

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: http://es.creativecommons.org/blog/licencias/

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: (c) (1) (a) https://creativecommons.org/licenses/?lang=en



Aplicación de la espectroscopia vibracional como técnica de control en la industria farmacéutica

Judit Puig Bertotto

Tesis Doctoral

Programa de Doctorat en Química

Directors: Prof. Santiago Maspoch* y Prof. Jordi Coello

Departament de Química

Facultat de Ciències

2022

^{*} jubilació a setembre de 2021

Memòria presentada per aspirar a	l Grau de Doctor per Judit Puig Bertotto
	Judit Puig Bertotto
Vist i plau	
	Prof. Santiago Maspoch
	Prof. Jordi Coello
Bellaterra, 7 de juliol de 2022	

"Recorda, si plou dins teu, que caminar cura les ferides. Que viure, viure no és estar vius. Viure és l'actitud d'omplir la vida. La vida, allò que passa quan rius."

Txarango

AGRAÏMENTS

Una part d'aquesta tesis va dedicada a la oportunitat que em van brindar des del grup de Quimiometria aplicada de la UAB. Gràcies al Prof. Santiago Maspoch i al Prof. Jordi Coello per deixar-me endinsar-me en el món del multivariable, totalment desconegut per mi uns anys enrere i ajudar-me en tot el camí d'aprenentatge. Gràcies també al Dr. Manel Alcalà que forma part d'aquesta família de la UAB.

Una de les parts mes boniques d'aquest recorregut ha estat compartir-lo amb les mateixes persones que vaig compartir seient el primer dia de classes de la universitat. Gràcies al destí per posar a una amiga del institut que compartia passió per la Química, Mariona, en el meu mateix horari i curs. A més, gràcies al *patatabrava* per posar en el meu camí a la Marta i trobar-nos el primer dia a la plaça cívica. Quina sensació aquella primera entrada a classe totes juntes, com si ens coneguéssim de tota una vida. Que segueixi passant el temps i seguim iguals. Només dirvos que les nits no moren mai, igual que les experiències. Gràcies també als que van formar part del camí des del principi o als que es van anar incorporant en ell: Laura, Claudia, Marc V, Arnau, Dani, Marcelo, Kevin. Igual que a la petita família que vam formar al departament d'analítica, les Reines Analítiques, gràcies per fer els dies més amens compartint esmorzars, cafès i pastissos, molts pastissos.

Una altre de les parts més boniques ha estat compartir tot el recorregut amb els companys del laboratori. Gràcies Diego, Laia, Glòria i Judit per transmetre'm l'experiència i el coneixement. Moltes gràcies Dr. Cruz per guiar-me i oferir-me sempre noves oportunitats i/o aventures. Però, el millor del laboratori ha estat descobrir i compartir amb persones que alleugeren el camí. Gràcies Elia per aparèixer i ser la companya ideal de laboratori amb qui afrontar problemes que no podia solucionar només una *estampita*. A la Claudia, por aparecer y compartir muchas horas durante meses. Alumnos como tu son los que te demuestran que lo que haces, vale la pena. Ah, y Roma (y otros destinos) nos esperan. A l'Aira per recolzar-me incondicionalment sempre i obrir-me les portes a un nou món. Només tinc paraules d'agraïment per a tu. I, sobretot, a la meva companya d'etapa, de confessions, de compartir-ho tot. Gràcies Desirèe per estar present en cada pas, per escoltar-me i ajudar-me, i sobretot, per estar dia rere dia al meu costat. Persones com tu son les que valen la pena. Ets i seràs el millor descobriment d'aquesta etapa. Seguirem sumant experiències juntes. Moltes gràcies amiga!

Gràcies també als que es van incorporar al camí més tard brindant-me una nova oportunitat. Gràcies Anna Peguero per acollir-me i ensenyar-me, junt a l'Aira, el món de la industria farmacèutica. I, gràcies a tots els que formen part d'aquest món.

Les hores i temps de dedicació son infinites, però sense l'ajuda dels més propers, no seria possible afrontar aquest repte. Per una banda, gràcies a les amigues de sempre, per seguir estant en tot moment, tindre la paciència d'aguantar-me i, sobretot, entendre'm. Soc qui soc avui, en part, gràcies a vosaltres. Gràcies Maria per acompanyar-me de la mà des de la guarderia, créixer amb tu, m'ha fet millor persona. La teva bondat és infinita. Gràcies Carla per aparèixer en l'època més adolescent, pel teu caràcter i la teva alegria constant. La teva força innata es contagia. *Flowerpowers*, gràcies per estar sempre. Gràcies també a la Judith, qui havia de dir que un entrenador (*paximpaxam*) seria la persona que ens unís. El teu esforç i perseverança sempre m'ha impressionat. De tú, he après a no rendir-me i que amb constància tot s'aconsegueix. Ah, i gràcies per fer-me tieta postissa. Gràcies també a la Marta per ser-hi sempre tot i la "distància". La teva tranquil·litat sempre alegra.

Per altre banda, gràcies a tota la meva família per recolzar-me en cada una de les meves decisions i estar allà sempre tot i que això fos un món completament nou per ells. Moltes gràcies als meus pares pel tot el seu esforç. Sense ells, res d'això hagués estat possible. Igual que al petit de la casa, però que a la vegada és el més gran. També, gràcies a la meva segona família per ser-hi. Però, sobretot, gràcies al meu company de vida, Marc, mai millor dit. Sense tu, el camí no hagués estat possible. Moltes gràcies per ajudar-me a créixer, per compartir, per ser i estar. Aquesta etapa ens serveix per demostrar que res no ens aturarà mai, que podem i afrontem tot el que vingui per davant, sigui el que sigui, sempre amb un somriure, però, sempre de la mà. Gràcies per ser la meva força. Itestmo!

En aquest camí, com en la vida, es difícil identificar totes les variables del procés sense les eines adequades. Per això, gràcies a tots els que han format part d'això. Moltes gràcies!.

RESUMEN

En esta tesis se han planteado distintas aplicaciones de las técnicas espectroscópicas vibracionales para su aplicación en la industria farmacéutica. En esta industria, se requiere garantizar la calidad de los productos tanto a su llegada como antes de su liberación. Por ello, es importante el tiempo de análisis requerido en cada una de las etapas de fabricación de un medicamento, para garantizar la liberación antes de la caducidad del producto. Las técnicas NIR y Raman son una gran alternativa a las técnicas más tradicionales para garantizar la calidad de los productos, reduciendo significativamente el tiempo de análisis. En este trabajo, se ha demostrado su aplicabilidad para la identificación de los materiales de partida, la identificación de producto acabado, la determinación del contenido del API en comprimidos y geles, así como la evaluación de la uniformidad de contenido del API.

La primera parte del trabajo demuestra la capacidad de la espectroscopia vibracional para la identificación de materiales de partida y producto acabado. Primero, se ha demostrado la aplicabilidad de la espectroscopia en el infrarrojo cercano (NIR) para la identificación de los materiales de partida recibidos en una industria farmacéutica. La identificación mediante esta técnica permite reducir el tiempo de entrada de los distintos materiales a fabrica, así como la verificación a tiempo real de todos ellos. Este método de identificación se ha construido mediante una biblioteca de espectros NIR representativo de cada uno de los productos recibidos, que ha sido validada posteriormente. Además, se han identificado los factores más importantes que contribuyen a la obtención de un resultado fuera de especificación (OOS) tras la realización de un análisis de riegos. Una vez identificados los factores, se han propuesto acciones mitigadoras de este riesgo. En segundo lugar, se ha demostrada la utilidad de instrumentos portátiles NIR y Raman junto con técnicas de reconocimiento de pautas (PRM) para la identificación de comprimidos fabricados por siete industrias farmacéuticas distintas.

La segunda parte del trabajo demuestra la capacidad de estas técnicas para la determinación del contenido de API tanto en comprimidos como en muestras de gel. Por un lado, se ha realizado un análisis univariante mediante una adición estándar para la determinación del API en geles comerciales. La obtención de los espectros se ha realizado con un Raman portátil equipado con una sonda que está en contacto con el gel. De esta forma, no es necesaria la preparación de un conjunto de muestras de calibración. Por otro lado, se ha evaluado la uniformidad de contenido (UC) en comprimidos. Esta evaluación se ha realizado mediante el uso de la intensidad Raman de un pico característico del API. La monitorización de esta intensidad permite obtener una

imagen de la distribución del API en la superficie del comprimido tras la definición de un mapeo de puntos. Además, se ha realizado una optimización del número de puntos necesarios para obtener una imagen fiable y representativa del API.

La última parte de este trabajo está basado en el desarrollo y validación un modelo de regresión por mínimos cuadrados (PLS) para la cuantificación de dos APIs en comprimidos farmacéuticos. Este modelo se ha construido con datos NIR obtenidos con un equipo de sobremesa y un equipo portátil, para así evaluar la utilidad de los instrumentos portátiles en aplicaciones de cuantificación.

LISTADO DE ABREVIACIONES

Acrónimo	Significado en inglés	Significado en español
1D	1st derivative	1ª derivada
2D	2nd derivative	2ª derivada
AL	-	Alter
ALP	-	Alprazolam
ALS	Asymmetric Least Squares	Mínimos cuadráticos asimétricos
AOTF	Acousto Optical Tuneable Filter	Filtros acústo-ópticos
API	Activa Pharmaceuticla Ingredient	Ingrediente activo farmacéutico
ASP	-	Aspecto
AV	Acceptance Value	Valor de aceptación
BLC	Baseline Correction	Corrección de línea base
BW	Baseline Wavelets	Corrección wavelet
CCD	Charged Coupled Device	Dispositivo de carga acoplada
CCF	Central Composite Face-center design	Diseño central cúbico centrado en las caras
CI	-	Cinfa
CV	Cross Validation	Validación cruzada
DOE	Design of Experiments	Diseño de experimentos
Dt	Detectability	Detectabilidad
EMA	European Medicines Agency	-
EP	European Pharmacopoeia	Farmacopea Europea
ESP	Specificity	Especificidad
EXA	-	Exactitud
EXC	Excipients	Excipientes
FDA	Food and Drug Administration	-
FMEA	Failure Mode and Effects Analysis	Análisis modal de fallos y efectos
FMECA	Failure Mode and Effects Critical Analysis	Análisis modal crítico de fallos y efectos
FN	False Negative	Falso negativo
FP	False Positive	Falso positivo
FT	Fourier Transform	Transformada de Fourier
HIG	-	Higroscopia
IBU	-	Ibuprofeno
IBUNa	-	Ibuprofeno sódico
ICH	International Council Harmonization	-
IR	Infrared	Infrarrojo
KPI	Key Perfomance Indicator	Indicador de clave de rendimiento
LD	Linear Discriminant	Lineales discriminantes
LDA	Linear Discriminant Analysis	Análisis lineal discriminante
LIS	-	Lisinopril
LOOCV	Leave One Out Cross Validation	-
MA	-	Mabo
MIR	Mid Infrared	Infrarrojo medio
MP	-	Material de partida
MY	-	Mylan
NIPALS	Nonlinear Iterative Partial Least Squares	Mínimos cuadrados parciales iterativos no lineales
NIR	Near Infrared	Infrarrojo cercano
NO	-	Normon
OOS	Out Of Specification	Fuera de especificación
P	Probability	Probabilidad

PCP Principal Component Componentes principales PCA Principal Component Analysis Análisis por componentes principales PCA-LDA Principal Component Analysis Análisis por componentes principales PCA-LDA Principal Component Analysis Análisis por componentes principales PCT - Paracetamol PE - Pensa PF - Pensa PF - Pensa PLS-DA Partial Least Squares PBA Partial Least Squares Piscriminant Analysis cuadrados PCS PCT - Polimeria Análisis discriminante por mínimos cuadrados PLS-DA Partial Least Squares-Discriminant Analysis cuadrados PCS PRE - Polimorfismo PRE - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Cross Validation RMSEC Root Mean Square Error of Cross Validation RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - Soft Independent Modelling of Class Analogy - True Positive Positivo verdadero TP True Positive Negativo verdadero TP True Positive Positivo verdadero TP True Positive Uniformity Uniformidad de contenido de validación VAR - Conjunto de validación verdiablidad VN Vector normalization Normalización vectorial	Acrónimo	Significado en inglés	Significado en español
PCA-LDA Principal Component Analysis PCA-LDA Principal Component Analysis-Linear Discriminant Analysis PCT	PAT	Process Analytical Technology	Tecnologia analítica de procesos
PCA-LDA Principal Component Analysis - Linear Discriminant Analysis PCT - Paracetamol PE - Pensa PF - Pensa PLS - Pensa PLS-DA Partial Least Squares - Regresión por mínimos cuadrados PLS-DA PRE - Polimorfismo PRE - Precisión PRE - Precisión PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Cross Validation Cruzada RMSEV Root Mean Square Error of Cross Validation Cruzada RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEN Sensibility Selectividad SEN Sensibility Sensibilidad SOR Savitzky-Golay - True Negative Prositive Positivo verdadero TP True Positive Positivo Verdadero TPR - True Positive Positivo Verdadero TPR - True Positive Positivo Verdadero TPR - True Positive Uniformity Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAI — Conjunto de validación Variabilidad VAR - Variabilidad Normalización vectorial	PC	Principal Component	Componentes principales
PCA-LDA Discriminant Analysis Componentes principales PCT PE PE PE PE PE Pensa PF PF PE Pensa PF PF PE Pensa Persial Regresión por mínimos cuadrados Regresión por mínimos cuadrados Regresión por mínimos cuadrados Regresión por mínimos cuadrados PLS-DA Partial Least Squares Poiscriminant Analysis POL POL PRE Polimorfismo PRE Precisión PRM Pattern Recognition Method RMSEC Root Mean Square Error of Cross Validation RMSECV Root Mean Square Error of Cross Validation ROC Relative Operating Characteristics RPN Risk Priority Number Risk Priority Number Risk Priority Number Risk Priority Number RSD Relative Standard Deviation Severity Severidad SEL Selectivity Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay Polimorfismo Precisión Error cuadrático medio de calibración cruzada Error cuadrático medio de predicción cruzada Error cuadrático medio de predicción cruzada Error cuadrático medio de predicción cruzada RMSECP Root Mean Square Error of Prediction ROC Relative Operating Characteristics Características operativas relativas RNN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy Standard Normal Variate Negativo verdadero TP True Negative Negativo verdadero TP True Positive Positivo verdadero TP True Positive Positivo verdadero TP True Positive Positivo verdadero TP Tramaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Ultraviolet Ultravisible VAL - Conjunto de validación Vormalización vectorial	PCA	Principal Component Analysis	Análisis por componentes principales
PCT - Paracetamol Pensa Pensa Pensa Pensa Pensa Pensa Pensa Personant Analysis Pensa	DCA I DA	Principal Component Analysis-Linear	Análisis lineal discriminante por
PCT PE - Pensa PF - Pensa PFLS Partial Least Squares Regresión por mínimos cuadrados PLS-DA Partial Least Squares-Discriminant Análysis discriminante por mínimos cuadrados POL - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Cross Validation Error cuadrático medio de calibración RMSEP Root Mean Square Error of Prediction Cruzada RMSEP Root Mean Square Error of Prediction PROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Negative verdadero TP True Positive Positive Positivo verdadero TP True Positive Positive Positivo verdadero TP True Positive Positive Verdadero TP True Positive Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido UN Ultraviolet Ultravisible VAL - Conjunto de validación Verdadero Variabilidad	PCA-LDA	Discriminant Analysis	componentes principales
PF - PLS Partial Least Squares Regresión por mínimos cuadrados PLS-DA Partial Least Squares-Discriminant Análisis discriminante por mínimos cuadrados POL - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Calibration Calibration Calibration Calibration Cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de calibración Cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de predicción ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - Soft Independent Modelling of Class Analogy Modelado suave independiente por analogía de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TP True Positive Positivo verdadero TP True Positive Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAAL - Conjunto de validación Variale Variablidad VAR - Variabilidad	PCT	-	Paracetamol
PLS Partial Least Squares Regresión por mínimos cuadrados PLS-DA Partial Least Squares-Discriminant Analysis cuadrados cuadrados POL - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Calibration Error cuadrático medio de calibración RMSECV Root Mean Square Error of Cross Validation cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de predicción ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa SEL Selectivity Selectividad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TP True Positive Positivo verdadero TP True Positive Positivo verdadero TP Tramaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAAL - Conjunto de validación Vormalización vectorial	PE	-	Pensa
PLS-DA Partial Least Squares-Discriminant Análisis discriminante por mínimos cuadrados POL - Polimorfismo PRE - Precisión RMSEC Root Mean Square Error of Calibration RMSEC Root Mean Square Error of Calibration RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy	PF	-	Pfizer
POL - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Calibration RMSECV Root Mean Square Error of Cross Validation RMSECV Root Mean Square Error of Cross Validation RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de predicción ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Selectividad SEL Selectivity Selectividad SEL Selectivity Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy Hodelado suave independiente por analogía de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TPR - Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad	PLS	Partial Least Squares	Regresión por mínimos cuadrados
POL - Polimorfismo PRE - Precisión PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Calibration RMSECV Root Mean Square Error of Cross Validation RMSECV Root Mean Square Error of Cross Validation RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de predicción ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Selectividad SEL Selectivity Selectividad SEL Selectivity Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy Hodelado suave independiente por analogía de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TPR - Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad	DICDA		
PRE PRM Pattern Recognition Method Método de reconocimiento de pautas RMSEC Root Mean Square Error of Calibration RMSECV Root Mean Square Error of Cross Pror cuadrático medio de validación cruzada RMSECV Validation Cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de validación cruzada RMSEP Root Mean Square Error of Prediction Error cuadrático medio de predicción ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Selectividad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy Modelado suave independiente por analogía de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Ultravisible VAL - Conjunto de validación VAR - Variabilidad VAR - Variabilidad Normalización vectorial	PLS-DA	_	cuadrados
PRMPattern Recognition MethodMétodo de reconocimiento de pautasRMSECRoot Mean Square Error of CalibrationError cuadrático medio de calibraciónRMSECVRoot Mean Square Error of Cross ValidationError cuadrático medio de validación cruzadaRMSEPRoot Mean Square Error of PredictionError cuadrático medio de predicciónROCRelative Operating CharacteristicsCaracterísticas operativas relativasRPNRisk Priority NumberNúmero de prioridad del riesgoRSDRelative Standard DeviationDesviación estándar relativaSSeveritySeveridadSELSelectivitySelectividadSENSensibilitySensibilidadSGSavitzky-Golay-SIMCASoft Independent Modelling of Class AnalogyModelado suave independiente por analogía de clasesSNVStandard Normal VariateVariable normal estándarTNTrue NegativeNegativo verdaderoTPTrue PositivePositivo verdaderoTPR-Tamaño de partículaTRA-TramadolUCContent UniformityUniformidad de contenido de dosis unitariasUVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	POL	-	Polimorfismo
RMSEC Root Mean Square Error of Cross Validation RMSEP Root Mean Square Error of Cross Validation ROC Relative Operating Characteristics RSD Relative Standard Deviation Severity Selectivity Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay SNV Standard Normal Variate TN True Negative TPP True Positive TRA - Tramadol UC Content Uniformity Uniformidad de contenido de validación cruzada Error cuadrático medio de validación cruzada Error cuadrático medio de validación cruzada Error cuadrático medio de validación Características operativas relativas Características operativas relativas Error cuadrático medio de validación Error cuadrático medio de validación Características operativas relativas Error cuadrático medio de validación	PRE	-	Precisión
RMSECV Root Mean Square Error of Cross Validation RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics RPN Risk Priority Number RSD Relative Standard Deviation Severity Selectividad SEN Sensibility Sensibility Sensibilidad SG Savitzky-Golay SNV Standard Normal Variate TN True Negative TP True Positive Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformity Dosages Units Uniformidad de calibración Error cuadrático medio de validación cruzada Error cuadrático medio de validación cruzada Error cuadrático medio de validación cruzada Error cuadrático medio de validación Características operativas relativas Error cuadrático medio de predicción Error cuadrático medio de predicción Características operativas relativas Error cuadrático medio de predicción Características operativas relativas Error cuadrático medio de predicción Error cuadrático medio de predicción Características operativas relativas Error cuadrático medio de perdicción Error cuadrático medio de perdicción Características operativas relativas Error cuadrático medio de perdicción Error cuadrático medio de perdicción Error cuadrático medio de perdicción Normalización vectorial	PRM	Pattern Recognition Method	Método de reconocimiento de pautas
RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad Vormalización vectorial	RMSEC	1	Error cuadrático medio de calibración
RMSEP Root Mean Square Error of Prediction ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad Vormalización vectorial	DMCECV	Root Mean Square Error of Cross	Error cuadrático medio de validación
ROC Relative Operating Characteristics Características operativas relativas RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy Modelado suave independiente por analogía de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	KIVISEC V	Validation	cruzada
RPN Risk Priority Number Número de prioridad del riesgo RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	RMSEP	Root Mean Square Error of Prediction	Error cuadrático medio de predicción
RSD Relative Standard Deviation Desviación estándar relativa S Severity Severidad SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	ROC	Relative Operating Characteristics	Características operativas relativas
SEL Selectivity Selectividad SEN Sensibility Sensibilidad SG Savitzky-Golay - SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	RPN	Risk Priority Number	Número de prioridad del riesgo
SELSelectivitySelectividadSENSensibilitySensibilidadSGSavitzky-Golay-SIMCASoft Independent Modelling of Class AnalogyModelado suave independiente por analogía de clasesSNVStandard Normal VariateVariable normal estándarTNTrue NegativeNegativo verdaderoTPTrue PositivePositivo verdaderoTPR-Tamaño de partículaTRA-TramadolUCContent UniformityUniformidad de contenidoUDUUniformity Dosages UnitsUniformidad de contenido de dosis unitariasUVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	RSD	Relative Standard Deviation	Desviación estándar relativa
SENSensibilitySensibilidadSGSavitzky-Golay-SIMCASoft Independent Modelling of Class AnalogyModelado suave independiente por analogía de clasesSNVStandard Normal VariateVariable normal estándarTNTrue NegativeNegativo verdaderoTPTrue PositivePositivo verdaderoTPR-Tamaño de partículaTRA-TramadolUCContent UniformityUniformidad de contenidoUDUUniformity Dosages UnitsUniformidad de contenido de dosis unitariasUVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	S	Severity	Severidad
SG Savitzky-Golay SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	SEL	Selectivity	Selectividad
SIMCA Soft Independent Modelling of Class Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	SEN	Sensibility	Sensibilidad
SINICA Analogy de clases SNV Standard Normal Variate Variable normal estándar TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units UV Ultraviolet VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	SG	Savitzky-Golay	-
TN True Negative Negativo verdadero TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	SIMCA		
TP True Positive Positivo verdadero TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	SNV	Standard Normal Variate	Variable normal estándar
TPR - Tamaño de partícula TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	TN	True Negative	Negativo verdadero
TRA - Tramadol UC Content Uniformity Uniformidad de contenido UDU Uniformity Dosages Units Uniformidad de contenido de dosis unitarias UV Ultraviolet Ultravisible VAL - Conjunto de validación VAR - Variabilidad VN Vector normalization Normalización vectorial	TP	True Positive	Positivo verdadero
UCContent UniformityUniformidad de contenidoUDUUniformity Dosages UnitsUniformidad de contenido de dosis unitariasUVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	TPR	-	Tamaño de partícula
UDUUniformity Dosages UnitsUniformidad de contenido de dosis unitariasUVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	TRA	-	Tramadol
UVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	UC	Content Uniformity	Uniformidad de contenido
UVUltravioletUltravisibleVAL-Conjunto de validaciónVAR-VariabilidadVNVector normalizationNormalización vectorial	UDU		Uniformidad de contenido de dosis unitarias
VAR - Variabilidad VN Vector normalization Normalización vectorial	UV		Ultravisible
VAR - Variabilidad VN Vector normalization Normalización vectorial	VAL	-	Conjunto de validación
VN Vector normalization Normalización vectorial		-	
	VN	Vector normalization	Normalización vectorial
	W	Weight	Peso

ÍNDICE

I.	INTR	ODUCCIÓN	15
1.	INTROI	DUCCIÓN	16
2.		ROSCOPIA DE ABSORCIÓN EN EL INFRARROJO CERCANO	
		oducción	
		damento teórico	
		rumentación	
	2.3.1.	Fuente de radiación	21
	2.3.2.	Selector de longitud de onda	22
	2.3.3.	Compartimento de la muestra	
	2.3.4.	Detector	
	2.4. Ava	nces en la Espectroscopia NIR	24
3.	ESPECT	ROSCOPIA RAMAN	25
	3.1. Prin	cipios básicos de la Espectroscopia Raman	25
	3.1.1.	Teoría clásica del efecto Raman	28
	3.2. Inst	rumentación	30
	3.2.1.	Láseres	31
	3.2.2.	Selector de longitud de onda	33
	3.2.2.	1. Espectrómetros Raman dispersivos	33
	3.2.2.	2. Espectrómetros FT-Raman	34
	3.2.3.	Detectores	35
	3.3. Ava	nces en la Espectroscopia Raman	36
	3.3.1.	Instrumentos portátiles	37
	3.4. Apl	icaciones de la espectroscopia Raman	37
4.	QUIMIC	OMETRIA	39
	4.1. Pers	pectiva histórica	40
	4.2. Pret	ratamientos espectrales	40
	4.2.1.	Corrección del efecto de la dispersión	41
	4.2.2.	Aplicación de derivadas	44
	4.3. Aná	lisis cualitativo	45
	4.3.1.	Medida de la similitud	45
	4.3.2.	Análisis de las componentes principales	47
	4.3.3.	Métodos discriminantes	48

4.3.4.	Métodos de modelado	50
4.3.5.	Parámetros para evaluar el rendimiento de la clasificación	52
4.4. Aná	lisis cuantitativo	54
4.4.1.	Calibración univariante.	54
4.4.2.	Calibración multivariante	55
4.5. Con	strucción de un modelo quimiométrico	58
4.5.1.	Recopilación de datos	59
4.5.2.	Calibración	62
4.5.3.	Validación	64
4.5.4.	Mantenimiento	65
5. BIBLIO	GRAFÍA	66
II. OBJE	TIVOS	85
1. OBJETI	VOS	86
III. MAT	ERIALES Y MÉTODOS	87
1. INSTRU	MENTACIÓN	88
1.1. Espe	ectroscopia Raman	88
1.1.1.	Espectrómetro MultiRaman	88
1.1.2.	Espectrómetro i-Raman ® Plus	89
1.2. Espe	ectroscopia NIR	90
1.2.1.	Espectrómetro Model 5000	90
1.2.2.	Espectrofotómetro MicroNIR	91
1.2.3.	Espectrómetro MPA-NIR	91
1.3. Técr	nica instrumental de referencia. HPLC.	92
1.4. Otro	s instrumentos	92
2. REACTI	VOS	93
3. MÉTOD	OS	94
3.1. Dete	erminación de TRA y PCT en comprimidos	94
3.2. Dete	erminación de IBU en geles	95
4. PROGRA	AMAS PARA LA ADQUISICIÓN Y TRATAMIENTO DE LOS DATOS .	96
5. BIBLIO	GRAFIA	97
IV. IDEN	TIFICACIÓN DE PRODUCTO ACABADO CON	
INSTRUM	ENTOS PORTÁTILES	99

	11 1 1	IROD	UCCIÓN	100
2.	MA	ATER	ALES	102
	2.1.	Mue	stras	102
	2.2.	Adqı	uisición de espectros	105
3.	RE	SULT	ADOS Y DISCUSIÓN	105
	3.1.	Anál	isis exploratorio de los datos	105
	3.2.	Mod	elos construidos con los datos NIR	106
	3.	2.1.	Construcción de los modelos NIR	108
	3.	2.2.	Validación de los modelos NIR construidos	111
	3.3.	Mod	elos construidos con los datos Raman	113
	3.	3.1.	Construcción de los modelos Raman	117
	3.	3.2.	Validación de los modelos Raman construidos	120
	3.4.	Cons	strucción de una biblioteca de espectros	121
4.	CO	NCLU	JSIONES	124
5.	BIE	BLIO	GRAFÍA	124
\mathbf{V}	C	ONS	TRUCCIÓN DE BIBLIOTECAS NIR PARA LA	
*			-	
TT	TIME	TIDI	CACIÓN DE MATEDIAI ES DE DADTIDA	127
II)EN	TIFI	CACIÓN DE MATERIALES DE PARTIDA	127
II	INT	ΓROD	UCCIÓN	128
	INT	ΓROD		128
1.	INT	ΓROD ATERI	UCCIÓN	128
1.	INT	ΓROD ATERI Mue	UCCIÓNIALES Y MÉTODOS	128
1.	IN7 MA 2.1.	TROD ATERI Mue Adqı	UCCIÓNIALES Y MÉTODOSstras	
1.	INT MA 2.1. 2.2. 2.3.	TROD ATERI Mue Adqı Softv	UCCIÓN	
1.	INT MA 2.1. 2.2. 2.3. 2.4.	TROD ATERI Mue Adqı Softv	UCCIÓN	
1.	INT MA 2.1. 2.2. 2.3. 2.4. 2.	TROD ATERI Mue Adqı Softv Desa	UCCIÓN	
1.	INT MA 2.1. 2.2. 2.3. 2.4. 2.	TROD ATERI Mue Adqu Softv Desa 4.1.	ALES Y MÉTODOS	
1.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2.	TROD ATERI Mue Adqu Softv Desa 4.1.	ALES Y MÉTODOS	
1.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2.	Mue Adqu Softv Desa 4.1. 4.2.	ALES Y MÉTODOS	
1.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2. 2.	Mue Adque Software A.1. 4.2. 4.3. 4.4. 4.5.	ALES Y MÉTODOS	
1. 2.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2. 2. RE	Mue Adqu Softv Desa 4.1. 4.2. 4.3. 4.4. 5. SULT	UCCIÓN ALES Y MÉTODOS stras disición de espectros vare rrollo de la biblioteca de espectros Optimización de los conjuntos de aprendizaje y validación Construcción de la biblioteca Validación interna Criterio de aceptación Validación externa	
1. 2.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2. 2. RES	Mue Adqu Softv Desa 4.1. 4.2. 4.3. 4.4. 5. SULT	UCCIÓN	
1. 2.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2. 2. RES 3.1. 3.	Mue Adqu Softv Desa 4.1. 4.2. 4.3. 4.4. 5. SULT Bibli	UCCIÓN ALES Y MÉTODOS. stras disición de espectros ware crollo de la biblioteca de espectros Optimización de los conjuntos de aprendizaje y validación Construcción de la biblioteca Validación interna Criterio de aceptación Validación externa ADOS Y DISCUSIÓN oteca Específica	
1. 2.	INT MA 2.1. 2.2. 2.3. 2.4. 2. 2. 2. RES 3.1. 3.	Mue Adqu Softv Desa 4.1. 4.2. 4.3. 4.4. 5. SULT Bibli 1.1.	ALES Y MÉTODOS	

	3.1.4.	Validación interna	142
	3.1.5.	Evaluación de la actualización	143
	3.2. Bibl	ioteca General	144
	3.2.1.	Histórico de resultados OOS	144
	3.2.2.	Construcción de la Biblioteca General	144
	3.2.3.	Problemas planteados durante la construcción de la biblioteca	146
	3.2.4.	Validación interna	150
	3.2.5.	Validación externa	151
	3.2.6.	Evaluación de la actualización	151
4.	CONCL	USIONES	152
5.	BIBLIO	GRAFÍA	152
V	I. ANÁI	LISIS DE RIESGO Y MANTENIMIENTO DE BIBLIOTE	CAS
		A MATERIALES DE PARTIDA	
Ι¶.			
1.		DUCCIÓN	
2.		IALES Y MÉTODOS	
3.		ACIÓN DEL ANALISIS DE RIESGOS	
	3.1. Valo	oración del riesgo	159
	3.1.1.	Evaluación del riesgo para los factores "Personal", "Instrumento" y	
	"Condi	ciones Ambientales"	
	3.1.2.	Evaluación del riesgo para el factor "Modelo de Identificación"	161
	3.1.3.	Evaluación del riesgo para el factor "Propiedades del MP"	164
		trol del riesgo	
		nprobación del riesgo	
4.	RESULT	TADOS Y DISCUSIÓN	166
	4.1. Aná	lisis de riesgo para la Biblioteca Específica	166
	4.1.1.	Evaluación del riesgo para los factores "Personal", "Instrumento" y	
	"Condi	ciones Ambientales"	
	4.1.2.	Evaluación del riesgo para el factor "Propiedades del MP"	167
	4.1.3.	Evaluación del riesgo para el factor "Modelo de Identificación"	170
	4.1.4.	Conclusiones del análisis de riesgo para la Biblioteca Específica	171
	4.2. Rev	alidación de la Biblioteca Específica	171
	4.2.1.	Construcción de la Biblioteca Específica revalidada	171
	4.2.2.	Validación interna	174

	4.2.3.	Validación externa	174
	4.2.4.	Evaluación de la acción mitigadora realizada	175
	4.3. Anál	lisis de riesgo de la Biblioteca General	179
	4.3.1.	Evaluación del riesgo asociado a los factores: "Propiedades del MP" y	
	"Model	o de Identificación"	179
	4.3.2.	Conclusiones del análisis de riesgos de la Biblioteca General	181
	4.4. Cons	strucción de las nuevas Bibliotecas Generales para APIs y EXCs	181
	4.4.1.	Problemas planteados durante la construcción de la biblioteca	182
	4.4.2.	Validación interna	183
	4.4.3.	Validación externa	184
	4.4.4.	Evaluación de la acción mitigadora realizada	184
5.	CONCLU	USIONS	185
6.	BIBLIO	GRAFÍA	186
V	II. DETE	RMINACIÓN DEL CONTENIDO DE IBUPROFENO EN	
		OMERCIALES	. 189
1.		OUCCIÓN	
2.		IALES Y MÉTODOS	
		stras	
2	-	uisición de espectros	
3.		ADOS Y DISCUSIÓN	
		cción del pretratamiento	
		mización de las condiciones hidrodinámicas	
		ño de experimentos	
		dio de la repetibilidad	
1		ultados de la adición estándar	
4. 5.		GRAFÍAGRAFÍA	
٤.			
V	III. I	ESTUDIO DE LA DISTRIBUCIÓN DEL CONTENIDO DE	API
E	N COMP	RIMIDOS UTILIZANDO ESPECTROSCOPIA RAMAN	209
1.	INTROD	OUCCIÓN	210
		IALES Y MÉTODOS	
		stras	

	2.2. Inst	rumento y software	217
	2.3. Adq	uisición de los espectros	217
3.	RESUL	TADOS Y DISCUSIÓN	218
	3.1. Res	ultados del DoE	219
	3.2. Pret	ratamiento espectral	221
	3.3. Dist	ribución del API	221
	3.3.1.	Estudio en la cara posterior y núcleo	225
	3.3.2.	Uniformidad de las unidades de dosificación	226
	3.3.3.	Estudio del número mínimo de espectros en unidades individuales	227
	3.4. Dist	ribución del API en las mitades	230
	3.4.1.	Uniformidad de las unidades de dosificación para mitades	230
	3.4.2.	Estudio del número mínimo de espectros en las mitades	231
4.	CONCL	USIONES	234
5.	BIBLIO	GRAFÍA	234
D	X. DETI	ERMINACIÓN DEL CONTENIDO DE DOS APIS EN	
		MIDOS MEDIANTE NIR	2/1
1.		DUCCIÓN	
2.		JALES Y MÉTODOS	
		estras	
		nprimidos preparados en el laboratorio	
	2.3. Ado	uisición de espectros	246
		lisis de los datos	
3.	RESUL	ΓADOS Y DISCUSIÓN	246
	3.1. Mod	delos PLS construidos usando el mismo intervalo de longitudes de onda ([1100 –
	1646 nm)		248
	3.1.1.	Construcción de los modelos	248
	3.1.2.	Validación de los modelos	251
	3.2. Mod	delos individuales construidos para ambos APIs y espectrofotómetros	252
	3.2.1.	Construcción de los modelos	252
	3.2.2.	Validación de los modelos	253
	3.2.2.		
	3.2.2.		
4.	CONCL	USIONES	256

5.	BIBLIOGRAFÍA	257
X.	CONCLUSIONES2	61
1.	CONCLUSIONES	262

I.	<u>INTRODUCCIÓN</u>	

1. INTRODUCCIÓN

El sector de la industria farmacéutica evoluciona constantemente implementando tecnologías innovadoras con el objetivo del control global del proceso de fabricación. Una de las principales guías con pautas para la llegar a este objetivo es la publicada por la FDA sobre la tecnología analítica de procesos (PAT, por sus siglas *Process Analytical Technology*) [1]. Esta guía intenta suministrar herramientas para mejorar el conocimiento del proceso de fabricación y controlar la calidad de los materiales de partida, intermedios y producto acabado a tiempo real. De esta forma, se podría establecer una estrategia de control del proceso de fabricación de un producto farmacéutico. Algunas de las herramientas más destacadas para lograr entender el proceso de fabricación implican el uso de tecnologías rápidas de análisis, acopladas a técnicas de tratamiento de gran cantidad de datos que permitan extraer información. Idealmente, con la aplicación de tecnología PAT se podría lograr la adquisición de datos entiempo real del proceso estudiado.

Las técnicas espectroscópicas vibracionales son una buena opción para la implementación de la tecnología PAT y se han convertido en un buen aliado para la industria farmacéutica, tanto en el proceso de fabricación como en el control de calidad de los productos. Estas técnicas presentan ciertas ventajas, como una reducción del tiempo de análisis y de preparación de las muestras, obtención de resultados muy rápidos, el ser no destructivas y que son sostenibles desde el punto de vista medioambiental (no es necesario el uso de disolventes, por ejemplo). La aplicación de estas técnicas permite lograr una mejora en el conocimiento del proceso, una mayor robustez y flexibilidad [2]. Las principales aplicaciones se basan en la identificación de los materiales de partida y producto acabado, la caracterización de formas polimórficas y la cuantificación de los APIs y excipientes de la formulación. Sin embargo, gracias al avance en la miniaturización de los instrumentos, se han empezado a usar estas técnicas para estudiar distintas etapas de un proceso de fabricación como la homogeneización, la granulación, el secado, el recubrimiento y la compresión, en caso de productos sólidos [3]. El uso de estas técnicas espectroscópicas solo es posible si se combina con el uso de la Quimiometria, que es la herramienta que permite entender y manipular todos los datos espectroscópicos generados.

La aplicación de estas técnicas en la industria farmacéutica ya está bien asentada gracias a distintas guías publicadas para su uso. Por ejemplo, en 2014, la Agencia Europea del Medicamente (EMA, por sus siglas *European Medicine Agency*) publicó una guía sobre

el uso de la técnica NIR en la industria farmacéutica y los documentos necesarios para su aprobación [4]. Además, la FDA publicó otra guía para la implementación de la técnica NIR en la industria farmacéutica [5]. Por otro lado, la Farmacopea Europea (EP) ha incluido monografías para estas técnicas, además de una monografía dedicada exclusivamente a la Quimiometria [6–8].

En las siguientes secciones de la introducción, se explicará extensamente las técnicas espectroscópicas utilizadas en esta tesis, NIR y Raman, así como las técnicas quimiométricas que se han utilizado para el tratamiento de los datos obtenidos.

2. ESPECTROSCOPIA DE ABSORCIÓN EN EL INFRARROJO CERCANO

2.1. Introducción

El descubrimiento de la región del infrarrojo (IR) se atribuye al astrónomo William Herschel. Un día, del año 1800, W. Herschel se preguntó qué color de la zona del espectro visible propagaría el calor proveniente del sol. Usó un prisma de cristal para generar un espectro mientras media los cambios de temperatura de cada color con un termómetro. Durante el experimento, observó ligeros aumentos de la temperatura en distintas regiones del espectro, en función del color. Al salir a comer, se dejó el termómetro sobre la mesa, justo fuera del espectro, junto la banda roja. Al volver, la temperatura había aumentado mucho. En ese momento, postuló una banda invisible de luz más allá del rojo, que la denominó rayo de calor (*Heat Ray* en inglés) [9]. Fue en 1835 cuando se confirmó que esta nueva banda invisible tenía una longitud de onda mayor que la del visible y pasó a llamarse infrarrojo. La **Figura I. 1** muestra a W. Herschel junto a una representación del experimento que efectuó.

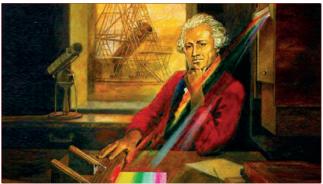


Figura I. 1. W. Herschel y el descubrimiento del infrarrojo. Reproducido de [10].

A pesar de que la radiación NIR se detectó antes que la radiación del infrarrojo medio (MIR por sus siglas en inglés *Mid-Infrared*), fue esta última la que ganó más importancia. David L. Wetzel, publicó un artículo titulado "*Near-Infrared Reflectance Analysis – Sleeper Among Spectroscopic Techniques*" en 1984 donde mencionaba las ventajas de esta técnica, así como sus posibles aplicaciones en la industria [11]. De este modo, la década de los 80s fue donde se vio impulsada la aplicabilidad de la espectroscopia NIR [12].

La **Figura I. 2** muestra algunos de los avances más importantes en la evolución de esta técnica.



Figura I. 2. Hitos destacados de la técnica NIR.

2.2. Fundamento teórico

La espectroscopia NIR se puede definir como una técnica de absorción de la radiación al interaccionar con la materia. Esta interacción produce cambios en los estados vibracionales de la molécula.

La mayoría de las moléculas se encuentran en el nivel fundamental a temperatura ambiente. Los átomos o grupos de átomos que participan en el enlace se desplazan juntamente con una frecuencia. Esta frecuencia está determinada por dos propiedades: la fuerza del enlace y la masa de cada uno de los átomos o grupos.

El modelo más simple para entender la interacción de la radiación y la materia en la zona espectral del NIR es mediante una molécula diatómica. Si consideramos la molécula como dos masas esféricas unidas mediante un muelle, que tendrá una determinada constante

de fuerza (k), la Ley de Hook establece que la energía de este sistema viene determinada por la **Ecuación I-1**, donde μ es la masa reducida de ambas masas [13–15].

$$E = \frac{h}{2\pi} \sqrt{\frac{k}{\mu}} \tag{I-1}$$

Mediante la suposición de un oscilador armónico se puede describir de una manera más simplificada la vibración molecular de un enlace. En el modelo armónico, la energía potencial (V) está directamente relacionada con la función del desplazamiento de los átomos según la **Ecuación I-2**.

$$V = \frac{1}{2}kx^2 \tag{I-2}$$

En la **Figura I. 3** se puede ver una representación del modelo armónico, donde se observa el comportamiento de la energía potencial como función del desplazamiento de los átomos. Como puede verse, esta relación es una función parabólica que se encuentra centrada en la distancia del equilibrio.

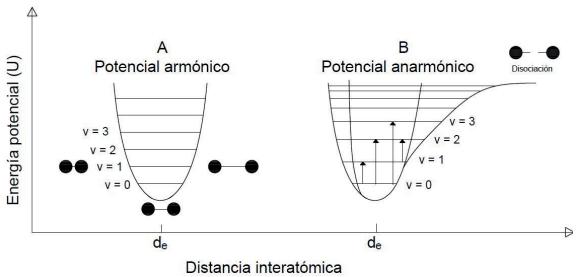


Figura I. 3. Energía potencial en función de la distancia interatómica para un modelo armónico.

Esta primera aproximación de un modelo armónico no tiene en consideración que un sistema molecular no puede estar regulado por un perfil de energía continua. El sistema solo puede tener algunos niveles de energía discretos (E_v) que se definen mediante la mecánica cuántica, según la **Ecuación I-3**, donde v representa el número cuántico vibracional, E_v la

energía asociada con el nivel cuántico vth, y v la frecuencia de la vibración fundamental, que se define mediante la **Ecuación I-4**.

$$E_{v} = \left(v + \frac{1}{2}\right)hv\tag{I-3}$$

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \tag{I-4}$$

En el modelo armónico, existe una restricción que indica que solo pueden ocurrir transiciones entre niveles adyacentes ($\Delta v = \pm 1$). Además, la diferencia de energía entre los niveles adyacentes es siempre la misma, como se aprecia en la **Figura I. 3**. Sin embargo, la energía de la radiación electromagnética que se absorbe al promover la excitación de la molécula a un nivel vibracional superior, debería ser la diferencia entre dos niveles adyacentes de energía (**Ecuación I-5**). En el oscilador armónico, como $\Delta v = 1$, la diferencia de energía entre niveles es siempre la misma.

$$\Delta E = E_{v2} - E_{v1} = \Delta v h v \tag{I-5}$$

La principal desventaja del modelo armónico es la restricción entre transiciones donde $\Delta v > 1$, de esta forma, la mayoría de las bandas observadas en la región NIR estarían prohibidas, como, por ejemplo, las correspondientes a los armónicos. Considerando esto, la **Figura I. 3B** muestra una situación más realista para una molécula diatómica, según un oscilador no armónico. A diferencia del modelo armónico, los niveles de energía ahora no son equidistantes y puede haber transiciones con $\Delta v > 1$, como por ejemplo $\Delta v = \pm 2, \pm 3, ...$ Estas transiciones que ahora están permitidas son las denominadas primero, segundo, y así sucesivamente, armónico. También se pueden observar bandas de combinación de distintas transiciones vibracionales. Sin embargo, la probabilidad de estas transiciones disminuye mucho y, por tanto, su intensidad. Las principales características de las bandas que aparecen en la región NIR son:

- (1) Bandas relacionadas con los armónicos y combinaciones, no sólo de niveles de energía adyacentes, sino que también combinaciones de segundo o tercer orden $(v_1 + 2v_2)$.
- (2) Difícil asignación dado el gran solapamiento de bandas.

- (3) La intensidad de las bandas disminuye al disminuir la longitud de onda, debido que en esta región aparecen las bandas de combinación de mayor orden. De esta forma, se puede dividir un espectro NIR en distintas secciones [16]:
 - Región del primer armónico y bandas de combinación (2000 2500 nm)
 - Región del segundo armónico y bandas de combinación (1100 2000 nm)
 - Región del tercer y cuarto armónico (700 1100 nm).
- (4) Muchas de las bandas en la región NIR son originadas por grupos funcionales que contienen un átomo de hidrogeno (OH, CH, NH). Esto está relacionado con la constante de anarmonicidad de un enlace XH.

En el espectro vibracional, pueden aparecer otras señales aparte de las ya comentadas relacionadas con los armónicos y bandas de combinación. Dos ejemplos son la resonancia de Fermi y la Darling-Dennison [14]. La de Fermi puede aparecer si dos niveles vibracionales pertenecen a la misma simetría de un producto y tienen una energía similar. Este efecto produce la aparición de dos bandas de absorción muy intensas a una frecuencia mayor o menor que la esperada [17,18]. Un ejemplo estudiado de este efecto se ha visto para el espectro del CO₂ [19,20]. En el caso de la resonancia de Darling-Denninson, tiene especial importancia en el espectro del agua, pero puede suceder para otras moléculas, en el caso de que tengan enlaces X-H simétricos equivalentes [14,21].

2.3. Instrumentación

Los instrumentos NIR pueden contener una gran variedad de dispositivos dependiendo de las características de la muestra que se va a analizar. Pero, esquemáticamente, un equipo NIR está formado por cinco elementos: una fuente de radiación, un selector de longitud de onda, un compartimento de contacto con la muestra, un detector y un ordenador o procesador de la señal.

2.3.1. Fuente de radiación

La fuente de radiación más popular es la que utiliza una radiación térmica generada al calentar un filamento mediante una carga de corriente eléctrica. Un ejemplo sería una lámpara halogenada que usa filamentos de tungsteno [13,14,22]. Otras de las fuentes de

radiación usadas son las LED. Su bajo consumo, precio, tamaño y larga vida hace que este segundo tipo de fuente sea usado en instrumentos NIR miniaturizados [23]. Además de estas dos fuentes, una más reciente son los láseres de diodo sintonizable.

2.3.2. Selector de longitud de onda

Los instrumentos NIR se pueden diferenciar según el selector de longitud de onda empleado: discreto o continuo. Los selectores discretos son los que irradian la muestra con un número determinado de longitudes de onda mediante filtros o LEDs. Los continuos incluyen una red de difracción, filtros acústo-ópticos (AOTF por sus siglas en ingles *Acousto-Optical Tuneable Filters*) o la instrumentación con Transforma de Fourier (FT) [24]. Algunas de las características principales de los selectores mencionados se presentan en la **Tabla I. 1** [14].

Sistema de selección Descripción Filtros de interferencia Estos filtros son de estructura sencilla, con una lámina fina dieléctrica que se coloca entre dos láminas semitransparentes. Filtros acusto-ópticos (AOTF) Capaz de realizar la selección de la longitud de onda mediante una modulación eléctrica externa. Modulan la intensidad a través de la interacción de las ondas generadas en un cristal de TiO₂ birrefringente.[25–27] Transformada de Fourier Mediante el uso de un interferómetro se modula la información de la fuente de radiación, permitiendo la adquisición simultánea de varias longitudes de onda. Monocromador Reparten la radiación en función de la longitud de onda seleccionada.

Tabla I. 1. Características de los principales selectores de longitud de onda.

2.3.3. Compartimento de la muestra

La espectroscopia NIR se puede usar para realizar medidas de distintos tipos de muestra como líquidos, sólidos, soluciones, gases, etc. Por ello, existen distintas formas de medida: reflectancia, transmitancia y transflectancia [28–31]. La selección del modo de adquisición depende de varios factores entre los cuales se encuentran las características de la muestra o la localización. La **Figura I. 4** muestra una representación de estas tres formas, donde la principal diferencia entre todas es la posición del detector [13,14,22,24].

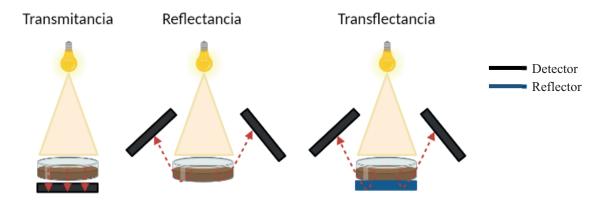


Figura I. 4. Modos de adquisición mediante la espectroscopia NIR.

En el modo de transmisión, el haz de luz se propaga a través de la muestra y se mide justo al salir de ésta, con el detector justo en el lado opuesto de la fuente de radiación. En el modo de reflectancia, el haz de luz y el detector se encuentran en el mismo lado de la muestra, pero los detectores se colocan a unos 45° para evitar la reflexión de la superficie. En algunos casos, también se incorpora una esfera integradora antes del detector para concentrar la luz reflejada de la muestra antes de su llegada al detector. La concentración de la luz se produce mediante la captura e integración de la luz de varias direcciones. En el caso del modo de transflectancia, se utiliza una combinación de los modos de transmitancia y reflectancia. La muestra se posiciona de igual manera que en el modo de reflectancia, pero se incluye un reflector en el lado opuesto de la fuente de radiación, justo detrás de la muestra. La función de este reflector es la de reflejar la luz, que es transmitida otra vez a través de la muestra para hacerla llegar al detector.

Dependiendo del modo de adquisición usado, la luz atenuada por la muestra se conoce como reflectancia o transmitancia. Sin embargo, normalmente estos parámetros se transforman a unidades de absorbancia aparente $(\log\left(\frac{1}{R}\right);\log\left(\frac{1}{T}\right))$.

2.3.4. Detector

Los detectores transforman la energía de la luz incidente en una señal eléctrica análoga, que se amplifica y transforma en digital y se procesa mediante un ordenador. Junto con los amplificadores, los detectores, se consideran las fuentes más comunes del ruido de los instrumentos, que se conoce como el ruido aleatorio. Para reducir al máximo este ruido aleatorio, la estrategia que se suele tomar es la realización de un promedio de espectros para una misma muestra, mejorando así la relación S/N [22,32].

Los principales detectores usados en la región espectral NIR son compuestos por materiales fotosensibles. En el rango de longitudes de onda de 400 – 1100 nm, los detectores más comunes son los de silicio (Si) [33]. Los detectores de sulfuro de plomo (PbS) o indiogalio-arseniuro (InGaAs) cubren hasta longitudes de onda mayores, 2500 nm para los PbS y 1700 nm para los InGaAs.

2.4. Avances en la Espectroscopia NIR

La instrumentación en la espectroscopia se puede dividir en espectrómetros de mesa, que son los que se pueden emprar solo en un laboratorio, y espectrómetros más autónomos que se pueden desplazar y usar *on-site*. Normalmente, este segundo tipo de espectrómetros se clasifican en transportables y de mano dependiendo de sus dimensiones y peso [14,34].

En la última década, gracias a la introducción de estos instrumentos miniaturizados, se ha abierto un gran abanico de posibilidad para la realización de aplicaciones innovativas de desarrollo y para la aplicación en la industria [34–38].

La **Figura I. 5** muestra un resumen de las publicaciones relacionadas con los instrumentos NIR hasta la actualidad. Como puede verse, su implementación ha ido aumentando con los años. Esto se debe principalmente a su mejora instrumental y a su gran abanico de posibilidad de utilización en distintos sectores farmacéuticos.

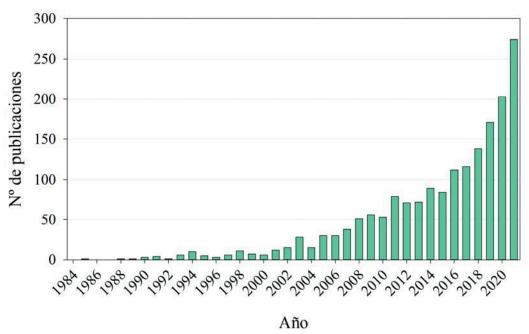


Figura I. 5. Evolución en el número de publicaciones que contienen como palabras clave NIR y Portátiles (*Portables*) o *Handheld* por año. Datos extraídos de la base de datos de *Scopus*.

3. ESPECTROSCOPIA RAMAN

En el 1921, durante un viaje por mar, el físico Indio Chandrasehara Venkata Raman, realizó algunos experimentos que publicó en forma de carta titulada "*The Colour of the Sea*" [39]. Los experimentos de C.V. Raman demostraron que el color del mar era un resultado directo de la dispersión molecular de la luz e independiente de la absorción de la reflexión de la luz del cielo [40]. Posteriormente, no mucho más tarde, volvió a publicar una carta donde explicaba la dispersión molecular de la luz en sólidos y líquidos. En base a estos hallazgos, junto con algunos colaboradores, empezaron una serie de experimentos acerca de la dispersión de la luz en distintas muestras de líquidos, así como teorías acerca del potencial de sus experimentos. Con todo esto, descubrieron el efecto de la dispersión inelástica, llamado Efecto Raman en 1928 [41]. Este hito le hizo ganar en 1930 el premio Nobel de Física.

El desarrollo y avance en la Espectroscopia Raman fue un proceso relativamente lento debido a diversos factores como la dificultad de adquisición, la fluorescencia y la descomposición de los productos. En 1970, gracias a la disponibilidad de nuevos láseres, esta técnica empezó a ganar importancia. Aún y así, seguía siendo complicado obtener un sistema óptico ajustado para una buena detección, por lo que se limitaron mucho las aplicaciones de la Espectroscopia Raman [42–44].

