



FACULTAD DE CIENCIAS  
EXACTAS Y NATURALES

Universidad Nacional de La Pampa

**ESTUDIO DE AUTENTICACIÓN Y CALIDAD  
DE PRODUCTOS AGROALIMENTARIOS  
MEDIANTE EL USO DE TÉCNICAS  
QUIMIOMÉTRICAS Y DATOS MULTI-  
INSTRUMENTALES**

MUCH, Diego Gabriel

**TESIS PRESENTADA PARA OBTENER EL GRADO  
ACADEMICO DE LICENCIADO EN QUÍMICA**

SANTA ROSA, LA PAMPA

ARGENTINA

2021

## PRÓLOGO

Esta tesina es presentada como parte de los requisitos para alcanzar el grado académico de licenciado en Química de la Universidad Nacional de La Pampa. Este es un trabajo original que no ha sido presentado previamente para la obtención de otro título en esta universidad ni en otra institución académica. Se llevó a cabo en el Departamento de Química-FCEyN-UNLPam, durante el periodo comprendido entre marzo de 2019 y abril de 2021, bajo la Dirección de la Dra. Mirta Raquel Alcaraz y la Co-Dirección de la Dra. Silvana Mariela Azcarate.

12 de mayo, 2021



---

MUCH, Diego Gabriel

## AGRADECIMIENTOS

Años atrás me dijeron “*si caminas solo vas a ir más rápido, pero si lo haces acompañado, vas a llegar más lejos*”. Parece simple, inclusive lógico, pero son de esas frases que te quedan grabadas y te ayudan a poder concientizar que no estamos solos, que cada situación o etapas de la vida las atravesamos acompañados, **¡y que fantástico que así sea!**; nada sería lo mismo sin ese condimento especial que le agrega a nuestra vida el compartir momentos con otros. El camino recorrido en esta hermosa etapa no fue la excepción, tuve el placer de vivirlo y compartirlo con personas maravillosas, que desde su lugar aportaron piezas para formar el rompecabezas final. Es por ello que utilizo este espacio para manifestar mi más modesto y sentido agradecimiento.

A Quela Alcaraz, mi directora, que sin conocerme aceptó dirigirme y se comprometió desde el primer momento para que esta tesina llegue a buen puerto. Por los conocimientos que me aportaste y por el enorme tiempo que me dedicaste, inclusive el destinado para levantarme el ánimo cuando más lo necesitaba.

A Silvana Azcarate, mi otra “directora”, que me dio la posibilidad de explorar este maravilloso mundo que es la quimiometría. Por haber confiado en mí desde un primer momento, por los conocimientos compartidos y por haberme tenido la suficiente paciencia durante todos estos años.

A mi familia, especialmente a mis padres Margarita y Pablo, que con su amor y apoyo incondicional hicieron este camino mucho más ameno y fácil de recorrer.

A Oli, mi gran sostén. Por las largas noches de estudio en donde su compañía siempre era gratificante y por, sobre todo, por los momentos que compartimos diariamente.

A Dai y Anita, amigas incondicionales que tuve el privilegio de conocer gracias a la carrera. Por su forma de ser para conmigo, y por, sobre todo, por los hermosos momentos que compartimos por fuera de la facultad.

A Viki y Moni, por esas charlas donde hablábamos sobre cosas que poco tenían que ver con la carrera, pero servían para crecer internamente.

A Mari, gran persona que tuve la suerte de conocer un poco más en detalle. Arrancó siendo mi primer “jefa”, y hoy ya es una gran amiga. Por sus palabras, los momentos compartidos en el bunker y los cafecitos de la mañana.

A todos los integrantes del Laboratorio de Espectrometrías Atómicas, particularmente a Flor, que con la mejor predisposición siempre estuvo para ayudar, despejar dudas y por,

sobre todo, tener charlas entretenidas, con la típica pizca de razonamiento lógico que la caracteriza.

A mis amigos de la secundaria, por los momentos de risas y diversión, inclusive en los momentos donde la rutina me sobrepasaba.


A la Facultad de Ciencias Exactas y Naturales, de la Universidad Nacional de la Pampa, por la educación de excelencia, gratuita y de calidad brindada.

A todos los docentes y profesionales que aportaron su granito de arena en mi crecimiento y desarrollo profesional.

Gracias a todos los que, de una forma u otra, estuvieron a lo largo de este camino recorrido.

Simplemente ¡GRACIAS!

12 de mayo, 2021



---

MUCH, Diego Gabriel

# ÍNDICE GENERAL

ÍNDICE DE FIGURAS .....	II
ÍNDICE DE TABLAS .....	III
RESUMEN .....	IV
SUMMARY .....	V
1 INTRODUCCIÓN .....	1
2 OBJETIVOS .....	11
2.1 OBJETIVOS GENERALES .....	12
2.2 OBJETIVOS ESPECÍFICOS .....	12
3 MATERIALES Y MÉTODOS .....	13
3.1 INSTRUMENTAL .....	14
3.2 MUESTRAS .....	15
3.2.1 Muestras de AOEV .....	15
3.2.2 Muestras de vinos .....	17
3.3 ANÁLISIS DE DATOS .....	18
3.4 PROGRAMAS INFORMÁTICOS .....	21
4. RESULTADOS Y DISCUSIÓN .....	22
4.1 ANÁLISIS DE ADULTERACIÓN EN AOEV .....	23
4.1.1 Consideraciones generales .....	23
4.1.2 Análisis quimiométrico de los datos .....	25
4.1.2.1 Análisis de componentes principales (PCA) .....	25
4.1.2.2 Estudios de predicción .....	30
4.2 CLASIFICACIÓN DE VINOS MONOVARIETALES MALBEC .....	35
4.2.1 Consideraciones generales .....	35
4.2.2 Análisis quimiométrico de los datos .....	37
4.2.2.1 Análisis de componentes principales (PCA) .....	37
4.2.2.2 Análisis de clasificación mediante PCA-DA .....	41
5. CONCLUSIÓN .....	44
6. REFERENCIAS .....	47

# ÍNDICE DE FIGURAS

<b>Figura 1.1.</b> Representación esquemática de la información instrumental según el orden de los datos obtenidos para una única muestra. ....	5
<b>Figura 1.2.</b> Esquema del proceso de fusión de datos de nivel bajo (azul), de nivel medio (naranja) y de nivel alto (verde) para el caso de un análisis de alimentos por diferentes técnicas espectroscópicas. ....	8
<b>Figura 1.3.</b> Proporciones de <b>A)</b> alimentos comúnmente adulterados y <b>B)</b> técnicas espectroscópicas utilizadas para su análisis, reportados en publicaciones de la última década (Gráficos adaptados de <i>esteki et al., 2018</i> ). ....	9
<b>Figura 3.1.</b> Esquematación del proceso de fusión de datos para el análisis de AOEV y de vinos. ....	20
<b>Figura 4.1.</b> Espectros de absorción UV-Vis de las muestras de AOEV analizadas. ....	23
<b>Figura 4.2.</b> Espectros de absorción NIR de las muestras de AOEV analizadas. ....	24
<b>Figura 4.3.</b> EEM de la muestra B3 de AOEV sin adulterar ( <b>A</b> ) y con 15 % de adulterante ( <b>B</b> ), mostrando las regiones entre 640 nm-700 nm ( <b>A.1</b> y <b>B.1</b> ) y entre 430 nm-640 nm ( <b>A.2</b> y <b>B.2</b> ). ....	25
<b>Figura 4.4.</b> Gráfico de <i>loadings</i> para los tres primeros PC obtenidos de PCA para los datos <b>a)</b> UV-Vis y <b>b)</b> NIR. ....	26
<b>Figura 4.5.</b> Gráfico de <i>scores</i> en el espacio tridimensional para los tres primeros PC de las diferentes matrices de datos instrumentales codificados según el % de adulteración (arriba) y según la marca de AOEV (abajo). Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional. ....	27
<b>Figura 4.6.</b> Gráfico de <i>scores</i> en el espacio tridimensional para los tres primeros PC de los diferentes arreglos de datos para los AOEV B3 ( <b>A.1, B.1, C.1</b> ) y B5 ( <b>A.2, B.2, C.2</b> ). Las líneas de guiones indican el eje central del espacio tridimensional. La línea negra sólida identifica la separación imaginaria de las muestras en grupos por % de adulterante. ....	29
<b>Figura 4.7.</b> Concentración de adulterante nominal vs. predicha, utilizando: <b>A)</b> UV-Vis, <b>B)</b> NIR, <b>C)</b> EEM desdobladas, <b>D)</b> fusión de nivel bajo (UV-vis -NIR-EEM desdobladas) y <b>E)</b> fusión de nivel medio. ....	32
<b>Figura 4.8.</b> Región elíptica de confianza conjunta en el plano pendiente-ordenada al origen, para un nivel de confianza del 95 %, indicándose el punto ideal [1, 0] (estrella negra) para pendiente y ordenada al origen, respectivamente, para los datos UV-Vis (rojo), NIR (azul), EEM (negro), fusión nivel bajo (verde) y fusión nivel medio (celeste). ....	33
<b>Figura 4.9.</b> Espectros UV-Vis de las 24 muestras de vino Malbec procedentes de tres regiones geográficas diferentes. ....	35
<b>Figura 4.10.</b> Espectros NIR de las 24 muestras de vino Malbec procedentes de tres regiones geográficas diferentes. ....	36
<b>Figura 4.11.</b> Espectros UV-vis originales ( <b>A.1</b> ) y pretratados con MC ( <b>A.2</b> ) y espectros NIR originales ( <b>B.1</b> ) y derivados + MC ( <b>B.2</b> ). ....	37
<b>Figura 4.12.</b> <b>A.1</b> y <b>A.2</b> Gráficos de <i>loadings</i> obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente. <b>B.1</b> y <b>B.2</b> gráficos de <i>scores</i> , obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente, agrupados según su procedencia geográfica. <b>C.1</b> y <b>C.2</b> gráficos de <i>scores</i> , obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente, codificados según la bodega de producción. Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional. ....	38
<b>Figura 4.13.</b> Datos UV-Vis-NIR fusionados originales ( <b>A</b> ) y pre-procesados ( <b>B</b> ), <i>loadings</i> ( <b>C</b> ) y <i>scores</i> del conjunto de datos fusionados, clasificados según su procedencia geográfica ( <b>D</b> ) y bodega de producción ( <b>E</b> ). Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional. ....	40

# ÍNDICE DE TABLAS

<b>Tabla 3.1.</b> Composición porcentual de AOEV adulterados, según el diseño experimental combinado construido. ....	15
<b>Tabla 3.2.</b> Origen geográfico, contenido alcohólico y pH de los vinos Malbec analizados.....	18
<b>Tabla 4.1.</b> Varianza explicada de los conjuntos de datos correspondientes a cada marca de AOEV .....	30
<b>Tabla 4.2.</b> Resultados de predicción obtenidos para cada modelo PLS construido. ....	31
<b>Tabla 4.3.</b> Parámetros estadísticos de los diferentes modelos PLS construidos. ....	31
<b>Tabla 4.4.</b> Matrices de confusión de los modelos PCA-DA. ....	42
<b>Tabla 4.5.</b> Índices de clasificación de los conjuntos de calibración y predicción para los diferentes arreglos de datos analizados.....	42

## RESUMEN

En la actualidad, el fraude vinculado a los alimentos es un problema creciente, donde situaciones de falsificación o el agregado intencional de sustancias adquieren cada vez más protagonismo y se tornan, a su vez, más sofisticados. En pos de superar estas dificultades, resulta necesario disponer de métodos analíticos rápidos y robustos, capaces de ser utilizados en los controles de rutina de productos alimentarios. En este contexto, diferentes técnicas espectroscópicas, en combinación con herramientas de análisis matemáticos y estadísticos son propuestas en la literatura como una opción prometedora para resolver cuestiones vinculadas al fraude alimentario.

En la presente tesina se realizó el estudio de dos alimentos consumidos masivamente: aceites de oliva extra virgen y vinos monovarietales Malbec. Las muestras se analizaron utilizando técnicas espectroscópicas de absorción molecular UV-Visible, de infrarrojo cercano y fluorescencia multidimensional basada en la generación de matrices de excitación-emisión. Adicionalmente, y con la finalidad de evaluar la sinergia de la información procedente de cada fuente instrumental, se implementaron estrategias de fusión de datos. Asimismo, se hizo partícipe a la quimiometria como herramienta de análisis, con el fin de desarrollar diferentes modelos de calibración y clasificación capaces de predecir la cantidad total de sustancia adulterante presente en el aceite de oliva extra virgen y discriminar los vinos según su procedencia geográfica. Además, se realizó un estudio comparativo entre modelos evaluando la eficiencia en términos de capacidad predictiva.

Los resultados demostraron el potencial uso de técnicas espectroscópicas en combinación con la quimiometria, como una opción confiable para los análisis de procedencia geográfica y autenticidad en vinos y aceites.



## SUMMARY

Nowadays, food fraud is a growing problem, where situations of counterfeiting or the intentional addition of substances acquire more importance and become more sophisticated too. To overcome these difficulties, fast and robust analytical methods, capable of being used in the routine controls of food products are necessary. In this context, different spectroscopic techniques, in combination with statistical analysis tools, have been proposed in the literature as a promising option for solving issues related to food fraud.

In this thesis, the study of two different food was carried out: Malbec wines and extra virgin olive oils. The data were obtained from UV-Vis spectroscopy, near infrared spectroscopy and excitation-emission fluorescence. Additionally, data fusion strategies have been implemented to evaluate the synergy of the information acquired from each instrumental source. Also, chemometrics was utilized as analysis tool to develop different calibration and classification models that allow predicting the amount of adulterant and discriminating wines according to their geographical origin. In addition, a comparative study between models was carried out evaluating the efficiency in terms of predictive capacity.

The results demonstrated the potential use of non-destructive spectroscopic techniques, in combination with chemometrics, as a reliable option for geographical origin and authenticity analyzes in wines and oils.

1

# Introducción

---

# 1 INTRODUCCIÓN

La calidad de los alimentos destinados al consumo humano se define no solo por el contenido nutricional, sino también por el sabor y la apariencia del producto. Asimismo, la frescura, los ingredientes y el origen son atributos importantes que aportan garantía de calidad. Durante la última década, el análisis de autenticación de productos agroalimentarios ha tomado mayor relevancia debido a un notable incremento en la ocurrencia de adulteraciones basadas, principalmente, en la incorporación deliberada de sustancias químicas que, bajo un marco legal establecido, se encuentran estrictamente reguladas, y en algunos casos, prohibidas. Los alimentos más propensos a ser manipulados intencionalmente son las especias, aceites comestibles, miel, leche y sus derivados, bebidas alcohólicas, café y productos cárnicos (*Rodriguez-Saona et al., 2011; Lakshmi, 2012*).

En la investigación por posible fraude alimentario, existen dos problemas principales: 1) la autenticación del origen de un producto, en términos de su procedencia geográfica u origen botánico/animal, y 2) el análisis de adulteración por la adición de alguna sustancia prohibida o por encima de los límites permitidos (*Caballero et al., 2020*).

La autenticación de productos según su origen geográfico reviste de interés ya que es una propiedad que incrementa su valor comercial, si bien no representa un riesgo para la salud humana. En este contexto, para garantizar la autenticidad del producto, existen leyes o regímenes que brindan protección a productos agrícolas y alimentarios como, por ejemplo, la denominación de origen protegida (DOP) o la indicación geográfica protegida (IGP) que garantizan y reconocen una calidad diferenciada del producto (*Azcarate et al., 2015; Callao et al., 2018; Rios Reina et al., 2020*). Por el contrario, las adulteraciones en los alimentos constituyen un alto riesgo para la salud humana. La adulteración de los productos alimenticios se encuentra ligada a un beneficio económico por parte de los productores que realizan el reemplazo de ciertos ingredientes por otros de menor valor o utilizan sustancias en exceso capaces de enmascarar algún defecto del producto, ya sea en su apariencia o su sabor (*Kelsey et al., 2021*).

Los análisis de autenticación y adulteración de un producto se orientan desde la perspectiva legal y se basan en una respuesta cualitativa o cuantitativa que permita garantizar la calidad o inocuidad del producto (*Medina et al., 2019*). El análisis de adulteración puede tener varios enfoques; por un lado, la determinación de sustancias permitidas en los alimentos se realiza mediante un análisis cuantitativo que permite obtener el contenido exacto de la sustancia de interés y, de esta manera, verificar que cumpla con los umbrales

establecidos por la legislación vigente. Por el contrario, en el caso de tratarse de una sustancia prohibida, un análisis cualitativo de respuesta binaria que permita la verificación de la presencia/ausencia de la sustancia evaluada es suficiente para establecer la calidad e inocuidad del producto. (Callao *et al.*, 2018).

Para hacer frente a la creciente tendencia en el fraude alimentario, es necesario disponer de métodos analíticos rápidos, confiables y competentes que aborden esta problemática y garanticen la calidad de los resultados. Entre los métodos comúnmente utilizados para la identificación de sustancias no permitidas y la autenticación de calidad de los alimentos se encuentran los experimentos físico-químicos (Danezis *et al.*, 2016), los análisis metabolómicos (Piovesana *et al.*, 2020) y la utilización de técnicas separativas (Esteki *et al.*, 2018). Si bien estas metodologías proporcionan resultados altamente confiables, requieren de largos tiempos de análisis, alto consumo de solventes y reactivos, de tediosos tratamientos de muestras y de personal entrenado.

