

M. MBUYI

M. MARAVALLE

Typologie des huiles brutes de pétrole

Les cahiers de l'analyse des données, tome 9, n° 3 (1984),
p. 301-314

http://www.numdam.org/item?id=CAD_1984__9_3_301_0

© Les cahiers de l'analyse des données, Dunod, 1984, tous droits réservés.

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

NUMDAM

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

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

TYPOLOGIE DES HUILES BRUTES DE PÉTROLE

[HUILES BRUTES]

par M. Mbuyi ⁽¹⁾ et M. Maravalle ⁽²⁾

0 Introduction

Les huiles brutes de pétrole sont des mélanges d'hydrocarbures d'origine naturelle, avec des traces variables de composés divers organiques ou non, comprenant du soufre, du phosphore, de l'azote, de l'oxygène, du calcium, du magnésium, du silicium, du nickel, du vanadium etc. Ces produits sont plus ou moins denses, plus ou moins fluides, il en est de lourds et de visqueux, de légers et de volatils.

Dans la présente étude, nous analysons 60 types d'huiles brutes obtenues par forage dans les gisements de différents pays. Les caractéristiques de chaque type de pétrole brut sont décrites par 11 paramètres.

1 Présentation des données

Les données de base constituent un tableau rectangulaire $K_0(I, J)$ où I est l'ensemble de 60 types d'huiles brutes et J l'ensemble des 11 variables qui sont mesurées par le chimiste sur chaque type d'huile brute.

On a affecté à chaque échantillon le poids (colonne Reser) qui représente pour chaque $i \in I$ l'importance relative $A(i)$ de son gisement. Ces données étant non homogènes, nous avons construit des histogrammes détaillés des variables afin de les découper en classes d'effectifs égaux et de faire un codage par modalités. Nous avons réalisé ainsi un ensemble homogène de données et toutes les lignes du tableau à analyser ont à peu près le même poids. Nous donnons ci-dessous le sigle, le nombre des modalités et la signification pour chaque variable.

1° DLE : % en poids des produits volatils obtenus après distillation sous conditions atmosphériques. Les valeurs associées à cette variable sont obtenues par la courbe de distillation (TBP) à 180° C, faite pour chaque type de pétrole brut (0,8 à 45,9%). Cette variable est éclatée en 4 modalités : DLE1, DLE2, DLE3, DLE4.

2° DTO : % en poids des produits volatils obtenus par le pic de la courbe de distillation à 370° C, (19,8 à 86,3%). Cette variable est éclatée en 4 modalités : DTO1, DTO2, DTO3, DTO4.

3° SM : Teneur en soufre présente dans les échantillons obtenus après distillation à 180° et 370° C, (0 à 2,1%). Elle est éclatée en 2 modalités SM1, SM2.

(1) Docteur 3° cycle en statistique. Université Pierre et Marie Curie.
(2) Université de L'AQUILA- AGIP (Italie)(ENI Groupe Project OAPECS).

4° FCC : % en poids des distillats obtenus après distillation sous vide, les valeurs associées à cette variable sont obtenues par la courbe de distillation du type (TBP) à 370° et 540° C et ce pour chaque type de brut, (0,8 à 35,7%). 4 modalités : FCC1, FCC2, FCC3, FCC4.

5° DEN : densité des distillats, à 15° C exprimée en FCC, (0,84 à 0,97). Elle est éclatée en deux modalités : DEN1 et DEN2.

6° TAR : % des résidus à courte chaîne (de molécules d'hydrocarbure CO), obtenus après distillation sous vide, (1,7 à 61,3%). 4 modalités : TAR1, TAR2, TAR3, TAR4.

7° ST : % teneur en soufre présente dans les résidus à courte chaîne, (0,06 à 9,7%). 3 modalités : ST1, ST2, ST3.

8° VBN : indice de viscosité pour les résidus à courte chaîne, (35,4 à 60%). Elle est éclatée en 4 modalités : VBN1, VBN2, VBN3, VBN4.

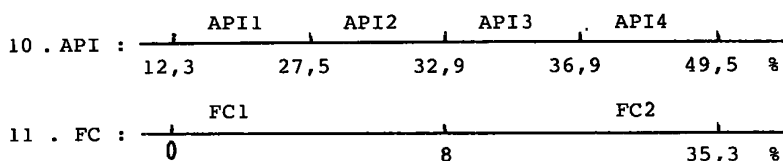
9° MET : teneur en nickel et vanadium des résidus à courte chaîne, (0 à 11,44%). Elle est éclatée en 4 modalités : MET1, MET2, MET3, MET4.

10° API : indice de gravité pour chaque type de brut. Degré API, du nom de l'American Petroleum Institute qui a créé cette échelle. Une densité plus élevée correspond à un faible API.

11° FC : on prend le chiffre correspondant à FCC si la valeur de la densité correspondante (DEN) est égale ou supérieure à 0,93, autrement FC est égale à zéro, (0 à 35,30). Deux modalités : FC1 et FC2.

Variables d'indication :

1. DLE :	DEL1	DEL2	DEL3	DEL4	
	0,80	16,5	20,9	25,6	45,9 %
2. DTO :	DTO1	DTO2	DTO3	DTO4	
	19,8	45,3	55,1	62,5	86,3 %
3. SM :	SM1	SM2			
	0	0,6		2,1	%
4. FCC :	FCC1	FCC2	FCC3	FCC4	
	8	21,7	25,1	26,9	35,7 %
5. DEN :	DEN1	DEN2			
	0,84	0,922		0,97	%
6. TAR :	TAR1	TAR2	TAR3	TAR4	
	1,70	10,3	17,2	25,8	61,3 %
7. ST :	ST1	ST2	ST3		
	0,06	1	3,9	9,7	%
8. VBN :	VBN1	VBN2	VBN3	VBN4	
	35,4	44,9	47,6	49,2	60 %
9. MET :	MET1	MET2	MET3	MET4	
	0	0,4	1	2,8	11,4 %



2 Les analyses effectuées

A partir du tableau $K_1(I, J)$ sous forme disjonctive nous avons construit le tableau de Burt en posant :

$$b_1(j, j') = \text{Card}\{i \in I ; K_1(i, j) = K_1(i, j') = 1\}$$

c'est-à-dire nombre des individus i de I , rentrant simultanément dans les modalités j et j' . En général on utilisera la formule :

$$b_1(j, j') = \sum \{K_1(i, j) \cdot K_1(i, j') / K_1(i) \mid i \in I\}$$

nous avons ainsi obtenu le tableau $B_1(J, J) = \{b_1(j, j')\}$.

