

## Caracterización de alimentos combinando reglas de decisión relacionales y lineales. Una aplicación al aceite de oliva virgen de Málaga

Por R. Aparicio\*, L. Ferreiro y J. L. Rodríguez

Instituto de la Grasa y sus Derivados.  
Avda. Padre García Tejero, 4. Sevilla.

### RESUMEN

**Caracterización de alimentos combinando reglas de decisión relacionales y lineales. Una aplicación al aceite de oliva virgen de Málaga.**

Los componentes químicos de un alimento cuantificados por cromatografía gaseosa, pueden ser usados para construir reglas de decisión relacionales y lineales que, combinadas mediante algoritmos matemáticos, podrían permitir su caracterización.

Este trabajo describe los métodos aplicados para diseñar y evaluar las reglas, que constituyen la base de conocimientos de un sistema experto deductivo, y el procedimiento usado para determinar la similitud entre una muestra desconocida y el alimento que definen las reglas.

Para explicar y desarrollar la metodología se estudia la caracterización del aceite de oliva virgen de la provincia de Málaga. Cincuenta y cuatro parámetros químicos pertenecientes a cinco series, ácidos grasos, metil-esteroles, esteroides, alcoholes alifáticos y triterpénicos e hidrocarburos alifáticos y terpénicos, más un grupo constituido por el fitol, el eritrodol y un hidroxialdehído triterpénico, se identificaron y midieron en 172 muestras de Andalucía. Los cromatogramas recuperados por ordenador permitieron construir 17 reglas entre relacionales y lineales. Las reglas relacionales (10) muestran mediante expresiones lingüísticas –mayor, menor, igual, etc.– las interrelaciones existentes entre uno o más componentes químicos de la misma o diferente serie. Las reglas lineales (7), una por cada una de las series indicadas más una usando simultáneamente todos los compuestos, se obtuvieron mediante la utilización de programas estadísticos discriminantes. Finalmente, para verificar el proceso descrito se usaron veintiseis muestras, de diferentes zonas, provincias o países.

**PALABRAS-CLAVE:** *Aceite de oliva virgen – Alimento (caracterización) – Málaga – Regla de decisión lineal – Regla de decisión relacional.*

### SUMMARY

**Characterization of foods merging relational and lineal decision rules. An application to virgin olive oil of Malaga.**

The chemical compounds of a food, quantified by GC, can be used to build relational and lineal decision rules, which merged by mathematical algorithms, could allow its characterization.

The paper describes the methods applied to design and evaluate those rules, which are kept in the knowledge base of a deductive expert system, and the procedure applied to know the similarity between an unidentified sample and the food defined by the rules.

The characterization of virgin olive oil of Malaga province has been used to explain the methodology. Thus, 54 chemical compounds

belonging to five different series, fatty acids, methylsterols, sterols, aliphatic and triterpénic alcohols, and aliphatic and terpénic hydrocarbons, besides a group composed by a hydroxi-aldehid-triterpénic, phytol, and erytrodol, were identified and quantified in each of the 172 samples collected over Andalusia. Their chromatograms, which were redrawn by computers, allowed to build seventeen rules among lineal and relational ones. These relational rules (10) show linguistic interrelations, such less, great, equal, etc., among two or more chemical compounds of the same or different series. On the other hand, statistical programs of discriminant analysis allowed to get the lineal rules (7) whose coefficients were taken from each one of the series, plus a rule where they were reached in the whole set of chemical compounds.

Finally, twenty-six samples collected in different zones, provinces or countries, or in the same area but from different crops, were used to measure the performance of the methodology.

**KEY-WORDS:** *Food (characterization) - Lineal decision rule - Malaga - Relational decision rule - Virgin olive oil.*

### 1. INTRODUCCION

En trabajos previos se ha descrito la ventaja y los inconvenientes de la caracterización de los alimentos mediante los sistemas expertos, frente a los métodos quimiométricos (1) (2).

Recientemente, se han venido aplicando sistemas expertos inductivos (3) en la identificación de aceites (4) u otros alimentos (5). Se trata de sistemas desarrollados a partir de una estructura arborescente (6) que utilizan una técnica de subdivisiones sucesivas de los valores de una variable para intentar la mejor separación entre dos y sólo dos posibles grupos. Es un método aceptable si los grupos son muy dispares, pero posiblemente no muy recomendable si éstos son algo similares o existen más de dos posibles conjuntos. Por otra parte, la dependencia directa del método con los datos que se analizan podría disminuir las posibilidades finales de los resultados, si éstos varían sustancialmente con los años. En cualquier caso, se trata de sistemas de los que existen un gran número de alternativas (6), que son fáciles de usar, transportables y útiles en análisis discriminante de productos dispares que tienen pocas variaciones con el tiempo.

Sin embargo, los autores optaron, desde el principio, por un sistema deductivo (1) que permite trabajar con más de dos conjuntos simultáneamente, sin más límite que la velocidad de la respuesta del sistema, y con una mayor precisión en sus resultados. Sin embargo, su base de conocimientos sólo es aplicable en el área para la que fue diseñado, en nuestro caso aceite de oliva virgen.

En el trabajo, se analiza la forma de construir la base de conocimientos del sistema experto, constituida por reglas de decisión. Las reglas que se analizan son de dos tipos, lineales, obtenidas a través de procedimientos estadísticos, en este caso SLDA, y relacionales, logradas a partir de la comparación entre patrones.

Esta estructura permite una cierta independencia de los resultados con la posible variabilidad de los datos, ya que se poseen reglas que describen relaciones lógicas entre parámetros químicos (reglas relacionales), poco dependientes de sus variaciones, y no se descarta la información obtenida con las reglas lineales, que son ecuaciones matemáticas que tienen una fuerte dependencia respecto a los datos usados para construirlas (7).

Por otra parte, los algoritmos de la teoría de la evidencia permiten combinar los resultados obtenidos con cada una de las reglas lineales o relacionales, de

forma independiente o combinando los resultados finales de ambas.

Para describir el tipo de diseño y el modo de operación del procedimiento, se analizan los resultados obtenidos al caracterizar los aceites de oliva virgen de variedad hojiblanco de Málaga, que constituyen, aproximadamente, la mitad de la producción total de aceite de esta variedad, segunda de España. En la bibliografía consultada no se describe ningún procedimiento o metodología que permita caracterizar el aceite de esta provincia.

Para el desarrollo del trabajo, se han analizado 172 muestras de aceite de oliva virgen. En cada una de ellas, se han identificado y cuantificado 54 parámetros químicos, pertenecientes a cinco series de compuestos, ácidos grasos, metilesteros, esteroides, alcoholes e hidrocarburos, más un grupo constituido por un hidroxialdehído-triterpénico, el fitol y el eritrodil. Con la información de parámetros químicos, agrupados por series de compuestos, se construyeron los cromatogramas patrones de esta provincia, y del resto de las muestras andaluzas. La comparación entre los patrones permitió obtener las reglas relacionales, mientras que las lineales fueron logradas a través del análisis discriminante (8). El trabajo describe esas reglas matemáticas lineales y relacionales, que han permitido clasificar e identificar esos aceites durante tres cosechas diferentes.