Actualmente, las investigaciones analíticas se orientan hacia el desarrollo de herramientas que brinden soluciones rápidas, no destructivas, de bajo costo y amigables con el ambiente. En este aspecto, las técnicas basadas en mediciones espectroscópicas se presentan como las alternativas más atractivas. En particular, estas técnicas han mostrado un notable crecimiento en el campo del análisis de alimentos debido a su principal ventaja que es la ausencia o minimización del pre-tratamiento de la muestra para su análisis. La espectroscopia infrarroja (IR) es una técnica versátil utilizada para el análisis cuantitativo y descriptivo de una amplia variedad de muestras. En espectroscopia IR del rango infrarrojo cercano (NIR), las señales de absorbancia están asociadas a vibraciones moleculares y fundamentales. Los enlaces químicos como  $-CO$ ,  $-OH$  y  $-NH$  tienen altas frecuencias de vibración, dando como resultado bandas detectables en la región del infrarrojo cercano, i.e. 780-2500 nm (Osborne *et al.*, 1993; Stuart, 2004). Este tipo de espectroscopia se usa habitualmente para el análisis cualitativo y cuantitativo de productos agrícolas y alimenticios con la gran ventaja de requerir poca o ninguna preparación de la muestra (Blanco y Villarroya, 2002). Por su parte, la espectroscopia de absorción molecular UV-Visible (UV-Vis) es una de las técnicas más utilizadas para la cuantificación de especies absorbentes en el análisis de muestras de diferente naturaleza (Cordella *et al.*, 2002). Por otro lado, la fluorescencia molecular es una técnica analítica que se basa en la detección de la luz emitida por un fluoróforo posterior a la absorción de una luz incidente (proceso de excitación). Por lo general, la información registrada en la espectroscopia de fluorescencia se presenta como espectros de emisión (Lakowicz, 2006). No obstante, la utilización de espectroscopia no

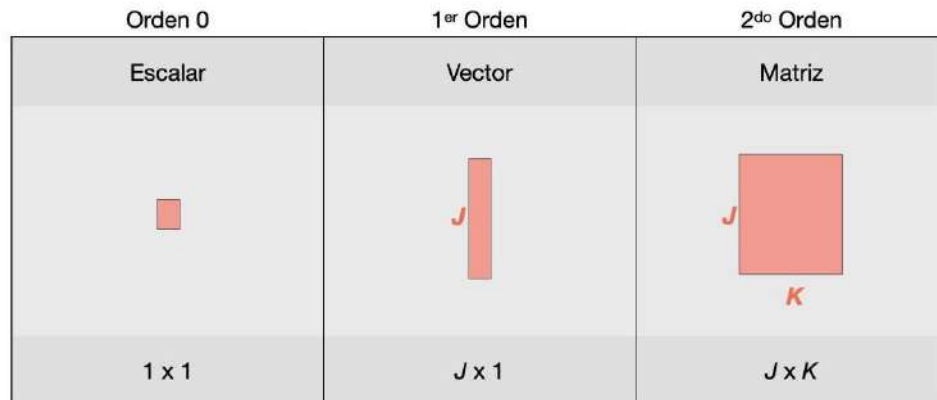
convencional de fluorescencia multidimensional, basada en la generación de matrices de excitación-emisión (EEM), se muestra como una poderosa herramienta de análisis debido a la gran cantidad de información colectada en un único registro. Para la obtención de una EEM se registran varios espectros de emisión a diferentes longitudes de onda de excitación en un rango determinado. Esta estructura de datos tiene la ventaja de detectar constituyentes de la muestra que presenten diferentes propiedades fluorescentes, por ejemplo, bandas de emisión máxima a diferentes longitudes de onda (Lakhal et al., 2014; Kumar et al., 2017).

Todas estas técnicas analíticas permiten la obtención de una gran cantidad de información que puede ser utilizada para realizar una caracterización global cuantitativa y cualitativa del sistema en estudio. Sin embargo, esta característica no siempre es una ventaja debido a las dificultades que se presentan durante el análisis y la manipulación de datos de grandes dimensiones, limitando la extracción de información relevante. La complejidad de los datos adquiridos obliga a utilizar herramientas de análisis matemático o estadístico a fin de explorar la información adquirida.

Una de las disciplinas más explotadas en el ámbito científico para el análisis de datos complejos, y que se está abriendo paso en el ámbito de la industria, es la quimiometría. Esta herramienta permite resolver problemas descriptivos y predictivos y ha sido definida por Luc Massart, en 1998, como “*La disciplina química que utiliza la matemática, la estadística y la lógica formal para diseñar o seleccionar procedimientos experimentales óptimos; proporcionar la máxima información química relevante a partir del análisis de datos químicos y obtener conocimiento de sistemas químicos*” (Massart et al., 1998). Las aplicaciones descriptivas se basan en el modelado de las propiedades químicas de los sistemas a fin de comprender su comportamiento, mientras que las aplicaciones predictivas se focalizan en el modelado de estas propiedades con el objetivo de poder predecirlas en sistemas nuevos de variada composición.

Para la construcción de modelos quimiométricos que permitan la extracción de información de un sistema en estudio es muy importante reconocer el número de modos o dimensiones instrumentales (orden o vía) que presenten. El orden de los datos y el arreglo en el que se disponen representan el primer condicionamiento en la implementación de las herramientas quimiométricas. Las dimensiones del dato dependerán del diseño de los experimentos, de los instrumentos utilizados y de la forma en que se organizan. (Alcaraz, 2016). Los datos generados para una sola muestra se clasifican según su *orden*, mientras que los datos adquiridos para un grupo de muestras se clasifican según el número de *vías* en los

que se disponen. La figura 1.1 muestra la clasificación que pueden recibir los datos instrumentales para una única muestra según el *orden*.



**Figura 1.1.** Representación esquemática de la información instrumental según el orden de los datos obtenidos para una única muestra.

Los datos de orden cero son generados por instrumentos que registran un solo valor numérico (escalar) por cada objeto (muestra) como, por ejemplo, los electrodos ion selectivos, los espectrofotómetros de absorción molecular a una única longitud de onda, entre otros. Al aumentar un orden en los datos, se generan los llamados datos de primer orden que son señales vectoriales adquiridas para cada objeto. Los ejemplos más claros de este tipo de señales son los espectros de absorción UV-Vis, los espectros de absorbancia NIR, espectros de emisión de fluorescencia, voltamperogramas, entre otros. Un orden de dato superior lo constituyen los datos de segundo orden que se representan como arreglos matriciales, conteniendo dos dimensiones instrumentales, o experimentales, diferentes. Estos datos pueden obtenerse directamente como señal matricial para una única muestra utilizando un único instrumento (espectroscopia de fluorescencia de EEM), mediante el acoplamiento de distintas técnicas analíticas (cromatografía acoplada a detección espectral) o por monitoreo de alguna propiedad de la muestra a diferentes condiciones experimentales (evoluciones espectrales a diferentes tiempos de adquisición para el monitoreo de un proceso evolutivo) (Alcaraz, 2016).

Por otro lado, para un conjunto de muestras, los datos pueden disponerse en arreglos del mismo orden o de orden superior, construyendo arreglos de datos de N-vías que contienen un modo o dimensión adicional correspondiente al número de muestras. Por ejemplo, los datos de primer orden pueden disponerse en arreglos vectoriales por concatenación unidireccional obteniendo un arreglo de 1-vía aumentada, o en arreglos matriciales por apilamiento de los vectores individuales generando un arreglo de 2-vías.

Asimismo, un conjunto de datos de segundo orden dispuestos en un arreglo tridimensional (o cubo) conforma un arreglo de 3-vías, pero, también, pueden disponerse en matrices apiladas conformando un arreglo de 2-vías aumentado y así sucesivamente (Alcaraz, 2016).

En un procedimiento experimental, la quimiometría está involucrada en varios de los pasos que lo constituyen, desde el diseño de los experimentos hasta la obtención e interpretación de resultados, siendo la etapa de análisis la más intervenida por esta disciplina. En este aspecto, los tratamientos aplicados en el manejo de datos pueden clasificarse, en grandes rasgos, dependiendo de la etapa en la que se lleva a cabo el análisis (Kharbach et al., 2020):

(i) **Diseño de experimentos:** la experimentación juega un papel fundamental en los campos de la investigación y el desarrollo. Es una metodología basada en la aplicación de herramientas matemáticas y estadísticas cuyo objetivo es obtener información de calidad ayudando al experimentador a seleccionar la estrategia experimental óptima que permita obtener la información buscada y evaluar los resultados experimentales obtenidos, garantizando la máxima fiabilidad en las conclusiones que se obtengan. Una de sus tantas aplicaciones radica en la preparación de muestras de calibración y validación para estimar una curva de calibración y posteriormente validar su ajuste.

(ii) **Pre-procesamiento:** tratamiento previo de datos necesario para reducir las interferencias no deseadas, sin afectar la información útil. Implementado con el fin de favorecer el posterior modelado de los datos (Rinnan et al., 2009; Lohumi et al., 2015). La importancia de este procedimiento radica en conseguir una mejora sustancial del dato, sin la introducción de señales no deseadas o la deformación del dato original (Anderson et al., 2010). Por ejemplo, substracción de ruido instrumental, eliminación de línea de base, corrección de señales de dispersión, entre otras.

(iii) **Tratamiento de datos propiamente dicho:** es la etapa más representativa del estudio e incluye una serie de análisis que, en combinación, llevan a la resolución del problema analítico. Entre ellos se pueden nombrar:

a. *Análisis exploratorio:* es un análisis que puede revelar la existencia de una relación entre las muestras, es decir, una agrupación natural o una tendencia en la distribución. La técnica exploratoria no supervisada más popular es el análisis de componentes principales (PCA) (Esbensen et al. 2009) que permite generar nuevas variables (componentes principales, PC) como una combinación lineal de las variables originales, reteniendo la máxima información del dato original (Lavine, 2006; Esbensen et al., 2009; Li Vigni et al., 2013).

b. *Construcción del modelo quimiométrico*: dependiendo del objetivo del análisis, y del orden y propiedades matemáticas de los datos adquiridos, se construyen modelos que expliquen la variabilidad y el comportamiento de los datos.

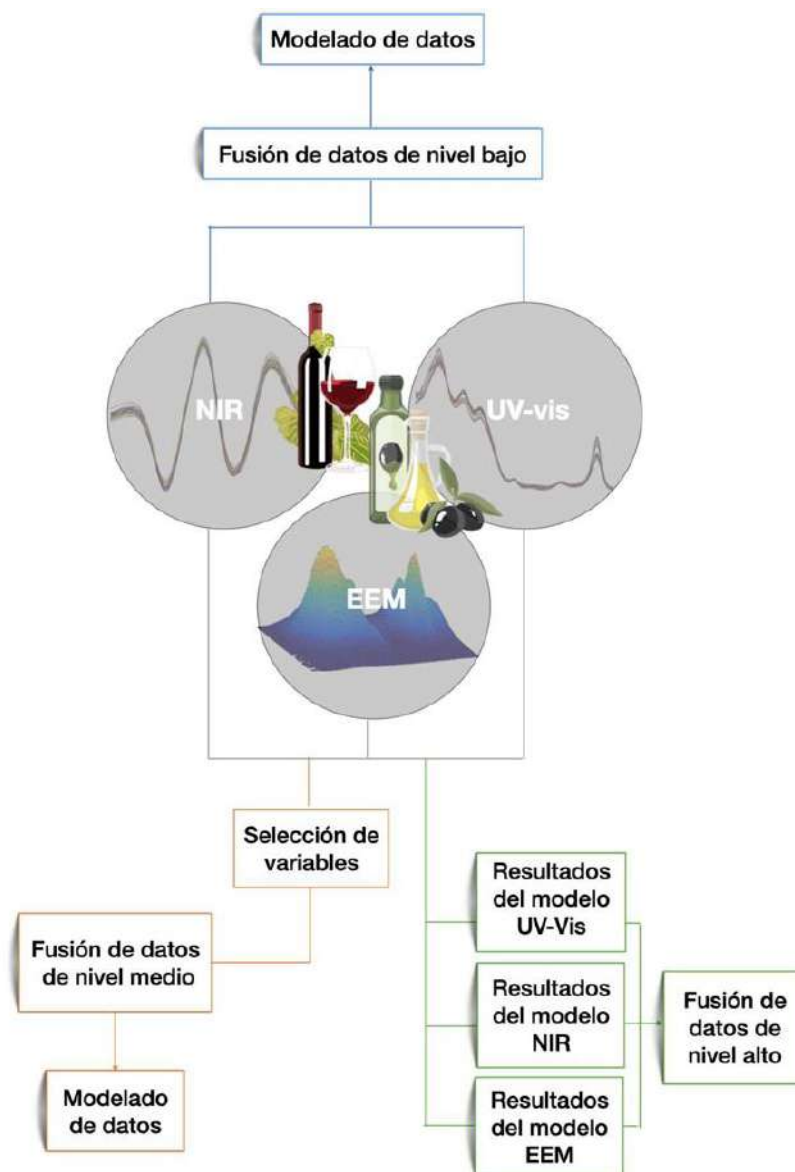
(iv) **Interpretación de los modelos construidos y los resultados alcanzados**: es la última etapa del análisis y depende, en gran parte, del criterio analítico y la experiencia del analista. Asimismo, permite reevaluar las etapas anteriores de manera transversal al procedimiento analítico, con la posibilidad de reestructurar el método desarrollo y el análisis posterior a fin de obtener resultados representativos y fiables del sistema en estudio.

En métodos de clasificación, uno de los principales objetivos finales es asignar una muestra desconocida a una (o ninguna) de las clases previamente designadas por un modelo ajustado y construido con muestras conocidas (*Callao et al., 2018*). Algunos de los métodos discriminantes más utilizados son el PCA y cuadrados mínimos parciales en combinación con análisis discriminante (PCA-DA y PLS-DA) (*Silvestri et al., 2014; Rácz et al., 2016*). Por el otro lado, en los métodos de calibración, el objetivo final es la predicción de la concentración de una sustancia de interés en una muestra particular. Para ello, se deben construir modelos que permitan relacionar la señal instrumental adquirida con la concentración conocida de la sustancia (*Brereton, 2000*). Posteriormente, con el modelo construido es posible predecir valores de concentración desconocidos correspondientes al analito en muestras de composición variada. Los métodos de calibración para datos de primer orden se denominan métodos de calibración multivariada y los modelos comúnmente utilizados son cuadrados mínimos parciales (PLS), proyecciones ortogonales a estructuras latentes (OPLS), entre otros (*Trygg et al., 2002; Rodriguez et al., 2019; Kharbach et al., 2020*).

En ocasiones, es posible realizar diferentes mediciones sobre una misma muestra y, si bien una única fuente puede brindar información analítica relevante, una muestra podría ser mejor caracterizada mediante el análisis conjunto de su información analítica multiplataforma, principalmente en muestras de alta complejidad, como, por ejemplo, en alimentos. Esta estrategia de análisis se conoce como fusión de datos y se refiere al análisis conjunto de datos fusionados de distinta naturaleza. Durante los últimos años, se registró un aumento en el uso de la fusión de los datos analíticos en el campo de la autenticidad de alimentos, (*Zhou et al., 2020*) debido a que, generalmente, arroja resultados más representativos y globales que los obtenidos del análisis de datos de técnicas individuales (*Borrás et al., 2015*). En química analítica, la fusión de datos puede clasificarse en tres categorías dependiendo de

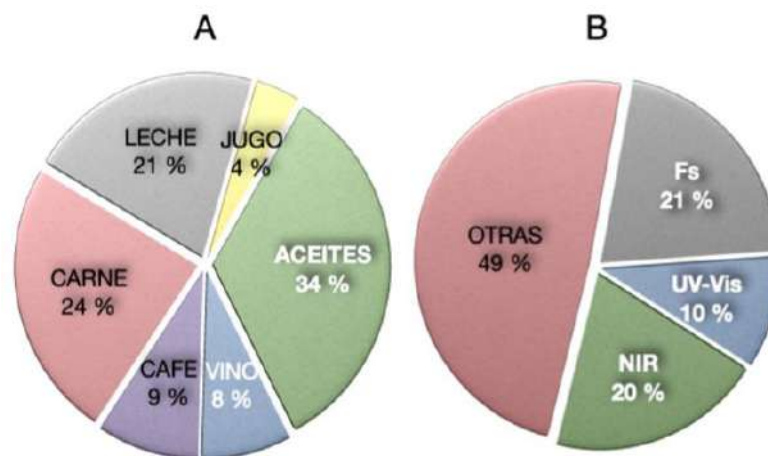


la combinación en la que se produzca la fusión (Godinho et al., 2015; Biancolillo et al., 2019). La fusión de datos de nivel bajo se lleva a cabo concatenando las diferentes matrices de datos en un solo arreglo, construyendo un único modelo final (Khaleghi et al., 2013; Borrás et al., 2015). En la fusión de nivel medio, las características más relevantes de las diferentes señales instrumentales se extraen por separado (mediante métodos de selección de variables o algoritmos de extracción) y luego se fusionan en una sola matriz que se utiliza para construir el modelo final (Spiteri et al., 2016). Por último, en la fusión de nivel alto, a partir de la información proveniente de cada fuente instrumental, se construyen modelos individuales y los resultados de cada uno se combinan para obtener la declaración de identidad final (Doeswijk et al., 2011). Estos niveles se representadas esquemáticamente en la figura 1.2.



**Figura 1.2.** Esquema del proceso de fusión de datos de nivel bajo (azul), de nivel medio (naranja) y de nivel alto (verde) para el caso de un análisis de alimentos por diferentes técnicas espectroscópicas.

El uso combinado de técnicas analíticas espectroscópicas en colaboración con la quimiometría resulta de gran utilidad para las investigaciones vinculadas a la autenticidad y adulteración de alimentos, reflejándose en el gran número de trabajos reportados en la bibliografía en los últimos diez años. Entre ellos, se reporta la utilización de espectroscopia NIR para detectar adulteraciones de melamina en leche a partir de modelos de regresión PLS (Chen *et al.*, 2017); la aplicación de fluorescencia molecular y UV-Vis junto con la habilidad predictiva de LDA para clasificar café *robusta* y café *arábica* en mezclas de café (Dankowska *et al.*, 2017); el uso de fluorescencia multidimensional junto a diferentes modelos predictivos para evaluar la presencia de caramelo en muestras de vinagres de vino (Rios Reina *et al.*, 2019); diferentes muestras de vino blanco se clasificaron en función del origen varietal y geográfico (Azcarate *et al.*, 2015); la aplicación de UV-Vis, NIR y fluorescencia con estrategias de fusión de datos para detectar aceites de oliva (AO) vencidos en AO extra virgen (AOEV) recién elaborado (Ozen *et al.*, 2019), entre otros. La figura 1.3 resume esquemáticamente la relación de trabajos publicados en el ámbito del análisis agroalimentario.



