



HAL
open science

PONDER puis MNDOPT. Une methodologie en classification automatique. Son application a l'analyse des effets de renforcements du reseau de transport a tres haute tension d'EDF

T. Cembrzynski

► **To cite this version:**

T. Cembrzynski. PONDER puis MNDOPT. Une methodologie en classification automatique. Son application a l'analyse des effets de renforcements du reseau de transport a tres haute tension d'EDF. RR-0726, INRIA. 1987. inria-00075826

HAL Id: inria-00075826

<https://inria.hal.science/inria-00075826>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
BP 105
78153 Le Chesnay Cedex
France

Tel. (1) 39 63 55 11

Rapports de Recherche

N° 726

**PONDER puis MNDOPT
UNE METHODOLOGIE EN
CLASSIFICATION
AUTOMATIQUE.
SON APPLICATION A
L'ANALYSE DES EFFETS DE
RENFORCEMENTS DU RESEAU DE
TRANSPORT A TRES
HAUTE TENSION D'EDF.**

Thierry CEMBRZYNSKI

SEPTEMBRE 1987

PONDER puis MNDOPT
Une méthodologie en Classification Automatique.
Son application à l'analyse des effets de renforcements du
réseau de transport à très haute tension d'EDF.

PONDER and MNDOPT
A methodology in Cluster Analysis.
Its application for studying effects of reinforcements on the
EDF very high voltage grid.

Thierry CEMBRZYNSKI

INRIA

Domaine de Voluceau

BP 105 - Rocquencourt -

78153 Le Chesnay Cedex - France -

Résumé

Nous présentons dans ces quelques pages la méthode de classification utilisée pour l'analyse des résultats du modèle de planification du réseau de transport à très haute tension d'EDF. Cette méthode comporte une étape de reconnaissance d'éléments aberrants par l'algorithme PONDER puis la phase de classification où la procédure MNDOPT adapte automatiquement le nombre de classes aux données à classer. Après une première partie qui décrit les méthodes de classification utilisées, nous présenterons un exemple d'application sur des données de fonctionnement du réseau de transport à très haute tension d'EDF.

Abstract

We present in some pages the clustering method used for analysing the results of the EDF very high voltage grid planning model. This method involves a first part for recognizing outliers with the PONDER algorithm and a clustering part in which the MNDOPT algorithm adjusts the number of clusters to the data. After an introduction to the clustering methods used, we will describe an example using data from the EDF very high voltage grid.

SOMMAIRE

INTRODUCTION

PREMIERE PARTIE

A : Présentation sommaire des méthodes de classification utilisées

- A.a : Généralités sur les méthodes de classification
- A.b : La classification hiérarchique
- A.c : L'algorithme des centres mobiles
- A.d : L'algorithme des transferts

B : La procédure de reconnaissance d'éléments aberrants : PONDER

- B.a : Description, origine et but
- B.b : L'algorithme suivi dans PONDER
- B.c : Les distances utilisables
- B.d : Les résultats fournis par PONDER

C : Une procédure de classification Intelligente : MNDOPT

- C.a : Description, origine et but
- C.b : L'algorithme suivi dans MNDOPT
- C.c : Cohérence des critères optimisés
- C.d : Les critères de sélection du nombre de classes
- C.e : La sélection des noyaux initiaux
- C.f : Les distances utilisables
- C.g : Les résultats fournis par MNDOPT
- C.h : Aspects informatiques

DEUXIEME PARTIE

Application de PONDER et de MNDOPT à l'analyse de l'effet du renforcement de la ligne Morbras-Villevaudé du réseau de transport à très haute tension d'EDF.

A: Analyse sans filtrage des surcoûts du renforcement (métrique $1/\sigma^2$)

B: Analyse sans filtrage des surcoûts du renforcement (métrique W^{-1})

C: Analyse avec filtrage des surcoûts du renforcement (métrique $1/\sigma^2$)

D: Analyse avec filtrage des surcoûts du renforcement (métrique W^{-1})

CONCLUSION

BIBLIOGRAPHIE

ANNEXES

INTRODUCTION

La méthodologie basée sur PONDER et MNDOPT que nous allons décrire, cherche à répondre de façon satisfaisante à l'objectif de nombreux utilisateurs en matière de classification automatique, à savoir: classer une population importante sans en connaître a priori le contenu et obtenir en un seul essai, une partition de bonne qualité; même si les données contiennent des éléments aberrants.

Dans une première partie, un peu théorique nous présenterons les méthodes retenues pour la reconnaissance avant classification, des données aberrantes, puis celles utilisées pour la phase de classification. Dans une seconde partie nous illustrerons cette méthodologie par un exemple d'application à l'étude des effets du renforcement d'une liaison du réseau de transport à très haute tension d'EDF, située au Nord Est de Paris.

PREMIERE PARTIE

A : Présentation sommaire des méthodes de classification utilisées

A.a : Généralités sur les méthodes de classification

Rappelons tout d'abord quelques généralités. Les techniques de classification automatique sont destinées à produire des groupements d'objets ou d'individus décrits par un certain nombre de variables ou de caractères. Il s'agit d'une branche de l'analyse des données en pleine effervescence donnant lieu à des publications nombreuses et variées.

Il serait bien ambitieux de vouloir dresser un bilan et de fournir un exposé synthétique des différentes méthodes existantes. Il est cependant possible en revanche, de faire quelques remarques critiques qui permettront (nous l'espérons) au lecteur non initié de cheminer sans se perdre dans la forêt des méthodes.

Pour l'essentiel, les techniques de classification font appel à une heuristique, et non aux calculs formalisés usuels. Alors que les valeurs des coefficients d'un modèle linéaire sont la solution d'une équation pouvant s'écrire sous une forme très condensée, (même si sa résolution est complexe), la définition des classes ne se fera qu'à partir d'une formulation algorithmique : une série d'opérations est définie de façon récursive et répétitive. Ceci explique que, plus encore que les techniques d'analyse linéaire, les techniques de classification soient tributaires de l'ordinateur. Il en découle aussi que la mise en oeuvre de la plupart d'entre elles fait surtout appel à des concepts récents.

Il existe plusieurs familles d'algorithmes de classification automatique. Nous nous limiterons ici qu'aux techniques utilisées dans notre application: les algorithmes ascendants ou agglomératifs qui procèdent à la construction des classes par agglomération successive des éléments deux à deux et qui fournissent une hiérarchie de partitions des objets; et les algorithmes conduisant directement à des partitions comme les Nuées Dynamiques (DIDAY 1971), particulièrement intéressantes dans le cas de grands ensembles de données.

A.b : La classification hiérarchique

L'algorithme fondamental de la classification ascendante hiérarchique se déroule de la façon suivante. On désigne par éléments soit les objets à classer eux-mêmes, soit les regroupements d'objets générés par la procédure.

A l'étape 0; il y a n éléments à classer qui sont les n objets.

On cherche les deux éléments les plus proches (au sens de la mesure de similarité choisie - il s'agit bien souvent de la perte d'information expliquée -) que l'on agrège en un nouvel élément.

On calcule les distances entre le nouvel élément et les éléments restants. On se retrouve alors dans les mêmes conditions qu'à l'étape 0 mais avec seulement $n-1$ éléments à classer cette fois.

On cherche de nouveau les deux éléments les plus proches que l'on agrège en un seul élément, on calcule de nouveau les distances et l'on réitère le processus jusqu'à ce qu'il n'y ait plus qu'un seul élément.

D'un point de vue informatique et donc pratique la complexité de ces algorithmes est proportionnelle à n^2v où n désigne le nombre d'objets à classer et v le nombre de variables qui décrivent ces objets.

A.c : L'algorithme des centres mobiles

La méthode de classification autour de centres mobiles est probablement la technique de partitionnement automatique la mieux adaptée actuellement aux vastes recueils de données qui seront si nécessaire conservées sur une mémoire auxiliaire qui sera lue plusieurs fois séquentiellement afin de ne jamais encombrer de zones importantes dans les mémoires des ordinateurs; quoique ces contraintes pratiques deviennent de moins en moins importantes depuis que l'on peut, même sur des minis ordinateurs (SPS7 par exemple) disposer de plusieurs méga-octets.

L'algorithme que nous présentons peut-être imputé principalement à FORGY (1965) ou à MACQUEEN (1967) bien que de nombreux travaux aient été menés parallèlement pour introduire des variantes ou des généralisations (DIDAY 1971).

Soit à partitionner un ensemble I de n individus caractérisés sur p variables quantitatives. On suppose que l'espace R^p dans lequel on représente les individus soit muni d'une distance notée δ . Pour constituer au maximum q classes, les étapes de l'algorithme d'agrégation autour des centres mobiles sont alors les suivantes:

Etape 0 : On détermine q centres de classes par tirage aléatoire (par exemple).

Ces centres C_{01}, \dots, C_{0q} induisent une partition de l'ensemble des objets I en q classes I_{01}, \dots, I_{0q} . Ainsi l'individu i appartient à la classe I_{0k} si le point i est plus proche de C_{0k} que de tout autre centre.

Etape 1 : On détermine q nouveaux centres de classes C_{11}, \dots, C_{1q} en prenant les centres de gravité des classes I_{01}, \dots, I_{0q} . Les nouveaux noyaux induisent une nouvelle partition construite avec la même règle et formée des classes I_{11}, \dots, I_{1q} .

Etape m : On détermine q nouveaux centres de classes C_{m1}, \dots, C_{mq} en prenant les centres de gravité des classes $I_{m-11}, \dots, I_{m-1q}$. Les nouveaux noyaux induisent une nouvelle partition construite avec la même règle et formée des classes I_{m1}, \dots, I_{mq} .

L'algorithme s'arrête soit lorsque deux itérations successives conduisent à la même partition, soit lorsqu'un critère convenablement choisi (il s'agit souvent de la variance intra-classes), cesse de décroître de façon sensible, soit encore parce qu'un nombre maximal d'itérations a été fixé a priori.

Dans tous les cas cependant, la partition finale dépendra très fortement du choix initial des noyaux de l'étape 0.

On peut facilement montrer que cet algorithme converge, mais bien entendu ce n'est pas la convergence mais surtout la vitesse de convergence qui justifie en pratique l'utilisation de ces méthodes. On constate de plus en plus, de nos jours que les techniques de classification les plus en vue sont en fait, celles qui ont donné lieu à l'écriture de programmes souples et performants.

D'un point de vue informatique et donc pratique la complexité de ces algorithmes non hiérarchiques est proportionnelle à nqv où n désigne le nombre d'objets à classer, v le nombre de variables qui décrivent ces objets et q le nombre de classes que l'on désire.

A.d : La méthode des Transferts

La méthode des transferts est une procédure de classification itérative d'un ensemble I de n objets décrits par v variables quantitatives en k classes disjointes. Comme pour les Nuées Dynamiques et autres algorithmes non hiérarchiques comme les centres mobiles, le nombre k de classes est un paramètre fixé initialement par l'utilisateur. Cette méthode n'est pas en général, utilisée directement pour construire une partition à partir d'une configuration quelconque, elle est plus adaptée à l'amélioration d'une partition déjà

construite par un autre algorithme. Cet algorithme est assez général et peut être adapté sans difficultés à d'autres critères que l'inertie.

Décrivons en l'algorithme fondamental tel qu'il est présenté dans les ouvrages d'HARTIGAN (1975) ou de VILLOING (1983).

A une étape donnée, l'algorithme transfère un élément d'une classe dans une autre à condition que la nouvelle partition soit meilleure au sens du critère que l'on cherche à optimiser (maximisation de l'inertie inter-classes par exemple), que la partition obtenue à l'étape précédente.

Le processus s'arrête quand aucun transfert n'améliore plus la partition; on a alors atteint un optimum relatif. Si le critère à optimiser permet de comparer des partitions n'ayant pas le même nombre de classes, on pourra alors vider des classes; sinon, on ne pourra pas le faire et on obtiendra à la fin une partition ayant le même nombre de classes que la partition de départ.

L'algorithme des transferts intégré dans les programmes MNDOPT, TRANSF et PONDER (commandes SICLA cf: Y.LECHEVALLIER, H.RALAMBONDRAINY, G.GOVAERT 1986) est très voisin de celui proposé par BEALE en 1969, cependant au lieu de travailler sur l'inertie perdue nous avons jugé préférable dans la procédure, au prix cependant de calculs plus lourds, de chercher à améliorer directement l'inertie expliquée; ceci afin de ne jamais rencontrer de problèmes d'inversion.

Précisons davantage la procédure mise en œuvre. Soit I un ensemble de n objets décrits par p paramètres quantitatifs, à partitionner en k classe; à l'élément i est associé le point x_i de R^p et on muni I d'une distance δ . On prend comme critère de qualité ϕ d'une partition, celui de l'inertie inter-classes :

$$\phi = \sum_{l=1}^k \delta^2(G_l, G) * \mu_l$$

G désigne le centre de gravité du nuage des points

μ_l le poids de la classe l

G_l le centre de gravité de la classe l.

Alors un calcul très simple montre que pour que l'élément i puisse être transféré de la première classe dans la seconde par exemple, c'est à dire pour que la partition : $\{C_1 - \{x_i\}, C_2 + \{x_i\}, C_3, \dots, C_k\}$ soit meilleure au sens de ϕ que la partition : $\{C_1, C_2, C_3, \dots, C_k\}$ il suffit, dans l'algorithme classique des Transferts que :

$$(P_{C_2} / P_{C_2+p_i}) * \delta^2(G_2, x_i) < (P_{C_1} / P_{C_1-p_i}) * \delta^2(G_1, x_i)$$

P_{C_1} poids de la classe 1

P_{C_2} poids de la classe 2

p_i poids de l'élément x_i

Cependant, l'expérience l'a montré, l'utilisation de ce critère basé sur la minimisation de l'inertie intra-classes, peut provoquer des inversions en cours d'optimisation. Il nous a semblé alors plus judicieux d'effectuer que les transferts qui améliorent directement l'inertie expliquée. Pour ce faire nous avons utilisé le critère suivant :

$$\delta^2(G_1, G) * P_{C1} + \delta^2(G_2, G) * P_{C2} < \delta^2(G_1^*, G) * (P_{C1} - p_i) + \delta^2(G_2^*, G) * (P_{C2} + p_i)$$

G_1^* centre de gravité de la classe 1 après la perte du point x_i

G_2^* centre de gravité de la classe 2 après le gain du point x_i

D'un point de vue informatique et donc pratique la complexité de l'algorithme non hiérarchique des Transferts est proportionnelle à nk^2v où n désigne le nombre d'objets à classer, v le nombre de variables qui décrivent ces objets et k le nombre de classes que l'on désire.

Cependant dans le cadre de la commande MNDOPT nous avons limité la complexité de cet algorithme en $3nv$ en ne considérant pour un objet i que les transferts dans les deux classes les plus proches de celle dont il fait initialement partie.

Ce choix se justifiait pour notre application où nous manipulons de volumineux ensembles de données; en effet l'algorithme des Transferts est employé en mode *optimisation* de la partition de départ obtenue par coupure de l'arbre hiérarchique à un niveau jugé significatif. La partition obtenue était alors de bonne qualité; l'expérience montra alors que le nombre total de transferts était très faible (<5%); dès lors il apparaissait inutile en fin d'optimisation d'explorer systématiquement nk transferts, en sachant parfaitement qu'il n'y en aurait pas.

B: La procédure de reconnaissance d'éléments aberrants : PONDER

B.a : Description, origine et but

PONDER est une procédure, basée sur l'algorithme des transferts, de reconnaissance d'éléments aberrants d'un ensemble I de n individus décrits par v variables quantitatives.

L'origine de cette procédure vient du fait qu'en Classification Automatique, les points aberrants introduisent un biais en particulier dans les métriques en conditionnant alors très fortement la mesure de ressemblance entre les objets à classer (problème fondamental en classification); ce qui rend alors les partitions finales instables. Confrontés à ce sérieux problème qui devient crucial lorsque l'on n'utilise plus les métriques Euclidiennes classiques, il nous est alors apparu intéressant de pouvoir les filtrer auparavant.

Dans le cadre de PONDER, nous avons essayé d'apporter un élément de réponse à ces problèmes en construisant alors une variable de pondération pouvant être utilisée par une procédure de classification automatique.

B.b : L'algorithme suivi dans PONDER

Pour repérer les points aberrants dans un ensemble I de n individus caractérisés par p variables quantitatives (on suppose que l'espace R^p dans lequel on représente les individus soit muni d'une distance); nous proposons dans le programme PONDER, la procédure de classification divisive suivante:

On définit un seuil critique en dessous duquel une classe sera considérée comme formée d'éléments aberrants; dans PONDER ce seuil est : $\min(0.5\% \text{ de l'effectif}, 10)$.

On calcule la métrique sur l'ensemble des points \bar{i} dont les poids sont non nuls.

A l'étape 0 : soit le point P_0 le plus éloigné du centre de gravité G du nuage. Partant de la partition formée des deux classes ($i - \{P_0\}$, $\{\bar{i} - P_0\}$), on cherche à améliorer celle-ci au regard du critère de l'inertie, par la méthode des Transferts.

Précisons davantage la démarche suivie dans la procédure.

On muni I d'une distance δ . On prend comme critère de qualité ϕ d'une partition, celui de l'inertie inter-classes. Si ϕ_0 désigne l'inertie expliquée par la partition formée des deux classes (I- $\{P_0\}$, $\{P_0\}$) on a :

$$\phi_0 = (M_1 - p_0) * \delta^2 (G_1, G) + p_0 * \delta^2 (P_0, G)$$

G désigne le centre de gravité du nuage des points

p_0 le poids du point P_0

G_1 le centre de gravité de la classe I- $\{P_0\}$

M_1 la masse de la classes I- $\{P_0\}$

On améliore par transferts la qualité de cette partition initiale ainsi : le point X sera affecté à la classe $\{P_0\}$ si la partition (I- $\{P_0, X\}$, $\{P_0, X\}$) est meilleure que la partition de départ (I- $\{P_0\}$, $\{P_0\}$) au regard de l'inertie expliquée, c'est à dire si :

$$(M_1 - p_0) * d^2 (G_1, G) + p_0 * d^2 (P_0, G) < (M_1 - (p_0 + p_x)) * d^2 (G^*_1, G) + (p_0 + p_x) * d^2 (P^*_0, G)$$

G^*_1 centre de gravité de la classe I- $\{P_0\}$ après la perte du point X

P^*_0 centre de gravité de la classe $\{P_0\}$ après le transfert du point X.

Quand la procédure ne permet plus aucune amélioration (i.e quand il n'y a plus aucun transfert d'éléments), on vérifie que l'effectif de la classe C_{P_0} engendrée initialement par le point P_0 soit bien inférieur au seuil critique fixé a priori.

Si oui, à l'étape 1: soit le point P_1 qui extrait le plus d'inertie intra-classes. Partant de la partition formée des trois classes (I- C_{P_0} - $\{P_1\}$, C_{P_0} , $\{P_1\}$) où C_{P_0} représente la classe engendrée à l'étape précédente par le point P_0 ; on cherche à améliorer la partition au regard du critère de l'inertie, par la méthode des transferts. A la convergence, on vérifie que les effectifs des classes engendrées initialement par les points P_0 et P_1 soient également inférieurs au seuil critique fixé au départ.

Si non, la procédure s'arrête quand la dernière classe créée est d'un effectif supérieur au seuil critique. Dans ce cas on associe un poids nul aux points reconnus comme aberrants, on recalcule la métrique sur les éléments restés actifs et on recommence l'opération de filtrage. PONDER s'arrête définitivement si entre deux filtrages consécutifs la procédure n'a pas pu identifier d'autres éléments aberrants; le programme crée alors une variable de pondération qui est une variable binaire si les éléments étaient initialement tous de même poids.

B.c : Les distances utilisables

Pour la classification et la reconnaissance des éléments aberrants, nous avons introduit la possibilité d'utiliser diverses distances, dans notre application à l'analyse du réseau électrique d'EDF, toutes ont fourni une partition intéressante d'un point de vue surtout pratique: mise en évidence de problèmes économiques ou plus électriques selon métriques utilisées. Nous expliquerons plus précisément cela dans le cadre de NMDOPT.

On peut dans PONDER, utiliser les métriques suivantes:

- la métrique identité : I (centrage des données).

- la métrique $1/\sigma^2 I$ dont l'opération de centrage et réduction des données rend les variables aussi discriminantes les unes que les autres.

Ces deux métriques sont très intéressantes, cependant elles ne permettent pas de prendre en compte les corrélations qui peuvent exister dans les données. De ce fait nous avons introduit la possibilité d'utiliser la distance de Mahalanobis (notre application étant relative à un graphe valué, nos variables étaient corrélées).

Ainsi dans PONDER on peut utiliser également:

- la métrique V^{-1} (métrique de Mahalanobis) qui, en prenant en compte les corrélations qui peuvent exister entre les variables; produit des classes généralement non sphériques mais plus allongées.

En fait cette métrique sert surtout à filtrer les éléments aberrants avant d'effectuer une classification avec les distances adaptatives globales (GOVAERT dans DIDAY 1979) dans MNDOPT. En effet si on utilise la métrique W^{-1} dans MNDOPT et encore plus si on utilise W_k^{-1} dans DISLOC (distances adaptatives locales) et que l'on a des points aberrants, on peut créer des classes d'effectif très faible, voire inférieur au nombre de variables actives de la classification. De ce fait les métriques peuvent ne plus être de plein rang, la convergence de la procédure n'est plus garantie ni numériquement ni théoriquement et finalement la qualité de la partition finale s'en trouve fortement compromise. L'introduction de cette métrique de Mahalanobis V^{-1} dans PONDER était alors impérative pour pouvoir utiliser sans risques les procédures basées sur les distances adaptatives. Nous reviendrons plus précisément sur ce point fondamental qui est en fait directement à l'origine de la procédure PONDER.

On constate alors que du fait des métriques utilisables, PONDER offre de larges champs d'applications.

B.d: Les résultats fournis par PONDER

L'objectif principal poursuivi dans la commande PONDER est de produire une variable de pondération (variable quantitative ou binaire) utilisable avec des procédures comme MNDOPT. Cependant à la fin de chaque étape de filtrage on décrit succinctement les éléments reconnus comme aberrants en mesurant leurs dispersions par rapport au centre de gravité du nuage avec les indices suivants:

$$\text{Dispersion (var } \langle j \rangle) = (\text{Moycl } \langle j \rangle - \text{Moy } \langle j \rangle) / \text{Et } \langle j \rangle$$

avec :

var $\langle j \rangle$: j ième variable active de la classification

Moycl $\langle j \rangle$: moyenne de la j ième variable dans la classe

Moy $\langle j \rangle$: moyenne de la j ième variable

Et $\langle j \rangle$: écart-type de la j ième variable

C : Une procédure de classification intelligente : MNDOPT

C.a : Description origine et but

MNDOPT est une procédure de classification itérative et principalement non hiérarchique, d'un ensemble de n objets décrits par v variables quantitatives en k classes disjointes. La différence principale avec les autres méthodes de classification non hiérarchiques est que le nombre de classes k est un paramètre non nécessairement fixé par l'utilisateur. Il ne s'agit pas d'une procédure à nombre de classes variable comme BOULES et ISODATA (cf: MODULAD); MNDOPT est en effet capable de détecter automatiquement un nombre de classes intéressant.

L'origine de cette procédure vient du fait que l'utilisateur désireux d'effectuer une classification d'une population donnée demande à une méthode de classification de l'aider dans le choix d'une partition, et en particulier d'un nombre de classes.

Or dans toutes les méthodes de classification ascendante hiérarchique, on se heurte principalement à deux problèmes; l'un essentiellement pratique et l'autre plus théorique.

Le problème pratique est que cette méthode est *difficilement utilisable pour le traitement de vastes recueils de données*, car elle impose le calcul, le stockage et la mise à jour de la demi-matrice des distances entre les objets à classer. Cependant certains pourront nous objecter que l'inaptitude de ces méthodes au traitement de populations nombreuses n'est peut-être plus justifiée depuis que, grâce à des algorithmes très performants on ait pu traiter plusieurs milliers d'individus sur quelques variables (BRUYNOOGHE 1978). Néanmoins il est possible de reprocher à ces méthodes d'imposer à la population à classer une structure hiérarchique, de ce fait les classes obtenues en haut de l'arbre, résultant donc d'un très grand nombre de regroupements hiérarchiques, ne risquent-elles pas d'être assez peu optimales, si telle n'est pas la structure naturelle de la population ?

Le problème théorique réside quant à lui dans le choix, par l'utilisateur, à la fin de classification, d'une partition en q classes parmi la hiérarchie des partitions possibles sans qu'il soit sûr, pour autant que la partition en q classes obtenue par la coupure de l'arbre de classification soit effectivement optimale - elle ne l'est d'ailleurs presque jamais-. Ce choix pourra alors se faire sur la valeur d'un critère ou d'un seuil de significativité, ou encore seulement d'après les caractéristiques des classes trouvées.

Restent alors les méthodes non hiérarchiques.

Pour les techniques d'agrégation autour de centres mobiles, les problèmes sont de nature radicalement différente.

Les algorithmes convergent toujours vers des optima locaux. Le problème de la recherche d'une partition en q classes (en prenant notamment comme critère à minimiser la variance intra-classes), n'a pas encore donné lieu à un algorithme vraiment satisfaisant. *Les partitions dépendent fortement des premiers centres choisis* et également de l'ordre des éléments à classer.

Ces algorithmes sont très efficaces pour décrire rapidement de vastes recueils de données, cependant *ils imposent à l'utilisateur de choisir a priori le nombre de classes* qu'il désire, et ce choix est bien souvent fort délicat, particulièrement lorsqu'on ne connaît pas le contenu de ses données.

Il existe bien des méthodes dites à nombre de classes variable comme BOULES et ISODATA (cf: MODULAD) mais l'utilisateur doit alors définir d'autres paramètres dont le choix est plus ou moins arbitraire, destinés à "piloter" l'élaboration de la partition, ce qui en rend l'utilisation fort délicate.

