



TESIS DOCTORAL

**Combinación de diferentes metodologías analíticas
con herramientas quimiométricas para la
caracterización y detección de fraudes en alimentos**

Programa de Doctorado en Modelización y
Experimentación en Ciencia y Tecnología

Jaime Domínguez Manzano

2024

Conformidad de los directores:

Isabel Durán
Martín-Merás

Arsenio Muñoz de la Peña
Castrillo

Olga Monago
Maraña

Esta tesis cuenta con la autorización de la directora y los codirectores de la misma y de la Comisión Académica del programa. Dichas autorizaciones constan en el Servicio de la Escuela Internacional de Doctorado de la Universidad de Extremadura.

AGRADECIMIENTOS

- Ministerio de Economía, Industria y Competitividad de España (Proyecto CTQ2017-82496-P) cofinanciado por el Fondo Europeo de Desarrollo Regional
- Junta de Extremadura: Ayudas a grupos de investigación FQM003 (GR15090, GR18041 y GR21048) y proyectos regionales de investigación (IB16058, IB16022 y IB20016) cofinanciados por el Fondo Europeo de Desarrollo Regional.
- Ministerio de Ciencia e Innovación de España (proyecto PID2020-112996GB-I00 financiado por MCIN/AEI/10.13039/501100011033) y cofinanciado por el Fondo Social Europeo.



ÍNDICE

I. RESUMEN	1
II. OBJETIVOS.....	7
III. INTRODUCCIÓN.....	11
1. MATRICES ALIMENTARIAS.....	13
1.1. Aceites de Oliva	13
1.2. Ciruelas	56
2. TÉCNICAS ANALÍTICAS EMPLEADAS	79
2.1. Fluorescencia molecular	79
2.2. Técnicas Cromatográficas	82
3. DESCRIPCIÓN DE LOS MÉTODOS QUIMIOMÉTRICOS EMPLEADOS	84
3.1. Algoritmos de dos vías para datos de primer orden.....	89
3.2. Algoritmos de tres vías para datos de segundo orden.....	96
4. TEORÍA DEL ESPACIO DE COLOR.....	111
5. REFERENCIAS	120
IV. EXPERIMENTAL.....	149
CAPÍTULO 1. SEGUIMIENTO Y CONTROL DE CALIDAD DE LOS ACEITES DE OLIVA Y DETECCIÓN DE FRAUDES.....	151
1.INTRODUCCIÓN.....	153
2. METODOLOGÍA	160
3. RESULTADOS Y DISCUSIÓN.....	171
4. REFERENCIAS	221

CAPÍTULO 2. SEGUIMIENTO Y CONTROL DEL ESTADO DE MADURACIÓN DE CIRUELAS A TRAVÉS DE SUS MATRICES DE FLUORESCENCIA Y DEL PROCESAMIENTO DE IMÁGENES DIGITALES.....	229
1. INTRODUCCIÓN.....	231
2. METODOLOGÍA	241
3. RESULTADOS Y DISCUSIÓN.....	254
4. REFERENCIAS	287
V. CONCLUSIONES.....	299
VI. ANEXOS.....	305
ANEXO I. ARTÍCULOS PUBLICADOS	229

I. RESUMEN

Esta tesis se centra en el desarrollo de nuevas metodologías analíticas dentro del sector agroalimentario, fundamentadas en el uso de técnicas no destructivas, accesibles y selectivas, como son la fluorescencia molecular y el procesamiento de imágenes digitales, ambas en combinación con métodos quimiométricos. Estas metodologías han permitido llevar a cabo un control de la calidad de varias matrices alimentarias como han sido los aceites de oliva y las ciruelas, así como la detección de fraudes.

La memoria se ha estructurado en diferentes partes: introducción, parte experimental y conclusiones. La parte experimental, a su vez, se ha dividido en dos capítulos, cada uno de ellos dedicado al desarrollo de las metodologías propuestas aplicadas al aceite de oliva y las ciruelas, respectivamente. Finalmente, se muestran las publicaciones derivadas de dichos estudios.

En primer lugar, la introducción describe los fundamentos que permiten entender y justificar el desarrollo de las metodologías propuestas dentro del campo agroalimentario. Se detalla la importancia de las dos matrices alimentarias objeto de esta tesis, tanto a nivel económico como territorial, destacando también sus características composicionales. Así mismo, la introducción recoge los fundamentos teóricos de las técnicas analíticas empleadas y de las herramientas quimiométricas necesarias para el análisis de los datos generados.

En segundo lugar, la parte experimental se encuentra dividida en dos capítulos. El primero está dedicado al seguimiento y control de calidad de los aceites de oliva y la detección de fraudes. En este capítulo se han desarrollado metodologías que han permitido evaluar el deterioro oxidativo de los aceites de oliva debido a factores como la exposición a la radiación electromagnética o el calentamiento, tanto suave como moderado, durante periodos de tiempo prolongados. También se han desarrollado

metodologías que han permitido detectar adulteraciones del aceite de oliva virgen extra con otros de menor calidad de acuerdo a sus compuestos fluorescentes, como pueden ser los polifenoles y las clorofilas. Estas metodologías han combinado el uso de la fluorescencia molecular (matrices de excitación – emisión) con herramientas quimiométricas de segundo orden.

En cuanto al segundo capítulo, se recogen en él las metodologías analíticas desarrolladas para el seguimiento y control del estado de maduración en ciruelas a través de su fluorescencia no-destructiva y del procesamiento de imágenes digitales tomadas de las mismas a lo largo del proceso de maduración. En ambos estudios se demostró que era posible hacer un seguimiento de la maduración de las frutas en base a su contenido en clorofilas, las cuales se degradan durante la maduración. Al igual que en el caso anterior, se emplearon algoritmos quimiométricos de primer y segundo orden para el procesamiento de los datos obtenidos mediante ambas técnicas.

Finalmente, se recogen las conclusiones de las investigaciones desarrolladas que han dado lugar a la publicación de cuatro artículos científicos mostrados en un anexo al final de la memoria de esta Tesis. Dichos artículos han sido los siguientes:

- **Capítulo 1:**
 - Detection and quantification of extra virgin olive oil adulteration by means of autofluorescence excitation-emission profiles combined with multi-way classification. **Talanta, 178, (2018) 751-762.**
 - Front-face fluorescence combined with second-order multiway classification, based on polyphenol and chlorophyll compounds, for virgin olive oil monitoring under different photo- and thermal-

oxidation Procedures. **Food Analytical Methods**, **12**, (2019) **1399-1411**.

- **Capítulo 2:**

- Second-order calibration in combination with fluorescence fibre-optic data modelling as a novel approach for monitoring the maturation stage of plums. **Chemometrics and Intelligent Laboratory Systems**, **199**, (2020) **103980**.
- Monitoring of chlorophylls during the maturation stage of plums by multivariate calibration of RGB data from digital images. **Chemosensors**, **11**, (2023) **9**.

II. OBJETIVOS

La presente Tesis Doctoral, enmarcada dentro del ámbito de la química analítica agroalimentaria, tiene como objetivo general el desarrollo de metodologías analíticas que cumplan necesidades, tales como la reducción del uso de sustancias auxiliares, reducción de residuos, reducción del consumo energético, así como la mejora en la eficiencia y la reducción de costes en la implementación de dichas metodologías.

Para ello, se ha abordado el desarrollo de metodologías analíticas mínimamente invasivas en matrices alimentarias tales como el aceite de oliva y las ciruelas.

Los objetivos específicos que se van a desarrollar en la presente Tesis Doctoral han sido los siguientes:

1. Aplicación de técnicas de fluorescencia molecular en combinación con algoritmos multivariantes de segundo orden para caracterizar aceites de oliva en función de su denominación, irrigación, estado de oxidación o grado de adulteración.
2. Identificación de compuestos característicos del grado de calidad de los aceites de oliva y seguimiento del deterioro oxidativo de estos a través de la aplicación de técnicas de fluorescencia molecular en combinación con algoritmos multivariantes de segundo orden.
3. Aplicación de técnicas de fluorescencia molecular no invasivas en combinación con algoritmos multivariantes en frutos de hueso, como las ciruelas, para la identificación y determinación de compuestos marcadores de su estado de maduración.
4. Implementación de técnicas de procesamiento de imagen digital fundamentadas en el estudio del espacio de color RGB en combinación con algoritmos multivariantes de primer y segundo orden para el estudio del estado de maduración en ciruelas.

III. INTRODUCCIÓN

Antes de proceder a detallar los resultados obtenidos en esta investigación, es necesario poner en contexto las matrices agroalimentarias estudiadas. Por lo que, en este apartado de la memoria, se comentarán las características de las matrices alimentarias seleccionadas, teniendo en cuenta tanto las similitudes como las diferencias entre estas, estructurando su estudio en base al impacto que tienen dichos alimentos sobre el sistema económico y productivo, tanto en el ámbito autonómico, como nacional e internacional. Se revisarán los compuestos químicos más importantes para la caracterización y seguimiento de la calidad, a través de los cuales se puedan establecer una serie de características analíticas generales, que permitan el desarrollo de métodos analíticos no destructivos y que puedan competir con los métodos tradicionales, que normalmente implican varias etapas previas de tratamiento de las muestras.

Seguidamente, se hará una revisión de las técnicas analíticas empleadas experimentalmente en esta memoria y su estado actual, así como de los métodos quimiométricos que han sido empleados en los diferentes estudios.

1. MATRICES ALIMENTARIAS

1.1. Aceite de oliva

1.1.1. Definición e Impacto Económico

El aceite de oliva, de acuerdo con el Consejo Oleícola Internacional [1], se puede definir como “el aceite procedente únicamente del fruto del olivo (*Olea europaea* L.), con exclusión de los aceites obtenidos por disolventes o por procedimientos de reesterificación y de toda mezcla con aceites de otra naturaleza”.

El Reglamento (UE) N° 1308/2013 [2], establece el uso de las denominaciones y definiciones de los aceites de oliva y de los aceites de orujo de oliva para la comercialización de dichos productos en la Unión Europea, así como en terceros países, con el fin de ofrecer a los consumidores una información adecuada del tipo de producto comercializado y de sus características nutricionales, estableciéndose así las siguientes categorías:

- **Aceite de oliva virgen:** es el aceite obtenido del fruto del olivo exclusivamente por medios mecánicos u otros procedimientos físicos, aplicados en condiciones que excluyan toda alteración del producto, y que no se ha sometido a ningún otro tratamiento que no sea su lavado, decantación, centrifugado o filtración, excluidos los aceites obtenidos con el uso de disolventes o de coadyuvantes de acción química o bioquímica, por un procedimiento de reesterificación, o como resultado de cualquier mezcla con aceites de otros tipos. El aceite de oliva virgen solo puede clasificarse y designarse de la forma siguiente:
 - a) *Aceite de oliva virgen extra (EVOO, del inglés extra virgin olive oil):* es aquel que tiene una acidez libre máxima, en ácido oleico, de 0,8 g por 100 g, cuyas otras características se ajustan a las establecidas para esta categoría.
 - b) *Aceite de oliva virgen (VOO del inglés virgin olive oil):* es el que tiene una acidez libre máxima, en ácido oleico, de 2 g por 100 g, cuyas otras características se ajustan a las establecidas para esta categoría.
 - c) *Aceite de oliva lampante:* tiene una acidez libre, en ácido oleico, de más de 2 g por 100 g y/o cuyas otras características se ajustan a las establecidas para esta categoría.

- **Aceite de oliva refinado** (*ROO del inglés refined olive oil*): es el aceite de oliva obtenido del refinado de aceite de oliva virgen, que tiene una acidez libre, expresada en ácido oleico, de no más de 0,3 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.
- **Aceite de oliva** (*OO del inglés olive oil*): es el que contiene exclusivamente mezclas de aceites de oliva virgen con aceites de oliva refinados, y tiene una acidez libre, expresada en ácido oleico, de no más de 1 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.
- **Aceite de orujo de oliva crudo**: se obtiene del orujo de oliva mediante un tratamiento con disolventes o empleando medios físicos, o que corresponde, salvo en determinadas características, al aceite de oliva lampante, y cuyas otras características se ajustan a las establecidas para esta categoría, excluido el aceite obtenido por un procedimiento de reesterificación o como resultado de una mezcla con aceites de otros tipos.
- **Aceite de orujo de oliva refinado**: es el aceite obtenido del refinado de aceite de orujo de oliva crudo, que tiene una acidez libre, expresada en ácido oleico, de no más de 0,3 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.
- **Aceite de orujo de oliva** (*POO, del inglés pomace olive oil*) es aquel que se obtiene mezclando aceite de orujo de oliva refinado y aceite de oliva virgen distinto del lampante, que tiene una acidez libre, expresada en ácido oleico, de no más de 1 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.

En la **Figura 1**, se presenta de manera esquematizada la clasificación de los distintos tipos de aceite de oliva desde su extracción hasta la obtención del aceite de oliva virgen y el aceite de orujo.

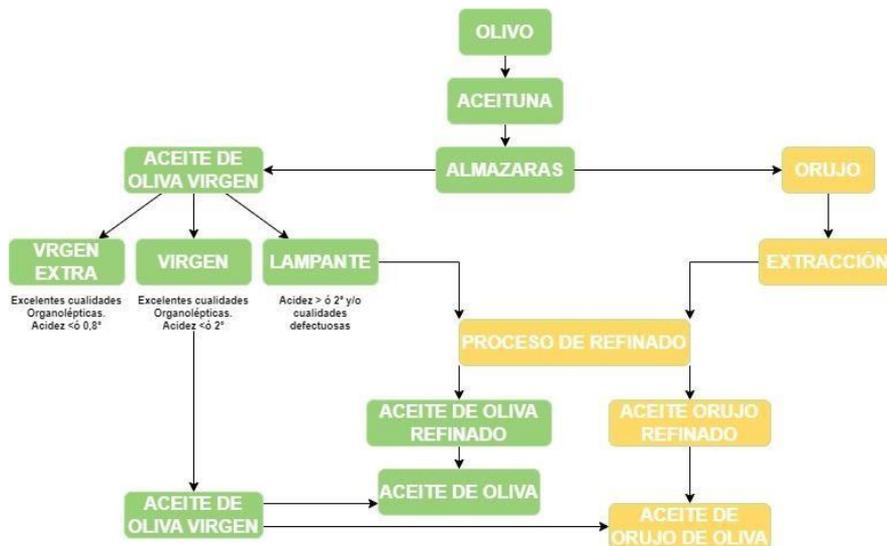


Figura 1. Clasificación de los diferentes tipos de aceites de oliva.

Adaptado del libro *El aceite de oliva virgen tesoro de Andalucía* [3].

Atendiendo a la diferenciación establecida según las diferentes denominaciones que contempla el Reglamento (UE) N° 1308/2013 de 17 de diciembre [2], cabe destacar que los aceites de oliva vírgenes deben su calidad diferenciada con respecto a otras denominaciones a aspectos tales como la calidad del fruto del que se extrae, que debe estar en perfectas condiciones de madurez y de salud, del correcto proceso de recolección y extracción del aceite, de tal manera que el producto no se deteriore durante dichos procesos, así como de las correctas condiciones de almacenamiento y envasado que eviten la pérdida de calidad para su comercialización. Todos estos aspectos se ven reflejados en cambios en la composición y en las características sensoriales del producto que han de ser analizadas a través de métodos analíticos adecuados para establecer dichas denominaciones.

Dentro del impacto económico que a nivel autonómico, nacional e internacional tiene el aceite de oliva, se especifican las siguientes consideraciones al respecto.

A nivel internacional, según los datos ofrecidos por la Organización de las Naciones Unidas para la Alimentación (FAO), España es el mayor productor de aceite de oliva virgen, con una producción total de 1.103.276 toneladas de este producto (FAOSTAT 2020) [4].

El impacto del cultivo del olivar en España queda recogido en los datos extraídos del Anuario de Estadística de 2021 del Ministerio de Agricultura, Pesca y Alimentación (MAPA 2021a.) [5], en el que se pueden destacar los siguientes datos de producción y superficie cultivada.

- Con respecto a la superficie cultivada, en el anuario del 2021 se recoge que hay 2.452.569 hectáreas dedicadas a la producción de aceite de oliva, de las cuales 1.902.408 son de secano y 550.161 están dedicadas a regadío, ofreciendo un rendimiento de 2.589 kg/ha para el primero y 5.684 kg/ha para el segundo.
- En cuanto a la producción, en la campaña 2020-2021 se obtuvo una cantidad de 1.356.411 toneladas de aceite de oliva, de las cuales el 44% fue virgen extra, el 36% virgen y el 20% restante lampante, siendo la comunidad de Extremadura la tercera en volumen de producción con un total de 50.374 toneladas, en su mayor parte aceite de oliva virgen extra. En cuanto a la producción total de aceite de orujo se produce un total de 101.232 toneladas en España.
- En cuanto al balance de mercado de este producto, se exportó una cantidad de 1.097.670 toneladas y 542.800 son para consumo interior, suponiendo las exportaciones un valor de 3144 millones de euros, lo que supone prácticamente el 1% de los ingresos totales en exportaciones.

La repercusión social, ambiental y territorial que la producción de aceite de oliva tiene en España es muy elevada, ya que según datos del Ministerio de Agricultura, Pesca y Alimentación [5] más de 350.000 agricultores se dedican al cultivo del olivar, el sector mantiene unos 15.000 empleos en la industria y genera más de 32 millones de jornales por campaña. Por ello, el establecimiento de normas de calidad diferenciada y procedimientos para garantizar dicha calidad al consumidor se hacen imprescindibles. Así, las autoridades competentes en materia agroalimentaria han establecido una normativa definida a nivel europeo por los reglamentos antes citados, que posteriormente han sido desarrollados y puestos en marcha por los organismos competentes a nivel nacional, estableciéndose así una serie de parámetros de calidad propios del aceite de oliva y los métodos analíticos validados que garantizan la composición nutricional y las características sensoriales del producto.

Por tanto, la calidad del aceite de oliva viene dada por una composición característica del producto, como consecuencia de la extracción del aceite de variedades características de aceitunas, sometidas a procesos agronómicos y de extracción propicios para obtener el producto de calidad diferenciada que establece la reglamentación.

1.1.2. Composición y Parámetros de Calidad

Con respecto a la composición de los aceites de oliva virgen, dada la importancia de esta en sus características, se diferencia entre fracción mayoritaria, la cuál es la fracción saponificable que supone un 98-99% del peso total del aceite, y la fracción minoritaria, la cuál es la fracción insaponificable que alcanzará el 1-2% restante. En la **Figura 2** se pueden observar las familias de compuestos que caracterizan ambas fracciones.



Figura 2. Especificación de las familias de compuestos presentes en las fracciones saponificable e insaponificable de los aceites de oliva [3].

Dentro de la fracción mayoritaria se encuentran los ácidos grasos, los cuales pueden encontrarse libres, aunque rara vez lo hacen ya que, de forma mayoritaria, suelen encontrarse esterificados a glicerina formando triglicéridos, y de forma minoritaria, a alcoholes, formando ceras, o a terpenos, formando ésteres de esteroides. El ácido graso principal que compone los ácidos grasos libres del aceite de oliva es el ácido oleico (68-81,5%).

Los efectos beneficiosos para la salud del aceite de oliva están asociados con la presencia de los ácidos grasos monoinsaturados $\omega-3$ y $\omega-9$ que, según Capurso C. et al., presentan propiedades antioxidantes y antiinflamatorias [6]. También se asocian con la presencia de otros compuestos pertenecientes a la fracción minoritaria, como son los compuestos fenólicos y los tocoferoles, los cuales se han relacionado con la prevención de enfermedades coronarias relacionadas con el daño oxidativo, el infarto o determinados tipos de cánceres [7].

En términos generales, en las **Tablas 1 y 2** se resumen la composición principal y valor nutricional asociado a la fracción

mayoritaria, o saponificable, donde se pueden observar los nutrientes principales que componen el aceite de oliva y la distribución porcentual en ácidos grasos, cuyo balance está asociado en gran medida a las propiedades antioxidantes previamente mencionadas.

Tabla 1. Datos sobre los nutrientes principales del aceite de oliva según la base de datos del Departamento de Agricultura de los Estados Unidos (USDA) [8].

Nutriente	Subclase	Contenido (por cada 100 g)
Minerales	Calcio	1 mg
	Hierro	0,56 mg
	Potasio	1 mg
	Sodio	2 mg
Vitaminas	Vitamina E (α -tocoferol)	14,35 mg
	Vitamina K (Filoquinona)	60,2 μ g
Lípidos	Ácidos grasos saturados	13,808 g
	Ácidos grasos monoinsaturados	72,961 g
	Ácidos grasos poliinsaturados	10,523 g

Como se puede observar en la **Tabla 1**, además de la ya mencionada composición rica en ácidos grasos monoinsaturados, también existen concentraciones importantes de minerales y vitaminas ya pertenecientes a la fracción minoritaria, entre los que destacan calcio, hierro, potasio, sodio, la vitamina E y la vitamina K.

En cuanto a la composición específica en ácidos grasos, en la **Tabla 2**, se resume la proporción de los diferentes ácidos grasos.

Tabla 2. Ácidos grasos en aceites de oliva [9].

Ácido Graso	Tipo de ácido graso	Proporción media (%)
Mirístico	C 14:0	0,010
Palmítico	C 16:0	12,09
Palmitoleico	C 16:1	1,15
Margárico	C 17:0	0,05
Margaroleico	C 17:1	0,10
Esteárico	C 18:0	3,01
Oleico	C 18:1	72,77
Linoleico	C 18:2	9,47
Linolénico	C 18:3	0,060
Araquidónico	C 20:0	0,36
Eicosanoico	C 22:0	0,23
Behénico	C 22:1	0,11
Lignocérico	C 24:0	0,05

Se observa que la proporción en ácidos grasos monoinsaturados está comprendida entre el 72 y 77% del contenido total en lípidos, la cuál es superior a la encontrada en cualquier otro aceite vegetal, así como un contenido de ácidos grasos poliinsaturados inferior al de otros aceites vegetales, lo cual se asocia a los conocidos beneficios que tiene la ingesta de aceite de oliva, ya que el alto contenido en ácido oleico se ha relacionado con la reducción del riesgo de padecer enfermedades coronarias [10], así como con un efecto hipotenso [11]. Además, el ácido oleico ha demostrado ser un excelente antiaterogénico, ya que se ha relacionado su ingesta con un decrecimiento en los niveles de colesterol LDL [12, 13]. Además, los ácidos grasos poliinsaturados presentes en el aceite de oliva, como el ácido linoleico, también presentan propiedades de prevención de enfermedades coronarias debido a que se asocian con la reducción de colesterol LDL en sangre [14] y con una menor incidencia de diabetes [15].

Por otro lado, la fracción minoritaria de los aceites de oliva la componen un grupo mucho más heterogéneo y numeroso de compuestos con gran diversidad de propiedades y estructuras químicas, entre los que destacan compuestos químicamente relacionados con ácidos grasos, como son los ya mencionados fosfolípidos, ceras o ésteres de esteroides, y compuestos catalogables como insaponificables como pueden ser los esteroides, tocoferoles, pigmentos, compuestos fenólicos o los compuestos volátiles, tal y como se indican en la **Figura 2**.

La gran diversidad estructural de los compuestos pertenecientes a la fracción no saponificable y la dependencia de estos con factores como la denominación de calidad, el origen geográfico, las prácticas agronómicas o la autenticidad de los aceites de oliva, hace necesario destacar las características no sólo funcionales, sino también fisicoquímicas de estos compuestos, que van a permitir el desarrollo de nuevas metodologías analíticas. Los compuestos de los que se van a tratar pues serán los siguientes:

Tocoferoles

Son un grupo de compuestos que constituyen la vitamina E de la que destaca su elevado poder antioxidante, impidiendo la peroxidación celular de las membranas celulares y de las lipoproteínas. Han demostrado un fuerte efecto en la prevención del daño oxidativo de los tejidos y la inflamación por la influencia en la expresión de genes homeostáticos, siendo capaz de atenuar el proceso de envejecimiento, de enfermedades degenerativas y cáncer. En cuanto a su estructura química son compuestos heteroácidos de alto peso molecular derivados del 6-cromanol que, junto con trienoles, conforman la vitamina E. Se han identificado cuatro tocoferoles de manera aislada, que se han designado como α , β , γ y δ -tocoferol. Todos tienen un núcleo de cromanol, sustituido con grupos fenólicos y metilos y una cadena lateral saturada, de 16 átomos de carbono,

con tres esqueletos de isopreno. De estos cuatro tocoferoles, se estima que, en los aceites de oliva, el α -tocoferol supone entre el 90-95% de los tocoferoles totales, el otro 10-5% lo conformarían el β y el γ -tocoferol, mientras que el δ -tocoferol aparece en concentraciones muy bajas o incluso se estima que no forme parte de los aceites de oliva [3]. Sus estructuras se muestran en la **Figura 3**:

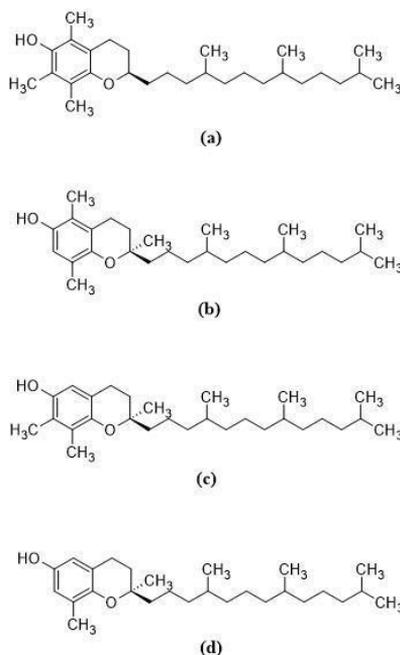


Figura 3: Estructura de los tocoferoles: α -tocoferol (a)
 β -tocoferol (b); γ -tocoferol (c) y δ -tocoferol (d).

Como se puede observar, la estructura química de los tocoferoles presenta aromaticidad, sustituyentes donadores de electrones y cierta rigidez en su núcleo de cromol. De cara a sus propiedades analíticas, las propiedades estructurales citadas se traducen en un aumento del rendimiento cuántico de fluorescencia, para longitudes de onda entre 284 y 500 nm, excitando entre 270-350 nm. Esta fluorescencia nativa permite identificarlos en los aceites de oliva de buena calidad, ya que el contenido en tocoferoles suele estar entre los 100 y los 300 mg/kg.

clasificar en cinco categorías, cuyos componentes principales, junto con su estructura y características de interés están resumidas en la **Tabla 3**.

Tabla 3. Fenoles mayoritarios en los aceites de oliva y características que afectan a su biodisponibilidad [17]–[20].

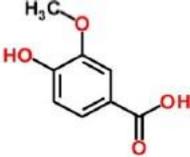
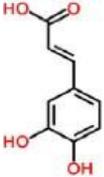
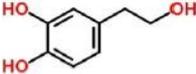
Familia	Compuesto Representativo	Estructura	Contenido en aceite mg/100 g	Masa molecular (g/mol)	Solubilidad en agua (mg/ml) 25°C	Polaridad ^a
Ácidos fenólicos	Ácido Vanílico		0,06	168,15	5,7	1,17-1,70
	Ácido cafeico		0,02	180,16	1,61	1,53-1,67
Alcoholes fenólicos	Hidroxitirosol		0,77	154,16	17,4	0,13-0,89

Tabla 3. Continuación

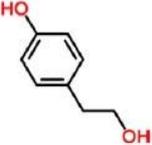
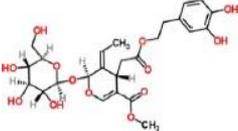
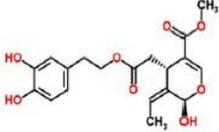
Familia	Compuesto Representativo	Estructura	Contenido en aceite mg/100 g	Masa molecular (g/mol)	Solubilidad en agua (mg/ml) 25°C	Polaridad ^a
	Tirosol		1,13	138,16	25,3	0,85-1,19
Secoiridoides	Oleuropeína		0,17	540,51	0,73	0,11-0,63
	Oleuropeína aglicona		7,22	378,37	0,042	1,88-2,16

Tabla 3. Continuación

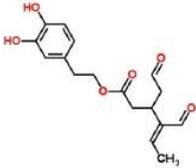
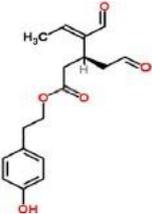
Familia	Compuesto Representativo	Estructura	Contenido en aceite mg/100 g	Masa molecular (g/mol)	Solubilidad en agua (mg/ml) 25°C	Polaridad ^a
Secoiridoides	3,4-DHPEA-EDA		25,16	320,34	0,042	1,53-2,71
	Oleocantal		14,28	304,34	0,016	1,83-2,95

Tabla 3. Continuación

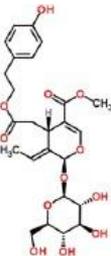
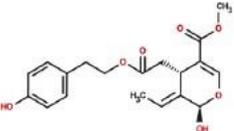
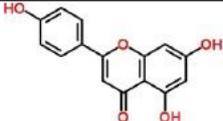
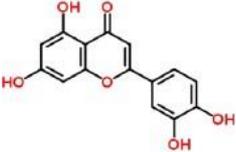
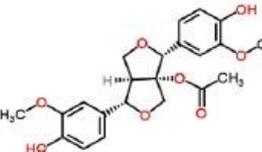
Familia	Compuesto Representativo	Estructura	Contenido en aceite mg/100 g	Masa molecular (g/mol)	Solubilidad en agua (mg/ml) 25°C	Polaridad ^a
Secoiridoides	Ligstrosida		1,56	524,51	0,72	0,41-0,77
	Ligstrosida aglicona		3,80	362,37	0,026	2,22-2,40
Flavonoides	Apigenina		1,17	270,24	0,12	2,71-3,07

Tabla 3. Continuación

Familia	Compuesto Representativo	Estructura	Contenido en aceite mg/100 g	Masa molecular (g/mol)	Solubilidad en agua (mg/ml) 25°C	Polaridad ^a
	Luteolina		0,36	286,24	0,14	2,40-2,73
	(+)-Pinoresinol		0,42	358,38	0,031	2,28-2,31
Lignan	1-Acetoxipinoesterol		0,66	416,42	0,08	2,53-3,73

^aAprox- estimación del log del coeficiente de partición entre octanol y agua estimada para cada compuesto individualmente.

Los principales compuestos fenólicos del aceite de oliva son el hidroxitirosol, el tirosol, los derivados de la oleuropeína y la ligstrosida, los cuales representan, según De la Torre-Carbot et al. [21], aproximadamente un 90% del contenido total en compuestos fenólicos en aceites de oliva vírgenes.

Se ha demostrado que la ingesta de hidroxitirosol, tirosol y oleuropeína ayuda a reducir el estrés oxidativo, a través de la síntesis de glutatión y de elementos de respuesta antioxidativa, siguiendo el mecanismo propuesto en la **Figura 5**.

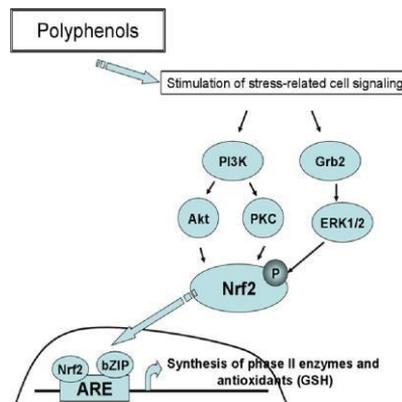


Figura 5. Mecanismo hipotético de acción de los fenoles de los aceites de oliva, en donde se muestra la síntesis de moléculas antioxidantes y desintoxicantes como el GSHl glutatión y GSH [22].

La concentración promedio en aceites de oliva de estos polifenoles oscila entre 50 y 200 mg/kg de aceite, en función de factores tanto endógenos como exógenos, como pueden ser el estado de maduración del fruto [23], la variedad, las condiciones climáticas y agronómicas y el procedimiento de extracción del aceite de oliva [24].

Como se ha ido viendo a lo largo de esta sección, la estructura característica de los compuestos fenólicos, fundamentada en la presencia de anillos aromáticos conjugados con grupos donadores de electrones, hacen que estos presenten también un alto rendimiento cuántico de fluorescencia, haciendo posible la detección y caracterización de estos compuestos por técnicas de fluorescencia molecular, lo cual permite obtener una huella digital que, en combinación con herramientas quimiométricas, puede permitir la caracterización y detección de fraudes.

Pigmentos

Los pigmentos son los componentes a los cuáles el aceite de oliva debe su color, el cual puede variar desde el verde oscuro al dorado, según la variedad y el grado de maduración del fruto. Estos componentes se ven sobre todo implicados en los mecanismos de auto-oxidación y foto-oxidación, por lo que su presencia se correlaciona con la calidad y durabilidad del producto, haciendo de estos compuestos una importante fuente de información para el análisis y seguimiento.

Dentro del grupo de los pigmentos, en los aceites de oliva abundan principalmente dos grupos, las clorofilas y feofitinas en un grupo, y los carotenoides en otro. La concentración y proporción entre ellos depende del estado de madurez, por lo que su determinación simultánea y clasificación resulta de gran utilidad para poder determinar aspectos relacionados con la caracterización e identificación de estos.

- Las **clorofilas** y **feofitinas** se clasifican en dos subgrupos *a* y *b*, su estructura básica se fundamenta en la presencia de un núcleo formado por cuatro anillos pirrólicos fuertemente conjugados, unidos a un radical metilo o carbonilo dependiendo de si se trata del subgrupo *a* o *b*, respectivamente, complejados con un átomo de magnesio en el caso de las clorofilas a los que se le adhiere una cadena de fitol de 20 átomos de

unión de cada una se invierte en el centro de la molécula. La **Figura 7** muestra la estructura básica de esta familia de compuestos.

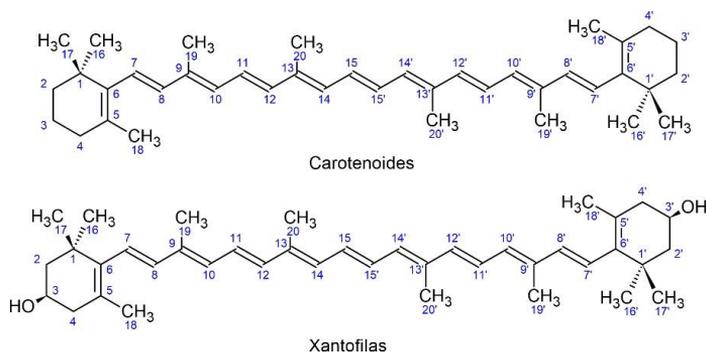


Figura 7. Estructura química de los carotenoides.

En el caso de los aceites de oliva, el contenido en carotenoides varía entre 1 y 20 mg/kg, siendo el mayoritario la luteína, pudiendo variar su contenido dependiendo de factores tales como los sistemas de extracción y molturación de la aceituna [3].

La absorción molecular de los carotenoides se debe a la presencia de dobles enlaces conjugados carbono-carbono, a través de los cuáles, los electrones presentes en dicha estructura pueden promocionar a estados excitados, produciendo tres bandas que dependen del cromóforo del carotenoide objeto de estudio, tal y como se puede observar en la **Figura 8**.

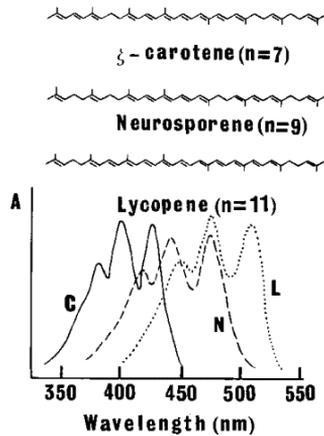


Figura 8. Bandas de absorción de los diferentes cromóforos de los carotenoides [25].

Por otro lado, con respecto a las propiedades fluorescentes de los carotenoides, es destacable mencionar que la intensidad de fluorescencia es débil en comparación con la de las clorofilas, ya que la transición de estados excitados a estados de menor energía suele ocurrir a través de la transición a estados “triplete”, los cuales no son fluorescentes.

Compuestos procedentes de la oxidación de los aceites de oliva

La alteración por vía oxidativa de los aceites da lugar a una serie de compuestos, que inicialmente no están presentes en cantidades significativas en los aceites vírgenes, pero que presentan también propiedades fluorescentes de especial relevancia de cara a la caracterización de los aceites de oliva según su calidad. Estos componentes son un claro indicador del deterioro oxidativo. Tal es el caso de dienos y trienos conjugados y derivados de la oxidación de fenoles y tocoferoles, cuyos máximos de emisión aparecen a longitudes de onda de 445, 520 y 681 nm [26]. La fluorescencia a estas longitudes de onda se ha correlacionado con el aumento del parámetro de absorbancia K_{270} , el cual es un parámetro

oficial estandarizado por el Comité Oleícola Internacional (COI) [27] que se relaciona con la presencia de productos de oxidación secundaria y que se detalla más adelante.

Pese a que la presencia de los compuestos de la fracción minoritaria citados previamente resultan mucho menor que otros, como es el caso del perfil de ácidos grasos, su estrecha relación con procesos oxidativos dependientes de la calidad y sus propiedades fluorescentes citadas de gran interés para esta memoria, hacen de estos compuestos un indicador excelente sobre los que desarrollar una serie de metodologías fundamentadas en la técnica de la fluorescencia molecular y el posterior tratamiento de los datos multivariantes derivados de esta, con fines de caracterización y detección de fraudes de manera rápida, selectiva, robusta y rentable. De este modo, se hace necesario introducir los parámetros correspondientes a la fluorescencia molecular, previamente estudiados y referenciados sobre este tipo de matrices, para su posterior desarrollo en el apartado experimental.

Las longitudes de onda de excitación y emisión de los compuestos que contribuyen a la fluorescencia nativa del aceite de oliva han sido previamente estudiadas por diversos investigadores. En la **Tabla 4** se resumen las longitudes de onda en diferentes disolventes según Eitenmiller et al., [28] (a), Ward et al. [29] (b), Udenfriend, [30] (c), Diaz et al., [31] (d) y Tena et al. [32] (e).

Tabla 4. Longitudes de onda de absorción, excitación y emisión de los compuestos responsables de la fluorescencia nativa en aceites de oliva

Vitamina E	Disolvente	λ_{abs} [nm]	ϵ [L mol ⁻¹ cm ⁻¹]	Disolvente	λ_{ex} [nm]	λ_{em} [nm]
α -tocoferol ^a	etanol	292	3265	n-hexano	295	320
β -tocoferol ^a	etanol	296	3725	n-hexano	297	322
δ -tocoferol ^a	etanol	298	3515	n-hexano	297	322
γ -tocoferol ^a	etanol	298	3809	n-hexano	297	322
α -tocotrienol ^a	etanol	292	3652	n-hexano	290	323
β -tocotrienol ^a	etanol	292	3540	n-hexano	290	323
δ -tocotrienol ^a	etanol	297	3430	n-hexano	292	324
γ -tocotrienol ^a	etanol	297	3737	n-hexano	290	324
Clorofilas						
Clorofila <i>a</i> ^{bcd}	acetona	430	94700	éter	436	668
		663	75000	acetona/ agua 9:1	405	669
Clorofila <i>b</i> ^{bcd}	acetona	455	131000	éter	436	648
		645	47100	acetona/ agua 9:1	405	652
Feofitina <i>a</i> ^{bcd}	acetona	409	101800	éter	436	673
		666	44500	acetona/ agua 9:1	406	671
Feofitina <i>b</i> ^{bcd}	acetona	434	145000	éter	436	661
		654	27800	acetona/ agua 9:1	435	658
Feoforbida <i>a</i> ^{bcd}	acetona	409	119200	-	-	-
		667	55200	-	-	-
Compuestos fenólicos						
Oleuropeína ^e	etanol/n- hexano	282	-	Etanol/ n- hexano	270	310
Ácido vanílico ^e					270	349
Ácido siríngico ^e					270	361
Ácido gálico ^e					270	382
Ácido <i>p</i> -cumárico ^e				metanol	-	416
Ácido <i>o</i> -cumárico ^e					270	426
Ácido cinámico ^e					270	420
Tirosol ^e					270	420
Ácido cafeico ^e					270	420

Los métodos establecidos para determinar las denominaciones de los aceites de oliva, en función de aspectos como la composición o las características sensoriales, vienen regulados por el reglamento de ejecución (UE) 2019/1604 de la comisión de 27 de septiembre de 2019 [33], los cuales se fundamentan en la determinación de los parámetros más significativos que se indican a continuación:

- **Grado de acidez:** Expresado como porcentaje de ácido oleico libre, determina la cantidad de ácidos grasos libres en el aceite de oliva procedentes de la hidrólisis de los triacilgliceroles por efecto de las lipasas. Valores altos de este parámetro, superiores al 2%, son indicativos de un deterioro del fruto.
- **Índice de peróxidos:** Expresado como equivalentes de O₂ por kilogramo de aceite, determina los productos de oxidación de las grasas y se asocia con la pérdida de valor nutricional por pérdida de compuestos como la vitamina E. Su límite para el consumo es de 20 meq/kg. Esta se ve promovida por la exposición del aceite al oxígeno del aire y también a la radiación electromagnética como la luz solar.
- **Absorbancia en el ultravioleta:** Se fundamenta en la medida espectrofotométrica del coeficiente de extinción molar a las longitudes de onda de 232 y 270 nm.
 - **K₂₃₂:** Es un valor indicativo de los productos de oxidación tempranos como dienos, trienos o peróxidos, por lo que su valor se correlaciona con el índice de peróxidos. Su límite máximo para aceites de oliva virgen extra es de 2,5 y para aceites vírgenes de 2,6.
 - **K₂₇₀:** Su valor se relaciona con la presencia de productos de oxidación secundaria como aldehídos, cetonas y trienos conjugados, algunos de los cuales provienen del proceso de refinación del aceite.

La medida del ΔK (definida como $K_m - \frac{K_{m-4} + K_{m+4}}{2}$, donde K_m es la extinción específica a la longitud de onda m) que no debe ser superior a 0,01, se relaciona con la intensidad de la refinación y es un parámetro utilizado para la determinación de adulteraciones.

- **Contenido en ésteres metílicos y etílicos:** Se trata de un parámetro relativamente reciente basado en la determinación cromatográfica del contenido en ésteres metílicos (FAMEs, del inglés *fatty acid methyl esters*) y etílicos (FAEEs, del inglés *fatty acid ethyl esters*) de los ácidos grasos del aceite de oliva, los cuales se forman por aceitunas fermentadas cuyo aceite ha de ser sometido a procesos de desodorización. Su uso es de especial utilidad para la determinación de mezclas fraudulentas y se mide en mg/kg de ésteres etílicos o metílicos, no siendo esta cantidad mayor de 15 mg/kg para ésteres etílicos y de 40 mg/kg para la suma de ambos ésteres, así como una relación mayor de 1 entre ésteres metílicos y etílicos.
- **Calidad sensorial:** Consiste en la evaluación estandarizada por las distintas normas del COI al respecto reflejada en el actual Reglamento (CEE) 2568/91 de 11 de Julio [34], que establece la clasificación de atributos positivos y negativos de parámetros olfato-gustativos y retro nasales de un panel de 8 a 12 catadores, en condiciones estandarizadas por dicha norma, para poder atribuir en base a una ficha de cata, la denominación que se le atribuye al aceite catado.

La obtención de aceites de calidad, entendida esta como el conjunto de cualidades deseables por el consumidor, ya sea en la componente nutricional, sanitaria u organoléptica de estos, depende de diversos factores. Estos pueden ser factores agronómicos, en los que se incluye la variedad, el medio agrológico, la recolección o el transporte, así como factores industriales, entre los que destaca el proceso de extracción, almacenamiento

y envasado, donde existen variables de carácter físico o químico, como pueden ser la exposición a agentes químicos, radiación, o temperatura, que a la postre influyen en su composición nutricional y en sus características sensoriales y, por ende, en su calidad.

Todos estos factores, de manera intrínseca o extrínseca, modifican la composición química del producto, pudiendo así establecerse patrones analíticos que sirven de parámetros de calidad, regulados por el Reglamento (CEE) 2568/91 de 11 de Julio [34], y que permiten caracterizar de manera precisa las denominaciones de calidad establecidas, con objeto de garantizar la calidad que demanda el consumidor a la hora de adquirir el producto.

Los parámetros establecidos por este reglamento se determinan con métodos de referencia altamente empíricos, por lo que su exactitud depende de factores experimentales, que a su vez se traducen en una inversión importante en equipos, así como de personal, tiempo o eluyentes. Todo ello, unido a las incipientes demandas en el ámbito de eficiencia, sostenibilidad y reducción de tiempos de análisis, recursos materiales y fungibles hacen necesario el desarrollo de nuevas herramientas analíticas. Para dar respuesta a las necesidades comentadas, la técnica elegida en esta Tesis Doctoral ha sido la fluorescencia molecular para caracterizar a los compuestos de la fracción minoritaria. Se han obtenido matrices de datos de excitación-emisión (EEMs, del inglés *excitation emission matrixes*) que se analizan con algoritmos quimiométricos de segundo orden, lo que ha permitido establecer métodos de caracterización rápidos, sin necesidad de emplear técnicas selectivas que implican un mayor coste en material e instrumentación, en presencia de interferencias debido a la posibilidad que estos algoritmos permiten poder aislar la información analítica de relevancia sin necesidad de separar los interferentes, hecho que se conoce como “ventaja de segundo orden”. Además, la posibilidad de realizar este

tipo de técnicas sin tratamientos previos de las muestras también ha sido posible con la incorporación de la modalidad de fluorescencia *front-face*, la cual consiste en variar el ángulo de incidencia del haz de luz para minimizar alteraciones en los análisis como el efecto de filtro interno.

Una vez descrita la composición de los aceites de oliva según su calidad, en base a los analitos presentes con propiedades fluorescentes, a continuación, se comentarán otros factores que afectan a estos compuestos durante todo el proceso de producción del aceite de oliva.

1.1.3. Factores críticos en la producción de los aceites de oliva y métodos de análisis.

Los factores que afectan a la calidad y composición de los aceites de oliva pueden clasificarse en tres categorías: factores de *pre-producción*, de *producción* y de *post-producción*. Los factores de *pre-producción* incluyen aquellas variables previas al proceso de producción que afectan a la calidad del aceite de oliva como, por ejemplo, la variedad, la edad del árbol, el área geográfica de cultivo y el relieve, el clima, el suelo, etc. Los factores del proceso de producción son aquellos que afectan a la producción del fruto, como son la madurez, las técnicas agronómicas, o el almacenamiento del fruto y la extracción del aceite y, en los factores de *post-producción*, se incluye el envasado y el almacenamiento del producto final

A. Factores de pre-producción

Variedad

El genotipo de la variedad cultivada puede considerarse como el principal factor que determina el perfil de ácidos grasos, compuestos fenólicos y volátiles. Bouckachabine et al. [35] llevaron a cabo un estudio donde caracterizaron, durante 6 y 8 años, 28 y 74 variedades de aceituna, respectivamente, y caracterizaron el efecto de la variedad sobre el perfil de

ácidos grasos, fenoles totales y amargor, concluyendo que la variedad era un parámetro más crítico que el estado de maduración o el año de cosecha. Tura et al. [36], en un trabajo sobre 18 variedades distintas analizadas, durante 4 años, en el mismo terreno de Italia, observaron que la influencia de la variedad sobre los compuestos antioxidantes, tocoferoles y ácidos grasos poliinsaturados era significativa en los parámetros estudiados. Por otra parte, Gómez-Rico et al. [37] detectaron una clara diferencia estadística en el contenido de oleuropeína entre variedades típicas de España como son Arbequina, Cornicabra, Morisca, Picolimón, Picudo y Picual. Asimismo, Franco et al. [38] estudiaron la presencia de polifenoles y tocoferoles en aceites de oliva provenientes de aceitunas de la comunidad autónoma de Extremadura, comprobando diferencias significativas en la proporción de estos entre aceites de siete variedades distintas, las cuales fueron Arbequina, Carrasqueña, Corniche, Manzanilla Cacereña, Morisca, Picual y Verdial.

Todo ello muestra la gran dependencia que el tipo de variedad puede tener sobre la composición del producto final y, en especial, sobre los componentes minoritarios fluorescentes objeto de análisis de esta memoria, como pueden ser los polifenoles, tocoferoles, clorofilas y carotenoides.

Edad del árbol

En términos de productividad, la edad del olivo es un factor a tener en cuenta, ya que los olivos de secano alcanzan una mayor productividad en edades de 100 años, mientras que, en los de regadío, la mayor productividad puede alcanzarse entre los 50 y los 65 años. No obstante, en términos de composición, Rouas et al. [39] han revelado que el contenido en fenoles es mayor en árboles de entre 30 y 50 años que en árboles menores de 30 años. Asimismo, Bouchaala et al. [40] investigaron la influencia de la edad del árbol en los parámetros físicoquímicos de los aceites de oliva en

la variedad “Oueslati” de Túnez y encontraron diferencias significativas en dichos parámetros según el estado de maduración. Mientras que los árboles de menor edad producían más aceite, en estados de maduración tardíos, los árboles de mayor edad producían aceite con parámetros de calidad mejores. Por otro lado, en una línea de investigación similar, Chtourou et al. [41], compararon el efecto de la edad de los olivos sobre el perfil de ácidos grasos saturados, escualeno, alfa tocoferol y tocoferoles totales, detectando diferencias significativas entre los árboles más jóvenes con diferentes grados de maduración, concluyendo que el estado de maduración óptimo depende de la edad del árbol.

Temporada de cultivo

La productividad de las cosechas en el cultivo del olivo suele seguir ciclos de dos años, en los que un año la cosecha resulta fructífera, seguida de otro año en el que resulta pobre. Esto se asocia al fenómeno de la vecería, el cual se debe a la tendencia a producir un mayor número de frutos y, al perder las reservas de energía y nutrientes, el árbol pasa a un estado de ahorro de energía, y este proceso se suele prevenir con la poda (aclareo de ramos fructíferos).

Este hecho también provoca una alteración en la composición del aceite de oliva que es importante tener en cuenta a la hora de analizar la calidad del producto obtenido. Un ejemplo es el impacto significativo en la producción de compuestos fenólicos observado por Salvador et al. [42] y por Mansouri et al. [43]. Estos últimos encontraron que las variables asociadas a los parámetros antioxidantes cambiaban significativamente en aceites procedentes de las mismas variedades entre las temporadas 2013/2014 y 2012/2013. Resultados similares encontraron Romero et al. [44] que detectaron diferencias significativas en el contenido total en fenoles de cultivos de las temporadas 1996/1997, 1997/1998 y 1998/1999

Origen Geográfico

El origen geográfico es un factor determinante en la composición del aceite de oliva, ya que este influye en la fracción insaponificable, lo que se traduce en aceites de diferentes características sensoriales. Al ser una especie adaptada al clima mediterráneo, el relieve, la altitud, la latitud y las características del suelo afectan a la maduración de la aceituna, ya que esta es sensible a las bajas temperaturas en estados tempranos de maduración.

De este factor, la composición en ácidos grasos es considerada como la variable más relevante de cara a discriminar el origen geográfico de los aceites de oliva. Beltrán et al. [45] han detectado la diferencia entre aceites de distintas zonas del oeste de España, a través del análisis de componentes trazas en los aceites de oliva y orujo, con un 93% de precisión. Borges et al. [46] han comparado aceites de la variedad “*Arbequina*” producidos en Brasil con otros producidos en España, a través de la evaluación del perfil de ácidos grasos y el color, encontrando diferencias significativas entre ellos, atribuibles a la altitud de determinadas zonas, a las precipitaciones y a la temperatura, factores que acaban afectando de manera notable a la composición final del aceite de oliva.

Factores Edafoclimáticos

El medio agrológico en el que se sitúa el cultivo del olivo está estrechamente relacionado con su origen geográfico, ya que este se caracteriza por las condiciones climatológicas y por el tipo de suelo. Estos factores influyen en la floración, el crecimiento y la maduración del fruto y, por tanto, en la composición final de los aceites de oliva.

Los factores climáticos incluyen la temperatura, las precipitaciones y la humedad, entre otros factores. Con respecto a las precipitaciones, se ha puesto de manifiesto que estas influyen de manera significativa en la composición en ácidos grasos [47], compuestos fenólicos y *o*-difenólicos.

Con respecto a la influencia de la temperatura, Romero et al. [47] han puesto de manifiesto la existencia de una relación significativa entre esta y la menor cantidad de compuestos volátiles, producidos vía lipoxigenasa, asociada a la evapotranspiración y al riego. En otro estudio desarrollado por Arafat S et al. [48] se estableció que la baja humedad relativa, en combinación con las altas temperaturas, conduce a un descenso en la síntesis de ácido linoleico. Sin embargo, durante el periodo de cosecha, las bajas temperaturas producen una disminución en clorofilas, carotenoides y en el contenido en α -tocoferol [49].

Por otro lado, el suelo proporciona el sustento necesario en nutrientes tanto orgánicos como inorgánicos, así como el soporte con la textura necesaria para que las raíces se fijen en él y capten el agua, por lo que el suelo influye de muchas maneras en la composición final de los aceites de oliva. Entre los estudios que muestran la influencia del tipo de suelo sobre los componentes fundamentales del aceite de oliva pueden destacarse los realizados por Rached et al. [50], sobre aceites procedentes de la variedad de Túnez "*Chemlali*", donde comprobaron que había diferencias significativas en las cantidades entre compuestos volátiles y esteroides, concluyendo que diferentes tipos de suelos afectaban a estas variables junto a la cantidad de α -tocoferol y polifenoles totales en rangos de 123 mg/kg en suelos pardos, a 284 mg/kg en suelos arcillosos, así como menores niveles de peróxidos en suelos arcillosos y mayor cantidad de ácidos grasos monoinsaturados. Por otro lado, para suelos pardos existe una mayor cantidad en el contenido de carotenoides (1,83 mg/kg) y una mayor viscosidad del aceite sobre otros tipos de suelos.

Prácticas de cultivo

Teniendo en cuenta que las prácticas de cultivo se realizan con objeto de mejorar la cantidad y la calidad de las cosechas, estas deben ser

diseñadas para adaptarse a las características de los diferentes cultivos. Dentro de estas prácticas, que terminan afectando de manera importante a la composición final, se pueden destacar el riego, la fertilización, el sistema de cultivo y la poda, los cuales, a su vez, afectan a la composición final en compuestos fenólicos.

- **Fertilización:** La fertilización produce un efecto similar al efecto que el tipo de suelo puede producir sobre los componentes fundamentales de los aceites de oliva, ya que esta se relaciona con los nutrientes accesibles al olivo, para que pueda realizar los procesos biosintéticos que terminan por determinar la composición final del aceite de oliva. Estudios realizados por Fernández-Escobar et al. [51] concluyen que la aplicación de nitrógeno produce un decrecimiento en el contenido total de polifenoles y, por lo tanto, en la estabilidad oxidativa de los aceites de oliva. Por otro lado, no se han detectado diferencias significativas en el contenido en clorofilas y carotenoides ni en la composición en ácidos grasos. Con respecto a los fertilizantes N, P y K, en un estudio realizado en Portugal por Jordao et al. [52] determinaron que la productividad del aceite se incrementaba debido al nitrógeno y la estabilidad oxidativa aumentaba tras una fertilización con fósforo.
- **Riego:** El régimen hídrico al que se somete el olivo afecta de manera profunda a la composición fenólica de este, así como a las características sensoriales y otros componentes que derivan en la calidad del fruto. El aspecto más importante que destacan los diferentes estudios es la disminución en el contenido total en polifenoles a medida que aumentaba el riego. Yamani et al. [53] estudiaron el contenido fenólico en aceites procedentes de la variedad "*Piccholine Marroquí*", donde concluyeron que el riego provocaba una disminución en los polifenoles totales. Con respecto al α -tocoferol, no se han detectado diferencias significativas con respecto a diferentes tratamientos de riego [54].

- **Poda:** El objetivo de la poda del olivar es mantener el equilibrio entre las funciones vegetativas y reproductivas del árbol, ya que este influye en la densidad vegetativa y, por ende, en la floración y en los procesos que configuran la composición del fruto. Además, el efecto de la sombra y el sol sobre los frutos hace que los procesos biosintéticos sean diferentes, situación que ha sido estudiada por varios autores, que han comprobado que en las capas altas se produce una mayor cantidad de polifenoles que en las bajas [55], mientras que García-Inza et al. [56] destacan que los frutos procedentes de las zonas de sombra no se ven afectados en su composición en ácido oleico.

Estas prácticas conforman un grupo variado y multifactorial de variables que pueden afectar a las características de calidad del aceite de oliva y que son objeto de análisis en esta Tesis Doctoral, a través de las técnicas y algoritmos que se explicarán en apartados posteriores.

B. Factores durante la fase de producción

Período de cosecha y estado de maduración

La maduración de la aceituna es un proceso lento que dura varios meses, los cuales se relacionan con varios factores como son la variedad, la edad y salud del árbol, el suelo, la fertilización y la incidencia de luz solar [57]. Por lo tanto, es importante saber el momento de cosecha de la aceituna ya que, a través del estado de madurez en el que se coseche, este será indicativo de la cantidad de ácidos grasos, compuestos fenólicos y cantidad de clorofilas.

En lo que afecta específicamente a la composición y calidad del aceite se pueden resaltar varios estudios en los que se puede concluir que el rendimiento de extracción del aceite aumenta en estados de maduración óptimo, aunque con un incremento de la acidez libre [58]–[60]. También se ha puesto de manifiesto que, con la maduración, el contenido en humedad

y las clorofilas disminuyen [60], mientras que el contenido en α -tocoferol [61] baja y el contenido en γ -tocoferol aumenta [62]. Con respecto al contenido en compuestos fenólicos, existen estudios contradictorios, unos que observan una relación positiva entre su contenido y el momento de cosecha [63], mientras que otros muestran resultados contrarios [58, 59].

En conclusión, se puede afirmar que el estado de maduración del fruto está fuertemente relacionado con las características del aceite, y este a su vez depende de variables geográficas, climáticas y genéticas del propio fruto. En este sentido, teniendo en cuenta todas las posibles variables, algunos artículos de investigación como el de Mahhou et al. [64] evidencian que la cosecha tardía (entre finales de noviembre y diciembre), es la óptima para producir aceite de oliva de alta calidad. Giuffré et al. [65] han corroborado este periodo de recolección analizando aceites obtenidos de la variedad "*Roggianella*". Sin embargo, como se ha comentado antes, esta estimación depende de factores como la variedad de la aceituna. Por ejemplo, Caporaso N. et al. [66] determinaron la fecha adecuada de cosecha para aceites de variedades españolas ("*Cornicabra*", "*Manzanilla Castellana*", "*Manzanilla Cacereña*", "*Picual*"), para un perfil de compuestos volátiles y análisis sensorial óptimos, a través de un análisis multifactorial, observando que la fecha de cosecha óptima variaba de manera significativa según la localización del cultivo y la variedad, lo cual es indicativo de la inmensa variabilidad que la determinación de calidad de los aceites de oliva.

Método de cosecha

Dentro de los métodos de cosecha de la aceituna se pueden diferenciar los tradicionales y los mecanizados, cuyo uso dependerá de la disponibilidad económica y las necesidades de producción. Se ha estudiado el impacto del método de cosecha en la calidad de los aceites de oliva. En Turquía, Saglam et al. [67] han comparado ambos métodos de recolección, observando que no había diferencias significativas en el contenido en acidez y valor de peróxidos para aceites de oliva de la variedad “*Gemlik*”.

Almacenamiento de las aceitunas

Tras la cosecha es preciso un almacenamiento correcto de las aceitunas previo a su tratamiento, ya que se pueden producir alteraciones negativas en las características sensoriales del aceite. También es importante el tiempo de almacenamiento, ya que este puede producir una alteración en la composición del aceite afectando a su estabilidad oxidativa. Stefanoudaki-Katzouraki [57] concluyeron que los aceites de oliva deberían ser procesados en las primeras 24 horas tras la cosecha para mejorar el rendimiento de extracción y, si no fuese posible, se recomendaba almacenar la aceituna en lugares bien ventilados a menos de 25°C y una humedad relativa menor al 75%.

