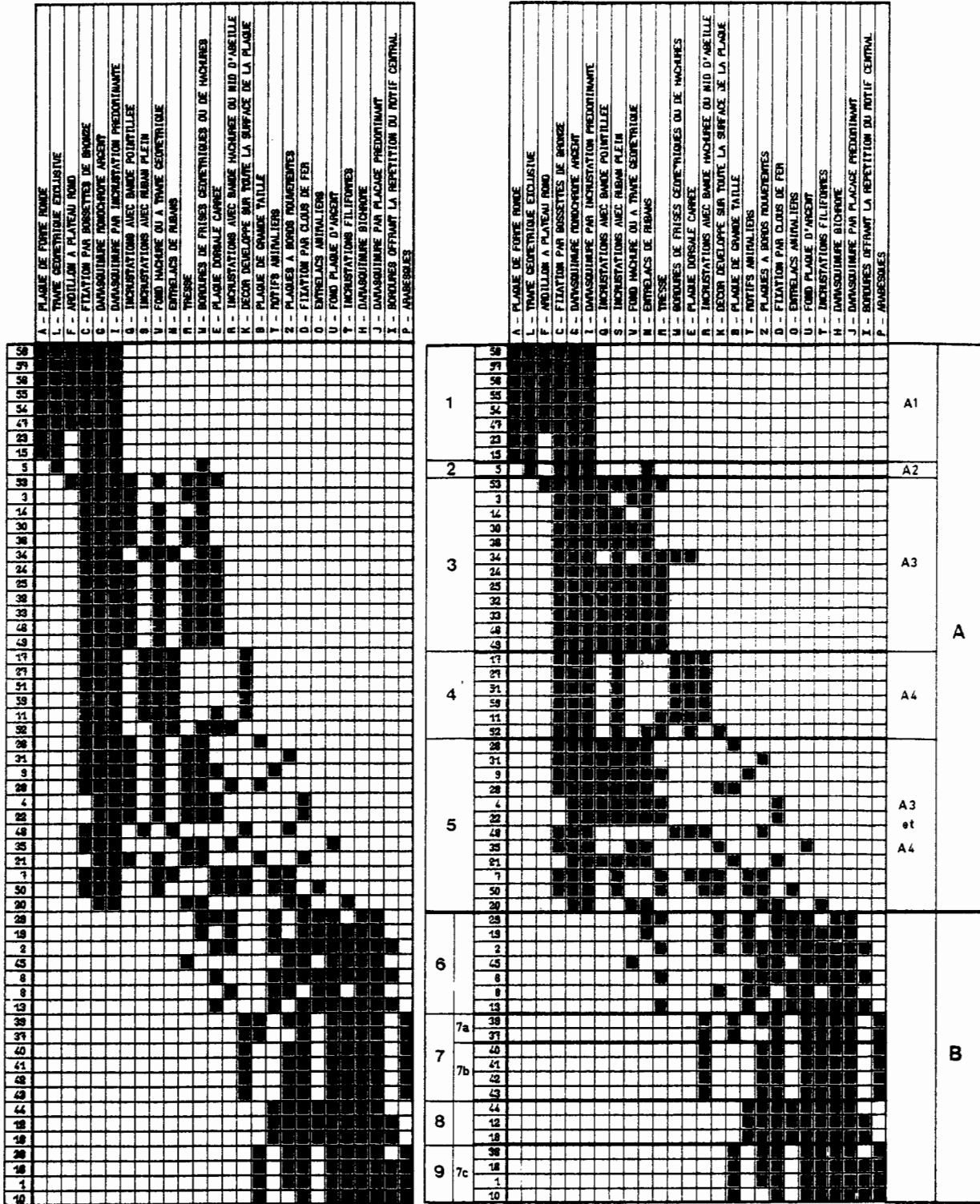


Figure 12

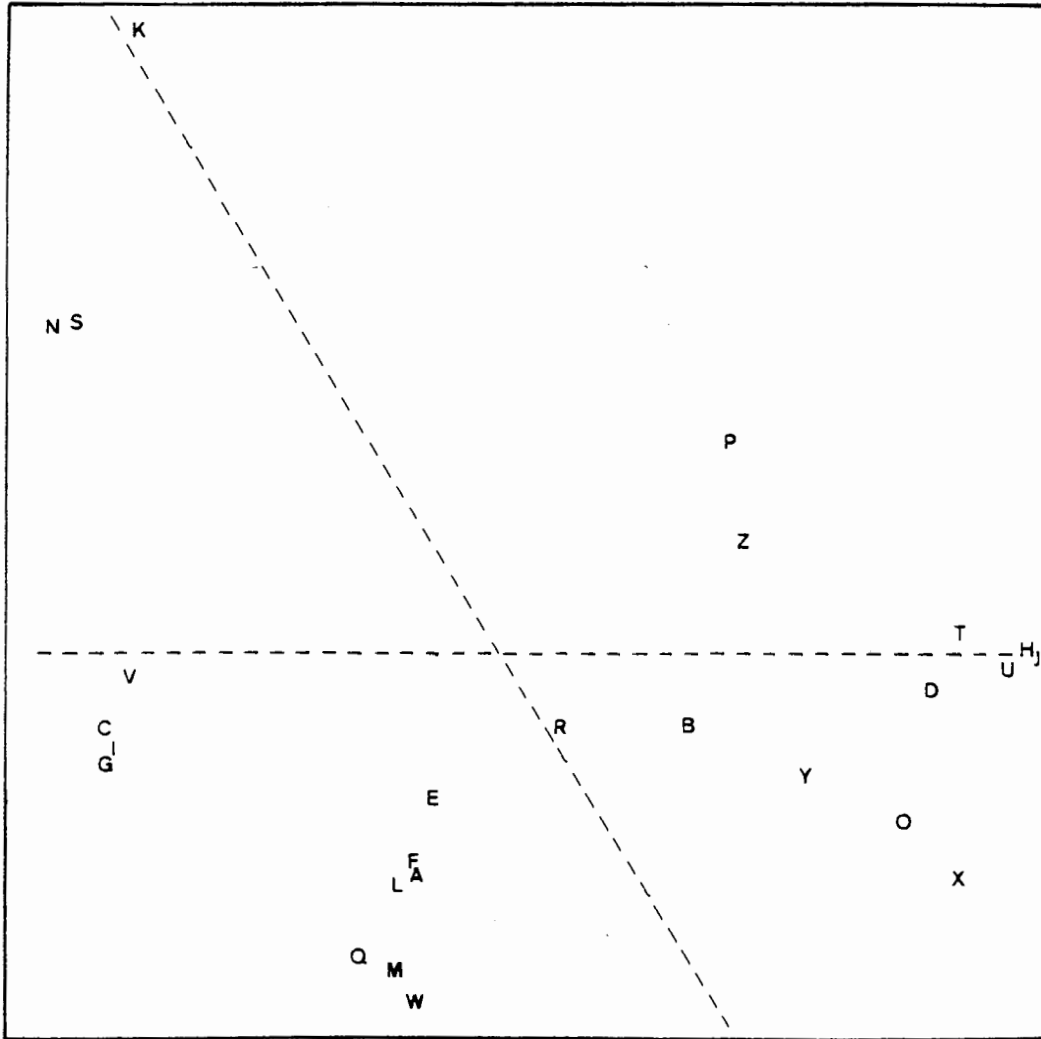


Figure 13

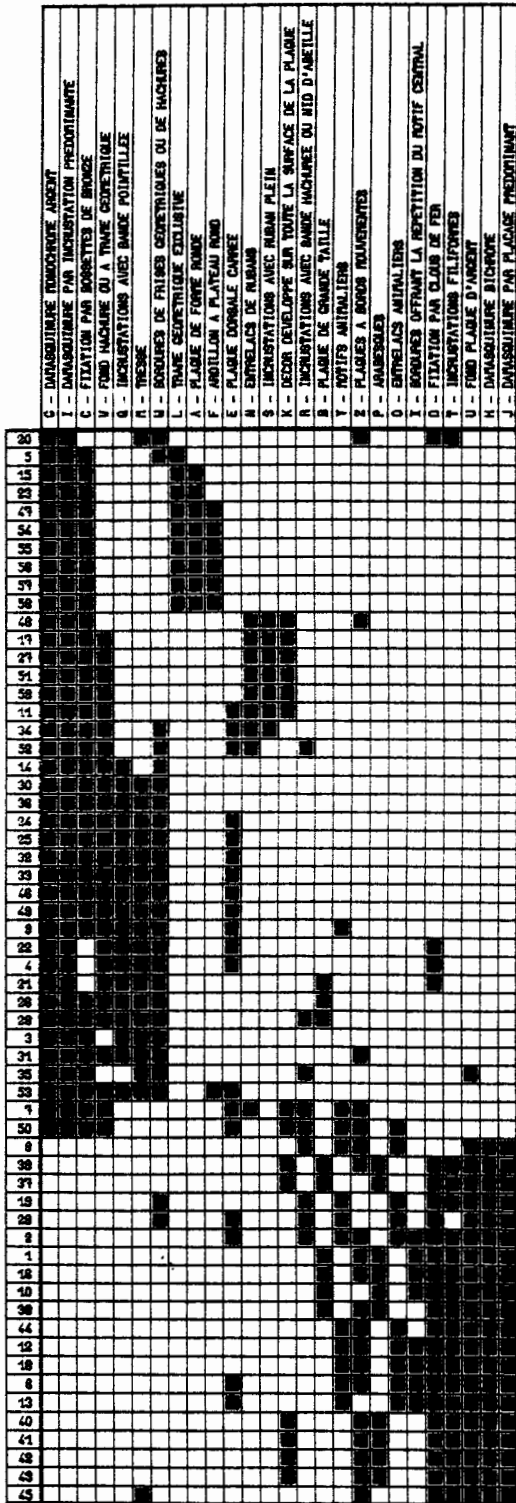


Figure 14

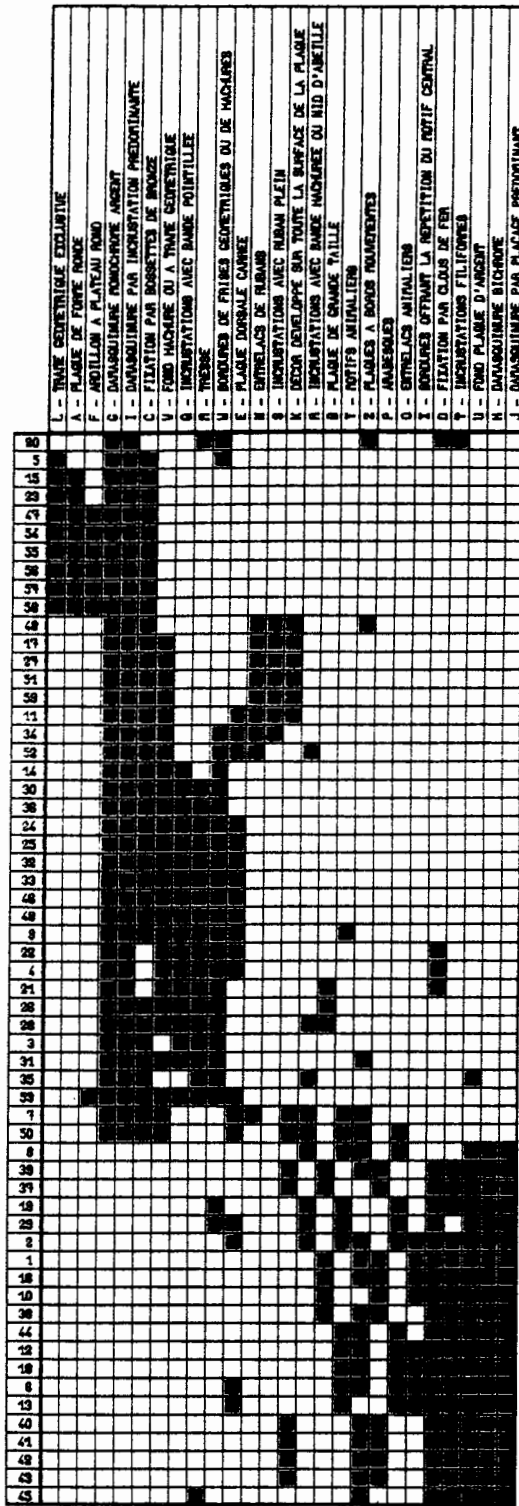


Figure 15

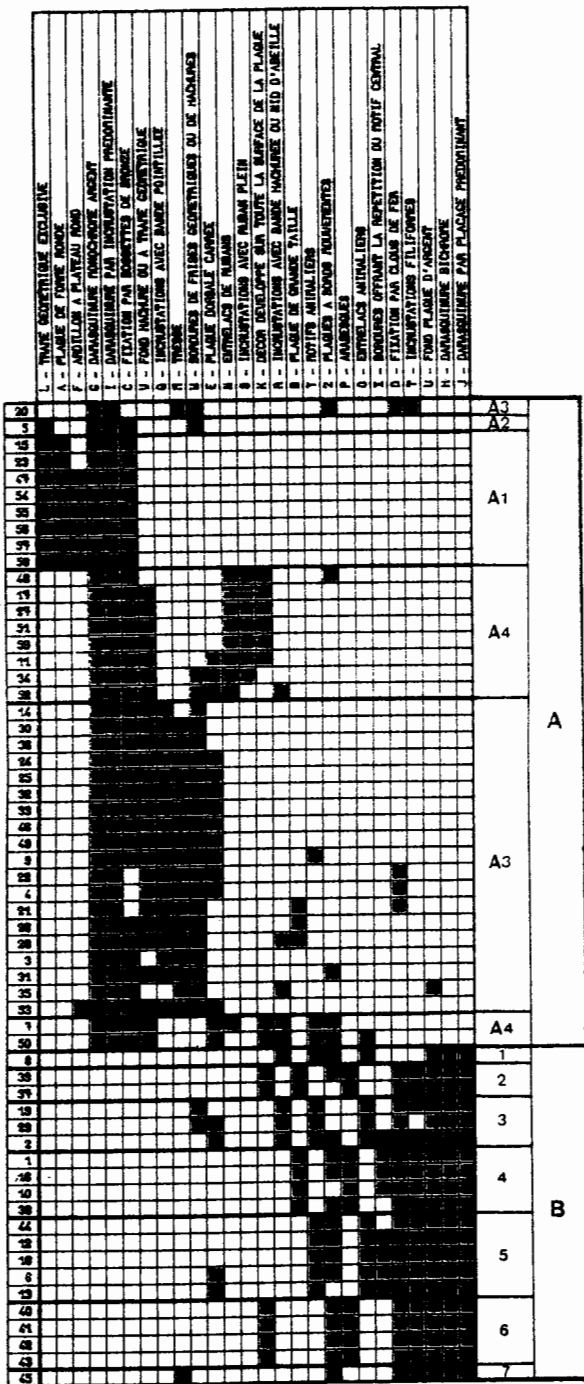


Figure 16

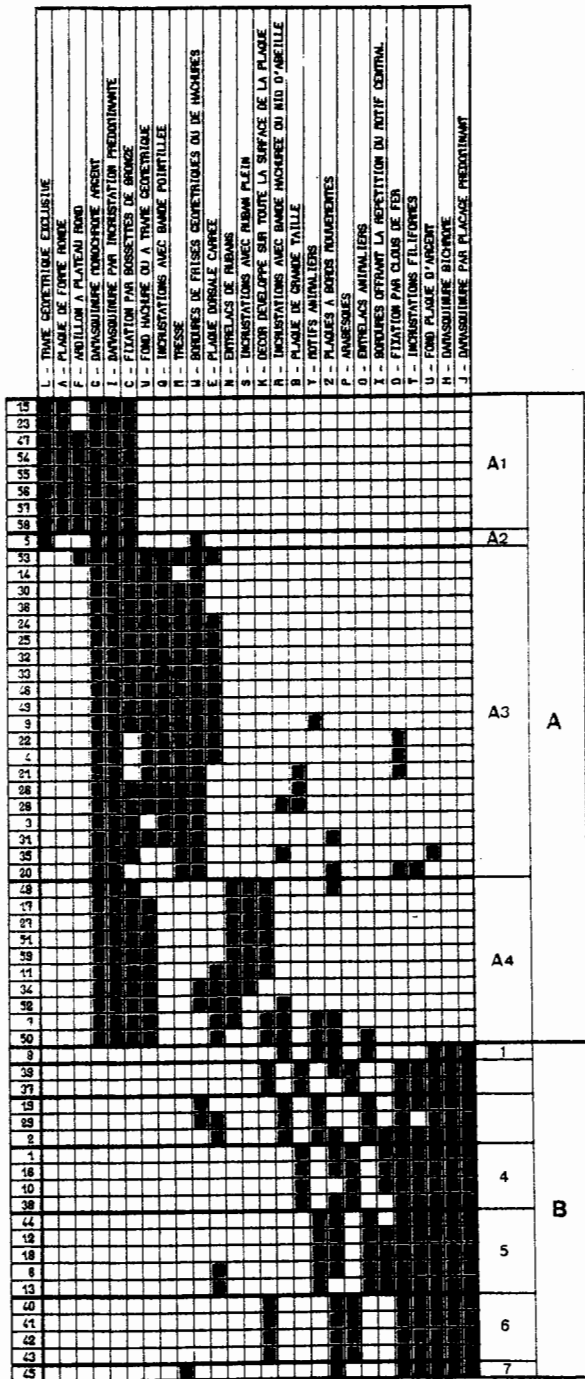


Figure 17

C. EXEMPLE EN ECONOMIE.

C.1. Introduction ; présentation des données.

Le modèle simpliste que nous présentons ici est d'inspiration Keynésienne, dans le cadre d'une économie fermée, mais il intègre le système bancaire.

Il s'agit de l'observation de 20 variables économiques sur 26 trimestres (figure 18). Les mouvements respectifs de la monnaie, du crédit et des taux d'intérêts résultent