**Figura 1.3.** Proporciones de A) alimentos comúnmente adulterados y B) técnicas espectroscópicas utilizadas para su análisis, reportados en publicaciones de la última década (Gráficos adaptados de esteki *et al.*, 2018).

En nuestro país, según el informe del Ministerio de Relaciones Exteriores, Comercio Internacional y Culto, entre los productos manufacturados más exportados se encuentran los AOEV y los vinos de diversos varietales, debido a la elevada calidad y prestigio internacional que presentan (Ministerio de relaciones exteriores, Comercio Internacional y Culto, 2020). Por un lado, entre los aceites vegetales, el AO virgen (AOV) y el AOEV presentan características nutricionales y sensoriales que los hacen componentes únicos y básicos de la dieta (Caramia *et al.*, 2012). De estos aceites, el AOEV es uno de los aceites comestibles más costoso en el mercado nacional, debido a su calidad y autenticidad (Aparicio *et al.*, 2013). Es por ello, que su

adulteración es una de las más frecuentes, generalmente por adición de aceites del mismo tipo, pero más económicos, como el AO refinado, o de otro tipo de aceites aún más económicos, como los refinados de girasol, maíz o soja, entre otros (Aparicio *et al.*, 2013). En Argentina, el Código Alimentario Argentino establece en el Art. 535 (CAA, 1953) que todo aceite declarado bajo el rótulo de Extra Virgen “*serán los obtenidos a partir del fruto del olivo exclusivamente*” y, por lo tanto, no se permite el agregado o presencia de cualquier otro tipo de aceite comestible. Por el otro lado, la autenticidad y el valor comercial de los vinos están vinculados, normalmente, a su origen geográfico. Por lo tanto, es esperable que aquellas regiones productoras de vinos de alto valor comercial utilicen categorías como DOP o IGP para demostrar su procedencia y garantizar su calidad (Azcarate *et al.*, 2015). Si bien el valor económico de los vinos está mediado por factores como el proceso de producción, la cosecha, la variedad de uva y la concentración de alcohol (Palade *et al.*, 2014), muchos de ellos obtienen un mayor valor agregado al estar reconocidos bajo una DOP, haciéndolos más vulnerables a sufrir fraude (Ohtsubo *et al.* 2008). Por todo ello, frente al actual contexto de un mercado cada vez más globalizado resulta imprescindible disponer de herramientas de análisis que permitan verificar la calidad e inocuidad de estos productos.

En resumen, es importante considerar no solo la amplia cantidad de información que se puede obtener a partir de la utilización de múltiples técnicas analíticas, sino también, la calidad de información que se puede extraer a partir de la implementación de modelos quimiométricos. De esta manera, las técnicas analíticas combinadas, en conjunto con la quimiometría, se presentan como una alternativa novedosa para el análisis de adulteración y autenticación de alimentos. En este trabajo se presenta el análisis integral de dos alimentos de gran valor económico en nuestro país: AOEV y vino Malbec, utilizando técnicas analíticas espectroscópicas NIR, UV-Vis y de fluorescencia en combinación con diferentes herramientas quimiométricas

2

# Objetivos

---

## 2 OBJETIVOS

### 2.1 OBJETIVOS GENERALES

Obtener modelos de clasificación y calibración que permitan evaluar la autenticidad y calidad de muestras alimentarias, a partir de la información obtenida por diversas técnicas espectroscópicas.

### 2.2 OBJETIVOS ESPECÍFICOS

Generar datos de primer orden a partir de técnicas espectroscópicas UV-Vis y NIR, y de segundo orden mediante la generación de EEM para el análisis de AOEV adulterados con aceites comestibles refinados.

Generar datos de primer orden utilizando espectroscopia de absorción molecular UV-Vis y NIR para el análisis de muestras de vinos Malbec procedentes de diversas áreas geográficas.

Aplicar recursos quimiométricos a los datos obtenidos en el análisis de las muestras de AOEV y obtener modelos cuantitativos y cualitativos eficientes para el estudio de adulteración.

Aplicar herramientas quimiométricas a los datos obtenidos en el análisis de las muestras de vino Malbec y obtener modelos eficientes que permitan identificar la procedencia geográfica.

3

**Materialles**  
**&**  
**Métodos**

---

## 3 MATERIALES Y MÉTODOS

### 3.1 INSTRUMENTAL

Las medidas de absorción UV-Vis se realizaron con un espectrofotómetro Ocean Optics CHEMUSB4 con sistema de detección por arreglo de diodo lineal y de configuración haz simple. El control del instrumento y la adquisición espectral se realizaron utilizando el programa informático OceanView (versión 1.6.7; Ocean Optics, The Netherlands). Se registraron los espectros en el rango espectral de 200 nm a 800 nm (dejando solo la región informativa entre 320 nm-745 nm para AOEV y entre 280 nm-650 nm para vinos) cada 0,21 nm, utilizando una cubeta de cuarzo de 1 cm de paso óptico.

Las mediciones NIR se realizaron cubriendo el rango espectral de 900 nm-1650 nm utilizando un espectrofotómetro FLAME-NIR Ocean Optics (Duiven, The Netherlands), equipado con celda de cuarzo con un paso óptico de 1 cm para las medidas de transmitancia en el análisis de AOEV y de 0,5 cm para el análisis de vino. El control del instrumento y la adquisición espectral se realizaron con el programa informático OceanView (versión 1.6.7; Ocean Optics, The Netherlands). Los espectros se adquirieron con una resolución de 6 nm. Para la calibración del instrumento se utilizó un cuerpo negro correspondiente al 0 % de transmitancia y la cubeta de cuarzo vacía correspondiente para la señal de 100 % de transmitancia.

Todas las medidas de fluorescencia se realizaron en un espectrofotómetro de fluorescencia Agilent Cary Eclipse (Agilent Technologies, Waldbronn, Alemania), utilizando cubeta de 1x1 cm de paso óptico, controlado con el programa Cary WinFLR (Agilent Technologies, Waldbronn, Alemania). Las EEM se adquirieron variando la longitud de onda de excitación en el rango espectral comprendido entre 280 nm y 580 nm, cada 3 nm, y registrando los espectros de emisión entre 420 nm y 700 nm, cada 2 nm, utilizando aperturas de rendija fijadas en 10 nm para excitación y emisión, 430 V de fotomultiplicador y una velocidad de barrido de 9600 nm s<sup>-1</sup>.

Las mediciones de pH se realizaron con un pH-metro F42 Horiba (Tokio, Japón) equipado con un electrodo combinado de vidrio y electrodo de referencia interna de Ag-AgCl.

## 3.2 MUESTRAS

### 3.2.1 Muestras de AOEV

Se analizaron cinco muestras comerciales de AOEV de diferentes marcas: Ciudad del Lago (B1), Oliovita (B2), El mistol (B3), La Toscana (B4) y Zuelo (B5). Los aceites refinados de maíz, girasol y soja utilizados como adulterantes fueron mezclas de 2 marcas diferentes para cada uno. Todos los aceites fueron comercialmente adquiridos de comercios locales.

Para la preparación de las muestras adulteradas se construyó un diseño experimental que consistió en la combinación de dos diseños D-optimal para mezclas cuaternarias. Para el primer diseño, se evaluaron concentraciones de adulterantes en cinco niveles de concentración oscilando en el rango de 0 %-15 % (% v/v), mientras que, para el segundo diseño, las concentraciones de los adulterantes se evaluaron entre 0 % y 7 % (% v/v), en cinco niveles de concentración. Los diseños se construyeron considerando tres tipos de adulterantes, con la condición de que la composición final de la mezcla no supere el valor 100. De esta manera, el diseño experimental completo consistió en un total de 47 muestras adulteradas y 5 no adulteradas, separadas en 5 bloques, donde cada bloque representa una marca de AOEV diferente. En la tabla 3.1 se detallan las muestras construidas.

**Tabla 3.1.** Composición porcentual de AOEV adulterados, según el diseño experimental combinado construido.

Muestra <sup>1</sup>	AOEV (mL)	Maíz (mL)	Girasol (mL)	Soja (mL)	% Adulteración
B1	10,00	0,00	0,00	0,00	0,0
B2	10,00	0,00	0,00	0,00	0,0
B3	10,00	0,00	0,00	0,00	0,0
B4	10,00	0,00	0,00	0,00	0,0
B5	10,00	0,00	0,00	0,00	0,0
B1_1	9,30	0,00	0,00	0,70	7,0
B1_2	9,80	0,00	0,00	0,20	2,0
B1_3	9,80	0,00	0,00	0,20	2,0
B1_4	9,55	0,23	0,22	0,00	5,5
B1_5	9,30	0,40	0,00	2,30	7,0
B2_6	9,80	0,20	0,00	0,00	2,0



Muestra	AOEV (mL)	Maíz (mL)	Girasol (mL)	Soja (mL)	% Adulteración
B2_7	9,55	0,00	0,45	0,00	4,5
B2_8	9,30	0,35	0,35	0,00	7,0
B2_9	9,30	0,70	0,00	0,00	7,0
B2_10	9,55	0,00	0,45	0,00	4,5
B3_11	9,55	0,00	0,00	0,45	4,5
B3_12	9,55	0,00	0,00	0,45	4,5
B3_13	9,80	0,00	0,20	0,00	2,0
B3_14	9,63	0,36	0,00	0,00	3,6
B3_15	9,30	0,00	0,70	0,00	7,0
B4_16	9,30	0,00	0,70	0,00	7,0
B4_17	9,30	0,70	0,00	0,00	7,0
B4_18	9,30	0,00	0,35	0,35	7,0
B4_19	9,30	0,00	0,35	0,35	7,0
B4_20	9,30	0,35	0,00	0,35	7,0
B5_21	9,46	0,00	0,53	0,00	5,3
B5_22	9,30	0,00	0,00	0,70	7,0
B5_23	9,30	0,00	0,00	0,70	7,0
B5_24	9,30	0,70	0,00	0,00	7,0
B1_25	8,50	0,75	0,75	0,00	15,0
B1_26	9,80	0,00	0,00	0,20	2,0
B1_27	8,50	0,00	0,00	1,50	15,0
B1_28	9,80	0,00	0,00	0,20	2,0
B1_28	8,93	0,00	1,07	0,00	10,7
B2_29	8,50	0,00	0,00	1,50	15,0
B2_30	8,50	0,00	1,50	0,00	15,0
B2_31	8,50	1,50	0,00	0,00	15,0
B2_32	9,80	0,20	0,00	0,00	2,0
B2_33	9,80	0,20	0,00	0,00	2,0
B3_34	8,50	0,75	0,00	0,75	15,0
B3_35	9,37	0,63	0,00	0,00	6,3
B3_36	9,15	0,00	0,00	0,85	8,5
B3_37	8,50	0,00	1,00	0,50	15,0
B3_38	9,15	0,00	0,00	0,85	8,5

Muestra	AOEV (mL)	Maíz (mL)	Girasol (mL)	Soja (mL)	% Adulteración
B4_39	9,37	0,00	0,63	0,00	6,3
B4_40	8,50	0,00	0,75	0,75	15,0
B4_41	8,50	1,50	0,00	0,00	15,0
B4_42	9,37	0,00	0,63	0,00	6,3
B4_43	8,50	0,00	1,50	0,00	15,0
B5_44	8,50	0,00	1,50	0,00	15,0
B5_45	9,80	0,00	0,20	0,00	2,0
B5_46	8,93	0,57	0,00	5,00	10,7
B5_47	9,80	0,00	0,20	0,00	2,0

<sup>1</sup> B1: Ciudad del Lago; B2: Oliyovita; B3: El mistol; B4: La Toscana; B5: Zuelo. Las muestras adulteradas se indican como B#\_XX.

Las muestras adulteradas se prepararon por adición directa de los aceites refinados al AOEV. Luego de la preparación, las muestras se agitaron mecánicamente en vortex durante 1 minuto y se almacenaron en la oscuridad a 4 °C hasta el análisis. Todas las muestras se analizaron al menos 24 h posterior a su preparación, a fin de favorecer la homogeneización de la solución preparada.

### 3.2.2 Muestras de vinos

Se analizó un total de 24 muestras de vinos monovarietales Malbec comerciales procedentes de tres regiones productoras de vinos argentinos: Mendoza este, Luján de Cuyo y Valle de Uco. Las muestras de vino fueron elaboradas por dos bodegas diferentes de la ciudad de Mendoza, Mendoza, Argentina: Bodega Ruca Malén y Bodega Doña Paula, que aplicaron idénticos protocolos de producción para la elaboración de los vinos.

El contenido alcohólico de las muestras osciló entre 12,2 % y 13,8 % (v/v). En la tabla 3.2 se resume la información referida a las muestras analizadas. Las muestras originales se fraccionaron y las fracciones se conservaron en un lugar fresco, protegidas de la luz hasta su análisis. Para las mediciones de absorbancia UV-Vis, las muestras se diluyeron con agua ultrapura (Millipore UV synergy System, Billerica, MA, USA) en un factor de 1/10. Las mediciones NIR se realizaron con las muestras originales sin diluir.

**Tabla 3.2.** Origen geográfico, contenido alcohólico y pH de los vinos Malbec analizados.

Muestra <sup>1</sup>	Región	% Etanol (v/v)	pH
1 RM	Mendoza este	11,9	3,75
2 RM	Mendoza este	12,9	3,73
3 RM	Luján de Cuyo	13,7	3,67
4 RM	Luján de Cuyo	13,4	3,68
5 RM	Valle de Uco	14,4	3,75
6 RM	Valle de Uco	13,0	3,66
7 RM	Luján de Cuyo	11,8	3,67
8 RM	Luján de Cuyo	14,5	3,67
9 RM	Luján de Cuyo	13,5	3,66
10 DP	Valle de Uco	14,3	3,86
11 DP	Valle de Uco	14,0	3,96
12 DP	Valle de Uco	13,6	4,17
13 DP	Valle de Uco	14,7	3,91
14 DP	Mendoza este	13,1	3,95
15 DP	Valle de Uco	13,5	3,88
16 DP	Valle de Uco	13,3	3,78
17 DP	Mendoza este	13,0	3,67
18 DP	Mendoza este	13,1	3,62
19 DP	Valle de Uco	14,6	3,76
20 DP	Valle de Uco	14,5	3,88
21 DP	Luján de Cuyo	14,0	3,86
22 DP	Mendoza este	12,4	3,76
23 DP	Mendoza este	13,9	3,78
24 DP	Valle de Uco	14,3	3,68

<sup>1</sup>RM: Bodega Ruca Malén; DP: Bodega Doña Paula

### 3.3 ANÁLISIS DE DATOS

Para el análisis de datos UV-Vis y NIR de AOEV se construyeron arreglos de 2-vías muestras  $\times$  variables, de dimensiones  $52 \times 2125$  y  $52 \times 128$ , respectivamente. Para vinos, los arreglos de 2-vías fueron de dimensiones  $24 \times 1836$  y  $24 \times 144$  para UV-Vis y NIR,

respectivamente. Las EEM de dimensiones  $141 \times 101$  para excitación y emisión, respectivamente, se desdoblaron de manera de obtener vectores  $1 \times 14241$  para cada muestra de AOEV. Posteriormente, se construyó un arreglo de 2-vías, apilando las señales vectorizadas de cada muestra, obteniéndose una matriz de datos de  $52 \times 14241$ .

En primer lugar, las señales fueron sometidas a un procedimiento de pre-tratamiento de datos. Para AOEV, las señales de UV-Vis obtenidas fueron corregidas con variación normal estándar (SNV), para corregir variaciones en la línea base, mientras que las NIR fueron sujetas a corrección de dispersión multiplicativa (MSC) con la finalidad de corregir desplazamientos espectrales debido a los diferentes componentes de las muestras. Las señales de fluorescencia fueron sometidas a un procedimiento de corrección de señales de dispersión aplicando la estrategia de conservación de señal (EEM\_cor *Chiappini et al., 2019*). Para los datos fusionados de nivel bajo se utilizó el método escalado por desvío estándar (VSS) debido que las escalas de las señales instrumentales son diferentes; de esta manera, todas las señales son comparables en escala, evitando errores en el modelo por diferencias en la sensibilidad instrumental. Por último, para los datos fusionados de nivel medio se utilizó valor absoluto (AV) y VSS.

Por otro lado, para el análisis de vino, a los datos UV-Vis se les aplicó centrado medio (MC) y los datos NIR fueron transformados utilizando la primera derivada a fin de eliminar los términos constantes a todas las longitudes de onda y sujetos, posteriormente a SNV. Por último, para el análisis de datos fusionados, se construyeron arreglos con los datos originales que luego fueron sujetos a la combinación SNV-MC como pretratamiento.