Dans le cadre du programme MNDOPT nous avons essayé d'apporter un élément de réponse à ces problèmes d'initialisation et du nombre de classes. Nous ne prétendons par avoir trouvé la solution miracle à ces problèmes; mais nous nous estimerons heureux si cette procédure peut permettre d'aider efficacement un utilisateur non nécessairement spécialiste de la Classification Automatique, dans l'analyse de ses données.

C.b : L'algorithme suivi dans MNDOPT

Dans le cadre de notre étude nous avons retenu les Nuées Dynamiques avec cependant une différence assez importante, qui réside dans l'utilisation d'une classification ascendante hiérarchique selon la variance (hiérarchie de Ward) pour "détecter"

automatiquement un nombre de classes statistiquement significatives et sélectionner également de bons noyaux initiaux. On adapte ainsi directement le nombre de classes aux données à classer.

La stratégie suivie dans notre programme de classification est apparentée à celle de MOLLIÈRE (cf : DIDAY et coll 1979), elle comporte :

Etape 1 phase d'apprentissage :

On effectue les Nuées Dynamiques sur un grand nombre de classes (cinquante classes pour être précis, ce qui est largement suffisant dans le cadre de notre application); et dans la mesure où le choix des noyaux de départ conditionne fortement la partition finale, on sélectionne initialement ceux-ci sous contraintes (cf: C.e La sélection des noyaux initiaux). Ce nombre de classes relativement important et cette sélection sous contraintes permettent alors d'espérer que l'on explorera plus profondément la structure réelle de la population et donc d'obtenir des classes plus significatives.

Cette partition est naturellement sauvegardée et peut-être bien évidemment réutilisée, en outre rien n'empêche d'utiliser pour la phase d'apprentissage la partition des formes fortes que l'on peut obtenir par ailleurs, en effectuant un grand nombre d'expériences des Nuées Dynamiques.

NB: Pendant la phase d'apprentissage, l'utilisateur ne peut utiliser que des métriques fixes ($I, I/\sigma^2, V^{-1}$).

Etape 2 phase de recherche du nombre de classes :

On effectue une classification ascendante hiérarchique pondérée suivant le critère de la variance (Hiérarchie de Ward avec maximisation du moment d'ordre deux d'une partition) sur les centres des classes de la partition obtenue à la fin de la précédente étape, pendant laquelle nous calculons divers critères statistiques - que nous allons présenter (cf § C.d : les critères de sélection du nombre de classes)- pour détecter un nombre de classes intéressant et calculer de bons noyaux initiaux. Cette étape est décisive pour le choix de la partition finale.

A la méthode des connexités descendantes qui présente le défaut d'induire un "effet de chaîne" important, dû à l'ultramétrie sous-dominante, nous avons préféré utiliser la hiérarchie de Ward basée sur le critère de la variance.

En effet les techniques d'agrégation selon la variance sont bien souvent plus fiables car elles cherchent à optimiser, à chaque étape, selon des critères liés à l'inertie, la partition obtenue par agrégation des deux éléments. La stratégie d'agrégation est alors la suivante : au lieu de chercher les éléments les plus proches, on cherche ceux dont la fusion provoque la plus petite perte d'inertie expliquée. Mathématiquement cela revient à considérer la perte d'inertie expliquée comme un indice de dissimilarité entre les objets à classer. Dans la hiérarchie de Ward, celle-ci est calculée ainsi :

$$\text{Perte } (i,j) = ((M_i * M_j) / (M_i + M_j)) * \delta^2 (i,j)$$

M_i : masse du point i

M_j : masse du point j

$\delta (i,j)$: distance entre i et j

C'est cet indice de dissimilarité que nous utilisons dans notre programme de typologie MNDOPT, pour la phase de recherche du nombre de classes.

Etape 3 phase d'optimisation de la partition :

L'étape 3 se propose simplement d'améliorer la partition obtenue à l'étape 2 par coupure de l'arbre hiérarchique à un niveau jugé significatif, en cherchant à optimiser un critère (inertie inter-classes). Suivant la métrique utilisée on peut utiliser un algorithme procédant par transferts d'individus entre classes (métriques $I, 1/\sigma^2 I, V^{-1}$) ou en utilisant les Nuées Dynamiques avec ces noyaux optimaux quand on décide d'utiliser la métrique W^{-1} . Il est alors en effet algorithmiquement impossible d'utiliser (avec la métrique W^{-1}) la méthode des Transferts qui comme celle des KMEANS, fait "bouger" les centres de gravité en cours d'affectation. Dès le premier transfert, la métrique n'étant plus adaptée au centre de gravité; la décroissance stricte de l'inertie intra-classes n'est alors plus garantie et la qualité de la partition finale peut s'en trouver fortement compromise.

Remarque: Dans tous les cas - au prix toutefois d'un tri indirect- l'ordre d'affectation des individus se fait par distances croissantes au centre de gravité de la classe à laquelle ils appartiennent. Cette option permet de s'affranchir de l'effet de séquentialité lié à la position des objets à classer, et d'affecter aux classes les individus les plus "douteux", qu'en fin d'itération.

Cette méthodologie assez complexe, nous est apparue répondre de façon satisfaisante à l'objectif de nombreux utilisateurs en matière de classification automatique, à savoir: classer une population importante sans en connaître a priori le contenu et obtenir en un seul essai, une partition de bonne qualité; même si les données contiennent des éléments aberrants.

C.c : Cohérence des critères optimisés

L'étape 2 (phase de recherche du nombre de classes) privilégie le critère de l'inertie inter-classes (agrégation selon la perte minimale d'inertie expliquée : méthode de Ward); dans ces conditions il nous a semblé plus cohérent de retenir pour l'étape 1 des Nuées Dynamiques (phase d'apprentissage) l'option "centres de gravité" et d'imposer également à l'algorithme d'optimiser le critère de l'inertie expliquée. Enfin la partition obtenue à l'étape 2 au niveau optimal de la hiérarchie, n'optimisait en effet le critère d'inertie que partiellement, en restant dans le cadre de la structure hiérarchique (à un niveau donné la partition obtenue est la meilleure du point de vue du critère parmi toutes celles pouvant résulter de la fusion de deux classes au niveau précédent).

Néanmoins on constate - en comparant en particulier les pourcentages d'inertie expliquée - que la partition résultant de la coupure (à un niveau apparu comme significatif) de la hiérarchie sur la partition d'apprentissage, est généralement déjà une bonne partition. De ce fait nous avons choisi d'utiliser pour l'étape 3 (phase d'optimisation) un algorithme du transfert pour fournir une partition correspondant, sinon à l'optimum absolu, du moins à un optimum local de bonne qualité au regard du critère de l'inertie expliquée.

C.d: Les critères de sélection du nombre de classes

Trois critères de jugement de la partition produite à un niveau donné de l'arbre des classifications ont été considérés et mis en œuvre.

Ces critères sont les suivants. Il s'agit :

- d'un **critère cubique de classification** (W.S SARLE 1983) dont la philosophie basée sur un test d'hypothèse cherche à estimer l'existence de classes significatives dans les données. On peut toujours, en effet exécuter un programme de partitionnement automatique non hiérarchique sur des données non classifiables (par exemple, tirées à partir d'une loi de probabilités uniforme), on obtiendra quand même une partition, bien

qu'elle ne présente a fortiori aucun sens. Le critère cherche précisément à éviter ce genre de problèmes.

On retiendra la règle de décision suivante :
Si $CCC > 0.0$ il doit y avoir des classes.
Si $CCC \leq 0.0$ il n'y a vraisemblablement pas de classes.

Ce critère empirique et quantitatif compare le R^2 de la classification à chaque niveau de l'arbre hiérarchique à un R^2 théorique calculé sous l'hypothèse d'absence de classes avec une formule élaborée empiriquement à partir de données générées par des simulations de Monte Carlo.

- d'une **statistique locale** (LERMAN 1982) qui cherche à mesurer le degré de dissemblance des éléments que l'on agrège à chaque étape de la hiérarchie.

Ce critère très général est celui fondé sur la "préordonnance". Si nous désignons par K l'ensemble à classifier qui peut correspondre, soit à l'ensemble V des variables, soit à l'ensemble O des objets, la préordonnance sur K est un préordre total sur l'ensemble $L = P_2(K)$ des paires d'éléments de K . Pour ce préordre que nous supposons ici -pour simplifier- un ordre total et strict, le rang d'une paire est une fonction croissante de la ressemblance entre ses composantes, mesurée par l'indice Q de proximité choisi :

$$\forall (p,q) \in L \times L, p < q \iff Q(p) < Q(q)$$

Nous représenterons dans $L \times L$ la préordonnance $\omega(K)$ par son graphe:

$$gr(\omega) = \{(p,q) / (p,q) \in L \times L, p < q \text{ et non } q < p \text{ pour } \omega\}$$

Une partition Π sera représentée dans $L \times L$ par le "rectangle" $R(\pi) \times S(\pi)$ ou $R(\pi)$ (resp. $S(\pi)$) est l'ensemble des paires réunies (resp. séparées) par la partition Π , ($R(\pi) < S(\pi)$ pour l'ordre quotient).

L'indice brut entre la préordonnance $\omega(K)$ et la partition Π est alors :

$$s(\omega, \pi) = \text{card} [gr(\omega) \cap (R(\pi) \times S(\pi))]$$

Ce cardinal peut se mettre - pour les calculs- sous la forme : $s(\omega, \pi) = \sum \{v(q) / q \in R(\pi)\}$ où $v(q)$ est le nombre de paires séparées, strictement à gauche de q pour la préordonnance.

Nous opérons une normalisation de cet indice en associant à la partition Π , une partition aléatoire Π^* dans l'ensemble -muni d'une probabilité uniformément répartie - $P(n,t)$ de toutes les partitions de même type cardinal que Π .

La forme la plus simple de l'indice normalisé qui prend le nom de statistique locale est alors la suivante:

$$S = [s(\omega, \pi) - [r(\pi).s(\pi) / 2] / [r(\pi).s(\pi).[r(\pi)+s(\pi) + 1] / 12]^{0.5}$$

où $r(\pi)$ est le cardinal de $R(\pi)$ et $s(\pi)$ celui de $S(\pi)$.

La suite des valeurs de ce critère local permet de reconnaître quels sont les principaux états d'équilibre dans la synthèse automatique, fournie niveau après niveau

dans l'arbre détaillé des classifications emboîtées.

On retiendra la règle de décision suivante :
Rechercher un maximum (relatif ou absolu) .

- de l'inertie expliquée qui cherche à estimer le degré de précision de la partition correspondante.

C'est probablement le plus classique qui sert d'ailleurs directement dans la formation de l'arbre hiérarchique ou comme critère à optimiser dans la plupart des méthodes non hiérarchiques; si $(P_0, P_1, \dots, P_{l-1}, \dots, P_m)$ est la suite des partitions produites aux différents niveaux de l'arbre $0, 1, \dots, l, \dots, m$, on associera à la partition P_l , indiquée par $\{I_h^1 / 1 \leq h \leq k_l\}$, la valeur du critère :

$$\sum_{1 \leq h \leq k_l} \eta_h // Gh - G //^2$$

où G_h désigne le centre de gravité de la h ième classe, η_h son poids et G le centre de gravité du nuage des éléments à classer.

Plusieurs remarques s'imposent toutefois:

1/ ce critère ne peut être utilisé que dans le cas où l'ensemble à classer peut être représenté de façon naturelle par un nuage de points dans un espace Euclidien;

2/ il y a de bons critères pour l'émergence des classes et il y a de bons critères pour l'évaluation des partitions obtenues et ce ne sont pas nécessairement les mêmes. D'ailleurs, dans le cas ci-dessus, ce critère ne fait que décroître sur la suite des niveaux de l'arbre des classifications qu'il a servi à bâtir. En fait, il importe que le critère d'évaluation soit relativement indépendant de la méthode de formation de l'arbre et ait un caractère très général, ce qui n'exclut pas d'utiliser l'inertie expliquée.

On retiendra la règle de décision suivante :
S'intéresser au gain marginal d'inertie expliquée entre deux partitions successives.

Dans notre programme de classification MNDOPT, nous cherchons le nombre de classes q qui satisfait simultanément aux trois critères de décision implémentés dans la procédure; en privilégiant cependant celui basé sur la préordonnance (LERMAN 1982).

C.e : La sélection des noyaux initiaux

Pour les techniques d'agrégation autour de centres mobiles, comme nous l'avons déjà dit, les partitions dépendent fortement des premiers centres choisis et également de l'ordre des éléments à classer. En fait le choix de ces noyaux initiaux est très important et conditionne fortement les résultats obtenus; c'est peut-être une des raisons pour lesquelles DIDAY proposa pour les Nuées Dynamiques la notion de formes fortes que l'on obtient après croisement de plusieurs partitions construites à partir de noyaux initiaux différents dans chacune des expériences.

Cependant ces formes fortes même si - à l'issue d'un grand nombre d'expériences des Nuées Dynamiques- elles restent en petit nombre par rapport aux individus de départ; elles sont généralement assez nombreuses, de ce fait la partition obtenue n'est que rarement exploitable commodément.

Dans la mesure où nous souhaitons obtenir en un essai, une partition de bonne

qualité, nous avons choisi d'utiliser pour sélectionner les n noyaux initiaux (au prix d'une heuristique assez lourde), la démarche suivante:

Le centre de gravité du nuage G est choisi arbitrairement comme premier noyau; et on choisit les autres noyaux i en cherchant une solution au problème d'optimisation suivant :

$$\text{Max } \left[\sum_{i=2}^k \sum_{l=1}^i d(i,l) \right] / \{ (Nb_max - k + 1) \} * \{ k.(k-1) \}$$

$$k \leq n = Nb_max_de_classes_imposé \quad (50)$$

$$\forall i \quad 1 < i \leq n \quad d(i,G) \leq dmax$$

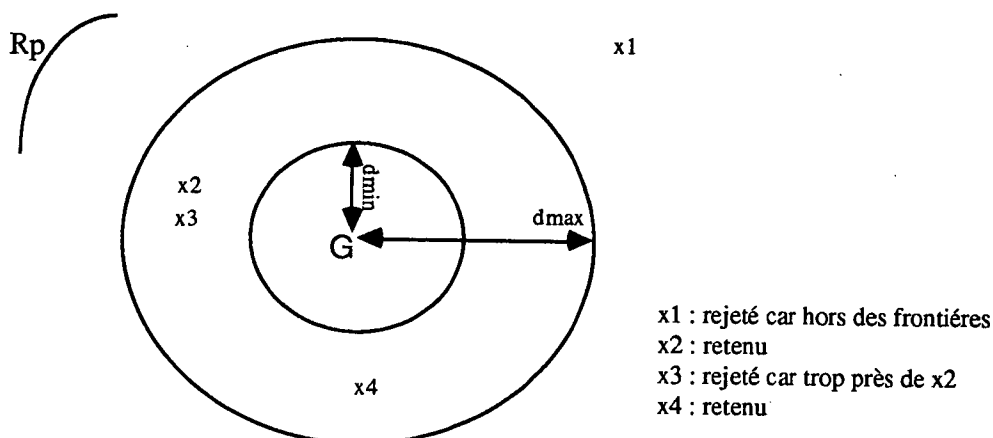
$$\forall i \quad 1 < i \leq n \quad d(i,G) \geq dmin$$

$$\forall i,j \quad i \neq j \quad d(i,j) \leq dmax$$

$$\forall i,j \quad i \neq j \quad d(i,j) \geq dmin$$

Géométriquement cette stratégie de sélection des noyaux de départ revient à conserver comme noyaux de départ, des points situés entre deux hypersphères dont les rayons sont $dmin$ et $dmax$; en leur imposant de surcroît une contrainte de non proximité.

Schéma de sélection des noyaux



Pendant ce mode de sélection impose le choix à priori des rayons de sélection. Pour ce faire on tire au hasard un échantillon d'au plus deux cents éléments de l'ensemble des objets à classer que l'on trie, $dmin$ est alors choisi comme premier quartile, $dmax$ comme troisième.

Les deux termes $A = \{ (Nb_max - k + 1) \}$ et $B = \{ k.(k-1) \}$ avec lesquels on divise la fonction à maximiser ont pour but :

1- pour A : d'obliger la sélection des Nb_max noyaux initiaux si cela est possible (populations très hétérogènes). C'est un moyen efficace de prendre en compte la première contrainte sur le nombre de noyaux à sélectionner.

2- pour B : de limiter le nombre de noyaux à introduire si contrairement au cas précédent, la population possède beaucoup de points identiques (ou très voisins) ou encore une structure en classes très forte. La quantité $2 * \left[\sum_{i=2}^k \sum_{l=1}^i d(i,l) \right] / \{ k.(k-1) \}$ représente en effet la moyenne des distances entre les k noyaux sélectionnés. Dans une population très facilement classifiable l'introduction d'un noyau supplémentaire peut alors faire décroître cette quantité de façon sensible, ceci permet alors d'empêcher la sélection de noyaux inutiles.

Enfin, il est certain que cette stratégie de sélection des noyaux pourrait être fort coûteuse en temps cpu avec de grands ensembles de données; ce qui serait évidemment désastreux. Cette étape n'étant qu'un préliminaire, on a choisi de limiter la "résolution" ou plus précisément la recherche de la solution de ce problème à un échantillon -tiré au hasard- d'au plus mille objets à classer.

C.f : Les distances utilisables

Dans tout problème de classification le choix de la mesure de ressemblance entre les objets à classer est fondamental; MNDOPT en propose plusieurs il convient alors de préciser leurs avantages, leurs inconvénients et leurs champs d'applications.

On peut dans MNDOPT, utiliser les métriques suivantes:

- la métrique identité : I qui favorise cependant les variables qui prennent de grandes valeurs, ceci malgré l'opération de centrage.
- la métrique $1/\sigma^2 I$ dont l'opération de centrage et réduction des données rend les variables aussi discriminantes les unes que les autres.

Ces deux métriques (distances Euclidiennes) sont très intéressantes, cependant elles ne permettent pas de prendre en compte les corrélations qui peuvent exister dans les données (métriques diagonales). De ce fait nous avons introduit la possibilité d'utiliser la distance de Mahalanobis afin de faire apparaître des classes plus discriminantes.

Ainsi dans MNDOPT on peut utiliser également:

- la métrique V^{-1} (métrique de Mahalanobis) qui, en prenant en compte les corrélations qui peuvent exister entre les variables; produit des classes généralement non sphériques mais plus allongées.
- la métrique W^{-1} (*distances adaptatives globales*) qui, plus encore que la métrique de Mahalanobis; produit des classes allongées. Cette métrique n'est cependant utilisée que pendant la phase d'optimisation de la partition; la phase d'apprentissage est quant à elle réalisée avec la métrique de Mahalanobis. Rien n'empêche d'ailleurs d'utiliser une partition d'apprentissage obtenue avec les métriques I ou $1/\sigma^2 I$, car la phase de détection du nombre de classes sera quand même réalisée avec la métrique de Mahalanobis V^{-1} , la partition obtenue par coupure de l'arbre hiérarchique à un niveau jugé significatif soulignera déjà des directions d'allongement que la phase d'optimisation réalisée avec la métrique W^{-1} ne fera que préciser.

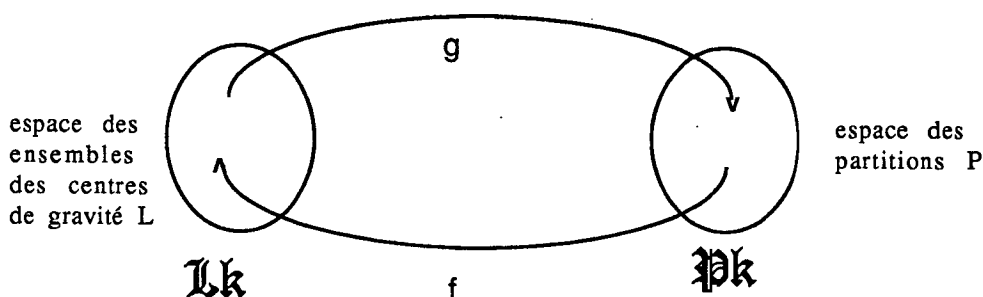
Nous avons souligné précédemment l'importance de la procédure PONDER dans le cadre d'une classification utilisant les distances adaptatives globales avec MNDOPT ou locales avec DISLOC. Il nous semble intéressant ici de préciser cela et pour ce faire nous allons présenter d'une manière un peu plus théorique la Méthode des Nuées Dynamiques dans sa version du centre de gravité (algorithme des centres mobiles) avec distances adaptatives.

La caractéristique principale de la Méthode des Nuées Dynamiques (M.N.D) est la problématique qu'elle introduit. Il s'agit en effet d'optimiser un critère W qui exprime l'adéquation entre une classification des objets et un mode de représentation des classes de cette partition. Le problème d'optimisation se pose alors en terme de "recherche simultanée de la partition $P = (P_1, P_2, \dots, P_k)$ dans l'espace des partitions \mathcal{P}_k et de la représentation de cette partition $L = (L_1, L_2, \dots, L_k)$ (L_i centre de gravité de la classe i) dans l'espace de représentations des partitions \mathcal{L}_k , qui aient la meilleure représentation

possible au sens du critère W qui est en fait une application de $\mathcal{L}_k \times \mathcal{P}_k$ dans \mathbb{R}^+ . Dans l'algorithme des centres mobiles, le critère W est l'inertie intra-classes qui est bien strictement positive et dont une petite valeur exprime une bonne adéquation entre L et P .

Pour résoudre ce problème de minimisation, l'algorithme des Nuées Dynamiques consiste alors à utiliser de façon itérative pour faire décroître le critère W jusqu'à stabilisation; ce qui donne une solution localement optimale au problème posé, une fonction de représentation $g : \mathcal{P}_k \rightarrow \mathcal{L}_k$ qui permet d'associer à une partition de \mathcal{P}_k un élément de \mathcal{L}_k ; et une fonction d'affectation $f : \mathcal{L}_k \rightarrow \mathcal{P}_k$ qui permet d'associer à un élément de \mathcal{L}_k une partition de \mathcal{P}_k .

Schéma de l'algorithme des Nuées Dynamiques



Pour initialiser l'algorithme des Nuées Dynamiques on utilise principalement une étape de représentation L_0 généralement tirée au hasard, ce qui n'est pas le cas dans MNDOPT (voir : C.e), ou secondairement une partition P_0 .

Le critère W s'écrit sous la forme $W(L,P) = \sum_{i=1}^k D(L_i, P_i)$ où D mesure l'adéquation entre le noyau (centre de gravité) L_i et la classe P_i , $L=(L_1, L_2, \dots, L_k)$ et $P = (P_1, P_2, \dots, P_k)$. On considère alors que l'application D s'exprime par :

$$\forall \lambda \in \mathcal{L}, \forall A \in \mathcal{P}, D(A, \lambda) = \sum_{w \in A} d(w, \lambda).$$

En général d mesure la "dissemblance" (souvent la distance) entre les éléments de I et un élément de \mathcal{L} . De ce fait le critère W s'écrit aussi :

$$\forall L \in \mathcal{L}_k, \forall P \in \mathcal{P}_k, W(L,P) = \sum_{i=1}^k D(L_i, P_i) = \sum_{i=1}^k \sum_{w \in P_i} d(w, L_i)$$

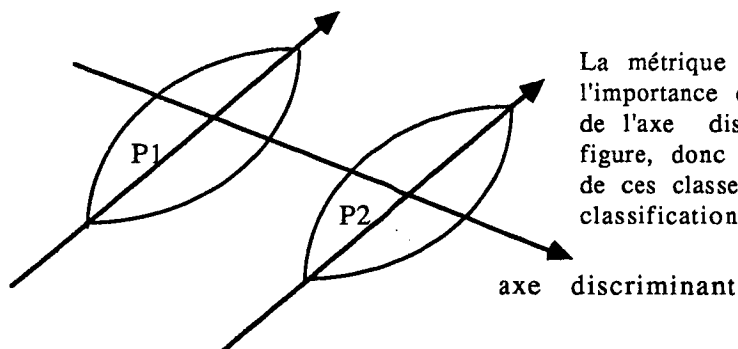
$L=(L_1, L_2, \dots, L_k)$ et $P = (P_1, P_2, \dots, P_k)$.

Lorsque les objets de l'ensemble I à classer sont décrits par p variables quantitatives, ils sont caractérisés par des vecteurs de \mathbb{R}^p que l'on muni généralement d'une métrique Euclidienne M (I ou $1/\sigma^2 I$); dans ce cas d est la distance Euclidienne usuelle.

$$\text{Ainsi } \forall (x,y) \in \mathbb{R}^p \quad d^2_M(x,y) = (x-y)^t M (x-y).$$

L'un des algorithmes de partitionnement basé sur la M.N.D les plus employés est celui des centres mobiles dont la popularité provient de sa simplicité et de sa rapidité

même sur de grands tableaux. Cependant on constate expérimentalement que ce choix de distance Euclidienne a tendance à donner des classes sphériques de mêmes volumes ce qui limite quelque peu ses champs d'applications. Pour lever ce problème on peut prendre $M=V^{-1}$ (matrice inverse des covariances) qui permet de reconnaître des classes de même allongement.



Certains auteurs se sont attachés à dépasser encore cette limitation de l'algorithme des centres mobiles (FRIEDMAN et RUBIN 1967 ou GOVAERT dans DIDAY 1979). Pour ce faire ils ont supprimé l'astreinte pour la mesure de ressemblance d'être fixe tout au long de l'algorithme tout en restant dans le cadre quadratique de façon à obtenir une mesure de dissemblance qui s'adapte aux données traitées.