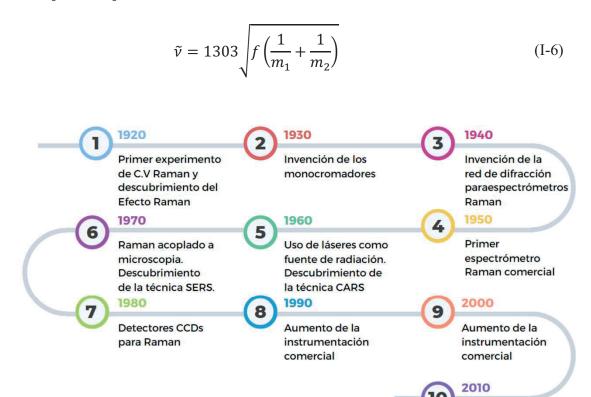
El cambio e inicio de la revolución que supuso la espectroscopia Raman llegó alrededor de los 90 gracias a la mejora de los láseres, de la tecnología de los detectores, del desarrollo de nuevos filtros ópticos, y en el perfeccionamiento de los programas de análisis de los datos.

La **Figura I. 6** muestra un esquema de la evolución de esta técnica con los años [44,45].

3.1. Principios básicos de la Espectroscopia Raman

Las moléculas están formadas por átomos de una determinada masa y que están conectados mediante enlaces elásticos, los cuales pueden realizar movimientos periódicos. Cada molécula tiene un número vibracional de grados de libertad. Por ejemplo, las moléculas poliatómicas con n átomos, contienen 3n-6 vibraciones normales que determinaran su espectro vibracional. El espectro dependerá también de la masa de los átomos, la fuerza de los enlaces y su geometría. La frecuencia de la vibración de una molécula diatómica viene

determinada por la **Ecuación I-6**, donde f es la constante de fuerza y m_1 , m_2 la masa de los átomos. [42,43,46]



Aumento de la aplicabilidad de los instrumentos portátiles

Figura I. 6. Avances en la espectroscopia Raman.

Según la mecánica cuántica, una molécula puede llegar a absorber una cantidad de energía (hv_s) para alcanzar el primer estado vibracional excitado. Si al irradiar una molécula ésta absorbe esa energía (hv_s) , el espectro presenta una banda de absorción a la frecuencia v_s . La dispersión Raman no exige que la energía de la radiación incidente sea igual a la diferencia de energía entre el estado fundamental y excitado, como en el caso de la absorción. Si la molécula es irradiada mediante una radiación monocromática de energía hv_0 , durante el impacto inelástico de la irradiación del haz de luz monocromática puede haber un intercambio de energía hv. Solo un pequeño número de fotones experimenta un cambio de energía tras la colisión con la molécula. Si el fotón incidente proporciona una cantidad de energía hv a la molécula, la energía del fotón dispersado se reduce a $h(v_0 - v)$. Por el contrario, si el fotón es el que recibe la energía de la molécula, la energía del fotón dispersado será de $h(v_0 + v)$. De esta forma, los fotones son dispersados con mayor o menor energía

según la **Ecuación I-7**. La dispersión en la que se produce un intercambio de energía entre el fotón y la molécula es lo que se denomina dispersión Raman y los valores $h\nu_R$ son la energía del espectro Raman.

$$h\nu_R = h\nu_0 \pm h\nu_s \tag{I-7}$$

Dependiendo si el fotón recibe o proporciona la energía, la dispersión se denomina, respectivamente, anti-Stokes o Stokes. La **Figura I. 7** muestra un esquema de la dispersión Raman y Rayleigh [43].

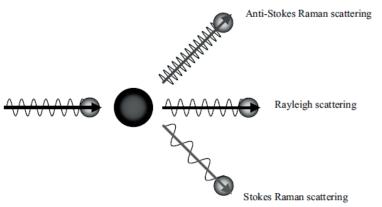


Figura I. 7. Representación de la dispersión Raman y Rayleigh. Reproducido de [43].

La gran mayoría de los fotones que chocan con la molécula no experimentan un impacto inelástico, sino un impacto elástico; dando lugar a una radiación que se denomina dispersión Rayleigh, con la misma energía que la de la luz incidente. Además de estos procesos, también puede suceder que haya moléculas irradiadas donde la energía de los fotones no sea suficientemente alta para permitir la transición de éstos al nivel electrónico excitado. En este caso, los fotones volverán al nivel fundamental con menor energía. Este proceso se conoce como fluorescencia y conlleva la obtención de un espectro Raman distorsionado [47].

La **Figura I. 8** muestra un resumen de estos tres procesos comentados, en el caso de absorción, dispersión y fluorescencia [48].

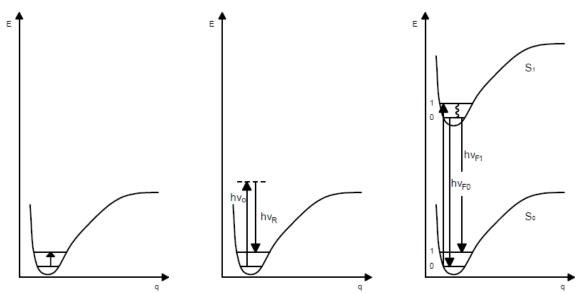


Figura I. 8. Observación de la excitación a un nivel vibracional desde el nivel fundamental (S₀) mediante la absorción IR, dispersión Raman, y la fluorescencia. Reproducido de [48].

3.1.1. Teoría clásica del efecto Raman

Para empezar con la descripción de la teoría clásica del efecto Raman, se debe considerar que cuando una molécula es irradiada por una onda electromagnética con frecuencia v_0 , el campo eléctrico oscilatorio, E, de la onda electromagnética produce un cambio en la distribución de los electrones de la molécula. Este hecho conlleva la inducción de un momento dipolar (μ) (**Ecuación I-8**) [49,50].

$$\mu = \alpha E \tag{I-8}$$

Donde el factor proporcional entre el campo eléctrico externo y el momento dipolar inducido es la polarizabilidad (α). Finalmente, tras distintas conversiones matemáticas, se acaba obteniendo la **Ecuación I-9** para el momento dipolar inducido. En esta ecuación, q representa el grupo de todos los modos normales individuales, después de realizar una expansión de la polarizabilidad mediante la serie de Taylor.

$$\mu(t) = \alpha_0 \cdot E_0 \cdot \cos(\nu_0 t) + \frac{1}{2} \left(\frac{\partial \alpha}{\partial q} \right)_{q_0} \cdot q_0 \cdot E_0 \cdot \cos[(\nu_0 - \nu_q)t]$$

$$+ \frac{1}{2} \left(\frac{\partial \alpha}{\partial q} \right)_{q_0} \cdot q_0 \cdot E_0 \cdot \cos[(\nu_0 + \nu_q)t]$$
(I-9)

Esta ecuación constituye el concepto de la dispersión Raman en el caso de una dispersión de luz general. El momento dipolar inducido dependiente del tiempo (t), contiene tres términos de frecuencia distintos:

- (1) El primer término, es el producto de la polarizabilidad y el campo eléctrico vibrando a la misma frecuencia, ν_0 , que la luz de incidencia. Por lo que corresponde a la dispersión Rayleigh.
- (2) El segundo término, oscila con diferencias entre el láser y la frecuencia del modo normal molecular. En este caso, se hace referencia a que la radiación emitida por el momento dipolar inducido se desplaza hacia el rojo con respecto a la frecuencia de excitación. Este término representa la dispersión inelástica y es la fuente de la dispersión Stokes, que se presenta normalmente en el espectro Raman.
- (3) El último término representa una fuente de radiación que se desplaza hacía el azul con respecto a la frecuencia del láser. Este caso es el denominado dispersión anti-Stokes, que tiene una intensidad mucho menor que la Stokes.

Con todo esto se puede ver como la **Ecuación I-9** revela algunos prerrequisitos que se deben cumplir para que la dispersión Raman suceda. Para que se produzca este efecto, el factor $\left(\frac{\partial \alpha}{\partial q}\right)_{q_0} \cdot q_0 \cdot E_0$ debe ser distinto a cero. Dado que ni q_0 ni E_0 son cero, la condición que se debe cumplir es que $\left(\frac{\partial \alpha}{\partial q}\right)_{q_0} \neq 0$. Por tanto, las vibraciones que cambien la polarizabilidad serán las denominadas "activas" en la espectroscopia Raman.

La polarizabilidad se puede definir como una forma de medir la deformidad de un enlace tras la irradiación de la molécula. Al irradiar una molécula, el campo eléctrico generado interactúa con los electrones del enlace entre ambos átomos, deformando así el enlace. Es por esto por lo que la polarizabilidad varía en función de la distancia entre los dos núcleos involucrados en el enlace. Por ejemplo, una molécula homonuclear no tiene un momento dipolar cuando se produce una vibración de estiramiento que varía la distancia entre los dos núcleos. Sin embargo, esta vibración sí que resulta en un cambio en la polarizabilidad del enlace entre los átomos. El valor de polarizabilidad será máximo cuando mayor sea la separación y mínimo cuando se encuentren más cercanos [51].

La principal diferencia entre ambos tipos de dispersión Raman es su intensidad, donde la anti-Stokes es mucho menor que la Stokes. Además, estos dos tipos de dispersión son mucho menos intensos que la dispersión Rayleigh. La intensidad de la radiación Raman dispersada, I_R , viene determinada por la **Ecuación I-10** [46].

$$I_R \propto v^4 I_0 N \left(\frac{\partial \alpha}{\partial O}\right)^2$$
 (I-10)

Donde I_0 es la intensidad del láser, N el número de moléculas dispersadas, v la frecuencia del láser, α la polarizabilidad y Q la amplitud de la vibración.

La teoría clásica de la espectroscopia Raman permite entender y obtener una visión de la dispersión Raman, pero no permite explicar todos los efectos, como la diferencia en la intensidad entre los dos tipos de dispersión: Stokes y anti-Stokes. En este caso, sería necesario realizar una profundización mediante la teoría cuántica del efecto Raman.

3.2. Instrumentación

Los instrumentos Raman se componen principalmente de cuatro elementos: la fuente de radiación (láser), configuración óptica, selector de longitud de onda y el detector.

La intensidad Raman depende de la potencia del láser usada para excitar la dispersión, de la polarizabilidad de la molécula y de la frecuencia del láser usada, tal y como se ha visto en la ecuación I-10 presentada anteriormente. Considerando que la polarizabilidad es una propiedad de la molécula, hay dos parámetros relacionados con el instrumento que sí se pueden seleccionar por el operario para obtener una mejor respuesta. Por ejemplo, la dispersión depende de la frecuencia del láser, por tanto, para mejorar la sensibilidad se debe usar la mayor frecuencia posible, lo que supondría trabajar en la región UV. Sin embargo, normalmente no se trabaja en esta región dado que muchos productos absorben la radiación UV y, existe un riesgo de degradación de la muestra [42].

En la espectroscopia Raman existen principalmente dos configuraciones ópticas para la colección de la luz, la configuración de 90° y 180° [52]. En la configuración de 90°, la luz dispersada es recogida en la dirección perpendicular de la excitación, En el caso de la de 180°, la colección se realiza en la dirección opuesta.

3.2.1. Láseres

Esta parte del instrumento Raman es la fuente para la generación de la dispersión Raman. Los avances en la tecnología y la mejora en componentes, permite que hoy en día haya una gran variabilidad de tipos de láser.

El primer láser que se usó en el primer instrumento Raman comercial fue el láser HeNe rojo de 632,8 nm en ondas continuas en el 1963 [53,54]. Sin embargo, este láser fue rápidamente sustituido por unos con mayor potencia (Ar⁺ y Kr⁺) que proporcionaban numerosas líneas de excitación en la región de 350 – 700 nm (región del visible y UV) sobre los 70s. En los 80s, llegaron los láseres de estados sólidos, como por ejemplo el YAG a 1064 nm, que fueron una revolución para la aplicación tanto de la espectroscopia Raman como NIR.

Generalmente, los instrumentos Raman que usan láseres con longitud de onda menor a 750 nm suelen ser dispersivos, mientras que los que usan una mayor longitud de onda se clasifican como instrumentos Raman de transformada de Fourier (FT). El rango de longitud de onda para los láseres Raman va desde la región UV (< 200 nm) hasta la región NIR (1064 nm). Las longitudes de onda de los láseres más frecuentes se presentan en la **Tabla I. 2** [43].

Tipo de láser	Longitud de onda (nm)
HeNe	632,8
$\mathrm{Kr}^{\scriptscriptstyle +}$	413,1 y 647,1
Ar^{+}	488,0 y 514,5
Diodo	660 - 880
Nd:YAG	1064
Doble frequencia Nd:YAG	532

Tabla I. 2. Láseres más frecuentes usados en la instrumentación Raman.

De la **Tabla I. 2** se puede ver que existen dos grupos de láseres dependiendo de la longitud de onda de excitación: en la región NIR o la región del visible. Además de estas opciones presentadas, se están empezando a usar también láseres a longitudes de onda 1280 y 1550 nm [42].

Actualmente, uno de los parámetros más importantes a considerar a la hora de seleccionar la frecuencia del láser es si la muestra puede presentar fluorescencia. En caso afirmativo, una de las mejores opciones es usar un láser diodo (780 nm) o un láser a mayor longitud de onda, como sería el Nd:YAG (1064 nm). En cambio, en el caso de que la muestra

no presente fluorescencia, la mejor opción es usar láseres de menor longitud de onda, por ejemplo, en el rango de 470 a 650 nm.

En el caso de los láseres con excitación en la región del visible, estos proporcionan una mejor relación señal-ruido (S/N) en el espectro Raman y están acoplados a detectores más económicos, dado el gran flujo de fotones que simplifican su detección. Sin embargo, la única desventaja de este grupo de láseres es que la excitación de la fluorescencia puede interferir en la señal Raman y que las muestran se pueden llegar a quemar si no se usa una potencia adecuada. En cambio, los láseres con excitación en la región NIR, proporcionan una señal Raman más débil al detector, por lo que se necesita que este sea más sofisticado y optimizado. Este tipo de láseres se usan para experimentos con muestra biológicas dada la gran probabilidad de fluorescencia de este tipo de muestras.

La **Tabla I. 3** presenta algunos de los láseres comerciales más usados, así como sus principales aplicaciones.

Tabla I. 3. Láseres comerciales usados en	los espectrómetros Raman. Reproducida de [55]	51.
---	---	-----

Fuente de excitación	Tipo de láser y longitud de onda	Técnicas Raman	Aplicaciones c
NIR	 Láser diodo: 785, 830 nm Láser estado sólido: Nd:YAG (1064 nm) 	Espectroscopia Raman FT Espectroscopia Raman Espectroscopia Raman mejorada de superficie	Muestras biológicas, polímeros y objetivos generales
Visible	 Láser ion: He-Ne (633 nm), He-Cd (442 nm), Ar⁺ (488, 514 nm) Láser estado sólido: Nd:YAG (532 nm) 	Espectroscopia Raman normal Espectroscopia Raman mejorada de superficie Time- resolved Raman spectroscopy Resonance Raman spectroscopy	Compuestos orgánicos, arte, arqueología, forenses, semiconductores, minerales y otros objetivos
UV	• Láser ion: He-Cd (325 nm), Ar ⁺ (244, 257 nm)	Espectroscopia Raman-UV Resonance Raman spectroscopy Time-resolved Raman spectroscopy	Proteínas y semiconductores

3.2.2. Selector de longitud de onda

La principal función del espectrómetro es realizar la separación de la luz dispersada según la longitud de onda para la obtención del espectro Raman y rechazar así la dispersión Rayleigh, que es mucho más intensa. Aproximadamente, la intensidad de la dispersión Rayleigh es más de un millón de veces la intensidad Raman [53]. Teniendo en cuenta esta gran diferencia en la intensidad entre ambos tipos de dispersión, es de gran importancia atenuar al máximo la dispersión Rayleigh antes de su llegada al espectrógrafo (medidor de la longitud de onda). Si esta dispersión no se viera reducida, enmascararía la dispersión Raman, interfiriendo en el espectro Raman obtenido.

Actualmente, existe una gran variedad de métodos para la selección de la longitud de onda, pero la gran mayoría de los instrumentos Raman usan espectrómetros dispersivos o de transformada de Fourier (FT). En el contexto histórico, los espectrómetros dispersivos fueron desarrollados primero usando detectores de canal simple fotomultiplicadores (single-channel photoomultiplier PMT), donde se median los fotones dispersados en una única longitud de onda. En cambio, los espectrómetros FT-Raman se desarrollaron a finales de 1986 debido a limitaciones tecnológicas [56,57].

3.2.2.1. Espectrómetros Raman dispersivos

Los elementos principales de un espectrómetro dispersivo son la rendija de apertura, espejo colimador, la red de difracción, el espejo de enfoque y la rendija de salida. La **Figura I. 9** muestra un esquema de este tipo de espectrómetro. La luz entra por la rendija de apertura, pasando por el espejo colimador, así la luz es colimada y direccionada a la red de difracción. Después de la separación, el haz de difracción se enfoca hacía la rendija de salida mediante el espejo de enfoque.

Su principal funcionalidad es realizar la separación espacial de los fotones con distinta energía que provienen de la muestra. El grado de separación (o dispersión), es una función del espacio de la ranura en la red de difracción. Cuanto más estrecha sea la ranura, mayor será la dispersión y la ventana espectral será más estrecha.

Además, como se ha comentado anteriormente, la dispersión Raman es mucho menor que la Rayleigh por lo que se necesita un mecanismo para atenuar esta dispersión. Este proceso se conoce como el rechazo de la luz en un espectrómetro.

Los principales filtros que se usan para obtener esta atenuación son los *notch filters*, que remueven la radiación de la misma frecuencia que el láser, y los *Edge filters*, que remueven la luz a una determinada frecuencia para que solo la dispersión Stokes y otra a una menor energía, como la fluorescencia, sean transmitidas [58–60]. Ambos filtros se han usado para atenuar la luz Rayleigh en un monocromador simple. Sin embargo, también es posible atenuar esta luz mediante un monocromador múltiple [61]. Como ejemplo, el uso de un monocromador triple es capaz de reducir esta dispersión en un 10¹², mientras que uno doble un 10⁸ y el uso de uno simple un 10⁴ [62]. Sin embargo, el uso de más de un monocromador resulta en un tamaño y configuración más complejos, así como un mayor coste.

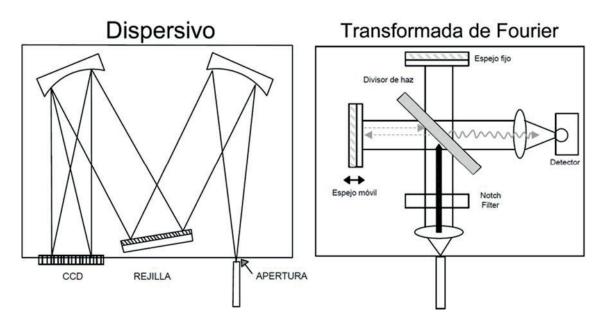


Figura I. 9. Ejemplo y comparativa de un espectrómetro dispersivo (Czerny-Turner) y un FT (Michelson interferómetro). Reproducido de [63].

3.2.2.2. Espectrómetros FT-Raman

A diferencia de los espectrómetros dispersivos, el FT-Raman consiste principalmente en un interferómetro. En este caso, la fuente de excitación del láser se direcciona hacia la muestra mediante un espejo óptico y una lente, que además recolecta la dispersión (Rayleigh y Raman) de la muestra. El espejo dicromático, transmite la dispersión Rayleigh de menor longitud de onda y reflecta la dispersión Raman de mayor longitud de onda hacía el divisor de haz. A través del divisor de haz, parte de la señal Raman es transmitida a un espejo fijo y la otra reflejada hacía un espejo móvil. Debido a la diferencia en el camino óptico, causada por el espejo móvil, los dos haces de luz reflejados de los dos espejos provocan una

interferencia constructiva y destructiva. La señal de todas las frecuencias es registrada en el detector simultáneamente obteniendo así el interferograma (**Figura I. 9**).

La resolución espectral ($\Delta\lambda$) de un espectrómetro FT-Raman se puede determinar considerando el rango máximo de movimiento del espejo ($\Delta\chi_{max}$), (**Ecuación I-11**) [64].

$$\Delta \lambda = \frac{1}{\Delta \chi_{max}} \tag{I-11}$$

Los FT-Raman presentan un mayor rendimiento, precisión y exactitud, además de una resolución espectral mayor. Además, si se usa como fuente de excitación el NIR (1064 nm), la fluorescencia se ve muy reducida. Por este motivo, este tipo de espectrómetros se usan especialmente en muestras que presentan una alta probabilidad de fluorescencia, como sería en el análisis forense o para aplicaciones farmacéuticas. Además, como las medidas de un FT-Raman se realizan en la región NIR, donde las muestras tienen una absorción mínima, se puede usar una potencia de láser mayor. Otra de las principales diferencias entre ambos tipos de espectrómetros es el ruido. Al realizar una adquisición de un espectro con un FT-Raman, se produce ruido de manera natural a lo largo de todo el rango espectral. Este aumento del ruido viene determinado por la capacidad de atenuación de la dispersión Rayleigh. Finalmente, otra diferencia entre ambos es la sensibilidad. Un espectrómetro dispersivo acoplado a un detector CCD tiene uno o dos órdenes de magnitud de sensibilidad mayor a los FT, en el caso de que no haya fluorescencia [65].

3.2.3. Detectores

Considerando la baja intensidad de la dispersión Raman, los detectores a usar deben ser muy sensibles. Los detectores más populares en los instrumentos Raman son los dispositivos de carga acoplada (CCD, *Charged Coupled Device*), que empezaron a usarse en aplicaciones Raman a finales del 1980 [66,67].

Los CCDs están formados por una matriz de elementos o pixeles, del cual la estructura principal es un diodo semiconductor de óxido de metal sobre un sustrato de silicio [68,69]. Cada elemento (o pixel) del detector es un capacitador fotoactivo que recolecta y mantiene una carga dependiendo del número de fotones dispersados que le llegan. Los detectores CCD están formados por una matriz de pixeles 2D. Los pixeles horizontales se calibran para que cada uno se corresponda con un número de onda del eje, mientras que los verticales son los miden la fuerza de la señal Raman. Lo que se obtiene es una figura electrónica de la señal

Raman que posteriormente se convierte en el espectro [43]. La principal función es el desplazamiento de la carga acumulada en cada uno de los pixeles simultáneamente a los pixeles adyacentes hasta que logran llegar al último nivel, donde se realiza su lectura individual, obteniendo el histograma presentado en la figura, como un ejemplo [63]. Cada pixel es un elemento de resolución especifica a cada número de onda y con una cantidad variable de fotones.

En este tipo de procesos, hay una variabilidad de parámetros que pueden influir en la respuesta, como, por ejemplo, el ruido asociado con cada elemento del detector, el ruido que se genera durante el desplazamiento y la lectura, la eficiencia de la matriz para detectar los fotones y la rapidez de lectura del detector. En este caso, el ruido asociado al detector se puede reducir disminuyendo la temperatura, normalmente de -20 a -70°C, donde los métodos más emprados son el enfriamiento criogénico con nitrógeno líquido y el enfriamiento termoeléctrico [68–70].

A parte de esta técnica, se han desarrollado otras para aumentar la eficiencia en la detección como por ejemplo los CCD intensificados (ICCD) y los CCDs con multiplicación de electrones (EMCCD) [71].

En el caso de la detección de la señal Raman en la región NIR, los detectores más emprados son el de InGaAs y el de Ge, que se usan junto con los espectrómetros FT-Raman.

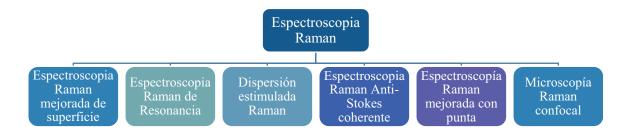


Figura I. 10. Esquema de las técnicas recientemente desarrolladas en la espectroscopia Raman.

3.3. Avances en la Espectroscopia Raman

Hoy en día existen muchos tipos de técnicas de espectroscopia Raman conocidas, gracias a los avances en la tecnología óptica. Algunas de las modalidades Raman más usadas se presentan en la **Figura I. 10.**

Las principales características y ventajas de estas técnicas han sido presentadas en distintos artículos publicados, así como sus principales aplicaciones en comparación a la espectroscopia Raman por retrodispersión [51,72–78].

3.3.1. Instrumentos portátiles

La tecnología usada para los espectrómetros Raman portátil empezó alrededor de los 90, especialmente con las cámaras CCDs que usan láseres diodo a 785 nm [66,79,80]. Muchos de los diseños de este tipo de instrumentos portátiles han sido reportados en libros y artículos [34,81–84]. Recientemente, J. Jehlička *et al.* publicaron un listado de algunos de los instrumentos comerciales disponibles donde incluían algunos de los parámetros más importantes como sus dimensiones, el láser, la resolución espectral y el tipo de conexión usado [85].

Existen distintos tipos de instrumentos portátiles, los de mano (handheld en inglés), los portátiles, los transportables y los del tamaño como la palma de una mano [86]. Sus principales ventajas en comparación a los instrumentos convencionales son sus dimensiones, así como su peso, y su menor coste. Pero, lo más importante es la posibilidad de llevar el espectrómetro a la muestra, en vez de la muestra al espectrómetro. Esta característica permite reducir el tiempo de análisis.

La **Figura I. 11**, muestra un resumen de las publicaciones relacionadas con los instrumentos Raman hasta la actualidad. Como puede verse, su implementación ha ido aumentando con los años. Esto se debe principalmente a su mejora instrumental y a su gran abanico de posibilidad de utilización en distintos sectores farmacéuticos.

3.4. Aplicaciones de la espectroscopia Raman

Considerando las características de esta técnica, puede implementarse en distintos sectores como, por ejemplo, en la industria farmacéutica, alimentaria, forense, para aplicaciones relacionadas con el arte y la arqueología, identificación de polímeros, colorantes y pigmentos, así como para la caracterización de materiales inorgánicos [73,87–95].

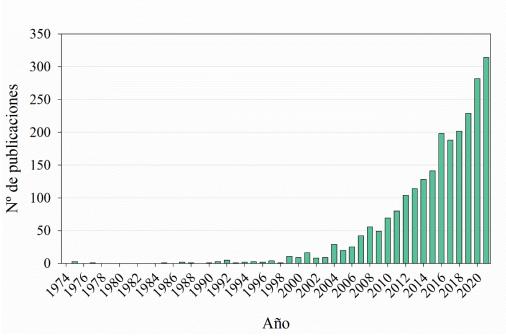


Figura I. 11. Evolución en el número de publicaciones que contienen como palabras clave Raman y Portátiles (*Portables*) o *Handheld* por año. Datos extraídos de la base de datos de *Scopus*.

En esta tesis, la espectroscopia Raman se ha usado para aplicaciones relacionadas con la industria farmacéutica. Las principales aplicaciones en esta industria están relacionadas con la identificación de los productos, tanto materiales de partida como producto acabado, la determinación de la uniformidad de contenido, la investigación y cuantificación de distintas formas polimórficas, o para la determinación de la homogeneización del API en comprimidos.

Además de todas estas aplicaciones, la espectroscopia Raman se ha convertido en una herramienta muy útil para la implementación PAT en la industria farmacéutica. Esto se debe a que se puede obtener información química y física durante distintas operaciones unitarias, pudiendo así obtener información a tiempo real del proceso de fabricación [77,96–98]. Algunas de las operaciones unitarias que pueden estudiarse son la síntesis, cristalización, homogeneización, granulación, compresión y recubrimiento [99–102]. El mayor peso de la espectroscopia Raman en la aplicación de operaciones unitarias está desarrollado para formas farmacéuticas sólidas. Por ejemplo, De Beer *et al* han aplicado esta técnica para determinar el punto final de la etapa de homogeneización durante el proceso de fabricación de un producto farmacéutico [103].

En relación con los instrumentos portátiles, su principal aplicabilidad está relacionada con la identificación de materiales de partida o producto acabado [104–107]. En referencia al análisis cuantitativo, estudios recientes han evaluado su capacidad para este tipo de aplicación [108].

4. QUIMIOMETRIA

La Quimiometria es una disciplina de la química que usa las matemáticas, estadística y la lógica formal para:

- (a) Diseñar o seleccionar procedimientos óptimos para el desarrollo experimental
- (b) Proporcionar la máxima y relevante información química mediante el análisis de los datos químicos
- (c) Obtener conocimiento de los sistemas químicos

De este modo, la aplicación de la Quimiometria resulta en una mejora de los procesos analíticos en sus distintas etapas, que se traduce en una mejora de la calidad de los resultados [109].

En la aplicación de la espectroscopia vibracional es posible la realización de un análisis univariante o multivariante de los datos. En el primer caso, se considera una única variable en el estudio; Por ejemplo la intensidad de una banda determinada a una longitud de onda del espectro obtenido [110]. En cambio, en el análisis multivariante, se emplean determinadas técnicas que permiten realizar el análisis de los datos de más de una variable a la vez, extrayendo y simplificando la información [111]. En el uso de la espectroscopia NIR, es difícil aplicar un análisis univariante dado el solapamiento y las bandas anchas del espectro. En la espectroscopia Raman puede ser posible aislar una banda concreta, pero también es habitual el estudio multivariante por la complejidad y gran información contenida en los espectros. La **Tabla I. 4** muestra algunos ejemplos de los tipos de análisis multivariable más aplicados en la espectroscopia NIR y Raman, así como los métodos más usados [31,112,113].

Tabla I. 4. Análisis de los datos NIR y	v Raman mediante el uso del análisis multivariable.

Análisis	Objetivo	Métodos usados
Procesamiento de los espectros	Reducir la información no	Suavizado
	deseada, como el ruido, y	MSC
	aumentar la relación S/N	SNC
	Magnificar pequeñas diferencias espectrales entre muestras	Normalización
Exploración de resultados	Investigar y encontrar relaciones entre muestras (tendencias)	Análisis de componentes principales (PCA)
Análisis cualitativo	Clasificar las muestras en	PLS-DA
	distintos grupos en base a	SIMCA
	diferencias espectrales	SVM
	Encontrar outliers o resultados	LDA
	fuera de especificación	Bibliotecas de espectros
Análisis cuantitativo	Correlacionar los datos	PLSR
	espectrales con parámetros	MLR
	cuantitativos de la muestra	PCR
	(contenido, humedad, etc)	ANN

4.1. Perspectiva histórica

Los inicios de la Quimiometría se establecen en 1969 tras una publicación relacionada con la aplicación de máquinas de aprendizaje (*learning machines*) para clasificar un espectro de masas de baja resolución [114–116]. Tras estas publicaciones se abrió una nueva posibilidad de transformar largas cantidades de datos en información útil. Posteriormente, en 1977, se introdujo el algoritmo del modelado suave independiente por analogía de clases (SIMCA en inglés) para el modelado de datos multivariables [117]. Mientras, en 1974, el interés ya había crecido tanto que justificó la creación de una sociedad de Quimiometría [118].

4.2. Pretratamientos espectrales

Para facilitar el análisis quimiométrico (exploratorio, cualitativo o cuantitativo) es habitual empezar con un preprocesado de los datos para mejorar sus características y eliminar efectos indeseados que complicarían su tratamiento posterior [119,120]. Los métodos de procesado más habituales en los datos espectroscópicos no requieren de un valor de referencia, por lo que se consideran independientes de referencia. La **Figura I. 12** muestra una posible clasificación de estos métodos en función de su principal objetivo [121].

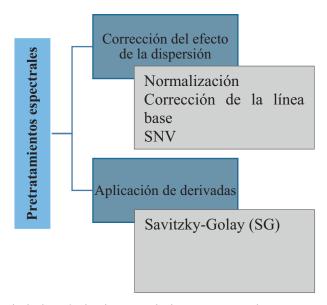


Figura I. 12. Clasificación de los principales métodos de procesado de espectros para datos espectroscópicos.

4.2.1. Corrección del efecto de la dispersión

La dispersión de la luz es un efecto muy común en las técnicas analíticas donde se usa como fuente de radiación un haz de luz. Este efecto puede ocasionar cambios en la intensidad de las bandas o en la pendiente del espectro. Algunos de los métodos de procesado más habituales para mitigar este efecto son los presentados en la **Figura I. 12**. A continuación, se detallarán los usados durante el desarrollo de esta tesis.

- <u>Variable normal estándar</u> (SNV). Este método corrige el espectro (x_{org}) mediante dos coeficientes, siendo uno el valor promedio del espectro (a_0) y el otro su desviación estándar (s). El espectro corregido (x_{SNV}) se calcula según la **Ecuación I-12.** Este pretratamiento permite mitigar variaciones multiplicativas entre espectros por diferencias en sus propiedades físicas, o diferencias derivadas de la preparación o presentación de la muestra, así como pequeñas variaciones debidas a la óptica del espectrómetro [122,123].

$$x_{SNV} = \frac{x_{org} - a_0}{s} \tag{I-12}$$

- <u>Corrección de la línea base</u>. Este pretratamiento es común para los datos espectroscópicos obtenidos con un instrumento Raman. El principal objetivo es

disminuir tanto la dispersión Rayleigh como el efecto de la fluorescencia [124]. Este tipo de métodos sustraen el desplazamiento de la línea base o la pendiente. Los más comunes se basan en una transformación *wavelet* o mediante la aplicación de un ajuste polinómico [125,126].

La transformada de *wavelet* descompone la señal original en distintas contribuciones, localizadas y caracterizadas por un valor escalado y una posición. Cada contribución a cada uno de los niveles o escalados, es la información de la señal original a distintas frecuencias [125,126]. En esta tesis, se ha utilizado un algoritmo basado en la transformada de *wavelet* llamado *baseline wavelets* (BW) y que ha sido detallado por Z. Zhang *et al*.[127]. Esta transformada, BW; consta de tres pasos: la detección de los máximos de los picos, la estimación de la amplitud de las bandas y, por último, el ajuste del fondo (*background*).

El algoritmo basado en el suavizado de Whittaker para la corrección de la línea base también ha sido usado en esta tesis. Este algoritmo se denomina también mínimos cuadrados ponderados, o mínimos cuadrados asimétricos (ALS por sus siglas Asymmetric Least Squares). Whittaker publicó en 1922 un algoritmo para el suavizado de los espectros que después mejoró Eilers [125,126]. Posteriormente, Cobat *et.al* introdujeron una corrección de línea base automática basándose en los algoritmos de Whittaker y Eilers [128]. Esta corrección se basa en encontrar el mejor ajuste de línea base a los datos obtenidos, mientras que también se penaliza la rugosidad de la línea base. La función (*F*) resultante que se minimiza para obtener la línea base es la presentada en la **Ecuación I-13.**

$$F = \sum_{i}^{N} \omega_{i} (y_{i} - z_{i})^{2} + \lambda \sum_{i}^{N-d} (\Delta^{d} z_{i})^{2}$$
 (I-13)

Donde y_i son los datos experimentales, z_i la línea base estimada, λ el factor de penalización para la segunda derivada, ω_i el peso, y Δ^d el operador de la diferencia finita de orden d. En la aproximación ALS, la asimetría se introduce mediante el parámetro p (0 < p < 1). Valores altos de este parámetro producen un mayor número de regiones con intensidades negativas en el espectro resultante. El factor de la penalización, λ , se puede asemejar al orden del polinomio, por lo que, mediante este factor se controla la curvatura de la línea base [129,130].

Por último, también se ha utilizado un algoritmo de corrección de línea base lineal. Este algoritmo realiza una corrección lineal entre las dos variables del espectro seleccionadas, corrigiendo su pendiente. De este modo, estas dos variables se transforman y convierten a un valor de y = 0, mientras que las demás variables se ajustan mediante esta corrección lineal a un nuevo valor determinado. Lo más importante de este pretratamiento es asegurar que las variables seleccionadas siempre son el valor mínimo en todos los espectros a pretratar [131].

Normalización. Este método de procesado tiene como objetivo la disminución de la influencia de variación no deseada entre muestras.

La normalización vectorial (VN), primero realiza el cálculo del espectro promedio (a_m) para todos los valores k de las muestras (N) en el rango de longitudes seleccionado. Este espectro promedio se sustrae de los individuales a'(k), centrando así todos los resultados alrededor de y = 0. Posteriormente, se calcula la suma de los cuadrados de todos los valores k y el espectro se divide por la raíz cuadrada de esta suma (a''(k)), según las **Ecuaciones I-14 – I-16**.[132]

$$a_m = \frac{\sum_k a(k)}{N} \tag{I-14}$$

$$a'(k) = a(k) - a_m \tag{I-15}$$

$$a''(k) = \frac{a'(k)}{\sqrt{\sum_{k} (a'(k))^2}}$$
 (I-16)

Otro tipo de normalización es la de normalizar al valor máximo observado para todas las variables de la muestra. En este caso, la respuesta es un vector con un valor máximo unidad. El espectro normalizado se calcularía mediante la **Ecuación I-17**, donde w_i es el peso de la normalización de la muestra i calculado según la **Ecuación I-18**, y x_i el vector de los valores observados de la muestra [133].

$$x_{i,norm} = x_i w_i^{-1} (I-17)$$

$$w_i = Max(x_i) (I-18)$$

Una alternativa a la normalización al valor máximo es la denominada normalización a área unidad. Este tipo de transformación lo que realiza es una normalización de los espectros mediante el cálculo de su área debajo de la curva, resultando así en la obtención de la misma área para todos los espectros de las muestras estudiadas [131].

4.2.2. Aplicación de derivadas

La corrección de los espectros mediante la aplicación de una derivada se usa comúnmente para reducir los efectos aditivos y multiplicativos. Los métodos más comunes son los filtros polinómicos de Savitzky-Golay (SG) y la derivada de Norris-Williams [134,135]. La **Figura I. 13** muestra una representación del efecto de la aplicación de una primera o segunda derivada a un pico gaussiano. En este caso, la primera derivada es capaz de sustraer una línea base constante, mientras que la segunda es capaz de corregir una línea de base con una cierta pendiente [136]. Este método de procesado es muy útil en casos donde hay un solapamiento de bandas, pero en determinados casos, el ruido se magnifica. Una de las formas para intentar reducir el ruido al aplicar una derivada es mediante el algoritmo SG.

El algoritmo de SG incluye un primer paso donde se realiza un suavizado. Para poder realizar la transformación de una primera deriva, se ajusta primero un polinomio mediante una ventana de puntos simétrica. Una vez calculado el ajuste, se calcula la derivada de la función. En la **Figura I. 13** se muestra una representación de este algoritmo.

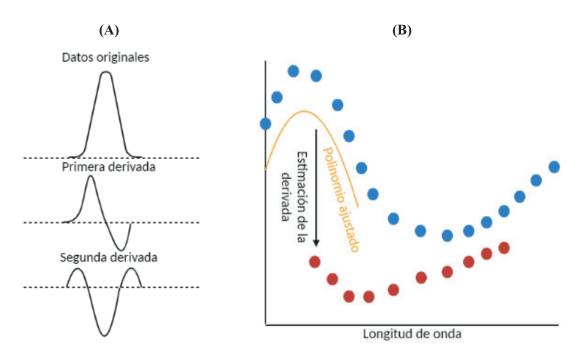


Figura I. 13. (A) Ejemplo de la aplicación de una derivada y segunda derivada en una banda gaussiana. (B) Ejemplo de la aplicación de una derivada mediante el algoritmo de SG. Adaptado de [136].

Tras la presentación de los métodos usados durante esta tesis para el procesamiento de los datos espectroscópicos, el paso más crítico es la selección del pretratamiento o combinación de pretratamientos a utilizar [137]. Existen distintas aproximaciones que han sido usadas para la selección de la mejor estrategia de procesado. La primera de ellas es la simple prueba y error, que consiste en la aplicación de distintas combinaciones para la realización del análisis cualitativo o cuantitativo. En este caso, tras el desarrollo de todos los modelos con cada uno de los pretratamientos, el que proporcione mejores resultados será el seleccionada para su propósito [121,138]. Otra posibilidad es la realización de una inspección visual de los datos antes y después de su procesado. La aplicación de esta posibilidad es más complicada, dado que depende de los datos a analizar. La última aproximación se basa en el uso de parámetros cualitativos para medir la calidad de los espectros preprocesados. El objetivo de esta última aproximación es evaluar los pretratamientos realizados para ver qué combinación resulta en un menor número de artefactos. Por ejemplo, se ha uso el coeficiente de correlación de Pearson para evaluar la similitud entre los espectros. Al obtener una mayor similitud entre los espectros tras el pretratamiento, la reducción de los artefactos habrá sido mayor [139,140]. Esta aproximación ha sido más empleada para datos cromatográficos que para datos espectroscópicos.

4.3. Análisis cualitativo

El análisis cualitativo de los datos espectroscópicos se realiza mediante el uso de métodos de reconocimiento de pautas (PRM por sus siglas *Pattern Recognition Methods*). Estos métodos se clasifican en supervisados o no supervisados dependiendo de si se conoce la clase perteneciente de la muestra (supervisados) o no (no supervisados). La **Figura I. 14** presenta un esquema de los PRM más usados con su correspondiente clasificación.

4.3.1. Medida de la similitud

Para establecer una medida cuantitativa que indique si muestras o variables se parecen entre sí en mayor o menor medida, se utiliza la similitud. La similitud se puede calcular a partir del coeficiente de correlación entre los valores de las variables o del cálculo de la distancia entre muestras o grupo de muestras.

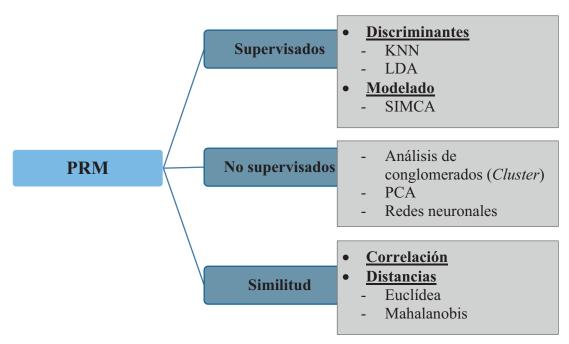


Figura I. 14. Clasificación de los PRM.

El método de correlación más usado es el de Pearson, que proporciona una medida de la correlación lineal entre dos vectores según la **Ecuación I-20**.

$$r = \frac{\sum_{i} (x_{i} - \bar{x})(y_{i} - \bar{y})}{\sqrt{\sum_{i} (x_{i} - \bar{x})^{2}} \sqrt{\sum_{i} (y_{i} - \bar{y})^{2}}}$$
(I-20)

Donde x_i y y_i son los elementos de los vectores de la intensidad que representan los espectros a comparar, y \bar{x} y \bar{y} los valores promedios de x_i y y_i , respectivamente [141,142]. Cuanto más cercano a 1 es el coeficiente de correlación, mayor es la similitud entre las variables. Esta medida de la similitud ha sido muy usado en la espectroscopia vibracional [143–145].

La medida de la similitud a partir de la distancia se basa en la idea de que cuanto más se parecen dos objetos, menor será la distancia entre ellos cuando se representan en un espacio multidimensional. Existen distintas maneras de calcular la distancia, según las propiedades de los objetos estudiados, siendo las más usadas la Euclídea y la de Mahalanobis.

- <u>Distancia Euclídea</u>. Es el tipo de distancia más sencillo y se define en la terminología vectorial según la **Ecuación I-21** [146].

$$d_{iA} = \sqrt{(x_{iA} - \bar{x}_A)(x_{iA} - \bar{x}_A)'}$$
 (I-21)

Donde A representa la clase, i la muestra, x_{iA} representa el vector que corresponde con la medida de la muestra A y \bar{x}_A el punto central.

 <u>Distancia de Mahalanobis</u>. Esta distancia, a diferencia de la Euclídea, considera que algunas variables pueden estar correlacionadas entre sí. La distancia entre dos objetos k y l viene dada por la **Ecuación I-22** [147].

$$d_{kl} = \sqrt{(x_k - x_l)C^{-1}(x_k - x_l)'}$$
 (I-22)

Donde C^{-1} representa la matriz de varianza-covarianza entre las variables.

4.3.2. Análisis de las componentes principales

La exploración de resultados se realiza comúnmente mediante métodos no supervisados, destacando el análisis de las componentes principales (PCA por sus siglas *Principal Component Analysis*). Mediante este análisis se logra una visualización del sistema bajo estudio pudiendo detectar similitudes entre muestras, identificar aglomeraciones o tendencias sistemáticas, encontrar las variables que tienen una mayor influencia o identificar posibles resultados anómalos (*outliers*).

El PCA es una técnica que se basa en la una reducción de las variables originales a un número inferior de variable que sean ortogonales entre sí y que se denominan Componentes Principales (PC). El objetivo principal de esta técnica es simplificar y reducir el número de variables de los datos originales, pero manteniendo la máxima información, y detectar patrones o relaciones entre muestras. Se debe hallar el número de componentes principales (PC), que explican la variabilidad existente en los datos originales analizados.

El procedimiento de reducción de las variables se resume en la **Figura I. 15**. Cada una de las muestras del conjunto de datos se puede representar como un punto en el espacio de las variables, donde cada una de las variables representa un eje de coordenadas. Una vez proyectadas las muestras en este espacio, se normalizan para centrarlas. A partir de aquí, se buscará el primer PC (PC1) que, según la aproximación de Pearson, es el que explica la mayor variabilidad de los datos (el de mayor varianza), por lo que este componente cruzará el centro del espacio de las variables. El PC2 será ortogonal al PC1 dado que explicará

variabilidad no incluida en el primero (son independientes) y cruzará también el centro del espacio. Este paso se repetiría sucesivamente para encontrar los demás PCs hasta llegar a cubrir el 100% de la varianza explicada. Una vez encontradas todas las direcciones, se proyectan las muestras en el espacio de los PCs, con unas nuevas coordenadas llamadas *scores*. Considerando esta aproximación, el espacio de los PCs acaba siendo una rotación del espacio de nuestras variables, pero con una menor dimensionalidad [148].

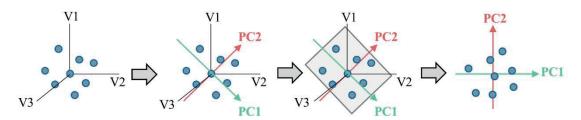


Figura I. 15. Interpretación gráfica de un PCA.

Matemáticamente, su interpretación se basa en una transformación de los datos originales $(X_{(N\times M)})$ a un nuevo eje de coordenadas ortogonal compuesto por los componentes principales (**Ecuación I-19**), donde los *scores* son representados por T, los *loadings* por P y los residuales por E.

$$X = TP^T + E (I-19)$$

La matriz de los *loadings* $(P_{(M\times F)})$ proporciona información de las variables de los datos originales $(X_{(N\times M)})$ donde F es el número de PCs. Matemáticamente, son la relación entre los ejes antiguos (variables) y los nuevos ejes (PCs), por lo que son los ángulos entre ambos. La matriz de *scores* $(P_{(M\times F)})$ proporciona información sobre las muestras y es la distancia de las muestras proyectadas al centro de los nuevos ejes. Por último, E es la matriz de los residuales.

4.3.3. Métodos discriminantes

En este tipo de métodos lo que se quiere lograr es una asignación de las muestras a una clase o clases representadas en el conjunto de entrenamiento mediante el establecimiento de unos límites.

Uno de los métodos más utilizados para este propósito es el análisis lineal discriminante (LDA por sus siglas *Linear Discriminant Analysis*). Es un método de clasificación muy básico que se basa en el uso de la información de las clases. Su principal

función es encontrar direcciones en los datos que maximicen la varianza entre clases y la minimicen dentro de una misma clase. Estas direcciones son las denominadas discriminantes lineales (LD). La principal diferencia con el PCA es que al ser éste un método no supervisado, busca la máxima varianza entre todas las muestras sin tener en cuenta la información de las clases. En cambio, el LDA contiene la información de las clases, por lo que la dirección de máxima varianza que encuentra puede ser distinta a la del PCA. La **Figura I. 16** muestra esta característica [149].

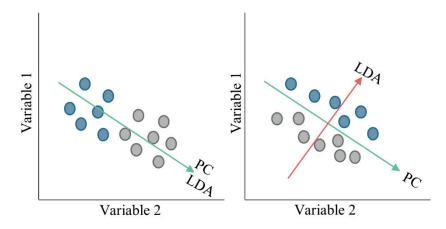


Figura I. 16. Interpretación gráfica de la dirección encontrada en un PCA y en un LDA.

El LDA, se basa en dos suposiciones. La primera es que todas las muestras de cada clase se encuentran distribuidas normalmente alrededor de su punto central (valor promedio), lo que significa que cada clase se distribuye como una gaussiana multivariable. La segunda es considerar que todas las clases tienen la misma matriz de varianza-covarianza. Si se consideran estas suposiciones, se puede calcular la probabilidad (p(g|x)) de que una muestra x pertenezca a una determinada clase g según la **Ecuación I-23** [150].

$$p(g|x) = \frac{\pi_g}{C} e^{-\frac{1}{2}(x - \bar{x}_g)^T S^{-1}(x - \bar{x}_g)}$$
(I-23)

Donde \bar{x}_g es el valor central de la clase g, S la matriz de varianza-covarianza dentro de la clase, π_g la probabilidad de observar una muestra de esa clase antes de realizar la medida, C la constante de normalización y el término $(x - \bar{x}_g)^T S^{-1}(x - \bar{x}_g)$ la distancia de Mahalanobis al cuadrado de la muestra respecto al centro de la clase. La asignación de la muestra se realiza en función del valor de probabilidad obtenido con una de las clases.

Las principales ventajas de este método son su facilidad de uso y robustez. Sin embargo, presenta dos claras desventajas. La primera está relacionada con que el número de muestras debe ser mayor al número de variables y la segunda con las dos suposiciones necesarias. La primera desventaja, se suele solventar mediante la realización previa de un PCA. De esta forma, se reduce primero la dimensionalidad y posteriormente se aplica el LDA a los *scores* del PCA obtenido. Esto se denominará LDA-PCA.

Otro método discriminante es el análisis discriminante por mínimos cuadrados (PLS-DA por sus siglas *Partial Least Squares Discriminant Analysis*). Este método discriminante se basa en el algoritmo de regresión PLS, que se explica en la Sección 3.4.2, con alguna modificación para poder realizar el análisis cualitativo. En este caso, se usa el método PLS2. El método PLS2 se usa cuando los predictores son una matriz con más de una variable. Esta matriz incluirá la información de la clase a la que pertenecen las muestras, pero en vez de usar una clasificación categórica, se usa una binaria con asignación de un valor 0 en caso de no pertenecer a la clase, o de 1 en el caso contrario (**Figura I. 17**). A partir de estas matrices, el método PLS buscará una relación entre ambas mediante la realización de un ajuste por regresión.

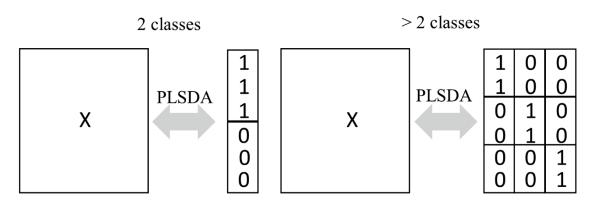


Figura I. 17. Interpretación de las matrices para un modelo PLS-DA.

4.3.4. Métodos de modelado

En estos tipos de métodos, la construcción se basa en el desarrollo de un modelo específicamente para una clase de interés y, posteriormente, combinar distintos modelos. Uno de los métodos más conocidos es el modelado suave independiente por analogía de clases (SIMCA por sus siglas *Soft Independent Modeling of Class Analogies*).

SIMCA se basa en la realización de una colección de PCAs, donde se desarrollará un PCA para cada una de las clases de interés. Tras la realización del PCA, se establecerá un umbral para pertenecer a esa clase y al predecir nuevas muestras, si estas pertenecen a la clase se encontrarán dentro del umbral, mientras que si no pertenecen serán identificadas como *outliers* [150,151]. La **Figura I. 18** muestra un resumen del método.

La identificación de la muestra como perteneciente o no a la clase se realiza mediante la detección de *Outliers* mediante los valores estadísticos: Hotelling's T² o los residuales Q. El primero detecta valores extremos pero pertenecientes al modelo o clase mientras que el segundo detecta valores que no pertenecen a este. Los residuales Q son el valor al cuadrado de los residuales del cálculo del PCA (**Ecuación I-19**). Cuanto mayor sea este valor Q, mayor será la probabilidad de que la muestra sea un *outlier*. El valor de T² es una medida de la distancia desde el origen en el espacio de los componentes principales.

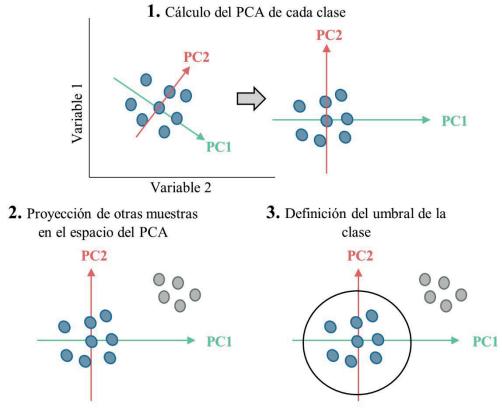


Figura I. 18. Representación de un modelo SIMCA y sus principales pasos.

El establecimiento de los límites para pertenecer a cada una de las clases se puede realizar de distintas maneras.

- El modelo SIMCA identificará la muestra *i* como perteneciente a la clase *g* si se cumplen las condiciones presentadas a continuación. Este método se basa principalmente en el establecimiento de unos límites estadísticos.

$$Q_{ig} \le Q_{0,95,g} \quad y \quad T_{ig}^2 \le T_{0,95,g}^2$$

Otra opción es mediante el cálculo de una distancia para definir el umbral (Ecuación I-24). Con este método se obtiene un umbral más ajustado a las muestras, pero el umbral establecido es fijo. Una opción para hacer el umbral móvil es modificar el criterio de que la distancia deba ser ≤ 1.

$$d_{ig} = \sqrt{\left(\frac{Q_{ig}}{Q_{0,95,g}}\right)^2 + \left(\frac{T_{ig}^2}{T_{0,95,g}^2}\right)^2} \le 1$$
 (I-24)

Las principales ventajas de este método es su facilidad de uso y interpretación dado que se basa en los modelos PCA. Sin embargo, si se quiere combinar más de una clase, suele haber muchas muestras que no se identifican como ninguna de las clases. Otra limitación es que se debe interpretar el modelo PCA para cada una de las clases.

Por último, otra opción de desarrollo de un modelo discriminante cualitativo es la construcción de las denominadas bibliotecas de espectros, que se describirá con detalle en los Capítulos V y VI de la tesis.

4.3.5. Parámetros para evaluar el rendimiento de la clasificación

El análisis de los métodos cualitativos se realiza mediante la llamada tabla o matriz de confusión, la cual incluye información sobre los resultados de la clasificación (positivos o negativos) para cada clase [152–154]. Esta matriz de confusión es una matriz cuadrada donde las filas y columnas representan las clases experimentales y predichas. De esta matriz, se puede extraer información como el número total de muestras, el número de muestras que corresponden a una clase, y el número de muestras que se predicen a una u otra clase.