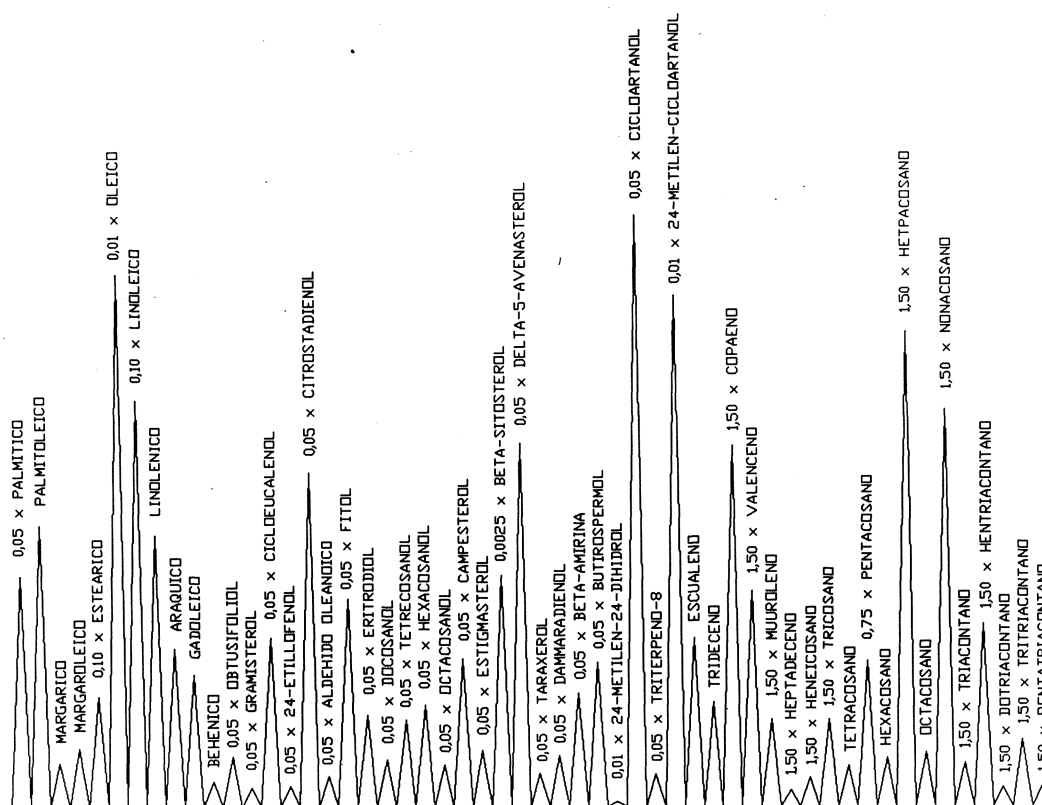


FIG 1 CROMATOGRAMA DE LA COOPERATIVA 'LA PURISIMA' (ARCHIDONA)

El conjunto de ambas reglas fue verificado con 26 muestras de diversas variedades, provincias y países. Se analizan los resultados del proceso de verificación tanto con muestras cuantificadas en iguales condiciones que las usadas con los patrones, como de algunas que fueron cuantificadas con columnas diferentes e incluso no usan todas las reglas almacenadas en la base de conocimientos.

## 2. MATERIALES Y MÉTODOS

### 2.1. Químicos

En la figura 1, se describen, sobre un cromatograma recuperado informáticamente, los nombres de los parámetros químicos que se han usado en el desarrollo del trabajo. Estos pertenecen a los ácidos grasos, alcoholes alifáticos y triterpénicos, esteroides, metil-esteroides, hidrocarburos alifáticos y terpénicos, el fitol, eritrodol y un hidroxialdehído-triterpénico.

La metodología y el instrumental químico para cuantificar los compuestos químicos se describe en (9) (10).

### 2.2. Datos

La base de datos del sistema está constituida por ciento setenta y dos muestras de Andalucía, de las cuales ocho, pertenecientes a la provincia de Málaga, representan completamente el universo a caracterizar. Con la información de esas muestras, se han construido las reglas matemáticas que caracterizan sus aceites.

Para verificar el sistema, se han analizado 18 muestras de las provincias de Málaga, Cáceres, Castellón, Córdoba, Jaén, Lérida, Teruel y Toledo, y 8 de la Liguria, Basilicata y Toscana de Italia, de cosechas que no coinciden con aquella con la que se construyeron las reglas. Algunas de ellas se han verificado con todos los parámetros químicos, mientras que otras sólo lo han sido con una parte.

### 2.3. Software

Se ha utilizado la librería estadística BMDP (11) para los estudios previos y la construcción de reglas lineales.

La base de datos relacional DBIII-Plus (12) para almacenar la información de los cromatogramas y detectar posibles errores. AUTOCAD, para dibujar los cromatogramas informáticos que aparecen en el trabajo (13). SEXIA (1), el sistema experto que usa las reglas construidas y que opera en FranzLisp sobre el sistema operativo Ultrix32.

### 2.4. Hardware

Se han utilizado dos ordenadores, con procesadores 80286 y 80386, para introducir los datos de los cromatogramas en la base de datos relacional y para construir y correr los programas en BMDP, DBIII+ y AUTOCAD, y un ordenador Digital mVAX II, en el que funciona el sistema experto SEXIA.

## 3. RESULTADOS Y DISCUSIÓN

### 3.1. Análisis informático de los cromatogramas

Los grupos de compuestos, cuantificados para el desarrollo del trabajo, han sido obtenidos en diferentes cromatogramas y algunos de ellos, como los n-Alcanoles y el Fitol, se han obtenido mediante la repartición del rascado de la placa en dos fases, que se analizan en dos cromatogramas diferentes (9). Esto implica la imposibilidad real de poder observar el cromatograma representativo de cada muestra, con lo que se imposibilita la aplicación de los métodos de reconocimiento de patrones, que son unos de los más utilizados en Química.

Como paso previo para poder tener patrones que representen a las variedades, zonas, etc., se construyó, por ordenador, la forma básica del cromatograma de cada muestra analizada, aunque sólo de los 54 parámetros químicos que se analizan.

En la figura 1 se representa uno de esos cromatogramas. Cada pico de ese cromatograma computerizado representa el valor de un parámetro químico, multiplicado por un coeficiente. La utilización de coeficientes permite observar simultáneamente todos los compuestos químicos analizados, ya que la cuantificación se ha realizado en mg/100g de aceite (10) y existe una gran diferencia cuantitativa entre algunos compuestos, por ejemplo, entre hidrocarburos y ácidos grasos. Por otra parte, en el cromatograma no se representan los tiempos de retención –el espacio entre picos es siempre igual– para poder tener una visión más concentrada del gráfico.