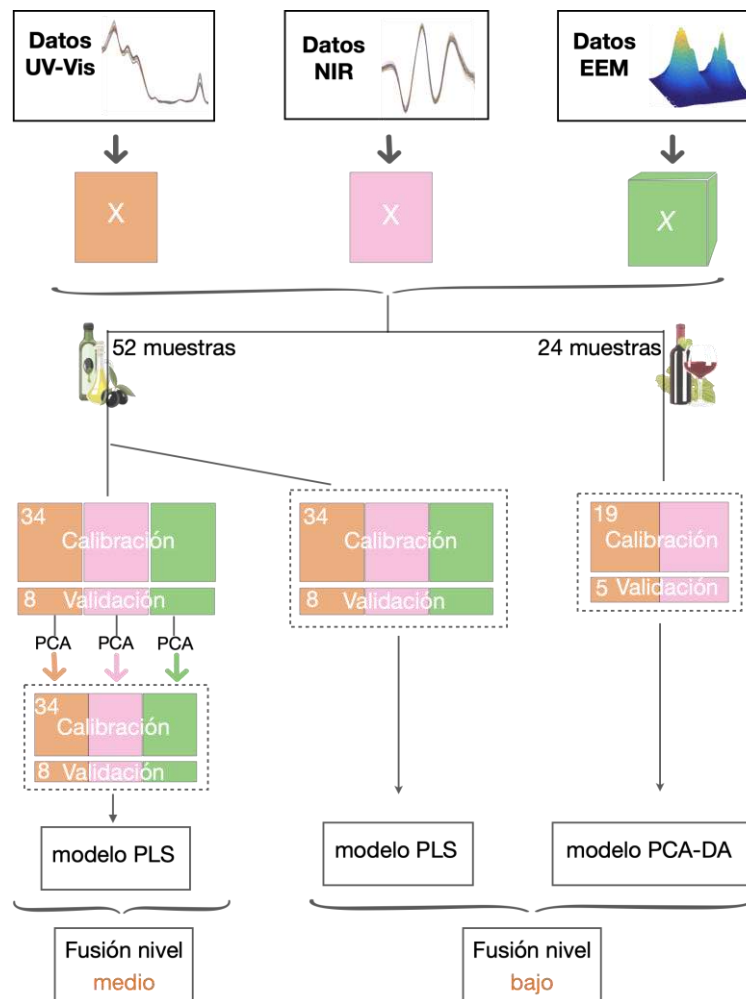
En todos los casos, antes del modelado propiamente dicho, se implementó PCA como análisis exploratorio sobre los datos espectrales completos. Con los resultados se realizaron diagramas de dispersión de los PC para realizar una inspección visual de las muestras, detectar valores atípicos y evaluar similitudes y diferencias. Adicionalmente se estudiaron los *loadings* para evaluar la información comprendida en los rangos espectrales.

Adicionalmente, para el estudio de adulteración de AOEV se construyeron modelos de calibración PLS dividiendo aleatoriamente el total de muestras en conjuntos de calibración y validación. La validación interna se realizó mediante validación cruzada con las muestras de calibración y la validación externa se realizó mediante un análisis de predicción de las muestras del conjunto de validación. El número de variables latentes (LV) se seleccionó teniendo en cuenta la raíz del error cuadrático medio de la calibración (RMSEC) y de la validación cruzada (RMSECV). Para el análisis mediante la estrategia de fusión, se realizaron arreglos de datos a nivel bajo. Para la calibración de las muestras de

AOEV se evaluó, por un lado, la estrategia de fusión de nivel bajo concatenando los bloques de datos UV-Vis, NIR y EEM desdoblados en una única matriz de  $52 \times 18017$  y construyendo el modelo de calibración mediante PLS. Asimismo, se implementó una estrategia de nivel medio extrayendo los *scores* de cada bloque de datos mediante PCA, que luego se concatenaron en una única matriz de  $52 \times 15$  para construir, finalmente, el modelo de calibración PLS.

Para el análisis de clasificación de vinos, las muestras se separaron aleatoriamente en el conjunto de calibración y el conjunto de predicción. El modelo de clasificación se construyó mediante PCA-DA. El modelo construido se validó internamente mediante validación cruzada y, externamente, utilizando el conjunto de predicción. Para la clasificación de las muestras de vinos se realizó, además, una fusión de datos utilizando los bloques procedentes de las mediciones de UV-Vis y NIR que se concatenaron en una única matriz de  $24 \times 1964$ , Posteriormente, se realizó un análisis exploratorio mediante PCA y se aplicó un PCA-DA para obtener el modelo final de clasificación.

Las estrategias de fusión para cada análisis se esquematizan en la figura 3.1.



**Figura 3.1.** Esquematización del proceso de fusión de datos para el análisis de AOEV y de vinos.

### 3.4 PROGRAMAS INFORMÁTICOS

El modelado de datos y análisis quimiométrico se realizaron utilizando el paquete quimiométrico PLS\_Toolbox, versión trial (Eigenvector Research Inc., Wenatchee, WA) y el paquete de algoritmos classification Toolbox escrito en entorno MATLAB v.8.5.0 (The Mathworks Inc., Natick, MA) disponible gratuitamente en internet en <https://michem.unimib.it/download/matlab-toolboxes/classification-toolbox-for-matlab/> (fecha de último acceso: 10 de mayo de 2021).

Los procedimientos de pre-procesamiento implementados para las señales UV-Vis y NIR, como SNV (Barnes et al., 1989), MSC (Geladi et al., 1985), VSS, CM y primera derivada de Savitzky-Golay (Savitzky&Golay, 1964) se realizaron utilizando el paquete PLS\_toolbox, mientras que la corrección de señales de dispersión de las EEM se realizó con el programa EEM\_corr (Chiappini *et al.*, 2019), asequible en internet accediendo a la página <http://www.fcb.unl.edu.ar/laboratorios/ladaq/> (fecha de último acceso: 05 de mayo de 2021).

Para la preparación de las muestras de AOEV adulteradas se construyó un diseño experimental utilizando el software Desing Expert 11 (Stat-Ease Inc., Minneapolis, Estados Unidos). Las muestras pertenecientes a cada set de calibración y validación se seleccionaron aleatoriamente utilizando el algoritmo de Kennard-Stone (*Kennard and Stone, 1969*) implementado desde una interfaz gráfica soportada en MATLAB v.8.5.0, desarrollada por un investigador colaborador del grupo de investigación.

4

Resultados  
&  
Discusiones

---

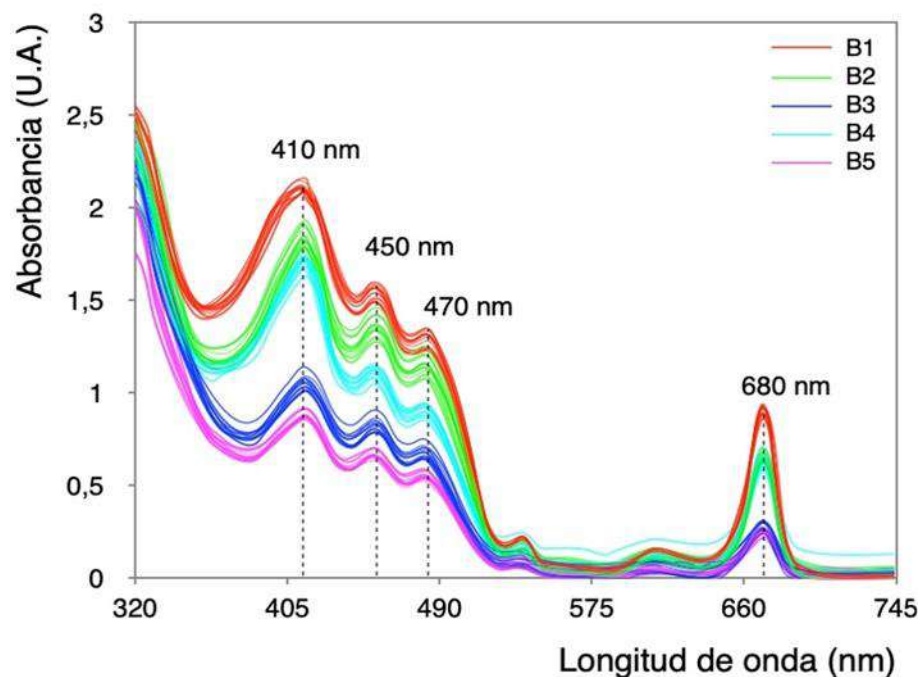
## 4 RESULTADOS Y DISCUSIÓN

### 4.1 ANÁLISIS DE ADULTERACIÓN EN AOEV

#### 4.1.1 Consideraciones generales

Los AOEV se obtienen del fruto del olivo (*Olea europaea L.*) y es reconocido por sus propiedades funcionales debido al alto contenido de ácidos grasos monoinsaturados y antioxidantes naturales. Además, en su composición, presentan una fracción saponificable, compuesta por triglicéridos y ácidos grasos libres, que representa casi la totalidad del aceite, y una fracción insaponificable, compuesta por otros componentes, representando un 1 %-3 % de la composición total. Sin embargo, esta última fracción es la responsable de las propiedades organolépticas y de las características de los AOEV que lo distingue frente a otros AO u otros aceites vegetales refinados. Por ejemplo, las elevadas cantidades de clorofila, carotenoides y derivados que presentan los AOEV es una característica que los diferencia de otros aceites de AO, como los AOV, AO refinados y otros aceites comestibles. Esta particularidad se presenta ya que los procesos involucrados en la elaboración de los aceites refinados afectan a estos compuestos orgánicos, disminuyendo así su concentración en el producto final, fenómeno que no ocurre en los AOEV que son elaborados por prensado en frío (Moncayo, 2016).

La figura 4.1 muestra los espectros UV-Vis de los diferentes analizados.

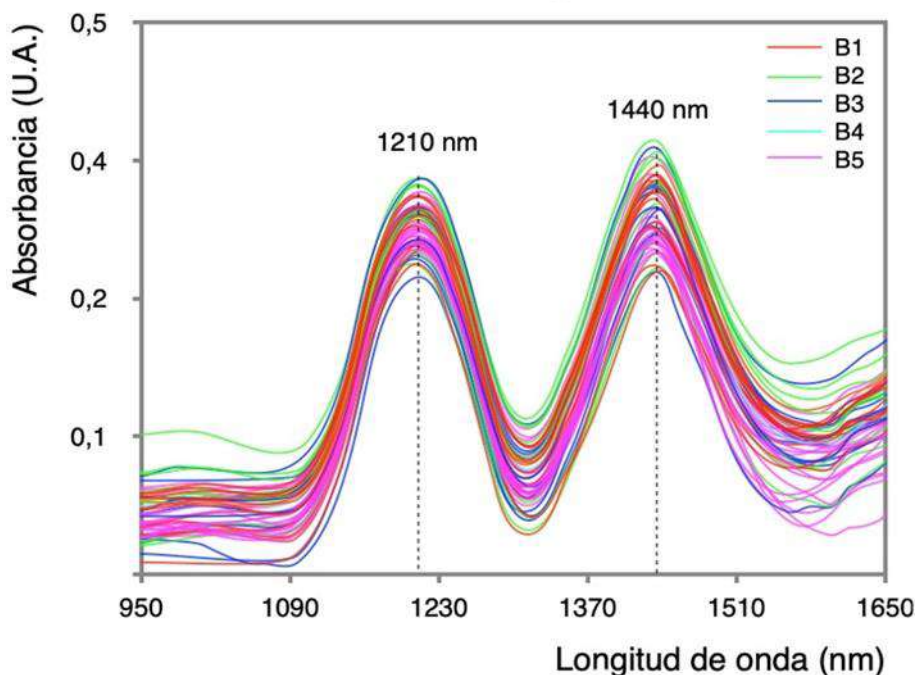


**Figura 4.1.** Espectros de absorción UV-Vis de las muestras de AOEV analizadas.



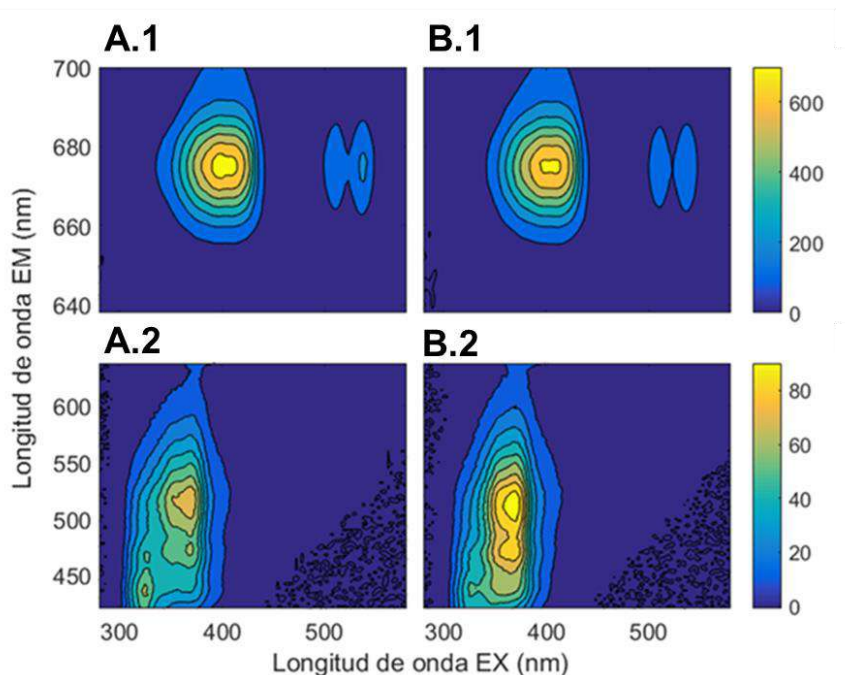
Las bandas espectrales del espectro UV-Vis características de AO se relacionan con la presencia de carotenoides y clorofila y sus derivados. Los carotenoides presentan bandas de absorción bien definidas a 410 nm, 450 nm y 470 nm, mientras que la banda espectral a 680 nm corresponde a las clorofilas y feofitinas (Tarakowski *et al.*, 2014) propias del AO.

Un típico espectro NIR de un AO (figura 4.2) presenta una banda de absorción a 1210 nm propia de los segundos armónicos vibracionales de estiramiento  $-CH$ , presentes en los triacilglicerol, mientras que las bandas de  $\sim 1450$  nm- son atribuibles a los primeros armónicos vibracionales de estiramiento  $-CH$  de grupos  $-CH_3$ ,  $-CH_2$  y  $HC=CH$  de las cadenas de ácidos grasos insaturados (Yang Li *et al.*, 2019),



**Figura 4.2.** Espectros de absorción NIR de las muestras de AOEV analizadas.

Por otro lado, la fluorescencia en los AOEV se atribuye a sus fluoróforos naturales, como los compuestos fenólicos y tocoferoles, productos de oxidación de ácidos grasos, vitamina E y la clorofila junto con sus productos de degradación con una banda característica de emisión a  $\sim 680$  nm (Zandomenighi *et al.*, 2005; Guzman *et al.*, 2015).



**Figura 4.3.** EEM de la muestra B3 de AOEV sin adulterar (A) y con 15 % de adulterante (B), mostrando las regiones entre 640 nm-700 nm (A.1 y B.1) y entre 430 nm-640 nm (A.2 y B.2).

En las señales UV-Vis de los AOEV adulterados, la intensidad de absorbancia entre 400 nm-500 nm disminuye a medida que aumenta la cantidad de aceite adulterante por efecto de dilución de los componentes naturales del AOEV. Este mismo efecto se evidencia en las bandas de emisión de fluorescencia correspondientes a las clorofilas y sus derivados, a 680 nm, que disminuye ante el aumento de la concentración de aceite refinado agregado. Sin embargo, la intensidad de emisión de fluorescencia en el rango entre 400 nm y 600 nm muestra un aumento al incrementar la concentración de adulterante, debido a la presencia de productos de oxidación de ácidos grasos relacionados a los procesos de refinación de los aceites refinados (Kongbonga *et al.*, 2011). Por el contrario, en los espectros NIR no se observa un patrón en dependencia a los diferentes % de adulteración que permita la diferenciación entre muestras.

## 4.1.2 Análisis quimiométrico de los datos

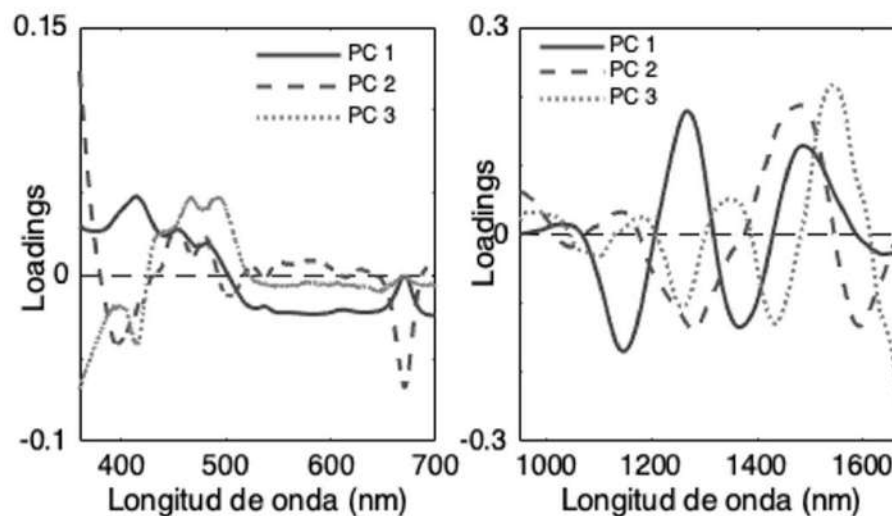
### 4.1.2.1 *Análisis de componentes principales (PCA)*

Un paso preliminar en el análisis de datos es la búsqueda de agrupaciones naturales entre las muestras (Rebolo *et al.*, 2000). Para ello, el estudio de adulteración de AOEV, se realizó un análisis por componentes principales (PCA) a cada set de datos instrumentales obtenidos. El análisis por PCA permite reducir la dimensionalidad de los datos a un pequeño

número de componentes con el fin de examinar posibles agrupamientos y visualizar diferencias entre muestras (Naes et al., 2002; Otto et al., 2007). Asimismo, como resultado del análisis PCA, se obtiene un sub-set de vectores, denominados *loadings*, que pueden entenderse como la variable de contribución de cada PC en particular. Un *loading* con valores positivos indica que la variable se correlaciona positivamente con el PC correspondiente y viceversa.