Los parámetros usados para definir la calidad del método de clasificación están basados en respuestas del tipo binario, donde principalmente existen cuatro situaciones:

- <u>Positivos verdaderos</u> (TP). Muestras que son clasificadas como positivas y, realmente lo son. Equivalente a decir que se clasifican como pertenecientes a la clase A y realmente lo son.

- <u>Negativos verdaderos</u> (TN). Muestras que son clasificadas como negativas y, realmente lo son.
- <u>Falsos positivos</u> (FP). Muestras que son clasificadas como positivas, pero realmente son negativas. Equivalente a decir que se clasifican como pertenecientes a la clase A, cuando en realidad no lo son.
- <u>Falsos negativos</u> (FN). Muestras que son clasificadas como negativas, pero realmente son positivas. Equivalente a decir que no se clasifican como pertenecientes a la clase A, cuando en realidad sí lo son.

Estos parámetros son los que se pueden extraer también de la matriz de confusión (**Figura I. 19**). Además de estas cuatro situaciones, se pueden usar otros parámetros para evaluar las propiedades analíticas del método. La sensibilidad (SEN), especificidad (ESP) y precisión (PRE) sirven para evaluar la clasificación obtenida para cada una de las clases individualmente [155]. La sensibilidad representa la habilidad de identificar correctamente las muestras pertenecientes a una clase; la precisión está relacionada con la pureza de la clase, lo que evalúa la habilidad de evitar malas predicciones, y la especificidad está relacionada con la habilidad de rechazar muestras de otras clases [156,157].

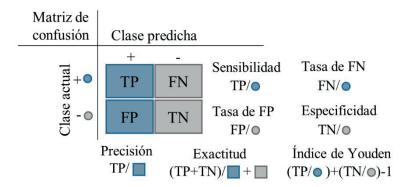


Figura I. 19. Representación de una matriz de confusión con sus parámetros.

Además de estos parámetros, también se pueden evaluar otros como la exactitud o el índice de Youden. La exactitud (EXA) proporciona una estimación del error del modelo. En el caso del Índice de Youden, expresa un compromiso entre la proporción de muestras que se identifican correctamente y las que no, proporcionando información sobre la fiabilidad del modelo de clasificar la clase A, por ejemplo. A continuación, se detallan las ecuaciones para el cálculo de estos parámetros (**Ecuaciones I-25** a **I-29**). Todos estos parámetros se expresan como un porcentaje o relación en la escala, normalmente, del 0 al 1.

$$SEN = \frac{TP}{TP + FN} \tag{I-25}$$

$$ESP = \frac{TN}{TN + FP} \tag{I-26}$$

$$PRE = \frac{TP}{TP + FP} \tag{I-27}$$

$$EXA = \frac{TN + TP}{TN + TP + FN + FP} \tag{I-28}$$

$$Youden = SEN + ESP - 1 (I-29)$$

La validación de un método cualitativo se basará en la estimación de estos parámetros analíticos donde el objetivo será lograr los mayores valores de estos parámetros, es decir, valores próximos a 1. Una forma de evaluación de estos métodos es mediante la construcción de curvas características operativas relativas (ROC por sus siglas *Relative Operating Characteristic*). En estas curvas se representan los valores de SEL y ESP en función del umbral establecido como límite a pertenecer a una clase. Mediante estas curvas se logra establecer el umbral óptimo, es decir, el umbral que permite una mejor clasificación de las muestras [155].

4.4. Análisis cuantitativo

La calibración es el proceso mediante el cual se puede relacionar los datos adquiridos de la muestra bajo estudio, por ejemplo, un espectro, con otras propiedades de interés, como la concentración. En este ámbito, existen dos tipos de calibración: univariante y multivariante. En la univariante, la información disponible de las muestras es única, por ejemplo, el valor de absorbancia a una determinada longitud de onda. En cambio, en la multivariante, se dispone de más de una variable por muestra, como sería el caso de un espectro.

4.4.1. Calibración univariante

Este tipo de calibración está representado por un vector x con dimensión $(1 \times n)$, donde n es el número de muestras disponibles. Asimismo, se dispone de un vector y de dimensión $(1 \times n)$, que contiene los valores de la propiedad analítica de interés (por ejemplo, concentración). La variable x contiene las variables independientes y y las dependientes. Al ajustar un modelo lineal con estos datos se obtiene la **Ecuación I-30**, donde

b es un escalar, el coeficiente de regresión que relaciona la respuesta con la propiedad de interés.

$$y = bx (I-30)$$

Al usar una calibración univariante, se deben cumplir dos requerimientos para poder predecir muestras desconocidas. El primero, es que la respuesta debe ser selectiva y, el segundo, que la relación entre x y y debe ser lineal. En el caso de que la respuesta no sea selectiva, es decir, que no solo dependa del analito de interés, las predicciones no serán apropiadas.

4.4.2. Calibración multivariante

La calibración multivariante se basa en el desarrollo de un modelo cuantitativo para la predicción de unas propiedades de interés a partir de unas variables predictoras. Un ejemplo de aplicación es en la espectroscopia vibracional, donde se quiere encontrar la concentración de un componente o componentes a partir de un espectro [158]. Esta calibración presenta una serie de ventajas en comparación a la univariante que se denominan de primer orden:

- Se usa un número mayor de variables (más de una longitud de onda, por ejemplo) durante el desarrollo del modelo
- El modelo no es dependiente de una única variable
- Se pueden realizar predicciones aún con una resolución de la señal baja
- Se puede realizar el análisis incluso con interferencias, dado éstas se han modelado en la calibración

El desarrollo de estos modelos multivariantes necesita de dos tipos de muestras. El primero es el denominado conjunto de calibración, del cual se tiene la información de la propiedad analítica bajo estudio. El segundo es el conjunto de validación, con muestras no usadas en el desarrollo del modelo, pero de las cuales se sabe la propiedad analizada, para verificar la predicción del modelo. La distribución de las muestras suele ser de un 70% para calibración y un 30% validación. modelo.

Existen distintos métodos para la realización de una calibración multivariante como el MLR (*Multi-linear Regression*), PCR (*Principal Component Regressión*) y PLS (*Partial Least Squares*) [159,160].

Actualmente, el método más empleado es el denominado regresión por mínimos cuadrados parciales (PLSR), y uno de los algoritmos más utilizados es originario de Herman Wold, cuando en 1975 propuso una modificación del algoritmo NIPALS para encontrar los componentes principales [161]. Su aplicabilidad en química analítica empezó por los 80s, y hoy en día, es uno de los métodos más útiles para resolver problemas con mayor número de variables x que objetos, como es el caso de la espectroscopia.

Este método matemático se basa en la transformación de los datos para encontrar la máxima covarianza entre los predictores (X) y la respuesta (Y), manteniendo la máxima correlación con Y. Esta es la principal diferencia con el otro método de reducción de variables presentado, el PCA, dado que el PLS tiene en cuenta la información de la respuesta (Y). Previamente a la descomposición, es aconsejable centrar los datos y, en el caso de que las variables tengan valores muy distintos o estén expresados en diferentes variables, también escalar los datos para darle el mismo peso a cada una de las variables (**Ecuaciones I-31** y **I-32**). El centrado permite mantener las unidades originales mientras que el método de escalado denominado "autoescalado a varianza unidad" realiza un ajuste cambiando las longitudes de los ejes. Este último pretratamiento se debe usar en el caso de que haya una gran diferencia en la varianza de las muestras.

$$x_{centrado} = x_{ik} - \bar{x}_k \tag{I-31}$$

$$x_{aut} = \frac{x_{ik} - \bar{x}_k}{S_{\nu}} \tag{I-32}$$

Donde x es el valor de la matriz en la fila i y columna k, y s_k la desviación estándar de la columna.

Para poder encontrar las variables latentes y calcular los modelos PLS, se pueden emplear diferentes algoritmos; uno de los más conocidos es el algoritmo basado en los mínimos cuadrados parciales iterativos no lineales (NIPALS por sus siglas *Nonlinear Iterative Partial Least Squares*)[162]. Este proceso se basa en un cálculo iterativo para encontrar los *scores* y *loadings* que maximizan la covarianza entre los *scores* de X y Y. Con este algoritmo, primero se descompone la matriz de las variables X según la **Ecuación I-33**, obteniendo así nuestro espacio de las variables normalizado. Sin embargo, en este caso, también disponemos de la matriz Y, la cual podemos proyectar también en el espacio (**Ecuación I-34**). Para encontrar la dirección con mayor correlación entre X y Y, se rotará el

plano que englobe los datos de la matriz X hasta obtener la correlación máxima. Para ello, se realiza una representación de los *scores* de ambos, para poder realizar las rotaciones hasta tener la máxima correlación entre los *scores* (**Ecuación I-35**). En este caso, los componentes principales se denominan variables latentes, dado que ya no son ortogonales entre ellas al buscar la máxima correlación entre *X* y *Y* (**Figura I. 20**) [163].

$$X = TP^T + E (I-33)$$

$$Y = UQ^T + F (I-34)$$

$$U = RT (I-35)$$

Donde X es la matriz de las variables, Y la matriz de la respuesta, T y U los scores, P y Q los loadings y E y F los residuales.

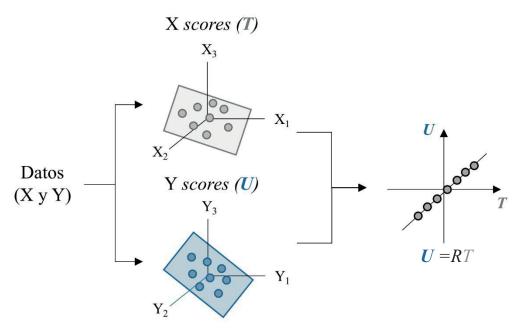


Figura I. 20. Representación de un modelo PLS. Adaptado de [164].

Los pasos de este algoritmo son:

- 1) Aleatorizar la matriz de U
- 2) Calcular el peso de la matriz U con respecto a la matriz X.
- 3) Mediante los pesos, se calculan los *scores* de *X*.
- 4) Evaluar mediante los *scores* calculados, los *loadings* de Y.
- 5) Recalcular la matriz de *U* mediante los *loadings* de *Y*.
- 6) Calcular los *loadings* de *X* considerando T (*scores* de *X*)

Estos pasos se resumen en la Figura I. 21.

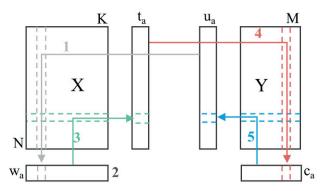


Figura I. 21. Representación del algoritmo NIPALS. Adaptado de [165].

4.5. Construcción de un modelo quimiométrico

El desarrollo y evaluación de los modelos quimiométricos, tanto cualitativos como cuantitativos, consisten en distintas etapas hasta alcanzar su propósito: la aplicación del mismo para conseguir resultados. La **Figura I. 22** muestra un esquema de las etapas, que se describirán con más detallas en las siguientes secciones.



Figura I. 22. Esquema de los pasos a seguir para la construcción de un modelo quimiométrico.

En el marco de esta tesis, la industria farmacéutica, existen distintas guías de agencias regulatorias que ofrecen indicaciones a seguir para cumplir con el marco regulatorio. Las principales guías publicadas de la EMA y FDA están destinadas al uso de la espectroscopia NIR, pero se pueden extender a otras técnicas espectroscópicas [4,5]. Además de estas guías específicas para NIR, se pueden usar las guías ICH sobre indicaciones de la validación de un método analítico (ICH Q2) y el nuevo borrador sobre el desarrollo de un proceso analítico (ICH Q14) [166,167]. Otras fuentes de información sobre el desarrollo de modelos quimiométricos se pueden encontrar en los capítulos específicos publicados en las farmacopeas, tanto la europea como la estadounidense, sobre Quimiometria, Raman, y NIR [6–8,168].

4.5.1. Recopilación de datos

La recopilación de datos engloba distintas etapas como la preparación de las muestras, el análisis por el método de referencia, la adquisición de los espectros y la separación de los conjuntos en uno de calibración y otro de validación.

(1) <u>Preparación de las muestras</u>. Uno de los pasos más críticos durante el desarrollo de un modelo quimiométrico es el diseño y preparación de un conjunto de muestras lo más representativas posibles del objetivo del estudio. Por este motivo, deben incluir la mayor variabilidad posible que pueda tener influencia en la señal analítica, como, por ejemplo, distinto proceso de fabricación o distintos proveedores del API.

Existen distintos procedimientos para la preparación de estas muestras representativas. El primero, pero el más difícil por temas económicos, es la obtención de las muestras en una planta piloto para simular lo más fielmente posible el proceso de fabricación del producto. Otra opción es la obtención de las muestras en el laboratorio mediante la mezcla de todos los ingredientes de la formulación a la concentración nominal del producto. En esta opción, se puede variar la concentración de las muestras para obtener un rango de concentración, pero será muy difícil incluir la variabilidad de fabricación en el modelo. Por último, otra alternativa es mediante sobre y baja dosificación del producto acabado para cubrir un intervalo de concentraciones alrededor del valor nominal [169–171].

Un procedimiento habitual para intentar incluir la variabilidad del proceso de fabricación en el modelo es la introducción de muestras de producción en el conjunto de calibración. De este modo, aunque las muestras se hayan preparado en un laboratorio, se incluye variabilidad del proceso en el modelo como, por ejemplo, el tamaño de partícula, fuerza de compresión entre otros parámetros físicos.

En el análisis cuantitativo, se incluirán las muestras preparadas según alguna de las estrategias comentadas más diversos lotes del producto acabado. En el caso del análisis cualitativo, solo será necesario incluir la variabilidad de distintos lotes de fabricación para mejorar la robustez de los métodos.

(2) Obtención del valor de referencia. Una vez preparado el conjunto de muestras, es necesario la obtención de su valor de referencia (concentración, contenido en agua, etc.) mediante un método analítico validado. Los métodos usados para la extracción

- de la información analítica suelen ser destructivos, por lo que, el primer paso es la obtención de los espectros de todas las muestras.
- (3) Adquisición de los espectros. En el caso de querer usar el modelo construido en la industria farmacéutica, tanto los conjuntos preparados como las muestras de producción se deben analizar usando el mismo instrumento. Normalmente, para la adquisición del espectro, se pueden configurar distintos parámetros en el instrumento, como el número de barridos o la potencia del láser, en el caso del Raman. Una de las opciones para encontrar las condiciones de adquisición óptimas es mediante un diseño de experimentos (DoE por sus siglas *Design of Experiments*). Esta herramienta permite estudiar la influencia de más de un factor simultáneamente para evaluar una determinada respuesta, así como la interacción entre factores.[172,173]
- (4) <u>Separación del conjunto de muestras</u>. El esquema generalizado usado para la separación del conjunto muestras para la creación del modelo de calibración y su evaluación se presenta en la **Figura I. 23**.[174]

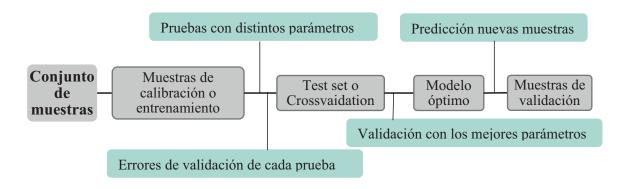


Figura I. 23. Esquema de la división del conjunto de muestras preparado para el desarrollo del modelo.

Como puede verse en la **Figura I. 23**, existen dos conjuntos de muestras que se utilizan durante el desarrollo del modelo. El primero lo constituyen las muestras de calibración/entrenamiento, que son las que se emplean para la construcción del modelo, ya sea cuantitativo o cualitativo. Este conjunto es el que contiene la mayoría de las muestras para incluir la mayor variabilidad y el mayor intervalo de concentración posible.

Como parte de la construcción del modelo, es necesario evaluar y optimizar la capacidad de predicción o clasificación del modelo construido. Si se dispone de suficientes muestras, se puede destinar una gran parte de ellas para ver cómo se comporta el modelo construido (lo que en algunos programas se denomina *Test set*). Lo más habitual es que el número de muestras disponible sea limitado, por lo que se verifica el comportamiento del modelo con una verificación interna, mediante el método de Validación cruzada (CV por sus siglas Cross-Validation). Con este método, se selecciona aleatoriamente un grupo de muestras, originalmente presentes en el conjunto de calibración, como muestras para evaluar el modelo creado con el resto de las muestras de calibración. El procedimiento se realiza repetidamente hasta que todas las muestras han quedado fuera del conjunto de calibración alguna vez, o hasta que se ha realizado un número prefijado de iteraciones. En cada una de las iteraciones el modelo es ligeramente distinto, ya que se ha realizado con una fracción de las muestras originales, y se ha evaluado también con otro conjunto de muestras distinto, calculando el error de predicción en la validación cruzada en cada iteración. Finalmente, se acumulan los resultados de cada uno de los modelos creados para hallar el óptimo. Existen distintas formas de realizar la selección aleatoria; una de ellas consiste en el método LOOCV por sus siglas Leave One Out Cross-Validation, donde, en cada iteración, se deja fuera del conjunto de calibración una de las muestras, calculando el modelo con el resto. Otra opción es utilizar un patrón fijo para seleccionar las n muestras que se quedarán fuera en cada iteración, o hacer una selección aleatoria de las mismas. La Figura I. 24 muestra estos tres tipos de CV comentados.

Por último, el conjunto de validación externa está formado por muestras independientes, que no se han utilizado durante la construcción del modelo. Al evaluar este conjunto de muestras externas con el modelo construido, se estudia la capacidad predictiva del modelo [175].

Existen distintos métodos para la separación del conjunto de muestras. Es habitual seleccionar aleatoriamente, entre todas las muestras disponibles, un grupo de ellas y utilizarlas como conjunto externo de validación, usando el resto de las muestras para el proceso de calibración. [176]. También se puede realizar esta selección utilizando un algoritmo (algoritmo de Kennard-Stone) que seleccione las

muestras más representativas para el conjunto de calibración. Considerando la distribución de los datos y seleccionando aquellos que son más representativos del conjunto para calibración [177].

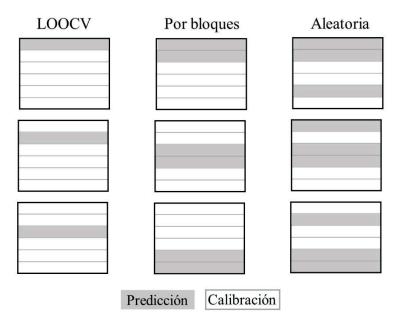


Figura I. 24. Representación de los tres tipos de CV. En gris, se indican las muestras que se seleccionan para la validación cruzada, calculando su resultado a partir de las muestras en blanco.

4.5.2. Calibración

En esta etapa del proceso de modelado es donde hay que seleccionar un intervalo de longitudes de onda, aplicar distintos pretratamientos, evaluar la presencia de *outliers* y construir el modelo cualitativo o cuantitativo.

Uno de los primeros pasos es la realización de un análisis exploratorio mediante un PCA. Esta herramienta nos permite realizar una visualización de los espectros para detectar posibles *outliers* y asegurar que las muestras seleccionadas para el conjunto de calibración son representativas de todo el conjunto. Pueden existir dos tipos de valores anómalos, los que se definen como valores extremos pero pertenecientes al modelo y los que no pertenecen al modelo. El primer tipo se podrá identificar por estar alejado en el gráfico de *scores*, mientras que el segundo por tener un residual alto. Si estos resultados se incluyeran en el modelo, introducirían mucha complejidad llegando a causar problemas de robustez. Por eso es tan importante realizar esta exploración de los datos previa al inicio del modelado. Después de la evaluación de los espectros, se deben encontrar las longitudes de onda asociadas a la presencia del analito de interés, para poder incluir estas regiones en el modelo.

Una vez descartados los espectros anómalos, identificadas las zonas del espectro asociadas al analito de interés, se puede empezar a construir el modelo quimiométrico.

La construcción del modelo se realiza mediante el uso del conjunto de muestras de calibración. En este punto, se ensayan distintos pretratamientos, intervalos de longitud de onda y se selecciona un determinado número de variables latentes en el caso de usar un método de reducción de variables (PCA, PLS, etc.). La selección del número de variables latentes que se va a usar en el modelo es crucial. En el caso de construir un modelo de reducción de variables cualitativo, el ajuste del número de variables latentes a usar se realizará mediante el gráfico de la varianza explicada frente al número de variables. Así, se podrá seleccionar como número óptimo aquel que no presente una mejora significativa al introducir un factor más. En el caso de un modelo PLS, no se evalúa solamente la varianza explicada por cada variable latente, sino que se necesita certificar que el modelo produce errores de predicción aceptables. Este proceso es el que se logra mediante la validación cruzada presentada anteriormente y el posterior cálculo del error cuadrático medio de predicción o de validación cruzada (RMSECV y RMSEP, por sus siglas *Root Mean Square Error of Cross Validation* o *Prediction*). El cálculo de estos parámetros se define en las **Ecuaciones 1-36** y **1-37** [178].

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2}{n_c}}$$
 (I-36)

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n_p}}$$
 (I-37)

Donde y_i es el valor de referencia, $\widehat{y_i}$ el valor predicho por el modelo, n_c el número de muestras de calibración y n_p las muestras de predicción.

Como se ha comentado anteriormente, la determinación del número de variables latentes es muy importante para el desarrollo de un modelo PLS. En el caso de seleccionar un menor número al ideal (o óptimo), habrá información fuera del modelo desarrollado, resultando en un modelo con falta de ajuste. En cambio, si se selecciona un número mayor al óptimo, se incluirá información no necesaria o redundante en el modelo. Esta sobreinformación implicará que las predicciones serán acuradas solo para los datos incluidos en el modelo. Esta segunda situación es la que se denomina un modelo sobre ajustado.[179]

Mediante una representación del valor de RMSECV vs el número de factores se puede escoger este último parámetro, donde se seleccionará el valor que presente el mínimo error.

4.5.3. Validación

La validación de los modelos quimiométricos desarrollados se debe realizar según los requerimientos de las agencias reguladoras [4,5,166,167]. Estas guías proporcionan información para la validación de los métodos, tanto cualitativos como cuantitativos, para poder aplicar estos modelos en el ámbito de una industria farmacéutica, dado que los modelos se deben presentar para su previa aprobación por las agencias.

En el nuevo borrador de la guía ICH Q2(R2), se ha presentado un ejemplo de los requerimientos que se deben cumplir para la determinación del contenido de un API mediante la espectroscopia NIR. Los parámetros analíticos para evaluar son la especificidad, exactitud, precisión, linealidad y robustez en un modelo cuantitativo, mientras que en el cualitativo solo es necesaria la evaluación de la especificidad. A continuación, se detallan los parámetros mencionados.

- **Especificidad o selectividad**. Evaluar que el modelo construido es capaz de identificar el analito en presencia de otros componentes, como podrían ser excipientes. Para un modelo cualitativo, este parámetro se medirá según el porcentaje obtenido de FP y FN en su validación. Para un modelo cuantitativo, lo que se evalúa es que no haya *outliers*, que se predice el parámetro bajo estudio y que la variabilidad del proceso se encuentra incluida en el conjunto de calibración.
- Linealidad y rango. Sirve para evaluar la relación lineal entre los valores predichos por el modelo y los valores de referencia. Normalmente se evalúa estadísticamente mediante la pendiente, la ordenada al origen y el valor del coeficiente de determinación (R²). Lo que se requiere, es que el pendiente sea estadísticamente igual a 1, y la ordenada a 0. Además, se sugiere que el valor de R² debe ser superior a 0,95. Sin embargo, existen otras herramientas estadísticas para la evaluación de este parámetro. Una opción es la alternativa que sugirieron Mulson y Rodbar, donde se determina si la región de confianza de la pendiente y la ordenada contienen los puntos [b=1, a=0] mediante la representación de una elipse [180,181].
- **Exactitud**. Este parámetro analítico sirve para evaluar el grado de similitud entre los valores predichos y los obtenidos mediante el método de referencia. En el caso de un

método multivariante, este parámetro se debe evaluar en el rango de concentraciones establecidos. Una herramienta útil para analizar este parámetro es la realización de una prueba t de datos pareados para establecer si existen diferencias significativas entre los valores encontrados y los de referencia.

- Precisión. La precisión se define como el grado de similitud entre una serie de medidas de una muestra obtenidas bajo las mismas condiciones (repetibilidad), con variaciones dentro del mismo laboratorio (precisión intermedia) o fuera del laboratorio (reproducibilidad). En el caso de los modelos multivariante, normalmente se evalúa la repetibilidad realizando un mínimo de seis determinaciones a la misma muestra, y la precisión intermedia, evaluando las diferencias entre distintos analistas.
- Robustez. La robustez expresa la capacidad del modelo de mantenerse constante a pesar de la presencia de pequeños cambios, es decir, evalúa la capacidad del modelo para no verse afectado por pequeñas variaciones. En este caso, se debe cubrir tanto la variabilidad física como química, así como variaciones en el proceso.

Mediante la evaluación de estos parámetros analíticos, se podrá considerar que el modelo cualitativo y/o cuantitativo han sido validados satisfactoriamente para su aplicación.

4.5.4. Mantenimiento

Los métodos quimiométricos deben ser evaluados regularmente para poder garantizar que su desarrollo está bajo control. A pesar de realizar esta monitorización, si se detecta un cambio en alguno de los parámetros críticos (proceso, proveedor, equipo, etc.) también se debe realizar un control del modelo. Mediante estos controles periódicos se quiere garantizar la robustez de los métodos. Este proceso se engloba dentro de lo establecido como gestión del ciclo de vida para los métodos multivariables. Recientemente, en el nuevo borrador de la ICH Q14 se ha propuesto un modelo de gestión del ciclo de vida para los métodos multivariables enfocado a la industria (**Figura I. 25**) [167].

En el caso de detectar que el modelo no se encuentra bajo control, se procederá a realizar una actualización del modelo para su reajuste.

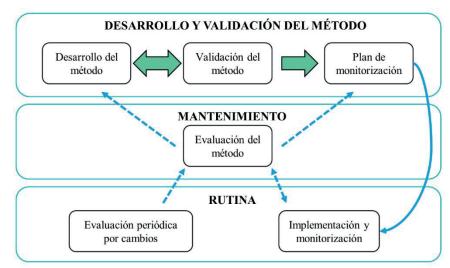


Figura I. 25. Diagrama del flujo de gestión del ciclo de vida de un método multivariante.

5. BIBLIOGRAFÍA

- [1] U.S. Department of Health and Human Service. Food and Drug Administration., Guidance for Industry PAT A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance., (2004) 16.
- [2] M. Panzitta, N. Calamassi, C. Sabatini, M. Grassi, C. Spagnoli, V. Vizzini, E. Ricchiuto, A. Venturini, A. Brogi, J. Brassier Font, L. Pontello, G. Bruno, P. Minghetti, M. Ricci, Spectrophotometry and pharmaceutical PAT/RTRT: Practical challenges and regulatory landscape from development to product lifecycle, Int. J. Pharm. 601 (2021) 120551. https://doi.org/10.1016/j.ijpharm.2021.120551.
- [3] E.J. Kim, J.H. Kim, M.-S. Kim, S.H. Jeong, D.H. Choi, *Process Analytical Technology Tools for Monitoring Pharmaceutical Unit Operations: A Control Strategy for Continuous Process Verification*, Pharm. . 13 (2021). https://doi.org/10.3390/pharmaceutics13060919.
- [4] EMA, Guideline on the use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data requirements for new Submissions and Variations, Eur. Med. Agency. 44 (2014) 1–28. https://doi.org/10.1213/ANE.0b013e318277a255.Cerebral.
- [5] FDA, Development and Submission of Near Infrared Analytical Procedures, (2021).
- [6] European Pharmacopoeia 10.0., Chapter 2.2.48. Raman spectroscopy, (2021).

- [7] European Pharmacopoeia 10.0., Chapter 2.2.40. Near-infrared spectroscopy, (2021).
- [8] European Pharmacopoeia 10.0., Chapter 5.21. Chemometrics Methods Applied to Analytical Data, (2021) 819–836.
- [9] W. Herschel, XIX. Experiments on the solar, and on the terrestrial rays that occasion heat; with a comparative view of the laws to which light and heat, or rather the rays which occasion them, are subject, in order to determine whether they are the same, or different, Philos. Trans. R. Soc. London. 90 (1800) 437–538. https://doi.org/10.1098/rstl.1800.0020.
- [10] El Sofista: El descubrimiento del infrarrojo, (n.d.). http://elsofista.blogspot.com/2010/10/el-descubrimiento-del-infrarrojo.html (accessed May 15, 2022).
- [11] D.L. Wetzel, *Near-Infrared Reflectance Analysis*, Anal. Chem. 55 (**1983**) 1165A-1176A. https://doi.org/10.1021/ac00262a718.
- [12] W.F. McClure, *Near-infrared spectroscopy. The giant is running strong*, Anal. Chem. 66 (**1994**) 43A-53A. https://doi.org/10.1021/ac00073a002.
- [13] C. Pasquini, Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications, J. Braz. Chem. Soc. 14 (2003) 198–219. https://doi.org/10.1590/S0103-50532003000200006.
- [14] Y. Ozaki, C. Huck, S. Tsuchikawa, S.B. Engelsen, Near-Infrared Spectroscopy Theory, Spectral Analysis, Instrumentation, and Applications, 1st ed. 20, **2021**. http://lib.ugent.be/catalog/ebk01:4100000011569121.
- [15] D.A. Burns, E.W. Ciurczak, Handbook of Near-Infrared Analysis, Third, 2007. https://www.routledge.com/Handbook-of-Near-Infrared-Analysis/Burns-Ciurczak/p/book/9780849373930 (accessed May 18, 2022).
- [16] P. Teppola, Near-infrared spectroscopy. Principles, instruments, applications. H. W. Siesler, Y. Ozaki, S. Kawata and H. M. Heise (eds), Wiley-VCH, Weinheim, 2002, ISBN 3-527-30149-6, 348 pp, £ 70.00, J. Chemom. 16 (2002) 636–638. https://doi.org/10.1002/cem.762.
- [17] E. Fermi, Über den Ramaneffekt des Kohlendioxyds, Zeitschrift Für Phys. 71 (1931)

- 250-259. https://doi.org/10.1007/BF01341712.
- [18] J.L. Duncan, *The determination of vibrational anharmonicity in molecules from spectroscopic observations*, Spectrochim. Acta Part A Mol. Spectrosc. 47 (**1991**) 1–27. https://doi.org/10.1016/0584-8539(91)80174-H.
- [19] siesler H.W, Ozaki Y., Kawata S., Heise H.M., *Near-infrared spectroscopy*. *Principles, instruments, applications*., J. Chemom. 16 (**2002**) 636–638. https://doi.org/10.1002/cem.762.
- [20] M. Buback, J. Schweer, H. Tups, Near Infrared Absorption of Pure Carbon Dioxide up to 3100 bar and 500 K. II. Wavenumber Range 5600 cm-1 to 7400 cm-1, Zeitschrift Fur Naturforsch. Sect. A-a J. Phys. Sci. Z NATURFORSCH SECT A. 41 (1986). https://doi.org/10.1515/zna-1986-0308.
- [21] B.T. Darling, D.M. Dennison, *The Water Vapor Molecule*, Phys. Rev. 57 (**1940**) 128–139. https://doi.org/10.1103/PhysRev.57.128.
- [22] L.E. Agelet, C.R. Hurburgh, *A Tutorial on Near Infrared Spectroscopy and Its Calibration*, Crit. Rev. Anal. Chem. 40 (**2010**) 246–260. https://doi.org/10.1080/10408347.2010.515468.
- [23] J. Malinen, M. Känsäkoski, R. Rikola, C.G. Eddison, LED-based NIR spectrometer module for hand-held and process analyser applications, Sensors Actuators B Chem. 51 (1998) 220–226. https://doi.org/10.1016/S0925-4005(98)00198-1.
- [24] M. Manley, V. Baeten, Chapter 3 Spectroscopic Technique: Near Infrared (NIR) Spectroscopy, in: D.-W.B.T.-M.T. for F.A. (Second E. Sun (Ed.), Academic Press, **2018**: pp. 51–102. https://doi.org/10.1016/B978-0-12-814264-6.00003-7.
- [25] C.D. Tran, *Acousto-optic devices. Optical elements for spectroscopy.*, Anal. Chem. 64 (1992) 971A-981A. https://doi.org/10.1021/ac00044a001.
- [26] J. Xu, R. Stroud, Acousto-optic devices: principles, design, and applications, Wiley, 1992.
- [27] G. Reich, Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications, Adv. Drug Deliv. Rev. 57 (2005) 1109–1143. https://doi.org/10.1016/j.addr.2005.01.020.

- [28] P. Williams, Sampling, sample pretreatment, and sample selectrion, in: Handb. Near-Infrared Anal., **2007**: pp. 267–296.
- [29] S. Kawano, Sampling and sample presentation, in: Near-Infrared Spectrosc. Pirnciples, Instrumentations, Appl., **2002**: pp. 115–124.
- [30] Y. Ozaki, J. Berry, Sampling techniques in near-infrared transmission spectroscopy, in: Handb. Vib. Spectrosc., **2002**: pp. 953–959.
- [31] F.W. Hong, K.S. Chia, *A review on recent near infrared spectroscopic measurement setups and their challenges*, Measurement. 171 (2021) 108732. https://doi.org/10.1016/j.measurement.2020.108732.
- [32] A. Rogalski, *History of infrared detectors*, Opto-Electronics Rev. 20 (**2012**) 279–308. https://doi.org/10.2478/s11772-012-0037-7.
- [33] E. Stark, K. Luchter, *NIR Instrumentation Technology*, NIR News. 16 (2005) 13. https://doi.org/10.1255/nirn.855.
- [34] R.A. Crocombe, *Portable Spectroscopy*, Appl. Spectrosc. Vol. 72, Issue 12, Pp. 1701-1751. 72 (**2018**) 1701–1751. https://www.osapublishing.org/as/abstract.cfm?uri=as-72-12-1701 (accessed October 5, 2021).
- [35] K.B. Beć, J. Grabska, C.W. Huck, *Principles and Applications of Miniaturized Near-Infrared (NIR) Spectrometers*, Chem. A Eur. J. 27 (2021) 1514–1532.
- [36] H. Yan, H.W. Siesler, *Hand-held near-infrared spectrometers: State-of-the-art instrumentation and practical applications*, NIR News. 29 (**2018**) 8–12. https://doi.org/10.1177/0960336018796391.
- [37] R. Deidda, P.Y. Sacre, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers, TrAC Trends Anal. Chem. 114 (2019) 251–259.
- [38] K.B. Beć, J. Grabska, H.W. Siesler, C.W. Huck, *Handheld near-infrared spectrometers: Where are we heading?*, NIR News. 31 (2020) 28–35. https://doi.org/10.1177/0960336020916815.
- [39] R. Singh, F. Riess, Sir C. V. Raman and the story of the Nobel prize, Curr. Sci. 75

- (1998) 965–971. http://www.jstor.org/stable/24101681.
- [40] C. V RAMAN, *The Colour of the Sea*, Nature. 108 (1921) 367. https://doi.org/10.1038/108367a0.
- [41] C.. Raman, K.. Krishnan, *A New Type of Secondary Radiation*, Nature. 121 (**1928**) 501–502. https://www.nature.com/nature/journal/v121/n3048/pdf/121501c0.pdf (accessed April 27, 2017).
- [42] E. Smith, G. Dent, Modern Raman Spectrocopy: A Practical Approach, 2005.
- [43] S. Sasic, S. Ekins, Pharmaceutical applications of Raman spectroscopy, Wiley-Interscience, **2008**.
- [44] R.S. Krishnan, R.K. Shankar, *Raman effect: History of the discovery*, J. Raman Spectrosc. 10 (1981) 1–8. https://doi.org/10.1002/jrs.1250100103.
- [45] D.A. Long, *Early history of the Raman effect*, Int. Rev. Phys. Chem. 7 (**1988**) 317–349. https://doi.org/10.1080/01442358809353216.
- [46] P.J. Larkin, Chapter 2 Basic Principles, in: P.J.B.T.-I. and R.S. (Second E. Larkin (Ed.), Infrared Raman Spectrocopy Princ. Spectr. Interpret., 2nd ed., Elsevier, **2018**: pp. 7–28. https://doi.org/10.1016/B978-0-12-804162-8.00002-1.
- [47] B. Valeur, M.N. Berberan-Santos, *A Brief History of Fluorescence and Phosphorescence before the Emergence of Quantum Theory*, J. Chem. Educ. 88 (2011) 731–738. https://doi.org/10.1021/ed100182h.
- [48] B. Schrader, D. Bougeard, Infrared and Raman spectroscopy: methods and applications, VCH, **1995**.
- [49] B. Dietzek, D. Cialla, M. Schmitt, J. Popp, Introduction to the Fundamentals of Raman Spectroscopy BT - Confocal Raman Microscopy, in: T. Dieing, O. Hollricher, J. Toporski (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2011: pp. 21–42. https://doi.org/10.1007/978-3-642-12522-5 2.
- [50] D.A. Long, The Raman effect: a unified treatment of the theory of Raman scattering by molecules, Wiley, **2002**.
- [51] N. John, S. George, Chapter 5 Raman Spectroscopy, in: S. Thomas, R. Thomas,

- A.K. Zachariah, R.K.B.T.-S.M. for N.C. Mishra (Eds.), Micro Nano Technol., Elsevier, **2017**: pp. 95–127. https://doi.org/10.1016/B978-0-323-46140-5.00005-4.
- [52] J.R. Ferraro, K. Nakamoto, C.W. Brown, Chapter 2 Instrumentation and Experimental Techniques, in: J.R. Ferraro, K. Nakamoto, C.W.B.T.-I.R.S. (Second E. Brown (Eds.), Introd. Raman Spectrosc., 2nd ed., Academic Press, San Diego, 2003: pp. 95–146. https://doi.org/10.1016/B978-012254105-6/50005-6.
- [53] D.A. Long, Raman spectroscopy, McGraw-Hill, New York, N.Y., 1977.
- [54] H. Kogelnik, S.P.S. Porto, Continuous Helium—Neon Red Laser as a Raman Source,
 J. Opt. Soc. Am. 53 (1963) 1446–1447. https://doi.org/10.1364/JOSA.53.001446.
- [55] Z. Li, M.J. Deen, S. Kumar, P.R. Selvaganapathy, *Raman spectroscopy for in-line water quality monitoring--instrumentation and potential.*, Sensors (Basel). 14 (2014) 17275–17303. https://doi.org/10.3390/s140917275.
- [56] D.B. Chase, Fourier transform Raman spectroscopy., J. Am. Chem. Soc. 108 (1986) 7485–7488. https://doi.org/10.1021/ja00284a007.
- [57] T. Hirschfeld, B. Chase, *FT-Raman Spectroscopy: Development and Justification*, Appl. Spectrosc. 40 (**1986**) 133–137. http://opg.optica.org/as/abstract.cfm?URI=as-40-2-133.
- [58] C.L. Schoen, S.K. Sharma, C.E. Helsley, H. Owen, *Performance of a Holographic Supernotch Filter*, Appl. Spectrosc. 47 (1993) 305–308. https://doi.org/10.1366/0003702934066776.
- [59] M. Futamata, Dielectric Filter for Highly Sensitive Raman Spectroscopy, Appl. Spectrosc. 50 (1996) 199–204. http://opg.optica.org/as/abstract.cfm?URI=as-50-2-199.
- [60] C.H. Munro, V. Pajceni, S.A. Asher, Dielectric Stack Filters for Ex Situ and In Situ UV Optical-Fiber Probe Raman Spectroscopic Measurements, Appl. Spectrosc. 51 (1997) 1722–1729. https://doi.org/10.1366/0003702971939424.
- [61] I.R. Lewis, H.G.M. Edwards, Handbook of Raman spectroscopy: from the research laboratory to the process line, Marcel Dekker, **2001**. https://www.crcpress.com/Handbook-of-Raman-Spectroscopy-From-the-Research-

- Laboratory-to-the-Process/Lewis-Edwards/p/book/9780824705572 (accessed July 18, 2017).
- [62] S.A. Asher, R. Bormett, Raman Instrumentation BT Raman Scattering in Materials Science, in: W.H. Weber, R. Merlin (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, **2000**: pp. 35–54. https://doi.org/10.1007/978-3-662-04221-2_2.
- [63] P. Motousek, M. D. Morris, Emerging Raman Applications and Techniques in Biomedical and Pharmaceutical Fields, **2010**.
- [64] R.L. McCreery, Raman Spectroscopy for Chemical Analysis, **2000**.
- [65] B. Chase, FT–Raman Spectroscopy: A Catalyst for the Raman Explosion?, J. Chem. Educ. 84 (2007) 75. https://doi.org/10.1021/ed084p75.
- [66] J.M. Williamson, R.J. Bowling, R.L. McCreery, *Near-Infrared Raman Spectroscopy with a 783-nm Diode Laser and CCD Array Detector*, Appl. Spectrosc. 43 (**1989**) 372–375. https://doi.org/10.1366/0003702894203048.
- [67] M.M. Carrabba, K.M. Spencer, C. Rich, D. Rauh, The Utilization of a Holographic Bragg Diffraction Filter for Rayleigh Line Rejection in Raman Spectroscopy, Appl. Spectrosc. 44 (1990) 1558–1561. https://doi.org/10.1366/0003702904417715.
- [68] T. Hardy, R. Murowinski, M.J. Deen, *Charge transfer efficiency in proton damaged CCD's*, IEEE Trans. Nucl. Sci. 45 (1998) 154–163. https://doi.org/10.1109/23.664167.
- [69] T. Hardy, M.J. Deen, R. Murowinski, *Effects of radiation damage on scientific charge coupled devices*, Adv. Imaging Electron Phys. 106 (**1999**) 1–96.
- [70] R.G. Murowinski, G. Linzhuang, M.J. Deen, Effects of space radiation damage and temperature on the noise in CCDs and LDD MOS transistors, IEEE Trans. Nucl. Sci. 40 (1993) 288–294. https://doi.org/10.1109/23.221053.
- [71] A. O'Grady, A comparison of EMCCD, CCD and emerging technologies optimized for low-light spectroscopy applications, in: Proc.SPIE, **2006**. https://doi.org/10.1117/12.644752.
- [72] R.R. Jones, D.C. Hooper, L. Zhang, D. Wolverson, V.K. Valev, *Raman Techniques:* Fundamentals and Frontiers, Nanoscale Res. Lett. 14 (2019) 231.

- https://doi.org/10.1186/s11671-019-3039-2.
- [73] A. Orlando, F. Franceschini, C. Muscas, S. Pidkova, M. Bartoli, M. Rovere, A. Tagliaferro, *A Comprehensive Review on Raman Spectroscopy Applications*, Chemosens. 9 (2021). https://doi.org/10.3390/chemosensors9090262.
- [74] R.S. Das, Y.K. Agrawal, *Raman spectroscopy: Recent advancements, techniques and applications*, Vib. Spectrosc. 57 **(2011)** 163–176. https://doi.org/10.1016/j.vibspec.2011.08.003.
- [75] R. Pilot, R. Signorini, C. Durante, L. Orian, M. Bhamidipati, L. Fabris, *A Review on Surface-Enhanced Raman Scattering*., Biosensors. 9 **(2019)**. https://doi.org/10.3390/bios9020057.
- [76] W.J. Tipping, M. Lee, A. Serrels, V.G. Brunton, A.N. Hulme, *Stimulated Raman scattering microscopy: an emerging tool for drug discovery*, Chem. Soc. Rev. 45 (2016) 2075–2089. https://doi.org/10.1039/C5CS00693G.
- [77] A. Paudel, D. Raijada, J. Rantanen, *Raman spectroscopy in pharmaceutical product design*, Adv. Drug Deliv. Rev. 89 (2015) 3–20. https://doi.org/10.1016/j.addr.2015.04.003.
- [78] K. Bērziņš, S.J. Fraser-Miller, K.C. Gordon, *A New Frontier for Nondestructive Spatial Analysis of Pharmaceutical Solid Dosage Forms: Spatially Offset Low-Frequency Raman Spectroscopy*, Anal. Chem. 93 (**2021**) 3698–3705. https://doi.org/10.1021/acs.analchem.0c04960.
- [79] M. Kim, H. Owen, P.R. Carey, *High-Performance Raman Spectroscopic System Based on a Single Spectrograph, CCD, Notch Filters, and a Kr+ Laser Ranging from the Near-IR to Near-UV Regions*, Appl. Spectrosc. 47 (**1993**) 1780–1783. https://doi.org/10.1366/0003702934066145.
- [80] Y. Wang, R.L. McCreery, Evaluation of a diode laser/charge coupled device spectrometer for near-infrared Raman spectroscopy, Anal. Chem. 61 (1989) 2647–2651. https://doi.org/10.1021/ac00198a012.
- [81] C. Rathmell, D. Bingemann, M. Zieg, D. Creasey, *Portable Raman Spectroscopy: Instrumentation and Technology*, Portable Spectrosc. Spectrom. (2021) 115–145.

- https://doi.org/10.1002/9781119636489.ch6.
- [82] R.L. McCreery, Raman Spectroscopy for Chemical Analysis, **2000**.
- [83] M. Pelletier, Raman instrumentation, in: Anal. Appl. Raman Spectrosc., **1999**: pp. 53–105.
- [84] B.A. Eckenrode, E.G. Bartick, S.D. Harvey, M.E. Vucilick, B.W. Wright, R.A. Huff, *Portable Raman Spectroscopy Systems for Field Analysis*, Forensic Sci. Commun. 3

 (2001). https://www.ojp.gov/ncjrs/virtual-library/abstracts/portable-raman-spectroscopy-systems-field-analysis (accessed May 22, 2022).
- [85] J. Jehlička, A. Culka, *Critical evaluation of portable Raman spectrometers: From rock outcrops and planetary analogs to cultural heritage A review*, Anal. Chim. Acta. 1209 (**2022**) 339027. https://doi.org/10.1016/j.aca.2021.339027.
- [86] P. Vandenabeele, H.G.M. Edwards, J. Jehlička, *The role of mobile instrumentation in novel applications of Raman spectroscopy: archaeometry, geosciences, and forensics*, Chem. Soc. Rev. 43 (**2014**) 2628–2649. https://doi.org/10.1039/C3CS60263J.
- [87] M. Arslan, H.E. Tahir, M. Zareef, J. Shi, A. Rakha, M. Bilal, H. Xiaowei, L. Zhihua, Z. Xiaobo, Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques, Trends Food Sci. Technol. 107 (2021) 80–113. https://doi.org/10.1016/j.tifs.2020.11.021.
- [88] M.A. Fikiet, S.R. Khandasammy, E. Mistek, Y. Ahmed, L. Halámková, J. Bueno, I.K. Lednev, Surface enhanced Raman spectroscopy: A review of recent applications in forensic science, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 197 (2018) 255–260. https://doi.org/10.1016/j.saa.2018.02.046.
- [89] S.-Y. Lin, W.-T. Cheng, *The use of hot-stage microscopy and thermal micro-Raman spectroscopy in the study of phase transformation of metoclopramide HCl monohydrate*, J. Raman Spectrosc. 43 **(2012)** 1166–1170. https://doi.org/10.1002/jrs.3155.
- [90] A. Hédoux, Y. Guinet, M. Descamps, *The contribution of Raman spectroscopy to the analysis of phase transformations in pharmaceutical compounds.*, Int. J. Pharm. 417

- (**2011**) 17–31. https://doi.org/10.1016/j.ijpharm.2011.01.031.
- [91] A. Hédoux, Y. Guinet, L. Paccou, F. Danède, P. Derollez, *Polymorphic transformation of anhydrous caffeine upon grinding and hydrostatic pressurizing analyzed by low-frequency raman spectroscopy.*, J. Pharm. Sci. 102 (**2013**) 162–170. https://doi.org/10.1002/jps.23346.
- [92] N. Chieng, T. Rades, J. Aaltonen, *An overview of recent studies on the analysis of pharmaceutical polymorphs*, J. Pharm. Biomed. Anal. 55 (**2011**) 618–644. https://doi.org/10.1016/j.jpba.2010.12.020.
- [93] S. Piqueras, L. Duponchel, R. Tauler, A. de Juan, *Monitoring polymorphic transformations by using in situ Raman hyperspectral imaging and image multiset analysis*, Anal. Chim. Acta. 819 (2014) 15–25. https://doi.org/10.1016/j.aca.2014.02.027.
- [94] H. Park, H. Nie, A. Dhiman, V. Tomar, Q.T. Zhou, Understanding Dynamics of Polymorphic Conversion during the Tableting Process Using In Situ Mechanical Raman Spectroscopy., Mol. Pharm. 17 (2020) 3043–3052. https://doi.org/10.1021/acs.molpharmaceut.0c00460.
- [95] P.-Y. Sacré, P. Lebrun, P.-F. Chavez, C. De Bleye, L. Netchacovitch, E. Rozet, R. Klinkenberg, B. Streel, P. Hubert, E. Ziemons, *A new criterion to assess distributional homogeneity in hyperspectral images of solid pharmaceutical dosage forms*, Anal. Chim. Acta. 818 (2014) 7–14. https://doi.org/10.1016/j.aca.2014.02.014.
- [96] K.A. Esmonde-White, M. Cuellar, C. Uerpmann, B. Lenain, I.R. Lewis, *Raman spectroscopy as a process analytical technology for pharmaceutical manufacturing and bioprocessing*, Anal. Bioanal. Chem. 409 (**2017**) 637–649. https://doi.org/10.1007/s00216-016-9824-1.
- [97] J. Rantanen, *Process analytical applications of Raman spectroscopy.*, J. Pharm. Pharmacol. 59 (**2007**) 171–177. https://doi.org/10.1211/jpp.59.2.0004.
- [98] K.L. Davis, M.S. Kemper, I.R. Lewis, *Raman Spectroscopy for Monitoring Real-Time Processes in the Pharmaceutical Industry*, Pharm. Appl. Raman Spectrosc. (2007) 117–162. https://doi.org/10.1002/9780470225882.ch5.

- [99] J. Radtke, H. Rehbaum, P. Kleinebudde, Raman Spectroscopy as a PAT-Tool for Film-Coating Processes: In-Line Predictions Using one PLS Model for Different Cores, Pharmaceutics. 12 (2020) 796. https://doi.org/10.3390/pharmaceutics12090796.
- [100] K. Knop, P. Kleinebudde, *PAT-tools for process control in pharmaceutical film coating applications*., Int. J. Pharm. 457 (2013) 527–536. https://doi.org/10.1016/j.ijpharm.2013.01.062.
- [101] M. Wirges, A. Funke, P. Serno, K. Knop, P. Kleinebudde, Development and in-line validation of a Process Analytical Technology to facilitate the scale up of coating processes, J. Pharm. Biomed. Anal. 78–79 (2013) 57–64. https://doi.org/10.1016/j.jpba.2013.01.037.
- [102] Z.K. Nagy, G. Fevotte, H. Kramer, L.L. Simon, Recent advances in the monitoring, modelling and control of crystallization systems, Chem. Eng. Res. Des. 91 (2013) 1903–1922. https://doi.org/10.1016/j.cherd.2013.07.018.
- [103] T.R.M. De Beer, C. Bodson, B. Dejaegher, B. Walczak, P. Vercruysse, A. Burggraeve, A. Lemos, L. Delattre, Y. Vander Heyden, J.P. Remon, C. Vervaet, W.R.G. Baeyens, *Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process.*, J. Pharm. Biomed. Anal. 48 (2008) 772–779. https://doi.org/10.1016/j.jpba.2008.07.023.
- [104] K. Dégardin, A. Guillemain, Y. Roggo, Comprehensive Study of a Handheld Raman Spectrometer for the Analysis of Counterfeits of Solid-Dosage Form Medicines, J. Spectrosc. 2017 (2017) 3154035. https://doi.org/10.1155/2017/3154035.
- [105] D. Yang, R. Thomas, *The Benefits of a High-Performance, Handheld Raman Spectrometer for the Rapid Identification of Pharmaceutical Raw Materials* | *American Pharmaceutical Review The Review of American Pharmaceutical Business & Technology*, Am. Pharm. Rev. (2012). https://www.americanpharmaceuticalreview.com/Featured-Articles/126738-The-Benefits-of-a-High-Performance-Handheld-Raman-Spectrometer-for-the-Rapid-Identification-of-Pharmaceutical-Raw-Materials/ (accessed May 22, 2022).

- [106] P.H. Ciza, P.-Y. Sacre, C. Waffo, L. Coïc, H. Avohou, J.K. Mbinze, R. Ngono, R.D. Marini, P. Hubert, E. Ziemons, Comparing the qualitative performances of handheld NIR and Raman spectrophotometers for the detection of falsified pharmaceutical products, Talanta. 202 (2019) 469–478. https://doi.org/10.1016/j.talanta.2019.04.049.
- [107] S. Assi, Raw material identification using dual laser handheld Raman spectroscopy, Eur. Pharm. Rev. 1360–8606 (2013).
- [108] M.A. Mansouri, P.-Y. Sacré, L. Coïc, C. De Bleye, E. Dumont, A. Bouklouze, P. Hubert, R.D. Marini, E. Ziemons, Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study, Talanta. 207 (2020) 120306. https://doi.org/10.1016/j.talanta.2019.120306.
- [109] M. Blanco, V. Cerdà, Temas Avanzados de Quimiometría, 1st ed., 2007.
- [110] A. Kuriyama, Y. Ozaki, Assessment of active pharmaceutical ingredient particle size in tablets by Raman chemical imaging validated using polystyrene microsphere size standards., AAPS PharmSciTech. 15 (2014) 375–387. https://doi.org/10.1208/s12249-013-0064-9.
- [111] M. Fransson, J. Johansson, A. Sparén, O. Svensson, Comparison of multivariate methods for quantitative determination with transmission Raman spectroscopy in pharmaceutical formulations, J. Chemom. 24 (2010) 674–680. https://doi.org/10.1002/cem.1330.
- [112] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*, J. Pharm. Biomed. Anal. 44 (2007) 683–700. https://doi.org/10.1016/j.jpba.2007.03.023.
- [113] Z. Xiaobo, Z. Jiewen, M.J.W. Povey, M. Holmes, M. Hanpin, Variables selection methods in near-infrared spectroscopy, Anal. Chim. Acta. 667 (2010) 14–32. https://doi.org/10.1016/j.aca.2010.03.048.
- [114] B.R. Kowalski, P.C. Jurs, T.L. Isenhour, C.N. Reilley, *Computerized learning machines applied to chemical problems. Multicategory pattern classification by least squares*, Anal. Chem. 41 (1969) 695–700. https://doi.org/10.1021/ac60275a026.