Método de extracción

Las metodologías de extracción del aceite de oliva son muy diversas y, entre ellas, se pueden destacar las siguientes: el proceso tradicional mediante molienda, batido, prensado y decantación; el proceso en tres fases, donde hay una molienda en molino de martillos, seguida del batido y la centrifugación de la pasta, por un lado, y, por otro lado, del agua; y el proceso en dos fases, en el que solo hay un proceso de centrifugación.

Independientemente del procedimiento utilizado, durante el proceso de extracción se pierden compuestos fenólicos. En un estudio realizado por Boussahel et al. [68] sobre aceitunas de la variedad “*Coratina*” se observó que el contenido en polifenoles fue de 441 mg/L cuando se utilizaba un molino metálico, frente a 228 mg/L en molino de piedra. Por otro lado, en relación con el contenido en volátiles tales como hexanal, trans-2-hexanal y los ésteres C6, Inglesse et al. [69] han descrito que aumenta su rendimiento con nuevos tipos de molinos como son los de cuchillas. Además, Preedy et al. [54], haciendo referencia a un estudio realizado por Sanz et al., indican que el contenido en carotenoides como la luteína, la zeaxantina y el β -caroteno se veían afectados por las temperaturas alcanzadas durante el proceso de molienda.

Finalmente, la calidad total de los aceites de oliva según el método de extracción puede diferenciarse analizando parámetros como acidez, PV (valor de peróxidos), fenoles, humedad, impurezas y coeficientes de absortividad K232 y K270. En la **Tabla 5**, recopilada por Vekiari S. et al. [70], se observan las diferencias significativas existentes entre el método de extracción tradicional y el de centrifugación, concluyéndose que el contenido en fenoles totales es mayor en el método tradicional.

Tabla 5. Análisis iniciales (media sobre nueve réplicas) de muestras de aceites de oliva obtenidos por diferentes sistemas de extracción [70].

Sistema de extracción	Acidez (%)	PV (meq/O ₂)	Fenoles (mg/L)	Humedad (%)	Impurezas (%)	K ₂₃₂	K ₂₇₀
Centrifugación	0,37 ^{a*}	2,75 ^a	130 ^a	0,321 ^a	0,218 ^a	1,39 ^a	0,127 ^a
Tradicional	0,39 ^b	3,37 ^b	196 ^b	0,39 ^b	0,431 ^b	1,40 ^a	0,131 ^b

Los valores seguidos de la misma letra en vertical no se diferencian estadísticamente con un $p = 0,05$

C. Factores de Post-Producción

Una vez el aceite de oliva es extraído del fruto, este se almacena hasta su consumo. No obstante, la duración del almacenamiento, así como el tipo de envase y las condiciones externas, como la humedad relativa y la temperatura, pueden afectar a la composición final y la calidad del producto. Dentro de estos factores, se pueden destacar los siguientes aspectos:

Embotellado del aceite de oliva

Dentro del embotellado del aceite de oliva, el tipo de material a envasar puede interactuar con los componentes propios del aceite de oliva o favorecer procesos de oxidación a través de la acción de la luz o la temperatura. Para entender cómo los envases podían causar pérdidas de calidad por estos motivos, Méndez y Falqué [71] han realizado un estudio en el que evaluaron el tiempo de almacenamiento (3 y 6 meses) y el tipo de envase sobre diversos aceites de oliva. Este estudio concluyó que entre el envasado con botellas de PET, botellas de PET cubiertas con papel de aluminio, botellas de vidrio, estaño y tetrabrik, los mejores envases eran el de estaño y tetrabrik, ya que, tras 6 meses de almacenamiento, las muestras que inicialmente tenían una cantidad de 285 ppm de ácido gálico acababan con 70 y 74 ppm en latas de estaño y tetrabrik, respectivamente, frente a 39 ppm en botellas de plástico, o 49 ppm en plástico opaco. Esto indica que la cinética de la oxidación de estos productos depende tanto del propio material como de la geometría del envase, ya que de esta depende la presión parcial de oxígeno inicial en el espacio de la cabeza de la botella que promueve los procesos de oxidación. De este modo, para preservar el aceite de los procesos de oxidación, los autores propusieron el uso de botellas diseñadas a partir de materiales plásticos innovadores capaces de capturar el oxígeno. Además, para reducir la presión de oxígeno en la cabeza de la

botella se propuso en otro estudio, realizado por del Nobile M.A. et al. [72], el envasado bajo atmósfera de nitrógeno.

Almacenamiento del aceite de oliva

Tanto el tiempo como las condiciones de almacenamiento son críticas para la conservación sin alterar la composición del aceite de oliva, ya que condiciones inadecuadas promueven las reacciones de oxidación que favorecen la aparición de productos de oxidación secundaria indeseables. Por este motivo, las condiciones de almacenamiento han de ser rigurosas para prevenir cualquier tipo de pérdida en la composición nutricional o en las características sensoriales del producto.

De hecho, la influencia de diferentes parámetros físico-químicos durante el tiempo de almacenamiento ha sido ampliamente estudiada. Abdalla et al. [73] han realizado un estudio con diferentes cooperativas localizadas al norte de Marruecos, investigando el efecto del tiempo de almacenamiento. Los resultados obtenidos pusieron de manifiesto que, a partir de los 12 meses, a temperatura ambiente, y en condiciones de obscuridad, se producían cambios significativos en parámetros físico-químicos como la acidez, que aumentaba del 0,31 a 0,69% o el valor de peróxidos que aumentaba de 9,57 a 14,64 meqO₂/kg. Stefanoudaki-Katzoura. E [57] observaron un incremento de los coeficientes K₂₃₂ (2.40) y K₂₇₀ (0.19) en muestras de aceite de oliva almacenadas en botellas abiertas al aire libre, comparadas con botellas almacenadas durante 2 meses bajo atmósfera de nitrógeno (K₂₃₂ = 1,86, K₂₇₀ = 0,14).

En cuanto a la fracción fenólica, esta también se ve afectada por las condiciones de almacenamiento. Di Stefano y Melilli [74] observaron que para aceites obtenidos a partir de aceitunas de la variedad "*Nocellara Messinese*" inmaduras, se producía una disminución en el contenido en oleaceína y oleochantal, mientras que se producía un incremento en tirosol e

hidroxitirosol, lo cual se asocia a la degradación de los derivados de la ligstrosida y de la oleuropeína. Además, también hay bibliografía que asocia las condiciones de almacenamiento al impacto en los pigmentos, produciendo una degradación de las clorofilas y los carotenoides en casos en los que se exponga el aceite a la luz, el oxígeno y la presencia de las propias enzimas que promueven los procesos de oxidación [75].

Temperatura

El último factor de interés que influye en la alteración de los componentes que dan su valor nutricional y sensorial y que conforman la calidad a los aceites de oliva, es la temperatura, ya que esta promueve la cinética de oxidación de estos, afectando a las cualidades nutricionales.

La fritura contribuye a la textura y sabor de los alimentos fritos, sin embargo, la elevada temperatura que alcanza el aceite de oliva produce una pérdida en la composición en compuestos de interés nutricional. Santos et al. [76] han estudiado los diferentes procesos de cocinado como son la fritura, el hervido, incluido en microondas, en condiciones operativas como son el tiempo de cocinado o las cantidades de alimento, proponiendo reducir el tiempo de calentamiento al mínimo independientemente de la técnica de cocinado, ya que la fracción fenólica y la capacidad antioxidante disminuían con el cocinado de verduras típicas de la dieta mediterránea.

Siendo más concretos, el proceso de cocinado afecta sobre todo a los polifenoles sencillos como el tirosol y el hidroxitirosol, mientras que la oleuropeína y la ligstrosida aglicona soportan mejor estos procesos. Peri C, et al. [77] y Brenes et al. [78] encontraron que la oxidación térmica de los aceites de oliva a 180°C producía una reducción significativa del tirosol y el hidroxitirosol, además del α -tocoferol (de 172 mg/kg a niveles de traza en 24 h en aceites de oliva de la variedad "*Picual*") y los triglicéridos. No obstante, los lignanos son los menos afectados por la termo oxidación a

tales temperaturas. Por otro lado, no se detectaron influencias significativas en la acidez libre y la cantidad de ácido mirístico. Sin embargo, el índice de estabilidad oxidativa, la actividad antirradicalaria y el contenido en fenoles disminuyeron en función del tiempo.

Con respecto a la formación de productos de oxidación secundaria, se ha descrito la presencia en aceites de oliva de algunos de ellos, como pueden ser aldehídos, acroleína y crotonal. La emisión de estos compuestos durante la fritura usando diferentes aceites de oliva fue investigada en estudios como los de Fullana et al. [79], quienes detectaron la presencia de 16 aldehídos, siete alcanales, ocho 2-alquenes y 2,4-heptadien en vapores de cocina.

Además, durante las etapas de formación de aldehídos también se producen hidroperóxidos en la primera etapa, lo que explica el elevado valor de peróxidos en aceites calentados a temperaturas inferiores a 180 °C. Además, estos compuestos, debido a su inestabilidad química, se descomponen originando aldehídos y cetonas [80].

Con todo ello, se observa que la calidad composicional y sensorial de los aceites de oliva es producto de un gran número de factores que, a su vez, dependen de diversas variables que convierten la caracterización y clasificación de estos en un problema de difícil solución y, con la necesidad de establecer costosos métodos analíticos de separación de aquellos componentes que permitan una correcta caracterización de estos. Por esto, uno de los objetivos de esta tesis doctoral es simplificar esta etapa a través de los modelos matemáticos basados en algoritmos quimiométricos.

Adulteraciones y fraudes en la comercialización de los aceites de oliva.

Otro aspecto de especial relevancia en el mercado del aceite de oliva son las adulteraciones y los fraudes, cuya detección rápida y eficaz es de especial importancia tanto para los consumidores como para los

empresarios. El Real Decreto 1945/1983 de 22 de junio, por el que se regulan las infracciones y sanciones en materia de defensa del consumidor y de la producción agro-alimentaria [81], define fraude como “*la elaboración de productos agrarios y alimentarios mediante tratamientos o procesos que no estén autorizados por la legislación vigente, así como la adición o sustracción de sustancias o elementos que modifiquen su composición con fines fraudulentos*”. También se consideran fraudes alteraciones en otros aspectos como la naturaleza, calidad, riqueza, peso, exceso de humedad o cualquier otra discrepancia que existiese entre las características reales de la materia o elementos que se trate y las ofrecidas por el productor, fabricante o vendedor. En el ámbito de los aceites de oliva, cualquier etiquetado establecido con una denominación no acorde a las especificaciones establecidas por la legislación vigente o cualquier mezcla fraudulenta detectable por los medios técnicos actuales disponibles se considera un fraude del producto.

Debido a que el aceite de oliva tiene un gran interés económico para todos los sectores interesados, se han llevado a cabo diferentes iniciativas con el fin de establecer un protocolo de seguimiento y alerta ante este tipo de prácticas fraudulentas, cuya incidencia se ha recogido en una revisión llevada a cabo a través de un estudio internacional realizado en 2021, dirigido por interesados en el sector oleícola, a través principalmente de la base de datos sistema de alerta rápida para la alimentación (RASFF) [82]. En dicha revisión se concluyó que los fraudes más comunes eran la venta de aceite de oliva virgen como aceite de oliva virgen extra y las adulteraciones de aceites de oliva con otros aceites vegetales (girasol, maíz, palma, lino, etc.), siendo estas vendidas con la denominación de aceite de oliva. También se puso el foco en nuevas tendencias en fraudes como es la adición de aceites desodorizados y la mezcla con aceites realizados bajo una centrifugación secundaria de la pasta del aceite de oliva (*remolido*).

En el ámbito analítico, Salah W y Nofal [83] hicieron una revisión de los principales métodos analíticos para la detección de fraudes y adulteraciones en distintos tipos de aceites comestibles, entre ellos pueden destacarse el uso del perfil de ácidos grasos para detectar adulteraciones de aceites de oliva con aceites vegetales de menor calidad a un nivel inferior al 5% [84], el uso de la constante ΔECN_{42} (definido como la diferencia entre el contenido teórico en triglicéridos y el determinado mediante HPLC) para detectar niveles bajos de aceite de girasol, soja o algodón en aceites de oliva [85], el uso de los glicerofosfolípidos para diferenciar aceites vírgenes de aceites refinados, ya que estos se encuentran en mayor medida formando aminas en los aceites refinados [86], así como la referencia a nuevas tendencias en métodos analíticos para la detección de fraudes y adulteraciones, como pueden ser el uso de fibras ópticas y espectroscopía infrarroja. En este último ámbito se ha centrado el desarrollo de los objetivos de la presente tesis.

1.2. Ciruelas

1.2.1. Definición e impacto económico

La segunda matriz alimentaria objeto de análisis en esta Tesis Doctoral es la ciruela, la cual en términos de consumo establecidos por la Norma CXS 242-2003, enmendada en 2017, del CODEX ALIMENTARIUS [87] se cataloga como fruta de hueso, a la cual se le aplican dichas normas de calidad para su venta y consumo.

Taxonómicamente, estas frutas de hueso se enmarcan dentro del género *Prunus* y, dentro de este, las ciruelas presentan diferentes variedades, de las cuales, comercialmente destacan la ciruela hexaploide europea (*Prunus domestica*) y la ciruela diploide japonesa (*Prunus. salicina* e híbridos de esta). La ciruela europea cuyo origen es aún incierto, presenta una diversidad palpable en la cantidad de variedades presentes en la

actualidad, mientras que la ciruela europea debe su extensión a la resistencia que ofrece a enfermedades como la enfermedad de *Sharka*. Por otro lado, las ciruelas japonesas se hibridan fácilmente, por lo que, a día de hoy, siguen en proceso de hibridación utilizando criterios de calidad, valor funcional, productividad y adaptación a enfermedades. De estas dos especies, la variedad en la que se ha enfocado esta investigación ha sido la variedad “*Friar*”, que es una variedad de la especie japonesa *Prunus Salicina Lindl*, cuyo cultivo es cada vez más extendido en regiones como Murcia, Badajoz o Valencia.

En términos de impacto a nivel productivo y económico se puede destacar que, según los datos recogidos por el sistema de registro de estadística de la Organización de las Naciones Unidas para la Alimentación, actualizados al año 2020 (FAOSTAT 2020) [4], España es el sexto productor a nivel europeo con una producción neta de 153.590 toneladas.

Por otro lado, a nivel del territorio español, las estadísticas respectivas al cultivo del ciruelo son recogidas por el Ministerio de Agricultura, Pesca y Alimentación en su Anuario de Estadística respectivo, en este caso a los datos registrados en 2020 (MAPA 2021a) [88]. De las 1.102.380 hectáreas destinadas en España a frutales no cítricos, el cultivo de los frutales de hueso ocupa un total de 206.216 hectáreas, siendo esta superficie una proporción del 13% del destinado al total. Dentro de la superficie destinada al cultivo de frutas de hueso, 14.406 hectáreas se dedican al cultivo de ciruelo, lo que supone un 7% del total del cultivo de fruta de hueso, donde la parte mayoritaria corresponde a la superficie destinada al cultivo de melocotonero, nectarina y cerezas, siendo, por último, destacable que en este tipo de frutales predomina el cultivo en regadío dadas las características agronómicas de dicho cultivo.

En cuanto a la distribución de la productividad por regiones conviene mencionar que, según el Anuario de Estadística, Extremadura, Andalucía, Murcia y la Comunidad Valenciana son los mayores productores, siendo Extremadura la región más destacada, con una producción neta de 77.600 toneladas según el último registro actualizado a 2020.

En cuanto a los términos de producción y rendimiento del ciruelo, en las **Figuras 9, 10 y 11**, correspondientes a la serie histórica desde 2010 hasta 2020, se representa la evolución de la superficie cultivada, la producción y el valor de la producción. Se puede observar un decrecimiento paulatino en el uso de superficie cultivada en consonancia con un descenso, algo más irregular, en la producción de ciruela y en el valor económico de esta.

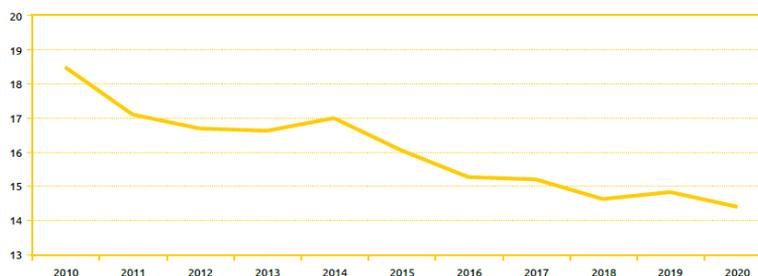


Figura 9. Evolución de la superficie destinada a la producción de ciruela desde 2010 hasta 2020 (miles de hectáreas) [88].

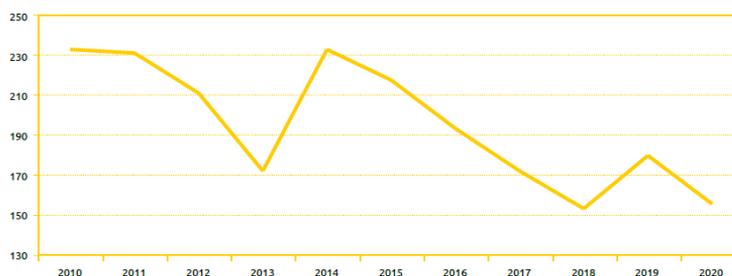


Figura 10. Evolución de la producción de ciruela desde 2010 hasta 2020 (miles de toneladas) [88].

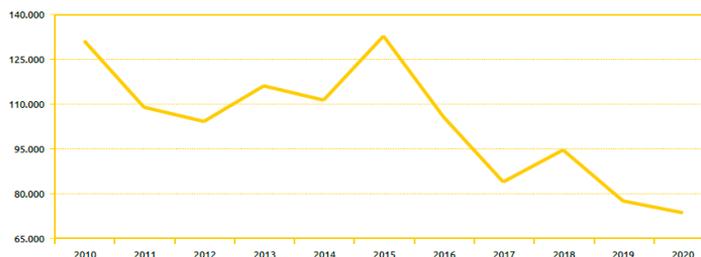


Figura 11. Evolución del valor de la ciruela desde 2010 hasta 2020 (miles de euros) [88].

La disminución de la superficie cultivada unida a años de condiciones climáticas adversas respecto a las características de crecimiento y floración del fruto en primavera (heladas, tormentas), han provocado el descenso en la productividad, a lo cual también se le suma un menor precio medio percibido por los agricultores, debido a factores como el aumento de la oferta con mercados exteriores y la concentración de la demanda entre los distribuidores.

Con el objetivo de destacar la calidad del producto, el Plan de Medidas para la mejora del sector de la fruta dulce promovido por el Ministerio de Agricultura, Pesca y Alimentación de 2018 (MAPA 2018 b) [89] establece una serie de medidas para promover una calidad diferenciada de estos productos, con objeto de equilibrar la demanda del producto y aumentar los precios percibidos por los agricultores. En este sentido, se propone mejorar la contribución de los programas operativos de desarrollo rural, el control y seguimiento de las nuevas plantaciones, fomentar la producción ecológica y fomentar la calidad a través del control de parámetros como el contenido mínimo de grados brix, consistencia de la fruta, períodos de almacenamiento adecuado, variedades recomendadas, variedades no recomendadas.

El desarrollo de figuras de calidad diferenciada en el sector de la fruta dulce, ya sean Denominaciones de Origen Protegidas (DOP) o Indicaciones

Geográficas Protegidas (IGP), constituyen una oportunidad para aumentar el valor de las producciones amparadas en estas figuras. En este ámbito, pese a que en otros países como Italia y Francia tienen reconocidas varias DOP para ciruelas, en España aún no existen estos reconocimientos. No obstante, se han promovido marcas de calidad por distintas asociaciones agrarias como es en el caso de Extremadura la marca CiEx® (Ciruela de Extremadura), impulsada por la asociación Afruex [90] (Asociación de Fruticultores de Extremadura), con el fin de impulsar la calidad diferenciada. Dentro de estas se enmarcan variedades de ciruela ‘*Angelino*’, ‘*Black Splendor*’, ‘*Crimson Golbe*’, ‘*Fortune*’, ‘*Golden Globe*’, ‘*Larry Am/*, ‘*Owen-T*’ y ‘*Primetime*’, todas ellas de la especie de ciruelo japonés *Prunus salicina Lindl.*

1.2.2. Composición y parámetros de calidad de las ciruelas

Al igual que otras frutas, las ciruelas presentan una composición en componentes mayoritarios caracterizada por un bajo contenido en grasas y por un elevado contenido en fibra dietética, compuesta principalmente por pectina, hemicelulosa, celulosa y lignina, lo cual se asocia a efectos beneficiosos como la prevención de enfermedades como el cáncer, la obesidad, la diabetes y patologías cardíacas.

Pese a que la composición de las ciruelas se ve afectada por factores como el clima, el suelo, las prácticas agronómicas o los factores genéticos atribuibles a la variedad, se puede establecer un valor promedio de los principales compuestos de estas, los cuales se resumen en la **Tabla 6**.

Tabla 6. Composición nutricional de las ciruelas por cada 100 g de porción comestible [91].

Componentes	Cantidades
Energía (Kcal)	51,0
Proteínas (g)	0,6
Lípidos totales (g)	Tr*
Hidratos de carbono (g)	11,0
Fibra (g)	2,1
Agua (g)	86,3
Minerales (mg):	
Ca	14
Fe	0,4
I	2,0
Mg	8,0
Zn	0,1
Na	2,0
K	214,0
P	19,0
Vitaminas:	
Tiamina (mg)	0,07
Riboflavina (mg)	0,05
Equivalentes de niacina (mg)	0,5
Vitamina B ₆ (mg)	0,05
Folatos (μg)	3
Vitamina B ₁₂ (μg)	0
Vitamina C (mg)	3
Vitamina A: Eq. Retinol (μg)	49,2
Vitamina D (μg)	0
Vitamina E (mg)	0,7

*Tr: Nivel Traza

Como se puede observar, además del bajo contenido en grasas (niveles traza) y el elevado contenido en agua y azúcares, la ciruela presenta grandes cantidades de mineral potasio y un alto contenido en vitamina A, que, junto a las vitaminas E, C y los folatos, le confieren una excelente actividad antioxidante.

En términos de actividad antioxidante, además de las mencionadas vitaminas, las ciruelas son una fuente de compuestos antioxidantes tales como flavonoides y ácidos fenólicos, los cuales se han relacionado con la reducción de padecer cáncer y otras enfermedades crónicas [92]–[94], perteneciendo estas sustancias al grupo de compuestos fenólicos ya estudiados previamente en el *apartado 1.1.2*.

Con respecto a la cantidad de compuestos fenólicos totales, esta es muy dependiente de la variedad de ciruela de la que se trate. Un estudio realizado por Siddiq et al., en 1994 [95], compararon este contenido en términos de compuestos fenólicos totales y ácido clorogénico, llegando a detectar diferencias de hasta 640 µg/g de compuestos fenólicos totales y de 28 µg/g de ácido clorogénico entre variedades *Beauty* y *Stanley*.

Según Jaiswal R. et al. [96] y Kim D. et al. [97] los principales compuestos fenólicos en estos frutos son los ácidos clorogénicos, las proantocianidinas, los flavonoides y las antocianinas.

Por otro lado, en cuanto a su potencialidad para la caracterización de estas frutas, hay que destacar que Cabrera M. et al. [98] caracterizaron la evolución de estos compuestos a través de sus propiedades fluorescentes, mediante calibración multivariante de segundo orden, según el período de cosecha.

Dentro de los compuestos minoritarios, además de los compuestos fenólicos ya citados o las vitaminas, es preciso destacar otros compuestos

fenólicos como son las antocianinas o, fuera de estos, los carotenoides, los cuales se pueden catalogar como pigmentos, ya que confieren un color rojo púrpura y amarillo anaranjado a estas frutas. Además, también se les ha asociado con propiedades protectoras frente a enfermedades crónicas según estudios epidemiológicos [99] y cuya presencia afecta de igual manera a las propiedades funcionales del alimento. Son también altamente dependientes de diversos factores, como son el estado de maduración o la variedad de la fruta u otros tratamientos pre y postcosecha. En este aspecto, Usenik et al. [100] investigaron el efecto de la maduración en la concentración de estos componentes a lo largo del tiempo para distintas variedades de ciruelo europeo *prunus domestica L.*, observando un incremento en todas las antocianinas desde estados tempranos hasta estados más tardíos de maduración del orden de 200 mg/kg. En la variedad *Friar*, Wang et al. en 2016 [101] observaron que la acumulación de estos pigmentos en total con respecto a la maduración eran dependientes de la temperatura, llegando a alcanzar valores máximos de unos 200 mg/kg, concentración característica para esta variedad.

Además de la variedad, existen otros parámetros que pueden alterar las características de color, como puede ser el pH o el estado de maduración, tal y como establecen Stintzing F. C. et al. [102]. Esta característica es de especial interés en el estudio de la calidad de las ciruelas, ya que es un indicador de su maduración y el desarrollo de metodologías analíticas fundamentadas en el procesamiento de imágenes digitales que puede usarse con fines de caracterización de la calidad del producto.

El contenido en carotenoides en distintas variedades de ciruela ha sido estudiado por Stacewicz-Sapuntzakis M. et al. [103] que observaron que los carotenoides mayoritarios en ciruelas europeas *prunus domestica L.* eran la luteína y el β -caroteno, con cantidades del orden de 2,4 y 4,3 mg/kg, respectivamente. Kaulmann et al. [104] investigaron la contribución de los

carotenoides a la capacidad antioxidante en diversas variedades de ciruelas, pertenecientes al género *prunus domestica*, y analizaron el contenido total en carotenoides, presentando contenidos que oscilaban entre 2,5 y 19,6 mg/kg, siendo los carotenoides mayoritarios el β -caroteno, que oscilaba entre los 0,8 y los 18,4 mg/kg, y la luteína, cuyos valores oscilaban entre 1 y 4,6 mg/kg, mientras que la β -cryptoxantina o la zeaxantina aparecen de manera minoritaria.

Con respecto al contenido en clorofilas, este es un parámetro de especial relevancia para caracterizar el estado de maduración entre la etapa de fructificación y la madurez fisiológica. En relación con estos compuestos, Vlaic R. et al. [105] investigaron en ciruelas de la variedad *Stanley*, *Vanat de Italia*, y *Tuleu Gras* la evolución de las clorofilas a lo largo de diferentes semanas de cosecha y diferentes posiciones del fruto en el árbol, observando en todas ellas un contenido máximo en la etapa de fructificación del orden de los 287,4 mg/kg, siendo en todos casos la clorofila a el analito mayoritario, con una concentración del orden de los 180 mg/kg. Esto también ha sido corroborado por Luo. Z, et al. [106] que analizaron el contenido total de clorofilas en ciruelas japonesas *Prunus salicina Lindl*, observando un decaimiento desde estados tempranos de maduración hasta los 12 días, de un 41%, desde los 300 mg/kg hasta los 170 mg/kg aproximadamente.

Los espectros de absorción de las clorofilas *a* y *b* presentan dos bandas, una en el rojo, correspondiente a su primer estado excitado y otra en el azul, correspondiente a su segundo estado excitado, puesto que presentan un sistema de anillos pirrólicos planos y conjugados que favorecen esta absorción sobre los 432 nm para el primer estado excitado y sobre 663 nm para el segundo estado excitado [107]. En estados de maduración tempranos, tal como se ha indicado, las clorofilas suelen ser

los pigmentos predominantes. Cuando los frutos maduran, la degradación de las clorofilas desenmascara la presencia de otros pigmentos como los carotenoides o las antocianinas, los cuáles a su vez cambian el color del fruto. Teniendo en cuenta estos aspectos concernientes al color de la ciruela, en términos de pigmentos dependientes del estado de maduración, esos son susceptibles de ser analizados a través de técnicas como el procesamiento de imágenes digitales y, por tanto, se puede llevar a cabo el procesamiento de los datos de la imagen obtenida a través de diferentes espacios de color.

Finalmente, con respecto a la maduración del fruto, es preciso destacar que dicho proceso de maduración va íntimamente ligado con la percepción de calidad del fruto, así como con la tecnología postcosecha que permita mantener dicha calidad a lo largo del tiempo, previa a la senescencia del fruto, lo cual repercute en un beneficio mutuo tanto para productores, como para distribuidores y consumidores. En este ámbito, las normas de comercialización del sector de las frutas y hortalizas establecidas por el Reglamento Delegado (UE) 2019/428 de la Comisión, de 12 de julio de 2018, que modifica el Reglamento de Ejecución (UE) n° 543/2011 [108], establece unos requisitos mínimos para su acondicionamiento y envasado y el artículo 5.1 del Reglamento (CE) 852/2004, del Parlamento Europeo y del Consejo del 29 de abril del 2004 [109], por su parte establece la obligación de los operadores de crear, aplicar y mantener procedimientos permanentes basados en los principios APPCC. Dentro de los procedimientos se pueden destacar los siguientes tipos de análisis realizados sobre frutas en particular y especificado para ciruelas en concreto:

- **Firmeza:** parámetro relacionado con la dureza de la pulpa del fruto y que es indicativo del estado de maduración, especialmente en frutas de hueso [110]. Dicha medida se realiza con un penetrómetro que mide la presión ejercida al introducir la punta del aparato en el fruto.

- **Contenido en sólidos solubles:** parámetro relacionado con la cantidad de azúcares presentes en el fruto. Estos son los principales compuestos solubles en la matriz alimentaria, a mayor contenido existe también una mayor correlación con el estado de maduración y, por ende, es un parámetro que se usa tanto como criterio de recolección en la cosecha, como de seguimiento del estado de maduración durante el almacenamiento. Se mide por refractometría y su valor se expresa en grados Brix, que se definen como la refracción provocada por un gramo de sacarosa en 100 g de disolución acuosa.
- **Acidez:** Junto con el contenido en sólidos solubles, es un parámetro empleado para establecer el índice de madurez de la fruta, ya que ambos están relacionados con los procesos enzimáticos que tienen lugar durante la madurez tanto fisiológica como organoléptica de la fruta. La relación entre el contenido en sólidos solubles y la acidez se emplea para tabular índices de madurez óptimos para la cosecha de la fruta. Para su medición se suele emplear una valoración potenciométrica con hidróxido sódico del zumo extraído de la fruta.
- **Colorimetría:** El cambio de color durante el proceso de maduración de la fruta es indicativo del cambio en la composición tanto externa como interna de los diferentes frutos objeto de estudio, por lo tanto, el color es la característica externa más importante en la determinación del punto de maduración y de la vida postcosecha y un factor determinante en la decisión de compra de los consumidores. En el caso de las ciruelas, la aparición del color rojo se debe a la síntesis de antocianinas características de la madurez del fruto y de la degradación de clorofilas. Para realizar la medición se utiliza un colorímetro. La función del colorímetro es describir la coloración de la epidermis de la pieza de fruta objeto de medición a partir del espacio de color definido por los parámetros L^* , a^* , b^* , siguiendo el estándar C.I.E. $L^*a^*b^*$, donde L es

la luminosidad, 0 representa el color negro y 100 el color blanco y los parámetros a^* y b^* representan la saturación y el tono. La descripción detallada de la teoría del espacio de color se especificará en apartados posteriores.

En relación con estos parámetros, Podestá. L. [111] realizó un estudio con ciruelas japonesas de las variedades *Black Amber Laroda* y *Angelino* donde evaluaron la intensidad de color de la piel, la firmeza de la pulpa, el contenido en sólidos solubles (CSS) y la acidez con el tiempo de maduración, estableciendo una correlación entre los parámetros analizados y el tiempo óptimo de cosecha, buscando de este modo el momento óptimo de cosecha. En dicho trabajo se observaban diferencias notables entre la evolución de los distintos parámetros con el tiempo para las tres variedades estudiadas, seleccionándose periodos aceptables de cosecha muy dispares, desde 28 días para las ciruelas de la variedad *Black Amber*, con cortos lapsos de tiempo para esta, debido al deterioro en la firmeza, hasta 70 días para las ciruelas de la variedad *Angelino* cuyos parámetros de calidad resultaban ser más estables a lo largo del tiempo en los modelos de correlación. Los valores típicos encontrados para la firmeza oscilaban entre las 16 y las 6 lb para la firmeza, los grados brix variaban entre 9 y 16 y la acidez entre 1,8 y 0,8%.

Los procedimientos utilizados para el análisis de los parámetros de calidad señalados anteriormente, a pesar de que pueden realizarse *in situ*, son destructivos, tediosos y requieren de instrumentación específica. Por ello, uno de los objetivos de esta tesis ha sido el desarrollo de nuevas metodologías analíticas no destructivas, fundamentadas en el desarrollo de modelos quimiométricos, utilizando las teorías del espacio de color, para el seguimiento del estado de maduración de las ciruelas. Para ello se han utilizado principalmente los compuestos indicadores de madurez como son las clorofilas que, debido a la fluorescencia que presentan, permitirán el uso

de fibras ópticas para establecer de manera precisa y rápida el estado óptimo de maduración, tanto para la recolección de las ciruelas, como para el alargamiento de su vida útil de cara a su posterior comercialización.

A continuación, se van a revisar los factores, tanto endógenos como exógenos, que hay que tener en cuenta a la hora de interpretar la información requerida para la caracterización del estado de maduración de las ciruelas y su momento óptimo de cosecha para la mejora de la calidad y vida comercial del producto.

1.2.3. Factores críticos en la fase de producción de las ciruelas

Respecto a la composición y los parámetros de calidad especificados en el apartado anterior, que determinan la percepción y la decisión de compra del consumidor respecto de estos productos, es preciso mencionar los factores que puedan afectar de manera positiva o negativa sobre estos. Estos los podemos clasificar como factores precosecha y postcosecha, de los cuales, los previos a la cosecha del fruto son el objeto de estudio de cara al desarrollo de metodologías analíticas para su evaluación.

A. Factores precosecha

Dentro de los factores previos a la cosecha del fruto caben destacar factores intrínsecos, como son los factores genéticos, propios de la variedad que se esté tratando, ya que, como se ha comentado en el estudio de Podestá. L. [111], se pone de manifiesto que las características del fruto y su consecuente tratamiento varían de manera crítica. En este sentido, se puede destacar la existencia de programas de mejora genética con objeto de mejorar características como la maduración temprana, la productividad, la calidad del fruto, la tendencia a la vecería y la resistencia a las condiciones climáticas adversas y al virus de la *Sharka* (PPV). Existen programas de este tipo en los países más afectados por este último virus. Entre ellos destacan Grecia o España, entre los que destaca el programa de mejora

genética del ciruelo japonés llevado a cabo en Murcia, en coordinación con el centro de edafología y biología aplicada del Segura (CEBAS-CSIC) y el Instituto Murciano de Investigación y Desarrollo Agrario y Medioambiental IMIDA en 2021 [112]. En cuanto a la edad del árbol es conveniente resaltar que, por norma general, en los árboles frutales, los árboles jóvenes presentan una débil producción, mientras que los árboles más antiguos tienen problemas de podredumbres debidas al desarrollo de chancros sobre las ramas.

Factores edafoclimáticos

Según Gascón et al, [113] el ciruelo es un cultivo que se adapta bien a distintas exposiciones, soportando bajas temperaturas, siempre que no sean extremas, siendo las variedades japonesas las más exigentes en este ámbito, ya que son más sensibles a las heladas tardías de primavera, razón por la cual hay años en los que la producción se reduce drásticamente. Por otro lado, las variedades europeas son más resistentes y requieren de una buena parada invernal. En cuanto a la floración, esta ocurre durante el mes de marzo, siendo quince días antes en las variedades japonesas.

En cuanto al suelo, este cultivo vegeta y prospera en una gran diversidad de suelos, ya que resiste el alto contenido en caliza sin que aparezca clorosis, como sucede con otros frutales de hueso. El suelo que más le conviene a este tipo de plantación es el arenoso-limoso, vegetando de manera óptima en suelos profundos. Para cultivarlo en suelos secos el patrón que suele utilizarse es el del almendro, mientras que el más resistente a la humedad es el *Mirabolano*.

En cuanto a cómo influye el clima y el suelo en la calidad del fruto, se pueden destacar estudios como los llevados a cabo por Vangdal E. et al. [114] donde observaron que los parámetros de calidad de ciruelas cultivadas en Noruega, estudiados con un clima templado y seco en Abril y

Junio y un clima fresco en Mayo, denotaban una mayor calidad con este tipo de clima en los meses estudiados durante 19 años distintos. Por otro lado, con respecto al tipo de suelo, un estudio reciente publicado por Oltenacu V. y Petrisor C. en 2021 [115], con ciruelas de las variedades '*Vânat Românesc*' y '*Stanley*', cultivadas en tres tipos distintos de suelo, puso de manifiesto que el tipo de suelo afectaba de manera significativa al contenido en antocianinas, al calibre de la fruta y al contenido en minerales, concluyendo que existen suelos con distintas características que se adaptan mejor a los requerimientos de calidad de las variedades estudiadas. Se concluyó que el suelo marrón es más adecuado para la variedad '*Vânat Românesc*' y el suelo arcilloso cubierto más adecuado para la variedad '*Stanley*'.

Prácticas de cultivo

Dentro de las prácticas de cultivo para el ciruelo se puede destacar el laboreo, el cual se suele recomendar como una labor fundamental en la parada invernal que perjudique lo menos posible al árbol y que se realiza con el fin de eliminar malas hierbas o suprimir brotes indeseados en cultivos injertados, además de las dos a cuatro labores superficiales en el transcurso de la vegetación. El uso de herbicidas también se recomienda en plantaciones en producción. Además de otros factores que afectan más concretamente a la calidad como son la fertilización, el riego y la poda que se comentan a continuación.

Fertilización

Las exigencias del ciruelo en cuanto a la fertilización son similares a los frutales de pepita, teniendo en cuenta un análisis detallado de las características del suelo que puedan aportar las necesidades establecidas por la planta. Los nutrientes más utilizados, dependiendo del estado de producción del árbol se resumen en la **Tabla 7**.

Tabla 7. *Requerimientos de fertilización del ciruelo según su etapa de producción [113].*

Fertilizantes	Kilogramos por hectárea		
	<i>Formación</i>	<i>Producción Media</i>	<i>Plena Producción</i>
Nitrógeno	30 a 60	240 a 480	450 a 900
Fósforo	80 a 160	160 a 320	300 a 600
Potasio	50 a 100	125 a 250	250 a 500

En lo referente a cómo el uso de fertilizantes afecta a la calidad de los frutos, existen diversos estudios para el caso de los principales nutrientes como son el nitrógeno, cuya aplicación excesiva puede causar impactos negativos en la firmeza y en el contenido en sólidos solubles en frutas de hueso [116]. Otros estudios indican que un exceso de nitrógeno provoca una disminución en el contenido de antocianinas [117], e incrementa el riesgo de enfermedades postcosecha [118]. El déficit de nitrógeno conlleva un descenso en el rendimiento y en el tamaño de la fruta, así como a un deterioro en las características sensoriales [119]. Por otro lado, el déficit de potasio afecta a la productividad de la fruta de hueso y la calidad y su vida útil se ven también perjudicadas [120, 121]. Un exceso de potasio aumenta la acidez de la fruta [122]. Finalmente, la fertilización con fósforo ha puesto de manifiesto que aumenta la productividad y el tamaño del fruto. No obstante, el exceso de fósforo produce un descenso en la calidad en términos de contenido en sólidos solubles [123].

Riego

En cuanto al riego, en la mayoría de los casos, el cultivo del ciruelo se viene implantando en suelos frescos y con un buen índice de retención de humedad; por esta circunstancia, es un cultivo que puede darse en buenas

condiciones de secano. No obstante, su respuesta en regadío es muy favorable, aumentando el rendimiento de las producciones.

Un claro ejemplo de cómo afectan las condiciones de riego a la calidad del fruto se demuestra en el estudio realizado por Hassan F et al. [124], en el cuál realizaron un experimento de campo en un cultivo de ciruelas '*African Rose*' en Sudáfrica, en la campaña 2019-2020, sometiendo a los árboles de la plantación a diferentes regímenes de riego, con un nivel de evapotranspiración del 100% para el control y del 80 y 60% para las muestras experimentales. Se observaron mejoras en los niveles de ácido abscísico, fenoles totales, antocianinas y sólidos solubles con los regímenes deficitarios, pero un peor rendimiento, acidez, tamaño y firmeza.

Poda

Con respecto a la poda, las variedades, tanto europeas como japonesas presentan dos formas diferentes de vegetar. Unas son propensas a emitir brotes anticipados, y otras presentan una tendencia a prolongarse en la parte terminal, por lo que la práctica de poda dependerá de la plantación, siendo la formación más empleada en el ciruelo el vaso. Además, para el caso del ciruelo japonés, se suele realizar el aclareo de frutos dada su tendencia a producir excesivamente, lo cual al final delimita la calidad del producto.

En términos de calidad, además del aclareo, la irradiación también juega un papel fundamental en el desarrollo de las características de calidad del fruto, ya que induce a la maduración de este como en todos los árboles frutales. Como consecuencia de la irradiación recibida, los frutos que maduran primero son los de la parte superior de la copa del árbol y los últimos los de la parte inferior, lo que puede provocar un desfase de maduración de 10 a 14 días. Murray X, et al. [125] estudiaron que en ciruelas japonesas de las variedades '*Laetita*' y '*Songold*', las diferencias

de maduración entre frutas sombreadas bajo determinadas condiciones de iluminación previas a la cosecha ofrecían parámetros de maduración muy inferiores a las establecidas por una iluminación completa usada como control.

B. Factores postcosecha

Posteriormente a la fase de producción, es preciso aplicar una serie de tecnologías a la conservación del material vegetal en cualquier tipo de producción, de cara a cuidar y mantener dicho material vegetal en las mejores condiciones posibles hasta su comercialización. Entre estas tecnologías, concretamente para la conservación de las ciruelas, se pueden destacar el método de cosecha, las temperaturas de almacenamiento, el uso de atmósferas modificadas para el envasado, el uso de inhibidores del etileno para retrasar la senescencia del fruto como es el 1-MCP, o los tratamientos térmicos. También tiene especial relevancia cómo estas tecnologías se encargan de prevenir defectos y patologías fisiológicas que surgen en el fruto previos a su cosecha o posteriores a esta.

Estado de maduración en el momento de la cosecha

Uno de los factores críticos por el cuál la calidad y la vida útil de las ciruelas se ven fuertemente condicionados y cuya caracterización es uno de los objetivos fundamentales de la presente memoria es el estado de maduración, a través del cual se selecciona el momento óptimo para la cosecha. Normalmente, la fecha para la cosecha del fruto se establece a través de análisis como son: el color de la piel y la firmeza de la pulpa. Este factor, junto al tiempo de almacenamiento resulta de especial relevancia de cara a la salida comercial del producto, por lo que existen estudios como el realizado por Guerra M. et al. [126], en el que evaluaron la vida útil, los cambios en la calidad y la aceptación del consumidor en ciruelas europeas de la variedad *Green Gage*. Observaron que tanto la fecha de cosecha, como

el tiempo de almacenamiento afectaban de manera significativa a los parámetros de calidad y sensoriales del producto, por lo que la vida útil era de 20 días en cosechas tempranas pero que estas tenían menor aceptación por parte del consumidor debido al menor ratio entre sólidos solubles y acidez titulable. También constataron que existía una correlación entre la aceptación por el consumidor y el parámetro a^* del espacio de color CIE $L^*a^*b^*$, siendo este una medida aproximada y no destructiva para la evaluación del periodo óptimo de cosecha.

Con respecto a la recolección, las ciruelas suelen ser recogidas a mano, colocadas en cajas y, posteriormente, en contenedores para el transporte hacia la envasadora. El tiempo de permanencia en el campo debe ser lo más corto posible, ya que el mantener la fruta en el campo por largos periodos de tiempo puede provocar un ablandamiento indeseado de la pulpa, lo que hace necesario el desarrollo de una serie de tecnologías de preenfriamiento, como el enfriamiento por aire forzado o el hidrogenfriamiento. Crisoto C. y Kader A. [127] analizaron, en ciruelas japonesas de la variedad *Friar*, el efecto que el estado de maduración (temprano, medio o tardío) tenía sobre las condiciones de almacenamiento a diferentes temperaturas y condiciones de humedad relativa, en términos de análisis de los defectos fisiológicos que aparecen como consecuencia de dicho almacenamiento, observando que la producción de etileno se retrasaba para temperaturas de 0, 2 y 25 °C, mientras que a 5 y 15°C se aceleraba, especialmente para ciruelas recolectadas en un estado de madurez avanzado.

Se puede concluir que el estado de maduración es un parámetro crítico, que hay que controlar de cara a un posterior acondicionamiento y almacenamiento del producto, para mantener sus propiedades asociadas a la calidad y su vida útil en las mejores condiciones para su comercialización.

Temperatura de almacenamiento

Después del almacenamiento de la fruta es preciso una correcta refrigeración, en condiciones de humedad relativa controlada, para extender lo máximo posible la vida útil del fruto. Sin embargo, hay que tener en cuenta que las ciruelas son frutas sensibles al '*daño por frío*', el cual causa lesiones en sus tejidos, como pueden ser la translucidez de la pulpa a consecuencia de la descomposición del tejido, pardeamiento y 'sangrado', síntomas que han sido ampliamente estudiados por multitud de autores y que concluyen que, pese a las diferencias debido a la variedad de la ciruela, las condiciones de refrigeración óptimas para reducir la producción de etileno y aumentar la vida útil del producto son 0°C y una humedad relativa del 95% [128].

Uso de atmósferas controladas y atmósferas modificadas

El uso de atmósferas controladas y atmósferas modificadas ha mostrado tener una efectividad significativa en el mantenimiento de la calidad y la vida útil postcosecha de multitud de frutas según Cantín. C.M. et al. [129], que estudiaron la influencia del envasado, en diferentes líneas de atmósfera modificada, sobre los parámetros de calidad de ciruelas de la variedad '*Friar*'. Los resultados encontrados fueron la conservación del color, el mantenimiento del peso respecto al control, el mantenimiento de la vida útil durante 45 días y no se observaron cambios en el resto de los parámetros de calidad, como la acidez titulable o el contenido en sólidos solubles. Resultados similares han sido obtenidos por Crisoto C. et al.

[127], que concluyeron que la aplicación de atmósferas controladas con 1-2% de O₂ y 3-5% de CO₂ conseguían mantener la firmeza y reducir los cambios de color.

Tratamiento con inhibidores de etileno 1-MCP

El 1-metilciclopropeno (1-MCP) es un potente inhibidor de la acción del etileno que retrasa los procesos senescentes y algunos desordenes fisiológicos, prolongando la vida útil de los vegetales y facilitando su transporte y comercialización. Fue propuesto por primera vez en 1994 por investigadores de la universidad de California y Carolina del Norte [130], y desde entonces se ha investigado su uso en un gran número de especies vegetales, con objeto de evaluar su acción en dichos efectos de prevención de desórdenes fisiológicos como efecto de la maduración y la senescencia.

Con respecto a su uso en las ciruelas se han ensayado aplicaciones del 1-MCP en un rango de concentración de 0,05 a 1 µL/L [131, 132], obteniéndose resultados satisfactorios en el retraso en la aparición de ablandamientos en el fruto, así como en la reducción de las enzimas encargadas de degradar la pared celular como son la poligalacturonasa, galactosidasa y endo-glucanasa/glucosidasa. Es importante remarcar que los efectos del 1-MCP se ven afectados por factores como el periodo de almacenamiento y también por el estado de maduración del fruto en el momento de su cosecha, observándose que dichos efectos se reducen con estados de maduración más avanzados. Además, el efecto del 1-MCP no se ve afectado con la temperatura y no sufre una reducción significativa a temperaturas superiores a las de refrigeración, evitando así el daño por frío [133, 134].

En estudios más recientes, como el realizado por Pan. H. et al. [135], emplearon métodos alternativos fundamentados en el uso de etileno y 1-MCP para el control del estado de maduración posterior al almacenamiento

en ciruelas japonesas de las variedades *Galaxian*, *Yuhuang* y *Aozhou 14*, tratando estas con tratamientos alternativos de 500 $\mu\text{L/L}$ de etileno o 2 $\mu\text{L/L}$ de 1-MCP durante 24 h a 25°C tras la refrigeración a 0°C y durante 8, 4 y 2 semanas, respectivamente, mostrando cada variedad diferentes patrones fisiológicos. En dicho estudio se observó además que el tratamiento con etileno incrementaba el cambio de color, mientras que el tratamiento con 1-MCP permitía mantener la firmeza y prevenía el enrojecimiento de la pulpa durante la vida útil normal, lo que mejoraba la aceptación comercial del producto.

Como se puede constatar, el tratamiento con inhibidores del etileno afecta de manera dispar a las ciruelas según el tipo de variedad y es otro factor a tener en cuenta de cara al mantenimiento de la vida útil postcosecha y a la selección óptima del estado de maduración para su aplicación.

Tratamientos Térmicos

Los tratamientos térmicos postcosecha han resultado ser efectivos en la prolongación de la vida comercial de estos productos, entre los que destaca la revisión de estos realizada por Lurie. S. [136]. En el caso de las ciruelas, Serrano M. et al. [138] desarrollaron un tratamiento térmico sobre ciruelas japonesas de la variedad *Blackstar*, que consistía en sumergir las ciruelas en agua a 45°C durante 10 minutos, con lo que reportaron una reducción de los defectos fisiológicos inducidos por el daño mecánico y una reducción también en los defectos inducidos por la producción de etileno. Estas mejoras en la prevención de defectos fisiológicos se han asociado con un incremento de la espermidina en la pared celular que induce una mayor estabilidad en esta y, por ende, en la firmeza de la pulpa, como fue monitorizada por Valero D. et al. [137]. Con respecto a este tipo de tratamientos, se ha estudiado también que la exposición a tales temperaturas en condiciones de precosecha puede provocar efectos similares en la

protección de la pared celular [138], por lo que resulta ser un factor interesante de cara a la mejora de las propiedades con fines de salida comercial del producto, tanto en la fase de producción como en los tratamientos postcosecha.

Otros tratamientos de interés

Además de los ya citados, se han investigado otras estrategias de tratamientos postcosecha a escala de laboratorio, cuyos resultados han arrojado una potencial utilidad a nivel industrial, como pueden ser el uso de ozono como agente oxidante, que previene el crecimiento microbiano [139], el tratamiento con poliaminas y calcio para la reducción del daño mecánico [137], [140], [141] recubrimientos con agentes comestibles como el lípido de hidroxipropilmetilcelulosa para la reducción del ablandamiento de la pulpa [142] o la fumigación con vapor de timol para reducir el crecimiento del hongo *Molinitina frutícola* [143], [144].

Todos los factores citados que afectan a la composición y a los parámetros de calidad estipulados para la comercialización de las ciruelas en sus múltiples variedades hacen que su análisis dependa de múltiples factores y, dado el gran interés económico y social que conlleva la producción y el consumo de estos frutos, es necesario el desarrollo de herramientas analíticas simples y accesibles a todos los sectores interesados.

A continuación, se exponen las técnicas, principios y teorías sobre las cuáles se han desarrollado las metodologías objeto de estudio en esta tesis doctoral.

2. TÉCNICAS ANALÍTICAS EMPLEADAS

En lo referente a las técnicas analíticas utilizadas para la caracterización y detección de fraudes alimentarios en las matrices alimentarias previamente indicadas en este apartado, se describirá el uso y aplicaciones de las técnicas analíticas empleadas en esta memoria. Estas técnicas han sido la fluorescencia molecular como técnica no destructiva, tanto en modo convencional como en modo *front-face* y acoplada a una fibra óptica para la detección de compuestos como tocoferoles, polifenoles y clorofilas, y la cromatografía líquida de alta resolución para el análisis cuantitativo de clorofilas en ciruelas. Así mismo, se han obtenido y procesado imágenes digitales, cuyo fundamento teórico se explicará en el **apartado 4** de esta introducción, y se ha realizado su procesamiento en el espacio de color RGB para la caracterización del estado de maduración en ciruelas. En todos los casos, para el análisis de la información obtenida con las distintas técnicas analíticas, se han utilizado distintos métodos quimiométricos.

2.1. Fluorescencia molecular

La fluorescencia es una propiedad intrínseca a los analitos que han sido objeto de estudio en las matrices alimentarias analizadas en esta tesis doctoral (aceites de oliva y ciruelas), de las cuales existe una extensa bibliografía sobre el desarrollo de técnicas fluorescentes a partir de las cuales se han desarrollado aplicaciones concretas.

La fluorescencia molecular es una técnica adecuada para desarrollar metodologías en combinación con análisis multivariante, ya que tiene un carácter inherentemente multidimensional. Las propiedades fluorescentes de cada compuesto se caracterizan por dos tipos de espectros, los espectros de excitación, que son aquellos que se obtienen debido a la absorción de radiación necesaria para proporcionar la energía entre el estado

fundamental y el estado excitado del fluoróforo, y el espectro de emisión, que es el resultado del proceso de desactivación radiante entre el estado excitado y el estado fundamental. Este carácter multidimensional de la fluorescencia molecular se emplea en el desarrollo de metodologías de análisis multivariante utilizando los espectros sincrónicos o las matrices de excitación-emisión, las cuales se describen a continuación:

- 1. Espectros sincrónicos:** esta modalidad fue propuesta por Lloyd et al. en 1971 [145] y consiste en registrar de manera simultánea, manteniendo constante la velocidad de ambos monocromadores, las longitudes de onda de excitación y emisión con la misma diferencia entre ellas, $\Delta\lambda = \lambda_{\text{ex}} - \lambda_{\text{em}}$. Esta técnica es muy útil en el análisis de mezclas de compuestos fluorescentes, ya que tanto las longitudes de onda de excitación como de emisión están incluidas en el mismo espectro.
- 2. Matrices de excitación-emisión (EEMs):** Esta técnica fue introducida inicialmente por Weber en 1961 [146]. Consiste en la obtención de una representación tridimensional de la intensidad de fluorescencia representada en función de las longitudes de onda de excitación y emisión, obteniendo de este modo una descripción cualitativa de la muestra a analizar. Pese a que se requiere una gran capacidad de procesamiento, almacenamiento y tiempo de análisis, dichas matrices proveen de una importante información, cuyo tratamiento matemático permite la identificación de los componentes objeto de estudio, su cuantificación y establecer métodos de clasificación de muestras en base a dichos componentes fluorescentes.

Además de la posibilidad de un registro multidimensional de la fluorescencia molecular utilizando los sistemas convencionales, actualmente existen otras posibilidades de trabajo, dado el desarrollo de

fibras ópticas o la posibilidad de utilizar la modalidad de fluorescencia “*front-face*”, cuya principal diferencia con respecto a la modalidad tradicional, es la modificación del ángulo de incidencia de la radiación sobre la muestra, pasando de un ángulo de 90° en la técnica tradicional a un ángulo próximo a 30° . Utilizando estas técnicas se facilita la obtención de información de muestras sólidas o muestras muy turbias, sin necesidad de pretratamientos previos ni aislamiento de los analitos de interés. Esto tiene la ventaja de no alterar la distribución de los componentes en la muestra y poder analizarlos en su estado natural. Estos dispositivos se muestran en la **Figura 12**.



Figura 12. Dispositivos para el uso de fibra óptica (A) y fluorescencia ‘*front-face*’ (B).

La modalidad de fluorescencia “*front-face*” permite realizar el análisis en muestras sin previa dilución, ahorrando así tiempo de análisis, ya que la geometría de 30° o 60° grados respecto al haz de radiación incidente hace que el camino óptico de este a lo largo de la muestra se reduzca considerablemente, reduciendo a su vez efectos indeseables debido a la concentración, como puede ser el efecto de filtro interno. Esta modificación de la disposición con respecto a la fluorescencia de ángulo recto ha sido empleada ampliamente en muestras homogéneas, como puede ser el análisis de aceite de oliva [147]. Por otro lado, el uso de fibras ópticas acopladas a los equipos de fluorescencia molecular ha tenido un profundo desarrollo en sus características de manejabilidad, versatilidad, precio y

miniaturización, lo que ha permitido desarrollar múltiples aplicaciones, sobre todo en el campo de la biomedicina [148], dado el carácter mínimamente invasivo de esta técnica. También ha sido utilizada en el campo del análisis de alimentos mínimamente procesados, como pueden ser los frutos que contengan clorofilas, ya que su uso también ha sido reportado para este fin [149].

Las características multidimensionales de la fluorescencia molecular y el desarrollo de métodos de análisis combinados con herramientas matemáticas han convertido a la fluorescencia molecular en una técnica de gran potencial para la implementación de metodologías analíticas rápidas, selectivas y sostenibles, con un mínimo procesamiento de muestra. Esto facilita y amplía su uso para el control de calidad en alimentos.

2.2. Técnicas Cromatográficas

Las técnicas cromatográficas han sido empleadas en esta tesis doctoral como métodos de referencia para la cuantificación de analitos de interés como pueden ser clorofilas, feofitinas, carotenoides o antocianinas, etc. Para explorar las diferentes posibilidades, se ha llevado a cabo una revisión bibliográfica de la que se han extraído las siguientes conclusiones:

- Los principales analitos referenciados en la bibliografía son los compuestos fenólicos, entre los que destacan los flavonoides y, más en concreto, las antocianinas, siendo el análisis de otros pigmentos como las clorofilas o los carotenoides meramente testimonial.
- Las principales técnicas utilizadas en este ámbito son la cromatografía líquida de alta eficacia (HPLC), con detectores fotométricos de serie de diodos (DAD), utilizando longitudes de onda de absorción de 520 nm para el análisis de antocianinas y 280 nm para el análisis de polifenoles.

- La separación de polifenoles, flavonoides y antocianinas se suele hacer en columnas C₁₈ en fase reversa con un tamaño de partícula de 5 µm, con una fase acuosa compuesta por una parte de ácido fórmico o fosfórico diluido y otra fase de metanol o acetonitrilo en modo gradiente.
- La bibliografía referente al uso de estas técnicas para la determinación de clorofilas en aspectos relacionados con la calidad de las ciruelas es escasa, habiendo referencias actualizadas con respecto a otras técnicas como la fluorescencia molecular, la cual se desarrollará en el siguiente apartado, o la absorción molecular [150], aunque hay que reseñar que existen otros estudios donde se desarrollan métodos cromatográficos de análisis de pigmentos en este tipo de matrices, como los realizados por Sousa C. [151].

La descripción pormenorizada del método cromatográfico empleado para el establecimiento del método de referencia desarrollado se especifica en el **Capítulo 2** de esta memoria.

3. DESCRIPCIÓN DE LOS MÉTODOS QUIMIOMÉTRICOS EMPLEADOS

Como se ha descrito a lo largo de esta introducción, el tratamiento de la información analítica a través de diferentes métodos matemáticos se ha vuelto una constante en las investigaciones en el campo de la química analítica en estos últimos años. Esto es debido al enorme potencial que las matemáticas tienen para poder resolver problemas analíticos complejos y al incremento exponencial que se ha venido dando en la capacidad de registro y tratamiento de la información analítica por parte de los métodos instrumentales actuales, así como el desarrollo conjunto de la capacidad de procesamiento de datos en estos.

El uso de métodos matemáticos con el objeto de extraer e interpretar la información procedente de las señales analíticas de un sistema químico complejo para resolver problemas de diversa tipología se conoce con el nombre de quimiometría, término acuñado en 1974 por Svante Wold [152]. Una de las principales ventajas que tiene el uso de la quimiometría en el tratamiento de la información de sistemas químicos complejos es su potencial en la resolución de problemas en mezclas de compuestos químicos. Estas herramientas tienen la capacidad de cuantificar o caracterizar los componentes de una mezcla a bajas concentraciones o en presencia de otras sustancias interferentes, sin necesidad de realizar separaciones previas, lo cual supone un ahorro en instrumental, tiempo y complejidad de los métodos analíticos.