de l'interaction entre le comportement des autorités, du système bancaire et du public (ménages + entreprises non bancaires). A ce dernier secteur correspond la sphère réelle du modèle, à savoir :

la production intérieure brute Y
la consommation des ménages CP
la formation brute de capital fixe FBCF.

Le bilan simplifié du système bancaire consolidé est représenté comme suit :

$$(RR + RL) + CB + EP = DV + DT$$

à savoir

réserves totales des banques (libres et obligatoires)
+ crédit bancaire
+ portefeuille d'effets publics, d'une part ;
dépôts à vue
+ dépôts à termes, d'autre part.

Le portefeuille d'effets publics essentiellement formé de bons du trésor est de nos jours d'un volume négligeable. Pour cette raison nous n'en tiendrons pas compte.

L'évolution du portefeuille de réescompte de la Banque de France (MCC) a été marquée au cours de la période d'étude par une transformation structurelle de grande ampleur tenant à la réforme des mécanismes d'intervention de la politique d'open-market plus affirmée. La variable DUMMY traduit la volonté des autorités monétaires de passer désormais en priorité par le marché monétaire en laissant progressivement le taux sur celui-ci (TAJJ) descendre en-dessous du taux de réescompte (TRESA).

Les autres variables instrumentales de la politique monétaire sont les coefficients des réserves obligatoires des banques sur les dépôts à vue (TRDV) et les dépôts à terme (TRDT).