Nous avons constitué le tableau $K_2(I, J)$ en pondérant la ligne $i \in I$ du tableau $K_1(I, J)$ par le coefficient $A(i)$ qui représente l'abondance du gisement (i).

$$K_2(I, J) = \{A(i) \cdot K_1(i, j) \mid i \in I, j \in J\}.$$

Et le tableau de Burt correspondant à $K_2(I, J)$.

$$b_2(j, j') = \sum \{K_2(i, j) \cdot K_2(i, j') / K_2(i) \mid i \in I\}.$$

$$B_2(J, J) = \{b_2(j, j')\}$$

2.1 Analyse du tableau $B_1(J, J)$ avec le tableau $K_1(I, J)$ en supplémentaires.

2.2 Classification ascendante hiérarchique sur le tableau $K_1(I, J)$ à l'aide du programme VACOR.

2.3 Analyse du tableau $B_2(J, J)$ avec le tableau $K_2(I, J)$ en supplémentaires.

2.4 Classification ascendante hiérarchique sur le tableau $K_2(I, J)$ interprétée à l'aide du programme VACOR.

2.5 Discrimination barycentrique entre les huiles moins soufrées et plus soufrées.

3 Analyse du tableau $B_1(J, J)$ avec $K_1(I, J)$ en supplémentaires

3.1 Le premier facteur : Ce facteur nettement détaché de celui qui le suit ($\tau_1 = 47,6\%$; $\tau_2 = 21,5\%$), est un facteur de niveau général opposant des huiles moins denses aux huiles plus denses, et moins soufrées aux huiles plus soufrées. Il est dominé par :

- Du côté négatif : DLE4 (plus de 27% de produits volatils obtenus après distillation à 180° C), DTO4 (plus de 62% de produits volatils obtenus après distillation à 370° C), SM1 (teneur en soufre après distillation à 180 et 370° C inférieure à 0,6%), TAR1 (% en poids des résidus à courte chaîne de CO entre 1,7 et 10,5%). A ces modalités s'associent les huiles suivantes : Zueitina de la Libye (ZUE),

Nigerian Bross du Nigéria (BRA), El Borma de la Tunisie (ELB), Malossa de l'Italie (MAL), Murban de l'Abu Dhabi (MUR), Sahara mélange de l'Algérie (SAH), Zarzaitine de l'Algérie (ZAR).

- Du côté positif : les caractéristiques opposées : SM2 (teneur en soufre supérieure à 0,7%), ST3 (teneur en soufre dans le résidu, supérieure à 4,1%), vont avec les huiles suivantes : Arabian Heavy de l'Arabie Saoudite (ARH), Basrah Heavy de l'Irak (BAH), Beloy Blend de l'Egypte (BEL). En bref, on note sur cet axe, l'opposition entre le pétrole de bonne qualité, huiles légères contenant peu de soufre à l'exemple de Zarzaitine de l'Algérie, Zueïtina de la Libye et Nigéria Brass du Nigéria aux huiles plus denses, plus lourdes et contenant plus de soufre à l'exemple de Bassah Heavy de l'Irak.

Il convient de signaler sur cet axe, l'opposition entre la modalité SM2 (teneur en soufre supérieure à 0,7% dans les échantillons, obtenue après distillation à 180° et 370° C) avec les gisements suivants : Ninian Blend du Royaume-Uni (NIN), Flotta du Royaume-Uni (FLO), Ekofisk de la Norvège (EKO) et Brent Blend du Royaume-Uni (BRN). Cette opposition s'explique par le fait que le pétrole de la mer du Nord contient peu de soufre.

Les huiles de l'Irak : Kirkuk Blend (KIR), Basrah Heavy (BAH), Bassah Light (BAL) et Basrah medium (BAM) qui ont une teneur élevée en soufre s'opposent aux modalités SM1 et ST1.

Cet axe montre enfin que la densité conditionne aussi la teneur en produits légers les plus volatils. Du côté $F_1 < 0$, les modalités DEN1, API4 (densité entre 0,84 et 0,92) en association avec les huiles légères comme Abuttifel de la Libye (BUA), El Borma de la Tunisie (ELB), Zueïtina de la Libye (ZUE). Du côté $F_1 > 0$, les modalités DEN2, API1 (densité supérieure à 0,93) vont avec les huiles plus lourdes : Tia-Juana H du Vénézuéla (TJH), Maya du Mexique (MAY), Souéie de la Syrie (SOU).