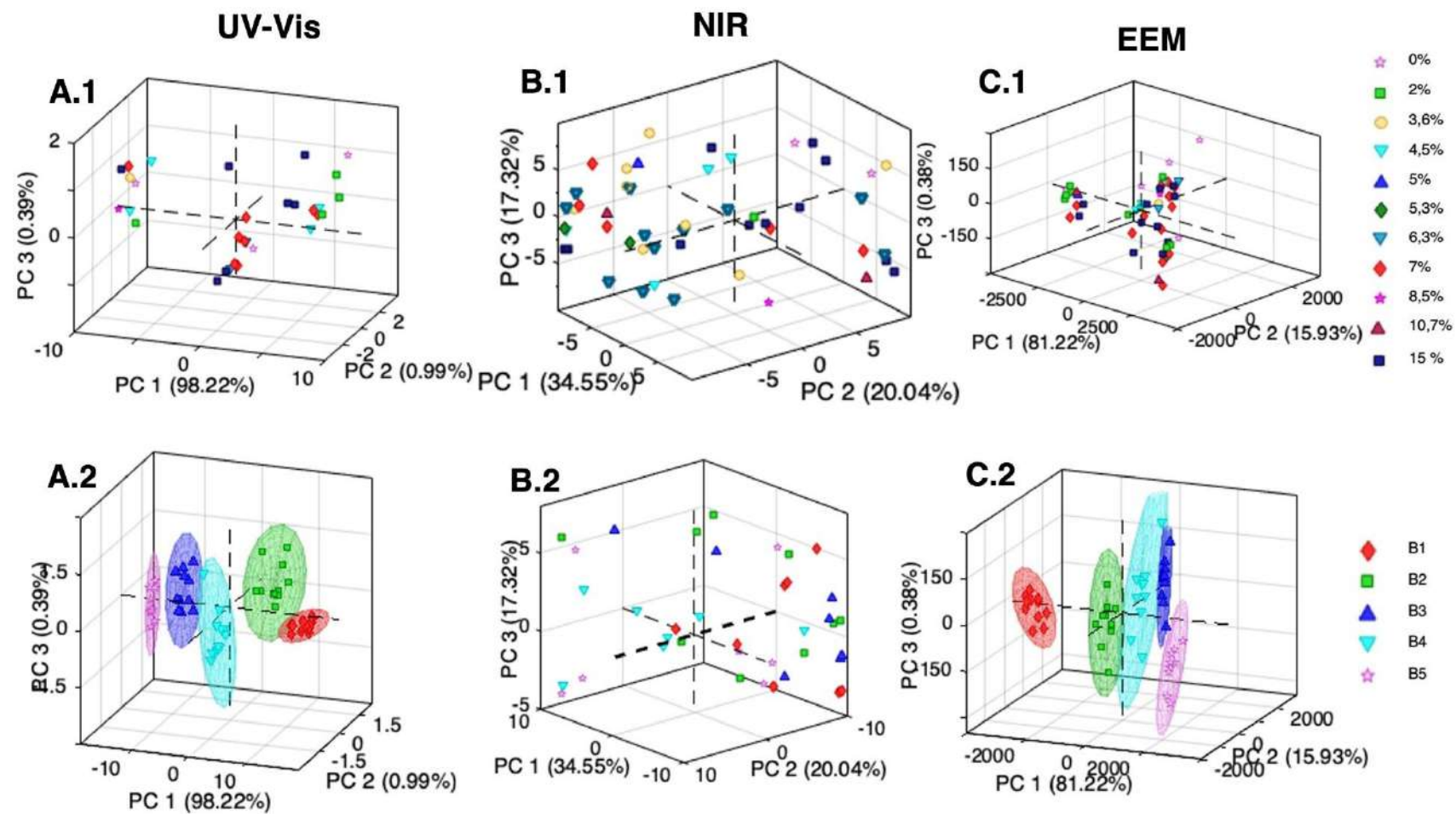
En primer lugar, con el objetivo de mejorar y corregir las señales instrumentales para el conjunto de datos UV-Vis y NIR, se aplicaron procedimientos de pre-procesamiento de datos. Para el primer set de datos se realizó MC y, para NIR, se implementó SNV.

La figura 4.4 muestra los *loadings* de los tres primeros PC obtenidos para cada señal instrumental. El análisis de los *loadings* demostró que en el rango de 320 nm-520 nm y 650 nm-670 nm para los espectros UV-Vis se encuentra la mayor contribución de la información espectral, mientras que, para NIR, todo el rango espectral aporta información relevante del sistema. Es importante mencionar que los *loadings* obtenidos de PCA para las EEM desdobladas no son de fácil interpretación, por lo que no es posible realizar una clara visualización de las variables que muestren la mejor diferenciación de las muestras.



**Figura 4.4.** Gráfico de *loadings* para los tres primeros PC obtenidos de PCA para los datos a) UV-Vis y b) NIR.

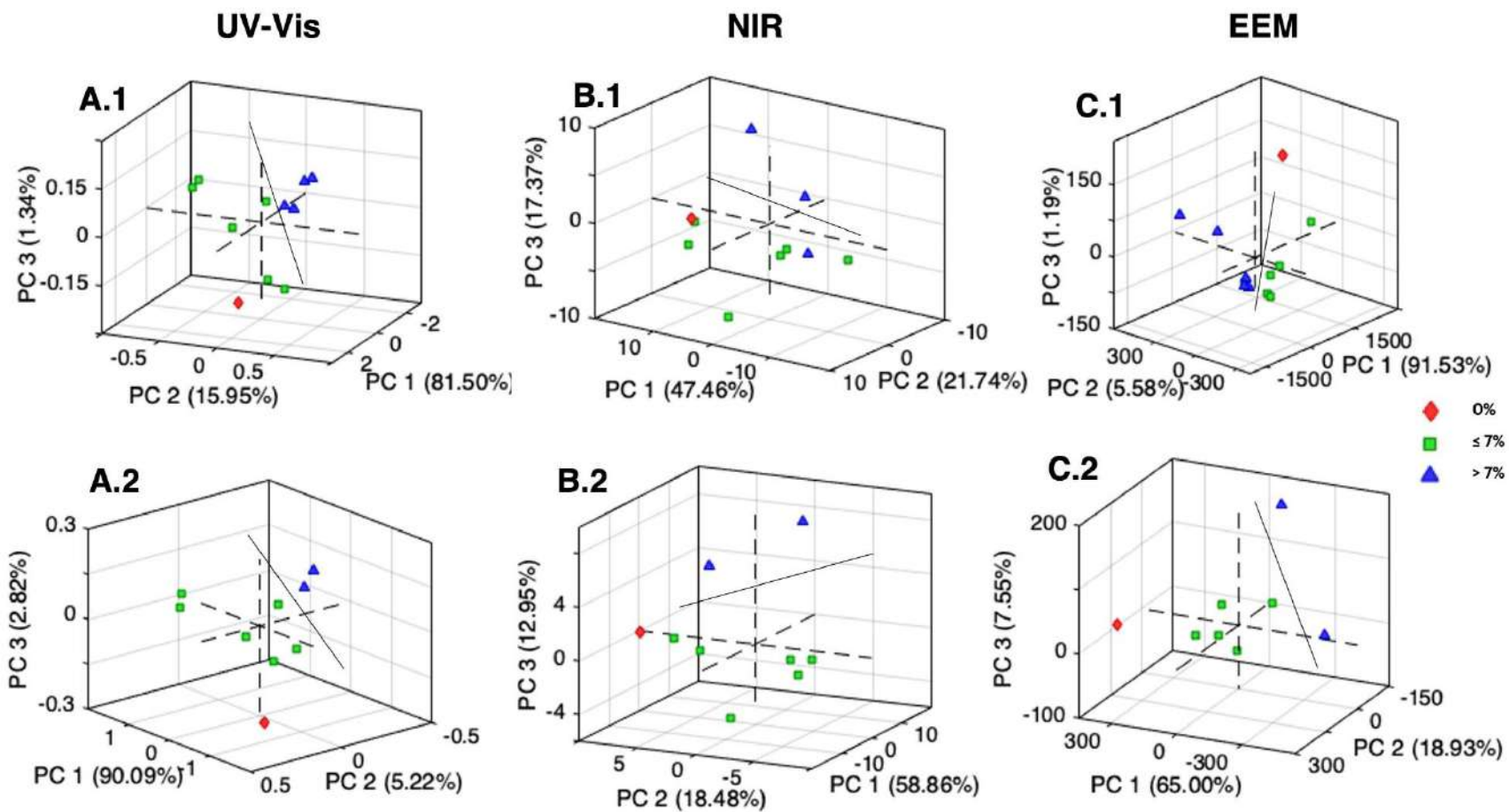
La figura 4.5 esquematiza el espacio construido con los scores de los primeros tres PC de cada modelo PCA para las diferentes técnicas instrumentales.



**Figura 4.5.** Gráfico de *scores* en el espacio tridimensional para los tres primeros PC de las diferentes matrices de datos instrumentales codificados según el % de adulteración (arriba) y según la marca de AOEV (abajo). Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional.

En los resultados obtenidos para UV-Vis se puede observar una clara agrupación de las muestras en cinco conjuntos, que coinciden con las diferentes marcas de los AOEV utilizadas evidenciado con la separación de las elipses de confianza construidos a un nivel del 95 % (figura 4.5 A.2, B.2 y C.2). Sin embargo, en cada agrupamiento, no se evidencia una separación con base en el % de adulteración o el tipo de adulterante utilizado (figura 4.5 A.1, B.1 y C.1). Curiosamente, los resultados obtenidos para los datos NIR (figura 4.5.B.1 y B.2) no muestran ninguna tendencia evidente a agruparse, por lo que no fue preciso, además, calcular las elipses de confianza para cada agrupación. Esta observación esclarece el hecho de que la técnica NIR, si bien genera datos reproducibles y robustos para el análisis de adulteración en AOEV a altos % de adulteración (Azizian et al., 2015; Yang Li, 2019), no muestra claras tendencias de agrupamiento en los niveles de adulteración evaluados en este trabajo. Finalmente, los resultados obtenidos del análisis de las EEM desdobladas están en concordancia con lo obtenido para UV-Vis, mostrando una buena separación entre las muestras de aceites según la marca. Asimismo, al igual que lo ocurrido para las otras técnicas, no se evidencian agrupamientos con base en el % de adulterante o el tipo de adulterante utilizado. Para todos los modelos se consideraron los tres primeros PCs, que alcanzaron una varianza explicada acumulada del 99,62%, el 99,89 % y el 98,25 % para UV-Vis, NIR y EEM, respectivamente.

Todas estas observaciones demuestran una alta variabilidad entre muestras que puede deberse, por un lado, a la composición natural de los componentes que son inherentes al tipo de aceituna utilizada en la producción de aceite y, por el otro, a los procesos de producción, propios de cada industria aceitera. En este sentido y para una mejor evaluación, se construyeron modelos PCA individuales para cada marca, con el fin de identificar agrupaciones con base en el % de adulterante. Los resultados obtenidos para dos muestras representativas se muestran en la figura 4.6 y las varianzas explicadas acumuladas de los modelos se resumen en la tabla 4.1.



**Figura 4.6.** Gráfico de *scores* en el espacio tridimensional para los tres primeros PC de los diferentes arreglos de datos para los AOEV B3 (A.1, B.1, C.1) y B5 (A.2, B.2, C.2). Las líneas de guiones indican el eje central del espacio tridimensional. La línea negra sólida identifica la separación imaginaria de las muestras en grupos por % de adulterante.

**Tabla 4.1.** Varianza explicada de los conjuntos de datos correspondientes a cada marca de AOEV

	Varianza explicada (%)				
	B1	B2	B3	B4	B5
UV-Vis	96,41	99,19	99,23	99,01	91,02
NIR	99,87	99,92	89,30	80,02	98,71
EEM	99,98	91,90	92,30	96,19	91,48

Si bien el número de muestras para este análisis no es lo suficientemente representativo del conjunto muestral, puede apreciarse una separación de los PC en al menos tres conjuntos de datos con base en el % de adulterante utilizado: muestra pura sin adulterar,  $\leq 7\%$  y  $> 7\%$ . Esta tendencia se observa en todas las marcas de AOEV analizada. En este sentido, el límite de separación entre clases se realizó trazando una línea imaginaria entre los grupos, considerando que la escasa cantidad de muestras para cada agrupamiento no hizo viable la confección de las elipses de confianza. Es importante resaltar que en todos los casos la muestra no adulterada mostró una clara separación de las adulteradas, siendo una primera evidencia de la posibilidad de la distinción binaria, incluso a bajos niveles de adulteración. Esta observación reviste de interés considerando que el sistema en estudio presenta una notable complejidad dada, principalmente, por la variabilidad demostrada entre las marcas evaluadas, entre la naturaleza de los aceites refinados utilizados como adulterantes y los bajos niveles de concentración evaluados.

#### 4.1.2.2 Estudios de predicción

Para determinar la cantidad de adulterante presente en las muestras de AOEV, se construyeron modelos predictivos mediante análisis PLS. Los modelos se construyeron con 35 muestras de calibración, seleccionadas aleatoriamente del total de muestras, y con 8 muestras de predicción. La capacidad predictiva de los modelos se evaluó a partir del análisis de los parámetros estadísticos, como los coeficientes de regresión ( $R^2$ ) de la calibración, los RMSEC, RMSECV, el error cuadrático medio de predicción (RMSEP) y el error porcentual de predicción (REP%) (Gurdeniz *et al.*, 2009). Además, se evaluó la exactitud de la metodología analítica desarrollada mediante un gráfico de concentración predicha con respecto a concentración nominal y la estimación de la región elíptica de confianza conjunta. Este último análisis consiste en graficar la región de confianza mutua de la pendiente y la

ordenada al origen para el gráfico de concentración predicha con respecto a concentración nominal, en el plano pendiente-ordenada al origen (Olivieri et al., 2014). Para este análisis se utilizaron las concentraciones predichas obtenidas en las muestras predicción, realizándose un ajuste de cuadrados mínimos ordinarios (OLS). Los resultados de precisión y los parámetros estadísticos de cada modelo se detallan en la tabla 4.2 y 4.3, respectivamente.

**Tabla 4.2.** Resultados de predicción obtenidos para cada modelo PLS construido.

adulterante nominal (%)	adulterante predicho (%)				
	UV-Vis	NIR	Fluorescencia	Fusion nivel bajo	Fusion nivel medio
2,0	3,9	1,7	2,5	3,5	1,8
4,5	5,3	2,8	4,7	4,0	4,4
7,0	5,4	8,7	7,7	6,3	8,7
7,0	8,3	0,9	9,1	7,5	6,2
7,0	10,8	6,2	5,5	8,5	9,0
15,0	11,3	17,0	12,7	11,6	14,7
15,0	12,0	9,0	16,4	14,7	15,7
15,0	13,0	10,5	15,5	14,7	14,4

**Tabla 4.3.** Parámetros estadísticos de los diferentes modelos PLS construidos.

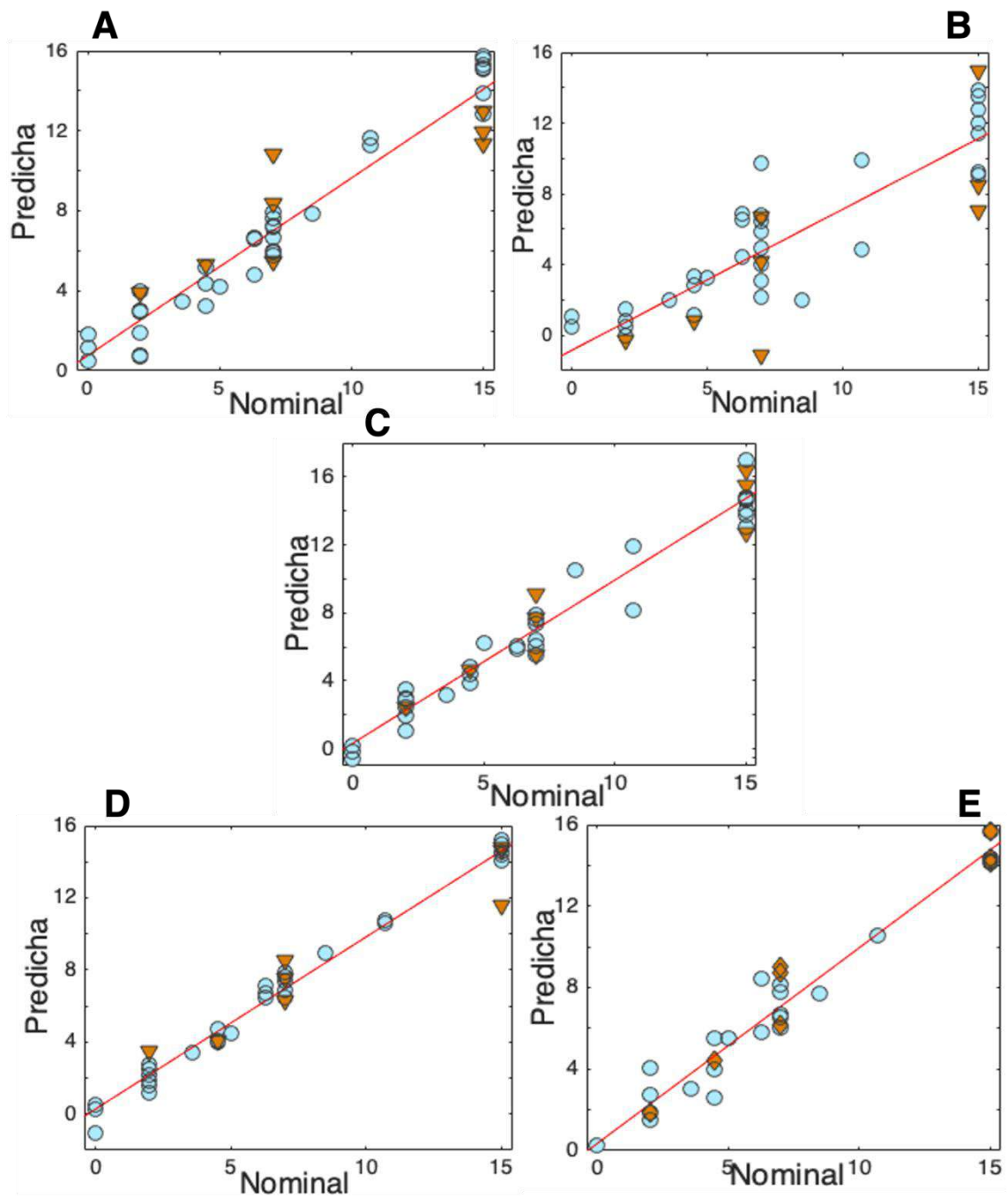
	Pre-tratamiento	Calibración		Validación			VL <sup>b</sup>
		R <sup>2</sup>	RMSEC	RMSECV	RMSEP	REP % <sup>a</sup>	
<b>UV-Vis</b>	SNV	0,95	1,0	2,6	2,5	36,2	8
<b>NIR</b>	MSC	0,81	2,0	3,2	3,6	52,2	9
<b>EEM</b>	EEM_cor	0,95	1,0	1,5	1,4	18,8	8
<b>Fusión nivel bajo</b>	VSS	0,98	0,5	1,5	1,0	14,5	5
<b>Fusión nivel medio</b>	AV, VSS	0,98	1,0	1,5	1,1	15,9	5

