

FACULTAD DE QUÍMICA
DEPARTAMENTO DE QUÍMICA ANALÍTICA



VNIVERSITAT
ID VALÈNCIA

NUEVOS DESAFÍOS EN ESPECTROSCOPIA VIBRACIONAL

Memoria presentada por
Mari Merce Cascant Vilaplana
para la obtención del título de Doctor

Programa de Doctorado:
Técnicas Experimentales en Química (R.D. 1393/2007)

Mayo 2017

Directores:

Prof. Dr. Salvador Garrigues Mateo
Prof. Dr. Miguel de la Guardia Cirugeda



D. Salvador Garrigues Mateo y D. Miguel de la Guardia Cirugeda, Catedráticos del Departamento de Química Analítica de la Universidad de Valencia,

CERTIFICAN

Que **Dña. Mari Merce Cascant Vilaplana** ha realizado la presente Tesis Doctoral, que lleva por título "*Nuevos desafíos en espectroscopia vibracional*" bajo su dirección en el Departamento de Química Analítica de la Universidad de Valencia y **AUTORIZAN** su presentación para optar a la obtención del Grado de Doctora por la Universitat de València.

Y para que así conste, a los efectos oportunos, firman la presente en Burjassot, a 25 de Mayo de 2017.

Salvador Garrigues Mateo

Miguel de la Guardia Cirugeda

A mi familia

AGRADECIMIENTOS/ ACKNOWLEDGEMENTS

En primer lugar quisiera agradecer a mis directores de Tesis, el Prof. Dr. Salvador Garrigues Mateo y el Prof. Dr. Miguel de la Guardia Cirugeda, por darme la oportunidad de trabajar a su lado. Por su gran ayuda, dedicación, constancia y paciencia durante estos 4 años. Muchas gracias por hacer posible este proyecto.

También quisiera agradecer a los demás componentes del grupo SOLINQUIANA, Agustín Pastor, María Luisa Cervera, Ángel Morales, Sergio Armenta y Francesc Esteve, ya que de una manera u otra han aportado su granito de arena.

A todo el equipo GREEN de la Universidad de Avignon, por los conocimientos transmitidos, ayuda y sobre todo, por hacer mi estancia verdaderamente fácil y agradable, merci beaucoup.

Al Dr. Gianni Gallelo, por su apoyo y por la confianza depositada en mí. Por las largas y tendidas conversaciones. Por sus ánimos y consejos, sobre todo en la última etapa del doctorado. Ha sido un placer trabajar contigo. Muchas gracias.

A la Dra. Julia Kuligowski, por guiarme y encaminarme en el mundo de la Quimiometría. Muchas gracias.

A mis compañeros de batalla, Alba, David, Manuela, Daniel Sanjuán y Daniel Gallart, por esos buenos momentos y no tan buenos, que sin vosotros hubiese sido difícil sobrellevar. Mención especial merece Alba. Mi amiga. Que a pesar de que hace tan sólo unos años que nos conocemos, sé que siempre podré contar contigo. A los compañeros de departamento, sobre todo a Enrique, por estar siempre ahí y por su ayuda desinteresada. A aquellos que hicieron de las comidas el momento perfecto para desconectar. Gracias.

Finalmente, y no por ello menos importante, quisiera agradecer a mis padres, Miguel y Merce, el esfuerzo y dedicación hacia sus hijas, por querer darnos siempre lo mejor, por hacer lo posible y lo imposible por ayudarnos. Gracias por la confianza depositada en mí. Gracias por estar siempre ahí. A mi hermana Mónica, a mis cuñados y sobrinos, a mis amigos. A todos vosotros, gracias por vuestro cariño, consejos, apoyo incondicional y por entenderme, sobre todo estos últimos meses de nerviosismo.

A Josele, gracias por tu paciencia, por entenderme, por aguantar esos meses separados y por los viajes que hacías por estar conmigo. Por tus consejos. Por esos momentos, por nuestros momentos. Por esa personita que llevo dentro de mí y que en unos meses conoceremos. Por emprender juntos la que será la carrera más bonita e importante de nuestras vidas. Gracias por estar ahí.

A todos vosotros, gracias.

Esta Tesis Doctoral se ha realizado en el grupo SOLINQUIANA del Departamento de Química Analítica de la Universitat de Valencia, en el marco de los Proyectos de Investigación “Greening analytical determinations” (Ministerio de Economía y Competitividad, ref. CTQ2011-25743), “Green Analytical Chemistry: desarrollo de herramientas sostenibles de análisis” (Generalitat Valenciana, ref. PROMETEO II/2014/077) y “Métodos de criba 'in situ' para resolver problemas en los sectores clínico-forense, medioambiental y alimentario” (Ministerio de Economía y Competitividad; ref. CTQ2014-52841-P).

Agradecer a la Secretaria de Estado de Investigación, Desarrollo e Innovación del Ministerio de Economía y Competitividad por la concesión de una ayuda predoctoral de Formación de Personal Investigador (BES-2012-055404) así como la concesión de una ayuda de movilidad predoctoral (EEBB-I-16-11854) para la realización de estancias breves en Centros I+D.

Acrónimos

Acrónimo	Término en inglés	Término en español
AA	Arachidonic acid	Ácido araquidónico
ALA	α -linolenic acid	Ácido α -linolénico
ATR	Attenuated Total Reflectance	Reflectancia Total Atenuada
CCA	Clean Air Act	Ley de Aire Limpio
CHD	Coronary Heart Disease	Enfermedad coronaria
COSMO-RS	Conductor-like Screening Model for Real Solvent	
CPME	Cyclopentyl Methyl Ether	Ciclopentil metil eter
CV	Cross Validation	Validación cruzada
DAGs	Diglycerides	Diglicéridos
DFT	Density Functional Theory	Teoría Funcional de Densidad
DHA	Docosahexaenoic acid	Ácido docosahexaenoico
DMC	Dimethyl Carbonate	Dimetilcarbonato
DPA	Docosapentaenoic acid	Ácido docosapentaenoico
DRIFT	Diffuse Reflectance Fourier Transformed	Reflectancia Difusa por Transformada de Fourier
EPA	Eicosapentaenoic acid	Ácido eicosapentaenoico
EPA	Environmental Protection Agency	Agencia de Protección Ambiental
EtAc	Ethyl Acetate	Etil Acetato
EtOH	Ethanol	Etanol
EHS	Environment, Human and Safety	Medio Ambiente, Salud y Seguridad
FAMES	Fatty Acids Methyl Ester	Ésteres metílicos de los ácidos grasos

Acrónimo	Término en inglés	Término en español
FD	First derivative	Primera derivada
FFAs	Free Fatty Acids	Ácidos grasos libres
FN	False Negative	Falso negativo
FP	False Positive	Falso Positivo
FT	Fourier Transform	Transformada de Fourier
GC/FID	Gas Chromatography with a Flame Ionization Detector	Cromatografía de gases con detector de ionización de llama
HAP	Hazardous Air Pollutants	Contaminante del aire peligroso
HCA	Hierarchical Cluster Analysis	Análisis jerárquico de conglomerados
HPTLC	High Performance Thin-Layer Chromatography	Cromatografía de capa fina de alta resolución
HSP	Hansen solubility parameters	Parámetros de Solubilidad de Hansen
IPA	Isopropanol	Isopropanol
IR	Infrared	Infrarrojo
K-S algorithm	Kennard-Stone algorithm	Algoritmo de Kennard-Stone
LV's	Latent Variables	Variables Latentes
LA	Linoleic Acid	Ácido linoleico
LCn-3PUFA	Long Chain Omega 3 Polyunsaturated Fatty Acids	Ácidos grasos poliinsaturados Omega-3 de cadena larga
LCn-6PUFA	Long Chain Omega 6 Polyunsaturated Fatty Acids	Ácidos grasos poliinsaturados Omega-6 de cadena larga
MAGs	Monoglycerides	Monoglicéridos
MC	Mean Center	Centrado en la Media
MCC	Matthews Correlation Coefficient	Coficiente de correlación de Matthews

Acrónimo	Término en inglés	Término en español
MeTHF	2-methyltetrahydrofurane	2- metiltetrahidrofurano
MIR	Mid Infrared	Infrarrojo medio
MSC	Multiplicative Scattering Correction	Corrección de la dispersión Multiplicativa
MUFAs	Monounsaturated Fatty Acids	Ácidos grasos monoinsaturados
NIR	Near Infrared	Infrarrojo cercano
OA	Oleic acid	Ácido Oleico
PA	Palmitic acid	Ácido Palmítico
PC	Principal Component	Componentes Principales
PCA	Principal Components Analysis	Análisis de Componentes Principales
PLS	Partial Least Squares	Mínimos Cuadrados Parciales
PLS-DA	Partial Least Squares-Discriminant Analysis	Mínimos Cuadrados Parciales-Análisis Discriminante
PUFAs	Poliunsaturated Fatty Acids	Ácidos grasos poliinsaturados
PLS-DA	Partial Least Squares-Discriminant Analysis	Mínimos Cuadrados Parciales-Análisis Discriminante
PUFAs	Poliunsaturated Fatty Acids	Ácidos grasos poliinsaturados
REACH	Registration, Evaluation, Authorisation and restriction of Chemicals	Registro, evaluación, autorización y restricción de sustancias químicas
RED	Relative Energy Difference	Diferencia de Energía Relativa
RMSEC	Root Mean Square Error of Calibration	Error cuadrático medio de calibración
RMSECV	Root Mean Square Error of Cross Validation	Error cuadrático medio de validación cruzada

Acrónimo	Término en inglés	Término en español
RPD	Residual Predictive Deviation	Desviación predictiva residual
RRMSEP	Relative Root Mean Square Prediction	Error cuadrático medio de Predicción
RSS	Residual Sum of Squares	Suma residual de cuadrados
SD	Second Derivative	Segunda Derivada
sd	Standard deviation	Desviación estandar
SNV	Standard Normal Variation	Variación Normal Estándar
SFAs	Saturated Fatty Acids	Ácidos grasos saturados
TAGs	Triglycerides	Triglicéridos
TP	True Positive	Verdadero Positivo
TN	True Negative	Verdadero Negativo
Wn	Wavenumber	Número de ondas

ÍNDICE GENERAL

RESUMEN	1
ABSTRACT (resumen traducido al inglés)	9
1. OBJETIVOS	17
2. INTRODUCCIÓN	21
2.1. Química Analítica Verde	23
2.2. Espectroscopia vibracional: Fundamentos e Instrumentación	28
2.2.1. Espectroscopia infrarroja	30
2.2.1.1. Instrumentación IR	31
2.2.1.2. Espectroscopia infrarroja en la zona del IR medio (MIR)	33
2.2.1.3. Espectroscopia infrarroja en la zona del IR cercano (NIR)	34
2.2.1.4. Métodos de medida	35
2.2.2. Espectroscopia Raman	39
2.2.2.1. Instrumentación Raman	41
2.2.2.2. Métodos de medida	42
2.2.3. Análisis cualitativo	44
2.2.4. Análisis cuantitativo	45
2.3. Quimiometría y Modelización	47
2.3.1. Generalidades	47
2.3.2. Técnicas multivariantes	47
2.3.2.1. Técnicas No Supervisadas	48
2.3.2.2. Técnicas Supervisadas	49
2.3.3. Construcción, validación y evaluación del modelo quimiométrico	50
2.3.3.1. Selección del set de calibración y validación	51
2.3.3.2. Pre-procesamiento de datos	52
2.3.3.3. Selección de la región espectral	53
2.3.3.4. Variables latentes	53
2.3.3.5. Muestras aberrantes o anómalas	54
2.3.3.6. Evaluación del modelo	54
2.4. Referencias Introducción	56

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN	61
PARTE 1. MUESTRAS DE ALIMENTOS	65
3.1. Capítulo 1. Determinación directa de componentes mayoritarios en alimentos	69
3.1.1. Introducción	71
3.1.2. Parte experimental	72
3.1.2.1. Instrumentos	72
3.1.2.2. Muestras	73
3.1.2.3. Métodos de referencia	74
3.1.3. Resultados y discusión	74
3.1.3.1. Espectroscopia NIR	74
3.1.3.1.1. Modelos PLS-NIR	75
3.1.3.1.2. Capacidad de predicción del modelo PLS-NIR	80
3.1.3.2. Espectroscopia MIR	82
3.1.3.2.1. Modelos PLS-ATR-FTIR	83
3.1.3.2.2. Capacidad de predicción del modelo PLS-ATR-FTIR	88
3.1.3.3. Comparación entre NIR y MIR	89
3.1.4. Conclusiones Capítulo 1	91
3.1.5. Referencias Capítulo 1	92
3.2. Capítulo 2. Determinación de compuestos polares en aceites	95
3.2.1. Introducción	97
3.2.2. Parte experimental	99
3.2.2.1. Instrumentos	99
3.2.2.2. Método de referencia	99
3.2.A. Capítulo 2.A. Comparación de la espectroscopia FT-NIR y FT-MIR para la determinación de compuestos polares en aceites	100
3.2.A.1. Muestras	100
3.2.A.2. Resultados y discusión	100
3.2.A.2.1. Espectroscopia NIR	100
3.2.A.2.2. Modelos PLS-NIR	101
3.2.A.2.3. Espectroscopia MIR	105
3.2.A.2.4. Modelos PLS-ATR-FTIR	106
3.2.A.2.5. Comparación MIR y NIR para la determinación de TPM	110
3.2.A.3. Conclusiones Capítulo 2.A	111

3.2.B. Capítulo 2.B. Estudios previos para el desarrollo de un instrumento de bajo coste para la evaluación de la calidad del aceite de fritura	112
3.2.B.1. Muestras	112
3.2.B.2. Resultados y discusión	112
3.2.B.2.1. Espectros NIR	112
3.2.B.2.2. Modelización	114
3.2.B.2.2.1. MLR y PLS con reducción de variables	115
3.2.B.2.2.1.1. Modelos MLR-NIR	115
3.2.B.2.2.1.2. Modelos PLS-NIR	117
3.2.B.2.2.2. Modelo PLS-NIR sin variables reducidas	118
3.2.B.2.3. Comparación de los diferentes modelos de calibración	120
3.2.B.3. Conclusiones Capítulo 2.B	122
3.2.3. Referencias Capítulo 2	123
3.3. Capítulo 3. Determinación de lípidos en aceite de salmón	125
3.3.1. Introducción	127
3.3.2. Parte experimental	131
3.3.2.1. Muestras	131
3.3.2.2. Extracción del aceite de salmón: Método Soxhlet	131
3.3.2.3. Métodos de referencia	132
3.3.2.3.1. HP-TLC	132
3.3.2.3.2. Cromatografía de gases	133
3.3.A. Capítulo 3.A. Un enfoque desde la Química Analítica Verde para el estudio de la extracción de lípidos: métodos de cálculo computacional para la selección de disolventes verdes como alternativa al hexano	134
3.3.A.1. Disolventes alternativos al n-hexano	134
3.3.A.2. Métodos computacionales	135
3.3.A.2.1. Hansen Solubility Parameters (HSP)	136
3.3.A.2.2. Conductor-like Screening Model Real Solvent (COSMO-RS)	138
3.3.A.3. Resultados y discusión	140
3.3.A.3.1. Cálculos mediante HSP	140
3.3.A.3.2. Cálculos con el COSMO-RS	142
3.3.A.3.3. Análisis HPTLC de las fracciones lipídicas extraídas	143
3.3.A.3.4. Análisis GC-FID de las fracciones lipídicas extraídas	146
3.3.A.4. Conclusiones Capítulo 3.A	150

3.3.B. Capítulo 3.B. Determinación de ácidos grasos y clases de lípidos en el aceite de salmón por espectroscopia de infrarrojo cercano	153
3.3.B.1. Resultados y discusión	153
3.3.B.1.1. Espectros NIR	153
3.3.B.1.2. Modelos PLS-NIR	154
3.3.B.1.3. Evaluación de la sostenibilidad del método propuesto	162
3.3.B.2. Conclusiones Capítulo 3.B	163
3.3.3. Referencias Capítulo 3	164
3.4. Capítulo 4. Control de calidad de los aditivos umami por Raman y espectroscopia infrarroja en la región cercana	169
3.4.1. Introducción	171
3.4.2. Parte experimental	173
3.4.2.1. Instrumentación	173
3.4.2.2. Muestras	173
3.4.3. Resultados y discusión	174
3.4.3.1. Espectroscopia Raman	174
3.4.3.1.1. Modelos PLS-Raman: Mezclas binarias (GMP + IMP)	175
3.4.3.1.2. Modelos PLS-Raman: Mezclas ternarias (MSG + GMP + IMP)	179
3.4.3.2. Espectroscopia NIR	179
3.4.3.2.1. Modelos PLS-NIR Mezclas binarias (GMP + IMP)	180
3.4.3.2.2. Modelos PLS-NIR: Mezclas ternarias (MSG + GMP + IMP)	181
3.4.3.3. Comparación entre el Raman y NIR para el análisis de mezclas umami	183
3.4.3.4. Evaluación de los métodos	183
3.4.4. Conclusiones Capítulo 4	184
3.4.5. Referencias Capítulo 4	186
PARTE 2. MUESTRAS AMBIENTALES Y DE INTERÉS AGRÍCOLA	189
3.5. Capítulo 5. Evaluación de compost para su posterior uso en la agricultura mediante espectroscopia infrarroja NIR y MIR	193
3.5.1. Introducción	195
3.5.2. Parte experimental	197
3.5.2.1. Instrumentos	197
3.5.A. Capítulo 5.A. Determinación de compuestos fenólicos totales en compostajes mediante NIR y MIR	198
3.5.A.1. Muestras	198
3.5.A.2. Método de referencia	198

3.5.A.3. Resultados y discusión	199
3.5.A.3.1. Espectroscopia NIR	199
3.5.A.3.1.1. Análisis exploratorio: PCA	200
3.5.A.3.1.2. Modelos PLS-NIR	201
3.5.A.3.2. Espectroscopia MIR	204
3.5.A.3.2.1. Análisis exploratorio: PCA	205
3.5.A.3.2.2. Modelos PLS-ATR-FTIR	206
3.5.A.4. Comparación MIR y NIR	209
3.5.A.5. Conclusiones Capítulo 5.A	209
3.5.B. Capítulo 5.B. Determinación del contenido de carbono orgánico y nitrógeno total	210
3.5.B.1. Muestras	210
3.5.B.2. Métodos de referencia	210
3.5.B.3. Resultados y discusión	211
3.5.B.3.1. Espectroscopia NIR	211
3.5.B.3.1.1. Modelos PLS-NIR: Carbono Orgánico	211
3.5.B.3.1.2. Modelos PLS-NIR: Nitrógeno Total	214
3.5.B.3.2. Espectroscopia MIR	217
3.5.B.3.2.1. Modelos PLS-MIR: Carbono orgánico	217
3.5.B.3.2.2. Modelos PLS-MIR: Nitrógeno Total	220
3.5.B.4. Comparación MIR y NIR	223
3.5.B.5. Conclusiones Capítulo 5.B	223
3.5.3. Referencias Capítulo 5	225
3.6. Capítulo 6. Determinación de benceno, tolueno, etilbenceno y xilenos en suelos mediante espectroscopia infrarroja en la región media en fase de vapor	229
3.6.1. Introducción	231
3.6.2. Parte experimental	233
3.6.2.1. Instrumentación	233
3.6.2.2. Patrones y reactivos	234
3.6.2.3. Generación de la fase vapor	234
3.6.2.4. Método de referencia: cromatografía de gases	235
3.6.3. Resultados y discusión	235
3.6.3.1. Evaluación de las condiciones experimentales e instrumentales	235
3.6.3.2. Espectro IR del benceno, tolueno, etilbenceno y xilenos en fase vapor	237
3.6.3.3. Modelos PLS	240
3.6.3.4. Análisis de muestras de suelo contaminado	242

3.6.4. Conclusiones Capítulo 6	246
3.6.5. Referencias Capítulo 6	248
3.7. Capítulo 7. Correcciones de vapor en espectroscopia infrarroja	253
3.7.1. Introducción	255
3.7.2. Parte experimental	256
3.7.2.1. Instrumentación	256
3.7.2.2. Medidas de las muestras	256
3.7.2.3. Espectros de referencia y corrección de muestras	257
3.7.3. Resultados y discusión	259
3.7.3.1. Espectro de FTIR en fase vapor en presencia del vapor de acetona	259
3.7.3.2. Evaluación de la corrección de vapor de la acetona	262
3.7.4. Conclusiones Capítulo 7	266
3.7.5. Referencias Capítulo 7	267
PARTE 3. MUESTRAS ARQUEOLÓGICAS	269
3.8. Capítulo 8. Predicción de elementos alcalinotérreos en restos óseos por espectroscopia de la región del infrarrojo cercano	273
3.8.1. Introducción	275
3.8.2. Parte experimental	276
3.8.2.1. Instrumentos	276
3.8.2.2. Muestras	277
3.8.2.3. Método de referencia	277
3.8.3. Resultados y discusión	278
3.8.3.1. Espectros FT-NIR	278
3.8.3.2. Exploración de los datos por PCA	281
3.8.3.3. Modelos PLS-NIR	283
3.8.3.4. Capacidad de predicción de los modelos PLS-NIR	287
3.8.4. Conclusiones Capítulo 8	289
3.8.5. Referencias Capítulo 8	290
3.9. Capítulo 9. Investigaciones forenses de huesos quemados mediante el uso de la espectroscopia infrarroja en la región cercana	295
3.9.1. Introducción	297
3.9.2. Parte experimental	299
3.9.2.1. Instrumentos	299
3.9.2.2. Muestras	300
3.9.2.3. Método de referencia	300

3.9.3. Resultados y discusión	301
3.9.3.1. Espectros NIR	301
3.9.3.2. Análisis exploratorio	303
3.9.3.2.1. Análisis de Componentes Principales (PCA)	303
3.9.3.2.2. Cluster	305
3.9.3.3. PLS-DA	307
3.9.3.4. Modelos PLS	309
3.9.4. Control de calidad (QC)	313
3.9.5. Evaluación del método propuesto	313
3.9.6. Aplicaciones forenses	314
3.9.7. Conclusiones	315
3.9.8. Referencias	317
4. CONCLUSIONES	323
4. FINAL CONCLUSIONS (conclusiones traducidas al inglés)	337
5. ANEXO: Artículos científicos publicados	349
5.1. Direct determination of major components in human diets and baby foods	351
5.2. A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane	365
5.3. Determination of total phenolic compounds in compost by infrared spectroscopy	381
5.4. Prediction of organic carbon and total nitrogen contents in organic wastes and their composts by Infrared spectroscopy and partial least square regression	389
5.5. Prediction of alkaline earth elements in bone remains by near infrared spectroscopy	399
5.6. Burned bones forensic investigations employing near infrared spectroscopy	409

RESUMEN

NUEVOS DESAFÍOS EN ESPECTROSCOPIA VIBRACIONAL

El principal objetivo de la Química Analítica Verde es la sustitución de metodologías contaminantes por otras más limpias, mediante:

- La reducción de reactivos y disolventes empleados en el pre-tratamiento y análisis de las muestras.
- Minimizar las cantidades y de la toxicidad de los compuestos empleados en las etapas de preparación y medida, especialmente mediante automatización y miniaturización.
- El desarrollo de metodologías directas de análisis sin uso de reactivos o disolventes y que, por tanto, no generen desechos tóxicos ni peligrosos.

Las técnicas espectroscópicas junto con el uso de la quimiometría pueden ser una buena alternativa a los métodos tradicionales, ya que se pueden realizar medidas directas de las muestras sin el uso de reactivos ni disolventes. Por ello, mediante el uso de espectroscopia de infrarrojo cercano (NIR), medio (MIR) y Raman y, con la ayuda de herramientas quimiométricas, se han desarrollado metodologías que cumplen con los objetivos de la Química Analítica Verde.

En aquellos casos en los que es necesaria una extracción de los analitos para su posterior determinación, se persiguió:

- Asegurar la alta calidad extracto/producto.
- El diseño de procesos de extracción reduciendo el consumo de energía.
- Uso de disolventes alternativos y productos renovables naturales, siendo los agro-solventes y bio-solventes los favoritos para reemplazar disolventes derivados del petróleo.

Esta Tesis Doctoral se ha estructurado en 3 apartados.

El primer apartado corresponde a los **Objetivos**, siendo el principal objetivo del proyecto, el desarrollo de una Química Analítica Verde a través de la simplificación de las formas de preparación de muestras y el empleo de metodologías directas mediante el uso de técnicas espectroscópicas en el análisis de muestras de diferente tipo.

El segundo apartado corresponde a la **Introducción**, considerando:

- **Química Analítica Verde.**

Se describen los conceptos fundamentales a partir de los 12 Principios de la Química Analítica Verde, así como la evaluación de los métodos analíticos.

- **Espectroscopia vibracional.**

Se resumen los fundamentos de la espectroscopia vibracional en el infrarrojo medio, cercano y Raman.

- **Quimiometría.**

Incluye las técnicas multivariantes utilizadas en esta Tesis Doctoral, ya sean supervisadas o no supervisadas, junto con la construcción, validación y evaluación de los modelos quimiométricos y descripción de los parámetros y pre-procesamientos más importantes.

En el tercer apartado se recogen los **Resultados Experimentales y Ámbito de aplicación** de esta Tesis Doctoral, estructurados en 3 Partes y subdivididos en 9 capítulos, constituidos por 12 trabajos; 6 de los cuales se han publicado y los 6 restantes se han enviado a revistas científicas para ser publicados. Al final de cada capítulo se exponen las conclusiones más relevantes obtenidas en cada uno de los estudios realizados.

Parte 1. Muestras de Alimentos

- En el **Capítulo 1** se desarrolla un método rápido para la determinación de grasas, proteínas, hidratos de carbono y valor energético de tarritos de bebé, comida rápida para niños y menús de comedores universitarios, mediante el uso de medidas NI) y MIR utilizando métodos de calibración multivariante.

(M.M Cascant, S. Garrigues, M. de la Guardia, Direct determination of major components in human diets and baby foods, Ana.l Bioanal. Chem., 407(2015) 1961:1972)

- En el **Capítulo 2** se describe la determinación de los compuestos polares en aceite de fritura a través de 2 trabajos:

2.A. Comparación de la espectroscopia FT-NIR y FT-MIR en asociación con la calibración de mínimos cuadrados parciales (PLS) como herramienta verde para la determinación directa de compuestos polares en aceites de fritura.

2.B. Estudios previos para el desarrollo de un instrumento de bajo coste para la evaluación de la calidad del aceite de fritura.

- En el **Capítulo 3**, se estudia la determinación del contenido de los lípidos en aceite de salmón, mediante 2 trabajos:

3.A. Estudio de la extracción de los lípidos del salmón mediante disolventes alternativos al hexano con la ayuda de dos métodos computacionales como son; el *Hansen Solubility Parameter* (HSP) y *Conductor-like Screening Model for Real Solvent* (COSMO-RS) y mediante el uso de estudios experimentales para comparar los disolventes alternativos empleados con el hexano.

Cabe mencionar que este trabajo fue realizado en una estancia predoctoral de tres meses en el *Groupe de Recherche en Eco-Extraction de produits Naturels (GREEN)* de la Universidad de Avignon (Francia) bajo la supervisión del Prof. Farid Chemat.

(M.M. Cascant, C. Breil, S. Garrigues, M. de la Guardia, A.S. Fabiano-Tixier, F. Chemat, A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane, Anal. Bioanal. Chem., 409(2017) 3527:3539)

3.B. Evaluación de la espectroscopia del infrarrojo cercano en combinación con los modelos PLS para la determinación simultánea de i) ácidos grasos como ácido oleico, ácido palmítico, ácido linolénico, ácido linoleico, ii) familias de ácidos grasos como omega-3, omega-6, ácidos grasos poliinsaturados, ácidos grasos monoinsaturados y ácidos grasos saturados, y iii) clases de lípidos como triacilglicéridos, diacilglicéridos, ácidos grasos libres y ergosterol en aceite de salmón

Parte de este trabajo, relativo a la obtención de los resultados de referencia, se realizó durante la estancia predoctoral antes mencionada en el grupo GREEN de la Universidad de Avignon (Francia).

- El **Capítulo 4** se basa en el Control de calidad de aditivos alimentarios mediante el uso de espectroscopia Raman e infrarrojo cercano (NIR)

Parte 2. Muestras ambientales y de interés agrícola

- El **Capítulo 5**, basado en la evaluación de composts para su posterior uso en la agricultura mediante espectroscopia NIR y MIR, se encuentra compuesto por 2 trabajos:

5.A. Determinación de compuestos fenólicos totales

(M.M. Cascant, M. Sisouane, S. Tahiri, M. El Krati, M.L. Cervera, S. Garrigues, M. de la Guardia, Determination of total phenolic compounds in compost by infrared spectroscopy, Talanta, 153(2016) 360:365)

5.B. Predicción del contenido de carbono orgánico y de nitrógeno total

(M. Sisouane, M.M. Cascant, S. Tahiri, S. Garrigues, M. El Krati, G.E. Boutchich, M.L. Cervera, M. de la Guardia, Prediction of organic carbon and total nitrogen contents in organic wastes and their composts by Infrared spectroscopy and partial least square regression, Talanta, 167(2017) 352:358)

En ambos casos las estrategias desarrolladas se basan en las medidas directas de las muestras sin tratamiento previo y sin el uso de reactivos ni disolventes. Los resultados obtenidos proporcionaron una capacidad predictiva aceptable para la determinación de compuestos fenólicos totales, mientras que para el caso del contenido de carbono orgánico y nitrógeno total los resultados obtenidos fueron aceptables para fines de cribado.

- El **Capítulo 6** se refiere al desarrollo de una metodología para la determinación rápida y directa de analitos volátiles en suelos, combinando el espacio de cabeza (HS) con la espectroscopia infrarroja (FTIR).
- El **Capítulo 7** se basa en la evaluación de un método flexible y automatizado para evitar las interferencias espectrales de vapores orgánicos en espectros FTIR medidos en fase de vapor, teniendo como compuesto diana la acetona, utilizando una metodología que combina el uso de espacio de cabeza con espectroscopia infrarroja para muestras de BTEX y EtOH.

Parte 3. Muestras arqueológicas

- El **Capítulo 8** describe un método analítico para la predicción de elementos alcalinotérreos y detección de procesos diagenéticos en muestras de huesos enterrados mediante espectroscopia de infrarrojo cercano.

(M.M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Prediction of alkaline earth elements in bone remains by near infrared spectroscopy, Talanta, 162(2017) 428:434)

- En el **Capítulo 9** se hace referencia al uso del infrarrojo cercano para el estudio y clasificación de huesos quemados. El Análisis de Componentes Principales y la clasificación de conglomerados se utilizaron como técnicas exploratorias, a

partir de los espectros de los huesos y de los sedimentos, para la selección de muestras de hueso quemado menos afectadas por procesos ambientales, para realizar adecuadamente estudios forenses. El Análisis Discriminante por Mínimos Cuadrados Parciales se utilizó para clasificar las muestras óseas según sus condiciones de combustión, y además, se construyeron modelos de regresión por mínimos cuadrados parciales para predecir la concentración de calcio, magnesio y estroncio de las muestras óseas.

(M.M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Burned bones forensic investigations employing near infrared spectroscopy, Vibrational Spectroscopy, 90 (2017) 21:30)

Para finalizar, se presentan las **Conclusiones generales** de esta Tesis Doctoral; indicando que la espectroscopia vibracional combinada con el uso de herramientas quimiométricas ofrece una buena alternativa a los métodos tradicionales, siguiendo los principales objetivos de la Química Analítica Verde.

ABSTRACT

NEW CHALLENGES IN VIBRATIONAL SPECTROSCOPY

The main objective of Green Analytical Chemistry is to replace polluting methodologies with cleaner ones by:

- Reducing reagents and solvents used in the pre-treatment and analysis of the samples.
- Minimizing the amounts and toxicity of the compounds used in the preparation and measurement steps, especially through automation and miniaturization.
- Developing direct analysis methodologies without the use of reagents or solvents which, therefore, do not generate toxic or hazardous waste.

Spectroscopic techniques together with the use of chemometrics can be a good alternative to traditional methods, since direct measurements of the samples can be made without the use of reagents or solvents. Therefore, through the use of near infrared (NIR), middle (MIR) and Raman spectroscopy combined with chemometric tools, methodologies have been developed in order to meet the objectives of Green Analytical Chemistry.

In those cases in which an extraction of the analytes is necessary for their subsequent determination, it was pursued to:

- Ensure high quality extract / product.
- Design extraction processes reducing energy consumption.
- Use alternative solvents and natural renewable products, being agro-solvents and bio-solvents the favorite to replace solvents derived from petroleum.

This Doctoral Thesis has been structured in 3 sections.

The first section corresponds to the **Objectives**, being the main objective of the project, the development of a Green Analytical Chemistry through the simplification of the sample preparation and the use of direct measurement methodologies by using spectroscopic techniques in the analysis of different kind of samples.

Second section corresponds to the **Introduction**, considering:

▪ **Green Analytical Chemistry.**

Fundamental concepts starting from the 12 Principles of Green Analytical Chemistry are described, as well as the greenness evaluation of analytical methods.

▪ **Vibrational spectroscopy.**

Vibrational spectroscopy fundamentals in the middle, near infrared and Raman are summarized.

▪ **Chemometrics.**

The multivariate techniques used in this Doctoral Thesis are comprehended here, whether supervised or unsupervised, together with the construction, validation and evaluation of the chemometric models and the description of the most important parameters and data pre-processing methods applied.

In the third section, the **Experimental Results and Scope** of this Doctoral Thesis are presented and discussed. This section is structured in 3 parts and subdivided into 9 chapters, consisting of 12 works; 6 of which have been already published and the remaining 6 have been sent to scientific journals to be published. At the end of each chapter the most relevant conclusions obtained in each of the studies are presented.

Part 1. Food Samples

- **Chapter 1** corresponds to the development of a fast method for the determination of fats, proteins, carbohydrates and energetic values for baby jars, children's fast food, and university canteens menus, by using NIR and MIR spectroscopy, and using multivariate calibration methods.

(M.M Cascant, S. Garrigues, M. de la Guardia, Direct determination of major components in human diets and baby foods, Anal. Bioanal. Chem., 407(2015) 1961:1972)

- **Chapter 2** describes the polar compounds determination in frying oil through 2 works:

2.A. Comparison of FT-NIR and FT-MIR spectroscopy in association with partial least squares (PLS) calibration models as a green tool for the direct determination of polar compounds in frying oils.

2.B. Previous studies for low cost instrument development for frying oil quality evaluation.

- In **Chapter 3**, the determination of the lipid content in salmon oil is studied by means of 2 works:

3.A. Study of lipid extraction from salmon using alternative solvents to hexane by using two computational methods such as; Hansen Solubility Parameter (HSP) and Conductor-like Screening Model for Real Solvent (COSMO-RS) and by using experimental studies to compare the alternative solvents used with hexane as a reference.

It should be mentioned that this work was carried out in a three-month predoctoral stay at the *Groupe de Recherche en Eco-Extraction of produits Naturels (GREEN)* of the University of Avignon (France) under the supervision of Prof. Farid Chemat.

(M.M. Cascant, C. Breil, S. Garrigues, M. de la Guardia, A.S. Fabiano-Tixier, F. Chemat, A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane, Anal. Bioanal. Chem. 409(2017) 3527:3539).

3.B. Evaluation of the near infrared spectroscopy in combination of PLS models for the simultaneous determination of: i) fatty acids as oleic acid, palmitic acid, linolenic acid, linoleic acid, ii) fatty acid families as omega-3, omega-6, polyunsaturated fatty acids, monounsaturated fatty acids and saturated fatty acids, and iii) lipid class as triacylglycerides, diacylglycerides, free fatty acids and ergosterol in salmon oil.

Part of this work, related to obtaining the reference results, was performed during the abovementioned predoctoral stay in the group *GREEN* of the University of Avignon (France).

- **Chapter 4** is based on the Quality Control of Food Additives through the use of Raman and NIR for direct measurement of samples contained into glass vials.

Part 2. Environmental and agricultural interest samples

- **Chapter 5**, is devoted to the evaluation of composts for later use in agriculture using NIR and MIR spectroscopy, and consists of 2 works:

5.A. Determination of total phenolic compounds

(M.M. Cascant, M. Sisouane, S. Tahiri, M. El Krati, M.L. Cervera, S. Garrigues, M. de la Guardia, Determination of total phenolic compounds in compost by infrared spectroscopy, Talanta, 153(2016) 360:365).

5.B. Prediction of organic carbon and total nitrogen content

(M. Sisouane, M.M. Cascant, S. Tahiri, S. Garrigues, M. El Krati, G.E. Boutchich, M.L. Cervera, M. de la Guardia, Prediction of organic carbon and total nitrogen contents in organic wastes and their composts by Infrared spectroscopy and partial least square regression, Talanta, 167(2017) 352:358).

In both cases developed strategies are based on the direct measurement of samples without previous pretreatment or the use of neither reagents nor solvents. Results obtained provided an acceptable predictive capability for determination of total phenolic compounds instead the poor capability for organic carbon and nitrogen content that is acceptable for screening purposes.

- **Chapter 6** deals with the development of a methodology for the rapid and direct determination of volatile analytes in soils, combining headspace (HS) with infrared spectroscopy (FTIR).
- **Chapter 7** is based on the evaluation of a flexible and automated method to avoid the spectral interferences of organic vapors in FTIR spectra measured in vapor phase, having acetone as a target compound, using a methodology that combines the headspace use with infrared spectroscopy for BTEX and EtOH samples.

Part 3. Archaeological samples

- **Chapter 8** describes an analytical method for the prediction of alkaline earth elements and detection of diagenetic processes in buried bones samples by near-infrared spectroscopy.

(M.M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Prediction of alkaline earth elements in bone remains by near infrared spectroscopy, Talanta, 162(2017) 428:434)

- **Chapter 9** refers to the use of the near infrared for the study and classification of burned bones. Bone and sediment spectra were treated by Principal Component Analysis and cluster classification as exploratory techniques to select burned

bone samples, less affected by environmental processes, to properly carry out forensic studies. Partial Least Square Discriminant Analysis was used to classify bone samples based on their burning conditions, and Partial Least Square regression models were built to predict calcium, magnesium and strontium concentration of bone samples.

(M.M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Burned bones forensic investigations employing near infrared spectroscopy, Vibrational Spectroscopy, 90 (2017) 21:30)

Finally, the **Final Conclusions** of this Doctoral Thesis are presented; indicating that vibrational spectroscopy with the use of chemometrics tools offers a good alternative to traditional methods, keeping the objectives of Green Analytical Chemistry.

1. OBJETIVOS

1. OBJETIVOS

El principal objetivo de esta Tesis es el uso de la espectroscopia vibracional como método de análisis para el desarrollo de una Química Analítica Verde en diferentes ámbitos de aplicación como pueden ser muestras de: alimentos, ambientales y arqueológicas.

Los métodos empleados habitualmente para el análisis de estas muestras implican un elevado tiempo de análisis, manipulación de las muestras, un gran consumo de disolventes y por consiguiente la generación de residuos, siendo así perjudiciales tanto para el operador como para el medio ambiente. Por tanto, es necesario disponer de metodologías analíticas que minimicen o eliminen estos efectos con el fin de aproximarse a la Química Analítica Verde. En este sentido, esta Tesis Doctoral evalúa la espectroscopia vibracional como posible alternativa a los métodos habituales con el fin de obtener métodos de análisis más “verdes”.

En cuanto los objetivos específicos de la presente Tesis se detallan a continuación, según su ámbito de aplicación:

Muestras de alimentos

- El desarrollo de metodologías para la determinación simultánea de componentes mayoritarios en alimentos.
- Análisis directo de componentes minoritarios y de interés en alimentos.
- Cuando el análisis directo no era aplicable, y fue necesaria una extracción de los analitos se intentó reducir el uso de disolventes o se utilizaron disolventes alternativos tales como los agro-solventes y bio-solventes para reemplazar disolventes derivados del petróleo.
- Desarrollo de sistemas de evaluación de bajo coste para su aplicación en el análisis de rutina.
- Control de calidad de productos acabados in situ, sin necesidad de transportarlos hasta el laboratorio.

Muestras ambientales

- Evaluación de parámetros en composts para su posterior uso en agricultura mediante espectroscopia NIR y MIR para ofrecer una alternativa no destructiva, rápida y relativamente barata a los métodos tradicionales.

- Determinación directa de analitos volátiles en suelos para evitar la manipulación de la muestra.
- Evaluación de un método flexible y automatizado para evitar las interferencias espectrales de los vapores orgánicos en los espectros FTIR medidos en fase vapor.

Muestras arqueológicas

- Detección de procesos diagenéticos que alteran las propiedades de los huesos enterrados.
- Predicción de elementos alcalinotérreos en restos óseos.
- Clasificación de huesos quemados con fines forenses.

En resumen, mediante esta Tesis Doctoral se pretende demostrar las distintas capacidades, ya sean cuantitativas o cualitativas, que presenta el empleo de la espectroscopia vibracional junto con la Quimiometría, en distintos ámbitos de aplicación, para minimizar la manipulación de muestras, tiempo de análisis así como el uso de reactivos y disolventes, proporcionando una alternativa sostenible a los métodos tradicionales.

2. INTRODUCCIÓN

2. INTRODUCCIÓN

2.1. Química Analítica Verde

La "Química Verde" propone el empleo de técnicas y metodologías capaces de reducir o incluso eliminar el consumo de materias primas, productos, subproductos, disolventes, reactivos y la generación de desechos, que son peligrosos para la salud del ser humano o para el medio ambiente [1]. En 1998, Anastas y Warner [1] propusieron los **12 Principios de la Química Verde**, que son un conjunto de criterios ampliamente aceptados para evaluar cuanto "verde" es un proceso, así como la aceptabilidad medioambiental de los procesos de fabricación de los productos químicos. En 2005, Tang, Smith y Poliakoff [2] propusieron una regla mnemotécnica, explicada en pocas palabras, para recordar de una manera más fácil los 12 Principios de la Química Verde con el objetivo de memorizar estos conceptos con más facilidad que se basaba en el acrónimo **PRODUCTIVELY** (Figura 1).

- P** - Prevent wastes
- R** - Renewable materials
- O** - Omit derivatization step
- D** - Degradable chemical products
- U** - Use safe synthetic methods
- C** - Catalytic reagents
- T** - Temperature, Pressure ambient
- I** - In-process Monitoring
- V** - Very few auxiliary substances
- E** - E-factor, maximise feed in product
- L** - Low toxicity of chemical products
- Y** - Yes, it is safe

Figura 1. 12 Principios de la Química Verde descritos de forma mnemotécnica: PRODUCTIVELY

En la década de los 90, se publicaron los primeros libros y ponencias, y se organizaron simposios sobre la química verde. Las primeras revistas dedicadas a la química verde fueron *Journal of Clean Processes and Products* (desde 1998) y *Green Chemistry* (desde 1999) [3]. La primera conferencia internacional sobre el tema fue organizada en septiembre de 1997 en Viena bajo el patrocinio de la IUPAC. Desde entonces, se han organizado periódicamente conferencias científicas similares [4].

En el año 2000, surge la idea de Química Analítica Verde a partir de la Química Verde [5], en la que los químicos analíticos trabajan para encontrar métodos más respetuosos con el medio ambiente [6-11].

En 2013, Galuszka, Migaszewski y Namiesnik [12] sugirieron que los 12 principios de la Química Verde debían ser revisados para encontrar una aplicación más completa en la Química Analítica. Para ello, utilizaron cuatro de los principios proporcionados por Anastas y Warner [1] y los complementaron con ocho nuevos principios con mayor importancia para la Química Analítica Verde (ver Figura 2).

12 Principios de la Química Analítica Verde

- 1- Aplicación de técnicas analíticas directas para evitar el tratamiento de la muestra.
- 2- Uso de la mínima cantidad de muestra y del número de muestras.
- 3- Realización de mediciones in situ.
- 4- Integración de procesos analíticos y operaciones para minimizar el uso de energía y de reactivos.
- 5- Seleccionar métodos automatizados y miniaturizados.
- 6- Debe evitarse la derivatización.
- 7- La generación de un gran volumen de residuos analíticos debe tener la gestión adecuada.
- 8- Se prefieren los métodos multi-analitos o multiparamétricos antes que los métodos que utilizan un solo analito.
- 9- El uso de energía debe ser minimizado.
- 10- Es preferible el uso de reactivos obtenidos a partir de fuentes renovables.
- 11- Los reactivos tóxicos deben ser eliminados o reemplazados.
- 12- Se debe aumentar la seguridad del operador.

Figura 2. 12 Principios de la Química Analítica Verde

Además, expresaron estos nuevos 12 Principios de la Química Analítica Verde de forma mnemotécnica mediante la palabra **SIGNIFICANCE** (Figura 3).

- S** - Select direct analytical technique
- I** - Integrate analytical processes and operations
- G** - Generate as little waste as possible and treat it properly
- N** - Never waste energy
- I** - Implement automation and miniaturization of methods
- F** - Favor reagents obtained from renewable source
- I** - Increase safety for operator
- C** - Carry out *in-situ* measurements
- A** - Avoid derivatization
- N** - Note that the sample number and size should be minimal
- C** - Choose multi-analyte or multi-parameter method
- E** - Eliminate or replace toxic reagents

Figura 3. 12 Principios de la Química Analítica Verde descritos de forma mnemotécnica: SIGNIFICANCE

En 2012, Gałuszka, Konieczka, Migaszewski y Namiesnik [13] propusieron la Eco-escala analítica como un nuevo enfoque para evaluar cuánto de “verde” es la metodología analítica usada. Este enfoque se basa en la asignación de puntos de penalización en las diferentes etapas del proceso analítico dependiendo de la cantidad de reactivos, peligros, energía y residuos utilizados. La suma de puntos de penalización para todo el procedimiento se incluirá en el cálculo de Eco-escala de acuerdo con la *Ecuación 1*, siendo 100 el método analítico “verde” ideal.

$$\text{Eco-escala analítica} = 100 - \text{Total puntos de penalización} \quad \text{Ecuación 1}$$

Esta evaluación de los métodos analíticos ha sido recientemente modificada por Armenta, de la Guardia y Namiesnik (2017) [14] mediante la asignación de puntos de penalización como una función de las cantidades de los reactivos y residuos, clasificando los métodos mediante el llamado “*Green Certificate*”. Dicha eco-escala se basa en la aplicación de códigos de colores y letras; de la A a la G, siendo la clase A la que corresponde con los métodos analíticos más “verdes”. La puntuación de los métodos evaluados se asocia con el número de puntos de penalización asignados según el uso de reactivos, así como la cantidad utilizada de los mismos, su peligrosidad, la energía consumida, los posibles riesgos laborales y los residuos generados.

Por lo tanto, de acuerdo con los 12 Principios de la Química Analítica Verde, el método analítico verde ideal se caracteriza por eliminar o minimizar el uso de reactivos, la energía y la generación de residuos. Teniendo en cuenta que estos criterios sólo se satisfacen por aquellos métodos que no requieran muestreo, transporte y preparación de las muestras, los métodos de detección remota, los métodos no invasivos y las mediciones de campo serían los más adecuados [15], [16]. Además, todos los desarrollos en quimiometría mejorarán las capacidades multiparamétricas de los instrumentos antes mencionados con el fin de proporcionar tanta información como sea posible con un consumo reducido de reactivos y pocas medidas [16].

Las técnicas espectroscópicas vibracionales como el MIR, NIR y Raman son, frecuentemente, no destructivas y podrían proporcionar un análisis muy rápido, sin ningún tratamiento previo de la muestra, pudiéndose utilizar para el análisis de muestras sólidas, líquidas o gases ya sea a nivel macroscópico o microscópico, y con poca cantidad. Estas técnicas permiten realizar análisis cualitativos mediante identificación de los analitos, análisis cuantitativos, así como el monitoreo y control de las distintas etapas del proceso de producción. Además, la espectroscopia vibracional permite la determinación de varios analitos de manera simultánea en una sola medida, sin la necesidad de separarlos previamente, proporcionando mayores beneficios en el sentido de menor consumo de reactivos, manipulación de muestras y velocidad de análisis [17]. Por otra parte, cuando no se requiere disolución de la muestra, se evita el empleo de reactivos potencialmente peligrosos, lo que claramente beneficia la salud de los operadores y el medio ambiente, reduciéndose el coste del tratamiento externo de los desechos químicos. Adicionalmente, la espectroscopia NIR y Raman, permiten el análisis de productos envasados a través de diferentes tipos de contenedores de muestras, como bolsas, viales o botellas de diversos materiales. Sin embargo, sólo es posible realizar este tipo de medidas si el material del envase es transparente o el envase tiene una ventana adecuada para la radiación. En este sentido presentan las siguientes ventajas: i) la medición directa sin manipulación de las muestras, ii) conservación de las muestras dentro de los viales después del análisis, lo que significa que pueden emplearse/consumirse o almacenarse y iii) ausencia de alteraciones por el operador o el entorno del laboratorio [15].

Por tanto, las técnicas espectroscópicas vibracionales, como el MIR, NIR y Raman, pueden ofrecer un análisis rápido, directo, con poca cantidad de muestra, sin el uso de reactivos ni disolventes, siendo respetuosa con el medio ambiente y proporcionando seguridad al analista. Sin embargo, para la determinación de los componentes de una muestra mediante estas técnicas no invasivas, es necesario el uso de la Quimiometría, que permite el tratamiento de los datos

proporcionando información química a partir de los datos espectrales y reduce la necesidad de calibraciones externas y procedimientos específicos para la determinación de cada una de las propiedades o componentes de una muestra, una vez establecidos los modelos multivariantes. Por lo tanto, la Quimiometría puede considerarse una estrategia “verde” que puede emplearse para mejorar los métodos de detección no invasiva o remota con el fin de obtener información precisa y exacta a partir de señales directas, evitando una gran cantidad de reactivos, energía y mano de obra [18].

2.2. Espectroscopia vibracional: Fundamentos e Instrumentación

La espectroscopia se basa en las interacciones de la materia con la radiación electromagnética, la cual se clasifica en varias regiones según su frecuencia, abarcando una enorme gama de longitudes de onda y por tanto, de energías (Figura 4). La radiación electromagnética se representa convenientemente como campos eléctricos y magnéticos que experimentan en fase oscilaciones sinusoidales en ángulos rectos entre sí y respecto a la dirección de propagación. En la Figura 4 se representa, en dos dimensiones, el componente eléctrico del haz, donde la *amplitud* (A) de la onda es la longitud del vector del campo eléctrico en el máximo de la onda, mientras que la *longitud de onda* (λ) es la distancia entre máximos sucesivos. El *periodo* (p) de la radiación es el tiempo en segundos que se requiere para el paso de máximo o mínimos sucesivos por un punto fijo en el espacio y la *frecuencia* (ν) es el número de oscilaciones del campo que ocurren por segundo, y es igual a $1/p$ [19].

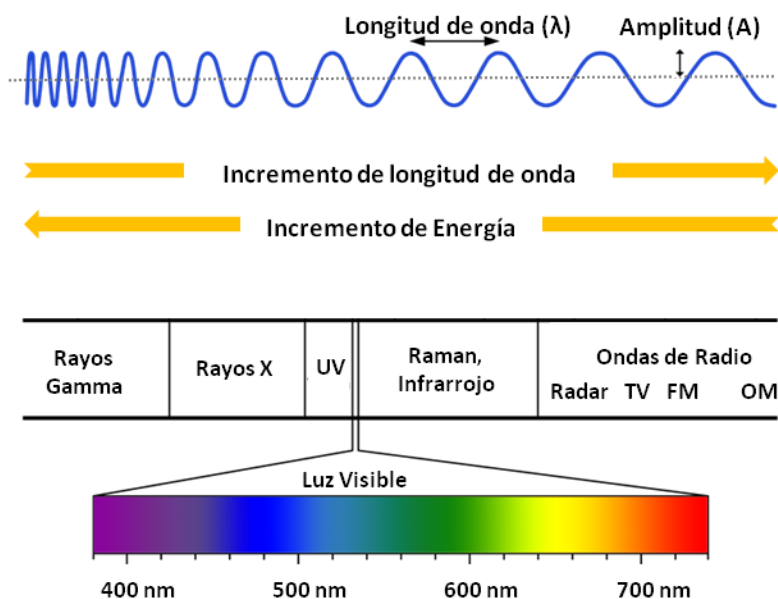


Figura 4. Espectro electromagnético.

El *número de onda* ($\bar{\nu}$) es otra forma de describir la radiación electromagnética, siendo este término más utilizado en espectroscopia NIR y Raman. El número de onda ($\bar{\nu}$), la longitud de onda (λ) y la frecuencia (ν) están relacionados por la *Ecuación 2* [20, 21]:

$$\bar{\nu} = 1/\lambda = \nu/c \quad \text{Ecuación 2}$$

donde c es la velocidad de la luz en el vacío.

La energía (E) que transportan las ondas electromagnéticas viene dada por la siguiente ecuación:

$$E = h\nu \quad \text{Ecuación 3}$$

siendo h la constante de Planck, con un valor de 6.6254×10^{-34} joule segundo, y ν la frecuencia, por lo que la diferencia entre las distintas espectroscopias que existen radica en el tipo de procesos a los que pueden dar lugar la energía de la radiación electromagnética.

Los métodos espectrométricos son aquellos métodos capaces de medir la intensidad de la radiación y están basados en la espectroscopia atómica y molecular. En la Tabla 1 se pueden observar los métodos espectrométricos clasificados según la región de la radiación electromagnética y el tipo de transición cuántica que se produce [19].

Tabla 1. Métodos espectroscópicos basados en la radiación electromagnética.

Tipo de espectroscopia	Intervalo de longitud de onda	Intervalo de número de onda (cm^{-1})	Tipo de transición cuántica
Emisión de Rayos Gamma	0.005 – 1.4 Å	-	Nuclear
Absorción, emisión, fluorescencia y difracción de Rayos X	0.1 – 100 Å	-	Electrón interno
Absorción ultravioleta en el vacío	10 - 180 nm	1×10^6 a 5×10^4	Electrones de enlace
Absorción, emisión y fluorescencia ultravioleta-visible	180 – 780 nm	5×10^4 a 1.3×10^4	Electrones de enlace
Absorción infrarroja y dispersión Raman	0.78 – 300 μm	1.3×10^4 a 3.3×10^2	Rotación/Vibración de moléculas
Absorción de microondas	0.75 – 375 mm	13 – 0.03	Rotación de moléculas
Resonancia de espín electrónico	3 cm	0.33	Espín de electrones en un campo magnético
Resonancia magnética nuclear	0.6 – 10 cm	1.7×10^{-2} a 1×10^3	Espín de núcleos en un campo magnético

Nota: $1 \text{ Å} = 10^{-8} \text{ cm}$; $1 \text{ nm} = 10^{-7} \text{ cm}$; $1 \mu\text{m} = 10^{-4} \text{ cm}$

La espectroscopia infrarroja (IR) y la espectroscopia Raman son metodologías basadas en vibraciones a nivel molecular que, aunque hay similitudes entre los espectros Raman y los espectros en el infrarrojo, hay suficientes diferencias entre las clases de grupos que son activos en uno u otro para hacer que las técnicas sean complementarias y no competitivas. Para obtener un espectro IR es necesario un cambio en el momento dipolar de la molécula durante la vibración [21], mientras que en el caso de la espectroscopia Raman es necesario un cambio en la polarizabilidad de la molécula [22].

2.2.1. Espectroscopia infrarroja

La espectroscopia de absorción IR se asocia a excitaciones vibracionales y rotacionales en moléculas, sin producirse transiciones electrónicas ya que la energía de radiación es insuficiente. Por tanto, los espectros infrarrojos se originan a partir de transiciones entre niveles vibracionales de la molécula en el estado electrónico básico (Figura 5).

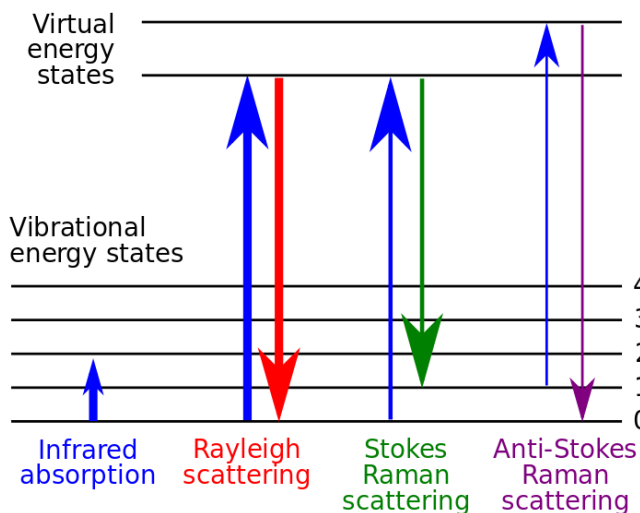


Figura 5. Diagrama de nivel de energía de los estados implicados en la señal IR y Raman.

Así mismo, tanto desde el punto de vista de las aplicaciones como de los instrumentos, el rango espectral IR se divide en tres regiones, infrarrojo lejano, medio y cercano, asociadas cada una de ellas a diferentes transiciones [23], siendo el IR medio y el cercano los más interesantes y utilizados.

- Región del IR lejano (FIR, del inglés *Far Infrared*), comprende entre 50 y 1000 μm , es decir de 200 a 10 cm^{-1} , produciéndose bandas de absorción a baja frecuencia, vibraciones esqueléticas, vibraciones de fonón cristalino y vibraciones intermoleculares.
- Región del IR medio (MIR, del inglés *Mid Infrared*), que abarca entre 2.5 y 50 μm , es decir entre 4000 y 200 cm^{-1} , dónde se producen vibraciones intramoleculares fundamentales.
- Región del IR cercano (NIR, del inglés *Near Infrared*), comprendida entre 0.7 y 2.5 μm , es decir entre 14300 y 4000 cm^{-1} , observándose vibraciones de armónicos y combinaciones de modos vibracionales fundamentales.

2.2.1.1. Instrumentación IR

Los primeros instrumentos IR utilizados eran en su mayoría de tipo dispersivo y usaban redes de difracción, hasta que, a principios de los años ochenta, se introdujeron los nuevos instrumentos de transformada de Fourier basados en el interferómetro de Michelson, aportando una mayor sensibilidad, una mejor relación señal/ruido, mejor precisión y exactitud y mayor velocidad de adquisición.

En los espectrómetros IR de transformada de Fourier (véase Figura 6) el haz de radiación infrarroja es generado en la fuente y después de su colimación se dirige hacia el interferómetro por medio de un espejo. El láser de He-Ne sigue a la radiación infrarroja a través del interferómetro con el fin de determinar el desplazamiento del espejo móvil y conocer así la longitud de onda a la que se produce la absorción de la radiación. Posteriormente, un espejo ajustable conduce el haz procedente del interferómetro hasta la muestra, llegando a través del compartimento de la muestra a un detector térmicamente estable. Un sistema de adquisición de datos registra la señal y la almacena en la memoria de una computadora en forma de interferograma. Un interferograma es básicamente la representación de la intensidad frente al tiempo, a partir del cual y mediante la operación matemática denominada transformada de Fourier, se obtiene un espectro en el que se representa la intensidad en función de la frecuencia [24]. Para obtener interferogramas satisfactorios y, por tanto, espectros satisfactorios, es necesario que la velocidad del espejo móvil sea constante y que su posición se conozca con exactitud en cualquier instante. También debe permanecer fijo el plano del espejo a lo largo de todo el recorrido.

Para obtener el espectro de una muestra con este tipo de instrumento, en primer lugar se obtiene un interferograma de fondo (o *background*) y luego es sometido a

la transformada de Fourier para obtener el espectro del *background*. A continuación se coloca la muestra en el accesorio de medida para obtener del mismo modo el espectro de la muestra. Sin embargo, este espectro contiene tanto información de la misma muestra como del *background*, por lo que el software del instrumento obtiene directamente el espectro de la muestra calculando el cociente entre la señal de la muestra y la del *background* [24].

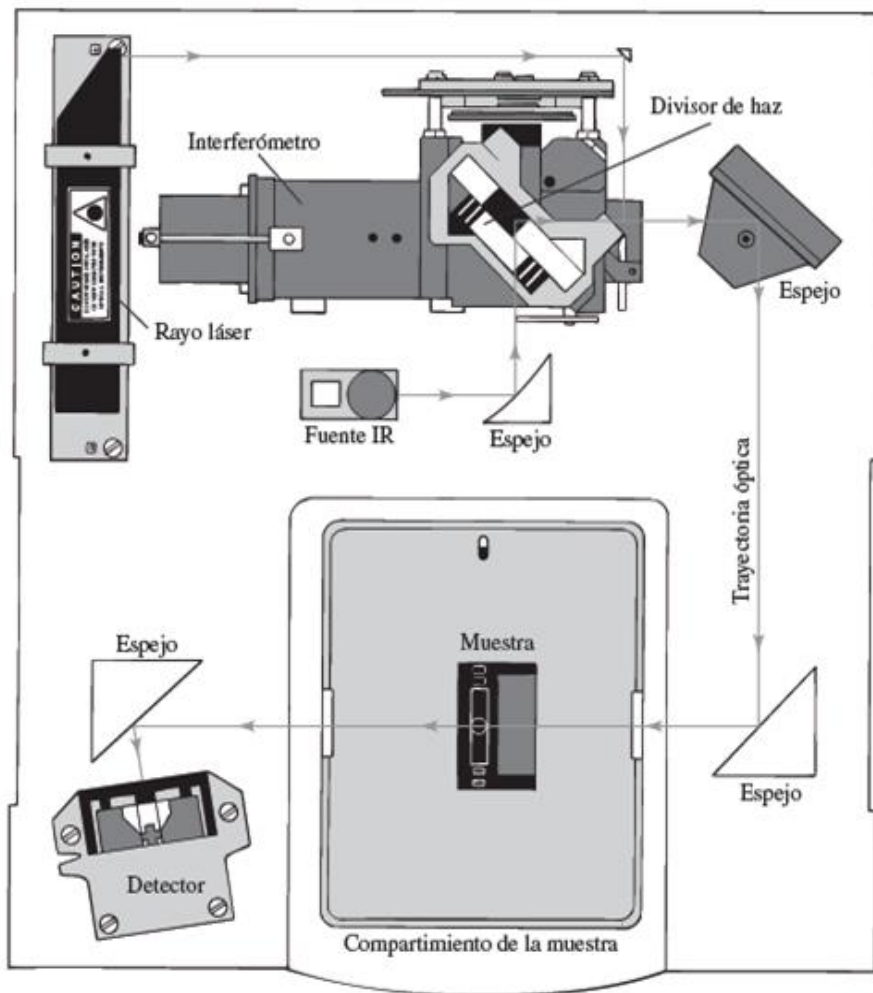


Figura 6. Diagrama de un espectrómetro básico de IR de transformada de Fourier

2.2.1.2. Espectroscopia infrarroja en la zona del IR medio (MIR)

Las fuentes de radiación más usadas en el MIR constan de un sólido inerte que se calienta eléctricamente a una temperatura comprendida entre 1500 K y 2200 K, y pueden ser de tipo globar o emisor de Nerst, filamento de nicromo además del arco de mercurio para el FIR, y la lámpara de filamento de tungsteno para el NIR o fuente láser de dióxido de carbono [19]. Además, existen tres tipos de detectores: los detectores térmicos, los detectores piroeléctricos y los detectores fotoconductores. Los dos primeros se encuentran por lo común en los fotómetros y en los espectrofotómetros dispersivos, mientras que los detectores fotoconductores se encuentran en los instrumentos multiplex o de transformada de Fourier.

El espectro MIR, el cual se encuentra en la región espectral comprendida entre 4000 y 200 cm^{-1} , puede dividirse aproximadamente en cuatro regiones [21]:

- Región comprendida entre 4000 y 2500 cm^{-1} , en la que aparecen las bandas de absorción características al estiramiento del enlace X-H.
- Región comprendida entre 2500 y 2000 cm^{-1} , denominada región del triple enlace, ya que es en ella donde aparecen sus bandas de absorción características.
- Región comprendida entre 2000 y 1500 cm^{-1} , considerada la región del doble enlace.
- Región comprendida entre 1500 y 600 cm^{-1} , o de la llamada **huella dactilar**.

Debido a la excitación de los grupos moleculares por el haz de radiación infrarroja los movimientos vibracionales que se generan en los enlaces de la molécula pueden ser de tensión o de flexión. Las tensiones provocan un estiramiento o estrechamiento del enlace y las flexiones generan una deformación del ángulo del enlace (Figura 7). En el caso de que un átomo forme más de un enlace, los movimientos también pueden presentarse como simétricos y antisimétricos.

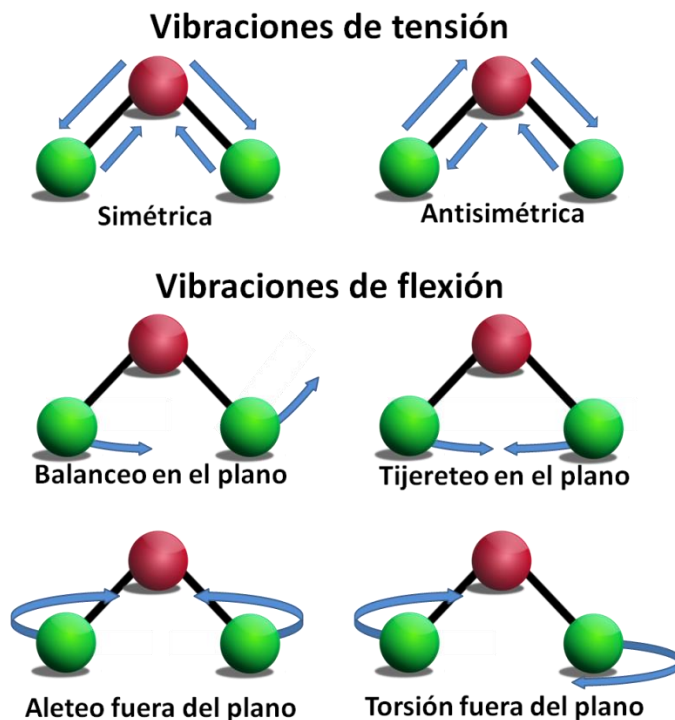


Figura 7. Modos de vibración de las moléculas.

2.2.1.3. Espectroscopia infrarroja en la zona del IR cercano (NIR)

La mayor parte de los espectrómetros NIR están equipados con lámparas de tungsteno-halógeno con ventanas de cuarzo. Los detectores varían desde fotoconductores de PbS y PbSe hasta fotodiodos de InSb e InAs; Además podemos encontrar detectores en serie, como los de InGaAs, para ser empleados en esta región.

Las absorciones observadas en la región del IR cercano, comprendido en la región espectral entre 14000 y 4000 cm^{-1} , están causadas por tres mecanismos diferentes:

- Sobretonos de las vibraciones fundamentales que ocurren en la región del infrarrojo medio ($4000\text{-}200\text{ cm}^{-1}$), producidos debido a que el comportamiento de las moléculas se acerca a un modelo de oscilador anarmónico en el que los niveles de energía no están homogéneamente espaciados, disminuyendo así la energía al aumentar la frecuencia fundamental de la vibración del enlace (ν).

$$\Delta E_{\text{vib}} = h \nu [1 - (2\nu + \Delta\nu + 1)y] \quad \text{Ecuación 4}$$

donde ν es la frecuencia fundamental de la vibración del enlace que genera una banda de absorción en el infrarrojo cercano, h es la constante de Planck e y es el factor de anarmonicidad. Las transiciones entre estados vibracionales no continuos dan lugar a las bandas de absorción denominadas sobretonos.

- Bandas de combinación de las vibraciones fundamentales, producidas en moléculas poliatómicas, donde dos o más modos vibracionales pueden interactuar, causando cambios de energía simultáneos y generando las bandas de absorción denominadas bandas de combinación.
- Absorciones electrónicas, causadas por el movimiento de los electrones desde una órbita a otra de mayor energía.

Estos mecanismos hacen que el espectro NIR sea tan complejo. La intensidad de las bandas NIR depende del cambio del momento dipolar y de la anarmonicidad del enlace.

2.2.1.4. Métodos de medida

Una de las grandes ventajas de la espectroscopia IR es su versatilidad, ya que permite estudiar prácticamente cualquier tipo de muestra ya sea líquida, sólida o gas y, dependiendo del estado físico de la muestra que se vaya a analizar, los espectros se pueden registrar mediante medidas de reflectancia o de transmisión.

- En las **medidas de transmisión** se establece la relación entre la potencia incidente y la transmitida, pasando el haz de radiación a través de la muestra antes de llegar al detector. Este método de medida se aplica a sustancias que no son opacas a la radiación IR.

Para la obtención de espectros en el MIR de muestras en fase gas mediante medidas en transmisión, se utiliza una celda de gases como la ilustrada esquemáticamente en la Figura 8. Este tipo de celda consiste en un cilindro de unos 45 mm de diámetro con dos conexiones que se puedan cerrar con llaves y que son resistentes al vacío, cerradas en los extremos con dos ventanas de un material transparente a la radiación IR.

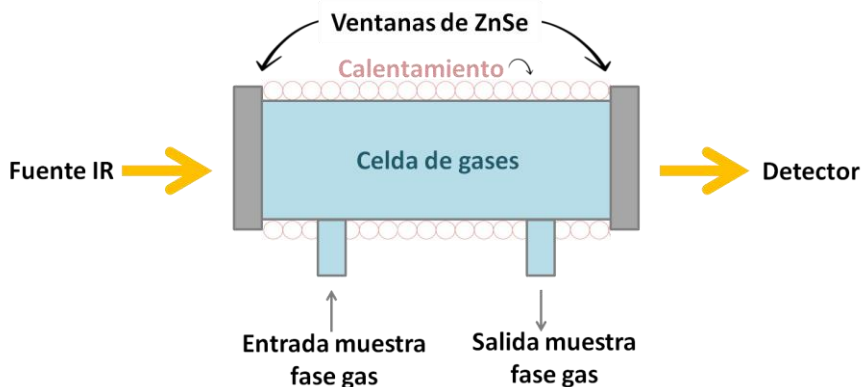


Figura 8. Esquema de una celda de gases para medidas de transmisión

Si la celda está en mal estado, las ventanas no están perfectamente paralelas o el grueso de las ventanas es del mismo orden que el paso óptico, estas medidas no van a poder ser practicables. Cuando hay que determinar trazas en gases poco absorbentes se usan celdas de multireflexión, que mediante un sistema de espejos permiten alcanzar caminos ópticos incluso de centenares de metros.

Para la obtención de espectros MIR de muestras líquidas, se utilizaban celdas de flujo como la ilustrada en la Figura 9, siendo los cristales de CaF_2 , ZnSe, KBr y KRS-5 (bromioduro de talio) los más usados [25]. Una vez cerradas pueden contener disolventes con puntos de ebullición por encima de 60°C , aunque hay que tener en cuenta que la muestra se calienta con el paso de la radiación y que el consiguiente aumento de presión puede traducirse en la evaporación parcial o completa de la muestra por fugas entre las ventanas y el espaciador. Normalmente se montan una vez y se reutilizan.

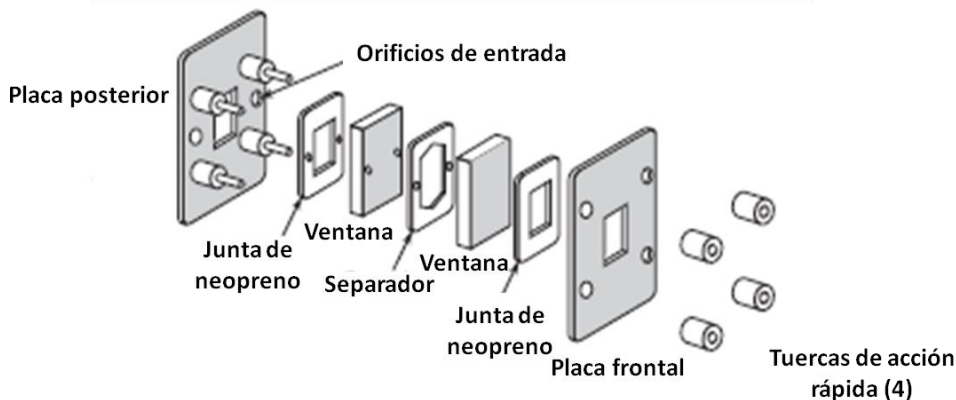


Figura 9. Representación gráfica de una celda de flujo

Una consideración importante en el uso de celdas es el tipo de material de las ventanas, ya que debe ser transparente a la radiación infrarroja incidente.

En la Tabla 2 se enumeran los distintos materiales más usados en las ventanas de las celdas para el infrarrojo medio [21].

Tabla 2. Materiales empleados en las ventanas de las celdas

Material de la ventana	Rango espectral (cm ⁻¹)	Índice de refracción	Propiedades
NaCl	40000 - 600	1.5	Soluble en agua, ligeramente soluble en alcohol, bajo coste, buena resistencia al choque mecánico y térmico, fácil pulido
KBr	43500 - 400	1.5	Soluble en agua y en alcohol, ligeramente soluble en éter, higroscópico, buena resistencia al choque mecánico y térmico
CaF ₂	77000 - 900	1.4	Insoluble en agua, resistente a ácidos y bases, útil para trabajos de alta presión
BaF ₂	66666 - 800	1.5	Insoluble en agua, soluble en ácidos y NH ₄ Cl, útil para trabajos de alta presión, sensible al choque mecánico y térmico
KCl	33000 - 400	1.5	Propiedades similares a NaCl pero menos soluble, higroscópico
CsBr	42000 - 250	1.7	Soluble en agua y ácidos, higroscópico
CsI	42000 - 20	1.7	Soluble en agua y alcohol, higroscópico

Las medidas de muestras líquidas mediante NIR se obtienen directamente utilizando viales de vidrio transparentes. En la Figura 10 se representa esquemáticamente la medida de transmisión de una muestra líquida a través de un vial.

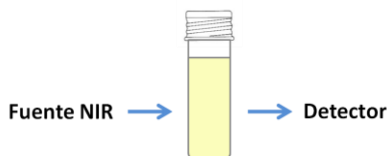


Figura 10. Ilustración esquemática del modo transmisión de la radiación IR a través de una muestra líquida situada en un vial de vidrio.

- Las **medidas de reflectancia** se pueden dividir en reflectancia total atenuada (ATR, del inglés *Attenuated Total Reflectance*) y en *reflectancia* difusa (DR, del inglés *Diffuse Reflectance*)

La técnica **ATR** fue desarrollada simultánea e independientemente por Harrick [26] y Fahrenfor [27] en los años 60. ATR es un tipo de espectroscopia de reflexión interna en la que la muestra se coloca en contacto con un elemento de reflexión interna (IRE) de alto índice de refracción. A continuación, la

radiación infrarroja se enfoca en el borde del IRE, se refleja a través del mismo, y luego se dirige a un detector adecuado [28]. En la Figura 11 se ha representado esquemáticamente una celda horizontal ATR. Su principal ventaja es que, al ser horizontal, una de las caras del IRE queda recubierta totalmente por la muestra. En el caso de muestras sólidas es necesario aplicar una presión para conseguir un buen contacto entre muestra y el IRE [29].

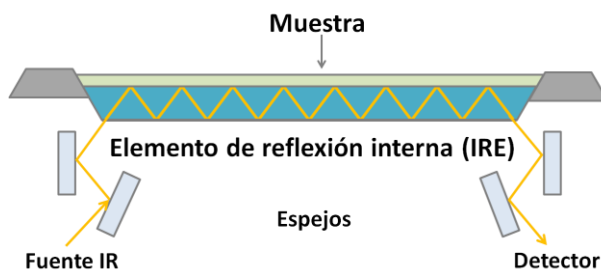


Figura 11. Esquema de una celda horizontal ATR

El número total de reflexiones (N) depende del ángulo de incidencia de la radiación, pero puede modificarse también según sea el grosor y la longitud del cristal. Un cristal más delgado y largo, proporcionará un mayor número de reflexiones según la Ecuación 5, aumentando así tanto la sensibilidad de la medida como la señal del disolvente [30]. Además se produce también un aumento en el ruido de fondo.

$$N = L / t \cot \alpha \quad \text{Ecuación 5}$$

donde L es la longitud del IRE, t el espesor y α el ángulo de incidencia.

Los cristales ATR tienen que tener un índice de refracción lo más alto posible (para que sea mayor que el de la mayoría de las muestras) y ser insolubles en el disolvente que se use (particularmente, en agua). Los materiales que más se emplean son el diamante, el seleniuro de zinc (ZnSe) y el germanio. El diamante es el material ideal ya que es resistente a las ralladuras, pH extremos y oxidantes y reductores, y a veces se usa en forma de película para proteger otro tipo de cristales. Otros posibles materiales son el telururo de cadmio (CdTe), el zafiro (Al_2O_3) o el dióxido de zirconio (ZrO_2).

El fenómeno de **reflectancia difusa** se produce, según Kubelka-Munk (K-M) [31], debido a la difusión de la radiación en todas direcciones después de incidir sobre las partículas de una muestra, como puede verse en la Figura 12, que

ilustra esquemáticamente la interacción de la radiación de luz con las partículas de una muestra y la generación de la reflectancia difusa.

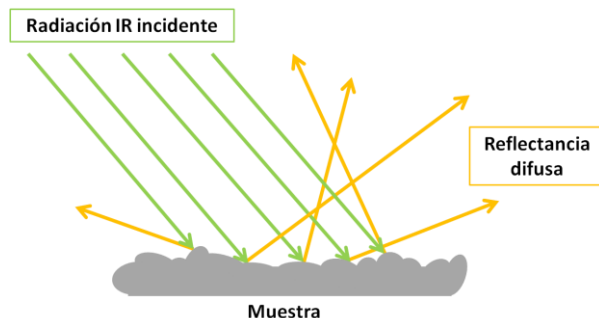


Figura 12. Ilustración esquemática de la interacción de la radiación IR con una muestra sólida para generar un espectro de reflectancia difusa.

Para obtener el espectro de una muestra mediante reflectancia difusa, en primer lugar se obtiene un *background* de la esfera integradora, la cual se caracteriza por las múltiples reflexiones de la radiación entrante en la misma. A continuación, se obtiene el espectro de la muestra ya sea mediante el uso de viales de cristal transparentes o mediante el uso de la sonda.

Dependiendo del tamaño de partícula, de los índices de refracción y de otros fenómenos físicos, la representación del espectro cambiará. Además de la reflectancia difusa, existe la reflectancia especular, que es la que se produce con ángulo igual al de incidencia de la radiación, que no aporta información sobre la composición de la muestra, por lo que solo contribuye al ruido.

Las **sondas de fibra óptica** transportan la radiación NIR directamente al contenedor de la muestra o a la muestra, pudiéndose realizar medidas rápidas e *in situ*, sin la necesidad de recoger las muestras y llevarlas al laboratorio. La luz reflejada o transmitida vuelve al detector para ser procesada.

2.2.2. Espectroscopia Raman

El fundamento de la técnica se basa en el efecto Raman, que consiste en la dispersión inelástica de una fracción de la luz incidente en la muestra, la cual sufre un cambio de frecuencia debido al intercambio de energía con la materia en dicha colisión, según la ecuación 3.

Al hacer incidir un haz de luz monocromático, de frecuencia ν_0 , sobre una muestra, éste se dispersa mayoritariamente de manera elástica (dispersión Rayleigh) y, también en una cantidad de luz dispersada inelásticamente. Mientras que la

radiación dispersada elásticamente tiene la misma frecuencia (longitud de onda) que la incidente y no aporta ninguna información sobre la composición de la muestra analizada, la radiación dispersada inelásticamente, llamada dispersión Raman, es la que proporciona información sobre la composición molecular de la muestra, habiendo dos tipos de dispersión (ver Figura 5):

- Dispersión Raman Stokes, dónde la radiación dispersada tiene menor energía que la incidente (la que tiene menor frecuencia).
- Dispersión Raman Anti-Stokes, la radiación dispersada tiene mayor energía que la incidente, es decir tiene una mayor frecuencia.

Microscópicamente, este efecto se puede describir como una excitación de la nube electrónica de la molécula debido a la interacción con el fotón incidente. Este provoca la excitación de la molécula a un estado excitado virtual, emitiendo otro fotón con el objetivo de regresar al estado fundamental, sin embargo, si el choque ha sido inelástico y se ha perdido o ganado energía, la molécula no volverá exactamente al mismo estado, sino que se encontrará en un estado rotacional o vibracional distinto del inicial. Según la frecuencia de los fotones dispersados, se pueden distinguir los casos representados gráficamente en la Figura 5.

Cabe destacar que las líneas Stokes se encuentran a longitudes de onda inferiores al pico Rayleigh y las líneas anti-Stokes a longitudes de onda superiores al de la fuente de excitación [22]. Además, estas últimas suelen ser menos intensas que las correspondientes líneas Stokes, siendo por ello la razón por la que normalmente se utilice la parte Stokes del espectro Raman (ver Figura 13).

Un espectro Raman está formado por una serie de bandas características, cada una de ellas asociada con un modo vibracional. Las frecuencias de vibración dependen de las masas de los átomos involucrados y de la fuerza de los enlaces. Los átomos pesados y los enlaces débiles presentan bajos desplazamientos Raman mientras que los átomos ligeros y los enlaces fuertes presentan altos desplazamientos Raman.

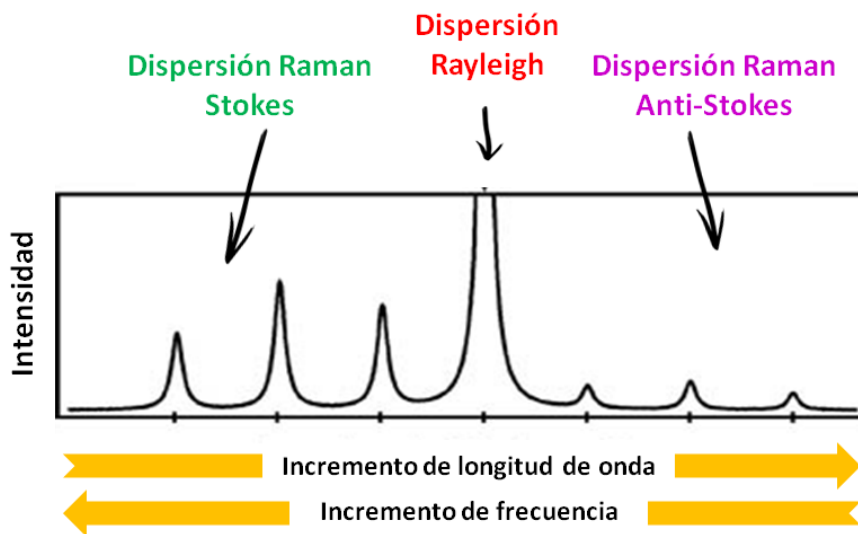


Figura 13. Representación esquemática de un espectro Raman

2.2.2.1. Instrumentación Raman

Los instrumentos para la espectroscopia Raman constan de tres componentes: una fuente láser, un sistema de iluminación de la muestra y un espectrómetro adecuado. La muestra se irradia con una fuente de rayos láser de radiación monocromática. La radiación dispersada a un cierto ángulo, se registra con la ayuda de un espectrómetro.

Las fuentes utilizadas en la espectrometría Raman son normalmente rayos laser debido a que su alta intensidad es necesaria para producir dispersión Raman lo suficientemente intensa como para poderse medir con una relación señal-ruido razonable. Los láseres que se utilizan en espectroscopia Raman se muestran en la Tabla 3 [19].

Tabla 3. Fuentes láser de la espectroscopia Raman

Tipo de láser	Longitud de onda (nm)
Ion argón	488 ó 514.5
Ion criptón	530.9 ó 647.1
Helio-Neón	632.8
De diodos	785 ó 830
Nd-YAG	1064

En cuanto al sistema de iluminación de muestra, la fuente láser se puede enfocar con facilidad sobre una pequeña zona de la muestra, enfocándose también la radiación emitida de manera eficaz sobre la rendija de entrada de un espectrómetro, pudiéndose analizar muestras muy pequeñas. En la espectroscopia Raman se puede usar vidrio para las ventanas, las lentes y otros componentes ópticos, en lugar de los haluros cristalinos que resultan más frágiles y menos estables en la atmósfera [19]. Además, la radiación visible o infrarroja cercana puede transmitirse a través de una distancia considerable por medio de fibra óptica. La mayoría de los espectrómetros Raman que se comercializan son instrumentos de transformada de Fourier que utilizan como transductores los detectores de germanio enfriados o instrumentos multicanal con dispositivos de acoplamiento de carga (CCD's).

En la actualidad, existe instrumentación analítica de altas prestaciones que se puede considerar portátil, con un reducido tamaño, que permite ser transportada hasta el lugar donde se encuentran las muestras y realizar los análisis directos, no destructivos, ni invasivos. Normalmente, estos equipos llevan incorporadas baterías que les proporcionan una gran autonomía sin necesidad de tener que hacer uso de fuentes externas de alimentación.

2.2.2.2. Métodos de medida

El espectro Raman recoge las bandas Raman Stokes y Anti-Stokes, situadas simétricamente a ambos lados de la banda Rayleigh (ver Figura 13), representando la intensidad óptica dispersada en función del número de onda ($\bar{\nu}$), que se expresa en cm^{-1} según la *Ecuación 2*. En la Figura 14 se representa esquemáticamente la medida de una muestra a través de las paredes de un vial y la generación de la dispersión simétrica y antisimétrica.

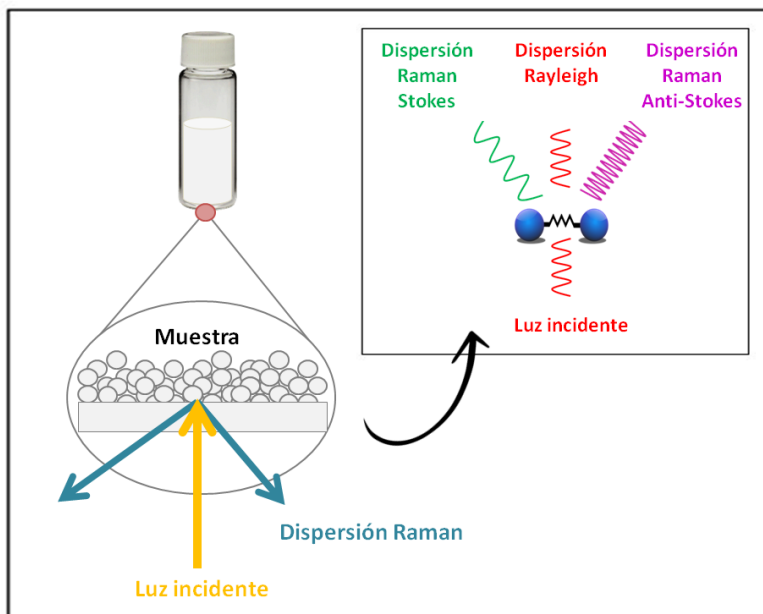


Figura 14. Ilustración esquemática de las medida Raman y las dispersiones producidas por la incidencia de luz sobre la muestra

Uno de los problemas que puede presentar la adquisición de espectros Raman es la presencia del ruido en la señal. Los ruidos más habituales se pueden clasificar en cinco grupos:

- Ruido *shot*, que es el resultado de la naturaleza aleatoria de la luz.
- Ruido generado por la muestra, que incluye emisiones ópticas no deseadas y generadas por la propia muestra, como es el caso de la fluorescencia. La fluorescencia es el fenómeno que se produce si al incidir un fotón sobre una molécula, éste es absorbido y la molécula pasa a un estado electrónico excitado donde permanece unas decenas de nanosegundos, para saltar a otro estado excitado pero de menor energía, liberando un fotón de frecuencia más baja que el incidente. La fluorescencia puede presentarse en el espectro Raman como una curvatura de la línea base, que puede llegar a enmascarar por completo la intensidad de las bandas Raman.

Además, el ruido generado por la muestra incluye los cambios de posición e intensidad de las bandas debido a que pueden modificarse en función de la temperatura de la muestra.

- Ruido generado por la instrumentación, en el que se incluye los ruidos introducidos por el detector, como los térmicos, el ruido de lectura o la dependencia de la eficacia cuántica del detector con la longitud de onda.
- Ruido computacional, referido al introducido en el proceso de digitalización de la señal de salida del detector.
- Ruido generado por fuentes externas, generalmente causado por alguna fuente de luz externa que contamina la señal en algún punto del equipo de medida, como por ejemplo la luz de los tubos fluorescentes o la luz solar. Las fuentes de ruido externo no-ópticos como los rayos cósmicos llegan al detector, y el resultado es un pico muy estrecho y de gran intensidad en el mismo espectro Raman. Es usual la presencia de ruido cósmico en detectores como las matrices CCD (dispositivos de carga acoplada).

En la Figura 15 se representa un espectro Raman con los diferentes tipos de ruido que se pueden encontrar.

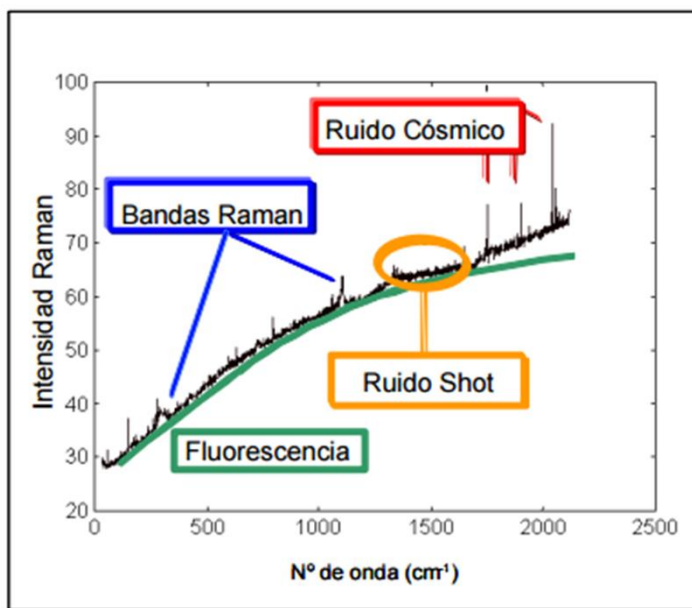


Figura 15. Causas de los ruidos más importantes en la espectroscopia Raman

2.2.3. Análisis cualitativo

La interacción de la radiación electromagnética con la materia provoca en la primera una alteración que guarda relación con los cambios en el estado vibracional de las moléculas. El espectro vibracional de una molécula se considera, por tanto, una propiedad física única, característica de la misma. Así, entre otras

aplicaciones, los espectros IR y Raman se pueden utilizar como “huella dactilar” en la identificación de muestras desconocidas mediante la comparación con espectros de referencia.

Debido a que se puede calcular la frecuencia (o número de onda) aproximada a la que un grupo funcional absorbe radiación en el infrarrojo a partir de las masas de los átomos y de la constante de fuerza del enlace entre ellos, se pueden utilizar gráficas/tablas de correlación para identificar los grupos funcionales presentes a partir de las señales de los espectros. Sin embargo, no es posible reconocer sin ambigüedad todas las bandas presentes en el espectro, por lo que estas gráficas/tablas de correlación sirven como punto de partida.

Para facilitar la identificación, existen colecciones y bibliotecas de espectros que ayudan a dicha identificación mediante una búsqueda y comparación de los espectros mediante el empleo del software incluido en los instrumentos.

2.2.4. Análisis cuantitativo

Absorción/Transmisión

El análisis cuantitativo se basa en la Ley de Beer-Lambert para establecer la relación entre las medidas de absorbancia y la concentración del analito en la muestra mediante la *Ecuación 6*:

$$A = \epsilon b c \quad \text{Ecuación 6}$$

dónde A es la absorbancia, ϵ el coeficiente de extinción molar a una determinada frecuencia o número de onda, b el paso óptico de la celda o camino recorrido por el haz de radiación al atravesar la muestra y c la concentración en la que se encuentra presente el analito.

Por otro lado, la transmitancia viene determinada por la *Ecuación 7*:

$$T = I/I_0 \quad \text{Ecuación 7}$$

dónde I es la cantidad de luz transmitida por la muestra e I_0 es la cantidad total de luz incidente [23], estando relacionadas absorbancia y transmitancia según la *Ecuación 8*:

$$A = \log_{10} T = \log_{10} I/I_0 \quad \text{Ecuación 8}$$

Por tanto, para poder llevar a cabo análisis cuantitativos es necesario que los pasos ópticos sean constantes y reproducibles. Este requisito es fácil de conseguir en el caso de análisis de líquidos, disoluciones y gases contenidos en el interior de celdas de medida con pasos ópticos constantes, pero no es tan sencillo en el caso de sólidos o películas delgadas.

ATR

Los espectros de ATR son similares, pero no idénticos a los espectros de absorción ordinarios. En general, se observan las bandas características a los mismos números de onda pero con distintas intensidades relativas. Las absorbancias dependen del ángulo de incidencia del haz de radiación, pero son independientes del espesor de la muestra debido a que la radiación sólo penetra unos pocos micrómetros en la muestra.

Reflectancia difusa

En la *Ecuación 8* se representa la ecuación de K-M [31], en la que se establece una relación lineal entre la intensidad y la concentración del analito en la muestra ya que la ley de Lambert-Beer, usada normalmente en medidas de transmisión, no es directamente aplicable para las medidas de reflectancia difusa. Las unidades de Kubelka-Munk eliminan cualquier efecto dependiente de la longitud de onda de la radiación incidente.

$$F = (1 - R)^2 / 2R = k / s \qquad \text{Ecuación 8}$$

dónde k es el coeficiente de absorción ($2.303 \cdot \epsilon \cdot c$), s el coeficiente de dispersión, R el índice de reflexión definido como espectro de muestra en polvo/espectro de dilución de polvo, de forma que si s es constante (propiedad intrínseca de un material y que depende solamente del tamaño de partícula) a una frecuencia determinada la función de Kubelka-Munk depende directamente de la concentración.

A diferencia de lo que ocurre con los espectros MIR o Raman, en los espectros NIR no aparecen bandas bien definidas y suele producirse un fuerte solapamiento de las señales. A pesar de la complejidad que puedan suponer los espectros, los métodos quimiométricos son una herramienta útil para la obtención de información analítica específica del mismo espectro, ya que es capaz de transformar las señales analíticas y datos complejos en información útil, tanto desde el punto de vista cualitativo como cuantitativo.

2.3. Quimiometría y Modelización

2.3.1. Generalidades

La Quimiometría puede describirse como la interacción entre métodos matemáticos y estadísticos con los procesos de medición química, transformando las señales analíticas en información.

Como se ha mencionado anteriormente, los espectros IR y Raman contienen información química y física de todos los componentes de la muestra, sin embargo debido a la complejidad de los mismos, en la mayoría de los casos no es posible diferenciar la información multivariable que contienen. La Quimiometría es capaz de extraer los parámetros de interés de los espectros a través de métodos matemáticos y estadísticos, relacionando las variables espectrales con las propiedades de los analitos y de las muestras, extrayendo información relevante y reduciendo la contribución de la irrelevante.

Los principales métodos quimiométricos de uso común son:

- Pre-tratamientos espectrales, para mejorar la información que interesa y disminuir la influencia de la información secundaria.
- Métodos de clasificación, para agrupar muestras según su espectro.
- Métodos de regresión, para relacionar el espectro con las propiedades cuantificables de las muestras.

Con el fin de facilitar el uso de los métodos quimiométricos se han diseñado varias herramientas estadísticas para obtener los resultados con mayor rapidez y sencillez. En esta Tesis se ha utilizado el software Matlab R2014a de Mathworks (Natick, MA, EEUU) junto con la herramienta PLS Toolbox 7.5.2 de Eigenvector Research Inc. (Wenatchee, WA, EEUU) para el tratamiento de los datos espectrales. En este capítulo se van a describir algunos métodos de calibración multivariable, haciendo especial énfasis en aquellos que han sido utilizados en esta memoria.

2.3.2. Técnicas multivariantes

Las técnicas multivariantes son un conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos compuestos por varias variables. Estas técnicas se pueden dividir en supervisadas o no supervisadas, que difieren entre sí en el sentido que las primeras intentan relacionar las variables con alguna propiedad de la muestra, mediante modelos que describen y explican fenómenos que están ocultos en el conjunto de datos y que pueden usarse para predecir el

valor de la propiedades, mientras que con los no supervisados no se requiere información previa sobre la propiedad, permitiendo obtener una información de tipo exploratorio sobre el conjunto de las muestras. Asimismo, los métodos supervisados pueden dividirse en métodos de clasificación y métodos de predicción. En la Figura 16 se muestra la clasificación de las técnicas quimiométricas utilizadas en este estudio.

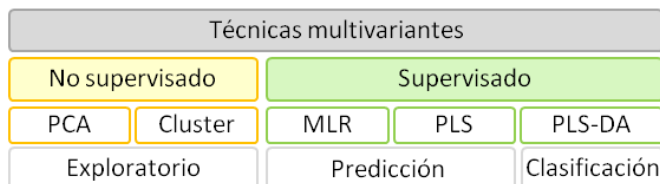


Figura 16. Técnicas quimiométricas utilizadas en este estudio

2.3.2.1. Técnicas No Supervisadas

El análisis de componentes principales (PCA, del inglés *Principal Component Analysis*) permite una reducción de variables y proporciona la proyección de datos en un nuevo espacio relacionado con el mínimo de variables independientes adecuadas para explicar una cantidad apropiada de varianza de datos, facilitando la exploración e interpretación de datos [32]. Los resultados del PCA se pueden representar mediante:

- Los *scores*, es decir, de los valores que toman las muestras en el nuevo espacio calculado siendo posible establecer relaciones entre ellas. Se les llama mapa de casos, cuando se representan en un diagrama los *scores* correspondientes a las muestras.
- Los *loadings*, indicando la importancia que tiene cada variable en cada componente. Se denominan, mapa de variables, cuando se representan los *loadings* de las muestras.

Estas representaciones permiten una interpretación eficaz del espacio de datos entero y su estudio es fundamental para la exploración de datos, aportando información sobre las dispersiones, agrupaciones, datos aberrantes y tendencias de los datos.

El análisis *cluster* difiere del PCA en su capacidad para detectar similitudes entre las muestras y definir grupos de muestras a partir de los datos de sus espectros. Las similitudes entre las muestras se estiman por medio de distancias: muestras similares se caracterizan por pequeñas distancias y muestras distintas por grandes distancias. La mejor interpretación gráfica del análisis *cluster* se representa

mediante un dendrograma, el cual ilustra la información en forma de diagrama de árbol. La división de las muestras se obtiene cortando el dendrograma en el nivel deseado de similitud [32].

2.3.2.2. Técnicas Supervisadas

La Regresión Lineal Múltiple (MLR, del inglés *Multiple Linear Regression*) intenta modelar la relación, entre dos o más variables explicativas y una variable de respuesta, ajustando una ecuación lineal a los datos observados. Para la construcción de un modelo MLR, el número de variables X debe ser menor que el número de muestras de calibración. Un problema típico en la calibración mediante el uso de la espectroscopia es el alto grado de correlación que pueden presentar las variables espectrales, ya que el uso de variables altamente colineales en MLR aumenta la posibilidad de sobredimensionar el modelo, lo que significa que la calibración puede encajar muy bien en los datos, pero falla cuando se usa para predecir las propiedades de nuevas muestras no empleadas para el conjunto de calibración. Por ello, un paso óptimo en la construcción del método MLR es la selección de variables que se incluirán en el modelo, ya que se podría evitar gran parte de la colinealidad [34]. Para este propósito, se suele utilizar una selección del subconjunto óptimo de variables por selección de las mismas paso a paso, antes de construir la calibración. Para este propósito existen distintas estrategias y una de ellas consiste en la *Forward Selection* (FS), en la que se comienza con un modelo de tamaño 0 y se procede agregando variables que cumplan un criterio definido. Típicamente, la variable a añadir en cada paso es aquella que minimiza los valores de la suma residual de cuadrados (RSS) de la calibración [35]. Otra alternativa es la *Backward Selection* (BS), en la que se comienza con el ajuste de un modelo con todas las variables de interés y se elimina la variable menos significativa. Los modelos se van reajustando hasta que todas las variables restantes sean estadísticamente significativas.

El método de Mínimos Cuadrados Parciales (PLS, del inglés *Partial Least Squares*) fue aplicado en 1983 por Martens y Jensen [36], para evaluar los espectros NIR. Actualmente se utiliza rutinariamente en el mundo académico y en la industria para correlacionar las medidas espectroscópicas (rápidas) con los datos químicos y/o físicos. El propósito de la regresión PLS es construir un modelo lineal que permita predecir una característica deseada (y) a partir de un espectro medido (x). Estos modelos de calibración se desarrollaron sobre la base de la modificación estadísticamente inspirada del algoritmo del método PLS (SIMPLS) [37].

El Análisis Discriminantes mediante mínimos cuadrados parciales (PLS-DA, del inglés *Partial Least Squares-Discriminant Analysis*) es un método de clasificación

lineal que combina las propiedades de la regresión de mínimos cuadrados parciales con el poder de discriminación de una técnica de clasificación como es el análisis discriminante [38]. El principal objetivo de PLS-DA es construir un modelo de calibración adecuado para ser aplicado en futuras clasificaciones. Se puede utilizar una serie de parámetros para evaluar el rendimiento de los modelos de clasificación como resultados verdaderos positivos (TP, del inglés *True Positive*) y verdaderos negativos (TN, del inglés *True Negative*), falsos positivos (FP, del inglés *False Positive*) y falsos negativos (FN, del inglés *False Negative*), sensibilidad, especificidad, precisión, eficiencia y coeficiente de correlación de Matthews (MCC). La tasa TP es la probabilidad de que una muestra positiva pueda clasificarse como positiva, TN implica que una muestra negativa puede clasificarse como negativa. Un criterio similar se emplea para determinar las tasas de FN y FP. La sensibilidad es la capacidad del modelo para clasificar correctamente las muestras auténticas dentro de su grupo usando la tasa TP y FN. La especificidad es la capacidad del modelo para identificar correctamente las muestras en relación con los valores de TN y FP. La precisión es la tasa de clasificación correcta, independientemente de la clase de muestra. La eficiencia y MCC resumen el desempeño del modelo. La eficiencia se determina como el promedio aritmético de los valores de sensibilidad y especificidad, donde un valor de 1 corresponde a una eficiencia del 100%. MCC da un valor entre -1 y +1, donde un valor de +1 representa la clasificación perfecta, 0 una clasificación errónea y -1 una clasificación inversa [39], [40].

2.3.3. Construcción, validación y evaluación del modelo quimiométrico

Para la construcción de un modelo quimiométrico es necesario realizar tres etapas para asegurar la robustez del mismo. En la Figura 17 se representan los pasos a seguir.

En primer lugar se debe de seleccionar un conjunto de muestras de calibración, que sea representativo de todas las muestras que constituyen la población, y otro conjunto de validación que permita poder evaluar la capacidad predictiva del modelo construido. Una vez seleccionados los conjuntos, se construye el modelo quimiométrico, teniendo en cuenta el pre-tratamiento de datos, la selección de la región espectral y de las variables latentes, junto con la eliminación de las muestras aberrantes. Para seleccionar el mejor modelo de calibración, es necesario evaluarlo mediante el empleo de un conjunto externo de validación o a través de una validación cruzada (CV, del inglés *cross validation*). La CV se emplea para comprobar el ajuste del modelo utilizando las muestras del conjunto de calibración y consiste en dividir en varios bloques el conjunto de calibración y

volver a construir el modelo tantas veces como bloques se hayan elegido, utilizando un bloque para comprobar los resultados y el resto para construir el modelo. Cuando el número de bloques es igual al número de muestras del conjunto de calibración, a esta CV se denomina *leave one out*.

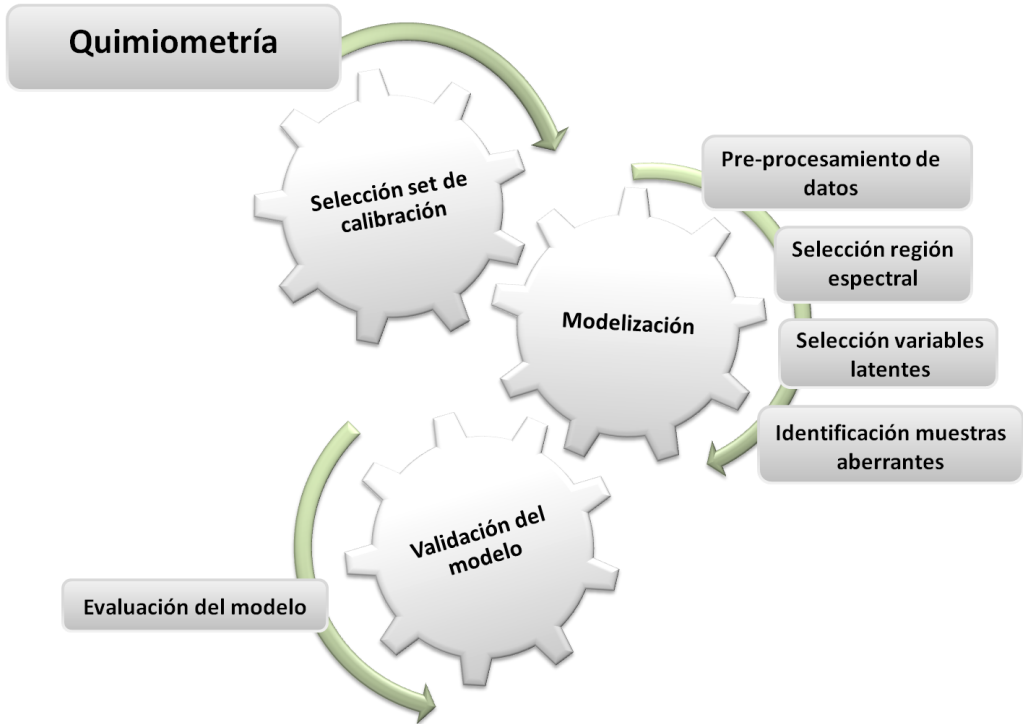


Figura 17. Etapas para la construcción y validación del modelo quimiométrico

2.3.3.1. Selección del set de calibración y validación

Debido a la complejidad de los espectros, para la construcción de modelos, se requiere utilizar un conjunto de muestras de referencia que tenga las mismas características que las muestras a analizar, esto es, que sean representativas del conjunto de la población objeto de estudio. Por ello, para construir los modelos, las muestras se dividen en dos conjuntos: uno para la calibración y el otro para la validación del modelo. Esta selección puede realizarse aplicando diferentes criterios o haciendo uso de algoritmos como pueda ser el de Kennard-Stone (K-S) [41]. Dicho algoritmo comienza con los dos objetos más lejanos como miembros del conjunto de calibración y va tomando los objetos basándose en la distancia euclidiana respecto de los objetos ya seleccionados. El conjunto de calibración debe ser representativo del total de muestras, por lo que además se deben

realizar otro tipo de estudio, como el análisis *cluster* o el PCA, para identificar los grupos que compartan características comunes y comprobar que los conjuntos de calibración seleccionados contienen muestras de los diferentes grupos en los que puede clasificarse la población. Además, el conjunto de calibración deberá comprender muestras uniformemente distribuidas preferiblemente en un rango mayor que el de las muestras que serán utilizadas para la evaluación y validación del modelo.

2.3.3.2. Pre-procesamiento de datos

Antes de utilizar cualquier técnica multivariante es necesario realizar una serie de tratamientos previos sobre la señal espectroscópica con el fin de corregir la línea base, la dispersión e incluso incrementar la relación señal/ruido. Los tratamientos más utilizados en el desarrollo de esta Tesis Doctoral son:

- Promediado de espectros
El espectro promedio de una muestra aumenta la relación señal/ruido ya que reduce la contribución del ruido.
- Derivadas
La primera y segunda derivadas se utilizan a menudo para el tratamiento de los datos ya que permiten acentuar las diferencias existentes en los datos espectrales. El uso de la primera derivada elimina los desplazamientos de línea base, mientras que la segunda derivada elimina el ruido de fondo lineal y constante. En nuestro caso se ha utilizado el método de Savitzky-Golay que incluye un suavizado previo a la derivada para disminuir el nivel de ruido [42].
- Variable normal estándar (SNV, *Standard Normal Variate*)
Elimina el efecto de dispersión (*scattering*) debido a los efectos producidos por el tamaño de partícula, aunque también permite reducir las diferencias en las intensidades globales de las señales [43].
- Corrección de Dispersión Multiplicativa (MSC, *Multiplicative Scattering Correction*)
Es un tratamiento que elimina los efectos debidos a la dispersión de la radiación [44], sin embargo el pre-tratamiento espectral depende del conjunto de muestras a la que se aplica, ya que considera el promedio de los espectros del conjunto de datos como espectro de referencia para corregir al resto el efecto de la dispersión.
- Centrado en la media (MC)

Todos los modelos se han realizado con centrado en la media, dónde se calcula el valor medio de cada variable de la matriz y se resta a cada punto de los espectros. El valor medio corresponde al centro del modelo quedando, por tanto, todos los valores de las variables referidos a este centro.

Estos tratamientos, por lo general, eliminan el ruido presente, minimizan las contribuciones de variables físicas que incorporan información irrelevante en los espectros y logran una mejor interpretación de los datos, permitiendo la construcción de modelos más simples y robustos.

2.3.3.3. Selección de la región espectral

Debido a que el uso del espectro completo no siempre proporciona resultados óptimos para la construcción de modelos, ya que en ocasiones incluye regiones sin información relevante o asociada a ruido, la selección de uno o más intervalos espectrales que contienen información química o estructural, así como la eliminación de rangos espectrales que sólo contienen ruido, son estrategias ampliamente utilizadas para mejorar los modelos de regresión PLS, PCA y *cluster* [45].

Como la selección de la región espectral es un paso crucial, en el presente trabajo se ha utilizado la estrategia de *Interval PLS* (iPLS) para la selección de conjuntos de variables contiguas [46]. iPLS es una extensión interactiva, que desarrolla modelos PLS locales sobre subintervalos equidistantes, definidos por el usuario, del espectro completo, proporcionando un cuadro general de la información relevante en diferentes subdivisiones espectrales, centrándose en aquellas regiones importantes y eliminando interferencias de otras regiones. El error cuadrático medio de la validación cruzada (RMSECV, *Root mean square error of cross validation*) (véase apartado 2.3.3.6) obtenido de cada modelo PLS local se compara con el RMSECV de un modelo global que utiliza todas las variables. Se considera que los intervalos que muestran un valor RMSECV inferior al modelo global contienen información útil y se seleccionan para construir el modelo PLS final.

2.3.3.4. Variables latentes

Los métodos basados en reducción de variables son métodos que permiten cuantificar un analito de una muestra sin la necesidad de conocer los otros componentes de la muestra. La información contenida en las variables medidas, se puede concentrar en un número menor de variables, denominadas variables latentes. La selección del número óptimo de variables latentes también es un paso crucial para obtener un modelo robusto, confiable e imparcial. Para ello, se utiliza

la validación cruzada (CV), de acuerdo con RMSECV. Un aumento de número de variables, aumenta la linealidad del modelo pero esto puede sobredimensionar el modelo (*over-fitting*) reduciendo su capacidad predictiva [47].

2.3.3.5. Muestras anómalas

Las muestras anómalas son aquellas que proporcionan señales discordantes con el resto del conjunto, por lo que la presencia de las mismas en la construcción del modelo puede llevar a modelos inadecuados y resultados erróneos, es por ello que resulta adecuado eliminarlas antes de la construcción del modelo.

En este trabajo se ha utilizado el gráfico del valor residual Q frente a los valores T² Hotteling para la identificación de muestras anómalas, ya que ambos parámetros, Q y T² Hotteling deben estar próximos a sus valores óptimos de 0% y 100%, respectivamente.

2.3.3.6. Evaluación del modelo

Una vez construido el modelo, se debe realizar la validación del mismo mediante el uso del set externo de validación, que no ha sido utilizado anteriormente para la calibración. Las figuras de merito utilizadas para la validación del modelo utilizadas en este trabajo son:

- Error cuadrático medio de la calibración (RMSEC, del inglés *Root Mean Square Error of Calibration*) es una medida para los modelos sobre el ajuste a los datos de calibración.
- Error cuadrático medio de la validación cruzada (RMSECV, del inglés *Root Mean Square Error of Cross Validation*), que corresponde con el que se obtiene cuando únicamente se tiene el conjunto de calibración y se realiza la validación cruzada aplicada a este conjunto.
- Error cuadrático medio de la predicción (RMSEP, del inglés *Root Mean Square Error of Prediction*), que es el que se obtiene cuando el modelo se aplica a un conjunto constituido por muestras diferentes de las incluidas en la calibración.

Estos errores vienen definidos por la *Ecuación 9*, y difieren en la determinación de \hat{y}_i .

$$RMSEC, RMSECV, RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad \text{Ecuación 9}$$

donde \hat{y}_i es el valor calculado por el modelo para las muestras del conjunto de calibración, validación cruzada o conjunto de validación, y_i es el valor conocido, n es el número total de objetos en el conjunto.

- Error relativo cuadrático medio (RRMSEP, del inglés *Relative Root Mean Square Error of Prediction*) viene dada por la Ecuación 10, y se utiliza para comparar la capacidad predictiva de distintos modelos.

$$RRMSEP = RMSEP / \bar{y} \times 100 \quad \text{Ecuación 10}$$

donde \bar{y} es el valor promedio de la variable de las muestras del conjunto externo de validación.

- Coeficiente de determinación de la predicción (R^2_{pred}) es un indicador de la bondad del ajuste del modelo aplicado al conjunto de validación. Este coeficiente determina qué porcentaje (en tanto por uno) de la varianza de la variable dependiente es explicado por el modelo de regresión.
- Desviación predictiva residual (RPD, del inglés *Residual Predictive Deviation*) calculado como el cociente entre RMSEP y la desviación estándar (SD, *standard deviation*) del conjunto de validación [48]. Generalmente, un valor RPD mayor que 3.0 se considera adecuado para propósitos analíticos con excelente capacidad de predicción, entre 2.5 y 3.0 implica que el modelo tiene una buena precisión y entre 2.0 y 2.5 indica que el modelo tiene una precisión aproximada. Un valor para la RPD entre 1.5 y 2.0 revela la posibilidad de distinguir entre valores altos y bajos, mientras que un RPD por debajo de 1.5 indica que la calibración no es utilizable [48]:

$$RPD = sd / RMSEP \quad \text{Ecuación 11}$$

donde sd es la desviación estándar de las variables del conjunto externo de validación.

Por tanto, los espectros IR y Raman junto con los métodos quimiométricos permiten extraer la información requerida sobre los componentes de interés de las muestras a partir del adecuado tratamiento de los datos, sin necesidad de su separación previa. Los métodos no supervisados, permiten visualizar las semejanzas y similitudes de las muestras sin necesidad de conocer previamente las características de las mismas, mientras que los métodos supervisados permiten realizar modelos, tanto de clasificación como de regresión, con la necesidad de conocer las características de las muestras y mediante el uso de un conjunto de calibración representativo de toda la población.

2.4. Referencias Introducción

- [1] P.T. Anastas, J.C. Warner, *Green Chemistry: Theory and Practice*, Oxford Univ. Press. New York, 1998; I. Horvath P. T. Anastas, *Chem. Rev.*, 2007, 107, 2167.
- [2] S.L.Y. Tang, R.L. Smith, M. Poliakoff, *Principles of green chemistry: PRODUCTIVELY,* *Green Chem.*, 7(2005) 761-762.
- [3] M. Tobiszewski, A. Mechlińska, B. Zygmunt, J. Namieśnik, *Green analytical chemistry in sample preparation for determination of trace organic pollutants*, *TrAC-Trends Anal. Chem.*, 28(2009) 943-951.
- [4] P. Tundo, P.T. Anastas, D. Black, T. Collins, S. Memoli, J. Miyamoto, M. Polyakoff, W. Tumas, *Pure Appl. Chem.*, 72(2000), 1207.
- [5] J. Namiesnik, *Trends in Environmental Analytics and Monitoring*, *Crit. Rev. Anal. Chem.*, 30(2010) 37–41.
- [6] S. Armenta, S. Garrigues, M. de la Guardia, *Green Analytical Chemistry*, *TrAC-Trends Anal. Chem.*, 27(2008) 497–511.
- [7] M. Koel, M. Kaljurand, *Application of the principles of Green Chemistry in Analytical Chemistry*, *Pure Appl. Chem.*, 78(2006) 1993–2002.
- [8] M. Koel, M. Kaljurand, *Green Analytical Chemistry*, The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 0WF, UK, 2010.
- [9] S. Garrigues, M. de la Guardia, *Challenges in Green Analytical Chemistry*, M. de la Guardia and S. Garrigues, Eds. The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 0WF, UK, 2011.
- [10] S. Garrigues, M. de la Guardia, *Handbook of Green Analytical Chemistry*, First Edition. Miguel de la Guardia and Salvador Garrigues, Eds. John Wiley & Sons, Ltd, 2012.
- [11] S. Armenta, M. de la Guardia, *Green Analytical Chemistry in Comprehensive Analytical Chemistry*, vol.57, D. Barceló, Ed., Elsevier, 2011.
- [12] A. Gałuszka, Z. Migaszewski, J. Namiesnik, *The 12 principles of green analytical chemistry and the SIGNIFICANCE mnemonic of green analytical practices*, *TrAC-Trends Anal. Chem.*, 50(2013) 78–84.
- [13] A. Gałuszka, P. Konieczka, Z. M. Migaszewski, *Analytical Eco-Scale for assessing the greenness of analytical procedures*, *TrAC-Trends Anal. Chem.*, 37(2012) 61–72.

-
- [14] S. Armenta, M. de la Guardia, J. Namiesnik, Green Microextraction. M. Valcárcel, S. Cárdenas, R. Lucena, Eds. Analytical Microextraction Techniques, Bentham Science, 2017.
- [15] M. de la Guardia, S. Garrigues, Direct Determination Methods Without Sample Preparation, in Challenges in Green Analytical Chemistry, M. de la Guardia, S. Garrigues, Eds. The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 0WF, UK, 2011.
- [16] M. de la Guardia, S. Garrigues, The Concept of Green Analytical Chemistry in Handbook of Green Analytical Chemistry, First Edition. Miguel de la Guardia and Salvador Garrigues, Eds. John Wiley & Sons, Ltd., 2012.
- [17] J. Moros, S. Garrigues, M. de la Guardia, Vibrational spectroscopy provides a green tool for multi-component analysis, TrAC-Trends Anal. Chem., 29(2010) 578-591.
- [18] M. de la Guardia, S. Garrigues, An Ethical Commitment and an Economic Opportunity, in Challenges in Green Analytical Chemistry, M. de la Guardia and S. Garrigues, Eds. The Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge CB4 0WF, UK, 2011.
- [19] D.A. Skoog, F. James Holler, S.R. Crouch, Introducción a los métodos espectrométricos in Principios de análisis instrumental, Sexta edic., Mexico, Cengage Learning, 2006.
- [20] B. Smith, Infrared Spectral Interpretation: A Systemetic Aproach. CRC Press, Florida, EEUU, 1999.
- [21] B.H. Stuart, Front Matter, in Infrared Spectroscopy: Fundamentals and Applications, John Wiley & Sons, Ltd, Chichester, UK, 2004
- [22] D.A. Skoog, F. James Holler, S. R. Crouch, Espectroscopía Raman, in Principios de análisis instrumental, Sexta edic., Mexico, Cengage Learning, 2006.
- [23] J.M. Chalmers, P.R. Griffiths, Handbook of Vibrational Spectroscopy, New York, John Wiley & Sons, 2002.
- [24] I. Sierra Alonso, S. Gómez Ruiz, D. Pérez Quintanilla, S. Morante Zarcero, Análisis Instrumental. La Coruña, España: Netbiblo, 2010.
- [25] D.A. Skoog, F. James Holler, S. R. Crouch, Aplicaciones de la espectrometría en el infrarrojo, in Principios de análisis instrumental, Sexta edic., Mexico, Cengage Learning, 2006.

- [26] N.J. Harrick, Surface chemistry from spectral analysis of totally internally reflected radiation, *J. Phys. Chem.*, vol. 64(1960) 1110-1114.
- [27] J. Fahrenfort, Attenuated Total Reflection: A new principle for the production of useful infrared reflection spectra of organic compound, *Spectrochim. Acta*, 17(1961) 698–709.
- [28] A.R. Hind, S.K. Bhargava, A. Mckinnon, At the solid/liquid interface: FTIR/ATR-the tool of choice, *Adv. Colloid Interface Sci.*, 93(2001) 91–114.
- [29] B. Smith, *Fundamentals of Fourier transform infrared spectroscopy*, Segunda ed. CRC Press, Taylor & Francis Group, 1996.
- [30] A. Smith, *Applied infrared spectroscopy: Fundamentals, techniques and analytical problem-solving*, 54(1979).
- [31] P. Kubelka, F. Munk, Ein Beitrag zur Optik der Farbanstriche., *Z. Technol. Phys.*, 12(1931) 593-620.
- [32] D. Ballabio, A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure, *Chemom. Intell. Lab. Syst.*, 149(2015) 1-9.
- [33] J.H. Ward, Hierarchical grouping to optimise an objective function, *J. Am. Stat. Assoc.*, 58(1963) 236–244.
- [34] S.N Jha, *Nondestructive Evaluation of Food Quality: Theory and Practice*, Springer Science & Business Media, 297(2010).
- [35] M. Cassoti, F. Grisoni, Variable selection methods: an introduction, In: www.moleculardescriptors.eu. Milano, Italia.
- [36] H. Martens, S. A. Jensen, Partial Least Squares Regression: A New Two Stage NIR Calibration Method, *Prog. Cereal Chem. Technol. J. Holas J. Krat. Eds.* Elsevier, Amsterdam, 1983, 607-647.
- [37] S. de Jong, SIMPLS: An alternative approach to partial least squares regression, *Chemom. Intell. Lab. Syst.*, 3 (1993) 251–263.
- [38] M.J. Hidalgo, D.C. Fechner, E. J. Marchevsky, R.G. Pellerano, Determining the geographical origin of *Sechium edule* fruits by multielement analysis and advanced chemometric techniques, *Food Chem.*, 210(2016) 228–234.
- [39] M.R. Almeida, C.H. Fidelis, L.E. Barata, R.J Poppi, Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation, *Talanta*, 117(2013) 305–311.

-
- [40] F. Bachion de Santana, L. Caixeta Gontijo, H. Mitsutake, S. J. Mazivila, L. M. de Souza, W. Borges Neto, Non-destructive fraud detection in rosehip oil by MIR spectroscopy and chemometrics, *Food Chem.*, 209(2016) 228–233.
- [41] R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Technometrics*, 11(1969) 137–148.
- [42] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedure, *Anal. Chem.*, 36(1964) 1627-1639.
- [43] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Appl. Spectrosc.*, 43(1989) 772- 777.
- [44] M.S. Dhanoa, S.J. Lister, R. Sanderson, R.J. Barnes, The Link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR spectra, *J. of Near Infrared Spectroscopy*, 2(1994) 43-47.
- [45] V. Centner, D.L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, Elimination of Uninformative Variables for Multivariate Calibration, *Anal. Chem*, 68(1996) 3851–3858.
- [46] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy, *Appl. Spectrosc.*, 54(2000) 413–419.
- [47] J.S. Shenk, M.O. Westerhaus, Routine Operation, Calibration, Development and Network System Management Manual. NIRSystems, Inc., 12101 Tech Road, Silver Spring MD 20904, EEUU; 1995.
- [48] P.C. Williams, D. Sobering, How do we do it: A brief summary of the methods we use in developing near infrared calibrations, A.M.C. Daves, P.C. Williams (Eds.), *Near infrared spectroscopy: The future waves*, NIR Publications, Chichester, Reino Unido (1995) 185–188.