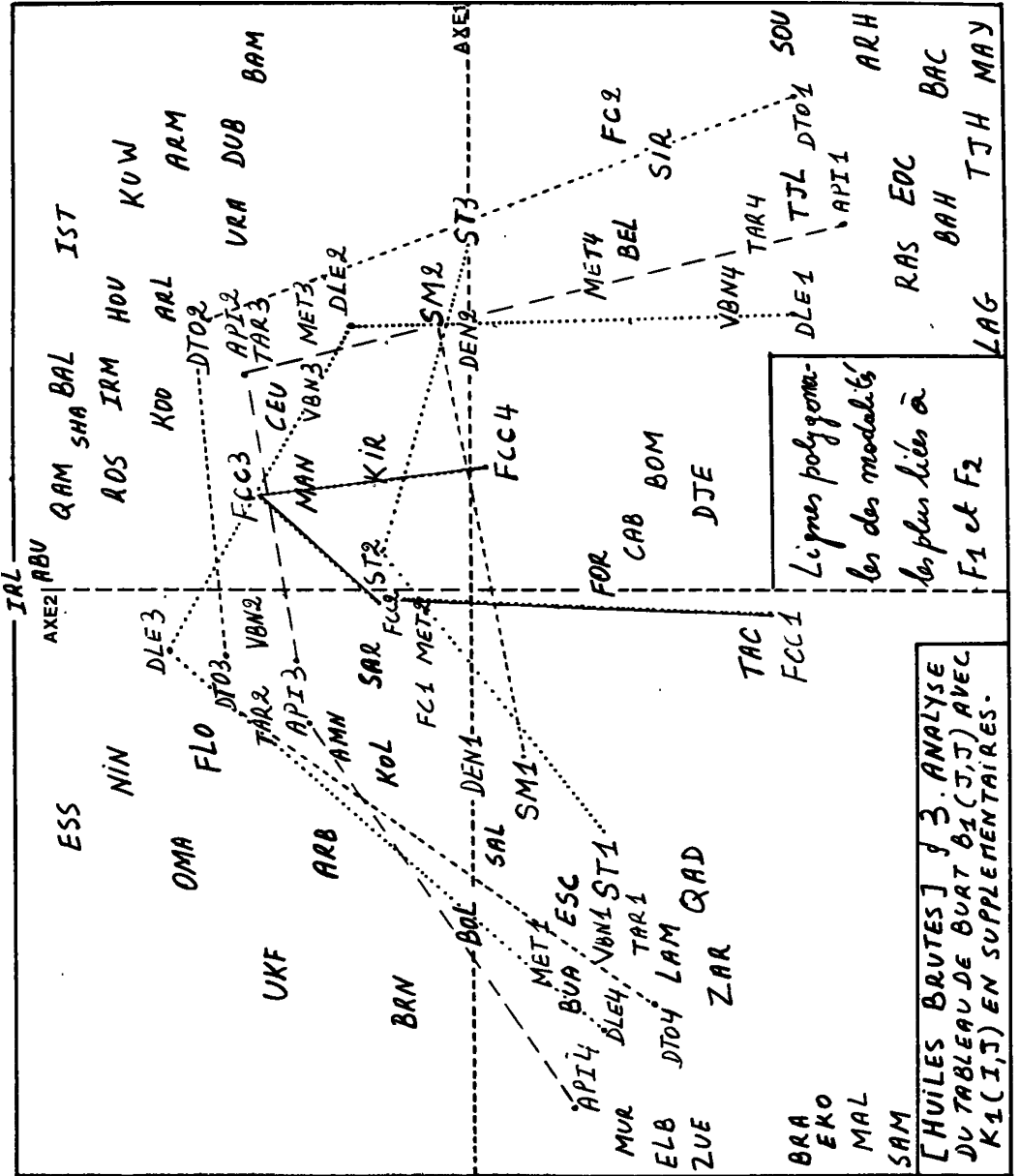
Dans cette analyse, il est inutile de parler de l'axe 2 car avec l'effet Guttman dans le plan 1×2 , il est difficile de donner un sens clair à cet axe. Il est également inutile de parler de l'axe 3 à cause du fait que ce sont les mêmes modalités ($F_3 > 0$ et $F_3 < 0$) qui contribuent à sa création : DLE2 ($F_3 < 0$), DLE3 ($F_3 > 0$), DTO2 ($F_3 < 0$), DTO3 ($F_3 > 0$), etc. . On réserve le grand commentaire au plan 1×2 .

Plan 1×2 : $\tau_1 + \tau_2 = 69\%$.

Le plan 1×2 explique 69% de l'inertie totale du nuage et permet de mieux représenter les huiles et leurs caractéristiques :

Dans le quadrant (F_1^+ , F_2^+) apparaissent les huiles ayant un pourcentage de soufre élevé dans l'échantillon obtenu après distillation à 370° C (SM2), une densité supérieure à 0,93 (DEN2, API2), contenant du nickel et du vanadium en quantité non négligeable (MET3). Ces caractéristiques sont associées aux huiles suivantes : Kirkuch Blend de l'Irak (KIR), Mandji du Gabon (MAN), Ceuta du Vénézuéla (CU), Basrah Medium de l'Irak (BAM), Dubaï de Dubaï (DUB), Kuwait du Koweït (KUW), Abu Koosh d'Abu Dhabi (KOO), Arabian Light de l'Arabie Saoudite (ARL). Il s'agit donc dans l'ensemble des huiles moyennes et ayant un pourcentage de soufre supérieur à 0,7%.

Le quadrant (F_1^- , F_2^+) est celui des huiles relativement légères, ayant une densité inférieure à 0,92 (DEN1, API3), possédant peu de nickel et de vanadium (MET2), ayant un pourcentage en poids des produits volatils obtenus après distillation à 180° C assez important, entre 19,4 et 25,4 % (DEL3). Ces huiles sont : Bonny Light du Nigéria (BOL), Brent Blend (BRN) du Royaume-Uni, Arabian Beni (ARB) de l'Arabie

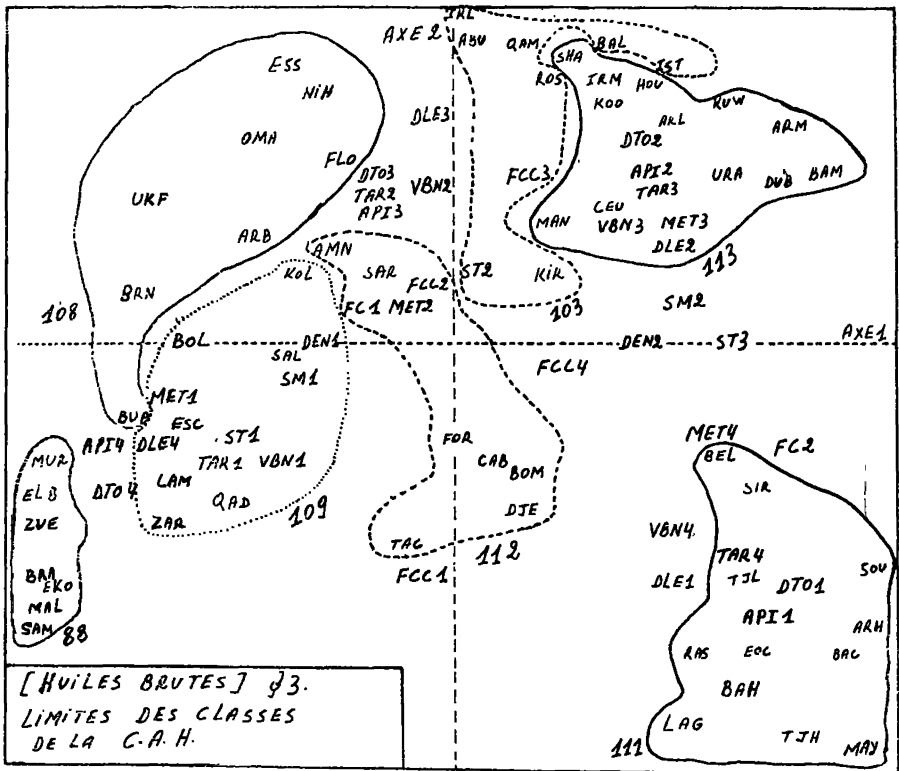


Saoudite, Sarir de la Libye (SAR), Oman de l'Oman (OMA), Kole du Cameroun (KOL) et Forties de l'UF (UKF).