- [115] B.R. Kowalski, P.C. Jurs, T.L. Isenhour, C.N. Reilley, Computerized learning machines applied to chemical problems. Interpretation of infrared spectrometry data, Anal. Chem. 41 (1969) 1945–1949. https://doi.org/10.1021/ac50159a026.
- [116] P.C. Jurs, B.R. Kowalski, T.L. Isenhour, *Computerized learning machines applied to chemical problems. Molecular formula determination from low resolution mass spectrometry*, Anal. Chem. 41 (1969) 21–27. https://doi.org/10.1021/ac60270a002.
- [117] S. Wold, M. Sjöström, SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy, in: Chemom. Theory Appl., American Chemical Society, 1977: pp. 12–243. https://doi.org/10.1021/bk-1977-0052.ch012.
- [118] D. Massart, B.G. Vandeginste, L.M.C. Buydens, S. De Jong, P. Lewi, J. Smeyers-Verbeke, Chapter 1 Introduction, in: D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J.B.T.-D.H. in S. and T. Smeyers-Verbeke (Eds.), Handb. Chemom. Qualimetrics Part A, Elsevier, 1998: pp. 1–20. https://doi.org/10.1016/S0922-3487(97)80031-7.
- [119] M. Casale, N. Sinelli, P. Oliveri, V. Di Egidio, S. Lanteri, *Chemometrical strategies* for feature selection and data compression applied to NIR and MIR spectra of extra virgin olive oils for cultivar identification, Talanta. 80 (2010) 1832–1837. https://doi.org/10.1016/j.talanta.2009.10.030.
- [120] J. Luypaert, S. Heuerding, Y. Vander Heyden, D.L. Massart, *The effect of preprocessing methods in reducing interfering variability from near-infrared measurements of creams*, J. Pharm. Biomed. Anal. 36 (2004) 495–503. https://doi.org/10.1016/j.jpba.2004.06.023.
- [121] M.D. Peris-Díaz, A. Krężel, A guide to good practice in chemometric methods for vibrational spectroscopy, electrochemistry, and hyphenated mass spectrometry, TrAC Trends Anal. Chem. 135 (2021) 116157. https://doi.org/10.1016/j.trac.2020.116157.
- [122] 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. https://doi.org/10.1366/0003702894202201.
- [123] T. Fearn, C. Riccioli, A. Garrido-Varo, J.E. Guerrero-Ginel, On the geometry of SNV

- *and MSC*, Chemom. Intell. Lab. Syst. 96 **(2009)** 22–26. https://doi.org/10.1016/j.chemolab.2008.11.006.
- [124] P.H.C. Eilers, P.M. Kroonenberg, *Modeling and correction of Raman and Rayleigh scatter in fluorescence landscapes*, Chemom. Intell. Lab. Syst. 130 (2014) 1–5. https://doi.org/10.1016/j.chemolab.2013.09.002.
- [125] O. Ryabchykov, S. Guo, T. Bocklitz, *Analyzing Raman spectroscopic data*, Phys. Sci. Rev. 4 (2019). https://doi.org/10.1515/psr-2017-0043.
- [126] H. Hu, J. Bai, G. Xia, W. Zhang, Y. Ma, Improved Baseline Correction Method Based on Polynomial Fitting for Raman Spectroscopy, Photonic Sensors. 8 (2018) 332–340. https://doi.org/10.1007/s13320-018-0512-y.
- [127] Z.-M. Zhang, S. Chen, Y.-Z. Liang, Z.-X. Liu, Q.-M. Zhang, L.-X. Ding, F. Ye, H. Zhou, An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy, J. Raman Spectrosc. 41 (2009) 659–669. https://doi.org/10.1002/jrs.2500.
- [128] J. Carlos Cobas, M.A. Bernstein, M. Martín-Pastor, P.G. Tahoces, A new general-purpose fully automatic baseline-correction procedure for 1D and 2D NMR data, J. Magn. Reson. 183 (2006) 145–151. https://doi.org/10.1016/j.jmr.2006.07.013.
- [129] V.I. Korepanov, Asymmetric least-squares baseline algorithm with peak screening for automatic processing of the Raman spectra, J. Raman Spectrosc. 51 (2020) 2061–2065. https://doi.org/10.1002/jrs.5952.
- [130] Eigenvector Research Documentation Wiki, *Advanced Preprocessing: Noise, Offset, and Baseline Filtering*, (n.d.). https://wiki.eigenvector.com/index.php?title=Advanced_Preprocessing:_Noise,_Offset, and Baseline Filtering (accessed May 22, 2022).
- [131] CAMO softwares, The Unscrambler X 10.3: User Manual, n.d.
- [132] Bruker, IDENT, in: Opus Spectrosc. Softw. User Man., Version 7, n.d.
- [133] Eigenvector Research Documentation Wiki, *Advanced Preprocessing: Sample Normalization*, (n.d.). https://wiki.eigenvector.com/index.php?title=Advanced Preprocessing: Sample No

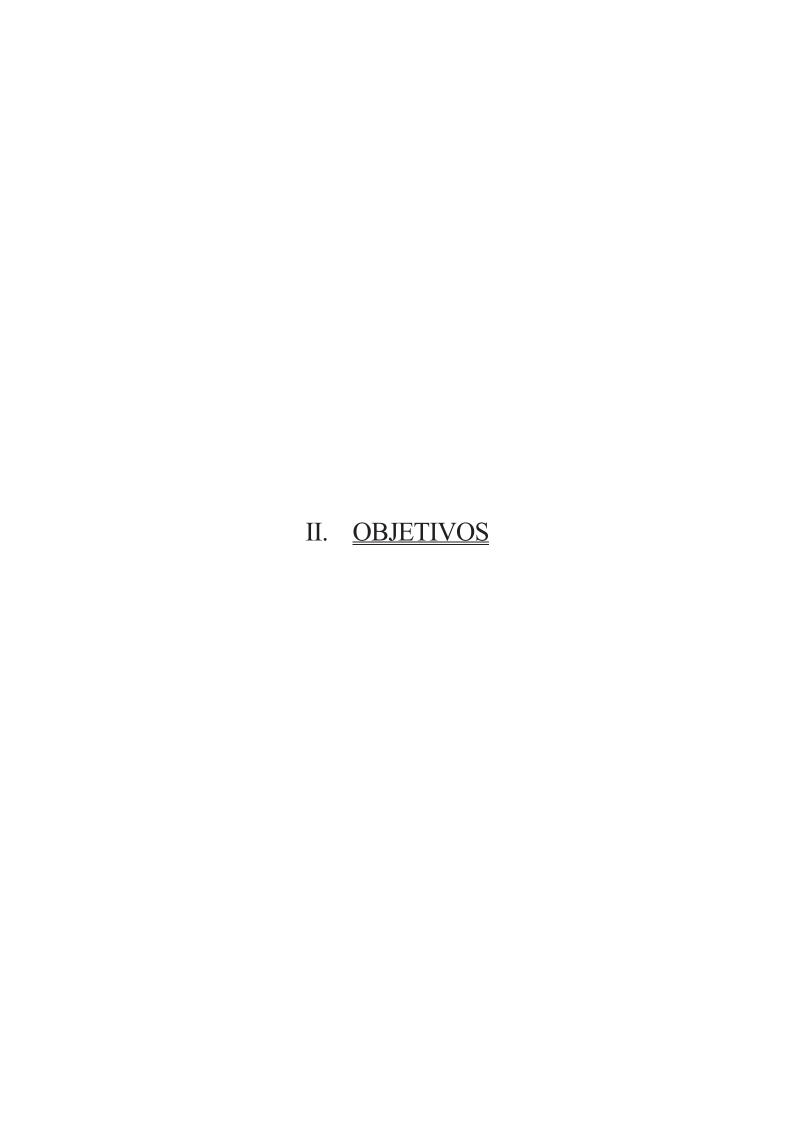
- rmalization (accessed May 22, 2022).
- [134] K.H. Norris, P.C. Williams, *Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. I. Influence of particle size*, Cereal Chem. 61 (1984) 158–165. http://europepmc.org/abstract/AGR/IND84039856.
- [135] A. Savitzky, M.J.E. Golay, *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.*, Anal. Chem. 36 (1964) 1627–1639. https://doi.org/10.1021/ac60214a047.
- [136] Åsmund Rinnan, F. van den Berg, S.B. Engelsen, *Review of the most common pre- processing techniques for near-infrared spectra*, TrAC Trends Anal. Chem. 28 (**2009**) 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007.
- [137] A. Famili, W.-M. Shen, R. Weber, E. Simoudis, *Data preprocessing and intelligent data analysis*, Intell. Data Anal. 1 (1997) 3–23. https://doi.org/10.1016/S1088-467X(98)00007-9.
- [138] J. Engel, J. Gerretzen, E. Szymańska, J.J. Jansen, G. Downey, L. Blanchet, L.M.C. Buydens, *Breaking with trends in pre-processing?*, TrAC Trends Anal. Chem. 50 (2013) 96–106. https://doi.org/10.1016/j.trac.2013.04.015.
- [139] T. Skov, F. van den Berg, G. Tomasi, R. Bro, Automated alignment of chromatographic data, J. Chemom. 20 (2006) 484–497. https://doi.org/10.1002/cem.1031.
- [140] W. Wu, M. Daszykowski, B. Walczak, B.C. Sweatman, S.C. Connor, J.N. Haselden, D.J. Crowther, R.W. Gill, M.W. Lutz, *Peak Alignment of Urine NMR Spectra Using Fuzzy Warping*, J. Chem. Inf. Model. 46 (2006) 863–875. https://doi.org/10.1021/ci050316w.
- [141] P. Louangrath, Correlation Coefficient According to Data Classification, SSRN Electron. J. (2014). https://doi.org/10.2139/ssrn.2417910.
- [142] X.-L. CHU, J.-Y. LI, P. CHEN, Y.-P. XU, Algorithms, Strategies and Application Progress of Spectral Searching Methods, Chinese J. Anal. Chem. 42 (2014) 1379– 1386. https://doi.org/10.1016/S1872-2040(14)60768-4.

- [143] F. Zapata, C. García-Ruiz, *The discrimination of 72 nitrate, chlorate and perchlorate salts using IR and Raman spectroscopy*, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 189 (**2018**) 535–542. https://doi.org/10.1016/j.saa.2017.08.058.
- [144] X. Tan, X. Chen, S. Song, A computational study of spectral matching algorithms for identifying Raman spectra of polycyclic aromatic hydrocarbons, J. Raman Spectrosc. 48 (2017) 113–118. https://doi.org/10.1002/jrs.4978.
- [145] H. Henschel, A.T. Andersson, W. Jespers, M. Mehdi Ghahremanpour, D. van der Spoel, Theoretical Infrared Spectra: Quantitative Similarity Measures and Force Fields, J. Chem. Theory Comput. 16 (2020) 3307–3315. https://doi.org/10.1021/acs.jctc.0c00126.
- [146] R.G. Brereton, Chemometrics: data driven extraction for science, 2nd ed., 2018.
- [147] R. De Maesschalck, D. Jouan-Rimbaud, D.L. Massart, *The Mahalanobis distance*, Chemom. Intell. Lab. Syst. 50 (2000) 1–18. https://doi.org/10.1016/S0169-7439(99)00047-7.
- [148] J. Shlens, *A Tutorial on Principal Component Analysis*, **(2014)**. http://arxiv.org/abs/1404.1100 (accessed May 31, 2016).
- [149] A. Tharwat, T. Gaber, A. Ibrahim, A.E. Hassanien, *Linear discriminant analysis: A detailed tutorial*, AI Commun. 30 (2017) 169–190. https://doi.org/10.3233/AIC-170729.
- [150] A. Biancolillo, F. Marini, *Chemometric Methods for Spectroscopy-Based Pharmaceutical Analysis*, Front. Chem. 6 (2018). https://doi.org/10.3389/fchem.2018.00576.
- [151] R. Brereton, *One-class classifiers*, J. Chemom. 25 (**2011**) 225–246. https://doi.org/10.1002/cem.1397.
- [152] C. Ferri, P.A. Flach, J. Hernández-Orallo, Improving the AUC of Probabilistic Estimation Trees BT, in: N. Lavrač, D. Gamberger, H. Blockeel, L. Todorovski (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2003: pp. 121–132. https://doi.org/10.1007/978-3-540-39857-8_13.
- [153] F. Provost, P. Domingos, Tree Induction for Probability-Based Ranking, Mach.

- Learn. 52 (2003) 199–215. https://doi.org/10.1023/A:1024099825458.
- [154] M.I. López, M.P. Callao, I. Ruisánchez, A tutorial on the validation of qualitative methods: From the univariate to the multivariate approach, Anal. Chim. Acta. 891 (2015) 62–72.
- [155] T. Fawcett, *An introduction to ROC analysis*, Pattern Recognit. Lett. 27 (**2006**) 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.
- [156] L. Cuadros-Rodríguez, E. Pérez-Castaño, C. Ruiz-Samblás, Quality performance metrics in multivariate classification methods for qualitative analysis, TrAC - Trends Anal. Chem. 80 (2016) 612–624. https://doi.org/10.1016/j.trac.2016.04.021.
- [157] C. de S. Gondim, R.G. Junqueira, S.V.C. de Souza, M.P. Callao, I. Ruisánchez, Determining performance parameters in qualitative multivariate methods using probability of detection (POD) curves. Case study: Two common milk adulterants, Talanta. 168 (2017) 23–30. https://doi.org/10.1016/j.talanta.2016.12.065.
- [158] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Chapter 36 Multivariate calibration, in: B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J.B.T.-D.H. in S. and T. Smeyers-Verbeke (Eds.), Handb. Chemom. Qualimetrics Part B, Elsevier, 1998: pp. 349–381. https://doi.org/10.1016/S0922-3487(98)80046-4.
- [159] A. Biancolillo, F. Marini, *Chemometric Methods for Spectroscopy-Based Pharmaceutical Analysis*, Front. Chem. 0 (2018) 576.
- [160] T. Næs, T. Isaksson, T. Fearn, T. Davies, A user-friendly guide to multivariate calibration and classification, in: **2002**.
- [161] H.O.A. Wold, Soft modelling: The Basic Design and Some Extensions, in: 1982.
- [162] M. Andersson, *A comparison of nine PLS1 algorithms*, J. Chemom. 23 (**2009**) 518–529. https://doi.org/10.1002/cem.1248.
- [163] S. Wold, M. Sjöström, L. Eriksson, *PLS-regression: a basic tool of chemometrics*, Chemom. Intell. Lab. Syst. 58 (2001) 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.
- [164] M. Bonet, Progresses in new carbon platgorms: graphene and carbon dots.

- Application to electrochemical sensors and (bio)electronic tongues), Universitat Autònoma de Barcelona, **2022**.
- [165] 6.7.7. How the PLS model is calculated Process Improvement using Data, (n.d.). https://learnche.org/pid/latent-variable-modelling/projection-to-latent-structures/how-the-pls-model-is-calculated (accessed May 22, 2022).
- [166] ICH, ICH Q2 (R2) Validation of Analytical Procedures, (2022).
- [167] ICH, Analytical Procedure Development Q14, (2022).
- [168] U.S. Pharmacopoeia, USP 40-NF 35 < 1039 > Chemometrics, (2016).
- [169] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, N. Pou, *Influence of the procedure used to prepare the calibration sample set on the performance of near infrared spectroscopy in quantitative pharmaceutical analyses*, Analyst. 126 (2001) 1129–1134. https://doi.org/10.1039/b102090k.
- [170] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, C. De La Pezuela, Strategies for constructing the calibration set in the determination of active principles in pharmaceuticals by near infrared diffuse reflectance spectrometry, Analyst. 122 (1997) 761–765. https://doi.org/10.1039/a700630f.
- [171] M.G. Ierapetritou, R. Ramachandran, Process Simulation and Data Modeling in Solid Oral Drug Development and Manufacture, **2016**.
- [172] M.A. Bezerra, R.E. Santelli, E.P. Oliveira, L.S. Villar, L.A. Escaleira, *Response surface methodology (RSM) as a tool for optimization in analytical chemistry*, Talanta. 76 (**2008**) 965–977. https://doi.org/10.1016/j.talanta.2008.05.019.
- [173] L. Vera Candioti, M.M. De Zan, M.S. Cámara, H.C. Goicoechea, *Experimental design and multiple response optimization. Using the desirability function in analytical methods development*, Talanta. 124 (2014) 123–138. https://doi.org/10.1016/j.talanta.2014.01.034.
- [174] Y. Xu, R. Goodacre, On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning, J. Anal. Test. 2 (2018) 249–262. https://doi.org/10.1007/s41664-018-0068-2.

- [175] K. Varmuza, P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, **2016**.
- [176] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap, 1993.
- [177] R.W. Kennard, L.A. Stone, *Computer Aided Design of Experiments*, Technometrics. 11 (**1969**) 137–148. https://doi.org/10.1080/00401706.1969.10490666.
- [178] M.M. Sena, M.R. Almeida, J.W.B. Braga, R.J. Poppi, Multivariate Statistical Analysis and Chemometrics, in: Spectrosc. Methods Food Anal., 1st ed., CRC Press, 2017: pp. 273–314. https://doi.org/10.1201/9781315152769-10.
- [179] N.M. Faber, R. Rajkó, *How to avoid over-fitting in multivariate calibration—The conventional validation approach and an alternative*, Anal. Chim. Acta. 595 (**2007**) 98–106. https://doi.org/10.1016/j.aca.2007.05.030.
- [180] P. Munson, D. Rodbard, Statistical comparison of two assay methods (weighted linear regression with errors in x and y), in: Qual. Cntrol Clin. Endocrinol., **1979**: pp. 89–98.
- [181] W.A. Sadler, *Joint parameter confidence regions improve the power of parametric regression in method-comparison studies*, Accredit. Qual. Assur. 15 (**2010**) 547–554. https://doi.org/10.1007/s00769-010-0674-9.



1. OBJETIVOS

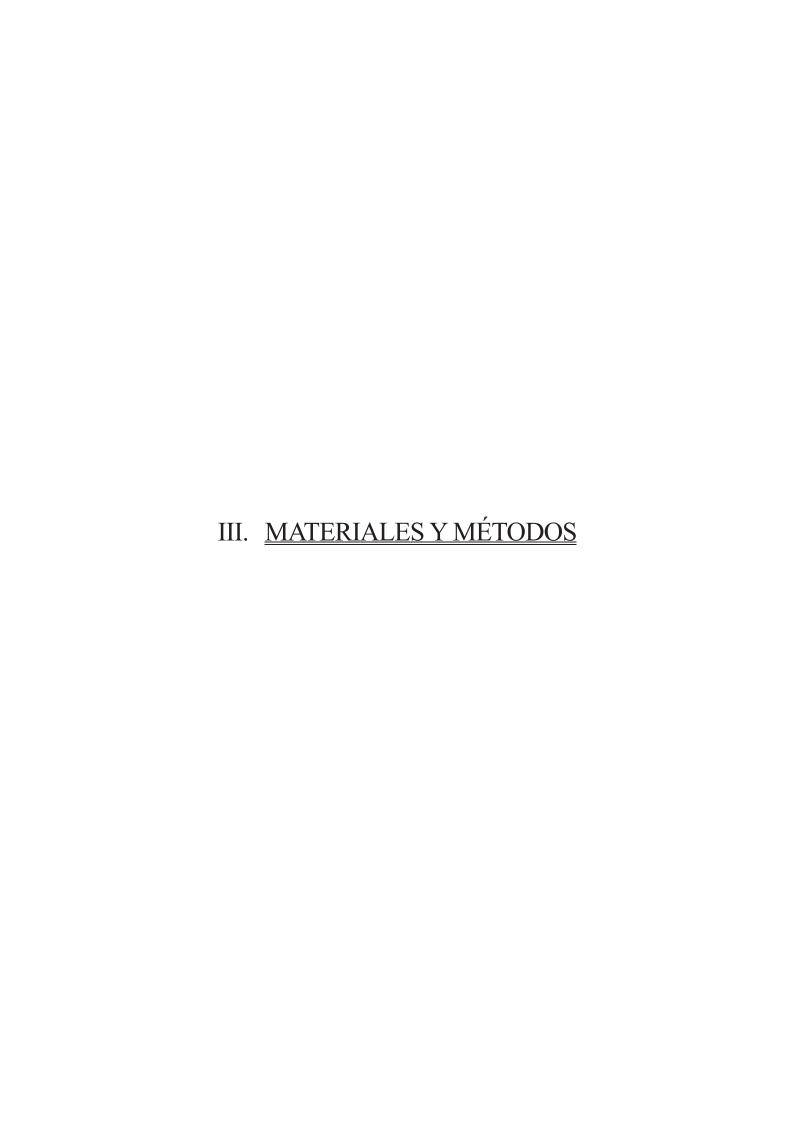
El objetivo principal de esta tesis es ofrecer una alternativa a los métodos tradicionales de análisis de identidad y cuantificación de los productos farmacéuticos mediante la aplicación de técnicas espectroscópicas y el uso de la Quimiometria. Estas técnicas ofrecen grandes ventajas en comparación a los métodos tradicionales como la rapidez de los resultados, la no destrucción de la muestra, y la no necesidad de disolventes ni de preparación de la muestra a analizar. A raíz del objetivo principal, se han creado distintos subobjetivos dependiendo del propósito: identificación o cuantificación.

1) Identificación.

- Evaluación de la idoneidad de dos instrumentos NIR y Raman portátiles para la identificación y discriminación entre fabricantes de un mismo producto farmacéutico.
- Puesta a punto de una biblioteca de espectros NIR para la identificación y cualificación de 398 materiales de partida recibidos en una industria farmacéutica.
- Diseño de un análisis de riesgos para poder asociar, a cada material de partida incluido en la biblioteca de espectros NIR, el riesgo de obtener un falso negativo en su identificación. A partir de los resultados se propondrán distintas acciones mitigadoras del riesgo.

2) Cuantificación.

- Puesta a punto de un método para determinar el contenido de API en un gel comercial mediante un instrumento Raman portátil y un análisis univariante.
- Desarrollo de un método multivariable para la cuantificación del contenido de dos APIs en un preparado farmacéutico con datos NIR obtenidos con un equipo de sobremesa y uno portátil.
- Optimización de un proceso de análisis Raman para la obtención de la uniformidad de contenido de comprimidos, utilizando un mapeo de puntos en la superficie de los comprimidos,



1. INSTRUMENTACIÓN

En esta sección se presentan los distintos instrumentos utilizados en esta tesis.

1.1. Espectroscopia Raman

1.1.1. Espectrómetro MultiRaman

El equipo MultiRaman de Bruker (Bruker, Billarica, MA, USA) es un espectrómetro equipado con tecnología de transformada de Fourier (FT) equipado con un láser Nd:YAG como fuente de excitación (1064 nm), que permite reducir los efectos derivados de la fluorescencia (**Figura III.1**). Además, el equipo contiene un recipiente acoplado al instrumento para introducir nitrógeno líquido, lo que permite tener controlada la temperatura del detector de Ge, proporcionando así una mayor sensibilidad (menor ruido de fondo). La configuración del equipo permite obtener información en el intervalo espectral de 3600 – 50 cm⁻¹.

La irradiación de la muestra se realiza en el compartimento llamado *autosampler*, el cual lleva integrado una platina y una cámara. Así, la adquisición de los espectros se realiza colocando la muestra encima de la platina y ajustando la posición de irradiación del láser desde el propio software mediante una previsualización de la cámara. De esta forma, se pueden realizar mediciones en distintas posiciones de la muestra (por mapeo) controlando los ejes x-y-z de la platina. Además, uno de los accesorios del equipo permite obtener el espectro de un producto sólido añadiendo una pequeña cantidad (mg) en una cubeta con la ayuda de otros accesorios que hacen la función de embudo.

Antes de empezar la adquisición de los espectros, es posible ajustar distintos parámetros con el fin de obtener bandas bien definidas. Se puede modificar tanto la potencia del láser como el número de barridos y el intervalo de longitud de onda. Además, es importante ajustar el eje z de la platina hasta que veamos en la cámara la muestra bien enfocada. Una vez ajustados los parámetros, se irradiará la muestra con el láser, el cual tiene un haz de iluminación de 500 µm de diámetro.



Figura III.1. Instrumentos Raman usados en la tesis. A la izquierda equipo de sobremesa, y a la derecha equipo portable.

1.1.2. Espectrómetro i-Raman ® Plus

El equipo i-Raman ® Plus de B&W TEK (BW TEK, Newark, DE, USA) es un espectrómetro Raman portable con unas dimensiones de 16,9 cm x 33,3 cm x 24,2 cm (**Figura III.1**). En este caso, la longitud de onda de excitación usada es de 785 nm, con una potencia de láser máxima de 320 mW. El detector es un dispositivo de carga acoplada de alta eficiencia cuántica (*CCD* por sus siglas en inglés), enfriado a -2°C y con alto rango dinámico. La configuración del equipo permite obtener información en el intervalo 3350 – 150 cm⁻¹, con una resolución inferior a 4,5 cm⁻¹.

La irradiación de la muestra se realiza a través de una fibra óptica que se acopla al instrumento. En esta tesis, se han usado dos fibras ópticas distintas: la BAC101 y la BAC102 (**Figura III. 2**). La sonda BAC101 está diseñada para realizar análisis in-situ, ya que el cuerpo es de acero inoxidable, lo que le permite que esté en contacto directo con la muestra. El diámetro de la sonda es de 12,7 mm y tiene una longitud de 203,2 mm; la fibra de excitación tiene un diámetro de núcleo de 105 μm y la fibra colectora de 200 μm con una distancia de trabajo de 0,8 mm en agua. Por otro lado, la sonda BAC102 no puede estar en contacto directo con la muestra, por ello lleva acoplado un regulador de distancia, para asegurar un mejor enfoque minimizando la contaminación de la lente. El diámetro de la sonda es de 9,42 mm con una longitud de 76,2 mm. Las fibras de excitación y colectora tienen el mismo diámetro que la sonda BAC101 y una distancia de trabajo de 5,5 mm.

Para la adquisición de los espectros, es necesario acoplar al instrumento i-Raman la sonda que vaya a usarse y ajustar la posición de ésta. Antes de la medida, se pueden ajustar tres parámetros mediante el software: la potencia del láser, el número de barridos, y el tiempo de integración (de 100 ms a 30 min). Además, una vez adquirido el espectro de la primera muestra, hay que realizar un espectro de la *dark current* para poder sustraer la influencia de la luz ambiental en los registros. Una vez finalizadas las medidas, se puede ajustar el intervalo de números de onda para guardar los datos.

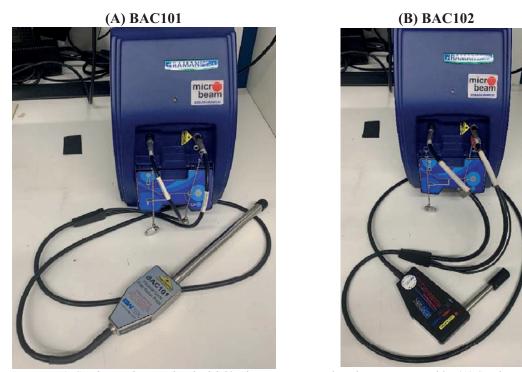


Figura III. 2. Sondas usadas para la adquisición de espectros con el equipo Raman portable. (A) Sonda BAC101; (B) Sonda BAC102.

1.2. Espectroscopia NIR

1.2.1. Espectrómetro Model 5000

Espectrómetro Model 5000 de FOSS (FOSS NIRSystem, Silver Spring, MD, USA) equipado con un módulo *Rapid Content Analyzer* (Figura III. 3). La radiación NIR obtenida está comprendida entre 1100 y 2500 nm, con una resolución de 2 nm. Para este instrumento, cada espectro es el resultado del promedio de 32 barridos.

Los espectros NIR de los comprimidos se adquirieron posicionando cada comprimido directamente sobre la ventana del instrumento, ajustando el iris para controlar la apertura de

la ventana por donde entra el haz de luz. Previamente al registro de las muestras, se adquiere un espectro del blanco, constituido por una placa de cerámica que incorpora el instrumento.





Figura III. 3. Instrumentos NIR usados en la tesis. A la izquierda equipo de sobremesa, y a la derecha equipo portable.

1.2.2. Espectrofotómetro MicroNIR

Espectrofotómetro MicroNIRTM OnSite (VIAVI, Milpitas, California, USA) (**Figura III. 3**). Tiene una fuente de irradiación formada por un par de lámparas de tungsteno integradas y como sistema de dispersión un filtro de variable lineal (LVF, siglas del inglés *Linear Variable Filter*), que se encuentra acoplado a un detector de diodos lineal. Con esta configuración, la radiación NIR está comprendida entre 950 nm y 1650 nm, con una resolución espectral de 6 nm.

En este caso, para la adquisición de los espectros, se colocó el instrumento dentro de una caja para evitar así la luz ambiental. Previamente a la adquisición de las muestras, se debe ajustar el tiempo de integración y el número de barridos y, realizar dos espectros: *dark current* y espectro de referencia (estándar con un 99% de reflectancia difusa).

1.2.3. Espectrómetro MPA-NIR

Espectrómetro MPA T (Bruker, Billarica, MA, USA) ubicado en las instalaciones de una industria farmacéutica. Tiene una fuente de irradiación formada por una lámpara de halógeno. Equipado con tecnología de FT con una esfera integradora y una sonda de fibra

óptica para la adquisición de los espectros. Con esta configuración, la radiación NIR está comprendida entre 12500 – 3750 cm⁻¹, con una resolución de 8 cm⁻¹.

Todas las medidas NIR usadas con este instrumento se han realizado en una cabina de muestreo localizada en el almacén de recepción de muestras de la industria. Una vez por semana, el equipo pasa una prueba de cualificación del funcionamiento (PQ, siglas del inglés *Performance Qualification*) y cada día se registra un espectro de referencia antes de empezar las medidas.

1.3. <u>Técnica instrumental de referencia. HPLC.</u>

La determinación del contenido de los distintos principios activos analizados se ha realizado mediante un cromatógrafo HPLC Agilent 1100 (Agilent, Santa Clara, CA, USA). El equipo cuenta con una bomba binaria G1312A, muestreador automático e inyector G1313A y un detector de diodos en línea G1315B DAD.

Para el desarrollo de este trabajo, se han usado dos métodos de referencia basados en HPLC. El primero para la cuantificación de Tramadol y Paracetamol en comprimidos farmacéuticos; y el segundo para la determinación del contenido de Ibuprofeno en muestras de gel comercial.

1.4. Otros instrumentos

Otros instrumentos que se han utilizado para el desarrollo de la tesis han sido una prensa y un homogeneizador (**Figura III. 4**).

- Prensa (Perkin Elmer, Waltham, MA, USA). Instrumento usado para compactar mezclas y así obtener comprimidos de dimensiones 1,3 mm aproximadamente mediante unos moldes específicos.
- Homogeneizador (Turbula T2C, WAB, Basel, Switzerland). Instrumento usado para la homogeneización de mezclas mediante rotaciones en los 3 ejes.





Figura III. 4. Otros instrumentos usados para el desarrollo del proyecto. A la izquierda, la prensa, y a la derecha el homogeneizador.

2. REACTIVOS

Durante el desarrollo de los distintos proyectos, ha sido necesario el uso de reactivos y patrones. En la **Tabla III. 1** se especifican todos los reactivos y patrones usados, así como la marca comercial, riqueza y uso principal.

Tabla III. 1. Descripción de los reactivos y productos usados.

Reactivo / Producto	Marca y calidad	Uso	
Acetonitrilo	Fisher (HPLC)	Disolvente HPLC	
Metanol	Fisher, ≥ 98% (HPLC)	Disolvente HPLC	
Ácido fosfórico Panreac 85%. PA,		HPLC	
Ihunrafana (IDII)	Sigma-Aldrich, Certified	Patrón	
Ibuprofeno (IBU)	Reference Material		
Sal sódica de IBU (IBUNa)	Sigma-Aldrich, ≥ 98%	Patrón	
Paragatamal (PCT)	Sigma-Aldrich, EP Reference	Patrón	
Paracetamol (PCT)	Standard		
Lisinopril (LIS)	Sigma-Aldrich, ≥ 98%	Patrón	
Clorhidrato de tramadol	Sigma-Aldrich, ≥ 99.5	Patrón Excipiente	
(TRAHCI)	Sigilia-Aluricii, 2 99.3		
Almidón de maíz	Sigma-Aldrich. PA		

3. MÉTODOS

3.1. Determinación de TRA y PCT en comprimidos

La cuantificación del contenido de TRA y PCT en comprimidos farmacéuticos se basó en los resultados de un método validado y publicado [1]. Sin embargo, como los primeros resultados obtenidos no produjeron valores de recuperación adecuados debido a una deficiente disolución de ambos APIs, se introdujeron algunos cambios para poder mejorarlo. Las condiciones finales usadas en el método HPLC se muestran en la **Tabla III. 2**. El análisis HPLC se realizó con una columna Symmetry-Shield RP18 (Waters) de 5 μm de tamaño de partícula, con unas dimensiones de 250 mm x 4,6 mm.

Tabla III. 2. Descripción de las condiciones de trabajo utilizadas en la determinación de TRA y PCT en comprimidos mediante HPLC.

Fase móvil:	40:60 v/v Metanol-Agua Fase estacionaria:		Intersil
	(ajustada a pH 3,4 con ácido		C18
	fosfórico)		
Flujo:	1 mL/min	Inyección:	20 μm
Longitud de	224 nm	Tiempo de análisis:	10 min
Ü			

La determinación del contenido de ambos APIs en los comprimidos farmacéuticos se realizó después de moler el comprimido, pesando aproximadamente 200 mg del polvo homogeneizado que se disolvió en metanol, enrasando en un matraz de 100 mL. Una vez disuelto, se trasvasaron 2 mL a un matraz aforado de 25 mL y se enrasó con la fase móvil para obtener una disolución de, aproximadamente, 130 y 15 ppm de PCT y TRA. Antes de la inyección, se filtraron todas las disoluciones con un filtro de membrana de 0.45 μm.

Para poder determinar la concentración de los APIs, se construyeron dos curvas de calibración, cada una con siete valores de concentración en el rango de 5-23 ppm para el TRA y de 60-170 ppm para el PCT. Los tiempos de retención para TRA y PCT fueron de 2,1 y 4,2 min, respectivamente. En la **Figura III. 5**, se muestra un cromatograma de uno de los comprimidos analizados.

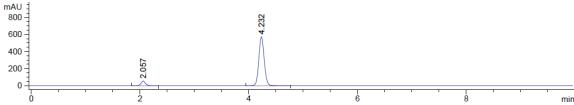


Figura III. 5. Cromatograma obtenido para un comprimido comercial.

3.2. Determinación de IBU en geles

La determinación del contenido de IBU en los geles comerciales es una modificación de un método de HPLC publicado [2]. La **Tabla III. 3** muestra las condiciones de trabajo utilizadas. En este caso, se usó una columna Symmetry C18 (Waters) de 3,5 μm, con dimensiones 150 mm x 4,6 mm. Antes de la inyección, se filtraron todas las disoluciones con un filtro Nylon de 0,45 μm (Fisherbrand).

Tabla III. 3. Descripción de las condiciones de trabajo utilizadas en la determinación de IBU en muestras de gel mediante HPLC.

Fase móvil:	40:60 v/v Disolución acuosa (20 Fase estacionaria:		Intersil
	mM ácido fosfórico) -Acetonitrilo		C18
Flujo:	1.5 mL/min	Inyección:	20 μL
Longitud de	220 nm	Tiempo de análisis:	7 min
onda:			

En la **Figura III.** 6 se muestra un ejemplo de un cromatograma de una muestra comercial de gel. El pico del ibuprofeno aparece a un tiempo de retención de 4 min, determinado después de la inyección del patrón.

Representando el área del pico frente a la concentración, se preparó una recta de calibrado inyectando 7 disoluciones en un intervalo de concentración entre 6,5 y 42 ppm, preparadas a partir de un patrón de IBU (≥ 98%, Sigma Aldrich).

Para realizar las adiciones estándar, se prepararon diariamente disoluciones a partir de sal sódica de ibuprofeno, IBUNa. Se pesaron 1,25 g de IBUNa y se diluyeron en un matraz de 5 mL con agua milliQ. Una vez disuelto, se preparó una disolución de 250 ppm en un matraz de 1 L con agua milliQ y, finalmente, se diluyó a 25 ppm con fase móvil en un matraz de 20 mL. Esta última disolución se inyectó en el HPLC para determinar la concentración real de IBUNa utilizado.

Para la cuantificación de las muestras, se pesó 1 g de producto y se disolvió en 50 mL de fase móvil. Con el fin de disolver bien el gel, la disolución estuvo 20 minutos en un baño de ultrasonidos. Posteriormente, se preparó una disolución de aproximadamente 22 ppm con fase móvil y se inyectó en el HPLC.

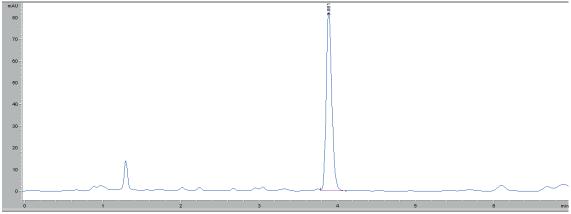


Figura III. 6. Cromatograma obtenido para una muestra de gel comercial.

4. PROGRAMAS PARA LA ADQUISICIÓN Y TRATAMIENTO DE LOS DATOS

Los espectros Raman, NIR y los datos de referencia fueran tratados usando los siguientes programas informáticos:

- OPUS (Bruker, Billerica, MA, USA) en sus versiones 6.5, 7.2 y 7.8. Programa diseñado para el control del espectrómetro MultiRaman y MPA-NIR. Permite el registro de los espectros, así como su visualización. Además, gracias al paquete incorporado Ident, permite la creación y desarrollo de bibliotecas de espectros.
- BW Spec (BWTEK ahora Metrohm). Programa diseñado para el control del Raman portable. Permite la adquisición de espectros, así como su visualización. Se puede optar por una adquisición en continuo, pudiendo ver a tiempo real el comportamiento de la muestra mediante los espectros adquiridos.
- VISION 2.5.1 (FOSS NIRSystems, Silver Spring, MD, USA). Programa diseñado para el control del espectrofotómetro NIR de sobremesa. Permite el registro de los espectros, su visualización, pretratamientos y la construcción de modelos, tanto de identificación como de cuantificación.
- MicroNIRTM ProSite v2.5. Programa diseñado para la adquisición de los espectros
 NIR usando el instrumento portátil mediante una conexión de puerto USB. Permite

- la visualización de los espectros, así como la aplicación de los pretratamientos espectrales más habituales (normalización y derivadas).
- The Unscrambler X 10.5 (Trondheim, Norway). Programa diseñado para la aplicación de algoritmos quimiométricos. Permite la aplicación de pretratamientos y la construcción de modelos tanto de identificación como de cuantificación.
- SOLO v8.2.1 (Eigenvector Research Inc, WA, USA). Programa diseñado para la aplicación de algoritmos quimiométricos. Permite la aplicación de pretratamientos y la construcción de distintos tipos de modelos cualitativos y cuantitativos.
- MATLAB® R2014b (MathWorks®, Massachusetts, USA). Programa diseñado para programar distintas aplicaciones.
- ChemStation. Programa para la manipulación del instrumento HPLC.

5. BIBLIOGRAFIA

- [1] R.M. Kamble, S.G. Singh, S. Singh, Validated RP-HPLC method for simultaneous estimation of paracetamol and tramadol hydrochloride in a commercial tablet, J. Pharm. Res. 4 (2011) 4038–4040.
- [2] M.B. Alsirawan, M. Mohammad, B. Al-kasmi, K. Alhareth, M. El-Hammadi, Development and validation of a simple HPLC method for the determination of ibuprofen sticking onto punch faces, Int. J. Pharm. Pharm. Sci. 5 (2013) 227–231.

IV. <u>IDENTIFICACIÓN DE PRODUCTO</u> <u>ACABADO CON INSTRUMENTOS</u> <u>PORTÁTILES</u>

1. INTRODUCCIÓN

La identificación de un producto farmacéutico acabado se realiza con alguno de los siguientes objetivos: el control de calidad del fármaco, realizado por el propio fabricante, con el objetivo de comprobar que tiene unas características iguales a la de su producción habitual; la identificación del fabricante que ha producido un fármaco determinado y la identificación de fraudes. En esta tesis se abordan los dos primeros casos.

La metodología empleada para la identificación de un producto farmacéutico acabado, utilizando técnicas espectroscópicas, se basa en la aplicación de un modelo quimiométrico cualitativo a la muestra desconocida que se desea identificar. Este modelo se debe haber desarrollado usando fuentes fiables e incluyendo la mayor variabilidad posible relacionada con el producto a identificar; por ello, es de gran utilidad tener la información química y física del producto, ya que se pueden obtener espectros distintos del mismo medicamento comercializado por distintas marcas, debido al uso de diferentes materiales de partida o a un proceso de fabricación distinto.

Una alternativa para el aumento de la discriminación es la utilización de algoritmos que permiten construir modelos de clasificación parciales, con distintas características para cada uno de ellos, hasta conseguir la discriminación final. Este es el mecanismo utilizado con las llamadas "bibliotecas de espectros" y que se aplica habitualmente para la identificación del producto acabado producido por un fabricante. Se puede comprobar que es distinto a todos los demás productos fabricados en la factoría (identificación) y que tiene las mismas características que el producido habitualmente (cualificación). En este caso, los modelos quimiométricos se aplican comparando el espectro de la muestra desconocida con una base de datos que contiene, como una clase de datos individual, los espectros de cada uno de los productos fabricados habitualmente, incluyendo su correspondiente variabilidad. Esta base de datos se denomina una biblioteca de espectros. En este capítulo solo se ofrecen los resultados del uso de una biblioteca para el problema concreto estudiado, sin entrar en una descripción profunda respecto a su creación y utilización. Estos aspectos se discuten en los capítulos V y VI de esta tesis, donde se abordará su uso en una aplicación a escala industrial.

La utilización de instrumentos espectroscópicos portátiles o de mano son una clara apuesta para la identificación de producto acabado [1–3]. Algunas de las ventajas más

destacables de estos instrumentos son su portabilidad, preció más económico que los de sobremesa y su aplicabilidad tanto in-situ como in-line. Dadas sus ventajas, estudios recientes han evaluado su uso para aplicaciones cualitativas o cuantitativas en la industria farmacéutica [4–7].

El objetivo de este capítulo es evaluar el desempeño de dos instrumentos portátiles, que utilizan la espectroscopia NIR y Raman, en la identificación de comprimidos recubiertos y en la discriminación del origen de su fabricación, ya que hay siete industrias farmacéuticas que comercializan el mismo producto. Para poder realizar la evaluación, se han estudiado diferentes métodos de clasificación (SIMCA, PCA-LDA y PLS-DA) y, con cada instrumento, se han registrado espectros de los productos de cada fabricante para la construcción de los modelos de clasificación, comparando, finalmente, los resultados obtenidos. También se ha ensayado la discriminación de las muestras por fabricantes utilizando el software comercial OPUS 7.8 para crear una biblioteca de espectros. Los espectros del producto de cada fabricante forman cada una de las distintas clases de una biblioteca.

Aunque todos ellos utilizan variables latentes para su construcción, la diferencia entre los modelos SIMCA, PLS-DA, PCA-LDA y la biblioteca de espectros es la manera de construirlos. Como se ha descrito en el capítulo anterior, para SIMCA, el objetivo de calcular los PCs individuales de cada clase es crear un modelo individual para cada una de ellas y exponer sus diferencias, con lo que una nueva muestra se clasificará como perteneciente al modelo más cercano, aunque también puede clasificarse como distinta a todos los demás (no pertenece a ninguna de las clases consideradas). Tanto en PLS-DA como en PCA-LDA, las variables latentes se calculan, con distintos algoritmos, conjuntamente para todas las clases, creando fronteras en el espacio que permitan una separación directa entre ellas, por lo que cualquier muestra nueva se identifica, obligatoriamente, a un lado u otro de la frontera. El uso de una biblioteca de espectros combina ambos enfoques, permitiendo la separación de algunas clases del conjunto total y estudiar posteriormente las clases que no se han conseguido separar cambiando los parámetros del algoritmo. El procedimiento se puede repetir hasta que se consigue la separación completa de las clases. Cada vez que se repite este procedimiento, se trabaja con un conjunto distinto de datos; se dice que cambia el nivel de la biblioteca, permitiendo que una muestra nueva se identifique en cascada por los distintos niveles hasta hallar la clase a la que pertenece. También se puede obtener como resultado que la nueva muestra no se identifica con ninguna de las clases existentes.

2. MATERIALES

2.1. Muestras

El producto farmacéutico analizado es Alprazolam (ALP). ALP es una benzodiazepina que se usa comúnmente para los trastornos de ansiedad. Este producto, es comercializado por distintas industrias farmacéuticas en dosis de 0,25; 0,50; 1 y 2 mg del principio activo. En este caso, se han analizado los comprimidos de 1 mg de siete fabricantes distintos: Alter® (AL), Cinfa® (CI), Mabo® (MA), Mylan® (MY), Normon® (NO), Pensa® (PE) y Pfizer® (PF).

Dependiendo del fabricante, el producto acabado presenta diferencias tanto en el aspecto como en la composición. Un parámetro que tienen en común todos ellos, es que los comprimidos son de color azul, más o menos intenso, dependiendo de la marca. Esto sirve para poder discriminar fácilmente entre las distintas presentaciones disponibles del mismo principio activo, ya que a cada concentración de API se le asigna un color en particular. El color de los comprimidos no tiene gran impacto en caso de la espectroscopia NIR, sin embargo, tiene un impacto directo en Raman debido a los efectos de la fluorescencia. En la **Tabla IV. 1** se muestran las características físicas (aspecto, color) de los comprimidos, así como información relacionada con su fabricación y comercialización. La **Tabla IV. 2**¡Error! No se encuentra el origen de la referencia. presenta su composición, para poder englobar las diferencias entre fabricantes.

Tabla IV. 1. Características físicas de los comprimidos recubiertos de cada uno de los fabricantes.

Marca	Imagen	Responsable fabricación	Titular de la comercialización
AL	90	Lacer S.A	Laboratorios Alter, S.A
CI		Laboratorios Cinfa, S.A	Laboratorios Cinfa, S.A

Marca	Imagan	Responsable	Titular de la
Marca	Imagen	fabricación	comercialización
MA		Laboratorios Cinfa, S.A	MABO-FARMA
		McDermott Laboratories	Mylan Pharmaceuticals, S.
MY		Ltd. t/a Gerard	L
		Laboratories	L
NO		Laboratorios Normon S.	Laboratorias Namasas C. A
NO		A	Laboratorios Normon S. A
PE		Laboratorios Dr. Esteve,	Pensa Pharma, S. A
112		S.A.	1 01100 1 11011110, 0, 11
PZ		Pfizer Italia S.r.l	Pfizer S. L

Como se puede observar en las Tablas, no todos los comprimidos tienen la misma forma ni el mismo tono azul. El colorante usado para proporcionar el tono azulado, en casi todas las marcas, es la Indigotina E-132, mientras que las marcas AL y PF usan distintos colorantes. Para poder apreciar mejor estas diferencias, en la **Figura IV. 1** se muestran una imagen de todos los comprimidos.



Figura IV. 1. Comparación del color para todos los comprimidos. El orden es el siguiente: MY, NO, PE, CI, AL, MA y PF.

A partir de la información proporcionada en la ¡Error! No se encuentra el origen de la referencia, se puede ver que hay excipientes que están presentes en todas las formulaciones, pero que hay otros que son característicos de una única marca. Por ejemplo, la marca AL es la única que presenta en su formulación Talco. Otro aspecto importante que resaltar es que la fabricación de CI y MA se realiza por el mismo laboratorio (Laboratorios Cinfa, S.A), aunque la comercialización es propia de cada marca. En consecuencia, su formulación es idéntica, por lo que la discriminación entre ambas marcas será imposible.

IDENTIFICACIÓN DE PRODUCTO ACABADO CON INSTRUMENTOS PORTÁTILES

Tabla IV. 2. Composición de los comprimidos recubiertos de cada uno de los fabricantes. Se señala con una X en caso de que el excipiente forme parte de la formulación.

	AL	CI	MA	MY	NO	PE	PF
Lactosa monohidrato	X	X	X	X	X	X	X
Celulosa micocristalina	X	X	X	X		X	X
Almidón de maíz	X	X	X	X	X	X	X
Talco	X						
Carboximetilalmidón sódico (Tipo A)	X	X	X				
Estearato de magnesio	X	X	X	X	X	X	X
Povidona			X	X			
Sílice coloidal anhidra		X	X	X			
Docusato sódico		X	X	X		X	
Benzoato sódico		X	X	X			
Óxido de aluminio hidratado		X	X				
Indigotina (E-132)		X	X	X	X	X	
Almidón de patata				X			
Polividona					X		
Dióxido de sílice coloidal						X	X
Laca azul de aluminio							X
Dioctil sulfosuccinato sódico 85 % con benzoato							V
sódico 15%							X
Amaranto (E-123)	X						
Colorante azul patente V (E-131)	X						

2.2. Adquisición de espectros

Los espectros de las muestras estudiadas en este apartado han sido registrados usando los instrumentos portátiles NIR (MicroNIR) y Raman (i-Raman plus) presentados en el capítulo III (Materiales y métodos). En el caso del instrumento Raman, las lecturas se realizaron usando la sonda BAC102.

La adquisición de los espectros de las muestras asignadas al conjunto de calibración se realizó registrando tres comprimidos de cada marca en cuatro días distintos, y siempre realizando un total de tres espectros por comprimido. Para PF sólo se han usado 31 espectros en el conjunto de calibración, debido a errores de adquisición en 5 espectros registrados. En el caso del conjunto de validación, se adquirieron espectros de 10 comprimidos extra para cada marca al cabo de un mes, realizando también tres mediciones por pastilla. En el caso de los espectros adquiridos con el MicroNIR, las repeticiones se lograron rotando el comprimido, mientras que, para el i-Raman Plus, cambiando la posición de la sonda.

Los parámetros de registro para el MicroNIR han sido un tiempo de integración de 10 ms y un total de 100 barridos. En el caso del i-Raman Plus, el tiempo de integración ha sido de 2000 ms, con un total de 1 barrido y la potencia del láser de un 80%, que equivale aproximadamente a 277 mW.

3. RESULTADOS Y DISCUSIÓN

3.1. Análisis exploratorio de los datos

Una vez adquiridos los espectros, el primer paso para construir un modelo cualitativo es realizar un análisis exploratorio de los datos. En la **Figura IV. 2**, se pueden observar los espectros sin procesar, obtenidos con ambas técnicas, de todas las marcas comerciales analizadas. Como se ha comentado anteriormente, el primer aspecto importante a resaltar es la gran influencia del color de los comprimidos en los espectros Raman. El color influye en la fluorescencia del espectro e implica líneas de base no horizontales de los espectros Raman, con distinta pendiente según el color. Por el contrario, en los datos sin procesar obtenidos con NIR, a simple vista no pueden apreciarse diferencias significativas entre marcas. A la vista de los espectros obtenidos, es necesaria la aplicación de pretratamientos espectrales con el objetivo de reducir ruido, corregir la dispersión de la luz en NIR o la fluorescencia en Raman y maximizar las diferencias entre marcas para su posible discriminación.

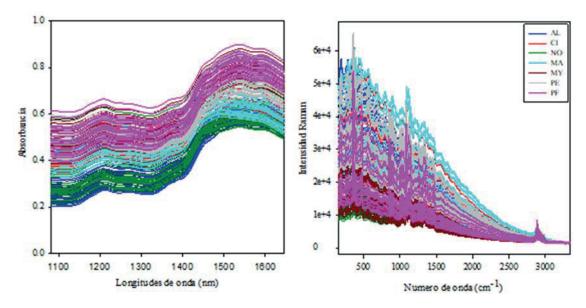


Figura IV. 2. Espectros NIR (izquierda) y Raman (derecha), sin procesar, de todas las marcas comerciales y en todo el rango espectral, para ambos instrumentos.

3.2. Modelos construidos con los datos NIR

Para la construcción de los modelos de clasificación, el primer paso es la selección del rango de longitudes de onda a utilizar, así como el pretratamiento aplicado a los espectros. Por ello, se han evaluado distintos pretratamientos en el intervalo de 1050 – 1676 nm, estudiando dos propiedades analíticas de un método cualitativo: la especificidad (ESP) y la sensibilidad (SEN). Estos parámetros evalúan la capacidad del método para identificar si cuando una muestra se compara con una clase, realmente no pertenece a ella (especificidad), o sí que pertenece a ella (sensibilidad). A partir de los valores obtenidos se ha escogido un único pretratamiento y se ha optimizado. En la construcción de los modelos basados en SIMCA y PLS-DA, para realizar la comparación de los distintos pretratamientos usados, se han escogido los componentes principales recomendados por el propio software (The Unscrambler X 10.5). Por otro lado, en el caso de los modelos PCA-LDA, se han escogido 7 componentes principales para todos ellos.

Los resultados de los parámetros de ESP y SEN se encuentran en la **Tabla IV. 3** y en la **Tabla IV. 4**. Cuanto más cercano a 1 sean los valores de ESP y SEN, mejor es el comportamiento del modelo de clasificación.

Tabla IV. 3. Valores de ESP y SEN obtenidos al aplicar distintos pretratamientos para cada marca con los modelos cualitativos construidos con datos NIR.

		SIN	ICA	PI	LS-DA		PCA-	-LDA
Pretratam	niento	ESP	SEN	Nº de factores	ESP	SEN	ESP	SEN
SNV				13				
A	L	1,00	0,97		1,00	1,00	1,00	1,00
C	I	0,83	0,11		1,00	0,97	0,97	0,72
M	[A	0,85	0,00		1,00	1,00	0,95	0,83
M	ΙΥ	1,00	0,97		1,00	1,00	1,00	1,00
N	O	1,00	0,97		1,00	1,00	1,00	1,00
P	Е	0,85	0,06		1,00	0,97	1,00	0,92
Pl	F	0,84	0,000		1,00	1,00	0,99	0,97
1D*				9				
A	L	1,00	1,00		1,00	1,00	1,00	1,00
C	I	1,00	1,00		1,00	1,00	1,00	1,00
M	ÍΑ	1,00	1,00		1,00	1,00	1,00	1,00
M	ΙΥ	1,00	1,00		1,00	0,97	1,00	1,00
N	O	1,00	1,00		1,00	1,00	1,00	1,00
P	E	0,95	1,00		1,00	0,97	1,00	1,00
Pl	F	1,00	0,68		1,00	1,00	1,00	1,00
2D*				8				
A	L	1,00	1,00		1,00	1,00	1,00	1,00
C	I	0,98	0,92		1,00	0,97	1,00	1,00
M	[A	0,99	0,89		1,00	1,00	1,00	1,00
M	ΙΥ	1,00	1,00		1,00	0,94	1,00	1,00
N	O	1,00	1,00		1,00	1,00	1,00	1,00
P	Е	0,97	1,00		0,99	1,00	1,00	1,00
P	F	1,00	0,81		1,00	0,97	1,00	1,00
SNV + 1	D*			12				
A	L	1,00	0,97		1,00	1,00	1,00	1,00
C	I	0,83	0,08		1,00	0,97	0,97	0,72
M	[A	0,84	0,00		1,00	1,00	0,95	0,81
M	ΙΥ	1,00	0,97		1,00	1,00	1,00	1,00
N	O	1,00	0,97		1,00	1,00	1,00	1,00
Pl	E	0,85	0,00		1,00	0,89	1,00	0,94
Pl	F	0,83	0,00		0,99	0,94	0,99	0,97

	SIN	ICA	PI	LS-DA		PCA-LDA		
Pretratamiento	ESP SEN		Nº de factores	ESP	SEN	ESP	SEN	
SNV + 2D*			9					
AL	1,00	1,00		1,00	1,00	1,00	1,00	
CI	0,83	0,14		1,00	0,94	0,98	0,78	
MA	0,85	0,00		0,99	1,00	0,96	0,89	
MY	1,00	1,00		1,00	1,00	1,00	1,00	
NO	1,00	0,97		1,00	1,00	1,00	1,00	
PE	0,87	0,00		1,00	1,00	1,00	0,94	
PF	0,83	0,10		1,00	0,97	0,99	0,97	

Tabla IV. 4. Valores globales (considerando todas las marcas) de ESP y SEN al aplicar distintos pretratamientos a los modelos cualitativos construidos con datos NIR.

