

CLASIFICACION DE LOS PETROLEOS EN EL YACIMIENTO PINA MEDIANTE TÉCNICAS DE LA ESTADÍSTICA MULTIVARIADA.

Olga Castro Castiñeira, Olga Pascual Fernández, Luisa Díaz Armas, Gloria Llabre Raurell

Centro de Investigaciones del Petróleo. Washington # 169 esq., a Churruca, Cerro, C. Habana

RESUMO

A partir da necessidade de se estudar a compartimentação vertical e horizontal do petróleo acumulado no campo de Pina, localizado no lado norte da parte central de Cuba, decidiu-se utilizar como uma ferramenta adicional a técnica de impressão cromatográfica (fingerprinting) da fração C_{15} .

Dos cromatogramas obtidos foram selecionadas determinadas relações de compostos que pudessem refletir, em gráficos dos tipos polar ou diagrama de estrelas, as semelhanças e/ou diferenças entre as amostras e, tornando-se assim possível estudá-las.

Outras técnicas de estatística multivariada, como por exemplo as medidas de agrupamento, ajudaram na correlação entre os grupos de óleos com composições similares.

Com o objetivo comparar geoquimicamente os óleos da área de Pina, este trabalho focaliza a utilização da Análise de Componentes Principais para o reconhecimento das relações entre os compostos e a suas respectivas classificações hierárquicas ascendentes. Para tanto foi necessário realizar, preliminarmente, um estudo profundo dos dados primários. A classificação dos óleos por este método confirma os resultados obtidos com os diagramas de estrelas, onde foram reconhecidas três famílias de óleos com origens similares, porém com níveis variáveis de biodegradação e fracionamento evaporativo.

ABSTRACT

This study focus on the oil field of Pina, discovered in 1990, and that it is located in the northern part of Central Cuba. The reservoirs of this accumulation are represented by the association of volcanic rocks and sediments that are structured by a SSW-NNE fault system. Such discontinuities provide various production compartments.

The vertical and lateral distributions of the oil in the Pina oil field were studied using the chromatographic pattern of the C_{15} fraction. The respective variations were based on selected ratios that reflected the variability in composition and that were easily recognized in statistical plots, such as star diagrams or dendrograms. The chromatographic results were corroborated by the use of PCA (Principal Component Analysis) that gave a hierarchical relationship among the parameters.

Considering that the sources of the studied oils is believed to be similar and that was identified three groups of oils, it is likely that biodegradation and evaporative fractionation were responsible for such dissimilarity. The results also have shown that the oil variability is not controlled by its geographical distribution (compartments) neither by the lithology of the correspondent reservoir.

Introducción

El yacimiento (campo) Pina fue descubierto en el año 1990, en la zona central de Cuba y abarca una extensión de 8 km² comenzando su explotación en el propio año (Figura 1). Está asociado a una secuencia vulcanógeno-sedimentaria, en una forma estructural positiva que conforma varios pliegues pequeños separados por fallas inversas y fallas SSO-NNE, las cuales los dividen en bloques. Pueden caracterizarse varios tipos de rocas que forman cuatro paquetes (capas u horizontes productivos): Conglomerados, Tobas Finas, Tobas Finas y Efusivos.

Es un yacimiento fracturado, con algunas fracturas cerradas, con porosidades entre 13 y 21 %, y permeabilidades muy bajas, de 3.4 a 14 mD por la vertical, y de 0.7 a 10.6 mD horizontalmente. La declinación experimentada ha planteado la necesidad de implementar métodos de recuperación mejorada.

Los petróleos de este yacimiento se caracterizan por presentar entre 13 y 35° API, altos contenidos de azufre, entre 1,37 y 2,79 %, contenidos de compuestos saturados entre 34,29 y 42,42 %.

El estudio cromatográfico de la fracción saturada de estos petróleos y de la fracción C₁₅ puso de manifiesto la existencia de fenómenos de biodegradación, lavado por agua y fraccionamiento evaporativo, con distintos niveles de desarrollo. Esto permitió establecer tres grupos:

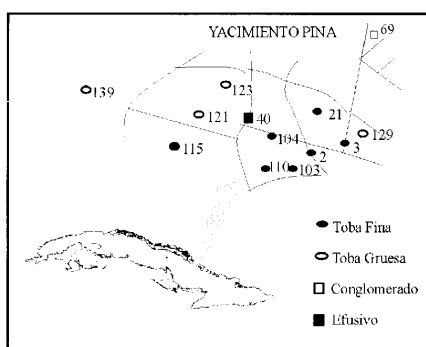


Figura 1. Yacimiento Pina ubicado en la Cuenca Central, de Cuba, dividido por bloques, con los pozos estudiados. La simbología representa los tipos de rocas atravesadas en los pozos.