Dans le quadrant (F_1^- , F_2^-) se trouve dans l'ensemble les huiles de meilleures qualités, très légères (DEN1, API4) et contenant peu de soufre (SM1, ST1); ces caractéristiques sont en association avec les huiles suivantes : Abuttifel de la Libye (ABU), Escravos du Nigéria (ESC), Lamar du Vénézuéla (LAM), Zarzaitine de l'Algérie (ZAR), Salawuti de l'Indonésie (SAL), Qatar Ducan de Qatar (QAD), Zueitina de la Libye (ZUE). Ce sont des huiles légères, de bonne qualité, possédant un pourcentage des produits volatils obtenus après distillation à 180° C (DLE4) et à 370° C (DT04) élevé.

Le quadrant (F_1^+ , F_2^-) est celui des huiles lourdes (DEN2, API1), ayant assez de soufre dans les résidus à courte chaîne (ST3). Ce sont les huiles suivantes : Lagunillas du Vénézuéla (LAG), Maya du Mexique (MAY), Tia-Juana H du Vénézuéla (TJH), Basrah Heavy de l'Irak (BAH), Ras Gharib de l'Egypte (RAS), Eocenie de la zone neutre (EOC), Djeno M du Congo (DJE), Bachaguero du Vénézuéla (BAC).

La classification automatique qui fait l'objet du § suivant (§ 4) confirme ces résultats.



4 Classification ascendante hiérarchique sur le tableau $K_1(I, J)$

Dans l'analyse du tableau de Burt $B_1(J, J)$, nous avons projeté en éléments supplémentaires les lignes du tableau disjonctif $K_1(I, J)$.

Nous nous proposons d'effectuer une classification hiérarchique sur les 60 huiles représentant les lignes du tableau disjonctif afin de confirmer les résultats de l'analyse factorielle. L'aide à l'interprétation de la classification a été faite en utilisant le programme VACOR. Nous nous sommes borné à expliquer les dichotomies afférentes aux 8 noeuds les plus hauts.

4.1 Noeud (119): $N(119) = N(118) \cup N(111)$; $V(N) = 0,364$. Ce noeud est le centre de gravité du nuage $N(I)$ pour l'ensemble des huiles brutes. Il se scinde en les classes (118) et (111). Les modalités API1 (densités plus élevées) et DTO1 (faible % des produits volatils à 370° C) sont en parfaite association avec la classe (111) qui contient les huiles lourdes : Ras Gharib de l'Égypte (RAS), Souedie de la Syrie (SOU), Tia Juana H du Vénézuéla (TJH), Maya du Mexique (MAY), Languillias du Vénézuéla (LAG), Basrah Heavy de l'Irak (BAH), Eocene de la zone neutre (EOC), Sirip de l'Iran (SIR), Belaym Blend de l'Égypte (BEL), Tia-Juana L du Vénézuéla (TJL), Arabian Heavy de l'Arabie Saoudite (ARH) et Bachaquero du Vénézuéla (BAC).

La classe (118) est constituée de toutes les autres huiles moyennes et légères. On retrouve cette opposition entre les classes (118) et (111) sur le premier axe factoriel.

4.2 Noeud (118): $N(118) = N(117) \cup N(114)$; $V(N) = 0,312$. La classe (114) est en parfaite association avec les huiles légères (API4++) ayant un pourcentage des produits volatils obtenu à 180° C (DLE4++) et à 370° C (DT04+++). Ce sont les huiles suivantes : Nigerian Brass du Nigéria (BRA), Ekofisk de la Norvège (EKO), Malossa de l'Italie (MAL), Sahara Mélange de l'Algérie (SAH), Zueitina de la Libye (ZUE), Murban de l'Abu Dhabi (MUR), El Borna de la Tunisie (ELB), Salawati de l'Indonésie (SAL), Kole du Cameroun (KOL), Bonny Light du Nigéria (BOL), Escravos du Nigéria (ESC), Qatar Ducan du Qatar (QAD), Zarzaitine de l'Algérie (ZAR) et Lamar du Vénézuéla (LAM). La classe (117) est associée aux modalités DTO2+, DTO3+, API2+ à cause de la présence des huiles moyennes : Abu Albu Koosk de l'Abu Dhabi (KOO), Kuwait du Koweït (KUW).

4.3 Noeud (117): $N(117) = N(116) \cup N(113)$; $V(N) = 0,233$. En ce troisième niveau de séparation, la classe (113) est en parfaite association avec les modalités qui caractérisent les huiles moyennes : API2+++, DTO2+++, DLE++, à cause de la présence des huiles suivantes : Arabian Light de l'Arabie Saoudite (ARL), Ural de l'URSS (URA), Arabian Medium de l'Arabie Saoudite (ARM), Kuwait du Koweït (KUW), Hout de la zone neutre (HOU). La classe (116) va avec les huiles brutes dont le pourcentage en poids des produits volatils obtenus par le pic à 370° C est entre 56 et 61,8% (DT03+) à cause de la présence des huiles suivantes : Oman de l'Oman (OMA), Flotta du Royaume-Uni (FLO), ES Siden de la Libye (ESS) Ninian Blend du Royaume-Uni (NIN).