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN

En esta sección se presentan los resultados obtenidos a partir de los trabajos realizados, dividiéndose en 3 partes de acuerdo con el ámbito de aplicación.

▪ **Parte 1. Muestras de alimentos.**

Centrada en la aplicación de la espectroscopia NIR, MIR y Raman para la determinación de los componentes mayoritarios presentes en muestras de alimentos, la determinación de parámetros nutricionales, así como el control de calidad en aditivos alimentarios.

▪ **Parte 2. Muestras ambientales y de interés agrícola.**

Aplicada la espectroscopia IR al estudio de muestras ambientales e interés agrícola, mediante la evaluación de composts para su posterior uso en la agricultura, determinación de analitos volátiles en suelos y la corrección de interferencias en fase vapor en espectroscopia infrarroja en fase gas.

▪ **Parte 3. Muestras arqueológicas.**

Aplicación del NIR al estudio de muestras arqueológicas, ya sea para la detección de procesos diagenéticos que alteran las propiedades de los huesos enterrados, predicción de elementos alcalinotérreos en restos óseos, así como la clasificación de huesos quemados con fines forenses.

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN

PARTE 1: Muestras de Alimentos

Parte 1. Muestras de alimentos

El empleo de la espectroscopia vibracional está muy extendido en el análisis de alimentos debido a su versatilidad, a la posibilidad de realizar análisis directos o con una mínima preparación de muestra, obteniendo, de forma rápida, información multicomponente a partir de los espectros, ya sea de modo cualitativo o cuantitativo. El análisis directo de las muestras es probablemente la forma más rápida y ecológica de realizar una determinación analítica ya que se puede evitar la parte principal de los pasos lentos y costosos, en términos de tiempo y manejo, de la manipulación de la muestra, así como prescindir del uso de reactivos y disolventes.

La creciente demanda de métodos analíticos más rápidos, más rentables y más respetuosos con el medio ambiente es un importante incentivo para mejorar los procedimientos clásicos utilizados para el análisis de alimentos. Muchas veces la preparación de la muestra es una necesidad, ya sea porque se necesite una calcinación, una digestión húmeda o una extracción de Soxhlet, siendo el paso más crucial de la mayoría de las metodologías analíticas, y también, la principal fuente de error en el proceso analítico.

La espectroscopia NIR, MIR y Raman proporcionan herramientas rápidas, no destructivas y, en muchos casos, libres de reactivos para el análisis multicomponente, siendo una gran alternativa verde a los métodos tradicionales, que requieren el uso de reactivos y de gran cantidad de tiempo. Además, los instrumentos portátiles, pueden proporcionar herramientas en línea, en tiempo real, optimizando la producción y garantizando la calidad y homogeneidad de los productos, lo que permite medidas de corrección oportunas durante la fabricación.

**3.1. CAPÍTULO 1:
Determinación directa de componentes
mayoritarios en alimentos**

3.1. Capítulo 1. Determinación directa de componentes mayoritarios en alimentos

3.1.1. Introducción

La nutrición humana depende, en gran medida, del valor energético de los alimentos y de una proporción equilibrada entre proteínas, carbohidratos y grasas, además de los micronutrientes encontrados en los alimentos, como los minerales y las vitaminas.

La nutrición durante los dos primeros años de vida es fundamental para garantizar un desarrollo óptimo de los niños, tanto físico como mental. La ingesta de nutrientes es necesaria para el crecimiento y obtención de energía, así como para mantener la salud. Las grasas y los carbohidratos se caracterizan por tener un alto valor energético, mientras que las proteínas participan en la construcción de las principales partes del cuerpo. Por ello, una mala nutrición aumenta el riesgo de enfermedad en la edad adulta, mientras que los déficits nutricionales en edad temprana, son reflejados a largo plazo en el crecimiento y la salud de la persona [1]. Se recomienda una dieta equilibrada, ya que el desequilibrio energético entre las calorías ingeridas y consumidas puede producir problemas de sobrepeso, obesidad, hipertensión y enfermedades cardiovasculares, debido a una elevada ingesta de alimentos hipercalóricos y/o dietas con altos valores de grasa, sal y azúcares junto con valores bajos de vitaminas, minerales y otros micronutrientes.

El etiquetado nutricional informa a los consumidores de las cualidades de los alimentos de acuerdo con el *Codex Alimentarius* [2], ya que la composición cualitativa y cuantitativa de los alimentos es esencial para asegurar que el producto cumple con la legislación vigente así como con las expectativas de los clientes y nutricionistas. Es por ello, que los consumidores deben ser informados de los valores nutricionales en alimentos procesados, ya que durante los procesos de calentamiento y cocción de dichos alimentos la composición y características nutricionales pueden ser alteradas.

Los requisitos nutricionales varían según la edad, por lo que este estudio se aplicará, en primer lugar, al estudio de diferentes tipos de alimentos que pueden ser consumidos tanto por adultos como por niños, en diferentes situaciones. Es por ello que se van a considerar los siguientes alimentos elaborados: i) los tarritos para bebés, ii) los menús infantiles ofrecidos en empresas de comida rápida y iii) los menús ofrecidos en comedores colectivos.

Normalmente, para la determinación de la composición de los alimentos se emplean los Métodos Oficiales de Análisis (AOAC) [3]. Estos análisis consisten en

muchos de los casos en la aplicación de métodos gravimétricos y volumétricos, que se caracterizan por consumir gran cantidad de tiempo y reactivos.

Las técnicas espectroscópicas pueden ofrecer alternativas interesantes a los métodos tradicionales de análisis de alimentos [4]. En ese sentido, la espectroscopia infrarroja en la región cercana (NIR) ha sido utilizada para la evaluación de la calidad de los alimentos [5], de cereales [6,7], productos lácteos [8], carnes [9], pescados [10], productos vegetales [11], y de alimentos procesados [12-14]. En cuanto el infrarrojo en la región media (MIR) se ha aplicado en la determinación de parámetros nutricionales en muestras de pan [6], bebidas [15], vegetales [11], pescados [16], productos lácteos [17,18] y alimentos procesados [6,14]. Sin embargo, sólo en unos pocos trabajos se ha determinado los macronutrientes en mezclas de alimentos mediante espectroscopia NIR [19-21], mientras que la espectroscopia MIR no ha sido utilizada para la determinación de dichos parámetros en dietas y en mezclas de alimentos.

El objetivo del presente estudio es la evaluación de la espectroscopia NIR y MIR, junto con herramientas quimiométricas, para la determinación directa de grasas totales, proteínas, carbohidratos y valor energético en mezclas de alimentos como son los tarritos de alimentación para bebés, menús infantiles de locales de comida rápida y menús para adultos en comedores colectivos.

3.1.2. Parte experimental

3.1.2.1. Instrumentos

Los espectros en el infrarrojo cercano se obtuvieron en un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Germany) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS.

Los espectros del infrarrojo medio se registraron con un espectrofotómetro modelo Tensor 27 FT-MIR de Bruker (Bremen, Alemania) equipado con un detector de sulfato de trigliceno dopado con L-alanina deuterado (DLaTGS). Los espectros de las muestras se obtuvieron mediante medidas en modo ATR, utilizando el accesorio para sólidos de Smiths Detection Inc. (Warrington, UK) con un cristal DuraDisk de diamante/ZnSe de tres reflexiones (véase Figura 1.1).

El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

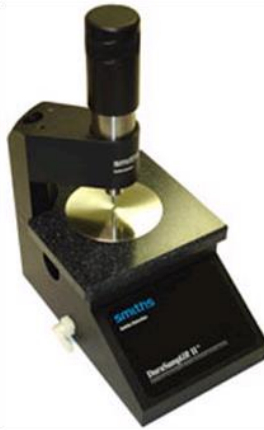


Figura 1.1. Accesorio con cristal de diamante/ZnSe de 3 reflexiones para muestras sólidas

3.1.2.2. Muestras

Se analizaron cincuenta y ocho muestras de alimentos elaborados con distintas combinaciones de carne, pescado, frutas, verduras, cereales, huevo, productos lácteos, salsas y aceites, las cuales se encontraban distribuidas en tres grupos: i) tarritos de comida para bebés (35 muestras), (ii) menús infantiles de locales de comida rápida (10 muestras), y (iii) menús completos para adultos de comedores universitarios (13 muestras).

Los tarritos para bebés fueron adquiridos en supermercados de la ciudad de Valencia, y estaban compuestos por agua empleada para la cocción, verduras y otros ingredientes, principalmente carne o pescado (29 muestras), y otros que contenían sólo frutas (6 muestras). Los menús infantiles de comida rápida fueron comprados en diferentes hamburgueserías de la ciudad de Valencia y estaban compuestos por hamburguesa de carne con queso y salsas, patatas fritas, un refresco y un postre lácteo. Los menús para adultos se compraron en las cantinas del campus de la Universidad de Valencia, y estaban formados por un entrante, un plato principal, y un postre, junto con una pieza de pan de 50-70-g.

Los alimentos fueron pesados, mezclados y triturados con un mezclador doméstico (Braun, Kronberg, Alemania), y posteriormente congelados a $-20\text{ }^{\circ}\text{C}$. Finalmente las muestras fueron liofilizadas durante un mínimo de 48 h en una cámara de 0.05 mbar usando un liofilizador Cryodos 50 de Telstar (Barcelona, España), y una vez deshidratadas, almacenadas en botellas de polietileno hasta su análisis.

3.1.2.3. Métodos de referencia

Para la determinación del contenido de grasa total se utilizó un analizador de grasa Soxtec System con una Unidad de Extracción HT2 1045 de Tecator (Burladingen, Alemania), equipado con un baño de aceite de silicona Soxtec System HT 1046, para el calentamiento. Las muestras fueron sometidas a hidrólisis ácida antes de realizarse la extracción con éter de petróleo, para obtener la grasa total (libre y unida). Una vez evaporado el disolvente, la fase lipídica fue determinada gravimétricamente [3].

El contenido de nitrógeno presente en las muestras fue determinado mediante el método Kjeldahl [3], utilizando un sistema automático compuesto por una unidad de digestión Büchi-426 y un analizador 2300 Kjeltec de Foss (Hillerod, Dinamarca). Para la determinación del contenido de proteínas, el nitrógeno encontrado en las muestras fue multiplicando por un factor de 6.25 [22].

Los carbohidratos totales fueron calculados como la diferencia entre 100 y la suma de los porcentajes de grasa total, proteína y ceniza [22]. La ceniza se determinó mediante calcinación de la muestra a 550 °C [3] utilizando un horno mufla de la marca Gallur (Valencia, España), con una temperatura máxima trabajo de 1100 °C.

El valor energético de las muestras se calculó según lo sugerido por el Departamento de Agricultura de los Estados Unidos (USDA) [23] mediante la ley de Etiquetado y Educación (NLEA) [24] para alimentos mixtos procesados. El valor energético sería la suma de las grasas totales, proteínas y carbohidratos totales multiplicados por factores de 37, 17 y 17 kJ por 100 g, respectivamente [25].

Todas las determinaciones se realizaron por duplicado.

3.1.3. Resultados y discusión

3.1.3.1. Espectroscopia NIR

Las muestras de los alimentos liofilizados se colocaron en viales de vidrio transparente de 22 mm de diámetro interno y 44 mm de altura para obtener directamente espectros NIR por reflectancia difusa, en unidades de Kubelka-Munk. Los espectros se adquirieron en la región espectral comprendida entre 14000 y 4000 cm^{-1} , con una resolución de 4 cm^{-1} , y acumulando 50 barridos por espectro. Cada una de las muestras se midió por triplicado, realizando rotaciones del vial entre medidas, y calculando, posteriormente, el promedio para fines quimiométricos. El espectro de fondo se adquirió a partir de la esfera de integración cerrada utilizando las mismas condiciones instrumentales que las empleadas para las muestras.

En la Figura 1.2 se puede observar el espectro NIR promedio, sin ningún tipo de pre-tratamiento, de todas las muestras medidas en este estudio. Los espectros NIR de las muestras de alimentos consideradas en este estudio, presentan bandas de absorbancia correspondientes a grupos funcionales asociados con lípidos, proteínas y carbohidratos. Las principales bandas corresponden a los sobretonos de estiramiento del enlace C-H de las grasas [14] y proteínas, junto con las bandas de combinación de los enlaces C-H, N-H y O-H de proteínas y carbohidratos [29,30]. Los espectros NIR obtenidos coinciden con la descripción disponible en la literatura para este tipo de muestras.

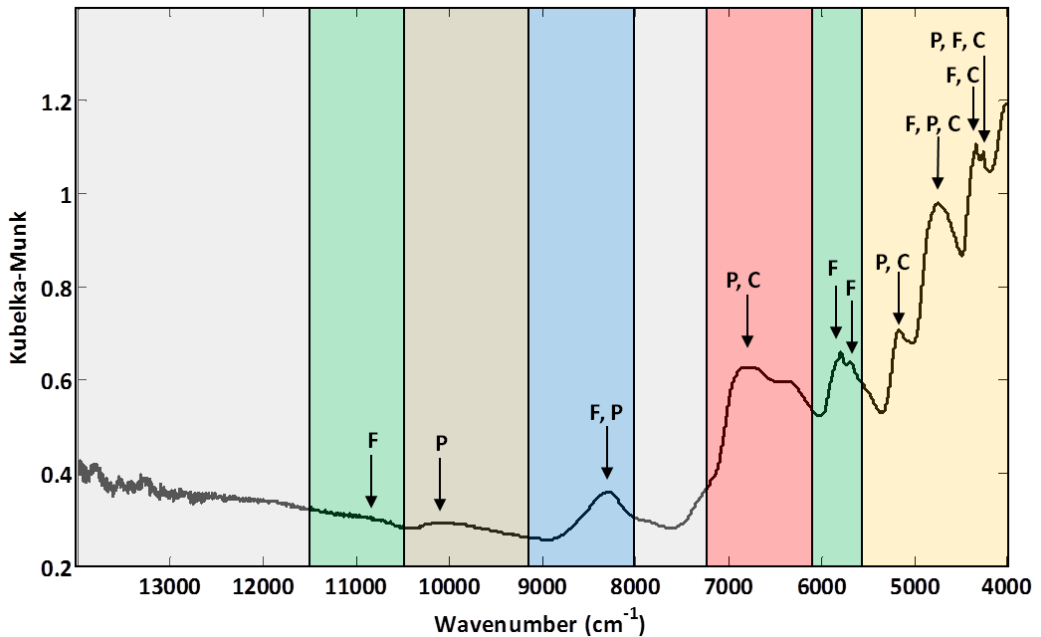


Figura 1.2. Espectro NIR promedio obtenido con reflectancia difusa en unidades de Kubelka-Munk
Nota: Regiones indicadas corresponden a bandas de grasa (F), proteínas (P) y carbohidratos (C)

3.1.3.1.1. Modelos PLS-NIR

El conjunto de datos compuesto por las 58 muestras de comida fue dividido mediante el algoritmo K-S para la selección de un conjunto de calibración (45 muestras) y un conjunto de validación (13 muestras). El conjunto de calibración quedó formado por 26 tarritos para bebés, 4 de ellos basados en frutas, 10 menús para adultos y 9 menús infantiles de comida rápida. Debido a que el conjunto de calibración debe ser representativo del conjunto de datos, se construyó un dendrograma a partir de los espectros de las muestras después de realizarle un pre-tratamiento de primera derivada (FD) (15 puntos) y centrado en la media

(MC), en la región comprendida entre 11000 y 4000 cm^{-1} , para la observación de similitudes y disimilitudes de las muestras.

En la Figura 1.3 se representa el dendrograma construido usando *Ward Linkage* como método de distancia, donde se puede observar que para un valor de corte mínimo de 0.055, todas las muestras se clasificaron en cuatro grupos principales, correspondientes a muestras con características comunes como los menús de comedores colectivos, comida rápida infantil y tarritos para bebé, considerando en este último las diferencias entre los alimentos a base de frutas y alimentos en general. El primer grupo está compuesto por los tarritos para bebés a base de frutas; el segundo por menús de adultos junto con una muestra de menú infantil de comida rápida. El tercer grupo está basado en los menús de comida rápida infantil, mientras que el último por los tarritos de bebés a base de carne, pescado y verduras.

Además, en esta misma figura, se pueden observar las muestras seleccionadas por el algoritmo de K-S para cada tipo de conjuntos, ya sea calibración o validación. Por lo tanto, a partir de dicha figura, se puede concluir que las muestras del conjunto de calibración son representativas del total de las muestras consideradas en este estudio.

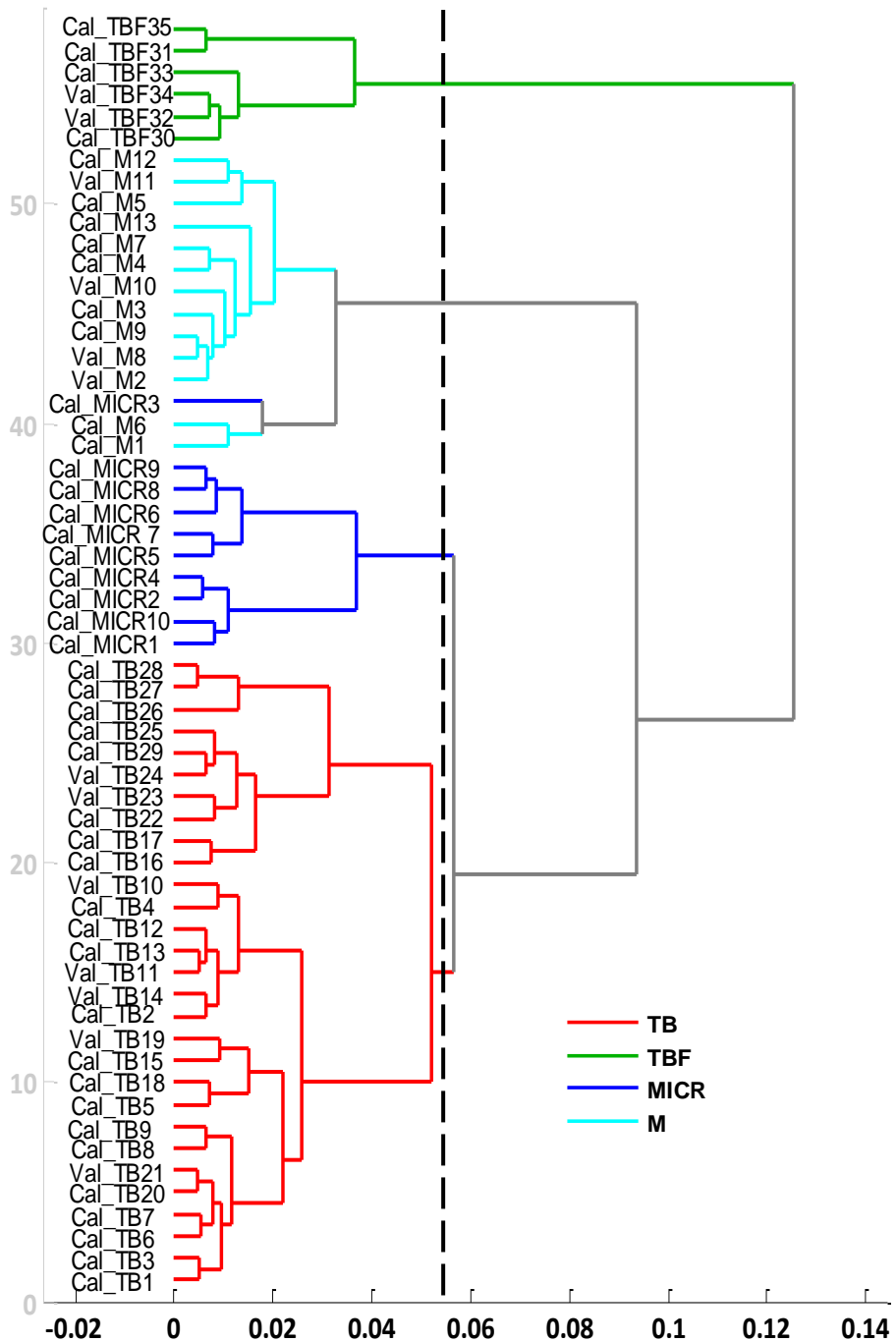


Figura 1.3. Clasificación dendrográfica, mediante el método *Ward Linkage*, de los espectros NIR de las muestras de tarritos para bebés (TB), tarritos para bebés de fruta (TBF), menús infantiles de comida rápida (MICR) y menús para adultos (M). Nota: Cal y Val pertenecen a los conjunto de calibración y validación, respectivamente, establecidos por el algoritmo K-S.

Línea discontinua: valor mínimo de corte de 0.055

En la Tabla 1.1 se muestran las principales características de los conjuntos de calibración y validación usados en los modelos PLS-NIR para los parámetros de grasa total, proteínas, carbohidratos y valor energético.

Tabla 1.1. Principales características de los conjuntos de calibración y validación empleados para la construcción de los modelos PLS-NIR

Analito	Conjunto	Muestras	Rango	Valor medio	SD
Grasa total (g por 100g)	Cal	45	0.05-18.6	10	5
Proteínas (g por 100g)		45	1.95-26.2	16	6
Carbohidratos (g por 100g)		45	55.5-94.9	70	10
Valor energético (KJ por 100g)		45	1660-2010	1850	90
Grasa total (g por 100g)	Val	11	0.07-15.1	10	5
Proteínas (g por 100g)		13	3.42-25.3	17	7
Carbohidratos (g por 100g)		11	59.2-93.7	70	10
Valor energético (KJ por 100g)		11	1660-2010	1900	90

Conjuntos de calibración (Cal) y validación (Val)
Rango: valor mínimo-máximo. SD desviación estándar

Una vez seleccionados los conjuntos de calibración y validación, se utilizó el algoritmo iPLS para seleccionar las regiones espectrales óptimas para cada uno de los parámetros a determinar. Para ello, el espectro fue dividido en intervalos de 207 variables, utilizando un máximo de 10 factores o variables latentes (LV's) para su selección. En la Figura 1.4 se pueden observar los espectros y las regiones seleccionadas mediante iPLS para cada uno de los cuatro parámetros a determinar.

Para el caso de la grasa total, las regiones seleccionadas fueron aquellas comprendidas en los intervalos 10803-10406, 6013-5217 y 4416-4019 cm^{-1} , correspondientes a las regiones de grasa. Para las proteínas, se seleccionó las bandas correspondientes a los intervalos 8807-8011, 6013-5616 y 5215-4817 cm^{-1} , pertenecientes al primer sobretono y a la banda de combinación de estiramiento

del enlace C-H. En cuanto las regiones seleccionadas para la determinación de los carbohidratos fueron aquellas comprendidas en los intervalos 7210-8011, 6013-5616 y 4816-4019 cm^{-1} , que corresponde con la banda de combinación del enlace N-H y la banda de combinación del estiramiento del enlace O-H y C-H. Para la determinación del valor energético, las regiones seleccionadas fueron aquellas comprendidas entre 10804-10406, 5217-4817 y 4416-4019 cm^{-1} , algunas de las cuales corresponden a las regiones seleccionadas para grasa total, proteínas y carbohidratos, lo que tiene sentido ya que el contenido de carbohidratos ha sido calculado de forma indirecta a partir de estos 3 parámetros.

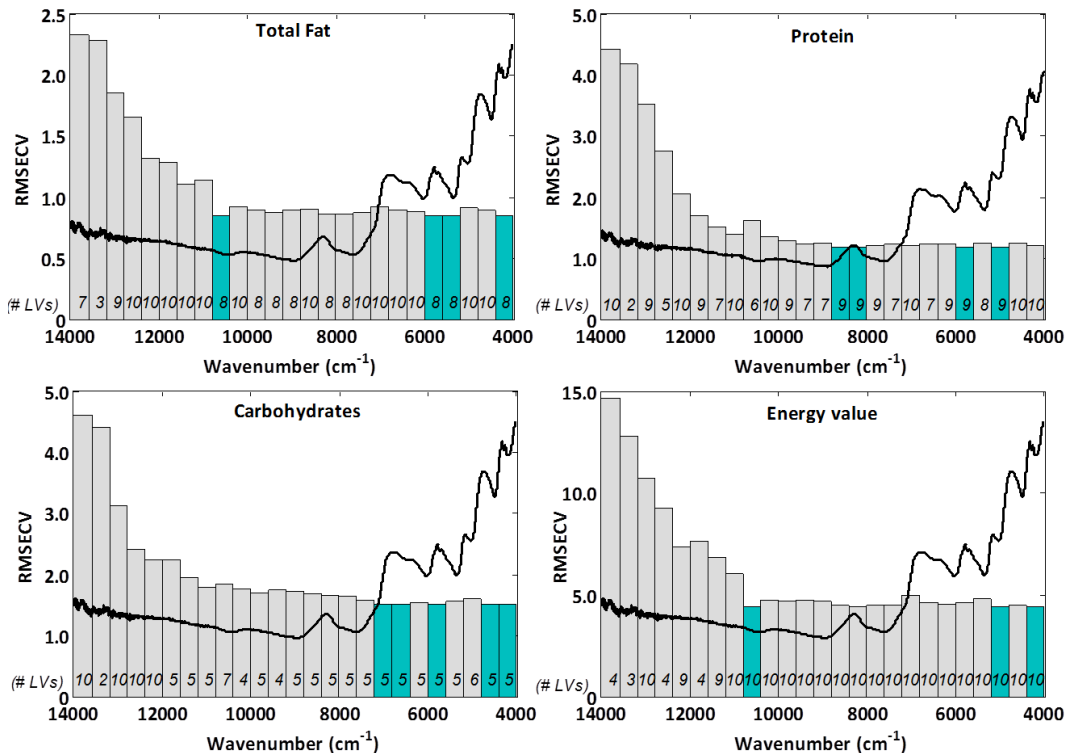


Figura 1.4. Espectro NIR con la selección de las regiones mediante iPLS para la determinación de grasa total, proteínas, carbohidratos y valor energético mediante la construcción de los modelos PLS-NIR. Nota: los valores en cursiva corresponden al número de variables latentes (LV)

Los modelos PLS-NIR se construyeron utilizando las regiones seleccionadas mediante el algoritmo iPLS y usando 6, 7, 4 y 7 LV's para la determinación de grasa total, proteínas, carbohidratos y valor energético, respectivamente. Para el caso de grasa total, proteínas y valor energético, se utilizó la primera derivada (FD) con un suavizado de 9 puntos y centrado en la media (MC) como pre-procesamiento de señal. En cambio, para los carbohidratos se utilizó, SNV, FD y MC. En la Tabla

1.2 se indican las LV's, los errores cuadráticos medios de calibración (RMSEC), validación cruzada (RMSECV), y los coeficientes de determinación (R^2) de calibración y validación cruzada de las muestras correspondientes al conjunto de calibración.

Tabla 1.2. Descripción de los modelos PLS-NIR empleados para la determinación de grasas totales, proteínas, carbohidratos y valor energético

	Pre-procesamiento	LV's	RMSEC	R^2 Cal	RMSECV	R^2 CV	RMSEP	RRMSEP (%)	R^2 Pred
Grasa total (g por 100g)	MC, FD	6	0.4	99.1	1.0	95.9	0.9	9.2	96.7
Proteínas (g por 100g)	MC, FD, SNV	7	0.6	99.0	1.2	95.8	1.1	6.1	98.1
Carbohidratos (g por 100g)	MC, FD	4	1.3	98.2	1.6	97.4	1.5	2.1	98.9
Valor energético (KJ por 100g)	MC, FD	7	2.0	99.3	5.1	95.2	4.6	1.0	96.5

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R^2 Cal: coeficiente de determinación de calibración; R^2 CV: coeficiente de determinación de validación cruzada; R^2 Pred: coeficiente de determinación de predicción. MC: centrado en la media; FD: primera derivada; SNV: Variación Normal Estándar

3.1.3.1.2. Capacidad de predicción del modelo PLS-NIR

Con el fin de confirmar que el modelo de calibración construido era adecuado para la predicción de nuevas muestras, se utilizó un conjunto de validación externo, compuesto por muestras no incluidas en el conjunto de calibración.

En la Figura 1.5 se representan los valores predichos de cada uno de los analitos estudiados frente sus valores de referencia. En el caso de grasas totales, carbohidratos y valor energético se eliminaron 2 muestras anómalas (1 menú y un tarrito para bebé) del conjunto de validación antes de la validación del modelo.

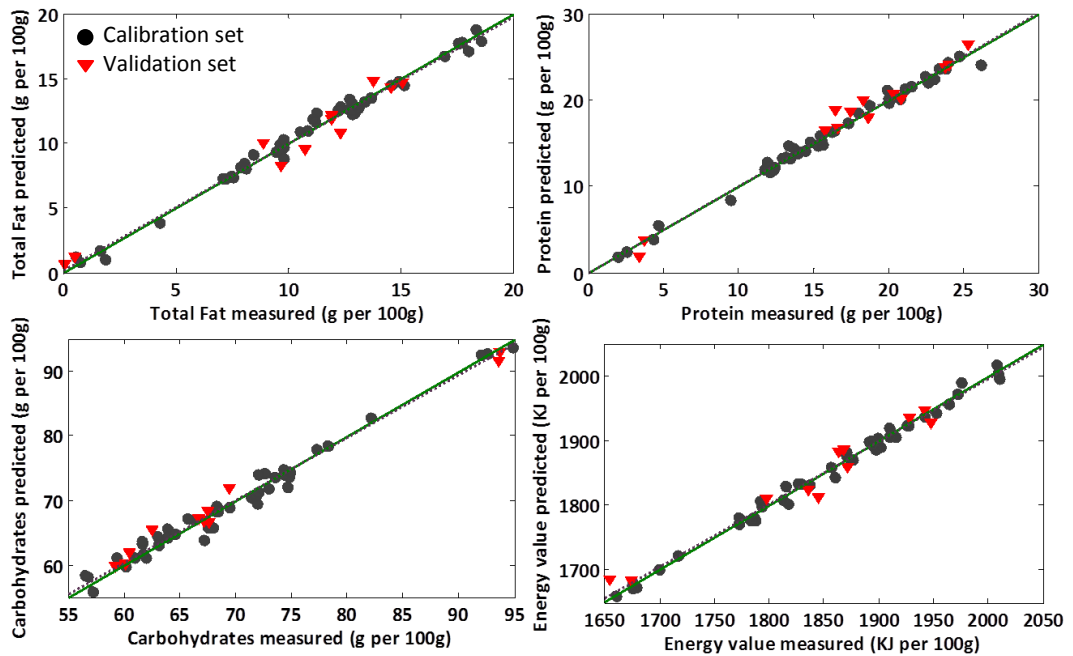


Figura 1.5. Valores predichos frente los valores de referencia para la determinación de grasa, proteínas, carbohidratos y valor energético mediante los modelos PLS-NIR construidos

En la Tabla 1.2 se muestran también los parámetros de predicción más importantes de los modelos PLS-NIR. Para el conjunto de validación se obtuvieron valores aceptables de error relativo de predicción, expresado como RRMSEP, para grasas, proteínas, carbohidratos y valor energético de 9.2, 6.1, 2.1, y 1.0 %, respectivamente, con coeficientes de determinación para la predicción (R^2 Pred) de 96.7, 98.1, 98.9, y 96.5.

Para conocer la exactitud del método se calculó el error relativo de las muestras del conjunto de validación a partir de los valores de referencia y los valores predichos por el método PLS seleccionado para cada uno de los parámetros considerados. La exactitud del método fue muy buena para carbohidratos y valor energético con errores comprendidos entre -1.9 y 1.0% y del -1.8 al 1.9%, respectivamente. Sin embargo, los errores relativos para grasas y proteínas obtenidos se encontraron entre -14.5 y 12.3% y -4.3 a 13.9%, respectivamente. Además se calcularon los coeficientes de variación para 11 muestras, con 3 repeticiones cada una, los cuales variaban entre 1.6 y 7.4, 1.0 y 4.3, 0.1 y 0.7, y 0.1 y 1.6% para grasa total, proteínas, carbohidratos y valor energético, respectivamente.

3.1.3.2. Espectroscopia MIR

Las medidas en el infrarrojo medio se realizaron depositando aproximadamente 100 mg de muestra directamente sobre el cristal ATR de diamante y presionando con el accesorio correspondiente. Los espectros se registraron en el intervalo entre 4000 y 600 cm^{-1} , acumulando 50 barridos por espectro, a una resolución de 4 cm^{-1} . El espectro del cristal ATR limpio, obtenido inmediatamente antes de adquirir el espectro de cada muestra se utilizó como *background*. Los espectros de tres porciones de cada muestra fueron normalizados por mínimo-máximo después de eliminar la región entre 2300 y 1900 cm^{-1} debido a la alta absorción de la célula ATR (diamante). Los espectros promedio de cada muestra se obtuvieron a partir de los normalizados y se utilizaron para construir los modelos de PLS y para la predicción de la muestra.

En la Figura 1.6 puede verse el espectro ATR promedio de las muestras de alimentos utilizadas en este estudio en la región entre 4000 y 600 cm^{-1} , con las distintas regiones pertenecientes a los principales constituyentes. La región entre 4000 y 3100 cm^{-1} pertenece al grupo O-H (3400 cm^{-1}) y a 2 modos vibracionales del estiramiento del enlace N-H. La región entre 3100 y 2800 cm^{-1} se relaciona con los enlaces C-H debido a los estiramientos tanto simétricos como asimétricos de los grupos CH_2 y CH_3 de grasas y carbohidratos. En la región comprendida entre 1800 y 1500 cm^{-1} se encuentran las bandas amida I (C=O estiramiento) y amida II (C-N estiramiento y N-H flexión), junto con la banda a 1653 cm^{-1} relacionada con la vibración de estiramiento del enlace N-H de las proteínas. En el caso de las grasas, la banda relacionada con el estiramiento del enlace C=O aparece alrededor de 1725–1745 cm^{-1} . La región mixta se encuentra entre 1500 y 1200 cm^{-1} , que contiene bandas características de las grasas, proteínas y carbohidratos. La región comprendida entre 1200 y 900 cm^{-1} se relaciona con el estiramiento de los enlaces C-C y C-O junto con la distorsión de los enlaces C-O-C y C-O-H de los carbohidratos, así como la banda debida a la flexión del enlace C=C-H de las grasas. El intervalo de 900 a 750 cm^{-1} corresponde a la región de la huella dactilar. A 720 cm^{-1} , aparece la banda de giro del enlace CH_2 de la grasa [31, 32].

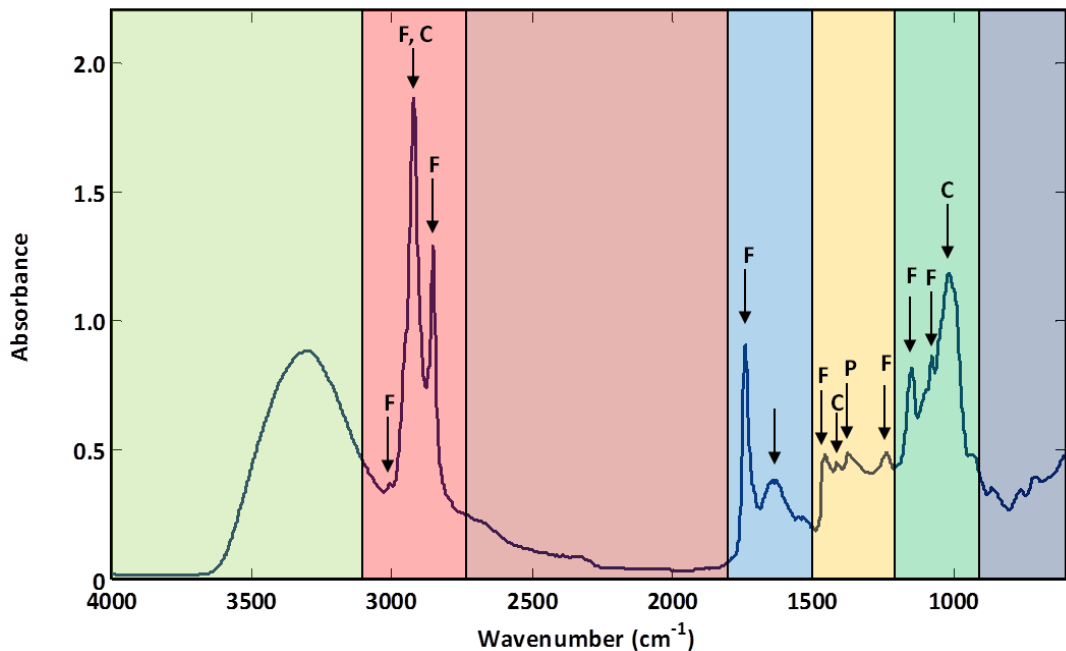


Figura 1.6. Espectro ATR promedio de las muestras de alimentos consideradas en el estudio, en la región comprendida entre 4000 y 600 cm^{-1} . Nota: Las regiones indicadas corresponden a las bandas características de grasa (F), proteínas (P) y carbohidratos (C). Para más detalles ver el texto

3.1.3.2.1. Modelos PLS-ATR-FTIR

La población de datos compuesta por las 58 muestras fue dividida mediante el algoritmo K-S para la selección de un conjunto de calibración y un conjunto de validación, con 45 y 13 muestras, respectivamente. Como en el modelo PLS-NIR, el conjunto de calibración se utilizó para construir los modelos, y el conjunto de validación para evaluar la capacidad predictiva del modelo y probar la robustez del mismo.

A partir de los espectros MIR obtenidos de las muestras se ha construido un dendrograma usando el método de *Ward Linkage* después de MSC, FD, y MC como pre-tratamiento de los espectros, utilizando la región espectral comprendida entre 3600 y 600 cm^{-1} . En la Figura 1.7 se muestra la similitud de las muestras establecidas por el dendrograma construido. Como se puede observar, para un valor de corte de 0.65 se pueden diferenciar 4 grupos: un grupo formado por tarritos para bebés de frutas, el otro basado en menús infantiles de comida rápida y los otros dos grupos adicionales basados en tarritos para bebés y menús para adultos. Por lo tanto, a partir del dendrograma se puede observar que las

muestras seleccionadas por el algoritmo K-S para el conjunto de calibración son representativas, estando formado por 26 muestras de tarritos para bebés, 4 de ellas de frutas, 9 menús para adultos y 10 menús de comida rápida infantil.

Las principales características de los conjuntos de calibración y validación empleados para la construcción de los modelos PLS-ATR-FTIR para todos los parámetros nutricionales considerados se muestran en la Tabla 1.3. En ella, se puede ver el número de muestras utilizadas después de rechazar posibles valores anómalos, el valor medio, la desviación y el rango de concentraciones de cada uno de los conjuntos de muestras.

Tabla 1.3. Principales características de los conjuntos de calibración y validación empleados para la construcción de los modelos PLS-ATR-FTIR

Analito	Conjunto	Muestras	Rango	Valor medio	SD
Grasa total (g por 100g)	Cal	43	0.07-18.4	11	5
Proteínas (g por 100g)		44	1.95-26.2	16	6
Carbohidratos (g por 100g)		42	55.5-94.8	70	10
Valor energético (KJ por 100g)		42	1650-2010	1900	90
Grasa total (g por 100g)	Val	12	0.05-18.6	11	5
Proteínas (g por 100g)		11	2.57-24.8	16	7
Carbohidratos (g por 100g)		13	57.2-95.0	70	10
Valor energético (KJ por 100g)		13	1660-2008	1900	100

Conjuntos de calibración (Cal) y validación (Val)
 Rango: valor mínimo-máximo. SD desviación estándar

Para la selección de regiones espectrales se utilizó iPLS. Para ello se usó la región espectral comprendida entre 3600 y 600 cm^{-1} , excepto la región entre 2300 y 1900 cm^{-1} debido a la absorción de la celda ATR de diamante, y se dividió en intervalos de 88 variables usando como máximo 10 LV's para la construcción del modelo. La combinación óptima de los intervalos espectrales se obtuvo a partir del menor error de la validación cruzada (RMSECV). En la Figura 1.8 se pueden observar los espectros y las regiones seleccionadas mediante iPLS para grasa, proteínas, carbohidratos y valor energético. Las regiones seleccionadas para la grasa fueron aquellas comprendidas entre 3148 y 2810 cm^{-1} , en las que aparecen las bandas características debidas a tensión tanto simétrica como asimétrica del enlace C-H de los grupos $-\text{CH}_2-$ y $-\text{CH}_3$, el intervalo entre 1450 y 1282 cm^{-1} correspondiente a la región mixta, y la región entre 1111 y 773 cm^{-1} donde se encuentran las bandas debidas a los doblamientos del grupo C=C-H. En cuanto a las proteínas, las regiones seleccionadas por el iPLS fueron aquellas comprendidas entre 1450 y 1282 cm^{-1} correspondientes a la región mixta y aquellas situadas entre 941 y 773 cm^{-1} de la región de la huella dactilar.

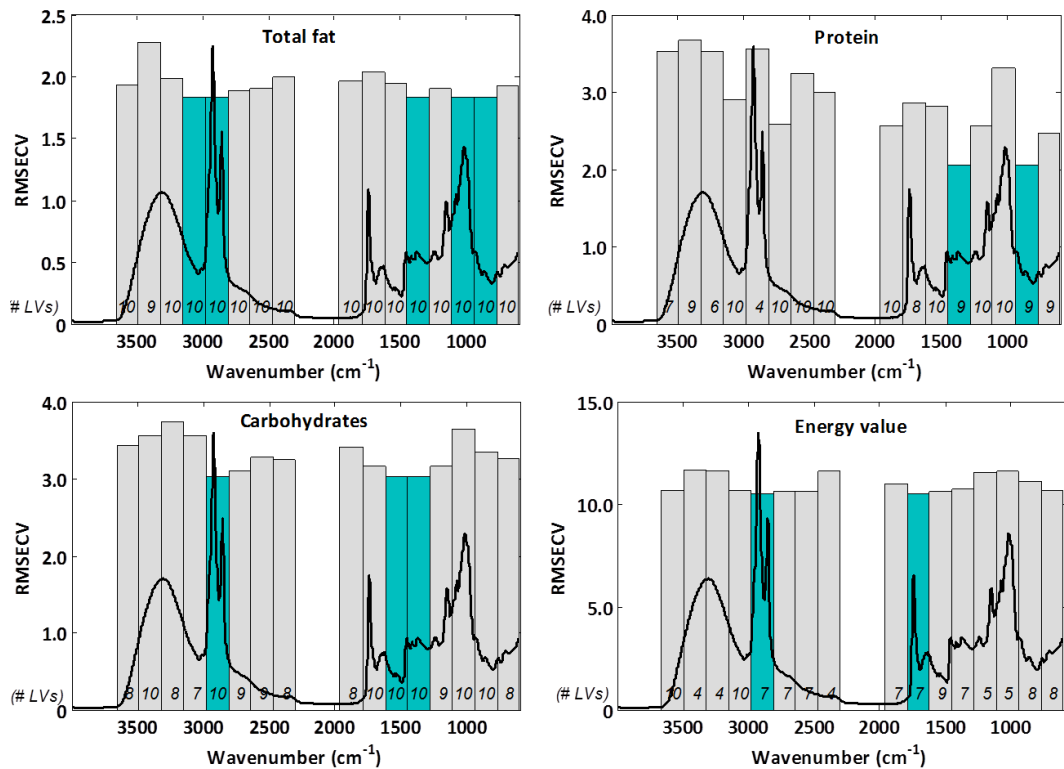


Figura 1.8. Espectro NIR con la selección de las regiones mediante iPLS para la determinación de grasa total, proteínas, carbohidratos y valor energético mediante la construcción de los modelos PLS-ATR-FTIR. Nota: los valores en cursiva corresponden al número de variables latentes (LV)

Las regiones seleccionadas para la predicción de los carbohidratos fueron las comprendidas entre 2978 y 2810 cm⁻¹, dónde se encuentran las bandas debidas al estiramiento del enlace C-H de los grupos CH₂ y CH₃, y la región mixta comprendida entre 1620 y 1282 cm⁻¹. En el caso del valor energético, se seleccionaron las regiones entre 2978 y 2810 cm⁻¹ correspondientes a los estiramientos simétricos y asimétricos del enlace C-H de los grupos CH₂ y CH₃ de la grasa y carbohidratos, y entre 1789 y 1622 cm⁻¹ donde se encuentran las bandas de estiramiento del grupo C=O de la grasa.

Los modelos PLS-ATR-FTIR se construyeron utilizando las regiones seleccionadas mediante iPLS, usando 6, 7, 5, y 7 LV's para grasa, proteínas, carbohidratos y valor energético. Como pre-tratamiento espectral para todos los casos se consideró FD y MC. Antes de construir el modelo, se eliminaron 2 muestras en el conjunto de calibración de grasa y 1 en proteínas, mientras que en carbohidratos y valor energético se eliminaron 3 muestras, por considerarse en todos los casos como

valores anómalos. En la Tabla 1.4 se indican las características más importantes de los modelos PLS-ATR-FTIR construidos para cada uno de los parámetros nutricionales, como son RMSEC, RMSECV y R² de calibración y validación cruzada.

Tabla 1.4. Descripción de los modelos PLS-NIR empleados para la determinación de grasas totales, proteínas, carbohidratos y valor energético

	Pre-procesamiento	LV's	RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	RRMSEP (%)	R ² Pred
Grasa total (g por 100g)	MC,FD	6	1.2	93.3	2.3	77.5	1.7	16.1	91.0
Proteínas (g por 100g)	MC,FD	7	1.4	94.2	2.4	84.0	1.9	11.4	93.0
Carbohidratos (g por 100g)	MC,FD	5	2.4	93.8	3.8	84.8	4.0	5.7	92.0
Valor energético (KJ por 100g)	MC,FD	7	24.9	92.3	42.1	78.4	42.6	2.3	84.1

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción. MC: centrado en la media; FD: primera derivada.

3.1.3.2.2. Capacidad de predicción del modelo PLS-ATR-FTIR

El conjunto de validación formado por 13 muestras independientes al conjunto de calibración, se utilizó para confirmar la capacidad predictiva del modelo PLS-ATR-FTIR para cada uno de los parámetros considerados en este estudio. En la Figura 1.9 se representa la regresión entre los resultados predichos por el modelo frente los valores de referencia de grasa, proteínas, carbohidratos y valor energético, tanto para el conjunto de calibración como para el de validación. Únicamente 1 muestra fue eliminada en el caso de la grasa y 2 en el caso de las proteínas.

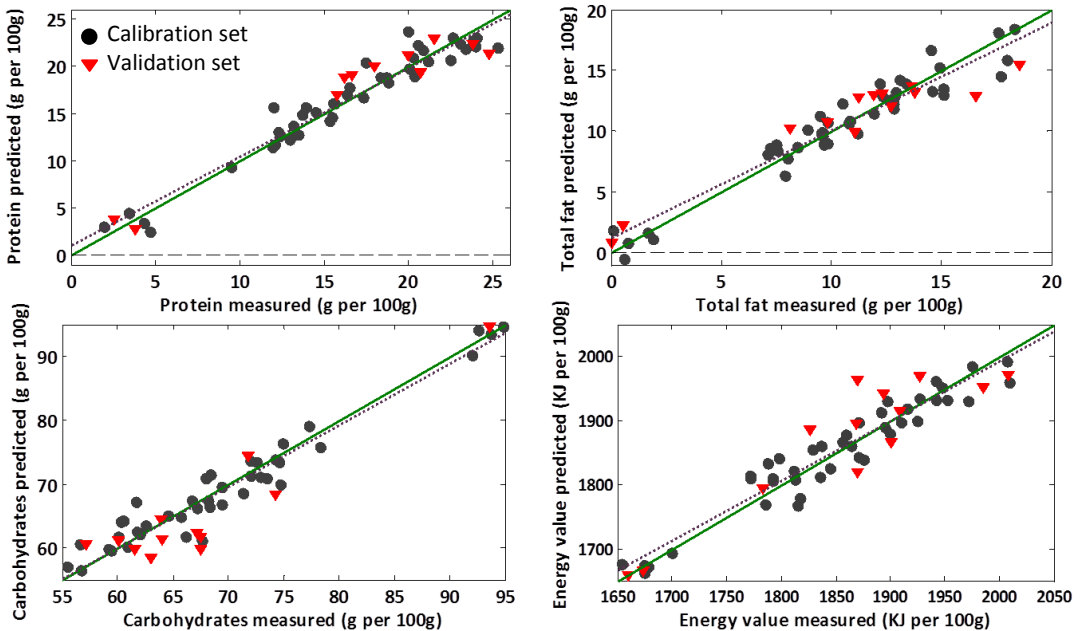


Figura 1.9. Valores predichos frente los valores de referencia para la determinación de grasa, proteínas, carbohidratos y valor energético mediante los modelos PLS-ATR-FTIR construidos

Para evaluar la exactitud de los modelos construidos, se calcularon los errores relativos entre los valores de referencia y los valores predichos del conjunto de validación. Para el caso de la grasa los errores encontrados variaban entre -16.8 y 25.4% , para las proteínas entre -13.8 y 16.2% , para los carbohidratos entre -11.5 y 13.8% , y para el valor energético entre -2.7 y 5% . Además, se calcularon los coeficientes de variación de 3 replicas de las 11 muestras del conjunto de validación, obteniendo valores entre 0.4 y 15.4% , 0.8 y 12.2% , 0.2 y 5.0% , y 0.1 y 3.0% para grasa, proteínas, carbohidratos y valor energético, respectivamente.

3.1.3.3. Comparación entre NIR y MIR

Los modelos desarrollados mediante el uso de la espectroscopia NIR y MIR para la determinación de grasa, proteínas, carbohidratos y valor energético fueron aceptables. Los errores obtenidos tanto de calibración como de validación cruzada fueron menores para los modelos PLS-NIR con valores de RMSEC entre 0.4 y 1.3 g por 100 g para grasa, proteínas y carbohidratos, y de 7.9 kJ por 100 g para el valor energético, mientras que para RMSECV, los valores obtenidos para grasa, proteínas y carbohidratos se encontraban entre 1.0 y 1.6 g por 100 g, mientras que para el valor energético el error fue de 20.5 kJ por 100 g. En el caso de los

modelos PLS-ATR-FTIR, los valores RMSEC obtenidos para grasa, proteínas y carbohidratos variaban entre 1.2 y 2.4 g por 100 g y de 24.9 kJ por 100 g en el valor energético, mientras que para RMSECV, dichos errores se encontraban entre 2.3 y 3.8 g por 100 g en el caso de grasa, proteínas y carbohidratos y de 42.1 kJ por 100 g para el valor energético. Como puede comprobarse los errores correspondientes a los modelos PLS-ATR-FTIR resultan ligeramente superiores a los obtenidos para la modelización PLS-NIR.

Con respecto la capacidad predictiva de los modelos construidos, el modelo PLS-NIR proporcionó R^2 Pred de 96.7, 98.1, 98.9, y 96.5 para grasa, proteínas, carbohidratos y valor energético, con valores de RRMSEP de 9.2, 6.1, 2.1, y 1.0 %, respectivamente. Para los modelos PLS-ATR-FTIR, los valores de R^2 Pred encontrados fueron de 91.0, 93.0, 92.0 y 84.1 para grasa, proteínas, carbohidratos y valor energético con valores de RRMSEP de 16.1, 11.4, 5.7 y 2.3 %, respectivamente. Como se puede observar, con los modelos PLS-NIR se obtienen mejores resultados que los encontrados por los modelos PLS-ATR-FTIR.

Para evaluar la capacidad de predicción de las técnicas NIR y MIR, se calcularon los errores relativos de la predicción de cada uno de los parámetros a determinar. Los errores relativos encontrados mediante los modelos PLS-NIR, variaban entre -14.5 y 12.3, -4.3 y 13.9, -1.9 y 1.0, y -1.8 y 1.9 % para grasa, proteínas, carbohidratos y valor energético, respectivamente, con coeficientes de variación entre 1.6 y 7.4, 1.0 y 4.3, 0.1 y 0.7, y 0.1 y 1.6 %. En cuanto los modelos PLS-ATR-FTIR, los errores relativos variaron entre -16.8 y 25.4, -13.8 y 16.2, -11.5 y 13.8, y -2.7 y 5 con coeficientes de variación entre 0.4 y 15.4, 0.8 y 12.2, 0.2 y 5.0, y 0.1 y 3.0 % para grasa, proteínas, carbohidratos y valor energético, respectivamente.

A partir de los resultados obtenidos en este estudio, se puede concluir que ambas técnicas, tanto NIR como MIR, pueden aplicarse para la determinación de grasa, proteínas, carbohidratos y valor energético en tarritos para bebés, menús infantiles de comida rápida y menús de comedores colectivos. Sin embargo, los modelos PLS-NIR presentan menores errores de calibración, validación cruzada y predicción que los modelos PLS-ATR-FTIR, y con mejores resultados de precisión, más adecuados para la determinación de grasas, proteínas, carbohidratos y valor energético en este tipo de muestras, lo que indica una mayor capacidad predictiva de las medidas en la región NIR que en la MIR. A ello contribuye en cierta medida la mayor precisión que se obtiene con las medidas de reflectancia difusa en el NIR, en las que no es necesario ningún tipo de manipulación de las muestras para su medida.

3.1.4. Conclusiones Capítulo 1

A partir de los resultados obtenidos en este estudio y teniendo en cuenta que la preparación de muestra empleada en este trabajo es una etapa común para obtener los resultados en materia seca, se puede concluir que los modelos PLS-NIR construidos después de la selección del set de calibración mediante el algoritmo K-S y una correcta selección de regiones espectrales, pre-procesado espectral y número de LV's, proporcionan una herramienta rápida y "verde" para el análisis de grasa, proteínas, carbohidratos y valor energético en tarritos para bebés, menús infantiles de comida rápida y menús de cafetería, en unos pocos minutos y sin la necesidad de utilizar reactivos ni disolventes. Es de destacar que en el caso de las medidas NIR éstas se realizan directamente sobre las muestras contenidas en viales de vidrio, por lo que la manipulación de las mismas es mínima, no produciéndose ninguna alteración ni modificación de las mismas.

3.1.5. Referencias Capítulo 1

- [1] World Health Organization (WHO) (2009) Infant and Young Child Feeding. Model Chapter for Textbooks for Medical Students and Allied Health Professionals.
- [2] World Health Organization (WHO) and Food and Agriculture Organization (FAO) of the United Nations CODEX Alimentarius, Food labelling, 5th edn. Rome, Italia, 2007.
- [3] W. Horwitz, Official Methods of Analysis of AOAC International, 17th edn, 2000.
- [4] J. Moros, S. Garrigues, M. de la Guardia, Vibrational spectroscopy provides a green tool for multi-component analysis, *TrAC Trends Anal. Chem.*, 29(2010) 578-591.
- [5] H. Cen, Y. He, Theory and application of near infrared reflectance spectroscopy in determination of food quality, *Trends Food Sci Technol.*, 18(2007) 72-83.
- [6] D.S. Ferreira, O.F. Galão, J.A.L. Pallone, R.J. Poppi, Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples, *Food Control*, 35(2014) 227-232.
- [7] J. Chen, X. Ren, Q. Zhang, X. Diao, Q. Shen, Determination of protein, total carbohydrates and crude fat contents of foxtail millet using effective wavelengths in NIR spectroscopy, *J. Cereal Sci.*, 58(2013) 241-247.
- [8] R. Karoui, A.M. Mouazen, E. Dufour, L. Pillonel, E. Schaller, J. de Baerdemaeker, J.O. Bosset, Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools. *Int. Dairy J.*, 16(2006) 1211-1217.
- [9] P. Ritthiruangdej, R. Ritthiron, H. Shinzawa, Y. Ozaki, Non-destructive and rapid analysis of chemical compositions in Thai steamed pork sausages by near-infrared spectroscopy, *Food Chem.*, 129(2011) 684-92.
- [10] K. Khodabux, M.S. L'Omelette, S. Jhaumeer-Laulloo, P. Ramasami, P. Rondeau, Chemical and near-infrared determination of moisture, fat and protein in tuna fishes, *Food Chem.*, 102(2007) 669-75.
- [11] C. Shiroma, L. Rodriguez-Saona, Application of NIR and MIR spectroscopy in quality control of potato chips, *J. Food Compos Anal* 22(2009) 596-605.

- [12] L.K. Sørensen, Application of reflectance near infrared spectroscopy for bread analyses, *Food Chem.*, 113(2009) 1318-22.
- [13] J. Moros, F.A. Iñón, S. Garrigues, M. de la Guardia, Near-infrared diffuse reflectance spectroscopy and neural networks for measuring nutritional parameters in chocolate samples, *Anal. Chim. Acta*, 584(2007) 215-222.
- [14] A. Vesela, A. Barros, A. Synytsya, I. Delgadillo, J. Copíková, M. Coimbra, Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder, *Anal. Chim. Acta*, 601(2007) 77-86.
- [15] J. Moros, F.A. Iñón, S. Garrigues, M. de la Guardia, Determination of the energetic value of fruit and milk-based beverages through partial-least-squares attenuated total reflectance-Fourier transform infrared spectrometry, *Anal Chim Acta*, 538(2005) 181-193.
- [16] M. Hernández-Martínez, T. Gallardo-Velázquez, G. Osorio-Revilla, N. Almaraz-Abarca, A. Ponce-Mendoza, M.S. Vásquez-Murrieta, Prediction of total fat, fatty acid composition and nutritional parameters in fish fillets using MID-FTIR spectroscopy and chemometrics, *LWT-Food Sci Technol.*, 52(2013) 12-20.
- [17] M. Khanmohammadi, A.B. Garmarudi, K. Ghasemi, S. Garrigues, M. de la Guardia, Artificial neural network for quantitative determination of total protein in yogurt by infrared spectrometry, *Microchem J.*, 91(2009) 47-52.
- [18] F.A. Iñón, S. Garrigues, M. de la Guardia, Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques, *Anal Chim Acta*, 513(2004) 401-412.
- [19] K. Almendingen, H.M. Meltzer, J.I. Pedersen, B.N. Nilsen, M. Ellekjær, Near infrared spectroscopy-a potentially useful method for rapid determination of fat and protein content in homogenized diets, *Eur. J. Clin. Nutr.*, 54(2000) 20-23.
- [20] Y. Kim, M. Singh, S.E. Kays, Near-infrared spectroscopic analysis of macronutrients and energy in homogenized meals, *Food Chem.*, 105(2007) 1248-1255.
- [21] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemom Intell Lab Syst.*, 58(2001) 109-130.
- [22] Food and Agriculture Organization of the United Nations, Food energy-methods of analysis and conversion factors, Roma, Italia, 2002.

- [23] U.S. Department of Agriculture, Agricultural Research Service, USDA Nutrient Data Laboratory, Composition of Foods Raw, Processed, Prepared, 2010.
- [24] Nutritional Labelling and Education Act, U.S. Department of Health & Human Services.
- [25] W.O. Atwater, C. Woods, The Chemical composition of American food materials. Washington, 1896.
- [26] A. Nawrocka, J. Lamorska, In S. Grundas, A. Stepniewski (ed) Advances in Agrophysical Research, Determination of Food Quality by Using Spectroscopic Methods, 2013.
- [27] J. Shenk, J. Jr. Workman, M. Westerhaus, In D. Burns, E. Ciurczak (ed) Handbook of Near-Infrared Analysis, Application of NIR spectroscopy to Agricultural Products, 2nd edn, 2001.
- [28] D. Naumann, Infrared Spectroscopy in Microbiology. Encyclopedia of Analytical Chemistry, 2006.
- [29] N. Vlachos, Y. Skopelitis, M. Psaroudaki, V. Konstantinidou, A. Chatzilazarou, E. Tegou, Applications of Fourier transform-infrared spectroscopy to edible oils, Anal. Chim. Acta, 573(2006) 459-465.

3.2. CAPÍTULO 2:
Determinación de compuestos polares en aceites

3.2. Capítulo 2. Determinación de compuestos polares en aceites

2.1. Introducción

Los aceites y grasas expuestas a altas temperaturas durante el proceso de fritura, en presencia de aire o humedad, sufren una serie de reacciones que afectan drásticamente a su calidad. La polimerización, la oxidación y las reacciones hidrolíticas generan una amplia gama de productos de degradación, tanto volátiles como no volátiles, que pueden tener importantes efectos fisiológicos sobre la salud humana [1]. Los materiales polares totales (TPM, del inglés *Total Polar Materials*) se consideran como el mejor indicador y el parámetro más común para comprobar la calidad y la evaluación de los aceites de fritura [2]. Algunos países han establecido un valor máximo de TPM en aceite de fritura entre 24 y 30% (m/m). En España el nivel máximo de TPM establecido por la legislación española está en torno al 25% en peso [3]. Por lo tanto, es de gran importancia la determinación de TPM en aceites de fritura para garantizar su calidad.

La determinación de TPM normalmente se realiza mediante separación por cromatografía en columna de gel de sílice y posterior análisis gravimétrico, siguiendo el método estándar aprobado por la Unión Internacional de Química Pura y Aplicada (IUPAC, del inglés *International Union of Pure and Applied Chemistry*) y la Asociación de Químicos Analíticos Oficiales (AOAC, del inglés *Association of Official Analytical Chemists*) [4]. Sin embargo, este método de referencia consume gran cantidad de tiempo y productos químicos y, es por ello, que se han realizado una serie de intentos para encontrar métodos alternativos, simples y rápidos, para la determinación de TPM en aceite de fritura usado.

Existen dispositivos comercializadas para determinar el contenido de TPM en aceites basados en la medida de la constante dieléctrica, ya que un aumento en la concentración de moléculas polares aumenta directamente la constante dieléctrica del aceite [5]. Sin embargo, estos instrumentos requieren, por lo general, de una calibración compleja, no son adecuadas para ser empleadas en todos los tipos de aceite y, además, la temperatura puede influir en las mediciones [6]. Además, cuando se emplean estos sistemas de medición, la presencia de agua, sal y minerales puede afectar la polaridad de las muestras, proporcionando una errónea lectura sobre el contenido de TPM en los aceites [5].

La espectroscopia con transformada de Fourier en el infrarrojo medio (FT-IR) y cercano (FT-NIR) se presentan como dos potentes herramientas para monitorear la calidad del aceite de fritura usado. Se han realizado diversos estudios sobre la calidad del aceite mediante el uso de FT-IR para la determinación de la acidez y del contenido de TPM [7] o para evaluar la degradación de aceites comestibles

durante el proceso de calentamiento, mediante la determinación de ácidos grasos cis-insaturados, trans y ácidos grasos libres [8]. Además existen otros trabajos donde se utiliza el FT-NIR para la determinación del contenido de compuestos polares totales, triacilglicéridos (TAG) polimerizados, acidez y anisidina en grasas y aceites de fritura [9], ácidos grasos libres (FFA), TPM, viscosidad y punto de humo en aceites refinados de avellana y cacahuete [10], TAG polimerizados en aceites vegetales de diferentes orígenes botánicos [11] y acidez, material polar total (TMT), triglicéridos diméricos y poliméricos (DPTG) en aceites de cantinas, restaurantes y cafeterías [12].

En los últimos años se han realizado esfuerzos para desarrollar métodos rápidos para evaluar la calidad del aceite de fritura, como el diseño de un sensor capacitivo mediante electrodos interdigitados (IDEs) [13] que permiten evaluar la degradación del aceite de fritura o el uso del análisis hiperespectral VIS/NIR con calibración PLS [14]. Estos últimos métodos parecen prometedores para el monitoreo continuo en línea de la calidad a gran escala, pero no para el sector de los servicios alimentarios y los restaurantes en pequeña escala debido a su volumen y coste.

Este estudio sobre el empleo de la espectroscopia vibracional para el análisis de aceites de fritura se encuentra dividido en 2 capítulos:

- **2.A.** Comparación de la espectroscopia FT-NIR y FT-MIR para la determinación de compuestos polares en aceites de fritura usados.
- **2.B.** Estudios previos para el desarrollo de un instrumento de bajo coste para la evaluación de la calidad del aceite de fritura.

Por tanto, los objetivos de esta parte de la Tesis son:

- El estudio y comparación del empleo de la espectroscopia FT-NIR y FT-MIR, en asociación con la calibración de mínimos cuadrados parciales (PLS), como herramientas analíticas verdes para la determinación directa de TPM en aceites de fritura.
- La evaluación del uso de la espectroscopia NIR junto con las herramientas quimiométricas para seleccionar las condiciones de medida y los números de onda más adecuados para la determinación directa de TPM en aceites fritos, con el fin de obtener datos preliminares para el desarrollo de un instrumento de bajo coste para la evaluación de la calidad del aceite de fritura.

3.2.2. Parte experimental

3.2.2.1. Instrumentos

Los espectros del infrarrojo cercano registrados en este trabajo fueron obtenidos con un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Germany) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS. El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

Los espectros del infrarrojo medio fueron registrados con un espectrofotómetro modelo Tensor 27 FT-MIR de Bruker (Bremen, Alemania) equipado con un detector de sulfato de triglicina dopado con L-alanina deuterada (DLATGS). Los espectros de las muestras se obtuvieron mediante medidas en modo ATR, utilizando el accesorio para líquidos de Smiths Detection Inc. (Warrington, UK) con un cristal DuraDisk de diamante/ZnSe de nueve reflexiones (véase Figura 2.1). Este instrumento también emplea el software Opus 6.5 de Bruker para su control instrumental y adquisición de los espectros.



Figura 2.1. Accesorio con cristal de diamante/ZnSe de 9 reflexiones para muestras líquidas

3.2.2.2. Método de referencia

En ambos estudios, para obtener los datos de referencia de los valores de TPM de las muestras empleadas, se utilizó un medidor de aceite de fritura Testo 270 (Testo Inc., Lenzkirch, Alemania) para medir la constante dieléctrica del aceite y relacionarla con el contenido total de polares en el aceite [5]. Osawa *et al.* [15] y Ögütcü *et al.* [10] encontraron una buena correlación entre los valores de TPM determinados por el procedimiento de referencia mediante separación cromatográfica y los obtenidos con el medidor Testo. La determinación de TPM se realizó por triplicado y se consideró el valor medio de dichos valores para el desarrollo de los estudios.

Capítulo 3.2.A. Comparación de la espectroscopia FT-NIR y FT-MIR para la determinación de compuestos polares en aceites

3.2.A.1. Muestras

El número total de muestras que se utilizaron en este estudio fue de 75, correspondientes a muestras de aceite frito usado y mezclas de ellas, las cuales pertenecen a aceites de girasol y de oliva (virgen y extra-virgen). Las muestras se pueden clasificar en tres grupos: i) 41 aceites fritos procedentes de uso doméstico y utilizados en la fritura de diferentes tipos de alimentos, ii) 5 muestras de aceites comestibles sin usar y, iii) 29 muestras mixtas creadas a partir de mezclas de las de los grupos anteriores.

3.2.A.2. Resultados y discusión

3.2.A.2.1. Espectroscopia NIR

Los espectros NIR en modo transmisión de las muestras de aceite se obtuvieron directamente en viales de vidrio de 1 ml, con un paso óptico de 6 mm, registrándose en el intervalo entre 14000 y 3500 cm^{-1} y usando el espectro de un vial de vidrio vacío como *background*. Los espectros fueron adquiridos acumulando 50 barridos y utilizando una resolución óptica de 4 cm^{-1} . Se obtuvieron tres medidas de cada una de las muestras, girando la posición del vial entre las repeticiones, para asegurar una buena representatividad de los espectros. Todos los espectros se registraron a 25 ± 1 °C, calculándose los espectros promedio de cada muestra para ser empleados en la construcción de modelos de regresión y para la predicción.

En la Figura 2.2 pueden verse los espectros NIR de muestras de aceite frito consideradas en este estudio sin ningún tipo de tratamiento previo de datos, en la región entre 14000 y 3500 cm^{-1} . Las bandas de absorción alrededor de 8261 cm^{-1} pertenecen al segundo sobretono de las vibraciones de estiramiento de los enlaces C-H, mientras que las bandas situadas a 7181 y 7065 cm^{-1} se deben a la banda de combinación de los enlaces C-H. Las bandas alrededor de 5792 y 5676 cm^{-1} se atribuyen al primer sobretono de las vibraciones de estiramiento de los enlaces C-H de los grupos CH_3 , CH_2 y $\text{CH}=\text{CH}$. Las bandas situadas a 4654 cm^{-1} se asocian con las bandas de combinación de los enlaces C-H y vibraciones de estiramientos de los enlaces C-O [16,17].

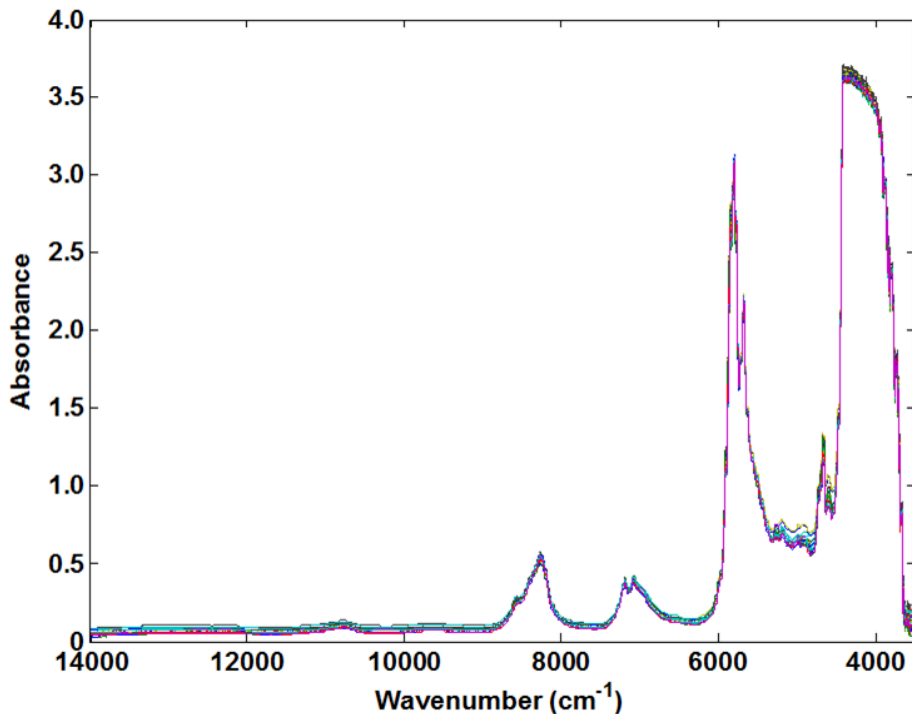


Figura 2.2. Espectros NIR de las muestras de aceite usado en la región comprendida entre 14000 y 3500 cm^{-1}

3.2.A.2.2. Modelos PLS-NIR

El conjunto completo de las muestras de aceite utilizadas en este estudio se dividió en un subconjunto de calibración y otro de validación con 60 y 15 muestras, respectivamente, utilizando el algoritmo K-S. El conjunto de calibración se utilizó para construir el modelo PLS, mientras que el otro conjunto se empleó para comprobar la capacidad predictiva del método desarrollado. En la Tabla 2.1 se indican las características de las muestras incluidas en estos dos conjuntos.

Para construir los modelos PLS, se probaron diferentes pre-procesamientos espectrales como la variación normal estándar (SNV), la corrección de la dispersión multiplicativa (MSC), la primera o segunda derivada y el centrado en la media, en diversas regiones espectrales. La validación cruzada "leave-one-out" fue utilizada para seleccionar el número óptimo de variables latentes (LV's) de los modelos de acuerdo con el error cuadrático medio de la validación cruzada (RMSECV).

Tabla 2.1. Características de los conjuntos de calibración y validación empleados para los modelos PLS-NIR

Modelo PLS	Conjunto	Número de muestras	TPM (%) (m/m)	
			Intervalo de concentración	Promedio \pm sd
NIR	Calibración	60 (1 anómala)	3.0 – 33.0	15.2 - 8.5
	Validación	15	4.0 – 23.0	8.2 – 5.6

sd: desviación estándar

Finalmente, la selección de la región espectral se realizó mediante iPLS, dividiendo el espectro en sub-intervalos equidistantes de 100 variables y considerando 10 LV's como máximo para la construcción de los modelos. Las regiones seleccionadas por el algoritmo iPLS para la determinación del contenido de TPM fueron aquellas comprendidas entre 12840-12456, 8983-8599, 7826-7442 y 5126-4742 cm^{-1} . Como pre-procesamiento se utilizó la primera derivada con un suavizado de 15 puntos y un centrado en la media, utilizando 5LV's. El conjunto de calibración presentaba una muestra anómala que fue retirada antes de construir el modelo PLS, ya que presentaba valores altos de Q y T^2 Hotelling. En la Tabla 2.2 se pueden ver los valores de R^2 Cal y R^2 CV de 0.99 para ambos, junto con los valores de RMSEC y RMSECV de 0.7 y 0.8, respectivamente.

Tabla 2.2. Parámetros de calibración y predicción del modelo PLS-NIR para la determinación de TPM en muestras de aceite.

Modelo PLS	LV's	Conjunto de Calibración				Conjunto de Validación			
		RMSEC	R^2 Cal	RMSECV	R^2 CV	RMSEP	RRMSEP	R^2 Pred	RPD
NIR	5	0.7	0.99	0.8	0.99	0.6	7.6	0.99	9.1

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R^2 Cal: coeficiente de determinación de calibración; R^2 CV: coeficiente de determinación de validación cruzada; R^2 Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

La capacidad predictiva del modelo PLS-NIR obtenido se evaluó utilizando el conjunto de validación externo de 15 muestras no utilizadas para la construcción

del modelo. La Figura 2.3 muestra la regresión entre los contenidos de TPM (%) predichos por el modelo PLS-NIR y los valores de referencia, tanto para el conjunto de calibración como para el de validación. Además, también se representa el ajuste de la recta teórico, es decir, en el caso de que los valores predichos y medidos fuesen los mismos, junto con el ajuste real perteneciente a los valores predichos por el modelo PLS construido.

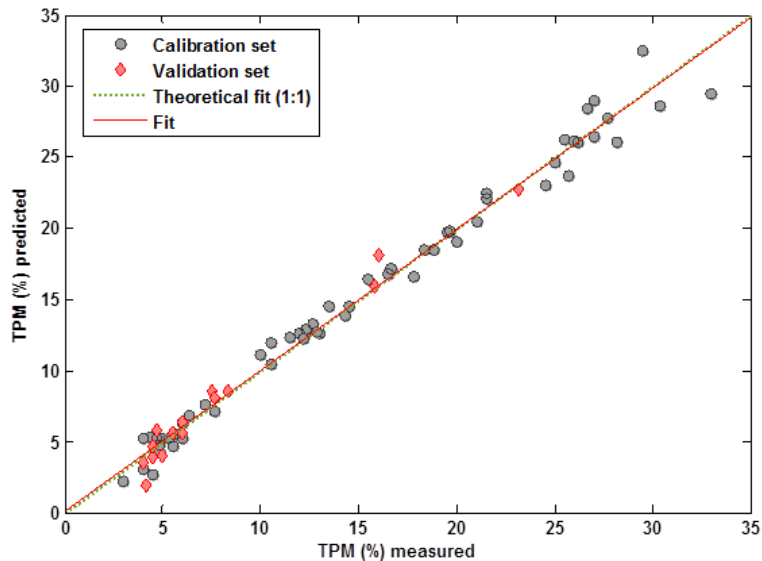


Figura 2.3. Regresión entre los valores de TPM (%) (m/m) predichos y de referencia mediante el uso del modelo PLS-NIR en muestras de aceite

Además, en la Tabla 2.2 también se resumen los parámetros de validación más importantes del modelo PLS-NIR que finalmente se seleccionó. Se puede observar un valor de RRMSEP de 7.6% y un coeficiente de determinación R^2_{pred} de 0.99 junto con un valor RPD de 9.1, por lo que se podría considerar el modelo es adecuado para fines analíticos con una excelente capacidad de predicción, obteniéndose para las muestras del conjunto de validación la siguiente línea de regresión entre los valores de referencia y los calculados por el modelo:

$$TPM_{(PLS-NIR)predicha} = 0.9892 TPM_{Referencia} - 0.2035$$

con unos valores de pendiente y de ordenada en el origen estadísticamente iguales a 1 y 0, respectivamente.

Por otra parte, se realizó un PCA para el análisis exploratorio de datos espectrales de todos los aceites considerados en el estudio. En la Figura 2.4a puede verse el gráfico de *scores* para el primer, segundo y tercer componente principal, después

del utilizar como tratamiento previo la FD y el MC. Estos tres PC's representan en su conjunto el 97.77% de la varianza explicada (PC1 corresponde al 56.17%, PC2 26.40% y PC3 15.20%).

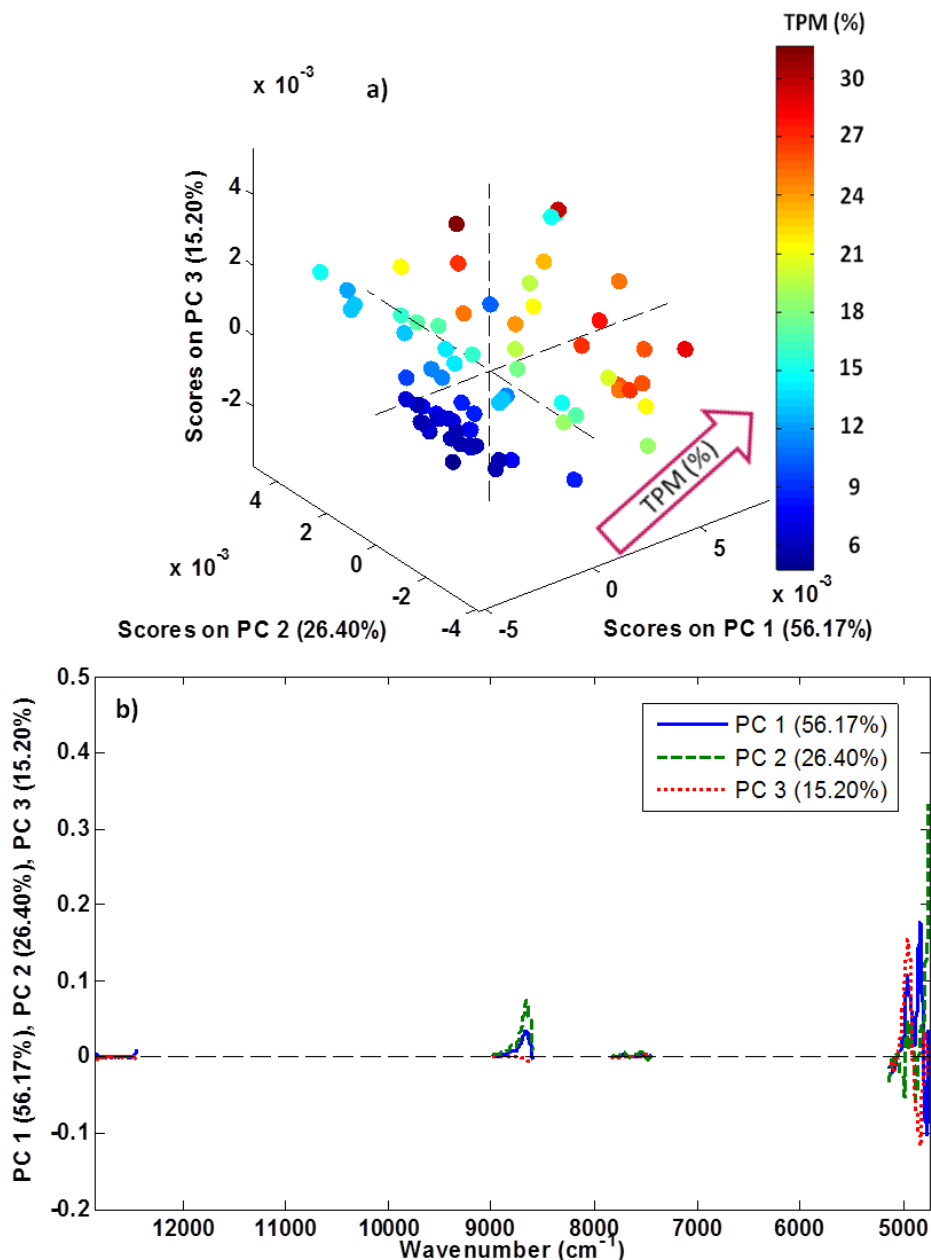


Figura 2.4. Análisis exploratorio PCA por NIR: a) Gráfico de *scores* de acuerdo a la clasificación del contenido de TPM (%) después de la primera derivada y centrado en la media y, b) Gráfico de *loadings* de PC1, PC2 y PC3 en las regiones seleccionadas por iPLS

Se puede observar que las muestras de aceite con bajo contenido de TPM están situadas hacia la izquierda del PC1 y en el centro de PC2 y las muestras con el mayor contenido de TPM se dispersan hacia la parte superior y a la derecha del gráfico (para más detalle ver la flecha en el gráfico de dirección de la distribución de las muestras en función de su contenido de TPM). Los números de onda responsables de la distribución de muestras de aceite frito en PCA pueden identificarse en los gráficos de *loadings* correspondientes (ver Figura 3b) con una clara identificación de bandas alrededor de 8700 cm^{-1} y entre 5100 y 4700 cm^{-1} .

3.2.A.2.3. Espectroscopia MIR

Las medidas ATR-FTIR de las muestras de aceite se realizaron colocando aproximadamente $50\ \mu\text{l}$ de muestra, con la ayuda de una pipeta pasteur, directamente sobre el cristal ATR de diamante. Los espectros se registraron en el intervalo comprendido entre 3500 y 500 cm^{-1} acumulando 50 barridos por espectro a una resolución de 4 cm^{-1} . El *background* se adquirió al cristal ATR limpio inmediatamente antes de realizar el espectro de cada muestra. El cristal ATR se limpiaba con papel de celulosa seco, y comprobando posteriormente que no permanecían restos de aceite en el cristal realizando un espectro del mismo. El promedio de los espectros triplicados de cada muestra se empleó para construir los modelos de regresión y predicción.

En la Figura 2.5 se muestran los espectros ATR de las muestras de aceite frito consideradas en este estudio, sin ningún tipo de pre-tratamiento de datos, en la región entre 3500 y 500 cm^{-1} . La banda de absorbancia observada a 3006 cm^{-1} se debe a las vibraciones de estiramiento del enlace $=\text{C-H}$, mientras que la banda situada alrededor de 2953 cm^{-1} , proviene de la vibración asimétrica de estiramiento del enlace CH de los grupos CH_3 . Las bandas situadas a 2924 y 2854 cm^{-1} se deben a las vibraciones de estiramiento, asimétricas y simétricas, respectivamente, del enlace CH de los grupos CH_2 . Por su parte, las bandas localizadas a 1746 y 1653 cm^{-1} surgen de estiramiento del enlace C=O y vibraciones de estiramiento C=C , respectivamente. Las bandas a 1463 y 1377 cm^{-1} pertenecen a vibraciones de tijereteo de los enlaces CH_2 y CH_3 , mientras que las de 1238 , 1163 , 1114 , 1099 cm^{-1} están asociadas con la vibración de estiramiento del grupo C-O. Finalmente, la banda de absorción a 721 cm^{-1} corresponde al modo de oscilación del grupo CH_2 [18]. La región comprendida entre 2240 y 1880 cm^{-1} se debe a ruido consecuencia de la absorción que presenta el diamante del elemento óptico del accesorio ATR.

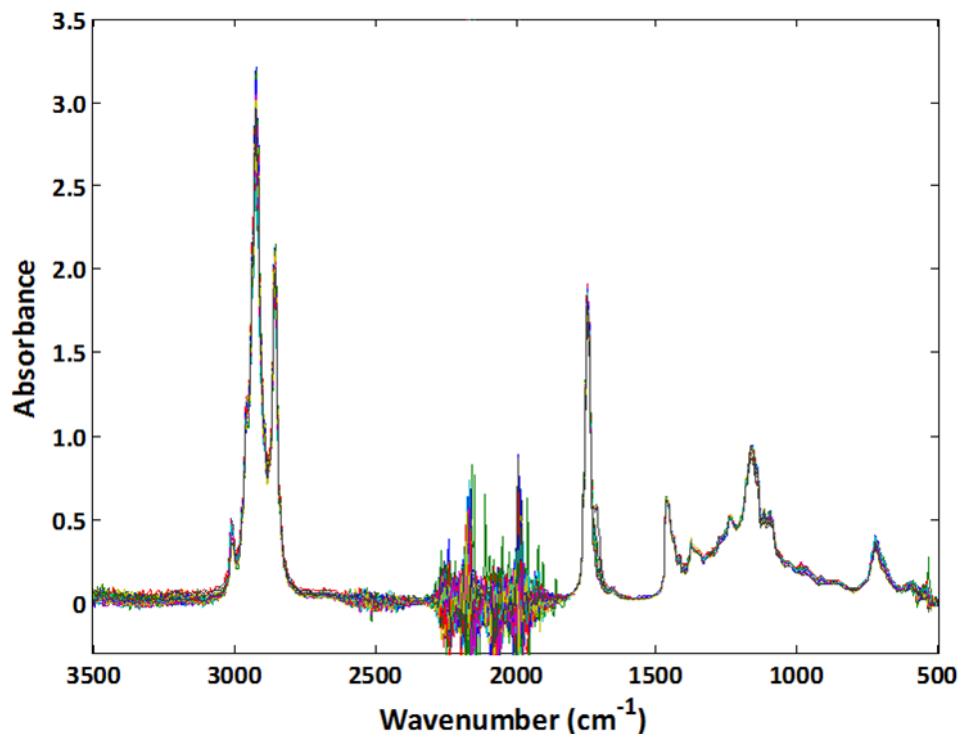


Figura 2.5. Espectros ATR-MIR de las muestras de aceite de este estudio, en la región entre 3500 y 500 cm^{-1}

3.2.A.2.4. Modelos PLS-ATR-FTIR

Para construir los modelos PLS-ATR-FTIR se utilizaron los mismos conjuntos de calibración y validación usados previamente para el modelo de regresión PLS-NIR con el fin de comparar las técnicas MIR y NIR para la predicción del contenido de TPM en aceites fritos. Las principales características de los conjuntos de calibración y validación empleados para la construcción del modelo PLS-ATR-FTIR se encuentran resumidas en la Tabla 2.3, dónde también se informa de las 4 muestras anómalas encontradas en el conjunto de calibración, con valores elevados de Q y T^2 Hotelling.

Para seleccionar las regiones espectrales mediante iPLS, el espectro se dividió en sub-intervalos equidistantes de 100 variables utilizando 10LVs como máximo para construir los modelos de calibración. Las regiones finalmente seleccionadas fueron aquellas comprendidas entre 1765-1381 y 993-802 cm^{-1} . Los mejores resultados se obtuvieron mediante el uso de primera derivada, con un suavizado de 15 puntos, y centrado en la media, con valores de 0.98 y 0.97 para $R^2\text{Cal}$ y $R^2\text{CV}$ y, RMSEC y RMSECV de 1.1 % m/m TPM y 1.6 % m/m TPM, respectivamente.

Tabla 2.3. Características de los conjuntos de calibración y validación empleados para los modelos PLS-MIR

Modelo PLS	Conjunto	Número de muestras	TPM (%) (m/m)	
			Intervalo de concentración	Promedio \pm sd
MIR	Calibración	60 (4 anómalas)	3.0–33.0	15.3–8.7
	Validación	15	4.0–23.0	8.2–5.6

sd: desviación estándar

En la Figura 2.6 se representa la regresión entre los valores predichos para el contenido de TPM obtenido por el modelo PLS-ATR-FTIR y los obtenidos por el método de referencia, encontrándose para el conjunto de validación una recta de regresión

$$\text{TPM}_{(\text{PLS-ATR-FTIR})} = 1.049 \text{TPM}_{\text{Referencia}} - 0.3395$$

con unos valores de pendiente y de ordenada en el origen estadísticamente iguales a 1 y 0, respectivamente.

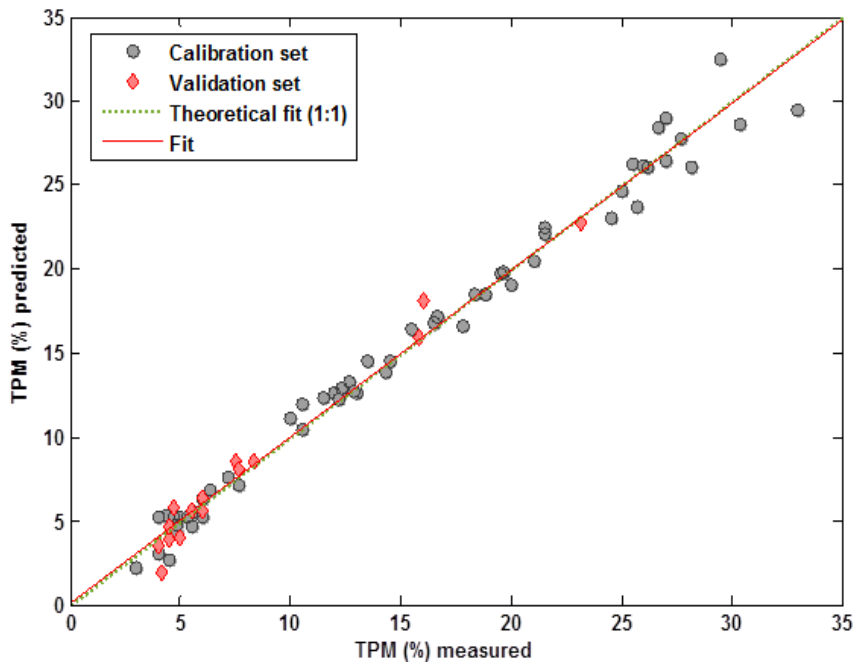


Figura 2.6. Regresión entre los valores de TPM (%) predichos y de referencia mediante el uso del modelo PLS-MIR en muestras de aceite

En la Tabla 2.4 se muestran los parámetros más importantes del modelo PLS-ATR-FTIR que ha proporcionado los mejores resultados, dónde se puede observar que el modelo proporcionó una buena correlación para la predicción, con valores de R^2 Pred y RRMSEP de 0.97 y 11.8%, respectivamente. Además, se obtuvo un valor RPD de 5.8, considerado adecuado para propósitos analíticos con excelente precisión de predicción.

Tabla 2.4. Parámetros de calibración y predicción del modelo PLS-MIR para la determinación de TPM en muestras de aceite

Modelo PLS	LV's	Conjunto de Calibración				Conjunto de Validación			
		RMSEC	R^2 Cal	RMSECV	R^2 CV	RMSEP	RRMSEP	R^2 Pred	RPD
MIR	6	1.1	0.98	1.6	0.97	1.0	11.8	0.97	5.8

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R^2 Cal: coeficiente de determinación de calibración; R^2 CV: coeficiente de determinación de validación cruzada; R^2 Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Por otro lado, la Figura 2.7 ilustra el análisis de datos exploratorio PCA de los datos ATR-MIR, coloreados según el contenido de TPM. La Figura 6a muestra el gráfico de *scores* para el primer, segundo y tercer componente principal después del pre-tratamiento de primera derivada y centrado en la media, representando un total de la varianza explicada del 94.88% (PC1 corresponde al 76.47%, PC2 13.71% y PC3 4.70%). En la Figura 2.7a, se observa que las muestras de aceite con bajo contenido de TPM están situadas hacia la izquierda del PC1 y PC2, mientras que las muestras con un contenido de TPM más alto se dispersan hacia la parte superior derecha del gráfico, tal como queda indicado por la flecha de dirección representada. En la Figura 2.7b se muestran las *loadings* correspondientes a los componentes principales PC1, PC2 y PC3, responsables de la distribución de muestras de aceite frito en el espacio PCA, con una clara identificación de bandas correspondientes a los intervalos de número de ondas comprendidos entre 1765-1381 y 993-802 cm^{-1} .

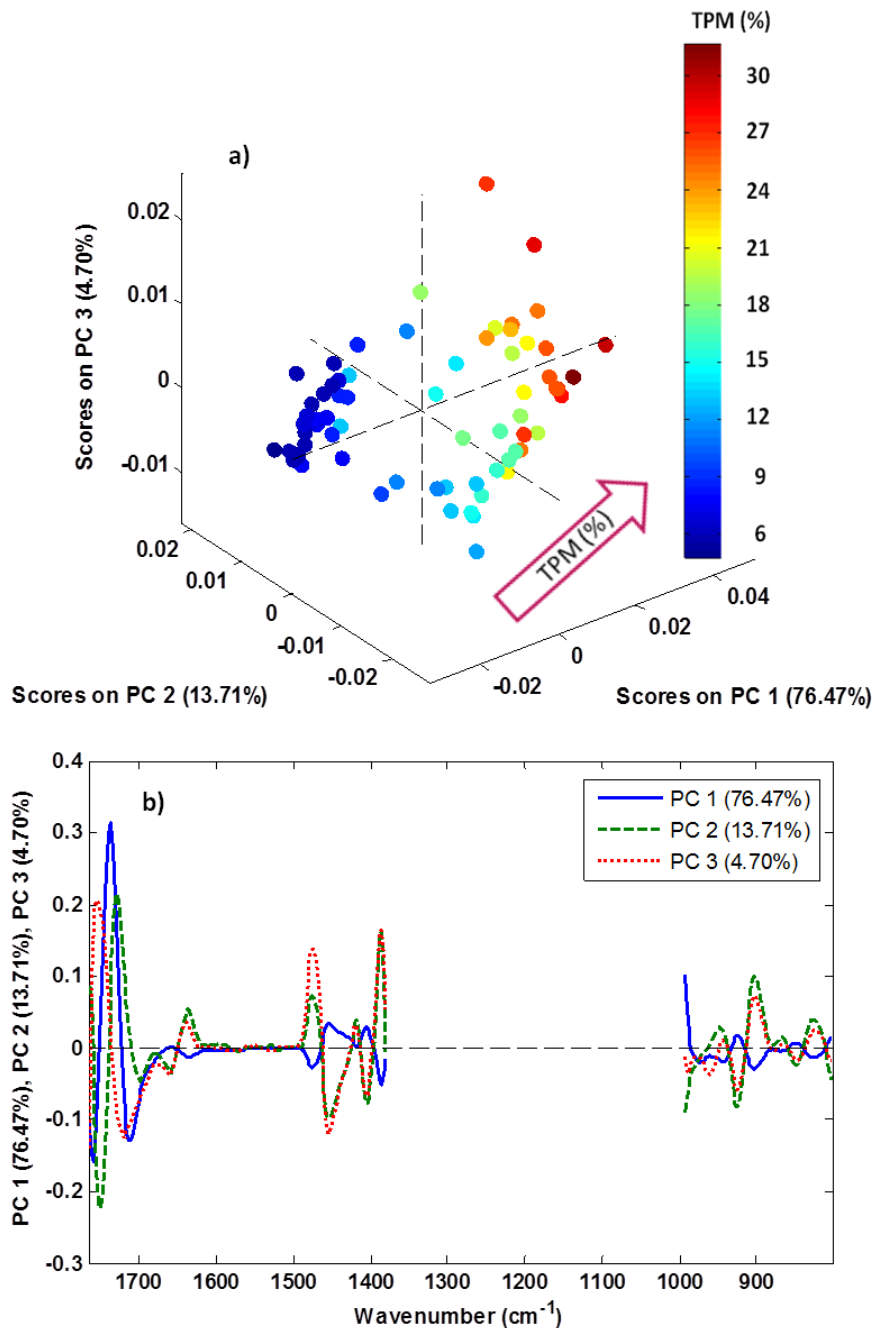


Figura 2.7. Análisis exploratorio PCA por MIR: a) Gráfico de *scores* de acuerdo a la clasificación del contenido de TPM (%) después de la primera derivada y centrado en la media y, b) Gráfico de *loadings* de PC1, PC2 y PC3 en las regiones seleccionadas por iPLS para el análisis de espectroscopia de infrarrojo medio.

3.2.A.2.5. Comparación MIR y NIR para la determinación de TPM

Con el fin de poder establecer una comparación de los modelos PLS construidos a partir de los espectros NIR y MIR para la determinación del contenido de TPM en aceites fritos, se utilizó la misma serie de muestras para seleccionar los conjuntos de calibración y validación y construir los mejores modelos PLS. En la Tabla 2.5 pueden compararse los resultados de predicción de los modelos PLS-NIR y PLS-ATR-FTIR que ofrecieron la mayor capacidad predictiva.

Tabla 2.5. Parámetros de calibración y predicción de los modelos PLS para la determinación de TPM en muestras de aceite.

Modelo PLS	LV's	Conjunto de Calibración				Conjunto de Validación			
		RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	RRMSEP	R ² Pred	RPD
NIR	5	0.7	0.99	0.8	0.99	0.6	7.6	0.99	9.1
MIR	6	1.1	0.98	1.6	0.97	1.0	11.8	0.97	5.8

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Como se puede observar, los valores encontrados tanto para RMSEC como sobre todo para RMSECV en el modelo PLS-ATR-FTIR son mayores que los obtenidos en el modelo PLS-NIR, y siendo los valores R² Cal y R² CV de este último mejores que los encontrados en PLS-ATR-FTIR.

Cabe destacar que la espectroscopia NIR proporcionó mejores resultados que la espectroscopia MIR ya que se obtuvieron menores errores de predicción de 0.6 y 1.0 % m/m, que suponen unos errores relativos de 7.6 % y 11.8 %, respectivamente. Sin embargo, tanto NIR y MIR, podrían ser consideradas adecuadas para fines analíticos con valores RPD de 9.1 y 5.8, respectivamente.

Comparando los resultados obtenidos en este estudio con los encontrados por otros autores, Bansal *et al.* [5] reportaron un valor de RRMSEP de 10.1% para la determinación de TPM utilizando FT-IR en dos tipos de aceite de colza y un valor de RPD de 7.1, con 104 muestras para el conjunto de calibración y 52 muestras para el conjunto de validación. Estos resultados son ligeramente mejores que los encontrados en este estudio, pero hay que destacar que en nuestro trabajo se han

considerado tres tipos de aceite (girasol, aceite virgen y extra virgen de oliva) y mezclas de los mismos, lo que supone una mayor heterogeneidad de la población de muestras. Gertz *et al.* [9] encontraron valores de RMSEP de 1.97 % m/m y R^2 Pred de 0.97 utilizando FT-NIR para la determinación de TPM, mientras que nuestros resultados fueron de 0.6 y 0.99, respectivamente. Ögütcü *et al.* [10] aplicaron un modelo de calibración PLS para determinar TPM por espectroscopia NIR usando también una sonda capacitiva Testo 265 como método de referencia, obteniendo un valor R^2 Pred de 0.93 frente a 0.99 obtenido en el presente estudio. Büning-Pfaue *et al.* [12] utilizaron las mediciones de constante dieléctrica como método de referencia obteniendo una relación SD/RMSECV de aproximadamente 5, mientras que el valor encontrado en el estudio realizado ha sido de 10.6.

3.2.A.3. Conclusiones Capítulo 2.A

Los resultados encontrados a través de este estudio indican que las medidas de espectroscopia NIR y MIR por transmisión y modo ATR, tratadas respectivamente por PLS, podrían aplicarse como un método rápido y verde para la determinación de TPM en aceites de fritura con diferentes niveles de uso.

En este trabajo el conjunto de muestras de aceite de fritura medidas en el NIR fue dividido, mediante el algoritmo de Kennard-Stone, en un conjunto de calibración y otro de validación, formados por 60 y 15 muestras, respectivamente. Para poder comparar los modelos PLS-NIR y PLS-ATR-FTIR, los mismos conjuntos de muestras seleccionados para el NIR, fueron utilizados para construir los modelos PLS-ATR-FTIR. A pesar de que se encontraron 4 muestras anómalas para el caso del ATR y únicamente 1 para el NIR, estos estudios evidencian que ambas técnicas ofrecen una alternativa a los métodos tradicionales empleando equipos de laboratorio para la evaluación de calidad de muestras de aceite frito con bajos errores de predicción. Hay que destacar especialmente que la técnica NIR ofrece mejores resultados y mayores ventajas que el MIR ya que, a pesar de que requiere una mayor cantidad de muestra, se necesita un menor tiempo de medida y presenta una menor probabilidad de contaminación debido a que las muestras se encuentran en viales de vidrio y no precisan de limpieza del accesorio antes de la medida de cada muestra.

Capítulo 3.2.B. Estudios previos para el desarrollo de un instrumento de bajo coste para la evaluación de la calidad del aceite de fritura

Según el estudio previo, en el que se comparaba la espectroscopia NIR y MIR, y a pesar que de los resultados obtenidos se pueden considerar que ambas son aptas para la determinación del contenido de TPM en muestras de aceite frito, la espectroscopia NIR ha presentado mejores resultados, con valores más bajos de RRMSEP y un mejor RPD. Es por ello, que el siguiente estudio se ha realizado únicamente mediante el uso de la espectroscopia NIR. Además, el objeto de este estudio es establecer las condiciones espectrales para poder llevar a cabo la determinación de TPM con el menor número de variables espectrales (número de ondas), que puedan ser aplicados mediante el empleo de filtros ópticos o fuentes de diodo (LED's). Para ello se va utilizar la regresión lineal múltiple (MLR, del inglés *Multiple Linear Regression*) para predecir el contenido de TPM en muestras de aceite de fritura utilizando más de una variable explicativa; ya que esto va a ofrecer la ventaja de utilizar más información en la construcción del modelo y, consecuentemente, realizar estimaciones más precisas, pero sin la necesidad de utilizar demasiadas variables, como es el caso de los modelos de mínimos cuadrados parciales (PLS, del inglés *Partial Least Square*) el cual se utilizará para conocer los mejores resultados que se podrían obtener utilizando esta técnica.

3.2.B.1. Muestras

Para este estudio se utilizaron un total de 105 muestras constituidas por diferentes aceites de fritura y mezclas de los mismos, que se pueden clasificar en 4 grupos: i) 17 aceites fritos fueron suministrados por el Centro Superior de Investigación en Salud Pública (CSISP) de Valencia, de restaurantes y servicios de catering a centros colectivos, ii) 41 aceites fritos procedentes de uso doméstico, iii) 5 muestras de aceites comestibles sin uso y, iv) 42 muestras preparadas a partir de mezclas de los diferentes aceites fritos.

3.2.B.2. Resultados y discusión

3.2.B.2.1. Espectros NIR

Las muestras de aceite se transfirieron directamente a viales de vidrio de 1 ml con un diámetro interno de 6 mm. Los espectros se registraron en el intervalo comprendido entre 14000 y 3500 cm^{-1} usando el espectro de un vial de vidrio vacío como *background*, promediando 50 barridos y utilizando una resolución espectral de 4 cm^{-1} . Se obtuvieron tres medidas de cada muestra girando la posición del vial de muestra entre las repeticiones para asegurar una buena

representatividad de los espectros, calculándose el promedio de cada muestra para su posterior empleo en la construcción de los modelos de regresión.

Con el fin de simular el uso de instrumentos basados en filtros, se ha considerado una estrategia de reducción variables basada en el cálculo del promedio cada 10 puntos del espectro (aproximadamente equivalente a 20 cm^{-1}) proporcionando así una reducción variable desde 5440 a 544 puntos por espectro. En la Tabla 2.6 se representa, a modo de ejemplo, algunas variables del espectro original junto con las correspondientes al espectro de las variables reducidas, de modo que el espectro original tendría 5440 variables, mientras que el espectro reducido contiene únicamente 544 variables.

Tabla 2.6. Comparación del espectro original y del espectro con las variables reducidas de una muestra de aceite de fritura.

Número de onda (cm^{-1})		Absorbancia	
Original	Reducida	Original	Reducida
13997.14	13988.46	0.032753	0.032834
13995.21	13988.46	0.033026	0.032834
13993.28	13988.46	0.033008	0.032834
13991.35	13988.46	0.032814	0.032834
13989.43	13988.46	0.032733	0.032834
13987.50	13988.46	0.032874	0.032834
13985.57	13988.46	0.033015	0.032834
13983.64	13988.46	0.032833	0.032834
13981.71	13988.46	0.032589	0.032834
13979.78	13988.46	0.032698	0.032834
13977.85	13969.18	0.033058	0.033842
13975.93	13969.18	0.033478	0.033842
13974.00	13969.18	0.033872	0.033842
13972.07	13969.18	0.034066	0.033842
13970.14	13969.18	0.033925	0.033842
13968.21	13969.18	0.033574	0.033842
13966.28	13969.18	0.033460	0.033842
13964.35	13969.18	0.033887	0.033842
13962.43	13969.18	0.034454	0.033842
13960.50	13969.18	0.034642	0.033842
...
...
...
3525.07	3502.89	0.156622	0.1447535
3523.14	3502.89	0.150101	0.1447535
3521.21	3502.89	0.146917	0.1447535
3519.28	3502.89	0.140769	0.1447535
3517.32	3502.89	0.12605	0.1447535
3515.42	3502.89	0.118789	0.1447535
3513.49	3502.89	0.129918	0.1447535
3511.57	3502.89	0.151298	0.1447535
3509.64	3502.89	0.163934	0.1447535
3507.71	3502.89	0.163137	0.1447535

En la Figura 2.8 se compara el espectro original de una muestra de aceite de fritura con un 33.5% m/m de contenido de TPM con el espectro obtenido después de la reducción variable llevada a cabo según el procedimiento descrito en anteriormente.

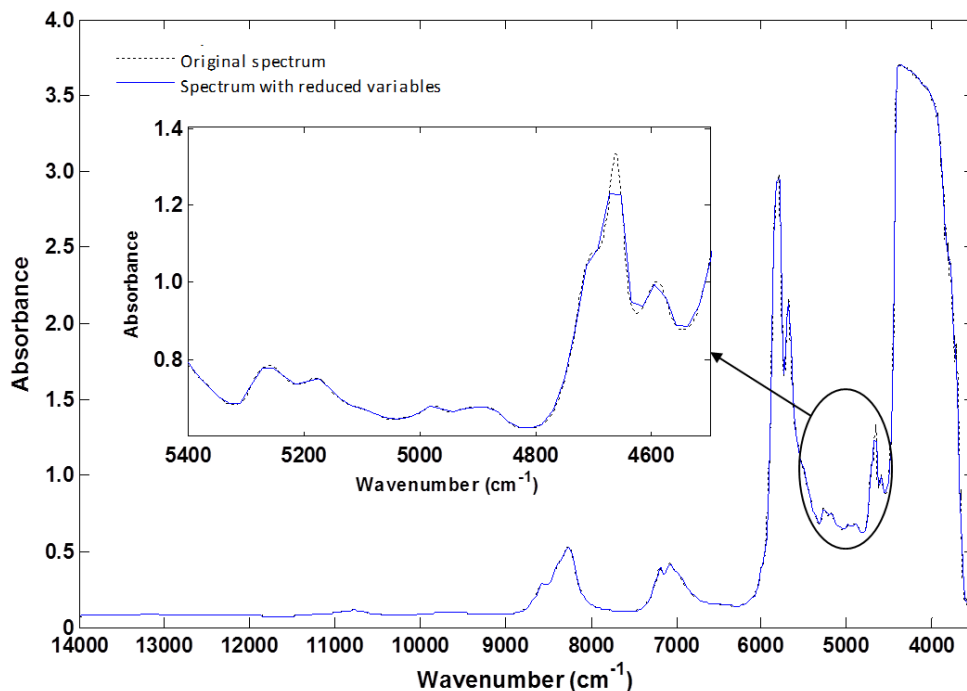


Figura 2.8. Comparación del espectro original de una muestra de aceite frito con una concentración de 33,5% m/m TPM con la obtenida después de la reducción de variables. Inserto: Detalle de las diferencias en la región de 5400-4500 cm^{-1} .

3.2.B.2.2. Modelización

Antes de construir los modelos de calibración mediante el empleo MLR o PLS, la población total de muestras de aceite se dividió en un subconjunto de calibración y otro de validación, con 73 y 32 aceites, respectivamente, seleccionados mediante el algoritmo de Kennard-Stone. Se van a utilizar los mismos conjuntos de calibración y validación en todos los supuestos que se van a considerar en este estudio, pudiendo encontrar las características de los mismos en la Tabla 2.7. Cabe decir, que las muestras incluidas en el conjunto de calibración son representativas de los diferentes grupos de aceite y cubren todo el intervalo de contenido de TPM.

Tabla 2.7. Características de los datos de referencia en los conjuntos de datos de calibración y validación empleados para construir los modelos MLR y PLS

Características	Conjunto	Número de muestras	TPM (%) (m/m)	
			Intervalo de concentración	Promedio \pm sd
MLR y PLS	Calibración	73	2.7 – 33.5	15 \pm 8
	Validación	32	3.5 – 28.5	12 \pm 7

sd: desviación estándar

3.2.B.2.2.1. MLR y PLS con reducción de variables

3.2.B.2.2.1.1. Modelos MLR-NIR

Para seleccionar las variables espectrales a partir de los espectros reducidos, se ha considerado que los espectros están divididos en 544 intervalos, cada uno de los cuales incluye 1 variable. Se construyeron 3 modelos diferentes mediante la validación cruzada "*leave-one-out*" considerando como máximo sólo 4, 3 ó 2 variables, usando en todos los casos el centrado en la media como tratamiento previo de los datos y la selección de variables paso a paso mediante forward. Los números de onda que proporcionaron los mejores resultados para el modelo MLR al considerar 4 variables fueron: 5792, 4808, 4634 y 4577 cm^{-1} . En el caso del modelo MLR con 3 variables, los números de onda seleccionados fueron 5792, 4808 y 4577 cm^{-1} , mientras que para los modelos MLR de 2 variables, los valores obtenidos fueron 5792 y 4808 cm^{-1} .

Los principales parámetros de calibración correspondientes a los tres modelos MLR-NIR desarrollados se resumen en la Tabla 2.8, indicando también los números de onda utilizados para cada uno de los modelos construidos, así como los valores de RMSEC, RMSECV, R^2 Cal y R^2 CV.

Los tres modelos MLR-NIR seleccionados fueron evaluados mediante el uso del conjunto de validación independiente formado por las restantes 32 muestras de aceite de fritura. Los principales parámetros de predicción de estos modelos MLR-NIR se resumen en la Tabla 2.8, donde se puede observar que se obtuvieron buenos resultados para el contenido de TPM de modelos MLR correspondientes a 3 y 4 variables, con valores de RRMSEP de 18.1 y 12.1% con R^2 pred por encima de 0.93.

Tabla 2.8. Parámetros de calibración y predicción de modelos MLR construidos para la determinación de TPM en muestras de aceite frito

Modelo	Pre-tratamiento	Número de onda (cm ⁻¹)/(LV's)	Conjunto de calibración				Conjunto de validación			
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² CV	RMSEP (%)	RRMSEP (%)	R ² Pred	RPD
MLR Variables reducidas	Centrado en la media	5792, 4808	4.4	0.69	4.7	0.65	2.8	23.0	0.86	2.6
		5792, 4808, 4577	3.1	0.85	3.6	0.83	2.2	18.1	0.93	3.3
		5792, 4808, 4634, 4577	2.4	0.91	2.6	0.89	1.5	12.1	0.96	5.0

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Los valores de RPD fueron de 3.3 y 5.0 para 3 y 4 variables, respectivamente, lo que indica que pueden utilizarse para la determinación cuantitativa de TPM en aceite de fritura, siendo el modelo MLR con 4 variables muy bueno. Para el modelo MLR con 2 variables, se obtuvieron resultados no satisfactorios con un valor de RRMSEP de aproximadamente 23% y un valor de RPD por debajo de 3.0, lo que indicaría que un número tan reducido de variables no resultaría adecuado para la correcta predicción del contenido de TPM en muestras de aceite de fritura

Obviamente cuanto mayor sea el número de variables considerado, la capacidad predictiva del modelo MLR-NIR aumenta, por lo que cabe esperar que un número de variables superior al máximo de 4 considerado proporcionaría aún mejores resultados. Sin embargo, considerando que el propósito de este estudio es el de conocer las condiciones espectrales que permitieran construir un instrumento de bajo coste, se ha considerado que 4 debería ser el número máximo de números de onda para no implicar un diseño excesivamente complejo del instrumento de medida.

3.2.B.2.2.1.2. Modelos PLS-NIR

En este caso se construyeron dos modelos PLS con diferentes regiones espectrales (número de variables) y pre-tratamiento basados en las variables reducidas. El primer modelo corresponde con el uso de sólo 4 variables para la construcción del modelo, con el fin de poder comparar su capacidad de predicción con la del mejor modelo MLR calculado en el apartado anterior. El segundo modelo, fue construido considerando intervalos a partir de las 544 variables para obtener la mejor capacidad predictiva mediante el uso de PLS con los espectros NIR con reducción de variables.

Para construir el modelo PLS con 4 variables, se empleó el iPLS, considerando el espectro reducido como 544 sub-intervalos equidistantes de 1 variable y usando 10LVs como máximo para la construcción del modelo. Como pre-tratamiento de los espectros se utilizó el centrado en la media. Los cuatro números de onda seleccionados por iPLS fueron 5792, 4828, 4808 y 4577 cm^{-1} , que coinciden con algunos de los seleccionados para la construcción de los modelos MLR.

Para construir el mejor modelo PLS en los espectros reducidos se consideraron intervalos espectrales en lugar de números de onda individuales, y para ello la región de espectro entre 14000 y 4356 cm^{-1} se dividió en 10 sub-intervalos equidistantes de 50 variables reducidas cada uno de ellos, utilizando 10 LV's como máximo para la obtención del modelo PLS. Se estudiaron diferentes pre-tratamientos de datos así como combinaciones de intervalos espectrales para

obtener el error mínimo de validación cruzada. El empleo de la primera derivada junto con el centrado en la media como pre-tratamiento espectral, proporcionó los mejores resultados, considerando como óptimas 5LV's y las regiones espectrales comprendidas entre 8203 y 5329 cm^{-1} .

Los principales parámetros de calibración de los modelos PLS-NIR considerando los espectros con variables reducidas, se resumen en la Tabla 2.9, donde se puede ver el número de variables latentes, y los valores de RMSEC, RMSECV, R^2 Cal y R^2 CV obtenidos para ambos modelos PLS-NIR obtenidos a partir de los espectros con variables reducidas.

El conjunto de validación externa se utilizó para confirmar la idoneidad de los modelos construidos, encontrándose en la Tabla 2.9, los principales parámetros de predicción. En los dos casos se obtuvo una buena capacidad predictiva para los modelos PLS establecidos. Para el modelo PLS construido con 4 variables, se obtuvo un valor RRMSEP del 14.4% con R^2 pred de 0.95. En el caso del modelo PLS construido con 5 LV's en la región entre 8203 y 5329 cm^{-1} , el RRMSEP obtenido fue del 7.0% con R^2 Pred de 0.99. Los valores de RPD encontrados fueron de 4.2 y 8.5, respectivamente.

A partir de los resultados obtenidos para estos modelos PLS-NIR, de acuerdo con los criterios reportados, ambos modelos de calibración pueden usarse para la determinación cuantitativa de TPM en aceite de fritura, siendo el modelo PLS basado en 5 LV's y en el rango comprendido entre 8203 y 5329 cm^{-1} , un excelente modelo para ser usado en cualquier aplicación.

3.2.B.2.2. Modelo PLS-NIR sin variables reducidas

El modelo PLS-NIR sin reducción de variables fue construido con el fin de evaluar cual sería la mejor capacidad predictiva que se podría obtener a partir de los espectros NIR considerando la totalidad de las variables espectrales.

Para ello, el iPLS se empleó para la selección de regiones espectrales mediante la división del espectro en sub-intervalos equidistantes de 100 variables y utilizando 10 LV's como máximo para la construcción de los modelos. Los mejores resultados se obtuvieron considerando como pre-tratamiento de los datos la primera derivada y el centrado en la media en las regiones consideradas entre 12840-12456, 8983-8599, 7826-7442 y 5126-4742 cm^{-1} . Los principales parámetros de calibración y validación cruzada del modelo PLS-NIR desarrollado se resumen en la Tabla 2.10. Como en los casos anteriores, el conjunto de validación fue utilizado para el estudio de la capacidad predictiva del modelo construido.

Tabla 2.9. Parámetros de calibración y predicción de modelos PLS con variables reducidas construidos para la determinación de TPM en muestras de aceite frito

Modelo	Pre-tratamiento	Número de onda (cm ⁻¹)/(LV's)	Conjunto de calibración			Conjunto de validación				
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² CV	RMSEP (%)	R ² Pred	RPD	
PLS Variables reducidas	Centrado en la media	5792, 4828, 4808, 4577 (4)	2.3	0.92	2.5	0.90	1.7	14.4	0.95	4.2
	Centrado en la media, Primera derivada	8203-5329 (5)	1.1	0.98	1.4	0.97	0.9	7.0	0.99	8.5

Tabla 2.10. Parámetros de calibración y predicción de modelos PLS construidos para la determinación de TPM en muestras de aceite frito

Modelo	Pre-tratamiento	Número de onda (cm ⁻¹)/(LV's)	Conjunto de calibración			Conjunto de validación				
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² CV	RMSEP (%)	R ² Pred	RPD	
PLS	Centrado en la media,	12840-12456, 8983-8599, 7826-7442 5126-4742 (6)	1.0	0.98	1.4	0.97	0.8	6.5	0.99	9.2
	Primera derivada									

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Los principales parámetros del modelo PLS-NIR desarrollado se indican también en la Tabla 2.10, donde se observa que el valor de RRMSEP se encuentra por debajo del 6.5% con valores de R^2 Cal, R^2 CV y R^2 Pred de 0.98, 0.97 y 0.99, respectivamente, indicando una buena precisión de este modelo PLS. El valor RPD encontrado para el modelo seleccionado fue de 9.2, que puede ser considerado adecuado para fines analíticos ya que supone una buena capacidad predictiva.

3.2.B.2.3. Comparación de los diferentes modelos de calibración

En la Tabla 2.11 se resumen los resultados obtenidos para los diferentes modelos de calibración MLR-NIR y PLS-NIR considerados en este estudio. Como puede verse, los mejores resultados obtenidos para la determinación de TPM en muestras de aceites de fritura por espectroscopia NIR fueron encontrados por PLS-NIR cuando se consideró el espectro NIR original, sin reducción en el número de variables. El uso de este modelo ofrece una alternativa para ser utilizado con equipos NIR de laboratorio para la determinación de TPM en aceites empleados para freír diferentes tipos de alimentos.

Con el fin de comparar los mejores resultados obtenidos en este estudio empleando el espectro NIR completo con los encontrados por otros autores, que investigan las posibilidades de utilizar NIR para este propósito, se calculó la relación entre la desviación estándar de los datos de calibración (SD) y el valor RMSECV. En nuestro estudio, la relación SD/RMSECV y los valores de R^2 pred fueron de 5.8 y 0.99, respectivamente, que se compara adecuadamente con los valores de aproximadamente 5 y 0.98 obtenidos por Büning-Pfaue *et al.* [12] que utilizó medidas de la constante dieléctrica como método de referencia. Por otra parte, Ögütcü *et al.* [10] aplicaron un modelo de calibración PLS para determinar TPM por espectroscopia NIR utilizando la sonda capacitiva Testo 265 como método de referencia, obteniendo un valor de R^2 pred de 0.93.

Para evaluar la posibilidad de diseño de un instrumento NIR de bajo coste para determinar el contenido de TPM en aceites de fritura, se puede considerar el uso de un modelo MLR-NIR basado en 4 números de onda seleccionados como la mejor opción, obteniéndose un valor RRMSEP de 12.1% y RPD de 5.0. En estas condiciones de trabajo, el contenido de TPM en aceites de fritura predicho por MLR-NIR mostró una correlación significativa, obteniéndose un valor R^2 Pred de 0.96 frente al 0.90 encontrado por Khaled *et al.* [13] que utiliza un sensor capacitivo para la obtención de los datos de referencia. Por lo tanto, la calibración MLR-NIR de los datos NIR encontrados en 5792, 4808, 4634 y 4577 cm^{-1} podrían utilizarse para la determinación cuantitativa de TPM en muestras de aceite de fritura con unos resultados adecuados.

Tabla 2.11. Comparación de los parámetros de calibración y predicción de los modelos MLR y PLS construidos para la determinación de TPM en muestras de aceite frito

Modelo	Pre-tratamiento	Número de onda (cm ⁻¹)/(LV's)	Conjunto de calibración				Conjunto de validación			
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² CV	RMSEP (%)	RRMSEP (%)	R ² Pred	RPD
MLR Variables reducidas		5792, 4808	4.4	0.69	4.7	0.65	2.8	23.0	0.86	2.6
	Centrado en la media	5792, 4808, 4577	3.1	0.85	3.6	0.83	2.2	18.1	0.93	3.3
		5792, 4808, 4634, 4577	2.4	0.91	2.6	0.89	1.5	12.1	0.96	5.0
PLS Variables reducidas	Centrado en la media	5792, 4828, 4808, 4577 (4)	2.3	0.92	2.5	0.90	1.7	14.4	0.95	4.2
	Centrado en la media, Primera derivada	8203-5329 (5)	1.1	0.98	1.4	0.97	0.9	7.0	0.99	8.5
	Centrado en la media, Primera derivada	12840-12456, 8983-8599, 7826-7442 5126-4742 (6)	1.0	0.98	1.4	0.97	0.8	6.5	0.99	9.2

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

3.2.B.3. Conclusiones Capítulo 2

El uso de modelos quimiométricos aplicado a los espectros NIR de transmisión de aceites de fritura usados permite, la determinación exacta de TPM con valores de RRMSEP de 6.5% y RPD de 9.2. Además el modelo MLR-NIR construido a partir de espectros NIR con variables reducidas proporcionó una buena correlación entre los 4 números de onda seleccionados y el contenido de TPM medido por el método de referencia, con un valor R^2_{pred} de 0.96 y un RPD de 5.0. Estos resultados son adecuados para fines analíticos con una precisión de predicción excelente pudiendo ser utilizados para diseñar un instrumento de bajo coste para una evaluación rápida y barata de la calidad del aceite de fritura. Estos instrumentos podrían fabricarse con un sistema de filtros o LED's que permitieran trabajar a los cuatro números de onda seleccionados, obteniéndose así un equipo portátil, de bajo coste, para la evaluación del aceite de fritura a partir de la estimación del contenido de TPM. Además, las medidas directas se realizarían directamente sobre las muestras contenidas en un vial de vidrio, evitando así la contaminación cruzada entre las mismas.

