

# **APLICABILIDAD DEL PERFIL DE ESTEROLES PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN ALIMENTOS**

TESIS DOCTORAL

**Estefanía Pérez Castaño**

Para optar al grado de:

**Doctora en Química por la Universidad de Granada**

PROGRAMA OFICIAL DE DOCTORADO EN QUÍMICA (P34 56 1)

Tesis doctoral dirigida por:

Dra. Mercedes Sánchez Viñas

Dr. Domingo Gázquez Evangelista

Dra. María Gracia Bagur González

**~ GRANADA, 2012 ~**

Editor: Editorial de la Universidad de Granada  
Autor: Estefanía Pérez Castaño  
D.L.: GR 884-2013  
ISBN: 978-84-9028-475-9



# **APLICABILIDAD DEL PERFIL DE ESTEROLES PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN ALIMENTOS**

por

**ESTEFANÍA PÉREZ CASTAÑO**

**DEPARTAMENTO DE QUÍMICA ANALÍTICA  
UNIVERSIDAD DE GRANADA**

Visado en Granada a 27 de Septiembre de 2012

Fdo. Dr. Dña. Mercedes Sánchez Viñas  
Profesora Titular del Departamento de Química Analítica  
Facultad de Ciencias. Universidad de Granada

Fdo. Dr. D. Domingo Gázquez Evangelista  
Profesor Titular del Departamento de Química Analítica  
Facultad de Ciencias. Universidad de Granada

Fdo. Dr. Dña. María Gracia Bagur González  
Profesora Titular del Departamento de Química Analítica  
Facultad de Ciencias. Universidad de Granada

**Memoria para optar al grado de Doctor en Química por la UGR:**

Fdo: Estefanía Pérez Castaño



La doctoranda Estefanía Pérez Castaño y los directores de la tesis Mercedes Sánchez Viñas, Domingo Gázquez Evangelista y M<sup>a</sup> Gracia Bagur González Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, a 27 de Septiembre de 2012

Director/es de la Tesis

Doctorando

Fdo.: Dra. Mercedes Sánchez Viñas

Fdo.: Estefanía Pérez Castaño

Fdo.: Dr. Domingo Gázquez Evangelista

Fdo.: Dra. María Gracia Bagur González



# **AGRADECIMIENTOS**





*A mis padres*  
*Manuel y Josefa*



A mis directores de tesis, a la Dra. Mercedes Sánchez, a la Dra. María Gracia Bagur y al Dr. Domingo Gázquez, por ayudarme a recorrer este camino haciéndolo más fácil a pesar de los inconvenientes que iban surgiendo, por sus ideas, por su paciencia y apoyo recibido y por compartir conmigo muchos momentos alegres.

Un agradecimiento especial al Dr. Luis Cuadros, por brindarme su ayuda cuando la necesitaba, por los consejos que me daba, por su cariño y ánimos recibos, por el surgimiento de una sólida amistad. Tendré muy presente aquella frase que un día me dijiste y que gran verdad dice “La suerte no se encuentra sino se busca”.

También quiero agradecer por la ayuda recibida al Dr. Antonio González, por ofrecerse en todo momento, en echar nuestros ratos de risa y como no, debatiendo qué equipo de fútbol español juega mejor, ¿verdad?

Entre los profesores del departamento, a pesar que con todos he tenido una excelente relación, debo destacar al Dr. Juan Manuel Bosque, al Dr. Carlos Jiménez por compartir aficiones, al Dr. Ignacio de Orbe, al Dr. Alberto Navalón, al Dr. Luis Fermín, a la Dra. María del Carmen Valencia, a la Dra. Paqui Molina, a la Dra. Natalia Navas, al Dr. Óscar Ballesteros y al Dr. Alberto Zafra. Gracias por vuestra ayuda y apoyo cuando os lo he pedido o simplemente por preocuparse por mí.

A mis compañeras de laboratorio, Cristina y Carmen por hacer más llevadero el día a día en el laboratorio. Echaré de menos los momentos que hemos vivido juntas como por ejemplo cuando al final de la tarde Carmen me cantaba o los viernes de complacencias cuando estaba Paulina con nosotras. A las nuevas incorporaciones, Lucía y Verónica que aunque hemos compartido poco tiempo juntas me habéis demostrado que sois unas personas estupendas. También a mis otras compañeras que estuvieron en nuestro grupo de investigación a Carmen, Irene y Marta.

A Julio y Alegría que juntos formamos el trío lalala. Julio, no voy a olvidar nuestros ratos de cháchara a última hora de la tarde cuando no quedaba casi nadie ya en el Departamento, contándonos nuestras penas y alegrías. Has sabido escucharme y entenderme. Gracias.

Alegría, qué voy a decir de ti, eres una gran persona y como compañera has estado muy pendiente preocupándote de mi día a día sobre todo en la recta final. Seguiremos siendo el trío lalala, eso no va a cambiar a pesar de que cada uno se encuentre en un sitio diferente. Buscaremos tiempo para recordar los momentos vividos juntos en estos 4 años que he compartido con vosotros.

A Ismael, que hemos congeniado muy bien desde el momento que nos conocimos. Ahora apenas nos vemos porque nos debemos a nuestro trabajo pero sé que puedo contar contigo y me lo has demostrado. Echaré de menos tu entrada sigilosa al laboratorio para pegarme un susto y las “historietas” que me contabas.

A María Porcel, por ser como eres, pero sobre todo como no, por mi PINGU. Te encomiendo una única tarea y es que lo trates muy bien. Qué no pase frío!!!



A Rosa, José Manuel (mi inspector favorito), Jesús, Isa, Salva, que sois increíbles, no sé qué deciros porque a veces las palabras sobran sólo con un simple gesto es suficiente por eso mil y mil gracias por creer en mí. Cada uno de nosotros estamos cumpliendo una meta y me alegra saber que la compartamos juntos. Isa ya sólo faltas tú por ser doctora, ánimo.

También a todos mis compañeros del departamento, María, Sonia, M<sup>a</sup> Luisa, Inma, Manuel, Alfonso, Alejandro, David, Fernando, Rocío, Elena, Marta, Santi, Paulina “la contri”, Barto, Noemí, Javi Espejo del departamento de Ingeniería Química, todos... A ti Alejandro, gracias por pasarte a mi laboratorio todos los sábados al final de la mañana para preguntar cómo iba con la escritura de la tesis. A Santi, por sacar tiempo y buscar un huequecillo para hacer varias macros en Matlab, haciéndolo así todo más fácil. A Javi Espejo por compartir esta etapa de la vida que emprendimos juntos al mismo tiempo y que ya ha llegado a su fin con muy buenos resultados.

Al Dr. Jesús Villén, por acogerme en su grupo de investigación durante mi estancia llevada a cabo en Albacete. Los tres meses que estuve allí son inolvidables para mí. A la Dr. Ana Vázquez y al Prof. Enrique. Me recibisteis como si fuera una más de vosotros. A mis compis Rosa, mi popi favorita y Álvaro, mi esbirro. Cada día que pasaba nuestra amistad se afianzaba. ¡Anda que no tenemos anécdotas que contar! ¿a qué sí?. Rosa siempre me acuerdo de ti cuando paso por la puerta de Planta Baja aquí en Granada. Álvaro ahora sí que es verdad que no te puedes ir al extranjero porque tienes una niña que criar, ¿te acuerdas? Y como no, nombrar a nuestra amiga griega Alexia, aún recuerdo el día que nos dijo una frase que había aprendido en clases de español y que nosotros le chinchábamos con cariño para que la pusiera en práctica. Me quedo con tu buen humor.

Al personal del Departamento de Aceites y Grasas en el Laboratorio Agroalimentario de Granada (calle Avda. de la Diputación, Atarfe, Granada), entre ellos a Juan Durán, a Dori, a Chema, a María, a Pepe, a el Nene, a Paco y a Alfonso.

A mis nuevos compis de trabajo de Laboratorios Farmacéuticos Rovi, S.A., que a pesar de que nos hemos conocido recientemente, me habéis animado en esta última etapa de escritura de la tesis.

Quiero agradecer especialmente a aquellas personas importantes que forman parte de mi vida y que han estado conmigo en todo momento, me han tendido su mano cada vez que me iba a caer y me han levantado. A Sergio, a Jose, a Carolina, a Ana José, a Alberto, a Antonio, a Ana María, a A.J., a Julia, a Inma y a Manuel. Sin vosotros hubiera sido difícil estar ahí al pie del cañón. Os llevaré en mi corazón, os quiero. Eso nunca lo dudéis. Estoy absolutamente convencida de que si de algo puedo presumir en esta vida es de teneros, lo que me hace sentir una persona muy afortunada. Joselillo muchisisisimas gracias por el diseño de la portada, me *SÚPER ENCANTA*. “El río fluye, va ensanchándose por su cauce y recorriendo recovecos que antes no podía”.

Y, por supuesto, el agradecimiento más profundo y sentido va para mi familia. A mis padres, Manuel y Josefa, por apoyarme SIEMPRE, por hacer de la persona que soy hoy en día, por su ejemplo de lucha y honestidad, por su paciencia ya que a veces soy de armas tomar. Quiero que os sintáis bastantes orgullosos por la hija que tenéis. A mi abuela, que aunque ya no está conmigo ha estado presente en casi todo este camino en el que este sueño pasa a ser una realidad. Al resto de familiares que han estado pendientes de mi preguntándome, ¿pero chiquilla cuanto te queda ya para ser doctora? Jajaja. Ahora os puedo contestar que ya ha llegado el momento y que ha empezado una nueva etapa en mi vida y que por supuesto quiero que forméis parte de ella.

***A TODOS Y CADA UNO DE VOSOTROS OS MANDO  
UNA SONRISA Y “NO PUEDO RESPONDER OTRA  
COSA QUE GRACIAS Y GRACIAS” (W. Shakespeare)***



Hoy puede ser un gran día,  
plantéatelo así,  
aprovecharlo o que pase de largo,  
depende en parte de ti.

Dale el día libre a la experiencia  
para comenzar,  
y recíbelo como si fuera  
fiesta de guardar.

No consientas que se esfume,  
asómate y consume  
la vida a granel.  
Hoy puede ser un gran día,  
duro con él.

Hoy puede ser un gran día  
donde todo está por descubrir,  
si lo empleas como el último  
que te toca vivir.

Saca de paseo a tus instintos  
y ventílalos al sol  
y no dosifiques los placeres;  
si puedes, derróchalos.

Si la rutina te aplasta,  
dile que ya basta  
de mediocridad.  
Hoy puede ser un gran día  
date una oportunidad.

Hoy puede ser un gran día  
imposible de recuperar,  
un ejemplar único,  
no lo dejes escapar.

Que todo cuanto te rodea  
lo han puesto para ti.  
No lo mires desde la ventana  
y siéntate al festín.

Pelea por lo que quieres  
y no desesperes  
si algo no anda bien.  
Hoy puede ser un gran día  
y mañana también.

Hoy puede ser un gran día  
duro, duro,  
duro con él.

**Joan Manuel Serrat**





# ÍNDICE



## ÍNDICE

	Página
RESUMEN.....	3
CONTEXTUALIZACIÓN	7
OBJETIVOS .....	11
INTRODUCCIÓN .....	15
METODOLOGÍA Y RESULTADOS	
CAPÍTULO I: Bases quimiométricas para el estudio del perfil de esteroides en aceites.....	81
I.1. RECONOCIMIENTO DE PAUTAS .....	83
I.2. MÉTODOS NO SUPERVISADOS.....	84
I.3. MÉTODOS SUPERVISADOS .....	98
I.4. PRE-PROCESADO DE DATOS.....	123
I.5. PRINCIPALES SOFTWARES UTILIZADOS EN QUIMIOMETRÍA.....	130
CAPÍTULO II: Materiales y Métodos .....	137
II.1. PATRONES DE FITOESTEROLES.....	138
II.2. OTRAS DISOLUCIONES .....	139
II.3. REACTIVOS Y DISOLVENTES .....	139
II.4. EQUIPOS DE LABORATORIO.....	140
II.5. MUESTRAS ANALIZADAS .....	143
II.6. PROCEDIMIENTO OPERATORIO DEL MÉTODO OFICIAL .....	151
CAPÍTULO III: Desarrollo de un método para la determinación de 4-desmetilesteroides en aceites vegetales, previa separación "off-line" mediante HPLC .....	155
III.1. EVALUACIÓN DE LA REACCIÓN DE DERIVATIZACIÓN DE LOS FITOESTEROLES PARA SU DETERMINACIÓN MEDIANTE GC .....	157
III.2. VERIFICACIÓN DE LA APTITUD DE LA ETAPA DE OBTENCIÓN DE LOS TRIMETILSILILDERIVADOS.....	161
III.3. SEPARACIÓN, MEDIANTE CROMATOGRAFÍA LÍQUIDA, DE LA FRACCIÓN DE FITOESTEROLES A PARTIR DEL EXTRACTO INSAPONIFICABLE. OPTIMIZACIÓN DE LA VENTANA DE TIEMPO.....	170
III.4. VALIDACIÓN DEL MÉTODO PROPUESTO ATENDIENDO A SU APTITUD .....	176
III.4.1. Evaluación de la incertidumbre asociada a la determinación de fitoesteroides.....	177

III.4.2.	Estimación de la veracidad.....	193
CAPÍTULO IV: Estudio quimiométrico del perfil de concentraciones de los 4-desmetilesteroles en aceites vegetales .....		203
IV.1	ANÁLISIS DE DIFERENTES MUESTRAS DE ACEITES VEGETALES MEDIANTE HPLC-GC(FID) OFF LINE .....	204
IV.2	ESTIMACION DE LOS LÍMITES DE DETECCIÓN ATENDIENDO A LA MATRIZ EN LA QUE SE ENCUENTRAN LOS ANALITOS.....	218
IV.3	PROCEDIMIENTO SEGUIDO PARA EL ESTUDIO QUIMIOMETRICO DE LOS PERFILES DE CONCENTRACION DE LOS 4-DESMITELESTEROLES DETERMINADOS EN ACEITES VEGETALES .....	219
IV.4	ESTIMACIÓN DE LA NORMALIDAD DE LOS DATOS: TRANSFORMACIÓN DE LOS PERFILES DE CONCENTRACIÓN A UNA DISTRIBUCIÓN NORMAL MEDIANTE LA TRANSFORMACIÓN DE BOX-COX.....	221
IV.4.1	Análisis estadístico de los perfiles de concentraciones en función de los tipos de aceites analizados .....	223
IV.4.2	Análisis de las gráficas de rayos obtenidas para los perfiles de concentraciones en función de los tipos de aceites analizados.....	224
IV.4.3	Normalización de los perfiles de concentraciones de los 4-desmetilesteroles .....	226
IV.5	TÉCNICAS NO SUPERVISADAS DE RECONOCIMIENTO DE PAUTAS: ANÁLISIS FACTORIAL DEL PERFIL DE CONCENTRACIONES NORMALIZADAS .....	233
IV.6	TÉCNICAS SUPERVISADAS DE RECONOCIMIENTO DE PAUTAS: ANÁLISIS DISCRIMINANTE LINEAL DEL PERFIL DE CONCENTRACIONES NORMALIZADAS .....	236
CAPÍTULO V: Estudio sobre la capacidad de la huella dactilar cromatográfica para diferenciar aceites vegetales .....		243
V.1	ACEITES UTILIZADOS.....	244
V.2	OBTENCIÓN DE LA MATRIZ DE DATOS ORIGINALES .....	245
V.3	PREPROCESADO DE LOS CROMATOGRAMAS.....	246
V.4	ANÁLISIS JERÁRQUICO DE CONGLOMERADOS.....	255
V.5	ANÁLISIS DE COMPONENTES PRINCIPALES .....	259
V.6	ANÁLISIS DISCRIMINANTE POR MÍNIMOS CUADRADOS PARCIALES (PLS-DA).....	264
CAPÍTULO VI: Cuantificación de aceite de oliva en vinagretas y aliños comerciales usando un modelo PLS desarrollado con mezclas binarias de aceites vegetales .....		277
VI.1	ACEITES UTILIZADOS PARA LA CREACIÓN DEL MODELO .....	279
VI.1.1	Muestras utilizadas como conjunto o set de calibración.....	279
VI.1.2	Muestras utilizadas como conjunto o set de validación.....	281
VI.1.3	Muestras utilizadas como conjunto o set de predicción.....	282
VI.2	DISEÑO PARA LA OBTENCIÓN DE LAS HUELLAS DACTILARES CROMATOGRÁFICAS.....	283

VI.3	PRE-PROCESADO DE LOS DATOS.....	285
VI.4	PRIMER MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES.....	286
VI.5	SEGUNDO MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES.....	294
VI.6	TERCER MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES.....	300
	CONCLUSIONES .....	313
	PERSPECTIVAS FUTURAS.....	317



# **RESUMEN**





## APLICABILIDAD DEL PERFIL DE ESTEROLES PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN ALIMENTOS

En esta memoria se reúnen los resultados obtenidos durante la realización de la Tesis Doctoral "Aplicabilidad del Perfil de Esteroles para la Cuantificación de Aceite de Oliva en Alimentos". Se ha dividido en tres bloques:

**Introducción**, que incluye información bibliográfica sobre los compuestos estudiados (fitoesteroles) y las diferentes técnicas separativas empleadas para su análisis (GC y HPLC con detección FID y UV-vis). Así mismo se describe el estado actual de los estudios quimiométricos realizados sobre la caracterización y autenticación de aceites vegetales.

El segundo bloque denominado **Metodología y Resultados**, consta de los siguientes capítulos:

*Capítulo I:* en este capítulo y tras una revisión bibliográfica, se recogen y resumen diferentes técnicas quimiométricas, así como, las distintas estrategias para procesar los datos experimentales obtenidos, y los programas informáticos más frecuentemente empleados.

*Capítulo II:* aquí se describen las disoluciones, reactivos e instrumentación utilizados a lo largo de todo el trabajo experimental desarrollado.

*Capítulo III:* describe el desarrollo y validación de un método de análisis para la determinación de los 4-desmetilesteroles que habitualmente aparecen en los aceites vegetales basado en la separación off-line de la fracción esterólica mediante cromatografía líquida de alta resolución (HPLC-UV) y determinación de sus trimetilisililéteres por cromatografía de gases (GC-FID).

*Capítulo IV:* en este capítulo se estudian 24 muestras de aceites vegetales y se realiza el análisis, mediante técnicas de reconocimiento de pautas, del perfil de concentraciones de los analitos con objeto de ver, si dicho perfil permite distinguir entre los diferentes tipos de aceites analizados.

*Capítulo V:* bajo un enfoque quimiométrico, se estudia la posibilidad de utilizar el cromatograma de los 4-desmetilesteroles como huella dactilar cromatográfica (fingerprint), para determinar el contenido de aceite de oliva en una mezcla de aceites. Partiendo de diferentes estrategias para el preprocesado de la señal, y utilizando 71 muestras de distintos tipos de aceites vegetales (oliva virgen extra, orujo de oliva, girasol, girasol alto contenido en ácido oleico, colza, canola, soja, maíz, lino, cacahuete, uva, sésamo y semillas), se estudia el posible agrupamiento natural y a partir de los resultados obtenidos, se propone un modelo PLS-DA para la cuantificación del porcentaje de aceite de oliva en una mezcla.

*Capítulo VI:* en este último capítulo se aborda la aplicación de la huella digital de los 4-desmetilesteroles para la cuantificación de aceite de oliva en productos alimenticios elaborados a partir de mezclas binarias de aceites vegetales (vinagretas). El método quimiométrico seleccionado fue la calibración multivariante.

El tercer y último bloque recoge las **Conclusiones** a las que se ha llegado.

# **CONTEXTUALIZACIÓN**



Esta Tesis Doctoral está enmarcada dentro de la línea de investigación "Técnicas Quimiométricas y Cualimétricas en Análisis Químico" perteneciente al Programa Oficial de Posgrado en Química (P34.56.1).

Se ha llevado a cabo en el Departamento de Química Analítica, de la Facultad de Ciencias de la Universidad de Granada dentro del grupo de investigación Análisis en Alimentación y Medio Ambiente (AnAMA, FQM-232) y ha sido realizada, bajo la dirección de los doctores Mercedes Sánchez Viñas, Domingo Gázquez Evangelista y María Gracia Bagur González.

Dentro de la línea de investigación, el trabajo se ha centrado en el estudio quimiométrico de la aplicabilidad del perfil cromatográfico de 4-desmetilesteroles como herramienta para la cuantificación de aceite de oliva en mezclas de aceites vegetales y en productos alimenticios, y está incluida dentro de las actividades y tareas del proyecto QuOLEO, que ha sido financiado con los siguientes fondos públicos:

- Cuantificación y control analítico de aceite de oliva en alimentos. Proyecto de Investigación de Excelencia, Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía (Ref.: P07-FQM-02667).
- Cuantificación de aceite de oliva en alimentos. Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2004-2007, Ministerio de Educación y Ciencia, Gobierno de España (MEC, Ref.: CTQ2006-15066-C02-02/BQU).
- Estudio de viabilidad de metodologías analíticas para la cuantificación de aceite de oliva en alimentos que contengan materias grasas. Convenio especial de colaboración, Consejería de Agricultura y Pesca (CAP), Junta de Andalucía. (Ref.: 9728-2007).



# **OBJETIVOS**





La determinación del contenido de fitoesteroles o esteroides en aceites de oliva y vegetales es uno de los parámetros analíticos que se utilizan para determinar su y establecer la calidad de los mismos. Es una determinación de rutina en laboratorios acreditados para detectar posibles adulteraciones ya que el aceite de oliva presenta efectos beneficiosos para la salud además de sus cualidades organolépticas. El precio del aceite de oliva a fecha de hoy es superior al resto de aceites que existen comercialmente, por este motivo en ocasiones se intenta mezclar con otros tipos de aceites de inferior calidad y precio.

Sobre la base del problema analítico de control de la calidad alimentaria, el principal objetivo del trabajo a realizar durante el periodo de elaboración de esta Tesis Doctoral, es estudiar la aplicabilidad del perfil de esteroides (compuestos mayoritarios del insaponificable) en la cuantificación del aceite de oliva presente en alimentos.

Para ello, se pretende proponer una metodología analíticas, genérica y fácilmente transferible, de uso rutinario en laboratorios de control de la calidad del aceite de oliva. Las técnicas que se aplicarán en los estudios experimentales, serán fundamentalmente la cromatografía líquida de alta resolución (HPLC) con detección ultravioleta-visible (Uv-Vis) y cromatografía de gases con detector de ionización de llama (GC-FID).

Se estudiarán los perfiles de concentración con objeto de verificar si con el uso de herramientas quimiométricas adecuadas dicho perfil permite discriminar entre los diferentes tipos de aceites vegetales.

Dada la naturaleza y complejidad de las muestras a analizar, uno de los objetivos principales se puede particularizar considerando la posibilidad de desarrollar un método que conlleve la disminución del tiempo de análisis y el consumo de reactivos.

Los diferentes objetivos planteados inicialmente para llevar a cabo esta Tesis Doctoral, fueron los siguientes:

1. Trabajar bajo un sistema de gestión de la calidad de la investigación con objeto de asegurar tanto ésta, como la transferibilidad de las metodologías analíticas desarrolladas.
2. Caracterizar la fracción esteróica del aceite de oliva y resto de aceites vegetales mediante la combinación de técnicas cromatográficas.

## OBJETIVOS

3. Mejorar los métodos de aislamiento y depuración de esteroides presentes en la fracción insaponificable en aceites y grasas.
4. Establecer el cálculo de incertidumbre aplicable al método analítico propuesto para la determinación de compuestos minoritarios, esteroides, en aceites vegetales.
5. Aplicar las huellas dactilares cromatográficas de fitoesteroides para identificar y autenticar aceites vegetales con ayuda de herramientas quimiométricas de reconocimiento de pautas no supervisadas y supervisadas.
6. Desarrollar un método de cuantificación del porcentaje de aceite de oliva en mezclas binarias con otros aceites vegetales (colza, cáñola, soja, maíz, girasol, semillas, etc) mediante las técnicas de cromatografía líquida de alta resolución (HPLC) y cromatografía de gases (GC) acopladas de modo off-line.
7. Aplicar las nuevas metodologías analíticas desarrolladas al análisis de muestras de aceites y alimentos, de diverso origen y procedencia, que contengan aceite de oliva: Verificar la autenticidad de la proporción de aceite de oliva declarada en la etiqueta de diferentes alimentos: salsas tipo vinagretas y aliños.
8. Finalmente, una vez alcanzados los objetivos anteriores se procederá a la elaboración de un procedimiento normalizado de trabajo que recoja todas las operaciones analíticas empleadas y sea de fácil transferencia y aplicación en un laboratorio de análisis de rutina.

# **INTRODUCCIÓN**



Los esteroides vegetales, también denominados "*fitoesteroides*", son esteroides derivados de plantas con estructura y propiedades similares a las presentadas por el colesterol en los vertebrados [1]. Este hecho se puede atribuir a que químicamente presentan pequeñas diferencias estructurales con respecto a este compuesto [2], pudiendo por ello ejercer un papel equivalente en la membrana de las células de organismos superiores al actuar como: (i) componentes estructurales de las membranas vegetales [3], (ii) intermediarios para la biosíntesis de productos vegetales tanto primarios como la celulosa, o secundarios como las hormonas esteroideas, los alcaloides, etc.

Desde el punto de vista de la salud humana, estudios médicos han demostrado que la ingesta diaria de fitosteroides dentro de la dieta normal puede reducir los niveles de colesterol en sangre y suero. Adicionalmente, se ha sugerido que estos compuestos pueden presentar propiedades antibacterianas, antifúngicas, antiulcerativas, antitumorales antiinflamatorias y antioxidativas.

Prácticamente se encuentran presentes en todos los alimentos vegetales, aunque aquellos en los que se encuentran en mayor concentración son los aceites, especialmente los de maíz, girasol, soja y colza/cánola (que contienen entre un 0.1 y un 0.8 %), siendo una excepción el aceite de palma, que pierde la mayoría de estos compuestos tras el proceso de refinado. También se encuentran en legumbres (0,2%) y, en menor cantidad, en frutos secos, pan y vegetales, por lo que se puede decir que, exceptuando los carbohidratos altamente refinados y los productos animales, casi todos los otros alimentos contribuyen de manera apreciable a la ingesta de esteroides vegetales.

En el caso de los aceites vegetales, el contenido de fitoesteroides depende de factores como el tipo de cultivo, año de la cosecha, grado de maduración del fruto o antigüedad de la semilla, tiempo de almacenamiento, métodos de extracción y refinado del aceite, entre otros, por lo que el análisis cuantitativo de estos compuestos, especialmente en el caso del aceite de oliva, presenta gran importancia para establecer su origen, autenticidad y calidad del aceite.

### 1. PROPIEDADES FÍSICAS Y QUÍMICAS.

Los fitoesteroles presentan alto punto de ebullición, son incoloros, algo inertes, solubles en disolventes orgánicos poco polares (cloroformo, etc.), menos solubles en alcoholes de bajo peso molecular, y funden sin descomponerse, tanto en su forma libre como esterificada.

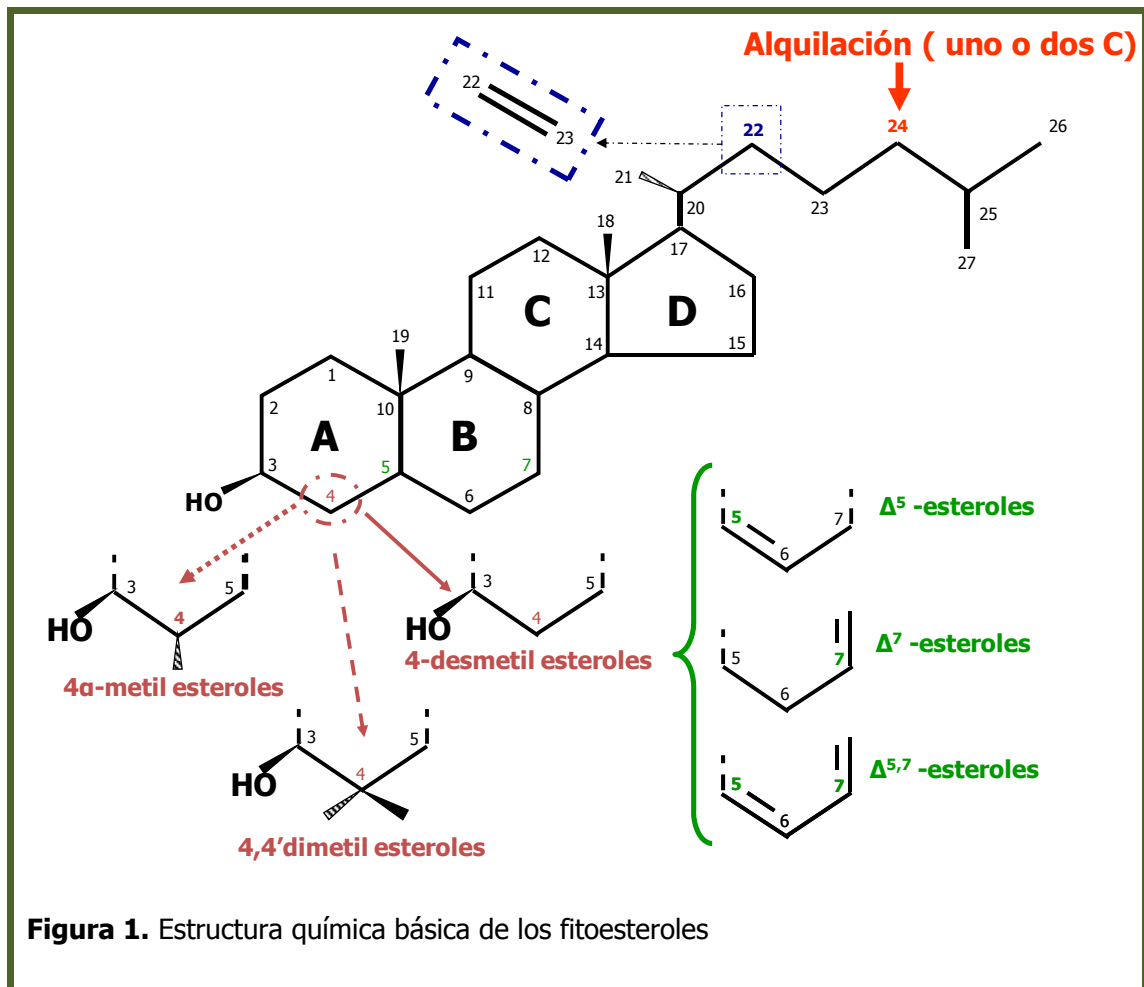
Químicamente, tal y como se muestra en la Figura 1., las moléculas de los fitoesteroles derivan del ciclopentano[a]perhidrofenantreno en el que, el carbono C-27 presenta una cadena lateral flexible. Estructuralmente, los cuatro anillos A,B,C y D (numerados de acuerdo con las recomendaciones establecidas en 1989 por la Unión Internacional de Química Pura y Aplicada (IUPAC) [4]) presentan una configuración *trans* que les otorga una geometría planar  $\alpha$  en la que en el grupo hidroxilo del carbono C-3 en la cadena lateral en C-27, y en los grupos metilo de los carbonos C-18 y C-19, se favorece una disposición angular sobre el plano que también les otorga lo que se conoce como estereoquímica  $\beta$ . Además, debido a que desde un punto de vista conformacional, el carbono C-20 presenta preferentemente la conformación R, los fitosteroles pueden originar superficies planas a ambos lados de sus moléculas. Este último hecho es el que permite biológicamente las múltiples interacciones hidrofóbicas que se producen entre el núcleo rígido de estos compuestos y la matriz de la membrana celular.

Entre las características de los fitoesteroles se pueden destacar las siguientes:

- a) La mayoría de ellos tienen un enlace  $\Delta^5$ , y uno o dos sustituyentes en el carbono C-24 de la cadena lateral, los cuales se introducen mediante reacciones de trans-metilación.
- b) También presentan un doble enlace *trans* en la posición  $\Delta^{22}$  de la cadena lateral.

Como también se puede apreciar en dicha figura, atendiendo a la presencia o ausencia de los grupos metilos en la posición del carbono C-4 perteneciente al anillo A, los fitoesteroles se pueden dividir estructural y biosintéticamente en tres clases:

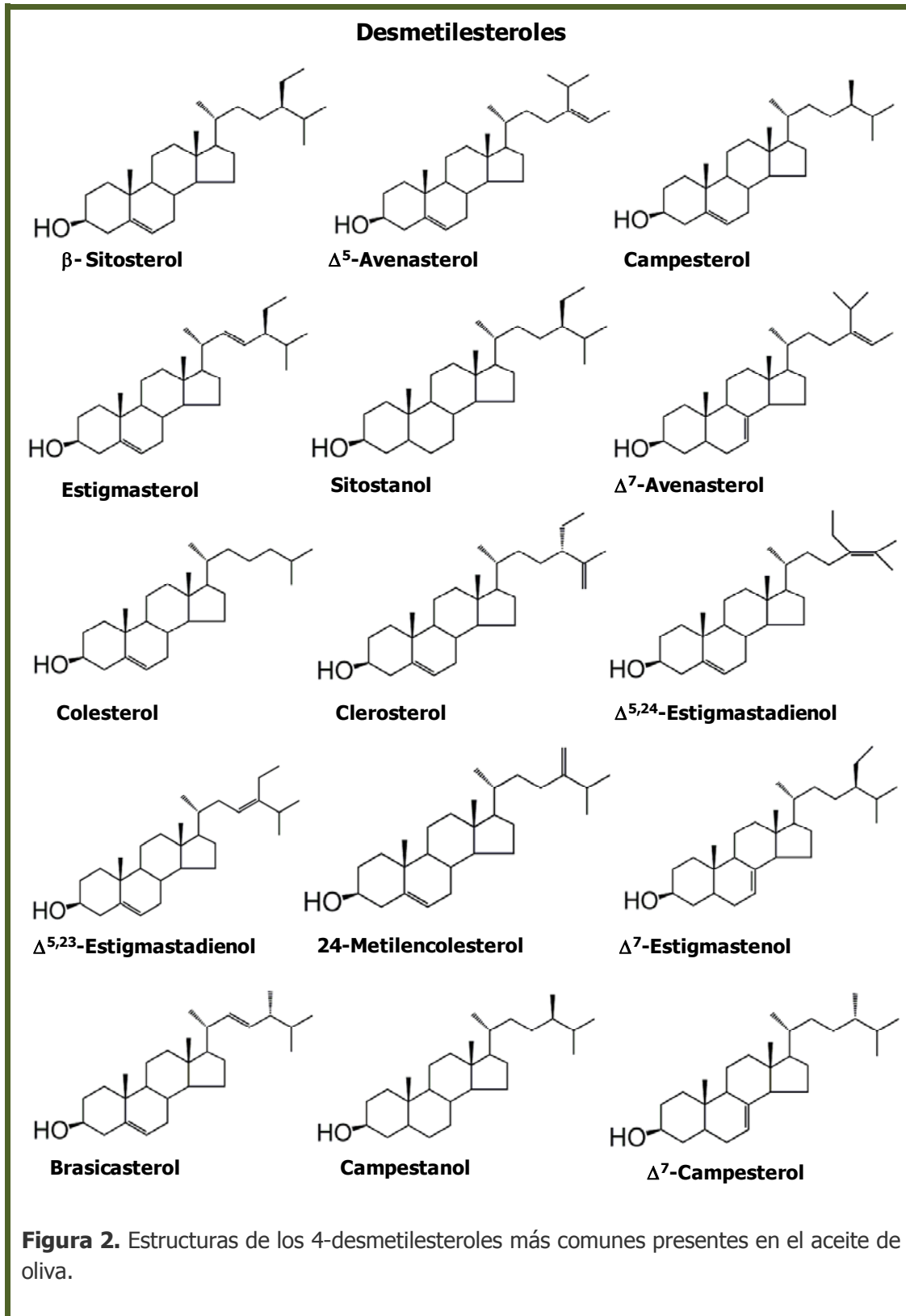
- i) 4-desmetilesteroles (sin grupo metilo).
- ii) 4a-monometilesteroles (un grupo metilo).
- iii) 4,4'-dimetilesteroles (dos grupos metilo). Dentro de éstos también se incluyen los alcoholes triterpénicos.



En el caso de los primeros (4-desmetilesteroides), y atendiendo a la posición y número de dobles enlaces en el anillo B, a su vez se clasifican en  $\Delta^5$ -esteroides,  $\Delta^7$ -esteroides y  $\Delta^{5,7}$ -esteroides.

La Figura 2 muestra las fórmulas estructurales de los 4-desmetilesteroides y los estanoles (productos correspondientes a la hidrogenación del doble enlace del carbono C5 de algunos fitoesteroides) estudiados a lo largo de esta memoria, los cuales a su vez se corresponden con los más comúnmente presentes en el aceite de oliva (una de las matrices estudiadas). La Tabla I [5] recoge los nombres sistemáticos y comunes de los 4-desmetilesteroides más frecuentes en aceites vegetales.





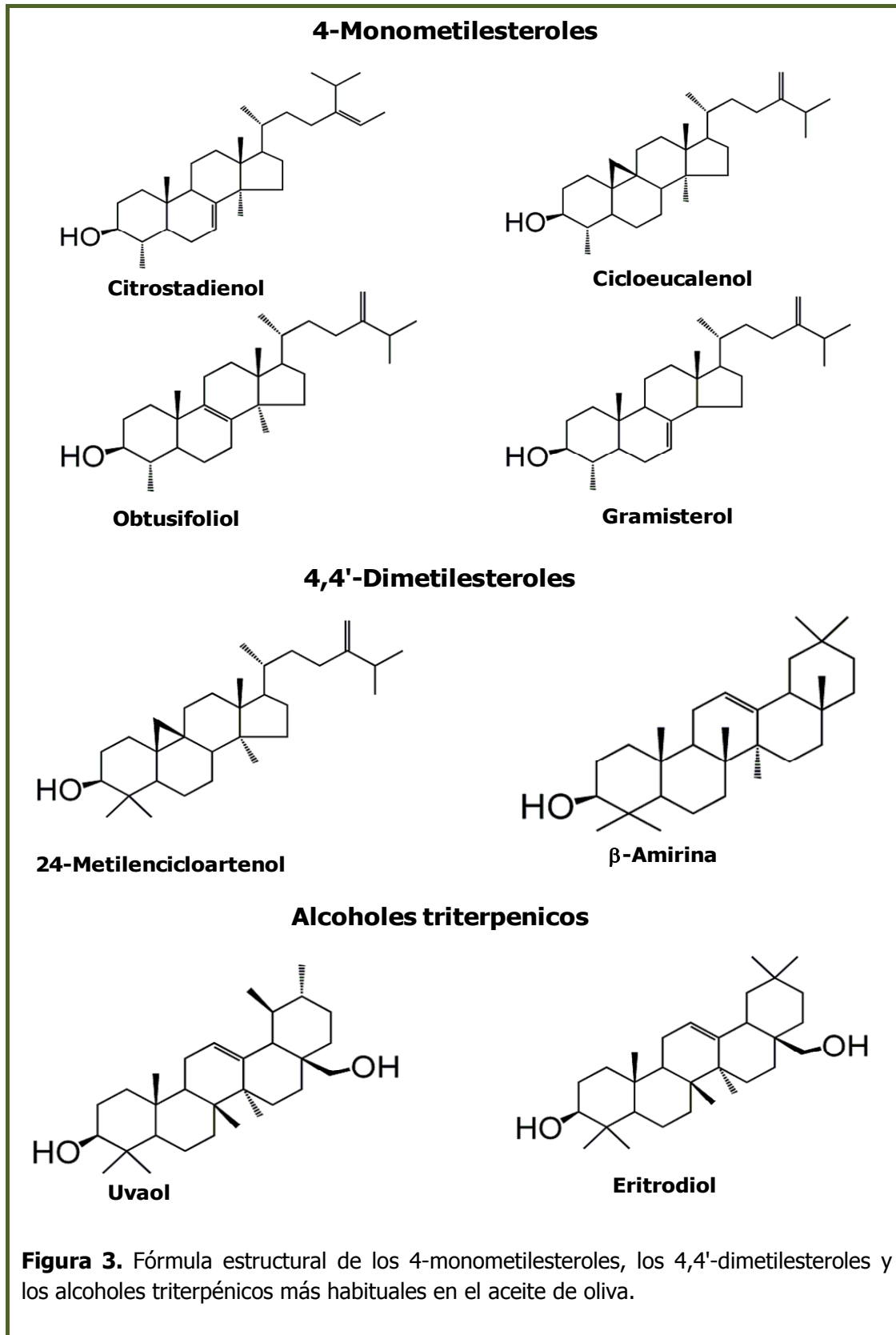
**Figura 2.** Estructuras de los 4-desmetilesteroles más comunes presentes en el aceite de oliva.

**Tabla I.** Nombres sistemáticos y comunes de los 4-desmetilesteroles analizados en esta memoria.

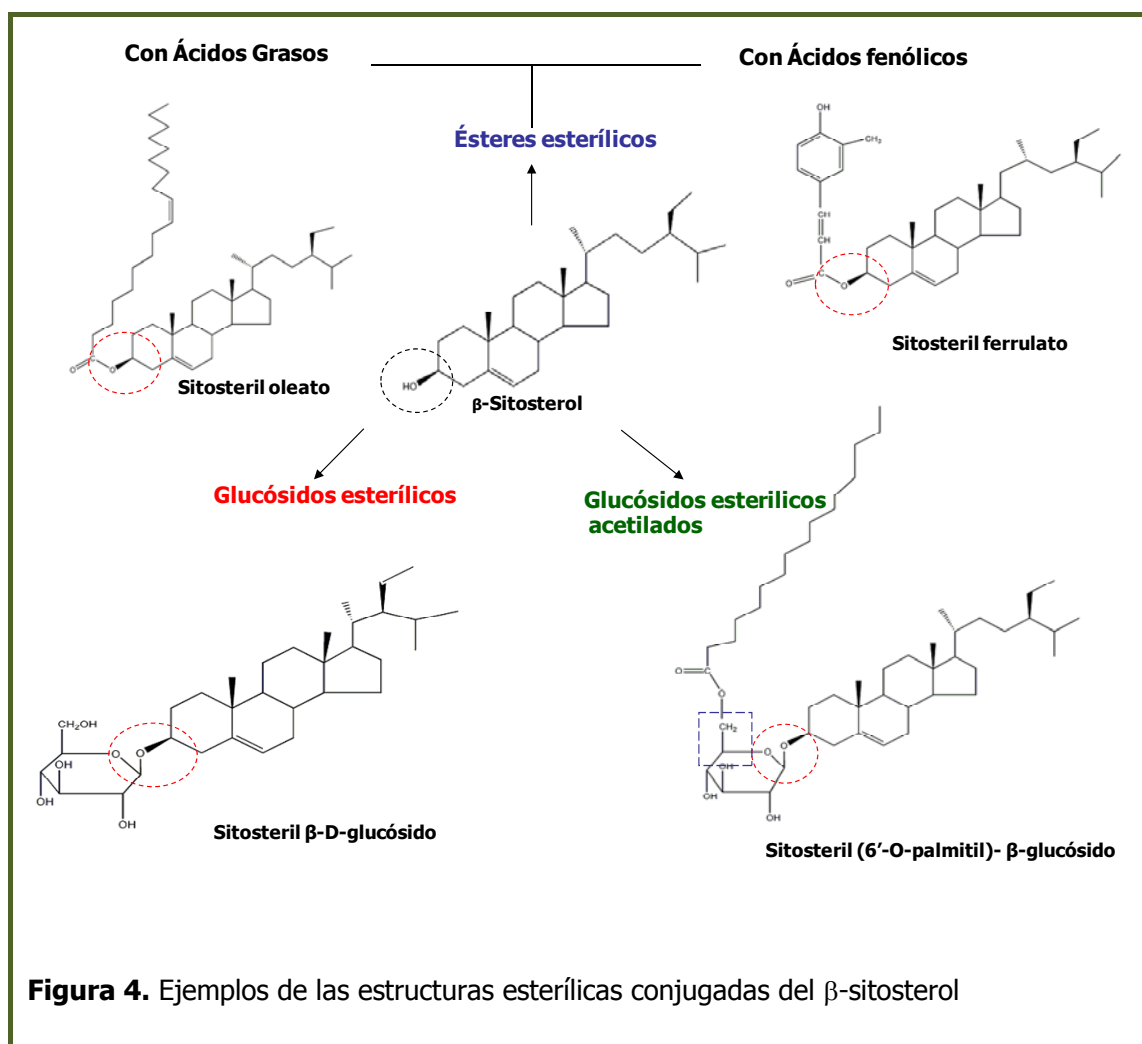
<b><u>Nombre sistemático</u></b>	<b><u>Nombre común</u></b>
$\Delta^5$ -colestén-3 $\beta$ -ol	Colesterol
(24S)-24-metil- $\Delta^{5,22}$ -colestadién-3 $\beta$ -ol	Brasicasterol
24-metilen- $\Delta^{5,24}$ -colestén-3 $\beta$ -ol	24-metilencolesterol
(24R)-24-metil- $\Delta^5$ -colestén-3 $\beta$ -ol	Campesterol
(24R)-24-metil-colestan-3 $\Delta^5$ -colestén-3 $\beta$ -ol	Campestanol
(24R)-24-metil- $\Delta^{5,22}$ -colestadién-3 $\beta$ -ol	Estigmasterol
(24R)-24-metil- $\Delta^7$ -colestan-3 $\beta$ -ol	$\Delta^7$ -campesterol
(24R,S)-24-etil- $\Delta^{5,23}$ -colestadién-3 $\beta$ -ol	$\Delta^{5,23}$ -estigmastadiénol
(24S)-24-etil- $\Delta^{5,25}$ -colestadién-3 $\beta$ -ol	Clerosterol
(24R)-24-etil- $\Delta^5$ -colestén-3 $\beta$ -ol	$\beta$ -sitosterol
(24R)-24-etil- colestan-3 $\beta$ -ol	Sitostanol (Estigmastanol)
(24Z)-24-etilen- $\Delta^5$ -colestén-3 $\beta$ -ol	$\Delta^5$ -avenasterol
(24R,S)-24-etil- $\Delta^{5,24}$ -colestadién-3 $\beta$ -ol	$\Delta^{5,24}$ -estigmastadiénol
(24R,S)-24-etil- $\Delta^7$ colestén-3 $\beta$ -ol	$\Delta^7$ -estigmastenol
(24Z)-24-etilen- $\Delta^5$ -colestén-3 $\beta$ -ol	$\Delta^7$ -avenasterol

La mayor parte de los 4-desmetilesteroles son sólidos cristalinos incoloros, por ejemplo,  $\beta$ -sitosterol, campesterol y estigmasterol tienen puntos de fusión de 140, 157/158 y 170 °C, respectivamente [6], y son más hidrofóbicos que el colesterol (en general, cuanto más grande es la cadena lateral, más hidrofóbico es el compuesto), aunque los dobles enlaces de las cadenas laterales hacen que aumente su carácter hidrofílico.

La Figura 3 muestra las fórmulas estructurales de los 4-monometilesteroles, los 4,4'-dimetilesteroles y los alcoholes triterpénicos que aparecen frecuentemente en el aceite de oliva.



En cuanto a su reactividad, y como ya se ha indicado previamente, al hidrogenarse forman los estanoles o fitoestanoles. Además, el grupo 3-hidroxilo de los fitoesteres libres puede ser esterificado por un ácido graso [7] o un ácido fenólico para dar esteril-ésteres, o puede, a través de un enlace  $\beta$  en la posición 1 de un hidrato de carbono, formar, bien esteril glicósidos o bien, glicósidos esteril acilados. En estos últimos, la posición 6 de los hidratos de carbono se encuentra esterificada con un ácido graso de cadena larga [8]. En la Figura 4. se recogen las estructuras correspondientes a los productos finales de las reacciones anteriormente mencionadas para el  $\beta$ -sitosterol.



Desde el punto de vista de las propiedades físicas cabe indicar que, aunque los esterolés libres y esteril ésteres son solubles en disolventes no polares como el hexano, es necesario el uso de modificadores polares para solubilizar los glicósidos esterilicos.

## Introducción

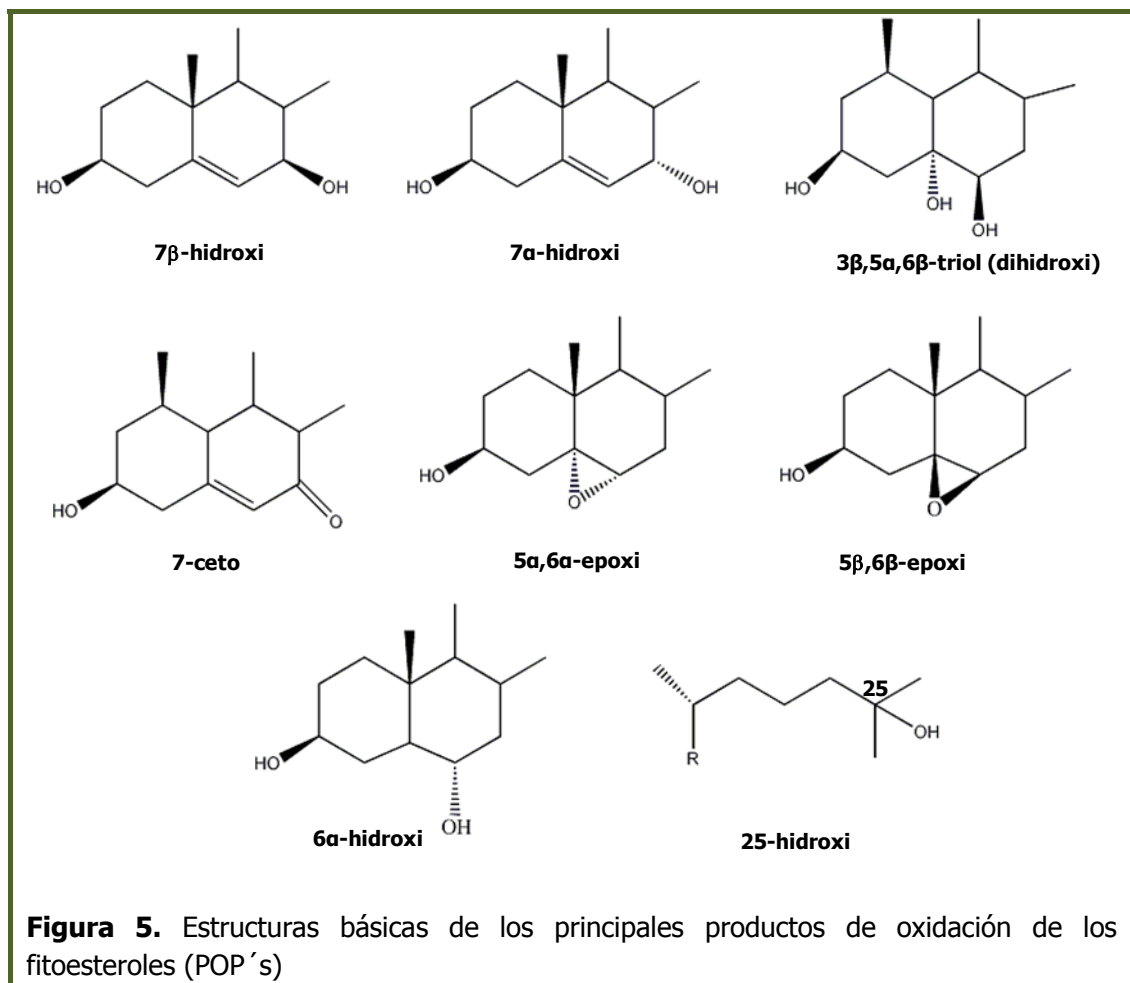
En relación con la estabilidad de los fitoesteroles, la reacción más importante a la que se ven sometidos estos compuestos es la oxidación, en la que el producto final se denomina genéricamente "oxifitoesterol". Aunque se ha sugerido [9] que algunos  $\Delta^5$ -esteroles siguen el mismo esquema de oxidación que el colesterol, se conoce poco acerca de los productos originados a partir de los fitoesteroles, los cuales también se suelen citar de forma abreviada como POPs (productos de oxidación de fitoesteroles). Diversos estudios indican que algunos oxifitoesteroles pueden ser carcinógenos, mutágenos o producir efectos tóxicos en los cultivos celulares [10].

En general, la formación de radicales en presencia de calor, luz, contaminantes metálicos, oxígeno, especies reactivas de oxígeno y enzimas oxidativas, facilita la oxidación de los fitoesteroles a través de un ataque al doble enlace en la estructura del anillo, dando lugar a una reacción autocatalítica en cadena de radicales libres que sigue la misma química que la oxidación de otros lípidos insaturados.

La Figura 5 recoge las estructuras básicas de los POPs, de los cuales, el principal es el 7-hidroxiperóxido, que se forma después de la eliminación de un átomo de hidrógeno alílico en el átomo de carbono C-7, posteriormente, el hidroperóxido alílico sufre una epimerización de la que se derivan los productos de oxidación más importantes como los 7 $\beta$ - y 7 $\alpha$ -hidroxi, y los compuestos 7-ceto.

Otros productos de oxidación importantes son los compuestos 5 $\alpha$ -, 6 $\alpha$ -epoxi, 5 $\beta$ -, 6 $\beta$ -epoxi y los 3 $\beta$ -, 5 $\alpha$ -, 6 $\beta$ -trioles. Dependiendo de la geometría y de las condiciones en las que se produce la oxidación, el oxígeno también puede atacar a los carbonos terciarios de la cadena lateral, originando pequeñas cantidades de compuestos 20 y 25-hidroxílicos.

Finalmente, la Tabla II recoge los nombres sistemático y común de algunos de los POPs [11].



**Tabla II.** Nomenclatura de algunos productos de oxidación de fitoesteros.

<b>Nombre sistemático</b>	<b>Nombre común</b>
(24S)-Etilcolest-5-en-3β, 7α-diol	7α-Hidroxibrasicasterol
(24R)-Metilcolest-5-en-3β, 7α-diol	7α-Hidroxicampesterol
(24S)-Etilcolest-5,22-dien-3β, 7α-diol	7α-Hidroxiestigmasterol
(24R)-Etilcolest-5-en-3β, 7α-diol	7α-Hidroxisitosterol
(24S)-Metilcolest-5,22-dien-3β,7β-diol	7β-Hidroxibrasicasterol
(24R)-Metilcolest-5-en-3β, 7β-diol	7β-Hidroxicampesterol
(24R)-Etilcolest-5-en-3β, 7β-diol	7β-Hidroxisitosterol

**Tabla II (Cont.).** Nomenclatura de algunos productos de oxidación de fitoesteroles.

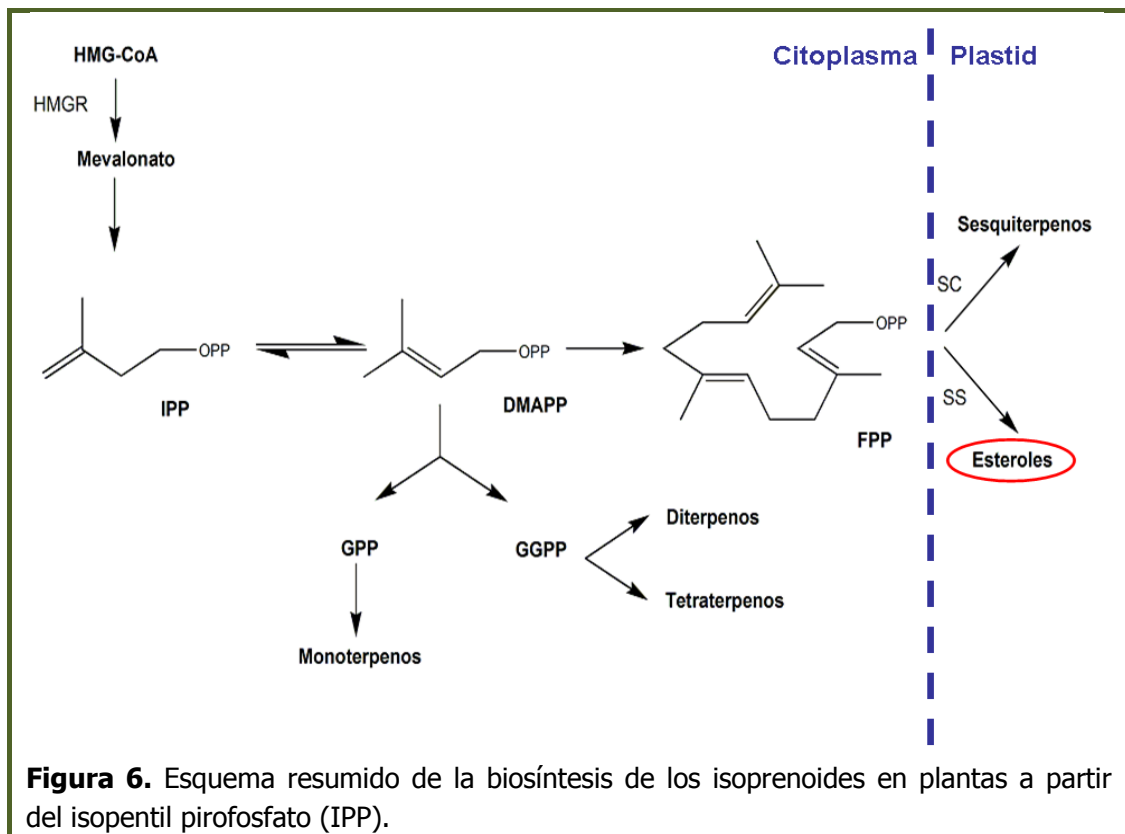
<b>Nombre sistemático</b>	<b>Nombre común</b>
(24R)-Metilcolestan-3 $\beta$ ,6 $\beta$ -diol	6 $\beta$ -Hidroxicampestanol
(24S)-Metilcolest-22-en-3 $\beta$ , 5 $\alpha$ , 6 $\beta$ -triol	Brasicastentriol (Dihidroxibrasicasterol)
(24R)-Metilcolestan-3 $\beta$ , 5 $\alpha$ , 6 $\beta$ -triol	Campestanotriol (Dihidroxicampesterol)
(24R)-Etilcolestan-3 $\beta$ , 5 $\alpha$ , 6 $\beta$ -triol	Sitostantriol (Dihidroxisitosterol)
(24R)-Etilcolestan-3 $\beta$ ,6 $\beta$ -diol	6 $\beta$ -Hidroxisitostanol
(24S)-5 $\beta$ ,6 $\beta$ -epoxi-24-metilcolest-22-en-3 $\beta$ -ol	5 $\alpha$ ,6 $\beta$ -Epoxibrasicasterol
(24R)-5 $\alpha$ ,6 $\alpha$ -epoxi-24-metilcolestan-3 $\beta$ -ol	5 $\alpha$ ,6 $\alpha$ -Epoxicampesterol
(24R)-5 $\beta$ ,6 $\beta$ -epoxi-24-metilcolestan-3 $\beta$ -ol	5 $\beta$ ,6 $\beta$ -Epoxicampesterol
(24R)-5 $\beta$ ,6 $\beta$ -epoxi-24-etilcolestan-3 $\beta$ -ol	5 $\beta$ ,6 $\beta$ -Epoxisitosterol
(24S)-Metilcolest-5,22-dien-3 $\beta$ -ol-7-ona	7-Cetobasicasterol
(24R)-Metilcolest-5-en-3 $\beta$ -ol-7-ona	7-Cetocampesterol
(24S)-Etilcolest-5,22-dien-3 $\beta$ -ol-7-ona	7-Cetoestigmasterol

## 2. BIOSÍNTESIS DE LOS FITOESTEROLES

De forma natural, los fitoesteroles se obtienen a través de la formación de un ciclo entre el (3S) 2,3-epóxido de escualeno [12,13] y el esqueleto tetracíclico del ciclopentanoperhidrofenantreno, procediendo por tanto de la ruta biosintética de los isoprenoides.

Tal y como se recoge de forma esquematizada en la Figura 6., el paso principal en la síntesis de los sesquiterpenos y triterpenos (grupo al que pertenecen los fitoesteroles) es la conversión del isopentil pirofosfato (IPP) en dimetilalquil difosfato (DMAPP) como primera etapa de elongación, a través de una prenil transferasa -la farnesil-pirofosfato-sintasa (FPP-sintasa)- que convierte el metabolito en farnesil pirofosfato (FPP). A nivel celular, la enzima encargada de regular la dirección dentro de la ruta de los esteroides en el citoplasma, es la escualeno sintetasa.

Cuando las plantas tienen aparato fotosintético, las etapas enzimáticas convierten el óxido de escualeno en cicloartenol, mientras que en organismos no-fotosintéticos, los hongos transforman el óxido de escualeno en lanosterol y finalmente en ergosterol. Ambos, cicloartenol y lanosterol, se pueden convertir en  $\Delta^5$ -24-alquilesteroles como productos finales.



En el caso de las plantas superiores, los principales productos finales de la síntesis tienen un grupo 3 $\beta$ -hidroxilo libre, no obstante, también se pueden encontrar fitoesteros esterificados con ácidos grasos de cadena larga o enlazados a glúcidos simples, generalmente a través de uniones  $\beta$  con el carbono situado en la posición 1 del monosacárido. Dependiendo de la naturaleza de estos últimos se pueden formar también acilesteril-glicósidos por esterificación en la posición 6 de la molécula de glúcido (generalmente glucosa).

De forma generalizada, se puede decir que, la biosíntesis de los fitoesteros en las plantas está formada por una secuencia de más de 30 reacciones catalizadas por enzimas presentes en las membranas de las mismas, si bien ni todas las etapas biosintéticas ni las reguladoras están perfectamente descritas y estudiadas. En el caso particular de la regulación de la biosíntesis de los fitoesteros, la mayoría de los estudios se han focalizado en el papel jugado por la 3-hidroximetil-3-glutaril coenzima A reductasa (HMGR), en la ruta del pre-escualeno para producir mevalonato (producto intermedio que conduce a la formación de IPP).

Atendiendo a la cadena alimentaria, la principal ruta biosintética lleva a un perfil de fitoesteros constituido por una mezcla de estigmasterol, sitosterol y campesterol.



### 3. FUNCIONES BIOLÓGICAS Y FISIOLÓGICAS

Debido a su carácter anfílico, los fitoesteroles se encuentran a nivel celular formando parte de las membranas, principalmente en el plasma de la membrana celular, la membrana externa de la mitocondria y en el retículo endoplasmático. Su función principal es regular la fluidez de estas, jugando un papel importante en su adaptación en función de la temperatura.

En el caso de los fitoesteroles libres, el grupo hidroxilo libre es el responsable de las interacciones específicas con fosfolípidos y proteínas en la membrana. En el caso de los esteril glicósidos y los esteril glicósidos acetilados, la orientación del grupo carbohidrato hacia la región acuosa permite su presencia junto con los esteroides libres en varias membranas.

Los fitoesteroles típicos de las plantas (sitosterol, estigmasterol y campesterol), actúan en la membrana para limitar el movimiento de las cadenas aciladas de los ácidos grasos y aunque como se ha dicho antes, todos los fitoesteroles son capaces de regular la fluidez de la membrana, la eficiencia depende del tipo de esteroides, de modo que  $\beta$ -sitosterol y campesterol son los más eficientes, comparándolos con el estigmasterol, debido a que éste tiene el doble enlace en el átomo de carbono C-22 orientado en la posición trans, reduciendo tal efecto.

Los fitoesteroles también juegan un papel importante en la diferenciación y proliferación celular, siendo su acumulación en semillas y aceites una reserva para el crecimiento de nuevas células y brotes.

Dada su similitud con el colesterol, determinadas moléculas de fitoesteroides pueden también participar en algunos acontecimientos de transducción de señales activando la  $\text{Na}^+/\text{K}^+$ -ATPasa.

Fisiológicamente, hay evidencias de que los fitoesteroides pueden actuar mostrando un efecto hipocolesterolemico importante reduciendo, tanto las concentraciones de colesterol total como las de colesterol LDL y en menor medida, mostrando una actividad anticancerígena frente a determinados tipos de células tumorales.

#### a) Efecto hipocolesterolemico de los fitoesteroides

Generalmente, se acepta que los fitoesteroides disminuyen la concentración de colesterol circulante en el organismo, suprimiendo su absorción intestinal, debido a la alta afinidad de dichos compuestos por las micelas en comparación con el colesterol. Como resultado, se produce una menor absorción de este último por parte de las lipoproteínas sintetizadas en el epitelio del intestino (quilomicrones).

El efecto reductor del contenido de colesterol en sangre debido a los fitoesteroles, comenzó a describirse hace cincuenta y nueve años por Pollack [14], quien demostró que la ingesta de grasas poliinsaturadas daba lugar a hipocolesterolemia. Por esa razón, con objeto de controlar la colesterolemia desde mediados de los años noventa se comercializan en diferentes países, alimentos enriquecidos con fitoesteroles, tales como margarinas funcionales, yogures, quesos, aperitivos, aderezos para las ensaladas, etc., [1]. Desde entonces, varios estudios clínicos han demostrado que la presencia de los 4-desmetilesteroles o estanoles en sangre, en cualquiera de sus dos formas, libre o esterificada, reduce el contenido total de colesterol y los niveles de colesterol-LDL, sin embargo, el nivel de protección que ejercen las lipoproteínas de alta densidad (colesterol-HDL), no disminuye por el consumo de fitoesteroles en la dieta.

La acción reductora del contenido en colesterol es más acusada con los 4-desmetilesteroles y estanoles, y se ha demostrado que los 4,4'-dimetilesteroles tienen poca o ninguna actividad en humanos. Los efectos específicos de los 4-monometilesteroles y los glicósidos de esteroides no son conocidos.

En un estudio con varios productos alimenticios enriquecidos, Katan y colaboradores [15] encontraron que con la dosis diaria óptima de fitoesteroles o estanoles de 2g/día, se reducía un 10% el nivel del LDL-Colesterol, mientras que una dosis más alta proporcionaba sólo un pequeño efecto benéfico adicional.

### b) Actividad anticancerígena de los fitoesteroides

Hoy en día, hay una clara evidencia de que los esteroides vegetales pueden ser aplicados en forma de suplementos para combatir enfermedades mortales en las que el sistema inmune necesita maduración. Su importancia en la lucha contra el cáncer es aún limitada y necesita más investigación, aunque algunos datos "*in vitro*" e "*in vivo*" sugieren que pueden ser eficaces, no solo en la prevención de la enfermedad, sino también en su tratamiento [16].

Diversas evidencias epidemiológicas y experimentales sugieren, que estos compuestos podrían ofrecer protección frente a los tipos más comunes de cáncer que se dan en las sociedades occidentales, como los de colon, pulmón, estómago, mama y próstata [17,18,19]. El efecto protector se basaría en una serie de mecanismos que incluyen efectos sobre: (i) la estructura y función de la membrana en las células tumorales y del huésped, (ii) las vías de transducción de señales que regulan el crecimiento del tumor y su apoptosis, (iii) la función inmunitaria del huésped, y (iv) el metabolismo del colesterol.

Así mismo, existen evidencias de que el consumo de fitoesteroides también puede

## Introducción

aumentar la actividad de enzimas antioxidantes y con ello, reducir el estrés oxidativo, relacionado con varias enfermedades, entre ellas el cáncer [20].

Por otro lado, a partir de la década de los 50, se desarrollaron un gran número de estudios clínicos relacionados con la prevención y reducción de colesterol en plasma, tanto en animales como en humanos. Sin embargo, su interés disminuyó posteriormente debido a que se descubrió, que era posible desarrollar fitosterolemia como consecuencia de un aumento del contenido de fitoesteroles en suero, especialmente el campesterol.

La fitosterolemia, también llamada sitosterolemia, es una enfermedad extremadamente rara del metabolismo de los lípidos, caracterizada por la presencia de un contenido elevado de fitoesteroles y estanoles en el plasma y tejidos, lo que conlleva al desarrollo de aterosclerosis y enfermedades coronarias a temprana edad. Dicha presencia es exclusivamente de origen dietético, ya que el hombre es incapaz de sintetizarlos. Los niveles plasmáticos de fitoesteroles en las personas que padecen esta enfermedad están entre un 7% y un 16% de la concentración total de esteroides en plasma [21].

Hay estudios [22,23], que sugieren que el consumo de sitostanol en pacientes con sitosterolemia podría ser una forma efectiva de reducir sus niveles de colesterol en suero y de fitoesteroles libres [24].

Finalmente, cabe indicar que un efecto no deseado de los fitoesteroles es su interferencia en la absorción de carotenoides y, por tanto, en la disminución de los niveles de los mismos en sangre. A su vez, no hay evidencias totalmente claras sobre los efectos negativos de las POPs sobre la salud humana.

## 4. FUENTES DE ESTEROLES

Como se ha comentado al inicio del capítulo, en la Tabla III se puede observar que aunque prácticamente todos los alimentos vegetales representativos contienen cantidades apreciables de fitoesteroles, la fuente principal la constituyen los aceites, seguidos de los cereales y los frutos secos y derivados. Por el contrario, las frutas y las hortalizas no son en general una buena fuente, y su contenido difiere dependiendo de las variedades y también de las condiciones de cultivo.

En general, en la mayoría de los alimentos los fitoesteroles están presentes en forma esterificada, mayoritariamente con ácidos grasos y como ésteres de ferulato [25], a diferencia de los esteroides glicosilados que son un componente minoritario

en los alimentos [26]. De hecho, los primeros constituyen cerca del 50% del total de esteroides en algunos alimentos, como el aceite de maíz [27]; no obstante, existen excepciones como las patatas en las que, el 82% de los fitoesteroides que contienen, son esteroides glicosilados.

**Tabla III.** Contenido de esteroides vegetales en alimentos representativos [28].

<b><u>Alimento</u></b>	<b><u>Esteroides vegetales</u></b> (mg/100 g porción comestible)
Aceite de maíz	952
Aceite de girasol	725
Aceite de semilla de soja	221
Aceite de oliva	176
Almendras	143
Alubias	76
Maíz	70
Trigo	69
Aceite de palma	49
Lechuga	38
Plátano	16
Manzana	12
Tomate	7

#### **4.1. Esteroides y aceites vegetales: El aceite de oliva.**

El hecho de que en los aceites vegetales comestibles, los fitoesteroides se puedan encontrar tanto en su forma libre como esterificada y en diferente proporción, provoca que los aceites puedan presentar efectos fisiológicos diferentes [29]. La Tabla IV recoge el contenido en 4-desmetilesteroides (libres y esterificados) de los aceites comestibles más utilizados [30, 31, 32]. A su vez, las Tablas V y VI [33] recogen el contenido en 4-monometilesteroides, y 4,4'-dimetilesteroides en diferentes tipos de aceites.

**Tabla IV.** Contenido de los 4-desmetilsteroles libres y esterificados (mg/100 g) en aceites vegetales.

Aceites	$\beta$ -Sitosterol	$\Delta^3$ -Avenasterol	Campesterol	Estigmasterol	Sitostanol	Brassicasterol	Campestanol	Colesterol
Maíz	L <sup>(1)</sup>	172.2	10.1	53.2	18.2	4.5	4.1	0.7
		242.1	===	70.8	22.6	===	===	===
E <sup>(2)</sup>		690.0	22.0	158.0	76.0	===	===	===
		290.8	25.0	70.9	34.1	9.3	===	0.8
		319.1	9.17	77.3	29.5	===	===	===
		28.7	1.0	10.7	6.5	<0.5	===	1.0
Palma	L	26.1	===	6.2	2.5	===	===	===
		30.35	trazas	10.0-18.0	6.0-7.0	===	===	===
E		10.7	0.6	4.1	2.2	<0.5	===	0.5
		20.7	===	10.0	6.6	===	===	===
L		47.2	3.0	8.3	8.1	1.6	0.9	<0.5
		100.0	===	17.3	16.1	===	===	===
E		67.8	9.9	15.4	3.9	0.9	<0.5	<0.5
		56.6	===	12.4	4.2	===	===	===
L		86.4	14.7	18.4	12.1	6.9	1.7	===
		34.3	3.6	9.8	2.1	2.7	0.9	===
L		145.5	3.6	56.2	7.6	2.3	3.7	1.0
		235.9	14.1	108.2	8.8	3.9	2.3	0.9
L		76.5	===	4.9	===	===	===	===
		54.8	===	2.7	===	===	===	===
L		101.4	0.5	7.5	2.5	1.8	0.7	===
		218.7	7.7	10.5	0.5	1.9	===	===
F		221.0	22.8	36.3	26.9	1.1	1.1	===
		114.4	30.6	40.1	9.8	0.9	===	===

(1) Libres  
(2) Esterificados

**Tabla IV (Cont.).** Contenido de los 4-desmetilsteroles libres y esterificados (mg/100 g) en aceites vegetales.

Aceites	$\beta$ -Sitosterol	$\Delta^5$ -Avenasterol	Campesterol	Estigmasterol	Sitosterol	Brassicasterol	Campestanol	Colesterol
Oliva	L(1) E(2)	6.7 9.4	2.2 2.1	1.6 1.1	1.8 1.6	*** ***	0.7 ***	< 0.5 < 0.5
Oliva Virgen Extra	L	15.2	3.4	0.9	0.9	***	0.7	< 0.5
		118.0-121.0	5.0	1.0	***	***	***	***
	E	27.1	1.1	***	0.9	***	***	< 0.5
Orujo de Oliva	L	221.0-236.0	9.0-10.0	5.0-6.0	***	***	***	***
	E	***	***	***	***	***	***	***
	L	79.34	9.61	12.99	1.0	0.6	1.2	< 0.5
		107.0	24.8	29.9	***	***	***	***
Girasol		465.0	69.0	75.0	***	***	***	***
	E	114.84	17.48	4.7	1.9	1.5	***	< 0.5
		97.2	14.9	4.1	***	***	***	***
Girasol alto oleico	L	143.3	20.0	22.5	***	***	***	***
	E	90.0	7.6	5.9	***	***	***	***
	L	117.9	48.7	55.2	4.1	0.7	2.4	0.9
		108.6	43.9	39.2	***	***	***	***
Soja		123.0-173.0	47.0-82.0	47.0-52.0	***	***	***	***
	E	40.3	6.4	4.0	0.7	***	***	***
		51.3	10.7	3.2	***	***	***	***
	F	117.4	63.2	2.0	0.5	35.7	***	1.2
		158.1	93.3	***	***	***	***	***
Colza		130.0-173.0	76.0-270.0	traces	***	27.0-54.0	***	***
	E	246.4	173.5	0.5	1.8	36.3	0.6	1.6
		254.7	191.1	***	***	***	***	***

(1) Libres  
(2) Esterificados

## Introducción

**Tabla V.** Contenido de 4 $\alpha$ -monometilesteroles (mg/100 g) en aceites vegetales.

Aceite	Obtus. <sup>(1)</sup>	Gram. <sup>(1)</sup>	Cycloeu. <sup>(1)</sup>	Citrost. <sup>(1)</sup>
Oliva	1.1-7.9	0.5-2.0	2.2-22.3	3.5-31.6
Soja	1.5-5.3	2.1-15.8	0.6-6.6	4.1-35.0
Colza	1.2-15.7	0.4-5.7	0.1-5.9	trazas-4.3
Maíz	12.9-51.3	16.1-37.1	trazas-5.6	10.9-26.8
Palma	0.5-6.1	0.4-0.6	0.8-24.1	0.2-3.2
Grano de Palma	0.3-0.5	trazas	0.6-1.3	0.6-1.2
Cacahuete	0.3-4.2	0.3-5.6	trazas-0.9	0.4-3.2
Cártamo	2.1-4.0	2.1-4.3	0.3-1.4	0.9-2.4

<sup>(1)</sup> **Obtus.:** Obtusifoliol  
**Gram.:** Gramisterol  
**Cycloeu.:** Cycloeucaalenol;  
**Citrost.:** Citrostadienol

**Tabla VI.** Contenido de 4-4'-dimetilesteroles (mg/kg) en aceites vegetales.

Aceite	24-metilencicloartanol	$\beta$ -amirina	$\alpha$ -amirina	<b>Total</b> (incluidos los desconocidos)
Oliva Virgen	64	5	===	===
Soja	30-109	10-126	20-42	hasta 840
Colza	hasta 200	hasta 11	hasta 25	hasta 540
Maíz	80-330	2-19	hasta 33	82-1150
Palma	7-109	trazas-2	2-3	9-320
Grano de Palma	8-94	hasta 29	hasta 209	8-720
Cacahuete	13-166	2-25	1-25	16-360
Cártamo	12-31	49-79	41-57	102-840
Algodón	15-158	hasta 34	hasta 37	15-480

Entre todos los aceites comestibles usados para la preparación de alimentos hay que destacar el aceite de oliva, ya que no solo contribuye aportando nutrientes, si no que además presenta efectos beneficiosos para la salud humana.

Entre sus cualidades cabe destacar que, además de ser considerado un tesoro gastronómico que conjuga como pocos el placer para el paladar con sus efectos beneficiosos, es efectivo para reducir el nivel de colesterol en sangre y prevenir el riesgo de enfermedades cardiovasculares, tiene un efecto beneficioso sobre la tensión arterial, facilita la digestión, previene la formación de coágulos y contiene propiedades antioxidantes que retardan el envejecimiento de las células [34]. Produce un perfil lipídico favorable que conlleva a una disminución del riesgo de padecer arterioesclerosis. También protege del cáncer de colon y mama y, como contiene vitaminas A y E, además su consumo contribuye a mejorar el aspecto y tersura [35].

Según el Reglamento (EC) 1513/2001 que modifica el Reglamento nº 136/66/CEE y el Reglamento (CE) nº 1638/98, la máxima categoría de aceite de oliva, es el aceite de *oliva virgen extra* que se obtiene del prensado de las aceitunas sin otra manipulación que la mecánica. Le siguen en la clasificación, *el aceite de oliva virgen*, *el aceite de oliva lampante*, *el aceite de oliva refinado* y *el aceite de oliva*. Este último, el más consumido, está compuesto por una mezcla de aceite de oliva refinado y virgen.

El aceite de oliva tiene la ventaja de que no sólo se utiliza directamente sino que se emplea para cocinar. Hay que tener en cuenta que los aceites, en general, se modifican durante este proceso, y si la temperatura sube por encima del punto de ebullición, se pueden producir ciertos productos que son tóxicos para nuestro organismo. Sin embargo, al aceite de oliva no le ocurre esto debido a que tiene un punto de ebullición alto y es más resistente por ejemplo, que los aceites de girasol o maíz, siendo este es uno de los motivos por el que se utiliza habitualmente en los países mediterráneos.

Desde el punto de vista de su composición, sus componentes se pueden agrupar en dos fracciones [36] una mayoritaria (98-99%), constituida por ácidos grasos mono y poliinsaturados (MUFAs y PUFAs), mono, di- y triglicéridos (MAGs, DAGs y TAGs), y otra minoritaria ( $\approx 2\%$ ) dentro de la cual se encuentran entre otros, los fitoesteroles. La Tabla VII recoge el contenido de 4-desmetilesteroles y alcoholes triterpénicos en diversos tipos de aceite de oliva [37].



## Introducción

**Tabla VII.** Contenido de 4-desmetilesteros y alcoholes triterpénicos en aceite de oliva expresado como mg estero/100 g de aceite

<b>Aceite</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>Total<sup>(*)</sup></b>	<b>(6)</b>
Oliva	≥ 93	≤ 4	≤ 4	≤ 0.1	≤ 0.5	≥ 100	≤ 4.5
Oliva Virgen	≥ 93	≤ 4	≤ 4	≤ 0.1	≤ 0.5	≥ 100	≤ 4.5
Oliva Virgen Extra	≥ 93	≤ 4	≤ 4	≤ 0,1	≤ 0.1	≥ 100	≤ 4.5
Oliva Lampante	≥ 93	≤ 4	-	≤ 0.1	≤ 0.1	≥ 100	≤ 4.5
Oliva Refinado	≥ 93	≤ 4	≤ 4	≤ 0.1	≤ 0.1	≥ 100	≤ 4.5
Orujo de Oliva Crudo	≥ 93	≤ 4	-	≤ 0.2	≤ 0.1	≥ 250	> 4.5
Orujo de Oliva Refinado	≥ 93	≤ 4	< 4	≤ 0.2	≤ 0.1	≥ 180	> 4.5
Orujo de Oliva	≥ 93	≤ 4	< 4	≤ 0.2	≤ 0.1	≥ 160	> 4.5

**(1) β-Sitosterol:** Expresado como ( $\Delta^{5,23}$  – estigmastadienol + Clerosterol + β-sitosterol + Sitostanol +  $\Delta^5$  – avenasterol +  $\Delta^{5,24}$  – estigmastadienol)  
**(2) Campesterol**  
**(3) Estigmasterol**  
**(4) Brasicasterol**  
**(5) Colesterol**  
**(6) Eritrodiol + Uvaol**  
<sup>(\*)</sup>mg estero/100 g aceite

### 4.2. Formas de administración

Los fitoesteros se consideraban inicialmente como agentes farmacológicos, pero con el conocimiento de que constituían parte de la dieta surgió la idea de que debían administrarse en forma de productos alimenticios convencionales. Al principio se suministraban en forma de preparados pulverulentos y en cantidades relativamente grandes para lograr un efecto significativo. Así, los preparados de soja inicialmente tuvieron un gran éxito, pero alrededor del año 1960 su interés empezó a ser menor, ya que eran difíciles de administrar debido a su naturaleza insoluble y, en consecuencia, creaban incertidumbre acerca de la cantidad ingerida. Peterson et al [38] recomendaron incorporar estos compuestos en la mantequilla, con una ingesta diaria de 5.7 g de fitoesteros.

Tradicionalmente, los alimentos ricos en grasas como las margarinas y los aceites se han considerado los vehículos ideales para el consumo de fitoesteros, debido a su naturaleza hidrofóbica. Hay estudios que indican que los fitoesteros que se han ido incorporando a otro tipo de alimentos como mayonesa, aderezos para ensaladas, leche y yogur disminuyen más efectivamente los niveles de colesterol LDL en comparación con otros productos a los que también se les ha

adicionado, como por ejemplo cruasanes, magdalenas, zumo de naranja, barritas de cereales y chocolate [39].

Finalmente, los fitoesteroles también se han combinado con otros componentes nutricionales beneficiosos, incluyendo el aceite de pescado, el aceite de oliva, fibras y frutos secos, con objeto de aumentar su efecto contra los factores de riesgo de las enfermedades cardiovasculares [40,41,42].

## 5. ANÁLISIS DE FITOESTEROLES

Dado el efecto beneficioso que presentan estos compuestos y considerando que pueden estar presentes de forma natural o bien, ser adicionados a alimentos funcionales, se han desarrollado muchos métodos analíticos para su determinación en una amplia gama de productos alimenticios.

La secuencia más comúnmente seguida para el análisis de fitoesteroles individuales implica la separación de la fracción grasa de la matriz (extracción de los lípidos), su saponificación, o hidrólisis ácida seguida de saponificación para liberar los fitoesteroles, extracción de la fracción insaponificable, separación o purificación parcial de los esteroides, la formación de derivados de los mismos y su análisis mediante cromatografía [30, 43, 44, 45]. En la mayoría de los casos, la determinación de estos analitos se realiza como alcoholes (fitoesteroles libres) en vez de ésteres esterilicos o estanilicos.

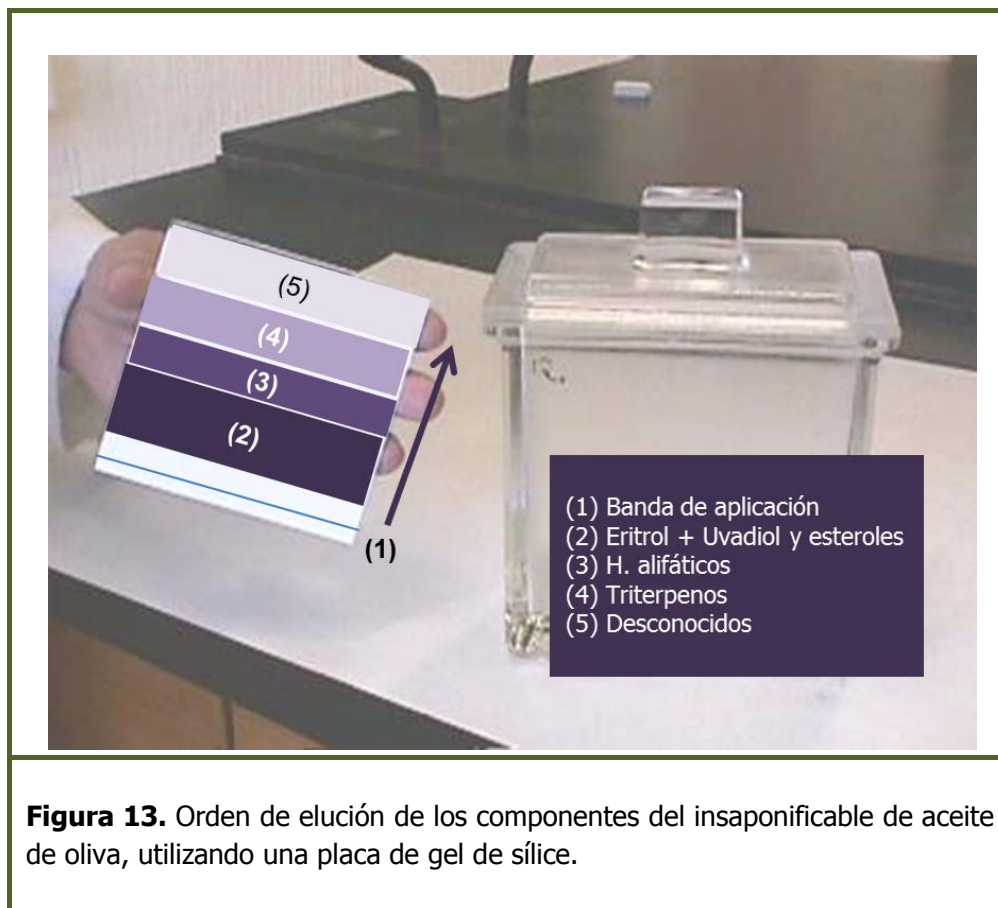
### 5.1. Aislamiento de fitosteroides para su posterior determinación cromatográfica

...Los procedimientos analíticos más utilizados para aislar los esteroides vegetales del resto de componentes de la fracción insaponificable han sido la cromatografía en capa fina (CCF), la extracción en fase sólida (SPE), la extracción en fluidos supercríticos (SFE), la cromatografía Flash y la cromatografía líquida de alta resolución (HPLC), en su modalidad preparativa.

La CCF [46] es la técnica cromatográfica que se emplea en el método oficial para la separación de esteroides. La fase estacionaria es una capa fina de gel de sílice dispuesta normalmente sobre un soporte de vidrio o aluminio y el eluyente o fase móvil está constituido por una mezcla de disolventes (n-hexano: éter dietílico).

## Introducción

En la Figura 13 se muestra el orden de elución de los compuestos presentes en la fracción insaponificable. Se puede observar que los compuestos más polares son los más fuertemente retenidos.



**Figura 13.** Orden de elución de los componentes del insaponificable de aceite de oliva, utilizando una placa de gel de sílice.

El principal inconveniente de este tipo de aislamiento es que resulta ser largo y tedioso.

Como una de las alternativas a la separación en capa fina, la extracción en fase sólida (SFE) se ha utilizado ampliamente para la obtención de distintas clases de lípidos antes de su análisis mediante HPLC, GC y GC-MS.

La Tabla VIII recoge los diferentes adsorbentes que se comercializan como fases estacionarias [47], de entre ellos, los que más se emplean en el análisis de los fitoesteroides son la gel de sílice, la alúmina y la C18.

Azadmard-Darmichi et al. [48,49] han propuesto un método rápido para la separación de las diferentes clases de fitoesteroides (4-desmetilesteroides, 4-monometilesteroides y 4,4'-dimetilesteroides) en diferentes aceites vegetales, con la

ventaja de que es más sencillo, gasta menos disolventes que la CCF y consume menos tiempo. Empleando gel de sílice como adsorbente y una mezcla hexano:éter-dietílico como eluyente, los 4,4'-dimetilesteroles son los que se retienen más débilmente, seguidos de los 4-monometilesteroles, mientras que los 4-desmetilesteroles son los más fuertemente retenidos.

**Tabla VIII.** Tipos de fases estacionarias utilizadas para la SPE.

<u>Tipo de Fase</u>	<u>Grupo</u>	<u>Modo de separación</u>
Silica	- <b>Si-OH</b>	Fase normal
Aminopropil (NH <sub>2</sub> )	- <b>Si-(CH<sub>2</sub>)<sub>3</sub>NH<sub>2</sub></b>	
Cianopropil (CN)	- <b>Si-(CH<sub>3</sub>)(CH<sub>2</sub>)<sub>3</sub>CN</b>	
Diol(2,3-hidroxi propoxipropil)	- <b>Si-(CH<sub>2</sub>)<sub>3</sub>OCH<sub>2</sub>CH(OH)CH<sub>2</sub>OH</b>	
C18 (octadecil)	- <b>Si-C<sub>18</sub>H<sub>37</sub></b>	Fase inversa
C8 (octal)	- <b>Si-C<sub>8</sub>H<sub>17</sub></b>	
C2 (etil)	- <b>Si-C<sub>2</sub>H<sub>5</sub></b>	
Fenil	- <b>Si-Ph</b>	Cambio iónico
Amina cuaternaria	- <b>Si-(CH<sub>2</sub>)<sub>3</sub>N<sup>+</sup>(CH<sub>3</sub>)<sub>3</sub></b>	
Propilbenzenosulfonato	- <b>Si-(CH<sub>2</sub>)<sub>3</sub>C<sub>6</sub>H<sub>4</sub>-SO<sub>3</sub><sup>-</sup></b>	
Propilsulfonato	- <b>Si-(CH<sub>2</sub>)<sub>3</sub>SO<sub>3</sub><sup>-</sup></b>	

En general, la SPE se ha propuesto como técnica de extracción de fitoesteroles, (tanto en su forma libre como esterificada) [50,51], y se ha utilizado no solo para la extracción y purificación de los fitoesteroles totales de otros constituyentes del insaponificable de aceites vegetales, sino también en la separación de los fitoesteroles libres de los triglicéridos [52].

Entre las aplicaciones de la extracción con fluidos supercríticos (SFE), técnica adecuada para el aislamiento y la purificación de compuestos de baja volatilidad,

## Introducción

se encuentra la extracción de los fitoesteroles y el fraccionamiento de aceites y grasas mediante CO<sub>2</sub> supercrítico [53, 54, 55, 56].

Lesellier [57] propone separar los diferentes compuestos del insaponificable, entre ellos los esteroides, mediante cromatografía de fluidos super/subcríticos como alternativa a la cromatografía de gases (GC) y a la cromatografía líquida de alta resolución (HPLC).

Cabe destacar la revisión sobre las aplicaciones de la SFE, usando CO<sub>2</sub>, en la extracción de ácidos grasos y lípidos realizada por Sahena et al. [58].

La cromatografía Flash [59], es una modalidad cromatográfica preparativa que se diferencia de la cromatografía en columna clásica tanto, por el menor tamaño de las partículas de fase estacionaria como por la presión a la que trabaja (media presión). Aunque se conocen muchas aplicaciones desarrolladas para el fraccionamiento de mezclas procedentes de síntesis orgánicas, en relación con el análisis de fitoesteroides solo cabe destacar la aplicación desarrollada por Kalo et al. [60] en la que se realiza el fraccionamiento secuencial de los ésteres esteroides, triacilglicéridos, esteroides libres, diacilglicéridos y triglicéridos residuales en grasas y aceites, combinando las siguientes fases móviles: diclorometano:hexano (22:78, v/v), diclorometano:hexano (65:35, v/v) y acetona:diclorometano (6:94, v/v).

Hoy día, la principal ventaja de esta técnica preparativa, radica en que las separaciones se pueden monitorizar en línea con sistemas de detección tales como, UV-visible, IR o ELSD.

Aunque la HPLC es una técnica fundamentalmente analítica, también se ha empleado como técnica preparativa para el aislamiento de los fitoesteroides al presentar diferentes ventajas como: (i) necesitar poca muestra (unos pocos microlitros), (ii) requerir menor tiempo para obtener la fracción de fitoesteroides, y (iii) poder llevar a cabo una detección continua de las fracciones eluidas. Con respecto a esta última, cabe indicar que entre los sistemas de detección más utilizados destacan el de ultravioleta-visible (UV-Vis) [61] y el de índice de refracción (IR) [62, 63]

Desde un punto de vista metodológico se puede usar tanto en la modalidad off-line, para recolectar la fracción esteroidea con vistas a la posterior derivatización y análisis por GC, como en la modalidad on-line, generalmente acoplada a un cromatógrafo de gases.

Cuando se utiliza la HPLC en su modalidad "off line", en la etapa de preparación de la muestra, la forma más común de trabajar implica, la recolección de la fracción de interés y posterior inyección en el cromatógrafo de gases, o la concentración y/o derivatización del eluido como paso previo a la GC.

Finalmente, cabe indicar que aunque todas las técnicas citadas anteriormente se emplean como alternativas a la CCF para la separación de los fitoesteroles del insaponificable de los aceites vegetales, también se puede inyectar directamente la muestra de aceite, diluida en el disolvente adecuado (generalmente orgánico), en el cromatógrafo de líquidos y analizarla mediante un equipo LC-GC [64].

### 5.2. Determinación cromatográfica de fitoesteroles

Como se puede ver en la Tabla IX, en la que se resume la búsqueda bibliográfica realizada para indicar el estado del arte de los procedimientos, técnicas y métodos aplicados a la determinación de fitoesteroles en muestras de aceite de oliva y otros aceites vegetales, la técnica que más se emplea para el análisis de los mismos es la GC con columnas capilares. Con este tipo de columnas se consiguen tiempos de análisis cortos, se presentan menos problemas de interferencias, se mejora la resolución cromatográfica y, desde el punto de vista instrumental, presentan mayor estabilidad térmica si se comparan con las columnas de relleno o empaquetadas.

Por otro lado, cabe indicar que aunque es posible separar los fitoesteroles sin derivatizar, es preferible transformarlos en trimetilsililderivados (TMS-esteroles) debido a su alto peso molecular y su baja volatilidad,.

La cromatografía líquida de alta resolución (HPLC) se puede emplear directamente para el análisis de los esteroides, en las modalidades de fase normal (con fases estacionarias tipo sílica, diol, amino, ciano), para separar los fitoesteroles en clases, como en fase invertida (con fases estacionarias tipo ODS, C8 o fenilo), para separar los compuestos individuales. Esta última, es la modalidad más ampliamente utilizada [65].

Tanto si se emplea la GC como la HPLC, es muy frecuente realizar la detección y cuantificación mediante el acoplamiento a un espectrómetro de masas. Aunque las modalidades de ionización por impacto electrónico (EI) y la fotoionización a presión atmosférica (APPI) son alternativas eficaces para los compuestos lipofílicos, la ionización química a presión atmosférica (APCI) es la que más se utiliza para este tipo de análisis, ya que se puede acoplar fácilmente al sistema HPLC [66,67,68].

**TABLA IX.** Procedimientos, técnicas y métodos aplicados a la determinación de fitosteroles en muestras de aceite de oliva y otros aceites vegetales.

<b>Muestra</b>	<b>FASE 1</b> (Sep. de lípidos apolares)	<b>FASE 2</b> (Sep. de esteroles)	<b>FASE 3</b> (Anál. Cromatográfico)	<b>[Ref.]</b>
Oliva (refinado y orujo crudo)	Saponificación/SPE		<b>TMS Derivados</b> - HPLC - HRGC 5300-FID	[69]
Oliva (virgen, refinado y oliva extraído)	Saponificación - Extracción	CCF	<b>TMS Derivados</b> - GC -FID - GC-MS	[70]
Oliva	Extracción (metilato sódico en metanol diluido con MTBE)		On-line - HPLC-HRGC	[71]
Vegetales (canola, oliva, cártamo, colza y girasol)	Saponificación - Extracción  HPLC Preparativa	CCF	<b>TMS Derivados</b> - GC-(EI)MS - GC-FID - HPLC	[72]
Vegetales, Oliva y grasas	Saponificación (En frío y en caliente)	CCF SPE	<b>TMS Derivados</b> - GC-FID - HPLC	[73]
Oliva y mezclas de aceite de avellana	-	-	<b>TMS Derivados</b> - GC-FID - HPLC-MS	[74]
Oliva, vegetales (girasol, soja y colza)	Saponificación- Extracción	CC	<b>TMS Derivados</b> - GC-FID	[75]
Oliva	Saponificación- Extracción	-	<b>TMS Derivados</b> - GC-FID	[76]

**TABLA IX. (Cont.).** Procedimientos, técnicas y métodos aplicados a la determinación de los fitoesteroles en muestras de aceite de oliva y otros aceites vegetales.

<b>Muestra</b>	<b>FASE 1</b> (Sep. de lípidos apolares)	<b>FASE 2</b> (Sep. de esteroles)	<b>FASE 3</b> (Anál. Cromatográfico)	<b>[Ref.]</b>
Oliva Virgen Extra, colza y algodón		On-Line LC-GC (DualChrom)		[77]
Oliva Virgen Extra y Avellana	<b>METODO 1</b> (Esteroles esterificados) LC + Saponificación en caliente <b>METODO 2</b> SPE+ Saponificación en frío	CCF	<b>TMS Derivados</b> - GC-FID - GC-MS	[51]
Oliva Virgen	Saponificación-Extracción	CCF	<b>TMS Derivados</b> - GC-FID	[78]
Monovarietal Oliva Virgen	Saponificación-Extracción	CCF	<b>TMS Derivados</b> - GC-FID	[79]
Acebuchina y Oliva Virgen	Saponificación-Extracción	Pre-HPLC 1050	<b>TMS Derivados</b> - GC-FID	[63]
PDO <sup>(1)</sup> Oliva	Saponificación- SPE	CCF	<b>TMS Derivados</b> - GC-FID	[80]
Oliva y avellana	Saponificación-Extracción	CCF SPE	<b>TMS Derivados</b> - GC-FID - GC-MS	[49]
Oliva	Saponificación-Extracción	CCF	<b>TMS Derivados</b> - GC-FID	[81]
Variedades de oliva y colza	Saponificación-Extracción - Centrifugación	-	- GC-FID	[82]

<sup>(1)</sup> Productos con denominación de origen.



## Introducción

**TABLA IX. (Cont.).** Procedimientos, técnicas y métodos aplicados a la determinación de los fitoesteroles en muestras de aceites de oliva y otros aceites vegetales.

<b>Muestra</b>	<b>FASE 1</b> (Sep. de lípidos apolares)	<b>FASE 2</b> (Sep. de esteroles)	<b>FASE 3</b> (Anál. Cromatográfico)	<b>[Ref.]</b>
Oliva	Saponificación- Extracción	CCF	- LC-(APCI)MS	<b>[68]</b>
Oliva (Virgen Extra, Virgen, Refinado, Orujo de Oliva y crudo)	Saponificación- Extracción	CCF	- LC-(APCI)MS - LC-(ESI)MS	<b>[98]</b>
Oliva Virgen y Avellana	Saponificación- Extracción	SPE	<b>TMS Derivados</b> - GC-FID - GC-MS	<b>[83]</b>
Oliva Virgen Extra y cacahuete refinado	Saponificación en Frío	CCF	<b>TMS Derivados</b> - GC-FID - GC-MS	<b>[84]</b>
Varietades Oliva (Tunecino)	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID	<b>[85]</b>
Oliva Virgen	Saponificación- Extracción	NP-LC	- GC-FID	<b>[86]</b>
Oliva	Saponificación- Extracción	-	<b>TMS Derivados</b> - GC-FID	<b>[87]</b>
Aceitunas de mesa	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID	<b>[88]</b>
Oliva Refinado	Saponificación	CCF (esteroles libres y esterificados)	<b>TMS Derivados</b> - GC-FID - GC-MS	<b>[89]</b>

**TABLA IX. (Cont.).** Procedimientos, técnicas y métodos aplicados a la determinación de los fitoesteroles en muestras de aceite de oliva y otros aceites vegetales.