Les 26 trimestres :

1	-	2ème trimestre 67
2	-	3ème trimestre 67
:		
:		
26	-	3ème trimestre 73

Variables mesurées

1 - CURR	Billets et monnaies divisionnaires
2 - CP	Consommation des ménages
3 - TDT	Taux débiteurs sur les dépôts à terme
4 - DV	Dépôts à vue
5 - CB	Crédit bancaire
6 - TCC	Taux créditaire partiqué par les banques
7 - TOBL	Taux des obligations
8 - TACT	Taux des actions
9 - MCG	Monnaies centrales gratuites
10 - TAJJ	Taux de l'argent au jour le jour
11 - FBCF	Formation brute de capital fixe
12 - RR+RL	Réserves totales auprès des banques (libres et obligatoires)
13 - DT	Dépôts à terme dans les banques
14 - MCC	Portefeuille de réescompte de la Banque de France
15 - TRES	Taux de réescompte de la Banque de France
16 - Y	Production intérieure brute
17 - DY/Y	Variation relative de la production
18 - TRDV	Coefficient des réserves obligatoires sur les dépôts à vue
19 - TRDT	Coefficient des réserves obligatoires sur les dépôts à terme
20 - DUMMY	Politique monétaire de la B. F. (modification du plafond de réescompte)

Figure 18.

Deux autres variables intéressantes sont prises en compte, à savoir la somme des billets et monnaie divisionnaire (CURR) et la monnaie centrale gratuite (or et devises) (MCG).

Enfin nous avons introduit un ensemble de taux d'intérêts représentant les coûts d'opportunités des agents sur les différents marchés financiers :

le taux de réescompte	TRES
le taux de l'argent au jour le jour	TAJJ
le taux créditeur pratiqué par les banques	TCC
le taux débiteur sur les dépôts à terme	TDT
le taux de rendement des obligations	TOBL
le taux de rendement des actions	TACT

L'étude porte sur des observations trimestrielles durant la période 1967-1973, et le modèle que nous venons de présenter ici a fait l'objet d'une recherche économétrique à laquelle nous avons participé à l'Institut Orléanais de Finance (cf. [2]).

C.2. Résultats comparés de MPATS, ACP et MPAGD.

Au delà de l'interprétation des résultats, ce qui nous intéresse est d'évaluer le comportement de la représentation euclidienne autour principalement des deux premiers pôles, fournie par MPATS ; par rapport à celle, techniquement plus élaborée, donnée par l'Analyse en Composantes Principales Normée et sous-tendue par les axes factoriels.

Avec extraction des pôles par la variance, l'algorithme MPATS fournit deux premiers pôles (MCG et TDT) entraînant deux axes non confondus avec

les deux premiers axes factoriels (cf. figure 19 et figure 20). Sur la figure 19 nous avons représenté la position des deux axes-pôles de notre analyse sur le premier plan factoriel. Nous pouvons constater sur la figure 20 que cette représentation est proche de celle de l'analyse factorielle, figure 19.

Sur la figure 21 nous avons projeté les variables sur le premier et le troisième pôle (MCG et DY/Y). Si nous comparons avec l'analyse factorielle, figure 22, où nous avons projeté le nuage sur les axes 1 et 3 la ressemblance des deux graphiques est encore très grande, les éléments proches du second axe étant les mêmes dans les deux cas.

Si maintenant nous comparons les représentations MPATS et ACP pour les observations, nous remarquons une très étonnante quasi-identité entre les projections sur les deux plans respectifs constitués par les deux premiers axes (figure 23 et figure 24). Malheureusement cette ressemblance ne se poursuit pas pour les autres pôles.

Comme nous pouvons le constater, l'algorithme de représentation de MPATS fournit des configurations très proches de celle de ACP. Il s'agit d'un résultat qu'on ne pouvait absolument pas prévoir a priori compte tenu du caractère élémentaire et simple de l'algorithme de MPATS par rapport à celui de ACP qui suppose une diagonalisation de matrice. Certes, ici, l'interprétation des résultats de traitement par MPATS n'apportera rien de plus par rapport à ceux par ACP ; toutefois les résultats de MPATS facilitent le dialogue avec le spécialiste non mathématicien et montrent de quoi dépendent réellement les résultats de ACP. Nous allons ici interpréter les représentations euclidiennes fournies par ACP ; alors qu'on aurait pu le faire à partir de celles fournies par MPATS.

Le premier axe factoriel (cf. figure 19) est entraîné par l'ensemble des grands agrégats économiques et en particulier par la production qui est croissante sur la période observée. Cet axe sous-tend donc également l'évolution dans le temps. C'est ce que nous retrouvons de manière très frappante le long du premier axe pour la projection des trimestres (figure 24). Le second axe lui (figure 19) oppose l'ensemble des taux d'intérêts aux taux des actions. Or, toujours sur la figure 24 (resp. 23) nous voyons apparaître une courbe continue de la forme \surd qui est entraînée par le temps au moyen du second axe factoriel (resp. par le second pôle d'attraction). Cette courbe correspond parfaitement à l'évolution des principaux taux d'intérêts sur la période en question : 2ème trimestre 67 (point 1) au 3ème trimestre 73 (point 26).

Signalons que sur l'ensemble des paramètres économiques, a été appliqué l'algorithme de classification MPAGD (méthode des pôles d'agrégation utilisant les distances) (cf. Chap. 8, § VIII) où après la première formation des classes ; ces dernières ont été recentrées, chacune autour de son "meilleur représentant", défini comme étant le point dont la somme des carrés des distances aux autres points de la classe est minimum. A partir du système de noyaux ainsi déterminé, l'ensemble des sommets est réaffecté conformément au critère du moment d'inertie où le carré de la distance entre deux points-variables reste conforme à l'indice de proximité Q entre variables avec lequel nous travaillons :

$$(\forall (v,w) \in V^2), d^2(v,w) = Q(v,v) + Q(w,w) - 2Q(v,w) .$$

L'application de l'algorithme nous donne une classification avec une

partition en 5 classes qui est excellente (figure 25). Le rapport inertie inter-classe / inertie totale du nuage est de 56,94 %, c'est la part d'inertie expliquée par cette partition.

La première classe correspond bien au système bancaire auquel s'est ajoutée la variable représentant les billets et monnaie divisionnaire (CURR).

La seconde classe regroupe l'ensemble de la sphère réelle y compris la variation relative de la production (DY/Y).

La troisième classe est celle de la politique monétaire des autorités.

Dans la quatrième classe on retrouve l'ensemble des taux d'intérêts que nous avons introduits dans le modèle.