<sup>a</sup> REP%: error relativo porcentual calculado como  $REP\% = 100 \frac{RMSEP}{\langle y \rangle}$ , siendo  $\langle y \rangle$  la concentración promedio utilizadas en la calibración (Olivieri, 2018)

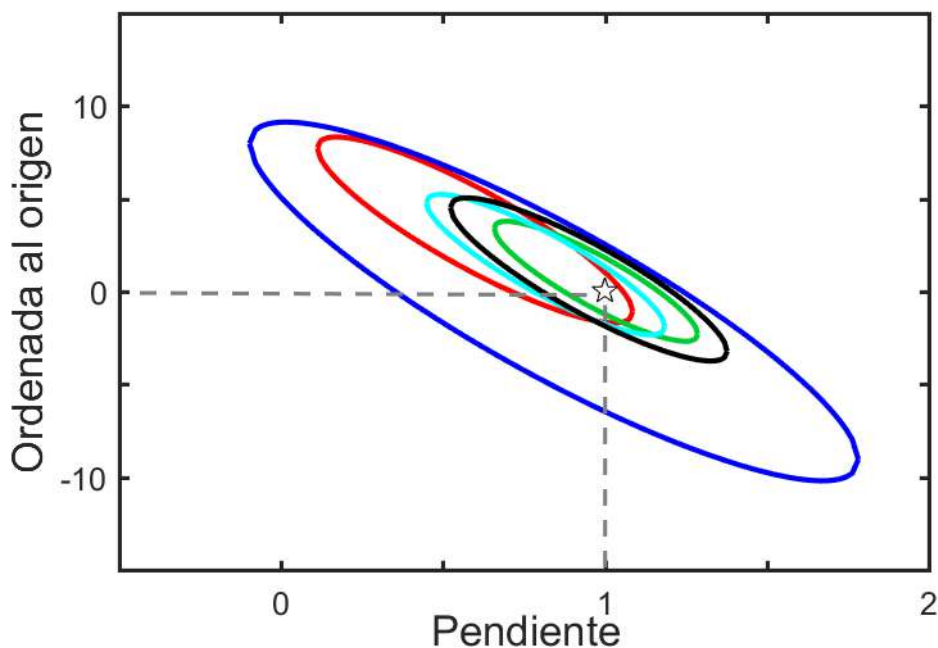
<sup>b</sup> VL: Variables latentes del modelo PLS

Los modelos de predicción construidos para los diferentes sets de datos se muestran en la figura 4.7 y las regiones elípticas de confianza conjunta en el plano pendiente-ordenada al origen, para un nivel de confianza del 95 %, se muestran en la figura 4.8.





**Figura 4.7.** Concentración de adulterante nominal vs. predicha, utilizando: **A)** UV-Vis, **B)** NIR, **C)** EEM desdobladas, **D)** fusión de nivel bajo (UV-vis -NIR-EEM desdobladas) y **E)** fusión de nivel medio.



**Figura 4.8.** Región elíptica de confianza conjunta en el plano pendiente-ordenada al origen, para un nivel de confianza del 95 %, indicándose el punto ideal [1, 0] (estrella negra) para pendiente y ordenada al origen, respectivamente, para los datos UV-Vis (rojo), NIR (azul), EEM (negro), fusión nivel bajo (verde) y fusión nivel medio (celestes).

Los resultados descritos en las tablas previas demuestran que el aumento de la multidimensionalidad de los datos lleva consigo una mejora sustancial en la eficiencia predictiva de los modelos respecto a los construidos con señales instrumentales individuales. Este aspecto se evidencia en las regiones elípticas de confianza conjunta que disminuyen su tamaño según aumenta la dimensionalidad del dato analizado (figura 4.8). Sin embargo, es notable que, si bien las elipses obtenidas para los datos espectrales de UV-Vis y NIR son las de mayor tamaño, para UV-Vis se observa una menor dispersión dada por la menor variabilidad en los resultados de predicción como consecuencia del aporte informativo que presenta cada técnica en el análisis de adulteración de AOEV. En este aspecto, la fluorescencia no convencional de EEM proporciona mayor información del sistema mediante el uso de un solo instrumento. Este aumento en la información química contenida en un mismo dato instrumental permite la obtención de modelos más sensibles que aprovechan esta información para lograr errores de predicción inferiores ( $REP\%=18,8\%$ ) a los obtenidos por las técnicas espectroscópicas convencionales de detección espectral ( $REP\%=36,2\%$  y  $52,2\%$  para UV-Vis y NIR, respectivamente). Esta observación valida, también, los resultados obtenidos para los modelos construidos con los datos fusionados de nivel bajo, en el que la combinación de la información instrumental, conlleva una mejora

sustancial de la eficiencia del modelo ( $REP\%=14,5\%$ ). Sin embargo, es importante mencionar que la fusión de datos de nivel medio no lleva a resultados superiores a los mencionados ( $REP\%=15,9\%$ ) para nivel bajo y EEM, fenómeno que se explica por la reducción de la dimensionalidad del dato a tratar. En este caso, si bien, el paso previo a la fusión se basó en la extracción de características significativas del sistema mediante PCA, esto conllevaría una posible pérdida de información relevante para el análisis posterior por PLS. Por último, es importante destacar que, como se puede observar en la figura 4.8, el punto ideal 1 y 0 para pendiente y ordenada al origen, respectivamente, se encuentran incluidos dentro de la región de confianza de cada elipse, para un nivel del 95 %, para todos los modelos construidos, demostrando la exactitud alcanzada.

De la figura 4.7, se destaca que el método de PLS construido con datos multidimensionales (EEM o fusionados) permiten distinguir las muestras no adulteradas con aquellas de muy bajo porcentaje de adulteración. Estos resultados están en concordancia con los obtenidos de los modelos PCA de la sección 4.1.2.1, lo que reafirma el hecho de que la diferenciación binaria entre muestras adulteradas/no adulteradas a bajos niveles de adulteración es posible con los métodos de análisis propuestos en esta tesis. En la figura 4.7 se puede observar, además, que las predicciones (CV y predicción) de las muestras presentan menor dispersión para las diferentes concentraciones frente a las obtenidos de los modelos UV-Vis y NIR, demostrando una vez más que el aumento de la dimensionalidad del dato conlleva la construcción de métodos más eficientes y confiables.

Si bien los resultados de eficiencia predictiva obtenidos en esta tesis resultan inferiores a los publicados en la bibliografía reciente (Jin Tan et al., 2018), es importante destacar que, hasta el momento, no se han reportado trabajos de investigación referidos al análisis de adulteración de AOEV con mezclas de aceites refinados, como girasol, maíz y soja, en rangos de niveles bajos de concentración (2 %-15 %). La mayoría de los trabajos publicados reportan rangos de concentraciones amplios, por ejemplo, entre 1 %-50 % (Jin Tan et al., 2018, Azizian et al., 2015, Borrás et al., 2016; Jiménez-Carvelo et al., 2019), que no revisten interés desde la perspectiva legal que establece la prohibición de agentes adulterantes en AOEV comercial (CAA, art 535, 1953).

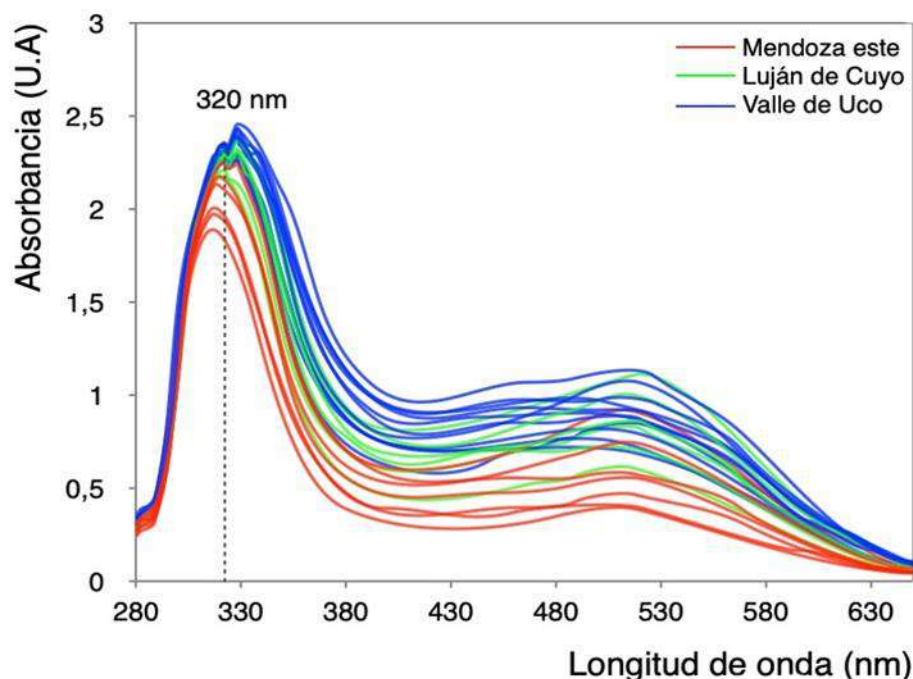
## 4.2 CLASIFICACIÓN DE VINOS MONOVARIETALES MALBEC

### 4.2.1 Consideraciones generales

El Malbec (*Vitis vinifera L.*), cepa de origen francés, llegó a mediados del siglo XIX a Argentina, donde se adaptó rápidamente a los diversos terruños y se desarrolló, incluso, mejor que en su región de origen, transformándose en la variedad emblemática de la producción vitivinícola nacional. Así, el Malbec argentino permite lograr excelentes vinos tintos que son apreciados y reconocidos por los consumidores a nivel nacional e internacional.

Entre los componentes fundamentales de los vinos Malbec, se encuentran las sustancias fenólicas que, dependiendo de su composición, determinan las características organolépticas del vino. Las antocianinas son compuestos fenólicos responsables del color rojo intenso, mientras que los taninos están relacionados al sabor y a las propiedades de gusto como amargor, aspereza, astringencia, entre otros. Estos compuestos son esenciales en la calidad y estilo de los vinos Malbec y pueden ser bien caracterizados mediante espectroscopia UV-Vis y NIR.

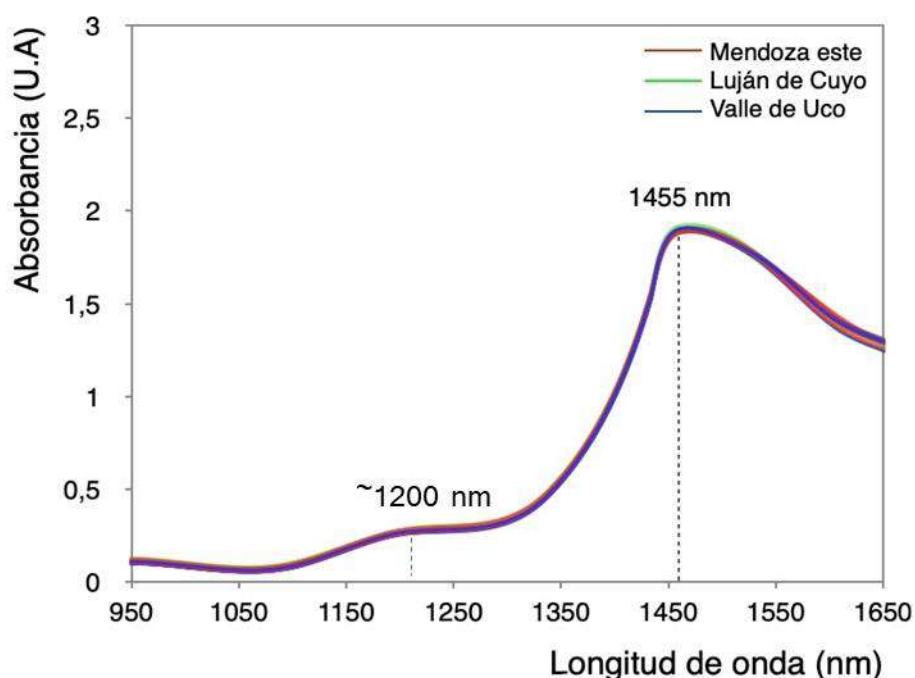
Los espectros de absorción UV-Vis de las muestras de vinos analizadas en este trabajo se muestran en la figura 4.9.



**Figura 4.9.** Espectros UV-Vis de las 24 muestras de vino Malbec procedentes de tres regiones geográficas diferentes.

Como puede apreciarse en la figura 4.9, los espectros UV-Vis muestran una prominente banda de absorción en el rango UV cercano (300 nm-400 nm) debida a la presencia de compuestos polifenólicos y aldehídos, entre otros (Cullere et al., 2019), que otorgan las características de aroma y sabor. Además, se observa una banda de absorción secundaria en el rango visible del espectro (> 400 nm) asociada a la presencia de antocianinas que, entre otros compuestos, otorgan las propiedades de color (Martelo-Vidal et al., 2014).

Los espectros NIR de los vinos Malbec analizados se muestran en la figura 4.10.



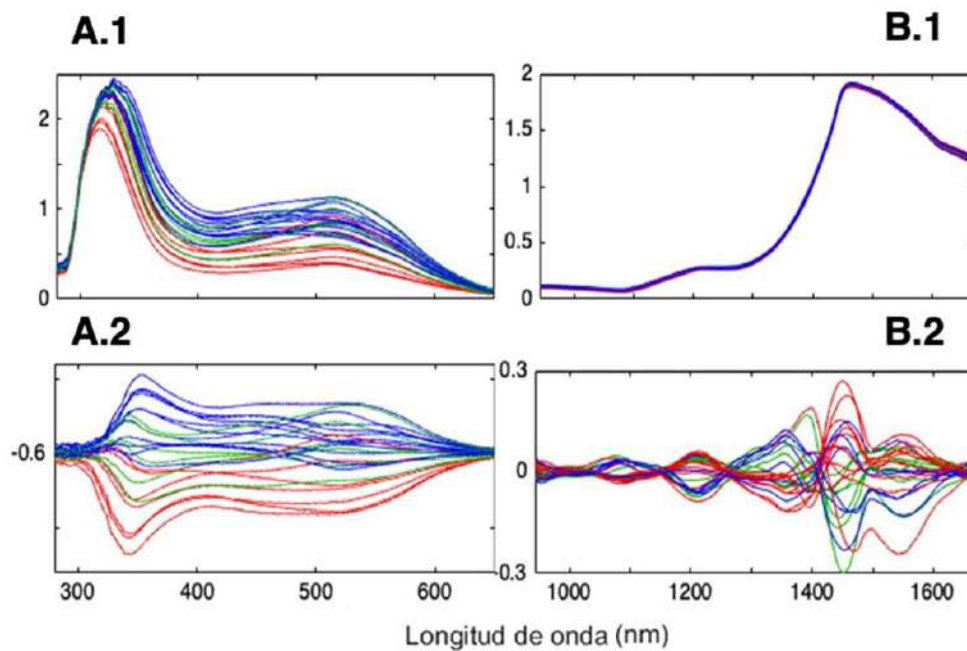
**Figura 4.10.** Espectros NIR de las 24 muestras de vino Malbec procedentes de tres regiones geográficas diferentes.

Como se puede observar, los espectros NIR muestran una banda a ~1200 nm correspondiente a los segundos sobretonos de los enlaces  $-CH$  asociados a los grupos aromáticos, como los taninos, y de los ácidos orgánicos y aldehídos. Además, se observa una banda de absorción bien definida a ~1455 nm (figura 4.9) que refleja la presencia de las primeras bandas armónicas atribuidas a las vibraciones  $-NH$  y  $-OH$  (Leqian Hu et al., 2018), de los aminoácidos libres y aminas biógenas.

## 4.2.2 Análisis quimiométrico de los datos

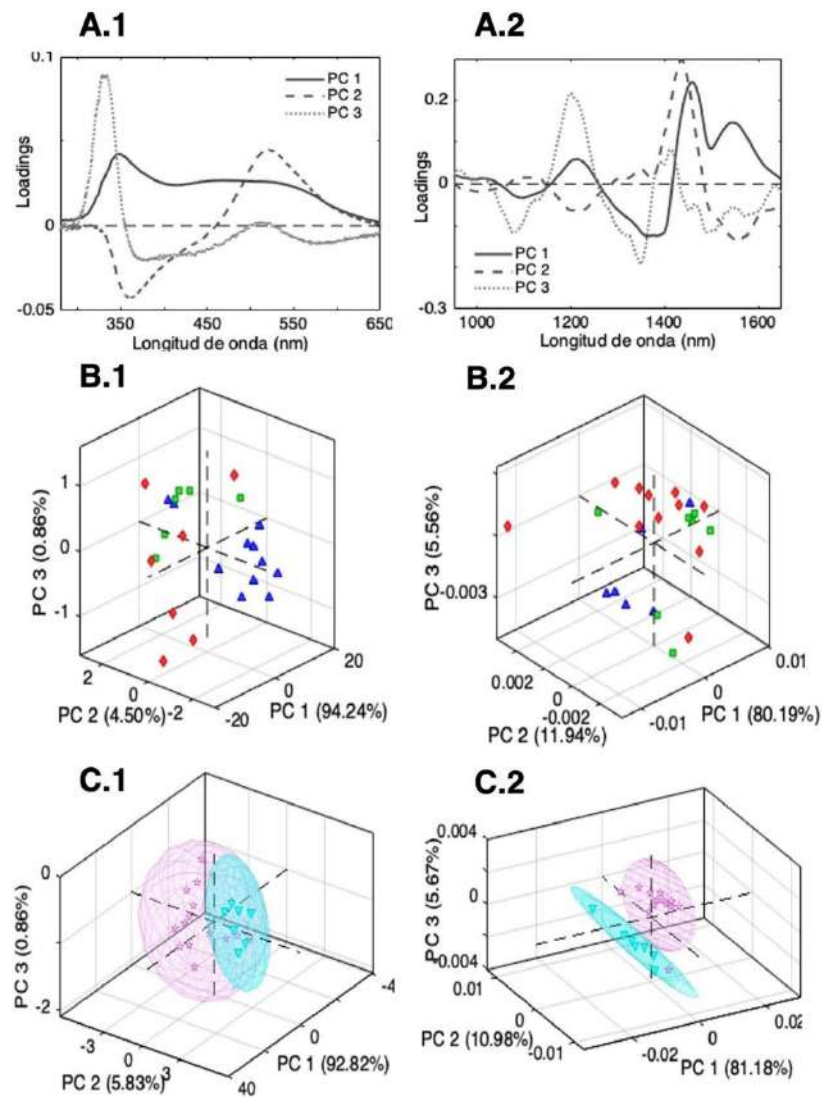
### 4.2.2.1 *Análisis de componentes principales (PCA)*

Previo a la construcción de los modelos quimiométricos, los datos se sometieron a diferentes pre-procesamientos, obteniéndose los mejores resultados mediante la aplicación de MC para datos UV-Vis y MC para los datos derivados de NIR (primera derivada) (figura 4.11).