<b>Muestra</b>	<b>FASE 1</b> (Sep. de lípidos apolares)	<b>FASE 2</b> (Sep. de esteroides)	<b>FASE 3</b> (Anál. Cromatográfico)	<b>[Ref.]</b>
Olivas de Túnez (variedad Meski)	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID - GC-MS	<b>[90]</b>
Vegetal (girasol, maíz, soja, nuez, etc.)	Saponificación- Extracción	-	- LC-(APCI)ITMS	<b>[91]</b>
Oliva (Hojiblanca y Manzanilla)	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID	<b>[92]</b>
Aceitunas de mesa	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID - GC-MS	<b>[93]</b>
Oliva (Arbequina y Empeltre)	Saponificación- Extracción	CCF	<b>TMS Derivados</b> - GC-FID	<b>[94]</b>
Oliva Virgen Extra	Saponificación- Extracción	SPE	- Nano LC-(ESI)MS	<b>[95]</b>
Oliva Virgen Extra (Aberquina, Borriolenca, Picual, Canetera, Farga y Serrana)	Saponificación- Extracción	CCF	- LC- (APPI)MS	<b>[96]</b>
Oliva Virgen	Saponification- Extraction	CCF	- GC-FID	<b>[97]</b>
Aceites vegetales	Saponificación- Extracción	CCF	- UPLC-(APCI)MS	<b>[100]</b>

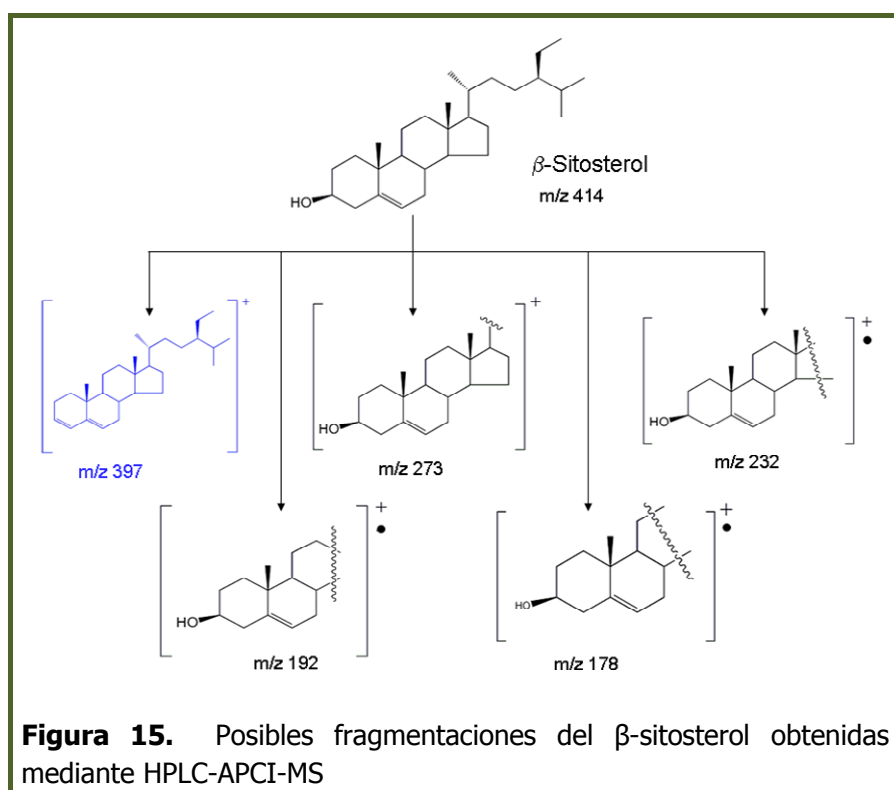
En el caso específico del análisis de aceite de oliva, tanto Cañavate et al. **[98]** como, Segura et al. **[99]**, indican que cada pico de fitoesterol libre presenta las

## Introducción

señales correspondientes a la relación masa/carga ( $m/z$ ) de los iones  $[M + H]^+$ ,  $[M - H_2O]^+$  y  $[M + H - H_2O]^+$ .

Recientemente, también se está empleando la ultra-cromatografía de líquidos (APCI-UPLC) ya que permite el uso de columnas con un tamaño de partícula de 1-2  $\mu\text{m}$  [100].

En la Figura 14 se muestran las posibles fragmentaciones que sufriría el  $\beta$ -sitosterol en un sistema HPLC-APCI-MS. Los iones que aparecen a  $m/z$  273 y 178 se deben a la pérdida de la cadena lateral y a la ruptura de los dos anillos adyacentes a la misma, respectivamente. De igual forma, las fragmentaciones que corresponden a los iones  $m/z$ , 232 y 192, se atribuyen a la ruptura de los enlaces entre C13-C17, C14-C15 y C12-C13 y C8-C14. El ion  $m/z$  397 puede deberse a la deshidratación del grupo OH presente en la molécula de fitoesterol libre.



En relación con el **acoplamiento HPLC-GC** aplicado al estudio de fitoesteroles en aceites vegetales, indicar que los primeros estudios que se conocen se iniciaron a finales de los ochenta, principios de los noventa cuando Grob et al. [101] determinaron fitoesteroles en aceite de oliva mediante el acoplamiento HPLC-GC después de la derivatización previa de los esteroides libres con ácido píválico. Los esteroides libres, los esterificados y los esteres de ceras, se separaban de la matriz del aceite mediante la modalidad NP-HPLC y se transferían al CG "on line". En un

trabajo posterior **[102]**, estos mismo autores sustituyeron la esterificación por la trimetilsililación, lo que permitía la determinación simultánea de escualeno, tocoferoles y epoxiesqualeno.

En 1993, Biedermann et al. **[103]**, propusieron un método para sustituir la etapa de saponificación por una transesterificación seguida del análisis "on line" HPLC-GC de esteroides.

A finales de los noventa, tanto Señorans et al., como Villén et al. **[104,105]**, utilizaron un vaporizador de temperatura programada (PTV) como interface para la transferencia entre el cromatógrafo de líquidos y el de gases. La separación en el cromatógrafo de líquidos, para la determinación de esteroides libres, se realizó en la modalidad de fase invertida y aunque la propuesta era novedosa, por inconvenientes técnicos el método desarrollado no se pudo automatizar.

Tras el desarrollo en 2001 del acoplamiento HPLC-GC con la interface TOTAD (transferencia a través de un horno de adsorción-desorción) **[106]**, en 2006 Cortés et al. **[107]** resolvieron el inconveniente y pusieron a punto un método automatizado RP-HPLC-GC "on line" a través de dicha interface. Dicho método, implicaba la inyección directa de la muestra (aceite) disuelta en el disolvente adecuado, y permitía dos alternativas: (i) la determinación conjunta de cuatro clases de lípidos (esteroides libres, tocoferoles, alcoholes triterpénicos y eritrol y uvaol) y (ii) la determinación conjunta de esteroides libres y tocoferoles.

Posteriormente, Toledano et al. **[108]** propusieron una modificación de este método en la que se puede llevar a cabo la determinación, tanto de esteroides libres como esterificados, en aceites. Así, para la determinación de esteroides libres, una alícuota de la muestra de aceite, diluida en un disolvente orgánico, se inyecta directamente en el HPLC donde los analitos son separados de los triglicéridos (TAGs), a continuación y de forma automatizada la fracción esterólica se transfiere al CG. Para determinar los esteroides totales se procede a la saponificación del aceite y a la extracción del insaponificable con éter dietílico, analizando posteriormente dicho extracto mediante RP-HPLC-GC.

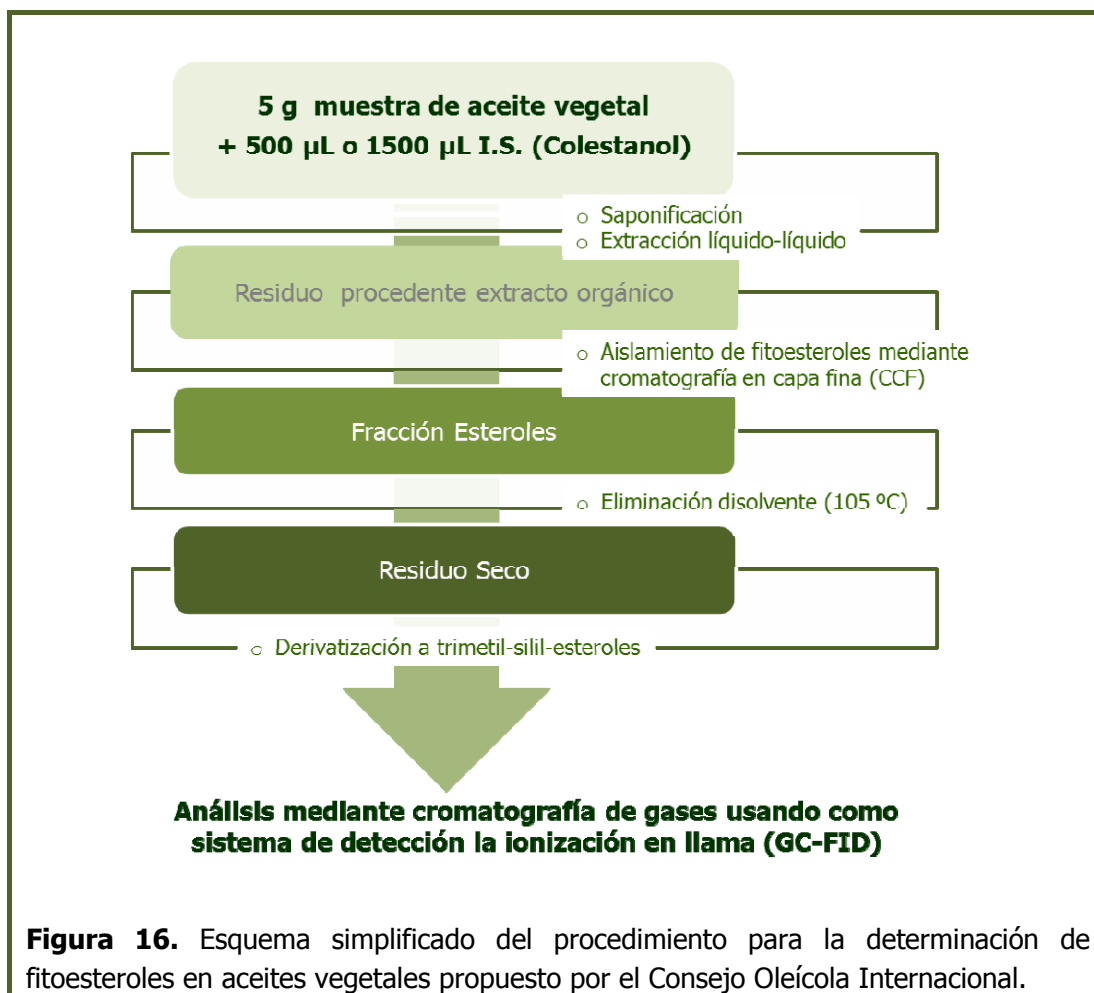
En 2012, estos mismos autores **[109]** proponen un método para la determinación de los fitoesteroides presentes en aceite de oliva, mediante NP-HPLC-GC. Lo más novedoso de dicho trabajo radica en el hecho de que es la primera vez que se propone la derivatización "on-line" de la fracción esterólica antes de entrar en el cromatógrafo de gases.

Finalmente, cabe indicar que no existen métodos oficiales para el análisis de esteroides en alimentos, a excepción del recogido en el Reglamento (CEE) 2568/91

## Introducción

de la Comisión Europea, de 11 de Julio de 1991 [110] y propuesto por el Consejo Oleícola Internacional (COI) [111,112,113], referido a la determinación de la composición y contenido de esteroides en aceites vegetales mediante cromatografía gaseosa con columna capilar, siendo este el que se suele utilizar normalmente en laboratorios oficiales o de rutina.

El procedimiento, esquematizado en la Figura 16, se basa en la saponificación de la muestra de aceite vegetal con una disolución etanólica de hidróxido potásico 2N, extracción de la fracción insaponificable con éter dietílico y aislamiento de los esteroides mediante cromatografía en capa fina (CCF) con gel de sílice. Posteriormente, los esteroides aislados se derivatizan, transformándolos en trimetilsililéteres, y se analizan mediante cromatografía de gases con detección de ionización de llama (GC-FID).



La siguiente Tabla muestra las modificaciones que se le han ido introduciendo al método analítico del COI.

**TABLA X.** Modificaciones sucesivas aplicadas al Método Oficial para la determinación de fitoesteroles.

Métodos	Modificaciones
Nº 2568/91	<ul style="list-style-type: none"> <li>- Determinación de la composición y del contenido de esteroides.</li> <li>- Disolución de colestanol en cloroformo</li> <li>- Separación de la fracción de esteroides mediante CCF</li> <li>- Condiciones cromatográficas recomendadas:               <ul style="list-style-type: none"> <li>Tª columna: 260 °C</li> <li>Tª Inyector: 280 °C</li> <li>Tª Detector: 290 °C</li> </ul> </li> <li>- <b>Contenido</b> de esteroides expresado en <b>mg/100 g de materia grasa y % de cada uno de ellos.</b></li> </ul>
COI 2001	<ul style="list-style-type: none"> <li>- Determinación de la composición y del contenido de esteroides.</li> <li>- Disolución de colestanol en cloroformo</li> <li>- Separación de la fracción de esteroides mediante CCF</li> <li>- Condiciones cromatográficas recomendadas:               <ul style="list-style-type: none"> <li>Tª columna: 260 ± 5 °C</li> <li>Tª Inyector: 280 °C</li> <li>Tª Detector: 290 °C</li> </ul> </li> <li>- <b>Contenido</b> de esteroides expresado en <b>mg/kg de materia grasa y % de cada uno de ellos.</b></li> </ul>
COI 2009	<ul style="list-style-type: none"> <li>- Determinación de la composición y del contenido de esteroides y dialcoholes triterpénicos.</li> <li>- Disolución de colestanol en acetato de etilo</li> <li>- Separación de la fracción de esteroides y dialcoholes triterpénicos mediante CCF</li> <li>- Condiciones cromatográficas recomendadas:               <ul style="list-style-type: none"> <li>Tª columna: 260 ± 5 °C</li> <li>Tª Inyector: 280 – 300 °C</li> <li>Tª Detector: 290 – 300 °C</li> </ul> </li> <li>- <b>Contenido de esteroides</b> expresado en <b>mg/kg de materia grasa, % de cada uno de ellos.</b></li> <li>- Determinación de ergosterol</li> </ul>

## Introducción

**TABLA X.(Cont.)** Modificaciones sucesivas aplicadas al Método Oficial para la determinación de fitoesteroles.

Métodos	Modificaciones
COI 2011	<ul style="list-style-type: none"><li>- Determinación de la composición y del contenido de esteroides y dialcoholes triterpénicos.</li><li>- Disolución de colestanol en acetato de etilo</li><li>- Separación de la fracción de esteroides y dialcoholes triterpénicos mediante CCF</li><li>- Condiciones cromatográficas recomendadas:<ul style="list-style-type: none"><li>Tª columna: 260 ± 5 °C</li><li>Tª Inyector: 280 – 300 °C</li><li>Tª Detector: 290 – 300 °C</li></ul></li><li>- <b>Contenido de esteroides</b> expresado en <b>mg/kg de materia grasa, % de cada uno de ellos.</b></li><li>- Determinación de ergosterol</li></ul>

A lo largo de sus modificaciones, y como se ha resaltado en la Tabla X, dicho método propone tres formas de expresar el resultado final:

a) En mg de cada uno de los fitoesteroides /100 g de muestra **[101]**, que viene dado por la siguiente expresión:

$$\text{fitosterol}_i \text{ (mg/100 g de muestra)} = \frac{A_x \times m_s \times 100}{A_s \times m} \quad [\text{Ec. 1}]$$

En la que:

- $A_x$  es el área del pico del fitosterol "i".
- $A_s$  es el área del pico del 5 $\alpha$ -colestano-3 $\beta$ -ol (I.S.).
- $m_s$  es la masa del patrón interno (I.S.), en miligramos.
- $m$  es la masa de la muestra de aceite en gramos.
- 100 es el factor utilizado para la transformación a porcentaje.

b) En porcentaje con respecto a los fitoesteroides totales **[101,102,103,104]**. En este caso, el porcentaje de cada uno de los fitoesteroides se calcula, de acuerdo con la Ec. 2, dividiendo el área del pico correspondiente entre la suma de las áreas de todos los picos de los fitoesteroides excluyendo la del I.S.:

$$\% \text{ fitoesterol}_i = \frac{A_i}{\sum A} \times 100 \quad [\text{Ec. } 2]$$

En la que:

- $A_x$  es el área del pico del fitoesterol "i".
- $\sum A$  es suma de las áreas de todos los picos (excluyendo el I.S.).

c) En mg de fitoesterol<sub>i</sub> por kilogramo de muestra de aceite (partes por millón de cada esterol) **[102,103,104]**.

$$\text{fitoesterol}_i \text{ (mg/kg de muestra)} = \frac{A_x \times m_s \times 1000}{A_s \times m} \quad [\text{Ec. } 3]$$

En la que:

- $A_x$  es el área del pico del fitoesterol "i".
- $A_s$  es el área del pico del 5 $\alpha$ -coleston-3 $\beta$ -ol (I.S.).
- $m_s$  es la masa del patrón interno (I.S.), en miligramos.
- $m$  es la masa de la muestra de aceite en gramos.
- 1000 es el factor utilizado para la transformación a partes por millón.

## 6. CUANTIFICACIÓN DE FITOESTEROLES

Reordenando la ecuaciones 1 y 3 del apartado anterior (ecuaciones 4 y 5) se puede observar que desde un punto de vista metrológico, la cuantificación empleada en el método oficial se corresponde, de acuerdo con Cuadros et al. **[114]**, con la aplicación de la calibración interna (participación de la muestra) en un solo punto o lo que es lo mismo, a un solo nivel, en la que se utiliza el 5 $\alpha$ -coleston-3 $\beta$ -ol (conocido como dihidrocolesterol o colestonol) como patrón interno (I.S.) o subrogado.

$$C_{\text{esterol}(i)} = 100 \times \frac{m_{\text{subrogado (I.S.)}}}{m_{\text{muestra}}} \times \frac{A_{\text{sterol}(i)}}{A_{\text{subrogado (I.S.)}}} \quad [\text{Ec. } 4]$$

$$C_{\text{esterol}(i)} = 1000 \times \frac{m_{\text{subrogado (I.S.)}}}{m_{\text{muestra}}} \times \frac{A_{\text{sterol}(i)}}{A_{\text{subrogado (I.S.)}}} \quad [\text{Ec. } 5]$$

En las determinaciones mediante técnicas cromatográficas, y tal como indican estos autores, este tipo de calibración-cuantificación presenta las siguientes ventajas:



## Introducción

(i) Permite realizar la cuantificación simultánea de varios analitos (generalmente de la misma familia) utilizando una única alícuota de muestra y uno o varios patrones internos o subrogados que representan a todos los analitos. La única condición que debe cumplir el patrón-es interno-s, es la de que el equipo de medida y ensayo (en este caso el cromatógrafo) sea capaz de distinguir entre las señales atribuidas al patrón-es interno-s (generalmente subrogados) y las atribuidas a cada uno de los analitos.

(ii) Produce una reducción -en muchos casos considerable- del tiempo de análisis, ya que solo se necesita de una preparación analítica para la calibración y la cuantificación, puesto que en la misma porción de ensayo, se obtienen la señal(es) analíticas, tanto del patrón(es) interno(s) como las de cada analito. Para el caso más simple, en el que solo se utilice un I.S., se obtiene un factor de calibración que se utiliza para cuantificar a los analitos a partir de sus señales individuales.

(iii) Minimiza la pérdida de analitos durante la preparación de la muestra y en menor medida la existencia de un efecto matriz.

Por estas razones, se considera que la calibración interna es la mejor opción para la cuantificación si, en el método analítico aplicado, concurren al menos una de las siguientes circunstancias:

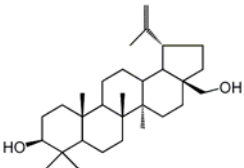
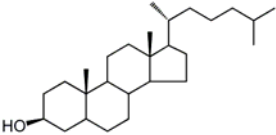
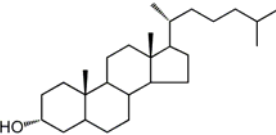
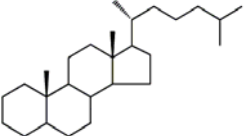
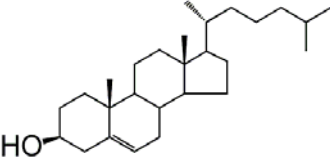
- (a) El proceso de preparación de la muestra es largo y complicado;
- (b) se requiere bastante tiempo para la medida (cromatogramas largos), y
- (c) no existen, o es prácticamente imposible disponer comercialmente de los patrones de todos los analitos que van a ser determinados.

Evidentemente, la limitación principal de esta metodología radica en la disponibilidad de un subrogado adecuado. En este sentido, tal y como se recoge en la Tabla XI, hay autores que proponen el uso de otras sustancias como subrogado alternativas al colestanol ( $5\alpha$ -colestano- $3\beta$ -ol).

Costin et al. proponen otra forma de cuantificación de estos compuestos, basada en la utilización de los llamados factores de corrección teóricos (TCFs) **[115]**.

De acuerdo con estos autores, dichos factores son bastante fáciles de determinar suponiendo que se conozca la estructura de los compuestos, pudiéndose usar para obtener un dato analítico fiable y exacto, así como para simplificar la preparación y el tratamiento de los datos experimentales.

**Tabla XI.** Diferentes subrogados utilizados para la determinación de fitoesteroles

Nombre y estructura del subrogado	[Ref.]
	[60, 80, 116, 117, 118, 119, 120, 121]
Betulina	
	[43, 51, 69, 74, 76, 78, 85, 86, 88, 90, 92, 122, 123, 123]
<u>5-α-Colestan-3β-ol/ Dihidrocolesterol</u>	
	[31, 43, 124, 125]
5-β-Colestan-3α-ol/ Epicoprostanol	
	[89, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 133, 134, 135]
5α-Colestanol	
	[ 60, 75, 81, 118, 123]
Colesterol	

Para el cálculo de los TCFs, hay que determinar en primer lugar, el número de carbonos activos del compuesto, es decir, todos los átomos de carbono excepto los de los carbonilos, ya que estos no dan señal en el FID. Los grupos trimetilsilil contienen 3

## Introducción

átomos de C y, como contribuyen también a la respuesta del detector, se incluyen en el cálculo de los carbonos activos.

El TCF se define como el cociente obtenido al dividir la suma de los pesos moleculares de todos los carbonos activos, entre el peso molecular de la molécula de fitoesterol.

El TCF relativo del estándar interno (I.S.) respecto al fitoesterol "X" ( $TCF_{I.S./X}$ ) se calcula de la siguiente manera:

$$TCF_{I.S./X} = \frac{TCF_{I.S.}}{TCF_X} = \frac{PM_{CARBONOS \text{ ACTIVOS}(I.S.)} / PM_{I.S.}}{PM_{CARBONOS \text{ ACTIVOS}(X)} / PM_X} \quad [\text{Ec. } 6]$$

siendo, PM el Peso molecular del patrón interno y del esterol X.

Una vez que se ha calculado el  $TCF_{I.S./X}$  para cada fitoesterol, la cantidad del mismo se puede determinar a partir de los datos cromatográficos, haciendo uso de la ecuación:

$$Q_X = \frac{A_X \times TCF_{I.S.} \times Q_{I.S.}}{A_{I.S.} \times TCF_X} = TCF_{I.S./X} \cdot Q_{I.S.} \cdot \frac{A_X}{A_{I.S.}} \quad [\text{Ec. } 7]$$

siendo:

- $Q_X$ : Cantidad de fitoesterol x
- $TCF_{I.S./X}$ : Factor de corrección teórico del I.S. con respecto al fitoesterol X.
- $Q_{I.S.}$ : Cantidad de patrón interno
- $A_X$ : Área del pico del esterol X
- $A_{I.S.}$ : Área del pico del patrón interno

Con arreglo a los criterios anteriores, la Tabla XII recoge los TCF relativos de varios trimetilsililderivados de los 4-desmetilesteroles que se van a estudiar en esta Memoria.

En el Capítulo III se han cuantificado los fitoesteroles mediante el procedimiento empleado en el método oficial y haciendo uso de los TCFs. No se han empleado los Factores de Corrección Experimentales (ECFs) debido a que para su determinación se requiere mucho tiempo y, principalmente, porque se necesitan estándares de elevada pureza, difíciles y, en algunos casos, imposibles de obtener.

**Tabla XII.** Factores de corrección teóricos calculados para los TMS-derivados de los 4-desmetilesteroles estudiados.

TMS-derivado de:	TCF <sub>IS/X</sub>
- Colestanol (I.S.)	1.0000
- Colesterol	0.9956
- Brasicasterol	0.9887
- 24-Metilencolesterol	0.9887
- Campesterol	0.9930
- Campestanol	0.9972
- Estigmasterol	0.9864
- $\Delta^7$ -Campesterol	0.9930
- $\Delta^{5,23}$ -Estigmastadienol	0.9864
- Clerosterol	0.9864
- $\beta$ -Sitosterol	0.9905
- Estigmastanol	0.9925
- $\Delta^5$ -Avenasterol	0.9864
- $\Delta^{5,24}$ -Estigmastadienol	0.9864
- $\Delta^7$ -Estigmastenol	0.9822
- $\Delta^7$ -Avenasterol	0.9864

## 7. CARACTERIZACIÓN Y AUTENTIFICACIÓN DE ACEITES VEGETALES Y MUESTRAS RELACIONADAS MEDIANTE EL TÁNDEM CROMATOGRAFÍA-QUIMIOMETRÍA

Existen una serie de problemas en la ciencia de la separación que requieren del uso de la quimiometría, tales como la evaluación de la calidad de la separación cromatográfica [16,17] o la optimización del sistema cromatográfico mediante un diseño experimental [18,19]. Además, el uso de los métodos multivariantes de reconocimiento de pautas combinados con la cromatografía, es importante para la monitorización del control de calidad, la calibración y la discriminación de datos analíticos [135, 136].

## Introducción

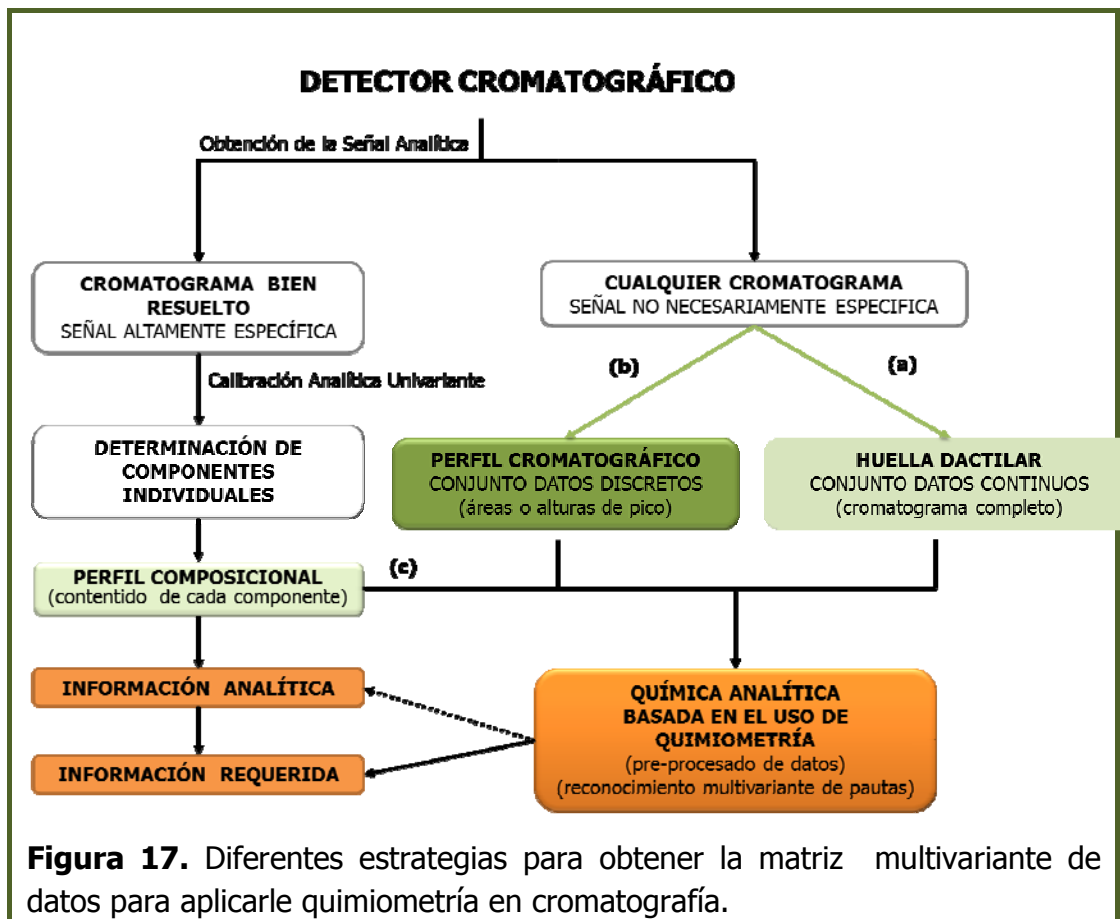
De una forma muy simple, las matrices de datos analíticos usadas en quimiometría, pueden venir de dos Fuentes: (i) señales que provienen directamente del cromatógrafo (huella cromatográfica o fingerprint); o (ii) datos obtenidos a partir de información tratada como, señales de intensidad medidas o resultados de composición/concentración (perfiles). La Figura muestra gráficamente los casos descritos.

En el primer caso (a), la matriz de datos la constituye el cromatograma complete manejado como una señal continua. Sin embargo, en la mayor parte de las ocasiones (b) lo que se utiliza es una serie de datos individuales previamente seleccionados por el analista. Habitualmente, esta serie de datos es un perfil analítico de una familia de compuestos con propiedades similares, por ejemplo, el perfil de compuestos volátiles de un alimento. El perfil anteriormente mencionado podría corresponder con las áreas o alturas de pico (perfil cromatográfico). Lo que se utiliza más frecuentemente es el caso (c) en el cual, la matriz está formada por los contenidos (concentraciones o porcentajes) de cada uno de los componentes de interés (perfil composicional), obtenidos después de un proceso previo de identificación, calibración y cuantificación de los analitos. Como complemento, los datos de la matriz se podrían aumentar añadiendo datos experimentales obtenidos a partir de medidas independientes [137].

Ambas técnicas se podrían aplicar de forma sinérgica. En el primer caso, la quimiometría podría mejorar los resultados de los cromatogramas ya que se podrían corregir componentes externos de la señal, como el ruido o la línea base [138]. Además, teniendo en cuenta que la cromatografía realiza la separación del analito en condiciones de cuasi-equilibrio, el entorno químico afecta las características de cada pico, es decir, diferentes analitos podrían variar con respecto a otros. El papel de la quimiometría podría ser crucial en la reducción de de este tipo de efectos indeseables. Es importante tener en cuenta que, incluso si solo se aplican las herramientas de preprocesado, los resultados serán mucho más fiables.

Sin embargo, la aplicación de las herramientas quimiométricas sobre los datos cromatográficos sin tratar, es decir, sobre la huella cromatográfica, no está libre de problemas, como se puede ver en dos tutoriales recientemente publicados [139]. Además, casi siempre se requieren los datos específicos de las técnicas de preprocesado. [136]. De hecho, la superación de problemas relacionados con el alineamiento de los picos [136, 139, 140] o la selección de variables [136,141] es particularmente crítica cuando la huella cromatográfica se usa como serie de datos. En especial, se ha demostrado recientemente que para resolver problemas de alineamiento de todos los cromatogramas en una serie de datos, el algoritmo icoshift, es una herramienta poderosa [141]. Estos problemas (línea base y

desalineación) podrían ser la razón por la que en la literatura hay más estudios que usan el área/altura de los picos (perfil de picos) o datos de concentración (perfiles de concentración o composicionales) que los que usan datos sin tratar (estrategia (c) en Fig 17 . [142] [143] [144].



En la revisión bibliográfica sobre la aplicación de las técnicas quimiométricas al análisis de fitoesteroides que aparece recogida en la Tabla XXXX se puede observar que la herramienta que más se ha utilizado hasta ahora es el Análisis de Componentes Principales (ACP), como método de exploración de datos multivariantes o de reconocimiento de pautas, si bien en algunos casos y de forma errónea, también se ha empleado como método de clasificación. También se han utilizado el análisis de clusters o conglomerados (CA o HCA) y diversos estudios asociados a la varianza (ANOVA, MANOVA). En relación a las herramientas quimiométricas para la clasificación, de entre las técnicas supervisadas de reconocimiento de pautas hay que destacar el Análisis Discriminante (AD o ADL).

## Introducción

**Tabla XIII.** Caracterización y autenticación de aceites vegetales mediante el tándem CROMATOGRAFÍA-QUIMIOMETRÍA

Matriz de datos	Tipo Cromatografía	Herramienta quimiométrica	Muestra	[Ref.]
Perfil en concentración de ácidos grasos, triglicéridos y esteroides	- GC-FID - HPLC-RI	-PCA (Rot. Varimax), -LDA -ANN	Aceite de oliva virgen extra de origen italiano	[145]
Perfil en concentración de esteroides y eritrodol + uvaol	- GC-FID	-DA	Aceite de oliva virgen (7 tipos o variedades de aceituna) de origen español	[146]
Perfil en concentración de ácidos grasos, triglicéridos, fitoesteroides y vitamina E (tocoferoles y tocotrienoles)	- GC-FID - HPLC	-PCA, -DA	Aceite de oliva origen portugués (3 variedades)	[147]
Perfil en señal de alcoholes, esteroides, hidrocarburos y tocoferoles	- RMN	-LDA , -PLS-DA, -SIMCA, -CART	Aceite de oliva virgen de España, Italia, Grecia, Turquía, Túnez y Siria	[148]
Datos discretos de fitoesteroides	-UPLC-(APCI)MS	-LDA	Aceite de oliva virgen extra de origen español	[149]
Perfil en concentración de ácidos grasos, alcoholes terpénicos, esteroides e hidrocarburos	- GC-FID - HPLC	-ANN	Aceite de oliva virgen de origen español, italiano y portugués	[150]
Perfil en concentración de ácidos grasos, tocoferoles y tocotrienoles, triglicéridos y fitoesteroides	- GC-FID	-ANOVA, -PCA	Aceite de avellana (China, Francia, Hungría, India, Italia, España y EEUU).	[151]
Perfil en concentración de ácidos grasos, triglicéridos, tocoferoles y fitoesteroides	- GC-FID - HPLC-UV	-ANOVA, -PCA, -HCA	Aceite de semilla de albaricoque de Turquía	[152]
Perfil en concentración de fitoesteroides, alcoholes triterpénicos entre otros	- GC-FID	-MANOVA - PCA, -DA	Aceitunas de mesa de las variedades hojiblanca y manzanilla	[153]
Perfil en concentración de fitoesteroides, alcoholes grasos y alcoholes triterpénicos	- GC-FID	-MANOVA, -PCA, -DA	Aceitunas de mesa	[154]
Perfil en señal de triglicéridos, $\beta$ -sitosterol,	-HPLC-APCI-MS	-PCA, -LDA	Aceite de oliva virgen extra de Italia y Grecia	[155]
Perfil en concentración de ácidos grasos, triglicéridos, fitoesteroides y alcoholes triterpénicos	- GC-FID - HPLC-UV	-MANOVA, -PCA, -LDA	Aceitunas de mesa	[156]

**Tabla XIII.** Caracterización y autenticación de aceites vegetales mediante el tándem CROMATOGRAFÍA-QUIMIOMETRÍA

Matriz de datos	Tipo Cromatografía	Herramienta quimiométrica	Muestras	[Ref.]
Perfil en concentración de ácidos grasos, tocoferoles y fitoesteroles	- GC-FID - HPLC-DAD	-PCA, -LDA, -QDA, -UNEQ, -SIMCA	Aceite de oliva virgen extra de origen italiano	[157]
Perfil en concentración de ácidos grasos y fitoesteroles	- GC-FID - HPLC	-LDA, -ANN	Aceite de oliva virgen extra de origen italiano	[158]
Perfil en concentración de tocoferoles, fitoesteroles y ácidos grasos	- GC-FID - HPLC-DAD	-MANOVA, -PCA, -CA	Aceite de oliva de 3 variedades portuguesas	[159]
Perfil en concentración de fitoesteroles libres y esterificados	- GC-FID	-PCA, -LDA	Aceite de oliva virgen lampante de España, Grecia, Italia, Marruecos, Turquía y Túnez	[160]
Perfil en concentración de ácidos grasos y fitoesteroles	- GC-FID - GC-MS	-PCA, -LDA	Aceite de semillas de pistacho de Italia, Grecia, Irán y Turquía	[161]
Perfil en concentración de triglicéridos y fitoesteroles	- GC-FID - HPLC-RI	-PCA, -SIMCA	Aceite de oliva virgen extra de España (Extremadura)	[162]
Datos discretos de aldehídos, terpenos, escualeno, cicloartenol y $\beta$ -sitosterol	- $^1\text{H}$ RMN	-ANOVA, -LDA, -CA	Aceite de oliva virgen extra de Italia	[163]
Perfil en concentración de triglicéridos y fitoesteroles	- HPLC-RI - GC-FID	-ANOVA, -PCA, -HCA	Aceite de oliva de Túnez	[164]
Perfil en concentración de triglicéridos, tocoferoles y fitoesteroles	- HPLC-(APCI)-MS - GC-FID	-ANOVA, -DA	Aceite de oliva y avellana	[165]
Perfil en concentración de triglicéridos, ácidos grasos y fitoesteroles	- GC-FID - HPLC-RI	-ANOVA, -PCA, -HCA, -DA	Aceite de oliva de Italia	[166]



## Introducción

De igual modo, el tipo de matriz de datos que más se ha tratado, es el que utiliza el perfil de concentración de los fitoesteroles combinado con datos de triglicéridos, ácidos grasos, tocoferoles y tocotrienoles, y aunque se han encontrado un par de casos en el que se emplean perfiles de datos discretos obtenidos a partir de la intensidad de iones fragmentados en los espectrómetros de masas acoplados a HPLC o GC, sorprendentemente, hasta el momento no se ha utilizado una matriz de datos continuos, construida a partir del cromatograma entero.

**BIBLIOGRAFÍA**

- [1] Oliver Palou, A.P.; Picó Segura, C.; Bonet Piña, M.L.; Oliver Vara, P.; Serra Vich, F.; Rodríguez Guerrero, A.M. & Ribot Riutort, J.. *Esteroles vegetales. Tipos, fuentes y mecanismo de acción* en *El libro Blanco de los esteroles vegetales* (2<sup>nd</sup> ed.) España, 2005. pp. 73-87.
- [2] Quinn, P.J. *Membranes as Targets of antimicrobial Lipids*, in Thomar, H. *Lipids and Essentials oils as antimicrobial agents*. John Wiley & Sons, Chichester (US), 2010, pp 1-24.
- [3] O'Brien, R.D. *Fats and Oils: Formulating and Processing for Applications* (3rd ed). CRC Press, Boca Raton (Florida, USA), 2009.
- [4] IUPAC, The nomenclature of steroids (recommendations 1989), International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology. Available. <http://www.chem.qmul.ac.uk/iupac/steroid/> (visitado Agosto 2012).
- [5] Boskou, D. *Composición del aceite de oliva* en: Boskou, D. *Química y Tecnología del aceite de oliva*. AMV Ediciones – Ediciones Mundi-Prensa, Libros, S.A. Madrid, 1998, pp 67-103.
- [6] Akihisa, T.; Kokke, W.C.M.C. & Tamura, T. *Naturally occurring sterols and related compounds from plants*, in *Physiology and Biochemistry of Sterols*. Ed. by Patterson, G.W. & Nes, W.D.; AOCS Press, Champaign, IL, 1991, pp 172-228.
- [7] Liu, K. *Soybeans as a Powerhouse of Nutrients and Phytochemicals*, in Liu, K., *Soybeans as Functional Foods and Ingredients*. AOCS Press, Champaign, (Illinois, USA), 2004, pp 12-34.
- [8] Toivo, J.; Phillips, K.; Lampi, A-M. & Piironen, V. *Determination of sterols in foods: Recovery of free, esterified, and glycosidic sterols*. *Journal of food composition and analysis*, 2001, 14, 631-643.
- [9] Dutta, P.C.; Przybylski, R.; Appelqvist, L.Å. & Eskin, N.A.M. *Formation and analysis of oxidized sterols in frying fat*, in *Deep Frying Practises, Chemistry and Nutrition*. Ed. by Perkins, E.G. & Erickson, M.D.; AOCS Press, Champaign, IL, 1996, 112-150.
- [10] Pizzoferrato, L.; Nicoli, S. & Lintas, C. *GC-MS Characterization and Quantification of Sterols and Cholesterol Oxidation Products*. *Chromatographia*, 1993, 35, pp 269-274.

- [11] Dutta, P.C. & Savage, G.P. *Formation and Content of Phytosterol Oxidation Products in Food*, in Guardiola, F.; Dutta, P.C.; Codony, R. & Savage, G.P. (ed) *Cholesterol and Phytosterol oxidation Products: Analysis, Occurrence, and biological Effects*. AOCS Press, USA, 2002, pp 319-334.
- [12] Boskou, D. *Other important minor constituents*, in *Olive Oil minor constituents and Health*. CRC Press, Boca Raton (Florida, USA), 2009, pp 45-54.
- [13] Parish, E.J.; Boos, T.L. & Li, S.R. *Chemistry of waxes and sterols* in Akoh, C.C.; Min, D.B. *Food Lipids. Chemistry, Nutrition and Biotechnology* (2nd ed). Marcel Dekker, New York (NY, USA), 2002, pp 103-131
- [14] Pollack, O.J. *Reduction of blood cholesterol in man*. *Circulation*, 1953, 7, 702-706.
- [15] Katan, M.B.; Grundy, S.M.; Jones, P.; Law, M.; Miettinen, T. & Paoletti, R. *Efficacy and safety of plant stanols and sterols in the management of blood cholesterol levels*. *Mayo Clin. Proc.*, 2003, 78, 965-978.
- [16] Awad, A.B. & Fink; C.S. *Phytosterols as anticancer dietary components: evidence and mechanism of action*. *J. Nutr.*, 2000, 130, 2127-2130.
- [17] Ikeda, L.; Kawasaki, A.; Samezima, K. & Sugano, M. *Antihypercholesterolemic activity of  $\beta$ -sitostanol in rabbits*. *J. Nutr. Sci. Vitaminol.*, 1981, 27, 243-251.
- [18] Bouic, P.J. *The role of phytosterols and phytosterolins in immune modulation: a review of the past 10 years*. *Curr. Opin. Clin. Nutr. Metab. Care.*, 2001, 4, 471-475.
- [19] Akoh, C.C. & Lai, O-M. (Eds). *Healthful Lipids*. AOCS Press, Urbana (Illinois, USA), 2005.
- [20] Woyengo, T.A. & Ramprasath, V.R. & Jones, P.J. *Anticancer effects of phytosterols*. *Eur. J. Clin. Nutr.*, 2009, 63, 813-820.
- [21] Kritchevsky, D. & Chen, S.C. *Phytosterols-health benefits and potential concerns: a review*. *Nutrition Research*, 2005, 25, 413-428.

- [22] Mellies, M.D.; Glueck, C.J.; Sweeney, C.; Fallat, R.W.; Tsang, T.C. & Ishikawa, T.T. *Plasma and dietary phytosterols in children*. *Pediatrics*, 1976, 57, 60-67.
- [23] Becker, M.; Staab, D. & Von Bergmann, K. *Treatment of severe familial hypercholesterolemia in childhood with sitosterol and siostanol*. *J. Pediatr.*, 1993, 122, 292-296.
- [24] Plat, J.; Brzezinka, H.; Lutjohann, D.; Mensink, R.P. & Von Bergmann, K.S. *Oxidized plant sterols in human serum and lipid infusions as measured by combined gas-liquid chromatography-mass spectrometry*. *J. Lipid Res.*, 2001, 42, 2030-2038.
- [25] Ostlund, R.E, Jr.; Racette, S.B.; Okeke, A. & Stenson, W.F. *Phytosterols that are naturally present in commercial corn oil significantly reduce cholesterol absorption in humans*. *Am. J. Clin. Nutr.*, 2002, 75, 1000-1004.
- [26] Clifton, P. *Plant sterol and stanols—comparison and contrasts. Sterols versus stanols in cholesterol-lowering: Is there a difference?* *Atheroscler.*, 2002, (Suppl. 3), 5-9.
- [27] Kochhar, S.P. *Influence of processing on sterols of edible vegetable oils*. *Prog. Lipid Res.*, 1983, 22, 161-188.
- [28] Weihrauch, J.L. & Gardner, J.M. *Sterol content of foods of plant origin*. *J. Am. Diet. Assoc.*, 1978, 73, 39-47.
- [29] Miettinen, T.A. & Gylling, H. *Regulation of cholesterol metabolism by dietary plant sterols*. *Curr. Opin. Lipidol.*, 1999, 10, 9-14.
- [30] Piironen, V.; Lindsay, D.G.; Miettinen, T.A.; Toivo, J. & Lampi, A.M. *Plant sterols: Biosynthesis, biological function and their importance to human nutrition*. *J. Sci. Food Agric.*, 2000, 80, 936-966.
- [31] Phillips, K.M.; Ruggio, D.M.; Toivo, J.I.; Swank, M.A. & Simpkins, A.H. *Free and sterified sterol composition of edible oils and fats*. *J. Food Comp. An.*, 2002, 15, 123-142.
- [32] Verleyen, T.; Forcades, M.; Verhe, R.; Dewettinck, K.; Huyghebaert, A.; & de Greyt, W. *Analysis of free and sterified sterols in vegetable oils*. *J. Am. Oil Chem. Soc.*, 2002, 79, 117-122.

- [33] Dionisi, F.; Hug, B. & Kamm, W. *Autenticidad de grasas y aceites* in *Fraudes alimentarios* Ducauze, C.J. Ed. Acrilia S.A., Zaragoza (España), 2006, pp 233-294.
- [34] [http://www.consumer.es/web/es/alimentacion/aprender a comer bien/alimentos a debate/20120/03/15/191750](http://www.consumer.es/web/es/alimentacion/aprender_a_comer_bien/alimentos_a_debate/20120/03/15/191750). (Visitado septiembre 2012)
- [35] <http://www.buenasalud.net/tag/oliva/page/2#>. (Visitado septiembre 2012)
- [36] Ramírez-Tortosa, M.; Granados, S. & Quiles, J.L. *Chemical Composition, Types and Characteristcs of Olive Oil* en Quiles, J.L.; Ramírez Tortosa, M.C. & Yaqoob, P. *Olive Oil and Health*. CAB International, Wallingford (UK), 2006. pp 45-62.
- [37] CEE regulation nº 2568/91, European commission 11/07/1991. (DO L248 5/9/1991, pp. 1.
- [38] Peterson, D.W.; Nichols, L.W.; Peck, N.F. & Chaikoff, I.L. *Depression of plasma cholesterol in human subjects consuming butter containing soy sterols*. Fed. Proc., 1956, 65, 569.
- [39] AbuMweis, S.S.; Barake, R. & Jones, P.J. *Plant sterols/stanols as cholesterol lowering agents: a meta-analysis of randomized controlled trials*. Food Nutr. Res., 2008, 52, 1-17.
- [40] Jenkins, D.J.; Kendall, C.W.; Nguyen, T.H.; Marchie, A.; Faulkner, D.A.; Ireland, C.; Josse, A.R.; Vidgen, E.; Trautwein, E.A.; Lapsley, K.G.; Holmes, C.; Josse, R.G.; Leiter, L.A.; Connelly, P.W. & Singer, W. *Effect of plant sterols in combination with other cholesterol-lowering foods*. Metabolism, 2008, 57, 130-139.
- [41] Khandelwal, S.; Demonty, I.; Jeemon, P.; Lakshmy, R.; Mukherjee, R.; Gupta, R.; Snehi, U.; Niveditha, D.; Singh, Y.; Van der Knaap, H.C.M.; Passi, S.J.; Prabhakaran, D. & Reddy, S. *Independent and interactive effects of plant sterols and fish oil n-3 long chain polyunsaturated fatty acids on the plasma lipid profile of mildly hyperlipidaemic Indian adults*. Br. J. Nutr., 2009, 102, 722-732.
- [42] Escurriol, V.; Cofan, M.; Serra, M.; Bullo, M.; Basora, J.; Salas-Salvado, J.; Corella, D.; Zazpe, I.; Martínez-Gonzalez, M.A.; Ruiz-Gutierrez, V.; Estruch, R. & Ros, E. *Serum sterol reponses to increasing plant sterol intake from natural foods in the Mediterranean diet*. Eur. J. Nutr., 2009, 48, 373-382.

- [43] Laakso, P. *Analysis of sterols from various food matrices. Eur. J. Lipid Sci. Technol.*, 2005, 107, 402-410.
- [44] Toivo, J.; Phillips, K.; Lampi, A.M. & Piironen, V. *Determination of sterols in foods: Recovery of free, esterified, and glycosidic sterols. J. of Food Comp. Anal.*, 2001, 14, 631-643.
- [45] Jekel, A.A.; Vaessen, H.A.M.G. & Schothorst, R.C. *Capillary gas chromatographic method for determinig non-derivatized sterols some results for duplicate 24h diet samples collect in 1994. Fresen. J. Anal. Chem.*, 1998, 360, 595-600.
- [46] Spangenberg, B.; Poole, C.F. & Weins, C. *Quantitative Thin-Layer Chromatography. A practical Survey.* Springer, Heidelberg (Germany), 2011.
- [47] Ruiz-Gutiérrez, V. & Pérez-Camino, M.C. *Review: Update on solid-phase extraction for the analysis of lipid classes and related compounds. J. Chromatogr. A*, 2000, 885, 321-341.
- [48] Azadmard-Darmichi, S. *Review : The use of phytosterols as a detection tool for adulteration of olive oil with hazelnut oil. Food Additives and Contaminants*, 2010, 27, 1-10.
- [49] Azadmard-Darmirchi, S. *Novel solid-phase extraction method to separate 4-desmethyl-,4-monomethyl and 4,4'-dimetilesterols in vegetable oils. Journal of Chromatography A*, 2006, 1108, 183-187.
- [50] Phillips, K.M.; Ruggio, D.M.; Toivo, J.I.; Swank, M.A. & Simpkins, A.H. *Free and esterified sterol composition of edible oils and fats. J. Food Compos. Anal.*, 2002, 15, 123-142.
- [51] Cercaci, L.; Rodríguez-Estrada, M.T. & Lercker, G. *Solid-phase extraction-thin-layer chromatography-gas-chromatography method for the detection of hazelnut oil in olive oils by determination of esterified sterols. J. Chromatogr. A*, 2003, 985, 211-220.
- [52] Ballesteros, E.; Gallego, M. & Valcárcel, M. *Simultaneous determination of sterols in edible oils by use of a continuous separation module coupled to a gas chromatograph. Anal. Chim. Acta*, 1995, 308, 253-260.

- [53] Snyder, J.M.; King, J.W.; Taylor, S.L. & Neese, A.L. *Concentration of Phytosterols for Analysis by supercritical fluid extraction*. J. Am. Oil Chem. Soc., 1999, 76, 717-721.
- [54] Brunner, G. *Fractionation of fats with supercritical carbon dioxide*. Eur. J. Lipid Sci. Technol., 2000, 240-244.
- [55] Mendes, M.F.; Pessoa, F.L.P.; Coelho, G.V. & Uller, A.M.C. *Recovery of the high aggregated compounds present in the deodorizer distillate of the vegetable oils using supercritical fluids*. The journal of supercritical fluids, 2005, 34, 157-162.
- [56] Herrero, M.; Mendiola, J.A.; Cifuentes, A. & Ibáñez, E. *Supercritical fluid extraction: Recent advances and applications - A review*. J. Chromatogr. A, 2010, 1217, 2495-2511.
- [57] Lesellier, E. *Analysis of non-saponifiable lipids by super-/subcritical-fluid chromatography*. J. Chromatogr. A, 2001, 936, 201-214.
- [58] Sahena, F.; Zaidul, L.S.M.; Jinap, S.; Karim, A.A.; Abbas, K.A.; Norulaini, N.A.N. & Omar, A.K.M. *Application of supercritical CO<sub>2</sub> in lipid extraction – A review*. Journal of Food Engineering, 2009, 95, 240-253.
- [59] Fair, J.D. & Kornos, C.M. *Flash column chromatograms estimated from thin-layer chromatography data*. J. Chromatogr. A, 2008, 1211, 49-54.
- [60] Kalo, P. & Kuuranne, T. *Analysis of free and esterified sterols in fats and oils by flash chromatography, gas chromatography and electrospray tandem mass spectrometry*. J. Chromatogr. A., 2001, 935, 237-248.
- [61] Sánchez-Machado, D.I.; López-Hernández, J.; Paseiro-Losada, P. & López-Cervantes, J. *An HPLC method for the quantification of sterols in edible seaweeds*. Biomedical Chromatography, 2004, 18, 183-190.
- [62] Cert, A.; Moreda, W. & García-Moreno, J. *Determinación de esteroides y alcoholes triterpénicos en aceite de oliva mediante separación de la fracción por cromatografía líquida de alta eficacia y análisis por cromatografía de gases. Estandarización del método analítico*. Grasas y aceites, 1997, 48(4), 207-218.

- [63] Espejo, J. *Estudio analítico comparado entre el aceite de acebuchina y el aceite de oliva virgen*. Tesis doctoral defendida en la Universidad de Sevilla. 2005.
- [64] Cortés, J.M.; Sánchez, R.; Villén, J.; Vázquez, A. *Analysis of unsaponifiable compounds of edible oils by automated on-line coupling reserved-phase liquid chromatography-gas chromatography using the through oven transfer adsorption desorption interface*. Journal of Agricultural and Food Chemistry, 2006, 54, 6963-6968.
- [65] Moreau, R.A.; Whitaker, B.D. & Hicks, K.B. *Review-Phytosterols, phytostanols, and their conjugates in foods: structural diversity, quantitative analysis, and health-promoting uses*. Progress in Lipid Research, 2002, 457-500.
- [66] Careri, M.; Elviri, L. & Mangia, A. *Liquid Chromatography-UV determination and liquid chromatography-atmospheric pressure chemical ionization mass spectrometric characterization of sitosterol and stigmasterol in soybean oil*. Journal of Chromatography A, 2001, 935, 249-257.
- [67] Zhan, L.F.; Liu, C.F. & Meng, Y.B. *Isolation and determination of sterols in rapeseed oil by HPLC-MS*. Analytical Letters, 2009, 42, 1650-1661.
- [68] Martínez-Vidal, J.L.; Garrido-Frenich, A.; Escobar-García, M.A. & Romero-González, R. *LC-MS Determination of sterols in olive oil*. Chromatographia, 2007, 65, 695-699.
- [69] Amelio, M.; Rizzo, R. & Varazini, F. *Determination of sterols, erythrodiol, uvaol and alkanols in olive oils using combined solid-phase extraction, high-performance liquid chromatographic and high-resolution chromatographic techniques*. J. Chromatogr., 1992, 606, 179-185.
- [70] Jiménez de Blas, O. & del Valle González, A. *Determination of sterols by capillary column gas chromatography. Differentiation among different types of olive oil: virgin, refined and solvent-extracted*. AOCS, 1996, 73, 1685-1689.
- [71] Lanuzza, F.; Micali, G. & Calabró, G. *On-line HPLC-HRGC coupling and simultaneous transfer of two different LC fractions: determination of aliphatic alcohols and sterols in olive oil*. J. High Resol. Chromatogr., 1996, 19, 424-448.
- [72] Aparicio, R. & Aparicio-Ruiz, R. *Authentication of vegetable oils by chromatographic techniques*. J. Chromatogr. A, 2000, 881, 93-104.



- [73] Lercker, G. & Rodríguez-Estrada, M.T. *Chromatographic analysis of unsaponifiable compounds of olive oils and fat-containing foods*. J. Chromatogr. A, 2000, 881, 105-129.
- [74] Parcerisa, J.; Casals, I.; Botella, J.; Codony, R. & Rafecas, M. *Analysis of olive and hazelnut oil mixtures by high-performance liquid chromatography-atmospheric pressure chemical ionisation mass spectrometry of triacylglycerols and gas-liquid chromatography of non-saponifiable compounds (tocopherols and sterols)*. J. Chromatogr. A., 2000, 881, 149-159.
- [75] Bohacenko, I. & Kopicová, Z. *Detection of olive oils authenticity by determination of their sterol content using LC/GC*. Czech J. Food Sci., 2001, 19, 97-103.
- [76] Giacometti, J. *Determination of aliphatic alcohols, squalene,  $\alpha$ -tocopherol and sterols in olive oils: direct method involving gas chromatography of the unsaponifiables fraction following silylation*. Analyst, 2001, 126, 472-475.
- [77] Kamm, W.; Dionisi, F.; Hischenhuber, C.; Schmarr, H.G.; Engel, K.H. *Rapid detection of vegetable oils in milk fat by on-line LC-GC analysis of  $\beta$ -sitosterol as marker*. Eur. J. Lipid Sci. Technol., 2002, 104, 756-761.
- [78] Sánchez-Casas, J.; Osorio Bueno, E.; Montañó García, A.M.; Martínez Cano, M. *Sterol and erythrodiol+uvaol content in virgin olive oils from cultivars of Extremadura (Spain)*. Food Chem., 2004, 87, 225-230.
- [79] Rivera del Álamo, R.M.; Fregapane, G.; Aranda, F.; González-Alonso, S.; Salvador, M.D. *Sterol and alcohol composition of Cornicabra virgin olive oil: the campesterol content exceeds the upper limit of 4% established by EU regulations*. Food Chem., 2004, 84, 533-537.
- [80] Alves, M.R.; Cunha, S.C.; Amaral, J.S.; Pereira, J.A.; Oliveira, M.B. *Classification of PDO olive oil on the basis of their sterol composition by multivariate analysis*. Anal. Chim. Acta, 2005, 549, 166-178.
- [81] Sivakumar, G.; Bricoli Bati, C.; Perri, E.; Uccella, N. *Gas chromatography screening of bioactive phytosterols from mono-cultivar olive oils*. Food Chemistry, 2006, 95, 525-528.

- [82]** Güll, M.K.; Secker, M. *Comparative analysis of phytosterols components from rapeseed (*Brassica napus* L.) and olive (*Olea europaea* L.) varieties.* Eur. J. Lipid Sci. Technol., 2006, 108, 759-765.
- [83]** Azadmard-Darmichi, S.; Dutta, P.C. *Free and sterified 4,4'-dimethylsterols in hazelnut oil and their retention during refining processes.* J. Amer. Oil Chem. Soc., 2007, 84, 297-304.
- [84]** Cercaci, L.; Passalacqua, G.; Poerio, A.; Rodríguez-Estrada, M.T.; Lercker, G. *Composition of total sterols (4-desmethyl-sterols) in extravirgin olive oils obtained with different extraction technologies and their influence on the oil oxidative stability.* Food Chem., 2007, 84, 297-304.
- [85]** Haddada, F.M.; Manai, H.; Oueslati, I.; Daoud, D.; Sánchez, J.; Osorio, E.; Zarrouk, M. *Fatty acid, triacylglycerol and phytosterol composition in six Tunisian olive varieties.* J. Agric. Food Chem., 2007, 55, 10941-10946.
- [86]** Ben Temime, S.; Manai, H.; Methenni, K.; Baccouri, B.; Abaza, L.; Daoud, D.; Sánchez-Casas, J.; Osorio Bueno, E.; Zarrouk, M. *Sterolic composition of Chètoui virgin olive oil: influence of geographical origin.* Food Chem., 2008, 110, 368-374.
- [87]** Jahouach-Rabai, W.; Trabelsi, M.; Van Hoed., V; Adams, A.; Verhé, R.; de Kimppe, N.; Frikha, M.H. *Influence of bleaching by ultrasound on fatty acids and minor compounds of olive oil. Qualitative and quantitative analysis of volatile compounds (By SPME coupled to GM/MS).* Ultrason. Sonochem., 2008, 15, 590-597.
- [88]** López-López, A.; Montaña, A.; Ruíz-Méndez, M.V.; Garrido-Fernández, A. *Sterols, fatty alcohols and triterpenic alcohols in comercial table olives.* J. Am. Oil Chem. Soc., 2008, 85, 253-262.
- [89]** Azadmard-Darmichi, A.; Dutta, P.C. *Stability of minor lipid components with emphasis on phytosterols during chemical interesterification of a blend of refined olive oil and palm stearin.* J. Am. Oil Soc., 2008, 85, 13-21.
- [90]** Sakouhi, F.; Absalon, C.; Harrabi, S.; Vitro, C.; Sebei, K.; Boukhchina, S.; Fouquet, E.; Kallel, H. *Dynamic accumulation of 4-desmethylsterols and phytosterols during ripening of Tunisian Meski olives (*Olea europaea* L.).* Food Chem., 2009, 112, 897-902.
- [91]** Zarrouk, W.; Carrasco-Pancorbo, A.; Zarrouk, M.; Segura-Carretero, A.; Fernández-Gutierrez, A. *Multicomponent analysis (sterols, tocopherols and*

- triterpenic dialcohols) of the unsaponifiable fraction of vegetable oils by liquid chromatography-atmospheric pressure chemical ionization-ion trap mass spectrometry. Talanta, 2009, 80, 924-934.*
- [92] López-López, A.; Rodríguez-Gomez, F.; Victoria Ruiz-Mendez, M.; Cortes-Delgado, A.; Garrido-Fernandez, A. *Sterols, fatty alcohol and triterpenic alcohol changes during ripe table olive processing. Food Chemistry, 2009, 117, 127-134.*
- [93] Sakouhi, F.; Absalon, C.; Sebeid, K.; Fouquet, E.; Boukhchina, S.; Kallel, H. *Gas chromatographic-mass spectrometric characterisation of triterpene alcohols and monomethylsterols in developing Olea europaea L. fruits. Food Chem., 2009, 116, 345-350.*
- [94] Gracia, M.S.; Royo, A.; Guillén, M. *Composición química de aceites de las variedades Arbequina y Empeltre cultivadas en regadío. Grasas y Aceites, 2009, 60, 321-329.*
- [95] Rocco, A.; Fanali, S. *Analysis of phytosterols in extra-virgin olive oil by nano-liquid chromatography. J. of Chromatogr. A, 2009, 1216, 7173-7178.*
- [96] Lerma-García, M.J.; Concha-Herrera, V.; Herrero-Martínez, J.M.; Simó-Alfonso, E.F. *Classification of extra virgin olive oil produced at La Comunitat Valenciana according to their genetic variety using sterol profiles established by high-performance liquid chromatography with mass spectrometry detection. J. Agric. Food Chem., 2009, 57, 10512-10517.*
- [97] Alonso-Salces, R.M.; Héberger, K.; Holland, M.V.; Moreno-Rojas, J.M.; Mariano, C.; Bellan, G.; Reniero, F.; Guillou, C. *Multivariate analysis of NRM fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes. Food Chem., 2010, 118, 956-965.*
- [98] Cañabate-Díaz, B.; Segura Carretero, A.; Fernández-Gutiérrez, A.; Belmonte Vega, A.; Garrido Frenich, A.; Martínez Vidal, J.L.; Duran Martos, J. *Separation and determination of sterols in olive oil by HPLC-MS. Food Chemistry, 2007, 102, 593-598.*
- [99] Segura Carretero, A.; Carrasco-Pancorbo, A.; Cortacero, S.; Gori, A.; Cerretani, L.; Fernández-Gutiérrez, A. *A simplified method for HPLC-MS analysis of sterols in vegetable oil. Eur. J. Lipid Sci. Technol., 2008, 110, 1142-1149.*
- [100] Lerma-García, M.J.; Simó-Alfonso, E.F.; Méndez, A.; Llibería, J.L.; Herrero-Martínez, J.M. *Fast separation and determination of sterols in vegetable*

*oils by ultra-performance liquid chromatography with atmospheric pressure chemical ionization mass spectrometry detection.* Journal of Agricultural and Food Chemistry, 2010, 58, 2771-2776.

- [101]** Grob, K.; Lanfranchi, M. & Mariani, C. *Determination of free and esterified sterols and wax esters in oils and fats by coupled liquid chromatography-gas chromatography.* J. Chromatogr. A, 1989, 471, 397-405
- [102]** Grob, K.; Lanfranchi, M. & Mariani, C. *Evaluation of olive oils through the fatty alcohols, the sterols and their esters by coupled LC-GC.* J. American Oil Chemists' Society, 1990, 67, 626-634
- [103]** Biedermann, M.; Grob, K. & Mariani, C. *Transesterification and on line LC-GC for determining the sum of free and esterified sterols in edible oils and fats.* Fat Sci. Technol., 1993, 95, 127-133.
- [104]** Señoráns, F.J.; Villén, J.; Tabera, J. & Herraiz, M. *Simplex optimization of the direct analysis of free sterols in sunflower oil by on-line coupled reversed phase liquid chromatography-gas chromatography.* J. Agric. Food Chem., 1998, 46, 1022-1026.
- [105]** Villén, J.; Blanck, G.P.; Ruiz del Castillo. M.L. & Herraiz, M. *Rapid and simultaneous analysis of free sterols, tocopherols and squalene in edible oils by coupled reserved phase liquid chromatography-gas chromatography.* J. Agric. Food Chem., 1998, 46, 1419-1422.
- [106]** Alario, J.; Pérez, M.; Vazquez, A. & Villén, J. *Very-Large Volume sampling of wáter in gas chromatography using the trough oven transfer adsorption desorption (TOTAD) interface for pesticide-residue analysis..*J. Chromatogr. Sci., 2001, 39, 65-69.
- [107]** Cortés, J.M.; Sánchez, R.; Villén, J. & Vázquez, A. *Analysis of unsaponifiable compounds of edible oils by automated on-line coupling reserved-phase liquid chromatography-gas chromatography using the through oven transfer adsorption desorption interface.* J. Agric. Food Chem., 2006, 54, 6963-6968.
- [108]** Toledano, R.M.; Cortés, J.M.; Rubio-Moraga, A.; Villén, J. & Vázquez, A. *Analysis of free and esterified sterols in edible oils by online reversed phase liquid chromatography-gas chromatography (RPLC-GC) using the through oven transfer adsorption desorption (TOTAD) interfase.* Food Chemistry, 2012, 135, 610-615.
- [109]** Toledano, R.; Cortés, J.M.; Andini, J.; Vázquez, A. & Villén, J. *On-line derivatization with on-line coupled normal phase liquid cromatography-gas*

- chromatography using the through oven transfer adsorption desorption interface: Application to the analysis of total sterols in edible oils.* J. Chromatogr. A, 2012, 1256, 191-196.
- [110] REGLAMENTO (CEE) 2568/91, de la Comisión, de 11 de julio (DOCE L 248, de 05.09.1991) relativo a las *características de los aceites de oliva y de los aceites de orujo de oliva y sobre sus métodos de análisis*.
- [111] COI/T.20/Doc. nº 10 / Rev.1 (2001). *Determinación de la composición y del contenido en esteroides mediante cromatografía de gases con columna capilar*. Consejo Oleícola Internacional, Madrid.
- [112] COI/T.20/Doc. nº 30 (2009) *Determinación de la composición y del contenido de esteroides y dialcoholes triterpénicos mediante cromatografía de gases con columna capilar*. Consejo Oleícola Internacional, Madrid.
- [113] COI/T.20/DOC. NO 30 (2011) *Determinación de la composición y del contenido de esteroides y dialcoholes triterpénicos mediante cromatografía de gases con columna capilar*. Consejo Oleícola Internacional, Madrid.
- [114] Cuadros-Rodríguez, L.; Bagur-González, M. G.; Sánchez-Viñas, M.; González-Casado, A.; Gómez-Sáez, A.M.; "Principles of analytical calibration/quantification for the separation sciences", J. of Chromatography A 1158 (2007) 33–46.
- [115] Costin, C.D.; Hansen, S.L.; Chambers, D.P. *Using theoretical correction factors for quantitative analysis of sterols and sterols concentrates.* J. Am. Oil Chem. Soc., 2009, 86, 117-118.
- [116] Covas, M.I.; Ruiz-Gutiérrez, V.; de la Torre, R.; Kafatos, A.; Lamuela-Raventós, R.M.; Osada, J.; Owen, R.W.; Francesco, V. *Minor components of olive oil: evidence to date of health benefits in humans.* Nutr. Rev., 2006, 64 (Suppl 1), 20-30.
- [117] Verleyen, T.; Forcades, M.; Verhe, R.; Dewettinck, K.; Huyghebaert, A.; de Greyt, W. *Analysis of free and esterified sterols in vegetable oils.* JAOCS, 2002, 79, 117-122.
- [118] Montanari, L.; Sensidoni, A.; Kikic, I., in Reverchon, E; Ed. Proceedings of the 4th Italian Conference on Supercritical Fluids and Their Applications, 1997, Università di Perugia, Perugia, Italy, pp. 89-95.

- [119] Toivo, J.; Piironen, V.; Kalo, P.; Varo, P. *Gas chromatographic determination of major sterols in edible oils and fats using solid-phase extraction in sample preparation. Chromatographia*, 1998, 48, 745-750.
- [120] Aitzetmüller, K.; Brühl, L.; Fiebig, H.J. *Analysis of sterol content and composition in fats and oils by capillary-gas liquid chromatography using an internal standard. Comments on the german sterol method. Fett/Lipid*, 1998, 100, 429-435.
- [121] Panagiotoupoulou, P.M.; Tsimidou, M. *Solid phase extraction: applications to the chromatographic analysis of vegetable oils and fats. Grasas y Aceites*, 2002, 53, 84-95.
- [122] Saavage, G.P.; McNeil, D.L.; Dutta, P.C. *Lipid composition and oxidative stability of oils in hazelnuts (Corylus avellana L.) grown in New Zealand. JAOCS*, 1997, 74, 755-759.
- [123] Plaza, L.; Sánchez-Moreno, C.; de Pascual-Teresa, S.; de Ancos, B.; Cano, M.P. *Fatty acids, sterols, and antioxidant activity in minimally processed avocados during refrigerated storage. J. Agric. Food Chem.*, 2009, 57, 3204-3209.
- [124] Lambelet, P.; Grandgirard, A.; Gregorio, S.; Juaneda, P.; Sebedio, J.L.; Bertoli, C. *Formation of modified fatty acids and oxyphytosterols during refining of low erucic acid rapessed oil. J. Agric. Food Chem.*, 2007, 55, 10787-10794.
- [125] Turan, S.; Topcu, A.; Karabulut, I.; Vural, H.; Hayaloglu, A.A. *Fatty acid, triacylglycerol, phytosterol and tocopherol variations in kernel oil of Malaysia apricots from Turkey. J. Agril. Food Chem.*, 2007, 55, 10787-10794.
- [126] Salta, F.N.; Kalogeropoulos, N.; Karavanou, N.; Andrilopoulos, N.K. *Distribution and retention of phytosterols in frying oils and fried potatoes during repeated deep and pan frying. Eur Food Res. Technol.*, 2008, 227, 391-400.
- [127] Tabee, E.; Azadmard-Damirchi, S.; Jaegerstad, M.; Dutta, .PC. *Effects of  $\alpha$ -tocopherol on oxidative stability and phytosterol oxidation during heating in some regular and high-oleic vegetable oils. J. Am. Oil Chem. Soc.*, 2008, 85, 857-867.

- [128] Winkler, J.K.; Warner, K. *The effect of phytosterol concentration on oxidative stability and thermal polymerization of heated oils*. Eur. J. Lipid Sci. Technol., 2008, 110, 455-464.
- [129] Tabee, E.; Jägerstad, M.; Dutta, P.C. *Lipid and phytosterol oxidation products in commercial potato crisps commonly consumed in Sweden*. Eur. Food Res Technol., 2008, 227, 745-755.
- [130] Mitei, Y.C.; Ngila, J.C.; Yeboah, S.O.; Wessjohann, L.; Schmidt, J. *Profiling of phytosterols, tocopherols and tocotrienols in selected seed oils from Botswana by GC-MS and HPLC*. J. Am. Oil Chem. Soc., 2009, 86, 617-625.
- [131] Dutta, P.C.; Appelqvist, L.A. *Studies on phytosterol oxides. I: Effect of storage on the content in potato chips prepared in different vegetable oils*. JAOCS, 1997, 74, 647-657.
- [132] Sato, H.; Ito, K.; Sakai, K.; Morinaga, Y.; Sukegawa, E.; Kitamura, T.; Shimasaki, H.; Itakura, H. *Effects of soybean-germ oil on reducing serum cholesterol level*. J Oleo Sci., 2001, 50, 649-655.
- [133] Durant, A.A.; Dumont, M.J.; Narine, S.S. *In situ silylation for the multicomponent analysis of canola oil by-products by gas chromatography-mass spectrometry*. Anal. Chim. Acta, 2006, 559, 227-233.
- [134] Sivakumar, G.; Bricoli Bati, C.; Perri, E.; Uccella, N. *Gas chromatography screening of bioactive phytosterols from mono-cultivar olive oils*. Food Chem., 2006, 95, 525-528.
- [135] Duarte, A.C. & Capelo, S. J. Liq. Chrom. Rel. Tech, 2006, 29, 1443.
- [136] Harynuk, J.; de la Mata, A.P. & Sinkov, N., in: Varmuza, K. (Ed.), *Application of chemometrics to the interpretation of analytical separations data*. Chemometrics, InTech – Open Acces Publishers, Rijeka, Croatia, 2012.
- [137] Marini, F.; Balestrieri, F.; Bucci, R.; Magri, A.D.; Magri, A.L. & Marini, D. Chemom. Intell. Lab. Syst., 2004, 73, 85.
- [138] Daszykowski, M.; Walczak, B. Trends Anal. Chem., 2006, 25, 1081.
- [139] Skov, T. & Bro, R. Anal, Bioanal. Chem. 2008, 390, 281.

- [140] Savorani, F.; Tomasi, G. & Engelsen, S.B. *J. Magn. Reson.* 2010, 202, 190.
- [141] Andersen, C.M. & Bro, R. *J. Chemometr.* 2010, 24, 78.
- [142] Daszykowski, M.; Walczak, B. *Use and abuse of chemometrics in chromatography.* *Trends in Analytical Chemistry*, 2006, 25(11), 1081-1096.
- [143] Amigo, J.M.; Skov, T.; Bro, R. *ChroMATHography: Solving Chromatographic issues with mathematical models and intuitive graphics.* *Chemical Reviews*, 2010, 110, 4582-4605.
- [144] Cserhàti, T. *Multivariate Methods in Chromatography: A Practical Guide.* John Wiley & Sons, Chichester (UK), 2008.
- [145] Bucci, R.; Magrí, A.D.; Magrí, A.L.; Marini, D.; Marini, F. *Chemical authentication of extra virgin olive oil varieties by supervised chemometric procedures.* *Journal of agricultural and food chemistry*. 2002, 50, 413-418.
- [146] Sánchez Casas, J.; Osorio Bueno, E.; Montañó García, A.M.; Martínez Cano, M. *Sterol and erythrodiol + uvaol content of virgin olive oils from cultivars of Extremadura (Spain).* *Food Chemistry*. 2004, 87, 225-230.
- [147] Amaral, S.S.; Mafra, I.; Oliveira, M.B.P.P. *Characterization of three Portuguese varietal olive oils based on fatty acids, triacylglycerols, phytosterols and vitamin E profiles: Application of chemometrics.* Chapter 63.
- [148] Alonso-Salces, R.M.; Héberger, K.; Holland, M.V.; Moreno-Rojas, J.M.; Mariani, C.; Bellan, G.; Reniero, F.; Guillou, C. *Multivariate analysis of NMR fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes.* *Food Chemistry*. 2010, 956-965.
- [149] Lerma-García, M.J.; Simó-Alfonso, E.F.; Méndez, A.; Lliberia, J.L.; Herrero-Martínez. *Classification of extra virgin olive oils according to their genetic variety using linear discriminant analysis of sterol profiles established by ultra-performance liquid chromatography with mass spectrometry detection.* *Food research international*. 2011, 44, 103-108.
- [150] García-González, D.L.; Luna, G.; Morales, M.T.; Aparacio, R. *Stepwise geographical traceability of virgin olive oils by chemical profiles using artificial neural network models.* *Eur. J. Lipid Sci. Technol.* 2009, 111, 1003-1013.



- [151] Crews, C.; Hough, P.; Godward, J.; Brereton, P.; Lees, M.; Guiet, S.; Winkelmann, W. Study of the main constituents of some authentic walnuts oils. *J. Agric. Food Chem.* 2005, 53, 4853-4860.
- [152] Turan, S.; Topcu, A.; Karabulut, I.; Vural, H.; Hayaloglu, A.A. Fatty acid, triacylglycerol, phytosterol, and tocopherol variation in kernel oil of Malatya apricots from Turkey. *J. Agric. Food Chem.* 2007, 55, 10787-10794.
- [153] López-López, A.; Rodríguez-Gómez, F.; Ruíz-Méndez, M.V.; Cortés-Delgado, A.; Garrido-Fernández, A. Sterols, fatty alcohol and triterpenic alcohol changes during ripe table olive processing. *Food chemistry.* 2009, 117, 127-134.
- [154] López-López, A.; Montañó, A.; Ruíz-Méndez, M.V.; Garrido-Fernández, A. Sterols, Fatty alcohols, and triterpenic alcohols in commercial table olives. *J. Am. Oil Chem. Soc.* 2008, 85, 253-262.
- [155] Nagy, K.; Bongiorno, D.; Avellone, G.; Agozzino, P.; Ceraulo, L.; Vékey, K. High performance liquid chromatography-mass spectrometry based chemometric characterization of olive oils. *J. Chromatography A.* 2005, 1078, 90-97.
- [156] Ruíz-Méndez, M.V.; López-López, A.; Garrido-Fernández, A. Characterization and chemometric study of crude and refined oils from table olive by-products. *Eur. J. Lipid Sci. Technolog.* 2008, 110, 537-546.
- [157] Boggia, R.; Zunin, P.; Lanteri, S.; Rossi, N.; Evangelisti, F. Classification and class-modeling of "Riviera Ligure" extra-virgin olive oil using chemical-physical parameters. *J. Agric. Food Chem.* 2002, 50, 2444-2449.
- [158] Marini, F.; Balestrieri, F.; Bucci, R.; Magrì, A.D.; Marini, D. Supervised pattern recognition to authenticate Italian extra virgin olive oil varieties. *Chemometrics and Intelligent Laboratory Systems*, 2004, 73, 85-94.
- [159] Matos, L.C.; Cunha, S.C.; Amaral, J.S.; Pereira, J.A.; Andrade, P.B.; Seabra, R.M.; Oliveira, B.P.P. Chemometric characterization of three varietal olive oils (Cvs. Cobrançosa, Madural and Verdeal Transmontana) extracted from olives with different maturation indices. *Food chemistry*, 2007, 102, 406-414.
- [160] Mariani, C.; Bellan, G.; Lestini, E.; Aparacio, R. The detection of the presence of hazelnut oil in olive oil by free and esterified sterols. *Eur. Food Res. Technol.* 2006, 223, 655-661.

- [161]** Arena, E.; Campisi, S.; Fallico, B.; Maccarone, E. Distribution of fatty acids and phytosterols as a criterio to discriminate geographic origin of pistachio seeds. *Food Chemistry*. 2007, 104, 403-408.
- [162]** Galeano Diaz, T.; Durán Merás, I.; Sánchez Casas, J.; Franco, M.F.A. Characterization of virgin olive oils according to its triglycerides and sterols composition by chemometric method. *Food Control*, 2005, 16, 339-347.
- [163]** Mannina, L.; Patumi, M.; Proietti, N.; Bassi, D.; Segre, A.L. Geographical characterization of Italian Extra Virgin Olive oils using high-field  $^1\text{H}$  NMR spectroscopy. *Journal of agricultural and food chemistry*. 2001, 6, 2688-2696.
- [164]** Ouni, Y.; Flamini, G.; Youssef, N.B.; Guerfel, M.; Zarrouk, M. Sterolic composition and triacylglycerols of Oueslati virgin olive oil: comparison among different geographic areas. *Int. J. Food Sci. Technol.* 46 (2011) 1747.
- [165]** Parcerisa, J.; Casals, I.; Botella, J.; Codony, R.; Rafecas, M. Analysis of olive and hazelnut oil mixtures by high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry of triacylglycerols and gas chromatography of non-saponifiable compounds (tocopherols and sterols). *J. Chromatogr. A* . 2000, 881, 149-158.
- [166]** Brescia, M.A.; Alviti, G.; Liuzzi, V.; Sacco, A. Chemometric Classification of olive cultivars based on compositional data of oils *J. Am. Oil Chem. Soc.* 2003, 80, 945.



# **METODOLOGÍA Y RESULTADOS**



# **CAPÍTULO I**



# Bases quimiométricas para el estudio del perfil de esteroides en aceites

La Quimiometría fue definida por The Chemometrics Society en 1975 como "la disciplina química que utiliza métodos matemáticos y estadísticos para diseñar o seleccionar procedimientos de medida y experimentos óptimos, y para proporcionar la máxima información química mediante el análisis de datos químicos".

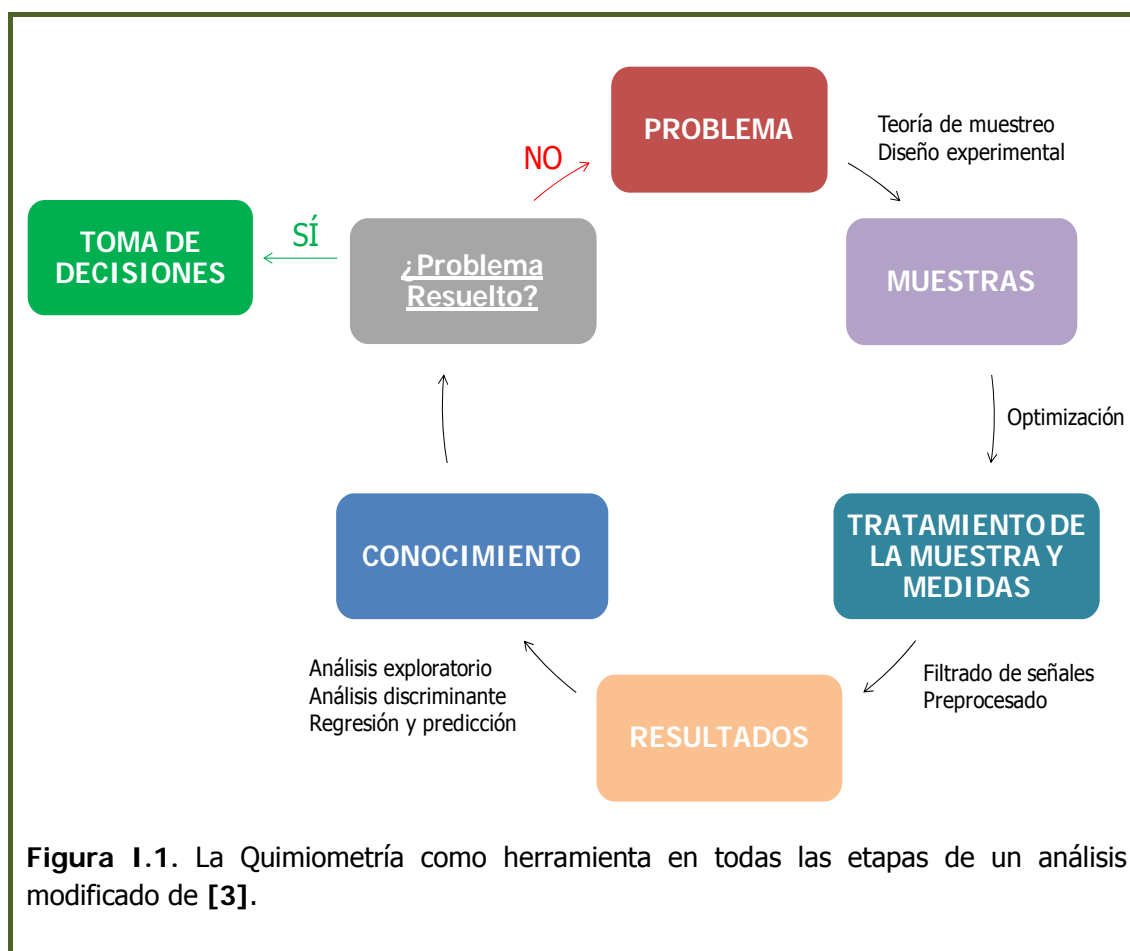
Otra definición fue dada por el profesor Massart, D.L. [1] en 1997 que la define como "la disciplina química que utiliza la Matemática, la Estadística y la lógica formal para diseñar o seleccionar procedimientos experimentales óptimos, proporcionar la máxima información química relevante a partir del análisis de señales analíticas y, finalmente, adquirir conocimiento de los sistemas químicos". Esta definición es la más global.

Siebert [2] ha proporcionado recientemente una definición sugerente: "La quimiometría es la aplicación de métodos matemáticos y estadísticos así como de los principios de la buena ciencia de la medida para extraer de forma eficiente información útil de datos químicos".

En definitiva la Quimiometría se sitúa en un campo interdisciplinar y su éxito depende de los problemas químicos que sea capaz de resolver. Se puede



considerar como una rama aplicada y especializada de ambas, la Química Analítica y la Estadística aplicada, con la función análoga a la que desempeñan disciplinas como la Biometría, Sociometría, Econometría o Psicometría en relación a la Biología, Sociología, Economía o Psicología, entre otras. Además, contribuye, a hacer realidad diversas tendencias de la química analítica actual: rapidez, abaratamiento de costes, miniaturización o transportabilidad de instrumentos. En la Figura I.1 se muestra un esquema de todas las etapas del uso de la Quimiometría para asegurar la calidad de los resultados.



**Figura I.1.** La Quimiometría como herramienta en todas las etapas de un análisis modificado de [3].

Se pueden encontrar métodos que ayuden a los químicos analíticos para hacer frente a todas las etapas de los procedimientos analíticos, a partir del diseño de un experimento para la extracción de la información y la toma de decisión final. Se aplica a cualquier tipo de experimento analítico y a cualquier tipo de señal instrumental pero, existen problemas asociados con los tipos específicos de señales instrumentales o con las técnicas analíticas que necesitan de un tratamiento especial y se deben de elegir bien los métodos a emplear.

El uso indebido y el abuso de la quimiometría aparecen cuando no se utiliza adecuadamente. Esto ocurre en muchos campos tales como, alimentación,

empresas farmacéuticas, etc. Se debe a que aplican de forma inadecuada los métodos, el modelo de validación es incorrecto, hay problemas con la representación de los datos, no hay buena comparación con diferentes métodos, etc. Ya que quieren "utilizar" la quimiometría sin tener una buena formación matemática o computacional, sin tener conocimiento previo del tema y querer recoger toda la información en poco tiempo.

## I.1. RECONOCIMIENTO DE PAUTAS

Los objetivos principales para aplicar métodos multivariantes en química analítica se dirigen al agrupamiento y clasificación de objetos (muestras, compuestos o materiales), así como en el modelado de las relaciones entre diferentes datos analíticos [4].

Se pueden dividir en:

- A. Agrupamiento o clustering.
- B. Clasificación.
- C. Calibración.
- D. Redes neuronales.
- E. Multi-vias.

Uno de los principales objetivos del uso de métodos multivariante consiste en permitir visualizar los datos con más de 3 variables medidas. La base matemática será la representación y manipulación de los datos por vectores y matrices. Se definirán las distancias y similaridades/similitudes en el espacio multidimensional y serán necesarios los métodos de proyección de datos multivariante de 2 o 3 dimensiones en el espacio.

En general, los datos analíticos se pueden organizar como una matriz  $X$  de  $N$  objetos (filas) y  $K$  características (columnas). Los objetos son las muestras. Las características o variables de esos objetos son las pautas elementales, espectros, cromatogramas, señal, características o rasgos estructurales o propiedades físicas. La matriz de datos  $X$  será del orden  $N \times K$ .

Se define clase como colección de objetos que tienen características similares. Un patrón o pauta de un objeto es la colección de rasgos característicos.

## I.2. MÉTODOS NO SUPERVISADOS

El agrupamiento de los datos analíticos es posible ya sea por medio de métodos de agrupamiento o por proyección de los datos de alta dimensionalidad a un espacio de menor dimensionalidad.

Los métodos de proyección están dirigidos a la proyección de los datos originales a partir de un espacio dimensional en una línea, un plano, o un sistema de coordenadas 3D.

En la actualidad, la proyección de datos se realiza principalmente por métodos como: Análisis de componentes principales (ACP), Análisis Factorial (AF), descomposición de valor singular (SVD), proyección de vector propio, (erradicación de rango).

### A). Análisis de Componentes Principales (ACP)

La idea clave es la aproximación de la matriz original  $X$  por un producto de dos matrices pequeñas (matrices de scores y loadings) de acuerdo a:

$$X = TP^t$$
$${}_N X^K = {}_N T^A {}_A P^{tK} + {}_N E^K$$

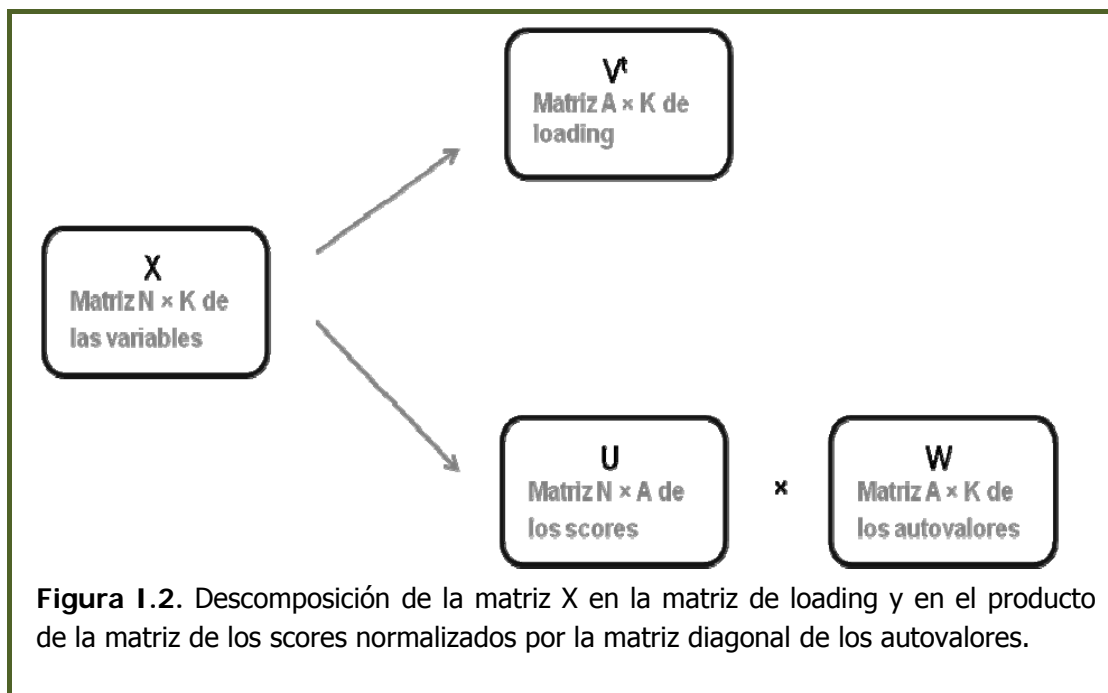
Donde  $X$  es la matriz de datos original formada por  $N$  filas (objetos) y  $K$  columnas (variables);  $T$  es la matriz de los *scores* con  $N$  filas y  $A$  columnas (número de componentes principales);  $P$  es la matriz de *loading* con  $K$  columnas y  $A$  filas;  $E$  es el error de la matriz que es  $N \times K$ ; y  $t$  es la transpuesta de la matriz.

El método más simple para el ACP es el algoritmo NIPALS (para diagonalizar la matriz), aunque existen otros como el algoritmo descomposición en valores singulares (SVD), que es una especie de análisis de vectores propios. La formulación del SVD es:

$$X = UWV^t$$

...Dónde  $\mathbf{U}$  contiene los mismos vectores columna como  $\mathbf{T}$  en la de arriba descrita pero normalizado a una longitud;  $\mathbf{W}$  es la matriz diagonal que contiene la raíz cuadrada de los autovalores o valores simples y  $\mathbf{V}^t$  es igual que  $\mathbf{P}^t$ .

Esta ecuación indica cómo se ha descompuesto  $\mathbf{X}$  en las matrices normalizadas de los *scores* y de *loading*,  $\mathbf{U}$  y  $\mathbf{V}$ , y en el conjunto de autovalores recogidos es  $\mathbf{W}$ , tal y como se esquematiza en la Figura I.2.



Existen otros algoritmos clásicos [5] que se emplean para hacer un ACP tales como el método de poder (POWER) y la descomposición del autovalor (EVD). Además de estos algoritmos clásicos también se pueden usar los algoritmos Kernel [6].

El algoritmo SVD es el que se emplea por defecto en el software PLS\_Toolbox. Además es más eficiente que el NIPALS, el más rápido, sencillo y fácil de programar y es un método no secuencial.

En resumen, el ACP trata de definir, a partir de un conjunto de variables (observables), otro conjunto de variables (no observables), obtenido por combinación lineal de las variables originales, pero que no aparecen correlacionadas entre sí, siendo posible extraer de ellas una información discriminada de algunas características del sistema en estudio. Se obtienen tantas

nuevas variables como variables originales existen, aunque posteriormente el número se pueda reducir, es decir, se utiliza el ACP como una forma de reducción de variables reduciendo el conjunto de datos grande original a uno más pequeño siendo más manejable para que pueda ser interpretado más fácilmente (Figura I.3). La finalidad del ACP es determinar la matriz de los pesos o coeficientes y los autovalores (varianzas de las nuevas variables).

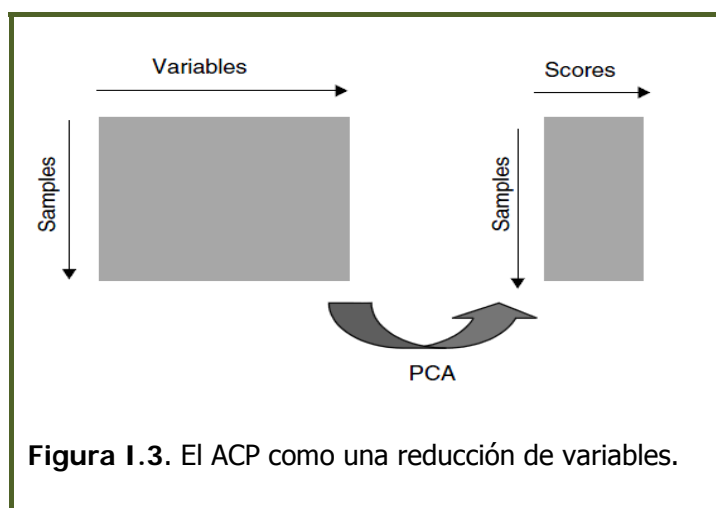


Figura I.3. El ACP como una reducción de variables.

### **Elección Número de componentes principales**

La elección del número de componentes principales es la primera etapa del ACP para llevarlo a cabo siendo la más importante y la más crucial. Es una decisión que dependerá de cuánta información se esté dispuesta a sacrificar, es una cuestión de criterio personal. Si el número de variables no es muy elevado, pueden intentarse varios tanteos incrementando el número de componentes extraídos hasta establecer el modelo más adecuado. Además es posible establecer algunos criterios para su elección:

- Porcentaje de varianza explicada.
- Criterio de autovalor 1.
- Test de sedimentación.
- Validación cruzada (cross-validation).

#### 1. Porcentaje de varianza explicada:

El porcentaje más adecuado varía ampliamente dependiendo del problema, pero este procedimiento puede ser útil cuando se realizan estudios sobre problemas de

un mismo tipo, para los cuales se ha establecido previamente un porcentaje adecuado de varianza explicada acumulada, tal como, por ejemplo un 90%.

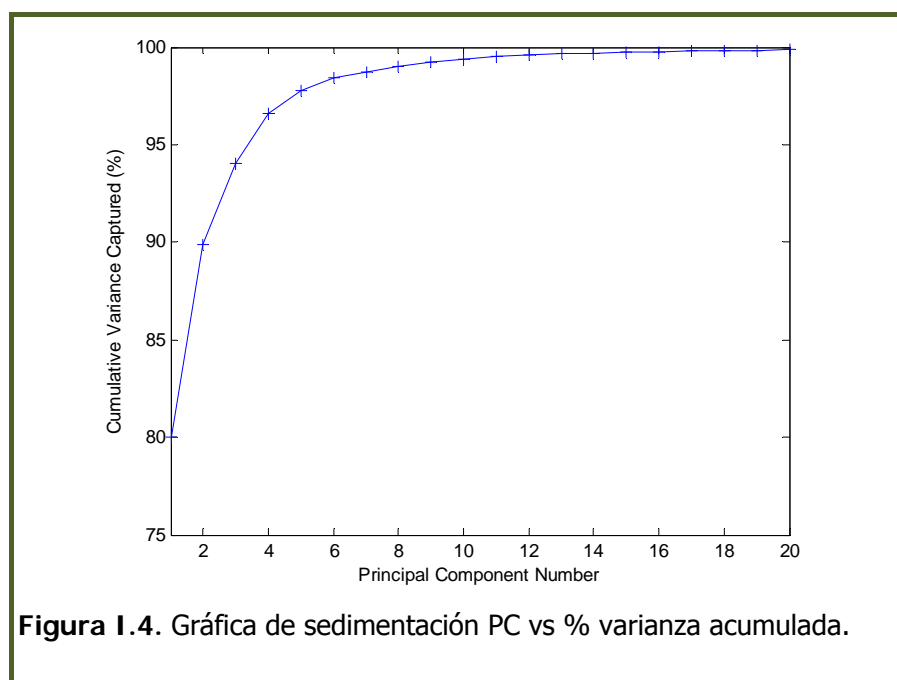
En la Tabla I.1 se muestra un ejemplo de cómo se seleccionaría el valor óptimo de número de componentes principales en función de este criterio.

**Tabla I.1.** Autovalores y varianza explicada para los 5 primeros componentes de un ejemplo en el que se hace el ACP a partir de las huellas dactilares de fitoesteroles de diferentes muestras de aceites vegetales.

Componente Principal	Autovalor	Varianza explicada, %	Varianza acumulada, %
1	$1.03 \times 10^6$	79.98	79.98
2	$1.27 \times 10^5$	9.89	89.87
3	$5.33 \times 10^4$	4.15	94.03
4	$3.26 \times 10^4$	2.54	96.57
5	$1.55 \times 10^4$	1.21	97.78

Según los datos que se muestran en la Tabla I.1, con 4 componentes principales sería suficiente explicar el ACP.

Esta decisión además se apoya visualizando el gráfico (Figura I.4) que aparece a continuación.

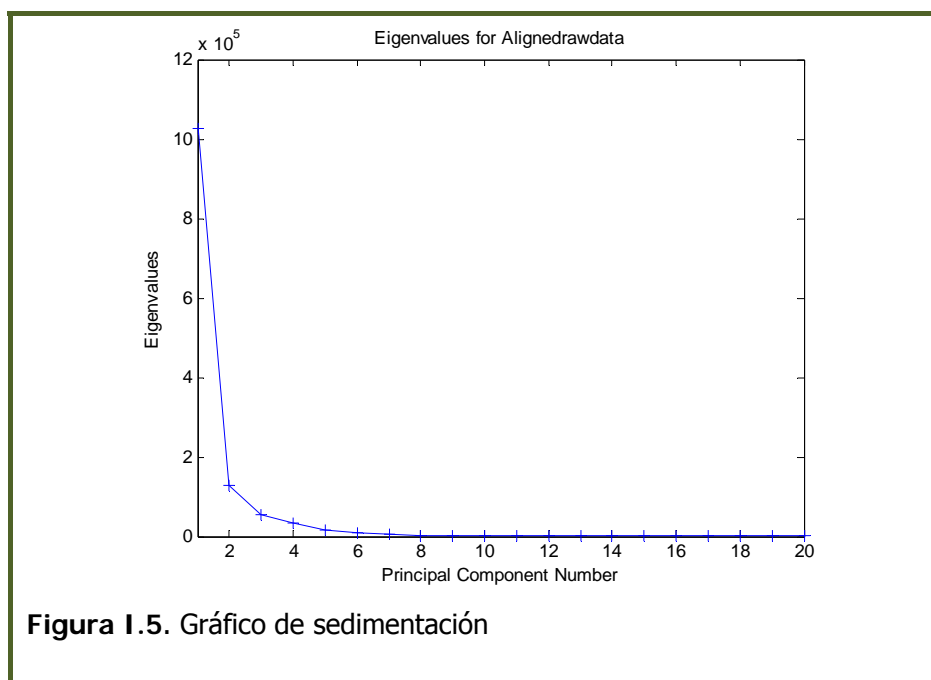


### 2. Criterio de autovalor 1

Este criterio se basa en el hecho de que el promedio del autovalor de los datos autoescalados sea justo uno. Los autovalores mayores que 1 son los que se consideran importantes.

### 3. Test de sedimentación

¿Cómo se puede aplicar este criterio? Mediante el gráfico de sedimentación de los autovalores. Este muestra una zona inicial de caída brusca, seguida de otra donde la caída es gradual. Se toma como valor óptimo de número de componentes principales el orden del autovalor o del componente donde se produce el cambio de caída brusca a gradual. En la Figura I.5 se representa este gráfico de sedimentación y cómo se puede observar, la zona de caída brusca de la varianza explicada finaliza en 3 como número de componentes principales, ya que las caídas posteriores son más suaves.