Con este tipo de cromatogramas computerizados, que podrían ser considerados como "la huella digital del aceite", se han construido los cromatogramas medios de los aceites de la agrupación hojiblanco de Málaga (en adelante hojiMálaga) y de aquellos analizados que no pertenecen a esa agrupación (en adelante, No-hojiMálaga). Cada uno de esos cromatogramas computerizados medios –en adelante patrones– han sido obtenidos con los valores medios de cada uno de los compuestos químicos. La representación espacial de estos dos cromatogramas, agrupados por series de compuestos, aparece en la figura 2 aunque No-hojiMálaga está dividida en las provincias de las que provienen las muestras. Esta representación gráfica permite detectar las características

más notables del tipo de aceite que se estudia. Por otra parte, un simple análisis visual de esos patrones, en relación con el histograma de una muestra desconocida, dará una primera información del grupo con el que posee mayor semejanza.

### 3.2. Reglas de decisión relacional

Con la información inicial que suministran los patrones es posible deducir reglas relacionales. Se consideran como reglas relacionales aquellas que, mediante expresiones lingüísticas, ponen de manifiesto relaciones entre los valores de los diferentes compuestos químicos de uno o varios de los patrones, que, generalmente, no se verifican en otros patrones, de los que se quieren distinguir. Las expresiones lingüísticas de esos operadores relacionales pueden ser mayor, menor, mucho mayor, casi igual, etc..

En este trabajo, las reglas relacionales están circunscritas a los términos y operadores: mayor (>), menor (<) y mucho menor (<<) o mucho mayor (>>), pudiéndose deducir otras mediante los algoritmos descritos en (14). En la figura 1, por ejemplo, se puede

observar que el pico del Butirospermol es mayor que el correspondiente al Hexacosanol. Esta misma relación también se observa en el patrón de hojiMálaga -figura 2-, a la vez que no se verifica en los otros patrones provinciales. Consiguientemente, la relación matemática se puede expresar como la siguiente regla relacional:

'Si el Butirospermol es mayor que el Hexacosanol la muestra es de hojiMálaga'

Sin embargo, esta regla no tiene ningún significado si no es cuantificable con un valor, que indique con qué probabilidad o posibilidad se asegura que es correcta. Aún a pesar de que en la tabla I se puede observar que sólo los aceites de Málaga y Sevilla tienen, de promedio, el valor del butirospermol mayor que hexacosanol.

Tabla I

Valores medios del Obtusifoliol y el Hexacosanol cuantificados con columnas capilares y agrupados por provincias y cosechas.

Provincia cosecha	muestras	valor medio Hexacosanol	valor medio Butirospermol
Málaga 1988	8	5.465	6.072
Sevilla 1988	8	3.551	3.970
Jaén 1988	97	10.881	5.216
Córdoba 1988	59	7.804	4.074
Toscana 1989	9	6.949	3.366
Basilicata 1989	26	5.675	4.034
Liguria 1989	2	4.846	3.737
Liguria 1988	6	4.989	4.539
Huelva 1989	17	5.578	3.626
Castellón 1989	22	8.282	3.823
Toledo 1989	19	5.060	2.961
Badajoz 1989	12	6.606	3.445
Teruel 1989	14	4.837	3.841
Cáceres 1989	7	9.290	3.358
Jaén 1989	55	8.866	3.389
Sevilla 1989	3	5.219	7.749
Málaga 1989	5	6.756	7.616

En estadística, las probabilidades asociadas con cada regla son obtenidas de las probabilidades condicionales y no-condicionales. La probabilidad no-condicional o 'a priori' se escribe, por ejemplo, como

$$P(\text{copaeno-alto}) \tag{3.1}$$

que se puede interpretar como "La probabilidad de que la muestra tenga el copaeno alto", sin considerar ninguna otra característica. Su probabilidad se calcula mediante el cociente entre los elementos del universo que cumplen esa característica (copaeno-alto) respecto a todo el universo estudiado (aceite de Andalucía),

$$P(\text{copaeno-alto}) = \frac{|\text{copaeno-alto}|}{|\text{universo}|} \tag{3.2}$$

donde, |A| son los elementos del conjunto A.

Cuando esa regla se combina con otra que debe cumplirse condicionada a la primera, se denomina probabilidad condicional o 'a posteriori' y se escribe, por ejemplo

$$P(\text{copaeno-alto} | \text{hojiMálaga}) \tag{3.3}$$

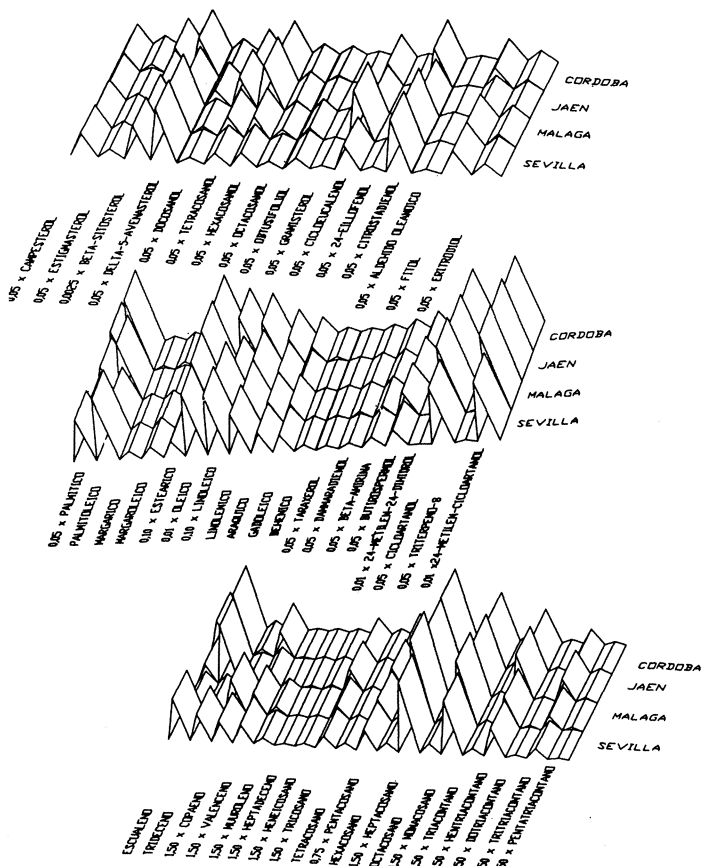


Figura 2 Series de compuestos de provincias.

Tabla II

Esquemas de las probabilidades condicionales calculadas en el trabajo.

Probabilidades condicionales:			
Tipo	Expresión:		
1	$P(E H)$ :	SI se verifica H	ENTONCES se verifica E
2	$P(H E)$ :	SI se verifica E	ENTONCES se verifica H
3	$P(E .not.H)$ :	SI se no verifica H	ENTONCES se verifica E
4	$P(H .not.E)$ :	SI se no verifica E	ENTONCES se verifica H

**Leyenda:**

**H:** Hipótesis a verificar

H1 el aceite es de la variedad hojiblanco

H2 el aceite es de hojiMálaga

**E:** eventos o características que deben cumplirse

**Eventos:** relaciones lingüísticas entre compuestos químicos

que se puede interpretar como "La probabilidad de que el copaeno sea alto cuando la muestra es de hojiMálaga" y se calcula, mediante operaciones entre conjuntos,

$$P(E|H) = \frac{P(\text{copaeno-alto} \mid \text{hojiMálaga})}{P(\text{hojiMálaga})} \quad (3.4)$$

La regla anterior se puede expresar, de forma más genérica, como, 'SI H, ENTONCES E'. La segunda parte se denomina evento, que en nuestro caso puede estar constituido por una o varias características del aceite de oliva, y a la primera se le llama hipótesis.