	SIN	ICA	PLS		PCA-LDA		
Pretratamiento	ESP SEN		Nº de factores	ESP	SEN	ESP	SEN
SNV	0,91	0,44	13	1,00	0,99	0,99	0,92
1D*	0,99	0,95	9	1,00	0,99	1,00	1,00
2D*	0,99	0,94	8	1,00	0,98	1,00	1,00
SNV + 1D*	0,91	0,43	12	1,00	0,97	0,99	0,92
SNV + 2D*	0,91	0,46	9	1,00	0,99	0,99	0,94

Como se puede observar en las **Tablas V.3** y **V.4**, los modelos construidos con primera y segunda derivada presentan una mayor especificidad y sensibilidad que el resto de los modelos construidos aplicando también SNV. Por este motivo, se descartan los modelos usando SNV. Respecto a los resultados obtenidos con las derivadas, tanto los valores de especificidad como sensibilidad son elevados y parecidos. Finalmente, se decide usar el pretratamiento basado solo en la aplicación de una segunda derivada porque necesita un modelo más sencillo (8 factores en lugar de 9) en el caso del modelo PLS-DA.

3.2.1. Construcción de los modelos NIR

La **Figura IV. 3** muestra los espectros del conjunto de calibración tras la aplicación de una segunda derivada (con 11 puntos de suavizado) en el rango 1050-1676 nm, así como la representación de los *scores*, para poder visualizar las diferencias entre las marcas comerciales.

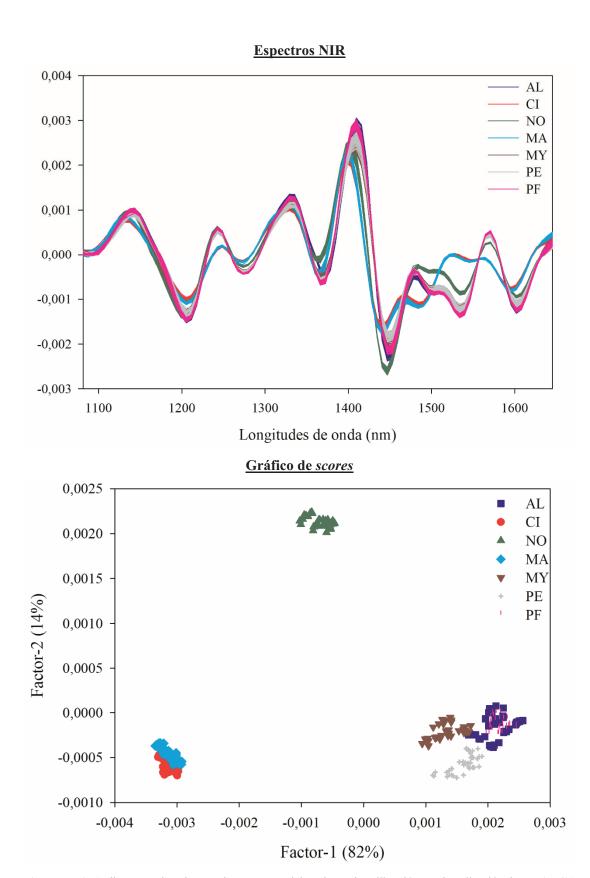


Figura IV. 3. Arriba se pueden observar los espectros del conjunto de calibración tras la aplicación de una 2D (11 puntos de suavizado) y abajo el gráfico de los *scores*, indicando el % de varianza explicada por cada componente principal.

Una vez escogido el pretratamiento, se ha ajustado el número de factores a usar con el fin de optimizar los modelos obtenidos. En el caso de SIMCA, cada marca tiene su propio número óptimo de factores para definir su modelo individual. Las condiciones utilizadas para la construcción de los modelos finales se encuentran en la **Tabla IV. 5** Además, la **Tabla IV. 6** muestra la tabla de contingencia de estos modelos, así como diversos parámetros analíticos evaluados. En esta tabla, TP significa "positivos verdaderos" es decir, muestras que se clasifican como pertenecientes a una clase concreta y realmente lo son. TN son los "negativos verdaderos", muestras que se clasifican como no pertenecientes a la clase y, realmente, no pertenecen a ella. FP son los "falsos positivos", muestras que se clasifican como pertenecientes a una clase concreta cuando en realidad no pertenecen a ella. FN son los "falsos negativos", muestras que no se clasifican como pertenecientes a una clase concreta, cuando en realidad sí que pertenecen a ella. La exactitud, EXA, precisión, PRE, y el índice de Youden se han descrito en la Sección 4.3.5 del capítulo anterior.

Tabla IV. 5. Resumen de los parámetros de los modelos SIMCA, PLS-DA y PCA-LDA utilizando datos NIR.

Parámetro		Resultado
Nº de muestras		247
Intervalo de longitudes	de onda	1050 - 1647 nm
Pretratamiento	1	2D (ventana de
		11 puntos)
Nº de factores		
SI	MCA	
	AL	1
	CI	8
	MA	6
	MY	3
	NO	1
	PE	4
	PF	5
PL	S-DA	8
PC	A-LDA	7

Tabla IV. 6. Parámetros analíticos de la evaluación de los modelos cualitativos de los datos NIR.

Parámetro										
TP T		FP FN		SEN ESP		PRE	EXA	Índice Youden		
36	211	0	0	1,00	1,00	1,00	1,00	1,00		
11	209	2	25	0,31	0,99	0,31	0,89	0,30		
34	186	25	2	0,94	0,88	0,94	0,89	0,83		
36	210	1	0	1,00	1,00	1,00	1,00	1,00		
	36 11 34	36 211 11 209 34 186	36 211 0 11 209 2 34 186 25	36 211 0 0 11 209 2 25 34 186 25 2	TP TN FP FN SEN 36 211 0 0 1,00 11 209 2 25 0,31 34 186 25 2 0,94	TP TN FP FN SEN ESP 36 211 0 0 1,00 1,00 11 209 2 25 0,31 0,99 34 186 25 2 0,94 0,88	TP TN FP FN SEN ESP PRE 36 211 0 0 1,00 1,00 1,00 11 209 2 25 0,31 0,99 0,31 34 186 25 2 0,94 0,88 0,94	TP TN FP FN SEN ESP PRE EXA 36 211 0 0 1,00 1,00 1,00 1,00 11 209 2 25 0,31 0,99 0,31 0,89 34 186 25 2 0,94 0,88 0,94 0,89		

						Paráme	tro			
Ma	arca	TP	TN	FP	FN	SEN	ESP	PRE	EXA	Índice Youden
	NO	36	211	0	1	0,97	1,00	0,97	1,00	0,97
	PE	35	205	6	1	0,97	0,97	0,97	0,97	0,94
	PF	25	216	0	6	0,81	1,00	0,81	0,98	0,81
PLS	S-DA									
	AL	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	CI	35	211	0	1	0,97	1,00	0,97	1,00	0,97
	MA	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	MY	34	211	0	2	0,94	1,00	0,94	0,99	0,94
	NO	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	PE	36	209	2	0	1,00	0,99	1,00	0,99	0,99
	PF	30	216	0	1	0,97	1,00	0,97	1,00	0,97
PCA	-LDA									
	AL	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	CI	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	MA	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	MY	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	NO	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	PE	36	211	0	0	1,00	1,00	1,00	1,00	1,00
	PF	31	216	0	0	1,00	1,00	1,00	1,00	1,00

Teniendo en cuenta los resultados obtenidos, se puede ver que los resultados obtenidos mediante SIMCA son peores que los obtenidos con PLS-DA y PCA-LDA. Los valores de sensibilidad y precisión son menores, lo que supone que el modelo no es capaz de predecir verdaderamente las muestras que pertenecen a un grupo determinado como, por ejemplo, PF. Esto se ve también reflejado en los resultados del Índice de Youden, que indican que el modelo SIMCA tiene un rendimiento bajo.

3.2.2. Validación de los modelos NIR construidos

La validación de los modelos se ha realizado acorde a las recomendaciones de las guías de las agencias reguladoras [8,9]. Así, se ha evaluado la capacidad de clasificación de los modelos prediciendo el resultado de la identificación de diez pastillas de cada marca comercial, las cuales no han sido usadas durante su desarrollo. Los mismos parámetros

analíticos evaluados durante la calibración se han usado también para la validación. La **Tabla IV. 7** muestra los resultados.

Tabla IV. 7. Resultados de la validación de los modelos cualitativos utilizando datos NIR.

					Parámo	etro			
Marca	TP	TN	FP	FN	SEN	ESP	PRE	EXA	Índice Youden
SIMCA									
AL	10	60	0	0	1,00	1,00	1,00	1,00	1,00
CI	8	54	6	2	0,80	0,90	0,80	0,89	0,70
MA	4	59	1	6	0,40	0,98	0,40	0,90	0,38
MY	10	59	1	0	1,00	0,98	1,00	0,99	0,98
NO	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PF	9	60	0	1	0,90	1,00	0,90	0,99	0,90
PLS-DA									
AL	10	60	0	0	1,00	1,00	1,00	1,00	1,00
CI	9	60	0	1	0,90	1,00	0,90	0,99	0,90
MA	10	59	1	0	1,00	0,98	1,00	0,99	0,98
MY	10	60	0	0	1,00	1,00	1,00	1,00	1,00
NO	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	10	59	1	0	1,00	0,98	1,00	0,99	0,98
PF	9	60	0	1	0,90	1,00	0,90	0,99	0,90
PCA-LD	4								
AL	10	60	0	0	1,00	1,00	1,00	1,00	1,00
CI	10	60	0	0	1,00	1,00	1,000	1,00	1,00
MA	10	60	0	0	1,00	1,00	1,00	1,00	1,00
MY	10	60	0	0	1,00	1,00	1,00	1,00	1,00
NO	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	10	59	1	0	1,00	0,98	1,00	0,99	0,98
PF	9	60	0	1	0,90	1,00	0,90	0,99	0,90

Los resultados de la validación muestran la misma tendencia que los obtenidos durante la construcción de los modelos con los datos NIR. El modelo SIMCA presenta valores de sensibilidad, precisión y exactitud menores que los otros dos modelos construidos. En el caso de los modelos PLS-DA y PCA-LDA, los resultados son parecidos, aunque el modelo PCA-LDA presenta resultados algo mejores para los parámetros analíticos estudiados.

Además, con los tres modelos, los parámetros evaluados son un poco peores para la marca PF. Esto es debido a la confusión de uno de los comprimidos, que siempre se considera, erróneamente, que pertenece a la marca PE. En resumen, la validación de los modelos PCA-LDA es satisfactoria.

3.3. Modelos construidos con los datos Raman

En la construcción de los modelos cualitativos con los datos de los espectros obtenidos con el instrumento i-Raman plus, se han probado distintos pretratamientos en el rango espectral de 1800 – 149 cm⁻¹. Dado el efecto de la fluorescencia causado por el color de los comprimidos, en este caso, se ha probado también un pretratamiento con el fin de reducir este efecto (baselinewavelets, BW) y otro para corregir el pendiente de la línea base (BLC). En el caso de la aplicación del BW, es necesario establecer dos parámetros para su aplicación, el umbral y la lambda. Los valores seleccionados han sido lambda 100 y un umbral de 0,3. Por otro lado, para la corrección de la línea base, se ha usado el algoritmo Automatic Whittaker Filter, con lamba igual a 1000 y un valor de p de 0,0001, proporcionado por el software SOLO. La información de estos pretratamientos, así como el significado de sus parámetros puede verse en la Sección 4.2.1 del Capitulo I. La **Figura IV. 4**, muestra los espectros del conjunto de calibración una vez aplicados estos dos pretratamientos. En ambos casos, puede verse que los espectros pertenecientes a una misma marca presentan mucha dispersión en la intensidad Raman.

En este caso, y siguiendo la misma pauta que en los modelos para los datos NIR, durante la construcción del modelo se ha usado el número de factores recomendado por el propio software en el caso de SIMCA y PLS-DA, mientras que para el modelo PCA-LDA se ha establecido el número de componentes principales en 7. La **Tabla IV. 8** y **Tabla IV. 9** muestran los resultados de ESP y SEN obtenidos al aplicar distintas combinaciones de pretratamientos para las muestras de calibración.

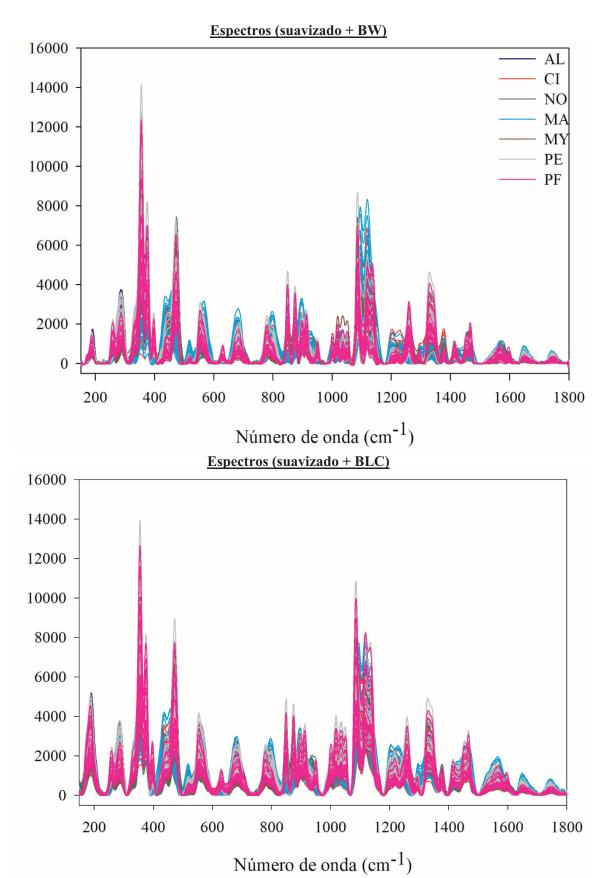


Figura IV. 4. Espectros del conjunto de calibración para los datos Raman al aplicar un suavizado y BW (arriba), o un suavizado y BLC (abajo). La leyenda de colores es la misma para ambas figuras.

Tabla IV. 8. Valores de ESP y SEN obtenidos al aplicar distintos pretratamientos para cada marca con los modelos cualitativos construidos con datos Raman.

		SIN	ICA	PI	S-DA		PCA	-LDA
Pretratamie	nto	ESP	SEN	Nº de factores	ESP	SEN	ESP	SEN
Suavizado + S	SNV			10				
	AL	1,00	0,97		1,00	1,00	1,00	1,00
-	CI	0,97	0,86		1,00	0,92	1,00	0,94
-	MA	0,98	0,78		1,00	0,97	0,99	0,97
-	MY	0,89	0,67		1,00	1,00	1,00	0,94
-	NO	1,00	0,94		1,00	1,00	1,00	1,00
-	PE	1,00	0,89		1,00	1,00	1,00	0,94
-	PF	0,93	0,08		1,00	1,00	0,98	0,96
Suavizado +	1D*			8				
	AL	1,00	0,81		1,00	1,00	1,00	1,00
-	CI	0,83	0,06		1,00	0,83	0,98	0,75
-	MA	0,84	0,06		1,00	0,83	0,96	0,86
-	MY	1,00	0,11		1,00	1,00	1,00	1,00
-	NO	1,00	0,92		1,00	1,00	1,00	1,00
-	PE	0,86	0,44		1,00	0,92	1,00	0,92
-	PF	0,76	0,00		1,00	1,00	0,99	1,00
BW + Suaviz	ado			8				
	AL	0,95	0,06		1,00	0,94	1,00	0,72
-	CI	0,83	0,00		0,98	0,42	0,94	0,69
-	MA	0,83	0,03		0,97	0,36	0,95	0,64
-	MY	0,89	0,14		0,99	0,67	0,96	0,67
-	NO	1,00	0,42		1,00	1,00	1,00	1,00
-	PE	0,72	0,03		0,98	0,69	0,94	0,67
-	PF	0,75	0,00		1,00	0,00	0,94	0,63
BW + Suavizad	o + 1D			10				
	AL	0,98	0,36		1,00	0,97	1,00	1,00
-	CI	0,83	0,00		0,99	0,83	0,97	0,75
-	MA	0,83	0,03		1,00	0,61	0,96	0,81
-	MY	1,00	0,11		1,00	0,89	1,00	0,92
-	NO	1,00	0,81		1,00	1,00	1,00	1,00
-	PE	0,73	0,22		1,00	0,94	1,00	0,97
-	PF	0,76	0,04		1,00	0,63	0,98	1,00

		SIN	ICA	PI	LS-DA		PCA-	-LDA
Pretratamie	ento	ESP	SEN	Nº de factores	ESP	SEN	ESP	SEN
Suavizado +	BLC			9				
	AL	0,94	0,44		1,00	1,00	1,00	1,00
	CI	0,85	0,08		1,00	0,75	0,97	0,75
	MA	0,85	0,11		0,99	0,64	0,96	0,81
	MY	1,00	0,11		1,00	1,00	1,00	0,92
	NO	1,00	0,97		1,00	1,00	1,00	1,00
	PE	0,80	0,17		1,00	0,81	1,00	0,97
	PF	0,79	0,00		0,99	0,92	0,98	1,00
Suavizado + BL	C + 1D			8				
	AL	1,00	0,72		1,00	1,00	1,00	1,00
	CI	0,84	0,06		1,00	0,92	0,98	0,69
	MA	0,84	0,08		1,00	0,81	0,95	0,86
	MY	1,00	0,11		1,00	1,00	1,00	1,00
	NO	1,00	0,92		1,00	1,00	1,00	1,00
	PE	0,84	0,42		1,00	0,89	1,00	0,92
	PF	0,76	0,00		1,00	1,00	0,99	1,00

Tabla IV. 9. Valores globales (considerando todas las marcas) de ESP y SEN al aplicar distintos pretratamientos a los modelos cualitativos construidos con datos Raman.

	SIN	ICA	PLS-DA			PCA-LDA		
Pretratamiento	ESP	SEN	Nº de factores	ESP	SEN	ESP	SEN	
Suavizado + SNV	0,97	0,74	10	1,00	0,98	0,99	0,97	
Suavizado + 1D	0,90	0,34	8	1,00	0,94	0,99	0,93	
BW + Suavizado	0,85	0,10	8	0,99	0,58	0,96	0,72	
BW + Suavizado + 1D	0,88	0,22	10	1,00	0,84	0,99	0,92	
Suavizado + BLC	0,89	0,27	9	1,00	0,87	0,99	0,95	
Suavizado + BLC + 1D	0,90	0,33	8	1,00	0,94	0,99	0,92	

Los espectros sin procesar obtenidos para todas las marcas (**Figura IV. 2**) presentaban mucha fluorescencia. En consecuencia, las bandas no se encuentran bien definidas, sino que se hallan enmascaradas por este efecto. Al corregir la fluorescencia aplicando el algoritmo de BW, no se observan diferencias significativas entre marcas. Por ello, los peores resultados son los obtenidos tras aplicar BW. En el caso de la corrección de la línea base, sucede un

poco lo mismo, aunque los valores de ESP y SEN son algo mejores. La discriminación entre marcas se vuelve más compleja, lo que se ve reflejado en los resultados.

Por el contrario, los mejores resultados de discriminación se consiguen cuando se aplica un suavizado para reducir el ruido espectral seguido de SNV. Considerando estos resultados, se puede concluir que, al eliminar la variabilidad asociada al color de las muestras, ya sea corrigiendo la fluorescencia o la línea base, la discriminación entre marcas empeora en el caso del instrumento i-Raman Plus. La influencia del color es tan grande que la identificación de las marcas se realiza por su propiedad física (el color) y no por su composición química.

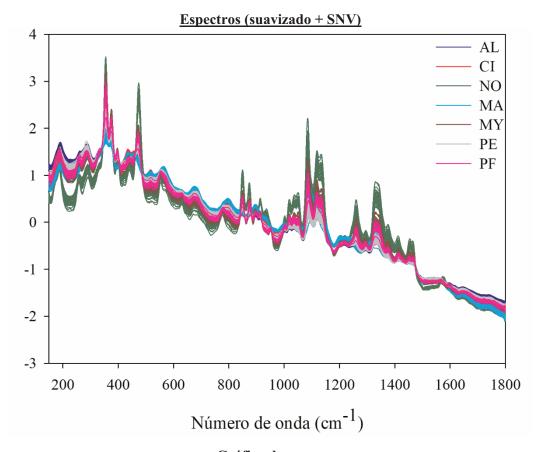
3.3.1. Construcción de los modelos Raman

Para poder visualizar las diferencias entre los espectros de las distintas marcas comerciales, en la **Figura IV. 5**, se muestran los espectros de calibración tras la aplicación de un suavizado (11 puntos) y SNV en el intervalo espectral de 149 – 1800 cm⁻¹, así como la representación de los *scores*. En el gráfico de *scores* presentado, puede verse que la marca comercial que se encuentra más alejada de las demás es NO, que resulta ser la que presenta un tono azulado más intenso en sus comprimidos (**Figura IV. 1**).

Una vez escogido el pretratamiento, se ha ajustado el número de factores a usar con el fin de optimizar los modelos obtenidos. Las condiciones utilizadas para la construcción de los modelos finales se encuentran en la **Tabla IV. 10**. Además, la **Tabla IV. 11** muestra la tabla de contingencia de estos modelos, así como diversos parámetros analíticos evaluados.

Tabla IV. 10. Resumen de los parámetros de los modelos SIMCA, PLS-DA y PCA-LDA utilizando datos Raman.

Parámetro		Resultado
Nº de muestras		
Intervalo de número de or	ıda	$149 - 1800 \text{ cm}^{-1}$
Pretratamiento		Suavizado (11 puntos) + SNV
Nº de factores		
SIMC	A	
	AL	2
_	CI	3
_	MA	3
_	MY	4
_	NO	5
_	PE	5
_	PF	5
PLS-D	Α	7
PCA-LI	DΑ	8



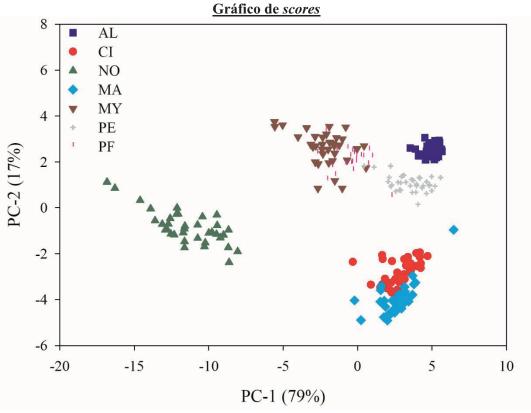


Figura IV. 5. Arriba se pueden observar los espectros del conjunto de calibración tras la aplicación de un suavizado (11 puntos) y SNV. Abajo, el gráfico de los *scores*.

Tabla IV. 11. Parámetros analíticos de la evaluación de los modelos cualitativos para el instrumento i-Raman Plus.

-						Paráme	tro			
Ma	arca	ТР	TN	FP	FN	SEN	ESP	PRE	EXA	Índice Youden
SIN	MCA									
	AL	35	204	0	1	0,97	1,00	0,97	1,00	0,97
	CI	31	202	2	5	0,86	0,99	0,86	0,97	0,85
	MA	32	200	4	4	0,89	0,98	0,89	0,97	0,87
	MY	24	184	20	12	0,67	0,90	0,67	0,87	0,57
	NO	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	PE	34	204	0	2	0,94	1,00	0,94	0,99	0,94
	PF	4	202	14	20	0,17	0,94	0,17	0,86	0,10
PLS	S-DA									
	AL	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	CI	29	200	4	7	0,81	0,98	0,81	0,95	0,79
	MA	33	202	2	3	0,92	0,99	0,92	0,98	0,91
	MY	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	NO	35	204	0	1	0,97	1,00	0,97	1,00	0,97
	PE	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	PF	24	216	0	0	1,00	1,00	1,00	1,00	1,00
PCA	-LDA									
	AL	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	CI	34	203	1	2	0,94	1,00	0,94	0,99	0,94
	MA	35	202	2	1	0,97	0,99	0,97	0,99	0,96
	MY	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	NO	36	204	0	0	1,00	1,00	1,00	1,00	1,00
	PE	34	203	1	2	0,94	1,00	0,94	0,99	0,94
	PF	23	214	2	1	0,96	0,99	0,96	0,99	0,95

Los resultados de los parámetros obtenidos durante la construcción de los modelos con los datos Raman muestran claramente que el modelo SIMCA no tiene una capacidad de discriminación adecuada para estos productos. En el caso de los modelos PLS-DA y PCA-LDA, los parámetros analíticos mejoran hasta el punto de que es posible realizar la identificación de todas las marcas, con la excepción de CI y MA que, tal y como era de prever, no se pueden separar entre sí.

3.3.2. Validación de los modelos Raman construidos

La validación de los modelos se ha realizado siguiendo las recomendaciones de las guías reguladoras, como en el caso de la validación de los modelos con el instrumento MicroNIR. La **Tabla IV. 12** muestra los resultados de los parámetros analíticos evaluados durante la calibración que se han usado también para la validación.

Tabla IV. 12. Resultados de la validación de los modelos cualitativos para los datos Raman.

						Paráme	tro			
Marca	ı	TP	TN	FP	FN	SEN	ESP	PRE	EXA	Índice Youden
SIMC	4									
AI		9	60	0	1	0,90	1,00	0,90	0,99	0,90
CI		10	59	1	0	1,00	0,98	1,00	0,99	0,98
M	A	9	60	0	1	0,90	1,00	0,90	0,99	0,90
M	Y	0	50	10	10	0,00	0,83	0,00	0,71	-0,17
NO)	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	E.	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PF	1	0	50	10	10	0,00	0,83	0,00	0,71	-0,17
PLS-D	A									
AI		10	60	0	0	1,00	1,00	1,00	1,00	1,00
CI		10	60	0	0	1,00	1,00	1,00	1,00	1,00
M	A	10	60	0	0	1,00	1,00	1,00	1,00	1,00
M	Y	10	60	0	0	1,00	1,00	1,00	1,00	1,00
NO)	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	E	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PF	1	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PCA-LI)A									
AI		10	60	0	0	1,00	1,00	1,00	1,00	1,00
CI		10	60	0	0	1,00	1,00	1,00	1,00	1,00
M	A	10	60	0	0	1,00	1,00	1,00	1,00	1,00
M	Y	10	60	0	0	1,00	1,00	1,00	1,00	1,00
NO)	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PE	E .	10	60	0	0	1,00	1,00	1,00	1,00	1,00
PF	1	10	60	0	0	1,00	1,00	1,00	1,00	1,00

Los resultados de la identificación de los diez comprimidos muestran que el mayor número de confusiones entre marcas se encuentran en el modelo SIMCA, donde no es posible la discriminación entre las marcas MY y PF, aunque su composición es ligeramente distinta. Por ejemplo, MY contiene povidona en su formulación mientras PF no. Por otro lado, los resultados de identidad con los modelos PLS-DA y PCA-LDA son satisfactorios para todas las marcas comerciales. En este caso, no se encuentran confusiones entre ninguna marca. Por ello, se considera que tanto el modelo PLS-DA como el PCA-LDA son útiles para su objetivo: discriminación entre marcas comerciales de un mismo producto farmacéutico, mientras que el modelo SIMCA no podría usarse para este propósito.

3.4. Construcción de una biblioteca de espectros

Cada una de las marcas de productos se ha introducido como una clase independiente dentro de la biblioteca. El procedimiento seguido para la construcción de la biblioteca ha sido encontrar el pretratamiento y rango espectral que proporcionara una mayor separación entre las distintas clases de muestras (marcas) y evaluar así su valor de selectividad, para poder determinar si es posible o no su separación. El valor de selectividad determina el grado de discriminación entre marcas. A mayor valor, mayor diferencia entre las clases. Al igual que en la construcción de modelos quimiométricos cualitativos, el conjunto de datos usado para la creación de las bibliotecas ha sido los espectros adquiridos de tres comprimidos en cuatro días distintos y la validación externa de la biblioteca se ha realizado con las lecturas de los diez comprimidos extras. Además, como se disponía de datos suficientes, se ha realizado una validación interna de la biblioteca. Para ello, los espectros de los 3 primeros días se han utilizado como conjunto de calibración y el del cuarto día como conjunto de predicción para realizar la validación interna. En el Capítulo V, se explica en más detallades la construcción de una biblioteca de espectros.

Tras la optimización de todos los parámetros, en la **Figura IV. 6** se muestra la estructura final obtenida para ambos instrumentos y, en las **Tabla IV. 13** y **Tabla IV. 14** se muestran los parámetros usados en cada uno de los niveles para el instrumento MicroNIR y MicroRaman, respectivamente. El nivel principal de la biblioteca es el denominado IP1, donde el número indica el nivel y, en el caso de las sub-bibliotecas, la letra diferencia entre las creadas para un mismo nivel. En este caso, se ha necesitado un nivel más durante la construcción de la biblioteca de espectros Raman, como puede verse en la estructura. Sin embargo, en el caso del NIR, con sólo tres niveles se ha podido discriminar entre ellas.



Figura IV. 6. Estructura de ambas bibliotecas de espectros. A la izquierda para los datos NIR y a la derecha para los datos Raman.

Tabla IV. 13. Parámetros de clasificación seleccionados para el nivel de identificación (nivel 1) y de cualificación para la construcción de la biblioteca NIR.

1	Nivel		Rango espectral (nm)	Pretratamiento	Número de factores
1			1081.5 - 1639.0	2D (9 puntos)	2 (2, 4)
	2.1		1608.1 - 1360.3	2D (21 puntos)	1 (1)
		3	1465.6 - 1385.1 1316.9 - 1236.4	2D (25 puntos) + VN	1 (2)
	2.2		1632.8 - 951.5	VN	1 (2)

Tabla IV. 14. Parámetros de clasificación seleccionados para el nivel de identificación (nivel 1) y de cualificación para la construcción de la biblioteca Raman.

	Nivel		Rango espectral (cm ⁻¹)	Pretratamiento	Método de discriminación
1			1703 - 419	VN	2 (2, 3)
	2.1		1525 - 741	VN	2 (2, 3)
	3		339 - 143	VN	1 (2)
		4	1600 - 488	VN	1 (2)
	2.2		1818 - 246	VN	1 (2)

Con estas condiciones, se consigue la separación entre todas las marcas, con la excepción de CI y MA. Para estas marcas, se ha decidido crear una categoría conjunta llamada Laboratorio Cinfa, S.A, dado que su composición es la misma. De esta manera, su identificación mediante las bibliotecas construidas será conjunta.

El histograma de selectividad obtenido se presenta en la **Figura IV. 7.** para ambos instrumentos. En OPUS 7.8, todos los resultados con selectividad menor a 1 son representados en rojo. A pesar de esta asignación del color, no hay confusión entre ninguna de las marcas incluidas en la biblioteca de espectros [10].

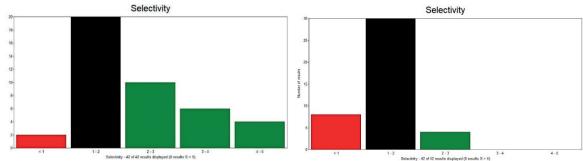


Figura IV. 7. Valores de selectividad obtenidos una vez validada la biblioteca de espectros NIR (izquierda) y Raman (derecha).

Las bibliotecas de espectros construidas se han validado realizando una prueba de identificación a los espectros de los diez comprimidos extra. Los resultados de la validación externa se muestran en la **Tabla IV. 15**.

Tabla IV. 15. Resultados de la identificación de las muestras de validación (n = 10) clasificando los resultados como muestras identificadas correctamente (OK), no identificadas o con un resultado negativo (No OK) para ambas bibliotecas de espectros.

	Re	sultados Micr	oNIR	Resultados i-Raman			
Marca	Resultado OK	Resultado No OK	No identificada	Resultado OK	Resultado No OK	No identificada	
AL	10	-	-	10	-	-	
Lab Cinfa ¹	20	-	-	20	-	-	
MY	8	-	2	10	-	-	
NO	10	-	-	10	-	-	
PE	10	-	-	10	-	-	
PF	9	-	1	10	-	-	

¹Incluye CI y MA.

Como se puede observar, ningún comprimido de una marca ha sido identificado como perteneciente a otra. Sin embargo, en el caso de los resultados obtenidos con los datos NIR, habría que optimizar el conjunto de calibración de las marcas MY y PF e introducir una mayor variabilidad para poder obtener una identificación satisfactoria de todos los comprimidos, ya que hay muestras que no son identificadas como pertenecientes a ninguna de ambas marcas ni a ninguna de las otras presentes en la biblioteca. La variabilidad introducida por estos comprimidos no identificados debería introducirse en los conjuntos de calibración para poder obtener una biblioteca de identificación más robusta.

El mismo comprimido de PF que no se ha identificado correctamente en la biblioteca es el mismo que fallaba cuando se utilizaba SIMCA o los modelos de discriminación lineal. Por otro lado, en el caso de la biblioteca de espectros Raman, es posible la discriminación y correcta identificación de todas las marcas y comprimidos, mejorando así los resultados

obtenidos con los modelos quimiométricos que realizan la identificación en una única etapa. Esta mejora en los resultados Raman puede estar relacionada con la aplicación del pretratamiento VN, que es ligeramente distinto al SNV usado en los modelos cualitativos.

4. CONCLUSIONES

En este trabajo, se ha comprobado que los espectrómetros portátiles, tanto NIR como Raman, son capaces de ofrecer resultados satisfactorios en la clasificación y discriminación de siete marcas comerciales de un mismo producto farmacéutico.

La construcción y validación de los modelos ha confirmado que los peores resultados de predicción se obtienen al utilizar la técnica SIMCA, ya que se confunden entre sí las distintas marcas. Por el contrario, los modelos PLS-DA y PCA-LDA han demostrado ser capaces de identificar y discriminar las marcas, a pesar de alguna confusión puntual de algún comprimido concreto. En el caso de la construcción de las bibliotecas, éstas han sido capaces de discriminar correctamente entre todas las marcas. Debido que se han fabricado igual (misma composición y proceso), aunque con distinta marca, los comprimidos de CI y MA se han agrupado en una única categoría que permite su identificación conjunta.

Al comparar ambas técnicas, en el caso de la espectroscopia NIR ha sido posible discriminar las marcas por su composición química, mientras que al utilizar espectros Raman, los cuales presentan bandas más bien definidas que los espectros NIR, la separación entre los comprimidos se ha realizado básicamente por la intensidad de su color, ya que la fluorescencia generada provoca cambios en el espectro mayores que los cambios debidos a la distinta composición química de cada marca.

Finalmente, se puede concluir que la identificación de productos utilizando bibliotecas de espectros en cascada, utilizada habitualmente en entornos industriales con muchas clases, resulta también muy útil en casos con pocas clases, aunque el software disponible tiene sistemas de pretratamiento de espectros bastante limitados.

5. BIBLIOGRAFÍA

- [1] K.B. Beć, J. Grabska, C.W. Huck, *Principles and Applications of Miniaturized Near-Infrared (NIR) Spectrometers*, Chem. A Eur. J. 27 (**2021**) 1514–1532.
- [2] R.A. Crocombe, *Portable Spectroscopy*, Appl. Spectrosc. Vol. 72, Issue 12, Pp. 1701-1751. 72 (**2018**) 1701–1751.

- https://www.osapublishing.org/as/abstract.cfm?uri=as-72-12-1701 (accessed October 5, 2021).
- [3] R. Deidda, P.Y. Sacre, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers, TrAC Trends Anal. Chem. 114 (2019) 251–259.
- [4] M. Alcalà, M. Blanco, D. Moyano, N. Broad, N. O'Brien, D. Friedrich, F. Pfeifer, H. Siesler, *Qualitative and quantitative pharmaceutical analysis with a novel handheld miniature near-infrared spectrometer*, J. Near Infrared Spectrosc. 21 (2013) 445. https://doi.org/10.1255/jnirs.1084.
- [5] H. Yan, H.W. Siesler, Quantitative analysis of a pharmaceutical formulation: Performance comparison of different handheld near-infrared spectrometers, J. Pharm. Biomed. Anal. 160 (2018) 179–186. https://doi.org/10.1016/J.JPBA.2018.07.048.
- [6] C.G. Pederson, D.M. Friedrich, C. Hsiung, M. von Gunten, N.A. O'Brien, H.-J. Ramaker, E. van Sprang, M. Dreischor, Pocket-size near-infrared spectrometer for narcotic materials identification, in: Proc.SPIE, **2014**.
- [7] K. Dégardin, A. Guillemain, Y. Roggo, Comprehensive Study of a Handheld Raman Spectrometer for the Analysis of Counterfeits of Solid-Dosage Form Medicines, J. Spectrosc. 2017 (2017) 3154035. https://doi.org/10.1155/2017/3154035.
- [8] ICH, ICH Q2 (R2) Validation of Analytical Procedures, (2022).
- [9] European Pharmacopoeia 10.0., Chapter 5.21. Chemometrics Methods Applied to Analytical Data, (2021) 819–836.
- [10] Bruker, Use Manual IDENT OPUS Spectroscopy Software, in: Version 7., 2011.

V. <u>CONSTRUCCIÓN DE BIBLIOTECAS NIR PARA</u> <u>LA IDENTIFICACIÓN DE MATERIALES DE</u> <u>PARTIDA</u>

1. INTRODUCCIÓN

Uno de los aspectos más importantes durante la fabricación de un producto farmacéutico, es la correcta identificación de los materiales de partida (MP). A parte del error catastrófico que supondría un error en la identificación del material, se ha de considerar que variaciones en propiedades químicas o físicas de los MP pueden provocar un cambio en la calidad final del producto, ya que la calidad de un producto farmacéutico depende, en gran parte, de la calidad de los MP. Por ello, el Anexo 8 complementa el Capítulo 6 de la de la *Guía de Normas de Correcta Fabricación de Medicamentos de Uso Humano y Veterinario*, y establece que "la identidad de un lote completo de materiales de partida sólo puede garantizarse si se toman muestras individuales de todos los envases y se lleva a cabo un ensayo de identidad en cada muestra. Se puede permitir tomar muestras de sólo una parte de los envases, cuando se haya establecido un procedimiento validado para garantizar que ningún envase individual de material de partida se haya etiquetado incorrectamente" [1].

Cada MP debe cumplir con las especificaciones marcadas en su propia monografía de la farmacopea europea (EP por sus siglas en inglés *European Pharmacopoeia*) o, en caso de no disponer de una monografía propia, debe cumplir con la monografía general *Substances for pharmaceutical use* [2]. Por este motivo, queda claro que cada recipiente recibido en la fábrica (bidón, bolsa, etc.) debe ser tratado según la monografía EP, que utiliza métodos tradicionales de análisis tales como valoraciones o HPLC y que necesitan de bastante tiempo de ejecución. Sin embargo, de acuerdo con las *General notices* de la propia EP, es posible usar métodos alternativos de análisis a los propuestos si lo aprueba la autoridad competente [3].

El uso de la espectroscopia NIR, combinado con algoritmos quimiométricos, presenta una serie de ventajas para la identificación de MP en comparación a los métodos tradicionales. Algunas de estas ventajas son: la reducción en el tiempo de análisis, la eliminación del uso de disolventes, el ser un método no destructivo, la obtención de la información at-line y la obtención de más información que usando métodos tradicionales, como, por ejemplo, información sobre propiedades físicas de la muestra.

Actualmente, la EP contiene una monografía propia para la aplicación de la técnica NIR en la industria farmacéutica [4]. En esta monografía, se contempla la posibilidad de

crear una biblioteca de espectros para la identificación y cualificación de productos y se establecen directrices sobre cómo se debe proceder para su construcción.

La identificación de diversos MPs mediante la construcción de una biblioteca de espectros NIR no está muy referenciado. Normalmente, se construyen métodos específicos de identificación para cada producto o, solo se discrimina entre unos pocos mediante el uso de métodos de reconocimiento de pautas (PRM por sus siglas en inglés Pattern Recognition Methods) [5–8]. Aún y así, hay artículos que prueban y proporcionan recomendaciones para la construcción de una biblioteca de espectros NIR para este fin [7,9]. Para la construcción de una biblioteca espectral, el primer paso es elaborar una biblioteca principal, donde cada producto presente en la misma estará definido por un conjunto de espectros NIR, formando el que se denomina una "clase", con el objetivo de permitir la identificación del mayor número de productos posibles. En este primer paso, normalmente se usa como criterio de similitud el coeficiente de correlación o el cálculo de distancias entre los espectros. Para que un producto pueda identificarse como perteneciente a una clase, se establece que ha de cumplir un criterio de aceptación. Aquellos MPs que no se hayan podido identificar en la biblioteca principal, es decir, que presentan ambigüedad entre sí, han de hacerlo en subbibliotecas, construidas usando un método cualitativo con mayor selectividad y que permita su identificación. Este procedimiento se denomina identificación en cascada. En la Figura V. 1 se muestra un ejemplo de la estructura que puede tener este tipo de biblioteca de espectros.

Como puede verse en la representación, a partir de la biblioteca principal se van estableciendo distintos niveles donde cada nivel contendrá un número distinto de subbibliotecas, dependiendo de la selectividad que haya entre los distintos materiales.

Cuando un producto no es identificado en la biblioteca, se clasifica como "fuera de especificación" (OOS por sus siglas en inglés *Out-of-specification*). En la inmensa mayoría de los casos, estos resultados OOS son productos que, una vez analizados por otras técnicas mucho más lentas y costosas, finalmente se identifican como el producto etiquetado. Realmente son falsos negativos, pero hasta obtener el resultado no se puede autorizar la entrada del lote de producto. La consecuencia de la obtención de este tipo de OOS es el aumento del tiempo necesario antes de liberar el producto, lo que se refleja en una pérdida económica para la empresa.

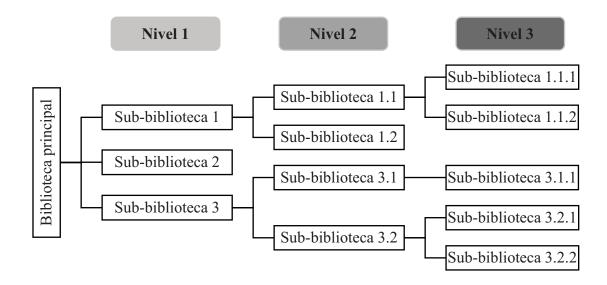


Figura V. 1. Esquema general de la construcción de una biblioteca de espectros mediante la estrategia de identificación en cascada.

En este trabajo, se ha realizado una revisión en profundidad y la revalidación de una biblioteca para la "Identidad por Recipiente" utilizando NIR. La primera versión de la biblioteca NIR se desarrolló en la industria el año 2018 por técnicos del departamento de control de calidad. Dada la gran cantidad de OOSs obtenidos, se revalidó la biblioteca en 2019, ensayando distintos pretratamientos para intentar mejorar la discriminación entre todos los MPs incluidos en la misma. El número de OOSs disminuyó, pero no lo suficiente, por lo que en 2020 se abordó el trabajo que se presenta en este capítulo. Para la nueva versión de la biblioteca, se decidió crear una Biblioteca Específica para los PMs utilizados como colorantes. El resto de los productos quedaron en lo que se denominará Biblioteca General. Para ambas bibliotecas, se ha optimizado los conjuntos de aprendizaje disponibles de cada producto y se ha mejorado los parámetros de discriminación para reducir el número de resultados OOS.

2. MATERIALES Y MÉTODOS

2.1. Muestras

La Biblioteca General contiene 398 clases de MPs y la Biblioteca Específica 34. Para cada clase se ha utilizado una cantidad variable de espectros NIR, ya que su número depende

del volumen de lotes recibidos y éste, de la demanda que haya del producto. En cualquier caso, se empezaron a registrar espectros NIR desde 2012.

En la Biblioteca General, se incluyen tanto APIs como excipientes, que pueden ser materiales orgánicos o inorgánicos, y engloba todos los MPs utilizados en la empresa cuya identidad puede ser hallada utilizando la técnica NIR. Hay ciertos compuestos, como pellets o materiales cerosos, cuyas características físicas dificultan la adquisición de los espectros y el NIR no es adecuado para su identificación. Los colorantes identificados en la Biblioteca Específica son de tres tipos (I, II y III), dependiendo del polímero sintético que contengan en su composición. Además, dentro de cada tipo de colorante puede haber diferentes productos, con distinto color y composición, aunque compartan el mismo polímero sintético como componente mayoritario.

Debido al acuerdo de confidencialidad firmado con la empresa, no se puede especificar el nombre de los MPs incluidos en ambas bibliotecas de espectros.

2.2. Adquisición de espectros

La adquisición de todos los espectros se ha realizado con el NIR de sobremesa Bruker, utilizando una sonda de fibra óptica y dentro de una cabina de muestreo. Los detalles del instrumento se encuentran en el Capítulo III (Sección 1.2.3).

Se debe hallar la "Identidad por Recipiente" mediante NIR de todos los bultos recibidos de cada lote, con excepción de aquellos productos que se encuentren homologados, y, por tanto, se puede realizar un muestreo reducido calculado mediante la **Ecuación V-1**, donde *m* representa el número de muestras a identificar y *N* el número de bultos recibidos.

$$m = \sqrt{N} + 1 \tag{V-1}$$

2.3. Software

El programa usado para la construcción y validación de las bibliotecas ha sido OPUS 7.2 (Bruker, Billerica, MA, USA). Este software permite la construcción y validación de bibliotecas de espectros mediante uno de sus paquetes incorporados (Ident). Los algoritmos y pretratamientos disponibles en este software se detallarán entre paréntesis, indicando el nombre especifico que usa el propio programa.

2.4. Desarrollo de la biblioteca de espectros

La creación de una biblioteca de espectros pasa por disponer de datos muy fiables de todos los productos que se incluirán en la misma. Para cada producto, estos datos fiables se dividen en dos bloques; el primero se denomina conjunto de aprendizaje y serán los datos usados para la construcción de la biblioteca, utilizando tantas clases como productos presentes en la misma. El segundo bloque servirá para comprobar el funcionamiento de ésta y se llama conjunto de validación.

Con el conjunto de aprendizaje, que contiene espectros de todos los productos de la biblioteca, se construye un modelo utilizando un método PRM, que comienza en el nivel principal y sigue analizando en cascada todos los productos, escogiéndose en cada nivel el método PRM y las condiciones apropiadas. Una vez diseñada la biblioteca, se realiza una validación interna para asegurar que se distinguen como productos distintos, sin ambigüedades, todos los productos incluidos en la misma. Finalmente, se realiza una validación externa con los espectros que no se habían incluido en el conjunto de aprendizaje, lo que constituye el conjunto de validación. El criterio de identificación de un espectro como perteneciente a una clase concreta se basa en la creación de una distancia límite desde el centro de la clase para cada producto incluido en la biblioteca. Si el espectro del producto desconocido está dentro del límite, se determinará que pertenece a esa clase. Una vez realizada la validación externa con resultados satisfactorios, se puede empezar a usar la biblioteca en rutina para analizar muestras desconocidas. La **Figura V. 2** muestra un esquema de los pasos a seguir para desarrollar y validar una biblioteca de espectros.

2.4.1. Optimización de los conjuntos de aprendizaje y validación

El conjunto de espectros de cada MP debe contener la mayor variabilidad posible para poder lograr el propósito de la biblioteca, identificar correctamente muestras desconocidas. La "Identidad por Recipiente" mediante NIR se trata de un análisis cualitativo, por lo que las fuentes de variabilidad asociadas a diferencias lote-lote, replicado-replicado y adquisición, deben estar representadas en el conjunto de aprendizaje para asegurar la construcción de un modelo robusto, aunque no existe una regla determinada para la selección del mínimo número de lotes que se deben incluir en el conjunto de aprendizaje. Además, también es necesario incluir la variabilidad aceptable relacionada con las propiedades físicas

y químicas de un mismo producto, como podría ser tamaño de partícula, contenido en agua, etc.

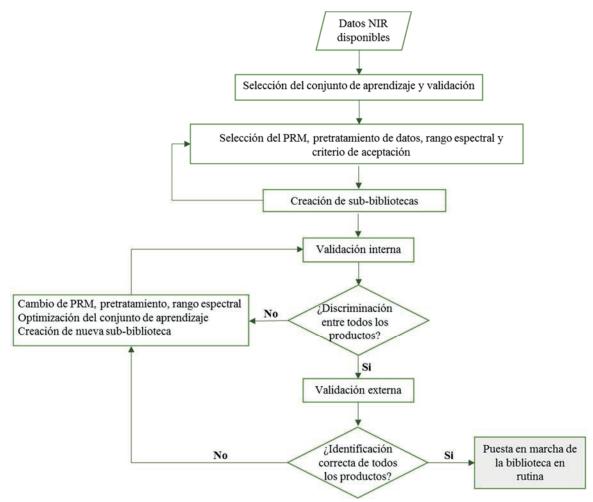


Figura V. 2. Diagrama de flujo para la creación y validación de una biblioteca de espectros.

Para elegir los espectros de cada MP que se utilizarán en el conjunto de aprendizaje de la biblioteca, se han representado gráficamente todos los espectros incluidos en la versión anterior de la misma, junto con nuevos espectros adquiridos durante el periodo de julio 2019 a marzo 2021. Estos espectros se han normalizado aplicando el método de normalización vectorial (*vector normalization*) para poder detectar posibles espectros anómalos (*outliers*). Una vez representados todos los espectros de un MP, se han seleccionado los localizados en las posiciones extremas y en el centro. También se ha intentado, en los casos en que era posible, incluir variabilidad de lotes adquiridos en años distintos. El resto de los espectros del MP se ha asignado al conjunto de validación.

2.4.2. Construcción de la biblioteca

El método seleccionado para asegurar la discriminación entre todos los MPs ha sido el cálculo de distancias en el espacio de las componentes principales (*Factorization*). Este método para el cálculo de similitudes es especialmente útil para la construcción de una biblioteca de espectros NIR, dado que es independiente del número de grupos, requiere pocos espectros para definir cada clase y, es muy poco sensible a las fluctuaciones del instrumento, lo que permite el desarrollo y validación de bibliotecas de espectros de una forma rápida.

El principal objetivo durante la construcción de una biblioteca de espectros es poder lograr una identificación única de todos los MPs incluidos en la misma. Para ello se ha usado la estrategia de identificación en cascada mencionada anteriormente. Así, después de la aplicación inicial de un conjunto de parámetros en el nivel principal de la biblioteca, que consigue separar algunos MPs individuales y la creación un grupo de MPs no separados, se ha aplicado otros parámetros de discriminación más restrictivos a este grupo, en lo que sería un nuevo nivel de identificación o sub-biblioteca. Esto se ha ido repitiendo hasta conseguir el objetivo final.

Sin embargo, hay MPs en la biblioteca que tienen la misma composición, pero son fabricados por distintos proveedores, o bien que tienen la misma composición, pero con ligeras diferencias en sus propiedades físicas. En estos casos, la diferencia entre ellos puede ser tan pequeña que no sea posible diferenciarlos, por lo que se crea una categoría que proporciona una identificación conjunta de este tipo de MPs. Cuando esto ocurre, el criterio utilizado para la creación de estas categorías debe ser siempre justificado.

2.4.3. Validación interna

Antes de poder utilizar la biblioteca, es necesario validarla para comprobar que las distintas clases que la componen están separadas entre sí. Este proceso se denomina validación interna y se basa en la comprobación de la selectividad (SEL) de la biblioteca.

Para cada una de las clases se define una distancia umbral, que marca la frontera entre la pertenencia al grupo o no. El umbral se calcula mediante la **Ecuación V-2**.

$$Umbral = Maximum \ hit + X \times SD \tag{V-2}$$

En esta Eq. 2, el valor $Maximum\ hit$ es la distancia entre el espectro promedio de la clase y el espectro más distante perteneciente a ella; el valor X es un factor multiplicativo que, por defecto, tiene un valor de $0,25\ y\ SD$ es la desviación estándar de todos los espectros que forman la clase. La selectividad comprobará que la distancia entre los conjuntos de espectros que definen cada MP es suficiente para el uso correcto de la biblioteca. En la **Ecuación V-3**, se relaciona la distancia (D) entre los espectros promedio de dos MPs, con sus respectivos valores umbral $U_1\ y\ U_2$

$$SEL = \frac{D}{U_1 + U_2}$$
 V-3

Dependiendo de los valores de SEL obtenidos, se puede considerar cuatro situaciones que se esquematizan en la **Figura V. 3**. En cualquier caso, cuanto menor es el valor de SEL, mayor será la similitud entre MPs. La situación ideal para evitar confusiones entre MPs es que el valor de selectividad sea mayor que 1 para todos los productos.

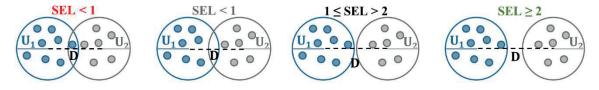


Figura V. 3. Representación de las cuatro posibles situaciones del software OPUS 7.2 dependiendo de los resultados de selectividad (SEL). D es la distancia entre los centros de ambos grupos y U_i el umbral de cada uno de ellos.

2.4.4. Criterio de aceptación

Una vez que se ha validado la biblioteca, puede ser utilizada en la identificación de espectros que no están presentes en ella. El resultado de la identificación del espectro de una muestra es positivo cuando el MP evaluado se identifica inequívocamente con uno de los grupos incluidos en la biblioteca. La identificación se realiza mediante el cálculo de la distancia entre las muestras a identificar y el grupo más cercano a ellas.

Así, para obtener una identificación positiva respecto a alguno de los grupos incluidos, la distancia calculada (*hit value*) para el espectro evaluado debe ser menor a alguno de los umbrales establecidos. Si no se cumple este criterio, no se obtendrá una identificación positiva con ningún producto incluido en la biblioteca. En el caso de que se obtuviera una distancia inferior a más de uno de los umbrales, no habrá una identificación única y, por tanto, la identificación no será concluyente.

2.4.5. Validación externa

Antes de utilizar la biblioteca para la identificación de muestras desconocidas, se realiza un último paso, evaluando con el modelo construido los lotes no utilizados durante el desarrollo de éste y de los que se conoce su identidad. Este proceso se denomina validación externa y servirá para tener confianza en el desempeño de la biblioteca.

3. RESULTADOS Y DISCUSIÓN

3.1. Biblioteca Específica

Antes de la creación de la biblioteca de los colorantes, en la segunda versión de la biblioteca NIR, hubo un total de 8 OOS de 9 colorantes distintos, para un total de 77 bultos analizados. Este resultado supone aproximadamente una tasa de fallo del 10%. La biblioteca de colorantes o Biblioteca Específica consta de un total de 34 colorantes, clasificados en tres tipos (I, II, III) dependiendo del polímero que contengan. En la **Figura V. 4** se presenta un esquema clasificando los 34 colorantes según su tipo (I, II, III) y su color.



Figura V. 4. Clasificación de los 34 colorantes usados para la construcción de la Biblioteca Específica. Entre paréntesis se muestra el número de colorantes pertenecientes a cada grupo de color y tipo.