Los métodos quimiométricos se fundamentan en la modelización de un número mayor de datos que en los métodos tradicionales en los que, comúnmente, se hacían medidas puntuales de señales químicas con respecto a una variable como puede ser la longitud de onda (univariantes), pasando a medidas dependientes de dos o más variables, como pueden ser

la utilización de los espectros de excitación y emisión de fluorescencia (multivariantes), aumentando así la selectividad de las señales de respuesta instrumentales en datos multivariantes.

En función del número de variables utilizadas para relacionarlas con el contenido de los analitos a determinar en una muestra, los datos pueden ser de orden cero, si sólo se relaciona el analito con una variable; de orden 1, cuando se proporciona un vector de datos, como puede ser un espectro; de orden 2, cuando las variables a relacionar con los analitos de interés son dos, como es el caso de una matriz de excitación-emisión (EEM, del inglés *excitation-emission matrix*) de fluorescencia, o de orden superior n , obteniendo cubos o hipercubos de datos, cuya modelización a su vez produce determinadas ventajas como el aumento de la sensibilidad, o lo que se conoce como “ventaja de segundo orden”, que consiste en poder cuantificar muestras en presencia de componentes interferentes que no entren en la calibración. La **Figura 13** muestra de manera gráfica la separación de este tipo de métodos en función de la complejidad de los datos procesados.

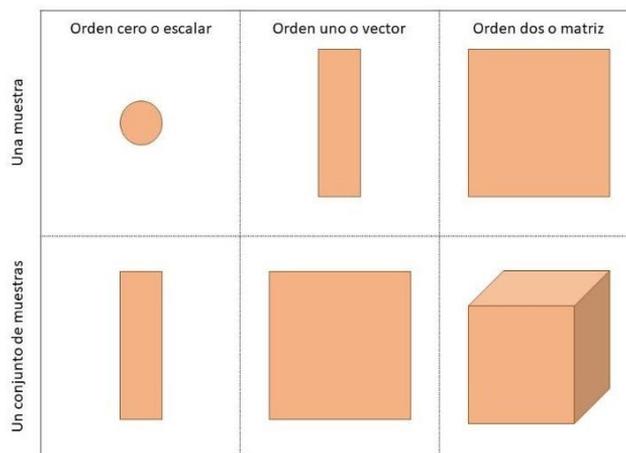


Figura 13. Tipos de conjuntos de datos multivariantes en función del número de variables [153].

Para la construcción de un modelo de calibración multivariante, que permita obtener e interpretar la información química relevante de la muestra objeto de estudio con fines de clasificación o de cuantificación, es preciso realizar previamente una serie de etapas comunes a todos los modelos matemáticos desarrollados:

- 1) *Definir un conjunto de calibración, también llamado de entrenamiento*, con las características representativas de los analitos y de sus interferencias, que permita establecer predicciones sobre un conjunto de muestras desconocidas semejantes a la del conjunto de calibración.
- 2) *Registrar la información analítica objeto de estudio.*
- 3) *Realizar un pretratamiento de los datos* en caso de ser necesario para minimizar las contribuciones no deseadas presentes en la señal analítica.
- 4) *Seleccionar el modelo óptimo* para la extracción de la información deseada, en función de las características de la muestra y el tipo de información que se quiera obtener. Posteriormente, se detallarán los modelos empleados en esta memoria y los algoritmos en los que se fundamentan, así como sus ventajas y desventajas de cara a los objetivos establecidos para la caracterización de las muestras estudiadas.
- 5) *Validar el modelo* en caso de ser necesario a través de un conjunto de muestras de validación obtenidas de manera similar al de calibración, pero con una concentración o categoría (en caso de utilizarse para fines de clasificación) conocidas, que nos permita saber si el modelo de calibración (o clasificación) describe correctamente el sistema.

- 6) *Establecer predicciones sobre las muestras* calculando la propiedad estudiada en nuevas muestras a las que previamente se les registra la señal analítica.

Finalmente, antes de abordar pormenorizadamente los algoritmos empleados en esta memoria, es conveniente hacer una breve revisión sobre las herramientas estadísticas descriptivas utilizadas para evaluar los errores y la capacidad predictiva de los modelos propuestos. Para ello, a continuación, se muestran los parámetros de interés.

Un parámetro frecuentemente usado es la suma de los cuadrados de los errores de predicción PRESS (del inglés *Prediction Residual Error Sum of Squares*), el cual resulta de sumar los cuadrados de los residuales pronosticados, y se calcula del siguiente modo:

$$PRESS = \sum_{i=1}^{m_p} (\hat{y}_j - y_{ij})^2 \quad (1)$$

Siendo \hat{y}_{ij} la concentración calculada por el modelo para la muestra i y para el analito j , siendo y_{ij} las concentraciones reales de las i muestras para el analito j y m_p el número de muestras de predicción. Otra manera de estimar el error de predicción sería a través dividir el PRESS por el número de muestras para así obtener lo que se conoce como el error cuadrático medio de predicción, MSEP (Standard Error of Prediction).

$$MSEP = \frac{PRESS}{m} \quad (2)$$

Para asegurar la bondad de un modelo en la predicción de muestras distintas a las utilizadas en la calibración es necesario un proceso previo de validación, la cual puede ser interna o externa. Para la validación externa se suelen usar dos grupos de muestras, uno llamado conjunto de calibración y otro llamado conjunto de prueba o de validación, de modo que a la hora de aplicar el modelo sobre el conjunto de prueba se pueda comprobar cómo se

comporta el modelo frente a dicho conjunto de muestras. Por otro lado, cuando el número de muestras disponible es limitado, se suelen utilizar métodos de validación interna, siendo la metodología mayormente empleada la validación cruzada [154], la cual utiliza muestras del propio conjunto de calibración para comprobar el modelo. Este procedimiento consiste en segmentar el conjunto de muestras de calibración en varios bloques. Posteriormente, el modelo se construye reservando uno de los segmentos como conjunto de datos para comprobar los resultados y el resto para construir el modelo. El proceso se repite en función del número de segmentos elegidos y finaliza calculando el valor de MSEP (del inglés *medium square error of prediction*) para cada segmento y acumulándolo de forma que se consigue una buena estimación del poder predictivo de las muestras de calibración, así, el MSEP utilizado en la validación cruzada recibirá el nombre de MSEC.V.

Otros parámetros importantes a la hora de evaluar los resultados del modelo son la raíz del error cuadrático medio de predicción, del inglés *root mean square error of prediction*, (RMSEP), el error relativo de predicción (%REP) del inglés *relative error of prediction*, expresado como porcentaje que puede usarse para comparar modelos, o el propio cuadrado del coeficiente de correlación una vez establecido el modelo:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{m_p} (y_j - y_{ij})^2}{m_p}} \quad (3)$$

$$\%REP = \frac{100}{\bar{y}} \sqrt{\frac{\sum_{i=1}^{m_p} (y_{ij} - \bar{y})^2}{m_p}} \quad (4)$$

$$R^2 = \frac{\sum_{i=1}^{m_p} (y_j - \bar{y})^2}{\sum_{i=1}^{m_p} (y_{ij} - \bar{y})^2} \quad (5)$$

Siendo \hat{y}_{ij} la concentración calculada por el modelo para la muestra i y el analito j , y_{ij} las concentraciones reales de las i muestras para el analito j , \bar{y}_j el valor medio de concentraciones del analito j y siendo m_p el número de muestras de predicción.

Una vez establecidas las generalidades que caracterizan los métodos quimiométricos y su uso en matrices complejas para la determinación de analitos o para la caracterización de muestras mínimamente tratadas, a continuación, se hará una descripción pormenorizada de los modelos y algoritmos empleados.

3.1. Algoritmos de dos vías para datos de primer orden

En este apartado se comentarán los principales algoritmos empleados para el análisis de datos de primer orden, que en el caso de la presente tesis doctoral, han sido fundamentalmente los histogramas del espacio de color RGB tomados del procesamiento de las imágenes de las ciruelas analizadas. Las herramientas quimiométricas empleadas para el análisis de datos de primer orden han sido el análisis por componentes principales (PCA), con el objeto de realizar un análisis exploratorio de los datos obtenidos y la regresión por mínimos cuadrados parciales (PLS), para la cuantificación de clorofilas totales a través de métodos de referencia por cromatografía.

Debido al gran volumen de información que proporcionan los métodos instrumentales, uno de los campos de estudio de la quimiometría ha sido el desarrollo de métodos capaces de reducir este gran volumen de información, sin pérdidas relevantes, en un número reducido de variables. Este es el caso de PCA, del cual existe una extensa bibliografía en cuanto a su desarrollo y aplicación a casos relacionados con problemas químicos [155]–[158].

Para explicar el funcionamiento de PCA se puede considerar un conjunto de datos organizados matricialmente, \mathbf{X} ($\mathbf{m} \times \mathbf{n}$), como un conjunto de m puntos en un espacio \mathbf{n} -dimensional (longitudes de onda). El objetivo de PCA es encontrar las direcciones, ortogonales entre sí, en las cuales existe la máxima variabilidad espectral. De este modo, en lo que consiste PCA es en la reducción de un sistema de coordenadas \mathbf{n} -dimensional, a otro \mathbf{r} -dimensional, en el cual se establezcan unos nuevos ejes denominados componentes principales, que son perpendiculares entre sí, y que se han obtenido como combinación lineal de las \mathbf{n} variables originales [159]. En la **Figura 14** se puede observar gráficamente la reducción de la dimensionalidad que el método PCA propone para el conjunto de datos \mathbf{X} .

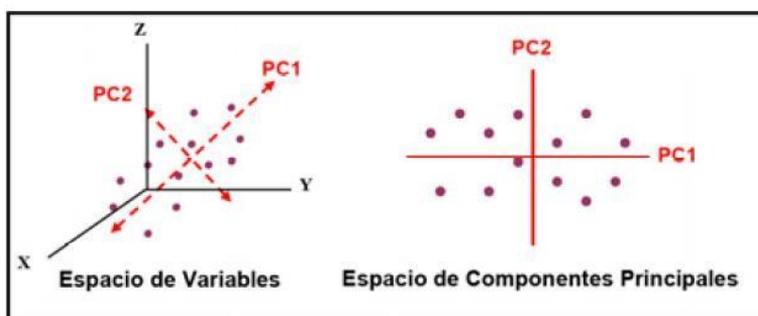


Figura 14. Representación gráfica de la reducción de la dimensionalidad llevada a cabo por PCA [159].

PCA se efectúa descomponiendo la matriz de datos iniciales \mathbf{X} en el producto de una matriz \mathbf{T} denominada “matriz de scores” y otra matriz \mathbf{P} denominada “matriz de loadings”, junto con la matriz de errores residuales \mathbf{E} de \mathbf{X} .

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (6)$$

La matriz \mathbf{T} recoge las coordenadas de las muestras originales en el nuevo sistema de ejes ortogonales. La matriz \mathbf{P}^T , recoge los cosenos

directores de los nuevos ejes respecto del sistema de ejes original. La dimensión r del nuevo subespacio viene determinada por el rango de la matriz \mathbf{X} .

Los diferentes componentes principales no contienen la misma información. Los primeros componentes describen las fuentes de variación más importantes de los datos, que se pueden asociar a la información más relevante. Para tener una reproducción perfecta de la matriz \mathbf{X} se necesitan k vectores de *loadings* y *scores*, sin embargo, es posible representar la matriz original con un número $a < k$ de vectores según la siguiente ecuación:

$$X = t_1 p_1^T + t_2 p_2^T + \dots + t_a p_a^T + E \quad (7)$$

ya que la información más relevante se encuentra en los a primeros componentes, lo que implica que la información del conjunto de datos \mathbf{X} , puede reducirse por un nuevo conjunto de variables llamadas componentes principales, las cuales son variables no correlacionadas entre sí en un nuevo sistema de ejes ortogonales.

La descomposición de esta matriz de datos de partida en su matriz de *scores* y *loadings* se puede realizar a partir de distintos algoritmos, siendo el más conocido y usado para este caso el algoritmo NIPALS (Nonlinear Iterative Partial Least-Squares) [160]. Independientemente del algoritmo empleado para la determinación de los componentes principales, el principal aspecto que hay que tener en cuenta inicialmente es la selección del número de óptimo de componentes que describen la matriz \mathbf{X} . Existen varios criterios para decidir el número de componentes, siendo los más empleados el porcentaje de varianza explicada y la validación cruzada.

- El **porcentaje de varianza explicada** o la fracción de varianza acumulada (s_e^2) se calcula a través de la relación de la suma de los

d valores propios respecto de la suma de todos los valores propios **p**, calculados a través de la expresión:

$$s_e^2 = \frac{\sum_{i=1}^{\infty} \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (8)$$

Si se representa la varianza explicada en función del número de componentes principales se suele escoger el número de componentes más bajo que presenta una variación significativa de la varianza.

- La **validación cruzada**, por el contrario, es un método iterativo en el cual se extrae el espectro de conjuntos de calibración y se calcula las matrices T y P con el resto del conjunto. Los espectros extraídos se predicen y se calcula el error residual utilizando, en cada iteración, un componente principal adicional. Al igual que en el caso de la varianza explicada, el número de componentes seleccionados es el mínimo que presenta una variación significativa en la varianza.

PCA permite hacer análisis exploratorios a través de herramientas de visualización, como es el análisis de conglomerados o *clusters* de los *scores* para cada muestra, haciendo una primera aproximación a posibles categorías que se han querido establecer, así como una caracterización de los datos espectrales de cada componente a través de sus *loadings*, todo ello asociado a los histogramas de color que se verán en el desarrollo de estudios realizados en el capítulo 2 de esta memoria dedicado al análisis de ciruelas.

Por otro lado, cuando se usan los datos de primer orden con fines de cuantificación, el método empleado ha sido el de la regresión por mínimos cuadrados parciales (PLS, del inglés, *partial least squares*), el cuál es

posiblemente el método de regresión multivariante más ampliamente utilizado [161], ya que presenta ventajas importantes como que permite trabajar con datos con alta colinealidad y proporciona una matriz de regresores robusta, ya que tiene en cuenta conjuntamente tanto la matriz de datos \mathbf{X} como la de calibración \mathbf{Y} , maximizando su covarianza. Este modelo fue desarrollado entre 1975 y 1982 por Wold et al. [162] y se basa en la obtención de variables latentes a través del siguiente proceso matemático.

En primer lugar, y antes de descomponer en factores las matrices de datos \mathbf{X} y de calibración \mathbf{Y} , se centran o autoescalan como en el caso del PCA. Posteriormente cada una de las matrices se descompone simultáneamente en una suma de a factores de forma que:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_{i=1}^a t_i \mathbf{p}_i^T + \mathbf{E} \quad (9)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} = \sum_{i=1}^a u_i \mathbf{q}_i^T + \mathbf{F} \quad (10)$$

Donde \mathbf{T} es la matriz de *scores*, \mathbf{P} es la matriz de los *loadings* y \mathbf{E} la matriz de residuales para la matriz de datos (matriz \mathbf{X}); \mathbf{U} es la matriz de *scores*, \mathbf{Q} la de *loadings* y \mathbf{F} la matriz de residuales para la matriz de datos (matriz \mathbf{Y}). Si tenemos m muestras, a factores, k variables y p analitos, la dimensionalidad de las matrices es la siguiente. \mathbf{T} y \mathbf{U} ($m \times a$), \mathbf{P}^T ($a \times k$) y \mathbf{Q}^T ($a \times p$). En este caso, los *loadings* no coinciden exactamente con la dirección de máxima variabilidad de las muestras, como en el caso de PCA, ya que están corregidos para obtener la máxima capacidad predictiva para la matriz \mathbf{Y} .

La descomposición de ambas matrices no es independiente, sino que se realiza de forma simultánea, estableciéndose una relación interna entre los *scores* de los bloques \mathbf{X} e \mathbf{Y} de forma que para cada componente a , se cumpla que:

$$\hat{y} = b_a t_a \quad (11)$$

Donde el símbolo $\hat{}$ indica que es un valor calculado y b_a es el coeficiente de regresión para cada uno de los componentes.

A partir de aquí, el valor de Y se calcula utilizando la relación interna \hat{u}_a

$$Y = BQ^T + F \quad (12)$$

Siendo B la matriz de regresores de dimensiones $(a \times a)$ y F la matriz de los residuales Y .

Una vez que el modelo de calibración ha sido establecido, la predicción de la propiedad modelada en un conjunto de muestras se puede obtener según la siguiente expresión:

$$y_i^T = b_0 + x_i^T B \quad (13)$$

Siendo x_i el vector que define la señal analítica de la muestra, y_i , el vector de concentraciones o propiedades a determinar y b_0^T , un vector de predicción de una muestra sin necesidad de descomponerla en matrices de *scores* y *loadings*.

En cuanto a la elección del número de factores para el modelo PLS, existen varios criterios, siendo el más popular el sugerido por Wold en 1978 [154], el cual consiste en representar el valor del error cuadrático medio de la validación cruzada (MSECV) frente al número de componentes y buscar el valor mínimo, ya que al aumentar los componentes el error disminuye hasta que llega a un valor en el que únicamente se modela el ruido y el MSECV vuelve a aumentar debido al sobreajuste del modelo.

Otro criterio es el descrito por Haaland y Thomas en 1988 [163], [164], en el cual, el número de componentes óptimos es aquel para el cual el valor del MSECV no es significativamente mayor al mínimo MSECV

del modelo, evitando de esta manera el sobreajuste. Este mínimo MSECv vendrá dado por un número de componentes al que denominamos a^* . Cada valor de MSECv obtenido con un número de componentes menor al a^* se compara con el valor de MSECv (a^*) mediante una prueba de Fisher, de este modo se pretende encontrar el modelo más pequeño cuyo MSECv no sea significativamente más grande que aquel modelo que proporciona el mínimo, así, para cada componente $a= 1, 2, \dots, a^*$, se calcula:

$$F(a) = \frac{MSECv(a)}{MSECv(a^*)} \quad (14)$$

Como número de componentes óptimo se escoge el menor a tal que $F(a) < F_{\alpha, m, m}$ donde $F_{\alpha, m, m}$ es el valor tabulado para una prueba F de una cola con un nivel de significación $(1-\alpha)$ y m grados de libertad, siendo 0,25 el valor de α recomendado de manera empírica por Haaland y Thomas.

Estos son los métodos empleados en esta memoria para el análisis exploratorio y para la regresión multivariante de los datos de primer orden obtenidos. No obstante, es preciso destacar que existen más métodos con fines clasificatorios como los árboles de decisión [165], el método de los K vecinos más próximos (K-NN)[166], el método de máquinas de vectores soporte, del inglés “*supporting vector machine*” (SVM)[167], las redes neuronales [168], etc. Así mismo, se dispone de otros métodos alternativos de regresión como pueden ser, mínimos cuadrados inversos, del inglés *inverse least squares* (ILS), regresión por componentes principales, del inglés *Principal Component Regression* (PCR) [169] o regresión lineal múltiple, del inglés *multiple linear regression* (MLR) [170].

Una vez detallado el fundamento teórico de los métodos de primer orden empleados, tanto para el análisis exploratorio como para la cuantificación de clorofilas, en el caso de las ciruelas, con fines de establecer patrones de seguimiento del estado de maduración de estas, es

preciso establecer una serie de apuntes sobre las herramientas gráficas y softwares empleados para la interpretación de los resultados obtenidos a partir de los datos adquiridos, referentes en este caso a los histogramas de color de las ciruelas.

- Para la realización de los cálculos de los datos de orden uno haciendo el análisis por componentes principales (PCA) en los histogramas obtenidos a través del software imageJ [171] y, posteriormente, la regresión multivariante de los histogramas con los datos cuantitativos de concentración de clorofilas obtenidos del método cromatográfico de referencia mediante PLS. Para ello, se utilizó el software The Unscrambler versión 6.11 (CAMO Software AS, Oslo, Noruega) [172].
- A la hora de realizar los análisis exploratorios en PCA y obtener los “clusters” de las muestras, analizadas en el espacio de los dos primeros componentes principales, se empleó una rutina programada en el entorno de Matlab, de representación gráfica de las muestras en el nuevo espacio configurado por los ejes obtenidos de los vectores directores de los componentes principales de PCA.

3.2. Algoritmos de tres vías para datos de segundo orden

Las EEMs obtenidas registrando espectros de emisión, a diferentes longitudes de onda de excitación, conforman una matriz de datos de segundo orden (muestras x longitud de onda de excitación x longitud de onda de emisión) y los algoritmos propuestos para analizarlos se denominan de tres vías. Estos datos de segundo orden cumplen hipotéticamente la condición de “bilinealidad” y, por tanto, la descomposición de los cubos de datos trilineales resulta única. Dependiendo del procedimiento utilizado, podemos clasificar los métodos de tres vías como (*N-way*) sí se obtienen tríadas de datos desdoblados en una matriz de *scores* y dos de *loadings* o

como métodos *unfolding*, si los cubos de datos se desdoblan en vectores de datos y se les aplica posteriormente algoritmos de primer orden.

En el caso que nos ocupa en este apartado, los métodos de tres vías empleados han sido PARAFAC y su combinación con otro método de análisis supervisado como es LDA.

Uno de los métodos *N-way* más utilizados es PARAFAC, propuesto de forma independiente por Harshman [173] y por Carroll y Chang [174] en 1970. Este modelo consiste en la descomposición de los datos del cubo X_{ijk} de columnas a_{if} , b_{jf} y c_{kf} , donde I son el número de muestras, J las longitudes de onda de emisión, K las longitudes de onda de excitación y F el número de componentes, en componentes trilineales, buscando el modelo que minimice la suma de cuadrados de los residuales e_{ijk} para el modelo:

$$X_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk} \quad (15)$$

El modelo PARAFAC se puede resolver aplicando el algoritmo *alternating least squares* (ALS), el cual consiste en establecer un ciclo iterativo para la matriz de datos, de tal modo que se repitan los siguientes pasos:

1. Optimización del número de componentes.
2. Inicialización de las matrices de *loadings* \mathbf{B} y \mathbf{C} , las cuales representan los espectros de los componentes modelizados a las longitudes de onda de excitación y emisión.
3. Estimación de la matriz de scores o coeficientes de proporcionalidad \mathbf{A} a partir de la matriz de datos \mathbf{X} , \mathbf{B} y \mathbf{C} por mínimos cuadrados.
4. Estimación de \mathbf{B} de la misma manera.

5. Estimación de C de la misma manera.
6. Se prosigue con el paso (3) hasta convergencia, es decir, hasta que la variación producida en iteraciones sucesivas entre los modelos no supera un criterio previamente fijado.

En el paso (3), para estimar el valor de la matriz A , la matriz de calibración X se desdobra en una matriz $IXJK$, mientras que la columna F de componentes, se descompone en la matriz Z de dimensiones $FxJK$. De modo que la estimación por mínimos cuadrados de la matriz A se hará del siguiente modo:

$$X = AZ \quad (16)$$

$$A = XZ^T(ZZ^T)^{-1} \quad (17)$$

Para estimar B y C se sigue el mismo procedimiento con la matriz A calculada previamente hasta converger a un mínimo global. Este mínimo global se alcanza por el uso de criterios de parada, tanto forzados, como puede ser la no negatividad de los *loadings*, en casos concretos en los que se haya utilizado un blanco corrector, como no forzados, como es el caso del cambio relativo en el ajuste entre dos iteraciones.

Por otro lado, es destacable también comentar previamente que, antes de aplicar el algoritmo e interpretar la información química presente en los datos de fluorescencia, hay que optimizar el número de factores, cuestión relevante que permite obtener además de la información sobre la aportación de los componentes a cada señal, el espectro de cada componente, por lo que se aporta una valiosa herramienta de cara a la identificación de analitos de interés a través del tratamiento matemático sobre las señales de fluorescencia sin tratamiento químico previo. Uno de los métodos para dicha optimización es el llamado *diagnóstico de la consistencia del modelo*

o “*core consistency diagnostic*”, CORCONDIA [175], cuya expresión es la siguiente:

$$Core\ consistency = 100 \left[\frac{1 - \sum_{d=1}^r \sum_{e=1}^r \sum_{f=1}^r (g_{def} - t_{def})^2}{\sum_{d=1}^F \sum_{e=1}^F \sum_{f=1}^F t_{def}^2} \right] \quad (18)$$

siendo **g** la superdiagonalidad y **t** la superidentidad.

Un modelado válido significa que **t** se aproxima a **g**, lo que a efectos prácticos supone que el número de componentes será el propicio cuando el valor del core consistency disminuya de manera abrupta debido a la sobredimensionalidad del modelo, lo que supondrá una reducción drástica de la superidentidad y un aumento de la correlación entre variables, por lo que el criterio de selección de componentes será dicho salto drástico, combinado también con la predicción establecida por los errores de regresión calculados también por el programa, los cuales coinciden en la gran mayoría de los casos con la predicción establecida para el “core consistency”. Cabe destacar que todos los cálculos han sido llevados a cabo a través del programa MatLab con la rutina MVC2 prediseñada por el Profesor Alejandro Olivieri del Instituto de Química de Rosario [176].

La comprensión del funcionamiento del modelo PARAFAC, y el modo de proceder de cara a la descomposición de la información analítica relevante con fines de caracterización en muestras complejas, hace necesaria a su vez la indagación en aspectos y características que justifiquen la idoneidad del modelo para llegar a los fines planteados en esta memoria. Por este motivo, se puede concluir que las características relevantes a destacar para el modelo PARAFAC son:

- Posee una solución única característica para un modelo trilineal, las cuales son los espectros representados en los *loadings* y los coeficientes de proporcionalidad representados en los *scores*, lo cual ofrece la

ventaja de obtener los espectros de los componentes fijados, siempre y cuando la relación señal ruido sea apropiada.

Esta ventaja nos permite la identificación de compuestos de interés, como pueden ser en el caso de los aceites la clorofila a, b o los diversos tocoferoles y compuestos fenólicos que se han mencionado previamente, que contribuyen a la fluorescencia de sus EEMs.

- Al no haber un orden intrínseco de los componentes, estos han de ser identificados después de la construcción del modelo. Además, al no haber una escala definida, también es necesario realizar un escalado con muestras de concentración conocida para poder cuantificar.

Finalmente, pese a que el método PARAFAC no está diseñado a priori con fines de clasificación, las diferencias en la composición química de las muestras a clasificar en los rangos de fluorescencia seleccionados permiten usar este algoritmo para dichos fines. Sólo hay que proceder a la representación tridimensional de los *scores* de la matriz **A** obtenida para los componentes que mayor contribuyen a la varianza explicada de estos.

PARAFAC también se puede utilizar con fines clasificatorios en combinación con el algoritmo de análisis lineal discriminante (LDA), el cuál fue propuesto por primera vez por Rao C. en 1948 para clasificar varias clases [177]. En comparación con PARAFAC, **LDA** es un método específico de clasificación supervisado fundamentado en el teorema de Bayes [178], el cual establece que la probabilidad de que una observación $\mathbf{X}=\mathbf{x}$ pertenezca a una categoría previamente definida k viene dada por la probabilidad a priori p_k de que dicha observación provenga de dicha categoría. Esto se asigna previamente con indicadores binarios y se obtiene la probabilidad dividiendo entre el número de muestras asignadas con dicho marcador, multiplicada por la función de densidad de probabilidad condicionada $f_k(\mathbf{x})$, la cuál es función de los scores obtenidos previamente

para los componentes determinados en cada muestra, y todo ello entre la probabilidad total, la cuál es la suma de la función de densidad por el número de observaciones, lo cual matemáticamente se expresa del siguiente modo:

$$P_k(x) = \frac{\pi_k f_k(x)}{\sum_{i=1}^K \pi_i f_i(x)} \quad (19)$$

Para que la clasificación basada en Bayes sea posible, se necesita conocer la probabilidad poblacional de que una observación cualquiera pertenezca a cada clase \mathbf{p}_k , y la probabilidad poblacional de que una observación que pertenece a la clase k adquiera el valor x en la función de densidad o predictor $f_k(x)$. En la práctica, raramente se dispone de esta información, por lo que los parámetros tienen que ser estimados a partir de la muestra. Como consecuencia, el clasificador *LDA* obtenido se aproxima al clasificador de Bayes.

La resolución del problema se reduce a la búsqueda de funciones discriminantes que minimicen la varianza dentro de un grupo previamente establecido, así como la maximización de la varianza entre grupos a través de la búsqueda de $z_i (i=1 \dots r)$ donde $r = \min(G-1, k)$ siendo G el número de poblaciones (en nuestro caso las 2 clasificaciones posibles) y k (el número de variables estudiadas (en nuestro caso los *scores* de los componentes obtenidos de PARAFAC), variables canónicas que se construyen como combinación lineal de los scores $z_i = \mathbf{a}_i \mathbf{x}$ de modo que

1. Se calcula la media de las poblaciones μ_g de las variables canónicas.
2. Se clasifica el punto en aquella población de cuya media se encuentre más próxima, con la distancia euclídea, en el espacio de variables canónicas:

$$(z_0 - z_i)'(z_0 - z_i) = \min(z_0 - z_g)'(z_0 - z_g) \quad (20)$$

A continuación, se trata el problema de la obtención de las variables canónicas, de modo que se busca un **vector u** tal que, cuando proyectamos los puntos sobre él, se obtiene la máxima variabilidad entre los grupos con respecto a variabilidad dentro de los grupos.

Para la obtención de las variables canónicas z que nos permitan distinguir entre poblaciones, el procedimiento llevado a cabo deriva en la obtención de las matrices **W** y **B** , ambas de dimensiones $k \times k$, siendo k el número de componentes de la muestra, las cuales representan la suma de los cuadrados dentro de grupos y la suma de cuadrados entre grupos, respectivamente, cuya representación matemática es la siguiente:

$$W = \sum_{j=1}^{n_g} \sum_{g=1}^G (x_{jg} - \bar{x}_g)(x_{jg} - \bar{x}_g)' \quad (21)$$

$$B = \sum n_g (\bar{x}_g - \bar{x})(\bar{x}_g - \bar{x})' \quad (22)$$

Siendo j los individuos de la muestra, y g los grupos. Las medias son obtenidas de los valores de los scores asignados tanto a cada grupo como al total. De este modo, se pretende obtener maximizar la siguiente función:

$$\Phi = \frac{u_1' B u_1}{u_1' W u_1} \quad (23)$$

Para ello se hace la derivada respecto del vector **u** e igualando a 0. Desarrollando esta derivada se llega a la expresión siguiente:

$$W^{-1} B u_1 = \Phi u_1 \quad (24)$$

Finalmente, mediante este procedimiento, asignando una matriz columna con dígitos binarios para particionar los subgrupos como explica Kemsley [179]:

$$B = (g - 1)^{-1}A^T Y(Y^T Y)^{-1}Y^T A \quad (25)$$

$$W = (i - g)^{-1}[A^T A - (g - 1)B] \quad (26)$$

Siendo A la matriz de scores de PARAFAC a la que aplicar el análisis lineal discriminante, podemos obtener las r variables canónicas aptas para la clasificación con la siguiente expresión:

$$z = U'x \quad (27)$$

El resultado obtenido resulta, como es de esperar en un método de clasificación supervisado, más eficiente a la hora de clasificar muestras con características de fluorescencia diferentes, ya que la maximización de la varianza produce una evidente separación en la atribución de las variables canónicas, cuya visualización en los *clusters* de datos ha sido contrastada para diversos tipos de clasificación.

Por otra parte, los métodos PLS de segundo orden como son *unfolded partial least squares* y *multiway partial least squares*, (U-PLS y N-PLS, respectivamente) también han sido empleados en esta memoria. Ambos han sido usados con fines de cuantificación en el caso de la determinación de clorofilas en ciruelas, a través de matrices de fluorescencia y análisis de imágenes digitales, como marcador de madurez, pero, en el caso de U-PLS, este también se ha empleado en combinación con el análisis discriminante (DA-UPLS), para poder discernir el grado de adulteración de aceites de oliva con otros aceites a partir de sus matrices de fluorescencia EEFs, por lo que, a continuación, se hará una breve descripción de ambos métodos aplicados a datos de segundo orden, para luego pormenorizar el caso del análisis discriminante en combinación con el método U-PLS.

Mínimos Cuadrados Parciales Multidimensionales (N-PLS)

El fundamento de este algoritmo es una extensión a señales de orden superior del de mínimos cuadrados parciales para datos de primer orden (PLS). Los datos se descomponen en un modelo similar al de PARAFAC para obtener triadas de datos. Concretamente $\underline{\mathbf{X}}$ se le llama al cubo de variables independientes de dimensiones $I \times J \times K$ y \mathbf{X} la matriz desdoblada $I \times JK$. $\underline{\mathbf{Y}}$ es el cubo de variables dependientes de dimensiones $I \times L \times M$ e \mathbf{Y} , en consecuencia, será la matriz desdoblada de dimensiones $I \times LM$, teniendo I, J, K, L y M el mismo significado que el dado en las dimensiones de PARAFAC. El modelo N-PLS puede escribirse de la siguiente manera:

$$\mathbf{X} = \mathbf{T}(\mathbf{W}^K | \otimes | \mathbf{W}^J)^T + \mathbf{E}_x \quad (28)$$

$$\mathbf{Y} = \mathbf{U}(\mathbf{Q}^M | \otimes | \mathbf{Q}^L)^T + \mathbf{E}_y \quad (29)$$

Donde \mathbf{T} y \mathbf{U} son las matrices de *scores*, \mathbf{W} y \mathbf{Q} las de *loadings* (el superíndice el orden considerado) y \mathbf{E}_x y \mathbf{E}_y los errores para las variables independientes y dependientes, respectivamente. El modelo de regresión se construye asumiendo la relación lineal interna entre las matrices de scores \mathbf{T} y \mathbf{U} .

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \mathbf{E}_u \quad (30)$$

Donde \mathbf{B} es la matriz diagonal de regresores.

Mínimos Cuadrados Parciales Desdoblados (U-PLS)

En este método, las matrices de datos de calibración inicialmente se vectorizan en vectores de dimensión $\mathbf{JK} \times \mathbf{I}$, para posteriormente, construir un modelo PLS convencional usando tanto los vectores de calibración como los de concentración y (de dimensiones $\mathbf{I} \times \mathbf{I}$) [180].

No obstante, el principal uso que se ha dado a la herramienta U-PLS en esta Tesis Doctoral ha sido, además del clásico uso en la cuantificación de analitos, el combinar con el análisis discriminante dicho método para la clasificación de muestras adulteradas de aceite de oliva. A modo explicativo y simplificado para datos de primer orden, se comentará el funcionamiento de dicho algoritmo con fines de clasificación según Bereton R.G. y Lloyd G.R. [181] .

El modelo PLS-DA implica establecer una regresión entre la matriz de datos X y la matriz de clasificación c tal y como se ilustra en la **Figura 15**. En dicho modelo c es un conjunto de muestras representada por números enteros (1 y 0). Las ecuaciones fundamentales de PLS-DA las podemos anotar del siguiente modo:

$$X = TP + E \quad (31)$$

$$c = Tq + f \quad (32)$$

Siendo T la matriz de *scores* P la matriz de los *loadings* y E y f , las matrices de los residuales.

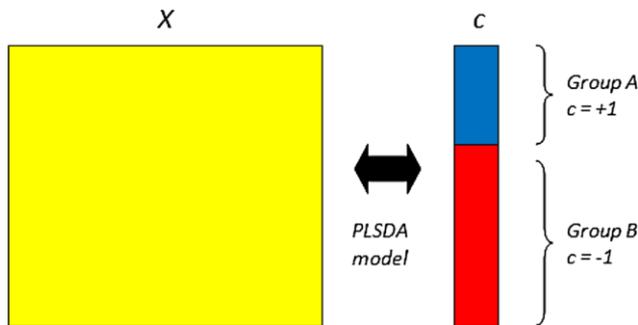


Figura 15. Modelo PLS-DA para la clasificación de dos grupos diferentes. Bereton R.G. y Lloyd G.R.[181].

Una vez construido el modelo, la predicción del valor de c para la clasificación puede hacerse del siguiente modo:

$$c = Xb + f = Tq + f \quad (33)$$

Donde b es un coeficiente de regresión, en este caso de primer orden un vector de dimensión $J \times I$ y, por lo tanto, una muestra de valor desconocido puede predecirse del siguiente modo:

$$\hat{c} = xb \quad (34)$$

La clase a la que pertenece la muestra problema queda determinada por el valor de \hat{c} . La regla de decisión en este caso será atribuirle el valor de 0 a una clase y el valor de 1 a la otra y ver en qué intervalo se sitúa la muestra seleccionada, con un intervalo de confianza establecido previamente.

Resolución de Curvas Multivariantes por Mínimos Cuadrados Alternantes (MCR-ALS)

El modelo de resolución de curvas para datos de primer orden se puede plasmar matemáticamente del siguiente modo según Tauler R. [182]:

$$D = XY^T + E \quad (35)$$

Donde D , de dimensiones $I \times J$, que corresponden al número de muestras y el número de variables, es la matriz de datos de señales analíticas X de dimensiones $I \times K$, mientras que Y^T , de dimensiones $K \times J$, son las matrices relacionadas con los perfiles de concentración y las respuestas unitarias a las especies químicas activas, respectivamente. La principal limitación de este modelo es que las soluciones de X e Y^T no presentan la cualidad de la unicidad, por lo que puede haber un infinito número de soluciones posibles. Por este motivo, el modelo MCR requiere de la aplicación de restricciones que reduzcan la ambigüedad rotacional de las matrices, y uno de los métodos más utilizados para aplicar MCR es el algoritmo de mínimos cuadrados alternantes (ALS), el cual desarrolla de

manera iterativa dos etapas de mínimos cuadrados, las cuales se pueden formular del siguiente modo:

$$\min_{Y^T_{restricciones}} ||D - XY^T|| \quad (36)$$

$$\min_{X_{restricciones}} ||D - XY^T|| \quad (37)$$

Siendo las soluciones sin restricciones a estas ecuaciones:

$$X = D(Y^T)^+ y Y^T = X^+D \quad (38)$$

Siendo $(Y^T)^+$ y X^+ las matrices pseudoinversas de Y^T y de X . En MCR-ALS, la resolución por mínimos cuadrados de las ecuaciones (36) y (37) se hace bajo restricciones como la no negatividad, la unimodalidad, la restricción de sistemas cerrados, o la restricción de igualdad. Posteriormente, se inicia ALS estimando inicialmente las matrices X e Y^T y el algoritmo resuelve las ecuaciones de forma iterativa, aplicando el criterio de mínimos cuadrados para ajustar la solución hasta adaptarse a las restricciones previamente marcadas.

Este modelo es extensible a datos de tres vías a través de la siguiente ecuación matricial:

$$D_K = XZ_KY^T + E_K \quad (39)$$

Esto significa que la descomposición de un cubo de datos \underline{D} en sus “submatrices” D_K compartirán exactamente las matrices X e Y^T y solo diferirán en los valores de la diagonal Z_k . Por lo tanto, habrá un número k de matrices diagonales Z_K que darán las cantidades relativas de los N componentes en cada matriz D_K . Con esta descomposición, posteriormente, se puede proceder al desarrollo del algoritmo ALS, teniendo en cuenta que cada columna del vector de la matriz X aumentada para cada componente articular se dobla en la dirección de la fila para dar el componente de la

matriz X^n , que posteriormente es descompuesto y truncado en sus primeros vectores singulares y, aplicando esta descomposición a las tres matrices del modelo, se aplica ALS hasta convergencia con las restricciones planteadas.

Los cálculos realizados para los diferentes algoritmos estudiados utilizados para los datos de segundo orden trabajados en la presente tesis doctoral, así como las representaciones gráficas correspondientes, fueron realizados gracias a las siguientes herramientas programadas dentro del entorno de Matlab:

- Las rutinas utilizadas empleadas para el uso de PARAFAC, se llevaron a cabo utilizando la rutina MVC2, una interfaz gráfica disponible en internet desarrollada por Olivieri A. et al. [176], [183, 184] que proporciona la representación de las matrices de fluorescencia antes de ejecutar las rutinas prediseñadas para la aplicación de PARAFAC, U-PLS, N-PLS y MCR-ALS.
- Para eliminar la distorsión de los datos producidos por la dispersión Rayleigh, en la región de los tocoferoles y compuestos fenólicos, se introdujo una rutina titulada *noray*, la cual permitía reemplazar las bandas producidas por dicha dispersión por valores interpolados a través del método de triangulación tridimensional Delaunay.
- Para aplicar el análisis discriminante lineal (LDA) [179] se empleó otra rutina previamente programada titulada *LDA_plot*, para la obtención de los vectores canónicos que, posteriormente, se empleaban para la discriminación de categorías previamente asignadas, así como la representación gráfica de los vectores canónicos en el espacio configurado por los componentes asignados al conjunto de muestras, para su clasificación en dos grupos diferentes.

- La representación gráfica de los *clusters* de los *scores* de PARAFAC y los vectores canónicos de LDA se efectuó a través de rutinas programadas manualmente tituladas *plot_2D* y *plot_3D*, dependiendo de si la representación de los “*clusters*” de datos de los “*scores*” o de los vectores canónicos se hacía en el espacio configurado por los dos o los tres componentes principales con mayor porcentaje de varianza explicada.
- Una vez establecidas las categorías en nuestras rutinas de interpretación gráfica como son *plot_2D* y *plot_3D*, para el establecimiento de un análisis confirmatorio a través de una interfaz gráfica de fácil interpretación, las rutinas *elipses_2D* y *elipses_3D* (dependiendo del espacio de dos o tres dimensiones configurado por los componentes que se quieren asignar) permite establecer un criterio objetivo y sencillo mediante la representación tridimensional de elipses en las tres dimensiones del espacio ocupado por los puntos asignados a cada categoría, en la cual, el contorno de dichas elipses se calcula a través del cálculo de una distancia media de todos los puntos hacia el contorno cuyo valor quede acotado por un valor de confianza extraído de un test estadístico *Chi-cuadrado* con un nivel de confianza del 95%. El intervalo de predicción para una distribución multivariante normal proporciona una elipse que consiste en vectores x que satisfacen la siguiente ecuación:

$$(x - \mu)^T \Sigma^{-1} (x - \mu) \leq \chi_k^2(p) \quad (40)$$

Donde μ es la media, Σ es la matriz de covarianzas y $\chi_k^2(p)$ es la función para la probabilidad p de la distribución Chi-cuadrado, con k grados de libertad, donde k es la dimensión de los datos. Los ejes se definen mediante los *eigenvectores* de la matriz de covarianzas y el radio de cada

eje es igual a 2,796 veces la raíz cuadrada del correspondiente eigenvalor. El valor 2,796 se obtiene de la raíz cuadrada de la distribución Chi-cuadrado con tres grados de libertad y un intervalo de confianza del 95% [185]. Introduciendo por orden las categorías dentro de la interfaz de la rutina prediseñada con estas premisas se obtiene la representación gráfica de dichas elipses, a través de las cuales se puede observar gráficamente que, si no existe solapamiento entre elipses, se puede afirmar con un 95% de confianza, que los grupos representados no pertenecen a la misma categoría.

- Para la realización del análisis discriminante por mínimos cuadrados parciales (DA-UPLS) se empleó la rutina prediseñada titulada ***DUPLS_plot***, a través de la cual se obtenían los valores predichos por el algoritmo para las dos categorías asignadas y la representación gráfica de los datos en las categorías asignadas separados por barras de error.

La separación se hacía utilizando el estadístico *t* de Student, con un nivel de significación del 95%, y $n-1$ grados de libertad, siendo n el número de muestras.

Si estos límites fijados no se solapan, se puede afirmar que las muestras clasificadas entre las distintas categorías son distintas y solo un 5% de veces, que se clasifican como distintos y no lo son, se debe al azar.

Todas estas herramientas, han sido puestas al servicio de los requerimientos previstos para alcanzar los objetivos propuestos en esta tesis doctoral, pudiendo constatar así el enorme potencial que presentan las herramientas de computación para multitud de campos de investigación aún por explorar y desarrollar, siendo esta tesis una parte muy concreta dentro del ámbito del control de la calidad de alimentos de origen vegetal, como son el caso de dos matrices alimentarias distintas como son los aceites de oliva y las ciruelas.

4. TEORÍA DEL ESPACIO DE COLOR

El desarrollo de metodologías fundamentadas en el procesamiento de imágenes digitales es una técnica relativamente moderna, cuyo carácter también multidimensional merece un desarrollo teórico previo de cara a la justificación de su uso en esta memoria. Por este motivo, en lo que sigue, se hará una descripción pormenorizada de la teoría del espacio de color y de su potencial para el desarrollo de metodologías sencillas que permitan alcanzar los objetivos planteados en esta tesis doctoral.

En este apartado se explica el fundamento teórico del procesamiento de imágenes digitales según los espacios de color seleccionados, siendo los más comunes en la bibliografía los espacios RGB, CIE L*a*b* y HSI.

Con el conocimiento de los distintos espacios de color es posible representar la información de una imagen de entrada adquirida por un dispositivo digital accesible como una cámara o un escáner. Estos dispositivos reconocen los tres colores primarios del espectro de luz visible, los cuales son el rojo, el azul y el verde y los transforman en un código binario de 8 bits. Las propiedades empleadas para diferenciar los colores son el brillo, el tono y la saturación, las cuales pueden obtenerse a partir de transformaciones de coordenadas de color dependiendo del modelo de color seleccionado. A continuación, se hace una breve descripción de los modelos más empleados en el ámbito del análisis de alimentos y de sus principales características y transformaciones.

En primer lugar, es conveniente establecer una primera clasificación de los espacios de color en función de la aplicación sobre las imágenes procesadas:

1. Modelos dependientes del dispositivo: son aquellos que se ven afectados por la señal del dispositivo. Ejemplos de estos espacios son el

RGB (del inglés *Red, Green, Blue*) o el CMY(K) (del inglés *Cyan, Yellow, Magenta, Black*).

2. **Modelos orientados por el usuario:** estos espacios de color pueden ser creados a partir de cualquier otro espacio de color, y se pueden personalizar para cumplir con ciertos requisitos o para optimizar la visualización de colores para una tarea específica. Ejemplos de este tipo de espacios son el HSI (del inglés '*Hue, Saturation, Intensity*')o el HSL (del inglés '*Hue, Saturation, Lightness*').
3. **Modelos independientes del dispositivo:** son aquellos que no dependen de la capacidad del dispositivo para mostrar o imprimir colores. Ejemplos de este espacio de color son el CIE L*a*b* o el CIE L*U*V*.

Una vez mencionados, a continuación se hace una breve descripción de estos espacios de color y sus potencialidades :

- **Espacio RGB**

El espacio RGB recibe su nombre de los tres colores aditivos procedentes del espectro electromagnético (rojo, azul y verde). Consiste en transformar cada pixel procedente de la imagen tomada en un código binario de 8 bits, que ofrece un total de 256 posibles valores para cada color aditivo y cuya combinación supone un total de 16.777.216 posibles valores.

El espacio RGB puede representarse como un cubo normalizado de los tres valores de RGB entre los rangos [0,1], con el valor del gris en la diagonal principal entre el negro (0,0,0) y el blanco (1,1,1) en la esquina contraria) [186], como se puede observar en la **Figura 16**.

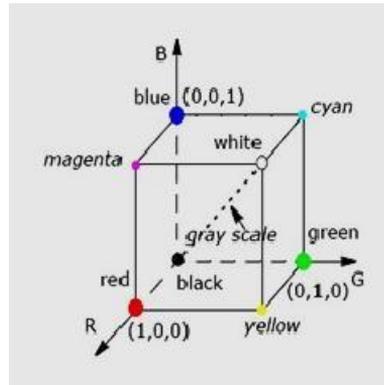


Figura 16. Representación cúbica del espacio de color RGB [187].

El modelo de color RGB se considera como la base de los espacios de color, ya que no requiere de ningún tipo de transformación para tratar la información procedente de dicho espacio, lo cual es una ventaja respecto a la velocidad de procesamiento. Además, es posible usarlo en imágenes en movimiento debido a la posibilidad de la aditividad. Por otro lado, presenta desventajas como cuando el objetivo de la aplicación es la determinación específica de los colores objeto de estudio.

- **Espacio HSI**

El espacio HSI está basado en el sistema visual humano. Este espacio emplea una transformación del espacio RGB en coordenadas cilíndricas. La importancia del modelo HSI reside en dos aspectos fundamentales; el componente I es separado del tono (H) y la saturación (S), que representan la cromaticidad de los componentes. En segundo lugar, los dos últimos componentes representan la “crominancia” y dependen de la percepción del ojo humano. La H representa el tono, cuyo significado es la “pureza del color”, la S indica la saturación, lo cual se entiende como el grado de blanco embebido en el color específico, y la I representa la Intensidad. La conversión entre el espacio RGB y HSI es no lineal [188] y puede verse representada en la **Figura 17**.

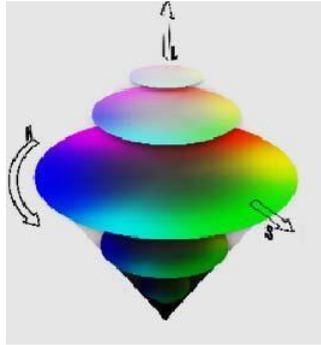


Figura 17. Espacio de color HSI [188].

En cuanto a su aplicabilidad, se pueden destacar ventajas del modelo HSI, como puede ser la asociación de las componentes H y S en relación a la percepción que se tiene de dicho color para el procesamiento de imágenes, o la separación de valores cromáticos de los acromáticos, lo cual puede utilizarse para una mejor diferenciación de las cualidades de las imágenes en relación a la característica analizada. Por el contrario, se pueden detectar desventajas como la indefinición de puntos acromáticos sensibles a las desviaciones de los valores de RGB o la no uniformidad de dicho espacio.

- **Espacio de color CIE $L^*a^*b^*$**

El espacio $L^*a^*b^*$ es derivado del espacio XYZ. Estas componentes dependen de la posición en el espacio de color 3D, la componente L^* es la luminosidad, que oscila entre un valor de 0 y 100, que corresponden al negro y al blanco, respectivamente, la componente a^* oscila para valores positivos en el rojo y para valores negativos en el verde, mientras que la componente b^* oscila para valores positivos en el amarillo y para valores negativos en el azul, tal y como ilustra la **Figura 18**.

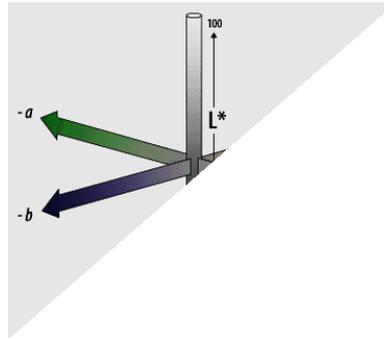


Figura 18: Espacio de color CIE $L^*a^*b^*$ [188].

Este espacio de color es muy usado en el análisis del control de maduración de la fruta “*in-situ*”, ya que posee la cualidad de ser uniforme y permite detectar con mejor precisión las diferencias de color entre muestras diferentes.

En cuanto al procesamiento y análisis de datos de este tipo de imágenes para fines analíticos cabe destacar que existen diferentes métodos de procesamiento de la información, dependiendo de las necesidades del método a desarrollar. Los programas de procesamiento de imágenes más utilizados son: **ImageJ**, Adobe Photoshop, **Matlab**, Java, LabVIEW, Pantone Studio, Digital Colorimeter y Color Pilot, entre otros [189].

Los softwares de procesamiento de imagen permiten obtener las coordenadas del espacio de color RGB para cada uno de los píxeles de la imagen. De este modo, se pueden extraer las características de los tres colores aditivos mediante un promedio de estas tres para comprobar su evolución o su correlación con la concentración de otros analitos [190]. El promedio de cada espacio de color se obtiene de la siguiente manera:

$$R_{ratio} = \frac{R}{R + G + B} \quad (41)$$

$$G_{ratio} = \frac{G}{R + G + B} \quad (42)$$

$$B_{ratio} = \frac{B}{R + G + B} \quad (43)$$

Otro parámetro analítico descrito en la bibliografía es la distancia euclídea [191].

$$ED = \sqrt{(R_s - R_b)^2 + (G_s - G_b)^2 + (B_s - B_b)^2} \quad (44)$$

Donde R_s , G_s y B_s son los valores promedios de la muestra y R_b , G_b y B_b son los valores promedio del blanco.

Una última manera de procesar la información directa obtenida del espacio RGB, a través de los dispositivos que transforman las señales ópticas en señales digitales, es la conversión de esta en escala de grises. Para realizar la conversión existen muchos algoritmos, siendo el más utilizado el de la luminosidad [189, 192], fundamentado en la sensibilidad del ojo humano a los tres colores básicos, a través de la cuál, los coeficientes aplicados para cada canal del espacio RGB se usan de la siguiente manera:

$$I = 0,3R + 0,59G + 0,11B \quad (45)$$

Existen, además, otros modos más sofisticados de procesamiento de la información analítica como puede ser el uso de los histogramas de los espacios de color. Estos consisten en expresar la información arrojada por el procesamiento de datos en diagramas de barras, cuya variable dependiente sería la frecuencia con la que un determinado valor de una característica de los espacios de color, como pueden ser las componentes R, G, B, H, S, I, u otras como la escala de grises, aparecen en la imagen seleccionada. Este tipo de procesamiento de datos sobre el color permite

obtener información analítica multivariante, cuyo uso ha sido ampliamente reportado en revisiones bibliográficas como la realizada por Díaz-Dinz et al. [193]. En ella se puede observar que el uso de histogramas en combinación con algoritmos quimiométricos, resultan de gran utilidad en el desarrollo de metodologías de control de calidad de alimentos, como pueden ser las ciruelas.

Estas metodologías han experimentado un extenso desarrollo en términos de publicaciones científicas tal y como recoge otra revisión bibliográfica realizada por Meenu. M. et al. [194], a partir de la base de datos de *Science Direct* como se puede observar en la **Figura 19**.

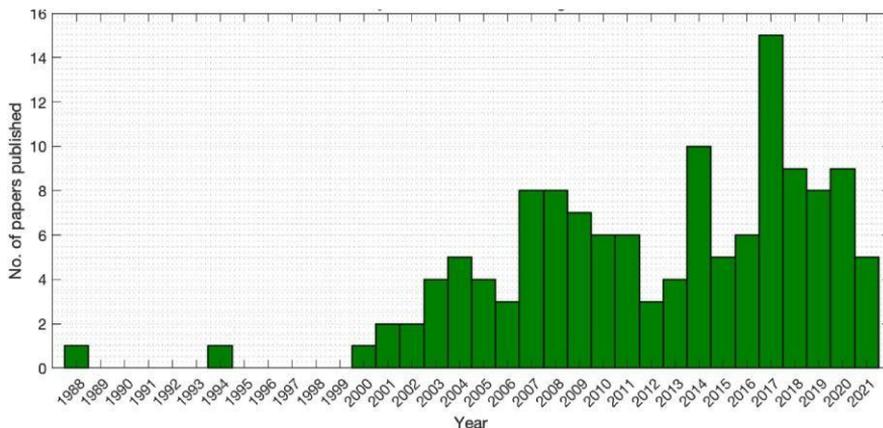


Figura 19. Número total de publicaciones anuales relacionadas con el control de calidad de alimentos a través de herramientas de procesamiento de imágenes digitales (DIP) Meenu. M. et al. [194]

Dentro de esta tendencia, que se viene dando y contrastando en la bibliografía, con respecto al uso de las herramientas de procesamiento de imagen digital y el tratamiento de datos de los espacios de color para el análisis, a continuación, se resumen los procedimientos y softwares empleados en el desarrollo de metodologías analíticas fundamentadas en los datos extraídos de las imágenes tomadas en el espacio de color RGB.

- **Obtención del valor medio de los canales R, G y B en las imágenes tomadas:** se empleó el entorno de Matlab (Matlab R2018a, versión 7.5.0.342, Natick, MA, USA) [195] utilizando para este caso el comando “*imread*”.
- **Obtención de los histogramas de color para análisis de datos de dos vías:** la ordenación de los datos correspondientes al espacio de color RGB en histogramas de orden uno representando la distribución de las 256 posibles intensidades correspondientes a cada color en diagramas de frecuencias fue realizada utilizando el software ImageJ (versión 1.53K, Wayne Rasband y colaboradores, Instituto Nacional de Salud de USA, dominio público) (utilizado el 15 de Enero de 2022) (<https://imagej.nih.gov/ij/>) [171].

El uso de estos histogramas para la obtención de señales multivariantes de primer orden y su tratamiento de datos a través de algoritmos de machine learning y deep learning puede resumirse en la **Figura 20**, obtenida de la revisión realizada por Dias-Dinz et al. en 2020 [193].

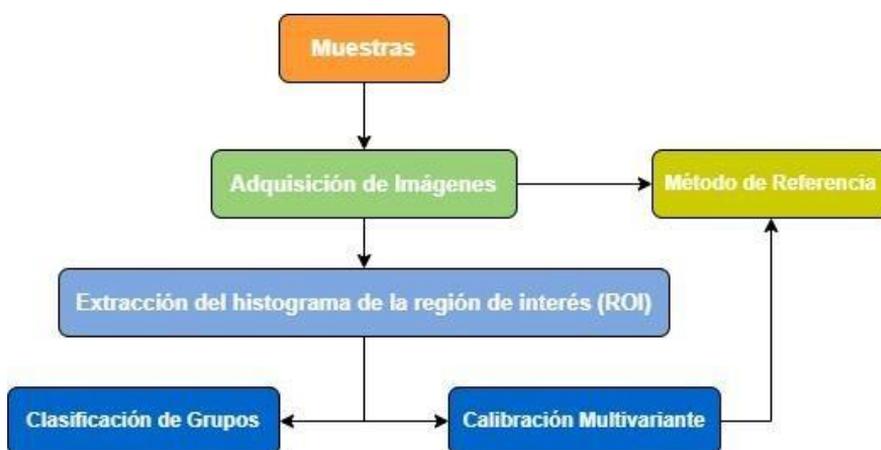


Figura 20. Diagrama de adquisición de datos y procesamiento de imagen en histogramas combinados con métodos quimiométricos.

Por otro lado, también existen estudios de combinación de datos de los diferentes espacios de color, con otras características relacionadas con el tamaño o la textura, que permiten aumentar el número de vías de datos para su análisis multivariante. Este es el caso de estudios realizados para evaluar la evolución del deterioro de distintas maderas con el tiempo [196], o estudios cinéticos de detección de nitritos en alimentos utilizando sensores basados en papel (PBS) [197]. En el caso de matrices alimentarias como las frutas, existen estudios realizados para evaluar los diferentes estados de maduración de los plátanos a través de la combinación de datos de color HSV y textura [198]. En el caso de esta tesis doctoral hay que mencionar que también se han ensayado métodos de tres vías, a través del tratamiento pormenorizado de los píxeles empleados en la región de interés (ROI), de las imágenes tomadas obtenidas a través del propio código “imread” de Matlab [195] mencionado previamente.

Todas estas herramientas han sido objeto de desarrollo para los objetivos planteados respecto al control de calidad y seguimiento del estado de maduración para las ciruelas. En este caso, aprovechando los cambios que los indicadores de madurez como las clorofilas provocan sobre el color de estas y así poder desarrollar métodos que permitan fijar estados de maduración óptimos para la cosecha, de cara a incrementar la calidad y la vida comercial del producto.

5. REFERENCIAS

- [1] COI/T.15/NC no 3/Rev. 7, "Norma Comercial Aplicable a los Aceites de Oliva y los Aceites de Orujo de Oliva," Noviembre de 2012.
- [2] Reglamento (UE) N° 1308/2013 del Parlamento Europeo y del Consejo de 17 de diciembre de 2013 sobre la organización común de mercados agrícolas y por el que se derogan los Reglamentos (CEE) N° 922/72, (CEE) N° 234/79, (CE) N° 1037/2001 y (CE) N° 1234/2007 del Consejo, Diario Oficial de la Unión Europea, L 347/671, 20 de diciembre de 2013.
- [3] A. Fernández Gutiérrez and A. Segura Carretero, "El aceite de oliva virgen tesoro de Andalucía," . Málaga. Servicio de Publicaciones de la Fundación Unicaja, 2009.
- [4] FAO, "FAOSTAT - Bases de datos de estadísticas alimentarias y agrícolas," Organización de las Naciones Unidas para la Alimentación y la Agricultura, [En línea]. Disponible en: <https://www.fao.org/faostat/es/#home>. [Fecha de acceso: 10 de octubre de 2022].
- [5] Ministerio de Agricultura, Pesca y Alimentación de España, "Aceite de Oliva y Aceituna de Mesa," [En línea]. Disponible en: <https://www.mapa.gob.es/es/agricultura/temas/producciones-agricolas/aceite-oliva-y-aceituna-mesa/aceite.aspx>. [Fecha de acceso: 11 de octubre de 2022]
- [6] C. Capurso, M. Massaro, E. Scoditti, G. Vendemiale, and A. Capurso, "Vascular effects of the Mediterranean diet Part I: Anti-hypertensive and anti-thrombotic effects," *Vascular Pharmacology*, vol. 63, no. 3. Elsevier Inc., pp. 118–126, Dec. 01, 2014.