**Figura 4.11.** Espectros UV-vis originales (A.1) y pretratados con MC (A.2) y espectros NIR originales (B.1) y derivados + MC (B.2).

Para cada conjunto de datos pre-procesados se construyó un modelo PCA a fin de evaluar agrupaciones y valores atípicos en la población muestral. Se analizaron los *loadings* obtenidos del modelo para evaluar el rango de longitudes de onda responsables de las mayores diferencias entre muestras. La figura 4.12 muestra los gráficos de *loadings* y *scores* para los tres primeros PC de ambos conjuntos de datos, indicando la diferenciación de acuerdo al origen geográfico y la bodega de producción.



**Figura 4.12.** A.1 y A.2 Gráficos de *loadings* obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente. B.1 y B.2 gráficos de *scores*, obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente, agrupados según su procedencia geográfica. C.1 y C.2 gráficos de *scores*, obtenidos para los tres primeros PC del conjunto de datos UV-Vis y NIR, respectivamente, codificados según la bodega de producción. Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional.

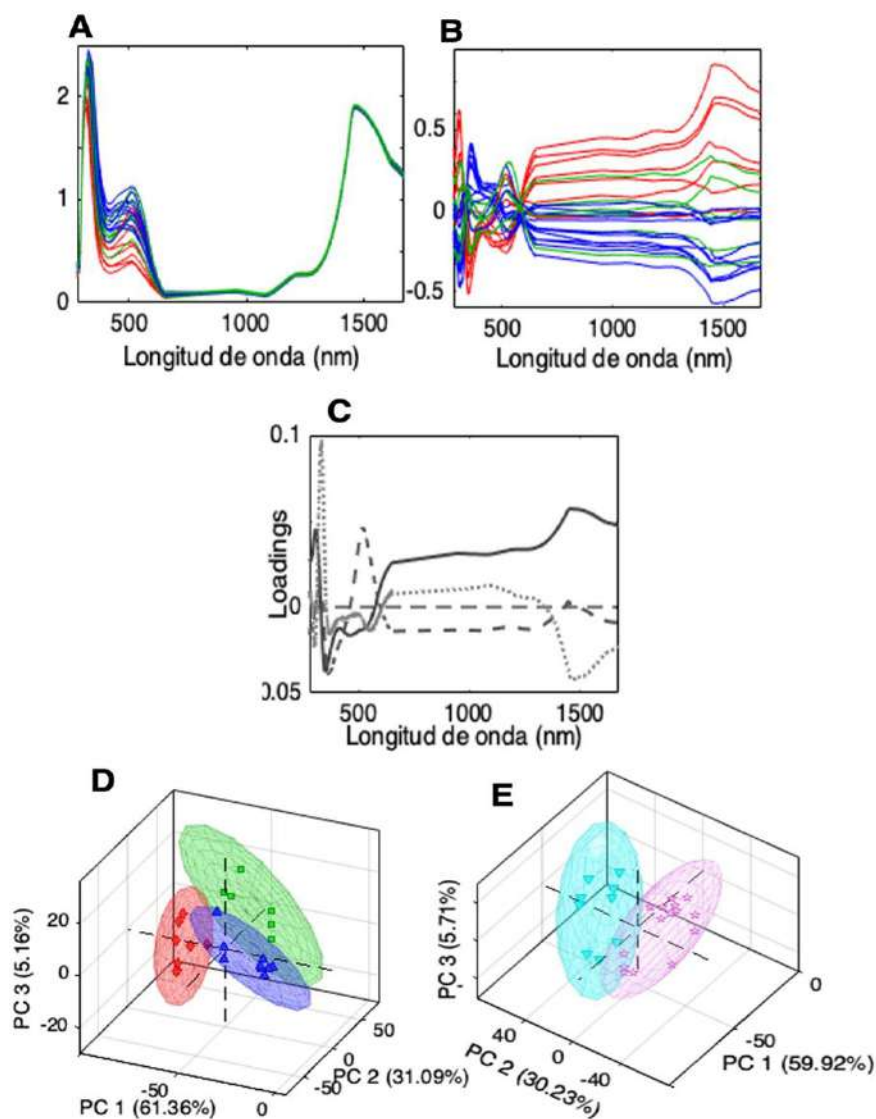
Los modelos PCA se construyeron con los tres primeros PC que explicaron el 99,6 % de la varianza total para el conjunto de datos UV-Vis y el 96,9 % para el conjunto de datos NIR. Como puede observarse en los gráficos de *loadings*, todo el rango espectral de UV-Vis y NIR presenta información relevante para la discriminación entre las muestras de vino. Estas diferencias entre muestras pueden ser causadas, en gran parte, por la variabilidad en la composición de compuestos polifenólicos del vino, tales como ácidos fenólicos, benzoicos

e hidroxicinámicos, que difieren según la composición natural de la uva, inherente a su terroir, y el proceso de elaboración del vino (Azcarate *et al.*, 2015).

Como puede observarse del gráfico de *scores*, no se distinguen agrupaciones evidentes con base en la procedencia geográfica de las muestras. Sin embargo, existe un agrupamiento natural de las muestras según la bodega productora del vino. Si bien se establecieron idénticos protocolos de fabricación para las diferentes bodegas, este análisis revela diferencias no esperadas en el proceso de elaboración de vino inherentes a cada bodega, como la molienda, el prensado, los recipientes utilizados para la fermentación y el control de temperatura durante el proceso, entre otros. Además, se realizó una prueba de hipótesis *t* para evaluar diferencias estadísticas entre ambas bodegas respecto al contenido alcohólico y al pH. Este análisis evidenció una diferencia estadísticamente significativa entre los valores medio de pH entre ambas bodegas ( $p_{\text{calculado,pH}}=0.007$   $p_{\text{calculado, alcohol}}= 0.133 < p_{\text{teórico}}=0.05$ ) lo que explicaría los resultados no esperados de diferenciación. Cabe destacar que el pH es uno de los factores químicos que mas afecta a los compuestos de la matriz, lo que llevaría consigo diferencias en la composición natural de los vinos (Ríos-Reina, 2020).

En pos de mejorar el modelo de clasificación, se aplicó la estrategia de fusión de datos de nivel bajo para el análisis conjunto de los conjuntos de datos instrumentales. Esta estrategia se ha implementado satisfactoriamente en recientes estudios de calidad agroalimentaria con rendimientos superiores a las estrategias clásicas de análisis de datos (Silvestri *et al.*, 2014; Teixeira dos Santos *et al.*, 2017, Ríos Reina *et al.*, 2020). Para este estudio, en primer lugar, se construyeron los arreglos de datos fusionados por concatenación vectorial. En este sentido, para cada muestra, se concatenó la señal correspondiente de ambas técnicas espectroscópicas, obteniéndose un espectro UV-Vis-NIR concatenado. El arreglo de datos final obtenido para todas las muestras se sometió luego a SNV y MC para construir el modelo PCA final La varianza explicada para los tres primeros PC fue de 95,9 %. La figura 4.13 muestra los gráficos de *loadings* y *scores* para los tres primeros PC de los datos fusionados.





**Figura 4.13.** Datos UV-Vis-NIR fusionados originales (A) y pre-procesados (B), *loadings* (C) y *scores* del conjunto de datos fusionados, clasificados según su procedencia geográfica (D) y bodega de producción (E). Las elipses muestran la región de confianza a un nivel de 95 %. Las líneas de guiones indican el eje central del espacio tridimensional.

A partir de la gráfica de *loadings* puede evidenciarse que ambas técnicas espectroscópicas brindan información relevante para la diferenciación entre muestras. Por otro lado, en el gráfico de *scores* se observa una clara agrupación de muestras según el origen geográfico, demostrando la superioridad en la eficiencia del análisis al usar datos fusionados. Asimismo, el uso de datos fusionados no altera la agrupación antes obtenida con base en la bodega productora.

Estos resultados brindan evidencias acerca de las bondades de las estrategias de análisis basadas en la combinación de información provista por diferentes técnicas

instrumentales, exaltando el efecto sinérgico de los datos fusionados para la extracción de información en sistemas complejos.

#### 4.2.2.2 Análisis de clasificación mediante PCA-DA

La discriminación de las muestras se realizó mediante un método supervisado, en este caso, PCA-DA, en el que se construyó un modelo de clasificación utilizando las muestras pre-categorizadas en categorías conocidas según la procedencia geográfica, que luego se utilizará para asignar muestras de procedencia desconocida en las categorías previamente asignadas. El modelo PCA-DA se basa en encontrar la regla de asignación que brinde la mejor eficiencia de clasificación. (Cocchi *et al.*, 2018)

Para cada conjunto de datos instrumentales se construyó un modelo discriminante lineal usando los *scores* obtenidos de los modelos PCA para cada conjunto de datos individual. Paralelamente, se construyó un modelo PCA-DA para datos fusionados de nivel bajo utilizando vectores UV-Vis-NIR concatenados para cada muestra analizada. Las muestras se categorizaron de acuerdo al origen geográfico como: ME (Mendoza Este), LC (Luján de Cuyo) y VU (Valle de Uco). El modelo se construyó utilizando 19 muestras en el conjunto de calibración, seleccionadas aleatoriamente del total, y se validó con las 5 muestras restantes. La tabla 4.4 muestra las matrices de confusión de los modelos PCA-DA para los conjuntos de datos individuales y fusionados, indicando el porcentaje de muestras correctamente clasificadas (NER%) para cada conjunto de datos. La matriz de confusión se arma incluyendo información de las clases nominales y predichas dispuestas en filas y columnas, respectivamente. Los elementos diagonales de la matriz contienen el número de muestras clasificadas correctamente, mientras que los elementos fuera de la diagonal incluyen el número de muestras mal clasificadas (Ballabio *et al.*, 2018).

Los índices más útiles utilizados para analizar los resultados de la clasificación son la sensibilidad (S), que describe la capacidad del modelo para reconocer correctamente muestras pertenecientes a una clase, y la especificidad (SP), que caracteriza la capacidad de una clase para rechazar las muestras de todas las demás. Además, la precisión (PR) es una cifra que permite evaluar la capacidad del modelo para evitar predicciones erróneas dentro de una clase (Ballabio *et al.*, 2018). Estos índices tienen valores entre 0 y 1 indicando la no clasificación y clasificación de clase perfecta, respectivamente. La tabla 4.5 resume los índices de clasificación obtenidos para cada modelo construido.

**Tabla 4.4.** Matrices de confusión de los modelos PCA-DA.

		<i>Conjunto de calibración</i>				<i>Conjunto de predicción</i>			
		VU	ME	LC	NER <sup>1</sup> (%)	VU	ME	LC	NER <sup>1</sup> (%)
<i>UV-Vis</i>	VU	7	0	2	77,8	2	0	0	100,0
	ME	0	5	1	83,3	0	1	0	100,0
	LC	1	0	3	75,0	1	1	0	0,0
<i>NIR</i>	VU	8	1	0	88,9	1	0	1	50,0
	ME	0	6	0	100,0	1	0	0	0,0
	LC	2	1	1	25,0	0	1	1	50,0
<i>UV-Vis-NIR</i>	VU	8	0	0	100,0	3	0	0	100,0
	ME	0	5	1	83,3	0	0	1	0,0
	LC	0	0	4	100,0	0	0	2	100,0

<sup>1</sup>NER %: porcentaje de muestras correctamente clasificadas  $NER\% = 100 \frac{n^{\circ} \text{muestras bien clasificadas}}{n^{\circ} \text{muestras totales}}$   
(Olivieri, 2016)

**Tabla 4.5.** Índices de clasificación de los conjuntos de calibración y predicción para los diferentes arreglos de datos analizados.

		<i>Conjunto de calibración</i>			<i>Conjunto de predicción</i>		
		S <sup>1</sup>	SP <sup>2</sup>	PR <sup>3</sup>	S <sup>1</sup>	SP <sup>2</sup>	PR <sup>3</sup>
<i>UV-Vis</i>	VU	0,78	0,90	0,88	1,00	0,67	0,67
	ME	0,83	1,00	1,00	1,00	0,75	0,50
	LC	0,75	0,80	0,50	0,00	1,00	0,00
<i>NIR</i>	VU	0,89	0,80	0,80	0,50	0,67	0,50
	ME	1,00	0,85	0,75	0,00	0,75	0,00
	LC	0,25	1,00	1,00	0,50	0,67	0,50
<i>UV-Vis-NIR</i>	VU	1,00	1,00	1,00	1,00	1,00	1,00
	ME	0,83	1,00	1,00	0,00	1,00	0,00
	LC	1,00	0,93	0,83	1,00	0,75	0,50

<sup>1</sup> S: Sensibilidad

<sup>2</sup> SP: Especificidad

<sup>3</sup> PR: Precisión

De un análisis global de los resultados mostrados en las tablas previas se puede concluir que el modelo construido con los datos fusionados es superior en eficiencia de

clasificación respecto a los modelos construidos con los conjuntos de datos individuales. Se puede observar que los valores de S, SP y PR son cercanos al valor de clasificación perfecta. No obstante, cabe destacar que el rendimiento de los modelos es superior en los conjuntos de calibración, obteniéndose valores pobres de clasificación para los conjuntos de validación. Esta observación refleja el hecho de que el número de muestras utilizado en el conjunto de validación no es lo suficientemente representativo para corroborar el poder clasificatorio de los modelos, por lo que se sugiere incrementar el número de muestras para análisis posteriores.

Todos estos resultados están en concordancia con lo publicado en la bibliografía, demostrando que la utilización de espectroscopia NIR para la clasificación de vinos según el origen geográfico arroja peores resultados, comparado con otras técnicas espectroscópicas (*Zhen Hu et al., 2019*). Asimismo, se demuestra una vez más que el uso de fusión de datos para análisis de clasificación conlleva una mejora en la capacidad de clasificación de los modelos. En este aspecto, se evidencia además que, si bien la información NIR no es lo suficientemente sensible para el análisis de clasificación de vinos según su procedencia geográfica, en combinación con otra información espectral, mejora los modelos clasificatorios (*Martelo-Vidal et al., 2012, Uricková et al, 2015, Leqian Hu et al., 2018*).

5

Conclusión

---

## 5 CONCLUSIÓN

Se desarrolló una metodología analítica basada en mediciones de absorbancia UV-Vis NIR y fluorescencia multidimensional de matrices de excitación-emisión acoplada a análisis quimiométrico para el estudio de adulteración de aceites de oliva extra virgen. Se evaluó la presencia de adulterantes de distinta naturaleza (aceites refinados de girasol, maíz y soja) en bajos niveles de concentración. Se realizó un primer análisis exploratorio de PCA con datos instrumentales individuales y fusionados que reveló tendencias de agrupamiento de acuerdo al fabricante (o marca). Posteriormente, sobre la base de estos resultados, se realizó un segundo análisis exploratorio para evaluar diferenciaciones entre niveles de adulterante por cada grupo previamente definido (marcas). Este análisis demostró la posibilidad de distinguir tres grupos de mezclas separadas según rangos de % de adulteración, destacándose la capacidad de diferenciar la muestra no adulterada del grupo. Posteriormente, se construyeron cinco modelos de calibración PLS con los datos recogidos por cada técnica espectroscópica individual, con datos fusionados de nivel bajo y datos fusionados de nivel medio. Los resultados obtenidos demostraron que el aumento en la combinación de información de los datos lleva consigo una mejora sustancial en la eficiencia predictiva de los modelos respecto a los construidos con señales instrumentales individuales, como así también, reafirmaron el hecho de poder distinguir las muestras no adulteradas de aquellas con muy bajo % de adulteración, superando la complejidad dada por la amplia variabilidad del sistema.

Además, se desarrolló una metodología analítica basada en mediciones de absorbancia UV-Vis y NIR acoplada a análisis quimiométrico para el estudio de clasificación de vinos Malbec según su procedencia geográfica. Para demostrar el potencial de las técnicas espectroscópicas utilizadas, se construyeron diferentes modelos quimiométricos utilizando distintas estrategias de análisis y resolución de datos. Los modelos se construyeron con datos instrumentales individuales y fusionados de nivel bajo. Del análisis exploratorio por PCA de los datos individuales se obtuvieron resultados que demostraron un agrupamiento natural de las muestras según la bodega productora del vino. Sin embargo, el análisis de clasificación mediante PCA-DA con datos fusionados arrojó resultados que permitieron, además, la clasificación de vinos según su procedencia geográfica, con la tasa de error más baja, cumplimentando, así, el objetivo original del estudio. Los resultados reflejaron la superioridad en eficiencia de clasificación de los modelos construidos con datos fusionados.