3.1.1. Construcción de la Biblioteca Específica

El PRM seleccionado ha sido el del cálculo de la distancia en el espacio de las componentes principales (*Factorization*), tanto para el nivel principal de la biblioteca, llamado IP1, como para todas las sub-bibliotecas creadas. La selección del algoritmo de discriminación, así como los pretratamientos y rango espectral, se ha basado en la estrategia de encontrar la combinación que proporcione una mejor discriminación en el primer nivel y que cree sub-bibliotecas lógicas según las propiedades físicas y químicas de los productos. Se deseaba una separación entre los tres tipos de colorantes en el nivel general de la biblioteca (IP1), dado que el polímero que contienen es distinto. Por ello, tras la aplicación

de una primera derivada (1D) con 9 puntos de suavizado y una normalización vectorial (VN), y seleccionando un total de 4 componentes principales, se logró separar entre el tipo I, II y III, con la excepción de los colorantes naranja y rosa del tipo III, que se confundían con colorantes del tipo II en la sub-biblioteca IP2_B, aunque en este nivel se podían separar en una nueva sub-biblioteca. La **Figura V. 5**, muestra los gráficos de *scores* obtenidos para este primer nivel, donde puede verse la separación en tres *clusters* formando, respectivamente, las tres sub-bibliotecas (IP2_A, IP2_B y IP2_C). El número que aparece en el nombre de cada sub-biblioteca indica el nivel de ésta. Los *clusters* se han identificado en la Figura. En el nivel general (IP1) ha sido posible separar tres colorantes individuales al representar los *scores* 4 vs 2 y 4 vs 3.

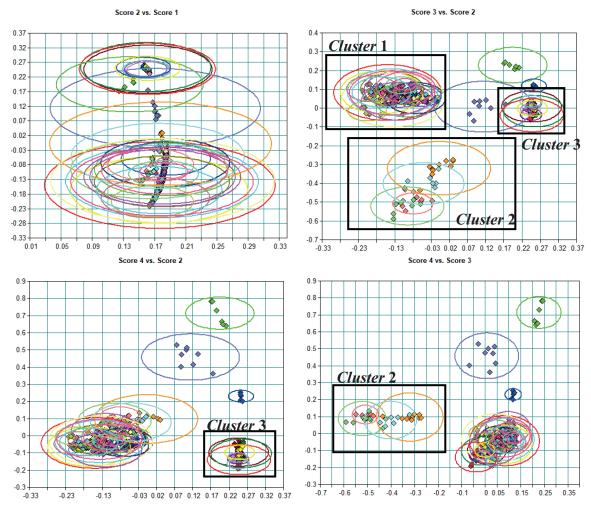


Figura V. 5. Gráfico de scores del nivel principal (IP1) de la Biblioteca Específica.

En la **Figura V. 6** se muestran los gráficos de scores de la sub-biblioteca IP2_B, mostrando la separación de los dos colorantes rosa y naranja de tipo III del resto de colorantes de tipo II.

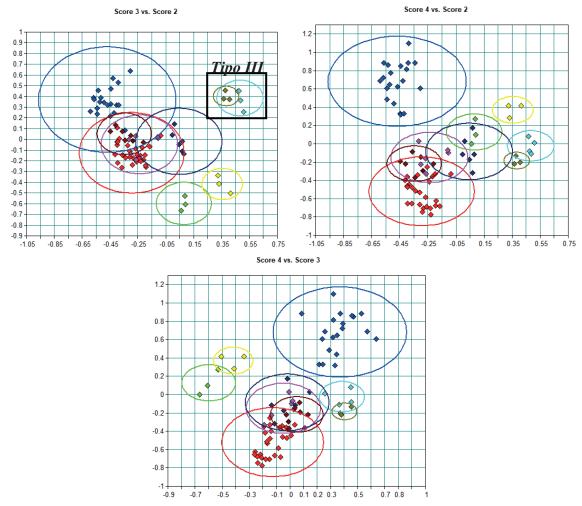


Figura V. 6. Gráfico de scores de la sub-biblioteca (IP2_B) de la Biblioteca Específica.

A partir de aquí, se han ido desarrollando las tres sub-bibliotecas creadas hasta la obtención de la estructura final de la Biblioteca Específica, que se presenta en la **Figura V.** 7, donde puede verse que se ha necesitado un total de 8 niveles y 20 sub-bibliotecas para poder discriminar entre todos los colorantes.

3.1.2. Problemas planteados durante la construcción de la biblioteca

El objetivo principal de esta Biblioteca Específica es la identificación inequívoca de los 34 colorantes usados en la industria. Esto se pudo lograr para 31 de los 34 productos incluidos en la biblioteca. Sin embargo, no fue posible lograr la discriminación de tres productos blancos pertenecientes al tipo I de colorante, debido a que en su composición

contienen el mismo polímero en alta concentración y comparten los mismos compuestos, aunque con distintas concentraciones, por lo que no hay diferencias entre sus espectros NIR. En este caso, se decidió usar la estrategia de la creación de una categoría para productos altamente similares, basándonos en su composición química, y que se ha denominado Colorante Blanco, conteniendo los tres productos.

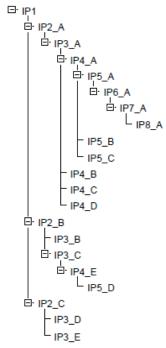


Figura V. 7. Estructura final de la Biblioteca Específica de colorantes.

En la ¡Error! No se encuentra el origen de la referencia. se detallan las diferencias en la composición de todos los compuestos de la formulación de la categoría. Como puede apreciarse, la diferencia en la composición de estos tres colorantes está en alrededor de un 3% en peso. La Figura V. 8 se presentan los espectros del conjunto de aprendizaje para estos tres colorantes, para mostrar que no es posible su discriminación individual.

Tabla V. 1. Composición (% m/m) de los tres colorantes blancos (Tipo I) que conforman la categoría Colorante Blanco creada.

Compuesto	Colorante Blanco 1	Colorante Blanco 2	Colorante Blanco 3
A	63,6	62,5	59,8
В	30	31,3	31,3
С	6,4	6,3	8,0
D	-	-	1,0

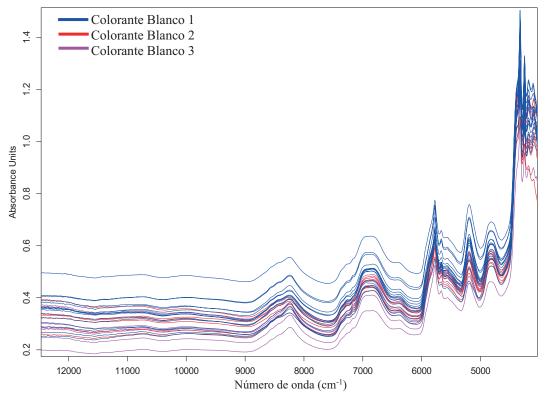


Figura V. 8. Espectros en absorbancia de los colorantes que constituyen la categoría Colorante Blanco en la Biblioteca Específica.

3.1.3. Validación interna

La validación interna de la biblioteca se ha realizado mediante el cálculo de la selectividad (SEL). En la **Figura V. 9**, puede verse el histograma de los valores de SEL de la comparación de 1122 pares de clases. Únicamente 6 colorantes presentan valores de SEL menores que 1, pero sin solapamiento, lo que quiere decir que, aunque ambos umbrales estén solapados, no se encuentra ningún espectro en el rango de intersección entre ambos. En estos 6 colorantes, se encuentran los tres colorantes blancos incluidos en la nueva categoría creada, Colorante Blanco, y, además otros 3 colorantes que presentan SEL bajas, por lo que, se deberá hacer un seguimiento exhaustivo del comportamiento de éstos cuando haya nuevas entradas para ver si se puede mejorar esta discriminación. Los tres colorantes que presentan SEL bajas son del tipo I. Esta SEL baja aparece en el nivel IP4_A. La **Figura V. 10** muestra los espectros tras la aplicación de una normalización vectorial en el rango espectral seleccionado, donde se ve que existen diferencias entre los espectros, y la **Figura V. 11** muestra el gráfico de los *scores*, donde se observa que los productos aparecen muy cercanos uno al otro.

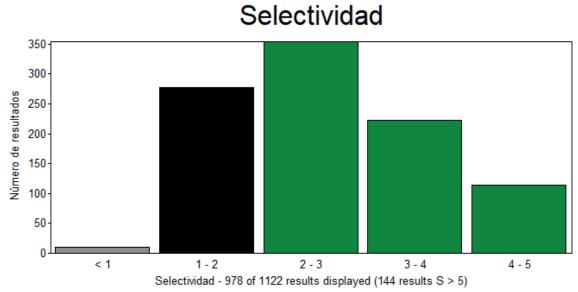


Figura V. 9. Histograma de selectividad obtenido tras la construcción de la Biblioteca Específica de colorantes.

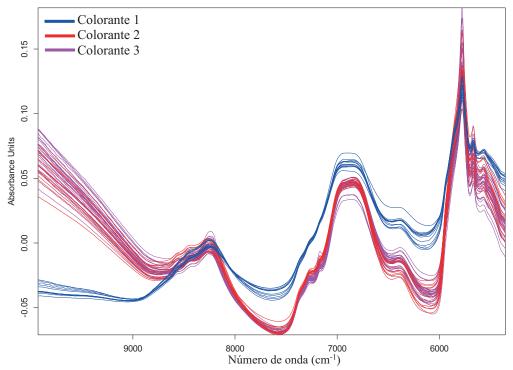


Figura V. 10. Espectros tras la aplicación de una normalización vectorial para los colorantes con valores de SEL baja en la Biblioteca Específica.

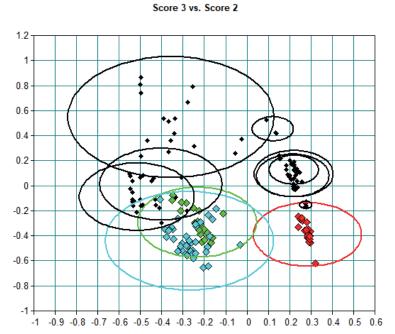


Figura V. 11. Gráfico de *scores* del nivel IP4_A de la Biblioteca Específica. Se resaltan en color los colorantes que tienen valores de SEL baja entre ellos.

El histograma de SEL muestra que no hay otras ambigüedades entre los productos incluidos en la biblioteca.

3.1.4. Validación interna

Una vez desarrollada la Biblioteca Específica NIR, se realizó la validación externa de la misma para comprobar su selectividad y robustez.

La validación externa de la Biblioteca Específica solo fue posible para 13 de los 34 colorantes incluidos en la misma. Esto es debido a los pocos lotes de estos colorantes que se reciben cada año en la industria, dado que se necesita muy poca cantidad para realizar el recubrimiento de los productos farmacéuticos. Por ejemplo, durante el desarrollo de esta nueva Biblioteca Específica, para 9 de los colorantes solo ha sido posible incluir en los conjuntos de aprendizaje de la biblioteca 3 espectros de un mismo lote y bulto, dado que se trataban de productos nuevos.

Basándonos en los resultados obtenidos, se puede concluir que la Biblioteca Específica de colorantes ha sido validada provisionalmente y puede usarse para identificar inequívocamente todos los colorantes usados en la industria. Sin embargo, la variabilidad incluida para alguno de los grupos debería incrementarse en futuras actualizaciones de la biblioteca para reducir así la posibilidad de obtener resultados OOS.

3.1.5. Evaluación de la actualización

La evaluación del impacto de la creación de una Biblioteca Específica para los colorantes se realiza mediante el análisis de los resultados de identidad por recipiente NIR obtenidos en el periodo de uso de la biblioteca. La **Figura V. 12** presenta los resultados considerando solo los colorantes.

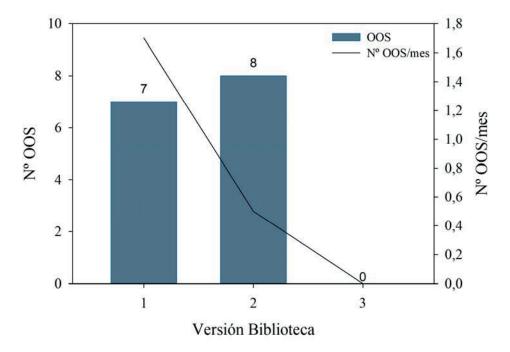


Figura V. 12. Evaluación de la creación de una Biblioteca Específica para la identidad por recipiente mediante NIR de los colorantes. En las versiones 1 y 2, los colorantes estaban incluidos en la misma biblioteca, junto con los demás productos.

Como puede verse, en la primera y segunda versión, donde los colorantes estaban incluidos con el resto de MP, el número de resultados no conformes (falsos negativos) era alto. Durante los cuatro meses en que ha estado funcionando la nueva Biblioteca Específica, presentada en este capítulo, se han analizado 13 colorantes y no se ha producido ningún resultado OOS. Por lo que, la separación de todos los colorantes en una única biblioteca facilita su identificación. A pesar de todo, muchas de las clases de los colorantes incluidos en la biblioteca tienen muy poca variabilidad y, cuando lleguen nuevos bultos, deberán añadirse sus espectros al conjunto de aprendizaje.

3.2. Biblioteca General

La Biblioteca General consta de un total de 497 MPs, muchos de distintos proveedores, y que se clasifican en APIs, excipientes y mezclas, con lo que se construye una biblioteca con 398 clases, incluyendo en la misma clase los productos con más de un proveedor. Como se ha mencionado anteriormente, se ha actualizado esta biblioteca realizando una optimización de los conjuntos de aprendizaje, de validación y de los parámetros de discriminación, con el fin de simplificar su estructura y reducir el número de resultados no conformes. Esto ha implicado partir de cero para la reconstrucción y revalidación de toda la biblioteca.

3.2.1. Histórico de resultados OOS

La primera versión de la Biblioteca General de MPs proporcionó bastantes falsos negativos en la prueba de Identidad por Recipiente, hasta unos 18 OOS/mes, por lo que fue necesaria su actualización. Tras la optimización de los parámetros realizada en la segunda versión, se disminuyó este parámetro hasta 4 OOS/mes.

En la nueva versión presentada en este capítulo, se han revisado todos los conjuntos de aprendizaje disponibles y optimizado los parámetros (rango espectral y pretratamiento). Estas acciones se han realizado con el fin de poder disminuir aún más el número de OOS detectados en la Identidad por Recipiente de todos los MPs de la industria.

3.2.2. Construcción de la Biblioteca General

El PRM seleccionado ha sido el mismo que el usado durante la construcción y revalidación de la Biblioteca Específica, el cálculo de la distancia en el espacio de las componentes principales.

El primer paso antes de la construcción de la biblioteca de espectros ha sido la revisión y optimización de los espectros de cada uno de los productos contenidos en el conjunto de aprendizaje. Para ello, se ha seguido la misma estrategia que para la Biblioteca Específica, es decir, representar todos los espectros disponibles, escalarlos, e incluir aquellos que presenten mayor variabilidad.

La estructura final de la biblioteca se muestra en la **Figura V. 13**. En este caso, al incluir muchos productos en una misma biblioteca, se necesita crear hasta 14 niveles y 106 sub-bibliotecas para poder asegurar la discriminación entre ellos. Además, en el nivel

principal (IP1) ha sido necesario usar 11 factores para obtener una buena separación de los materiales y, así, obtener un total de 31 sub-bibliotecas de nivel 2, que proporcionan una buena separación de los productos incluidos en el nivel principal.

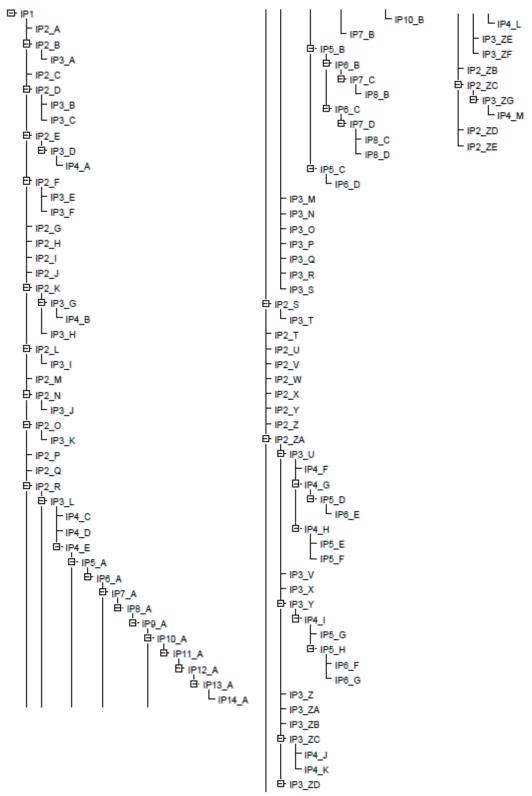


Figura V. 13. Estructura final tras la actualización de la Biblioteca General.

3.2.3. Problemas planteados durante la construcción de la biblioteca

El objetivo de la actualización de la Biblioteca General era identificar inequívocamente todos los MPs usados, pero hay algunos casos en que no ha sido posible discriminar entre todos ellos debido a la alta variabilidad del producto de cada uno de los proveedores, o a dificultades en el registro de los espectros, o a las pequeñas diferencias existentes en las propiedades físicas y químicas de los MPs. Estos casos se han debido abordar individualmente, después de un estudio cuidadoso de las circunstancias implicadas en cada uno de ellos y ha sido necesario usar diferentes estrategias.

El API A es un ejemplo de un producto que presenta una gran variabilidad. Es un producto de origen natural y muy higroscópico que se adquiere de tres proveedores. Las especificaciones del contenido de agua son las mismas para los tres proveedores, pero puede ser que haya importantes diferencias entre lotes, por lo que la variabilidad esperada es muy alta (ver **Figura V. 14**) y se ha decidido crear una clase distinta para cada uno de los proveedores.

El API B es un ejemplo de problemas en el registro del espectro. El espectro de este producto puede registrarse a través de la bolsa de polietileno o, abriendo la bolsa y colocando una funda en la sonda NIR para que no esté en contacto directo con el producto. La **Figura V. 15** muestra los espectros del conjunto de aprendizaje del API B adquiridos de ambas formas. Como puede verse, en el caso de los espectros adquiridos a través del material de envase, las bandas donde absorbe el polietileno presenten intensidades mayores. Estas intensidades se asignan al primer y segundo armónico de los enlaces C-H presentes en su estructura. En cambio, en los espectros adquiridos con la funda en la sonda NIR, las bandas asociadas al propio producto son las que presentan intensidades mayores, mientras que la intensidad de las bandas donde también absorbe el polietileno es mucho menor. En la biblioteca se han creado dos clases para el mismo producto según el método de adquisición de los espectros.

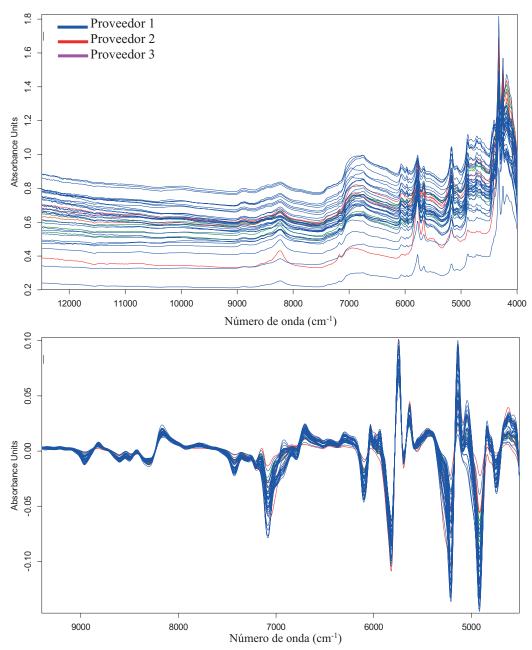


Figura V. 14. Espectros de los distintos proveedores del API A en absorbancia (arriba) y tras la aplicación de una primera derivada (ventana de 13 puntos) más una normalización vectorial.

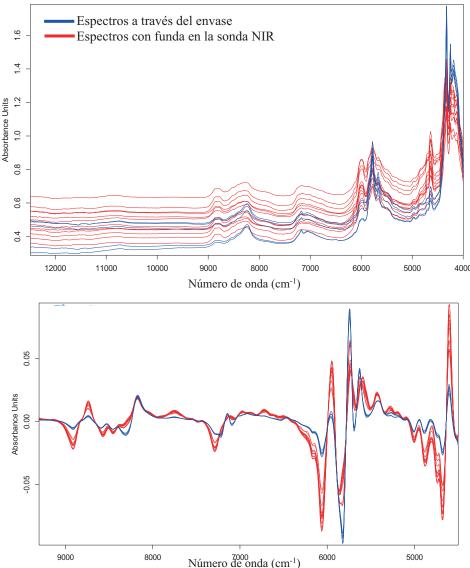


Figura V. 15. Espectros de API B en absorbancia (arriba) y tras la aplicación de una primera derivada (13 pts) para ambos modos de adquisición.

Al identificar materiales con una composición muy similar, se han creado categorías en la biblioteca que incluyen algunos de estos productos. En total, se han creado 39 categorías, donde se pueden encontrar grupos con mucha variabilidad espectral y pequeñas diferencias químicas, lo que impide su discriminación mediante la técnica NIR, o productos con igual composición, pero con grandes diferencias en el tamaño de partícula o en el contenido de humedad.

A modo de ejemplo, en la **Figura V. 16** se muestran los espectros de los productos que constituyen una de las categorías, llamada Ácido Acetilsalicílico. Esta categoría está compuesta por tres productos con igual composición pero que presentan diferencias en el

tamaño de partícula. Como puede verse en la Figura, sí que se observa un cambio en la pendiente para el producto con un tamaño de partícula menor. Al aplicar la segunda derivada, que es el pretratamiento seleccionado en su nivel de identificación (IP3_B), también se pueden apreciar pequeñas diferencias en la intensidad. En el gráfico de *scores* obtenido para este nivel, donde solo se ha usado un único factor, puede verse como el producto con menor tamaño de partícula se encuentra separado de los demás. Sin embargo, una vez se establece el umbral para la clase, dada la gran variabilidad de los productos con mayor tamaño de partícula, no es posible lograr la separación entre ellos. El gráfico de *scores* obtenido se muestra en la **Figura V. 17**.

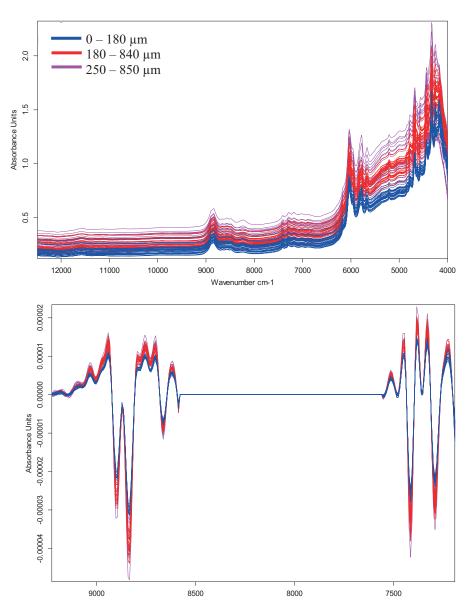


Figura V. 16. Espectros en absorbancia (arriba) y, espectros tras la aplicación de la segunda derivada, con una ventana de 9 puntos, en el rango espectral del nivel IP3 D (abajo).

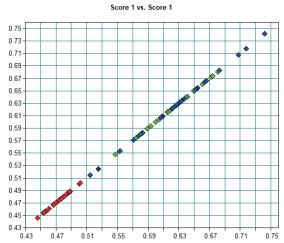


Figura V. 17. Gráfico de *scores* para el nivel IP3_B, donde se encuentran los productos de la categoría Ácido Acetilsalicílico.

3.2.4. Validación interna

El histograma de SEL obtenido tras la actualización de la Biblioteca General se muestra en la **Figura V. 18** donde pueden verse los valores de SEL de la comparación de 158006 pares de clases. De todos estos valores obtenidos, solo un 1,5% se encuentra en el rango de SEL < 1. La mayoría de estos resultados son debidos a la creación de las 39 categorías de la biblioteca. Finalmente, se ha decidido que las clases incluidas en la Biblioteca General no presentan superposiciones.

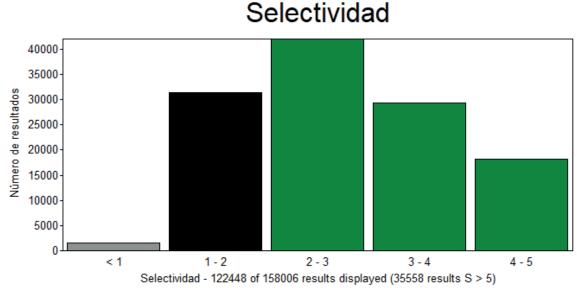


Figura V. 18. Histograma de selectividad obtenido tras la actualización de la Biblioteca General.

3.2.5. Validación externa

La validación externa de la Biblioteca General se realizó evaluando únicamente espectros obtenidos a partir de 2019. Por ello, solo ha sido posible realizar la validación externa de 119 clases de las 398 incluidas en la Biblioteca General. El resto, son productos con una entrada anual de menos de un lote. Todos los productos ensayados han sido identificados correctamente por la biblioteca.

3.2.6. Evaluación de la actualización

La **Figura V. 19** presenta el número de OOS totales y la evolución del tiempo promedio para tener un resultado no conforme en la identidad por recipiente mediante NIR en las diferentes versiones de la Biblioteca General. No se han considerado los OOS de los colorantes.

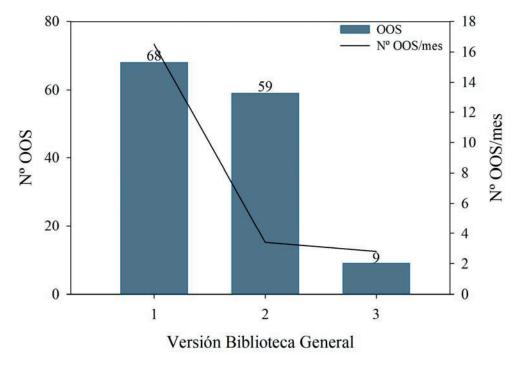


Figura V. 19. Evaluación de las acciones mitigadoras realizadas en la Biblioteca General.

Como puede verse, la primera acción de mantenimiento presentada en este capítulo (versión 3) ha conseguido disminuir el número de OOS mensuales en un 20% respecto a la versión anterior, de 3,4 OOS/mes (con 32 análisis diarios), a 2,8 OOS/mes (con 35 análisis

diarios). A pesar de esta reducción, sigue habiendo un número elevado de productos que presentan un fallo en su identidad por recipiente mediante NIR.

4. CONCLUSIONES

En este trabajo, se han desarrollado y validado bibliotecas de espectros NIR para la identificación de los materiales de partida usados en una industria farmacéutica.

El primer paso para la mejora de la biblioteca existente ha sido separar los colorantes en una Biblioteca Específica, dejando el resto de los productos en la Biblioteca General. En ambas bibliotecas, ha sido necesaria la creación de categorías para productos con composición similar o con la misma composición, pero ligeras diferencias físicas.

Una vez desarrolladas las bibliotecas, primero se ha realizado una validación interna y, posteriormente, una validación externa con muestras no incluidas en el conjunto de aprendizaje. Debido a la poca entrada en la fábrica de algunos productos, la validación externa de todas las clases no ha sido posible. Para los productos habituales, los resultados de la validación externa son satisfactorios. Finalmente, se ha logrado obtener una Biblioteca Específica para identificar 34 colorantes, y una Biblioteca General para identificar 398 materiales de partida, ambas suficientemente robustas para poder identificar muestras de los productos habituales, pero con la posibilidad de ofrecer resultados OOS para productos identificados esporádicamente.

El uso de las nuevas versiones de estas bibliotecas ha reducido de forma muy importante el número de falsos negativos detectados en las versiones anteriores, aunque requieren de una actualización periódica que mejore la variabilidad incluida en la biblioteca.

Es necesario incorporar los nuevos lotes de MPs infrecuentes que puedan llegar en un futuro.

5. BIBLIOGRAFÍA

- [1] Ministerio de Sanidad y Política Social, Anexo 08: Toma de muestras de materiales de partida y acondicionamiento, in: Guía Normas Correcta Fabr. Medicam. Uso Hum. y Vet., n.d.
- [2] Substances for pharmaceutical use, in: Eur. Pharmacopoeia 10.3, **2021**: pp. 4817–4819.
- [3] 1. General notices, in: Eur. Pharmacopoeia 10.7, n.d.: pp. 6341–6350.

- [4] 2.2.40 Near-Infrared spectroscopy, in: Eur. Pharmacopoeia 10.7, **70AD**.
- [5] M. Blanco, M.A. Romero, Near-infrared libraries in the pharmaceutical industry: a solution for identity confirmation, Analyst. 126 (2001) 2212–2217. https://doi.org/10.1039/B105012P.
- [6] L. Dass, Validation of Near-Infrared systems for raw material identification, in: Pract. Approaches to Method Valid. Essent. Instrum. Qualif., 2010: pp. 75–91.
- [7] A. Miró Vera, M. Bernardez, Near-Infrared Spectroscopy in Identification of Pharmaceutical Raw Materials, in: Encycl. Anal. Chem., **2017**.
- [8] M. Alcalà, M. Blanco, D. Moyano, N.W. Broad, N. O'Brien, D. Friedrich, F. Pfeifer, H.W. Siesler, Qualitative and Quantitative Pharmaceutical Analysis with a Novel Hand-Held Miniature near Infrared Spectrometer, J. Near Infrared Spectrosc. 21 (2013) 445–457.
- [9] A. Candolfi, D.L. Massart, *Model Updating for the Identification of NIR Spectra from a Pharmaceutical Excipient*, Appl. Spectrosc. 54 (**2000**) 48–53.

VI. <u>ANÁLISIS DE RIESGO Y MANTENIMIENTO</u> <u>DE BIBLIOTECAS NIR PARA MATERIALES DE</u> <u>PARTIDA</u>

1. INTRODUCCIÓN

Los métodos quimiométricos, tanto cualitativos como cuantitativos, deben ser evaluados regularmente con el fin de garantizar el cumplimiento de su propósito. Este proceso es el que se denomina como gestión de ciclo de vida de los modelos quimiométricos desarrollados [1,2]. La **Figura VI. 1** muestra un esquema del ciclo de vida para un modelo NIR cualitativo, basado en la construcción de bibliotecas de espectros. El ciclo de vida viene determinado por cada una de las etapas presentadas, dado que un modelo NIR es iterativo y continuo.



Figura VI. 1. Esquema del ciclo de vida de un modelo cualitativo NIR.

Durante el ciclo de vida de un modelo NIR, se debe establecer cómo se gestionará la etapa de mantenimiento. En la industria farmacéutica, el mantenimiento de las bibliotecas de espectros para la identificación de MP es fundamental debido a la gran cantidad de MP que se reciben y a la homologación constante de nuevos proveedores y materiales. Como parte de la mejora continua de la robustez y fiabilidad de las bibliotecas, se debe realizar una

acción de mantenimiento después de un periodo de tiempo preestablecido [3]. El principal objetivo de este mantenimiento es la introducción de nueva variabilidad para así evitar la obtención de falsos negativos y falsos positivos durante el análisis en rutina, y obtener modelos fiables y robustos para periodos largos de tiempo.

Una de las posibles herramientas a usar para la identificación de la necesidad de una nueva validación es la realización de un análisis de riesgo, tal como recomiendan varias agencias. Mediante este procedimiento, se identifican los factores que pueden tener influencia en el resultado final de un proceso y se evalúa la incidencia que tendría un posible fallo de los mismos en el resultado final. Se deben proponer medidas correctoras para aquellos factores cuyo fallo conllevaría un error importante en el resultado del proceso estudiado. Estas medidas correctoras mitigarán el riesgo de fallos en el proceso. La Figura VI. 2 muestra un esquema de estas etapas. La primera, valoración del riesgo, consiste en la identificación de los factores que pueden significar un riesgo y en la evaluación de éstos. Una herramienta típica para la identificación de las causas que pueden llevar a un fallo es la realización de un Diagrama de Ishikawa también llamado "Diagrama Causa-Efecto" o "Diagrama de Espina de Pescado" en el que se representan gráficamente las relaciones entre un resultado especifico y sus causas. El efecto o problema estudiado representaría la cabeza del pescado, mientras que las causas potenciales o sub-potenciales serían el esqueleto. Se trata de una representación simple para identificar y entender los factores que influyen en el proceso y sus interacciones [4]. Una vez identificados todos los efectos, se pueden cuantificar usando el análisis modal de fallos y efectos (FMEA, por sus siglas en inglés Failure Mode and Effects Analysis) o, el análisis crítico modal de fallos y efectos (FMECA, por sus siglas en inglés Failure Mode and Effects Critical Analysis) [4–6]. En estos tipos de análisis, se logra identificar los factores que suponen un mayor riesgo para el proceso o problema bajo estudio y cuantificarlos numéricamente. A partir de la cuantificación, se deben establecer acciones para reducir el riesgo de los factores que tengan una mayor influencia para nuestro proceso o problema.

La segunda etapa, control del riesgo, se basa en la búsqueda de acciones mitigadoras para reducir el riesgo de los factores evaluados como riesgo alto, y en la aceptación de riesgos residuales. La última etapa, comprobación del riesgo, se basa en establecer un criterio para la monitorización de los factores y comprobar la eficacia de las acciones mitigadoras propuestas [7].



Figura VI. 2. Esquema de los pasos para la realización de un análisis de riesgos.

El objetivo principal de este trabajo es la reducción del riesgo de obtener falsos resultados OOS para los MP identificados con una biblioteca de espectros NIR. Actualmente, hay estudios donde se han identificado los factores que conllevan un riesgo para el desarrollo de modelos cuantitativos NIR, pero ninguno para la construcción y actualización de bibliotecas de espectros [8–10].

Se realizará un análisis de riesgo para las bibliotecas desarrolladas y validadas en el **Capítulo V** de esta tesis (Biblioteca Específica y Biblioteca General), con el fin de poder realizar acciones que las mejoren cuando se realice su mantenimiento. Una vez identificados y evaluados los factores de riesgo y propuestas unas acciones mitigadoras, se aplicarán revalidando las bibliotecas. Finalmente, se estudiará la eficacia de estas acciones mitigadores evaluando los resultados OOS obtenidos durante el uso en rutina de las nuevas bibliotecas validadas.

2. MATERIALES Y MÉTODOS

Se han utilizado la Biblioteca General y la Biblioteca Específica validadas en el Capítulo V, que contienen, respectivamente, 514 productos y 38 colorantes.

Los nuevos espectros se han registrado con el mismo instrumento y en las mismas condiciones que los utilizados en el capítulo anterior. El software utilizado para la construcción y validación de las nuevas bibliotecas ha sido también el mismo

3. REALIZACIÓN DEL ANALISIS DE RIESGOS

La **Figura VI. 2** muestra el esquema seguido en el análisis de riesgos para ambas bibliotecas.

3.1. Valoración del riesgo

El primer paso para efectuar un análisis de riesgos de una biblioteca NIR es identificar los factores que pueden ser responsables de generar un fallo en la identificación, analizarlos y evaluarlos.

La identificación de los factores se ha realizado mediante un Diagrama de Ishikawa. El esquema utilizado se presenta en la **Figura VI. 3**. Como se observa, se han seleccionado como posibles factores de riesgo los relacionados con el Personal, el Instrumento utilizado, las Condiciones Ambientales de registro del espectro, las Propiedades específicas de cada MP y el Modelo utilizado en la biblioteca de identificación. Para cada uno de estos factores, se han estudiado las posibles influencias de otras variables relacionadas con ellos, identificadas también en el diagrama.

En esta tesis, la evaluación del impacto que conllevan los factores estudiados se ha realizado usando dos tipos distintos de análisis de riesgos: el FMEA y el FMECA. El FMEA se ha usado para evaluar los factores: Personal, Instrumento y Condiciones Ambientales, dado que su impacto en la identificación es el mismo para todos los productos. Sin embargo, en el caso de las propiedades del MP y el modelo construido de la biblioteca, la evaluación del riesgo se ha basado en un FMECA, dado que los resultados son únicos para cada uno de los productos incluidos.

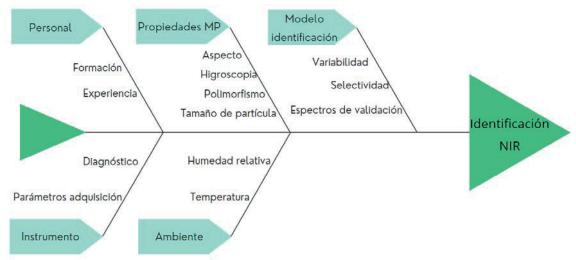


Figura VI. 3. Diagrama de Ishikawa identificando los factores y variables de riesgo de una biblioteca de espectros NIR.

3.1.1. Evaluación del riesgo para los factores "Personal", "Instrumento" y "Condiciones Ambientales"

Como se ha comentado, la evaluación de estos factores se ha realizado mediante un FMEA. El FMEA está basado en la evaluación del impacto de cada variable del factor según los parámetros de severidad (S), probabilidad de aparición (P) y detectabilidad (Dt) del fallo. Según su propia escala subjetiva, a cada uno de los parámetros se le asigna un valor numérico que servirá para el cálculo del número de prioridad del riesgo (RPN, por sus siglas en inglés *Risk Priority Number*) y que será el valor del riesgo de cada variable. Para ello se utiliza la **Ecuación VI-1**.

$$RPN = S \times P \times Dt$$
 (VI-1)

Si el valor de RPN obtenido se encuentra por debajo de un umbral preestablecido, se considera que esa variable no constituye un riesgo para la identificación del producto. Sin embargo, en el caso de que el valor de RPN se encuentre por encima del umbral, se deberá estudiar la variable para encontrar alguna forma de mitigar o reducir su impacto en la identificación de los MP. La descripción de los valores asignados a la severidad, probabilidad, detectabilidad y RPN se encuentran detallados en las **Tabla VI. 1** y **VI.2**, respectivamente.

Tabla VI. 1. Descripción de los valores asignados a la S, P y Dt durante la evaluación de las variables de los factores Personal, Instrumento y Condiciones Ambientales de una biblioteca NIR.

Parámetro	Valor	Descripción
	1	El fallo potencial tiene un impacto bajo en esta variable
S	3	El fallo potencial tiene un impacto intermedio en esta
S		variable
	5	El fallo potencial tiene un gran impacto en esta variable
	1	Muy poco probable
P	3	Ocasional
	5	Frecuente
	1	Capacidad alta de detección
Dt	3	Capacidad media de detección
	5	Capacidad muy baja de detección

Tabla VI. 2. Descripción de la clasificación de los valores de RPN obtenidos para las variables de los factores Personal, Instrumento y Condiciones Ambientales de la biblioteca.

RPN	Acción			
Bajo (RPN < 32)	No influye en la identificación del producto.			
Medio $(32 \le RPN \le 64)$	Poca influencia en la identificación del producto.			
Alto (64 ≤ RPN < 96)	Se requieren acciones dada la influencia en la identificación.			
Muy alto (RPN≥ 96)	Se requieren acciones dada la gran influencia en la identificación.			

3.1.2. Evaluación del riesgo para el factor "Modelo de Identificación"

Este factor se debe estudiar para cada producto de la biblioteca. Su evaluación se ha realizado mediante un FMECA. El objetivo principal de este tipo de análisis de riesgo es el mismo que para el FMEA, con una diferencia en el cálculo de la severidad de cada variable. En este caso, la severidad (S), se calcula según la **Ecuación VI-2**, donde W_i es la importancia, el peso, que se le da a la severidad de cada una de las variables bajo estudio (S_i) y el sumatorio se extiende a las 3 variables relacionadas con el factor (V=3). Los valores asignados a W_i se presentan en la **Tabla VI. 3**.

$$S = \frac{\sum_{i=1}^{V} S_i \times W_i}{\sum_{i=1}^{V} W_i}$$
 (VI-2)

Tabla VI. 3. Descripción de los valores asignados al peso (W_i) para el análisis de riesgo del factor Modelo de Identificación en una biblioteca.

Parámetro	Valor	Descripción		
	1	Impacto bajo para la identificación		
Wi	3	Impacto intermedio para la		
VV 1		identificación		
5 Impacto elevado para la identificació				

Todas las variables relacionadas con el factor "Modelo de Identificación" tendrán un impacto alto en la identificación y el valor de W_i será igual a 5 para todas ellas.

Cada variable que afecta al factor "Modelo de Identificación" es independiente del resto y se evalúa de una forma diferente, por lo tanto, el valor asignado a la severidad de cada una de ellas es también independiente. A continuación, se detalla la asignación de la Si para cada variable bajo estudio.

Variabilidad (VAR). Para cada uno de los productos incluidos en la biblioteca NIR, la evaluación de esta variable se realiza mediante la ponderación del número de años y lotes incluidos en el conjunto de aprendizaje. Esta ponderación se indica en la Ecuación VI-3. Para todos los productos, se asigna el valor de la severidad de la (Tabla VI. 4).

$$VAR = 0.6 \times n^{\circ} \ a\tilde{n}os + 0.4 \times n^{\circ} \ lotes \tag{VI-3}$$

Tabla VI. 4. Descripción de los valores asignados a la Si para la variabilidad de cada producto de la biblioteca.

VAR	Wi	SVAR	Descripción
VAR ≥ 6		1	Conjunto de aprendizaje con gran variabilidad incluida
$3 \ge VAR < 6$	5	3	Conjunto de aprendizaje con variabilidad intermedia incluida
VAR ≤ 3		5	Conjunto de aprendizaje con poca variabilidad incluida

Selectividad (SEL). Esta variable se evalúa mediante los resultados de SEL obtenidos para cada producto con el software OPUS. La asignación de la Si se basa en la Tabla VI. 5. De esta forma, lo que se evalúa es la distancia que hay entre los distintos productos incluidos en la biblioteca NIR. Así, en el caso de que dos productos tengan un valor de SEL bajo, es más probable que puedan llegar a confundirse y, en consecuencia, el riesgo asociado debe ser mayor.

Tabla VI. 5. Descripción de los valores asignados a la Si para la selectividad (SEL) de cada producto de la biblioteca.

SEL	Wi	SSEL	Descripción
\geq 90% con SEL \geq 1.5		1	Buena selectividad
Resto de situaciones	5	3	Selectividad intermedia
≥ 5% con SEL < 1		5	Baja selectividad

Conjunto de validación (VAL). Este parámetro se evalúa para cada producto considerando el número de lotes del conjunto de validación que se han identificado correctamente mediante la biblioteca de espectros construida. Los valores asignados a la Si de la Biblioteca Específica se muestran en la Tabla VI. 6 y los de la Biblioteca General en la Tabla VI. 7.

Tabla VI. 6. Descripción de los valores asignados a la S_i para cada producto del conjunto de validación de la Biblioteca Específica.

VAL	Wi	Sval	Descripción	
$VAL \ge 4$		1	Gran número de lotes evaluados	
$2 \ge VAL < 4$	5	3	Número intermedio de lotes evaluados	
VAL < 2		5	Bajo número de lotes evaluados	

Tabla VI. 7. Descripción de los valores asignados a la S_i para cada producto del conjunto de validación de la Biblioteca General.

VAL	$\mathbf{W_{i}}$	SVAL	Descripción	
VAL ≥ 10		1 Gran número de lotes evaluados		
$5 \ge VAL < 10$	5	3	Número intermedio de lotes evaluados	
VAL < 5		5	Bajo número de lotes evaluados	

La evaluación de la probabilidad (P) de que haya un resultado OOS se ha basado en datos históricos de los OOS obtenidos con versiones anteriores de la biblioteca. Para ello, utilizando la **Ecuación VI-4**, se ha calculado un indicador clave de rendimiento (KPI, por sus siglas en inglés *Key Performance Indicator*).

$$KPI = \frac{n^{\circ} de \ OOS}{bultos \ analizados} \times 100$$
 (VI-4)

Este indicador se ha calculado para la versión 2 de las bibliotecas. Una vez calculado este KPI, se ha asignado a cada producto el valor de probabilidad que aparece en la **Tabla** VI. 8¡Error! No se encuentra el origen de la referencia.

Además, también se le otorga el valor máximo de probabilidad, P=5, a cualquier producto que se incluya por primera vez en la biblioteca NIR para su identificación, ya que no hay histórico de resultados y poca variabilidad. En el caso de que no haya habido entradas de un producto durante el periodo de vigencia de una versión de la biblioteca, no se modificará su valor de probabilidad para la revisión, dado que no se habrá podido evaluar el cambio en el riesgo asignado.

Tabla VI. 8. Asignación de los valores de probabilidad para el FMECA, donde KPI se ha calculado a partir de la última versión de la biblioteca antes de la actualización.

P	Descripción
1	KPI = 0
2	0 > KPI < 25
3	25 ≥ KPI < 50
4	50 ≥ KPI < 75
5	KPI ≥ 75

Por último, para la asignación de los valores de la detectabilidad (Dt), se han considerado los valores presentados en la **Tabla VI. 9**.

Tabla VI. 9. Descripción de los valores asignados a la detectabilidad (Dt) de cada producto para el factor Modelo de Identificación.

Dt	Descripción	
1	Muy alta	
3	Media	
5	Baja	

En este caso, como cualquier fallo en la identificación de un producto es detectado automáticamente por los operadores, se ha asignado, para todos los productos incluidos en la biblioteca, un valor de detectabilidad constante e igual a 1.

El riesgo final se calcula mediante la obtención del valor de RPN (**Ecuación VI-1**), aunque en este caso, la severidad se ha calculado con la **Ecuación VI-2**. Para poder visualizar mejor los resultados, el valor de RPN se ha normalizado a un valor máximo de 100.

3.1.3. Evaluación del riesgo para el factor "Propiedades del MP"

La evaluación de este factor se ha realizado de la misma forma que para el "Modelo de Identificación". Sin embargo, la principal diferencia se basa en la asignación del peso, W_i, de cada variable evaluada. En este caso, el peso asignado no es el mismo para todas las variables, ya que cada una de ellas tiene un impacto distinto en la evaluación.

 Aspecto (ASP). La evaluación de esta variable se realiza mediante la clasificación presentada en la **Tabla VI. 10**, basada en las dificultades de adquisición de los espectros NIR según la presentación del material.

Tabla VI. 10. Descripción de los valores asignados a la Si para el aspecto del producto.

ASP	W_i	SASP
Polvo		1
Gránulo		2
Escamas o cristales	3	3
Pellets		4
Ceras		5

Higroscopia (HIG). La clasificación para determinar la influencia de la higroscopia de los productos en los espectros NIR está basada en la clasificación presentada en la monografía 5.11 de la Farmacopea Europea [11]. En la monografía se describe una asignación dependiendo de porcentaje de incremento en masa (% Δm) tras la realización de una prueba de pérdida por desecación. Los detalles se presentan en la **Tabla VI. 11**.

Tabla VI. 11. Descripción de los valores asignados a la Si para la higroscopia.

HIG	Wi	SHIG	Descripción
Muy baja		1	No se realiza prueba de contenido de agua o
Witty Daja		1	$\% \Delta m < 0.2 \%$
Baja		2	Poco higroscópico: $0.2 \ge \% \Delta m \le 2$
Media	3	3	Higroscópico: $2 \ge \% \Delta m < 15$
Alta		4	Muy higroscópico: % Δm ≥ 15
Mary alta		5	Delicuescente: se absorbe suficiente agua como
Muy alta		3	para formar un líquido.

Tamaño de partícula (TPR). La asignación de la severidad para esta variable se ha basada en si se realiza una estimación de la distribución del tamaño de partícula o no, es decir, si el producto bajo estudio puede presentar distintos tamaños de partícula (Tabla VI. 12).

Tabla VI. 12. Descripción de los valores asignados a la Si para el tamaño de partícula.

TPR	$\mathbf{W_{i}}$	STPR	Descripción
Bajo		1	No tiene especificación de tamaño de partícula
Medio	1	3	Tiene como especificación una única granulometría
Alto		5	Tiene especificación de más de una granulometría

Polimorfismo (POL). Algunos de los productos identificados mediante la biblioteca de espectros NIR pueden presentar polimorfismo según su monografía de la Farmacopea Europea. Como estos productos pueden existir en distintas formas cristalinas, pueden presentar pequeñas diferencias en su perfil espectral, motivo por el cual se ha decidido tener en cuenta esta variable a la hora de asignar los posibles riesgos. La **Tabla VI. 13** muestra la asignación de los valores de severidad para la variable polimorfismo.

Tabla VI. 13. Descripción de los valores asignados a la Si para el polimorfismo.

POL	Wi	SPOL
No presenta	1	1
Puede presentar	1	5

3.2. Control del riesgo

La segunda etapa del análisis de riesgo está basada en la aceptación y reducción del riesgo. A partir de los resultados obtenidos mediante el cálculo del RPN, se debe establecer un criterio para constatar si alguna de las variables estudiadas tiene un impacto tan grande como para obtener un resultado OOS. El criterio utilizado se basa en una representación gráfica de todos los valores de RPN obtenidos. De esta forma, según la tendencia de los resultados, se establecerá un criterio sobre cómo se debe proceder con los mismos.

3.3. Comprobación del riesgo

Una vez identificadas las variables con un mayor impacto en la obtención de resultados OOS, se aplicarán las acciones destinadas a reducir al máximo este riesgo. Una vez aplicadas estas acciones, se comprobará sí han tenido efecto mediante la evaluación de los resultados OOS obtenidos durante el periodo de uso de las bibliotecas.

4. RESULTADOS Y DISCUSIÓN

4.1. Análisis de riesgo para la Biblioteca Específica

Tras la creación y validación de la Biblioteca Específica presentado en el Capítulo IV, se ha realizado un análisis de riesgos para la evaluación del riesgo de obtención de resultados OOS de los colorantes incluidos en ella. En la **Figura VI. 3** se mostraba el Diagrama de

Ishikawa con las variables que se han identificado como posibles fuentes de un resultado OOS para este análisis de riesgos.

4.1.1. Evaluación del riesgo para los factores "Personal", "Instrumento" y "Condiciones Ambientales"

Como se ha mencionado anteriormente, para la evaluación de las variables que tienen un mismo impacto en la identificación de todos los colorantes (Personal, Instrumento y Condiciones Ambientales), se ha realizado un FMEA, Considerando los valores y las descripciones presentadas en las **Tabla VI. 1**y **VI.2**, la **Tabla VI. 14** muestra la matriz de riesgo para estas variables.

Las variables que tienen un impacto mayor en la obtención de un resultado OOS son: la experiencia y la formación de los operarios, la adquisición del espectro de referencia, y las condiciones ambientales de humedad y temperatura de la sala donde se registran los espectros NIR. Para todas ellas se han realizado acciones mitigadoras para disminuir su impacto en los resultados. Tras la realización de las acciones, ninguno de estos parámetros supone un riesgo para el propósito de la biblioteca construida. Sin embargo, según estudios publicados, se ha encontrado que la variabilidad de adquisición entre distintos operarios es una de las mayores fuentes de variabilidad al usar una fibra óptica para el registro de los espectros NIR [12]. Por ello, será crítico la introducción de variabilidad asociado al operador en los conjuntos de calibración, con el fin de reducir la obtención de resultados OOS por esta causa.

4.1.2. Evaluación del riesgo para el factor "Propiedades del MP"

En este caso, todos los colorantes comparten las mismas propiedades de aspecto, higroscopia, tamaño de partícula y polimorfismo, por lo que el impacto de las variables asociadas a ellas es el mismo para todos; los colorantes tienen el mismo aspecto, un polvo con la misma granulometría, sin ninguna particularidad respecto al contenido de agua o polimorfismo y con un bajo impacto en el riesgo de OOS. Por ello, estas variables se han excluido del análisis de riesgo para esta biblioteca.

Tabla VI. 14. Análisis de riesgo para los factores "Personal", "Instrumento" y "Condiciones Ambientales".

	Identificación del riesgo	Consecuencia	S	P	Dt	RPN	Acciones mitigadoras	S	P	Dt	Evaluación del riesgo residual
Personal	Experiencia	Mayor variabilidad espectral entre replicados de bultos y lotes.		3	3	45	Formaciones recordatorias al personal. Inclusión en los espectros del conjunto de aprendizaje del producto variabilidad asociada al operador.	5	1	3	Bajo
	Formación	Error en la colocación de la sonda durante la adquisición de los espectros.	5	3	3	45	Formación al personal sobre una buena adquisición de espectros NIR	5	1	3	Bajo
	Diagnóstico	El instrumento no puede usarse	1	1	1	1	Acciones no requeridas	1	-	1	Bajo
Instrumento	Espectro de referencia	Una incorrecta adquisición del espectro de referencia implicaría la aparición de bandas negativas en el espectro NIR.	5	3	3	45	Formaciones recordatorias al personal sobre cómo detectar este error <i>in-situ</i>	5	1	1	Bajo
	Parámetros de adquisición	El espectro de la muestra sería distinto si cambian los parámetros de adquisición.	3	1	5	15	Acciones no requeridas	-	-	-	Bajo

	Identificación del riesgo	Consecuencia	S	P	Dt	RPN	Acciones mitigadoras	S	P	Dt	Evaluación del riesgo residual
s ambientales	Humedad relativa	La humedad del producto podría verse afecta produciendo cambios en el espectro NIR debido a la absorción de trazas de humedad.	5	3	3	45	Adquisición de los espectros NIR en una cabina de muestreo controlando la humedad relativa.		1	1	Bajo
Condicione	Temperatura	La temperatura puede influenciar cambios en las propiedades de la muestra, resultando en cambios en el espectro NIR.	5	3	3	45	Adquisición de los espectros NIR en una cabina de muestreo controlando la temperatura.		1	1	Bajo

4.1.3. Evaluación del riesgo para el factor "Modelo de Identificación"

La evaluación de este factor se ha realizado mediante un FMECA. El peso, W_i, y los valores de severidad, S_i, de cada variable, así como su KPI, se ha calculado como se indica en la **Sección 3.1.2.** El valor asignado a la detectabilidad, Dt, ha sido 1 (muy alta) en todos los casos, como se ha descrito previamente.

En la **Figura VI. 4** se representa el valor de RPN (normalizado a un valor máximo de 100) para los 34 colorantes incluidos en la Biblioteca Específica. Viendo la tendencia de los valores RPN, se han asignado tres niveles de riesgo asociado: bajo, medio y alto, que se detallan en la **Tabla VI. 15**. Así se puede predecir qué colorantes corren un riesgo mayor de obtener un fallo en su identificación utilizando esta biblioteca específica.

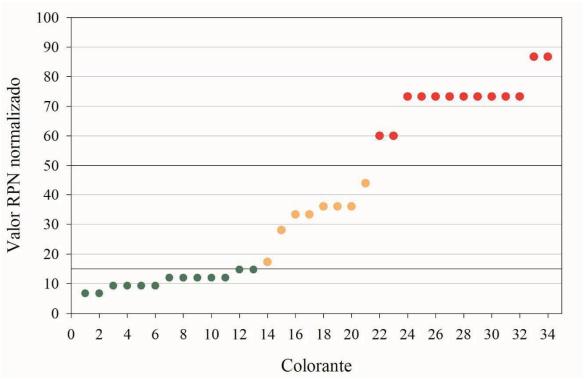


Figura VI. 4. Valores RPN obtenidos durante la evaluación de las propiedades del modelo de identificación de la Biblioteca Específica. Los resultados se muestran en orden ascendente.