L'objectif d'une distance adaptative unique (métrique W^{-1} dans MNDOPT) est de reconnaître des classes ayant le même type de dispersion mais possédant des directions d'allongement inconnues (cf figure précédente). Les deux classes à reconnaître sont assimilables à des ellipses de mêmes directions principales. Ce type de classes ne peut pas être reconnues par l'algorithme des centres mobiles utilisant les métriques Euclidiennes pour peu qu'elle soient un peu proches. Pour reconnaître ce genre de structure on peut essayer la métrique V^{-1} , malheureusement elle ne conduira pas toujours au résultat escompté.

La solution proposée par GOVAERT (cf : DIDAY 1979) notamment, consiste à associer à une partition une métrique de telle sorte que la somme des distances de chaque individu de I au centre de gravité de sa classe soit la plus petite possible. Dans ce cadre une partition P de l'ensemble des éléments à classer I sera représentée par k centres de gravité g_1, \dots, g_k et une métrique M de déterminant 1 (la métrique ne doit pas prendre une forme dégénérée).

Notons \mathcal{M} l'espace des matrices carrées symétriques définies positives de dimension p et de déterminant 1. L'espace de représentation est ici $\mathcal{L}_k = (\mathbb{R}^p)^k \times \mathcal{M}$ (on n'utilise pas une métrique Euclidienne).

Le critère $W : \mathcal{L}_k \times \mathcal{P}_k \rightarrow \mathbb{R}^+$ est défini par :

$$\forall L \in \mathcal{L}_k, \forall P \in \mathcal{P}_k, W(L, P) = \sum_{i=1}^k \sum_{x_i \in P_i} d^2_{M}(x_i, g_i)$$

où $L = (g_1, g_2, \dots, g_k, M)$ et $P = (P_1, P_2, \dots, P_k)$.

Il représente la somme des inerties de chaque classe par rapport à son centre de gravité au sens de la distance d_{M^*} .

La fonction d'affectation $f : \mathcal{L}_k \rightarrow \mathcal{P}_k$ est définie par:

La fonction f associée à $L = (g_1, g_2, \dots, g_k, M)$ la partition $P = (P_1, P_2, \dots, P_k)$ minimisant

$W(\cdot, L)$. On voit facilement que cette partition est obtenue de la façon suivante:

$$\forall i=1, k, P_i = \{x_i \in I / d_M(x_i, g_i) < d_M(x_i, g_m) \text{ avec } l > m \text{ en cas d'égalité}\}.$$

Pourquoi $l > m$ en cas d'égalité ? Dans MNDOPT les classes sont triées par effectif décroissant cela permet d'éviter partiellement la dégénérescence de la métrique.

La fonction de représentation $g : \mathcal{P}_k \rightarrow \mathcal{L}_k$ est définie par:

La fonction g associée à la partition $P=(P_1, P_2, \dots, P_k)$ la représentation $L=(g_1, g_2, \dots, g_k, M)$

de manière à minimiser le critère $W(P, \cdot)$. Il est facile de voir que $\forall i=1, k, g_i$ est le centre de gravité (algorithme des centres mobiles) de la classe P_i .

La matrice M : la détermination de M est plus complexe. On cherche $M \in \mathcal{M}$ telle que:

$$W(L, P) = \sum_{i=1}^k \sum_{x_i \in P_i} (x_i - g_i)^t M (x_i - g_i) \text{ soit minimum.}$$

Il s'agit du même problème que l'on rencontre en Analyse Discriminante Linéaire pour choisir la métrique la plus discriminante.

La solution est fournie par la matrice M définie par $M = \det(W)^{1/p} \cdot W^{-1}$

où $W = \sum_{i=1}^k \sum_{x_i \in P_i} (x_i - g_i)^t (x_i - g_i)$; W est la matrice d'inertie intra-classes associée à la partition P .

Présentons succinctement la démonstration de ce résultat. Mis sous forme matricielle le problème revient à chercher la distance quadratique d définie par M telle que: $\det(M)=1$ et $\sum_x d(x, g)$ soit minimum (g centre de gravité de l'ensemble des points x).

$$\text{On a } \sum_x d(x, g) = \sum_y Y^t M Y \text{ avec } Y = X - G$$

La quantité $Y^t M Y$ est un scalaire; d'où $Y^t M Y = \text{trace}(Y^t M Y) = \text{trace}(Y^t Y M)$ et enfin $\sum_y Y^t M Y = \text{trace}((\sum_y Y^t Y) M)$.

On note $Q = \sum_y Y^t Y$ et $R = Q M$. On est donc ramené à trouver la matrice R telle que $\text{trace}(R)$ soit minimum avec $\det(R) = \det(Q) = \alpha$ (par hypothèse $\det(M)=1$).

On sait que R est symétrique (Q et M sont symétriques) et réelle donc diagonalisable, donc $\det(R) = \prod$ (valeurs propres de R). Le produit des valeurs propres étant fixé, pour minimiser la somme $\sum_y Y^t M Y$, il faut que toutes les valeurs propres de R soient égales à $\alpha^{1/p}$ ($\alpha = \det(Q)$). R ne peut donc être que la matrice $\alpha^{1/p} I_p$.

$$\text{Enfin } M = \alpha^{1/p} Q^{-1} = \det(Q)^{1/p} Q^{-1} \text{ avec } Q = \sum_y Y^t Y.$$

L'algorithme des Nuées Dynamiques (version des centres mobiles) ainsi défini fait décroître le critère d'inertie W à chaque itération sous réserve que la matrice W donc la métrique M soit bien conditionnée.

Cependant les méthodes basées sur les distances adaptatives sont de loin les procédures les plus coûteuses car elles imposent l'inversion de la matrice de covariances intra-classes (dimension $p \times p$) à chaque itération dans MNDOPT ou encore pire k inversions par itération dans DISLOC (la métrique est W^{-1}). D'autre part pour obtenir des résultats assez fiables, le nombre d'individus doit être assez grand par rapport à p (nombre de

variables) et k (nombre de classes). On comprend alors le problème que posent les points aberrants tant d'un point de vue numérique (matrice de covariances non de plein rang donc non inversible) que théorique (dégénérescence de la métrique). C'est alors que la procédure PONDER prend toute son importance car elle permet d'être pratiquement certain que l'on ne pourra plus créer de classe de faible effectif.

Cependant qu'elle métrique employer dans PONDER pour un filtrage efficace avant une méthode basée sur les distances adaptatives ? Dans DISLOC on effectue trois itérations de l'algorithme des centres mobiles en utilisant la métrique I avant de commencer l'algorithme des distances adaptatives; ceci uniquement parce que ce type de procédure donne des résultats qui dépendent beaucoup plus de l'initialisation que d'autres algorithmes plus simples. Ce choix ne nous semble pas cohérent, nous avons précisé que les métriques Euclidiennes tendent à donner des classes sphériques contrairement à l'objectif poursuivi dans les distances adaptatives. De ce fait il nous a semblé plus logique d'initialiser la procédure pendant la recherche des points aberrants avec PONDER ou pendant celle de recherche du nombre de classes dans MNDOPT, par une métrique fixe qui permette aussi de reconnaître des classes de même allongement. C'est pourquoi nous avons introduit la possibilité d'utiliser la métrique V^{-1} dans ces deux méthodes.

Remarque: Dans MNDOPT, pour les distances adaptatives on peut fort bien partir d'une partition d'apprentissage lue sur l'archive et obtenue avec les métriques Euclidiennes car la phase de recherche du nombre de classes sera quand même réalisée avec la métrique V^{-1} (C'est ce que nous avons fait dans l'exemple II.B). Certes la hiérarchie ne donnera pas des pourcentages d'inertie expliquée considérables à cause du changement de métrique sous jacent; mais la partition obtenue par coupure de l'arbre hiérarchique présentera quand même des directions d'allongement. L'initialisation des distances adaptatives sera ainsi cohérente avec l'objectif poursuivi par la procédure (recherche de classes allongées).

C.g: Les résultats fournis par MNDOPT

L'objectif principal poursuivi dans la commande MNDOPT est l'obtention d'une bonne partition sans contraindre l'utilisateur.

Pour la phase de détection automatique, le programme imprime les valeurs des critères de décision et en représente les graphes. L'utilisateur peut alors, après examen de ceux-ci, fort bien recommencer la classification en imposant alors lui-même un autre nombre de classes. Il n'est plus alors obligé de recommencer la phase d'apprentissage.

A la fin de la classification on décrit succinctement la partition obtenue en mesurant la dispersion des classes par rapport au centre de gravité du nuage avec les indices suivants:

$$\text{Dispersion (var } \langle j \rangle) = (\text{Moycl } \langle j \rangle - \text{Moy } \langle j \rangle) / \text{Et } \langle j \rangle$$

avec :

- var $\langle j \rangle$: j ième variable active de la classification
- Moycl $\langle j \rangle$: moyenne de la j ième variable dans la classe
- Moy $\langle j \rangle$: moyenne de la j ième variable
- Et $\langle j \rangle$: écart-type de la j ième variable

Remarque: Il faut noter que ces indices de dispersion ne sont pas réellement bien adaptés aux partitions issues des distances adaptatives puisque les critères optimisés ne sont pas au sens de la métrique Euclidienne usuelle.

C.h : Aspects informatiques

La place mémoire occupée par ce programme dépend du nombre d'observations et du

nombre de variables. Le temps calcul dépend lui aussi de l'importance de l'échantillon à classer. Cependant les algorithmes ont été optimisés : la complexité de ce programme est en $N.V$ avec les métriques Euclidiennes et en $N.V^2$ avec les métriques de Mahalanobis.

C'est surtout la phase d'apprentissage qui est très coûteuse, mais n'oublions pas que l'on peut fort bien ne l'effectuer ne l'exécuter qu'une seule fois au cours de l'étude.

Pendant la troisième phase, l'algorithme procède par améliorations successives de la qualité de la partition initiale. Il est naturellement évident que le nombre de possibilités à explorer systématiquement croît rapidement avec la taille de l'échantillon et bien sûr avec lui le coût d'exploitation. Pour compenser ce défaut des procédures visant à optimiser les transferts ont été introduites dans le programme, ce qui permet une exécution efficace.

Tous les calculs sont faits en double précision.

Selon le nombre d'observations à classer, le programme travaille en mémoire centrale ou en lecture directe (version séquentialisée paginée).

Les programmes PONDER et MNDOPT sont écrits en Fortran77 aux normes MODULAD, et ont été intégrés dans le logiciel SICLA développé à l'INRIA.

DEUXIEME PARTIE

Application de PONDER et de MNDOPT à l'analyse de l'effet du renforcement de la ligne Morbras-Villevaudé du réseau de transport à très haute tension d'EDF.

A: Analyse sans filtrage des surcoûts du renforcement métrique $1/\sigma^2 I$

Les variables retenues pour celle ci furent les surcoûts d'exploitation nucléaire (DCNU), charbon (DCCH), hydraulique (DCHY), fioul et gaz (DCFG) et de défaillance (DCDF). Il s'agit des deltas de coûts d'exploitation avec réseau. Ainsi la variable DCNU représente l'écart entre le coût d'exploitation du Nucléaire avec réseau (pour le réseau non renforcé) et le coût d'exploitation Nucléaire avec réseau (pour le réseau renforcé).

Le renforcement de la liaison Morbr172Vleva sera alors d'autant plus intéressant que pour chaque situation étudiée le surcoût d'exploitation global (nucléaire + charbon + hydraulique + fioul+ défaillance) sera important.

La classification montre les résultats qui suivent.

Le graphe de la statistique locale souligne très nettement une partition en 6 ou 12 classes. On constate en effet un maximum absolu pour ce critère, à 12 classes. Résultat confirmé par le graphe de l'inertie expliquée qui souligne des gains d'inertie expliquée très minimes au-delà. Le critère cubique de classification favorise plutôt une partition en 7 classes. Finalement la procédure retient une partition en 6 classes.

Décrivons maintenant la partition obtenue après cinq étapes d'améliorations de la partition obtenue par coupure de l'arbre hiérarchique; les aides à l'interprétation incluses dans la procédure vont nous y aider.

L'examen des graphes des indices de dispersion de chaque variable active de cette classification sur les surcoûts d'exploitation du au renforcement, montre dans les classes les faits qui suivent :

La **classe 1** comporte 1330 aléas; où l'on trouve les situations où le surcoût d'exploitation global est en moyenne de 5.1 kf. Il s'agit des situations où le renforcement de la liaison Morbr172Vleva n'a presque rien changé.

La **classe 2** comporte 125 aléas; où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 117 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant celle du charbon.

La **classe 3** comporte 42 aléas; où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 258 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du charbon disponible en réduisant celle du fioul.

La **classe 4** comporte 2 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 87 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du charbon disponible et de l'hydraulique en réduisant celle du fioul.

La **classe 5** comporte 2 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 291 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant très fortement celles de l'hydraulique et du fioul.

La classe 6 met en évidence l'aléa où le surcoût d'exploitation global est en moyenne de 964 kf. On trouve dans cette classe un élément de défaillance. Le renforcement de la liaison Morbrl72Vlewa a permis de supprimer un cas de défaillance réseau par une meilleure valorisation du nucléaire et du charbon.

On notera quand même que la procédure de classification MNDOPT est également capable de reconnaître des éléments aberrants. Elle a sélectionnée une partition en six classes dont les trois dernières ont un effectif très faible.

Le lecteur trouvera dans les quelques pages qui suivent une représentation graphique tridimensionnelle des surcoûts d'exploitation, qui montre la classification dans les espaces :

($x=DCNU$, $y=DCCH$, $z=DCFG$) et ($x=DCNU$, $y=DCCH$, $z=DCHY$)

et une représentation graphique qui montre les divers paliers du surcoût d'exploitation du thermique global du renforcement dans l'espace :

($x=DCNU$, $y=DCCH$, $z=DCEX$).

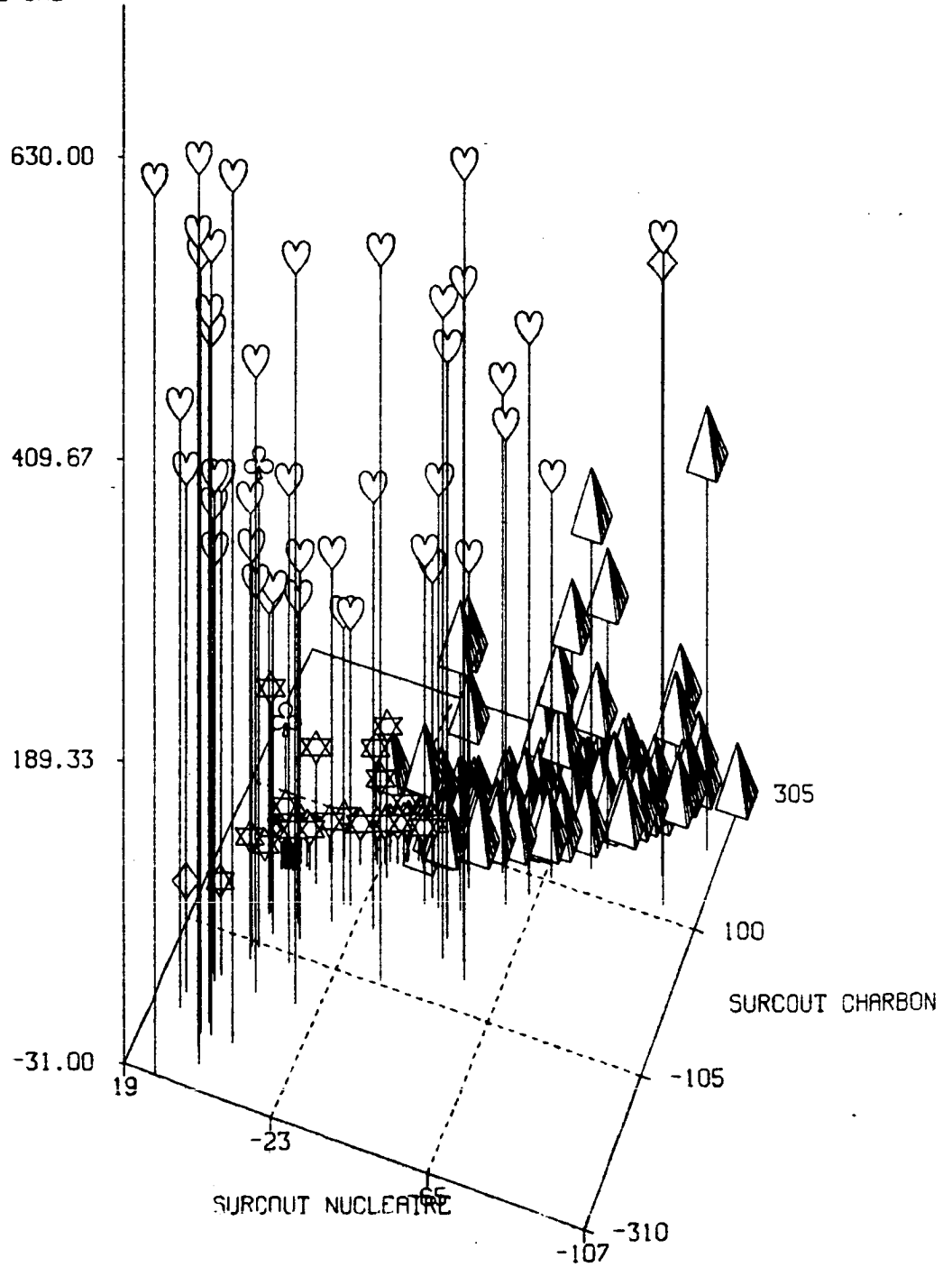
On utilise la légende suivante pour repérer les différentes classes :

- 1 : étoile => (sans changement)
- 2 : pyramide => (nucléaire / charbon)
- 3 : cœur => (charbon / fioul fort)
- 4 : trèfle => (hydraulique & charbon / fioul fort)
- 5 : carreau => (nucléaire / hydraulique & fioul fort)
- 6 : pique => (défaillance)

La défaillance n'est pas représentée sur les graphiques pour des raisons de lisibilité. Pour des raisons de temps calcul tous les points n'ont pas été représentés.

ANALYSE DU RENFORCEMENT (I/S2)

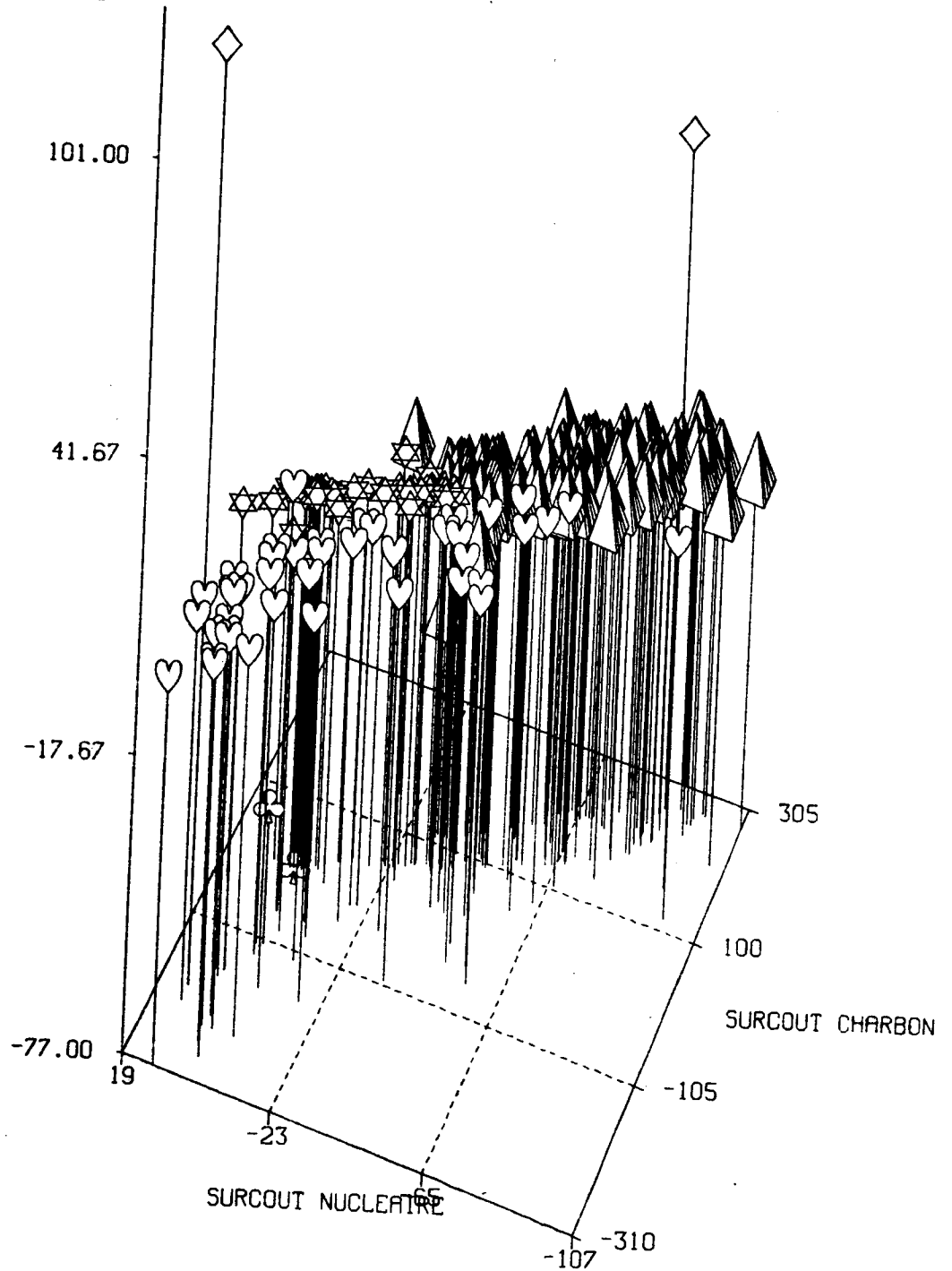
SURCOUT FIOUL & GAZ



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

ANALYSE DU RENFORCEMENT (I/S2)

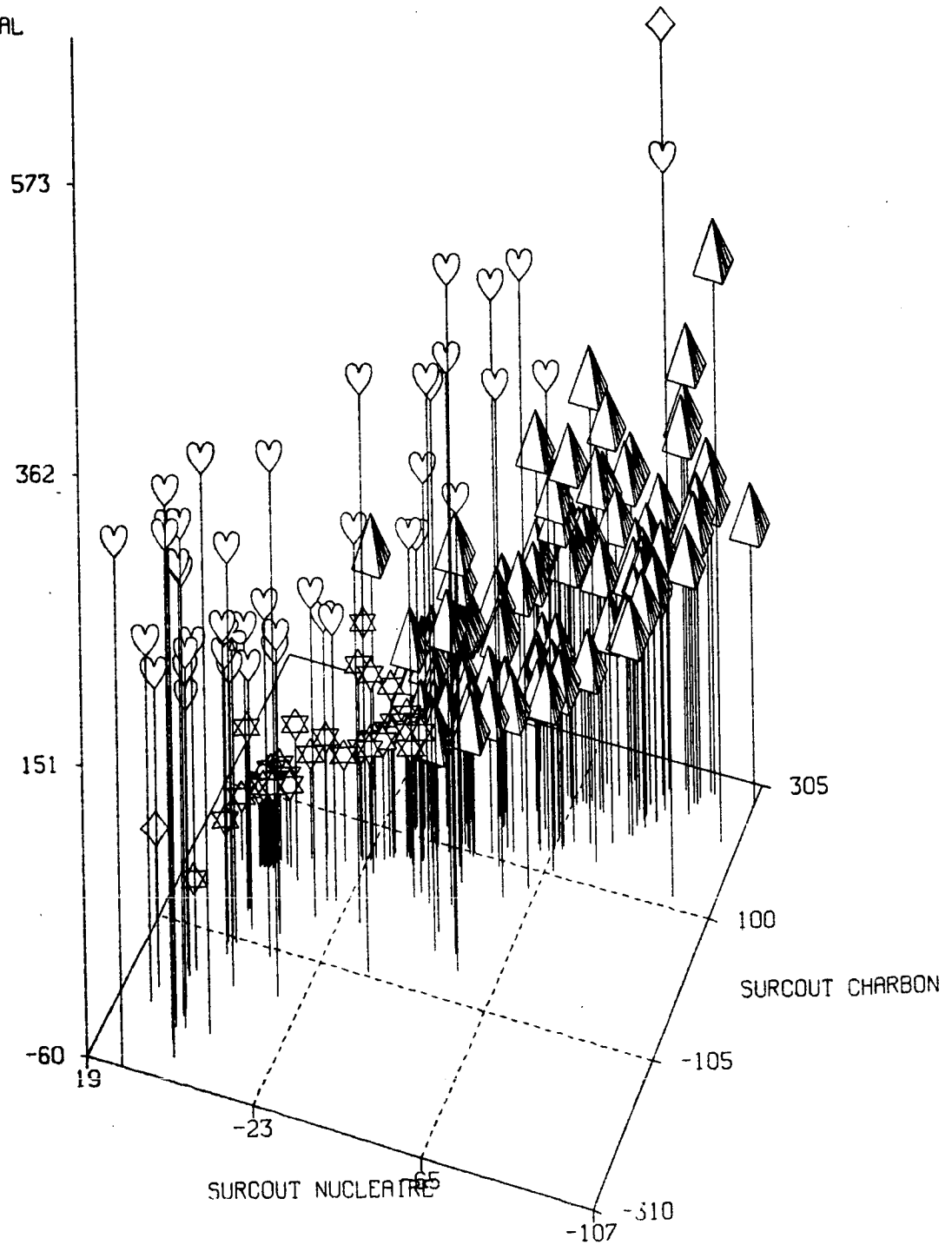
SURCOUT HYDRAULIQUE



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

ANALYSE DU RENFORCEMENT (I/S2)

SURCOUT GLOBAL



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

B: Analyse sans filtrage des surcoûts du renforcement (métrique W^{-1})

La classification montre les résultats qui suivent.