### 4. Validación cruzada

Es una herramienta útil en quimiometría con 2 funciones:

- Permite una evolución de la complejidad del modelo (por ejemplo, el número de componentes principales (CP) en un ACP o modelo de regresión por

componentes principales, RCP, o el número de variables latentes, VL, en un modelo de regresión por mínimos cuadrados parciales, PLS).

- Permite establecer una estimación del alcance de un modelo cuando se aplica a datos desconocidos.

Como hay varios métodos de modelado en quimiometría, también hay varios métodos de validación cruzada, y estos varían en función de cómo un subconjunto de diferentes muestras se seleccionan para estos experimentos de sub-validación.

El caso más simple es que de todos los objetos de la matriz  $X$ , se retira un conjunto de datos y se hace un modelo con los datos restantes. Los datos que se eliminan se predicen por medio del ACP y la suma de la raíz cuadrada de los residuales se calcula. Este es el método Leave-one-out.

En la Tabla I.2 se muestra un resumen de los 5 diferentes métodos de validación cruzada.

### **Interpretación gráfica de los componentes principales**


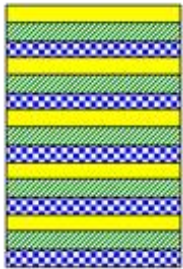


La interpretación de los resultados de un ACP se lleva a cabo mediante visualización de los scores/puntuaciones y loading/carga. Para ello hay software que dan la opción de obtener graficas de 2 o 3 dimensiones.

En la gráfica de los *scores*, se representa la proyección lineal de los objetos, que representa la parte principal de la variación total de los datos. En la Figura I.6. se muestra un ejemplo de una gráfica de *scores* de los aceites vegetales.

En la gráfica de los *loadings* se representa los *scores* de cada una de las componentes principales versus muestras. Con esta gráfica se ve qué componente influye más, es decir, tiene más peso sobre las muestras.



Tabla I. 2. Diferentes métodos de validación cruzada.

Métodos	Características	Esquema selección muestra
Leave one out	Fácil. Recomendado usar para $n^* \leq 20$ . No se recomienda para el diseño de experimentos	
Venetian Blinds	Fácil y relativamente rápido. Buena elección cuando se tiene objetos distribuidos al azar.	
Contiguous Block	Fácil y relativamente rápido. Buena elección cuando se tiene objetos distribuidos al azar. Útil para la previsibilidad entre lotes.	
Random subsets	Fácil. Puede ser lento si el número de iteraciones es grande. Selecciona subconjuntos desconocidos	
Custom	Flexible. Requiere tiempo.	Los subset / subconjuntos son definidos por el usuario. Se puede forzar objetos específicos en cada conjunto de datos (test set) o excluirlos del procedimiento.

\* n es el número de objetos

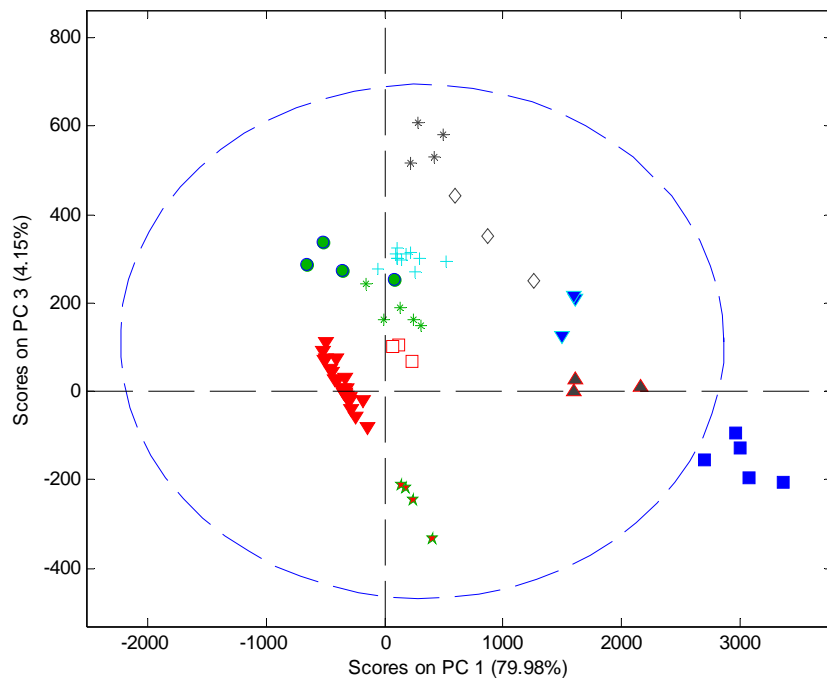
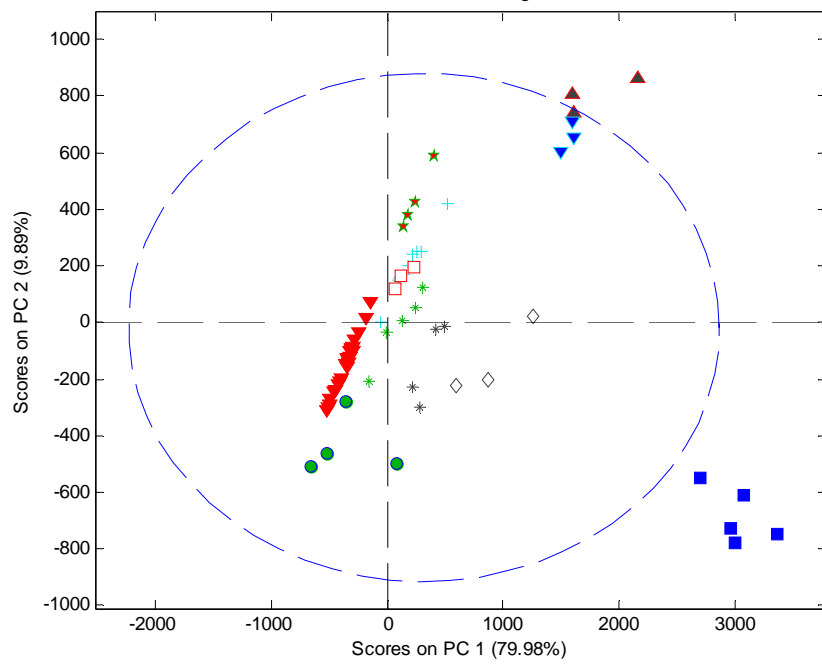


Figura I.6. Gráfica de los scores. A). PC 1 vs PC 2. B). PC 1 vs PC 3.

## **B). Análisis Factorial (AF)**

El análisis factorial [7] es un método descriptivo multivariante que pretende explicar, según un modelo lineal, la configuración de correlaciones de un conjunto de variables experimentales (observables) mediante un número reducido de variables hipotéticas denominadas factores.

La finalidad del análisis factorial es delimitar y obtener el menor número de factores capaces de explicar la máxima correlación entre las variables. El modelo se simplifica estandarizando las variables experimentales, transformando la distribución original  $X$  de cada variable en la normal reducida  $X^*$ ; así los factores comunes también serán variables reducidas. Esta técnica surge ante la necesidad de explicar en determinados sectores algunas variables conceptuales no observables directamente.

## **C). Relación entre ACP y AF**

Tanto el análisis factorial como el ACP se usan como reductores de la dimensionalidad del espacio, identificando un pequeño número de factores capaces de explicar la máxima correlación posible entre las variables experimentales. Los factores capaces de explicar las correlaciones yacen latentes y no son directamente observables, las situaciones que pretende resolver el análisis factorial son más ocultas que las que se presentan en el ACP donde sólo se pretende obtener un nuevo conjunto de variables no correlacionadas para excluir el problema de la colinealidad en un análisis de regresión por ejemplo. Estas técnicas están relacionadas pero son conceptualmente distintas. El análisis factorial engloba de forma genérica todas aquellas técnicas de reducción de dimensionalidad del espacio, para describir el sistema de forma más parsimoniosa a partir de un número menor de variables.

## **D). Análisis de Conglomerados (cluster)**

Es una técnica descriptiva multivariable que pretende agrupar un conjunto de muestras por afinidades definidas en función de las variables estudiadas, examinando las relaciones o similitudes entre ellas, de forma que los elementos de una misma agrupación se asemejen entre sí y discrepen de los restantes. Se utiliza como método exploratorio para agrupar las variables en subgrupos, catalogados en función de su similitud, cuando no se tiene ninguna idea preconcebida sobre la semejanza de las variables.

Los métodos de análisis de conglomerados se dividen en: métodos jerárquicos

(acumulativos y disgregativos) y métodos no jerárquicos.

En los *métodos jerárquicos* no se especifica el número de conglomerados que se van a formar, y se verifican fusiones sucesivas (acumulativos) o divisiones (disgregativos) para definirlos, presentando los resultados en forma de diagrama de árbol (dendograma). El número de conglomerados que se eligen se define a posteriori en función del resultado obtenido a lo largo de todo el desarrollo. La entrada y salida de un elemento son irreversibles.

En los *métodos no jerárquicos o de partición* se debe definir a priori el número K de conglomerados en que se quiere subdividir el conjunto de objetos, sin que se establezca estructura jerárquica alguna. Con estos métodos se pretende clasificar un conjunto inicial de casos en K subgrupos que mantengan una homogeneidad de acuerdo con las variables de partida, de forma que los elementos dentro de cada subgrupo se parezcan entre sí en base a las características del conjunto de variables observadas y, en base a esas mismas características, los K subgrupos estén alejados.

Se tienen que establecer criterios de asociación (Tabla I.3) de los elementos con medidas de distancia y medidas de comparación.

A continuación se explica brevemente en qué consiste cada una de las distintas medidas de asociación.

La distancia euclídea es la suma de los cuadrados de las diferencias entre los valores de los elementos (o su raíz cuadrada).

La distancia Mahalanobis por un lado es invariante ante los cambios de escala y no depende, por tanto de las unidades de medida. Esta es la diferencia con respecto a la distancia euclídea.

La distancia Bloque o Manhattan es la suma de las diferencias absolutas entre valores de los elementos.

La distancia Minkowski es la raíz n-ésima de la suma de las potencias n-ésimas de las diferencias absolutas entre los valores de los elementos.

La distancia Chebychev es la mayor de las diferencias absolutas entre las coordenadas de los elementos.

Correlación de Pearson es el coeficiente de correlación lineal entre dos variables.

El Coseno es el coseno del ángulo entre dos vectores de valores.

En nuestro caso se ha estudiado tanto la distancia euclídea como la distancia Mahalanobis, obteniendo mejores resultados con ésta última.

**Tabla I.3.** Diferentes medidas de asociación

Disimilaridades		Similaridades	
Distancia euclídea al cuadrado	$\sum_{j=1}^j (x_{ij} - x_{kj})^2$	Correlación de Pearson	$\frac{\sum_{j=1}^j (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k)}{\sqrt{\sum_{j=1}^j (x_{ij} - \bar{x}_i)^2 \sum_{j=1}^j (x_{kj} - \bar{x}_k)^2}}$ Varia de +1 a -1
Distancia euclídea	$\sqrt{\sum_{j=1}^j (x_{ij} - x_{kj})^2}$	Coseno	$\frac{\sum_{j=1}^j x_{ij}x_{kj}}{\sqrt{\sum_{j=1}^j x_{ij}^2 \sum_{j=1}^j x_{kj}^2}}$ Varia de +1 a -1
Distancia Mahalanobis	$\sqrt{(x_i - x_j)^t C^{-1} (x_i - x_j)}$		
Distancia Bloque o Manhattan	$\sum_{j=1}^j  x_{ij} - x_{kj} $		
Distancia Minkowski	$\sqrt[n]{\sum_{j=1}^j (x_{ij} - x_{kj})^n}$ de orden $n$ e igual a la euclídea si $n=2$ . Si $n=1$ igual a la distancia Manhattan		
Distancia Chebychev	$\max  x_{ij} - x_{kj} $		
Distancia Canberra	$\sqrt{\sum_{j=1}^j \frac{ x_{ij} - x_{kj} }{x_{ij} + x_{kj}}}$		

Una vez que se ha seleccionado la medida de proximidad, el siguiente paso es elegir el algoritmo adecuado para la formación de los conglomerados (Tabla I.4)

**Tabla I.4:** Criterios de enlace

Criterios de enlace
Enlace simple (Vecino más próximo)
Enlace completo (Vecino más lejano)
Promedio (vinculación inter o intra-grupos)
Agrupación de centroides
Agrupación de medianas
Método de Ward

A continuación se explica brevemente en qué consiste cada criterio:

Distancia mínima: Los grupos se unen considerando la menor de las distancias existentes entre los miembros más cercanos de distintos grupos.

Distancia máxima: Los grupos se unen considerando la menor de las distancias existentes entre los miembros más lejanos de distintos grupos.

Promedio: Mide la distancia entre dos conglomerados por la media aritmética de todas las posibles distancias entre los pares de individuos, de forma que cada componente del par pertenezca a un conglomerado distinto. Conduce a soluciones intermedias entre las de los enlaces simple y completo.

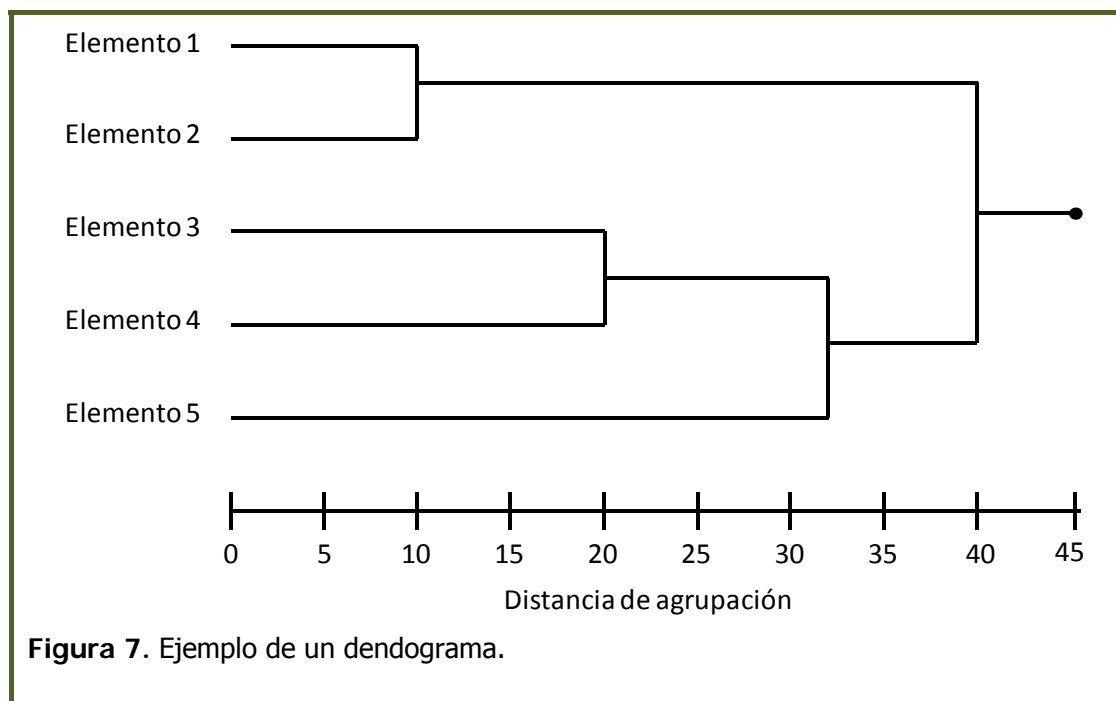
Agrupación de centroides: Es una variante del enlace promedio, en el que se mide la distancia entre los centroides de los conglomerados. Se reemplaza cada agrupación por su centroide y se forman aquellos grupos que tengan la menor distancia posible entre los centroides.

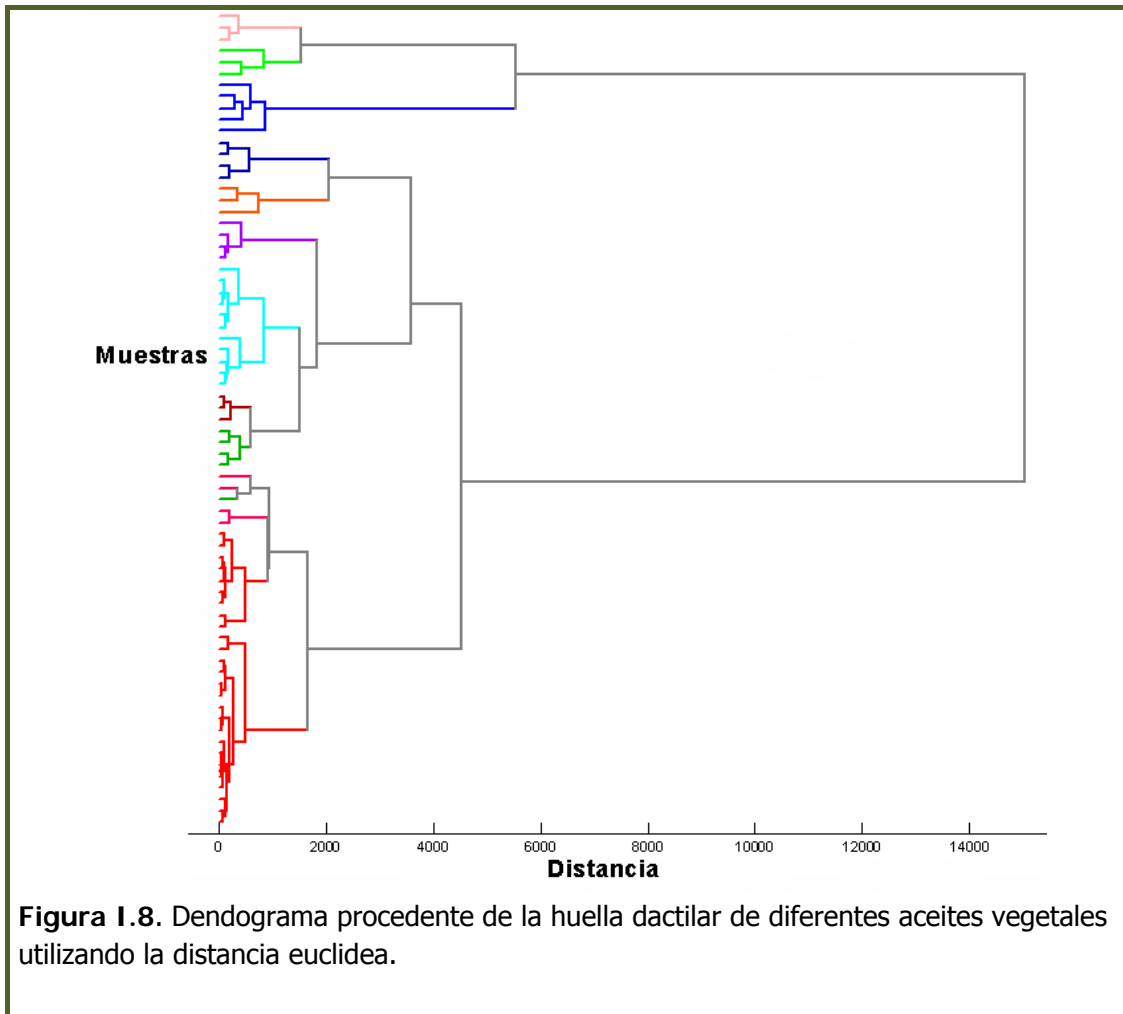
Agrupación de medianas: Se considera como nuevo centro de un conglomerado la media entre el centro anterior y el nuevo elemento o nuevo centro incorporado. Funciona de forma similar a la de los centroides.

Método de Ward: Se trata de ir agrupando de forma jerárquica elementos de modo que se minimice una determinada función objetivo. Este método no calcula distancias entre conglomerados sino que intenta minimizar la varianza dentro de ellos. Con este fin se computa como medida de disimilaridad la suma total de cuadrados de las distancias ( $SC_{DENTRO}$ ) dentro de los conglomerados desde cada objeto hasta su centroide, y en cada etapa se forma aquel conglomerado que tenga la menor ( $SC_{DENTRO}$ ).

El método de Ward es el método seleccionado para llevar a cabo nuestros análisis quimiométricos aplicados al estudio de la huella dactilar de los trimetilsililderivados de los esteroides mediante off-line HPLC-GC-FID.

Finalmente, tras elegir la distancia adecuada y el criterio de enlace se obtiene un *dendograma* (Figuras I.7 y I.8) que es la representación gráfica o diagrama de los datos en forma de árbol que organiza los datos en subcategorías que se van dividiendo en otros hasta llegar al nivel de detalle deseado (asemejándose a las ramas de un árbol que se van dividiendo en otras sucesivamente). Este tipo de representación permite apreciar claramente las relaciones de agrupación entre los datos e incluso entre grupos de ellos aunque no las relaciones de similitud o cercanía entre categorías.





**Figura I.8.** Dendrograma procedente de la huella dactilar de diferentes aceites vegetales utilizando la distancia euclídea.

### E). Análisis de Coordenadas Principales (PCO)

El PCO [8,9] se considera una versión generalizada de ACP. Es una técnica poderosa que permite una gama más amplia de enfoques para la visualización de datos de la matriz de disimilitud, y puede ser empleado como método de exploración. Además permite una flexibilidad adicional en el caso de las variables binarias. Los componentes obtenidos a partir del PCO se pueden utilizar como entrada a los clasificadores al igual que para los scores del ACP. A veces también se llama escalamiento multidimensional (MDS). Una ventaja que tiene el PCO sobre el ACP es que puede hacer frente a cualquier medida de disimilaridad y así es más flexible pero tiene la desventaja de que está diseñado sólo para examinar similitudes entre muestras y no es fácil de demostrar la influencia o la importancia de las variables.



## I.3. MÉTODOS SUPERVISADOS

El objetivo de los métodos de reconocimiento de pautas supervisados es utilizar una serie de objetos para encontrar una regla para asignar un nuevo objeto de grupo desconocido al grupo correcto.

### A). Clasificación

Para construir un modelo de clasificación supervisada es necesario disponer de una muestra de objetos cuya clase sea conocida y para los que también se conozcan los valores de las variables predictoras.

La pertenencia de los objetos a las clases puede ser supuesta, esto es, puede tratarse de una hipótesis a comprobar. La asignación de los objetos a las clases debe ser exhaustivo (todos los objetos pertenecen a alguna clase) y mutuamente exclusiva (ningún objeto pertenece a más de una clase). Estos objetos forman lo que se denomina el "conjunto de entrenamiento" con el cual se construye el modelo de clasificación. Una vez construido, el modelo se utiliza para predecir la clase a que pertenecen nuevos objetos a partir de la medida de la variables predictoras. Es usual utilizar muestras adicionales pertenecientes a clases conocidas, que no forman parte del conjunto de entrenamiento original, para comprobar el modelo y controlar su precisión y exactitud. De manera general en una clasificación supervisada se siguen los siguientes pasos:

1. Modelación del conjunto de entrenamiento.
2. Validación del modelo.
3. Mejoramiento de los datos (si es necesario).
4. Aplicación del modelo.

Existe una variedad de técnicas de clasificación entre las que se destaca: El análisis discriminante lineal y cuadrático (ADL y ADC), el método de los k vecinos más próximos (K-NN), el método de la máquina de aprendizaje lineal, los métodos de modelación de clases como el SIMCA y UNEQ entre otros.

### A.1. Análisis Discriminante

El ADL [10] fue la primera técnica de clasificación presentada en literatura e introducida por Fisher en 1936. Es una de las técnicas más usadas por su simplicidad y robustez.

En análisis discriminante se utiliza un algoritmo que busca funciones o vectores discriminantes, es decir, combinaciones lineales (ADL) o cuadráticas (ADC) de las variables originales que maximizan la varianza entre clases, a la vez que minimizan las varianzas inter-clases. Otra función que se emplea es la llamada función de clasificación de Bayesian [11].

Para construir el modelo, es necesario asignar los objetos del conjunto de entrenamiento a una clase dada. Se añade una variable categórica a la matriz de datos conteniendo tantas clases como sean necesarias.

El punto de partida del ADL es encontrar una función discriminante lineal,  $Y$ , que sea una combinación lineal de las variables originales,  $X_1$ ,  $X_2$ , etc.

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$$

Las  $n$  medidas originales para cada objeto se combinan en un único valor de  $Y$ , de manera que los datos se han reducido a  $n$  dimensiones a una dimensión. Los coeficientes de los términos se eligen de manera que  $Y$  refleje la diferencia entre los grupos tanto como sea posible: los objetos en el mismo grupo tendrán valores similares de  $Y$  y los objetos en grupos diferentes tendrán valores muy diferentes de  $Y$ . En consecuencia, la función discriminante lineal, proporciona un medio de discriminación entre los dos grupos.

Tras determinar la función discriminante lineal,  $Y$ , se puede establecer fronteras explícitas que separan las regiones del espacio que pertenecen a cada categórica. Cada clase se modela en torno a su centroide mediante una distribución normal de probabilidad. Esa probabilidad decrece por igual en todas las direcciones del espacio  $m$ -dimensional de las variables originales siendo alta para puntos situados cerca del centroide de la clase considerada, y baja para los puntos alejados.

El ADL tiene dos aplicaciones principales [12]: a). Evaluación del poder discriminante de las variables con respecto a la separación de los grupos y b) la asignación de cada objeto a uno de los grupos.

La clasificación de objetos desconocidos se lleva a cabo mediante la inserción de los datos en las funciones discriminantes con el fin de transformar las coordenadas en la misma forma que los datos originales. Así, el objeto se asigna a esa clase para la cual su centroide tiene la menor distancia euclídea.

Una desventaja de este método es el hecho de que la relación de un objeto a varias clases no se detecta y que los espurios/outliers no pertenecen a ninguna

clase. Debido a una serie de limitaciones que presenta se ha ideado el PLS-DA que será descrito en el apartado A.7.

El ADC supone que los dos grupos tienen distribuciones normales multivariantes pero con varianzas diferentes.

La diferencia entre un análisis de cluster y el AD está en que este último asigna objetos a grupos que fueron definidos por adelantado mientras que el cluster es bastante diferente porque establece los grupos.

## **A.2. Modelado Independiente Suave de Analogía de Clase, SIMCA**

SIMCA [13,14] como método de reconocimiento de pautas fue introducido por Svante Wold [15] en el año 1976 y es uno de los métodos de clasificación más utilizados en el control de calidad en las industrias alimenticias. Este método utiliza un análisis de componentes principales para desarrollar un modelo de cada grupo o clase dentro del conjunto de entrenamiento. Esas clases se modelan por una o series de estructuras lineales (una recta, un plano etc.) dependiendo del número de componentes requeridos para reproducir los datos de la clase. Posteriormente integra cada una de las clases y calcula sus límites con una probabilidad dada del 95%. Es posible definir superficies enlazadas alrededor de estas estructuras lineales basadas en los residuos de los datos después de fijar los componentes, nuevos objetos pueden ser cuantitativamente probados por sus miembros en las clases definidas.

Las principales etapas del SIMCA son:

- ACP. Cada grupo es modelado independientemente usando PCA.
- Distancia de clase. Se puede calcular como la distancia geométrica de los modelos de componentes principales.
- Poder del modelado. Varía entre 1 y 0.
- Poder discriminatorio entre dos clases.

## **A.3. Método de los K vecinos más próximos, K-NN.**

Este método no paramétrico de clasificación es conceptualmente simple para decidir cómo clasificar un objeto desconocido, nuevo, cuando hay dos o más grupos de objetos de clase conocida. No hace ninguna suposición sobre la distribución en las clases y se puede utilizar cuando los grupos no puedan ser separados en el plano (Figura I.9). En su forma más simple un objeto desconocido

es asignado a la clase de su vecino más próximo. Se toman los  $K$  vecinos más próximos, siendo un valor arbitrario, normalmente bajo e impar, tal como 3 o 5.

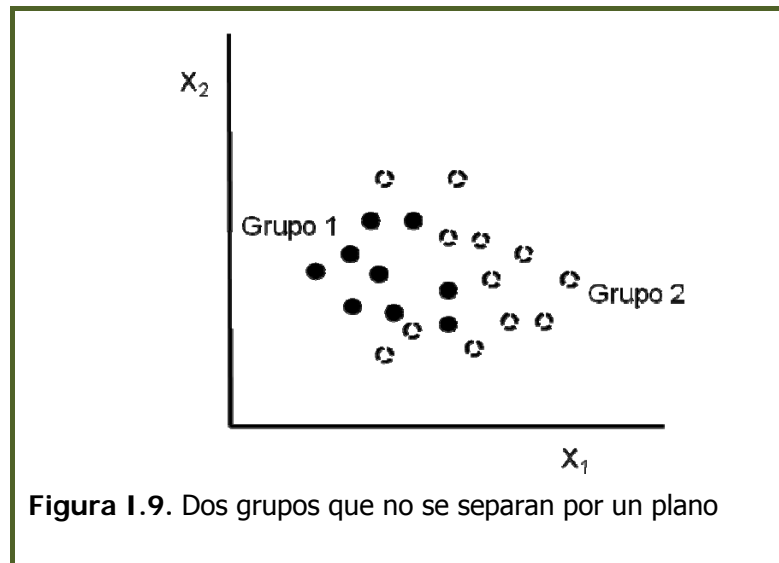


Figura I.9. Dos grupos que no se separan por un plano

Para aplicar el método se deben de calcular las distancias sobre la matriz objetos-variables sin escalar o autoescalada, y también antes o después de realizar una rotación propia. Si se realiza la rotación propia, se tiene la ventaja de trabajar en un espacio no correlacionado, en donde las distancias euclídeas tienen un significado correcto.

Las distancias se ordenan de menor a mayor, y se asigna el objeto nuevo a la clase a la que pertenecen los  $K$  objetos con las distancias menores. En el caso de que no todos los  $K$  objetos pertenezcan a la misma clase, decide la mayoría.

El mayor inconveniente del método  $K$ -NN es que no calcula probabilidades de pertenencia del objeto nuevo a cada categoría. Pero por otra parte, presenta varias ventajas: Permite clasificar nuevos objetos, independientemente del número de clases y la calidad de las predicciones depende poco de la forma que tengan las clases. Por ello, el método puede clasificar bien aunque las categorías no sean linealmente, ni cuadráticamente separables.

#### **A.4. Desigualdad de la Clase de Modelado, UNEQ**

Este método fue propuesto por Derde y Massart en los años 80 y está basado en la distribución Gausiana. Cada grupo es modelado con una distribución normal Multivariante, por este motivo se puede considerar como análogo al modelado de ADC.

El modelo se construye mediante el centroide de la clase, teniendo forma de hiper-elipsoide, siendo el centroide la media de la población y definiendo su tamaño mediante el valor de la distancia crítica. Teniendo en cuenta que el modelo de cada clase se construye mediante el cálculo de las distancias de Mahalanobis de forma individual, siendo la matriz de dispersión distinta para cada clase. UNEQ es especialmente adecuado para clases que presenten diferente dispersión y además como se basa en la estimación de los parámetros de una distribución gaussiana multivariante, presenta limitaciones, en términos del número de muestras con el número de variables.

### **A.5. Máquina de Aprendizaje Lineal, LLM**

Este método [16] encuentra fronteras lineales entre las clases que pueden calcularse a través del sistema de prueba y error en el cual los errores de la prueba inicial permiten que la próxima prueba pueda ser desarrollada más inteligentemente. Para un espacio  $m$ - dimensional la superficie de decisión tiene que ser de una dimensión menor que  $m$ . Este método puede ser extendido a la separación de clases múltiples a través de una secuencia de superficies que producen separaciones binarias.

### **A.6. Máquinas de Soporte Vectorial, SVM**

Las máquinas de Vectores de Soporte (Support Vector Machine, por sus siglas en inglés SVM) son un nuevo sistema de aprendizaje el cual ha tenido un desarrollo muy significativo en los últimos años tanto en la generación de nuevos algoritmos como en las estrategias para su implantación. Fue descubierto recientemente en el año 2004 como clasificador por Zomer. SVM [17,18] es un sistema de aprendizaje robusto que se basa en el uso de un espacio de hipótesis de funciones lineales en un espacio de mayor dimensión inducido por un Kernel, en el cual las hipótesis son entrenadas por un algoritmo tomado de la teoría de optimización el cual utiliza elementos de la teoría de generalización. Su uso está dirigido para solucionar problemas que presentan dificultades como por ejemplo los límites entre clases son complejos y el conjunto de datos se caracteriza por un gran número de variables. SVM es un sistema para entrenar máquinas de aprendizaje lineal y no lineal eficientemente tanto para resolver problemas de clasificación binaria como para regresión con el inconveniente de que está creado para solucionar problemas de dos clases.

Se han encontrado muchas aplicaciones como clasificación de imágenes, reconocimiento de caracteres, detección de proteínas, clasificación de patrones, identificación de funciones, etc.

### **A.7. PLS-DA**

Debido a que el ADL está sujeto a la restricción de que el número de variables no debe de exceder al número de objetos, se propone a este problema una solución que consiste en aplicar métodos como el de PLS (descrito en el apartado B.4).

El PLS-DA se lleva a cabo en dos pasos: 1. Una regresión clásica PLS dónde la variable respuesta es binaria y expresa una pertenencia a una clase. 2. AD para clasificar los objetos a partir del resultado de la regresión obtenida con el modelo PLS. Por tanto, PLS-DA no permite atribuir una muestra a otros grupos que están definidos. Como consecuencia, todas las variables medidas juegan el mismo papel con respecto a la clase que se asignan. Actualmente, las variables latentes PLS se construyen encontrando un compromiso propio entre dos propuestas: describir el set de variables exploratorias y la predicción de las respuestas.

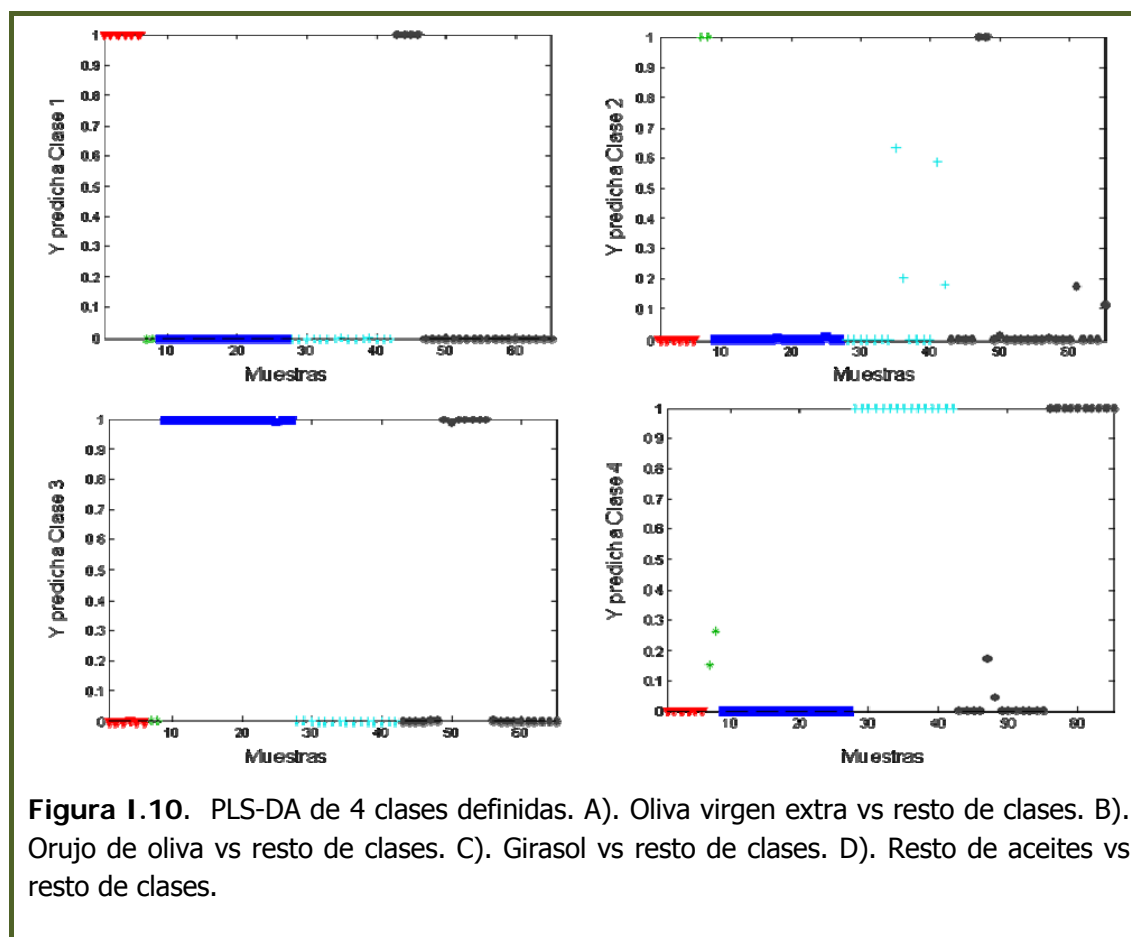
En el PLS-DA igual que en el PLS hay dos bloques: El bloque **X** que es la que tiene las variables independientes y el bloque **Y** que es la variable respuesta. El bloque **Y** se construye formando una matriz con 0 y 1. Constará de tantas columnas como clases haya. Se pondrá el valor 1 para la clase a la que pertenezca correspondiente al orden en que se define y 0 al resto de muestras que no pertenezcan a esa clase en concreto.

Las puntuaciones/loadings del bloque **X** se calculan a partir de las coordenadas/scores del bloque **Y**, mientras que los loadings del bloque **Y** se determinan a partir de los scores del bloque **X**. Se desarrolla un modelo para cada clase.

Para su entendimiento, a continuación se muestra un ejemplo basado en la clasificación de diferentes aceites vegetales pertenecientes a distintas clases partiendo de una matriz de datos obtenidas a partir de la señal cromatografica. Se consideran un total de 67 muestras pertenecientes a 4 clases (aceite de oliva virgen extra, orujo de oliva, girasol y resto de aceites vegetales (maíz, soja, cacahuate, uva, sésamo, lino, canola y colza)). Se hace el modelo partiendo de un set de entrenamiento de muestras correspondiente al 75% del total para luego hacer una validación interna con el 25% restante, cogiendo las muestras al azar. 5 componentes principales fueron necesarios explicando un 98.31 % de varianza. Las clases definidas fueron: Clase 1: Aceite de oliva virgen extra. Clase 2: Orujo de oliva, Clase 3: Girasol y Clase 4. Resto de aceites vegetales.

Los resultados se muestran en la Figura I.10 en dónde se puede observar que hay una buena predicción. Todas las muestras son clasificadas correctamente a la

clase a la que pertenecen. Los resultados estadísticos obtenidos fueron:  $R^2$  de calibración 0.961., error de calibración: 0.0689, error de validación cruzada: 0.0909 y error de predicción: 0.462



## B). CALIBRACIÓN

Existen diferentes métodos de calibración: univariante y multivariante. Esta última tiene muchas aplicaciones en espectroscopía y en cromatografía. La calibración univariante implica relacionar dos variables individuales entre sí, y se llama a menudo regresión lineal.

La principal ventaja de la calibración multivariante frente a la univariante es que no requiere medidas instrumentales selectivas. La muestra puede contener, junto con el analito, otros componentes que contribuyen a la señal de forma no constante. Otra ventaja es que la señal multivariante permite detectar muestras discrepantes en predicción (*outliers*). Esto no es posible en calibración univariante.

Hay distintos niveles de complejidad dependiendo del tipo de problema que se quiera resolver. La calibración más sencilla es la que permite predecir la concentración de un solo compuesto utilizando un equipo instrumental. Otra consiste en una situación más compleja basada en la determinación de varios componentes en una mezcla cuyo objetivo es determinar las concentraciones individuales de forma fiable de cada elemento a partir por ejemplo de datos espectroscopicos. Otra opción sería que el objetivo de la calibración no fuese determinar la concentración de un compuesto sino calcular diferentes parámetros estadísticos.

A continuación se van a explicar los diferentes métodos de calibración que son más empleados a fecha de hoy.

### **B.1. Regresión Lineal Múltiple (RLM)**

En calibración lineal simple es habitual usar el modelo de predicción:

$$\hat{y} = b_0 + b_1 C$$

Dónde  $\hat{y}$  es la señal instrumental ajustada o predicha por el modelo, y  $C$  es la concentración de una especie química. Al construir el modelo,  $C$  es la variable predictora e  $\hat{y}$  es la variable respuesta. Este modelo calcula, para el análisis de muestras con más de un componente, una relación lineal entre la señal y la concentración de cada componente mediante un ajuste de mínimos cuadrados. Hay dos variantes de RLM: Clásica e inversa

Para la regresión lineal múltiple clásica se asume el cumplimiento de la ley de Lambert-Beer para cada uno de los componentes de una mezcla en todo el intervalo de trabajo y la actividad de las absorbancias en las muestras. Cuando se trata de predicción múltiple y multivariante, se utilizan siempre modelos en los que la variable independiente,  $C$ , es siempre la predictora tanto en la etapa de calibración como en la de predicción, y la variable dependiente,  $\hat{y}$ , es siempre la respuesta.

La ecuación que figura a continuación es la que se emplea para el modelo de RLM. Se utiliza para predecir una única respuesta,  $\hat{y}$ , teniendo en cuenta que son varias las variables predictoras,  $C_1, C_2, \dots, C_j, \dots, C_m$ . y se emplea el error asociado a la  $i$ -ésima observación,  $\epsilon$ .



$$\hat{y} = b_0 + b_1 C_1 + \dots + b_j C_j + \dots + b_m C_m + \varepsilon = \sum b_m C_m + \varepsilon$$

Cada variable predictora está multiplicada por su pendiente o coeficiente de regresión parcial,  $b_j$ .

Este método de regresión lineal múltiple, también se llama mínimos cuadrado inversa, ILS.

La ventaja de este modelo es que cada respuesta se puede determinar por separado, sin que sea necesario de conocer el resto de respuestas, ni establecer sus correspondientes coeficientes de regresión.

El modelo lineal multivariante es un modelo que tendrá tantas ecuaciones como respuestas interesa para predecir. Esto se debe a que puede haber la presencia de distintas fuentes de varianza que influyen sobre ellas. Para predecir simultáneamente hasta  $k$  respuestas, el modelo es:

$$\hat{y}_1 = b_{01} + b_{11} C_1 + b_{21} C_2 + \dots + b_{j1} C_j + \dots + b_{m1} C_m + \varepsilon$$

..

$$\hat{y}_k = b_{0k} + b_{1k} C_1 + b_{2k} C_2 + \dots + b_{jk} C_j + \dots + b_{mk} C_m + \varepsilon$$

## **B.2. Regresión en Componentes Principales, RCP**

La RCP es un método introducido por Massy en 1965. Aprovecha las propiedades de la descomposición en componentes principales, ACP, realizando una regresión múltiple inversa, ILS, de la propiedad a determinar sobre los *scores* obtenidos en el ACP en lugar de realizarla sobre los datos originales. No existe pérdida de información, ya que los *scores* contienen la misma información que los datos originales pero habiendo eliminado el ruido.

El primer paso del RCP consiste en realizar una descomposición de la matriz  $X$  en sus componentes principales.

$$X = TP^t + E = \sum_{a=1}^A t_a p_a^t + E$$

Tras elegir el número de  $A$  componentes principales para la matriz  $X$ , ésta se puede representar

$$T = XP$$

Hasta aquí se ha hecho un ACP, obteniendo a partir de la matriz de datos **X** la matriz de *scores* **T** y la de *loadings* **P**. Luego se calcula la matriz de regresores que se halla por mínimos cuadrados.

#### **B.4. Regresión Lineal por Mínimos Cuadrados Parciales, PLS.**

La regresión por mínimos cuadrados parciales (**PLS**) como técnica de regresión multivariante es un método para la construcción de modelos de predicción y sus principios generales son bastantes sencillos. Es un modelo muy utilizado para clasificar los datos, encontrar similitudes entre muestras o detección de muestras extrañas (outliers). Fue desarrollado por primera vez por Hermann Wold en 1975 en las áreas de econometría y ciencias sociales. Posteriormente, su hijo Svante Wold lo popularizó en el área de química y se convirtió en una de las técnicas de regresión más utilizadas en quimiometría. El PLS puede resolver muchos problemas prácticos que no son fáciles de tratar por otros medios. Fue modificado ligeramente para adaptarse a los datos de la ciencia y tecnología.

En PLS, para nuestro caso en concreto, el perfil cromatográfico de una muestra (intensidades registradas en diferentes tiempos) constituyen un conjunto de variables independientes (o, mejor, las predictoras) (variables de entrada) que, en general, se llama **bloque-X**. La variable que ha de ser predicha es la variable dependiente (o variable respuesta) (variables de salida) y se llama **bloque-Y**. El **bloque Y**, debe incluir tantas variables respuesta como fuentes de varianza significativas e independientes que influyan sobre las variables del **bloque X**, en nuestro caso se trata de las concentraciones de los analitos [3].

En realidad, el PLS se puede utilizar para predecir no sólo una variable, sino varias, de ahí el término "**bloque**". Por este motivo surge la idea por Hermann Wold de cómo utilizar un conjunto de variables para predecir el comportamiento de otras.

El funcionamiento del PLS es supervisado y consta de las siguientes etapas: calibración, cross-validación/validación-cruzada, validación y predicción.

En la etapa de **calibración** se establecen las variables que deben incluirse en los **bloques X** e **Y**. El bloque Y contiene tantas variables como fuentes de varianza relevantes influyen sobre el **bloque X**.

Las variables del **bloque X** contienen varianza correlacionada con todas y cada una de las variables del **bloque Y**.

Además las variables de ambos bloques están poco correlacionadas entre sí cubriendo los diferentes intervalos de calibración, que corresponden a las situaciones extremas que vayan a encontrarse en las muestras problema.

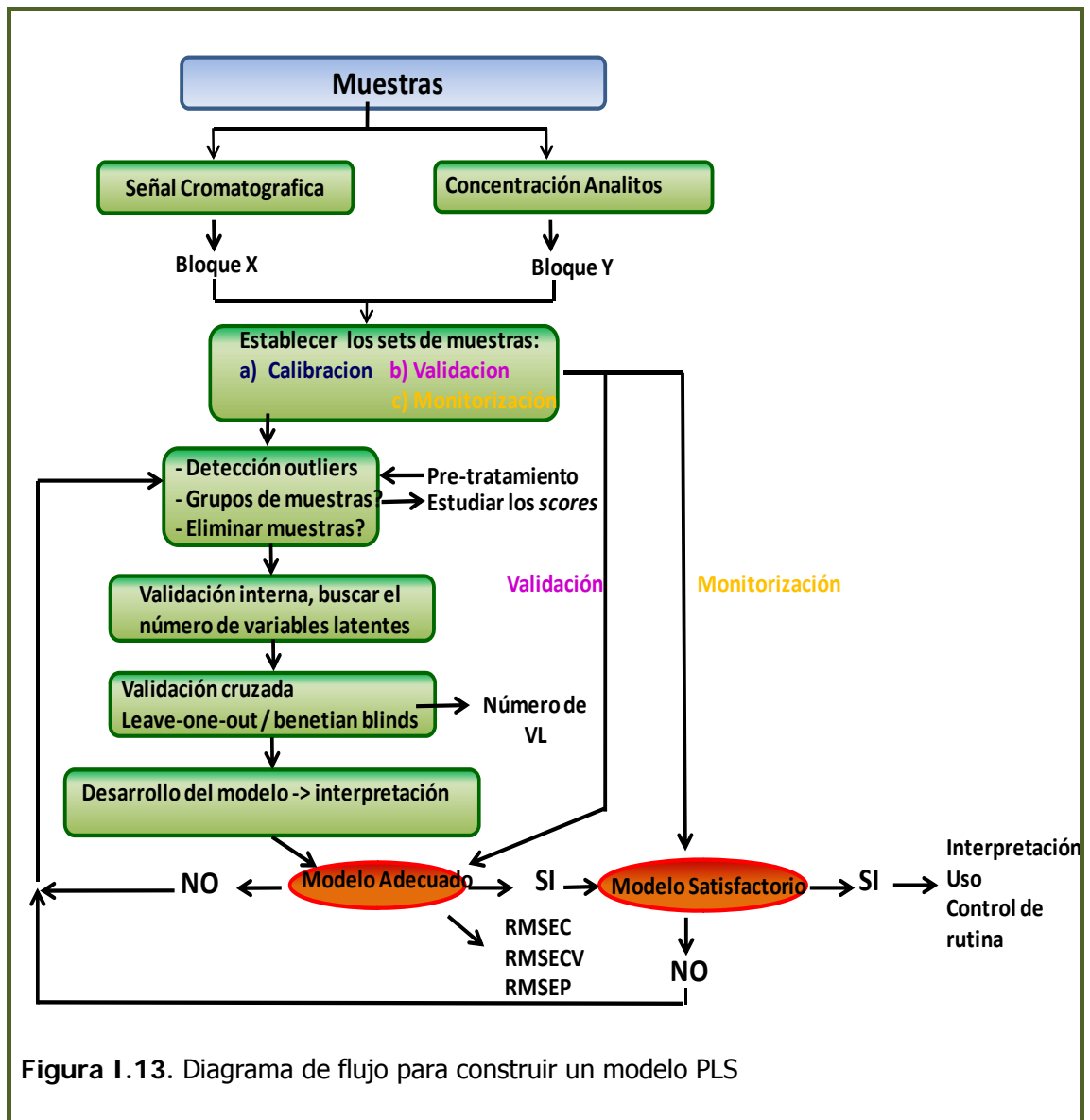
En la etapa de **validación**, se establece la capacidad del modelo para realizar predicciones exactas y precisas. La validación puede ser interna o externa. Si es interna lo que se hace es coger un subset del conjunto de variables del bloque X que se ha empleado para calibrar. Si la validación es externa, se toma un set de muestras no incluidas en la calibración y que son perfectamente conocidas.

En la etapa de **predicción** se establecen las variables problema, sin limitación en el número de objetos/muestras.

En definitiva en el PLS se utilizan como variables predictoras, vectores que explican menos varianza que los componentes principales, pero a cambio están más correlacionados con las respuestas. Los componentes principales se obtienen en orden de máxima varianza explicada decreciente, por lo que no es posible aumentar la correlación de los vectores con la respuesta sin perder varianza. Este razonamiento también puede hacerse a la inversa: si se combinan linealmente las variables del **bloque X** con el único criterio de obtener vectores correlacionados con las respuestas, es posible que los vectores resultantes expliquen poca varianza.

En PLS se construyen vectores que establecen un compromiso razonable entre ambas situaciones. Dichos vectores se suelen denominar "variables latentes", variable predictora para predecir una respuesta dada (método PLS1, para regresión múltiple), o para predecir un pequeño grupo de respuestas (método PLS2, para regresión multivariante).

Las principales etapas que se requieren para obtener un modelo PLS se presentan en el siguiente diagrama de flujo (Figura I.13).



### Algoritmos que se emplean en el PLS

Existen varios algoritmos que se emplean en el PLS entre los que destaca el algoritmo en mínimos cuadrados parciales iterativo no lineal (**NIPALS**), singular value decomposition (SVD), el SIMPLS, etc.

El **NIPALS** calcula los scores/coordenadas  $T$  y los loadings/cargas/pesos  $P$  (similar que en el PCR o RCP) y además un set adicional de vectores conocidos como pesos,  $W$  (con la misma dimensionalidad como los loadings,  $P$ ). La adición de los pesos en el PLS se requiere para mantener los scores/muestras/objetos ortogonales.

La principal diferencia del porqué no se usa el **NIPALS** es porque éste extrae componentes de uno en uno y se puede detener tras elegir el número de factores [19]

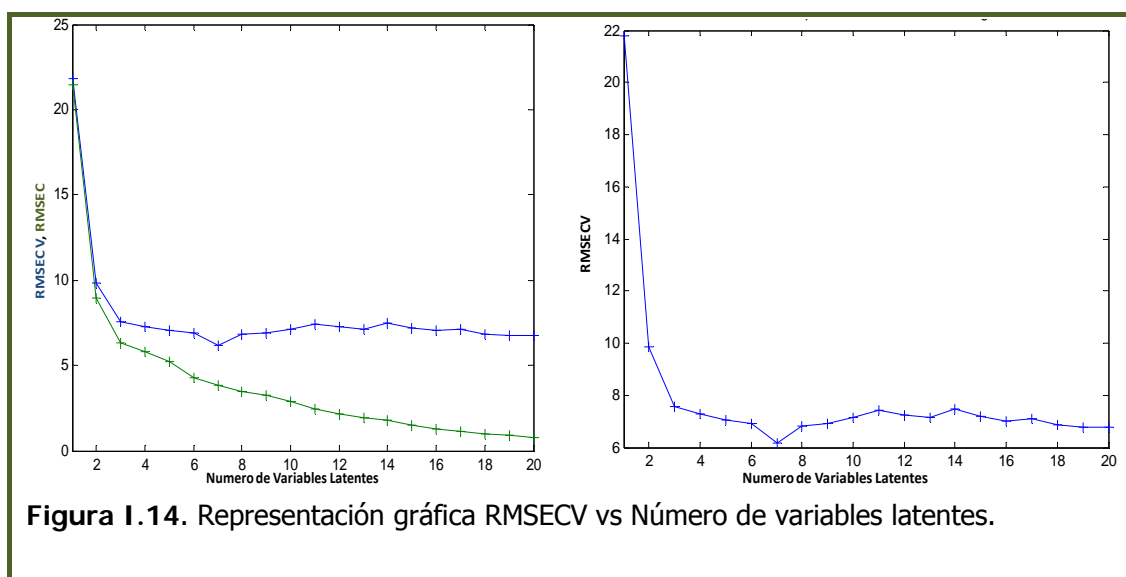
Hay evidencias de que el algoritmo **SIMPLS** es actualmente ligeramente superior al algoritmo **NIPALS** porque su objeto explícito es maximizar el criterio de la covarianza y el **NIPALS** no lo hace. Además es más rápido y las matrices residuales son diferentes [20,21].

El software PLS\_Toolbox emplea por defecto el algoritmo **SIMPLS** aunque también da la opción de usar el **NIPALS** pero debido a lo último que se ha comentado el algoritmo empleado ha sido el **SIMPLS**.

### Elección del número de variables latentes:

Para la selección del número óptimo de variables latentes se usa el método de validación cruzada. Se representa la gráfica (Figura I.14) "Número de variables latentes" vs "Errores de calibración y calibración cruzada de la media de la desviación estándar", RMSEC (root-mean-square error of calibration /error de calibración de la raíz cuadrada de la media) y RMSECV (error de calibración cruzada de la raíz cuadrada de la media) respectivamente.

Una buena regla general es la de no incluir variables latentes adicionales a menos que se mejore el RMSECV al menos un 2%. Por este motivo, en este ejemplo se eligen 3 variables latentes y no 4 variables latentes.



Otra regla es que se debe de utilizar un modelo con la complejidad más baja posible para la construcción de modelos predictivos, es decir, elegir el menor número de variables latentes posibles. Al realizar el PLS aparecerá una tabla como se muestra a continuación (Tabla I.5) donde aporta la información necesaria a tener en cuenta para elegir el número de variables latentes.

**Tabla I.5.** Elección del número de factores.

Factores	% Varianza			
	Bloque X		Bloque Y	
	LV	Acumulada	LV	Acumulada
1	67.95	67.95	54.72	54.72
2	13.48	81.43	34.99	89.71
3	4.65	86.08	6.42	96.13
4	5.24	91.33	1.04	97.18
5	3.79	95.12	0.41	97.59
6	1.05	96.17	0.64	98.22
7	1.55	97.72	0.20	98.43
8	0.44	98.16	0.43	98.85
9	0.52	98.68	0.13	98.99
10	0.21	98.89	0.25	99.23

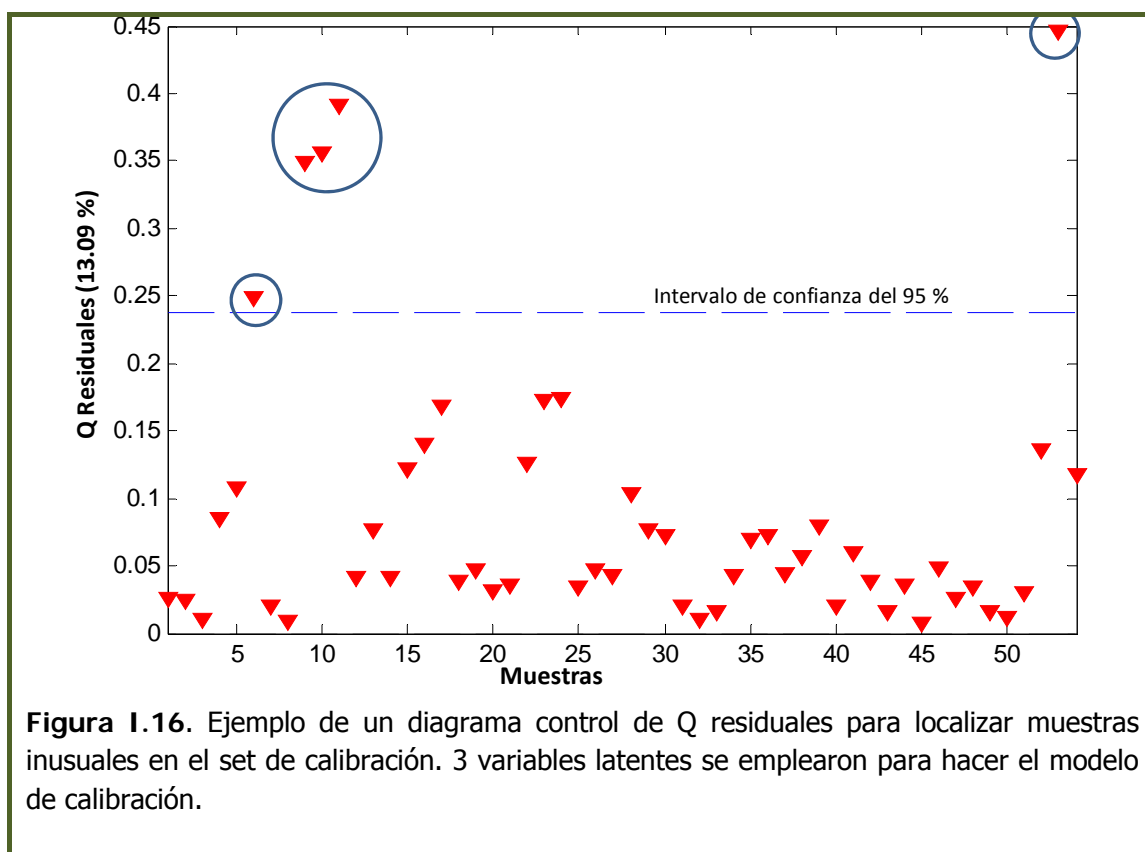
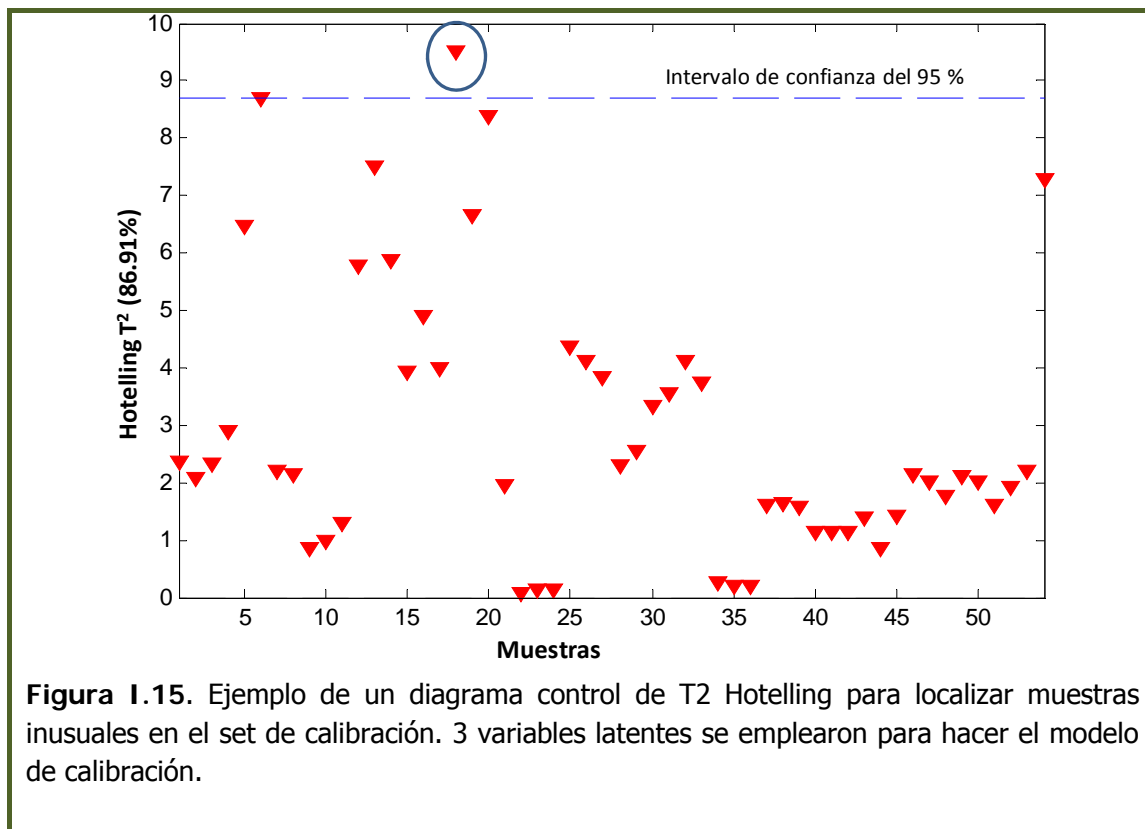
### Detección de muestras extrañas o espurios

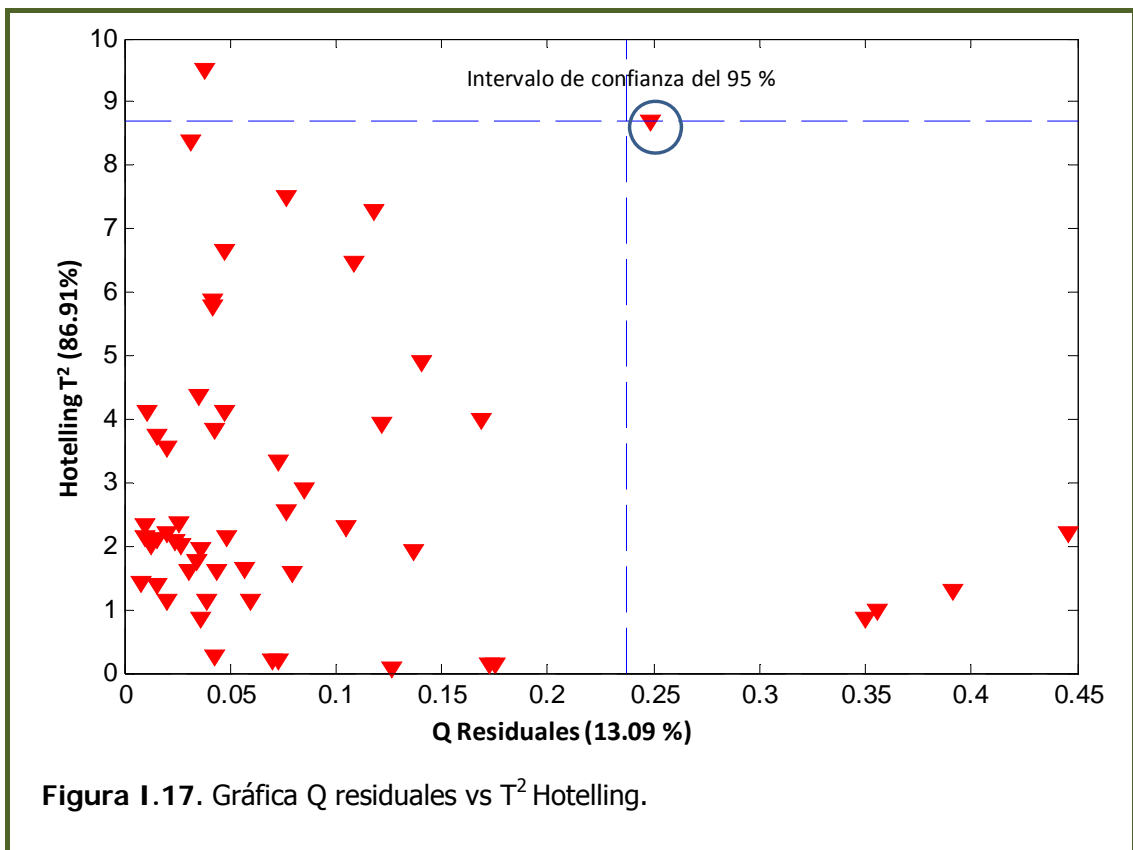
Hay varias formas de analizar las distintas muestras de estudio.

1. Gráfica de los "scores" obtenida tras elegir el número de variables latentes. Con esta gráfica se conoce el % de varianza total explicada de los datos. De cada muestra se puede obtener información acerca de los valores  $Q$  residuales y  $T^2$  Hotelling. Si hay muestras que presentan valores elevados tanto en  $Q$  residuales como  $T^2$  deben considerarse como muestras inusuales y habría que estudiarlas más a fondo.

2. Otra forma de ayudar a ver claramente las muestras inusuales sería representar  $T^2$  Hotellings versus muestras (Figura I.15),  $Q$  residuales versus muestras (Figura I.16) o bien  $T^2$  Hotellings versus  $Q$  residuales (Figura I.17). La  $T^2$  Hotellings se utiliza para verificar, en un sentido multivariante, si una muestra pertenece a una distribución cuyo promedio está dado por el vector promedio de las muestras de calibración. Es una generalización multivariante de la distribución *t de Student*.

El término "residuales" significa la diferencia entre el valor verdadero y el predicho.





Las muestras dudosas serían aquellas que sus valores tanto de Q residuales como de T<sup>2</sup> Hotellings están fuera del límite de confianza del 95%. Véase Figura I.17. La línea azul horizontal discontinua es la que se refiere al límite de confianza. La muestra que está marcada con un círculo sería una muestra sospechosa ya que está justo en el límite.

3. Gráfica Leverage or h-statistic versus Y studentized Residual (Figura I.18). Studentized Residual indica la falta de ajuste de los valores y de la muestra. Se usa para hacer frente a la posible variación en los errores de medición cuando las muestras son diferentes de la media. El término Leverage es la influencia de una muestra en el modelo. Si hay muestras con ambos valores elevados nos indica que esas muestras tienen mucha influencia sobre el modelo pero no se sabe si puede ser buena o mala.

El valor de leverage se calcula de la siguiente manera:

$$\text{Leverage} \succ \frac{2A}{I} \text{ o } \text{Leverage} \succ \frac{3A}{I}$$

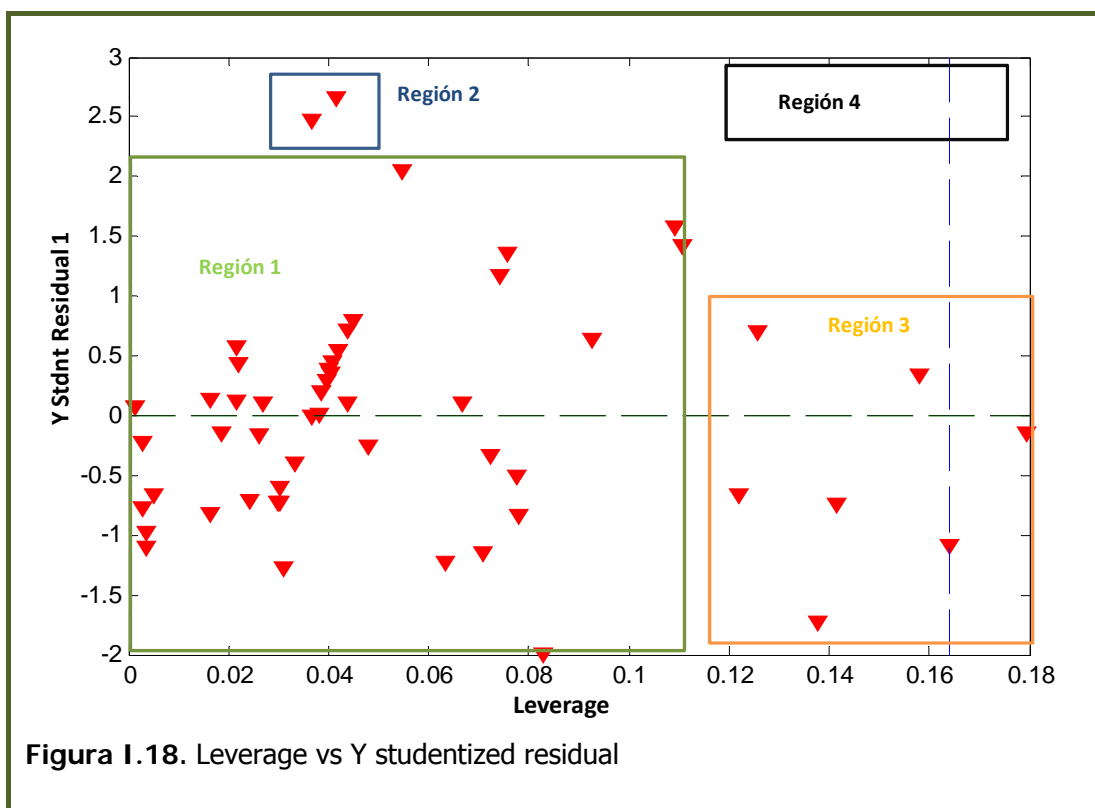
siendo,



A, número de variables latentes elegidas. I, número total de muestras usadas en el set de calibración.

Se considera que un valor de  $\pm 3$  en Studentized Residual indica que las muestras deberían de ser controladas.

Con la gráfica que se ha representado se pueden estudiar varias regiones: La región 1 corresponde a muestras que tienen bajos valores residuales y bajos valores de leverage. La región 2 corresponde a muestras caracterizadas por tener un alto valor residual pero bajo Leverage. La región 3 corresponde a muestras con alto leverage pero bajo residual. La región 4 correspondería a muestras que tienen alto nivel de valores residuales y de leverage. En este ejemplo en concreto no hay muestras situadas en esa región.



### Parámetros de calidad y capacidad predictiva del modelo

Diferentes parámetros de calidad son estudiados [22]:

1. Error de calibración de la raíz cuadrada de la media, RMSEC

$$\text{RMSEC} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n - K - 1}};$$

Siendo

$\hat{y}_i$ : los valores de la variable predicha cuando todas las muestras están incluidas en el modelo de formación.

n: número de muestras en la calibración.

K: número de variables latentes

## 2. Error de validación cruzada de la raíz cuadrada de la media, RMSECV

Se define igual que RMSEC pero con la diferencia de que  $\hat{y}_i$  corresponde a las predicciones de muestras que no están incluidas en el modelo de formación.

## 3. Error residual predicho de la suma de los cuadrados, PRESS

$$\text{RMSECV}_k = \sqrt{\frac{\text{PRESS}_k}{n}}$$

siendo,

K: número de variables latentes usado en el modelo

n: número de muestras de calibración

Los valores de RMSECV y PRESS se refieren cuando se hace validación cruzada dónde los datos de calibración se dividen en diferentes set de entrenamiento y test para evaluar el modelo construido cuando se le aplica nuevos datos.

## 4. Error de predicción de la raíz cuadrada de la media, RMSEP

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

$\hat{y}_i$ : los valores de la variable para muestras de predicción que no están incluidas en el modelo de de calibración o validación cruzada.

n: número de muestras de predicción.

## 5. Coeficiente de correlación lineal al cuadrado, R<sup>2</sup>

Evalúa el nivel de asociación lineal entre la variable respuesta y las regresoras, que oscila entre  $\pm 1$  (fuerte asociación lineal positiva o negativa). Si el valor es 0 no existe correlación lineal.

Para poder asegurar la capacidad predictiva de un modelo en la predicción de muestras distintas a las utilizadas en calibración / set de entrenamiento, es necesario un proceso de validación. Esta validación puede ser externa o interna. Para la validación externa de los modelos, se suele utilizar dos grupos de muestras, uno llamado propiamente el de calibración y el otro llamado conjunto de prueba o test set. Las muestras del test set deben ser independientes de las del conjunto de calibración pero representativas y de las futuras muestras a analizar. Se conoce toda la información acerca de este test set para comprobar cómo se comporta el modelo frente a muestras distintas a las utilizadas en la construcción del mismo.

La validación interna, también llamada validación cruzada es la que utiliza para comprobar el modelo muestras del propio conjunto de calibración. Normalmente se cogen el 25% de las muestras del conjunto de calibración.

### **Comparación con otras técnicas de regresión**

El PLS presenta algunas ventajas con respecto a otras como RLM y RCP.

1. La colinealidad entre las variables X (correlación de los predictores) no es un problema. Las variables latentes en el PLS, resume la información más relevante de todo el conjunto de datos, tomando combinaciones lineales de las variables X que son ortogonales y los pesos que son ortonormales asignando un peso extra a las variables que muestran una alta correlación con las variables respuesta para que sean más efectivas en la predicción[23]. En RLM hay problemas de colinealidad ya que el número máximo posible de componentes principales (PCs) o variables latentes debe ser igual al número menor de variables respuesta y al número de muestras de calibración. Además no puede ser nunca mayor que el número de muestras de calibración.
2. El PLS genera modelos más parsimoniosos que el RCP porque requiere un número menor de factores para obtener un rendimiento predictivo similar.
3. En PLS se utilizan como variables predictoras vectores o como se suelen denominar "variables latentes" que explican menos varianza que los componentes principales, pero a cambio están más correlacionados con las respuestas. Los componentes principales se obtienen en orden de máxima varianza explicada decreciente, por lo que no es posible aumentar la correlación de los vectores con la respuesta sin perder varianza. Por lo tanto la ventaja que presenta es que el número óptimo de variables

latentes en PLS suele ser menor que en PCR ganando sencillez sin perder exactitud en las predicciones.

4. Admite errores tanto en los bloques X como Y.

En el modelo RCP se buscan, al igual que en un ACP, las componentes principales que mejor describen la varianza de las variables predictoras. Con estas componentes principales se monta un modelo de correlación entre las variables predictoras y las predichas (en RCP solo se puede predecir una variable por modelo).

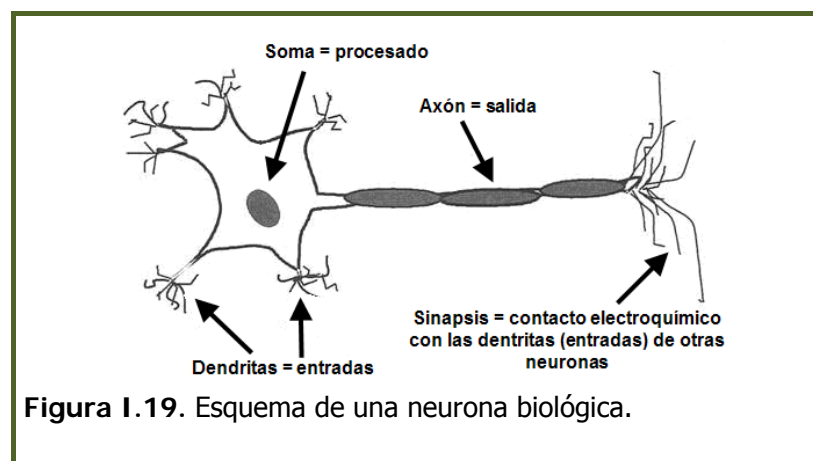
El RLM busca un factor simple que maximice la correlación con la variable a predecir.

El PLS suele funcionar mejor que RCP o RLM en la mayoría de las situaciones, motivo por el que se decidió utilizar este algoritmo para correlacionar la señal Cromatográfica de los fitoesteroles en aceites vegetales.

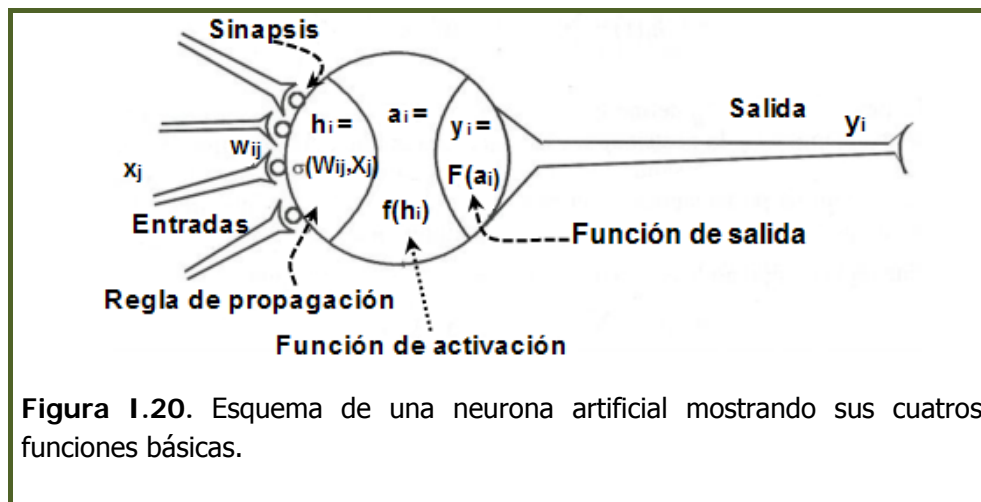
### C). MÉTODO DE LAS REDES NEURONALES ARTIFICIALES

En general, el método de las redes neuronales es un procedimiento sistemático de procesamiento de datos, basado en la imitación del funcionamiento de los sistemas nerviosos biológicos, y que es capaz de predecir con precisión variables categóricas y de escala, diferentes clases. Se basa en un conjunto de "nodos" o "neuronas artificiales" interconectados entre sí en forma de red. Se usan tanto como métodos de reconocimiento de pautas supervisados y no supervisados.

En la Figura I.19 se muestra un esquema de las neuronas biológicas con sus cuatro componentes básicos: dendritas, soma, axón y sinapsis.



La neurona artificial contiene 2 algoritmos, uno calcula la suma ponderada de los valores que le llega por las conexiones de entrada, y el otro es el que genera una única respuesta o salida que se comunica con otras neuronas (Figura I.20).



Diferentes tipos de redes neuronales se pueden emplear: Las redes neuronales de Teoría de Resonancia Adaptiva, ART [24] y las redes neuronales de Kohonen [25]. Estas últimas lo que hacen son comparar de forma repetida los vectores que representan los datos de la muestra con los vectores de peso usando una distancia como por ejemplo la distancia euclídea para encontrar grupos de objetos similares.

La red para que funcione bien debe ser sometida a la etapa de entrenamiento para que luego las neuronas se alimenten con los valores de las variables originales para el nuevo objeto.

La **red neuronal de Kohonen (Mapas autoorganizados, SOMs)** se emplea como método de reconocimiento de patrones no supervisado. Es un enfoque alternativo para la visualización de los datos de ACP y análisis de coordenadas principales, PCO. Fue descrito por primera vez por Kohonen en la década de los 80 como un método de visualización.

Tiene la característica de organizar mapas topológicos. El mapa que presenta la red a partir de una situación inicial aleatoria muestra las relaciones existentes entre los diferentes patrones presentados a la red. Esta red muestra de forma efectiva la idea básica de tener una red neuronal artificial que organice un mapa topológico, constituyendo como tal una primera aproximación a los mapas topológicos de los fenómenos motores y sensoriales existentes en la superficie del cerebro humano.

Utiliza la arquitectura paralela de las redes neuronales y provee una representación gráfica de las relaciones entre los patrones (Ver Figura I.21). Un aspecto diferenciador de la red SOM de otras redes es que aprende sin supervisión, de ahí su nombre en inglés. Presenta una topología constituida por dos capas. La primera capa de la red es la capa de entrada y la segunda capa, llamada capa Kohonen está organizada en una rejilla de dos dimensiones. Las dos capas están totalmente interconectadas y cada una de las conexiones tiene asociado un peso que será modificado a lo largo de la sesión de entrenamiento.

Otra herramienta utilizada para la visualización entre neuronas obtenidas en el mapa que permite interpretar los resultados es The Minimal Spanning Tree (MST).

Finalmente, para que la red funcione correctamente es necesario someterla a la etapa de aprendizaje para que así las neuronas de la capa de entrada se alimenten con los valores de las variables manifiestas para el nuevo objeto. Los resultados de la capa de entrada se procesan en la capa escondida, y los resultados generados por ésta se procesan de nuevo en la capa de salida.

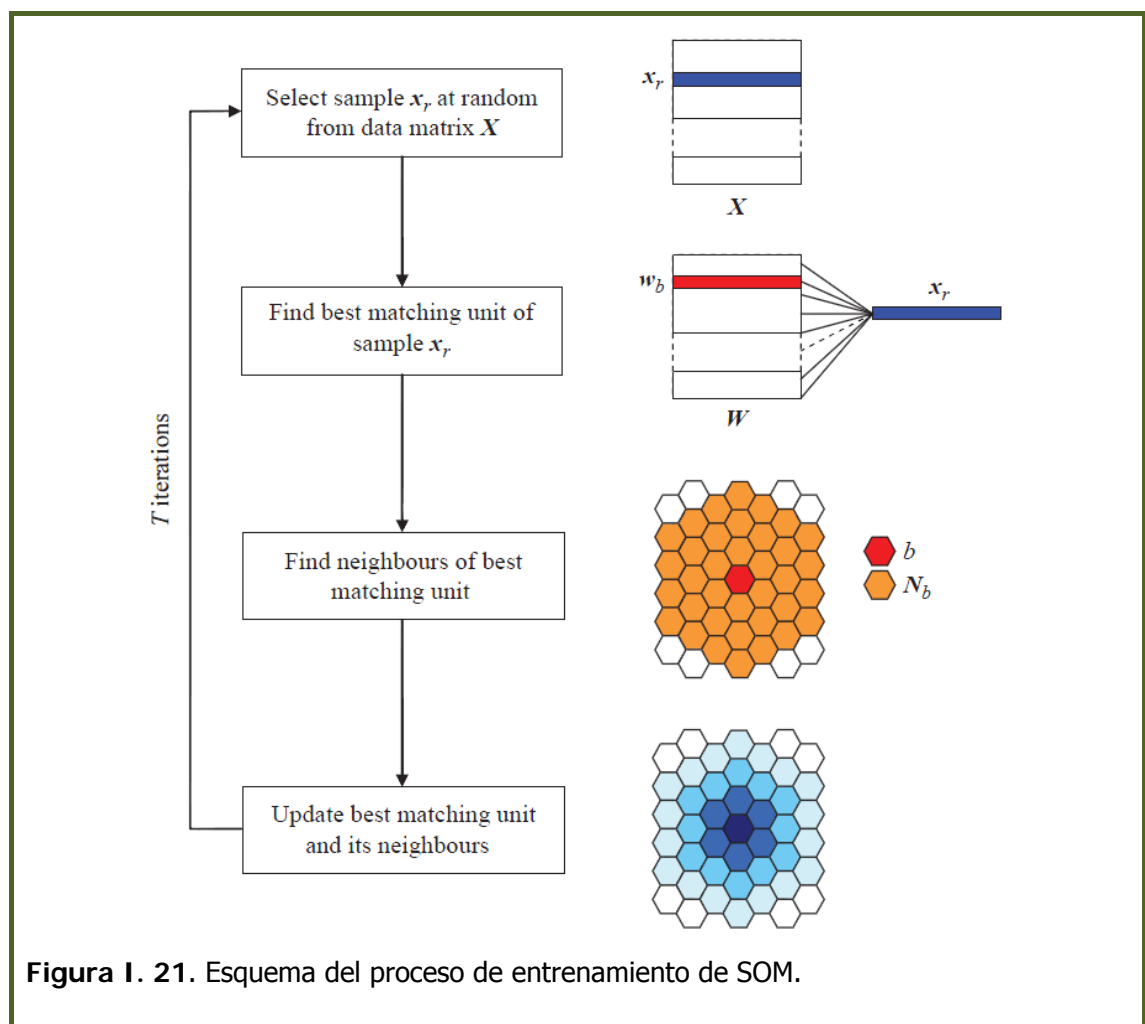
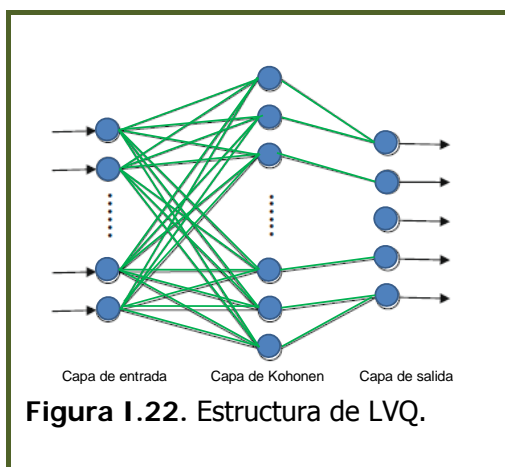


Figura I. 21. Esquema del proceso de entrenamiento de SOM.

Otro tipo de red neuronal es la **LVQ** [8,26] que contiene una capa de entrada, una capa de Kohonen que aprende y realiza la clasificación, y una capa de salida.

La capa de entrada contiene un nodo para cada conexión de entrada, la capa de salida contiene un nodo para cada clase como se muestra en la Figura I.22.

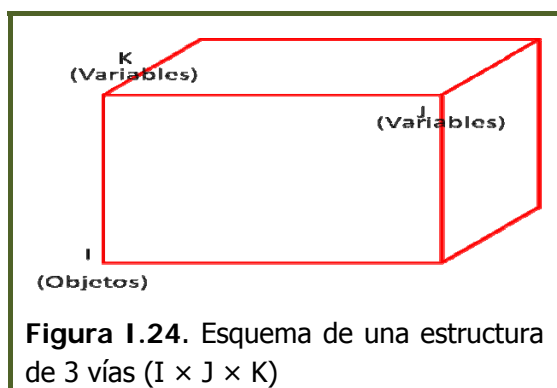
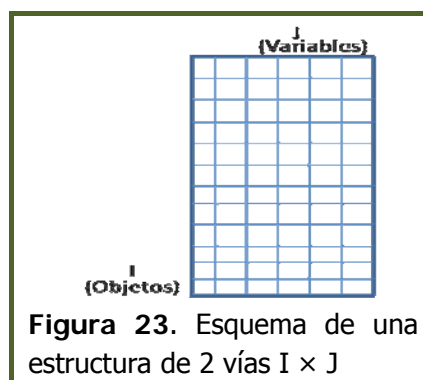


#### D). MULTIPLES VIAS

Un análisis de múltiples vías [27] es un análisis de datos de múltiples vías. Se puede hablar de datos de 1 vía, 2 vías, 3 vías y de un número mayor de vías.

Un ejemplo de datos de 1 vía sería una sola medida en un sistema químico que genera un único número y esta medida se repite, así que el resultado es la secuencia de números.

Si se sube de nivel, los datos de 2 vías son generados a través de instrumentos (GC-FID, UV-vis, etc.) que generan una secuencia de números. Se generaría un matriz **X** de dimensiones  $I \times J$  (Figura I.23).



Si un instrumento sencillo genera una tabla de números para cada muestra se obtendría una matriz **X** ( $J \times K$ ). Si se toman **I** medidas en diferentes ocasiones, genera una matriz de 3 vías  $I \times J \times K$  (Figura I.24).

También es posible generar datos de 4 vías o 5 vías o en general múltiples vías. Hoy en día, el análisis de múltiples vías es ampliamente utilizado en la cromatografía [28] en el que se usan modelos como PARAFAC, PARAFAC2, Tucker y N-PLS para tratar los datos obtenidos a partir por ejemplo de un cromatografo de gases (GC).

### **D.1. Análisis Paralelo de Factores, PARAFAC**

El PARAFAC es un método de descomposición de matrices tridimensionales. Fue desarrollado por Harshman en 1970 e independientemente en el mismo año por Carroll y Chang, los cuales lo llamaron CANDECOMP (descomposición canónica). Es un método de resolución de curvas que descompone la matriz tridimensional  $\mathbf{M}$  ( $\mathbf{I} \times \mathbf{J} \times \mathbf{K}$ ) en tres sub-matrices bidimensionales  $\mathbf{A}$  ( $\mathbf{I} \times \mathbf{R}$ ),  $\mathbf{B}$  ( $\mathbf{J} \times \mathbf{R}$ ) y  $\mathbf{C}$  ( $\mathbf{K} \times \mathbf{R}$ ) siendo cada componente  $\mathbf{R}$  definido por tres vectores de *loadings*, los cuales representan o indican el comportamiento de los componentes  $\mathbf{R}$  en cada una de las tres dimensiones.

El PARAFAC lo que busca es el mejor modelo trilineal que minimice la suma de cuadrados de los residuales en el modelo. Su estructura debe ser obligatoria y absolutamente trilineal. Esto es lo que hace que limite su correcta utilización. Una de las ventajas principales de este modelo es la unicidad de su solución, obtenida mediante optimización por mínimos cuadrados alternados.

Este método ha ganado interés en quimiometría por diferentes razones: simplicidad y posibilidades, por el incremento de la complejidad de los datos obtenidos con instrumentos cada vez más sofisticados y por el incremento computacional de los ordenadores de hoy en día que permiten procesar grandes cantidades de información.

### **D.2. TUCKER3 o ACP a tres vías.**

Este modelo es un modelo más flexible que el conocido modelo PARAFAC. Es útil para análisis, exploración y comprensión de múltiples vías de datos que en ocasiones no es aplicable con PARAFAC. Al igual que en ACP, el modelo Tucker3 tiene rotación libre y cualquier modelo puede rotar sin cambiar el ajuste del modelo.

Al igual que el PARAFAC éste reduce los datos de tres vías de  $\mathbf{M}$  ( $\mathbf{I} \times \mathbf{J} \times \mathbf{K}$ ) a tres matrices:  $\mathbf{A}$ ,  $\mathbf{B}$  y  $\mathbf{C}$ . Se puede considerar como una extensión del APC bilineal para el caso multidimensional.



Una propiedad que tiene este modelo es que no requiere que el número de factores de **A**, **B** y **C** sean los mismos. Los vectores **P**, **Q** y **R** como se describen en el apartado anterior, no tienen que ser necesariamente los mismos y pueden interactuar.

### **D.3. Regresión por Mínimos Cuadrados Parciales Multilineal, N-PLS**

El modelo N-PLS es ajustado describiendo la covarianza de las variables dependientes ( $y$ ) e independientes ( $x$ ). Las variables independientes se modelan de modo que se enfatiza en la variación que es especialmente relevante para predecir las variables dependientes.

La terminología general de N-PLS depende del orden de los datos: el prefijo griego indica el orden de **X** y de un sufijo árabe después de que PLS indica el orden de **Y**. En tri-PLS1, por lo tanto, cada muestra de calibración es caracterizada por una matriz  $X_i$  de dimensiones ( $J \times K$ ), para construir un cubo de datos **X** de dimensiones  $I \times J \times K$ , y para cada muestra hay un valor de mesurando conocido tal que se prediga su valor mediante variables independientes. Estos valores se recogen en una matriz general **Y** de las dimensiones  $I \times L \times M$ . Cuando los mesurando son concentraciones, se recogen en un vector  $I \times 1$  llamado  $y$ . Si más de una variable es encontrada, éstas son recogidas en una matriz **Y** de dimensiones  $I \times L$ , donde  $L$  es el número de analitos (tri-PLS2, tri-PLS3, etc.).

Cuando estamos en presencia de más de una variable dependiente, todos los vectores  $y_i$  pueden ser modelados simultáneamente (como en PLS2) y cada variable dependiente puede ser modelada con el algoritmo PLS1.

La principal característica del algoritmo N-PLS es que produce vectores de scores, en el sentido trilineal, que tienen la máxima covarianza con la parte no explicada de la variable dependiente. La solución es fácil de interpretar, hecho importante sobre todo cuando se tiene un gran número de variables y el algoritmo es rápido debido a la poca cantidad de parámetros que hay que calcular.

El modelo N-PLS es único en el sentido de que consiste en estimar modelos sucesivos de un componente.

## I.4. PRE-PROCESADO DE DATOS

El preprocesado de datos es una etapa previa al análisis de los mismos, siendo su mayor objetivo el linealizar las respuestas y eliminar fuentes extrañas de variabilidad. En esta memoria se han llevado a cabo diferentes tipos de preprocesado aplicados en función de si se trata de datos cuantitativos (concentraciones) o datos de señal analítica (huella dactilar instrumental).

### A). MATRIZ DE DATOS

Cuando la matriz de datos que tenemos se trata de datos de concentración obtenida a partir de la huella dactilar de un instrumento en nuestro caso del GC-FID, se llevan a cabo los siguientes preprocesados:

#### A.1. Centrado de Media.

Se calcula la media de cada columna,  $\bar{X}_k$  y se le resta a cada variable,  $X_{ij}$  de esa columna:

$$X_{ik}^* = X_{ik} - \bar{X}_k$$

donde i es el número de fila y k el número de columna.

#### A.2. Autoescalado

El autoescalado se hace por desviación estándar

$$X_{ik}^* = \frac{X_{ik} - \bar{X}_k}{s_k} \text{ siendo } s_k = \sqrt{\frac{\sum_{i=1}^N (X_{ik} - \bar{X}_k)^2}{N - 1}}$$

N es el número de objetos.

#### A.3. Transformación de BOX-COX

La transformación de potencias de Box-Cox se ha utilizada con el fin de transformar un conjunto de datos a una distribución normal mediante la obtención de una función óptima de transformación que ajuste la variable no lineal a un modelo de regresión lineal, tal que:

$$\mathbf{W} = \beta_0 + \beta_1 \mathbf{X} + \varepsilon$$

donde

W es la variable transformada; X la variable de regresión;  $\beta_0$  y  $\beta_1$  ordenada en el origen y pendiente poblacional en el modelo y  $\varepsilon$  es el error.

La relación entre la variable sin transformar (Y) y la variable transformada W viene dada por la ecuación:

$$\mathbf{W} = 1 + \frac{(\mathbf{Y} + \lambda_2)^{\lambda_1}}{\lambda_1 \mathbf{G}^{\lambda_1 - 1}} \quad \text{para } \lambda_1 \neq 0$$

$$\mathbf{W} = \mathbf{G} \ln(\mathbf{Y} + \lambda_2) \quad \text{para } \lambda_1 = 0$$

Siendo  $\lambda_1$  y  $\lambda_2$  los parámetros de transformación y G la media geométrica de Y +  $\lambda_2$

La transformación óptima se obtiene minimizando el error cuadrático medio de la función calculada W.

## B). MATRIZ DE SEÑAL DE DATOS

### B.1. Alineamiento de la Señal

Para alinear de forma independiente la señal existen varios algoritmos en bibliografía como por ejemplo entre los que se destaca el COW [29] (correlation optimized warping), RSPA [30] (recursive segment-wise peak alignment), MSPA (multiscale peak alignment) [31] y ICOSHIFT [32].

El algoritmo ICOSHIFT, es el que se ha utilizado para eliminar el posible desplazamiento de los cromatogramas debido al uso de la metodología de GC-FID pudiendo permitir el alineamiento de esos cromatogramas (Figura I.25). Este es lo suficientemente rápido para trabajar el conjunto de datos casi en tiempo real.

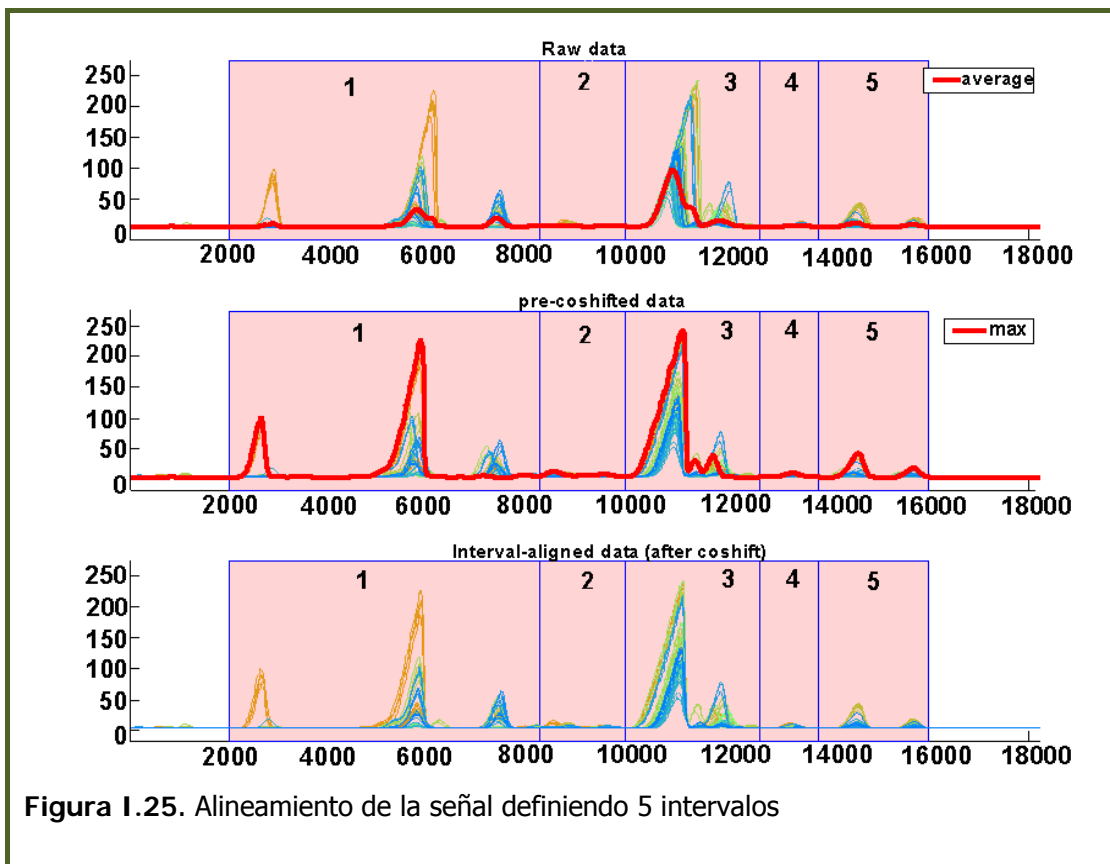


Figura I.25. Alineamiento de la señal definiendo 5 intervalos

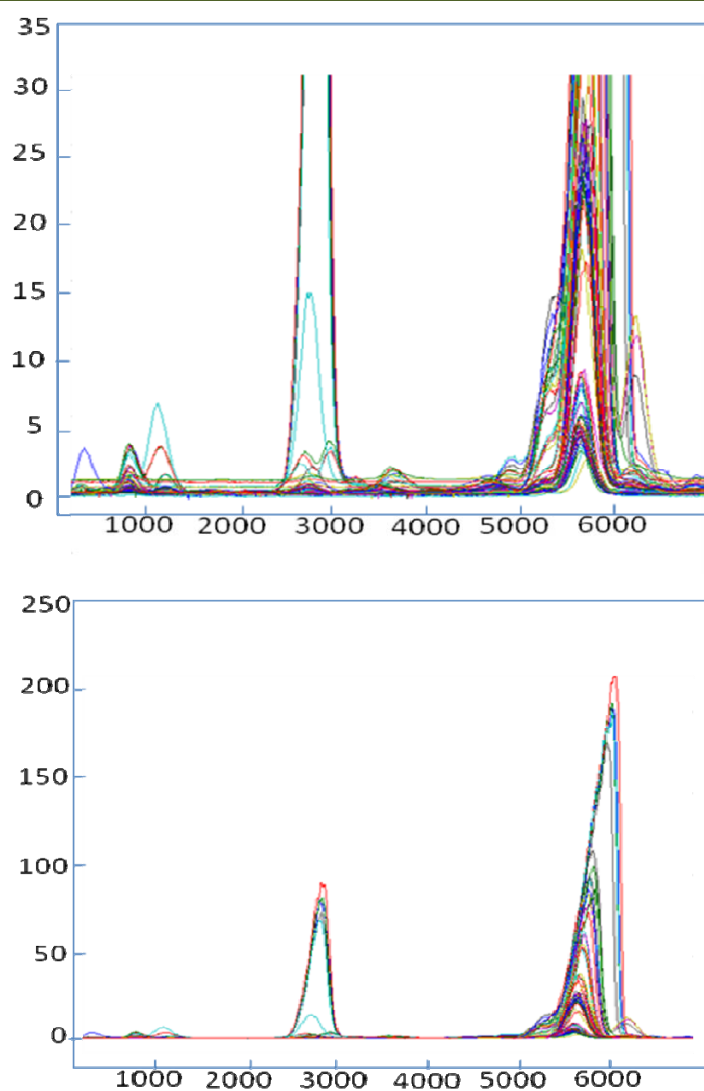
## **B.2. Ajuste de la línea base**

Hay dos opciones:

1. Algoritmo de mínimos cuadrados ponderados (WLS) [33]

Permite establecer los puntos que son más parecidos a los correspondientes de la línea base. Se realiza de forma iterativa permitiendo determinar qué variables están por encima de la línea base o por debajo de esta. El efecto final es la eliminación del fondo evitando crear picos negativos. La línea base está aproximada a un polinomio de bajo orden, normalmente de orden 2, el cual se usa para sustraer la línea base.