3.2.3. Referencias Capítulo 2

- [1] S. Casal, R. Malheiro, A. Sendas, B. P. P. Oliveira, J.A. Pereira, Olive oil stability under deep-frying conditions, *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.*, 48(2010) 2972–2979.
- [2] M.C. Dobarganes, J. Velasco, A. Dieffenbacher, Determination of polar compounds, polymerized and oxidized triacylglycerols, and diacylglycerols in oils and fats: results of collaborative studies and the standardized method (Technical report), *Pure Appl. Chem.*, 72(2000) 1563–1575.
- [3] Quality Standard for Oils and Fats Heated, dated 26th of January of 1989.
- [4] D. Firestone (Ed), *Official Methods and Recommended Practices of the AOCS*, 6th Edn., 3rd printing, 2013. Determination of Polar Compounds in Frying Fats. AOCS Official Method Cd 20-91.
- [5] G. Bansal, W. Zhou, P.J. Barlow, P.S. Joshi, H.L. Lo, Y. K. Chung, Review of Rapid Tests Available for Measuring the Quality Changes in Frying Oils and Comparison with Standard Methods, *Crit. Rev. Food Sci. Nutr.*, 50(2010) 503–514.
- [6] M. Saltmarsh, Rapid analysis of fats and oils: thermal degradation, Presented at Western Food Industries Conference 1990; UC Davis.
- [7] J.Y. Chen, H. Zhang, J. Ma, T. Tuchiya, Y. Miao, Determination of the Degree of Degradation of Frying Rapeseed Oil Using Fourier-Transform Infrared Spectroscopy Combined with Partial Least-Squares Regression, *Int. J. Anal. Chem.*, 2015(2015) e185367.
- [8] J. Moros, M. Roth, S. Garrigues, M. de la Guardia, Preliminary studies about thermal degradation of edible oils through attenuated total reflectance mid-infrared spectrometry, *Food Chem.*, 114(2009) 1529–1536.
- [9] C. Gertz, H.J. Fiebig, J.N.S. Hancock, FT-near infrared (NIR) spectroscopy–Screening analysis of used frying fats and oils for rapid determination of polar compounds, polymerized triacylglycerols, acid value and anisidine value [DGF C-VI 21a (13)], *Eur. J. Lipid Sci. Technol.*, 115(2013) 1193–1197.

- [10] M. Ögütcü, B. Aydeniz, M.B. Büyücan, E. Yilmaz, Determining Frying Oil Degradation by Near Infrared Spectroscopy Using Chemometric Techniques, *J Am Oil Chem Soc*, 89(2012) 1823–1830.
- [11] J. Kuligowski, D. Carrión, G. Quintás, S. Garrigues, M. de la Guardia, Direct determination of polymerised triacylglycerides in deep-frying vegetable oil by near infrared spectroscopy using Partial Least Squares regression, *Food Chem.*, 131(2012) 353–359.
- [12] H. Büning-Pfaue, S. Kehraus, Application of near infrared spectroscopy (NIRS) in the analysis of frying fats, *Eur. J. Lipid Sci. Technol.*, 103(2001) 793–797.
- [13] A.Y. Khaled, S.A. Aziz, F.Z. Rokhani, Capacitive sensor probe to assess frying oil degradation, *Inf Process Agric*, 2(2015)142–148.
- [14] S. Kazemi, N. Wang, M. Ngadi, S.O. Prasher, Evaluation of Frying Oil Quality Using VIS/NIR Hyperspectral Analysis, *International Commission of Agricultural Engineering (CIGR, Commission Internationale du Genie Rural) E-Journal*, 7(2005).
- [15] C.C. Osawa, L.A.G. Gonçalves, H.F. Gumerato, F.M. Mendes, Study of the effectiveness of quick tests based on physical properties for the evaluation of used frying oil, *Food Control*, 26(2012)525–530.
- [16] N. Sinelli, L. Cerretani, V. di Egidio, A. Bendini, E. Casiraghi, Application of near (NIR) and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity, *Food Res. Int.*, 43(2010) 369–375.
- [17] S. Armenta, S. Garrigues, M. de la Guardia, Determination of edible oil parameters by near infrared spectrometry, *Anal. Chim. Acta*, 596(2007)330–337.

3.3. CAPÍTULO 3:
Determinación de lípidos en aceite de salmón

3.3. Capítulo 3. Determinación de lípidos en aceite de salmón

3.3.1. Introducción

La determinación de los lípidos y la evaluación de la composición de los ácidos grasos en los alimentos resultan de gran importancia para la salud humana. Los lípidos se clasifican generalmente en lípidos no polares (triglicéridos (TAG), diglicéridos (DAG), monoglicéridos (MAG) y esteroides) y lípidos polares (ácidos grasos libres (FFA), fosfolípidos y esfingolípidos) [1], siendo el impacto de ácidos grasos (FA) sobre los efectos beneficiosos de la salud humana uno de los temas más investigados. El pescado y el marisco son fuentes importantes de ácidos grasos poliinsaturados de cadena larga del tipo Omega-3 (LCn-3PUFA, del inglés *Long Chain Omega 3 Polyunsaturated Fatty Acids*). Se ha comprobado que una elevada ingesta de LCn-3PUFAs produce efectos beneficiosos sobre la salud humana, reduciendo el impacto de las enfermedades cardiovasculares, la diabetes, el cáncer, la obesidad, las enfermedades autoinmunes, la artritis reumatoide, el asma e, incluso, la depresión [2-4].

Los ácidos LCn-3PUFA más importantes son el ácido eicosapentaenoico (EPA, C20: 5n-3) y el ácido docosahexaenoico (DHA, C22: 6n-3) [5], junto con el ácido docosapentaenoico (DPA, C22: 5n-3) que también representa una relativa importancia [6]. Cabe indicar que el precursor tanto del EPA como del DHA es el ácido α -linolénico (ALA, C18: 3n-3) [7]. Por otra parte, las altas ingestas de ácidos grasos poliinsaturados de cadena larga del tipo Omega-6 (LCn-6PUFA) tienen propiedades opuestas a las de los LCn-3PUFAs [3], [4], siendo el ácido linoleico (LA; C18: 2n-6) el precursor del ácido araquidónico (AA, 20: 4n-6), considerado uno de los principales ácidos grasos de la familia de los omega 6 [7].

A diferencia de los ácidos grasos Omega-3 y Omega-6, la ingesta de ácidos grasos del tipo Omega-9 no resulta tan esencial ya que los seres humanos son capaces de sintetizar uno de los principales Omega 9, el ácido oleico (OA, C18: 1n-9), a partir del ácido esteárico [8]. El OA es el principal ácido graso monoinsaturado (MUFA, del inglés *Monounsaturated Fatty Acid*) presente en el aceite de salmón y, tal como se ha indicado, pertenece a la familia de los ácidos grasos Omega-9. En la literatura científica algunos autores constatan que existe una correlación entre el alto contenido de MUFA del aceite de pescado con los efectos cardioprotectores que presenta [9]. Por el contrario, Wu *et al.* [10] demostraron que el ácido palmítico (PA, C16: 0) y el OA afectaban al colesterol y al metabolismo de los ácidos grasos de una manera opuesta. Por su parte, el PA es el principal ácido graso saturado (SFA, del inglés *Saturated Fatty Acid*) presente en el aceite de salmón. Se ha podido comprobar que las dietas ricas en PA, promueven la desregulación hipotalámica y contribuyen a la obesidad a través de la interrupción

de las señales reguladoras del apetito [11]. Los SFA afectan a los mecanismos implicados en el metabolismo del colesterol [12]. Por tanto, por las razones antes mencionadas y desde el punto de vista de la salud humana, es necesario un conocimiento más detallado de la presencia y concentración de los diferentes tipos o familias de lípidos en el aceite de pescado.

A pesar de su importancia, la determinación de los lípidos en alimentos no puede realizarse directamente, sino que es necesaria una extracción previa. Un método de extracción comúnmente usado para determinar los lípidos en alimentos es el método de Soxhlet [13], utilizando n-hexano como disolvente [14]. El n-hexano es un disolvente con bajo punto de ebullición, baja polaridad y estabilidad química, sin embargo se obtiene a partir de petróleo, un recurso no renovable. Este disolvente se clasifica en el Reglamento de Registro, evaluación, autorización y restricción de sustancias químicas (REACH, del inglés *Registration, Evaluation, Authorisation and restriction of CHemicals*) como Categoría 2 de tóxicos para ambientes acuáticos y para la reproducibilidad [15]. La Agencia de Protección Ambiental de los Estados Unidos de América (EPA, del inglés *Environmental Protection Agency*) regula el n-hexano bajo la Ley de Aire Limpio (CAA, del inglés *Clean Air Act*) y lo ha clasificado como un contaminante del aire peligroso (HAP, del inglés *Hazardous Air Pollutants*). El n-hexano se evapora muy fácilmente al aire y se disuelve muy ligeramente en el agua [16]. Por lo tanto, para prevenir la contaminación producida por el n-hexano es necesario reemplazarlo por otro tipo de disolvente que resulten más inocuos y menos peligrosos. Se han realizado varios intentos para encontrar disolventes alternativos que permitan reemplazar el hexano como disolvente cuando se emplea para realizar las extracciones o como disolvente para los análisis químicos [17-22]. En ese sentido se ha contemplado la posibilidad de sustituirlo por disolventes con baja toxicidad, fácil de reciclar, inertes y que no contamine el producto, es decir por un disolvente considerado verde.

Particularizando para el proceso de separación de las grasas presentes en pescados, una vez realizada la extracción de los lípidos con n-hexano, el disolvente es evaporado mediante el uso de un rotavapor y, posteriormente, se hace uso de las técnicas cromatográficas para determinar la composición de los ácidos grasos y de los lípidos presentes en el extracto obtenido.

Las técnicas cromatográficas habitualmente empleadas para el estudio de los ácidos grasos y lípidos presentes en la fracción extraída son:

- Cromatografía en capa fina de alto rendimiento (HPTLC), que se utiliza para la separación, identificación de los lípidos individuales y determinación cuantitativa en los alimentos. Su uso en el análisis de lípidos implica los

siguientes pasos: i) preparación de la muestra, ii) acondicionamiento de la placa de TLC, iii) aplicación de la muestra en la placa de TLC, iv) preparación de la fase móvil, v) inmersión de la placa con la fase estacionaria en la fase móvil, vi) desarrollo cromatográfico del eluyente automático, vii) secado final, viii) revelado y visualización y, ix) evaluación cuantitativa de los datos densiométricos generados.

- Cromatografía de gases (GC), que se usa para la determinación individual de los ácidos grasos. Antes del análisis por GC, es necesaria una etapa previa en la que los ácidos grasos se liberan de los glicéridos y se derivatizan, convirtiéndolos en sus respectivos ésteres metílicos (FAME), que son más volátiles que los FFA.

Ambas técnicas cromatográficas requieren un gran consumo de tiempo y de dedicación del operador, así como de energía, reduciendo su aplicabilidad para el control de calidad de una manera sostenible, según los principios de la Química Analítica Verde. Además el trabajo con las placas de TLC presentan otros inconvenientes, como la posible oxidación de los lípidos debido a la exposición de la placa al oxígeno atmosférico, así como la limitación de las aplicaciones preparativas en la placa.

Considerando la importancia que supone el encontrar alternativas más verdes al uso de disolventes como el n-hexano y centrándonos en el estudio sobre el extracto de aceite procedente del salmón, en este capítulo se van a considerar las siguientes aplicaciones:

- **3.A.** Estudio de la extracción de los lípidos del salmón mediante disolventes alternativos al n-hexano.
- **3.B.** Desarrollo de un método analítico verde para la determinación rápida de los ácidos grasos y de las clases de lípidos mediante el uso de la espectroscopia en el infrarrojo cercano.

Por tanto, los objetivos de estos dos trabajos son:

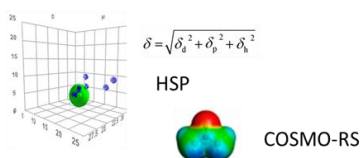
- Investigar y evaluar el efecto que supone reemplazar el n-hexano por disolventes alternativos para la extracción de la fracción lipídica del salmón. Para ello, se llevaron a cabo diversos estudios sobre la solubilidad relativa de los componentes grasos del pescado (salmón, *Salmon Salar*) en diferentes disolventes mediante el empleo de 2 métodos computacionales, como son: el *Hansen Solubility Parameter* (HSP) y el *Conductor-like Screening Model for Real Solvent* (COSMO-RS). Además se realizaron estudios experimentales para comparar los disolventes alternativos empleados con el hexano. En la Figura 3.1 se ilustra gráficamente el procedimiento de este trabajo, dividido en las

siguientes tres etapas: i) estudios computacionales mediante el uso de HSP y Cosmo-RS sobre la solubilidad relativa de los componentes grasos del salmón en diferentes disolventes, ii) comparación de la capacidad de extracción de los distintos disolventes considerados a partir de la determinación de las clases de lípidos extraídos y medidos mediante HPTLC y la cuantificación de los ácidos grasos a partir de sus derivados metilados mediante el uso de la CG y, iii) interpretación de los resultados mediante el uso de herramientas quimiométricas como el PCA o el análisis de agrupamiento jerárquico (HCA, *hierarchical cluster analysis*).

Soxhlet extraction



Computational methods: Solvent selection



Experimental study: Determination



Chemometric tools: Interpretation



Figura 3.1. Esquema del procedimiento experimental seguido en este estudio.

- Evaluar la técnica NIR, en combinación con los modelos PLS, para la determinación simultánea de i) ácidos grasos como OA, PA, ALA, LA, ii) familias de ácidos grasos como Omega-3, Omega-6, PUFAs, MUFAs y SFAs, y iii) clases de lípidos como TAGs, DAGS, FFAs y ergosterol en los extractos de aceite de salmón. La Figura 3.2 ilustra la idea general que se persigue en este trabajo, tratando de aportar las ventajas que supone el empleo de la espectroscopia NIR, desde el punto de vista de la sencillez y rapidez, para la determinación de los ácidos grasos y las clases de los lípidos mencionados anteriormente, haciendo uso de las técnicas quimiométricas. Finalmente se hace referencia una la comparación desde el punto de vista de la sostenibilidad en Química

Analítica Verde del procedimiento propuesto frente los métodos tradicionales basados en el empleo de la cromatografía.

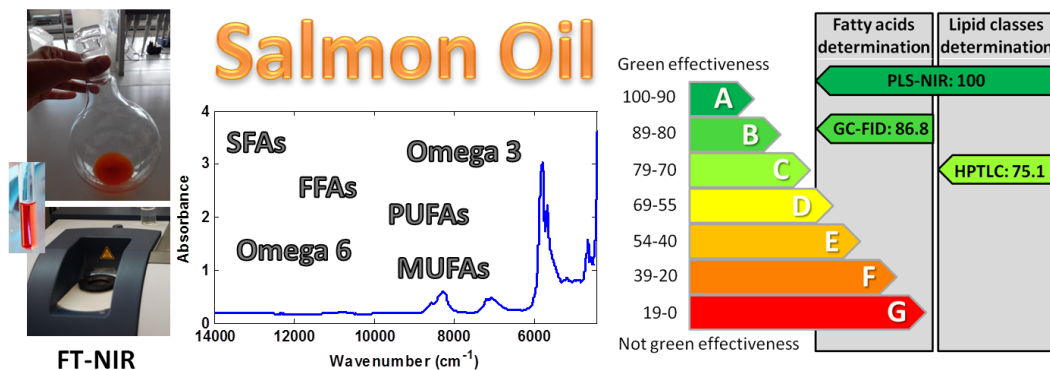


Figura 3.2. Representación gráfica del procedimiento empleado y comparación con los métodos tradicionales

3.3.2. Parte experimental

3.3.2.1. Muestras

En este estudio se han empleado muestras de salmón atlántico (*Salmon salar*) que se obtuvieron en el mercado local. Una vez adquiridas se prepararon adecuadamente en el laboratorio y se almacenaron a -20°C hasta su posterior utilización. Las muestras congeladas se liofilizaron, durante un mínimo de 48 horas a una presión de cámara de 0.053 mbar, utilizando un liofilizador Cosmos de la marca Cryotec (Saint-Gély-du-Fesc, Francia). Las muestras liofilizadas se pulverizaron con un mezclador doméstico Blixer 2 de Robot Coupe (Montceau-en-Bourgogne, Francia) y se almacenaron en bolsas herméticas hasta la extracción. Para poder comparar la capacidad de extracción de cada disolvente, se mezclaron todas las muestras de salmón para obtener así una cantidad suficiente, homogénea y representativa de la totalidad del salmón utilizado en este estudio. La pérdida de agua durante el proceso de liofilización nos indicó que el contenido de agua en el salmón original era del 68.5% m/m.

3.3.2.2. Extracción del aceite de salmón: Método Soxhlet

Veinticinco gramos de salmón liofilizado se extrajeron a reflujo, durante 8 h, con 300 ml del disolvente considerado. Después de este tiempo, los restos de disolvente encontrado junto al extracto en el matraz de destilación, se evaporaron a presión reducida utilizando un rotavapor Laborota 4000 de Heidolph (Schwabach, Alemania).

3.3.2.3. Métodos de referencia

3.3.2.3.1. HP-TLC

La determinación de las clases de lípidos extraídos por los disolventes se realizó mediante el uso de un sistema automatizado de cromatografía de capa fina de alto rendimiento (HPTLC) de CAMAG (Muttentz, Suiza).

Las soluciones estándar se prepararon usando cloroformo como disolvente, con una concentración de 0.2 mg mL^{-1} de lípido en cada una de las disoluciones, empleando una mezcla conocida de los fosfolípidos de soja, compuesta por L-A lisofosfatidilcolina, fosfatidilcolina, fosfatidiletanolamina y fosfatidilinositol, suministrada por Supelco (Bellefonte, EE.UU.), patrones de ergosterol (ergo), ácidos grasos libres (FFA, ácido palmítico), monoglicerol (MAG, DL- α -palmitina), diacilglicerol (DAG, 1,3-dipalmitato de glicerilo) y triacilglicerol (TAG, tripalmitato de glicerilo) de Sigma-Aldrich (Taufkirchen, Alemania). En cuanto las muestras, se preparó aproximadamente 1 mg de extracto puro en 1 ml de cloroformo. Todas las soluciones se almacenaron en la oscuridad a $-20 \text{ }^{\circ}\text{C}$ hasta su análisis.

Los lípidos se detectaron mediante un densitómetro de barrido TLC de CAMAG 3, utilizando para la identificación de las clases, los patrones conocidos de lípidos. El extracto lipídico del salmón fue introducido automáticamente sobre placas de TLC de 20x10 cm de gel de sílice 60 F254 suministradas por la empresa Merck (Darmstadt, Alemania), utilizando un muestreador automático TLC ATS5 de CAMAG (Muttentz, Suiza). Las placas TLC fueron previamente acondicionadas con una mezcla de cloroformo/metanol (2:1, v/v) y secadas a una temperatura de $110 \text{ }^{\circ}\text{C}$ durante 60 min utilizando un calentador de placas TLC (CAMAG, Muttentz, Suiza).

Para separar los lípidos polares se utilizó una mezcla de acetato de metilo/isopropanol/cloroformo/metanol/KCl (solución al 0.25%) en una relación de 25:25:25:10:9 (v/v/v/v/v), mientras que para la separación de los lípidos neutros se empleó una mezcla de n-hexano/éter dietílico/ácido acético glacial en una proporción de 70:30:2 (v/v/v). En ambos casos el desarrollo cromatográfico del eluyente se realizó hasta una altura de 7cm. Una vez secadas, las placas se sumergen durante 2 segundos en una disolución de 10 mg de primulina en 160 ml de acetona y 40 ml de agua. A continuación las placas se escanearon utilizando un TLC Scanner 3 con el software WinCATs de CAMAG. La identificación y cuantificación de las clases de los lípidos se realizó mediante una curva de calibración de cada una de las moléculas testigo, correlacionando las áreas de pico con la cantidad de clases de lípidos. Los valores obtenidos para los lípidos totales se expresaron como porcentaje m/m de cada una de las clases de lípidos

consideradas. El factor de retención (Rf) y el área de los picos se calcularon utilizando el software winCATS-Planar Chromatography Manager (versión 1.3.3) de CAMAG.

3.3.2.3.2. Cromatografía de gases

Los ésteres metílicos de los ácidos grasos (FAME) se separaron e identificaron mediante un sistema de cromatografía de gases (GC) 7820 A de Agilent Technologies (California, EE.UU.), equipado con un detector de ionización de llama (GC-FID) y un inyector automático. El análisis se realizó empleando una columna capilar BD-EN14103 (30 m × 0,32 mm × 0,25 μm), usando helio como un gas portador a la velocidad de 33 cm s⁻¹. Se inyectaron 2 μl de muestras en modo *split* (1:20) a una temperatura de 250 °C, con el programa de temperatura del horno de que se indica a continuación: i) temperatura inicial a 50° C durante un minuto, ii) incremento de la temperatura a una velocidad de 20 °C por minuto hasta un valor de 180 °C, iii) incremento de la temperatura hasta 230 °C a una velocidad de 2 °C por minuto, y iv) mantenimiento de la temperatura de 230 °C durante 10 min.

Los extractos fueron preparados para su análisis mediante la trans-metilación siguiendo el procedimiento del trabajo de Morrison *et al.* [23]. Una disolución de 400 μl de disolución de triheptadecanoin (TAG, C17:0) de Sigma Aldrich (Taufkirchen, Alemania) que contenía 2 mg mL⁻¹ en n-hexano, se utilizó como patrón interno, adicionándola a aproximadamente 20 mg del extracto de aceite de salmón. A continuación se le añadió 1 ml de ácido sulfúrico metanólico (5% v/v) y la mezcla homogenizada se calentó a 85 °C en una placa durante 90 minutos en un tubo cerrado. Una vez transcurrido el tiempo y alcanzada la temperatura ambiente, se añadieron 1.5 ml de una solución de cloruro de sodio al 0.9% m/v y 1 ml de n-hexano. El tubo se agitó vigorosamente durante 30 s antes de la centrifugación de la mezcla resultante a 4000 rpm durante 2 min. La capa orgánica se transfirió a un vial de GC para su posterior inyección en el cromatógrafo de gases. Los datos cromatográficos se registraron y se trataron mediante el software EZChrom Elite de Agilent. Las FAMEs se identificaron a partir de los patrones purificados obtenidos de Sigma Co. (St.Louis, MO, EE.UU.).

3.3.A. Capítulo 3.A. Un enfoque desde la Química Analítica Verde para el estudio de la extracción de lípidos: métodos de cálculo computacional para la selección de disolventes verdes como alternativa al hexano

3.3.A.1. Disolventes alternativos al n-hexano

Los disolventes alternativos considerados adecuados desde la perspectiva de la Química Verde son aquellos que presentan las siguientes características: i) baja toxicidad, ii) son fáciles de reciclar, iii) son inertes y iv) no contaminan el producto. En general, los disolventes considerados como “verdes” son los incluidos en una de estas 7 clases según establece Chemat y Abert [17]: i) bio-solventes obtenidos a partir de recursos renovables, ii) ecológicos, iii) agua, (iv) polímeros líquidos que son no volátiles, biocompatibles y no tóxicos, v) disolventes fluorados que no son inflamables y no tóxicos, vi) líquidos iónicos y mezclas eutécticas, como disolventes no volátiles y térmicamente estables, y vii) fluidos supercríticos inertes, reciclables, no inflamables y no tóxicos. Además, el disolvente alternativo debe mantener las mismas propiedades de extracción que el disolvente a reemplazar, por lo que la elección del disolvente es un paso crucial en el proceso de sustitución del disolvente.

Por tanto, teniendo en cuenta estas premisas, los disolventes que se evaluaron para reemplazar el n-hexano, fueron: 2-metiltetrahidrofurano (MeTHF), ciclopentil metil éter (CPME), carbonato de dimetilo (DMC), isopropanol (IPA), etanol (EtOH), acetato de etilo (EtAc), p-cimeno, d-limoneno, junto con el n-hexano que se utilizó como disolvente de referencia. IPA, EtOH, EtAc, p-cimeno, d-limoneno y MeTHF se consideran bio-solventes que provienen de algunas materias primas renovables y sin emisión de compuestos orgánicos volátiles; son biodegradables y no tóxicos. IPA y EtOH son alcoholes que se obtienen a partir de la fermentación de azúcares (principalmente caña de azúcar o remolacha azucarera), por hidrólisis enzimática del almidón (obtenido de maíz o trigo) o a partir de la materia prima lignocelulósica (como son la madera, hierbas y residuos agrícolas). EtAc es un éster de acético, ácido orgánico que hoy en día se obtiene por la deshidrogenación de bioetanol producido por fermentación [24]. El limoneno es un terpeno biodegradable de bajo coste y baja toxicidad presente en los desechos agrícolas, principalmente derivados de las cáscaras de los cítricos, mientras que el p-cimeno puede obtenerse a partir del limoneno bajo tratamiento de irradiación por microondas [25]. MeTHF se deriva de recursos renovables como maíz y bagazo, y puede reemplazar al tetrahidrofurano (THF) y al dietil éter (DEE) como disolvente [26] [27]. Por otra parte, CPME y DMC son disolventes obtenidos a partir de síntesis química pero que pueden derivarse de subproductos de la industria

petroquímica o también pueden obtenerse mediante síntesis química “verde”. Hoy en día, DMC se obtiene por transesterificación del carbonato de propileno o por reacción del monóxido de carbono con metanol y oxígeno [28]; y es por su no toxicidad y por sus propiedades de biodegradabilidad por las que el DMC se reconoce como un disolvente “verde”. CPME no es producido por materiales renovables, pero es un buen candidato para la sustitución de éteres debido a que este disolvente supera a otros disolventes en términos de características de Medio Ambiente, Salud y Seguridad (EHS, del inglés *Environment, Health and Safety*), siendo comparativamente más favorable que otros éteres [29].

En la Tabla 3.1 se muestran las propiedades, la eficacia energética y la toxicidad de los disolventes estudiados: 2 ésteres (EtAc y DMC), 2 éteres (MeTHF y CPME), 2 alcoholes (IPA y EtOH), 2 terpenos (p-cimeno y d-limoneno) y el n-hexano, usado como disolvente de referencia.

Tabla 3.1. Propiedades, eficacia energética y toxicidad de los disolventes estudiados

Disolvente	Propiedades				Eficacia energética				Toxicidad		
	Peso Molecular (g/mol)	Viscosidad (25 °C Cp)	Log P	Punto inflamación (°C)	Punto ebullición (°C)	δH vap (kJ/mol)	Cp, liq (kJ/mol K)	Energía de evaporación de 1 Kg disolvente (kWh)	Fuente	CMR	Indice de Toxicidad
Hexano	86.2	0.31	3.94	-23.3	68.5	28.9	0.226	0.121	Petroleo	2	5
MeTHF	86.1	0.6	0.82	-11.1	79.9	30.7	0.151	0.126	Cereales	No	4
CPME	100.2	0.55	1.41	6.5	105.3	33	0.182	0.132	Síntesis química	No	4
p-cimeno	134.2	0.79	4.02	47.2	173.9	39.3	0.236	0.154	Madera	No	5
d-limoneno	136.2	0.923	4.45	42.8	175.4	39.5	0.249	0.157	Cereales	No	5
Etil Acetato	88.1	0.42	0.71	-3.3	73.9	31.9	0.194	0.127	Cereales	No	5
DMC	90.1	0.62	0.15	18.3	90.5	33	0.455	0.194	Síntesis química	No	5
IPA	60.1	2.04	0.16	11.7	73	39.8	0.156	0.219	Cereales	No	5
EtOH	46.1	1.095	-0.19	8.9	72.6	38.6	0.112	0.265	Cereales	No	5

3.3.A.2. Métodos computacionales

Para evaluar la potencia de solvatación del disolvente existen métodos computacionales como el método de *Hansen Solubility Parameter* (HSP) [30] y el *Conductor-like Screening Model for Real Solvent* (COSMO-RS) [31] que pueden utilizarse como una herramienta de decisión real para la elección de disolventes alternativos. El enfoque semi-empírico del método Hansen correlaciona y predice la capacidad de un disolvente para disolver los componentes principales de la muestra y COSMO-RS determina y predice las propiedades termodinámicas de los

disolventes sin la necesidad de datos experimentales. Las moléculas de soluto se consideran incrustadas en una cavidad que está rodeada por un medio virtual y, posteriormente, se cuantifican las energías de interacción entre las moléculas testigo y los disolventes mediante el uso de la termodinámica estadística, proporcionando predicciones de solubilidad de la molécula de referencia en el disolvente de estudio.

Para utilizar estos métodos computacionales, es necesario conocer la estructura de los componentes principales presentes en las muestras. En la Figura 3.3 se representan los compuestos utilizados en este estudio, presentes en las muestras de salmón, en la que se incluyen 4 triacilglicéridos (TAG 1 (C18:1, C18:1, C18:1), TAG 2 (C18:1, C18:1, C18:2) TAG 3 (C16:0, C18:1, C18:1) y TAG 4 (C16:0, C18:1, C18:2)), 4 diacilglicéridos (DAG1 (C18:1, C18:1), DAG2 (C18:1, C18:2), DAG3 (C18:2, C18:2) y DAG4 (C16:0, C18:1)), 3 ácidos grasos libres (FFA1 (C18:1), FFA2 (C18:2) y FFA3 (C16:0)) y el ergosterol.

3.3.A.2.1. Hansen Solubility Parameters (HSP)

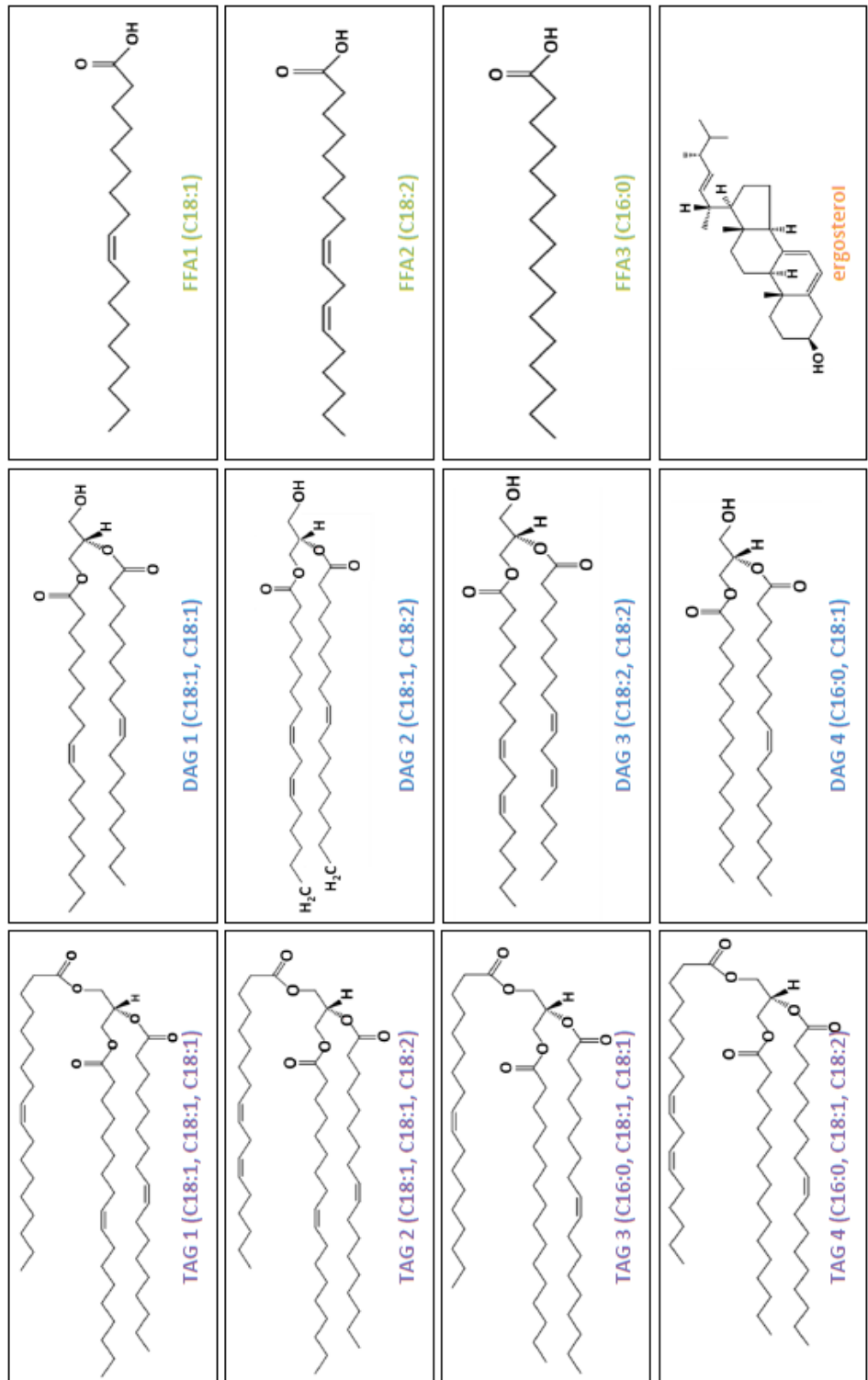
La predicción por el método de Hansen proporciona una evaluación de la capacidad de un disolvente para disolver los componentes principales de una muestra, definiéndose el parámetro de solubilidad mediante la siguiente ecuación:

$$\delta_{total}^2 = \delta_D^2 + \delta_P^2 + \delta_H^2 \quad (\text{Ecuación 3.1})$$

donde δ_{total} es el parámetro de solubilidad total de Hansen, δ_D el término dispersivo, δ_P el término polar y δ_H el término correspondiente a los enlaces de hidrógeno. Esta capacidad se evalúa mediante la diferencia de energía relativa (RED) calculada como la relación entre R_a y R_o . La relación R_a/R_o se llama *número RED*, y refleja la diferencia de la energía relativa, siendo R_a la “distancia” entre dos materiales basada en sus respectivos componentes de parámetro de solubilidad parcial, y R_o la distancia experimentalmente determinada del radio de la esfera de solubilidad [32].

Los valores de RED iguales a 0 significan que no tienen ninguna diferencia de energía, los valores RED menores que 1.0 indican una elevada afinidad; RED igual o cercano a 1.0 indican una afinidad media; y los valores de RED muy altos indican afinidades bajas [30].

Figura 3.3. Principales compuestos lipídicos del aceite de salmón considerados en este estudio

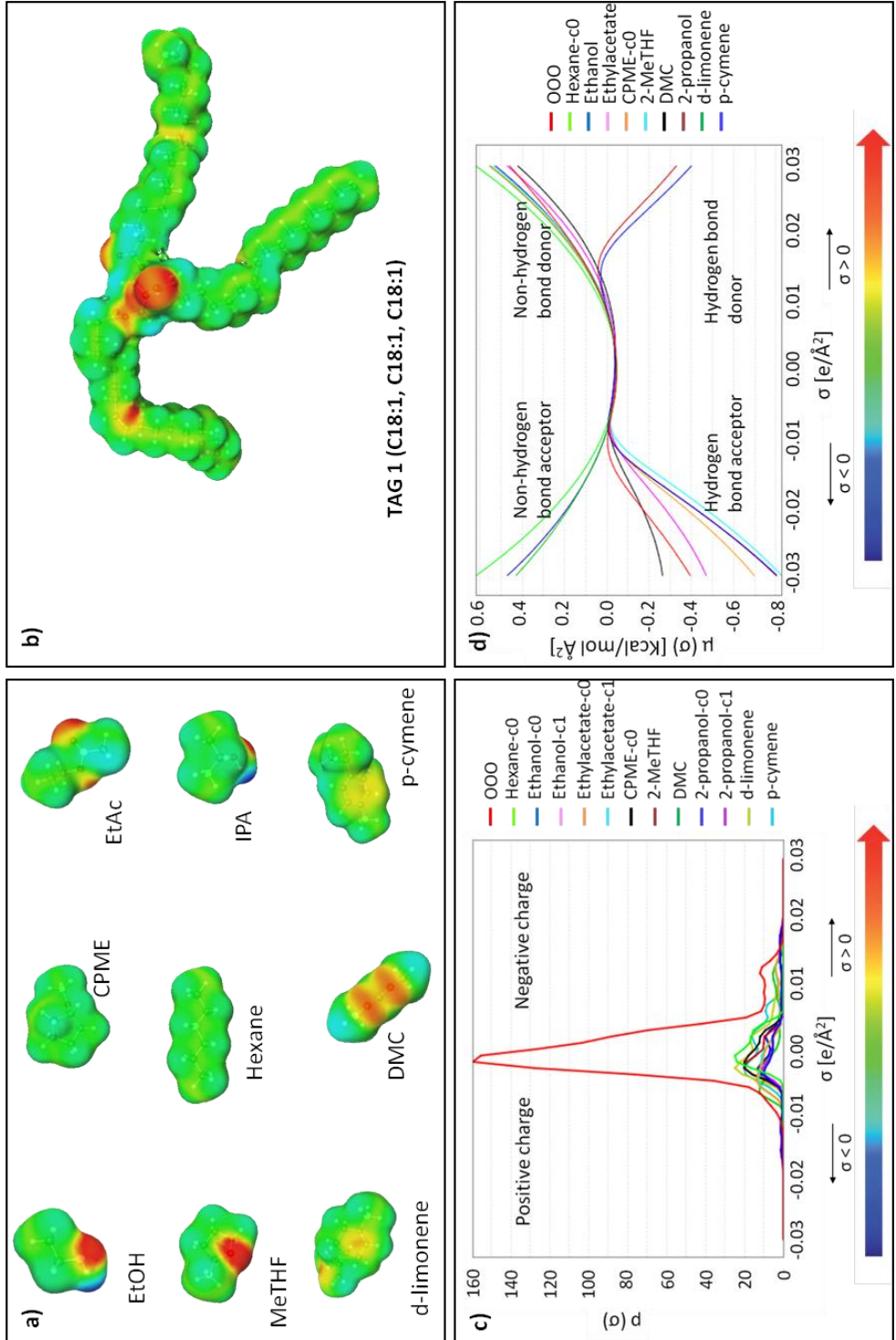


3.3.A.2.2. Conductor-like Screening Model for Real Solvent (COSMO-RS)

COSMO-RS combina la química cuántica (COSMO) y la termodinámica estadística (RS) para determinar y predecir propiedades termodinámicas sin la necesidad de disponer de datos experimentales. Es un método de gran alcance para la descripción molecular y la investigación del disolvente. En la Figura 3.4 se puede ver un ejemplo de las distintas etapas del cálculo por el método COSMO-RS aplicado al TAG1: a) cálculo de la σ -superficie de los distintos disolventes estudiados; b) construcción de la σ -superficie del triacilglicérido TAG1; c) cálculo de las energías de interacciones superficiales locales entre s-perfiles de TAG1 y disolventes y, d) determinación de las s-potenciales de TAG1 y disolventes.

El proceso de cálculo lleva a cabo las siguientes consideraciones. La molécula de soluto es sumergida en un medio continuo con una constante dieléctrica ϵ . En este entorno, se construye una superficie alrededor de la molécula que genera un gran número de cargas electrostáticas. La estructura y la distribución de las cargas se optimizan para encontrar la energía mínima del sistema (molécula en su estado más estable) a través de cálculos basados en algoritmos, denominados Teoría Funcional de Densidad (DFT, del inglés Density Functional Theory). Esta densidad de carga es la llamada σ -superficie (Figura 3.4a y 3.4b). El código de color de la superficie σ proporciona información sobre la densidad de carga en cada punto de la molécula, ya que la zona verde corresponde a la densidad de carga cero, el azul caracteriza una densidad de carga positiva δ^+ y el color rojo significa densidades de carga negativa δ^- . En otras palabras, las áreas verdes normalmente están ligadas al esqueleto de carbono, el área roja a los grupos oxigenados o nitrogenados y las áreas azules a los átomos de hidrógeno. Esta distribución de carga óptima si se segmenta y se reduce a un histograma obtenemos el llamado σ -perfil (Figura 3.4c). El σ -perfil traduce la información contenida en la superficie 3D en información 2D, así, la σ -superficie y el σ -perfil permiten caracterizar una molécula de referencia (energía mínima, molécula estable) como un soluto en un conductor perfecto. Estos perfiles se utilizan para interpretar la afinidad de un disolvente, entender la interacción entre el compuesto y disolvente en estado mixto, así como para estimar las propiedades termodinámicas del sistema. La termodinámica de las interacciones moleculares, basada en el σ -perfil obtenido, se utiliza para calcular el potencial químico del segmento superficial (σ -potencial) (Figura 3.4d), utilizando para ello el programa COSMOthermX C30 release 13.01).

Figura 3.4. Etapas del cálculo: a) σ -superficie de los distintos disolventes estudiados; b) σ -superficie del triacilglicerido TAG1; c) energías de interacciones superficiales locales entre s-perfiles de TAG1 y disolventes y, d) s-potenciales de TAG1 y disolventes.



El σ -potencial describe la probabilidad de interactuar los compuestos considerados con los disolventes con polaridad y enlaces de hidrógeno. Para cada molécula, la parte de la carga negativa de la molécula se localiza en el lado derecho (enlaces de hidrógeno aceptor) con valores σ positivos mientras que la parte de la carga positiva se localiza en el lado izquierdo (enlaces de hidrógeno del donante) con valores σ negativos. Generalmente, la región $\sigma \pm 0.01 \text{ eA}^{-2}$ se considera no polar o débilmente polar.

Además, el COSMOthermX también calcula la solubilidad relativa entre un compuesto sólido y el disolvente, en términos de logaritmo de la solubilidad en fracciones molares ($\log_{10}(X_{\text{solub}})$). El logaritmo de la mejor solubilidad se aproxima a 0.

3.3.A.3. Resultados y discusión

3.3.A.3.1. Cálculos mediante HSP

En la Tabla 3.2 se resumen los valores RED que estiman la capacidad de los disolventes seleccionados para solvatar los TAGs, DAGs, FFAs y ergosterol estudiados, comparando la solubilidad de los principales componentes lipídicos del aceite de salmón en cada disolvente con los valores de referencia correspondientes al n-hexano. De acuerdo con los resultados, teóricamente los disolventes CPME, MeTHF y, especialmente, d-limoneno y p-cimeno, son mejores que el n-hexano para la solvatación de los compuestos mencionados. El EtOH, EtAc, DMC e IPA presentaron una peor capacidad que la del n-hexano en el estudio teórico de HSPs para los TAGs. Además, para IPA y EtOH se obtuvieron peores resultados para la solvatación de DAGs y FFAs, siendo EtAc, CPME, MeTHF, DMC, p-cimeno y limoneno teóricamente mejores que el n-hexano para estas familias de compuestos. En el caso del ergosterol, con EtOH, DMC e IPA se obtuvieron los peores resultados teóricos para la solvatación de los mencionados compuestos.

A partir de los datos calculados por el modelo se realizó un análisis por componentes principales de los 9 disolventes y de los compuestos considerados en este estudio. La Figura 3.5 muestra el gráfico *biplot* (valores de puntuación y de carga) del PCA para los valores RED de los compuestos estudiados con los disolventes. Los dos primeros componentes principales representan el 99.49% de la varianza explicada, siendo el 94.21 y el 5.28% para PC1 y PC2, respectivamente. De acuerdo con los resultados, el d-limoneno y el p-cimeno presentan un comportamiento similar al del n-hexano para solvatar a los TAGs. IPA y EtOH están situados en la parte opuesta de los TAGs y ergosterol, indicando que estos disolventes son menos afines a estos compuestos.

Tabla 3.2. Valores de Diferencias Energéticas Relativas (RED) calculadas mediante Hansen Solubility Parameters (HSP)

Solvent	TAGs				DAGs				FFAs			ERGO
	TAG1	TAG2	TAG3	TAG4	DAG1	DAG2	DAG3	DAG4	FFA1	FFA2	FFA3	
Hexane	1.12	1.14	1.09	1.12	1.91	1.84	1.80	1.94	1.79	1.95	1.84	1.76
EtOH	4.14	4.1	4.13	4.08	3.06	3.13	3.20	3.02	3.26	3.16	3.16	3.90
EtAc	1.53	1.49	1.50	1.47	0.66	0.68	0.73	0.59	0.81	0.86	0.67	1.65
CPME	0.55	0.52	0.52	0.48	0.65	0.58	0.50	0.68	0.48	0.65	0.53	0.93
MeTHF	0.89	0.86	0.87	0.84	0.63	0.57	0.52	0.63	0.51	0.62	0.50	1.09
DMC	2.07	2.04	2.04	2.01	1.39	1.40	1.42	1.33	1.47	1.49	1.37	2.18
IPA	3.23	3.18	3.21	3.17	2.14	2.21	2.28	2.09	2.35	2.26	2.24	3.04
p-cymene	0.36	0.36	0.40	0.40	1.18	1.12	1.03	1.23	0.93	1.0	1.06	0.52
d-limonene	0.22	0.18	0.24	0.21	0.95	0.89	0.81	1.01	0.73	0.84	0.85	0.60

Por otro lado, EtAc, MeTHF y DMC se sitúan más cerca de DAGs y FFAs que de los TAGs y el ergosterol, de manera que se podría esperar que estos tres disolventes pudieran solvatar más fácilmente los dos primeros compuestos antes que los dos últimos. CPME está situado en el centro de todos los compuestos estudiados, indicando que este disolvente es menos específico que los demás y que teóricamente podría solvatar a todos los compuestos por igual.

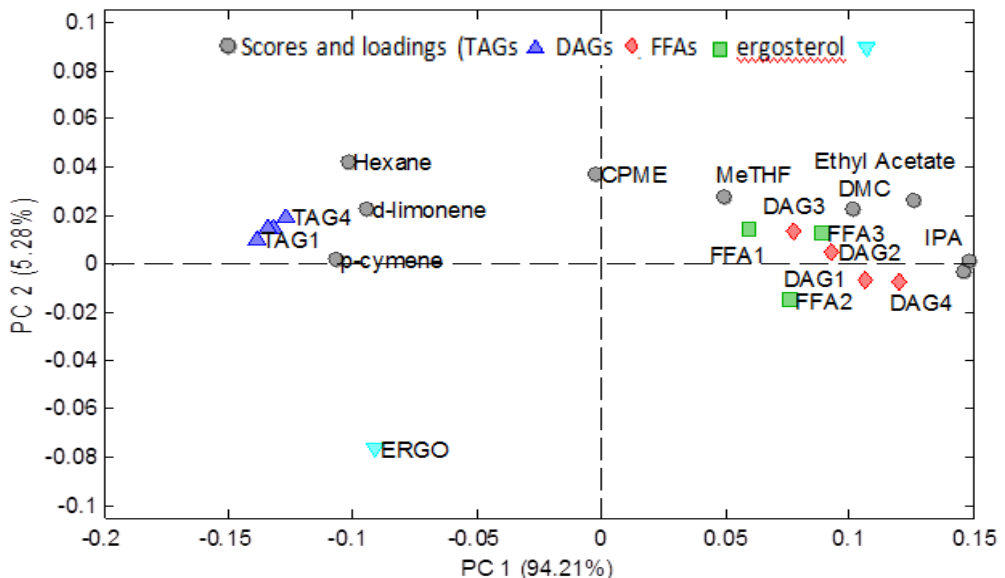


Figura 3.5. Análisis de componentes principales biplot (valores de puntuación y de carga) de los compuestos estudiados y disolventes, representados por PC2 vs PC1 para los valores RED de Hansen

3.3.A.3.2. Cálculos con el COSMO-RS

La Tabla 3.3 muestra la solubilidad de los 4 TAGs, 4 DAGs, 4 FFAs y el ergosterol encontrados en el salmón con los diferentes disolventes estudiados, expresada como solubilidad relativa, $\log_{10}(X_{\text{solub}})$, y estableciendo el logaritmo de la mejor solubilidad como 0. Considerando el $\log_{10}(X_{\text{solub}})$ de los TAGs, los disolventes como MeTHF, CPME, EtOAc, p-cimeno y d-limoneno han obtenido iguales o mejores resultados que para el n-hexano. Por su parte, DMC, IPA y EtOH teóricamente no son buenos sustitutos del n-hexano para la solvatación de TAGs y algunos DAGs; sin embargo, para otros componentes podrían considerarse como una mejor opción que el n-hexano. En resumen, EtAc, MeTHF y CPME parecen ser, teóricamente a partir de los resultados obtenidos por COSMO-RS, los mejores disolventes alternativos al n-hexano para la extracción de todos los compuestos estudiados.

Tabla 3.3. Valores de solubilidad relativa ($\log_{10}(X_{\text{solub}})$) de los principales compuestos del aceite de salmón en varios disolventes obtenidos por modelización empleando el programa COSMO-RS.

Solvent	TAGs				DAGs				FFAs			ERGO
	TAG1	TAG2	TAG3	TAG4	DAG1	DAG2	DAG3	DAG4	FFA1	FFA2	FFA3	
Hexane	0.0	0.0	-0.17	-0.001	-1.17	-0.44	-1.63	-0.64	-1.34	-1.31	-1.27	-0.81
EtOH	-2.52	-2.53	-2.19	-2.55	-1.03	-1.52	-0.78	-1.1	-0.05	0.0	-0.001	-0.56
EtAc	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CPME	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MeTHF	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DMC	-1.51	-1.46	-1.21	-1.41	-0.99	-0.92	-0.39	-1.12	-0.51	-0.22	-0.56	-0.82
IPA	-1.49	-1.51	-1.24	-1.52	-0.34	-0.79	-0.21	-0.48	0.0	0.0	0.0	-0.21
p-cymene	0.0	0.0	0.0	0.0	-0.75	-0.19	-0.83	-0.44	-1.03	-0.9	-1.03	-0.71
d-limonene	0.0	0.0	0.0	0.0	-0.69	-0.13	-0.88	-0.37	-0.98	-0.88	-0.96	-0.65

Al igual que en el caso del programa Hansen, para el COSMO-RS se realizó también un análisis de componentes principales y en la Figura 3.6 se muestra el gráfico *biplot* del PCA obtenidos para los valores $\log_{10}(X_{\text{solub}})$ de los compuestos y disolventes estudiados. Los dos primeros componentes principales representan el 98.41% de la varianza total explicada, siendo 96.05% y 2.36% para PC1 y PC2, respectivamente. El hexano, el d-limoneno y el p-cimeno se encuentran juntos y cerca de los TAGs, indicando que como disolventes tienen un comportamiento similar para la extracción de los estos lípidos. DMC, IPA y EtOH se encuentran en la

parte opuesta de los TAGs, sin embargo, se sitúan cerca de FFAs, DAGs y ergosterol, lo que pondría de manifiesto su afinidad. MeTHF, CPME y EtAc están situados en el medio, indicando que estos disolventes teóricamente podrían extraer por igual tanto TAGs, DAGs, FFAs y ergosterol de las muestras de salmón.

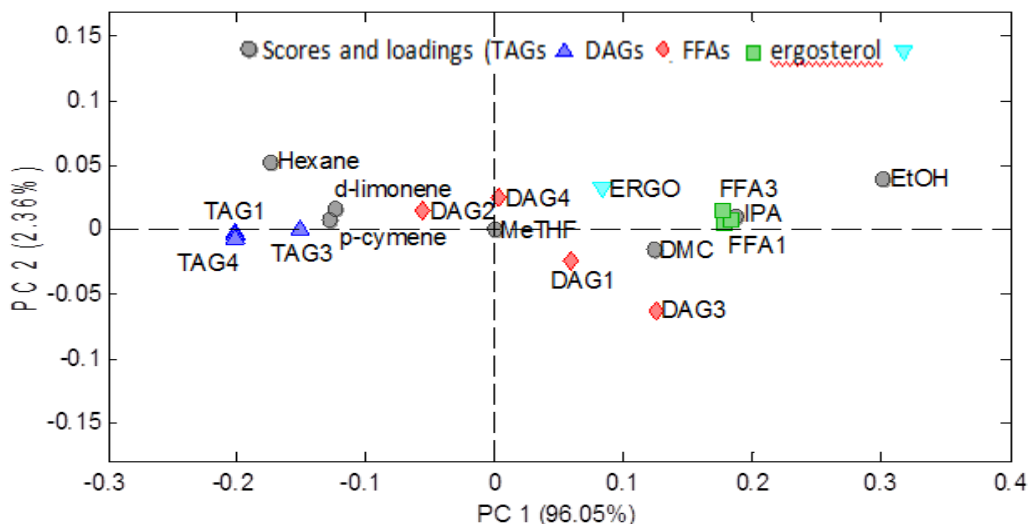


Figura 3.6. Análisis de componentes principales biplot (valores de puntuación y de carga) de los compuestos estudiados y disolventes, representados por PC2 vs PC1 para los valores $\log_{10}(X_{\text{solub}})$ obtenidos utilizando COSMO-RS

3.3.A.3.3. Análisis HPTLC de las fracciones lipídicas extraídas

Tal como se ha indicado en la parte experimental de este capítulo, utilizando una misma muestra de salmón liofilizado se procedió a realizar una extracción de la fracción lipídica considerando cada uno de los disolventes indicados. Las extracciones se realizaron por triplicado. Para establecer la composición del aceite de salmón recuperado, cada uno de los extractos se analizaron aplicando los procedimientos de referencia explicados.

En la Figura 3.7 se han representado los porcentajes de clases de lípidos determinados mediante la técnica HPTLC en los 9 disolventes estudiados, indicando que el 73-77% de los constituyentes extraídos son TAGs seguidos por DAGs, FFAs y ergosterol con porcentajes del 15-20 %, 5-6% y 2-3%, respectivamente. Los MAGs, PCs y PE no fueron detectados en las muestras de salmón.

Se puede ver que las cantidades de los lípidos extraídos por el p-cimeno y d-limoneno son ligeramente diferentes que en el caso de los otros disolventes estudiados, ya que dan lugar a recuperaciones de cantidades de TAGs más bajas y cantidades más altas de DAGs. Una posible explicación podría ser la elevada temperatura de ebullición de estos disolventes frente a los otros, la cual puede afectar al rendimiento de la extracción (véase datos de la Tabla 3.1). Hay que tener en cuenta que la principal reacción responsable de la degradación de los lípidos es la oxidación de los ácidos grasos insaturados [33] y ésta puede ser aumentada por trabajar a temperaturas por encima de 100 °C [34]. Durante el proceso de calentamiento se puede propiciar la generación/transformación a una amplia variedad de TAGs polimerizados (PTGs), incluyendo dímeros y oligómeros [35]. Particularmente los PUFAs son más susceptibles a la degradación oxidativa, especialmente cuando se encuentran en su forma libre [36]. Por otra parte, pudo verificarse visualmente que el extracto obtenido mediante el uso de p-cimeno y d-limoneno, presentaba un color más oscuro que en el caso de los otros extractos, siendo eso un posible indicador de la degradación de parte de los componentes inicialmente presentes en el aceite de salmón [37].

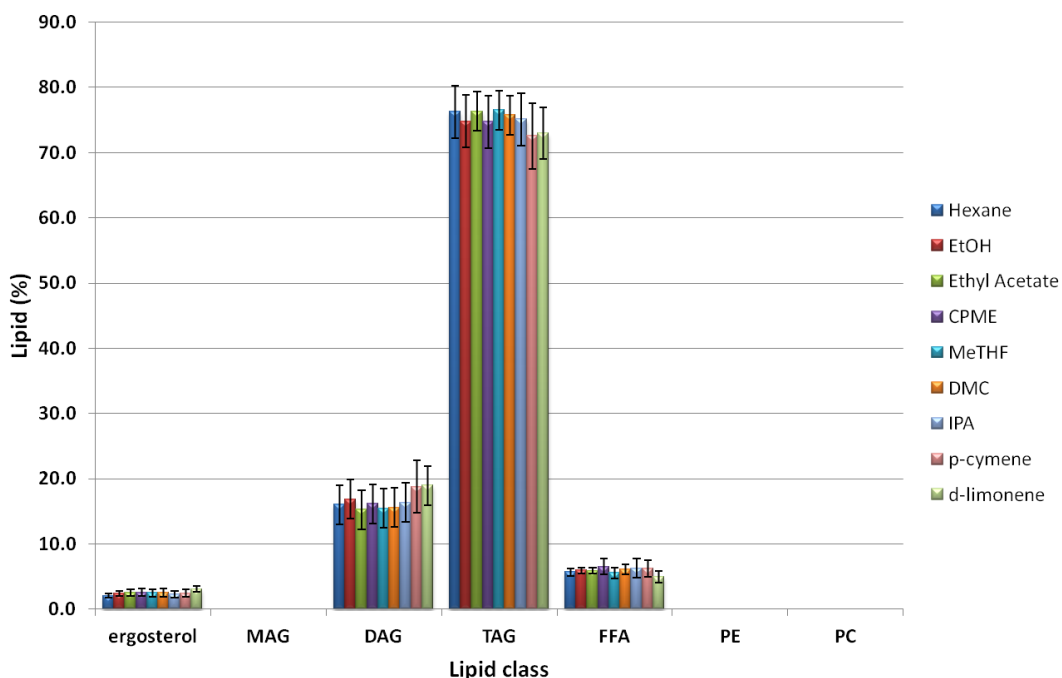


Figura 3.7. Clases de lípidos determinadas por la técnica HP-TLC en los extractos de aceite de salmón obtenidos por los diferentes disolventes estudiados

En esta ocasión se realizó un análisis de componentes principales considerando los resultados experimentales obtenidos, representándose en la Figura 3.8 el gráfico *biplot* correspondiente a los datos obtenidos por HP-TLC para las clases de lípidos y los disolventes estudiados. Los dos primeros componentes principales representan un 91.99% y un 6.55% de la varianza explicada. Se puede observar que el d-limoneno y el p-cimeno están situados cerca de los DAGs y en el lado opuesto de los TAGs, lo que resulta coherente con las recuperaciones observadas experimentalmente. El ETOH se encuentra cerca del ergosterol, mientras que CPME e IPA se sitúan cerca de los FFAs. Por su parte, el MeTHF, EtAc, DMC y hexano están más próximos a los TAGs, lo que significaría una mayor afinidad entre ellos. En resumen, y coherentemente con los resultados observados, esto podría evidenciar la existencia de problemas en el uso de los terpenos como disolventes de extracción para el aceite de salmón, a pesar de lo que se podía haber deducido a partir de los datos teóricos previos en cuanto a su poder de solvatación.

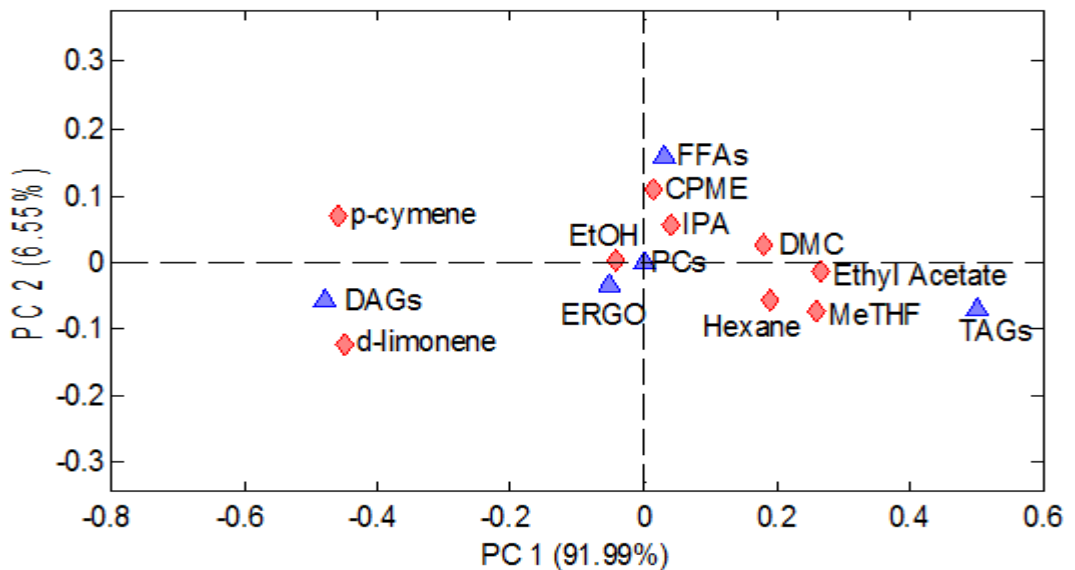


Figura 3.8. Análisis de componentes principales biplot (valores de puntuación y de carga) representados por PC1 vs PC2 para a) las clases de lípidos obtenidos por HPTLC y b) ácidos grasos obtenidos por GC-FID, y los disolventes estudiados

3.3.A.3.4. Análisis GC-FID de las fracciones lipídicas extraídas

Con el fin de comparar y evaluar la capacidad de extracción de disolventes alternativos al hexano, se determinaron mediante GC-FID las cantidades de los diferentes ácidos grasos extraídos por cada uno de los disolventes considerados. En la Tabla 3.4 se puede observar la composición lipídica del aceite de salmón extraído, observándose que los ácidos grasos predominantes son el ácido oleico (OA; C18:1) (38-40 %), el ácido linoleico (LA; C18:2) (16-17%) y el ácido palmítico (PA; C16:0) (11-12%), los cuales representan en su conjunto más del 70% de los ácidos grasos totales en el aceite extraído. Existen otros ácidos grasos con menor concentración como el ácido linolénico (ALA; C18:3) alrededor del 5-6%, y el ácido esteárico (C18:0), ácido palmitoleico (C16:1), ácido eicosenoico (C20:1), ácido eicosapentaenoico (EPA, C20:5) y el ácido docosahexaenoico (DHA, C22:6) con cerca del 3%. Otros como ácido docosapentaenoico (DPA, C22:5), el ácido eicosadienoico (C20:2) y el ácido mirístico (C14:0) se encuentran presentes en los extractos con valores comprendidos entre el 1 y 2%.

Considerándolos por familias, en general, los SFAs se encuentran entre el 19.5 y 19.9% en el extracto de aceite de salmón, mientras que los MUFAs se encuentran entre el 43.5 y 44.9% y los PUFAs entre el 31.2 y 34.6%. Para todos los disolventes estudiados se obtuvieron resultados similares excepto para el limoneno y, especialmente, el p-cimeno. Los extractos de los disolventes de p-cimeno y limoneno presentan porcentajes reducidos de algunos PUFAs debido, probablemente, al hecho, como se ha explicado, que estos ácidos grasos insaturados son más susceptibles a la degradación oxidativa térmica que los MUFAs. Se puede observar que el valor más bajo de los PUFAs se encuentra en el extracto de p-cimeno con un 31.2% seguido por el limoneno con un 32.8%. Al mismo tiempo, se observa un aumento del ácido esteárico (C18:0) junto con un ácido graso desconocido en el extracto de p-cimeno. En el caso del extracto de limoneno, este aumento se observa para el ácido esteárico (C18:0) y el ácido oleico (C18:1).

Al igual que para los resultados por HP-TLC, se realizó un análisis PCA y en la Figura 3.9 podemos ver el gráfico *biplot* de los ácidos grasos determinados por de GC a partir de los extractos de los diferentes disolventes. Los dos primeros componentes principales representan el 95.29% de la varianza explicada, siendo 81.57% y 13.72% para PC1 y PC2, respectivamente.

Tabla 3.4. Composición de ácidos grasos determinados en el aceite de salmón extraído con n-hexano y los disolventes alternativos usando GC-FID

FFA	Extracted amounts of fatty acids (%) [w/w]									
	Hexane	EtOH	Ethyl Acetate	CPME	MeTHF	DMC	IPA	p-cymene	d-limonene	
C13:0	No Detected	No Detected	No Detected	0.170 ± 0.010	0.66 ± 0.05	No Detected	No Detected	No Detected	No Detected	
C14:0	2.050 ± 0.017	2.05 ± 0.03	2.049 ± 0.011	2.079 ± 0.006	2.049 ± 0.0052	2.057 ± 0.0165	2.07 ± 0.05	2.12 ± 0.02	2.153 ± 0.010	
C15:0	0.198 ± 0.002	0.195 ± 0.003	0.1969 ± 0.0009	0.1981 ± 0.0005	0.1963 ± 0.0014	0.193 ± 0.0046	0.193 ± 0.0009	0.215 ± 0.010	0.270 ± 0.004	
C16:0	11.26 ± 0.05	11.55 ± 0.08	11.363 ± 0.006	11.45 ± 0.03	11.352 ± 0.010	11.41 ± 0.08	11.61 ± 0.14	11.58 ± 0.11	11.592 ± 0.013	
C16:1	2.81 ± 0.04	2.86 ± 0.04	2.847 ± 0.007	2.88 ± 0.03	2.828 ± 0.0088	2.85 ± 0.05	2.86 ± 0.04	2.86 ± 0.03	2.893 ± 0.006	
C18:0	3.42 ± 0.07	3.26 ± 0.09	3.274 ± 0.003	3.28 ± 0.05	3.264 ± 0.010	3.25 ± 0.18	3.2 ± 0.2	3.6 ± 0.2	3.5 ± 0.3	
C18:1n9	38.91 ± 0.04	38.22 ± 0.16	39.16 ± 0.12	39.07 ± 0.15	38.81 ± 0.15	38.81 ± 0.10	38.5 ± 0.4	38.2 ± 1.2	39.51 ± 0.17	
C18:2n6	17.28 ± 0.04	17.06 ± 0.09	17.14 ± 0.03	17.18 ± 0.02	17.071 ± 0.0164	17.21 ± 0.07	17.15 ± 0.12	16.5 ± 0.2	17.29 ± 0.04	
unknown	0.1706 ± 0.002	0.1704 ± 0.0013	0.173 ± 0.002	0.1663 ± 0.0009	0.172 ± 0.0041	0.20 ± 0.04	0.1670 ± 0.0011	3.7 ± 1.2	0.12 ± 0.08	
C18:3n6	0.240 ± 0.002	0.2375 ± 0.0005	0.2392 ± 0.0003	0.2365 ± 0.0013	0.2355 ± 0.0013	0.2360 ± 0.0005	0.2318 ± 0.0010	0.222 ± 0.003	0.283 ± 0.016	
unknown	0.272 ± 1E-03	0.2707 ± 0.0018	0.2696 ± 0.0006	0.2696 ± 0.0015	0.261 ± 0.0034	0.263 ± 0.005	0.262 ± 0.004	0.235 ± 0.009	0.346 ± 0.016	
C18:3n3	5.506 ± 0.018	5.46 ± 0.04	5.46 ± 0.03	5.515 ± 0.017	5.42 ± 0.06	5.50 ± 0.03	5.54 ± 0.05	5.19 ± 0.02	5.32 ± 0.08	
unknown	0.652 ± 0.004	0.644 ± 0.007	0.643 ± 0.006	0.642 ± 0.003	0.635 ± 0.0083	0.643 ± 0.006	0.645 ± 0.009	0.57 ± 0.02	0.655 ± 0.007	
unknown	0.2370 ± 6E-04	0.242 ± 0.003	0.236 ± 0.003	0.2348 ± 0.0014	0.232 ± 0.0032	0.245 ± 0.005	0.242 ± 0.004	0.13 ± 0.09	0.268 ± 0.004	
C20:0	0.36 ± 0.03	0.369 ± 0.011	0.387 ± 0.007	0.383 ± 0.015	0.3743 ± 0.0006	0.48 ± 0.12	0.372 ± 0.020	0.36 ± 0.02	0.398 ± 0.017	
C20:1	2.42 ± 0.04	2.40 ± 0.03	2.471 ± 0.016	2.437 ± 0.014	2.430 ± 0.017	2.53 ± 0.09	2.41 ± 0.03	2.35 ± 0.13	2.47 ± 0.02	
C20:2	1.165 ± 0.008	1.160 ± 0.012	1.179 ± 0.004	1.1685 ± 0.0004	1.1647 ± 0.0013	1.177 ± 0.009	1.162 ± 0.004	1.09 ± 0.04	1.124 ± 0.005	
C20:3n6	0.425 ± 0.004	0.4324 ± 0.0018	0.429 ± 0.002	0.4265 ± 0.0017	0.422 ± 0.003	0.425 ± 0.006	0.421 ± 0.006	0.383 ± 0.004	0.487 ± 0.015	
C20:4n6	0.4196 ± 0.002	0.483 ± 0.003	0.4264 ± 0.0014	0.431 ± 0.003	0.426 ± 0.005	0.425 ± 0.005	0.450 ± 0.009	0.395 ± 0.014	0.418 ± 0.006	
C20:3n3	0.408 ± 0.003	0.4054 ± 0.0016	0.4089 ± 0.0011	0.4097 ± 0.0017	0.404 ± 0.002	0.408 ± 0.002	0.4026 ± 0.0014	0.363 ± 0.008	0.483 ± 0.013	
unknown	0.800 ± 0.004	0.796 ± 0.004	0.791 ± 0.004	0.779 ± 0.003	0.785 ± 0.007	0.788 ± 0.006	0.789 ± 0.006	0.678 ± 0.014	0.734 ± 0.004	
C20:5n3	3.075 ± 0.010	3.29 ± 0.05	3.07 ± 0.03	3.01 ± 0.03	3.03 ± 0.02	3.13 ± 0.05	3.24 ± 0.06	2.62 ± 0.13	2.79 ± 0.03	
C22:0	0.789 ± 0.012	0.621 ± 0.017	0.637 ± 0.009	0.616 ± 0.010	0.621 ± 0.016	0.61 ± 0.02	0.60 ± 0.03	0.71 ± 0.15	0.64 ± 0.02	
C23:0	0.204 ± 0.012	0.196 ± 0.004	0.193 ± 0.002	0.1909 ± 0.0011	0.198 ± 0.003	0.195 ± 0.003	0.199 ± 0.003	0.11 ± 0.08	0.06 ± 0.08	
C24:0	1.493 ± 0.006	1.565 ± 0.005	1.497 ± 0.004	1.453 ± 0.003	1.4814 ± 0.0014	1.50 ± 0.02	1.529 ± 0.018	1.21 ± 0.05	1.29 ± 0.02	
C22:5	1.493 ± 0.006	1.565 ± 0.005	1.497 ± 0.004	1.453 ± 0.003	1.4814 ± 0.0014	1.50 ± 0.02	1.529 ± 0.018	1.21 ± 0.05	1.29 ± 0.02	
C22:6	3.81 ± 0.04	4.49 ± 0.16	3.97 ± 0.03	3.87 ± 0.10	4.00 ± 0.04	3.94 ± 0.10	4.26 ± 0.11	3.3 ± 0.2	3.34 ± 0.04	
Σ SFAs	19.8	19.8	19.6	19.6	19.5	19.7	19.8	19.9	19.9	
Σ MUFA	44.3	43.5	44.5	44.4	44.1	44.2	43.8	43.5	44.9	
Σ PUFA	33.8	34.6	33.8	33.7	33.7	33.9	34.4	31.2	32.8	

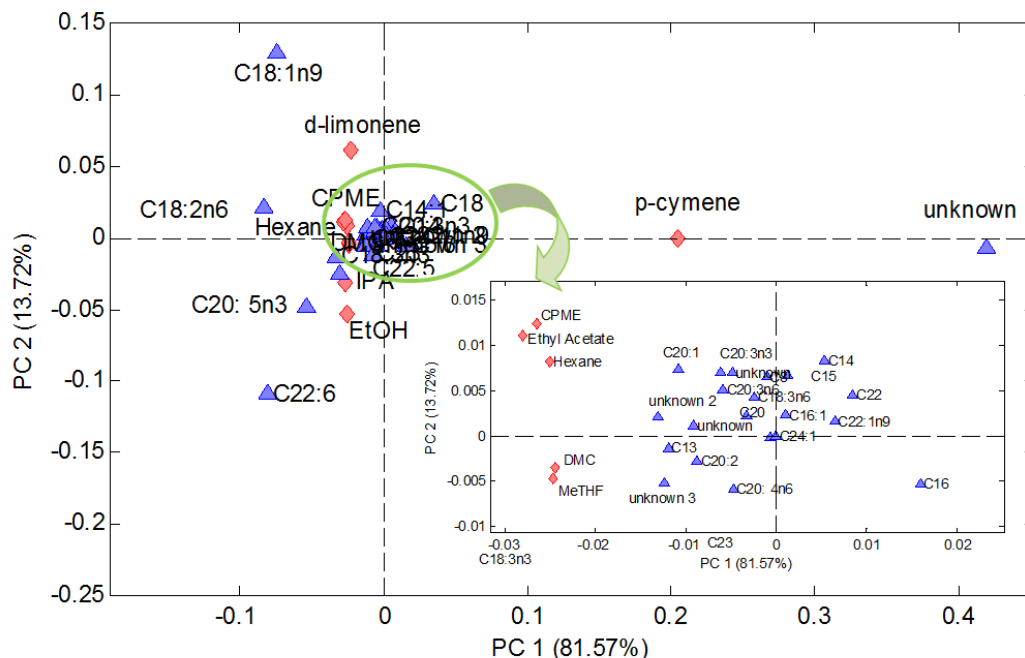


Figura 3.9. Análisis de componentes principales biplot (valores de puntuación y de carga) representados por PC1 vs PC2 para a) las clases de lípidos obtenidos por HPTLC y b) ácidos grasos obtenidos por GC-FID, y los disolventes estudiados

Los resultados de PCA evidenciaron una separación clara del p-cimeno de los otros disolventes estudiados en la dirección de PC1, encontrándose cerca de un FFA desconocido. Sin embargo, el hexano, el EtOH, el EtAc, el CPME, el MeTHF, el DMC, el IPA y el limoneno se encuentran en el centro de todos los FFAs, apreciándose que el MeTHF se encuentra más cercano al C13:0 y el limoneno al C18:1. También, el EtOH y el IPA están más cerca del C22:5, C20:5 y C22:6 que el hexano, EtAc, CPME y DMC. Para obtener una visualización de las similitudes del poder de extracción de cada disolvente, se realizó el análisis de agrupamiento jerárquico (HCA), utilizando el método *Ward's Linkage*. Como puede verse en la Figura 3.10, para un valor de corte mínimo de 1.0, existen 4 grupos diferentes de extractos.

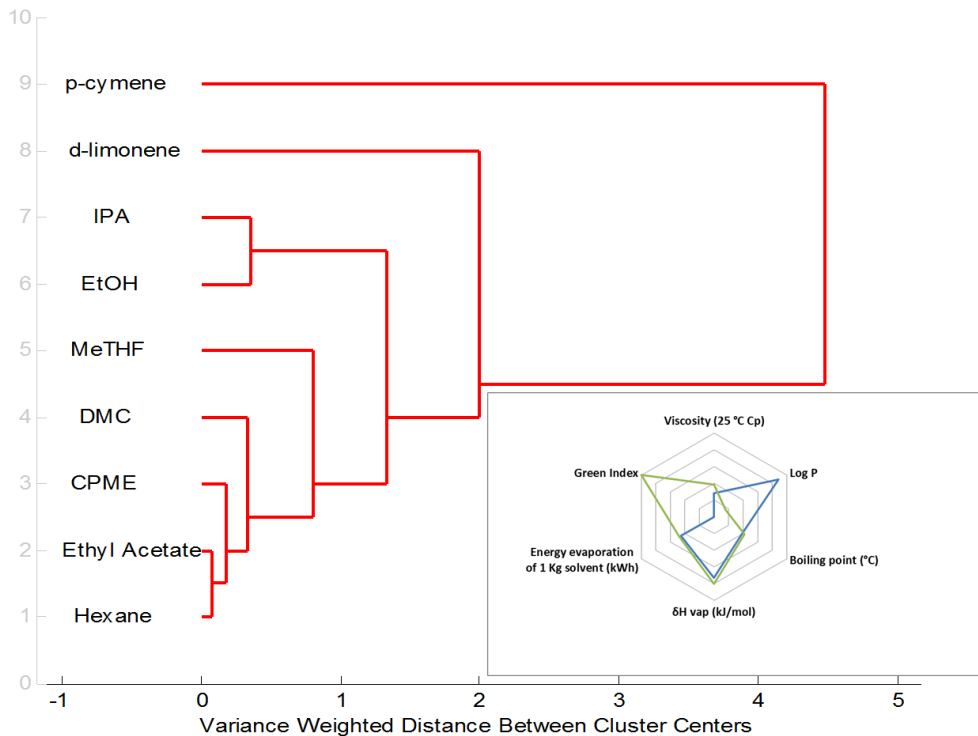


Figura 3.10. Análisis de agrupamiento jerárquico de los disolventes empleados a partir de los ácidos grasos determinados por GC.
 Inserto: Comparación de propiedades de hexano y EtAc

El primer grupo está compuesto por el n-hexano, EtAc, CPME, DMC y MeTHF, el segundo incluye los alcoholes (IPA y EtOH) mientras que los dos grupos adicionales se refieren a los terpenos. Puede observarse que el disolvente más próximo al n-hexano es el EtAc seguido por el CPME, lo que indica que los extractos obtenidos con ambos disolventes presentan una composición similar a la obtenida con el empleo de n-hexano. Por otro lado, el p-cimeno y el limoneno se encuentran situados en el lado opuesto del n-hexano, lo que significa un comportamiento diferente en lo que se refiere a la extracción de los lípidos presentes en el salmón, tal como se deduce de los datos encontrados en el análisis por GC.

En general, se puede concluir que los extractos de los ésteres y éteres son similares a los encontrados para el n-hexano, especialmente cuando se emplea el EtAc, seguido de los alcoholes EtOH e IPA. Sin embargo, los extractos obtenidos utilizando los terpenos d-limoneno y p-cimeno son significativamente diferentes de los obtenidos con el resto de los disolventes. El recuadro de la Figura 3.10 representa las propiedades atribuibles al n-hexano y al EtAc, ya que éste último

supone la mejor alternativa al n-hexano, para realizar la extracción de los aceite presentes en el salmón.

En la Figura 3.11 se representan los rendimientos lipídicos totales (cantidad de aceite extraído de las muestras de salmón) encontrados para cada uno de los disolventes estudiados en este trabajo. Los valores promedio obtenidos se encuentran entre el 28.9 y 33.9% de lípidos respecto a la cantidad de materia seca (salmón liofilizado), sin diferencias significativas entre el n-hexano y los disolventes estudiados para un nivel de confianza del 95%, excepto para el extracto de IPA que proporcionó un valor de rendimiento lipídico más bajo. Es de destacar también la mayor variabilidad y menor reproducibilidad obtenida para la extracción de la fracción lipídica cuando se emplearon los dos terpenos, tal como puede verse con las barras de error representadas en el gráfico de barras de la Figura 3.11.

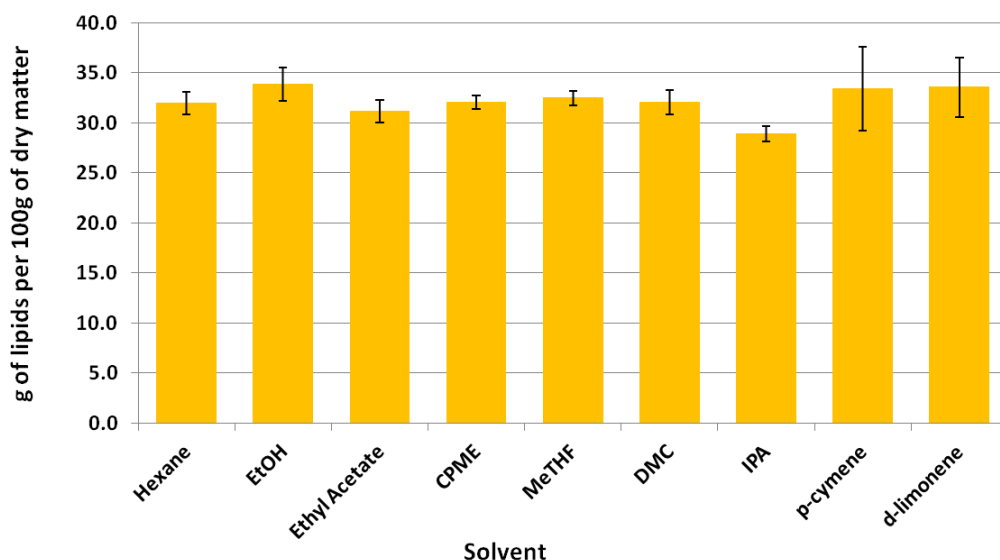


Figura 3.11. Comparación de los rendimientos de extracción de la fracción lipídica total del salmón usando los disolventes estudiados

3.3.A.4. Conclusiones Capítulo 3.A

Los estudios de simulación de Hansen y COSMO-RS sobre el carácter de solvatación de los diferentes disolventes para una serie de clases de lípidos presentes en el aceite de salmón indicaron que el d-limoneno y el p-cimeno deberían presentar un comportamiento similar al del n-hexano para solvatar los TAGs. Además, los valores RED de solubilidad de Hansen predijeron que el CPME podría ser el mejor disolvente para solubilizar todos los compuestos considerados,

siendo el resto de los disolventes evaluados más específicos según la clase de lípidos considerada. Sin embargo, los valores obtenidos mediante el programa COSMO-RS indicaban que el CPME, EtAc y MeTHF podrían solvatar de forma similar que el n-hexano a los TAGs, DAGs, FFAs y el ergosterol.

Al considerar los datos experimentales encontrados en los análisis por cromatografía de los distintos extractos de aceite de salmón, los resultados por HP-TLC indican que los TAGs es la clase predominante de los lípidos extraídos, con 73-77%, seguido de DAGs (15-20%), FFAs (5-6%) y el ergosterol (2-3%). Sin embargo, se encontró que los extractos obtenidos con el p-cimeno y d-limoneno contenían mayores contenidos de DAGs y cantidades más reducidas de TAGs que los extractos obtenidos con los otros disolventes, a pesar de los prometedoros valores de predicción calculados mediante HSP y COSMO-RS, que indicaban que p-cimeno y d-limoneno eran teóricamente los mejores disolventes para disolver TAGs. Este fenómeno podría explicarse como consecuencia de la posible degradación de la fracción lipídica debida al uso de temperaturas más altas cuando se trabaja estos los dos terpenos estudiados. Por otra parte, las cantidades de los ácidos grasos extraídos, obtenidos por GC-FID, indicaron que todos los disolventes proporcionan un perfil de ácidos grasos similar excepto el d-limoneno y, especialmente, el p-cimeno, que proporcionaron pequeñas cantidades de algunos ácidos grasos insaturados, probablemente debido al hecho de que los enlaces dobles presentes en los PUFAs son más propensos a la oxidación térmica como consecuencia de las temperaturas más altas implicadas en el uso de los dos terpenos estudiados.

La estrategia planteada a través de este estudio podría ser utilizada como una guía práctica para buscar un disolvente verde alternativo al n-hexano para la extracción de lípidos en otro tipo de muestras.

En resumen, se puede concluir que existen algunas diferencias entre los valores de solvatación teóricos predichos y los estudios experimentales sobre el uso de disolventes alternativos para las extracciones de lípidos, pero que estas diferencias podrían deberse al hecho que las simulaciones computacionales no consideran algunos aspectos como el tiempo y la temperatura de extracción, u otros parámetros como la entalpía de vaporización o los puntos de fusión de los compuestos considerados. Estas consideraciones en los modelos predictivos pueden modificar la fracción extraída lipídica. En general, los resultados experimentales mostraron que el EtAc era el disolvente más adecuado para reemplazar al hexano, debido a sus propiedades para extraer todas las clases de lípidos, tales como TAGs, DAGs, FFAs y ergosterol, en las muestras de salmón liofilizado. Si además se consideran otros parámetros como punto de ebullición,

Log P, categoría de toxicidad y energía requerida para la evaporación del disolvente, se puede concluir que EtAc podría ser considerado como un buen disolvente alternativo para la extracción de la fracción lipídica del salmón. Finalmente debe tenerse en cuenta que el EtAc es un solvente biodegradable y no tiene efectos tóxicos, por lo que puede ser reconocido como un disolvente más adecuado y alternativo al empleo de n-hexano.

Por otro lado, la estrategia expuesta en este trabajo podría ser también útil para evaluar disolventes alternativos en otros campos con el fin de mantener la capacidad de extracción de los disolventes clásicos, pero mejorando la sostenibilidad de los procesos y, de esta manera, evitar efectos secundarios.

3.3.B. Capítulo 3.B. Determinación de ácidos grasos y clases de lípidos en el aceite de salmón por espectroscopia de infrarrojo cercano

Dado que se disponía de un conjunto de muestras de extractos de aceite de salmón, de las que se conocía su composición cualitativa y cuantitativa, se valoró la posibilidad de desarrollar un método sencillo y directo, basado en la espectroscopia infrarroja NIR, que permitiera analizar de forma no destructiva la composición de fracción lipídica del extracto de salmón. Para ello, a partir de las muestras de extractos de salmón obtenidas en el estudio anterior, se realizaron mezclas hasta obtener un total de 54 muestras de aceite de salmón.

3.3.B.1. Resultados y discusión

3.3.B.1.1. Espectros NIR

A partir de las muestras de los extractos contenidas en viales de vidrio de 6 mm de diámetro interno se obtuvieron los espectros NIR de transmitancia en el intervalo comprendido entre 14000 y 4500 cm^{-1} . Cada espectro fue adquirido acumulando 50 barridos y usando una resolución óptica de 4 cm^{-1} , empleando como *background* el espectro de un vial de vidrio vacío bajo las mismas condiciones instrumentales que las empleadas para las mediciones de la muestra. Se obtuvieron tres medidas de cada muestra girando la posición del vial entre las repeticiones. Se emplearon los valores promedios de los espectros triplicados de cada muestra para construir modelos de calibración PLS y predicción.

En la Figura 3.12 pueden verse los espectros NIR por transmitancia de todas las muestras de aceite de salmón consideradas en este estudio en el rango espectral comprendido entre 14000 y 4500 cm^{-1} . Se pueden observar unas bandas de absorción alrededor de 8560 y 8261 cm^{-1} que están asociadas con los segundos armónicos de las vibraciones de estiramiento del enlace C-H de los grupos CH_3 y CH_2 , respectivamente, mientras que las señales situadas a 7181 y 7065 cm^{-1} se deben a las bandas de combinación de los enlaces C-H. Las bandas más intensas del espectro son las que están situadas a 5792 y 5676 cm^{-1} , que corresponden con el primer sobretono de las vibraciones de estiramiento del enlace C-H de los grupos CH_3 , CH_2 y $\text{CH}=\text{CH}$; mientras que las pequeñas bandas situadas a 4654 cm^{-1} están asociadas con las bandas de combinación de las vibraciones de estiramiento de los enlaces C-H y C-O [38].

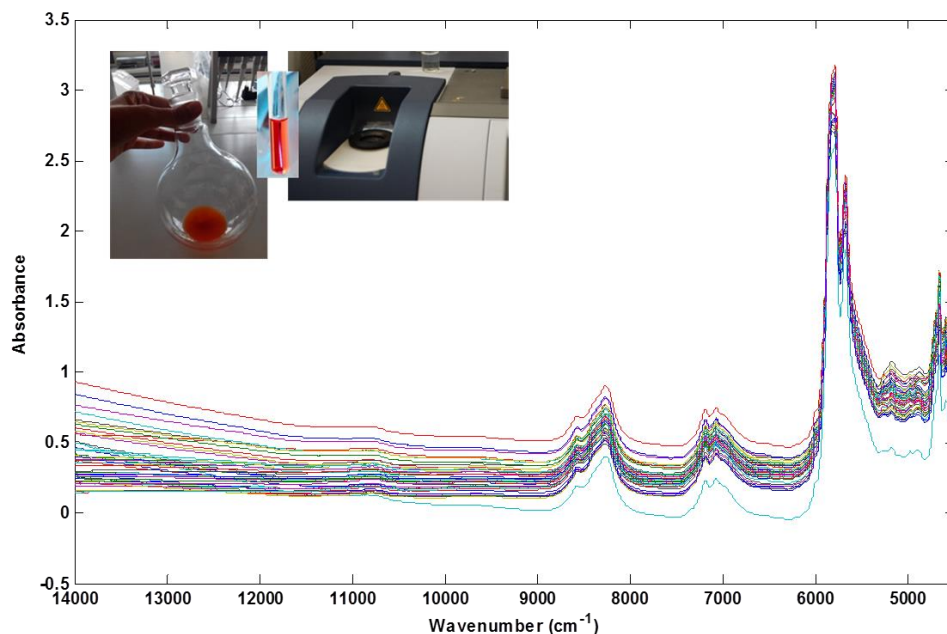


Figura 3.12. Espectros NIR de transmisión de las muestras de aceite de salmón en la región entre 14000 y 4500 cm^{-1} sin ningún tratamiento previo de datos. Inserto: Balón de extracción y vial de medición del NIR.

3.3.B.1.2. Modelos PLS-NIR

Se desarrollaron diferentes modelos PLS-NIR para evaluar las posibilidades de predicción de i) los ácidos grasos principales en el aceite de salmón como son PA, OA, ALA y LA; ii) la familias de ácidos grasos como Omega-3, Omega-6, PUFAs, MUFAs y SFAs y, iii) clases de lípidos como TAGs, DAGs, FFAs y el ergosterol.

Para la familia de Omega-3 se ha considerado la suma de ALA, ácido eicosatrienoico (ETE, C20:3n-3), EPA, DPA y DHA; mientras que la Omega-6 corresponde a la suma de LA, ácido gamma-linolénico (GLA; C18:3n-6), ácido eicosadienoico (C20: 2n-6), ácido dihomo-gamma-linolénico (DGLA, C20:3n-6) y AA. Los PUFAs se establecieron a partir de la suma de todos los ácidos grasos que contienen más de un doble enlace en su esqueleto; los MUFAs se calcularon como la suma de ácidos grasos que tienen un doble enlace en su cadena, y los valores de SFAs considerados corresponden a la suma de todos los ácidos grasos con enlaces simples.

Para construir los modelos PLS-NIR se utilizaron un total de 54 muestras de extractos de aceite de salmón y mezclas de ellas, de las cuales 38 muestras seleccionadas mediante el algoritmo K-S, se utilizaron como conjunto de

calibración y las 16 restantes, como conjunto de validación externo. Diferentes pre-tratamientos espectrales con diferentes regiones fueron probados para encontrar los mejores modelos PLS mediante validación cruzada, siendo el pre-tratamiento escogido en todos los casos SNV, FD con un suavizado de 13 puntos y MC. iPLS se empleó para seleccionar la región espectral y, para ello, los espectros se dividieron en intervalos de 494 variables. En la Figura 3.13a puede verse el espectro NIR promedio de todas las muestras de aceite de salmón, con las regiones seleccionadas automáticamente usando el iPLS y considerando valores correspondientes a los compuestos de Omega-3. Para construir el modelo PLS-NIR usando el rango espectral antes mencionado, se seleccionó el número óptimo de LV's de acuerdo con los valores de RMSECV (véase la Figura 3.13b).

La Figura 3.13c presenta la regresión entre los valores predichos y de referencia para los compuestos de Omega-3, obteniéndose una buena distribución de los puntos de calibración y validación, ya que se sitúan próximos a la línea de regresión óptima con una intersección de 0 y una pendiente de 1. La Figura 3.13d presenta el gráfico de puntuaciones para el primer, segundo y tercer componente principal, obtenidos en el PCA de las muestras de aceite de salmón después del tratamiento previo de SNV, FD y MC con la región seleccionada por el iPLS, construido como método exploratorio. Los tres primeros componentes principales representan el 96.5% de la varianza explicada, siendo 71.79, 14.43 y 10.28% para PC1, PC2 y PC3, respectivamente. El PCA indica que las muestras de conjunto de calibración pueden ser utilizadas para construir el modelo PLS-NIR para la determinación de Omega-3, ya que son representativas de todo el conjunto de datos.

Los valores de concentración media, mínima y máxima de los compuestos estudiados obtenidos mediante el uso CG-FID y HP-TLC, utilizados como datos de referencia, se muestran en la Tabla 3.5, diferenciando los correspondientes conjuntos de calibración y de validación.

Figura 3.13. Estrategia para construir el modelo de regresión PLS-NIR aplicado a los compuestos Omega-3. a) Selección de la región espectral utilizando el intervalo del mínimo cuadrado parcial (iPLS). b) Selección del número de variables latentes según el error cuadrático medio de validación cruzada (RMSECV). c) Regresión entre los valores predichos y de referencia. d) Diagrama de las scores del PCA utilizado para asegurar la representatividad del conjunto de calibración.

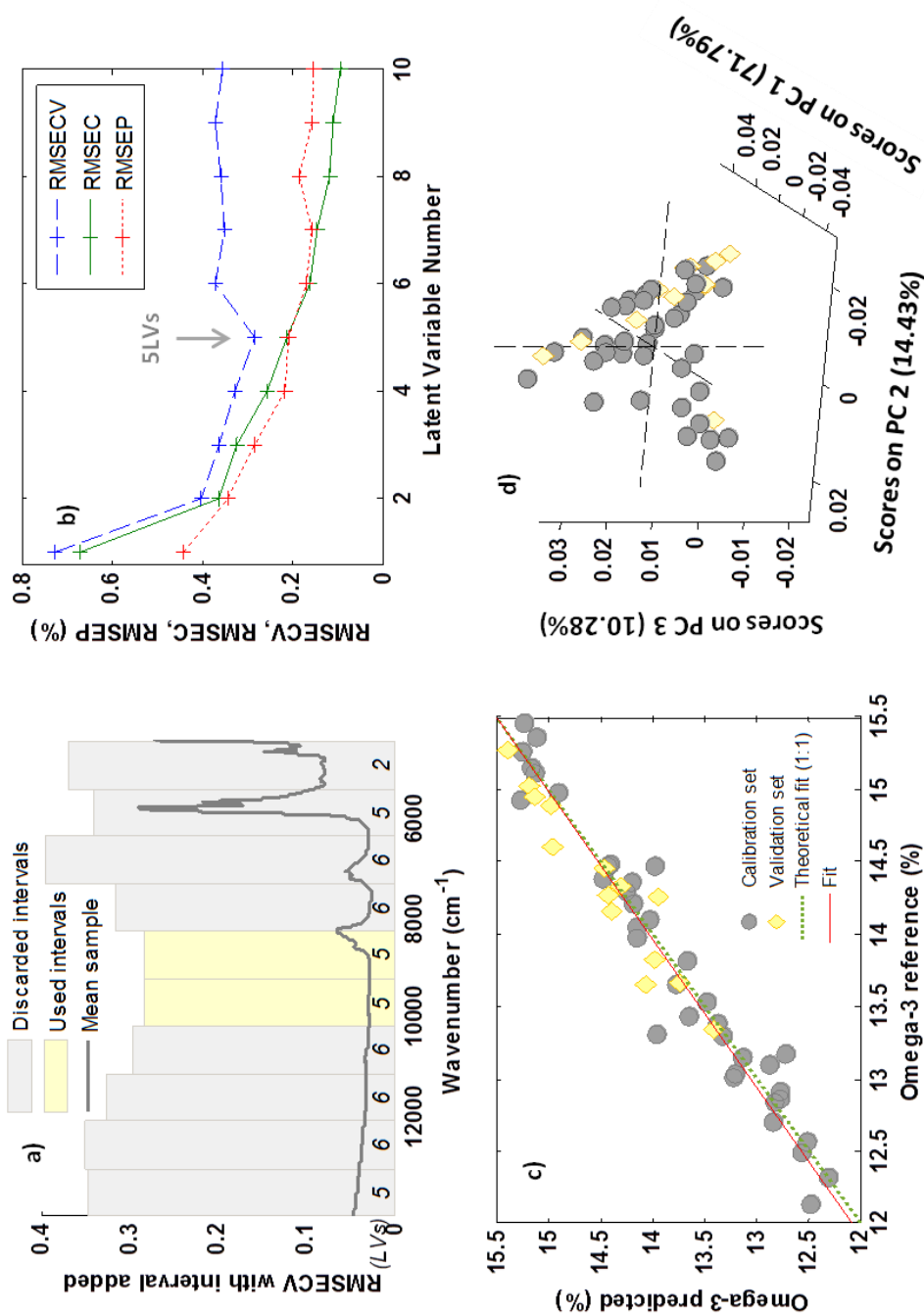


Tabla 3.5. Descriptores de los conjuntos de calibración y de validación empleados para la modelización PLS-NIR de los contenidos de ácidos grasos y clases de lípidos en las muestras de extractos de aceite de salmón.

Analito	Conjunto	Número de muestras	Rango (% m/m)	Promedio ± SD
Ácido Palmítico (PA; C16:0)	Cal	37	11.19-11.75	11.51 ± 0.13
	Val	14	11.28-11.82	11.48 ± 0.14
Ácido Oleico (OA; C18:1)	Cal	36	37.98-39.63	38.94 ± 0.42
	Val	14	37.9-39.44	38.89 ± 0.37
Ácido Linoleico (LA; C18:2n6)	Cal	37	16.17-17.34	17.01 ± 0.26
	Val	14	16.76-17.33	17.11 ± 0.15
Ácido Linolenico (ALA; C18:3n3)	Cal	37	5.16-5.57	5.36 ± 0.13
	Val	14	5.24-5.61	5.44 ± 0.11
Omega-3	Cal	37	12.12-15.46	13.77 ± 0.94
	Val	14	13.35-15.27	14.33 ± 0.58
Omega-6	Cal	37	18.21-19.62	19.25 ± 0.33
	Val	14	18.94-19.58	19.36 ± 0.17
Ácidos Grasos Poliinsaturados (PUFAs)	Cal	37	30.78-34.76	33.05 ± 1.15
	Val	14	32.59-34.85	33.72 ± 0.62
Ácidos Grasos Monoinsaturados (MUFAs)	Cal	36	43.24-45.23	44.31 ± 0.51
	Val	14	43.18-44.99	44.20 ± 0.42
Ácidos Grasos Saturados (SFAs)	Cal	37	19.62-20.22	19.89 ± 0.17
	Val	14	19.48-20.17	19.80 ± 0.20
Triacilgliceridos (TAGs)	Cal	37	72.63-76.01	74.16 ± 1.22
	Val	14	73.20-76.01	75.06 ± 0.89
Diacilgliceridos (DAGs)	Cal	37	16.06-18.94	17.43 ± 1.17
	Val	14	15.06-18.53	16.57 ± 0.93
Ácidos Grasos Libres (FFAs)	Cal	37	4.98-6.44	5.85 ± 0.41
	Val	14	5.16-6.44	5.93 ± 0.32
Ergosterol	Cal	37	2.22-3.12	2.51 ± 0.26
	Val	14	2.22-3.0	2.44 ± 0.18

Cal: Calibración; Val: Validación; SD: desviación estándar

Cabe mencionar que se eliminó una muestra de todos los conjuntos de calibración y dos muestras de todos los conjuntos de validación debido a sus altos valores de Q residual y T^2 de Hotelling, lo que sugería que se trataban de muestras anómalas. En el caso de OA y MUFAs, se eliminó una muestra adicional de los conjuntos de calibración por la misma razón. Los parámetros de calibración más importantes de los modelos PLS-NIR desarrollados para determinación de los ácidos grasos mencionados y clases lipídicas se muestran resumidos en la Tabla 3.6. En esta tabla se indican la región espectral y el número de LV's utilizado para el modelo seleccionado para cada compuesto junto con los valores de RMSEC, RMSECV y los coeficientes de determinación de la calibración (R^2 Cal) y validación cruzada (R^2 CV). Como puede verse a partir de los datos de la Tabla 3.6, la mayoría de los parámetros podrían predecirse basándose en un número reducido de variables latentes. Sin embargo, para las clases de lípidos, que involucran varios compuestos, se requieren 6 LV's en el caso de FFA, y 7 LV's para los SFAs.

Los coeficientes de determinación para los mejores modelos de calibración encontrados estaban comprendidos entre 0.72 y 0.99, independientemente de si se trataba de un solo compuesto o una clase de lípidos, con bajos valores de R^2 para la validación cruzada de SFAs y PA.

Tabla 3.6. Parámetros de los modelos PLS-NIR desarrollados para la determinación de los ácidos grasos y clases de lípidos en el aceite de salmón

Analyte	Región espectral (cm ⁻¹)	LV	Conjunto de Calibración			Conjunto de Validación			
			RMSEC (% w/w)	R ² Cal	RMSECV (% w/w)	R ² CV	RMSEP (% w/w)	R ² Pred	RRMSEP (%)
Ácido Palmítico (PA; C16:0)	13044.5-12093.7	3	0.07	0.72	0.09	0.55	0.58	0.78	1.54
	5422.9-4472.2								
Ácido Oleico (OA; C18:1)	6375.7-5424.9	5	0.14	0.88	0.18	0.80	0.65	0.75	1.27
	5422.9-4472.2								
Ácido Linoleico (LA; C18:2-n6)	5422.9-4472.2	4	0.06	0.95	0.09	0.86	0.70	0.50	1.76
	13997.1-13046.4								
Ácido Linolénico (ALA; C18:3-n3)	11139.1-10188.3	4	0.05	0.85	0.07	0.7	0.60	1.35	1.48
	7328.4-6377.6								
	5422.9-4472.2								
Omega-3	10186.4-8282.9	5	0.21	0.95	0.28	0.91	0.92	1.41	2.84
	5423.0-4472.2								
Omega-6	5423.0-4472.2	4	0.07	0.96	0.11	0.89	0.73	0.51	1.71
Ácidos Grasos Poliinsaturados (PUFAs)	6375.7-4472.2	5	0.19	0.97	0.35	0.91	0.91	0.76	2.41
Ácidos Grasos Monoinsaturados (MUFAs)	13997.1-4472.2	3	0.19	0.86	0.25	0.75	0.61	0.69	1.39
Ácidos Grasos Saturados (SFAs)	11139.1-10188.3 7328.4-6377.6	7	0.08	0.74	0.13	0.40	0.63	0.62	1.62
Triglicéridos (TAGs)	5422.9-4472.2	4	0.25	0.96	0.34	0.92	0.94	0.29	4.00
Diglicéridos (DAGs)	11139.1-10188.3 5422.9-4472.2	4	0.27	0.95	0.40	0.88	0.90	1.78	3.16
Ácidos Grasos Libres (FFAs)	13997.1-4472.2	6	0.05	0.99	0.09	0.95	0.96	1.32	4.10
Ergosterol	13997.1-4472.2	3	0.04	0.98	0.05	0.97	0.98	1.12	6.44

Para evaluar la capacidad predictiva de los modelos PLS-NIR, se utilizaron las muestras del conjunto de validación que no se utilizaron durante la etapa de calibración. Las ecuaciones de regresión entre los valores predichos PLS-NIR y los valores de referencia, tanto para el conjunto de calibración como el de validación, se resumen en la Tabla 3.7. En líneas generales las pendientes y ordenadas en el origen se acercan a valores teóricos de 1 y 0, respectivamente. Sin embargo, en el caso de los MUFAs, la ordenada en el origen de la regresión entre los valores predichos y de referencia también fue considerablemente alto, con un valor de 18.8.

Los principales parámetros de validación de los modelos PLS-NIR desarrollados se indican en la Tabla 3.6.

Dentro de cada modelo, se obtuvieron valores similares para RMSECV y el RMSEP, lo que indica una buena correlación para la determinación de los ácidos grasos y las clases de lípidos antes mencionadas. Se obtuvieron valores de RRMSEP entre 0.29 y 1.78% m/m, con valores de RPD en el intervalo de 1.27 y 6.44. En muchos casos, estos resultados demuestran una buena capacidad predictiva de los modelos PLS-NIR construidos para la determinación de los principales compuestos de los extractos de aceite de salmón. Los mejores resultados se obtuvieron para las clases de lípidos, especialmente las de ergosterol, FFAs, TAGs y DAGs, seguido por las familias de determinaciones de ácidos grasos como Omega-3 y PUFAs.

Tabla 3.7. Rectas de regresión entre los valores de PLS-NIR predicho y los valores de referencia correspondientes a los conjuntos de calibración y de validación, en la determinación de ácidos grasos libres y clases lipídicas en extractos de aceite de salmón.

Analito	Regresión	
	Calibración	Predicción
Ácido Palmítico (PA; C16:0)	$y = 0.715x + 3.28$	$y = 0.551x + 5.13$
Ácido Oleico (OA; C18:1)	$y = 0.887x + 4.39$	$y = 0.832x + 6.32$
Ácido Linoleico (LA; C18:2n-6)	$y = 0.949x + 0.862$	$y = 0.7380x + 4.45$
Ácido Linolénico (ALA; C18:3n-3)	$y = 0.847x + 0.818$	$y = 0.435x + 3.09$
Omega-3	$y = 0.950x + 0.685$	$y = 0.996x + 0.173$
Omega-6	$y = 0.957x + 0.830$	$y = 0.774x + 4.33$
Ácido Grasos Poliinsaturados (PUFAs)	$y = 0.972x + 0.918$	$y = 1.126x - 4.13$
Ácido Grasos Monoinsaturados (MUFAs)	$y = 0.860x + 6.21$	$y = 0.570x + 18.83$
Ácido Grasos Saturados (SFAs)	$y = 0.739x + 5.20$	$y = 0.659x + 6.76$
Triglicéridos (TAGs)	$y = 0.957x + 3.17$	$y = 0.885x + 8.64$
Diglicéridos (DAGs)	$y = 0.947x + 0.921$	$y = 0.802x + 3.28$
Ácido Grasos Libres (FFAs)	$y = 0.987x + 0.074$	$y = 0.855x + 0.829$
Ergosterol	$y = 0.976x + 0.064$	$y = 0.982x + 0.057$

3.3.B.1.3. Evaluación de la sostenibilidad del método propuesto

La metodología PLS-NIR propuesta se evaluó utilizando la escala ecológica propuesta por Galuszkay *et al.* [39] y modificada por Armenta *et al.* [40], para establecer un "Certificado Verde" del método. La Figura 3.11 muestra las puntuaciones de acuerdo con el Certificado Verde para evaluar el método FT-NIR propuesto y los métodos CG-FID y HP-TLC utilizados como metodologías de referencia. Esta escala ecológica se basa en la aplicación de códigos de color y letras (A a G) para designar categorías, siendo la clase A la correspondiente a los métodos analíticos más ecológicos.

Sobre la base de una puntuación de 100 puntos, se asignaron cero puntos de penalización en la escala ecológica al procedimiento FT-NIR propuesto para determinación simultánea de ácidos grasos y clases de lípidos. Esto se debe a la falta de consumo de reactivos, generación de residuos y emisiones nocivas, junto con el bajo consumo de energía del instrumento empleado. Con respecto al método de referencia no es posible realizar la determinación simultánea de los ácidos grasos y de las clases de lípidos en un solo ensayo, es por ello que se realizaron dos evaluaciones diferentes. Para la determinación de ácidos grasos por CG-FID se asignaron 13.2 puntos de penalización debido al uso de reactivos (5.7), energía consumida por el sistema cromatográfico (3) y la cantidad de residuos generados (4.5), obteniéndose una puntuación total final de 86.8. Para la determinación de las clases de lípidos por HP-TLC, la puntuación total fue 75.1 debido a la asignación de 24.9 puntos de penalización (ver detalle en la Figura 11). Por lo tanto, puede concluirse que el procedimiento PLS-NIR proporciona un método de Certificado Verde de categoría A con la máxima puntuación de 100 puntos para la determinación de ácidos grasos y clases lipídicas, mientras que al método CG-FID de referencia utilizado para la determinación de ácidos grasos correspondería asignarle la categoría B y, finalmente, la determinación de las clases de lípidos por HP-TLC se le asigna a la categoría C de la escala de certificación verde.

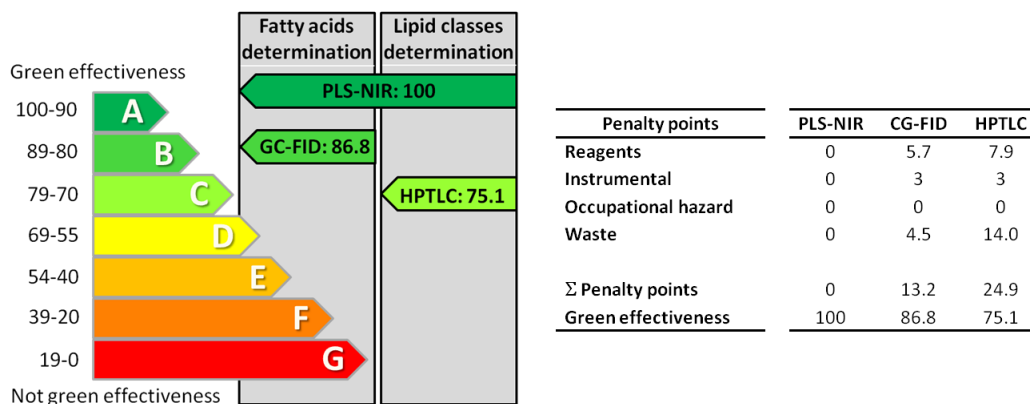


Figura 3.11. Evaluación verde del método PLS-NIR propuesto y su comparación con las metodologías de referencia CG-FID y HP-TLC utilizadas para la determinación de ácidos grasos y clases lipídicas en los extractos de aceite de salmón.

3.3.B.2. Conclusiones Capítulo 3.B.

El método PLS-NIR desarrollado permite una adecuada cuantificación simultánea de varios ácidos grasos, como OA, PA, ALA y LA, así como Omega-3, Omega-6, PUFAs, MUFAs y SFAs, además de las clases de lípidos como TAGs, DAGs, FFAs y ergosterol en extractos de aceite de salmón. El método propuesto evitando el uso de reactivos tóxicos o peligrosos y supone una drástica reducción del tiempo requerido y del consumo de energía. Sin embargo, la exactitud en la capacidad predictiva del método únicamente mantiene el nivel de detección para el ácido palmítico, linolénico y los MUFAs.

En resumen, se puede concluir que los modelos PLS-NIR ofrecen una herramienta de alto valor de detección para la caracterización de muestras de extractos de aceite de salmón con una excelente capacidad predictiva de ergosterol, FFAs, DAGs, TAGs, Omega-3 y PUFAs. A esto hay que destacar la prácticamente nula manipulación de las muestras, que se trata de una técnica directa y no destructiva, con una gran rapidez en su aplicación.

3.3.3. Referencias Capítulo 3

- [1] P. Manirakiza, A. Covaci, P. Schepens, Comparative Study on Total Lipid Determination using Soxhlet, Roese-Gottlieb, Bligh & Dyer, and Modified Bligh & Dyer Extraction Methods, *J. Food Compos. Anal.*, 14(2001) 93–100.
- [2] C.J. Lavie, R.V. Milani, M.R. Mehra, H.O. Ventura, Omega-3 Polyunsaturated Fatty Acids and Cardiovascular Diseases, *Journal of the American College of Cardiology*, (2009) 585–594.
- [3] A.P. Simopoulos, The importance of the ratio of omega-6 / omega-3 essential fatty acids, *Biomed. Pharmacother.*, 56(2002) 365–379.
- [4] E. Patterson, R. Wall, G. F. Fitzgerald, R. P. Ross, and C. Stanton, Health Implications of High Dietary Omega-6 Polyunsaturated Fatty Acids, *J. Nutr. Metab.*, 2012(2012).
- [5] P.C. Calder, P.Yaqoob, Understanding Omega-3 Polyunsaturated Fatty Acids, *Postgrad. Med.*, 121(2009) 148–157.
- [6] G. Kaur, D. Cameron-Smith, M. Garg, A.J. Sinclair, Docosapentaenoic acid (22:5n-3): A review of its biological effects, *Prog. Lipid Res.*, 50(2011) 28–34.
- [7] K. Kim, Y.A. Nam, H. Sik, A.W. Hayes, B. Lee, α -Linolenic acid : Nutraceutical , pharmacological and toxicological evaluation, *Food Chem. Toxicol.*, 70(2014) 163–178.
- [8] G.E. Delgado, B.K. Krämer, S. Lorkowski, W. März, C. von Schacky, M.E. Kleber, Individual omega-9 monounsaturated fatty acids and mortality — The Ludwigshafen Risk and Cardiovascular Health Study, *J. Clin. Lipidol.*, 11(2017) 126:135.
- [9] M.I. Covas, V. Ruiz-Gutierrez, R. de la Torre, A. Kafatos, R.M. Lamuela-Raventós, J. Osada, R.W. Owen, F. Visioli, Minor Components of Olive Oil: Evidence to Date of Health, *Nutr. Rev.*, 64(2006) S20–S30.
- [10] Y. Wu, P. Song, J. Xu, M. Zhang, M. Zou, Activation of Protein Phosphatase 2A by Palmitate Inhibits AMP-activated Protein Kinase, *J. Biol. Chem.*, 282(2007) 9777–9788.
- [11] C.M. Duffy, J.P. Nixon, T.A. Butterick, Molecular and Cellular Neuroscience Orexin A attenuates palmitic acid-induced hypothalamic cell death, *Mol. Cell. Neurosci.*, 75(2016) 93–100.
- [12] R. P. Mensink, Health Effects of Saturated Fatty Acids, *Ref. Modul. Biomed. Sci. Encycl. Hum. Nutr.*, Third Ed., vol. 2, pp. 215–219, 2013.

- [13] F. Sahena, I.S.M. Zaidul, S. Jinap, N. Saari, H.A. Jahurul, K.A. Abbas, N.A. Norulaini, PUFAs in fish: Extraction, fractionation, importance in health, *Compr. Rev. Food Sci. Food Saf.*, 8(2009) 59–74.
- [14] M. Abdulkadir, G.I. Abubakar, a Mohammed, Production and characterization of oil from fishes, *J. Eng. Appl. Sci.*, 5(2010) 1–5.
- [15] Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures.
- [16] Public Health Statement n-Hexane, Public Heal. Serv. Agency Toxic Subst. Dis. Regist.
- [17] F. Chemat, M. Abert, *Alternative Solvents for Natural Products Extraction*, Springer. 2014.
- [18] M. Virost, V. Tomao, C. Ginies, F. Chemat, Total Lipid Extraction of Food Using d-Limonene as an Alternative to n-Hexane, *Chromatographia*, 68(2008) 311–313.
- [19] S. Bertouche, V. Tomao, A. Hellal, C. Boutekedjiret, F. Chemat, First approach on edible oil determination in oilseeds products using alpha-pinene, *J. Essent. Oil Res.*, 25(2013) 439–443.
- [20] C. Breil, A. Meullemiestre, M. Vian, F. Chemat, Bio-based solvents for green extraction of lipids from oleaginous yeast biomass for sustainable aviation biofuel, *Molecules*, 21(2016) 1–14.
- [21] A.G. Sicaire, M. Abert Vian, F. Fine, P. Carré, S. Tostain, F. Chemat, Experimental approach versus COSMO-RS assisted solvent screening for predicting the solubility of rapeseed oil, *OCL - Oilseeds fats, Crop. Lipids*, vol. 22(2015).
- [22] Y. Li, F. Fine, A.S. Fabiano-Tixier, M. Abert-Vian, P. Carre, X. Pages, F. Chemat, Evaluation of alternative solvents for improvement of oil extraction from rapeseeds, *Comptes Rendus Chim.*, 17(2014) 242–251.
- [23] W.R. Morrison, L.M. Smith, Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride-methanol, *J. Lipid Res.*, 5(1964) 600–608.
- [24] S.W. Colley, C.R. Fawcett, C. Rathmell, M.W.M. Tuck, Process for the preparation of ethyl acetate, US6809217 B1, 2004.

- [25] M. Yates, M.J. Martí, B. Casal, M. Iglesias, M. Esteban, E. Ruiz-hitzky, Synthesis of p -cymene from limonene , a renewable feedstock, 81(2008) 218–224.
- [26] V. Pace, P. Hoyos, M. Fernandez, J.V. Sinisterra, A.R. Alcantara, '2-Methyltetrahydrofuran as a suitable green solvent for phthalimide functionalization promoted by supported KF, Green Chemistry, 12(2010) 1380–1382.
- [27] M. Nardi, G. Sindona, P. Costanzo, M. Oliverio, A. Procopio, Eco-friendly stereoselective reduction of a , b -unsaturated carbonyl compounds by Er (OTf) 3 / NaBH 4 in 2-MeTHF, 71(2015) 1132–1135.
- [28] P. Kumar, V. Chandra, I. Mani, Dimethyl carbonate synthesis by transesterification of propylene carbonate with methanol: Comparative assessment of Ce-M (M ¼ Co , Fe , Cu and Zn) catalysts, 88(2016).
- [29] R.K. Henderson, C. Jiménez-González, D.J.C. Constable, S.R. Alston, G.G.A. Inglis, G. Fisher, J. Sherwood, S.P. Binks, A.D. Curzons, Expanding GSK's solvent selection guide – embedding sustainability into solvent selection starting at medicinal chemistry, Green Chem, 13(2011) 854–862.
- [30] C. M. Hansen, Hansen Solubility Parameters, Second edi. 2007.
- [31] A. Klamt, The COSMO and COSMO-RS solvation models, Wiley Interdiscip. Rev. Comput. Mol. Sci., 1(2011) 699–709.
- [32] C.M. Hansen, The Universality of the Solubility Parameter, Ind. Eng. Chem. Prod. Res. Dev., 8(1969) 2–11.
- [33] B. Muik, B. Lendl, A. Molina-Díaz, M.J. Ayora-Cañada, Direct monitoring of lipid oxidation in edible oils by Fourier transform Raman spectroscopy, Chem. Phys. Lipids, 134(2005) 173–182.
- [34] H. Vaskova, M. Buckova, Thermal Degradation of Vegetable Oils: Spectroscopic Measurement and Analysis, Procedia Eng., 100(2015) 630–635.
- [35] J. Kuligowski, G. Quintás, S. Garrigues, M. De La Guardia, Monitoring of Polymerized Triglycerides in Deep-Frying Oil by On-Line GPC-FTIR Spectrometry Using the Science Based Calibration Multivariate Approach, Chromatographia, 71(2010) 201–209.
- [36] T. Loftsson, B. Ilievska, G.M. Asgrimsdottir, O.T. Ormarsson, E. Stefansson, Fatty acids from marine lipids: Biological activity, formulation and stability, J. Drug Deliv. Sci. Technol., 34(2016) 71–75.

- [37] C. Gertz, Chemical and physical parameters as quality Deep-Frying Process – Changes at Elevated Temperature, *Quality*, 102(2000) 566–572.
- [38] N. Sinelli, L. Cerretani, V. Di Egidio, A. Bendini, E. Casiraghi, Application of near (NIR) and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity, *Food Res. Int.*, 43(2010) 369–375.
- [39] A. Gałuszka, P. Konieczka, Z.M. Migaszewski, Analytical Eco-Scale for assessing the greenness of analytical procedures, *Trends Anal. Chem.*, 37(2012) 61–72.
- [40] S. Armenta, M. de la Guardia, J. Namiesnik, Green Microextraction. M. Valcárcel, S. Cárdenas, R. Lucenas, eds. *Analytical Microextraction Techniques*. Bentham Science, 2017.

3.4.CAPÍTULO 4:
Control de calidad de los aditivos umami por Raman y
espectroscopia infrarroja en la región cercana

3.4. Capítulo 4. Control de calidad de los aditivos umami por espectroscopia Raman e infrarroja en la región cercana

3.4.1. Introducción

Umami es el quinto sabor sugerido por Ikeda [1] para referirse a una nueva sensación de sabor independiente de los cuatro sabores básicos clásicos: amargo, dulce, salado y ácido. Las especiales características del sabor umami se han atribuido al sinergismo entre el glutamato monosódico (MSG, del inglés *Monosodium glutamate*), la guanosina 5'-monofosfato (GMP, del inglés *Guanosine-5'-monophosphate*) y la inosina-5'-monofosfato (IMP, del inglés *Inosine-5'-monophosphate*), estos dos últimos encontrados normalmente en una proporción 1:1, que, cuando se mezclan, aumentan la intensidad del sabor umami en los alimentos [2, 3]. El Codex Alimentarius ha reconocido el MSG (E621), el GMP (E627) y el IMP (E631) como aditivos alimentarios, clasificándolos como potenciadores de sabor, que sujetos a las disposiciones de esta Norma se utilizarán bajo las condiciones de buenas prácticas de fabricación (BPF) [4]. El Comité Mixto FAO/OMS de Expertos en Aditivos Alimentarios (JECFA, del inglés *Joint FAO/WHO Expert Committee on Food Additives*) no especificó una ingesta diaria aceptable (ADI, *Acceptable Daily Intake*) para estos aditivos [5], indicando que tienen una toxicidad baja, no representan ningún peligro para la salud humana, y no se considera necesario establecer una ADI numérica. El Comité Científico de la Alimentación Humana (SCF, del inglés *Scientific Committee on Food*) de la Unión Europea (UE) llegó a una conclusión equivalente al JECFA, al establecer un grupo no especificado de ADI para el ácido L-glutámico y sus sales [6]. Sin embargo, algunos estudios afirman que altas concentraciones de estos compuestos pueden causar problemas de salud en las personas que son sensibles a ellos [7] [8]. Estudios toxicológicos no han demostrado efectos tóxicos específicos de MSG a pesar de las sospechas tempranas [9]. Se publicaron informes anecdóticos de síntomas subjetivos tras la ingestión de comidas chinas [10], siendo el MSG inicialmente implicado como agente causante del llamado "Síndrome del restaurante chino" [11]. La revisión de Kenney [12] de los estudios y encuestas experimentales existentes concluyó que las evidencias no pudieron demostrar que el MSG era un agente causal en la provocación de la gama completa de síntomas del "Síndrome del restaurante chino".

Las sustancias umami se agregan con frecuencia, como premezcla de sales, en una gran variedad de alimentos tales como carne, pescados, mariscos, vegetales y productos elaborados. Así que el análisis cuantitativo de las mezclas industriales de aditivos umami es un primer paso crítico durante el proceso de fabricación de las mismas para asegurar la calidad de los productos finales producidos. El análisis

de la mezcla umami puede proporcionar especificaciones completas y precisas desde el punto de vista de la calidad y la seguridad alimentaria. Además, la identificación de materiales no conformes puede mejorar el desempeño operacional y ahorrar tiempo y costes durante el proceso de fabricación. El control de procesos requiere monitorizar, a poder ser en tiempo real, los puntos críticos de procesamiento mediante el uso de métodos analíticos rápidos y precisos para asegurar la calidad del producto que se está fabricando, siendo de interés en el caso de los productos de aditivos umami.

Los métodos tradicionales para determinar los tres potenciadores de sabor de una manera simultánea están basados en el empleo de las técnicas cromatográficas; implican procedimientos de múltiples etapas, con gran consumo de tiempo y un coste considerable, además de requerir disolución previa de la muestra [13] [14] [15]. Debido al gran interés en encontrar métodos analíticos alternativos con el fin de obtener procedimientos más rápidos, económicos y respetuosos con el medio ambiente, se ha propuesto el uso de la espectrofotometría con red de diodos empleando la quimométrica para la determinación simultánea de MSG, GMP e IMP, n en cubitos de caldo concentrado [16] y en preparaciones alimenticias [17]. Sin embargo, en estos métodos se requiere también la disolución de la muestra, aumentando así la manipulación de la muestra.

La lengua electrónica (e-Tongue) también se ha utilizado para evaluar el sabor umami en muestras líquidas y comparar diferentes potenciadores del sabor mediante el análisis de componentes principales (PCA) [18] o para evaluar la intensidad del sabor usando el PLS [19] [20]. Dependiendo del instrumento, se puede requerir mucho tiempo para limpiar los sensores o incluso ser necesaria una gran cantidad de muestra centrifugada para ser analizada y, además, el protocolo de limpieza del sensor puede ser limitado y puede producir una deriva en el mismo.

La espectroscopia Raman y el infrarrojo cercano (NIR) son herramientas no destructivas bien establecidas para el análisis directo multicomponente en el control de calidad de alimentos. Estas técnicas resultan complementarias y permiten tanto el estudio de la composición química de las muestras, como la determinación de algunos de los compuestos presentes. Bagnasco *et al.* [20] correlacionaron los espectros NIR de transmisión de mezclas de aminoácidos/péptidos y muestras de ribonucleótidos con el atributo umami en formulaciones de potenciadores de sabor, obteniendo un RMSECV del 21%. Sin embargo, en la revisión bibliográfica realizada, hasta nuestro conocimiento, no hay precedentes sobre el empleo de la espectroscopia Raman y NIR para la determinación de MSG, GMP e IMP.