La phase d'apprentissage est effectuée avec la métrique $1/\sigma^2 I$ (il s'agit de la partition d'apprentissage de la précédente classification avec métrique $1/\sigma^2 I$). On remarquera sur la matrice de covariances (inverse de la métrique) que le nucléaire est corrélé négativement avec tous les autres combustibles. Néanmoins la phase de détection du nombre de classes est quant à elle réalisée avec la métrique V^{-1} .

Le graphe de la statistique locale souligne très nettement une partition en 8 ou 13 classes. On constate en effet un maximum absolu pour ce critère, à 13 classes et un maximum relatif à 8 classes. Résultat confirmé par le graphe de l'inertie expliquée qui souligne des gains d'inertie expliquée assez minimes au-delà. Le critère cubique de classification favorise plutôt une partition en 9 classes. Finalement la procédure retient une partition en 8 classes.

On remarquera que l'utilisation d'une partition d'apprentissage obtenue avec la métrique $1/\sigma^2 I$ pour la hiérarchie ne donne pas d'excellents pourcentages d'inertie expliquée (maximum à 77.81%).

Décrivons maintenant la partition obtenue après deux itérations utilisant la métrique W^{-1} pour la phase d'optimisation; les aides à l'interprétation incluses dans la procédure vont nous y aider. Rappelons encore, que celles-ci sont calculées avec la métrique Identité parce qu'il n'est pas possible de calculer un écart-type avec les métriques de Mahalanobis. Ces indices ne sont donc pas parfaitement adaptés à l'interprétation de partitions obtenues avec les distances adaptatives. Le pourcentage d'inertie expliquée de 99.4% est calculé avec la métrique W^{-1} ; il est excellent.

L'examen des graphes des indices de dispersion de chaque variable active de cette classification sur les surcoûts d'exploitation du au renforcement, montre dans les classes les faits qui suivent :

La classe 1 comporte 1320 aléas; où l'on trouve les situations où le surcoût d'exploitation global est en moyenne de 5.14 kf. Il s'agit des situations où le renforcement de la liaison Morbr172Vleva n'a presque rien changé. On constate cependant que la métrique W^{-1} a affecté dans cette classe d'éléments où les problèmes rencontrés se sont conservés, pratiquement autant de situations que la métrique $1/\sigma^2 I$.

La classe 2 (resp 3) comporte 67 aléas (resp 61 aléas); où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 79 kf (resp 139 kf). On trouve dans ces deux classes des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant celle du charbon. On remarque que les graphes de dispersion dans ces deux classes sont très voisins. cependant il existe une différence sensible sur le fioul.

La classe 4 comporte 32 aléas ; où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 232 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du charbon disponible (et un peu de nucléaire) en réduisant très fortement celle du fioul.

La classe 5 comporte 13 aléas ; où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 337 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant très fortement celle du fioul et également celle du charbon.

L'effet principal du renforcement de la liaison Morbr172Vleva se situe donc dans 173 situations qui représentent pratiquement l'essentiel du gain sur le coût d'exploitation.

La classification avec métrique W^{-1} permet de retrouver également les cinq situations que l'on peut qualifier d'aberrantes d'un point de vue statistique; mais pas du point de vue du planificateur. Il s'agit des classes 6,7 et 8. Ces situations sont si particulières qu'avec la métrique $1/\sigma^2 I$, il fallait choisir une partition finale en 5 classes afin que notre procédure MNDOPT de classification automatique commence à les faire disparaître.

La classe 6 comporte 2 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 87 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du charbon disponible et de l'hydraulique en réduisant celle du fioul.

La classe 7 comporte 2 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 291 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant très fortement celles de l'hydraulique et du fioul.

La classe 8 met en évidence 1 aléa où le surcoût d'exploitation global est en moyenne de 964 kf. On trouve dans cette classe un élément de défaillance. Le renforcement de la liaison Morbr172Vleva a permis de supprimer un cas de défaillance réseau par une meilleure valorisation du nucléaire et du charbon.

Le lecteur trouvera dans les quelques pages qui suivent une représentation graphique tridimensionnelle des surcoûts d'exploitation, qui montre la classification dans les espaces :

($x=DCNU$, $y=DCCH$, $z=DCFG$) et ($x=DCNU$, $y=DCCH$, $z=DCHY$)

et une représentation graphique qui montre les divers paliers du surcoût d'exploitation du thermique global dans l'espace :

($x=DCNU$, $y=DCCH$, $z=DCEX$).

On utilise la légende suivante pour repérer les différentes classes :

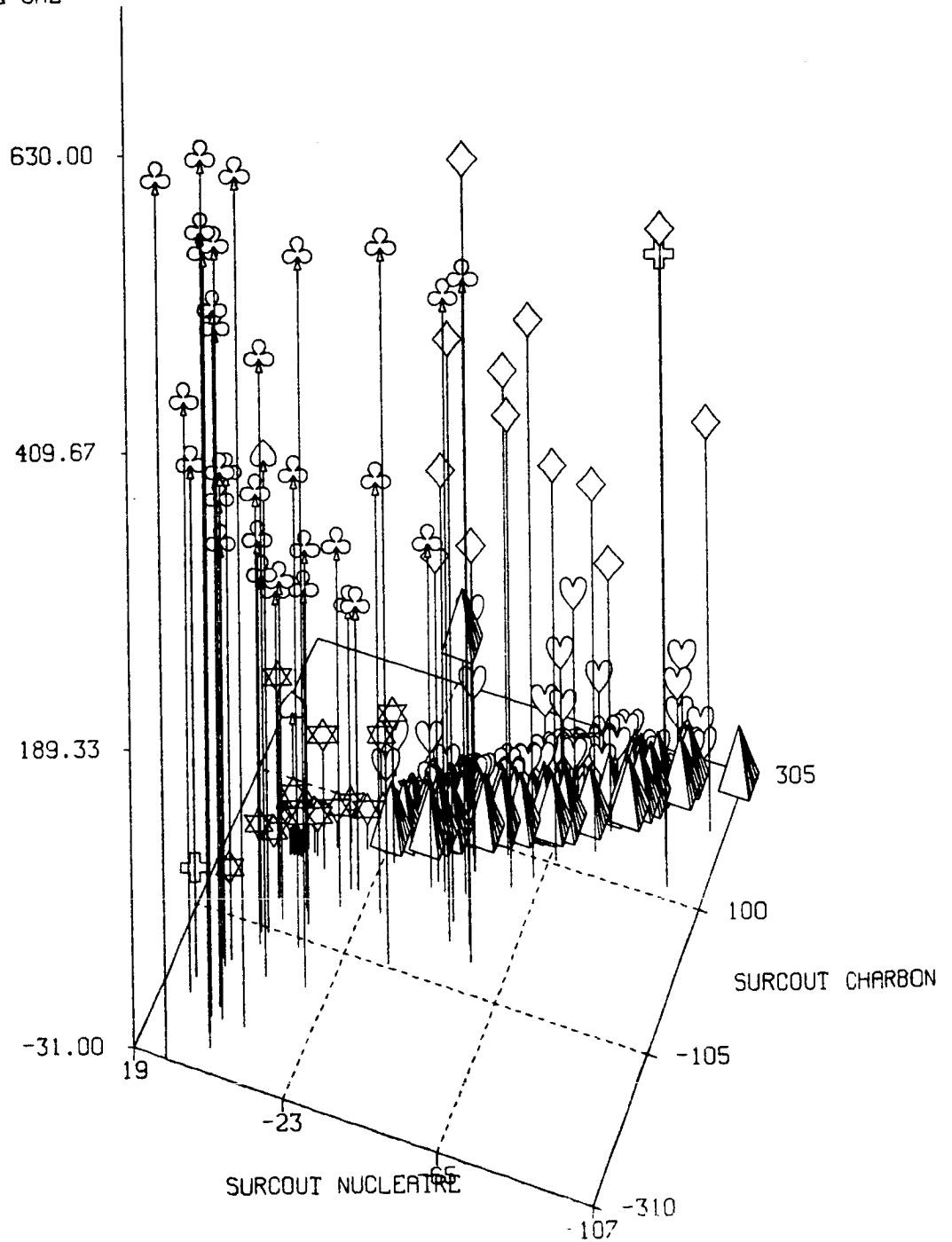
- 1 : étoile => (sans changement)
- 2 : pyramide => (nucléaire / charbon)
- 3 : cœur => (nucléaire / charbon + fioul faible)
- 4 : trèfle => (charbon / fioul fort)
- 5 : carreau => (nucléaire / fioul)
- 6 : pique => (hydraulique & charbon / fioul fort)
- 7 : croix => (nucléaire / hydraulique & fioul fort)
- 8 : carré => (défaillance)

La défaillance n'est pas représentée sur les graphiques pour des raisons de lisibilité.

NB : Comme précédemment pour des raisons de temps calcul, tous les points ne sont pas représentés.

ANALYSE DU RENFORCEMENT (W-1)

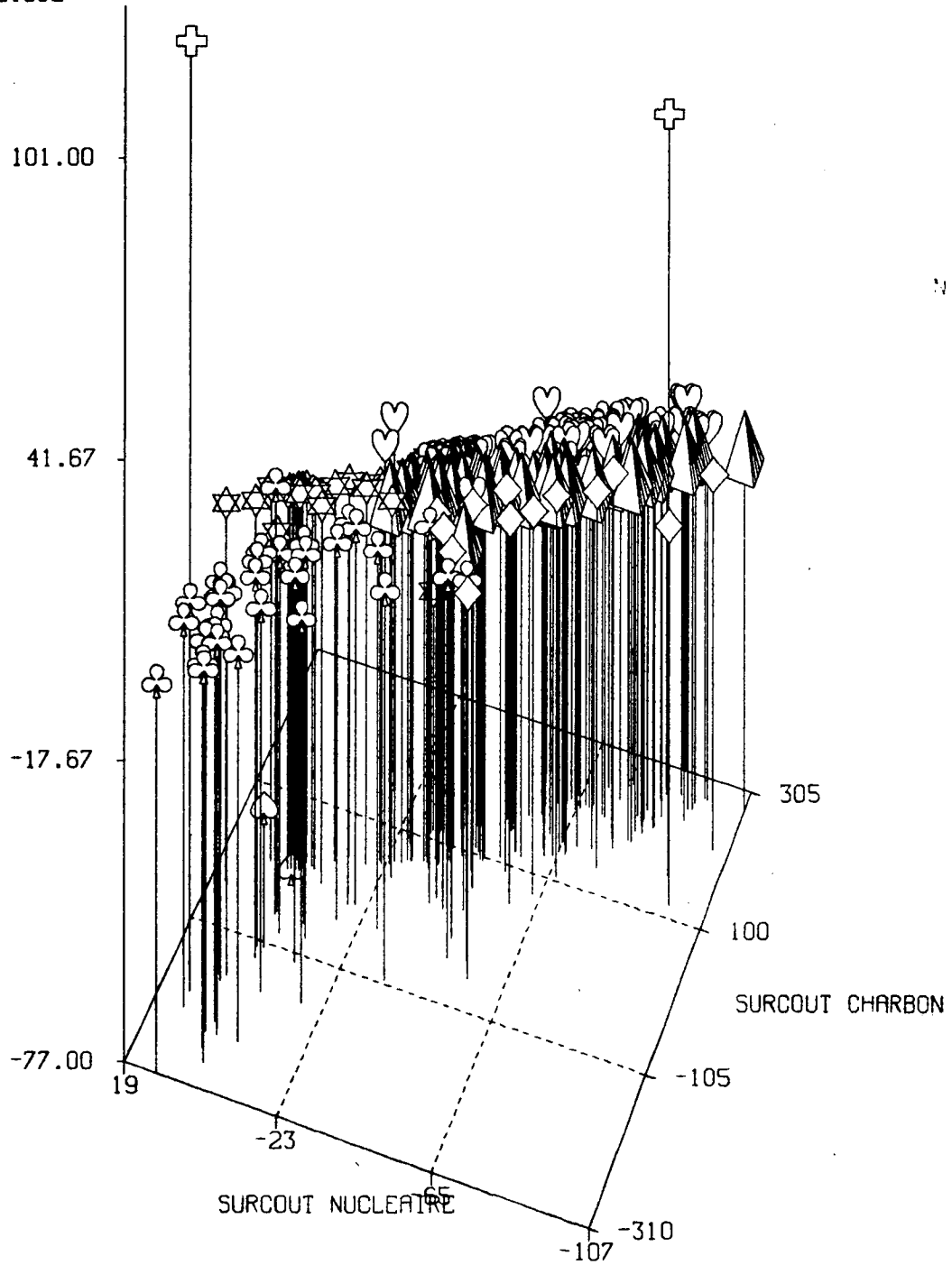
SURCOU FIOUL & GAZ



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

ANALYSE DU RENFORCEMENT (W-1)

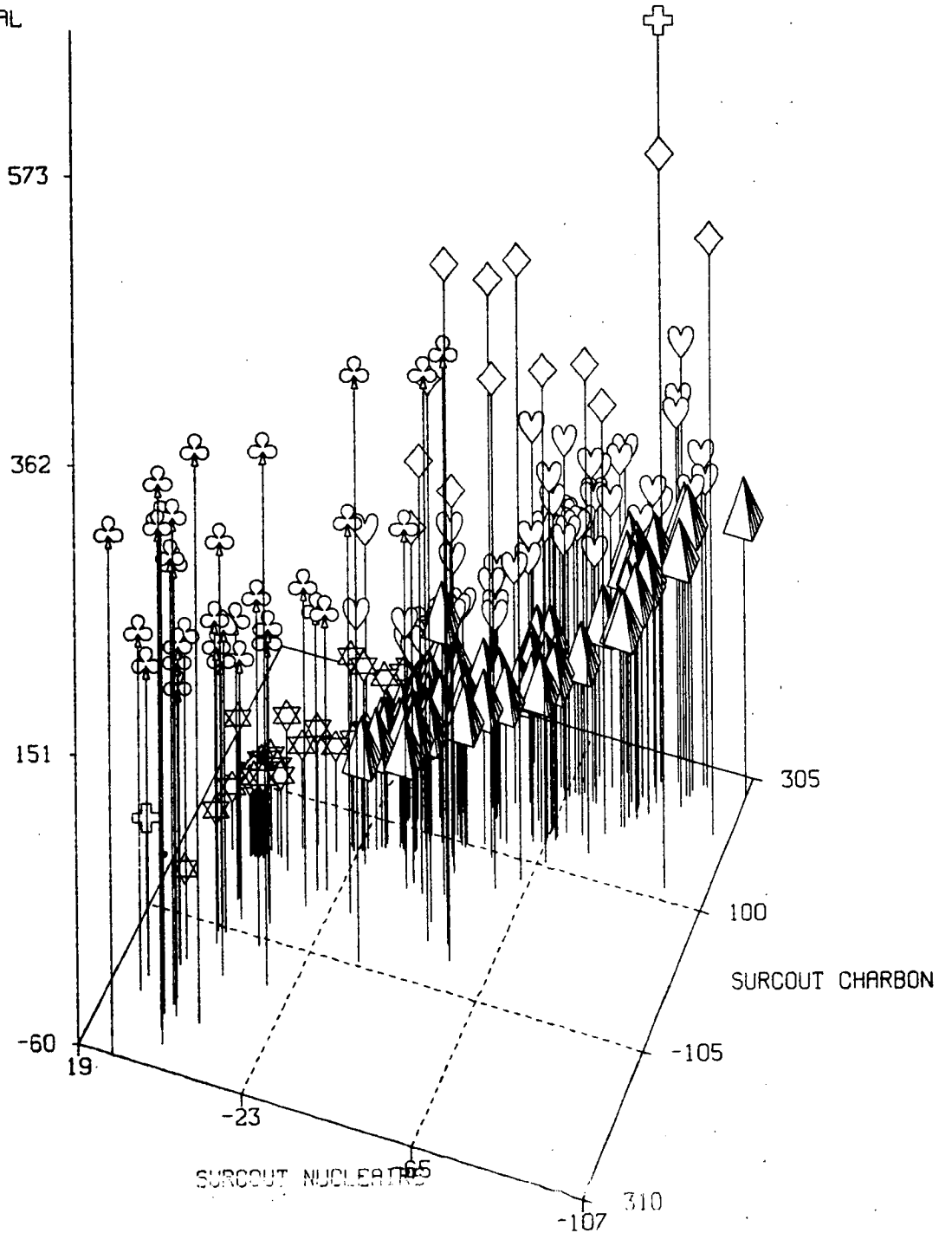
SURCOUT HYDRAULIQUE



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

ANALYSE DU RENFORCEMENT (W-1)

SURCOUT GLOBAL



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

C: Analyse avec filtrage des surcoûts du renforcement métrique $1/\sigma^2 I$

Comme précédemment, les variables retenues pour celle ci furent les surcoûts d'exploitation nucléaire (DCNU), charbon (DCCH), hydraulique (DCHY), fioul et gaz (DCFG) et de défaillance (DCDF).

La phase de filtrage avec PONDER s'est exécutée en cinq étapes, la dernière n'ayant pas donné lieu à la création de pondération.

La première étape de filtrage permet, d'isoler une situation de défaillance avant renforcement. Il s'agit de la situation 1404 où le renforcement a réduit le coût d'exploitation du thermique de 964 kf. En conservant la même métrique, PONDER permet de filtrer deux situations (1366) où le renforcement a diminué le coût d'exploitation du thermique de 291 kf (meilleure valorisation du nucléaire et diminution du fioul et de l'hydraulique au delà de l'hypothèse de turbinage prévue à coût nul). Enfin cette première étape permet de filtrer encore deux situations (1046) où le renforcement a diminué le coût d'exploitation du thermique de 87 kf (meilleure valorisation du charbon et de l'hydraulique et diminution du fioul).

La deuxième étape de filtrage permet, après recalcul de la métrique, d'isoler deux situations (101) où le renforcement a diminué le coût d'exploitation du thermique de 28 kf (meilleure valorisation du nucléaire et de l'hydraulique et diminution du charbon).

La troisième étape de filtrage permet, après recalcul de la métrique, d'isoler deux situations (452) où le renforcement a diminué le coût d'exploitation du thermique de 80 kf (meilleure valorisation du charbon et diminution de l'hydraulique et du fioul). Enfin cette troisième étape permet de filtrer encore trois situations (1483) où le renforcement a diminué le coût d'exploitation du thermique de 15 kf (meilleure valorisation de l'hydraulique et diminution du fioul).

La quatrième étape de filtrage permet, après recalcul de la métrique, d'isoler neuf situations (580) où le renforcement a diminué le coût d'exploitation du thermique de 60 kf (meilleure valorisation du charbon et de l'hydraulique et diminution du fioul). Cependant peut on encore considérer ces situations comme aberrantes ? Notre seuil critique avait été fixé à dix, une classe comportant neuf éléments est alors considérée comme formée d'éléments aberrants. Ceci montre bien que l'introduction de seuil critique est toujours un angoissant dilemme. Peut-être vaudrait il mieux se montrer moins sévère et prendre un seuil critique de cinq, et laisser le soin à la commande MNDOPT de repérer toute seule des classes de faibles effectifs puisqu'elle en est capable.

La classification montre les résultats qui suivent.

La phase d'apprentissage s'est effectuée en trois étapes. La partition finale est sauvegardée sur l'archive; elle extrait 97,4% d'inertie.

Le graphe de la statistique locale souligne très nettement une partition en 7 ou 13 classes. On constate en effet un maximum absolu pour ce critère, à 13 classes. Résultat confirmé par le graphe de l'inertie expliquée qui souligne des gains d'inertie expliquée très minimes au-delà. Le critère cubique de classification favorise plutôt une partition en 8 ou 10 classes. Finalement la procédure MNDOPT retient une partition en 7 classes.

Décrivons maintenant la partition obtenue; les aides à l'interprétation incluses dans la procédure vont nous y aider.

L'examen des graphes des indices de dispersion de chaque variable active de cette

classification sur les surcoûts d'exploitation du au renforcement, montre dans les classes les faits qui suivent :

La classe 1 comporte environ 1230 aléas; où l'on trouve les situations où le surcoût d'exploitation global est en moyenne de 8.29 kf. Il s'agit comme dans les précédentes classifications des situations où le renforcement de la liaison Morbrl72Vleva n'a presque rien changé.

La classe 2 (resp 3) comporte 129 aléas (resp 67 aléas); où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 54.2 kf (resp 145 kf). On trouve dans ces deux classes des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant celle du charbon.

La classe 4 (resp 5) comporte 19 aléas (resp 15 aléas); où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 216 kf (resp 101kf). On trouve dans ces deux classes des situations où le renforcement a permis une meilleure utilisation du charbon disponible en réduisant celle du fioul.

La classe 6 comporte 12 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 312 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire et du charbon disponible en réduisant celle du fioul.

La classe 7 comporte enfin 6 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 399 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant très fortement celle du fioul.

La partition après filtrage souligne donc des résultats parfaitement comparables à ceux obtenus avec la classification sans filtrage. Il nous semble cependant plus intéressant de filtrer les éléments peu représentatifs avant classification afin d'obtenir une partition beaucoup plus stable.

Le lecteur trouvera dans les quelques pages qui suivent une représentation graphique tridimensionnelle des surcoûts d'exploitation, qui montre la classification dans les espaces :

$(x=DCNU, y=DCCH, z=DCFG)$ et $(x=DCNU, y=DCCH, z=DCHY)$

et une représentation graphique qui montre les divers paliers du surcoût d'exploitation du thermique global dans l'espace :

$(x=DCNU, y=DCCH, z=DCEX)$.

On utilise la légende suivante pour repérer les différentes classes :

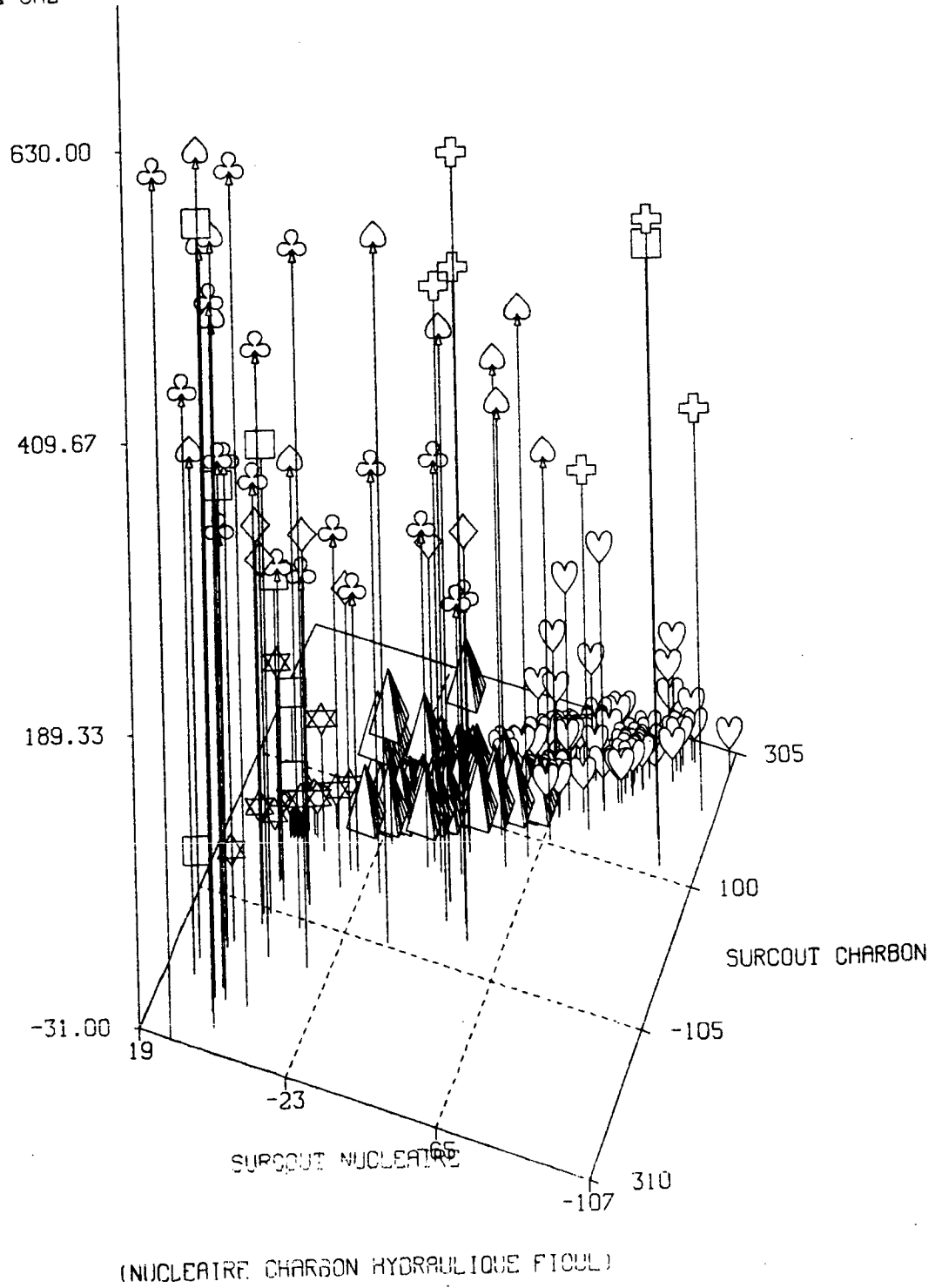
- 1 : étoile => (sans changement)
- 2 : pyramide => (nucléaire / charbon)
- 3 : cœur => (nucléaire / charbon)
- 4 : trèfle => (charbon / fioul fort)
- 5 : carreau => (charbon / fioul)
- 6 : pique => (nucléaire & charbon / fioul fort)
- 7 : croix => (nucléaire / fioul fort)
- 8 : carré => (points filtrés)

La défaillance n'est pas représentée sur les graphiques pour des raisons de lisibilité.