A continuación se muestra un ejemplo en el que se aplica el algoritmo WLS a la matriz de datos cromatográficos originales solapados de diferentes muestras (Figura I.26).



**Figura I.26.** Efecto del ajuste de la línea base con WLS. Los datos originales arriba y los datos corregidos en la línea base abajo.

## 2. Función implantada en Bioinformatics Tool\_Box

El ajuste de la línea base a través de una función implantada en el programa Matlab llamada "Bioinformatics Tool\_box".

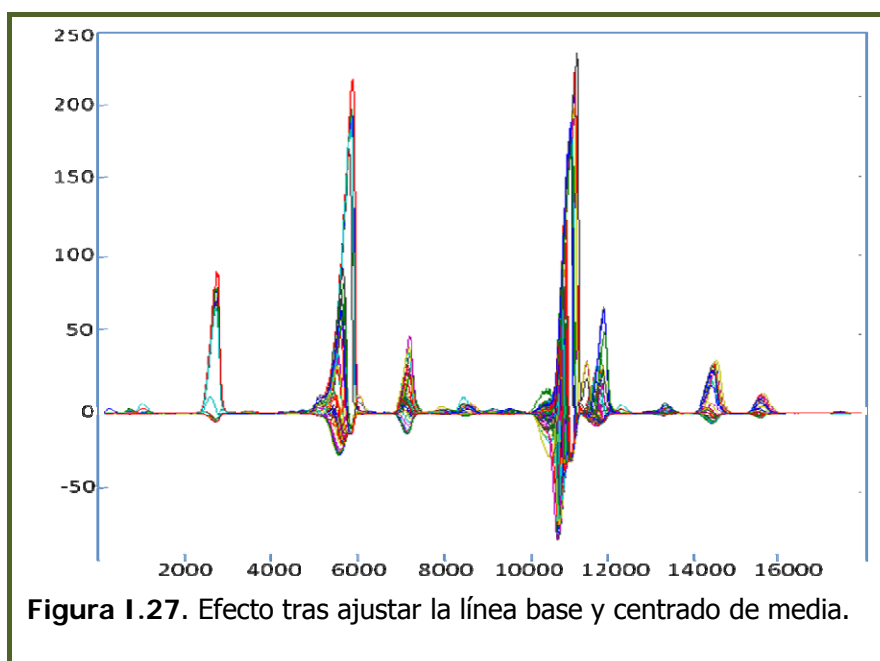
Esta función ajusta la línea base siguiendo 3 pasos: ajuste del desplazamiento de los picos estableciendo una anchura de ventana, ajuste de la línea base usando una aproximación spline y ajuste de la línea base de la señal de los picos por intensidad. El resultado de este algoritmo depende de la elección del tamaño de ventana (350 unidades de separación) que se mueve a través del cromatograma para corregir los picos desalineados y el tamaño de

paso que es 100. En los extremos de la ventana se preserva la altura de los picos.

### **B.3. Centrado de media**

Igual que el apartado 3.1.1.

En la Figura I.27 se muestra un ejemplo de cómo quedarían los datos tras haberle hecho un ajuste de línea base seguido de un centrado de media.



### **B.4. Normalización interna**

La normalización interna que se propone para esta memoria consiste en dividir cada una de las variables que forman parte de la huella dactilar de los TMS-esteroles, por la de mayor intensidad que corresponde con el pico mayoritario en los aceites vegetales, el  $\beta$ -sitosterol. De tal forma que así el valor máximo será 1 y no habrá ningún tipo de problema con la cantidad de muestra tratada ya que no es lo mismo tratar por ejemplo 1 gramo de aceite de oliva virgen extra que 1 gramo de un alimento como por ejemplo una vinagreta.

Cuando se analizan diferentes muestras partiendo de distintas cantidades de muestra existe el problema que al obtener su huella dactilar instrumental, es decir, el cromatograma entero, no tiene porqué presentar la misma escala

independientemente de los picos que aparezcan. Por este motivo, se decide dividir cada punto del cromatograma por el valor máximo.

### C). TEOREMA DE MUESTREO DE NYQUIST

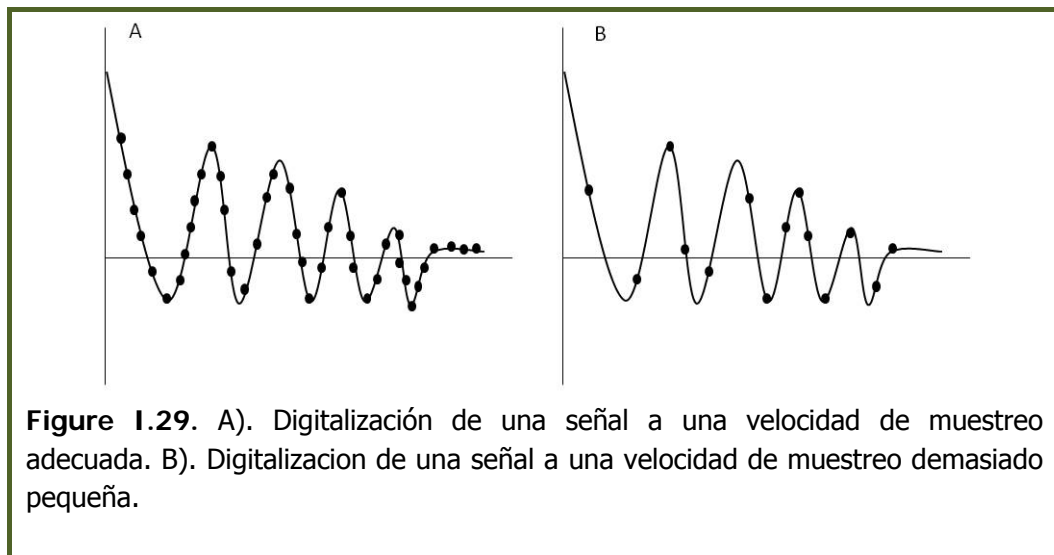
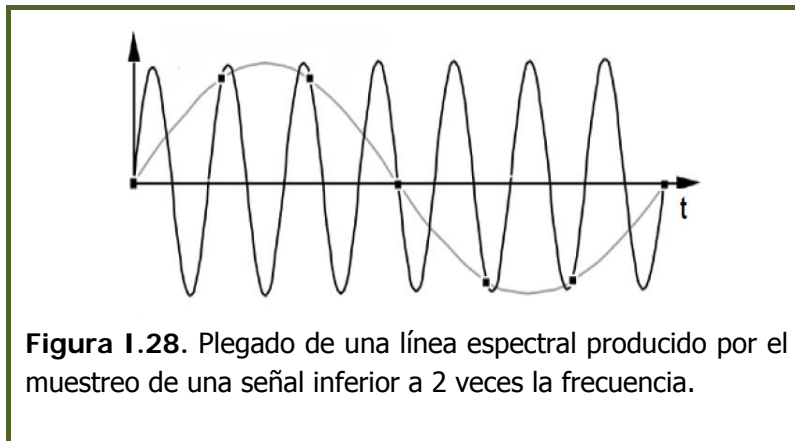
Este teorema fue formulado en forma de conjetura por primera vez por Harry Nyquist en 1928 y fue demostrado formalmente por Claude E. Shannon en 1949. Se enuncia de la siguiente manera:

Una señal  $x(t)$  con un espectro limitado a la frecuencia  $f_B$  ( $|f| < f_B$ ) puede ser muestreada sin pérdida de información si la frecuencia de muestreo  $f_S$  supera la cantidad  $2f_B$ , es decir,  $f_S \geq 2f_B$ .

Esto es, que la frecuencia de la adquisición de datos debe ser de al menos  $2f = \frac{1}{(2\Delta t)}$ , donde  $\Delta t$  es el intervalo de tiempo [34,35].

- Si no se muestrea como mínimo a esa frecuencia tiene lugar el fenómeno denominado "aliasing" o plegamiento espectral.
- Si se muestrea de acuerdo al teorema existe un proceso de reconstrucción que garantiza la reproducción exacta de la señal continua  $x(t)$  sobre el tiempo total de muestreo a partir de sus muestras  $x[n]$ . Este teorema nos garantiza que no se pierde información de la señal original.

Si el periodo de muestreo es demasiado grande (frecuencia de muestreo pequeña) la señal estará pobremente representada en el dominio del tiempo por el conjunto de puntos digitalizados (Figura I.28); se perderá la evolución temporal de la señal entre dos puntos consecutivos y la reconstrucción de la señal a partir de los datos digitalizados. El efecto de esa inadecuada velocidad de muestreo se indica por la línea gris que aparece en la figura I.29. Mientras que cuanto más pequeño sea el período de muestreo tanto más fielmente estará representada la señal en el dominio del tiempo (Figura I.29).



Es aconsejable muestrear de manera reproducible, es decir, empezar a muestrear cada vez en el mismo punto en cada una de las ondas espectrales sucesivas.

Este teorema de muestreo fue aplicado como se verá más adelante en otros capítulos a los perfiles cromatográficos obtenidos de los fitoesteroides en aceites vegetales reduciendo el número de variables de 18201 a 911 sin perder ningún tipo de información. De cada 20 datos se toma 1, es decir, se muestrea cada 20 datos. El muestreo se aplica tras seleccionar el intervalo de tiempo de interés correspondiente entre los 15 y 30 minutos de los 45 minutos que dura el registro del perfil cromatográfico.



## I.5. PRINCIPALES SOFTWARES UTILIZADOS EN QUIMIOMETRÍA

Los diferentes softwares que se utilizan en Quimiometría son:

- Pirouette.
- PLS-Toolbox<sup>1</sup>.
- Unscrambler.
- PARVUS.
- SIMCA-P.
- Sirius.
- Extract.
- MVA.

<sup>1</sup> El PLS\_Toolbox consiste en un paquete de herramientas independientes que trabajan en un ambiente Matlab.

El programa que se ha empleado en esta memoria fue MATLAB versión 7.8.0347 R2009a, (Mathworks Inc., Natick, MA, USA) y el PLS-Toolbox versión 6.2. (Eigenvector Research, Inc., West Eaglerock Drive, Wenatchee, WA).

## BIBLIOGRAFÍA

- [1] Massart, D.L.; Vandeginste, B.G.M.; Buydens, L.C.M.; De Jong, S.; Lewi, P.J. & Smeyers-Verbeke, J. *Handbook of Chemometrics and Qualimetrics: Part A. 1st edition, ed. Elsevier, 1997.*
- [2] Siebert, K.J. *Chemometrics in Brewing. A Review.* J. Am. Soc. Brew Chem. 2001, 59(4), 147-156.
- [3] Ramis Ramos, G.; & García Álvarez-Coque, M.C. *Quimiometría.* Editorial Síntesis, Madrid (España), 2001.
- [4] Kellner, R.; Mermet, J.M.; Otto, M.; Valcárcel, M. & Widmer, H.M. *Analytical Chemistry: A modern approach to analytical science (2<sup>nd</sup> ed).* Wiley-VCH, Weinheim (Germany), 2004.
- [5] Wu, W.; Massart, D.L. & de Jong, S. *The Kernel PCA algorithms for wide data. Part I: Theory and algorithms.* Chemometrics and Intelligent Laboratory Systems, 1997, 36, 165-172.
- [6] Varmuza, K. & Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics.* CRC Press, Boca Raton (Florida, USA), 2009.
- [7] Abollino, O.; Malandrino, M.; Giacomino, A. & Mentasti, E. *The role of chemometrics in single and sequential extraction assays. A review Part I: Extraction, procedures, uni-and bivariate techniques and multivariate variable reduction techniques for pattern recognition.* Analytica Chimica Acta, 2011, 688, 104-121.
- [8] Brereton, R.G. *Chemometrics for Pattern Recognition.* John Wiley & Sons, Chichester (UK), 2009.
- [9] Chae, S.S. & Warde, W.D. *Effect of using principal coordinates and principal components on retrieval of clusters.* Computational Statistics & Data analysis, 2006, 50, 1407-1417.
- [10] Marini, F. *Classification Methods in Chemometrics.* Current Analytical Chemistry, 2010, 6, 72-79.
- [11] Berrueta, L.A.; Alonso-Salces, R.M. & Héberger, K. *Review Supervised pattern recognition in food analysis.* J. Chromatogr. A., 2007, 1158, 196-214.

- [12] Giacomino, A.; Abollino, O.; Malandrino, M. & Mentasti, E. *The role of chemometrics in single and sequential extraction assays: A review. Part II. Cluster analysis, multiple linear regression, mixture resolution, experimental design and other techniques.* Analytica Chimica Acta, 2011, 688, 122-139.
- [13] Flaten, G.R.; Grung, B. & Kvalheim, O.M. *A method for validation of reference sets in SIMCA modelling.* Chemometrics and Intelligent Laboratory Systems, 2004, 72, 101-109.
- [14] Branden, K.V.M. & Hubert, M. *Robust classification in high dimensions based on the SIMCA Method.* Chemometrics and Intelligent Laboratory Systems, 2005, 79, 10-21.
- [15] Wold, S. *Pattern recognition by means of disjoint principal components models.* Pattern Recognit., 1976, 8, 127-139.
- [16] Talavera Bustamante, I. & Rodriguez Hierrezuelo, J.L. *Estado del arte de reconocimiento de patrones en la Quimiometría.* 2008. <http://www.cenatav.co.cu>. (consultada por última vez en junio de 2012).
- [17] Izenman, A.J. *Modern Multivariate Statistical Techniques. Regression, Clasification, and Manifold Learning.* Springer Science, New York (NY, USA), 2008.
- [18] Luts, J.; Ojeda, F.; Van de Plas, R.; De Moor, B.; Van Huffel, S. & Suykens, J.A.K. *A tutorial on support vector machine-based methods for classification problems in chemometrics.* Analytica Chimica Acta, 2010, 665, 129-145.
- [19] Brereton, R.G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant.* John Wiley & Sons, Chichester (UK), 2003.
- [20] De Jong, S. *SIMPLS: An alternative approach to partial least squares regression.* Chemometrics and Inteligent Laboratory Systems, 1993, 18, 251-263.
- [21] Kachbi, A.; Benamor, M.; Aguerssif, N. *Application of SIMPLS Algorithm for simultaneous determination of Co(II), Ni(II) and Cu(II) in industrial alloys. Multivariate analysis and chemometry applied to environment and cultural heritage 2nd ed.,* Ventotene Island, Italy, Europe, 1-4 junio 2008.

- [22] Andrade Garda, J.M.; *Basic Chemometric Techniques in Atomic Spectroscopy*. Royal Society of Chemistry (RSC). Cambridge, 2009.
- [23] Miller, J.N. & Miller, J.C. *Análisis Multivariante* en Miller, J.N. & Miller, J.C. *Estadística y Quimiometría para Química Analítica (4th ed)*. Prentice Hall, Madrid (España), 2002, pp 221-244.
- [24] Kohonen, T. *Self –Organization and Associative Memory*,(3th ed). Springer Verlag, New York, 1989.
- [25] Wienke, D. & Bydens, L. *Adaptive resonance theory based neural Networks- the ART of real- time pattern recognition in chemical process monitoring*. Trends Anal.Chem., 1995, 14, 398-406.
- [26] Liu, J.; Zuo, B.; Zeng, X.; Vroman, P. & Rabenasolo, B. *Nonwoven uniformity identification using wavelet texture analysis and LVQ neural network*. Expert Systems with Applications, 2010, 37, 2241-2246.
- [27] Smilde, A.; Bro, R. & Geladi, P. *Multi-way analysis with applications in the chemical sciences*. John Wiley & Sons, Hoboken (NJ, USA), 2004.
- [28] Bro, R. *Review on multiway analysis in chemistry-2000-2005*. Critical Reviews in Analytical Chemistry, 2006, 36, 279-293.
- [29] Hendriks, M.M.W.B.; Cruz-Juarez, L.; De Bont, D. & Hall, R.D. *Preprocessing and exploratory analysis of chromatographic profiles of plant extracts*. Anal. Chim. Acta, 2005, 545, 53-64.
- [30] Veselkov, K.A.; Lindon, J.C.; Ebbels, T.M.D.; Crockford, D.; Volynkin, V.V.; Holmes, E.; Davies, D.B. & Nicholson, J.K. *Recursive segment-wise peak alignment of biological H-1 NMR spectra for improved metabolic biomarker recovery* . Analytical Chemistry, 2009, 81, 56-66.
- [31] Zhang, Z.M.; Liang, Y.Z.; Lu, H.M.; Tan, B.B.; Xu, X.N. & Ferro, M. *Multiscale peak alignment for chromatographic datasets*. J. Chromatogr. A., 2012, 93-106.
- [32] Savorani, F.; Tomasi, G. & Engelsen, S.B. *icoshift: A versatile tool for the rapid alignment of 2D NMR spectra*. Journal of Magnetic Resonance, 2010, 202, 190-202.

- [33] Wise, B.M.; Gallagher, N.B.; Bro, R.; Shaver, J.M.; Windig, W. & Koch, R.S. *PLS\_Toolbox 4.0 Manual for use with MATLAB™*, Eigenvector Research, Inc. Wenatchee, 2006.
- [34] Skoog, D.A.; Holler, F.J. & Nieman, T.A. *Señales y ruido* en Skoog, D.A.; Holler, F.J. & Nieman, T.A. *Principios de Análisis Instrumental. 5 ed. Mc Graw Hill, Madrid*, 2000, pp 103-120.
- [35] Felinger, A. *Data acquisition* in Felinger. A. *Data analysis and signal processing in chromatography*. Elsevier, Amsterdam (The Netherlands), 1998, pp 80-96.

## **CAPÍTULO II**



## **Materiales y Métodos**

En este capítulo se relacionan los reactivos, patrones, disoluciones, muestras, instrumentación y metodología experimental empleados a lo largo de la memoria.



## II.1. PATRONES DE FITOESTEROLES

Se dispuso de los patrones recogidos en la Tabla II.1. Se procedió a la determinación de su pureza por triplicado a partir de las correspondientes disoluciones en cloroformo, mediante cromatografía de gases, en función del porcentaje del área del pico correspondiente con respecto a la suma de las áreas de todos los picos que aparecen en el cromatograma.

Tabla II.1. Patrones de fitoesteroles.

Esteroles	Fabricante	Pureza Declarada Fabricante	Pureza Obtenida <sup>1</sup>
Campesterol	Xpectrix (Cromasped)	99%	98.28%
5 $\alpha$ -colestán-3 $\beta$ -ol (colestanol)	Sigma Chemical Co.	~95%	99.61%
Colesterol	Sigma Chemical Co.	95%	97.34%
Eritrodiol	Sigma/Fluka	≥97%	100%
Estigmasterol	Sigma Chemical Co.	95%	97.45%
Estigmastanol	Sigma Chemical Co.	95%	94.16%
$\beta$ -sitosterol	Extrasynthèse	≥80%	75.97%
Uvaol	Sigma-Aldrich	98.4%	100%

<sup>1</sup> Pureza obtenida experimentalmente mediante GC

Se prepararon las siguientes **disoluciones madre** de los fitoesteroles mencionados anteriormente.

- *Disolución de campesterol al 0.05 % (m/m) en cloroformo.* Se pesan  $5 \pm 0.01$  mg de analito y se añade cloroformo hasta  $10 \pm 0.01$  g, agitando a continuación en un vortex hasta su completa disolución. De igual manera se prepararon las siguientes disoluciones.
- *Disoluciones de colesterol, estigmasterol, estigmastanol (sitostanol) y  $\beta$ -sitosterol al 1 % (m/m) en cloroformo.*

- *Disolución de uvaol al 0.4 % (m/m) en cloroformo.*
- *Disolución de eritrodiol al 0.1 % (m/m) en cloroformo.*
- *Disolución de 5 $\alpha$ -colestanol al 0.2 % (m/m) en cloroformo.*
- *Disolución de  $\alpha$ -colestanol al 0.2 % (m/v) en acetato de etilo.* Se pesan  $20 \pm 0.01$  mg. Se añade 10 mL de acetato de etilo. Se agita en el vortex hasta su completa disolución.

Estas disoluciones se conservaron a  $-20$  °C en viales topacio de 15 mL con tapón de rosca, comprobándose su estabilidad de forma periódica mediante cromatografía de gases.

Las **disoluciones de trabajo** se prepararon, por dilución con cloroformo, a partir de las disoluciones madre.

## II.2. OTRAS DISOLUCIONES

- *Disolución 2M de hidróxido potásico en etanol.* Se añade 20 mL de agua desionizada a 13 g de hidróxido potásico, se agita y se lleva la mezcla, con etanol hasta un volumen de 100 mL. Esta disolución se debe conservar en frasco oscuro de vidrio, un máximo de dos días.
- *Mezcla derivatizante:* Disolución de piridina/hexametildisilazano/trimetilclorosilano en la proporción 9:3:2 (V/V/V), conservada en vial topacio con tapón de rosca.

## II.3. REACTIVOS Y DISOLVENTES

- Hidróxido potásico (85%). Para análisis. Panreac.
- Sulfato sódico anhidro (99%). Para análisis. Panreac.
- Papel indicador de pH. Merck.
- Piridina seca (99.5%). DS-ACS. Panreac.

## Materiales y Métodos

- Hexametildisilazano (97%). Para síntesis. Sigma.
- Trimetilclorosilano (97%). Para cromatografía de gases. Sigma.
- Cloroformo (estabilizado con etanol) (99%). Para análisis instrumental. Panreac.
- Acetonitrilo (99.8%). Para HPLC. Panreac.
- Ácido acético ( $\geq 99.7\%$ ). Para análisis. Panreac.
- Metanol (99.8%). Panreac.
- Etanol (96%). Para análisis. Panreac.
- Eter dietílico (99.5%). Para HPLC. Merck.
- n-Hexano (97%). Para HPLC. Merck.
- Tert-butlimetiléter (TBME) ( $\geq 99.8\%$ ). Para HPLC. Merck.
- Acetona (99.5%). Para análisis. Panreac.
- Agua desionizada (MilliQ 18.2 M $\Omega$ cm).
- Isopropanol (99.9%). Para análisis. Panreac.
- Helio AIR LIQUID ALFA. 99.9999% de pureza.
- Nitrógeno AIR LIQUID ALFA. 99.9999% de pureza.

## II.4. EQUIPOS DE LABORATORIO

### A). Instrumentos de Medida

- **Balanza analítica** verificada capaz de proporcionar medidas fiables en un intervalo de 0.001 – 410 g con una resolución mínima de 0.001 g.

- **Cromatógrafo de líquidos de alta resolución** HEWLETT-PACKARD, serie 1050, equipado con:
  - ❖ Detector espectrofotométrico de longitud de onda variable.
  - ❖ Inyector manual RHEDOYNE 7125 con un bucle de 20  $\mu\text{L}$ .
  - ❖ Microjeringa HAMILTON para HPLC, modelo 1810 RNE, con 100  $\mu\text{L}$  de capacidad.
  - ❖ Columnas:
    - LiChrospher® 100 RP-8, (125 x 4 mm i.d., 5  $\mu\text{m}$ ).
    - LiChrospher® 100 CN, (244 x 4.4 mm i.d., 5  $\mu\text{m}$ ), equipada con una precolumna (10 x 4.6 mm i.d.) empaquetada con la misma fase estacionaria.

Previamente a cada sesión de trabajo con el cromatógrafo de líquidos, se preparaban, se filtraban y se desgasificaban las disoluciones que se iban a utilizar en la composición de la fase móvil.

- **Cromatógrafo de gases** AGILENT TECHNOLOGIES, modelo 6890, equipado con:
  - ❖ Detector de ionización de llama (FID).
  - ❖ Sistema de inyección Split/splitless manual.
  - ❖ Microjeringa HAMILTON para GC, modelo 1701 N, con 10  $\mu\text{L}$  de capacidad.
  - ❖ Columna DB-5 (25 m x 0.32 mm x 0.25  $\mu\text{m}$ ) (J&W, Folsom, CA) recubierta interiormente con una fase de fenilmetilsilicona al 5%.

Así mismo, antes de cada sesión de trabajo, se sometía la columna tanto para HPLC como para GC a un proceso de equilibración con la fase móvil correspondiente hasta obtener una línea base suficientemente estable.

## B). Instrumentos Generales

- Mezclador VORTEX HEIDOLPH, modelo Reax 2000.

- Centrífuga BHG, modelo FIXETTE II.
- Bomba de agua para vacío BUCHI B-169 Vacuum System.
- Placa agitadora / calefactora SELECTA modelo AGIMATC-N.
- Rotavapor BÜCHI.
- Manta eléctrica SELECTA.
- Estufa SELECTA.
- Homogenizador WIGGEN HAUSER.

### C). Material de Laboratorio

- Kit de filtración de disolventes MILLIPORE de 47 mm de diámetro.
- Filtros de membrana MILLIPORE FG FGLP04700 para filtración de disolventes orgánicos, hidrofóbicos, con trama de 0.20  $\mu\text{m}$  de tamaño de poro.
- Refrigerante de reflujo.
- Matraces de fondo redondo de diferentes capacidades.
- Matraces de forma de corazón de 10 mL.
- Embudos de decantación de distintos tamaños.
- Probetas de distintos tamaños.
- Vasos de precipitado de diferentes capacidades.
- Pipetas graduadas y verificadas de tipo 1 y 2, de diferentes capacidades.
- Micropipetas verificadas con volumen nominal de 200, 1000 y 5000  $\mu\text{L}$ .
- Viales de vidrio transparentes y de diferentes capacidades, de color ámbar y con tapones de rosca.

- Tubos de centrifuga de 10 mL de capacidad.
- Papel de filtro de cenizas conocidas.
- Desecador.

#### D). Material Informático

Programas utilizados:

- Statgraphics Centurion XV (Versión 15.2.05 n) for Windows (Statpoint Technologies Inc.).
- Matlab (versión 7.8.0347 R2009a) The Mathworks Inc. Natick, MA, USA.
- PLS\_Toolbox (versión 6.2., Eigenvector Research Inc., West Eaglerock Drive, Wenatchee, WA).
- ChemDraw Ultra 8.0.

### II.5. MUESTRAS ANALIZADAS

A lo largo del trabajo desarrollado en esta memoria se han analizado las siguientes muestras:

- **Aceites vegetales** del tipo oliva virgen extra, virgen, refinado, orujo de oliva, girasol, girasol alto oleico, soja, maíz, colza, cáñola, sésamo, pepita de uva, lino, cacahuete y semillas, adquiridos en diferentes supermercados y procedentes de España, Francia, Bélgica, Italia, México, E.E.U.U. y Japón. (Tabla II.2).
- **Aceites Vegetales** facilitados por el Laboratorio Agroalimentario de Granada (Atarfe). Dicho laboratorio nos suministró dos muestras de aceite de orujo de oliva (AT-POO-1 y AT-POO-2), dos de aceite de oliva virgen extra (AT-EVOO-1 y AT-EVOO-2) y una de aceite de girasol alto oleico (AT-SFO(o)) que fueron utilizadas en las etapas de verificación y validación del método cromatográfico propuesto.

**Tabla II.2** Listado muestras de aceites vegetales analizadas.

<b>CÓDIGO DE ACEITE</b>	<b>CÓDIGO DE MUESTRA</b>	<b>TIPO ACEITE - MARCA</b>	<b>ORIGEN</b>
Ore(i)-HA	ROO-1	Oliva Refin. Sabor Intenso - <i>Hacendado</i>	España
Ore(s)-HA	ROO-2	Oliva Refin. Sabor Suave - <i>Hacendado</i>	España
Ore(s)-OS	ROO-3	Oliva Refin. - <i>Oli Sone</i>	España
Ore(s)-MA	ROO-4	Oliva Refin. - <i>La Masia</i>	España
Ore(s)-CO	ROO-5	Oliva Refin. - <i>Covirán</i>	España
Oru-AO	POO-1	Orujo - <i>As de Oros</i>	España
Oru-MA	POO-2	Orujo - <i>La Masia</i>	España
Oru-CA	POO-3	Orujo - <i>Capicua</i>	España
Oru(s)-CA	POO-4	Orujo Sabor Suave - <i>Capicua</i>	España
Oru-LU	POO-5	Orujo - <i>Lucioliva</i>	España
Ove(pic)-CT	EVOO-1	Oliva Virgen Extra Picual - <i>Castillo de Tabernas</i>	España
Ove(pic)-CS	EVOO-2	Oliva Virgen Extra Picual - <i>Coosur</i>	España
Ove(pic)-FR	EVOO-3	Oliva Virgen Extra Picual - <i>Fragata</i>	España
Ove(pic)-OC	EVOO-4	Oliva Virgen Extra Picual - <i>OleoCazorla</i>	España
Ove(pic)-SÑ	EVOO-5	Ove(Picual) - <i>Señorío de Segura</i>	España
Ove(pic)-CA	EVOO-6	Ove(Picual) - <i>Carbonel</i>	España
Ove(arb)-CC	EVOO-7	Oliva Virgen Extra Arbequina - <i>Castillo Canena</i>	España
Ove(arb)-OL	EVOO-8	Oliva Virgen Extra Arbequina - <i>Oleaurum Verd</i>	España
Ove(arb)-VA	EVOO-9	Oliva Virgen Extra Arbequina - <i>Valderrama</i>	España

Tabla II.2 (Cont.) Listado muestras de aceites vegetales analizadas.

CÓDIGO DE ACEITE	CÓDIGO DE MUESTRA	TIPO ACEITE - MARCA	ORIGEN
Ove(arb)-FR	EVOO-10	Oliva Virgen Extra Arbequina - <i>Fragata</i>	España
Ove(arb)-VE	EVOO-11	Oliva Virgen Extra Arbequina - <i>Vea</i>	España
Ove(cor)-HI	EVOO-12	Oliva Virgen Extra Cornicabra - <i>Hipercor</i>	España
Ove(cor)-NA	EVOO-13	Oliva Virgen Extra Cornicabra - <i>Sierra Nambroca</i>	España
Ove(cor)-AZ	EVOO-14	Oliva Virgen Extra Cornicabra - <i>Arzuaga</i>	España
Ove(cor)-1	EVOO-15	Oliva Virgen Extra Cornicabra - <i>1</i>	España
Ove(cor)-2	EVOO-16	Oliva Virgen Extra Cornicabra - <i>2</i>	España
Ove(hoj)-AN	EVOO-17	Oliva Virgen Extra Hojiblanca - <i>Antequera</i>	España
Ove(hoj)-CA	EVOO-18	Oliva Virgen Extra Hojiblanca - <i>Carbonell</i>	España
Ove(hoj)-CL	EVOO-19	Oliva Virgen Extra Hojiblanca - <i>Columela</i>	España
Ove(hoj)-CS	EVOO-20	Oliva Virgen Extra Hojiblanca - <i>Coosur</i>	España
Ove(hoj)-FR	EVOO-21	Oliva Virgen Extra Hojiblanca - <i>Fragata</i>	España
Ove(hoj)-AU	EVOO-22	Oliva Virgen Extra Hojiblanca - <i>Auchan</i>	España
Ove(man)-DE	EVOO-23	Oliva Virgen Extra Manzanilla - <i>Deguste</i>	España
Ove(man)-FR	EVOO-24	Oliva Virgen Extra Manzanilla - <i>Fragata</i>	España
Ove(roy)-BU	EVOO-25	Oliva Virgen Extra Royal - <i>Burrunchel</i>	España
Ove(roy)-FR	EVOO-26	Oliva Virgen Extra Royal - <i>Fragata</i>	España



**Tabla II.2 (Cont.)** Listado muestras de aceites vegetales analizadas.

<b>CÓDIGO DE ACEITE</b>	<b>CÓDIGO DE MUESTRA</b>	<b>TIPO ACEITE - MARCA</b>	<b>ORIGEN</b>
Ove-AU	EVOO-27	Oliva Virgen Extra- <i>Auchan</i>	España
Ove-SS	EVOO-28	Oliva Virgen Extra - <i>Sierra Segura</i>	España
Ove-FR	EVOO-29	Oliva Virgen Extra - <i>Fragata</i>	España
Ove-TP	EVOO-30	Oliva Virgen Extra - <i>Terras de Portugal</i>	España
Ove-CT	EVOO-31	Oliva Virgen Extra – <i>Castillo de Tabernas</i>	España
Ove-US	EVOO-32	Oliva Virgen Extra - <i>Uscar</i>	España
Ovi-MA	EVOO-33	Oliva Virgen - <i>Mariscal</i>	España
Ovi-MC	EVOO-34	Oliva Virgen - <i>Marruecos</i>	Marruecos
Ovi-AN	EVOO-35	Oliva Virgen - <i>N<sup>a</sup> Sra. Angustias</i>	España
Ovi-SE	EVOO-36	Oliva Virgen – <i>San Sebastián</i>	España
Cac-LE	PeaO-1	Cacahuete - <i>Lesieur</i>	Francia
Cac-CA	PeaO-2	Cacahuete- <i>Casino</i>	Francia
Cac-MO	PeaO-3	Cacahuete – <i>Monoprix</i>	Francia
Cac-AU	PeaO-4	Cacahuete – <i>Auchan</i>	Francia
Cac-VIT	PeaO-5	Cacahuete - <i>Vitasia</i>	Bélgica
Can-CR	CanO-1	Cánola - <i>Crisco</i>	Estados Unidos
Can-PW	CanO-2	Cánola - <i>Pure Wesson</i>	Estados Unidos

Tabla II.2 (Cont.) Listado muestras de aceites vegetales analizadas.

CÓDIGO DE ACEITE	CÓDIGO DE MUESTRA	TIPO ACEITE – MARCA	ORIGEN
Can-GH	CanO-3	Cánola – <i>Golden Hills</i>	México
Can-MA	CanO-4	Cánola - <i>Maravilla</i>	Estados Unidos
Col -1	RO-1	Colza - <i>1</i>	Francia
Col-CA	RO-2	Colza - <i>Casino</i>	Francia
Col-LE	RO-3	Colza - <i>Lesieur</i>	Francia
Col-MI	RO-4	Colza - <i>Mille et une huiles</i>	Francia
Gir-CO	SFO-1	Girasol – <i>Coosol</i>	España
Gir- ER	SFO-2	Girasol – <i>Eroski</i>	España
Gir-AL	SFO-3	Girasol – <i>Aliada</i>	España
Gir-CI	SFO-4	Girasol – <i>Corte Inglés</i>	España
Gir-CV	SFO-5	Girasol – <i>Covirán</i>	España
Gir-FU	SFO-6	Girasol – <i>Fuensol</i>	España
Gir-KO	SFO-7	Girasol – <i>Koipesol</i>	España
Gir-LA	SFO-8	Girasol – <i>Lanisol</i>	España
Gir-SS	SFO-9	Girasol – <i>Supersol</i>	España
Gir-VE	SFO-10	Girasol – <i>Végé</i>	España

**Tabla II.2 (Cont.)** Listado muestras de aceites vegetales analizadas.

<b>CÓDIGO DE ACEITE</b>	<b>CÓDIGO DE MUESTRA</b>	<b>TIPO ACEITE - MARCA</b>	<b>ORIGEN</b>
Gir(o)-SU	SFO(o)-1	Girasol alto oleico – <i>Suavesol</i>	España
Gir(o)-CO	SFO(o)-2	Girasol alto oleico – <i>Coosol</i>	España
Gir(o)-ER	SFO(o)-3	Girasol alto oleico – <i>Eroski</i>	España
Mai-AS	CO-1	Maíz – <i>Asua</i>	España
Mai-MA	CO-2	Maíz - <i>La Masia</i>	España
Mai-GH	CO-3	Maíz - <i>Golden Hills</i>	México
Mai-HA	CO-4	Maíz - <i>Hacendado</i>	España
Lin-BI	FO-1	Lino - <i>Biolasi</i>	España
Lin-FI	FO-2	Lino – <i>La Finesta sul cielo</i>	Italia
Lin-SO	FO-3	Lino – <i>Sol Natural</i>	España
Sem-ER	SeeO-1	Semillas – <i>Eroski</i>	España
Sem-HA	SeeO-2	Semillas – <i>Hacendado</i>	España
Sem-MA	SeeO-3	Semillas – <i>La Masia</i>	España
Sem-TI	SeeO-4	Semillas – <i>Titan</i>	España
Sem-SOL	SeeO-5	Semillas - <i>Dalysol</i>	España
Ses-CH	SesO-1	Sésamo – <i>Chee seng</i>	Singapur
Ses-EC	SesO-2	Sésamo – <i>Eco Salim</i>	España

Tabla II.2 (Cont.) Listado muestras de aceites vegetales analizadas.

CÓDIGO DE ACEITE	CÓDIGO DE MUESTRA	TIPO ACEITE - MARCA	ORIGEN
Ses-JP	SesO-3	Sésamo – <i>Japonés</i>	Japón
Ses-VIT	SesO-4	Sésamo – <i>Vitasia</i>	Bélgica
Soj-CL	SyO-1	Soja – <i>Clearspring</i>	Italia
Soj-MA	SyO-2	Soja – <i>Maurel</i>	Francia
Soj-CR	SyO-3	Soja - <i>Cristal</i>	Marruecos
Soj-BI	SyO-4	Soja – <i>Biolasi</i>	España
Soj-GH	SyO-5	Soja – <i>Golden Hills</i>	México
Uva-AU	GO-1	Uva – <i>Auchan</i>	Francia
Uva-DI	GO-2	Uva - <i>Dietisa</i>	España
Uva-BI	GO-3	Uva- <i>Biosan</i>	España
Uva-EV	GO-4	Uva - <i>Emile Noël</i>	España

- **Vinagretas**, como materiales de referencia terciarios y comerciales (Tablas II.3 y II.4)).

**Materiales de referencia terciarios:**

La elaboración de los materiales de referencia terciarios se llevó a cabo siguiendo una receta clásica para preparar vinagretas en la que se emplea una parte de vinagre por tres de aceite: Se pesa un vial de 4 mL de capacidad, y se añaden, con una pipeta automática de 5 mL, 2.25 mL de aceite vegetal. Se anota el peso y a continuación se añaden 0.75 mL de vinagre y se vuelve a

anotar el peso. Finalmente, se añaden  $0.001 \pm 0.001$  g de sal y con un homogeneizador se mezclan bien los ingredientes durante al menos 10 minutos para asegurarse de que se forme la emulsión. De esta forma se prepararon los materiales de referencia recogidos en la Tabla II.3.

Estos materiales deben ser analizados justamente después de su elaboración.

Tabla II.3. Materiales de Referencia Terciarios.

CÓDIGO	ACEITES EMPLEADOS EN LA ELABORACIÓN DE LAS SALSAS
VRM1	100 % Oliva
VRM2	100 % Oliva
VRM3	50 % Oliva : 50 % Girasol
VRM4	25 % Oliva : 75 % Girasol
VRM5	75 % Oliva : 25 % Colza
VRM6	60 % Oliva : 40 % Soja
VRM7	90 % Oliva : 10 % Maíz
VRM8	20 % Oliva : 80 % Canola
VRM9	40 % Oliva : 60 % Maíz
VRM10	10 % Oliva : 90 % Soja
VRM11	50 % Canola : 50 % Soja
VRM12	75 % Semillas : 25 % Girasol
VRM13	25 % Maíz : 75 % Girasol
VRM14	60 % Orujo : 40 % Soja
VRM15	80 % Orujo : 20 % Girasol

### Vinagretas comerciales:

Se adquirieron diferentes vinagretas y mezclas para aliño, que se mantuvieron en el frigorífico a 4 °C hasta el momento de su análisis (Tabla II.4).

Una vez que la muestra se encuentra a temperatura ambiente, es agitada vigorosamente con objeto de homogenizarla. A continuación, se filtra a través de un colador de tela de 12 cm de diámetro y el filtrado (Figura II.1) se recoge sobre el vaso de una batidora, ayudándose de una varilla de vidrio debido a que algunas vinagretas son muy espesas.



Figura II.1. Filtración Vinagreta.

Después de emulsionar la muestra filtrada con una batidora, se somete al proceso de saponificación y se continúa con el procedimiento hasta obtener la fracción esterólica.

**Tabla II.4.** Vinagretas y mezclas para aliño comerciales.

CÓDIGO	TIPO ALIMENTO - MARCA	ORIGEN
ALI	Aliño – <i>Calvé</i>	España
VIN-MOD	Vinagreta al Vinagre de Módena - <i>Borges</i>	España
VIN-MOS	Vinagreta de Mostaza – <i>Borges</i>	España
VIN-SEC	Vinagreta de Frutos Secos al Vinagre de Módena Blanco – <i>Borges</i>	España
VIN-ORI	Vinagreta al Toque Oriental – <i>Borges</i>	España

## II.6. PROCEDIMIENTO OPERATORIO DEL MÉTODO OFICIAL [1]

### Preparación de la fracción insaponificable

En un matraz de 250 mL, se introducen 500  $\mu\text{L}$  (para aceite de oliva) o 1500  $\mu\text{L}$  (para aceites vegetales) de disolución patrón de  $\alpha$ -colestanol al 0.2% (m/v) en cloroformo. Se lleva a sequedad mediante corriente de  $\text{N}_2$  y a continuación se pesan  $5 \pm 0.1$  g de aceite. Se añade 50 mL de disolución etanólica de hidróxido potásico 2M, se coloca un refrigerante de reflujo y se calienta a ebullición suave durante 1 hora.

Se deja enfriar ligeramente el matraz de reacción (30°C aprox.), se añaden 100 mL de agua, se transfiere la disolución a un embudo de decantación con 80 mL de éter etílico, se agita vigorosamente durante unos 30 s y se deja decantar (Si se forma una emulsión que no desaparece, se añaden pequeñas cantidades de etanol). La fase inferior (acuosa) se trasvasa a un segundo embudo de decantación y se extrae nuevamente con 60-70 mL de éter etílico. Esta operación se realiza 2 veces.

[1] COI/T.20/Do. nº. 10/Rev.1. 2001. Determinación de la composición y del contenido de esteroides mediante cromatografía de gases en columna capilar. Consejo Oleícola Internacional, Madrid.

Los extractos de éter etílico de ambos embudos se juntan y se lavan tres veces, con agua (50 mL cada vez) hasta que su pH sea neutro.

Posteriormente, la disolución de éter etílico se seca y se filtra sobre sulfato sódico anhidro, y se evapora en el rotavapor hasta que queden unos pocos mL. Se pasa una corriente de N<sub>2</sub> durante un par de minutos asegurándose de que se haya eliminado todo el disolvente. Se completa el secado en una estufa a 100 °C durante 15 minutos.

### Separación de la fracción de esteroides

Sobre residuo obtenido de la fracción insaponificable, previamente pesado, se añade un volumen tal de cloroformo que la disolución obtenida sea del 5% aproximadamente. Con una microjeringa de 100 µL se depositan 300 µL de dicha disolución en una placa de gel de sílice para proceder a la separación de la fracción de esteroides. Se coloca la placa en una cubeta de desarrollo, previamente equilibrada con una fase móvil de composición hexano: éter etílico 65:35 (v/v), y se tapa para que se produzca la elución hasta que el frente del disolvente se sitúe a 1 cm aproximadamente del borde superior de la placa.

Se pulveriza la placa con una disolución de 2,7-diclorofluoresceína, tras haber evaporado el disolvente. Se identifica la banda de los esteroides a la luz ultravioleta mediante comparación con la mancha obtenida a partir de una disolución de referencia. Se raspa el gel de sílice contenido en el área delimitada, se introduce en un embudo filtrante, se añaden 10 mL de cloroformo caliente y a continuación se lava tres veces el sólido con éter etílico, empleando cada vez 10 mL. El filtrado se lleva a sequedad mediante rotavapor, se pasa una corriente de N<sub>2</sub>, se introduce en una estufa a 105 °C durante 10 minutos, se deja enfriar en el desecador y se pesa.

### Preparación de los trimetilsililéteres (TMS-éteres)

A la fracción de esteroides obtenida se le añade el reactivo de silanización formado por una mezcla de piridina:hexametildisilazano:clorotrimetilsilano en las proporciones 9:3:1 (v/v/v), a la razón de 50 µL / mg esteroides, evitando toda absorción de humedad.

Finalmente, se agita la disolución con los TMS-éteres, se deja reposar 15 min a temperatura ambiente y se centrifuga durante un par de minutos.

La disolución ya está lista para el análisis mediante cromatografía de gases.

## **CAPÍTULO III**





# Desarrollo de un método para la determinación de 4-desmetilesteroles en aceites vegetales, previa separación "off-line" mediante HPLC

Como se ha comentado previamente en la introducción, la forma más habitual de realizar la determinación de fitoesteroles totales (suma de fitoesteroles libres más esterificados), implica una etapa previa de aislamiento de los mismos a partir de la fracción insaponificable del aceite vegetal. En la mayoría de los casos, los procedimientos desarrollados, basados a su vez en métodos oficiales, emplean como técnica separativa la cromatografía en capa fina (CCF).

Con objeto de soslayar esta etapa tediosa, que aumenta el tiempo de análisis, se conocen diferentes alternativas entre las que se pueden citar, la extracción en fase sólida (SPE) **[1,2,3]**, la cromatografía flash **[4]** o la cromatografía líquida de alta resolución (HPLC).

En relación con esta última, cabe indicar que:

- (i) en la modalidad "off-line", solo se conocen dos referencias **[5,6]** en las que, para la separación de la fracción fitoesterólica en aceites, se utiliza la modalidad de fase normal (NP-HPLC) usando como fase estacionaria una columna Lichrospher Si 60 de 250 mm y como fase

móvil una mezcla de n-hexano: éter dietílico, 50:50 (v/v). En ambos casos se trabaja en régimen isocrático y la detección se realiza mediante un detector de índice de refracción. No obstante, existen ligeras diferencias en las condiciones utilizadas. Así, Cert et al. [5], seleccionan un flujo de fase móvil de 1 mL/min, consiguiendo, a partir del minuto diez del cromatograma, separar en dos picos diferenciados los  $\Delta^5$ - y  $\Delta^7$ -esteroles. Por su parte Peinado [6] lleva a cabo la separación conjunta de los analitos utilizando un flujo de fase móvil de 1.5 mL/min, manteniendo constante la temperatura de la columna a 40 °C.

(ii) en la modalidad "on-line" acoplada con cromatografía de gases (LC-GC), cabe indicar que, tal y como aparece recogido en la introducción de esta memoria, se han desarrollado métodos en los que se emplean tanto las modalidades en fase normal [7, 8, 9, 10] como en fase invertida [11, 12, 13].

En relación a la forma en la que se introduce la muestra en el sistema cromatográfico, varía dependiendo de la finalidad de la determinación (esteroles totales o libres y esterificados). Así, Toledano et al. han desarrollado métodos en los que, o bien se introduce directamente la muestra de aceite diluida en el disolvente adecuado, o como una alícuota de la fracción insaponificable [12, 14].

Tras estos antecedentes, este estudio tiene como objetivo el desarrollo de un método off-line, mediante cromatografía líquida de alta resolución (HPLC) con detección ultravioleta visible (UV-Vis), para recolectar la fracción que contiene los 4-desmetilesteroles, con vistas a su posterior derivatización y análisis por GC.

Para ello se ha tomado como base el método propuesto por el COI [15], en el cual, como ya se ha comentado en la introducción, la fracción de esteroles se obtiene mediante cromatografía en capa fina y posteriormente los analitos se transforman en trimetilsililéteres para su cuantificación mediante CG-FID, ya que los trimetilsililderivados son mucho más volátiles, menos polares y más estables térmicamente que los compuestos originales.

Con objeto de asegurar una separación y cuantificación reproducibles, se procedió, en primer lugar, a la evaluación de la reacción de derivatización y,

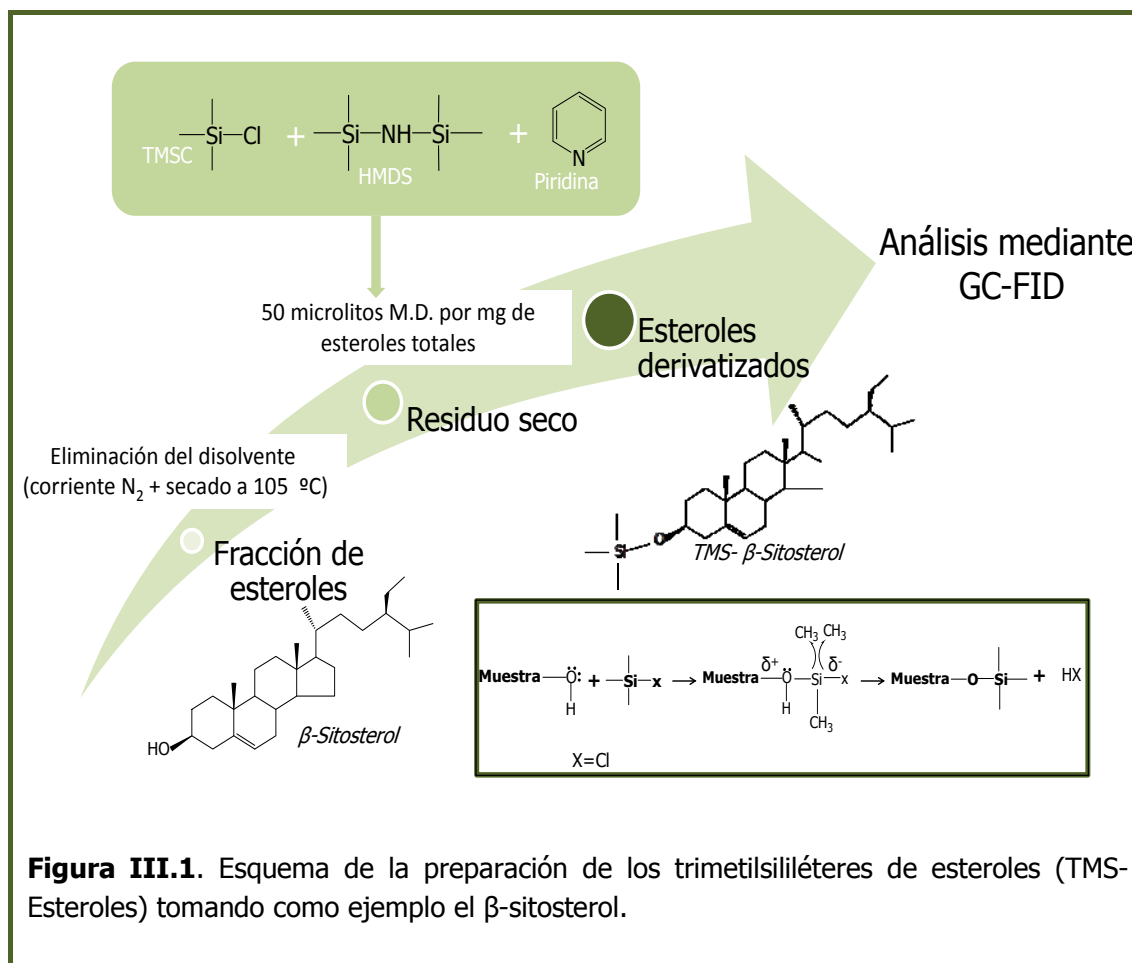
posteriormente, a la forma de cuantificación de los analitos. Una vez optimizado el método, se evaluó su incertidumbre y se validó.

A lo largo del trabajo recogido en esta memoria, la determinación de los trimetilsililderivados se ha realizado mediante GC-FID, empleando las condiciones operatorias indicadas a continuación:

- <b>Equipo:</b>	HP 6890
- <b>Columna:</b>	DB-5 (25m×0.32mm×0.25µm) recubierta interiormente con una fase de fenilmetilsilicona al 5%
- <b>Gas portador:</b>	Helio
- <b>Flujo:</b>	1.4 mL/min
- <b>Tª Inyector:</b>	280 °C
- <b>Tª Horno:</b>	260 °C (régimen isotérmico)
- <b>Tª Detector:</b>	290 °C
- <b>Modo de inyección:</b>	Split 1:20
<b>Cantidad inyectada:</b>	2 µL
<b>Tiempo de análisis:</b>	45 min

### III.1. EVALUACIÓN DE LA REACCIÓN DE DERIVATIZACIÓN DE LOS FITOESTEROLES PARA SU DETERMINACIÓN MEDIANTE GC

En el método oficial del Consejo Oleícola Internacional [15], los fitoesteroles se derivatizan a trimetilsililéteres (TMS) con una mezcla de piridina: hexametildisilazano (HMDS): trimetilclorosilano (TMCS), en proporción 9:3:1 (V/V/V), a razón de 50 µL de mezcla derivatizante (M.D.) por mg de esterol, de acuerdo con el esquema recogido en la Figura III.1 en el que se ha tomado como ejemplo el β-sitosterol.



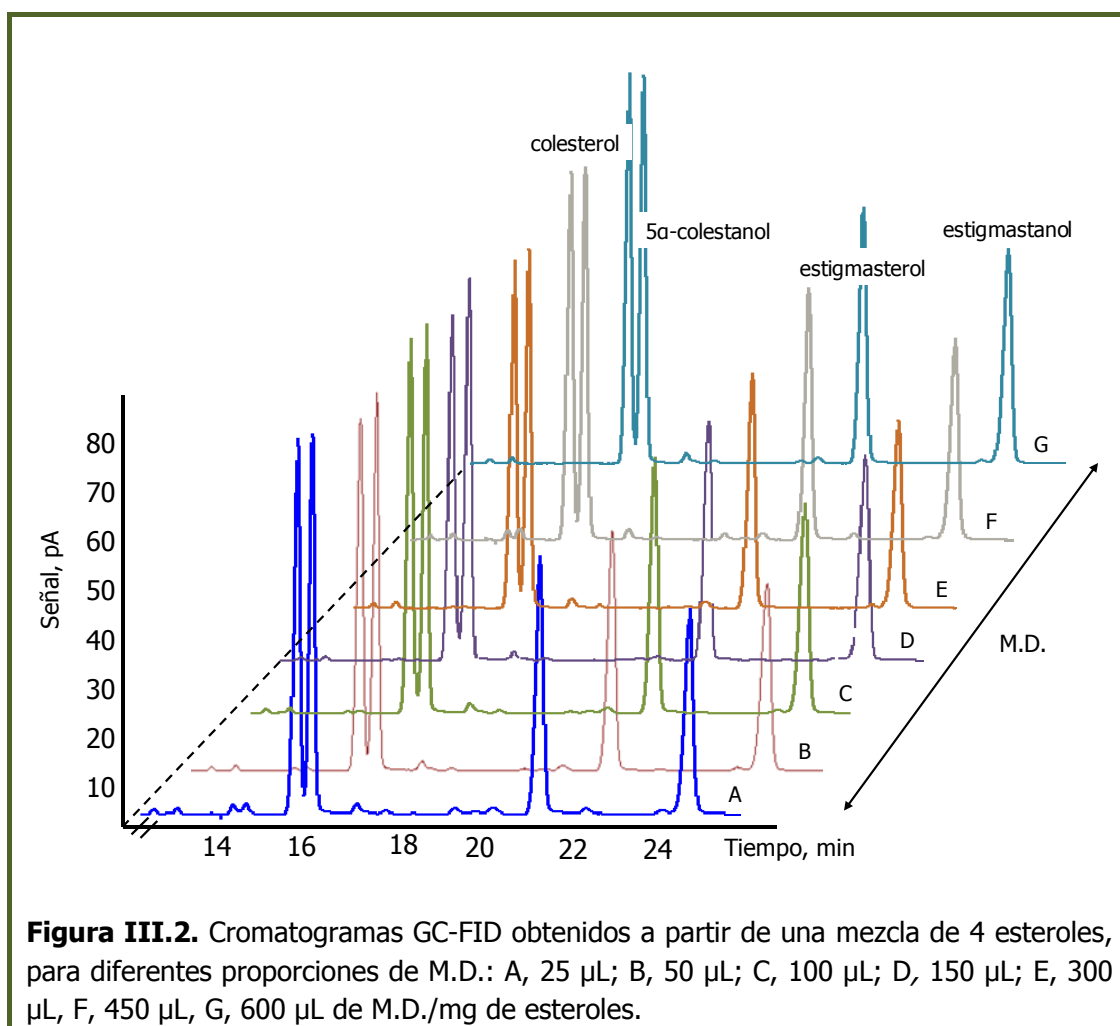
**Figura III.1.** Esquema de la preparación de los trimetilsililéteres de esteroides (TMS-Esteroides) tomando como ejemplo el  $\beta$ -sitosterol.

Para estudiar la influencia de la relación  $V_{M.D.}(\mu\text{l})/(\text{mg esteroides totales})$  sobre la resolución, el área y la altura de los picos cromatográficos, se seleccionaron cuatro esteroides (colesterol,  $\alpha$ -colestanol, estigmasterol y estigmastanol), entre los que se incluye el estándar interno (colestanol, I.S.), que aparecen en posiciones distintas del cromatograma: colesterol, que presenta un pico con tiempo de retención muy parecido al del I.S., estigmasterol, cuyo pico está situado en una posición central del cromatograma, y estigmastanol, que aparece casi al final del mismo.

Se preparó una mezcla de los mismos a partir de disoluciones de concentración 0.1% (m/m) en cloroformo, conteniendo la misma cantidad de cada uno de ellos para que la disolución de trabajo contuviera siempre 1 mg de esteroides totales. Alícuotas de la misma se derivatizaron con cantidades crecientes, entre 25 y 600  $\mu\text{L}$ , de M.D. y, en todos los casos, se llevó con piridina a un volumen final de 600  $\mu\text{L}$ , tal como se recoge en la Tabla III.1. Los cromatogramas obtenidos se muestran en la Figura III.2.

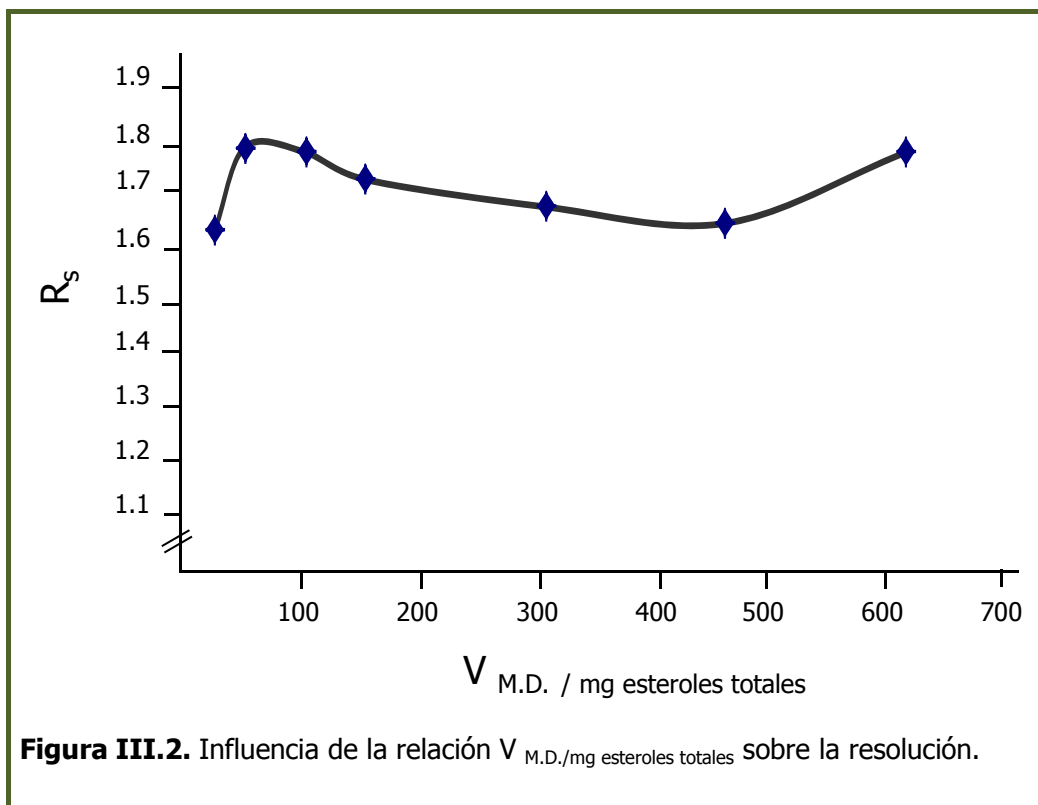
**Tabla III.1.** Volumen añadido de mezcla derivatizante (M.D.)

M.D. μL/mg esteroides totales	Piridina (μL)
25	575
50	550
100	500
150	450
300	300
450	150
600	0



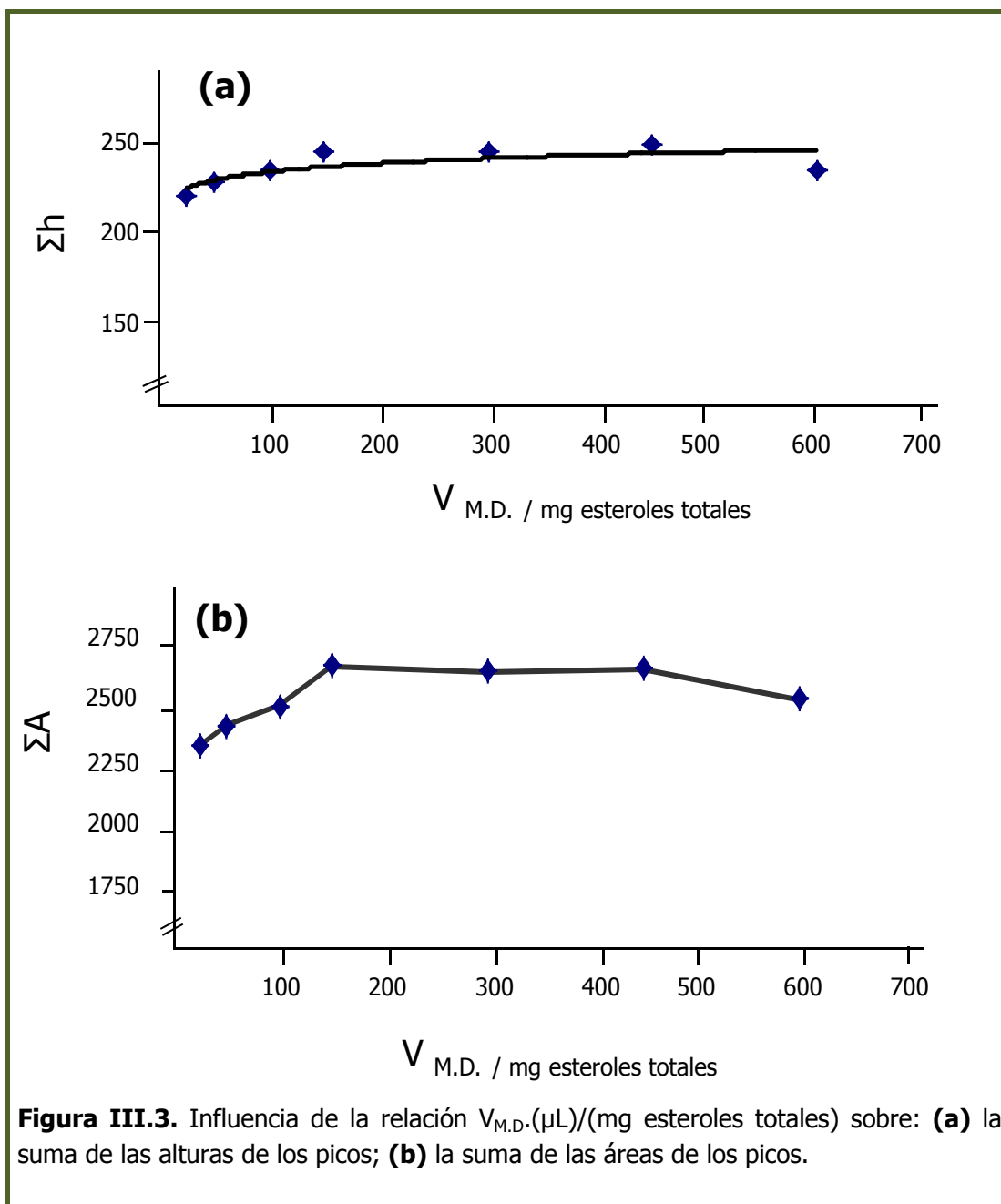
**Figura III.2.** Cromatogramas GC-FID obtenidos a partir de una mezcla de 4 esteroides, para diferentes proporciones de M.D.: A, 25 μL; B, 50 μL; C, 100 μL; D, 150 μL; E, 300 μL, F, 450 μL, G, 600 μL de M.D./mg de esteroides.

El cálculo de la resolución entre los picos que salen más próximos, colesterol y  $5\alpha$ -colestano- $3\beta$ -ol (colestanol), dio lugar a los resultados representados en la Figura III.2, donde se observa que ambos picos están suficientemente resueltos para una correcta cuantificación ( $R_s > 1.5$ ).



A continuación se estudió la influencia de la relación  $V_{M.D.}/\text{mg esteroides totales}$  sobre la suma de áreas y alturas de los picos cromatográficos. Los resultados se han representado en la Figura III.3.

Como se puede observar, se obtiene un valor constante en la suma de áreas y alturas en el intervalo de  $V_{M.D.}(\mu\text{L})/(\text{mg esteroides totales})$  comprendido entre 150 y 450  $\mu\text{L}/\text{mg}$ , por lo que para experiencias posteriores se decidió trabajar adicionando 200  $\mu\text{L}$  de M.D. por cada mg de esteroides totales contenido en la muestra.



### III.2. VERIFICACIÓN DE LA APTITUD DE LA ETAPA DE OBTENCIÓN DE LOS TRIMETILSILILDERIVADOS

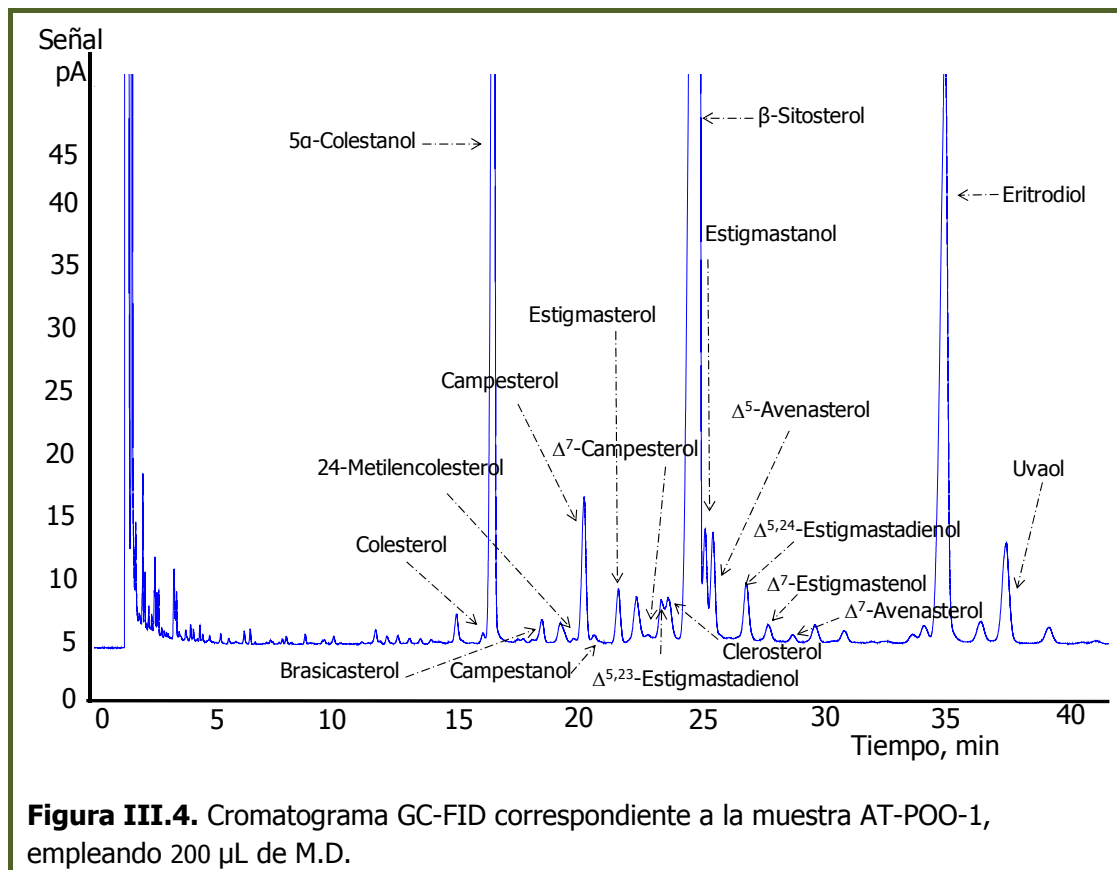
Con objeto de verificar si la cantidad de mezcla derivatizante optimizada se podía considerar "apta" para el propósito establecido inicialmente (en inglés "fitness for purpose"), se llevó a cabo la determinación del contenido de esteroides en tres muestras de aceites vegetales, oliva virgen extra (AT-EVOO-1), orujo de oliva (AT-



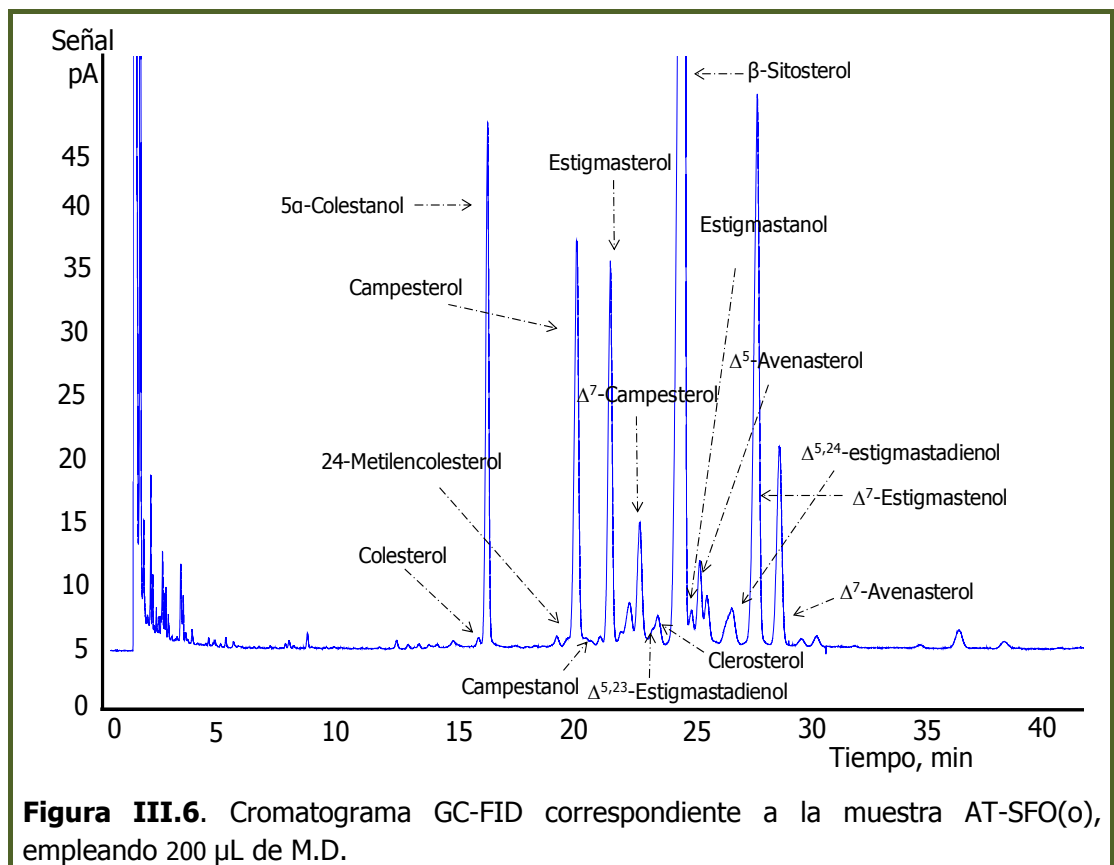
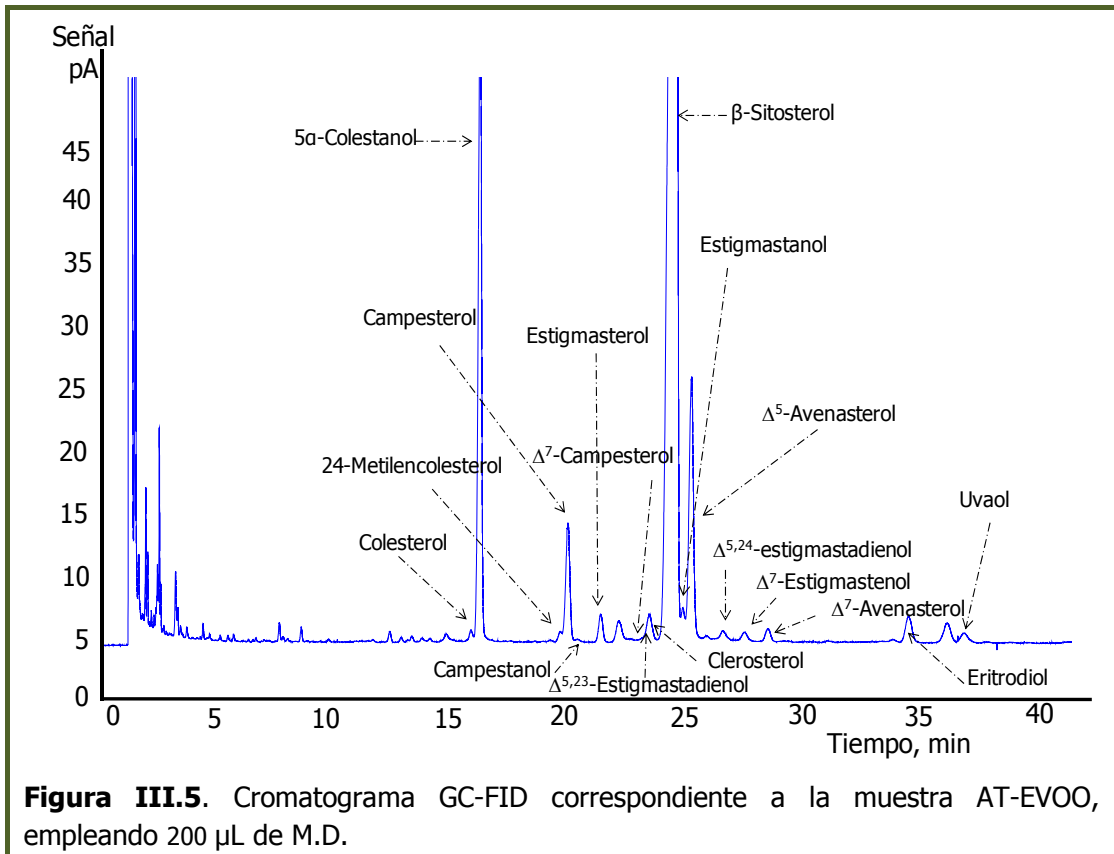
POO-1) y girasol con alto contenido en ácido oleico (AT-SFO(o)), proporcionados por el Laboratorio Agroalimentario de Atarfe (Granada).

El tratamiento de muestra se realizó según se especifica en el método oficial, separando la fracción que contiene los esteroides mediante cromatografía de capa fina. De cada muestra de aceite se prepararon paralelamente dos fracciones esterólicas, con objeto de comparar los resultados obtenidos según la cantidad de M.D. adicionada: 50  $\mu\text{L}$  M.D./( $\text{mg}$  esteroides totales), según el COI (datos suministrados por el Laboratorio de Atarfe); 200  $\mu\text{L}$  M.D./( $\text{mg}$  esteroides totales), según la modificación que se propone en esta memoria.

En las Figuras III.4 a III.6 se pueden ver los cromatogramas GC-FID obtenidos al adicionar 200  $\mu\text{L}$  de M.D.



Como se puede observar en dichos cromatogramas, el orden de elución es el siguiente: colesterol, colestanol (I.S.), brasicasterol, 24-metilencolesterol, campesterol, campestanol, estigmasterol,  $\Delta^7$ -estigmasterol,  $\Delta^{5,23}$ -estigmastadienol, clerosterol,  $\beta$ -sitosterol, estigmastanol/sitostanol,  $\Delta^5$ -avenasterol,  $\Delta^{5,24}$ -estigmas-tadienol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.



Los esteroides que se separan antes son aquellos que en la cadena lateral, en la posición del átomo de carbono 24, no tienen ningún grupo metilo, etilo, ni doble enlace. A continuación se separan los que presentan un doble enlace en dicha cadena (grupo etileno), seguidos de los que tienen un grupo metilo y, finalmente, los que tienen un grupo etilo junto con un doble enlace. Además, los que presentan un doble enlace en  $\Delta^5$  se separan antes que los que lo presentan en la posición  $\Delta^7$ . Dentro del grupo de los estanoles, se eluye primero el campestanol y a continuación el sitostanol (estigmastanol).

La Figura III.7 recoge las concentraciones de los analitos encontradas en las tres muestras de aceite.

También se puede observar que los esteroides mayoritarios de cada aceite son:

- *Aceite de Orujo de Oliva*:  $\beta$ -sitosterol, eritrodioleol, uvaol, campesterol, sitostanol,  $\Delta^5$ -avenasterol y clerosterol.
- *Aceite de Oliva Virgen Extra*:  $\beta$ -sitosterol,  $\Delta^5$ -avenasterol y campesterol.
- *Aceite de Girasol con alto contenido en Ácido Oleico*:  $\beta$ -sitosterol,  $\Delta^7$ -estigmastanol, campesterol, estigmasterol,  $\Delta^7$ -avenasterol,  $\Delta^7$ -campesterol y  $\Delta^5$ -avenasterol.

A partir de ahora, en las tablas de resultados, se utilizarán los códigos recogidos en la siguiente tabla (Tabla III.2) para referirnos a los analitos determinados.

**Tabla III.2.** Códigos asignados a los analitos

<b>Analito</b>	<b>Código</b>
colesterol	(1)
brassicasterol	(2)
24-metilen-colesterol	(3)
campesterol	(4)
campestanol	(5)
estigmasterol	(6)
$\Delta^7$ -campesterol	(7)
$\Delta^{5,23}$ -estigmastadienol	(8)
clerosterol	(9)
$\beta$ -sitosterol	(10)
Sitostanol o estigmastanol	(11)
$\Delta^5$ -avenasterol	(12)
$\Delta^{5,24}$ -estigmastadienol	(13)
$\Delta^7$ -estigmastanol	(14)
$\Delta^7$ -avenasterol	(15)

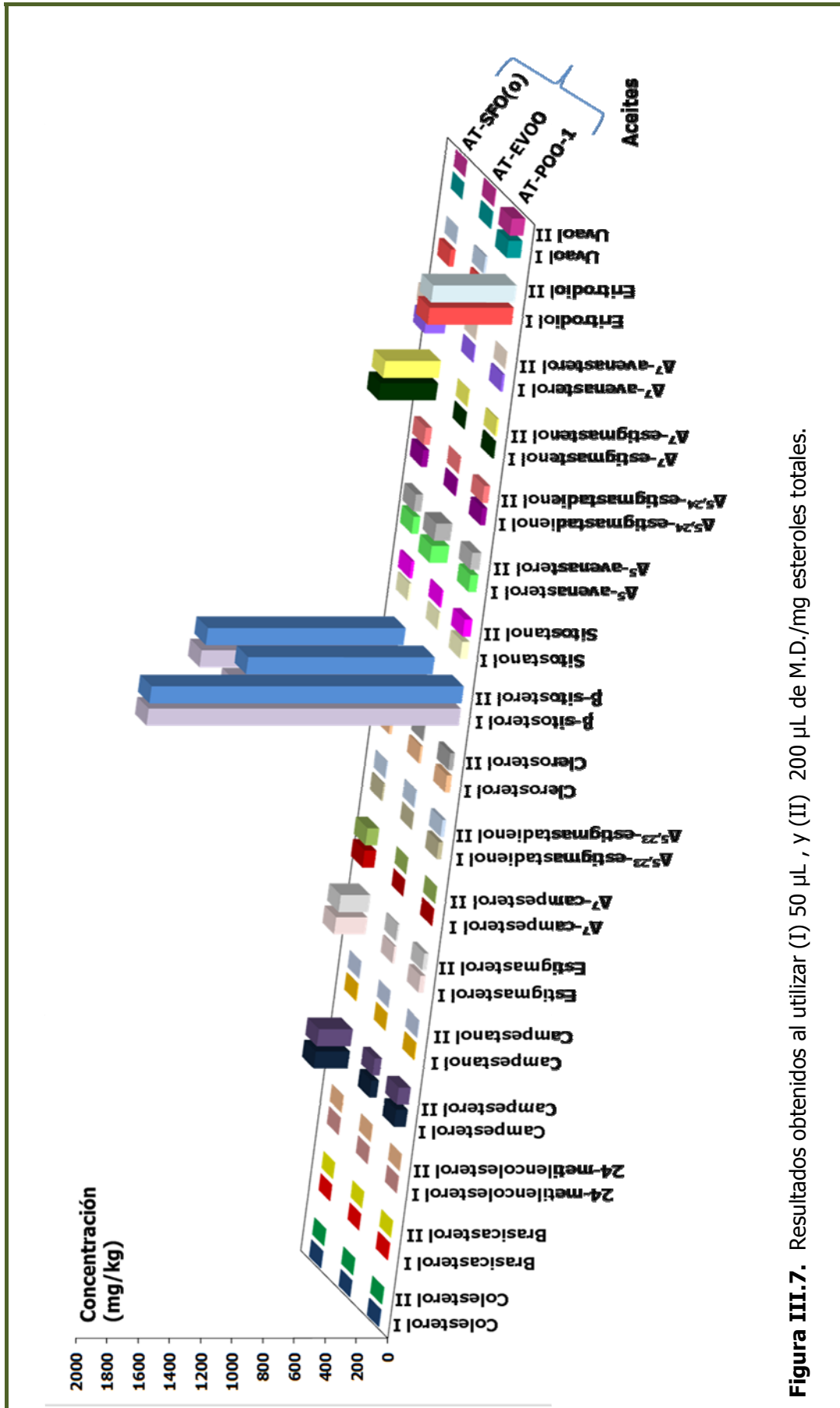


Figura III.7. Resultados obtenidos al utilizar (I) 50 µL , y (II) 200 µL de M.D./mg esteroides totales.

Con objeto de comprobar si había alguna diferencia estadística significativa (validación del resultado analítico) en las concentraciones de analitos encontradas al variar el volumen de mezcla derivatizante adicionado, se procedió a la aplicación del test de la t de Student para la comparación de dos medias. En este caso, se plantearon las siguientes hipótesis:

- Hipótesis nula ( $H_0$ ). No existe diferencia significativa entre las concentraciones encontradas al añadir 50 o 200  $\mu$ L de M.D.
- Hipótesis alternativa ( $H_1$ ). Hay diferencia significativa entre las concentraciones encontradas.

Dado que los resultados suministrados por el Laboratorio de Atarfe se obtuvieron a partir de dos réplicas, los valores de "s" eran, en algunos casos, mayores que los obtenidos en nuestro laboratorio. Aunque, a pesar de eso, se obtenía homogeneidad de varianzas, se consideró que no la había, por lo que el test se aplicó con las desviaciones estándar obtenidas en cada caso. Los resultados para cada uno de los aceites analizados se recogen en las Tablas III.3 a III.5.

**Tabla III.3.** Resultados obtenidos en la verificación de la muestra AT-POO-1.

Analito	C <sub>50</sub> (mg/kg)	s (n =2)	C <sub>200</sub> (mg/kg)	s (n = 3)	t <sub>cal</sub>	t <sub>tab</sub> (g.l.)	¿Existen diferencias?
(1)	6.51	0.44	3.72	0.24	8.177	12.71(1)	NO
(2)	5.60	0.30	1.6	0.19	16.643	4.30(2)	SI
(3)	2.99	0.30	1.72	0.13	5.586	12.71(1)	NO
(4)	75.84	3.47	77.18	0.80	0.537	12.71(1)	NO
(5)	4.66	1.07	4.17	0.75	0.569	4.30(2)	NO
(6)	32.83	2.38	27.15	0.19	3.373	12.71(1)	NO
(7)	5.05	0.70	3.66	2.23	1.007	3.18(3)	NO
(8)	28.38	0.70	23.27	2.37	3.516	4.30(2)	NO
(9)	32.97	0.96	29.46	2.35	2.309	3.18(3)	NO
(10)	2053.75	74.04	2055.78	13.14	0.038	12.71(1)	NO
(11)	47.40	2.61	50.56	0.44	1.693	12.71(1)	NO
(12)	50.78	1.92	57.86	3.28	3.034	3.18(3)	NO
(13)	31.91	4.77	37.61	0.43	1.687	12.71(1)	NO
(14)	11.74	1.15	11.85	0.26	0.126	12.71(1)	NO
(15)	15.33	2.32	4.59	0.10	6.559	12.71(1)	NO

Los analitos se han codificado según Tabla III.2.

**Tabla III.4.** Resultados obtenidos en la verificación de la muestra AT-EVOO-1.

Analito	C <sub>50</sub> (mg/kg)	s (n =2)	C <sub>200</sub> (mg/kg)	s (n = 3)	t <sub>cal</sub>	t <sub>tab</sub> (g.l.)	¿Existen diferencias?
(1)	1.56	0.30	3.38	0.07	8.41	12.71(1)	NO
(2)	N.D.	===	N.D.	===	===	===	===
(3)	3.35	0.01	3.06	0.14	3.62	4.30(2)	NO
(4)	47.50	0.68	45.41	0.47	3.79	4.30(2)	NO
(5)	1.11	0.45	1.07	0.20	0.12	12.71(1)	NO
(6)	13.78	0.48	10.52	0.08	9.52	12.71(1)	NO
(7)	0.81	0.20	0.53	0.28	1.28	3.18(3)	NO
(8)	1.90	0.67	6.43	0.21	9.26	12.71(1)	NO
(9)	15.05	1.45	6.46	0.48	8.09	12.71(1)	NO
(10)	1270.55	21.34	1195.19	12.34	4.52	12.71(1)	NO
(11)	7.80	0.33	10.05	0.20	8.71	12.71(1)	NO
(12)	113.14	1.39	103.27	0.21	9.96	12.71(1)	NO
(13)	7.31	0.79	4.34	0.24	5.18	12.71(1)	NO
(14)	5.24	0.33	4.32	0.18	3.58	12.71(1)	NO
(15)	7.03	0.19	5.97	0.08	7.50	12.71(1)	NO

Los analitos se han codificado según Tabla III.2.

**Tabla III.5.** Resultados obtenidos en la verificación de la muestra AT-SFO(o).

Analito	C <sub>50</sub> (mg/kg)	s (n =2)	C <sub>200</sub> (mg/kg)	s (n = 3)	t <sub>cal</sub>	t <sub>tab</sub> (g.l.)	¿Existen diferencias?
(1)	4.88	0.44	3.42	0.13	4.626	12.71(1)	NO
(2)	N.D.	==	N.D.	==	==	==	==
(3)	3.54	0.44	3.72	0.07	0.572	12.71(1)	NO
(4)	218.69	1.76	216.63	5.27	0.627	3.18(3)	NO
(5)	5.21	2.79	2.56	1.38	1.242	12.71(1)	NO
(6)	201.32	3.25	190.24	3.53	3.607	4.30(2)	NO
(7)	75.29	4.17	75.38	1.04	0.029	12.71(1)	NO
(8)	10.05	0.22	5.79	3.58	2.054	4.30(2)	NO
(9)	19.08	1.23	22.16	6.54	0.794	4.30(2)	NO
(10)	1288.88	24.04	1267.18	18.40	1.083	4.30(2)	NO
(11)	13.79	3.76	18.7	2.14	1.676	12,71(1)	NO
(12)	43.22	2.98	46.89	1.68	1.579	12,71(1)	NO
(13)	37.51	2.40	40.69	2.99	1.314	3.18(3)	NO
(14)	368.15	7.18	356.14	7.12	1.837	4.30(2)	NO
(15)	129.87	2.74	126.63	3.31	1.188	3.18(3)	NO

Los analitos se han codificado según Tabla III.2.

Como se puede observar, no hay diferencia significativa entre los resultados obtenidos, salvo para el brasicasterol en el aceite AT-POO-1, por lo que a partir de

entonces, se consideró que 200 µL de M.D. por mg de esteroides totales era un volumen apto para llevar a cabo la derivatización ya que, con este volumen, se mejoran tanto la señal analítica como la repetibilidad de los resultados.

A continuación se procedió a evaluar si la corrección teórica de los factores de respuesta (TCFs) de cada uno de los analitos **[16]** producía alguna diferencia estadísticamente significativa con respecto a la forma de cuantificación seguida por el COI, en la que se asume que todos los analitos presentan la misma respuesta frente al sistema de detección. Dado que los datos obtenidos en cada caso proceden de los mismos cromatogramas, se aplicó el test de la t de Student para la comparación de dos medias, utilizando la desviación estándar promedio ( $s_p$ ) y comparando la  $t_{calculada}$  con el valor del estadístico t que, para cuatro grados de libertad y una probabilidad del 95 % ( $p = 0.95$ ), vale 2.78. Los resultados obtenidos se recogen en las Tablas III.6 a III.8.

**Tabla III.6** Resultados obtenidos al comparar la concentración calculada mediante el método oficial del COI con la obtenida utilizando los factores de corrección teóricos (TCFs). Muestra AT-POO-1.

Analito	C <sub>(COI)</sub> (mg/kg)	S <sub>(COI)</sub> (n = 2)	C <sub>(TCF)</sub> (mg/kg)	S <sub>(TCF)</sub> (n = 3)	S <sub>p</sub>	t <sub>cal</sub>	¿Existen diferencias?
(1)	3.72	0.24	3.70	0.24	0.240	0.10	NO
(2)	1.6	0.19	1.58	0.19	0.194	0.13	NO
(3)	1.72	0.13	1.71	0.13	0.131	0.09	NO
(4)	77.18	0.80	76.64	0.80	0.801	0.83	NO
(5)	4.17	0.75	4.16	0.75	0.751	0.02	NO
(6)	27.15	0.19	26.78	0.18	0.186	2.44	NO
(7)	3.66	2.23	3.63	2.33	2.284	0.02	NO
(8)	23.27	2.37	22.96	2.33	2.349	0.16	NO
(9)	29.46	2.35	29.06	2.32	2.338	0.21	NO
(10)	2055.78	13.14	2036.25	13.01	13.076	1.83	NO
(11)	50.56	0.44	50.18	0.43	0.438	1.06	NO
(12)	57.86	3.28	57.07	3.24	3.261	0.30	NO
(13)	37.61	0.43	37.10	0.42	0.429	1.46	NO
(14)	11.85	0.26	11.63	0.26	0.262	1.03	NO
(15)	4.59	0.10	4.52	0.09	0.096	0.89	NO

Los analitos se han codificado según Tabla III.2.

**Tabla III.7.** Resultados obtenidos al comparar la concentración calculada mediante el método oficial del COI con la obtenida utilizando los factores de corrección teóricos (TCFs). Muestra AT-EVOO-1.

Analito	C <sub>(COI)</sub> (mg/kg)	S <sub>(COI)</sub> (n =2)	C <sub>(TCF)</sub> (mg/kg)	S <sub>(TCF)</sub> (n = 3)	S <sub>p</sub>	t <sub>cal</sub>	¿Existen diferencias?
(1)	3.38	0.07	3.37	0.07	0.071	0.17	NO
(2)	n.d.	==	n.d.	==	==	==	==
(3)	3.06	0.14	3.03	0.14	0.139	0.27	NO
(4)	45.41	0.47	45.09	0.47	0.474	0.83	NO
(5)	1.07	0.20	1.07	0.20	0.198	0.00	NO
(6)	10.52	0.08	10.38	0.08	0.076	2.25	NO
(7)	0.53	0.28	0.53	0.29	0.286	0.00	NO
(8)	6.43	0.21	6.35	0.21	0.210	0.47	NO
(9)	6.46	0.48	6.37	0.47	0.475	0.23	NO
(10)	1195.19	12.34	1183.84	12.22	12.284	1.13	NO
(11)	10.05	0.20	9.98	0.19	0.195	0.44	NO
(12)	103.27	0.21	102.87	0.21	0.209	2.35	NO
(13)	4.34	0.24	4.28	0.24	0.242	0.30	NO
(14)	4.32	0.18	4.240	0.17	0.176	0.56	NO
(15)	5.97	0.08	5.890	0.08	0.078	1.25	NO

Los analitos se han codificado según Tabla III.2.

**Tabla III.8.** Resultados obtenidos al comparar la concentración calculada mediante el método oficial del COI con la obtenida utilizando los factores de corrección teóricos (TCFs). Muestra AT-SFO(o).

Analito	C <sub>(COI)</sub> (mg/kg)	S <sub>(COI)</sub> (n =2)	C <sub>(TCF)</sub> (mg/kg)	S <sub>(TCF)</sub> (n = 3)	S <sub>p</sub>	t <sub>cal</sub>	¿Existen diferencias?
(1)	3.42	0.13	3.43	0.13	0.131	0.09	NO
(2)	n.d.	==	n.d.	==	==	==	==
(3)	3.72	0.07	3.70	0.05	0.063	0.39	NO
(4)	216.63	5.27	216.43	5.22	5.244	0.05	NO
(5)	2.56	1.38	2.57	1.38	1.379	0.01	NO
(6)	190.24	3.53	188.82	3.47	3.500	0.50	NO
(7)	75.38	1.04	75.32	1.08	1.056	0.07	NO
(8)	5.79	3.58	5.74	3.52	3.551	0.02	NO
(9)	22.16	6.54	21.97	6.45	6.495	0.04	NO
(10)	1267.18	18.40	1262.93	18.13	18.265	0.28	NO
(11)	18.7	2.14	18.69	2.12	2.131	0.01	NO
(12)	46.89	1.68	46.54	1.66	1.670	0.26	NO
(13)	40.69	2.99	40.39	2.94	2.967	0.12	NO
(14)	356.14	7.12	351.96	7.01	7.067	0.72	NO
(15)	126.63	3.31	125.67	3.26	3.288	0.36	NO

Los analitos se han codificado según Tabla III.2.



A la vista de los resultados, se puede concluir que los valores de concentración encontrados usando el método oficial son, en general, ligeramente más altos que los encontrados usando los TCFs. Sin embargo, considerando las concentraciones individuales de cada analito, la aplicación del tratamiento estadístico indica que las concentraciones encontradas en ambos casos no difieren significativamente. Dado que la utilización de los TCFs no producía ninguna mejora en el cálculo de las concentraciones, a partir de entonces, la cuantificación se llevó a cabo siguiendo el método oficial.

### III.3. SEPARACIÓN, MEDIANTE CROMATOGRAFÍA LÍQUIDA, DE LA FRACCIÓN DE FITOESTEROLES A PARTIR DEL EXTRACTO INSAPONIFICABLE. OPTIMIZACIÓN DE LA VENTANA DE TIEMPO.

Con la finalidad de establecer las condiciones cromatográficas más adecuadas para la obtención de la fracción esterólica de un aceite vegetal, se realizaron diversas experiencias usando las dos modalidades de HPLC más habituales.

Partiendo de una mezcla sintética de esteroides formada por colesterol, estigmasterol,  $\beta$ -sitosterol, campesterol y sitostanol, disueltos en cloroformo, se probó primero la modalidad de fase invertida, empleando una columna LICHROSPHER<sup>®</sup> 100 RP-8, en régimen isocrático y detección espectrofotométrica a la longitud de onda de 208 nm. Se trabajó con un flujo de fase móvil de 0.8 mL/min y se inyectó un volumen de muestra de 20  $\mu$ L. Las fases móviles que se ensayaron se recogen en la Tabla III.9.

**Tabla III.9.** Fases móviles ensayadas (Fase invertida)

Composición	%(v/v)
Acetonitrilo:Agua .....	86:14
Acetonitrilo: Agua (0.01% ácido acético) .....	70:30
Acetonitrilo .....	100
Acetonitrilo: Metanol.....	70:30
Acetonitrilo: Metanol: Agua (0.01% ácido acético) .....	48.5:48.5:3
Metanol: Acetonitrilo .....	80:20

Puesto que ninguna de las experiencias llevadas a cabo en esta modalidad cromatográfica dio resultados satisfactorios, ya que no se conseguía una separación clara de la familia de esteroides, se optó por cambiar a la modalidad de fase normal.

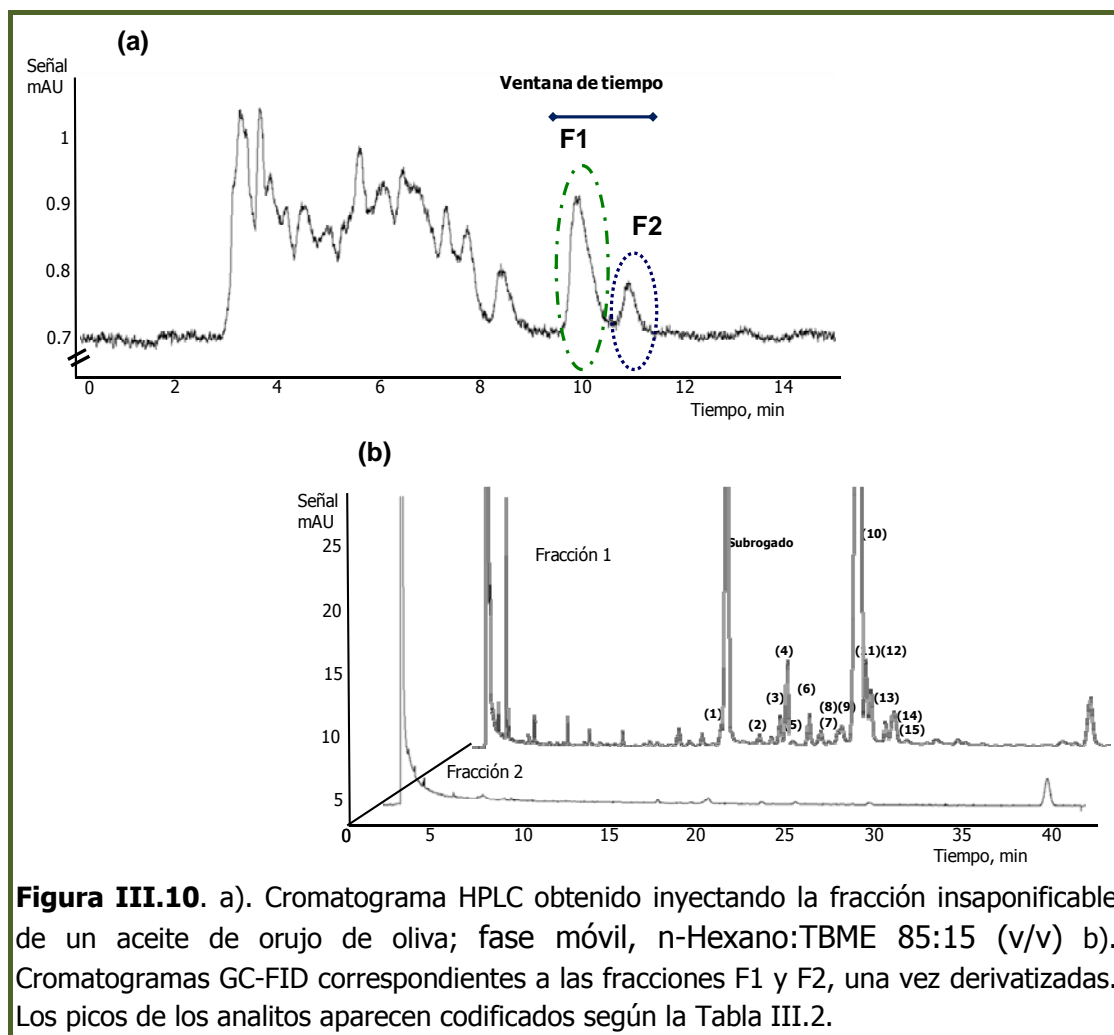
Aunque la fase normal más común es la que emplea sílice como fase estacionaria, con el fin de evitar los problemas inherentes a la misma (baja reproducibilidad, retenciones irreversibles, etc. [17]), se seleccionó una fase de polaridad intermedia como es la ciano, concretamente una columna LICHROSPHER® 100 CN (250 x 4 mm i.d., 5 µm). En este caso, se trabajó con disoluciones de los cinco esteroides antes citados, disueltos en la fase móvil que iba a ser utilizada en la separación cromatográfica. Los cromatogramas se registraron en las mismas condiciones utilizadas en el caso anterior. La Tabla III.10 recoge las fases móviles ensayadas.

**Tabla III.10.** Fases móviles ensayadas (Fase normal)

Composición	%(v/v)
Acetonitrilo .....	100
Acetonitrilo: n-Hexano .....	50:50
n-Hexano: Metanol .....	90:10
n-Hexano: TBME .....	75:25
n-Hexano: TBME .....	80:20
n-Hexano: TBME .....	85:15

Se observó que cuando se utilizaba la mezcla n-Hexano: TBME, se obtenía un pico cromatográfico asignable a los analitos. Cuando la composición de la fase móvil era n-Hexano:TBME 85:15 (v/v), dicho pico aparecía entre los minutos 9.5 a 11.5.