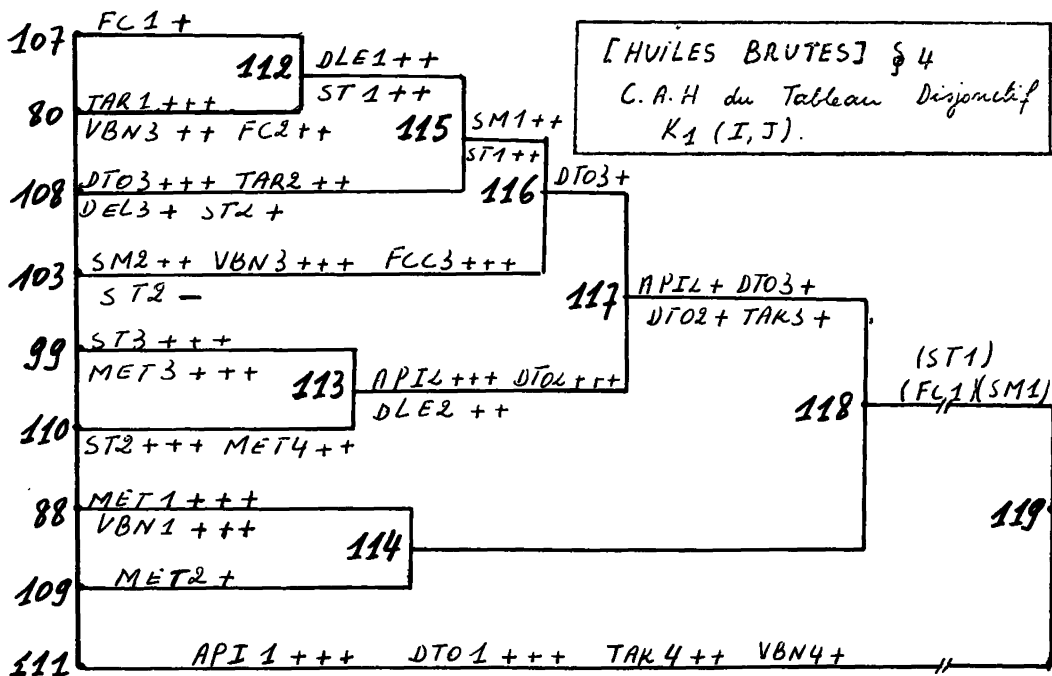
4.4 Noeud (116): $N(116) = N(115) \cup N(103)$; $V(N) = 0,141$. Au quatrième niveau de séparation, on retrouve l'opposition entre les huiles peu soufrées avec un teneur en soufre obtenue après distillation à 180° et 370° C inférieure à 0,6% (SM1) et les huiles plus soufrées : entre 0,7 et 2,1% (SM2). En effet la classe (103) est en parfaite association avec SM2++, FCC3+++, VBN3+++ à cause de la présence dans cette classe des huiles suivantes : Isthmus du Mexique (IST), Qatar marine de Qatar (QAM), Abu Dhabi de l'Abu Dhabi (ABU), Rostam Blend de l'Iran (ROS), Kirkuk Blend de l'Irak (KIR), Iranian Light de l'Iran (IRL).

La classe (115) est en bonne association avec SMI++, FCC4+, ST1++ qui se justifie par la présence des huiles peu soufrées : Forties de l'UF (UKF), ES Sider de la Libye (ESS), Brent Blend (BRN) du Royaume-Uni, Abuttifel de la Libye (BUA), Sarir de la Libye (SAR), TA-Ching de la Chine (TAC).

4.5 Noeud [115] : $N(115) = N(112) \cup N(108)$; $V(N) = 0,104$. Le pourcentage des produits volatils obtenus à 180° C compris entre 19,4 et 25,4% (DLE3+), à 370° C compris entre 5% et 61,8% (DIO3+++), et la teneur en soufre dans le résidu à courte chaîne entre 1,4 et 3,9% sont en association avec la classe (108) qui se justifie par la présence des huiles suivantes : Oman de l'Oman (OMA), Flotta du Royaume-Uni (FLO), Ninian Blend du Royaume-Uni (NIN), Arabian Berri de l'Arabie Saoudite (ARB), Abuttifel de la Libye (ABU), ES Sider de la Libye (ESS).

La classe (112) va avec les huiles ayant un faible pourcentage des produits volatils obtenus à 180° C. (DLE1++) et ayant la faible teneur en soufre dans les résidus à courte chaîne (ST1++) à cause de la présence des huiles suivantes : TA-Ching de la Chine (TAC), Cabinda de l'Angola (CAB), Djeno M du Congo (DJE), Amna de la Libye (AMN), Sarir de la Libye (SAR), Bonny M du Nigéria (BOM), Nigerian Forcados du Nigéria (FOR).

4.6 Noeud [114] : $N(114) = N(88) \cup N(109)$; $V(N) = 0,072$. A ce noeud s'opposent les huiles ayant une teneur en nickel et vanadium dans les résidus à courte chaîne faible (MET1) et moyenne (MET2). En effet la classe (109) est associée à MET2+ à cause de la présence des huiles Bonny Light du Nigéria (BOL), Escravos du Nigéria (ESC), Lamar du Vénézuéla (LAM), Qatar Ducan de Qatar (QAD), Kole du Cameroun (KOL) et Salawati de l'Indonésie (SAL). Par contre, la classe (88) est en parfaite association avec MET1+++ et VBN+++ car elle contient : Zuetino de la Libye (ZUE), Sahara mélange de l'Algérie (SAH), Murban de l'Abu Dhabi (MUR), Malossa de l'ITALIE (MAL) et Nigéria Brass du Nigéria (BRA).



4.7 Noeud [113] : $N(113) = N(99) \cup N(110)$; $V(N) = 0,068$. On retrouve l'opposition entre la classe (99) associée à MET3++ et ST3+++ à cause des huiles suivantes : Arabian Light de l'Arabie Saoudite (ARL), Kuweit du Koweit (KUW), Abu Albu Koosk de l'Abu Dhabi (KOO), hout de la zone neutre (HOU) ; et la classe (110) associée à ST2+++ et MET4++ à cause de la présence dans cette classe des huiles Shoab Aly de l'Egypte (SHA), Iranian M de l'Iran (IRM), Ceuta du Vénézuéla (CEU), Dubaï de Dubaï (DUB).

4.8 Noeud [112] : $N(112) = N(107) \cup N(80)$; $V(N) = 0,064$. Au huitième niveau, la classe (80) est en association avec les huiles ayant un faible pourcentage en résidus à courte chaîne d'hydrocarbures CO (TAR1) et dont l'indice de viscosité est non négligeable (VBN3, FC2) à cause de la présence des huiles Bonny M du Nigéria (BOM) et Nigerian Forcados du Nigéria (FOR). Par contre la classe (107) avec les huiles ayant l'indice de viscosité faible (FC1) à cause de la présence des huiles Amna de la Libye (AMN), Sarir de la Libye (SAR).

Nous arrêtons l'interprétation de la C.A.H. après le 8-ème noeud car les indices de niveau de la hiérarchie deviennent très faibles.

4.9 Description des 9 classes de la partition de l'ensemble I :

107 : TAC : TA-Ching (Chine), CAB (Cabinda (Angola), DJE : Djeno M (Congo), AMN : Amna (Libye), SAR : Sarir (Libye).

80 : BOA : Bonny M (Nigeria), FOR : Nigéria Forcados (Nigéria).

108 : OMA : Oman (Oman), FLO : Flotta (U.K), NIN : Ninian Blend (U.K), ARB : Arabian Berri (Arabie Saoudite), BUA : Abultifel (Libye), BRN : Brent Blend (U.K), ESS : ES Siser (Libye), UKF : Forties (U.F).