En este trabajo, se han calculado dos probabilidades 'a priori', las probabilidades de cada hipótesis y de cada evento. La primera a través de la información sobre las producciones; por ejemplo, la probabilidad de la 'Variedad hojiblanco' (hipótesis H1, Tabla II) ha sido calculada dividiendo la producción media de la variedad hojiblanco por la producción total del universo estudiado, aceite de Andalucía. Se trata, pues, de una posibilidad sujeta a variaciones, en función del universo analizado. También puede considerarse como la probabilidad con que acertaríamos que una muestra desconocida es de variedad hojiblanco, sin necesidad de hacer ningún análisis. El valor de esta probabilidad 'a priori' disminuye en la misma medida en que aumente el universo a estudiar. Por ejemplo, la probabilidad de H1 para el universo de Andalucía es mayor que la de H1 para el universo de España.

La segunda probabilidad se calculó dividiendo el número de muestras que verifica el evento por el total del universo. Esta probabilidad también debe actualizarse al cambiar el universo analizado y al final de un proceso de identificación de muestras desconocidas, que es un proceso de autoaprendizaje sólo posible con sistemas expertos. Se trata, como en el caso anterior, de la probabilidad con que acertaríamos las características de uno o algunos de los parámetros químicos de una muestra, sin necesidad de hacer ningún análisis químico. En este caso, si el universo aumenta o disminuye, respecto del inicial, no se puede 'a priori' deducir en qué forma afectará a la probabilidad asociada con el evento.

En relación con las probabilidades condicionales, éstas pueden ser de dos tipos:

- SI ocurre un evento (compuesto por una característica o un grupo de ellas) ENTONCES se verifica una hipótesis, o
- SI se supone una hipótesis ENTONCES se verifica un evento.

En términos de este trabajo podrían expresarse, por ejemplo, como:

$$P(E \mid H) = P(\text{copaeno-alto} \mid \text{hojiMálaga}) =$$

SI la muestra es de hojiMálaga ENTONCES el copaeno es alto. (3.5)

$$P(H \mid E) = P(\text{hojiMálaga} \mid \text{copaeno-alto}) =$$

SI el copaeno es alto ENTONCES la muestra es de hojiMálaga. (3.6)

la primera probabilidad condicional puede ser obtenida siguiendo el proceso indicado anteriormente, mas no así la segunda que se calcula aplicando el Teorema de Bayes:

$$P(H \mid E) = \frac{P(H)P(E \mid H)}{P(E)} \quad (3.7)$$

donde las probabilidades 'a priori' y 'a posteriori' se calculan como se ha indicado anteriormente. Por tanto, esta probabilidad condicional es función del universo analizado, que en el caso que analiza este trabajo es el aceite de Andalucía.

Para construir las reglas relacionales, se ha utilizado la estructura de las reglas del tipo IF-THEN o Causa-Efecto, indicadas en (3.5) y (3.6), con las dos hipótesis que, analiza el trabajo, 'el aceite es de hojiMálaga' y 'el aceite es de variedad hojiblanco', Tabla II. De cada hipótesis se pueden deducir ocho posibles reglas, que se corresponden con las posibles combinaciones de hipótesis y eventos. Sin embargo, es fácil deducir que, al tratarse de la teoría de la probabilidad las ocho combinaciones se reducen a cuatro, siendo las otras cuatro sus complementos a la certeza.

Por otra parte, la Tabla III muestra los eventos más importantes encontrados al analizar los patrones

mostrados en la figura 2. Estos eventos, o sus relaciones lingüísticas entre compuestos químicos, no constituyen la totalidad de las que se pueden construir, aunque posiblemente sí sean los que posean mayor nivel de probabilidad de entre los posibles. En cualquier caso, el sistema almacena toda la información de los cromatogramas para usarlos en la identificación de reglas con nuevas categorías de variedades, comarcas o naciones. Esa información es almacenada en forma matricial para acelerar la construcción de nuevas reglas. La Tabla IV describe, en cada línea, la probabilidad condicional de cada tipo de regla para cada una de las hipótesis y sus eventos.

Tabla III

Tipos de eventos, relacionados con el trabajo, que se consideran más discriminantes.

Eventos	Descripción del evento
E1	Palmitoleico << 10*Margaroleico .and. Palmitoleico < 10*Margárico
E2	Eritrodiol > Docosanol
E3	Cicloeucaleol > Citrostadienol
E4	Trideceno < Copaeno .and. Octacosanol > Tricosano
E5	Trideceno < Copaeno .and. Valenceno > Hexacosano
E6	(Palmitoleico < Linolénico .and. Margaroleico > Trideceno) .or. (1.5*Hexacosano < Muuroleno .and. Muuroleno > 1.5*Tricosano)
E7	Copaeno >> Pentacosano
E8	Cicloeucaleol > Delta_5-Avenasterol
E9	Margaroleico > Tricosano .and. Obtusifoliol > Octacosanol .and. Fitol > Hexacosanol
E10	Fitol > Campesterol .and. Butirospermol > Hexacosanol .and. Obtusifoliol > Octacosanol .and. Fitol > Hexacosanol

Tabla IV

Probabilidad, truncada a números enteros, con cada una de las reglas relacionales construidas.

Tipo reglas	Eventos	Probabilidad Hipótesis							
		H1				H2			
1,2,3,4	E1	48	86	1	1	75	43	5	4
1,2,3,4	E2	80	42	19	4	99	17	24	0
1,2,3,4	E3	24	25	12	13	63	21	12	2
1,2,3,4	E4	52	99	0	8	38	23	4	3
1,2,3,4	E5	56	99	0	7	75	43	5	1
1,2,3,4	E6	24	99	0	5	75	99	0	1
1,2,3,4	E7	52	99	0	8	75	46	4	1
1,2,3,4	E8	20	25	10	13	50	20	10	3
1,2,3,4	E9	48	99	0	8	63	42	4	2
1,2,3,4	E10	40	83	1	9	50	33	5	3

Nombre hipótesis	Significado de la hipótesis
H1	el aceite es de la variedad hojiblanco
H2	el aceite es de hojiMálaga

Así, el tipo 1 (Tablas II, IV) muestra la condición necesaria pero no suficiente que debe cumplir una

hipótesis. Por ejemplo, si la hipótesis fuese 'hojiMálaga' y el evento 'Eritrodiol >Docosanol' y aunque la probabilidad fuese del 100%, pueden existir muestras que cumplan el evento pero que no verifiquen la hipótesis. Sin embargo, siempre que la muestra sea hojiMálaga se verificará el evento. Se observa que el tipo 1 es menos restrictivo para la hipótesis H1 que para la hipótesis H2. Esto es debido a que el grupo de la hojiMálaga está asociado con un área no dispersa de esa provincia y donde la existencia de otras variedades distintas del hojiblanco es menor del 5%. Ese tipo de regla muestra la homogeneidad de la variedad. El proceso se invierte al analizar el tipo 2 (Tablas II, IV), debido a que el conjunto de muestras que verifica la hipótesis H2 es un subconjunto de los elementos de la hipótesis H1. Existiendo muestras de cooperativas próximas a hojiMálaga que poseen similares características químicas, debido a la artificialidad de los límites provinciales en relación con el aceite de oliva. Por ejemplo, algunas muestras de las provincias de Córdoba, Penibética, y Sevilla, Agrupación OleoEstepa.