Para establecer el intervalo de tiempo (**ventana de tiempo**) en el que los analitos debían ser recogidos a la salida de la columna, el residuo insaponificable obtenido a partir de una muestra de aceite de orujo de oliva se disolvió en 1 mL de la fase móvil antes citada y se inyectó en el cromatógrafo de líquidos. En la Figura III.10 (a) se puede ver que, en el citado intervalo de tiempo, aparecen dos picos (F1 y F2). En una nueva inyección de 20 µL del insaponificable, se recogieron las fracciones correspondientes a estos picos, se derivatizaron y se inyectaron en el GC-FID. Como puede observarse en la figura III.10. (b), los fitoesteroides coeluyen en la fracción F1, una vez que han salido los compuestos menos polares, mientras que la fracción F2 da lugar a un cromatograma sin picos de esteroides.



**Figura III.10.** a). Cromatograma HPLC obtenido inyectando la fracción insaponificable de un aceite de orujo de oliva; fase móvil, n-Hexano:TBME 85:15 (v/v) b). Cromatogramas GC-FID correspondientes a las fracciones F1 y F2, una vez derivatizadas. Los picos de los analitos aparecen codificados según la Tabla III.2.

Para optimizar las condiciones de la separación de la familia de esteroides mediante cromatografía líquida (obtener un cromatograma en el menor tiempo posible sin pérdida de resolución), se realizaron diversas pruebas con la mezcla n-Hexano: TBME, en las dos modalidades, elución en gradiente e isocrática. En el primer caso se ensayaron los perfiles de gradiente recogidos en la Tabla III.11. Aunque se observó que la fracción de esteroides eluía a un tiempo más corto, no se seleccionó esta modalidad debido a que se necesitaban 20 minutos para la estabilización del sistema, lo cual alargaba excesivamente el tiempo de análisis.

En la modalidad isocrática, se probaron distintas proporciones de la mezcla n-Hexano:TBME: 75:25; 80:20; 85:15 (v/v). A partir de los cromatogramas obtenidos (Figura III.11) se comprobó que se producía un desplazamiento de la ventana de tiempo en función de la composición de la fase móvil, considerándose que una fase móvil de composición n-Hexano:TBME 80:20 (v/v) era adecuada para la separación de la fracción esteróica, puesto que daba lugar, en un tiempo aceptable, a un cromatograma con buena resolución. La ventana de tiempo en este caso estaba comprendida entre los minutos 7.5 y 9.5.

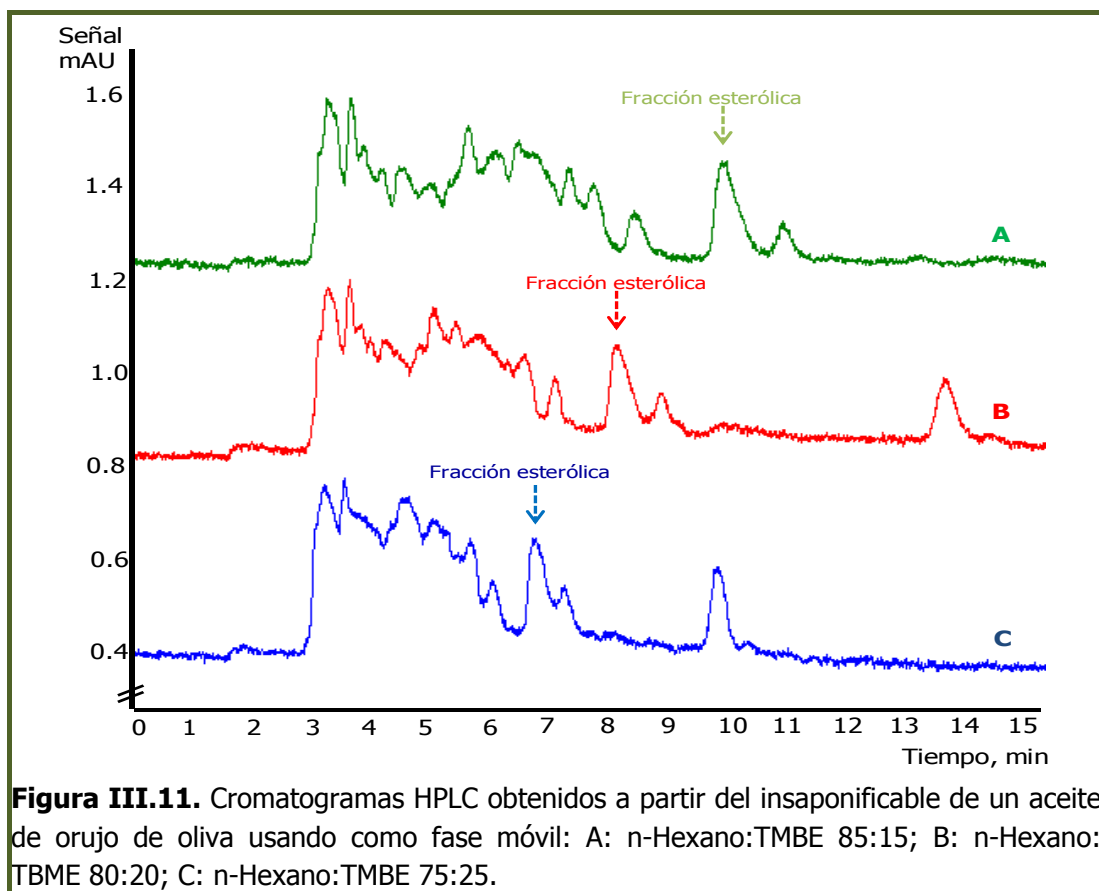
**Tabla III.11.** Perfil de gradientes n-hexano: TBME

<u>Gradiente 1</u>		<u>Gradiente 2</u>		<u>Gradiente 3</u>	
Tiempo (min)	% B(*)	Tiempo (min)	% B	Tiempo (min)	% B
0	15	0	15	0	15
8	15	8	15	6	15
15	50	11	50	9	50
17	15	13	15	12	50
20	15			15	15

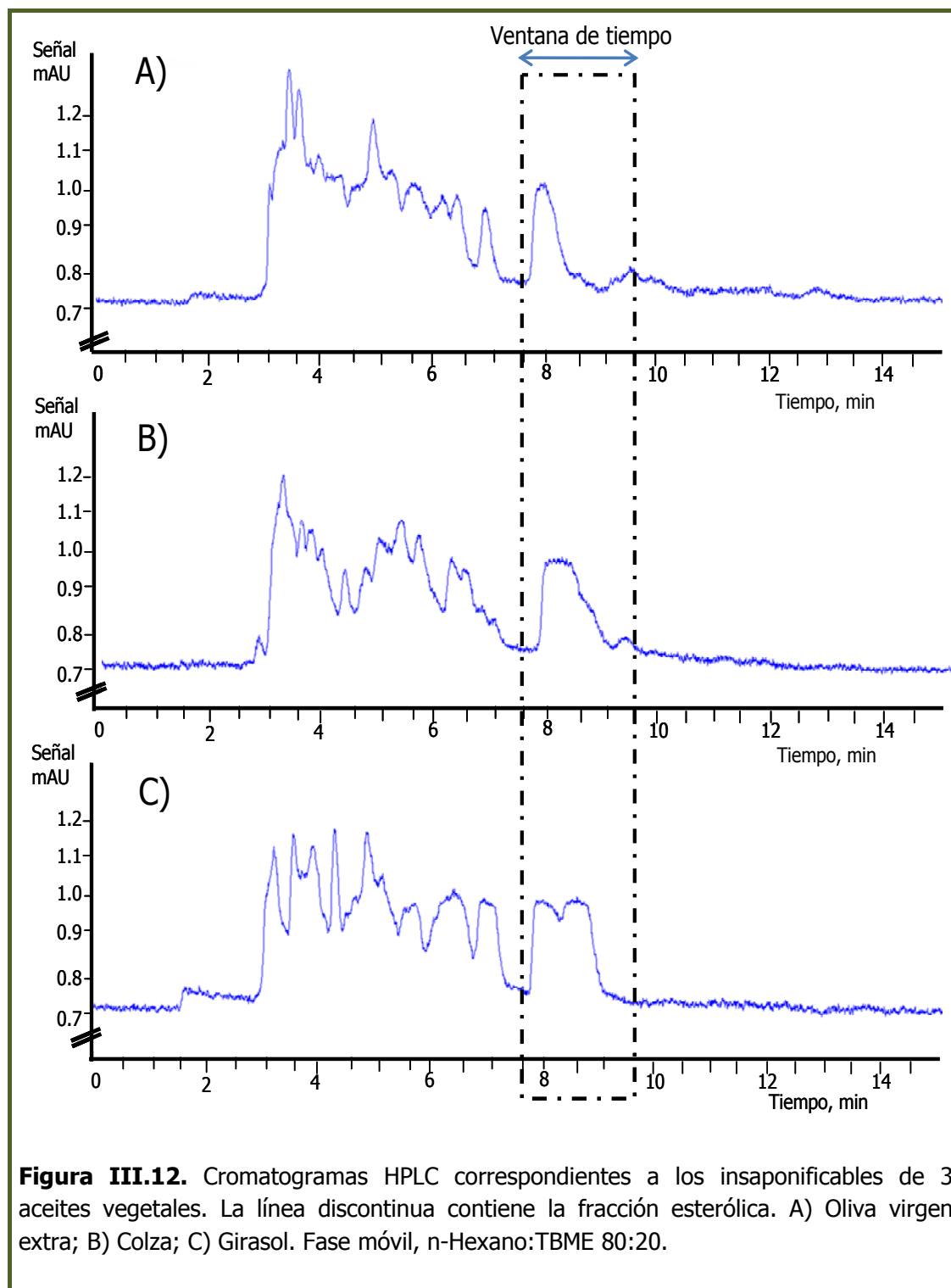
<u>Gradiente 4</u>		<u>Gradiente 5</u>	
Tiempo (min)	% B*	Tiempo (min)	% B*
0	20	0	25
6	20	6	25
9	50	9	50
12	50	12	50
15	20	15	25

(\*) tert-butilmetileter (TBME)



**Figura III.11.** Cromatogramas HPLC obtenidos a partir del insaponificable de un aceite de orujo de oliva usando como fase móvil: A: n-Hexano:TMBE 85:15; B: n-Hexano:TBME 80:20; C: n-Hexano:TMBE 75:25.

Una vez optimizadas las condiciones cromatográficas, se procedió al estudio de la fracción insaponificable de diferentes tipos de aceites vegetales. La Figura III.12 muestra los cromatogramas obtenidos para tres de estos aceites (oliva virgen extra, colza y girasol).



**Figura III.12.** Cromatogramas HPLC correspondientes a los insaponificables de 3 aceites vegetales. La línea discontinua contiene la fracción esterólica. A) Oliva virgen extra; B) Colza; C) Girasol. Fase móvil, n-Hexano:TBME 80:20.

Como se puede observar, el perfil del pico de esteroides cambia en función del tipo de aceite, pero la ventana de tiempo se mantiene prácticamente constante en todos los casos. Para posteriores análisis, y con objeto de asegurar la recogida de todos los esteroides, se seleccionó como ventana de tiempo adecuada la comprendida entre los minutos 7.5 y 9.5, aproximadamente.

Finalmente, cabe indicar que de acuerdo con la bibliografía [5], el cambio de perfil observado al variar el tipo de aceite puede atribuirse a que cambia la concentración relativa de los diferentes 4-desmetilesteroides.

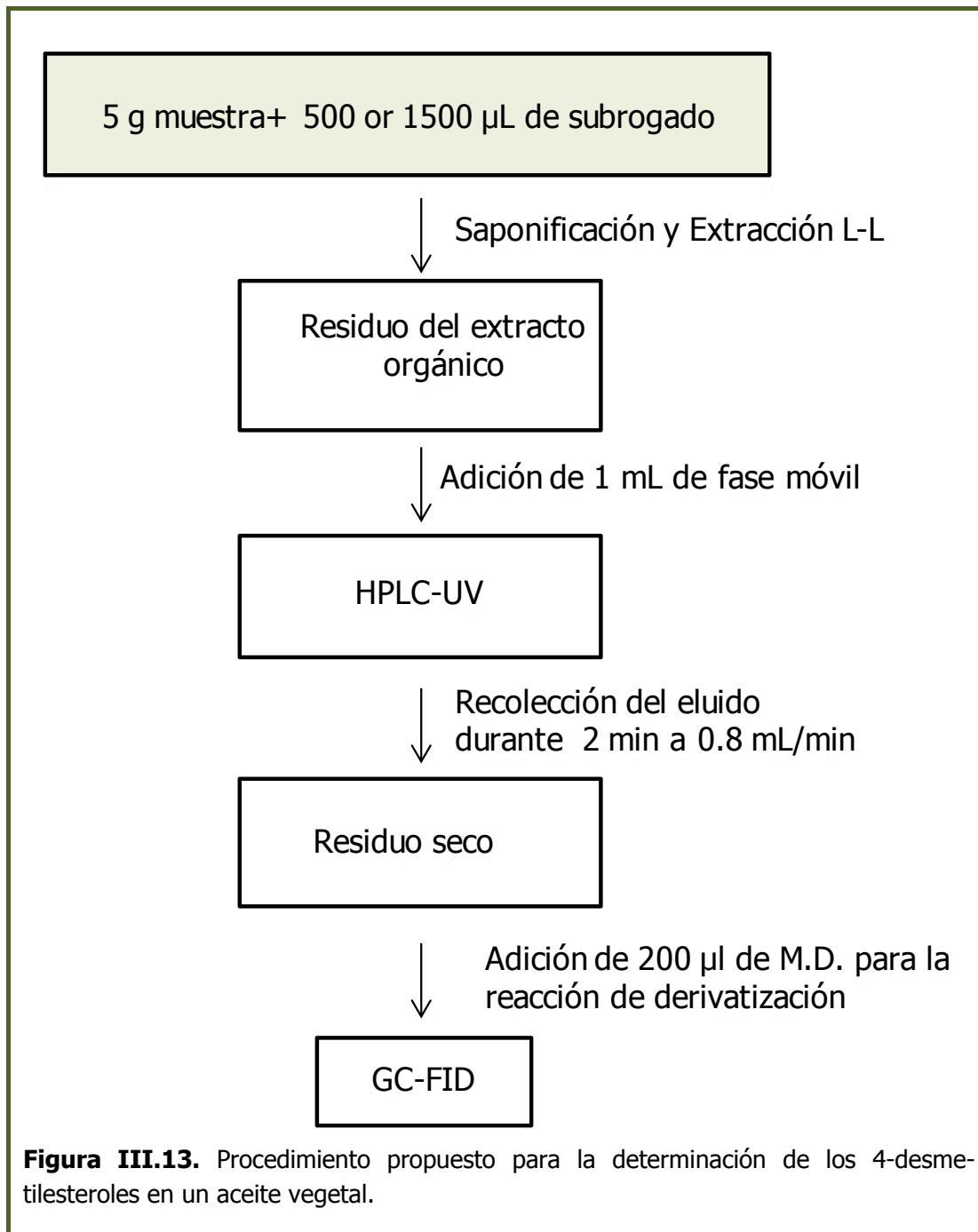
Por tanto, una vez optimizadas las variables cromatográficas para la obtención de la fracción de fitoesteroides, el **procedimiento operatorio** propuesto es el siguiente:

Partiendo de la disolución que contiene la fracción insaponificable de la muestra de aceite, se evapora el disolvente, se disuelve el residuo en 1 mL de la fase móvil y se inyecta en el cromatógrafo de líquidos, empleando los siguientes parámetros cromatográficos:

- **Fase estacionaria:** Columna CIANO LICHROSPHER® 100 CN (250 x 4 mm i.d., 5 µm), equipada con una precolumna LiChroCART® (10 x 4.6 mm i.d.) empaquetada con la misma fase estacionaria.
- **Fase móvil:** n-Hexano:TBME 80:20 (v/v).
- **Flujo de fase móvil:** 0.8 mL/min.
- **Detección:** Espectrofotométrica a 208 nm.
- **Volumen de inyección:** 20 µL.

Entre los minutos 7.5 y 9.5 se recolecta la fracción eluida a la salida del detector, se elimina el disolvente mediante el paso de una ligera corriente de nitrógeno seguido de un calentamiento en estufa a 105 °C, durante 10 minutos. El residuo obtenido se enfría en un desecador y, a continuación, se procede a la derivatización de los fitoesteroides para su posterior análisis mediante GC-FID.

En el esquema recogido en la siguiente página (Figura III.13) se representa el procedimiento completo para la determinación de los 4-desmetilesteroides en un aceite vegetal.



#### III.4. VALIDACIÓN DEL MÉTODO PROPUESTO ATENDIENDO A SU APTITUD

De acuerdo con la norma ISO/IEC 17025:2005 la validación de un método analítico es la confirmación, mediante el examen y el aporte de evidencias objetivas, de que se cumplen los requisitos particulares para un uso específico

previsto (aptitud). En este caso, implicaría evaluar la determinación de 4-desmetil-esteroles en aceites vegetales mediante GC-FID, previa separación off-line mediante HPLC, una vez establecidos los requisitos. Dado que la forma de cuantificación del método conlleva una calibración interna a un solo nivel con un subrogado en la que no se pueden estimar los parámetros de calidad habituales de un método analítico, se procedió a la evaluación global mediante la estimación de:

- a) la incertidumbre de la determinación de cada analito.
- b) la veracidad, a través de la comparación de los resultados obtenidos al analizar con el método muestras de aceites suministradas por un laboratorio acreditado, que aunque no constituyen materiales de referencia certificados propiamente dichos, a escala metrológica, se pueden considerar valores de referencia adecuados.

#### **III.4.1. Evaluación de la incertidumbre asociada a la determinación de fitoesteroles**

De acuerdo con la última versión de la Guía para la Expresión de la Incertidumbre en la medida (GUM) [18] y con lo que algunos autores han señalado [19], para determinar la incertidumbre de los resultados analíticos usando una aproximación "bottom up", se deben de seguir las siguientes etapas:

- 1) Definir el procedimiento de medida y el mesurando, lo que en este caso se corresponde con, el procedimiento analítico completo y la concentración de 4-desmetilesteroles en los aceites vegetales.
- 2) Establecer un modelo matemático a partir del cual se pueda obtener el contenido de los analitos. En este método se utiliza una calibración interna a un nivel, usando colestanol como subrogado.
- 3) Asignar valores a todos los posibles parámetros que puedan afectar el resultado final del análisis y determinar la incertidumbre estándar de cada uno de ellos.
- 4) Aplicar los principios de la propagación de incertidumbre y, por último,
- 5) Expresar el resultado final como *resultado ± incertidumbre expandida* (factor K)

Bagur et al. [20] consideran que las principales fuentes de incertidumbre de un método analítico son:



- Incertidumbre operacional o de trabajo ( $u_{\text{operacional}}$ ), debida a factores tales como: efectos instrumentales, pureza de los reactivos, condiciones de medida y de manejo de la muestra, entre los más importantes.
- Incertidumbre de recuperación ( $u_{\text{recuperación}}$ ), la cual se deriva del error asociado con el método.
- Incertidumbre inherente ( $u_{\text{inherente}}$ ) que afecta directamente a los resultados analíticos y que proviene de factores no controlados por el operador. En términos generales, esta incertidumbre tiene dos componentes que se pueden relacionar con dos aspectos del proceso químico de medida:
  - i) Las etapas químicas que dependen de los parámetros químicos dan lugar a una incertidumbre que estos autores [21] definieron previamente, como "incertidumbre intrínseca ( $u_{\text{intrínseca}}$ )".
  - ii) El proceso de calibración química generado por la transformación de la señal analítica en concentración y, la aceptación de una distribución normal en la generación de la señal analítica, dan lugar a una componente de la incertidumbre denominada "incertidumbre de calibración química ( $u_{\text{cal química}}$ )".

En este caso, la incertidumbre de la calibración química se podría redefinir como "incertidumbre de cuantificación ( $u_{\text{cuantificación}}$ )", considerando que el uso de la calibración interna a un nivel con un subrogado, implica que la calibración está implícita en el proceso de cuantificación. En la estimación de esta fuente de incertidumbre es necesario considerar que las dos señales analíticas, el área de pico del analito y el área de pico del subrogado, están relacionadas tal como se pone de manifiesto en la siguiente ecuación:

$$C_{\text{esterol}(i)} = 1000 \times \frac{m_{\text{subrogado}}}{m_{\text{muestra}}} \times \frac{A_{\text{esterol}(i)}}{A_{\text{subrogado}}} \quad [\text{Ec.1}]$$

dónde:

-  $C_{\text{esterol}(i)}$  es la concentración del estero "i" en la muestral de aceite analizada, expresada en  $\text{mg} \cdot \text{kg}^{-1}$ .

-  $m_{\text{subrogado}}$  es la masa de subrogado expresada en mg.

- $m_{\text{muestra}}$  es la masa de aceite, expresada en g.
- $A_{\text{esterol}(i)}$  es el área de pico del estero*i*, expresada en unidades arbitrarias.
- $A_{\text{subrogado}}$  es el área de pico del subrogado, expresada en unidades arbitrarias.
- 1000 es el factor de conversión para expresar la concentración de los analitos en  $\text{mg}\cdot\text{kg}^{-1}$ .

Las diferentes contribuciones a la incertidumbre total asociada al método propuesto, se representan en el diagrama de Ishikawa (diagrama causa-efecto) recogido en la Figura III.14.

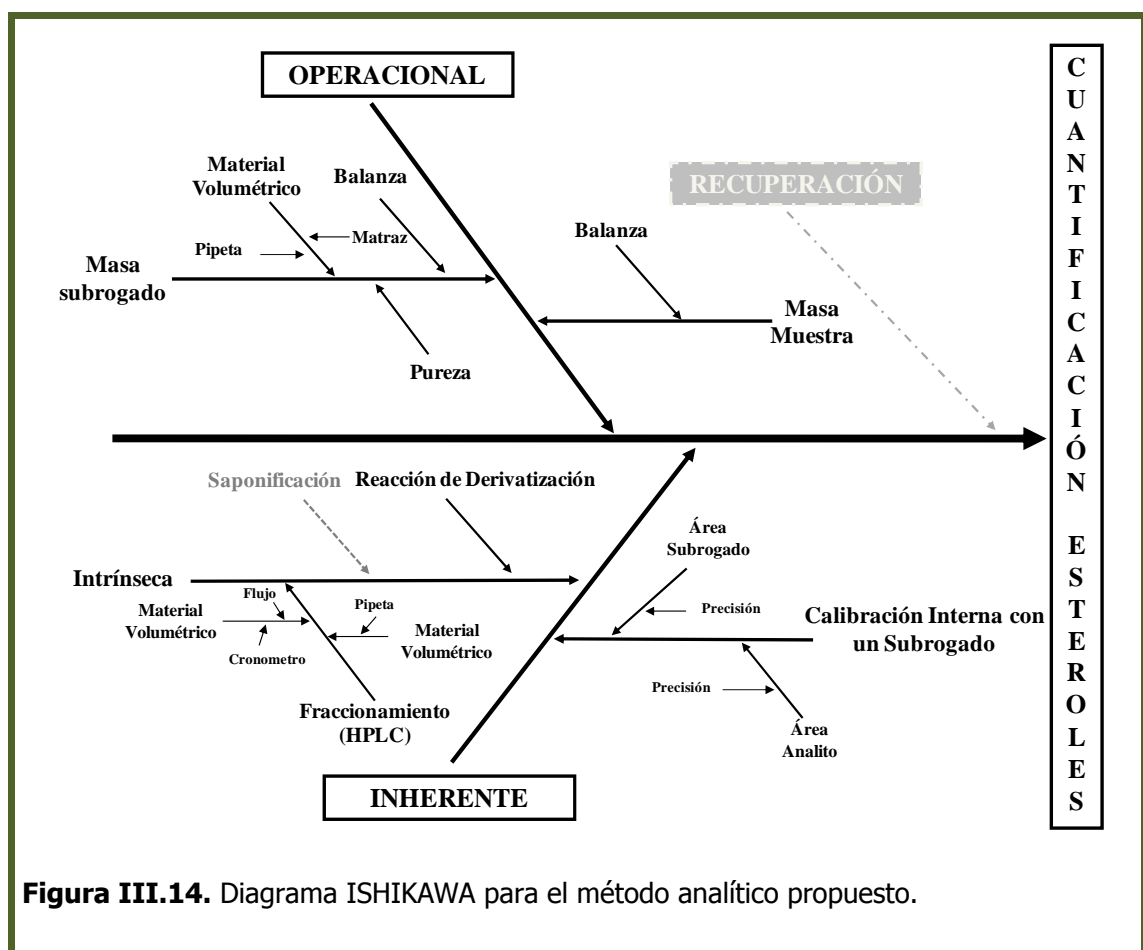


Figura III.14. Diagrama ISHIKAWA para el método analítico propuesto.

A continuación se describe la forma en la que se calcula cada una de las diferentes fuentes de incertidumbre:

■ Incertidumbre operacional o de trabajo ( $u^{\text{rel}}_{\text{op.}}$ )

La ecuación para esta fuente de incertidumbre viene dada por:

$$u_{op}^{rel} = \left( \left( u_{\text{masa-subrogado}}^{rel} \right)^2 + \left( u_{\text{masa-muestra}}^{rel} \right)^2 \right)^{1/2} \quad [\text{Ec. 2}]$$

(1) Incertidumbre asociada a la masa del subrogado ( $u_{\text{masa-subrogado}}^{rel}$ )

Como se puede ver en la Figura III.14, esta incertidumbre se ve afectada por las incertidumbres asociadas a la preparación de la disolución madre del subrogado y a la adición de un volumen de esta disolución a la muestra de aceite, antes de la etapa de saponificación. Expresada como incertidumbre relativa, viene dada por la siguiente expresión:

$$u_{\text{masa - subrogado}}^{rel} = \left( \left( u_{\text{dis. madre subrogado}}^{rel} \right)^2 + \left( u_{\text{pipeta}}^{rel} \right)^2 \right)^{1/2} \quad [\text{Ec. 3}]$$

El primer término, *incertidumbre asociada a la preparación de la disolución madre del subrogado* depende, entre otros factores, de la pureza del subrogado usado como patrón interno. Puesto que los fabricantes no suministran ningún dato a este respecto, se calcula teniendo en cuenta la última cifra significativa del valor de la pureza indicada por el fabricante. Además, asumiendo que la masa molar del colestanol no tiene una influencia significativa sobre la incertidumbre, la expresión que permite calcular la incertidumbre estándar relativa para la concentración de la disolución madre sería:

$$u_{\text{dis. madre subrogado}}^{rel} = \left( \left( u_{\text{patrón subrogado}}^{rel} \right)^2 + \left( u_{\text{pureza patrón}}^{rel} \right)^2 + \left( u_{\text{matraz}}^{rel} \right)^2 \right)^{1/2} \quad [\text{Ec. 4}]$$

es decir

$$u_{\text{dis. madre subrogado}}^{rel} = \left( \frac{u^2(m_{\text{patrón subrogado}})}{m_{\text{patrón subrogado}}^2} + \frac{u^2(\text{PUR})}{\text{PUR}^2} + \frac{u^2(V_{\text{matraz}})}{V_{\text{matraz}}^2} \right)^{1/2} \quad [\text{Ec. 5}]$$

Cada término expresado como incertidumbre relativa se ha calculado de acuerdo con las siguientes expresiones:

a) Incertidumbre asociada a la masa de subrogado pesada para preparar la disolución madre

$$u_{\text{patrón subrogado}}^{\text{rel}} = \left( \frac{\left( \frac{\text{Error máximo balanza}_{\text{proceso verific.}}}{\sqrt{3}} \right)^2}{m_{\text{patrón subrogado}}^2} \right)^{1/2} \quad [\text{Ec. 6}]$$

siendo,

- $\text{Error máximo}_{\text{proceso verific.}}$ , el error máximo de la balanza usada para pesar el subrogado. Se obtiene del plan de verificación del equipo y se calcula sumando, en valor absoluto, los valores de la corrección y de la incertidumbre expandida obtenida a partir del proceso de verificación.

$$(\text{Error máximo})_{\text{verificación}} = |\text{Corrección}| + \text{Incertidumbre expandida} \quad [\text{Ec. 7}]$$

- $M_{\text{patrón subrogado}}$  es la masa de subrogado necesaria para preparar 10 mL de una disolución madre al 0.2 % (m/v).

b) Incertidumbre asociada a la pureza del patrón de subrogado:

$$u_{\text{PUR}}^{\text{rel}} = \left( \frac{\left( \frac{\text{orden del último decimal PUREZA (especific. fabricante)}}{2\sqrt{3}} \right)^2}{\text{PUR}^2} \right)^{1/2} \quad [\text{Ec. 8}]$$

dónde PUR es la pureza del subrogado declarada por el fabricante.

c) Incertidumbre asociada al matraz utilizado para preparar la disolución madre de subrogado

$$u_{\text{matraz}}^{\text{rel}} = \left( \frac{u_{\text{standar}}^2}{V_{\text{matraz}}^2} \right)^{1/2} \quad [\text{Ec. 9}]$$

dónde:

- $u_{\text{estándar}}$  es la incertidumbre estándar asociada al matraz empleado para preparar 10 mL de la disolución de trabajo.

-  $V_{\text{matraz}}$  es 10 mL.

Teniendo en cuenta que el volumen de disolución madre añadido a la muestra de aceite depende del tipo de aceite analizado, el segundo término de la ecuación 3, es la *incertidumbre asociada al volumen de subrogado añadido a la muestra de aceite* y se ha calculado utilizando las siguientes ecuaciones:

a) Adición de 500  $\mu\text{L}$  de subrogado (aceites de oliva virgen extra y oliva refinado)

$$u_{\text{pipeta}(500)}^{\text{rel}} = \left( \frac{\left( \frac{\text{Error máximo}_{\text{proceso verific.}}}{\sqrt{3}} \right)^2}{V_{\text{pipeta}(500)}^2} \right)^{1/2} \quad [\text{Ec. 10}]$$

b) Adición de 1500  $\mu\text{L}$  de subrogado (aceite de orujo de oliva y resto de aceites vegetales)

$$u_{\text{pipeta}(500)}^{\text{rel}} = \left( \frac{u_{\text{estándar}}^2}{V_{\text{pipeta}(1500)}^2} \right)^{1/2} \quad [\text{Ec. 11}]$$

dónde:

-  $u_{\text{estándar}}$  es la incertidumbre estándar de la pipeta empleada para adicionar 1.5 mL de disolución a la muestra.

(2) Incertidumbre asociada a la masa de muestra ( $u_{\text{masa-muestra}}^{\text{rel}}$ )

La incertidumbre relativa asociada a la masa de muestra pesada, se calcula de la siguiente manera:

$$u_{\text{masa-muestra}}^{\text{rel}} = \left( \frac{\left( \frac{\text{Error máximo balanza}_{\text{proceso verific.}}}{\sqrt{3}} \right)^2}{m_{\text{masa-muestra}}^2} \right)^{1/2} \quad [\text{Ec. 12}]$$

En todos los casos, los valores de  $u^{\text{rel}}$  se calcularon considerando: i) las especificaciones del fabricante, ii) el plan de calidad interno seguido en el laboratorio (balanza analítica y las pipetas usadas para la adición del subrogado a

las muestras de aceite de oliva virgen extra y refinado), y iii) los datos recogidos en bibliografía, para el resto del material volumétrico (pipetas y matraces). Los valores utilizados se recogen en la Tabla III.12.

**Tabla III.12.** Valores de Incertidumbre para el equipo básico usado en el método.

<b>(a) Equipos verificados en el laboratorio de acuerdo con el plan de calidad interno</b>				
	<b><u>Precisión (g)</u></b>	<b><u>Peso (g)</u></b>	<b><u>Corrección (g)</u></b>	<b><u>U<sup>(1)</sup>(g)</u></b>
Balanzas	0.0001	1	0.0001	0.0003
	0.001	10	Not needed	0.002
	<b><u>Rango (µL)</u></b>	<b><u>Volumen (mL)</u></b>	<b><u>Corrección (mL)</u></b>	<b><u>U (mL)</u></b>
Pipetas mecánicas	100-1000	0.1	0.000	0.002
	100-1000	0.2	-0.001	0.002
	100-1000	0.3	-0.001	0.002
	100-1000	0.5	0.001	0.002
	100-1000	0.9	-0.005	0.002
	100-1000	1.0	-0.005	0.002
<b>(b) Equipos utilizados directamente del fabricante<sup>(2)</sup></b>				
	<b><u>Volumen (mL)</u></b>	<b><u>u<sup>(3)</sup> para disolventes orgánicos (mL)</u></b>		
Pipetas	2	0.0092		
Matraces	10	0.025		
<sup>(1)</sup> Incertidumbre global expandida				
<sup>(2)</sup> LGC/VAM/2000/053. Barwick, V.J.; et al. Evaluating Confidence in Analytical Measurement. Part (d): Studies of uncertainty in laboratory operations. VAM.LGC.2000.				
<sup>(3)</sup> Incertidumbre estándar				

La siguiente tabla (Tabla III.13) muestra la influencia, expresada como porcentaje, de cada término sobre la incertidumbre operacional en dos muestras de aceite, una de oliva extra virgen (EVOO-1) y otra de orujo de oliva (POO-1), seleccionadas como ejemplo para los cálculos. Se puede ver que en ambos casos, el mayor peso en la incertidumbre corresponde a la preparación y adición de la disolución madre de subrogado.

Por otro lado, aunque se podía esperar que la influencia de la masa de muestra sobre la incertidumbre fuera constante puesto que prácticamente siempre se pesa

la misma cantidad de la misma, se han observado ligeras diferencias en función de la cantidad de muestra pesada.

**Tabla III.13.** Contribución de la adición de subrogado y pesada de la muestra a la incertidumbre operacional relativa, para dos de las muestras analizadas.

	<b><u>EVOO-1</u></b>	<b><u>POO-1</u></b>
Masa subrogado ( $u^{rel}_{\text{masa-subrogado}}$ )	99.965	99.970
Masa muestra ( $u^{rel}_{\text{masa-muestra}}$ )	0.035	0.030

■ **Incertidumbre de recuperación** ( $u^{rel}_{\text{recuperación}}$ )

De acuerdo con las normas de la ISO [8], en la definición de incertidumbre se indica que la presentación de resultados debe estar libre de errores sistemáticos (sesgos proporcional y constante). En general, la evidencia del sesgo proporcional se comprueba mediante estudios de recuperación en los que se añaden, normalmente a diferentes niveles, cantidades conocidas de analito a la muestra, antes de su tratamiento. Sin embargo a veces, como en este caso, no se pueden llevar a cabo los estudios de recuperación porque, como se ha indicado anteriormente, es prácticamente imposible adquirir patrones de los analitos y, por tanto, se realiza la cuantificación mediante calibración a un solo nivel con un subrogado. Por estas razones, aunque este término figura en el diagrama de Ishikawa, no se calcula.

■ **Incertidumbre inherente** ( $u^{rel}_{\text{inherente}}$ )

Como se ha indicado anteriormente, las principales contribuciones a esta incertidumbre son:

(1) **Incertidumbre intrínseca** ( $u^{rel}_{\text{intrínseca}}$ )

La incertidumbre intrínseca asociada a la cuantificación de los 4-desmetilesteroles en aceites vegetales depende de diferentes etapas químicas como son: la saponificación de la muestra, el fraccionamiento de los fitoesteroles mediante HPLC off-line y la reacción de derivatización de los mismos a trimetil-sililesteroles para su determinación mediante GC-FID. De todas ellas, la incertidumbre que se

deriva de la saponificación no se ha considerado, debido principalmente a que los reactivos y disolventes empleados están en gran exceso con respecto al subrogado y los analitos, lo que implica un gran número de procesos que no están "bajo control metrológico".

La expresión usada en este caso para calcular la incertidumbre intrínseca es la siguiente:

$$u_{\text{intrínseca}}^{\text{rel}} = \left( (u_{\text{HPLC}}^{\text{rel}})^2 + (u_{\text{derivatización}}^{\text{re}})^2 \right)^{1/2} \quad [\text{Ec. 13}]$$

a) Incertidumbre asociada al fraccionamiento de los 4-desmetilesteroles mediante HPLC ( $u_{\text{HPLC}}^{\text{rel}}$ )

Para su estimación se ha tenido en cuenta que el residuo obtenido al final de la etapa de saponificación se disuelve en 1 mL de la fase móvil (añadido con una pipeta) y que 20  $\mu\text{L}$  de dicha disolución, se inyectan en el cromatógrafo de líquidos. En un tubo graduado se recoge la fracción eluida entre los minutos 7.5 y 9.5, y se analiza posteriormente mediante GC-FID. Además, se ha considerado que la incertidumbre asociada a la inyección cromatográfica es insignificante, comparada con la derivada del uso de otro tipo de material volumétrico.

El cálculo se ha realizado mediante la siguiente expresión:

$$u_{\text{HPLC}}^{\text{rel}} = \left( \frac{u^2(V_{\text{pipeta}})}{V_{\text{pipeta}}^2} + \frac{u^2(V_{\text{eluido}})}{V_{\text{eluido}}^2} \right)^{1/2} \quad [\text{Ec. 14}]$$

La incertidumbre asociada al  $V_{\text{eluido}}$  se calculó considerando la masa del volumen recolectado durante la ventana de tiempo (2 minutos), usando un flujo de 0.8 mL/min. El volumen recolectado (1.54 mL) se determinó a partir de la masa y la densidad de la fase móvil (0.6675 kg/L), empleando la siguiente expresión:

$$\frac{u^2(V_{\text{eluido}})}{V_{\text{eluido}}^2} = \frac{s_{\Delta m}^2}{\Delta m^2} \quad [\text{Ec. 15}]$$

siendo,

-  $s_{(\Delta m)}$  la desviación estándar de la masa del volumen recolectado.



- $\Delta m$  es la diferencia entre la masa del tubo de ensayo con la fracción eluida y el tubo de ensayo vacío.

Finalmente, la ecuación que se ha usado para calcular la incertidumbre asociada a la adición de 1 mL de fase móvil al residuo final de la saponificación es la siguiente:

$$u_{\text{pipeta}(1 \text{ mL})}^{\text{rel}} = \left( \frac{\left( \frac{\text{Error máximo}_{\text{proceso verific.}}}{\sqrt{3}} \right)^2}{V_{\text{pipeta}(1 \text{ mL})}^2} \right)^{1/2} \quad [\text{Ec. 16}]$$

dónde:

- $\text{Error máximo}_{\text{proceso verific.}}$  es el máximo error de la pipeta empleada para adicionar 1 mL de la fase móvil.
- $V_{\text{pipeta}}$ : 1 mL.

#### b) Incertidumbre asociada a la reacción de derivatización ( $u^{\text{rel}}_{\text{derivatización}}$ )

Se debe principalmente al uso de pipetas en la preparación de la mezcla derivatizante y a la adición de 200  $\mu\text{L}$  de la misma al residuo seco obtenido tras la elución. Este componente viene dado por:

$$u_{\text{derivatización}}^{\text{rel}} = \left( \left( u_{\text{(M.D.)}}^{\text{rel}} \right)^2 + \frac{u^2(V_{\text{pipeta}})}{V_{\text{pipeta}}^2} \right)^{1/2} \quad [\text{Ec. 17}]$$

Las ecuaciones empleadas para calcular cada término son:

##### b.1) Preparación de la mezcla derivatizante

$$\left( u_{\text{(M.D.)}}^{\text{rel}} \right)^2 = \frac{u^2(V_{\text{pipeta}}(1))}{V_{\text{pipeta}}^2(1)} + \frac{u^2(V_{\text{pipeta}}(2))}{V_{\text{pipeta}}^2(2)} + \frac{u^2(V_{\text{pipeta}}(3))}{V_{\text{pipeta}}^2(3)} \quad [\text{Ec. 18}]$$

dónde:

- $V_{\text{pipeta}}$  (1) es el volumen añadido de piridina (900  $\mu\text{L}$ ).
- $V_{\text{pipeta}}$  (2) es el volumen añadido de hexametildisilazano (300  $\mu\text{L}$ ).
- $V_{\text{pipeta}}$  (3) es el volumen añadido de trimetilclorosilano (100  $\mu\text{L}$ ).

En todos los casos, las pipetas usadas han sido calibradas en el laboratorio, y la ecuación utilizada es similar a la Ec. 16.

b.2) Adición de la M.D. al residuo seco obtenido del fraccionamiento mediante HPLC.

Análogamente al caso anterior, la incertidumbre asociada a la adición de la M.D. se evalúa mediante la Ec. 16.

En la Tabla III.14 se puede ver, para las dos muestras seleccionadas, la influencia de cada término (expresado como porcentaje) sobre la incertidumbre intrínseca relativa. El factor que mas contribuye, es el referido a la preparación y adición de la mezcla derivatizante.

**Tabla III.14.** Contribución de las etapas de fraccionamiento por HPLC y derivatización , a la incertidumbre intrínseca relativa ( $u^{\text{rel}}_{\text{intrínseca}}$ ).

	<b>EVOO-1</b>	<b>POO-1</b>
Fraccionamiento con HPLC de 4-desmetilesteroles <b>(<math>u^{\text{rel}}_{\text{HPLC}}</math>)</b>	11.103	11.103
Incertidumbre asociada con la reacción de derivatización <b>(<math>u^{\text{rel}}_{\text{derivatización}}</math>)</b>	88.896	88.896

(2) Incertidumbre asociada a la cuantificación ( $u^{\text{rel}}_{\text{cuantificación}}$ )

La incertidumbre asociada a la cuantificación, usando una calibración interna a un nivel, con un subrogado, depende del cociente entre las áreas de pico del subrogado (colestanol) y de cada analito ( $\text{esterol}_i$ ), estando ambas correlacionadas. Utilizando  $A_{\text{REL}}$  para referirse a la relación de áreas y aplicando la regla general de propagación de incertidumbre [22], el cálculo se hace mediante las siguientes expresiones:

$$\begin{aligned}
 u^2(A_{REL}) = & \left( \frac{\partial A_{REL}}{\partial A_{esterol(i)}} \right)^2 \times u^2(A_{esterol(i)}) + \left( \frac{\partial A_{REL}}{\partial A_{subrogado}} \right)^2 \times u^2(A_{subrogado}) + \\
 & + 2 \times r_{(A_{esterol(i)}, A_{subrogado})} \times \left( \frac{\partial A_{REL}}{\partial A_{esterol(i)}} \right) \times \left( \frac{\partial A_{REL}}{\partial A_{subrogado}} \right) \quad [Ec.19]
 \end{aligned}$$

$$\begin{aligned}
 \frac{u^2(A_{REL})}{A_{REL}^2} = & \frac{s^2(A_{esterol(i)})}{A_{esterol(i)}^2} + \frac{s^2(A_{subrogado})}{A_{subrogado}^2} - 2 \times r_{(A_{esterol(i)}, A_{subrogado})} \times \\
 & \times \frac{s(A_{esterol(i)})}{A_{esterol(i)}} \times \frac{s(A_{subrogado})}{A_{subrogado}} \quad [Ec.20]
 \end{aligned}$$

dónde:

- $s(A_{esterol(i)})$  es la desviación estándar del área de pico del esterol (i), obtenida de "n" cromatogramas GC-FID registrados en condiciones de repetibilidad.
- $s(A_{subrogado})$  es la desviación estándar del área de pico del subrogado, obtenida de "n" cromatogramas GC-FID registrados en condiciones de repetibilidad.
- $A_{esterol(i)}$  es la media aritmética de las áreas de pico del esterol (i) correspondientes a "n" cromatogramas.
- $A_{subrogado}$  es la media aritmética de las áreas de pico del subrogado correspondientes a "n" cromatogramas.
- $r_{(A_{esterol(i)}, A_{subrogado})}$  es el coeficiente de correlación de las áreas, cuyo valor es 1 en el caso de distribuciones multivariantes.

Las ecuaciones finales usadas son:

$$\frac{s^2(A_{esterol(i)})}{A_{esterol(i)}^2} = \frac{\left( \frac{s_{repet}(A_{esterol(i)})}{\sqrt{n}} \right)^2}{A_{esterol(i)}^2} \quad [Ec. 21]$$

$$\frac{s^2(A_{\text{subrogado}})}{A_{\text{subrogado}}^2} = \frac{\left( s_{\text{repet}}(A_{\text{subrogado}}) / \sqrt{n} \right)^2}{A_{\text{subrogado}}^2} \quad [\text{Ec. 22}]$$

$$\frac{s(A_{\text{esterol}(i)})}{A_{\text{esterol}(i)}} = \sqrt{\frac{\left( s_{\text{repet}}(A_{\text{esterol}(i)}) / \sqrt{n} \right)^2}{A_{\text{esterol}(i)}^2}} \quad [\text{Ec. 23}]$$

$$\frac{s(A_{\text{subrogado}})}{A_{\text{subrogado}}} = \sqrt{\frac{\left( s_{\text{repet}}(A_{\text{subrogado}}) / \sqrt{n} \right)^2}{A_{\text{subrogado}}^2}} \quad [\text{Ec. 24}]$$

dónde “n” es el número de inyecciones replicadas. Para la estimación de esta incertidumbre, el número de réplicas osciló entre 7 y 9, dependiendo del tipo de aceite.

La Tabla III.15 muestra los resultados obtenidos en el cálculo de  $u^{\text{rel}}_{\text{cuantificación}}$  para los 4-desmetilesteroles encontrados en las dos muestras seleccionadas. Se puede ver que, a excepción de campesterol, estigmasterol,  $\beta$ -sitosterol y  $\Delta^{5,24}$ -estigmastadienol, los valores de incertidumbre son mayores para los esteroides encontrados en la muestra POO-1 (aceite de orujo de oliva). Este hecho se puede atribuir a la mayor o menor concentración de los analitos, lo que afecta a su resolución cromatográfica.

Finalmente, en la Tabla III.16 se recoge el peso que las incertidumbres intrínseca relativa y de cuantificación, tienen sobre la incertidumbre inherente. Como dicha influencia depende de cada analito, solo se dan los datos que corresponden al campesterol, estigmasterol,  $\beta$ -sitosterol y  $\Delta^{5,24}$ -estigmastadienol. Como se puede ver, en la mayor parte de los casos, la principal contribución se debe a la incertidumbre intrínseca (especialmente la que proviene de la reacción de derivatización), excepto para el campesterol en el aceite de orujo de oliva en el cual, la cuantificación predomina sobre el resto de los factores químicos.

**Tabla III.15.** Valores de incertidumbre relativa obtenidas en la cuantificación de los 4-desmetilesteroles en las dos muestras analizadas.

Analito <sup>(1)</sup>	<b>EVOO-1</b>	<b>POO-1</b>
	$u^{rel}_{cuantificación}$	$u^{rel}_{cuantificación}$
<b>1</b>	$1.42 \times 10^{-2}$	$1.73 \times 10^{-2}$
<b>2</b>	===	$1.48 \times 10^{-2}$
<b>3</b>	$1.00 \times 10^{-2}$	$1.98 \times 10^{-2}$
<b>4</b>	$5.75 \times 10^{-3}$	$2.80 \times 10^{-4}$
<b>5</b>	$4.81 \times 10^{-3}$	$2.00 \times 10^{-2}$
<b>6</b>	$6.33 \times 10^{-3}$	$4.01 \times 10^{-3}$
<b>7</b>	===	$1.61 \times 10^{-2}$
<b>8</b>	===	$1,56 \times 10^{-2}$
<b>9</b>	$1.21 \times 10^{-3}$	$1.99 \times 10^{-2}$
<b>10</b>	$3.65 \times 10^{-3}$	$1.28 \times 10^{-3}$
<b>11</b>	$4.83 \times 10^{-4}$	$1,10 \times 10^{-2}$
<b>12</b>	$3.65 \times 10^{-3}$	$6,84 \times 10^{-3}$
<b>13</b>	$3.68 \times 10^{-3}$	$3.29 \times 10^{-3}$
<b>14</b>	$7.65 \times 10^{-3}$	$2.03 \times 10^{-2}$
<b>15</b>	$1.48 \times 10^{-3}$	$9.63 \times 10^{-3}$

(1) Los analitos se han codificado según Tabla III.2.

**Tabla III.16.** Contribución de las incertidumbres intrínseca y de cuantificación a los valores de incertidumbre inherente, para los analitos seleccionados.

Analito <sup>(1)</sup>		<b>EVOO-1</b>	<b>POO-1</b>
		<b>(4)</b>	$(u^{rel}_{intrínseca})$
	$(u^{rel}_{cuantificación})$	10.087%	57.064%
<b>(6)</b>	$(u^{rel}_{intrínseca})$	88.015%	94.816%
	$(u^{rel}_{cuantificación})$	11.987%	5.174%
<b>(10)</b>	$(u^{rel}_{intrínseca})$	88.518%	99.447%
	$(u^{rel}_{cuantificación})$	11.471%	0.553%
<b>(13)</b>	$(u^{rel}_{intrínseca})$	95.601%	96.456%
	$(u^{rel}_{cuantificación})$	4.399%	3.543%

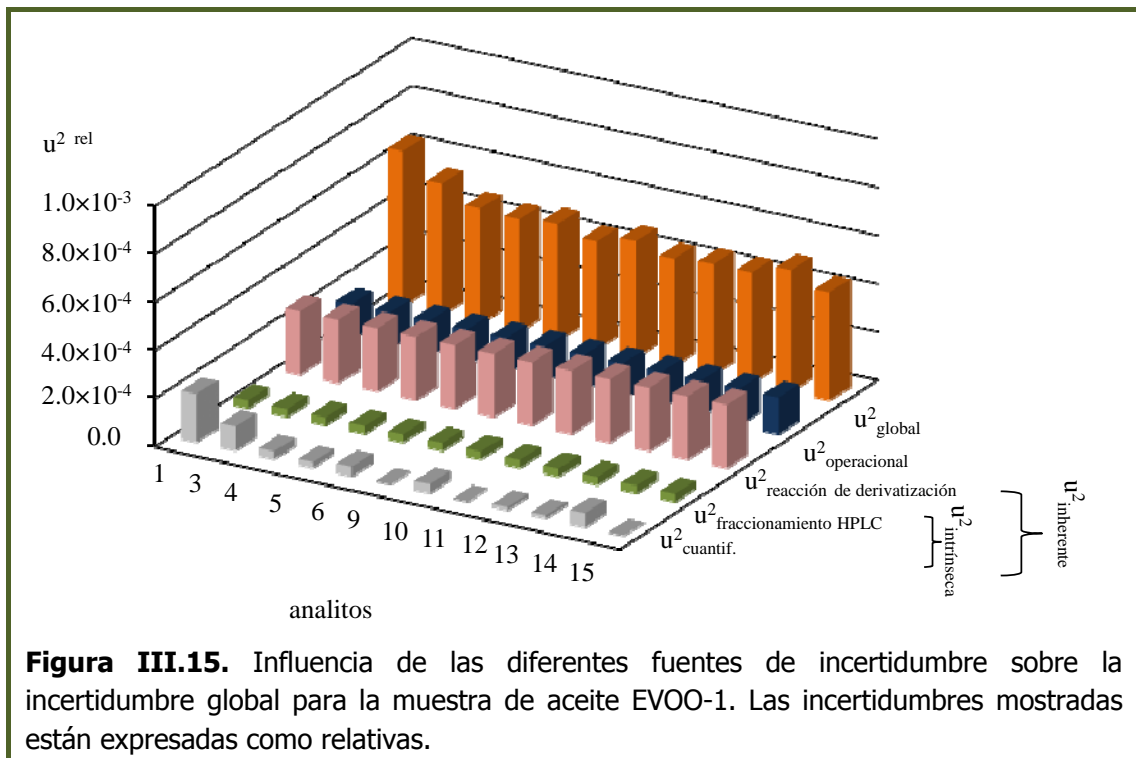
(1) Los analitos se han codificado según Tabla III.2.

■ **Incertidumbre global relativa ( $u^{rel}_{global}$ )**

Se obtiene a partir de la expresión:

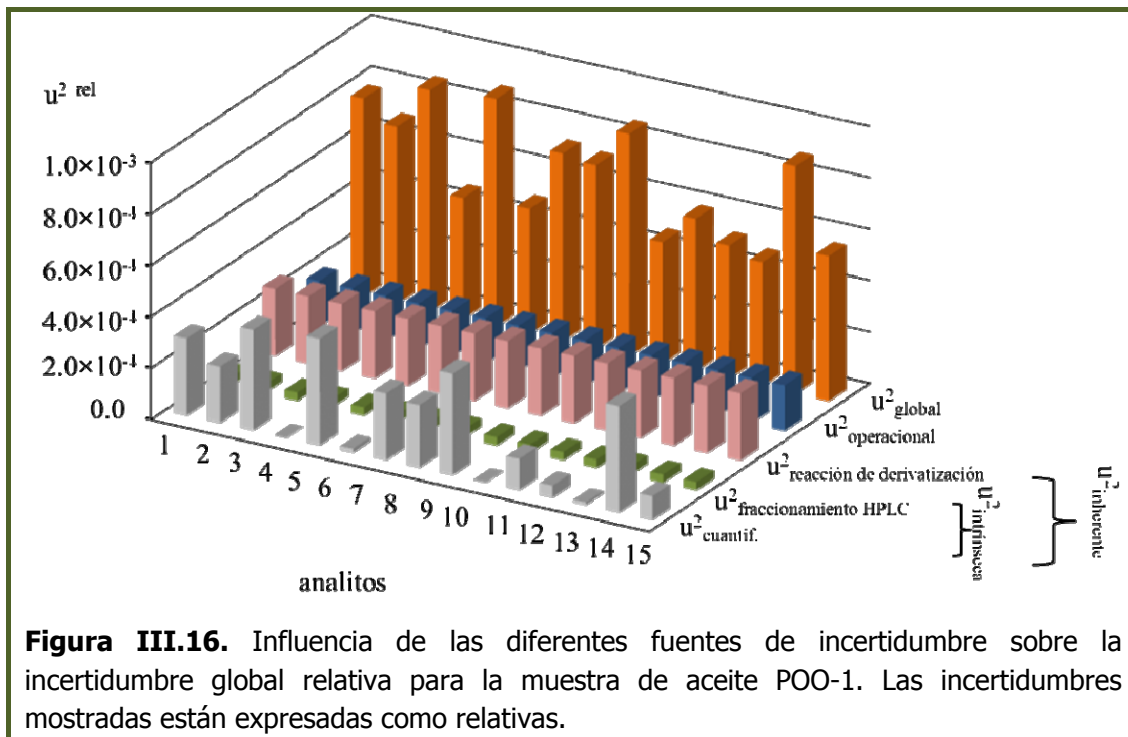
$$u^{rel}_{global}(\text{esterol } i) \frac{u_{global}(\text{esterol } i)}{C(\text{esterol } i)} = \sqrt{(u^{rel}_{op})^2 + (u^{rel}_{inherente})^2} \quad [\text{Ec. 25}]$$

En donde  $C(\text{esterol}_i)$  es la concentración estimada para cada analito. La contribución de las diferentes fuentes de incertidumbre en las dos muestras seleccionadas se muestra en las Figuras III.15 y III.16.



Se puede ver que para el aceite de orujo de oliva los valores de la incertidumbre global son mayores o del mismo orden de magnitud que para el aceite de oliva virgen extra.

En ambas muestras, ninguna de las fuentes principales es despreciable, siendo en todos los casos la incertidumbre inherente relativa, la predominante. En relación con esta fuente de incertidumbre, aunque su influencia varía dependiendo del 4-desmetilesterol, de manera general se podría decir que el aceite de oliva virgen se ve principalmente afectado por la contribución de los aspectos químicos, como la reacción de derivatización, mientras que el aceite de orujo de oliva se ve principalmente afectado por la etapa de cuantificación.



La evaluación de la incertidumbre asociada a la determinación de 4-desmetil-esteroles en aceites vegetales aplicando un método analítico basado en la separación de los mismos, a partir de la fracción insaponificable, mediante un fraccionamiento por HPLC on-line, posterior derivatización como TMS derivados y determinación mediante GC-FID, permite deducir que:

- De las fuentes de incertidumbre evaluadas, la principal contribución está ligada a los aspectos químicos del método (incertidumbre inherente).
- Dependiendo del tipo de aceite, así como del 4-desmetilesterol considerado, la etapa de cuantificación juega un papel más o menos importante en la incertidumbre global relativa.

■ Incertidumbre global expandida relativa ( $U_{\text{global}}^{\text{rel}}$ )

Con objeto de proporcionar un nivel de confianza del 95 % a los resultados finales, se obtuvieron las incertidumbres globales expandidas relativas ( $U_{\text{global}}^{\text{rel}}(\text{esterol}_i)$ ) multiplicando las incertidumbres globales por un factor de cobertura  $k = 2$ . La Tabla III. 17 recoge los valores estimados en las dos muestras seleccionadas como ejemplo de cálculo.

**Tabla III.17.** Valores de incertidumbre global expandida estimados en las muestras de aceite EVOO-1 y POO-3.

<u>Analito<sup>(1)</sup></u>	<b>EVOO-1</b> ± <b>U<sup>rel</sup><sub>global</sub></b>	<b>POO-1</b> ± <b>U<sup>rel</sup><sub>global</sub></b>
<b>1</b>	0.051	0.056
<b>2</b>	===	0.053
<b>3</b>	0.047	0.059
<b>4</b>	0.044	0.043
<b>5</b>	0.043	0.060
<b>6</b>	0.044	0.044
<b>7</b>	===	0.054
<b>8</b>	===	0.054
<b>9</b>	0.042	0.059
<b>10</b>	0.044	0.044
<b>11</b>	0.042	0.049
<b>12</b>	0.043	0.046
<b>13</b>	0.043	0.044
<b>14</b>	0.045	0.060
<b>15</b>	0.042	0.048

<sup>(1)</sup> Los analitos se han codificado según Tabla III.2.

### III.4.2. Estimación de la veracidad

Para verificar la veracidad de los resultados, se han utilizado dos aceites vegetales (oliva virgen extra, AT-EVOO-2 y orujo, AT-POO-2) analizados por el Laboratorio Agroalimentario de Atarfe, acreditado para la realización de ensayos de productos alimenticios. A su vez, el test de veracidad utilizado es una adaptación del propuesto en la Guía ISO 33:2000 sobre "Usos de materiales de referencia certificados" y viene dado por la siguiente expresión:

$$|\bar{X}_1 - \bar{X}_2| \leq K \times \sqrt{2 \times U^2(C_{\text{esterol}(i)})} \quad [\text{Ec. } 26]$$

dónde

-:  $\bar{X}_1$  valor de referencia del material analizado por el laboratorio acreditado.



- $\bar{X}_2$ : valor medio obtenido al aplicar el método en estudio.
- K: factor de cobertura. Al igual que en el estudio de la incertidumbre, se ha utilizado el valor de  $k = 2$  ( $p = 0.95$ )
- $U(C_{\text{esterol}(i)})$ : incertidumbre asociada a la concentración de analito determinada por el método en estudio.

Los resultados obtenidos para cada muestra se recogen en las Tablas III.17 y III.18.

**Tabla III.17.** Resultados obtenidos para la muestra AT-EVOO-2.

<b>Muestra:AT-EVOO-1</b>					
<b>Analito</b>	<b>X<sub>1</sub><sup>(a)</sup></b>	<b>X<sub>2</sub><sup>(b)</sup></b>	<b>U<sup>(c)</sup></b>	$ \bar{X}_1 - \bar{X}_2  \leq K \sqrt{2U^2(C_{\text{esterol}(i)})}$	<b>¿Es veraz?</b>
1	1.58	1.55	0.08	$0.01 \leq 0.11$	SI
2	n.d. <sup>(1)</sup>	n.d.	n.c. <sup>(2)</sup>	-	-
3	3.35	3.27	0.14	$0.08 \leq 0.20$	SI
4	47.50	48.93	1.92	$1.43 \leq 2.72$	SI
5	1.42	1.50	0.06	$0.08 \leq 0.08$	SI
6	13.78	13.56	0.60	$0.22 \leq 0.85$	SI
7	0.81	<b>n.d.</b>	<b>n.c.</b>	=====	No verificado
8	1.43	<b>n.d.</b>	<b>n.c.</b>	=====	No verificado
9	15.05	15.97	0.68	$0.15 \leq 0.96$	SI
10	1262.50	1201.76	52.90	$60.74 \leq 74.81$	SI
11	<b>8.03</b>	<b>10.92</b>	<b>0.46</b>	<b><math>1.71 \leq 0.65</math></b>	<b>NO</b>
12	113.14	107.77	4.62	$5.37 \leq 6.53$	SI
13	7.49	7.82	0.34	$0.33 \leq 0.48$	SI
14	5.47	5.75	0.28	$0.28 \leq 0.40$	SI
15	6.90	7.06	0.30	$0.16 \leq 0.42$	SI

(a) valor de referencia  
 (b) valor medio obtenido mediante el método  
 (c) Incertidumbre global expandida

(1) n.d.: no detectado  
 (2) n.d.: no calculado

**Tabla III.18.** Resultados obtenidos para la muestra AT-POO-2.

<b>Muestra: AT-POO-2</b>						
<b>Analito</b>	<b>X<sub>1</sub><sup>(a)</sup></b>	<b>X<sub>2</sub><sup>(b)</sup></b>	<b>U<sub>esterol(i)</sub><sup>(c)</sup></b>	<b><math> \bar{X}_1 - \bar{X}_2  \leq K \sqrt{2U^2(C_{esterol(i)})}</math></b>	<b>¿Es veraz?</b>	
<b>1</b>	6.51	6.37	0.35	0.14 ≤ 0.49	SI	
<b>2</b>	5.40	5.80	0.31	0.40 ≤ 0.44	SI	
<b>3</b>	2.99	3.03	0.18	0.04 ≤ 0.25	SI	
<b>4</b>	75.84	78.45	3.41	2.61 ≤ 4.82	SI	
<b>5</b>	5.42	5.85	0.35	0.43 ≤ 0.49	SI	
<b>6</b>	32.83	31.60	1.40	1.23 ≤ 1.98	SI	
<b>7</b>	5.05	5.37	0.29	0.32 ≤ 0.41	SI	
<b>8</b>	28.38	26.50	1.42	1.88 ≤ 2.09	SI	
<b>9</b>	32.97	33.91	2.00	0.94 ≤ 2.83	SI	
<b>10</b>	2053.75	2004.35	87.23	49.40 ≤ 123.36	SI	
<b>11</b>	47.41	50.23	2.45	2.82 ≤ 3.46	SI	
<b>12</b>	52.14	55.59	2.53	3.45 ≤ 3.58	SI	
<b>13</b>	35.28	37.60	1.65	2.32 ≤ 2.33	SI	
<b>14</b>	11.74	12.69	0.76	0.82 ≤ 1.07	SI	
<b>15</b>	13.70	13.12	0.62	0.58 ≤ 0.88	SI	

(a) valor de referencia  
 (b) valor medio obtenido mediante el método  
 (c) Incertidumbre global expandida

En el caso de la muestra AT-EVOO-2, se puede observar que, tras aplicar el método evaluado el 80% de las concentraciones determinadas son estadísticamente iguales a las suministradas por el Laboratorio de Atarfe. Para el 20% restante, dado que no se detectan ni  $\Delta^7$ -campesterol (7) ni  $\Delta^{5,23}$ -estigmastadienol (8) no se ha podido aplicar el test de veracidad, y hay que señalar que se ha encontrado una diferencia estadística significativa en el caso del estigmastanol/sitostanol (11). Este último hecho se puede relacionar con la posición que ocupa su pico en el cromatograma (entre dos picos mayoritarios en la muestra,  $\beta$ -sitosterol (10) y  $\Delta^5$ avenasterol (12)), lo que puede dar lugar a que, en función de las concentraciones presentadas por estos analitos, el sitostanol pueda ser cuantificado por exceso o incluso, como contrapartida, no detectado.

En el caso de la muestra AT-POO-2, se puede observar que, tras aplicar el método evaluado, el 100% de las concentraciones determinadas son estadísticamente iguales a las suministradas por el Laboratorio de Atarfe.

Cabe comentar que aunque Cert et al. [5] indican que el fraccionamiento mediante HPLC origina una mejora en la cuantificación de los  $\Delta^7$ -desmetilesteroles con respecto a la utilización de la CCF (se obtienen valores más elevados de estos analitos, verificados mediante ensayos interlaboratorio), este hecho no se ha podido comprobar completamente en ambas muestras. Se puede decir que en la muestra AT-EVOO-2, se confirma para los  $\Delta^7$ -desmetilesteroles encontrados,  $\Delta^7$ -estigmastenol (14) y  $\Delta^7$ -avenasterol (15), mientras que en la muestra AT-POO-2, solo se ha podido comprobar en dos de ellos, el  $\Delta^7$ -estigmasterol (7) y el  $\Delta^7$ -estigmastenol (14).

Por otro lado, y de acuerdo con dichos autores, los valores de concentración ligeramente superiores encontrados en el 73.3% de los 4-desmetilesteroles cuantificados en la muestra AT-POO-2 y en el 53.8% de los encontrados en la muestra AT-EVOO-2, podrían atribuirse al hecho de que el fraccionamiento mediante HPLC, minimiza las posibles pérdidas de subrogado ocasionadas en el fraccionamiento mediante CCF.

Finalmente, con objeto de completar la evaluación de la aptitud del método propuesto se procedió a estimar la repetibilidad, tanto del fraccionamiento mediante HPLC (replicando el fraccionamiento de alícuotas del insaponificable), como de la inyección cromatográfica (haciendo inyecciones replicadas de alícuotas de cada uno de los volúmenes recogidos en el eluido). Las réplicas utilizadas oscilaron entre 3 y 9. En ambos casos, los valores de % DER encontrados estaban comprendidos entre 0.7% y un 12%, correspondiendo los valores más altos a aquellos analitos que pueden presentar problemas de resolución, independientemente de que se hagan réplicas del fraccionamiento o de la inyección, si bien es verdad que cuando se hacen réplicas de la inyección, los % DER encontrados son ligeramente inferiores sin que las diferencias sean estadísticamente significativas.

A la vista de los resultados obtenidos a lo largo de este capítulo, se puede considerar que:

**Ajustadas las condiciones cromatográficas para la determinación mediante CG-FID, optimizado el volumen de mezcla derivatizante utilizado para obtener los TMS-derivados y la ventana de tiempo para la recolección de la fracción de 4-desmetilesteroles mediante HPLC, al haber**

**estimado y cuantificado las principales fuentes de incertidumbre que afectan a los resultados, y realizado la verificación de los mismos con dos muestras procedentes de un laboratorio acreditado, se puede concluir que el método desarrollado es apto para la determinación del perfil de concentraciones de los principales 4-desmetilesteroles presentes en aceites vegetales y mezclas de los mismos.**

**La aplicabilidad del mismo a la cuantificación de aceite de oliva en alimentos se abordará en los siguientes capítulos.**

## BIBLIOGRAFÍA

- [1] Toivo, J.; Piironen, V.; Kalo, P. & Varo, P. *Gas chromatographic determination of major sterols in edible oils and fats using solid phase extraction sample preparation*. *Chromatographia*, 1998, 48, 745-750.
- [2] Phillips, K.M.; Ruggio, D.M.; Toivo, J.I.; Swank, M.A. & Simpkins, A.H. *Free and Esterified Sterol Composition of Edible Oils and Fats*. *J. Food Comp. An.*, 2002, 15, 123-142.
- [3] Azadmard-Darmirchi, S. *Novel solid-phase extraction method to separate 4-desmethyl-, 4-monomethyl and 4,4'-dimetilesterols in vegetable oils*. *J. Chromatogr. A*, 2006, 1108, 183-187.
- [4] Kalo, P. & Kuuranne, T. *Analysis of free and esterified sterols in fats and oils by flash chromatography, gas chromatography and electrospray tandem mass spectrometry*. *J. Chromatogr. A*, 2001, 935, 237-248.
- [5] Cert, A.; Moreda, W. & García-Moreno, J. *Determinación de esteroides y alcoholes triterpénicos en aceite de oliva mediante separación de la fracción por cromatografía líquida de alta eficacia y análisis por cromatografía de gases. Estandarización del método analítico*. *Grasas y aceites*, 1997, 48(4), 207-218.
- [6] Espejo, J. *Estudio analítico comparado entre el aceite de acebuchina y el aceite de oliva virgen*. Tesis doctoral defendida en la Universidad de Sevilla. 2005.
- [7] Grob, K.; Lanfranchi, M. & Mariani, C. *Determination of free and esterified sterols and wax esters in oils and fats by coupled liquid chromatography-gas chromatography*. *J. Chromatogr. A*, 1989, 471, 397-405.
- [8] Grob, K.; Lanfranchi, M. & Mariani, C. *Evaluation of olive oils through the fatty alcohols, the sterols and their esters by coupled LC-GC*. *J. American Oil Chemists' Society*, 1990, 67, 626-634.
- [9] Biedermann, M.; Grob, K. & Mariani, C. *Transesterification and on line LC-GC for determining the sum of free and esterified sterols in edible oils and fats*. *Fat Sci. Technol.*, 1993, 95, 127-133.

- [10] Toledano, R.; Cortés, J.M.; Andini, J.; Vázquez, A. & Villén, J. *On-line derivatization with on-line coupled normal phase liquid chromatography-gas chromatography using the through oven transfer adsorption desorption interface: Application to the analysis of total sterols in edible oils*. J. Chromatogr. A, 2012, 1256, 191-196.
- [11] Cortés, J.M.; Sánchez, R.; Villén, J. & Vázquez, A. *Analysis of unsaponifiable compounds of edible oils by automated on-line coupling reserved-phase liquid chromatography-gas chromatography using the through oven transfer adsorption desorption interface*. J. Agric. Food Chem., 2006, 54, 6963-6968.
- [12] Villén, J.; Blanck, G.P.; Ruiz del Castillo, M.L. & Herraiz, M. *Rapid and simultaneous analysis of free sterols, tocopherols and squalene in edible oils by coupled reserved phase liquid chromatography-gas chromatography*. J. Agric. Food Chem., 1998, 46, 1419-1422.
- [13] Señoráns, F.J.; Villén, J.; Tabera, J. & Herraiz, M. *Simplex optimization of the direct analysis of free sterols in sunflower oil by on-line coupled reversed phase liquid chromatography-gas chromatography*. J. Agric. Food Chem., 1998, 46, 1022-1026.
- [14] Toledano, R.M.; Cortés, J.M.; Rubio-Moraga, A.; Villén, J. & Vázquez, A. *Analysis of free and esterified sterols in edible oils by online reversed phase liquid chromatography-gas chromatography (RPLC-GC) using the through oven transfer adsorption desorption (TOTAD) interface*. Food Chemistry, 2012, 135, 610-615.
- [15] COI/T.20/Doc. nº 30 (2009) *Determinación de la composición y del contenido de esteroides y dialcoholes triterpénicos mediante cromatografía de gases con columna capilar*. Consejo Oleícola Internacional, Madrid.
- [16] Costin, C.D.; Hansen, S.L. & Chambers, D.P. *Using theoretical correction factors for quantitative analysis of sterols and sterols concentrates*. J. Am. Oil Chem. Soc., 2009, 86, 117-118.
- [17] Snyder, L.R.; Kirkland, J.J. & Dolan, J. W. *Introduction to Modern Liquid Chromatography*, 3rd Edition, Wiley-Blackwell: Hoboken, NJ, 2010.

- [18] Ellison, S.L.R. & Williams, A. (Eds). Eurachem/CITAC guide: "*Quantifying Uncertainty in Analytical Measurement*", Third edition, (2012) ISBN 978-0-948926-30-3. Disponible en [www.eurachem.org](http://www.eurachem.org).
- [19] Konieczka, P. & Namiésnik, J. *Review Estimating uncertainty in analytical procedures base on chromatographic techniques*. J. Chromatogr. A., 2010, 1217, 882-891.
- [20] Bagur, G.; Sánchez-Viñas. M.; Gázquez, D.; Ortega, M. & Romero R. *Estimation of the uncertainty associated with the standard addition methodology when a matrix effect is detected*. Talanta, 2005, 66, 1168–1174.
- [21] Romero, R.; Bagur, M.G.; Gázquez, D., Sánchez-Viñas, M.; Cuadros-Rodríguez, L. & Ortega, M. *Estimation of the Main Sources of Uncertainty in Chromatographic Analysis: Determination of Biogenic Amines*. LC-GC N. Am., 2004, 22, 150-164.
- [22] JCGM (Joint Committee for Guides in Metrology) 2008. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement*. (GUM), JCGM 100:2008.

## **CAPÍTULO IV**





# Estudio quimiométrico del perfil de concentraciones de los 4-desmetil-esteroles en aceites vegetales

El objetivo fundamental del estudio quimiométrico desarrollado en este capítulo se basa en la evaluación de la viabilidad de utilizar el perfil de concentraciones de los 4-desmetilesteroles más habituales en aceites vegetales a partir de su determinación por aplicación del método cromatográfico basado en la combinación "off line" de la HPLC con la CG-FID desarrollado en el capítulo anterior.

Aunque la revisión bibliográfica realizada en la Introducción de esta Memoria, y de la que aquí se citan las publicaciones destacadas[1,2,3,4,5,6,7,8,], indica que se conocen estudios quimiométricos aplicados a fitoesteroles presentes en aceites, la mayoría de ellos se basan en la utilización conjunta de estos compuestos con otras clases de lípidos también presentes en aceites (triglicéridos, eritrol y uvaol, ácidos grasos, tocoferoles, etc.), y sobre todo en la utilización de estos perfiles de concentración, para realizar estudios de caracterización, detección de adulteraciones, denominaciones de origen, etc. No habiéndose encontrado a priori, ningún estudio enfocado a su uso para la posible diferenciación o discriminación entre diferentes tipos de aceites, entre los cuales, dadas sus propiedades, el aceite de oliva juega un papel importante al ser un alimento destacado en la dieta mediterránea.

Por ello, parece adecuado realizar un estudio quimiométrico para comprobar si el perfil de concentraciones de los 4-desmetilesteroles podría utilizarse como "huella dactilar discreta" para poder discriminar entre aceites vegetales.

#### **IV.1. ANÁLISIS DE DIFERENTES MUESTRAS DE ACEITES VEGETALES MEDIANTE HPLC-GC(FID) OFF-LINE**

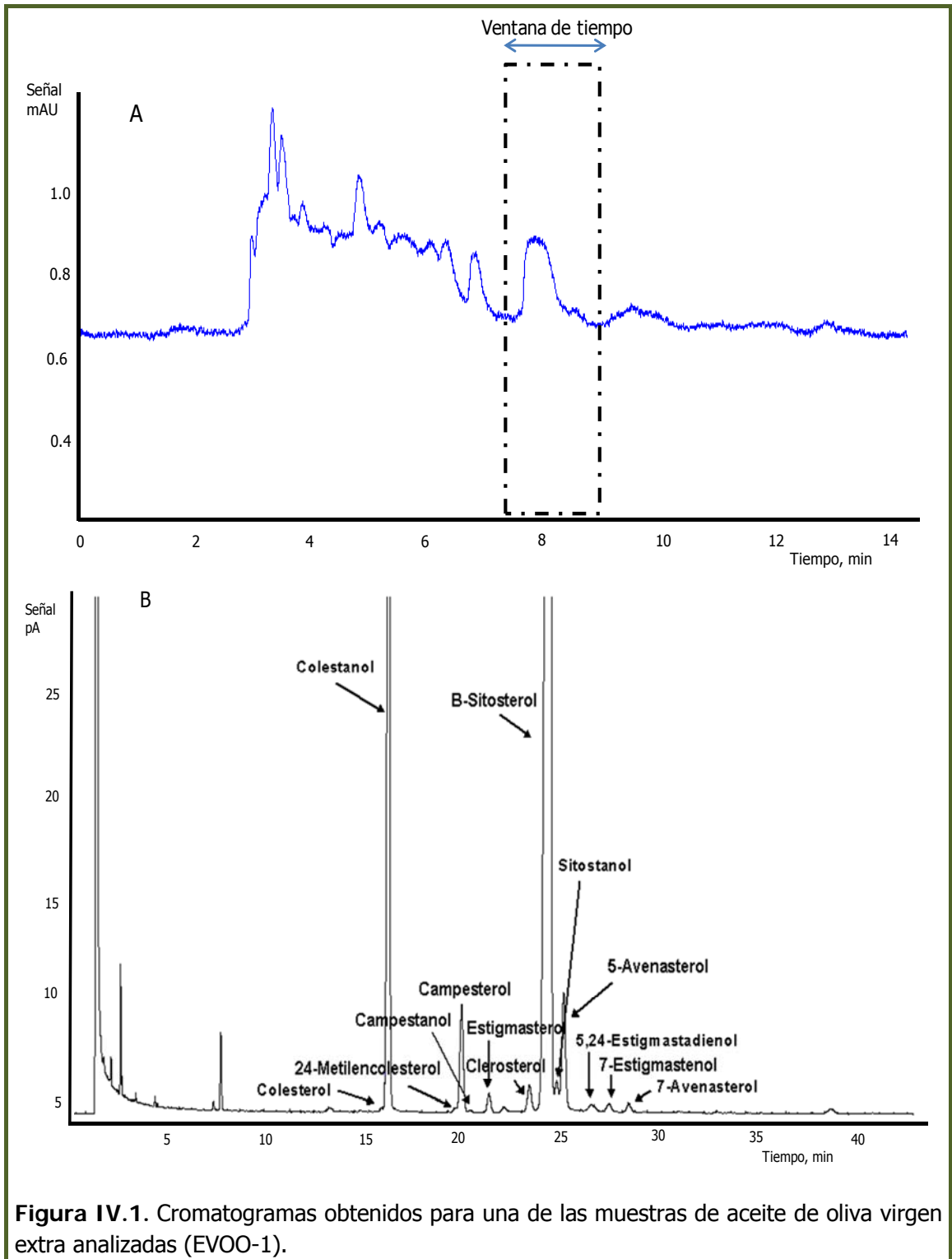
De acuerdo con la validación realizada en el capítulo anterior relativa a la determinación de 4-desmetilesteroles mediante la determinación cromatográfica "oof-line" NP-HPLC-GC(FID) , la cuantificación utilizada consituye un ejemplo típico de calibración interna a un solo nivel en la cual, se usa un único patrón interno o subrogado (5 $\alpha$ -colestanol) para la cuantificación de los analitos.

Tomando esta premisa como punto de partida y con objeto de evaluar la aplicabilidad del método desarrollado, se llevaron a cabo los análisis de las 24 muestras de aceites vegetales recogidas en la Tabla IV.1.

**Tabla IV. 1.** Muestras de aceites vegetales analizadas.

<b>Tipo de aceite</b>	<b>Código</b>	<b>Cantidad</b>
Oliva Virgen Extra	EVOO-1; EVOO-27;EVOO-28	3
Oliva Refinado	ROO-1	1
Orujo de Oliva	POO-1;POO-2; POO-3	3
Girasol	SFO	1
Girasol alto oleico	SFO(o)-1; SFO(o)-2	2
Colza	RO-1; RO-2	2
Canola	CanO-1; CanO-2	2
Soja	SyO-1; SyO-2	2
Maíz	CO-1; CO-2	2
Cacahuete	PeaO-1; PeaO-2	2
Uva	GO-1; GO-2	2
Sésamo	SesO-1; SesO-2	2

En las siguientes figuras (Figura IV.1 a IV.10), y como ejemplo ilustrativo de los diferentes tipos de aceites analizados, se muestran sus perfiles cromatográficos. (A): HPLC-UV; (B): GC-FID.



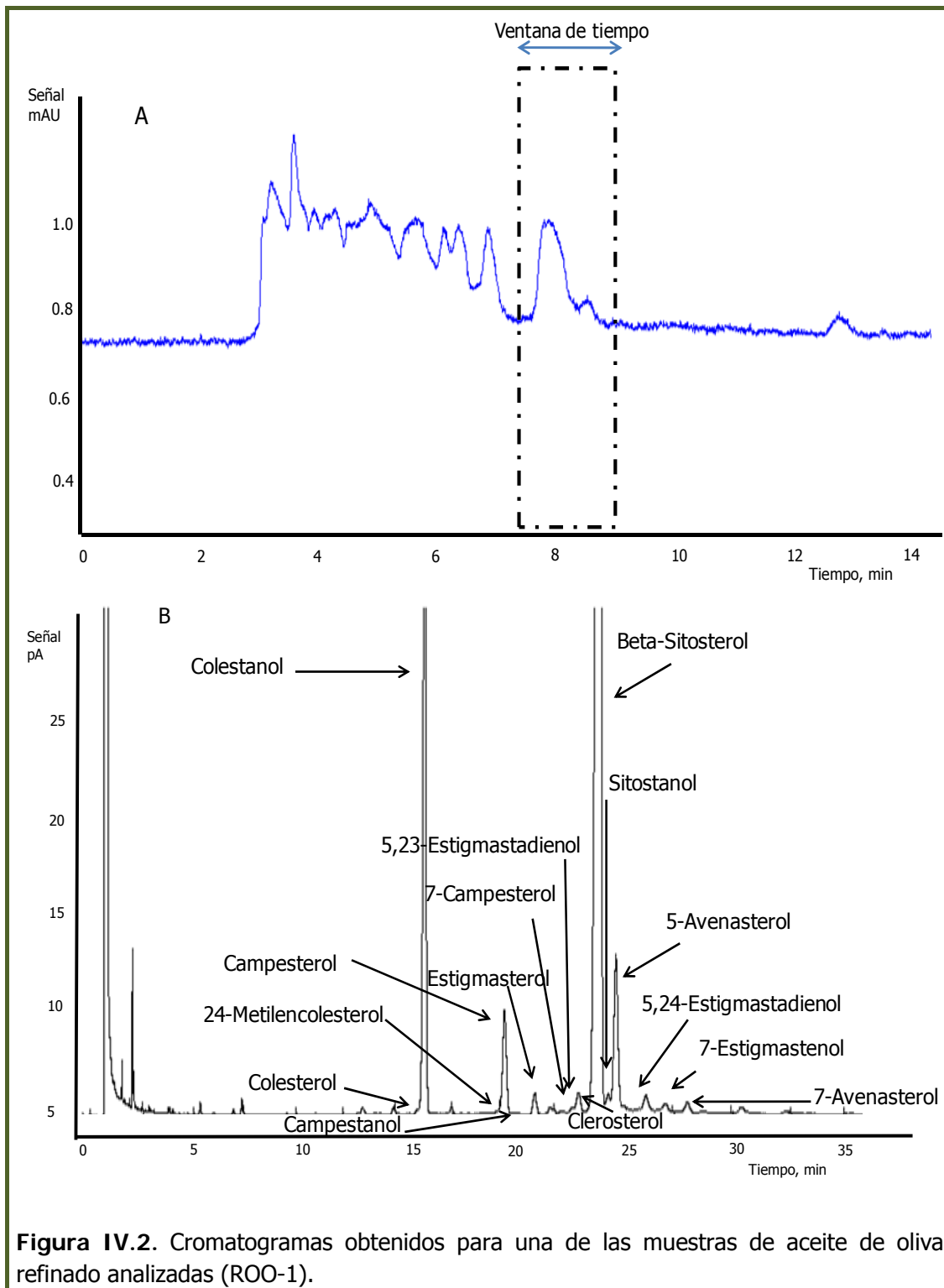
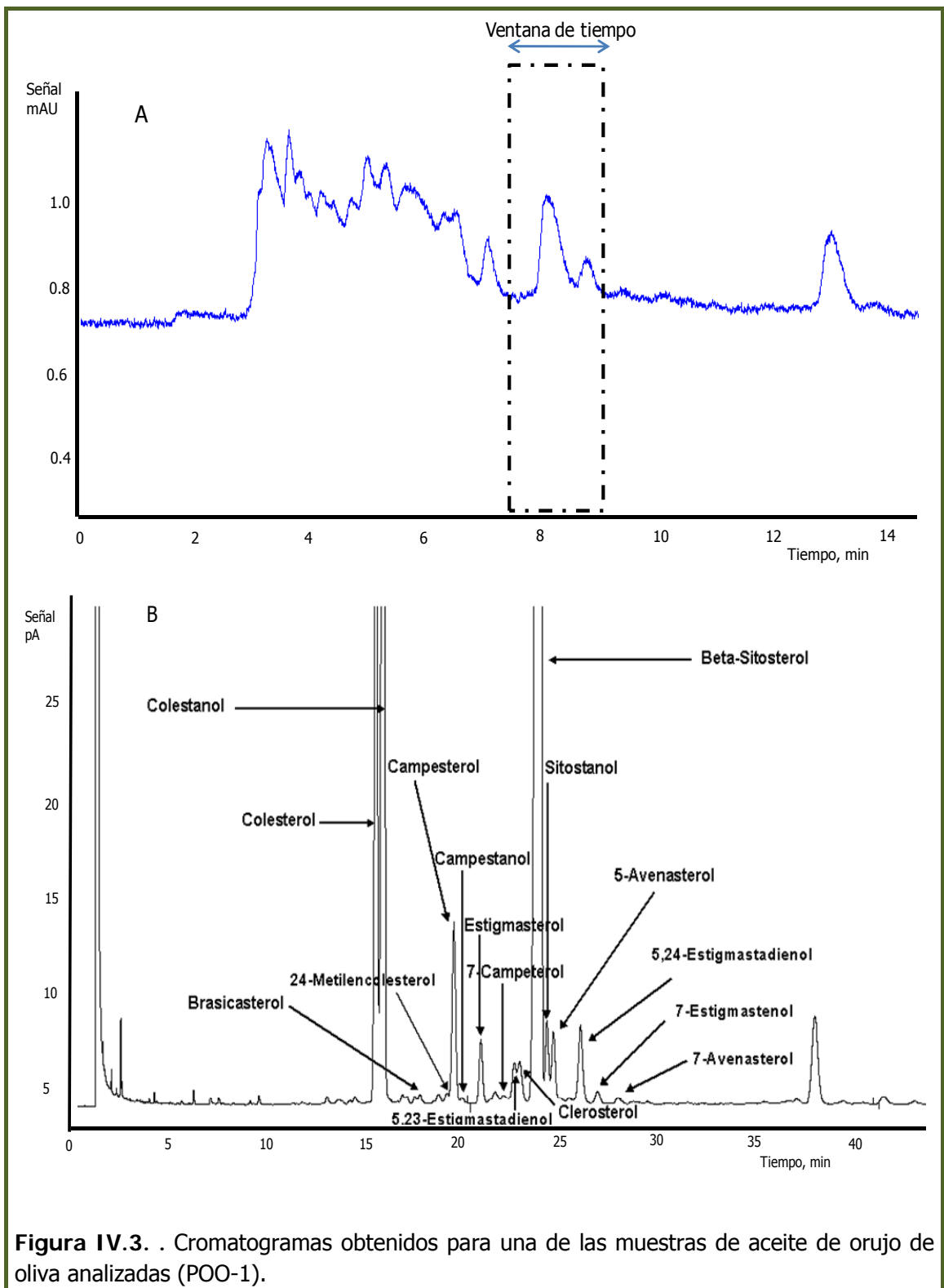


Figura IV.2. Cromatogramas obtenidos para una de las muestras de aceite de oliva refinado analizadas (ROO-1).



**Figura IV.3.** . Cromatogramas obtenidos para una de las muestras de aceite de orujo de oliva analizadas (POO-1).

## Estudio quimiométrico del perfil de concentraciones ...

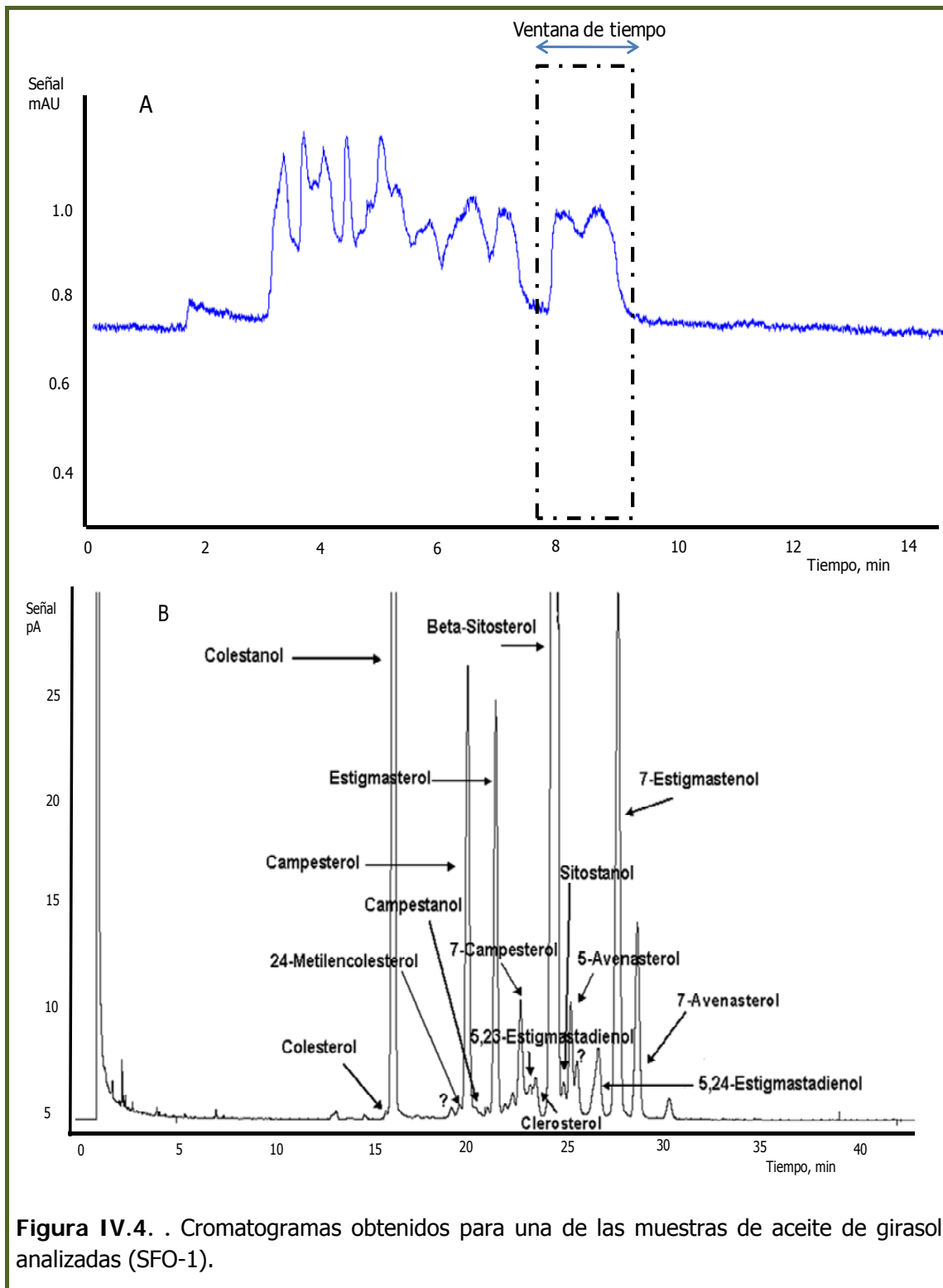
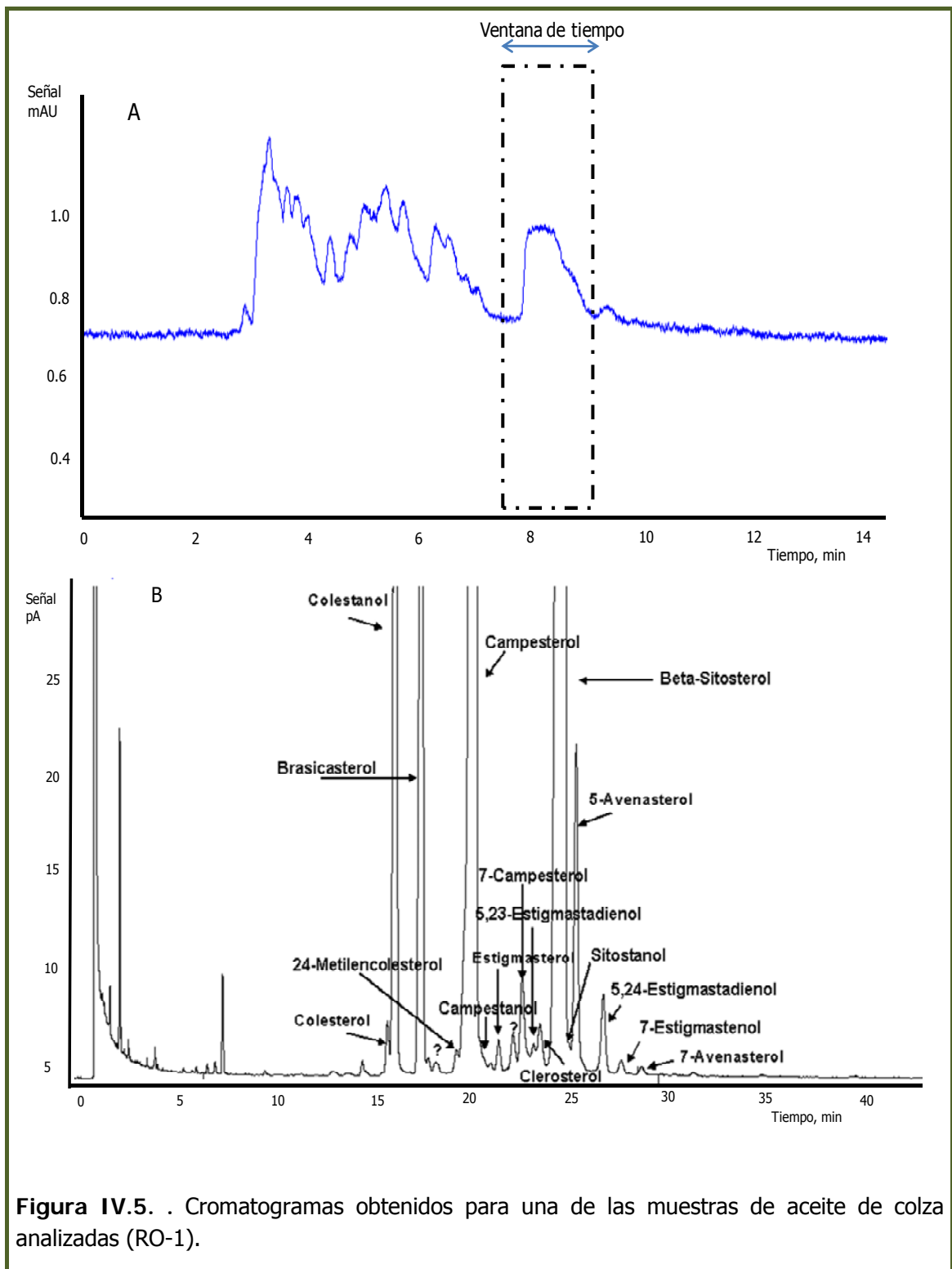


Figura IV.4. . Cromatogramas obtenidos para una de las muestras de aceite de girasol analizadas (SFO-1).





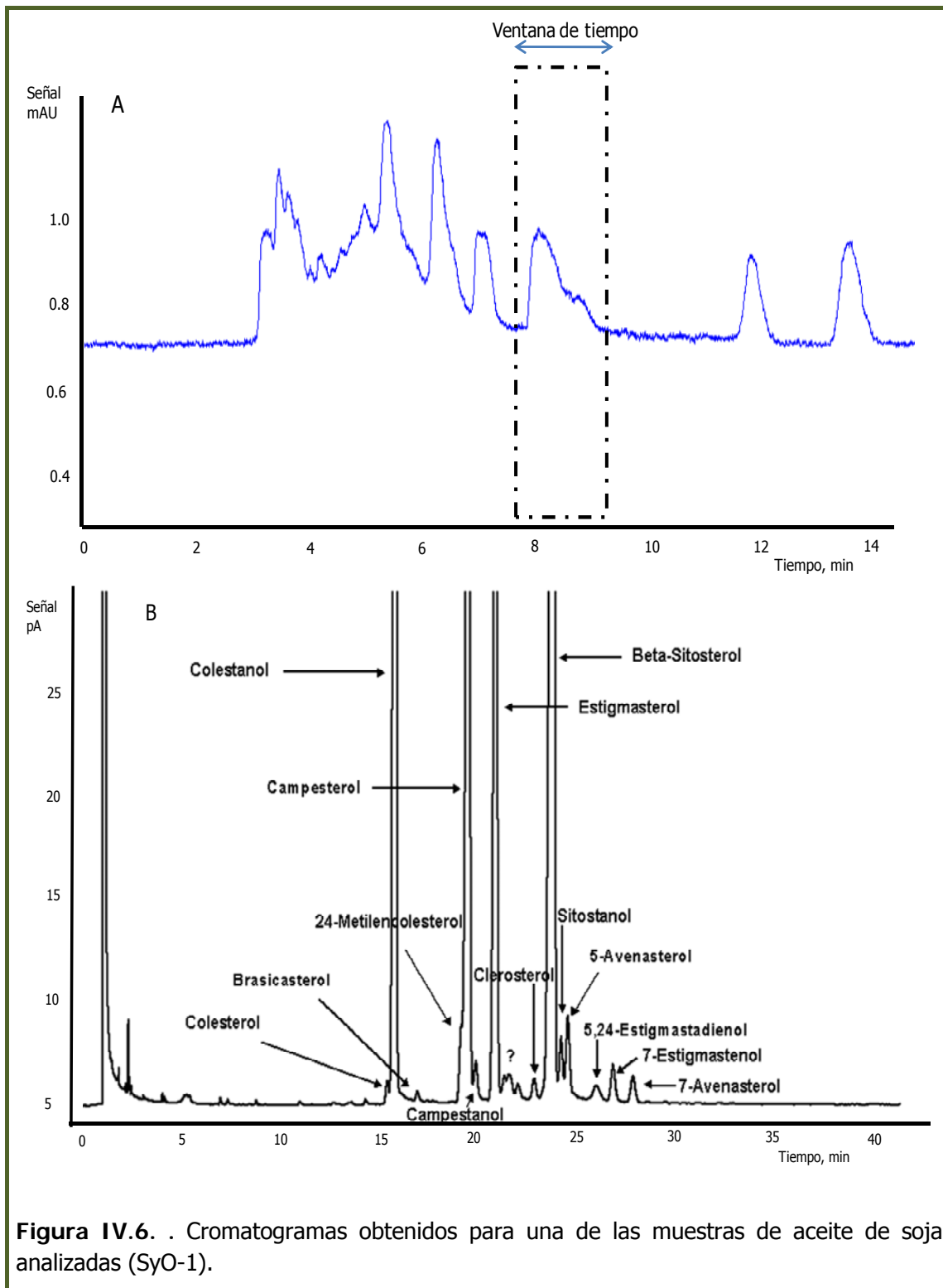
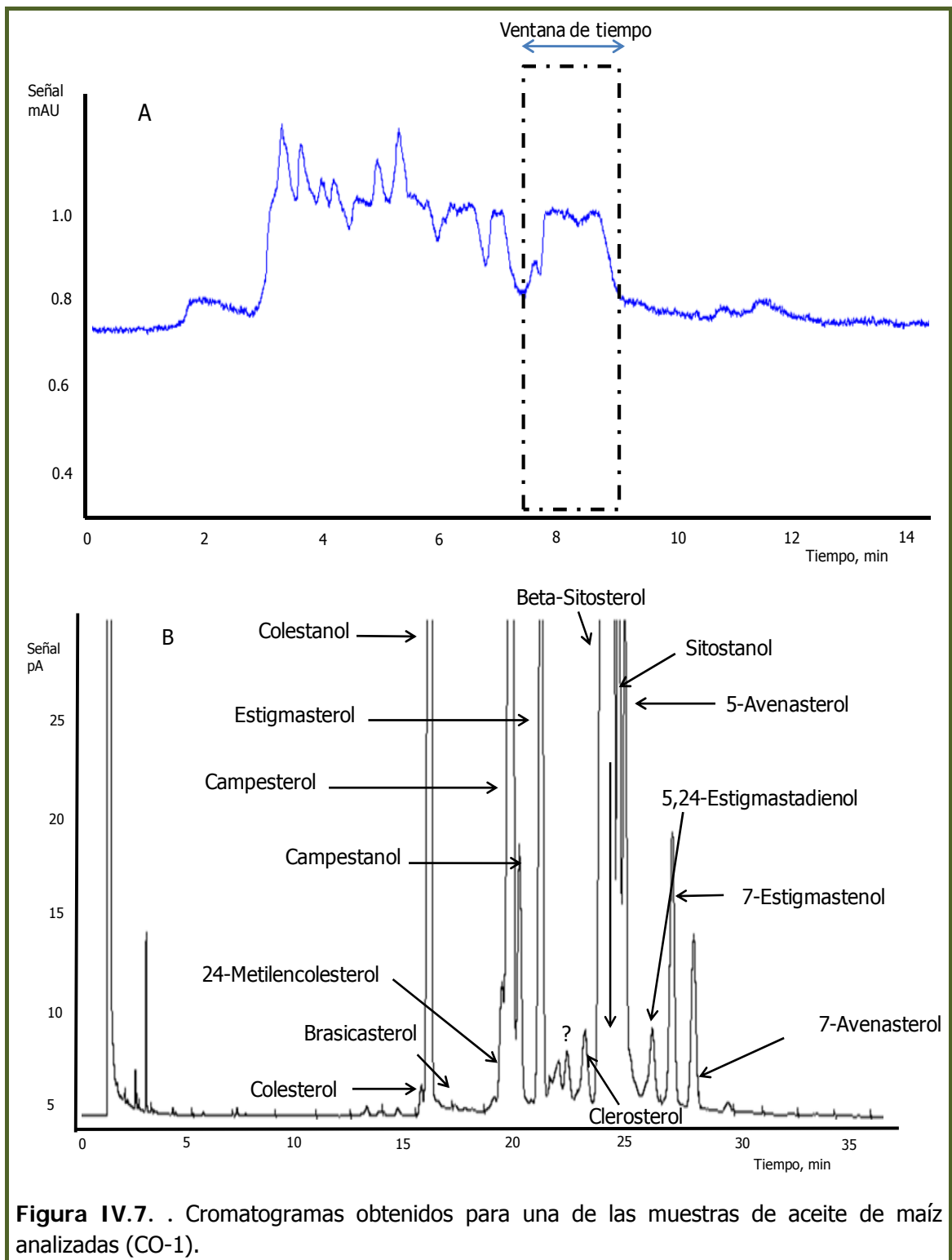


Figura IV.6. . Cromatogramas obtenidos para una de las muestras de aceite de soja analizadas (SyO-1).