- [7] N. Condelli, M. C. Caruso, F. Galgano, D. Russo, L. Milella, and F. Favati, “Prediction of the antioxidant activity of extra virgin olive oils produced in the Mediterranean area,” *Food Chemistry*, vol. 177, pp. 233–239, Jun. 2015.
- [8] Departamento de Agricultura de los Estados Unidos (USDA), [En línea]. Disponible en: <https://www.usda.gov/>. [Fecha de acceso: 11 de octubre de 2022].
- [9] M. Monfreda, L. Gobbi, and A. Grippa, “Blends of olive oil and sunflower oil: Characterisation and olive oil quantification using fatty acid composition and chemometric tools,” *Food Chemistry*, vol. 134, no. 4, pp. 2283–2290.
- [10] E. Lopez-Huertas, “Health effects of oleic acid and long chain omega-3 fatty acids (EPA and DHA) enriched milks. A review of intervention studies,” *Pharmacological Research*, vol. 61, no. 3. pp. 200–207.
- [11] S. Teres *et al.*, “Oleic acid content is responsible for the reduction in blood pressure induced by olive oil,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 37, pp. 13811–13816, 2008.
- [12] T. H. Adams, R. L. Walzem, D. R. Smith, S. Tseng, and S. B. Smith, “Hamburger high in total, saturated and trans-fatty acids decreases HDL cholesterol and LDL particle diameter, and increases TAG, in mildly hypercholesterolaemic men,” *British Journal of Nutrition*, vol. 103, no. 1, pp. 91–98.

-
- [13] Hioki, H., Miura, T., Miyashita, Y., Ebisawa, S., Motoki, H., Izawa, A., ... & Ikeda, U. (2016). Circulating eicosapentaenoic acid to oleic acid ratio and risk for cardiovascular events in patients with coronary artery disease: A sub-analysis of the SHINANO registry. *IJC Metabolic & Endocrine*, 10, 1-6.
- [14] K. L. Fritsche, "Linoleic acid, vegetable oils & inflammation," *Mo Med*, vol. 111, no. 1, p. 41, 2014.
- [15] F. B. Hu *et al.*, "Diet, lifestyle, and the risk of type 2 diabetes mellitus in women," *New England journal of medicine*, vol. 345, no. 11, pp. 790–797, 2001.
- [16] O. R. Alara, N. H. Abdurahman, and C. I. Ukaegbu, "Extraction of phenolic compounds: A review," *Current Research in Food Science*, vol. 4. Elsevier B.V., pp. 200–214.
- [17] T. Bohn, "Dietary factors affecting polyphenol bioavailability," *Nutrition Reviews*, vol. 72, no. 7, pp. 429–452, 2014.
- [18] V. Neveu *et al.*, "Phenol-Explorer: an online comprehensive database on polyphenol contents in foods.," *Database (Oxford)*, vol. 2010, bap024. [19] J. A. Rothwell *et al.*, "Phenol-Explorer 2.0: a major update of the Phenol-Explorer database integrating data on polyphenol metabolism and pharmacokinetics in humans and experimental animals.," *Database (Oxford)*, vol. 2012.
- [20] D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, ... y A. Scalbert, "HMDB 3.0—the human metabolome database in 2013," *Nucleic Acids Research*, vol. 41, no. D1, pp. D801-D807, 2013.

- [21] K. De La Torre-Carbot, O. Jauregui, E. Gimeno, A. I. Castellote, R. M. Lamuela-Raventós, and M. C. López-Sabater, “Characterization and quantification of phenolic compounds in olive oils by solid-phase extraction, HPLC-DAD, and HPLC-MS/MS,” *J Agric Food Chem*, vol. 53, no. 11, pp. 4331–4340.
- [22] F. Visioli and E. Bernardini, “Extra virgin olive oil’s polyphenols: biological activities,” *Curr Pharm Des*, vol. 17, no. 8, pp. 786–804, 2011.
- [23] M. N. Franco, T. Galeano-Díaz, J. Sánchez, C. De Miguel, and D. Martín-Vertedor, “Total phenolic compounds and tocopherols profiles of seven olive oil varieties grown in the south-west of Spain,” *J Oleo Sci*, vol. 63, no. 2, pp. 115–125, 2014.
- [24] K. Kotsiou and M. Tasioula-Margari, “Monitoring the phenolic compounds of Greek extra-virgin olive oils during storage This work is in memory of Maria Tasioula-Margari, who was the thesis supervisor.,” *Food Chem*, vol. 200, pp. 255–262, Jun. 2016.
- [25] G. Britton, S. Liaaen-Jensen, y H. Pfaender, "Carotenoids. Volume 1A: Isolation and analysis," *Acta Physiologiae Plantarum*, vol. 18, pp. 91-91, 1996.
- [26] N. B. Kyriakidis and P. Skarkalis, “Fluorescence spectra measurement of olive oil and other vegetable oils,” *Journal of AOAC International*, vol. 83, no. 6, pp. 1435–1439, 2000.
- [27] Consejo Oleícola Internacional, "Prueba de espectrofotometría-UV," España: International Olive Council, 2019, pp. 1–7.
- [28] Eitenmiller, R. R., Landen Jr, W. O., & Ye, L. (2016). *Vitamin analysis for the health and food sciences*. CRC Press.

- [29] K. Ward, R. Scarth, J. K. Daun, and C. T. Thorsteinson, "A comparison of high-performance liquid chromatography and spectrophotometry to measure chlorophyll in canola seed and oil," *Journal of American Oil Chemistry Society*, vol. 71, pp. 931–934, 1994.
- [30] S. Udenfriend, *Fluorescence assay in biology and medicine*, vol. 2. Academic Press, 2014.
- [31] T. Galeano Díaz, I. Durán Merás, C. Arturo Correa, B. Roldán, and M. Isabel Rodríguez Cáceres, "Simultaneous Fluorometric Determination of Chlorophylls a and b and Pheophytins a and b in Olive Oil by Partial Least-Squares Calibration," *Journal of agricultural and food chemistry*, vol. 51, no. 24, pp. 6934–6940.
- [32] N. Tena, D. L. García-gonzález, and R. Aparicio, "Evaluation of virgin olive oil thermal deterioration by fluorescence spectroscopy," *Journal of agricultural and food chemistry*, vol. 57, no. 22, pp. 10505–10511, Nov. 2009.
- [33] Comisión Europea, "Reglamento de Ejecución (UE) 2019/1604 de la Comisión de 27 de septiembre de 2019."
- [34] Comisión Europea, "Reglamento (CEE) no 2568/91 de la Comisión, de 11 de julio de 1991, relativo a las características de los aceites de oliva y de los aceites de orujo de oliva y sobre sus métodos de análisis," *Diario Oficial de las Comunidades Europeas*, vol. L 248, pp. 1-83, 1991.
- [35] N. Boukachabine, H. Ajana, and A. El Antari, "A study of fatty acids and triglycerides oil composition and quality parameters of five autochthon olive varieties in Morocco," *Lebanese Science Journal*, vol. 12, no. 2, pp. 45–65, 2011.

- [36] D. Tura, C. Gigliotti, S. Pedò, O. Failla, D. Bassi, and A. Serraiocco, “Influence of cultivar and site of cultivation on levels of lipophilic and hydrophilic antioxidants in virgin olive oils (*Olea Europea* L.) and correlations with oxidative stability,” *Sci Horti*c, vol. 112, no. 1, pp. 108–119, 2007.
- [37] A. Gómez-Rico, G. Fregapane, and M. D. Salvador, “Effect of cultivar and ripening on minor components in Spanish olive fruits and their corresponding virgin olive oils,” *Food Research International*, vol. 41, no. 4, pp. 433–440, 2008.
- [38] M. N. Franco, T. Galeano-Díaz, J. Sánchez, C. De Miguel, and D. Martín-Vertedor, “Total phenolic compounds and tocopherols profiles of seven olive oil varieties grown in the south-west of Spain,” *J Oleo Sci*, vol. 63, no. 2, pp. 115–125, 2014.
- [39] S. Rouas, M. Rahmani, A. El Antari, D. J. Idrissi, A. Souizi, and N. Maata, “Effect of geographical conditions (altitude and pedology) and age of olive plantations on the typicality of olive oil in Moulay Driss Zarhoun,” *Mediterranean Journal of Biosciences*, vol. 1, no. 3, pp. 128–137, 2016.
- [40] F. C. Bouchaala, A. Lazzez, H. Jabeur, L. Daoud, and M. Bouaziz, “Physicochemical characteristics of extra virgin olive oil in function of tree age and harvesting period using chemometric analysis,” *Sci Horti*c, vol. 180, pp. 52–58, Dec. 2014.
- [41] F. Chtourou, E. Valli, A. Bendini, A. Lazzez, T. G. Toschi, and M. Bouaziz, “Effects of Olive Trees Age on the Minor Components of Oueslati Virgin Olive Oils Produced from Olives Harvested at Different Ripening Degrees,” *Journal of american oil chemistry Society*, vol. 94, no. 3, pp. 435–447, 2017.

- [42] M. D. Salvador, F. Aranda, and G. Fregapane, "Contribution of chemical components of Cornicabra virgin olive oils to oxidative stability. A study of three successive crop seasons," *JAOCS, Journal of the American Oil Chemists' Society*, vol. 76, no. 4, pp. 427–432, 1999.
- [43] F. Mansouri *et al.*, "Effect of crop season on the quality and composition of extra virgin olive oils from Greek and Spanish varieties grown in the Oriental region of Morocco," *Emirates journal of food and agriculture*, vol. 30, no. 7, pp. 549–562.
- [44] M. P. Romero, M. J. Tovar, T. Ramo, and M. J. Motilva, "Effect of crop season on the composition of virgin olive oil with protected designation of origin 'Les Garrigues,'" *Journal of american oil chemistry society*, vol. 80, pp. 423–430, 2003.
- [45] M. Beltrán, M. Sánchez-Astudillo, R. Aparicio, and D. L. García-González, "Geographical traceability of virgin olive oils from southwestern Spain by their multi-elemental composition," *food chemistry*, vol. 169, pp. 350–357, 2015.
- [46] T. H. Borges, J. A. Pereira, C. Cabrera-Vique, L. Lara, A. F. Oliveira, and I. Seiquer, "Characterization of Arbequina virgin olive oils produced in different regions of Brazil and Spain: Physicochemical properties, oxidative stability and fatty acid profile," *Food Chemistry*, vol. 215, pp. 454–462, 2017.
- [47] N. Romero, J. Saavedra, F. Tapia, B. Sepúlveda, and R. Aparicio, "Influence of agroclimatic parameters on phenolic and volatile compounds of Chilean virgin olive oils and characterization based on geographical origin, cultivar and ripening stage," *J Sci Food Agric*, vol. 96, no. 2, pp. 583–592, 2016.

- [48] S. M. Arafat, A. M. M. Basuniy, M. E. Elsayed, and H. M. Soliman, "Effect of pedological, cultivar and climatic condition on sterols and quality indices of olive oil," *Scientia Agric*, vol. 13, no. 1, pp. 23–29, 2016.
- [49] M. P. Romero, M. J. Tovar, T. Ramo, and M. J. Motilva, "Effect of crop season on the composition of virgin olive oil with protected designation of origin 'Les garrigues,'" *Journal of the american oil chemistry Society*, vol. 80, no. 5, pp. 423–430, 2003.
- [50] Rached, M. B., Galaverna, G., Cirlini, M., Boujne, D., Zarrouk, M., & Guerfel, M. Pedologic factors affecting virgin olive oil quality of "Chemlali" olive trees (*Olea europaea* L.). *Journal of Oleo Science*, 66(8), 907-915, 2017.
- [51] R. Fernández-Escobar, G. Beltrán, M. A. Sánchez-Zamora, J. García-Novelo, M. P. Aguilera, and M. Uceda, "Olive oil quality decreases with nitrogen over-fertilization," *HortScience*, vol. 41, no. 1, pp. 215–219, 2006.
- [52] P. V. Jordão, M. Encarnação Marcelo, J. I. Lopes, y P. Simoes Lopes, "A long-term experiment on olive tree with nitrogen, phosphorus, and limestone fertilization," in *VI International Symposium on Mineral Nutrition of Fruit Crops 868*, 2008, pp. 313-318.
- [53] M. El Yamani, E. H. Sakar, A. Boussakouran, and Y. Rharrabti, "Influence of ripening index and water regime on the yield and quality of 'moroccan Picholine' virgin olive oil," *OCL - Oilseeds and fats, Crops and Lipids*, vol. 27, 2020.
- [54] V. R. Preedy and R. Watson, *Olives and olive oil in health and disease prevention*. Academic press, 2020.

- [55] F. Grilo, S. Sedaghat, V. Di Stefano, R. Sacchi, T. Caruso, and R. Lo Bianco, "Tree planting density and canopy position affect 'cerasuola' and 'koroneiki' olive oil quality," *Horticulturae*, vol. 7, no. 2, pp. 1–12.
- [56] G. P. García-Inza, D. N. Castro, A. J. Hall, and M. C. Rousseaux, "Responses to temperature of fruit dry weight, oil concentration, and oil fatty acid composition in olive (*olea europaea* L. var. 'arauco')," *European Journal of Agronomy*, vol. 54, pp. 107–115, 2014.
- [57] Stefanoudaki-Katzouraki, E. (2004). *Factors affecting olive oil quality*. Cardiff University (United Kingdom).
- [58] B. Baccouri *et al.*, "Influence of fruit ripening and crop yield on chemical properties of virgin olive oils from seven selected oleasters (*Olea europea* L.)," *Journal of Agronomy*, vol. 6, no. 3, pp. 388–396, 2007.
- [59] A. Dag, Z. Kerem, N. Yogev, I. Zipori, S. Lavee, and E. Ben-David, "Influence of time of harvest and maturity index on olive oil yield and quality," *Science Horticulturae*, vol. 127, no. 3, pp. 358–366.
- [60] M. I. Al-Maaitah, K. M. Al-Absi, and A. Al-Rawashdeh, "Oil quality and quantity of three olive cultivars as influenced by harvesting date in the middle and southern parts of Jordan," *Int. J. Agric. Biol*, vol. 11, no. 3, pp. 266–272, 2009.
- [61] D. Boskou, *Olive oil: chemistry and technology*. AOCS Publishing, 2006.
- [62] D. Boskou, *Olive oil: minor constituents and health*. CRC press, 2008.

- [63] C. Mahfouz and A. Bassal, "Influence de l'altitude et de la date de récolte sur la qualité de l'huile d'olive au Liban sud," *Annales de recherche scientifique*, no. 7, pp. 97–103, 2007.
- [64] A. Mahhou, A. Jermmouni, A. Hadiddou, A. Oukabli y A. Mamouni, "Période de récolte et caractéristiques de l'huile d'olive de quatre variétés en irrigué dans la région de Meknès au Maroc," *Revue Marocaine des Sciences Agronomiques et Vétérinaires*, vol. 2, no. 2.
- [65] A. M. Giuffrè, A. Piscopo, V. Sicari, and M. Poiana, "The effects of harvesting on phenolic compounds and fatty acids content in virgin olive oil (cv Roggianella)," *Riv. Ital. Sostanze Gr*, vol. 87, no. 1, pp. 14–23, 2010.
- [66] N. Caporaso, A. Genovese, M. A. Pérez-Jiménez, R. Olivero-David, and R. Sacchi, "Impact of Olive Harvesting Date on Virgin Olive Oil Volatile Composition in Four Spanish Varieties," *European Journal of Lipid Science and Technology*, vol. 123, no. 9, 2021.
- [67] C. Saglam, Y. T. Tuna, U. Gecgel, and E. S. Atar, "Effects of Olive Harvesting Methods on Oil Quality," *APCBEE Procedia*, vol. 8, pp. 334–342, 2014.
- [68] S. Boussahel, V. Di Stefano, C. Muscarà, M. Cristani, and M. G. Melilli, "Phenolic compounds characterization and antioxidant properties of monocultivar olive oils from northeast algeria," *Agriculture (Switzerland)*, vol. 10, no. 11, pp. 1–13, 2020.
- [69] P. Inglese, F. Famiani, F. Galvano, M. Servili, S. Esposito, and S. Urbani, "3 factors affecting extra-virgin olive oil composition," *Hortic Rev (Am Soc Hortic Sci)*, vol. 38, p. 83, 2011.

- [70] S. A. Vekiari, P. Papadopoulou, and A. Kiritsakis, “Effects of processing methods and commercial storage conditions on the extra virgin olive oil quality indexes,” *Grasas y aceites*, vol. 58, no. 3, pp. 237–242, 2007.
- [71] A. I. Méndez and E. Falqué, “Effect of storage time and container type on the quality of extra-virgin olive oil,” *Food Control*, vol. 18, no. 5, pp. 521–529, 2007.
- [72] M. A. Del Nobile, S. Bove, E. La Notte, and R. Sacchi, “Influence of packaging geometry and material properties on the oxidation kinetic of bottled virgin olive oil,” *J Food Eng*, vol. 57, no. 2, pp. 189–197, 2003.
- [73] I. I. H. Abdalla, M. Khaddor, A. Boussab, D. El Garrouj, M. Ayadi, and T. H. M. Bakheet, “The effect of storage time on the quality of olive oil produced by cooperatives for olive growers in the north of Morocco,” *Asian Journal of Agriculture and Food Sciences*, vol. 2, no. 2, 2014.
- [74] V. Di Stefano and M. G. Melilli, “Effect of storage on quality parameters and phenolic content of Italian extra-virgin olive oils,” *Nat Prod Res*, vol. 34, no. 1, pp. 78–86, 2020.
- [75] S. El Qarnifa, A. El Antari, and A. Hafidi, “Effect of maturity and environmental conditions on chemical composition of olive oils of introduced cultivars in Morocco,” *J Food Qual*, vol. 2019, pp. 1–14, 2019.
- [76] C. S. P. Santos, R. Cruz, S. C. Cunha, and S. Casal, “Effect of cooking on olive oil quality attributes,” *Food Research International*, vol. 54, no. 2, pp. 2016–2024, 2013.

- [77] C. Peri, *The extra virgin olive oil handbook*. Wiley Online Library, 2014.
- [78] M. Brenes, A. García, M. C. Dobarganes, J. Velasco, and C. Romero, "Influence of thermal treatments simulating cooking processes on the polyphenol content in virgin olive oil," *Journal of Agricultural Food Chemistry*, vol. 50, no. 21, pp. 5962–5967, 2002.
- [79] A. Fullana, A. A. Carbonell-Barrachina, and S. Sidhu, "Comparison of volatile aldehydes present in the cooking fumes of extra virgin olive, olive, and canola oils," *J Agric Food Chem*, vol. 52, no. 16, pp. 5207–5214, 2004.
- [80] A. M. Giuffrè, C. Zappia, and M. Capocasale, "Effects of High Temperatures and Duration of Heating on Olive Oil Properties for Food Use and Biodiesel Production," *JAACS, Journal of the American Oil Chemists' Society*, vol. 94, no. 6, pp. 819–830, 2017.
- [81] P. del Gobierno, "Real Decreto 1945/1983 de 22 de Junio por el que se regulan las infracciones y sanciones en materia de defensa del consumidor y de la producción agro-alimentaria," *BOE*. pp. 1–17, 1983.
- [82] E. Casadei, E. Valli, F. Panni, J. Donarski, J. F. Gubern, P. Lucci, ... y T. G. Toschi, "Emerging trends in olive oil fraud and possible countermeasures," *Food Control*, vol. 124, p. 107902, 2021.
- [83] W. A. Salah and M. Nofal, "Review of some adulteration detection techniques of edible oils," *Journal of the Science of Food and Agriculture*, vol. 101, no. 3. John Wiley and Sons Ltd, pp. 811–819, 2021.

- [84] W. Carrillo, C. Carpio, D. Morales, E. Vilcacundo, M. Álvarez, and M. Silva, "Content of fatty acids in corn (*Zea mays* L.) oil from Ecuador," *Asian Journal of Pharmaceutical and Clinical Research*, vol. 10, no. 8, pp. 150–153, 2017.
- [85] S. Azadmard-Damirchi and M. Torbati, "Adulterations in Some Edible Oils and Fats and Their Detection Methods.," *J Food Qual Hazards Control*, vol. 2, no. 2, 2015.
- [86] I. Criado-Navarro, A. Mena-Bravo, M. Calderón-Santiago, and F. Priego-Capote, "Determination of glycerophospholipids in vegetable edible oils: Proof of concept to discriminate olive oil categories," *Food Chem*, vol. 299, 2019.
- [87] Comisión del Codex Alimentarius, "Norma para las frutas de hueso en conserva 1 CXS 242-2003," pp. 1-8, 2003.
- [88] P. y A. Ministerio de Agricultura, "Anuario de Estadística del Ministerio de Agricultura, Pesca y Alimentación." 2021.
- [89] P. y A. de E. Ministerio de Agricultura, "Plan de Fruta Dulce," Ministerio de Agricultura, Pesca y Alimentación.
- [90] B. G. López, J. Rodrigo, and M. E. Guerra, "Situación y perspectivas del cultivo del ciruelo japonés (híbridos de *Prunus salicina* Lindl.): la marca de calidad CiEx® en Extremadura," *Revista de fruticultura*, no. 65, pp. 82–91, 2018.
- [91] S. H. Quiliche Estacio, "Efectos biológicos de los antioxidantes y fitoquímicos de alimentos que presentan pigmentos antociánicos," 2013.

- [92] A. Subar and B. Patterson, "Fruit, Vegetables, and Cancer Prevention: A Review of the Epidemiological Evidence," *Nutrition and Cancer*, vol. 18, no. 1, pp. 1–29, 1992.
- [93] L. J. Machlin, "Critical assessment of the epidemiological data concerning the impact of antioxidant nutrients on cancer and cardiovascular disease," *Critical Reviews in Food Science and Nutrition*, vol. 35, no. 1–2, pp. 41–49, 1995.
- [94] B. N. Ames, L. S. Gold, and W. C. Willett, "The causes and prevention of cancer.," *Proceedings of the National Academy of Sciences*, vol. 92, no. 12, pp. 5258–5265, 1995.
- [95] M. Siddiq, J. F. Arnold, N. K. Sinha, and J. N. CASH, "Effect of polyphenol oxidase and its inhibitors on anthocyanin changes in plum juice," *Journal of Food Processing and Preservation*, vol. 18, no. 1, pp. 75–84, 1994.
- [96] R. Jaiswal *et al.*, "Identification of phenolic compounds in plum fruits (*prunus salicina* l. and *prunus domestica* l.) by high-performance liquid chromatography/tandem mass spectrometry and characterization of varieties by quantitative phenolic fingerprints," *Journal of Agricultural Food Chemistry*, vol. 61, no. 49, pp. 12020–12031, 2013.
- [97] D. O. Kim, O. K. Chun, Y. J. Kim, H. Y. Moon, and C. Y. Lee, "Quantification of Polyphenolics and Their Antioxidant Capacity in Fresh Plums," *Journal of Agricultural Food Chemistry*, vol. 51, no. 22, pp. 6509–6515, 2003.

-
- [98] M. Cabrera-Bañegil *et al.*, “Evolution of polyphenols content in plum fruits (*Prunus salicina*) with harvesting time by second-order excitation-emission fluorescence multivariate calibration,” *Microchemical Journal*, vol. 158, 2020.
- [99] P. M. Kris-Etherton *et al.*, “Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer,” *American Journal of Medicine*, vol. 113, no. 9, Supplement 2, pp. 71–88, 2002.
- [100] V. Usenik, D. Kastelec, R. Veberič, and F. Štampar, “Quality changes during ripening of plums (*Prunus domestica* L.),” *Food Chemistry*, vol. 111, no. 4, pp. 830–836, 2008.
- [101] J. Wang, H. Pan, R. Wang, K. Hong, and J. Cao, “Patterns of flesh reddening, translucency, ethylene production and storability of ‘Friar’ plum fruit harvested at three maturity stages as affected by the storage temperature,” *Postharvest Biology and Technology*, vol. 121, pp. 9–18, 2016.
- [102] F. C. Stintzing, A. S. Stintzing, R. Carle, B. Frei, and R. E. Wrolstad, “Color and antioxidant properties of cyanidin-based anthocyanin pigments,” *Journal of Agricultural Food Chemistry*, vol. 50, no. 21, pp. 6172–6181, 2002.
- [103] M. Stacewicz-Sapuntzakis, P. E. Bowen, E. A. Hussain, B. I. Damayanti-Wood, and N. R. Farnsworth, “Chemical composition and potential health effects of prunes: A functional food?” *Critical Reviews in Food Science and Nutrition*, vol. 41, no. 4, pp. 251–286, 2001.

- [104] A. Kaulmann, M. C. Jonville, Y. J. Schneider, L. Hoffmann, and T. Bohn, "Carotenoids, polyphenols and micronutrient profiles of Brassica oleraceae and plum varieties and their contribution to measures of total antioxidant capacity," *Food Chemistry*, vol. 155, pp. 240–250, 2014.
- [105] R. A. Vlaic *et al.*, "The changes of polyphenols, flavonoids, anthocyanins and chlorophyll content in plum peels during growth phases: From fructification to ripening," *Notulae Botanicae Horti Agrobotanici Cluj Napoca*, vol. 46, no. 1, pp. 148–155, 2018.
- [106] Z. Luo, J. Xie, T. Xu, and L. Zhang, "Delay ripening of 'Qingnai' plum (*Prunus salicina* Lindl.) with 1-methylcyclopropene," *Plant Science*, vol. 177, no. 6, pp. 705–709, Dec. 2009.
- [107] B. Schoefs, "Chlorophyll and carotenoid analysis in food products. Properties of the pigments and methods of analysis," *Trends in Food Science & Technol*, vol. 13, no. 11, pp. 361–371, 2002.
- [108] C. Europea, "Reglamento Delegado (UE) 2019/428 de la Comisión, de 12 de julio de 2018, que modifica el Reglamento de Ejecución (UE) no 543/2011 en lo que atañe a las normas de comercialización en el sector de las frutas y hortalizas," *Diario Oficial de la Unión Europea*, vol. L 77, no. 2, 2019.
- [109] P. E. y C. de la U. Europea, "Reglamento (CE) no 852/2004 del Parlamento Europeo y del Consejo, de 29 de abril de 2004, relativo a la higiene de los productos alimenticios," *Diario Oficial de la Unión Europea*, vol. L 139, pp. 1-54, 2004.
- [110] C. H. Crisosto, "Optimum procedures for ripening stone fruit," *Management of Fruit Ripening. University of California, Davis, Postharvest Horticulture Series*, vol. 9, pp. 28–30, 1999.

- [111] Podestá, L. Ciruela japonesa: parámetros de madurez y calidad. 2002.
- [112] D. Ruiz *et al.*, “New promising Japanese plum cultivars for warm areas from CEBAS-CSIC/IMIDA breeding programme,” in *XII International Symposium on Plum and Prune Genetics, Breeding and Pomology 1322*, 2021, pp. 55–60.
- [113] Ministerio de Agricultura, Pesca y Alimentación. "Información sobre Prunus." Sitio web del Ministerio de Agricultura, Pesca y Alimentación de España. https://www.mapa.gob.es/ministerio/pags/biblioteca/hojas/hojas_materia.asp?materiaid=ZZ0043585&materia=PRUNUS. Consultado el [30/10/2022].
- [114] E. Vangdal, M. Meland, F. MÅ¥ge, and A. DÅ, ving, “Prediction of fruit quality of plums (*Prunus domestica* L.),” in *III International Symposium on Applications of Modelling as an Innovative Technology in the Agri-Food Chain; MODEL-IT 674*, 2005, pp. 613–617.
- [115] V. OLTENACU and C. PETRIȘOR, “YIELD AND FRUIT CHEMICAL COMPOSITION OF SOME PLUM CULTIVARS AFFECTED BY DIFFERENT SOIL TYPES.,” *Scientific Papers. Series B. Horticulture*, vol. 65, no. 1, 2021.
- [116] M. A. Rettke, T. R. Pitt, N. A. Maier, and J. A. Jones, “Quality of fresh and dried fruit of apricot (cv. Moorpark) in response to soil-applied nitrogen,” *Australian Journal of Experimental Agriculture*, vol. 46, no. 1, pp. 123–129, 2006.

- [117] C. H. Crisosto, F. G. Mitchell, and S. Johnson, “Factors in fresh market stone fruit quality,” *Postharvest News and Information*, vol. 6, no. 2, pp. 17–21, 1995.
- [118] K. Daane *et al.*, “Excess nitrogen raises nectarine susceptibility to disease and insects,” *California Agriculture (Berkeley)*, vol. 49, no. 4, pp. 13–18, 1995.
- [119] K. C. Taylor, “Cultural management of the bearing peach orchard,” University of Georgia, 2009.
- [120] I. T. Chatzitheodorou, T. E. Sotiropoulos, and G. I. Mouhtaridou, “Effect of nitrogen, phosphorus, potassium fertilization and manure on fruit yield and fruit quality of the peach cultivars ‘Spring Time’ and ‘Red Haven,’” *Agronomy research*, vol. 2, no. 2, pp. 135–143, 2004.
- [121] R. Ruiz, “Effects of different potassium fertilizers on yield, fruit quality and nutritional status of Fairlane nectarine trees and on soil fertility,” in *V International Symposium on Mineral Nutrition of Fruit Plants 721*, 2005, pp. 185–190.
- [122] A. A. Kader and R. S. Rolle, *The role of post-harvest management in assuring the quality and safety of horticultural produce*, vol. 152. Food & Agriculture Org., 2004.
- [123] J. Ystaas and O. Frøynes, “Plum tree nutrition: effects of phosphorus, liming and some other elements on vigour, yield, fruit weight and fruit quality of ‘Mallard’ plum (*Prunus domestica* L.) grown on a virgin, acid soil,” *Norwegian Journal of Agricultural Sciences*, vol. 9, pp. 217–230, 1995.

- [124] I. F. Hassan, M. S. Gaballah, H. M. El-Hoseiny, M. E. El-Sharnouby, and S. M. Alam-Eldein, "Deficit irrigation to enhance fruit quality of the 'African Rose' Plum under the Egyptian semi-arid conditions," *Agronomy*, vol. 11, no. 7, p. 1405, 2021.
- [125] X. J. Murray, D. M. Holcroft, N. C. Cook, and S. J. E. Wand, "Postharvest quality of 'Laetitia' and 'Songold' (*Prunus salicina* Lindell) plums as affected by preharvest shading treatments," *Postharvest Biology and Technology*, vol. 37, no. 1, pp. 81–92, 2005.
- [126] M. Guerra, M. A. Sanz, and P. A. Casquero, "Influence of harvest dates on quality, storage capacity and sensory attributes of European plum cv. green gage," *Food Science and Technology International*, vol. 15, no. 6, pp. 527–534, 2009.
- [127] C. H. Crisosto and A. A. Kader, "Plum and fresh prune postharvest quality maintenance guidelines," *Department of Plant Sciences, University of California, Davis, CA*, vol. 95616, pp. 1–8, 2000.
- [128] C. H. Crisosto, E. J. Mitcham, and A. A. Kader, "Recommendations for Maintaining Postharvest Quality Produce Facts Plum," 2002.
- [129] C. M. Cantín, C. H. Crisosto, and K. R. Day, "Evaluation of the effect of different modified atmosphere packaging box liners on the quality and shelf life of 'Friar' plums," *Horttechnology*, vol. 18, no. 2, pp. 261–265, 2008.
- [130] M. Serek, E. C. Sisler, and M. S. Reid, "Novel gaseous ethylene binding inhibitor prevents ethylene effects in potted flowering plants," *Journal of the American Society for Horticultural Science*, vol. 119, no. 6, pp. 1230–1233, 1994.

- [131] L. Dong, H.-W. Zhou, L. Sonogo, A. Lers, and S. Lurie, “Ripening of ‘Red Rosa’ plums: effect of ethylene and 1-methylcyclopropene,” *Functional Plant Biology*, vol. 28, no. 10, pp. 1039–1045, 2001.
- [132] G. A. Manganaris, A. R. Vicente, C. H. Crisosto, and J. M. Labavitch, “Effect of dips in a 1-methylcyclopropene-generating solution on ‘Harrow Sun’ plums stored under different temperature regimes,” *Journal of Agricultural Food Chemistry*, vol. 55, no. 17, pp. 7015–7020, 2007.
- [133] A. M. Menniti, R. Gregori, and I. Donati, “1-Methylcyclopropene retards postharvest softening of plums,” *Postharvest Biology and Technology*, vol. 31, no. 3, pp. 269–275, 2004.
- [134] Z. Luo, J. Xie, T. Xu, and L. Zhang, “Delay ripening of ‘Qingnai’ plum (*Prunus salicina* Lindl.) with 1-methylcyclopropene,” *Plant Science*, vol. 177, no. 6, pp. 705–709, Dec. 2009.
- [135] H. Pan, R. Wang, L. Li, J. Wang, J. Cao, and W. Jiang, “Manipulation of ripening progress of different plum cultivars during shelf life by post-storage treatments with ethylene and 1-methylcyclopropene,” *Scientia Horticulturae*, vol. 198, pp. 176–182, 2016.
- [136] S. Lurie, “Postharvest heat treatments,” *Postharvest Biology and Technology*, vol. 14, no. 3, pp. 257–269, 1998.
- [137] D. Valero, A. Pérez-Vicente, D. Martínez-Romero, S. Castillo, F. Guillen, and M. Serrano, “Plum storability improved after calcium and heat postharvest treatments: role of polyamines,” *Journal of Food Science*, vol. 67, no. 7, pp. 2571–2575, 2002.

- [138] A. B. Woolf and I. B. Ferguson, "Postharvest responses to high fruit temperatures in the field," *Postharvest Biology and Technology*, vol. 21, no. 1, pp. 7–20, 2000.
- [139] A. Pascual, I. Llorca, and A. Canut, "Use of ozone in food industries for reducing the environmental impact of cleaning and disinfection activities," *Trends in Food Science and Technology*, vol. 18, no. SUPPL. 1, 2007.
- [140] M. Serrano, D. Martínez-Romero, S. Castillo, F. Guillén, and D. Valero, "Role of calcium and heat treatments in alleviating physiological changes induced by mechanical damage in plum," *Postharvest Biology and Technology*, vol. 34, no. 2, pp. 155–167, 2004. [141] A. Pérez-Vicente *et al.*, "Role of polyamines in extending shelf life and the reduction of mechanical damage during plum (*Prunus salicina* Lindl.) storage," *Postharvest Biology and Technology*, vol. 25, no. 1, pp. 25–32, 2002.
- [142] M. B. Perez-Gago, C. Rojas, and M. A. Del Rio, "Effect of hydroxypropyl methylcellulose-lipid edible composite coatings on plum (cv. Autumn giant) quality during storage," *Journal of Food Science*, vol. 68, no. 3, pp. 879–883, 2003.
- [143] A. M. Svircev, R. J. Smith, T. Zhou, M. Hernandez, W. Liu, and C. L. Chu, "Effects of thymol fumigation on survival and ultrastructure of *Monilinia fructicola*," *Postharvest Biology and Technology*, vol. 45, no. 2, pp. 228–233, 2007.
- [144] W. T. Liu, C. L. Chu, and T. Zhou, "Thymol and acetic acid vapors reduce postharvest brown rot of apricots and plums," *HortScience*, vol. 37, no. 1, pp. 151–156, 2002.

- [145] J. B. F. LLOYD, “Synchronized Excitation of Fluorescence Emission Spectra,” *Nature Physical Science*, vol. 231, no. 20, pp. 64–65, 1971.
- [146] G. WEBER, “Enumeration of Components in Complex Systems by Fluorescence Spectrophotometry,” *Nature*, vol. 190, no. 4770, pp. 27–29, 1961.
- [147] F. Guimet, J. Ferré, R. Boqué, M. Vidal, and J. Garcia, “Excitation–emission fluorescence spectroscopy combined with three-way methods of analysis as a complementary technique for olive oil characterization,” *Journal of Agricultural Food Chemistry*, vol. 53, no. 24, pp. 9319–9328, 2005.
- [148] B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. M. Cheung, and M. J. Schnitzer, “Fiber-optic fluorescence imaging,” *Nature Methods*, vol. 2, no. 12, pp. 941–950, 2005.
- [149] E. H. Murchie and T. Lawson, “Chlorophyll fluorescence analysis: A guide to good practice and understanding some new applications,” *Journal of Experimental Botany*, vol. 64, no. 13, pp. 3983–3998, 2013.
- [150] R. A. Vlaic, A. E. Muresan, C. C. Muresan, G. S. Petrut, V. Muresan, and S. Muste, “Quantitative analysis by HPLC and FT-MIR prediction of individual sugars from the plum fruit harvested during growth and fruit development,” *Agronomy*, vol. 8, no. 12, 2018.
- [151] C. Sousa, “Anthocyanins, carotenoids and chlorophylls in edible plant leaves unveiled by tandem mass spectrometry,” *Foods*, vol. 11, no. 13, p. 1924, 2022.

- [152] S. Wold, “Spline functions in data analysis,” *Technometrics*, vol. 16, no. 1, pp. 1–11, 1974.
- [153] J. A. Arancibia, P. C. Damiani, G. A. Ibañez, and A. C. Olivieri, “Recent applications of first-and second-order multivariate calibration to analytical chemistry,” *JAOAC Int*, vol. 97, no. 1, pp. 39–49, 2014.
- [154] S. Wold, “Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models,” *Technometrics*, vol. 20, no. 4, pp. 397–405, 1978.
- [155] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1–3, pp. 37–52, 1987.
- [156] J. B. Kruskal, “Factor analysis and principal components: Bilinear methods.” *International Encyclopedia of Statistics*, New York: Free Press, 1978.
- [157] G. W. Johnson and R. Ehrlich, “State of the art report on multivariate chemometric methods in environmental forensics,” *Environ Forensics*, vol. 3, no. 1, pp. 59–79, 2002.
- [158] J. E. Jackson, “Principal components and factor analysis: part I—principal components,” *Journal of Quality Technology*, vol. 12, no. 4, pp. 201–213, 1980.
- [159] A. C. Peinado Amores, *Aplicación y desarrollo de técnicas quimiométricas para el seguimiento de fermentaciones alcohólicas*. Universitat Autònoma de Barcelona, 2007.

- [160] H. Wold, “Estimation of principal components and related models by iterative least squares,” *Multivariate analysis*, pp. 391–420, 1966.
- [161] H. Martens and T. Naes, *Multivariate calibration*. John Wiley & Sons, 1992.
- [162] H. Wold, “Systems analysis by partial least squares,” 1983.
- [163] D. M. Haaland and E. V Thomas, “Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information,” *Analytical Chemistry*, vol. 60, no. 11, pp. 1193–1202, 1988.
- [164] D. M. Haaland and E. V Thomas, “Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data,” *Analytical Chemistry*, vol. 60, no. 11, pp. 1202–1208, 1988.
- [165] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Taylor & Francis, 1984. [Online]. Available: <https://books.google.es/books?id=JwQx-WOmSyQC>
- [166] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE Trans Inf Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [167] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach Learn*, vol. 20, pp. 273–297, 1995.
- [168] S. S. Haykin, *Neural Networks: A Comprehensive Foundation*. in International edition. Prentice Hall, 1999. [Online]. Available: <https://books.google.es/books?id=bX4pAQAAAMAJ>
- [169] J. E. Jackson, *A user’s guide to principal components*. John Wiley & Sons, 2005.

-
- [170] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
- [171] W. S. , C. T. J. Rasband, “ImageJ. Versión 1.53K. Editor: Instituto Nacional de Salud de USA. Dominio público. Disponible en: <https://imagej.nih.gov/ij/>.”
- [172] CAMO Software AS, “The Unscrambler version 6.11 [Software]. Oslo, Noruega.” 2012.
- [173] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multimodal factor analysis," 1970, pp. 1-84.
- [174] J. D. Carroll and J.-J. Chang, “Analysis of individual differences in multidimensional scaling via an N-way generalization of ‘Eckart-Young’ decomposition,” *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
- [175] R. Bro and H. A. L. Kiers, “A new efficient method for determining the number of components in PARAFAC models,” *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 17, no. 5, pp. 274–286, 2003.
- [176] A. C. Olivieri, H. L. Wu, and R. Q. Yu, “MVC2: A MATLAB graphical interface toolbox for second-order multivariate calibration,” *Chemometrics and Intelligent Laboratory Systems*, vol. 96, no. 2, pp. 246–251, 2009.
- [177] C. R. Rao, “The utilization of multiple measurements in problems of biological classification,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 10, no. 2, pp. 159–203, 1948.

- [178] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*, vol. 112. Springer, 2013.
- [179] E. K. Kemsley, “A genetic algorithm (GA) approach to the calculation of canonical variates (CVs),” *TrAC Trends in Analytical Chemistry*, vol. 17, no. 1, pp. 24–34, 1998.
- [180] D. B. Gil, A. M. De La Peña, J. A. Arancibia, G. M. Escandar, and A. C. Olivieri, “Second-order advantage achieved by unfolded-partial least-squares/residual bilinearization modeling of excitation-emission fluorescence data presenting inner filter effects,” *Analytical Chemistry*, vol. 78, no. 23, pp. 8051–8058, 2006.
- [181] R. G. Brereton and G. R. Lloyd, “Partial least squares discriminant analysis: Taking the magic away,” *Journal of Chemometrics*, vol. 28, no. 4, pp. 213–225, 2014.
- [182] R. Tauler, “Multivariate curve resolution applied to second order data,” *Chemometrics and intelligent laboratory systems*, vol. 30, no. 1, pp. 133–146, 1995.
- [183] A. C. Olivieri, “On a versatile second-order multivariate calibration method based on partial least-squares and residual bilinearization: Second-order advantage and precision properties,” *Journal of Chemometrics*, vol. 19, no. 4, pp. 253–265, 2005.
- [184] F. A. Chiappini, A. M. de la Peña, H. C. Goicoechea, and A. C. Olivieri, “An upgrade of MVC2, a MATLAB graphical user interface for second-order multivariate calibration: Beyond trilinear models,” *Chemometrics and Intelligent Laboratory Systems*, vol. 237, p. 104814, 2023.

-
- [185] A. Muñoz de la Peña, N. Mujumdar, E. C. Heider, H. C. Goicoechea, D. Muñoz de la Peña, and A. D. Campiglia, “Nondestructive Total Excitation–Emission Fluorescence Microscopy Combined with Multi-Way Chemometric Analysis for Visually Indistinguishable Single Fiber Discrimination,” *Analytical Chemistry*, vol. 88, no. 5, pp. 2967–2975, 2016.
- [186] A. Ford and A. Roberts, “Colour space conversions,” *Westminster University, London*, vol. 1998, pp. 1–31, 1998.
- [187] N. A. Ibraheem, M. M. Hasan, R. Z. Khan, and P. K. Mishra, “Understanding color models: a review,” *ARPJ Journal of science and technology*, vol. 2, no. 3, pp. 265–275, 2012.
- [188] K. Plataniotis and A. N. Venetsanopoulos, *Color image processing and applications*. Springer Science & Business Media, 2000.
- [189] L. F. Capitán-Vallvey, N. López-Ruiz, A. Martínez-Olmos, M. M. Erenas, and A. J. Palma, “Recent developments in computer vision-based analytical chemistry: A tutorial review,” *Analytical Chimica Acta*, vol. 899, pp. 23–56, 2015.
- [190] V. Rotich, D. F. Al Riza, F. Giametta, T. Suzuki, Y. Ogawa, and N. Kondo, “Thermal oxidation assessment of Italian extra virgin olive oil using an UltraViolet (UV) induced fluorescence imaging system,” *Spectrochimica Acta Part A Molecular and Biomolecular Spectroscopy*, vol. 237, p. 118373, 2020.
- [191] M. Aberrahim, S. Arribas, and L. Condezo-Hoyos, “A novel high-throughput image based rapid Folin-Ciocalteu assay for assessment of reducing capacity,” 2016.

- [192] J. Passaretti Filho, J. F. Da Silveira Petrucci, and A. A. Cardoso, “Development of a simple method for determination of NO₂ in air using digital scanner images,” *Talanta*, vol. 140, pp. 73–80, 2015.
- [193] P. H. Gonçalves Dias Diniz, “Chemometrics-assisted color histogram-based analytical systems,” *J Chemom*, vol. 34, no. 12, p. e3242, 2020.
- [194] M. Meenu, Q. Cai, and B. Xu, “A critical review on analytical techniques to detect adulteration of extra virgin olive oil,” *Trends in Food Science and Technology*, vol. 91. Elsevier Ltd, pp. 391–408, 2019.
- [195] MathWorks, “Matlab R2018a. Versión 7.5.0.342. Editor: MathWorks. Natick, MA, USA, 2018. Software .exe. Disponible en: <https://www.mathworks.com/products/matlab.html>.” 2018.
- [196] J. Sandak, A. Sandak, and M. Cocchi, “Multi-sensor data fusion and parallel factor analysis reveals kinetics of wood weathering,” *Talanta*, vol. 225, p. 122024, 2021.
- [197] Z. Almasvandi, A. Vahidinia, A. Heshmati, M. M. Zangeneh, H. C. Goicoechea, and A. R. Jalalvand, “Coupling of digital image processing and three-way calibration to assist a paper-based sensor for determination of nitrite in food samples,” *RSC Advances*, vol. 10, no. 24, pp. 14422–14430, 2020.
- [198] J. Zhuang *et al.*, “Assessment of external properties for identifying banana fruit maturity stages using optical imaging techniques,” *Sensors*, vol. 19, no. 13, p. 2910, 2019.

IV. EXPERIMENTAL

**CAPÍTULO 1: SEGUIMIENTO Y CONTROL DE
CALIDAD DE LOS ACEITES DE OLIVA Y DETECCIÓN
DE FRAUDES.**

1. INTRODUCCIÓN

El aceite de oliva es un producto agroalimentario que, dada su singular composición química, rica en ácidos grasos monoinsaturados, compuestos fenólicos y vitamina E, entre otros, presenta unas características sensoriales y funcionales que lo hacen de especial interés para los consumidores. Por este motivo, y dado el impacto económico, territorial y ambiental que genera su producción y comercialización en los principales países productores (*véase apartado 1.1.1 de la introducción de esta memoria*), se hace necesario establecer una normativa general, impulsada por el consejo oleícola internacional (COI), y legislada por los distintos estamentos competentes (*véase legislación citada previamente*), que garantice al consumidor la calidad que se espera por el producto que adquiere, regulando así las condiciones de producción y las características sensoriales y composicionales mínimas que han de cumplir los aceites producidos para ser vendidos bajo dichas condiciones.

La producción limitada de aceites de oliva de calidad diferenciada, como son los aceites de oliva virgen extra, unida a la gran demanda que suscita dicho producto, hacen cada vez más necesaria la inversión en sistemas de control de calidad que garanticen las características bajo las cuales se comercializan, incentivando de este modo una producción de calidad a través de todas las fases del proceso y evitando prácticas fraudulentas como puede ser la adulteración de los aceites de oliva virgen extra con otros aceites de menor calidad.

Por este motivo, se requiere el desarrollo de metodologías analíticas que ofrezcan ventajas competitivas frente a los métodos tradicionales de análisis, garantizando la rapidez y automatización de los procesos, la reducción de costes instrumentales y material fungible y el aumento del volumen de muestras por unidad de tiempo o la robustez.

Actualmente, el desarrollo técnico de los equipos instrumentales de análisis y la capacidad de procesamiento de los equipos informáticos, unidos al uso de herramientas matemáticas de modelización y análisis multivariante de datos, ha abierto un campo de investigación muy extenso en el que se hace posible la mejora de los procesos de análisis de alimentos. Las publicaciones referidas a la aplicación de métodos quimiométricos para la detección de fraudes en aceites de oliva en combinación con técnicas analíticas tales como, espectroscopía Raman [1, 2], electroforesis capilar [3, 4], cromatografía líquida de alta eficacia (HPLC) [5] o cromatografía de gases (GC) [6], han experimentado un auge durante la última década, a través del análisis de diferentes analitos como pueden ser el perfil de ácidos grasos, esteroides, compuestos volátiles o polifenoles.

A continuación, se detalla la investigación objeto de estudio de este capítulo sobre el desarrollo de metodologías analíticas, aplicando métodos quimiométricos de segundo orden sobre las matrices de fluorescencia de excitación-emisión (EEMs), las cuales se han obtenido a partir de aceites de oliva sin tratamiento previo, usando la técnica de fluorescencia *'front-face'*.

Aprovechando la fluorescencia intrínseca de los compuestos indicadores de la calidad y del grado de deterioro oxidativo de los propios aceites de oliva, como son los tocoferoles, los polifenoles, los productos de oxidación secundaria (dienos, trienos conjugados y peróxidos) y las clorofilas, se establecen patrones de discriminación para aceites de oliva de diferente calidad, así como para la detección de fraudes o deterioro oxidativo a través de diferentes estrategias.

Existe una extensa bibliografía relacionada con la aplicación de la fluorescencia nativa de los aceites de oliva para el desarrollo de metodologías analíticas, fundamentadas en métodos quimiométricos de

análisis discriminante o regresión multivariante para la detección de fraudes o del deterioro oxidativo. En este sentido se han publicado numerosas revisiones [7]–[11] donde se detalla el desarrollo de diversos métodos de interés. En la **Tabla 1** se hace referencia a los estudios más recientes en este campo para observar las tendencias actuales.

Tabla 1. Resumen de los métodos de fluorescencia molecular usados para la determinación de la calidad de los aceites de oliva virgen extra (EVOO) y aceites de oliva virgen (VOO).

Muestras	Número de muestras	Técnica	Análisis de datos	Objetivos	Resultados	Referencias
EVOO	65	EEFS	PCA, PARAFAC, LDA, DN-PLSR	Autenticidad según origen geográfico de EVOO de dos regiones diferentes de Malta.	73, 80 y 93% de clasificaciones correctas de muestras de EVOO de acuerdo con su origen geográfico aplicando respectivamente LDA, PARAFAC y DN-PLS.	[12]
EVOO y VOO	87	EEFS	PCA, LDA, SVM y k-NN	Autenticidad según origen geográfico de EVOO y VOO de cinco regiones de Italia.	94%, 96% y 90% de precisión en la autenticación de EVOO y de VOO de acuerdo a su origen geográfico, altitud y variedad, respectivamente.	[13]

Tabla 1. Continuación

Muestras	Número de muestras	Técnica	Análisis de datos	Objetivos	Resultados	Referencias
VOO	41	FFFS	PCA, FDA, CCSWA y PLSR	<p>Autenticidad de cuatro variedades distintas de aceitunas de Italia.</p> <p>Autenticidad según origen geográfico de EVOO de cinco regiones de Marruecos.</p> <p>Autenticidad de EVOO de cuatro variedades distintas.</p> <p>Predicción de FA, PV, K_{232} y K_{270} y del nivel de clorofilas.</p>	<p>Discriminación de muestras de VOO de acuerdo a su origen geográfico y la variedad con precisión de 96,72 y 95,12 %, respectivamente.</p> <p>Predicción de valores de FA, PV, K_{232}, K_{270} y nivel de clorofilas con coeficientes R^2, no inferiores a 0,88.</p>	[14]
EVOO	36	FFFS	PARAFAC	Control del riego en aceite de oliva.	Buena discriminación entre el modo de riego con coeficientes r entre 0,85 y 0,89.	[15]

Tabla 1. Continuación

Muestras	Número de muestras	Técnica	Análisis de datos	Objetivos	Resultados	Referencias
EVOO, VOO y POO	12 EVOO, 12 VOO y 12 POO	SFS	SPA y LDA	Clasificación de diferentes clases de aceites de oliva.	Niveles de discriminación entre muestras con una precisión del 99%.	[16]
EVOO	99	SFS	PLSR	Detección del grado de adulteración de EVOO con ROO con niveles de entre el 2 y el 20%.	Detección del nivel de adulteración con correlaciones de R ² de 0,94.	[17]
EVOO	63	EEFS	PLSR y PCR	Detección de adulteraciones de EVOO con ROO.	Detección de adulteraciones con una precisión marcada por el R ² de 0,98 tanto para PLSR como para PCR.	[18]

CCSWA: *Common Components and Specific Weights Analysis* (Componentes comunes y análisis de peso específico); DN-PLSR: *multiway discriminant partial least squares* (Discriminación multivía por regresión por mínimos cuadrados parciales); EEFS: *excitation-emission fluorescence spectroscopy* (Espectroscopía de fluorescencia de excitación-emisión); EVOO: *Extra virgin olive oil* (Aceite de oliva virgen extra); FA: *fatty acids* (ácidos grasos); FDA: *Factorial discriminant analysis* (Análisis factorial discriminante); FFFS: *Front-face fluorescence spectroscopy* (Espectroscopía front-face); k-NN: *K nearest neighbours* (K-vecinos más cercanos); LDA: *Linear discriminant analysis* (Análisis lineal discriminante); OO: *Olive oil* (Aceite de oliva); PARAFAC: *Parallel factors analysis* (Análisis de factores paralelos); PCA: *Principal component analysis* (Análisis por componentes principales); PLSR: *Partial least squares regression* (Regresión por mínimos cuadrados parciales); POO: *Pomace olive oil* (Aceite de orujo de oliva); PV: *peroxid value* (valor de peróxidos) SFS: *Synchronous fluorescence spectroscopy* (Espectroscopía de fluorescencia sincrónica); SPA: *Successive projections algorithm* (Algoritmo de proyecciones sucesivas); SVM: *supporting vector machine* (Máquinas de vector de soporte); VOO: *Virgin olive oil* (Aceite de oliva virgen).

Una vez revisadas las tendencias actuales del desarrollo de la fluorescencia molecular en sus múltiples variantes, se puede destacar el uso de matrices de excitación-emisión por fluorescencia front-face (FFFS) o la espectroscopía de fluorescencia sincrónica (SFS), en combinación con diferentes algoritmos, tanto de aprendizaje supervisado como no supervisado, con fines tanto exploratorios, como de discriminación y de regresión.

Teniendo en cuenta estos antecedentes, el objetivo de esta investigación ha sido el desarrollo de métodos no invasivos para el control del deterioro oxidativo de los aceites de oliva vírgenes por acción de factores que afectan a su estabilidad y calidad. Estos factores han sido la exposición a la radiación electromagnética y a la luz solar, tratamientos térmicos prolongados a diferentes temperaturas, así como el desarrollo de métodos de detección y cuantificación de adulteraciones. En todos los casos se utiliza la técnica de fluorescencia front-face con muestras intactas y se aprovecha la fluorescencia nativa de los compuestos presentes en los aceites de oliva como son los compuestos fenólicos, tocoferoles o las clorofilas.

Las herramientas quimiométricas empleadas con estos fines han sido herramientas de análisis no supervisado como PARAFAC, con fines exploratorios y de identificación de los analitos objeto de estudio, herramientas de clasificación como PARAFAC supervisado por análisis lineal discriminante (PARAFAC-LDA), análisis discriminante acoplado a mínimos cuadrados parciales desdoblados (del inglés *discriminant unfolded partial least squares*) (DA-UPLS), o U-PLS con fines de cuantificación sobre las EEMs de aceites adulterados.

Los modelos desarrollados permitieron establecer discriminaciones entre aceites de oliva virgen extra, sometidos y no sometidos a radiación solar, así como detectar cambios en el perfil fluorescente de los compuestos

estudiados como consecuencia de distintos tratamientos térmicos realizados sobre aceites de oliva virgen extra y detectar adulteraciones de EVOO con POO y OO con altos niveles de precisión, llegándose a desarrollar modelos de regresión para detectar el grado de adulteración.

2. METODOLOGÍA

2.1. Muestras

Las muestras empleadas han sido aceites de oliva virgen extra, obtenidas de mercados locales de la ciudad de Badajoz de distintas marcas comerciales y procedentes de distintas variedades (*hojiblanca*, *manzanilla cacereña*, *cornicabra* y *picual*) durante la campaña 2016. En todos ellos los parámetros de calidad se corresponden con los establecidos por la norma aplicable a los EVOO (grado de acidez libre entre 0,4 y 0,6°). Los OO y los POO también se adquirieron en diferentes locales, teniendo en cuenta los criterios de calidad y composición de estos.

2.2. Instrumentación

Las EEMs se obtuvieron con un espectrofotómetro de fluorescencia modelo Varian Cary Eclipse, equipado con dos monocromadores, una lámpara de xenón, y un tubo fotomultiplicador como detector, conectado a un PC con una interfaz IEE 488 (GPIB). Para la adquisición de datos se usó el software Cary Eclipse 1.0 y las medidas fueron realizadas en modo *front-face*, lo cual minimiza el fenómeno de dispersión de la luz reflejada y el efecto de filtro interno, permitiendo así trabajar con muestras líquidas sin ningún tratamiento previo. El ángulo de incidencia, definido como el ángulo entre el rayo de luz de excitación y la perpendicular a la superficie de la celda, fue de 34°. Las medidas de fluorescencia fueron tomadas en celdas de cuarzo de 10 mm a temperatura ambiente. Las rendijas de excitación y emisión de los monocromadores fueron de 5 nm cada una. La

sensibilidad del tubo fotomultiplicador fue de 550 V y la velocidad de registro del instrumento fue de 300 nm·min⁻¹.

Para estudiar la influencia que la radiación UV tiene sobre los principales fluoróforos de los aceites de oliva y utilizarlos como marcadores de dicha exposición del deterioro en su calidad, se empleó una lámpara de mercurio de alta presión Osram de 200-W modelo Oriel 8500 (Spectra-Physics, Newport, USA). El dispositivo consta de una caja negra con un ventilador, una lámpara de mercurio y una lente de cuarzo. El proceso de irradiación se realizó mediante una celda de cuarzo con tapón de plástico colocada en un receptáculo a 30 cm de la lámpara de mercurio. El aceite de oliva situado en una celda de cuarzo era agitado magnéticamente durante la irradiación, siendo cada muestra irradiada a diferentes tiempos de 1, 5, 15, 35, 70, 120 y 180 min. Posteriormente, y para poder llevar a cabo una extrapolación más adecuada a las condiciones reales de almacenamiento, se hicieron pruebas análogas irradiando con luz solar natural en las mismas celdas de cuarzo tapadas y agitadas magnéticamente durante tiempos de 30, 60 y 120 minutos.

Para llevar a cabo el estudio de la influencia de la temperatura sobre el perfil fluorescente de los aceites de oliva virgen extra, se empleó un baño de arena termostático P-Selecta de 220W modelo 2.1.1. Las muestras de aceite se disponían en cápsulas de porcelana y se calentaban a 40, 60 y 80 °C durante 30, 60 y 120 minutos.

2.3. Selección de los parámetros óptimos para la adquisición de matrices de excitación-emisión (EEMs)

Para obtener la máxima información analítica de las EEMs, se optimizaron los parámetros instrumentales. Para ello, se ha tenido en cuenta que los datos obtenidos van a ser tratados con diferentes algoritmos matemáticos con fines de clasificación y cuantificación. En este sentido es

importante tener en cuenta la capacidad de procesamiento del equipo de fluorescencia y el tiempo que lleva registrar las EEMs en un amplio intervalo de longitudes de onda. De este modo, por los criterios de compromiso entre un tiempo óptimo de análisis y la recogida del mayor volumen de datos posibles, sin verse alterados por fenómenos instrumentales como la dispersión Rayleigh, se optó por realizar dos registros independientes que permitían obtener información de especies químicas fluoróforas muy alejadas en términos espectrales. Estas dos regiones espectrales abarcaban, por un lado, los tocoferoles y compuestos fenólicos, los cuales presentan longitudes de onda de excitación en zonas inferiores a 350 nm y, por otro lado, los pigmentos como las clorofilas cuya excitación se produce a longitudes de onda superiores a los 350 nm y la emisión a longitudes de onda superiores a 600 nm.