Tabla VI. 15. Ratio de valores RPN para la evaluación del factor Modelo de Identificación para la Biblioteca Específica.

RPN	Consecuencia
RPN < 15	Riesgo bajo a obtener un fallo en la identificación NIR
$15 \ge RPN < 50$	Riesgo medio a obtener un fallo en la identificación NIR
$RPN \ge 50$	Riesgo alto a obtener un fallo en la identificación NIR

4.1.4. Conclusiones del análisis de riesgo para la Biblioteca Específica

La evaluación del riesgo residual obtenido para los factores Personal, Instrumento y Condiciones Ambientales, indican que su impacto para la obtención de un fallo en la identificación NIR es bajo. Sin embargo, en la evaluación del Modelo de Identificación, se ha visto que solo 13 colorantes de un total de 34 presentan un riesgo bajo a obtener un error en su identificación, 8 colorantes presentan un riesgo medio y 13 un riesgo alto. Como este riesgo asociado a obtener un resultado OOS solo está relacionado con las variables evaluadas para el Modelo de Identificación, las acciones mitigadoras que se pueden aplicar están directamente relacionadas con la mejora de los conjuntos de aprendizaje y validación, y con la mejora de la selectividad mediante la optimización de los parámetros de discriminación.

A partir de estos resultados, se decidió realizar una revalidación de la Biblioteca Específica al cabo de tres meses de su utilización en rutina para incluir nueva variabilidad en el modelo y optimizar los parámetros de discriminación. El objetivo es reducir el riesgo asociado a obtener un fallo en la identificación de los colorantes estudiados y mejorar así la robustez y fiabilidad de la biblioteca.

4.2. Revalidación de la Biblioteca Específica

Durante los 3 meses siguientes a la construcción de la Biblioteca Específica, entraron en la empresa cuatro nuevos colorantes, por lo que en esta revalidación se discriminará entre 38 colorantes, clasificados en cuatro categorías (I, II, III y IV), dado que ha entrado un nuevo tipo de colorante, Tipo IV, que tiene en su composición uno o más de los polímeros principales de los otros tres, conteniendo, además, sacarosa. Como resumen, ha entrado un nuevo colorante azul de Tipo I, y tres de Tipo IV, de colores blanco, claro y translúcido, respectivamente.

4.2.1. Construcción de la Biblioteca Específica revalidada

Como primera etapa de la construcción de la nueva biblioteca, se han actualizado los conjuntos de aprendizaje y validación disponibles, usando la estrategia de seleccionar aquellos que se encuentren en las posiciones extremas y centrales. Como pretratamiento de los espectros se ha calculado la primera derivada, con una ventana de 13 puntos, seguida de una normalización vectorial.

Una vez optimizados los conjuntos de aprendizaje y validación, se ha empezado el desarrollo del modelo de identificación. Igual que en la versión anterior, como PRM se ha usado la distancia en el espacio de los componentes principales (*Factorization*), estudiando el pretratamiento, rango espectral y número de factores del nivel principal, que permiten obtener mayores valores de selectividad. En esta revalidación, se ha cambiado el pretratamiento del nivel principal (IP1) y el número de factores, consiguiendo la separación entre tipos de colorantes en el tercer nivel, con valores de SEL > 1 entre la mayor parte de compuestos. El pretratamiento seleccionado en el nivel principal, en esta revalidación, ha sido la aplicación de una primera derivada (17 puntos de suavizado) en el rango de 9658,3 - 4204,3 cm⁻¹ con 5 factores. Mientras que en la primera versión de la Biblioteca Especifica se usó una primera derivada (con 9 puntos de suavizado) y una normalización vectorial en el rango de 9658,3 a 4204,3 cm⁻¹ con 4 factores. Este cambio ha permitido mejorar la SEL entre prácticamente todos los colorantes.

Sin embargo, dada la gran similitud en la composición de los tres colorantes blancos de Tipo I, no ha sido posible su discriminación individual y se ha mantenido la categoría Colorante Blanco que ya existía en la biblioteca anterior.

La **Tabla VI. 16** presenta los detalles (% en peso) de la composición de dos colorantes Tipo II (blanco y rosa) que, a pesar de sus claras diferencias en el color, no se pueden discriminar (ya que el espectrómetro NIR disponible no incluye las longitudes de onda del visible) y para los cuales ha sido necesaria la creación de una categoría denominada Colorante II.

Compuesto	Colorante II Blanco	Colorante II Rosa
A	40.0	40.0
В	25.0	21.9
С	20.2	20.2
D	14.8	14.8
E (colorante rojo)	-	1.6
F (colorante amarillo)	-	1.2
G (colorante rojo)	-	0.3

Tabla VI. 16. Composición de los colorantes Tipo II blanco y rosa.

La principal diferencia en la composición de ambos colorantes es la concentración del compuesto B y la presencia de pigmentos para el colorante II rosa. El compuesto B es un compuesto inorgánico que presenta una absorbancia baja en la región del infrarrojo medio,

por lo que su señal en la zona NIR es aún menor. Además, los pigmentos presentes en la formulación del producto rosa representan cada uno menos de un 2% m/m de la composición del producto. Estas pequeñas diferencias entre ambos productos se encuentran en los límites de sensibilidad del instrumento NIR usado. En la **Figura VI. 5** se presentan los espectros NIR de ambos colorantes después de la aplicación de una normalización vectorial, donde se puede ver la similitud en el perfil espectral NIR.

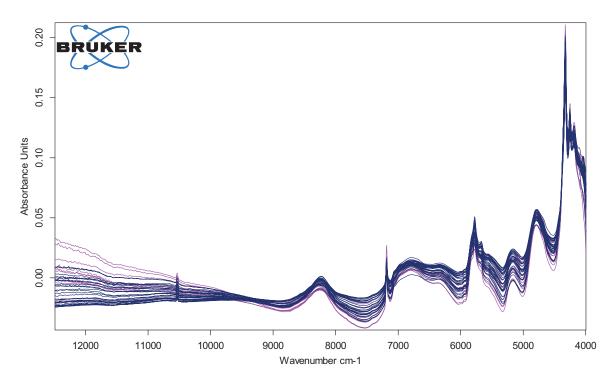


Figura VI. 5. Espectros NIR del Colorante II Blanco (azul oscuro) y Colorante II Rosa (Rosa) en el rango espectral completo del instrumento después de aplicar una normalización vectorial.

Considerando esto, se eliminarán ambos colorantes en la próxima actualización de la biblioteca, dado que la técnica NIR no es adecuada para la identificación de estos materiales, y la identidad por recipiente se realizará mediante una técnica alternativa.

La estructura final de la nueva biblioteca se muestra en la **Figura VI. 6**.

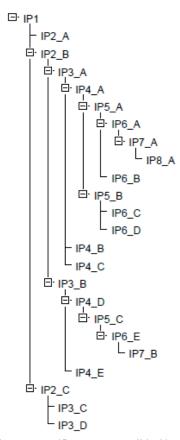


Figura VI. 6. Estructura final de la biblioteca específica tras su revalidación.

4.2.2. Validación interna

El histograma de selectividad obtenido muestra que todos los colorantes tienen una selectividad mayor que 1, con la excepción de los 5 colorantes mencionados anteriormente y para los que se han creado categorías específicas. Con estos resultados, se ha conseguido mejorar la SEL de todos los colorantes incluidos en la biblioteca de espectros NIR. Por ello, el histograma de SEL muestra que no hay ambigüedades entre los grupos incluidos en la biblioteca (**Figura VI. 7**).

4.2.3. Validación externa

La validación externa de la Biblioteca Específica solo fue posible para 14 de los 38 colorantes incluidos en la misma. Además, para un total de 9 colorantes, el conjunto de aprendizaje actualmente disponible es muy pobre y, por ello, el riesgo asociado que tienen a obtener un fallo en la identificación NIR sigue siendo elevado.

Basándonos en los resultados obtenidos, se puede concluir que la biblioteca específica de colorantes ha sido revalidada y, aunque puede usarse para la identificación inequívoca de

los colorantes usados en la industria, existe un alto riesgo de obtener falsos OOS. En futuras actualizaciones de la biblioteca debería incrementarse la variabilidad de alguno de los grupos para reducir así la posibilidad de obtener falsos resultados OOS.

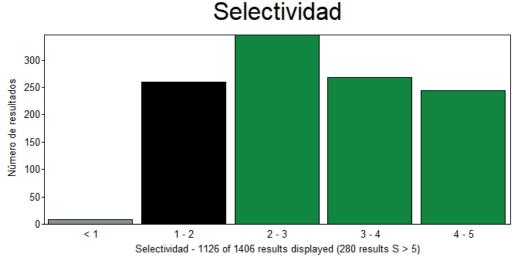


Figura VI. 7. Histograma de selectividad obtenido tras la revalidación de la biblioteca específica como acción mitigadora.

4.2.4. Evaluación de la acción mitigadora realizada

Con el fin de evaluar si la revalidación de la biblioteca específica al cabo de tres meses ha servido para disminuir el riesgo de obtener fallos en la identificación NIR, se ha repetido el análisis de riesgo para el factor Modelo de Identificación mediante un FMECA. De este modo, se quiere evaluar si, tras la introducción de nueva variabilidad y optimización de los parámetros de discriminación, el riesgo asociado a cada colorante ha disminuido.

El análisis de riesgo FMECA se ha realizado del mismo modo que el presentado en la **Sección 3**. La matriz del análisis de riesgo se presenta en la **Tabla VI. 17** para ambas versiones de las bibliotecas, así como el detalle de las acciones realizadas para cada colorante.

ANÁLISIS DE RIESGO Y MANTENIMIENTO DE BIBLIOTECAS NIR PARA MATERIALES DE PARTIDA

Tabla VI. 17. Matriz de análisis de riesgos obtenido tras la revalidación de la Biblioteca Específica.

								Acci	ón mitigad	lora.			
Tipo de colorante	VAR $(W = 5)$	SEL $ (W = 5)$	VAL $(W = 5)$	S	P	RPN ⁽¹⁾	Común: opti	Común: optimización de los parámetros de discriminac			ació	n .	
	(W - 5)	(W - 5)	(W - 5)				Acción	VAR $(W = 5)$	$ SEL \\ (W = 5) $	VAL (W = 5)	S	P	RPN ⁽¹⁾
Colorante I Naranja	1	5	1	2,3	1	9,3	Incluir variabilidad	1	3	1	1,7	1	6,7
Colorante I Blanco	1	3	1	1,7	1	6,7	-	1	3	1	1,7	1	6,7
Colorante I Amarillo	1	1	3	1,7	1	6,7	-	1	1	3	1,7	1	6,7
Colorante I Amarillo	1	3	1	1,7	2	33,3	Incluir variabilidad	1	1	1	1,0	1	12,0
Colorante I Beige	1	3	5	3,0	1	36,0	Incluir variabilidad	1	3	5	3,0	1	36,0
Colorante I Blanco	1	3	3	2,3	1	9,3	Incluir variabilidad	1	3	3	2,3	1	9,3
Colorante I Blanco	1	3	5	3,0	1	12,0	Incluir variabilidad	1	3	5	3,0	1	12,0
Colorante I Blanco	1	1	5	3,0	1	12,0	-	1	3	5	3,0	1	12,0
Colorante I Blanco	3	3	5	3,7	5	73,3	-	3	1	5	3,0	5	60,0
Colorante II Blanco	1	1	3	2,3	1	9,3	Incluir variabilidad	1	3	3	2,3	1	9,3
Colorante II Amarillo	1	1	3	2,3	1	9,3	-	1	3	3	2,3	1	9,3
Colorante II Azul	3	1	5	3,7	4	73,3	-	3	1	5	3,0	4	60,0
Colorante II Rosa	5	3	5	4,3	5	86,7	-	5	1	5	3,7	5	73,3
Colorante II Rojo	5	1	5	3,7	5	73,3	-	5	1	5	3,7	5	73,3
Colorante II Rosa	3	1	3	3,0	5	12,0	-	3	3	3	3,0	5	12,0
Colorante II Rosa	3	1	5	3,0	5	36,0	-	3	3	5	3,7	5	44,0
Colorante II Rosa	5	1	5	3,7	5	73,3	Incluir variabilidad	5	1	5	3,7	5	73,3
Colorante II Rosa	3	1	5	3,0	5	60,0	Incluir variabilidad	3	3	5	3,7	5	73,3
Colorante II Amarillo	5	1	5	3,7	4	44,0	-	5	1	5	3,7	4	44,0
Colorante II Amarillo	3	1	5	3,0	1	36,0	-	3	1	5	3,0	1	36,0
Colorante I Naranja	1	3	1	1,7	2	33,3	Incluir variabilidad	1	3	1	1,7	1	20,0
Colorante I Naranja	5	3	5	4,3	5	86,7	-	5	1	5	3,7	5	73,3
Colorante III Naranja	5	1	3	3,0	5	60,0	-	5	1	3	3,0	5	60,0
Colorante I Blanco	3	1	5	3,0	5	12,0	-	3	1	5	3,0	5	12,0
Colorante I Rosa	5	1	5	3,7	5	73,3	-	5	3	5	4,3	5	86,7
Colorante I Rosa	5	1	5	3,7	5	14,7	-	5	1	5	3,7	3	14,7

								Acci	ión mitigad	lora.			
Tipo de colorante	VAR	<u> </u>		S P RPN ⁽¹⁾		RPN ⁽¹⁾	Común: optimización de los parámetros de discriminación.						1.
	(W = 5)	(W = 5)	(W = 5)				Acción	VAR (W = 5)	SEL (W = 5)	VAL (W = 5)	S	P	RPN ⁽¹⁾
Colorante I Rosa	1	1	5	2,3	1	28,0	-	1	3	3	2,3	1	28,0
Colorante III Marrón	5	1	5	3,7	5	73,3	-	5	1	5	3,7	5	73,3
Colorante III Rosa	5	1	5	3,7	5	73,3	-	5	1	5	3,7	5	73,3
Colorante III Rojo	5	1	5	3,7	5	73,3	-	5	1	5	3,7	5	73,3
Colorante I Amarillo	5	3	5	4,3	5	17,3	-	5	1	5	3,7	3	14,7
Colorante I Amarillo	1	1	5	3,0	1	12,0	Incluir variabilidad	1	3	5	3,0	1	12,0
Colorante I Amarillo	5	1	5	3,7	5	14,7	-	5	1	5	3,7	3	14,7
Colorante II Amarillo	5	1	5	3,7	5	73,3	Incluir variabilidad	5	1	5	3,7	5	73,3
Colorante I Azul	-	-	-	-	-	-	Nuevo colorante	5	1	5	3,7	5	73,3
Colorante IV Translucido	-	-	-	-	-	-	Nuevo colorante	5	1	5	3,7	5	73,3
Colorante IV Blanco	-	-	-	-	-	-	Nuevo colorante	5	1	5	3,7	5	73,3
Colorante IV Claro	-	-	-	-	-	-	Nuevo colorante	5	1	5	3,7	5	73,3

¹El valor de RPN se ha normalizado a un valor máximo de 100.

En la mayoría de los resultados presentados en la **Tabla VI. 17**, en los que ha sido posible incluir variabilidad, el valor de RPN ha disminuido tras la revalidación de la Biblioteca Específica. Sin embargo, hay algún caso en el que este se ha visto aumentado, como es el caso de un Colorante II Rosa. Este aumento en su riesgo viene determinado por la optimización de los parámetros de discriminación (rango espectral y pretratamiento) efectuados. El propósito de modificar estos parámetros era mejorar la SEL entre todos los colorantes. Esta mejora en la SEL se ha visto para la mayoría de ellos, pero concretamente para este colorante no se ha podido lograr una mejora y como consecuencia tiene un riesgo mayor.

Para poder evaluar si el riesgo residual ha disminuido de forma general, la **Figura VI.**8 muestra un histograma con la comparación entre ambas bibliotecas.

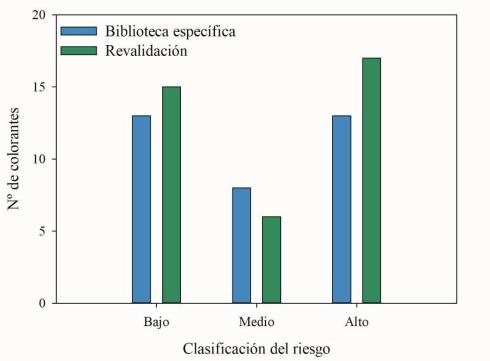


Figura VI. 8. Valores RPN normalizados obtenidos durante la evaluación de las propiedades del modelo de identificación para la biblioteca específica (azul) y tras su revalidación (verde).

Como puede observarse en la **Figura VI. 8**, tras la revalidación de la Biblioteca Específica, el número de colorantes clasificados como riesgo bajo ha aumentado, disminuyendo así el riesgo medio asociado. En el caso del riesgo alto ha aumentado debido a los nuevos cuatro colorantes registrados, pero se ha mantenido constante considerando el resto. En definitiva, sigue existiendo un riesgo elevado de obtener resultados OOS en la identificación NIR de algunos colorantes. La causa es el bajo número de envíos que se

reciben de estos productos, en muchos casos un solo lote y un único bulto, lo que dificulta mucho la inclusión de nueva variabilidad en el conjunto de aprendizaje para construir un modelo robusto. Se deberá realizar un mantenimiento de esta Biblioteca Específica para seguir ampliando los conjuntos de aprendizaje y validación hasta conseguir que el modelo de identificación sea lo más robusto posible. El mantenimiento deberá hacerse tras la llegada de una cantidad suficiente de este tipo de productos.

4.3. Análisis de riesgo de la Biblioteca General

La identificación de las variables que pueden suponer un riesgo para la identidad por recipiente NIR se ha basado en el diagrama de Ishikawa presentado en la **Figura VI. 3**. Sin embargo, en este caso, el factor asociado a las Propiedades de los MPs sí que tiene un impacto importante, dado que cada producto incluido en la biblioteca general tiene unas propiedades físicas distintas.

El riesgo asociado a los factores Personal, Instrumento y Condiciones Ambientales, es el mismo que el evaluado para la Biblioteca Específica (**Sección 3.1.1**), ya que es común para ambas bibliotecas.

Debido a la falta de datos, sólo se ha podido realizar el análisis de riesgos para 409 productos de los 497 incluidos en la tercera versión de la Biblioteca General.

4.3.1. Evaluación del riesgo asociado a los factores: "Propiedades del MP" y "Modelo de Identificación"

Tal y como se ha presentado para la Biblioteca Específica, el análisis de riesgo realizado para la evaluación de estas variables se ha basado en un FMECA. La severidad asignada a las variables Variabilidad, Selectividad y Conjunto de Validación sigue los mismos conceptos que en el caso de la Biblioteca Específica y, debido al mayor número de lotes implicados, los valores concretos se han indicado en la (**Tabla VI. 7**). También se ha utilizado el mismo criterio para la definición de la Probabilidad de aparición del fallo (**Tabla VI. 8**) y la Detectabilidad (**Tabla VI. 9**). La Severidad y el peso asignado a las variables evaluadas se han detallado en las **Secciones 3.1.2** y **3.1.3**.

Una vez estudiada la influencia de todas las variables, se realiza el cálculo de los valores RPN para poder determinar el riesgo asociado a obtener un fallo en la identificación NIR de los productos incluidos en la Biblioteca General. Utilizando los límites de RPN

establecidos en la **Tabla VI. 15**, 275 de los productos tienen un "riesgo bajo", 56 un "riesgo medio" y 78 un "riesgo alto".

De esta clasificación, se han identificado cuántos de los productos se trataban de APIs y cuántos de excipientes (EXCs). La **Figura VI. 9** muestra los resultados obtenidos. En la categoría denominada como "riesgo alto", prevalecen los EXCs mientras que en el "riesgo medio" los APIs. Cabe destacar que, de los 409 productos estudiados, un total de 192 son APIs y 217 son EXCs.

Tras los resultados obtenidos mediante el análisis de riesgos de la Biblioteca General, se puede concluir que un total de 134 productos, que representan un 32,8% del total, tiene un riesgo medio/alto de obtener una identificación NIR no correcta, mientras que 275 productos tienen un riesgo de fallo bajo. La proporción de APIs y EXCs en los riesgos más críticos está desproporcionada siendo los EXCs los productos mayoritarios con una clasificación de "riesgo alto". Este hecho se puede relacionar a la mayor variabilidad de estos productos al ser, principalmente, mezclas.

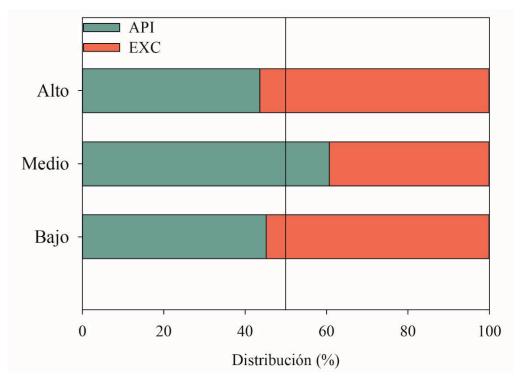


Figura VI. 9. Distribución del % de APIs y EXC clasificados en los tres riesgos (bajo, medio y alto).

4.3.2. Conclusiones del análisis de riesgos de la Biblioteca General

Durante los 95 días que ha estado utilizándose la versión 3 se han realizado 3280 análisis y se han obtenido nueve OOS, todos ellos falsos negativos. Cinco de ellos eran de productos identificados de "riesgo alto", uno de "riesgo medio" y tres de "riesgo bajo". Los tres productos que presentan un riesgo bajo se caracterizan por tener pocos datos, ya sea en el conjunto de aprendizaje o en el de validación, con un valor alto de la severidad. A pesar de ello, como nunca había habido ningún resultado OOS para estos materiales, su probabilidad establecida era baja. No parece razonable modificar los valores establecidos de las variables de riesgo ni las fronteras actuales de RPN.

Como acción mitigadora del riesgo de obtener un fallo en la Identidad por Recipiente NIR, se propone dividir la biblioteca general en dos, una para los APIs y otra para EXCs y mezclas. Esta acción se determina considerando el número total de productos incluidos en la biblioteca y en la distribución presentada en la **Figura VI. 9**.

4.4. Construcción de las nuevas Bibliotecas Generales para APIs y EXCs

Para cumplir con la acción mitigadora, se han creado dos bibliotecas, con 266 productos para la biblioteca de APIs y 248 productos para la biblioteca de EXCs, que incluyen 7 nuevos productos llegados a la empresa antes de esta acción.

Además, se ha realizado una búsqueda exhaustiva de todos los espectros registrados desde 2012 para los productos incluidos en las bibliotecas generales de APIs y de EXCs. Así se ha podido ampliar el conjunto de validación de ambos tipos de materiales de partida con espectros que anteriormente no estaban contemplados.

Con el fin de poder visualizar la mejora al dividir en dos la biblioteca general, en la **Figura VI. 10** se presentan las nuevas estructuras de ambas bibliotecas. La mayor simplificación respecto a la Biblioteca General anterior (con 14 niveles y 106 subbibliotecas) ocurre en la Biblioteca de APIs, que ahora tiene únicamente 6 niveles y 43 subbibliotecas. En el caso de la biblioteca de los EXCs, la estructura sigue siendo compleja, dado que la mayoría de ellos no son productos puros y presentan una mayor variabilidad espectral que los APIs. Por ello, se necesitan 10 niveles y 84 sub-bibliotecas.

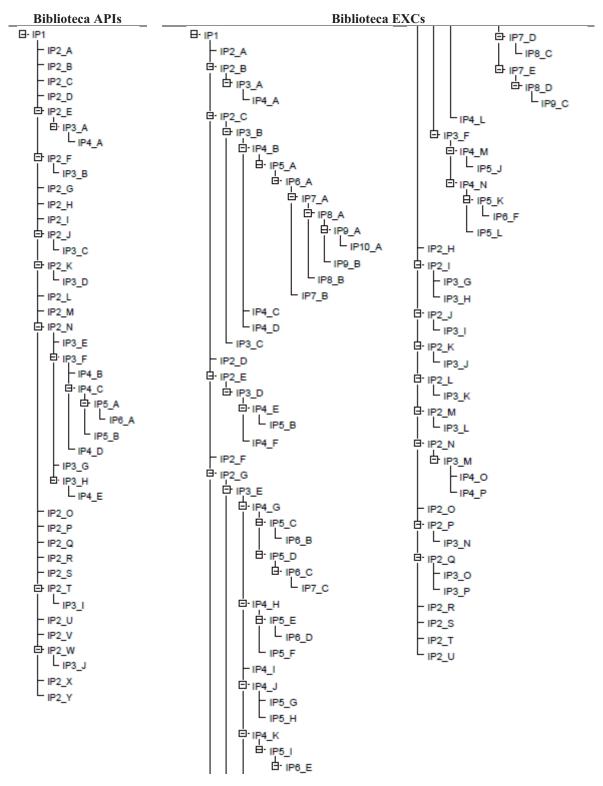


Figura VI. 10. Estructura final de las Bibliotecas Generales para APIs y Excipientes.

4.4.1. Problemas planteados durante la construcción de la biblioteca

En ambas bibliotecas, ha sido necesario crear distintas categorías para poder discriminar entre todos los productos incluidos, tal y como sucedía con la Biblioteca General

única. Concretamente, para la biblioteca de APIs, se han creado 22 categorías y para la de EXCs, 28. Todas las categorías se han basado en la similitud química entre los productos incluidos.

4.4.2. Validación interna

Los histogramas de selectividad obtenidos para ambas bibliotecas específicas se muestran en la **Figura VI. 11**, donde pueden verse los valores de SEL al comparar 32942 pares de clases en la Biblioteca de APIs y 53592 pares en la Biblioteca de EXCs. Sumando ambas bibliotecas, solo un 1,2% de los valores obtenidos se encuentra en el rango de SEL < 1. Todos estos valores bajos de SEL corresponden a los productos que se encuentran agrupados en las distintas categorías. Se puede concluir que no hay ambigüedades entre las clases que forman cada biblioteca.

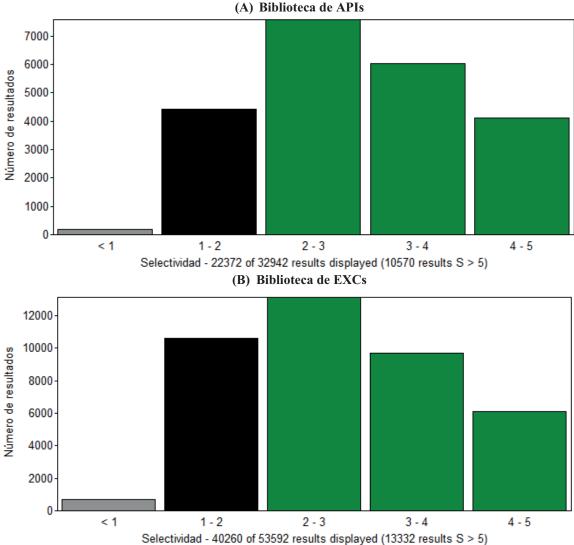


Figura VI. 11. Histogramas de selectividad obtenidos para las bibliotecas de (A) APIs y (B) excipientes.

4.4.3. Validación externa

La validación externa de ambas bibliotecas, APIs y EXCs, se ha realizado evaluando espectros obtenidos a partir de 2012, mejorando así el conjunto de validación para la mayoría de los productos incluidos, ya que en la Biblioteca General estudiada anteriormente solo contenía espectros de 2019 a 2021. Se ha pasado de 109 productos distintos, en el conjunto de validación de la Biblioteca General, a 209 APIs y 196 excipientes. En concreto, para la validación externa, se han evaluado un total de 24193 espectros de APIs y 21736 de excipientes, todos con resultados correctos.

Basándonos en los resultados obtenidos, se puede concluir que ambas Bibliotecas han sido desarrolladas y validadas para identificar inequívocamente tanto los APIs como los excipientes utilizados en la planta.

4.4.4. Evaluación de la acción mitigadora realizada

La evaluación de la acción mitigadora realizada (división de la Biblioteca General en dos), se ha realizado analizando la disminución del número de OOS hallados. Para esta comparación, se han evaluado los resultados OOS generados por la versión desarrollada y validada en el Capítulo V (versión 3) y los obtenidos con la Biblioteca de EXCs y de APIs (versión 4).

La Figura VI. 12 presenta el número de OOS totales durante el tiempo de utilización de las versiones 3 y 4 de la Biblioteca General. La versión 4 de la Biblioteca General contempla su división en dos, una biblioteca exclusivamente de APIs y otra de EXCs. En la figura, también se muestra la evolución del tiempo promedio de aparición de un OOS (sumando los de ambas bibliotecas de la versión 4). La versión 4 se ha utilizado durante 125 días, realizando 7387 análisis, y el número de OOS ha caído prácticamente a la mitad (cinco casos), aunque se ha utilizado casi un 50% más, ya que se ha pasado de analizar 35 espectros/día con la versión 3 a 59 espectros/día con la versión 4. Se han obtenido únicamente 2 casos de OOS en productos de "riesgo alto", ninguno de "riesgo medio" y siguen habiendo 3 de "riesgo bajo", distintos a los detectados con la versión 3. Como en el análisis de la versión anterior, los clasificados como "riesgo bajo" presentan una severidad alta, dado que tienen un número bajo o ningún espectro en el conjunto de validación. Sin embargo, al no tener un histórico previo, la probabilidad que se les había asignado era baja.

La tendencia de la biblioteca será la disminución de los valores de S, debido a la introducción de más datos en los conjuntos de aprendizaje y/o validación y a la disminución de P debido al menor número de OOSs. Para cambiar los valores actuales que se utilizan en el análisis de riesgos para nuevas versiones de la biblioteca, se debería esperar más tiempo y analizar las causas de los pocos OOS que se vayan produciendo hasta el momento.

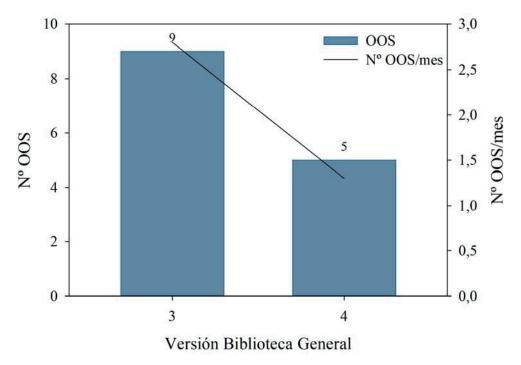


Figura VI. 12. Evaluación de las acciones mitigadoras realizadas en la Biblioteca General.

5. CONCLUSIONS

En este trabajo se ha utilizado el análisis de riesgo para identificar factores que tienen un impacto directo durante el análisis de identidad por recipiente de los productos incluidos en las bibliotecas NIR de una industria.

Tras la valoración de los riesgos identificados, se ha establecido un criterio para poder determinar qué factores y productos son los que tienen un riesgo más elevado a obtener un fallo en su identificación. Con los resultados, se han aplicado acciones mitigadoras para poder reducir este riesgo asociado. En el caso de la Biblioteca Específica, se ha revalidado optimizando los conjuntos de aprendizaje y validación, y se han optimizado los parámetros de discriminación para mejorar la selectividad entre los distintos colorantes. En el caso de la Biblioteca General se ha decidido dividirla en dos bibliotecas, una para APIs y otra para excipientes, para poder mejorar las características del modelo de identificación. Una vez

construidas las nuevas bibliotecas, se han vuelto a validar ampliando los conjuntos de validación, incluyendo datos desde 2012, con lo que se ha conseguido una mayor fiabilidad en la identificación de muestras desconocidas.

La eficacia de estas acciones mitigadoras ha sido comprobada mediante la evaluación de la reducción del número de resultados OOS obtenidos. De la versión 3 a la 4 de la biblioteca, la reducción en los resultados OOS es, aproximadamente, de un 50% y el número de análisis el doble. Por ello, se puede concluir que la ayuda del análisis de riesgos para entender la causa de los OOS y la estrategia a seguir en la construcción de bibliotecas ha permitido una mejora en la robustez y fiabilidad de las bibliotecas NIR construidas.

6. BIBLIOGRAFÍA

- [1] EMA, Guideline on the use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data requirements for new Submissions and Variations, Eur. Med. Agency. 44 (2014) 1–28.
- [2] European Pharmacopoeia 10.0., Chapter 5.21. Chemometrics Methods Applied to Analytical Data, (2021) 819–836.
- [3] M.S. Kemper, L.M. Luchetta, *A Guide to Raw Material Analysis Using near Infrared Spectroscopy*, J. Near Infrared Spectrosc. 11 (2003) 155–174. https://doi.org/10.1255/jnirs.363.
- [4] L. Liliana, A new model of Ishikawa diagram for quality assessment, {IOP} Conf. Ser. Mater. Sci. Eng. 161 (2016) 12099. https://doi.org/10.1088/1757-899X/161/1/012099.
- [5] Z. Bluvband, P. Grabov, Failure analysis of FMEA, in: 2009 Annu. Reliab. Maintainab. Symp., 2009: pp. 344–347. https://doi.org/10.1109/RAMS.2009.4914700.
- [6] D.. Stamatis, Failure Mode Effects and Criticality Analysis (FMECA), in: Risk Manag. Using Fail. Mode Eff. Anal., **2019**: pp. 97–112.
- [7] The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use., ICH Q9 on quality risk management, in: **2006**.
- [8] Y. Kato, D. Sasakura, T. Miura, A. Nagatomo, K. Terada, Evaluation of risk and

- benefit in the application of near-infrared spectroscopy to monitor the granule coating process., Pharm. Dev. Technol. 13 (2008) 205–211. https://doi.org/10.1080/10837450801949434.
- [9] H. Lim, V.S. Dave, L. Kidder, E. Neil Lewis, R. Fahmy, S.W. Hoag, *Assessment of the critical factors affecting the porosity of roller compacted ribbons and the feasibility of using NIR chemical imaging to evaluate the porosity distribution*, Int. J. Pharm. 410 (**2011**) 1–8. https://doi.org/10.1016/j.ijpharm.2011.02.028.
- [10] C.C. Corredor, R. Lozano, X. Bu, R. McCann, J. Dougherty, T. Stevens, D. Both, P. Shah, Analytical Method Quality by Design for an On-Line Near-Infrared Method to Monitor Blend Potency and Uniformity, J. Pharm. Innov. 10 (2015) 47–55. https://doi.org/10.1007/s12247-014-9205-2.
- [11] 5.11 Characters section in monographs, in: Eur. Pharmacopoeia 10.7, n.d.: p. 6531.
- [12] K.B. Bradfield, R.A. Forbes, Development and Validation of an Analytical Method for Identification of Granulated Nicarbazin by near Infrared Reflectance Spectroscopy, J. Near Infrared Spectrosc. 5 (1997) 41–65. https://doi.org/10.1255/jnirs.99.

VII. <u>DETERMINACIÓN DEL CONTENIDO DE</u> <u>IBUPROFENO EN GELES COMERCIALES</u>

1. INTRODUCCIÓN

Los geles son preparaciones sólidas o semisólidas que contienen una base líquida gelificada mediante el uso de un gelificante adecuado. En este tipo de muestras, el principio activo puede encontrarse disuelto, suspendido o emulsionado en la base, con tal de adaptar el mejor medio para el API. Normalmente, este tipo de productos tienen un alto contenido en agua (más de un 80% m/m aproximadamente), que se usa para agilizar la absorción del API en el tejido.

Los análisis cuantitativos de este tipo de muestras resultan ser laboriosos y requieren de un gran tiempo de análisis, dada la necesidad de la preparación previa de las mismas. Los métodos más usados para la determinación del contenido del API suelen basarse en determinaciones cromatográficas [1,2]. Sin embargo, existen distintos estudios donde se han usado técnicas como el NIR, IR, y UV para estos análisis. La **Tabla VII. 1** muestra un resumen de algunas de estas aplicaciones.

Tabla VII. 1. Resumen de artículos con estudio de la determinación del API en geles o derivados [3–9].

Técnica	Resumen	Referencia
NIR	Determinación del contenido de Ketoprofeno en muestras de gel translúcidas preparadas en el laboratorio.	[3]
NIR	Determinación del contenido de Dexketoprofeno en un hidrogel,	[4]
NIR	Determinación del contenido de Ketoprofeno y dos conservantes en un hidrogel.	[5]
NIR	Determinación del contenido de Nistatina y Metrodinazol en una crema vaginal.	[6]
NIR	Determinación del contenido de API en seis formulaciones distintas (crema y vaselina).	[7]
UV	Determinación del contenido de Indometacina y Ácido benzoico en geles durante su proceso de fabricación.	[8]
NIR	Determinación del contenido de dos APIs en dos formulaciones de geles distintas, probando distintas configuraciones ópticas.	[9]

La aplicación de NIR e IR como técnicas alternativas a las tradicionales llama la atención, ya que en ambas técnicas el agua absorbe con gran intensidad, lo que puede llegar a ser una desventaja dependiendo de la zona de absorción del API. Por ello, una buena alternativa seria la aplicación de la espectroscopia Raman, dado que el agua no presenta apenas señal (dispersión débil), al contrario que en el caso del NIR e IR. Sin embargo, esta técnica no ha sido muy empleada para esta finalidad. T.R.M. de Beer *et.al.* desarrollaron y validaron un modelo univariante mediante la preparación de distintos estándares de la muestra a distintos niveles de concentración, para determinar el contenido de ácido salicílico en cremas [10]. S. Mazurek *et al.* demostraron la posibilidad de la aplicación de la espectroscopia Raman para la determinación del contenido de cuatro APIs en dos geles tópicos y un ungüento [11]. Los resultados que obtuvieron muestran la eficiencia de la construcción de un modelo PLS para la determinación del API mediante la espectroscopia Raman.

El objetivo principal del presente trabajo es evaluar la capacidad de un instrumento Raman portátil para la determinación del contenido del API (Ibuprofeno) en geles comerciales mediante la aplicación de una adición estándar. Al realizar la cuantificación por una adición estándar, lo que se quiere es reducir el tiempo de análisis al no realizar la preparación previa de un conjunto de muestras de calibración y validación para la construcción de un modelo cuantitativo. Mediante la adición de un patrón con concentración conocida de API, se extrapolará la concentración de éste en la muestra.

2. MATERIALES Y MÉTODOS

2.1. Muestras

El gel analizado en este trabajo es Novidol 50 mg/g gel, comercializado por Laboratorios Menarini S.L. Este producto contiene como principio activo ibuprofeno (IBU) en una proporción aproximada del 5% m/m, dado que cada tubo de gel (60 g) contiene 50 mg de IBU por gramo de producto. Como agente gelificante contiene hidroxietilcelulosa y alcohol bencílico e isopropanol como excipientes, además de agua.

Para la realización del trabajo, se han usado tres lotes de producto acabado. Concretamente los lotes usados han sido: P016, P019, P020. Además, diariamente se ha preparado el patrón para la realización de la adición estándar que está compuesto de una sal

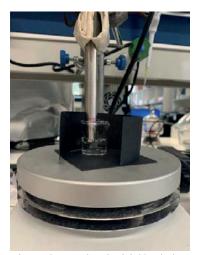
sódica de ibuprofeno (IBUNa). Los detalles de los productos usados se encuentran en el capítulo III (sección 2).

Tanto los lotes comerciales como el estándar preparado se han analizado por HPLC siguiendo el método presentado en el Capítulo III (Sección 1.4.1) con el fin de obtener la concentración real de ibuprofeno.

2.2. Adquisición de espectros

Los espectros de las muestras de gel han sido adquiridos con el instrumento i-Raman plus equipado con la sonda BAC101. Como la sonda BAC101 puede estar en contacto con la muestra, se ha fijado una posición para que quede siempre a la misma altura.

En la **Figura VII. 1** se puede observar el montaje usado para la realización de los experimentos. Se ha usado una cartulina negra para cubrir el vaso de precipitados con la muestra para minimizar la posible interferencia causada por la luz ambiental.



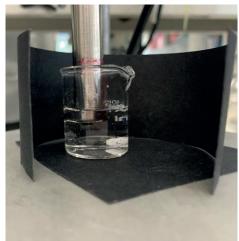


Figura VII. 1. Montaje usado para la adquisición de los espectros de las muestras de gel con el instrumento i-Raman plus.

Los parámetros de registro usados han sido: 1 único barrido, con un tiempo de integración de 15 s, y la potencia de láser ajustada a un 80%, que equivale aproximadamente a 277 mW, en el rango espectral de 2000 – 500 cm⁻¹, con un intervalo de número de onda de 1 cm⁻¹. Estas condiciones de registro se han determinado a partir de los resultados de estudios previos realizados, donde se concluyeron los parámetros más importantes por controlar en el instrumento i-Raman plus [12,13].

3. RESULTADOS Y DISCUSIÓN

El primer paso para la realización de una adición estándar utilizando Raman, es garantizar que hay bandas relacionadas con el API que se pueden identificar en el producto acabado; ya que, al tratarse de un análisis univariante, lo que se hará será relacionar directamente la intensidad Raman de un pico concreto con un valor de concentración.

En la **Figura VII. 2**, se muestra el espectro puro de IBU en forma sólida obtenido con un instrumento de mesa (MultiRaman), así como el espectro de una muestra de gel adquirida con el i-Raman plus. En el espectro del gel comercial, se puede ver que la banda que presenta una mayor intensidad se encuentra a un número de onda de 815 cm⁻¹. Esta banda tan intensa no puede asociarse a la presencia de IBU en la muestra, sino a la presencia del excipiente isopropanol, y corresponde al enlace C-C-O de la cadena alifática de esta molécula [11]. Una de las bandas que se pueden identificar claramente tanto en el API puro como en la muestra comercial es la que se encuentra alrededor de 1611 cm⁻¹, que corresponde al anillo aromático del API [14]. Sin embargo, hay que tener en cuenta que el alcohol bencílico también presenta una banda a 1606 cm⁻¹, que corresponde también al anillo aromático. En la **Figura VII. 3** se presenta el espectro del API solapado con el espectro del isopropanol y alcohol bencílico. Como se observa, la banda de alrededor de 1606 cm⁻¹ del alcohol bencílico, es una de las bandas menos intensas en su espectro Raman, por lo que la contribución de este producto a la banda aromática en el producto comercial será muy baja. En el caso de la hidroxietilcelulosa, como no presenta ningún anillo aromático, no interfiere en la intensidad del pico de 1611 cm⁻¹.

La realización del estudio univariante mediante la adición estándar se realizará siguiendo la intensidad de la banda aromática (1611 cm⁻¹) identificada anteriormente. Antes de poder relacionar la intensidad de la banda con la concentración de IBU es necesario realizar un pretratamiento de la señal para eliminar el efecto de la línea de base del espectro.

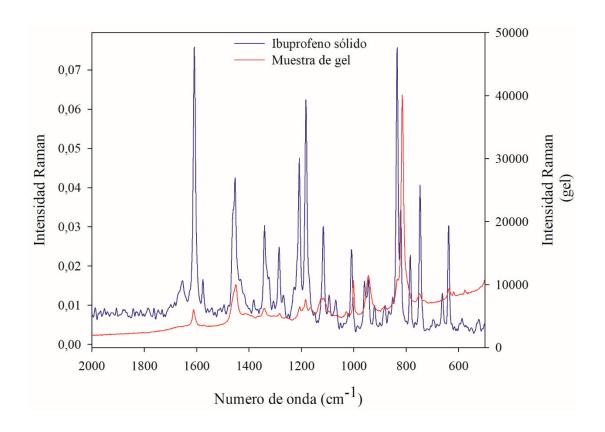


Figura VII. 2. Comparación del espectro puro de IBU con el espectro de una muestra del gel comercial.

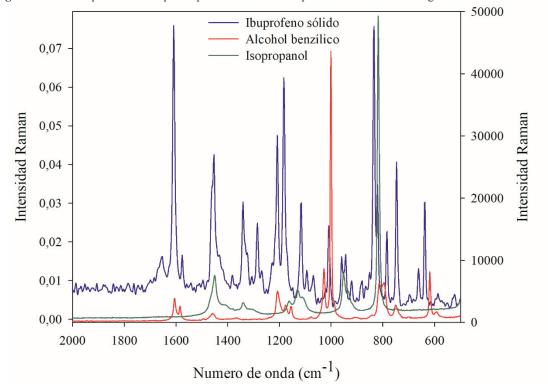


Figura VII. 3. Comparación del espectro puro de IBU con el espectro del alcohol bencílico e isopropanol adquiridos con el i-Raman.

3.1. Selección del pretratamiento

El pretratamiento de los espectros se ha realizado con el software SOLO (Eigenvector Research). Como puede apreciarse en la **Figura VII. 2.**, el espectro del gel comercial presenta una ligera pendiente, disminuyendo con el número de onda del desplazamiento Raman. Como el objetivo del presente trabajo es relacionar la intensidad de la banda de 1611 cm⁻¹ directamente con la concentración de IBU, es necesario corregir la pendiente de la línea base. Además, ya que se va a ir añadiendo una cantidad de IBUNa en la muestra, indirectamente se irá diluyendo la concentración de excipientes, dado que el patrón sólo incluye agua como disolvente. Por ello, se debe encontrar una combinación de pretratamientos que permita corregir tanto el efecto de la línea base como el efecto de dilución de la muestra.

La combinación de pretratamientos empleada ha sido la siguiente:

- (1) Suavizado del espectro utilizando el método de Savitzky-Golay, con un polinomio de orden 2 y una ventana móvil de 15 puntos en todo el intervalo del espectro (2000 500 cm⁻¹), con el fin de minimizar el ruido del registro.
- (2) Corrección de la curvatura de la línea base en el intervalo 1680 700 cm⁻¹, mediante la corrección *Automatic Whittaker Filter*, Este método se utiliza cuando en el espectro hay algunas variables que sólo dependen del fondo y permiten saber la cantidad de señal que es debida a esa línea de base y que puede eliminarse del resto del espectro. Utiliza un método de mínimos cuadrados por partes (*piecewise*) para eliminar automáticamente el fondo sin crear picos negativos. Los parámetros utilizados en este caso para el control del ajuste son λ= 1000, para el control de la curvatura que se permite que tenga la línea de base (cuanto menor es λ se permite una mayor curvatura) y un valor P de 0,001. Este valor P controla la asimetría; cuanto mayor es el valor, más fácil es que aparezcan picos negativos.
- (3) Normalización del espectro, dividiendo su valor por la intensidad del pico más intenso (814 cm⁻¹) que corresponde al isopropanol. De esta forma se corrige el efecto de la dilución de los excipientes en cada adición.
- (4) Una nueva corrección de la línea de base en el intervalo 1635 1590 cm⁻¹, utilizando la misma técnica y parámetros que en el caso anterior, para asegurar que la línea de base en la región donde se encuentra el pico deseado es completamente plana.

La **Figura VII. 4** Figura VII. 4. Comparación de un espectro sin pretratar y del mismo espectro después de la aplicación de pretratamientos presentada. muestra un espectro de una muestra de gel sin pretratar y el mismo espectro tras la aplicación de la combinación de pretratamientos.

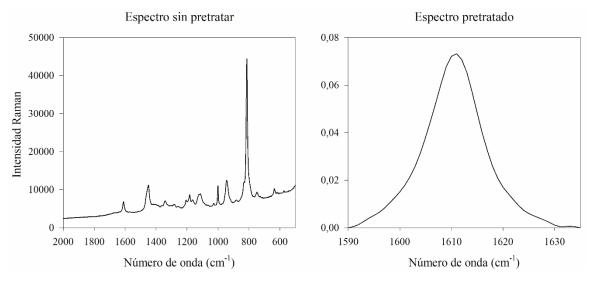


Figura VII. 4. Comparación de un espectro sin pretratar y del mismo espectro después de la aplicación de pretratamientos presentada.

3.2. Optimización de las condiciones hidrodinámicas

Debido a que el gel es viscoso, es necesario estudiar el tiempo necesario para que, después de realizar cada adición, se estabilice la intensidad del pico Raman estudiado y evaluar el tiempo que se ha de esperar entre adiciones. Otro factor importante es el efecto de la agitación en las medidas del gel puro. Para el estudio de estas variables, se preparó una muestra de gel comercial y se configuró un experimento para que se fueran adquiriendo espectros en continuo cada 35 s, usando un agitador mediano y a una velocidad de 500 rpm Se adquirieron espectros solamente agitando el gel puro durante 40 min; luego se realizó una primera adición de 200 µL de un patrón de IBUNa (0,25 g/mL) y se agitó durante otros 40 min. Finalmente, se realizó una segunda adición del mismo patrón y se agitó de nuevo durante 30 min. La variación de la intensidad del pico frente al tiempo, tras la aplicación en cada espectro del pretratamiento descrito anteriormente se presenta en la **Figura VII. 5**.

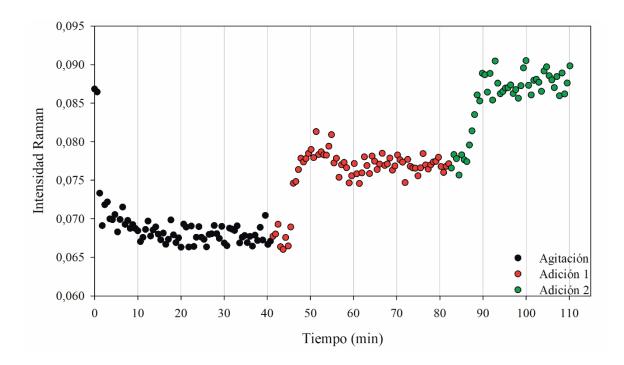


Figura VII. 5. Gráfico de serie de tiempo evaluando la intensidad Raman (1611 cm⁻¹) con el tiempo.

Los resultados muestran que la intensidad del pico disminuye al agitar la muestra, este hecho se puede correlacionar con la obtención de una capa de gel gruesa en la superficie de la sonda. Esta disminución se estabiliza a partir de los primeros 20 min de agitación. Además, tras la primera adición, se puede observar el incremento en la intensidad de la banda, pero que este incremento no es instantáneo, sino que tarda alrededor de otros 20 min en estabilizarse. Para la segunda adición, se observa el mismo patrón que para la primera, con la única diferencia de que la intensidad se estabiliza antes, más o menos tras unos 10 min.

El tiempo necesario para poder observar el incremento en la intensidad disminuye a medida que realizamos un mayor número de adiciones del patrón. Esto puede estar relacionado con la disminución de la viscosidad de la muestra, que favorece la desaparición de la capa de gel localizada en la sonda de medida cuando se agita la muestra.

Considerando estos factores, se decidió que era necesario un tiempo de agitación de la muestra inicial de 20 min antes de realizar la primera adición y del mismo tiempo antes de la segunda. A partir de aquí, con solo 10 min era suficiente para la estabilización de la señal del pico estudiado.

3.3. Diseño de experimentos

Una vez que se ha logrado identificar una banda que se puede relacionar directamente con la concentración de IBU en la muestra y se han establecido unas condiciones de pretratamiento, se ha realizado un estudio para encontrar las mejores condiciones para la homogeneización de la muestra. Para ello se ha realizado un diseño de experimentos (DOE, por sus siglas en inglés *Design of Experiments*). Se ha utilizado un DOE basado en un diseño central cúbico centrado en las caras (CCF, por sus siglas en inglés *Central Composite Face-center design*) [15,16]. Los dos factores considerados para la construcción del diseño han sido la velocidad de agitación (revoluciones por minuto, rpm) y el tamaño del agitador magnético. Cada uno de ellos se ha estudiado a tres niveles distintos. Para la velocidad de agitación, los niveles han sido 300, 500 y 700 rpm; y, en el caso del tamaño del agitador magnético, pequeño, mediano y grande, codificados en la tabla como -1, 0 y 1, respectivamente. (**Figura VII. 6.**).



Figura VII. 6. Tamaños del agitador magnético usados de menor a mayor tamaño.

De los tres agitadores usados, el mediano tiene un abultamiento en la parte central que no existe en los otros dos tamaños. Con todos estos parámetros, se construyó el diseño CCF realizando tres repeticiones en el punto central, obteniendo un total de 11 experimentos a realizar (**Figura VII. 7**). Para la realización de éstos, se ha seguido el procedimiento experimental presentado en la **Figura VII. 8**., efectuando una única adición del patrón IBUNa (0,25 g/mL). Para cada uno de los experimentos, se han registrado 5 espectros Raman tras cada uno de los tres pasos experimentales realizados.

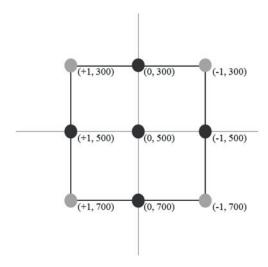


Figura VII. 7. Diseño CCF construido para la optimización de la homogeneización del gel.



Figura VII. 8. Esquema del procedimiento experimental llevado a cabo para el DoE. En gris oscuro se representa la sonda Raman, en blanco el agitador, y en gris claro la muestra de gel.

Figura VII. 9 muestra el incremento de la intensidad del pico tras la adición del patrón de IBUNa al utilizar cada tipo de agitador y a distintas velocidades, así como el valor de la intensidad de la muestra inicial. Como puede apreciarse en la Figura VII. 9., al usar el agitador magnético más pequeño, el aumento en la intensidad Raman tras la adición del patrón de IBUNa es casi inapreciable, dado que como máximo se observa un incremento de 0,002 unidades de intensidad al usar la máxima velocidad de agitación. Con este agitador también se observa una disminución de la intensidad a bajas revoluciones, lo que indica la baja efectividad de la agitación con este tamaño de agitador. Se consigue el mismo resultado

(un incremento de 0,01 unidades de intensidad) utilizando el agitador de mayor tamaño a 300 rpm y el agitador mediano a 500 rpm. El uso de más revoluciones empeora el resultado.

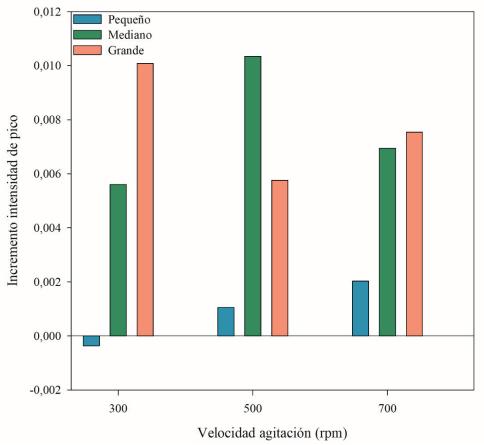


Figura VII. 9. Incremento en la intensidad del pico tras la aplicación del pretratamiento espectral para los experimentos realizados del DoE.

3.4. Estudio de la repetibilidad

Se estudió la repetibilidad de la medida usando las dos condiciones que, a partir del DOE, proporcionaban un aumento mayor en la intensidad del pico.

Como recomienda la guía ICH Q2(R2), se han realizado seis determinaciones del incremento de la intensidad, es decir, se ha realizado el mismo experimento seis veces usando las mismas condiciones [17].