1) P-103, P-104 y P-110 con mayor nivel de biodegradación, lavado por agua y fraccionamiento evaporativo;

2) P-2, P-69 y P-40 con características intermedias;

3) El resto de los petróleos, menos afectados por los fenómenos antes relacionados.

En los petróleos, tanto el estudio de las características físico – químicas como de biomarcadores ponen de manifiesto un origen marino, carbonático y anóxico.

Un factor importante a tener en cuenta en la aplicación de métodos de recuperación mejorada es la continuidad hidrodinámica del yacimiento. Para evaluar este aspecto, se estudió la composición de hidrocarburos, utilizando la huella cromatográfica de la fracción C₁₅, en un cromatógrafo de gas Chrompack CP-9001, con columna capilar DB-1 de 30 m x 0.25 mm de diámetro y 0.25 mm de espesor. Las condiciones del ensayo fueron: temperatura del inyector 300°C, temperatura del detector 320°C, y el programa de la columna 30°C durante 5 min, 30-50°C a 20°C/min, 50 a 160°C a 2°C/min, 160-310°C a 20°C/min, y 5 min a 310°C. La data fue analizada automáticamente, empleando el software Maestro II, versión 2.4 de Chrompack International B.V. 1995.

Se escogieron 17 relaciones de picos. Los compuestos se identificaron mediante la data de retención por comparación con patrones puros. Las diferencias promedios obtenidas para estas relaciones en análisis duplicado fue de 2 a 5 %, en correspondencia con lo reportado en la literatura. Mayores diferencias fueron obtenidas al comparar petróleos de diferentes pozos, los que superan el 10 %. Esto indica diferencias en la composición hidrocarbonada de las muestras, y pone de manifiesto la pobre comunicación vertical y horizontal en el yacimiento.

La confiabilidad de los resultados obtenidos se evidencia por la repetibilidad de los valores de las relaciones de picos, algunos de los cuales son mostrados en la Tabla 1:

Este trabajo permitió establecer la pobre comunicación vertical (entre horizontes productivos) y horizontalmente en el yacimiento (entre bloques).

Tabla 1. Repetibilidad de las medidas de alturas de pico.

Run No.	19/20	35/34	42/41	42/45	61/69
1	0.721	0.751	0.939	0.802	0.500
2	0.715	0.730	0.932	0.800	0.518
3	0.704	0.739	0.940	0.809	0.521
4	0.685	0.750	0.935	0.812	0.502
X	0.706	0.742	0.936	0.806	0.510
R%	4.473	2.813	1.043	1.602	3.048

Como una vía para corroborar los resultados obtenidos por el método anteriormente planteado, se empleó la técnica del análisis de cluster muy utilizada en las correlaciones de petróleos. El análisis de cluster agrupa a los petróleos basado en las similitudes entre los mismos (Kaufman, 1987 y 1990). Usualmente se usa la distancia euclidiana como medida de las similitudes entre muestras en un espacio n-dimensional (donde n es el número de relaciones usadas). El resultado de este análisis es un gráfico, conocido comúnmente como dendrograma, que ilustra gráficamente los grupos formados. Este análisis es particularmente útil, cuando el número de petróleos y/o relaciones es grande. Las variables utilizadas en el análisis fueron relaciones de compuestos integradas por parafinas normales e isoparafinas, cicloparafinas normales y ramificadas, y compuestos aromáticos.

Datos de observación

Se utilizaron datos que corresponden a 17 relaciones entre diferentes picos tomados de los análisis cromatográficos para 13 pozos del área. Los elementos de los picos son los que aparecen a continuación, y las relaciones utilizadas se identificarán en adelante por la que tienen sus números de orden (código):

Estos datos se agrupan en una matriz donde la primera columna se refiere al pozo (identificador), la segunda al paquete litológico (variable nominal ilustrativa) y la tercera al bloque (idem). El resto corresponde a las relaciones mencionadas, siendo magnitudes adimensionales (variables continuas). (Ver Tabla 2)