Por tanto, los rangos de longitudes de onda de excitación y emisión utilizados para el registro de las EEMs en cada una de estas regiones fueron las siguientes:

- **Rango de los tocoferoles y compuestos fenólicos:** como se ha comentado anteriormente en referencias previas [19], y como se refleja en la **Tabla 4** de la introducción de esta memoria, los aceites de oliva presentan fluorescencia en una zona de longitudes de onda cortas en las que se encuentran los espectros de fluorescencia de la vitamina E y de los polifenoles. Por tanto, el rango de medida empleado fue de 270 a 350 nm, a intervalos de 5 nm, para la excitación, y de 284 a 500 nm, con intervalos de 0,5 nm, para la emisión, para el estudio de la evolución de estos compuestos.
- **Rango de las clorofilas y derivados:** la fluorescencia de estos pigmentos se asocia a longitudes de onda largas tal y como se indicó en la **Tabla 4** de la introducción de esta tesis. Para la visualización de estos

compuestos se ha seleccionado un rango de longitudes de onda de excitación entre 360 y 500 nm, registrando los espectros de emisión a intervalos de 5 nm en un rango de longitudes de onda entre 640 y 700 nm, a intervalos de 0,5 nm.

2.4. Diseño experimental

Además de la optimización de las longitudes de onda de excitación y emisión, otro de los aspectos fundamentales en el desarrollo de metodologías basadas en el uso de métodos quimiométricos, es un diseño experimental óptimo para obtener una información representativa de todas las muestras analizadas. De este modo, y dentro de los estudios diferenciados objeto de esta memoria, como son el estudio de la influencia de la irradiación UV, de la irradiación con luz solar y la influencia del tiempo de calentamiento, en las **Tablas 1.1** y **1.2** se resumen los diseños experimentales utilizados.

Tabla 1.1 Diseño experimental propuesto para el estudio de la evolución de las EEMs de aceites de oliva sometidas a irradiación UV e irradiación con luz solar.

Set de muestras de aceites sometidos a diversos tiempos de irradiación					
Lámpara UV			Irradiación Solar		
Set	Muestras	Tiempo (min)	Set	Muestras	Tiempo (min)
1	20	1	1	20	30
2	20	5	2	20	60
3	20	15	3	20	120
4	20	35			
5	20	70			

Tabla 1.2. Diseño experimental propuesto para el estudio de la evolución de las EEMs de aceites de oliva sometidas a diferentes temperaturas y tiempos de calentamiento.

Set	Muestras	Temperatura (°C)	Tiempo (min)
1	16		30
2	16	40	60
3	16		80
4	16		30
5	16	60	60
6	16		80
7	16		30
8	16	80	60
9	16		80

2.5. Herramientas quimiométricas

Para la realización de los cálculos correspondientes al desarrollo de los métodos quimiométricos de segundo orden se empleó el entorno de MatLab (MATLAB R2008a versión 7.6.0.324) [20]. Los cálculos realizados con el algoritmo PARAFAC, disponible en internet gracias a Bro. R. [21], se llevaron a cabo a través de la interfaz gráfica MVC2, desarrollada previamente por Olivieri et al. en 2009 y 2014 [22, 23] (<http://www.iquir-conicet.gov.ar/esp/div5.php?area=12>), la cual permite también utilizar U-PLS y cargar de manera simple las EEMs en el entorno de Matlab y analizarlas según los parámetros propios de cada algoritmo. El análisis lineal discriminante supervisado sobre los *scores* de PARAFAC, se realizó a través de la rutina programada *LDA_plot*, desarrollada en base al algoritmo de Kemsley. E. [24].

Por su parte, el análisis discriminante por mínimos cuadrados parciales fue realizado a través de la rutina titulada *DUPLS_plot*, las

representaciones gráficas de los scores de PARAFAC y los vectores canónicos de LDA a través de las rutinas *plot_3D* y *plot_2D*, y la discriminación entre grupos asignados fue realizada mediante las rutinas *elipses_3D* y *elipses_2D*, a través de las elipses de contorno cuya distancia viene calculada a través de las leyes de la estadística inferencial previamente descritas en la introducción de esta tesis.

El procedimiento empleado para obtener los resultados gráficos y numéricos a través de la interfaz de las rutinas empleadas en Matlab fue el siguiente:

1. Exportar los datos de las EEMs recopiladas de cada prueba procedentes del software Cary Eclipse 1.0 en formato CSV.
2. Trasladar los datos al ordenador personal a documento de texto en formato MS-DOS.
3. Abrir el entorno de MatLab R2008a y aplicar la rutina *noray.m* para eliminar la dispersión Rayleigh de las EEMs con dicho efecto.
4. Construir los conjuntos de datos según los diseños experimentales previamente propuestos y generar otro fichero de texto (titulado por defecto "*calfiles.txt*") con los nombres asignados a las muestras correspondientes a cada conjunto para su tratamiento en la interfaz de la rutina MVC2 de MatLab, cuya interfaz se puede visualizar en la *Figura 1.1*.



Figura 1.1. Interfaz de la rutina MVC2 de Matlab desarrollada para el cálculo de datos de segundo orden.

5. Se realiza el cálculo asignando previamente el número de componentes y las restricciones como el criterio de no negatividad y se genera una nueva carpeta, titulada *temp 2*, donde se recogen los archivos que proporcionan las matrices desdobladas de PARAFAC, tituladas A_PARAFAC, B_PARAFAC y C_PARAFAC. Para la determinación del número de componentes se representan estos frente al valor del *core consistency* [25] (apartado 5.2 de la introducción) y la evolución de los residuales.
6. Las matrices de *scores* (tituladas por defecto A-PARAFAC) se emplean en la rutina *plot_3D* (en caso de disponer de más de dos componentes) para su representación tridimensional, previa asignación en la propia rutina de las categorías que se estén estudiando a través de la asignación del orden establecido dentro del fichero *calfiles*, ya sean tipos de aceite, grado de adulteración, alteraciones físicas, etc. Su interfaz en el entorno de MatLab puede visualizarse en la **Figura 1.2**.

```

1 - figure
2 - hold on
3 - grid on
4 - box on
5 - %N=numero de medidas de la clase
6 - N=10
7 - %M=numero total de medidas
8 - M=20
9 - for i=1:N
10 -     formato(i) = 'sr';
11 - end
12
13 - for i=(N+1):M
14 -     formato(i) = 'ob';
15 - end

```

Figura 1.2. Interfaz de la rutina *Plot_3D*.

7. Las matrices de *loadings* (tituladas por defecto B_PARAFAC y C_PARAFAC), previamente eliminados los datos de la primera fila repetidos, ya que se emplean como iniciación del algoritmo, se recogen y se representan en función de las longitudes de onda de excitación y emisión para obtener los perfiles fluorescentes de los componentes obtenidos de manera matemática a través del desdoblamiento de las EEMs mediante PARAFAC.
8. Cuando se pretende hacer un análisis discriminante de las categorías previamente seleccionadas, se genera un fichero donde se asigna un código binario de valores, 1 y 2, atribuibles a las categorías seleccionadas para agruparlas. Posteriormente se carga dicho fichero (llamado por defecto *ycal.txt*), junto con el fichero *calfiles.txt* en la interfaz de la rutina *LDA_plot* con los componentes previamente seleccionados y esta te devuelve un fichero con la matriz de vectores canónicos de LDA, titulado por defecto 't.txt' para su representación del mismo modo que los 'scores' de PARAFAC. El procedimiento gráfico dentro de la interfaz de Matlab viene descrito en la **Figura 1.3**.

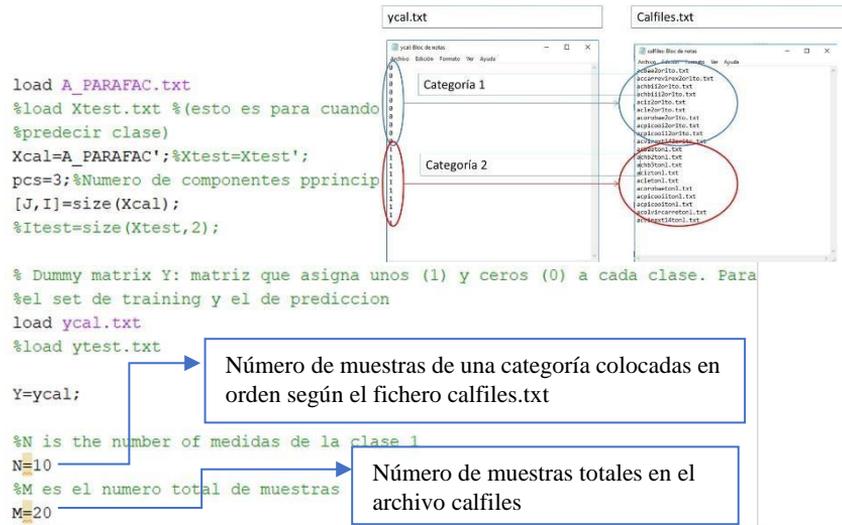


Figura 1.3. Interfaz de la rutina LDA-plot antes de ejecutarla para obtener los vectores canónicos a partir de la matriz A-PARAFAC.

9. Finalmente, para el establecimiento de un análisis confirmatorio, la rutina *elipses_3D* permite establecer un criterio objetivo y sencillo mediante la representación tridimensional de elipses en las tres dimensiones del espacio ocupado por los puntos asignados a las muestras de cada categoría. Para ello, el proceso a seguir es similar al llevado a cabo para la obtención de los vectores canónicos de LDA. En primer lugar, se abre la rutina *elipses-3D* en el entorno de Matlab y se cargan las matrices *A-PARAFAC.txt* y *t.txt* en la ventana de comandos para poder ser interpretadas por la rutina. Posteriormente, dentro de esta, se vuelve a elegir el número **N** de muestras que la rutina debe asignar a la primera categoría y el número **M** de muestras totales para poder asignar el resto a la segunda categoría, se programa el intervalo de confianza dentro de la tabla *chi cuadrado* y se lanza la rutina. La **Figura 1.4** muestra los pasos del proceso dentro de la interfaz.

Se cargan las matrices `A_PARAFAC.TXT` y `t.txt` en la ventana de comandos de manera análoga a como se explicó en la representación de los scores de PARAFAC mediante plot 3D

Cargamos la rutina en la carpeta `temp2`

Programar la rutina

Elegimos el número de muestras de la primera clase en `N` y el número total en `M`

Buscamos en la tabla chi cuadrado valor par el intervalo de confianza que queremos aplicar a la elipse

Elegimos el color y la forma que le queremos dar a los puntos y el color de la elipse. `r=red` `b=blue` `s=square` `o=circle`

Figura 1.4. Programación de la representación de las elipses en la Interfaz de la rutina `elipses-3D`.

10. Por último, para el análisis discriminante a través de U-PLS mediante la rutina *DUPLS_plot*, es preciso generar previamente cuatro ficheros, dos con las matrices de datos ordenadas según las dos categorías a asignar, el primero llamado por defecto *calfiles* con un número mayor de muestras que el segundo, denominado *valfiles*, y otros dos archivos donde cada fila represente la categoría para cada archivo. Es decir, un archivo titulado por defecto *ycal* que tenga el mismo número de filas que el archivo *calfiles*, y otro archivo titulado *yval* con el mismo número de filas que el archivo *valfiles*, en los que se establezca mediante asignación binaria (1 y 2) la categoría a la que pertenece la matriz asignada a cada fila del archivo, y una vez cargados dichos archivos se les aplica U-PLS en la interfaz de MVC2 para obtener los valores de concentración atribuidos a cada matriz, montando posteriormente con los valores predichos los ficheros *predcal* y *predval*. Estos ficheros se cargan sobre la rutina *DUPLS_Plot*. La interfaz de esta rutina consiste en programar el intervalo de confianza para la tabla *t* de *student*, para el nivel de significación y los grados de libertad correspondientes, y cargar las matrices obtenidas de la aplicación de U-PLS a la predicción de la categoría establecida por el algoritmo, según los datos de fluorescencia para obtener el gráfico de error de predicción que representa el valor nominal frente al valor predicho con sus franjas de error delimitadas por el valor del intervalo de confianza asignado. La interfaz de trabajo en Matlab viene representada en la **Figura 1.5**.

```

%Rutina para hacer figura DPLS
clear all, clc, close all

load predcal.txt
load predval.txt

%Estos archivos contienen cuatro columnas en la primera hay 1, en la
%tercera 2, y en la segunda y cuarta las predicciones para cada clase
t= 2.31; %poner el valor con n-1 grados de libertad y 0.05

s1=std(predcal(:,2));
x1=mean(predcal(:,2));
l1i=x1-t*s1;l1s=x1+t*s1;

s2=std(predcal(:,4));
x2=mean(predcal(:,4));
l2i=x2-t*s2;l2s=x2+t*s2;

figure(1), plot(predcal(:,1),predcal(:,2),'ob')

```

Matrices de calibración y validación a introducir para la representación del gráfico de predicción por DA-UPLS

Valor previsto para la tabla t-student según el nivel de significación y los grados n-1 grados de libertad

Figura 1.5. Programación de la representación del gráfico de predicción frente al valor nominal con su franja de significación para el contraste de hipótesis en la interfaz **DUPLS_plot**.

Una vez establecidas las condiciones de partida para el desarrollo de las investigaciones, en los siguientes apartados se comentarán los resultados obtenidos de los análisis descritos en este apartado, obteniendo las conclusiones pertinentes en lo referente al uso de la fluorescencia en combinación con métodos quimiométricos establecidos para el control de calidad y detección de fraudes en aceites de oliva.

3. RESULTADOS Y DISCUSIÓN

3.1. Caracterización de las matrices de fluorescencia de aceites de oliva según su denominación de calidad

3.1.1. EEMs de los distintos tipos de aceites de oliva

La primera fase para el diseño de los estudios propuestos para el seguimiento de la calidad y el control de fraudes en aceites de oliva fue la caracterización de los perfiles de excitación y emisión de los aceites de oliva virgen extra (EVOO), aceites de oliva (OO) y aceites de orujo de oliva (POO).

En la **Figura 1.6**, se representan las EEMs obtenidas para tres aceites de oliva, perteneciente a cada una de las categorías (EVOO, OO y POO), y para los dos intervalos de longitudes de onda seleccionadas.

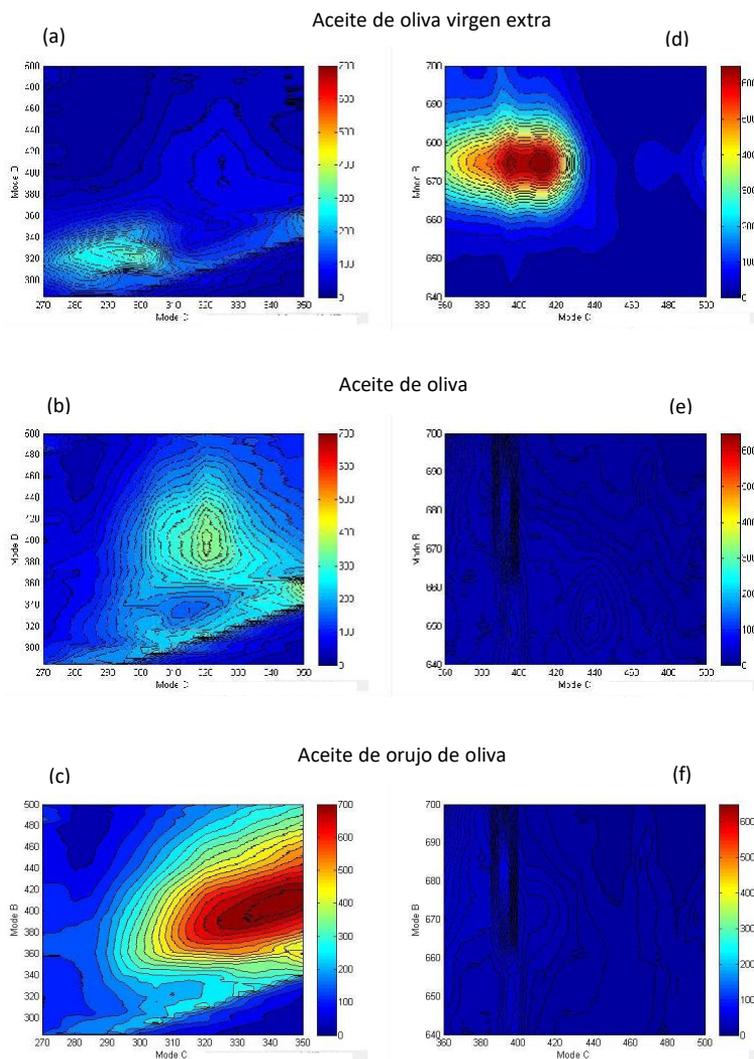


Figura 1.6. Matrices de fluorescencia de aceites de oliva de diferentes categorías, en el rango de longitudes de onda de los tocoferoles y polifenoles (a, b, c) (excitación (mode C) = 270-350 nm y emisión (mode B) = 280-500 nm), y en la región de los pigmentos (d, e, f) (excitación (mode C) = 360-500 nm y emisión (mode B) = 640-700 nm).

Tal como se aprecia, existen grandes diferencias en el perfil fluorescente de los diferentes tipos de aceites de oliva. En el caso de los EVOO, se puede observar que, en el rango de longitudes de onda de los tocoferoles y polifenoles, la zona de mayor intensidad de fluorescencia se encuentra en un intervalo de longitudes de onda de excitación entre 280 y 310 nm y longitudes de onda de emisión entre 270 y 360 nm (**Figura 1.6a**). Esta zona se relaciona con la presencia de α -tocoferol y de compuestos fenólicos y resulta de suma utilidad para la caracterización y diferenciación de estos tipos de aceites, dada su sensibilidad y la relación que existe entre la presencia de los compuestos detectados y la calidad de los diferentes tipos de aceites de oliva.

En el caso del perfil fluorescente del OO, obtenido mezclando aceite de oliva virgen con aceites de oliva refinados (**Figura 1.6b**), se observa que aparece una amplia banda de fluorescencia con un máximo a una longitud de onda de excitación próxima a los 300 nm y con un máximo de emisión situado a 400 nm, longitudes de onda superiores a las atribuibles a los tocoferoles y polifenoles. Esta banda de fluorescencia se puede atribuir a la presencia de productos de oxidación producidos durante el proceso de refinado del aceite, como consecuencia de las altas temperaturas alcanzadas durante la decoloración (100 °C aproximadamente) y la desodorización (200-250 °C aproximadamente). Principalmente pueden ser dienos y trienos conjugados, productos de hidrólisis, ácidos grasos oxidados productos de la peroxidación lipídica, etc. [19, 26]. Esta banda resulta de gran utilidad para la diferenciación y caracterización de aceites mediante métodos quimiométricos.

En el caso de los POO (**Figura 1.6c**) las matrices obtenidas en el rango de los tocoferoles y polifenoles son muy similares a la de los aceites refinados, puesto que son sometidos a los mismos procesos, que incluyen los procesos de desodorización y de blanqueo, los cuales producen la

pérdida de gran parte de polifenoles y tocoferoles, que disminuyen la estabilidad oxidativa del aceite.

En cuanto a las EEMs obtenidas en los rangos de longitudes de onda de los pigmentos (clorofila *a*, clorofila *b*, feofitina *a*, feofitina *b*, etc.), se observa en los aceites de oliva virgen extra una zona de elevada intensidad de fluorescencia correspondiente a longitudes de onda de excitación comprendidas entre 400 y 420 nm y una longitud de onda de emisión próxima a 675 nm (**Figura 1.6d**). Sin embargo, en los aceites sometidos a procesos de refinado como el aceite de oliva y el aceite de orujo de oliva (**Figuras 1.6e y 1.6f**) esta señal es muy débil ya que los pigmentos se encuentran muy degradados debido a los procesos térmicos a los que son sometidos. Como consecuencia de esto, se produce una pérdida de la capacidad antioxidante, puesto que los pigmentos clorofílicos del aceite de oliva (clorofilas *a* y *b*, y sus derivados feofitina *a* y *b*) han mostrado tener una gran capacidad antioxidante en aceites almacenados en la oscuridad y a bajas temperaturas [27]. Estas diferencias pueden resultar una herramienta útil para diferenciar estos tipos de aceites de aceites vírgenes extra no alterados.

El contenido en clorofilas también es un indicador de la calidad del aceite de oliva, ya que influye en la prevención de su autooxidación. No obstante, este contenido está sujeto a otros factores endógenos y exógenos que hacen que su presencia esté sujeta a una mayor variabilidad, entre los que se encuentran:

- Efectos del procesamiento (rotura de las paredes celulares, aumento de la temperatura), senescencia de los tejidos vegetales, aumento de la actividad enzimática y aceleración de la degradación de las moléculas de clorofila *a* y *b* a feofitina, la cual es una molécula fotodinámica que en presencia de luz y oxígeno rompe la estructura de porfirina por la

rotura del enlace entre dos anillos pirrolidínicos transformándose en una molécula sin propiedades fluorescentes [28].

- Momento de la temporada de recolección de las aceitunas, puesto que al recogerse de forma temprana las aceitunas se encontrarán en menor estado de maduración y poseen mayor cantidad de clorofilas.
- La incidencia de radiación durante la producción y el almacenamiento. Esta variabilidad intrínseca dentro del rango de las clorofilas puede afectar a la caracterización del tipo de aceite, pero es importante tenerla en cuenta para el control de su calidad en términos de almacenamiento y conservación.

3.1.2. Clasificación de muestras de aceite de oliva según su categoría

Una vez vistas las diferencias existentes en los perfiles de fluorescencia de aceites de oliva de diferentes categorías, se procedió a estudiar la capacidad de los algoritmos quimiométricos para clasificar y diferenciar entre EVOO y aceites de oliva refinados (ROO, del inglés *refined olive oil*).

Como ya se ha mencionado, una de las ventajas de PARAFAC es la posibilidad de obtener los datos cualitativos de los propios componentes de la muestra analizada. Para ello, en primer lugar, una vez diseñado el modelo experimental, el siguiente paso será la determinación del número de componentes que se ajuste mejor a los criterios del ‘*core consistency*’ [29] y de la mínima variación en los errores residuales, herramientas proporcionadas por el propio entorno de MatLab.

Inicialmente, con el objetivo de evaluar la capacidad de las EEMs para caracterizar aceites de oliva de diferentes categorías de calidad, a través de sus perfiles de fluorescencia, se modelizaron a través de PARAFAC seis conjuntos de datos correspondientes a los aceites de oliva

virgen extra (EVOO), correspondientes a las dos zonas de longitudes de onda estudiadas.

Para el caso del EVOO, los conjuntos de datos presentaron unas dimensiones de [16 x 433 x 17] y de [16 x 121 x 29] para el rango de fluorescencia correspondiente a la región de los tocoferoles, polifenoles y productos de oxidación y de clorofilas, respectivamente. En el caso de los OO las dimensiones fueron [10 x 433 x 17] y [10 x 121 x 29] para los mismos rangos seleccionados y, por último, los conjunto de datos de los POO presentaban unas dimensiones de [4 x 433 x 17] y [4 x 121 x 29]. En todos los casos, el primer término se corresponde con el número de muestras, el segundo término el número de datos correspondientes a las longitudes de onda de emisión y el tercer término el número de datos de longitudes de onda de excitación.

A continuación, se llevó a cabo el análisis utilizando PARAFAC con todas las muestras, representando los *scores* en dos categorías, el conjunto de EVOO y el conjunto de aceites refinados (no EVOO), analizando los datos también los dos tangos de longitudes de onda estudiados. Por último, estos *scores* del modelo PARAFAC son utilizados para llevar a cabo un análisis clasificatorio mediante análisis lineal discriminante.

En las siguientes secciones se muestran los resultados obtenidos para dichos análisis.

A. Análisis de las EEMs mediante PARAFAC para la caracterización de los diferentes aceites

En primer lugar, se estudió el rango espectral de los tocoferoles, compuestos fenólicos y productos de oxidación de cada uno de los tres grupos de aceites mencionados anteriormente ($\lambda_{ex} = 270-350 \text{ nm}$, $\lambda_{em} = 280-500 \text{ nm}$). Para ello, se aplicó el algoritmo PARAFAC sobre cada uno de los conjuntos de muestras. Se tuvo en cuenta la restricción de no-

negatividad, ya que las concentraciones y las señales de fluorescencia son siempre positivas y posteriormente se obtuvieron los perfiles de fluorescencia de los *'loadings'* de PARAFAC y los correspondientes *scores*.

El primer paso fue seleccionar el número de componentes óptimos aplicando los criterios del *core consistency* y de la desviación estándar de los residuales tanto para el conjunto de los EVOO, como de los POO y de los OO tal y como se puede observar en la **Figura 1.7**. En esta figura se muestra la variación del *core consistency* y de los residuales en función del número de componentes para el conjunto de 16 muestras de EVOO (A, B), el de 4 muestras de POO (C, D) y el de 10 muestras de OO (E, F) analizadas.

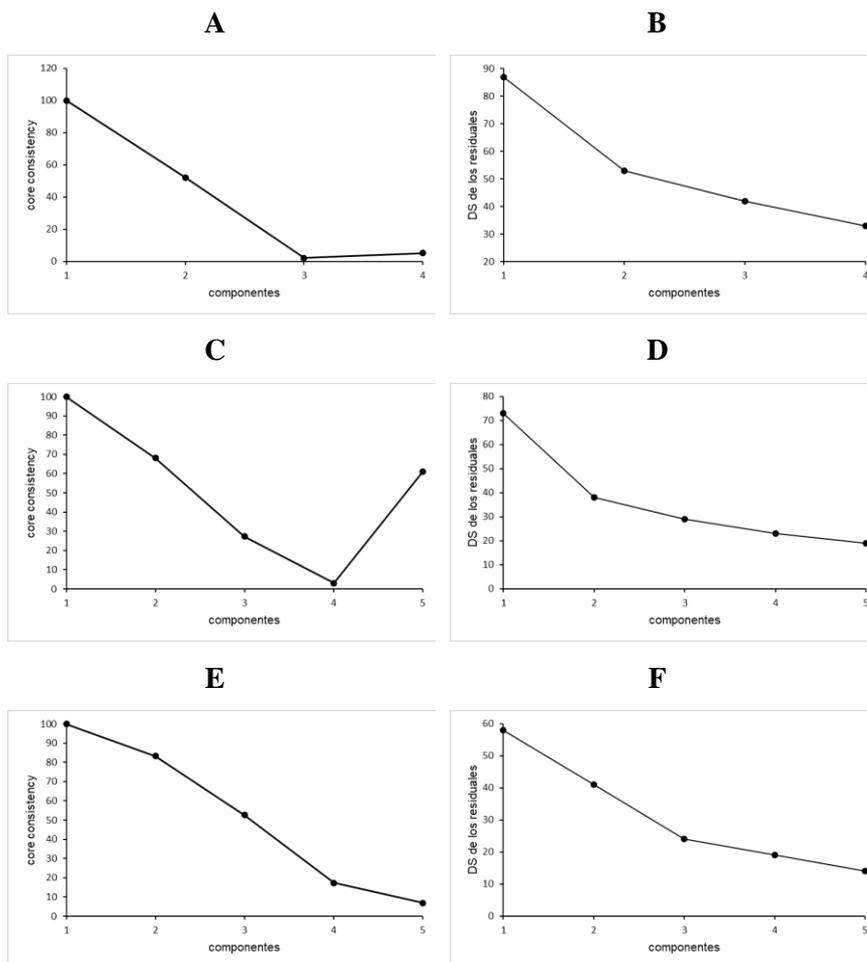


Figura 1.7. Representación del valor del core consistency (A, C y E) y de la desviación estándar de los residuales (B, D y F) frente al número de componentes, para el conjunto de 16 muestras de EVOO (A y B), 4 muestras de POO (C y D) y 10 muestras de OO (E y F), respectivamente. Región espectral: $\lambda_{ex} = 270-350 \text{ nm}$, $\lambda_{em} = 280-500 \text{ nm}$.

De estos resultados se puede observar que, para los EVOO (A) y POO (C), el número de componentes óptimo sería de 2, ya que el valor del core consistency para ambos disminuye de 50 a partir del tercer

componente, y el valor de la desviación estándar de los residuales se estabiliza a partir del segundo componente, mientras que para las muestras de OO (E), ambos criterios concluyen que el número óptimo de componentes en este caso sería de tres, lo cual establece una diferencia con respecto a los POO que experimentan un proceso de refinado semejante.

Una vez seleccionado el número de componentes, el algoritmo PARAFAC proporciona los *loadings* en cada una de las dimensiones (excitación y emisión), de los componentes seleccionados, pudiendo correlacionarse estos con los posibles componentes o con familias de compuestos, que contribuyen a la fluorescencia nativa de las muestras. Estos perfiles pueden corresponderse con los ya mencionados tocoferoles y compuestos fenólicos, presentes en los aceites de buena calidad, así como los compuestos derivados de la oxidación que afectan a los atributos de calidad y funcionales, tal y como se corrobora en varios estudios [30]–[35]. Estos estudios muestran EEMs en rangos de longitudes de onda similares a los de esta tesis confirmándose que la aparición de estas regiones espectrales se debe a productos de oxidación secundaria como son aldehídos cetonas o dienos conjugados.

En la **Figura 1.8** se pueden observar los perfiles de excitación y emisión de los dos componentes obtenidos de la descomposición de las matrices de los conjuntos de EVOO, POO y OO.

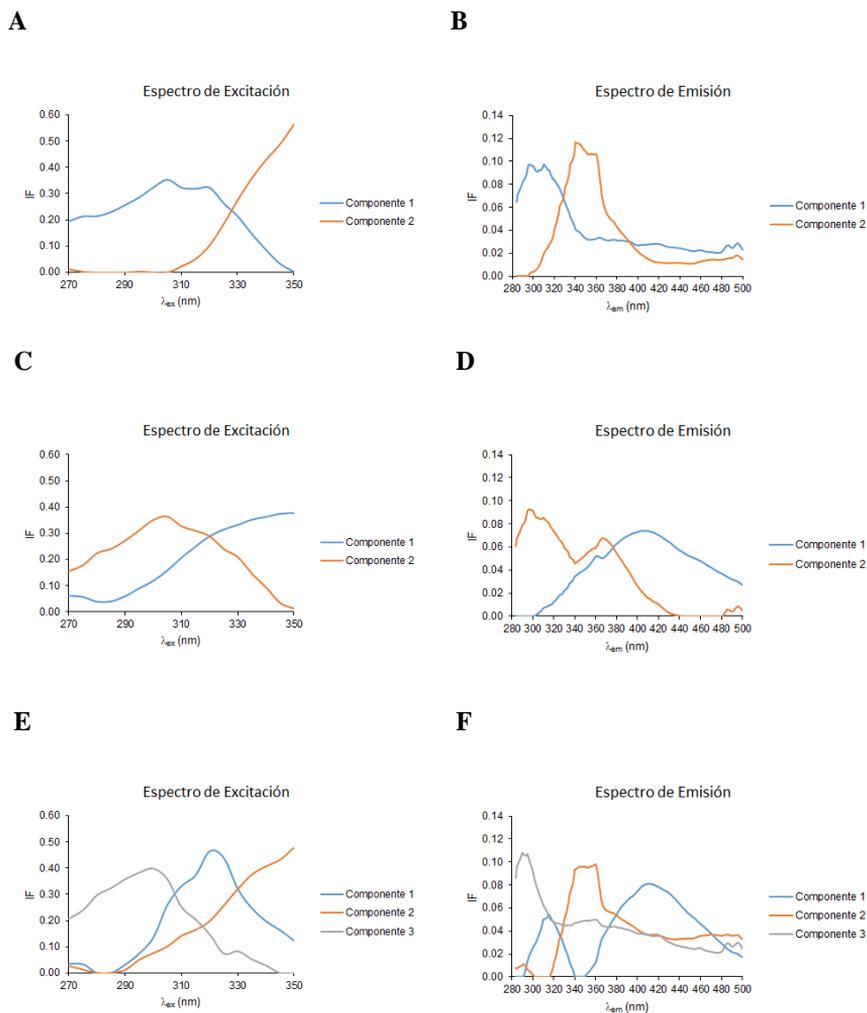


Figura 1.8. Representación de los loadings de excitación ($\lambda_{ex} = 270\text{-}350$ nm) de los EVOO (A), POO (C) y OO (E) y los loadings de emisión ($\lambda_{em} = 280\text{-}500$ nm) de los EVOO (B), los POO (D) y los OO (F).

Haciendo la comparativa de los perfiles de excitación y emisión de los tres tipos de aceites de oliva estudiados en esta investigación se puede observar que los EVOO presentan una fisionomía bastante diferente a la de los aceites refinados, ya sean POO o OO.

En el análisis pormenorizado de los perfiles de excitación se observa que, en los EVOO, el primer componente queda bien definido por una extensa banda que se extiende desde los 270 hasta los 350 nm, alcanzando dos máximos puntuales sobre los 300 y los 315 nm, mientras que el segundo componente no queda del todo definido dentro del rango de longitudes de onda estudiado. Por otro lado, para los POO, parece existir una inversión de los componentes, cobrando mayor relevancia la banda no definida de los EVOO, lo que hace indicar que es representativa de compuestos de oxidación, mientras que, en los OO, ocurre algo similar, ya que, a pesar de haber tres componentes, parece que el componente 2 de los POO se asemeja a la suma del segundo y tercer componente de los OO.

En los perfiles de emisión, las bandas quedan bien definidas en todos ellos, pudiendo observarse tendencias claras contrastables con los espectros puros de determinados compuestos consultados en la bibliografía. Este es el caso de los componentes 1 y 2 de los EVOO, que quedan definidos por dos bandas estrechas con sus máximos localizados sobre los 300 y los 350 nm, respectivamente. Dichos máximos son comparables a los registrados por Tena et al. en 2009 [19] y que se corresponden con el α -tocoferol y el ácido gálico, respectivamente, aunque este último no está del todo claro ya que hay multitud de compuestos fenólicos que presentan fluorescencia sobre esta zona del espectro. Por su parte, en el caso de los OO y los POO, el primer componente presenta una banda ancha cuyo máximo se centra sobre los 420 nm, lo cual coincide con las estimaciones realizadas por Kyriakidis [26] y que lo relaciona con compuestos procedentes de la oxidación de ácidos grasos, ya que registró esta banda del espectro de emisión para aceites sometidos a diferentes condiciones de oxidación, observando que se producía este crecimiento como consecuencia de la oxidación de ácidos grasos mono y poliinsaturados.

Por lo tanto, queda demostrado que el rango seleccionado es propicio para la caracterización de la calidad de los aceites de oliva de manera cualitativa, ya que da una información de especial relevancia respecto a las características composicionales de los distintos tipos de aceites sin necesidad de realizar procedimientos destructivos para su determinación.

Una vez estudiado el primer rango de longitudes de onda, se pasó a estudiar el rango espectral de las clorofilas ($\lambda_{ex} = 360-500$ nm, $\lambda_{em} = 640-700$ nm). Siguiendo el mismo procedimiento, se seleccionaron los componentes óptimos para cada grupo de aceites. En la **Figura 1.9** se observan las correspondientes representaciones.

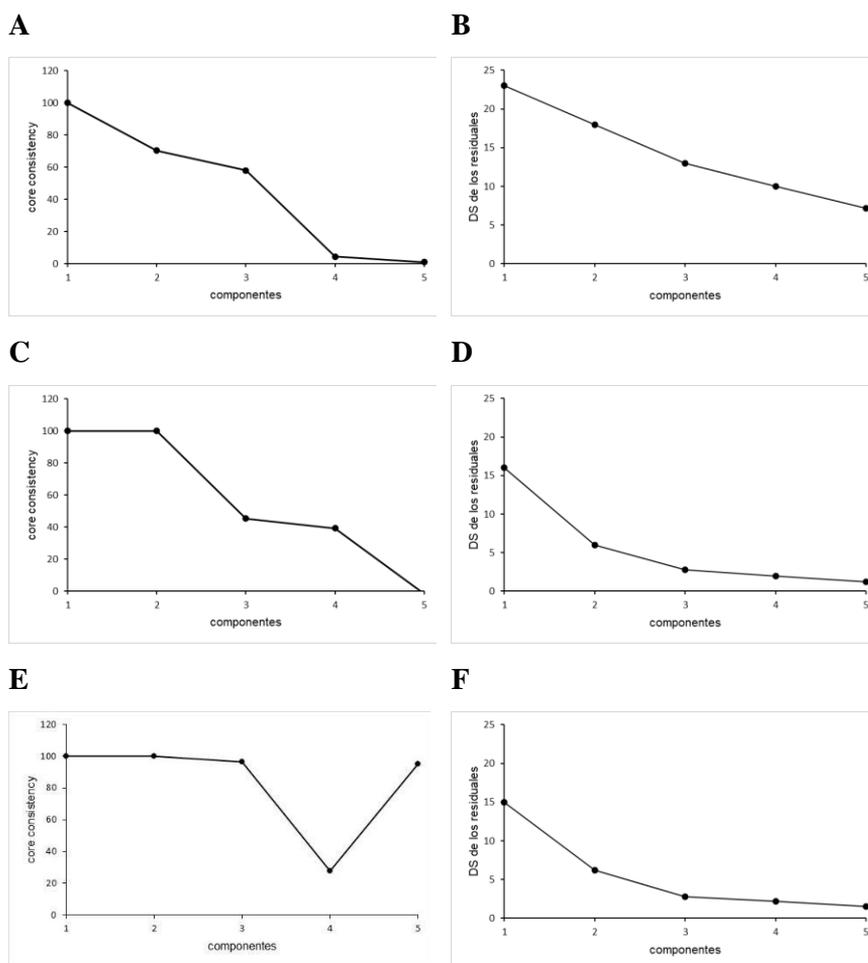


Figura 1.9. Representación del valor del core consistency (A, C y E) y de la desviación estándar de los residuales (B, D y F) frente al número de componentes, para el conjunto de 16 muestras de EVOO (A y B), 4 muestras de POO (C y D) y 10 muestras de OO (E y F), respectivamente. Región espectral: $\lambda_{ex} = 360-500 \text{ nm}$ y $\lambda_{em} = 640-700 \text{ nm}$.

Como se puede observar, en la **Figura 1.9**, el número de componentes adecuados a los criterios establecidos del ‘*core consistency*’ y de la desviación de los residuales es 3 para los EVOO (A y B), 2 para los POO (C y D) y 3 para los OO (E y F). En estos últimos, pese a que la estabilización de los residuales no se observa con tanta facilidad entre los componentes 2 y 3, sí se puede comprobar con claridad que el *core consistency* (E) baja de 50 para el cuarto componente, cosa que para los POO (C) no ocurre. De este modo, los perfiles obtenidos de dicha descomposición mediante PARAFAC puede observarse en la **Figura 1.10**.

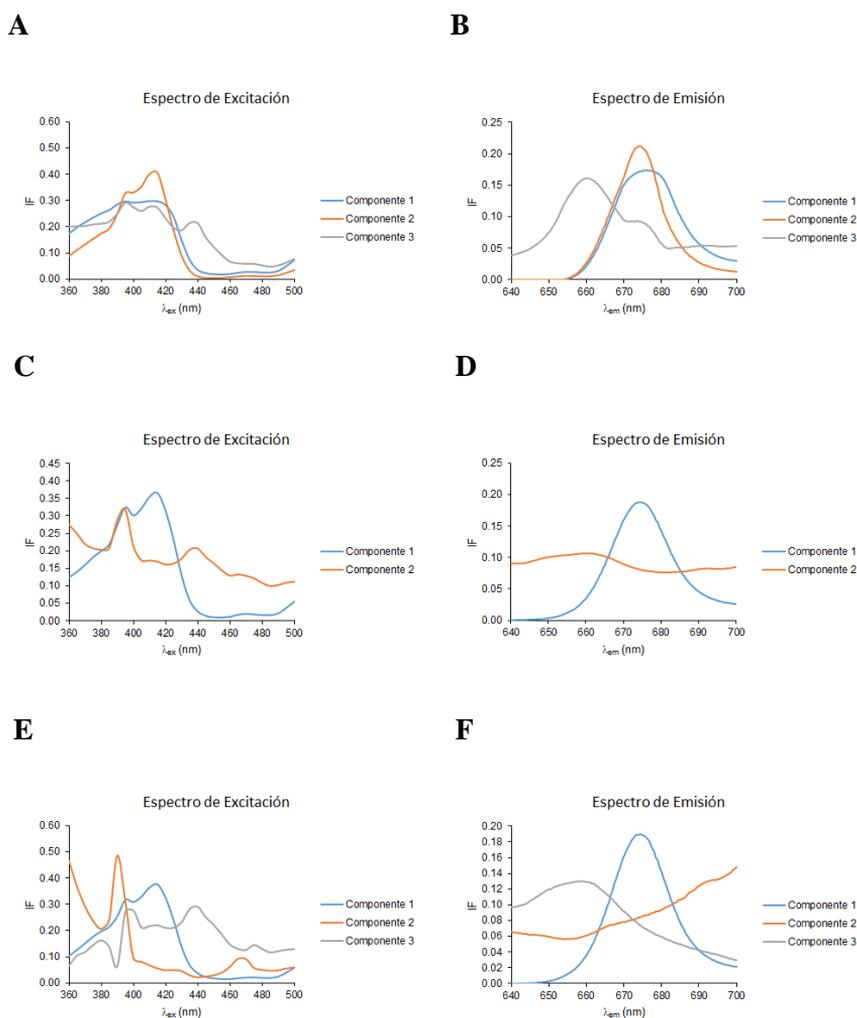


Figura 1.10. Representación de los loadings de excitación ($\lambda_{ex} = 360-500$ nm) de los EVOO (A), los POO (C) y los OO (E) y los loadings de emisión ($\lambda_{em} = 640-700$ nm) de los EVOO (B), los POO (D) y los OO (F).

Al hacer el respectivo análisis de los componentes de los EVOO, POO y OO en la **Figura 1.10** se observa que, en los EVOO, el primer componente, con mayor influencia en el modelo, presenta una banda de emisión ancha con un máximo en los 420 nm, y una banda ancha en el espectro de emisión también cuyo máximo se observa sobre los 680 nm.

Este espectro puede correlacionarse con el de la *clorofila a*, ya que, en aceites de calidad, como los EVOO, su presencia es mayor y su alto rendimiento cuántico de fluorescencia hace que su presencia aumente sensiblemente la señal de las EEMs. Además, el espectro coincide con el estudiado por Galeano Díaz et al. [36].

Por otro lado, el segundo componente, con menor presencia en los EVOO y con unos perfiles de fluorescencia muy similares, con su máximo de excitación localizado sobre los 415 nm y su máximo de emisión sobre los 675 nm, resulta ser el principal componente en los POO y en los OO, por lo que, dada su similitud con el espectro de la *feofitina a*, su mayor resistencia a la degradación oxidativa, y su abundante presencia en los aceites de oliva [37], dicho componente se ha asociado con la *feofitina a*.

En cuanto al tercer componente presente en los EVOO (A, B) y en los OO (E, F), muestran también una especial similitud entre ellos y con el segundo de los POO (C, D), con una extensa banda de excitación con máximos puntuales entre 390 y 420 nm, y una banda de emisión bien definida con su máximo situado sobre 660 nm, por lo que debido a su similitud con la *clorofila b* [36], y su decaimiento en los aceites de menor calidad, como consecuencia de la degradación oxidativa, se ha correlacionado con esta, aunque en este caso los indicios no son tan claros.

De este modo, la caracterización de los perfiles de los diferentes aceites analizados para posteriores estudios da una información de especial relevancia respecto a las características composicionales de los compuestos fluorescentes que contribuyen a las EEMs.

A continuación, se llevó a cabo el análisis exploratorio y clasificatorio de los datos teniendo en cuenta todo el conjunto de muestras.

B. Clasificación de muestras de aceite según su categoría en los rangos estudiados aplicando PARAFAC y LDA-PARAFAC

Una vez obtenidos los perfiles de fluorescencia a través de los *loadings* de las matrices B_PARAFAC y C_PARAFAC, el análisis de conglomerados (*clustering*) de los *scores* de PARAFAC permite comprobar de manera gráfica y sencilla la posible existencia de categorías en base a las distancias existentes entre los *scores* obtenidos para cada muestra, en el espacio conformado por los componentes obtenidos. De este modo, se conformó un conjunto de datos con las 30 muestras analizadas, 16 de EVOO, 10 de OO y 4 de POO (no EVOO), para, a través del modelo PARAFAC realizado en los conjuntos anteriores por separado, hacer un primer análisis exploratorio de la tendencia de los *scores* en el espacio conformado por las proyecciones de sus componentes obtenidos mediante los criterios anteriores, y los resultados obtenidos pueden visualizarse en la **Figura 1.11** para ambos rangos:

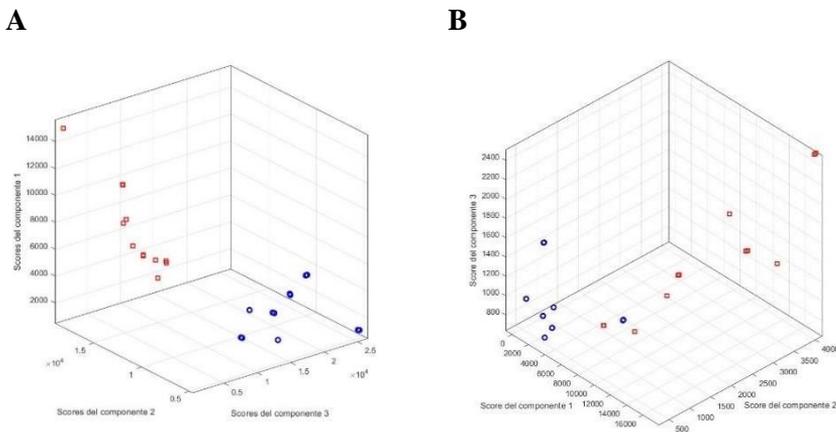


Figura 1.11. Clusters de los ‘scores’ de PARAFAC para el conjunto de datos correspondiente a los 16 EVOO (círculos rojos) 14 no EVOO (círculos azules) para el rango de los tocoferoles polifenoles y productos de oxidación (A) ($\lambda_{ex} = 270\text{-}350\text{ nm}$, $\lambda_{em} = 280\text{-}500\text{ nm}$) y el rango de clorofilas (B) ($\lambda_{ex} = 360\text{-}500\text{ nm}$, $\lambda_{em} = 640\text{-}700\text{ nm}$).

En el análisis de los *clusters* separados según la categoría de EVOO, o de aceites refinados (no EVOO) ya sean POO o OO, se puede observar que, para el primer rango estudiado, se observa una separación muy notable entre ellos, ya que estos componentes son los que más se relacionan con los atributos de calidad de los aceites de oliva y su estabilidad oxidativa, la cual, como resultado de determinados procesos de degradación correspondientes al refinado se pierden. Sin embargo, el segundo rango, correspondiente a las clorofilas, al ser esta dependiente de más factores, como puede ser el momento de cosecha, o las condiciones de almacenamiento, etc. no es tan decisiva.

De este modo, y para confirmar la tendencia detectada, la aplicación del análisis lineal discriminante sobre los ‘scores’ de PARAFAC (LDA-PARAFAC), permitió obtener los vectores canónicos de la función discriminante, separándolos en sus respectivos *clusters* a través de la rutina *elipses-3D*. Los resultados obtenidos para los rangos estudiados pueden visualizarse en la **Figura 1.12**.

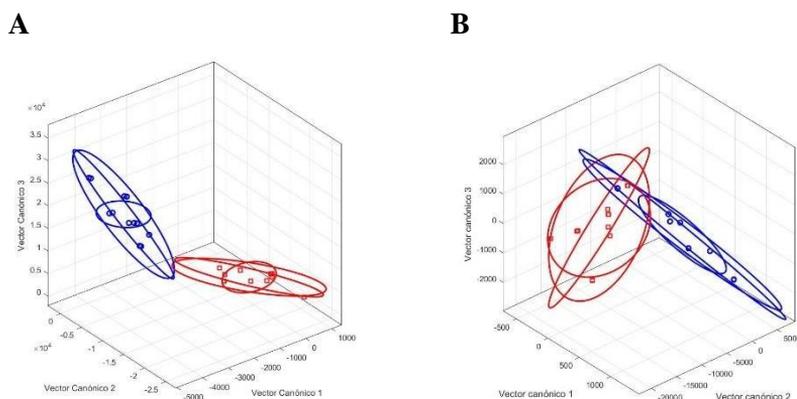


Figura 1.12. *Elipses de discriminación de categorías sobre los vectores canónicos de LDA-PARAFAC para el conjunto de datos correspondiente a los 16 EVOO (círculos rojos) y 14 no EVOO (círculos azules) para el rango de los tocoferoles polifenoles y productos de oxidación, (A) ($\lambda_{ex} = 270-350\text{ nm}$, $\lambda_{em} = 280-500\text{ nm}$) y el rango de clorofilas (B) ($\lambda_{ex} = 360-500\text{ nm}$, $\lambda_{em} = 640-700\text{ nm}$).*

Como se puede comprobar, para la discriminación entre grados de calidad, el mejor rango es, como se puede observar según el criterio establecido por la proyección de las elipses sobre los vectores canónicos de LDA, el rango atribuible a los tocoferoles, compuestos fenólicos y productos de oxidación secundaria ($\lambda_{ex} = 270-350\text{ nm}$, $\lambda_{em} = 280-500\text{ nm}$), tal y como se ha venido justificando a lo largo de este apartado. No obstante, la alta sensibilidad de las clorofilas en estados primarios de oxidación también resulta de gran utilidad para el control de sus condiciones de almacenamiento, tal y como veremos en los siguientes apartados, una vez establecida la caracterización básica de las EEMs de los aceites de diferente calidad obtenidas por medios no destructivos.

3.2. Estudio de la influencia de variables externas sobre la calidad de los aceites de oliva a través de sus EEMs

Para estudiar cómo se ve afectado el perfil de fluorescencia de los EVOO, en condiciones de conservación no adecuadas, como pueden ser con exposición a la luz o sometidos a temperaturas elevadas, se realizaron los experimentos que se detallan a continuación.

3.2.1. Estudio de la influencia de la radiación ultravioleta

Para llevar a cabo el estudio de la influencia de la radiación UV sobre la fluorescencia del EVOO, se registraron las EEMs de diferentes muestras de aceite de oliva virgen extra mantenidas en la oscuridad y, a continuación, alícuotas de cada aceite se irradiaron con luz UV, utilizando una lámpara de mercurio durante diferentes periodos de tiempo, comprendidos entre 1 y 70 min. Por cada muestra se obtuvieron dos EEMs, en los rangos de longitudes de onda previamente seleccionados. En la *Figura 1.13*, se muestran las EEMs de una de las muestras irradiada a diferentes tiempos, registradas en los dos intervalos de longitudes de onda seleccionados.

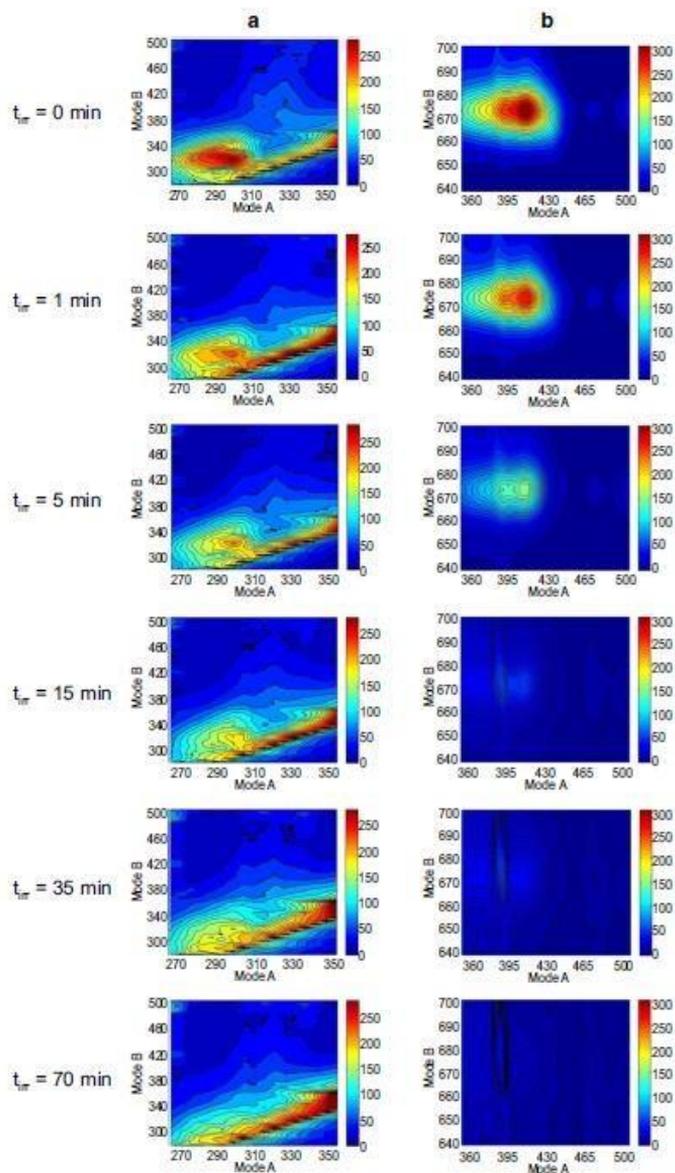


Figura 1.13. Matrices de fluorescencia de una muestra de aceite de oliva virgen extra sometida a irradiación UV durante diferentes tiempos: 0, 1, 5, 15, 35 y 70 minutos, en el rango espectral de tocoferoles, compuestos fenólicos y productos de oxidación (a) ($\lambda_{ex} = 270\text{-}350\text{ nm}$, $\lambda_{em} = 370\text{-}500\text{ nm}$) y de clorofilas y feofitinas (b) ($\lambda_{ex} = 360\text{-}500\text{ nm}$, $\lambda_{em} = 640\text{-}700\text{ nm}$).

En el rango de longitudes de onda correspondientes a los tocoferoles y a los compuestos fenólicos ($\lambda_{ex} = 270-350 \text{ nm}$, $\lambda_{em} = 284-500 \text{ nm}$) se observa cómo, después de 5 min de irradiación, el perfil de fluorescencia disminuye considerablemente y, a partir de 15 min, prácticamente desaparece la fluorescencia en esta región espectral.

Por otro lado, la fluorescencia correspondiente a la zona de los pigmentos presentes en el aceite de oliva también se ve afectada por la irradiación y, a partir de los primeros 5 min de irradiación, se pueden observar cambios que se acentúan con el tiempo de irradiación, desapareciendo prácticamente la banda de fluorescencia centrada a 675 nm a los 35 minutos de irradiación.

Este primer estudio se ha realizado con el objetivo de poder observar el modo en que la radiación UV afecta a los componentes fluorescentes del aceite de oliva que se monitorizan en las matrices de fluorescencia, comprobándose que, tanto en el rango de los tocoferoles y polifenoles, como en el rango de longitudes de onda de las clorofilas, disminuyen las intensidades de las bandas de fluorescencia. Este hecho se atribuye a que estos compuestos se degradan por acción de la luz a otros compuestos que no presentan propiedades fluorescentes. En el caso de los tocoferoles, por ejemplo, la radiación provoca reacciones radicalarias en las cadenas hidrocarbonadas y en los grupos OH de la estructura química de los tocoferoles, llegándose a formar quinonas, semiquinonas y aniones radicales [38]. En cuanto a las clorofilas, como ya se ha indicado, la acción de la luz sobre estos compuestos provoca la rotura de la estructura de la porfirina que deriva en la pérdida de las propiedades fluorescentes [28].

Una vez realizados estos primeros estudios en condiciones artificiales de irradiación con luz ultravioleta, se procedió a realizar un estudio similar en condiciones más realistas, sometiendo las muestras de

aceite a la radiación natural solar para poder obtener conclusiones extrapolables a posibles condiciones de almacenamiento reales. Para ello, se realizaron ensayos exponiendo las muestras de EVOO, durante tiempos prolongados, de hasta cinco horas, a la luz solar de forma directa. También se han estudiado en condiciones de luz menos intensas, dejando las muestras expuestas a la luz del laboratorio. En ambos casos se registraron las matrices en los dos rangos de longitudes de onda fijados para esta investigación. En estas condiciones la región espectral más afectada corresponde a la de los pigmentos. La evolución de las EEMs de muestras de aceite expuestas a la luz solar se muestra en la **Figura 1.14**.

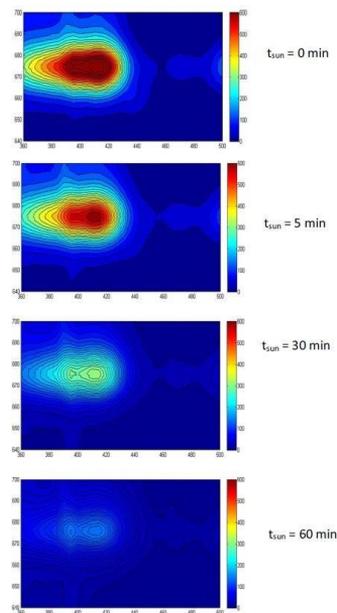


Figura 1.14. Influencia de la radiación solar en el perfil de fluorescencia de una muestra de aceite de oliva virgen extra en el rango de longitudes de clorofilas, inalterada y sometida a periodos de radiación solar de 5, 30 y 60 minutos.

Como se puede comprobar en la **Figura 1.14**, las clorofilas se ven fuertemente afectadas ante la radiación solar, incluso en condiciones de radiación ambiental, ya que los pigmentos son más sensibles a la radiación que los compuestos fenólicos en los procesos de oxidación del aceite de oliva, tal y como ensayaron Poulli K.I. et al. [39] en la medición de compuestos de oxidación totales en los aceites (TOTOX) como consecuencia de la exposición a la luz a 80°C, así como Jenisova. Z. et al.[40], que midieron el decaimiento del contenido en pigmentos en relación al tiempo de exposición a la radiación UV. Con esto, tras una primera exploración y de manera cualitativa, se puede decir que las clorofilas son más sensibles a la exposición a la radiación solar que los tocoferoles y polifenoles. Esto influye de manera importante en la calidad de los aceites, ya que el contenido en clorofilas influye en la autooxidación de los aceites y, por tanto, retrasan la aparición de compuestos procedentes de la oxidación de ácidos grasos, como son los peróxidos que dan lugar posteriormente a la aparición de productos de oxidación secundaria, como son aldehídos, cetonas y ácidos de bajo peso molecular que afectan a las características sensoriales y funcionales de los EVOO.

Una vez comprobado que la irradiación afecta significativamente al perfil fluorescente de los aceites de oliva, se estableció un modelo quimiométrico que permitiera diferenciar entre muestras de aceite bien conservadas de aquellas que han sido expuestas a la luz UV.

Se utilizó un conjunto de 20 muestras para cada tiempo de irradiación, lo que permitió construir matrices de datos de segundo orden de dimensiones [20×17×433] (*muestras* × λ_{ex} × λ_{em}) para el rango de los tocoferoles, polifenoles y productos de oxidación, y de dimensiones [20×29×31] (*muestras* × λ_{ex} × λ_{em}) para el rango espectral de las clorofilas. Se descompusieron las matrices utilizando PARAFAC y se utilizan los

valores de los *scores* para estudiar la capacidad de PARAFAC para diferenciar entre aceites.

Posteriormente, siguiendo la misma metodología que en el apartado anterior para hacer un análisis discriminante de muestras irradiadas frente a no irradiadas, se aplicó LDA sobre los scores de PARAFAC. Las representaciones tridimensionales de los vectores canónicos rodeados de la proyección estadística de significación de las elipses en el espacio de los tres componentes para las dos zonas espectrales utilizadas, se presentan en las *Figura 1.15*.