NB : Comme précédemment pour des raisons de temps calcul, tous les points ne sont pas représentés.

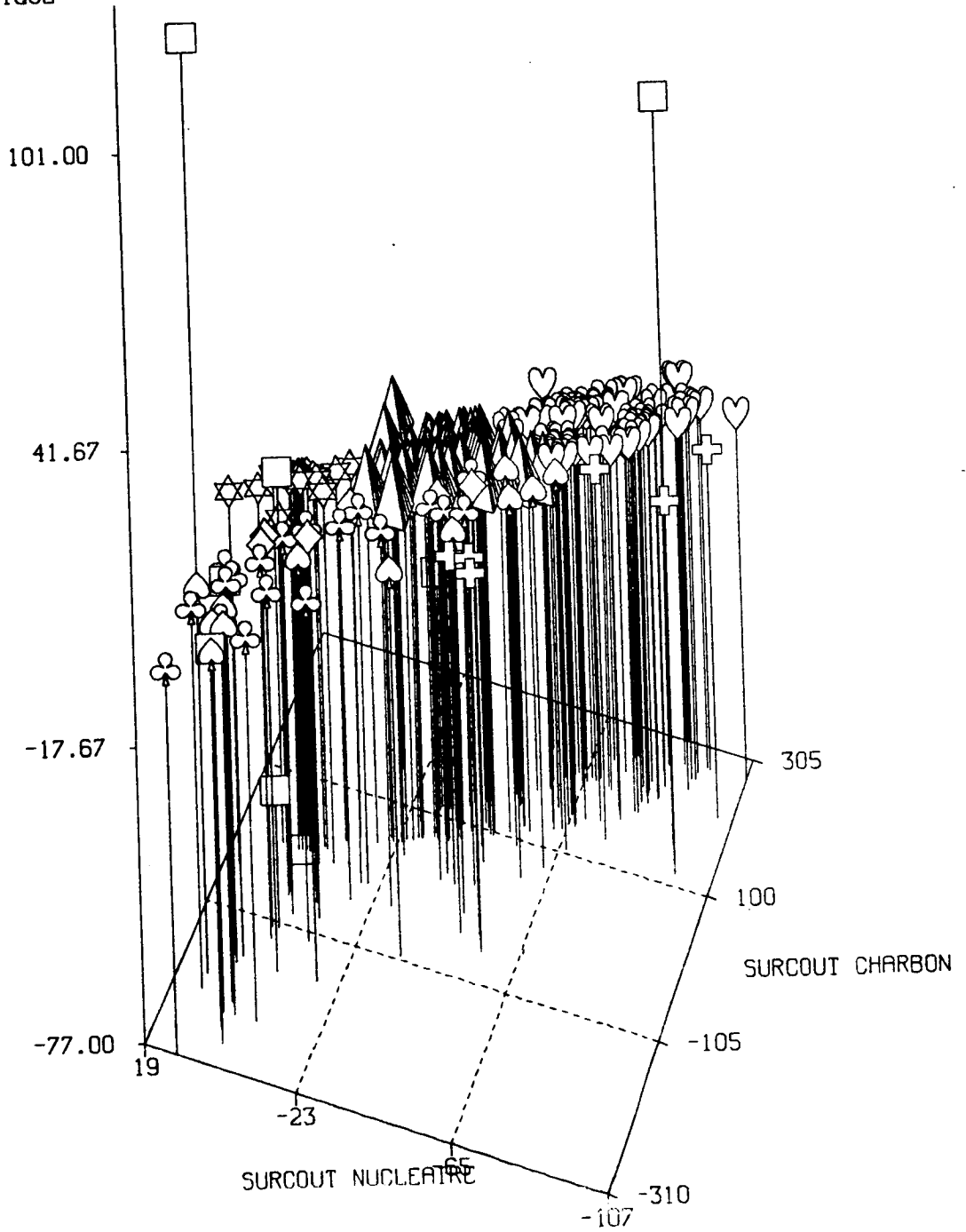
ANALYSE DU RENFORCEMENT (I/S2)

SURCOUT FIOUL & GAZ



ANALYSE DU RENFORCEMENT (I/S2)

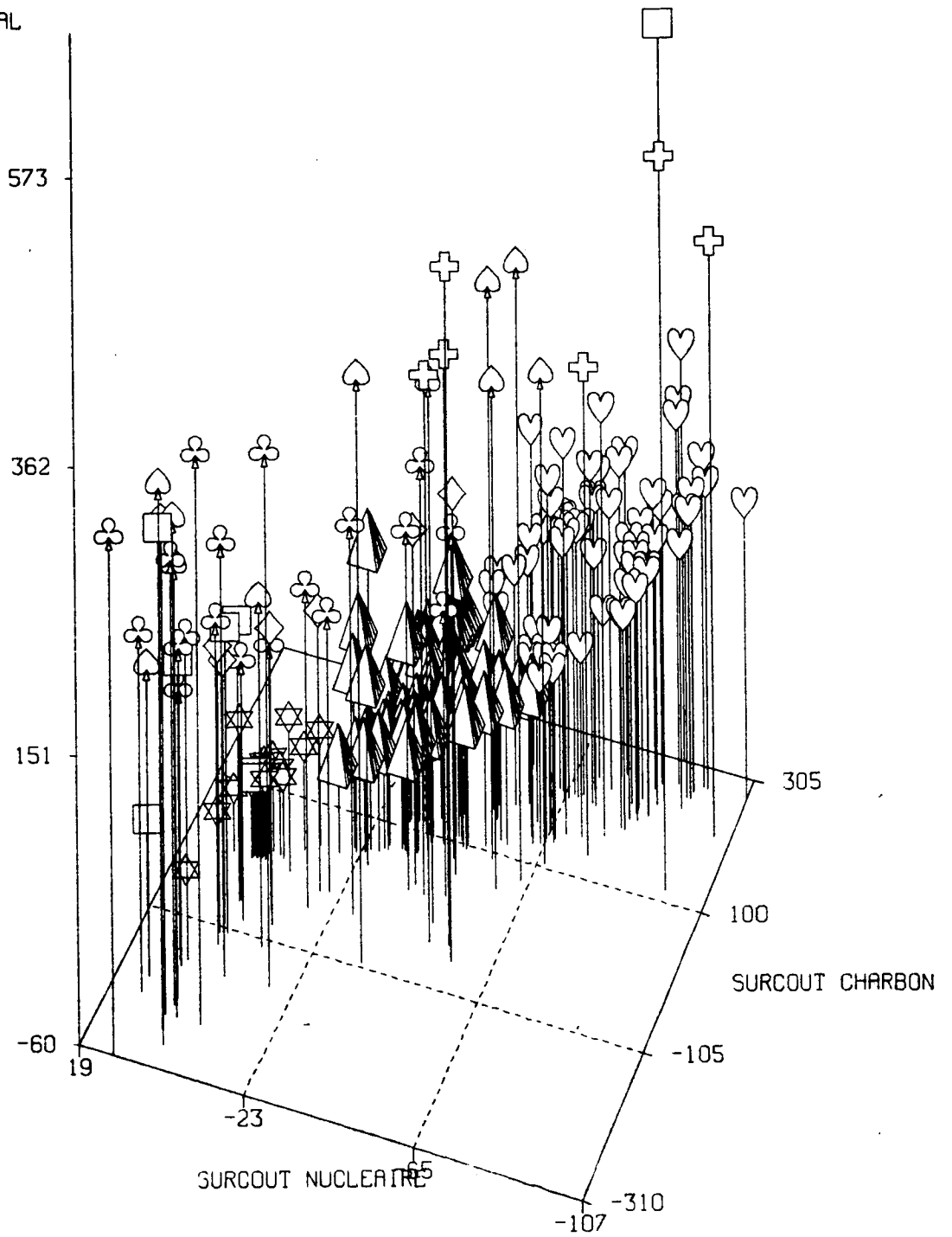
SURCOUT HYDRAULIQUE



(NUCLEAIRE CARBON HYDRAULIQUE FIOUL)

ANALYSE DU RENFORCEMENT (I/S2)

SURCOUT GLOBAL



(NUCLEAIRE CHARBON HYDRAULIQUE FIOUL)

D: Analyse avec filtrage des surcoûts du renforcement métrique W^{-1}

Comme précédemment, les variables retenues pour cette classification furent les surcoûts d'exploitation nucléaire (DCNU), charbon (DCCH), hydraulique (DCHY), fioul et gaz (DCFG) et de défaillance (DCDF).

La phase de filtrage avec PONDER s'est exécutée en quatre étapes, la dernière n'ayant pas donné lieu à la création de pondération. Nous avons fixé le seuil critique à cinq éléments contrairement à dix dans le cas précédent. La métrique utilisée pour le filtrage est celle de Mahalanobis : V^{-1} .

La première étape de filtrage permet, d'isoler une situation de défaillance avant renforcement. Il s'agit de la situation 1404 où le renforcement a réduit le coût d'exploitation du thermique de 964 kf. En conservant la même métrique, PONDER permet de filtrer deux situations (1366) où le renforcement a diminué le coût d'exploitation du thermique de 291 kf (meilleure valorisation du nucléaire et diminution du fioul et de l'hydraulique au delà de l'hypothèse de turbinage prévue à coût nul). Enfin cette première étape permet de filtrer encore deux situations (1046) où le renforcement a diminué le coût d'exploitation du thermique de 87kf (meilleure valorisation du charbon et de l'hydraulique et diminution du fioul). Cette phase de filtrage est exactement la même que celle réalisée avec la métrique $1/\sigma^2 I$.

La deuxième étape de filtrage permet, après recalcul de la métrique, d'isoler deux situations (101) où le renforcement a diminué le coût d'exploitation du thermique de 28 kf (meilleure valorisation du nucléaire et de l'hydraulique et diminution du charbon). En conservant la même métrique, PONDER permet de filtrer encore deux situations (452) où le renforcement a diminué le coût d'exploitation du thermique de 80 kf (meilleure valorisation du charbon et diminution de l'hydraulique et du fioul).

La troisième étape de filtrage permet, après recalcul de la métrique, d'isoler trois situations (1483) où le renforcement a diminué le coût d'exploitation du thermique de 15 kf (meilleure valorisation de l'hydraulique et diminution du fioul).

Finalement on constate que les points reconnus comme aberrants par la procédure PONDER le sont quelque soit la métrique utilisée ($1/\sigma^2 I$ ou V^{-1}). Cette fois cependant nous en avons filtré moins car le seuil critique avait été choisi plus faible.

La classification montre les résultats qui suivent.

La phase d'apprentissage s'est effectuée avec la métrique V^{-1} , en trois étapes. La partition finale est sauvegardée sur l'archive; elle extrait 90,1% d'inertie.

Le graphe de la statistique locale souligne très nettement une partition en 8 ou 16 classes. On constate en effet un maximum absolu pour ce critère, à 16 classes. Résultat confirmé par le graphe de l'inertie expliquée qui souligne des gains d'inertie expliquée très minimes au-delà. Le critère cubique de classification favorise plutôt une partition en 7 ou 12 classes. Finalement la procédure MNDOPT retient une partition en 8 classes.

Décrivons maintenant la partition obtenue après dix itérations utilisant la métrique W^{-1} pour la phase d'optimisation; les aides à l'interprétation incluses dans la procédure vont nous y aider. La convergence fut très délicate à obtenir car la matrice W n'est pas de plein rang (après filtrage de l'élément 1404, il n'y a plus de défaillance la variable DCDF prend une seule et même valeur : zéro).

Rappelons encore, que les aides à l'interprétation de la partition finale sont calculées avec la métrique identité parce qu'il n'est pas possible de calculer un écart-type

avec les métriques de Mahalanobis. Ces indices ne sont donc pas parfaitement adaptés à l'interprétation de partitions obtenues avec les distances adaptatives. Le pourcentage d'inertie expliquée de 86,4% est calculé avec la métrique W^{-1} ; il est excellent par comparaison à celui de la hiérarchie à huit classes (71,28%), ceci s'explique par le fait que nous avons filtré les points aberrants et que la métrique W^{-1} accentue les profils soulignés par la métrique V^{-1} .

L'examen des graphes des indices de dispersion de chaque variable active de cette classification sur les surcoûts d'exploitation du au renforcement, montre dans les classes les faits qui suivent :

La classe 1 comporte 897 aléas; où l'on trouve les situations où le surcoût d'exploitation global est en moyenne de 4.97 kf. Il s'agit comme dans les précédentes classifications des situations où le renforcement de la liaison Morbr172Vleva n'a presque rien changé.

La classe 2 (resp 3) comporte 209 aléas (resp 208 aléas); où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 12.2 kf (resp 5.98 kf). On trouve dans ces deux classes des situations où le renforcement de la liaison Morbr172Vleva n'a aussi presque rien changé.

La classe 4 (resp 5) comporte 76 aléas (resp 48 aléas); où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 73.9 kf (resp 172kf). On trouve dans ces deux classes des situations où le renforcement a permis une meilleure utilisation du nucléaire disponible en réduisant fortement celle du charbon.

La classe 6 (resp 7) comporte 27 aléas (resp 13 aléas) où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 194 kf (resp 307kf). On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du nucléaire et du charbon disponible en réduisant fortement celle du fioul.

La classe 8 comporte enfin 9 aléas où l'on trouve des situations où le surcoût d'exploitation global est en moyenne de 60.4 kf. On trouve dans cette classe des situations où le renforcement a permis une meilleure utilisation du charbon et de l'hydraulique (mais faiblement) disponible en réduisant celle du fioul.

CONCLUSION

Voici illustrées par quelques exemples les procédures de classification automatique PONDER et MNDOPT; les travaux que nous venons de présenter ont été élaborés pour des spécialistes de la planification des réseaux qui ne sont pas en revanche, des spécialistes de l'analyse des données. La méthodologie de classification élaborée à leur intention n'a pas la prétention de fournir la réponse définitive aux problèmes que pose la planification des réseaux; cependant au regard de l'expérience acquise, elle nous apparaît intéressante pour les utilisateurs désireux d'appliquer les techniques de classification automatique à leurs problèmes, sans posséder une grande compétence en Analyse de Données.

Les programmes PONDER et MNDOPT présentés sont disponibles dans le logiciel SICLA développé à l'INRIA de Rocquencourt.

Remerciements

L'auteur tient à remercier monsieur Y. LECHEVALLIER pour les nombreux conseils qu'il lui a donnés et pour l'aide qu'il lui a apportée tout au long de ces travaux.

BIBLIOGRAPHIE

E.M.L BEALE (1969) *Euclidean cluster analysis*.
Bull. Instit. Intern. Statist. 43(2)

M.BRUYNNOOGHE (1978)
Classification ascendante hiérarchique des grands ensembles de données. L'algorithme rapide fondé sur la construction de voisinages réductibles.
Les cahiers de l'analyse des Données 1. Dunod

T.CEMBRZYNSKI (1987)
Méthodologie d'analyse automatique des résultats du modèle de planification du réseau de transport à très haute tension d'EDF.
INRIA Rapports de recherche N°651

R.M. CORMACK (1971) *A review of classification*.
J. Royal. Statis. Soc. A N°134

E.DIDAY (1971) *La méthode des nuées dynamiques*.
Rev Stat App vol 19 n:2 (p19 à 34)

DIDAY et Coll (1979) *Optimisation en Classification Automatique (2 vol)*.
INRIA

H.P. FRIEDMAN et J.RUBIN (1967) *On some invariant criteria for grouping data*
JASA Vol 62

J.A HARTIGAN (1975) *Clustering algorithms*
John Wiley & Sons

Y.LECHEVALLIER, H.RALAMBONDRAIN, G.GOVAERT (1986)
Le système SICLA : Principes et Architecture
INRIA Rapport de Recherche N°500

I.C.LERMAN (1985) *Cours de DEA "traitement de l'information"*.
Université de Rennes I BEAULIEU

I.C.LERMAN (1982) *Classification et Analyse Ordinale des Données*.
Dunod

MODULAD (1986) *brochure d'utilisation*

W.S. SARLE (1983) *Cubic Clustering Criterion*.
SAS Institute Inc.
SAS Technical Report A-108

P.VILLOING (1983) *Méthodes de classification automatique non hiérarchiques*
CELAR / CCSA note d'étude L.A,N°56

ANNEXES

Application de PONDER et de MNDOPT à l'analyse de l'effet du renforcement de la ligne Morbras-Villevaudé du réseau de transport à très haute tension d'EDF.

1 à 8: Analyse sans filtrage des surcoûts du renforcement (métrique $1/\sigma^2$)

9 à 18: Analyse sans filtrage des surcoûts du renforcement (métrique W^{-1})

19 à 42: Analyse avec filtrage des surcoûts du renforcement (métrique $1/\sigma^2$)

43 à 56: Analyse avec filtrage des surcoûts du renforcement (métrique W^{-1})

```
*****  
* SICLA                Cembrzynski                Date : 25/05/87 *  
* S.D. : fusion.sdo          Archive : fusion.sar          *  
* Listing : liq09_45        Mode interactif          *  
*****  
* SICLA                Commande NUEDYN            Version 25/02/87 *  
*****  
* donnees : fusion.sdo      individus :1500 * variables : 165 *  
* _____ *  
*                               quantitatives : 93  qualitatives : 72 *  
*****
```

variables selectionnees :

dcnu dcch dchy dcfg dcd f dcex

Lecture sur l archive :fusion.sar

Constituant numero : 1

Nom du constituant
abrege :P3

Libelle :partition avec NUEDYN 982 %.

Type du constituant : partition
Date creation : 25/05/87
Nom de la S.D : fusion.sdo

obs	dcnu	dcch	dchy	dcfg	dcdf	dcex
1	-0.6178E-01	.3913	-0.6865E-02	0.1487E-01	0.0000E+00	.3124
2	-1.377	4.908	0.0000E+00	-0.7729E-01	0.0000E+00	3.527
3	-3.420	8.720	-0.4000E-01	-1.220	0.0000E+00	4.100
4	-21.03	54.69	0.0000E+00	-.4286	0.0000E+00	33.23
5	-27.16	82.72	0.0000E+00	-4.240	0.0000E+00	51.44
6	-3.320	14.72	0.4000E-01	.6400	0.0000E+00	12.24
7	-14.00	35.88	0.0000E+00	2.250	0.0000E+00	24.29
8	-7.870	23.30	0.0000E+00	1.826	0.0000E+00	17.35
9	-35.83	79.33	0.0000E+00	0.0000E+00	0.0000E+00	43.78
10	-66.67	213.3	0.0000E+00	10.39	0.0000E+00	157.2
11	-0.5882E-01	-11.82	.2353	30.24	0.0000E+00	18.18
12	-82.13	209.3	0.0000E+00	0.0000E+00	0.0000E+00	127.1
13	-31.07	111.2	0.0000E+00	4.571	0.0000E+00	84.79
14	7.462	-18.31	0.0000E+00	-4.231	0.0000E+00	-15.00
15	-89.50	260.4	0.0000E+00	-.9167	0.0000E+00	170.2
16	-40.70	124.9	0.0000E+00	8.400	0.0000E+00	92.80
17	-49.40	108.5	0.0000E+00	0.0000E+00	0.0000E+00	58.90
18	-64.50	142.2	0.0000E+00	0.0000E+00	0.0000E+00	77.70
19	-5.300	-38.00	.1000	118.7	0.0000E+00	75.70
20	-56.17	42.50	.8333	368.5	0.0000E+00	354.5
21	4.000	-174.5	.1667	367.0	0.0000E+00	196.7
22	-4.167	-85.33	2.167	234.5	0.0000E+00	146.8
23	-20.00	38.00	0.0000E+00	56.40	0.0000E+00	74.80
24	-55.20	149.4	0.0000E+00	-2.800	0.0000E+00	91.40
25	-5.400	-10.40	-3.000	62.20	0.0000E+00	43.60
26	1.800	-235.4	.4000	524.2	0.0000E+00	290.4
27	-43.20	39.40	.4000	201.6	0.0000E+00	198.4
28	-24.60	39.40	.4000	82.80	0.0000E+00	98.20
29	-54.50	179.5	0.0000E+00	0.0000E+00	0.0000E+00	125.3
30	-29.00	-23.67	0.0000E+00	279.0	0.0000E+00	226.7
31	-89.00	293.0	0.0000E+00	29.33	0.0000E+00	234.0
32	-74.33	152.3	0.0000E+00	193.0	0.0000E+00	271.3
33	.6667	-143.3	0.0000E+00	307.7	0.0000E+00	165.7
34	1.333	-275.7	.3333	618.0	0.0000E+00	343.7
35	-2.333	6.333	-8.000	18.67	0.0000E+00	15.00
36	-54.00	-51.50	0.0000E+00	523.5	0.0000E+00	418.0
37	-100.0	163.0	0.0000E+00	379.0	0.0000E+00	442.5
38	-26.00	-132.0	.5000	515.0	0.0000E+00	357.0
39	-15.00	-7.000	.5000	136.5	0.0000E+00	115.5
40	-1.000	-55.50	-63.00	206.5	0.0000E+00	87.00
41	-54.00	239.0	0.0000E+00	14.00	0.0000E+00	199.0
42	-13.00	6.000	0.0000E+00	0.0000E+00	0.0000E+00	-6.000
43	1.000	-14.00	12.00	11.00	0.0000E+00	9.000
44	-94.00	157.0	78.00	432.0	0.0000E+00	573.0
45	19.00	-112.0	101.0	0.0000E+00	0.0000E+00	8.000
46	0.0000E+00	6.000	0.0000E+00	-13.00	0.0000E+00	-7.000
47	-16.00	63.00	-17.00	0.0000E+00	0.0000E+00	29.00
48	15.00	-44.00	0.0000E+00	-31.00	0.0000E+00	-60.00
49	-35.00	-66.00	1.000	366.0	698.0	964.0
50	-31.00	114.0	-28.00	0.0000E+00	0.0000E+00	55.00

nci	% inertie	stat-loc	ccc
1	0.000	8.14469	0.000E+00
2	27.530	8.10572	-3.87
3	48.170	8.21619	4.41
4	65.860	8.52283	15.2
5	76.180	8.80956	19.2
6	81.370	9.08111	25.6
7	85.420	9.03520	42.7
8	88.770	9.08900	67.0
9	90.310	9.11536	82.7
10	91.730	8.95977	92.3
11	92.880	9.24811	110.
12	93.530	9.32593	129.
13	94.170	9.20554	151.
14	94.760	9.17413	174.
15	95.280	9.04973	200.
16	95.720	9.04574	229.
17	95.990	9.01481	262.
18	96.230	8.87073	298.
19	96.440	8.78109	334.
20	96.650	8.76523	379.
21	96.830	8.61917	429.
22	97.000	8.48760	486.
23	97.170	8.31299	550.
24	97.290	8.12993	623.
25	97.390	7.95618	707.
26	97.480	7.90067	814.
27	97.550	7.70785	926.
28	97.620	7.57556	0.106E+04
29	97.690	7.38539	0.121E+04
30	97.750	7.16250	0.144E+04
31	97.820	7.01677	0.166E+04
32	97.870	6.79528	0.191E+04
33	97.930	6.69115	0.160E+04
34	97.980	6.64597	0.186E+04
35	98.020	6.46968	0.216E+04
36	98.070	6.27533	0.256E+04
37	98.100	6.05949	0.305E+04
38	98.130	5.84674	0.367E+04
39	98.160	5.58197	0.447E+04
40	98.180	5.31438	0.549E+04
41	98.200	5.02266	0.682E+04
42	98.220	4.75964	0.861E+04
43	98.240	4.42403	0.110E+05
44	98.250	4.08359	0.143E+05
45	98.260	3.76469	0.194E+05
46	98.270	3.41315	0.260E+05
47	98.280	2.97059	0.350E+05
48	98.290	2.41248	0.470E+05
49	98.290	1.72215	0.609E+05
50	98.290	0.00000	0.732E+05

* graphe de l'inertie expliquée *



	0.00	98.29
1	i	
2	i.....	
3	i.....	
4	i.....	
5	i.....	
6	i.....	
7	i.....	
8	i.....	
9	i.....	
10	i.....	
11	i.....	
12	i.....	
13	i.....	
14	i.....	
15	i.....	
16	i.....	
17	i.....	
18	i.....	
19	i.....	
20	i.....	
21	i.....	
22	i.....	
23	i.....	
24	i.....	
25	i.....	
26	i.....	
27	i.....	
28	i.....	
29	i.....	
30	i.....	
31	i.....	
32	i.....	
33	i.....	
34	i.....	
35	i.....	
36	i.....	
37	i.....	
38	i.....	
39	i.....	
40	i.....	
41	i.....	
42	i.....	
43	i.....	
44	i.....	
45	i.....	
46	i.....	
47	i.....	
48	i.....	
49	i.....	
50	i.....	
	0.00	98.29

* statistique locale *



0.00

9.33

- 1 i.....
- 2 i.....
- 3 i.....
- 4 i.....
- 5 i.....
- 6 i.....
- 7 i.....
- 8 i.....
- 9 i.....
- 10 i.....
- 11 i.....
- 12 i.....
- 13 i.....
- 14 i.....
- 15 i.....
- 16 i.....
- 17 i.....
- 18 i.....
- 19 i.....
- 20 i.....
- 21 i.....
- 22 i.....
- 23 i.....
- 24 i.....
- 25 i.....
- 26 i.....
- 27 i.....
- 28 i.....
- 29 i.....
- 30 i.....
- 31 i.....
- 32 i.....
- 33 i.....
- 34 i.....
- 35 i.....
- 36 i.....
- 37 i.....
- 38 i.....
- 39 i.....
- 40 i.....
- 41 i.....
- 42 i.....
- 43 i.....
- 44 i.....
- 45 i.....
- 46 i.....
- 47 i.....
- 48 i.....
- 49 i.....
- 50 i.....