Donde:	1:	3-Metilpentano
	2:	n-Hexano
	3:	Metilciclohexano
	4:	Benceno+Ciclohexano
	5:	2-Metilhexano
	6:	3-Metilhexano
	7:	1,tri-dimetilciclohexano+3-Etilpentano
	8:	n-Heptano
	9:	Metilciclohexano
	10:	2,4-Dimetilhexano
	11:	Tolueno
	12:	2-Metilheptano
	13:	1-Tri-2-Dimetilciclohexano
	14:	n-Octano
	15:	Propilciclohexano
	16:	2,5-Dimetilheptano
	17:	(meta-para) Xileno
	18:	3-Metiloctano
	19:	n-Nonano
	20:	Isopropilbenceno
	21:	1,2,4-Trimetilbenceno
	22:	n-Decano
	23:	2,6-Dimetilnonano
	24:	1,4-Dimetil2-etilbenceno+3-Metildecano
	25:	n-Undecano
	26:	3-Metilundecano
	27:	n-Dodecano
	28:	2,6-Dimetilundecano

Lectura preliminar y análisis de la Tabla de Datos

En este tipo de análisis se trata de destacar a simple vista si hay diferencias entre las variables para los distintos pozos, buscándose alguna forma de contrastar los datos para que resalte el problema que se examina. Para ello se señalaron los valores máximo y mínimo en la Tabla 2 para cada variable (columna) y se estableció la cantidad de máximos y mínimos existentes por individuo (pozo-fila).

A simple vista se observa:

- Una concentración de valores máximos en los pozos: Pina 103, 104 y 110, donde las variables con los máximos no se repiten de un pozo a otro.

- Una concentración de valores mínimos en los pozos: Pina 3 y Pina 139, que tampoco se repiten de un pozo a otro.

- Un grupo de pozos donde no son notables los máximos y mínimos.

Todo esto conlleva a la identificación de tres grupos de pozos con características diferentes en sus muestras de petróleo.

Tabla 2. Datos de las relaciones entre los picos de los análisis de cromatografía en los pozos del Yacimiento.

POZO	PAQUETE	BLOQUE	(1/2)	(3/2)	(4/5)	(5/6)	(7/8)	(10/9)
P-21	1	5	0.5209	0.3562	0.7272	1.0226	0.1864	0.1434
P-2	1	7	0.7436	0.4352	1.0316	1.0326	0.3769	0.1416
P-3	1	6	0.4681	0.5191	0.9176	1.1227	0.2032	0.1500
P-103	1	7	0.0111	2.2455	1.2159	0.8248	6.3834	0.1595
P-104	1	3	2.6421	1.1631	1.3990	0.6854	0.0000	7.1194
P-110	1	3	9.2168	5.0251	1.2917	0.7280	26.9494	0.1643
P-115	1	2	0.7625	0.3094	0.6779	1.0684	0.2766	0.1607
P-121	2	1	0.6150	0.3211	0.7119	1.3262	0.2638	0.1943
P-123	2	1	0.5292	0.2676	0.6504	1.2791	0.2307	0.1950
P-129	2	6	0.4796	0.2741	0.7135	1.1889	0.1715	0.1820
P-139	2	9	0.5471	0.3051	0.6208	0.2423	0.2252	0.1896
P-69	3	8	0.7399	0.4000	0.6169	1.1387	0.4943	0.2144
P-40	4	4	0.7444	0.5991	0.8560	1.6958	0.4746	0.1457
Media			1.386	0.94	0.8793	1.0273	2.79	0.705

Tabla 2. Continuación

POZO	PAQUETE	BLOQUE	(11/9)	(12/11)	(13/14)	(16/15)	(18/17)	(20/19)
P-21	1	5	0.4765	1.1273	0.1740	0.3974	0.6996	0.0671
P-2	1	7	0.6379	0.6789	0.3601	0.4790	0.5931	0.2315
P-3	1	6	0.5817	0.0000	0.0000	0.0000	0.3802	0.5609
P-103	1	7	0.4401	0.3288	1.4890	0.3895	0.3887	0.5789
P-104	1	3	0.5309	0.3537	1.4827	0.4215	0.4554	1.6885
P-110	1	3	0.3948	0.3184	1.5486	0.3581	0.4296	0.5171
P-115	1	2	0.5814	0.8681	0.2432	0.3963	0.6424	0.0680
P-121	2	1	0.4921	0.8502	0.2307	0.4133	0.6774	0.0969
P-123	2	1	0.4788	0.9949	0.2087	0.3950	0.7034	0.0933
P-129	2	6	0.5286	1.1296	0.1468	0.3469	0.6945	0.0605
P-139	2	9	0.4641	0.3866	0.1927	0.8149	0.6927	0.0687
P-69	3	8	0.4437	1.0617	0.4677	0.5441	1.1891	0.3963
P-40	4	4	0.4400	0.5271	0.4636	0.4277	0.4285	0.1750
Media			0.4993	0.663	0.539	0.4141	0.6134	0.354