Sin embargo, la información sobre la variedad hojiblanco es una consecuencia de las reglas sobre hojiMálaga, al ser ésta un subconjunto de la anterior. Por tanto, no tienen esas reglas por qué coincidir con las obtenidas exclusivamente para caracterizar esa variedad, que no se analizan en este trabajo. Por tanto, ni las probabilidades de la Tabla IV ni las reglas de la Tabla III han sido creadas por SEXIA para identificar si una muestra desconocida pertenece a la variedad hojiblanco, únicamente para mostrar la interdependencia entre una variedad y su subconjunto.

El tipo 3 muestra el caso de que no verificándose la hipótesis sí se verifique el evento. Este tipo no es el complementario del tipo 1, por lo que ambos no suman 100; el complemento del tipo 1 es

$$\text{'IF H THEN .NOT. A'} \tag{3.8}$$

y el del tipo 3

$$\text{'IF .NOT. H THEN .NOT. A'} \tag{3.9}$$

Un razonamiento similar se puede expresar sobre los tipos cuarto y segundo.

La información suministrada por los diferentes eventos para cualquiera de los tipos de la Tabla II, permite conocer dos parámetros que dan información sobre la importancia del evento, su nivel de suficiencia (LS) y el de necesidad (LN). Así, para las reglas "SI E, entonces H" (tipo 2, tabla II), el primer parámetro es formulado como:

$$LS=P(H | E) / P(No-H | E) \tag{3.10}$$

mientras que el segundo lo es como:

$$LN=P(H | No_E) / P(No_H | No_E) \tag{3.11}$$

La consideración de suficiencia, en el parámetro LS, se explica dado que un elevado valor de este parámetro implica que la existencia del evento incrementa la probabilidad de verificar la hipótesis. Mientras que un valor pequeño de LN significa que la no existencia del evento disminuye la probabilidad de verificar la hipótesis.

El mayor nivel de suficiencia de acuerdo con estos parámetros –Tabla V–, aparece en las reglas que utilizan los hidrocarburos (Eventos E6, E7), en consonancia con la importancia de esta serie de compuestos en la caracterización de hojiMálaga. Tan sólo el evento E1, que no usa compuestos de esa serie, aparece entre los de mayor LS. Por el contrario y lógicamente, el grado más alto de necesidad, valor de LN más bajo, pertenece a una regla que no aplica los hidrocarburos (Evento E2). Se deduce, pues, que la utilización de los hidrocarburos incrementa en gran manera la probabilidad de verificar la hipótesis H2, pero que las reglas que combinan los ácidos grasos y los hidrocarburos poseen el más alto nivel de suficiencia, pudiéndose deducir que esta última serie, por sí sola, no podría caracterizar a hojiMálaga.

Sin embargo, atendiendo a consideraciones generales, la probabilidad asociada con los tipos primero y tercero no variaría si se conociera completamente el universo descrito por la hipótesis. Sin embargo, las probabilidades asociadas con los tipos segundo y cuarto podrían cambiar, aún conociéndose completamente el universo descrito por la hipótesis (p.e., aceite de hojiMálaga), al incrementarse el universo conocido (p.e., aceite de España) con otras muestras que pueden estar o no incluidas dentro del ámbito de la hipótesis (p.e., aceites de Lérica).

Evento	LS	LN
E1	15.0	0.26
E2	4.1	0.01
E3	5.4	0.42
E4	8.7	0.65
E5	15.0	0.26
E6	>750	0.25
E7	17.4	0.26
E8	5.0	0.56
E9	14.5	0.39
E10	10.4	0.53

Tabla V

Grado de Suficiencia y de Necesidad asociado con cada evento para una regla del tipo 2, "SI E, ENTONCES H2"

### 3.3. Reglas de decisión lineales.

Independientemente de esas reglas relacionales, los programas estadísticos permiten construir ecuaciones matemáticas de primer orden a partir de los datos de los cromatogramas. Un análisis discriminante por pasos ha sido el procedimiento utilizado para obtener este tipo de reglas de decisión (15). Las reglas están basadas en ecuaciones lineales cuyo límite máximo de probabilidad ha sido obtenido mediante el método de verificación de Jackknife. Las variables químicas seleccionadas en cada análisis discriminante han sido acotadas a aquellas que poseen un F-Enter mayor de 5.58, y el número total de éstas es inferior al propuesto en (16), con el fin de evitar que las soluciones que se presentan se deban a la casualidad.

En la Tabla IV se describen las ecuaciones que constituyen las reglas de decisión lineales que, de

Tabla VI

Ecuaciones matemáticas para clasificar las muestras de hojiMálaga.

Evento	Descripción del evento
E1	$-1.05 \cdot \text{Palmitico} + 0.62 \cdot \text{Linoleico} + 6.73$
E2	$-0.38 \cdot \text{Hexacosanol} + 1.01 \cdot \text{Octacosanol} + 2.58 \cdot \text{Taraxerol} - 1.36$
E3	$51.12 \cdot \text{Copaeno} - 64.39 \cdot \text{Muuroleno} - 1.94$
E4	$-0.76 \cdot \text{Campesterol} + 0.51 \cdot \text{Delta5-Avenasterol} - 0.74$
E5	$3.20 \cdot \text{Aldehido Oleanólico} - 2.26$
E6	$-1.93 \cdot \text{Escualeno} + 0.13 \cdot \text{Cicloeucaalenol} + 47.6 \cdot \text{Copaeno} - 60.6 \cdot \text{Muuroleno} - 1.9$
E7	$0.37 \cdot \text{Cicloeucaalenol} - 1.99$

Tabla VII

Probabilidad asociada con cada una de las reglas lineales.