0.00

9.33

* critere cubique de classification *

	-3.87	706.69
1	i	
2	i	
3	i	
4	i.	
5	i..	
6	i..	
7	i....	
8	i.....	
9	i.....	
10	i.....	
11	i.....	
12	i.....	
13	i.....	
14	i.....	
15	i.....	
16	i.....	
17	i.....	
18	i.....	
19	i.....	
20	i.....	
21	i.....	
22	i.....	
23	i.....	
24	i.....	
25	i.....	
	-3.87	706.69

* noyaux finaux *

obs	dcnu	dcch	dchy	dcfg	dcdf	dcex
1	-1.603	3.800	-0.1997E-01	2.290	0.0000E+00	4.474
2	-52.20	147.4	-.3061	2.082	0.0000E+00	97.12
3	-26.28	-69.48	.5435	358.2	0.0000E+00	262.8
4	-1.000	-55.50	-63.00	206.5	0.0000E+00	87.00
5	-37.50	22.50	89.50	216.0	0.0000E+00	290.5
6	-35.00	-66.00	1.000	366.0	698.0	964.0
nb echanges :	24	i inter :	0.7422783E+04	i intra :	0.1577217E+04	
nb echanges :	6	i inter :	0.7431236E+04	i intra :	0.1568764E+04	
nb echanges :	2	i inter :	0.7431784E+04	i intra :	0.1568216E+04	
nb echanges :	4	i inter :	0.7432516E+04	i intra :	0.1567484E+04	
nb echanges :	0	i inter :	0.7432516E+04	i intra :	0.1567484E+04	

partition en 6 classes
 inertie expliquée : 7432.516
 inertie perdue : 1567.484
 % inertie extrait : 82.58350
 fisher (5;1494) : 1416.816
 nombre d iterations : 5
 nombre d echanges : 0

* impression de la partition *

impression de la classe : 1 effectif : 0.133E+04

impression de la classe : 2 effectif : 125.

9	16	17	23	71	72	84	101	104	115	131	134	136	154	166	173
181	189	196	205	207	215	227	263	275	285						
299	336	338	343	359	374	375	400	408	426	433	437	455	461	567	569
576	604	605	609	618	630	641	666	675	676						
692	693	694	695	698	717	725	727	732	738	748	749	771	773	777	786
820	822	852	858	871	874	881	886	902	912						
914	949	960	963	966	978	995	1003	1028	1063	1071	1082	1083	1104	1110	1132
1154	1160	1167	1173	1209	1227	1230	1233	1243	1265						
1272	1289	1306	1308	1328	1337	1349	1362	1390	1391	1401	1407	1412	1415	1421	1426
1428	1429	1456	1467	1479											

impression de la classe : 3 effectif : 42.0

146	161	197	213	234	295	300	431	464	465	485	499	526	542	550	651
686	711	722	787	916	970	971	1061	1100	1153						
1208	1263	1277	1279	1303	1329	1348	1359	1386	1397	1416	1431	1450	1471	1490	1493

impression de la classe : 4 effectif : 2.00

1046 1198

impression de la classe : 5 effectif : 2.00

480 1366

impression de la classe : 6 effectif : 1.00

1404

Ecriture sur l archive : fusion.sar

Constituant numero : 3

Nom du constituant
abrege : p 6

Libelle : partition avec NUEDYN 825 %.

Type du constituant : partition
Date creation : 25/05/87
Nom de la S.D : fusion.sdo

```

*****
* SICLA                      Cembrzynski                      Date : 25/05/87      *
* S.D. : fusion.sdo          Archive : fusion.sar              *
* Listing : lis10_03        Mode interactif                  *
*****
* SICLA                      Commande NUEDYN                  Version 25/02/87    *
*****
* donnees : fusion.sdo      individus :1500 * variables : 165      *
* _____                _____                        *
*                               quantitatives : 93  qualitatives : 72 *
*****

```

variables selectionnees :

dcnu dcch dchy dcfg dddf dcex

Lecture sur l archive :fusion.sar

Constituant numero : 1

Nom du constituant
abrege :P3

Libelle :partition avec NUEDYN 982 %.

Type du constituant : partition
Date creation : 25/05/87
Nom de la S.D : fusion.sdo

impression de l inverse de la metrique

dcnu	339.87					
dcch	-882.19	3167.9				
dchy	-3.0493	-0.12823E-01	17.595			
dcfg	-223.52	-1311.6	12.705	4831.0		
dcdf	-12.849	-37.938	.46285	163.80	324.59	
dcex	-782.24	938.42	27.687	3469.7	438.06	4091.0

* barycentres *

obs	dcnu	dcch	dchy	dcfg	dcdf	dcex
1	-0.6178E-01	.3913	-0.6865E-02	0.1487E-01	0.0000E+00	.3124
2	-1.377	4.908	0.0000E+00	-0.7729E-01	0.0000E+00	3.527
3	-3.420	8.720	-0.4000E-01	-1.220	0.0000E+00	4.100
4	-21.03	54.69	0.0000E+00	-.4286	0.0000E+00	33.23
5	-27.16	82.72	0.0000E+00	-4.240	0.0000E+00	51.44
6	-3.320	14.72	0.4000E-01	.6400	0.0000E+00	12.24
7	-14.00	35.88	0.0000E+00	2.250	0.0000E+00	24.29
8	-7.870	23.30	0.0000E+00	1.826	0.0000E+00	17.35
9	-35.83	79.33	0.0000E+00	0.0000E+00	0.0000E+00	43.78
10	-66.67	213.3	0.0000E+00	10.39	0.0000E+00	157.2
11	-0.5882E-01	-11.82	.2353	30.24	0.0000E+00	18.18
12	-82.13	209.3	0.0000E+00	0.0000E+00	0.0000E+00	127.1
13	-31.07	111.2	0.0000E+00	4.571	0.0000E+00	84.79
14	7.462	-18.31	0.0000E+00	-4.231	0.0000E+00	-15.00
15	-89.50	260.4	0.0000E+00	-.9167	0.0000E+00	170.2
16	-40.70	124.9	0.0000E+00	8.400	0.0000E+00	92.80
17	-49.40	108.5	0.0000E+00	0.0000E+00	0.0000E+00	58.90
18	-64.50	142.2	0.0000E+00	0.0000E+00	0.0000E+00	77.70
19	-5.300	-38.00	.1000	118.7	0.0000E+00	75.70
20	-56.17	42.50	.8333	368.5	0.0000E+00	354.5
21	4.000	-174.5	.1667	367.0	0.0000E+00	196.7
22	-4.167	-85.33	2.167	234.5	0.0000E+00	146.8
23	-20.00	38.00	0.0000E+00	56.40	0.0000E+00	74.80
24	-55.20	149.4	0.0000E+00	-2.800	0.0000E+00	91.40
25	-5.400	-10.40	-3.000	62.20	0.0000E+00	43.60
26	1.800	-235.4	.4000	524.2	0.0000E+00	290.4
27	-43.20	39.40	.4000	201.6	0.0000E+00	198.4
28	-24.60	39.40	.4000	82.80	0.0000E+00	98.20
29	-54.50	179.5	0.0000E+00	0.0000E+00	0.0000E+00	125.3
30	-29.00	-23.67	0.0000E+00	279.0	0.0000E+00	226.7
31	-89.00	293.0	0.0000E+00	29.33	0.0000E+00	234.0
32	-74.33	152.3	0.0000E+00	193.0	0.0000E+00	271.3
33	.6667	-143.3	0.0000E+00	307.7	0.0000E+00	165.7
34	1.333	-275.7	.3333	618.0	0.0000E+00	343.7
35	-2.333	6.333	-8.000	18.67	0.0000E+00	15.00
36	-54.00	-51.50	0.0000E+00	523.5	0.0000E+00	418.0
37	-100.0	163.0	0.0000E+00	379.0	0.0000E+00	442.5
38	-26.00	-132.0	.5000	515.0	0.0000E+00	357.0
39	-15.00	-7.000	.5000	136.5	0.0000E+00	115.5
40	-1.000	-55.50	-63.00	206.5	0.0000E+00	87.00
41	-54.00	239.0	0.0000E+00	14.00	0.0000E+00	199.0
42	-13.00	6.000	0.0000E+00	0.0000E+00	0.0000E+00	-6.000
43	1.000	-14.00	1.000	11.00	0.0000E+00	9.000
44	-94.00	157.0	78.00	432.0	0.0000E+00	573.0
45	19.00	-112.0	101.0	0.0000E+00	0.0000E+00	8.000
46	0.0000E+00	6.000	0.0000E+00	-13.00	0.0000E+00	-7.000
47	-16.00	63.00	-17.00	0.0000E+00	0.0000E+00	29.00
48	15.00	-44.00	0.0000E+00	-31.00	0.0000E+00	-60.00
49	-35.00	-66.00	1.000	366.0	698.0	964.0
50	-31.00	114.0	-28.00	0.0000E+00	0.0000E+00	55.00

nci	% inertie	stat-loc	ccc
1	0.000	7.91348	0.000E+00
2	16.660	8.20603	-.239
3	30.220	8.25403	3.18
4	41.520	8.17490	14.1
5	51.810	8.46008	23.4
6	58.920	8.49781	26.3
7	63.900	8.76224	38.1
8	65.970	8.80139	49.1
9	67.520	8.62950	82.5
10	68.950	8.62396	92.5
11	70.160	8.90862	108.
12	71.130	8.84100	128.
13	71.830	8.94883	149.
14	72.510	8.82936	173.
15	73.130	8.68324	198.
16	73.670	8.66619	227.
17	74.170	8.58681	259.
18	74.530	8.41502	295.
19	74.870	8.37227	334.
20	75.190	8.32945	385.
21	75.510	8.18994	436.
22	75.820	8.04663	494.
23	75.990	7.92462	560.
24	76.140	7.96575	637.
25	76.290	7.82138	723.
26	76.420	7.65592	821.
27	76.560	7.45497	920.
28	76.690	7.39677	0.105E+04
29	76.810	7.17446	0.120E+04
30	76.930	6.99298	0.137E+04
31	77.020	6.92789	0.158E+04
32	77.120	6.71944	0.182E+04
33	77.200	6.51372	0.212E+04
34	77.270	6.34366	0.247E+04
35	77.330	6.09883	0.220E+04
36	77.390	6.05119	0.253E+04
37	77.440	5.85456	0.301E+04
38	77.480	5.59254	0.363E+04
39	77.530	5.37906	0.441E+04
40	77.570	5.14654	0.544E+04
41	77.610	4.95268	0.677E+04
42	77.640	4.63130	0.861E+04
43	77.680	4.29545	0.110E+05
44	77.710	4.05005	0.143E+05
45	77.740	3.71656	0.190E+05
46	77.760	3.29140	0.260E+05
47	77.790	2.93791	0.350E+05
48	77.810	2.41849	0.470E+05
49	77.810	1.72781	0.609E+05
50	77.810	0.00000	0.732E+05

* graphe de l'inertie expliquée *

	0.00	77.81
1 i		
2 i.....		
3 i.....		
4 i.....		
5 i.....		
6 i.....		
7 i.....		
8 i.....		
9 i.....		
10 i.....		
11 i.....		
12 i.....		
13 i.....		
14 i.....		
15 i.....		
16 i.....		
17 i.....		
18 i.....		
19 i.....		
20 i.....		
21 i.....		
22 i.....		
23 i.....		
24 i.....		
25 i.....		
26 i.....		
27 i.....		
28 i.....		
29 i.....		
30 i.....		
31 i.....		
32 i.....		
33 i.....		
34 i.....		
35 i.....		
36 i.....		
37 i.....		
38 i.....		
39 i.....		
40 i.....		
41 i.....		
42 i.....		
43 i.....		
44 i.....		
45 i.....		
46 i.....		
47 i.....		
48 i.....		
49 i.....		
50 i.....		
	0.00	77.81

* statistique locale *

-----*

	0.00	8.95
1 i	
2 i	
3 i	
4 i	
5 i	
6 i	
7 i	
8 i	
9 i	
10 i	
11 i	
12 i	
13 i	
14 i	
15 i	
16 i	
17 i	
18 i	
19 i	
20 i	
21 i	
22 i	
23 i	
24 i	
25 i	
26 i	
27 i	
28 i	
29 i	
30 i	
31 i	
32 i	
33 i	
34 i	
35 i	
36 i	
37 i	
38 i	
39 i	
40 i	
41 i	
42 i	
43 i	
44 i	
45 i	
46 i	
47 i	
48 i	
49 i	
50 i		
	0.00	8.95

critere cubique de classification *

	-0.24	722.96
1 i		
2 i		
3 i		
4 i.		
5 i..		
6 i..		
7 i...		
8 i....		
9 i.....		
10 i.....		
11 i.....		
12 i.....		
13 i.....		
14 i.....		
15 i.....		
16 i.....		
17 i.....		
18 i.....		
19 i.....		
20 i.....		
21 i.....		
22 i.....		
23 i.....		
24 i.....		
25 i.....		
	-0.24	722.96

impression de l inverse de la metrique

dcnu 89.108
 dcch -248.66 891.26
 dchy 1.6136 -5.6362 1.6100
 dcfg -14.762 -306.62 -1.7898 865.44
 dcd f 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -173.08 331.36 -4.2133 541.43 0.00000E+00 695.70

* noyaux finaux *

obs	dcnu	dcch	dchy	dcfg	dcd f	dcex
1	-2.397	6.305	-0.5306E-01	2.212	0.0000E+00	6.075
2	-56.02	132.7	0.0000E+00	-.2373	0.0000E+00	76.46
3	-62.46	199.2	0.0000E+00	6.577	0.0000E+00	143.5
4	-69.09	94.36	.4545	322.5	0.0000E+00	347.8
5	-12.83	-121.0	.5714	369.4	0.0000E+00	236.1
6	-1.000	-55.50	-63.00	206.5	0.0000E+00	87.00
7	-37.50	22.50	89.50	216.0	0.0000E+00	290.5
8	-35.00	-66.00	1.000	366.0	698.0	964.0

totale : 0.1279022E+07 i inter : 0.1272029E+07 i intra : 0.6993331E+04

impression de l inverse de la metrique

dcnu 58.515
 dcch -163.51 673.05
 dchy 1.6221 -5.7491 1.6103
 dcfg -11.288 -370.09 -1.6077 970.31
 dcd f 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -115.29 135.49 -4.1352 586.42 0.00000E+00 603.10

totale : 0.1300822E+07 i inter : 0.1293016E+07 i intra : 0.7806142E+04

impression de l inverse de la metrique

dcnu 53.545
 dcch -150.11 645.56
 dchy 1.6381 -5.8230 1.6104
 dcfg -0.63715E-02 -417.88 -1.5655 991.21
 dcd f 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -95.614 73.675 -4.1507 570.92 0.00000E+00 545.61

partition en 8 classes

inertie expliquée : 1293016.
 inertie perdue : 7806.142
 % inertie extrait : 99.39000
 fisher (7;1492) : 35305.15
 nombre d iterations : 2
 nombre d echanges : 0

description de la classe : 1

effectif : 0.132E+04

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-2.00	5.68	i-
dcch	15.5	56.3	5.30	16.7	i
dchy	0.533E-02	4.19	-0.545E-01	1.12	i
dcfg	14.0	69.5	1.89	13.7	i
dcd	.465	18.0	0.000E+00	0.000E+00	i
dcex	22.6	64.0	5.14	14.5	-i

description de la classe : 2

effectif : 67.0

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-55.2	21.5	—*—*—i
dcch	15.5	56.3	129.	60.7	i—*—*
dchy	0.533E-02	4.19	0.000E+00	0.000E+00	i
dcfg	14.0	69.5	2.49	22.0	i
dcd	.465	18.0	0.000E+00	0.000E+00	i
dcex	22.6	64.0	76.0	39.6	i—

description de la classe : 3

effectif : 61.0

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-58.0	20.1	—*—*—i
dcch	15.5	56.3	185.	60.1	i—*—*—*
dchy	0.533E-02	4.19	0.000E+00	0.000E+00	i
dcfg	14.0	69.5	12.0	33.3	i
dcd	.465	18.0	0.000E+00	0.000E+00	i
dcex	22.6	64.0	139.	45.4	i—*—

description de la classe : 4

effectif : 32.0

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-8.13	16.4	i
dcch	15.5	56.3	-139.	84.9	—*—*—i
dchy	0.533E-02	4.19	.594	1.78	i
dcfg	14.0	69.5	379.	138.	i—*—*—*—*—>
dcd	.465	18.0	0.000E+00	0.000E+00	i
dcex	22.6	64.0	232.	82.4	i—*—*—*

description de la classe : 5

effectif : 13.0

variable moypop stdpop moycla stdcla

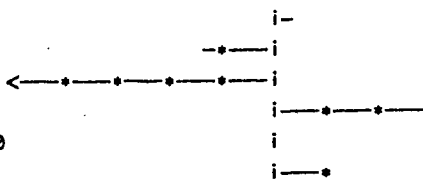
dcnu	-7.39	18.4	-63.5	19.3	*—*—*—i
dcch	15.5	56.3	74.2	72.1	i—*
dchy	0.533E-02	4.19	.538	.499	i
dcfg	14.0	69.5	326.	108.	i—*—*—*—*
dcd	.465	18.0	0.000E+00	0.000E+00	i
dcex	22.6	64.0	337.	82.4	i—*—*—*—*

description de la classe : 6

effectif : 2.00

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-1.00	1.00
dcch	15.5	56.3	-55.5	55.5
dchy	0.533E-02	4.19	-63.0	14.0
dcfg	14.0	69.5	207.	125.
dcdf	.465	18.0	0.000E+00	0.000E+00
dcex	22.6	64.0	87.0	82.0

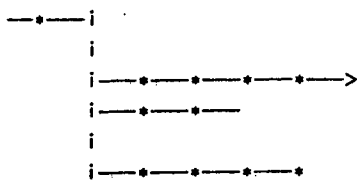


description de la classe : 7

effectif : 2.00

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-37.5	56.5
dcch	15.5	56.3	22.5	135.
dchy	0.533E-02	4.19	89.5	11.5
dcfg	14.0	69.5	216.	216.
dcdf	.465	18.0	0.000E+00	0.000E+00
dcex	22.6	64.0	291.	283.

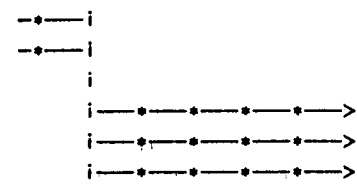


description de la classe : 8

effectif : 1.00

variable moypop stdpop moycla stdcla

dcnu	-7.39	18.4	-35.0	0.000E+00
dcch	15.5	56.3	-66.0	0.000E+00
dchy	0.533E-02	4.19	1.00	0.000E+00
dcfg	14.0	69.5	366.	0.000E+00
dcdf	.465	18.0	698.	0.000E+00
dcex	22.6	64.0	964.	0.000E+00



* impression de la partition *

impression de la classe : 1 effectif : 0.132E+04

impression de la classe : 2 effectif : 67.0

9	16	23	71	72	104	131	148	164	173	181	189	215	226	296	299
338	374	397	400	433	437	455	461	569	576						
618	649	655	666	692	727	738	748	758	786	820	822	849	850	852	886
902	914	966	995	999	1021	1071	1082	1083	1154						
1160	1186	1240	1247	1265	1289	1347	1362	1390	1391	1412	1415	1421	1429	1467	

impression de la classe : 3 effectif : 61.0

17	115	134	136	154	166	196	205	207	227	263	275	285	336	375	408
426	604	605	609	630	675	676	694	695	698						
717	725	732	749	771	773	777	858	871	874	904	912	949	960	1003	1028
1063	1110	1132	1167	1173	1209	1227	1230	1233	1243						
1308	1328	1337	1349	1401	1407	1426	1456	1479							

impression de la classe : 4 effectif : 32.0

146	161	197	213	295	300	431	464	485	499	526	542	550	711	722	916
971	1100	1153	1208	1263	1277	1279	1303	1359	1386						
1397	1416	1431	1450	1471	1490										

impression de la classe : 5 effectif : 13.0

234 465 651 686 693 787 963 970 978 1061 1329 1348 1493

impression de la classe : 6 effectif : 2.00

1046 1198

impression de la classe : 7 effectif : 2.00

480 1366

impression de la classe : 8 effectif : 1.00

1404

Ecriture sur l'archive : fusion.sar

Constituant numero : 4

Nom du constituant

abrege : p 8

Libelle : partition avec NUEDYN 993 %.

Type du constituant : partition
 Date creation : 25/05/87
 Nom de la S.D : fusion.sar

```
*****
* SICLA                      Cembrzynski          Date : 20/05/87  *
* S.D. :                      Archive :           *
* Listing : lis14_18          Mode interactif     *
* version 2.2 : (faire help sicla)                *
* _____ *
*****
```

```
*****
* SICLA                      Commande PONDER      Version 10/04/87 *
*****
* donnees : renfo.sdo        individus :1500 * variables : 6    *
* _____ *
*                               quantitatives : 6  qualitatives : 0 *
*****
```

```
-----*
* filtrage d elements douteux *
*-----*
```

le point le plus eloigne : 1404

nb de classes :	2	nb transferts :	0 % inertie	:	19.41
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 1366

nb de classes :	3	nb transferts :	1 % inertie	:	30.18
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 1046

nb de classes :	4	nb transferts :	1 % inertie	:	35.41
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 711

nb de classes :	5	nb transferts :	36 % inertie	:	57.02
-----------------	---	-----------------	--------------	---	-------

```
-----*
* filtrage d elements douteux *
*-----*
```

le point le plus eloigne : 101

nb de classes :	2	nb transferts :	2 % inertie	:	13.54
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 711

nb de classes :	3	nb transferts :	41 % inertie	:	43.09
-----------------	---	-----------------	--------------	---	-------

```
-----*
* filtrage d elements douteux *
*-----*
```

le point le plus eloigne : 452

nb de classes :	2	nb transferts :	1 % inertie	:	10.24
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 1483

nb de classes :	3	nb transferts :	2 % inertie	:	18.28
-----------------	---	-----------------	-------------	---	-------

creation d un noyau avec : 711

nb de classes :	4	nb transferts :	40 % inertie	:	47.85
-----------------	---	-----------------	--------------	---	-------

•-----•
• filtrage d elements douteux •
•-----•

le point le plus eloigne : 711

nb de classes : 2 nb transferts : 38 % inertie : 32.38

creation d un noyau avec : 580

nb de classes : 3 nb transferts : 14 % inertie : 39.04

•-----•
• filtrage d elements douteux •
•-----•

le point le plus eloigne : 711

nb de classes : 2 nb transferts : 41 % inertie : 35.96

creation d un noyau avec : 693

nb de classes : 3 nb transferts : 110 % inertie : 68.39

liste des points aberrants

97 101 235 452 468 480 526 570 580 597 602 744 920
1002 1046 1198 1225 1263 1277 1366 1404 1483

```
*****
* SICLA                      Cembrzynski                      Date : 20/05/87 *
* S.D. :                      Archive :                      *
* Listing : lis14_18          Mode interactif              *
* version 2.2 : (faire help sicla)                          *
* _____ *
*****
* SICLA                      Commande PONDER                  Version 10/04/87 *
*****
* donnees : renfo.sdo        individus :1500 * variables : 6 *
* _____ *
*                               quantitatives : 6 qualitatives : 0 *
*****
```

description des variables

obs	moyenne	e. type	maximum	minimum
dcnu	-7.387	18.44	20.00	-107.0
dcch	15.53	56.28	305.0	-310.0
dchy	0.5333E-02	4.195	101.0	-77.00
dcfg	13.98	69.51	630.0	-78.00
dcd	.4653	18.02	698.0	0.0000E+00
dcex	22.61	63.96	964.0	-60.00

partition sur : 1500 individus
 en : 2 classes
 inertie totale: 9000.000
 inertie inter : 1746.838
 inertie intra : 7253.162
 pourcentage d inertie expliquee : 19.400

description de la classe : 1
 effectif : 0.150E+04
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-7.37	i
dcch	15.5	56.3	15.6	i
dchy	0.533E-02	4.19	0.467E-02	i
dcfg	14.0	69.5	13.7	i
dcd	.465	18.0	0.000E+00	i
dcex	22.6	64.0	22.0	i

description de la classe : 2
 effectif : 1.00
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-35.0	-----i
dcch	15.5	56.3	-66.0	-----i
dchy	0.533E-02	4.19	1.00	i
dcfg	14.0	69.5	366.	i----->
dcd	.465	18.0	698.	i----->
dcex	22.6	64.0	964.	i----->

partition sur : 1500 individus
 en : 3 classes
 inertie totale: 9000.000
 inertie inter : 2716.127
 inertie intra : 6283.873
 pourcentage d inertie expliquée : 30.170

description de la classe : 1
 effectif : 0.150E+04

variable moypop stdpop moycla

dcnu	-7.39	18.4	-7.33	i
dcch	15.5	56.3	15.6	i
dchy	0.533E-02	4.19	-.115	i
dcfg	14.0	69.5	13.5	i
dcd	.465	18.0	0.000E+00	i
dcex	22.6	64.0	21.6	i

description de la classe : 2
 effectif : 2,00

variable moypop stdpop moycla

dcnu	-7.39	18.4	-37.5	—•—i
dcch	15.5	56.3	22.5	i
dchy	0.533E-02	4.19	89.5	i—•—•—•—•—•—>
dcfg	14.0	69.5	216.	i—•—•—
dcd	.465	18.0	0.000E+00	i
dcex	22.6	64.0	291.	i—•—•—•—•—

description de la classe : 3
 effectif : 1,00

variable moypop stdpop moycla

dcnu	-7.39	18.4	-35.0	—•—i
dcch	15.5	56.3	-66.0	—•—i
dchy	0.533E-02	4.19	1.00	i
dcfg	14.0	69.5	366.	i—•—•—•—•—•—>
dcd	.465	18.0	698.	i—•—•—•—•—•—>
dcex	22.6	64.0	964.	i—•—•—•—•—•—>

partition sur : 1500 individus
 en : 4 classes
 inertie totale: 9000.000
 inertie inter : 3187.211
 inertie intra : 5812.789
 pourcentage d inertie expliquée : 35.410

description de la classe : 1
 effectif : 0.150E+04
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-7.34	i
dcch	15.5	56.3	15.7	i
dchy	0.533E-02	4.19	-0.308E-01	i
dcfg	14.0	69.5	13.2	i
dcdf	.465	18.0	0.000E+00	i
dcex	22.6	64.0	21.5	i

description de la classe : 2
 effectif : 2.00
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-37.5	i
dcch	15.5	56.3	22.5	i
dchy	0.533E-02	4.19	89.5	i
dcfg	14.0	69.5	216.	i
dcdf	.465	18.0	0.000E+00	i
dcex	22.6	64.0	291.	i

description de la classe : 3
 effectif : 2.00
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-1.00	i
dcch	15.5	56.3	-55.5	i
dchy	0.533E-02	4.19	-63.0	i
dcfg	14.0	69.5	207.	i
dcdf	.465	18.0	0.000E+00	i
dcex	22.6	64.0	87.0	i

description de la classe : 4
 effectif : 1.00
 variable moypop stdpop moycla

dcnu	-7.39	18.4	-35.0	i
dcch	15.5	56.3	-66.0	i
dchy	0.533E-02	4.19	1.00	i
dcfg	14.0	69.5	366.	i
dcdf	.465	18.0	698.	i
dcex	22.6	64.0	964.	i

partition sur : 1500 individus
 en : 5 classes
 inertie totale: 9000.000
 inertie inter : 5131.448
 inertie intra : 3868.552
 pourcentage d inertie expliquee : 57.010

description de la classe : 1
 effectif : 0.146E+04

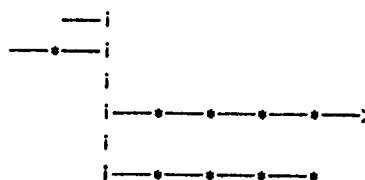
variable moypop stdpop moycla

dcnu	-7.39	18.4	-6.95
dcch	15.5	56.3	18.5
dchy	0.533E-02	4.19	-0.466E-01
dcfg	14.0	69.5	3.49
dcd	.465	18.0	0.000E+00
dcex	22.6	64.0	15.0

description de la classe : 2
 effectif : 37.0

variable moypop stdpop moycla

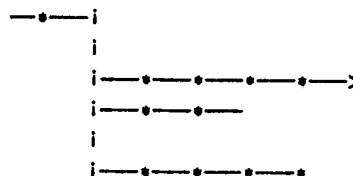
dcnu	-7.39	18.4	-22.8
dcch	15.5	56.3	-95.4
dchy	0.533E-02	4.19	.595
dcfg	14.0	69.5	397.
dcd	.465	18.0	0.000E+00
dcex	22.6	64.0	279.