Tabla 2. Continuación

POZO	PAQUETE	BLOQUE	(21/22)	(23/22)	(24/25)	(26/27)	(28/27)	
P-21	1	5	0.2073	0.4106	0.1808	0.4265	0.4602	Máximo
P-2	1	7	0.4320	0.8090	0.7460	0.6188	0.7534	
P-3	1	6	0.1081	0.3471	0.6748	0.4622	0.4592	
P-103	1	7	4.2589	7.3542	1.2495	1.6337	2.5158	Mínimo
P-104	1	3	1.5314	2.9979	1.4309	1.1321	1.6130	
P-110	1	3	5.8242	10.3899	1.1310	0.9660	0.6228	
P-115	1	2	0.3333	0.4106	0.4250	0.3362	0.5878	
P-121	2	1	0.2346	0.4318	0.3954	0.3477	0.5043	
P-123	2	1	0.2359	0.4312	0.1962	0.4476	0.4633	
P-129	2	6	0.4464	0.7351	0.2821	0.4578	0.4393	
P-139	2	9	0.3977	0.8415	0.4180	1.6170	0.4124	
P-69	3	8	0.3448	0.8834	1.0567	0.7723	0.9514	
P-40	4	4	0.3615	1.4558	0.5146	0.3809	0.5238	
Media			1.132	2.115	0.669	0.738	0.793	

A partir de los datos anteriores, se pudo obtener la siguiente información resumida en la Tabla 3

Como una comprobación del análisis anterior, se trabajó con los valores de las medianas de las variables (M). Esta medida de tendencia central, por ser robusta, no está influida por los valores extremos como ocurre con la media. A partir de la construcción de histogramas del tipo steam and leaf, se preparó la codificación de la

Tabla 4 y se separaron las relaciones que resultaron mayores y menores que la mediana, y la cantidad de estos valores por pozo. La información obtenida puede resumirse como sigue:

Examinando la información a simple vista, se destaca:

Tabla 3. Cantidad de valores máximos y mínimos por pozo y relaciones que los presentan.

POZO	PAQUETE	BLOQUE	CANTIDAD DE VALORES MÁXIMOS	RELACIONES QUE PRESENTAN VALORES MÁXIMOS	CANTIDAD DE VALORES MÍNIMOS	RELACIONES QUE PRESENTAN VALORES MÍNIMOS
P-21	Toba Fina	V	-	-	1	24/25
P-2	Toba Fina	?	1	11/9	1	10/9
P-3	Toba Fina	VI	-	-	6	12/11 13/14 16/15 18/17 21/22 23/22
P-103	Toba Fina	?	2	26/27 28/27	1	1/2
P-104	Toba Fina	III	4	4/5 10/9 20/19 24/25	1	7/8
P-110	Toba Fina	III	6	1/2 3/2 7/8 13/14 21/22 23/22	1	11/9
P-115	Toba Fina	II	-	-	1	26/27
P-121	Toba Gruesa	I	-	-	-	-
P-123	Toba Gruesa	I	-	-	1	3/2
P-129	Toba Gruesa	VI	1	12/11	1	20/19
P-139	Toba Gruesa	IX	1	16/15	2	5/6 28/27
P-69	Glomerado	VIII	1	18/17	1	4/5
P-40	Efusivo	IV	1	5/6	-	-

-Los pozos Pina 103, 104 y 110 con valores superiores a la mediana, y que coinciden en una serie de relaciones (13/14, 20/19, 21/22, 23/22, 24/25 y 26/27); le sigue el Pina 69 (20/19 y 24/25).

-Con valores inferiores a la mediana, para estos casos: Pina 3 y 139.

-El resto de los pozos presenta características bastante homogéneas.

Puede apreciarse entonces, que en este análisis también aparecen tres grupos de pozos conformados casi por los mismos individuos, y además, con cierta regularidad en el

comportamiento de las variables. Nótese la frecuencia de aparición que tienen en orden las relaciones 20/19 (8), 24/25 (6), 13/14 (4), 21/22 (4), y 23/22 (4).