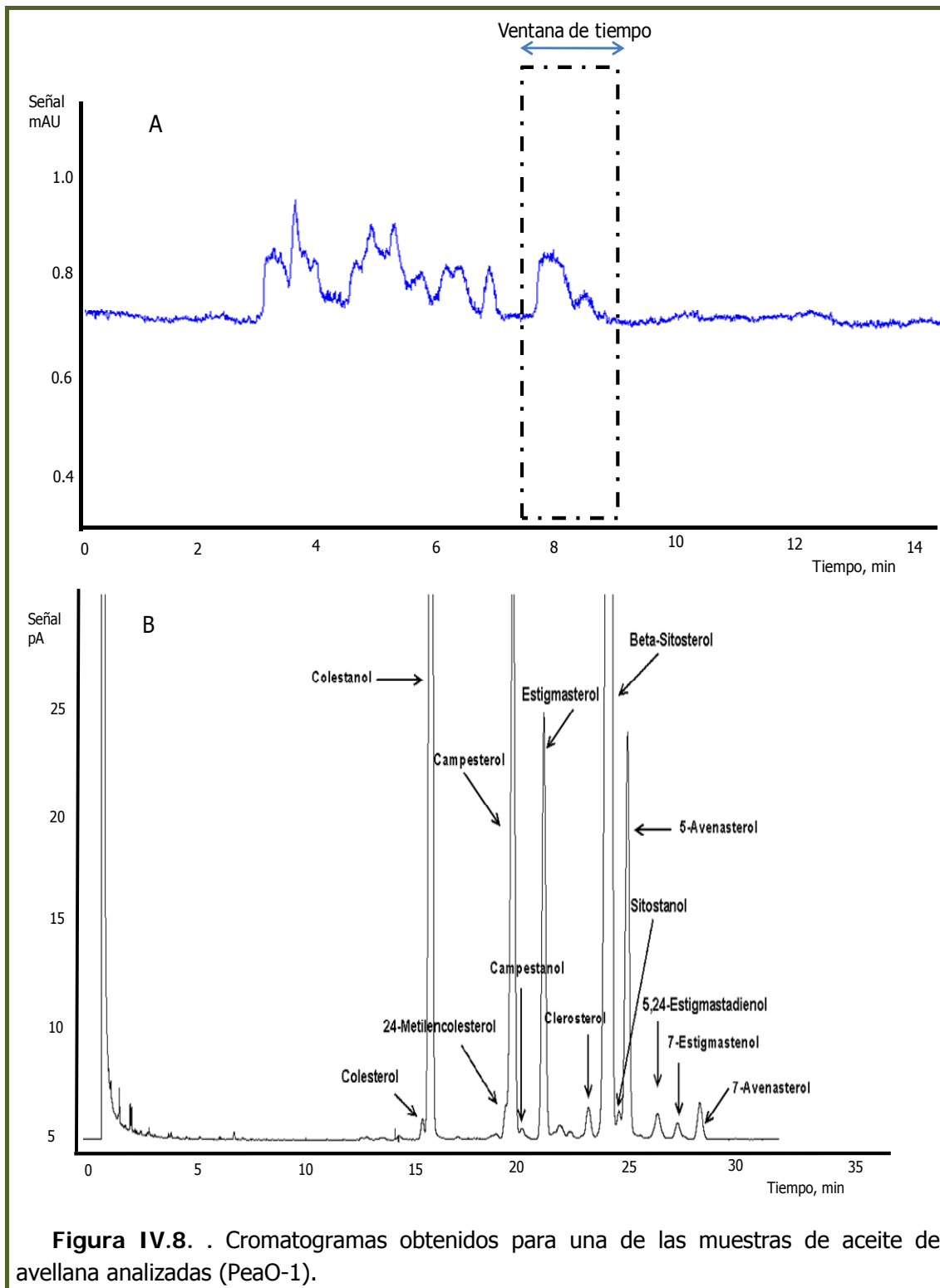
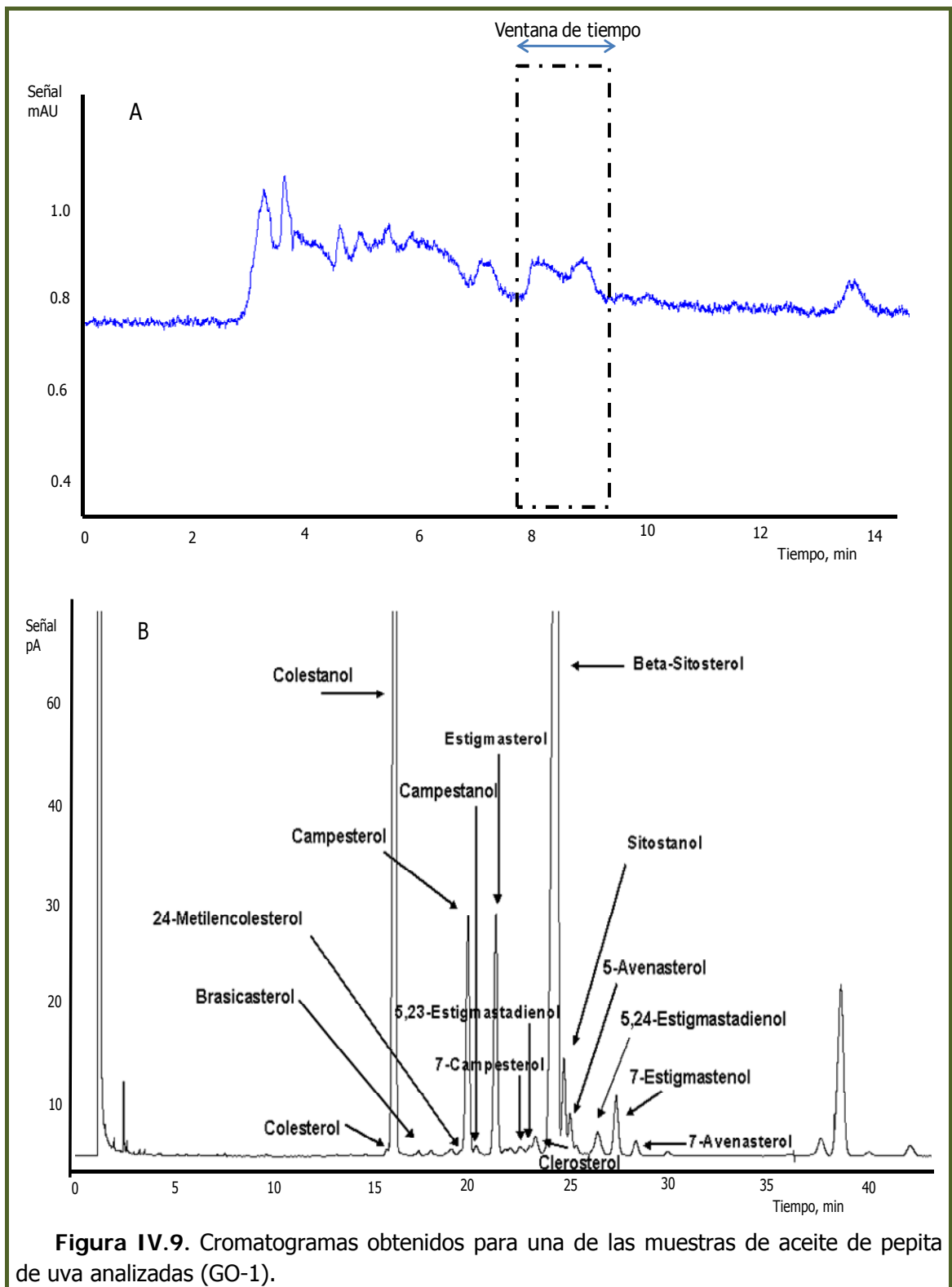
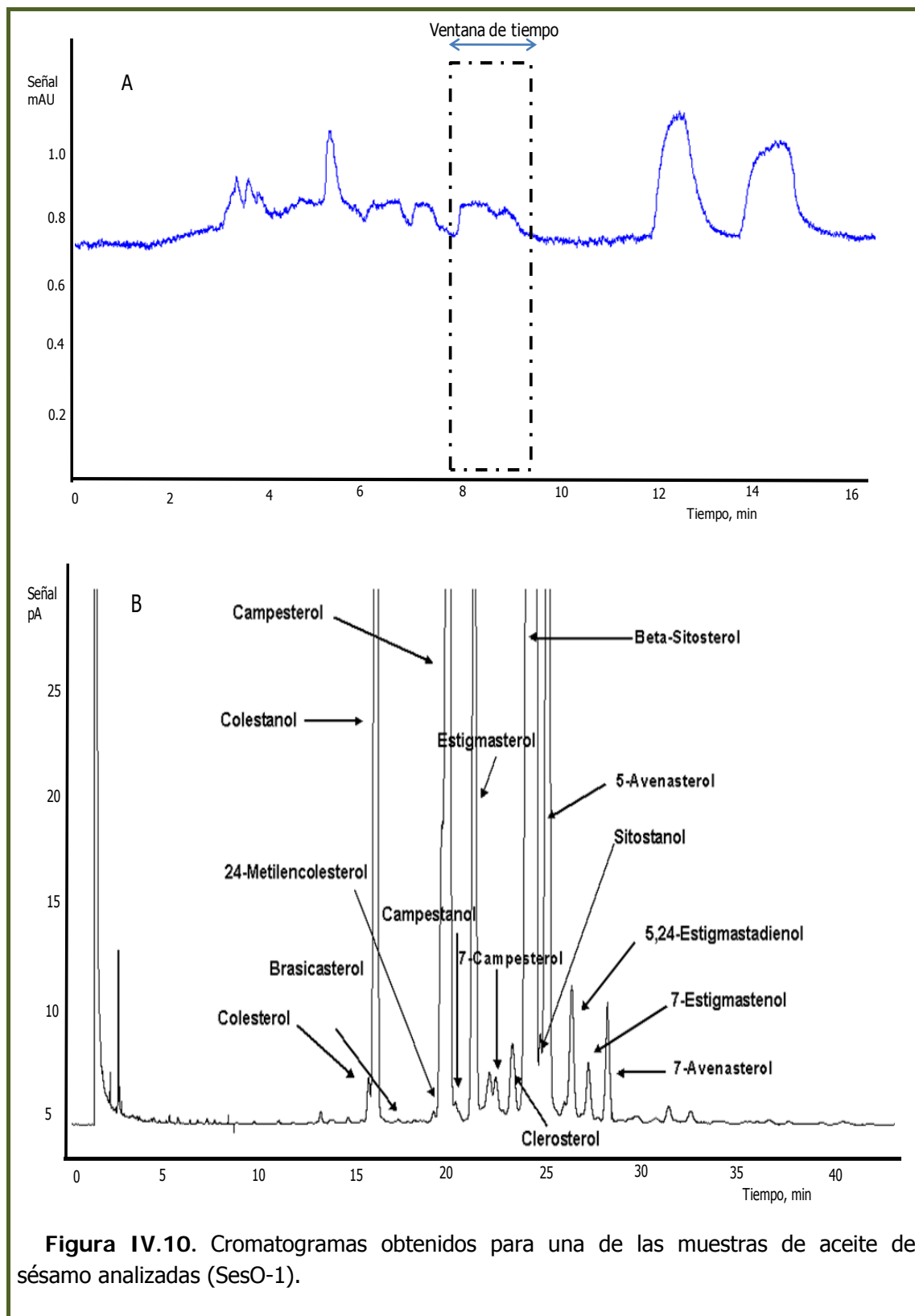


Figura IV.8. . Cromatogramas obtenidos para una de las muestras de aceite de avellana analizadas (PeaO-1).





Las concentraciones de 4-desmetilesteroles encontradas en las 24 muestras analizadas se recogen en la Tabla IV.2.

Tabla IV.2. Contenido de los 4-desmetilesteroles en aceites vegetales ± U (mg/kg).

Análitos	Oliva Virgen Extra			Oliva Refinado			Orujo de Oliva		
	EVOO-1	EVOO-27	EVOO-28	ROO-1	ROO-2	ROO-3	FOO-1	FOO-2	FOO-3
1	1.9±0.1	2.1±0.1	2.0±0.1	2.9±0.1	5.4±0.3	328.5±18.3	5.4±0.3	5.7±0.3	10.3±0.6
2	< L.O.D	2.1±0.1	< L.O.D	< L.O.D	2.3±0.1	5.7±0.3	2.3±0.1	5.7±0.3	3.9±0.2
3	1.8±0.1	< L.O.D	3.1±0.1	1.8±0.1	4.2±0.2	4.9±0.3	4.2±0.2	4.9±0.3	2.9±0.2
4	45.8±2.0	3.3±0.2	60.8±2.7	41.8±1.8	112.2±4.9	91.3±4.0	112.2±4.9	91.3±4.0	81.0±3.5
5	< L.O.D	48.9±2.1	1.2±0.1	1.5±0.1	6.7±0.4	4.5±0.3	6.7±0.4	4.5±0.3	3.6±0.2
6	8.0±0.4	2.2±0.1	7.5±0.3	11.2±0.5	30.6±1.4	28.8±1.3	30.6±1.4	28.8±1.3	37.7±1.7
7	< L.O.D	17.3±10.8	< L.O.D	< L.O.D	6.7±0.4	3.8±0.2	6.7±0.4	3.8±0.2	7.6±0.4
8	< L.O.D	< L.O.D	< L.O.D	< L.O.D	26.5±1.4	25.7±1.4	26.5±1.4	25.7±1.4	7.6±0.4
9	13.9±0.6	< L.O.D	11.4±0.5	12.6±0.5	33.9±2.0	30.6±1.8	33.9±2.0	30.6±1.8	26.9±1.6
10	1253.1±55.2	13.2±0.6	895.4±39.4	936.8±41.2	2004.4±87.2	1632.9±71.1	2004.4±87.2	1632.9±71.1	2178.0±94.8
11	13.0±0.5	1012.3±44.6	9.0±0.4	9.0±0.4	50.2±2.4	41.4±2.0	50.2±2.4	41.4±2.0	48.3±2.4
12	55.7±2.2	15.2±0.6	69.3±3.0	76.4±3.3	61.6±2.8	38.3±1.7	61.6±2.8	38.3±1.7	81.9±3.7
13	6.0±0.3	101.4±4.3	6.6±0.3	12.6±0.5	44.7±2.0	53.4±2.3	44.7±2.0	53.4±2.3	22.1±1.0
14	5.2±0.3	8.7±0.4	3.9±0.2	7.7±0.4	17.9±1.1	9.5±0.6	17.9±1.1	9.5±0.6	16.6±1.0
15	4.5±0.2	6.9±0.3	4.3±0.2	6.9±0.3	7.9±0.4	5.2±0.2	7.9±0.4	5.2±0.2	4.9±0.2

- Los analitos están codificados de acuerdo con la Tabla.III.2.

Tabla IV.2 (Cont.) . Contenido de los 4-desmetilesteroles en aceites vegetales ± U (mg/kg).

Analitos	Girasol		Girasol Alto Oleico		Colza		Canola		Soja	
	SFO-1	SFO(o)-1	SFO(o)-2	RO-1	RO-2	CanO-1	CanO-2	SyO-1	SyO-2	
1	3.4±0.2	2.4±0.1	4.1±0.2	24.2±1.3	27.8±1.5	26.4±1.5	26.8±1.5	9.0±0.5	11.0±0.6	
2	< L.O.D	1.0±0.1	1.8±0.1	713.2±37.5	843.0±44.3	703.4±37.0	711.5±37.4	6.7±0.4	7.6±0.4	
3	9.0±0.5	4.8±0.3	7.0±0.4	16.8±1.0	< L.O.D	57.4±3.4	< L.O.D	30.6±1.8	23.4±1.4	
4	275.8±12.0	189.2±8.2	247.1±10.7	2602.9±113.1	3074.4±133.6	2286.2±99.3	2632.2±114.4	581.2±25.3	524.6±22.8	
5	7.2±0.4	13.4±0.8	6.2±0.4	7.9±0.3	6.7±0.4	11.6±0.7	11.7±0.7	23.3±1.4	21.8±1.3	
6	239.0±10.6	161.0±7.1	217.8±9.6	22.0±1.0	28.0±1.2	30.8±1.4	32.5±1.4	437.1±19.3	473.1±20.9	
7	86.6±4.7	65.8±3.6	82.9±4.5	80.3±4.3	78.8±4.3	119.3±6.5	104.2±5.6	< L.O.D	21.6±1.2	
8	26.9±1.4	9.6±0.5	< L.O.D	22.2±1.2	19.8±1.1	46.4±2.5	38.6±2.1	< L.O.D	< L.O.D	
9	30.2±1.8	19.7±1.2	34.6±2.0	42.2±2.5	51.2±3.0	63.8±3.8	67.8±4.0	17.7±1.0	22.3±1.3	
10	1723.4±75.0	1093.1±47.6	1433.8±62.4	3385.3±147.3	4018.9±174.9	3610.6±157.1	3687.9±160.5	1075.5±46.8	1399.8±60.9	
11	22.5±1.1	34.8±1.7	19.6±1.0	12.3±0.6	12.3±0.6	12.4±0.6	10.8±0.5	33.7±1.6	43.8±2.1	
12	74.8±3.4	48.7±2.2	88.5±4.0	206.4±9.4	244.9±11.2	159.5±7.3	259.5±11.8	45.8±2.1	72.0±3.3	
13	79.2±3.5	39.8±1.7	47.1±2.1	51.7±2.3	59.2±2.6	105.1±4.6	86.5±3.8	20.9±0.9	30.2±1.3	
14	399.6±23.5	299.7±17.8	374.4±22.3	12.5±0.7	9.2±0.5	25.9±1.5	20.2±1.2	25.0±1.5	92.3±5.5	
15	141.0±6.7	111.9±5.3	151.4±7.2	7.4±0.4	7.0±0.3	11.2±0.5	12.6±0.6	17.6±0.8	52.6±2.5	

- Los analitos están codificados de acuerdo con la Tabla.III.2

Tabla IV.2 (Cont.). Contenido de los 4-desmetilesteroles en aceites vegetales ± U (mg/kg).

Análitos	Maíz		Cacahuete		Pepita de Uva		Sésamo	
	CO-1	CO-2	PeaO-1	PeaO-2	GO-1	GO-2	SesO-1	SesO-2
1	12.4±0.7	16.2±0.9	9.0±0.5	8.6±0.5	5.0±0.3	4.1±0.2	8.6±0.5	4.3±0.2
2	4.7±0.2	2.2±0.1	8.7±0.5	< L.O.D	4.7±0.2	5.6±0.3	3.5±0.2	1.9±0.1
3	64.8±3.8	81.5±4.8	10.3±0.6	14.0±0.8	4.9±0.3	5.5±0.3	114.4±6.7	96.6±5.7
4	1082.0±47.0	1289.9±56.0	332.9±14.5	354.0±15.4	192.8±8.4	224.6±9.8	927.0±40.3	912.0±39.6
5	144.9±8.6	81.4±4.8	7.1±0.4	7.6±0.5	11.3±0.7	10.5±0.6	10.3±0.6	15.3±0.9
6	431.6±19.1	483.4±21.4	189.6±8.4	214.5±9.5	195.7±8.6	216.3±9.6	358.6±15.8	382.9±16.9
7	< L.O.D	< L.O.D	5.0±0.3	6.8±0.4	14.2±0.8	20.2±1.1	26.8±1.5	29.0±1.6
8	< L.O.D	< L.O.D	< L.O.D	< L.O.D	9.0±0.5	8.7±0.5	< L.O.D	< L.O.D
9	83.5±4.9	63.2±3.7	20.1±1.2	22.1±1.3	24.2±1.4	24.7±1.4	53.0±3.1	53.8±3.2
10	4448.7±193.6	4258.3±185.3	1398.5±60.9	1430.9±62.3	1399.4±60.9	1591.2±69.2	3052.4±132.8	3133.7±136.4
11	359.5±17.5	187.8±9.2	15.1±0.7	14.3±0.7	82.5±4.0	73.4±3.6	20.0±1.0	16.4±0.8
12	292.6±13.3	291.9±13.3	164.0±8.4	222.6±10.1	38.2±1.7	45.6±2.1	663.6±30.2	679.8±31.0
13	85.0±3.7	47.5±2.1	20.7±0.9	23.1±1.0	39.5±1.7	37.7±1.7	75.7±3.3	80.8±3.6
14	191.7±11.4	50.8±3.0	13.5±0.8	14.7±0.9	63.4±3.8	89.1±5.3	35.1±2.1	62.9±3.7
15	119.2±5.7	58.5±2.8	23.4±2.1	24.8±1.2	17.6±0.8	28.2±1.3	60.3±2.9	70.6±3.4

- Los analitos están codificados de acuerdo con la Tabla.III.2



## IV.2. Estimación de los límites de detección atendiendo a la matriz en la que se encuentran los analitos

Hay que destacar, que cuando la concentración de un analito está por debajo del límite de detección (L.D.) en la práctica es habitual informar que los valores para esos datos son simplemente  $< \text{L.D.}$ . Para llevar a cabo el análisis quimiométrico y no obtener resultados sesgados, lo más habitual es sustituir estos datos no cuantificados por el valor numérico del límite de detección, determinado como parámetro de calidad del método o técnica instrumental. Algunas veces el valor que se introduce en la matriz de datos originales es la mitad de dicho límite.

Debido a ello, el valor del L.D. se ha calculado para cada tipo de aceite seleccionando una zona del cromatograma sin señal/picos atribuibles a los analitos, calculando el ruido (N) siendo la diferencia entre el máximo y mínimo de señal en altura. Las ecuaciones que recogen este cálculo se describen a continuación:

$$N = h_2 - h_1 \quad [\text{Ec. } 1]$$

Siendo

- $h_2$ : la altura máxima de la línea base del cromatograma.
- $h_1$ : la altura mínima de la línea base del cromatograma.

$$\text{L.D.} = \frac{3 \times N \times C_{\text{I.S.}}}{h_{\text{I.S.}}} \quad [\text{Ec. } 2]$$

Siendo

- N: Ruido estimado
- $h_{\text{I.S.}}$ : la altura del I.S. (5 $\alpha$ -colestanol)
- $C_{\text{I.S.}}$ : Concentración de patrón interno

Para obtener una estimación fiable se realizaron 9 determinaciones y se consideró como L.D. para cada tipo de aceite el valor medio obtenido.

A continuación, la Tabla IV.3 muestra los valores correspondientes de L.D. para cada tipo de aceite estudiado.

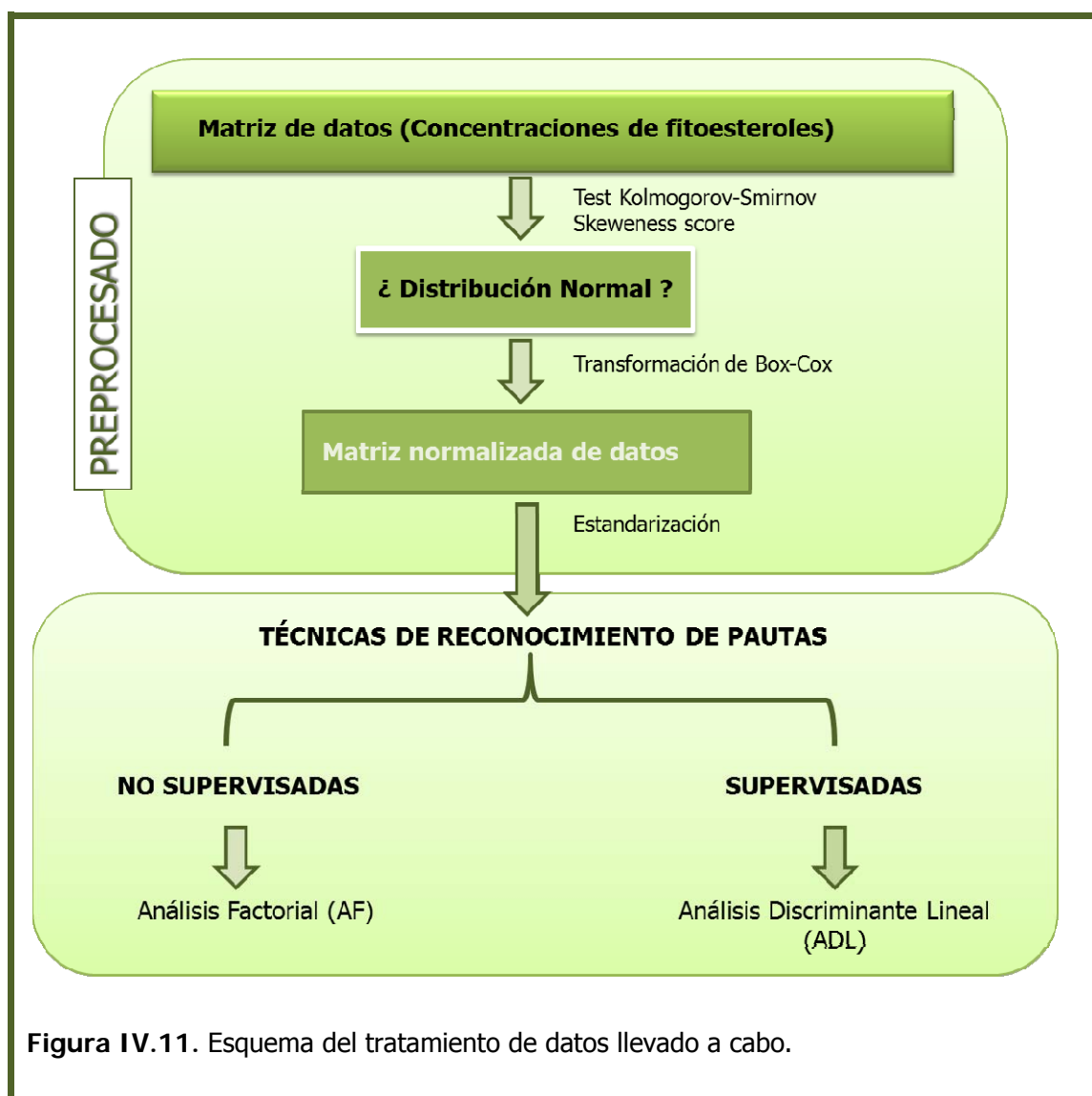
**Tabla IV.3.** Valores de límites de detección (L.D.)

Tipo de Muestra	Código	L.D.
Oliva Virgen Extra	EVOO-1	1.6
	EVOO-27	0.7
	EVOO-28	0.6
Oliva Refinado	ROO-1	0.4
Orujo de Oliva	POO-1	1.0
	POO-2	1.1
	POO-3	2.2
Girasol	SFO-1	2.1
Girasol Alto Oleico	SFO(o)-1	0.8
	SFO(o)-2	1.0
Colza	RO-1	2.4
	RO-2	1.0
Canola	CanO-1	1.2
	CanO-2	1.2
Soja	SyO-1	1.3
	SyO-2	0.9
Maíz	CO-1	0.9
	CO-2	1.0
Cacahuete	PeaO-1	0.9
	PeaO-2	1.1
Pepita de Uva	GO-1	0.9
	GO-2	0.8
Sésamo	SesO-1	0.9
	SesO-2	0.9

### IV.3. PROCEDIMIENTO SEGUIDO PARA EL ESTUDIO QUIMIOMÉTRICO DE LOS PERFIL DE CONCENTRACIÓN DE LOS 4-DESMETILESTEROLES DETERMINADOS EN ACEITES VEGETALES

El tratamiento estadístico que se realiza en este capítulo consiste en un análisis factorial (AF) como análisis exploratorio de datos de las variables implicadas en el estudio usando componentes principales con una rotación Varimax, la comprobación del ajuste a una distribución normal de los datos analizados (concentraciones), la normalización de los mismos mediante una transformación de potencias y un análisis multivariante lineal basado en la aplicación de métodos de

reconocimiento de pautas. En la Figura IV.11 se muestra un esquema resumen del tratamiento estadístico llevado a cabo.



El programa estadístico utilizado para el tratamiento de los datos ha sido el STATGRAPHICS Centurion XV Versión 15.2.05.

La dimensión de la matriz de datos originales de la que se partió en este estudio fue de (24x15), en la que como se ha indicado anteriormente, las concentraciones de los analitos no detectadas se han sustituido por la mitad del límite de detección estimado en el apartado anterior.

#### IV.4. ESTIMACIÓN DE LA NORMALIDAD DE LOS DATOS: TRANSFORMACIÓN DE LOS PERFILES DE CONCENTRACIÓN A UNA DISTRIBUCIÓN NORMAL MEDIANTE LA TRANSFORMACIÓN DE BOX-COX.

El tratamiento estadístico llevado a cabo con las concentraciones de los 4-desmetilesteroles recogidas en la Tabla IV.4 comenzó con un análisis exploratorio univariante (AED univariante). La Tabla IV.4 resume los resultados obtenidos e incluye medidas de tendencia central, de variabilidad, y de forma. En relación a esta última cabe indicar que los valores de sesgo y curtosis estandarizados encontrados indicaban que:

(i) atendiendo a la curtosis estandarizada (medida de la simetría de la distribución de datos y para una distribución normal su valor es cero), las concentraciones de colesterol, brasicasterol, 24-metilcolesterol, campesterol, campestanol,  $\Delta^7$ -campesterol,  $\Delta^{5,23}$ -estigmastadienol, sitostanol,  $\Delta^5$ -avenasterol,  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol no seguían una distribución normal.

(ii) atendiendo al sesgo estandarizado (medida de la falta de simetría o de forma equivalente del ajuste de los datos a una distribución normal), las concentraciones de colesterol, brasicasterol, campestanol, sitostanol,  $\Delta^5$ -avenasterol y  $\Delta^7$ -estigmastenol. ide el grado de simetría o la falta de esta

Adicionalmente, la aplicación del test de Kolmogorov-Smirnov y el puntuaciones de asimetría para evaluar el ajuste de los datos a una distribución normal para un nivel de confianza del 95% confirmó que no todos estaban normalizados por lo que tras el estudio estadístico por muestras (análisis estadístico por filas de la matriz de datos) y el análisis de los gráficos de rayos se procedió a la transformación de la matriz de datos en una matriz de datos normalizados, i.e. en una matriz que recogiera los perfiles de concentraciones de los 4-desmetilesteroles normalizados.

Tabla IV.4. Parámetros estadísticos para los 4-desmetilesteroles determinados en aceites vegetales.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Recuento	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
Promedio	23.18	126.58	23.55	758.78	17.48	177.29	31.92	10.53	34.84	2168.92	46.22	170.97	45.15	76.99	39.80
Desviación Estándar	65.56	282.55	33.38	938.02	31.44	169.43	38.91	13.62	20.27	1178.44	76.58	176.11	28.37	117.14	46.42
Coefficiente de Variación, %	282.89	223.22	141.73	123.62	179.85	95.56	121.91	129.26	58.17	54.33	158.82	103.01	62.86	152.14	116.64
Mínimo	1.88	0.39	0.99	41.81	1.16	7.53	0.39	0.39	11.43	895.43	8.98	38.15	5.99	3.93	4.27
Máximo	328.47	843.04	114.43	3074.44	144.86	483.36	119.28	46.39	83.51	4448.69	359.54	679.82	105.14	399.61	151.38
Rango	326.59	842.65	113.44	3032.63	143.70	475.83	118.89	46.00	72.08	3553.26	350.56	641.67	99.15	395.68	147.11
Seigo Estandarizado	9.54	3.89	3.40	2.96	7.02	1.22	2.079	2.66	1.75	1.49	6.84	4.12	0.87	4.05	2.78
Curtosis Estandarizada	23.12	2.01	1.78	0.96	12.77	-1.08	-0.43	0.83	-0.20	-1.00	12.65	4.12	-0.74	3.02	0.72

- Los analitos están codificados de acuerdo con la Tabla.III.2

#### IV.4.1. Análisis estadístico de los perfiles de concentraciones en función de los tipos de aceites analizados.

Este análisis se llevó a cabo para comparar la posible similitud de los perfiles de 4-desmetilesteroles entre las muestras de aceites vegetales analizados del mismo tipo (oliva virgen extra, oliva refinado, orujo de oliva, girasol, colza/cánola, soja, maíz, cacahuete, pepita de uva y sésamo).

La Figura IV.12 muestra el diagrama de dispersión obtenido para las veinticuatro muestras de aceites analizados. Se puede observar que independientemente del tipo de aceite hay un analito mayoritario en todas las muestras, el  $\beta$ -sitosterol que además desde un punto de vista estadístico se puede considerar como dato anómalo porque dada su elevada concentración, "nivela" las concentraciones del resto de los analitos en los aceites estudiados.

Además en dicho diagrama se puede observar que aparece una clara diferencia entre los aceites de oliva y girasol y el resto de aceites estudiados. En el primer grupo el único fitoesterol claramente mayoritario es el  $\beta$ -sitosterol, mientras que en el segundo grupo además de este también se observan como mayoritarios otros como el campesterol y  $\Delta^5$ -avenasterol.

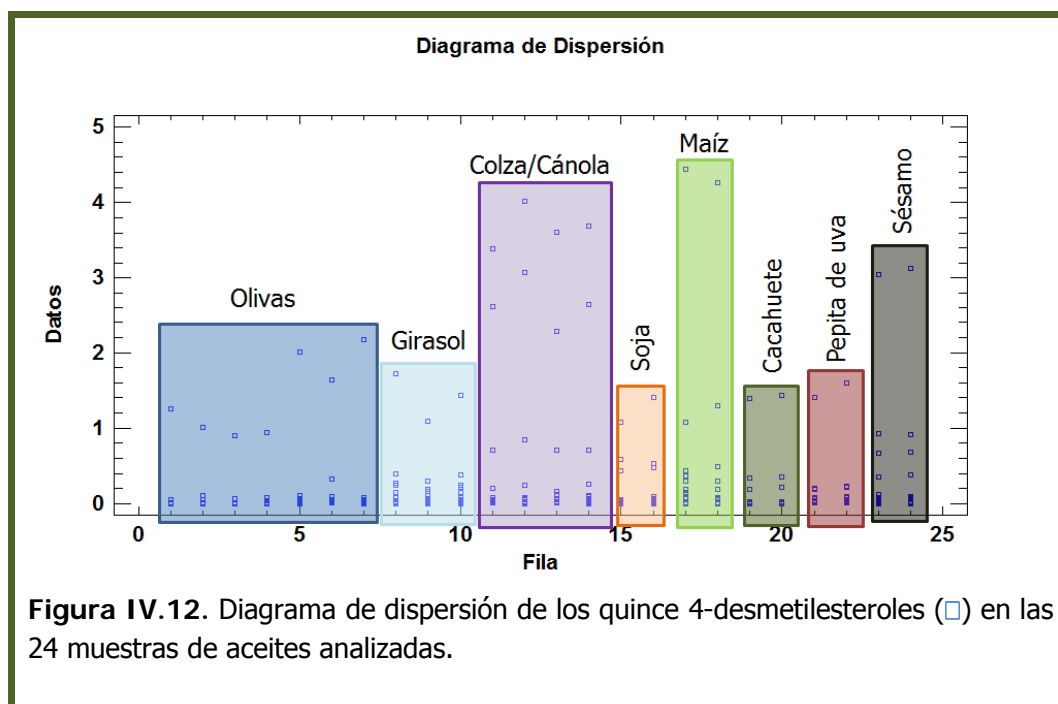
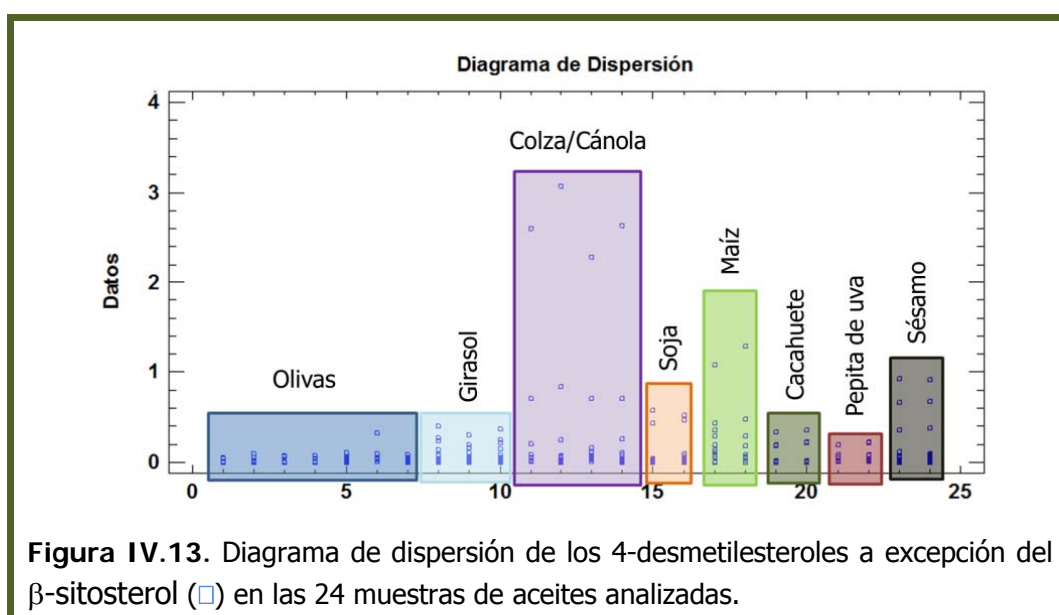


Figura IV.12. Diagrama de dispersión de los quince 4-desmetilesteroles (□) en las 24 muestras de aceites analizadas.

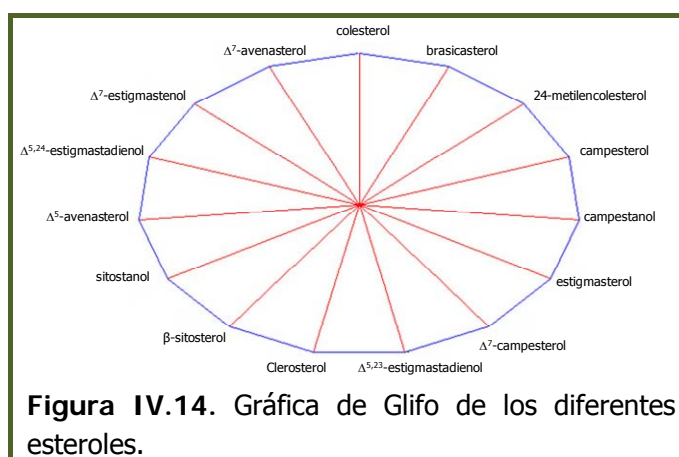
Con objeto de intentar evaluar si existen diferencias en el perfil de concentraciones de los otros fitoesteroles, se decidió evaluar el diagrama de dispersión de las muestras de aceites eliminando la concentración de  $\beta$ -sitosterol. Los resultados obtenidos se muestran en la Figura IV.13.

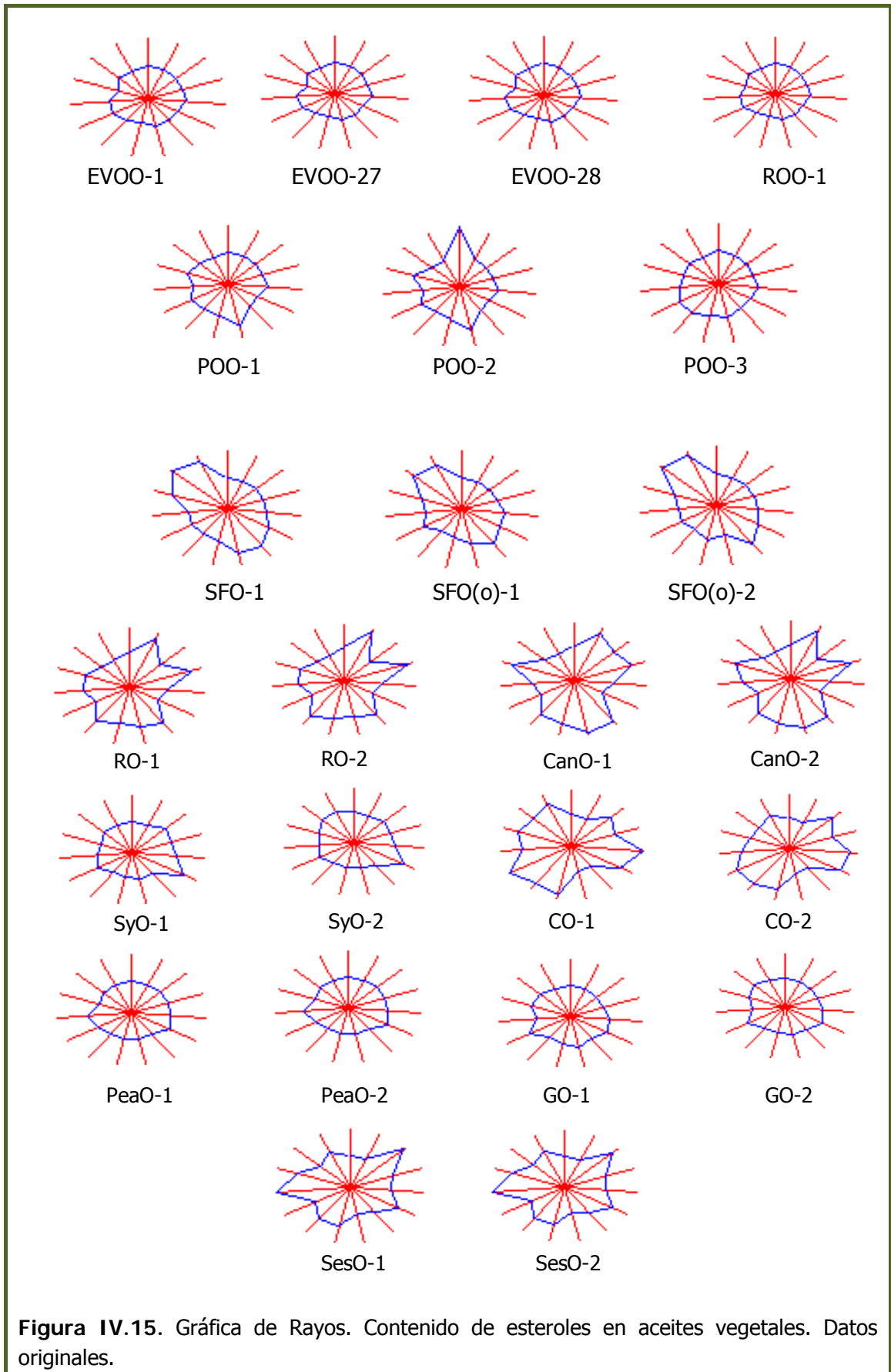
De forma cualitativa se puede observar que atendiendo a la dispersión de los resultados los tipos de aceites analizados muestran un comportamiento ligeramente diferente siendo similar para las mayoría de las muestras de aceites del mismo tipo, si bien es verdad que a priori las muestras de los diferentes tipos de aceite de oliva (virgen extra, refinado y orujo) no se diferencian mucho entre sí.



#### IV.4.2. Análisis de las gráficas de rayos obtenidas para los perfiles de concentraciones en función de los tipos de aceites analizados

...Con objeto de evaluar de forma rápida y visual los perfiles de concentraciones individuales obtenidos la Figura IV.15 muestra el gráfico de rayos obtenido del análisis estadístico. Para su mejor entendimiento, la Figura IV.14 recoge el gráfico de rayos utilizado





**Figura IV.15.** Gráfica de Rayos. Contenido de esteroles en aceites vegetales. Datos originales.



Observando detenidamente la gráfica de rayos se confirma que:

- ✓ Los aceites de oliva virgen extra y oliva refinado son los aceites que menos contenido en esteroides totales presenta, manteniendo los tres últimos esteroides ( $\Delta^{5,24}$ -estigmastadienol,  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol) con la misma concentración.
- ✓ El aceite de orujo de oliva es el aceite que contiene todos los desmetilesteroides. El aceite de girasol se caracteriza por tener alto contenido en  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol. Además presenta alto contenido al igual que en el aceite de cáñola y colza del  $\Delta^7$ -campesterol.
- ✓ El aceite de soja contiene alto contenido en estigmasterol.
- ✓ Los aceites de cáñola y colza se caracterizan por presentar un alto contenido en brasicasterol a diferencia del resto de aceites que lo presentan a nivel de trazas.
- ✓ También presentan un alto contenido en campesterol,  $\Delta^7$ -campesterol y  $\Delta^{5,23}$ -estigmastadienol. En aceite de cáñola el  $\Delta^{5,24}$ -estigmastadienol es mayoritario.
- ✓ El aceite de maíz entre todos los aceites analizados es el que mayor contenido en sitostanol, campestanol, clerosterol y  $\beta$ -sitosterol presenta, siendo este último esteroide el más bajo en aceite de oliva.
- ✓ El aceite de sésamo es el aceite que mayor contenido en 24-metilcolesterol y  $\Delta^5$ -avenasterol tiene.
- ✓ Los aceites de cacahuete y uva tienen un contenido en esteroides similar con la diferencia de que el aceite de uva tiene más contenido en sitostanol,  $\Delta^7$ -estigmastenol y  $\Delta^{5,23}$ -estigmastadienol, y el cacahuete en  $\Delta^5$ -avenasterol y campesterol.

### IV.4.3. Normalización de los perfiles de concentraciones de los 4-desmetilesteroides

Entre las transformaciones descritas en bibliografía, la transformación de Box-Cox ha sido usada ampliamente para transformar un conjunto o matriz de datos en

otro que siga una distribución normal. El procedimiento de transformación fue diseñado para originar una transformación óptima que ajustara una variable Y a un modelo lineal:

$$W = \beta_0 + \beta_1 X + \varepsilon \quad [\text{Ec. } 3]$$

Donde:

- W es la variable normalizada o transformada,
- X es la variable de regresión
- $\beta_0$  y  $\beta_1$  son los parámetros desconocidos, y
- $\varepsilon$  es el error.

La variable dependiente W se relaciona con la variable Y (en este caso la concentración de los 4-desmetilesteroles) de acuerdo con las siguientes ecuaciones:

$$W = 1 + \frac{(Y + \lambda_2)^{\lambda_1} - 1}{\lambda_1 G^{\lambda_1 - 1}} \quad \text{for } \lambda_1 \neq 0 \quad [\text{Ec. } 4]$$

$$W = G \ln(Y + \lambda_2) \quad \text{for } \lambda_1 = 0 \quad [\text{Eq. } 5]$$

Donde:

- $\lambda_1$  y  $\lambda_2$  son los parámetros de la transformación
- G es la media geométrica de  $(Y + \lambda_2)$ .

Considerándose que la transformación óptima es la que minimiza el error cuadrático medio (MSE) W.

Dado que la mayoría de las concentraciones no seguían una distribución normal se decidió aplicar esta transformación para normalizarlas. Los valores obtenidos para los parámetros de la transformación se recogen en la Tabla IV.5 en la que también se recoge la disminución provocada en el MSE de cada analito y los valores del test de Shapiro-Wilkis aplicado a los datos transformados para verificar el ajuste de los mismos a una distribución normal.

**Tabla IV.5.** Parámetros óptimos obtenidos para aplicar la transformación de Box-Cox a las concentraciones de fitoesteroles en las muestras analizadas.

Analito <sup>(1)</sup>	Parámetros óptimos de la transformación Box-Cox	Disminución MSE(%)	¿normalidad?
(1)	$\lambda_1 = -0.345$ $\lambda_2 = 0.0$ G=8.163	98.33	$W_{(Shapiro-Wilks)}: 0.967$ (p=0.605)
(2)	$\lambda_1 = -0.297$ $\lambda_2 = 0.0$ G=6.252	99.83	$W_{(Shapiro-Wilks)}: 0.943$ (p=0.195)
(3)	$\lambda_1 = -0.148$ $\lambda_2 = 0.0$ G=9.120	85.34	$W_{(Shapiro-Wilks)}: 0.959$ (p=0.427)
(4)	$\lambda_1 = -0.033$ $\lambda_2 = 0.0$ G=339.204	75.30	$W_{(Shapiro-Wilks)}: 0.947$ (p=0.235)
(5)	$\lambda_1 = -0.134$ $\lambda_2 = 0.0$ G=8.445	91.24	$W_{(Shapiro-Wilks)}: 0.952$ (p=0.313)
(6)	$\lambda_1 = 0.163$ $\lambda_2 = 0.0$ G=85.013	50.05	$W_{(Shapiro-Wilks)}: 0.880$ (p=0.007)
(7)	$\lambda_1 = 0.082$ $\lambda_2 = 0.0$ G=9.581	78.67	$W_{(Shapiro-Wilks)}: 0.920$ (p=0.058)
(8)	$\lambda_1 = -0.167$ $\lambda_2 = 0.0$ G=3.514	82.89	$W_{(Shapiro-Wilks)}: 0.859$ (p=0.003)

<sup>(1)</sup> Los analitos están codificados de acuerdo con la Tabla III.2.

**Tabla IV.5 (Cont.).** Parámetros óptimos obtenidos para aplicar la transformación de Box-Cox a las concentraciones de fitoesteroles en las muestras analizadas.

Analito <sup>(1)</sup>	Parámetros óptimos de la transformación Box-Cox	Disminución MSE(%)	¿normalidad?
(9)	$\lambda_1 = -0.083$ $\lambda_2 = 0.0$ G=29.687	27.50	$W_{(Shapiro-Wilks)}: 0.960$ (p=0.455)
(10)	$\lambda_1 = -0.431$ $\lambda_2 = 0.0$ G=1893.49	30.72	$W_{(Shapiro-Wilks)}: 0.923$ (p=0.069)
(11)	$\lambda_1 = 0.566$ $\lambda_2 = 0.0$ G=26.932	91.30	$W_{(Shapiro-Wilks)}: 0.954$ (p=0.340)
(12)	$\lambda_1 = -0.380$ $\lambda_2 = 0.0$ G=116.017	70.31	$W_{(Shapiro-Wilks)}: 0.940$ (p=0.168)
(13)	$\lambda_1 = 0.495$ $\lambda_2 = 0.0$ G=34.867	12.32	$W_{(Shapiro-Wilks)}: 0.962$ (p=0.487)
(14)	$\lambda_1 = -0.21$ $\lambda_2 = 0.0$ G=30.628	88.28	$W_{(Shapiro-Wilks)}: 0.968$ (p=0.623)
(15)	$\lambda_1 = -0.194$ $\lambda_2 = 0.0$ G=20.236	73.16	$W_{(Shapiro-Wilks)}: 0.919$ (p=0.056)

<sup>(1)</sup> Los analitos están codificados de acuerdo con la Tabla III.2.

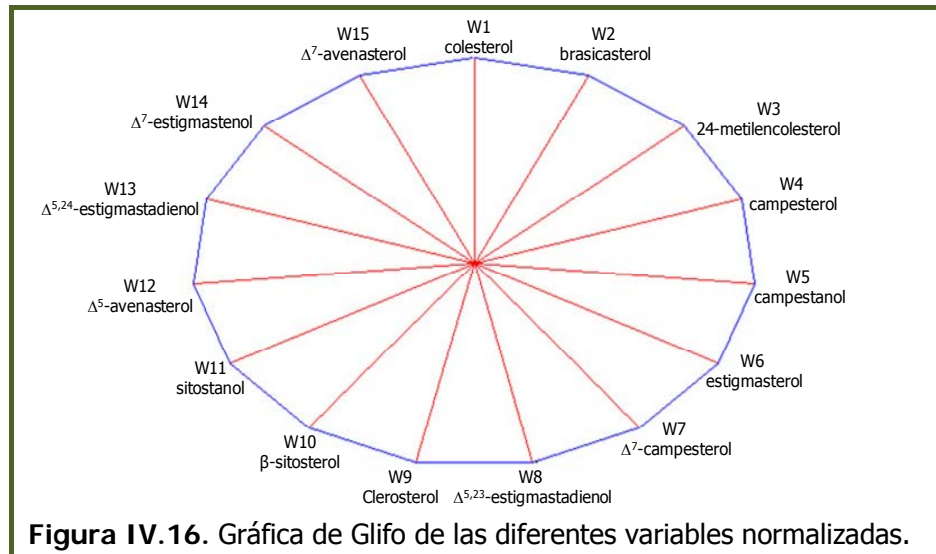
La Tabla IV.6 recoge los valores de las nuevas concentraciones normalizadas de los 4-desmetilesteroles (representadas como  $W_i$ )

Tabla IV.6. Valores normalizados mediante la transformación de potencia de BOX-COX para las variables estudiadas.

Tipode Muestra	Código	W <sub>1</sub>	W <sub>2</sub>	W <sub>3</sub>	W <sub>4</sub>	W <sub>5</sub>	W <sub>6</sub>	W <sub>7</sub>	W <sub>8</sub>	W <sub>9</sub>	W <sub>10</sub>	W <sub>11</sub>	W <sub>12</sub>	W <sub>13</sub>	W <sub>14</sub>	W <sub>15</sub>
Oliva Virgen Extra	EVO-1	10.70	5.97	8.12	1478.07	6.39	73.68	4.87	3.06	94.00	103328.15	236.03	1456.41	8.14	88.56	48.31
	EVO-27	12.03	-2.85	14.84	1501.77	9.41	94.96	-1.61	-0.59	92.36	107822.51	242.11	1538.56	8.98	100.77	56.31
	EVO-28	11.38	-4.94	14.17	1580.25	3.02	71.77	-3.15	-1.60	87.68	107509.98	219.40	1488.56	8.37	75.38	47.07
Oliva Refinado	ROO-1	16.01	-10.35	8.12	1444.90	5.44	83.29	-6.58	-3.29	90.88	107627.02	219.40	1502.07	9.68	105.31	59.42
	POO-1	22.54	8.95	17.35	1798.00	19.87	109.06	15.02	11.94	121.17	102990.10	274.44	1471.56	11.29	136.95	62.75
	POO-2	43.21	15.95	18.91	1725.23	16.31	107.62	11.07	11.86	118.15	108894.13	270.58	1394.70	11.45	113.72	52.17
Orujo de Oliva	POO-3	27.99	13.06	13.45	1682.75	14.23	113.90	15.87	8.46	114.31	109440.85	273.70	1511.40	10.82	134.35	50.60
	SFO-1	17.81	8.18	24.72	2109.77	20.50	150.27	30.74	11.98	117.76	109001.75	255.21	1499.18	11.75	215.17	116.37
	SFO(o)-1	13.73	1.00	18.70	1980.26	25.64	143.38	29.21	9.16	104.87	108009.85	266.73	1435.29	11.18	209.87	113.08
Olivo	SFO(o)-2	19.82	6.81	22.38	2072.18	19.19	148.69	30.50	0.01	121.77	108624.21	250.93	1521.51	11.34	214.00	117.35
	RO-1	33.56	32.13	30.16	2848.70	21.29	101.06	30.33	11.49	127.54	110156.85	233.74	1614.62	11.42	124.17	61.15
Cela	RO-2	34.32	32.38	0.83	2901.35	19.87	106.95	30.22	11.19	133.07	110400.62	233.74	1630.05	11.53	112.46	59.78
	CanO-1	34.04	32.11	39.52	2807.46	24.48	109.21	32.49	13.28	139.26	110250.52	234.08	1589.38	11.93	149.14	70.95
Canoila	CanO-2	34.12	32.12	3.27	2852.25	24.55	110.47	31.76	12.85	140.95	110280.75	228.08	1635.06	11.81	141.04	73.59
	SyO-1	26.95	16.66	34.94	2361.25	29.94	159.97	3.07	2.11	101.56	107970.76	265.97	1425.26	10.45	148.01	80.77
Soja	SyO-2	28.48	17.42	32.85	2327.05	29.38	161.17	22.62	0.54	108.65	108572.71	271.75	1493.91	10.90	184.53	101.28
	CO-1	29.34	14.37	40.35	2566.23	41.82	159.78	0.16	0.54	146.67	110536.71	296.91	1645.14	11.80	200.99	114.00
Maiz	CO-2	31.14	8.57	41.89	2623.44	38.36	161.50	1.00	1.00	138.99	110478.85	292.03	1644.89	11.35	169.07	103.05
	PeaO-1	26.95	18.20	25.94	2173.84	20.38	146.28	13.00	0.54	105.48	108570.71	241.87	1603.67	10.43	126.99	86.53
Cacahuete	PeaO-2	26.58	2.01	28.62	2194.68	20.96	148.43	15.12	1.41	108.38	108619.88	239.80	1621.56	10.58	130.06	87.87
	GO-1	21.80	14.37	18.91	1986.77	24.27	146.84	19.98	8.97	111.13	108572.10	282.64	1394.24	11.17	175.03	80.77
Pepita de Uva	GO-2	19.82	15.53	20.05	2039.40	23.67	148.57	22.21	8.87	111.13	108841.63	280.91	1424.54	11.13	183.66	90.14
	SeoO-1	26.58	12.27	44.07	2515.63	23.51	156.89	23.95	0.54	134.05	110000.81	251.57	1702.53	11.72	158.48	103.55
Sésamo	SeoO-2	20.31	7.30	43.00	2510.27	26.68	157.92	24.42	0.54	134.47	110041.10	244.88	1703.96	11.76	174.82	106.09

Con objeto de verificar que la transformación aplicada no modificaba los perfiles de concentración de los 4-desmetil esteroides se procedió a estudiar visualmente el gráfico de rayos de cada una de las muestras. Los gráficos individuales se muestran en la Figura IV.17.

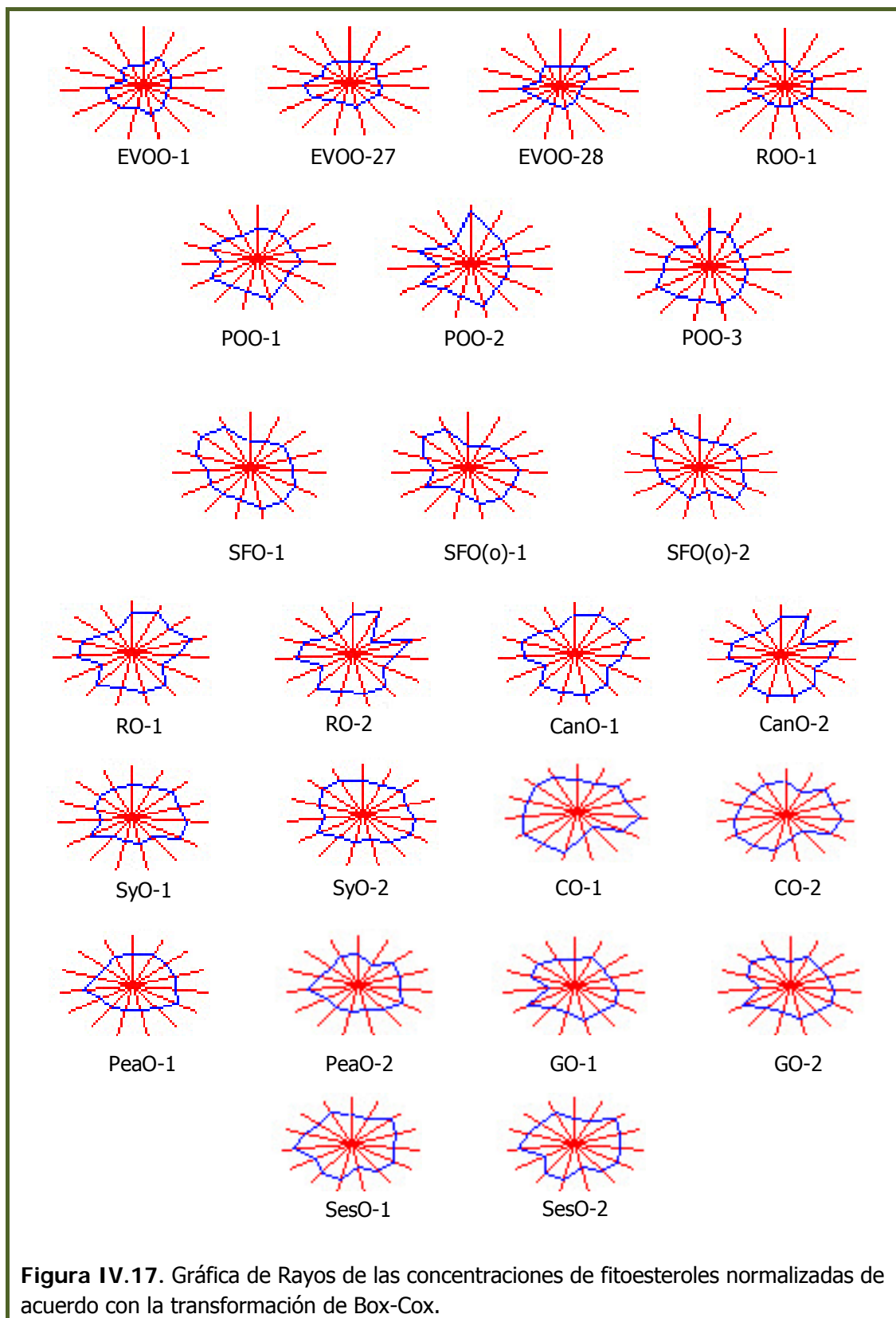
La siguiente figura (Figura IV.16) muestra el glifo asociado al gráfico de rayos.



**Figura IV.16.** Gráfica de Glifo de las diferentes variables normalizadas.

Como se puede observar los perfiles normalizados mantienen la misma tendencia individual que los valores no normalizados por lo que tras el estudio de los mismos se pueden destacar los siguientes hechos:

- ✓ Entre todos los aceites estudiados el POO-2 mostraba la mayor concentración de colesterol.
- ✓ El perfil de concentraciones de los 4-desmetilesteroides en las muestras de aceite de oliva virgen (EVOO<sub>i</sub>) y de oliva refinado ROO-1 era muy similar exceptuando el caso de la concentración del  $\Delta^{5,24}$ estigmastadienol que es mayor en el oliva refinado.
- ✓ El perfil de concentraciones de los 4-desmetilesteroides de las muestras de orujo de oliva es diferente a los perfiles de las muestras de aceites de oliva extra virgen y refinado tanto en las concentraciones encontradas como en el contenido de  $\Delta^{5,23}$ estigmastadienol. Esa misma diferencia se ha encontrado en dos de las muestras de orujo analizadas (POO-1 y POO-3).



- ✓ Los Perfiles de concentración encontrados en las muestras de aceites de cacahuete (PeaO<sub>i</sub>) y pepita de uva (GO<sub>i</sub>) son ligeramente superiores a los mostrados por los aceites de oliva, especialmente para el  $\Delta^5$ avenasterol en los de cacahuete y sitostanol en los aceites de pepita de uva.
- ✓ Los Perfiles de concentración más diferentes se encontraron en los siguientes tipos de aceite: (i) soja (SyO<sub>i</sub>) con alto contenido en estigmasterol, (ii) colza (RO<sub>i</sub>) y cáñola (CanO<sub>i</sub>) con valores altos de brasicasterol y campesterol. Existían evidencias de que estos tipos de aceites podrían diferenciarse atendiendo a su contenido en their  $\Delta^{5,24}$ estigmastadienol, (iii) Una muestral de aceite de girasol (SFO-3) con valores muy altos de  $\Delta^7$ estigmastenol y  $\Delta^7$ avenasterol, (iv) los aceites de maíz (CO<sub>i</sub>) con las mayores concentraciones de, y finalmente, (v) los aceites de sésamo (SesO<sub>i</sub>) con las mayores concentraciones de  $\Delta^5$ avenasterol y 24-metilen-colesterol.

#### **IV.5. TÉCNICAS NO SUPERVISADAS DE RECONOCIMIENTO DE PAUTAS: ANÁLISIS FACTORIAL DEL PERFIL DE CONCENTRACIONES NORMALIZADAS.**

Este análisis multivariante se aplicó a los datos normalizados (15 variables) con la finalidad de comparar la composición de los aceites analizados e identificar los factores que influyen en cada uno de ellos.

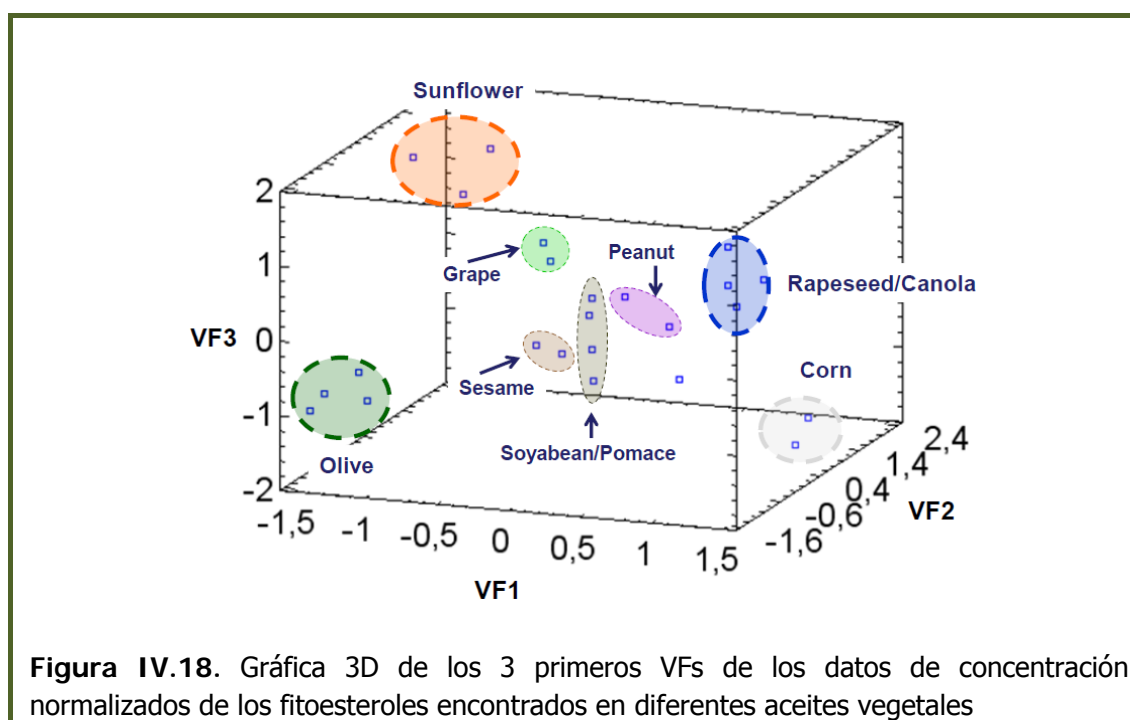
El análisis factorial se aplicó sobre las componentes principales y se utilizó como método de rotación, la rotación VARIMAX, que redistribuye la varianza para obtener un patrón de factores con mayor significado. Este criterio se centra en simplificar al máximo los vectores columna de la matriz de factores, su interés radica en que permite interpretar los factores más fácilmente, al indicar una asociación positiva o negativa clara entre la variable y el factor. Esta rotación fue fijada según los ejes definidos por los componentes principales tomando como criterio que el autovalor sea  $> 1$ . Los resultados indicaron que cuatro varifactores (VFs) eran adecuados para explicar el 89.48% de la varianza total.

El análisis de las contribuciones de las distintas variables normalizadas sobre cada uno de los varifactores permitió establecer que:



- a) El VF1 estaba influenciado positivamente por las concentraciones normalizadas del colesterol, brasicasterol, clerosterol y  $\beta$ -sitosterol, y negativamente por las de  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol.
- b) El VF2 estaba influenciado positivamente por las concentraciones de estigmasterol,  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol.
- c) El VF3 estaba influenciado positivamente por las concentraciones normalizadas de estigmasterol,  $\Delta^{5,23}$ -estigmastadienol,  $\Delta^{5,24}$ -estigmastadienol y  $\Delta^7$ -estigmastenol, y negativamente por el 24-metilencolesterol y sitostanol.
- d) El VF4 estaba influenciado positivamente por las concentraciones normalizadas de campesterol y  $\Delta^5$ -avenasterol y negativamente por  $\Delta^{5,23}$ -estigmastadienol y sitostanol.

La Figura IV.18 muestra el gráfico 3D para la dispersión de las muestras estudiadas y los tres primeros varifactores



A la vista de la figura se puede observar que de los nueve tipos de aceites iniciales, atendiendo a su agrupamiento natural solo se podrían distinguir claramente entre cuatro tipos de aceites:

- (1) Grupo I, constituido por las muestras de aceite de oliva (virgen extra y oliva refinado),
- (2) Grupo II, constituido por las muestras de aceite de colza y cáñola,
- (3) Grupo III, constituido por las muestras de aceite de maíz, y
- (4) Grupo IV, constituido por las muestras de aceite de girasol (tanto normal como en alto contenido en ácido oleico).

Podría considerarse también la existencia de un Grupo V, caracterizado por perfiles de concentración intermedios constituido por el resto de los aceites, es decir por aceites de sésamo, soja, orujo de oliva, cacahuete, y pepita de uva.

La Figura IV.19 muestra las puntuaciones obtenidas para cada varifactor en las muestras analizadas.

Se puede observar que el Grupo I (aceites de oliva virgen extra y refinado) se podría caracterizar por valores negativos de los varifactores VF1 a VF3, (baja concentración de  $\Delta^7$ -estigmastenol,  $\Delta^7$ -avenasterol, 24-metilcolesterol y sitostanol) y valores positivos del VF4 (alta concentración de campesterol). El aceite de orujo de oliva, excluido de forma natural de este grupo solo se podría diferenciar de los otros tipos de oliva por los varifactores VF1 y VF4.

Por otro lado las diferencias encontradas en los grupos II y IV se deberían principalmente a los varifactores VF1 y VF3.

El Grupo III se diferenciaría por las puntuaciones de los varifactores VF2 y VF3.

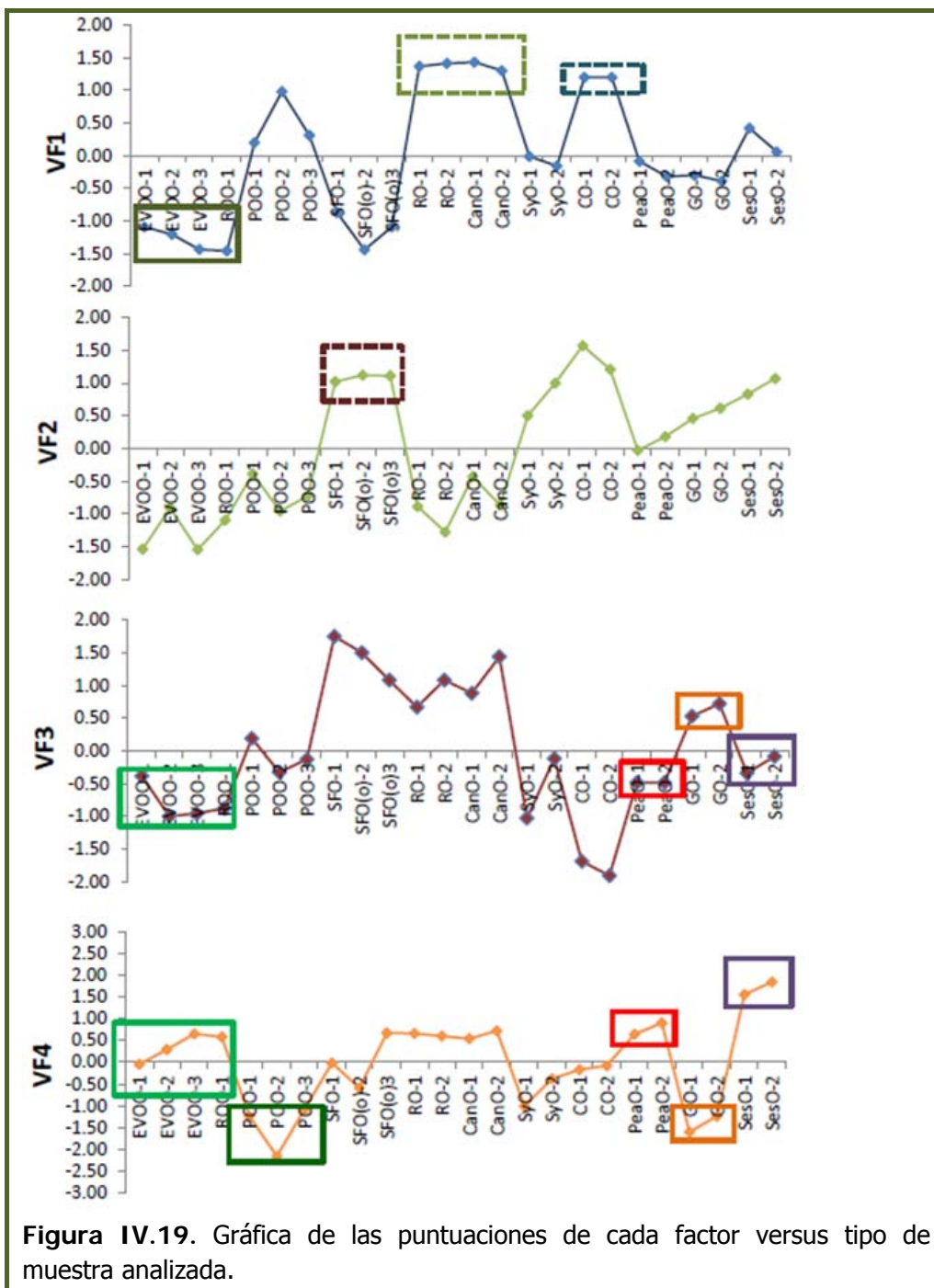


Figura IV.19. Gráfica de las puntuaciones de cada factor versus tipo de muestra analizada.

#### IV.6. TÉCNICAS SUPERVISADAS DE RECONOCIMIENTO DE PAUTAS: ANÁLISIS DISCRIMINANTE LINEAL DEL PERFIL DE CONCENTRACIONES NORMALIZADAS.

Este tipo de análisis está diseñado para desarrollar un conjunto de funciones discriminantes el cual puede ayudar a predecir la "Clase" tomando como referencia

las variables cuantitativas consideradas para el establecimiento del modelo. Esto se hace construyendo funciones discriminantes que son combinaciones lineales de las variables. Para tal fin, se emplearon 24 muestras para desarrollar un modelo que intentara discriminar entre las nueve "Clases" definidas en la Tabla IV.7.

**Tabla IV.7.** Clases asignadas para el análisis discriminante lineal.

Clase	Tipo de aceite
1	Oliva virgen extra y refinado
2	Orujo de oliva
3	Girasol
4	Colza y cáñola
5	Soja
6	Maíz
7	Cacahuete
8	Pepita de uva
9	Sésamo

Se empleó un algoritmo de selección por pasos que incluía solamente algunas variables que eran estadísticamente significativas para discriminar sobre los grupos/clases y se determinó que 7 variables de las 15 eran predictoras significativas de "Clase".

Estas variables correspondían a: los valores normalizados de **campesterol**, **estigmasterol**,  $\Delta^7$ -**campesterol**, **sitostanol**,  $\Delta^5$ -**avenasterol**,  $\Delta^{5,24}$ -**estigmastadienol** y  $\Delta^7$ -**avenasterol**.

La opción del análisis desarrollada fue la selección paso a paso hacia delante que consiste en comenzar con el modelo implicando todas las variables especificadas en la caja de diálogo entrada de datos y elimina una sola variable a la vez basando en la significancia estadística del modelo actual. En cada paso, el algoritmo elimina del modelo la variable que podría ser la de menor significancia estadística.

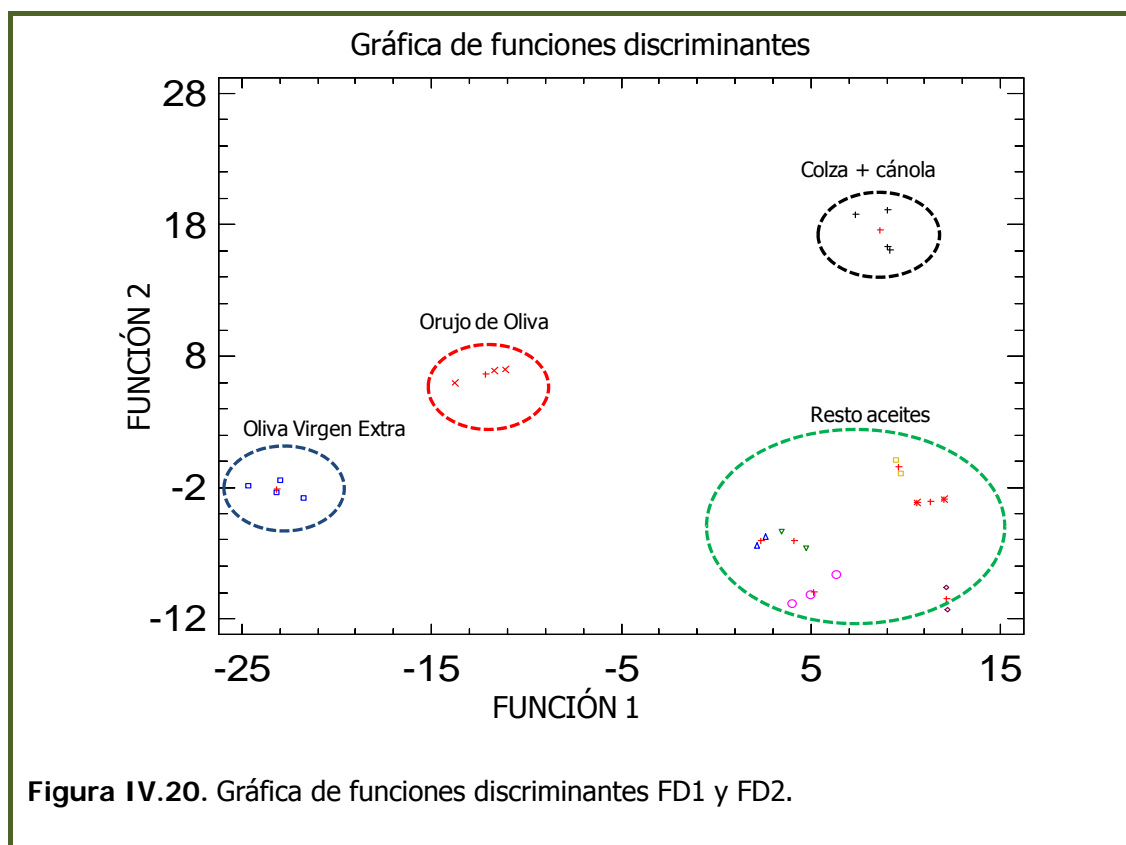
Se necesitaron 6 funciones discriminantes con valores de p menores que 0.05 siendo estadísticamente significativas con un nivel de confianza del 95%. La primera función discriminante estandarizada es:

$$FD1 = -0.0710438 \times W11 - 0.979747 \times W12 - 0.266992 \times W13 + 0.327256 \times W15 + 1.01 \times W4 + 1.02285 \times W6 + 0.375105 \times W7$$

siendo,

- W11: valor normalizado de la concentración de sitostanol.
- W12: valor normalizado de la concentración de  $\Delta^5$ -avenasterol.
- W13: valor normalizado de la concentración de  $\Delta^{5,24}$ -estigmastadienol.
- W15: valor normalizado de la concentración de  $\Delta^7$ -avenasterol.
- W4: valor normalizado de la concentración de campesterol.
- W6: valor normalizado de la concentración de estigmasterol.
- W7: valor normalizado de la concentración de  $\Delta^7$ -campesterol.

La Figura IV.20 muestra la gráfica para las dos primeras funciones discriminantes. El ajuste del modelo para las 24 muestras fue del 100%, clasificándose todas las muestras correctamente.



Aunque el porcentaje de clasificación del modelo fue del 100% en todos los casos, a la vista de la figura se podría indicar que verdaderamente el modelo permitiría discriminar entre cuatro clases consideradas estas como:

- Clase I: Oliva virgen extra y refinado.
- Clase II: Orujo de oliva
- Clase III: Colza/cánola
- Clase IV: Otros aceites.

Aunque puede parecer que los resultados encontrados no fueron satisfactorios, considerando que cinco de las nueve clases estaban constituidas solo por dos muestras y que a priori el aceite de oliva mostraba un comportamiento diferente al del resto de los aceites analizados, el estudio quimiométrico realizado sobre los perfiles de concentración de los 4-desmetilesteroles permitió como conclusión el planteamiento de la siguiente hipótesis:

**Si existen diferencias entre los perfiles de concentración del aceite de oliva con respecto al de otros aceites vegetales ¿se podría utilizar la huella dactilar cromatográfica asociada a este aceite para cuantificar su presencia en mezclas binarias con otros aceites vegetales?**

## BIBLIOGRAFÍA

- [1] Bucci, R.; Magrí, A.D.; Magrí, A.L.; Marini, D.; Marini, F. *Chemical authentication of extra virgin olive oil varieties by supervised chemometric procedures*. Journal of agricultural and food chemistry. 2002, 50, 413-418.
- [2] Sánchez Casas, J.; Osorio Bueno, E.; Montaña García, A.M.; Martínez Cano, M. *Sterol and erythrodiol + uvaol content of virgin olive oils from cultivars of Extremadura (Spain)*. Food Chemistry. 2004, 87, 225-230.
- [3] Lerma-García, M.J.; Simó-Alfonso, E.F.; Méndez, A.; Lliberia, J.L.; Herrero-Martínez. *Classification of extra virgin olive oils according to their genetic variety using linear discriminant analysis of sterol profiles established by ultra-performance liquid chromatography with mass spectrometry detection*. Food Research International. 2011, 44, 103-108.
- [4] Alonso-Salces, R.M.; Héberger, K.; Holland, M.V.; Moreno-Rojas, J.M.; Mariani, C.; Bellan, G.; Reniero, F.; Guillou, C. *Multivariate analysis of NMR fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes*. Food Chemistry. 2010, 956-965.
- [5] Ruíz-Méndez, M.V.; López-López, A.; Garrido-Fernández, A. *Characterization and chemometric study of crude and refined oils from table olive by-products*. Eur. J. Lipid Sci. Technol. 2008, 110, 537-546.
- [6] Sánchez Casas, J.; Osorio Bueno, E.; Montaña García, A.M.; Martínez Cano, M. *Sterol and erythrodiol + uvaol content of virgin olive oils from cultivars of Extremadura (Spain)*. Food Chemistry. 2004, 87, 225-230.
- [7] Amaral, S.S.; Mafra, I.; Oliveira, M.B.P.P. *Characterization of three Portuguese varietal olive oils based on fatty acids, triacylglycerols, phytosterols and vitamin E profiles: Application of chemometrics*.
- [8] Alonso-Salces, R.M.; Héberger, K.; Holland, M.V.; Moreno-Rojas, J.M.; Mariani, C.; Bellan, G.; Reniero, F.; Guillou, C. *Multivariate analysis of NMR fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes*. Food Chemistry. 2010, 956-965.

## **CAPÍTULO V**





# Estudio sobre la capacidad de la huella dactilar cromatográfica para diferenciar aceites vegetales

En este estudio nos propusimos como objetivo evaluar la utilización de la huella dactilar (fingerprint), obtenida a partir de la metodología desarrollada en el Capítulo III de esta memoria, de los perfiles cromatográficos de los fitoesteroles de diferentes muestras de aceites vegetales para observar, si dicha huella exhibía las características adecuadas para permitir, mediante la aplicación de técnicas de reconocimiento de pautas:

- (1) Obtener información sobre como dicho perfil podía agrupar o diferenciar de forma natural los distintos tipos de aceites vegetales,
- (2) Desarrollar un posible modelo de clasificación de los diferentes tipos de aceites vegetales utilizados en alimentación para poder llevar a cabo la evaluación del contenido en aceite de oliva en mezclas o alimentos elaborados con él.

Con objeto de mantener el perfil cromatográfico inherente a los fitoesteroles que habitualmente acompañan a los aceites vegetales se decidió que para la obtención de la huella cromatográfica, siguiendo el procedimiento recogido en el Capítulo II de esta memoria (seguido por el método oficial), en la etapa de adecuación de la muestra de aceite a la etapa de medida se eliminaría la adición del subrogado necesario para la cuantificación de los fitoesteroles.

Así, para obtener el insaponificable de un aceite vegetal se sigue el procedimiento citado anteriormente con la diferencia de que no se adiciona patrón interno porque el objetivo de este tratamiento de muestra no es identificar y cuantificar los desmetilesteroles, sino obtener el cromatograma entero, huella dactilar, para aplicarle diferentes tipos de técnicas quimiométricas.

## V.1. ACEITES UTILIZADOS

Las muestras analizadas se recogen en la Tabla V.1. en dónde se muestra además el código asignado para cada tipo de aceite y cuya procedencia viene descrita en el Capítulo II de esta memoria.

**Tabla V.1.** Muestras de aceites vegetales

<b>Tipo de muestra</b>	<b>Código</b>	<b>Cantidad</b>
Oliva Virgen Extra (Arbequina, cornicabra, hojiblanca, picual, manzanilla, royal)	EVOO	26
Orujo de Oliva	POO	4
Girasol	SFO	9
Girasol alto contenido en ácido oleico	SFO(o)	2
Colza	RO	3
Canola	CanO	2
Soja	SyO	4
Maíz	CO	3
Lino	FO	3
Cacahuete	PeaO	5
Uva	GO	3
Sésamo	SesO	3
Semillas	SeeO	4

## V.2. OBTENCIÓN DE LA MATRIZ DE DATOS ORIGINALES

El perfil cromatográfico de los TMS-esteros (cromatograma entero) de cada muestra analizada que se obtiene dura 45 minutos con un total de 54001 datos/variables. Cada matriz de datos está compuesta por dos columnas (tiempo y señal) para cada cromatograma. Estos datos se exportan a una hoja de Excel para convertirlos en un fichero con extensión .csv para poder ser leídos en el programa MATLAB y trabajar con el software PLS-toolbox.

Los datos exportados se procesan eliminando los 15 primeros y últimos minutos del cromatograma reduciendo la matriz de datos de 54001 a 18201 ya que esas regiones del principio y final del cromatograma no aportaban ningún tipo de información química. En la Figura V.1, se observa la región de interés seleccionada que va desde el intervalo 17800 a 36000. Por tanto, se genera una matriz de datos donde cada una de las filas corresponde a los datos de un cromatograma, es decir, un único vector de datos correspondiente a los valores de señal dada por los analitos. Esta matriz que se ha generado es la que se va a utilizar para su posterior tratamiento.

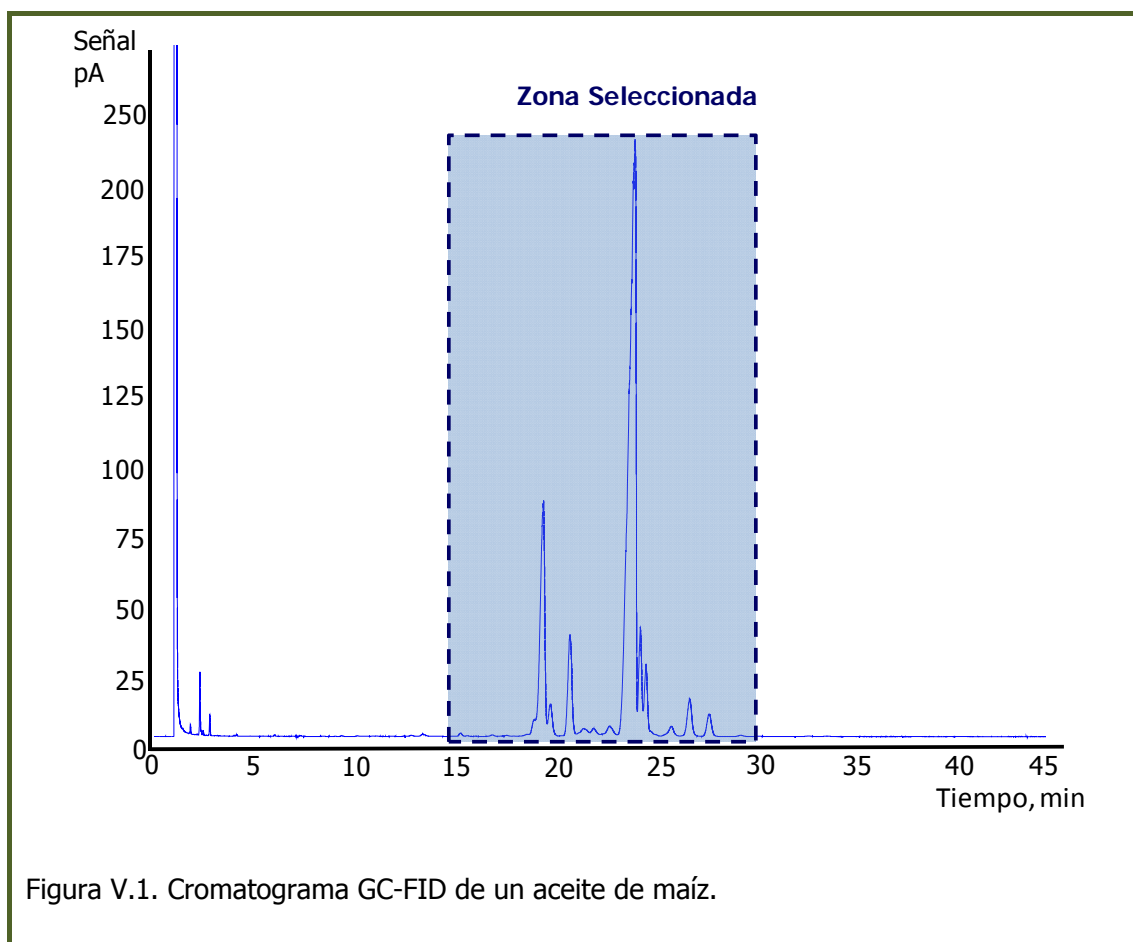


Figura V.1. Cromatograma GC-FID de un aceite de maíz.

### V.3. PREPROCESADO DE LOS CROMATOGRAMAS

Una vez solapados todos los cromatogramas de las 71 muestras y seleccionada la zona de interés, se lleva a cabo el pre-procesamiento de los datos [1, 2, 3, 4] para linealizar la respuesta y eliminar las posibles fuentes extrañas de variabilidad.

Se llevaron a cabo 2 tipos de estrategias con el fin de comparar los resultados obtenidos. A continuación, se muestra la Tabla V.2 de los diferentes preprocesados llevado a cabo.

Tabla V.2. Preprocesado de la señal.

Estrategia I	
Opción A	Opción B
<ol style="list-style-type: none"> <li>1. Ajuste de la línea base "Bioinformatics Toolbox"</li> <li>2. Alineamiento de los picos</li> <li>3. Centrado de Media</li> </ol>	<ol style="list-style-type: none"> <li>1. Alineamiento de los picos</li> <li>2. Ajuste de la línea base "Bioinformatics Toolbox"</li> <li>3. Centrado de Media</li> </ol>
Estrategia II	
Opción A	Opción B
<ol style="list-style-type: none"> <li>1. Ajuste de la línea base (WLS)</li> <li>2. Alineamiento de los picos</li> <li>3. Centrado de Media</li> </ol>	<ol style="list-style-type: none"> <li>1. Alineamiento de los picos</li> <li>2. Ajuste de la línea base (WLS)</li> <li>3. Centrado de Media</li> </ol>

La "Estrategia N° I" consistía en corregir la línea base usando como función la incluida en el "Bioinformatics Toolbox" de Matlab. Esta función, "msbackadj", es específica para tratar datos de señales tales procedentes de técnicas separativas como espectroscopia, NMR, electroforesis, cromatografía o espectrometría de masas para mejorar la calidad de los mismos. Es fundamental ajustar especialmente dos de los diferentes parámetros: ancho de ventana de desplazamiento, y tamaño del paso de ventana de desplazamiento. Además se mantiene la altura de los picos más altos del cromatograma al hacer un suavizado en los extremos de la ventana definida. Se empleó el algoritmo ICOSHIFT, en dónde se definieron 5 intervalos para mejorar el alineamiento entre los picos mayoritarios presentes en los aceites vegetales destacando:

24-metilcolesterol, campesterol, y campestanol que salen "desalineados" y corresponden al intervalo nº 1 y el  $\beta$ -sitosterol, sitostanol y  $\Delta^5$ -avenasterol correspondiente al intervalo nº 3.

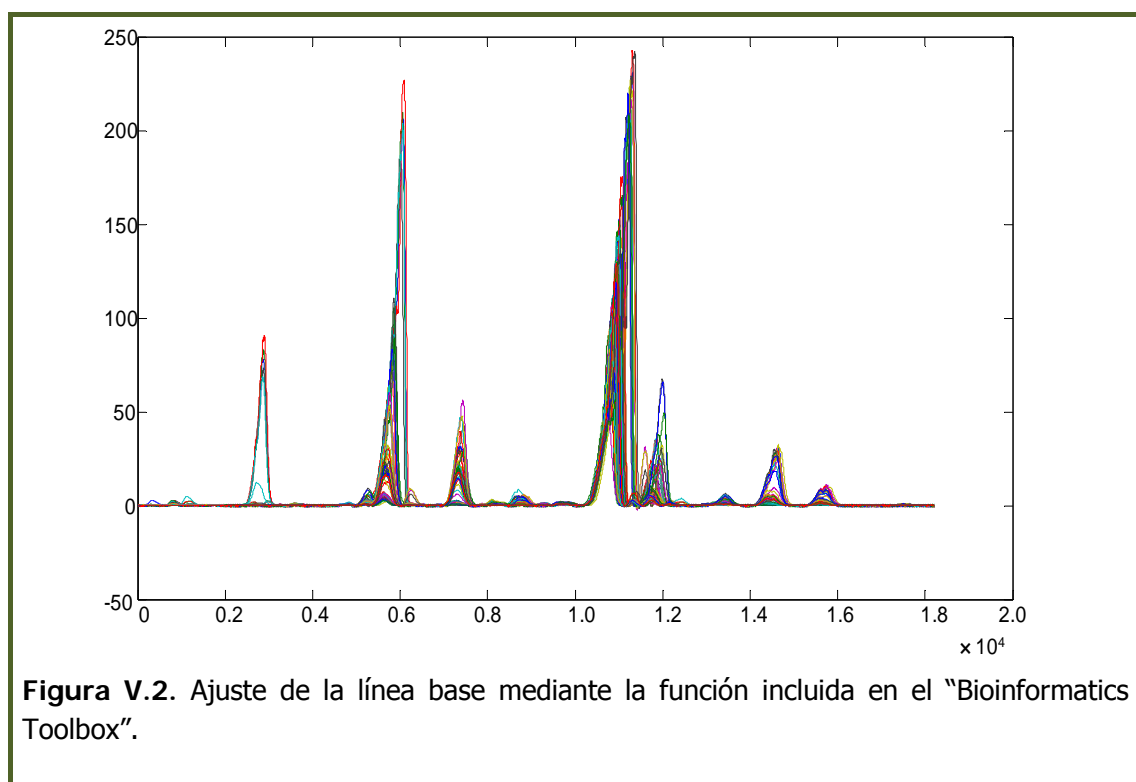
Cada señal era alineada usando el pico máximo como una referencia. Los intervalos definidos que se usaron para alinear fueron (2000, 8200, 9900, 12600, 13800, 16000). Se hizo un centrado de media.

Con esta estrategia se quiso comprobar el efecto que producía al llevar a cabo el ajuste de la línea base con la función definida anteriormente antes o después del alineamiento de la señal.

▪ **Opción A, Estrategia I.**

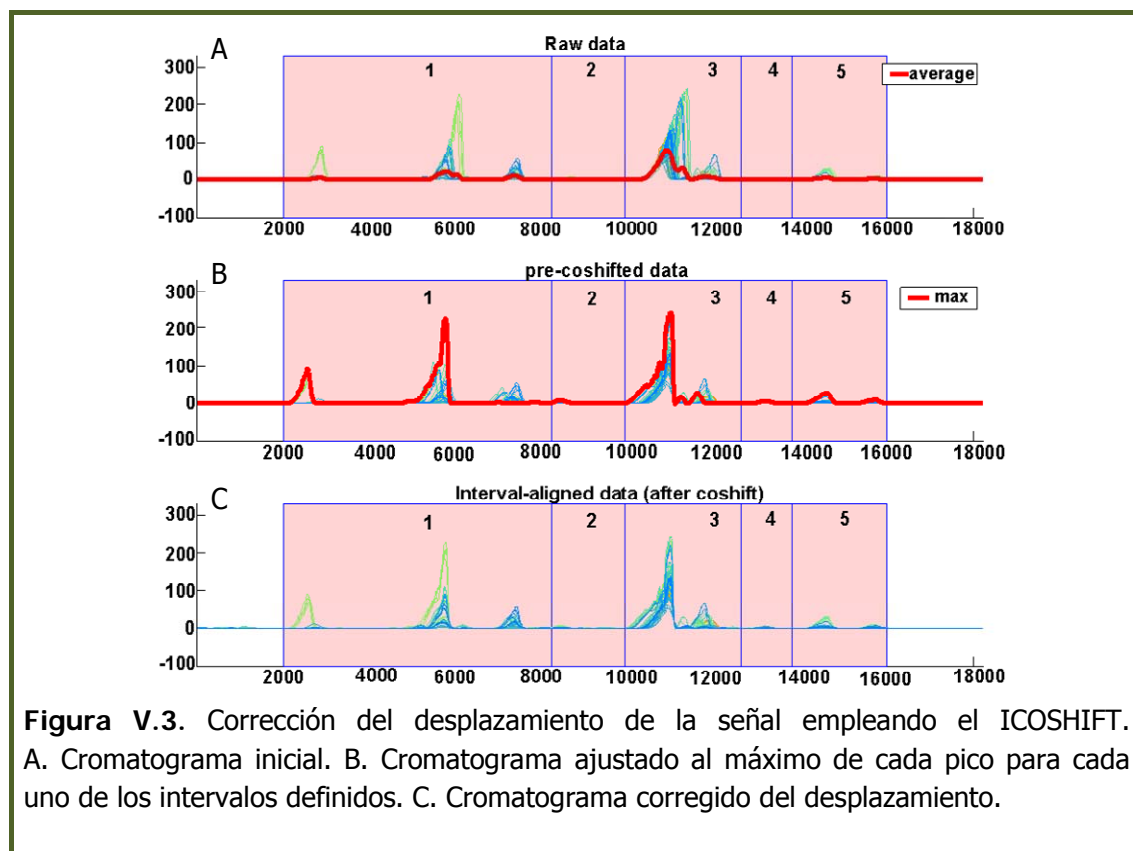
Los resultados que se obtienen tras aplicar este tipo de preprocesado se muestran en las Figuras V.2, V.3 y V.4.

En la Figura V.2 se muestra la superposición de las diferentes huellas dactilares en las que se hace un ajuste de la línea base mediante la función incluida en el programa Matlab llamada "Bioinformatics Toolbox".