El objeto de este estudio ha sido el desarrollo de un método analítico para la determinación de MSG, GMP e IMP en los aditivos alimentarios industriales, potenciadores del sabor umami.

3.4.2. Parte experimental

3.4.2.1. Instrumentación

Los espectros Raman fueron obtenidos mediante el uso de un equipo Raman portátil de la marca StellarNet Inc. (Tampa, FL, EUA), equipado con un láser de diodo de 785 nm para excitación, con una potencia de salida máxima de 120 mW, y un detector de dispositivo de carga acoplada termoelectricamente. Para el control instrumental y de medición, así como para la adquisición de datos, se utilizó el software SpectraWiz de StellarNet Inc. (Tampa, FL, EUA).

Los espectros del infrarrojo cercano registrados en este trabajo fueron adquiridos en un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Germany) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS. El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

3.4.2.2. Muestras

Las muestras utilizadas en el estudio se obtuvieron de una empresa fabricante de mezcla de aditivos alimentarios empleados en distintas aplicaciones agroalimentarias. Se emplearon 35 muestras binarias de GMP e IMP, que fueron divididas en tres conjuntos: calibración (19), validación (9) y predicción (7). Además, se emplearon 26 muestras de mezclas ternarias de MSG, GMP e IMP divididas en calibración (14), validación (6) y predicción (6). En este último caso, las mezclas de aditivos alimentarios contenían concentraciones de MSG comprendidas entre el 80 y el 99% m/m, siendo el porcentaje restante el correspondiente a mezclas binarias de GMP e IMP en una proporción aproximada 1:1 m/m.

El conjunto de calibración fue utilizado para la construcción del modelo PLS, el conjunto de validación se empleó en este caso para seleccionar el mejor modelo predictivo con respecto el menor error de validación, y el conjunto de predicción se utilizó, una vez escogido el modelo, para comprobar la capacidad predictiva del modelo construido.

Las mezclas se colocaron en viales de vidrio transparente de 11 mm de diámetro interno y 25 mm de altura, para obtener directamente los espectros Raman y NIR.

3.4.3. Resultados y discusión

3.4.3.1. Espectroscopia Raman

Los espectros Raman se obtuvieron en el intervalo entre 160 y 2850 cm^{-1} de desplazamiento Raman, acumulando 1 barrido de 6000 ms por registro. Todos los espectros Raman se realizaron a una potencia de 120 mW, a una distancia de 5.4 mm entre el vial de muestra y la sonda. Para ello, se diseñó un accesorio de aluminio para medir directamente a través del fondo del vial de la muestra. Este dispositivo nos permitía de manera fácil variar la distancia entre la muestra y la sonda con el objeto de conseguir la máxima señal (ver Figura 1). La señal de fondo debida al vial de vidrio se corrigió midiendo un patrón de NaCl en las mismas condiciones que las muestras y restando posteriormente las señales. Antes de las medidas de cada muestra se realizó un espectro de la oscuridad como referencia. Se obtuvieron tres medidas de cada muestra y se empleó el promedio de los espectros triplicados de cada muestra para fines quimiométricos.

En la Figura 4.1 pueden verse los espectros Raman, sin ningún tratamiento previo, en la región entre 2850 y 160 cm^{-1} , que corresponden a patrones puros de MSG, GMP e IMP, junto con mezclas de GMP + IMP y MSG + GMP + IMP. Se puede observar que las bandas más intensas aparecen en el intervalo de desplazamiento Raman de 1700 a 200 cm^{-1} . Las bandas más características del espectro de MSG se encuentran alrededor de 1434 y 1401 cm^{-1} y son debidas al tijereteo del grupo CH_2 y del estiramiento simétrico del grupo COO^- , respectivamente. Las bandas presentes cerca de 940 cm^{-1} están relacionadas con los modos de estiramiento de los enlaces C-C, mientras que las que se observa alrededor de 860 cm^{-1} correspondería a los modos de deformación de grupo COOH [21] [22].

Cabe destacar que los espectros de GMP e IMP son muy similares entre sí, siendo muy difícil la identificación/diferenciación visual de dichas sustancias, pero que sí se distinguen bien del MSG, ya que éste último presenta bandas más pronunciadas en el intervalo de 1400 a 900 cm^{-1} .

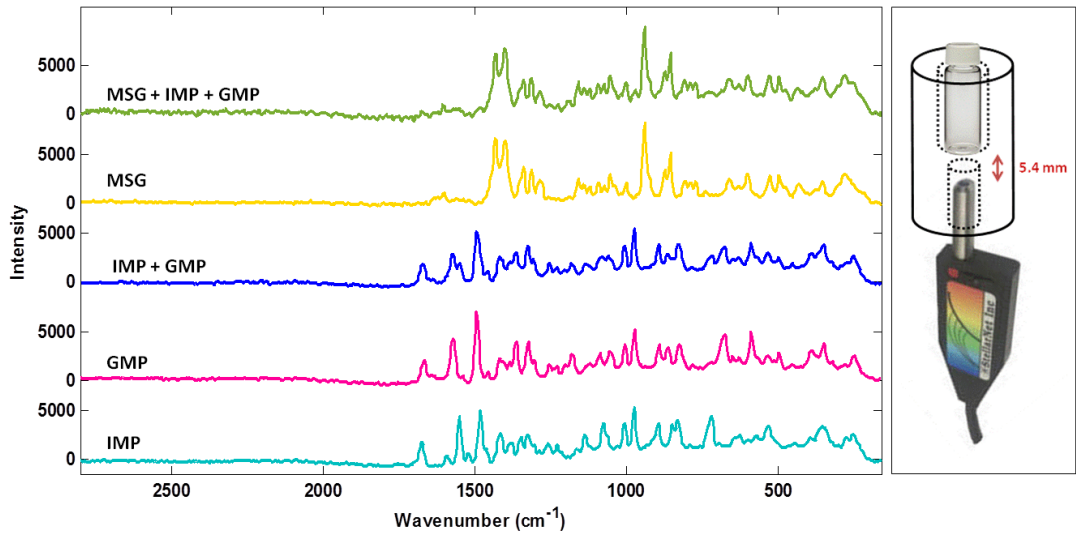


Figura 4.1. a) Espectros Raman de glutamato monosódico (MSG), guanósina 5'-monofosfato (GMP), inosina-5'-monofosfato (IMP) y mezclas de estos aditivos en la región entre 2850 y 160 cm^{-1} sin ningún tratamiento previo de los datos. b) Detalle del accesorio para realizar las medidas directamente sobre el vial utilizando una sonda de fibra óptica.

3.4.3.1.1. Modelos PLS-Raman: Mezclas binarias (GMP + IMP)

Las principales características, como valores promedio, mínimos y máximos de las concentraciones, para los conjuntos de calibración, validación y predicción de las mezclas binarias de GMP e IMP y ternarias de MSG, GMP e IMP, se resumen en la Tabla 4.1. En el análisis de mezclas binarias de GMP e IMP los mejores modelos PLS-Raman se obtuvieron usando el rango espectral completo, con un pre-tratamiento de los datos consistente en la primera derivada (FD), con un suavizado de 11 puntos, y el centrado en la media (MC), utilizando 5 LV's tanto para GMP como para IMP.

Tabla 4.1. Rangos de concentración de glutamato monosódico (MSG), guanosina-5'-monofosfato (GMP) e inosina-5'-monofosfato (IMP) en los conjuntos de calibración, validación y predicción para las formulaciones binarias y ternarias

Modelo	Conjunto	Muestras	Analito	Rango (%)	Promedio (%) ± SD
GMP + IMP	Calibración	19	GMP	0 - 99.77	49.9 ± 25.4
			IMP	0 - 99.5	49.8 ± 25.3
	Validación	9	GMP	7.6 - 89.5	49.6 ± 25.9
			IMP	10.2 - 91.9	50.0 ± 25.8
	Predicción	7	GMP	15.2 - 89.6	51.9 ± 24.3
			IMP	10.1 - 84.35	47.8 ± 24.2
MSG + GMP + IMP	Calibración	14	MSG	79.5 - 99.75	89.9 ± 6.5
			GMP	0 - 9.9	4.9 ± 3.2
			IMP	0 - 10.3	5.0 ± 3.3
	Validación	6	MSG	80.6 - 98.5	89.1 ± 7.2
			GMP	0.5 - 10.1	5.3 ± 3.7
			IMP	0.7 - 9.6	5.4 ± 3.6
	Predicción	6	MSG	79.6 - 98.7	88.0 ± 7.5
			GMP	0.5 - 10.1	5.8 ± 3.8
			IMP	0.5 - 1.0	5.9 ± 3.7

SD: Desviación estándar

En la Figura 4.2 puede verse la regresión entre los valores predichos por el modelo PLS-Raman y los valores de referencia tanto para GMP como para IMP. Los parámetros de los modelos de calibración y validación más importantes de los modelos seleccionados se resumen en la Tabla 4.2, en la que se puede observar que las capacidades predictivas de los modelos PLS-Raman fueron buenas tanto para la determinación de GMP e IMP, con altos valores de los coeficientes de determinación (R^2 Cal y R^2 Val) y valores bajos del error cuadrático medio del conjunto de validación (RMSEV).

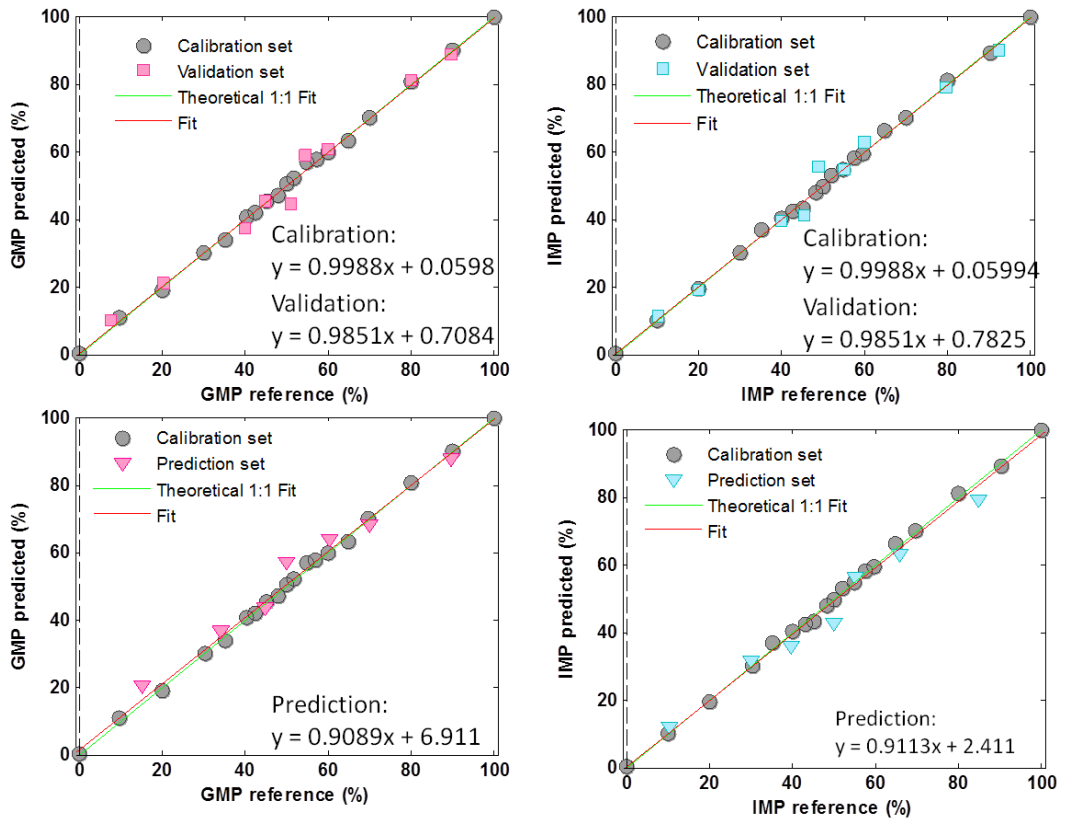


Figura 4.2. Valores de predicción vs valores de referencia para la determinación de GMP e IMP en mezclas GMP + IMP mediante PLS-Raman para muestras empleadas en los conjuntos de calibración, validación y predicción

Para evaluar la capacidad de predicción de los modelos PLS-Raman construidos, se utilizó un conjunto de predicción independiente, no empleado durante el paso de calibración para predecir GMP e IMP en las mezclas binarias. Se obtuvieron valores de R^2 Pred para GMP e IMP de 0.984, errores relativos de predicción (RRMSEP) de 7.6 y 7.9%, respectivamente, y valores de RPD de 6.1 para GMP y 6.5 para IMP (véase Tabla 4.2).

Tabla 4.2. Parámetros de los modelos PLS-Raman para la determinación de glutamato monosódico (MSG), guanosina-5'-monofosfato (GMP) e inosina-5'-monofosfato (IMP) en preparaciones binarias y ternarias

RAMAN		Conjunto de Calibración				Conjunto de Validación			Conjunto de Predicción				
Modelo	Analito	Pre-procesamiento	LV	RMSEC	R ² Cal	RMSECV	R ² CV	RMSEV	R ² Val	RMSEP	R ² Pred	RRMSEP (%)	RPD
GMP + IMP	GMP	FD, MC	5	0.85	0.999	5.63	0.948	2.98	0.985	3.96	0.984	7.6	6.1
	IMP	FD, MC	5	0.85	0.999	5.61	0.948	2.97	0.985	3.75	0.984	7.9	6.5
MSG + GMP + IMP	MSG	SNV, SD, MC	5	0.14	0.999	3.66	0.656	3.57	0.735	5.39	0.841	6.1	1.4
	GMP	SNV, SD, MC	5	0.07	0.999	1.77	0.670	1.85	0.737	2.72	0.838	46.6	1.4
	IMP	SNV, SD, MC	5	0.08	0.999	1.90	0.670	1.73	0.737	2.69	0.843	45.6	1.4

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEV: error cuadrático medio de validación; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Val: coeficiente de determinación de validación; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Tabla 4.3. Regresión entre los valores de referencia para la determinación de MSG, GMP e IMP en mezclas de aditivos umami mediante PLS-Raman

Raman	Regresión	
Modelo	Analito	Predicción
MSG + GMP + IMP	MSG	$y = 0.9995x + 0.0466$ $y = 0.689x + 28.74$ $y = 0.6533x + 34.95$
	GMP	$y = 0.9995x + 0.002237$ $y = 0.6627x + 1.161$ $y = 0.6372x - 0.1029$
	IMP	$y = 0.9994x + 0.002901$ $y = 0.7154x + 1.123$ $y = 0.6692x - 0.26$

3.4.3.1.2. Modelos PLS-Raman: Mezclas ternarias (MSG + GMP + IMP)

Las características de los conjuntos de calibración, validación y predicción de las mezclas ternarias de MSG + GMP + IMP se resumen en la Tabla 4.1. En la Tabla 4.2 se indican los parámetros de los mejores modelos de regresión de PLS-Raman obtenidos para MSG, GMP e IMP usando toda la región espectral, Variación Normal Estándar (SNV, del inglés *Standard Normal Variation*), segunda derivada (SD) con suavizado de 11 puntos, y MC como pre-tratamiento de datos, con un valor óptimo de 5 LV's. Los resultados obtenidos para la predicción del contenido de MSG en el conjunto de validación (ver Tabla 4.2) fueron mejores que los encontrados para la predicción de GMP e IMP, lo que resulta bastante evidente debido al hecho que el MSG siempre constituye el compuesto principal en aditivos umami, con una elevada concentración, mientras que GMP e IMP se emplean en pequeñas cantidades debido también a su coste más elevado. Por este motivo los modelos construidos para los tres compuestos considerados proporcionaron errores de predicción relativamente altos y valores bajos de RPD de 1.4, que evidencia una capacidad predictiva deficiente. La Tabla 4.3 muestra las ecuaciones de regresión entre los valores predichos mediante PLS-Raman para MSG, GMP e IMP y sus valores de referencia.

3.4.3.2. Espectroscopia NIR

Los espectros NIR se obtuvieron directamente por reflectancia difusa en unidades Kubelka-Munk, en la región comprendida entre 14000 y 4000 cm^{-1} , acumulando 50 barridos por espectro y usando una resolución espectral de 4 cm^{-1} . El *background* se adquirió de la esfera de integración cerrada bajo las mismas condiciones instrumentales que las empleadas para las mediciones de la muestra. Se obtuvieron tres medidas de cada muestra girando la posición del vial de muestra entre las repeticiones. Los valores promedios de los espectros triplicados de cada muestra se utilizaron para construir modelos de calibración PLS y para la predicción de las muestras.

En la Figura 4.3 pueden verse, en unidades Kubelka-Munk, los espectros NIR medios en la región entre 14000 y 4000 cm^{-1} sin pre-tratamiento, correspondientes a los patrones de MSG, GMP e IMP, mezclas binarias de GMP e IMP y mezclas ternarias de MSG+GMP+IMP. Las bandas más intensas aparecen en las regiones entre 4000-4750 cm^{-1} (bandas de combinación), 5600-5900 cm^{-1} (primer sobretono) y 8000-9000 cm^{-1} (segundo sobretono).

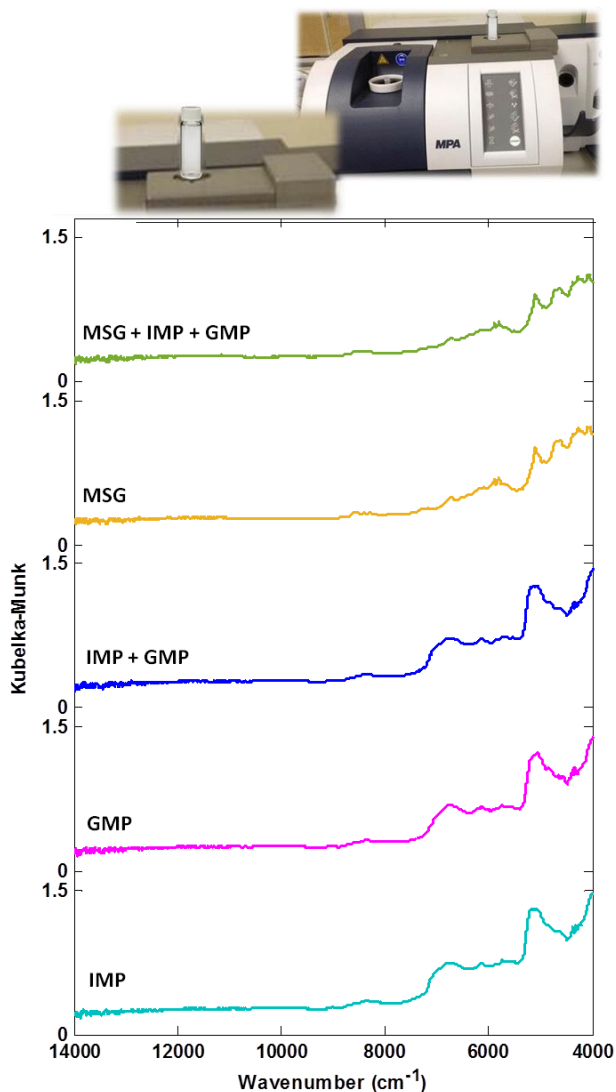


Figura 4.3. Espectros NIR del glutamato monosódico (MSG), guanosina 5'-monofosfato (GMP), inosina-5'-monofosfato (IMP) y mezclas binarias y ternarias de dichos compuestos en la región entre 14000 y 4000 cm^{-1} sin pre-tratamiento.

3.4.3.2.1. Modelos PLS-NIR Mezclas binarias (GMP + IMP)

Para comparar capacidad de predicción de los modelos PLS-Raman y PLS-NIR, se utilizaron los mismos conjuntos de calibración, validación y predicción en ambos casos (véase la Tabla 4.1). Diferentes pre-tratamientos de datos y distintas regiones espectrales se ensayaron para seleccionar los mejores modelos, encontrando los mejores resultados utilizando la región comprendida entre 10000 y 4000 cm^{-1} , y empleando FD (13 puntos) y MC como pre-procesamiento de la señal, con 5 LV's, sin la necesidad de eliminar ninguna muestra ya que no se identificaron valores anómalos.

La Tabla 4.4 muestra las ecuaciones de regresión entre los valores predichos mediante PLS-NIR para GMP e IMP y sus valores de referencia. Se obtuvieron elevados valores de R^2 Cal, R^2 CV y R^2 Val (ver Tabla 4.5). La capacidad de predicción de dichos modelos PLS-NIR se evaluó con el conjunto de predicción externo obteniendo buenos resultados en ambos casos, para GMP e IMP, con valores de RRMSEP de 3.7 y 4.0% para GMP e IMP, y RPD de 12.6 en ambos casos. Los parámetros más importantes de calibración, validación y predicción de los modelos PLS-NIR se resumen en la Tabla 4.5, demostrando una buena capacidad predictiva de los modelos PLS-NIR desarrollados para la determinación de GMP e IMP en mezclas binarias.

Tabla 4.4 Regresión entre los valores predichos y los valores de referencia para la determinación de glutamato monosódico (MSG), guanosina-5'-monofosfato (GMP) e inosina-5'-monofosfato (IMP) en mezclas binarias y ternarias de aditivos de sabor umami mediante PLS-NIR

NIR		Regresión		
Modelo	Analito	Calibración	Validación	Predicción
GMP + IMP	GMP	$y = 0.9999x + 0.006451$	$y = 0.9775x + 0.7425$	$y = 1.008x - 1.173$
	IMP	$y = 0.9999x + 0.006434$	$y = 0.9775x + 1.495$	$y = 1.008x + 0.3334$
MSG + GMP + IMP	MSG	$y = 0.9928x + 0.6466$	$y = 1.007x - 1.032$	$y = 0.9789x + 1.383$
	GMP	$y = 0.9918x + 0.0402$	$y = 0.9803x + 0.2616$	$y = 0.9604x + 0.4301$
	IMP	$y = 0.9922x + 0.03889$	$y = 1.034x + 0.05922$	$y = 0.9972x + 0.2898$

3.4.3.2.2. Modelos PLS-NIR: Mezclas ternarias (MSG + GMP + IMP)

Al igual que en el caso anterior, los conjuntos de calibración, validación y predicción empleados para construir los modelos PLS-Raman se utilizaron también para los modelos PLS-NIR (ver Tabla 4.1). Los mejores modelos PLS-NIR para la predicción de MSG, GMP e IMP se obtuvieron utilizando FD (13 puntos) y MC como pre-procesamiento de los datos espectrales, en la región entre 10000 y 4000 cm^{-1} y empleando 3 LVs. En la Tabla 4.4 se pueden ver las rectas de regresión entre los datos predichos y los valores de referencia, mientras que los principales parámetros de los mejores modelos PLS-NIR se incluyen en la Tabla 4.5. En todos los casos se obtuvieron altos valores de R^2 y bajos valores de RMSEV.

Tabla 4.5. Parámetros de los modelos PLS-NIR para la determinación de glutamato monosódico (MSG), guanosina-5'-monofosfato (GMP) e inosina-5'-monofosfato (IMP) en preparaciones binarias y ternarias

NIR		Conjunto de Calibración				Conjunto de Validación		Conjunto de Predicción					
Modelo	Analito	Pre-procesamiento	LV	RMSEC	R ² Cal	RMSECV	R ² CV	RMSEV	R ² Val	RMSEP	R ² Pred	RRMSEP (%)	RPD
GMP + IMP	GMP	FD, MC	5	0.28	0.999	2.97	0.986	1.23	0.998	1.93	0.994	3.7	12.6
	IMP	FD, MC	5	0.28	0.999	2.96	0.986	1.22	0.998	1.92	0.994	4.0	12.6
MSG + GMP + IMP	MSG	FD, MC	3	0.53	0.992	1.14	0.967	0.62	0.995	0.61	0.997	0.7	12.2
	GMP	FD, MC	3	0.27	0.992	0.58	0.964	0.28	0.995	0.29	0.998	4.9	13.0
	IMP	FD, MC	3	0.28	0.992	0.58	0.966	0.37	0.995	0.35	0.996	5.8	10.8

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEV: error cuadrático medio de validación; RRMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Val: coeficiente de determinación de validación; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

El conjunto de predicción independiente se utilizó para evaluar la capacidad predictiva de los modelos PLS-NIR para MSG, GMP e IMP en mezclas ternarias y esta información también se incluye en las Tablas 4.4 y 4.5. Los valores de RRMSEP encontrados fueron de 0.7, 5.0 y 5.9% con RPD de 12.2, 13.0 y 10.8, para MSG, GMP e IMP, respectivamente. Esto indica que la técnica PLS-NIR resulta adecuada para el análisis directo de mezclas de estos tres potenciadores de sabor.

3.4.3.3. Comparación entre el Raman y NIR para el análisis de mezclas umami

Los resultados obtenidos mediante el uso de modelos PLS para determinar MSG, GMP e IMP en aditivos umami utilizando espectros Raman y NIR se compararon con el fin de evaluar la mejor alternativa.

Para las mezclas GMP e IMP, las técnicas Raman y NIR proporcionaron buenos resultados con valores máximos de RRMSEP de 7.9 y 4.0%, y valores RPD de 6.5 y 12.6, respectivamente. Sin embargo, en el caso de mezclas ternarias, se encontraron valores de RRMSEP altos y valores bajos de RPD usando PLS-Raman, mientras que los modelos PLS-NIR proporcionaron valores de RPD para MSG, GMP e IMP de 12.2, 13.0 y 10.8, respectivamente, siendo una metodologías con una buena capacidad predictivas [23].

3.4.3.4. Evaluación de los métodos

Para evaluar el carácter verde de la metodología propuesta, se utilizó la escala ecológica propuesta por Galuszka *et al.* [20] y modificada por Armenta *et al.* [21]. La Figura 4.4 muestra los resultados de acuerdo con el Certificado Verde para evaluar la metodología propuesta FT-NIR y el método LC-UV utilizado por Chiang *et al.* [13] como metodología de referencia. Se asignaron cero puntos de penalización a la determinación simultánea de MSG, GMP e IMP utilizando el procedimiento FT-NIR propuesto en la escala ecológica debido a la ausencia de consumo de reactivos, generación de residuos y emisiones nocivas.

Con respecto al método de referencia no es posible realizar la determinación simultánea de MSG, GMP e IMP en un solo ensayo ya que se requiere trabajar a niveles de dilución diferentes y, por ello, se realizaron dos evaluaciones diferentes; la primera para evaluar la determinación de MSG y la segundo para la determinación de simultánea de GMP e IMP. Para la determinación de MSG por LC-UV, se asignaron 25.7 puntos de penalización debido al uso de reactivos (7.9), energía consumida (1) y residuos generados (16.8), obteniéndose una puntuación total de 74.3. Para la determinación de GMP e IMP por LC-UV, la puntuación total

fue 86 debido a que se asignaron 14 puntos de penalización (ver detalles en la tabla de la Figura 4.4). Por lo tanto, para la determinación de MSG, GMP e IMP por PLS-NIR se proporciona un método de Certificado Verde de categoría A con puntuación de 100 puntos. Los métodos LC-UV de referencia utilizados para la determinación de MSG corresponden a la categoría C y a la categoría B los que corresponden a la determinación de GMP e IMP. Por lo tanto, se puede concluir que el método propuesto es una alternativa sostenible y respetuosa con el medio ambiente a los métodos disponibles en la literatura para el control de calidad de aditivos umami, pudiéndose determinar los 3 componentes de manera simultánea.

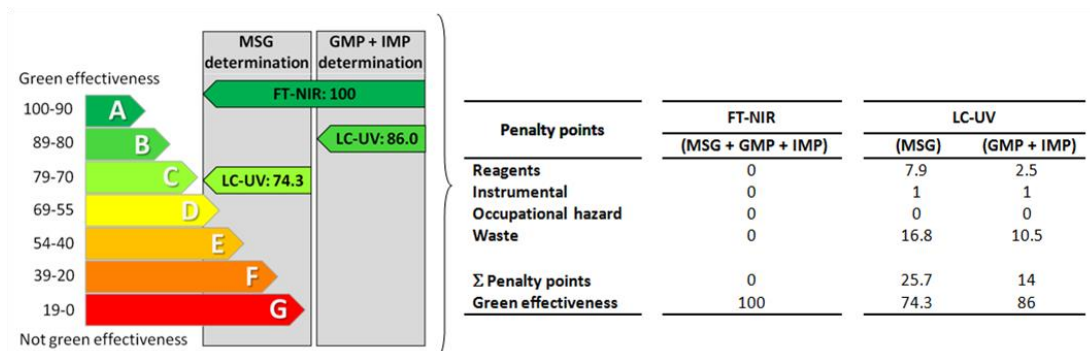


Figura 4.4. Evaluación, puntos de penalización por muestra y etiqueta de Certificado Verde correspondiente a la metodología propuesta por NIR en comparación con los procedimientos de referencia por LC-UV.

3.4.4. Conclusiones Capítulo 4

Las técnicas de espectroscopia Raman y NIR proporcionan alternativas rápidas y ecológicas para determinar simultáneamente MSG, GMP e IMP en aditivos umami. Para las mezclas binarias compuestas por GMP e IMP, ambas técnicas combinadas con PLS ofrecen procedimientos precisos y exactos para el análisis cuantitativo de estas muestras. Sin embargo, para las mezclas ternarias compuestas por MSG, GMP e IMP, los resultados obtenidos mediante la espectroscopia NIR fueron mejores que los encontrados por Raman. En resumen, estos estudios ponen de manifiesto que para el análisis de aditivos umami en mezclas de potenciadores de sabor, PLS-NIR ofrece una herramienta precisa y exacta para la determinación de MSG, GMP e IMP, siendo una clara alternativa verde frente a los métodos tradicionales basados en determinaciones cromatográficas.

En resumen, la espectroscopia NIR permite obtener resultados prácticamente en tiempo real para mediciones in situ y en línea, sin pre-tratamiento de muestras, siendo un buen método alternativo para la inspección rápida de formulaciones de aditivos alimentarios para asegurar la calidad de las mezclas umami que se comercializan para su uso en la industria agroalimentaria y en el procesado de alimentos elaborados.

3.4.5. Referencias Capítulo 4

- [1] K. Ikeda, On a new seasoning, *Journal of Tokyo Chemistry Society*, 30 (1909) 820–836.
- [2] A. Kuninaka, Studies on taste of ribonucleic acid derivatives. *J. Agric. Chem. Soc. Jpn.*, 34 (1960) 487–492.
- [3] A. Kuninaka, The Nucleotides, a Rationale of Research on Flavor Potentiation. *Symposium on Flavor Potentiation*, (1964) 4–9. Arthur D. Little, Cambridge, MA.
- [4] CODEX Alimentarius: General Standard for Food Additives Codex Stan 192-1995.
- [5] Joint FAO/WHO Expert Committee on Food Additives (JECFA), 1988.
- [6] Scientific Committee for Food (SCF), Reports of the Scientific Committee for Food on a First Series of Food Additives of Various Technological Functions, Commission of the European Communities, 25th Series. Brussels, Belgium (1991).
- [7] N.J. Arruda, J.L. Filho, M.C. Montenegro, A.N. Araújo, V.L. Silva, Simple and Inexpensive Flow L-Glutamate Determination Using Pumpkin Tissue, *J. Agric. Food Chem.*, 51(2003) 6945–6948.
- [8] *Textbook of Biochemistry with Clinical Correlations*, T. M. Devlin (fourth ed.) Wiley-Liss, New York (1997) 489–523. [Online]. Available: <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-EHEP000229.html>.
- [9] P.D. Chiang, C.T. Yen, J.L. Mau, Non-volatile taste components of various broth cubes, *Food Chem.*, 101(2007) 932–937.
- [10] B.E. Demirhan, B. Demirhan, C. Sönmez, H. Torul, U. Tamer, G. Yentür, Monosodium glutamate in chicken and beef stock cubes using high-performance liquid chromatography, *Food Addit. Contam. Part B*, 8(2015) 63–66.
- [11] High-Performance Liquid Chromatography for Joint Determination of the Flavour Enhancers Monosodium Glutamate, Disodium Guanylate and Disodium Inosinate in Dehydrated Bouillons, *Dtsch. Lebensm.-Rundsch. DLR Anal. Forsch. Prozesse Recht*, 98(2002).
- [12] C.C. Acebal, A.G. Lista, B.S. Fernández, Simultaneous determination of flavor enhancers in stock cube samples by using spectrophotometric data and multivariate calibration, *Food Chem.*, 106(2008) 811–815.

- [13] I. Durán-Merás, A. Muñoz de la Peña, A. Espinosa-Mansilla, F. Salinas, Multicomponent determination of flavour enhancers in food preparations by partial least squares and principal component regression modelling of spectrophotometric data, *Analyst*, 118(1993) 807–813.
- [14] Y. Yang, Q. Chen, C. Shen, S. Zhang, Z. Gan, R. Hu, J. Zhao, Y. Ni, Evaluation of monosodium glutamate, disodium inosinate and guanylate umami taste by an electronic tongue, *J. Food Eng.*, 116(2013) 627–632.
- [15] K. Beullens, P. Mészáros, S. Vermeir, D. Kirsanov, A. Legin, S. Buysens, N. Cap, B. M. Nicolai, J. Lammertyn, Analysis of tomato taste using two types of electronic tongues, *Sens. Actuators B Chem.*, 131(2008) 10–17.
- [16] L. Bagnasco, M. E. Cosulich, G. Speranza, L. Medini, P. Oliveri, S. Lanteri, Application of a voltammetric electronic tongue and near infrared spectroscopy for a rapid umami taste assessment, *Food Chem.*, 157(2014) 421–428.
- [17] D. Lin-Vien, N.B. Colthup, W.G. Fateley, J.G. Grasselli H.B. Jovanovich (Ed.), *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules*, Academic Press Inc., San Diego (1991), p. 140.
- [18] N. Peica, C. Lehene, N. Leopold, S. Schlücker, W. Kiefer, Monosodium glutamate in its anhydrous and monohydrate form: Differentiation by Raman spectroscopies and density functional calculations, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, 66(2007) 604–615.
- [19] Chapter 8: Implementation of Near-Infrared Technology in Near-Infrared Technology in the Agricultural and Food Industry, Second Edition. P. C. Williams, Phil Williams and Karl Norris eds, 2004, pages 145-169.
- [20] A. Galuszka, Z.M. Migaszewski, P. Konieczka, J. Namiesnik, Analytical eco-scale for assessing the greenness of analytical procedures. *Trends Anal. Chem.* 37(2012) 61–72.
- [21] S. Armenta, M. de la Guardia, J. Namiesnik, Green Microextraction. M. Valcárcel, S. Cárdenas, R. Lucenas, eds. *Analytical Microextraction Techniques*. Bentham Science, 2017.

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN

PARTE 2: Muestras Ambientales y de Interés Agrícola

Parte 2. Muestras Ambientales y de Interés Agrícola

La espectroscopia infrarroja presenta grandes ventajas debido a su gran versatilidad y posibilidad de realizar medidas *in situ* y en tiempo real. Es por ello, que se trata de una técnica adecuada para el análisis de muestras ambientales, ya sea, para el control y evaluación de las características de los compost, desde el inicio del compostaje hasta el final del mismo, análisis de suelos contaminados, así como medidas en fase gas.

Mediante el uso de la espectroscopia vibracional es posible el seguimiento de la maduración de los compost en el procedimiento del compostaje, para que puedan ser utilizados en la agricultura como fertilizantes y enmendantes del suelo, contribuyendo a mejorar las características físico-químicas y biológicas del mismo. Al mismo tiempo, la espectroscopia permite la determinación cuantitativa de los compuestos fenólicos, que están estrechamente relacionados con la calidad del compost, así como la determinación de carbono orgánico (OC) y de nitrógeno total (TN) que también se consideran parámetros químicos para evaluar las posibilidades de éxito del proceso de compostaje.

El control de los suelos contaminados puede considerarse en la actualidad como uno de los mayores problemas-retos que nos podemos encontrar desde el punto de vista medio ambiental. La espectroscopia infrarroja con generación de fase vapor, es capaz de proporcionar una herramienta adecuada para la determinación simultánea de compuestos BTEX en suelos contaminados de forma rápida y sin ningún pre-tratamiento químico de la muestra, reduciendo el tiempo de análisis de las muestras.

Por otro lado, en los espectros IR es posible encontrarse con la presencia de interferencias causadas por gases y vapores que estén presentes en la atmósfera de medida. Dichas interferencias perjudican la calidad de las medidas, especialmente cuando dichos gases y/o vapores ambientales presentan bandas de absorción en la zona espectral de interés, produciéndose un solapamiento de las señales, siendo necesaria la corrección de los espectros mediante métodos de corrección. Es por ello que se necesita un método flexible y automatizado para corregir y evitar las interferencias espectrales de vapores orgánicos presentes en los espectros FTIR medidos.

3.5. CAPÍTULO 5:
Evaluación de compost para su posterior uso en la
agricultura mediante espectroscopia infrarroja NIR y
MIR

3.5. Capítulo 5. Evaluación de compost para su posterior uso en la agricultura mediante espectroscopia infrarroja NIR y MIR

3.5.1. Introducción

Los compost se caracterizan por ser materiales de naturaleza orgánica procedentes de residuos de origen animal y/o vegetal, obtenidos por fermentación aeróbica controlada. Transcurrido un cierto tiempo de estabilización de estos residuos, suficiente para considerarlos como maduros, pueden utilizarse en agricultura como fertilizantes y enmendantes del suelo, contribuyendo a mejorar las características físico-químicas y biológicas del mismo. Por tanto, el compostaje es una de las soluciones sostenibles para el manejo y gestión de los desechos sólidos orgánicos, siendo una técnica de biorremediación que ofrece las ventajas de la degradación de los residuos orgánicos junto con la producción de fertilizantes adecuados para la modificación del suelo [1], siendo en su conjunto un proceso biooxidativo controlado, que se desarrolla sobre sustratos orgánicos heterogéneos en estado sólido, debido a la actividad secuencial de una gran diversidad de microorganismos.

Según las características del compost, puede emplearse como enmendante orgánico, que actuaría principalmente sobre las propiedades físico-químicas y biológicas del suelo, o como abono, mejorando además la nutrición mineral de las plantas y ahorrando el uso de fertilizantes minerales.

Existen estudios que relacionan el nivel de compuestos fenólicos totales (TPC) en un compost con la calidad del mismo, ya que niveles bajos estos compuestos indican alta estabilidad del producto, mientras que la alta concentración de fenoles evidencia su inestabilidad [1]. Palm *et al.* [2] consideran la concentración de los compuestos fenólicos como un buen parámetro para abordar la modificación de la materia orgánica en la agricultura mediante el empleo de composts. El proceso de maduración del compost reduce el volumen del producto final pero a pesar de ello disminuye gradualmente la concentración de compuestos fenólicos [1].

Desde un punto de vista práctico, el compost constituye una fuente importante de carbono, fósforo orgánico, nitrógeno, y también aumenta la capacidad de intercambio catiónico del suelo [3], por lo que los contenidos de carbono orgánico (OC) y de nitrógeno total (TN) también se consideran parámetros químicos para evaluar las posibilidades de éxito del proceso de compostaje [4]. De hecho, como se ha indicado, la calidad del compost para su aplicación en el suelo depende de su madurez y estabilidad [3], y éstas se suelen evaluar teniendo en cuenta la relación OC/TN ya que se utiliza en el control de la madurez final del producto obtenido [5]

[6]. Sin embargo, tanto para la determinación de los compuestos fenólicos como para la determinación de OC y TN en suelos, composts y residuos orgánicos es necesario una gran cantidad de reactivos y de tiempo para el análisis, junto con mano de obra calificada, ya que se utilizan métodos químicos clásicos como: i) la valoración, en el caso de TPC, que incluye una etapa previa de extracción; ii) mineralización a elevadas temperaturas y valoración para la determinación de OC y, iii) destilación después de una mineralización previa con ácido sulfúrico, para la determinación del TN [7-9].

La espectroscopia infrarroja tanto en la región del infrarrojo cercano (NIR) como en la del medio (MIR), puede ofrecer una alternativa no destructiva, rápida y relativamente barata a los métodos tradicionales para el control de este tipo de parámetros en muestras de diferente naturaleza. Existen algunos estudios donde se utiliza la espectroscopia infrarroja tanto NIR como MIR para la determinación de compuestos fenólicos en alimentos [10-13], y OC y TN en plantas [14], suelos [15, 16] y composts [3, 17]. En nuestro caso se pretende realizar un estudio comparativo sobre las posibilidades del empleo de la espectroscopia NIR y MIR para en análisis y control de los composts, se modo que estas técnicas puedan ser planteadas como alternativas a los métodos tradicionales. Como parámetros de referencia se van a considerar los TPC, el OC y el TN.

En este sentido, el presente capítulo se encuentra subdividido en 2 partes diferenciadas en función de los analitos estudiados:

- **5.A.** Determinación de compuestos fenólicos totales.
- **5.B.** Predicción del contenido de carbono orgánico y de nitrógeno total.

Teniendo en cuenta lo comentado, los objetivos fijados en esta parte del trabajo son:

- Desarrollo de un método rápido, simple y no destructivo para la determinación directa de TPC en muestras de compostajes mediante el uso de los modelos de regresión PLS a partir de los espectros NIR y MIR, tomando como valores de referencia los obtenidos a partir del método basado en la utilización del reactivo de Folin-Ciocalteu.
- Encontrar una metodología de cribado para la determinación de OC y TN en muestras orgánicas como compostajes, desechos biogénicos y mezclas de residuos de compostajes utilizando las metodologías MIR y NIR combinadas con la regresión PLS.

3.5.2. Parte experimental

3.5.2.1. Instrumentos

Los espectros del infrarrojo cercano registrados en este trabajo fueron obtenidos con un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Germany) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS. El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

Los espectros del infrarrojo medio fueron registrados con un espectrofotómetro modelo Tensor 27 FT-MIR de Bruker (Bremen, Alemania) equipado con un detector de sulfato de trigliceno dopado con L-alanina deuterado (DLaTGS). El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

3.5.A. Capítulo 5.A. Determinación de compuestos fenólicos totales en compostajes mediante NIR y MIR

3.5.A.1. Muestras

Este estudio se ha realizado utilizando 95 muestras de compost, obtenidas a partir de: (i) lodos de alcantarillado, (ii) residuos sólidos municipales y (iii) diversos residuos biológicos vegetales, tales como serrín, paja, caña de Provençe, caña de azúcar, corteza de palma, estiércol de caballo, etc. Todas las muestras de compost empleadas se obtuvieron mediante fermentación aeróbica (compostaje) utilizando experimentos a escala de laboratorio, pudiéndose clasificar en los siguientes tipos: (i) compost mineral maduro ($C_m=9$ muestras), (ii) compost orgánico maduro ($C_o=68$ muestras), y (iii) compost vegetal inmaduro ($C_v=18$ muestras), de modo que todos ellos representarían diferentes etapas del proceso de compostaje. Las muestras fueron secadas a $105\text{ }^\circ\text{C}$ durante 12 h y almacenadas en todo momento en viales de vidrio con tapón roscado hasta el momento de su análisis.

3.5.A.2. Método de referencia

La determinación de TPC se realizó de acuerdo con el método colorimétrico Folin-Ciocalteu siguiendo el procedimiento que se recoge en la literatura [18]. Para ello, se prepararon extractos de todas las muestras obtenidos a partir de 1 g de muestra seca tratado mediante agitación magnética durante 30 min con 10 ml de metanol puro. Los extractos se mantuvieron en contacto con el compost durante 24 h a $4\text{ }^\circ\text{C}$ y se separaron por centrifugación utilizando una centrífuga Digicen 20 de Ortoalresa (Madrid, España), que trabajaba a 3000 rpm durante 10 min. Una vez separados los sobrenadantes se almacenaron en tubos de plástico a $4\text{ }^\circ\text{C}$ hasta su análisis. Para el análisis, se añadió un volumen de extracto de cada muestra adecuadamente diluido, a 0.5 mL de agua destilada y 0.125 mL de reactivo de Folin-Ciocalteu. La mezcla se agitó y se dejó en reposo durante 6 min, antes de la adición de 1.25 mL de Na_2CO_3 al 7%. Después la solución se ajustó con agua destilada hasta un volumen final de 3 mL y se homogenizó. La absorbancia se midió a 760 nm frente a un blanco preparado de la misma manera que las muestras, después de mantenerlas en un armario y cubiertas con papel de aluminio en la oscuridad. El contenido fenólico total de todas las muestras se expresó como mg de equivalente de ácido gálico por gramo de masa seca ($\text{mg GAE g}^{-1}\text{ DW}$), usando una curva de calibración establecida a partir de disoluciones patrón de ácido gálico. Todas las mediciones se realizaron por triplicado.

La concentración mínima encontrada de TPC fue 0.07 mg GAE g⁻¹ DW mientras que 4.83 fue el valor máximo medido. Las variaciones en el contenido de TPC se deben a la naturaleza de los principales ingredientes que componen los diferentes tipos de composts, como son sustancias minerales, vegetales o residuos biológicos. Los valores más bajos de TPC estuvieron presentes en los composts obtenidos a partir de compuestos minerales con un bajo porcentaje de materia orgánica (menor del 30% m/m); por el contrario, se encontraron altas concentraciones de TPC en los composts procedentes de residuos biogénicos (con un contenido de materia orgánica superior al 70% m/m). En resumen, los TPC encontrados en los composts de naturaleza mineral variaban entre 0.07 y 0.15 mg GAE g⁻¹ DW, mientras que los presentes en composts orgánicos y composts vegetales se encontraban entre 0.10-0.80 y 0.48-4.83 mg GAE g⁻¹ DW, respectivamente.

3.5.A.3. Resultados y discusión

3.5.A.3.1. Espectroscopia NIR

Las muestras se colocaron en viales de vidrio transparente de 22 mm de diámetro interno y 44 mm de altura para obtener directamente espectros NIR por reflectancia difusa, en unidades de Kubelka-Munk. Todos los espectros NIR se adquirieron en la región espectral comprendida entre 14000 y 4000 cm⁻¹, acumulando 50 barridos por espectro y utilizando una resolución óptica de 4 cm⁻¹. El *background* se adquirió a partir de la esfera de integración cerrada usando las mismas condiciones instrumentales que las empleadas para la medida de las muestras. Se obtuvieron tres medidas de cada muestra girando aproximadamente 120° la posición del vial entre las réplicas, para asegurar una mejor representatividad de los espectros. Los valores promedios de los espectros triplicados de cada muestra se emplearon para construir modelos de calibración PLS y predicción.

En la Figura 5.1 se representan los espectros de reflectancia difusa NIR originales de las muestras de compost utilizadas en este estudio en la región entre 14000 y 4000 cm⁻¹, diferenciadas en las tres clases: Cm, Co y Cv. Las principales bandas que pueden verse en estos los espectros de los composts se encuentran alrededor de 8300, 6850 y 5160 cm⁻¹ y están relacionadas con los grupos O-H de los grupos fenol [19,20], obteniéndose una mayor intensidad de la señal de estas bandas en los espectros correspondientes a las muestras de la clase Cv, seguidas por las clases Co y Cm. Además, en los espectros pueden distinguirse también, pero con menos intensidad, las bandas situadas aproximadamente a 5800 y 4670 cm⁻¹, que son debidas a los armónicos de la vibración de estiramiento de los enlaces C-H, y

que normalmente son atribuidas a la presencia de fenoles, lípidos, aceites y grupos aromáticos en plantas [21].

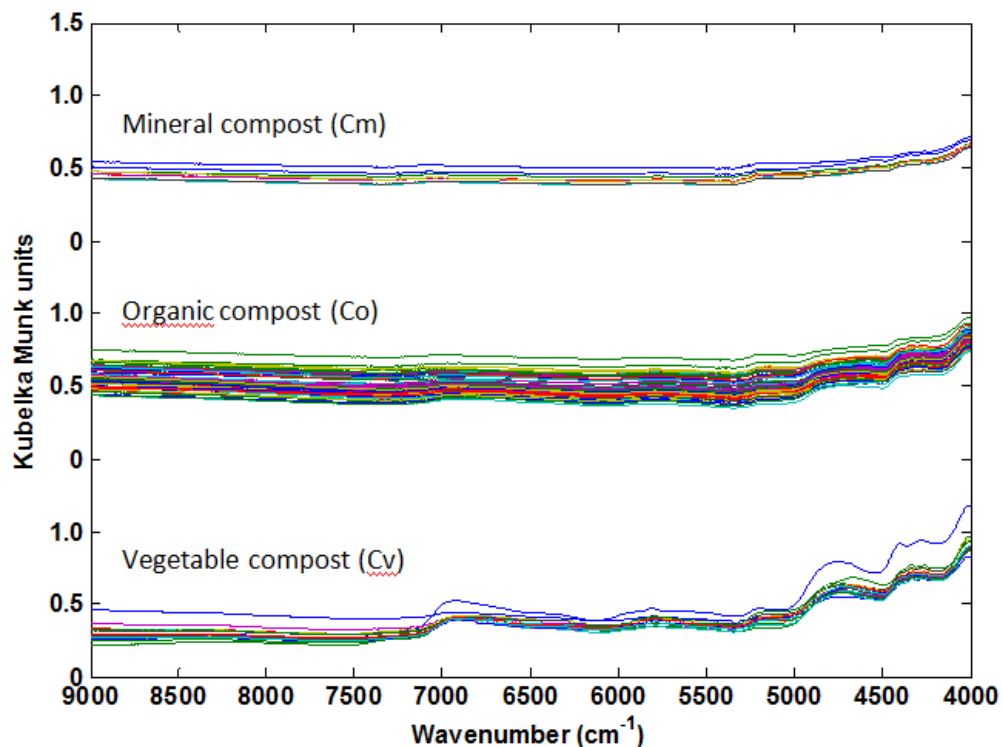


Figura 5.1. Espectros NIR de las muestras de compost en el intervalo entre 9000 y 4000 cm^{-1}
 Nota: Los espectros correspondientes a cada uno de los grupos de compost se desplazaron sobre el eje de la absorbancia para evitar la superposición espectral.

3.5.A.3.1.1. Análisis exploratorio: PCA

Con el fin de ver si es posible diferenciar los diferentes tipos de compost en función de los espectros NIR de las muestras, se realizó un análisis de componentes principales (PCA) de los mismos. Para ello se utilizó un pretratamiento de datos basado MC y SNV, considerando la región espectral comprendida entre 9000 y 4000 cm^{-1} . Los dos primeros componentes principales calculados explican una varianza total de 95.29%. En la Figura 5.2 puede verse el gráfico de *scores* para PC2 (8.7 % de varianza explicada) frente a PC1 (86.59% de la varianza). En este gráfico se pueden observar las muestras de composts etiquetadas según sea su origen mineral, orgánico o vegetal, apreciándose una separación de muestras en dos grupos a lo largo del PC1, encontrándose la mayoría de las del tipo Cv situadas hacia la izquierda del PC1 y el resto de composts situados más hacia la derecha, de acuerdo con la concentración de TPC.

Teniendo en cuenta el PC2, existe una pequeña diferencia entre los espectros NIR de las muestras de los grupos Cm y Co, que el PC1 no era capaz de separar.

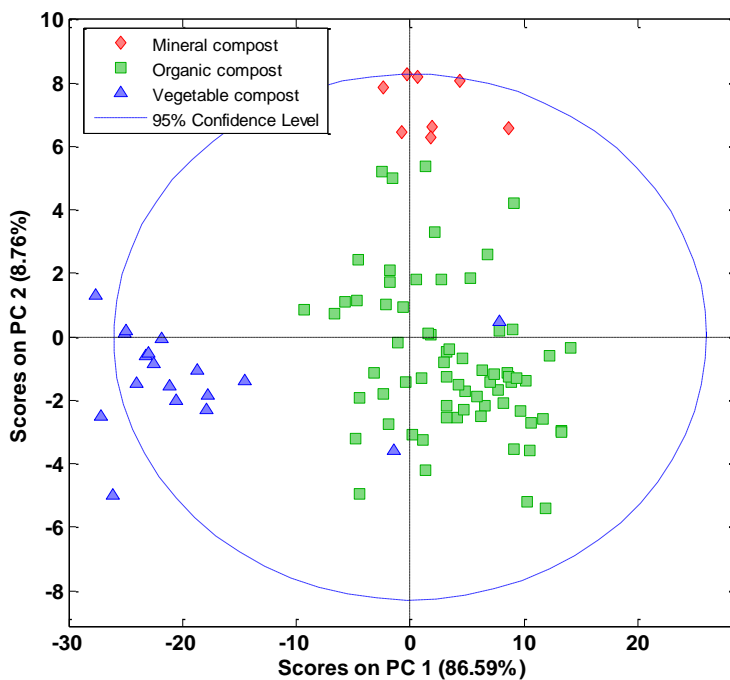


Figura 5.2. Resultados del análisis del componente principal basado en espectros DR-NIR de las muestras etiquetadas como Cm, Co y Cv

3.5.A.3.1.2. Modelos PLS-NIR

Con el fin de construir y evaluar los modelos PLS-NIR para la determinación del contenido de TPC en las muestras de compost, la población de muestras se dividió, mediante el algoritmo de K-S, en un conjunto de calibración y un segundo de validación, con 76 y 19 muestras en cada uno de ellos. Debido a que el conjunto de calibración debe ser representativo y que el contenido de TPC debe incluir a las muestras de validación se representó el gráfico de PCA pero diferenciando las muestras según pertenezcan al conjunto de calibración o de validación (véase Figura 5.3), pudiéndose observar con claridad que la distribución de muestras del conjunto de calibración contiene también a las de validación. Además, los contenidos de TPC de los conjuntos de calibración y validación se encontraron en el intervalo entre 0.07 y 4.83 mg GAE g⁻¹ DW y entre 0.11 y 4.48 mg GAE g⁻¹ DW, respectivamente.

Con el fin de determinar el mejor modelo PLS, se ensayaron y compararon diferentes regiones espectrales con distintos pre-tratamientos de los datos,

utilizando las 76 muestras seleccionadas por el algoritmo de K-S como conjunto de calibración para construir el modelo y teniendo en cuenta el menor error cuadrático medio de la validación cruzada *leave one out* para la selección de las variables latentes.

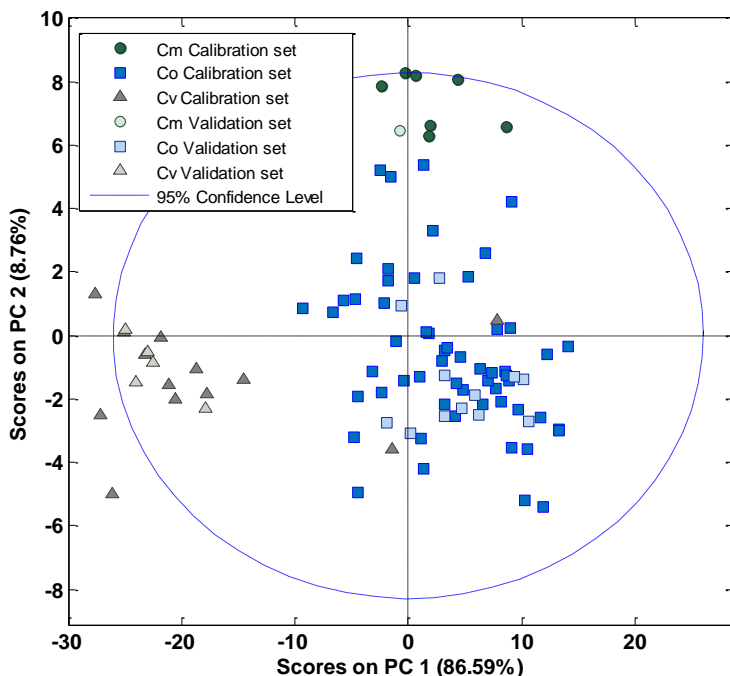


Figura 5.3. Resultados del análisis del componente principal basado en espectros DR-NIR de las muestras etiquetadas como Cm, Co y Cv y diferenciadas en los conjuntos de calibración y validación.

Una vez seleccionado el modelo, se evaluó su capacidad predictiva mediante el uso del set externo de validación, formado por 19 muestras. En la Tabla 5.1 pueden verse las características y los parámetros de los diferentes modelos PLS estudiados. Los mejores resultados se obtuvieron utilizando la región entre 9000 y 4000 cm^{-1} con los pre-procesamientos de MC y SNV y 7LV's, con valores de RMSEC de 0.16 $\text{mg GAE g}^{-1} \text{DW}$ y $R^2 \text{ Cal}$ de 0.97, RMSECV $\text{mg GAE g}^{-1} \text{DW}$ de 0.29 con $R^2 \text{ CV}$ de 0.91 y un RMSEP de 0.19 $\text{mg GAE g}^{-1} \text{DW}$ con $R^2 \text{ Pred}$ de 0.99 y RPD de 8.26.

El modelo construido con MC y primera derivada en la región comprendida entre 6100 y 5385 cm^{-1} , con 7LV's, presenta valores similares a los obtenidos con el pre-procesado y región espectral indicados anteriormente, con valores de RMSEC de 0.11 $\text{mg GAE g}^{-1} \text{DW}$ y $R^2 \text{ cal}$ de 0.98, RMSECV $\text{mg GAE g}^{-1} \text{DW}$ de 0.31 con $R^2 \text{ CV}$ de 0.90 y un RMSEP de 0.19 $\text{mg GAE g}^{-1} \text{DW}$ con $R^2 \text{ Pred}$ de 0.99 y RPD de 8.22.

Tabla 5.1. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-NIR para la determinación de TPC en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración			Conjunto de Validación			
			RMSEC (mg GAE g ⁻¹ DW)	R ² Cal	RMSECV (mg GAE g ⁻¹ DW)	R ² CV	RMSEP (mg GAE g ⁻¹ DW)	R ² pred	RPD
MC, SNV	9000-4000	7	0.16	0.97	0.29	0.91	0.19	0.99	8.26
MC, FD	9000-4000	2	0.27	0.91	0.31	0.88	0.38	0.98	4.2
MC, MSC, FD	9000-4000	6	0.16	0.97	0.22	0.94	0.22	0.98	7.34
MC, SNV	6100-5385	4	0.27	0.91	0.40	0.81	0.24	0.99	6.73
MC, FD	6100-5385	7	0.11	0.98	0.31	0.90	0.19	0.99	8.22
MC, MSC, FD	6100-5385	5	0.32	0.87	0.40	0.80	0.26	0.98	6.21

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

En resumen, los resultados obtenidos a partir de los modelos PLS-NIR indican que el procedimiento desarrollado para la determinación del contenido de TPC en composts es adecuado para fines analíticos con valores RPD de 8.2.

3.5.A.3.2. Espectroscopia MIR

Aproximadamente 100 mg de muestra, secada a 105 °C durante 12 h, se depositó sobre el accesorio de ATR para sólidos y se prensó, para registrar los espectros se en el intervalo comprendido entre 4000 y 550 cm^{-1} , acumulando 50 barridos por espectro y a una resolución de 4 cm^{-1} . Se utilizó como *background* un espectro del cristal ATR limpio, obtenido inmediatamente antes de adquirir el espectro de cada muestra. El cristal ATR se limpió mediante el uso de un papel de celulosa seco. Los espectros promedios de las medidas realizadas por triplicado, para cada una de las muestras se utilizaron para construir los modelos quimiométricos.

La Figura 5.4 muestra los espectros ATR-MIR de todas las muestras, en la región entre 4000 y 550 cm^{-1} , y clasificadas en las tres clases de compost: Cm, Co y Cv. La región comprendida entre 2240 y 1880 cm^{-1} no se consideró para tratamiento numérico de los datos debido al elevado nivel de ruido como consecuencia de la alta absorbancia del cristal ATR. Las bandas intensas detectadas en la región de 3500-3100 cm^{-1} son debidas al grupo O-H del agua [22, 23] y al estiramiento del enlace N-H [24]. Las bandas encontradas en la región entre 2950 y 2850 cm^{-1} están relacionadas con el contenido de materia orgánica y las bandas a 2920 y 2850 cm^{-1} corresponden al estiramiento asimétrico y simétrico del grupo metileno, respectivamente. Las bandas más intensas en los espectros se detectaron entre 1760 y 800 cm^{-1} , que incluye información de los compuestos fenólicos, tales como la banda de estiramiento de los grupos carbonilo C=O en la región 1712-1704 cm^{-1} y las bandas de estiramiento del C=C a 1612-1608 y 1519-1516 cm^{-1} [11]. Las bandas centradas en 1635 cm^{-1} pueden ser debidas a grupos O-H de los polifenoles [25]. La absorbancia a 1200-1050 cm^{-1} se ha asignado a compuestos fenólicos y la banda a 1385 cm^{-1} se ha atribuido al grupo COOH [22]. La banda centrada a 875 cm^{-1} presentes en los composts minerales puede usarse como un indicador de mineralización avanzada [23].

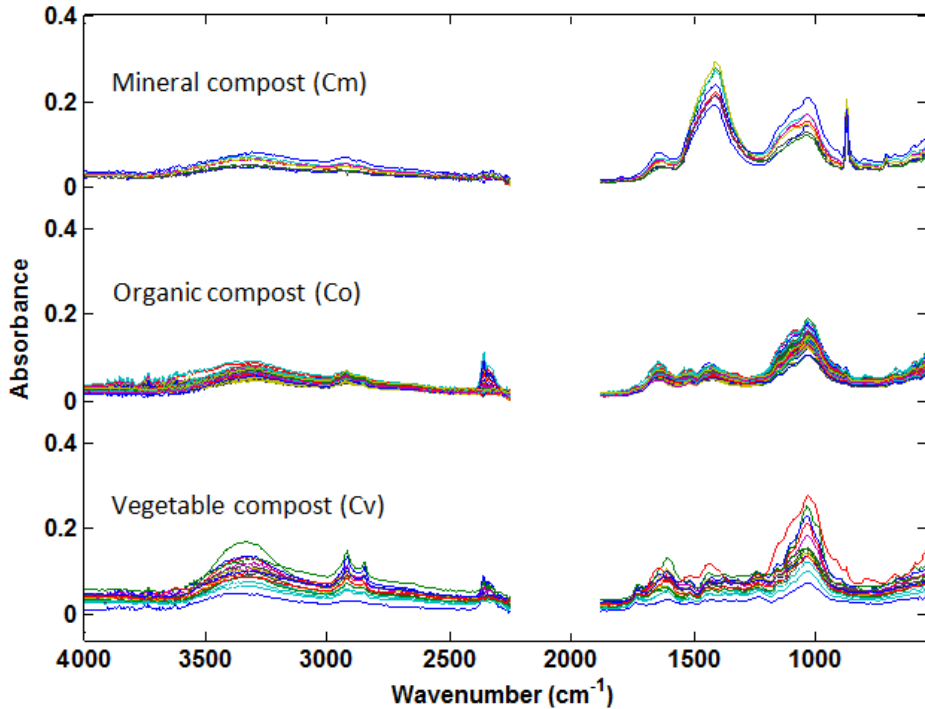


Figura 5.4. Espectros ATR-MIR de las distintas clases de compost en la región entre 4000 y 550 cm^{-1}

Nota: La región entre 2240 y 1880 cm^{-1} ha sido eliminada debido a la alta absorbancia del cristal ATR. Los espectros se desplazaron sobre el eje de absorbancia y se mostraron en grupos para evitar la superposición espectral

3.5.A.3.2.1. Análisis exploratorio: PCA

El PCA se empleó para visualizar las muestras de compost a partir de sus espectros ATR-MIR usando la región espectral entre 1760 y 550 cm^{-1} y un pretratamiento de datos basado en MC y SNV. Los 5 primeros componentes principales explican un varianza total del 96.97%.

En la Figura 5.5 se representa el gráfico de puntuaciones de las dos primeras componentes principales del modelo (PC1 con 68.16% de varianza y PC2 con un 18.68%), permitiendo ver como las muestras se agrupan en tres grupos diferentes de acuerdo con su naturaleza. A lo largo del PC1, se puede observar a los composts Cm situados más hacia la derecha, mientras que Cv y Co se encuentran situados más hacia la izquierda, según la concentración de fenoles totales. El PC2 es capaz de distinguir con bastante claridad los grupos Co y Cv, que no era posible mediante PC1. El grupo Cm presenta la mayor diferencia respecto al resto de los

composts evaluados debido a su alta estabilidad y grado de mineralización, que llevan asociada una baja concentración de fenoles. En comparación con el PCA realizado a partir de los espectros NIR, el PCA con los espectros MIR proporciona una mejor separación y agrupación de los diferentes tipos de compost, recomendándose el uso de las medidas ATR-MIR para el análisis y clasificación de las muestras de compost para diferencias según su origen.

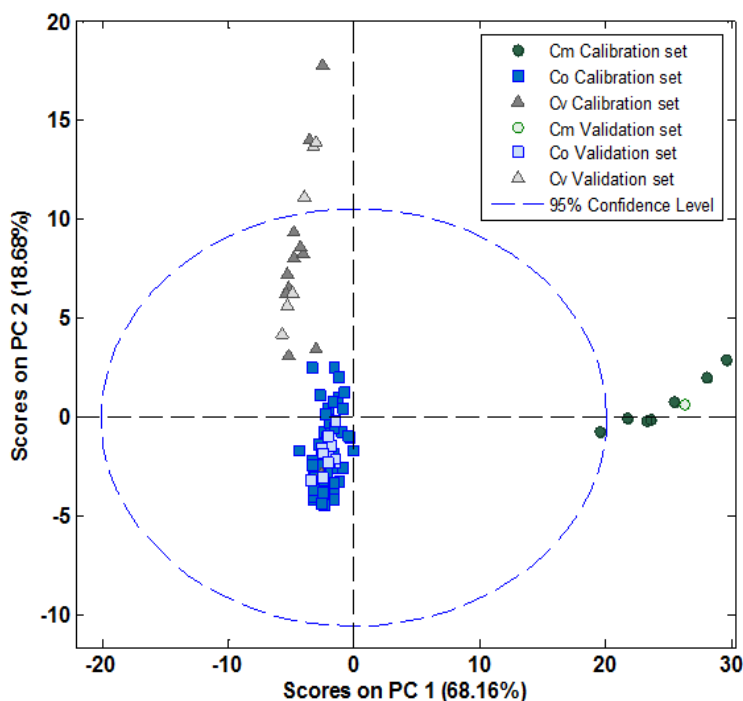


Figura 5.5. Resultados del análisis del componente principal basado en los espectros de ATR-MIR.

3.5.A.3.2.2. Modelos PLS-ATR-FTIR

Con el fin de comparar ambas regiones MIR y NIR, se emplearon los mismos conjuntos de calibración y validación que se utilizaron anteriormente para el modelado PLS-NIR. Se consideraron diferentes regiones espectrales y estrategias de pretratamiento (ver Tabla 5.2), encontrándose los mejores resultados usando MC y SNV como pre-tratamiento y la región comprendida entre 1760 y 550 cm^{-1} , que contiene las bandas más importantes concernientes a TPC. Para validar los modelos se utilizó un conjunto independiente de muestras de compost no consideradas para la calibración. Como en los espectros NIR, se comprobó en el

gráfico de PCA la correcta distribución de las muestras incluidas en los conjuntos de calibración y validación, verificando que las muestras del conjunto de calibración son representativas de la población total (véase Figura 5.4).

Utilizando 4 LV's, se obtuvieron valores de RMSEC de $0.23 \text{ mg GAE g}^{-1} \text{ DW}$ y R^2_{cal} de 0.94, RMSECV de $0.28 \text{ mg GAE g}^{-1} \text{ DW}$ y R^2_{CV} de 0.91 y, finalmente, RMSEP de $0.27 \text{ mg GAE g}^{-1} \text{ DW}$ y R^2_{pred} de 0.98; el RPD obtenido fue de 5.83. La gran similitud entre los errores de calibración y de validación cruzada es indicativo de la robustez del modelo, con un RPD de 5.8 que indica una buena capacidad predictiva, considerándose el método propuesto para la determinación de TPC en compost adecuado para propósitos analíticos.

Tabla 5.2. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-MIR para la determinación de TPC en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración				Conjunto de Validación			
			RMSEC (mg GAE g ⁻¹ DW)	R ² Cal	RMSECV (mg GAE g ⁻¹ DW)	R ² CV	RMSEP (mg GAE g ⁻¹ DW)	R ² pred	RPD	
MC, SNV	3000-2800 & 1760-550	4	0.22	0.94	0.28	0.90	0.33	0.97	4.76	
MC, FD	3000-2800 & 1760-550	5	0.23	0.94	0.31	0.88	0.31	0.96	5.08	
MC, MSC, FD	3000-2800 & 1760-550	3	0.22	0.94	0.27	0.91	0.40	0.96	4.03	
MC, SNV	1760-550	4	0.23	0.94	0.28	0.91	0.27	0.98	5.83	
MC, FD	1760-550	4	0.26	0.92	0.31	0.88	0.33	0.96	4.80	
MC, MSC, FD	1760-550	3	0.22	0.94	0.24	0.93	0.33	0.97	4.89	

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R²Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

3.5.A.4. Comparación MIR y NIR

Para poder establecer una comparación entre los modelos PLS-NIR y PLS-MIR seleccionados para la determinación del contenido de TPC en composts, se utilizaron los mismos conjuntos de calibración y de validación para la construcción de los modelos. Los valores encontrados para RMSEC en el modelo PLS-MIR fueron mayores que los obtenidos en el modelo PLS-NIR, con valores de R^2 Cal mejores que los encontrados en los modelos basados en los espectros MIR. En cuanto la capacidad predictiva, el modelo PLS-NIR proporcionó mejores resultados con RPD de 8.2, pero cabe decir, que ambas técnicas pueden considerarse adecuadas para fines analíticos.

3.5.A.5. Conclusiones Capítulo 5.A

Los métodos de calibración multivariable desarrollados utilizando tanto la espectroscopia NIR como MIR, ofrecen una alternativa sostenible para la determinación de TPC en muestras de compost ya que basan únicamente en la obtención de los espectros IR sin necesidad del empleo de reactivos ni disolventes. Sin embargo, al comparar las principales características analíticas de los dos métodos desarrollados, se puede concluir que el NIR proporciona una capacidad de predicción mejor en comparación con el MIR; basado en los valores de RPD y del error para el conjunto de validación (RMSEP) que fue de 0.19 y 0.27 mg GAE g⁻¹ DW para NIR y MIR, respectivamente.

En cualquier caso, se pueden emplear las dos metodologías propuestas para la determinación de TPC, siendo las medidas NIR más rápidas y seguras debido a que las muestras se introducen en viales de vidrio, mientras que en el MIR es necesario tener en cuenta un protocolo de limpieza adecuado para evitar la contaminación cruzada entre muestras.

3.5.B. Capítulo 5.B. Determinación del contenido de carbono orgánico y nitrógeno total

3.5.B.1. Muestras

Para esta parte del estudio se amplió la población a un total de 121 muestras de residuos orgánicos, compost y mezclas de residuos y compost. Todas las muestras de compost empleadas se obtuvieron por fermentación aeróbica en un reactor a escala de laboratorio, mientras que las mezclas se prepararon añadiendo varias cantidades de compost maduros a desechos orgánicos y adecuada homogenización. Las principales clases de muestras evaluadas en este estudio son: i) compost de lodos de depuradora y residuos sólidos urbanos (13 muestras); ii) diferentes residuos orgánicos (9 muestras), y iii) mezclas entre residuos orgánicos y compost maduros (99 muestras).

El total de muestras de compost se dividió en un conjunto de calibración y otro de validación, con 91 y 30 muestras, respectivamente, mediante el algoritmo de Kennard-Stone. La concentración de CO varió entre 9.23% y 55.0% m/m para el conjunto de calibración y del 11.68 al 41.9% en el conjunto de validación. En el caso de TN, las concentraciones cubrían el intervalo de 0.54 a 4.35% m/m para la calibración y del 1.82 al 4.39% para la validación.

3.5.B.2. Métodos de referencia

Los contenidos de referencia de OC y TN se determinaron de acuerdo con los métodos propuestos por la Asociación de Normalización francés (AFNOR, del francés *Association française de Normalisation*).

El OC se determinó de acuerdo con norma AFNOR (X31-109) [26], dónde 0.05 g de muestra seca se oxida a una temperatura controlada de 135° C, en presencia de $K_2Cr_2O_7$ y ácido sulfúrico, durante 5 min. Después de enfriar, la disolución resultante se ajusta con agua destilada hasta un volumen final de 200 mL y se homogeniza. A 25 mL de la solución antes mencionada se añaden 5 ml de ácido ortofosfórico y unas pocas gotas de indicador difenilamina. El exceso de $K_2Cr_2O_7$ se valora con una disolución normalizada (1 N) de la sal de Mohr. Para cada muestra se realizan tres réplicas y se considera el valor promedio de las mismas.

El TN se determinó de acuerdo con la norma AFNOR (X31-111) [26]. 1 g de muestra seca en polvo se mineraliza durante 2h con ácido sulfúrico en presencia de una mezcla de catalizadores, enfriándola posteriormente a temperatura ambiente. A la disolución resultante se le añaden 50 mL de NaOH (40%) y se

destila, recogiendo el amoniaco destilado sobre un volumen de una disolución de ácido bórico (4%). Finalmente se valora el borato formado con una disolución normalizada de ácido sulfúrico (0.2 N) en presencia del indicador Tashiro. El análisis se realizó por triplicado y se consideró el valor promedio.

Las muestras del estudio cubren un intervalo de contenido OC comprendido entre el 9.23% y el 45.36% m/m, referido a la muestra de compost seca, y del 0.55 al 4.40% m/m para TN.

3.5.B.3. Resultados y discusión

Al igual que en el estudio anterior, una vez registrados los espectros IR, se procedió al estudio y modelización de los mismos.

3.5.B.3.1. Espectroscopia NIR

3.5.B.3.1.1. Modelos PLS-NIR: Carbono Orgánico

Considerando los valores de referencia determinados para el CO se construyeron distintos modelos PLS considerando diferentes pre-procesamientos de los datos y distintas regiones espectrales. Los mejores modelos obtenidos se incidan en la Tabla 5.3 y, como puede verse, se caracterizan por valores de R^2_{cal} comprendidos entre 0.80 y 0.85, valores de RMSEC de 3.23 a 3.82 % m/m, R^2_{CV} de 0.69 a 0.73, con RMSECV de 4.43 a 4.79 % m/m, valores de R^2_{Pred} de 0.41 a 0.61 y RMSEP entre 4.09 y 4.92 % m/m, con RPD desde 1.33 hasta 1.60 para la predicción de un conjunto de muestras independiente del utilizado para la calibración.

Los mejores resultados se encontraron con el modelo basado en la segunda derivada y centrado en la media como pre-procesado, en la región entre 6200 y 4400 cm^{-1} con 4LVs. En estas condiciones se obtuvo un RMSEC de 3.23% m/m con R^2_{cal} igual a 0.85, siendo el RMSECV 4.79 y RMSEP de 4.09% m/m, con R^2_{cv} y R^2_{pred} de 0.69 y 0.61, respectivamente.

Tabla 5.3. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-NIR seleccionados para la determinación de OC en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración				Conjunto de Validación		
			RMSE C (%)	R ² Cal	RMSECV (%)	R ² cv	RMSEP (%)	R ² pred	RPD
MC-FD	6200-4400	5	3.82	0.80	4.51	0.72	4.92	0.41	1.35
MC-SD	6200-4400	4	3.23	0.85	4.79	0.69	4.09	0.61	1.60
MC-FD	8600-8000 and 6200-4400	5	3.52	0.83	4.43	0.73	4.85	0.44	1.33

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

En la Figura 5.6 se representan los valores predichos frente los valores de referencia del OC encontrado en las muestras de residuos orgánicos y composts para la predicción del conjunto de validación.

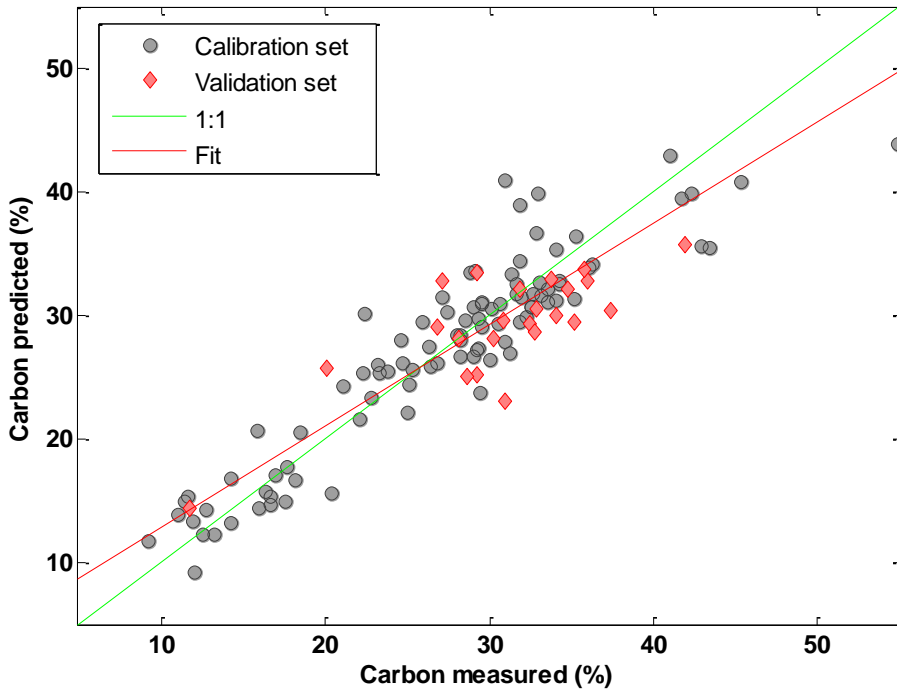


Figura 5.6. Correlación PLS-NIR entre los valores predichos y los valores de referencia del contenido de carbono orgánico en residuos orgánicos y compost.

La Figura 5.7 corresponde con la gráfica de cargas para las dos primeras LV, que indican que la parte principal de la información fue proporcionada por la región de 5600 y 5000 cm^{-1} para la LV1 (con una varianza explicada del 81.13%) mientras que la LV2 (varianza explicada del 6.14%) también resulta significativa es la misma región además de las variables alrededor de 4987, 4528 y 4476 cm^{-1} .

La regresión entre los resultados obtenidos por PLS-NIR y los datos de OC medidos en las muestras utilizando el método químico no fue tan buena, proporcionando una ecuación de $y = 0.860x + 3.8$.

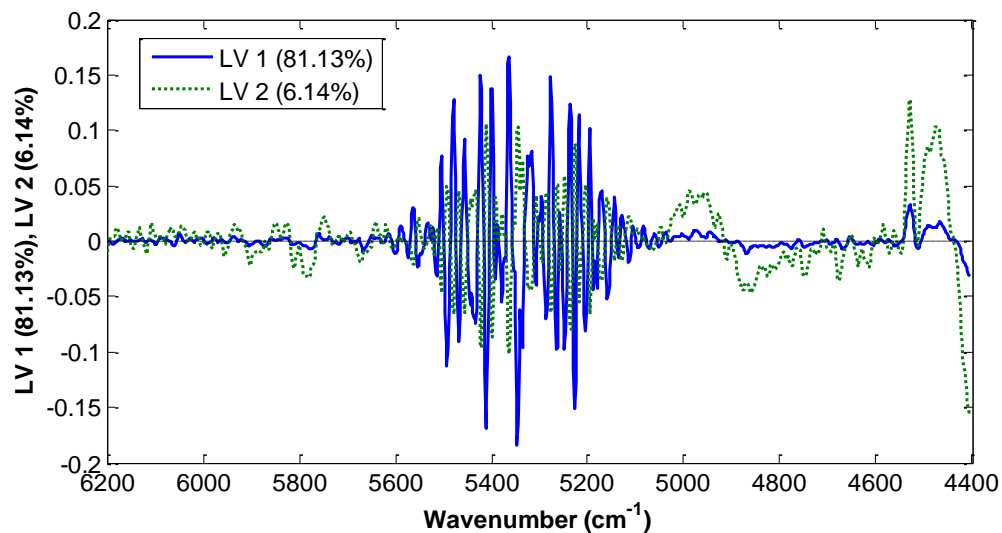


Figura 5.7. Mapa de cargas de las dos primeras variables latentes seleccionadas para construir los mejores modelos de determinación de OC.

3.5.B.3.1.2. Modelos PLS-NIR: Nitrógeno Total

En la Tabla 5.4 se indican las principales características de los modelos PLS-NIR finalmente considerados para la determinación del nitrógeno total, contruidos teniendo en cuenta el uso de diferentes pre-tratamientos y regiones espectrales.

El mejor resultado fue para el modelo de segunda derivada y centrado en la media de los datos en la región espectral de 7400 a 5002 cm^{-1} , usando 5 LVs. Los valores de R^2_{cal} y R^2_{pred} fueron 0.89 y 0.82, con un RMSEC de 0.25% m/m y un RMSEP de 0.34% m/m.

Tabla 5.4. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-NIR seleccionados para la determinación de TN en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración				Conjunto de Validación			
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² cv	RMSEP (%)	R ² pred	RPD	RPD
FD-MC	10045-4000	6	0.23	0.91	0.43	0.69	0.40	0.68	1.67	
FD-MC	7400-5002	7	0.26	0.88	0.40	0.74	0.38	0.74	1.79	
SD-MC	7400-5002	5	0.25	0.89	0.50	0.74	0.34	0.82	2.00	

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R²Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

En la Figura 5.8 se representan los valores predichos frente los valores de referencia del nitrógeno total encontrado en calibración mediante validación cruzada y en la validación de las muestras de composts, obteniéndose para esta última una recta de ecuación $y = 0.898x + 0.25$.

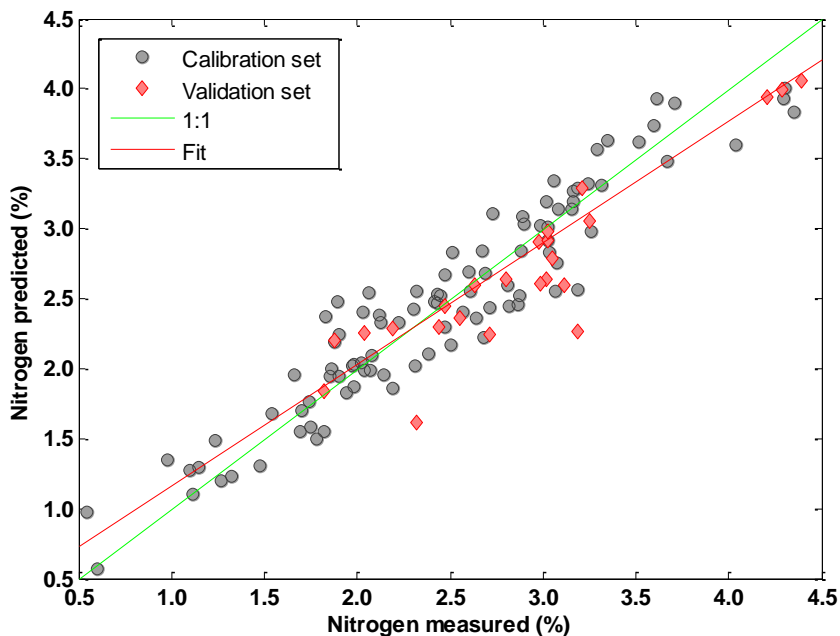


Figura 5.8. a) Correlación PLS-NIR entre los valores predichos y los valores de referencia del contenido de nitrógeno total en desechos orgánicos y compost.

En cuanto el gráfico de cargas (véase Figura 5.9) LV1 explica el 86.27 de la varianza total y está ligada a las regiones de 7400 a 6900 cm^{-1} y de 5600 a 5100 cm^{-1} . Sin embargo, como se puede ver en esta figura, la información sobre TN en espectros DR-NIR no está asociada a ninguna banda característica en especial.

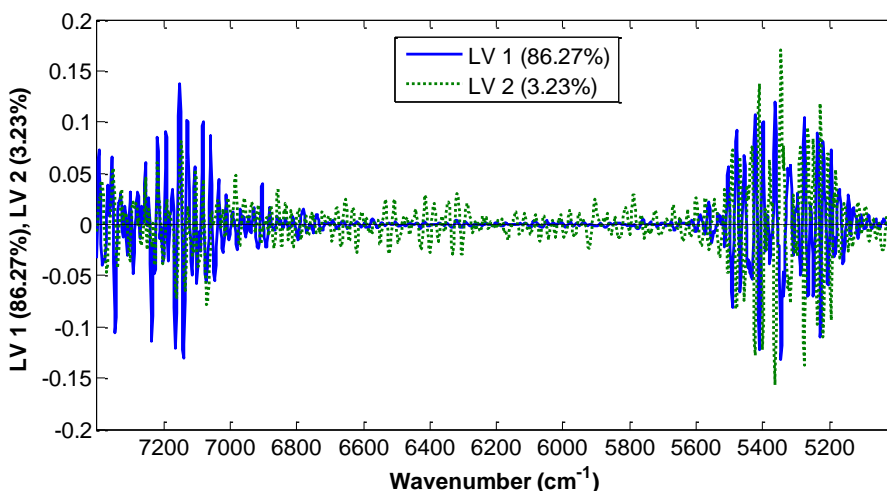


Figura 5.9. Mapa de cargas de las dos primeras variables latentes seleccionadas para construir los mejores modelos de determinación de NT

3.5.B.3.2. Espectroscopia MIR

3.5.B.3.2.1. Modelos PLS-MIR: Carbono orgánico

Los mejores modelos ensayados para la determinación del contenido de OC en muestras de composts, basados en el uso de diferentes métodos de pre-procesamiento en distintas regiones del espectro ATR-MIR (ver Tabla 5.5), se caracterizan por valores de R^2_{cv} de 0.62 a 0.71 y RMSECV de 4.26% a 5.50% m/m para la validación cruzada y R^2_{pred} de 0.40 a 0.76 y RMSEP entre 2.20% y 4.55% m/m, con valores de RPD desde 0.97 hasta 2.00, para el conjunto independiente de muestras utilizado para el método de validación. La mejor opción de modelización fue mediante el uso de OSC y centrado en la media en las regiones de 3014-2300 y 1776-550 cm^{-1} , con un valor de RPD de 2.0, que se considera adecuado para la cuantificación de OC en muestras desconocidas.

Tabla 5.5. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-MIR seleccionados para la determinación de OC en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración			Conjunto de Validación			
			RMSEC (%)	R ² cal	RMSECV (%)	R ² cv	RMSEP (%)	R ² pred	RPD
MC-FD	1602-1247 and 534-400	5	4.16	0.78	4.95	0.69	3.7	0.4	1.2
MC-OSC	3014-2300 and 1776-549	2	1.24	0.98	4.26	0.67	2.2	0.76	2.0
MC-OSC	1751-549	2	2.95	0.89	5.15	0.67	2.84	0.61	1.57
MC-FD	3014-2300 and 1776-549	3	4.92	0.69	5.5	0.62	3.74	0.31	1.19
MC-OSC	2966-2625 and 1250-900	4	4.16	0.78	4.8	0.71	4.55	0.4	0.97

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R²Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

La Figura 5.10 representa los valores predichos frente los valores de referencia del carbono orgánico encontrado en las muestras de composts, mientras que en la Figura 5.11 se ilustra la gráfica de cargas para las dos primeras LV's, que explican, respectivamente, el 42.66% y el 46.07% de la varianza total, con un mayor peso en los rangos de $3014\text{-}2300\text{ cm}^{-1}$ y especialmente en bandas de 1406 y 871.8 cm^{-1} para LV1 y de 1034 cm^{-1} para LV2. Usando 2 variables latentes el RMSEC obtenido fue de 1.24% m/m con un valor de R^2 Cal de 0.98 , siendo el RMSECV igual a 4.26% m/m y RMSEP de 2.20% m/m con R^2 CV y R^2 Pred de 0.67 y 0.76 .

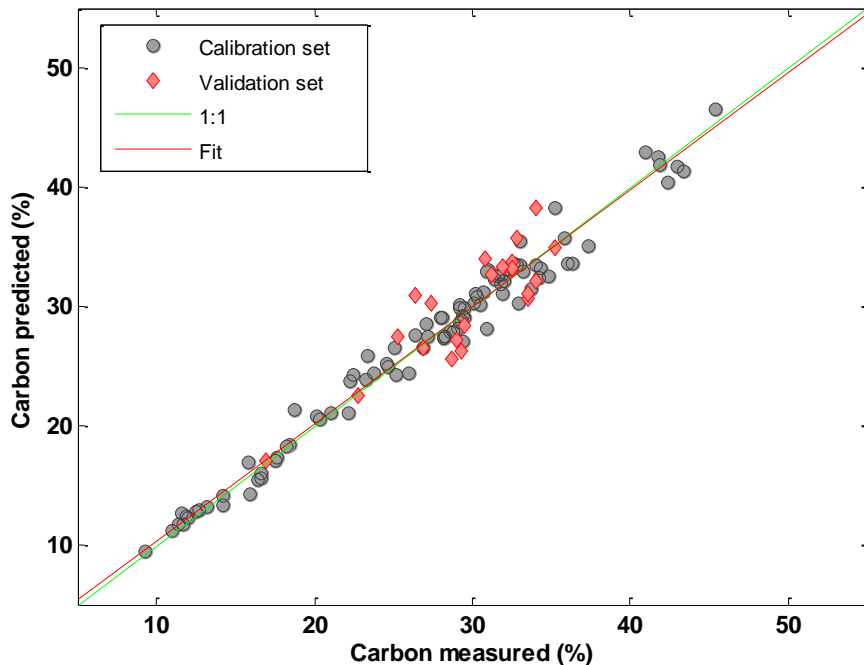


Figura 5.10. Correlación PLS-MIR entre los valores predichos y los valores de referencia del contenido de carbono orgánico en desechos orgánicos y compost.

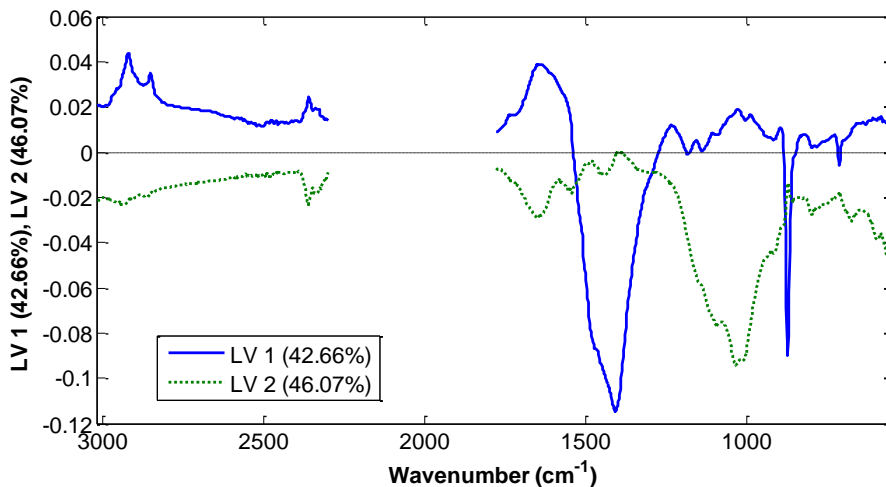


Tabla 5.11. Mapa de cargas de las dos primeras variables latentes seleccionadas para construir los mejores modelos de determinación de OC

3.5.B.3.2.2. Modelos PLS-MIR: Nitrógeno Total

Diferentes modelos de PLS, basados en los espectros ATR-MIR, se construyeron para la determinación de TN siguiendo el mismo procedimiento descrito en los casos anteriores según puede verse en los resultados incluidos en la Tabla 5.6, el modelo que proporcionó la mejor calibración y capacidad predictiva en la validación fue mediante el uso del pre-procesado SNV y centrado en la media, de las regiones de 2966-2625 y 1593-565 cm^{-1} y utilizando 5 LV's.

Tabla 5.6. Parámetros de calibración, validación cruzada y validación externa de los modelos PLS-MIR seleccionados para la determinación de TN en composts utilizando diferentes regiones espectrales y pre-procesamientos de los datos.

Pre-procesado	Región (cm ⁻¹)	LVs	Conjunto de Calibración			Conjunto de Validación			
			RMSEC (%)	R ² Cal	RMSECV (%)	R ² cv	RMSEP (%)	R ² Pred	RPD
MC-FD	2966-2625 and 1593-565	5	0.33	0.83	0.39	0.78	0.21	0.78	2.28
MC-SD	2966-2625 and 1593-565	6	0.29	0.87	0.38	0.78	0.27	0.64	1.77
MC-SD	1593-565	5	0.33	0.83	0.39	0.77	0.25	0.67	1.92
MC-SNV	2966-2625 and 1593-565	5	0.38	0.78	0.42	0.74	0.20	0.82	2.39

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R²Cal: coeficiente de determinación de calibración; R²CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

La Figura 5.12 presenta la regresión entre los resultados obtenidos para la determinación de TN por PLS-MIR y los valores medidos utilizando el método de referencia en muestras de compost y de residuos orgánicos. Además, en la Figura 5.13 se puede ver el gráfico de cargas para las dos primeras LV, que explican, respectivamente, el 78.85% y el 6.54% de la varianza total, que presentan un mayor peso y significancia en la región de 1400 a 500 cm^{-1} . Usando 5 variables latentes el RMSEC fue de 0.38% m/m con R^2_{cal} de 0.78, siendo el RMSECV igual a 0.42% m/m y RMSEP de 0.20 m/m con R^2_{cv} y R^2_{pred} de 0.74 y 0.82, respectivamente.

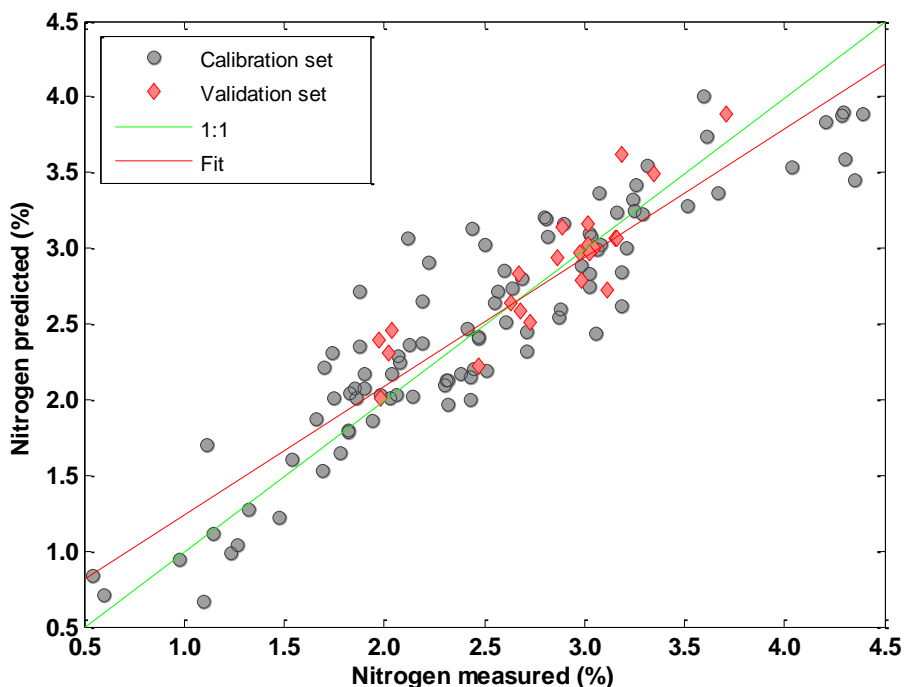


Figura 5.12. Correlación PLS-MIR entre los valores predichos y los valores de referencia del contenido de nitrógeno total en desechos orgánicos y compost.

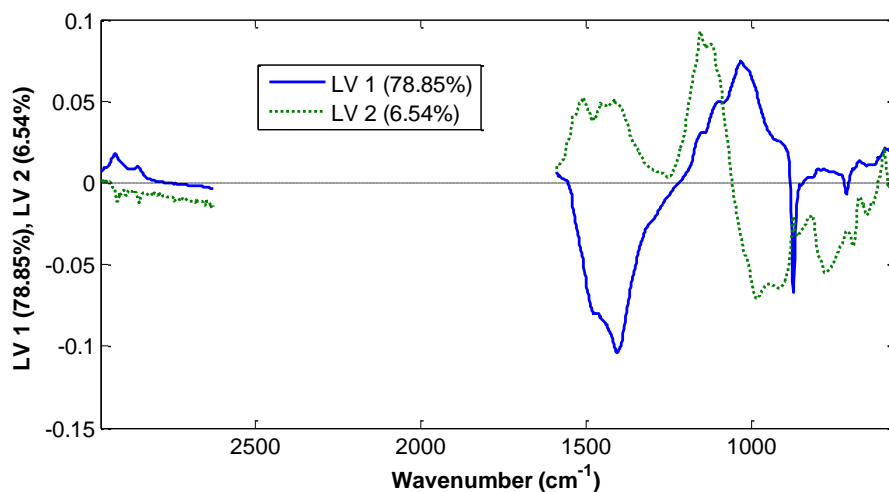


Tabla 5.13. Mapa de cargas de las dos primeras variables latentes seleccionadas para construir los mejores modelos de determinación de NT

3.5.B.4. Comparación MIR y NIR

Los modelos PLS construidos a partir de los espectros NIR y MIR para la determinación del contenido de OC y NT en residuos orgánicos y composts, se construyeron con los mismos conjuntos de calibración y validación, con el fin de poder establecer una comparación entre los modelos PLS. Los valores encontrados para el modelo PLS-MIR son mejores que los obtenidos en el modelo PLS-NIR, ya que presentan menores errores de predicción tanto para la determinación de carbono orgánico como para la determinación de nitrógeno total en muestras de residuos orgánicos y composts. En el caso de los modelos PLS-NIR, los valores RPD obtenidos fueron de 1.6 y 2, para OC y NT, respectivamente, mientras que para los modelos PLS-MIR dichos valores fueron de 2 y 2.39, si bien es cierto reconocer que en ningún caso puedan considerarse plenamente satisfactorio.

3.5.B.5. Conclusiones Capítulo 5.B

El presente estudio demuestra la utilidad de la espectroscopia infrarroja, MIR y NIR, para la predicción de los contenidos de carbono orgánico y nitrógeno total en muestras de compost y residuos orgánicos. De los resultados obtenidos se puede concluir que los métodos basados en PLS-MIR proporcionan una mejor capacidad predictiva que el PLS-NIR. No obstante, en su conjunto, podríamos considerar que ambos resultan ser métodos alternativos verdes y sostenibles para la

determinación de OC y TN en composts por lo menos con fines de cribado. Cabe destacar que el procedimiento NIR está libre del uso de reactivos o disolventes, no generándose residuos, mientras que el procedimiento MIR requiere pequeñas cantidades de disolventes de limpieza, pero produciéndose volúmenes reducidos de residuos.

3.5.3. Referencias Capítulo 5

- [1] A.E. Ghaly, B. Zhang, D. Dave, Biodegradation of phenolic compounds in creosote treated wood waste by a composting microbial culture augmented with the fungus *Thermoascus aurantiacus*, *Am. J. Biochem. Biotechnol.*, 2(2011) 90–103.
- [2] C.A. Palm, C.N. Gachengo, R.J. Delve, G. Cadisch, Ken E. Giller, Organic inputs for soil fertility management in tropical agroecosystems: application of an organic resource database, *Agric. Ecosyst. Environ.*, 83(2001) 27–42.
- [3] R. Albrecht, R. Joffre, R. Gros, J. Le Petit, G. Terron, C. Périssol, Efficiency of nearinfrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process, *Bioresour. Technol.*, 99(2008) 448–455.
- [4] O. Fourti, N. Jedidi, A. Hassen, Physico-chemical aspects during the composting of municipal solid wastes and sewage sludge in a semi-industrial composting plant, *Afr. J. Microbiol. Res.*, 7(2013) 1055–1068.
- [5] G.F. Huang, Q.T. Wu, J.W.C. Wong, B.B. Nagar, Transformation of organic matter during co-composting of pig manure with sawdust, *Bioresour. Technol.*, 97(2006) 1834–1842.
- [6] M.F. Hirai, V. Chanyasak, H. Kubota, A standard method for measurement of compost maturity, *Bio Cycles*, 24(1983) 54–56.
- [7] A. Lakhdar, H. Falleh, Y. Ouni, S. Oueslati, A. Debez, R. Ksouri, C. Abdelly, Municipal solid waste compost application improves productivity, polyphenol content, and antioxidant capacity of *Mesembryanthemum edule*, *J. Hazard. Mater.*, 191(2011) 373–379.
- [8] F.Z. El Ouaquodi, L. El Fels, L. Lemée, A. Amblès, M. Hafidi, Evaluation of lignocelullose compost stability and maturity using spectroscopic (FTIR) and thermal (TGA/TDA) analysis, *Ecol. Eng.*, 75(2015) 217–222.
- [9] E. Compaoré, L.S. Nanéma, Compostage et qualité du compost de déchets urbains solides de la ville de Bobo-Dioulasso, Burkina Faso, *Tropicultura*, 28(2010), 232–237.
- [10] Z. Wu, E. Xu, J. Long, X. Pan, X. Xu, Z. Jin, A. Jiao, Comparison between ATR-IR, Raman, concatenated ATR-IR and Raman spectroscopy for the determination of total antioxidant capacity and total phenolic content of Chinese rice wine, *Food Chem.*, 194(2016) 671–679.

- [11] X. Lu, J. Wang, H.M. Al-Qadiri, C.F. Ross, J.R. Powers, J. Tang, B.A. Rasco, Determination of total phenolic content and antioxidant capacity of onion (*Allium cepa*) and shallot (*Allium oschaninii*) using infrared spectroscopy, *Food Chem.*, 129(2011) 637–644.
- [12] M.J. Martelo-Vidal, M. Vázquez, Determination of polyphenolic compounds of red wines by UV–vis–NIR spectroscopy and chemometrics tools, *Food Chem.*, 158(2014) 28–34.
- [13] R. Ferrer-Gallego, J.M. Hernández-Hierro, J.C. Rivas-Gonzalo, M.T. Escribano-Bailón, Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy, *LWT – Food Sci. Biotechnol.*, 44(2011) 847–853.
- [14] R. Gislum, E. Micklander, J.P. Nielsen, Quantification of nitrogen concentration in perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and chemometrics, *Field Crop. Res.*, 88(2004), 269–277.
- [15] V. Bellon-Maurel, A. McBratney, Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – critical review and research perspectives, *Soil Biol. Biochem.*, 43(2011) 1398–1410.
- [16] Y. Ge, J.A. Thomasson, C.L.S. Morgan, Mid-infrared attenuated total reflectance spectroscopy for soil carbon and particle size determination, *Geoderma*, 213(2014) 57–63.
- [17] L. Galvez-Sola, R. Moral, M.D. Perez-Murcia, A. Perez-Espinosa, M.A. Bustamante, E. Martinez-Sabater, C. Paredes, The potential of near infrared reflectance spectroscopy (NIRS) for the estimation of agroindustrial compost quality, *Sci. Total Environ.*, 408(2010), 1414–1421.
- [18] A. Lakhdar, H. Falleh, Y. Ouni, S. Oueslati, A. Debez, R. Ksouri, C. Abdelly, Municipal solid waste compost application improves productivity, polyphenol content, and antioxidant capacity of *Mesembryanthemum edule*, *J. Hazard. Mater.*, 191(2011) 373–379.
- [19] C.N.T. Frizon, G.A. Oliveira, C.A. Perussello, P.G. Peralta-Zamora, A.M.O. Camlofski, Ü.B. Rossa, R. Hoffmann-Ribani, Determination of total phenolic compounds in yerba mate (*Ilex paraguariensis*) combining near infrared spectroscopy (NIR) and multivariate analysis, *LWT – Food Sci. Technol.*, 60(2015) 795–801.

- [20] W. Li, H. Qu, Rapid quantification of phenolic acids in Radix Salvia Miltorrhiza extract solutions by FT-NIR spectroscopy in transreflective mode, *J. Pharm. Biomed. Anal.*, 52(2010) 425–431.
- [21] C.E. Millier, Chemical principals of near infrared technology, in: P.C. Williams, K.H. Norris (Eds.), *Near Infrared Technology in the Agricultural And Food Industries*, 2nd ed., American Association of Cereal Chemists, St. Paul, MN, 2001, 19–39.
- [22] J.S. Ribeiro, S. Garrigues, M. de la Guardia, FTIR tentative characterization of humic acids extracted from organic materials, *Spectrosc. Lett.*, 34(2001) 179–190.
- [23] C. Peltre, S. Bruun, C. Du, I.K. Thomsen, L.S. Jensen, Assessing soil constituents and labile soil organic carbon by mid infrared photoacoustic spectroscopy, *Soil. Biol. Biochem.*, 77(2014) 41–50.
- [24] M.M. Cascant, S. Garrigues, M. de la Guardia, Direct determination of major components in human diets and baby foods, *Anal. Bioanal. Chem.*, 407(2015) 1961–1972.
- [25] E. Smidt, K. Meissl, The applicability of modern analytical tools in waste management Fourier Transform Infrared (FT-IR) spectroscopy, *Waste Manag.*, 27(2006) 268–276.
- [26] AFNOR, *Qualité des sols, Méthodes d'analyses*. 1ère édition, 1987.

3.6. CAPÍTULO 6:

**Determinación de benceno, tolueno, etilbenceno y xilenos
en suelos mediante espectroscopia infrarroja en la región
media en fase de vapor**

3.6. Capítulo 6. Determinación de benceno, tolueno, etilbenceno y xilenos en suelos mediante espectroscopia infrarroja en la región media en fase de vapor

3.6.1. Introducción

Las sustancias químicas peligrosas presentes en los suelos contaminados pueden suponer un riesgo para la salud humana y el medio ambiente. La Directiva 2008/98/CEE del Parlamento Europeo y del Consejo, del 19 de Noviembre del 2008 [1], establece el marco jurídico para la generación y gestión de residuos, incluyendo las medidas adecuadas para prevenir o reducir los efectos adversos. Los suelos contaminados en España están regulados por el Real Decreto del 14 de Enero del 2005 (RD 9/2005) [2] y la Ley 22/2011, del 28 de Julio, de Desechos y suelos contaminados [3] que establecen la lista de actividades potencialmente contaminantes del suelo y los criterios y valores para la declaración de suelos contaminados, incluyendo una lista de contaminantes y niveles de referencia genéricos para la protección de la salud humana [2]. Según la información recogida en esta lista, los niveles de benceno, tolueno, etilbenceno y xilenos (BTEX) son unos parámetros a tener en cuenta para ser evaluados en la contaminación del suelo [4 - 8].

El BTEX es una mezcla de compuestos orgánicos volátiles aromáticos ampliamente distribuidos en el medio ambiente debido a fuentes naturales como plantas, algas y plancton, y fuentes antropogénicas, como derrames y fugas de productos petrolíferos. Los compuestos BTEX tienen puntos de ebullición entre 80 y 150 °C y presiones de vapor entre 0.8 y 10 KPa a 20 °C. Debido a su alta volatilidad y su solubilidad en el agua, estos compuestos son fuentes potenciales de contaminación de suelos y aguas subterráneas. La exposición a BTEX puede ocurrir por ingestión, inhalación o absorción a través de la piel, ya que la ingesta de BTEX puede ocurrir a través del consumo de agua contaminada o la ingestión de productos agrícolas. Algunos de los compuestos BTEX pueden ser tóxicos, mutagénicos o carcinógenos [9, 10]

Los métodos analíticos para la determinación de residuos BTEX en muestras ambientales se basan generalmente en la cromatografía de gases acoplada a la espectrometría de masas (GC-MS) [4-6], [11], cromatografía de gases con detector de foto-ionización (GC-PID) [12,13], cromatografía de gases con detector de ionización de llama (GC-FID) [7,8] [14-17], cromatografía de gases con espectroscopia de infrarrojos (GC-IR) [18] y cromatografía líquida [19]. PID permite la discriminación de compuestos BTEX excepto para meta- y para-xileno que deben ser identificados y cuantificados juntos debido a su co-elución. Un detector FID es un detector no selectivo y tiene interferencias potenciales, presentes en las

muestras. Además, en los últimos años la sensibilidad de los métodos de GC para la determinación de BTEX se ha incrementado mediante la incorporación de un paso de micro extracción anterior [7, 8, 11, 13, 14].

En cuanto al uso de técnicas vibracionales para la determinación de BTEX en muestras ambientales, sólo unos pocos trabajos están disponibles en la literatura [20-24] y corresponden al uso de medidas de reflexión total atenuada (ATR) en la región de IR medio (MIR) en suelos [20], en mezclas aceite-agua [21] y muestras de agua [22, 23], o el uso de medidas de transmisión en el infrarrojo cercano (NIR) para el análisis de muestras de agua [24].

Los métodos utilizados para la determinación de BTEX implican, frecuentemente, tratamientos previos de separación como la extracción con disolvente, el calentamiento térmico o la purga de la muestra con un gas inerte. La purga y atrapamiento (P&T) con extracción con metanol es la técnica más frecuentemente aplicada para la extracción y pre-concentración de BTEX a partir de muestras de suelo y agua [5]. Los métodos para la determinación de BTEX en el suelo se han basado en la técnica de espacio de cabeza combinado con la micro-extracción en fase sólida (SPME) [7, 8]. Sin embargo, puede ocurrir la contaminación por arrastre del sistema de purga y trampa cuando se analizan muestras de alta y baja concentración en una misma secuencia [5], siendo el espacio de cabeza una mejor técnica alternativa.

Los métodos de espectroscopia vibracional han sido considerados como alternativas analíticas verdes a la cromatografía, ya que son rápidos y no destructivos y las muestras pueden analizarse sin o con una preparación mínima de las mismas [25]. La espectroscopia NIR y MIR se ha aplicado para la determinación de compuestos aromáticos en la refinería y productos petroquímicos, como la gasolina y diesel [26-28], siendo el IR una excelente herramienta para la determinación cuantitativa de compuestos volátiles o semi volátiles basados en la generación de la fase vapor del analito [29]. Esta técnicas de fase vapor en el infrarrojo también se ha aplicado para determinación de sustancias volátiles en diferentes matrices de muestra [30-33]. Una limitación del empleo de la fase vapor en el infrarrojo, especialmente en la región MIR, es la presencia de CO₂ y H₂O en las muestras, que pueden proporcionar interferencias espectrales y afectar a las determinaciones analíticas que se están realizando, pero es posible utilizar diferentes funciones de corrección de agua y CO₂ desarrolladas por los fabricantes de instrumentación e incluidas en sus paquetes de software [34-36]. Con el fin de evitar o reducir este problema de la presencia de CO₂ y H₂O también se propuso el uso de adsorbentes sólidos para retener el agua [20].

Por tanto, el objetivo de este trabajo es el desarrollo de una metodología simple que combine el uso del muestreo en espacio en cabeza para la generación de fases vapor con la detección IR para una determinación rápida y directa de BTEX en muestras de suelo, evitando el uso de disolventes extractantes y de reactivos, y sin la necesidad de una separación cromatográfica. Para ello se considerará el tratamiento quimiométrico de los espectros IR en fase vapor para la determinación del benceno, tolueno, etilbenceno, orto-, meta- y para-xileno, eliminando las interferencias espectrales del agua y del CO₂, utilizando una metodología de corrección de los espectros desarrollada en nuestro laboratorio [37, 38].

3.6.2. Parte experimental

3.6.2.1. Instrumentación

Los espectros en el infrarrojo medio fueron obtenidos con un espectrofotómetro modelo Tensor 27 FT-MIR de Bruker (Bremen, Alemania) equipado con un detector de sulfato de trigliceno dopado con L-alanina deuterado (DLaTGS). El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

Las muestras de suelo fueron medidas en modo transmisión, utilizando una mini celda de gases de New Era Enterprises, Inc (Vineland, NJ, USA) con ventanas de ZnSe (25 mm diam x 2 mm grosor) y un volumen interno de 30mL, con un paso óptico de 10 cm. Esta celda dispone de dos válvulas de politetrafluoroetileno (PTFE) para controlar la entrada y la salida de la muestra a la misma celda. Para calentar a una temperatura controlada las muestras situadas en viales de vidrio de 10 mL se utilizó como sistema calentador de bloque metálico seco de JP Selecta (Barcelona, España) y un tubo de acero inoxidable, de un diámetro interno de 0.5 mm y térmicamente controlado, fue utilizado como línea de transferencia entre el vial y la celda de gases. Además, para transferir la muestra en fase vapor desde el vial a la celda se utilizó una mini bomba de vacío de KnF Laboport (Trenton, NJ, USA). Un esquema del montaje experimental empleado en este estudio se puede ver en la Figura 6.1.

Para la determinación de BTEX en las muestras de suelo por el método de referencia, se empleó un cromatógrafo de gases (GC) Modelo 7890A de Agilent (California, EE.UU.) equipado con un automuestreador de espacio de cabeza 7697A (HS), equipado con un detector de espectrometría de masas (MS).

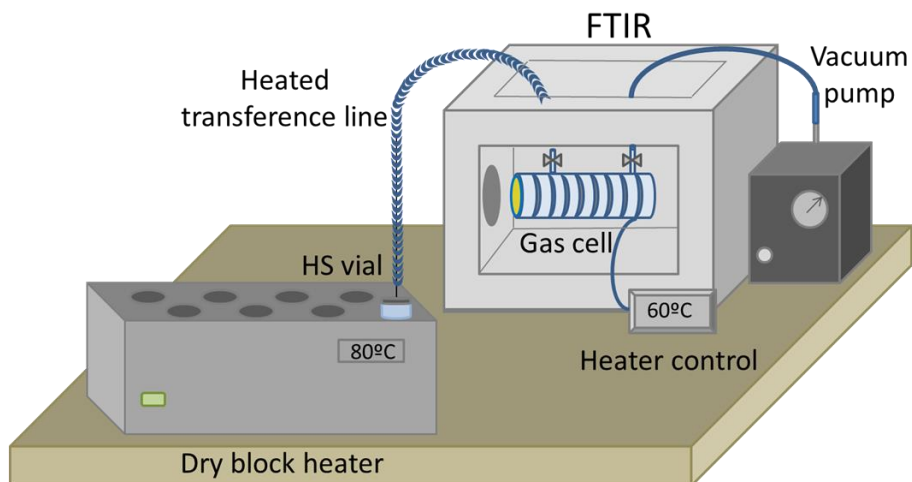


Figura 6.1. Montaje experimental para la determinación de muestras gaseosas mediante FT-MIR

3.6.2.2. Patrones y reactivos

Se emplearon patrones de benceno y tolueno de Scharlau (Barcelona, España) y, etilbenceno y orto, meta y para-xileno de Fluka (Buchs, Suiza) para preparar los estándares y muestras adicionadas utilizadas en este estudio. Para ello, se prepararon mezclas de BTEX utilizando microjeringas de diferentes volúmenes, y para la preparación de las muestras de suelo adicionadas se usó una microjeringa Hamilton 7001 (cero volumen muerto).

3.6.2.3. Generación de la fase vapor

Se pesaron 200 mg de muestra de suelo dentro de los viales de espacio de cabeza de 10 mL, que se cerraron herméticamente y se colocaron en el calentador de bloque metálico seco, a una temperatura constante durante un tiempo fijo. Los viales de muestra se conectaron a la celda de gases IR usando un tubo de acero inoxidable. Estas líneas de transferencia y la celda de gases se calentaron mediante una cinta térmica controlada por un termopar, para evitar condensaciones de vapor. Las válvulas de Teflón de la celda de gas IR de entrada y salida se conectaron al tubo que proviene del vial y a una mini bomba de vacío, respectivamente. Mientras se calienta el vial de muestra, manteniendo la llave de entrada a la celda cerrada, se hace el vacío dentro de la celda de gases a través de la válvula de salida y después de alcanzar un vacío constante, se cierra esta válvula y se apaga la mini bomba. Una vez que los compuestos se evaporan, la fase vapor se transfiere a la celda de gases por succión, abriendo la válvula de entrada y rompiendo el vacío. En este momento, se registran los espectros IR de la fase

gaseosa usando un *background* de la celda vacía registrada en las mismas condiciones instrumentales que las empleadas para los patrones y muestras. Todos los espectros se adquirieron en la región comprendida entre 4000 y 500 cm^{-1} , acumulando 50 barridos por espectro, a una resolución óptica de 4 cm^{-1} .

Para la calibración se introdujeron 200 mg de una matriz de suelo con distintos volúmenes de una mezcla estándar de BTEX y se midieron de la misma manera que las muestras. Los suelos utilizados para preparar las muestras adicionadas se midieron previamente para comprobar que no contenían BTEX.

3.6.2.4. Método de referencia: cromatografía de gases

Para el análisis de muestras de suelo contaminado por BTEX, 200 mg de suelo, recogido cerca de derrames de combustible, se introdujeron dentro de un vial HS de 10 ml y se añadió tolueno-d8 como patrón interno (IS). El vial se cerró herméticamente y se calentó a 200 °C durante 10 min en el horno del muestreador del HS. 1 ml de la fase de vapor generada se inyectó en el sistema GC en modo *split* (1:8) con una temperatura del inyector de 250 °C. Se utilizó una columna capilar Zebron ZB-5MS de Phenomenex (Torrance, CA, EEUU) (30 m x 0.32 mm id y espesor de película de 0.25 μm) empleando He como gas portador a un flujo constante de 1.3 ml min^{-1} . El programa de temperatura utilizado para el horno fue el siguiente: temperatura inicial 40 °C, que se mantuvo durante 8 min, luego se incrementó a una velocidad de calentamiento de 20 °C min^{-1} hasta alcanzar 200 °C y finalmente se mantuvo durante 2 min. Se empleó un detector de espectrometría de masas de simple cuadrupolo, que funcionaba en modo de impacto de electrones (EI) a 70 eV. Se seleccionó un rango de barrido de relación de masas de 40 a 400 m/z , para una adquisición en modo *full scan*. Las temperaturas de la fuente de iones y de la línea de transferencia fueron de 250 y 210 °C, respectivamente. Para el control de instrumentos y la adquisición de datos, se utilizó el software MSD de Agilent (California, EE.UU.)

Se prepararon curvas para calibración externa con suelos no contaminados con BTEX. Para ello se utilizaron mezclas estándar BTEX a diferentes niveles de concentración preparadas en n-hexadecano, usando tolueno-d8 como IS.

3.6.3. Resultados y discusión

6.3.1. Evaluación de las condiciones experimentales e instrumentales

En la etapa de generación de vapor de la muestra, se estudió la temperatura y el tiempo de calentamiento del vial para asegurar la eficacia y reproducibilidad de la

vaporización del analito. Debido a la diferente volatilidad de los compuestos considerados en este estudio, para evaluar las condiciones experimentales se seleccionaron el benceno y el p-xileno como moléculas testigo, ya que tienen temperaturas de ebullición de 80.1 y 138 °C, respectivamente. Como respuesta analítica se usaron las áreas de pico del espectro IR, comprendidas entre 679 y 667 cm^{-1} para el benceno y de 800 a 789 cm^{-1} para el p-xileno, en ambos casos corregidas con una línea de base establecida entre estos mismos valores. Estos intervalos de número de onda fueron seleccionados en lugar de considerar todas las bandas del espectro, con el fin de mejorar la relación lineal entre las señales y las concentraciones de ambos compuestos considerados. Se realizaron tres repeticiones para cada medida y se promediaron los valores del área del pico.

Para estudiar el efecto de la temperatura en la etapa de volatilización, se añadieron por separado 1 μl de benceno y de p-xileno puros a 200 mg de suelo dentro de los viales de vidrio HS y se calentaron a diferentes temperaturas durante un tiempo fijo de 6 minutos. La Figura 6.2a muestra el efecto del aumento de las temperaturas sobre los valores del área del pico del benceno y p-xileno. Como se puede ver, el aumento de la temperatura desde 50 hasta 80° C proporcionó un incremento en los valores del área del pico del benceno que permanecen prácticamente constantes para temperaturas más altas de 80 °C. El p-xileno presenta un comportamiento similar y, es por ello, que se seleccionó 80 °C para el calentamiento de la muestra con el fin de obtener la mejor relación señal/ruido.

Para la evaluación del tiempo de calentamiento, los suelos con 1 μl de benceno y p-xileno se calentaron a 80 °C durante diferentes tiempos. La Figura 6.2b pueden verse los valores medios de área de pico para los dos compuestos testigo, observándose que no se existen diferencias significativas al aumentar el tiempo de calentamiento, siendo escogidos 6 minutos para obtener medidas precisas y una volatilización completa de BTEX en presencia de suelo.

La línea de transferencia y la celda de gases también se calentaron para evitar la condensación del vapor de agua y la retención de los analitos debidos a los cambios de temperatura entre el vial de HS. Se estudió el efecto de la temperatura de calentamiento entre 50 y 80 °C y los resultados obtenidos indican que no es un parámetro relevante si la temperatura de trabajo es igual o mayor que el mínimo requerido para evitar la condensación de agua. En estas condiciones se seleccionó 60 °C para obtener una buena repetitividad de las medidas. Una vez estudiados los diferentes parámetros experimentales, para la determinación de BTEX en muestras de suelo se utilizaron unas condiciones consistentes en calentar las

muestras de suelo a una temperatura constante de 80 °C, durante 6 minutos, con una temperatura de la línea de transferencia y la celda de gases de 60 °C.

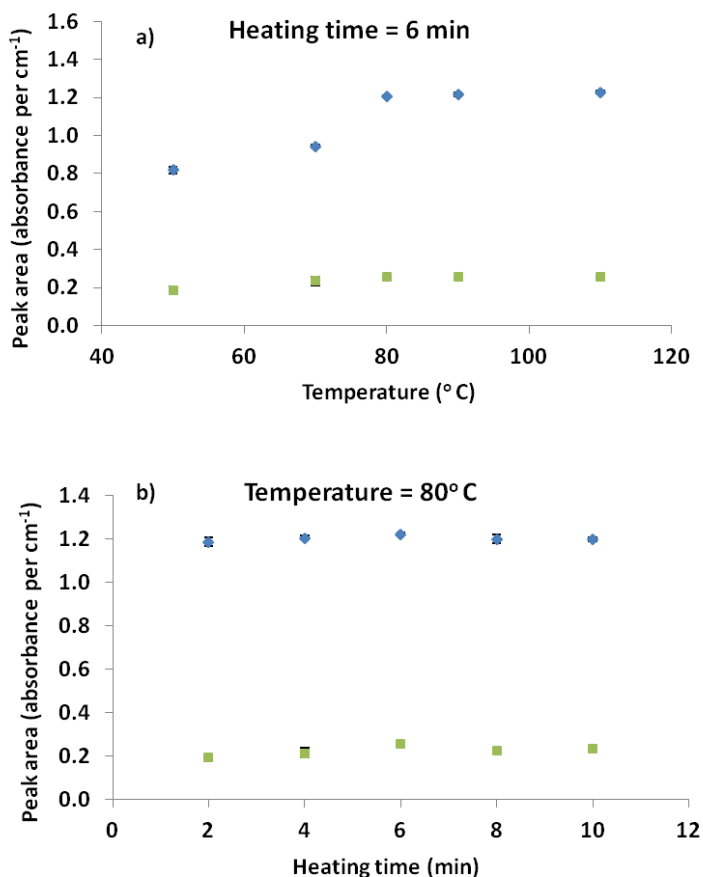


Figura 6.2. Efecto de la temperatura (a) y del tiempo de calentamiento (b) para la generación de fase vapor HS del benceno y para-xileno. Los valores medios del área del pico corresponden a tres repeticiones medidas entre 679 y 667 cm⁻¹ para el benceno (□) y entre 800 y 789 cm⁻¹ para el p-xileno (◇)

3.6.3.2. Espectro IR del benceno, tolueno, etilbenceno y xilenos en fase vapor

La Figura 6.3 muestra los espectros IR en fase vapor del benceno, tolueno, etilbenceno, xilenos y una mezcla de BTEX obtenidos en las condiciones experimentales anteriormente seleccionadas. Como puede apreciarse, los compuestos estudiados presentan bandas entre 3100 y 2800 cm⁻¹ características de las vibraciones de estiramiento de los enlaces =CH, -CH₂- y -CH₃. Alrededor de 1450 cm⁻¹, los compuestos alifáticos presentan bandas de absorción debido a la

deformación antisimétrica de los grupos $-\text{CH}_3$ y la deformación simétrica de los grupos $-\text{CH}_2-$. Adicionalmente en el intervalo de 1450 a 1600 cm^{-1} se pueden observar diferentes bandas debidas a las vibraciones de estiramiento de los enlaces carbono-carbono del anillo aromático. También se observan absorciones débiles en la región de 1660 a 2000 cm^{-1} debido a armónicos de los enlaces aromáticos. En la región de la huella dactilar, de 850 a 600 cm^{-1} , todos los compuestos mostraron bandas características de absorción con una superposición clara que dificulta la determinación directa de cada compuesto utilizando modelos convencionales de regresión monoparamétrica. Como alternativa a los métodos clásicos de calibración, el uso de la calibración multivariable permite cuantificar los analitos en presencia de interferencias con espectros similares, cuando se introduce información espectral del mismo tipo de muestras en la calibración del modelo.