Selección del método para el análisis multivariado de las variables

Ya con el análisis precedente hay elementos para definir el tipo de análisis multivariado que se puede utilizar para obtener el resultado que se busca. No obstante, cualquier tipo de análisis que se escoja, lo primero a considerar es que el número de pozos resulta insuficiente para

Tabla 4. Codificación de acuerdo a los valores excedentarios, deficitarios o similares a la mediana de cada variable o relación.

Pozo	Paq.	Bl.	1/2	3/2	4/5	5/6	7/8	10/9	11/9	12/11	13/14	16/15	18/17	20/19	21/22	23/22	24/25	26/27	28/27
21	T.F.	V																	
2	T.F.	?																	
3	T.F.	VI								\$---	\$---	\$---		++	--	-			
103	T.F.	?	\$---	++++			++++\$				++++			++	++++\$	++++\$	+	++	+++
104	T.F.	III	+++	+			\$---				++++			++++\$	+++	++	+	+	++
110	T.F.	III	++++\$	++++\$			++++\$				++++			+	++++\$	++++\$	+	+	
115	T.F.	II												-					
121	T.G.	I																	
123	T.G.	I																	
129	T.G.	VI												-					
139	T.G.	IX					---							+					
69	C.	VIII												+				+	
40	E.	IV																	
	M		0.615	0.400	0.727	1.068	0.260	0.164	0.479	0.679	0.243	0.397	0.624	0.175	0.362	0.809	0.515	0.462	0.524

Simbología: M = Mediana;

2-3M = + 1/2-1/3M = -
 3-4M = ++ 1/3-1/4M = --
 4-5M = +++ 1/4-1/5M = ---
 5-6M = ++++ 1/5-1/6M = ----
 6-7M = +++++ 1/6-1/7M = -----
 >7M = +++++\$ <1/7M = \$-----

Tabla 5. Cantidad de valores mayores y menores que la mediana y relaciones que los presentan.

POZO	PAQUETE	BLOQUE	CANTIDAD DE VALORES > M	RELACIONES QUE PRESENTAN VALORES > M	CANTIDAD DE VALORES < M	RELACIONES QUE PRESENTAN VALORES < M
P-21	Toba Fina	V	-	-	1	24/25
P-2	Toba Fina	?	-	-	-	-
P-3	Toba Fina	VI	1	20/19	5	12/11 13/14 16/15 21/22 23/22
P-103	Toba Fina	?	9	3/2 7/8 13/14 20/19 21/22 23/22 24/25 26/27 28/27	1	1/2
P-104	Toba Fina	III	9	1/2 3/2 7/8 13/14 20/19 21/22 23/22 24/25 26/27 28/27	1	7/8
P-110	Toba Fina	III	9	1/2 3/2 7/8 13/14 20/19 21/22 23/22 24/25 26/27	-	-
P-115	Toba Fina	II	-	-	1	20/19
P-121	Toba Gruesa	I	-	-	-	-
P-123	Toba Gruesa	I	-	-	1	24/25
P-129	Toba Gruesa	VI	-	-	1	20/19
P-139	Toba Gruesa	IX	1	16/15	2	5/6 20/19
P-69	Conglomerado	VIII	2	20/19 24/25	-	-
P-40	Efúsivo	IV	-	-	-	-

representar a cada paquete o bloque, existiendo desproporción en los primeros y excesiva división en los segundos, mientras que el número de variables supera en un tercio al de individuos o pozos.

La estrategia clásica consiste en el Análisis de Componentes Principales (ACP), clasificación jerárquica ascendente y partición de las muestras a partir de sus coordenadas factoriales; caracterización de las clases de la partición y confrontación de la tipología de las muestras con la repartición de las mismas (pозos) por paquetes y/o bloques. El procesamiento se realizó mediante el programa SPAD3.5, que fue el software estadístico utilizado como ayuda en este trabajo.

El ACP consiste en la descomposición factorial de la inercia de la matriz de correlaciones entre las variables activas, conceptualmente homogéneas, pero de todas formas centradas y reducidas (Crivisqui y Droesbeke, 1993; Morineau, 1984). Se obtuvo la matriz de correlación entre las variables, donde aunque no se establece una buena correlación lineal entre todas las variables, hay un grupo de ellas bastante representativo en que si parece existir.

a) Análisis de Componentes Principales

Al procesar los datos, se obtuvo como resultado que el primer plano factorial (formado por el primer y el segundo eje factorial) absorbe el 66 % de la inercia total, mientras que en los tres primeros ejes se concentra el 79 %. El resto de los ejes, hasta 13, contribuyen muy poco, por lo que solamente se toman en cuenta para los posteriores análisis los tres primeros ejes.