Tipo	Evento	Probabilidad		Límites Histogramas					
		hoji Málaga	No_hoji Málaga	hojiMálaga			No_hojiMálaga		
				min	med	max	min	med	max
1	E1	99.0	94.5	( 1.7	3.2	5.6)	(-5.0	-0.2	3.2)
1	E2	87.5	86.6	( 1.0	2.1	3.2)	(-2.4	-0.1	3.5)
1	E3	87.5	96.3	( 2.2	4.8	8.6)	(-1.5	-0.2	3.4)
1	E4	87.5	82.9	( 0.7	1.9	2.7)	(-2.0	-0.1	2.7)
1	E5	87.5	81.1	( 0.7	1.4	2.2)	(-2.3	-0.7	4.6)
1	E6	87.5	98.2	( 2.6	5.6	7.8)	(-2.0	-0.3	3.0)
1	E7	62.5	92.7	( 0.2	2.6	6.1)	(-1.6	-0.1	3.9)

acuerdo con el esquema de la Tabla II, pertenecen al tipo 1. Los coeficientes de las ecuaciones han sido obtenidos usando, cada vez, sólo los compuestos de cada una de las series analizadas y, en sólo una ocasión, simultáneamente todos los compuestos químicos. En la Tabla VII aparecen los niveles máximos de probabilidad y los límites máximos, medios y mínimos de las distribuciones de los datos al aplicar esas ecuaciones. No obstante, a pesar de describirse siete reglas, Tabla VI, sólo las seis primeras serán aplicadas en este trabajo, pues la séptima no tiene una probabilidad predictiva suficiente y podría introducir errores (ruido de fondo).

Las probabilidades asociadas con las reglas lineales confirman que la serie de los hidrocarburos poseen la mayor capacidad discriminatoria o de decisión sobre si una muestra "no es" de hojijMálaga (96.3%, 98.2%) que "sí es" de hojijMálaga (87.5%). Tan sólo el grupo de los ácidos grasos posee un nivel superior de probabilidad en hojijMálaga (99.0%), con los que también se confirma la importancia de los ácidos grasos en la caracterización de estos aceites. Respecto a las otras reglas, tan sólo una de las muestras de hojijMálaga no es clasificada en su grupo, se trata alternativamente de las muestras MANT06 –Fuente Piedra– y MARC03 –Cuevas de San Marcos– que se encuentran en los límites de la provincia de Málaga.

Por otra parte, y en relación a las muestras de otras zonas de Andalucía, aparecen como similares a hojijMálaga aquellas que tiene una mayor proximidad geográfica y varietal, aceites de Sevilla (OleoEstepa) o de la comarca Penibética de Córdoba y, en raras ocasiones de zonas más distantes, como la Sierra del Segura, y siempre con baja probabilidad de pertenencia a hojijMálaga.

Este grupo de reglas, del mismo modo que las reglas relacionales, permiten al usuario decidir el número de parámetros químicos a cuantificar, que puede oscilar entre una sola serie y la totalidad de los compuestos indicados en la figura 1. La decisión podría tomarse en virtud de la relación costo/tiempo del análisis y el nivel máximo de probabilidad que, inicialmente, tiene cada serie de compuestos.

Sin embargo, si se aplicase más de una regla, bien por que se cuantificase más de una serie de compuestos o porque una de las series permitiese más de una regla, las probabilidades tendrían que poder combinarse para obtener la probabilidad final de clasificación. Esto implicaría la necesidad de poseer algoritmos matemáticos que pudiesen operar con las probabilidades obtenidas, tanto por ecuaciones matemáticas lineales como relacionales.

### 3.4. Combinación de reglas.

Cada regla relacional o lineal tiene capacidad para identificar una muestra desconocida como perteneciente a la categoría que se analiza de hojijMálaga. Sin

embargo, el resultado de aplicar sólo una regla puede no ser satisfactorio para los límites mínimos de probabilidad que un analista se haya fijado. Y, por otra parte, la teoría de la posibilidad ofrece diferentes tipos de algoritmos capaces de combinar los resultados obtenidos con las reglas de decisión.

En este trabajo se ha optado por la aplicación de uno de ellos perteneciente a la teoría de la evidencia, en la misma forma en que se desarrolló en (1). No obstante, la aplicación de esta teoría a una discriminación entre dos grupos, sólo puede considerarse a título de ejemplo ilustrativo de sus posibilidades cuando se aplica sobre un número mayor de grupos.

Cada regla, en esta teoría, tiene asociado un nivel de credibilidad (MB) o incredibilidad (MD) que se calcula considerando lo que aporta la regla sobre la probabilidad 'a priori' de la hipótesis,

$$MB(H|E) = \begin{cases} \frac{[\text{Max}[P(H|E), P(H)] - P(H)]}{[1 - P(H)]}, & \text{si } P(H) = 1 \\ 1, & \text{si } P(H) = 0 \end{cases} \quad (3.12)$$

$$MD(H|E) = \begin{cases} \frac{[P(H) - \text{Min}[P(H|E), P(H)]]}{[1 - P(H)]}, & \text{si } P(H) = 0 \\ 1, & \text{si } P(H) = 1 \end{cases} \quad (3.13)$$

Los grados de MB y MD asociados con cada regla se combinan mediante la T-norma S2, para determinar la credibilidad o incredibilidad final asociada con la hipótesis. Se ha elegido la T-norma S2 por encontrarse su valoración entre la optimista de S0 y la pesimista de S3 (17). El algoritmo matemático de esta T-norma, en términos de probabilidad, es

$$P(H|E1 \cup E2) = \frac{P(H|E1) + P(H|E2) - P(H|E1) * P(H|E2)}{P(H|E2)} \quad (3.14)$$

que expresado en términos de la teoría de la evidencia es:

$$MB(H|E1 \cup E2) = \frac{MB(H|E1) + MB(H|E2) - MB(H|E1) * MB(H|E2)}{MB(H|E2)} \quad (3.15)$$

$$MD(H|E1 \cup E2) = \frac{MD(H|E1) + MD(H|E2) - MD(H|E1) * MD(H|E2)}{MD(H|E2)} \quad (3.16)$$

A partir de la credibilidad e incredibilidad se calcula el factor de certeza asociado a la hipótesis

$$CF(H|E) = MB(H|E) - MD(H|E) \quad (3.17)$$

con el que se valorarán las muestras desconocidas para conocer la bondad de las reglas descritas en las Tablas III y VI.



### 3.5. Verificación de las reglas

Previamente al análisis de los resultados de verificación de las reglas deben tenerse en cuenta dos consideraciones. Las dos clases de reglas, relacionales y lineales, siguen dos procesos diferentes en el cálculo de las valoraciones finales asociadas con ellas. Las reglas relacionales sólo indican la probabilidad con que una muestra puede verificar la hipótesis de pertenencia a hojimaálaga. Por el contrario, las reglas lineales muestran la probabilidad tanto de pertenencia como de no pertenencia a la misma, aunque sólo del tipo 1, frente a las relacionales que permiten obtener los cuatro tipos de la Tabla II.

Por otra parte, no todas las reglas propuestas en la Tablas III y VI han sido aplicadas a todas las muestras con las que se verificará el sistema, tan sólo una pequeña parte de las muestras usadas para verificar el sistema poseen información sobre las series

Para verificar la utilidad de las reglas de decisión descritas se seleccionaron 26 muestras de diferentes variedades, provincias y naciones.