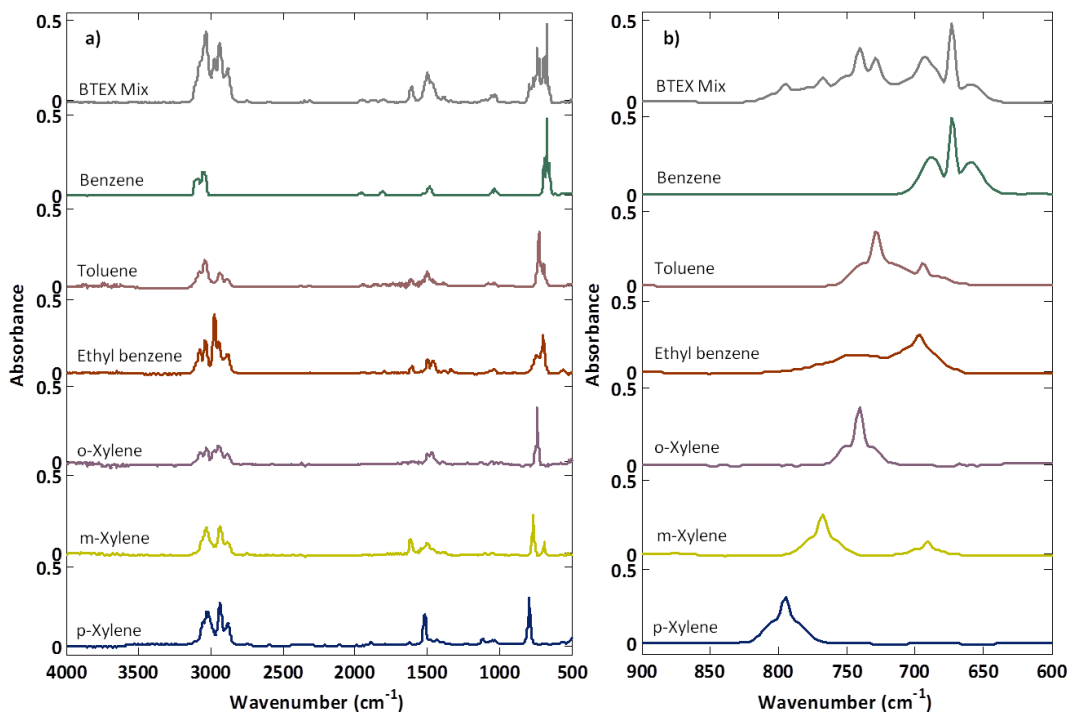


Figura 6.3. Espectros IR en fase vapor de BTEX y patrones puros, a) Espectro completo y b) región detallada de la huella dactilar.

Uno de los principales problemas para determinar el BTEX por espectrometría IR en fase de vapor es la contribución espectral de dos gases atmosféricos producidos durante el calentamiento del suelo: CO_2 y, especialmente, agua, ya que pueden dificultar la identificación exacta y la cuantificación correcta de los analitos. Con el fin de evitar esta contribución, a las señales obtenidas se aplicó un algoritmo de corrección desarrollado en nuestro grupo [37]. El algoritmo se basa en una primera estimación de la contribución de los gases atmosféricos seguido de la substracción simple y secuencial de los espectros del CO_2 y del vapor de H_2O , previamente medidos en las mismas condiciones instrumentales [38].

La Figura 6.4 muestra los espectros de una muestra de BTEX antes y después de la corrección del vapor de H_2O y CO_2 y, como se puede ver, la compensación de agua fue realizada en las regiones alrededor de 3500 cm^{-1} y $1900\text{--}1400\text{ cm}^{-1}$, mientras que la contribución del CO_2 fue eliminada correctamente en las regiones alrededor de 660 cm^{-1} y de 2350 cm^{-1} . Por lo tanto, puede comprobarse en los espectros resultantes que la corrección realizada evita, en gran medida, las interferencias asociadas a las contribuciones de los principales gases atmosféricos presentes y que absorben en el IR medio, siendo especialmente valiosa cuando se alcanzan señales bajas de analitos.

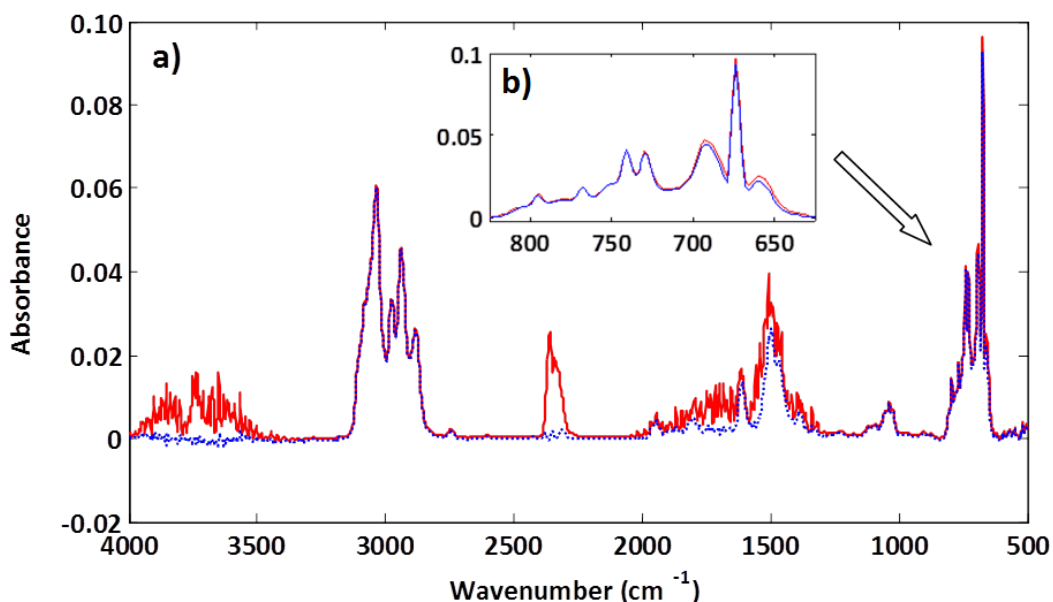


Figura 6.4. Espectros infrarrojos de una mezcla BTEX obtenida de una muestra de suelo a) antes (línea continua) y después (línea punteada) de la corrección de vapor del H_2O y del CO_2 . b) Detalle de los espectros en la región de huella dactilar

3.6.3.3. Modelos PLS

Para este estudio se utilizaron suelos naturales obtenidos de zonas de estacionamiento público y suelos de jardín. Las muestras se analizaron previamente para asegurarse que estaban libres de BTEX. Para obtener un calibrado adecuado, las muestras de suelo fueron adicionadas con cantidades conocidas de BTEX a diferentes proporciones y niveles de concentración y se utilizaron para construir los modelos multivariantes para cada compuesto estudiado. Los modelos fueron construidos utilizando 27 muestras para la calibración y validados mediante un conjunto independiente de validación formado por 12 muestras. Las principales características de los conjuntos de datos de calibración y validación se resumen en la Tabla 6.1.

Tabla 6.1. Intervalos de concentración en mg Kg^{-1} de benceno, tolueno, etilbenceno y xilenos en los conjuntos de datos de calibración y de validación

Analito	Conjunto	Muestras	Intervalo de concentración	Promedio \pm sd
Benceno	Calibración	27	0-3600	1290 \pm 1060
Tolueno			0-3300	1220 \pm 990
Etil benceno			0-2410	700 \pm 680
orto-Xileno			0-740	400 \pm 210
meta-Xileno			0-720	340 \pm 210
para-Xileno			0-720	390 \pm 200
Benceno	Validación	12	60-2660	940 \pm 990
Tolueno			50-3300	1710 \pm 1140
Etil benceno			160-2360	780 \pm 670
orto-Xileno			70-720	320 \pm 200
meta-Xileno			70-580	310 \pm 140
para-Xileno			130-580	280 \pm 140

Sd: desviación estándar

Se utilizaron distintos pre-procesamientos de los datos y distintas regiones espectrales con el fin de obtener los mejores modelos de predicción. Para ello se utilizó la validación cruzada, encontrándose los mejores resultados mediante el centrado en la media y el uso de la primera derivada, con un suavizado de 15 puntos, y, seleccionando toda la región del espectro ($4000\text{-}500\text{ cm}^{-1}$). En la Tabla

6.2 se resumen las principales figuras de mérito de los mejores modelos PLS obtenidos para la determinación de benceno, tolueno, etilbenceno, orto-, meta- y para-xileno.

Tabla 6.2. Parámetros de calibración y predicción de los modelos PLS empleados para la determinación de benceno, tolueno, etilbenceno, orto-, meta- y para-xileno en muestras de suelo considerando toda la región del espectro infrarrojo ($4000 - 500 \text{ cm}^{-1}$)

Analito	LV	Conjunto de Calibración		Conjunto de Validación		
		RMSEC	RMSECV	RMSEP	RRMSEP (%)	R ² Pred
Benceno	4	72	96	87	9.3	99.2
Tolueno	3	75	91	104	6.1	99.1
Etil benceno	3	93	114	76	9.8	99.5
o-Xileno	5	28	48	41	12.9	96.1
m- Xileno	5	18	29	21	6.9	98.4
p- Xileno	5	27	44	17	6.0	98.9

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

Los mejores resultados obtenidos para la determinación del benceno se lograron utilizando 4 LV's (explicando el 98.7% de la varianza total), para el tolueno y el etilbenceno se requirieron 3 LV's (que explicaron un 97.7% y un 98.1% de la varianza total, respectivamente), mientras que para orto-, meta- y para-xileno, en todos los casos los mejores resultados se obtuvieron utilizando 5 LV's (que explicaron un 99.1%, 99.1 y 99.0% de la varianza total, respectivamente). Por otro lado, los modelos quimiométricos seleccionados predicen el contenido de benceno, tolueno, etilbenceno, orto-, meta- y para-xileno con coeficientes de determinación para la predicción de 0.992, 0.991, 0.995, 0.961, 0.984 y 0.989, respectivamente, obteniéndose valores de RRMSEP comprendidos entre el 6 y el

13%, proporcionando así un método preciso para la determinación BTEX en muestras de suelo contaminado.

3.6.3.4. Análisis de muestras de suelo contaminado

Para evaluar la capacidad predictiva del método HS-FTIR propuesto para la determinación de BTEX en muestras reales se recogió un suelo cerca de un derrame de combustible y se analizó mediante el procedimiento propuesto.

Comparando los espectros de la muestra y de los patrones (ver Figura 6.5), existen algunas diferencias significativas, especialmente en la región entre 900 y 1400 cm^{-1} , probablemente debido a la presencia de otros hidrocarburos distintos de los del BTEX. En estas condiciones, y utilizando el modelo PLS para toda la región IR (4000 a 500 cm^{-1}), los resultados obtenidos presentaron valores anómalos, especialmente para el tolueno. Con el fin de evaluar la similitud entre los espectros de la muestra de suelo contaminado y las muestras de calibración incluidas en la construcción del modelo PLS, se realizó un PCA después de considerar la primera derivada y centrado en la media de los datos y usando la región espectral completa de 4000 a 500 cm^{-1} .

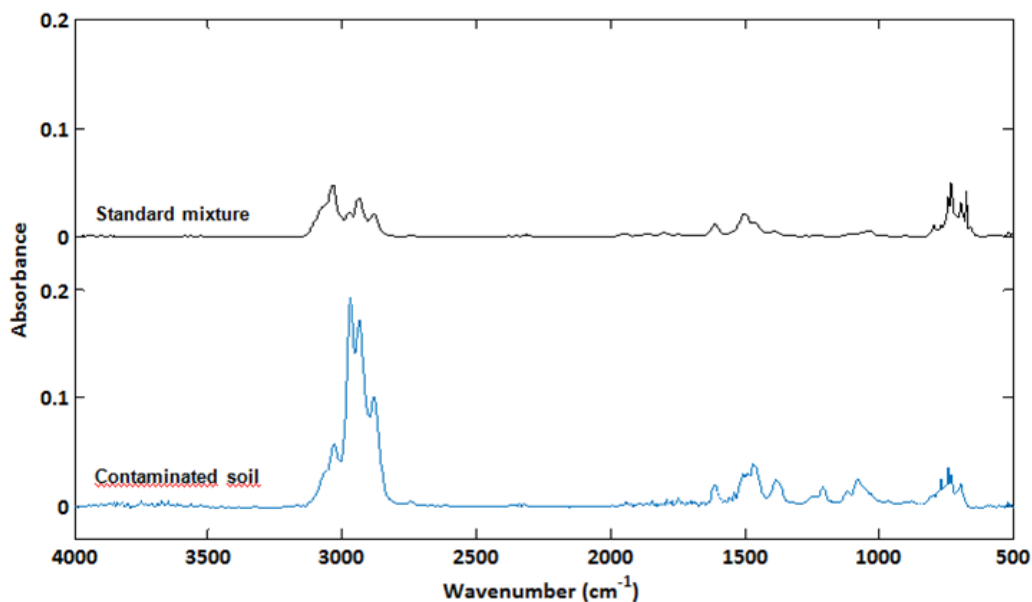


Figura 6.5. Espectro en fase vapor obtenido a partir de un suelo contaminado comparado con el correspondiente a una mezcla patrón BTEX.

En la Figura 6.6 se ilustra el gráfico de *scores* donde se representa PC2 frente al PC1, y en el que se puede observar que las réplicas de la muestra de suelo contaminado aparecen claramente separadas del grupo compuesto por los

patrones BTEX. Esto indica que el uso de este modelo PLS no sería recomendable para el análisis de esta muestra.

Para mejorar la capacidad de predicción del método HS-FTIR, se generaron modelos PLS adicionales considerando diferentes regiones espectrales: i) 913 a 567 cm^{-1} (una parte de la región de la huella dactilar de los compuestos BTEX), ii) 3258 a 2567 cm^{-1} , iii) 3258 a 3001 cm^{-1} (excluyendo así, en este caso, la región característica de las vibraciones de estiramiento $-\text{CH}_2-$ y $-\text{CH}_3$ inferiores a 3000 cm^{-1} , junto con varias combinaciones de las regiones seleccionadas y distintos pre-tratamientos de los datos.

Para seleccionar el modelo más adecuado, se analizaron replicas de una muestra de suelo contaminado mediante HS-FTIR utilizando los modelos PLS propuestos y los resultados se compararon con los obtenidos mediante un método de referencia HS-GC-MS. La Tabla 6.3 resume las principales figuras de mérito de los modelos PLS seleccionados de acuerdo con los mejores resultados obtenidos para cada analito a través de la validación cruzada y del análisis de un conjunto de muestras de validación. Como puede verse, para el benceno, tolueno, etilbenceno, orto-xileno y para-xileno, se puede recomendar el uso de las regiones comprendidas entre 3258 y 3001 y de 913 a 567 cm^{-1} , con un pre-tratamiento consistente en el centrado en la media y la primera derivada, mientras que para la determinación del meta-xileno es más adecuada la región de 913 a 567 cm^{-1} y el uso del centrado en la media y de la segunda derivada.

En todos los modelos PLS construidos se obtuvieron coeficientes de determinación elevados, encontrándose los valores dentro del intervalo comprendido entre 0.961 y 0.995, que pertenecen al orto-xileno y meta-xileno, respectivamente. Los errores relativos de predicción encontrados para el tolueno, meta- y para-xileno no superaron el 6%, mientras que el orto-xileno y etilbenceno superaron el 13%. Por lo tanto, los resultados obtenidos indican que el procedimiento desarrollado proporciona una herramienta adecuada para la determinación simultánea de compuestos BTEX en suelos contaminados de forma rápida y sin ningún pre-tratamiento químico de la muestra.

Tabla 6.3. Parámetros de calibración y predicción de los modelos PLS seleccionados para la determinación de benceno, tolueno, etilbenceno, orto-, meta- y para-xileno en muestras de suelo considerando diferentes regiones del espectro.

Análito	Región (cm ⁻¹)	pre- procesado	LV's	Conjunto de Calibración				Conjunto de Validación			
				RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	R ² pred	RRMSEP (%)	
Benceno			3	74	0.995	93	0.992	90	0.991	9.5	
Tolueno			2	82	0.993	94	0.991	105	0.991	6.1	
Etil benceno	3258-3001; 913-567	FD, MC	3	133	0.961	157	0.945	106	0.981	13.6	
o-Xileno			5	28	0.981	42	0.959	42	0.961	13.1	
p-Xileno			5	28	0.981	44	0.954	15	0.988	5.3	
m-Xileno	913-567	SD, MC	5	24	0.985	37	0.967	19	0.995	6.1	

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción; RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada; R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

En las condiciones anteriormente mencionadas, los gráficos PCA (véanse las figuras 6.6b y 6.6c) ponen de manifiesto que las réplicas de la muestra y de los patrones de BTEX aparecen distribuidas en un mismo grupo.

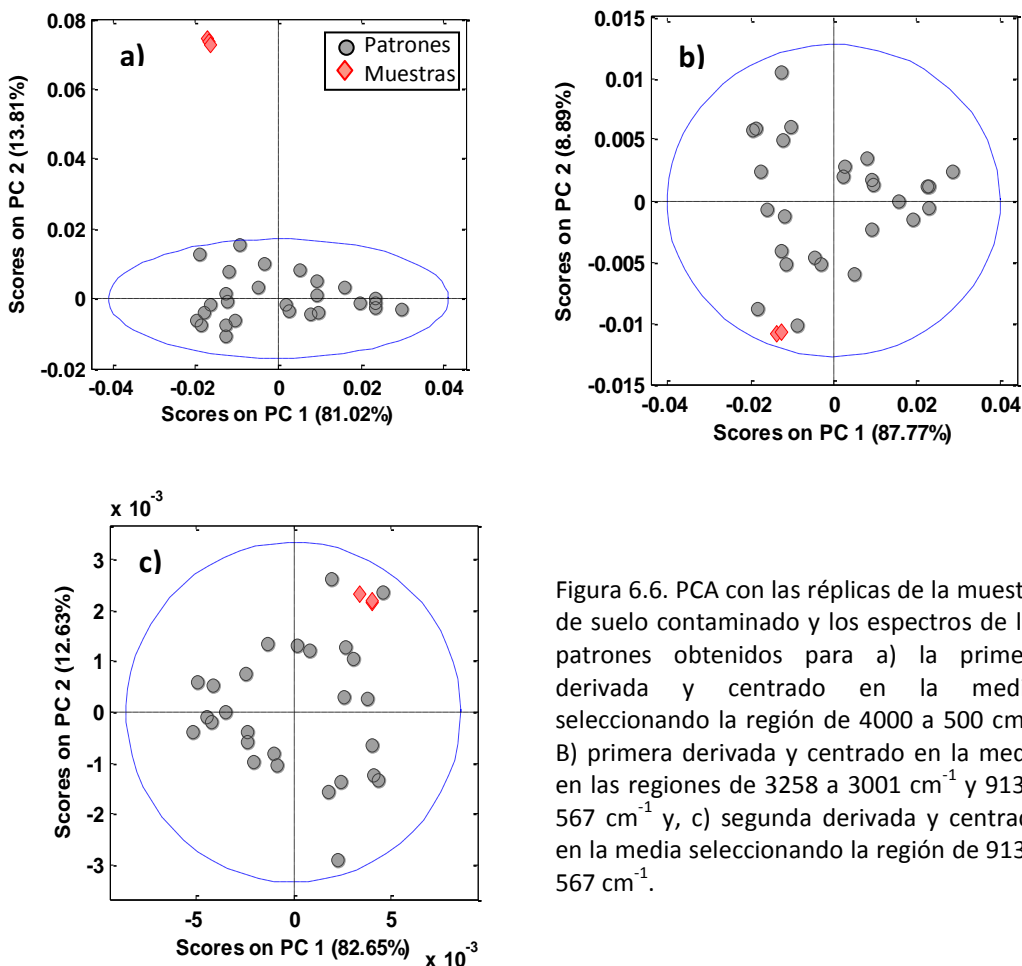


Figura 6.6. PCA con las réplicas de la muestra de suelo contaminado y los espectros de los patrones obtenidos para a) la primera derivada y centrado en la media, seleccionando la región de 4000 a 500 cm^{-1} ; B) primera derivada y centrado en la media en las regiones de 3258 a 3001 cm^{-1} y 913 a 567 cm^{-1} y, c) segunda derivada y centrado en la media seleccionando la región de 913 a 567 cm^{-1} .

Los resultados obtenidos para el análisis de una muestra de suelo contaminado usando los modelos PLS seleccionados y el método HS-GC-MS de referencia se muestran en la Tabla 6.4.

Tabla 6.4. Concentración de benceno, tolueno, etilbenceno y xilenos en una muestra de suelo contaminado obtenida por CG-MS y por el procedimiento IR propuesto.

Analito	Promedio \pm sd	
	HS-GC-MS	HS-FT-IR
Tolueno	650 \pm 30	470 \pm 20
Etil benceno	520 \pm 20	780 \pm 60
Orto-xileno	440 \pm 10	470 \pm 30
meta- más para-Xileno	1010 \pm 40	1100 \pm 20

sd: desviación estándar

Como puede verse, para la muestra de suelo analizada, el benceno no fue detectado por HS- FTIR y se consideró la suma de isómeros de meta- y para-xileno para comparar los resultados del método propuesto con los datos proporcionados por el método HS-GC-MS de referencia empleado. Con el fin de comprobar el rendimiento de los modelos de predicción seleccionados, se aplicó la prueba t de Student para comparar los resultados HS-FT-IR y HS-GC-MS, obteniéndose que ambos métodos son estadísticamente comparables para un nivel de confianza del 99%.

3.6.4. Conclusiones Capítulo 6

La espectroscopia infrarroja con generación de fase vapor, combinada con técnicas de calibración multivariante, proporciona una herramienta adecuada para la determinación simultánea de compuestos BTEX en suelos contaminados de forma rápida y sin ningún pre-tratamiento químico de la muestra.

El uso del espacio de cabeza, para la generación del vapor del analito, reduce los errores experimentales asociados con este paso y disminuye el tiempo de análisis a 7 minutos, incluyendo el tiempo de calentamiento y de medición para cada muestra. Además, el tiempo de análisis puede reducirse cuando varias muestras se calientan juntas. Por lo tanto, se ha demostrado que el HS-FTIR ofrece una técnica verde, ya que no requiere el empleo ni de reactivos ni de disolventes, adecuada para la determinación precisa de BTEX mediante calibración PLS, con coeficientes de variación entre 0,9-11.2, 0.4-4.7, 1.2-7.2, 0.9-9.3, 2.0- 7.0 y 0.4-8.0% para el benceno, tolueno, etilbenceno, orto-, meta- y para-xileno, respectivamente. Para

una muestra real de suelo contaminado, tomada de zona próxima a un área de derrame de combustible, los modelos PLS en la región 3258-3001 y 913-567 cm^{-1} y pre-tratamiento de primera derivada y centrado en la media fueron la mejor opción para determinar benceno, tolueno, etilbenceno, orto- y para-xileno, mientras que para la determinación del meta-xileno fue la segunda derivada y el centrado en la media como pre-tratamiento en la región de 913-567 cm^{-1} la que proporcionó los mejores resultados.

3.6.5. Referencias

- [1] The European Parliament and the Council of the European Union, Directive 2008/98/EC on wastes and repealing certain directives, Official Journal of the European Union, OJ L 312 (2008) 3–30.
- [2] Spanish Government, Royal decree of 14th January of 2005 (RD 9/2005), BOE, 15 (2005) 1833-1843. (Available at <https://www.boe.es/boe/dias/2005/01/18/pdfs/A01833-01843.pdf>)
- [3] Spanish Government, Law 22/2011, of 28th July, of waste and contaminated soils, BOE, 181 (2011) 85650-85705 (Available at <https://www.boe.es/boe/dias/2011/07/29/pdfs/BOE-A-2011-13046.pdf>)
- [4] F.A. Esteve-Turrillas, S. Armenta, S. Garrigues, A. Pastor, M. de la Guardia, Headspace–mass spectrometry determination of benzene, toluene and the mixture of ethylbenzene and xylene isomers in soil samples using chemometrics, *Anal. Chim. Acta*, 587(2007) 89–96.
- [5] M. Rosell, S. Lacorte, D. Barceló, Simultaneous determination of methyl tert-butyl ether, its degradation products and other gasoline additives in soil samples by closed-system purge-and-trap gas chromatography–mass spectrometry, *J. Chrom. A*, 1132(2006) 28–38.
- [6] J.L. Pérez-Pavón, S. Herrero-Martín, C. García-Pinto, B. Moreno-Cordero, Programmed temperature vaporizer based method for the sensitive determination of trihalomethanes and benzene, toluene, ethylbenzene and xylenes in soils, *J. Chrom. A*, 1216(2009) 6063–6070.
- [7] O. Ezquerro, G. Ortiz, B. Pons, M.T. Tena, Determination of benzene, toluene, ethylbenzene and xylenes in soils by multiple headspace solid-phase microextraction, *J. of Chrom. A*, 1035(2004) 17–22.
- [8] A. Abbas-Matin, P. Biparva, M. Gheshlaghi, K. Farhadi, A. Gheshlaghi, Environmental monitoring of complex hydrocarbon mixtures in water and soil samples after solid phase microextraction using PVC/MWCNTs nanocomposite fiber, *Chemosphere*, 93(2013) 1920–1926.
- [9] S. Wilbur, S. Bosch, Interaction Profile: Benzene, Toluene, Ethylbenzene, and Xylenes (BTEX), Agency for Toxic Substances and Disease Registry (ATSDR), 2004. (available at: <http://www.atsdr.cdc.gov/interactionprofiles/IP-btex/ip05.pdf>)

- [10] International Agency for Research on Cancer (IARC), Monographs on the Evaluation of Carcinogenic Risk to humans, Summaries & Evaluations, vol. 77, IARC Press, Geneva, 2000.
- [11] R.S. Dungan, Headspace solid-phase microextraction (HS-SPME) for the determination of benzene, toluene, ethylbenzene, and xylenes (BTEX) in foundry molding sand, *Anal. Lett.*, 38(2005) 2393-2405.
- [12] K. Liu, J. Quan, Y. Mu, Q. Zhang, J. Liu, Y. Gao, P. Chen, D. Zhao, H. Tian, Aircraft measurements of BTEX compounds around Beijing city, *Atmos. Environ.*, 73(2013) 11-15.
- [13] Carneiro, P.M.; Firmino, P.I.M.; Costa, M.C.; Lopes, A.C.; dos Santos, A.B. Multivariate optimization of headspace-GC for the determination of monoaromatic compounds (benzene, toluene, ethylbenzene and xylenes) in waters and wastewaters. *J. of Separation Science*, 37(2014) 265-271.
- [14] A. Sarafraz-Yazdi, A.H. Amiri, Z. Es'haghi, BTEX determination in water matrices using HF-LPME with gas chromatography-flame ionization detector. *Chemosphere*, 71(2008) 671-676.
- [15] A. Gaujac, E. Soares-Emídio, S. Navickiene, S.L. Costa-Ferreira, H. Silveira-Dóre, Multivariate optimization of a solid phase microextraction-headspace procedure for the determination of benzene, toluene, ethylbenzene and xylenes in effluent samples from a waste treatment plant, *J. Chrom. A*, 1203(2008) 99-104.
- [16] R. Kubinec, J. Adamuscin, H. Jurdáková, M. Foltin, I. Ostrovský, A. Kraus, L. Soják, Gas chromatographic determination of benzene, toluene, ethylbenzene and xylenes using flame ionization detector in water samples with direct aqueous injection up to 250 μ l, *J. Chrom. A*, 1084(2005) 90-94.
- [17] A. Przyjazny, J.M. Kokosa, Analytical characteristics of the determination of benzene, toluene, ethylbenzene and xylenes in water by headspace solvent microextraction, *J. Chrom. A*, 977(2002), 143-153.
- [18] J.W. Diehl, J.W. Finkbeiner, F.P. DiSanzo, Determination of benzene, toluene, ethylbenzene, and xylenes in gasolines by gas chromatography/deuterium isotope dilution Fourier transform infrared spectroscopy, *Anal. Chem.*, 65(1993) 2493-2496.
- [19] A. Bahrami, H. Mahjub, M. Sadeghian, F. Golbabei, Determination of benzene, toluene and xylene (BTX) concentrations in air using HPLC developed

method compared to gas chromatography, *International J. of Occupational Hygiene*, 3(2011) 12-17.

- [20] A. González, S. Garrigues, S. Armenta, M. de la Guardia, Headspace-liquid phase microextraction for attenuated total reflection infrared determination of volatile organic compounds at trace levels, *Anal. Chem.*, 87(2010) 3045–3051.
- [21] B. Pejčić, L. Boyd, M. Myers, A. Ross, Y. Raichlin, A. Katzir, R. Lu, B. Mizaikoff, Direct quantification of aromatic hydrocarbons in geochemical fluids with a mid-infrared attenuated total reflection sensor, *Organic Geochemistry*, 55(2013) 63–71.
- [22] B. Pejčić, M. Myers, A. Ross, Mid-infrared sensing of organic pollutants in aqueous environments, *Sensors*, 9(2009) 6232-6253.
- [23] M. Karłowatz, M. Kraft, B. Mizaikoff, Simultaneous quantitative determination of benzene, toluene, and xylenes in water using mid-infrared evanescent field spectroscopy, *Anal. Chem.*, 76(2004) 2643-2648.
- [24] K.M.G. Lima, I.M. Raimundo Jr., M.F. Pimentel, Simultaneous determination of BTX and total hydrocarbons in water employing near infrared spectroscopy and multivariate calibration, *Sensors and Actuators B: Chemical*, 160(2011) 691– 697.
- [25] J. Moros, S. Garrigues, M. de la Guardia, Vibrational spectroscopy provides a green tool for multi-component analysis, *TrAC-Trends Anal. Chem.*, 29(2010) 579-591.
- [26] M. Khanmohammadi, A. Bagheri Garmarudi, M. de la Guardia, Characterization of petroleum based products by infrared spectroscopy and chemometrics, *TrAC-Trends Anal. Chem.*, 35(2012) 135-149.
- [27] C. C. Felício, L.P. Brás, J.A. Lopes, L. Cabrita, J.C. Menezes, Comparison of PLS algorithms in gasoline and gas oil parameter monitoring with MIR and NIR, *Chem. Intell. Lab. Syst.*, 78(2005) 74 – 80.
- [28] E. López-Anreus, S. Garrigues, M. de la Guardia, Vapour generation-Fourier transform infrared spectrometric determination of benzene, toluene and methyl tert-butyl ether in gasolines, *Anal. Chim. Acta*, 333(1996) 157-165.
- [29] E. López-Anreus, S. Garrigues, M. de la Guardia, Vapour generation Fourier transform infrared spectrometry. A new analytical technique, *Anal. Chim. Acta*, 308(1995) 28-35.

- [30] S. Armenta, F.A. Esteve-Turrillas, G. Quintás, S. Garrigues, A. Pastor, M. de la Guardia, Development of a simple and low cost device for vapour phase Fourier Transform Infrared spectrometry determination of ethanol in mouthwashes, *Anal. Chim. Acta*, 569(2006) 238–243.
- [31] S. Armenta, N.M.M. Coelho, R. Roda, S. Garrigues, M. de la Guardia, Seafood freshness determination through vapour phase Fourier transform infrared spectroscopy, *Anal. Chim. Acta*, 580(2006) 216–222.
- [32] M.P. Gómez-Carracedo, J.M. Andrade, M. Calviño, E. Fernández, D. Prada, S. Muniategui, Multivariate prediction of eight kerosene properties employing vapour-phase mid-infrared spectrometry, *Fuel*, 82(2003) 1211–1218.
- [33] S. Armenta, S. Garrigues, M. de la Guardia, On-line vapor-phase generation combined with Fourier transform infrared spectrometry, *TrAC-Trends Anal. Chem.*, 27(2008) 15-23.
- [34] Bruker Corporation. Atmospheric Compensation Command. Reference Manual for OPUS Spectroscopy Software, Ettlingen, Germany: Bruker Optik, 2007.
- [35] R.A. Hault, Suppression of undesired components in measured spectra. European Patent EP0982582B1. Filed 1998. Issued 2005.
- [36] Perkin Elmer. Atmospheric vapor compensation on Spectrum Two and Frontier FTIR spectrometers, Technical Note, 2010. (Available at http://www.perkinelmer.com/CMSResources/Images/44-74824TCH_AVCIR.pdf)
- [37] G. Quintás, J. Kuligowski, M. de la Guardia, S. Garrigues, Método de compensación de gases y/o vapores ambientales en espectros de infrarrojos. Spanish patent ES2440893 A2. Failed: 2012. Issued: 2014.
- [38] D. Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M. de la Guardia, Atmospheric Compensation in Fourier transform infrared spectra of clinical samples, *Appl. Spectrosc.*, 67(2013) 1339-1342.

3.7. CAPÍTULO 7:
Correcciones de vapor en espectroscopia infrarroja

3.7. Capítulo.7. Correcciones de vapor en espectroscopia infrarroja

3.7.1. Introducción

La espectroscopia infrarroja por transformada de Fourier (FTIR) basada en la generación de fase vapor ha demostrado ser una buena herramienta para la determinación cualitativa y cuantitativa de compuestos volátiles como amonio y nitrógeno orgánico en productos agroquímicos [1], penicilamina en cápsulas [2], benceno, tolueno y metil terc-butyl éter en la gasolina [3] o contenido de etanol en diferentes matrices como el cloroformo, bebidas alcohólicas, sangre y enjuagues bucales [4-7]. Sin embargo, la presencia en la atmósfera de los laboratorios de moléculas gaseosas que absorben en el IR mediopodría producir interferencias en los espectros de la muestra. Para evitar este tipo de interferencias espectrales se utilizan diferentes estrategias como: i) mantener cerrado el interferómetro y la cámara de muestra o dispositivo de muestreo o ii) purgar el sistema con aire seco o N_2 [8]. En muchos casos, es difícil eliminar la contribución de la atmósfera del laboratorio a las mediciones de la muestra, por lo que, todavía existe interés en la eliminación de los restantes compuestos interferentes [9, 10]. Nuestro grupo desarrolló un método de corrección para la compensación de la contribución espectral de los principales componentes atmosféricos empleando un algoritmo, basado en la medición de un espectro que contenía la molécula a eliminar, seguida de la estimación automática de la contribución espectral de ésta sobre la muestra a corregir, compensándose así la contribución mediante una simple substracción espectral [11]. Esta estrategia de compensación se aplicó para la corrección de la contribución espectral del vapor de H_2O y del CO_2 en los espectros ATR-FTIR de muestras clínicas [12] y para la determinación directa de benceno, tolueno, etilbenceno y xilenos en suelos por espectroscopia infrarroja con espacio de cabeza [13], proporcionando muy buenos resultados. Sin embargo, el método propuesto no se limita a la compensación de los gases atmosféricos de CO_2 y H_2O ya que también se podría realizar la corrección de cualquier otro gas presente en el intervalo de frecuencias en el que se esté trabajando.

La acetona es una de las sustancias que pueden interferir en la determinación de compuestos volátiles por espectroscopia infrarroja y puede ser falsamente identificada y medida como etanol por algunos alcoholímetros [14, 15], ya que el vapor de acetona puede ser detectado en el aliento humano de personas con diabetes [16] o que presentan ataques de epilepsia, siendo la determinación de acetona en la respiración un control diagnóstico en pacientes diabéticos [17]. Otra situación en la que la acetona puede producir interferencia espectral es en la determinación de compuestos aromáticos, como benceno, tolueno y xilenos, cuando la acetona está presente en la atmósfera de laboratorio. Los compuestos

carbonílicos y aromáticos son los principales contaminantes orgánicos en el aire urbano [18], por lo que la eliminación de la contribución de la acetona del espectro de la muestra FTIR podría ser muy interesante especialmente cuando las determinaciones se realizan por espectroscopia infrarroja en fase de vapor.

Por tanto, el objetivo de este estudio es la evaluación de un método flexible y automatizado para evitar las interferencias espectrales de los vapores orgánicos en los espectros FTIR medidos en fase de vapor, tomando acetona como compuesto de referencia. Para ello se utilizó una metodología verde, que combina el uso del muestreo de espacio de cabeza (HS) con la espectroscopia de infrarrojo por transformada de Fourier (FTIR) [13] para el análisis de muestras que contienen benceno, tolueno y para-xileno y/o etanol.

Para simular las contribuciones del gas atmosférico, se introdujo vapor de acetona dentro de la cámara de medida para ser corregido más tarde. Como evaluación de la capacidad del método para la corrección, se realizó una comparación de las similitudes espectrales de los registros de la muestra en ausencia de interferencia y con los espectros corregidos. Como parámetros de evaluación se utilizaron: i) el coeficiente de correlación (R), ii) la transformación de Fisher Z y iii) el nivel de ruido espectral.

3.7.2. Parte experimental

3.7.2.1. Instrumentación

Las medidas IR de transmitancia en fase vapor fueron llevadas a cabo utilizando un espectrómetro FTIR modelo IFS 66 de Bruker (Bremen, Alemania) equipado con un detector de sulfato de triglicina deuterada dopado de L-alanina (DLaTGS), y mediante el uso de una mini celda de gas de New Era Enterprises, Inc. (Vineland, NJ, EE.UU.) con ventanas de ZnSe (25 mm de diámetro x 2 mm de espesor) y un volumen interno de 30 ml, con un paso óptico de 10 cm. Los espectros se adquirieron en el intervalo comprendido entre 4000 y 500 cm^{-1} , acumulando 50 barridos por espectro, con una resolución 4 cm^{-1} .

3.7.2.2. Medidas de las muestras

Para la generación de fase de vapor y medida por espectroscopia FTIR, las muestras de suelo se introdujeron dentro de viales de espacio de cabeza de 10 ml, se cerraron herméticamente, se colocaron a calentar en bloque metálico seco a temperatura constante. El vial se conectó a la celda IR de gases mediante una línea de transferencia constituida por un tubo capilar metálico. Una vez generada la

fase de vapor, los analitos volátiles se transfirieron desde el vial a la celda de gases IR y se adquirieron los espectros FTIR en fase de vapor, mediante medidas de transmisión, usando como *background* la celda vacía purgada con aire seco. El montaje experimental es el mismo que el descrito en el Capítulo 6. El procedimiento experimental consiste en tres pasos: i) adquisición de espectros de muestra sin interferencia de acetona, ii) adquisición de espectros de las muestras con diferentes niveles de acetona en la cámara de muestras y iii) adquisición de espectros de la misma muestra en ausencia de acetona interferente. Estas medidas permitirán conocer las máximas similitudes entre espectros de la muestra obtenidos sin interferencia de acetona y los espectros corregidos. En todos los casos se adquirieron tres espectros de cada muestra para ser corregidos posteriormente.

3.7.2.3. Espectros de referencia y corrección de muestras

Para el desarrollo del método de corrección de los espectros es necesaria la adquisición de un espectro de referencia del compuesto interferente, en las mismas condiciones instrumentales y experimentales que las muestras. En nuestro caso, se creó una contaminación atmosférica de acetona introduciendo vapores de acetona en el compartimiento de muestras del espectrofotómetro; por lo que se adquirió el espectro de acetona además de los espectros interferentes de los gases atmosféricos (CO_2 y H_2O), para la eliminación de sus contribuciones en los espectros. Para la medición de los espectros de referencia en fase gaseosa de acetona, H_2O y CO_2 , se utilizó como fondo la señal de la celda de gases vacía purgada con aire seco.

Los espectros de la acetona se obtuvieron colocando pequeñas alícuotas de acetona en el compartimiento de la muestra, mientras que los espectros de H_2O y CO_2 se obtuvieron a partir de aire húmedo manteniendo abierto el compartimiento de muestra. La Figura 7.1 muestra los pasos de la estrategia de compensación desarrollada para evitar las interferencias de la acetona.

Como puede verse, debido a la presencia de CO_2 y H_2O , la compensación de estos dos compuestos debe realizarse previamente a la de otros vapores en todos los espectros, tanto de referencia como de muestras. A continuación, y como se muestra en la *Ecuación 7.1*, se seleccionan dos números de onda en el espectro de referencia del acetona para calcular la Absorbancia (RA) a partir del máximo y mínimo de absorbancia del gas interferente en regiones sin ninguna contribución de bandas en las muestras de benceno, tolueno y para-xileno o etanol en nuestro caso.

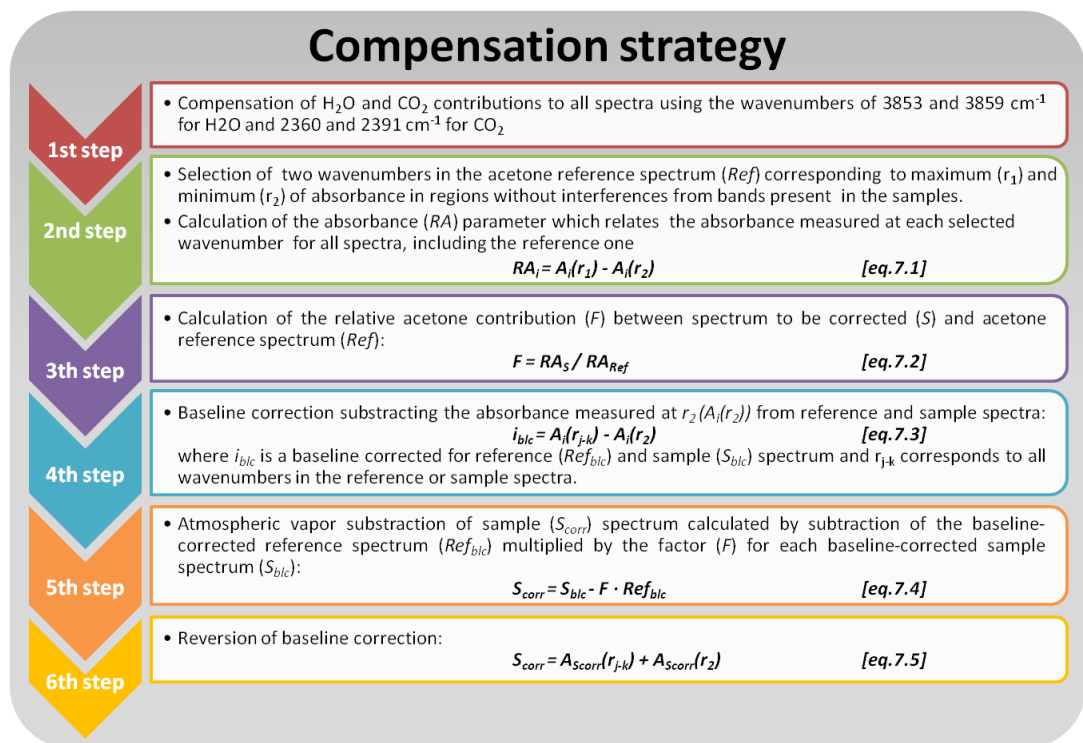


Figura 7.1. Esquema de las etapas de la estrategia de compensación

Para la corrección de H₂O, se seleccionó la absorbancia relativa entre 3853 y 3859 cm⁻¹ y, finalmente, los valores entre 4000 y 3766, 3569 y 2412, 1809 y 1249 y, 605 y 500 cm⁻¹ se reemplazaron por ceros. La contribución espectral del CO₂ presente en el espectro de referencia de H₂O se eliminó reemplazando por ceros los valores de absorbancia entre 2948 y 2123 y 1186 y 500 cm⁻¹.

En las determinaciones FTIR en fase de vapor, todos los espectros contienen vapores de CO₂ y H₂O debido al calentamiento de las muestras y, por este motivo, la contribución espectral de H₂O y CO₂ presentes en los espectros de acetona y de la muestra se corrigieron usando los pares de números de referencia 3853 y 3859 cm⁻¹ para H₂O y, 2360 y 2391 cm⁻¹ para CO₂. Estas correcciones en las muestras deben hacerse antes de la corrección de vapor de acetona para evaluar cuantitativamente la calidad del algoritmo de compensación de acetona. Después de la compensación atmosférica de H₂O y CO₂, el vapor de acetona presente en las muestras se corrige con los pares de números de onda 1737 y 1859 cm⁻¹ (*r*₁ y *r*₂).

3.7.3. Resultados y discusión

3.7.3.1. Espectro de FTIR en fase vapor en presencia del vapor de acetona

La Figura 7.2A muestra los espectros FTIR en fase vapor de la acetona y etanol y la Figura 7.2B los espectros de acetona y benceno, tolueno y para-xileno en la región entre 4000 y 500 cm^{-1} . La acetona presenta bandas en las regiones entre 3100 y 2800 cm^{-1} características de las vibraciones de estiramiento de los enlaces $-\text{CH}_3$ y una serie de bandas en el rango de 2200 a 500 cm^{-1} debido al estiramiento del grupo $\text{C}=\text{O}$, cercana a 1740 cm^{-1} , y de las flexiones de los enlaces $-\text{CH}_3$, alrededor de 1440 y 1365 cm^{-1} , entre otras. Una contaminación por acetona daría lugar a importantes contribuciones espectrales en las principales regiones del espectro de la muestra, creando problemas para determinar etanol y benceno, tolueno y para-xileno por espectroscopia FTIR en fase vapor. A partir de los espectros mencionados, se seleccionaron números de onda de 1737 debido a que la acetona tiene un máximo de absorbancia y 1895 cm^{-1} debido a que en esta zona las muestras no absorben.

Las Figuras 7.3A y 7.3B muestran los espectros FTIR en fase vapor del etanol y benceno, tolueno y para-xileno en la región entre 4000 y 500 cm^{-1} , antes y después de la corrección del vapor de acetona. A partir de una comparación visual, se puede ver que la metodología desarrollada aparentemente proporciona una compensación correcta del vapor de acetona.

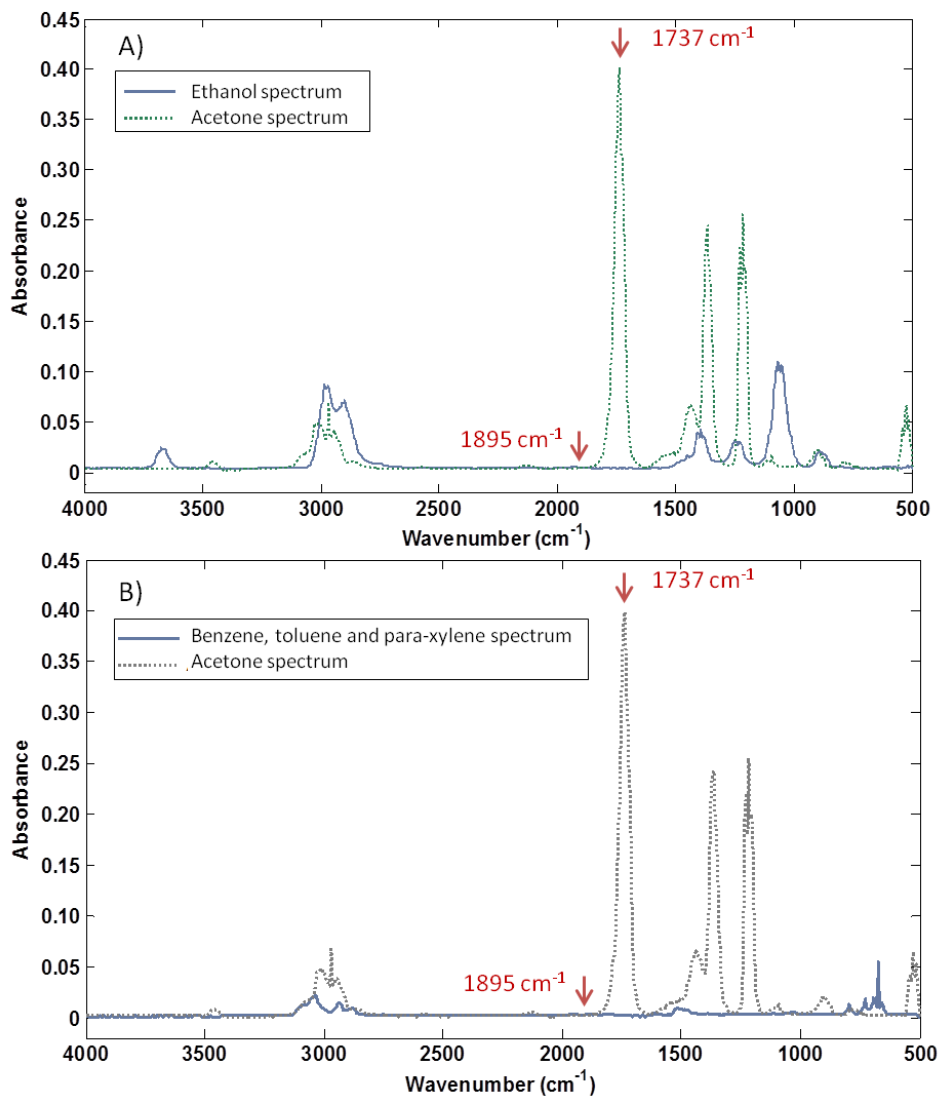


Figura 7.2. Espectro FTIR en fase vapor de la acetona junto al del etanol (2A) y los del benceno, tolueno y para-xileno (2B), en la región entre 4000 y 500 cm⁻¹

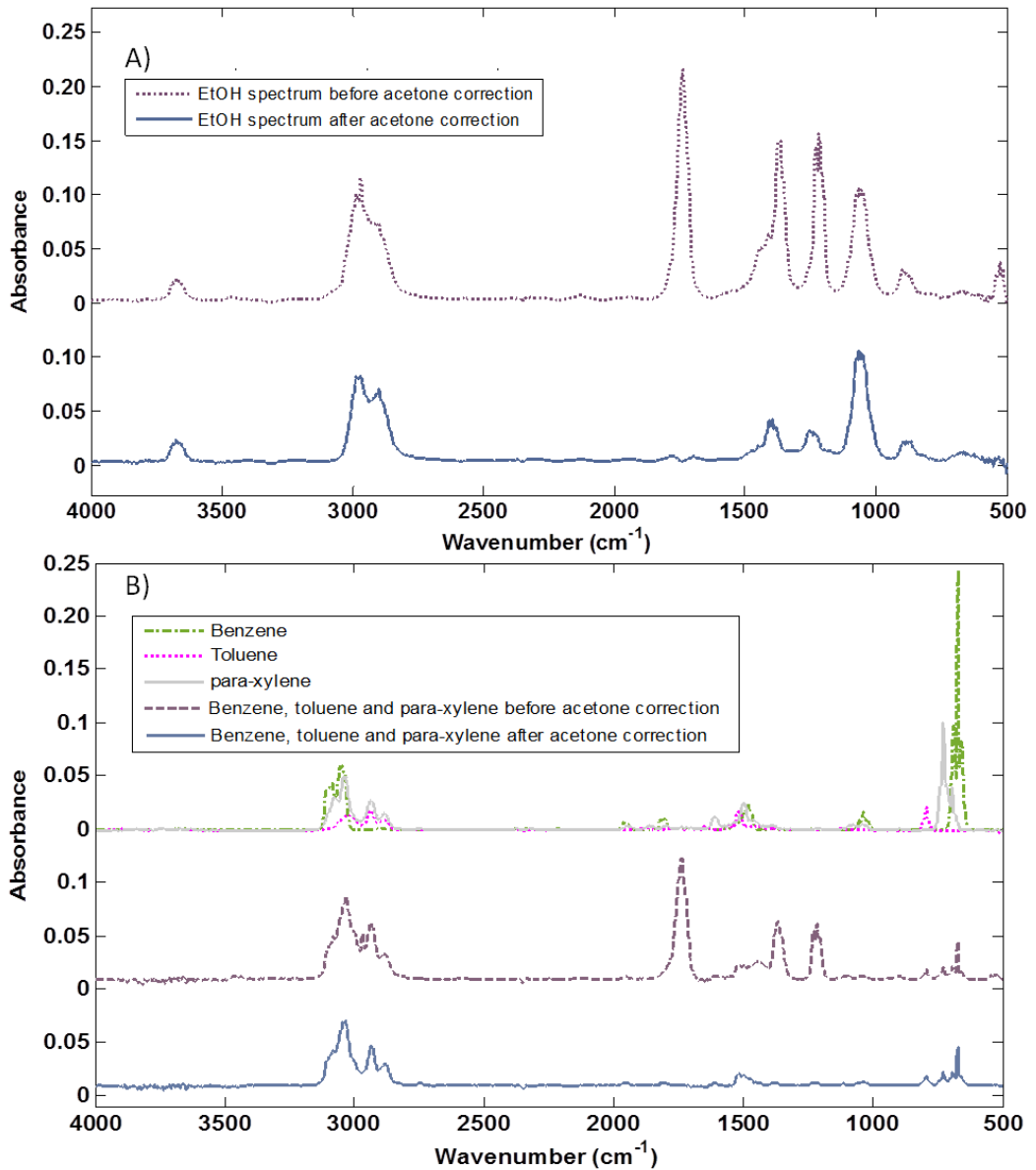


Figura 3. Espectro FTIR del a) etanol y b) muestra de benceno, tolueno y para-xileno obtenidos en presencia de vapor de acetona, antes y después de la corrección de vapor de acetona utilizando el método propuesto. Nota: los espectros se desplazaron sobre el eje y.

3.7.3.2. Evaluación de la corrección de vapor de la acetona

Para comparar similitudes espectrales, se utilizaron tres parámetros: el coeficiente de correlación (R), la transformación de Fisher Z [19] y el nivel de ruido de los espectros, antes y después de la corrección.

El coeficiente de correlación entre dos espectros se define como la relación entre la covarianza y el producto de sus desviaciones estándar. Por lo tanto, R indica el grado de dependencia de dos espectros, de modo que un valor de R igual a 1 indica que se trata de espectros idénticos.

Los valores de Z se calcularon a partir de los de R mediante la siguiente expresión:

$$Z = (1/2) * \ln (1 + R) / (1-R) \qquad \text{Ecuación 7.6}$$

siendo el valor de Z más alto cuanto mayor es la similitud espectral.

Los valores de R y Z se calcularon comparando diferentes regiones espectrales entre los espectros de fase de vapor obtenidos en ausencia de interferencia de la muestra y: i) el espectro final obtenido al final análisis, ii) los espectros de muestra con interferencia de la acetona y iii) espectro de la muestra después de la corrección de la acetona. El ruido espectral se calculó también en diferentes regiones espectrales, donde absorbe la acetona pero no los analitos, mediante el software Matlab para evaluar el rendimiento del método de corrección aplicado.

En las Tablas 7.1 y 7.2 pueden verse los valores de los tres parámetros mencionados anteriormente para evaluar la corrección de los espectros en las muestras de etanol y del benceno, tolueno y para-xileno, respectivamente. La evaluación de la corrección de los espectros se realizó en tres regiones diferentes: i) todo espectro, entre 4000 y 500 cm^{-1} , ii) vibraciones de estiramiento características de los enlaces =CH, -CH₂ y -CH₃, entre 3200 y 2623 cm^{-1} y iii) entre 1600 y 500 cm^{-1} . A partir de los datos reportados, se puede observar que los mejores resultados se obtuvieron en las correcciones de etanol y benceno, tolueno y para-xileno en la región entre 3200 y 2623 cm^{-1} con altos valores de R y Z.

Tabla 7.1. Parámetros de rendimiento cuantitativos de los métodos de corrección utilizados para evitar la interferencia de acetona en muestras de etanol en diferentes regiones espectrales

	4000-500 cm^{-1}		3200-2623 cm^{-1}		1600-500 cm^{-1}		Ruido
	R	Z	R	Z	R	Z	
Muestra Final	0.99975 ± 0.00005	4.5 ± 0.1	0.9999709 ± 0.0000016	5.57 ± 0.03	0.99971 ± 0.00009	4.44 ± 0.14	1.65 ± 0.03
Antes de la corrección	0.82 ± 0.07	1.2 ± 0.2	0.994 ± 0.003	3.0 ± 0.3	0.84 ± 0.07	1.2 ± 0.2	19 ± 6
Después de la corrección	0.998 ± 0.003	4.4 ± 0.3	0.99993 ± 0.00008	5.4 ± 0.6	0.998 ± 0.003	3.8 ± 0.8	1.7 ± 0.6

R: coeficiente de correlación

Z: Transformación de Fisher Z

El nivel de ruido se estableció en términos de desviación estándar en 3508-3412, 3167-3070, 2172-2085, 1859-1651, 1632-1537, 1344-1298, 557-500 cm^{-1} . El ruido de la muestra inicial fue 1.73.

Tabla 7.2.: Parámetros de rendimiento cuantitativos de los métodos de corrección utilizados para evitar la interferencia de acetona en muestras de benceno, tolueno y para-xileno en diferentes regiones espectrales

	4000-500 cm ⁻¹		3200-2623 cm ⁻¹		1600-500 cm ⁻¹		
	R	Z	R	Z	R	Z	
Muestra Final	0.996 ± 0.003	3.2 ± 0.4	0.99991 ± 0.00003	5.04 ± 0.15	0.9978 ± 0.0003	3.41 ± 0.07	0.62 ± 0.12
Antes de la corrección	0.44 ± 0.02	0.48 ± 0.03	0.967 ± 0.004	2.05 ± 0.06	0.35 ± 0.02	0.36 ± 0.03	16.0 ± 1.3
Después de la corrección	0.991 ± 0.003	2.71 ± 0.17	0.99984 ± 0.00008	4.7 ± 0.2	0.993 ± 0.003	2.85 ± 0.19	0.63 ± 0.19

R: coeficiente de correlación

Z: Transformación de Fisher Z

El nivel de ruido se estableció en términos de desviación estándar en 3508-3412, 3167-3070, 2172-2085, 1859-1651, 1632-1537, 1344-1298, 557-500 cm⁻¹. El ruido de la muestra inicial fue 0.68.

Para comprobar la adecuación y validez del método de corrección de la acetona, se utilizó la prueba t de Student para comparar el valor medio de los espectros de la muestra sin interferencia de acetona y después de la corrección de la interferencia (ver Tabla 7.3). Se empleó una prueba F con anterioridad para evaluar la homogeneidad entre las dos varianzas.

Tabla 7.3. Prueba F y t de Student utilizadas para comparar los valores medios entre el espectro de la muestra sin interferencias y el espectro de la muestra después de la corrección de la interferencia de acetona en tres regiones diferentes para etanol y benceno, tolueno y para-xileno muestras

		4000-500 cm^{-1}		3200-2623 cm^{-1}		1600-500 cm^{-1}	
		exp	crit	exp	crit	exp	crit
Etanol	F	2659	19	2555	19	1341	19
	t	1	4.3	0.83	4.3	1.04	4.3
Benceno, tolueno y para- xileno	F	1.21	19	8.22	19	79.43	19
	t	1.97	2.77	1.64	2.77	3.01	4.3

F: Prueba F de valores experimentales (exp) y críticos (crit). $F_{\text{exp}} > F_{\text{crit}}$, indicó que los valores de varianza eran diferentes y $F_{\text{exp}} < F_{\text{crit}}$ correspondió a una varianza idéntica.

T: valores obtenidos en la prueba experimental de t de Student (exp) y críticos (crit). $t_{\text{exp}} > t_{\text{crit}}$ indicó que los dos valores medios difieren significativamente a un nivel de probabilidad del 95% y $t_{\text{exp}} < t_{\text{crit}}$ muestra que los dos valores medios no difieren significativamente a este grado de probabilidad.

En el caso de la muestra de etanol, se realizó la prueba t de Student, suponiendo varianzas no homogéneas, debido al hecho de que F_{exp} fue mayor que F_{crit} , mientras que para muestras de benceno, tolueno y para-xileno las varianzas resultaron ser iguales. Como se puede ver, en todos los casos, los dos valores medios no difieren significativamente para el nivel de confianza del 95%, siendo $t_{\text{exp}} < t_{\text{crit}}$, confirmándose así la alta similitud espectral en todas las regiones obtenida entre los espectros en ausencia de interferencia y después de la corrección.

Por otra parte, el rendimiento de la corrección del espectro de acetona se evaluó midiendo el ruido en términos de desviación estándar en las regiones influenciadas por este compuesto pero donde las muestras no absorben. Para el etanol, las regiones seleccionadas fueron 3508-3412, 3167-3070, 2172-2085, 1859-1651, 1632-1537, 1344-1298 y 557-500 cm^{-1} . Para el benceno, tolueno y para-xileno, las regiones consideradas seleccionadas fueron 3508-3423, 2172-2087, 1761-1691, 1358-1305, 1263-1167, 943-852 y 559-500 cm^{-1} . Como puede verse en los espectros de las Figuras 7.2A y 7.2B, las regiones seleccionadas proporcionan un nivel de ruido creado por la presencia de vapor de acetona en la zona de medición. Tanto para el espectro del etanol como el de la mezcla de benceno, tolueno y para-xileno, el nivel de ruido después de la corrección de interferencia de acetona por el método propuesto fue similar al obtenido sin interferencias de acetona (véanse Tablas 7.1 y 7.2), evidenciando así una buena adecuación de la estrategia propuesta para la corrección de la interferencia espectral debida a la presencia de contaminaciones de acetona.

3.7.4. Conclusiones Capítulo 7

El método propuesto para la compensación de acetona aplicado a los espectros de muestras de benceno, tolueno y para-xileno y etanol en fase vapor proporciona una buena correlación entre los espectros en ausencia de interferencia y los espectros corregidos. Los valores de R son cercanos a 1, siendo los valores de Z después de la corrección entre 3.8 y 5.4 para el etanol y entre 2.71 y 4.7 para la mezcla de benceno, tolueno y para-xileno, dependiendo de la región espectral. Los valores de la t de Student obtenidos para la comparación entre los espectros corregidos y los mismos con interferencia de acetona, no difieren estadísticamente de una manera significativa para el nivel de confianza del 95%, confirmándose la alta similitud espectral en todas las regiones estudiadas de los espectros corregidos.

Por lo tanto, este estudio demostró que el método de compensación propuesto para las interferencias de gas y vapor no se limita únicamente a la compensación del CO_2 y vapor de H_2O presentes en el aire, sino que puede aplicarse para la corrección de cualquier otro gas presente en la celda de medida o en la trayectoria óptica del haz de radiación desde la fuente hasta el detector.

3.7.5. Referencias Capítulo 7

- [1] N.M.M. Coelho, S. Garrigues, M. de la Guardia, Determination of nitrogen in hydrolyzed protein formulations by continuous vapour phase FTIR, *Talanta*, 68(2006) 836–841.
- [2] M. Zeeb, M.R. Ganjali, P. Norouzi, S.R. Moeinossadat, Selective determination of penicillamine by on-line vapor-phase generation combined with Fourier transform infrared spectrometry, *Talanta*, 78(2009) 584–589.
- [3] E. López-Anreus, S. Garrigues, M. de la Guardia, Vapour generation-Fourier transform infrared spectrometric determination of benzene, toluene and methyl tert.-butyl ether in gasolines, *Anal. Chim. Acta*, 333(1996) 157–165.
- [4] E. López-Anreus, S. Garrigues, M. De La Guardia, Vapour generation Fourier transform infrared spectrometry. A new analytical technique, *Anal. Chim. Acta*, 308(1995) 28–35.
- [5] A. Pérez-Ponce, S. Garrigues, and M. de la Guardia, Vapour generation–Fourier transform infrared direct determination of ethanol in alcoholic beverages, *Analyst*, 121(1996) 923–928.
- [6] A. Pérez-Ponce, S. Garrigues, L. G. De, Direct vapor generation Fourier transform infrared spectrometric determination of ethanol in blood, *Anal. Chim. Acta*, 336(1996) 123–129.
- [7] S. Armenta, F. A. Esteve-Turrillas, G. Quintás, S. Garrigues, A. Pastor, M. de la Guardia, Development of a simple and low cost device for vapour phase Fourier Transform Infrared spectrometry determination of ethanol in mouthwashes, *Anal. Chim. Acta*, 569(2006) 238–243.
- [8] S. Armenta, S. Garrigues, M. de la Guardia, On-line vapor-phase generation combined with Fourier transform infrared spectrometry, *TrAC Trends Anal. Chem.*, 27(2008) 15–23.
- [9] S.W. Bruun, A. Kohler, I. Adt, G. D. Sockalingum, M. Manfait, H. Martens, Correcting attenuated total reflection-Fourier transform infrared spectra for water vapor and carbon dioxide, *Appl. Spectrosc.*, 60(2006) 1029–1039.
- [10] G. Erik, R. Jean-Marie, Subtraction of atmospheric water contribution in Fourier transform infrared spectroscopy of biological membranes and proteins, *Spectrochim. Acta Part Mol. Spectrosc.*, 50(1994) 2137–2144.
- [11] G. Quintás, J. Kuligowski, M. de la Guardia, S. Garrigues, Método de compensación de gases y/o vapores ambientales en espectros de infrarrojos., *Span. Pat. P201200792 Fail. 2012 Issued Pending.*

- [12] David Perez-Guaita, J. Kuligowski, G. Quintás, S. Garrigues, M. de la Guardia, Atmospheric Compensation in Fourier Transform Infrared (FT-IR) Spectra of Clinical Samples., *Appl. Spectrosc.*, 67(2013) 1339–42.
- [13] M.M. Cascant Vilaplana, S. Garrigues, M. de la Guardia, Headspace sample handling coupled with infrared spectroscopy, XVIII Reunión Soc. Esp. Quím. Analítica, 2013.
- [14] R.G. Gullberg, The frequency of apparent acetone in a group of breath alcohol data: statistical treatment and forensic implications, *Forensic Sci. Int.*, 67(1994) 65–72.
- [15] J.H. Watterson, Assessment of Response of the Intoxilyzer® 8000C to Volatiles of Forensic Relevance In Vitro, Part I: Acetone, Isopropanol, and Methanol, *Journal of Analytical Toxicology*, 33(2009) 109-117.
- [16] M.J. Henderson, B.A. Karger, G.A. Wren Shall, Acetone in the breath; a study of acetone exhalation in diabetic and nondiabetic human subjects, *Diabetes*, 1(1952) 188–193.
- [17] M.J. Sulway, J. M. Malins, Acetone in diabetic ketoacidosis, *The Lancet*, 296(1970) 736–740.
- [18] Y. Zhang, Y. Mu, P. Liang, Z. Xu, J. Liu, H. Zhang, X. Wa, J. Gao, S. Wang, F. Chai, A. Mellouki, Atmospheric BTEX and carbonyls during summer seasons of 2008–2010 in Beijing, *Atmos. Environ.*, 59(2012) 186–191.
- [19] F. Kondagula, K. Molt, Infrared Spectrometric Purity Control of Organic Liquids and Water, *CLEAN – Soil Air Water*, 37(2009) 955–962.

3. RESULTADOS EXPERIMENTALES Y ÁMBITO DE APLICACIÓN

PARTE 3: Muestras Arqueológicas

Parte 3. Muestras arqueológicas

La determinación de los elementos minerales en restos de esqueletos arqueológicos puede revelar los hábitos alimentarios, así como las condiciones ambientales y los hábitos de vida de la población. La principal herramienta utilizada para desentrañar el pasado es la química arqueológica (arqueoquímica), que permite descubrir las prácticas y los estilos de vida de las poblaciones humanas pasadas, incluyendo sus hábitos de alimentación.

Los huesos y los dientes son considerados como indicadores de la exposición a los elementos minerales y, aunque los huesos presentan numerosas dificultades analíticas con respecto a la separación y caracterización de los minerales constituyentes, se utilizan para determinar la alimentación de nuestros antepasados. Sin embargo, debido a su morfología, el esqueleto humano presenta una menor cantidad de hueso cortical junto una gran cantidad de hueso esponjoso, por lo que es más susceptible a los cambios post mortem producidos por el suelo. Por lo tanto, la reconstrucción de la dieta basada en la composición mineral de los huesos requiere que el proceso diagenético no haya alterado la composición ósea con el tiempo.

Normalmente el ICP-OES se utiliza para determinar los elementos minerales en suelos y huesos, ya que se trata de una técnica de análisis multielemental. Sin embargo, requiere de una digestión previa y disolución de las muestras, suponiendo un perjuicio para el operador desde el punto de vista de la manipulación y utilización de reactivos y ácidos, además de generar residuos que requieren de una posterior gestión. Es por ello, que resultaría de gran interés desarrollar una estrategia de bajo coste, rápida y limpia para estudiar y entender la posible degradación post-mortem de los huesos enterrados causada por el entorno, además de desarrollar un método para la determinación de diferentes elementos en restos óseos, mediante el uso de la espectroscopia en la región del infrarrojo cercano con transformada de Fourier (FT-NIR) junto con modelos quimiométricos y multivariantes.

3.8. CAPÍTULO 8:

**Predicción de elementos alcalinotérreos en restos óseos por
espectroscopia de la región del infrarrojo cercano**

3.8. Capítulo 8. Predicción de elementos alcalinotérreos en restos óseos por espectroscopia de la región del infrarrojo cercano

3.8.1. Introducción

Los huesos contienen entre el 50 y el 70% de minerales, entre el 50 y el 40% de materia orgánica, principalmente colágeno, el 5-10% de agua y menos del 3% de contenido en lípidos. La hidroxiapatita $[\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2]$ es el mineral mayoritario del hueso, junto con pequeñas cantidades de carbonato, magnesio y otros fosfatos, y la presencia de algunos grupos hidroxidos [1]. En condiciones fisiológicas, la hidroxiapatita es el único mineral estable, compuesta por un 38% de calcio y un 18% de fósforo, con trazas de sodio (0.6%), magnesio y pequeñas cantidades de otros elementos [2]. El número de sustituciones posibles en la apatita biológica es menor que en la apatita geológica debido al limitado número de elementos disponibles en el cuerpo. Las sustituciones de los iones en el hueso y en los dientes que se conocen son aquellas producidas por los iones F^- , Cl^- , Na^+ , K^+ , Fe^{2+} , Zn^{2+} , Sr^{2+} , Mg^{2+} , citrato, y carbonato [3]. Los minerales son ingeridos a través de los alimentos o involuntariamente absorbidos desde el entorno del individuo al mismo.

La composición mineral de los huesos es comúnmente empleada para investigar las patologías, nutrición y otros fines bio-arqueológicos y forenses. La composición mineral de los huesos puede ser modificada post-mortem debido a procesos post-deposicionales, llamados diagénesis. Desde hace unas décadas, varios investigadores han estudiado estos mecanismos naturales: por lo que algunos autores han evaluado el efecto de la contaminación del suelo en los huesos [4-9] y la diagénesis y efectos de la degradación en la matriz ósea [10-13].

Como se ha indicado, el calcio, el oxígeno y el hidrógeno son los constituyentes mayoritarios de los huesos minerales como la hidroxiapatita [14]. Por otro lado, el estroncio es un elemento no esencial que compite para reemplazar el calcio [15] y además se relaciona con los hábitos de consumo alimentario [16]. La composición del suelo y las condiciones ambientales juegan un papel crucial en la diagénesis, ya que afectan a las concentraciones de los elementos en los huesos enterrados.

Normalmente el ICP-OES se utiliza para determinar elementos minerales en suelos y huesos [17-21], ya que se trata de una técnica de análisis multielemental. Sin embargo, requiere de una digestión previa y disolución de las muestras, utilizando ácidos fuertes y, por lo tanto, lo que supone un perjuicio para el operador además de generar residuos que requieren de una posterior gestión.

Una posible alternativa podría ser el empleo de la espectroscopia infrarroja (IR), la cual proporciona un método de adquisición de los espectros rápido, sencillo y económico. Además, al tratar de de medidas directas, la técnica NIR puede considerarse dentro del grupo de los métodos sostenibles, ya que no utiliza reactivos ni disolventes y puede emplearse directamente sobre muestras sólidas a través de las medidas por reflectancia difusa. La combinación de la espectroscopia IR con el tratamiento y análisis quimiométrico de los espectros ofrece una herramienta alternativa para la determinación de los componentes mayoritarios de los huesos y sedimentos, incluyendo elementos minerales. La espectroscopia FT-IR ha sido utilizada en diversas aplicaciones para el estudio y análisis de muestras de huesos, entre las que se incluyen la determinación de colágeno, del tamaño de cristal del hueso, del contenido de carbonatos y la estructura cristalina [22–25], la tafonomía geológica [26], además de su uso como herramienta de criba para el estudio de la alteración diagenética [27, 28]. Solo en unos pocos estudios se ha empleado la espectroscopia NIR en el estudio y análisis de suelos [29] y de huesos [30] en contextos arqueológicos.

Los objetivos principales de este estudio es el desarrollo de una estrategia de bajo coste, rápida y limpia para estudiar y entender la posible degradación post-mortem de los huesos enterrados causada por el entorno de los mismos y ver cómo afecta a las distintas partes y tipos de huesos del esqueleto, además de definir un proceso para la selección de huesos mejor preservados y su posterior estudio. Al mismo tiempo, se ha desarrollado un método para la determinación de algunos elementos en restos óseos mediante el uso de la espectroscopia en la región del infrarrojo cercano con transformada de Fourier (FT-NIR) junto con herramientas quimiométricas como PCA y modelos multivariantes PLS.

3.8.2. Parte experimental

3.8.2.1. Instrumentos

Para la obtención de los espectros utilizó un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Alemania) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS. El equipo dispone de una esfera integradora que permite las medidas mediante reflectancia difusa de las muestras contenidas en viales de vidrio de 11 mm de diámetro interno y 25 mm de altura. El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos.

3.8.2.2. Muestras

Se recogieron 271 muestras de huesos de 78 individuos diferentes, pertenecientes a individuos adultos y jóvenes de la necrópolis romana tardía de la calle Virgen de la Misericordia [19] y de la calle En Gil [31], ambas ubicadas en la ciudad de Valencia (España), en el que el rito de sepultura romana consistía en la inhumación. Las tumbas pertenecen al período comprendido desde el siglo I después de Cristo hasta el comienzo del siglo V. Los huesos recogidos eran aproximadamente de un tamaño de 5 cm y una masa entre 15 y 20 gramos. Huesos como el fémur, la tibia, el húmero, el radio y los huesos parietales, compuestos principalmente por una parte cortical, se clasificaron como "corticales", y las costillas, principalmente compuestas por una matriz esponjosa, se clasificaron como "esponjosas". Los primeros 2 mm de hueso en contacto directo con los sedimentos se consideraron parte externa de los huesos y se denominaron como tal.

Todos los huesos fueron muestreados evitando los puntos osteométricos, usando un bisturí y una micro-espátula de cuchara. Para proporcionar resultados reproducibles, comparables y compatibles con la sensibilidad de los métodos analíticos empleados y poder relacionar los espectros NIR con los datos de referencia obtenidos, las muestras fueron previamente mineralizadas en una mufla empleando el siguiente programa de temperaturas : i) 30 minutos a 150 °C; ii) incremento de la temperatura hasta 450 °C a una velocidad de 1 °C min⁻¹; iii) mantenimiento de la temperatura a 450 °C durante 24 h; iv) desconexión y enfriamiento hasta temperatura ambiente. Una vez mineralizadas las muestras, se pulverizaron y se homogeneizaron con la ayuda de un mortero de ágata.

Se analizaron 243 muestras de la calle Virgen de la Misericordia, divididas en tres grupos: i) parte interna de los huesos, compuesto por la parte interna de los huesos corticales y esponjosos (95 muestras), ii) parte externa de los huesos, compuesta por la primera capa externa de los huesos corticales y esponjosos (94 muestras) y, iii) sedimentos que estaban directamente en contacto con los huesos (54 muestras). Adicionalmente y siguiendo la clasificación mencionada, 28 muestras de la calle En Gil fueron analizadas. Estas muestras se dividen en los tres grupos: i) parte interna de los huesos (8 muestras), ii) parte externa de los huesos (10 muestras) y iii) sedimentos (10 muestras).

3.8.2.3. Método de referencia

Los valores de referencia de la composición mineral de las muestras de huesos y sedimentos fueron determinados mediante el uso de un ICP-OES, modelo Optima

5300 DV, de Perkin Elmer (Norwalk, CT, EE.UU.), equipado con un autoamplificador AS 93-plus y un nebulizador de flujo cruzado. Las muestras se calcinaron previamente en una mufla modelo Biometra Lenton ECF 12145A (Lanera, España) y la digestión ácida de todas las muestras se realizó utilizando una placa de calentamiento C-Mag HS7 de Ika (Staufen, Alemania).

El procedimiento de digestión utilizado fue el método descrito por Gallelo *et al.* [19], con unas pequeñas modificaciones consistente en adicionar a 0.5 g de muestra mineralizada, 3.5 mL de una mezcla 1:1 de HCl y HNO₃ (usando HCl al 37% y HNO al 69% de pureza) y digerir en un baño de agua a 100 °C durante 40 minutos. Esta disolución concentrada (A), se diluyó 1:250 obteniéndose la disolución (B), para medir la concentración de Mg y Sr. La disolución (C) se obtuvo para analizar Ca por dilución de la disolución (A) 1:2000. La calibración se preparó a partir de una disolución madre multi-elemental de Ca, Mg y Sr de 100 µg ml⁻¹. Para evitar la obstrucción del nebulizador las disoluciones de las muestras fueron filtradas, empleando papel de filtro (WhatmanTM N.1 de 70 mm),. Los patrones se adquirieron de Scharlab S.L. (Barcelona, España). Como estándar interno se utilizó Re. La precisión estándar de las medidas oscilaron entre el 0% y 2% para los elementos considerados.

Para evaluar y controlar la exactitud del método analítico se utilizaron como materiales de referencia la ceniza de hueso NIST 1400 y el suelo GBW07408. El contenido de elementos minerales de las muestras analizadas variaron entre 72 y 421 mg g⁻¹ para el Ca, entre el 998 y 10964 µg g⁻¹ para el Mg, y entre 102 y 2100 µg g⁻¹ para el Sr, todos ellos referidos a la muestra mineralizada.

3.8.3. Resultados y discusión

3.8.3.1. Espectros FT-NIR

Las muestras una vez mineralizadas y pulverizadas en mortero de ágata se colocaron en viales de vidrio transparente de 11 mm de diámetro interno y 25 mm de altura para obtener directamente los espectros NIR por reflectancia difusa en unidades de Kubelka-Munk. Los espectros se registraron entre 14000 y 4000 cm⁻¹, acumulando 50 barridos por espectro y utilizando una resolución óptica de 4 cm⁻¹. El *background* fue adquirido antes de cada serie de medidas, con la esfera de integración cerrada y utilizando las mismas condiciones instrumentales que las empleadas para las muestras. Se obtuvieron tres réplicas de cada muestra, girando la posición del vial entre las réplicas para asegurar una buena representatividad. El promedio de los espectros triplicados de cada muestra fue

empleado para la exploración de datos y para construir los modelos quimiométricos.

En la Figura 8.1 se pueden ver, en unidades de Kubelka Munk, los espectros NIR promedios de cada uno de los grupos de muestras considerados en este estudio, correspondientes a la parte interna de los huesos (Figura 8.1a), la parte externa de los huesos (Figura 8.1b) y los sedimentos (Figura 8.1c), sin ningún tipo de pre-tratamiento espectral y en la región entre 9000 y 4000 cm^{-1} . Estos espectros presentan diferencias en la intensidad y posición de ciertas bandas, pero todos tienen como denominador común la presencia de bandas de absorción alrededor de 7200 y 5220 cm^{-1} , relacionadas con la combinación de las vibraciones de los enlaces H-O-H y O-H de agua.

La principal diferencia que presenta el espectro de los huesos internos es una débil vibración que aparece cerca de 4655 cm^{-1} (sin asignación clara) y una absorción localizada sobre 6977 cm^{-1} , debida al primer sobretono del estiramiento de grupo OH de la hidroxiapatita [36], que es más intensa que la que se observa en el espectro de la parte externa de los huesos. Además, en el caso de las muestras de sedimento, dicha banda está ausente. A 5278 cm^{-1} el espectro de la parte interna y el de la parte externa del hueso, pero en menor intensidad, presentan una banda de absorción relacionada con la apatita [30]. Por otro lado, en las muestras de sedimento aparece una banda a 7070 cm^{-1} relacionada con el primer sobretono de la vibración de estiramiento del enlace O-H en los compuestos del tipo metal-OH [37]. Además, las bandas de vibración cercanas a 4530 cm^{-1} son debidas a minerales arcillosos, como esmectita y illita, y la banda a 4265 cm^{-1} está asociada con la contribución de la calcita [37], que disminuye en intensidad en el espectro de la parte externa del hueso.

En el caso de parte interna de los huesos, la banda situada a 4265 cm^{-1} no aparece y la banda a 4530 cm^{-1} es menos intensa que en el espectro correspondiente a la parte externa de los huesos. Por lo tanto, en la parte externa de los huesos aparecen bandas típicas de muestras de sedimentos relacionadas con el enlace metal-OH (7070 cm^{-1}), calcita (4265 cm^{-1}) y minerales arcillosos (4530 cm^{-1}), y bandas menos intensas de hidroxiapatita (6977 cm^{-1}) y apatita (5278 cm^{-1}), siendo la degradación de hidroxiapatita en la parte externa de los huesos un gran indicador del proceso diagenético. Los sedimentos alrededor de los huesos son cruciales para la conservación de los mismos, debiéndose llevar a cabo el análisis de los mismos con el fin de evaluar su contribución en la modificación de la composición química del hueso.

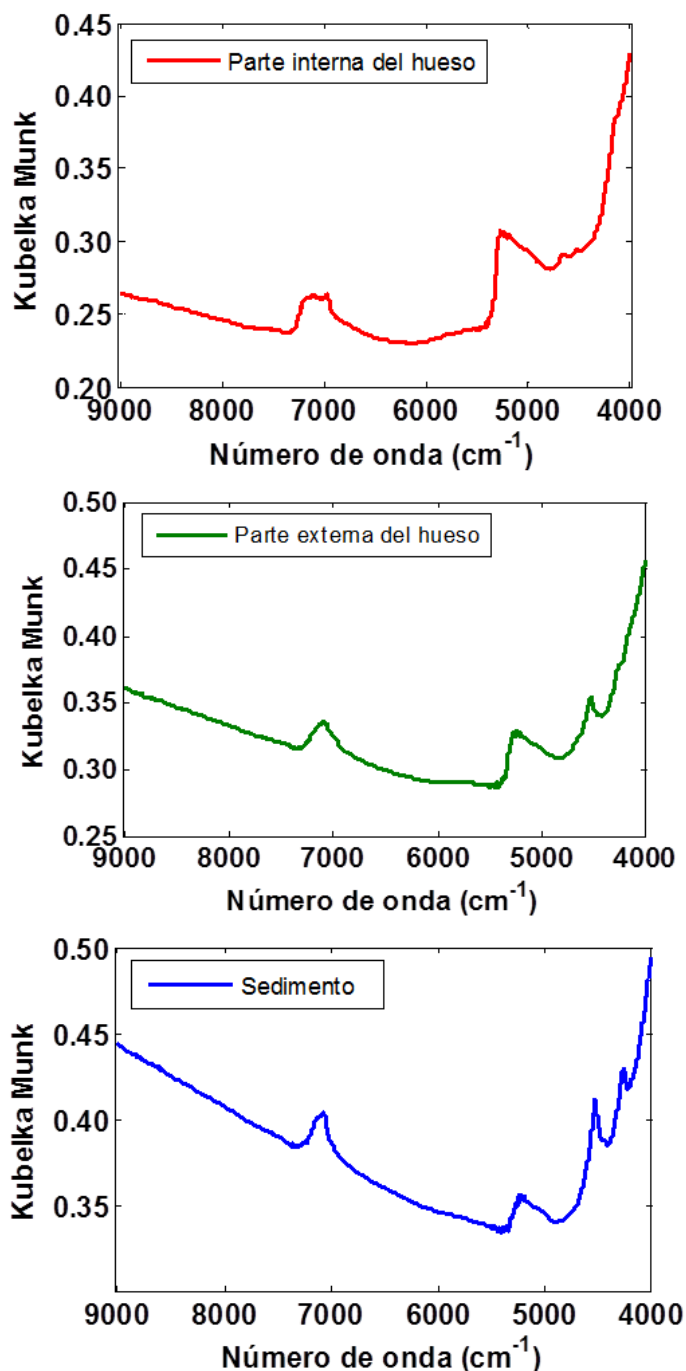


Figura 8.1. Espectros NIR promedio de las muestras mineralizadas en la región comprendida entre 9000 y 4000 cm⁻¹ y correspondientes a los grupos de muestras: a) parte interna de los huesos, b) parte externa de los huesos, y c) sedimento.

3.8.3.2. Exploración de los datos por PCA

Antes de construir los modelos de calibración, el PCA se utilizó para estudiar las diferencias espectrales entre las muestras analizadas. La Figura 8.2 presenta el gráfico de puntuaciones para las dos primeras componentes principales del PCA para las muestras de las Necropolis de calle Misericordia y calle En Gil, después de los pre-tratamientos: a) FD y MC y b) MSC, FD y MC, en la región de 9014 a 4000 cm^{-1} . En la Figura 8.2a, los dos primeros componentes principales representan el 86.07% de la varianza total explicada, siendo del 76.93% para PC1 y del 9.14% para PC2. En el gráfico de puntuaciones, se observa que no existe una clara separación de las muestras en los tres grupos, pero se puede apreciar que la posición de los sedimentos es claramente diferente de la parte interna de los huesos, de modo que las muestras de la parte externa de los huesos se encuentran situadas entre ambas, en la dirección del PC1. En la Figura 8.2b, puede observarse que el primer y segundo componente principal, después de realizarse el pre-tratamiento espectral de MSC, FD y MC, explican un total del 85.72% de la varianza explicada, con 80.49% para el PC1 y 5.23% para el PC2. En este gráfico de scores, las muestras de la parte interna de los huesos están situadas hacia la derecha del PC1, mientras que los sedimentos están situados hacia la izquierda. Las muestras de la parte externa de los huesos se encuentran situadas entre ambos grupos de muestras en la dirección PC1. Además, el PCA indica que las muestras de la calle En Gil son similares a las muestras de calle Misericordia, por lo que en principio las muestras de calle Misericordia podrían utilizarse como conjunto de calibración para predecir los elementos alcalinotérreos de la Necrópolis de la calle En Gil.

Además, se representaron los gráficos de PCAs con los mismos pre-tratamientos espectrales anteriores pero diferenciando tanto las partes internas como externas en huesos esponjosos y corticales. La Figura 8.2c presenta el gráfico de puntuaciones después de FD y MC y la Figura 8.2d después de MSC, FD y MC. No se detectaron diferencias entre los espectros de las partes internas de huesos corticales y esponjosos. Sin embargo, en cuanto la parte externa de los huesos, los huesos esponjosos se encuentran cerca de las muestras de sedimentos y los huesos corticales están situados más cerca de las partes internas de los huesos. Una posible explicación es que los huesos esponjosos son más susceptibles a la diagénesis que los huesos corticales [38], debido a la fragilidad y porosidad de su estructura.

Los números de onda responsables de estas diferencias se pueden ver en los gráficos de cargas representados en la Figura 8.2e después del pre-tratamiento de FD y MC, y en la Figura 8.2f después de MSC, FD y MC, siendo las más influyentes las bandas encontradas a 7000 cm^{-1} y entre 5500 y 4000 cm^{-1} .

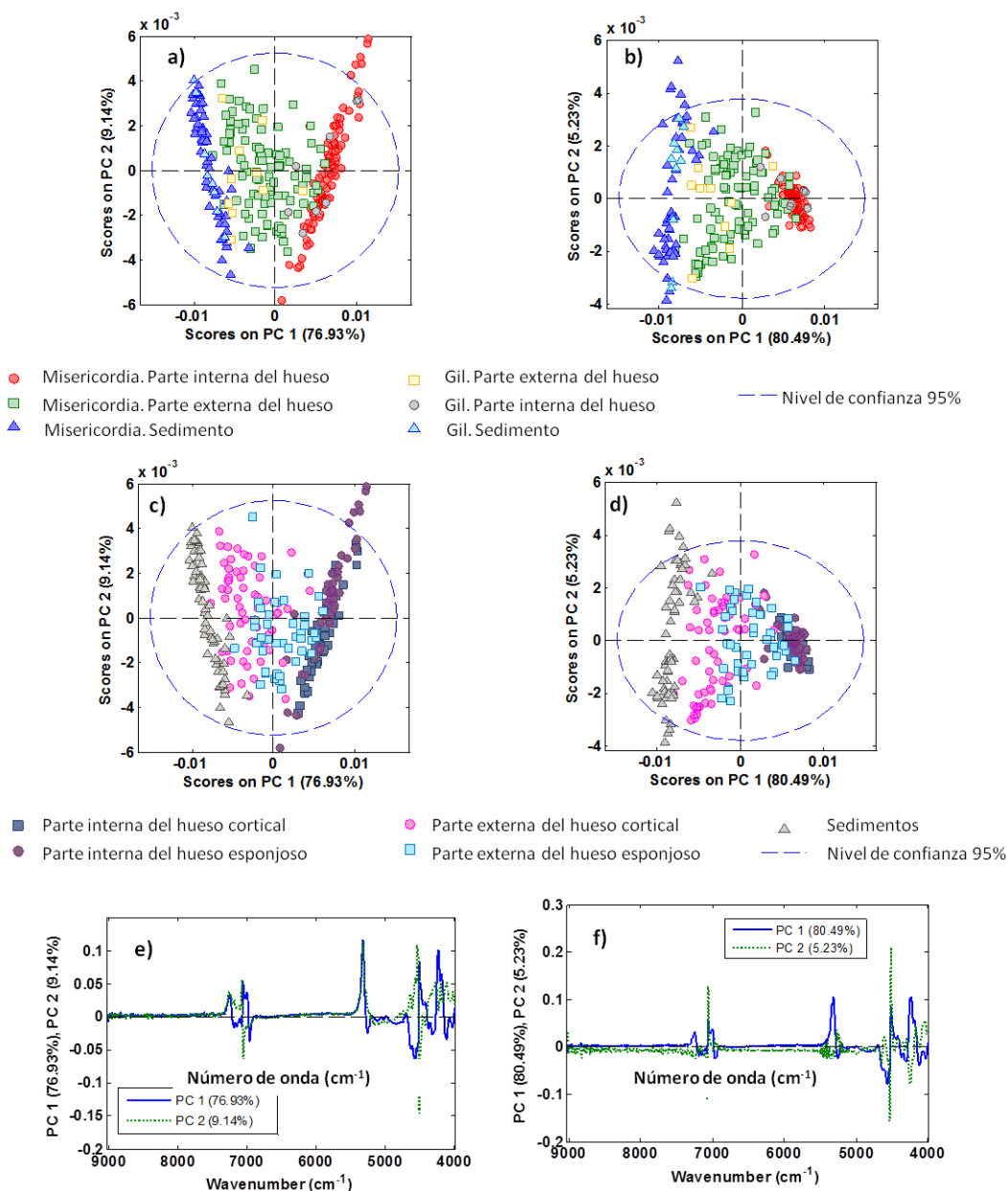


Figura 8.2. Gráficos de PCA obtenidos utilizando diferentes pre-tratamientos espectrales, clasificando las muestras en Necrópolis de calle Misericordia y calle En Gil, después de: a) FD y MC, y b) MSC, FD y MC. Gráficos de PCA obtenidos clasificando los huesos en cortical y esponjoso después de: c) FD y MC, y d) MSC, FD y MC. Cargas de las variables en PC1 y PC2 después de: e) FD y MC, f) MSC, FD y MC.

Por lo tanto, se puede concluir que el estudio por PCA, utilizando los espectros NIR, proporciona una herramienta rápida y verde para identificar los cambios producidos por la diagénesis en los distintos huesos. Estos resultados indican que la parte externa de los huesos sufre una mayor degradación que la parte interna, posiblemente producida por los factores diagenéticos, especialmente significativa en el caso de los huesos esponjosos debido a su alta porosidad. Por lo tanto, cabe esperar que dependiendo del tipo y de la parte (interna o externa) de hueso que se utilice para el análisis se obtendrían resultados distintos, por lo que, para evitar errores, sería interesante el uso del análisis PCA como método exploratorio, permitiendo así una selección previa de las muestras no afectadas por los procesos diagenéticos. Además, el PCA es un buen indicador para confirmar que el conjunto de calibración sea representativo de la población total de muestras, y que en su distribución se encuentre también contenido el conjunto de las muestras de validación, de modo que se puedan construir modelos PLS adecuadamente representativos.

3.8.3.3. Modelos PLS-NIR

Para construir los modelos PLS, las muestras de la Necrópolis de la calle Misericordia se dividieron en un conjunto de calibración y otro de validación, con 183 y 60 muestras, respectivamente. Para ello se utilizó el algoritmo K-S para la selección del conjunto de calibración, estando finalmente compuesto por 62 espectros de las partes externas de los huesos, 70 espectros de las partes internas de los huesos y 51 espectros correspondientes a muestras de sedimento. Antes de construir el modelo, se eliminaron las muestras aberrantes tanto del conjunto de calibración como de validación, ya que presentaban valores elevados de Q residuales y T^2 Hotelling.

Una vez establecidos los modelos, se probó la capacidad predictiva de los mismos utilizando un tercer conjunto, denominado conjunto de predicción, compuesto por 26 muestras de hueso pertenecientes a la Necrópolis de la calle En Gil.

Las principales características de los conjuntos de calibración, validación y predicción para la determinación de calcio, magnesio y estroncio mediante los modelos PLS-NIR se indican en Tabla 8.1.

Para construir los mejores modelos PLS-NIR y evaluar su capacidad de predicción, se ensayaron varias regiones y diferentes estrategias de pre-procesamiento. Para todos los elementos, los mejores resultados fueron obtenidos utilizando la región comprendida entre 9014 y 4000 cm^{-1} .

Tabla 8.1. Rangos de concentración de los conjuntos de calibración, validación y predicción de los elementos minerales determinados.

Analito	Conjunto	Muestras	Rango	Valor promedio	SD
Calcio (mg g ⁻¹)	Calibración	181	96 - 421	260	100
Magnesio (µg g ⁻¹)		181	1050 - 10500	4967	3091
Estroncio (µg g ⁻¹)		181	153 - 2100	823	486
Calcio (mg g ⁻¹)	Validación	58	72 - 400	312	76
Magnesio (µg g ⁻¹)		59	1070 - 8500	3226	2545
Estroncio (µg g ⁻¹)		58	124 - 2010	1058	405
Calcio (mg g ⁻¹)	Predicción	26	112 - 421	233	114
Magnesio (µg g ⁻¹)		26	998 - 10964	6723	3864
Estroncio (µg g ⁻¹)		26	102 - 1350	518	442

SD: desviación estándar

Para construir modelos PLS-NIR utilizando el rango espectral seleccionado, se emplearon 3, 4 y 4 LV's para Ca, Mg y Sr, respectivamente, con el fin de minimizar el RMSECV, que explicaban un 89.1%, 90.8% y 88.5% de la varianza total, en para cada uno de ellos. Para Ca y Mg, se seleccionó el pre-tratamiento de FD con una suavizado de 11 puntos y un polinomio de segundo orden, y MC. Para el Sr, se eligieron MSC, FD y MC. La presencia de algunas muestras aberrantes fue identificada por altos valores de Q residual y T² de Hotelling, eliminándose 2 muestras del conjunto de calibración para los tres elementos, 2 muestras del conjunto de validación para los modelos del Ca y del Sr, mientras que sólo una muestra del conjunto de validación de Mg.

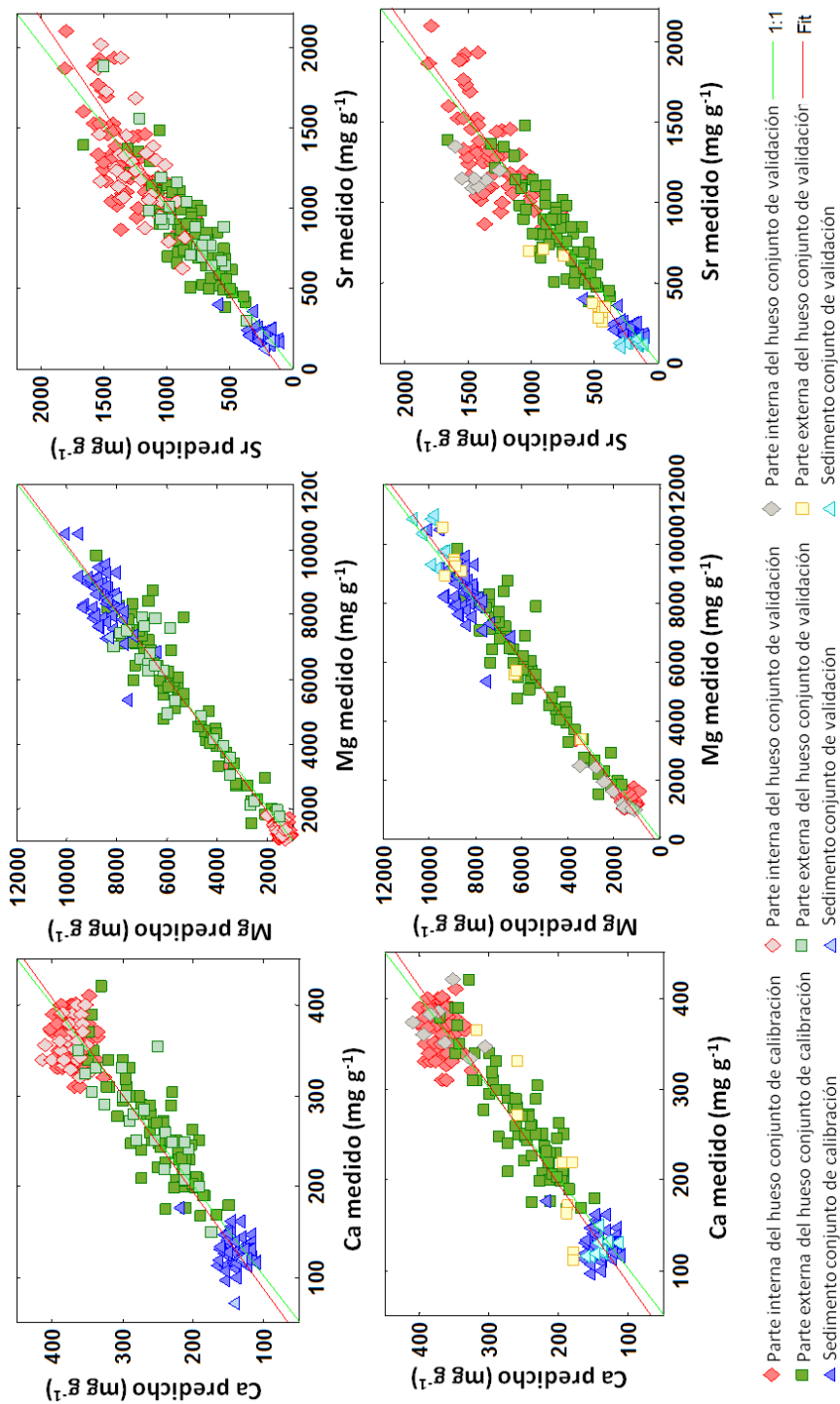
En la Figura 8.3 se pueden ver las regresiones entre los valores predichos PLS-NIR para Ca, Mg y Sr y los obtenidos por el método de referencia ICP-OES. Se puede observar que las muestras de calibración y validación se distribuyeron cerca de la línea de regresión óptima entre los valores predichos y medidos para los tres elementos, obteniendo altos coeficientes de determinación tanto para la calibración (R² Cal), la validación cruzada (R² CV) y la predicción (R² Pred). La capacidad de predicción de los modelos PLS-NIR resultó aceptable para todos los elementos, obteniendo valores de RRMSEP del 10, 15 y 19% para el Ca, Mg y Sr, respectivamente.

Los parámetros de calibración y validación más importantes de los modelos PLS-NIR desarrollados se resumen en la Tabla 8.2. Se puede observar que las muestras de la parte interna de los huesos presentan mayores contenidos de calcio ($368 \pm 24 \text{ mg g}^{-1}$) y estroncio ($1306 \pm 235 \text{ } \mu\text{g g}^{-1}$) y menor contenido de magnesio ($1393 \pm 175 \text{ } \mu\text{g g}^{-1}$) que las partes externas (con $261 \pm 47 \text{ mg g}^{-1}$ de Ca, $844 \pm 203 \text{ } \mu\text{g g}^{-1}$ de Sr y $5599 \pm 1684 \text{ } \mu\text{g g}^{-1}$ de Mg), y muestras de sedimentos (con $133 \pm 17 \text{ mg g}^{-1}$ de Ca, $239 \pm 60 \text{ } \mu\text{g g}^{-1}$ de Sr y $8165 \pm 858 \text{ } \mu\text{g g}^{-1}$ de Mg).

Como se mencionó al comienzo de este capítulo, en condiciones fisiológicas, la hidroxilapatita está compuesta de un 38% de Ca, con trazas de Mg y pequeñas cantidades de otros elementos como Sr, por lo que los resultados obtenidos en este estudio indican que la parte interna de los huesos está prácticamente intacta, observando valores de Ca de alrededor del 38%. El enriquecimiento de Mg en la parte externa de los huesos indica que este elemento es incorporado de los sedimentos a los huesos, probablemente, durante el proceso de diagénesis. La parte externa de los huesos, especialmente en el caso de los restos óseos esponjosos, son más sensibles a los procesos diagenéticos causados por el medio ambiente, lo que indica que la degradación no afecta en la misma medida a las diferentes partes esqueléticas.

Debido al hecho de que los huesos esponjosos tienen mayor porosidad y corteza más delgada que los huesos corticales, los primeros son menos resistentes a los factores diagenéticos, sufriendo mayores cambios químicos. Los resultados obtenidos por los modelos PLS-NIR indicaron que la espectroscopia NIR tiene un buen potencial para predecir de forma rápida, sencilla y directa el contenido de elementos alcalinotérreos en los restos óseos. Tal como se ha mencionado, este método tiene muchas ventajas sobre las técnicas analíticas empleadas convencionalmente porque es rápido, barato, no destructivo para muestras pulverizadas y no requiere el uso de reactivos químicos ni de disolventes.

Figura 8.3. Regresión entre los valores predichos y de referencia obtenidos para la determinación de calcio, magnesio y estroncio en restos óseos y sedimentos



3.8.3.4. Capacidad de predicción de los modelos PLS-NIR

Para evaluar la capacidad de predicción de los modelos PLS-NIR construidos para la determinación de las concentraciones de Ca, Mg y Sr en huesos y sedimentos, se utilizó un conjunto de validación independiente, formado por 26 muestras de la necrópolis de calle En Gil, que contenían parte externa y parte interna de los huesos, así como muestras de sedimentos. Los valores correspondientes al valor promedio, el intervalo de concentración y la desviación estándar para la concentración de elementos analizados empleados para el conjunto de predicción se resumen en la Tabla 8.1. Los valores predichos de los analitos frente a los obtenidos por los métodos de referencia se muestran en la Figura 8.3 y, los parámetros de calibración y predicción más importantes de los modelos PLS-NIR desarrollados se resumen en la Tabla 8.2. Se obtuvieron valores RRMSEP aceptables para calcio y magnesio con un 15 y 11%, respectivamente, y para estroncio se obtuvo un valor de 37%, con coeficientes de correlación de 90.4, 97.4 y 97.4, y valores RPD de 3.2, 5.33 y 2.3, respectivamente. Estos resultados demuestran una buena capacidad predictiva de los modelos PLS-NIR desarrollados para evaluar elementos alcalinotérreos en muestras de huesos y sedimentos. Además, debe indicarse que los modelos PLS fueron elaborados a partir de muestras de hueso de distinto origen y fecha, lo que confirma la viabilidad de nuestra innovadora propuesta metodológica para la evaluación de elementos alcalinotérreos en restos arqueológicos. Los resultados obtenidos indican una buena capacidad predictiva para Mg en muestras que no pertenecen a la misma población, mientras que Ca y Sr quedan limitados como herramienta de cribado. El alto error relativo del 37% para el Sr se debe, probablemente, a los bajos niveles de este elemento encontrados en algunas muestras no compatibles con la sensibilidad de NIR.

Tabla 8.2. Descripción de los mejores modelos PLS-NIR utilizados para la determinación de Ca, Mg y Sr y sus principales figuras de mérito

Elemento	Pre-procesamiento	LV	Muestras de calle Misericordia										Muestras de calle En Gil				
			RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	R ² Pred	RRMSEP (%)	RPD	RMSEP	R ² Pred	Slope	Intercept	Bias	RRMSEP (%)	RPD
Calcio (mg g ⁻¹)	FD, MC	3	27.5	92.4	29.0	91.6	31.7	83.4	10.2	2.4	35.6	90.4	0.8299	41.4	1.7	15.3	3.2
			686.3	95.0	750.2	94.1	496.6	96.3	15.4	5.1	723.7	97.3	0.9504	246.2	-81.3	10.8	5.3
Stroncio (µg g ⁻¹)	MSC, FD, MC	4	150.1	90.4	169.8	87.7	198.9	76.2	18.8	2.0	198.9	97.4	0.9041	78.9	155.0	37.3	2.3

LV's: número de variables latentes
 RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción;
 RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada;
 R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

3.8.4. Conclusiones Capítulo 8

Este estudio ha demostrado que las medidas de reflectancia difusa por espectroscopía FT-NIR proporciona un método analítico rápido, barato y verde para la predicción del contenido de calcio, magnesio y estroncio en muestras de hueso enterrados y sedimentos, que podría ser muy útil para entender los cambios post mortem de los huesos causados por el medio ambiente y para seleccionar las muestras de hueso con señales biogenéticas bien conservadas.

El PCA ha demostrado que el perfil de elementos alcalinotérreos de las muestras de hueso y sedimentos, como era de esperar, es claramente diferente. Además, la superficie externa de los huesos esponjosos es más similar a los sedimentos que a la superficie de los huesos corticales. Esto confirma que la capa ósea externa de los huesos esponjosos está más alterada que la superficie cortical y la presencia de Ca, Mg y Sr es similar a la de las muestras de suelo de los sitios estudiados. Por lo tanto, se puede concluir que los factores ambientales tienen un impacto principal en los tejidos esponjosos, que son más susceptibles a los procesos diagenéticos, que los huesos corticales. A estas conclusiones se puede llegar a partir de los cambios en los espectros NIR en los que se puede observar la pérdida de la banda de absorción correspondiente a la apatita (el compuesto óseo principal) y a la presencia de bandas de arcilla, especialmente en la superficie del hueso esponjoso.

En conclusión, la espectroscopia FT-NIR puede emplearse para predecir los contenidos de Ca, Mg y Sr en restos óseos y muestras de suelos, encontrándose que el contenido de Ca y Sr era mayor en la parte interna de los huesos, frente a los contenidos de Mg que eran más altos en la superficie externa de los huesos y en los suelos, y más bajos en la parte interna de los huesos. Nuestro enfoque metodológico desarrollado combina la espectroscopía NIR y el análisis estadístico, permitiendo predecir la composición mineral ósea de los minerales alcalinotérreos, abriendo una nueva perspectiva para la identificación de aquellas muestras mejor conservadas en estudios bioarqueológicos e investigaciones de la ciencia forense. Respecto a las limitaciones del estudio, debe observarse que se basó únicamente en la parte mineral de los restos óseos tratados térmicamente y pulverizados antes de las mediciones, perdiendo así información de la parte orgánica de los huesos y suelos. Además, la falta de sensibilidad de la espectroscopia NIR afecta a la capacidad de estas mediciones para ser empleadas para la determinación de elementos traza, pudiéndose determinar únicamente componentes principales del hueso, como son Ca, Mg y Sr.

3.8.5. Referencias Capítulo 8

- [1] B. Clarke, Normal bone anatomy and physiology, *Clin. J. Am. Soc. Nephrol.* 3 (2008) Suppl 3, S131–139.
- [2] J. Burton, Bone chemistry and trace element analysis. in M.A. Katzenberg, S.R. Saunders (Eds.) *Biological anthropology of the human skeleton*. 2nd ed. John Wiley & Sons, Inc., Hoboken, NJ (2008) 443–460.
- [3] B. Wopenka and J. D. Pasteris, A mineralogical perspective on the apatite in bone, *Mater. Sci. Eng. C*, 25 (2005) 131–143.
- [4] R. E. M. Hedges and A. R. Millard, Bones and Groundwater: Towards the Modelling of Diagenetic Processes, *J. Archaeol. Sci.*, 22 (1995) 155–164.
- [5] R. E. M. Hedges, Bone diagenesis: an overview of processes, *Archaeometry*, 44 (2002) 319–328.
- [6] W. Querido, A. L. Rossi, and M. Farina, The effects of strontium on bone mineral: A review on current knowledge and microanalytical approaches, *Micron*, 80 (2016), 122–134.
- [7] J. B. Lambert, L. Xue, and J. E. Buikstra, Inorganic analysis of excavated human bone after surface removal, *J. Archaeol. Sci.*, 18 (1991) 363–383.
- [8] F. Donald Pate and J. T. Hutton, The Use of Soil Chemistry Data to Address Post-mortem Diagenesis in Bone Mineral, *J. Archaeol. Sci.*, 15 (1988) 729–39.
- [9] J. Zapata, C. Pérez-Sirvent, M. Martínez-Sánchez, and P. Tovar, Diagenesis, not biogenesis: Two late Roman skeletal examples, *Sci. Total Environ.*, 369 (2006) 357–368.
- [10] I. Reiche, L. Favre-Quattropani, Calligaro, T. J. S., H. Bocherens, L. Charlet, M. Menu, Trace element composition of Archaeological Bones and postmortem alteration in the burial environment. *Nuclear Instruments and Methods. Physics Research B*, 150 (1999) 656-662.
- [11] I. Reiche, L. Favre-Quattropani, C. Vignaud, H. Bocherens, L. Charlet, M. Menu, A multi-analytical study of bone diagenesis: the Neolithic site of Bercy (Paris, France). *Meas. Sci. Technol.*, 14 (2003)1608-1619.
- [12] T.H. Schmidt-Schultz, M. Schultz, Intact protein molecules in archaeological bones – Bone matrix as a treasure chest of ancient diseases and living conditions. *American Journal of Physical Anthropology*, Annual Meeting Issue 1999, Wiley-Liss, p. 230.

- [13] T. A. Surovell and M. C. Stiner, Standardizing Infra-red Measures of Bone Mineral Crystallinity: an Experimental Approach, *J. Archaeol. Sci.*, 28(2001) 633–642.
- [14] F. Ignac, G. Gopinath, and V. der W. Hans, *Radionuclide and Hybrid Bone Imaging*. Springer Heidelberg New York Dordrecht London, 2012.
- [15] M. L. Carvalho and A. F. Marques, Diagenesis evaluation in Middle Ages human bones using EDXRF, *X-Ray Spectrom.*, 37 (2008) 32–36.
- [16] D. Guimarães, A. A. Dias, M. Carvalho, M. L. Carvalho, J. P. Santos, F. R. Henriques, F. Curate, and S. Pessanha, Quantitative determinations and imaging in different structures of buried human bones from the XVIII-XIXth centuries by energy dispersive X-ray fluorescence – Postmortem evaluation, *Talanta*, 155 (2016) 107–115.
- [17] G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, and J. Bernabeu, Biological mineral content in Iberian skeletal remains for control of diagenetic factors employing multivariate statistics, *J. Archaeol. Sci.*, 40(2013) 2477–2484.
- [18] A.-F. Maurer, M. Gerard, A. Person, I. Barrientos, C. R. del, V. Darras, C. Durllet, V. Zeitoun, M. Renard, and B. Faugère, Intra-skeletal variability in trace elemental content of Precolumbian Chupicuaro human bones: The record of post-mortem alteration and a tool for palaeodietary reconstruction, *J. Archaeol. Sci.*, 38 (2011) 1784–1797.
- [19] G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, and J. Bernabeu, Chemical element levels as a methodological tool in forensic science, *J Forensic Res*, 6 (2014) 1000264.
- [20] A.-F. Maurer, A. Person, T. Tütken, S. Amblard-Pison, and L. Ségalen, Bone diagenesis in arid environments: An intra-skeletal approach, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 416 (2014) 17–29.
- [21] S. Dal, L. Maritan, D. Usai, I. Angelini, and G. Artioli, Bone diagenesis at the micro-scale: Bone alteration patterns during multiple burial phases at Al Khiday (Khartoum, Sudan) between the Early Holocene and the II century AD, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 416 (2014) 30–42.
- [22] S. W. Keenan, A. S. Engel, A. Roy, G. Lisa Bovenkamp-Langlois, Evaluating the consequences of diagenesis and fossilization on bioapatite lattice structure and composition, *Chem. Geol.*, 413 (2015) 18–27.

- [23] C. N. Trueman, K. Privat, and J. Field, Why do crystallinity values fail to predict the extent of diagenetic alteration of bone mineral, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 266 (2008) 160–167.
- [24] T. J. U. Thompson, M. Islam, K. Piduru, and A. Marcel, An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 299(2011) 168–174.
- [25] T. J. U. Thompson, M. Islam, and M. Bonniere, A new statistical approach for determining the crystallinity of heat-altered bone mineral from FTIR spectra, *J. Archaeol. Sci.*, 40 (2013) 416–422.
- [26] G. Iliopoulos, N. Galanidou, S. A. Pergantis, V. Vamvakaki, and N. Chaniotakis, Identifying the geochemical taphonomy of the osteological material from Katsambas rockshelter, *J. Archaeol. Sci.*, 37 (2010) 116–123.
- [27] L. E. Wright and H. P. Schwarcz, Infrared and Isotopic Evidence for Diagenesis of Bone Apatite at Dos Pilas, Guatemala: Palaeodietary Implications, *J. Archaeol. Sci.*, 23 (1996) 933–944.
- [28] D. Alfano, A. R. Albuja, O. Motta, and A. Proto, Detection of diagenetic alterations by Spectroscopic Analysis on Archaeological Bones from the Necropolis of Poseidonia (Paestum): A case study, *J. Cult. Herit.*, 10 (2009) 509–513.
- [29] J. Linderholm and P. Geladi, Classification of archaeological soil and sediment samples using near infrared techniques, *NIR News*, 23 (2012) 6.
- [30] D. Thomas, C. McGoverin, A. Chinsamy, and M. Manley, Near infrared analysis of fossil bone from the Western Cape of South Africa, *J. Infrared Spectrosc.*, 19(2011) 151.
- [31] G. Gallelo, S. Silvia, J. Kuligowski, B. Fulvio, M. Francesco, and P. Agustín, Variación química intraesquelética relacionada con la diagénesis en los restos óseos de c/ en Gil (Valencia), *SAGVNTVM*, 47 (2015) 175–186.
- [32] S. de Jong, SIMPLS: An alternative approach to partial least squares regression, *Chemom. Intell. Lab. Syst.*, 3 (1993) 251–263.
- [33] R. W. Kennard and L. A. Stone, Computer Aided Design of Experiments, *Technometrics*, 11 (1969) 137–148.

- [34] P.C. Williams, D. Sobering, How do we do it: A brief summary of the methods we use in developing near infrared calibrations in A.M.C. Daves, P.C. Williams, Near infrared spectroscopy: The future waves (Eds.), NIR Publications, Chichester, UK (1995) 185–188.
- [35] W. Saeys, A. M. Mouazen, and H. Ramon, Potential for Onsite and Online Analysis of Pig Manure using Visible and Near Infrared Reflectance Spectroscopy, *Biosyst. Eng.*, 91(2005) 393–402.
- [36] Y. Ning, J. Li, W. Cai, and X. Shao, Simultaneous determination of heavy metal ions in water using near-infrared spectroscopy with preconcentration by nano-hydroxyapatite, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, 96 (2012) 289–294.
- [37] V. Aranda, A. Domínguez-Vidal, F. Comino, J. Calero, and M. J. Ayora-Cañada, Agro-environmental characterization of semi-arid Mediterranean soils using NIR reflection and mid-IR-attenuated total reflection spectroscopies, *Vib. Spectrosc.*, 74 (2014) 88–97.
- [38] J.B. Lambert, S.M. Vlasak, A.C. Thometz, J.E. Buikstra, A comparative study of the chemical analysis of ribs and femurs in Woodland population, *Am. J. Phys. Anthropol.*, 59 (1982) 289-294.

3.9. CAPÍTULO 9:

Investigaciones forenses de huesos quemados mediante el uso de la espectroscopia infrarroja en la región cercana

3.9. Capítulo 9. Investigaciones forenses de huesos quemados mediante el uso de la espectroscopia infrarroja en la región cercana

3.9.1. Introducción

El análisis de restos humanos quemados es de gran interés entre los forenses y antropólogos debido a los problemas asociados a su correcta recuperación, identificación y clasificación [1]. Las muestras de hueso quemado proporcionan conocimiento de los cambios producidos en el hueso y esta información puede ser utilizada posteriormente para los estudios antropológicos y forenses [2]. Consecuencia de este interés, durante las últimas décadas, el número de artículos publicados sobre el análisis de huesos quemados ha aumentado [3-5].

Los huesos quemados pueden clasificarse como carbonizados o como calcinados, dependiendo de la exposición térmica a la que han sido sometidos. Pueden exhibir un aspecto que varía del negro (si se trata de huesos carbonizados, debido al material esquelético y tejido blando) al blanco (cuando se trata de huesos calcinados). Los huesos quemados se consideran incinerados cuando la alteración térmica es tan prolongada e intensa que todos los materiales orgánicos y la humedad se pierden [6, 7], siendo la estimación de la temperatura máxima de exposición un factor crucial para la correcta interpretación de los huesos quemados; es por este motivo que se utiliza el cambio de color como método indicador del rango de temperatura a las que los restos óseos han sido sometidos [6, 8]. El color ha sido descrito cualitativamente mediante comparación visual con gráficos estándar, como los denominados *Munsell Soil Color Charts* [9]: sin embargo, diferentes percepciones o cambios en las condiciones de iluminación puede modificar significativamente los resultados observados [6]. Otro método utilizado para la determinación del color fue sugerido por Devlin *et al.* [10] que propone la utilización del sistema de coordenadas CIE L*a*b* (CIELAB) para el registro de los datos de color de la superficie ósea. Recientemente se ha hecho hincapié en los principales errores cometidos en las metodologías empleadas para la clasificación y el análisis de los huesos quemados, ya que algunos autores indican que el color de los fragmentos óseos son combinaciones de temperatura, duración de los tratamientos a temperatura elevada, presencia de oxígeno durante la combustión así como las condiciones y contexto en el que se realiza la incineración, razón por la que todos estos factores deben ser considerados para un correcto estudio de los restos óseos quemados [11, 12]. Nicholson encontró que los cambios de color variaban enormemente entre los huesos quemados, atribuyendo estas observaciones a diferentes contenidos orgánicos y a la química ósea [13]. Symes *et al.* indican que se pueden encontrar diferentes alteraciones de

color dentro de un solo esqueleto e incluso en un solo hueso, especialmente en casos de quema de restos carnosos [14]. Además diversos estudios inciden en que otros factores, como los de tipo externos y diagenéticos, podrían modificar el color del hueso, produciendo cambios similares a los causados por las altas temperaturas [15-17]. En general, el estudio de restos quemados es un desafío interpretativo difícil para forenses y antropólogos debido a los diferentes factores y aspectos a considerar.

Desde un punto de vista químico y cristalográfico, a medida que la temperatura aumenta la hidroxiapatita, presente en los huesos, tiene una mayor cristalinidad y forma cristales más estables y más grandes [18], comenzando la transformación más importante a temperaturas de 600 °C [11]. Lanting *et al.* [19] percibieron el potencial de utilizar los huesos calcinados para la datación por radiocarbono, indicando que todos los intentos anteriores realizados con huesos quemados fracasaron debido a que fueron tratados como huesos carbonizados. Los procesos de intercambio de iones bicarbonato disueltos en las aguas del suelo producen una contaminación en la bio-apatita que da como resultado fechas demasiado jóvenes en la datación por el método del ^{14}C , caso que no ocurre con los huesos calcinados [20]. Durante la última década se ha producido un incremento de los estudios de datación por radiocarbono en huesos calcinados [21-24]. Además, se han llevado a cabo varios estudios para comprender el comportamiento de los elementos minerales óseos en condiciones térmicas y la influencia de los procesos post-deposicionales en los restos esqueléticos [25-28]. Algunos trabajos han demostrado que los procesos post-mortem afectan en igual medida a los huesos antiguos cremados y no cremados [27, 28]. Grupe *et al.* realizaron estudios de composición elemental sobre huesos incinerados, concluyendo que la reconstrucción de la paleodieta es limitada debido a las pérdidas por volatilización y las modificaciones cristalinas causadas por las altas temperaturas [25]. Sin embargo, para la relación Sr/Ca, que es el método más utilizado para la reconstrucción paleodietaria, es posible utilizar el análisis de regresión.

La presencia de grupos hidroxilo en la apatita ósea sigue siendo un tema de controversia. Algunos estudios creen que una serie de grupos hidroxilo están presentes en los huesos arqueológicos [29, 30], mientras que otros argumentan que la apatita ósea no contiene concentraciones detectables de grupos hidroxilo [31, 32]. La hidroxilación de la apatita ósea se atribuye a la presencia de CO_3^{2-} , de acuerdo con dos mecanismos de sustitución de los sitios ocupados por el OH^- , a través de la llamada sustitución "tipo A", o de los sitios del PO_4^- , denominada sustitución "tipo B" [31]. También se producen otros intercambios, como la sustitución de fluoruro (F^-) por hidróxido (OH^-), y estroncio (Sr^{2+}) y otros cationes (Zn^{2+} , Pb^{2+}) presentes a niveles de trazas por el ión calcio (Ca^{2+}) [33].

A pesar de lo comentado anteriormente, se necesitan estudios adicionales para obtener información profunda sobre la composición mineral de los restos óseos, cristalinidad de las muestras e influencia de los intercambios de componentes del suelo. En este sentido, la espectroscopia infrarroja (IR) se ha utilizado para analizar las modificaciones de la apatita ósea durante las calcinaciones, sobre todo en muestras de hueso calentadas experimentalmente [34-36]. Sin embargo, los impactos relativos que suponen la temperatura y el tiempo de exposición permanecen poco claros, por lo que la utilidad del empleo de la espectroscopia IR podría limitarse a la evaluación de la presencia en los restos óseos de elementos minerales más importantes.