description de la classe : 3
 effectif : 2.00

variable moypop stdpop moycla

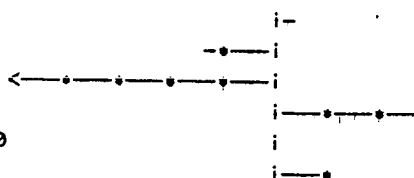
dcnu	-7.39	18.4	-37.5
dcch	15.5	56.3	22.5
dchy	0.533E-02	4.19	89.5
dcfg	14.0	69.5	216.
dcd	.465	18.0	0.000E+00
dcex	22.6	64.0	291.



description de la classe : 4
 effectif : 2.00

variable moypop stdpop moycla

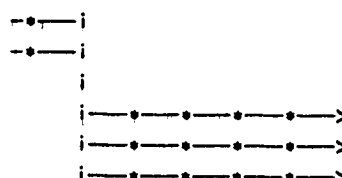
dcnu	-7.39	18.4	-1.00
dcch	15.5	56.3	-55.5
dchy	0.533E-02	4.19	-63.0
dcfg	14.0	69.5	207.
dcd	.465	18.0	0.000E+00
dcex	22.6	64.0	87.0



description de la classe : 5
 effectif : 1.00

variable moypop stdpop moycla

dcnu	-7.39	18.4	-35.0
dcch	15.5	56.3	-66.0
dchy	0.533E-02	4.19	1.00
dcfg	14.0	69.5	366.
dcd	.465	18.0	698.
dcex	22.6	64.0	964.



description des variables

obs	moyenne	e. type	maximum	minimum
dcnu	-7.337	18.30	20.00	-107.0
dcch	15.67	56.03	305.0	-310.0
dchy	-0.3077E-01	1.090	101.0	-77.00
dcfg	13.22	67.65	630.0	-78.00
dcdf	0.0000E+00	0.0000E+00	698.0	0.0000E+00
dcex	21.53	57.39	964.0	-60.00

partition sur : 1500 individus
 en : 2 classes
 inertie totale: 7475.000
 inertie inter : 1012.332
 inertie intra : 6462.668
 pourcentage d inertie expliquée : 13.540

description de la classe : 1
 effectif : 0.149E+04

variable moypop stdpop moycla

dcnu	-7.34	18.3	-7.32	i
dcch	15.7	56.0	15.6	i
dchy	-0.308E-01	1.09	0.938E-02	i
dcfg	13.2	67.6	13.2	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	21.5	i

description de la classe : 2
 effectif : 3.00

variable moypop stdpop moycla

dcnu	-7.34	18.3	-15.7	-i
dcch	15.7	56.0	58.7	i
dchy	-0.308E-01	1.09	-20.0	i
dcfg	13.2	67.6	5.67	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	28.3	i

partition sur : 1500 individus
 en : 3 classes
 inertie totale: 7460.000
 inertie inter : 1363.865
 inertie intra : 6096.135
 pourcentage d inertie expliquee : 18.280

description de la classe : 1
 effectif : 0.149E+04

variable moypop stdpop moycla

dcnu	-7.32	18.3	-7.34	i
dcch	15.6	56.0	15.7	i
dchy	0.938E-02	.566	0.108E-01	i
dcfg	13.2	67.7	13.1	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	21.5	i

description de la classe : 2
 effectif : 3.00

variable moypop stdpop moycla

dcnu	-7.32	18.3	-2.33	i-
dcch	15.6	56.0	6.33	i
dchy	0.938E-02	.566	-8.00	i ←
dcfg	13.2	67.7	18.7	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	15.0	i

description de la classe : 3

effectif : 2.00

variable moypop stdpop moycla

dcnu	-7.32	18.3	0.000E+00	i-
dcch	15.6	56.0	-39.5	i
dchy	0.938E-02	.566	11.0	i →
dcfg	13.2	67.7	109.	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	80.0	i

partition sur : 1500 individus
 en : 4 classes
 inertie totale: 7460.000
 inertie inter : 3569.777
 inertie intra : 3890.223
 pourcentage d inertie expliquee : 47.850

description de la classe : 1
 effectif : 0.145E+04

variable moypop stdpop moycla

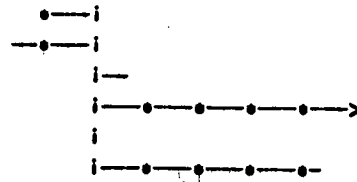
dcnu	-7.32	18.3	-6.79
dcch	15.6	56.0	18.3
dchy	0.938E-02	.566	0.692E-03
dcfg	13.2	67.7	2.80
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	14.3

|
|
|
|
|
|

description de la classe : 2
 effectif : 41.0

variable moypop stdpop moycla

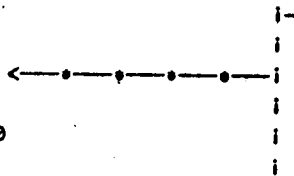
dcnu	-7.32	18.3	-26.6
dcch	15.6	56.0	-75.1
dchy	0.938E-02	.566	.366
dcfg	13.2	67.7	376.
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	275.



description de la classe : 3
 effectif : 3.00

variable moypop stdpop moycla

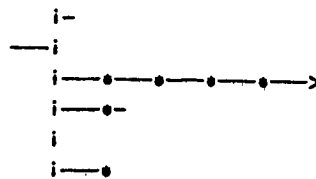
dcnu	-7.32	18.3	-2.33
dcch	15.6	56.0	6.33
dchy	0.938E-02	.566	-8.00
dcfg	13.2	67.7	18.7
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	15.0



description de la classe : 4
 effectif : 2.00

variable moypop stdpop moycla

dcnu	-7.32	18.3	0.000E+00
dcch	15.6	56.0	-39.5
dchy	0.938E-02	.566	11.0
dcfg	13.2	67.7	109.
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	80.0



description des variables

obs	moyenne	e. type	maximum	minimum
dcnu	-7.340	18.33	20.00	-107.0
dcch	15.68	56.05	305.0	-310.0
dchy	0.1076E-01	.1677	101.0	-77.00
dcfg	13.10	67.64	630.0	-78.00
dcdf	0.0000E+00	0.0000E+00	698.0	0.0000E+00
dcex	21.45	57.43	964.0	-60.00

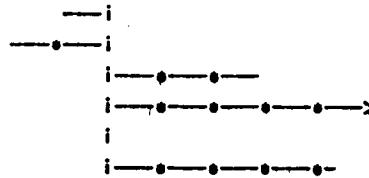
partition sur : 1500 individus
 en : 2 classes
 inertie totale: 7435.000
 inertie inter : 2407.357
 inertie intra : 5027.643
 pourcentage d inertie expliquée : 32.370

description de la classe : 1
 effectif : 0.145E+04

variable	moypop	stdpop	moycla
dcnu	-7.34	18.3	-6.88
dcch	15.7	56.0	18.3
dchy	0.108E-01	.168	-0.207E-02
dcfg	13.1	67.6	3.26
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	14.7

description de la classe : 2
 effectif : 39.0

variable	moypop	stdpop	moycla
dcnu	-7.34	18.3	-24.5
dcch	15.7	56.0	-82.7
dchy	0.108E-01	.168	.487
dcfg	13.1	67.6	378.
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	271.



description des variables

obs	moyenne	e. type	maximum	minimum
dcnu	-7.385	18.38	20.00	-107.0
dcch	16.07	55.55	305.0	-310.0
dchy	0.1827E-01	.1339	101.0	-77.00
dcfg	12.50	65.77	630.0	-78.00
dcdf	0.0000E+00	0.0000E+00	698.0	0.0000E+00
dcex	21.22	56.98	964.0	-60.00

partition sur : 1500 individus
 en : 2 classes
 inertie totale: 7390.000
 inertie inter : 2657.290
 inertie intra : 4732.710
 pourcentage d inertie expliquee : 35.950

description de la classe : 1
 effectif : 0.144E+04

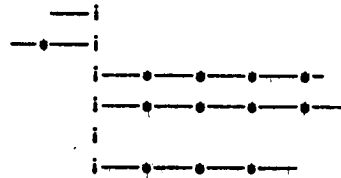
variable moypop stdpop moycla

dcnu	-7.38	18.4	-6.96	i
dcch	16.1	55.5	18.6	i
dchy	0.183E-01	.134	0.696E-03	i
dcfg	12.5	65.8	3.16	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	14.8	i

description de la classe : 2
 effectif : 42.0

variable moypop stdpop moycla

dcnu	-7.38	18.4	-21.8	i
dcch	16.1	55.5	-70.8	i
dchy	0.183E-01	.134	.619	i
dcfg	12.5	65.8	332.	i
dcdf	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	240.	i



partition sur : 1500 individus
 en : 3 classes
 inertie totale: 7390.000
 inertie inter : 5053.787
 inertie intra : 2336.213
 pourcentage d inertie expliquée : 68.380

description de la classe : 1
 effectif : 0.133E+04

variable	moypop	stdpop	moycla
dcnu	-7.38	18.4	-2.59
dcch	16.1	55.5	6.60
dchy	0.183E-01	.134	-0.293E-19
dcfg	12.5	65.8	1.75
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.2	57.0	5.78

description de la classe : 2
 effectif : 107.

variable moypop stdpop moycla

variable	moypop	stdpop	moycla
dcnu	-7.38	18.4	-62.3
dcch	16.1	55.5	172.
dchy	0.183E-01	.134	0.000E+00
dcfg	12.5	65.8	17.3
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.2	57.0	127.

description de la classe : 3
 effectif : 44.0

variable moypop stdpop moycla

variable	moypop	stdpop	moycla
dcnu	-7.38	18.4	-18.5
dcch	16.1	55.5	-78.5
dchy	0.183E-01	.134	.614
dcfg	12.5	65.8	325.
dcdf	0.000E+00	0.000E+00	0.000E+00
dcex	21.2	57.0	228.

liste des points aberrants

97 101 235 452 468 480 526 570 580 597 602 744 920
 1002 1046 1198 1225 1263 1277 1366 1404 1483

creation de la variable :POID Ponderation avec PONDER (metrique 1/s2)
 — SICLA ————— FIN COMMANDE : PONDER

```

*****
* SICLA                      Cembrzynski          Date : 20/05/87  *
* S.D. :                      Archive :          *
* Listing : lis14_53          Mode interactif   *
* version 2.2 : (faire help sicla)              *
* _____ *
*****
* SICLA                      Commande NUEDYN      Version 25/02/87 *
*****
* donnees : renfo.sdo        individus :1500 * variables : 7 *
* _____ *
*                               quantitatives : 7 qualitatives : 0 *
*****

```

```

totale : 0.7390000E+04 i inter : 0.7197136E+04 i intra : 0.1928641E+03
iteration : 1 i inter : 0.7202434E+04 i intra : 0.1875663E+03
iteration : 2 i inter : 0.7203982E+04 i intra : 0.1860179E+03

```

Ecriture sur l archive :renfo.sar

Constituant numero : 1

Nom du constituant
abrege :p 50

Libelle :partition avec NUEDYN 974 %.

```

Type du constituant : partition
Date creation       : 20/05/87
Nom de la S.D      : renfo.sdo

```

* barycentres *

obs	dcnu	dcch	dchy	dcfg	dcdf	dcex
1	-0.9091E-02	.3429	0.0000E+00	0.2987E-01	0.0000E+00	.3000
2	-1.106	3.282	0.0000E+00	0.0000E+00	0.0000E+00	2.326
3	-2.316	7.772	0.0000E+00	0.5263E-01	0.0000E+00	5.623
4	-5.000	17.65	0.0000E+00	0.0000E+00	0.0000E+00	12.62
5	1.560	-5.680	0.0000E+00	.9200	0.0000E+00	-3.360
6	-28.35	92.53	0.0000E+00	-1.647	0.0000E+00	62.71
7	-9.375	30.44	0.0000E+00	.7500	0.0000E+00	22.19
8	-35.07	107.1	0.0000E+00	4.571	0.0000E+00	76.64
9	-21.93	47.14	0.0000E+00	0.0000E+00	0.0000E+00	25.29
10	-35.36	74.57	0.0000E+00	0.0000E+00	0.0000E+00	39.64
11	-82.07	207.5	0.0000E+00	0.0000E+00	0.0000E+00	125.4
12	-63.62	207.8	0.0000E+00	-1.000	0.0000E+00	143.4
13	-64.17	142.8	0.0000E+00	-1.167	0.0000E+00	77.42
14	-27.82	62.91	0.0000E+00	0.0000E+00	0.0000E+00	35.00
15	-48.64	107.0	0.0000E+00	0.0000E+00	0.0000E+00	58.18
16	.2000	7.100	0.0000E+00	-16.50	0.0000E+00	-9.100
17	-6.900	-11.50	1.000	80.00	0.0000E+00	62.50
18	-16.10	34.00	0.0000E+00	0.0000E+00	0.0000E+00	18.00
19	-93.89	260.7	0.0000E+00	-1.444	0.0000E+00	165.6
20	-37.22	132.2	0.0000E+00	0.0000E+00	0.0000E+00	95.22
21	-17.67	57.67	0.0000E+00	0.0000E+00	0.0000E+00	40.11
22	-13.67	43.33	0.0000E+00	0.0000E+00	0.0000E+00	29.56
23	-21.33	70.67	0.0000E+00	-1.667	0.0000E+00	47.89
24	-75.00	245.1	0.0000E+00	.2500	0.0000E+00	170.4
25	-8.750	-51.50	0.0000E+00	179.6	0.0000E+00	119.4
26	-4.571	-20.57	0.0000E+00	75.43	0.0000E+00	50.57
27	-54.67	67.00	1.000	296.3	0.0000E+00	308.8
28	-40.00	18.20	0.0000E+00	232.6	0.0000E+00	211.2
29	-20.60	24.60	0.0000E+00	84.00	0.0000E+00	88.20
30	-51.00	159.6	0.0000E+00	0.0000E+00	0.0000E+00	108.6
31	11.40	-26.40	0.0000E+00	0.0000E+00	0.0000E+00	-14.80
32	-2.200	-121.8	1.000	288.4	0.0000E+00	165.0
33	-21.20	43.00	0.0000E+00	48.00	0.0000E+00	69.80
34	3.750	-153.3	0.0000E+00	318.3	0.0000E+00	169.3
35	-.2500	-12.25	0.0000E+00	30.75	0.0000E+00	17.75
36	-63.25	201.8	0.0000E+00	53.50	0.0000E+00	192.0
37	-6.000	11.00	0.0000E+00	18.75	0.0000E+00	24.00
38	-81.50	166.5	0.0000E+00	216.3	0.0000E+00	301.8
39	1.250	-243.3	1.000	548.3	0.0000E+00	306.3
40	-89.00	293.0	0.0000E+00	29.33	0.0000E+00	234.0
41	-52.00	-49.67	0.0000E+00	501.7	0.0000E+00	399.7
42	-12.33	6.000	0.0000E+00	0.0000E+00	0.0000E+00	-6.000
43	-7.000	-185.7	0.0000E+00	482.7	0.0000E+00	289.7
44	2.000	-274.0	0.0000E+00	612.0	0.0000E+00	340.0
45	15.50	-48.50	0.0000E+00	-15.50	0.0000E+00	-48.50
46	-41.00	-59.00	1.000	461.0	0.0000E+00	361.5
47	11.00	-219.0	0.0000E+00	414.0	0.0000E+00	205.0
48	-97.00	117.0	0.0000E+00	472.0	0.0000E+00	492.0
49	-19.00	102.0	0.0000E+00	-78.00	0.0000E+00	4.000
50	-43.00	108.0	0.0000E+00	84.00	0.0000E+00	149.0

ncl	% inertie	stat-loc	ccc
1	0.000	8.26751	0.000E+00
2	35.330	8.37960	-1.26
3	60.710	8.35452	6.69
4	71.870	8.48927	14.8
5	81.600	8.51847	26.3
6	84.560	8.60370	43.5
7	87.020	8.68588	57.3
8	89.120	8.61275	77.8
9	91.110	8.68518	93.2
10	92.680	8.81754	127.
11	93.580	8.92335	162.
12	94.360	8.96771	171.
13	94.890	9.00546	200.
14	95.410	8.87096	234.
15	95.730	8.85437	267.
16	96.040	8.75223	308.
17	96.340	8.63271	354.
18	96.600	8.66212	406.
19	96.840	8.53117	464.
20	97.020	8.59907	538.
21	97.150	8.43324	613.
22	97.290	8.28765	741.
23	97.410	8.28163	844.
24	97.520	8.16243	923.
25	97.620	8.01501	0.105E+04
26	97.710	7.80743	0.120E+04
27	97.790	7.64954	0.137E+04
28	97.860	7.47582	0.158E+04
29	97.920	7.35054	0.181E+04
30	97.980	7.19181	0.208E+04
31	98.040	7.04488	0.240E+04
32	98.090	6.81811	0.278E+04
33	98.130	6.75881	0.323E+04
34	98.180	6.55065	0.361E+04
35	98.210	6.46454	0.442E+04
36	98.240	6.34296	0.522E+04
37	98.270	6.11387	0.621E+04
38	98.290	5.86478	0.742E+04
39	98.310	5.62306	0.895E+04
40	98.330	5.35656	0.109E+05
41	98.350	5.10777	0.133E+05
42	98.360	4.79474	0.165E+05
43	98.380	4.46795	0.206E+05
44	98.390	4.11830	0.255E+05
45	98.400	3.81283	0.323E+05
46	98.410	3.43155	0.408E+05
47	98.420	2.95915	0.512E+05
48	98.430	2.41248	0.629E+05
49	98.430	1.71367	0.746E+05
50	98.440	0.00000	0.836E+05

* graphe de l'inertie expliquée *

	0.00	98.44
1	i	
2	i	
3	i	
4	i	
5	i	
6	i	
7	i	
8	i	
9	i	
10	i	
11	i	
12	i	
13	i	
14	i	
15	i	
16	i	
17	i	
18	i	
19	i	
20	i	
21	i	
22	i	
23	i	
24	i	
25	i	
26	i	
27	i	
28	i	
29	i	
30	i	
31	i	
32	i	
33	i	
34	i	
35	i	
36	i	
37	i	
38	i	
39	i	
40	i	
41	i	
42	i	
43	i	
44	i	
45	i	
46	i	
47	i	
48	i	
49	i	
50	i	
	0.00	98.44

* statistique locale *



	0.00	9.01
1	i.....	
2	i.....	
3	i.....	
4	i.....	
5	i.....	
6	i.....	
7	i.....	
8	i.....	
9	i.....	
10	i.....	
11	i.....	
12	i.....	
13	i.....	
14	i.....	
15	i.....	
16	i.....	
17	i.....	
18	i.....	
19	i.....	
20	i.....	
21	i.....	
22	i.....	
23	i.....	
24	i.....	
25	i.....	
26	i.....	
27	i.....	
28	i.....	
29	i.....	
30	i.....	
31	i.....	
32	i.....	
33	i.....	
34	i.....	
35	i.....	
36	i.....	
37	i.....	
38	i.....	
39	i.....	
40	i.....	
41	i.....	
42	i.....	
43	i.....	
44	i.....	
45	i.....	
46	i.....	
47	i.....	
48	i.....	
49	i.....	
50	i.....	