En la Tabla VIII aparecen los resultados, tanto de las muestras sobre las que fueron cuantificados los 54 parámetros químicos (grupo H), como de aquellas en las que no fueron analizados los hidrocarburos (grupo SH).

La primera columna corresponde al coeficiente de certeza obtenido con las reglas lineales del tipo 2. En el caso de las muestras del grupo H se han aplicado todos los eventos de la Tabla VI, mientras que sólo E1, E4 y E5 han sido aplicados a las muestras del grupo SH.

El coeficiente de certeza, ec. (3.17), está comprendido entre  $[-1, +1]$ , el valor negativo corresponde a la negación absoluta de la hipótesis y el positivo a su aceptación. Sin embargo, ese margen ha sido

Tabla VIII

Coeficientes de certeza de las muestras de verificación. Grupo: H=con hidrocarburos, SH=sin hidrocarburos.

Código	Población	Zona	Cosecha	CF		Grupo
				lineales	relacionales	
ITAL-1	Ventimiglia	Liguria	1988	0.51	0.30	H
ITAL-7	Lorenzo	Liguria	1988	0.34	0.00	H
C-H87	Córdoba	Córdoba	1987	0.28	0.83	H
C-H88	Córdoba	Córdoba	1988	0.67	0.90	H
CA19-1	Puente Genil	Córdoba	1988	0.44	0.39	H
CA19-2	Puente Genil	Córdoba	1988	0.39	0.00	H
MA-36	Antequera	Málaga	1987	0.66	0.99	H
MA-46	Archidona	Málaga	1987	0.79	0.99	H
ITAL-11	Aliano	Basilicata	1989	0.60	0.13	SH
ITAL-12	Stigliano	Basilicata	1989	0.39	0.39	SH
ITAL-52	Pistoia	Toscana	1989	0.41	0.00	SH
ITAL-53	Vinci	Toscana	1989	0.09	0.00	SH
ITAL-54	Grosseto	Toscana	1989	0.45	0.00	SH
ITAL-55	Arezzo	Toscana	1989	0.41	0.28	SH
CC1-4	Montanches	Cáceres	1989	0.34	0.00	SH
TE-5	Andorra	Teruel	1989	0.45	0.00	SH
TO-6	Mora Toledo	Toledo	1989	0.44	0.13	SH
CS-4-1	Segorbe	Castellón	1989	0.56	0.43	SH
MARC01	Archidona	Málaga	1989	0.64	0.51	SH
MARC03	V.Trabuco	Málaga	1989	0.44	0.43	SH
MARC04	C.S.Marcos	Málaga	1989	0.52	0.59	SH
MANT01	Antequera	Málaga	1989	0.55	0.51	SH
MANT06	Fuentepiedra	Málaga	1989	0.44	0.51	SH
SEST02	Herrera	Sevilla	1989	0.66	0.39	SH
SEST06	Gilena	Sevilla	1989	0.39	0.13	SH
SEST08	Corcoya	Sevilla	1989	0.44	0.59	SH

de compuestos de los hidrocarburos. Esta limitación, no obstante, permitirá observar la importancia de los hidrocarburos para desechar una muestra que no verifique la hipótesis de pertenencia a hojimaálaga. Así como, en el sentido opuesto, determinar su capacidad discriminatoria con las muestras que son de hojimaálaga o muy similares químicamente, caso de OleoEs-tepa (Sevilla).

En la verificación del sistema, sólo se analizará el tipo segundo, de entre los cuatro calculados, por considerarse el más interesante desde el punto de vista de identificación de muestras desconocidas en relación a unos patrones. En el caso de las reglas lineales, este tipo se deduce mediante la aplicación del Teorema de Bayes.

trasladado, para su mejor comprensión, a los límites de la teoría de la probabilidad  $[0, 1]$ .

Analizando los resultados de las reglas de decisión lineales y sólo del grupo H, se puede observar que sólo la muestra C-H88 (hojiblanco, Parque de variedades de Córdoba, cosecha 1987-88) ha sido erróneamente clasificada como hojimaálaga. Mientras que lo han sido correctamente las muestras MA-46 (hojimaálaga, Antequera, cosecha 1986-87) y MA-36 (hojimaálaga, Archidona, cosecha 1986-87).

En cambio en las muestras del grupo SH, se deduce que la no aplicación de las reglas lineales cuyos coeficientes son hidrocarburos produce una disminución del nivel de certeza. Esta disminución es

mayor en relación a la capacidad por determinar si las muestras de hojimaálaga son de ese grupo (MANT06, Fuentepiedra, y MARC03, Cuevas de San Marcos) que de aquellas que no pertenecen a esa variedad y zona, ITAL-11 (Basilicata, Italia), CS-4-1 (Segorbe, Castellón) y SEST02 (Herrera, Sevilla). Sin embargo, hay que resaltar que las dos muestras de hojimaálaga están en los límites de la provincia de Málaga, una en las proximidades de Sevilla y la otra en la de Córdoba, y son las causantes de que las probabilidades de la Tabla VII no sean mejores.

Respecto a las muestras de Italia y Castellón, hay que tener en cuenta que las reglas se construyeron para distinguir aceites dentro de Andalucía y, en esos y otros casos, se han aplicado sobre unos aceites de los que el sistema no tiene información. No obstante, sólo ha cometido tres errores con un bajo nivel de asignación, exceptuando ITAL-11.

La segunda columna muestra el coeficiente de certeza asociado con las reglas relacionales. Estas poseen dos ventajas en la caracterización de hojimaálaga. La primera discriminar aquellas muestras que no son de hojimaálaga o próximas a ella. La segunda, la poca incidencia de la variación de los datos, con las cosechas, en los niveles de probabilidad calculados. Sin embargo, tienen un inconveniente, los bajos niveles de probabilidad de las tres reglas que no utilizan los hidrocarburos.

Atendiendo a esas consideraciones, las muestras C-H87, C-H88 (variedad hojiblanco, Parque de variedades de Córdoba) del grupo H aparecen, erróneamente, como pertenecientes a hojimaálaga. Mientras que en el grupo SH sólo aparece SEST08 (variedad hojiblanco, Corcoya). Aunque la muestra MARC03, también mal clasificada con las reglas lineales, aparece como no perteneciente a hojimaálaga.

En resumen, la base de conocimientos de SEXIA, aplicada a la clasificación de los aceites de oliva vírgenes de hojimaálaga, permite clasificar correctamente sus aceites en el 85.7% de los casos, usando las reglas relacionales, y en el 71.4%, con las lineales. El 15.8% de los aceites que no pertenecían a hojimaálaga fueron erróneamente clasificados en ese grupo, usando las mismas reglas relacionales, y el 26.3% usando las lineales.

No obstante, la verificación del procedimiento podría considerarse incompleta si no se discutieran los resultados con los obtenidos con sólo métodos estadísticos. Sin embargo, para ese contraste se debería elegir la mejor regla, sólo es posible aplicar una, y cualquiera que ésta fuese, sería una de las existentes en la base de conocimientos, con lo que el sistema siempre podría usarla junto con otras más, mejorando su probabilidad.