El objetivo del presente estudio ha sido seleccionar las muestras más adecuadas para llevar a cabo estudios forenses y, para ello, se ha evaluado la utilización de la espectroscopia de infrarrojo cercano mediante el uso de herramientas quimiométricas, como PCA, *Cluster Analysis* y PLS-DA. Los huesos carbonizados son más propensos que los calcinados a la degradación producidos por procesos de intercambio con el suelo, por lo que una clasificación de huesos permite asegurar los resultados de los análisis de las muestras más adecuadas. En ese sentido, para una mejor comprensión de los procesos de modificación ósea post-mortem, se analizaron tanto la parte interna como la parte externa de los huesos quemados así como muestras de sedimentos. Además, se evaluó la determinación simultánea de calcio, magnesio y estroncio en los huesos quemados por PLS-NIR con el fin de proporcionar una herramienta para evaluar la presencia de elementos minerales en los huesos que resulte rápida, directa, sencilla y respetuosa con el medio ambiente.

3.9.2. Parte experimental

9.2.1. Instrumentos

Los espectros del infrarrojo cercano registrados en este trabajo fueron adquiridos en un espectrofotómetro FT-NIR modelo Multi Purpose Analyzer (MPA) de Bruker (Bremen, Germany) equipado con una fuente NIR, un divisor de haz de cuarzo y un detector de PbS. El equipo dispone de una esfera integradora que permite las medidas mediante reflectancia difusa de las muestras contenidas en viales de vidrio de 11 mm de diámetro interno y 25 mm de altura. El software Opus 6.5 de Bruker se utilizó tanto para el control instrumental, como para la adquisición de datos

3.9.2.2. Muestras

Se estudiaron un total de 32 muestras de huesos quemados junto con 6 muestras de sedimentos, procedentes de la necrópolis de Corral de Saus [28], fechado entre los siglos III y II a.C., período en el que las cremaciones se colocaban en urnas donde cada una contenía los restos de un solo individuo. Los huesos se dividieron en i) "carbonizados", huesos quemados en atmósfera reductora, ii) "calcinados", quemados en atmósfera oxidante, y iii) "desconocidos", para los huesos de condiciones de combustión no identificadas. Además, se obtuvieron huesos de animales no quemados y sedimentos del interior de las urnas, mezclados con cenizas, carbón y huesos humanos. La distribución detallada de las 38 muestras del estudio corresponde a 15 partes internas de los huesos (7 carbonizados y 8 calcinados), 7 de huesos externos (obtenidos de los primeros 2 mm de hueso directamente en contacto con los sedimentos), 7 huesos desconocidos, 3 huesos de animales no quemados y 6 muestras de sedimento. Adicionalmente, para evaluar la capacidad y validez del método analítico desarrollado, se han considerado 31 muestras desconocidas procedentes de la Necrópolis de Las Peñas [37] y los materiales certificados de referencia correspondientes a hueso (NIST 1400) y suelo (NIM GBW07408).

3.9.2.3. Método de referencia

Las muestras fueron sometidas a digestión ácida antes de ser analizadas mediante la técnica de ICP-OES. El método de digestión consistió en la adición de 1.5 ml de HCl conc. y 1.5 ml de HNO₃ conc. a 0.5 g de muestra, tanto para los huesos como para el suelo, en tubos de vidrio, llevándose a cabo en un baño de agua a 100 °C durante 40 min. Las soluciones digeridas se transfirieron cuantitativamente a tubos de plástico y se diluyeron hasta un volumen final de 15 ml con agua destilada. Los intervalos de concentración de las diluciones de la solución digerida se adaptaron a la sensibilidad de las mediciones ICP-OES para cada elemento. A partir de una disolución patrón multi-elemental que contiene Ca, Mg y Sr a una concentración individual de 100 µg ml⁻¹, se prepararon patrones de calibración mediante la dilución apropiada, añadiendo HNO₃ y HCl a todas las muestras. Además, se utilizaron los materiales de referencia NIST 1400 de ceniza de hueso y NIM GBW07408 de suelo, para evaluar la exactitud del método analítico. Una disolución patrón de Re fue utilizada como estándar interno para las determinaciones por ICP-OES [28]. Los contenidos de minerales de las muestras, determinados por el método de referencia, se encontraba comprendidos entre 250 y 461 mg g⁻¹ para el calcio, entre 790 y 34257 µg mg⁻¹ para el magnesio, y entre 161 y 1068 µg mg⁻¹ para el estroncio.

3.9.3. Resultados y discusión

9.3.1. Espectros NIR

Las muestras mineralizadas y pulverizadas en un mortero de ágata se colocaron en viales de vidrio transparente de 11 mm de diámetro interno y 25 mm de altura para obtener directamente los espectros NIR por reflectancia difusa en unidades de Kubelka-Munk. Los espectros se registraron entre 14000 y 4000 cm^{-1} , acumulando 50 barridos por espectro y utilizando una resolución óptica de 4 cm^{-1} . El *background* fue adquirido antes de cada serie de medidas, con la esfera de integración cerrada y utilizando las mismas condiciones instrumentales que las empleadas para las muestras. Se obtuvieron tres réplicas de cada muestra, girando la posición del vial entre las réplicas para asegurar una buena representatividad. El promedio de los espectros triplicados de cada muestra fue empleado para la exploración de datos y para construir los modelos quimiométricos.

En la Figura 9.1 pueden verse los espectros NIR promedio, en la región entre 9000 y 4000 cm^{-1} sin ningún tipo de tratamiento previo de los datos, de las muestras correspondientes a los grupos de: a) hueso calcinado, b) hueso carbonizado, c) parte superficial del hueso, d) hueso animal y e) muestras de sedimento. Las bandas que se observan en la región de los espectros comprendida entre 5500 y 8000 cm^{-1} corresponde con el primer sobretono de la vibración de estiramiento del enlace -OH que aparece en el infrarrojo medio; las bandas de la región entre 4000 y 5500 cm^{-1} se atribuyen a los armónicos de los enlace OH del agua y a las bandas de combinación del grupo carbonato; finalmente, las absorciones presentes en la región entre 8000 y 12000 cm^{-1} se relacionan con el segundo sobretono de las vibraciones de estiramiento de los grupos OH, en las que también se incluyen las bandas correspondientes a las transiciones electrónicas resultantes de la presencia de iones metálicos de transición en la estructura de los compuestos presentes en las muestras [43].

Una descripción de las posiciones y asignaciones de las distintas bandas se incluye en la Tabla 9.1 [44-47]. Las principales diferencias encontradas entre los huesos quemados, en función de su tratamiento, se basan en la banda situada a 6977 cm^{-1} , asignada a las vibraciones de estiramiento del grupo OH en la hidroxiapatita, presentando una mayor intensidad en los huesos calcinados que en los carbonizados. Adicionalmente, la banda situada a 5275 cm^{-1} también está relacionada con la apatita, estando presente en los huesos carbonizados y no en los calcinados. Estas dos diferencias pueden indicar la sustitución de CO_3^{2-} en el lugar ocupado por los grupos OH debido a la pérdida del grado de hidroxilación

durante el proceso de carbonización, siendo por tanto estos huesos menos resistentes a la alteración diagenética que los huesos calcinados.

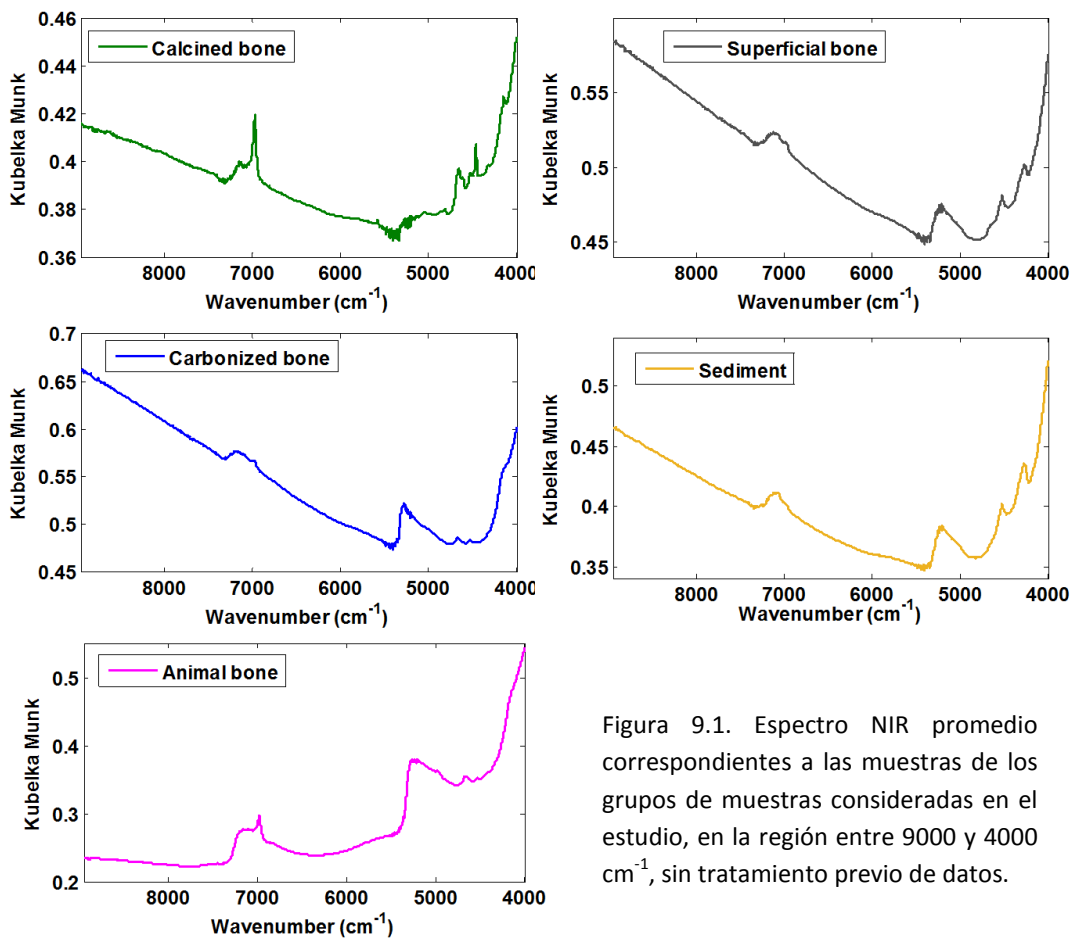


Figura 9.1. Espectro NIR promedio correspondientes a las muestras de los grupos de muestras consideradas en el estudio, en la región entre 9000 y 4000 cm^{-1} , sin tratamiento previo de datos.

Tabla 9.1. Asignación de las principales bandas NIR presentes los espectros correspondientes a los restos esqueléticos y a las muestras de sedimentos.

Wavenumber (cm ⁻¹)	CLN	CRB	EXT	ANM	SED	Assignments	Ref
7197	•	•	•	•	•	1st overtone of Mg-OH stretching in dioctahedral layers	[34]
7070	-	-	•	-	•	1st overtone of the O-H stretch vibration in metal-O-H	[32]
6977	•	•	•	•	•	1st overtone stretching vibrations of OH group in hydroxyapatite	[31]
5275	-	•	-	•	•	Apatite	[33]
5190	•	•	•	•	•	Combination vibrations of H-O-H bend and O-H stretch of water	[32]
4966	-	-	-	•	-		
4812	•	-	-	-	-		
4655	•	•	•	•	-		
4619	•	-	-	-	-		
4530	-	•	•	•	•	clay minerals, like smectite and illite	[31]
4468	•	-	-	-	-		
4326	•	-	-	-	-	OH combination (Mg-OH, stretching + bending)	[34]
4265	-	-	•	-	•	contribution of calcite	[31]
4150	•	•	-	•	-		

CLN: Hueso calcinado; CRB: Hueso carbonizado; EXT: Hueso externo; ANM: Hueso animal; SED: sedimentos

3.9.3.2. Análisis exploratorio

3.9.3.2.1. PCA

Como se ha venido indicando a lo largo de la memoria de esta Tesis Doctoral, el PCA es hoy en día la estrategia quimiométrica más común para el análisis exploratorio sin supervisión de datos experimentales, proporcionando el gráfico de *scores* una buena visión de las tendencias de los diferentes componentes y su relación. La Figura 9.2a muestra el gráfico de puntuaciones, para el primer y segundo componente principal, obtenido a partir de los espectros NIR de las muestras de las Necrópolis Corral de Saus y Las Peñas, después del pre-tratamiento FD y MC, y seleccionando la región entre 9000 y 4000 cm⁻¹. Los dos primeros componentes principales representan el 72.61% de la varianza explicada, siendo el 41.39% y el 31.22% para PC1 y PC2, respectivamente.

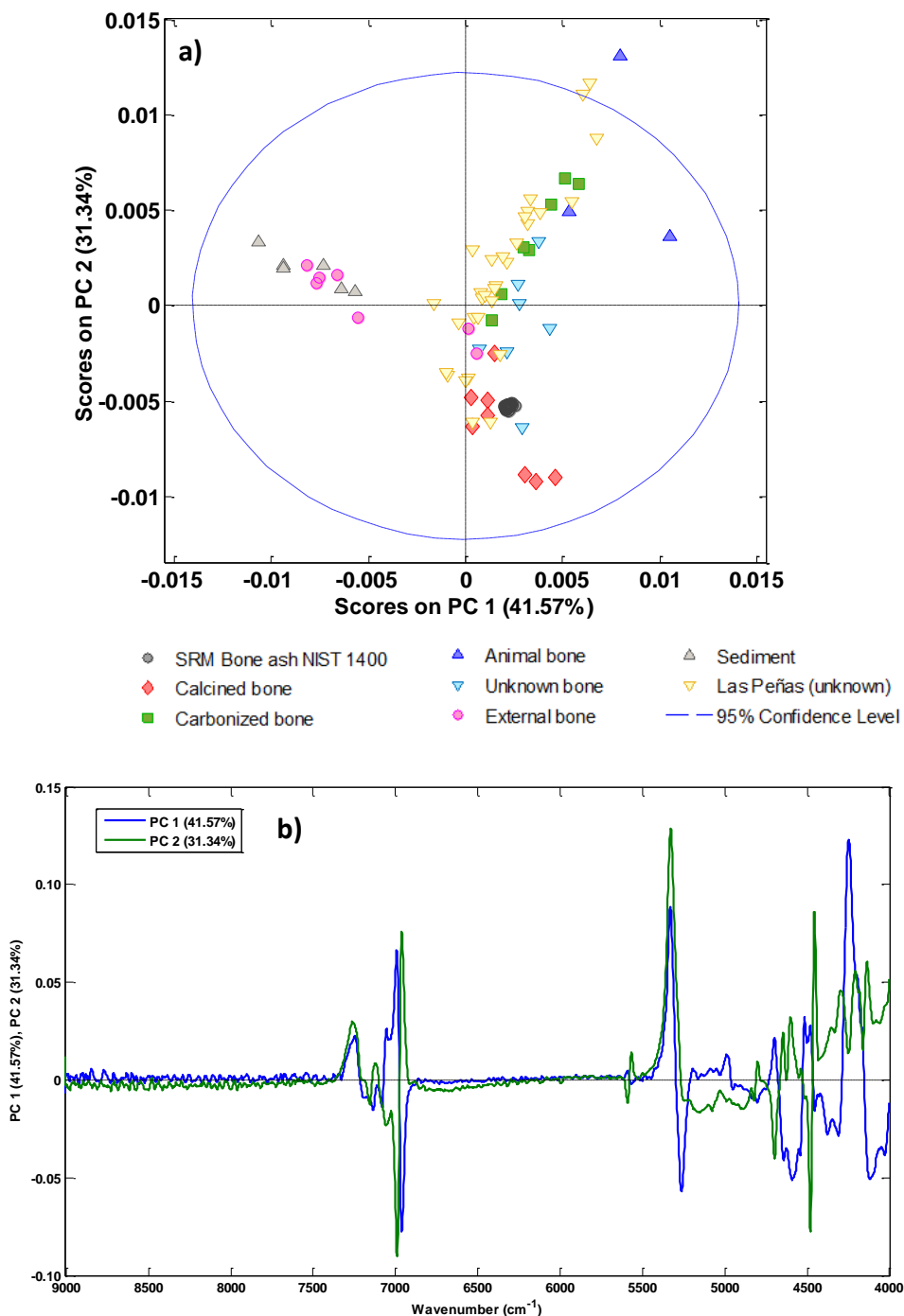


Figura 9.2. a) Gráfico de *scores* y b) gráfico de *loadings* obtenidos a partir de los espectros NIR de las muestras de las Necrópolis de Corral de Saus y Las Peñas, después de FD y MC en la región comprendida entre 9000 y 4000 cm⁻¹.

En la dirección de PC1, se puede apreciar que los espectros NIR de las muestras de sedimento son claramente diferentes de los correspondientes a las muestras de la parte interna de los huesos. Además unos pocos espectros de la parte externa de los huesos están situados más cerca de los sedimentos, mientras que otros se encuentran más cerca de las muestras correspondientes a la parte interna de los huesos. Como ya se ha evidenciado en un estudio previo [28] esto indica que la parte externa de los huesos ha sufrido cambios más fuertes que la parte interna de los huesos debido, principalmente, a los factores digenéticos. En la dirección de PC2 se puede observar que los espectros de huesos calcinados de Corral de Saus están situados más hacia la zona baja del gráfico mientras que los huesos carbonizados más hacia arriba. Por otra parte, los espectros de los huesos de la Necrópolis de Las Peñas se localizan cerca de la parte interna de los huesos de Corral de Saus, indicando similitudes entre estos dos grupos de muestras; este mismo comportamiento se observa para los espectros materiales certificados de referencia correspondientes a hueso (NIST 1400). La Figura 9.2b muestra el gráfico de *loadings* para una mejor comprensión de la distribución de las muestras en el gráfico del PCA, indicando que los números de onda responsables de la clasificación obtenida son aquellos encontrados alrededor de 7000 cm^{-1} y entre 5500 y 4000 cm^{-1} . En consecuencia, se puede observar que el análisis PCA utilizando los espectros NIR ofrece una herramienta rápida, sencilla, directa y sostenible para identificar los cambios en los huesos causados por los factores ambientales.

3.9.3.2.2. Cluster

La Figura 9.3 muestra el dendrograma obtenido a partir de los espectros de las muestras de Corral de Saus, después de un tratamiento previo con FD y MC de y seleccionando la región de 9000 a 4000 cm^{-1} , utilizando el promedio pareado como medida de distancia. Se observa que para un valor de corte mínimo de 0.016, todas las muestras se clasificaron en dos grupos principales, que corresponden a sedimentos y huesos. El primer grupo incorpora todos los sedimentos, junto con 5 muestras de la parte externa de los huesos, probablemente debido a que estas muestras fueron alteradas por factores ambientales. El segundo grupo se basa en las muestras de la parte interna de los huesos, donde se pueden ver animales, huesos carbonizados y calcinados junto con muestras desconocidas y 2 partes externas de los huesos. Para un valor de corte de 0.012 las muestras se clasificaron como sedimentos, huesos de animales y parte interna de los huesos. Para un valor de corte de 0.0075, las muestras se clasificaron como huesos calcinados y carbonizados. Por tanto, se puede concluir que existen muestras similares a los sedimentos, como son la parte externa de los

huesos, y muestras diferentes al suelo, como los huesos calcinados que aparecen situados en la parte opuesta del dendrograma, de modo que el análisis cluster indica que los huesos calcinados presentan menos cambios estructurales que la parte carbonizada y externa de los huesos.

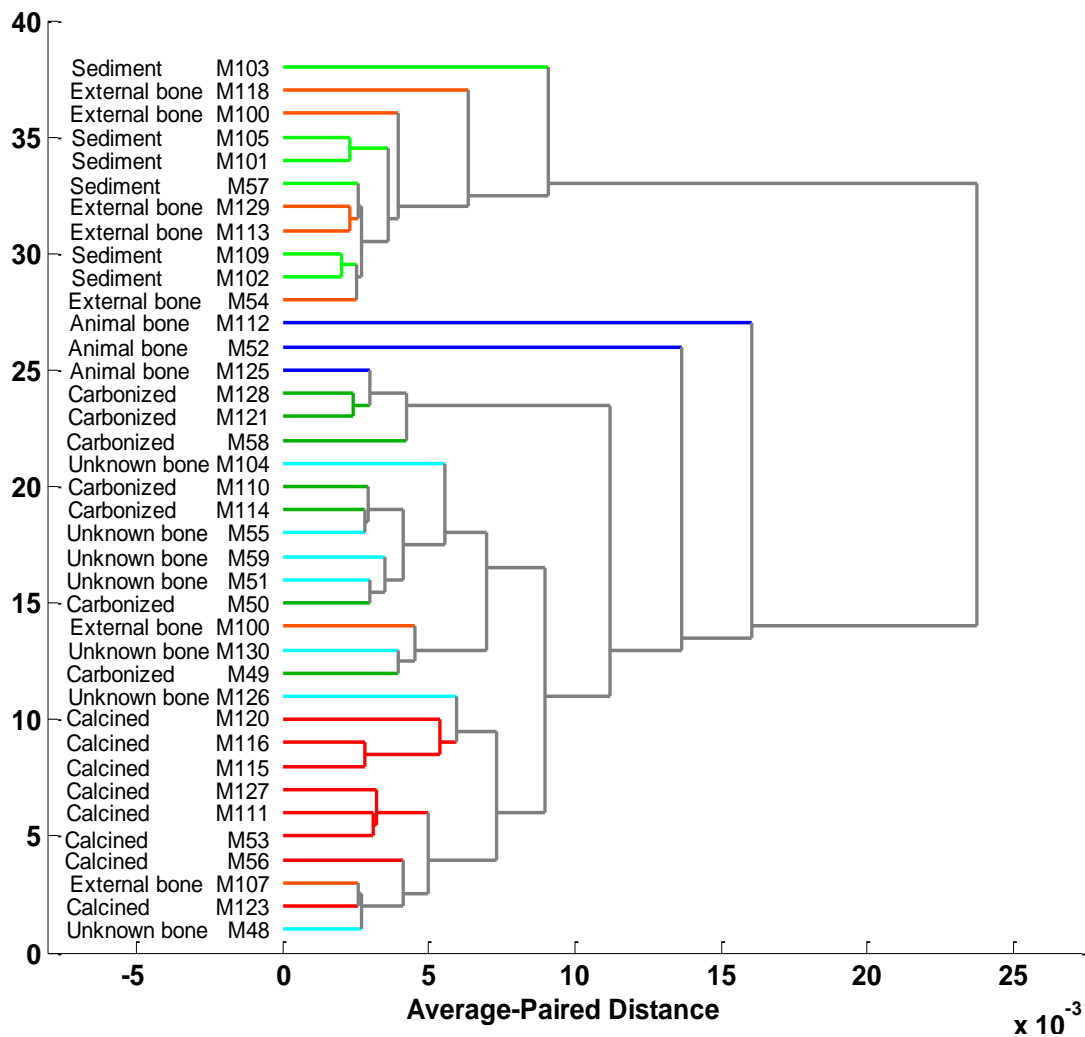


Figura 9.3. Clasificación dendrográfica obtenida a partir del análisis clúster de las muestras de huesos y sedimentos, a partir de sus espectros NIR después del tratamiento previo de FD y MC, en la región de 9000 a 4000 cm^{-1} , en relación con la clasificación de huesos calcinados, carbonizados, de animal, parte externa de huesos y sedimentos, utilizando el promedio pareado como medida de distancia

En la Figura 9.4 se ha incluido un dendrograma establecido a partir de espectros de la parte interna de los huesos y construido mediante el criterio de Kmeans nearest como medida de distancia. En este caso, se puede apreciar que para un valor de corte mínimo de 0.008, las muestras se clasificaron en los dos grupos principales, correspondientes a los de los huesos calcinados y carbonizados.

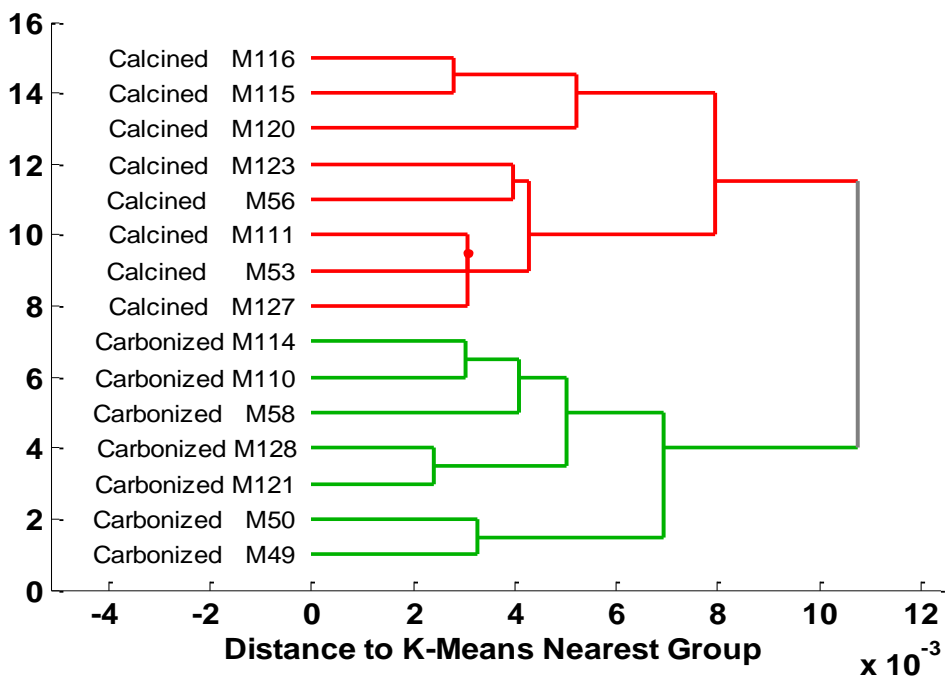


Figura 9.4. Clasificación dendrográfica obtenida a partir del análisis clúster de las muestras de huesos y sedimentos, a partir de sus espectros NIR después del tratamiento previo de FD y MC, en la región de 9000 a 4000 cm^{-1} , en relación con la clasificación de huesos calcinados y carbonizados usando el grupo K-means nearest como método de medida de la distancia

3.9.3.3. PLS-DA

El principal objetivo que se persigue con el PLS-DA de los espectros de las muestras es el de construir un modelo de calibración/predicción que pueda aplicarse en la clasificación de futuras muestras, ya que las asignaciones de las clases de huesos quemados no siempre es posible únicamente a partir de la comparación del color de los restos. El PLS-DA se puede aplicar para clasificar las muestras de hueso, dependiendo de sus condiciones de combustión, utilizándose para ello un conjunto de calibración formado por los huesos internos de Corral de

Saus, diferenciados en "calcinados" y "carbonizados", y aplicando el método de validación cruzada con el criterio "leave-one-out" en el conjunto de calibración para seleccionar el número óptimo de variables latentes del modelo PLS-DA. La Figura 9.5 ilustra la clasificación para PLS-DA, después del tratamiento previo de FD y MC para la región de los espectros comprendida entre 9000 y 4000 cm^{-1} .

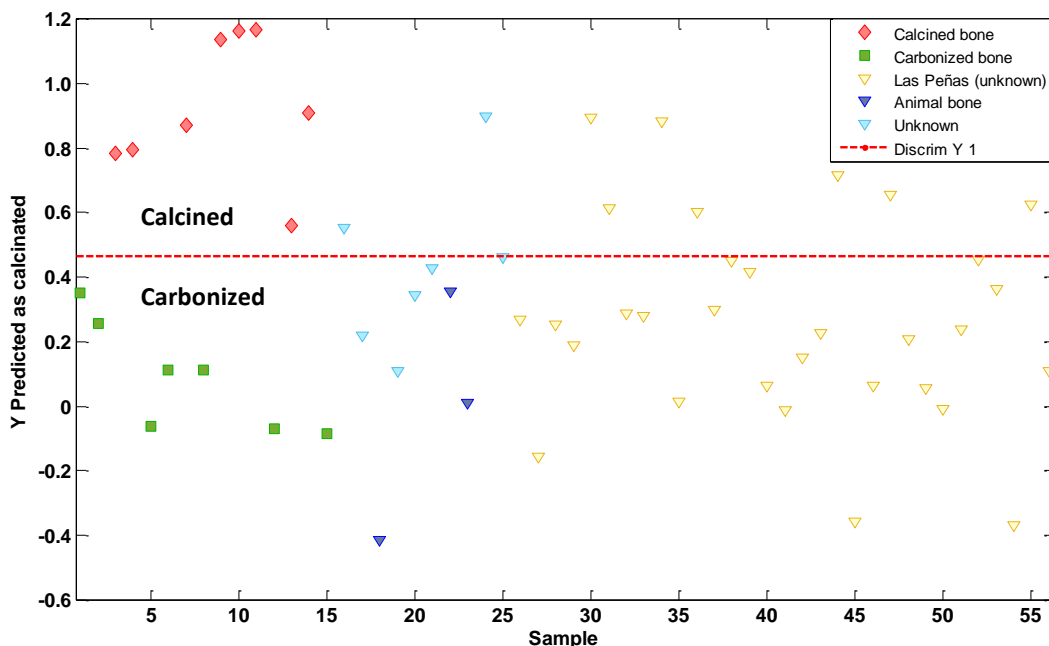


Figura 9.5. Clasificación por PLS-DA de los espectros NIR de los huesos, después del pre-tratamiento FD y MC en la región de 9000 a 4000 cm^{-1} .

Se puede observar que existe una buena separación de clases entre los huesos calcinados y carbonizados para el conjunto de calibración, utilizando únicamente una variable latente, que explica el 70.83% de la varianza total del bloque de datos X y el 83.56% del bloque de datos Y. Los resultados de sensibilidad y especificidad para la clasificación mediante PLS-DA fueron del 100% para el conjunto de calibración y para la validación cruzada, obteniéndose resultados satisfactorios para la clasificación de las muestras en verdaderos positivos y negativos, lo que indica una excelente separación de las clases. Tanto la eficiencia como la precisión fueron del 100% y el MCC tuvo un valor de 1. Además, el modelo PLS-DA fue validado empleando pruebas de permutación, usando 30 iteraciones y una variable latente. Se obtuvieron p-valores de 0.010 y 0.008 para la auto-predicción y la validación cruzada, respectivamente, mediante el empleo de una prueba de

asignación al azar para evaluar los residuales, lo que confirma la validez del modelo PLS-DA original a un nivel de confianza del 95%.

El modelo PLS-DA seleccionado se aplicó al estudio de los espectros NIR de muestras óseas sometidas a condiciones de combustión desconocidas, procedentes del Corral de Saus y la Necrópolis de Las Peñas, junto con muestras de huesos de animales no quemados. Como se puede apreciar en los resultados mostrados en la Figura 4, todas las muestras de hueso animal aparecen distribuidas de similar a los huesos carbonizados, mientras que algunos de los huesos desconocidos se asignaban a la clase de huesos calcinados y otros a la clase de huesos carbonizados.

Por lo tanto, se puede concluir que el PLS-DA permite clasificar de forma adecuada los huesos en calcinados o carbonizados, evitando las confusiones que en ocasiones supone la interpretación visual. Dado que los huesos calcinados son más resistentes a la alteración diagenética que los carbonizados debido a su alta cristalinidad, el PLS-DA permite asegurar la selección adecuada de huesos para los estudios forenses, tanto biológicos como los basados en el método del radiocarbono, además de otros estudios sobre la composición mineral de restos humanos relativos a aspectos culturales, dietas o hábitos de vida.

3.9.3.4. Modelos PLS

Los modelos de regresión PLS se construyeron para predecir, a partir de sus espectros NIR, la concentración de los elementos más importantes encontrados en las muestras de hueso estudiadas. Se utilizaron 38 muestras de la Necrópolis Corral de Saus como conjunto de calibración y, para confirmar la idoneidad de la estrategia del pre-procesamiento elegida, se utilizaron 31 muestras de la Necrópolis de Las Peñas como conjunto de validación externa. En la Tabla 9.2 aparecen indicados los valores promedio, la desviación estándar y el rango de concentraciones de los elementos Ca, Mg y Sr, tanto para el conjunto de calibración como el de validación. Además, una muestra de hueso Ceniza NIST 1400 se utilizó como material de referencia para evaluar la exactitud y precisión del procedimiento analítico.

Para construir los mejores modelos se estudiaron diferentes pre-procesamientos y regiones espectrales, siendo elegida la región comprendida entre 9000 y 4000 cm^{-1} y la FD (con una ventana de 15 puntos y utilizando un polinomio de segundo orden) junto con el tratamiento MC como pre-procesamiento de la señal para todos los elementos.

Tabla 9.2. Rangos de concentración de los conjuntos de calibración y validación de los elementos minerales determinados.

Analito	Conjunto	Muestras	Rango	Promedio	SD
Calcio (mg g^{-1})		35	250 - 461	360	49
Magnesio ($\mu\text{g g}^{-1}$)	Calibración	36	790 - 34257	8091	10820
Stroncio ($\mu\text{g g}^{-1}$)		35	161 - 1061	494	215
Calcio (mg g^{-1})		31	341 - 397	363	16
Magnesio ($\mu\text{g g}^{-1}$)	Validación	31	1213 - 4116	2014	711
Stroncio ($\mu\text{g g}^{-1}$)		29	179 - 1068	503	176

SD: desviación estándar

Se seleccionaron 3, 8 y 3 LVs para los modelos del calcio, magnesio y estroncio, respectivamente, con una varianza explicada de 72.62%, 95.9% y 87.61% para variables X y 72.77%, 99.61% y 80.90% para variables Y, respectivamente. Se identificaron como aberrantes, debido a sus valores elevados de Q residuales y T^2 Hotelling, 3 muestras en los modelos de calibración de Ca y Sr y 2 en el caso del Mg, que fueron eliminadas antes de la construcción de los modelos PLS. Además, se retiraron 2 muestras para el conjunto de validación de Sr.

La Figura 9.6 muestra la correlación entre los valores predichos y de referencia para las concentraciones de Ca, Mg y Sr para los conjuntos de calibración (muestras de la Necrópolis Corral de Saus) y validación (muestras de la Necrópolis de Las Peñas) y en la Tabla 9.3 se resumen Los principales parámetros de calibración y validación de los modelos PLS-NIR desarrollados, con unos buenos valores de los coeficientes de determinación R^2_{cal} y R^2_{CV} . Como puede verse, se obtuvieron valores aceptables de RRMSEP para el calcio, próximos al 5%; sin embargo, la predicción de los contenidos de Mg y Sr proporcionó valores de RRMSEP mucho mayores, especialmente en el caso del Mg. Los valores de RPD fueron 0.86, 0.24 y 1.08, para Ca, Mg y Sr, respectivamente.

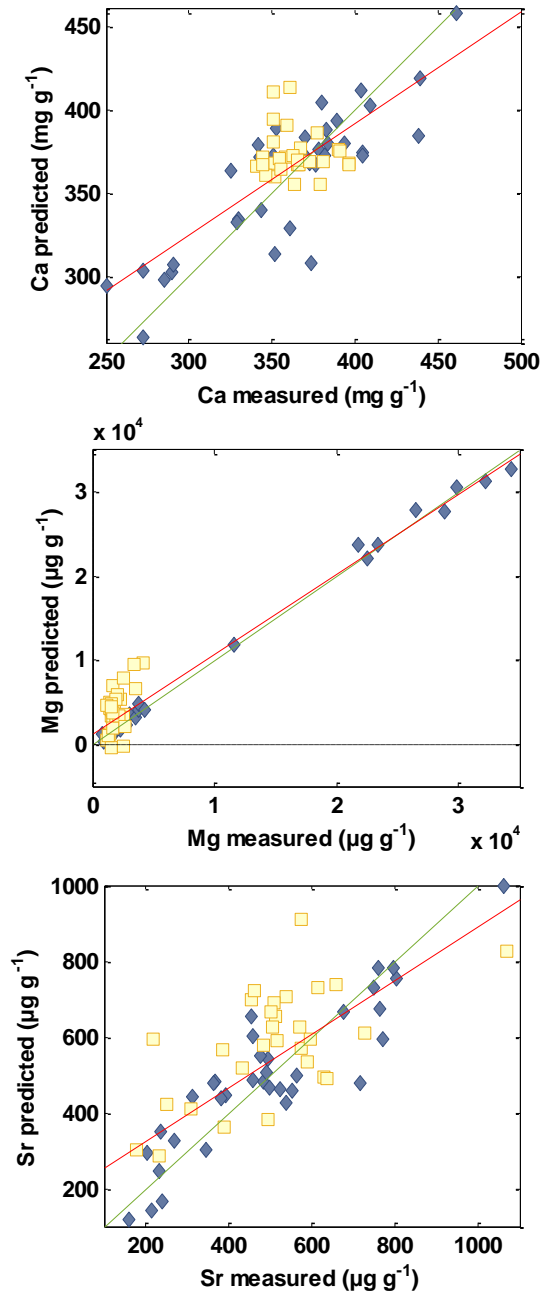


Figura 9.6. Regresiones entre los valores predichos y de referencia obtenidas para la determinación de calcio, magnesio y estroncio en restos óseos

- ◆ Conjunto de Calibración Corral
- Conjunto de Validación Las Peñas

Tabla 9.3. Parámetros de calibración y validación de los modelos PLS-NIR desarrollados para la determinación de calcio, magnesio y estroncio en restos óseos

Elemento	Pre-procesamiento	LV's	Conjunto de Calibración				Conjunto de Validación		
			RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	RRMSEP (%)	RPD
Calcio (mg g ⁻¹)	FD, MC	3	25.4	72.8	29.9	62.2	18.9	5.2 %	0.86
Magnesio (µg g ⁻¹)	FD, MC	8	665.7	99.6	3115.7	92.4	2941.5	146 %	0.24
Estroncio (µg g ⁻¹)	FD, MC	3	92.5	80.9	112.6	71.8	162.57	32.3 %	1.08

LV's: número de variables latentes

RMSEC: Error cuadrático medio de calibración; RMSECV: Error cuadrático medio de validación cruzada; RMSEP: Error cuadrático medio de predicción;

RRMSEP: Error relativo cuadrático medio de predicción; R² Cal: coeficiente de determinación de calibración; R² CV: coeficiente de determinación de validación cruzada;

R² Pred: coeficiente de determinación de predicción; RPD: desviación predictiva residual.

3.9.4. Control de calidad (QC)

El material de referencia certificado de ceniza de hueso NIST 1400 se midió en las mismas condiciones que las muestras, durante 18 días y en tres momentos diferentes de cada sesión, coincidentes con el comienzo, la mitad y el final de la misma. Con el fin de evaluar el método propuesto, se utilizaron la precisión inter-día y la exactitud de las mediciones obtenidas para esta muestra de referencia. La precisión se expresó como el coeficiente de variación de los resultados obtenidos durante 18 días diferentes y la exactitud se expresó como el sesgo entre el valor predicho y el valor certificado dividido por el valor certificado, expresado como porcentaje. Los coeficientes de variación (CV) varían de 0.2% hasta 1.8% para Ca y Sr, siendo los datos encontrados para Mg entre 4.7% y 13.8%. Por otro lado, el sesgo obtenido para el material de referencia fue inferior al 10% en el caso de determinación de Ca y Sr, con un intervalo de errores relativos entre -4.3% y -2.3% para Ca y desde el 6.9% hasta 9.9% para Sr, mientras que en el caso del Mg variaron del -26,5% hasta el -8,4% (véase resultados de la Tabla 9.4). Además, el PCA permitió evaluar cualitativamente la precisión del método mediante la visualización de la proyección de datos, debido al hecho de que las repeticiones se colocaron todas juntas y cerca de los huesos calcinados.

Tabla 9.4. Parámetros de precisión y exactitud de la metodología PLS-NIR propuesta para la determinación de calcio, magnesio y estroncio en el material de referencia ceniza de hueso NIST 1400.

	Ca	Mg	Sr
CV (%)	0.2 – 1.5	4.7 – 13.8	0.3 – 1.8
Er (%)	(-4.3) – (-2.3)	(-26.5) – (-8.4)	6.9 – 9.9

3.9.5. Evaluación del método propuesto

Como se ha explicado en las secciones anteriores, la metodología NIR desarrollada nos permitió identificar las muestras mejor conservadas y predecir Ca, Mg y Sr en los huesos quemados. Sin embargo, es necesario evaluar las ventajas y desventajas de la metodología propuesta frente al método de referencia, pudiéndose concluir que la metodología de espectroscopia NIR es más rápida que la ICP-OES para la preparación y medición de las muestras y para la obtención de los resultados. Sin embargo, la falta de sensibilidad del NIR es la principal limitación del método, siendo determinados únicamente los elementos mayoritarios de los huesos como Ca, Mg y Sr.

El procedimiento NIR propuesto y la metodología de referencia ICP-OES se evaluaron mediante comparación según el llamado Certificado Verde, ya explicado en capítulos precedentes de esta Tesis. La Figura 9.7 muestra los resultados obtenidos en esta evaluación verde de los procedimientos FT-NIR e ICP-OES.

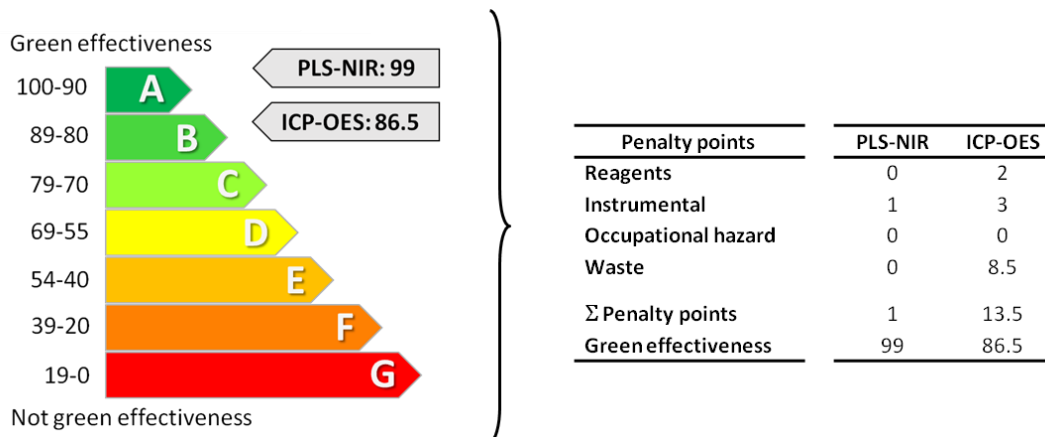


Figura 9.7. Evaluación verde de los procedimientos PLS-NIR e ICP-OES para la determinación de Ca, Mg y Sr en los restos óseos.

Se asignó un punto de penalización al método FT-NIR debido al consumo de energía por la mineralización de muestras en el horno de mufla previa a la obtención de los espectros de reflectancia difusa, obteniéndose una clasificación final en la categoría A del Certificado Verde, con una puntuación de 99. Para el ICP-OES, la puntuación final fue de 86.5 debido a que se asignaron 2 puntos de penalización por el uso de HNO_3 y HCl en el tratamiento de la muestra, el uso de soluciones estándar y de argón para la generación de plasma; 3 puntos de penalización por el consumo de energía por el horno de mufla y el ICP-OES, y 8.5 puntos de penalización por los residuos generados. Por lo tanto, el método ICP-OES debe ser considerado como una clase B en el ranking del *Green Certificate*.

3.9.6. Aplicaciones forenses

Los antropólogos forenses correlacionan la temperatura máxima de exposición de los restos óseos con los cambios de color para obtener una interpretación correcta de los huesos quemados. La estimación de esta temperatura podría ser una tarea difícil, dado que la observación de los cambios de color es un método subjetivo y confuso, debido a los diferentes factores que deben ser considerados como la

propia temperatura, el tiempo, la disponibilidad de oxígeno y factores externos y diagenéticos. Los huesos calcinados son más resistentes a alteraciones ambientales que los carbonizados, por lo que los huesos calcinados son más adecuados para los estudios de investigación sobre la datación con radiocarbono y la reconstrucción de la dieta mediante el uso de la proporción Sr/Ca. El análisis de PCA utilizando los espectros NIR ofrece una herramienta rápida y verde para identificar los cambios en los huesos causados por factores ambientales, que permitiría una adecuada selección de las muestras para estudios posteriores, evitando errores interpretativos relacionados con la estructura y composición química de los huesos que pueden haber sido modificados post-mortem por factores diagenéticos.

Por otra parte, el uso de PLS-NIR como metodología de cribado para la estimación de Ca, Mg y Sr ofrece un valor adicional al empleo de los espectros NIR en los estudios de restos óseos.

3.9.7. Conclusiones Capítulo 9

Se obtuvieron resultados interesantes utilizando PCA, análisis cluster y PLS-DA para clasificar los huesos quemados a partir de los espectros NIR. El PCA, utilizando dos componentes principales, fue capaz de discriminar los restos óseos afectados por los procesos diagenéticos post-deposicionales de los relativamente libres, debido a la falta de intercambios con los componentes del suelo. El HCA permitió diferenciar claramente huesos calcinados y carbonizados y podría ser útil para entender de una manera más o menos clara el tratamiento térmico que hayan podido experimentar restos óseos desconocidos. Por su parte, el PLS-DA proporcionó una herramienta exacta para discriminar entre huesos calcinados y carbonizados, permitiendo clasificar con precisión los huesos quemados desconocidos.

A partir de un estudio previo sobre huesos enterrados de dos necrópolis romanas tardías [51], el uso de la espectroscopia NIR combinada con la quimiometría proporcionó un método rápido y adecuado para determinar la concentración de calcio, magnesio y estroncio en huesos quemados, con el objetivo de comprender la necrosis post-mortem a partir de los cambios producidos como consecuencia de la degradación ambiental.

En resumen, se puede concluir que las metodologías propuestas, basadas en el uso de la espectroscopia NIR combinada con herramientas quimiométricas, proporcionan un enfoque rápido y sostenible para seleccionar las muestras más adecuadas para realizar estudios forenses y evaluar los elementos minerales más

importantes en restos óseos. Este interés resulta muy importante debido que los huesos carbonizados son más propensos a los procesos post-deposicionales que los calcinados, produciendo importantes errores a partir de las interpretaciones de los análisis realizados sobre los restos óseos. Además, los modelos de regresión PLS construidos a partir de los espectros NIR proporcionaron una herramienta de predicción, rápida y sencilla, para estimar los contenidos de Ca, Mg y Sr en los huesos quemados.

3.9.8. Referencias Capítulo 9

- [1] S.A. Symes, D.C. Dirkmaat, S.Ousley, E. Chapman, L. Cabo, *The Analysis of Burned Human Remains*, San Diego, US Department of Justice, 2012.
- [2] T.J.U. Thompson, Heat-induced Dimensional Changes in Bone and their Consequences for Forensic Anthropology, *J Forensic Sci*, 50(2005) 1008–15.
- [3] D. Gonçalves, T.J.U. Thompson, E. Cunha, Implications of heat-induced changes in bone on the interpretation of funerary behaviour and practice, *J. Archaeol. Sci.*, 38(2011) 1308–1313.
- [4] T.J.U. Thompson, Recent advances in the study of burned bone and their implications for forensic anthropology, *Forensic Sci. Int.*, 146(2004) Supplement, S203–S205
- [5] D.H. Ubelaker, The forensic evaluation of burned skeletal remains: A synthesis, *Forensic Sci. Int.*, 183(2009) 1–5.
- [6] S.T.D. Ellingham, T.J.U. Thompson, M. Islam, G. Taylor, Estimating temperature exposure of burnt bone — A methodological review, *Sci. Justice*, 55(2015) 181–188.
- [7] S.A. Symes, C.W. Rainwater, E.N. Chapman, D.R. Glipson, A.L. Piper, Chapter 2. Patterned thermal destruction of human remains in a forensic setting, in *The Analysis of Burned Human Remains*, San Diego: Academic Press, 2008.
- [8] P. Shipman, G. Foster, M. Schoeninger, Burnt bones and teeth: an experimental study of color, morphology, crystal structure and shrinkage, *J. Archaeol. Sci.*, 11(1984) 307–325.
- [9] M. Color, *Munsell Soil Color Charts*. New Windsor, N.Y, Macbeth Division of Kollmorgan Instruments Corp, NY, 1994.
- [10] J. B. Devlin, N. P. Herrmann, Chapter 6. Bone color as an interpretive tool of the depositional history of archaeological cremains, in *The Analysis of Burned Human Remains*, San Diego: Academic Press, 2008.
- [11] R.F. Castillo, D.H. Ubelaker, J.A.L. Acosta, G.A.C. de la Fuente, Effects of temperature on bone tissue. Histological study of the changes in the bone matrix, *Forensic Sci. Int.*, 226(2013) 33–37.
- [12] P.L. Walker, K.P. Miller, Time, temperature, and oxygen availability: an experimental study of the effects of environmental conditions on the color and organic content of cremated bone, *Am. J. Phys. Anthropol. Suppl.* 40 (2005) 216–217.

- [13] R. A. Nicholson, A Morphological Investigation of Burnt Animal Bone and an Evaluation of its Utility in Archaeology, *J. Archaeol. Sci.*, 20(1993) 411–428.
- [14] S.A. Symes, C.W. Rainwater, E.N. Chapman, D.R. Gipson, A.L. Piper, Patterned thermal destruction of human remains in a forensic setting, in: C.W. Schmidt, S.A. Symes (Eds.), *The Analysis of Burned Human Remains*, Academic Press, London, 2008, 15–54.
- [15] J.B. Devlin, N.P. Herrmann, Bone color as an interpretive tool of the depositional history of archaeological cremains, in: C.W. Schmidt, S.A. Symes (Eds.), *The Analysis of Burned Human Remains*, Academic Press, London, 2008, 109–128.
- [16] R. Shahack-Gross, O. Bar-Yosef, S. Weiner, Black-coloured bones in Hayonim Cave, Israel: differentiating between burning and oxide staining, *J. Archaeol. Sci.*, 24(1997), 439–446.
- [17] C.K. Brain, A. Sillen Evidence from the Swartkrans cave for the earliest use of fire, *Nature*, 336 (1988), 464–466.
- [18] K.A. Gross, C.C. Berndt, Biomedical Application of Apatites, *Rev. Mineral. Geochem.*, 48(2002) 631–672.
- [19] J.N. Lanting, A.T. Aerts-Bijma, J. van der Plicht, Dating cremated bone, *Radiocarbon* 43(2001) 249–254.
- [20] T.A. Surovell, Radiocarbon dating of bone apatite by step heating, *Geoarchaeology*, 15(2000) 591–608.
- [21] G. Quarta, L. Calcagnile, M. D’Elia, L. Maruccio, V. Gaballo, A. Caramia, A combined PIXE–PIGE approach for the assessment of the diagenetic state of cremated bones submitted to AMS radiocarbon dating, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.*, 294(2013) 221–225.
- [22] A. Zazzo, J.-F. Saliège, Radiocarbon dating of biological apatites: A review, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 310(2011) 52–61.
- [23] J. Olsen, K. M. Hornstrup, J. Heinemeier, P. Bennike, H. Thrane, Chronology of the Danish Bronze Age Based on ¹⁴C Dating of Cremated Bone Remains, *Radiocarbon*, 53(2011) 261–275.
- [24] J. Olsen, J. Heinemeier, P. Bennike, C. Krause, K. Margrethe Hornstrup, and H. Thrane, Characterisation and blind testing of radiocarbon dating of cremated bone, *J. Archaeol. Sci.*, 35(2008) 791–800.

- [25] G. Grupe, S. Hummel, Trace element studies on experimentally cremated bone. I. Alteration of the chemical composition at high temperatures, *J. Archaeol. Sci.*, 18(1991) 177-186.
- [26] J.J. Schultz, M.W. Warren, J.S. Krigbaum, Chapter 4. Analysis of human cremains: gross and chemical methods, in *The Analysis of Burned Human Remains*, San Diego: Academic Press, 2008.
- [27] M. E. Subira, A. Malgosa, The effect of cremation on the study of trace elements, *Int. J. Osteoarchaeol.*, 3(1993) 115–118.
- [28] G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, J. Bernabeu, Biological mineral content in Iberian skeletal cremains for control of diagenetic factors employing multivariate statistics, *J. Archaeol. Sci.*, 40(2013) 2477–2484.
- [29] L.D. Mkukuma, J.M.S. Skakle, I.R. Gibson, C.T. Imrie, R.M. Aspden, D.W.L. Hukins, Effect of the Proportion of Organic Material in Bone on Thermal Decomposition of Bone Mineral: An Investigation of a Variety of Bones from Different Species Using Thermogravimetric Analysis coupled to Mass Spectrometry, High-Temperature X-ray Diffraction, and Fourier Transform Infrared Spectroscopy, *Calcif. Tissue Int.*, 75(2004) 321–328.
- [30] T. Leventouri, Synthetic and biological hydroxyapatites: Crystal structure questions, *Biomaterials*, 27(2006) 3339–3342.
- [31] J.D. Pasteris, B. Wopenka, J.J. Freeman, K. Rogers, E. Valsami-Jones, J.A.M. van der Houwen, M.J. Silva, Lack of OH in nanocrystalline apatite as a function of degree of atomic order: implications for bone and biomaterials, *Biomaterials*, 25(2004) 229–238.
- [32] B. Wopenka, J.D. Pasteris, A mineralogical perspective on the apatite in bone, *Mater. Sci. Eng. C*, 25(2005) 131–143.
- [33] *Apatites and their Synthetic Analogues-Synthesis, Structure, Properties and Applications*, Petr Ptacek, 2016.
- [34] T.J.U. Thompson, M. Gauthier, M. Islam, The application of a new method of Fourier Transform Infrared Spectroscopy to the analysis of burned bone, *J. Archaeol. Sci.*, 36(2009) 910–914.
- [35] T.J.U. Thompson, M. Islam, K. Piduru, A. Marcel, An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 299(2011) 168–174.

- [36] M. Lebon, I. Reiche, J.J. Bahain, C. Chadeaux, A.M. Moigne, F. Fröhlich, F. Sémah, H.P. Schwarcz, C. Falguères, New parameters for the characterization of diagenetic alterations and heat-induced changes of fossil bone mineral using Fourier transform infrared spectrometry, *J. Archaeol. Sci.*, 37(2010) 2265–2276.
- [37] J.M. Martínez García, La necrópolis ibérica de Las Peñas (Zarra, Valencia), *Archivo de Prehistoria Levantina.*, vol. XIX. Valencia, 1989.
- [38] D. Ballabio, A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure, *Chemom. Intell. Lab. Syst.*, 149(2015), Part B, 1–9.
- [39] M.J. Hidalgo, D.C. Fechner, E.J. Marchevsky, R.G. Pellerano, Determining the geographical origin of *Sechium edule* fruits by multielement analysis and advanced chemometric techniques, *Food Chem.*, 210(2016) 228–234.
- [40] M.R. Almeida, C.H.V. Fidelis, L.E.S. Barata, R.J. Poppi, Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation, *Talanta*, 117(2013) 305–311.
- [41] F.B. de Santana, L.C. Gontijo, H. Mitsutake, S.J. Mazivila, L.M. de Souza, W. Borges Neto, Non-destructive fraud detection in rosehip oil by MIR spectroscopy and chemometrics, *Food Chem.*, 209(2016) 228–233.
- [42] P.C. Williams, D. Sobering How do we do it: A brief summary of the methods we use in developing near infrared calibrations A.M.C. Daves, P.C. Williams (Eds.), *Near infrared spectroscopy: The future waves*, NIR Publications, Chichester, UK (1995) 185–188.
- [43] R.L. Frost, B.J. Reddy, S. Bahfenne, J. Graham, Mid-infrared and near-infrared spectroscopic study of selected magnesium carbonate minerals containing ferric iron—Implications for the geosequestration of greenhouse gases, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, 72(2009) 597–604.
- [44] Y. Ning, J. Li, W. Cai, X. Shao, Simultaneous determination of heavy metal ions in water using near-infrared spectroscopy with preconcentration by nano-hydroxyapatite, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, 96(2012) 289–294.
- [45] V. Aranda, A. Domínguez-Vidal, F. Comino, J. Calero, M.J. Ayora-Cañada, Agro-environmental characterization of semi-arid Mediterranean soils using NIR reflection and mid-IR-attenuated total reflection spectroscopies, *Vib. Spectrosc.*, 74(2014) 88–97.

- [46] D. Thomas, C. McGoverin, A. Chinsamy, M. Manley, Near infrared analysis of fossil bone from the Western Cape of South Africa, *J. Infrared Spectrosc.*, 19(2011) 151, 2011.
- [47] M. Mora, M. Isabel López, M. Ángeles Carmona, C. Jiménez-Sanchidrián, J. Rafael Ruiz, Study of the thermal decomposition of a sepiolite by mid- and near-infrared spectroscopies, *Polyhedron*, 29(2010) 3046–3051.
- [48] K. Van Aken, L. Strekowski, L. Patiny, EcoScale, a semi-quantitative tool to select an organic preparation based on economical and ecological parameters *Beilstein Journal of Organic Chemistry*, 2 (2006), 3.
- [49] A. Galuszka, P. Konieczka, Z.M. Migaszewski, J. Namiésnik, Analytical eco-scale for assessing the greenness of analytical procedures, *Trac-Trend in Analytical Chemistry*, 37 (2012), 61–72.
- [50] S. Armenta, M. de la Guardia, J. Namiesnik. Green Microextraction. In R. L. M. Valcárcel, S. Cárdenas, ed. *Analytical Microextraction Techniques*. Bentham Science, 2017.
- [51] M. M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Prediction of alkaline earth elements in bone remains by near infrared spectroscopy, *Talanta*, 162(2017) 428–434.

4. CONCLUSIONES

Conclusiones

De los estudios realizados durante el desarrollo de esta Tesis para la aplicación de la espectroscopia MIR, NIR y Raman en tres ámbitos diferentes, como son las muestras de alimentos, muestras ambientales y muestras arqueológicas, se extraen las siguientes conclusiones generales y específicas.

Conclusiones generales

Una de las grandes ventajas de la espectroscopia vibracional, es su versatilidad, ya que permite el análisis de una gran variedad de muestras con independencia del estado físico en el que se encuentren.

Cuando no es posible disponer de calibrados externos adecuados, el análisis de muestras mediante la espectroscopia MIR, NIR o Raman, combinada con el tratamiento quimiométrico de los datos, requiere utilizar un conjunto de muestras de referencia que tenga las mismas características que las muestras a analizar, siendo todas ellas representativas del conjunto de datos. Además, si se trata de un análisis cuantitativo es necesario un método de referencia para conocer las propiedades y concentraciones de los analitos, siendo normalmente metodologías que precisan de etapas previas de pre-tratamiento de la muestra, uso de reactivos y disolventes, generan residuos y suponen un coste y consumo de tiempo elevados. Como alternativa a estas metodologías se han desarrollado procedimientos que combinan la espectroscopia vibracional con métodos quimiométricos que permiten extraer la información requerida a partir del tratamiento de datos y determinar, de manera simultánea, varios analitos, con escasa preparación y manipulación de las muestras.

La espectroscopia NIR, MIR y Raman permiten la inspección y el análisis de las muestras prácticamente en tiempo real ya sea en la línea de producción o a través del producto final, de una manera rápida y sencilla, tratándose en muchos casos de análisis no destructivos, por lo que se consideran metodologías ideales para el control de calidad de materias primas y control de calidad de productos acabados, en el laboratorio y durante el proceso de fabricación. La espectroscopia NIR es menos útil para la identificación de compuestos que la espectroscopia MIR, y más útil para el análisis cuantitativo. Si bien la espectroscopia MIR ha sido la técnica vibracional más utilizada en años anteriores para la identificación de materiales, está ganando fuerza en el análisis cuantitativo. En cuanto la espectroscopia NIR, es necesario trabajar con mayor cantidad de muestra que la requerida en el MIR, pero al igual que el Raman, pueden utilizarse sin la necesidad de manipular o poner en contacto las muestras con el material de medida, empleando contenedores de vidrio y evitando así una posible contaminación de las muestras.

Debido a que no se requieren ni reactivos ni disolventes, no se crea ningún residuo peligroso, por lo que también se puede reducir el coste de las pruebas, aún más si se trata de análisis de rutina.

Además de las conclusiones antedichas también se incluyen las conclusiones específicas ordenadas según los trabajos que integran la presente memoria:

Conclusiones específicas

Parte 1. Muestras de alimentos

Capítulo 1. Determinación directa de componentes mayoritarios en alimentos

El objetivo fijado de este estudio fue evaluar la espectroscopia NIR y MIR junto con herramientas quimiométricas, para la determinación directa de componentes mayoritarios (grasa total, proteínas y carbohidratos) y valor energético total de alimentos elaborados como tarritos de comida para bebés, menús infantiles de comida rápida y menús para adultos de comedores colectivos.

A partir de los resultados obtenidos se puede concluir que los modelos PLS-NIR construidos después de la selección del conjunto de calibración mediante el algoritmo K-S y una correcta selección de regiones espectrales, pre-procesado y número de variables latentes, proporcionan una herramienta rápida y sostenible para el análisis de grasa, proteínas, carbohidratos y valor energético en muestras sin la necesidad de utilizar reactivos ni disolventes.

Capítulo 2. Determinación de compuestos polares en aceites

En el estudio sobre la comparación de la espectroscopia FT-NIR y FT-MIR para la determinación de compuestos polares en aceites de fritura usados, los resultados encontrados indican que tanto las medidas por transmisión en el NIR como las de reflectancia total atenuada en el MIR, combinadas con el uso de la calibración multivariante PLS, pueden ser consideradas como una alternativa rápida, directa y sostenible para la determinación de TPM en aceites de fritura. Además, el método PCA facilita la clasificación de las muestras distinguiendo aquellas con un alto contenido de TPM de las de bajo contenido a partir del gráfico de puntuaciones. En resumen, estos estudios evidencian que el MIR y especialmente la técnica NIR ofrecen una alternativa para la evaluación de calidad de muestras de aceite frito con errores de predicción aceptables.

El estudio detallado de la determinación de TPM en aceites a través de las medidas de transmisión en el NIR ha permitido, a partir de la reducción del número de variables de los espectros, comparar las técnicas de MLR y PLS, para la selección de un número reducido de variables que permita escoger únicamente 4 números de onda para llevar a cabo la determinación con una adecuada precisión y capacidad predictiva.

El uso de modelos quimiométricos, aplicado a los espectros NIR de transmisión de aceites fritos, permite la determinación exacta de TPM con valores de RRMSEP de 6.5% y RPD de 9.2. Además el modelo MLR-NIR, construido a partir de espectros NIR con variables reducidas, proporcionó una buena correlación entre los 4 números de onda seleccionados y el contenido de TMP medido por el método de referencia, con un valor R^2_{pred} de 0.96 y un RPD de 5.0. Estos resultados son adecuados para fines analíticos con una precisión de predicción excelente pudiendo ser utilizados para diseñar un instrumento de bajo coste fabricado con un sistema de filtros o LED's que permitiera trabajar a los cuatro número de onda seleccionados para una evaluación rápida y barata de la calidad del aceite de fritura

Capítulo 3. Determinación de lípidos en aceite de salmón

En el estudio centrado en buscar alternativas al hexano como disolvente para la extracción de lípidos en salmón, se llevaron a cabo estudios sobre la solubilidad relativa de los componentes del pescado en diferentes disolventes mediante los métodos de *Hansen Solubility Parameter* (HSP) y *Conductor-like Screening Model for Real Solvent* (COSMO-RS). Además se realizaron estudios experimentales para comparar la eficiencia de disolventes alternativos al hexano.

Los estudios de simulación de Hansen y Cosmo-RS sobre el carácter de solvatación de los disolventes, para una serie de clases del aceite salmón indicaron que el d-limoneno y el p-cimeno eran similares al hexano para solvatar los TAGs frente a otros compuestos lipídicos. Además, los valores RED de Solubilidad de Hansen predijeron que el CPME podría ser el mejor disolvente para solubilizar todos los compuestos considerados, resultando el resto de los disolventes más específicos para las distintas clases de lípidos. Sin embargo, los valores obtenidos de COSMO-RS indican que CPME, EtAc y MeTHF podrían solvatar de manera similar a TAGs, DAGs, FFAs y al ergosterol.

Al considerar los datos experimentales obtenidos en el análisis por HPTLC y GC-FID, los resultados por HPTLC indican que los TAGs es la clase predominante de los lípidos extraídos con una riqueza entre 73-77%, seguido de DAGs (15-20%), FFAs

(5-6%) y ergosterol (2-3%). Sin embargo, se encontró que los extractos obtenidos empleando p-cimeno y limoneno contenían mayores niveles de DAGs y, por el contrario, cantidades más reducidas de TAGs que los extractos correspondientes a los otros disolventes. Este fenómeno podría explicarse por la degradación lipídica debida al uso de las temperaturas para trabajar con p-cimeno y limoneno, a pesar de los prometedoros valores de predicción proporcionados por HSP y COSMO-RS, que indicaban que tanto el p-cimeno y como el limoneno eran, teóricamente, los mejores disolventes para extraer los TAGs del salmón. Por otra parte, las cantidades de los ácidos grasos extraídos, determinados por GC-FID, indicaron que todos los disolventes proporcionan un perfil de ácidos grasos similar, excepto el limoneno y, especialmente, el p-cimeno, que daban lugar a la presencia de pequeñas cantidades de algunos ácidos grasos insaturados, probablemente debido al hecho de que los dobles enlaces de los PUFAs son más propensos a la oxidación debido a las temperaturas más elevadas implicadas en el uso de limoneno y p-cimeno.

La estrategia empleada en este estudio podría utilizarse como una guía para buscar disolventes verdes alternativos al hexano para la extracción de lípidos.

En resumen, se puede concluir que existen algunas diferencias entre los valores de solvatación teóricos predichos y los estudios experimentales sobre el uso de disolventes alternativos para la extracción de lípidos, pero que estas diferencias podrían deberse a que las simulaciones computacionales no consideran algunos aspectos como el tiempo y temperatura de extracción, u otros parámetros como la entalpía de vaporización o el punto de fusión de los compuestos considerados. Estas condiciones pueden modificar la fracción lipídica extraída.

En general, los resultados experimentales mostraron que EtAc era el disolvente más adecuado para reemplazar el hexano, debido a sus propiedades para extraer todas las clases de lípidos, tales como TAGs, DAGs, FFAs y ergosterol, en las muestras de salmón. Dado que se han considerado otros parámetros como punto de ebullición, Log P, categoría de toxicidad y energía requerida para la evaporación del disolvente; se puede concluir que EtAc podría ser un buen disolvente alternativo para la extracción de lípidos del salmón, ya que EtAc no proporciona emisiones de compuestos orgánicos volátiles, es un disolvente biodegradable y no tiene efectos tóxicos.

La estrategia expuesta podría ser útil para evaluar disolventes alternativos en otros campos con el fin de mantener la capacidad de extracción de los disolventes clásicos, pero mejorando la sostenibilidad de los procesos, y evitando efectos secundarios.

Considerando los extractos lipídicos del salmón obtenidos en la parte experimental del estudio previo, se consideró la posibilidad de evaluar la técnica NIR en combinación con los modelos PLS para la determinación simultánea de i) ácidos grasos, como oleico, palmítico, linoleico y linolénico, ii) familias de ácidos grasos, como omega-3, omega-6, ácidos grasos poliinsaturados, ácidos grasos monoinsaturados y ácidos grasos saturados, y iii) clases de lípidos, como triacilglicéridos, diacilglicéridos, ácidos grasos libres y ergosterol en aceite de salmón.

El método PLS-NIR desarrollado se basa en medidas de transmisión, no destructivas y permite la cuantificación simultánea de varios ácidos grasos, como OA, PA, ALA y LA así como omega-3, omega-6, PUFAs, MUFAs y SFAs, además de clases de lípidos como TAGs, DAGs, FFAs y ergosterol en el aceite de salmón, evitando el uso de reactivos tóxicos o peligrosos junto con una drástica reducción en el tiempo y el consumo de energía. Sin embargo, la exactitud y la robustez del método mantienen el nivel de detección para el ácido palmítico, linolénico y MUFAs. En resumen, se puede concluir que los modelos PLS-NIR ofrecen una herramienta de alto valor de detección para la caracterización de muestras de aceite de pescado de salmón con una buena capacidad predictiva para ergosterol, FFAs, DAGs, TAGs, omega-3 y PUFAs.

Capítulo 4. Control de calidad de los aditivos umami por Raman y espectroscopia infrarroja en la región cercana

El objeto de este estudio fue el desarrollo de métodos analíticos para la determinación de MSG, GMP e IMP en aditivos alimentarios industriales de Umami, basados en medidas directas por espectroscopia vibracional.

Las técnicas de espectroscopia NIR y Raman proporcionan alternativas rápidas y ecológicas para determinar simultáneamente MSG, GMP e IMP en aditivos umami. Para mezclas binarias compuestas por GMP e IMP, ambas técnicas combinadas con PLS ofrecen procedimientos precisos para el análisis cuantitativo de estas muestras. Sin embargo, para las mezclas ternarias compuestas por MSG, GMP e IMP, los resultados obtenidos mediante la espectroscopia NIR fueron mejores que los encontrados por Raman. En resumen, estos estudios ponen de manifiesto que para el análisis de aditivos umami, PLS-NIR ofrece una herramienta precisa para la determinación de MSG, GMP e IMP, siendo una clara alternativa verde a los métodos tradicionales basados en el empleo de las técnicas cromatográficas, que presentan un gran consumo de tiempo y un coste considerable, además de requerir la disolución previa de la muestra.

La espectroscopia NIR permite obtener resultados en un periodo de tiempo relativamente bajo y sin pre-tratamiento de muestras, siendo un buen método alternativo para la inspección rápida de preparados de aditivos alimentarios para asegurar la calidad de las mezclas umami antes de su uso en la elaboración de alimentos.

Parte 2. Muestras ambientales y de interés agrícola

Capítulo 5. Evaluación de los compost para su posterior uso en la agricultura mediante espectroscopia infrarroja NIR y MIR

Los métodos de calibración multivariante desarrollados utilizando tanto la espectroscopia MIR como NIR, ofrecen una alternativa sostenible para la determinación de compuestos fenólicos totales en muestras de compost. Sin embargo, al comparar las principales características analíticas de los dos métodos desarrollados, se puede concluir que el NIR proporciona un poder de predicción mejor en comparación con el MIR, si se tienen en cuenta los valores de RPD y el RMSEP para el conjunto de validación, que fue de 0.19 mg de equivalente de ácido gálico por gramo de masa seca ($\text{mg GAE g}^{-1} \text{DW}$) para NIR y de 0.27 $\text{mg GAE g}^{-1} \text{DW}$ para MIR. La principal conclusión es que se pueden emplear las dos metodologías propuestas para la determinación de TPC, siendo las medidas NIR más rápidas y seguras debido a que las muestras se introducen en viales de vidrio, mientras que en el MIR es necesario aplicar un protocolo de limpieza adecuado para evitar la contaminación cruzada entre muestras durante el proceso de medida sobre el cristal ATR. Sin embargo, está claro que los modelos con espectros NIR proporcionaron una mejor modelización (valores más bajos de calibración, validación cruzada y predicción) que los obtenidos a partir de espectros ATR-MIR.

Por lo que se refiere a la predicción de los contenidos de carbono orgánico y de nitrógeno total en muestras de compost y residuos orgánicos por espectroscopia infrarroja, se puede concluir que los métodos basados en PLS-MIR proporcionan mejores resultados que el PLS-NIR, sin que en ninguno de los dos casos puedan considerarse excelentes. En ese sentido habría que destacar que ofrecen alternativas verdes sostenibles para la determinación de OC y TN, adecuadas para fines de cribado. El procedimiento NIR está libre del uso de reactivos y residuos, mientras que el procedimiento MIR requiere pequeñas cantidades de disolventes de limpieza, produciéndose volúmenes reducidos de residuos.

Capítulo 6. Determinación de benceno, tolueno, etilbenceno y xilenos en suelos mediante espectroscopia infrarroja en la región media en fase de vapor

El objetivo de este trabajo era el de desarrollar una metodología simple que combine el uso del muestreo HS con la detección IR para una determinación rápida y directa de BTEX en muestras de suelo, evitando el uso de disolventes y reactivos y sin la necesidad de una separación cromatográfica. Para ello se utilizó el tratamiento quimiométrico de los espectros medidos en fase vapor para la determinación del benceno, tolueno, etilbenceno, orto-, meta- y para-xileno eliminando las interferencias espectrales del agua y del CO₂, utilizando una metodología desarrollada en nuestro laboratorio.

Los resultados indicaron que la espectroscopia infrarroja, combinada con técnicas de calibración multivariante, proporciona una herramienta adecuada para la determinación simultánea de compuestos BTEX en suelos contaminados de forma rápida y sin ningún pre-tratamiento químico de la muestra, si bien es cierto que su sensibilidad no es comparable a la de los métodos cromatográficos.

El uso del espacio de cabeza, para la generación del vapor del analito, reduce los errores experimentales asociados a la etapa de preparación y disminuye el tiempo de análisis a sólo 7 minutos, incluyendo el requerido para el calentamiento y medida de cada muestra. Además, el tiempo de análisis puede reducirse cuando se calientan juntas varias muestras. Por lo tanto, se ha demostrado que el HS-FT-IR ofrece una técnica verde adecuada para la determinación precisa de BTEX mediante calibración PLS con coeficientes de variación entre 0.9-11.2, 0.4-4.7, 1.2-7.2, 0.9-9.3, 2.0-7.0 y 0.4-8.0 para el benceno, tolueno, etilbenceno, orto-, meta- y para-xileno, respectivamente. Para una muestra real de suelo contaminado, tomada de un área de derrame de combustible, los modelos PLS en la región 3258-3001 y 913-567 cm⁻¹ y pre-tratamiento de primera derivada y centrado en la media fueron la mejor opción para determinar benceno, tolueno, etilbenceno, orto- y para-xileno, mientras que para la determinación del meta-xileno se requiere el empleo de la segunda derivada y el centrado en la media de los datos como pre-tratamiento en la región de 913-567 cm⁻¹.

Capítulo 7. Correcciones de vapor en espectroscopia infrarroja

La presencia de interferencias espectrales supone un problema para la obtención de resultados exactos cuando se emplea la espectroscopia infrarroja. Por ese motivo, el objetivo de este estudio era evaluar un método flexible y automatizado para evitar las interferencias espectrales de los vapores orgánicos en los espectros

FTIR medidos en fase de vapor, tomando acetona como compuesto modelo. Para ello se utilizó una metodología verde, que combina el uso del muestreo de espacio de cabeza (HS) con las medidas de transmisión por espectroscopia FTIR para el análisis de muestras conteniendo benceno, tolueno y para-xileno y etanol.

Los resultados obtenidos evidenciaron que el método propuesto para la compensación de la presencia de acetona, aplicado a los espectros de muestras de benceno, tolueno, para-xileno y etanol en fase vapor proporciona una buena correlación entre los espectros en ausencia de interferencia y los espectros corregidos. Los valores de R son cercanos a 1, siendo los valores de Z después de la corrección entre 3.8 y 5.4 y entre 2.71 y 4.7 para etanol y benceno, tolueno y para-xileno, respectivamente, dependiendo de la región espectral. Los valores de la t de Student obtenidos para la comparación entre los espectros corregidos y los mismos sin interferencia de acetona, no difieren significativamente para el nivel de confianza del 95%, confirmándose la alta similitud espectral en todas las regiones estudiadas. Como conclusión, este estudio demostró que el método de compensación propuesto para evitar las interferencias de gas y vapor, no se limita a la compensación de aire CO₂ y H₂O y puede aplicarse para la corrección de cualquier otro gas presente en la trayectoria óptica del sistema.

Parte 3. Muestras arqueológicas

Capítulo 8. Predicción de elementos alcalinotérreos en restos óseos por espectroscopia de la región del infrarrojo cercano

Los objetivos principales de este estudio eran el desarrollo de una estrategia de bajo coste, rápida y limpia para entender la degradación post-mortem de los huesos enterrados causada por el entorno de los mismos y evaluar cómo afecta a las distintas partes y tipos de huesos del esqueleto, además de definir una estrategia para la selección de huesos mejor preservados y desarrollar estudios adicionales más detallados. Al mismo tiempo, se ha desarrollado un método para la determinación de diversos elementos en restos óseos mediante el uso de la espectroscopia FT-NIR mediante medidas directas por reflectancia junto con herramientas quimiométricas como PCA y modelos multivariantes PLS.

Los resultados indican que FT-NIR proporciona un método analítico rápido, directo, sostenible y de bajo coste para la predicción del contenido de calcio, magnesio y estroncio en muestras de hueso enterrados y sedimentos, que podría ser muy útil para entender los cambios post mortem de los huesos causados por el medio ambiente, al tiempo que permitirían seleccionar las muestras de hueso con

señales biogénicas bien conservadas que se pudieran emplear en estudios posteriores.

El PCA ha demostrado que el perfil de elementos alcalinotérreos de las muestras de hueso y sedimentos, tal como cabía esperar, es claramente diferente. Además, la superficie externa de los huesos esponjosos es más similar a los sedimentos que a la superficie de los huesos corticales. Esto confirma que la capa ósea externa de los huesos esponjosos está más alterada que la capa superficial de los huesos corticales y que la presencia de Ca, Mg y Sr en los mismos es similar a la de las muestras de suelo. Por lo tanto, se puede concluir que los factores ambientales tienen un mayor impacto en los tejidos esponjosos, que son más susceptibles a los procesos diagenéticos que en los huesos corticales. Esta conclusión se alcanzó tras la observación de los espectros NIR en los que se distingue la pérdida de la banda de absorción de la apatita (el compuesto óseo principal) y la presencia de bandas características de la arcilla, especialmente en el estudio de la superficie de huesos esponjosos.

Se ha demostrado que la espectroscopia FT-NIR puede emplearse para predecir los contenidos de Ca, Mg y Sr en restos óseos y muestras de suelos, encontrándose que los valores de Ca y Sr eran mayores en la parte interna de los huesos, frente a los contenidos de Mg que eran más altos en la superficie externa de los huesos y en los suelos, y más bajos en la parte interna de los huesos.

En resumen, el enfoque metodológico desarrollado combina la espectroscopía NIR y el análisis estadístico, permitiendo predecir la composición mineral ósea de los elementos alcalinotérreos, abriendo una nueva perspectiva para la identificación de muestras mejor conservadas en estudios bioarqueológicos e investigaciones de la ciencia forense.

Respecto a las limitaciones del estudio, debe observarse que se basó únicamente en la parte mineral de los restos óseos tratados térmicamente y pulverizados antes de medir, perdiendo así información de la parte orgánica de los huesos y suelos. Además, la falta de sensibilidad de la espectroscopia NIR afecta a la capacidad de estas mediciones para la determinación de elementos traza, pudiéndose determinar únicamente componentes mayoritarios del hueso, como Ca, Mg y Sr.

Capítulo 9. Investigaciones forenses de huesos quemados mediante el uso de la espectroscopia infrarroja en la región cercana

El objetivo del estudio fue el de seleccionar las muestras de hueso más adecuadas para llevar a cabo estudios forenses y, para ello, se utilizó la evaluación de la espectroscopia NIR mediante el uso de herramientas quimiométricas como PCA, HCA y PLS-DA.

Los huesos carbonizados son más propensos que los calcinados a la degradación producida por procesos de intercambio con el suelo, por lo que una correcta clasificación de los huesos permite asegurar los resultados de los análisis posteriores. En este estudio se han analizado tanto la parte interna como la parte externa de los huesos quemados, así como muestras de sedimentos, para una mejor comprensión de los procesos de modificación ósea post-mortem. Además, se evaluó la determinación simultánea de calcio, magnesio y estroncio en los huesos quemados por PLS-NIR con el fin de proporcionar herramientas de detección respetuosas con el medio ambiente para evaluar la presencia de estos elementos minerales en los huesos.

Se obtuvieron resultados interesantes, utilizando tanto PCA, HCA y PLS-DA, para clasificar los huesos quemados a partir de los espectros NIR. El PCA, considerando únicamente dos componentes principales. El procedimiento fue capaz de discriminar los restos óseos afectados por los procesos diagenéticos post-deposicionales de los relativamente libres, debido a la falta de intercambios con los componentes del suelo. El HCA permitió separar claramente huesos calcinados y carbonizados y podría ser útil para predecir, de una forma bastante clara, el tratamiento térmico de restos desconocidos. El PLS-DA proporcionó una herramienta precisa para discriminar entre huesos calcinados y carbonizados, permitiendo clasificar con precisión las muestras desconocidas de huesos quemados.

El uso de la espectroscopia NIR combinada con la quimiometría ha demostrado ser una alternativa rápida y sencilla para estimar la concentración de calcio, magnesio y estroncio también en huesos quemados, con el objetivo de comprender la necrosis post mortem de los restos quemados causada por los cambios mediante la degradación ambiental.

En resumen, se puede concluir que las metodologías propuestas, basadas en el uso de la espectroscopia NIR combinada con diferentes herramientas quimiométricas, proporcionan una forma rápida y sostenible para seleccionar las muestras más adecuadas para realizar estudios forenses y evaluar los elementos minerales mayoritarios en los restos óseos. Esto resulta muy importante debido a

que los huesos carbonizados son más propensos a los procesos post-deposicionales que los calcinados, pudiéndose producir errores durante las interpretaciones de los datos. Además, los modelos de regresión PLS construidos a partir de espectros NIR proporcionaron una herramienta adecuada para predecir los niveles Ca, Mg y Sr en los huesos quemados.

4. FINAL CONCLUSIONS

4. Final Conclusions

From studies carried out during the development of this Thesis for the application of MIR, NIR and Raman spectroscopy in three different areas, such as food samples, environmental samples and archaeological samples, the following General and Specific conclusions are drawn.

General conclusions

One of the great advantages of vibrational spectroscopy is its versatility due to these techniques allow the analysis different samples regardless of its physical state.

If adequate external calibration is not possible, sample analysis using MIR, NIR or Raman spectroscopy combined with the chemometric tools, requires a reference samples set with the same characteristics representative of the total data set. In addition, in quantitative analysis is necessary a reference method to know the properties and concentrations of the analytes, being usually methodologies that require sample pre-treatment steps, the use of reagents and solvents, generate residues and usually are costly and time consuming. As alternative to these reference methodologies, procedures by using vibrational spectroscopy with chemometric methods have been developed. These methodologies allow the extraction of the required information from the data processing and simultaneous determination of several analytes, with reduced or any preparation and manipulation of the samples.

NIR, MIR and Raman spectroscopy allow the inspection and analysis of the samples practically in real time, in the production line or through the final product, in a fast and simple way. In many cases these techniques provide a non-destructive analysis, and are considered ideal methodologies for quality control of raw materials and quality control of finished products, in the laboratory and during the manufacturing process. NIR spectroscopy is less useful for compound identification than MIR spectroscopy, but appropriate for quantitative analysis. Although MIR spectroscopy has been the most used vibrational technique in previous years for the identification of materials, it is gaining strength in the quantitative analysis. Usually NIR spectroscopy needs to work with more sample quantity than required in the MIR, but like the Raman, can be used without manipulation or contact the samples with the measurement material, using containers of glass and avoiding possible contamination of the samples.

No reagents or solvents are required, thus no hazardous waste is created, reducing the testing cost, being lesser in the case of routine testing.

In addition to the aforementioned conclusions, the specific conclusions are also included, according to the works included in this report:

Specific conclusions

Part 1. Food samples

Chapter 1. Direct determination of major components in food

The aim of this study was to evaluate NIR and MIR spectroscopy combined with chemometric tools for the direct determination of major components (total fat, proteins and carbohydrates) and energy value of processed foods covering from baby food, infant fast food and adult menus.

From the results obtained through this study it can be concluded that the PLS-NIR models built after Kennard-Stone selection of a correct calibration set and careful selection of spectral ranges, data preprocessing techniques, and latent variables, provide a fast and sustainable tool for the analysis of fat, proteins, carbohydrates and energy value in samples without using reagents or solvents.

Chapter 2. Determination of polar compounds in oils

In the study on the comparison of FT-NIR and FT-MIR spectroscopy for the determination of polar compounds in frying oils, the results indicated that both NIR and MIR techniques combined with the use of PLS multivariate calibration can be considered as a fast, direct and sustainable alternative method for the determination of TPM in frying oils. In addition, the PCA method facilitates the sample classification by distinguishing the high or low TPM content. In summary, these studies show that the MIR and especially the NIR technique offer an alternative for the quality evaluation of fried oil samples with acceptable prediction errors.

The detailed study of the TPM determination in fried oils by transmission measures in the NIR spectroscopy has allowed, from the reduction of the number of variables of the spectra, to compare the techniques of MLR and PLS. The selection of a number of reduced variables with only 4 wavenumbers chosen allowed carrying out the TPM determination with adequate precision and predictive capability.

The use of chemometric models, applied to fried oils transmission NIR spectra, allows the determination of TPM content with RRMSEP values of 6.5% and RPD of 9.2. Moreover, the MLR-NIR model built from NIR spectra with reduced variables, provided a good correlation between the 4 selected wave numbers and the TMP content measured by the reference method, with a R^2 Pred value of 0.96 and a RPD of 5.0. These results are adequate for analytical purposes with excellent prediction accuracy and obtained results can be used to design a low cost instrument manufactured with a filter system or LEDs that allows working with the four selected wavenumbers for a fast and cheap evaluation of frying oil quality.

Chapter 3. Lipid determination in salmon oil

In the study focused on finding alternatives to hexane as solvent for lipid extraction in salmon, studies on the relative solubility of fish components in different solvents were carried out using Hansen Solubility Parameter (HSP) and Conductor-like Screening Model for Real Solvent (COSMO-RS) methods. In addition experimental studies were carried out to compare the efficiency of alternative solvents to hexane.

Hansen and Cosmo-RS simulation studies of the solvation character of different solvents for a series of lipid classes present in salmon fish oil indicated that d-limonene and p-cymene were similar to hexane for solvating TAGs over other lipidic compounds. Also, Hansen solubility RED values predicted that CPME could be the best solvent to solubilize all compounds considered, being more specific than the rest of solvents evaluated for each lipid class. However, obtained COSMO-RS values indicate that CPME, EtAc, and MeTHF could solvate in the same extent TAGs, DAGs, FFAs, and ergosterol compounds.

On considering the experimental data found by HPTLC and GC-FID for the analysis of salmon fish oil extracts, HPTLC indicate that TAGs were the predominant class of the lipid extracted with 73–77% followed by DAGs (15–20%), FFAs (5–6%), and ergosterol (2–3%). However, it was found that p-cymene and limonene extracts contained increased levels of DAGs and reduced amounts of TAGs than other solvent extracts. This phenomenon could be explained by the lipid degradation due to the use of temperatures to work using p-cymene and limonene in spite of the promising values found by prediction of RED values of HSP and COSMO-RS, which indicate that p-cymene and limonene are theoretically the best solvents to dissolve TAGs. Moreover, the amounts of fatty acids extracted by using each solvent obtained by GC-FID indicated that all solvents provide similar fatty acids profile except limonene and especially p-cymene extracts, which had low amounts

of some unsaturated fatty acids probably because double bonds present in PUFAs are prone to oxidation due to high temperatures involved using of limonene and p-cymene.

The strategy employed in this study could be used as a guide to find alternative green solvents to hexane for lipid extraction from different samples.

In short, it can be concluded that there are some differences between theoretical predicted solvation values of lipid class and experimental studies on the use of alternative solvents for lipid extractions, and these differences could be attributable to computational simulations that do not consider some aspects as time and temperature of extraction or other parameters such as enthalpy of vaporization or melting point of considered compounds. These conditions can modify the lipidic extracted fraction.

In general, experimental results showed that EtAc was the most suitable solvent to replace hexane due to its properties to extract all classes of lipids such as TAGs, DAGs, FFAs, and ergosterol in salmon samples. Given that other parameters as boiling point, Log P, toxicity category, and energy required for the evaporation of the solvent have been considered, it can be concluded that EtAc could be a good alternative solvent for the extraction of salmon oil lipids. EtAc does not provides emissions of volatile organic compounds, is a biodegradable solvent, and has nontoxic effects.

Alternatively, the exposed strategy could be also useful to evaluate alternative solvents in other fields in order to maintain the extraction capability of classic solvents but improving the sustainability of the processes by avoiding side effects.

Considering the lipid extracts of salmon obtained in the experimental part of the previous study, it was considered the possibility of evaluating the NIR technique in combination with the PLS models for the simultaneous determination of i) fatty acids, such as oleic, palmitic, linoleic and linolenic, ii) families of fatty acids, such as omega-3, omega-6, polyunsaturated fatty acids, monounsaturated fatty acids and saturated fatty acids, and iii) lipid classes such as triacylglycerides, diacylglycerides, free fatty acids and ergosterol in salmon oil.

The PLS-NIR method developed based on NIR transmission measures, non-destructive, allows the simultaneous quantification of several fatty acids, as OA, PA, ALA and LA as well omega-3, omega-6, PUFAs, MUFAs and SFAs, in addition to lipid classes as TAGs, DAGs, FFAs and ergosterol in salmon oil. This strategy avoids the use of toxic or dangerous reagents together with a drastic reduction in time and energy consumption. However, the accuracy and robustness of the method remains at the screening level for palmitic and linolenic acids and MUFAs. So, it

can be concluded that PLS-NIR offers a high value screening tool for characterization of salmon fish oil samples with some good predictive capability for ergosterol, FFAs, DAGs, TAGs, omega-3 and PUFAs.

Chapter 4. Quality control of umami additives by Raman and infrared spectroscopy in the near region

The aim of this study was the development of analytical methods for the determination of MSG, GMP and IMP in Umami industrial food additives based on direct measurements by vibrational spectroscopy.

Raman and NIR spectroscopy techniques provide fast and green alternatives to determine simultaneously MSG, GMP and IMP in umami additives. For binary mixtures composed by GMP and IMP, both techniques combined with PLS offer accurate procedures for the quantitative analysis of these samples. However, for ternary mixtures composed by MSG, GMP and IMP, results obtained by using NIR spectroscopy were better than those found by Raman. In short, these studies evidenced that for umami additive analysis, PLS-NIR offers an accurate and precise tool for the determination of MSG, GMP and IMP analysis, being a clear green alternative to traditional methods, based on the use of chromatographic techniques, which present great time consumption and a considerable cost, besides requiring the previous dissolution of the sample.

NIR spectroscopy allows obtaining real-time results without any sample pre-treatment, being a good alternative method for rapid inspection of food additive formulations to ensure the quality of umami mixtures prior to their use in food processing.

Part 2. Environmental and agricultural interest samples

Chapter 5. Evaluation of compost for its posterior use in agriculture using infrared spectroscopy NIR and MIR

Multivariate calibration methods developed using MIR and NIR spectroscopy offer a sustainable alternative for the determination of total phenolic compounds in compost samples. However, on comparing the main analytical features of the two developed methods it can be conclude that NIR provides an improved prediction power as compared with MIR; based on the RPD values and RMSEP for the validation set of samples was 0.19 mg of gallic acid equivalent per gram of dry mass (mg GAE g⁻¹ DW) for NIR and 0.27 mg GAE g⁻¹ DW for MIR. In any cases, the two proposed methodologies can be employed for the determination of TPC,

being the NIR measurements faster and earlier than MIR ones due to the samples are inside of glass vials while MIR needs a clean protocol to avoid cross-contamination in ATR measurements. However, it is clear that models with NIR spectra provided a better modeling (lower values of calibration, cross-validation and prediction) than those obtained from ATR-MIR spectra.

Concerning the prediction of organic carbon and total nitrogen content in compost and organic waste samples by infrared spectroscopy, it can be concluded that PLS-MIR models provide better results than PLS-NIR, without which in neither case can be considered excellent. In this sense, it should be noted that NIR and MIR offer sustainable green alternatives for the determination of OC and TN, suitable for screening purposes. NIR procedure is free from the use of reagents and wastes, whereas the MIR ones require small amounts of cell cleaning solvents and provides reduced volumes of sample wastes.

Chapter 6. Determination of benzene, toluene, ethylbenzene and xylenes in soils by infrared spectroscopy in the middle region in vapor phase

The aim of this work was to develop a simple methodology that combines the use of HS sampling with IR detection for a fast and direct determination of BTEX in soil samples, avoiding the use of solvents and reagents, and without chromatographic separation. For this purpose, the chemometric treatment of vapor phase spectra was used for benzene, toluene, ethylbenzene, ortho-, meta- and para-xylene determination by eliminating the spectral interferences from water and CO₂ by using a methodology developed in our laboratory.

The results indicated that infrared spectroscopy combined with multivariate calibration techniques provides a suitable tool for simultaneous determination of BTEX compounds in contaminated soils in a fast way without any sample chemical pre-treatment, although it is true that their sensitivity is not comparable to that of chromatographic methods.

The use of headspace for the generation the analyte vapor from samples reduces the experimental errors associated with the preparation step and reduces the analysis time to 7 minutes, including the heating and measurement time of each sample. Additionally, the analysis time can be reduced when several samples are heated together. Therefore, it has been shown that the HS-FT-IR offers a suitable green technique for the accurate determination of BTEX by PLS calibration with coefficients of variation between 0.9-11.2, 0.4-4.7, 1.2-7.2 0.9-9.3, 2.0- 7.0 and 0.4-8.0% for benzene, toluene, ethylbenzene, ortho-, meta- and para-xylene, respectively. For a real sample of contaminated soil, taken from a fuel spill area,

PLS models in the 3258-3001 and 913-567 cm^{-1} region and first derivative and mean centered pre-treatment were the best option to determine benzene, toluene, ethylbenzene, ortho- and para-xylene, whereas the determination of meta-xylene requires the use of the second derivative and the mean center as data pre-treatment in the region of 913-567 cm^{-1} .

Chapter 7. Vapor Corrections in Infrared Spectroscopy

The presence of spectral interferences is a problem to obtain accurate results when using infrared spectroscopy. For this reason, the aim of this study was to evaluate a flexible and automated method for avoiding the spectral interferences of organic vapors in FTIR spectra measured in vapor phase, taking acetone as a model compound. For this purpose, a green methodology was used, combining the use of head space sampling (HS) with FTIR spectroscopy transmission for the samples analysis containing benzene, toluene and para-xylene and ethanol.

The results obtained evidenced that the proposed method for the compensation of the presence of acetone, applied to the spectra of samples of benzene, toluene, para-xylene and ethanol in vapor phase provides a good correlation between spectra in the absence of interference and corrected spectra. The values of R are close to 1, the values of Z after correction being between 3.8 and 5.4 and between 2.71 and 4.7 for ethanol and benzene, toluene and para-xylene respectively, also depending on the spectral region. The Student t values obtained for the comparison between corrected spectra and the same ones without acetone interference do not differ significantly for the 95% confidence level, confirming the high spectral similarity in all the studied regions. In conclusion, this study demonstrated that the proposed compensation method to avoid gas and vapor interference is not limited to CO_2 and H_2O air compensation and can be applied for the correction of any other gas present in the optical path of the system.

Part 3. Archaeological samples

Chapter 8. Prediction of alkaline earth elements in bone remains by spectroscopy of the near infrared region

The main objectives of this study were to develop a low-cost, fast and clean strategy to understand post-mortem degradation in buried bones caused by the environment and to evaluate how it affects the different skeletal parts and types of bones, in addition to defining a strategy to select bones better preserved and to develop further detailed studies. At the same time, a method has been developed

for the determination of various elements in skeletal remains by the use of FT-NIR spectroscopy by using direct measurements by reflectance along with chemometric tools such as PCA and PLS multivariate models.

The results indicate that FT-NIR provides a fast, direct, sustainable and low cost analytical method for the prediction of calcium, magnesium and strontium content in buried bone and sediments samples, and it could be very useful to understand post mortem changes of bones caused by the environment, at the same time allowing bone samples selection with well-preserved biogenetic signals that could be used in later studies.

PCA has demonstrated that the alkaline earth elements profile of bone and sediment samples, as expected, is clearly different. In addition, the external surface of spongy bones is more similar to sediments than to the surface of cortical bones. This confirms that the outer bone layer of spongy bones is more altered than the cortical surface and that the presence of Ca, Mg and Sr in the same is similar to that of the soil samples. Therefore, it can be concluded that environmental factors have a greater impact on spongy tissues, which are more susceptible to diagenetic processes than to cortical bones. This conclusion was reached after observation of the NIR spectra in which the loss of the absorption band of apatite (the main bone compound) and the presence of clay bands, especially in the study of the surface of spongy bones.

It has been shown that FT-NIR spectroscopy can be used to predict the Ca, Mg and Sr contents in skeletal remains and soil samples. Ca and Sr values were found to be higher in the internal bones of bone samples compared to Mg contents that were higher in the external bone surface and soils, and lower in the internal part of bones.

In summary, the methodological approach developed combines NIR spectroscopy and statistical analysis, allowing to predict the mineral bone composition of the alkaline earth elements, opening a new perspective for the identification of better preserved samples in bioarchaeological studies and forensic science investigations.

Regarding the limitations of the study, it should be noticed that it was based only on the mineral part of the bone remnants which were treated thermally and pulverized before measurements, thus losing information from the organic part of the bones and soils. Additionally, the lack of sensitivity of NIR spectroscopy affects the capability of these measurements to determine trace elements, being able to determine only major components of the bone, such as Ca, Mg and Mr.

Chapter 9. Burned bones forensic investigations by the use of infrared spectroscopy in the near region

The aim of the study was to select the most suitable bone samples to carry out forensic studies and, to do it, the evaluation of NIR spectroscopy was employed being used chemometric tools as PCA, HCA and PLS-DA.

Carbonized bones are more prone than calcined ones to degradation produced by exchange processes with the soil, so a correct classification of the bones allows ensuring the subsequent analysis results. In this study internal part, external part of burned bones and sediment samples were analysed for a better understanding of post-mortem bone modification processes. Furthermore, it has been evaluated the simultaneous determination of calcium, magnesium and strontium in burned bones by PLS-NIR in order to provide environmentally friendly screening tools to evaluate the presence of mineral elements in bones.

Interesting results were obtained using PCA, HCA and PLS-DA to classify burned bones. PCA, using two principal components, was able to discriminate bone remains affected by post depositional diagenetic processes to those relatively free due to their lack of exchanges with the soil components. HCA permitted to clearly separate calcined and carbonized bones and could be useful to roughly understand the thermal treatment of unknown remains. PLS-DA provided an accurate tool to discriminate between calcined and carbonized bones and to classify accurately unknown burned bones.

The use of NIR spectroscopy combined with chemometrics has been shown to be a fast and simple alternative for estimating the calcium, magnesium and strontium concentration in burned bones in order to understand post mortem necrosis of burned remains caused by the environmental degradation changes.

In short, it can be concluded that the proposed methodologies, based on the use of NIR spectroscopy combined with different chemometric tools, provide a fast and sustainable way to select the most suitable samples for forensic studies and to evaluate the majority mineral elements in the bone remains. It is very important because the carbonized bones are more prone than the calcined ones to the post-depositional processes, and those can produce errors during data interpretations. In addition, PLS regression models built from NIR spectra provided a screening tool to predict Ca, Mg and Sr levels in burned bones.

5. ANEXO: Artículos científicos publicados

Artículo 1. Direct determination of major components in human diets and baby foods

Analytical and Bioanalytical Chemistry 407(2015) 1961-1972

Direct determination of major components in human diets and baby foods

Mari Merce Cascant · Salvador Garrigues · Miguel de la Guardia

Received: 8 October 2014 / Revised: 23 December 2014 / Accepted: 5 January 2015 / Published online: 18 January 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract A fast method has been developed for the determination of fat, proteins, carbohydrates, and energy value in baby food, infant fast food, and canteen menus, using near-infrared (NIR) and mid-infrared (MIR) spectroscopy measurements and multivariate calibration methods based on partial least square regression. Reference standard methods were employed to build and validate the infrared methods for direct determination of nutrients. Coefficients of determination obtained between predicted values and reference ones for total fat, proteins, carbohydrates, and energy value were 96.7, 98.1, 98.9, and 96.5 for NIR and 91.0, 93.0, 92.0, and 84.1 for MIR, respectively, with relative root mean square error of prediction (RRMSEP) below or equal to 9 % for NIR and 16 % for MIR. Results obtained indicate that both NIR and MIR techniques have good predictive capabilities, with the NIR method being the most accurate and simple.

Keywords Near-infrared spectroscopy · Mid-infrared spectroscopy · Total fat · Proteins · Carbohydrates · Energy value · Human diet · Baby food · Infant fast food · Canteen menus · Partial least square (PLS)

Introduction

Human nutrition strongly depends on the energy value of foods and from an equilibrated proportion of proteins, carbohydrates, and fats, being also dependent on the presence in foods of micronutrients like minerals and vitamins. Adequate

nutrition during the first and second year of life is critical to ensure optimal physical and mental development of infants and young children. Young children, like everyone else, need many different nutrients for growth and development, to provide energy, and to keep healthy. Fats and carbohydrates are characterized for their high energetic value, while proteins are used as construction materials and major parts of body compounds. On the other hand, poor nutrition increases the risk of illness in adulthood. Early nutritional deficits are also linked to long-term impairment in growth and health [1].

The energy imbalance between calories consumed and expended can produce overweight and obesity, hypertension, and cardiovascular diseases, due to high ingestion of hypercaloric foods and/or diets with high values of fat, salt, and sugars and low values in vitamins, minerals, and other micronutrients. Therefore, the recommended human diet should be varied so that it can provide all the nutrients required by the body needed for proper functioning, and the complementary foods must be introduced from young years [1]. Nutrition labeling informs consumers of food nutritional qualities, and in the case of baby food labels, the nutrient contents should appear according to the Codex Alimentarius [2]. In processed food, the composition can be altered and consumers must be informed on the labels of foods at the time of product acquisition. Therefore, qualitative and quantitative composition of food is essential to ensure that the product conforms to the expectations of customers and nutritionists. However, problems related to the balance of major components in foods increase on increasing the age of consumers and the diversity of foods that can be consumed. So, for kids, children's menus offered by fast food companies and, for adults, the canteen menus can be taken as examples of human diets which concern one of the main meals consumed during a day.

The Official Methods of Analysis of AOAC International uses gravimetric and volumetric analysis to characterize and quantify the components of foodstuffs [3], but these

Parts of this work were presented at JAI 2014 held in Barcelona, October 2014.

M. M. Cascant · S. Garrigues (✉) · M. de la Guardia
Department of Analytical Chemistry, University of Valencia, 50 Dr.
Moliner Street, Research Building, 46100 Burjassot, Valencia, Spain
e-mail: salvador.garrigues@uv.es

methodologies are time and reagent consuming, and because of these, spectroscopy techniques offer interesting alternatives to traditional chemical analysis for the quantitative analysis of major components in foods [4]. Near-infrared (NIR) spectroscopy provides information about untreated samples from overtones and their combinations, and it has been applied to food quality evaluation [5] in breadstuff [6, 7], dairy products [8], meat [9] and fish products [10], vegetables products [11], and processed food [12–14]. Meanwhile, mid-infrared (MIR) spectroscopy provides information from frequencies of fundamental molecular vibrations, and it has been applied for the determination of nutritional parameters in breadstuff [6], beverages [15], vegetables [11], fish [16], dairy products [17, 18], and processed food [6, 14]. However, only a few works based on the determination of macronutrients by NIR spectroscopy in homogenized meals were found [19–21], while in our knowledge, MIR has not been employed for the determination of major components in total diets and mixed foods.

The aim of the present study has been the evaluation of MIR and NIR spectroscopy for the direct determination of total fat, proteins, total carbohydrates, and energy value of mixed foods covering from baby food, infant fast food, and adult menus, by using NIR and MIR spectrometry, in association with chemometric tools, in order to provide screening tools for the evaluation of the human diet.

Experimental

Instruments

A Fourier transform near-infrared (FT-NIR) spectrometer, model Multipurpose Analyzer (MPA) from Bruker (Bremen, Germany), equipped with an integrating sphere, was employed for diffuse reflectance near-infrared spectra acquisition. This instrument is equipped with a NIR source, a quartz beam splitter, and PbS detector.

A Fourier transform middle infrared, model Tensor 27 (Bruker, Bremen, Germany), equipped with a DLATGS detector was employed for IR spectral measurements. Samples were measured in attenuated total reflectance (ATR) mode using a dry air-purged in-compartment DuraSampleIR accessory for solids from Smiths Detection Inc. (Warrington, UK) equipped with a three-reflection diamond/ZnSe DuraDisk plate.

Samples

Fifty-eight homogenized meals, containing two or more food groups, were analyzed. Ingredients included in these meals are meat, fish, fruits and vegetables, cereal products, shellfish, eggs, dairy products, sauces, and oils in varied combinations. These meals concern three groups: (i) commercial jar baby

food (35 samples), (ii) infant fast food (10 samples), and (iii) complete canteen menus (13 samples).

Prepared baby food samples were obtained from national supermarket chains in the city of Valencia (Spain). Some of them contain a mixture of cooking water, vegetables, and other ingredients, mainly meat or fish (29 samples), and others contain only mixed fruits (six samples). Ten infant fast foods were purchased from different burgers of Valencia City and are composed of meat burger with cheese and sauce, chips, a soft drink, and a dairy dessert. Thirteen different menus were purchased from canteens of the campus of the University of Valencia. Menus were composed of a starter, a main dish, and a dessert and also included a 50–70-g piece of bread. Samples were chosen to include all the foods offered to the students in the canteen. Different starters and main courses were selected randomly between those offered each day in order to have a representative description of the diet followed in the area of study.

All samples were handled and stored in a refrigerator at 4 °C until their analysis for a period of less than 24 h. The edible part of the foods was weighed and mixed, crushed and homogenized with a domestic mixer (Braun, Kronberg, Germany), and finally frozen at –20 °C in a freezer. Afterward, they were lyophilized for a minimum of 48 h at a chamber pressure of 0.05 mbar by using a Cryodos 50 lyophilizer from Telstar (Barcelona, Spain). The dried samples were pulverized with a domestic mixer, and the powdered samples stored in polyethylene bottles until their analysis.

Software and algorithms

For instrumental and measurement control of both infrared spectrometers as well as for data acquisition, Opus 6.5 software from Bruker was used. Data treatment was carried out using in-house written functions employing Matlab 7.14.0.739 (R 2012a) from MathWorks (Natick, MA, USA), and for partial least square (PLS) model calculation and validation, interval PLS (iPLS), and hierarchical cluster analysis (HCA), the PLS Toolbox 6.7.1 from Eigenvector Research Inc. (Wenatchee, WA, USA) was used. Detailed information on PLS regression and other related parameters employed in this work can be obtained from the literature [21].

Reference methods for macronutrient determination

Official methods of analysis were used to obtain reference data for samples under study in the analysis of general parameters of all samples considered in this study. All determinations were performed in duplicate. Total fat content, which includes free and bound fat, was determined by using a Fat Analyzer Soxtec System HT2 1045 Extraction Unit from Tecator (Burladingen, Germany) equipped with a silicone bath Soxtec System HT 1046 Service Unit. In this study, total fat was determined following acid hydrolysis prior to the

extraction with petroleum ether and gravimetric measurement of the extracted lipid fraction [3]. To determine the nitrogen content of the samples, the Kjeldahl method was used [3], using an automatic system composed of a Büchi-426 Digestion Unit and an Analyzer Unit 2300 Kjeltec from Foss (Hillerød, Denmark). Protein content was obtained from nitrogen data multiplied by a factor of 6.25 [22]. Total carbohydrates were calculated as the difference between 100 and the sum of the percentages of total fat, protein, and ash [22]. Ash was determined by dry ashing at 550 °C [3] using an oven with 46 L of capacity and maximum temperature of 1100 °C, from Hornos Gallur (Valencia, Spain). The energy content of the samples was calculated as suggested by the United States Department of Agriculture (USDA) [23] by the Nutritional Labelling and Education Act (NLEA) [24] for multi-ingredient processed foods, by the sum of the energetic value of total fat, protein, and total carbohydrates, using factors of 37, 17, and 17 kJ per 100 g, respectively [25]. Nutrient contents of samples determined by the reference method were in the range of 0.05–18.6, 1.9–26.2, and 55.5–95.0 g per 100 g and 1650–2010 kJ per 100 g for total fat, proteins, carbohydrates, and energy values, with standard deviations below 0.5, 1.6, 2.2, and 5.2 %, respectively.

NIR procedure

Lyophilized samples were placed in clear glass vials of 22 mm internal diameter and 44 mm height to directly obtain NIR spectra by diffuse reflectance. Diffuse reflectance near-infrared spectra were obtained in Kubelka-Munk units. All FT-NIR spectra were acquired between 14,000 and 4000 cm^{-1} by co-adding 50 scans and using an optical resolution of 4 cm^{-1} . The background spectrum was acquired from the closed integrating sphere using the same instrumental conditions than those employed for samples. Three measurements of each sample were obtained by rotating the sample vial position between replicates in order to ensure a better representative. Means of the triplicate spectra of each sample were calculated and employed to build PLS calibration models and for sample predictive purposes.

ATR-FTIR procedure

Direct ATR measurements were performed by depositing approximately 100 mg of sample directly onto the diamond ATR crystal and pressed with the corresponding accessory. Spectra were recorded in the range between 4000 and 600 cm^{-1} , by accumulating 50 scans per spectrum, at a resolution of 4 cm^{-1} . A spectrum of the clean ATR crystal, obtained immediately before acquiring the spectrum of each sample, was used as a background.

Spectra from three portions of each sample were normalized by minimum-maximum after eliminating the region

between 2300 and 1900 cm^{-1} due the high absorption of the ATR cell (diamond). Mean spectra of each sample were obtained from the individual normalized ones and used to build the PLS models and for sample prediction.

Data treatment

To build PLS models, samples were divided into two sets: one used for calibration and a second one employed as a validation set. The Kennard-Stone (K-S) algorithm was used for the calibration set selection [26]. The K-S algorithm starts with the two most distant objects as members of the calibration set. The algorithm takes the objects into the training set based on their Euclidean distance from the already selected objects. Therefore, the K-S algorithm for each NIR and MIR measurement was employed to establish the optimal samples to be selected for each PLS model. Calibration models for prediction of percentage of total fat, protein, and total carbohydrates, as well as energy value, in kilojoules, from NIR and MIR spectra were developed using the training set selected by K-S. The calibration set must be representative of the total samples; thus, a cluster analysis was also made to identify groups of observations that share common characteristics and was used to check that the selected calibration sets contain samples of the different sample groups. The outcome of hierarchical cluster analysis was represented graphically as a dendrogram that was established using Ward linkage method [27].

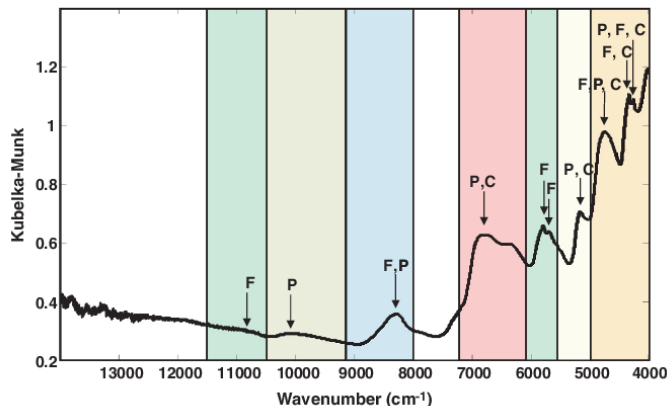
The selection of regions is a crucial step to build PLS models. In this work, two strategies were used for variable selection: one based on a specific region for each component and their characteristic absorption bands and the second one based on iPLS [28] that selects a subset of variables which will give superior prediction compared to using all the variables in the data set. In all the cases, leave-one-out cross validation (CV) was employed for the selection of the optimum number of latent variables (LVs) of the models according to the root mean square error of cross validation (RMSECV) values. Different preprocessing data treatments, such as mean centering, standard normal variate (SNV), multiplicative scattering correction (MSC), and first or second derivative, were studied for both NIR and MIR measurements. The proposed PLS models were built using the statistically inspired modification of the PLS method (SIMPLS) algorithm with the same preprocessing used in iPLS.

Results

FT-NIR spectra

Figure 1 shows, in Kubelka-Munk units, the mean NIR spectra in the region between 14,000 and 4000 cm^{-1} , corresponding to all food samples considered in this study without any

Fig. 1 NIR average spectrum of food samples considered in this study obtained as mean of all diffuse reflectance of untreated spectra in Kubelka-Munk units. *Note:* Regions indicated correspond to fat (F), protein (P), and carbohydrate (C) bands



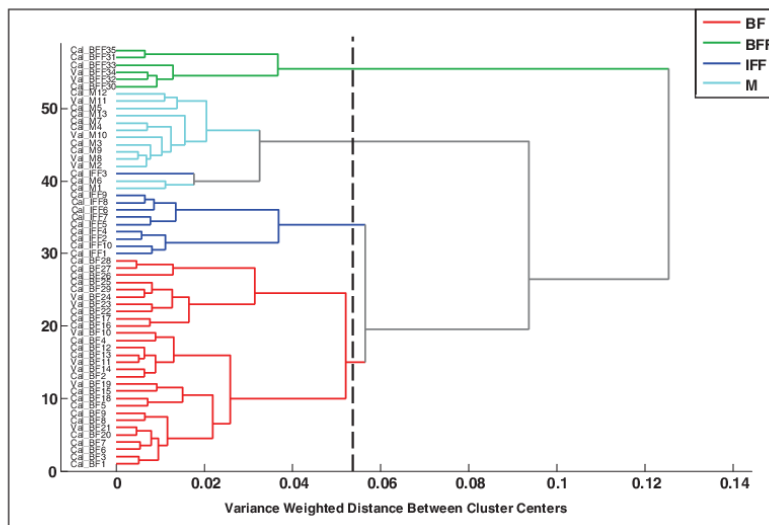
data pretreatment. Characteristic NIR spectra of these foods present absorbance bands corresponding to functional groups associated with lipids, proteins, and carbohydrates. The main bands correspond to overtone of C-H stretching of fats [14] and proteins, together with combination bands of C-H, N-H, and O-H stretching in proteins and carbohydrates [29, 30]. These NIR spectra agree with those expected from the description available in the literature for this type of samples.

PLS-NIR models

The complete data set was divided into a calibration and a validation subset with 45 and 13 objects, respectively, using the K-S algorithm for the calibration set selection. In order that

the calibration set could be a representative of all the samples considered to obtain models with a good predictive capability, it was composed of 26 baby food, 4 of these based on fruits, 10 adult menus, and 9 infant fast foods. Moreover, a dendrogram of the sample spectra obtained after the first derivative and mean center pretreatment and selecting the region from 11,000 to 4000 cm^{-1} was built to evaluate the similarity of the samples. The dendrogram of the samples established using Ward linkage method is shown in Fig. 2, and it can be seen that for a minimum cutoff value of 0.055, all samples were classified into four main groups of spectra, which correspond to samples with common characteristics, like canteen, infant fast food, and baby foods, considering in the last type the differences between fruit-based foods and general foods.

Fig. 2 Dendrographic classification of NIR spectra of samples concerning baby food (BF), baby food based in fruits (BFF), adult menus (M), and infant fast food (IFF) established using Ward linkage method. *Note:* Cal and Val denote calibration and validation samples chosen using Kennard-Stone algorithm. *Dashed line:* minimum cutoff value of 0.055



The first group is composed of baby food based on fruits and the second one is based on adult canteen menus; however, it also included one sample of infant fast food. The third group is based on infant fast food, and the last one is composed of baby food based on meat, fish, and vegetables. In this figure, the samples selected for calibration and validation were observed, based on the use of the K-S algorithm. From that, it can be concluded that the calibration set is a representative of all the different samples considered in this study. The main characteristics of the calibration and validation sets selected for all nutrient determination by PLS-NIR are indicated in Table 1. iPLS was employed for the selection of spectral regions, and to do it, the spectrum was divided into intervals of 207 variables using 10 LVs as maximum in each one. Figure 3 shows the NIR spectra of human diets and baby food with regions selected for iPLS considering each type of main nutrients. The optimal combination of intervals was established from the lowest RMSECV. For fat determination, regions selected were those between 10,803 and 10,406, 6013 and 5217, and 4416 and 4019 cm^{-1} , corresponding to the aforementioned fat regions. In the case of proteins, regions selected were those within 8807–8011, 6013–5616, and 5215–4817 cm^{-1} , corresponding to the first overtone and C-H stretching combination. Regions selected for carbohydrate determination were between 7210 and 8011, 6013 and 5616, and 4816 and 4019 cm^{-1} , corresponding to N-H and O-H stretching combination and C-H stretching combination. To determine the

energy value of mixed food samples, included regions selected were between 10,804 and 10,406, 5215 and 4817, and 4416 and 4019 cm^{-1} , corresponding to some of the regions selected for total fat, proteins, and carbohydrates, because this parameter was calculated based on these three parameters. To build PLS-NIR models using the selected spectral range, 6, 7, 4, and 7 LVs were employed for total fat, protein, total carbohydrate, and energy value, respectively, in order to minimize the RMSECV, explaining 93.0, 96.5, 93.0, and 93.2 % of the total variance of the X data block and 99.1, 99.0, 98.3, and 99.3 % of the Y data block, respectively. For total fat, proteins, and energy value, a Savitzky-Golay first derivative (FD), with a window of 9 points and a second order polynomial, and mean centering treatment were adopted as signal preprocessing. For carbohydrates, SNV, first derivative, and mean center were chosen. The first derivative effectively removes any offset from the sample and de-emphasizes low-frequency signals. Additionally, concentration data were mean centered as well in all cases. The most important calibration parameters of the developed PLS-NIR models are summarized in Table 2, also indicating the number of latent variables used for the model built for each analyte, root mean square error of calibration (RMSEC), the RMSECV, and coefficients of determination (R^2) of calibration and cross validation of sample sets.

PLS-NIR prediction capability

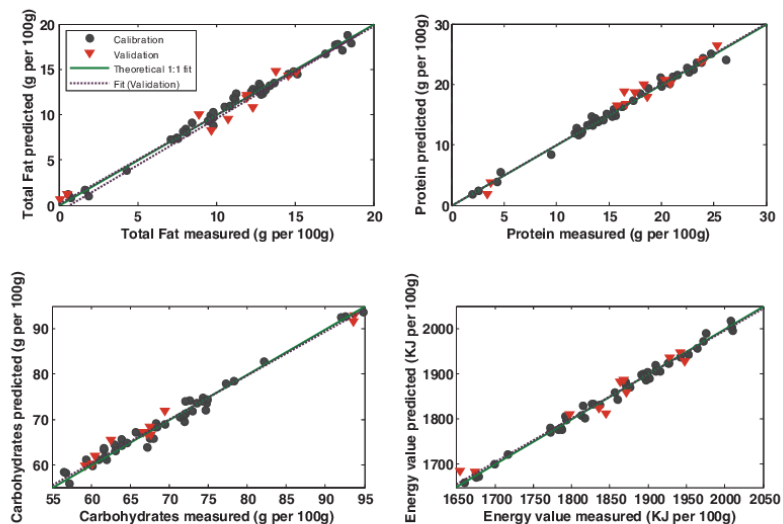
To confirm the suitability of the chosen preprocessing for model building, an external validation set was used. Outliers can be identified by Q residual versus the Hotelling T^2 values. Two samples were removed, corresponding to a canteen menu and a baby food, from the validation set of total fat, carbohydrates, and energy value prior to the calculations and validation of the final PLS models of those analytes. Predicted values of the analytes versus those obtained by the reference methods are depicted in Fig. 4, and it can be seen that calibration and validation sample points were closely distributed near the optimum regression line between predicted and measured values for total fat, protein, total carbohydrate, and energy. The most important calibration and prediction parameters of the developed PLS-NIR models are summarized in Table 2. Acceptable relative prediction errors (relative root mean square error of prediction (RRMSEP)) of models were obtained with average values of 9.2, 6.1, 2.1, and 1.0 % for total fat, proteins, carbohydrates, and energy, with coefficients of determination of 96.7, 98.1, 98.9, and 96.5, respectively. Low cross validation and prediction bias values as well as high coefficients of determination supported the quality of model performance. Table 1 summarizes the main characteristics of the validation set for all nutrients predicted by the PLS-NIR model. In terms of accuracy, the relative error was calculated between the reference value and the predicted ones for the set of validation samples. The accuracy of the method was very

Table 1 Statistics for calibration and validation sets employed for PLS-NIR determination of food nutritional parameters

Analyte (units)	Set	Samples	Range	Mean value	SD
Total fat (g per 100 g)	Cal	45	0.05–18.6	10	5
Protein (g per 100 g)		45	1.95–26.2	16	6
Total carbohydrates (g per 100 g)		45	55.5–94.9	70	10
Energy values (kJ per 100 g)		45	391–478	440	20
Total fat (g per 100 g)	Val	11	0.07–15.1	10	5
Protein (g per 100 g)		13	3.42–25.3	17	7
Total carbohydrates (g per 100 g)		11	59.2–93.7	70	10
Energy values (kJ per 100 g)		11	1660–2010	1900	90
Total fat (g per 100 g)	Pred	11	0.08–14.8	10	5
Protein (g per 100 g)		13	3.3–26.4	18	7
Total carbohydrates (g per 100 g)		11	59.7–93.0	70	10
Energy values (kJ per 100 g)		11	1650–1950	1800	100

For calibration (Cal) and validation (Val) datasets, reference values are reported. For this last dataset, predicted values (Pred) are also reported. Range: concerns minimum-maximum value
SD standard deviation

Fig. 4 Predicted versus reference values for PLS-NIR determination of total fat, proteins, carbohydrates, and energy value using regions selected by iPLS



and C-O-C and C-O-H distortion from carbohydrates and to C=C-H bending of fat. The interval between 900 and 750 cm^{-1} corresponds to the fingerprint region. At 720 cm^{-1} , there appears a CH_2 twist vibrational mode from fat [31, 32].