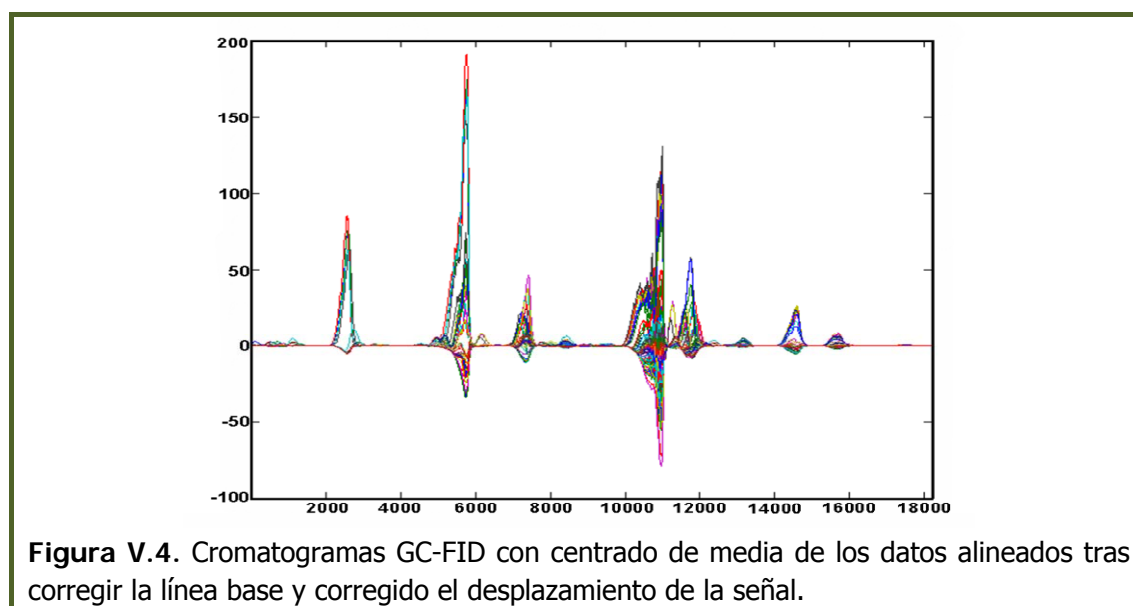


**Figura V.2.** Ajuste de la línea base mediante la función incluida en el "Bioinformatics Toolbox".

En la Figura V.3. se muestra el alineamiento de la señal mediante el algoritmo ICOSIFT tras hacerle un ajuste de la línea base a las huellas dactilares.

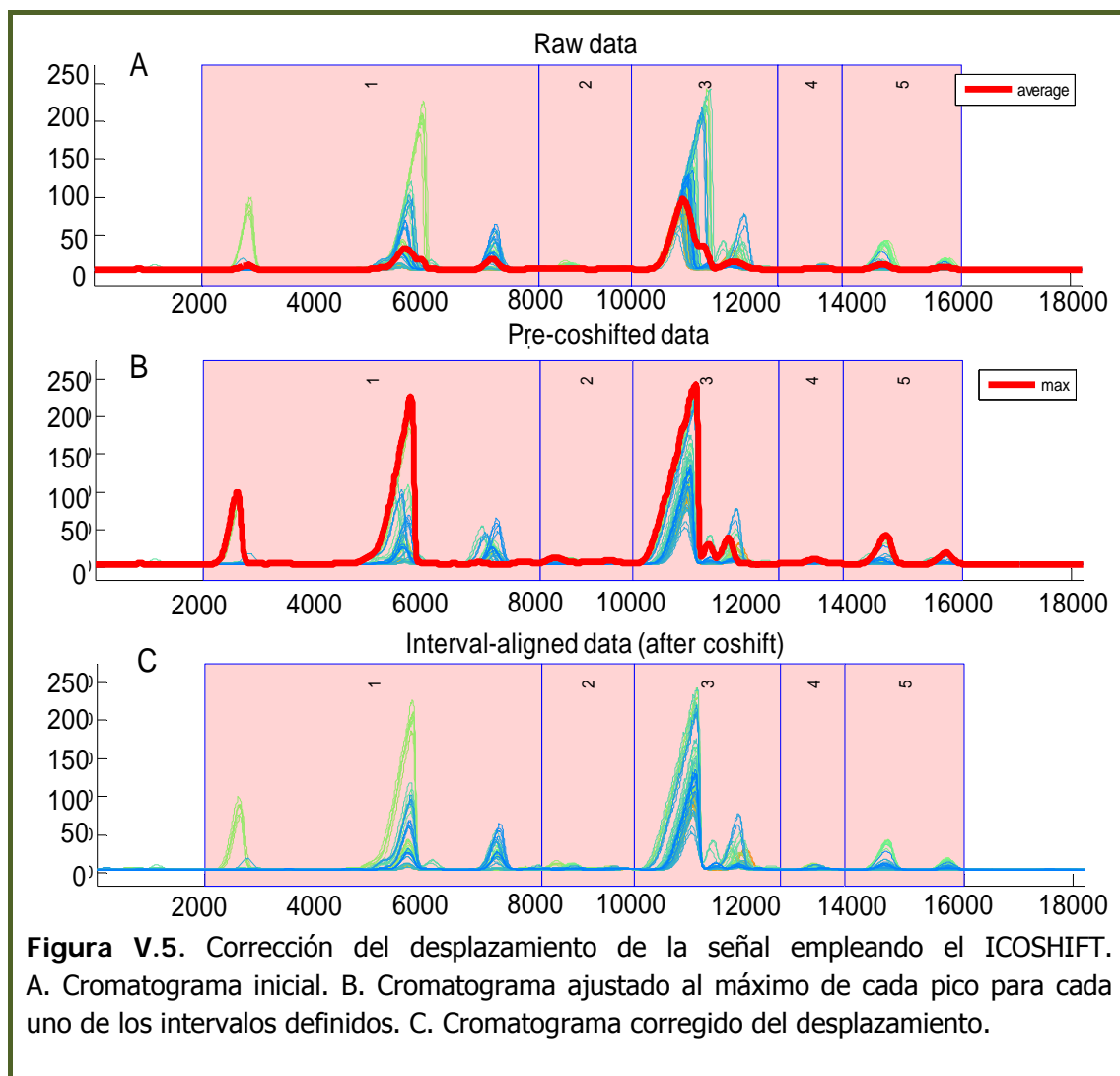


Finalmente, la Figura V.4 muestra el centrado de la media tras hacerle los otros pre-procesados citados anteriormente.



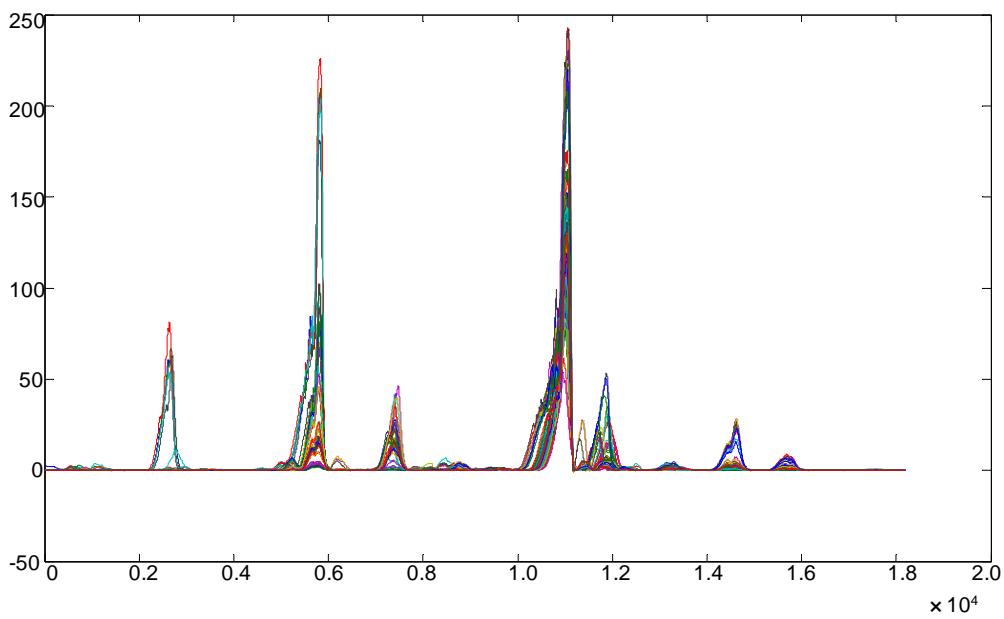
▪ Opción B, Estrategia I.

Siguiendo con la estrategia I se lleva a cabo la Opción B en dónde se altera el orden de preprocesado. Primero se alinea la señal como se puede ver en la Figura V.5, segundo, se ajusta la línea base como se muestra en la Figura V.6 y por último en la Figura V.7 se muestra el centrado de media. Este último paso es común con la opción anterior y con la estrategia I.

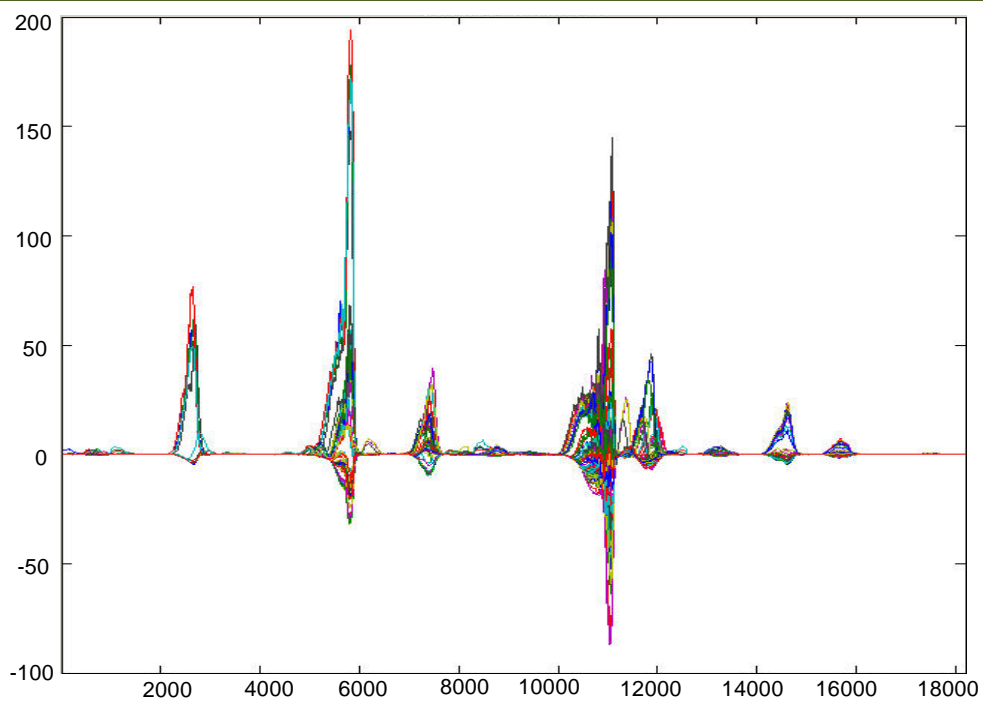


**Figura V.5.** Corrección del desplazamiento de la señal empleando el ICOSHIFT. A. Cromatograma inicial. B. Cromatograma ajustado al máximo de cada pico para cada uno de los intervalos definidos. C. Cromatograma corregido del desplazamiento.





**Figura V.6.** Ajuste de la línea base mediante la función incluida en el "Bioinformatics Toolbox".



**Figura V.7.** Cromatogramas GC-FID con centrado de media de los datos alineados tras corregir el desplazamiento de la señal y la línea base.

Al aplicar esta estrategia, en la opción A, se observa que aumenta la asimetría del perfil cromatográfico de los picos  $\beta$ -sitosterol y campesterol, a diferencia de cuando se emplea la opción B.

De igual modo, en ambos casos sale solapado el 24-metilcoleseterol con el campesterol.

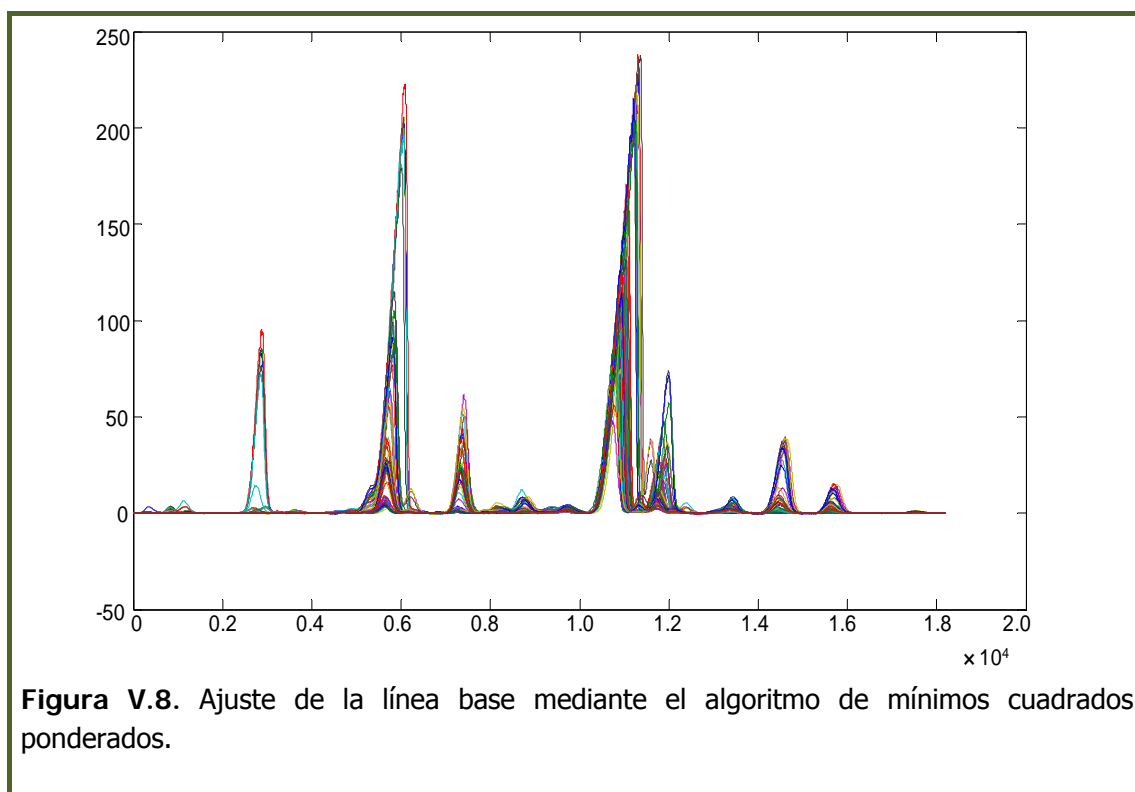
Por tanto, se obtienen mejores resultados cuando primero se alinea la señal y a continuación se ajusta la línea base.

Siguiendo con el esquema de preprocesado se lleva a cabo la "Estrategia N° II" en la que se corrigió la línea base usando el algoritmo de mínimos cuadrados ponderados (WLS) programado desde Matlab. Alineamiento de la señal empleando el algoritmo ICOSHIFT y centrado de media de los datos alineados.

- **Opción A, Estrategia II.**

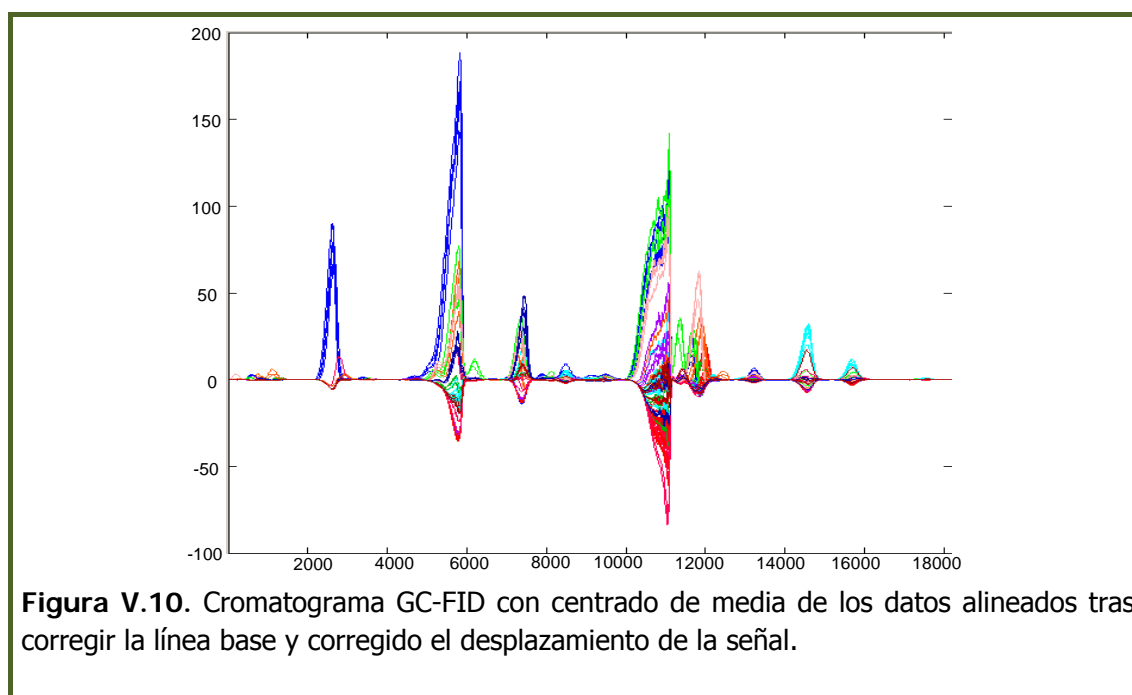
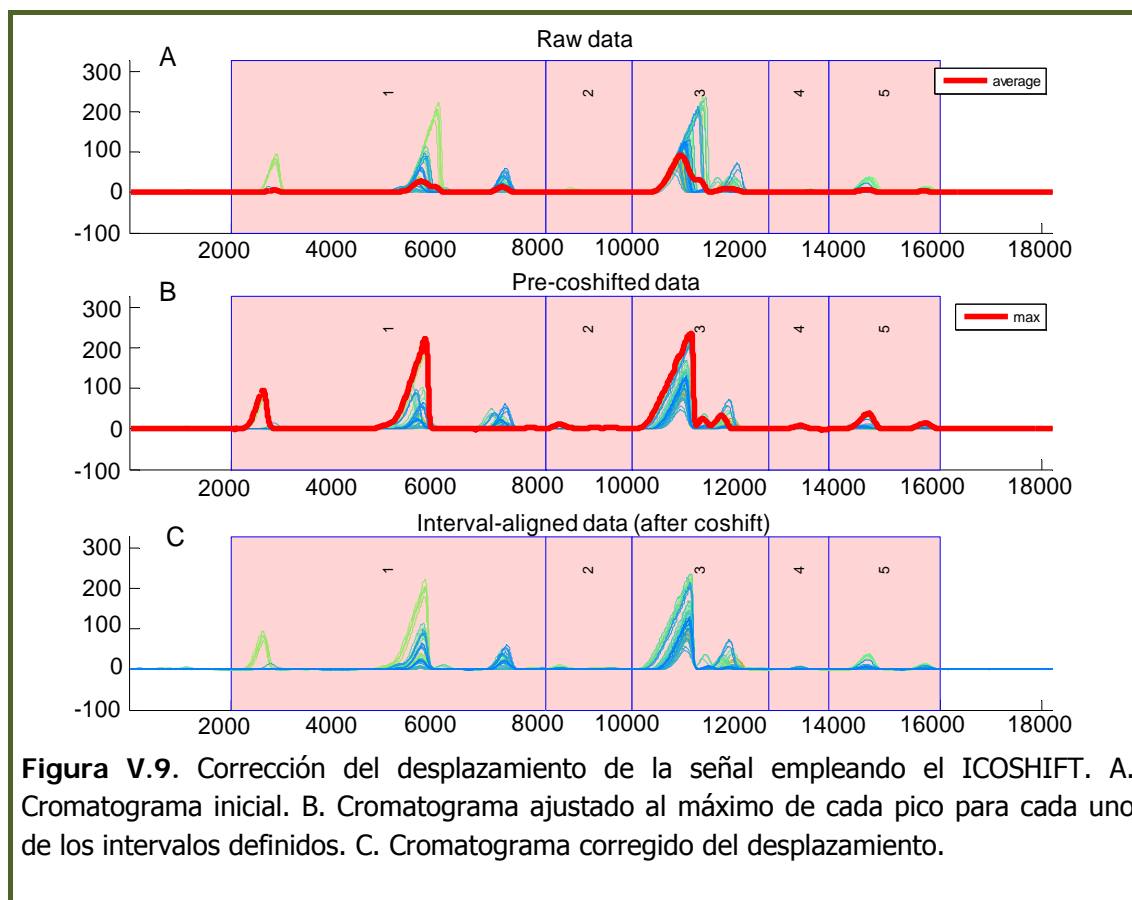
Los resultados obtenidos tras aplicar la opción A para la estrategia 2 se recogen en las siguientes figuras por orden de preprocesado.

En la Figura V.8 se muestra el ajuste de la línea base empleando el algoritmo de mínimos cuadrados parciales (WLS) programándolo desde Matlab.



**Figura V.8.** Ajuste de la línea base mediante el algoritmo de mínimos cuadrados ponderados.

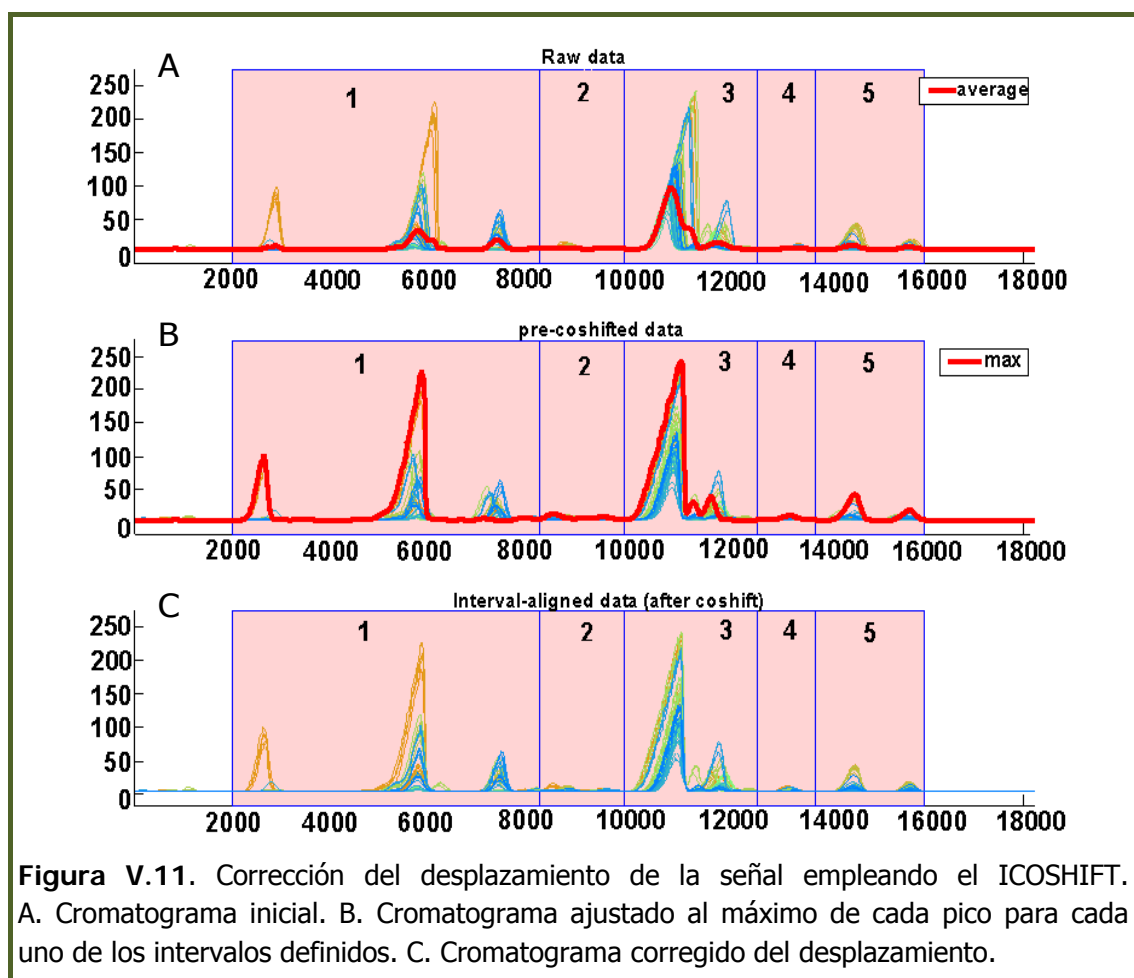
En la Figura V.9. se muestra el alineamiento de las señales y en la Figura V.10 el centrado de la media.

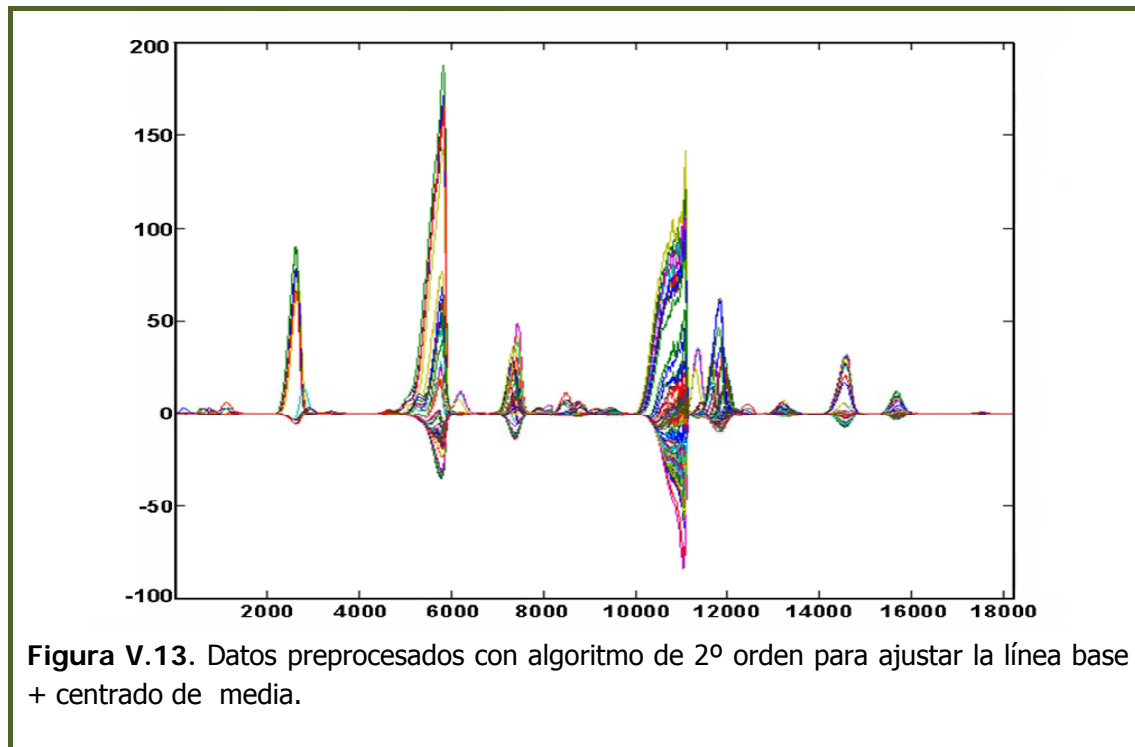
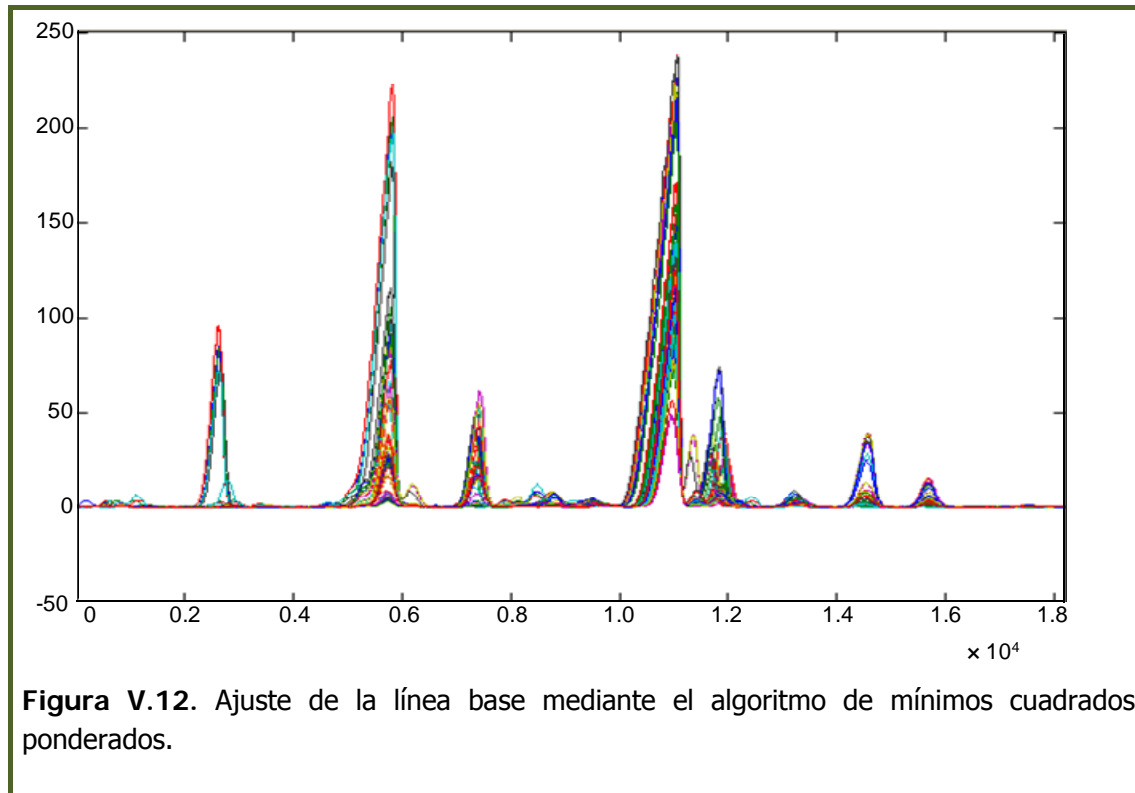


▪ Opción B, Estrategia II.

En esta opción se altera el orden con respecto a la opción A de esta misma estrategia, es decir, primero se alinea las señales y luego se ajusta la línea base con el algoritmo que nos proporciona el programa PLS-Toobox..

En la Figura V.11 se muestra el alineamiento de las señales, en la Figura V.12 se muestra el ajuste de la línea base tras alinear las señales y la Figura V.13 muestra el centrado de la media.





A la vista de los resultados obtenidos, en principio daría igual usar una u otra función para ajustar la línea base, sin embargo, se decide preprocesar los datos según la **Estrategia N° II, opción B**. Debido a que al llevar a cabo el alineamiento de los picos no aparece ningún tipo de asimetría en el perfil cromatográfico y el ajuste de la línea base y el centrado de media se lleva a cabo usando el programa PLS\_TOOLBOX, sin tener que programar desde MATLAB, siendo más fácil su manejo.

**Una vez eligida esta opción como estrategia para el preprocesado de los datos, la matriz generada (71×18201), fue la que se utilizó para aplicar diferentes técnicas quimiométricas de análisis.**

#### V.4. ANÁLISIS JERÁRQUICO DE CONGLOMERADOS

Se realizó un análisis de cluster o de conglomerados jerárquico (HCA) [5, 6, 7, 8, 9, 10] con objeto de analizar si la huella dactilar instrumental obtenida a partir de los cromatogramas completos de los TMS-fitoesteroles de 71 muestras diferentes de aceites vegetales permite observar un agrupamiento natural de los mismos atendiendo al tipo de aceite declarado en la etiqueta. De acuerdo, con esta se trabajó con 10 tipos diferentes de aceites vegetales, que en este caso se identificaron como:

Tipo 1: Oliva Virgen Extra.

Tipo 2: Cacahuete.

Tipo 3: Colza y Canola.

Tipo 4: Girasol.

Tipo 5: Lino.

Tipo 6: Orujo de Oliva.

Tipo 7: Semillas.

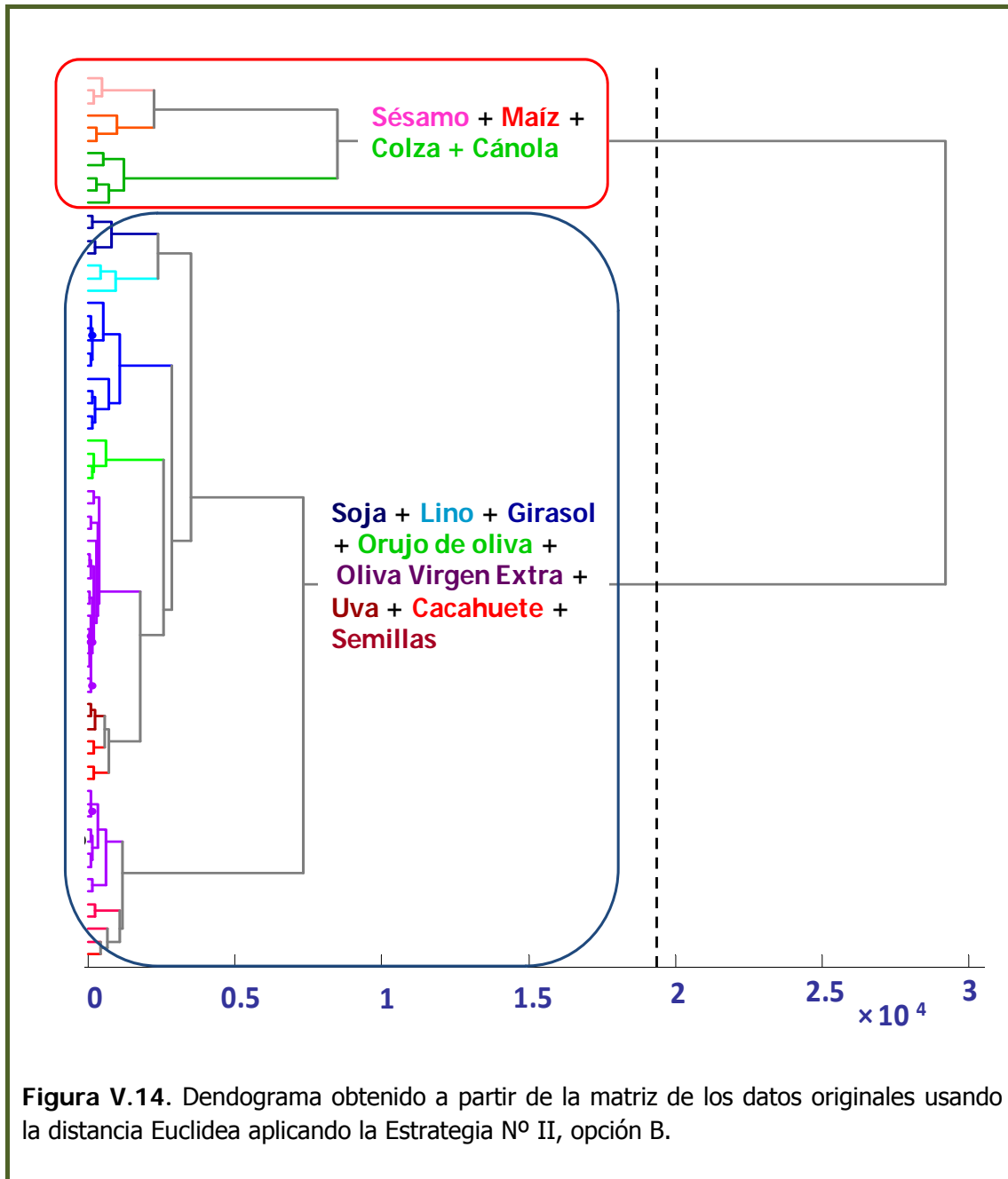
Tipo 8: Sésamo.

Tipo 9: Soja.

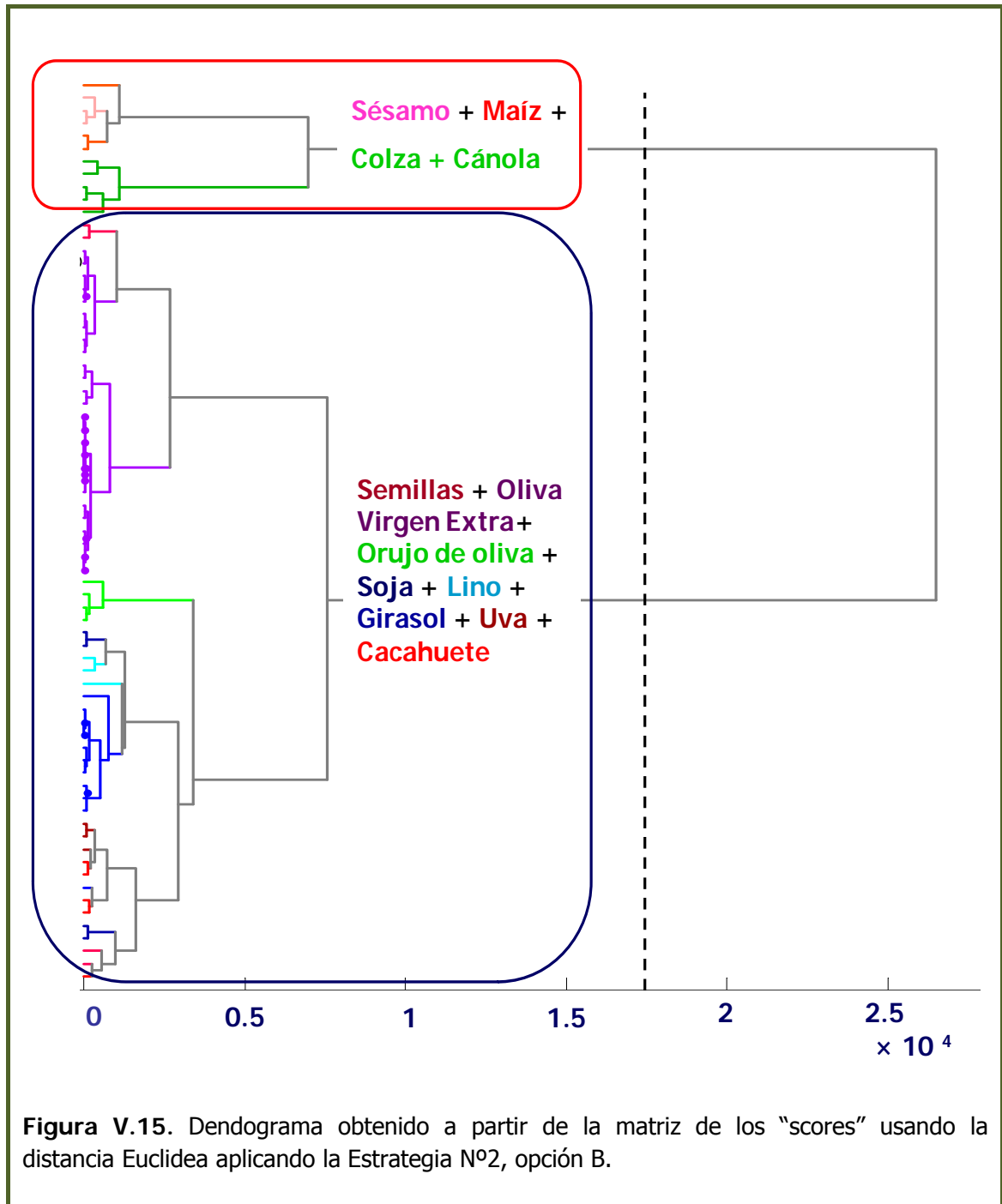
Tipo 10: Pepita de Uva.

Con la matriz de datos inicial (71×18201) obtenida tras preprocesar los mismos, se le aplica el método de Ward y se evalúa con la distancia de aglomeración seleccionando la distancia Euclídea, para el análisis de la matriz de datos originales y la matriz de puntuaciones o "*scores*" obtenida al realizar un ACP previo. Para este último caso se ha ensayado además con la distancia Mahalanobis [11].

Los resultados obtenidos se muestran en varias figuras. La Figura V.14 muestra el dendograma obtenido a partir de la matriz de los datos originales usando la distancia Euclídea



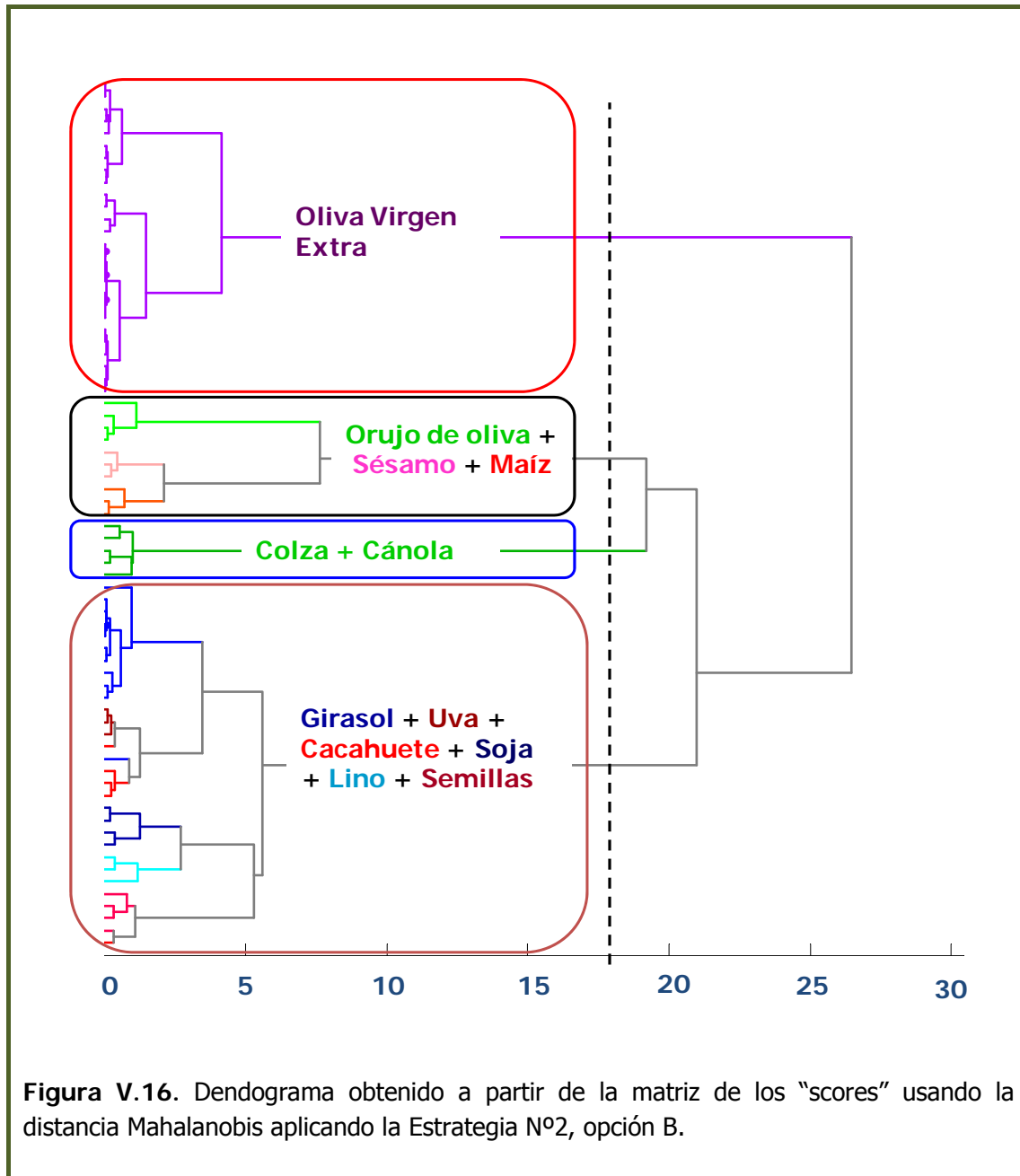
La Figura V.15 muestra el dendograma obtenido a partir de la matriz de los "scores" usando la distancia Euclídea.



**Figura V.15.** Dendograma obtenido a partir de la matriz de los "scores" usando la distancia Euclídea aplicando la Estrategia N°2, opción B.



La Figura V.16 muestra el dendograma obtenido a partir de la matriz de los "scores" usando la distancia Mahalanobis.



**Figura V.16.** Dendograma obtenido a partir de la matriz de los "scores" usando la distancia Mahalanobis aplicando la Estrategia N°2, opción B.

El estudio comparativo de todos los dendogramas obtenidos, indica que los mejores resultados se obtienen cuando se utiliza la distancia de Mahalanobis como distancia de aglomeración. Utilizando como distancia para la selección del clusters  $D_{\text{linkage}} = 0.66 D_{\text{máxima}}$ , se obtienen cuatro agrupamientos, dos de los cuales corresponden uno a las muestras de aceite de oliva virgen extra y el otro a las

muestras de aceites de colza y cáñola, lo que confirma que la utilización de la huella dactilar instrumental obtenida en el preprocesado de los perfiles cromatográficos, constituye una fuente de datos adecuada para la aplicación de técnicas de reconocimiento de pautas a la caracterización de las muestras analizadas.

Además se concluye que la distancia Mahalanobis es mejor que la distancia Euclídea ya que ésta minimiza la variabilidad. Cabe destacar que, al emplear la distancia Mahalanobis, se emplea una matriz de datos con un número de variables inferior a la matriz original facilitando su uso al tener que utilizar una matriz reducida. Para ello se requirió de 3 componentes principales que explicaban un 95.05% de la varianza total.

## V.5. ANÁLISIS DE COMPONENTES PRINCIPALES

Se llevó a cabo un modelo de ACP [12, 13, 14, 15, 16, 17, 18] en el que con un 97.49 % de varianza se explicaban 4 componentes principales tomando como criterio que el autovalor sea mayor que 1 (Figura V.17). En la Tabla V.3 se indican los autovalores y el porcentaje de varianza explicada para cada una de las cuatro componentes principales obtenidas. Observando las gráficas de los pesos, se confirma que el primer y tercer componente son los que más influencia presentan dado que la componente 4 ejerce una contribución mínima. El peso de la componente 1 (PC1) (Figura V.18) corresponde a los principales picos de los esteroides:  $\beta$ -sitosterol, campesterol, brasicasterol, y en menor proporción  $\Delta^5$ -avenasterol, los cuales aportan información en la parte positiva. Los aceites de colza y canola se identifican por la presencia de brasicasterol y los aceites de lino y sésamo por ejemplo se separan por esta componente ya que tienen el mismo perfil de campesterol,  $\beta$ -sitosterol y  $\Delta^5$ -avenasterol. El peso de la componente 2 (Figura V.19) corresponde a los picos del campesterol y brasicasterol en la parte positiva y en la parte negativa se ve influenciada por los picos de  $\beta$ -sitosterol, estigmasterol y  $\Delta^5$ -avenasterol. El peso de la componente 3 (Figura V.20) según la parte positiva del cromatograma tiene mayor influencia el estigmasterol, campesterol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.

La relación de alturas del  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol es muy característica de aceites como soja, girasol y maíz. Finalmente, el peso de la componente 4 (Figura V.21) corresponde en la parte positiva a los esteroides  $\Delta^5$ -avenasterol, campesterol y estigmasterol. Los aceites de cacahuete y uva tienen el mismo perfil pero no guardan la relación campesterol /estigmasterol que indica la componente 3 sino que es al revés. Se definen por la parte positiva de la componente 3 y por la parte negativa de la componente 1. Los aceites de soja se definen por la

componente 3 en la parte positiva ya que mantienen la relación del  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol y también viene definido por la componente 1 tanto en la parte positiva como en la parte negativa. Los aceites de girasol se definen por la parte positiva tanto de la componente 1 como la componente 3. El aceite de orujo de oliva se define por la parte positiva y negativa de la componente 1 y la parte negativa de la componente 3, presentando más campesterol que estigmasterol. Este tipo de aceites sí que tiene  $\Delta^{5,23}$ -estigmastadienol a diferencia del resto de aceites. El aceite de oliva virgen extra se define por la componente 1 pero presenta poco  $\beta$ -sitosterol así como campesterol.

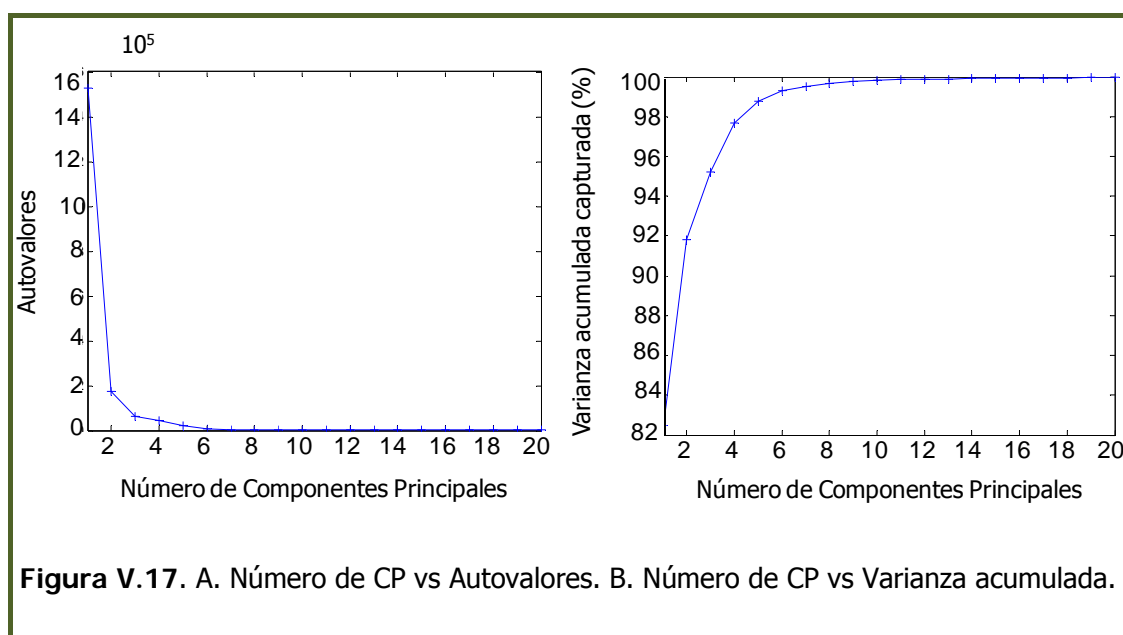


Figura V.17. A. Número de CP vs Autovalores. B. Número de CP vs Varianza acumulada.

Tabla V.3. Autovalores y varianza explicada para los datos analizados según Estrategia N° 2, opción B.

Componente	Autovalor de covarianza	% Varianza explicada	% Varianza acumulada
1	$4.49 \times 10^6$	82.13	82.13
2	$1.75 \times 10^5$	9.67	91.80
3	$5.88 \times 10^4$	3.25	95.05
4	$4.41 \times 10^4$	2.44	97.49

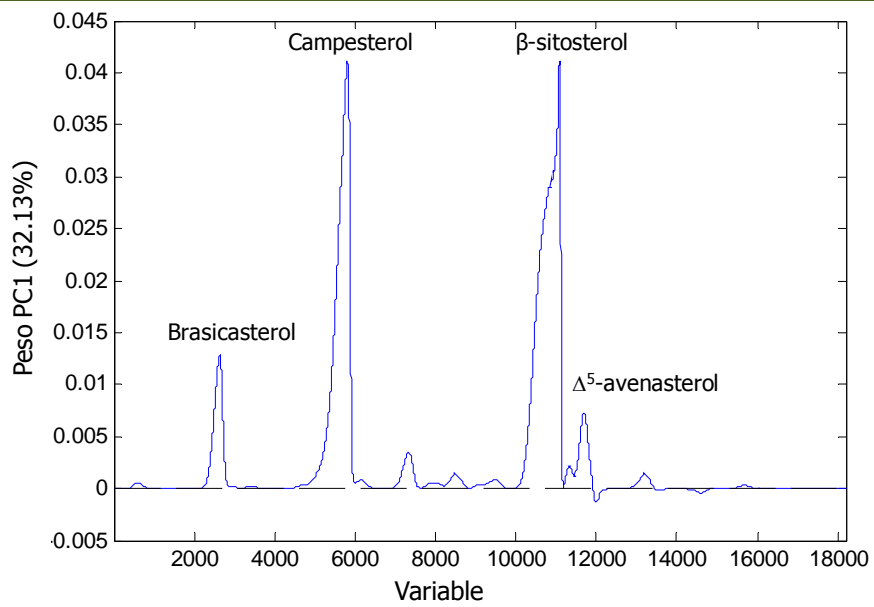


Figura V.18. Peso del ACP de la PC1.

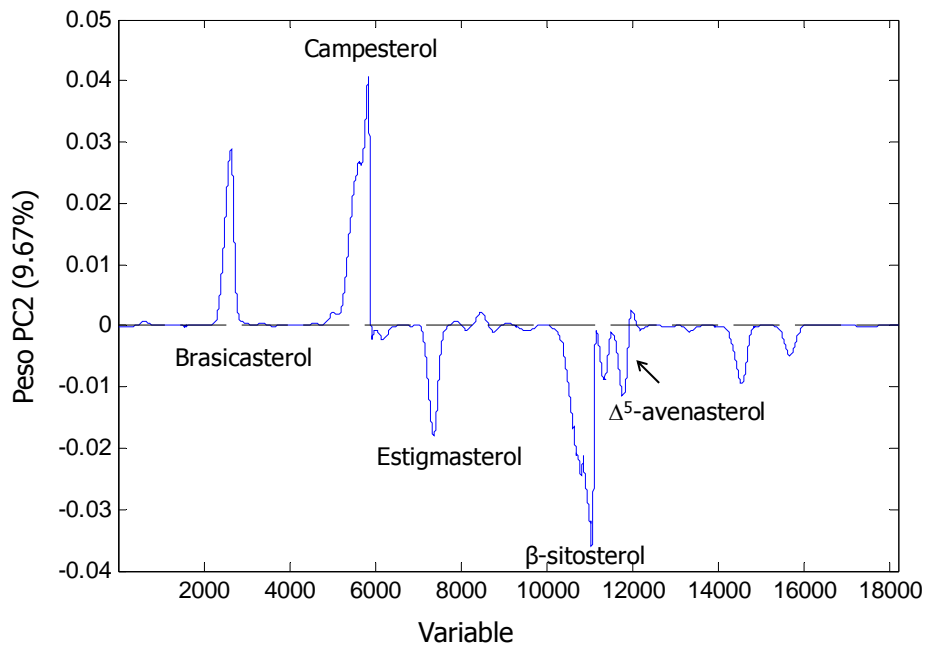


Figura V.19. Peso del ACP de la PC2.

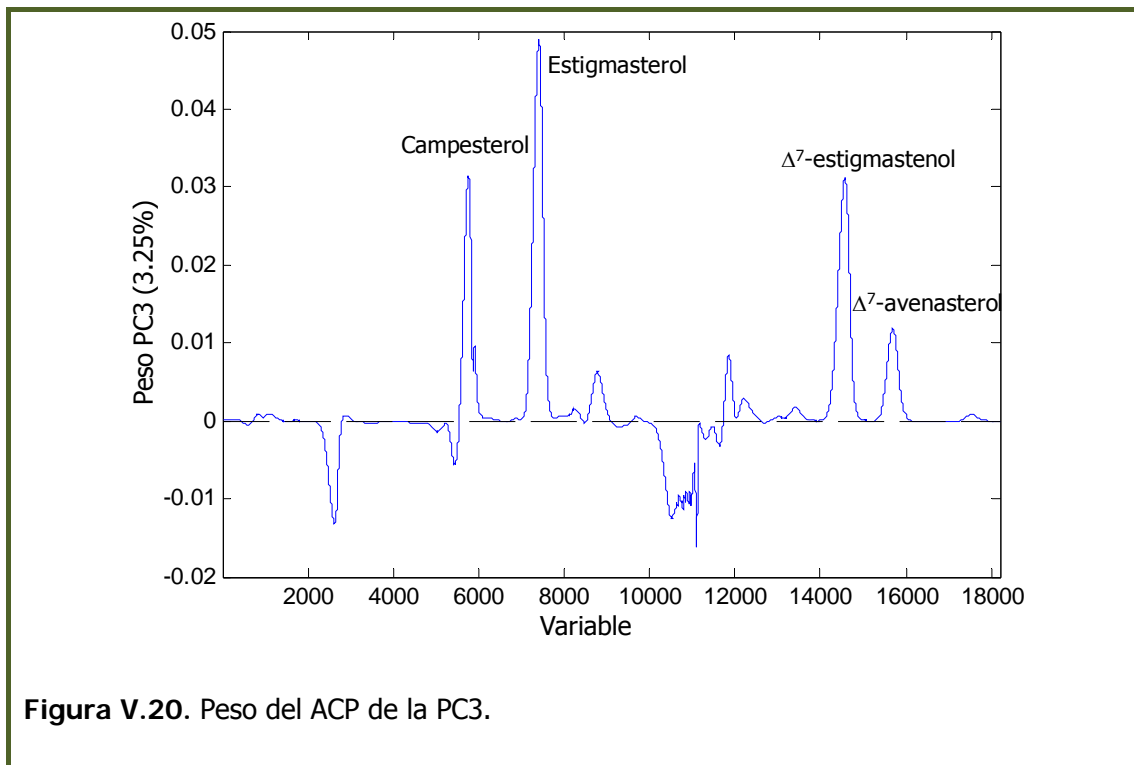


Figura V.20. Peso del ACP de la PC3.

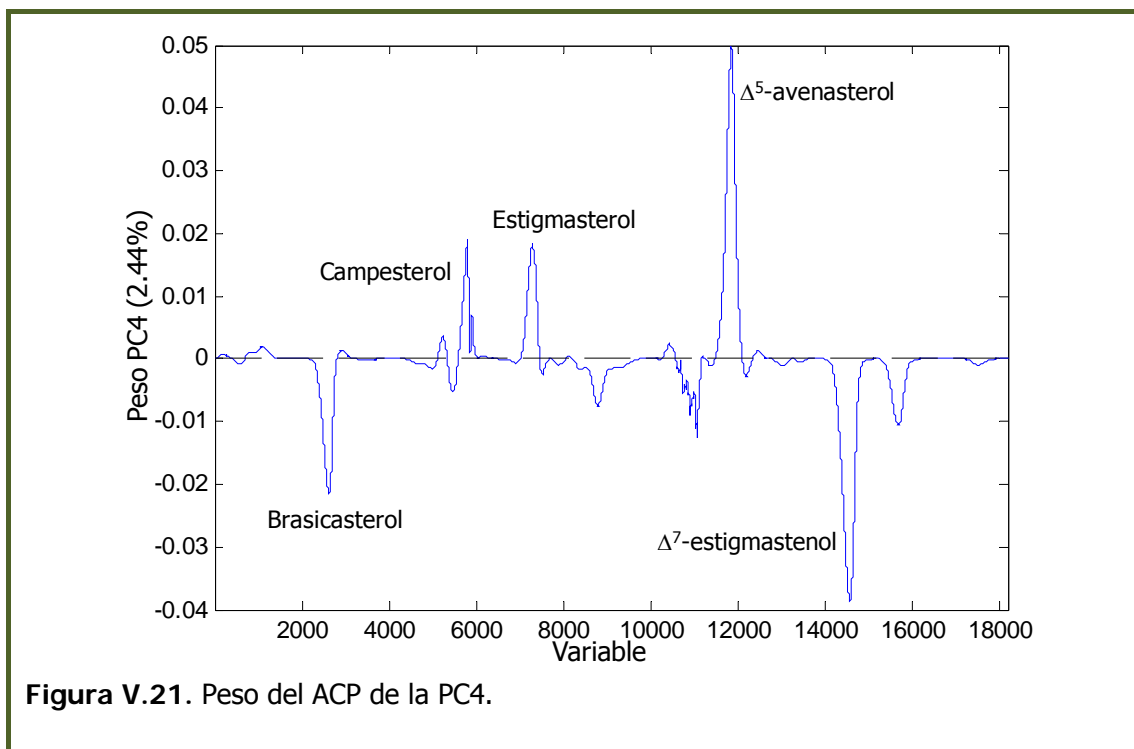
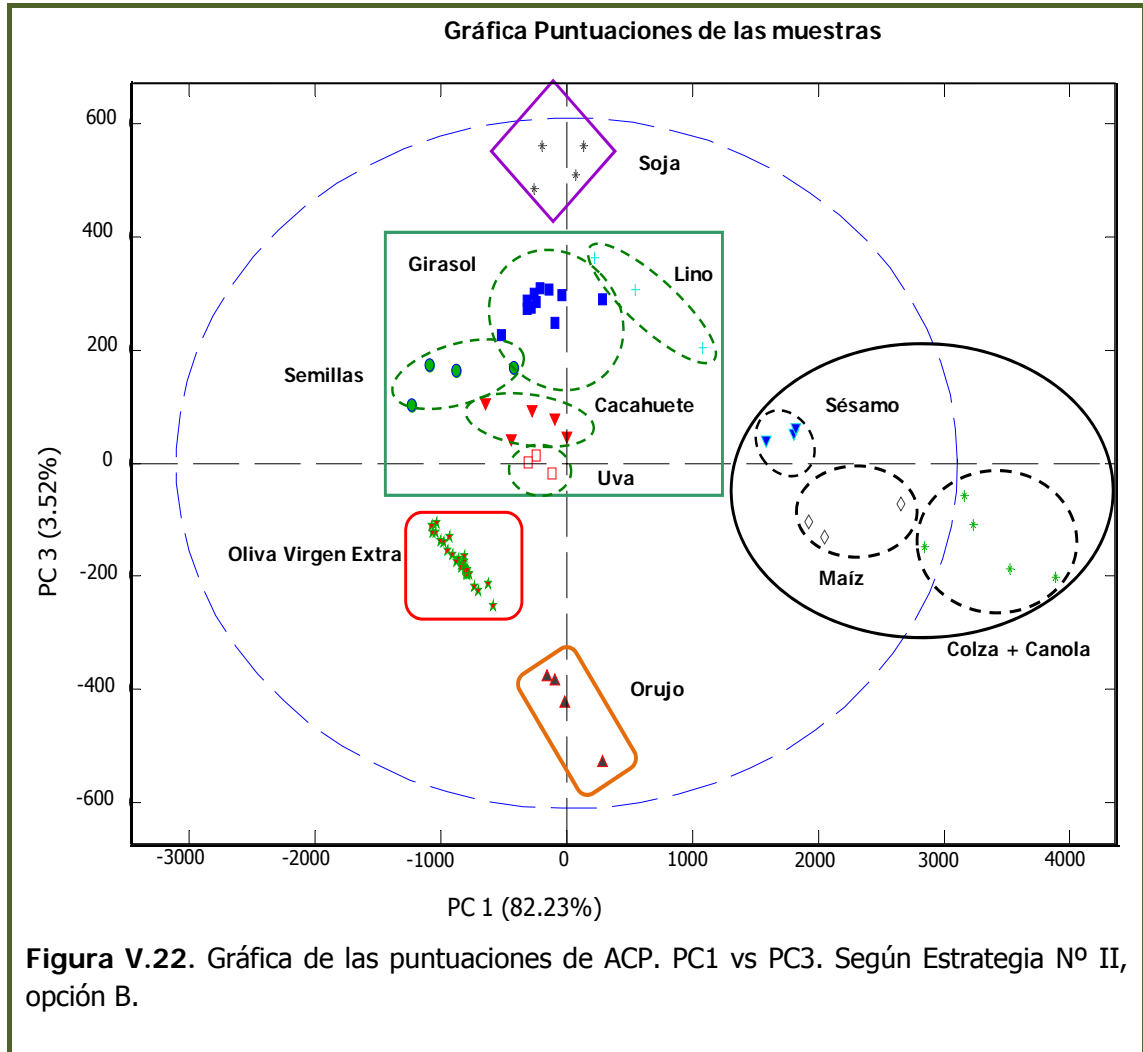


Figura V.21. Peso del ACP de la PC4.

Tras analizar los diferentes pesos de cada una de las 4 componentes se observa la gráfica de las puntuaciones. En la Figura V.22 se observa que se pueden distinguir claramente 5 grupos: Grupo 1: Colza, canola, maíz y sésamo; Grupo 2:

Lino, girasol, semillas, cacahuete y uva; Grupo 3: Soja; Grupo 4: Oliva virgen extra; Grupo 5: Orujo de oliva.



La matriz generada a partir de las puntuaciones del ACP realizado es la que se ha empleado para realizar el HCA con ambas distancias (Euclidea y Mahalanobis).

**Se decide aplicar un método de clasificación, PLS-DA, como técnica de reconocimiento de pautas supervisadas a las siguientes clases de aceites vegetales (soja, girasol, colza y canola, oliva virgen extra y orujo de oliva) ya que tras aplicar el ACP se observan que estos grupos están bien diferenciados.**

## V.6 ANÁLISIS DISCRIMINANTE POR MÍNIMOS CUADRADOS LINEALES PARCIALES (PLS-DA)

Para llevar a cabo este análisis discriminante por mínimos cuadrados parciales [19] se tuvo en cuenta los resultados obtenidos atendiendo al agrupamiento natural de los datos analizados mediante un ACP. De tal forma que se observaba que existían 5 grupos de ahí que para realizar el modelo de PLS-DA se definan 5 clases constituidas por: Clase 1: Ac. Girasol; Clase 2: Ac. Orujo de Oliva; Clase 3: Ac. Oliva Virgen Extra; Clase 4: Ac. Colza y Canola y Clase 5: Ac. Soja.

Las muestras que se analizaron se recogen en la **Tabla V.4**.

**Tabla V.4:** Muestras de aceites vegetales.

Tipo de muestra	Código	Cantidad
Oliva Virgen Extra (Arbequina, cornicabra, hojiblanca, picual, manzanilla, royal)	EVOO	26
Orujo de Oliva	POO	4
Girasol	SFO	9
Girasol alto contenido en ácido oleico	SFO(o)	2
Colza	RO	3
Canola	CanO	2
Soja	SyO	4

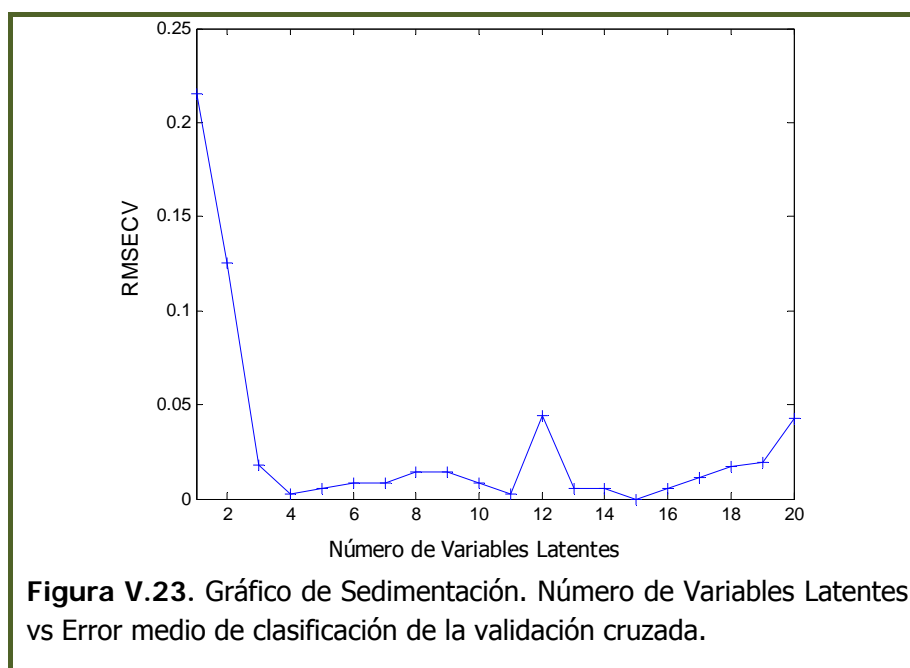
El proceso que se llevo a cabo para realizar el PLS-DA se resume en varias etapas:

1. Se solapan los cromatogramas.
2. Se selecciona la zona de interés químico.
3. Se preprocesan los cromatogramas empleando la Estrategia N° II, opción B, descrita en el Apartado V.4 de este Capítulo, en la que se alinea la señal, se ajusta la línea base y se hace un centrado de media.
4. Se lleva a cabo el PLS-DA con el algoritmo SIMPLS, como técnica de clasificación.

5. Se interpretan los pesos para identificar las regiones del cromatograma de mayor utilidad.
6. Se evalúan los resultados obtenidos.

Para construir el modelo se ha empleado un set de calibración compuesto por 38 muestras y 12 muestras se emplearon para validarlo externamente constituyendo el 25% del conjunto total de muestras que eran 50. Simultáneamente se ha validado usando una validación cruzada tipo Venetian Blinds con un Split de 6. 4 variables latentes fueron necesarias para construir el modelo PLS-DA bajo la regla general de que no se incluyen más variables latentes adicionales a menos que se mejore el RMSECV por encima de un 2%.

En la Figura V.23 se muestra el gráfico de sedimentación para elegir el número de variables latentes.



Estas 4 variables latentes explican el 99.20 % de la varianza en el **Bloque X** y el 91.20 % en el **Bloque Y**.

El **bloque X** del set de calibración estaba compuesto por 38 muestras  $\times$  18201 variables y el **bloque Y** por una matriz de 38 filas  $\times$  5 columnas, tantas filas como muestras haya y tantas columnas como clases definidas, compuestas por 1s y 0s colocando el valor de 1 a la columna que le corresponda ese tipo de clase. La matriz del **bloque Y** se muestra en la Tabla V.5.



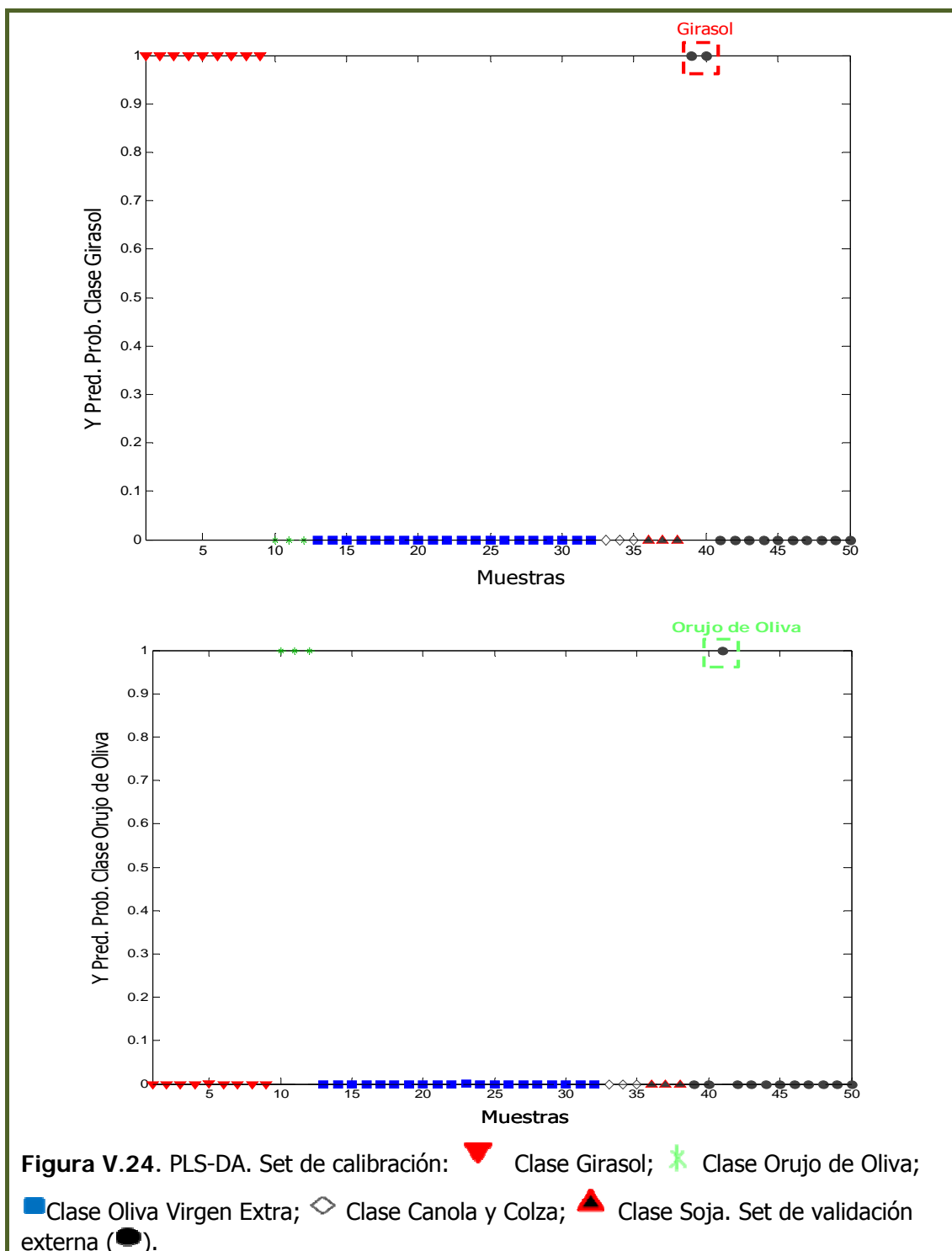
El **bloque X** del set de validación estaba compuesto por 12 muestras  $\times$  18201 variables.

**Tabla V.5.** Matriz del **Bloque Y** del set de Calibración compuesta por 38 filas  $\times$  5 columnas.

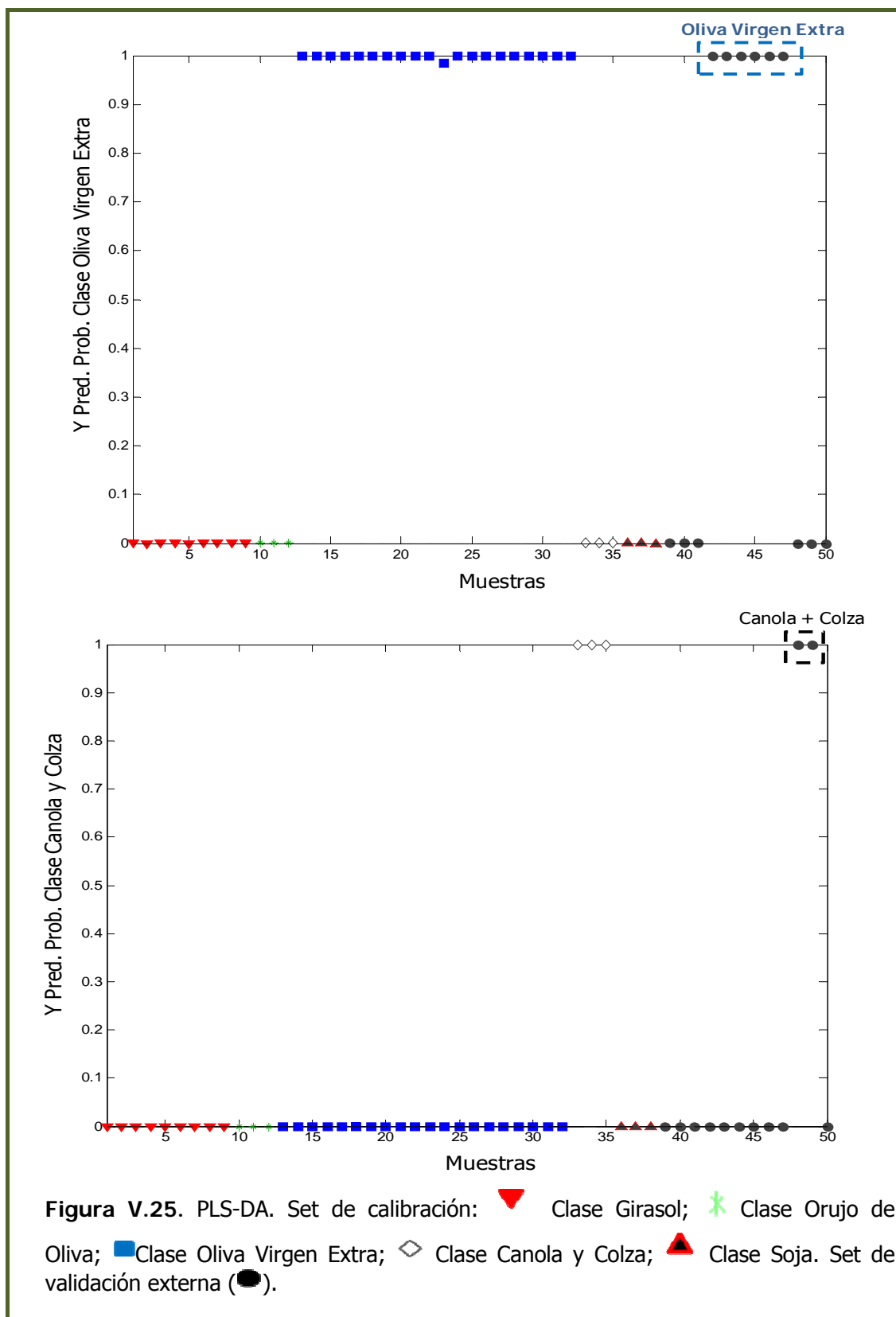
Clase 1	Clase 2	Clase 3	Clase 4	Clase 5
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
1	0	0	0	0
0	1	0	0	0
0	1	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	1	0
0	0	0	1	0
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1

Como se puede ver en las Figuras V.24, V.25 y V.26, todas las muestras están perfectamente clasificadas, discriminando entre aceites de oliva virgen extra, orujo de oliva, girasol, colza más canola y soja.

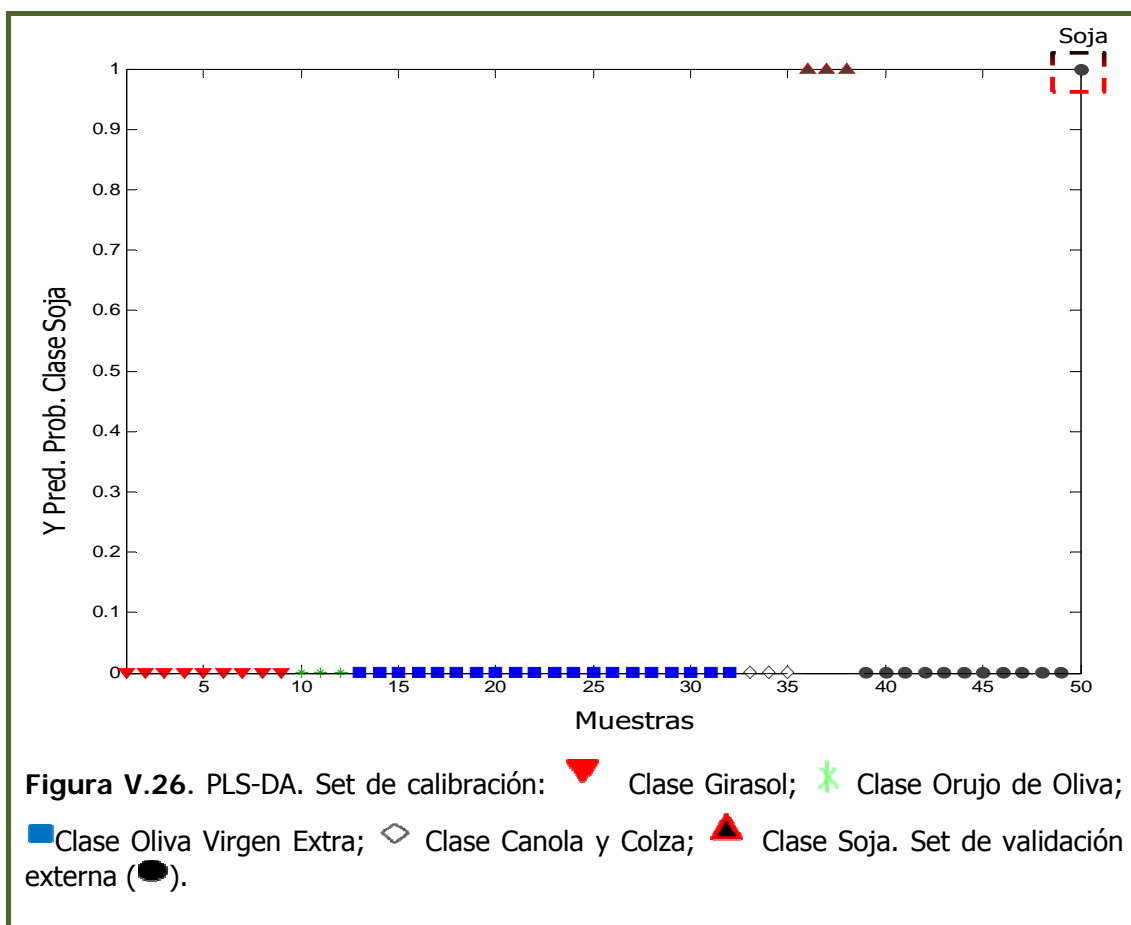
La Figura V.24 muestra la clasificación de las muestras de Girasol y Orujo de oliva.



La Figura V.25 muestra la clasificación de las muestras de aceite de oliva virgen extra y canola y colza.



La Figura V.26 muestra la clasificación de las muestras de aceite de oliva virgen extra y canola y colza.



Los resultados se apoyan con los parámetros estadísticos obtenidos evaluando el modelo con los valores de  $R^2$ , RMSEC, RMSECV y RMSEP que se muestran en las Tablas V.6 y V.7 siendo satisfactorios.

**Tabla V.6.** Resultados estadísticos del modelo entrenado.

	Clase 1	Clase 2	Clase 3	Clase 4	Clase 5
Nº muestras	9	3	20	3	3
RMSEC	0.019	0.053	0.143	0.168	0.043
RMSECV	0.028	0.066	0.171	0.199	0.058
$R^2$ Cal.	0.995	0.985	0.720	0.887	0.974
$R^2$ CV	0.989	0.976	0.610	0.843	0.954

Cal. Calibración

Tabla V.7. Resultados estadísticos de la predicción del modelo.

	Clase 1	Clase 2	Clase 3	Clase 4	Clase 5
Nº muestras	2	1	6	2	1
RMSEP	0.032	0.037	0.114	0.123	0.056
R <sup>2</sup> Pred.	0.993	0.991	0.840	0.953	0.992

Pred. Predicción

Al observar los pesos de cada variable latente (VL) (Figura V.27) se comprueba que las VL que más peso tienen son la VL1 y VL4, ya que el peso de la VL1 corresponde a los picos de  $\beta$ -sitosterol, campesterol y brasicasterol, los cuales aportan información en la parte positiva y el peso de la LV4 corresponde al pico del estigmasterol en la parte positiva y a los picos  $\Delta^7$ -estigmastenol y  $\Delta^7$ -avenasterol en la parte negativa.

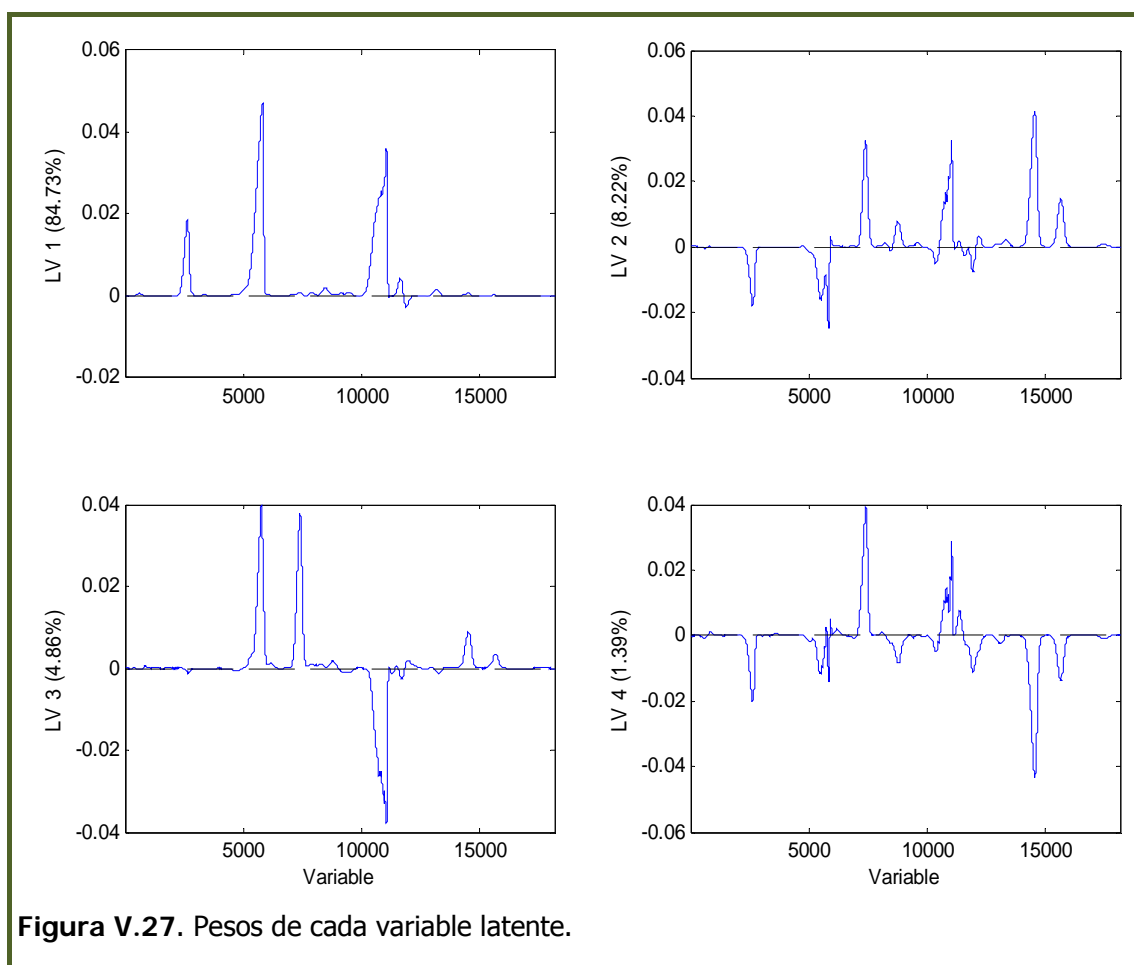


Figura V.27. Pesos de cada variable latente.

Tras aplicar un modelo clasificatorio a las diferentes muestras de aceites vegetales disponibles y comprobar que de todas las muestras analizadas el modelo de clasificación permite diferenciar perfectamente entre el 100 % de las clases consideradas.

Este hecho permitió establecer una nueva hipótesis según la cual la huella dactilar cromatográfica de los 4-desmetilesteroles más habitualmente presentes en aceites vegetales podría emplearse para la cuantificación de aceite de oliva virgen extra en mezclas con otros aceites tales como orujo de oliva, girasol, colza/ canola y soja. Si además se considera las agrupaciones obtenidas mediante la aplicación del análisis de componentes principales se podría cuantificar aceite de oliva en aquellos casos en los que el lugar de utilizar algunos de los aceites mencionados anteriormente se emplearan maíz como sustituto de colza y cáñola y/o aceites de semillas en vez de aceites de girasol.

## BIBLIOGRAFÍA

- [1] Wise, B.M.; Gallager, N.B.; Bro, R.; Shaver, J.M.; Windig, W.; Koch, R.S. PLS\_Toolbox 4.0 Manual for use with MATLABTM, Eigenvector Research, Inc. Wenatchee. 2006.
- [2] Brereton, R.G. *Chemometrics: Data analysis for the Laboratory and Chemical Plant*, Wiley, Chichester, 2003, pp 119.
- [3] Wold, S.; Antti, H.; Lindgren, F. & Ohman, J. *Orthogonal signal correction of near-infrared spectra*. Chemom. Intell. Lab. Syst. 1998, 44, 175-185.
- [4] Berrueta, L.A.; Alonso-Salces, R.M. & Héberger, K. *Review Supervised pattern recognition in food analysis*. J. Chromatogr. A., 2007, 1158, 196-214.
- [5] Lavine, B.K. *Pattern Recognition*. Crit. Rev. Anal. Chem., 2006, 36, 153-161.
- [6] Moller, S.F.; Von Frese, J. & Bro, R. *Robust methods for multivariate data analysis*. J. Chemom., 2005, 19, 549-563.
- [7] Matos, L.C.; Cunha, S.C.; Amaral, J.S.; Pereira, J.A.; Andrade, P.B.; Seabra, R.M. & Oliveira, B.P.P. *Chemometric characterization of three varietal olive oils (Cvs. Cobrancosa, Madural, and Verdeal Transmontana) extracted from olives with different maturation indices*. Food Chemistry, 2007, 102, 406-414.
- [8] Lavine, B.K. & Davidson, C.E. *Classification and Pattern recognition in Gemperline, P. Practical Guide Chemometrics 2 nd.*, Taylor Francis, 2006, pp 139-178.
- [9] Hill, T. & Lewicki, P. *Cluster Analysis* in Hill, T. & Lewicki, P., *Statistics Methods and Applications*, StatSoft, 2005, pp 124-135.
- [10] Iñón, F.A.; Garrigues, J.M.; Garrigues, S.; Molina, A.; de la Guardia, M. *Selection of calibration set samples in determination of olive oil acidity by partial least squares-attenuated total reflectance- Fourier transform infrared spectroscopy*. Analytica Chimica Acta, 2003, 489, 59-75.
- [11] De Maesschalck, R.; Jouan-Rimbaud, D.; Massart, D.L. *Tutorial-The mahalanobis distance*. J. Chemom. Intell. Lab. Syst., 2000, 50, 1-18.

- [12] Galeano Díaz, T.; Duran Merás, I.; Sánchez Casas, J. & Alexandre Franco, M.F. *Characterization of virgin olive oils according to its triglycerides and sterols composition by chemometric methods*. Food Control, 2005, 16, 339-347.
- [13] López-Feria, S.; Cárdenas, S.; García-Mesa, J.A. & Valcárcel, M. *Classification of extra virgin olive oils according to the protected designation of origin, olive variety and geographical origin*. Talanta, 2008, 75, 937-943.
- [14] Nagy, K.; Bongiorno, D.; Avellone, G.; Agozzino, P.; Ceraulo, L. & Vékey, K. *High performance liquid chromatography-mass spectrometry based chemometric characterization of olive oils*. J. Chromatogr. A, 2005, 1078, 90-107.
- [15] Zhang, G.; Ni, Y.; Churchill, J. & Kokot, S. *Authentication of vegetable oils on the basis of their physic-chemical properties with the aid of chemometrics*. Talanta, 2006, 70, 293.
- [16] Varmuza, K. & Filzmoser, P. *Principal Component Analysis* in Varmuza, K.; Filzmoser, P., *Introduction to multivariate statistical analysis in chemometrics*, Taylor & Francis Group, London, England, 2009, pp 74-117.
- [17] Livingstone, D. *Unsupervised learning* in Livingstone, D., *A practical guide to scientific data analysis*, John Wiley & Sons, Ltd, United Kingdom, 2009, pp 119-144.
- [18] Gemperline, P J. *Principal component analysis* in Gemperline, P. *Practical Guide to Chemometric* (2<sup>nd</sup> ed.), Taylor & Francis Group, London (UK), 2006, pp 69-83.
- [19] Dupuy, N.; Galtier, O.; Ollivier, D.; Vanloot, P. & Artaud, J. *Comparison between NIR, MIR, concatenated NIR and MIR analysis and hierarchical PLS model. Application to virgin olive oil analysis*. Analytica Chimica Acta, 20120, 666, 23-31.





# **CAPÍTULO VI**



# Cuantificación de aceite de oliva en vinagretas y aliños comerciales usando un modelo PLS desarrollado con mezclas binarias de aceites vegetales

Es bien conocido, que el aceite de oliva se consume directamente (sólo o mezclado con otros aceites) o, como ingrediente en numerosos alimentos (diversas conservas, salsas y aliños, galletas, embutidos, margarinas y mantequillas, patatas fritas, pastillas para caldo, etc.).

El contenido de aceite de oliva en productos alimenticios viene recogido en el Reglamento de la UE Nº 29/2012 de la Comisión de 13 de enero de 2012 **[1]** sobre las normas de comercialización del aceite de oliva. El artículo 6 de dicho Reglamento, indica que:

1. Cuando se comercializan mezclas de aceites y se menciona (en la forma que sea) que dicha mezcla contiene aceite de oliva (de cualquier categoría (oliva virgen extra, oliva virgen, oliva y orujo de oliva), el contenido de éste **debe estar expresamente indicado en su etiqueta** y su porcentaje deber ser siempre superior al 50% de la mezcla utilizada.

2. Si el producto alimenticio, contiene aceite de oliva, su **contenido debe ser expresado explícitamente**. En ningún caso el reglamento obliga a identificar o indicar la composición de los otros aceites que puedan estar presentes. Están excluidos del reglamento, aquellos alimentos que están regulados por reglamentos específicos, como las conservas en sardinas (Reglamento CEE 2136/89) y atún y bonito (Reglamento CEE 1536/92), en las que sólo se puede indicar la presencia de aceite de oliva cuando el medio de cobertura sea aceite de oliva sin mezclas. Por lo tanto, la denominación de venta de la mezcla será la siguiente: "Mezcla de aceites vegetales (o nombres específicos de esos aceites vegetales) y de aceite de oliva", seguida directamente de la indicación del porcentaje de aceite de oliva en la mezcla.

*En dicho Reglamento se indica además que "...es necesario evitar que los productos alimenticios que contienen aceite de oliva engañen al consumidor poniendo de relieve la reputación del aceite de oliva sin señalar la composición real del producto. Por consiguiente, en las etiquetas debe aparecer claramente una indicación del porcentaje de aceite de oliva y determinadas indicaciones en el caso de los productos constituidos exclusivamente por una mezcla de aceites vegetales. Además, es necesario tener en cuenta las disposiciones especiales previstas en determinados reglamentos específicos sobre productos con aceite de oliva..."*

Con estos antecedentes normativos, y siguiendo con el objetivo de evaluar la aplicabilidad del perfil cromatográfico de fitoesteroles obtenido de acuerdo con la metodología desarrollada en el Capítulo III para la cuantificación de aceite de oliva en alimentos, este capítulo tiene como objetivo desarrollar y validar un modelo PLS [2, 3] para la cuantificación de dicho aceite en salsas vinagretas y aliños comerciales.

Para su desarrollo se han tomado como base herramientas quimiométricas que han sido utilizadas para comparar, discriminar, clasificar y predecir la autenticidad de aceites vegetales [4, 5, 6, 7].

El algoritmo empleado en el PLS ha sido el SIMPLS en lugar del NIPALS ya que este último es más lento y más difícil de interpretar, adicionalmente se ha considerado que dicho algoritmo resulta más adecuado ya que maximiza la covarianza y el NIPALS no lo hace [8, 9].

## VI.1. ACEITES UTILIZADOS PARA LA CREACIÓN DEL MODELO

Para la construcción del modelo se seleccionaron los cincuenta y tres aceites vegetales que aparecen recogidas en la Tabla VI.1 en la que además se muestra el código asignado a cada variedad o tipo de aceite y el número de aceites de cada tipo. Como se puede observar, el modelo se construyó a partir de un 47.2% de aceites de oliva y un 52.8% de aceites vegetales. La procedencia de las muestras se recoge en el Capítulo II de esta Memoria.

**Tabla VI.1.** Aceites utilizados para la construcción del modelo PLS.

Tipo de aceite	Código	Cantidad
Oliva Virgen Extra (arbequina, cornicabra, hojiblanca, picual, manzanilla, royal)	EVOO	13
Oliva Virgen	VOO	4
Oliva Refinado	ROO	4
Orujo de Oliva	POO	4
Girasol	SFO	8
Colza	RO	4
Canola	CanO	2
Soja	SyO	5
Maíz	CO	3
Semillas	SO	6

### VI.I.1. Muestras utilizadas como conjunto o set de calibración:

El conjunto de calibración se diseñó a partir de mezclas binarias de los aceites de oliva virgen extra y aceites vegetales recogidos en la tabla anterior. La composición de las mismas, se varió entre un 25 y un 75% de oliva, además se introdujeron muestras constituidas por 100% de oliva virgen extra y 100% de aceites vegetales (girasol, soja, maíz, colza y canola). En la Tabla VI.2 se detalla la composición individual de cada mezcla o aceite ensayados.

**Tabla VI.2.** Composición las mezclas y aceites del set de calibración.

<u>Tipo de mezclas</u>	<u>Composición</u>	<u>Cantidad</u>
Oliva Virgen Extra: <b>Aceite Vegetal</b>	EVOO: <b>SFO</b> 25:75% (v/v)	3
	EVOO: <b>SFO</b> 50:50% (v/v)	3
	EVOO: <b>SFO</b> 75:25% (v/v)	3
	EVOO: <b>CanO</b> 25:75% (v/v)	3
	EVOO: <b>CanO</b> 50:50% (v/v)	3
	EVOO: <b>CanO</b> 75:25% (v/v)	3
	EVOO: <b>SyO</b> 25:75% (v/v)	3
	EVOO: <b>SyO</b> 50:50% (v/v)	3
	EVOO: <b>SyO</b> 75:25% (v/v)	3
	EVOO: <b>CO</b> 25:75% (v/v)	3
	EVOO: <b>CO</b> 50:50% (v/v)	3
	EVOO: <b>CO</b> 75:25% (v/v)	3
Aceites vegetales (100%)	SFO	3
	CanO	3
	CO	3
	SyO	3
Aceites de oliva virgen extra (100%)	EVOO	12

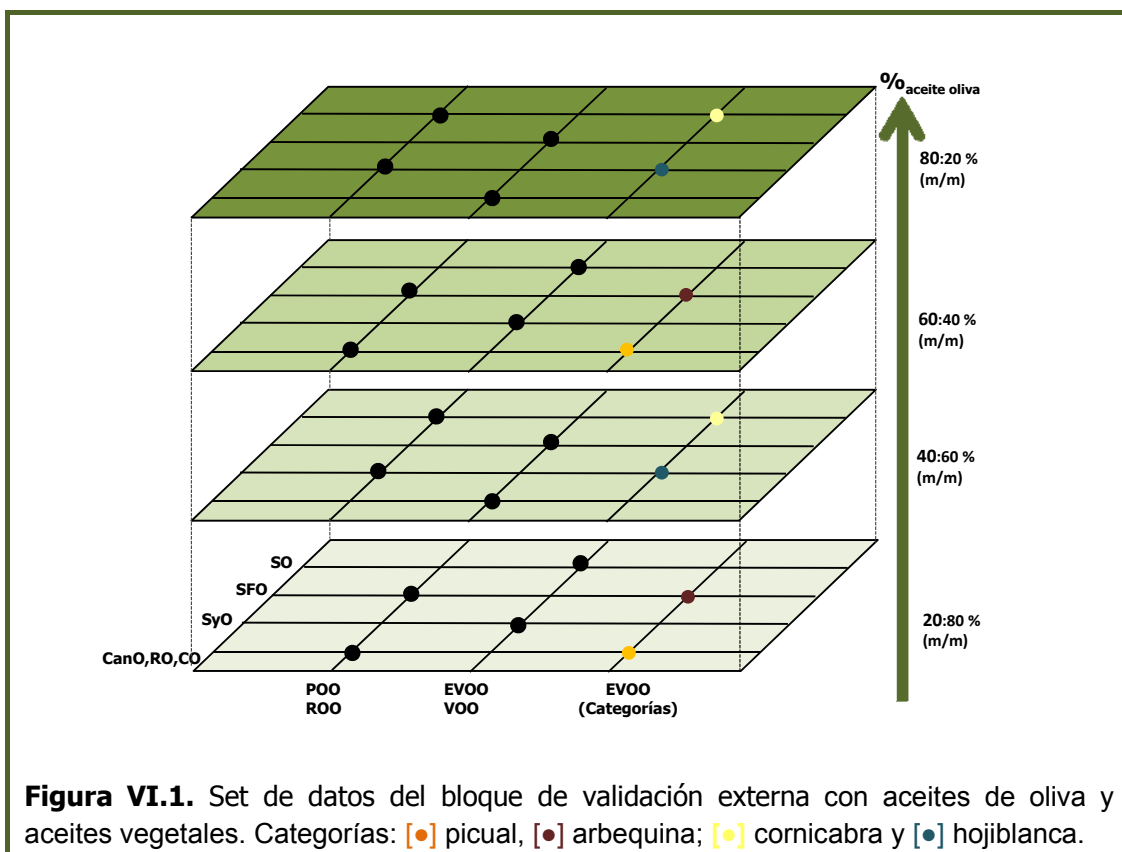
### VI.I.2. Muestras utilizadas como conjunto o set de validación

Para la validación se optó por realizar una validación externa con veinticuatro muestras diferentes a las ensayadas en la calibración. Para diseñar el conjunto de validación se decidió establecer cuatro niveles de porcentaje de aceite de oliva (20%, 40%, 60% y 80%) que de acuerdo con la Figura VI.1 se pueden considerar como pisos del bloque de validación.

A su vez cada piso se diseñó de tal forma que:

(a) En el eje X se establecieron tres niveles: (1) aceite de oliva refinado (ROO) y aceite de orujo de oliva (POO), (2) aceite de oliva virgen extra no monovarietal (EVOO) y oliva virgen (VOO), y (3) aceite de oliva virgen extra monovarietal (picual, arbequina, hojiblanca y cornicabra), y

(b) en el eje Y se consideraron cuatro niveles atendiendo a los diferentes tipos de aceites vegetales utilizados: (1) canola (CanO), colza (RO) y maíz (CO), (2) soja (SyO), (3) girasol (SFO) y (4) semillas (SO).