Seule la dernière classe n'a pas grande signification, si ce n'est qu'elle permet de regrouper les deux variables qui se sont avérées non-explicatives et non-significatives en tant que variables exogènes dans l'analyse économétrique du modèle.

En guise de conclusion, terminons en soulignant que, quelle que soit l'importance des résultats obtenus, l'analyse des faits économiques ne peut se dispenser d'une référence à la théorie.

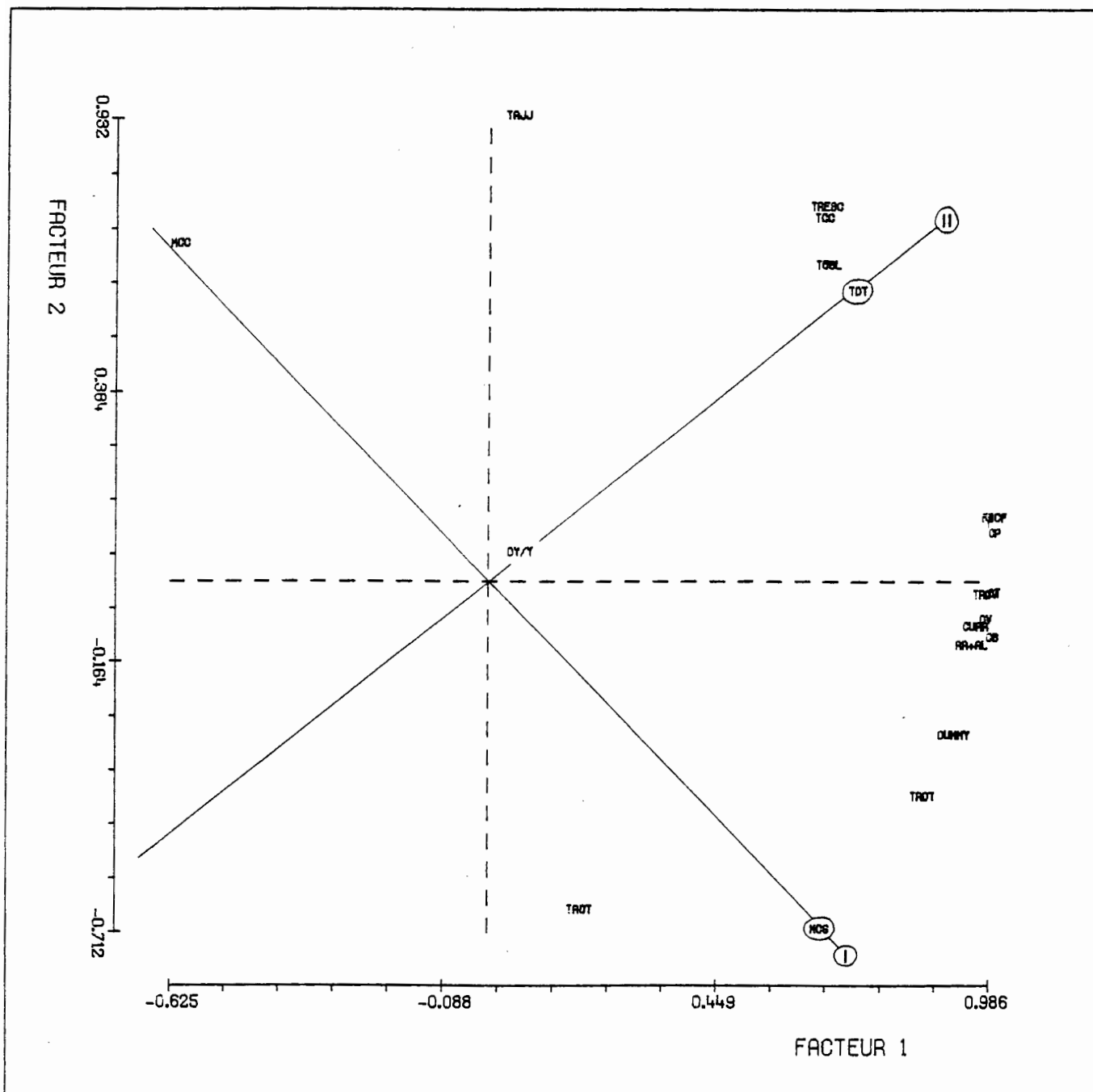


Fig. 19.