PLS-ATR-FTIR models

The ATR-FTIR-obtained data set was divided into a calibration and a validation subset with 45 and 13 objects, respectively, using the K-S algorithm. The calibration set was used to build the models, and the other one was employed to evaluate the predictive capability of the model and to test the robustness of the models for an independent number of samples. Figure 6 shows the similarity of samples established by

HCA from a dendrogram built using Ward linkage method after MSC, first derivative, and mean center as pretreatment of spectra, using the region between 3600 and 600 cm^{-1} . As can be seen from the figure, the four different types of samples considered could be identified for a minimum cutoff value of 0.65. One group is composed of baby food based on fruits, the other based on infant fast food, and the additional two groups based on baby food and adult menus. So, it can be seen that the samples selected by the K-S algorithm for calibration are representatives of those selected for validation. The calibration set consisted of 26 baby food samples, 4 of them based on fruits, 9 adult menus, and 10 infant fast foods. The main characteristics of the calibration and validation sets employed to build the PLS-ATR-FTIR models for all nutrients considered are shown in Table 3. From the table, the number of samples

Fig. 5 Mean ATR spectrum corresponding to all samples of mixed foods considered in the study. Note: Regions indicated correspond to fat (F), protein (P), and carbohydrate (C) bands. For details, see text

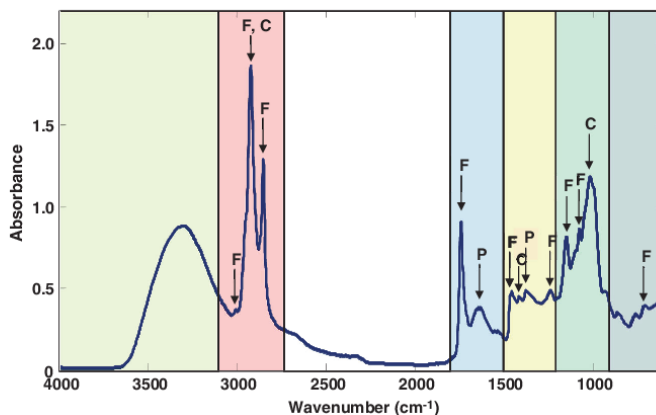
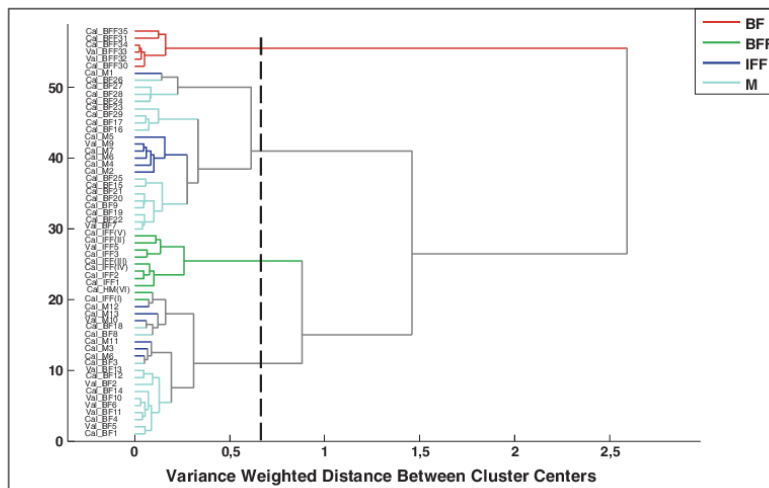


Fig. 6 Dendrographic classification of ATR spectra of samples concerning baby food (BF), baby food based in fruits (BFF), adult menus (M), and infant fast food (IFF) established using Ward linkage method. *Note:* Cal and Val denote calibration and validation samples chosen using the Kennard-Stone algorithm. *Dashed line:* minimum cutoff value of 0.65



included in each set and the range and standard deviation of the reference values for each considered parameter can also be seen. All spectra were preprocessed by minimum-maximum normalization.

iPLS was used for the selection of spectral regions. MIR spectra between 3600 and 600 cm^{-1} were used, without the region between 2300 and 1900 cm^{-1} corresponding to the

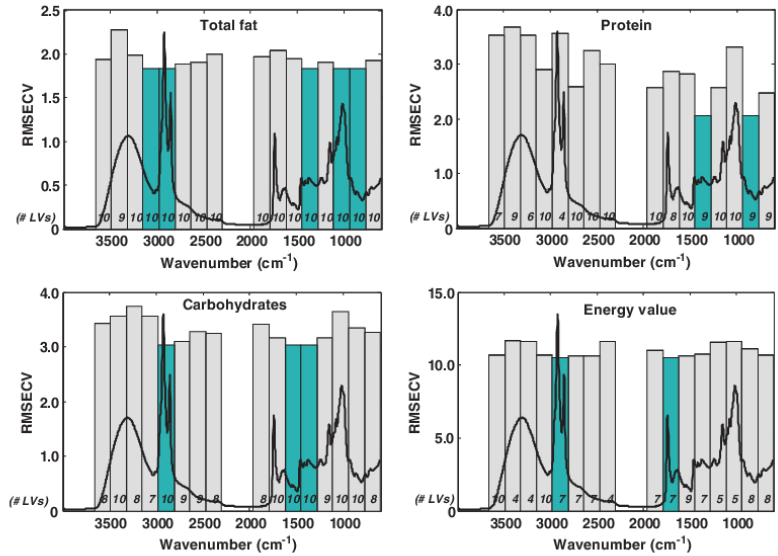
absorption of the ATR cell, and were divided into equidistant subintervals of 88 variables using 10 LVs as maximum for model building in each case. The optimal combination of intervals was obtained from the lowest RMSECV. Figure 7 displays the results obtained from iPLS for variable selection for the models of total fat, protein, carbohydrates, and energy values with the spectral regions associated to these constituents. The regions selected for total fat were the ones between 3148 and 2810 cm^{-1} due to symmetrical and asymmetrical C-H stretching of CH_2 and CH_3 , the region between 1450 and 1282 cm^{-1} that corresponds to mixed region, and the region between 1111 and 773 cm^{-1} where C=C-H bending is found. For proteins, the regions selected were the ones between 1450 and 1282 cm^{-1} that correspond to the mixed region and the region between 941 and 773 cm^{-1} that corresponds to the fingerprint region. In the case of carbohydrates, the regions selected were the ones between 2978 and 2810 cm^{-1} due to C-H stretching of CH_2 and CH_3 and the mixed region found between 1620 and 1282 cm^{-1} . In the case of energy value, the regions selected were the ones between 2978 and 2810 cm^{-1} that correspond to C-H groups for asymmetrical and symmetrical stretching of CH_2 and CH_3 of fat and carbohydrates and 1789 and 1622 cm^{-1} where the band due to C=O stretching of fat appears at 1725–1745 cm^{-1} . Regions selected in iPLS to build the PLS-ATR-FTIR models were used with 7, 7, 5, and 7 LVs for the sets of percentage of total fat, protein, total carbohydrate, and energy value (in kJ), respectively, according to RMSECV results. For all cases, first derivative and mean center were chosen and concentration data were mean centered as well in all cases. It has to be remarked that two and one sample had to be removed from the calibration set of fat and proteins, respectively, and three samples from carbohydrates and energy value, due to high values of Q residual and/or the Hotelling T^2 values. In the validation set, one sample was

Table 3 Statistics for calibration and validation sets employed for PLS-ATR-FTIR determination of food nutritional parameters

Analyte (units)	Set	Samples	Range	Mean value	SD
Total fat (g per 100 g)	Cal	43	0.07–18.4	11	5
Protein (g per 100 g)		44	1.95–26.2	16	6
Total carbohydrates (g per 100 g)		42	55.5–94.8	70	10
Energy values (kJ per 100 g)		42	1650–2010	1900	90
Total fat (g per 100 g)	Val	12	0.05–18.6	11	5
Protein (g per 100 g)		11	2.57–24.8	16	7
Total carbohydrates (g per 100 g)		13	57.2–95.0	70	10
Energy values (kJ per 100 g)		13	1660–2008	1900	100
Total fat (g per 100 g)	Pred	12	0.06–15.4	10	5
Protein (g per 100 g)		11	2.8–22.8	17	7
Total carbohydrates (g per 100 g)		13	58.3–95.1	70	10
Energy values (kJ per 100 g)		13	1660–1971	1900	107

For calibration (Cal) and validation (Val) datasets, reference values are reported. For this last dataset, predicted values (Pred) are also reported. Range: concerns minimum-maximum value
SD standard deviation

Fig. 7 Mean ATR-FTIR spectrum of mixed food samples employed through this study with indication of spectral interval selection for total fat, proteins, carbohydrates, and energy. *Note:* regions between 4000–600 and 2300–1900 are not considered (see text for details). *Values in italic* correspond to the number of latent variables (# LVs)



removed from fat and two samples from proteins. Total variance is explained with 98.8, 96.1, 98.6, and 99.2 % of the X data block and 93.3, 94.2, 93.8, and 92.3 % of the Y data block for total fat, proteins, carbohydrates, and energy value, respectively. The most important calibration parameters of the developed PLS-ATR-FTIR models are summarized in Table 4, where the number of latent variable for each analyte, RMSEC, RMSECV, and R^2 of calibration and cross validation can also be seen.

PLS-ATR-FTIR prediction capability

An external validation set was used to confirm the suitability of the model built after choosing the appropriate preprocessing. Figure 8 shows the regression between the results obtained by the PLS-ATR-FTIR for the determination of total fat, protein, total carbohydrate, and energy value in human diets

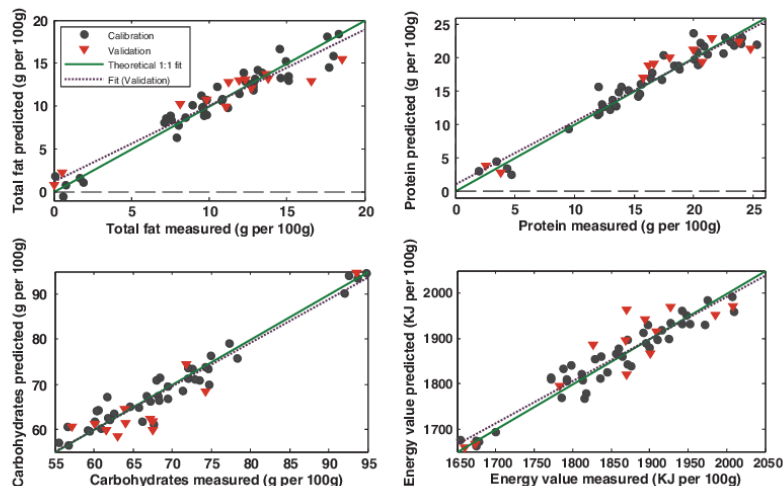
and baby foods versus those determined using the reference methods, indicating that calibration and validation sample points are closely distributed near the optimum regression line between predicted and measured values. Only one sample was removed from the validation set of fat and two samples of proteins, prior to the calculation and validation of the final PLS models of those analytes. Outliers can be identified by the Q residual versus the Hotelling T^2 values. The most important calibration and prediction parameters of the developed PLS-ATR-FTIR models are summarized in Table 4. Acceptable RRMSEP of 5.7 % was achieved in the cases of carbohydrates and the error of energy value was 2.3 %. In the case of total fat and proteins, the relative prediction errors were 16 and 11 %, respectively. Within each model, similar values for the RMSECV and RMSEP were obtained, indicating good model performance with correlation of determination of

Table 4 Main characteristics of PLS-ATR-FTIR models for the determination of the major components of mixed foods

Analyte (units)	Preprocessing	LVs	RMSEC	R^2 cal	RMSECV	R^2 CV	RMSEP	RRMSEP	R^2 pred
Total fat (g per 100 g)	Mean center	6	1.2	93.3	2.3	77.5	1.7	16.1	91.0
	1st derivative								
Protein (g per 100 g)	Mean center	7	1.4	94.2	2.4	84.0	1.9	11.4	93.0
	1st derivative								
Total carbohydrates (g per 100 g)	Mean center	5	2.4	93.8	3.8	84.8	4.0	5.7	92.0
	1st derivative								
Energy values (kJ per 100 g)	Mean center	7	24.9	92.3	42.1	78.4	42.6	2.3	84.1
	1st derivative								

LV number of latent variables, RMSEC root mean square error of calibration, RMSECV root mean square error of cross validation, RMSEP root mean square error of prediction, RRMSEP relative root mean square error of prediction, R^2 Cal coefficient of determination of calibration, R^2 CV coefficient of determination of cross validation, R^2 pred coefficient of determination of prediction

Fig. 8 Regression between predicted and reference values of total fat, proteins, carbohydrates, and energy value in human diets and baby food by PLS-ATR-FTIR using regions selected by iPLS



91.0, 93.0, 92.0, and 84.1 for total fat, proteins, carbohydrates, and energy value, respectively. Table 3 includes the main characteristics of the validation set parameters for all nutrients predicted by PLS-ATR-FTIR models. In order to evaluate the accuracy of the proposed procedure, relative errors were calculated between the reference value and the predicted value for the set of validation samples. Accuracy relative errors obtained for the prediction of total fat, proteins, carbohydrates, and energy value vary between -16.8 and 25.4 , -13.8 and 16.2 , -11.5 and 13.8 , and -2.7 and 5 %, respectively. Variation coefficients by replicates of three individual spectra of 11 samples of the validation set were determined, and values between 0.4 and 15.4 , 0.8 and 12.2 , 0.2 and 5.0 , and 0.1 and 3.0 % for total fat, proteins, carbohydrates, and energy value, respectively, were obtained.

Comparison between NIR and MIR for the analysis of mixed foods

Multivariate calibration methods developed using NIR and MIR spectroscopy for the determination of major components of human diets and baby foods, concerning total fat, proteins, total carbohydrates, and energy, were quite acceptable. However, according to the data summarized in Tables 2 and 3, calibration and cross validation errors were lower for PLS-NIR models in all cases, with values of fat, proteins, and carbohydrates between 0.4 and 1.3 g per 100 g and 7.9 kJ per 100 g of energy value for RMSEC, while for RMSECV, the values were from 1.0 to 1.6 g per 100 g of fat, proteins, and carbohydrates and 20.5 kJ per 100 g of energy value. In the case of PLS-ATR-FTIR models, the values of RMSEC were from 1.2 to 2.4 g per 100 g of fat, proteins, and carbohydrates

and an energy value of 24.9 kJ per 100 g and between 2.3 and 3.8 g per 100 g in the case of fat, proteins, and carbohydrates and an energy value of 42.1 kJ per 100 g for RMSECV. Additionally, regarding the predictive capability of PLS-NIR, the coefficients of determination for total fat, proteins, carbohydrates, and energy value were 96.7 , 98.1 , 98.9 , and 96.5 , with RRMSEP values of 9.2 , 6.1 , 2.1 , and 1.0 %, respectively, while for the PLS-ATR-FTIR models, values of 91.0 , 93.0 , 92.0 , and 84.1 for coefficients of determination for total fat, proteins, carbohydrates, and energy value and RRMSEP values of 16.1 , 11.4 , 5.7 , and 2.3 % were obtained, which clearly shows the best performance of NIR data compared to MIR ones. In order to evaluate the accuracy of the NIR and MIR techniques, accuracy relative errors for the prediction of each analyte were compared. Accuracy relative errors obtained by NIR varied between -14.5 and 12.3 , -4.3 and 13.9 , -1.9 and 1.0 , and -1.8 and 1.9 % for total fat, proteins, carbohydrates, and energy value, respectively. For the abovementioned parameters, variation coefficients of $1.6-7.4$, $1.0-4.3$, $0.1-0.7$, and $0.1-1.6$ % were obtained, respectively. In front of that, the accuracy relative errors obtained by ATR-FTIR were between -16.8 and 25.4 , -13.8 and 16.2 , -11.5 and 13.8 , and -2.7 and 5 , with variation coefficients between 0.4 and 15.4 , $0.8-12.2$, $0.2-5.0$, and $0.1-3.0$ % for total fat, proteins, carbohydrates, and energy value, respectively.

From the results obtained through this work, it can be concluded that both NIR and MIR can be applied to determine quality parameters in baby food, infant fast food, and adult menus. However, it is clear that NIR presents lower calibration, cross validation, and prediction errors than MIR, with better coefficients of determination for all analytes. For all

the cases, NIR models present the best results of accuracy, with the PLS-NIR models more appropriate than PLS-ATR-FTIR for the determination of total fat, proteins, carbohydrates, and energy value in mixed foods.

Conclusions

From the results obtained through this study and taking into consideration that sample preparation employed in this work is a common step if we can refer the final results to the dried sample, it can be concluded that PLS-NIR models built after Kennard-Stone selection of a correct calibration set and careful selection of spectral ranges, preprocessing, and LVs provide predictive abilities of the major components in mixed foods in an easy, quick, and environment-friendly way, verifying that the models built are suitable to be used for the analysis of total fat, proteins, carbohydrates, and energy values in only a few minutes without using any chemicals or heating steps.

Acknowledgments The authors gratefully acknowledge the financial support of Ministerio de Economía y Competitividad and FEDER (Projects CTQ2011-25743 and CTQ2012-38635). M.C. acknowledges the "FPI grant" provided by the Ministerio de Economía y Competitividad of the Spanish government (ref. BES-2012-055404).

References

- World Health Organization (WHO) (2009) Infant and young child feeding. Model chapter for textbooks for medical students and allied health professionals
- World Health Organization (WHO) and Food and Agriculture Organization (FAO) of the United Nations (2007) Codex Alimentarius, food labelling, 5th edn. Rome
- Horwitz W (2000) Official methods of analysis of AOAC International, 17th edn
- Moros J, Garrigues S, de la Guardia M (2010) Vibrational spectroscopy provides a green tool for multi-component analysis. *TrAC Trends Anal Chem* 29:578–591
- Cen H, He Y (2007) Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends Food Sci Technol* 18:72–83
- Ferreira DS, Galão OF, Pallone JAL, Poppi RJ (2014) Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples. *Food Control* 35:227–232
- Chen J, Ren X, Zhang Q, Diao X, Shen Q (2013) Determination of protein, total carbohydrates and crude fat contents of foxtail millet using effective wavelengths in NIR spectroscopy. *J Cereal Sci* 58: 241–247
- Karoui R, Mouazen AM, Dufour É, Pillonel L, Schaller E, De Baerdemaeker J, Bosset J-O (2006) Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools. *Int Dairy J* 16:1211–1217
- Ritthiruangdej P, Ritthiron R, Shinzawa H, Ozaki Y (2011) Non-destructive and rapid analysis of chemical compositions in Thai steamed pork sausages by near-infrared spectroscopy. *Food Chem* 129:684–692
- Khodabux K, L'Omelette MS, Jhaumeer-Laulloo S, Ramasami P, Rondeau P (2007) Chemical and near-infrared determination of moisture, fat and protein in tuna fishes. *Food Chem* 102:669–675
- Shiroma C, Rodriguez-Saona L (2009) Application of NIR and MIR spectroscopy in quality control of potato chips. *J Food Compos Anal* 22:596–605
- Sørensen LK (2009) Application of reflectance near infrared spectroscopy for bread analyses. *Food Chem* 113:1318–1322
- Moros J, Iñón FA, Garrigues S, de la Guardia M (2007) Near-infrared diffuse reflectance spectroscopy and neural networks for measuring nutritional parameters in chocolate samples. *Anal Chim Acta* 584: 215–222
- Vesela A, Barros A, Synytsya A, Delgadillo I, Copíková J, Coimbra M (2007) Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder. *Anal Chim Acta* 601:77–86
- Moros J, Iñón FA, Garrigues S, de la Guardia M (2005) Determination of the energetic value of fruit and milk-based beverages through partial-least-squares attenuated total reflectance-Fourier transform infrared spectrometry. *Anal Chim Acta* 538:181–193
- Hernández-Martínez M, Gallardo-Velázquez T, Osorio-Revilla G, Almaraz-Abarca N, Ponce-Mendoza A, Vásquez-Murrieta MS (2013) Prediction of total fat, fatty acid composition and nutritional parameters in fish fillets using MID-FTIR spectroscopy and chemometrics. *LWT Food Sci Technol* 52:12–20
- Khanmohammadi M, Garmanudi AB, Ghasemi K, Garrigues S, de la Guardia M (2009) Artificial neural network for quantitative determination of total protein in yogurt by infrared spectrometry. *Microchem J* 91:47–52
- Iñón FA, Garrigues S, de la Guardia M (2004) Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. *Anal Chim Acta* 513:401–412
- Almendingen K, Meltzer HM, Pedersen JI, Nilsen BN, Ellekjær M (2000) Near infrared spectroscopy—a potentially useful method for rapid determination of fat and protein content in homogenized diets. *Eur J Clin Nutr* 54:20–23
- Kim Y, Singh M, Kays SE (2007) Near-infrared spectroscopic analysis of macronutrients and energy in homogenized meals. *Food Chem* 105:1248–1255
- Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 58:109–130
- Food and Agriculture Organization of the United Nations (2002) Food energy—methods of analysis and conversion factors. Rome
- U.S. Department of Agriculture, Agricultural Research Service, USDA Nutrient Data Laboratory (2010) Composition of foods raw, processed, prepared
- Nutritional Labelling and Education Act, U.S. Department of Health & Human Services
- Atwater WO, Woods C (1896) The chemical composition of American food materials. Washington
- Kennard RW, Stone LA (1969) Computer aided design of experiments. *Technometrics* 11:137–148
- Ward JH (1963) Hierarchical grouping to optimise an objective function. *J Am Stat Assoc* 58:236–244
- Norgaard L, Saudland A, Wagner J, Nielsen JP, Munck L, Engelsen SB (2000) Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. *Appl Spectrosc* 54:413–419
- Nawrocka A, Lamorska J (2013) In: Grundas S, Stepniowski A (eds) Advances in agrophysical research, determination of food quality by using spectroscopic methods
- Shenk J, Workman J Jr, Westerhaus M (2001) In: Burns D, Ciurczak E (eds) Handbook of near-infrared analysis,

- application of NIR spectroscopy to agricultural products, 2nd edn.
31. Naumann D (2006) Infrared spectroscopy in microbiology. *Encycl Anal Chem*. doi:10.1002/9780470027318.a0117/abstract
32. Vlachos N, Skopelitis Y, Psaroudaki M, Konstantinidou V, Chatzilarou A, Tegou E (2006) Applications of Fourier transform-infrared spectroscopy to edible oils. *Anal Chim Acta* 573–574:459–465

Artículo 2. A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane

Analytical and Bioanalytical Chemistry 409(2017) 3527-3539



RESEARCH PAPER

A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane

Mari Merce Cascant^{1,2} · Cassandra Breil¹ · Salvador Garrigues Miguel de la Guardia² · Anne Silvie Fabiano-Tixier¹ · Farid Chemat¹

Received: 19 December 2016 / Revised: 14 February 2017 / Accepted: 28 February 2017 / Published online: 7 April 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract There is a great interest in finding alternatives and green solvents in extraction processes to replace petroleum based solvents. In order to investigate these possibilities, computational methods, as Hansen solubility parameters (HSP) and conductor-like screening model for real solvent (COSMO-RS), were used in this work to predict the solvation power of a series of solvents in salmon fish lipids. Additionally, experimental studies were used to evaluate the performance in lipids extraction using 2-methyltetrahydrofuran, cyclopentyl methyl ether, dimethyl carbonate, isopropanol, ethanol, ethyl acetate, p-cymene and d-limonene compared with hexane. Lipid classes of extracts were obtained by using high performance thin-layer chromatography (HPTLC), whereas gas chromatography with a flame ionization detector (GC/FID) technique was employed to obtain fatty acid profiles. Some differences between theoretical and experimental results were observed, especially regarding the behavior of p-cymene and d-limonene, which separate from the predicted capability. Results obtained from HPTLC indicated that p-cymene and d-limonene extract triglycerides (TAGs) and diglycerides (DAGs) at levels of 73 and 19%, respectively, whereas the other studied extracts contain between 75 and 76% of TAGs and between 16 and 17% of DAGs. Fatty acid profiles, obtained by using GC-FID, indicated that saturated fatty acids (SFAs) between 19.5 and 19.9% of extracted oil, monounsaturated fatty

acids (MUFAs) in the range between 43.5 and 44.9%, and PUFAs between 31.2 and 34.6% were extracted. p-Cymene and limonene extracts contained lower percentages than the other studied solvents of some PUFAs due probably to the fact that these unsaturated fatty acids are more susceptible to oxidative degradation than MUFAs. Ethyl acetate has been found to be the best alternative solvent to hexane for the extraction of salmon oil lipids.

Keywords Green alternative solvent · n-Hexane · Cosmo-RS · Hansen parameters · Salmon fish · Lipid extraction

Introduction

Determination of lipids is an important task in food chemistry because of the effect on human health. Lipids are usually classified into nonpolar lipids [triglycerides (TAGs), diglycerides (DAGs), monoglycerides (MAGs), and sterols] and high polar lipids (free fatty acids (FFAs), phospholipids, and sphingolipids) [1], being the impact of fatty acids (FA) on the beneficial effects on human health investigated in many studies. Salmon fish has received particular attention with emphasis on the long-chain Omega-3 polyunsaturated fatty acids (LCn-3PUFAs) [2–4] because they can reduce the risk of primary cardiovascular events, including sudden cardiac death, coronary heart disease (CHD), and congestive heart failure [5–8].

A commonly used extraction method to determine lipids in food is the Soxhlet method [9], using n-hexane as solvent [10]. n-Hexane is a major solvent for the extraction of natural products in the food industry because of its low boiling point, low polarity, and chemical stability. However, it is obtained from petroleum, a nonrenewable resource, and it is classified under

✉ Miguel de la Guardia
miguel.delaguardia@uv.es

✉ Farid Chemat
farid.chemat@univ-avignon.fr

¹ Université d'Avignon et des Pays de Vaucluse, INRA, UMR408, GREEN Team Extraction, 84000 Avignon Cedex, France

² Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, Research Building, 46100 Burjassot, Valencia, Spain

the REACH Regulation as a category 2 reprotoxic and as a category 2 aquatic chronic toxic [11]. The Environmental Protection Agency (EPA) regulates n-hexane under the Clean Air Act (CAA) and has classified n-hexane as a hazardous air pollutant (HAP). n-Hexane evaporates very easily into the air and it gets dissolved only slightly in the water [12]. Therefore, in order to prevent the aforementioned pollution, the replacement of this solvent should be tried; thus, there are many attempts to find alternative solvents to substitute hexane as solvent for extraction and analysis [13–18].

Alternative solvents suitable for green chemistry are those that (1) have low toxicity, (2) are easy to recycle, (3) are inert, and (4) do not contaminate the product. In general, solvents considered as ‘green’ are the ones included in one of these categories: (1) biosolvents obtained from renewable resources, (2) eco-friendly; those with no effects on environment, human beings, and safety, the so called (EHS), (3) water, (4) liquid polymers that are nonvolatile, biocompatible, nontoxic, (5) fluorinated solvents that are nonflammable, nontoxic, (6) ionic liquids and eutectic mixtures as nonvolatile and thermally stable solvents, and (7) supercritical fluids that are inert, recyclable, nonflammable, and nontoxic [13]. Also, the alternative solvent should retain the same extraction properties as the reference ones, as the choice of the solvent is a crucial step in the substitution solvent process. To evaluate the solvation power, there are computational methods as Hansen solubility parameters (HSP) [19] and conductor-like screening model for real solvent (COSMO-RS) [20], which can be used as a real decision-making tool for the choice of alternative solvents. Semi-empirical Hansen approach correlates and predicts the ability of a solvent to dissolve major components of the sample, and COSMO-RS determines and predicts thermodynamic properties of solvents without experimental data. In these approaches, the solute molecules are considered as embedded in a cavity that is surrounded by a virtual medium and later on, the energies of interaction between target molecules and solvents are quantified with statistical thermodynamics, providing solubility predictions.

Therefore, the aim of the present study was to investigate the effect of replacing the use of hexane with alternative solvents for lipid extraction in salmon fish. Various studies on the relative solubility of fish components in different solvents by HSP and COSMO-RS were carried out, and the performances of these solvents were experimentally studied in comparison with hexane. Solvents such as 2-methyltetrahydrofuran (MeTHF), cyclopentyl methyl ether (CPME), dimethyl carbonate (DMC), isopropanol (IPA), ethanol (EtOH), ethyl acetate (EtAc), p-cymene, d-limonene, and hexane were evaluated as reference solvent. IPA, EtOH, EtAc, p-cymene, d-limonene, and MeTHF are considered bio-based solvents that come from some renewable raw materials and offer no emission of volatile organic compounds, are biodegradable, and

nontoxic. IPA and EtOH are alcohols that are obtained from the fermentation of sugars (sugar cane or sugar beet), by enzymatic hydrolysis of starch (from corn or wheat), or from lignocellulosic raw material (wood, herbs, agricultural waste) [21]. EtAc is an organic acid ester, which is obtained by the dehydrogenation of bio-ethanol [22] produced by fermentation. Limonene is a low cost, low toxicity biodegradable terpene present in agricultural wastes derived from citrus peels, and p-cymene can be obtained from limonene under microwave irradiation [23]. MeTHF is derived from renewable resources such as corncobs and bagasse, and it can replace tetrahydrofuran (THF) and diethyl ether (DEE) as solvent [24, 25]. On the other hand, CPME and DMC are solvents obtained from chemical synthesis that can be derived from by-products of the petrochemical industry or obtained by ‘green’ chemical synthesis. DMC is obtained by transesterification of propylene carbonate or by reaction of carbon monoxide, methanol, and oxygen [26], and because its nontoxicity and biodegradability properties, DMC is recognized as a ‘green’ solvent. CPME is not produced by renewable materials but it is a good candidate for replacement of ethers because this solvent outperforms other solvents in terms of EHS characteristics and is comparatively more favorable than other ethers [27].

Experimental

Materials, samples, and reagents

n-Hexane (HPLC grade), methanol, sulfuric acid, sodium chloride, chloroform, methyl acetate, acetic acid, diethyl ether, and potassium chloride were of analytical grade and were supplied by VWR International (Darmstadt, Germany). Dimethyl carbonate (DMC) (purity 99%) was from Alfa Aesar (Karlsruhe, Germany). Cyclopentyl methyl ether (CPME) (reagent plus 99.5%), 2-methyltetrahydrofuran (2-MeTHF) (purity 99.5%), isopropyl alcohol (IPA) (purity 99.5%), and p-cymene (purity 99%) were provided by Sigma Aldrich (Taufkirchen, Germany). Ethyl acetate (purity 99.5%) was sourced from Merck (Darmstadt, Germany). Ethanol (EtOH) (96%) was acquired from Fisher Scientific (Hampton, NH, USA) and d-limonene was supplied by Acros (Hampton, NH, USA).

Atlantic salmon (*Salmon salar*) samples were bought at the local market and were stored at $-20\text{ }^{\circ}\text{C}$. Samples were lyophilized for a minimum of 48 h at a chamber pressure of 0.053 mbar by using a Cosmos lyophilizer from Cryotec (Saint-Gély-du-Fesc, France). The dried sample was pulverized with a Blixer 2 domestic mixer from Robot Coupe (Montceau-en-Bourgogne, France), and the powdered sample stored in hermetic bags until extraction. The water content of salmon was 68.5%.

Computational methods

Hansen solubility parameters (HSP)

Hansen prediction gives an evaluation of the ability of a solvent to dissolve major components of the sample, which defined solubility parameters by the following equation:

$$\delta_{\text{total}}^2 = \delta_{\text{D}}^2 + \delta_{\text{P}}^2 + \delta_{\text{H}}^2 \quad (1)$$

where δ_{total} is Hansen's total solubility parameter, δ_{D} the dispersive term, δ_{P} the polar term, and δ_{H} the hydrogen bonding term. This ability is evaluated by the relative energy difference (RED) calculated as the ratio between R_a and R_o . The ratio R_a/R_o has been called the *RED number*, reflecting the relative energy difference, being R_a the “distance” between two materials based on their respective partial solubility parameter components and R_o , the distance experimentally determined radius of the solubility sphere [28].

A RED number of 0 corresponds to no energy difference, RED numbers less than 1.0 indicate high affinity; RED equal or close to 1.0 is a boundary condition; and very high RED numbers indicate progressively low affinities [19].

For the calculations of the solubility parameters, the HSPiP software ver. 4.0 (Hansen-Solubility, Hørsholm, Denmark) offers four groups of contribution methods to determine HSP values: (1) Stefani-Panayiotou method, (2) Van

Krevelen method, (3) Hoy method, and (4) Yamamoto-Molecule Breaking (Y-MB) method.

In this work, solubility parameters were determined by using HSPiP software and Y-MB method. Besides HSP values, Y-MB can also estimate physicochemical parameters (melting point, boiling point, density, vapor pressure...). Y-MB method can give satisfying HSP prediction for complex inter-group interactions. Moreover, as the breaking is automatically generated, this method can be taken as a reference for the calculations of HSP values.

Conductor-like screening model for real solvent (COSMO-RS) calculations

COSMO-RS combines quantum chemical considerations (COSMO) and statistical thermodynamics (RS) to determine and predict thermodynamic properties without experimental data. It is a powerful method for molecular description and solvent screening based on a quantum chemical approach.

Figure 1 shows an example of the calculation step with COSMO-RS: (a) s-surface of studied solvents; (b) s-surface of triacylglycerid TAG1; (c) energies of local surface interactions between s-profiles of TAG1 and solvents; (d) s-potentials of TAG1 and solvents.

The molecule of solute was immersed in a continuous medium of dielectric constant ϵ . In this environment, a surface was built around the molecule that generated a large number

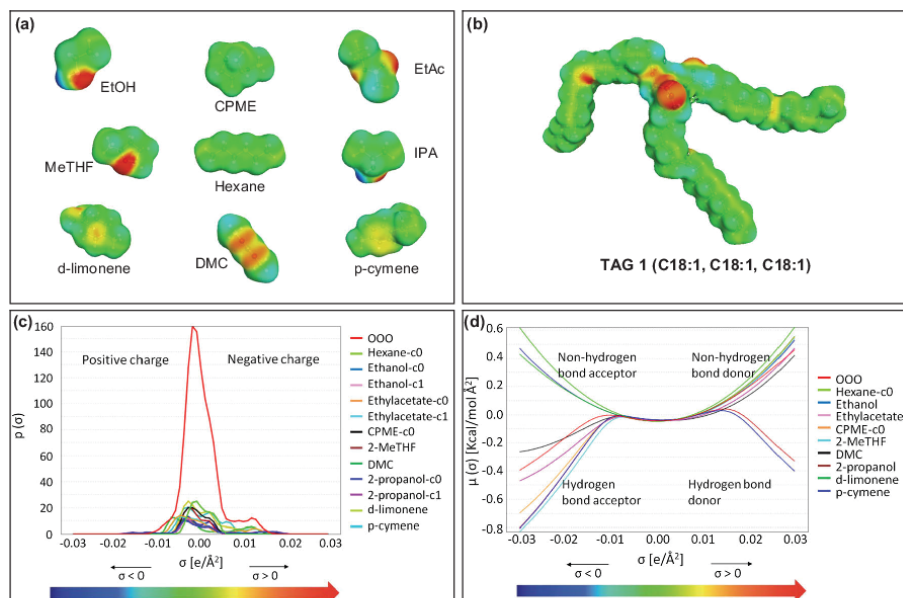


Fig. 1 Step calculation with COSMO-RS: (a) solvents s-surface; (b) triacylglycerid (TAG1) s-surface; (c) energies of local surface interactions between s-profiles of TAG1 and solvents; (d) s-potentials of TAG1 and solvents

of electrostatic charges. The structure and the distribution of charges were optimized in order to find the minimum energy of the system (molecule in its most stable state) through calculations based on algorithms, called density functional theory (DFT). This charge density is called σ -surface (Fig. 1a and b). The color code of the σ -surface gives information on the charge density at each point of the molecule. The green zone corresponding to zero charge density, the blue characterizes a density of positive charge $\delta+$, and the red color means densities of negative charge $\delta-$. In other words, green areas were usually linked to the carbon skeleton, red areas to oxygenated or nitrogenous groups, and blue areas to hydrogen atoms. This optimum charge distribution was segmented and reduced to a histogram called σ -profile (Fig. 1c). The σ -profile noted $p(\sigma)$ translates into 2D the information contained in the 3D surface. Thus, the σ -surface and the σ -profile permit characterizing a reference molecule (minimum energy, stable molecule) as a solute in a perfect conductor. Therefore, these profiles were used to interpret the affinity of a solvent, so as to understand the interaction between the compound and solvent in a mixed state and to estimate the thermodynamic properties of the system.

The thermodynamics of the molecular interactions, based on the obtained σ -profile, were then used to calculate the chemical potential of the surface segment (σ -potential) (Fig. 1d) using COSMOthermX program (ver. C30 release 13.01). The σ -potential described the likeliness of the considered compounds being able to interact with the solvents with polarity and hydrogen bonds. For each molecule, the part of the negative charge of the molecule was located on the right side (acceptor hydrogen bonds) with positive σ values, whereas the part of the positive charge was located on the left side (donor hydrogen bonds) with negative σ values. Generally, the region $\sigma \pm 0.01 \text{ e A}^{-2}$ was considered to be nonpolar or weakly polar.

In addition, the COSMOthermX also calculated the relative solubility between a solid compound and the solvent in terms of the logarithm of the solubility in mole fractions [$\log_{10}(X_{\text{solub}})$]. The logarithm of the best solubility was set to 0 and all other solvents were given relative to the best solvent.

Lipid extraction: conventional Soxhlet procedure

Twenty-five grams of lyophilized salmon samples were weighted inside two 30 mm \times 80 mm cellulose thimble (Macherey-Nagel, Düren, Germany) put in series and were placed in the extraction chamber of a 125 mL Soxhlet apparatus fitted with a condenser, which was placed on a 500 mL distillation flask containing 300 mL of n-hexane. Samples were extracted under reflux during 8 h. After this time, the content of the distillation flask was evaporated under reduced pressure by using Laborota 4000 rotary evaporator from

Heidolph (Schwabach, Germany). The described procedure was repeated with different solvents such as MeTHF, CPME, DMC, IPA, EtOH, and EtAc.

In the cases of p-cymene and d-limonene, due to their boiling point around 180 °C, solvent elimination was extracted by the addition of 50% (V/V) water in the solvents and then evaporated under reduced pressure.

Extractions with all solvents were performed in triplicate and the mean weights of the extracted salmon oils were calculated. Performance of extracts was expressed as a percentage of the total weight of lipids obtained after extraction relative to the dry weight salmon samples.

High performance thin layer chromatography (HPTLC)

To obtain the amount of lipid classes extracted with each solvent, the extracts were analyzed by using a high performance thin layer chromatography (HPTLC) from CAMAG (Muttenz, Switzerland).

Lipids were detected by charring and quantified using a CAMAG 3 TLC scanning densitometer with identification of the classes against known polar and neutral lipid standards. Typically, lipid extract was loaded as a spot onto 20 \times 10 cm silica gel 60 F254 HP-TLC plates from Merck (Darmstadt, Germany) using an ATS 5 automatic TLC sampler. HPTLC plates silica gel 60 F254 were washed by pre-development with a mixture of chloroform/methanol (2:1, v/v), followed by drying at 110 °C for 60 min on the TLC Plate Heater (CAMAG, Muttenz, Switzerland).

The first eluent to separate polar lipids was a mixture of methyl acetate/isopropanol/chloroform/methanol/KCl (0.25% solution) in a ratio of 25:25:25:10:9 (v/v/v/v/v) running to a height of 7 cm from the origin. On another plate, the second eluent was a mixture of n-hexane/diethyl ether/glacial acetic acid in a ratio of 70:30:2 (v/v/v) to a height of 7 cm from the origin to separate neutral lipids. After drying, the plate was dipped for 2 s in a modified (10 mg of primuline, 160 mL of acetone, 40 mL of water), then scanned using a TLC Scanner 3 with WinCATs software (CAMAG). Lipid classes were identified and quantified attributable the calibration curve of each target molecule. Values obtained were expressed as percentage of lipid class in total lipids.

Standard solutions were prepared using chloroform as solvent obtaining 0.2 mg mL⁻¹ of lipid in each solution. For sample solutions, approximately 1 mg mL⁻¹ of pure oil was prepared in chloroform. All solutions were stored in the dark at -20 °C until their analysis.

The different lipid classes were identified based on known standards such as phospholipids mixture from soybean composed for L-A lysophosphatidylcoline, phosphatidylcoline, phosphatidylethanolamine, and phosphatidylinositol from Supelco (Bellefonte, PA, USA), ergosterol (ergo), free fatty acids (FFA; palmitic acid), monoglycerol (MAG; DL- α -

palmitin), diacylglycerol (DAG; glyceryl 1,3-dipalmitate), and triacylglycerol (TAG; glyceryl trilpalmitate) from Sigma-Aldrich, (Germany). Retention factor (Rf) and the area of the peaks were calculated using winCATS-Planar Chromatography Manager (ver. 1.3.3; Camag). Calibration curves, with correlated peak areas and amount of lipid classes, were used to calculate the quantitative amounts of the different lipid classes [16, 17]

Gas chromatography

Fatty acids methyl esters (FAMES) were separated and identified by a 7820 A gas chromatography (GC) system from Agilent technologies (CA, USA) using flame ionization detector (GC-FID) and autosampler. GC was performed by a BD-EN14103 capillary column (30 m × 0.32 mm × 0.25 μm) using helium as a carrier gas at the speed of 33 cm s⁻¹. Two μL of samples were injected in split mode (split ratio: 1:20) at 250 °C. The oven temperature program was operated as follows: initial temperature at 50 °C for one min, increasing at a rate of 20 °C min⁻¹ to 180 °C, and at a rate of 2 °C min⁻¹ from 180 to 230 °C, held isothermally at 230 °C for 10 min.

Samples were prepared from extracted oils using acid-catalyzed trans-methylation [29]. Four hundred μL of triheptadecanoin (C17:0 TAG, Sigma Aldrich) solution containing 2 mg mL⁻¹ in hexane was used as internal standard. One mL of methanolic sulfuric acid (5% v/v) was added to approximately 20 mg of extracted oil and the mixture was heated at 85 °C for 90 min. Once solution reached room temperature, 1.5 mL of sodium chloride (0.9%) solution and 1 mL of n-hexane were added. The flask was shaken vigorously during 30 s before centrifugation at 4000 rpm for 2 min, then the organic layer was transferred into a GC vial before direct injection in the gas chromatograph. Data were collected with Agilent EZChrom Elite software and FAMES were identified compared with purified FAMES standards (Sigma Co., St. Louis, MO, USA) [16, 17].

Chemometric data treatment

Data treatment was done in Matlab 8.3.0.532 (R2014a) from Mathworks (Natick, MA, USA) using PLS Toolbox 7.5.2 from Eigenvector Research Inc. (Wenatchee, WA, USA) for principal components analysis (PCA) and dendrogram analysis.

PCA allows a reduction of variables and facilitates data exploration and interpretation of multivariate datasets. By projecting the objects of the data set into the space of the first few components, it is possible to visualize the differences among the various objects [30]. Prior to chemometric analysis, data was preprocessed by using mean center (MC) pretreatment.

Hierarchical cluster analysis (HCA) dendrogram involves a measurement of the similarity between objects to be clustered. Similarities among samples were estimated by means of distances: similar samples are characterized by small distances and the opposite for dissimilar samples [30].

Results and discussion

Computational methods

Computational simulations were conducted to evaluate the solubility of major components of salmon fish in various solvents. Major components structure present in the samples are depicted in Fig. 2 and include four triacylglycerids [TAG 1 (C18:1, C18:1, C18:1), TAG 2 (C18:1, C18:1, C18:2), TAG 3 (C16:0, C18:1, C18:1), TAG 4 (C16:0, C18:1, C18:2)], four diacylglycerids [DAG 1 (C18:1, C18:1), DAG 2 (C18:1, C18:2), DAG 3 (C18:2, C18:2), DAG 4 (C18:1, C16:0)], three free fatty acids [FFA1 (C18:1), (FFA2 (C18:2), (FFA3 (C16:0)], and ergosterol.

Table 1 shows the properties, energy efficacy, and toxicity of studied solvents: two esters (EtAc and DMC), two ethers (MeTHF and CPME), two alcohols (IPA and EtOH), two terpenes (p-cymene and d-limonene), and hexane, which was used as reference solvent.

Solute-solvent solubility by HSPs

RED values, which estimate the capacity of selected solvents to dissolve TAGs, DAGs, FFAs, and ergosterol are summarized in Table 2. The solubility of the main lipid components of salmon oil in each solvent was compared with n-hexane as reference. According to the results CPME, MeTHF, and especially d-limonene and p-cymene are better solvents than n-hexane for the solvation of these compounds. EtOH, EtAc, DMC, and IPA were worse than n-hexane in the HSPs theoretical study for TAGs. Moreover, IPA and EtOH solvents are worse solvents for DAGs and FFAs; EtAc, CPME, MeTHF, DMC, p-cymene and limonene were theoretically better than hexane for these compounds. In the case of ergosterol, EtOH, DMC, and IPA, the worst results were obtained. So, the aforementioned conclusions must be taken into consideration for the assay of extraction solvents to remove the lipids from salmon fish.

Figure 3a shows principal component analysis biplot (score and loading values) for RED values of studied compounds and solvents. The two first principal components represent 99.49% of the explained variance, being 94.21% and 5.28% for PC1 and PC2, respectively. According to the results, d-limonene and p-cymene present a similar behavior to that of hexane to theoretically solvate TAGs. IPA and EtOH are situated in opposite side

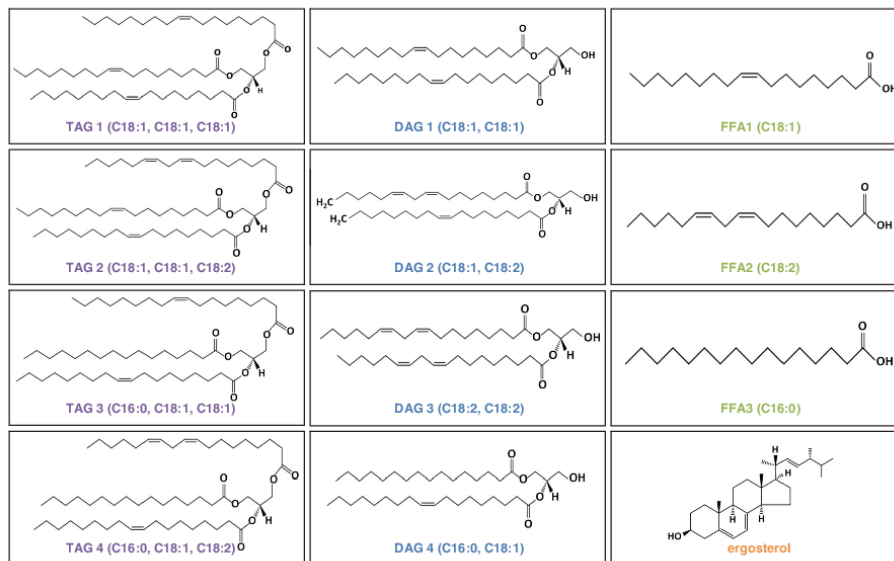


Fig. 2 Major lipidic compounds of salmon oil

of TAGs and ergosterol, thus indicating that the aforementioned solvents are less affine to these compounds. On the other hand, EtAc, MeTHF, and DMC were located closer to DAGs and FFAs than TAGs and ergosterol, so that it could be expected that the lipids could be easily solvated in the first two aforementioned compounds than in the last ones. CPME was located in the middle of the studied compounds, indicating that this solvent is less specific than the other ones and that theoretically could solvate all compounds equally.

COSMO-RS prediction

Table 3 shows the solubility of four TAGs, four DAGs, four FFAs, and ergosterol of salmon fish in the different studied solvents expressed in relative solubility, $\log_{10}(X_{\text{solub}})$, establishing the logarithm of the best solubility as 0. Taking hexane as the reference, the solubility of only two of the four TAGs considered are equal to 0. Nevertheless, all DAGs, FFAs, and ergosterol solubility values are below zero, showing that probably n-hexane is not the best solvent to be used for extraction of these

Table 1 Properties, energy efficacy, and toxicity of studied solvents

Solvents	Propierties				Energy efficacy				Toxicity		
	Molecular weight (g/mol)	Viscosity (25 °C Cp)	Log P	Flash point (°C)	Boiling point (°C)	δH vap (kJ/mol)	Cp, liq (kJ/mol K)	Energy evaporation of 1 Kg solvent (kWh)	Resource	CMR	Toxicity index
Hexane	86.2	0.31	3.94	-23.3	68.5	28.9	0.226	0.121	Petroleum	2	5
EtOH	46.1	1.095	-0.19	8.9	72.6	38.6	0.112	0.265	Cereal crop	No	5
EtAc	88.1	0.42	0.71	-3.3	73.9	31.9	0.194	0.127	Cereal crop	No	5
CPME	100.2	0.55	1.41	6.5	105.3	33	0.182	0.132	Chemical synthesis	No	4
MeTHF	86.1	0.6	0.82	-11.1	79.9	30.7	0.151	0.126	Cereal crop	No	4
DMC	90.1	0.62	0.15	18.3	90.5	33	0.455	0.194	Chemical synthesis	No	5
IPA	60.1	2.04	0.16	11.7	73	39.8	0.156	0.219	Cereal crop	No	5
p-Cymene	134.2	0.79	4.02	47.2	173.9	39.3	0.236	0.154	Wood	No	5
d-Limonene	136.2	0.923	4.45	42.8	175.4	39.5	0.249	0.157	Cereal crop	No	5

Table 2 Relative energy differences (RED) values for Hansen solubility parameters (HSP) assisted selection of alternative solvent to n-hexane for the solvation of salmon oil

Solvent	TAGs				DAGs				FFAs			ERGO
	TAG1	TAG2	TAG3	TAG4	DAG1	DAG2	DAG3	DAG4	FFA1	FFA2	FFA3	
Hexane	1.12	1.14	1.09	1.12	1.91	1.84	1.80	1.94	1.79	1.95	1.84	1.76
EtOH	4.14	4.1	4.13	4.08	3.06	3.13	3.20	3.02	3.26	3.16	3.16	3.90
EtAc	1.53	1.49	1.50	1.47	0.66	0.68	0.73	0.59	0.81	0.86	0.67	1.65
CPME	0.55	0.52	0.52	0.48	0.65	0.58	0.50	0.68	0.48	0.65	0.53	0.93
MeTHF	0.89	0.86	0.87	0.84	0.63	0.57	0.52	0.63	0.51	0.62	0.50	1.09
DMC	2.07	2.04	2.04	2.01	1.39	1.40	1.42	1.33	1.47	1.49	1.37	2.18
IPA	3.23	3.18	3.21	3.17	2.14	2.21	2.28	2.09	2.35	2.26	2.24	3.04
p-cymene	0.36	0.36	0.40	0.40	1.18	1.12	1.03	1.23	0.93	1.0	1.06	0.52
d-limonene	0.22	0.18	0.24	0.21	0.95	0.89	0.81	1.01	0.73	0.84	0.85	0.60

Reference or equivalent (grey), Better than reference (green), Lesser than reference (red)

Triglycerides (TAGs): TAG 1 (C18:1, C18:1, C18:1), TAG 2 (C18:1, C18:1, C18:2), TAG 3 (C16:0, C18:1, C18:1), TAG 4 (C16:0, C18:1, C18:2)

Di-glycerides (DAGs): DAG 1 (C18:1, C18:1), DAG 2 (C18:1, C18:2), DAG 3 (C18:2, C18:2), DAG 4 (C18:1, C16:0)

Free fatty acids (FFAs): FFA1 (C18:1), (FFA2 (C18:2), (FFA3 (C16:0)

Ergosterol (ERGO)

compounds. Considering the $\log_{10}(X_{\text{solub}})$ of TAGs, solvents such as MeTHF, CPME, EtOAc, p-cymene, and d-limonene are set to 0. DMC, IPA, and EtOH are theoretically not good substitutes to n-hexane for the extraction of TAGs and some DAGs. Nevertheless, for other components of salmon oil fish they could be considered as better option than n-hexane. EtAc, MeTHF, and CPME appear to theoretically be the best alternative solvents to n-hexane for the extraction of all studied compounds.

Figure 3b shows principal component analysis biplot for $\log_{10}(X_{\text{solub}})$ values of studied compounds and solvents. The two first principal components represent 98.41% of the explained variance, being 96.05% and 2.36% for PC1 and PC2, respectively. Hexane, d-limonene and p-cymene were together and were near to TAGs indicating that these solvents have a similar behavior when extracting the studied lipids. DMC, IPA, and EtOH are located in the opposite side of TAGs. Nevertheless they are placed near to FFAs, DAGs, and ergosterol than hexane. MeTHF, CPME, and EtAc were located in the middle, indicating that these solvents theoretically could extract TAGs, DAGs, FFAs, and ergosterol equally on salmon fish.

Experimental study: Soxhlet extractions

HPTLC results

Figure 4 shows the amount of lipid classes determined by HPTLC technique after extraction of salmon fish oil in the

studied solvents, indicating that 73–77% of the constituents extracted with these solvents are TAGs followed by DAGs, FFAs, and cholesterol with 15–20%, 5–6%, and 2–3%, respectively. MAGs, PCs, and PEs were not detectable in these samples. It can be seen that class lipids extracted by p-cymene and d-limonene are different than in the case of the other studied solvents because they contain low TAGs amounts and at the same time high amounts of DAGs. A possible explanation could be the need of a higher temperature for the boiling of these solvents over the other ones assayed, which could affect their extraction power (see Table 1). Oxidation of unsaturated fatty acids is the main reaction responsible of the degradation of lipids [31], and it can be increased by heating above 100 °C [32]. During the heating process, reactions of unsaturated fatty acyl groups in the different TAGs could happen, thus providing a wide variety of polymerized TAGs (PTG), including dimers and oligomers [33]. Particularly PUFAs are more susceptible to oxidative degradation, especially in their free form [34]. Moreover, the oil extracted by using p-cymene and d-limonene present a darker color than the other extracts, which is a possible indicator of oil degradation [35].

Figure 5a shows PCA biplot of data found by HPTLC for the lipid class and studied solvents represented by PC1 versus PC2 with 91.99 and 6.55%, respectively, explaining a total variance of 98.54%. It can be seen that d-limonene and p-cymene were close to DAGs and were located in the opposite

Fig. 3 Principal component analysis biplot (score and loading values) of studied compounds and solvents represented by PC2 versus PC1 for **(a)** RED values of Hansen and **(b)** log₁₀(X_{solub}) values obtained by using COSMO-RS scores and loadings (TAGs [blue triangle]; DAGs [pink diamond]; FFAs [green square]; ergosterol [inverted light blue triangle])

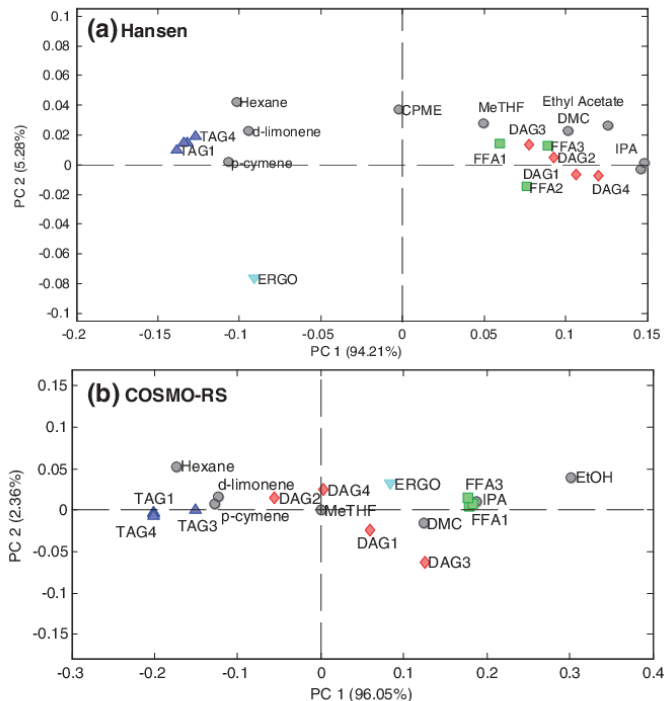


Table 3 Conductor-like screening model for real solvent (COSMO-RS) assisted solvent selection: relative solubility (log₁₀(X_{solub})) of major compounds of salmon oil in several solvents

Solvent	TAGs				DAGs				FFAs			
	TAG1	TAG2	TAG3	TAG4	DAG1	DAG2	DAG3	DAG4	FFA1	FFA2	FFA3	ERGO
Hexane	0.0	0.0	-0.17	-0.001	-1.17	-0.44	-1.63	-0.64	-1.34	-1.31	-1.27	-0.81
EtOH	-2.52	-2.53	-2.19	-2.55	-1.03	-1.52	-0.78	-1.1	-0.05	0.0	-0.001	-0.56
EtAc	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CPME	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MeTHF	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DMC	-1.51	-1.46	-1.21	-1.41	-0.99	-0.92	-0.39	-1.12	-0.51	-0.22	-0.56	-0.82
IPA	-1.49	-1.51	-1.24	-1.52	-0.34	-0.79	-0.21	-0.48	0.0	0.0	0.0	-0.21
p-cymene	0.0	0.0	0.0	0.0	-0.75	-0.19	-0.83	-0.44	-1.03	-0.9	-1.03	-0.71
d-limonene	0.0	0.0	0.0	0.0	-0.69	-0.13	-0.88	-0.37	-0.98	-0.88	-0.96	-0.65

Reference or equivalent Better than reference Lesser than reference

Triglycerides (TAGs): TAG 1 (C18:1, C18:1, C18:1), TAG 2 (C18:1, C18:1, C18:2), TAG 3 (C16:0, C18:1, C18:1), TAG 4 (C16:0, C18:1, C18:2)

Di-glycerides (DAGs): DAG 1 (C18:1, C18:1), DAG 2 (C18:1, C18:2), DAG 3 (C18:2, C18:2), DAG 4 (C18:1, C16:0)

Free fatty acids (FFAs): FFA1 (C18:1), (FFA2 (C18:2), (FFA3 (C16:0)

Ergosterol (ERGO)

logical and could evidence potential trouble on using terpenes as extraction solvents for salmon fish oil in spite of the theoretical data predicted for their solvation power.

GC-FID results

In order to compare and evaluate the extraction power of alternative solvents to hexane, the amounts of fatty acids extracted by each solvent were determined by using GC-FID. Table 4 shows the lipid composition of salmon oil extracted. It can be seen that the predominant fatty acids were oleic (C18:1) (38–40%), linoleic (18:2) (16–17%), and palmitic acid (C16:0) (11–12%), which represent more than 70% of the total fatty acids in extracted oil. There are some fatty acids with low concentration such as linolenic acid (C18:3) around 5–6%, and stearic acid (C18:0), palmitoleic acid (C16:1), eicosenoic acid (C20:1), eicosapentaenoic acid (EPA; C20:5), and docosahexaenoic acid (DHA; C22:6) near of 3%. Others such as docosapentaenoic acid (DPA; C22:5), eicosadienoic acid (C20:2), and myristic acid (C14:0) were present in extracts with values between 1 and 2%.

In general, total saturated fatty acids (SFAs) concern between 19.5 and 19.9% of extracted oil, whereas monounsaturated fatty acids (MUFAs) are in the range between 43.5 and 44.9% and PUFAs between 31.2 and 34.6%. For all studied solvents, similar results were obtained except on using limonene and especially p-cymene. Extracts with p-cymene and limonene have reduced percentages of some PUFAs due probably to the fact that these unsaturated fatty acids are more susceptible to oxidative degradation than MUFAs. It can be seen that the lowest value of extracted PUFAs in the oil was for p-cymene extract with 31.2% followed by limonene with 32.8%. At the same time, an increase of stearic acid (C18:0) together with an unknown methyl fatty acid in p-cymene extract were observed. In the case of limonene extract, this increase was observed for stearic acid (C18:0) and oleic acid (C18:1).

Figure 5b shows the PCA biplot regarding GC data of lipid extraction by using different solvent. The two first principal components represent 95.29% of the explained variance, which are 81.57% and 13.72% for PC1 and PC2, respectively. PCA results provided a clear separation of p-cymene from the other solvents assayed in the direction of PC1 being located near to unknown FFA. However, hexane, EtOH, EtAc, CPME, MeTHF, DMC, IPA, and limonene are located in the middle of all FFAs, appreciating that MeTHF seems closer to C13:0 and limonene to C18:1. Also, EtOH and IPA are closer to C22:5, C20:5, and C22:6 than hexane, EtAc, CPME, and DMC solvents. However, as it can be seen the PCA biplot obtained was highly overlapped by the main part of lipids determined and solvents considered.

In order to obtain a visualization of the solvent extraction power similarities, hierarchical cluster analysis (HCA) using Ward's linkage distance method was performed. As can be

seen from Fig. 6, for a minimum cutoff value of 1.0, there are four different groups of extracts. The first group is composed by hexane, EtAc, CPME, DMC, and MeTHF; the other groups are based on alcohols (IPA and EtOH) and the additional two groups concern terpenes. It can be seen that solvent situated nearest to hexane was EtAc, followed by CPME, indicating that the extracts obtained with the aforementioned solvent contain a similar composition as those obtained with hexane. On the other hand, p-cymene and limonene were located in the side opposite of hexane, indicating a different behavior on the extraction of salmon fish lipid based on data found by GC. In general, it can be concluded that ester and ether extracts are similar to those found with hexane, especially using EtAc, followed by alcohols such as EtOH and IPA. However, extracts obtained by using terpenes, as d-limonene and p-cymene, are clearly different than those obtained with the rest of solvents. Inset in Fig. 6 shows the properties of hexane and EtAc are the best alternative to hexane for extractions of salmon fish oil.

Figure 7 shows lipid yields found in the extracted oil. Values obtained were around 28.9–33.9% in dry matter, with no significant difference between hexane and studied solvents at 95% confidence level except for IPA extract, which provided a low lipid yield value.

Conclusions

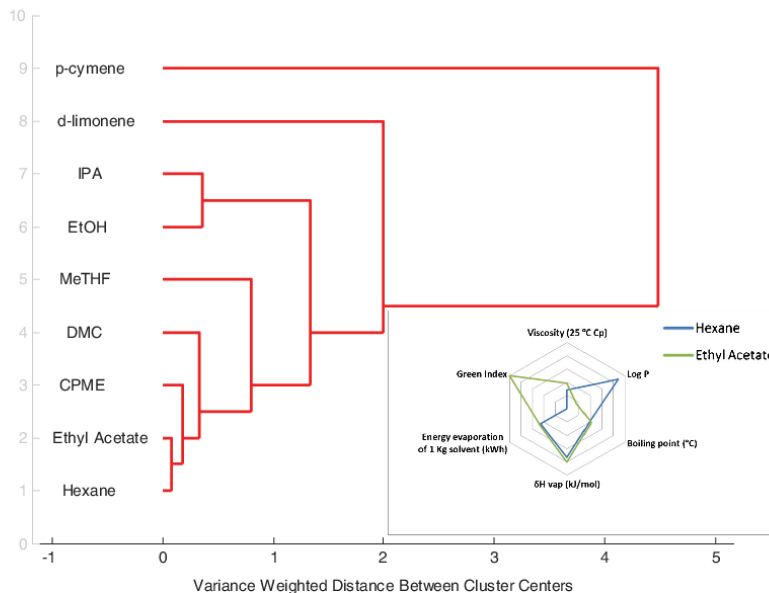
Hansen and Cosmo-RS simulation studies of the solvation character of different solvents for a series of lipid classes present in salmon fish oil indicated that d-limonene and p-cymene were similar to hexane for solvating TAGs over other lipidic compounds. Also, Hansen solubility RED values predicted that CPME could be the best solvent to solubilize all compounds considered, being more specific than the rest of solvents evaluated for each lipid class. However, obtained COSMO-RS values indicate that CPME, EtAc, and MeTHF could solvate in the same extent TAGs, DAGs, FFAs, and ergosterol compounds.

On considering the experimental data found by HPTLC and GC-FID for the analysis of salmon fish oil extracts, HPTLC indicate that TAGs were the predominant class of the lipid extracted with 73–77% followed by DAGs (15–20%), FFAs (5–6%), and ergosterol (2–3%). However, it was found that p-cymene and limonene extracts contained increased levels of DAGs and reduced amounts of TAGs than other solvent extracts. This phenomenon could be explained by the lipid degradation due to the use of high temperatures to work using p-cymene and limonene in spite of the promising values found by prediction of RED values of HSP and $\log_{10}(X_{\text{solub}})$ of COSMO-RS, which indicate that p-cymene and limonene are theoretically the best solvents to dissolve TAGs. Moreover, the amounts of fatty acids extracted by using each solvent obtained by GC-FID indicated

Table 4 Fatty acid composition of salmon oil extracted with *n*-hexane and alternative solvents determined by using GC-FID

FFA	Extracted amounts of fatty acids (%) (w/w)									
	Hexane	EtOH	Ethyl acetate	CPME	MeTHF	DMC	IPA	p-Cymene	d-Limonene	
C13:0	No Detected	No Detected	No Detected	0.170±0.010	0.66±0.05	No Detected	No Detected	No Detected	No Detected	
C14:0	2.050±0.017	2.05±0.03	2.049±0.011	2.079±0.006	2.049±0.017	2.057±0.017	2.07±0.05	2.12±0.02	2.153±0.010	
C15:0	0.198±0.002	0.195±0.003	0.1969±0.0009	0.1981±0.0005	0.1963±0.0014	0.193±0.005	0.193±0.009	0.215±0.010	0.270±0.004	
C16:0	11.26±0.05	11.55±0.08	11.363±0.006	11.45±0.03	11.352±0.010	11.41±0.08	11.61±0.14	11.58±0.11	11.592±0.013	
C16:1	2.81±0.04	2.86±0.04	2.847±0.007	2.88±0.03	2.828±0.009	2.85±0.05	2.86±0.04	2.86±0.03	2.893±0.006	
C18:0	3.42±0.07	3.26±0.09	3.274±0.003	3.28±0.05	3.264±0.010	3.25±0.18	3.2±0.2	3.6±0.2	3.5±0.3	
C18:1n9	38.91±0.04	38.22±0.16	39.16±0.12	39.07±0.15	38.81±0.15	38.81±0.10	38.5±0.4	38.2±0.2	39.51±0.17	
C18:2n6	17.28±0.04	17.06±0.09	17.14±0.03	17.18±0.02	17.071±0.016	17.21±0.07	17.15±0.12	16.5±0.2	17.29±0.04	
unknown	0.1706±0.002	0.1704±0.0013	0.173±0.002	0.1663±0.0009	0.172±0.004	0.20±0.04	0.1670±0.0011	3.7±1.2	0.12±0.08	
C18:3n6	0.240±0.002	0.2375±0.0005	0.2392±0.0003	0.2365±0.0013	0.2355±0.0013	0.2360±0.0005	0.2318±0.0010	0.222±0.003	0.283±0.016	
unknown	0.272±0.001	0.2707±0.0018	0.2696±0.0006	0.2696±0.0015	0.261±0.003	0.263±0.005	0.262±0.004	0.235±0.009	0.346±0.016	
C18:3n3	5.506±0.018	5.46±0.04	5.46±0.03	5.515±0.017	5.42±0.06	5.50±0.03	5.54±0.05	5.19±0.02	5.32±0.08	
unknown	0.652±0.004	0.644±0.007	0.643±0.006	0.642±0.003	0.635±0.008	0.643±0.006	0.645±0.009	0.57±0.02	0.655±0.007	
unknown	0.2370±0.0006	0.242±0.003	0.236±0.003	0.2348±0.0014	0.232±0.003	0.245±0.005	0.242±0.004	0.13±0.09	0.268±0.004	
C20:0	0.36±0.03	0.369±0.011	0.387±0.007	0.383±0.015	0.3743±0.0006	0.48±0.12	0.372±0.020	0.36±0.02	0.398±0.017	
C20:1	2.42±0.04	2.40±0.03	2.471±0.016	2.437±0.014	2.430±0.017	2.53±0.09	2.41±0.03	2.35±0.13	2.47±0.02	
C20:2	1.165±0.008	1.160±0.012	1.179±0.004	1.1685±0.0004	1.1647±0.0013	1.177±0.009	1.162±0.004	1.09±0.04	1.124±0.005	
C20:3n6	0.425±0.004	0.4324±0.0018	0.429±0.002	0.4265±0.0017	0.422±0.003	0.425±0.006	0.421±0.006	0.383±0.004	0.487±0.015	
C20:4n6	0.4196±0.002	0.483±0.003	0.4264±0.0014	0.431±0.003	0.426±0.005	0.425±0.005	0.450±0.009	0.395±0.014	0.418±0.006	
C20:3n3	0.408±0.003	0.4054±0.0016	0.4089±0.0011	0.4097±0.0017	0.404±0.002	0.408±0.002	0.4026±0.00014	0.363±0.008	0.483±0.013	
unknown	0.800±0.004	0.796±0.004	0.791±0.004	0.779±0.003	0.785±0.007	0.788±0.006	0.789±0.006	0.678±0.014	0.734±0.004	
C20:5n3	3.075±0.010	3.29±0.05	3.07±0.03	3.01±0.03	3.03±0.02	3.13±0.05	3.24±0.06	2.62±0.13	2.79±0.03	
C22:0	0.789±0.012	0.621±0.017	0.637±0.009	0.616±0.010	0.621±0.016	0.61±0.02	0.60±0.03	0.71±0.15	0.64±0.02	
C23:0	0.204±0.012	0.196±0.004	0.193±0.002	0.1909±0.0011	0.198±0.003	0.195±0.003	0.199±0.003	0.11±0.08	0.06±0.08	
C24:0	1.493±0.006	1.565±0.005	1.497±0.004	1.453±0.003	1.4814±0.0014	1.50±0.02	1.529±0.018	1.21±0.05	1.29±0.02	
C22:5	1.493±0.006	1.565±0.005	1.497±0.004	1.453±0.003	1.4814±0.0014	1.50±0.02	1.529±0.018	1.21±0.05	1.29±0.02	
C22:6	3.81±0.04	4.49±0.16	3.97±0.03	3.87±0.10	4.00±0.04	3.94±0.10	4.26±0.11	3.3±0.2	3.34±0.04	
Σ SFAs	19.8	19.8	19.6	19.6	19.5	19.7	19.8	19.9	19.9	
Σ MUFA	44.3	43.5	44.5	44.4	44.1	44.2	43.8	43.5	44.9	
Σ PUFA	33.8	34.6	33.8	33.7	33.7	33.9	34.4	31.2	32.8	

Fig. 6 Cluster dendrographic classification of solvents employed based on their methyl fatty acid analysis by GC after MC pretreatment established using Ward's linkage distance. Inset: Comparison of properties of hexane and EtAc

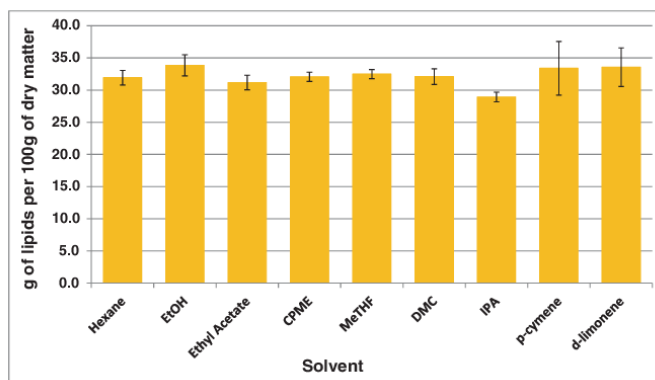


that all solvents provide similar fatty acids profile except limonene and especially p-cymene extracts, which had low amounts of some unsaturated fatty acids probably because double bonds present in PUFAs are prone to oxidation due to high temperatures involved using of limonene and p-cymene.

The strategy employed through this study could be employed as a guide to look for green alternative solvents to the use of hexane for the extraction of lipids from salmon fish. In short, it can be concluded that there are some differences between theoretical predicted solvation values of lipid class and experimental studies on the use of alternative solvents for lipid extractions, and these differences could be attributable to computational

simulations that do not consider some aspects as time and temperature of extraction or other parameters such as enthalpy of vaporization or melting point of considered compounds. These conditions can modify the lipidic extracted fraction. In general, experimental results showed that EtAc was the most suitable solvent to replace hexane due to its properties to extract all classes of lipids such as TAGs, DAGs, FFAs, and ergosterol in salmon samples. Given that other parameters as boiling point, Log P, toxicity category, and energy required for the evaporation of the solvent have been considered, it can be concluded that EtAc could be a good alternative solvent for the extraction of salmon oil lipids. EtAc does not impart emissions of volatile

Fig. 7 Comparison of lipid extraction yields obtained by using the studied alternative solvents



organic compounds, is a biodegradable solvent, and has nontoxic effects. Therefore, it can be a greener solvent than hexane.

Alternatively, the exposed strategy could be also useful to evaluate alternative solvents in other fields in order to maintain the extraction capability of classic solvents but improving the sustainability of the processes by avoiding deleterious side effects.

Acknowledgements M.C. acknowledges the FPI grant (BES-2012-055404) and (EEBB-I-16-11854) provided by the Ministerio de Economía y Competitividad of the Spanish government. The financial support of the Generalitat Valenciana (Project PROMETEO II 2014/077) is also acknowledged.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Samples from salmon already harvested were acquired in the super-market. Thus, no ethical permits were needed (Directive 2010/63/EU).

References

- Manirakiza P, Covaci A, Schepens P. Comparative study on total lipid determination using Soxhlet, Roese-Gottlieb, Bligh & Dyer, and Modified Bligh & Dyer extraction methods. *J Food Compos Anal.* 2001;14:93–100.
- Henriques J, Dick JR, Tocher DR, Bell JG. Nutritional quality of salmon products available from major retailers in the UK: content and composition of n-3 long-chain PUFA. *Br J Nutr.* 2014;112:964–75.
- Pando ME, Auborg SP. Concentrating n-3 fatty acids from crude and refined commercial salmon oil. *Czech J Food Sci.* 2014;32:169–76.
- Vinagre J, Rodríguez A, Larraín MA, Auborg SP. Chemical composition and quality loss during technological treatment in coho salmon (*Oncorhynchus kisutch*). *Food Res Int.* 2011;44:1–13.
- Lavie CJ, Milani RV, Mehra MR, Ventura HO. Omega-3 polyunsaturated fatty acids and cardiovascular diseases. *J Am Coll Cardiol.* 2009;54:585–94.
- Strobel C, Jahreis G, Kuhn K. Survey of n-3 and n-6 polyunsaturated fatty acids in fish and fish products. *Lipids Health Dis.* 2012;11:144.
- Calder PC. Very long chain omega-3 (n-3) fatty acids and human health. *Eur J Lipid Sci Technol.* 2014;116:1280–300.
- Kaur G, Cameron-Smith D, Garg M, Sinclair AJ. Docosapentaenoic acid (22:5n-3): a review of its biological effect. *Prog Lipid Res.* 2011;50:28–34.
- Sahena F, Zaidul ISM, Jinap S, et al. PUFAs in fish: extraction, fractionation, importance in health. *Compr Rev Food Sci Food Saf.* 2009;8:59–74.
- Abdulkadir M, Abubakar GI, Mohammed. Production and characterization of oil from fishes. *J Eng Appl Sci.* 2010;5:1–5.
- Regulation (EC) No. 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labeling, and packaging of substances and mixtures.
- Public Health Statement, n-Hexane. Public Health Service Agency Toxic Substance Dis. Registration.
- Chemat F, Abert Vian M (eds). *Alternative solvents for natural products extraction, green chemistry and sustainable technology.* Springer-Verlag, Berlin Heidelberg; 2014.
- Virost M, Tomao V, Ginies C, Chemat F. Total lipid extraction of food using d-limonene as an alternative to n-hexane. *Chromatographia.* 2008;68:311–3.
- Bertouche S, Tomao V, Hellal A, Boutekedjiret C, Chemat F. First approach on edible oil determination in oilseeds products using alpha-pinene. *J Essent Oil Res.* 2013;25:439–43.
- Breil C, Meullemiestre A, Vian M, Chemat F. Bio-based solvents for green extraction of lipids from oleaginous yeast biomass for sustainable aviation biofuel. *Molecules.* 2016;21:1–14.
- Sicaire AG, Abert Vian M, Fine F, Carré P, Tostain S, Chemat F. Experimental approach versus COSMO-RS assisted solvent screening for predicting the solubility of rapeseed oil. *OCL.* 2015;22:D404.
- Li Y, Fine F, Fabiano-Tixier AS, et al. Evaluation of alternative solvents for improvement of oil extraction from rapeseeds. *Comptes Rendus Chim.* 2014;17:242–51.
- Hansen CM. Hansen solubility parameters. A User's Handbook. 2nd ed. Boca Raton: CRC press, Taylor & Francis Group; 2007.
- Klamt A. The COSMO and COSMO-RS solvation models. *Wiley Interdiscipl Rev Comput Mol Sci.* 2011;1:699–709.
- Gupta AVJP. Sustainable bioethanol production from agro-residues: a review. *Renew Sustain Energy Rev.* 2015;41:550–67.
- Colley SW, Fawcett CR, Rathmell C, Tuck MWM. Process for the preparation of ethyl acetate. US Patent 6,809,217. Davy Process Technology Limited. 2004.
- Yates M, Marti MJ, Casal B, Iglesias M, Esteban M, Ruiz-Hitzky E. Synthesis of p-cymene from limonene, a renewable feedstock. *Appl Catal B Environ.* 2008;81:218–24.
- Pace V, Hoyos P, Fernandez M, Sinisterra JV, Alcantara AR. 2-Methyltetrahydrofuran as a suitable green solvent for phthalimide functionalization promoted by supported KF. *Green Chem.* 2010;12:1380–2.
- Nardi M, Sindona G, Costanzo P, Oliverio M, Procopio A. Eco-friendly stereoselective reduction of α , β -unsaturated carbonyl compounds by $\text{Er}(\text{OTf})_3/\text{NaBH}_4$ in 2-MeTHF. *Tetrahedron.* 2015;71:1132–5.
- Kumar P, Chandra V, Mani I. Dimethyl carbonate synthesis by transesterification of propylene carbonate with methanol: comparative assessment of Ce-M (M = Co, Fe, Cu, and Zn) catalysts. *Renew Energy.* 2016;88:3.
- Henderson RK, Jiménez-González C, Constable DJC, et al. Expanding GSK's solvent selection guide—embedding sustainability into solvent selection starting at medicinal chemistry. *Green Chem.* 2011;13:854–62.
- Hansen CM. The universality of the solubility parameter. *Ind Eng Chem Prod Res Dev.* 1969;8:2–11.
- Morrison WR, Smith LM. Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride-methanol. *J Lipid Res.* 1964;5:600–8.
- Ballabio DA. MATLAB toolbox for principal component analysis and unsupervised exploration of data structure. *Chemom Intell Lab Syst.* 2015;15:1–9.
- Muik B, Lendl B, Molina-Díaz A, Ayora-Cañada MJ. Direct monitoring of lipid oxidation in edible oils by Fourier transform Raman spectroscopy. *Chem Phys Lipids.* 2005;134:173–82.
- Vaskova H, Buckova M. Thermal degradation of vegetable oils: spectroscopic measurement and analysis. *Procedia Eng.* 2015;100:630–5.
- Kuligowski J, Quintás G, Garrigues S, de la Guardia M. Monitoring of polymerized triglycerides in deep-frying oil by on-line GPC-FTIR spectrometry using the science based calibration multivariate approach. *Chromatographia.* 2010;71:201–9.
- Lofthsson T, Ilievska B, Asgrimsdottir GM, Ormarsson OT, Stefansson E. Fatty acids from marine lipids: biological activity, formulation, and stability. *J Drug Deliv Sci Technol.* 2013;34:71–5.
- Gertz C. Chemical and physical parameters as quality deep-frying process—changes at elevated temperature. *Quality.* 2000;102:566–72.

Artículo 3. Determination of total phenolic compounds in compost by infrared spectroscopy

Talanta 153(2016) 360-365



Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Determination of total phenolic compounds in compost by infrared spectroscopy



M.M. Cascant^a, M. Sisouane^b, S. Tahiri^b, M. EL Krati^b, M.L. Cervera^a, S. Garrigues^{a,*},
M. de la Guardia^a

^a Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, Research Building, 46100 Burjassot, Valencia, Spain

^b Laboratory of Water and Environment, Department of Chemistry, Chouaib Doukkali University, Faculty of Sciences El Jadida, Morocco

ARTICLE INFO

Article history:

Received 20 January 2016

Received in revised form

29 February 2016

Accepted 5 March 2016

Available online 9 March 2016

Keywords:

Near-infrared spectroscopy

Mid-infrared spectroscopy

Total phenolic compounds

Compost

Partial least square

ABSTRACT

Middle and near infrared (MIR and NIR) were applied to determine the total phenolic compounds (TPC) content in compost samples based on models built by using partial least squares (PLS) regression. The multiplicative scatter correction, standard normal variate and first derivative were employed as spectra pretreatment, and the number of latent variable were optimized by leave-one-out cross-validation. The performance of PLS-ATR-MIR and PLS-DR-NIR models was evaluated according to root mean square error of cross validation and prediction (RMSECV and RMSEP), the coefficient of determination for prediction (R_{pred}^2) and residual predictive deviation (RPD) being obtained for this latter values of 5.83 and 8.26 for MIR and NIR, respectively.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Phenolic compounds are widely distributed in the nature. They are mostly derived from various agricultural and industrial activities, including waste discharge from wood preservatives, coking, textiles, plastics, dyes, paper, herbicides industries and the partial degradation of phenoxy contaminants in remediation processes [1]. Palm et al. (2001) considered the concentration of phenolic compounds as a good parameter to manage the organic matter amendment in agriculture [2]. Composting is a bioremediation technique that offers the advantage of degrading the organic wastes and provides suitable fertilizers for soils. Through mineralization and humification, the organic wastes can be transformed in a humic acid rich substance that can offer plant nutrients. The compost process reduces the volume of the final product and decreased gradually the concentration of phenolic compounds [3]. The contaminants in the waste are degraded during composting, and the final product could be used as a soil amendment [4]. Several researches in literature reported the level of phenolic compounds as a conventional parameter of compost quality, and also as an indication about the stability of the product which is related to low levels whereas the high concentration of phenols evidences their instability [3].

The concentration of total phenolic compounds (TPC) in vegetables and in biowastes is frequently determined by using classical chemical methods like titration including a previous extraction step. These using methods are accurate but they need a big amount of reagents, skilled labor and time for analysis. Therefore, a rapid and simple alternative is required.

Infrared spectroscopy in both, near (NIR) and middle (MIR) regions offers non-destructive, fast and relatively cheap possibilities for qualitative and quantitative analysis of different compounds in food and environmental samples, being thus employed for the determination of phenolic compounds in different vegetal materials [5–8]. Attenuated total reflectance (ATR) MIR spectrum provides information about the frequencies of fundamental molecular vibrations and generally exhibits sharp absorption bands and distinct spectral [5]. On the other hand NIR included the part of the electromagnetic spectrum comprised between 14,000 and 4000 cm^{-1} , related with overtones and combination of the fundamental vibrations of C–H, O–H and N–H bonds in organic molecules [7]. So, both infrared methodologies could be complementary to evaluate the presence of molecular compounds in samples of different nature.

However, in our knowledge, there is no reported research work to determine the TPC in compost by NIR or MIR and thus, the main objective of this study has been the development of a fast, non destructive and simple method for the direct determination of TPC in compost samples based on the built of partial least square (PLS) models from samples analyzed by the Folin-Ciocalteu reference method.

* Corresponding author.

E-mail address: salvador.garrigues@uv.es (S. Garrigues).

<http://dx.doi.org/10.1016/j.talanta.2016.03.020>

0039-9140/© 2016 Elsevier B.V. All rights reserved.

2. Experimental

2.1. Chemical reagents

Pure methanol HPLC grade (99.99% purity) from Scharlau (Barcelona, Spain) was used for the extraction of phenolic compounds; gallic acid (3,4,5-trihydroxybenzoic acid) was used as reference compound to build the calibration by using Folin-Ciocalteu phenol reagent from Scharlau (Barcelona, Spain) and sodium carbonate from Guinama (Valencia, Spain) was employed for pH adjustment before measurement.

2.2. Instruments

Fourier transform-NIR spectra of all samples were obtained in the diffuse reflectance (DR) mode using a Multipurpose Analyzer (MPA) from Bruker (Bremen, Germany), equipped with an integrating sphere, a quartz beam splitter, and PbS detector.

MIR spectra were obtained by using Fourier transform middle infrared, model Tensor 27 from Bruker equipped with a DLaTGS detector. Sample spectra were measured by ATR mode using a DuraSampleIR accessory for solids from Smiths Detection Inc. (Warrington, UK) with three-reflection diamond/ZnSe DuraDisk plate.

An Hewlett Packard 8452A diode array spectrophotometer (Palo Alto, USA) was employed to perform the determination of TPC content by the Folin-Ciocalteu colorimetric method.

2.3. Samples

95 Samples of compost, obtained from three different classes sources were evaluated through this study. Samples cover a wide range of TPC from 0.07 to 4.83 mg of gallic acid equivalent per gram of dry weight (mg GAE g^{-1} DW). The aforementioned samples were originated from processes of composting and co-composting mixtures of: (i) sewage sludge, (ii) municipal solid waste and, (iii) various vegetable biowastes; such as sawdust, straw, provence cane, bagasse from sugarcane, palm bark, horse manure, etc. Some of these composts have been studied in deep in our laboratory [9] and other ones are under research. All samples of compost employed were obtained by aerobic fermentation (composting) by using laboratory-scale experiment in laboratory.

The three main classes of compost evaluated in this study include: (i) compost original from mineral mature (Cm=9 samples); (ii) maturity organic compost (Co=68 samples), and (iii) vegetable origin immature compost (Cv=18 samples), all of them representing different stages of the composting process. All samples were analyzed by the reference method and IR spectra recorded without any chemical sample treatment.

2.4. Reference method for total phenolic compounds determination

The determination of TPC was performed according to the Folin-Ciocalteu colorimetric method following the previously published procedure [10]. The extracts of all samples were obtained by magnetic stirring of 1 g dry powder sample with 10 ml pure methanol for 30 min. Extracts were kept for 24 h at 4 °C and separated by centrifugation using a Digicen 20 centrifuge from Ortoalresa (Madrid, Spain) working at 3000 rpm for 10 min. The supernatants were stored at 4 °C until analysis. A volume of diluted sample extract was added to 0.5 ml of distilled water and 0.125 ml of the Folin-Ciocalteu reagent. The mixture was shaken and allowed to stand for 6 min, before addition of 1.25 ml of 7% Na_2CO_3 . The solution was then adjusted with distilled water to a final volume of 3 ml and mixed thoroughly. After incubation in the dark, the absorbance at 760 nm was measured in front of a blank

prepared in the same way that samples. Total phenolic content of all samples was expressed as mg GAE g^{-1} DW using a calibration curve established with gallic acid standard solutions. All measurements were performed in triplicate.

2.5. DR-NIR procedure

Samples, dried at 105 °C during 12 h, were placed in clear glass vials of 22 mm external diameter and 48 mm height to obtain NIR spectra by DR. DR-NIR spectra was obtained in Kubelka-Munk units. All NIR spectra were acquired between 14,000 and 4000 cm^{-1} by co-adding 50 scans and using an optical resolution of 4 cm^{-1} . The background spectrum was acquired from the closed integrating sphere using the same instrumental conditions than those employed for sample measurement. Three measurements of each sample were obtained by rotating 120° approximately, the sample vial position between replicates in order to ensure a better representativity of spectra. The mean values of the triplicate spectra of each sample were employed to build PLS calibration models and prediction.

2.6. ATR-MIR procedure

100 mg of sample, dried at 105 °C during 12 h, were deposited on the ATR accessory and pressed. Spectra were recorded in the range between 4000 and 550 cm^{-1} , by accumulating 50 scans per spectrum, at a resolution of 4 cm^{-1} . A spectrum of the clean ATR crystal, obtained immediately before acquiring the spectrum of each sample, was used as background. Mean spectra of triplicate measurement for each sample were used to build the PLS models and for sample prediction.

2.7. Data treatment

The Opus 6.5 software from Bruker was used for data acquisition in both, FT-IR and FT-NIR instruments. Data treatment was carried out using in-house written functions employing Matlab 8.3.0.532 (R 2014a) from MathWorks (Natick, MA, USA). For principal component analysis (PCA) and for partial least square (PLS) calibration and validation model, the PLS Toolbox 6.7.1 from Eigenvector Research Inc. (Wenatchee, WA, USA) was used.

An infrared spectrum consists of many related variables (wavenumbers) which are difficult to be analyzed. The PCA is a well-known technique for reducing the dimensionality of multivariate data while preserving most of the variances and this technique has been used to identify correlations among a set of variables and to transform the original set of variables to a new set of uncorrelated ones called principal components (PCs) [6]. For this purpose the PCA was applied to IR spectra to extract the most informative data set and to identify the different classes of samples.

On the other hand, PLS is a linear model frequently used in chemometrics to establish the relationship between predicted values of a property (axis y) and reference values (axis x) based on the modelization of signals obtained for a reference population of well identified samples [11]. Before building PLS models, spectra of samples were split in calibration and validation subsets with 76 and 19 objects respectively. To select the most appropriate calibration set the Kennard Stone (K-S) algorithm was used [12]. K-S algorithm starts by spectra preprocessing with the two most distant objects as members of the calibration set. The algorithm takes the objects into the training set based on their Euclidean distance. From already selected objects, calibration set was used to build the models and validation set was employed to evaluate their predictive capability.

To select the best number of latent variable leave-one-out cross validation method was employed being evaluated the root mean

square of cross validation (RMSECV) and for evaluating the prediction capability of the models, the root mean square of prediction (RMSEP) was used. Residual predictive deviation (RPD) is a statistical parameter used to evaluate the performance of models builds for the prediction of unknown samples. It was calculated as the ratio between standard deviation (SD) of TPC values of the prediction set and the RMSEP result and it is employed as an additional parameter to evaluate the methods. Generally, an RPD value greater than 3.0 is considered adequate for analytical purposes with excellent prediction accuracy, between 2.5 and 3.0 implies that the model has a good precision and between 2.0 and 2.5 indicates that the model has an approximate precision [5].

3. Results and discussion

3.1. Total phenolic concentration in compost samples

The TPC of samples, obtained from different origin, was determined by using the colorimetric method of Folin-Ciocalteu and interval of data found are summarized in Table 1. The minimum concentration found was $0.07 \text{ mg GAE g}^{-1} \text{ DW}$ and 4.83 was the maximum value of this parameter. Variations were due to the major ingredients composite materials as mineral substances, plants or biowaste. The lowest values were present in mineral composts with a low percentage of organic matter, less than 30%, and high concentrations of phenolic compounds were found in biogenic wastes with a reach proportion of organic matter, higher than 70%. Results obtained agree well with the concentration of organic fraction present in the composts samples.

3.2. DR-NIR spectra of compost samples

Fig. 1 shows the original NIR spectra of samples in the region between $14,000$ and 4000 cm^{-1} , corresponding to all samples from the three classes Cm, Co and Cv. All spectra were characterized by a similar profile. However, it can be indicated the differences between spectra in terms of intensity. The main bands of samples were present around at 8300 , 6850 and 5160 cm^{-1} being obtained the highest intensity observed in spectra of class Cv samples due to their high concentration of phenolic compounds followed by Co and Cm sample class.

The bands around 8300 , 6850 and 5160 cm^{-1} have been reported in the literature as related to the O–H groups of phenol [19,20]. Additionally, other bands, located approximately at 5800 and 4670 cm^{-1} , were reported as overtones of C–H stretch vibration, usually ascribed to the presence of phenol, lipid, oils and aromatic groups in plant [21], are present in the NIR spectra of compost samples at different intensity levels.

3.2.1. PCA of DR-NIR compost spectra

PCA has been used to evaluate the spectra of all samples from different classes of compost. The PCA using a data pretreatment based on mean center (MC) and standard normal variation (SNV)

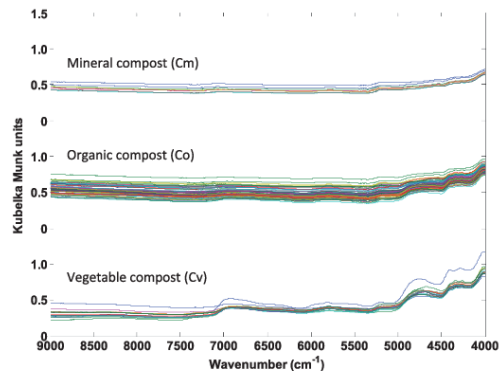


Fig. 1. DR-NIR spectra of compost samples in the range $9000\text{--}4000 \text{ cm}^{-1}$. Note: spectra were shifted on the absorbance-axis to avoid spectral overlapping.

in the region between 9000 and 4000 cm^{-1} was built. A PCA explained variance of 98.61% can be found by using 4 principal components. Fig. 2 shows the scores plot for PC2 versus PC1. PC1, with an explained variance of 86.59%, presents the aggregation of samples into two groups according to the concentration of TPC and no essential differences among the samples from classes Cm and Co were found due to small differences of TPC between them. Whereas PC2, with 8.76% of additionally explained variance, it confirmed the small difference in NIR spectra between samples concerning Cm and Co.

3.2.2. PLS-DR-NIR model

In order to build PLS-DR-NIR models suitable to evaluate the TPC of compost samples, the whole data set was split into calibration and validation set, with 76 and 19 objects respectively. The K–S algorithm was used to select the most appropriate calibration set. The calibration set must include the best representative samples assigned regarding their phenolic content based on the IR spectra. The main parameters of calibration and validation sets correspond to TPC range between 0.07 and $4.83 \text{ mg g}^{-1} \text{ GAE}$ for the calibration set and between 0.11 and $4.48 \text{ mg g}^{-1} \text{ GAE}$ for the validation one.

In order to determine the best quantitative model, to evaluate TPC in compost, different spectral regions with different data preprocessing were compared, including first derivative, SNV and multiplicative standard correction (MSC) (see Table 2). In our study, best results were obtained using the region between 9000 and 4000 cm^{-1} by using MC and SNV as data preprocessing. The correlation between predicted and reference values for TPC in compost samples can be seen in Fig. 3.

Using 7LVs, a RMSEC of 0.16 with an R_{cal}^2 value of 0.97 was found, being RMSECV 0.29 with R_{CV}^2 of 0.91 and RMSEP of 0.19, being R_{pred}^2 0.99 and RPD value of 8.26, which evidences that accurate results can be obtained on using this method.

Table 1

Total phenolic compounds present in standard samples as a function of their origin and ingredients.

Compost origin	Ingredients	Total phenolic (mg GAE g ⁻¹ DW)	Number of samples
Mineral compost (Cm)	Biowaste and biowaste additives: various sewage sludges, animal fertilizer, bio-waste, sawdust, palm bark	0.07–0.15	9
Organic compost (Co)	Including a mixture containing organic and mineral matter	0.10–0.80	68
Vegetable compost (Cv)	Biowaste additives: various provence cane, leaves, bagasse from sugarcane, shoots	0.48–4.83	18
...			
Total samples used for PLS modeling (calibration and validation)			95

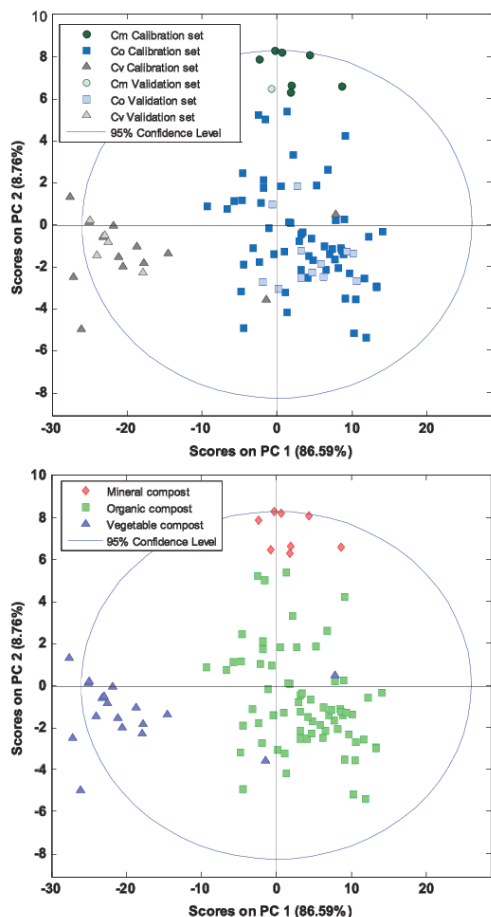


Fig. 2. Scores plot of principal component analysis based on DR-NIR spectra.

Additionally, the coefficients of determination obtained were statistically significant for data set used in calibration and validation and highly correlated with the reference concentration of total phenols and predicted value obtained from DR-NIR spectra [22].

Table 2

Parameters of calibration, cross validation and external validation of PLS DR-NIR method for the determination of TPC in composts using different spectral preprocessing methods applied to the NIR spectral data.

Preprocessing	Regions (cm ⁻¹)	LVs	RMSEC (mg GAE g ⁻¹ DW)	R _{cal} ²	RMSECV (mg GAE g ⁻¹ DW)	R _{CV} ²	RMSEP (mg GAE g ⁻¹ DW)	R _{pred} ²	RPD
MC, SNV	9000–4000	7	0.16	0.97	0.29	0.91	0.19	0.99	8.26
MC, FD	9000–4000	2	0.27	0.91	0.31	0.88	0.38	0.98	4.20
MC, MSC, FD	9000–4000	6	0.16	0.97	0.22	0.94	0.22	0.98	7.34
MC, SNV	6100–5385	4	0.27	0.91	0.40	0.81	0.24	0.99	6.73
MC, FD	6100–5385	7	0.11	0.98	0.31	0.90	0.19	0.99	8.22
MC, MSC, FD	6100–5385	5	0.32	0.87	0.40	0.80	0.26	0.98	6.21

Note: MC: mean centering; SNV: Standard normal variate; FD: first derivative; MSC: multiplicative scatter correction; R²: coefficient of determination for calibration (cal), cross-validation (CV) and prediction (pred).

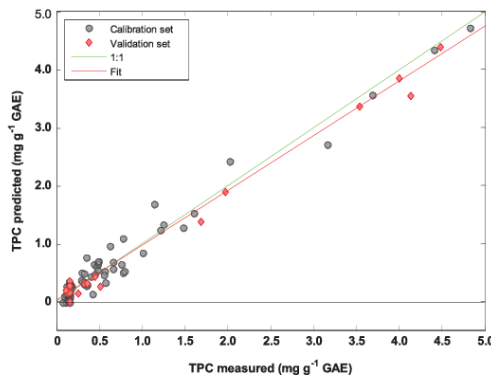


Fig. 3. Correlation between predicted and measured total phenolic compounds by the PLS DR-NIR.

3.3. ATR-MIR spectra of compost samples

Fig. 4 shows the ATR-MIR spectra of all samples separated in the three classes of composts: Cm, Co and Cv in the region between 4000 and 550 cm⁻¹. The region between 2240 and 1880 cm⁻¹ was not considered for further data treatment due to high absorbance of ATR crystal. As can be seen the spectra were characterized by similar profile and typical relative intensities of the mean bands. The intense bands detected in the 3500–3100 cm⁻¹ region are originated from O–H of molecule water [13,14] and of N–H stretching [15]. Bands found in the region between 2950 and 2850 cm⁻¹ are related to the organic matter content and due to the aliphatic methylene bands at 2920 and 2850 cm⁻¹ corresponding to asymmetric and symmetric stretching, respectively. The most intense bands were detected in the region between 1760 and 800 cm⁻¹, this region includes the information concerns of phenolic molecules, such as the stretching band of carbonyl C=O groups in the region 1712–1704 cm⁻¹ and C=C stretching bands at 1612–1608 and 1519–1516 cm⁻¹ [6]. The bands centered at 1635 cm⁻¹ could be assigned to O–H groups of polyphenol [16]. The absorbance at 1200–1050 cm⁻¹ has been assigned to phenolic compounds and to the C–O stretching vibration of polysaccharide [6]. The band at 1385 cm⁻¹ has been attributed to O–H and C–O of phenolic, and COO⁻, C–H of methylene [13]. The peak centered at 875 cm⁻¹ present in class of mineral compost can be used as an indicator of advanced mineralization [14]. The similar region between classes was located between 800 and 550 cm⁻¹. It is dominated by several vibrations of minerals from carbonates, quartz, etc., but contains also bands characteristics of C–H aromatic and N–H amines.

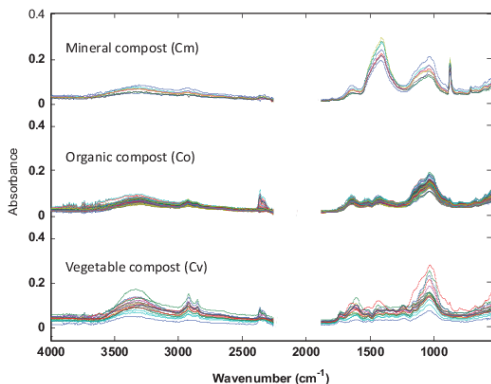


Fig. 4. ATR-MIR spectra of different classes of compost between 4000 to 550 cm^{-1} of all composts. Note: The region between 2240 and 1880 cm^{-1} was deleted due to the high absorbance of the ATR crystal. Spectra were shifted on the absorbance-axis and displayed in groups to avoid spectral overlapping.

3.3.1. PCA of ATR-MIR compost spectra

PCA permits the extraction of systematic variations in data set and could be used for classification of samples and interpretation of their differences. In this study, it was employed to classify compost samples from their ATR-MIR spectra using different wavenumber ranges from 4000 to 550 cm^{-1} such as 3000–2800, 1760–550 cm^{-1} and combinations between the different regions. The PCA built using a data pretreatment based on MC and SNV provided results in order to clearly identify the origin of samples. A PCA explained variance of 96.97% can be found by using 5 principal components. Fig. 5 shows the scores plot of the first two PCs of the model, which permit the aggregation of samples in three different groups in accordance with their nature. PC1 explains 68.16% of the total variance, which illustrates the aggregation of samples according to the concentration of total phenols. Whereas the PC2 explains 18.68% of variance and it can be clearly distinguished the differences between the Co and Cv groups. The Cm group presents

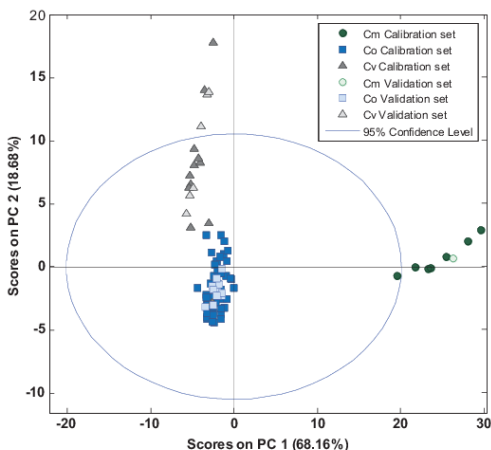


Fig. 5. Scores plot of principal component analysis based on ATR-MIR spectra.

the highest difference between composts evaluated due to the influence of their high stability and mineralization associated with the low concentration of total phenols concentrations.

As compared with DR-NIR spectra the PCA of compost samples provide a best separation of the different compost types, being recommended the use of ATR-MIR measurements for evaluative analysis and classification of compost samples according their origin.

3.3.2. PLS ATR-MIR models

In order to build PLS-ATR-MIR models suitable to evaluate the TPC of compost samples, the same calibration and validation sets that previously used for PLS-DR-NIR modeling were employed in order to compare both, MIR and NIR regions for PLS. Different spectral regions and pretreatment strategies were considered (see Table 3) that included the first derivative, MSC or SNV. In our study, best results were obtained using a MC and SNV pretreatment and the region between 1760 and 550 cm^{-1} , that contains the most important bands concerning TPC. To validate the models an independent set of compost samples not considered for calibration was used. To confirm the adequate use of previously defined distribution of samples in calibration and validation sets, a PCA were considered. Results (shown in Fig. 5) indicate that both sample sets were comparable.

Fig. 6 shows the regression between results obtained by using PLS-ATR-MIR for determination of TPC in compost, and those found by using the reference Folin-Ciocalteu method. Using 4 latent variables the RMSEC was 0.23 with R_{cal}^2 of 0.94, being the RMSECV equal to 0.28 and RMSEP of 0.27 well R_{CV}^2 and R_{pred}^2 were of 0.91 and 0.98 respectively, and the RPD value was 5.83. These characteristics are reported in the literature [17] indicating that the method provides a relative of robust model. On the other hand, the R_{pred}^2 equal to 0.98 indicates that predicted values of total phenol are close related to the measured total phenolic. The R_{cal}^2 was greater than 0.8 being considerate as acceptable [18]. In our case R_{cal}^2 equal to 0.94 can be considerate as a good parameter of the PLS model. Additionally, a RPD of 5.83 is clearly greater than 3. So, it can be considered that the method proposed for TPC determination in compost, based on ATR-MIR spectra is adequate for analytical purposes and with adequate prediction accuracy [5].

3.4. Comparison between MIR and NIR models

Multivariate calibration methods developed using MIR and NIR spectroscopy offer a sustainable alternative for the determination of TPC in compost samples. However, on comparing the main analytical features of the two developed methods it can be concluded that NIR provides an improved prediction power as compared with MIR; based on the RPD values. However, the RMSEP for the validation set of samples was 0.19 for NIR and 0.27 for MIR. In any cases, the two proposed methodologies can be employed for the determination of TPC, being the NIR measurements early obtained by using glass vials in front of the need of a suitable clean protocol to avoid cross-contamination in obtaining ATR-MIR spectra.

From the results obtained through in this study, it can be concluded that both, MIR and NIR can be applied to determine the TPC in compost samples. However, it is clear that models with NIR created from DR-NIR measurements provided a better modeling (lower values of calibration, cross-validation and prediction) than those obtained from ATR-MIR spectra.

Table 3

Parameters of calibration, cross validation and external validation of PLS ATR-MIR model for the determination of TPC in composts, using the different spectral preprocessing methods applied to the spectral ATR-MIR data.

Preprocessing	Regions (cm ⁻¹)	LVs	RMSEC (mg GAE g ⁻¹ DW)	R ² _{cal}	RMSECV (mg GAE g ⁻¹ DW)	R ² _{CV}	RMSEP (mg GAE g ⁻¹ DW)	R ² _{pred}	RPD
MC, SNV	3000–2800 and 1760–550	4	0.22	0.94	0.28	0.90	0.33	0.97	4.76
MC, FD	3000–2800 and 1760–550	5	0.23	0.94	0.31	0.88	0.31	0.96	5.08
MC, MSC, FD	3000–2800 and 1760–550	3	0.22	0.94	0.27	0.91	0.40	0.96	4.03
MC, SNV	1760–550	4	0.23	0.94	0.28	0.91	0.27	0.98	5.83
MC, FD	1760–550	4	0.26	0.92	0.31	0.88	0.33	0.96	4.80
MC, MSC,FD	1760–550	3	0.22	0.94	0.24	0.93	0.33	0.97	4.89

Note: MC: mean centering; SNV: Standard normal variate; FD: first derivative; MSC: multiplicative scatter correction; R²: coefficient of determination for calibration (cal), cross-validation (CV) and prediction (pred).

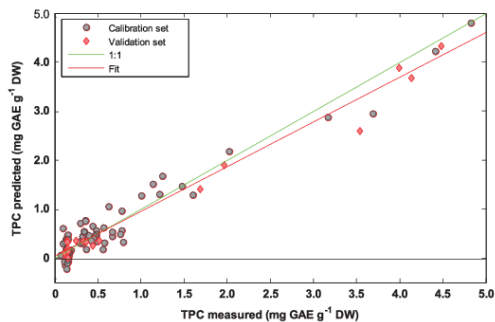


Fig. 6. Diagram of correlation between predicted and measured total phenols obtained by PLS ATR-MIR.

Acknowledgments

Authors gratefully acknowledge the financial support of Ministerio de Economía y Competitividad and FEDER (Projects CTQ2014-52841-P and CTQ2012-38635). M.C acknowledges the “FPI grant” provided by the Ministerio de Economía y Competitividad of the Spanish government (ref. BES-2012-055404). M.S. acknowledges the “Becas Jóvenes Investigadores 2015 de países en vías de desarrollo del Programa de Cooperación 0’7” by, University of Valencia.

References

[1] Z. Yaoyu, T. Lin, Z. Guangming, Z. Yi, L. Zhen, L. Yuanyuan, C. Jun, Y. Guide, Z. Lu, Z. Sheng, Simultaneous determination of hydroquinone and catechol in compost bioremediation using a tyrosinase biosensor and artificial neural networks, *Anal. Methods* 6 (2014) 2371–2378.
 [2] C.A. Palm, C.N. Gachengo, R.J. Delve, G. Cadisch, Ken E. Giller, Organic inputs for soil fertility management in tropical agroecosystems: application of an organic resource database, *Agric. Ecosyst. Environ.* 83 (2001) 27–42.
 [3] A.E. Ghaly, B. Zhang, D. Dave, Biodegradation of phenolic compounds in creosote treated wood waste by a composting microbial culture augmented with the fungus *thermoascus aurantiacus*, *Am. J. Biochem. Biotechnol.* 2 (2001) 90–103.
 [4] H. Borazjani, S. Diehl, H.A. Stewart, Composting of wood wastes: plywood and sawmill residue, *Res. Adv.* 5 (2000).
 [5] Z. Wu, E. Xu, J. Long, X. Pan, X. Xu, Z. Jin, A. Jiao, Comparison between ATR-IR, Raman, concatenated ATR-IR and Raman spectroscopy for the determination of total antioxidant capacity and total phenolic content of Chinese rice wine,

Food Chem. 194 (2016) 671–679.
 [6] X. Lu, J. Wang, H.M. Al-Qadiri, C.F. Ross, J.R. Powers, J. Tang, B.A. Rasco, Determination of total phenolic content and antioxidant capacity of onion (*Allium cepa*) and shallot (*Allium oschaninii*) using infrared spectroscopy, *Food Chem.* 129 (2011) 637–644.
 [7] M.J. Martelo-Vidal, M. Vázquez, Determination of polyphenolic compounds of red wines by UV-vis-NIR spectroscopy and chemometrics tools, *Food Chem.* 158 (2014) 28–34.
 [8] R. Ferrer-Gallego, J.M. Hernández-Hierro, J.C. Rivas-Gonzalo, M.T. Escribano-Ballón, Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy, *LWT – Food Sci. Biotechnol.* 44 (2011) 847–853.
 [9] G. El Kadiri Boutchich, S. Tahiri, M. Mahi, D. Gallart-Mateu, M. de la Guardia, A. Aarfane, E.K. Lhadi, M. El Krati, Characterization of activated sludge from domestic sewage treatment plants and their management using composting and co-composting in aerobic silos, *J. Mater. Environ. Sci.* 8 (2015) 2206–2220.
 [10] A. Lakhdar, H. Falleh, Y. Ouni, S. Oueslati, A. Debez, R. Ksouri, C. Abdely, Municipal solid waste compost application improves productivity, polyphenol content, and antioxidant capacity of *Mesembryanthemum edule*, *J. Hazard. Mater.* 191 (2011) 373–379.
 [11] D. Bertrand, La spectroscopie proche infrarouge et ses applications dans les industries de l'alimentation animale, *INRA Prod. Anim.* 15 (2002) 209–219.
 [12] R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Technometrics* 11 (1) (1969) 137–148.
 [13] J.S. Ribeiro, S. Garrigues, M. de la Guardia, FTIR tentative characterization of humic acids extracted from organic materials, *Spectrosc. Lett.* 34 (2001) 179–190.
 [14] C. Peltre, S. Bruun, C. Du, I.K. Thomsen, L.S. Jensen, Assessing soil constituents and labile soil organic carbon by mid infrared photacoustic spectroscopy, *Soil. Biol. Biochem.* 77 (2014) 41–50.
 [15] M.M. Cascant, S. Garrigues, M. de la Guardia, Direct determination of major components in human diets and baby foods, *Anal. Bioanal. Chem.* 407 (2015) 1961–1972.
 [16] E. Smidt, K. Meissl, The applicability of modern analytical tools in waste management Fourier Transform Infrared (FT-IR) spectroscopy, *Waste Manag.* 27 (2006) 268–276.
 [17] M. Kyraleou, C. Pappas, E. Voskidi, Y. Kotseridis, M. Basalekou, P.A. Tarantilis, S. Kallithraka, Diffuse reflectance Fourier transform infrared spectroscopy for simultaneous quantification of total phenolics and condensed tannins contained in grape seeds, *Ind. Crop. Prod.* 74 (2015) 784–791.
 [18] S.D. Silva, R.P. Feliciano, L.V. Boas, M.R. Bronze, Application of FTIR-ATR to Moscatel dessert wines for prediction of total phenolic and flavonoid contents and antioxidant capacity, *Food Chem.* 150 (2014) 489–493.
 [19] C.N.T. Frizon, G.A. Oliveira, C.A. Perussello, P.G. Peralta-Zamora, A.M. O. Camlowski, Ü.B. Rossa, R. Hoffmann-Ribani, Determination of total phenolic compounds in yerba mate (*Ilex paraguariensis*) combining near infrared spectroscopy (NIR) and multivariate analysis, *LWT – Food Sci. Technol.* 60 (2015) 795–801.
 [20] W. Li, H. Qu, Rapid quantification of phenolic acids in *Radix Salvia Miltorrhiza* extract solutions by FT-NIR spectroscopy in transreflective mode, *J. Pharm. Biomed. Anal.* 52 (2010) 425–431.
 [21] C.E. Millier, Chemical principals of near infrared technology, in: P.C. Williams, K.H. Norris (Eds.), *Near Infrared Technology in the Agricultural And Food Industries*, 2nd ed., American Association of Cereal Chemists, St. Paul, MN, 2001, pp. 19–39.
 [22] M. Kyraleou, C. Pappas, E. Voskidi, Y. Kotseridis, M. Basalekou, P.A. Tarantilis, S. Kallithraka, Diffuse reflectance Fourier transform infrared spectroscopy for simultaneous quantification of total phenolics and condensed tannins contained in grape seeds, *Ind. Crop. Prod.* 74 (2015) 784–791.

Artículo 4. Prediction of organic carbon and total nitrogen contents in organic wastes and their composts by Infrared spectroscopy and partial least square regression

Talanta 167(2017) 352-358



Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Prediction of organic carbon and total nitrogen contents in organic wastes and their composts by Infrared spectroscopy and partial least square regression



M. Sisouane^a, M.M. Cascant^b, S. Tahiri^{a,*}, S. Garrigues^b, M. EL Krati^a, G.EL Kadiri Boutchich^a, M.L. Cervera^b, M. de la Guardia^b

^a Laboratory of Water and Environment, Department of Chemistry, Chouaib Doukkali University, Faculty of Sciences, P.O. Box 20, El Jadida 24000, Morocco

^b Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, Research Building, 46100 Burjassot, Valencia, Spain

ARTICLE INFO

Keywords:

MIR and NIR
Organic carbon
Total nitrogen
Organic wastes
Composts
Partial Least Squares (PLS)

ABSTRACT

Middle and near infrared (MIR and NIR) were employed to determine organic carbon (OC) and total nitrogen (TN) in different soil organic amendments including wastes, composts and mixtures of composts and organic wastes. Prediction models based on partial least squares (PLS) regression from the spectra of untreated samples were built. Different spectra preprocessing strategies were adopted and the best number of latent variable was evaluated using leave-one-out cross-validation. Attenuated total reflectance (PLS-ATR-MIR) and diffuse reflectance (PLS-DR-NIR) models were built and evaluated from root mean square error of cross validation and prediction (RMSECV and RMSEP), coefficients of determination for prediction (R^2_{pred}) and residual predictive deviation (RPD). ATR-MIR provided a better prediction capability than DR-NIR with RMSEP, R^2_{pred} and RPD values of 2.2%, 0.76 and 2.0 for OC and values of 0.2%, 0.82 and 2.4 for TN, respectively.

1. Introduction

Composting is one of the sustainable solutions for management of organic solid wastes. It is a bioremediation technique that offers the advantages of organic waste degradation together with the production of suitable fertilizers for the soil amendment [1]. Compost constitutes an important source of carbon, organic phosphorus, nitrogen, and also increases the cation exchange capacity of soil [2], a lot of biogenic wastes being used as departure substrate in composting processes.

Organic carbon (OC) and total nitrogen (TN) are the main chemical parameters used to evaluate the possibilities of successful composting process [3]. Indeed, compost quality for land application depends on its maturity and stability [2], the ratio OC/TN being used in control of final maturity of the obtained product [4,5].

OC and TN in soils, compost and in organics wastes are frequently determined by using classical chemical methods like mineralization at hot temperature and titration for OC, and distillation after a previous mineralization with sulfuric acid in the presence of a catalyst for TN [6–8]. Those methods are accurate but they need a big amount of reagents, skilled labor and time. Therefore, rapid and simple alternative procedures are required in this field.

Infrared spectroscopy in both, middle (MIR) and near (NIR)

ranges, provided non-destructive, fast and relatively cheap methodologies for qualitative and quantitative analysis of different compounds in all type of samples, being employed for the determination of OC and TN in plants, soils [9–11], and composts [2,12].

Attenuated total reflectance (ATR) MIR spectrum provides information about the frequencies of fundamental molecular vibrations and generally exhibits specific absorption bands for the major components of samples [13]. NIR, in the wavenumber range from 14,000 to 4000 cm^{-1} concerns overtones and combination bands of the fundamental vibrations of C-H, O-H and N-H bonds [14]. So, both infrared methodologies could be complementary to evaluate the presence of compounds in samples of different nature.

The main objective of this study has been to develop a screening methodology for the determination of OC and TN in organic samples as composts, biogenic wastes, and compost-waste mixtures by using MIR and NIR methodologies. Partial least squares (PLS) was used to build the corresponding calibration and prediction models, based on the use of a series of samples previously analyzed for OC and TN using the reference AFNOR methods [15].

* Corresponding author.

E-mail address: tahiris@ucd.ac.ma (S. Tahiri).

<http://dx.doi.org/10.1016/j.talanta.2017.02.034>

Received 13 January 2017; Received in revised form 16 February 2017; Accepted 17 February 2017

Available online 20 February 2017

0039-9140/ © 2017 Elsevier B.V. All rights reserved.

2. Materials and methods

2.1. Chemical reagents

Sulfuric acid (98% purity) from Somaprol (Casablanca, Morocco) was used to mineralize the samples for organic carbon and total nitrogen determination; $K_2Cr_2O_7$ from Somaprol was used for oxidation of organic carbon; $(NH_4)_2 [Fe(SO_4)_2] \cdot 6H_2O$ was employed to titrate the excess of $K_2Cr_2O_7$ in the presence of orthophosphoric acid and using diphenylamine as indicator, from Somaprol. A Mixture of K_2SO_4 , $CuSO_4$ and selenium was used as catalyst in the mineralization step of total nitrogen. H_3BO_3 and sodium hydroxide (> 97%) were employed for total nitrogen determination using Taschiro indicator being all the reagents obtained from Prolabo (Casablanca, Morocco).

2.2. Instruments

MIR spectra were obtained by using Fourier transform middle infrared, model Tensor 27 from Bruker (Bremen, Germany) equipped with a DLATGS detector. Sample spectra were measured by ATR mode using a DuraSampleIR accessory for solids from Smiths Detection Inc. (Warrington, UK) with three-reflection diamond/ZnSe DuraDisk plate.

Fourier transform-NIR spectra of all samples were obtained in the diffuse reflectance (DR) mode using a Multipurpose Analyzer (MPA) from Bruker, equipped with an integrating sphere, a quartz beam splitter, and PbS detector.

A Kjeltect™ 2300 Analyzer Unit and Kjeltect™ System 1002 distillation unit were employed to perform the mineralization and distillation of samples for TN determination through conventional Kjeldahl method.

2.3. Sample preparation

121 samples from organic wastes, composts and waste-compost mixtures were studied through this work. Samples cover a wide range of OC from 9.23% to 45.36% based on dry weight and 0.55–4.40% for TN. The aforementioned samples were prepared from composts and organic wastes. All samples of compost employed were obtained by aerobic fermentation (co-composting) in a laboratory-scale reactor [16]. The mixture samples were prepared by adding various amounts of mature composts to organic wastes.

The three main classes of samples evaluated in this study include: i) mature composts from sewage sludge and urban solid waste (13 samples); ii) different organic wastes (9 samples), and iii) mixtures between organic wastes and mature composts (99 samples). All samples were analyzed by the conventional reference methods and directly by IR spectroscopy, in this case without any chemical sample treatment.

2.4. Reference analysis

The OC and TN reference contents were determined according to the AFNOR methods. The OC was determined according to standard AFNOR (X31-109) norm [15]. 0.05 g of dry powdered sample was oxidized at a controlled temperature of 135 °C in the presence of $K_2Cr_2O_7$ and sulfuric acid for 5 min. After cooling, the solution was adjusted with distilled water to a final volume of 200 mL and mixed thoroughly. To 25 mL of the aforementioned solution 5 mL orthophosphoric acid and few drops of the indicator (diphenylamine) were added. The excess of $K_2Cr_2O_7$ was titrated by a 1 N Mohr salt solution making three replicates for each sample.

The TN was determined based on the steam distillation according to the standard AFNOR (X31-111) norm [15]. 1 g of dry powdered sample was mineralized during 2 h with sulfuric acid in the presence of a catalyst mixture. The solution obtained was kept for 2 h at ambient temperature for cooling. Distillation was carried in the presence of

50 mL of NaOH (40%) and the distillate trapped with boric acid (4%), in the presence of the Tashiro indicator. The titration was performed by triplicate using a 0.2 N sulfuric acid solution.

2.5. ATR-MIR procedure

Approximately 100 mg of sample, dried at 105 °C for 12 h, were deposited on the ATR accessory and pressed. Spectra were recorded in the range between 4000 and 550 cm^{-1} , by accumulating 50 scans per spectrum, at a resolution of 4 cm^{-1} . Mean spectra of triplicate measurement for each sample were used to build the PLS models and for sample prediction.

The ATR crystal must be cleaned after each measurement by using cellulose tissue. A spectrum of the empty clean ATR crystal, obtained immediately before acquiring the spectrum of each sample, was used as background.

2.6. DR-NIR procedure

Samples, dried at 105 °C during 12 h with an average particle size of 250 μm , were placed in clear glass vials of 22 mm external diameter and 48 mm height to obtain their NIR spectra. Measurements were obtained in Kubelka-Munk units in the range between 14,000 and 4000 cm^{-1} by co-adding 50 scans and using an optical resolution of 4 cm^{-1} . The background spectrum was acquired from the closed integrating sphere using the same instrumental conditions than those employed for sample measurement. Three measurements of each sample were obtained. To do it samples were shaken and the sample vial position was rotate 120° between replicates in order to ensure a better representativity of signals. The mean values of the triplicate spectra of each sample were employed to build PLS calibration models and for prediction.

2.7. Data treatment

Infrared spectrum consists of many variables related to each wavenumber measurement, thus containing a lot of information from which the concentration of target analytes are difficult to be extracted directly [17]. For this purpose partial least squares (PLS), a chemometric method used to derive prediction models from IR spectra, was used [18]. The combination of spectroscopy measurements and PLS was introduced by Haaland and Thomas in 1988 [19] for the quantification of chemical compounds. Software Opus 6.5 from Bruker was used for data acquisition in both, FT-MIR and FT-NIR instruments. Data treatment was carried out using in-house written functions for Matlab 8.3.0.532 (R 2014a) from MathWorks (Natick, MA, USA). The PLS Toolbox 6.7.1 from Eigenvector Research Inc. (Wenatchee, WA, USA) was used for calibration and validation partial least square (PLS) model.

In order to build PLS models, spectra of all samples were divided into two groups: one employed for calibration and the second one for validation. The Kennard Stone (K-S) algorithm [20] was used in our study to select the most appropriate calibration set. This algorithm starts by spectra preprocessing with the two most distant objects as members of the calibration set and takes the objects into the training set based on their Euclidean distance. From already selected objects, calibration set was used to build the models and the validation set was employed to evaluate their predictive capability. Sample spectra were split in calibration and validation subsets (3/4) and (1/4) of the objects, respectively.

Two strategies were employed for selecting the region of spectra that contains main information about OC and TN: the first was based on the consideration of a specific region for each component and their characteristic absorption bands and the second on the use of iPLS that selects a subset of variables which will give superior prediction compared to using all variables in the data set [21].