Con base en lo expuesto, se puede concluir que los métodos propuestos resultan altamente adecuados para el control de calidad de muestras agroalimentarias de una forma simple, rápida y eficiente, sin requerir instrumentación sofisticada y/o de alto costo, evitando largas y tediosas etapas de tratamiento de muestras. Además, y en concordancia con lo reportado en la bibliografía consultada, los métodos propuestos resultarían idóneos para el estudio de autenticación y adulteración de vinos y aceites de oliva extra virgen, respectivamente.

Por todo lo mencionado, se puede concluir que la quimiometría es una herramienta poderosa que permite, entre otras cosas, la obtención de información relevante acerca de un sistema de estudio, la reducción de tiempos de análisis y la disminución del consumo de reactivos, aunque no puede prescindir del criterio analítico y de la interpretación del analista para lograr el éxito en el análisis.

Por último, pero no menos importante, se debe remarcar que este trabajo representa un valioso aporte al estudio integral de muestras agroalimentarias de consumo masivo en nuestro país y de querer explorar mayores bondades de los sistemas en estudio es requisito incrementar el tamaño muestral a fin de obtener modelos más robustos y resultados representativos.

6

# Referencias



---

## 6 REFERENCIAS

- Alcaraz M.R. (2016). *Nuevas estrategias analíticas para el análisis de muestras de composición compleja basadas en generación y modelado de datos multidimensionales*. Tesis doctoral.
- Anderson C.M., Bro R. (2010). *Variable selection in regression*. Journal of Chemometrics 24, 728-737.
- Aparicio R., Morales M.T., Aparicio-Ruiz R., Tena N., Garcia-Gonzalez D.L. (2013). *Authenticity of olive oil: Mapping and comparing official methods and promising alternatives*. Food research International 54, 2025-2038.
- Azcarate S.M, Araujo Gomes A., Alcaraz M.R, Ugulino de Araujo M.C, Camiña J.M., Goicoechea H.C (2015). *Modeling excitation-emission fluorescence matrices with pattern recognition algorithms for classification of Argentine white wines according grape variety*. Food chemistry, 184, 214-219.
- Azcarate S.M. (2016). *Estudio clasificatorio en muestras de interés alimenticio basado en el análisis multivariado de datos*. Tesis doctoral.
- Azizian H, Mossoba M, Fardin-Kia A, Delmonte P, Karunathilaka S, Kramer J (2015). *Novel, Rapid identification, and quantification of adulterants in extra virgin olive oil using near-infrared spectroscopy and chemometrics*. Lipids, 50, 705-718.
- Ballabio D, Grisoni F., Todeschini R. (2018). *Multivariate comparison of classification performance measures*. Chemometrics and Intelligent Laboratory Systems, 174 33-44.
- Barnes R.J., Dhanoa S., Lister S.J. (1989). *Standard Normal Variate transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra*. Applied Spectroscopy, 43, 772-777.
- Biancolillo A., Bucci R., Marini F. (2019). *Data-fusion strategies in food analysis*. Data Handling Science Technology 31, 271-310.
- Blanco, M., Villarroya, I., 2002. *NIR spectroscopy: A rapid-response analytical tool*. TrAC - Trends Analytical. Chemistry. 21, 240–250.
- Brereton, R. G. (2000) *Introduction to Multivariate Calibration in Analytical Chemistry*. Analyst 125, 2125–2154.

- Borrás E., Ferre J., Boque R., Mestres M., Acena L., Bustos O. (2015). *Data fusion methodologies for food and beverage authentication and quality assesment*. Analytica Chimica Acta 891, 1-14.
- Borrás E., Ferré J., Boqué R., Mestres M., Aceña L., Calvo A., Busto O. (2016). *Prediction of olive oil sensory descriptors using instrumental data fusion and partial least squares (PLS) regression*. Talanta, 155, 116-123.
- CAA (1953). *Capítulo VII: alimentos grasos y aceites alimenticios*. <https://www.argentina.gob.ar/sites/default/files/anmat-capitulo-vii-grasosactualiz-2018-12.pdf>. Último acceso, 10 de Abril 2021.
- Caballero D., Rios-Reina R., Amigo J.M. (2020). *Chemometrics and food traceability*. Comprehensive Foodomics, 387-406.
- Callao M.P., Ruisanchez I. (2018). *An overview of multivariate qualitative methods for food fraud detection*. Food Control 86, 283-293.
- Caramia G., Gori A., Valli E., Cerretani L. (2012). *Virgin olive oil in preventive medicine: From legend to epigenetics*. European Journal of Lipid Science and Technology 114, 375-388.
- Chen H., Tan C., Lin Z., Wu T. (2017). *Detection of melamine adulteration in mil by near-infrared spectroscopy and one-class partial least squares*. Spectrochimical Acta part A: Molecular and Biomolecular Spectroscopy, 173, 832-836.
- Chiappini F., Alcaraz M.R., Goicoechea H.C, Olivieri A.C. (2019). *A graphical user interface as a new tool for scattering correction in fluorescence data*. Chemometrics and intelligent laboratory systems, 193, 103810.
- Cocchi M., Biancolillo A., Marini F. (2018). *Chemometric Methods for Classification and Feature Selection*. Comprehensive Analytical Chemistry 82 265–299.
- Cordella C., Moussa I., Martel A.C., Sbirrazzouli N., Lizzani-Cuvelier L. (2002). *Recent developments in food characterization and adulteration detection: technique-oriented perspectives*. Journal of agricultural and food chemistry 50, 1751-1764.
- Culleré L., Lopez R., Ferreira V. (2019). *The instrumental analysis of aroma-active compounds for explaining the flavor of red wines*. Red Wine Technology. Chapter 20, 283-307.
- Danezis G.P., Tsagkaris A.S., Brusica V., Georgiou C.A. (2016). *Food authentication: state of the art and prospects*. Food Science 10, 22-31.

- Dankowska A., Domagala A., Kowalewski W. (2017). *Quantification of coffea arabica and coffea canephora var. robusta concentration in blends by means of synchronous fluorescence and UV-Vis spectroscopies*. *Talanta*, 172, 215-220.
- Doeswijk T.G., Smilde A.K., Hageman J.A., Westerhuis J.A., Eeuwijk van, F.A. (2011). *On the increase of predictive performance with high-level data fusion*. *Analytical Chemistry Acta* 705, 41-47.
- Esbensen K.H., Geladi P. (2009). *Comprehensive chemometrics: chemical and biochemical data analysis*. Vol. 2.
- Esteki M., Simal-Gandara J., Shahsavari Z., Zandbaaf S., Dashtaki E., Heyden Y.V. (2018). *A review on the application of chromatographic methods, coupled to chemometrics, for food authentication*. *Food control* 93, 165-182.
- Geladi P., MacDougall D., Martens H. (1985). *Linearization and Scatter-Correction for Near-Infrared reflectance Spectra of Meat*. *Applied Spectroscopy*, 39, 491-500.
- Godinho M.S., Blanco M.R., Gambarra Neto F.F, Liao L.M., Sena M.M., Tauler R., Oliveira A.E. (2015). *Evaluation of transformer insulating oil quality using NIR, fluorescence, and NMR spectroscopic data fusion*. *Talanta*, 129, 143-149.
- Gurdeniz, G., & Ozen, B. (2009). *Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data*. *Food Chemistry*, 116, 519–525.
- Guzman, E., Baeten, V., Pierna, J. A. F., & Garcia-Mesa, J. A. (2015). *Evaluation of the overall quality of olive oil using fluorescence spectroscopy*. *Food Chemistry*, 173, 927-934.
- Jian Tan, Li R., Jiang Z.T., Shi M., Xiao Y.Q., Jia B., Tian-Xiang L. Wang H. (2018). *Detection of Extra Virgin Olive Oil Adulteration with edible oils using front-face fluorescence and visible spectroscopies*. *Journal of the American Oil chemist Society* 95, 535-546.
- Jimenez-Carvelo A.M, Lozano V., Olivieri A. (2019). *Comparative chemometrics analysis of fluorescence and near infrared spectroscopies for authenticity confirmation and geographical origin of Argentinean extra virgin olive oils*. *Food control*, 96, 22-28.
- Kelsey R., Dean M., Haughey S., Elliott C. (2021). *A comprehensive review of food fraud terminologies and food fraud mitigation guides*. *Food control*, 120, 107516.
- Kennard R., Stone L.A (1969). *Computer-aided design of experiments*. *Technometrics*, 11, 137-148.

- Kharbach M., Marmouzi I., El Jemli M., Bouklouze A., Vander Heyden Y. (2020). *Recent advances in untargeted and targeted approaches applied in herbal-extracts and essential-oils fingerprinting- A review*. Journal of pharmaceutical and biomedical Analysis, 177, 112849.
- Khaleghi B., Khamis A., Karray F.O., Razavi S.N. (2013). *Multisensor data fusion: A review of the state-of-the-art*. Information Fusion 14, 28-44.
- Kongbonga, Y. G. M., Ghalila, H., Onana, M. B., Majdi, Y., Lakhdar, Z. B., Mezlini, H., & Sevestre-Ghalila, S. (2011). *Characterization of Vegetable Oils by Fluorescence Spectroscopy*. Food and Nutrition Sciences, 2, 692-699
- Kumar K., Tarai M., Mishra A.K. (2017). *Unconventional steady-state fluorescence spectroscopy as an analytical technique for analyses of complex-multifluorophoric mixtures*. Trends in analytical Chemistry 97, 216-243.
- Lakhil L., Acha V., Aussenac T. (2014). *Resolution of fluorophore mixtures in biological media using fluorescence spectroscopy and Monte Carlo Simulation*. Applied Spectroscopy 68, 697-711.
- Lakowicz J.R. (2006). *Principles of fluorescence spectroscopy*. Springer Science, New York, 158-204.
- Lakshmi V. (2012). *Food adulteration*. International journal of science 1, 106-113.
- Lavine B. K. (2006), *Pattern recognition, Critical Review*. Analytical Chemistry 36, 153–161.
- Leqian H., Yin C., Ma S., Liu Z. (2018). *Rapid detection of three quality parameters and classification of wine based on Vis-NIR spectroscopy with wavelength selection by ACO and CARS algorithms*. Spectrochimica Acta 205, 574-581.
- Li Vigni M., Durante C., Cocchi M. (2013). *Exploratory data analysis*. Data Handling in Science and Technology 28, 55-126.
- Lohumi S., Lee S., Lee H., Kwan Cho B. (2015). *A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration*. Trends in Food Science & Technology 46, 85-98.
- Massart D.L., Vandeginste B.G.M., Buydens L.C.M., de Jong S., Lewi P.J., Smeyers-Verbeke J. (1998). *Handbook of chemometrics and qualimetrics: Part A. Data Handle*. Science Technology, 20.
- Martelo-Vidal M.J, Vazquez M., (2012). *Ultraviolet/visible/near-infrared spectral analysis and chemometric tolos for the discrimination of wines between subzones*

*inside a controlled designation of origin: a case study of Rias Baixas*. Australian Journal of grape and wine research, 19, 62-67.

- Martelo-Vidal M.J, Vazquez M., (2014). *Determination of polyphenolic compounds of red wines by UV-Vis-NIR spectroscopy and chemometrics tools*. Food chemistry, 158, 28-34.
- Medina S., Perestrelo R., Silva P., Pereira J., Camara J.S. (2019). *Currents trends and recent advances on food authenticity technologies and chemometric approaches*. Food Science & Technology, 85, 163-176.
- Ministerio de Relaciones Exteriores, Comercio Internacional y Culto (2020). *Relevamiento de oportunidades*. <https://www.cancilleria.gob.ar/userfiles/ut/mercados-agroalimentarios-relevamiento-de-oportunidades.pdf>. Último acceso, 25 de Abril 2021.
- Moncayo L.P. (2016). *Análisis de la calidad del aceite de oliva virgen: relación entre la estabilidad oxidativa y la composición fenólica*. Tesis de grado.
- Naes T., Isaksson T., Fearn T., Davies T. (2002). *User-friendly guide to multivariate calibration and classification* (1st ed.). Chichester, U.K.: NIR Publications.
- Ohtsubo, K., Suzuki, K., Haraguchi, K., Nakamura, S., (2008). *Novel method for preparation of the template DNA and selection of primers to differentiate the material rice cultivars of rice wine by PCR*. Journal of Biochemical Biophysical Methods 70, 1020–1028.
- Olivieri A.C. (2014). *Analytical Figures of Merit: from univariate to multiway calibration*. Chemical reviews, 114, 5358-5378.
- Osborne B.G., Fearn T., Hindle P.T. (1993). *Practical NIR spectroscopy with applications in food and beverage analysis*. Longman Scientific and technical.
- Otto M. (2007). *Chemometrics, Statistics and Computer Application in Analytical Chemistry* (1st ed.). Weinheim: Wiley-VCH.
- Ozen B., Uncu O. (2019). *A comparative study of mid-infrared, UV-Visible and fluorescence spectroscopy in combination with chemometrics for the detection of adulteration of fresh olive oils with old olive oils*. Food Control, 105, 209-218.
- Palade L., Duta D., Butu A., Kefalas P., Popa E. (2014). *Characterization of the volatile profile of Romanian Cabernet Sauvignon wines through GC-MS and electronic nose detection*. 3rd international iseki food conference.
- Piovasana S., Cavaliere C., Cerrato A., Montone C., Lagana A., Capriotti A.L. (2020). *Developments and pitfalls in the characterization of phenolic compounds in food:*

*from targeted analysis to metabolomics-based approaches*. TrAC. Trends in Analytical Chemistry, 133, 116083.

- Racz A., Bajusz D., Fodor M., Heberger K. (2016) *Comparison of classification methods with “n-class” receiver operating characteristic curves: A case study of energy drink*. Chemometrics and Intelligent Laboratory Systems, 151, 34-43.
- Rebolo S., Peña R.M., Latorre M.J., García S, Botana A.M., Herrero C. (2000). *Characterization of Galician (NW Spain) Ribeira Sacra wines using pattern recognition analysis*. Analytical Chemistry Acta, 417, 211-220.
- Rinnan A., Berg F.V.D., Engelsen S.B. (2009). *Review of the most common preprocessing techniques for near-infrared spectra*. Trends in Analytical Chemistry 28, 1201-1222.
- Rios Reina R., Azcarate S.M., Camiña J.M., Callejón R.M. (2020). *Assessment of UV-visible spectroscopy as a useful tool for determining grape-must caramel in high-quality wine and balsamic vinegars*. Food Chemistry, 323, 126792.
- Rios Reina R., Ocaña J.A., Azcarate S.M., Perez-Bernal J.L., Villar-Navarro M., Callejón R.M. (2019). *Excitation-emission fluorescence as a tool to assess the presence of grape-must caramel in PDO wine vinegars*. Food Chemistry, 287, 115-125.
- Rodriguez-Saona L.E., Allendorf M.E. (2011). *Use of FTIR for rapid authentication and detection of adulteration of food*. Annual Review of Food Science and Technology 2, 467-483.
- Rodríguez S. D., Gagneten M., Farroni A. E., Percibaldi N. M., Buera P. (2019). *FT-IR and untargeted chemometric analysis for adulterant detection in chia and sesame oils*. Food Control 105, 78-85.
- Savitzky A., Golay M.J (1964). *Smoothing and differentiation of data by simplified Least Squares Procedures*. Analytical Chemistry, 36, 1627-1639.
- Silvestri M., Elia A., Bertelli D., Salvatore E., Durante C., Li Vigni M., Marchetti A., Cocchi M. (2014). *A mid level data fusion strategy for the varietal classification of Lambrusco PDO wines*. Chemometrics and Intelligent Laboratory Systems, 137, 181-189.
- Spiteri M. (2016). *Data fusion between high resolution 1 H-NMR and mass spectrometry: a synergetic approach to honey botanical origin characterization*. Bioanalytical Chemistry. 408, 1-13.

- Stuart B.H. (2004). *Infrared Spectroscopy: Fundamentals and applications*. Methods (Vol. 8). John Wiley & Sons, Ltd.
- Urickva V., Sadecka J. (2015). *Determination of geographical origin of alcoholic beverages using ultraviolet, visible and infrared spectroscopy: a review*. Spectrochemical Acta Part A: Molecular and biomolecular Spectroscopy, 148, 131-137.
- Tarakowski, R., Malanowski, A., Kościeszka, R., & Siegoczyński, R. M. (2014). *VIS spectroscopy and pressure induced phase transitions - Chasing the olive oils quality*. Journal of Food Engineering, 122, 28-32
- Teixeira dos Santos C.A., Ricardo N.M.J, Sarraguca M.C., Porto P.A.L.S., Cerdeira A.L., Gonzalez-Saiz J.M., Pizarro C., Lopes J.A. (2017). *Merging vibrational spectroscopic data for wine classification according to the geographic origin*. Food Research International, 102, 504-510.
- Trygg J., Wold S. (2002). *Orthogonal projections to latent structures (O-PLS)*. Journal of Chemometrics, 16, 119–128.
- Yang L., Xiong, Y., & Min, S. (2019). *Data fusion strategy in quantitative analysis of spectroscopy relevant to olive oil adulteration*. Vibrational Spectroscopy, 101, 20–27.
- Zandomenighi, M., Carbonaro, L., & Caffarata, C. (2005). *Fluorescence of Vegetable Oils: Olive Oils*. Journal of Agricultural and Food Chemistry, 53, 759-766.
- Zhen Hu X., Liu S., Hong Li, X., Xian Wang C., Lu Ni X., Liu X., Wang Y., Liu Y., Hua Xu C. (2019). *Geographical origin traceability of cabernet sauvignon wines based on infrared fingerprint technology combined with chemometrics*. Scientific reports, 1-9.
- Zhou L., Zhang C., Zhengjun Q., Yong H. (2020). *Information fusion of emerging non-destructive analytical techniques for food quality authentication*. Trends in Analytical Chemistry 127, 115901.