103 : IST : Isthimus (Mexico), ROS : Rostam Blend (Iran), BAL : Basrah Light (Irak), ABU : Abu Dhabi (Dhabu Dhabi), QAM : Qatar marine (Qatar), IRL : Iranian Light (Iran), KIR : Kirkuk Blend (Irak).

99 : ARL : Arabian Light (Arabie Saoudite), BAM : Basrah medium (Irak), ARM : Arabia medium (Arabie Saoudite), KUW : Kuweit (Koweit), HOU : hout (zone neutre), KOO : Abu Albu Koosk (Abu Dhabi).

110 : SHA : Shaab Aly (Egypte), IRM : Iranian M. (Iran), CEU : Ceuta (Vénézuéla), MAN : Mandji (Gabon), DUB : Dubaï (Dubaï), URA : Ural (URSS).

88 : BRA : Nigerian Bross (Nigéria), EKO : Ekofisk (Norvège), MAL : Malossa (Italie), SAH : Sahara mélange (Algérie), ZUE : Zueitina (Libye), MUR : Murban (Abu Dhabi), ELB : El Borma (Tunisie).

109 : Salarbati (Indonésie), KOL : Kole (Cameroun), BOL : Bonny Light (Nigéria), ESC : Escravos (Nigéria), QAD : Qatar Ducan (Qatar), ZAR : Zarzaitina (Algérie), LAM : Lamar (Vénézuéla).

111 : RAS : Ras Garib (Egypte) BAH : Basrah Heavy (Irak), EOC : Eocene (zone neutre), SIR : Sirip (Iran), BEL : Belaym Blend (Egypte), SOU : Souedie (Syrie), TJL : Tia-Juana L (Vénézuéla), LAG : lagunillas (Vénézuéla), MAY : Maya (Mexico), ARH : Arabian Heavy (Arabie Saoudite), BAC : Bachaquero (Vénézuéla), TJH : Tia-Juana H (Vénézuéla).

5 Analyse du tableau $B_2(J,J)$ avec $K_2(I,J)$ en supplémentaires

Nous soumettons à l'analyse factorielle le tableau de Burt $B_2(J,J)$ avec le tableau $K_2(I,J)$ obtenu après pondération des lignes par l'abondance des gisements (a(i)) en supplémentaires.

5.1 Le premier facteur : Comme l'expérience nous permettait de l'entendre, le premier facteur est un facteur de niveau général opposant les huiles moins denses aux huiles plus denses, moins soufrées aux huiles plus soufrées. Ce facteur explique 41,6% de l'inertie totale du nuage (moins que pour l'analyse sans pondération des lignes par l'abondance des gisements 47,6%).

Les modalités qui contribuent à la création du premier axe factoriel sont les mêmes que dans l'analyse du tableau $B_1(J,J)$ (cf. § 3).

Du côté négatif : DLE4, DTO4, SM1, DEN1, TAR1, MET1, API4. Il s'agit des modalités qui caractérisent les huiles relativement légères, contenant peu de soufre.

Du côté positif : SM2, DEN2. Les modalités qui vont avec les huiles plus lourdes et plus soufrées. La pondération des lignes par l'abondance des gisements n'a rien changé à la constitution du premier axe factoriel.

5.2 Deuxième facteur : L'étude de ce facteur qui explique 31% montre qu'il est dominé comme dans le cas de l'analyse du tableau $B_1(J,J)$ (§ 3) par les modalités suivantes :

Du côté négatif : DLE1, DTO1, TAR4, VBN4, API1, FC2. La seule différence qu'on note ici est l'opposition des huiles ayant un indice de viscosité très important pour les résidus à courte chaîne (VBN4) à la création du deuxième axe. Cette modalité est en association avec les huiles lourdes : Bashah Heavy de l'Irak (BAH), Arabian Heavy de l'Arabie Saoudite (ARH), Tia-Juana H du Vénézuéla (TJH).

Du côté positif : On retrouve DTO2 comme dans l'analyse du § 3.

Nous soumettons au lecteur le plan 1×2 (72% de l'inertie totale) qui offre la meilleure représentation des associations entre modalités et huiles brutes et qui bien que ne différant pas beaucoup du plan 1×2 de l'analyse sans pondération (§ 3), présente des variations pour les huiles moyennes (FOR, CAB, DJE, etc.).

La classification automatique sur le tableau $K_2(I,J)$ pondéré confirme les résultats obtenus par l'analyse factorielle.

5.2 Description des 8 classes de la partition de l'ensemble I

112 : ABU, QAM, ESS, OMA, FLO, NIN, IRL, ROS, IST, KIK, BAL.

3 : ARB.

105 : AMN, BRN, UKF, SAR, BOM, FOR, CAB, DJE, TAC.