Leave-one-out cross validation method was carried out to evaluate all models in order to obtain the best number of latent variables. Evaluation of prediction capability of models was based on root mean square error of cross validation (RMSECV) and root mean square error of prediction (RMSEP). Residual predictive deviation (RPD) was mainly as a statistical parameter used to evaluate the performance of PLS models for the prediction of unknown samples [22]. Moreover, Tekin et al. [23] and Chang and Laird [24] classify RPD values into three classes: i) an RPD greater than or equal to 2.0 associate models with good predictive ability, ii) with $1.4 < \text{RPD} < 2.0$ concern intermediate models and iii) RPD below 1.4 corresponds to a model with no predictive ability.

3. Results and discussion

3.1. Organic carbon and total nitrogen concentrations in samples

The levels of OC found in samples by using the reference method ranged between 9.23% as the minimum value and 45.36% as the maximum of this parameter. These variations were due to the composted materials as mineral substances, mature compost and organic wastes. Maturity of compost contributes to relatively low percentage of organic matter, less than 50%, but high concentrations of OC were found in organic wastes before the composting process.

The TN content in samples ranged from 0.55% to 4.4%, and it must be taken into consideration that composting process of biogenic wastes contributes to the decomposition of nitrogen containing compounds as amino acids and proteins.

3.2. Characteristics of ATR-MIR spectra of samples

Fig. 1 shows the ATR spectra of samples employed in this work in the wavenumber range between 4000 and 550 cm^{-1} with the indication of the main bands attributed to organic carbon and nitrogen containing compounds identified in the literature. The region located between 2240 and 1880 cm^{-1} was excluded for further data treatment due to high absorbance of the crystal employed in this ATR accessory. All spectra obtained were characterized by a similar profile, but with different intensities of bands especially for spectra of mature composts. According to previous overview literature based on vibrational functional groups, the intense bands found in the $3500\text{--}3100 \text{ cm}^{-1}$ range were assigned to O-H of water and of N-H stretching [25,26]. The range between 2950 and 2850 cm^{-1} contains bands of aliphatic methylene at 2920 and 2850 cm^{-1} . The region between 1760 and 800 cm^{-1} includes intense bands generally attributed to C=C vibration of alkenes and aromatic compounds centered at around 1640 cm^{-1} , and also to the C=O vibration of amides and carboxylated groups [27]. Indeed, the weak bands detected at 1385 cm^{-1} could be attributed to O-H and C-O of phenolic compounds, and to COO⁻ and C-H of methylene [25]. Bands in the region between 1380 and 1050 cm^{-1} are characterized by C-N stretch vibration of amines around 1320 cm^{-1} , and C-O of carboxylic acids and C-N of amides at 1260 and 1240 cm^{-1} , respectively. The peak obtained at 875 cm^{-1} has been assigned to inorganic compounds obtained at advanced decomposition process [28]. On the other hand, the region located between 800 and 550 cm^{-1} is dominated by several vibrations of minerals and contains bands characteristics of C-H aromatic and N-H amines [17].

3.3. PLS ATR-MIR models

Spectra depicted in Fig. 1 were employed to evaluate both, OC and TN, in the studied samples; being selected the most appropriated model in each case.

3.3.1. Organic carbon determination

The procedure described in Section 2.7, was applied in all cases to

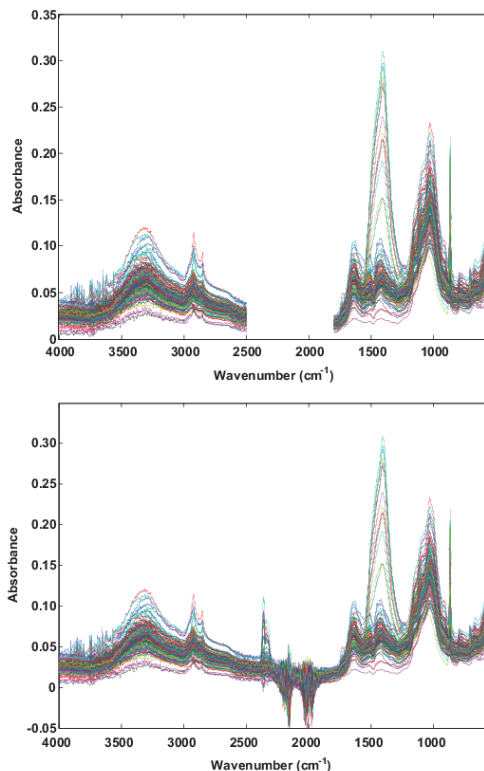


Fig. 1. ATR-MIR spectra of samples in the region from 4000 to 550 cm^{-1} .

build PLS models based on ATR-MIR for OC prediction. The whole data set was split into calibration and validation sets, with 91 and 30 objects respectively. The OC concentration of calibration and validation sets employed ranged between 9.29% and 55.0% for calibration and between 16.9% and 36.02% for validation. iPLS was adopted for the selection of spectral regions, and to do it, spectra were divided into intervals of 178 variables using 10 LVs as maximum in each one.

The assayed models, based on different preprocessing methods in different regions (see Table 1), are characterized by R^2_{cv} from 0.62 to 0.71 and RMSECV from 4.26% to 5.50% for the cross validation, and R^2_{pred} from 0.40 to 0.76, RMSEP between 2.20% and 4.55%, RPD values from 0.97 till 2.00 were obtained for the independent set of samples used for validation method.

Models assayed included first derivative (FD), orthogonal signal correction (OSC) and mean centering as spectra pretreatment (see Table 1). Based on the coefficient of determination, root mean squares error and ratio of predicted deviation it was selected the model that provided the best performance for calibration and validation which was established as the lowest RMSEP.

OSC and mean center combining regions of 3014–2300 and 1776–550 cm^{-1} was the best modelling option. Hence, this model report a RPD value reached 2.0 comprised between 2.0 and 2.5 that is considered adequate for analyte quantification of OC in unknown samples [29].

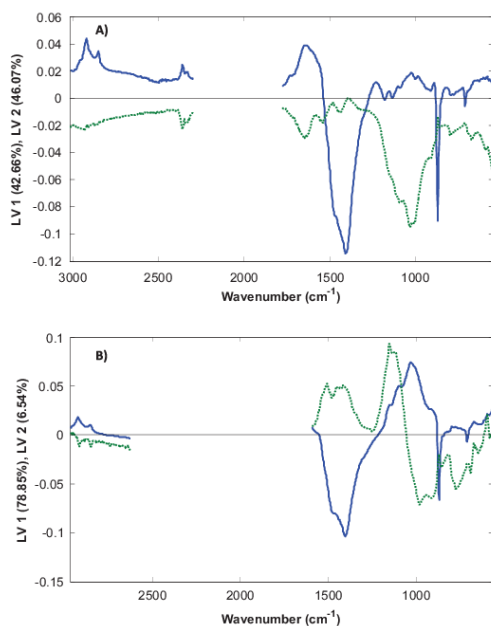
The selected model required the use of 2 latent variables and, as it can be seen in Fig. 2A the loading concerning the first and second

Table 1

Parameters of calibration, cross validation and external validation of PLS-ATR-MIR method for the determination of OC in organic wastes and compost samples using different pretreatment and spectra ranges.

Preprocessing	Regions (cm ⁻¹)	LVs	RMSEC (%)	R ² _{cal}	RMSECV (%)	R ² _{cv}	RMSEP (%)	R ² _{pred}	RPD
MC-FD	1602–1247 and 534–400	5	4.16	0.78	4.95	0.69	3.7	0.4	1.2
MC-OSC	3014–2300 and 1776–549	2	1.24	0.98	4.26	0.67	2.2	0.76	2.0
MC-OSC	1751–549	2	2.95	0.89	5.15	0.67	2.84	0.61	1.57
MC-FD	3014–2300 and 1776–549	3	4.92	0.69	5.5	0.62	3.74	0.31	1.19
MC-OSC	2966–2625 and 1250–900	4	4.16	0.78	4.8	0.71	4.55	0.4	0.97

Note: MC: Mean Center; FD: First Derivative; OSC: Orthogonal Signal Correction.

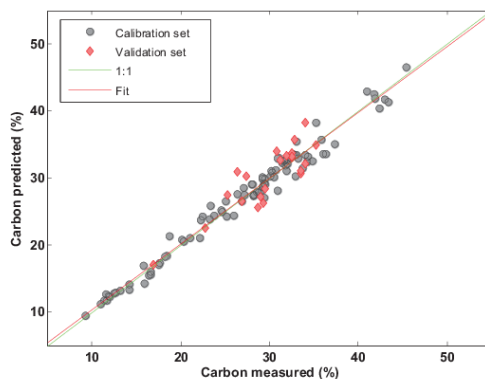

Fig. 2. Loadings of the first two latent variables selected to build the best models for OC (2 A) and TN (2B) determination by PLS-ATR-MIR.

variable, which explain respectively 42.66% and 46.07% of the data variance focused on the wavenumber ranges of 3014–2300 cm⁻¹ and specially on bands at 1406 and 871.8 cm⁻¹ for LV1 and that at 1034 cm⁻¹ for LV2.

The regression between results obtained by PLS-ATR-MIR for OC determination and those measured using the reference method is shown in Fig. 3. Using 2 latent variables the RMSEC was 1.24% with R²_{cal} of 0.98, being the RMSECV equal to 4.26% and that of prediction 2.20 with R²_{cv} and R²_{pred} of 0.67 and 0.76 respectively as indicated in Table 1.

3.3.2. Total nitrogen determination

Different PLS models, based on the ATR-MIR spectra, were developed for TN determination by adopting the same procedure described in the case of OC. Spectra were split in two subsets for calibration and validation and the mean spectrum was divided into intervals of 178 variables by using iPLS selection variable in the region between 4000 and 550 cm⁻¹. The main characteristics of calibration and validation sets involve of TN contents that ranges between 0.54% and 4.39% for the calibration set and between 1.89% to 3.71% for the


Fig. 3. Correlation between predicted PLS-ATR-MIR and measured organic carbon content of organic wastes and composts.

validation set. Based on the different preprocessing strategies, that included the first, second derivative and standard normal variate (SNV) and mean centering treatment; two models provided the best calibration and predictive validation accuracy. These models were established with combination between regions of 2966–2625 and 1593–565 cm⁻¹. Using 5 LVs, the first one based on SNV and mean center treatment and the second one based on first derivative mean center, respectively. A RMSECV of 0.38% was found in the first case being RPD value of 2.39 (Table 2).

The loadings plot of the two first LVs (See Fig. 2B) indicated that the bands at regions 2966–2625 cm⁻¹ for both LV and those at 1406 and 871.8 cm⁻¹ for LV1, and 1034 cm⁻¹ for LV2 are the most important to modelling TN from ATR-MIR spectra, being around 78.85% and 6.54% of the variance explained to LV1 and LV2, respectively.

Fig. 4 presents the regression between results obtained for TN determination by PLS-ATR-MIR in composts and different biogenic waste samples, and those measured using the reference method. 5 latent variables with MC-SNV-FD preprocessing of spectral data in combined regions of 2966–2625 and 1593–565 cm⁻¹ provided the best model for TN determination by using ATR-MIR spectra, being the root mean square error of calibration 0.38% with R²_{cal} of 0.78, and a root mean square error of cross validation equal to 0.42%, being that of prediction 0.20%, and R²_{cv} and R²_{pred} of 0.74 and 0.82, respectively.

3.4. Characteristics of DR-NIR spectra

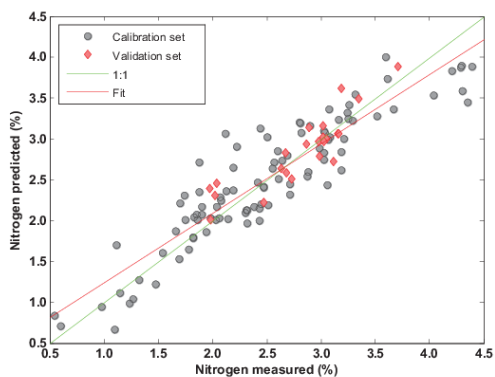
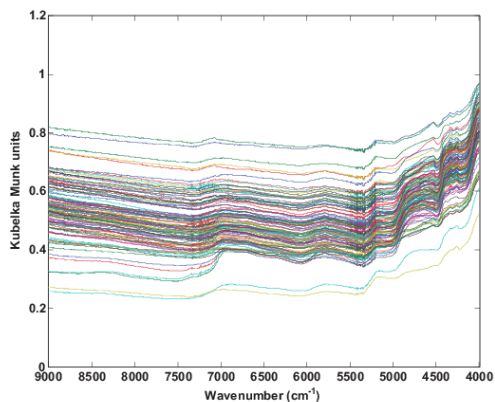
Near infrared reflectance spectroscopy measures the intensity of the absorption of near infrared light for a sample and for each wavelength [2]. So, it is also a method commonly applied in green analysis for direct determination of major compounds in organic samples which allows knowing the chemical composition much more quickly than

Table 2

Parameters of calibration, validation and external validation of PLS-ATR-MIR method for the determination of TN in organic wastes and composts samples using different pretreatment and spectra ranges.

Preprocessing	Regions (cm^{-1})	LVs	RMSEC (%)	R_{cal}^2	RMSECV (%)	R_{cv}^2	RMSEP (%)	R_{pred}^2	RPD
MC-FD	2966–2625 and 1593–565	5	0.33	0.83	0.39	0.78	0.21	0.78	2.28
MC-SD	2966–2625 and 1593–565	6	0.29	0.87	0.38	0.78	0.27	0.64	1.77
MC-SD	1593–565	5	0.33	0.83	0.39	0.77	0.25	0.67	1.92
MC-SNV	2966–2625 and 1593–565	5	0.38	0.78	0.42	0.74	0.20	0.82	2.39

Note: MC: Mean Center; FD: First Derivative; SD: Second Derivative; SNV: Standard Normal Variate.

**Fig. 4.** Predicted vs. measured TN contents in compost and organic waste samples by using PLS-ATR-MIR and the reference procedure.**Fig. 5.** DR-NIR spectra of compost and organic waste samples in the region between 14000–4000 cm^{-1} .

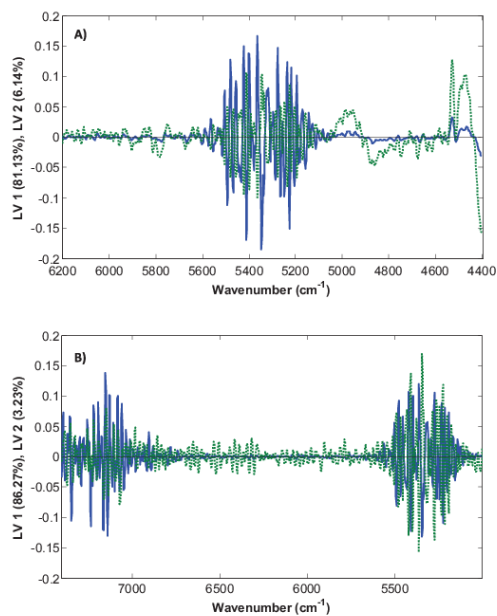
classic biochemical tools [30]. Fig. 5 shows the original NIR spectra of

Table 3

Parameters of calibration, validation and external validation of PLS-DR-NIR method for the determination of OC in organic waste and compost samples using different pretreatment and spectra ranges.

Preprocessing	Regions (cm^{-1})	LVs	RMSEC (%)	R_{cal}^2	RMSECV (%)	R_{cv}^2	RMSEP (%)	R_{pred}^2	RPD
MC-FD	6200–4400	5	3.82	0.80	4.51	0.72	4.92	0.41	1.35
MC-SD	6200–4400	4	3.23	0.85	4.79	0.69	4.09	0.61	1.60
MC-FD	8600–8000 and 6200–4400	5	3.52	0.83	4.43	0.73	4.85	0.44	1.33

Note: MC: Mean Center; FD: First Derivative; SD: Second Derivative.

**Fig. 6.** Loadings of the first two latent variables selected to build the best models for OC (6A) and TN (6B) determination by PLS-DR-NIR.

samples in the region between 9000 and 4000 cm^{-1} . Characteristic of DR-NIR spectra of all samples present four important absorbance bands around 8300, 6850, 5160 and 4670 cm^{-1} . These bands are related to the presence of functional groups containing bonds as C–H, O–H and N–H [2]. Moreover, according to the literature, NIR bands observed are mainly associated to the O–H and C–H group [31,32]. The high correlation between O–H and C–H groups was reported by Ben-Dor et al. [33]. On other hand, the bands located at 5800 and 4670 cm^{-1} were assigned to overtones of C–H stretch vibration, usually ascribed to the presence of aromatic groups and other major compound in plant [34] and O–H, N–H and C–H stretching bands could be also associated to the presence of proteins, carbohydrates and related

Table 4

Parameters of calibration, validation and external validation of PLS-DR-NIR method for determination of TN in organic waste and compost samples using different pretreatments and spectra ranges.

Preprocessing	Regions (cm ⁻¹)	LVs	RMSEC (%)	R ² _{cal}	RMSECV (%)	R ² _{cv}	RMSEP (%)	R ² _{pred}	RPD
FD-MC	10045–4000	6	0.23	0.91	0.43	0.69	0.40	0.68	1.67
FD-MC	7400–5002	7	0.26	0.88	0.40	0.74	0.38	0.74	1.79
SD-MC	7400–5002	5	0.25	0.89	0.50	0.74	0.34	0.82	2.00

Note: MC: Mean Center; FD: First Derivative; SD: Second Derivative.

compounds [35].

3.5. PLS DR-NIR models

As it was made for data obtained by ATR-MIR, DR-NIR was used for modelling the OC and TN present in compost and organic wastes being selected for each analyte the best spectra preprocessing, wave-number region and number of latent variables.

3.5.1. Organic carbon

Following the procedure indicated in Sections 2.6 and 2.7, DR-NIR spectra were processed for direct determination of OC in compost samples being iPLS used for the selection of variables. To do it, the spectral region between 9000 and 4000 cm⁻¹ was divided into intervals of 311 variables and different pretreatment strategies were assayed. Concentration of OC in the considered samples employed in both, calibration and validation sets, ranged between 9.23–55.0% and 11.68–41.9%, respectively. The best RPD value found was equal to 1.6 for PLS-DR-NIR based on the second derivative mean center in the region between 6200 and 4400 cm⁻¹ as it can be seen in Table 3.

The loading plot for the first two LV of NIR spectra (See Fig. 6A) indicate that the main part of information was provided by LV1 with an explained variance of data of 81.13% and concerns the bands between 5500 and 5000 cm⁻¹ and those around to LV2 (explained variance of 6.14%) are also located in this range and at 4987, 4528 and 4476 cm⁻¹. So, it can be concluded that band in the NIR range are not so much specific of OC.

Coefficients of determination for calibration, varied from 0.85 to 0.80, and RPD values obtained for all the models assayed indicated the moderate accuracy of prediction (1.4 < RPD < 2) [22,24,36]. The regression between results obtained by PLS-DR-NIR and data of OC measured in samples using the chemical method was not so good providing an equation of $y=0.8596x+3.8172$. With 4 latent variables, a RMSEC of 3.23% with R²_{cal} equal 0.85 was obtained, being the root mean square error of cross validation equal to 4.79 and that of prediction 4.09%, R²_{cv} and R²_{pred} were 0.69 and 0.61, respectively. So, it can be concluded that PLS-DR-NIR provides a worse prediction capability than PLS-ATR-MIR for OC determination in compost and organic wastes.

3.5.2. Total nitrogen determination

Predictive models for direct determination of TN in organic samples were established by using different preprocessing strategies of DR-NIR spectra. Based on the absorption of C-H and N-H groups, iPLS selection of variables was made and the mean spectrum was divided into intervals of 311 variables. First, second derivative and mean center pretreatment of DR-NIR spectra provided two classes of models, the first with RPD values around 2.0 and the second with RPD comprised between 1.4 and 2.0. The concentration values of % TN corresponding to calibration and validation sample sets were from 0.54% to 4.35% for the calibration set and between 1.82% and 4.39% for the validation one.

From results listed in Table 4, the best PLS-DR-NIR model was based on second derivative and mean centering pretreatment in the region of 7400–5002 cm⁻¹, using 5 LVs. The correlation coefficients of

calibration and validation were 0.89 and 0.82, being a RMSEC 0.25% and RMSEP 0.34%, respectively. Moreover, the accuracy of this model, with a RPD value of 2.0, could be considered a good performance [37,38] for direct determination of TN. Although, near infrared spectroscopy has been reported as efficient and convenient tool for the estimation of TN content in soils [39], the prediction capability was lower than that obtained based on PLS-ATR-MIR strategy, with a regression line between results obtained by PLS-DR-NIR and reference chemical data of $y=0.8977x+0.2519$.

Regarding the loadings plot for NIR (See Fig. 6B) LV1 explains 86.27 of data variance and as linked to the regions from 9000 until 7000 cm⁻¹ and from 5500 to 5000 cm⁻¹, being this second region that involves a strongest influence on the LV2 which explains 3.23% of the variance. However, as it can be seen in this figure the information about TN in DR-NIR spectra is not associated to any characteristic band specially.

4. Conclusions

The present study demonstrates the usefulness of infrared spectroscopy, MIR and NIR, to be employed for the prediction of organic carbon and total nitrogen contents in compost and organic waste samples. It can be seen that methods based on PLS-ATR-MIR provided better performance than PLS-DR-NIR. However, both offer sustainable green alternative methods for the determination of OC and TN in organic samples being suitable for screening purposes. The NIR procedures are free from the use of reagents and wastes and the MIR ones require small amounts of cell cleaning solvents and provide reduced volumes of sample wastes.

Acknowledgments

Authors gratefully acknowledge the financial support of Ministerio de Economía y Competitividad and FEDER (Project CTQ2014-52841). M.S. acknowledges the "Becas Jóvenes Investigadores 2015 de países en vías de desarrollo del Programa de Cooperación 077" by University of Valencia. M.C acknowledges the "FPI grant" provided by the Ministerio de Economía y Competitividad of the Spanish government (ref. BES-2012-055404). The Prometeo II (2014/077) project of the Generalitat Valenciana is also acknowledged.

References

- [1] A.E. Ghaly, B. Zhang, D. Dave, Biodegradation of phenolic compounds in creosote treated wood waste by a composting microbial culture augmented with the fungus *Thermoascus aurantiacus*, *Am. J. Biochem. Biotechnol.* 2 (2011) 90–103.
- [2] R. Albrecht, R. Joffre, R. Gros, J. Le Petit, G. Terron, C. Périssol, Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process, *Bioresour. Technol.* 99 (2008) 448–455.
- [3] O. Fourti, N. Jedidi, A. Hassen, Physico-chemical aspects during the composting of municipal solid wastes and sewage sludge in a semi-industrial composting plant, *Afr. J. Microbiol. Res.* 7 (12) (2013) 1055–1068.
- [4] G.F. Huang, Q.T. Wu, J.W.C. Wong, B.B. Nagar, Transformation of organic matter during co-composting of pig manure with sawdust, *Bioresour. Technol.* 97 (2006) 1834–1842.
- [5] M.F. Hirai, V. Chanyasak, H. Kubota, A standard method for measurement of compost maturity, *Bio Cycles* 24 (1983) 54–56.
- [6] A. Lakhdar, H. Falleh, Y. Ouni, S. Oueslati, A. Debez, R. Ksour, C. Abdely,

- Municipal solid waste compost application improves productivity, polyphenol content, and antioxidant capacity of *Mesembryanthemum edule*, *J. Hazard. Mater.* 191 (2011) 373–379.
- [7] F.Z. El Onagouh, L. El Fels, L. Lemée, A. Amblys, M. Hafidi, Evaluation of lignocellulose compost stability and maturity using spectroscopic (FTIR) and thermal (TGA/TDA) analysis, *Ecol. Eng.* 75 (2015) 217–222.
- [8] E. Compaoré, L.S. Nandéma, Compostage et qualité du compost de déchets urbains solides de la ville de Bobo-Dioulasso, Burkina Faso, *Tropicicultura* 28 (2010) 232–237.
- [9] V. Bellon-Maurel, A. McBratney, Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – critical review and research perspectives, *Soil Biol. Biochem.* 43 (2011) 1398–1410.
- [10] R. Gislum, E. Micklander, J.P. Nielsen, Quantification of nitrogen concentration in perennial ryegrass and red fescue using near-infrared reflectance spectroscopy (NIRS) and chemometrics, *Field Crop. Res.* 88 (2004) 269–277.
- [11] Y. Ge, J.A. Thomasson, C.L.S. Morgan, Mid-infrared attenuated total reflectance spectroscopy for soil carbon and particle size determination, *Geoderma* 213 (2014) 57–63.
- [12] L. Galvez-Sola, R. Moral, M.D. Perez-Murcia, A. Perez-Espinosa, M.A. Bustamante, E. Martinez-Sabater, C. Paredes, The potential of near infrared reflectance spectroscopy (NIRS) for the estimation of agroindustrial compost quality, *Sci. Total Environ.* 408 (2010) 1414–1421.
- [13] Z. Wu, E. Xu, J. Long, X. Pan, X. Xu, Z. Jin, A. Jiao, Comparison between ATR-IR, Raman, concatenated ATR-IR and Raman spectroscopy for the determination of total antioxidant capacity and total phenolic content of Chinese rice wine, *Food Chem.* 194 (2016) 671–679.
- [14] M.J. Martelo-Vidal, M. Vázquez, Determination of polyphenolic compounds of red wines by UV–VIS–NIR spectroscopy and chemometrics tools, *Food Chem.* 158 (2014) 28–34.
- [15] APNOR, *Qualité des sols, Méthodes d'analyses. 1ère édition, 1987.*
- [16] G. El. Kadiri Boutchich, S. Tahiri, M. Mahi, D. Gallart-Mateu, M. de la Guardia, A. Aarfane, E.K. Lhadi, M. El Krati, Characterization of activated sludge from domestic sewage treatment plants and their management using composting and co-composting in aerobic silos, *J. Mater. Environ. Sci.* 8 (2015) 2206–2220.
- [17] M.M. Cascant, M. Sisouane, S. Tahiri, M.E.L. Krati, M.L. Cervera, S. Garrigues, M. de la Guardia, Determination of total phenolic compounds in compost by infrared spectroscopy, *Talanta* 153 (2016) 360–365.
- [18] M. Zimmermann, J. Leifeld, J. Fuhrer, Quantifying soil organic carbon fractions by infrared-spectroscopy, *Soil Biol. Biochem.* 39 (2007) 224–231.
- [19] D.M. Haaland, E.V. Thomas, Partial least-squares methods for spectral analyses. 1. relation to other quantitative calibration methods and the extraction of qualitative information, *Anal. Chem.* 60 (1988) 1193–1202.
- [20] R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Technometrics* 11 (1969) 137–148.
- [21] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy, *Appl. Spectrosc.* 54 (2000) 413–419.
- [22] A. Shiferaw, C. Hergarten, Visible near infra-red (VisNIR) spectroscopy for predicting soil organic carbon in Ethiopia, *J. Ecol. Nat. Environ.* 6 (3) (2014) 126–136.
- [23] Y. Tekin, Z. Tuimsavas, A.M. Mouazen, Comparing the artificial neural network with partial least squares for prediction of soil organic carbon and pH at different moisture content levels using visible and near-infrared spectroscopy, *R. Bras. Ci. Solo.* 38 (2014) 1794–1804.
- [24] C.W. Chang, D.A. Laird, Near-infrared reflectance spectroscopic analysis of soil C and N, *Soil. Sci. Soc. Am.* 167 (2002) 110–116.
- [25] J.S. Ribeiro, S. Garrigues, M. de la Guardia, FTIR tentative characterization of humic acids extracted from organic materials, *Spectrosc. Lett.* 34 (2001) 179–190.
- [26] E. Smidt, K. Meissl, The applicability of modern analytical tools in waste management Fourier transform infrared (FT-IR) spectroscopy, *Waste Manag.* 27 (2006) 268–276.
- [27] E. Smidt, K. Meissl, M. Schwanninger, P. Lechner, Classification of waste materials using Fourier transform infrared spectroscopy and soft independent modeling of class analogy, *Waste Manag.* 28 (2008) 1699–1710.
- [28] C. Peltre, S. Bruun, C. Du, L.K. Thomsen, L.S. Jensen, Assessing soil constituents and labile soil organic carbon by mid infrared photoacoustic spectroscopy, *Soil Biol. Biochem.* 77 (2014) 41–50.
- [29] D. Cozzolino, D. Schultz, K. Allder, J. Eglinton, S. Roumeliotis, Feasibility study on the use of attenuated total reflectance mid-infrared spectroscopy for the analysis of malt quality parameters in wort, *J. Inst. Brew.* 120 (2014) 385–389.
- [30] D. Bertrand, La spectroscopie proche infrarouge et ses applications dans les industries de l'alimentation animale, *INRA Prod. Anim.* 15 (2002) 209–219.
- [31] C.N.T. Frizon, G.A. Oliveira, C.A. Perussello, P.G. Peralta-Zamora, A.M.O. Camlóski, Ú.B. Rossa, R. Hoffmann-Ribani, Determination of total phenolic compounds in yerba mate (*Ilex paraguariensis*) combining near infrared spectroscopy (NIR) and multivariate analysis, *LWT - Food Sci. Technol.* 60 (2015) 795–801.
- [32] W. Li, H. Qu, Rapid quantification of phenolic acids in *Radix Salvia Miltorrhiza* extract solutions by FT-NIR spectroscopy in transfective mode, *J. Pharm. Biomed. Anal.* 52 (2010) 425–431.
- [33] E. Ben-Dor, Y. Inbar, Y. Chen, The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process, *Remote Sens. Environ.* 61 (1997) 1–15.
- [34] C.E. Millier, Chemical principals of near infrared technology, in: P.C. Williams, K.H. Norris (Eds.), *Near Infrared Technology in the Agricultural and Food Industries 2nd ed.*, American Association of Cereal Chemists, St. Paul, MN, 2001, pp. 19–39.
- [35] A. Nawrocka, J. Lamorska, S. Grundas, A. Stepniewski (Eds.), *Determination of Food Quality by Using Spectroscopic Methods, Advances in Agrophysical Research*, 2013.
- [36] C.W. Chang, D. Laird, M.J. Matusbach, C.R. Hurburgh, Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties, *Soil Sci. Soc. Am. J.* 65 (2001) 480–490.
- [37] R.A. Viscarra-Rossel, Robust modelling of soil diffuse reflectance spectra by bagging-partial least squares regression, *J. Infrared Spectrosc.* 15 (2007) 39–47.
- [38] R. Karoui, A.M. Mouazen, É. Dufour, L. Pillonel, E. Schaller, J. De Baerdemaeker, J.O. Bosset, Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools, *Int. Dairy J.* 16 (2006) 1211–1217.
- [39] X. An, M. Li, L. Zheng, H. Sun, Eliminating the interference of soil moisture and particle size on predicting soil total nitrogen content using a NIRS-based portable detector, *Comput. Electron. Agric.* 112 (2015) 47–53.

Artículo 5. Prediction of alkaline earth elements in bone remains by near infrared spectroscopy

Talanta 162(2017) 428-434



Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Prediction of alkaline earth elements in bone remains by near infrared spectroscopy



Mari Merce Cascant, Sonia Rubio, Gianni Gallelo*, Agustín Pastor, Salvador Garrigues, Miguel de la Guardia

Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, Research building, 46100 Burjassot, Valencia, Spain

ARTICLE INFO

Keywords:

NIR
Multivariate statistics
Alkaline earth elements
Buried bones
Diagenesis
Soil

ABSTRACT

An innovative methodological approach has been developed for the prediction of the mineral element composition of bone remains. It is based on the use of Fourier Transform Near Infrared (FT-NIR) diffuse reflectance measurements. The method permits a fast, cheap and green analytical way, to understand post-mortem degradation of bones caused by the environment conditions on different skeletal parts and to select the best preserved bone samples. Samples, from the Late Roman Necropolis of Virgen de la Misericordia street and En Gil street located in Valencia (Spain), were employed to test the proposed approach being determined calcium, magnesium and strontium in bone remains and sediments. Coefficients of determination obtained between predicted values and reference ones for Ca, Mg and Sr were 90.4, 97.3 and 97.4, with residual predictive deviation of 3.2, 5.3 and 2.3, respectively, and relative root mean square error of prediction between 10% and 37%. Results obtained evidenced that NIR spectra combined with statistical analysis can help to predict bone mineral profiles suitable to evaluate bone diagenesis.

1. Introduction

Bone is composed of 50–70% mineral, 20–40% organic matrix, mainly collagen, 5–10% water, and less than 3% lipids. The mineral content of bone is mostly hydroxylapatite $[\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2]$, with small amounts of carbonate, magnesium, and acid phosphate, with missing hydroxyl groups that are normally present [1]. Under physiological conditions, hydroxylapatite is the only stable mineral in bones and is composed of 38% of calcium and 18% of phosphorous, with trace of sodium (0.6%), magnesium and small amount of other elements [2]. The number of ionic substitutions possible in biological apatite is smaller than in geologic apatites due to the limited number of available elements in the body. Among the substituting ions that are known and/or reported in bone and tooth mineral are F^- , Cl^- , Na^+ , K^+ , Fe^{2+} , Zn^{2+} , Sr^{2+} , Mg^{2+} , citrate, and carbonate [3].

Minerals are ingested from food or involuntary absorbed from the environment. Chemical composition of bones mineral is commonly employed to investigate pathologies, nutrition, injuries and other bioarchaeological and forensic issues.

The mineral composition of bones can be post-mortem modified by post-depositional processes, called as well diagenesis, and since many decades researchers have been intensively studied these natural mechanisms. Some authors have evaluated post-mortem soil contam-

ination in bones [4–9]. Other authors have studied diagenesis and degradation effects on bone matrix [10–13].

Calcium, oxygen and hydrogen are major constituents of mineral bones as hydroxylapatite [14]. Strontium is a non-essential trace element that competes to replace calcium [15] and has been linked to food consumption habits [16]. The soil composition and environmental conditions play a crucial role in diagenesis, affecting the element concentration of the buried bones.

The general aims of this study has been to develop a low cost, clean and fast strategy to understand how post-mortem degradation in bones caused by the environment could affect different skeletal parts and define an approach to select the best preserved bone samples. Therefore, a method for the determination of mineral elements in bone remains has been developed employing Fourier Transform Near Infrared (FT-NIR) spectroscopy by using chemometric tools as principal component analysis (PCA) and Partial Least Square (PLS). Inductively-Coupled Plasma Optical Emission Spectroscopy (ICP-OES) results have been used as reference data to create prediction models.

ICP-OES has been commonly used as an accurate procedure to determine mineral elements in soils and bones [17–21]. ICP-OES technique is very useful for multi-element analysis. However it requires a previous sample digestion and dissolution, involving the use of strong

* Corresponding author.

E-mail address: gianni.gallelo@uv.es (G. Gallelo).

<http://dx.doi.org/10.1016/j.talanta.2016.10.071>

Received 25 July 2016; Received in revised form 16 October 2016; Accepted 18 October 2016

Available online 19 October 2016

0039-9140/© 2016 Elsevier B.V. All rights reserved.

acids and providing non-degradable wastes. A green alternative to the aforementioned technique could be infrared (IR) spectroscopy which provides fast spectral acquisition, a cheap acquisition and maintenance cost and a sustainable method, since it does not use reagents and can be employed directly on solid samples. Combination of IR spectroscopy with chemometric data analysis could offer a significant tool for the determination of major components of bones and sediments including mineral elements.

Fourier Transform infrared (FT-IR) spectroscopy of bones has been used for various applications as collagen determination, bone crystal size, some carbonates and crystalline structure [22–25], and geochemical taphonomy [26] and as screening tool for diagenetic alteration [27,28]. On the other hand, only few studies have employed NIR for soils [29] and bones analysis [30] in archaeological contexts.

In the present study, two hundred seventy one samples obtained from at least 78 individuals, have been analyzed. Samples belong to adult and young individuals and are from the Late Roman Necropolis of Virgen de la Misericordia Street [19] and en Gil Street [31], both located in the city of Valencia (Spain). The Late Roman burial rite consisted in inhumation. At this sites, the tombs belong to the period from the I century A.D. until the beginning of the V century A.D.

Bone samples were collected from femur, tibia, humerus, radius and parietal bones, which were classified as cortical bones, and ribs as spongy bones. Furthermore, bone samples from the outer bone layer and soil samples were analyzed. Outer bone layer samples were obtained from the external bone surface directly in contact with the sediments. Principal Component Analysis (PCA) was applied to identify bone samples with a well preserved elemental composition. Therefore for the first time Partial Least Square (PLS) regression models were built to predict the concentration of calcium, magnesium and strontium in skeletal remains by NIR spectra of the mineralized and homogenized solid samples.

2. Materials and methods

2.1. Sample collection

Bone samples were collected from 78 different individuals, sampling around 5 cm of each bone and the weight of the collected sample ranged between 15 and 20 g depending on the type of bone. Bones as femur, tibia, humerus, radius and parietals, mainly composed by a cortical matrix were classified as “cortical”, and ribs mainly composed by a spongy matrix were grouped as “spongy”. The first 2 mm of bone directly in contact with the sediments was called external part of bones and was obtained employing a bistoury to scrape. All bones were sampled avoiding the osteometric points, and the sampling was carried out using a cutting toll and a micro spoon spatula.

To provide reproducible and comparable results compatible with the sensitivity of the analytical methods employed and so appropriately relate NIR obtained data with the reference method data all samples were previously mineralized in a muffle furnace, (temperature programme employed: I. 30 min at 150 °C; II. 1°/min up to 450 °C; III. 24 h at 450 °C; IV. 30 °C), and pulverized and homogenized with an agate mortar.

A total of 243 samples from Virgen de la Misericordia were analyzed. Bone samples were divided in three groups: i) internal part of bones, composed by the internal part of cortical and spongy ones (95 samples), ii) external part of bones, composed by the first external layer of cortical bones and spongy ones (94 samples) and, iii) sediments that were directly in contact with the bones (54 samples). Additionally following the mentioned classification, 28 samples from En Gil internal part of bones (8 samples), external part of bones (10 samples) and sediments (10 samples) were also analyzed.

2.2. Apparatus and methods

For diffuse reflectance near infrared spectra acquisition, a Fourier transform near infrared spectrometer, model Multipurpose Analyzer (MPA) from Bruker (Bremen, Germany), equipped with an integrating sphere was employed. This instrument is equipped with a NIR source, a quartz beamsplitter and PbS detector. For instrumental and measurement control as well as for data acquisition, Opus 6.5 software from Bruker was used.

References values about the mineral composition of samples were determined by using an Optima 5300 DV ICP-OES Perkin Elmer (Norwalk, CT, USA) equipped with an autosampler AS 93-plus and a cross flow nebulizer. Samples were previously calcinated inside a muffle furnace Biometa Lenton ECF 12145A (Lanera, España) and acid digested using a heating plate Ika C-Mag HS7.

2.3. Reference method

The digestion method was developed modifying the process described by Gallelo et al., 2013 and 2014 [19] and consisted in addition to 0.5 g of mineralized sample of 1:1 HCl and HNO₃ (using 37% HCl and 69% HNO₃ high purity stock) and digested in a water bath at 100 °C. This concentrated solution (A), was diluted 1:250 obtaining solution (B), to measure Mg and Sr. Solution (C) was obtained to analyze Ca by diluting 1:2000 solution (A). Concentrations of HCl and HNO₃ have been maintained constant in all solutions. A multi-element stock solution containing Ca, Mg and Sr at a concentration of 100 µg ml⁻¹ was employed for the preparation of the calibration standards in 50 ml volumetric flasks. To avoid the obstruction of the nebulizer system samples were filtered employing filter paper (Whatman™ N.1 of 70 mm). Concentrations ranging between 0 and 20 µg ml⁻¹ were used for Ca, Mg and Sr. Standards were obtained from Sharlab S.L. (Barcelona, Spain). The standard error of readings during the analysis ranged from 0% to 2% for the major elements considered. Bone ash NIST 1400 and soil GBW07408 were used as standard reference materials for evaluating the accuracy of the analytical method and Re was used as internal standard. Mineral element content of studied samples, determined by reference method, varied between 72 and 421 mg g⁻¹ for calcium, 998 and 10,964 µg g⁻¹ for magnesium and, 102 and 2100 µg g⁻¹ for strontium.

2.4. NIR procedure

Pulverized mineralized samples were placed in clear glass vials of 11 mm internal diameter and 25 mm height to directly obtain their NIR spectra by diffuse reflectance in Kubelka–Munk units. Spectra were collected between 14,000 and 4000 cm⁻¹ by averaging 50 scans and using an optical resolution of 4 cm⁻¹. A background spectrum was acquired before each series, from the closed integrating sphere using the same instrumental conditions than those employed for samples measurement. Three measurements of each sample were obtained by rotating the sample vial position between replicates in order to ensure a good reliability. The average of the triplicate spectra of each sample was employed for data exploration and to build the chemometric models.

2.5. Chemometric data treatment

Data treatment was carried out using in-house written functions employing Matlab 8.3.0.532 (R2014a) from Mathworks (Natick, MA, USA) being employed for Principal Analysis Components (PCA) and Partial least squares (PLS) regression model, the PLS Toolbox 7.5.2 from Eigenvector Research Inc. (Wenatchee, WA, USA).

PCA applied to IR spectra was used as data exploration; based on the distance between samples the evaluation of their similarity. PLS models were applied to spectral data to develop prediction models for Ca, Mg and Sr. These calibration models were developed based on the

statistically inspired modification of the PLS method (SIMPLS) algorithm [32]. To select the most appropriate sample calibration set, Kennard–Stone (KS) algorithm [33] was used, thus selecting a representative subset to ensure training samples spread evenly throughout the sample space. For building the best PLS models, different spectral regions and spectra pre-treatments were tested as multiplicative scatter correction (MSC), standard normal variation (SNV), a Savitzky-Golay first (FD) and second derivative (SD), and mean center (MC) also combination of them. The performance of PLS-NIR models was evaluated according to the root mean square error of cross validation and prediction (RMSECV and RMSEP) values, the coefficient of determination for prediction (R^2_{pred}), relative root mean square error of prediction (RRMSEP) and residual predictive deviation (RPD), calculated this last parameter as the ratio between standard deviation (SD) of the prediction set and the RMSEP values [34]. Generally, RPD value greater than 3.0 is considered adequate for analytical purposes with excellent prediction accuracy, between 2.5 and 3.0 implies that the model has a good precision and between 2.0 and 2.5 indicates that the model has an approximate precision. A value for the RPD between 1.5 and 2.0 reveals a possibility to distinguish between high and low values and a RPD 1.5 indicates that the calibration is not usable [35]. The optimum number of latent variables (LV's) was determined by cross-validation using leave one out sample in order to obtain the minimum value of RMSECV.

3. Results and discussion

3.1. FT-NIR spectra

Fig. 1 shows, in Kubelka Munk units, the averaged NIR spectra of a) internal part, b) external part of bones and c) sediment samples, considered in this study without any data pre-treatment, in the region between 9000 and 4000 cm^{-1} . These sample spectra present differences in the intensity and position of certain bands, but all show absorption bands around 7200 and 5220 cm^{-1} related to combination vibrations of H–O–H bend and O–H stretch of water. The main difference which presents internal bones is the weak vibration which appear near 4655 cm^{-1} (without a clear assignment) and the absorption near 6977 cm^{-1} assigned to the first overtone of the stretching vibrations of OH group in hydroxyapatite [36], that is higher than in external part of bones. In the case of the sediment samples, the aforementioned band is absent. Additionally, at 5278 cm^{-1} , internal part of bone and external part of bone in minor intensity, present an absorption band related to apatite [30]. On the other hand, in sediment samples it is present a band at 7070 cm^{-1} related to the first overtone of the O–H stretch vibration in metal–O–H [37]. Moreover, vibrations bands near of 4530 cm^{-1} are due to clay minerals, like smectite and illite, and the band at 4265 cm^{-1} is associated with the contribution of calcite [37], which decreases in intensity in external bone. In the case of internal part of bone samples, the band located at 4265 cm^{-1} not appears and the band at 4530 cm^{-1} is less intense than in external part of bones. So that, in external part of bones appear typical bands of sediment samples related with metal-OH (7070 cm^{-1}), calcite (4265 cm^{-1}) and clay minerals (4530 cm^{-1}), and lower bands of hydroxyapatite (6977 cm^{-1}) and apatite (5278 cm^{-1}). Degradation of hydroxyapatite in external part of bones is a great indicator of diagenetic process. Sediments around bones are crucial role to conservation of them, and sediments analysis should be carried out in order to assess their contribution to the modification of the bone chemical composition.

3.2. Data exploration by PCA

Before building the calibration models, PCA was used for exploratory data analysis to study the spectral differences between considered samples. Fig. 2 presents the scores plot for first and the second

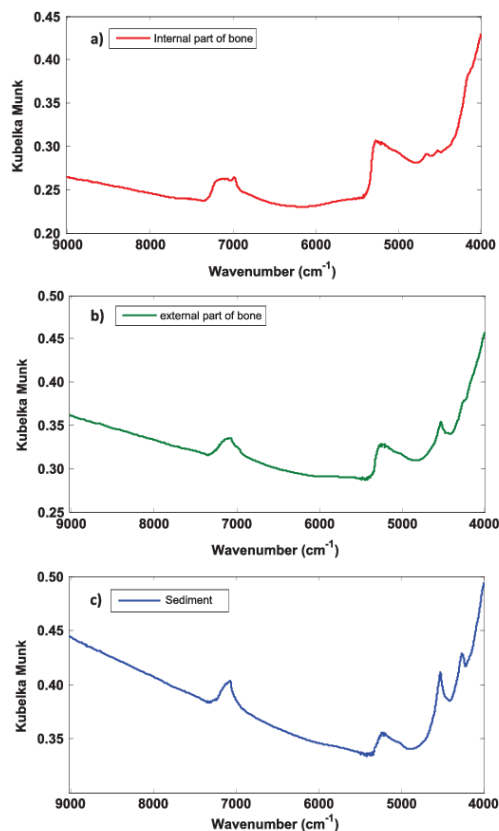


Fig. 1. Averaged NIR spectra of mineralized bone and sediment samples comparing a) internal part, b) external part, and c) sediment samples. Note: Spectra were in Kubelka Munk units in the region between 9000 and 4000 cm^{-1} .

principal components obtained from PCA for Misericordia and En Gil Necropolis, after a) FD and MC and b) MSC, FD and MC pre-treatment, being selected the region from 9014 to 4000 cm^{-1} . In Fig. 2a, the two first principal components after spectra pretreatment by using FD and MC represent 86.07% of the explained variance, being 76.93% and 9.14% for PC1 and PC2, respectively. In the scores plot, there is not a clear-cut grouping of samples, but it can be appreciated that sediment samples spectra are clearly different from those of internal part of bone samples and that, external part of bone samples are located between sediment and internal part of bone ones in the direction of PC1. In Fig. 2b, it can be seen that the scores plot for first and the second principal components after MSC, FD and MC pretreatment, represent 85.72% of the explained variance (PC1 corresponds to 80.49% and PC2 corresponds to 5.23% of the variance). It can be seen that internal part of bone samples are located towards the right of the PC1 and sediment samples are dispersed in the left. External part of bone samples are situated between these two groups in the direction of PC1. Moreover, PCA indicates that samples derived from En Gil are similar to Misericordia samples, and because of that samples from Misericordia can be used as calibration set to predict alkaline earth elements in En Gil samples.

The aforementioned PCAs were built using differentiated internal

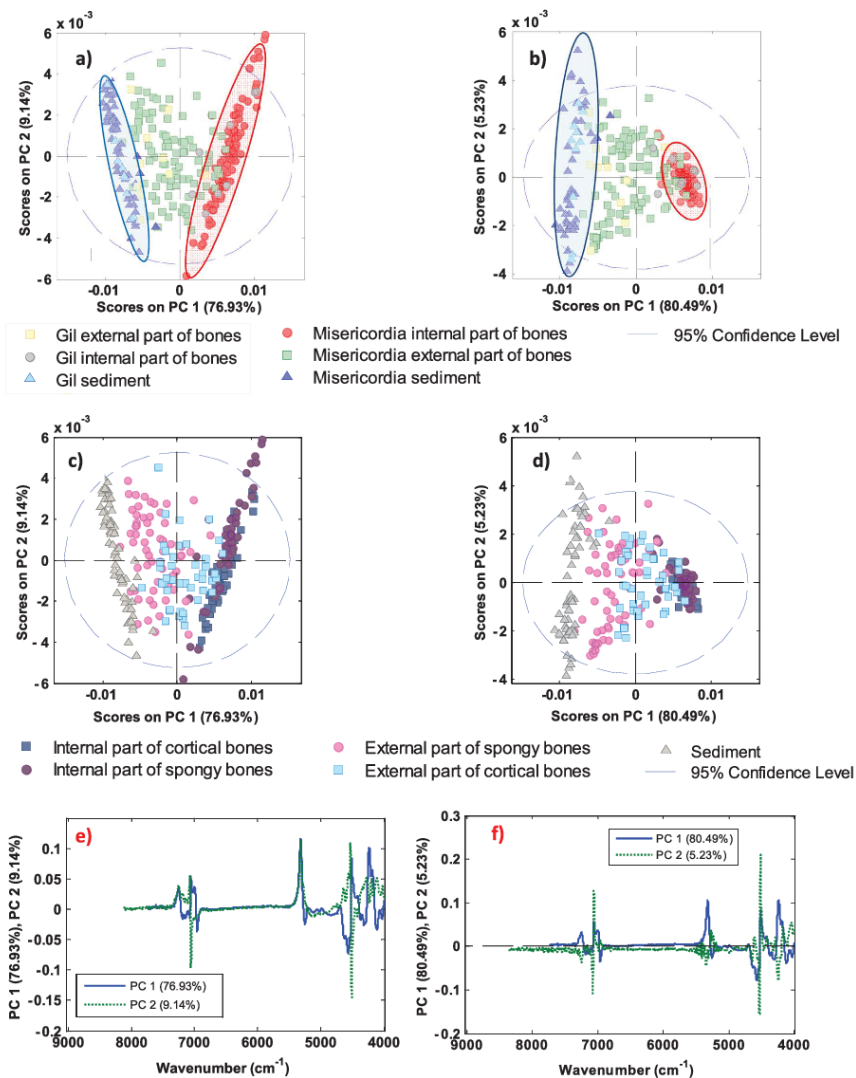


Fig. 2. PLS scores and loadings graphs obtained after different spectra treatments. Score plots according Misericordia and En Gil Necropolis sample classification after: a) FD and MC, and b) MSC, FD and MC pretreatment. Score plots considering the bone parts after: c) FD and MC, and d) MSC, FD and MC pretreatment. Loadings of PC1 and PC2 after: e) FD and MC, f) MSC, FD and MC pretreatment.

and external parts of bones in both, spongy and cortical remains. Fig. 2c presents the scores plot after FD and MC and Fig. 2d after MSC, FD and MC pretreatment. No differences were detected between internal parts of cortical and spongy bones. However if we look at the external part of bones, spongy group is located near to sediment samples, and external part of cortical bones are located close to internal part of bones because spongy tissues are more susceptible to diagenesis than cortical [38] due to the fragility and porosity of their structure. Wavenumbers responsible for the distinction of the samples in PCA can be identified in the corresponding loading plots (Fig. 2e and f) with a

clear identification of bands around 7000 cm⁻¹ and between 5500 and 4000 cm⁻¹.

So, it can be concluded that PCA analysis using NIR spectra data provides a fast and green tool to identify changes caused by the environment in bone samples. These results indicate that external part of bones suffer an increased degradation produced by diagenetic factors than internal part of bones, especially in the case of spongy bones external part due to their high porosity. Consequently, conclusions about identification of biogenetic signals using major element ratios could change depending of the bone class evaluated. To avoid

Table 1

Concentration ranges of determined mineral elements in calibration, validation and prediction data sets.

Analyte	Set	Samples	Range	Mean value	SD
Calcium (mg g^{-1})	Calibration	181	96–421	260	100
Magnesium ($\mu\text{g g}^{-1}$)		181	1050–10,500	4967	3091
Strontium ($\mu\text{g g}^{-1}$)		181	153–2100	823	486
Calcium (mg g^{-1})	Validation	58	72–400	312	76
Magnesium ($\mu\text{g g}^{-1}$)		59	1070–8500	3226	2545
Strontium ($\mu\text{g g}^{-1}$)	58	124–2010	1058	405	
Calcium (mg g^{-1})	Prediction	26	112–421	233	114
Magnesium ($\mu\text{g g}^{-1}$)		26	998–10,964	6723	3864
Strontium ($\mu\text{g g}^{-1}$)		26	102–1350	518	442

SD: Standard deviation.

mistakes, the use of PCA analysis, as exploratory method, permits a prior selection of samples, not affected by diagenetic processes, to be used in different bioarchaeological and forensic studies involved in bone analyses. Additionally, PCA is a good indicator to test if the calibration set contains the validation samples in order to be used in PLS models without errors due to extrapolation of data.

3.3. PLS-NIR models

To build PLS models, Misericordia samples were divided into a calibration and a validation subset with 183 and 60 objects, respectively, using K-S algorithm for the training set selection, composed by 62 external part of bones, 70 internal part of bones and 51 sediment samples.

The main characteristics of calibration and validation sets selected for the three elements determination by PLS-NIR are indicated in Table 1. Several regions and different pre-processing strategies were assayed to build the best models and to evaluate their prediction capability. For all elements, regions selected were those between 9014 and 4000 cm^{-1} . To build PLS-NIR models using the selected spectral range, 3, 4 and 4 LVs were employed for Ca, Mg and Sr, respectively, in order to minimize the RMSECV, explaining 89.1%, 90.8% and 88.5% of the total variance of the X data block and 92.4%, 95.0% and 90.4% of the Y data block, respectively. For Ca and Mg, FD with a window of 11 points and a second order polynomial, and MC treatment were adopted as signal pre-processing. For strontium, MSC, FD and MC were chosen. Additionally, concentration data were mean centered as well in all cases. Outliers can be identified by Q residual versus the Hotelling T^2 values, being removed 2 samples from the calibration set of Ca, Mg and Sr and 2 samples from the validation set of Ca and Sr and only one sample of Mg, prior to do the calculations and validation of the final PLS models for those elements. Fig. 3 shows the regression between PLS-NIR predicted values for Ca, Mg and Sr and those obtained by the reference ICP-OES method. It can be seen that calibration and validation sample points were closely distributed near the optimum regression line between predicted and measured values for the three elements, obtaining high coefficients of determination for calibration (R^2_{cal}), cross validation (R^2_{CV}) and prediction (R^2_{pred}) in all elements. The prediction capability of PLS-NIR models were good for all elements, obtaining RRMSEP values of 10%, 15% and 19% for calcium, magnesium and strontium, respectively. The most important calibration and validation parameters of the developed PLS-NIR models are summarized in Table 2.

It can be seen that samples of internal part of bones contain higher calcium ($368 \pm 24\text{ mg g}^{-1}$) and strontium ($1306 \pm 235\text{ }\mu\text{g g}^{-1}$) levels and lower magnesium content ($1393 \pm 175\text{ }\mu\text{g g}^{-1}$) than external parts (with $261 \pm 47\text{ mg g}^{-1}$ of Ca, $844 \pm 203\text{ }\mu\text{g g}^{-1}$ of Sr and $5599 \pm 1684\text{ }\mu\text{g g}^{-1}$ of Mg) and sediment samples (with $133 \pm 17\text{ mg g}^{-1}$ of Ca, $239 \pm 60\text{ }\mu\text{g g}^{-1}$ of Sr and $8165 \pm 858\text{ }\mu\text{g g}^{-1}$ of Mg). As aforementioned

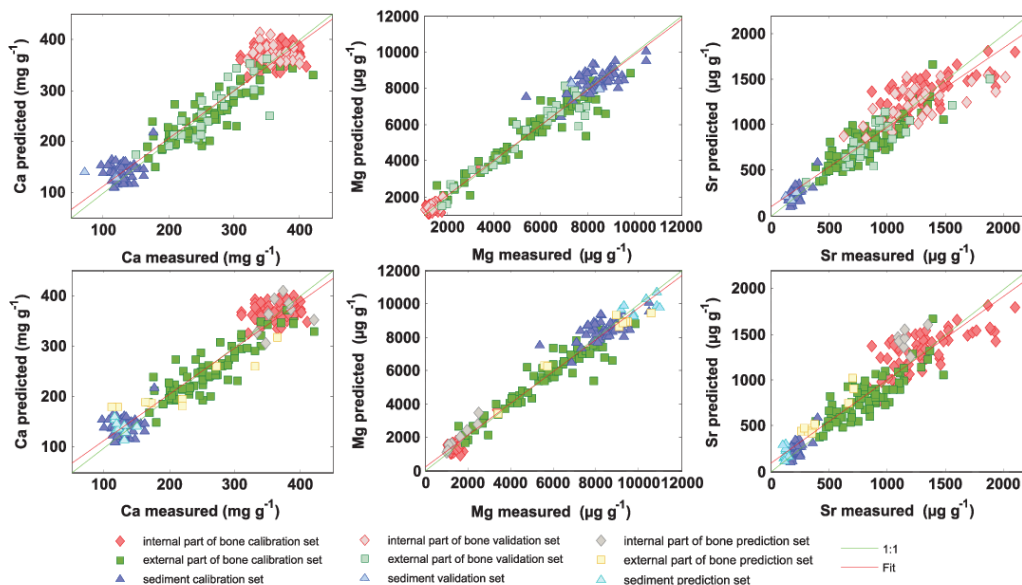


Fig. 3. Regression between predicted and reference values obtained for the PLS-NIR determination of calcium, magnesium and strontium in bone remains and sediments.

Table 2
Description of the best PLS-NIR models employed for the determination of Ca, Mg and Sr and their main figures of merit.

Element	Pre-process	Misericordia samples										en Gil samples									
		LV	RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	R ² Pred	RRMSEP (%)	RPD	RMSEP	R ² Pred	Slope	Intercept	Bias	RRMSEP (%)	RPD				
Calcium (mg g ⁻¹)	FD, MC	3	27.5	92.4	29.0	91.6	31.7	83.4	10.2	2.4	35.6	90.4	41.4	1.7	15.3	3.2					
Magnesium (µg g ⁻¹)	FD, MC	4	686.3	95.0	750.2	94.1	496.6	15.4	15.4	723.7	97.3	246.2	-81.3	10.8	5.3						
Strontium (µg g ⁻¹)	MSC, FD, MC	4	150.1	90.4	169.8	87.7	198.9	76.2	18.8	198.9	97.4	78.9	0.9041	155.0	2.3						

LV: Number of latent variables; RMSEC: Root mean square error of cross validation; RMSECV: Root mean square error of cross validation; RMSEP: Relative root mean square error of prediction; R² Cal: coefficient of determination of calibration; R² CV: coefficient of determination of cross validation; R² Pred: coefficient of determination of prediction.

tioned, in physiologic conditions, hydroxylapatite is composed of 38% Ca, with traces of Mg and small amounts of other elements like Sr, indicating that internal part of bones are intact looking at Ca values that are around the 38%. Mg enrichment in external part of bones indicates that this element is probably incorporated into the bones from sediments during diagenesis process.

External part of bones, especially in the case of spongy bone remains are more sensitive to diagenetic processes caused by the environment indicating that degradation not affects in the same extent different skeletal parts. Due to the fact that spongy bones have higher porosity and thinner cortex than cortical bones, the first type is less resistant to diagenetic factors suffering chemical changes.

Results obtained by PLS-NIR models indicated that NIR spectroscopy has a good potential to predict alkaline earth elements content in bone remains. As mentioned this method has many advantages over the conventional employed analytical techniques because it is quick, inexpensive, non-destructive for pulverized samples and does not require the use of chemical reagents nor solvents.

3.4. PLS-NIR prediction capability

To evaluate the PLS-NIR prediction capability of models built to predict Ca, Mg and Sr concentrations in bones and sediments, an independent validation set, not employed during the calibration step, was used. For this propose external part, internal part of bones and sediment samples of En Gil Necropolis were used as external set. The mean, concentration range and standard deviation for the concentration of analyzed elements employed for prediction set are summarized in Table 1. Predicted values of the analytes versus those obtained by the reference methods are shown in Fig. 3 and the most important calibration and prediction parameters of the developed NIR-PLS models are summarized in Table 2. Acceptable RRMSEP values were obtained for calcium and magnesium with 15% and 11%, respectively, and for strontium a value of 37% was obtained, with correlation coefficients of 90.4, 97.4 and 97.4, and RPD values of 3.2, 5.33 and 2.3, respectively. These results demonstrate a good predictive capability of the PLS-NIR models developed to evaluate alkaline earth elements in bones and sediment samples.

Furthermore it must be indicated the PLS models were made from bone samples of different origin, and date of those of the training set, confirming the feasibility of our innovative methodological proposal for the evaluation of alkaline earth elements in archaeological remains.

The obtained results indicate a good predictive capability for Mg in samples that not belong to the same population while Ca were predicted with relative errors of 15% and Sr 37%, respectively, being thus limited as a screening tool. Especially for Sr the high relative error of 37% is probably due to the lower levels of this element found in some samples not compatible with the sensitivity of NIR.

4. Conclusions

Fourier Transform Near Infrared (FT-NIR) provides a fast, cheap and green analytical method for the prediction of the content of calcium, magnesium and strontium in buried bone and sediment samples, and it could be very useful to understand post-mortem changes of bones caused by the environment which can affects different skeletal parts and for selecting bone samples with well preserved biogenetic signals. PCA has shown that the alkaline earth element profile of bone and soil samples, as expected, is clearly different. Furthermore the external surface of spongy bones is more similar to the soils than to the surface of cortical bones. This confirm that the outer bone layer of spongy bones is more altered than the cortical surface and the presence of Ca, Mg and Sr is similar to that of soil samples of the studied sites. Therefore it can be concluded that environmental factors have a main impact on spongy tissues, which are more susceptible to diagenetic processes, than on the cortical

bones. It is indicated in the NIR spectra by the loss of apatite band (the main bone compound) and the presence of clay bands especially on spongy bone surface.

A milestone in the study of bone remains is settled in this work by the development of PLS-NIR models, indicating that FT-NIR spectroscopy can be employed to predict Ca, Mg and Sr contents in bones remains and soils samples. Ca and Sr contents were found to be higher in the internal part of bone samples, opposite to Mg contents that were higher in the external bone surface and soils, and lower in internal part of bones. Our developed methodological approach combines NIR spectroscopy and statistical analysis and this allowed us to predict bone alkaline earth mineral composition opening a new perspective for the identification of better preserved samples in bioarchaeological studies and forensic science investigations.

Regarding the limitations of the study it must be noticed that it was based only on the mineral part of skeletal remains which were treated thermally and pulverized before measurements thus losing information from the organic part of bones and soils. Additionally, the lack of sensitivity of NIR spectroscopy seriously affected the capability of these measurements to be employed for trace element determination and just important elements present as major components of bones, like Ca, Mg and Sr could be determined.

Acknowledgements

Authors acknowledge the financial support of Generalitat Valenciana (Project PROMETEO II/2014/077) and Ministerio de Economía y Competitividad-Feder (Projects CTQ 2014–52841-P and CTQ 2012–38635). M.C acknowledges the FPI Grant (BES-2012–055404) provided by the Ministerio de Economía y Competitividad of the Spanish government.

The authors would like to thanks all the students of Chemistry and Archaeology which have contributed to the realization of this study.

References

- B. Clarke, Normal bone anatomy and physiology, *Clin. J. Am. Soc. Nephrol.* 3 (Suppl 3) (2008) S131–S139.
- J. Burton, Bone chemistry and trace element analysis, in: M.A. Katzenberg, S.R. Saunders (Eds.), *Biological Anthropology of the Human Skeleton* 2nd ed., John Wiley & Sons, Inc, Hoboken, NJ, 2008, pp. 443–460.
- B. Wopenka, J.D. Pasteris, A mineralogical perspective on the apatite in bone, *Mater. Sci. Eng. C* 25 (2005) 131–143.
- R.E.M. Hedges, A.R. Millard, Bones and groundwater: towards the modelling of diagenetic processes, *J. Archaeol. Sci.* 22 (1995) 155–164.
- R.E.M. Hedges, Bone diagenesis: an overview of processes, *Archaeometry* 44 (2002) 319–328.
- W. Querido, A.L. Rossi, M. Farina, The effects of strontium on bone mineral: a review on current knowledge and microanalytical approaches, *Micron* 80 (2016) 122–134.
- J.B. Lambert, L. Xue, J.E. Buikstra, Inorganic analysis of excavated human bone after surface removal, *J. Archaeol. Sci.* 18 (1991) 363–383.
- F. Donald Pate, J.T. Hutton, The use of soil chemistry data to address post-mortem diagenesis in bone mineral, *J. Archaeol. Sci.* 15 (1988) 729–739.
- J. Zapata, C. Pérez-Sirvent, M. Martínez-Sánchez, P. Tovar, Diagenesis, not biogenesis: two late Roman skeletal examples, *Sci. Total Environ.* 369 (2006) 357–368.
- I. Reiche, L. Favre-Quattrapani, T.J.S. Calligaro, H. Bocherens, L. Charlet, M. Menu, Trace element composition of Archaeological Bones and postmortem alteration in the burial environment, *Nucl. Instrum. Methods Phys. Res. B* 150 (1999) 656–662.
- I. Reiche, L. Favre-Quattrapani, C. Vignaud, H. Bocherens, L. Charlet, M. Menu, A multi-analytical study of bone diagenesis: the Neolithic site of Bercy (Paris, France), *Meas. Sci. Technol.* 14 (2003) 1608–1619.
- T.H. Schmidt-Schultz, M. Schultz, Intact protein molecules in archaeological bones – Bone matrix as a treasure chest of ancient diseases and living conditions. *American Journal of Physical Anthropology. Annual Meeting Issue 1999*, Wiley-Liss, p. 230
- T.A. Surovell, M.C. Stiner, Standardizing infra-red measures of bone mineral crystallinity: an experimental approach, *J. Archaeol. Sci.* 28 (2001) 633–642.
- F. Ignac, G. Gopinath, V. der W. Hans, Radionuclide and Hybrid Bone Imaging, Springer, Heidelberg, New York, Dordrecht, London, 2012.
- M.L. Carvalho, A.F. Marques, Diagenesis evaluation in middle ages human bones using EDXRF, *X-Ray Spectrom.* 37 (2008) 32–36.
- D. Guimarães, A.A. Dias, M. Carvalho, M.L. Carvalho, J.P. Santos, F.R. Henriques, F. Curate, S. Pessanha, Quantitative determinations and imaging in different structures of buried human bones from the XVIII-XIXth centuries by energy dispersive X-ray fluorescence – postmortem evaluation, *Talanta* 155 (2016) 107–115.
- G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, J. Bernabeu, Biological mineral content in Iberian skeletal cremains for control of diagenetic factors employing multivariate statistics, *J. Archaeol. Sci.* 40 (2013) 2477–2484.
- A.-F. Maurer, M. Gerard, A. Person, I. Barrientos, C.R. del V. Darras, C. Durllet, V. Zeitoun, M. Renard, B. Faugère, Intra-skeletal variability in trace elemental content of Precolombian Chupicuaro human bones: the record of post-mortem alteration and a tool for palaeodietary reconstruction, *J. Archaeol. Sci.* 38 (2011) 1784–1797.
- G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, J. Bernabeu, Chemical element levels as a methodological tool in forensic science, *J. Forensic Res.* 6 (2014) 1000264.
- A.-F. Maurer, A. Person, T. Tüfken, S. Amblard-Pison, L. Ségalen, Bone diagenesis in arid environments: an intra-skeletal approach, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 416 (2014) 17–29.
- S. Dal, L. Maritan, D. Usai, I. Angelini, G. Artigoli, Bone diagenesis at the micro-scale: Bone alteration patterns during multiple burial phases at Al Khiday (Khartoum, Sudan) between the Early Holocene and the II century AD, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 416 (2014) 30–42.
- S.W. Keenan, A.S. Engel, A. Roy, G. Lisa Bovenkamp-Langlois, Evaluating the consequences of diagenesis and fossilization on bioapatite lattice structure and composition, *Chem. Geol.* 413 (2015) 18–27.
- C.N. Trueman, K. Privat, J. Field, Why do crystallinity values fail to predict the extent of diagenetic alteration of bone mineral?, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 266 (2008) 160–167.
- T.J.U. Thompson, M. Islam, K. Piduru, A. Marcel, An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 299 (2011) 168–174.
- T.J.U. Thompson, M. Islam, M. Bonniere, A new statistical approach for determining the crystallinity of heat-altered bone mineral from FTIR spectra, *J. Archaeol. Sci.* 40 (2013) 416–422.
- G. Iliopoulos, N. Galanidou, S.A. Pergantis, V. Vamvakaki, N. Chaniotakis, Identifying the geochemical taphonomy of the osteological material from Katsambas rockshelter, *J. Archaeol. Sci.* 37 (2010) 116–123.
- L.E. Wright, H.P. Schwarz, Infrared and isotopic evidence for diagenesis of bone apatite at Dos Pilas, Guatemala: Palaeodietary implication, *J. Archaeol. Sci.* 23 (1996) 933–944.
- D. Alfano, A.R. Albuñia, O. Motta, A. Proto, Detection of diagenetic alterations by Spectroscopic analysis on Archaeological bones from the Necropolis of Poseidonía (Paestum): a case study, *J. Cult. Herit.* 10 (2009) 509–513.
- J. Linderholm, P. Geladi, Classification of archaeological soil and sediment samples using near infrared techniques, *NIR News* 23 (2012) 6.
- D. Thomas, C. McGovern, A. Chinsamy, M. Manley, Near infrared analysis of fossil bone from the Western Cape of South Africa, *J. Infrared Spectrosc.* 19 (2011) 151.
- G. Gallelo, S. Silvia, J. Kuligowski, B. Pulvio, M. Francesco, P. Agustín, Variación química intraesquelética relacionada con la diagenesis en los restos óseos de c/ en Gil (Valencia), *SAGVNTVM* 47 (2015) 175–186.
- S. de Jong, SIMPLS: an alternative approach to partial least squares regression, *Chemom. Intell. Lab. Syst.* (3) (1993) 251–263.
- R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Techonometrics* 11 (1969) 137–148.
- P.C. Williams, D. Sobering, How do we do it: a brief summary of the methods we use in developing near infrared calibrations, in: A.M.C. Daves, P.C. Williams (Eds.), *Near Infrared Spectroscopy: The Future Waves*, NIR Publications, Chichester, UK, 1995, pp. 185–188.
- W. Saeys, A.M. Mouazen, H. Ramon, Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy, *Biosyst. Eng.* 91 (2005) 393–402.
- Y. Ning, J. Li, W. Cai, X. Shao, Simultaneous determination of heavy metal ions in water using near-infrared spectroscopy with precontamination by nano-hydroxyapatite, *Spectrochim. Acta A. Mol. Biomol. Spectrosc.* 96 (2012) 289–294.
- V. Aranda, A. Dominguez-Vidal, F. Comino, J. Calero, M.J. Ayora-Cañada, Agro-environmental characterization of semi-arid Mediterranean soils using NIR reflection and mid-IR-attenuated total reflection spectroscopies, *Vib. Spectrosc.* 74 (2014) 88–97.
- J.B. Lambert, S.M. Vlasak, A.C. Thometz, J.E. Buikstra, A comparative study of the chemical analysis of ribs and femurs in Woodland population, *Am. J. Phys. Anthropol.* 59 (1982) 289–294.

Artículo 6. Burned bones forensic investigations employing near infrared spectroscopy

Vibrational Spectroscopy 90(2017) 21-30

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Vibrational Spectroscopy

journal homepage: www.elsevier.com/locate/vibspec

Burned bones forensic investigations employing near infrared spectroscopy



Mari Merce Cascant^a, Sonia Rubio^a, Gianni Gallelo^{a,b,*}, Agustín Pastor^a, Salvador Garrigues^a, Miguel de la Guardia^a

^a Department of Analytical Chemistry, University of Valencia, 50 Dr. Moliner Street, Research Building, 46100 Burjassot, Valencia, Spain

^b Department of Archaeology, University of York, King's Manor, York YO1 7EP, UK

ARTICLE INFO

Article history:

Received 21 January 2017

Received in revised form 21 February 2017

Accepted 22 February 2017

Available online 7 March 2017

Keywords:

Burned bones

FT-NIR

Chemical elements

Statistics

ABSTRACT

The use of near infrared (NIR) spectroscopy was evaluated, by using chemometric tools, for the study of the environmental impact on burned bones. Spectra of internal and external parts of burned bones, together with sediment samples, were treated by Principal Component Analysis and cluster classification as exploratory techniques to select burned bone samples, less affected by environmental processes, to properly carry out forensic studies. Partial Least Square Discriminant Analysis was used to build a model to classify bone samples based on their burning conditions, providing an efficient and accurate method to discern calcined and carbonized bone. Additionally, Partial Least Square regression models were built to predict calcium, magnesium and strontium concentration of bone samples from their NIR spectra, being obtained an accurate root mean square error of prediction of 5.2% for calcium. Furthermore a screen methodology, for magnesium and strontium prediction, with a RPD of 0.24 and 1.08 respectively, was developed.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The analysis of burned human remains is of great interest among forensics and anthropologists due to the problems associated to their recovery, identification and classification [1]. Burned bone samples provide knowledge of changes in bone in order to use this information for anthropological and forensic studies [2]. Consequently, during the last decades, the number of published papers on burned bone analysis has increased [3–5].

Burned bones could be classified as carbonized and calcined depending on the thermal exposure, exhibiting a black appearance (carbonized) due to skeletal material and soft tissue or a white appearance (calcined). Burned bones are incinerated when the thermal alteration is so prolonged and intense that all the organic materials and moisture were lost [6,7]. Estimation of the maximum exposure temperature is a crucial factor for the correct interpretation of burned bones, being the change of colour an employed method as temperature range indicator [6,8]. Colour has been

described qualitatively by visual comparison with standard charts, such as the Munsell Soil Color Charts [9]. However different perceptions or changes in lighting conditions can significantly modify the results [6]. Another method used to colour determination was suggested by Devlin et al. [10] using CIE L*a*b* (CIELAB) uniform colour space for the recording of bone surface colour data. However, mistakes on the methodologies employed for burned bone classification and analysis have recently been stressed. Some authors indicate that colour of bone fragments is a combination of temperature, duration of elevated temperature treatments, oxygen presence during combustion and context of incineration. All those factors should be considered [11,12]. Nicholson [13] found the colour changes greatly varied between bones, attributing these observations to different organic content and bone chemistry. Symes et al. [14] indicate that different colour alterations can be found within a single skeleton and even on a single bone, especially in cases of burning of fleshed remains. Also, other aspects, as external and diagenetic factors, could modify the bone colour producing similar changes than those caused by high temperatures [15–17]. Overall, study of burned remains is a difficult interpretative challenge for forensics and anthropologist due to different factors and aspects to be considered.

* Corresponding author at: Department of Archaeology, University of York, King's Manor, York YO1 7EP, UK.

E-mail address: gianni.gallelo@uv.es (G. Gallelo).

<http://dx.doi.org/10.1016/j.vibspec.2017.02.005>

0924-2031/© 2017 Elsevier B.V. All rights reserved.

As the temperature increases, the hydroxyapatite present in bones has an increased crystallinity and forms stable and large crystals [18], starting the most important transformation at 600 °C [11]. Lanting et al. [19] perceived the potential of calcined bones for radiocarbon dating, indicating that all previous attempts to date burned bones failed because they were treated as carbonized bones. Exchange processes of bicarbonate ions dissolved in soil waters produce a bio-apatite contamination which results in too young ^{14}C dates, not occurring that on calcined bones [20]. During the last decade an increment of calcined bones radiocarbon dating studies has been produced [21–24].

Several studies were carried out to understand the bone mineral elements behaviour under thermal conditions and the influence of post-depositional processes in skeletal remains [25–28]. Some works have shown that post-mortem processes equally affect cremated and non-cremated ancient bones [27,28]. Elemental composition studies made on incinerated bones were carried out by Grupe et al. [25], concluding that paleodietary reconstruction is limited due to volatilization and crystal modifications caused by high temperatures. However, for Sr/Ca, a ratio which is the most used method for palaeodietary reconstruction, it is possible to use regression analysis.

Indeed, the presence of hydroxyl groups in bone apatite continues to be a topic of controversy. Some studies believe that a number of hydroxyl groups are present in archaeological bones [29,30] while others argue that bone apatite does not contain detectable concentrations of hydroxyl groups [31,32]. Hydroxylation of bone apatite is attributed to the presence of CO_3^{2-} , according to two substitution mechanisms in the OH^- site, through the so called “A-type” substitution, or in the PO_4^- site, based on “B-type” substitution [31]. Other exchanges, such as fluoride (F^-) substitution for hydroxide (OH^-), and strontium (Sr^{2+}) and other trace cations (Zn^{2+} , Pb^{2+}) substitution for calcium (Ca^{2+}), are also produced [33].

Additional studies are needed to get deep information on the mineral composition of remains, crystallinity of samples and the influence of soil component exchanges. Infrared spectroscopy (IR) has been used to assay the modifications of bone apatite during calcinations, mostly on experimentally heated bone [34–36]. However, the relative impacts of temperature and exposure time remain unclear and IR could be also useful to evaluate the presence of major mineral elements in bone remains.

The aim of the present study has been to select the most suitable samples to carry out forensic studies and, to do it, the evaluation of near infrared spectroscopy has been employed being used chemometric tools as Principal Component Analysis (PCA), Cluster Analysis (CA) and Partial Least Squares-Discriminant Analysis (PLS-DA). Carbonized bones are more prone than calcined ones to degradation by exchange processes with soil, for that a classification of bones allows ensuring analysis results. Internal part, external part of burned bones and sediment samples were analysed for a better understanding of post-mortem bone modification processes. Furthermore, it has been evaluated the simultaneous determination of calcium, magnesium and strontium in burned bones by PLS-NIR in order to provide environmentally friendly screening tools to evaluate the presence of mineral elements in bones.

2. Material and methods

2.1. Samples

Thirty-eight burned bone samples, from Corral de Saus Necropolis were studied [28]. This site was dated between the III and II centuries B.C., a period in which cremations were placed in urns where each one contained the remains of a single individual.

Bones were divided into “carbonized”, bones fired in reducing atmosphere, “calcined”, those fired in oxidizing atmosphere and “unknown”, for bones of unidentified burning conditions. Additionally, unburned animal bones and sediments were obtained from the inside of urns, mixed with ashes, carbons and human bones. To carry out the study fifteen internal part of bones (seven carbonized and eight calcined), seven external part of bones (being obtained from the first 2 mm of bone directly in contact with the sediments), seven unknown bones, four unburned animal bones and six sediment samples were used. Additionally, 31 unknown samples from Las Peñas Necropolis [37] were analysed and the certified material Bone ash NIST 1400 and soil GBW07408 were used for evaluating the analytical method.

2.2. Apparatus and methods

A Fourier transform near infrared (FT-NIR) spectrometer, model Multipurpose Analyzer (MPA) from Bruker (Bremen, Germany) equipped with an integrating sphere, was employed for diffuse reflectance near infrared spectra acquisition. For instrumental and measurement control of the spectrometer, as well as for data acquisition, Opus 6.5 software from Bruker was used.

Optima 5300 DV Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) Perkin Elmer (Norwalk, CT, USA) equipped with an autosampler AS 93-plus and a cross flow nebuliser was used to obtain reference values on the mineral composition of samples. Samples were previously calcined in a muffle furnace Biometa Lenton ECF 12145A (Lanera, España) and digested with acids using a heating plate Ika C-Mag HS7 (Staufen, Germany).

2.3. Reference method

Samples were analysed after acid digestion using ICP-OES. The digestion method consisted of the addition of 1.5 ml HCl and 1.5 ml HNO_3 to 0.5 g of sample in glass tubes, for both, bones and soil. Digestion of samples was carried out in a water bath at 100 °C for 40 min. The digested solutions were quantitatively transferred to plastic tubes and diluted to a final volume of 15 ml with distilled water. Concentration ranges of the dilutions from the digested solution were adapted to the sensitivity of the ICP-OES measurements for each element. From a multi-elemental stock solution containing Ca, Mg and Sr of 100 $\mu\text{g ml}^{-1}$ calibration standards were prepared by the appropriate dilution, also adding HNO_3 and HCl at the same level of samples. Standards were obtained from Sharlab S. L. (Barcelona, Spain). Reference materials bone ash NIST 1400 and soil NIM GBW07408 were used to evaluate the accuracy of the analytical method. Re was used as internal standard in ICP-OES [28]. Mineral content of samples, determined by reference method, were between the range of 250 and 461 mg g^{-1} for calcium, 790 and 34257 $\mu\text{g g}^{-1}$ for magnesium and, 161 and 1068 $\mu\text{g g}^{-1}$ for strontium.

2.4. NIR procedure

NIR spectra were directly obtained by diffuse reflectance in Kubelka–Munk units from pulverized samples placed inside glass vials of 11 mm internal diameter and 25 mm height. Sample spectra were obtained from 14000 to 4000 cm^{-1} by averaging 50 scans per spectrum using a nominal resolution of 4 cm^{-1} , as instrumental conditions. The background spectrum was acquired from the integrating sphere under the same instrumental conditions than those employed for sample measurements. Three measurements of each sample were obtained by rotating the sample vial position between replicates in order to ensure a better reliability. The means of the triplicate spectra of each sample were employed for chemometric treatment.

2.5. NIR quality control (QC) NIR quality control (QC)

Bone samples were measured in triplicate. Certified reference material Bone ash NIST 1400 was measured under the same conditions than samples in three different times during each measurement session during 18 days: at the beginning of the session, in the middle and at the end of measurements. In order to evaluate the proposed method, inter-day precision and accuracy of measurements obtained for Bone ash NIST 1400 were used. Precision was expressed as the coefficient of variation of the results obtained during 18 different days and accuracy was expressed as the bias between the predicted and the certified value divided by the certified value, expressed as percentage.

2.6. Chemometric data treatment

Data treatment was done with Matlab 8.3.0.532 (R2014a) from Mathworks (Natick, MA, USA) using PLS Toolbox 7.5.2 from Eigenvektor Research Inc. (Wenatchee, WA, USA) for Principal Component Analysis (PCA), cluster classification, Partial Least Square Discriminant Analysis (PLS-DA) and Partial Least Squares (PLS) regression in order to do a correct identification of samples and for the quantitative prediction of Ca, Mg and Sr concentrations in bones from sample NIR spectra.

PCA allows a reduction of variables and provides projection of data in a new space related to the minimum of independent variables suitable to explain an appropriate amount of data

variance, so that, enhancing and easing data exploration and interpretation [38].

Cluster Analysis differs from PCA in its capability to detect similarities between samples and to define groups of samples from their spectra. Similarities among samples are estimated by means of distances: similar samples are characterized by small distances and the opposite for dissimilar ones. The dendrograms obtained encode the cluster structure of data and the partition of samples is obtained by cutting the dendrogram at the desired level of similarity [38].

PLS-DA is a linear classification method that combines the properties of partial least squares regression with the discrimination power of a classification technique [39]. The main goal of PLS-DA is to build a calibration model suitable to be applied in future classifications. A series of parameters can be used to evaluate the performance of classification models as true positive (TP) and true negative (TN), false positive (FP) and false negative (FN) classification results, sensitivity, specificity, accuracy, efficiency, and Matthews correlation coefficient (MCC). The TP rate is the probability that a positive sample can be classified as positive, TN involves that a negative sample can be classified as negative. A similar criterion is employed to determine FN and FP rates. Sensitivity is the ability of the model to correctly classify authentic samples inside their group using the TP and FN rate. Specificity is the capacity of the model to correctly identify samples relative to the values of TN and FP. The accuracy is the rate of correct classification, independently of the class of sample. The efficiency

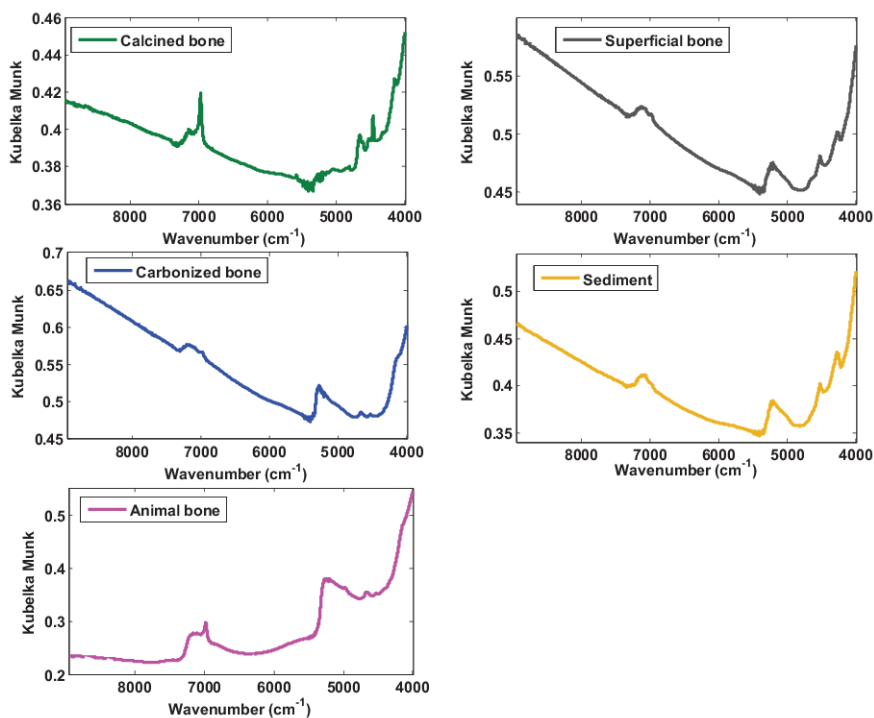


Fig. 1. NIR spectra of calcined, carbonized, animal, superficial part of bones and sediment samples in the region between 9000 and 4000 cm^{-1} without data pre-treatment.

and MCC summarize the model performance. The efficiency was determined as the arithmetic average of the values of sensitivity and specificity, where a value of 1 corresponds to an efficiency of 100%. MCC gives a value between -1 and +1, where a value of +1 represents perfect classification, 0 an erroneous classification and -1 an inverse classification [40,41].

PLS regression was applied to NIR data in order to develop prediction models for calcium, magnesium and strontium content in bone samples. Different spectra pre-processing, including multiplicative scatter correction (MSC), standard normal variation (SNV), a Savitzky-Golay first (FD) and second derivative (SD), and mean center (MC), were tested to treat raw data prior to the PLS regression using different spectral regions. The performance of PLS-NIR models was evaluated according to the root mean square error of cross validation and prediction (RMSECV and RMSEP), the correlation coefficient of prediction (R^2_{pred}), relative root mean square error of prediction (RRMSEP) and residual predictive deviation (RPD), calculated as the ratio between standard deviation (SD) of the prediction set and the RMSEP values [42].

3. Results and discussion

3.1. NIR spectra

Fig. 1 shows the NIR spectra of: a) calcined bone, b) carbonized bone, c) superficial part of bone, d) animal bone and, e) sediment samples, in the region between 9000 and 4000 cm^{-1} without any data pretreatment. The spectra region between 5500 and 8000 cm^{-1} corresponds to the first fundamental overtone of the mid-infrared OH stretching vibration, the region between 4000 and 5500 cm^{-1} is ascribed to the water OH overtones and carbonate combination bands, and the region between 8000 and 12000 cm^{-1} is related to the second fundamental overtone of the OH stretching vibrations and also includes electronic bands resulting from the presence of transition metal ions in the structure [43]. Position and assignation of the different bands is given in Table 1 [44–47]. Main differences found between burned bones as a function of their treatment are based on the band located at 6977 cm^{-1} , assigned to the first overtone of the stretching vibrations of OH group in hydroxyapatite. This band is higher in calcined bones than in carbonized ones. Additionally, the band at 5275 cm^{-1} is related to apatite, present in carbonized bones and not in calcined ones. These two differences can indicate the substitution of CO_3^{2-} in the OH^- site through A-type substitution due to the loss of hydroxylation degree during the carbonization process, being these bones less resistant to diagenetic alteration than calcined ones.

Table 1
Assignation of NIR bands present in skeletal remains and sediment samples.

Wavenumber (cm^{-1})	CLN	CRB	EXT	ANM	SED	Assignments	Ref
7197	●	●	●	●	●	1st overtone of Mg—OH stretching in dioctahedral layers	[47]
7070	—	—	●	—	●	1st overtone of the O—H stretch vibration in metal—O—H	[45]
6977	●	●	●	●	●	1st overtone stretching vibrations of OH group in hydroxyapatite	[44]
5275	—	●	—	●	●	Apatite	[46]
5190	●	●	●	●	●	Combination vibrations of H—O—H bend and O—H stretch of water	[45]
4966	—	—	—	●	—		
4812	●	—	—	—	—		
4655	●	●	●	●	—		
4619	●	—	—	—	—		
4530	—	●	●	●	●	clay minerals, like smectite and illite	[44]
4468	—	—	—	—	—		
4326	●	—	—	—	—	OH combination (Mg—OH, stretching + bending)	[47]
4265	—	—	●	—	●	contribution of calcite	[44]
4150	●	●	—	●	—		

CLN: Calcined bone; CRB: Carbonized bone; EXT: External bone; ANM: Animal bone; SED: Sediments.

3.2. Exploratory analysis

3.2.1. Principal component analysis (PCA)

PCA is nowadays the most common chemometric strategy for unsupervised exploratory data analysis. Scores plot often gives a good view of how the trends from different components relate to each other. Fig. 2 shows the scores plot for first and second principal components obtained from PCA treatment of NIR spectra of samples from Corral de Saus and Las Peñas Necropolis. After FD and MC pre-treatment and, selecting the region from 9000 to 4000 cm^{-1} , the two first principal components represent 72.61% of the explained variance, being 41.39% and 31.22% explained by PC1 and PC2, respectively. It can be appreciated that NIR spectra of sediment samples are clearly different from those of internal part of bone samples in the direction of PC1 and that, a few spectra of external part of bone samples are located near of sediment and others near of the internal part of bones. As It has been already evidenced [28] this indicates that external part of bones had undergone stronger changes than internal ones due to diagenetic factors. Also, in the direction of PC2 it can be seen that calcined bone sample spectra from Corral de Saus are located down, and carbonized bone spectra are placed up. Additionally, Las Peñas bone spectra are located close to these of internal part of bones from Corral samples It indicates similarities between these two groups. The same can be observed for SRM Bone ash NIST 1400 spectrum. Fig. 3 shows loading plots in order to understand the distribution of samples in PCA, indicating that responsible wavenumbers are ones around 7000 cm^{-1} and between 5500 and 4000 cm^{-1} . Consequently, it can be seen that PCA analysis using NIR offers a fast and green tool to identify changes in bones caused by environmental factors.

3.2.2. Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis (HCA) involves a measurement of the similarity between objects to be clustered, and thus samples with the maximum similarities in their NIR spectra were clustered together preferentially. Fig. 4a shows the dendrogram obtained from samples of Corral de Saus after FD and MC pre-treatment of NIR spectra and selecting the region from 9000 to 4000 cm^{-1} , using the average-paired as a distance measure. It can be seen that for a minimum cut off value of 0.016, all samples were classified into two main groups, which correspond to samples with common characteristics, like sediment and bones. The first group incorporates all sediments. However, it also included 5 samples of external part of bones, probably because these samples were altered by environmental factors. The second group is based on internal part of bone spectra, where it can be seen animal ones together with carbonized and calcined bones with also unknown samples and 2

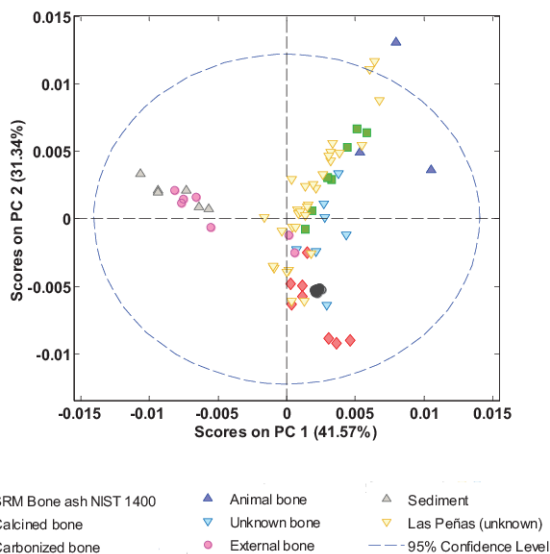


Fig. 2. Scores plot for first and second principal components obtained by PCA from NIR spectra of samples of Corral de Saus and Las Peñas Necropolis, after FD and MC pre-treatment using the region from 9000 to 4000 cm^{-1} .

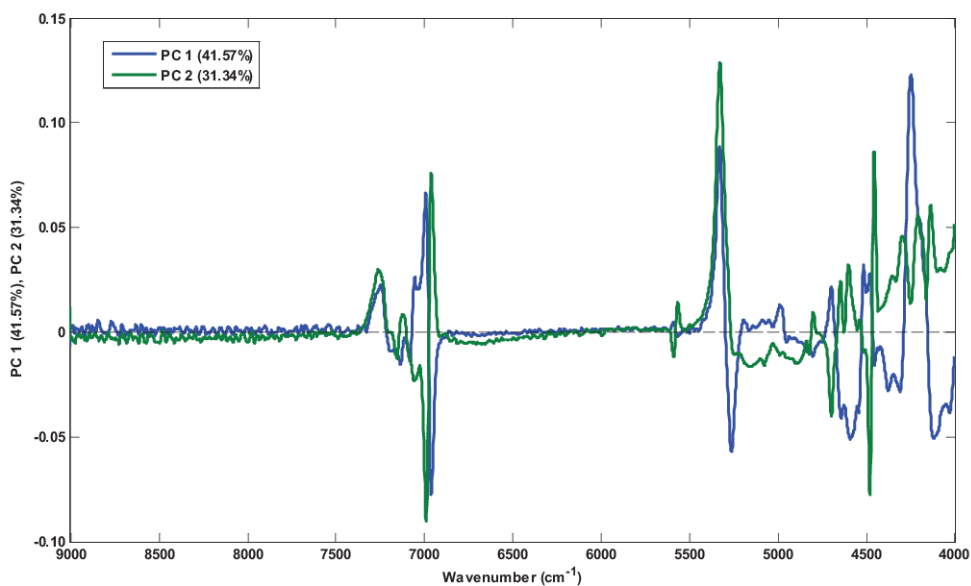


Fig. 3. Loadings of PC1 and PC2 after FD and MC pre-treatment of spectra using the region from 9000 to 4000 cm^{-1} .

external parts of bones. Additionally, it can be appreciated general trends in the behaviour of samples. For a cut-off value of 0.012 samples were classified as sediments, animal bones and internal part of bones. For a cut-off value of 0.0075, samples were classified

as calcined and carbonized bones. From the aforementioned facts it can be concluded that samples similar to sediments, as external part of bones, are included in this group, and samples different to the soil, as calcined bones, are placed in the opposite part of the

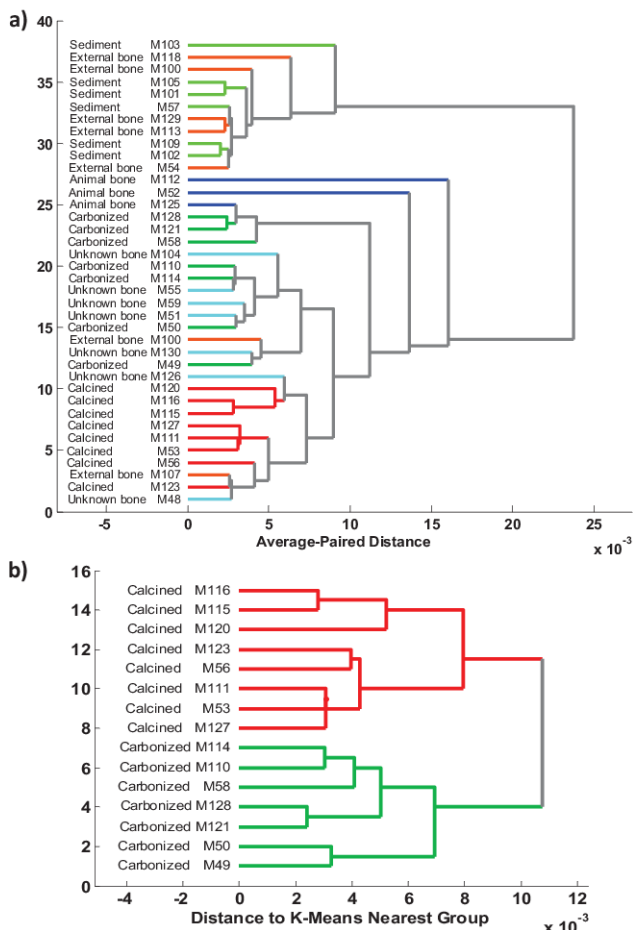


Fig. 4. Cluster dendrographic classification of bone and sediment samples based on their NIR spectra after FD and MC pre-treatment and selecting the region from 9000 to 4000 cm^{-1} concerning a) calcined, carbonized, animal, external part of bones and sediments established using average-paired distance and, b) calcined and carbonized bones classification using K-means nearest group as distance method.

dendrogram obtained from NIR spectra. So that, HCA indicates that calcined bones presents fewer structural changes than carbonized and external part of bones.

Furthermore, Fig. 4b shows a dendrogram established from spectra of internal part of bone samples that was build using a K-means nearest group as distance measure. Here, it can be appreciated that for a minimum cut off value of $8 \cdot 10^{-3}$, samples were classified into two main groups corresponding to calcined and carbonized bones, respectively.

3.3. Partial least squares for discriminant analysis (PLS-DA)

The main goal of PLS-DA is to build a calibration model that could be applied in future classifications. Burned bone class assignments are not always possible by looking at the remains colour. PLS-DA can be applied to classify bone samples depending

from their burning conditions. A Calibration set was built using NIR spectra of internal bones of Corral de Saus assigned to the classes "calcined" and "carbonized". Leave-one-out cross-validation method was used in the training set to select the number of latent variable in the PLS-DA. Fig. 5 illustrates the classification for PLS-DA, after FD and MC pre-treatment and selecting the region from 9000 to 4000 cm^{-1} . It can be seen a good class separation between calcined and carbonized bones, obtained using one latent variable which explains 70.83% of the total variance of the X data block and 83.56% of the Y data block. Sensitivity and specificity results for classification using PLS-DA were 100% for both, calibration and cross validation sets, obtaining satisfactory results in the classification of samples into true positive and negative classes. It indicates excellent separation of the classes. The efficiency and accuracy were 100% and the MCC had a value of 1. Additionally, PLS-DA model was validated employing

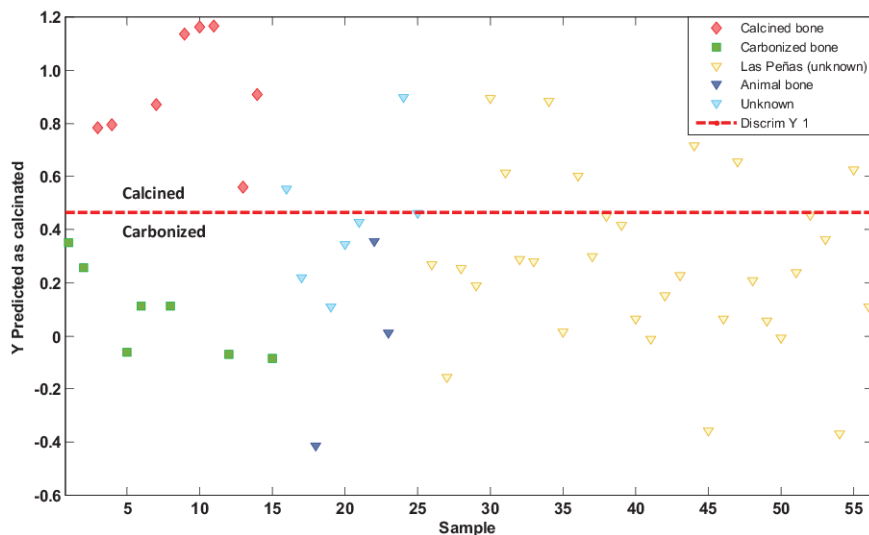


Fig. 5. Classification by PLS-DA of NIR spectra of bones, after FD and MC pre-treatment and selecting the region from 9000 to 4000 cm^{-1} .

permutation testing using 30 iterations and one latent variable. p-values of 0.010 and 0.008 were obtained for self-prediction and cross-validation, respectively, by employing a randomization *t*-test for evaluating residuals, thus confirming the significance of the original PLS-DA model at a 95% confidence level.

The aforementioned model was applied to evaluate the NIR spectra of bone samples of unknown burning conditions derived from Corral de Saus and Las Peñas Necropolis, together with unburned animal bone samples. As it can be appreciated from Fig. 5, all animal bone samples were similar to carbonized bones, whereas some of the unknown bones were assigned to the class of calcined bones and others to the class of carbonized bones.

Therefore, it can be concluded that PLS-DA allows a screening classification to assign bone class as calcined or carbonized, avoiding the confusing visual interpretation. So that, as the calcined bones are more resistant to diagenetic alteration due to their high crystallinity, PLS-DA permits to do an adequate selection of bones in order to ensure forensic studies that could comprise biological or radiocarbon dating investigations, additionally than studies based on the mineral composition of human remains concerning cultural aspects as diets or life habits.

3.4. Prediction of major elements by partial least squares (PLS) regression model/Prediction of major elements by partial least squares (PLS) regression model

PLS regression models were built to predict the presence of major elements in bone samples from their NIR spectra. A total of 38 samples of Corral de Saus Necropolis were used as calibration set. To confirm the suitability of the chosen pre-processing strategy for model building, 31 samples of Las Peñas Necropolis were employed as external validation set. Table 2 shows the descriptive statistics mean, standard deviation (SD) and range of concentration for the determination of Ca, Mg and Sr in samples included in the calibration and validation sets. Additionally, a sample of Bone ash NIST 1400 was used as standard reference material for evaluating accuracy of the analytical procedure.

Table 2
Concentration ranges of major elements present in samples corresponding to calibration, and validation data set.

Analyte	Set	Samples	Range	Mean value	SD
Calcium (mg g^{-1})	Calibration	35	250–461	360	49
Magnesium ($\mu\text{g g}^{-1}$)		36	790–	8091	10820
			34257		
Strontium ($\mu\text{g g}^{-1}$)		35	161–1061	494	215
Calcium (mg g^{-1})	Validation	31	341–397	363	16
Magnesium ($\mu\text{g g}^{-1}$)		31	1213–4116	2014	711
Strontium ($\mu\text{g g}^{-1}$)		29	179–1068	503	176

SD: Standard deviation.

Different spectral regions and pre-processing strategies were applied to build the best models to evaluate Ca, Mg and Sr concentration in burned bones. For all elements determination, regions selected were those between 9000 and 4000 cm^{-1} . In all cases, FD was calculated with a window of 15 points and second order polynomial, together with MC treatment, were adopted as signal pre-processing. Models with 3, 8 and 3 latent variables (LVs) were selected for calcium, magnesium and strontium, respectively and these LVs achieved an explained variance of 72.62%, 95.9% and 87.61% for X variables and 72.77%, 99.61% and 80.90% for Y variables, respectively. Outliers can be identified by Q residual versus the Hotelling T2 values, being removed 3 samples prior to build calibration models of Ca and Sr and 2 samples, in the case of Mg. Moreover 2 samples were removed for validation set of Sr. Fig. 6 shows the correlation among the predicted and reference values for the concentrations of Ca, Mg and Sr for the calibration and validation sets, with good coefficients of determination of calibration (R^2_{cal}), cross validation (R^2_{CV}) and prediction (R^2_{pred}). Acceptable RRMSEP values were obtained for calcium with 5% mean values. However, average relative prediction, for Mg and Sr, provided high values of RRMSEP, with RPD values of 0.86, 0.24 and 1.08, respectively for the three elements under study. The most

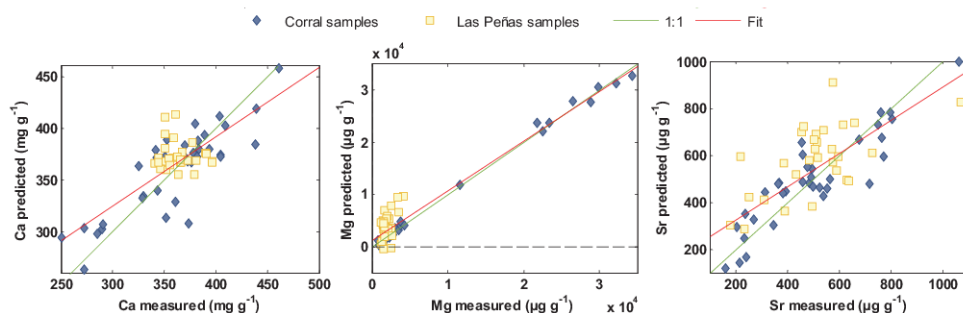


Fig. 6. Predicted vs reference values for PLS-NIR determination of calcium, magnesium and strontium in burned bone samples.

Table 3

Description and validation parameters of PLS-NIR models developed for the determination of calcium, magnesium and strontium.

Element	Pre-process	LV's	RMSEC	R ² Cal	RMSECV	R ² CV	RMSEP	RRMSEP (%)	RPD
Calcium (mg g ⁻¹)	FD, MC	3	25.4	72.8	29.9	62.2	18.9	5.2%	0.86
Magnesium (μg g ⁻¹)	FD, MC	8	665.7	99.6	3115.7	92.4	2941.5	146%	0.24
Strontium (μg g ⁻¹)	FD, MC	3	92.5	80.9	112.6	71.8	162.57	32.3%	1.08

LV: number of latent variables; RMSEC: Root mean square error of calibration; RMSECV: Root mean square error of cross validation; RMSEP: Root mean square error of prediction; RRMSEP: Relative root mean square error of prediction; R² Cal: coefficient of determination of calibration; R² CV: coefficient of determination of cross validation; R² Pred: coefficient of determination of prediction.

important calibration and validation parameters of the developed PLS-NIR models are summarized in Table 3.

3.5. NIR quality assurance and quality control (QA/QC)

Inter-day precision and accuracy of PLS regression models to predict Ca, Mg and Sr were evaluated from the analysis of archaeological samples and a reference one. CV varies from 0.2% till 1.8% for Ca and Sr, being data found for Mg between 4.7% and 13.8%. On the other hand the bias obtained for the reference material was lower than 10% in the case of Ca and Sr determination being obtained a range of relative errors between -4.3% and -2.3% for Ca and from 6.9% till 9.9% for Sr. In the case of Mg the accuracy relative errors varied from -26.5% till -8.4%. Moreover, PCA could evaluate the precision of the method by visualization of data projection due to the fact that replicates were placed all together and near of calcined bones (Table 4).

3.6. Evaluation of methods

The developed NIR methodology allowed us to identify the best preserved samples and to predict Ca, Mg and Sr in burned bones. However it is necessary to evaluate advantages and drawbacks of the proposed methodology versus the reference one and in fact it can be concluded that NIR spectroscopy methodology is faster than

Table 4

Inter-day precision and accuracy parameters of the PLS-NIR methodology for the determination of calcium, magnesium and strontium in Bone ash NIST 1400.

	Ca	Mg	Sr
CV (%)	0.2-1.5	4.7-13.8	0.3-1.8
Er (%)	(-4.3)-(-2.3)	(-26.5)-(-8.4)	6.9-9.9

CV: coefficient of variation.
Er: relative error.

ICP-OES for sample preparation and measurement and to obtain results. However, the lack of NIR sensitivity is the main limitation of the method, being only the major components of bones as Ca, Mg and Sr able to be determined and not the trace ones.

Moreover, evaluation of the proposed NIR procedure and the ICP-OES reference methodology was made by comparing them according by the so called Green Certificate. The greenness of the analytical methodologies were evaluated by using the eco-scale proposed by Van Aken et al. [48], developed by Galuszka et al. [49] and modified by Armenta et al. [50] to establish the Green Certificate. It includes a new criterion to quantify the penalty point values and associate the eco-scale value to a category class from A to G. Fig. 7 shows the results of the green evaluation of FT-NIR and ICP-OES. One penalty point was assigned to FT-NIR due to the consume of energy by the mineralization of samples in muffle furnace. The NIR procedure has a category class A on the Green Certificate with a mark of 99. For ICP-OES, the final mark was of 86.5 due to 2 penalty points assigned for the use of HNO₃ and HCl in the sample treatment, the use of standard solutions and argon for plasma generation; 3 penalty points due to the consume of energy by the muffle furnace and ICP-OES and 8.5 penalty points due to the wastes. So, ICP-OES can be considered as a B class in Green Certificate ranking.

3.7. Forensic applications

Forensic anthropologists correlate the maximum exposure temperature of bones remains with colour changes to obtain a correct interpretation of burned bones. This estimated temperature could be a difficult task being a subjective and confusing method due to different factors that should be considered such as; time, temperature, oxygen availability and external and diagenetic factors. PLS-DA by using NIR spectra permits an accurate burned bone classification avoiding the confusing colour interpretation.

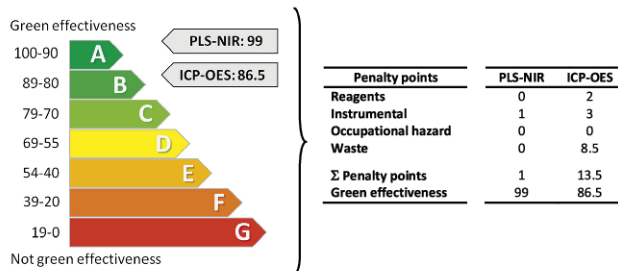


Fig. 7. Green evaluation of PLS-NIR vs ICP-OES for Ca, Mg and Sr determination in bone remains.

Calcined bones are more resistant to environmental alterations than carbonized ones being thus calcined bones more adequate for investigation studies about radiocarbon dating and diet reconstruction by using Sr/Ca ratio. PCA analysis using NIR offers a fast and green tool to identify changes in bones caused by environmental factors. Therefore, an appropriate selection of samples could avoid interpretative errors related to the structure and chemical composition of bones that could be post-mortem modified by diagenetic factors.

On the other hand, the use of PLS-NIR as a screening methodology for Ca, Mg and Sr estimation offer an additional value to the use of NIR spectra in bone remains studies.

4. Conclusions

Interesting results were obtained using PCA, HCA and PLS-DA to classify burned bones. PCA, using two principal components, was able to discriminate bone remains affected by post depositional diagenetic processes to those relatively free due to their lack of exchanges with the soil components. HCA permitted to clearly separate calcined and carbonized bones and could be useful to roughly understand the thermal treatment of unknown remains. PLS-DA multivariate provided an accurate tool to discriminate between calcined and carbonized and to classify accurately unknown burned bones.

From a prior study concerning buried bones from two late roman necropolises [51] the use of NIR spectroscopy combined with chemometrics provides a rapid and cost efficient method to screen the concentration of calcium, magnesium and strontium also in burned bones aiming to understand post-mortem changes by environmental degradation.

In short, it can be concluded that the proposed methodologies, based on the use of NIR spectroscopy combined with chemometric tools provide fast and green approaches to select the most suitable samples for forensic studies and evaluation of major mineral elements in bone remains. It is very important due to the fact that carbonized bones are more prone than calcined ones to post-depositional processes and those can produce mistakes during data interpretations. Additionally, PLS regression models built from NIR spectra provided a screening tool to predict Ca, Mg and Sr in burned bones.

Acknowledgements

Authors acknowledge the financial support of Generalitat Valenciana (Project PROMETEO II/2014/077) and Ministerio de Economía y Competitividad-Feder (Projects CTQ 2014-52841-P). M.C acknowledges the FPI grant (BES-2012-055404) provided by the Ministerio de Economía y Competitividad of the Spanish government.

The authors would like to thanks all the students of Chemistry and Archaeology which have contributed to the realization of this study.

References

- [1] S.A. Symes, D.C. Dirkmaat, S. Ousley, E.D. Chapman, L.L. Cabo, Recovery and Interpretation of Burned Human Remains, 2012. NIJ Report, Grant #2008-DN-BX-K131. Available at <https://www.ncjrs.gov/pdffiles1/nij/grants/237966.pdf>.
- [2] T.J.U. Thompson, Heat-induced dimensional changes in bone and their consequences for forensic anthropology, *J. Forensic Sci.* 50 (2005) 1008–1015.
- [3] D. Gonçalves, T.J.U. Thompson, E. Cunha, Implications of heat-induced changes in bone on the interpretation of funerary behaviour and practice, *J. Archaeol. Sci.* 38 (2011) 1308–1313.
- [4] T.J.U. Thompson, Recent advances in the study of burned bone and their implications for forensic anthropology, *Forensic Sci. Int.* 146 (Suppl.) (2004) S203–S205.
- [5] D.H. Ubelaker, The forensic evaluation of burned skeletal remains: a synthesis, *Forensic Sci. Int.* 183 (2009) 1–5.
- [6] S.T.D. Ellingham, T.J.U. Thompson, M. Islam, G. Taylor, Estimating temperature exposure of burnt bone – A methodological review, *Sci. Justice* 55 (2015) 181–188.
- [7] S.A. Symes, C.W. Rainwater, E.N. Chapman, D.R. Glipson, A.L. Piper, Chapter 2. Patterned thermal destruction of human remains in a forensic setting, *The Analysis of Burned Human Remains*, Academic Press, San Diego, 2008.
- [8] P. Shipman, G. Foster, M. Schoeninger, Burnt bones and teeth: an experimental study of color, morphology, crystal structure and shrinkage, *J. Archaeol. Sci.* 11 (1984) 307–325.
- [9] M. Color, Munsell Soil Color Charts, Macbeth Division of Kollmorgan Instruments Corp, New Windsor, N.Y. NY, 1994.
- [10] J.B. Devlin, N.P. Herrmann, Chapter 6. Bone color as an interpretive tool of the depositional history of archaeological cremains, *The Analysis of Burned Human Remains*, Academic Press, San Diego, 2008.
- [11] R.F. Castillo, D.H. Ubelaker, J.A.L. Acosta, G.A.C. de la Fuente, Effects of temperature on bone tissue: Histological study of the changes in the bone matrix, *Forensic Sci. Int.* 226 (2013) 33–37.
- [12] P.L. Walker, K.P. Miller, Time, temperature, and oxygen availability: an experimental study of the effects of environmental conditions on the color and organic content of cremated bone, *Am. J. Phys. Anthropol.* 40 (Suppl.) (2005) 216–217.
- [13] R.A. Nicholson, A morphological investigation of burnt animal bone and an evaluation of its utility in archaeology, *J. Archaeol. Sci.* 20 (1993) 411–428.
- [14] S.A. Symes, C.W. Rainwater, E.N. Chapman, D.R. Glipson, A.L. Piper, Patterned thermal destruction of human remains in a forensic setting, in: C.W. Schmidt, S.A. Symes (Eds.), *The Analysis of Burned Human Remains*, Academic Press, London, 2008, pp. 15–54.
- [15] J.B. Devlin, N.P. Herrmann, Bone color as an interpretive tool of the depositional history of archaeological cremains, in: C.W. Schmidt, S.A. Symes (Eds.), *The Analysis of Burned Human Remains*, Academic Press, London, 2008, pp. 109–128.
- [16] R. Shahack-Gross, O. Bar-Yosef, S. Weiner, Black-coloured bones in Hayonim Cave, Israel: differentiating between burning and oxide staining, *J. Archaeol. Sci.* 24 (1997) 439–446.
- [17] C.K. Brain, A. Sillen, Evidence from the Swartkrans cave for the earliest use of fire, *Nature* 336 (1988) 464–466.
- [18] K.A. Gross, C.C. Berndt, Biomedical application of apatites, *Rev. Mineral. Geochem.* 48 (2002) 631–672.
- [19] J.N. Lanting, A.T. Aerts-Bijma, J. van der Plicht, Dating cremated bone, *Radiocarbon* 43 (2001) 249–254.
- [20] T.A. Surovell, Radiocarbon dating of bone apatite by step heating, *Geoarchaeology* 15 (2000) 591–608.
- [21] G. Quarta, L. Calcagnile, M. D'Elia, L. Maruccio, V. Gaballo, A. Caramia, A combined PIXE/PIGE approach for the assessment of the diagenetic state of

- cremated bones submitted to AMS radiocarbon dating, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* 294 (2013) 221–225.
- [22] A. Zazzo, J.-F. Salâgé, Radiocarbon dating of biological apatites: a review, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 310 (2011) 52–61.
- [23] J. Olsen, K.M. Hornstrup, J. Heinemeier, P. Bennike, H. Thirane, Chronology of the danish bronze age based on ¹⁴C dating of cremated bone remains, *Radiocarbon* 53 (2011) 261–275.
- [24] J. Olsen, J. Heinemeier, P. Bennike, C. Krause, K. Margrethe Hornstrup, H. Thirane, Characterisation and blind testing of radiocarbon dating of cremated bone, *J. Archaeol. Sci.* 35 (2008) 791–800.
- [25] G. Gruppe, S. Hummel, Trace element studies on experimentally cremated bone. I. Alteration of the chemical composition at high temperatures, *J. Archaeol. Sci.* 18 (1991) 177–186.
- [26] J.J. Schultz, M.W. Warren, J.S. Krigbaum, Chapter 4. Analysis of human remains: gross and chemical methods, *The Analysis of Burned Human Remains*, Academic Press, San Diego, 2008.
- [27] M.E. Subira, A. Malgosa, The effect of cremation on the study of trace elements, *Int. J. Osteoarchaeol.* 3 (1993) 115–118.
- [28] G. Gallelo, J. Kuligowski, A. Pastor, A. Diez, J. Bernabeu, Biological mineral content in Iberian skeletal remains for control of diagenetic factors employing multivariate statistics, *J. Archaeol. Sci.* 40 (2013) 2477–2484.
- [29] L.D. Mkukuma, J.M.S. Skakle, L.R. Gibson, C.T. Imrie, R.M. Aspden, D.W.L. Hukins, Effect of the proportion of organic material in bone on thermal decomposition of bone mineral: an investigation of a variety of bones from different species using thermogravimetric analysis coupled to mass spectrometry, high-temperature X-ray diffraction, and fourier transform infrared spectroscopy, *Calcif. Tissue Int.* 75 (2004) 321–328.
- [30] T. Leventouri, Synthetic and biological hydroxyapatites: crystal structure questions, *Biomaterials* 27 (2006) 3339–3342.
- [31] J.D. Pasteris, B. Wopenka, J.J. Freeman, K. Rogers, E. Valsami-Jones, J.A.M. van der Houwen, M.J. Silva, Lack of OH in nanocrystalline apatite as a function of degree of atomic order: implications for bone and biomaterials, *Biomaterials* 25 (2004) 229–238.
- [32] B. Wopenka, J.D. Pasteris, A mineralogical perspective on the apatite in bone, *Mater. Sci. Eng. C* 25 (2005) 131–143.
- [33] *Apatites and their Synthetic Analogues—Synthesis, Structure, Properties and Applications*, Petr Ptacek, 2016.
- [34] T.J.U. Thompson, M. Gauthier, M. Islam, The application of a new method of Fourier Transform Infrared Spectroscopy to the analysis of burned bone, *J. Archaeol. Sci.* 36 (2009) 910–914.
- [35] T.J.U. Thompson, M. Islam, K. Piduru, A. Marcel, An investigation into the internal and external variables acting on crystallinity index using Fourier Transform Infrared Spectroscopy on unaltered and burned bone, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 299 (2011) 168–174.
- [36] M. Lebon, et al., New parameters for the characterization of diagenetic alterations and heat-induced changes of fossil bone mineral using Fourier transform infrared spectrometry, *J. Archaeol. Sci.* 37 (2010) 2265–2276.
- [37] J.M. Martínez García, *La Necrópolis Ibérica De Las Peñas* (Zarra, Valencia), vol. XIX, Archivo de Prehistoria Levantina, Valencia, 1989.
- [38] D. Ballabio, A MATLAB toolbox for Principal Component Analysis and unsupervised exploration of data structure, *Chemom. Intell. Lab. Syst.* 149 (2015) 1–9 (Part B).
- [39] M.J. Hidalgo, D.C. Fechner, E.J. Marchevsky, R.G. Pellerano, Determining the geographical origin of *Sechium edule* fruits by multielement analysis and advanced chemometric techniques, *Food Chem.* 210 (2016) 228–234.
- [40] M.R. Almeida, C.H.V. Fidelis, L.E.S. Barata, R.J. Poppi, Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation, *Talanta* 117 (2013) 305–311.
- [41] F.B. de Santana, L.C. Gontijo, H. Mitsutake, S.J. Mazivila, L.M. de Souza, W. Borges Neto, Non-destructive fraud detection in rosehip oil by MIR spectroscopy and chemometrics, *Food Chem.* 209 (2016) 228–233.
- [42] P.C. Williams, D. Sobering, How do we do it: a brief summary of the methods we use in developing near infrared calibrations, in: A.M.C. Daves, P.C. Williams (Eds.), *Near Infrared Spectroscopy: The Future Waves*, NIR Publications, Chichester, UK, 1995, pp. 185–188.
- [43] R.L. Frost, B.J. Reddy, S. Bahfenne, J. Graham, Mid-infrared and near-infrared spectroscopic study of selected magnesium carbonate minerals containing ferric iron—implications for the geosequestration of greenhouse gases, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* 72 (2009) 597–604.
- [44] Y. Ning, J. Li, W. Cai, X. Shao, Simultaneous determination of heavy metal ions in water using near-infrared spectroscopy with preconcentration by nano-hydroxyapatite, *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* 96 (2012) 289–294.
- [45] V. Aranda, A. Domínguez-Vidal, F. Comino, J. Calero, M.J. Ayora-Cañada, Agro-environmental characterization of semi-arid Mediterranean soils using NIR reflection and mid-IR-attenuated total reflection spectroscopies, *Vib. Spectrosc.* 74 (2014) 88–97.
- [46] D. Thomas, C. McGovern, A. Chinsamy, M. Manley, Near infrared analysis of fossil bone from the Western Cape of South Africa, *J. Infrared Spectrosc.* 19 (2011) (2011) 151.
- [47] M. Mora, M. Isabel López, M. Ángeles Carmona, C. Jiménez-Sanchidrián, J. Rafael Ruiz, Study of the thermal decomposition of a sepiolite by mid- and near-infrared spectroscopies, *Polyhedron* 29 (2010) 3046–3051.
- [48] K. Van Aken, L. Streckowski, L. Patiny, EcoScale, a semi-quantitative tool to select an organic preparation based on economical and ecological parameters, *Beilstein J. Org. Chem.* 2 (2006) 3.
- [49] A. Galuszka, P. Konieczka, Z.M. Migaszewski, J. Namiésnik, Analytical eco-scale for assessing the greenness of analytical procedures, *Trac-Trend Anal. Chem.* 37 (2012) 61–72.
- [50] S. Armenta, M. de la Guardia, J. Namiesnik, green microextraction, in: M. Valcarcel (Ed.), *Analytical Microextraction Techniques*, Bentham Science, 2017.
- [51] M.M. Cascant, S. Rubio, G. Gallelo, A. Pastor, S. Garrigues, M. de la Guardia, Prediction of alkaline earth elements in bone remains by near infrared spectroscopy, *Talanta* 162 (2017) 428–434.