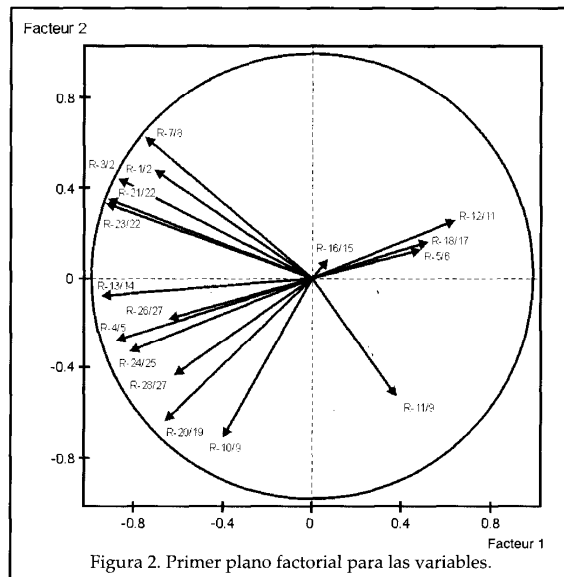


Figura 2. Primer plano factorial para las variables.

En la representación gráfica del primer plano factorial de las variables (Figura 2), se aprecia la correlación entre ellas por el ángulo agudo

que forman entre sí los vectores que las representan, además de la longitud de su norma. Las variables que tienen la representación más mala en el primer plano factorial dado por sus coordenadas y correlaciones con los ejes, son:

Variable	Coordenadas	
	1	2
5/6	0.49	0.13
11/9	0.39	-0.54
16/15	0.06	0.07
18/17	0.52	0.15

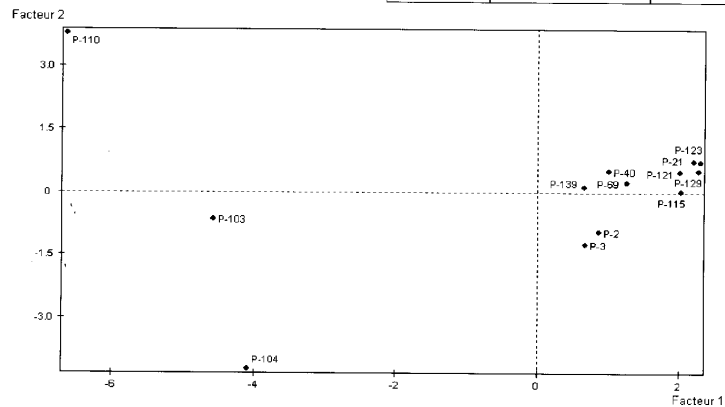


Figura 3. Representación del primer plano de los individuos o pozos formado por el primero y segundo ejes factoriales.

Nótese que algunas de estas variables tienen poca incidencia o no están representadas en las tablas de contraste de las medianas.

En el primer plano de los individuos (pozos), los menos contributivos a la inercia proyectada en este plano, y que están peor representados son (Figura 3):

Individuos	Distancia	Coordenadas		Contribución		Cosenos Cuadrados	
		1	2	1	2	1	2
Pina 139	19.55	0.68	0.12	0.4	0.0	0.02	0.00
Pina 69	13.15	1.27	0.24	1.5	0.2	0.12	0.00
Pina 3	15.61	0.68	-1.25	0.4	4.3	0.03	0.10

Es decir, los pozos Pina 139, 69 y 3 no son contributivos a la inercia proyectada sobre el primer plano dado por los cosenos cuadrados, que tienen valores mayores en el segundo y tercer ejes factoriales

y por tanto sus contribuciones son mayores en otros planos. Estos individuos difieren mucho del resto de la población. Muy contributivos resultan:

Individuos	Distancia	Coordenadas		Contribución		Cosenos Cuadrados	
		1	2	1	2	1	2
Pina 110	60.39	-6.65	3.75	40.3	38.3	0.73	0.23
Pina 104	38.02	-4.09	-4.22	15.3	48.5	0.44	0.47
Pina 103	30.89	-4.58	-0.66	19.1	1.2	0.68	0.01

Estos individuos también difieren del resto. Si se analiza la representación gráfica del primer, segundo y tercer planos (Figuras 3, 4 y 5), se verá la posición que mantienen uno con respecto a otro los pozos Pina 103 y 104, y Pina 69 y 139, así como la unión entre los pozos Pina 21, 115, 121, 123 y 129. Esta última relación se aprecia muy bien.