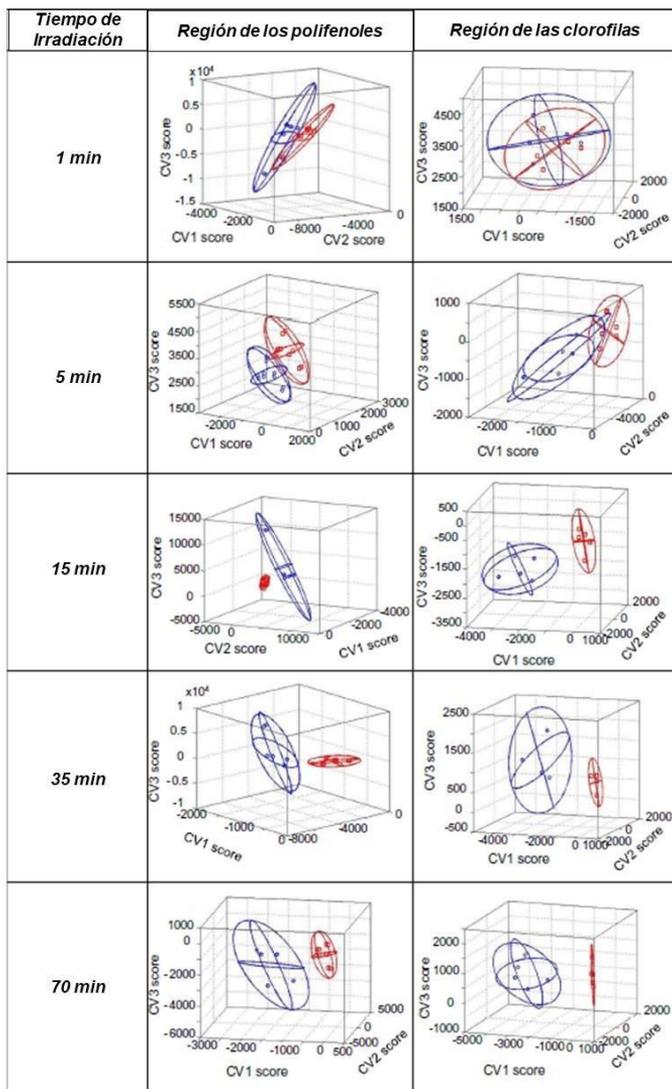


Figura 1.15. Elipses de LDA aplicados a aceites sin irradiar (azules) e irradiados (rojos) a los tiempos de irradiación de 1, 5, 15, 25 y 70 minutos.

Analizando los datos se observa que, a medida que aumenta el tiempo de irradiación, se incrementa la posibilidad de diferenciar entre aceites irradiados y aceites sin irradiar. Las elipses proyectadas sobre los vectores canónicos de LDA muestran que, utilizando el rango de longitudes de onda de los tocoferoles y polifenoles, a partir de los 15 min de irradiación

es posible llegar a observar una discriminación entre aceites irradiados y aceites sin irradiar de manera inequívoca. Sin embargo, esta separación es más efectiva utilizando el rango de las clorofilas, puesto que se puede llegar a detectar la separación entre aceites irradiados 5 min, de aceites sin alterar. Con esto, se puede concluir que el rango de clorofilas es el adecuado para detectar si un aceite ha iniciado el proceso de oxidación que afecta a sus características de calidad.

3.2.2. Influencia de la temperatura en el perfil fluorescente de los aceites de oliva

Otra de las variables externas que afectan a la calidad del aceite de oliva virgen extra es la temperatura de almacenamiento y esto también se ve reflejado en el perfil fluorescente de las muestras.

En este estudio, y siguiendo la misma sistemática que en el caso anterior, se registran las EEMs de muestras calentadas durante diferentes tiempos. Las temperaturas y tiempos de calentamiento se resumen en la *Tabla 1.2*.

Las EEMs de cada muestra fueron registradas usando los mismos rangos de longitudes de onda empleados en los estudios previos. En la *Figura 1.16* se presentan las matrices de fluorescencia de una muestra de aceite sin haber sido sometida a calentamiento y después de calentarla a 80 °C durante 120 min. Se observa que, después del calentamiento, aparece una nueva zona espectral de fluorescencia en un rango de longitudes de onda de emisión de entre 380 y 500 nm, y un rango de longitudes de onda de excitación de entre 300 y 350 nm y, además, la fluorescencia de la zona de los polifenoles, situada en un rango de longitudes de onda de emisión de entre 300 y 350 nm, y un rango de longitudes de onda de excitación de entre 280 y 310, prácticamente ha desaparecido. Esta región se corresponde con los productos de oxidación secundaria de los aceites refinados, ya

mencionados y referenciados en el **apartado 3.1** de este capítulo. Por otro lado, con respecto a la afectación que sufren las clorofilas bajo estas condiciones de calentamiento ensayadas, cabe mencionar que sus EEMs no sufrieron cambios significativos a simple vista. Así mismo, su tratamiento quimiométrico y análisis exploratorio posterior no ofrecieron resultados prometedores para su aplicación en el seguimiento del deterioro oxidativo por calentamiento, por lo que en este apartado no se muestran sus resultados experimentales.

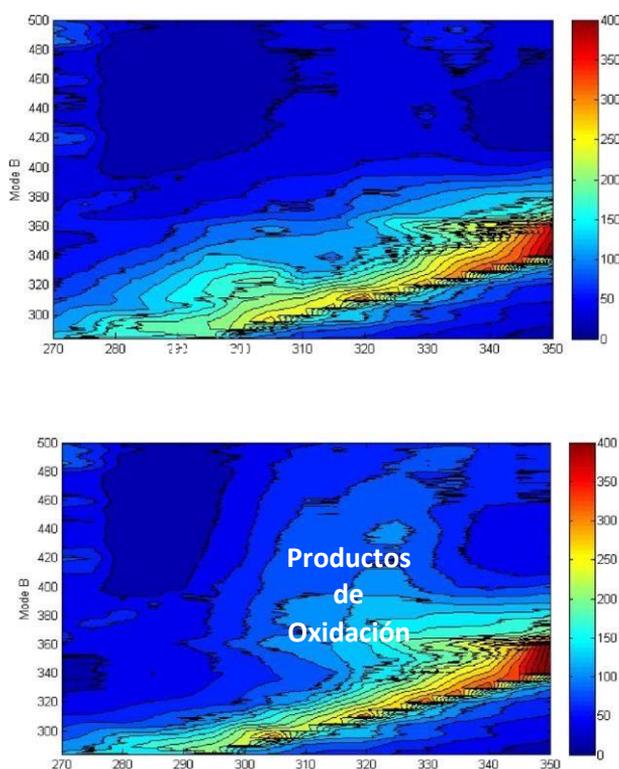


Figura 1.16. EEMs de un EVOO en la zona espectral de los polifenoles, tocoferoles y productos de oxidación ($\lambda_{ex} = 270-350$ nm, $\lambda_{em} = 284-500$ nm), de manera inalterada (A) y sometida a un calentamiento de 80°C durante 120 minutos (B).

Se utilizaron 9 conjuntos diferentes de 16 muestras, cada uno de ellos correspondiente a los diferentes tiempos de calentamiento (30, 60 y 120 min) a una misma temperatura (40, 60 y 80°C). Se llevó a cabo la descomposición mediante PARAFAC de los cubos de datos en el rango de los tocoferoles y polifenoles de dimensiones $[16(\text{muestras}) \times 17(\lambda_{ex}) \times 433(\lambda_{em})]$, obteniéndose un número de componentes óptimo de tres.

Atendiendo a las observaciones previas, en este caso, se optó por subdividir la EEM en dos rangos espectrales diferentes, siendo el primero el correspondiente a la zona de los polifenoles y tocoferoles con longitudes de onda de emisión comprendidas entre 310 y 400 nm y excitación entre 270 y 310 nm y el segundo rango correspondiéndose con los productos de oxidación secundaria, con rangos de emisión entre 380 y 500 nm y de excitación entre 300 y 350 nm.

Los perfiles fluorescentes obtenidos a partir de las matrices de los *loadings* para el primer componente de PARAFAC, una vez desdoblada la matriz para las distintas temperaturas ensayadas, pueden observarse en la **Figura 1.17**.

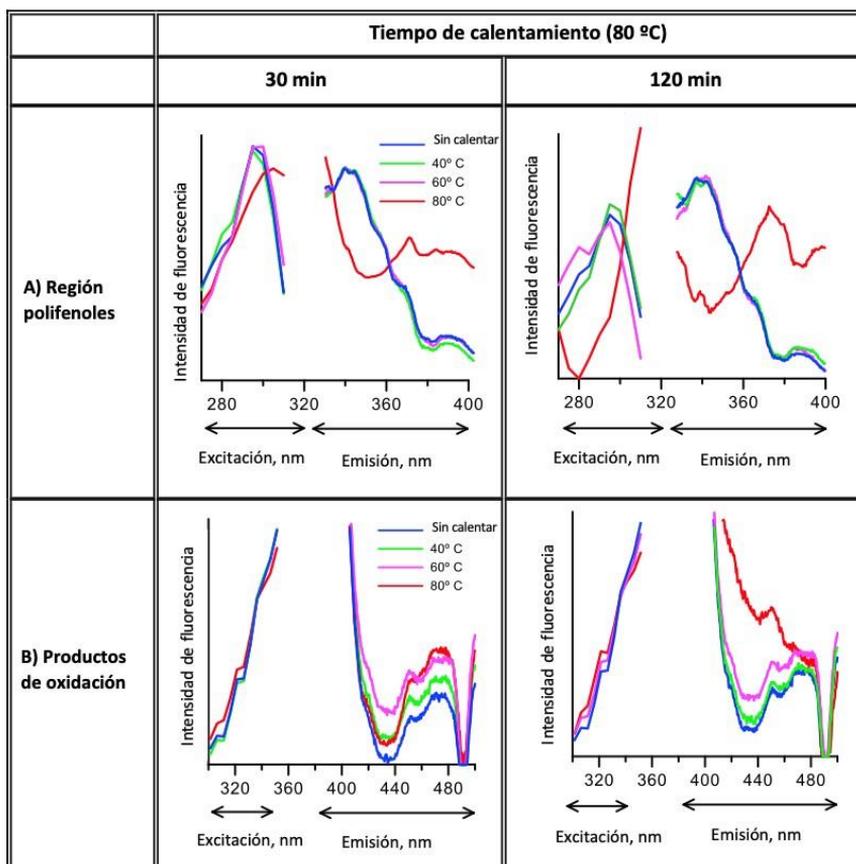


Figura 1.17. Influencia del tiempo de calentamiento (30 y 120 min) en los perfiles de excitación y emisión de PARAFAC para el primer componente en la zona espectral de los polifenoles (A) y en la zona espectral de los productos de oxidación (B), para los aceites ensayados sin alterar y calentados a temperaturas de 40, 60 y 80°C.

Se puede observar que para temperaturas inferiores a 60°C, la fisonomía de los *loadings* del primer componente no se altera significativamente, independientemente del tiempo de calentamiento, tanto en la zona espectral de los compuestos fenólicos y tocoferoles, como en la de los productos de oxidación. Sin embargo, para temperaturas de 80°C, los perfiles de excitación y emisión del primer componente cambian de manera

significativa. Para el caso del rango de los polifenoles (*Figura 1.17a*) se observa que, a 80°C, después de calentar la muestra durante 30 min, se produce un decaimiento en el máximo de emisión situado sobre los 350 nm, así como un incremento de la banda cuyo máximo se encuentra sobre 370 nm, situación que se vuelve mucho más pronunciada para tiempos de calentamiento de 120 min. La intensidad de fluorescencia del máximo de emisión situado a 370 nm aumenta considerablemente probablemente debido al aumento en la producción de productos de oxidación secundaria. En cuanto a los *loadings* de excitación, también se observa una disminución de la intensidad de fluorescencia en el máximo de la banda que está entre 270 y 300 nm. Además, en el rango de los productos de oxidación secundaria (*Figura 1.17b*), también se puede observar la intensificación de bandas de fluorescencia en la zona de emisión para la banda que se encuentra entre 420 y 480 nm, desplazándose su máximo hacia longitudes de onda más cortas, situación también atribuible a la mayor presencia de estos productos de oxidación secundaria con sus máximos de emisión situados sobre estas zonas.

De acuerdo a estas observaciones cualitativas que los perfiles de PARAFAC pueden proporcionar, se puede concluir que no sólo existe una degradación de los polifenoles como consecuencia de temperaturas moderadamente elevadas, sino que además se produce la formación de productos de oxidación secundaria, tal como se recoge en la bibliografía, según los estudios llevados a cabo por Guimet et al. [30]–[35], y Tena et al. [41]. En dichos estudios se especifica que la aparición de los productos de oxidación favorece la aparición de bandas de fluorescencia a longitudes de onda similares a las detectadas en esta investigación (λ_{em} entre 420 y 480 nm) a temperaturas muy elevadas, superiores a los 170 °C. No obstante, en esta investigación se ha podido detectar dicha banda en condiciones menos

drásticas, de temperaturas del orden de 80°C y tiempos de calentamiento de 120 min.

Por último, para terminar con los estudios planteados para el seguimiento de la alteración producida en los EVOO como consecuencia del calentamiento, se hizo un análisis supervisado por LDA sobre los *scores* de PARAFAC con su respectivo test de elipses, para comprobar si existían diferencias significativas entre los aceites sometidos a elevadas temperaturas y los que no habían sido previamente calentados. En la **Figura 1.18** se muestran las representaciones tridimensionales correspondientes a los vectores canónicos de LDA obtenidos en el rango espectral de los compuestos fenólicos y tocoferoles. Se representan los valores para los dos grupos de muestras: muestras no calentadas y muestras sometidas a una temperatura de 80 °C, durante 30, 60 y 120 minutos. Como en representaciones anteriores se incluye la proyección de las elipses, para un 95 % de confianza, sobre los tres planos definidos por sus correspondientes ejes.

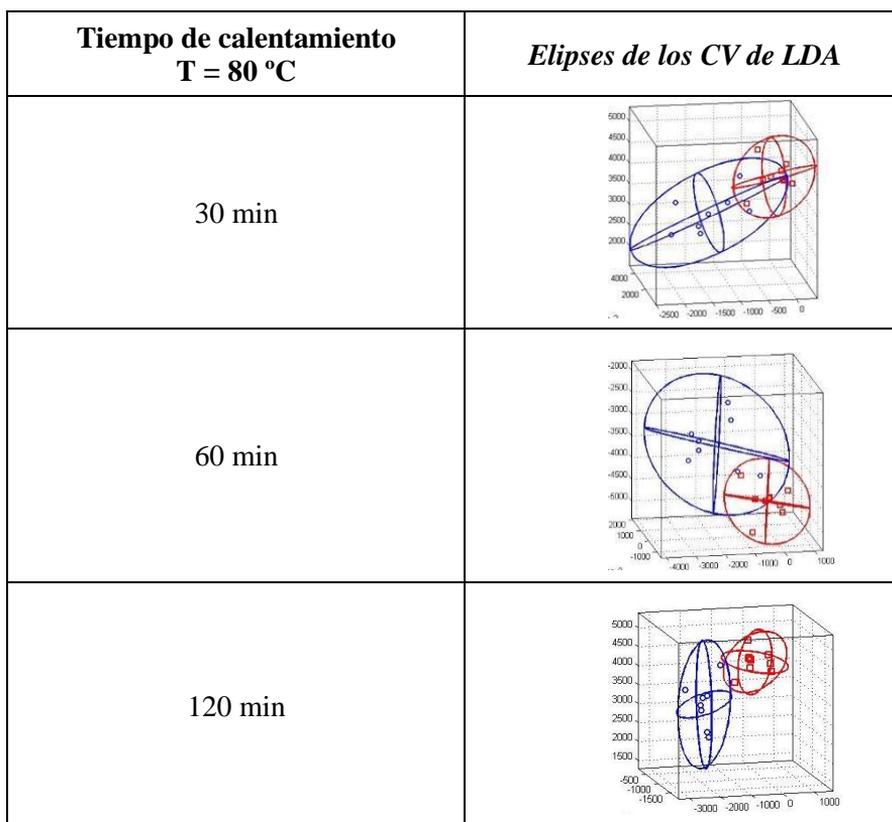


Figura 1.18. Elipses aplicadas sobre los vectores de LDA para las 16 muestras registradas de EVOO, de las cuales 8 no fueron sometidas a calentamiento (círculos azules) y las otras 8 fueron sometidas a un calentamiento de 80 °C durante 30, 60 y 120 min (cuadrados rojos).

Como se puede observar, para tiempos de calentamiento a 80 °C inferiores a una hora, no es posible diferenciar entre los dos grupos de muestras. Sin embargo, existe una clara separación entre los grupos sometidos a calentamiento durante un tiempo de 120 minutos. Esto permite concluir que los algoritmos empleados son capaces de discriminar tratamientos térmicos moderadamente elevados, aspecto que no es factible que se dé en condiciones normales de almacenamiento, pero sí que es indicativo de tratamientos térmicos más intensos como es el cocinado de alimentos.

3.3. Detección y determinación del grado de adulteración de aceites de oliva virgen extra (EVOO), con aceites de oliva (OO) y aceites de orujo de oliva (POO)

La última parte de este capítulo está dedicada al desarrollo de métodos para la detección de fraudes en el aceite de oliva virgen extra, así como en la determinación del grado de adulteración de EVOO con OO y POO, utilizando las herramientas quimiométricas previamente desarrolladas para el análisis exploratorio y el análisis discriminante, como son PARAFAC y PARAFAC-LDA. En este caso, también se ha empleado el análisis discriminante DA-UPLS. Posteriormente, se ha validado el modelo a través de la predicción de un conjunto de muestras de validación de aceites adulterados con distintos grados de OO y POO.

Para la realización de estos estudios se disponía de muestras comerciales de aceite de oliva virgen extra y de aceites de oliva refinados, como aceites de oliva y aceites de orujo de oliva. Se prepararon muestras de aceites de oliva virgen extra adulteradas con diferentes proporciones de los aceites refinados. La composición de estas muestras se recoge en la **Tabla 1.3**.

Tabla 3. Diseño experimental propuesto para el estudio del nivel de adulteración de aceites de oliva virgen extra con aceites de menor calidad.

Sets de calibración de los EVOO adulterados con OO		
Set de calibración	EVOO	EVOO/OO
1	10 muestras	10 muestras de EVOO adulteradas al 33% con VOO de distinta clase
2	10 muestras	10 muestras de EVOO adulteradas al 16% con VOO de distinta clase
3	10 muestras	10 muestras de EVOO adulteradas al 10% con VOO de distinta clase
4	10 muestras	10 muestras de EVOO adulteradas al 5% con VOO de distinta clase
Set de calibración de los EVOO adulterados con POO		
Set de calibración	EVOO	POO
1	10 muestras	10 muestras de EVOO adulteradas al 33% con POO de distinta clase
2	10 muestras	10 muestras de EVOO adulteradas al 16% con POO de distinta clase
3	10 muestras	10 muestras de EVOO adulteradas al 10% con POO de distinta clase
4	10 muestras	10 muestras de EVOO adulteradas al 5% con POO de distinta clase
5	10 muestras	10 muestras de EVOO adulteradas al 3% con POO de distinta clase

Siguiendo la sistemática indicada en las experiencias anteriores para el desarrollo de modelos quimiométricos, con el fin de discriminar con un 95 % de nivel de significación el grado de adulteración de EVOO con

distintas proporciones de OO y de POO, en primer lugar, se registraron las EEMs en las dos regiones espectrales. En la **Figura 1.19** se comparan las EEMs obtenidas de muestras de EVOO y de muestras adulteradas con diferentes porcentajes de los aceites refinados, entre un 5% y un 33% de OO y POO.

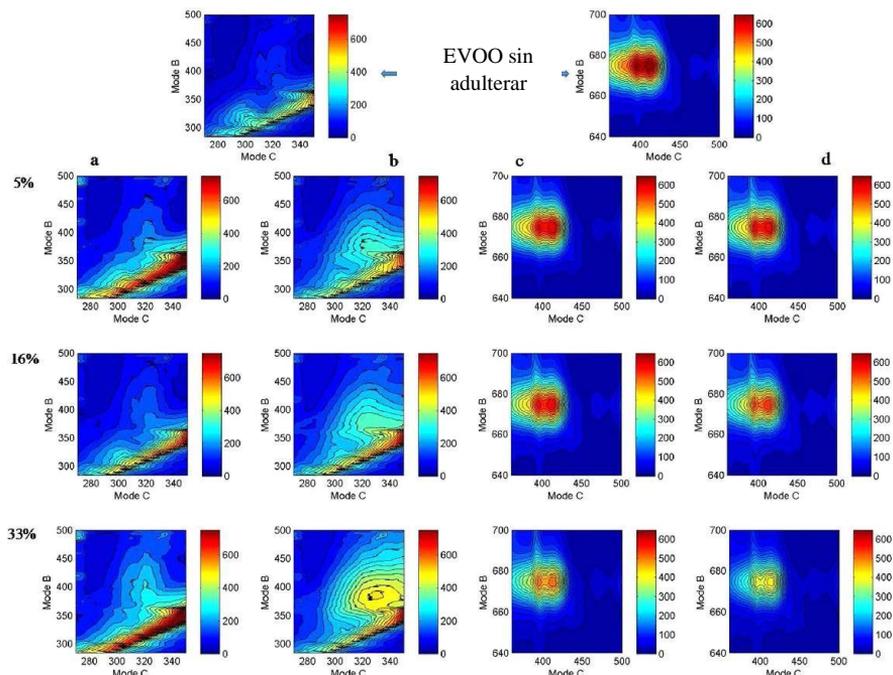


Figura 1.19. EEMs de muestra de EVOO adulterada con diferentes niveles (5, 16 y 33%) de OO (a y c) y POO (b y d) en el rango de longitudes de onda de tocoferoles y polifenoles (a y b), y el rango de longitudes de onda de las clorofilas (c y d).

Como se puede observar, a medida que aumenta el grado de adulteración con OO (*Figura 1.19a*) aparece una banda de fluorescencia en la zona a longitudes de onda de excitación entre 310 y 335 nm y entre 370 y 440 nm en el rango de emisión. Cuando se añade POO como adulterante (*Figura 1.19b*), la banda de fluorescencia aparece en la misma zona pero se observa con mayor intensidad, cuestión lógica respecto a la caracterización previa de las EEMs de dichos tipos de aceite de oliva, ya que pese a que los POO y los OO son sometidos a procesos parecidos de desodorización (entre 200 y 250 °C) y decoloración (sobre 100°C), los OO son el resultado de mezclas de aceites de oliva virgen (VOO) con aceites de oliva refinados (ROO), por lo que, por efecto de la dilución, su contenido en compuestos oxidados fluoróforos será menor, lo que a su vez lleva a que la detección del grado de adulteración en estos últimos sea menos sensible.

Con respecto a la zona espectral de los pigmentos, se observa que, tanto para las adulteraciones realizadas con OO (*Figura 1. 19c*) como para las realizadas con POO (*Figura 1.19d*), al aumentar el grado de adulteración, disminuye proporcionalmente la banda de fluorescencia localizada en la región de las clorofilas (rango de emisión entre 660 y 680 nm y rango de excitación entre 360 y 420 nm), ya que los EVOO contienen altas concentraciones de pigmentos que con los procesos de refinado, producen una degradación en estos compuestos, lo cual se traduce en la reducción de la intensidad de fluorescencia de las propias EEMs.

3.3.1. Descomposición espectral de las EEMs y determinación del número de componentes

Una vez caracterizadas las EEMs y observados los cambios producidos como consecuencia de las adulteraciones de las EVOO con otros tipos de aceites de menor calidad y composición diferenciada en las zonas espectrales estudiadas, a continuación, se realizaron la

descomposición espectral de las EEMs. Los primeros modelos se construyeron de manera independiente para cada uno de los aceites utilizados como adulterante. Las regiones espectrales correspondientes a tocoferoles y polifenoles y pigmentos se trataron de forma independiente. Así, los datos de fluorescencia se dispusieron en dos estructuras 3D independientes, el primero de dimensiones $[80 \times 433 \times 17]$ (muestras \times emisión \times excitación) en la región espectral de los tocoferoles, compuestos fenólicos y productos de oxidación, y el segundo de dimensiones $[80 \times 121 \times 29]$, correspondiente a la región espectral de los pigmentos. El número óptimo de componentes se calculó utilizando el criterio de la desviación estándar de los residuales de la matriz de PARAFAC y, como se observa en la **Figura 1.20**, el menor número de componentes para los cuales el error residual no es estadísticamente diferente al ruido instrumental es de tres componentes en todos los casos.

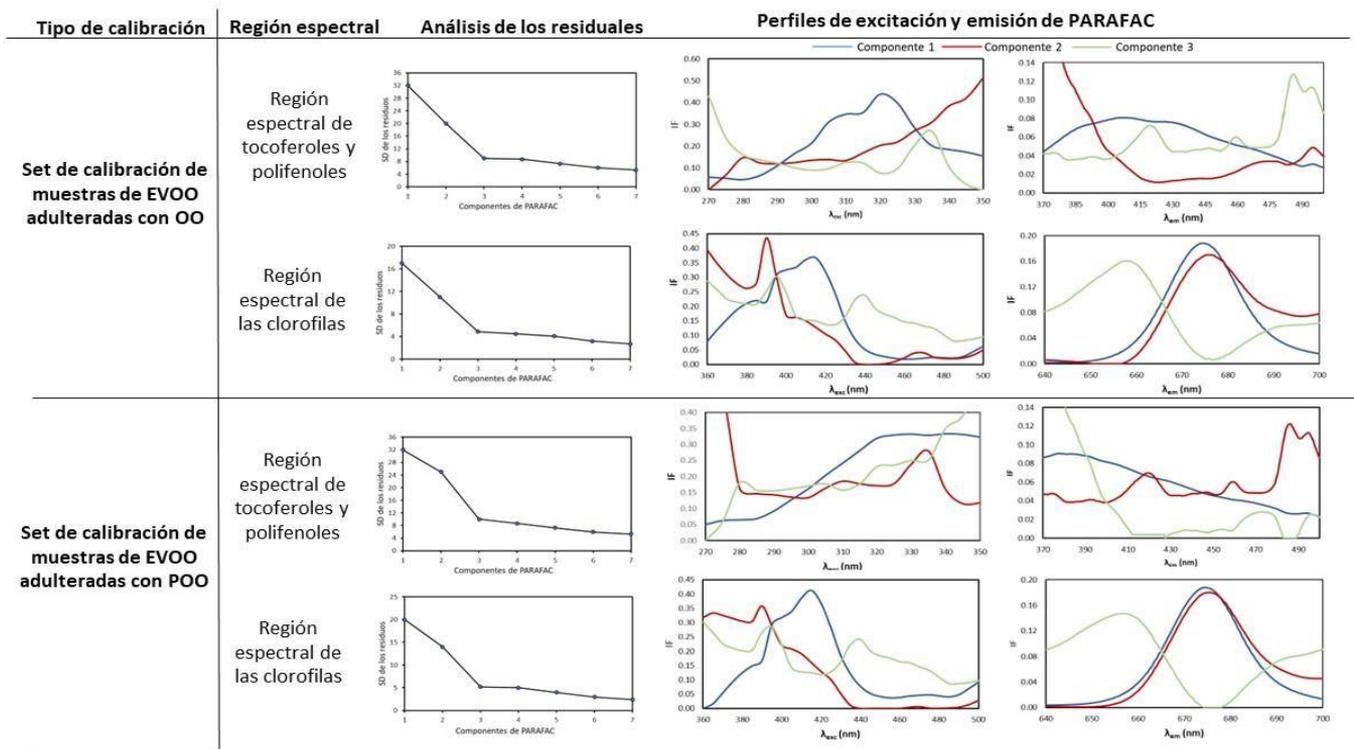


Figura 1.20. Desviación estándar de los errores de los residuales en función del número de componentes y perfiles de fluorescencia de excitación y emisión de los conjuntos de muestras de EVOO adulterados con OO y POO, en las regiones espectrales de los tocoferoles y polifenoles y las clorofilas.

Como se puede observar en la **Figura 1.20**, para este número de componentes se han obtenido los perfiles de excitación y de emisión en ambas regiones. Se observa que, en el caso del rango de los tocoferoles y los polifenoles, la fisonomía del primer componente es muy similar para ambas adulteraciones. Sin embargo, dependiendo del tipo de adulteración, el componente 2 y el componente 3 ofrecen una contribución distinta a la matriz general. El componente 1 presenta características espectrales muy similares en el caso de las adulteraciones con OO y con POO, presentando un máximo de excitación a 320 nm y una amplia banda de emisión entre los 370 y los 420 nm. En el caso del perfil de emisión del primer componente concuerda con el perfil correspondiente a los productos de oxidación secundaria para longitudes de onda de excitación de 310 nm descritas por Guimet et al. [32]. Por otro lado, en estos perfiles de excitación también es posible encontrar la contribución de los tocoferoles a las mismas longitudes de onda [31].

La fisonomía del segundo componente obtenido para el conjunto de EVOO adulterado con OO es similar a la del tercer componente obtenido para el conjunto de EVOO adulterado con POO. Sin embargo, las características de estos componentes espectrales no están aún bien definidas. Los perfiles de excitación presentan un crecimiento continuo con un pequeño máximo situado a 280 nm, mientras que los perfiles de emisión presentan un hombro a 475 nm y un máximo situado a 490 nm. Estos componentes pueden relacionarse con productos de degradación secundaria producidos como consecuencia de los tratamientos realizados sobre los OO y los POO, ya que sus valores de *scores* son mayores con el grado de adulteración.

Por último, para el tercer componente, los perfiles de excitación y emisión del modelo, para los EVOO adulterados con POO, son similares a los obtenidos para el segundo componente del conjunto de los EVOO

adulterados con OO. El perfil de excitación presenta dos máximos centrados a 315 y 335 nm y en el de emisión también se observan dos máximos a 420 y 480 nm. Las observaciones realizadas para los ‘scores’ de estos componentes también revelan que los valores más altos obtenidos son para el caso de las adulteraciones de POO que en el caso de las adulteraciones con OO como consecuencia de que el proceso de elaboración de los POO es más drástico en la formación de productos de oxidación secundaria que en el caso de los OO.

En la región espectral de los pigmentos, los perfiles obtenidos en ambos conjuntos son muy similares, independientemente del tipo de adulteración, siendo el orden de los tres componentes y, por lo tanto, su contribución al modelo, el mismo en los dos *conjuntos* de calibración estudiados, lo que puede indicar que la escasa presencia de pigmentos verdes en los OO y en los POO no contribuyen al modelo PARAFAC en una magnitud semejante a los EVOO. El primer componente en ambos conjuntos tiene un perfil de fluorescencia con máximos de excitación y emisión en 418 y 675 nm, respectivamente, longitudes de onda que se corresponden con el perfil atribuido a la *clorofila a*, tal y como se pudo comprobar en el **apartado 3.1**. El segundo componente, con un máximo de excitación centrado a una longitud de onda más corta, concretamente a 390 nm, y un máximo de emisión localizado a una longitud de onda ligeramente superior, 678 nm, se puede atribuir a la presencia de feofitina *a*. Finalmente, el tercer componente detectado por el modelo presenta un máximo de emisión a 655 nm y puede estar asociado a la *clorofila b*. Las longitudes de onda de excitación y emisión obtenidas por PARAFAC para estos pigmentos son ligeramente diferentes a los que aparecen en la bibliografía, donde las matrices de fluorescencia fueron obtenidas en disolventes puros, mientras que los espectros obtenidos en este caso corresponden a los obtenidos en muestras de aceites inalterados, lo cual supone un medio no

polar y de alta viscosidad, condiciones que pueden desplazar las bandas de fluorescencia por dichos efectos de los disolventes.

3.3.2. Discriminación del grado de adulteración mediante PARAFAC-LDA

Una vez obtenidas las matrices desdobladas de *scores* y *loadings* de PARAFAC, siguiendo la sistemática empleada en los estudios previos para detectar alteraciones en los EVOO, se procedió a someter a los *scores* de PARAFAC al proceso de análisis lineal discriminante (LDA), para poder obtener los vectores canónicos a través de dicho proceso de aprendizaje supervisado y obtener la separación por categorías previamente seleccionadas. Para la realización de este estudio se utilizaron 10 muestras de EVOO, que estuviesen perfectamente envasadas. A estos aceites se les adicionaron cantidades diferentes de tres tipos de aceites de oliva, de inferior calidad, de los cuales dos eran OO y uno de POO, para que las concentraciones de los aceites fraudulentos fuesen, en porcentaje/volumen, 33,3%, 16,7%, 10%, 5% y 3,33%, respectivamente. A cada una de las muestras se les registró las EEMs en los rangos de longitudes de onda seleccionados. Los resultados obtenidos en la representación de los vectores canónicos de LDA, en el espacio de los tres componentes principales, para el rango espectral de los tocoferoles y polifenoles ($\lambda_{em}=360-480nm$, $\lambda_{ex}=284-350nm$), así como su clasificación según el criterio de las elipses al 95% de significación, pueden observarse en las **Figuras 1.21** y **1.22**, para aceites de oliva virgen extra contaminadas con OO y POO, respectivamente.

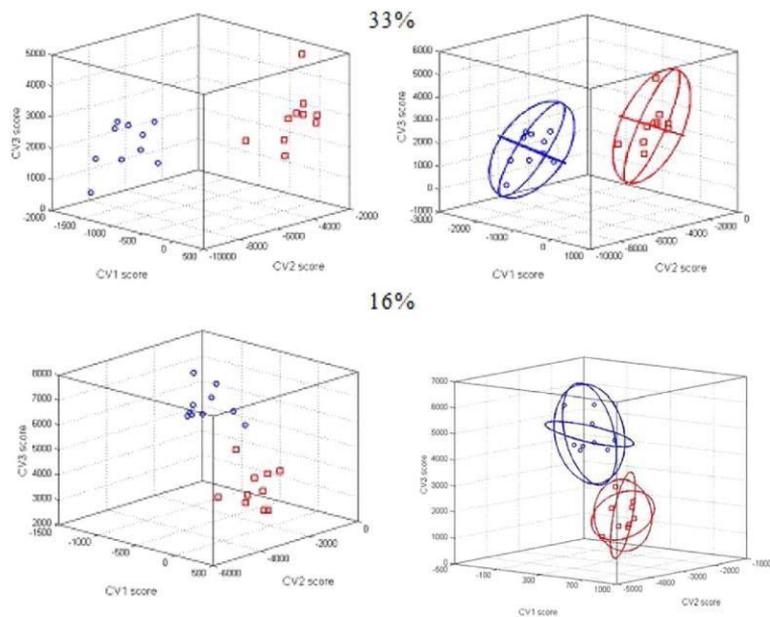


Figura 1.21. ‘Clusters’ de LDA-PARAFAC (izquierda) y elipses al 95% de significación (derecha) en la región espectral de los tocoferoles y polifenoles ($\lambda_{em}=360-480nm$, $\lambda_{ex}=284-350nm$) para la discriminación entre EVOO \circ y EVOO adulterado con diferentes cantidades de OO \square ; en diferentes proporciones (33% y 16%).

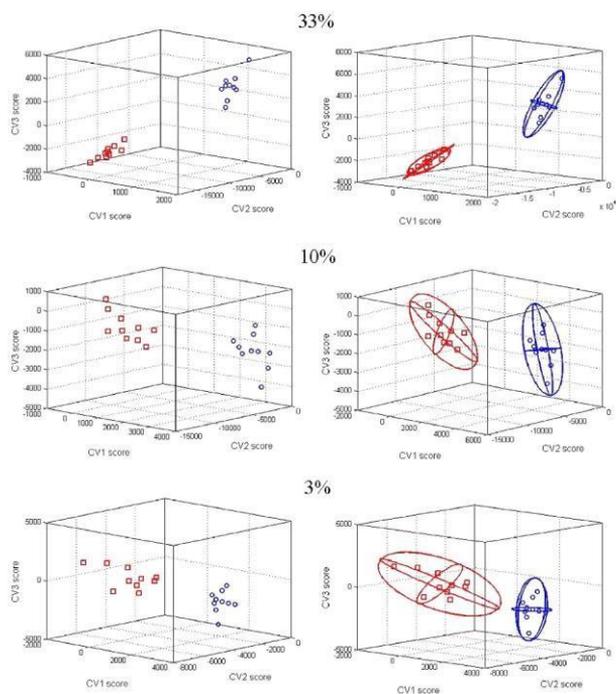


Figura 1.22. ‘Clusters’ de LDA-PARAFAC (izquierda) y elipses al 95% de significación (derecha) en la región espectral de los tocoferoles y polifenoles ($\lambda_{em}=360-480nm$, $\lambda_{ex}=284-350nm$) para la discriminación entre EVOO (○) y EVOO adulterado con diferentes cantidades de POO (□); en diferentes proporciones (33%, 10% y 3%).

Atendiendo a la discriminación establecida según la proyección de los elipses, en el espacio de coordenadas de los tres primeros componentes principales configurado por los vectores canónicos de LDA-PARAFAC, se concluye que, con un nivel de significación del 95%, se pueden diferenciar entre aceites EVOO y aquellos contaminados con solo un 3% v:v con aceites de orujo de oliva. No es así en el caso de adulteraciones con aceites de oliva a base de mezclas de aceites refinados, los cuales solo se pueden distinguir en los *clusters* de datos de LDA-PARAFAC a partir de un 16% de adulteración.

En la región espectral de los pigmentos, las diferencias solo son observables para adulteraciones con cualquiera de los dos tipos de aceites utilizados superiores al 33% v/v. Esta peor capacidad de discriminación en el rango de clorofilas se debe a que los aceites refinados no aportan información a las EEMs de los EVOO adulterados, ya que solo disminuye la señal de las clorofilas por efecto de la dilución.

3.3.3. Discriminación del grado de adulteración mediante DA-UPLS

El algoritmo DA-UPLS puede proporcionar mejores resultados cuando el número de grupos es pequeño y, además, puede proporcionar información adicional acerca de su capacidad predictiva.

Este estudio se ha llevado a cabo solo con las matrices registradas en el rango de tocoferoles y polifenoles, dada la poca discriminación observada en la zona espectral de los pigmentos. Para aplicar DA-UPLS, en primer lugar, al igual que en PARAFAC, hay que seleccionar el número de componentes. Su selección se hace según el número de componentes que hace que el valor del PRESS, (**apartado 3** de la introducción de esta tesis) se minimice, sin producir un sobreajuste en la matriz adicionando factores que den ruido. Este valor no es el mínimo PRESS, sino aquel número de factores que proporcionen un PRESS que no sea significativamente mayor al mínimo PRESS, según el parámetro estadístico F . Este criterio fue propuesto por Haaland y Thomas (1988) [42, 43] y es el que se utiliza en este trabajo para la elección del número de factores. En el caso del conjunto de muestras adulteradas con OO, el número óptimo de factores fue dos, mientras que, en el caso de muestras contaminadas con POO, el número óptimo de componentes en el rango estudiado fue de cinco.

Este método fue probado con las dos regiones espectrales previamente ensayadas con PARAFAC-LDA. Para la construcción del modelo, puesto que requiere subdividir los conjuntos de muestras en un

conjunto de entrenamiento y otro de validación, se dividió el conjunto de 20 muestras en uno de 12 muestras, con 6 muestras adulteradas y 6 sin adulterar y en un conjunto de validación de 8 muestras restantes, con 4 muestras adulteradas y 4 sin adulterar, tanto en los conjuntos iniciales de EVOO con EVOO adulterado con OO, como con el conjunto de EVOO y EVOO adulterado con POO.

Posteriormente, y dadas las indicaciones que aparecen en el **apartado 2.6** de este capítulo, en cada conjunto, se codificaron las muestras adulteradas con un 2 y las muestras sin adulterar con un 1, para poder obtener la clasificación de las muestras de acuerdo a la categoría, mostradas en la **Figura 1.23**, en la cual se representa el valor predicho frente a los códigos nominales para los diferentes niveles de adulteración de los aceites EVOO con OO y POO. El intervalo de confianza para ambas categorías queda marcado por las bandas horizontales, las cuales fueron estimadas como el producto de la desviación estándar calculada de los resultados obtenidos para cada categoría predicha en la muestra de entrenamiento y el valor de la *tabla t de Student* (con un nivel de confianza del 95%) y con $n-1$ grados de libertad para cada categoría. La capacidad de dicho algoritmo para la discriminación se evalúa a través del solapamiento producido entre las dos categorías establecidas.

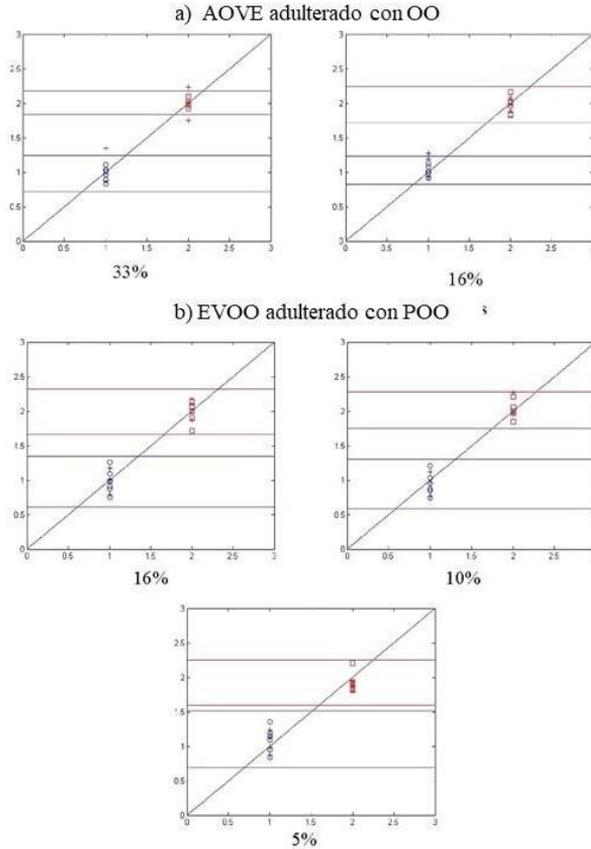


Figura 1.23. Gráficos de DA-UPLS para la discriminación entre EVOO (símbolos azules) y EVOO adulterada con diferentes proporciones de POO (símbolos rojos) en el rango espectral de tocoferoles y polifenoles. □ conjunto de calibración; + conjunto de validación

Como puede observarse en la **Figura 1.23a**, para muestras de EVOO adulteradas con diferentes niveles de OO, y en la **Figura 1.23b**, para EVOO adulterados con diferentes porcentajes de POO, los límites para una discriminación aceptable fueron niveles de adulteración superiores al 16% para la contaminación con aceites de oliva y del 5% para la contaminación con aceites de orujo de oliva, siendo pues los resultados similares a los obtenidos utilizando LDA-PARAFAC, no pudiéndose distinguir en el caso de las adulteraciones con POO adulteraciones inferiores al 5%.

3.3.4. Validación de los métodos de clasificación. Cuantificación mediante U-PLS

Como última parte de este estudio, se aplicó U-PLS en un intento de cuantificar el nivel de adulteración de diferentes muestras EVOO contaminadas con otros distintos tipos de OO y POO. Para ello, se prepararon dos nuevos conjuntos de muestras de aceites de EVOO adulteradas con diferentes proporciones y tipos de OO y POO, tal como se recoge en la **Tabla 1.4**. Las predicciones de los porcentajes de adulteración fueron muy satisfactorias con ambos tipos de aceites. En ambos casos, los parámetros estadísticos se evaluaron mediante el error relativo de predicción (REP%) y el error cuadrático medio de predicción (RMSEP). Estos valores fueron similares para ambos tipos de adulteraciones, siendo de un 15,4% y un 1,8%, respectivamente, para el caso de los EVOO adulterados con OO y un 15,5% y 1,4%, para el caso de los EVOO adulterados con POO.

Tabla 1.4. Predicciones obtenidas y parámetros estadísticos de los resultados cuantitativos del análisis relativos al porcentaje de adulteración de EVOO con OO y POO utilizando U-PLS.

Tipo de adulterante	Muestra	Concentración nominal (% v:v)	Concentración predicha (% v:v)	Recuperación (%)	
OO	OO 1	33,3	31,0	93,1	
		16,7	16,1	96,3	
	OO 2	33,3	32,7	98,4	
		16,7	18,0	107,7	
	Rec \pm SD				99 \pm 6
	RMSEP				1,3
	REP (%)				15,5
POO	POO 1	33,3	32,4	97,2	
		16,7	20,0	119,9	
		10,0	11,6	116,1	
		5,0	4,6	91,3	
		7,0	6,4	91,4	
	POO 2	33,3	29,9	89,9	
		16,7	18,9	113,2	
		10,0	8,3	83,5	
		5,0	3,5	70,6	
		7,0	6,5	92,8	
		Rec \pm SD			
RMSEP				1,8	
REP (%)				15,4	

Posteriormente, con objeto de comprobar la robustez del método se prepararon 10 muestras ciegas y se analizaron mediante U-PLS. En la **Figura 1.24** se representa el porcentaje predicho mediante U-PLS frente al nominal, observándose una muy buena correlación para ambos tipos de aceite adulterante.

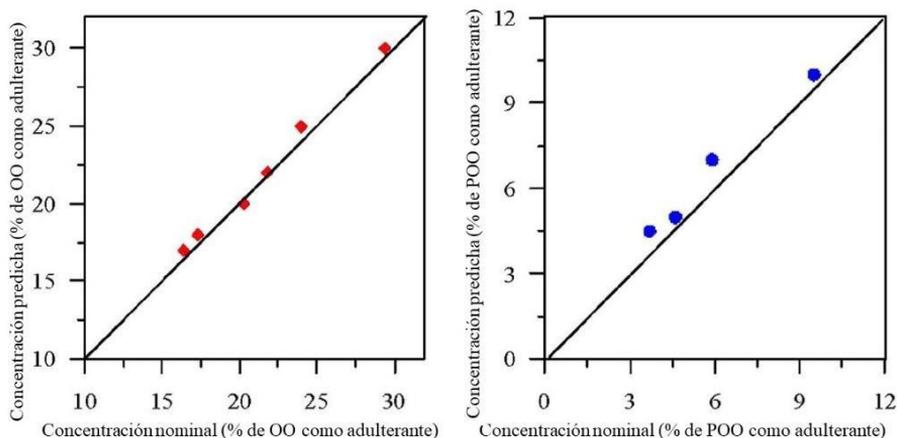


Figura 1.24. Representación del porcentaje de adulteración predicho para muestras ciegas frente al porcentaje nominal mediante U-PLS.

Adulterante (■) aceite de oliva (OO); (●) aceite de orujo (POO).

Por otro lado, el conjunto de 10 muestras ciegas fue procesado y optimizado, según los modelos DA-UPLS y U-PLS, para su análisis a través del análisis discriminante, previa visualización de las EEMs para catalogar el tipo de adulteración del que se trata, pudiendo predecir, con arreglo a los criterios estadísticos utilizados para este modelo, si los aceites estaban adulterados o no en el 100 % de los casos.

4. REFERENCIAS

- [1] M. Vasconcelos, L. Coelho, A. Barros, and J. M. M. M. de Almeida, “Study of adulteration of extra virgin olive oil with peanut oil using FTIR spectroscopy and chemometrics,” *Cogent Food & Agriculture*, vol. 1, no. 1, p. 1018695.
- [2] A. Rohman and Y. B. C. Man, “Fourier transform infrared (FTIR) spectroscopy for analysis of extra virgin olive oil adulterated with palm oil,” *Food Research International*, vol. 43, no. 3, pp. 886–892, 2010.
- [3] L. Sánchez-Hernández, M. Castro-Puyana, M. Luisa Marina, and A. L. Crego, “Determination of betaines in vegetable oils by capillary electrophoresis tandem mass spectrometry—application to the detection of olive oil adulteration with seed oils,” *Electrophoresis*, vol. 32, no. 11, pp. 1394–1401, 2011.
- [4] L. Sánchez-Hernández, M. L. Marina, and A. L. Crego, “A capillary electrophoresis–tandem mass spectrometry methodology for the determination of non-protein amino acids in vegetable oils as novel markers for the detection of adulterations in olive oils,” *Journal of Chromatography A*, vol. 1218, no. 30, pp. 4944–4951, 2011.
- [5] M. G. Bagur-González, E. Pérez-Castaño, M. Sánchez-Viñas, and D. Gázquez-Evangelista, “Using the liquid-chromatographic-fingerprint of sterols fraction to discriminate virgin olive from other edible oils,” *Journal of Chromatography A*, vol. 1380, pp. 64–70, 2015.

- [6] H. Jabeur, A. Zribi, J. Makni, A. Rebai, R. Abdelhedi, and M. Bouaziz, “Detection of Chemlali Extra-Virgin Olive Oil Adulteration Mixed with Soybean Oil, Corn Oil, and Sunflower Oil by Using GC and HPLC,” *Journal of Agricultural and Food Chemistry*, vol. 62, no. 21, pp. 4893–4904, 2014.
- [7] E. Sikorska, I. Khmelinskii, and M. Sikorski, “Analysis of olive oils by fluorescence spectroscopy: methods and applications,” *Olive oil-constituents, quality, health properties and bioconversions*, pp. 63–88, 2012.
- [8] E. v. Avramidou, A. G. Doulis, and P. v. Petrakis, “Chemometrical and molecular methods in olive oil analysis: A review,” *Journal of food processing and preservation*, vol. 42, no. 11, p. e13770, 2018.
- [9] M. Meenu, Q. Cai, and B. Xu, “A critical review on analytical techniques to detect adulteration of extra virgin olive oil,” *Trends in Food Science and Technology*, vol. 91. Elsevier Ltd, pp. 391–408, 2019.
- [10] H. Zaroual, C. Chénè, E. M. el Hadrami, and R. Karoui, “Application of new emerging techniques in combination with classical methods for the determination of the quality and authenticity of olive oil: a review,” *Critical Reviews in Food Science and Nutrition*, vol. 62, no. 16, pp. 4526–4549. 2022.
- [11] C. Tang, J. Qiao, Y. Wen, Z. Zeng, S. Shao, and S. Dong, “Quality control of woody edible oil: The application of fluorescence spectroscopy and the influencing factors of fluorescence,” *Food Control*, vol. 142, p. 109275, 2022.

- [12] F. Lia, J. P. Formosa, M. Zammit-Mangion, and C. Farrugia, “The first identification of the uniqueness and authentication of Maltese extra virgin olive oil using 3D-fluorescence spectroscopy coupled with multi-way data analysis,” *Foods*, vol. 9, no. 4, p. 498, 2020.
- [13] D. F. al Riza, N. Kondo, V. K. Rotich, C. Perone, and F. Giametta, “Cultivar and geographical origin authentication of Italian extra virgin olive oil using front-face fluorescence spectroscopy and chemometrics,” *Food Control*, vol. 121, p. 107604, 2021.
- [14] H. Zaroual, E. M. el Hadrami, and R. Karoui, “A preliminary study on the potential of front face fluorescence spectroscopy for the discrimination of Moroccan virgin olive oils and the prediction of their quality,” *Analytical Methods*, vol. 13, no. 3, pp. 345–358, 2021.
- [15] M. Cabrera-Bañegil, D. Martín-Vertedor, E. Boselli, and I. Durán-Merás, “Control of olive cultivar irrigation by front-face fluorescence excitation-emission matrices in combination with PARAFAC,” *Journal of Food Composition and Analysis*, vol. 69, pp. 189–196, 2018.
- [16] A. Dankowska, M. Malecka, and W. Kowalewski, “Discrimination of edible olive oils by means of synchronous fluorescence spectroscopy with multivariate data analysis,” *Grasas y Aceites*, vol. 64, no. 4, pp. 425–431, 2013.
- [17] A. Dankowska, M. Malecka, and W. Kowalewski, “Discrimination of edible olive oils by means of synchronous fluorescence spectroscopy with multivariate data analysis,” *Grasas y Aceites*, vol. 64, no. 4, pp. 425–431, Jul. 2013.

- [18] K. Wójcicki *et al.*, “Spectroscopic techniques and chemometrics in analysis of blends of extra virgin with refined and mild deodorized olive oils,” *European Journal of Lipid Science and Technology*, vol. 117, no. 1, pp. 92–102, 2015.
- [19] N. Tena, D. L. García-gonzález, and R. Aparicio, “Evaluation of virgin olive oil thermal deterioration by fluorescence spectroscopy,” *Journal of Agricultural and Food Chemistry*, vol. 57, no. 22, pp. 10505–10511, 2009.
- [20] MathWorks, “MATLAB R2008a versión 7.6.0.324, MathWorks, Natick, MA, USA, 2008.” Natick, 2008.
- [21] R. Bro, “Multiway analysis in the food industry. Models, algorithms and applications,” *Ph.D. dissertation, University of Amsterdam, Amsterdam*, 2001.
- [22] A. C. Olivieri, H. L. Wu, and R. Q. Yu, “MVC2: A MATLAB graphical interface toolbox for second-order multivariate calibration,” *Chemometrics and Intelligent Laboratory Systems*, vol. 96, no. 2, pp. 246–251. 2009.
- [23] A. Olivieri and G. M. Escandar, *Practical three-way calibration*. Elsevier, 2014.
- [24] E. K. Kemsley, “A genetic algorithm (GA) approach to the calculation of canonical variates (CVs),” *TrAC Trends in Analytical Chemistry*, vol. 17, no. 1, pp. 24–34, 1998.
- [25] A. Muñoz de la Peña, A. Espinosa Mansilla, D. González Gómez, A. C. Olivieri, and H. C. Goicoechea, “Interference-free analysis using three-way fluorescence data and the parallel factor model. Determination of fluoroquinolone antibiotics in human serum,” *Analytical chemistry*, vol. 75, no. 11, pp. 2640–2646, 2003.

- [26] N. B. Kyriakidis and P. Skarkalis, "Fluorescence spectra measurement of olive oil and other vegetable oils," *Journal of AOAC International*, vol. 83, no. 6, pp. 1435–1439, 2000.
- [27] Y. Endo, R. Usuki, and T. Kaneda, "Antioxidant effects of chlorophyll and pheophytin on the autoxidation of oils in the dark. II. The mechanism of antioxidative action of chlorophyll," *Journal of the American Oil Chemists' Society*, vol. 62, no. 9, pp. 1387–1390, 1985.
- [28] J. W. Heaton and A. G. Marangoni, "Chlorophyll degradation in processed foods and senescent plant tissues," *Trends in Food Science & Technology*, vol. 7, no. 1, pp. 8–15, 1996.
- [29] R. Bro and H. A. L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 17, no. 5, pp. 274–286, 2003.
- [30] F. Guimet, R. Boqué, and J. Ferré, "Cluster Analysis Applied to the Exploratory Analysis of Commercial Spanish Olive Oils by Means of Excitation–Emission Fluorescence Spectroscopy," *Journal of agricultural food chemistry*, vol. 52, no. 22, pp. 6673–6679, 2004.
- [31] F. Guimet, J. Ferré, R. Boqué, and F. X. Rius, "Application of unfold principal component analysis and parallel factor analysis to the exploratory analysis of olive oils by means of excitation–emission matrix fluorescence spectroscopy," *Analytica Chimica Acta*, vol. 515, no. 1, pp. 75–85, 2004.

- [32] F. Guimet, J. Ferré, and R. Boqué, “Rapid detection of olive–pomace oil adulteration in extra virgin olive oils from the protected denomination of origin ‘Siurana’ using excitation–emission fluorescence spectroscopy and three-way methods of analysis,” *Analytica Chimica Acta*, vol. 544, no. 1, pp. 143–152, 2005.
- [33] F. Guimet, J. Ferré, and R. Boqué, “Study of oils from the protected denomination of origin ‘Siurana’ using excitation–emission fluorescence spectroscopy and three-way methods of analysis,” *Grasas y Aceites*, vol. 56, no. 4, pp. 292–297, 2005.
- [34] F. Guimet, R. Boqué, and J. Ferré, “Application of non-negative matrix factorization combined with Fisher’s linear discriminant analysis for classification of olive oil excitation–emission fluorescence spectra,” *Chemometrics and Intelligent Laboratory Systems*, vol. 81, no. 1, pp. 94–106, 2006.
- [35] F. Guimet, J. Ferré, R. Boqué, M. Vidal, and J. Garcia, “Excitation–emission fluorescence spectroscopy combined with three-way methods of analysis as a complementary technique for olive oil characterization,” *Journal of Agricultural and Food Chemistry*, vol. 53, no. 24, pp. 9319–9328, 2005.
- [36] T. Galeano Díaz, I. Durán Merás, C. Arturo Correa, B. Roldán, and M. Isabel Rodríguez Cáceres, “Simultaneous Fluorometric Determination of Chlorophylls a and b and Pheophytins a and b in Olive Oil by Partial Least-Squares Calibration,” *Journal of Agricultural and Food Chemistry*, vol. 51, no. 24, pp. 6934–6940, 2003.

- [37] A. Fernández Gutiérrez and A. Segura Carretero, "El aceite de oliva virgen tesoro de Andalucía," . Málaga. Servicio de Publicaciones de la Fundación Unicaja, 2009.
- [38] C. M. Sabliov, C. Fronczek, C. E. Astete, M. Khachatryan, L. Khachatryan, and C. Leonardi, "Effects of temperature and UV light on degradation of α -tocopherol in free and dissolved form," *Journal of the American Oil Chemists' Society*, vol. 86, no. 9, pp. 895–902, 2009.
- [39] K. I. Poulli, G. A. Mousdis, and C. A. Georgiou, "Monitoring olive oil oxidation under thermal and UV stress through synchronous fluorescence spectroscopy and classical assays," *Food Chemistry*, vol. 117, no. 3, pp. 499–503, 2009.
- [40] Z. Jenisová, J. Branisa, K. Jomová, and M. Porubská, "Variations of some nutrition values of olive oil by household using," *Journal of Microbiology, Biotechnology and Food Sciences*, vol. 3, p. 221, 2014.
- [41] N. Tena, D. L. García-González, and R. Aparicio, "Evaluation of Virgin Olive Oil Thermal Deterioration by Fluorescence Spectroscopy," *Journal of Agricultural and Food Chemistry*, vol. 57, no. 22, pp. 10505–10511, 2009.
- [42] D. M. Haaland and E. V Thomas, "Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information," *Analytical chemistry*, vol. 60, no. 11, pp. 1193–1202, 1988.
- [43] D. M. Haaland and E. V Thomas, "Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data," *Analytical Chemistry*, vol. 60, no. 11, pp. 1202–1208, 1988

**CAPÍTULO 2. SEGUIMIENTO Y CONTROL DEL
ESTADO DE MADURACIÓN DE CIRUELAS A TRAVÉS
DE SUS MATRICES DE FLUORESCENCIA Y DEL
PROCESAMIENTO DE IMÁGENES DIGITALES**

1. INTRODUCCIÓN

Como se ha comentado en el *apartado 1.2.3* de la introducción de esta tesis, uno de los factores fundamentales que condicionan la calidad de las frutas climatéricas, como las ciruelas, así como su vida útil, aceptabilidad y salida comercial de cara al mercado, es el control del estado de maduración durante la fase de producción. Esto es debido a que una cosecha demasiado prematura produce un descenso en la calidad percibida a través de análisis sensoriales y de los parámetros de calidad, y una cosecha demasiado tardía puede comprometer la vida útil del producto durante la frigoconservación, debido a la aparición de daños por magulladuras o a la aceleración de la tasa de producción de etileno, lo que precipita la senescencia de la fruta tal y como han puesto de manifiesto Guerra M. et al. [1].

De este modo, para un tratamiento postcosecha óptimo y una máxima aceptabilidad por parte del cliente, existen numerosas guías que establecen una serie de valores de compromiso de parámetros de calidad, como la acidez titulable, pH, contenido en sólidos solubles, color y textura, siendo estos muy variables dependiendo de la variedad de ciruela estudiada.