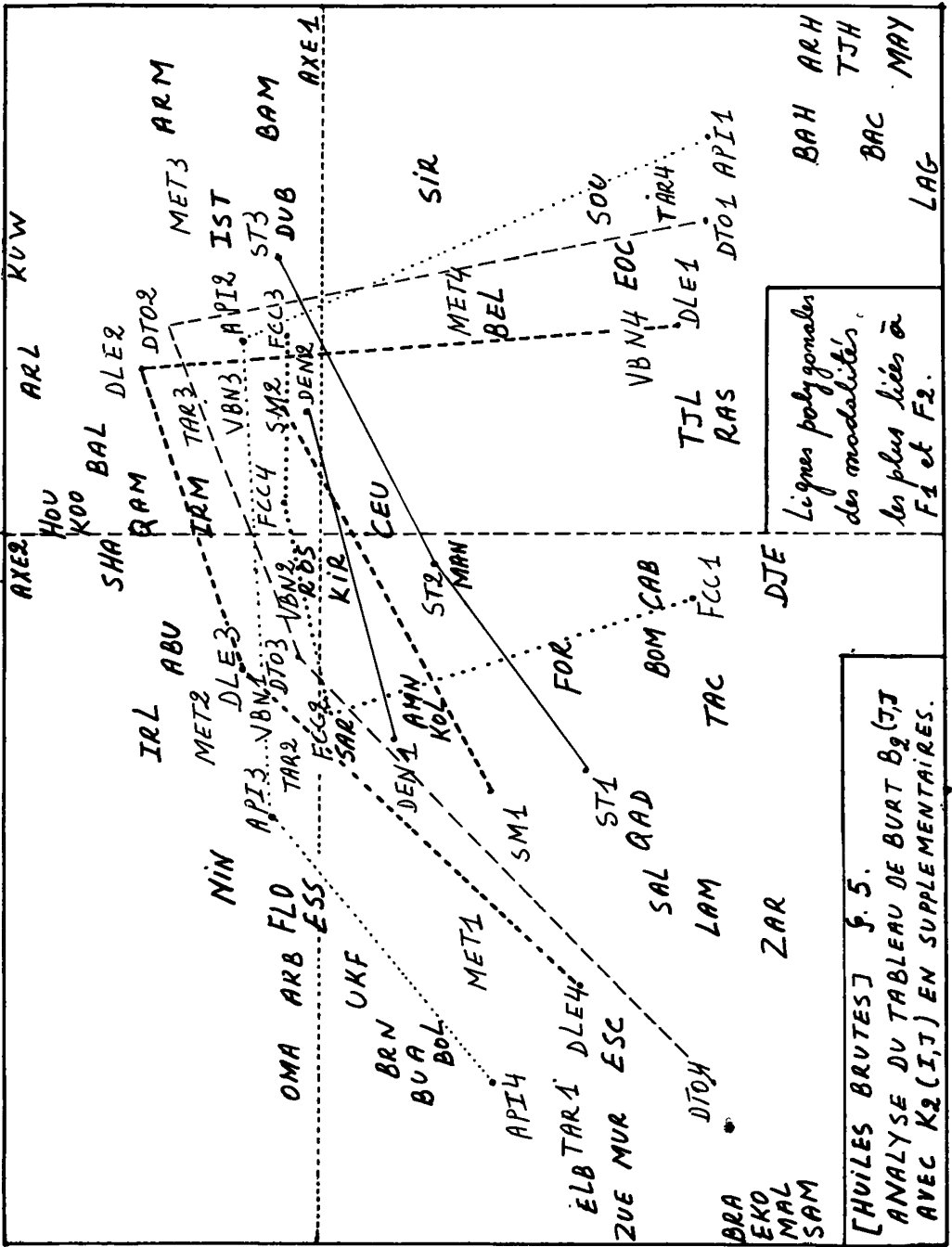
96 : IRM, CEU, MAN, HOU, SHA.

75 : ARL, BAM.

109 : KUW, ARM, KOO, DUB, URA.

110 : KOL, SAL, MUR, BOL, ESC, QAD, ZAR, LAM, BRA, EKO, MAL, SAM, BUA, ELB, ZUE.

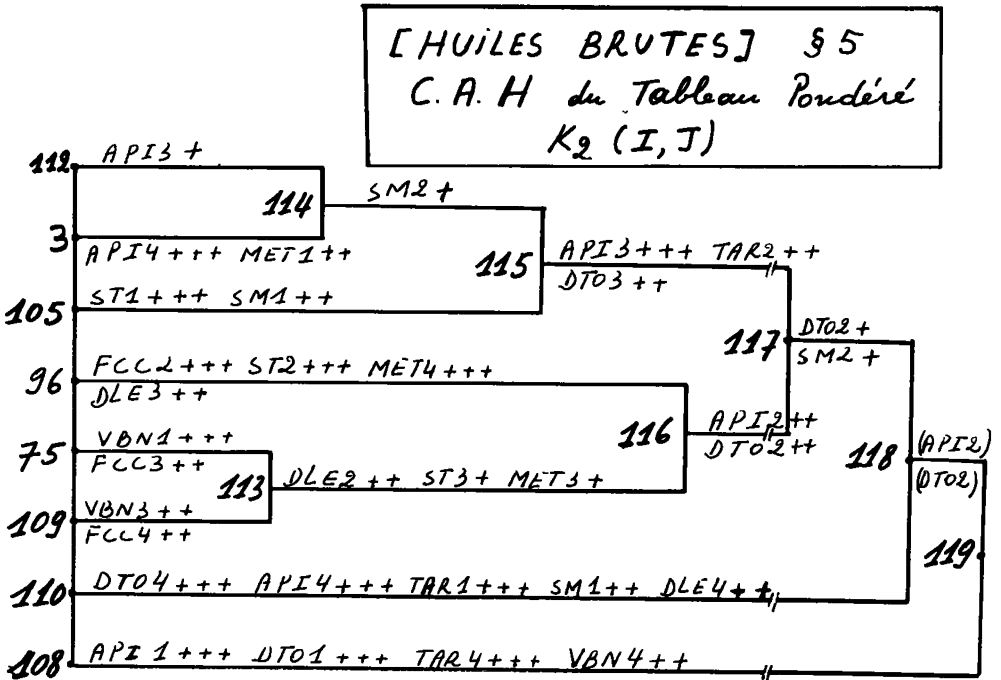
108 : ARH, BEL, SOU, BAC, TJH, TJL, BAH, RAS, MAY, EOC, LAG, SIR.



Lignes polygonales
des modalités
les plus liées à
F1 et F2.

[HUILES BRUTES] §. 5.
ANALYSE DU TABLEAU DE BURT B2 (TJ)
AVEC K2 (I,J) EN SUPPLEMENTAIRES.

BAR ARH
EKO TJH
MAL BAC
SAM MAY
LAG



6 Discrimination barycentrique

On se propose de séparer deux sous-ensembles d'huiles brutes : moins soufrées ISM1 et plus soufrées ISM2. Pour cela on considère le tableau disjonctif $K_1(I, J)$, 60×37 , l'ensemble I ou ensemble des huiles brutes est partagé en deux classes d'intersection vide, ISM1 et ISM2. La première correspond à la modalité SM1 (% de soufre compris entre 0 et 0,6), la deuxième correspond à la modalité SM2 (% de soufre compris entre 0,7 et 2,1%).