Clasificación Jerárquica Ascendente.

Esta clasificación de los individuos se hace

por el método de clúster utilizando las distancias euclidianas (Aldenderfer y Blashfield, 1987). Se diferencia de un método simple de clúster en que las agrupaciones no se hacen con las variables directamente, previa selección de las mismas, sino a partir de los ejes factoriales escogidos por ACP, donde en cada eje hay una contribución mayor o menor de cada variable como se examinó anteriormente. Como se analizó anteriormente, se tomaron los 3 primeros ejes por contener el mayor por ciento de la inercia.

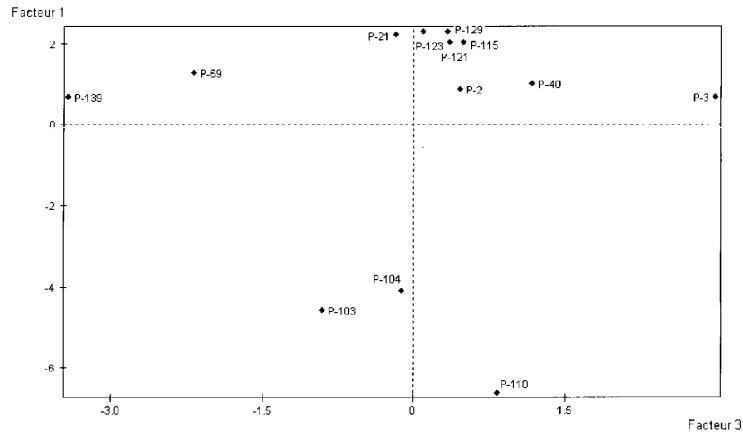


Figura 4. Representación del segundo plano de los individuos o pozos formado por el primer y el tercer ejes factoriales.

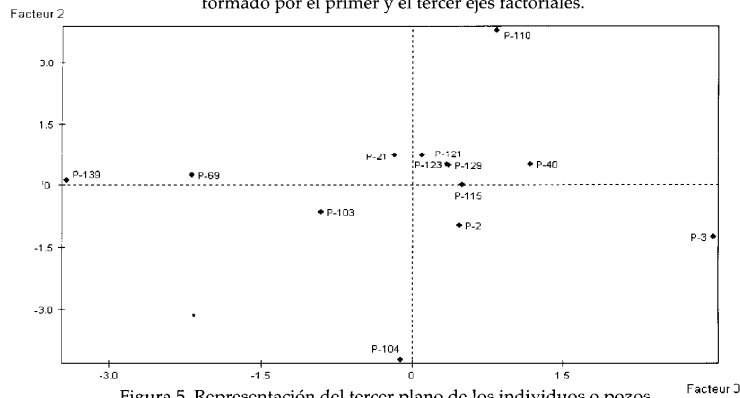


Figura 5. Representación del tercer plano de los individuos o pozos formado por el segundo y el tercer ejes factoriales.

En este caso (Figura 6), se aprecia el agrupamiento ascendente de los pozos Pina 21 y 129; 121 y 123; éstos con el 115, formando claramente el grupo mencionado en el ACP y que se destacaba por determinadas características de contraste de los datos primarios. Otra agrupación evidente es la de los pozos Pina 103 y 104 y de éstos con el 110, pero este grupo no tiene a sus individuos en el grado de cercanía de la agrupación anterior, pero si difiere bastante para diferenciarse de la primera. El resto de los pozos, en orden de su parecido o diferencia se van incorporando al primer grupo, pudiendo considerarse los más cercanos el 40 y luego el 2, por último el 69, el 3 y el 139, como los más separados, que se comportaban como algo diferente en el análisis de ACP. El trazado de la línea para las particiones puede hacerse a voluntad.

a) Partición de la muestra.