Estos índices, o parámetros, de calidad han demostrado ser muy útiles para categorizar el grado de madurez tanto en el árbol, como para caracterizar la evolución de la calidad organoléptica durante el proceso de frigoconservación. No obstante, dichos indicadores guardan escasa relación con aspectos fisiológicos del fruto o con características organolépticas más complejas. Por este motivo, se tiende a buscar indicadores complementarios que permitan obtener más información acerca de las características fisiológicas del fruto como puede ser la producción de etileno o la intensidad respiratoria [2].

En este ámbito, el estudio de la evolución de la concentración de pigmentos en las frutas ha sido investigado como otro posible indicador relacionado con la madurez fisiológica en el árbol y puede tener especial relevancia como criterio de selección de la cosecha [3]. Unos de los pigmentos más representativos en el cambio del estado de maduración son las clorofilas, cuyo decaimiento a lo largo del periodo de maduración ha sido ampliamente estudiado [4]. Sin embargo, los procedimientos tradicionales de determinación de clorofilas implican normalmente la extracción de estos pigmentos a través de procedimientos tediosos, largos, costosos y sujetos a múltiples fuentes de errores experimentales [5, 6], por lo que su determinación de manera selectiva, rápida y sencilla supondría un importante avance en la utilización de estos compuestos como indicadores del momento óptimo de recolección.

En primer lugar, cabe destacar que, con respecto a la aplicación de técnicas basadas en la fluorescencia molecular para el control del estado de maduración de frutas en general, y de ciruelas en particular, la bibliografía revisada es escasa. Entre la información más relevante, están descritas aplicaciones para estimar la madurez de cítricos a través de la aplicación de métodos quimiométricos utilizando las EEMs obtenidas de extractos de estos [7] y también para estimar la madurez fenólica de vinos [8]. En las ciruelas, se ha propuesto el uso de la fluorescencia de la fracción fenólica para la determinación de la fecha de cosecha, según los estudios realizados por Cabrera-Bañegil. M. et al. [9] y Monago-Maraña. O. et al. [10]. Otros estudios sobre el uso de técnicas de fluorescencia no invasivas basadas en la fluorescencia de las clorofilas del mango han sido propuestos para la determinación de su grado de madurez [11], o la utilización de sensores manuales multiparamétricos para determinar la maduración de las ciruelas a través de su contenido en clorofilas [12].

En el estudio desarrollado en esta memoria se propone el uso de la fluorescencia molecular y de imágenes digitales, como técnicas no destructivas, para la caracterización del estado de maduración de ciruelas de la variedad *Friar*, obtenidas de una plantación de la provincia de Badajoz, durante su fase de producción, en el período de mayo a agosto de 2018. La información obtenida se analizará con diferentes algoritmos quimiométricos.

Utilizando una fibra óptica acoplada a un espectrofluorímetro se obtienen las matrices de excitación-emisión (EEMs), registradas sobre la piel de las ciruelas, sin un tratamiento previo de la muestra. Se utilizan herramientas quimiométricas con fines exploratorios y de cuantificación de clorofilas. Los algoritmos empleados fueron PARAFAC [13], para el análisis exploratorio, y PARAFAC, N-PLS [14], U-PLS [15] y MCR-ALS [16] para la cuantificación de clorofilas.

Por otro lado, la segunda vía empleada para la obtención de datos multivariantes, relacionados con el contenido en clorofilas, ha sido el procesamiento de imágenes digitales de fotografías de las ciruelas tomadas en condiciones de luminosidad controladas con la cámara de un *smartphone*.

Las imágenes tomadas con dispositivos digitales transforman la radiación electromagnética de los tres colores aditivos del espectro (rojo, verde y azul), que captan a través de sus sistemas ópticos, en señales eléctricas fundamentadas en código binario de 8 bits. De este modo, se pueden obtener hasta 167.777.216 valores diferentes para los colores aditivos rojo (R), verde (G) y azul (B), distribuidos en el espacio conformado por los 256 valores posibles para cada eje. Este sistema de coordenadas conforma el espacio de color RGB, que dependiendo de los parámetros que se quieran analizar para la imagen en cuestión pueden

tomarse como: escalares, si se toma el valor medio de cada color; vectores, si se usan los datos de frecuencias de distintas intensidades en forma de histograma; o incluso matrices, si se toman los datos de intensidad de cada pixel por separado para los colores analizados.

A partir del espacio RGB pueden surgir nuevos sistemas de coordenadas fundamentados en otros criterios como la percepción del ojo humano, como pueden ser los espacios HSI (del inglés '*Hue Saturation, Intensity*') o HSL (del inglés '*Hue Saturation, Lightness*') o espacios que no dependan del dispositivo, como pueden ser el CIE $L^*a^*b^*$ o el CIE $L^*U^*B^*$. Todos estos espacios aparecen de manera recurrente en la bibliografía, no obstante, no han sido utilizados en los estudios que se han llevado a cabo en esta tesis.

Con respecto al procesamiento de imágenes digitales, existe una amplia bibliografía empleando distintos espacios de color, en combinación con métodos quimiométricos de análisis de información tanto univariante, como de dos y tres vías, para distintos tipos de aplicaciones, tanto en el ámbito agroalimentario como en el control de parámetros medioambientales.

Entre todos estos métodos, los más extendidos son los de primer orden, que emplean como datos los histogramas de los distintos espacios de color, ya que estos ofrecen una información muy precisa y poco sesgada de las características superficiales de la región estudiada con respecto a su color, por lo que para una región de interés (ROI) representativa de las características superficiales de la muestra objeto de estudio, se puede extraer una gran información respecto a los colores predominantes y el contraste que existe entre ellos. Los histogramas son una herramienta muy potente bajo las que se puede aplicar métodos multivariantes de análisis exploratorio y de regresión, aunque también es de destacar que existen

referencias bibliográficas en las cuales se combinan otras características superficiales como puede ser la textura, para formar tensores representativos de primer orden que abarquen una información más amplia y específica, como es el caso del estudio realizado por Zhuang. J. et al. [17] en la maduración de plátanos. Incluso, se han diseñado métodos de tres vías fundamentados en parámetros como la espectroscopía NIR, el espacio de color CIE L*a*b* y la cinética a lo largo del tiempo para el estudio del deterioro de maderas [18].

Las aplicaciones más representativas del uso de los histogramas de espacios de color con fines exploratorios, discriminantes y de cuantificación, se resumen en la **Tabla 2.1**.

Tabla 2.1. Resumen de los métodos de procesamiento de imagen digital (DIP) usados para la determinación de la calidad y detección de fraudes en matrices alimentarias a través de técnicas mínimamente invasivas.

Producto	Número de muestras	Espacio de color	Análisis de datos	Objetivos	Resultados	Referencias
Zumo de uva	22	Histogramas de RGB	PCA, ICA y MCR-ALS	Identificación y cuantificación de compuestos bioactivos.	Detección de compuestos fenólicos, antocianinas monoméricas y contenido en taninos con unos coeficientes R de 0,99; 0,97 y 0,98 respectivamente.	[19]
Colorante en caramelos	238	Histogramas de RGB	PLS	Determinación de colorante <i>azo allura rojo</i> en caramelos.	Obtención de un método de regresión preciso, lineal, sensible y no sesgado con RMSEP de 6,1 mg·Kg ⁻¹	[20]

Tabla 2.1. Continuación

Producto	Número de muestras	Espacio de color	Análisis de datos	Objetivos	Resultados	Referencias
Mangos	34	Índices de RGB	PLS	Correlación de datos de imagen hiperespectral y análisis de color con parámetros relacionados con el índice de madurez como el contenido en sólidos solubles, acidez titulable, clorofila a, clorofila b, clorofilas totales y carotenoides.	Correlaciones dadas por R^2 para la validación del modelo de 0,75, 0,57 y 0,80 para clorofila a, clorofila b y clorofilas totales.	[21]
Hamburguesas de pollo	74	Histogramas de escala de grises RGB y HSI	SPA-LDA y PLS	Cuantificación e identificación de adulteraciones en el contenido en grasas de hamburguesas de pollo.	Resultados en la cuantificación con un R^2 de 0,95, RMSEP de 2,01% y REP de 7,26%	[22]

Tabla 2.1. Continuación

Producto	Número de muestras	Espacio de color	Análisis de datos	Objetivos	Resultados	Referencias
Vino	Muestras puras: 25 Muestras adulteradas: 78 Muestras de validación: 52	Histogramas de RGB	PCA y PLS	Detección de adulteraciones de vinos gran reserva con vinos jóvenes	El análisis exploratorio con PCA permitió observar la diferencia entre las muestras de vino gran reserva con vino joven y la regresión PLS permitió obtener predicciones con un RMSEP de 1,6%.	[23]
Mango, tomate, naranja y manzanas	Muestras entrenamiento: 100 muestras maduras y 100 muestras verdes. Muestras de prueba: 20 maduras y 20 verdes.	Histogramas RGB, HSL y CIE L*a*b*	SVM	Determinación del grado de maduración para distintos tipos de frutas.	El mejor espacio de color para aplicar el algoritmo SVM fue el espacio HSL, con una precisión del 0,80 y un valor de F de 0,78.	[24]

Tabla 2.1. Continuación

Producto	Número de muestras	Espacio de color	Análisis de datos	Objetivos	Resultados	Referencias
Ketchup	63 muestras adulteradas y 17 muestras no adulteradas	Histogramas RGB, HSI y escala de grises.	PCA, SPA-LDA, iSPA-PLS	Determinación de adulteraciones en ketchup con el colorante Sudan I.	Clasificación correcta de todas las muestras adulteradas y predicción del nivel de adulteración con PLS con un RMSEP de 11,64 mg·Kg ⁻¹ .	[25]
Plátanos	441 muestras de diferentes estados de maduración	Tensor de media de espacios de color H,S,a*,b*,c y h, así como características de textura y forma.	NB, LDA, SVM y PCA	Clasificación de los diferentes estados de maduración del plátano.	Se consiguió clasificar con una precisión del 100% para LDA las muestras de validación para cada uno de los cuatro estados de maduración caracterizados.	[17]

ICA: *Independent component analysis* (Análisis por componentes independientes); iSPA-PLS: *partial least squares coupled with successive projections algorithm for interval selection* (Mínimos cuadrados parciales acoplado a algoritmo de sucesiones progresivas para un intervalo seleccionado MCR-ALS: *Multivariate curve resolution-alternating least squares* (Resolución multivariante de curvas mediante mínimos cuadrados alternantes); NB: *Navie Bayes* (Clasificador Bayesiano Ingenuo) LDA: *Linear discriminant analysis* (Análisis lineal discriminante); PCA: *Principal component analysis* (Análisis por componentes principales); PLS: *Partial least squares* (mínimos cuadrados parciales); SPA-LDA: *successive projections algorithm coupled with linear discriminant analysis* (Algoritmo de proyecciones sucesivas acoplado a análisis lineal discriminante) SVM: *suporting vector machine* (Máquinas de vector de soporte).

Tras lo anteriormente expuesto, los objetivos que se persiguen en este capítulo y para los que se expone la metodología y resultados analíticos a continuación fueron:

Para el caso concreto del análisis del estado de maduración de las ciruelas a través de la fluorescencia nativa de las mismas, obtener las EEMs mediante fibra óptica en la superficie, identificar y caracterizar el perfil fluorescente de las clorofilas registradas a través de las EEMs de la superficie de la piel de las ciruelas, y el análisis cuantitativo del contenido en clorofilas a través de métodos de regresión multivariante de segundo orden como PARAFAC, U-PLS, N-PLS y MCR-ALS.

En cuanto a la aplicación del procesamiento de imágenes digitales al análisis multivariante de clorofilas, el objetivo fue demostrar la viabilidad de dichos métodos para el desarrollo de metodologías de control del estado de maduración en ciruelas, sin tratamientos previos de las muestras o el uso de equipos sofisticados de análisis. Para este fin, se utilizó el espacio de color RGB, del cual se extrajo la información analítica de distintos modos para obtener datos univariantes, datos multivariantes de primer orden y datos de segundo orden, a través de los cuales se caracterizaron las clorofilas y su evolución a lo largo del periodo de maduración mediante PCA y PARAFAC y se cuantificaron las clorofilas aplicando regresión lineal univariante por mínimos cuadrados, PLS, U-PLS y N-PLS

2. METODOLOGÍA

2.1. Muestreo

Para el estudio del proceso de maduración desde las etapas tempranas del crecimiento del fruto en el árbol hasta etapas próximas a la madurez organoléptica, se emplearon ciruelas de la variedad *Friar* cosechadas de una plantación de la localidad de Badajoz. Las muestras fueron cosechadas desde la última semana de mayo hasta agosto de 2018, siendo recolectadas en tandas semanales con muestras recogidas aleatoriamente de dicha plantación durante el proceso de maduración. De dichas tandas se seleccionaron, a su vez, tres muestras aleatorias para proceder a su análisis por fluorescencia y mediante el procesamiento de imágenes digitales. También se llevó a cabo la extracción de los pigmentos, tanto en la piel, como en la pulpa, para su posterior cuantificación mediante (HPLC-FLD). Los datos cuantitativos de clorofilas totales (suma de clorofila a y clorofila b) obtenidos, fueron posteriormente empleados como datos de referencia para la validación de los modelos quimiométricos construidos sobre los datos de segundo orden de las EEMs obtenidas mediante fibra óptica y sobre los datos univariantes y multivariantes de primer y segundo orden obtenidos del procesamiento de las imágenes de las ciruelas en el espacio de color RGB.

2.2. Optimización del proceso de extracción de pigmentos

Para la extracción y posterior análisis cromatográfico del contenido en clorofila a y clorofila b en ciruelas, se ensayaron tres métodos diferentes reflejados en la bibliografía como son el expuesto por Hart y Scott en 1995 [26] para la determinación cromatográfica de clorofilas y carotenoides en diferentes matrices alimentarias entre las que se encuentran las frutas de hueso, el método de Wright y Kader de 1997 [27] diseñado para la extracción de carotenoides de melocotones y caquis para el seguimiento de

su estado de maduración y el propuesto por Mellado-Ortega y Hornero-Méndez en 2018 [28], para estudiar el efecto de la esterificación de la luteína en la distribución de carotenoides en el germen y el endospermo de granos de *tritordeum*.

Los mejores resultados se obtuvieron con el procedimiento descrito por Hart y Scott [26], el cuál fue modificado para que se ajustase a los parámetros cromatográficos fijados y a los requerimientos de reducción de gasto en disolventes para el procedimiento de extracción, consiguiendo extractos cuyas concentraciones estuvieran dentro del rango lineal del método de patrón externo construido sobre una recta multipatrón de 1 a 10 $\mu\text{g}\cdot\text{g}^{-1}$ de concentración.

El método propuesto se aplica tanto al análisis de la pulpa como de la piel y fue el siguiente: se pesan 4 g de la piel recién obtenida, o 10 g de pulpa, y se añaden 15,0 mL de la disolución extractante, formada por tetrahidrofurano (THF):metanol (MeOH) (1:1) con butilhidroxitolueno (BHT) al 0,1 %. La extracción se repite dos veces más y los extractos una vez combinados se filtran a vacío con papel de filtro nº1242 (Filter-Lab, Anovia, Barcelona, España) en un embudo Büchner. El eluato es reextraído con 15 mL de éter de petróleo y 15 mL de NaCl al 10%. Una vez separadas las fases, la fase orgánica se evapora a sequedad en un rotavapor a 35 °C y el residuo se redissuelve en 5 mL de la disolución extractante de THF:MeOH. A continuación, 1 mL de esta disolución se diluye con la disolución extractante THF:MeOH en un factor de 1:250 para poder ser analizado dentro del rango de sensibilidad del método cromatográfico. Finalmente, los extractos se filtraban con un filtro de 0,25 μm *Cromafil filters* (Düren, Alemania) previamente a su inyección en el sistema cromatográfico. Los otros 4,0 mL restantes son saponificados con 4 mL de una disolución metanólica al 10 % de KOH, para posteriormente ser lavados tres veces hasta pH neutro con 4 mL de disolución de NaCl al 10%

y reextraídos en 12 mL de éter de petróleo tres veces para su posterior evaporación a sequedad en rotavapor y redisolución en 1 mL de fase extractante THF:MeOH (1:1) con BHT al 0,1% para su filtración a través de filtro de 0,25 μm e introducción al sistema cromatográfico para el análisis de clorofilas con el detector FLD con las longitudes de onda de excitación y emisión de 420 y 660 nm respectivamente.

2.3. Optimización del método cromatográfico de referencia

Para la optimización de los parámetros cromatográficos se utilizaron disoluciones patrón de clorofila a, y clorofila b, las cuales, fueron seleccionadas para los objetivos planteados dadas su íntima relación con el estado de maduración y a su elevado rendimiento cuántico de fluorescencia, el cual es sensiblemente superior al de otros pigmentos indicadores del estado de maduración como son los carotenoides o las antocianinas [29]–[33].

El equipo utilizado ha sido un cromatógrafo UFLC Shimadzu Prominence LC-20AD, compuesto por un módulo de distribución del disolvente con una bomba cuaternaria, una válvula de gradiente LC-20AD/20AT y un módulo de muestreo automático SIL-20A, SIL-20AC, SIL-20AHT, SIL-20ACHT, con un detector fotométrico de diodos UV-visible SPD-M20A y un detector espectrofluorimétrico RF-20A/20Axs, con una columna Kinetex C18 (150 \times 4,6 mm, 5 μm) termostatzada con un horno de columna CTO-10ASvp. Todos ellos proporcionados por la casa comercial Shimadzu.

Para la optimización del método cromatográfico de referencia se usaron patrones de clorofilas (clorofila a y clorofila b) obtenidos de la casa comercial *Sigma-Aldrich*, y sus disoluciones de trabajo fueron preparadas disolviendo 1 mg de cada clorofila en 25 mL de acetona obtenida de la casa

comercial Merck (Darmsadt, Alemania) y almacenadas en la oscuridad a -4 °C.

Para la selección de las longitudes de onda para el análisis de clorofilas con el detector de fluorescencia (FLD) del cromatógrafo en los cuatro canales disponibles, se tuvieron en cuenta los datos bibliográficos resumidos en la **Tabla 2.2** para localizar los máximos de excitación y emisión de los espectros de fluorescencia de las clorofilas según las referencias indicadas en la tabla, *a* [34], *b* [35] y *c* [36].

El método cromatográfico de referencia seleccionado para su posterior optimización fue el propuesto por Hart y Scott en 1995 [26], el cuál fue modificado posteriormente en 2009 por Días M. et al. [37] para el análisis tanto de clorofilas, como de carotenoides, en diversas frutas entre las que se podía destacar el melocotón, el cuál es una fruta de hueso de características composicionales similares a la que puede tener la ciruela, por lo cual se optó por iniciar por este método de partida.

Tabla 2.2. Longitudes de onda de absorción y de fluorescencia para los compuestos seleccionados para el método cromatográfico de referencia para el seguimiento del estado de maduración de las ciruelas.

Clorofilas	Disolvente	λ_{abs}	ϵ	Disolvente	λ_{ex}	λ_{em}
		nm	[L mol ⁻¹ cm ¹]		nm	nm
Clorofila <i>a</i> ^{abc}	acetona	430	94700	éter	436	668
				acetona	405	669
		663	75000	acetona/agua 9:1	430	669
Clorofila <i>b</i> ^{abc}	acetona	455	131000	éter	436	648
				acetona	405	652
		645	47100	acetona/agua 9:1	458	653
Feofitina <i>a</i> ^{abc}	acetona	409	101800	éter	436	673
				acetona/agua	406	671
		666	44500	9:1		
Feofitina <i>b</i> ^{abc}	acetona	434	145000	éter	436	661
				acetona/agua	435	658
		654	27800	9:1		
Feoforbida <i>a</i> ^{abc}	acetona	409	119200	-	-	-
		667	55200	-	-	-

Para llevar a cabo la separación cromatográfica una vez optimizados los parámetros a la separación de clorofilas en la matriz alimentaria estudiada, se empleó una columna analítica Kinetex C18 (150 x 4,6mm, 5 μ m) termostaticada a 30°C, con un volumen de inyección de 5 μ L. La fase móvil estuvo compuesta de acetonitrilo y una disolución metanólica de acetato de amonio 5mM con una proporción (v/v) de 95:5, con un flujo de 1,0 mL \cdot min⁻¹. La detección se lleva a cabo con el detector de fluorescencia RF-20A/20Axs con longitudes de onda de excitación y emisión de 420 y

660 nm, respectivamente. Para la cuantificación se empleó el método de calibración externa utilizando como señal analítica el área de los picos obtenidos a los tiempos de retención de 15,2 y 8,6 min para la *clorofila a* y la *clorofila b*, respectivamente.

2.4. Obtención de las EEMs

Se optimizaron las variables instrumentales para la obtención de las EEMs tanto en la superficie de las ciruelas utilizando una fibra óptica como en los extractos de la piel de ciruelas y de la pulpa mediante fluorescencia front-face. Las rendijas de los monocromadores de excitación y emisión se fijaron a 2,5 y 5 nm respectivamente, y las longitudes de onda de excitación se variaron entre 360 y 500 nm con incrementos de 5 nm, y los espectros de emisión se registraron entre 600 y 700 nm con intervalos de 1 nm. La velocidad de registro fue fijada a 300 nm·min⁻¹ y el tiempo total de registro fue aproximadamente de 5 minutos. Para el detector se fijó una sensibilidad de 550 V para registrar la EEMs de los extractos y de 700 V para las EEMs de ciruelas enteras mediante fibra óptica.

2.5. Adquisición de imágenes digitales

La adquisición de las imágenes digitales de las ciruelas para su procesamiento al espacio de color RGB, se hizo en unas condiciones experimentales reproducibles de distancia, nitidez y luminosidad, en las, asegurando así la precisión, fiabilidad y robustez mínimas requeridas para los métodos de análisis de imagen a desarrollar.

Se usó un *smartphone* BQ (Aquarius modelo M5) con un sistema operativo Android versión 5.0 equipado con una cámara de 13 megapixels de resolución y una distancia focal de 2.0, para obtener las fotografías que posteriormente serían procesadas a través de los softwares pertinentes para la extracción de los datos correspondientes a los espacios de color para la caracterización del estado de maduración de las ciruelas.

Por otro lado, para mantener constantes las variables experimentales que afectan a la toma de imágenes, como pueden ser la iluminación externa, la nitidez o la distancia al objeto de referencia, se usó un dispositivo para el aislamiento de la iluminación exterior y el control de la iluminación interna consistente en una caja de dimensiones 25 cm de alto, ancho y largo construida sobre madera conglomerada acabada en blanco y rodeada en su parte superior por una tira de catorce LEDs (Iluminante D65 y 6500 K de temperatura de color), utilizada previamente por Cerrato. M. et al. para la medida del ozono troposférico por análisis de imagen digital mediante su reacción con indigotrisulfonato [38]. Seleccionaron el iluminante D65 debido a su extendido uso en aplicaciones de colorimetría, además de ser el establecido por la comisión internacional de iluminación (CIE) como iluminante estándar en este tipo de aplicaciones [39]. La cámara del smartphone se fijaba en la parte superior de la caja frente al soporte de plástico colocado en la parte inferior del interior de la caja.

2.6. Calibración de segundo orden empleando las EEMs

Para el tratamiento de los datos de segundo orden obtenidos al registrar las EEMs con la fibra óptica, así como de los extractos, se empleó el software Matlab (Matlab R2008b, versión 7.5.0.342) [40], con las rutinas de PARAFAC [41] (<http://www.models.kvl.dk/source/>) y se ha utilizado la interfaz gráfica MVC2 para la representación gráfica de las EEMs y la aplicación de los cálculos de segundo orden con los métodos PARAFAC, U-PLS, N-PLS y MCR-ALS, usadas en el capítulo anterior y citadas previamente en estudios anteriores [42]–[44].

Tal y como se observa en la *Figura 2.1*. Para obtener una información representativa de la fluorescencia de cada una de las 9 tandas que se analizaron a lo largo de todo el periodo de maduración se procedió del siguiente modo.

En primer lugar, se tomaron 3 réplicas de cada tanda de las diferentes semanas. A continuación, se analizaban las cuatro caras de cada réplica a través de la fibra óptica para poder obtener una información representativa de la superficie total de cada ciruela, ya que la maduración se produce de manera irregular entre los extremos y las caras laterales. Posteriormente se extraían los pigmentos de las tres réplicas por duplicado en piel y en la totalidad de la ciruela y se analizaban los extractos mediante fluorescencia *front-face* y se cuantificaban las clorofilas mediante HPLC-FDL. Finalmente, se agrupaban los datos de fluorescencia obtenidos en diferentes cubos de datos para su descomposición mediante PARAFAC.

Para la cuantificación de clorofilas se realizó en primer lugar la regresión de los scores del primer componente de PARAFAC para las EEMs obtenidas de los extractos de la piel y del fruto entero con los valores de clorofilas, posteriormente, se compararon los resultados con los obtenidos por los algoritmos MCR-ALS, U-PLS y N-PLS.

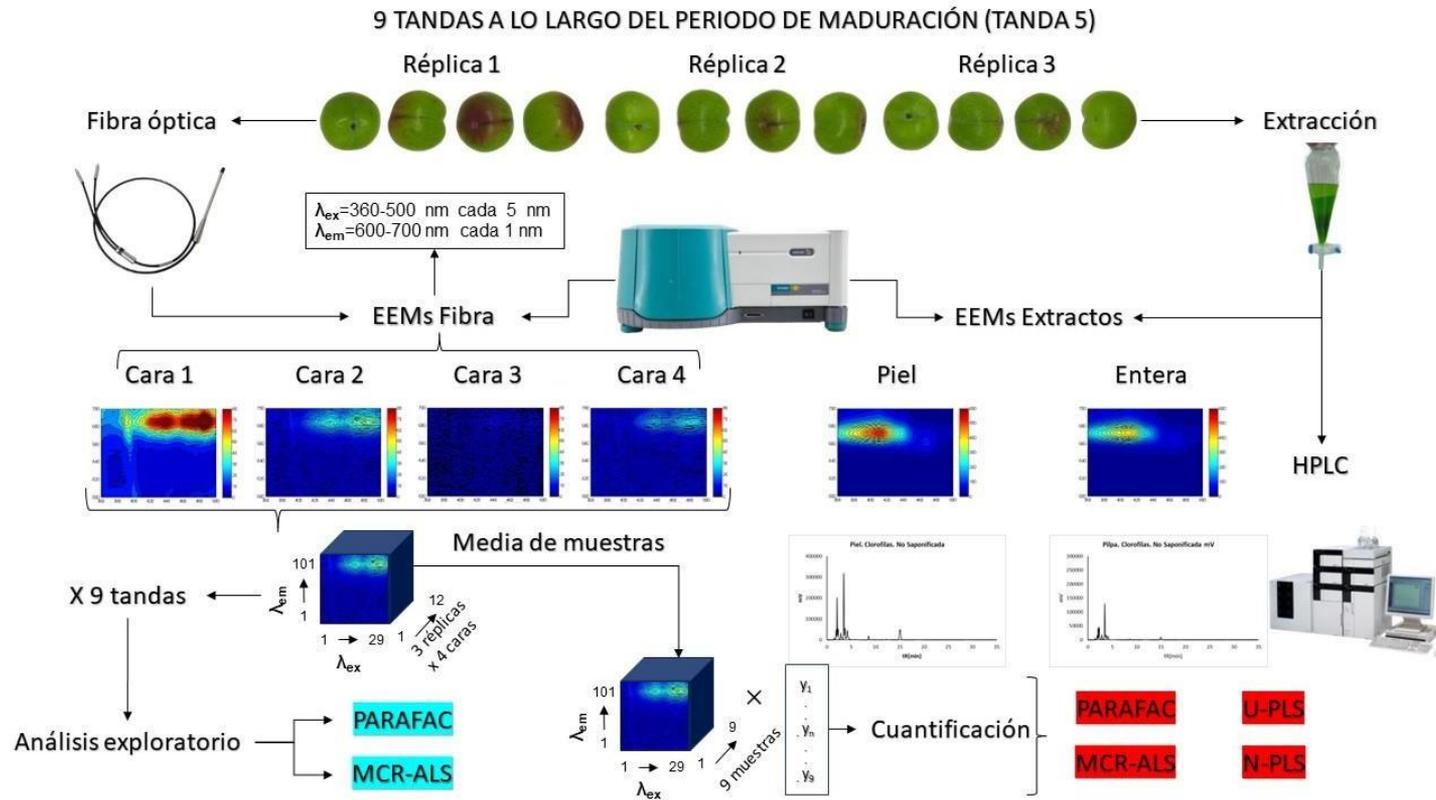


Figura 2.1. Esquema de la adquisición de las EEMs obtenidas con la fibra óptica y posterior tratamiento de los datos

2.7. Tratamiento de las imágenes digitales

Una vez obtenidas las imágenes en formato JPEG, estas fueron transferidas a un ordenador en el que, en primer lugar, se seleccionó la ROI para todas ellas, posteriormente se procesaron para la obtención de los datos correspondientes al espacio de color RGB, y finalmente se ordenaron los datos para convertirlos en escalares (ratios de color R, G y B)), vectores (histogramas de color R, G, B y RGB) y matrices (matrices de color R, G y B). El procedimiento para la extracción de la información puede resumirse a través del esquema mostrado en la **Figura 2.2**

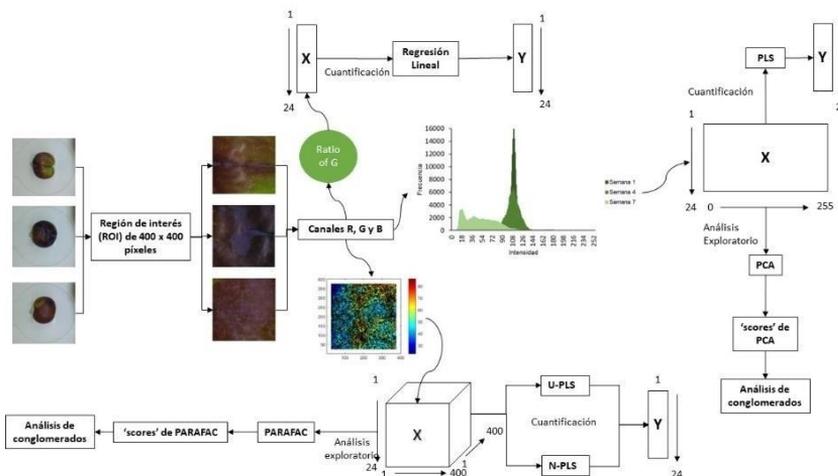


Figura 2.2. Esquema de la adquisición de imágenes y procesamiento de datos obtenidos del espacio de color RGB.

Para la extracción del valor numérico asociado al código de 8 bits del espacio RGB en el tratamiento de imágenes digitales, se empleó el entorno de Matlab (Matlab R2018a, versión 7.5.0.342, Natick, MA, USA) [45] utilizando para este caso el comando “*imread*” para extraer el valor asociado a los 400 x 400 píxeles obtenidos de la ROI de la fotografía. Dicha información fue empleada para obtener los escalares correspondientes a cada color aditivo, rojo (R), verde (G) y azul (B), y poder calcular las ratios

(datos de orden 0), que fueron agrupados en una matriz de datos de 400 x 400 para cada pixel, para su posterior tratamiento como datos de segundo orden.

Los cálculos realizados para la regresión univariante con datos de orden 0 fueron realizados utilizando las hojas de cálculo de Microsoft Excel 2016 [46] instalado con el paquete de Microsoft Office 2016. La descomposición factorial de la matriz de segundo orden y la regresión multivariante de los espacios de color de las matrices de segundo orden utilizando los valores de concentraciones de clorofilas obtenidas con el método cromatográfico de referencia, se llevaron a cabo con la interfaz gráfica MVC2 (<http://www.iquir-conicet.gov.ar/descargas/mvc2.rar>), actualizada para la versión de Matlab R2018a.

Por otra parte, la ordenación de los datos correspondientes al espacio de color RGB en histogramas de orden uno, representando la distribución de las 256 posibles intensidades correspondientes a cada color en diagramas de frecuencias, fue realizada utilizando el software **imageJ** (versión 1.53K, Wayne Rasband y colaboradores, Instituto Nacional de Salud de USA, dominio público) (utilizado el 15 de enero de 2022) [47] (<https://imagej.nih.gov/ij/>).

Por último, se realizaron los cálculos con los datos de orden uno, empleando análisis por componentes principales (PCA) de los histogramas y, posteriormente, se realizó la regresión multivariante mediante PLS de los histogramas con los datos cuantitativos de concentración de clorofilas obtenidos del método cromatográfico de referencia. Para ello se utilizó el software The Unscrambler versión 6.11 (CAMO Software AS, Oslo, Noruega) [48].

2.7.1. Análisis univariante de datos de color de las ciruelas

Para llevar a cabo el análisis univariante, en primer lugar, se extrajo el valor medio de cada canal del espacio de color RGB para la matriz íntegra de la ROI de 400 x 400 pixels. Posteriormente, y con objeto de evitar posibles desviaciones en la intensidad de los canales como consecuencia de posibles variaciones en la luminosidad recibida por el objetivo de la cámara digital, se tomó la ratio de dichos canales R, G y B, los cuales se definen según las siguientes ecuaciones para cada uno de ellos:

$$\text{Ratio } R = \frac{R}{R + G + B} \quad (1)$$

$$\text{Ratio } G = \frac{G}{R + G + B} \quad (2)$$

$$\text{Ratio } B = \frac{B}{R + G + B} \quad (3)$$

La evolución de estas ratios, junto a la dispersión de las réplicas de las tandas representativas de la semana de maduración objeto de estudio fue analizada para la caracterización cualitativa del estado de maduración. Posteriormente, se realizó la correlación con la concentración de clorofilas totales determinadas mediante HPLC con el fin de obtener un indicador cuantitativo de la evolución del proceso de maduración, teniendo en cuenta la influencia que estos compuestos tienen sobre el proceso de maduración de las frutas.

2.7.2. Análisis multivariante de datos de color de primer orden

Para el procesamiento y ordenación sistemática de los datos numéricos de intensidades de los valores de R, G y B en los histogramas de color representativos de las imágenes de las ROI, se empleó el software de acceso libre *ImageJ* [47], a través del cual fue posible la obtención de cada uno de los histogramas correspondientes a los valores de los canales R, G, B y RGB en conjunto. De este modo, se extrajeron cada uno de los histogramas correspondientes a los canales seleccionados del espacio RGB para las 3 caras de las 3 réplicas representativas de cada una de las 8 semanas objeto de estudio, para posteriormente sacar el valor medio de las 3 caras obteniendo un único histograma para las 3 réplicas de cada semana. Se obtuvieron así un número total de 24 muestras representativas con intensidades de 0 a 256 para valores de R, G y B, y 768 para el canal conjunto RGB. También se obtuvieron las frecuencias normalizadas de 0 a 160000 (de los 400 x 400 píxels de la ROI) para los canales de R, G y B y de 0 a 480000 (de los 400 x 400 píxels x 3 características de color de la ROI) para el canal conjunto RGB.

El conjunto de datos correspondiente a los histogramas fue empleado para realizar un análisis exploratorio inicial mediante PCA con el fin de explorar la variabilidad de los datos y la detección de los componentes representativos de los histogramas medios, que permitan obtener información cualitativa, para después efectuar un modelo de regresión mediante PLS. Para evaluar la eficacia del modelo se empleó el método de validación cruzada, y los parámetros de calidad establecidos para la validación del modelo fueron el coeficiente de determinación (R^2), el error cuadrático medio de predicción (RMSEP) y el error relativo de predicción (REP), todos ellos obtenidos a través del uso del software The Unscrambler versión 6.11.[48].

2.7.3. Análisis multivariante de datos de color de segundo orden

Los datos de segundo orden extraídos de las imágenes digitales fueron las propias matrices obtenidas del código ‘*imread*’ del software Matlab R2018a, versión 7.5.0.342 [40]. El tratamiento previo de estas ha sido análogo al realizado para los datos de primer orden, obteniendo de este modo las matrices correspondientes a cada una de las caras seleccionadas de cada una de las tres réplicas representativas de las semanas de cosecha. Se obtiene el valor medio de las tres caras, para finalmente obtener una única ‘matriz media’ para cada réplica. Estas matrices se analizan mediante PARAFAC, y posteriormente se realiza la calibración multivariante mediante U-PLS y N-PLS.

3. RESULTADOS Y DISCUSIÓN

3.1. Evolución de la concentración de clorofilas durante el periodo de maduración

La concentración de clorofilas totales (suma de clorofila a y clorofila b) es un excelente indicador del proceso de maduración fisiológica de la ciruela en el árbol, y como tal, su estudio a través del método cromatográfico de referencia indicado en el **apartado 2.3** de este capítulo ha sido de especial interés.

La evolución de los cromatogramas correspondientes a las clorofilas a y b, en relación al estado de maduración de las ciruelas puede observarse en la **Figura 2.3**, mientras que la evolución cuantitativa de la concentración de clorofilas totales frente al período de maduración estudiado puede observarse en la **Figura 2.4**.

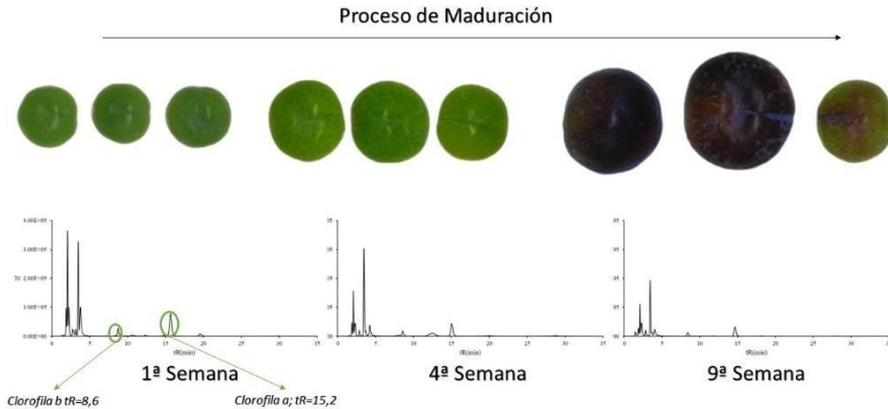


Figura 2.3. Evolución de los cromatogramas correspondientes a la 1^a, a la 4^a y 9^a semana de cosecha de las ciruelas de la variedad ‘Friar’ (fotografías correspondientes encima) correspondientes al análisis de clorofila a y clorofila b en la piel de las ciruelas.

Como se puede observar en la **Figura 2.3**, el proceso de maduración observable a simple vista durante el periodo de cosecha de las ciruelas, caracterizado por la desaparición de las tonalidades verdes intensas y la aparición progresiva de tonalidades rojo, púrpuras en su piel, se ve reflejado en la reducción del contenido tanto en clorofila a (tR = 15,2 min) como en clorofila b (tR = 8,6 min), en los correspondientes cromatogramas de los extractos de la piel de dichas ciruelas.

Por otro lado, la cuantificación de clorofilas totales mediante el método cromatográfico de referencia permitió contrastar, tal y como se observa en la **Figura 2.4**, que el proceso de maduración lleva consigo una disminución del contenido en clorofilas dentro del periodo estudiado, tal y como se ha mencionado en anteriores trabajos citados como el de Vlačić et al. [49], o el de Luo et al. [50].

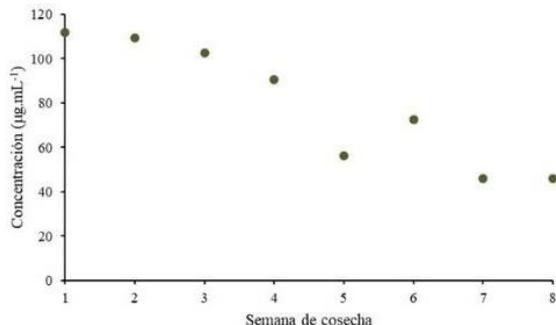


Figura 2.4. Evolución de las concentraciones de clorofilas totales de las muestras de ciruelas referentes a las semanas de cosecha del periodo de maduración elegido.

3.2. Evolución de las EEMs obtenidas de los extractos de las ciruelas

La disminución del contenido en clorofilas también ha sido estudiada en las propias EEMs de los extractos realizados tanto en la totalidad de la ciruela, como en la piel de estas, pudiendo observarse los resultados de la evolución de dichas EEMs a la par que los cromatogramas de estas en ambos extractos en la **Figura 2.5**.

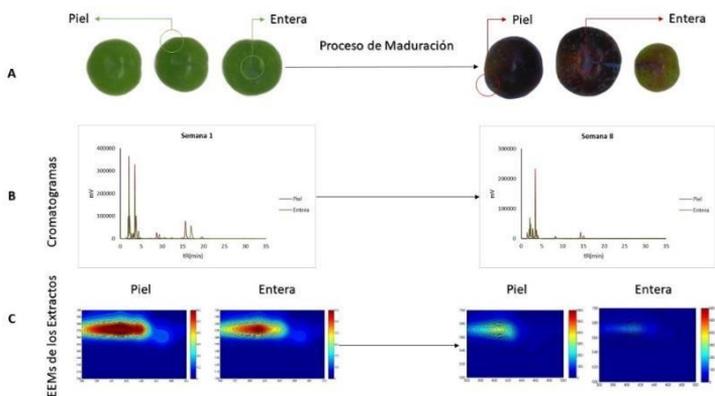


Figura 2.5. Comparación de los cromatogramas (B) de la piel y la ciruela entera, las EEMs (C) a lo largo del proceso de maduración de 9 semanas objeto de estudio (A).

Como se puede comprobar, tanto los resultados experimentales obtenidos del análisis cromatográfico y el registro de las EEMs de los extractos, observados en la **Figura 2.5**, como los antecedentes bibliográficos revisados [51, 52] muestran que la concentración promedio de fitoquímicos, incluidas las clorofilas, es significativamente mayor en la piel que en la pulpa de las ciruelas y, por tanto, justifica el desarrollo de metodologías no invasivas que puedan aplicarse sobre la propia superficie de la ciruela en relación a dichos indicadores, como es el ya mencionado uso de la fibra óptica.

3.3. Aplicación de métodos quimiométricos de segundo orden a las EEMs de las ciruelas registradas con fibra óptica para el seguimiento de su estado de maduración

3.3.1. Caracterización de las EEMs de los extractos de las ciruelas y de las ciruelas intactas

Una vez optimizados los parámetros instrumentales para la obtención de las EEMs haciendo uso de una fibra óptica, en la **Figura 2.6**, se muestran tanto las fotografías como las EEMs registradas en cuatro caras de cada ciruela de frutos recogidos al inicio y al final del proceso de maduración.

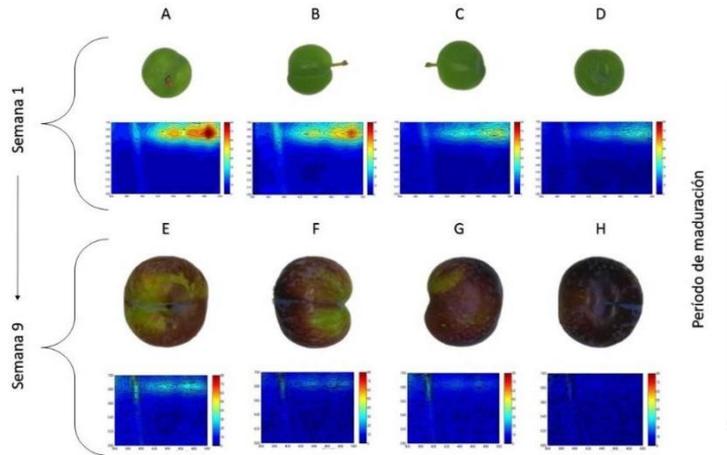


Figura 2.6. Fotografías de las caras frontal (A), laterales (B y C) y trasera (D) de una muestra de ciruela en la tercera primera de maduración y sus respectivas EEMs, y sus fotografías análogas en la novena semana de maduración (E, F G y H) registradas con la fibra óptica.

Como se puede observar, haciendo la comparativa de la apariencia visual que tienen las ciruelas a simple vista con respecto a su estado de maduración y sus respectivas EEMs, se aprecia que la fluorescencia intrínseca de las clorofilas de la piel de las ciruelas se relaciona con el color verde característico de la fruta en estados tempranos de maduración, lo que también se puede corroborar al hacer el análisis pormenorizado de las diferentes caras, en donde se puede ver que las zonas superficiales donde se aprecia una tonalidad más verde, presentan una mayor intensidad de fluorescencia en sus EEMs. Esto permite relacionar, en primer lugar, de manera cualitativa, y en segundo lugar, de manera cuantitativa, a través de la calibración multivariante de segundo orden, el estado de maduración, con la variación del contenido en clorofilas a través de las EEMs medias de todas las caras de las ciruelas, de modo que se evite la variabilidad intrínseca a la maduración irregular que se produce en las diferentes caras

de las ciruelas, como se ha podido comprobar en el análisis de la **Figura 2.6**.

De este modo, en la **Figura 2.7** se muestran solo las EEMs medias del inicio y del final de la maduración para comprobar la diferencia general entre los distintos estados de maduración. En el análisis pormenorizado de las EEMs se observa que en la matriz correspondiente al inicio del periodo de maduración, las muestras exhiben su mayor señal fluorescente en un rango de emisión entre 670 y 700 nm, con una señal de excitación máxima localizada entre los 400 y 500 nm. Esta señal fluorescente decae como consecuencia la disminución del contenido en clorofilas a medida que aumenta el estado de maduración [49, 50], haciéndose prácticamente inapreciable tras la novena semana de estudio, semana en la cual el cambio de color en la piel de la fruta de verde al rojo característico de las antocianinas [53]–[55] se hace evidente.

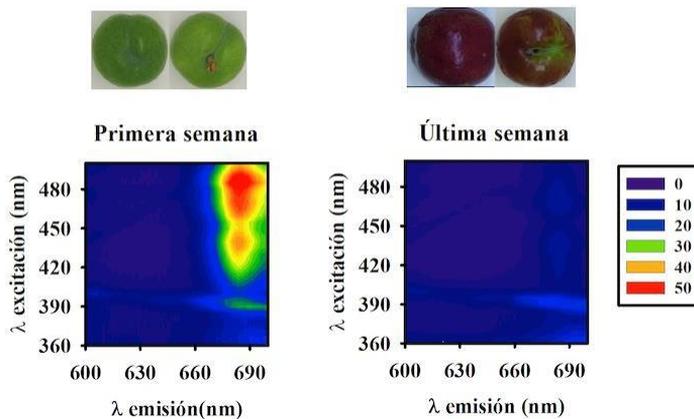


Figura 2.7. Representación de los perfiles fluorescentes de las EEMs obtenidas con el uso de la fibra óptica y fotografías correspondientes a las muestras de las EEMs correspondientes a la primera semana (Mayo de 2018) y la última semana (Agosto de 2018) del estudio.

El uso de las EEMs, como “*huellas dactilares*” del estado de maduración de las ciruelas, muestra una información apreciable a simple vista sobre la evolución de la concentración de las clorofilas en la piel de la ciruela, por lo que, el uso de algoritmos quimiométricos como los que se desarrollan a continuación, proporcionarán información más específica que permite caracterizar de manera efectiva el estado de maduración de las ciruelas.

3.3.2. Estudio cualitativo de la evolución del estado de maduración a partir de los perfiles de fluorescencia obtenidos con PARAFAC

Por cada una de las nueve semana correspondientes al proceso de maduración se registraron 12 EEMs correspondientes a tres ciruelas seleccionadas al azar (3 ciruelas por semana x 4 caras por cada ciruela). Se obtuvo, como consecuencia, un conjunto de datos tridimensionales de dimensiones 12 x 101 x 29 (muestras x 101 longitudes de onda de emisión x 29 longitudes de onda de excitación).

Los cubos tridimensionales fueron descompuestos utilizando inicialmente PARAFAC [56], aplicando el criterio diagnóstico del *core consistency* (CORCONDIA) [41], la desviación estándar de los residuales de la matriz de error de PARAFAC [57] y la evaluación visual del perfil de los *loadings* para la optimización del número de componentes [44].

Se construyó un modelo con las muestras independientes de cada semana, con un número creciente de factores de 1 a 5 para observar el cambio del *core consistency* y el error estándar de los residuales y seleccionar el número óptimo de componentes en base al análisis de dichos parámetros, y se aplicó, en todos los casos, los criterios de no negatividad para no atribuir en este caso concentraciones negativas a los valores espectrales obtenidos. Los resultados obtenidos a lo largo de las 9 semanas de maduración pueden observarse en la **Tabla 2.3**.

Tabla 2.3. Core, desviación estándar y varianza explicada obtenidas durante las 9 semanas objeto de estudio.

Componentes	Semana 1		Semana 2		Semana 3		Semana 4		Semana 5		Semana 6		Semana 7		Semana 8		Semana 9	
	core	SD	core	SD	core	SD	core	SD	core	SD	core	SD	core	SD	core	SD	core	SD
1	100	2.3	100	2.3	100	2.7	100	3.1	100	3.3	100	3.1	100	2.4	100	3.1	100	2.0
2	100	0.97	100	0.94	99	1.1	99	1.3	100	1.4	99	1.3	100	1.1	99	1.4	97	1.2
3	48	0.86	43	0.80	38	0.90	33	1.1	38	1.1	33	1.1	58	0.89	57	1.0	57	0.64
4	24	0.75	9	0.71	7.7	0.74	15	1.0	23	1.0	15	1.0	18	0.81	26	0.87	4.9	0.58
5	6	0.86	3	0.64	0.94	0.66	0.46	0.87	0.74	0.87	0.5	0.87	1.0	0.72	2.3	0.79	3.0	0.55
Varianza explicada (%) ^a	99.5		99.4		99.3		99.4		99.3		98.9		99.3		99.1		98.9	

^apara el número óptimo de componentes.

Como se puede comprobar, durante las primeras seis semanas, y usando el criterio del *core consistency*, se comprueba que al pasar al tercer componente el *core consistency* baja del 100% hasta aproximadamente el 40%, mientras que la desviación estándar de los residuales decrece de 3 al 1 aproximadamente cuando el número de componentes baja del primero al segundo, para posteriormente mantenerse estable hasta el quinto componente estudiado. Por este motivo, con ambos criterios, el valor óptimo del número de componentes fue de dos.

Por otro lado, a partir de la séptima semana, el valor del *core consistency*, decae del tercer al cuarto componente aproximadamente un 40%, en concreto del 60 al 20%, siendo la varianza explicada en este caso también similar para ambos números de componentes. Esta tendencia se mantiene a lo largo del resto de semanas de estudio, deduciéndose que en este caso se detecta inequívocamente un tercer componente y que coincide con la aparición de los colores rojos en el propio análisis visual de la fruta.

Una vez detectado el número de componentes correspondiente a los conjuntos representativos de cada semana de maduración, a continuación, en las **Figura 2.8A, 2.8B y 2.8C**, se pueden observar los perfiles de excitación y emisión obtenidos de los *loadings* de PARAFAC, obtenidos una vez aplicado el algoritmo, correspondientes a la segunda, cuarta y séptima semana de cosecha.

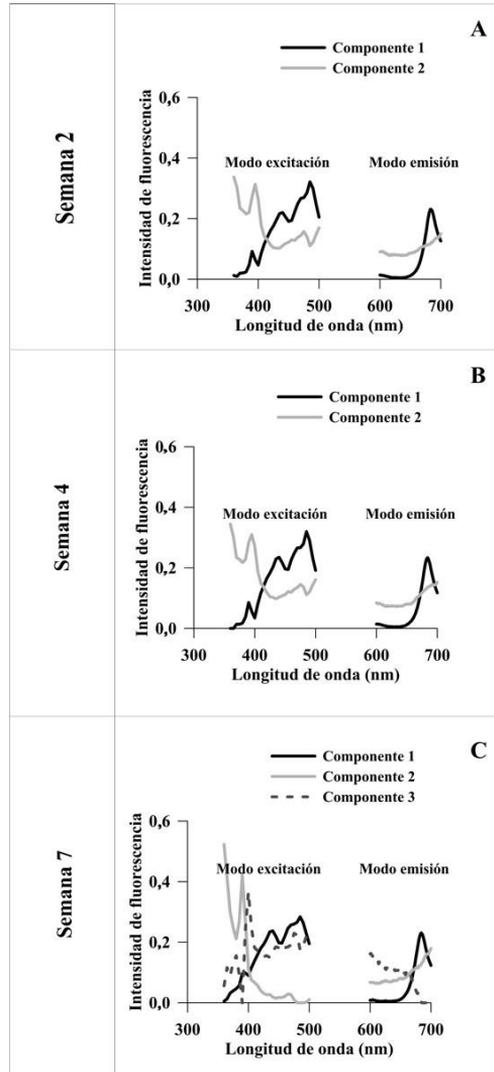


Figura 2.8. *Perfiles de excitación y emisión de PARAFAC de las muestras pertenecientes a los estados de maduración correspondientes a la segunda (A) , cuarta (B) y séptima semana (C).*

Como se puede observar, los *loadings* correspondientes al primer componente presentan perfiles muy similares durante todas las semanas de cosecha estudiadas, mostrando una forma definida con un máximo de emisión a 685 nm y dos máximos de excitación a 440 y 485 nm. Estos perfiles se pueden asociar con la presencia de compuestos del grupo de las

clorofilas presentes en las ciruelas en estados tempranos de maduración, ya que en la bibliografía consultada, los máximos de excitación y emisión están sobre los 458/653 nm cuando el espectro de fluorescencia se registra en medio acetona-agua [36]. La posición del máximo puede cambiar a diferentes longitudes de onda cuando se obtienen los espectros de muestras sólidas intactas en vez de disoluciones, debido a la variación que sufren las moléculas en estos medios, tal y como se ha estudiado previamente también [58].

Por otro lado, con respecto al segundo componente, los perfiles de emisión no muestran información relevante, sin embargo, el perfil de excitación está bien definido y presenta un máximo a 395 nm, el cual crece ligeramente con el proceso de maduración. Por último, el tercer componente que aparece a partir de la séptima semana, **Figura 2.8C**, presenta una forma definida con su máximo de excitación en 400 nm y dos máximos de emisión a 600 y 650 nm, sin embargo, la identificación de este componente resulta compleja.

Posteriormente, y teniendo en cuenta la variabilidad de las diferentes EEMs a lo largo del proceso de maduración, se construyeron dos modelos tridimensionales con objeto de estudiar la variación que se produce en los componentes atribuibles al estado de maduración en todas las muestras.

En primer lugar, se construyó un modelo con las matrices de cada semana (obteniendo la media de las 4 caras registradas en cada una de las tres réplicas correspondientes a las nueve semanas), siendo las dimensiones de dicho modelo 27 x 101 x 29 (siendo 27 el número de muestras [3 ciruelas x 9 semanas], 101 el número de longitudes de onda de emisión [desde 600 a 700 nm cada 1 nm] y siendo 29 el número de longitudes de onda de excitación [desde 360 a 500 nm cada 5 nm]). Luego, como en los casos anteriores, se utilizaron los diferentes criterios de selección para establecer

el número óptimo de componentes, siendo dos los suficientes para explicar el 99,4 % de la varianza. En la **Figura 2.9A** se muestra la evolución del valor de los scores de PARAFAC obtenidos para esta matriz a lo largo del periodo de maduración estudiado, pudiéndose observar que el valor de los scores para el primer componente crece durante las primeras semanas y empieza a decrecer a partir de la sexta, mientras que para los scores del segundo componente no se observa una tendencia al crecimiento o decrecimiento del mismo durante el periodo estudiado.

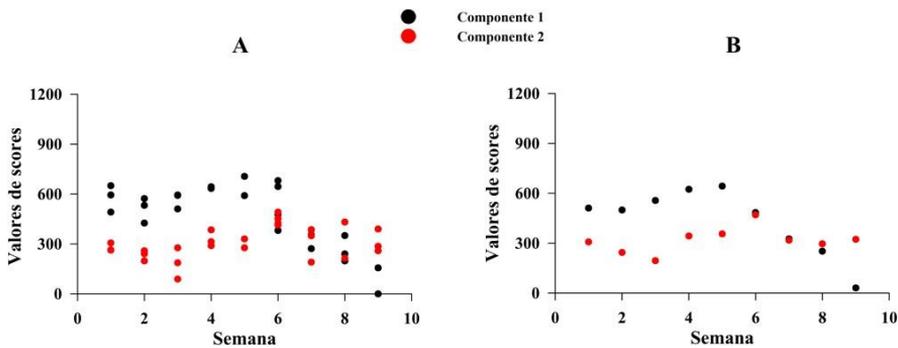


Figura 2.9. Evolución de los valores de los scores del primer y el segundo componente de PARAFAC, obtenidos de la descomposición de la matriz de correspondiente al valor medio de las 4 caras para las 3 réplicas de las 9 muestras representativas de cada semana de maduración (A), y considerando la media de las 3 réplicas correspondientes a cada muestra (B) en función del tiempo (semanas).

Por otro lado, en la **Figura 2.9B**, para poder normalizar la variabilidad dentro de las mismas réplicas, se hizo la media de las réplicas en base al valor medio de sus cuatro caras, de este modo se obtuvo un último modelo de dimensiones de $9 \times 101 \times 29$ (siendo el número de muestras 9 [valor medio de cada semana], 101 el número de longitudes de onda de emisión [desde 600 a 700 nm cada 1 nm] y siendo 29 el número de longitudes de onda de excitación [desde 360 a 500 nm cada 5 nm]). En esta

representación se observa mejor la tendencia de la evolución de los dos componentes obtenidos de la aplicación de PARAFAC, observándose una tendencia parecida para el primer componente, donde este crece hasta la quinta semana y partir de la sexta decrece continuamente hasta llegar a valores cercanos a cero.

3.3.3. Estudio cuantitativo de la evolución del estado de maduración a partir de la cuantificación de clorofilas.

Una vez realizada la identificación de los componentes de PARAFAC en función de la concordancia de los perfiles de fluorescencia obtenidos de los *loadings* con los espectros de las clorofilas consultados en la bibliografía, el siguiente paso fue confirmar dicha hipótesis a través de la correlación del valor de los *scores* del primer componente de PARAFAC con la concentración de clorofilas obtenida a partir de la aplicación del método HPLC-FLD de referencia, tanto sobre la piel, como sobre la totalidad de las ciruelas analizadas. Las correlaciones obtenidas entre los *scores* del primer componente y la concentración de clorofilas totales obtenidas de los extractos ofrecieron buenos resultados, obteniendo unos coeficientes de correlación para el caso de la concentración en la piel de 0,899 y para el caso de la concentración en la totalidad de la fruta fue de 0,863 (**Tabla 2.4**). La correlación de PARAFAC se puede observar en la **Figura 2.11A** y **Figura 2.11C**.

Para la aplicación de MCR-ALS se obtuvo la matriz aumentada en la dirección de las longitudes de onda de excitación. El número de componentes en cada matriz aumentada se estimó mediante PCA y se justificó teniendo en cuenta los analitos correspondientes, las posibles interferencias y la señal de fondo. Se aplicaron las restricciones de no negatividad tanto en los modos de emisión como de excitación y se aplicó la restricción de unimodalidad a los analitos correspondientes y no a la señal

de fondo. Tras la optimización mediante ALS para cada muestra, con ayuda de las correspondientes curvas de calibración pseudounivariantes, los constituyentes fueron identificados. El número óptimo de componentes fue de 3, los cuales explicaban un total del 99,7% de la varianza. En la **Figura 2.10**, se pueden ver los perfiles de excitación y emisión a los componentes del modelo MCR-ALS al que se ha hecho referencia.