0.00

9.01

critere cubique de classification *

	-1.26	1051.65
1	i	
2	i	
3	i	
4	i.	
5	i.	
6	i..	
7	i...	
8	i.....	
9	i.....	
10	i.....	
11	i.....	
12	i.....	
13	i.....	
14	i.....	
15	i.....	
16	i.....	
17	i.....	
18	i.....	
19	i.....	
20	i.....	
21	i.....	
22	i.....	
23	i.....	
24	i.....	
25	i.....	
	-1.26	1051.65

• noyaux finaux •

obs	dcnu	dcch	dchy	dcfg	dcdf	dcex
1	-.4916	1.737	0.0000E+00	0.4538E-01	0.0000E+00	1.287
2	-28.29	73.31	0.0000E+00	6.777	0.0000E+00	51.91
3	-77.27	227.4	0.0000E+00	5.451	0.0000E+00	155.7
4	-5.333	-48.27	1.000	149.5	0.0000E+00	96.67
5	-11.35	-98.17	0.0000E+00	302.6	0.0000E+00	193.1
6	-33.75	-57.42	1.000	407.8	0.0000E+00	316.8
7	-72.38	79.25	0.0000E+00	355.3	0.0000E+00	362.3
nb echanges :		56	i inter :	0.6575620E+04	i intra :	0.8143803E+03
nb echanges :		4	i inter :	0.6576348E+04	i intra :	0.8136525E+03
nb echanges :		0	i inter :	0.6576348E+04	i intra :	0.8136525E+03

partition en 7 classes
 inertie expliquée : 6576.348
 inertie perdue : 813.6525
 % inertie extrait : 88.98980
 fisher (6;1493) : 2011.196
 nombre d iterations : 3
 nombre d echanges : 0

description de la classe : 1
 effectif : 0.123E+04
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-7.85	2.47	i-
dcch	16.1	55.5	2.15	7.50	-i
dchy	0.183E-01	.134	0.000E+00	0.000E+00	i
dcfg	12.5	65.8	1.01	10.5	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	2.38	8.29	-i

description de la classe : 2
 effectif : 129.
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-29.8	10.4	*—i
dcch	16.1	55.5	79.3	29.5	i—*
dchy	0.183E-01	.134	0.000E+00	0.000E+00	i
dcfg	12.5	65.8	4.67	20.4	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	54.2	24.4	i—

description de la classe : 3
 effectif : 67.0
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-73.7	12.9	—*—*—*—i
dcch	16.1	55.5	209.	42.4	i—*—*—*—
dchy	0.183E-01	.134	0.000E+00	0.000E+00	i
dcfg	12.5	65.8	8.84	32.1	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	145.	43.3	i—*—*—

description de la classe : 4
 effectif : 19.0
 variable moypop stdpop moycla stdcla

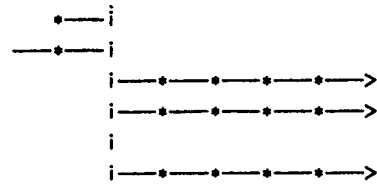
dcnu	-7.38	18.4	-12.7	18.7	—i
dcch	16.1	55.5	-109.	103.	*—*—i
dchy	0.183E-01	.134	0.000E+00	0.000E+00	i
dcfg	12.5	65.8	337.	137.	i—*—*—*—*—
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	216.	67.4	i—*—*—*—

description de la classe : 5
 effectif : 15.0
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-11.1	15.3	i
dcch	16.1	55.5	-21.9	51.4	-i
dchy	0.183E-01	.134	1.00	0.000E+00	i—*—*—*—*—>
dcfg	12.5	65.8	133.	87.6	i—*—
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.2	57.0	101.	68.4	i—*—

description de la classe : 6
 effectif : 12.0
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-26.6	28.8
dcch	16.1	55.5	-90.4	139.
dchy	0.183E-01	.134	1.00	0.000E+00
dcfg	12.5	65.8	428.	106.
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.2	57.0	312.	67.7



description de la classe : 7
 effectif : 6.00
 variable moypop stdpop moycla stdcla

dcnu	-7.38	18.4	-72.0	21.7
dcch	16.1	55.5	52.0	106.
dchy	0.183E-01	.134	0.000E+00	0.000E+00
dcfg	12.5	65.8	419.	113.
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.2	57.0	399.	59.9



```
*****
* SICLA                      Cembrzynski                      Date : 03/06/87 *
* S.D. :                      Archive :                      *
* Listing : lis08_44          Mode interactif              *
* version 2.2 : (faire help sicla)                          *
* _____ *
*****
```

```
*****
* SICLA                      Commande PONDER                Version 10/04/87 *
*****
* donnees : part_renf.sd      individus :1500 * variables : 12 *
* _____ *
*                               quantitatives : 7   qualitatives : 5 *
*****
```

```
-----*
* filtrage d elements douteux *
*-----*
```

```
le point le plus eloigne : 1404
nb de classes : 2 nb transferts : 0 % inertie : 16.67

creation d un noyau avec : 1366
nb de classes : 3 nb transferts : 1 % inertie : 26.98

creation d un noyau avec : 1046
nb de classes : 4 nb transferts : 1 % inertie : 32.27

creation d un noyau avec : 722
nb de classes : 5 nb transferts : 35 % inertie : 45.34
```

```
-----*
* filtrage d elements douteux *
*-----*
```

```
le point le plus eloigne : 101
nb de classes : 2 nb transferts : 2 % inertie : 13.61

creation d un noyau avec : 452
nb de classes : 3 nb transferts : 1 % inertie : 16.39

creation d un noyau avec : 1348
nb de classes : 4 nb transferts : 74 % inertie : 30.85
```

```
-----*
* filtrage d elements douteux *
*-----*
```

```
le point le plus eloigne : 1483
nb de classes : 2 nb transferts : 2 % inertie : 16.83

creation d un noyau avec : 1348
nb de classes : 3 nb transferts : 74 % inertie : 31.35
```

```
-----*
* filtrage d elements douteux *
*-----*
```

```
le point le plus eloigne : 597
nb de classes : 2 nb transferts : 20 % inertie : 12.99

creation d un noyau avec : 1348
nb de classes : 3 nb transferts : 76 % inertie : 28.16
liste des points aberrants
```

97 101 452 468 480 602 920 1046 1198 1277 1366 1404 1483

```

*****
* SICLA                      Cembrzynski                      Date : 03/06/87 *
* S.D. : part_renf.sdo      Archive : part_renf.sar          *
* Listing : lis08_44        Mode interactif                *
*****
* SICLA                      Commande NUEDYN                 Version 25/02/87 *
*****
* donnees : part_renf.sd    individus :1500 * variables : 13    *
* _____                _____                _____ *
*                               quantitatives : 8    qualitatives : 5    *
*****

```

impression de l inverse de la metrique

```

dcnu 336.15
dcch -877.71      3141.6
dchy -.24449     -.83041      0.28129E-01
dcfg -197.05     -1304.4      3.9896      4575.1
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -739.37     960.78      2.9306      3074.9      0.00000E+00 3298.6

```

totale : 0.7435000E+04 i inter : 0.6667252E+04 i intra : 0.7677479E+03

```

iteration : 1 i inter : 0.6685440E+04 i intra : 0.7495600E+03
iteration : 2 i inter : 0.6701041E+04 i intra : 0.7339593E+03

```

Ecriture sur l archive :part_renf.sar

Constituant numero : 1

Nom du constituant
abrege : p 50

Libelle :partition avec NUEDYN 901 %.

```

Type du constituant : partition
Date creation      : 03/06/87
Nom de la S.D     : part_renf.sdo

```

* barycentres *

obs	dcnu	dcch	dchy	dcfg	dcd	dcex
1	-0.8405E-01	.2392	0.0000E+00	0.3017E-01	0.0000E+00	.1853
2	-.2809	2.213	0.0000E+00	.3596	0.0000E+00	1.292
3	-.3253	2.229	0.0000E+00	0.1205E-01	0.0000E+00	1.916
4	-1.240	1.890	0.0000E+00	0.4000E-01	0.0000E+00	.6900
5	-1.064	3.574	0.0000E+00	-.2660	0.0000E+00	3.245
6	-1.222	-.4815	0.0000E+00	3.241	0.0000E+00	2.537
7	-25.22	57.30	0.0000E+00	0.0000E+00	0.0000E+00	32.07
8	.5926	1.481	0.0000E+00	0.0000E+00	0.0000E+00	2.074
9	-1.720	8.080	0.0000E+00	.9200	0.0000E+00	7.280
10	-1.400	1.050	0.0000E+00	0.0000E+00	0.0000E+00	1.650
11	-5.850	11.60	0.0000E+00	-.9000	0.0000E+00	4.850
12	-12.74	38.95	0.0000E+00	9.789	0.0000E+00	37.00
13	-77.67	184.8	0.0000E+00	0.0000E+00	0.0000E+00	107.2
14	-66.72	214.4	0.0000E+00	7.444	0.0000E+00	155.2
15	-9.889	30.61	0.0000E+00	1.333	0.0000E+00	22.06
16	-2.000	13.81	0.0000E+00	-2.188	0.0000E+00	9.625
17	-40.14	131.9	0.0000E+00	6.000	0.0000E+00	97.71
18	-38.15	78.08	0.0000E+00	0.0000E+00	0.0000E+00	39.92
19	-33.08	99.77	0.0000E+00	1.692	0.0000E+00	68.46
20	-21.23	66.08	0.0000E+00	9.923	0.0000E+00	54.77
21	0.0000E+00	0.0000E+00	0.0000E+00	0.0000E+00	0.0000E+00	0.0000E+00
22	-12.92	-2.750	1.000	103.1	0.0000E+00	88.33
23	-2.091	12.00	0.0000E+00	0.0000E+00	0.0000E+00	8.273
24	-25.00	80.60	0.0000E+00	-7.800	0.0000E+00	46.80
25	-37.10	117.6	0.0000E+00	1.400	0.0000E+00	83.00
26	-87.20	246.8	0.0000E+00	-1.300	0.0000E+00	158.6
27	-68.33	160.9	0.0000E+00	0.0000E+00	0.0000E+00	91.44
28	-4.375	-14.88	0.0000E+00	65.63	0.0000E+00	46.38
29	-70.71	248.7	0.0000E+00	6.857	0.0000E+00	185.0
30	-22.00	43.86	0.0000E+00	7.714	0.0000E+00	30.57
31	-26.00	56.43	0.0000E+00	6.000	0.0000E+00	35.43
32	-1.143	-2.857	-1.286	15.57	0.0000E+00	10.71
33	-8.333	-52.50	0.0000E+00	175.3	0.0000E+00	114.7
34	-1.000	-220.7	0.0000E+00	514.3	0.0000E+00	292.3
35	-2.200	-121.8	1.000	288.4	0.0000E+00	165.0
36	-49.60	115.6	0.0000E+00	-2.800	0.0000E+00	64.80
37	-34.00	4.400	0.0000E+00	245.4	0.0000E+00	216.0
38	-78.00	166.8	0.0000E+00	192.0	0.0000E+00	281.4
39	-57.80	60.20	1.000	350.6	0.0000E+00	352.8
40	-34.00	51.00	0.0000E+00	36.20	0.0000E+00	54.20
41	.5000	5.250	0.0000E+00	-15.50	0.0000E+00	-9.750
42	3.750	-153.3	0.0000E+00	318.3	0.0000E+00	169.3
43	1.250	-243.3	1.000	548.3	0.0000E+00	306.3
44	11.33	-13.00	0.0000E+00	0.0000E+00	0.0000E+00	-1.667
45	-17.00	3.000	0.0000E+00	102.7	0.0000E+00	88.00
46	8.667	-34.67	0.0000E+00	-10.33	0.0000E+00	-36.33
47	-76.00	27.50	0.0000E+00	522.5	0.0000E+00	474.0
48	-50.50	-43.50	0.0000E+00	466.0	0.0000E+00	371.5
49	-35.50	194.0	0.0000E+00	7.000	0.0000E+00	165.5
50	4.000	-211.0	-1.000	442.0	0.0000E+00	234.5
51	-35.00	-99.00	1.000	510.0	0.0000E+00	377.0

ncl	% inertie	stat-loc	ccc
1	0.000	7.53749	0.000E+00
2	14.460	7.49696	2.00
3	28.450	7.60953	3.94
4	41.340	7.76160	16.1
5	51.470	7.81214	28.2
6	61.000	7.69039	41.4
7	67.060	7.53537	62.3
8	71.280	7.76138	69.0
9	73.810	7.69110	85.9
10	76.250	7.61130	105.
11	78.410	7.48719	124.
12	80.510	7.55212	179.
13	81.940	7.70769	219.
14	83.130	7.68747	245.
15	84.170	7.84822	284.
16	85.090	8.01406	327.
17	85.870	7.90684	375.
18	86.560	7.73558	406.
19	87.170	7.78818	464.
20	87.780	7.62517	529.
21	88.370	7.55893	602.
22	88.940	7.52998	684.
23	89.470	7.51954	778.
24	89.840	7.36865	884.
25	90.170	7.26170	992.
26	90.480	7.24620	0.113E+04
27	90.780	7.11115	0.129E+04
28	91.040	6.92459	0.147E+04
29	91.270	6.73103	0.169E+04
30	91.470	6.49944	0.193E+04
31	91.670	6.30227	0.222E+04
32	91.860	6.11062	0.256E+04
33	92.030	5.88810	0.296E+04
34	92.190	5.70626	0.344E+04
35	92.340	5.52082	0.405E+04
36	92.480	5.53808	0.461E+04
37	92.610	5.35503	0.544E+04
38	92.740	5.08784	0.645E+04
39	92.840	5.01149	0.765E+04
40	92.910	4.71660	0.920E+04
41	92.990	4.45597	0.112E+05
42	93.050	4.17276	0.137E+05
43	93.100	3.84234	0.169E+05
44	93.150	3.50283	0.211E+05
45	93.190	3.08072	0.264E+05
46	93.230	2.62743	0.332E+05
47	93.260	2.07003	0.418E+05
48	93.280	1.39404	0.523E+05
49	93.290	0.49586	0.640E+05
50	93.290	0.00000	0.755E+05
51	93.290	0.00000	0.843E+05

* graphe de l'inertie expliquée *



0.00

93.29

- 1 i
- 2 i
- 3 i
- 4 i
- 5 i
- 6 i
- 7 i
- 8 i
- 9 i
- 10 i
- 11 i
- 12 i
- 13 i
- 14 i
- 15 i
- 16 i
- 17 i
- 18 i
- 19 i
- 20 i
- 21 i
- 22 i
- 23 i
- 24 i
- 25 i
- 26 i
- 27 i
- 28 i
- 29 i
- 30 i
- 31 i
- 32 i
- 33 i
- 34 i
- 35 i
- 36 i
- 37 i
- 38 i
- 39 i
- 40 i
- 41 i
- 42 i
- 43 i
- 44 i
- 45 i
- 46 i
- 47 i
- 48 i
- 49 i
- 50 i
- 51 i

0.00

93.29

• statistique locale •



0.00

8.01

- 1 i.....
- 2 i.....
- 3 i.....
- 4 i.....
- 5 i.....
- 6 i.....
- 7 i.....
- 8 i.....
- 9 i.....
- 10 i.....
- 11 i.....
- 12 i.....
- 13 i.....
- 14 i.....
- 15 i.....
- 16 i.....
- 17 i.....
- 18 i.....
- 19 i.....
- 20 i.....
- 21 i.....
- 22 i.....
- 23 i.....
- 24 i.....
- 25 i.....
- 26 i.....
- 27 i.....
- 28 i.....
- 29 i.....
- 30 i.....
- 31 i.....
- 32 i.....
- 33 i.....
- 34 i.....
- 35 i.....
- 36 i.....
- 37 i.....
- 38 i.....
- 39 i.....
- 40 i.....
- 41 i.....
- 42 i.....
- 43 i.....
- 44 i.....
- 45 i.....
- 46 i.....
- 47 i.....
- 48 i.....
- 49 i.....
- 50 i.....
- 51 i.....

0.00

8.01

critere cubique de classification *

	0.00	992.15
1 i		
2 i.		
3 i		
4 i.		
5 i.		
6 i..		
7 i....		
8 i.....		
9 i.....		
10 i.....		
11 i.....		
12 i.....		
13 i.....		
14 i.....		
15 i.....		
16 i.....		
17 i.....		
18 i.....		
19 i.....		
20 i.....		
21 i.....		
22 i.....		
23 i.....		
24 i.....		
25 i.....	0.00	992.15

impression de l inverse de la metrique

dcnu 122.69
 dcch -347.09 1286.1
 dchy -0.32360E-01 0.71925E-01 0.20815E-02
 dcfg -54.960 -458.49 0.93012E-01 1451.3
 dcd f 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -279.29 480.19 .13367 937.91 0.00000E+00 1138.6

* noyaux finaux *

obs	dcnu	dcch	dchy	dcfg	dcd f	dcex
1	-2.268	5.242	0.0000E+00	4.979	0.0000E+00	7.956
2	-1.623	6.693	0.0000E+00	-0.7035E-01	0.0000E+00	3.965
3	-4.096	11.41	0.0000E+00	1.777	0.0000E+00	10.19
4	-47.68	112.3	0.0000E+00	2.475	0.0000E+00	67.13
5	-67.41	213.2	0.0000E+00	36.13	0.0000E+00	182.1
6	-17.96	-52.33	1.000	264.3	0.0000E+00	194.5
7	-1.143	-2.857	-1.286	15.57	0.0000E+00	10.71
8	-20.92	-148.2	-1.667	495.6	0.0000E+00	326.2
totale :	0.3976869E+05	i inter :	0.3313155E+05	i intro :	0.6637145E+04	

impression de l inverse de la metrique

dcnu 111.80
dcch -300.43 1023.9
dchy-0.26721E-01-0.13376E-01-0.24231E-03
dcfg -25.541 -347.39 .27289 977.82
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -214.03 375.38 .23223 605.20 0.00000E+00 766.34
totale : 0.3071278E+06 i inter : 0.2856681E+06 i intra : 0.2145975E+05
impression de l inverse de la metrique

dcnu 112.78
dcch -285.61 1097.0
dchy-0.15939E-01 .10276 -0.11576E-03
dcfg -46.853 -680.88 -0.72244E-01 1915.8
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -219.51 129.98 0.14284E-01 1187.7 0.00000E+00 1097.5
totale : 0.5044668E+06 i inter : 0.4774941E+06 i intra : 0.2697263E+05
impression de l inverse de la metrique

dcnu 108.92
dcch -255.97 1206.6
dchy-0.19757E-02 .28110 0.74077E-04
dcfg -84.402 -1165.9 -.59054 3323.6
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -231.23 -215.49 -.31134 2071.9 0.00000E+00 1623.8
totale : 0.2136902E+05 i inter : 0.1622938E+05 i intra : 0.5139639E+04
impression de l inverse de la metrique

dcnu 106.36
dcch -271.48 920.01
dchy-0.67192E-01 -.37906 0.20822E-02
dcfg -25.782 -398.48 1.5056 1116.3
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -190.79 249.05 1.0596 693.18 0.00000E+00 751.71
totale : 0.6480518E+05 i inter : 0.5620575E+05 i intra : 0.8599431E+04
impression de l inverse de la metrique

dcnu 59.346
dcch -143.57 623.97
dchy-0.31424E-01 0.76221E-01-0.17903E-03
dcfg -27.364 -540.02 0.97102E-01 1431.4
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -110.21 -62.264 .13904 861.11 0.00000E+00 684.43
totale : 0.4957425E+06 i inter : 0.4620001E+06 i intra : 0.3374235E+05
impression de l inverse de la metrique

dcnu 80.728
dcch -179.18 937.43
dchy-0.96171E-02 .25367 0.31892E-04
dcfg -72.309 -1073.7 -.47231 3011.4
dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
dcex -169.97 -316.71 -.22972 1862.6 0.00000E+00 1372.8
totale : 0.2260999E+05 i inter : 0.1649139E+05 i intra : 0.6118601E+04

impression de l inverse de la metrique

dcnu 95.287
 dcch -245.89 753.99
 dchy 0.39346E-01 -0.50687 -0.18698E-01
 dcfg -4.8216 -197.62 1.0471 478.14
 dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -155.30 309.30 .56627 276.98 0.00000E+00 431.30
 totale : 0.4485276E+05 i inter : 0.3404334E+05 i intra : 0.1080942E+05
 impression de l inverse de la metrique

dcnu 118.57
 dcch -313.04 985.24
 dchy-0.67783E-01 -0.22601 0.45871E-02
 dcfg -12.498 -252.69 1.1100 676.47
 dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -206.85 418.56 .81839 412.28 0.00000E+00 624.21
 totale : 0.4425676E+05 i inter : 0.3322781E+05 i intra : 0.1102894E+05
 impression de l inverse de la metrique

dcnu 117.68
 dcch -312.07 1016.1
 dchy-0.31440E-01-0.21292E-01 0.86688E-03
 dcfg -16.105 -307.70 .32873 820.58
 dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -210.36 395.55 .27616 497.19 0.00000E+00 682.21
 totale : 0.7999494E+05 i inter : 0.6914919E+05 i intra : 0.1084576E+05
 impression de l inverse de la metrique

dcnu 111.04
 dcch -286.40 1005.2
 dchy-0.22726E-01 0.25629E-01-0.20012E-03
 dcfg -28.151 -471.71 .15853 1289.6
 dcdf 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00 0.00000E+00
 dcex -203.36 246.36 .16098 789.79 0.00000E+00 832.40

partition en 8 classes
 inertie expliquée : 69149.19
 inertie perdue : 10845.76
 % inertie extrait : 86.44000
 fisher (7;1492) : 1358.933
 nombre d iterations : 10
 nombre d echanges : 0

description de la classe : 1

effectif : 897.

variable moypop stdpop moycla stdcla

dcnu	-7.34	18.3	-1.73	5.71	i-
dcch	15.7	56.0	4.64	17.9	i
dchy	0.108E-01	.168	0.000E+00	0.000E+00	i
dcfg	13.1	67.6	2.05	17.9	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	4.97	17.3	-i

description de la classe : 2

effectif : 209.

variable moypop stdpop moycla stdcla

dcnu	-7.34	18.3	-5.12	10.3	i
dcch	15.7	56.0	12.2	31.3	i
dchy	0.108E-01	.168	0.000E+00	0.000E+00	i
dcfg	13.1	67.6	3.93	24.5	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	12.2	24.0	i

description de la classe : 3

effectif : 208.

variable moypop stdpop moycla stdcla

dcnu	-7.34	18.3	-2.52	7.50	i-
dcch	15.7	56.0	8.32	21.7	i
dchy	0.108E-01	.168	0.000E+00	0.000E+00	i
dcfg	13.1	67.6	1.21	13.2	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	5.98	15.9	-i

description de la classe : 4

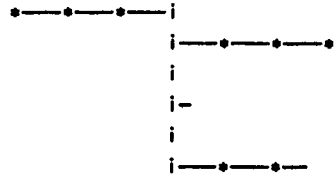
effectif : 76.0

variable moypop stdpop moycla stdcla

dcnu	-7.34	18.3	-53.0	24.8	-i
dcch	15.7	56.0	125.	70.9	i
dchy	0.108E-01	.168	0.000E+00	0.000E+00	i
dcfg	13.1	67.6	2.20	20.7	i
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00	i
dcex	21.5	57.4	73.9	46.6	i

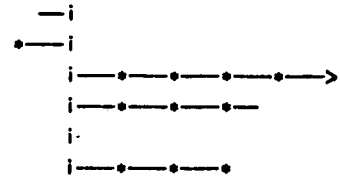
description de la classe : 5
 effectif : 48.0
 variable moypop stdpop moycla stdcia

dcnu	-7.34	18.3	-62.6	18.8
dcch	15.7	56.0	188.	70.9
dchy	0.108E-01	.168	0.000E+00	0.000E+00
dcfg	13.1	67.6	46.0	88.8
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	172.	58.7



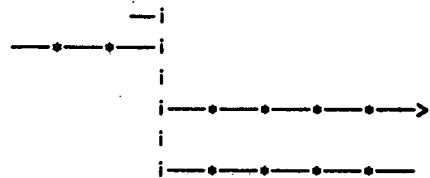
description de la classe : 6
 effectif : 27.0
 variable moypop stdpop moycla stdcia

dcnu	-7.34	18.3	-18.0	23.6
dcch	15.7	56.0	-52.3	106.
dchy	0.108E-01	.168	1.00	0.000E+00
dcfg	13.1	67.6	264.	175.
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	194.	125.



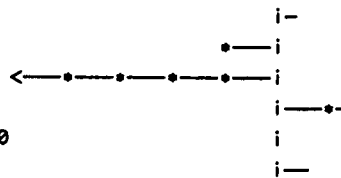
description de la classe : 7
 effectif : 13.0
 variable moypop stdpop moycla stdcia

dcnu	-7.34	18.3	-19.3	32.0
dcch	15.7	56.0	-140.	106.
dchy	0.108E-01	.168	0.000E+00	0.000E+00
dcfg	13.1	67.6	466.	94.5
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	307.	100.



description de la classe : 8
 effectif : 9.00
 variable moypop stdpop moycla stdcia

dcnu	-7.34	18.3	0.000E+00	2.75
dcch	15.7	56.0	-49.1	90.2
dchy	0.108E-01	.168	-1.22	.416
dcfg	13.1	67.6	110.	190.
dcdf	0.000E+00	0.000E+00	0.000E+00	0.000E+00
dcex	21.5	57.4	60.4	102.



* impression de la partition *

impression de la classe : 1 effectif : 897.

impression de la classe : 2 effectif : 209.

7 11 16 35 44 46 48 56 60 70 72 77 79 80 81 84
 94 113 125 138 144 153 155 156 172 174
 201 203 218 222 245 250 253 260 279 288 293 298 301 302 306 308
 311 320 322 327 332 340 343 347 349 351
 352 367 375 376 379 380 387 392 406 409 415 421 428 470 471 476
 482 488 490 497 498 516 520 523 524 525
 531 532 541 542 551 563 571 578 585 598 600 607 612 634 640 650
 655 665 669 671 677 678 690 691 708 714
 718 743 747 753 754 767 769 773 781 783 789 790 796 797 798 804
 816 817 821 825 833 838 847 851 853 860
 862 867 868 874 881 885 887 894 895 899 913 924 925 936 942 958
 964 965 974 975 985 990 1014 1021 1026 1032
 1036 1038 1039 1064 1072 1081 1086 1094 1104 1113 1140 1156 1170 1171 1211 1213
 1228 1236 1237 1238 1244 1258 1262 1264 1286 1295
 1297 1300 1301 1319 1321 1331 1334 1339 1382 1396 1399 1405 1420 1434 1440 1442
 1447 1448 1456 1463 1464 1468 1469 1472 1480 1482
 1499

impression de la classe : 3 effectif : 208.

19 24 27 32 36 45 49 55 57 58 73 86 92 109 110 117
 119 121 129 132 150 158 165 166 177 178
 179 182 184 198 204 206 238 246 248 254 257 261 262 272 280 281
 283 284 296 310 323 339 345 357 361 366
 369 372 386 391 393 395 402 413 422 424 438 440 447 486 489 494
 505 517 521 527 530 535 546 549 553 554
 555 562 565 592 594 622 623 625 631 639 656 662 670 679 683 697
 705 712 713 715 719 723 731 748 762 764
 765 766 785 793 799 800 801 805 806 823 826 828 832 837 840 858
 859 864 870 875 876 878 884 898 901 909
 910 915 918 923 931 938 941 969 981 988 992 999 1005 1006 1009 1010
 1024 1029 1033 1045 1049 1056 1079 1084 1087 1095
 1101 1108 1121 1136 1137 1143 1144 1145 1148 1155 1172 1180 1194 1221 1223 1229
 1247 1248 1250 1266 1278 1292 1293 1298 1299 1305
 1312 1316 1325 1340 1341 1342 1343 1355 1361 1365 1367 1392 1393 1410 1423 1424
 1430 1439 1445 1452 1454 1461 1481 1485 1494 1495
 impression de la classe : 4 effectif : 76.0

9 21 23 71 95 104 131 133 148 164 173 181 189 215 226 233
 299 307 338 368 374 397 400 433 437 443
 455 461 569 576 618 649 666 692 695 716 727 733 738 758 786 820
 822 849 850 852 869 886 902 914 966 995
 1003 1071 1082 1083 1138 1154 1160 1186 1209 1230 1240 1265 1274 1289 1347 1362
 1390 1391 1412 1415 1421 1429 1462 1467
 impression de la classe : 5 effectif : 48.0

17 115 136 154 196 197 205 207 227 275 336 408 426 604 605 609
 630 675 676 686 693 694 698 717 732 749
 771 871 904 949 960 963 978 1063 1132 1167 1173 1227 1233 1243 1308 1328
 1337 1349 1359 1401 1426 1479
 impression de la classe : 6 effectif : 27.0

122 161 234 295 300 304 431 465 485 550 614 624 651 701 711 839
 857 940 970 1061 1100 1303 1348 1373 1450 1471
 1493

impression de la classe : 7 effectif : 13.0

146 213 499 722 787 916 971 1208 1329 1386 1397 1431 1490
 impression de la classe : 8 effectif : 9.00

235 526 570 580 597 744 1002 1225 1263

Ecriture sur l archive :part_renf.sar

Constituant numero : 2

Nom du constituant
abrege :p 8

Libelle :partition avec NUEDYN 864 %.

Type du constituant : partition
Date creation : 03/06/87
Nom de la S.D : part_renf.sdo

— SICLA ————— FIN COMMANDE : NUEDYN —————