Se realizó la partición en 2 y 4 grupos, pero este fue un trabajo puramente formal, ya que la

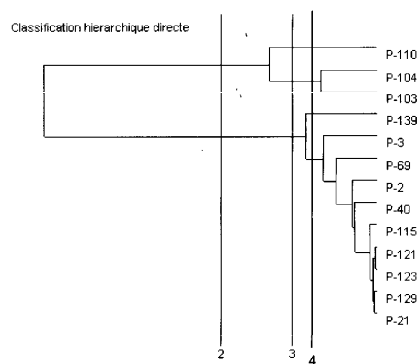


Figura 6. Clasificación jerárquica ascendente de los individuos después del Análisis de Componentes Principales.

estructura de la muestra quedó bien establecida en los pasos anteriores, donde se vio como se forman las agrupaciones. Solamente puede destacarse en este paso las variables características para cada grupo, la mayoría de las cuales ya se habían detallado en la lectura preliminar y el análisis de la tabla de datos. Estas se expresan en la Tabla 6.

Obsérvese la representación de las variables características, coincidentes en todas las agrupaciones, además de su repetitividad y coincidencia con las que mejor se definen en los datos primarios y juegan un papel importante en los ejes factoriales del ACP.

Tabla 6. Partición en clases, pozos representados en cada una y variables características.

Partición	Clase	Pozos	Variables Características
En 2 Clases	1/2	Pina 121, 115, 123, 129, 21,2,40,69, 3 y 139	20/19, 28/27, 3/2, 24/25, 21/22, 23/22, 4/5 y 13/14
	2/2	Pina 103, 104 y 110	13/14, 4/5, 23/22, 21/22, 24/25, 3/2, 28/27 y 20/19
En 4 Clases	1/4	Pina 121, 115, 123, 129, 21,2,40,69 y 3	5/6, 21/22, 23/22, 13/14 y 26/27
	2/4	Pina 139	-
	3/4	Pina 104 y 103	28/27, 20/19, 13/14, 24/25, 4/5 y 10/9
	4/4	Pina 110	-

Conclusiones

1. El análisis de datos primarios destaca el contraste de tres grupos de pozos con características diferentes de sus petróleos, donde hay un grupo de variables que los definen.

2. El ACP define muy bien el grupo de pozos constituido por Pina 21, 115, 121, 123 y 129, con características comunes en sus petróleos así como la diferenciación del resto, de los petróleos en los pozos Pina 103, 104 y 110. Se diferencian por otra parte, Pina 69 y 139.

3. Con el ACP se pudieron establecer las variables más contributivas a los ejes factoriales principales.

4. Mediante las técnicas de clasificación por cluster se confirmó la agrupación realizada por el ACP.

5. En todo el análisis no se obtuvo asociación entre el tipo de petróleo y los bloques donde se ubican los pozos; tampoco se encontró relación directa con el paquete litológico.

6. Estos resultados concuerdan con los obtenidos por los métodos de la huella de petróleo (fingerprinting), donde el ploteo polar (di-

agrama de estrella) de las relaciones utilizadas separa dos grupos principales: Pina 115, 121, 123, 129, 21 y 3 de Pina 103, 104 y 110. Existe un grupo integrado por Pina 2, 40, 69 y 139 con características intermedias entre los dos primeros.

7. Estudios paralelos realizados sobre estos petróleos indican similar origen, pero han puesto de manifiesto la existencia de fenómenos secundarios como biodegradación y fraccionamiento evaporativo, provocando cambios más o menos intensos en la composición de los mismos, en dependencia del nivel alcanzado por estos fenómenos.

Bibliografía

1. Kaufman, R. L.; Ahmed, A. S.; Hemkins, W. B. "A new technique for the analysis of commingled oils and its application to production allocation calculations". Proceeding Indonesian Petroleum Association. Sixteen Annual Convention, October 1987.

2. Kaufman, R. L.; Ahmed, A. S. "Gas chromatography as development and production tool for finger printing oils from individual reservoirs: Applications in the Gulf of México".

Gessepm Foundation Ninth Annual Research. Conference Proceeding, October 1990.

3. Crivisqui, E.; Droesbeke, J. J. "Le concept mathématique de distance et la comparaison des éléments d'un tableau de données". Cultures et Sociétés, (3), pp.17-32, 1993.

4. Morineau, A.; "Note sur la

caractérisation statistique d'une classe et les valeurs-tests". Bulletin Technique du Centre de Statistique et d'Informatique Appliquées, 2, pp. 20-27, 1984.

5. Aldenderfer, M. S; Blashfield, R. K. "Cluster Analysis". California, Sage Publications, Inc. 1987.