**Figura VI.1.** Set de datos del bloque de validación externa con aceites de oliva y aceites vegetales. Categorías: [●] picual, [●] arbequina; [●] cornicabra y [●] hojiblanca.

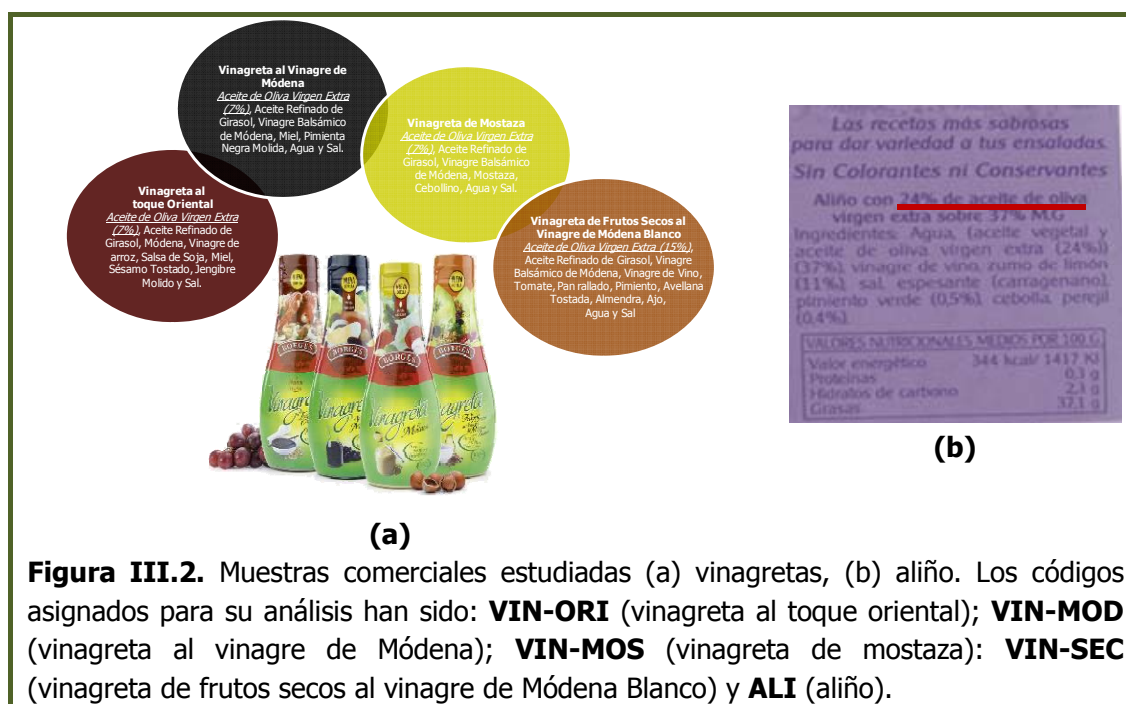


En todos los casos las mezclas binarias se realizaron utilizando siempre aceites diferentes.

### VI.I.3. Muestras utilizadas como conjunto o set de predicción

El set de predicción se diseñó a partir de veinte muestras entre las que se utilizaron materiales de referencia terciarios elaborados en el laboratorio según el procedimiento descrito en el Capítulo II de esta Memoria, cuatro vinagretas comerciales y un aliño.

La información suministrada en la etiqueta de estas muestras se recoge en la Figura III.2. En ella se puede observar que el contenido de aceite de oliva virgen extra declarado en las vinagretas es del 7% o del 15%. En el caso del aliño comercial el contenido declarado es del 24% de aceite de oliva virgen extra sobre un 37% de materia grasa (M.G.) declarada.



**Figura III.2.** Muestras comerciales estudiadas (a) vinagretas, (b) aliño. Los códigos asignados para su análisis han sido: **VIN-ORI** (vinagreta al toque oriental); **VIN-MOD** (vinagreta al vinagre de Módena); **VIN-MOS** (vinagreta de mostaza); **VIN-SEC** (vinagreta de frutos secos al vinagre de Módena Blanco) y **ALI** (aliño).

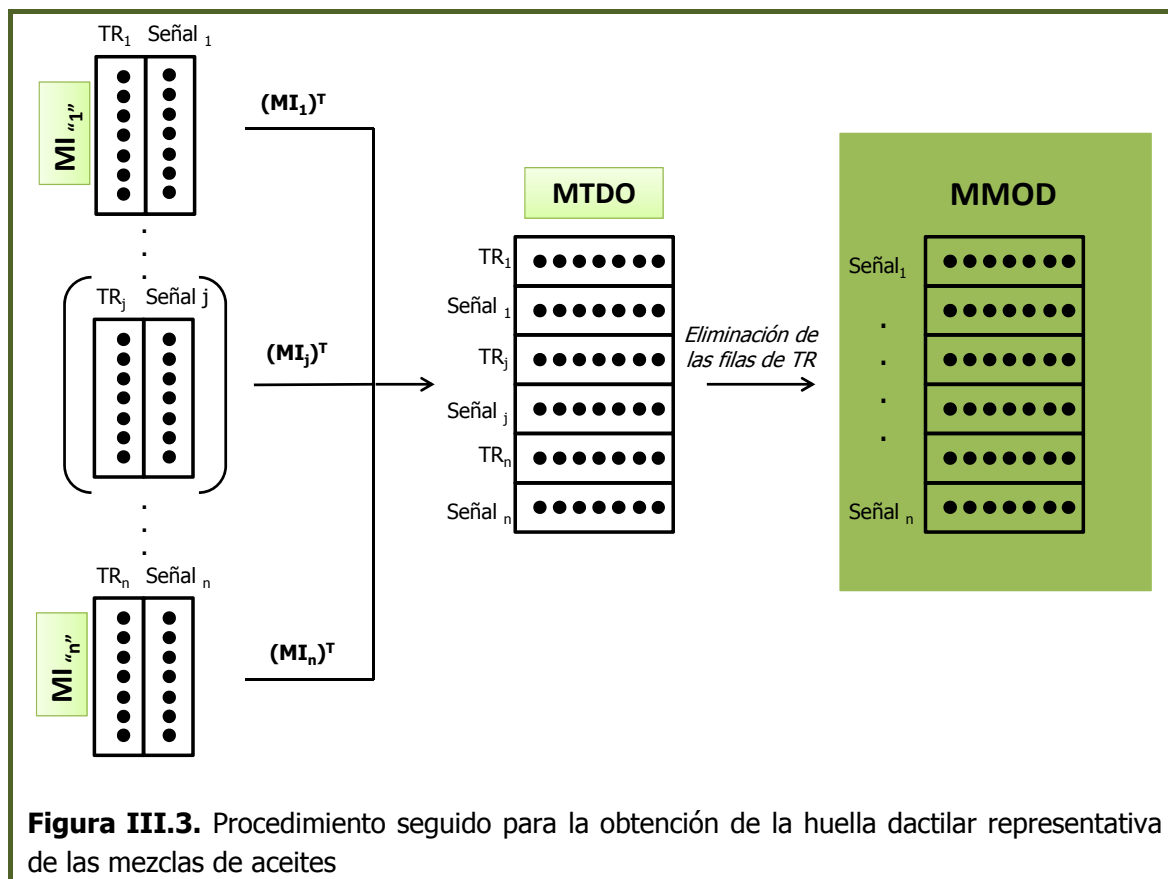
La Tabla VI.3 recoge el código y la composición de los distintos materiales de referencia terciarios elaborados. Aunque aparece descrito en el Capítulo II, indicar que el procedimiento seguido para su elaboración se basó en la receta más habitual que aparece tanto en blogs [10] como en recetarios de cocina [11], según la cual para su elaboración se emplea una parte de vinagre por tres partes de aceite o mezcla de aceites y una pizca de sal.

**Tabla VI.3.** Descripción de las vinagretas empleadas como materiales de referencia terciarios.

<b>Código</b>	<b>Composición (v:v,%)</b>
VRM1	EVOO (100)
VMR2	EVOO (100)
VMR3	EVOO:SFO (50:50)
VMR4	EVOO:SFO (25:75)
VMR5	EVOO:RO (75:25)
VMR6	EVOO:SyO (60:40)
VMR7	EVOO:CO (90:10)
VMR8	EVOO:CanO (20:80)
VMR9	EVOO:CO (40:60)
VMR10	EVOO:SyO (10:90)
VMR11	CanO:SyO (50:50)
VMR12	SO: SFO (75:25)
VMR13	CO:SFO (25:75)
VMR14	POO:SyO (60:40)
VMR15	POO:SFO (80:20)

## VI.2.DISEÑO PARA LA OBTENCIÓN DE LAS HUELLAS DACTILARES CROMATOGRÁFICAS

La Figura VI.3 recoge de forma resumida el procedimiento seguido para la obtención de la huella dactilar cromatográfica a emplear en el modelo de cuantificación [12].



**Figura III.3.** Procedimiento seguido para la obtención de la huella dactilar representativa de las mezclas de aceites

Inicialmente, para cada comatograma de GC-FID de aceite o mezcla de aceites obtenido mediante el método HPLC-GC-FID off line desarrollado, se recogió una **matriz inicial**<sub>muestra "i"</sub> ( $MI_{"i"}$ ) (matriz del cromatograma registrado para la muestra) de dos columnas, la primera constituida por los valores crecientes del tiempo de análisis o valores de tiempos registrados (TR) y la segunda constituida por la señal cromatográfica registrada a esos tiempos. Para poder operar con el software, dichas matrices se transpusieron antes de ser combinadas para generar la denominada **matriz total de datos originales** (MTDO).

En este caso, y a partir de lo establecido en el Capítulo V de esta Memoria, dicha matriz estaba constituida por 116 filas 18201 columnas, en las que, en las filas y de forma alternativa, aparecían los TR y las señales cromatográficas de las 58 muestras analizadas.

Finalmente, se construyó una nueva matriz denominada **matriz para el modelado** (MMOD) en la que las variables finales seleccionadas fueron las señales cromatográficas una vez que se eliminaron las filas de los TR. Esta matriz se utilizó como matriz base para realizar el modelo de cuantificación de aceite de oliva en salsas y aliños comerciales.

Antes de seguir con el establecimiento del modelo de cuantificación se observó que al superponer varios perfiles cromatográficos obtenidos tratando diferentes cantidades de muestra en la etapa de saponificación (1 o 5 gramos de aceite o mezcla binaria de aceites), y tras seguir el método analítico desarrollado, el perfil cromatográfico era dependiente del tamaño de la muestra.

Con objeto de resolver este problema, sin tener que limitar el procedimiento a un tamaño de muestra fijo, se optó por realizar una normalización interna de los perfiles cromatográficos dividiendo cada una de las variables de la huella dactilar por la variable de mayor valor del perfil, la cual aparece al tiempo de retención del  $\beta$ -sitosterol, fitoesterol más abundante en todos los aceites vegetales, incluido el aceite de oliva.

De esta forma se consiguió una **huella dactilar normalizada** (todas las variables alcanzan un valor menor o igual a uno en el caso del  $\beta$ -sitosterol) para cada muestra analizada. Tras esta normalización, y para posteriores estudios, **la MMOD final se consideró constituida por las huellas dactilares normalizadas de las muestras utilizadas en cada caso (MMODNOR).**

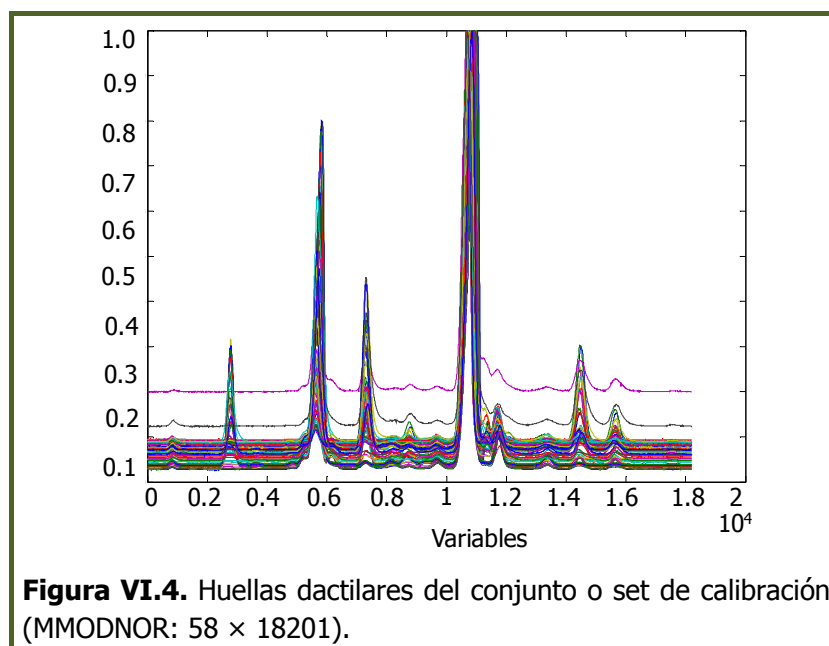
### VI.3. PRE-PROCESADO DE LOS DATOS

Como ya se ha comentado en el capítulo anterior, el pre-procesado de los datos es la etapa previa al tratamiento de los mismos para alinear las respuestas y eliminar fuentes extrañas de variabilidad, siendo necesario en casi todos los casos ya que de acuerdo con lo establecido por Wise et al. [13] en el Manual de manejo del PLS\_Toolbox 4.0<sup>®</sup> un número excesivo de fuentes extrañas de variabilidad puede incrementar la dificultad en la estimación del modelo y además puede entorpecer el poder aislar la varianza de interés. El pre-procesamiento llevado a cabo en todos los modelos desarrollados en este capítulo ha sido el mismo que el aplicado en el modelo clasificatorio desarrollado en el Capítulo V de esta Memoria, y que ha constado de las siguientes etapas:

1. Alineamiento de la señal mediante el algoritmo ICOSHIFT definiendo 5 intervalos [14,15].
2. Ajuste de la línea base mediante un algoritmo de segundo orden.
3. Centrado de media.

#### VI.4. PRIMER MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES

En la Figura VI.4 se muestra la superposición de las diferentes huellas dactilares normalizadas obtenidas para el conjunto de calibración utilizado en la creación del modelo.

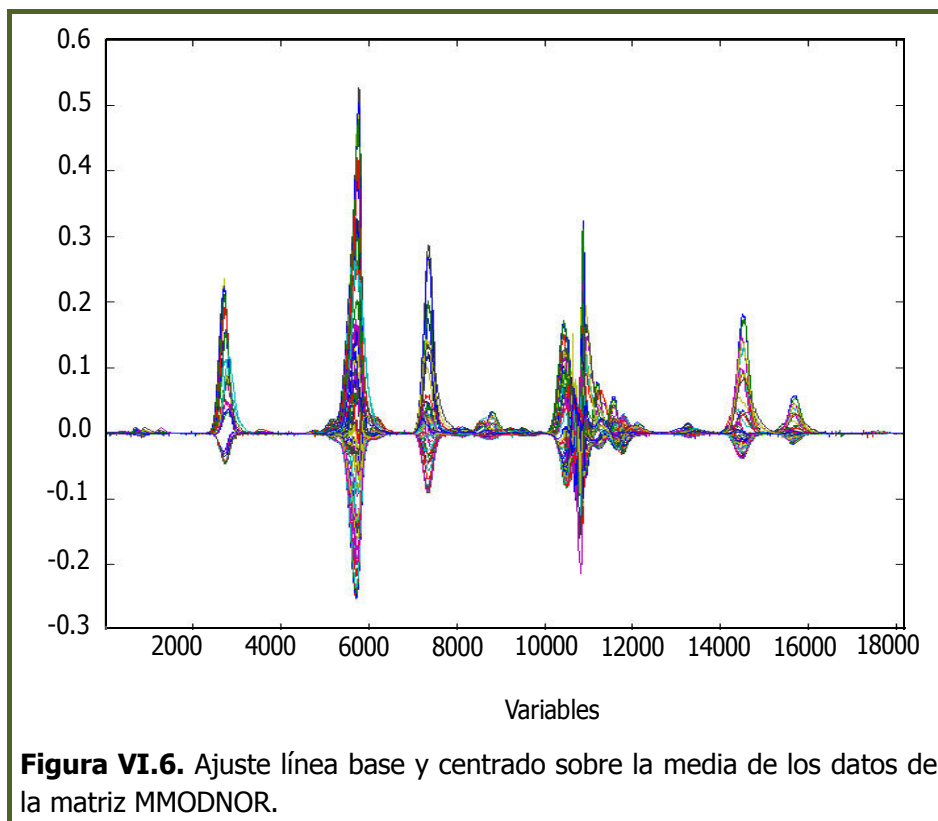
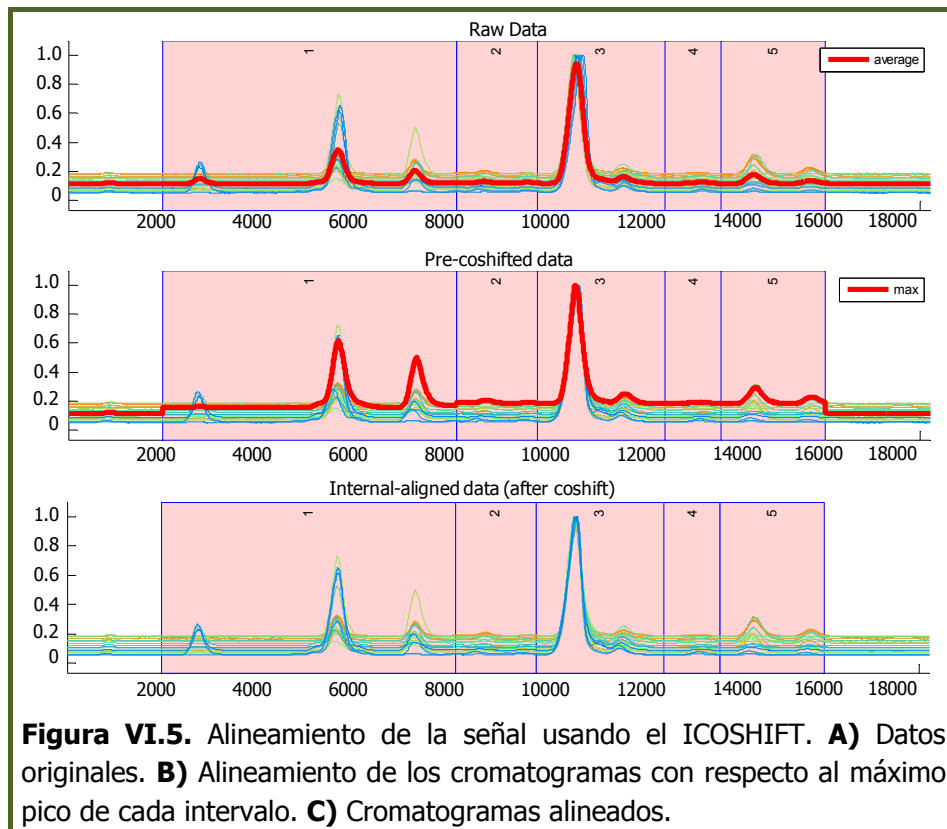


**Figura VI.4.** Huellas dactilares del conjunto o set de calibración (MMODNOR: 58 × 18201).

El preprocesado de las mismas se muestra en las Figuras VI.6 y VI.7 en las que se muestran respectivamente el alineamiento de la señal mediante el algoritmo ICOSHIFT para los cinco intervalos considerados, el ajuste de la línea base mediante el algoritmo de mínimos cuadrados ponderados de segundo orden (WLS) y el centrado sobre la media.

En la elaboración del modelo PLS se empleó el algoritmo SIMPLS para encontrar el número de variables latentes significativas y validar el modelo mediante una validación cruzada tipo "Venetian Blinds". Los bloques establecidos para el modelado de calibración XY fueron:

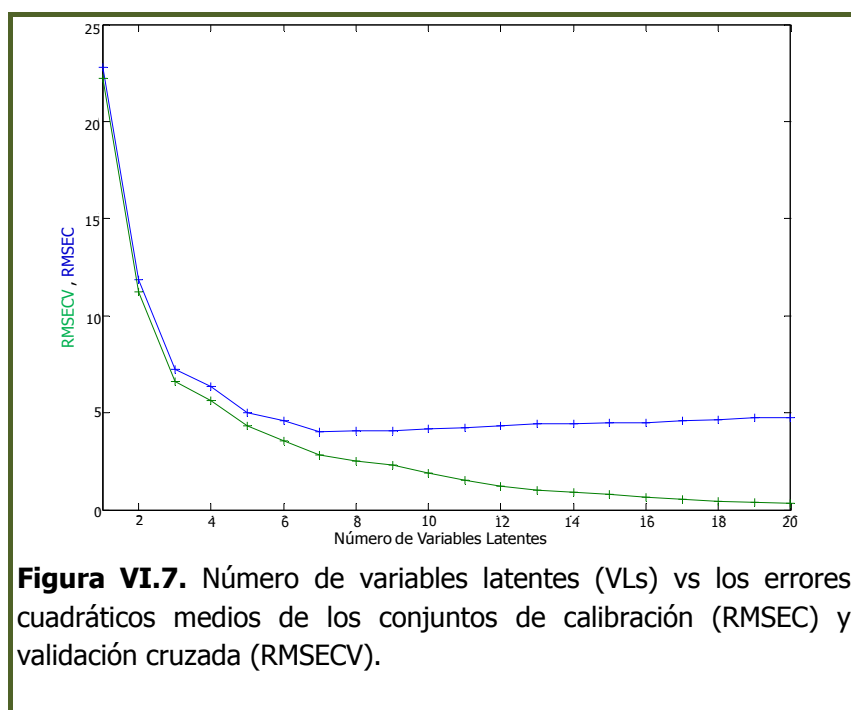
- Bloque de variables de entrada o bloque X: Matriz MMODNOR.
- Bloque de variables de salida o bloque Y: Matriz de porcentajes de aceite de oliva utilizados en las mezclas ensayadas para la cuantificación (correspondería a una matriz en la que solo se identifica cada nivel por su porcentaje de aceite de oliva, independientemente del aceite vegetal utilizado en la mezcla binaria. Los niveles seleccionados aparecen recogidos en la segunda columna de la Tabla VI.2.)



## Cuantificación de Aceite de Oliva en Vinagretas

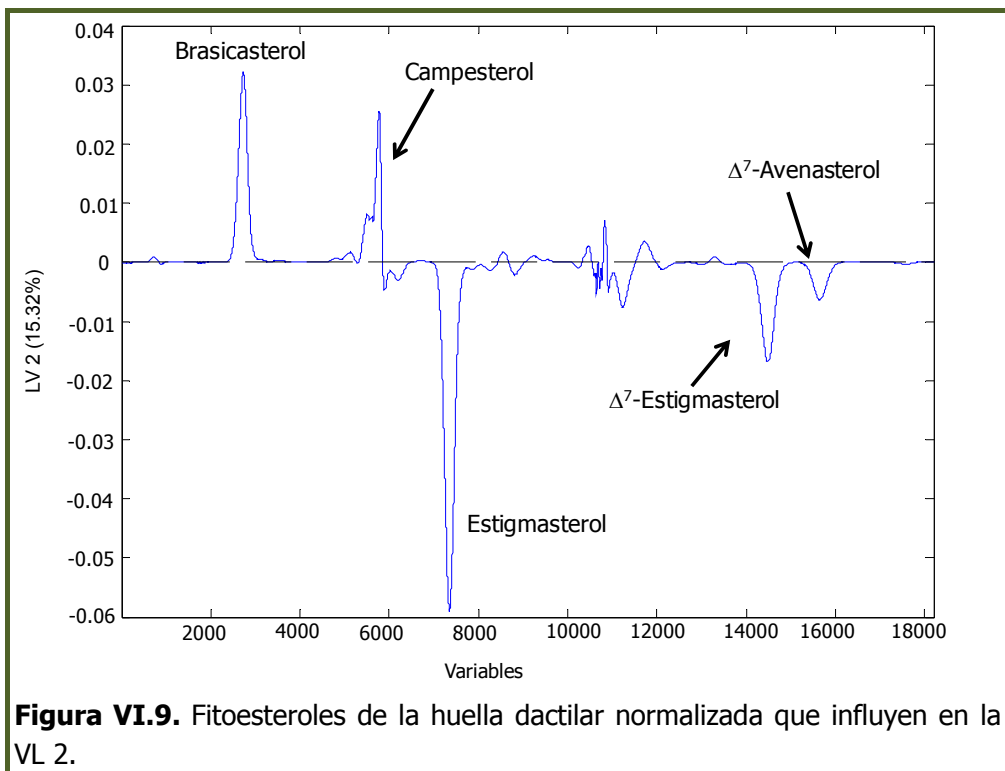
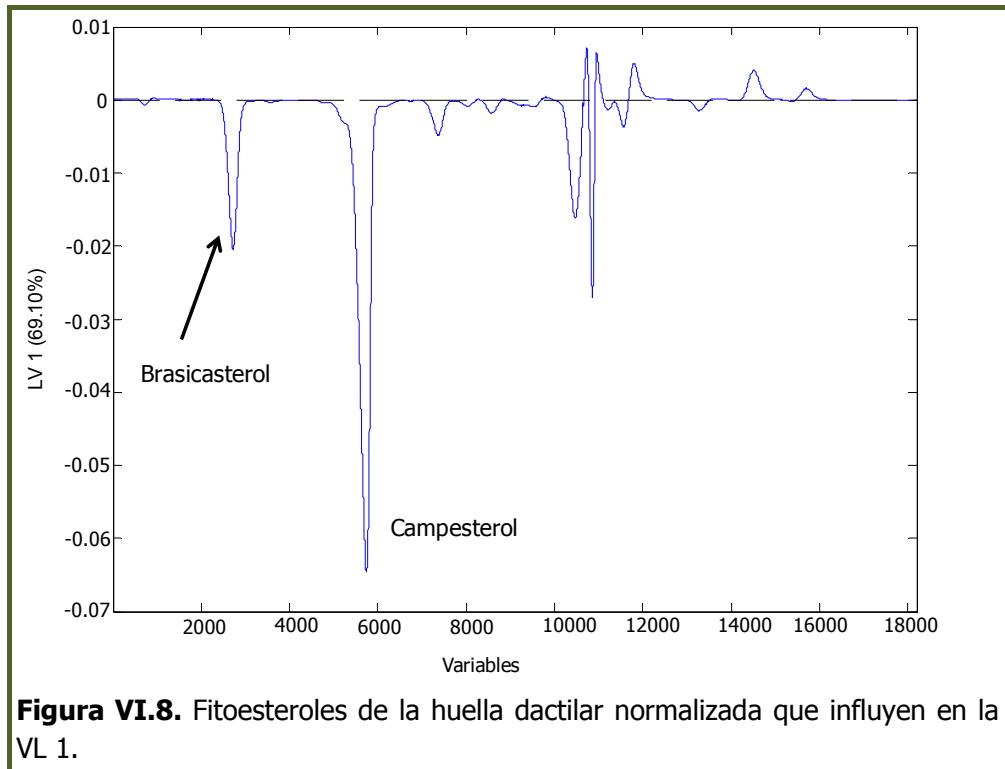
El número óptimo de variables latentes (VLs) se seleccionó considerando los gráficos de sedimentación para los errores cuadráticos medios de los conjuntos de calibración (RMSEC) y validación cruzada (RMSECV) realizada con el set de validación. Como se puede observar en la Figura VI. 7 el número seleccionado fue de 3, considerando además que seleccionar un número mayor no mejoraría sustancialmente la varianza acumulada.

Con estas tres variables latentes se explica un un 90.87% de la varianza del Bloque X y un 96.36% de la varianza del Bloque Y. Dado el número de muestras utilizadas en el set de calibración, se necesitaron 7 splits para el tipo de validación cruzada seleccionado.



A continuación se procedió a evaluar el peso o carga (en inglés "loadings") que los distintos fitoesteroles presentes en las huellas dactilares normalizadas ejercen sobre las variables latentes seleccionadas.

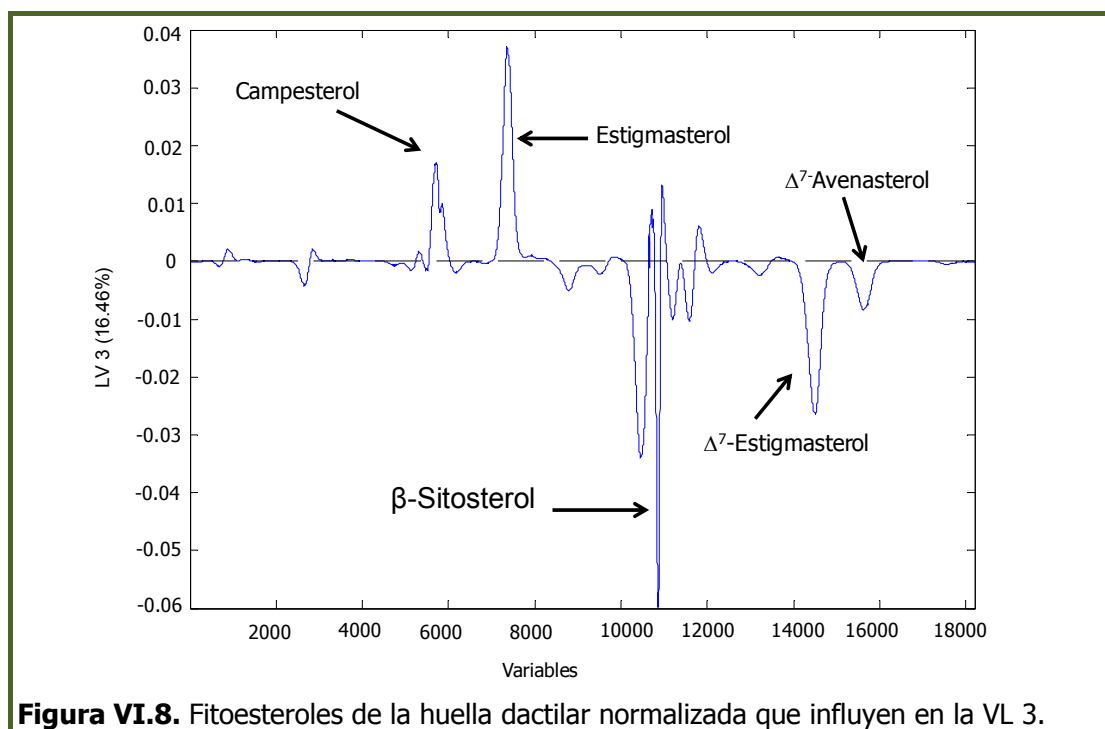
En la Figura VI.8 se puede observar que la VL1 está influenciada negativamente por brasicasterol y campesterol (mezclas con menor contenido en dichos fitoesteroles), en la figura VI.9 se puede observar que la VL2 está influenciada positivamente por el brasicasterol y el campesterol y negativamente por la parte de la huella dactilar correspondiente a estigmasterol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.



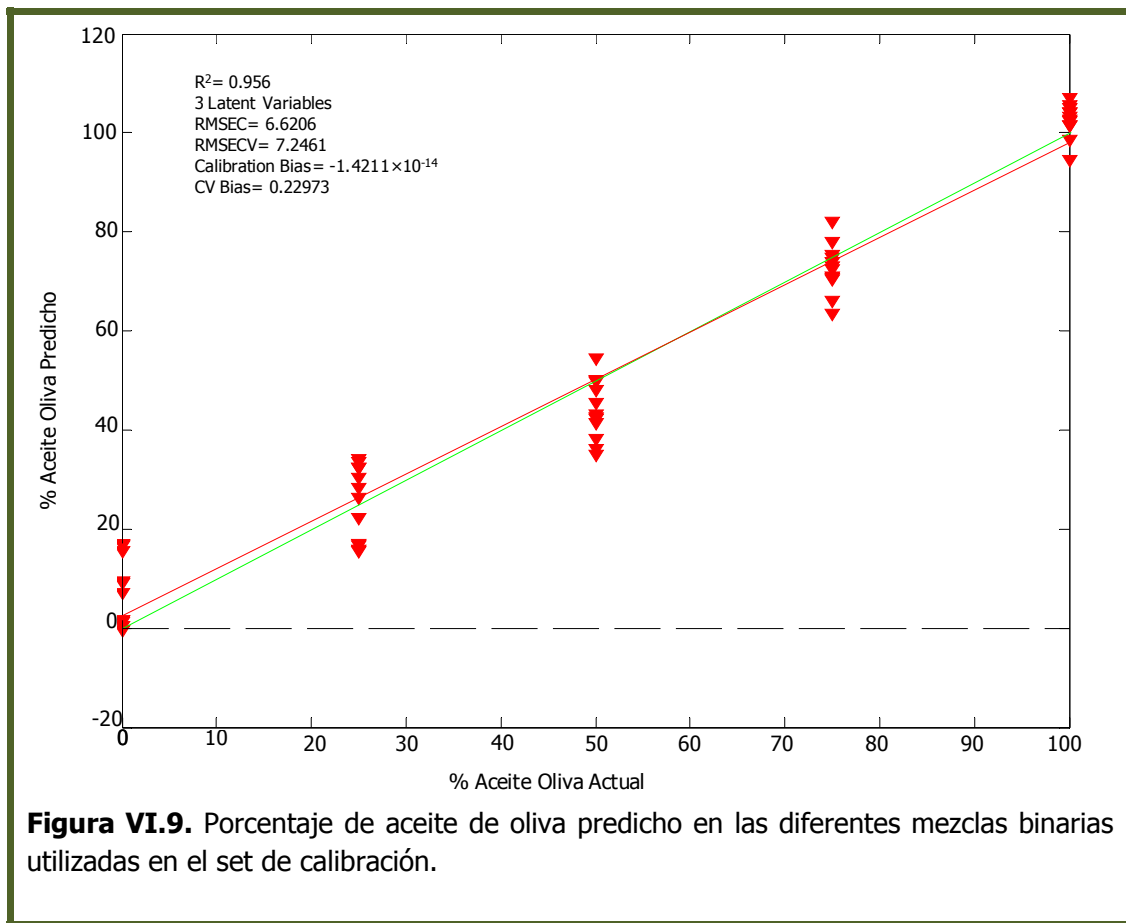


## Cuantificación de Aceite de Oliva en Vinagretas

Finalmente, en la Figura VI.10 se puede observar que la VL 3 está influenciada positivamente por la parte de la huella dactilar en la que se encuentran el campesterol y el estigmasterol y negativamente por la zona de la huella dactilar en la que se encuentran el  $\Delta^7$ -estigmasterol y el  $\Delta^7$ -avenasterol (zona terminal de la huella), y por la zona del  $\beta$ -sitosterol. Hecho este último importante, ya que permite confirmar que aunque es un fitoesterol mayoritario en todos los aceites, las diferencias en su contenido permitirían diferenciar el contenido en aceite de oliva en una mezcla.



La Figura VI.9 recoge la representación de los porcentajes de aceite de oliva predichos frente a los porcentajes reales. Como se puede observar los valores obtenidos para  $R^2(0.956)$  y  $RMSECV(7.246)$  indican que la capacidad de predicción del modelo es adecuada cuando la validación cruzada se realiza con muestras del propio set de calibración, así, el modelo establecido permitiría cuantificar aceite de oliva en mezclas binarias con otros aceites vegetales utilizando las huellas dactilares cromatográficas normalizadas obtenidas en mezclas cuya composición varía entre 0% de aceite de oliva (muestras de aceites vegetales individuales tales como girasol, soja, canola, colza y maíz) y 100% de aceite de oliva virgen extra. Siendo el rango ensayado para las mezclas el comprendido entre 20 y 80% de aceite de oliva en la mezcla.

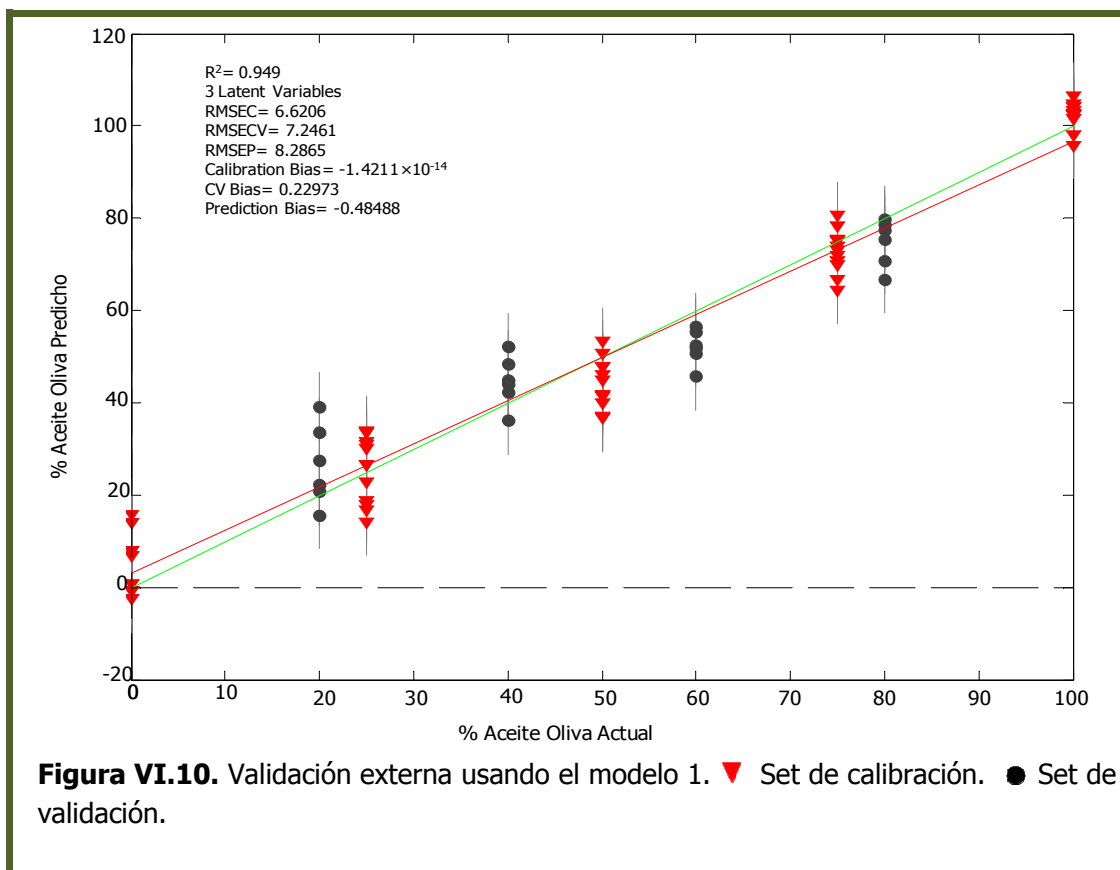


**Figura VI.9.** Porcentaje de aceite de oliva predicho en las diferentes mezclas binarias utilizadas en el set de calibración.

A continuación, el modelo se validó de forma externa, utilizando para ello un set de validación de veinticuatro mezclas no incluidas en el entrenamiento del modelo. Como se ha comentado al inicio del capítulo, estas muestras se elaboraron de acuerdo con el Bloque de validación externa diseñado para tal fin.

La Figura VI.10 muestra los resultados obtenidos. Como se puede observar los valores obtenidos en este caso para  $R^2(0.949)$  y  $RMSECV(8.286)$  indican que la capacidad de predicción del modelo es adecuada cuando la validación cruzada se realiza con muestras externas al set de calibración, si bien es verdad que ambos parámetros empeoran ligeramente con respecto a la validación cruzada interna realizada previamente. No obstante el valor de  $RMSEP$  inferior a un 10% indicaría que el modelo establecido permitiría cuantificar aceite de oliva en mezclas binarias de composición desconocida.

## Cuantificación de Aceite de Oliva en Vinagretas



Los datos obtenidos como resultado del entrenamiento y validación del modelo se recogen en la Tabla VI.4.

**Tabla VI.4.** Resumen de los parámetros más importantes del modelo 1 desarrollado para la determinación de aceite de oliva en mezclas binarias.

VL	$R^2_{Cal}$	RMSEC	$R^2_{cv}$	RMSECV	$R^2_p$	RMSEP	Bias
3	0.9636	6.6206	0.9564	7.2461	0.8787	8.2865	$-1.4 \times 10^{-14}$

Una vez conocido el error de predicción del modelo se procedió a su aplicación en muestras reales (vinagretas y aliños) analizando por una parte materiales de referencia terciarios elaborados con mezclas de aceites, siguiendo la receta habitual empleada en la elaboración de este tipo de salsas y por otra parte diversas vinagretas y aliño comerciales. Los resultados se recogen en la Tabla VI.5.

**Tabla VI.5.** Cuantificación de aceite de oliva en vinagretas.

Muestras	Código	% (m/m) de aceite de oliva	% (m/m) aceite de oliva predicho
	VRM1	100	102.31
	VMR2	100	107.51
	VMR3	50	55.18
	VMR4	25	39.98
	VMR5	75	69.36
	VMR6	60	57.16
Materiales de Referencia Terciarios	VMR7	90	84.22
	VMR8	20	42.02
	VMR9	40	37.50
	VMR10	10	8.97
	VMR11	0	28.51
	VMR12	0	31.91
	VMR13	0	25.65
	VMR14	60	59.30
	VMR15	80	64.89
	ALI	24	59.24
Vinagretas Comerciales	VIN-ORI	7	39.91
	VIN-SEC	7	39.99
	VIN-MOD	15	38.27
	VIN-MOS	7	35.10

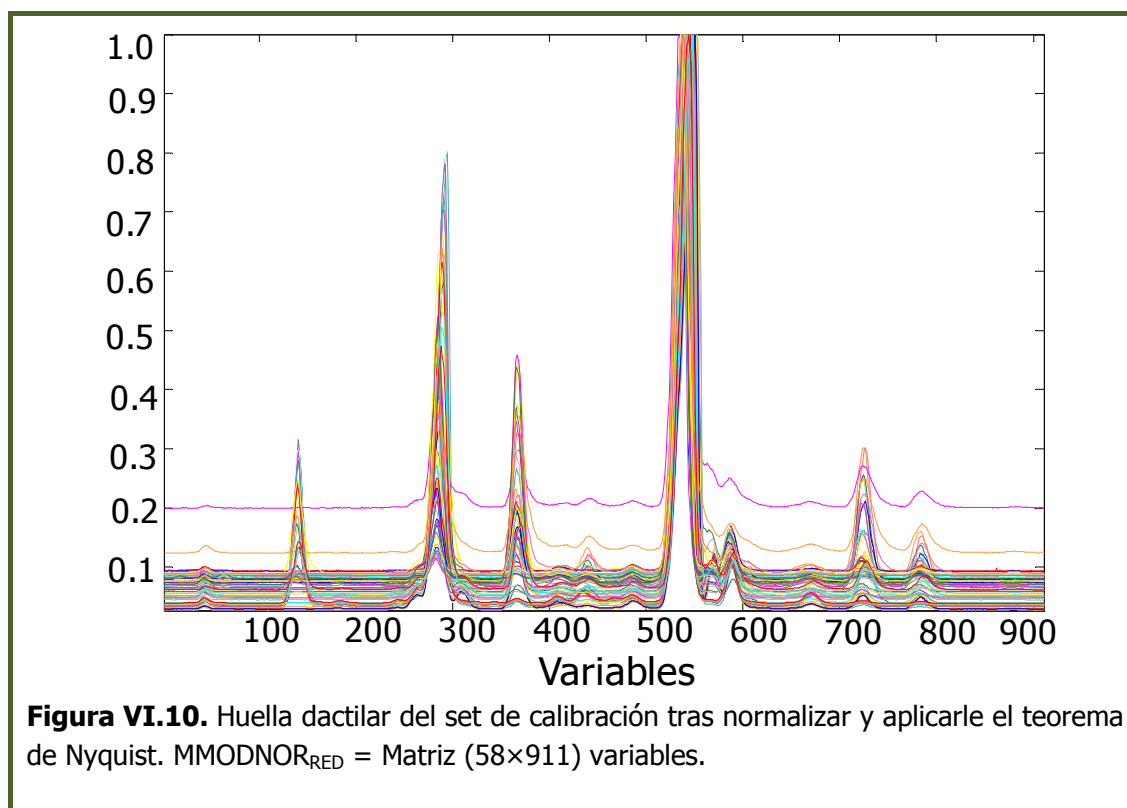
Como se puede observar, en el caso de las vinagretas elaboradas en el laboratorio (materiales de referencia terciarios) el modelo predice de forma adecuada el 60% de las mismas. Destacando que para el 40% restante, el modelo no predice bien en el caso de que la vinagreta contenga una mezcla de aceites vegetales y no contenga aceite de oliva (VMR11, VMR12 y VMR13). Este hecho puede atribuirse a que en el set de calibración no se contempló la posibilidad de incluir mezclas binarias para el 0% de aceite de oliva, y dado que el perfil de esteroides de este último está incluido en los perfiles de esteroides de los aceites vegetales, el modelo puede confundirlo.

En el caso de las salsas comerciales, el modelo predice mal en el 100% de los casos. Aquí, la razón puede encontrarse en la composición de las salsas ya que según se especifica en la etiqueta además de aceite de oliva, contienen semillas y trazas de frutos secos (almendras, avellanas, etc.) que pueden interferir en el perfil cromatográfico de los TMS-esteroides. Este hecho se justifica porque en todos los

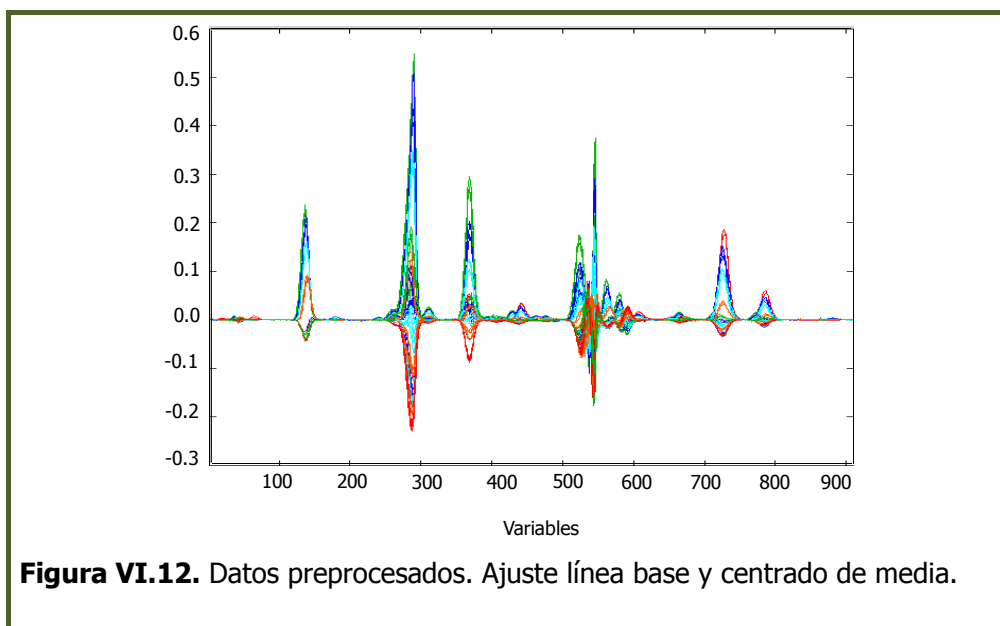
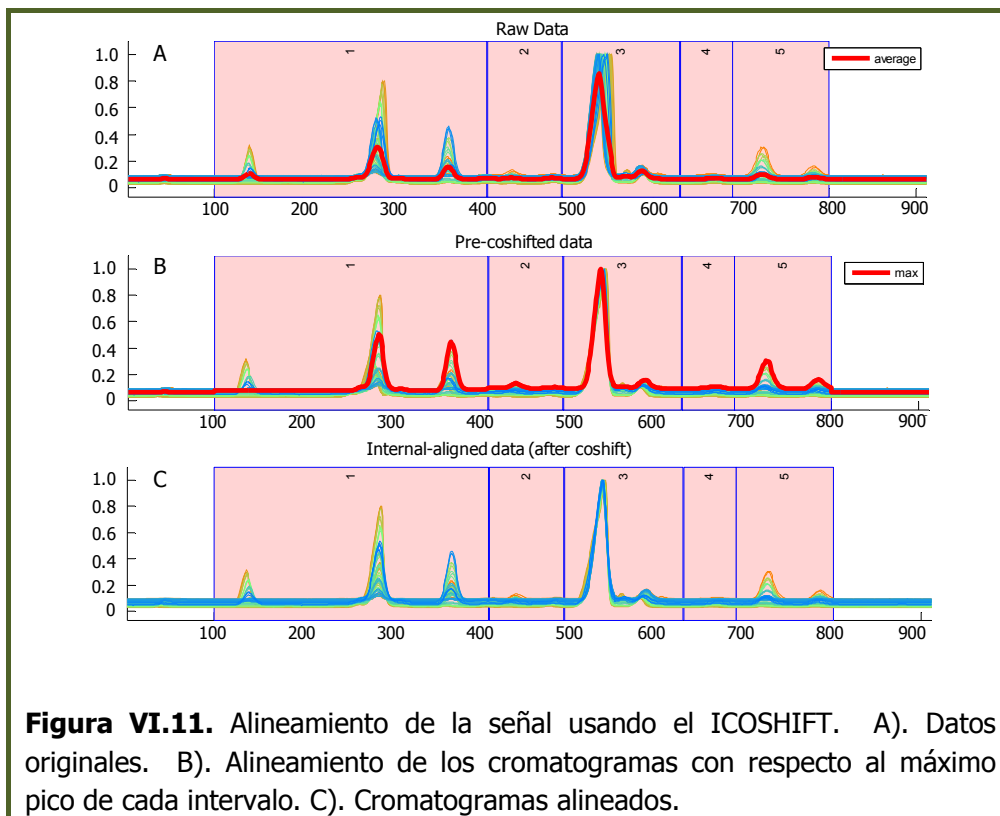
casos el valor encontrado ha sido siempre mayor que el declarado por el fabricante.

### VI.5. SEGUNDO MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES

Con objeto de mejorar el modelo anterior y a la vez facilitar el cálculo, se pensó en la posibilidad de reducir el número de variables normalizadas. Para tal fin se realizó un muestreo de la señal de acuerdo con lo establecido en el Teorema de Nyquist [16, 17]. Tras probar diferentes secuencias de muestreo de la señal (1 de cada 10, de cada 20 o de cada 30), los mejores resultados se obtuvieron cuando sobre la matriz MMODNOR se eliminaron 19 de cada veinte variables (muestreo 1 de 20) con lo que esta matriz se redujo de 58x18201 a una matriz MMODNOR<sub>RED</sub> de 58x911 variables sin una pérdida sustancial de información. La Figura VI.10 muestra la superposición de las diferentes huellas dactilares normalizadas obtenidas para el conjunto de calibración utilizado una vez realizado el muestreo.



El preprocesado realizado antes de desarrollar el modelo (alineamiento con ICOSHIFT, ajuste de la línea base y centrado de media) se muestra en las Figuras VI.11 y VI.12.

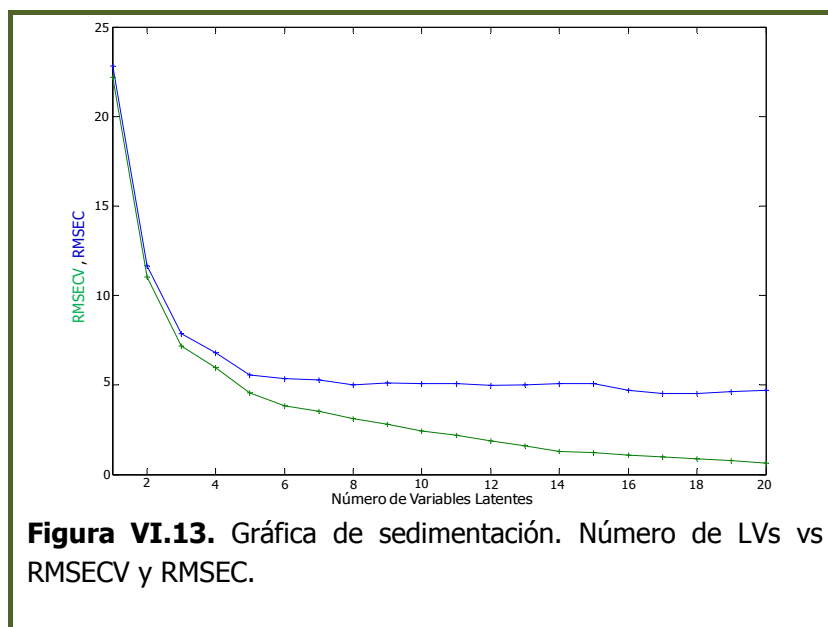


## Cuantificación de Aceite de Oliva en Vinagretas

En la elaboración del segundo modelo PLS (Modelo 2) se empleó el algoritmo SIMPLS para encontrar el número de variables latentes significativas y validar el modelo mediante una validación cruzada tipo "Venetian Blinds". Los bloques establecidos para el modelado de calibración XY fueron:

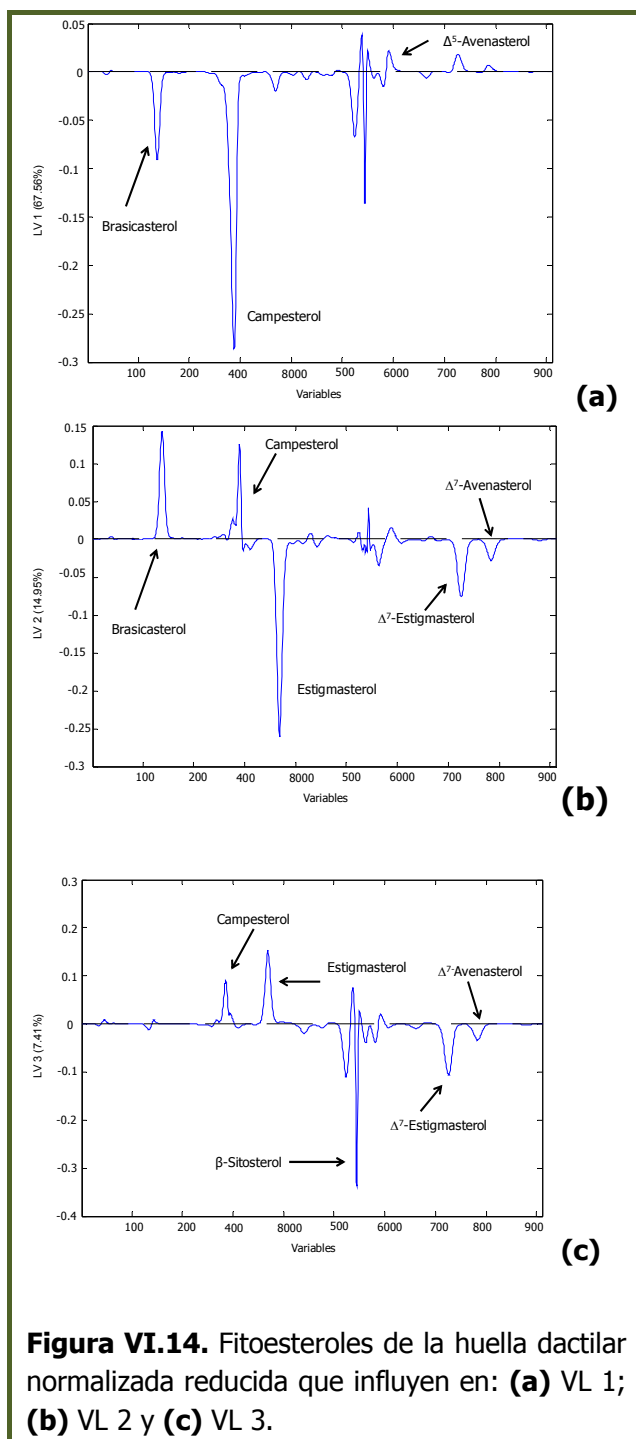
- Bloque de variables de entrada o bloque X: Matriz  $MMODNOR_{RED}$ .
- Bloque de variables de salida o bloque Y: Matriz de porcentajes de aceite de oliva utilizados en las mezclas ensayadas para la cuantificación. Esta matriz es la misma que la utilizada para el Modelo 1.

Al igual que en el Modelo 1, el número óptimo de variables latentes (VLs) se seleccionó considerando los gráficos de sedimentación para los errores cuadráticos medios de los conjuntos de calibración (RMSEC) y validación cruzada (RMSECV) realizada con el set de validación. Como se puede observar en la Figura VI.13, el número óptimo de VLs fue también de 3. En este caso, estas tres variables latentes explican un 89.91% de la varianza del Bloque X y un 95.71% de la varianza del Bloque Y. Al igual que en el modelo anterior, se necesitaron 7 splits para el tipo de validación cruzada seleccionado.



**Figura VI.13.** Gráfica de sedimentación. Número de LVs vs RMSECV y RMSEC.

A continuación se procedió a evaluar el peso o carga que los distintos fitoesteroles presentes en las huellas dactilares normalizadas ejercen sobre las variables latentes seleccionadas.



Se puede observar que, al igual que en el Modelo 1:

**(a)** La VL 1 está influenciada negativamente por brasicasterol y campesterol.

**(b)** La VL 2 está influenciada positivamente por el brasicasterol y el campesterol y negativamente por la parte de la huella dactilar correspondiente a estigmasterol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.

**(c)** La VL 3 está influenciada positivamente por la parte de la huella dactilar en la que se encuentran el campesterol y el estigmasterol y negativamente por la zona de la huella dactilar en la que se encuentran el  $\Delta^7$ -estigmasterol y el  $\Delta^7$ -avenasterol (zona terminal de la huella), y por la zona del  $\beta$ -sitosterol.

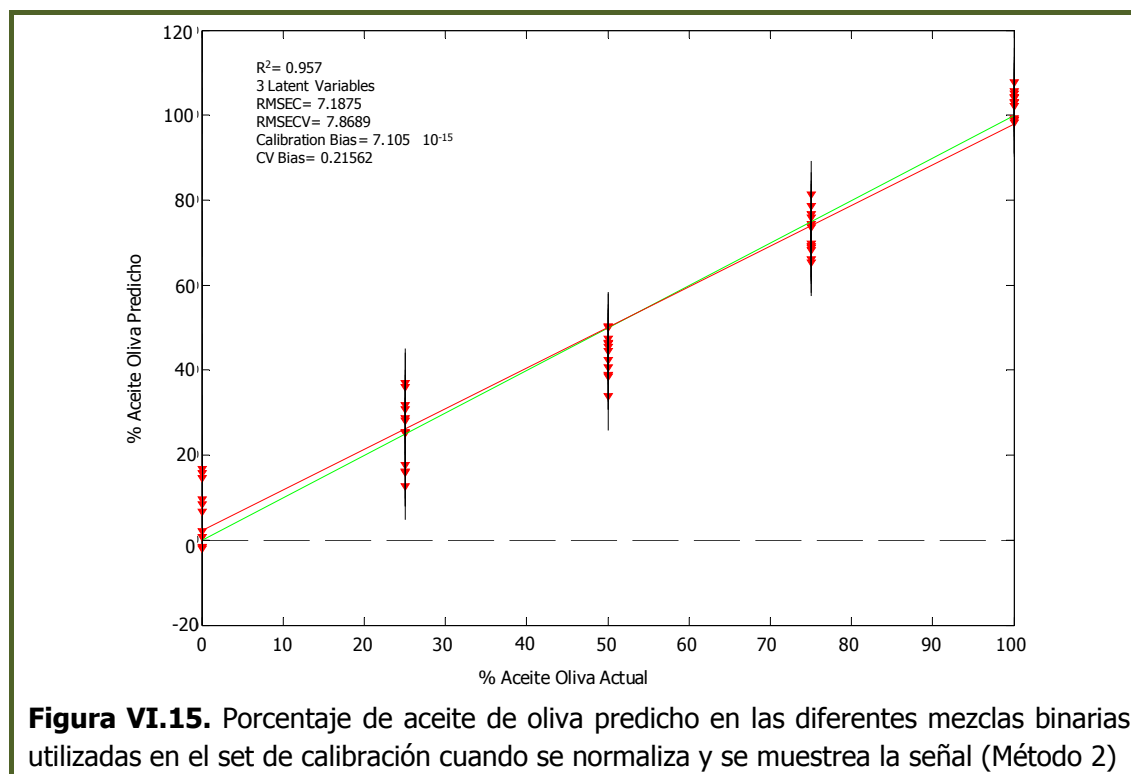
En todos los casos el peso de cada "huella individual" asociada a los fitoesteroles presenta un valor más pequeño en comparación con el mostrado en el modelo anterior.

La Figura VI.15 recoge la representación de los porcentajes de aceite de oliva predichos frente a los porcentajes reales en el caso de la validación cruzada realizada sobre el set de calibración. Como se puede observar los valores obtenidos para  $R^2(0.957)$  y  $RMSECV(7.869)$  indican que la capacidad de predicción del modelo es también adecuada, hay que resaltar que el suavizado de la señal provoca una ligera mejora en la correlación/linealidad del modelo y empeora



## Cuantificación de Aceite de Oliva en Vinagretas

ligeramente el error cuadrático medio para la validación cruzada con el set de calibración.



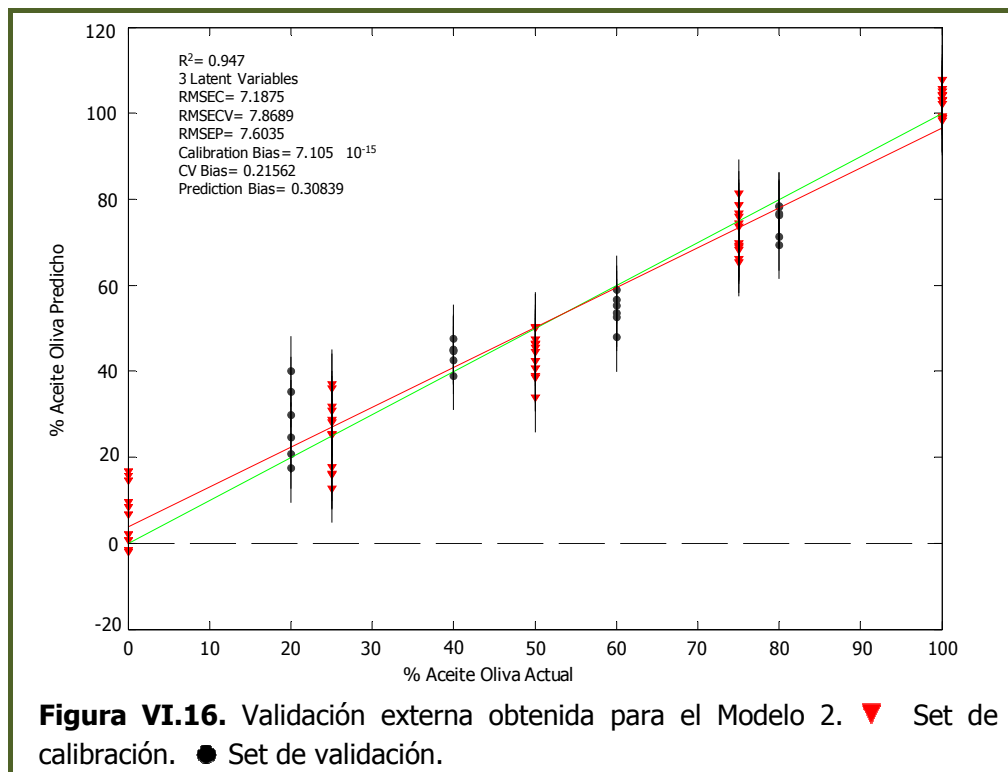
**Figura VI.15.** Porcentaje de aceite de oliva predicho en las diferentes mezclas binarias utilizadas en el set de calibración cuando se normaliza y se muestra la señal (Método 2)

A su vez, la Figura VI.16 muestra la representación de los porcentajes de aceite de oliva predichos frente a los porcentajes reales en el caso de la validación cruzada realizada sobre el set de validación. Como se puede observar los valores obtenidos para  $R^2(0.9103)$  y  $RMSECV(7.6035)$  indican que la capacidad de predicción del modelo es también adecuada.

Los datos obtenidos como resultado del entrenamiento y validación del método se recogen en la Tabla VI.6.

**Tabla VI.6.** Resumen de los parámetros más importantes del modelo desarrollado PLS para la determinación de aceite de oliva en muestras de vinagretas.

VL	$R^2_{Cal}$	RMSEC	$R^2_{cv}$	RMSECV	$R^2_p$	RMSEP	Bias
3	0.9571	7.1875	0.9486	7.8690	0.9103	7.6035	$7.1 \times 10^{-15}$



Por último, y de forma análoga a la seguida en el caso anterior, se aplicó el modelo a los perfiles normalizados y suavizados de las muestras que constituían el set de predicción de muestras reales (materiales de referencia terciarios, vinagretas y aliños comerciales). Los resultados obtenidos se recogen en la Tabla VI.7.

Como se puede observar, los resultados encontrados al aplicar este modelo son equivalentes a los obtenidos con el modelo anterior, sin embargo, aunque se obtiene el mismo porcentaje de aciertos en la predicción (60% en el caso de las vinagretas elaboradas en el laboratorio) y a su vez predice con una diferencia de  $\pm 1\%$ , el tiempo invertido en la obtención del resultado es menor.

Al igual que el modelo anterior, en el caso de las salsas comerciales, este modelo actual predice mal en el 100% de los casos.

**Tabla VI.7.** Cuantificación de aceite de oliva en vinagretas.

Muestras	Código	% (m/m) de aceite de oliva	% (m/m) aceite de oliva predicho	
	VRM1	100	102.70	
	VMR2	100	107.37	
	VMR3	50	58.8	
	VMR4	25	41.79	
	VMR5	75	70.15	
	VMR6	60	55.69	
Materiales de Referencia Terciarios	VMR7	90	86.93	
	VMR8	20	44.74	
	VMR9	40	44.12	
	VMR10	10	8.06	
	VMR11	0	31.00	
	VMR12	0	32.93	
	VMR13	0	28.69	
	VMR14	60	62.05	
	VMR15	80	67.15	
		ALI	24	61.28
	Vinagretas Comerciales	VIN-ORI	7	42.38
		VIN-SEC	7	42.77
		VIN-MOD	15	39.49
		VIN-MOS	7	36.21

### VI.5. TERCER MODELO PARA LA CUANTIFICACIÓN DE ACEITE DE OLIVA EN SALSAS Y ALIÑOS COMERCIALES

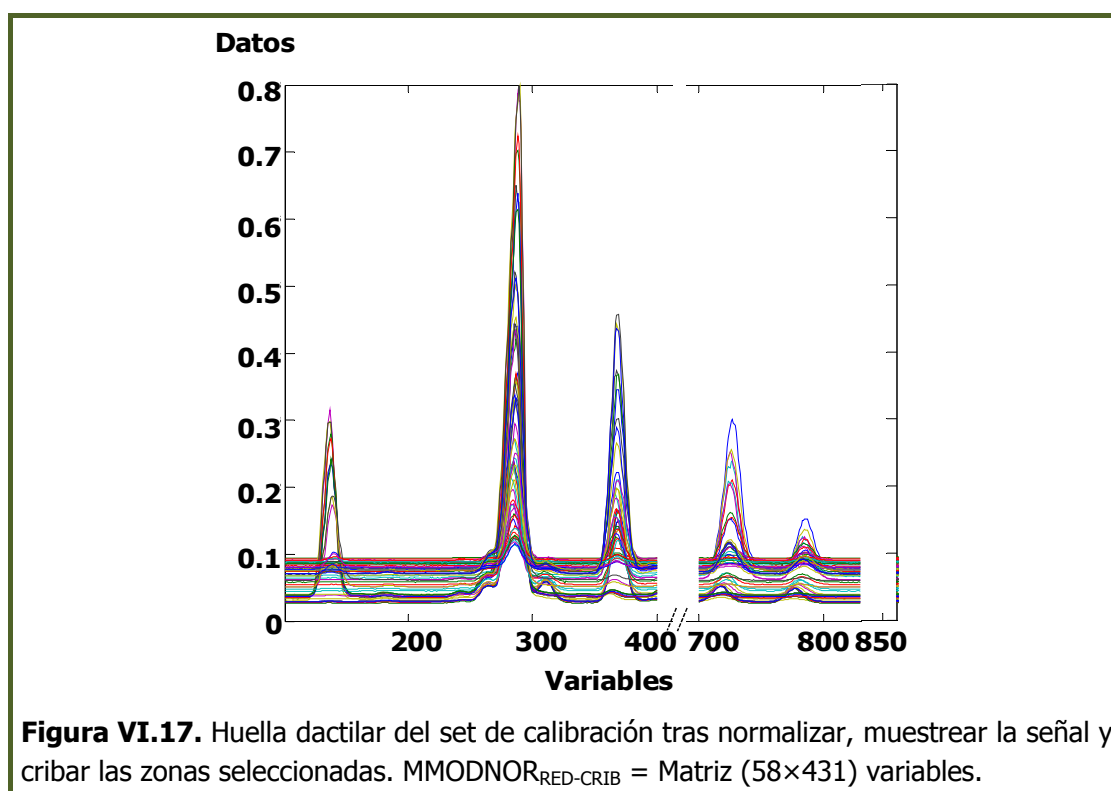
Dado que la reducción de la dimensión de la matriz de datos aplicada en el Modelo 2 no introducía una mejora sustancial a la hora de intentar cuantificar el contenido de aceite de oliva en productos alimenticios tipo salsas y aliños, se decidió establecer otro modelo de cuantificación mediante PLS (Modelo 3).

Para ello, tomando como partida la matriz  $MNOR_{RED}$  definida tras el suavizado de la señal cromatográfica una vez realizado el muestreo, y aplicado el algoritmo ICOSHIFT para el alineamiento de la huella dactilar de todas las muestras, se procedió a realizar un cribado de la misma eliminando diferentes columnas. Esta eliminación se realizó tomando como referencia, tanto el agrupamiento no supervisado mediante ACP obtenido para el estudio individual de los aceites

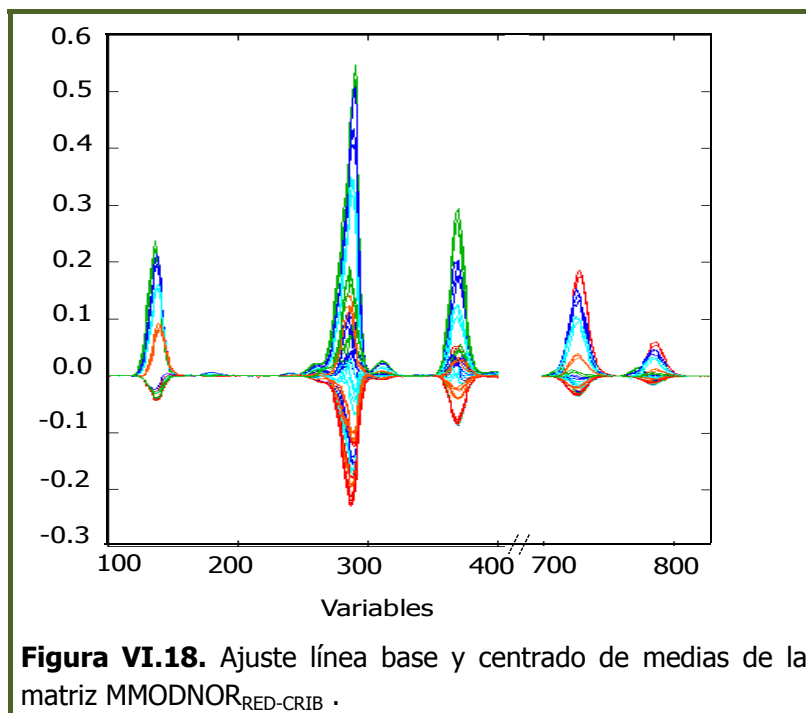
seleccionados, como el peso de cada uno de los perfiles individuales de los fitoesteroles en las tres variables latentes seleccionadas en los modelos anteriores. Con este criterio y tras optimizar las zonas a eliminar, se decidió a partir de los resultados obtenidos en diferentes ensayos, trabajar con una nueva matriz de datos  $\text{MMODNOR}_{\text{RED-CRIB}}$  en la que solo se contemplaba la huella dactilar normalizada y suavizada en los intervalos 100-400 y 700-831 y que en consecuencia originaba una huella dactilar parcial de 431 variables.

En esta **huella dactilar parcial** solo aparecen los perfiles normalizados de brasicasterol, 24-metilencolesterol, campesterol, campestanol, estigmasterol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol y se eliminan el resto de picos, entre ellos, el del  $\beta$ -sitosterol.

La Figura VI.17 muestra las huellas dactilares normalizadas parciales de la nueva matriz  $\text{MMODNOR}_{\text{RED-CRIB}}$  utilizada para entrenar el nuevo modelo.



En este caso, tras obtener la huella dactilar parcial, se procede al nuevo alineamiento de la línea base y a un centrado de medias, el resultado final del mismo se muestra en la Figura VI.18.

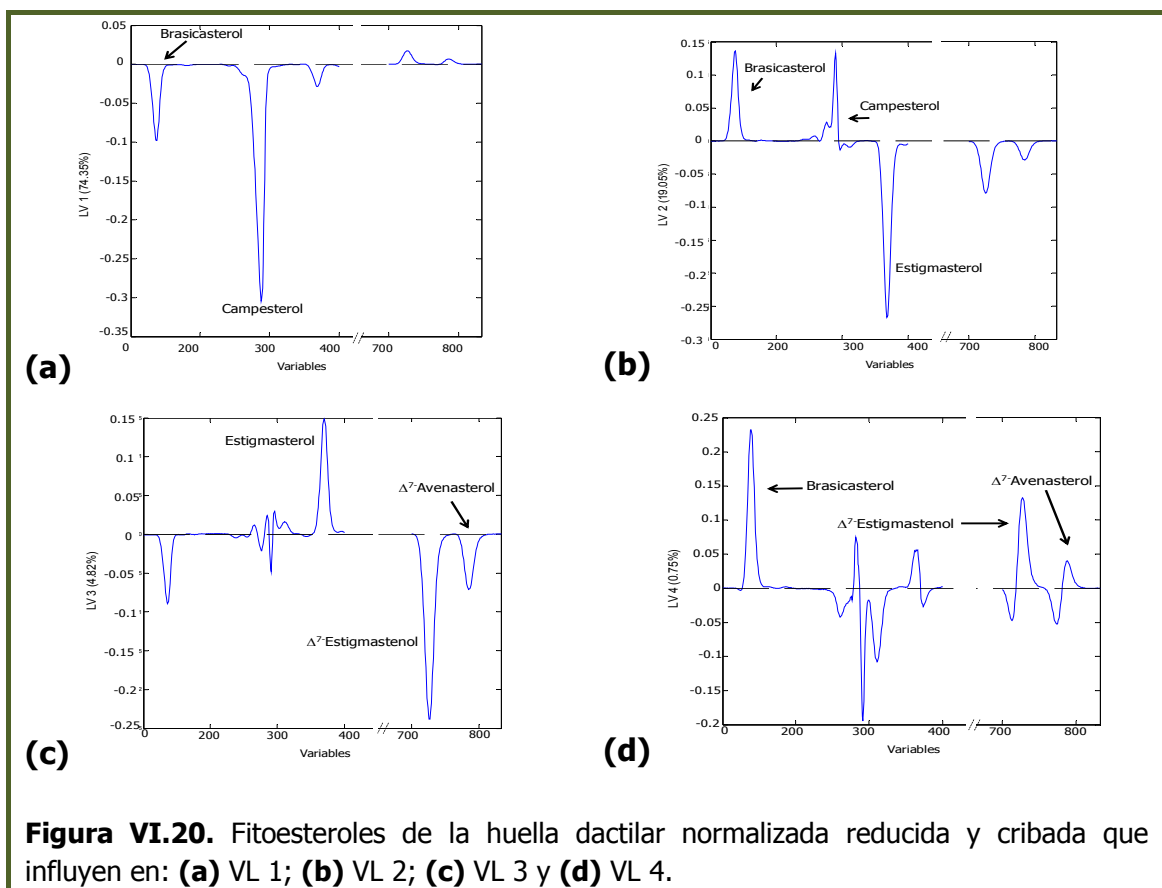
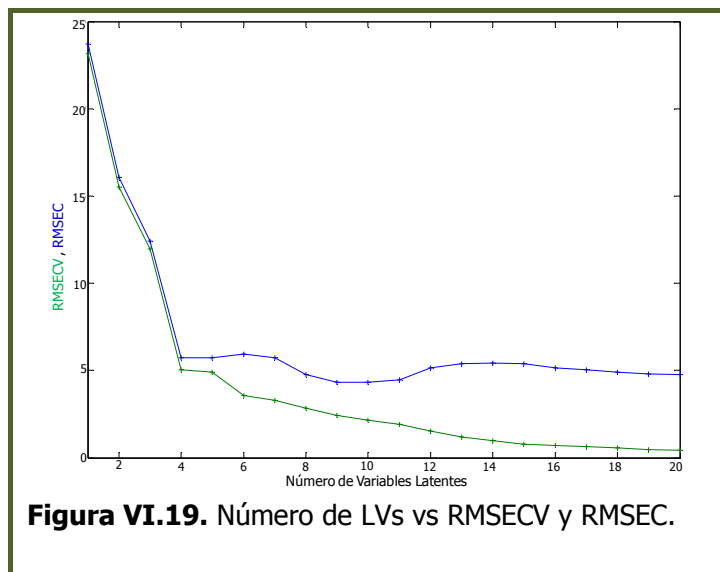


En la elaboración de este último modelo se emplearon de nuevo el algoritmo SIMPLS para encontrar el número de variables latentes significativas y la validación cruzada tipo "Venetian Blinds". Los bloques establecidos para el modelado de calibración XY fueron:

- Bloque de variables de entrada o bloque X: Matriz  $MMODNOR_{RED-CRIB}$ .
- Bloque de variables de salida o bloque Y: Matriz de porcentajes de aceite de oliva utilizados en las mezclas ensayadas para la cuantificación. Esta matriz es la misma que la utilizada en los modelos anteriores.

El número óptimo de variables latentes (VLs) seleccionado, al igual que en los modelos anteriores, a partir los gráficos de sedimentación para los errores cuadráticos medios de los conjuntos de calibración (RMSEC) y validación cruzada (RMSECV) realizada con el set de validación.

En este caso, y como se puede observar en la Figura VI.19, el número óptimo de VLs seleccionadas fue de 4. Estas cuatro variables latentes explican un 98.98% de la varianza del Bloque X y un 96.71% de la varianza del Bloque Y. Al igual que en los modelos anteriores, se necesitaron 7 splits para el tipo de validación cruzada seleccionado.



Tras analizar la Figura VI.20 se puede observar que en este modelo:

**(a)** La LV 1 está influenciada negativamente por el brasicasterol y campesterol.

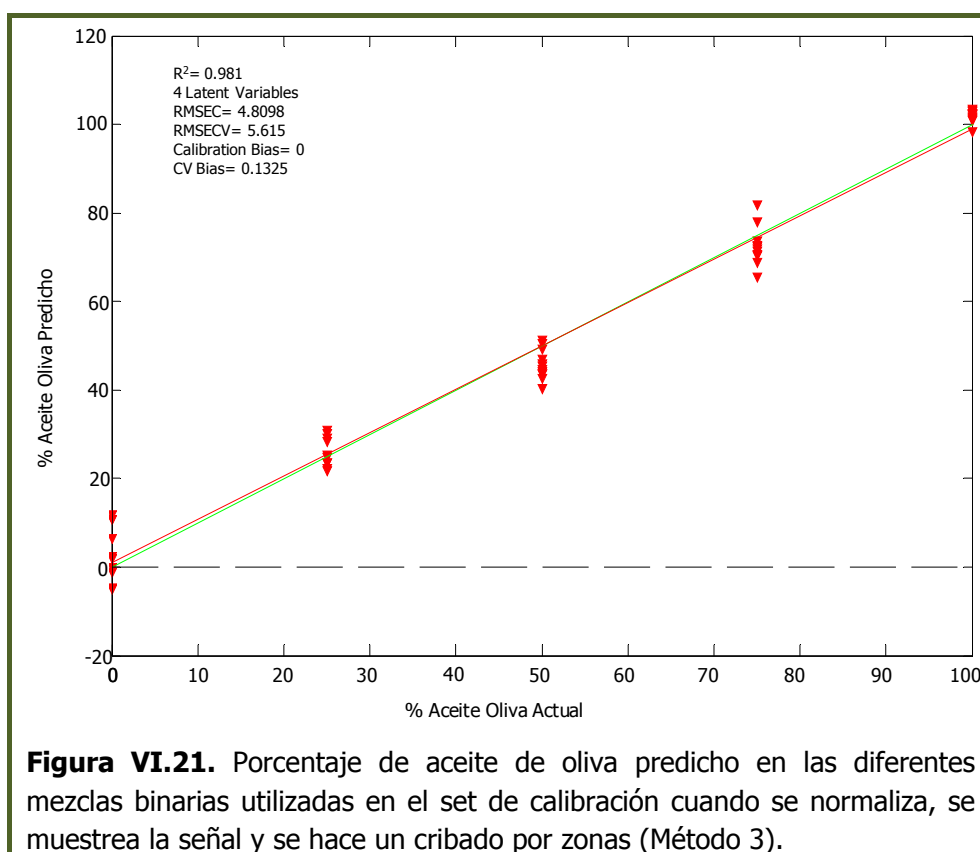
## Cuantificación de Aceite de Oliva en Vinagretas

(b) La LV 2 está influenciada positivamente por el brasicasterol y el campesterol y negativamente por el estigmasterol.

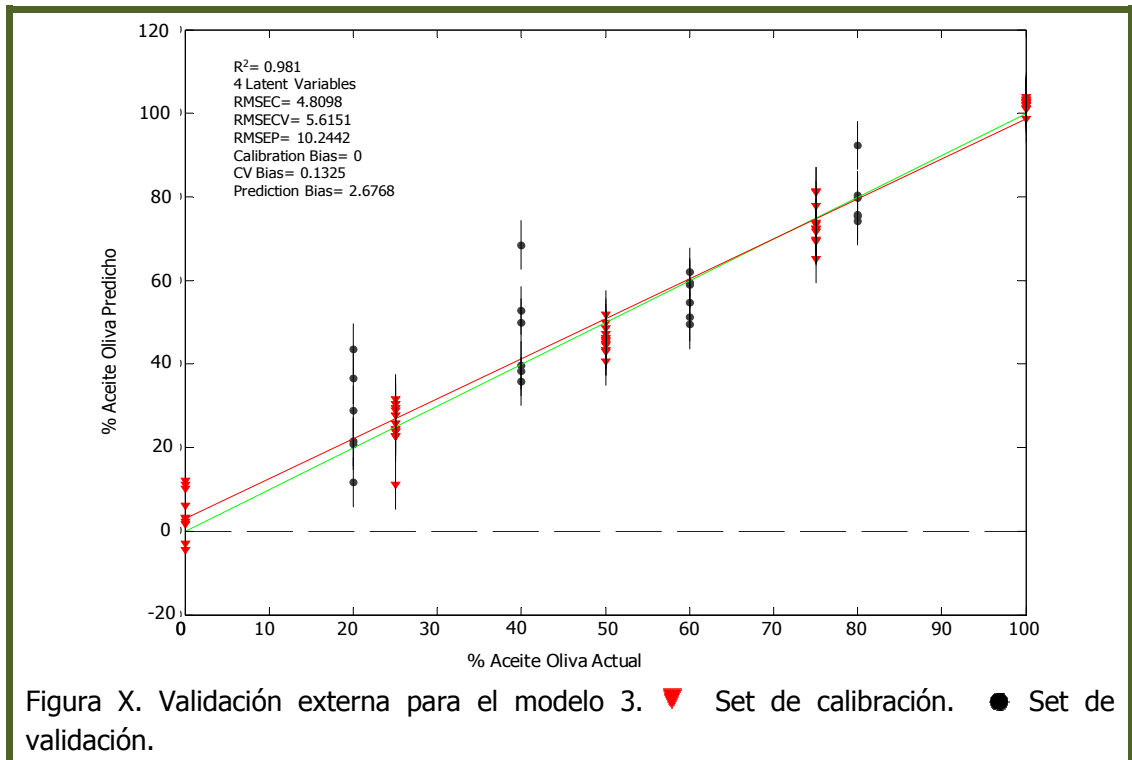
(c) La LV 3 está influenciada positivamente por el estigmasterol y negativamente por  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.

(d) La LV 4 está influenciada positivamente por el brasicasterol,  $\Delta^7$ -estigmasterol y  $\Delta^7$ -avenasterol.

La Figura VI.21 recoge la representación de los porcentajes de aceite de oliva predichos frente a los porcentajes reales en el caso de la validación cruzada realizada sobre el set de calibración. Como se puede observar los valores obtenidos para  $R^2(0.9817)$  y  $RMSECV(5.615)$  indican que la capacidad de predicción del modelo simplificado mejora sustancialmente con respecto a los modelos anteriores a la vez que simplifica el procesado informático. Por tanto es de resaltar que **el suavizado de la señal y el cribado de la huella dactilar parcial mejora sustancialmente tanto la correlación/linealidad del modelo como el error cuadrático medio para la validación cruzada con el set de calibración.**



Al igual que en los modelos anteriores, una vez entrenado el modelo se procedió a su validación con el set de validación externa. La Figura VI.21 muestra el porcentaje de aceite de oliva predicho en las diferentes mezclas binarias utilizadas en dicho set de validación cuando se normaliza, se muestrea la señal y se hace un cribado por zonas, es decir se trabaja con la huella dactilar parcial.



Los principales parámetros de calidad de este modelo se recogen en la siguiente Tabla. Cabe indicar que en este caso el error cuadrático medio de predicción es superior a los obtenidos en los modelos anteriores, aunque oscila en torno a un 10%, valor aceptable para este tipo de cuantificaciones.

**Tabla VI.8.** Resumen de los parámetros más importantes del modelo desarrollado PLS para la determinación de aceite de oliva en muestras de vinagretas.

VL	R <sup>2</sup> <sub>Cal</sub>	RMSEC	R <sup>2</sup> <sub>CV</sub>	RMSECV	R <sup>2</sup> <sub>p</sub>	RMSEP	Bias
4	0.9817	4.8098	0.9773	5.6151	0.8025	10.2442	0



## Cuantificación de Aceite de Oliva en Vinagretas

Finalmente, los resultados obtenidos al utilizar este modelo para la cuantificación de aceite de oliva en alimentos tipo salsas se muestran en la Tabla VI.9.

**Tabla VI.9.** Cuantificación de aceite de oliva en vinagretas usando el Modelo 3.

Muestras	Código	% (m/m) de aceite de oliva	% (m/m) aceite de oliva predicho
	VRM1	100	97.72
	VMR2	100	102.36
	VMR3	50	41.61
	VMR4	25	28.31
	<b>VMR5</b>	<b>75</b>	<b>56.27</b>
	VMR6	60	59.74
Materiales de Referencia Terciarios	VMR7	90	81.72
	<b>VMR8</b>	<b>20</b>	<b>37.75</b>
	VMR9	40	30.43
	<b>VMR10</b>	<b>10</b>	<b>-12.60</b>
	VMR11	0	4.01
	<b>VMR12</b>	<b>0</b>	<b>20.21</b>
	VMR13	0	8.64
	VMR14	60	71.87
	VMR15	80	82.30
		<b>ALI</b>	<b>24</b>
Vinagretas Comerciales	<b>VIN-ORI</b>	<b>7</b>	<b>23.61</b>
	<b>VIN-SEC</b>	<b>7</b>	<b>31.97</b>
	<b>VIN-MOD</b>	<b>15</b>	<b>38.67</b>
	<b>VIN-MOS</b>	<b>7</b>	<b>30.23</b>

Como se puede observar, en el caso de las vinagretas elaboradas en el laboratorio (materiales de referencia terciarios) el modelo predice de forma adecuada el 73% de las mismas, frente al 60% predicho en los modelos anteriores. Entre las nuevas vinagretas cuantificadas de forma correcta caben destacarse las VMR11 y VMR13 correspondientes a vinagretas de referencia elaboradas con aceites que no son de oliva, no predichas en los modelos anteriores. No obstante, y dado que las muestras VMR8 y VMR10 siguen sin ser predichas correctamente se podría concluir que en este tipo de muestras el modelo predice de forma adecuada el contenido en aceite de oliva en aquellas mezclas en las que supere el 30%.

Por otro lado, dado que el modelo desarrollado predice bien el 66.67% de las vinagretas de referencia elaboradas sin aceite de oliva, la hipótesis anterior de que el modelo construido a partir de aceites individuales produce "confusión" puede descartarse, si bien sería recomendable completar el entrenamiento del modelo aumentando la casuística en la preparación de mezclas binarias.

Dadas las características del modelo y considerando la composición de las salsas comerciales (contenido en aceite de oliva < 30% y presencia de diferentes semillas ricas en fitoesteroles). Cabe pensar que el conjunto de muestras comerciales utilizado no resultó ser el más adecuado, sin embargo tras una búsqueda exhaustiva no se encontró ninguna salsa comercializada con un % de aceite de oliva superior.

A la vista de los resultados obtenidos a lo largo de este capítulo, se puede considerar que:

**Una vez eliminada la dependencia que el tamaño de la muestra analizada ejerce sobre la huella dactilar cromatográfica de los fitoesteroles más habituales en aceites vegetales mediante la normalización interna de las variables implicadas (dividiéndolas por la mayor señal cromatográfica registrada al tiempo de retención del  $\beta$ -sitosterol) y, establecidos los tres modelos PLS para la cuantificación de aceite de oliva en mezclas binarias de aceites vegetales que se enumeran en la Tabla VI.10 en la que se además se reúnen sus principales características,**

**Tabla VI.10.** Características más relevantes empleadas en el diseño de los diferentes modelos de cuantificación mediante PLS desarrollados en este capítulo.

<u>Modelo</u>	<u>Tratamiento datos</u> (Preprocesado)	<u>Variabes</u>
1	<b>Normalización</b> (Alineamiento de los picos + Ajuste línea base + Centrado en la media)	<b>18201</b>
2	<b>Normalización + Muestreo basado en el Teorema de Nyquist</b> (Alineamiento de los picos + Ajuste línea base + Centrado en la media)	<b>911</b>
3	<b>Normalización + Muestreo basado en el Teorema de Nyquist</b> (Alineamiento de los picos + Selección de zonas + Ajuste línea base + Centrado en la media)	<b>431</b>

**Se pueden extraer las siguientes conclusiones:**

**(1) Al realizar un muestreo de la señal cromatográfica se ha mejorado notablemente la cantidad de datos para procesar con el software adecuado, reduciendo la matriz de datos original de 18201 variables (Modelo 1) a una matriz de 911 (Modelo 2), sin alterar la información obtenida pero facilitando su manejo.**

**(2) Al realizar un muestreo y acortar parte de la huella dactilar cromatográfica se ha reducido considerablemente la cantidad de datos para procesar informáticamente, reduciendo la matriz de datos original de 18201 variables (Modelo 1) a una matriz de 431 variables (Modelo 3), sin alterar la información original y acortando el tiempo de obtención e interpretación de resultados.**

**(3) La cuantificación del aceite de oliva presente en mezclas binarias, materiales de referencia caseros y productos alimenticios tipo vinagretas o aliños se realiza adecuadamente siempre y cuando el porcentaje real de aceite de oliva en el caso de los dos primeros tipos de muestras, o el porcentaje especificado en la etiqueta, el caso de los dos últimos sea superior al 30%.**

**(4) Para todas las muestras comerciales se ha obtenido un valor sobreestimado con respecto al suministrado en las etiquetas. Este hecho se puede atribuir a que, tal y como se declara en las mismas, para su elaboración, además de aceites se han utilizado otras semillas (sésamo, soja, almendra, trigo, etc.) que pueden ocasionar un aporte de fitoesteroles imposible de cuantificar.**

## BIBLIOGRAFÍA

- [1] Reglamento de Ejecución (UE) Nº 29/2012 de la comisión de 13 de enero de 2012, sobre las normas de comercialización del aceite de oliva. DO L 12 de 14.1.2012.
- [2] Andrade-Garda, J.M.; Boqué-Martí, R.; Ferré-Baldrich, J. & Carlosena-Zubieta, A. in Andrade-Garda, J.M. (Ed.), *Basic Chemometric Techniques in Atomic Spectroscopy*, Royal Society of Chemistry (RCS), Cambridge, 2009, pp 181-243.
- [3] Brereton, R.G. *Chemometrics for Pattern Recognition*, John Wiley & Sons, Chichester, UK, 2009.
- [4] Brodnjak-Voncina, D.; Kodbab, Z.C. & Novic, M. *Multivariate data analysis in classification of vegetable oils characterized by the content of fatty acids*. Chemometric. Intell. Lab. Syst., 2005, 75, 31-43.
- [5] Maggio, R.M.; Cerretani, L.; Chiavaro, E.; Kaufman, T.S. & Bendini, A. *A novel chemometric strategy for the estimation of extra virgin olive oil adulteration with edible oils*. Food Control, 2010, 21, 890-895.
- [6] Dourtoglou, V.G.; Dourtoglou, Th.; Antonopoulos, A.; Stefanou, E.; Lalas, S. & Poulos, C. *Detection of olive oil adulteration using principal component analysis applied on tota and region FA content*. J. AOCS, 2003, 80(3), 203..
- [7] Peña, F.; Cárdenas, S.; Gallego, M. & Valcarcel, M. *Direct olive oil authentication: Detection of adulteration of olive oil with hazelnut oil by direct coupling of heaspace and mass spectrometry, and multivariate regression techniques*. J. Chromatogr. A, 2005, 1074, 215-221.
- [8] de Jong, S. *SIMPLS: an alternative aproach to partial least squares regression*. Intell. Lab. Syst., 1993, 18, 251-263.
- [9] Varmuza, K.; Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press, Boca-Raton, FL, 2009.
- [10] Las Salsas: Vinagretas, en <http://www.gastronomiaycia.com/2008/05/26/las-salsas-vinagretas/> (consultada por última vez en agosto de 2012).

- [11] Ortega, S. *1080 recetas de cocina*. El libro de bolsillo. Alianza Editorial, S.A., Madrid 1988 (receta 89).
- [12] Lucio-Gutiérrez, J.R.; Garza-Juárez, A.; Coello, J.; Maspoch, S.; Salazar-Cavazos, M.L.; Salazar-Aranda, R. & Waksman de Torres, N. *Multi-wavelength high-performance liquid chromatographic fingerprints and chemometrics to predict the antioxidant activity of Turnera diffusa as part of its quality control*. J. Chromatogr. A., 2012, 1235, 68-76.
- [13] Wise, B.M.; Gallagher, N.B.; Bro, R.; Shaver, J.V.; Windig, W.; Koch, R.S. *PLS\_Toolbox 4.0 Manual for use with Matlab™*, Eigenvector Research, Inc. Wenatchee, 2006.
- [14] Savorani, F.; Tomasi, G. & Engelsen, S.B. *icoshift: A versatile tool for the rapid alignment of 1D-NMR-spectra*. J. Magn. Reson., 2010, 202, 190-202.
- [15] Tomasi, G.; Savorani, F. & Engelsen, S.B. *icoshift: An effective tool for the alignment of chromatographic data*. J. Chromatogr. A., 2011, 1218, 7832-7840.
- [16] Skoog, D.A.; Holler, F.J. & Nieman, T.A., en: Skoog, D.A.; Holler, F.J. & Nieman, T.A. (Eds). *Principios de análisis instrumental*, fifth ed., Mc. Graw Hill, Madrid, 2000, pp 103-120.
- [17] Felinger, A. in Felinger, A. (Eds.). *Data analysis and signal processing in chromatography*, Elsevier, Amsterdam (The Netherlands), 1998, pp 80-96.

## **CONCLUSIONES**



Como resultado de esta Tesis Doctoral, se pueden establecer las siguientes conclusiones:

1. Se ha llevado a cabo una revisión bibliográfica de los métodos utilizados para la determinación de fitoesteroles en aceites vegetales. Paralelamente, se hizo sobre el uso conjunto, de cromatografía y quimiometría para el análisis de estos compuestos en aceites y grasas. Sorprendentemente, no se ha encontrado hasta el momento el uso de perfiles en señales como “huellas dactilares” posiblemente por la falta de conocimiento de muchos cromatografistas. Además, se han encontrado casos en los que se ha hecho un mal uso sobre datos discontinuos.
2. Se ha desarrollado y puesto a punto una metodología de identificación de fitoesteroles en aceites vegetales y alimentos mediante un análisis off-line HPLC-GC/FID, evitando el uso de la CCF. Los compuestos esterólicos se aislaron mediante HPLC como técnica semi-preparativa empleando una columna de polaridad media tipo ciano, siendo posible detectar y aislar la fracción de esteroles en menos de 12 minutos. Estudios de repetibilidad y reproducibilidad permiten que este análisis pueda llevarse a cabo como análisis de rutina en cualquier laboratorio.
3. Se han optimizado las condiciones cromatográficas así como, la cantidad de mezcla derivatizante necesaria para favorecer la formación de los derivados siliados de los fitoesteroles en aceites vegetales, para su determinación en un equipo GC-FID
4. Se ha propuesto de forma detallada el cálculo de la incertidumbre asociada a la cuantificación de los TMS-fitoesteroles en aceites vegetales. Se han tenido en cuenta todas las fuentes de incertidumbre relacionadas con la heterogeneidad de la muestra, la calibración interna con un subrogado, la pureza de los reactivos, la reacción de derivatización, los errores aleatorios asociados a todo proceso de medida químico, etc.
5. Se ha llevado a cabo la cuantificación de los TMS-4-desmetilesteroles mediante el método oficial propuesto por el Consejo Oleícola Internacional, COI, que emplea una calibración interna con un subrogado (5-colestanol) y, mediante los factores de corrección teóricos (TCF). En este último caso, se ha podido realizar la cuantificación debido a que se conoce la estructura de estos compuestos determinando el número de carbonos activos.



6. Se ha llevado a cabo el análisis de 75 muestras de aceites vegetales (oliva virgen extra (picual, hojiblanca, arbequina, cornicabra, manzanilla y royal), oliva virgen, oliva refinado, orujo de oliva, girasol, girasol alto oleico, maíz, soja, colza, canola, semillas, pepita de uva, cacahuete, sésamo y lino); 39 muestras preparadas de mezclas de aceite de oliva con otros aceites vegetales en diferentes porcentajes (20-80 %), cubriendo el mayor rango posible; 15 vinagretas como materiales de referencia terciarios, elaboradas en el laboratorio y 5 vinagretas y aliños comerciales. El tratamiento de estas muestras no ha requerido la adición de subrogado. Se ha usado el cromatograma completo para aplicar diferentes herramientas quimiométricas.
7. Se ha potenciado el uso de las huellas dactilares cromatográficas (fingerprint) para agrupar, clasificar y cuantificar los aceites de oliva y aceites vegetales obteniendo buenos resultados. Se han clasificado de forma correcta todas las muestras diferenciando el aceite de oliva del resto de aceites vegetales.
8. Se ha mejorado el pre-procesado de los datos, etapa clave para el tratamiento de los mismos evitando resultados erróneos. Siguiendo varias estrategias para encontrar el pretratamiento más adecuado. Además, se ha reducido el número de variables mediante el Teorema de Nyquist.
9. Se ha desarrollado un modelo quimiométrico para la cuantificación de aceite de oliva en alimentos elaborados con mezclas de aceites vegetales (soja, girasol, colza, canola y semillas) preparado en el laboratorio y adquiridos de forma comercial como vinagretas o aliños, a partir de la huella dactilar de los TMS-4-desmetilesteros. Este modelo de cuantificación, PLS, fue entrenado y validado con conjunto de muestras diferentes en cada caso. Se obtuvieron unos errores de predicción inferiores al 10%.

# **PERSPECTIVAS FUTURAS**



A partir de los estudios y discusiones presentados en esta Tesis Doctoral, se han abierto ante nosotros posibilidades para futuros trabajos, de entre las cuales, se mencionan algunas a continuación:

- ⊙ Puesta a punto de un método para la separación de esteroides con el equipo KONIK HPLC HPLC+GC-MS (K2Q12), con interfase TOTAD ("Through Oven Thermal Absorption Desorption"), provisto de un detector de espectrometría de masas con analizador de cuadrupolo y de un detector de ionización de llama (FID). Este equipo fue recepcionado a mediados del año 2011. Con la adquisición de este equipo se pretende seguir mejorando la etapa de tratamiento de muestra para identificar y cuantificar esteroides sin necesidad de derivatizarlos. Se analizaría el insaponificable de los diferentes aceites vegetales tratados transfiriendo la fracción esteróica aislada directamente por el HPLC a través de la interfase al cromatógrafo de gases. De igual modo, sería interesante analizar el aceite directamente sin tener que saponificarlo.
  
- ⊙ Uso de las huellas dactilares (cromatogramas completos, o espectros de masas) para desarrollar modelos multivariantes para la caracterización y cuantificación del aceite de oliva, aceites vegetales, grasas y productos alimenticios a partir de la metodología arriba indicada.