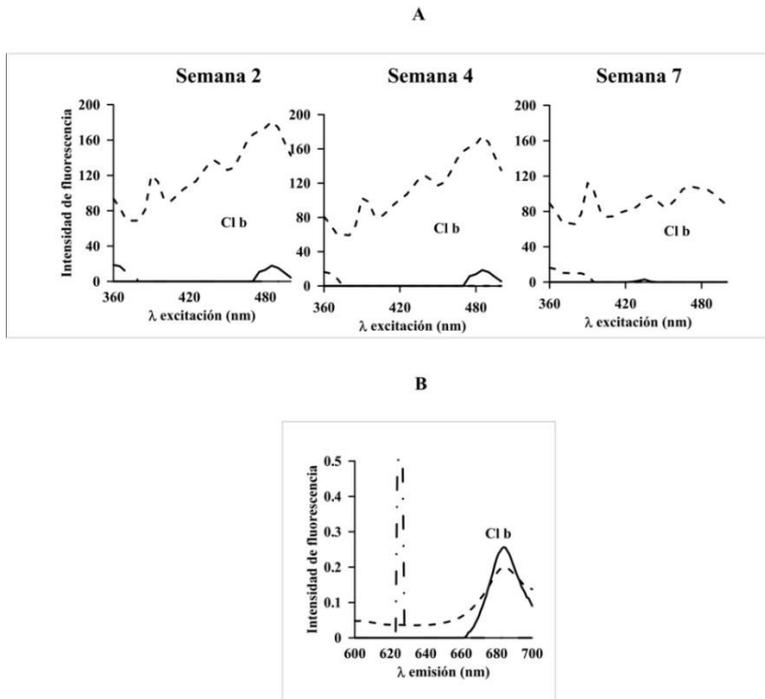


Figura 2.10. (A) Perfiles de excitación obtenidos del modelo MCR-ALS para las distintas semanas 2, 4 y 7. (B) Espectro de emisión obtenido del modelo MCR-ALS. Las líneas punteadas corresponden a los perfiles de excitación y emisión obtenidos mediante MCR-ALS para los compuestos desconocidos y la señal de fondo. Las líneas continuas se corresponden con los espectros de excitación y emisión obtenidos de MCR-ALS para la clorofila.

Como se puede observar, al comparar los perfiles del primer componente de MCR-ALS con los obtenidos de PARAFAC, se observa que estos son muy similares tanto en excitación como en emisión. Para comparar la similitud de dichos perfiles, se llevó a cabo un análisis estadístico de correlación entre ambos (análisis de regresión de Pearson) que ofreció unos coeficientes de regresión de 0,989 y 0,964 para la emisión y la excitación, respectivamente, y con un nivel de confianza del 95%. Los coeficientes de correlación entre los valores de los *scores* del primer componente frente a las concentraciones obtenidas mediante HPLC en la piel de las ciruelas fue de 0,9581, mientras que en el caso de la totalidad de la fruta fue de 0,8851.

Seguidamente, se aplicaron los algoritmos U-PLS y N-PLS, utilizando el criterio de Haaland y Thomas [59, 60], para determinar el número óptimo de variables latentes para el modelo, las cuales son aquellas que dado un número de componentes determinado, no arroja un valor del PRESS estadísticamente inferior al mínimo valor obtenido para un número determinado de componentes (ratio F de probabilidad debajo de 0,75), siendo en este caso también un número óptimo de variables total de 3. En este caso, la predicción de la concentración de clorofilas mediante ambos algoritmos arrojaron coeficientes de correlación entre las concentraciones predichas y las reales obtenidas mediante el método HPLC de 0,940 y 0,936 para U-PLS y N-PLS, respectivamente, cuando se realizó la comparación con las concentraciones obtenidas en la piel de las ciruelas, y de 0,877 y 0,909, respectivamente, en el caso de la validación realizada con las concentraciones de referencia obtenidas del análisis de la totalidad de la fruta. En la **Figura 2.11** se pueden observar los resultados obtenidos para los modelos de validación establecidos para los algoritmos ensayados entre las concentraciones reales y predichas mediante validación cruzada, así como la aplicación del test de la región elíptica de confianza EJCR para un

nivel de confianza del 95 % [61], para la evaluación de la pendiente y la ordenada en el origen correspondientes a la regresión lineal entre las concentraciones predichas por los algoritmos y la concentración real obtenida del método HPLC de referencia.

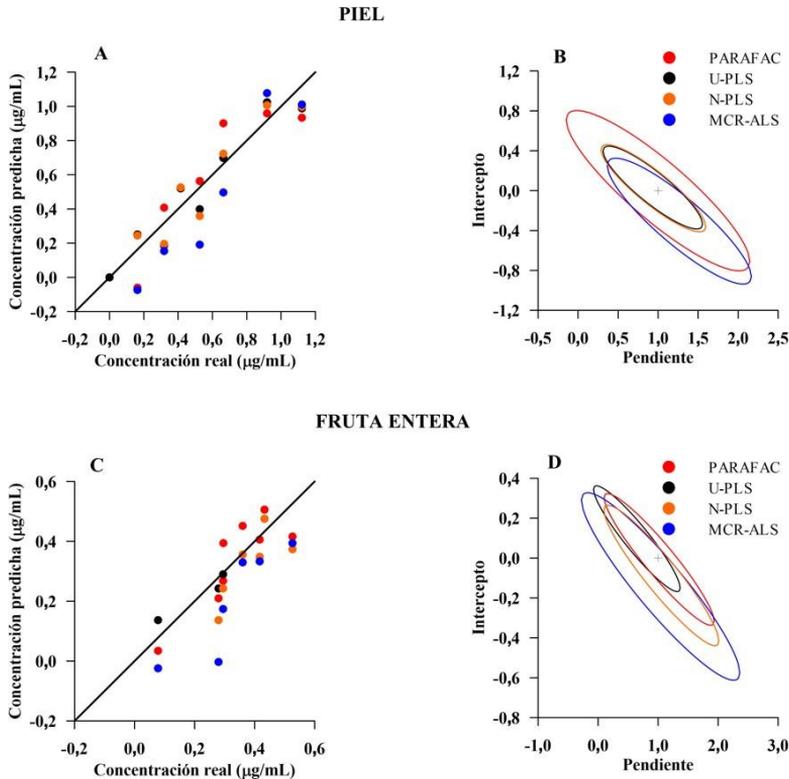


Figura 2.11. Representación de la predicción del contenido en clorofilas para los diferentes algoritmos en función de la concentración real obtenida mediante HPLC (A y C) y la correspondiente EJCR para un 95% de nivel de confianza para las pendientes y ordenadas en el origen de las regresiones (B y D). El punto teórico (ordenada en el origen=0; y pendiente=1) está marcada en la figura con la línea negra cruzada.

Se observa que el ajuste del modelo a la concentración teórica es muy preciso a simple vista, ya que, en todos los casos, los puntos se acercan a la línea del punto teórico, especialmente en el caso de la piel, lo cual queda

remarcado en las **Figuras 2.11 B y 2.11 D**, donde al comprobar el test EJC se observan claramente que las elipses para los modelos U-PLS y N-PLS aplicados sobre la piel de las ciruelas son más pequeñas, lo que sugiere una mayor precisión de estos algoritmos para estos casos.

Por otro lado, las figuras de mérito obtenidas de la aplicación de los algoritmos a la predicción de la concentración total de clorofilas en las muestras pueden observarse en la **Tabla 2.4**, donde se pueden comparar los valores del número de componentes, pendiente, ordenada en el origen, coeficiente determinación (R^2), desviación estándar, error cuadrático medio de predicción RMSEP ($\text{mg}\cdot\text{g}^{-1}$) y error relativo de predicción REP (%) para la validación realizada tanto en la piel como en la fruta entera, para los algoritmos ensayados PARAFAC, MCR-ALS, U-PLS y N-PLS.

Tabla 2.4. Figuras de mérito para los diferentes algoritmos empleados.

Piel							
Algoritmo	Componentes	Pendiente	Ordenada	R^2	Sr	RMSEP ($\text{mg}\cdot\text{g}^{-1}$)	REP (%)
PARAFAC	3	1,0	-0,001	0,809	0,2	0,2	26
U-PLS	3	0,93	0,03	0,884	0,1	0,1	20
N-PLS	3	0,94	0,03	0,877	0,1	0,1	21
MCR-ALS	3	1,3	-0,31	0,918	0,8	0,2	44
Pulpa							
PARAFAC	3	1,0	-1e-7	0,744	0,1	0,1	37
U-PLS	3	0,67	0,09	0,769	0,2	0,1	24
N-PLS	3	1,0	-0,09	0,826	0,2	0,1	38
MCR-ALS	3	1,1	-0,15	0,784	0,4	0,1	74

En el análisis pormenorizado que podemos hacer de las figuras de mérito reflejadas en la **Tabla 2.4**, se puede comprobar que los mejores resultados se obtienen a través de los algoritmos U-PLS y N-PLS, arrojando estos los valores más bajos de sus respectivos errores cuadráticos medios de predicción (RMSEP) y errores relativos de predicción (REP), que junto a su mayor precisión en el test EJCR reflejan que para la aplicación requerida de correlacionar los valores obtenidos de las EEMs en muestras intactas con la cantidad de clorofilas son los más adecuados.

Como conclusión general se puede decir que los resultados son similares en los diferentes algoritmos, siendo más potentes los algoritmos U-PLS y N-PLS, destacando así el potencial que los algoritmos de segundo orden presentan para aplicaciones como el seguimiento del estado de maduración y la adecuación del uso de la fluorescencia molecular registrada mediante fibra óptica para el desarrollo de una metodología rápida, sencilla, selectiva y no destructiva.

3.4. Aplicación de métodos quimiométricos a los datos de espacios de color obtenidos del procesamiento de imágenes digitales de ciruelas para el seguimiento de su estado de maduración

Una vez mostrada la potencialidad del uso de datos de fluorescencia molecular en la caracterización del estado de maduración, en este apartado se detallará el estudio realizado a partir del procesamiento y análisis de imágenes digitales.

3.4.1. Análisis univariante

En primer lugar, con objeto de simplificar el procedimiento de análisis de las imágenes digitales se realizó inicialmente el estudio los parámetros de color del espacio RGB usando inicialmente el valor promedio de los canales R, G y B obtenidos del código “imread” de Matlab.

Se empleó dicho valor promedio para las imágenes obtenidas de las tres caras representativas en cada una de las tres réplicas correspondientes a las distintas semanas de maduración y posteriormente se calculó la ratio de cada canal a través de las ecuaciones (1), (2) y (3) especificadas en el **apartado 2.7.1.**

Para el estudio cualitativo del grado de maduración respecto a los parámetros univariantes establecidos a través del valor de la ratio de los canales R, G y B se representó la evolución de dichas ratios con las diferentes semanas de maduración, obteniendo los resultados que pueden observarse en la **Figura 2.12.**

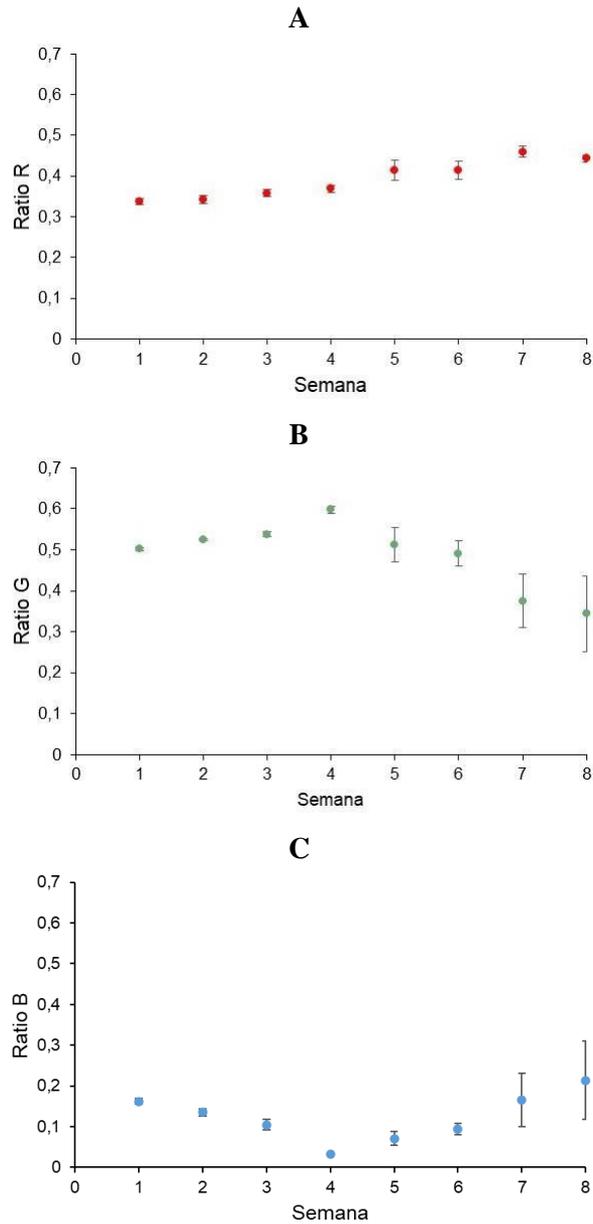


Figura 2.12. Evolución de las ratios medias de los canales R, G y B junto a sus barras de error durante las 8 semanas el periodo de maduración objeto de estudio (la ratio media fue obtenida considerando las tres réplicas de las muestras y las barras de error, siendo la desviación media de las tres réplicas).

En esta primera aproximación se puede destacar que a partir de la cuarta semana de maduración la ratio del canal verde (**Figura 2.12B**) alcanza un máximo mientras que la ratio del canal azul (**Figura 2.12C**) alcanza un mínimo. También es destacable mencionar que este cambio en la distribución del valor medio de las ratios de color no se aprecia a simple vista si comparamos con la propia imagen sin tratar, por lo que se puede comprobar que el procesamiento de las imágenes arroja en una primera aproximación cualitativa, datos que muestran determinadas tendencias en el propio color. Estos resultados se atribuyen a los cambios en el proceso de maduración, ya que, el proceso madurativo del fruto se relaciona con los cambios del contenido de fitoquímicos como son las clorofilas, los carotenoides y las antocianinas, cuyas propiedades dan la pigmentación característica de la fruta en sus diferentes etapas de maduración [62, 63]. Además, como se puede ver en las barras de error, también se puede comprobar que la variabilidad en la ratio de color se hace mayor conforme aumenta el estado de maduración, mientras que por otro lado, como se puede observar en la **Figura 2.12A**, la *ratio* del canal rojo muestra un ligero incremento a lo largo del periodo de maduración estudiado. Esto podría estar asociado al decrecimiento en el contenido de clorofilas, ya que, como se ha podido mencionar en otros estudios [25], el rojo es el color complementario al verde, indicando una posible correlación de este con el decrecimiento en el contenido en clorofilas totales, lo cual es objeto de la segunda fase del estudio, que sería asociar el estado de maduración a un indicador cuantitativo como es el contenido en clorofilas totales.

Posteriormente, se correlacionaron los valores de las ratios de los canales rojo, verde y azul con el contenido total en clorofilas a través de regresión lineal univariante cuyas rectas se representan en la **Figura 2.13**. Y cuyas figuras de mérito se recogen en la **Tabla 2.5**.

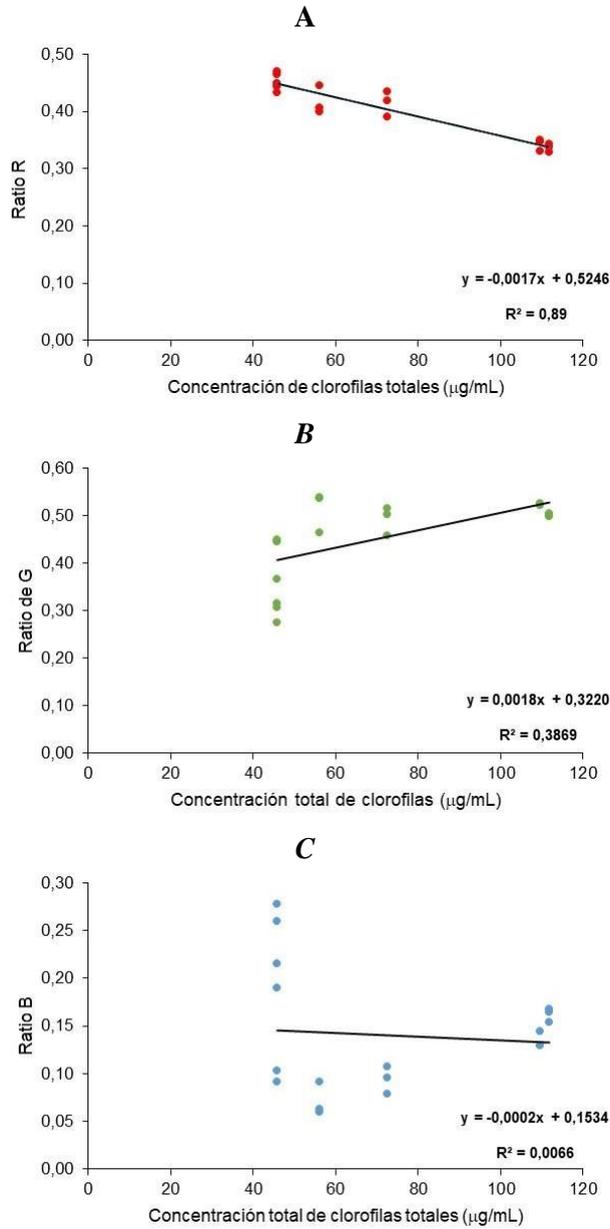


Figura 2.13. Curvas de regresión lineal para los valores de ratio de los canales R, G y B y el contenido total en clorofilas (los valores medios se tomaron considerando las tres imágenes de cada muestra correspondiente a cada semana de manera individual).

Como se puede observar, el uso del método de regresión lineal no es adecuado para los canales G y B (**Figura 2.13 B** y **Figura 2.13C**), no obstante, en el canal rojo (**Figura 2.13 A**) se obtienen unos resultados muy aceptables, con unos valores de $R^2=0,89$, $RMSEP= 9,42\mu\text{g/mL}$ y un $REP=13\%$, como se puede observar en la **Tabla 2.5**.

Tabla 2.5. Figuras de mérito para la regresión univariante de los diferentes canales.

Canal	Pendiente	Ordenada	R^2	RMSEP ($\mu\text{g/mL}$)	REP (%)
R	0,98	0,57	0,89	9,4	13
G	0,24	101,4	0,39	53,8	73
B	-4,98	635,7	0,01	304,9	414

R: rojo; G: verde; B: azul; R^2 : coeficiente de determinación; RMSEP: Raíz del error cuadrático medio de predicción; REP: error relativo de predicción.

Estos resultados concuerdan con otros estudios realizados en este ámbito como en el caso de la estimación de clorofilas totales en plantas de patatas [64], hojas de remolacha azucarera [65], y hojas de betel [66], usando la *ratio* de color de los canales R, G y B. Además, los resultados obtenidos sobre las muestras ensayadas también están conformes con los citados en la sección anterior utilizando la autofluorescencia propia de las clorofilas presentes en la piel de la ciruela mediante el uso de la fibra óptica [67], donde se pudo observar que los factores obtenidos mediante el uso de PARAFAC sobre las EEMs de las ciruelas, en relación a las clorofilas a y b, muestran una tendencia similar a lo largo del proceso de maduración, lo que refuerza el argumento de que el incremento del valor de la *ratio* del color rojo se correlaciona con la disminución en la concentración de clorofilas totales con unas figuras de mérito en dicha correlación bastante reseñables. Por tanto, este método resulta una propuesta analítica simple y fiable de aplicación de los datos obtenidos del procesamiento de imágenes digitales sobre las ciruelas.

3.2.2. Análisis multivariante de primer orden

Una vez contrastados los resultados obtenidos mediante análisis de datos univariantes, el siguiente paso fue comprobar si dichos resultados pueden ser mejorados a través de la aplicación de métodos de análisis exploratorio y calibración multivariante de orden superior a partir del procesamiento y obtención de datos en este caso de primer orden (histogramas en el espacio de color RGB).

Para este primer estudio, en primer lugar, se obtuvo el histograma medio de las tres caras analizadas para las tres réplicas representativas de las muestras pertenecientes a cada una de las semanas de maduración estudiadas para los canales R, G, B y los tres canales en conjunto RGB. De este modo, tal y como se ha especificado en el **apartado 2.7.2**, se obtuvo una matriz de 24 x 256 (siendo 24 el número de muestras [8 muestras x 3 réplicas] y 256 las distintas intensidades de cada color para cada uno de los 400 x 400 píxeles de la ROI]. Con todo ello, se procedió al análisis exploratorio de los datos referentes a los histogramas de las 24 muestras recopiladas para cada uno de los canales estudiados, obteniendo los gráficos que pueden visualizarse en la **Figura 2.14 A-D**.

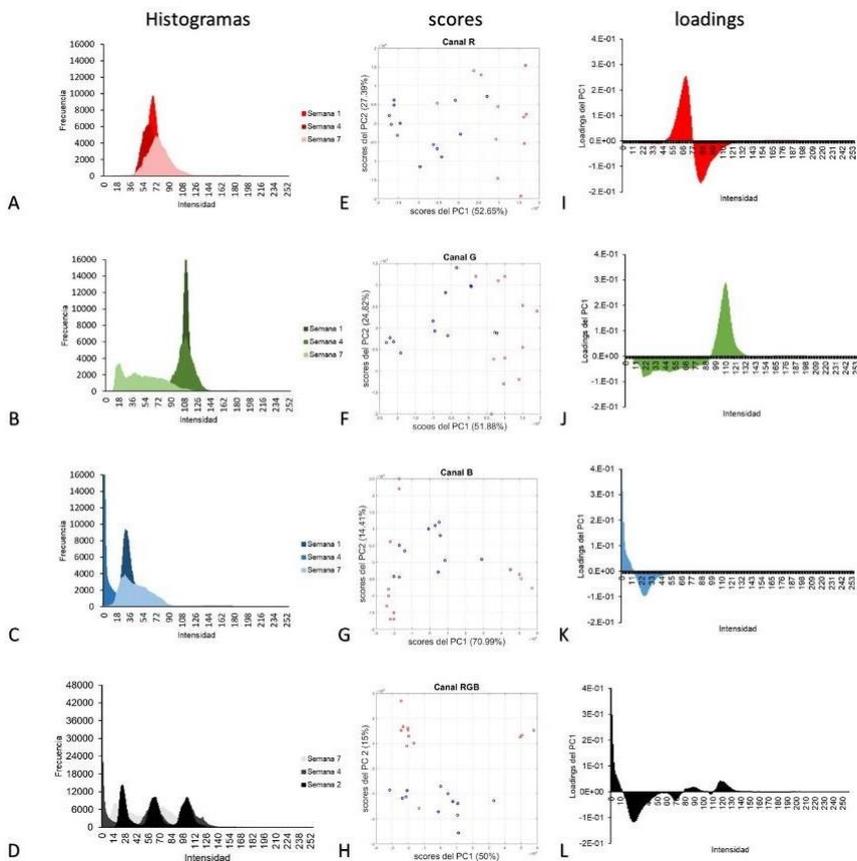


Figura 2.14. Histogramas (A-D), valores de los scores de los dos primeros componentes de PCA, siendo los puntos de color rojo las muestras anteriores a la cuarta semana de cosecha, y los puntos de color azul las muestras posteriores a la cuarta semana de cosecha (E-H) y perfiles de los loadings del primer componente de PCA (I-L) para los canales R, G, B y RGB durante el proceso de maduración.

Analizando en primer lugar el canal verde (**Figura 2.14B**), se puede observar que la distribución de la frecuencia de sus histogramas, con respecto al tiempo, evoluciona hacia intensidades más bajas. Inicialmente, empieza en una intensidad máxima de 112, para posteriormente, alcanzar el máximo en la séptima semana en intensidades alrededor de 20,

observándose además una mayor distribución de las distintas intensidades en concordancia a la variabilidad observada en el análisis univariante. Esta evolución se hace evidente debido al proceso de maduración de la ciruela en el que a simple vista se observa cómo el color verde deja paso a tonalidades más oscuras y colores como los colores rojo y violeta que aparecen con la biosíntesis de antocianinas y la reducción del contenido en clorofilas, provocando así una mayor absorción de longitudes de onda en la zona del verde y reflejando las longitudes de onda en el rojo, tendencia que se puede complementar con las observaciones realizadas para la evolución de los histogramas en el canal rojo (**Figura 2.14A**). En esta figura se observa una evolución contraria, en el desplazamiento del pico de máxima frecuencia de intensidades más bajas (sobre 64) a intensidades más altas de rojo (sobre 80 para la séptima semana), pero observándose la misma variabilidad de tonalidades con el paso del tiempo. Por otro lado, en el caso del canal azul, este desplazamiento de las frecuencias para intensidades características no se hace tan evidente, siendo el comportamiento menos predecible.

Una vez analizada la tendencia de distribución de frecuencias de las diferentes intensidades de los canales R, G, B y RGB en los histogramas correspondientes a cada semana, se aplicó el análisis PCA con objeto de identificar y caracterizar los componentes que más contribuyen a la diferenciación del estado de maduración para los canales analizados, pudiendo así distinguir entre diferentes estados de maduración de acuerdo al valor que adquieren los *scores* (**Figuras 2.14E-H**), y caracterizando los componentes principales a partir del perfil de sus *loadings* (**Figura 2.14 I-L**). De este modo, para determinar el estado de maduración, en la representación gráfica del valor de los *scores* para los componentes obtenidos, se colorearon estos según una fecha límite de cosecha, la cual fue la cuarta semana, donde se detectaron cambios sustanciales en las

tendencias de las ratios de color en el apartado anterior. De este modo, las muestras pertenecientes a fechas anteriores a la cuarta semana se colorearon de rojo y las posteriores de azul, tal y como se puede comprobar en la **Figura 2.14 E-H**. Los resultados obtenidos muestran que, sobre todo, para el canal G (**Figura 2.14 F**) se observa una tendencia a la separación en la distribución de las muestras según el valor de los *scores* de sus dos primeros componentes principales, que en el caso del canal G explicaban un 76% del total de la varianza. La tendencia a la separación se hace más evidente en el *cluster* efectuado en los tres canales conjuntos RGB (**Figura 2.14 H**). Por otro lado, los *loadings* para el primer componente mostrados en la **Figura 2.14 I-L**, muestran las variables que se asocian a dichos componentes, pudiéndose observar que para el caso de los *loadings* del canal RGB (**Figura 2.14 L**), las intensidades donde se da la máxima frecuencia son coincidentes con las de su histograma, en torno a 26-30, 66-68 y 110-112. Estos resultados concuerdan con las estimaciones realizadas conforme al comportamiento del color observado durante las cuatro primeras semanas de cosecha, cuya morfología en los histogramas y cuya evolución según la *ratio* de R, G, y B visto en el apartado anterior cambian drásticamente a partir de dicha semana.

Finalmente, y una vez confirmada la tendencia al cambio del estado de maduración conforme al PCA realizado sobre las características del espacio de color RGB de los histogramas, el siguiente paso fue, al igual que en el análisis de datos univariantes, establecer un criterio cuantitativo de seguimiento del estado de maduración por correlación mediante análisis multivariante de los histogramas obtenidos con respecto a las clorofilas totales determinadas mediante análisis HPLC-FLD. De este modo, se realizaron modelos de calibración multivariante aplicando PLS para los distintos canales estudiados, utilizando la validación cruzada para establecer dicho modelo y determinar el número de componentes

seleccionados de acuerdo con el criterio de la varianza explicada. La **Figura 2.15 A-D** muestra los coeficientes de regresión de PLS para los distintos canales ensayados, observándose bastante concordancia con los *loadings* de PCA principalmente en los canales R y G, y la **Figura 2.15 E-H** muestra la concentración de clorofilas totales predicha frente a la concentración real obtenida a partir del análisis de referencia HPLC, observándose una buena concordancia entre dichos valores con el mejor coeficiente de determinación (0,92) obtenido para el canal verde. Las figuras de mérito para la aplicación de esta regresión PLS sobre los histogramas se pueden observar en la **Tabla 2.6**, en la cual se muestra, además del coeficiente de determinación, otras figuras de mérito que para los canales verde y RGB superan las obtenidas por la regresión univariante que se muestra en la **Tabla 2.5**.

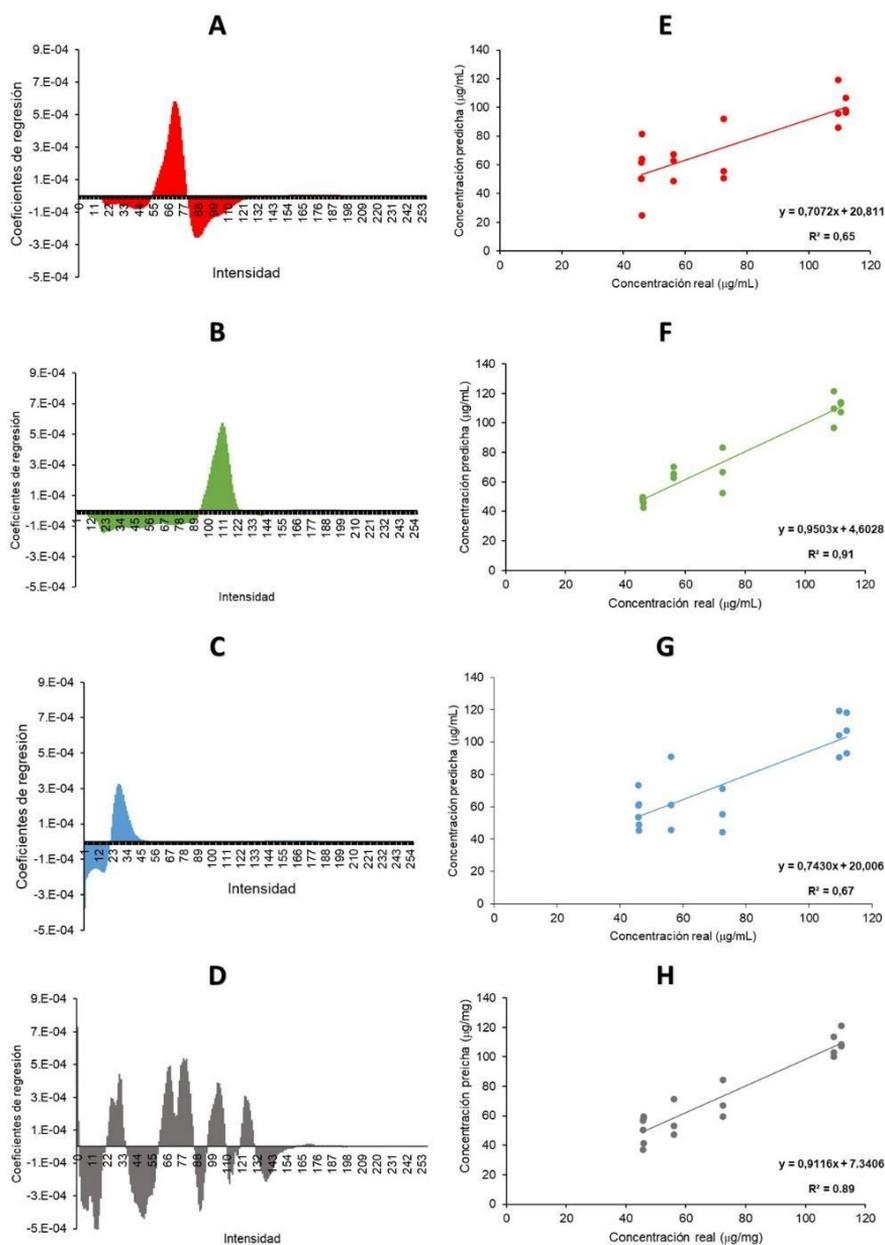


Figura 2.15. Representación de los coeficientes de regresión de PLS (A-D) y concentración predicha frente a concentración real de clorofilas totales (E-H) para los canales R, G, B y RGB.

Tabla 2.6. Figuras de mérito para los modelos de regresión PLS obtenidos sobre los histogramas de primer orden para los diferentes canales estudiados.

Canal	Componentes	Pendiente	Ordenada	R ²	RMSEP (µg/mL)	REP (%)
R	3	0,70	29,8	0,68	16,6	23
G	5	0,97	4,60	0,92	8,4	11
B	3	0,74	20,0	0,70	16,0	21
RGB	7	0,71	7,34	0,89	9,1	12

R: rojo; G: verde; B: azul; RGB: rojo; verde; azul; R²: coeficiente de determinación; RMSEP: Raíz del error cuadrático medio de predicción; REP: error relativo de predicción.

3.4.3. Análisis multivariante de segundo orden

Por último, una vez concluido el tratamiento de histogramas de color de primer orden, el siguiente objetivo fue comparar la adecuación de dichos resultados con los obtenidos a través del tratamiento de datos de segundo orden, de acuerdo a la metodología especificada en el **apartado 2.7.3**. De este modo, cada matriz obtenida de los píxeles individuales de cada ciruela arrojó cubos de datos de 24 x 400 x 400, (siendo 24 el número de muestras [8 muestras x 3 réplicas] y 400 el número de píxeles por cada dimensión de la superficie de la ROI fotografiada). Posteriormente, se calculó la matriz media correspondiente a cada una de las tres réplicas de las muestras representativas de cada semana de cosecha del mismo modo que el anterior, haciendo la media de cada uno de los 400 x 400 píxeles correspondientes a cada una de las tres caras estudiadas. Una vez obtenida la matriz media, inicialmente se realizó el mismo análisis exploratorio que en el caso de los datos de primer orden aplicando en este caso PARAFAC para obtener los correspondientes *scores* después de la determinación del número óptimo de componentes, el cual se determinó usando del criterio del *core consistency*

(CORCONDIA) [41]. A través de este análisis exploratorio y representando los scores de los dos principales componentes de PARAFAC para cada muestra no se pudieron obtener resultados tan satisfactorios a la hora de detectar diferencias entre grupos como en el caso de los histogramas. Este hecho posiblemente se pueda deber a la heterogeneidad en el tamaño y la superficie de muestras estudiadas, las cuales no son modeladas dentro de los parámetros de color en una matriz de datos tan grande donde se recogen los datos de cada uno de los pixeles.

Por otro lado, para llevar a cabo el análisis cuantitativo y obtener los correspondientes modelos de calibración, se aplicaron los algoritmos U-PLS y N-PLS a los datos de segundo orden procedentes de los cubos de matrices medias de las tres caras de las ciruelas analizadas para correlacionarlos con las concentraciones de clorofilas totales obtenidas mediante el método de referencia HPLC, comprobando la adecuación de estos modelos a la determinación cuantitativa de dichas clorofilas. De este modo, aplicando estos algoritmos y representando dicha predicción frente a la concentración obtenida a través del método HPLC de referencia se obtuvieron los resultados que pueden observarse en la **Figura 2.16**.

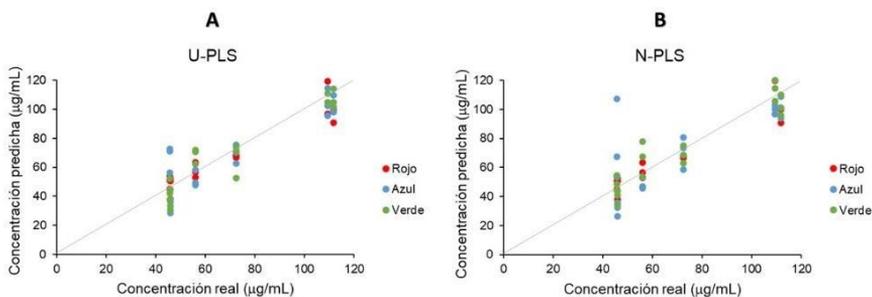


Figura 2.16. Concentraciones reales frente a concentraciones predichas por los modelos U-PLS (A) y N-PLS (B) para los canales rojo, verde y azul con sus respectivas líneas de regresión.

Como se puede observar en las líneas de regresión, los resultados para los canales R y G son adecuados con respecto a las concentraciones reales obtenidas mediante HPLC, mientras que el canal azul no arroja resultados acordes a los valores de referencia. En la **Tabla 2.7** se puede contrastar cuantitativamente, a través de las figuras de mérito respectivas a los canales estudiados para los algoritmos aplicados, la adecuación de los métodos de regresión propuestos, viendo que los resultados para el canal verde son todavía más precisos que los obtenidos con el modelo construido con datos de primer orden, con un R^2 de 0,96 y 0,95, un RMSECV de 7,50 y 9,64 $\mu\text{g/mL}$ y un REP del 10 y el 12 % para U-PLS y N-PLS, respectivamente.

Tabla 2.7. Figuras de mérito para los modelos de regresión PLS obtenidos sobre los histogramas de primer orden para los diferentes canales estudiados.

Algoritmo	Canal	Componentes	Pendiente	Ordenada	R^2	RMSEP ($\mu\text{g/mL}$)	REP (%)
U-PLS	R	3	0,93	5,06	0,95	9,1	12
	G	3	0,95	3,52	0,96	7,5	10
	B	3	0,90	6,44	0,95	12,2	17
N-PLS	R	4	0,93	6,35	0,92	10,9	15
	G	3	0,93	4,67	0,95	9,64	12
	B	2	0,77	14,7	0,78	18,9	25

R: rojo; G: verde; B: azul; R^2 : coeficiente de determinación; RMSEP: Raíz del error cuadrático medio de predicción; REP: error relativo de predicción. U-PLS: unfolded-PLS; N-PLS: multi-way-PLS

Como se puede observar, esta metodología aplica los algoritmos quimiométricos a cada pixel de manera individual, lo cual añade más información y se traduce en datos más exactos en ámbitos cuantitativos, pero menos fiables en el caso del análisis exploratorio.

Como conclusión podemos extraer que el procesamiento de imágenes digitales a través del espacio de color RGB en combinación con métodos quimiométricos han demostrado ser una metodología simple, fiable y precisa para el seguimiento del estado de maduración durante su periodo de cosecha, con potenciales beneficios en sus características comerciales.

A través del desarrollo de estos métodos se obtuvieron los mejores resultados para el canal R, en el caso de la regresión univariante, el cual es el complementario al color observado correlacionado con las clorofilas, estos resultados concuerdan con otros estudios previamente señalados.

En el caso del análisis de los histogramas, realizados mediante PCA, se pudo observar que los cambios en los tres colores aditivos R, G, B, representados en los histogramas, presentaban variaciones a lo largo del tiempo que se podían visualizar fácilmente y tratar a través de herramientas gráficas para la representación de los componentes de PCA. Además, mediante el análisis por PLS se obtuvo una buena correlación con el contenido en clorofilas totales, obteniendo excelentes resultados reflejados en sus figuras de mérito obtenidas, principalmente para el canal G.

Por otro lado, empleando los datos de segundo orden (con cada pixel arrojando un valor de intensidad de R, G o B), se han obtenido resultados más acertados en las cuantificaciones a través de los algoritmos U-PLS y N-PLS.

Con estos resultados, se puede concluir que las metodologías mostradas presentan grandes ventajas respecto a metodologías tradicionales, así como un gran potencial para continuar investigado en este campo de estudio, utilizando un volumen mayor de muestras para dar a los modelos planteados una mayor robustez.

4. REFERENCIAS

- [1] M. Guerra, M. A. Sanz, and P. A. Casquero, “Influence of harvest dates on quality, storage capacity and sensory attributes of European plum cv. green gage,” *Food Science and Technology International*, vol. 15, no. 6, pp. 527–534, 2009.
- [2] M. Knee, “Pome fruits,” in *Biochemistry of Fruit Ripening*, G. B. Seymour, J. E. Taylor, and G. A. Tucker, Eds., Dordrecht: Springer Netherlands, 1993, pp. 325–346.
- [3] M. Bajcar *et al.*, “METHOD FOR DETERMINING FRUIT HARVESTING MATURITY,” *Journal of microbiology, biotechnology and food sciences*, vol. 6, no. 2, pp. 773–776, 2016.
- [4] C. H. Crisosto, “Stone fruit maturity indices: a descriptive,” *Postharvest News and Information*, vol. 5, no. 6, pp. 65N-68N, 1994.
- [5] B. Schoefs, “Chlorophyll and carotenoid analysis in food products. Properties of the pigments and methods of analysis,” *Trends in food science & technology*, vol. 13, no. 11, pp. 361–371, 2002.
- [6] R. Delgado-Pelayo, L. Gallardo-Guerrero, and D. Hornero-Méndez, “Chlorophyll and carotenoid pigments in the peel and flesh of commercial apple fruit varieties,” *Food Research International*, vol. 65, pp. 272–281, 2014.
- [7] K. Itakura, Y. Saito, T. Suzuki, N. Kondo, and F. Hosoi, “Estimation of citrus maturity with fluorescence spectroscopy using deep learning,” *Horticulturae*, vol. 5, no. 1, pp.2, 2019.

- [8] Z. G. Cerovic, N. Moise, G. Agati, G. Latouche, N. Ben Ghazlen, and S. Meyer, “New portable optical sensors for the assessment of winegrape phenolic maturity based on berry fluorescence,” *Journal of Food Composition and Analysis*, vol. 21, pp. 650–654, 2008.
- [9] M. Cabrera-Bañegil *et al.*, “Evolution of polyphenols content in plum fruits (*Prunus salicina*) with harvesting time by second-order excitation-emission fluorescence multivariate calibration,” *Microchemical Journal*, vol. 158, p. 105299, 2020.
- [10] O. Monago-Maraña, M. Cabrera-Bañegil, N. L. Rodas, A. Muñoz de la Peña, and I. Durán-Merás, “First-order discrimination of methanolic extracts from plums according to harvesting date using fluorescence spectra. Quantification of polyphenols,” *Microchemical Journal*, vol. 169, p.106533, 2021.
- [11] M. Lechaudel, L. Urban, and J. Joas, “Chlorophyll Fluorescence, a Nondestructive Method To Assess Maturity of Mango Fruits (Cv. ‘Cogshall’) without Growth Conditions Bias,” *Journal of agricultural food chemistry*, vol. 58, no. 13, pp. 7532–7538, 2010.
- [12] A. Salama, M. Neumüller, and D. Treutter, “Preliminary study on non-destructive assessment of European plum (*Prunus domestica* L.) maturity,” in *II International Symposium on Horticulture in Europe 1099*, pp. 257–262, 2012.
- [13] R. A. Harshman, “Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multimodal factor analysis,” 1970.

- [14] S. Wold, P. Geladi, K. Esbensen, and J. Ohman, "Multi-way principal components-and pls-analysis", *Journal of chemometrics*, vol 1, no1, pp. 41-56 1987.
- [15] D. B. Gil, A. M. De La Peña, J. A. Arancibia, G. M. Escandar, and A. C. Olivieri, "Second-order advantage achieved by unfolded-partial least-squares/residual bilinearization modeling of excitation-emission fluorescence data presenting inner filter effects," *Analytical Chemistry*, vol. 78, no. 23, pp. 8051–8058, 2006.
- [16] R. Tauler, "Multivariate curve resolution of multiway data using the multilinearity constraint," *Journal of Chemometrics*, vol. 35, no. 2, 2021.
- [17] J. Zhuang *et al.*, "Assessment of external properties for identifying banana fruit maturity stages using optical imaging techniques," *Sensors*, vol. 19, no. 13, 2019.
- [18] J. Sandak, A. Sandak, and M. Cocchi, "Multi-sensor data fusion and parallel factor analysis reveals kinetics of wood weathering," *Talanta*, vol. 225, 2021.
- [19] K. K. Beltrame *et al.*, "Digital images and independent components analysis in the determination of bioactive compounds from grape juice," *LWT*, vol. 152, 2021.
- [20] B. G. Botelho, K. C. F. Dantas, and M. M. Sena, "Determination of allura red dye in hard candies by using digital images obtained with a mobile phone and N-PLS," *Chemometrics and Intelligent Laboratory Systems*, vol. 167, pp. 44–49, 2017.

- [21] S. Elsayed, H. Galal, A. Allam, and U. Schmidhalter, “Passive reflectance sensing and digital image analysis for assessing quality parameters of mango fruits,” *Scientia Horticulturae*, vol. 212, pp. 136–147, 2016.
- [22] D. D. de S. Fernandes *et al.*, “Quantification and identification of adulteration in the fat content of chicken hamburgers using digital images and chemometric tools,” *LWT*, vol. 100, pp. 20–27, 2019.
- [23] C. Herrero-Latorre, J. Barciela-García, S. García-Martín, and R. M. Peña-Crecente, “Detection and quantification of adulterations in aged wine using RGB digital images combined with multivariate chemometric techniques,” *Food Chem X*, vol. 3, p. 100046, 2019.
- [24] J. Pardede, M. G. Husada, A. N. Hermana, and S. A. Rumapea, “Fruit Ripeness Based on RGB, HSV, HSL, L*a*b* Color Feature Using SVM,” in *2019 International Conference of Computer Science and Information Technology, (ICoSNIKOM) 2019*, Institute of Electrical and Electronics Engineers Inc., p. 1-5. Nov. 2019.
- [25] C. G. Reile, M. S. Rodríguez, D. D. de S. Fernandes, A. de A. Gomes, P. H. G. D. Diniz, and C. V. Di Anibal, “Qualitative and quantitative analysis based on digital images to determine the adulteration of ketchup samples with Sudan I dye,” *Food chemistry*, vol. 328, p. 127101, 2020.
- [26] D. J. Hart and K. J. Scott, “Analytical, Nutritional and Clinical Methods Section Development and evaluation of an HPLC method for the analysis of carotenoids in foods, and the measurement of the carotenoid content of vegetables and fruits commonly consumed in the UK,” *Food Chemistry*, vol. 54, pp. 101–111, 1995.

- [27] K. P. Wright and A. A. Kader, "Effect of controlled-atmosphere storage on the quality and carotenoid content of sliced persimmons and peaches," *Postharvest Biology and Technology*, vol. 10, no. 1, pp. 89–97, 1997.
- [28] E. Mellado-Ortega and D. Hornero-Méndez, "Effect of lutein esterification on the differential distribution of carotenoids in germ and endosperm fractions from tritordeum grains," *Journal of Cereal Science*, vol. 79, pp. 462–468, 2018.
- [29] A. Oliveira, E. M. C. Alexandre, M. Coelho, R. M. Barros, D. P. F. Almeida, and M. Pintado, "Peach polyphenol and carotenoid content as affected by frozen storage and pasteurization," *LWT - Food Science and Technology*, vol. 66, pp. 361–368, 2016.
- [30] P. Orazem, M. Mikulic-Petkovsek, F. Stampar, and M. Hudina, "Changes during the last ripening stage in pomological and biochemical parameters of the 'Redhaven' peach cultivar grafted on different rootstocks," *Scientia Horticulturae*, vol. 160, pp. 326–334, 2013.
- [31] F. Delgado-Vargas, A. R. Jiménez, and O. Paredes-López, "Natural pigments: carotenoids, anthocyanins, and betalains—characteristics, biosynthesis, processing, and stability," *Critical Reviews in Food Science and Nutrition*, vol. 40, no. 3, pp. 173–289, 2000.
- [32] C. Scattino *et al.*, "Post-harvest UV-B irradiation induces changes of phenol contents and corresponding biosynthetic gene expression in peaches and nectarines," *Food Chemistry*, vol. 163, pp. 51–60, 2014.

- [33] S. Jayasankar, C. Dowling, and D. K. Selvaraj, *Plums and Related Fruits*, 1st ed. Elsevier Ltd., 2016.
- [34] K. Ward, R. Scarth, J. K. Daun, and C. T. Thorsteinson, “A comparison of high-performance liquid chromatography and spectrophotometry to measure chlorophyll in canola seed and oil,” *Journal of the American Oil Chemists Society*, vol. 71, pp. 931–934, 1994.
- [35] S. Udenfriend, *Fluorescence assay in biology and medicine*, vol. 2. Academic Press, 2014.
- [36] T. Galeano Díaz, I. Durán Merás, C. Arturo Correa, B. Roldán, and M. Isabel Rodríguez Cáceres, “Simultaneous Fluorometric Determination of Chlorophylls a and b and Pheophytins a and b in Olive Oil by Partial Least-Squares Calibration,” *Journal of agricultural food chemistry*, vol. 51, no. 24, pp. 6934–6940, 2003.
- [37] M. G. Dias, M. F. G. F. C. Camões, and L. Oliveira, “Carotenoids in traditional Portuguese fruits and vegetables,” *Food Chemistry*, vol. 113, no. 3, pp. 808–815, 2009.
- [38] M. Cerrato-Alvarez, S. Frutos-Puerto, C. Miró-Rodríguez, and E. Pinilla-Gil, “Measurement of tropospheric ozone by digital image analysis of indigotrisulfonate-impregnated passive sampling pads using a smartphone camera,” *Microchemical Journal*, vol. 154, p. 104535, 2020.
- [39] L. F. Capitán-Vallvey, N. López-Ruiz, A. Martínez-Olmos, M. M. Erenas, and A. J. Palma, “Recent developments in computer vision-based analytical chemistry: A tutorial review,” *Analytical Chimica Acta*, vol. 899, pp. 23–56, 2015.

- [40] MathWorks, “MATLAB R2008a versión 7.6.0.324, MathWorks, Natick, MA, USA, 2008.” Natick, 2008.
- [41] R. Bro and H. A. L. Kiers, “A new efficient method for determining the number of components in PARAFAC models,” *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 17, no. 5, pp. 274–286, 2003.
- [42] A. C. Olivieri, “On a versatile second-order multivariate calibration method based on partial least-squares and residual bilinearization: Second-order advantage and precision properties,” *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 19, no. 4, pp. 253–265, 2005.
- [43] A. C. Olivieri, H. L. Wu, and R. Q. Yu, “MVC2: A MATLAB graphical interface toolbox for second-order multivariate calibration,” *Chemometrics and Intelligent Laboratory Systems*, vol. 96, no. 2, pp. 246–251, 2009.
- [44] A. Olivieri and G. M. Escandar, *Practical three-way calibration*. Elsevier, 2014.
- [45] MathWorks, “Matlab R2018a. Versión 7.5.0.342. Editor: MathWorks. Natick, MA, USA, 2018. Software .exe. Disponible en: <https://www.mathworks.com/products/matlab.html>.” 2018.
- [46] Microsoft Corporation, “‘Microsoft Excel 2016’, Versión 16.0.4266.1001, Redmond, WA, Microsoft Corporation, 2015.” Redmond, 2015.
- [47] W. S. , C. T. J. Rasband, “ImageJ. Versión 1.53K. Editor: Instituto Nacional de Salud de USA. Dominio público. Disponible en: <https://imagej.nih.gov/ij/>.”

- [48] CAMO, “The Unscrambler.” CAMO Software AS., Oslo, Norway, 2007.
- [49] R. A. Vlaic *et al.*, “The changes of polyphenols, flavonoids, anthocyanins and chlorophyll content in plum peels during growth phases: From fructification to ripening,” *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, vol. 46, no. 1, pp. 148–155, 2018.
- [50] Z. Luo, J. Xie, T. Xu, and L. Zhang, “Delay ripening of ‘Qingnai’ plum (*Prunus salicina* Lindl.) with 1-methylcyclopropene,” *Plant Science*, vol. 177, no. 6, pp. 705–709, 2009.
- [51] P. R. Patel and T. V. R. Rao, “Physiological changes in relation to growth and ripening of khirni [*Manilkara hexandra* (Roxb.) Dubard] fruit,” *Fruits*, vol. 64, no. 3, pp. 139–146, 2009.
- [52] V. Usenik, F. Stampar, and D. Kastelec, “Phytochemicals in fruits of two *Prunus domestica* L. plum cultivars during ripening,” *Journal of the Science of Food and Agriculture*, vol. 93, no. 3, pp. 681–692, 2013.
- [53] V. Usenik, D. Kastelec, R. Veberič, and F. Štampar, “Quality changes during ripening of plums (*Prunus domestica* L.),” *Food chemistry*, vol. 111, no. 4, pp. 830–836, 2008.
- [54] F. C. Stintzing, A. S. Stintzing, R. Carle, B. Frei, and R. E. Wrolstad, “Color and antioxidant properties of cyanidin-based anthocyanin pigments,” *Journal of agricultural and food chemistry*, vol. 50, no. 21, pp. 6172–6181, 2002.

- [55] J. Wang, H. Pan, R. Wang, K. Hong, and J. Cao, "Patterns of flesh reddening, translucency, ethylene production and storability of 'Friar' plum fruit harvested at three maturity stages as affected by the storage temperature," *Postharvest Biology and Technology*, vol. 121, pp. 9–18, 2016.
- [56] R. Bro, "PARAFAC. Tutorial and applications," *Chemometrics and intelligent laboratory systems*, vol. 38, no. 2, pp. 149–171, 1997.
- [57] A. Muñoz de la Peña, A. Espinosa Mansilla, D. González Gómez, A. C. Olivieri, and H. C. Goicoechea, "Interference-free analysis using three-way fluorescence data and the parallel factor model. Determination of fluoroquinolone antibiotics in human serum," *Analytical Chemistry*, vol. 75, no. 11, pp. 2640–2646, 2003.
- [58] O. Monago-Maraña, C. E. Eskildsen, A. Muñoz de la Peña, T. Galeano-Díaz, and J. P. Wold, "Non-destructive fluorescence spectroscopy combined with second-order calibration as a new strategy for the analysis of the illegal Sudan I dye in paprika powder," *Microchemical Journal*, vol. 154, p. 104539, 2020.
- [59] D. M. Haaland and E. V Thomas, "Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information," *Analytical Chemistry*, vol. 60, no. 11, pp. 1193–1202, 1988.
- [60] D. M. Haaland and E. V Thomas, "Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data," *Analytical Chemistry*, vol. 60, no. 11, pp. 1202–1208, 1988.
- [61] A. G. González, M. A. Herrador, and A. G. Asuero, "Intra-laboratory testing of method accuracy from recovery assays," 1999.

- [62] R. Vlaic *et al.*, “The Changes of Polyphenols, Flavonoids, Anthocyanins and Chlorophyll Content in Plum Peels during Growth Phases: from Fructification to Ripening,” *Not Bot Horti Agrobot Cluj Napoca*, vol. 46, pp. 148–155, Sep. 2017.
- [63] P. R. Patel and T. V. R. Rao, “Growth and Ripening in Black Plum [*Syzygium cumini* (L.) Skeels],” *International Journal of Fruit Science*, vol. 14, no. 2, pp. 147–156, 2014.
- [64] S. P. Yadav, Y. Ibaraki, and S. D. Gupta, “Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis,” *Plant Cell Tissue Organ-Culture (PCTOC)*, vol. 100, no. 2, pp. 183–188, 2010.
- [65] L. F. Sánchez-Sastre, N. M. Alte da Veiga, N. M. Ruiz-Potosme, P. Carrión-Prieto, J. L. Marcos-Robles, L. M. Navas-García & P. Martín-Ramos, “Assessment of RGB Vegetation Indices to Estimate Chlorophyll Content in Sugar Beet Leaves in the Final Cultivation Stage,” *AgriEngineering*, vol. 2, no. 1, pp. 128–149, 2020.
- [66] A. K. Dey, M. Sharma, and M. R. Meshram, “An Analysis of Leaf Chlorophyll Measurement Method Using Chlorophyll Meter and Image Processing Technique,” in *Procedia Computer Science*, Elsevier B.V., 2016, pp. 286–292.

- [67] O. Monago-Maraña, J. Domínguez-Manzano, A. Muñoz-de la Peña, and I. Durán-Merás, “Second-order calibration in combination with fluorescence fibre-optic data modelling as a novel approach for monitoring the maturation stage of plums,” *Chemometrics and Intelligent Laboratory Systems*, vol. 199, p. 103980, 2020.

V. CONCLUSIONES

Los resultados obtenidos del desarrollo de la presente tesis doctoral han permitido extraer las siguientes conclusiones respecto a los objetivos planteados:

1. En primer lugar, se ha llevado a cabo la detección y cuantificación de adulteración de aceite de oliva virgen extra empleando sus matrices de excitación – emisión (EEMs) combinadas con clasificación multivariante. Se ha demostrado la posibilidad de emplear señales no destructivas para medir la fluorescencia nativa de estos aceites en base a dos grupos de compuestos y sus respectivas regiones de fluorescencia: polifenoles y productos de oxidación, y clorofilas.

La combinación de los datos obtenidos de dichas EEMs con herramientas quimiométricas como PARAFAC, PARAFAC-LDA y U-PLS-DA ha permitido diferenciar aceites de oliva virgen extra puros de aquellos adulterados con un 16% y un 3% de aceite de oliva o aceite de oliva de orujo, respectivamente. Además, mediante U-PLS se ha podido cuantificar dicho nivel de adulteración, demostrando el potencial de esta metodología para una rápida detección de la adulteración.

2. En segundo lugar, se ha propuesto una metodología basada en fluorescencia *front-face*, en combinación con herramientas quimiométricas de segundo orden, para monitorizar aceite de oliva virgen extra sometido a diferentes procedimientos de irradiación UV, irradiación con luz solar y calentamiento térmico.

Se ha demostrado la capacidad de esta metodología combinando las EEMs con PARAFAC-LDA para discriminar entre muestras irradiadas y no irradiadas. Se ha obtenido una buena discriminación para tiempos de irradiación UV superiores a 15 minutos y tiempos de irradiación solar superiores a 30 minutos, siempre en el rango espectral de las clorofilas. En el caso del rango espectral de los compuestos fenólicos, la

discriminación entre muestras irradiadas y no irradiadas se ha llevado a cabo con efectividad para muestras irradiadas por encima de 15 minutos.

Con respecto a la temperatura, se ha obtenido buena discriminación para aquellas muestras calentadas durante 120 minutos a 80°C, con respecto a muestras no calentadas, no siendo apreciable discriminación para muestras calentadas a una menor temperatura y durante un menor tiempo.

La principal ventaja de esta metodología ha sido la posibilidad de estudiar los diferentes procesos oxidativos que pueden sufrir los aceites de oliva mediante metodologías no destructivas.

3. En tercer lugar, con respecto al estudio de la calidad de ciruelas, se ha desarrollado una metodología no destructiva que ha permitido monitorizar la maduración de las mismas a lo largo de diferentes semanas. Para ello, se han obtenido las EEMs de dichas muestras empleando una fibra óptica. Los datos generados han sido tratados con algoritmos de calibración de segundo orden.

La caracterización no destructiva de estas muestras ha sido empleada por primera vez en este estudio, y los resultados ofrecidos por dicha metodología mostraron a la clorofila como un buen indicador del proceso de maduración.

Además, se emplearon diferentes algoritmos de segundo orden para cuantificar las clorofilas en la piel de las ciruelas, obteniéndose buenos resultados para todos los algoritmos empleados, pero menores errores de predicción en el caso de U-PLS (RMSEP = 0,1 mg/g) y N-PLS (RMSEP= 0,1 mg/g), con respecto a PARFAFAC (RMSEP= 0,2 mg/g) y MCR-ALS (RMSEP= 0,2 mg/g).

4. Finalmente, la última metodología desarrollada en esta tesis doctoral, aplicada también sobre muestras de ciruelas, ha consistido en el procesamiento de imágenes digitales tomadas directamente de estas muestras. Para ello, se han empleado como datos los valores RGB obtenidos desde las imágenes de las ciruelas.

Mediante un análisis univariante de las ratios R, G y B para poder cuantificar el contenido en clorofila, previamente demostrado como un indicador de la maduración de las ciruelas, los mejores resultados se obtuvieron para el canal R, el canal complementario al que se observa visualmente (el verde).

El análisis multivariante de primer orden de los histogramas RGB permitió observar una distribución de las muestras diferenciadas entre las cuatro primeras semanas y las posteriores al aplicar PCA. Estos histogramas permitieron la cuantificación de clorofilas obteniéndose buenos resultados para el canal G y RGB, con coeficientes de determinación de 0,92 y 0,89, respectivamente, y REPs% de 11 y 12%, respectivamente.

Finalmente, se empleó la información correspondiente para cada pixel, obteniendo así una matriz para cada muestra en cada canal y se aplicaron algoritmos de segundo orden. En este caso, se obtuvieron buenos resultados de cuantificación cuando se emplearon U-PLS y N-PLS en los tres canales, obteniéndose menores REPs% para el caso del canal G (10 y 12%, respectivamente).

Con todo esto, se pudo concluir que esta metodología es adecuada para seguir el proceso de maduración de las ciruelas y la cuantificación de su contenido en clorofilas a lo largo del proceso de maduración.