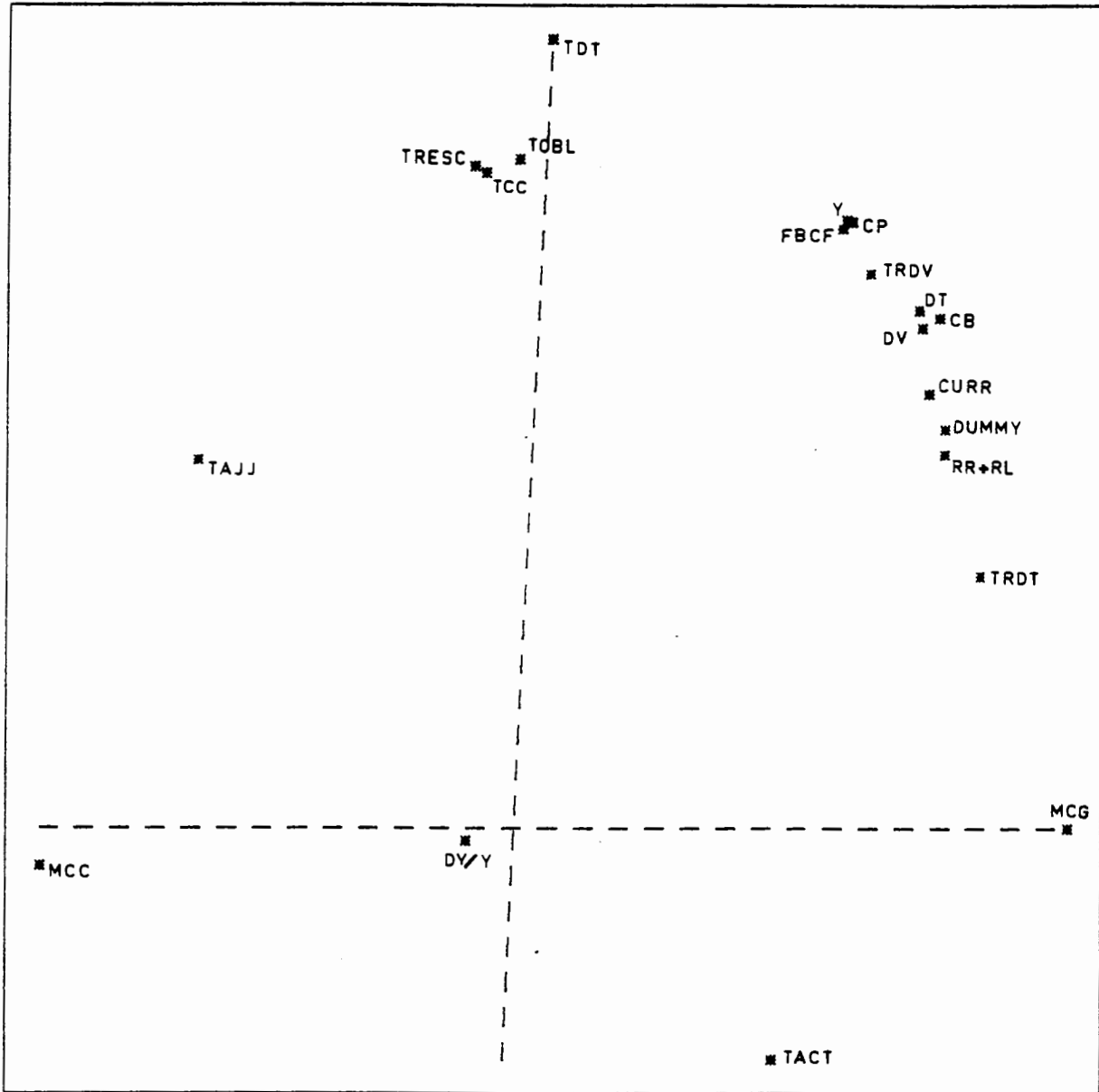


Fig. 20.

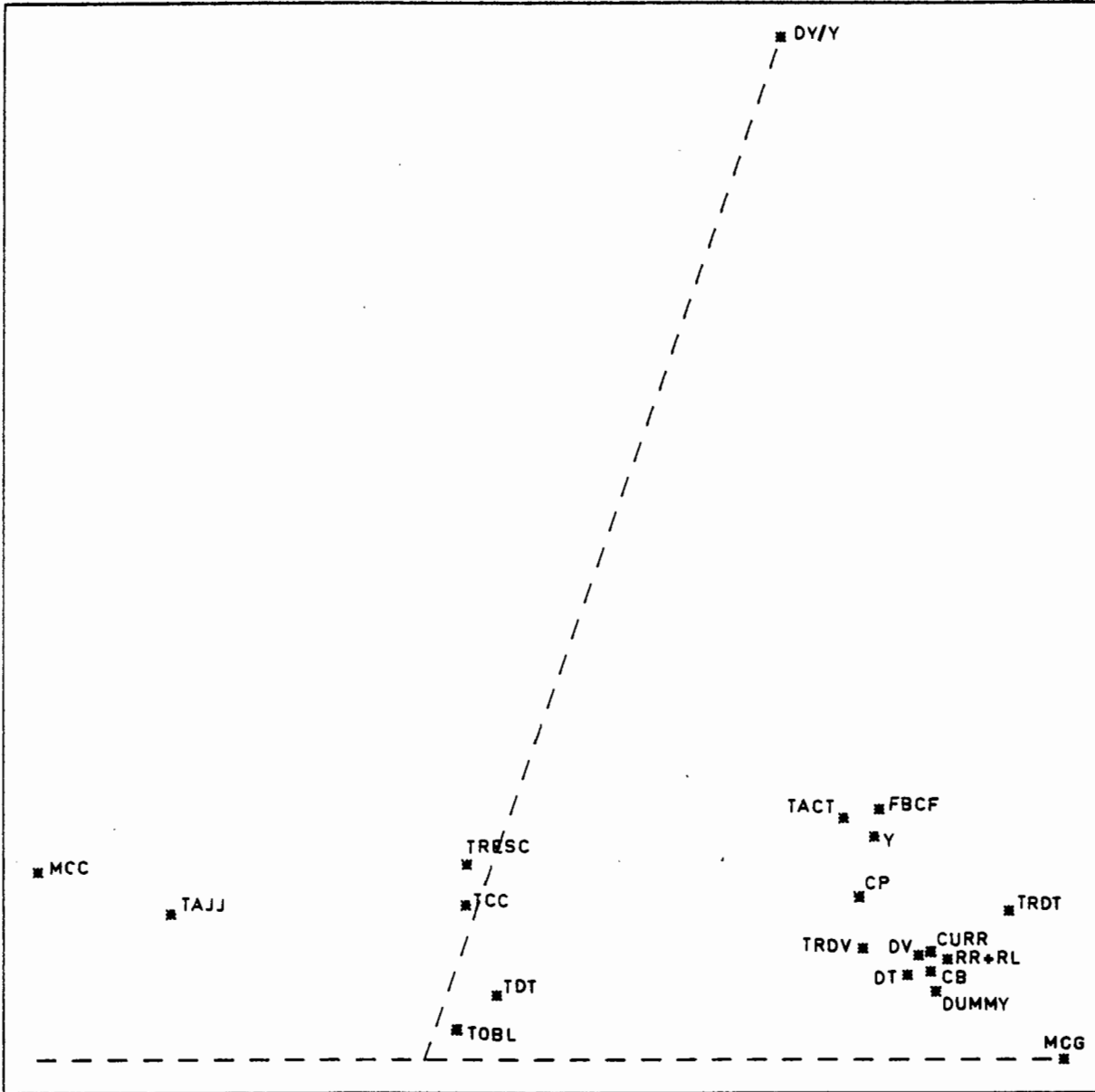


Fig. 21.