#### 4. CONCLUSIONES

Para el análisis del modo de funcionamiento y construcción de parte de la base de conocimientos del sistema experto SEXIA, se ha elegido un estudio sobre aceites de la provincia de Málaga.

Se querían poner de manifiesto las posibilidades del sistema cuando se desea caracterizar un aceite de oliva virgen de una zona que posee cierta similitud varietal, geográfica (altitud, clima) y edafológica con comarcas próximas, pertenecientes a otras provincias, pero de las que el sistema posee una información muestral representativa.

Por otra parte, se han verificado las reglas con muestras cuya información química es, en algunos casos, incompleta en relación a los patrones.

Finalmente, el universo al que se circunscriben las reglas, aceite de Andalucía, es un subconjunto del universo constituido por las muestras de aceite de España e Italia, al que pertenecen algunas muestras usadas en el proceso de verificación. Se ha hecho, en consecuencia, una extensión de las posibilidades de aplicación de las reglas más allá de sus límites.

No obstante esas dificultades, introducidas para conocer la posibilidad de extensión del método y la posibilidad de aplicación del sistema usando sólo una parte de su base de conocimientos, los resultados pueden considerarse aceptables. En consecuencia, se pueden deducir algunas conclusiones de carácter particular, relativas a la caracterización de hojimaálaga, y otras generales, relacionadas con la forma de construir la base de conocimientos del sistema experto.

Entre las primeras, se observa,

- que cuando se usan los hidrocarburos se pierde capacidad discriminadora entre la variedad hojiblanco y el caso particular de hojimaálaga y,
- que en la combinación de ácidos grasos e hidrocarburos -regla relacional E6- está el mayor poder discriminante del sistema en relación a hojimaálaga.

Entre las segundas,

- que la extensión de la aplicación a otro universo parece no afectar, en términos generales, a los resultados y, en particular, a las reglas de decisión relacionales.
- que el usuario puede elegir las series de compuestos que desea analizar para identificar una muestra y, en consecuencia, aplicar todas o algunas de las reglas relacionales o lineales, y analizar los resultados independiente o conjuntamente y,

- a medida que la base de datos de SEXIA se vaya incrementando con los resultados de otras campañas oleícolas o con otras series de compuestos, mejorarán las posibilidades de la caracterización de estos aceites.

### AGRADECIMIENTO

Los autores desean mostrar su agradecimiento al Dr. D. Arturo Cert Ventulá por la puesta a punto de los métodos analíticos, organización del laboratorio de análisis e inestimable ayuda en la detección y corrección de errores. Al Centro de Investigación y Desarrollo Agrario "Las Torres" cuyos becarios Ldos. D. Manuel León, D. Manuel Martín Polvillo y Lda. Dña. Mercedes Sánchez Navarro realizaron los análisis con la colaboración de la Lda. Dña. Soledad Ferreiro Almeda y de Dña. Rosario González Cordones, Dña. Teresa Rubio Oliver y D. Manuel Rodríguez Aguilar.

Al Prof. Forina del Instituto de Farmacología y Tecnología Alimentaria de la Universidad de Génova (Italia), a las cooperativas oleícolas participantes en el proyecto SEXIA y a Metaponto-Agrobios (Italia) por colaborar desinteresadamente con muestras de aceite de oliva virgen.

Este proyecto ha sido realizado con fondos de la Dirección General de Investigación y Extensión Agraria de la Junta de Andalucía (RAEA, A6.87-1), Comisión Interministerial de Ciencia y Tecnología (ALI88-0187-C02-02) y Convenio Marco de Investigación España-Italia (PR.3-1)

### REFERENCIAS

1. Aparicio, R.- "Characterization of foods by inexact rules: The SEXIA expert system".- *J. Chemometrics* **3** (1988) 175-192.
2. Aparicio, R.; Ferreiro, L.; Cert, A., y Lanzón, A.- "Caracterización de aceites de oliva vírgenes andaluces".- *Grasas y Aceites* **41** (1990) 23-39.
3. Breiman, L.; Friedman, J. H.; Olshen, R. A., and Stone, C. J.- "Classification and Regression Trees".- Wadsworth International Group, Belmont CA, 1984.
4. Derde, M.P.; Buydens, L.; Guns C.; Massart, D.L., and Hopke, P.K.- "Comparison of rule-building expert systems with pattern recognition for the classification of analytical data".- *Anal Chem.* **59** (1987) 1868-1871.
5. Frank, I.E., and Lanteri, S.- "Classification Models: Discriminant Analysis, SIMCA, CART".- *Chemlab* **5** (1989) 247-256.
6. Shirai, Y., and Tsujii, J.- *Artificial Intelligence: Concepts, Techniques and Applications*. John Wiley and Sons, Chichester 1985.
7. Scarponi, G; Moret, I; Capodaglio, G; Cescon, P.- "Multiple Discriminant Analysis in the Analytical Differentiation of Venetian Wines. 3. A reevaluation with addition of data from samples of 1979 Vintage Prosecco wine".- *J. Agric. Food Chem.* **30** (1982) 1135-1140.
8. Aparicio, R.; Albi, T.; Cert, A., y Lanzón, A.- "Sistema experto SEXIA: Ecuaciones canónicas para caracterizar aceites de oliva españoles por variedades".- *Grasas y Aceites* **39** (1988) 219-228.
9. Aparicio, R., Ferreiro, L., Cert, A.- "Informe final Proyecto OF.A6-87, 1.-"Dirección General de Investigación y Extensión Agraria. Junta de Andalucía.
10. Cert, A.- "Aplicación de la cromatografía gaseosa con columnas capilares al análisis de la fracción alcohólica del insaponificable de los aceites".- XXIII Asamblea del Instituto de la Grasa, Abril, 1989. (Sin publicar).
11. "BMDP Statistical software".- University of California, Los Angeles, 1981.
12. Jones, E.- "Aplique el DBase III Plus".- Osborne/McGraw Hill, 1987.
13. Johnson, N.- "AutoCAD Manual de Referencia".- Osborne/McGraw Hill, 1989.
14. Mandami, E.H., and Gaines, B.R.- "Fuzzy Reasoning and its applications".- Academic Press, London
15. Aparicio, R., and Albi, T.- "The SEXIA, An Expert System Identifying Olive Oils".- II World Congress of Food Technology, Prosema Ed., Barcelona (1987), p. 2915-2923.
16. Lavine, B.K., Jurs, P.C., Henry, D.R., Van der Meer, R.K., Pino, J.A., and McMurry, J.E.- "Pattern Recognition Studies of Complex Chromatographic Data Sets: Design and Analysis of Pattern Recognition Experiments".- *Chemometrics and Intell. Lab. Sys.* **3** (1988) 79-89
17. Godo L.I., López de Mantaras R., Sierra C., Verdaguer A.- "Managing Linguistically expressed uncertainty in MILORD- Application to Medical Diagnosis".- Proc. 7th Int. Workshop on Expert Systemn & their Applications (1987) 571-597.

(Recibido: Julio 1990)