A partir du tableau $K_1(I, J)$, on construit un tableau à deux lignes par cumul et équipondération : $K(C, J)$ avec $C = \{N, P\}$

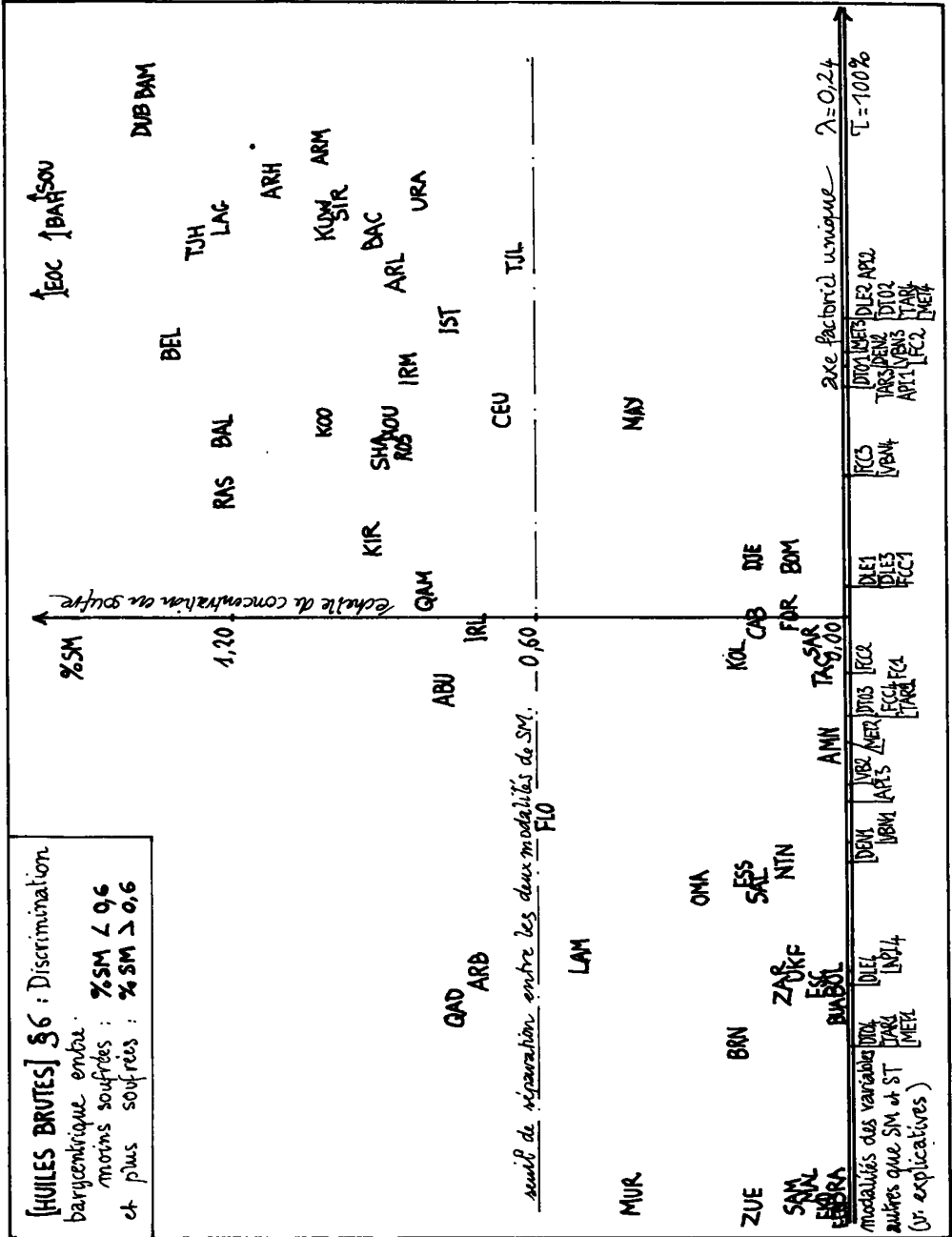
$$K(N, j) = \sum \{K(i, j) \mid i \in \text{ISM1}\}$$

$$K(P, j) = \sum \{K(i, j) \mid i \in \text{ISM2}\}$$

Seules sont retenues comme variables explicatives de la discrimination les 9 variables dont la définition ne fait pas intervenir explicitement le soufre : en d'autres termes on élimine les variables SM et ST.

Le tableau principal 2×32 a donc deux lignes et 22 colonnes ; on adjoint à ce tableau les 60 lignes du tableau $K_1(I, J)$ en éléments supplémentaires pour les projeter sur l'axe factoriel.

La représentation unidimensionnelle ainsi obtenue aboutit à une bonne discrimination entre huiles moins soufrées associées à ISM1 et plus soufrées associées à ISM2.



On en jugera d'après la figure où le facteur (unique) est en abscisse et la teneur en soufre, SM, est l'ordonnée. On a tracé la ligne $SM = 0,6$, qui sépare les deux modalités SM1 (inf.) et SM2 (sup.). La plupart des huiles ayant un facteur $F > 0$ rentrent dans la modalité SM2 (% $SM > 0,6$) tandis que $F < 0$ implique généralement la modalité SM1 (% $SM < 0,6$). Il y a sur 60 huiles analysées 3 exceptions nettes MAY ($F > 0$ et SM1) et QAD et ARB ($F < 0$ et SM2). D'autre part la place de quelques points au voisinage de $F = 0$ ne permet pas de discriminer valablement entre SM1 et SM2 ; ce qui correspond à 4 erreurs (ABU, IRL, DJE, BOM).

Quant au nuage J des modalités des variables explicatives, on constate que API, MET, TAR, DTO contribuent particulièrement à expliquer la teneur en soufre : les huiles très soufrées ont un faible indice de gravité API, (ce qui signifie qu'elles sont lourdes) une forte teneur MET en nickel et vanadium des résidus à courte chaîne ; un fort pourcentage DTO de produits volatils obtenus à 370° C.

On conclura donc que l'ensemble des 9 variables retenues pour la description des huiles et non explicitement liées au soufre, apportent en fait une information importante relativement au soufre ; permettant de discriminer entre huiles très soufrées et peu soufrées dans plus de 80% des cas.

Conclusion

Les analyses factorielles, les classifications ascendantes hiérarchiques et la discrimination barycentrique effectuées dans cette étude ont permis de classer les huiles brutes de pétrole en fonction de leurs caractéristiques et rationaliser les différents aspects du problème taxinomique.

La réserve totale de brut pour chaque gisement qui a été introduite n'a pas beaucoup changé les résultats des analyses.

REFERENCES :

- 1 . M. MBUYI : "Contribution à l'étude par analyse factorielle et classification automatique de la production du pétrole, cartographie européenne et africaine de la consommation des produits raffinés (Thèse de 3° cycle, Paris 1983).
- 2 . M. MBUYI : "[PETRO MONDE GUTTMAN DOUBLE]" ; in C.A.D. n° 1 ; 1984.
- 3 . "Les Cahiers de l'Analyse des Données". Revue trimestrielle DUNOD 1976-1983 (publication éditée avec le concours du C.N.R.S.).