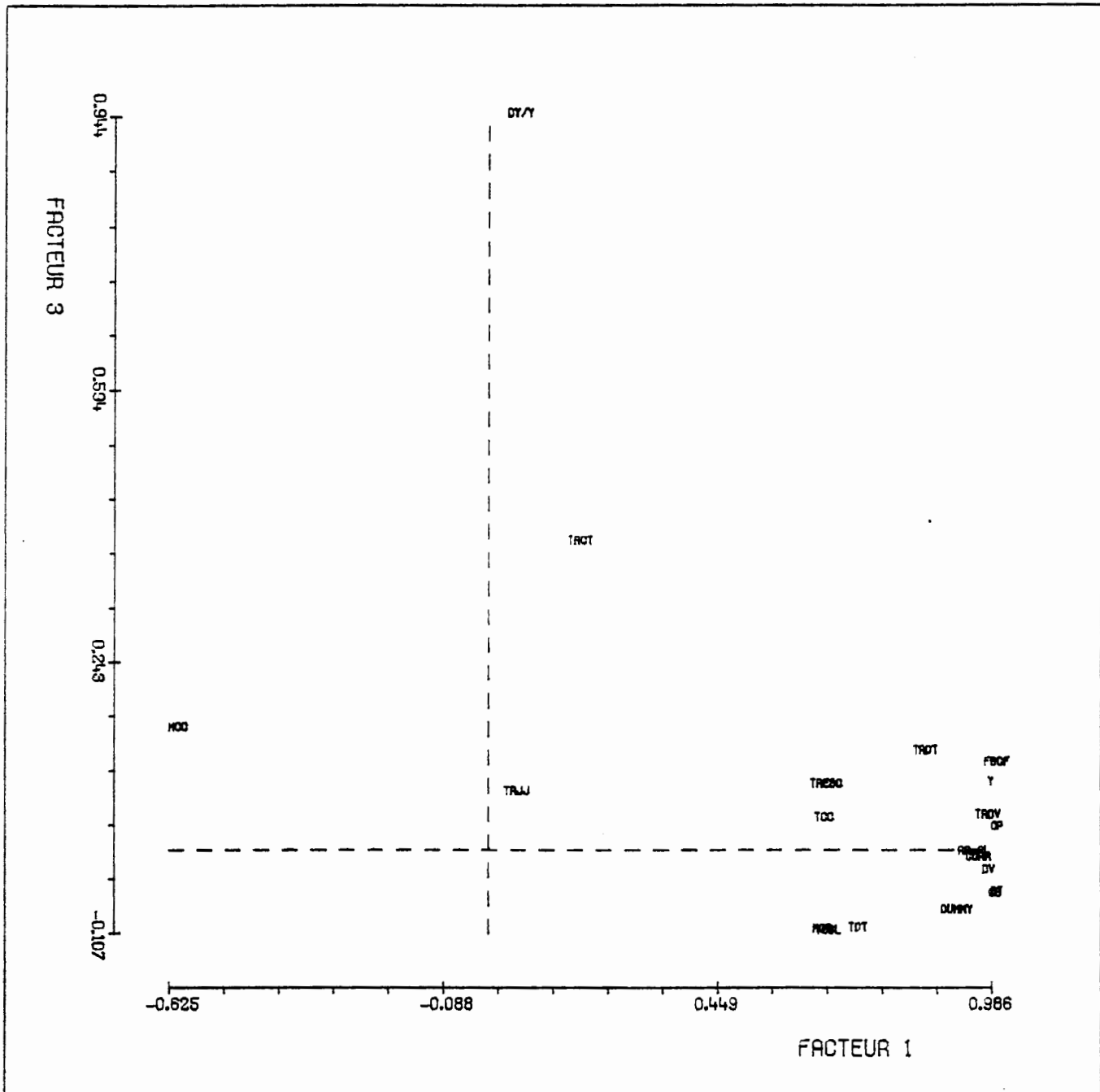


Fig. 22.

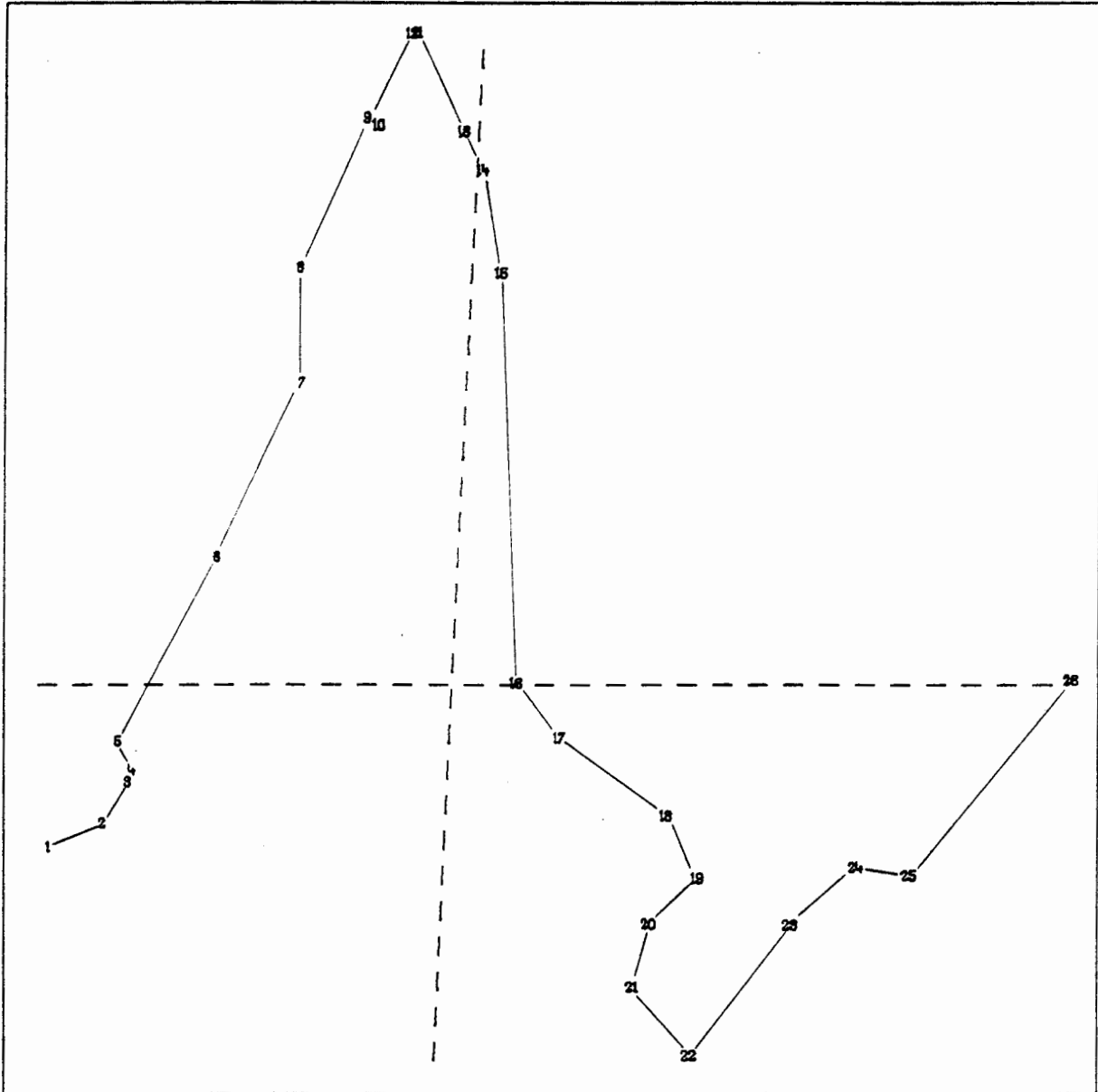


Fig. 23.

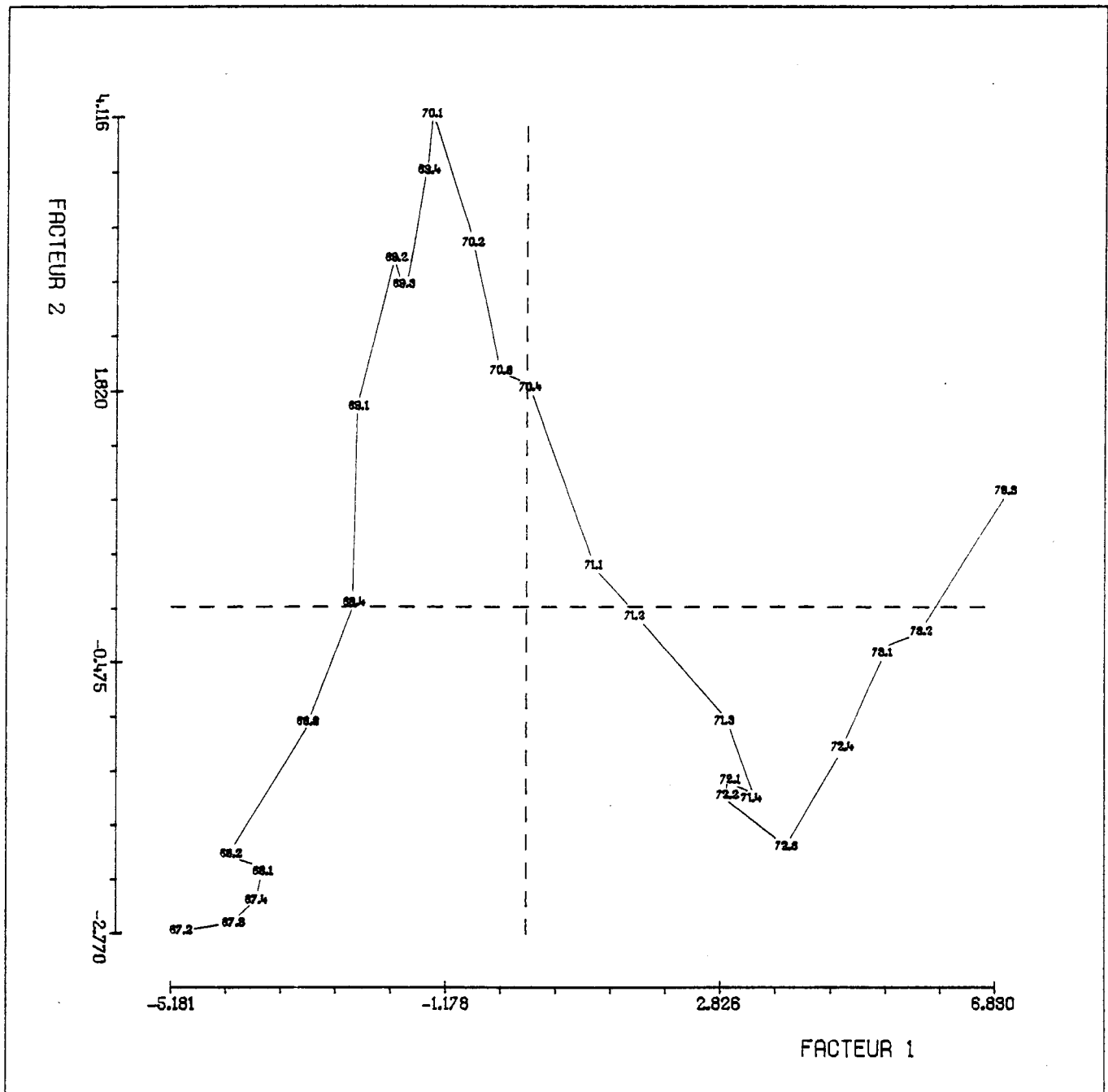


Fig. 24.

NOMBRE DE CLASSES DE LA PARTITION. 5 CLASSES	
Part d'inertie expliquée : 58,94 % Critère d'adéquation : 9,68	
CLASSE 1	
DT	Dépôts à terme dans les banques
CB	Crédits bancaire
DY	Dépôts à vue
CURR	Billets et monnaies divisonnaires
RR+RL	Réserves totales auprès des banques (libres et obligatoires)
CLASSE 2	
Y	Production intérieure brute
FBCF	Formation brute de capital fixe
CP	Consommation des ménages
DY/Y	Variation relative de la production
CLASSE 3	
TRDY	Coefficient des réserves obligatoires sur les dépôts à vie
TRDT	Coefficient des réserves obligatoires sur les dépôts à terme
DUMMY	Politique monétaire de la B. F. (modification du plafond de réescompte)
CLASSE 4	
TRESC	Taux de réescompte de la Banque de France
TCC	Taux créditaire pratiqué par les banques
TOBL	Taux des obligations
TDT	Taux débiteurs sur les dépôts à terme
TAJJ	Taux de l'argent au jour le jour
MCC	Portefeuille de préescompte de la Banque de France
CLASSE 5	
MCS	Marchés centrales gratuites
TACT	Taux des actions

Figure 25

BIBLIOGRAPHIE

- [1] LEREDDE H. ; *"Les Méthodes des Pôles d'Attraction"*, thèse en cours de rédaction, à soutenir en 1979, Univ. Paris VI, Institut de Programmation.
- [2] OUTREVILLE J.-F. ; *"Un modèle macroéconomique du secteur Français"*, rapport interne, Institut Orléanais de Finance, 1976.
- [3] PERIN P. ; *"La datation des tombes méroviengiennes"*. Historiques, Méthodes, Applications. Paris, DROZ, 1980.
- [4] LEREDDE H., PERIN P. ; *"Les plaques-boucles mérovingiennes"*. Les dossiers de l'Archéologie, n° 42, mars-avril 1980, p. 83-87.

IMPRIMERIE GAUTHIER VILLARS, 70, rue de Saint-Mandé, 93100 MONTREUIL

Imprimeur : 2445

198070-81

Dépôt légal, 4^e trimestre 1981

Imprimé en France