Primero, se realizó el estudio con el agitador grande y una velocidad de 300 rpm. Los resultados obtenidos para la intensidad del pico se muestran en la **Figura VII. 10.** Para cada experimento, las barras de error indican la desviación estándar de los 5 espectros registrados. Se observa una enorme diferencia entre los valores obtenidos al sumergir directamente la sonda en el gel, sin agitar, y los obtenidos cuando se realiza la agitación, Esto indica que es

muy importante asegurarse de una agitación correcta y reproducible para evitar la acumulación de gel en la sonda. Al realizar una adición de 200 μ L del patrón de IBUNa (0,25 g/mL) el incremento medio es de 0,01 unidades.

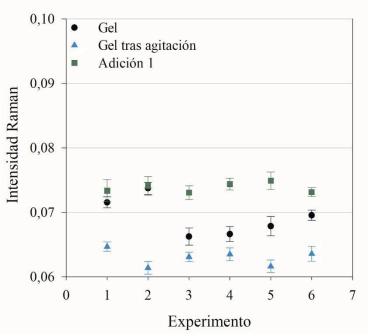


Figura VII. 10. Representación de la intensidad Raman del pico bajo estudio usando el agitador magnético grande y 300 rpm. Las barras son representadas por la desviación estándar de los 5 espectros registrados.

Al realizar el estudio de la repetitividad con el agitador mediano y a 500 rpm se decidió realizar dos adiciones para ver el comportamiento. La representación del incremento en la intensidad se presenta en la **Figura VII. 11**, tal y como se ha hecho en el caso anterior. En este caso, las diferencias entre las medidas realizadas en el gel original con y sin agitación son menores que al utilizar el agitador grande a 300 rpm. El incremento de la señal al realizar una adición es el mismo para ambas condiciones experimentales. En el caso de agitar la muestra original, la RSD es de un 2% al utilizar el agitador grande y un 0,8% al utilizar el mediano.

En ambos experimentos, puede verse que la mayor variabilidad entre resultados se encuentra siempre en las muestras de gel sin agitación. En cambio, una vez agitadas las muestras durante 20 min, esta variabilidad se estabiliza para todos los experimentos.

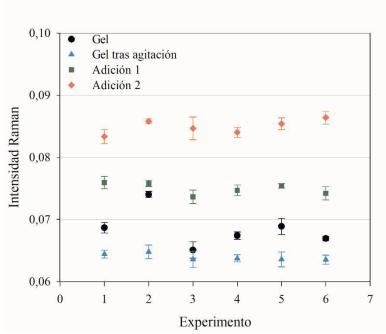


Figura VII. 11. Representación de la intensidad Raman del pico bajo estudio usando el agitador magnético mediano y 500 rpm. Las barras son representadas por la desviación estándar de los 5 espectros registrados.

El orden de magnitud del incremento de la señal Raman tras las adiciones es el mismo para ambos agitadores, con un valor aproximado de 0,01 unidades de incremento. Con el fin de poder determinar qué condiciones usar para la realización de la adición estándar, se ha comparado la desviación estándar relativa (% RSD) para ambas condiciones. Los resultados se presentan en la **Tabla VII. 2**.

Tabla VII. 2. Resultados de RSD (%) obtenidos para las dos condiciones experimentales bajo estudio.

	Agitador mediano	Agitador grande
Gel con agitación	0,83	1,99
Primera adición	1,24	1,03
Segunda adición	1,35	-

Los valores de RSD muestran resultados similares utilizando ambos agitadores, aunque en el caso del gel original (con agitación), la desviación es apreciablemente más baja en las condiciones del agitador mediano. Por ello, se ha decidido utilizar el agitador mediano y una velocidad de 500 rpm para llevar a cabo la adición estándar del patrón en la muestra de gel comercial.

3.5. Resultados de la adición estándar

Se han realizado seis experimentos de adición estándar, utilizando un patrón de IBUNa, en los que se han analizado tres lotes distintos de gel comercial (P016, P019 y P020). La primera etapa ha sido encontrar el valor de la concentración real de IBU, tanto en los geles comerciales como en las disoluciones patrón de IBUNa utilizadas. Para ello, se usó el método HPLC descrito en el Capítulo III (Sección 1.4.1). Los resultados obtenidos mediante el método de referencia (HPLC) se muestran en la **Tabla VII. 3**.

 Tabla VII. 3. Resultados HPLC obtenidos para los tres lotes de gel comercial y los patrones de IBUNa usados.

Lote gel comercial	Resultado HPLC (% IBU m/m)	Patrón IBUNa
P016	$5,0 \pm 0,3$	1
P019	$4,9 \pm 0,2$	2
P020	$4,9 \pm 0,3$	3

Para cada uno de los lotes usados, se ha realizado el experimento por duplicado, obteniendo así el resultado de la concentración de cada lote como el promedio de ambos resultados.

En cada análisis utilizando Raman, se han realizado 5 adiciones estándar, añadiendo 200 μL del patrón en cada una de ellas (aproximadamente 40 mg de IBU). Tal y como se ha comentado anteriormente, antes de la adquisición del espectro Raman, se ha agitado la muestra durante 20 min, utilizando el agitador mediano y a 500rpm, para la medida inicial del gel y para la primera adición; las siguientes adiciones se han realizado después de una agitación de 10 min. Así, el tiempo aproximado necesario para la realización de un experimento completo es de 90 min.

La **Tabla VII. 4** muestra los resultados de la concentración de IBU obtenidos tras la extrapolación de la recta de calibrado, así como el patrón usado, y los parámetros de regresión obtenidos. Como se ha visto anteriormente, la mayor dispersión de resultados se daba al registrar los espectros del gel sin agitación, debido a la deposición de gel en la sonda. Por tanto, para la recta de calibrado, se ha considerado como concentración inicial las muestras de gel tras los primeros 20 min de agitación. Como ejemplo, en la **Figura VII. 12** se muestran los espectros obtenidos para el experimento 3, una vez pretratados según la sección 3.1 y la recta de calibrado obtenida.

Exp.	Lote	Patrón IBUNa	Pendiente	Intersección	R2	% IBU (m/m)
1	P016	1	$2,46\cdot10^{-4}\pm9,66\cdot10^{-6}$	$6,25\cdot10^{-2}\pm1,17\cdot10^{-3}$	0,99	4.9 ± 0.3
2*	P016	3	$2,27 \cdot 10^{-4} \pm 6,30 \cdot 10^{-5}$	$6,22\cdot10^{-2}\pm7,91\cdot10^{-3}$	0,96	$5,4 \pm 1,2$
3	P019	1	$2,52 \cdot 10^{-4} \pm 1,86 \cdot 10^{-5}$	$6,32\cdot10^{-2}\pm2,26\cdot10^{-3}$	0,99	$5,0 \pm 0,5$
4**	P019	3	$2,42\cdot10^{-4}\pm3,69\cdot10^{-5}$	$6,19\cdot10^{-2}\pm3,74\cdot10^{-3}$	0,99	$5,0 \pm 1,0$
5	P020	1	$2,67 \cdot 10^{-4} \pm 1,55 \cdot 10^{-5}$	$6,40\cdot10^{-2}\pm1,89\cdot10^{-3}$	0,99	$4,8 \pm 0,4$
6	P020	2	$2.62 \cdot 10^{-4} \pm 9.95 \cdot 10^{-6}$	$6.53 \cdot 10^{-2} \pm 1.19 \cdot 10^{-3}$	0.99	4.8 ± 0.3

Tabla VII. 4. Resultados de la concentración de IBU tras la adición estándar.

^{**}Eliminando la última adición de la recta por un error experimental

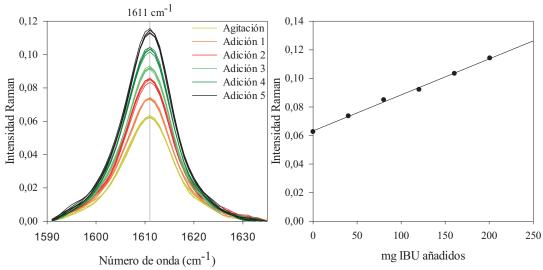


Figura VII. 12. Espectros Raman obtenidos en la adición estándar del experimento 3 tras la aplicación del conjunto de pretratamientos establecidos, y su recta de regresión obtenida.

Los resultados obtenidos son satisfactorios dado que la concentración de IBU es muy próxima a la obtenida con el método de referencia (HPLC).

En la **Tabla VII. 5** se presentan los resultados de la concentración de IBU para los tres lotes, tanto hallada con el método de referencia como por adición estándar.

Tabla VII. 5. Comparativa de los resultados obtenidos por HPLC y por adición estándar (n=2).

Lote	% IBU (m/m) por HPLC	% IBU (m/m) por adición estándar
P016	$5,0 \pm 0,3$	$5,2 \pm 0,8$
P019	$4,9 \pm 0,2$	$5,0\pm0,5$
P020	$4,9 \pm 0,3$	4.8 ± 0.4

Considerando estos resultados, se puede concluir que el método de adición estándar y el uso de una sonda Raman sirven para determinar el contenido de IBU en geles comerciales.

^{*}El contenido de IBU parece algo elevado.

Los resultados obtenidos para los tres lotes se encuentran dentro de la especificación de un 95 – 105% de la concentración nominal de API, que sería el rango de concentración entre 4,8 – 5,3 % m/m, mostrando la posible aplicabilidad de esta técnica en rutina. Sin embargo, aunque la espectroscopia Raman es considerada una técnica rápida y no destructiva, al realizar este experimento, es necesario un tiempo aproximado de 90 min dada la necesidad de realizar cinco adiciones del patrón. Aún y así, la cuantificación del API sería más rápida que si se tuviera que realizar mediante técnicas cromatográficas, si se considera el tiempo de preparación del instrumento y la preparación de las muestras. Mediante HPLC, si se considera la preparación de las muestras, el tiempo de acondicionado del equipo y la columna, y la inyección de las muestras, el tiempo de análisis aproximado sería de 3-4 horas.

A pesar de las pocas muestras utilizadas en la validación, se ha querido determinar si hay diferencias significativas entre los resultados de la concentración de IBU obtenidos mediante HPLC y Raman mediante una prueba t pareada, para un nivel de significación $\alpha = 0$, 05. El resultado muestra que no hay diferencia entre ambas técnicas (p = 0,529), aunque habría que ampliar el número de muestras analizado para que la validación fuera concluyente.

4. CONCLUSIONES

En este trabajo, se ha demostrado que la espectroscopia Raman es una técnica adecuada para el análisis cuantitativo de Ibuprofeno en un gel de alto contenido acuoso y que la utilización de un método de adición estándar permite obtener su concentración a partir de datos univariantes. Este método permite el análisis sin necesidad de disolventes orgánicos.

La alta viscosidad de la muestra de gel ha hecho necesario una buena selección de las condiciones de agitación, que son críticas en la obtención de una señal reproducible.

Se ha conseguido identificar uno de los picos del espectro Raman como una señal característica del API a determinar. También se ha conseguido eliminar el ruido y el efecto de la curvatura en la línea base del espectro Raman utilizando pretratamientos de filtrado y de ajuste de la señal. La dilución del API en la muestra debido a las adiciones consecutivas de disolución patrón se ha solventado normalizando el espectro, dividiéndolo por la señal de una la longitud de onda a la que existe dispersión Raman de uno de los excipientes.

El contenido de IBU en geles comerciales, obtenido por el método de adición estándar y la señal Raman pretratada, se ha comparado con el hallado con el método HPLC de referencia. Se ha comprobado que no hay diferencias significativas entre ambos métodos.

El mayor inconveniente del método es el tiempo necesario para el análisis debido a la necesidad de asegurar la homogeneización de la muestra antes del registro del espectro Raman.

5. BIBLIOGRAFÍA

- [1] C.. Ueda, V.P. Shah, K. Derszinski, G. Ewing, G. Flynn, H.I. Maibach, M.R.C. Marques, H. Rytting, S. Shaw, K.D. Thakker, *Topical and Transdermal Drug Products*, Pharmacopoeial Forum. 35 (2009) 750–764.
- [2] K. Wu, T. Yeoh, Y.-L. Hsieh, D.W. Osborne, Quality Assessment of API in Semisolid Topical Drug Products BT The Role of Microstructure in Topical Drug Product Development, in: N. Langley, B. Michniak-Kohn, D.W. Osborne (Eds.), Springer International Publishing, Cham, **2019**: pp. 109–154. https://doi.org/10.1007/978-3-030-17355-5_4.
- [3] M.S. Kemper, E.J. Magnuson, S.R. Lowry, W.J. McCarthy, N. Aksornkoae, D.C. Watts, J.R. Johnson, A.J. Shukla, *Use of FT-NIR transmission spectroscopy for the quantitative analysis of an active ingredient in a translucent pharmaceutical topical gel formulation*, AAPS PharmSci. 3 (2001) E23–E23. https://doi.org/10.1208/ps030323.
- [4] M. Blanco, M.A. Romero, *Near infrared transflectance spectroscopy:*Determination of dexketoprofen in a hydrogel, J. Pharm. Biomed. Anal. 30 (2002) 467–472. https://doi.org/10.1016/S0731-7085(02)00093-6.
- [5] M. Blanco, M. Alcalá, M. Bautista, Pharmaceutical gel analysis by NIR spectroscopy: Determination of the active principle and low concentration of preservatives, Eur. J. Pharm. Sci. 33 (2008) 409–414. https://doi.org/10.1016/j.ejps.2008.01.011.
- [6] S.C. Baratieri, J.M. Barbosa, M.P. Freitas, J.A. Martins, *Multivariate analysis of nystatin and metronidazole in a semi-solid matrix by means of diffuse reflectance NIR spectroscopy and PLS regression*, J. Pharm. Biomed. Anal. 40 (2006) 51–55.

- https://doi.org/10.1016/j.jpba.2005.05.025.
- [7] L.B. Schlegel, M. Schubert-Zsilavecz, M. Abdel-Tawab, *Quantification of active ingredients in semi-solid pharmaceutical formulations by near infrared spectroscopy*, J. Pharm. Biomed. Anal. 142 (**2017**) 178–189. https://doi.org/10.1016/j.jpba.2017.04.048.
- [8] J. Peerapattana, Y. Hattori, M. Otsuka, Simultaneous quantitative analysis of indomethacin and benzoic acid in gel using ultra-violet-visible spectrophotometry and chemometrics, Biomed. Mater. Eng. 30 (2019) 73–84. https://doi.org/10.3233/BME-181034.
- [9] É. Szabó, S. Gergely, T. Spaits, T. Simon, A. Salgó, Near-infrared spectroscopy-based methods for quantitative determination of active pharmaceutical ingredient in transdermal gel formulations, Spectrosc. Lett. 52 (2019) 599–611. https://doi.org/10.1080/00387010.2019.1681459.
- [10] T.R. De Beer, W.R. Baeyens, Y. V Heyden, J.P. Remon, C. Vervaet, F. Verpoort, Influence of particle size on the quantitative determination of salicylic acid in a pharmaceutical ointment using FT-Raman spectroscopy, Eur J Pharm Sci. 30 (2007) 229–235. https://doi.org/10.1016/j.ejps.2006.11.009.
- [11] S. Mazurek, R. Szostak, *Quantitative analysis of topical gels and ointments by FT-Raman spectroscopy*, Vib. Spectrosc. 83 (**2016**) 1–7. https://doi.org/10.1016/j.vibspec.2015.12.005.
- [12] C. Masucci, Evaluation of a portable Raman spectrometer as an analytical tool in the pharmaceutical industry, Università di Roma La Sapienza, **2020**.
- [13] I. Avilés, Avaluació d'un Raman Portàtil per al seu ús amb preparacions farmacèutiques, Universitat Autònoma de Barcelona, **2019**.
- [14] D.A. Prystupa, A. Anderson, B.H. Torrie, *Raman and infrared study of solid benzyl alcohol*, J. Raman Spectrosc. 25 (**1994**) 175–182. https://doi.org/10.1002/jrs.1250250206.
- [15] S.A. Weissman, N.G. Anderson, *Design of Experiments (DoE) and Process Optimization. A Review of Recent Publications*, Org. Process Res. Dev. 19 (2015)

DETERMINACIÓN DEL CONTENIDO DE IBUPROFENO EN GELES COMERCIALES

- 1605-1633. https://doi.org/10.1021/op500169m.
- [16] B. Ait-Amir, P. Pougnet, A. El Hami, 6 Meta-Model Development, in: A. El Hami,
 P.B.T.-E.M.S. 2 (Second E. Pougnet (Eds.), ISTE, 2020: pp. 157–187.
 https://doi.org/10.1016/B978-1-78548-190-1.50006-2.
- [17] ICH, ICH Q2 (R2) Validation of Analytical Procedures, (2022).

VIII. <u>ESTUDIO DE LA DISTRIBUCIÓN DEL</u> <u>CONTENIDO DE API EN COMPRIMIDOS</u> <u>UTILIZANDO ESPECTROSCOPIA RAMAN</u>

1. INTRODUCCIÓN

Uno de los parámetros críticos en cualquier especialidad farmacéutica es la uniformidad de contenido (CU) [1]. En el caso de los comprimidos, la etapa de mezclado es un parámetro crítico de la calidad del proceso de fabricación, por lo que es muy importante poder monitorizar el proceso de homogeneización de la mezcla antes de realizar la compactación. Sin embargo, existe un riesgo a segregación durante el tránsito de la mezcla del mezclador a la compactadora, por lo que, generalmente, la CU se evalúa tras la compactación de los comprimidos. La CU es un parámetro crítico para todos los comprimidos, pero especialmente para los que pueden dividirse para su dosificación.

La división de los comprimidos se ha convertido en una práctica frecuente, especialmente para la administración en adultos y niños de una dosis específica. Esta práctica permite modificar la dosis, reducir el coste económico y facilitar su ingesta [2–5]. Sin embargo, hay estudios que muestran una mayor variabilidad en la dosis al tomar comprimidos divididos que enteros [6–9]. Esta variabilidad puede estar relacionada con la variación de masa en las mitades, la falta de homogeneidad del API o la menor estabilidad del producto una vez dividido. Además, existen diferencias según cómo se realiza la división, de forma manual o mecánica [10]. Considerando esta variabilidad, la distribución del API en todo el comprimido debería tratarse como un parámetro crítico en la calidad del producto.

En relación con la división de los comprimidos, la FDA publicó en marzo de 2013 una guía donde presentaba recomendaciones para evaluar estos productos [11]. La guía concluye que se debe cumplir con los mismos requerimientos que los productos acabados, como, por ejemplo, la prueba de uniformidad de contenido de dosis unitarias (UDU por sus siglas Uniformity Dosages Units). La **Tabla VIII. 1** muestra un resumen de los principales criterios para evaluar el contenido y variación de masa en fracciones de un comprimido [12–14].

Por otro lado, el primer documento donde se introdujo una guía para comprobar la uniformidad de contenido se publicó en 2003 por la FDA. En esta guía, se describe la prueba de UDU, que posteriormente fue harmonizado por la USP y EP. En 2011, la EP realizó una modificación en los parámetros de la prueba para poder analizar un mayor número de muestras. Tras esta propuesta, lo que se proponía era analizar 30 muestras y cuantificar como mínimo 10 de ellas. El objetivo principal de esta prueba es calcular un valor de aceptación

(AV), el cual considera una constante de aceptabilidad (k) que depende del número de muestras (n). Este nuevo capítulo incluye valores de k para un número de muestras de hasta n=10000. A pesar de estas modificaciones, el proyecto de la uniformidad de contenido se apartó en 2013 por no encajar con la mentalidad de la agencia. Hasta 2015 no se incluyó una nueva sugerencia para la realización de esta prueba. Esta sugerencia está basada en la realización del cálculo de la AV para 10 comprimidos y, en caso de obtener un resultado fuera de límites (AV < 15), se debe ampliar la población hasta n=30 y recalcular el AV.

Tabla VIII. 1. Resumen de los criterios de la FDA para evaluar el contenido y la variación de peso de las fracciones de un comprimido divisible.

Formas de dosificación oral sólida de liberación modificada (con tecnología de matriz y con componentes recubiertos de película comprimida) USP General Chapter <905> Uniformity of Dosage Units.

- USP General Chapter <905> Uniformity of Dosage Units.
 Realización de la prueba de uniformidad de contenido (UC) o variación de masa (MV) dependiendo del contenido de API. La prueba de MV se puede aplicar en comprimidos divisibles si contienen 25 mg o más del producto que representan un 25% o más (de peso) de la fracción. Si no, se debe realizar el UC.
- USP General Chapter <1216> Tablet
 Friability.
 La pérdida de peso tras la evaluación de 15
 comprimidos debe ser menor al 3% entre las
 fracciones al compararlo con el comprimido
 entero.
- La prueba de disolución debe ser el mismo que el del comprimido entero.

- Todos los requisitos anteriores
- FDA Guideline: Dissolution Testing of Immediate Release Solid Oral Dosage Forms.
 El factor de similitud (f₂) debe cumplir con el criterio establecido (f₂ cercano a 100)
- FDA Guideline: Dissolution Testing of Immediate Release Solid Oral Dosage Forms.
 Los perfiles de disolución de los gránulos deben cumplir con el criterio de similitud de (f₂) durante la compresión.

La UDU se puede demostrar por dos métodos distintos, la variación de masa (MV) o la uniformidad de contenido (UC). El método de MV se puede aplicar en cápsulas, comprimidos sin recubrimiento o recubiertos, que contengan 25 mg o más del API y que represente un 15% o más del peso de la dosis unitaria. El test de MV se realiza para una muestra representativa del lote usando un método analítico adecuado. La UC se basa en la realización del ensayo en cada dosis individual para determinar si ésta se encuentra dentro de los límites establecidos y es la prueba estudiada en este trabajo por las características de los comprimidos.

La determinación del contenido del API en productos farmacéuticos suele realizarse mediante procesos offline, como sería el análisis por HPLC [15–17]. La principal desventaja de estos métodos es su elevado coste, el tiempo necesario para realizar el análisis, la destrucción de las muestras analizadas, y el uso de disolventes, lo que genera una gran cantidad de residuos. Los métodos espectroscópicos pueden ser una alternativa rápida para el estudio de la distribución del API frente a los métodos tradicionales, como puede ser HPLC. El análisis hiperespectral de la superficie de todo el comprimido parece ser una de las mejores soluciones. La imagen hiperespectral proporciona información química y espacial de los componentes del producto estudiado, por lo que permite visualizar la distribución espacial de todos ellos y calcular así su distribución de homogeneidad [18–20]. Sin embargo, este análisis requiere instrumentos muy específicos de elevado coste y de un tratamiento de datos complejo [21–23]. Una alternativa a la imagen hiperespectral es la realización de un mapeo apropiado de puntos en la superficie del comprimido. Este mapeo permite obtener una imagen química de la distribución del API usando instrumentación analítica más convencional y con un tratamiento de los datos más sencillo y rápido. La espectroscopia Raman es una técnica muy útil para este propósito [24,25].

El diseño de una cuadricula para la realización del mapeo de puntos permite obtener una distribución del API en el comprimido y calcular la UC tanto en comprimidos enteros como en mitades. Normalmente, la estimación de la distribución del API se realiza mediante el cálculo de la concentración del API mediante una regresión utilizando técnicas multivariables. La **Tabla VIII. 2** muestra un resumen de este tipo de aplicaciones en la literatura.

La espectroscopia Raman se ha usado y descrito para la determinación de la UC. Sin embargo, como el diámetro del láser de este tipo de instrumentos suele ser menor que el tamaño de partícula del API, obtener una distribución aproximada del API en los comprimidos requiere de un gran número de puntos de muestreo. Sin embargo, un número de puntos de muestreo mayor implica un tiempo de análisis más elevado al tener que adquirir un número de espectros Raman mayor, por lo que la espectroscopia Raman no es muy útil para un análisis rutinario. El diámetro del láser y el número de puntos de muestreo es crítico para obtener una buena distribución del API en la superficie del comprimido. Dado que el primer parámetro viene determinado por el instrumento, el único parámetro que se puede optimizar es el número y distribución de los puntos de mapeo a realizar en la superficie. Otro

factor importante es la extracción de la información cuantitativa, como la concentración del API, del espectro Raman obtenido. Esta información se suele obtener mediante un proceso de regresión multivariable (por ejemplo, PLSR). Una de las mayores desventajas de este proceso es la necesidad de preparar un conjunto de muestras de calibración y de la actualización y mantenimiento del modelo durante el ciclo de vida del producto.

La espectroscopia Raman permite la realización de medidas de forma rápida y no destructiva, sin la necesidad de realizar una preparación previa de la muestra y proporciona información tanto química como física de la misma[26]. Una de sus ventajas es que los espectros Raman están formados por picos estrechos bien definidos; además, la mayoría de los APIs contienen grupos aromáticos, que proporcionan picos con gran intensidad en el espectro Raman y que no se superponen a los de los excipientes, que normalmente son compuestos alifáticos. De esta forma, existe una correlación directa entre la intensidad del pico Raman y la concentración del API. Dado que el objetivo principal de realizar la UC es comparar el contenido del API entre distintas medidas en las formas farmacéuticas sólidas, se puede utilizar directamente la intensidad del pico sin tener que desarrollar y validar un método de calibración.

El objetivo principal de este trabajo es demostrar la capacidad de la técnica Raman tras la optimización del número de puntos de mapeo de la superficie del comprimido, para realizar el cálculo de la UC. Como el producto farmacéutico usado en este trabajo son comprimidos que contienen una línea ranurada, la evaluación se realiza en los comprimidos enteros y en sus mitades. El cálculo de la UC se ha realizado mediante la intensidad de un pico aromático asociado al API. Como objetivo secundario, se propone el uso de la espectroscopia Raman para verificar la uniformidad de la distribución del API en todo el comprimido y un procedimiento para estimar cuantitativamente la homogeneidad de la distribución.

ESTUDIO DE LA DISTRIBUCIÓN DEL CONTENIDO DE API EN COMPRIMIDOS UTILIZANDO ESPECTROSCOPIA RAMAN

Tabla VIII. 2. Resumen de aplicaciones de un mapeo Raman para productos farmacéuticos [19,20,35–41,27–34].

Problema	Tipo de muestra	Instrumento	Área muestrada	Tratamiento de los datos	Ref
Monitorizar la concentración y distribución del API en micro comprimidos	Micro-comprimidos (2 mm diámetro)	Imagen hiperespectral	Toda la superficie 0,1mm/scan	PLS y LS-SVM para cuantificación Intensidad de pico para cualificación	22
Localización y cuantificación de los componentes químicos de un Sistema multicomponente.	Comprimidos	Microscopio Raman	150 x 150 μm	PCA MCR-ALS PLS	23
Evaluar el área de mapeo del FT- Raman para determinar los componentes presentes en un comprimido	Comprimidos	FT-Raman mapeo	1000 x 1000 μm	MCR	16
Comparar los resultados obtenidos usando distintos métodos de resolución multivariable de curvas	Comprimidos	Microscopio Raman	-	CLS MCR-ALS Factorización positiva de matriz Sistema lineal de identificación Simplex	24
Determinar la formación de complejos de inclusión, distribución espacial de los ingredientes y cuantificar las trazas residuales	Formulación de API y ciclodextrinas	Microscopio Raman	0,8 mm x 0,8 mm 1,0 mm x 1,0 mm	CLS	25
Estudiar la fracción de una sustancia amorfa en una dispersión sólida	Dispersiones sólidas (API y excipientes)	Convencional	Radio iluminado 3 mm	Intensidad de pico	26
Caracterizar los gránulos durante el proceso de granulación de un producto farmacéutico	Gránulos secos	Microscopio Raman	3,5 x 3,5 mm	Intensidad de pico	27
Evaluación de cuatro técnicas quimiométricas para la cuantificación de un producto a bajo concentración en mezclas sólidas	Mezclas sólidas	Microscopio Raman	29 x 29 puntos 8410 espectros/mapa Score maps using Raman signals		28
Distribución de los componentes en la superficie y en el núcleo	Formas farmacéuticas sólidas	Convencional	Promedio de tres registros	Área de pico	29

Problema	Tipo de muestra	Instrumento	Área muestrada	Tratamiento de los datos	Ref
Distribución especial de las trazas de un API usando SERS	Comprimidos	Microscopio Raman	-	MCR-ALS	30
Distribución especial del API y excipientes en comprimidos de seis distribuidores	Comprimidos	Microscopio Raman	175 x 175 μm (mapeo confocal)	Intensidad y área de pico	31
Evaluación de la homogeneidad de las muestras producidas mediante distintos procesos de fabricación	Comprimidos	Microscopio Raman	-	Análisis de macropíxeles utilizando el método de mosaico de nivel discreto, con mosaicos que no se superponen, o el método de bloque móvil de nivel continuo.	15
Metodología para la evaluación de la distribución espacial de la homogeneidad	Comprimidos	Espectrómetro Raman dispersivo 16 mm² CLS PCA			32
Detección de aglomerados y distribución especial del API	Comprimidos	Espectrómetro Micro-Raman	10 x 10 μm	PLS-DA	33
Uniformidad de contenido y evaluación de distintas formas cristalinas del API	Comprimidos	Raman de baja frecuencia	25100 espectros	PCA	34
Investigación de la composición de los comprimidos	Comprimidos	Espectrómetro Raman acoplado a un microscopio	6000 espectros	MCR-ALS	35
Análisis de la falsificación de productos farmacéuticos	Comprimidos	Espectrómetro Raman acoplado a un microscopio	3000 x 3000 μm ²	MCR-ALS CLS	36

^aArea muestreada y/o número de espectros adquiridos, si consta.

^bConsiderando solo la cualificación (distribución, detección) de la muestra

2. MATERIALES Y MÉTODOS

2.1. Muestras

Las muestras analizadas han sido adquiridas en una farmacia. En total, se han usado cinco lotes distintos, denominados A, B, C, D y E, de comprimidos de Lisinopril, comercializados por Mylan Pharmaceuticals S.L. Cada lote contiene dos blísteres con 14 comprimidos y cada comprimido contiene 20 mg del API (Lisinopril, LIS), que representa, aproximadamente, un 10% m/m de su peso. Los excipientes presentes en la formulación son dihidrogenofosfato de calcio dihidrato, manitol, almidón de maíz pregelatinizado, povidona, croscarmelosa de sodio, estearato de magnesio /lauril sulfato de sodio (94/6), óxido de hierro rojo y dióxido de silicio. La cara superior de los comprimidos contiene una ranura para facilitar la división, lo que sugiere la posibilidad de dividirlo en dos mitades para su administración, y la cara posterior tiene grabado el número 20.

El lote E se ha usado para obtener las condiciones experimentales de la adquisición de los espectros Raman.

Las partículas del API puro tienen una forma rectangular, como prismas de aproximadamente 20 µm de longitud y 4 µm de amplitud. La **Figura VIII. 1Figura VIII.**2. muestra la imagen SEM obtenida a partir de la cual se han obtenido estos valores.

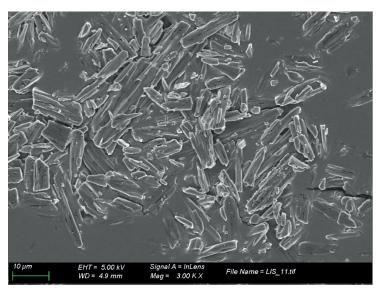


Figura VIII. 1. Imagen microscópica de las partículas del API (Lisinopril).

El producto farmacéutico estudiado, LIS, es una angiotensina usada para tratar la hipertensión, ataques de corazón, y para prevenir complicaciones de la diabetes. Algunos estudios publicados han evaluado el efecto y la variación en peso obtenido entre la dosificación de los comprimidos enteros o por mitades.[42,43]

2.2. Instrumento y software

Los espectros Raman se han adquirido mediante el espectrómetro MultiRaman de Bruker (Bruker, Billerica, MA, USA) presentado y descrito en el Capítulo III. El diámetro del láser de este instrumento es de 500 µm. El programa usado para controlar el equipo y extraer los datos ha sido Opus 6.5.

El procesado de los datos se ha realizado mediante el software The Unscrambler 10.5 (Trondheim, Norway), y los cálculos matemáticos mediante MATLAB R2014b (Natick, Massachusetts, USA).

2.3. Adquisición de los espectros

Los comprimidos de LIS se colocaron en una platina que se puede controlar visualmente mediante el software. Para cada una de las mitades de la cara superior se adquirieron 8 espectros Raman siguiendo un esquema prediseñado (**Figura VIII. 2**). En total, se obtuvieron 16 espectros para cada comprimido, que supone un tiempo de análisis aproximado de 12 min por muestra. Además, se adquirieron 7 espectros en la cara inferior de los comprimidos y 8 en el núcleo, siguiendo el diagrama de la figura. La división de los comprimidos se ha realizado mediante un cortador de pastillas y la superficie del núcleo se ha rascado con un cúter para que fuera lo más regular y plana posible.

Para evitar posibles efectos sistemáticos no controlados, los espectros se registraron aleatoriamente y en días diferentes. También se calculó la reproducibilidad tratando las medidas realizadas en el mismo comprimido durante 10 días.

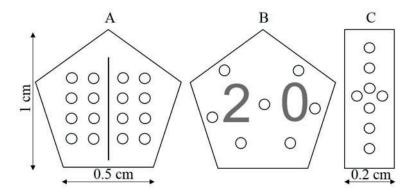


Figura VIII. 2. Dimensiones y esquema donde se han adquirido los espectros Raman individuales para (A) Cara superior; (B) Cara posterior y (C) Núcleo de los comprimidos.

3. RESULTADOS Y DISCUSIÓN

El primer paso para el desarrollo de este trabajo ha sido encontrar las condiciones óptimas de registro de los espectros mediante un diseño de experimentos (DoE). A partir de aquí, se ha optimizado el pretratamiento espectral y se ha evaluado la homogeneidad del API en los comprimidos.

Una de las características de realizar un mapeo punto a punto por Raman es la observación y aparición de diferencias en la intensidad de los picos en los distintos puntos estudiados. Estas diferencias se pueden atribuir a tres factores: el ruido inherente del instrumento usado, las diferencias en el número de partículas del API en la zona irradiada (error de submuestreo) y a la dispersión, que incluye cambios en la potencia del láser, enfoque, radiación dispersada, etc. La varianza esperada (s²) durante la realización del mapeo se presenta en la **Ecuación VIII-1**, donde s_a^2 representa la variabilidad analítica, que depende de las condiciones establecidas para el instrumento, s_m^2 es la variabilidad producida por diferencias en el muestreo y s_s^2 la relacionada con los factores del efecto de la dispersión. El último factor, la varianza s_s^2 , se puede considerar prácticamente no significativo si el espectro se normaliza, obteniendo así una simplificación en la ecuación (Ecuación VIII-2) [44].

$$s^2 = s_a^2 + s_m^2 + s_s^2$$
 (VIII-1)
 $s^2 = s_a^2 + s_m^2$ (VIII-2)

$$s^2 = s_a^2 + s_m^2 \tag{VIII-2}$$

3.1. Resultados del DoE

Las condiciones óptimas de adquisición de los espectros se han obtenido mediante un diseño factorial completo a tres niveles (3³) usando como variables la potencia del láser (400, 600 y 800 mW), la resolución (4, 8 y 16 cm⁻¹) y el número de barridos (32, 64 y 128). Éstas tres variables son las que se pueden modificar mediante el software. Las respuestas del diseño han sido el tiempo de adquisición de los espectros (un espectro individual) y la relación señal ruido (S/N) de cada uno de los experimentos realizados. El objetivo de este diseño es encontrar las condiciones de registro que impliquen el mejor compromiso entre un tiempo de análisis bajo y una buena relación S/N.

La **Figura VIII. 3** muestra los espectros Raman de los lotes estudiados sin procesar y el espectro del principio activo, con un pico a 1603 cm⁻¹ que se atribuye a la extensión (*stretching*) del doble enlace C=C del anillo aromático presente en la estructura química del LIS. Este C=C aromático no lo contiene ningún otro producto presente en la formulación, por lo que los excipientes no interfieren en la señal analítica seleccionada. Se ha marcado este pico en los espectros del producto acabado. Entre los lotes estudiados, se ha encontrado una diferencia en la región de 900 – 570 cm⁻¹. Esta diferencia no se ha estudiado dado que probablemente se deba a un cambio de proveedor o cambio de algún excipiente.

Una vez seleccionado el pico a estudiar, se puede determinar la relación S/N para cada uno de los experimentos a efectuar dividiendo la intensidad del pico característico por la desviación estándar del ruido espectral. En este caso, el ruido espectral se ha estimado usando el rango de 2104 a 1718 cm⁻¹. Con todo esto, se han realizado un total de 27 experimentos.

Los resultados del diseño se muestran en la **Figura VIII. 4** representando el tiempo de adquisición *vs* la relación S/N. De los resultados obtenidos, se han marcado los que proporcionan una mayor respuesta en la representación. Entre ellos, el experimento C3 se ha descartado por su necesidad de un tiempo de análisis mayor, que viene determinado por el número de barridos usados (128 barridos). Respecto a los experimentos C1 (800 mW, 128 barridos y 16 cm⁻¹) y C2 (800 mW, 64 barridos y 16 cm⁻¹), ambos proporcionan una buena respuesta con una ligera diferencia en el tiempo de adquisición por la diferencia en el número de barridos. Finalmente, se han seleccionado las condiciones del experimento C2 por esta pequeña reducción en el tiempo de análisis (de 1 a 0,7 min).

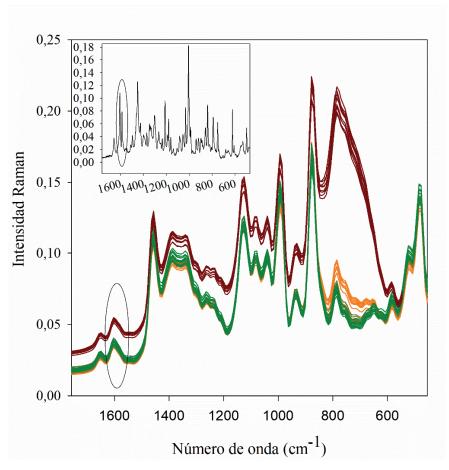


Figura VIII. 3. Espectros Raman de todas las muestras y del API en el rango espectral de 1757 a 453 cm⁻¹. La elipse marca el pico característico del API (1603 cm⁻¹). La resolución de los espectros de la muestra es de 16 cm⁻¹, mientras que la del API 2 cm⁻¹. Rojo oscuro: lote A;

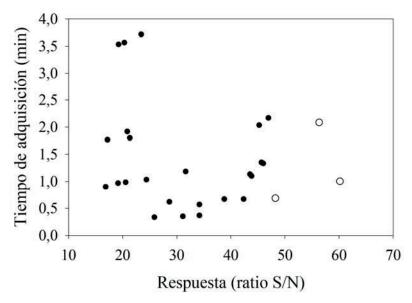


Figura VIII. 4. Representación del tiempo de adquisición vs la relación S/N para los 27 experimentos del DoE. Los círculos (C1, C2 y C3) corresponden a los resultadas de las siguientes condiciones experimentales: C1: 800 mW, 128 scans and 16 cm⁻¹; C2: 800 mW, 64 scans and 16 cm⁻¹; C3: 800 mW, 128 scans and 8 cm⁻¹.

3.2. Pretratamiento espectral

El pico analítico estudiado (1603 cm⁻¹) se ha analizado usando dos pretratamientos distintos. El primero se basa solo en la aplicación de una corrección de la línea base mediante la corrección *Linear Baseline Correction* (BLC) en el intervalo de 1695 a 1562 cm⁻¹. El segundo pretratamiento consta de dos etapas; primero una corrección de la línea base mediante una BLC en el rango de 1703 a 962 cm⁻¹ y luego una normalización a área unidad de 1703 a 546 cm⁻¹ (BLC-N). La varianza esperada para estos dos pretratamientos viene definida por la Ecuación IX-1 y Ecuación IX-2, respectivamente.

La precisión, expresada como desviación estándar relativa de los diez espectros registrados en días consecutivos, muestra que, tras la aplicación de una normalización, la varianza total se reduce prácticamente la mitad. Al aplicar solo BLC, la variabilidad de la intensidad del pico analítico es de 1,96% mientras que al normalizar se reduce hasta un 0,83%.

3.3. Distribución del API

En el análisis de la uniformidad de contenido de distintos comprimidos mediante Raman, se pueden esperar dos fuentes de variabilidad: la relacionada con el proceso de fabricación (intrínseca) y la relacionada con el error de submuestreo. La varianza relacionada con el error del submuestreo en un sólido sigue una distribución binomial.[45] Sin embargo, considerando el gran número de partículas muestreadas (N) y la proporción del API en el comprimido (aprox 10% m/m), se cumple la condición de que el número de partículas, N, por la probabilidad de muestreo, p, es mayor que 5 y se puede aproximar que el error de submuestreo sigue una distribución normal [46].

La **Figura VIII. 5** muestra el promedio de la intensidad con la desviación estándar (n=10) de los comprimidos de todos los lotes para ambos pretratamientos estudiados. En esta representación, todas las intensidades se encuentran en el rango de 0,013-0,016 cm⁻¹ al aplicar BLC y en 0,0047-0,0050 cm⁻¹ con BLC-N. Sin embargo, si se consideran los valores de intensidad Raman individuales, la variabilidad observada es mayor, como se puede observar en las barras de error presentadas mediante el cálculo de la desviación estándar.

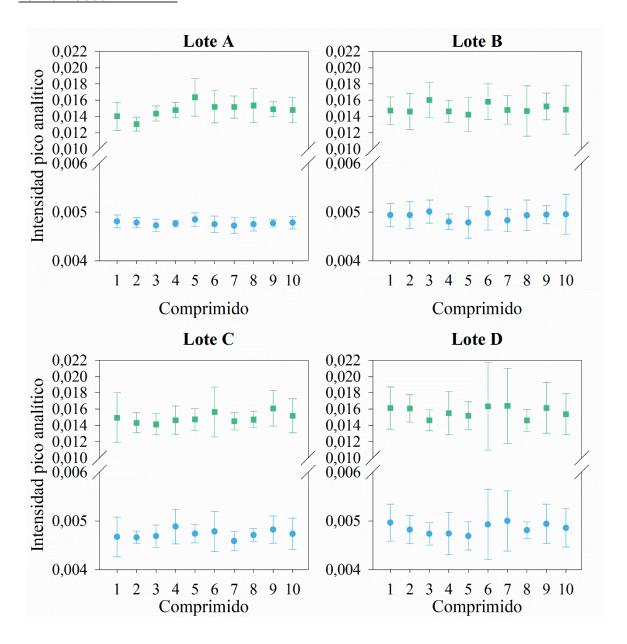


Figura VIII. 5. Intensidad promedio (n=16) del pico característico de LIS (1603 cm⁻¹) para los diez comprimidos analizados de cada lote. En verde se muestra la intensidad promedio tras la corrección BLC, y en azul BLC-N.

En la **Figura VIII.** 6 se presenta el histograma de probabilidad de las intensidades del pico analítico estudiado para todos los comprimidos analizados después de la aplicación de ambos pretratamientos. El histograma muestra que la mayoría de los resultados se encuentran agrupados alrededor de un valor central con poca dispersión. Sin embargo, se observa en que hay resultados en los extremos, lo que resulta en una menor o mayor intensidad del pico analítico. Este patrón se repite en todos los lotes estudiados y con ambos

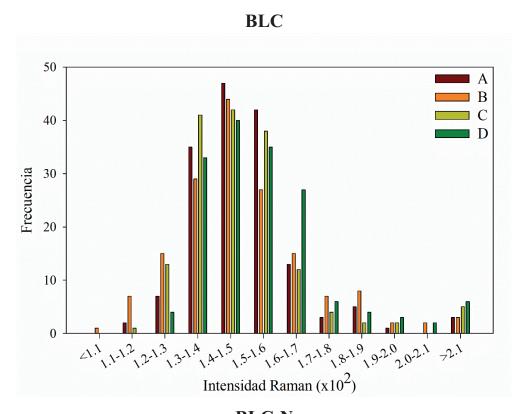
pretratamientos, lo que sugiere una tendencia sistemática del API a formar aglomerados o la presencia del API en distintos tamaños de partícula [47,48].

A partir de los resultados presentados en el histograma, se calculó un índice de agregación del API para cada uno de los lotes, siguiendo la **Ecuación VIII-3**. Considerando los valores promedio y la desviación estándar obtenida para cada lote, se estimó lo que sería una distribución normal ideal. Para los lotes A, B y C el intervalo de la distribución normal es de 0,011-0,017, y para el lote D de 0,012-0,018, en el caso de corregir solamente la línea base. En el caso de normalizar también los espectros, el rango de distribución sería de 0,0042-0,0052 para el lote A, 0,0042-0,0058 para los lotes B y D, y 0,0042-0,0056 para el C. A partir de aquí, se ha calculado el índice de agregación para poder realizar una estimación de la fracción del API que está formando agregados en el comprimido, por lo que la intensidad Raman en algunos puntos analizados es mayor al promedio.

$$Indice = 1 - \frac{Puntos \ distribución \ normal - Puntos \ no \ distribución \ normal}{Total \ de \ puntos} \tag{VIII-3}$$

Los resultados obtenidos para el índice de agregación han sido de un 7,5% del contenido promedio total del API con la aplicación de la BLC, y de un 2,5% con BLC-N. Aunque el índice promedio obtenido es bajo, puede suponer un problema durante la estimación de la uniformidad de contenido, tanto para comprimidos enteros como para en sus mitades. Como puede verse en los resultados, al normalizar los espectros, se corrige bastante el efecto de la segregación.

Aparte de la estimación de la segregación, se realizó un ANOVA de una vía con los datos de la cara superior para determinar, si en un mismo lote, todos los comprimidos analizados son idénticos. Al corregir los espectros aplicando solo la BLC, los resultados muestran que no hay diferencias significativas entre los comprimidos de los lotes B, C y D. En cambio, para el lote A, hay diferencias en uno de los comprimidos analizados que pueden estar relacionadas con la menor intensidad del pico característico de este comprimido en comparación a los demás. Por otro lado, al normalizar los espectros, no hay diferencias significativas entre los comprimidos para ninguno de los lotes.



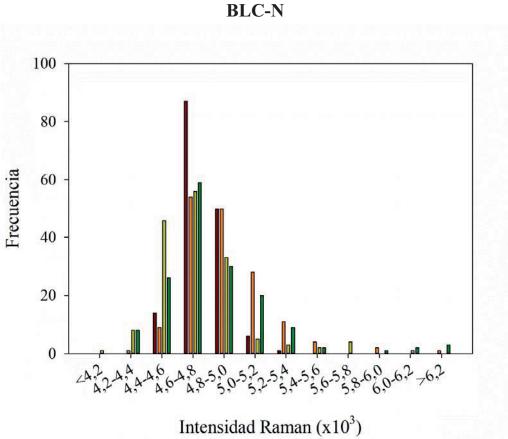


Figura VIII. 6. Histograma de la probabilidad de las intensidades del pico analítico para todas las medidas después de la aplicación de ambos pretratamientos estudiados. La leyenda de colores es la misma para ambas figuras.

3.3.1. Estudio en la cara posterior y núcleo

La distribución del API en la cara posterior y el núcleo se ha evaluado comparando el valor promedio de la intensidad del pico analítico y su desviación estándar. La **Figura VIII.** 7 muestra la comparación de los resultados obtenidos. Como puede verse, sin la aplicación de una normalización existe una mayor dispersión en los resultados obtenidos para el núcleo. Sin embargo, una vez normalizados los espectros, no se ven diferencias significativas entre los valores medios de las caras.

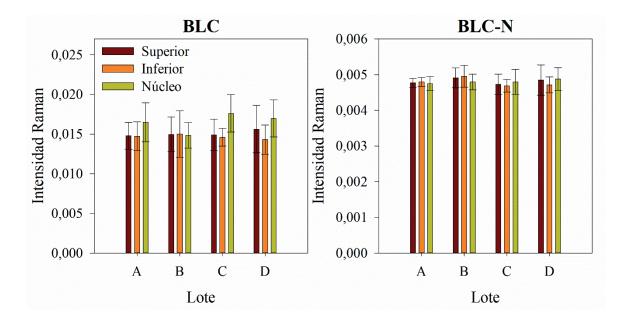


Figura VIII. 7. Valor promedio y desviación estándar de la intensidad Raman del pico analítico para todos los comprimidos y caras después de la aplicación de BLC o BLC-N.

Para poder evaluar estadísticamente si existen diferencias significativas entre las tres caras estudiadas, se ha realizado un ANOVA de una vía. Los resultados han mostrado que para tres de los cuatro lotes estudiados al realizar solo una BLC hay diferencias significativas entre las tres caras. En cambio, tras la normalización de los espectros, ninguno de los lotes presenta diferencias significativas. De las tres caras, las diferencias siempre vienen dadas con la intensidad del pico analítico obtenidas para el núcleo, dado que la cara superior e inferior no presentan diferencias significativas tras la realización de una prueba t. Esta diferencia obtenida entre caras está relacionada con la superficie de los comprimidos. En el caso de la cara superior y posterior, la superficie es totalmente lisa mientras que, al dividir los espectros, la superficie del núcleo puede no estar del todo pulida. La diferencia en la

superficie de los comprimidos provoca diferencias en la dispersión resultando en valores de intensidades distintos. Al aplicar la normalización este efecto se reduce obteniendo intensidades parecidas para las tres caras estudiadas. Este cambio en la intensidad promedio también puede apreciarse en la **Figura VIII. 1.Figura VIII. 7**.

3.3.2. Uniformidad de las unidades de dosificación

Una vez se ha verificado que la intensidad del pico analítico de los 16 espectros Raman individuales obtenidos para la cara superior es representativa de la distribución del API en la superficie, se ha determinado la uniformidad de contenido siguiendo el criterio de la EP, que calcula un valor de aceptación (AV), para las primeras 10 dosis a evaluar, utilizando la **Ecuación VIII-4**. Si AV < 15, se supera la prueba; en caso contrario es necesario utilizar las siguientes 20 dosis y recalcular AV.

$$AV = |M - \bar{X}| + k \cdot s \tag{VIII-4}$$

Donde AV es el valor de aceptación, \overline{X} el promedio de los valores individuales (x₁, x₂...x_n) expresados como el porcentaje del valor nominal, M el valor de referencia, s la desviación estándar y k la constante de aceptabilidad, que depende del número de muestras (n). En este caso, k vale 2,4. Considerando el objetivo de este trabajo, se ha asumido que las diferencias entre el valor de referencia y el promedio de los individuales es menor a 1,5%. De esta forma, se puede simplificar el criterio para determinar AV siguiendo la **Ecuación VIII-5**, tal y como establece la farmacopea.

$$AV = k \cdot s \tag{VIII-5}$$

En este caso, en vez de usar la concentración de API en la muestra, se ha utilizado el promedio de la intensidad Raman del pico analítico, que es proporcional a la concentración. Se ha usado la intensidad del pico de diez comprimidos de cada lote para calcular el valor de aceptación (AV) siguiendo la Ecuación IX-5. En este caso, como se ha comentado anteriormente, los mejores resultados de distribución del API se encuentras al normalizar los espectros. Por ello, la uniformidad de contenido solo se evaluará con este pretratamiento (BLC-N).

Los resultados de los AV obtenidos se muestran en la ¡Error! No se encuentra el origen de la referencia. Como puede verse, en todos los casos el valor obtenido se encuentra dentro de los límites establecidos por la EP ($AV \le 15$).

Tabla VIII. 3. Resultados de AV usando el promedio de la intensidad del pico analítico de 16 o 12 espectros Raman individuales. k = 2,4.

Lote -	BLC-N			
	16 espectros	12 espectros		
A	1,90	2,16		
В	3,75	3,43		
C	4,34	4,32		
D	5,31	5,64		

3.3.3. Estudio del número mínimo de espectros en unidades individuales

Los resultados más fiables se obtienen tras la realización de un mapeo de espectros individuales que abarque la mayor superficie del comprimido, lo que lleva a la adquisición de muchos espectros Raman individuales para cada una de las muestras. El principal inconveniente de usar esta estrategia es que implica un tiempo de análisis muy elevado para poder realizarse en rutina, aunque este incremento del tiempo no sería mayor al de las técnicas rutinarias usadas para este propósito, como puede ser la cromatografía liquida.

Se ha realizado un estudio para determinar el número de puntos de mapeo necesario para obtener un resultado de AV fiable. El algoritmo usado para la realización de este estudio se presenta en la **Figura VIII. 8**, donde se han calculado los promedios de intensidades de todas las combinaciones de medidas de los espectros Raman adquiridos en la cara superior de los comprimidos (**Figura VIII. 2**). En el algoritmo, K es el número máximo de medidas individuales realizadas (16 en este trabajo), y n el número de medidas elegidas para el cálculo del promedio de la intensidad del pico analítico. El número de combinaciones de K elementos según el número n de grupos (sin repetición) es $\left(\frac{n}{K}\right) = Z_n$. El índice j es el contador para los grupos con n medidas. En cada caso, de n = 1 a K, se ha calculado el valor de intensidad promedio para Z_n posibles combinaciones de n medidas, y se ha representado el valor máximo y mínimo obtenido dado el gran número de combinaciones posibles. Por ejemplo, al realizar todas las posibles combinaciones de promediar hasta 8 de los 16 espectros individuales para un comprimido, se obtienen un total de 128700 combinaciones.

El algoritmo se ha aplicado para todos los comprimidos de todos los lotes. Los resultados obtenidos al aplicar BLC-N se muestran en la **Figura VIII. 9**.

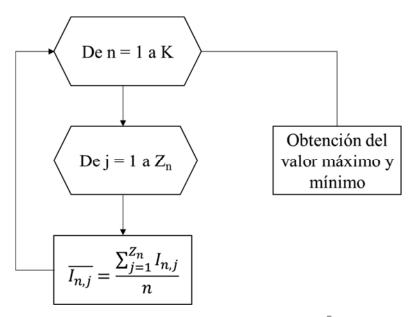


Figura VIII. 8. Algoritmo usado para el cálculo del promedio de las intensidades $\bar{I}_{(n,j)}$ de todas las combinaciones de medidas realizadas.

Para determinar el número óptimo de espectros se han establecido límites de control para las intensidades promedio como $\bar{x} \pm 3s$, donde \bar{x} es la intensidad media de los 16 espectros Raman registrados en cada tableta del lote y s su desviación estándar, El número mínimo de medidas se obtiene cuando todas las combinaciones posibles de los resultados están dentro de los límites. Se ve en la figura que en todos los lotes se deben adquirir, al menos, 12 espectros individuales para garantizar la homogeneidad del API.

También se puede calcular la varianza obtenida en cada combinación de grupos para ver si existen diferencias significativas entre ellas. En la **Figura VIII. 9B** se representa la relación del coeficiente de varianzas indicado en la **Ecuación VIII-6**, donde el numerador es la varianza hallada con el número de puntos, n, que se selecciona para el cálculo de todas las posibles combinaciones de cada grupo. El denominador es la varianza de las combinaciones con 16 puntos. El límite de control está determinado por el valor teórico de F considerando n-1=15 grados de libertad en el numerador e ∞ grados de libertad en el denominador, para un nivel de significación $\alpha = 0.05$. El valor de $F_{(15, \infty, 0.05)} = 1.67$.

Teniendo en cuenta los límites de control para el valor medio y la varianza, se deben adquirir al menos 12 espectros individuales para garantizar la homogeneidad del API.

$$s_n^2/s_{16}^2 \tag{VIII-6}$$

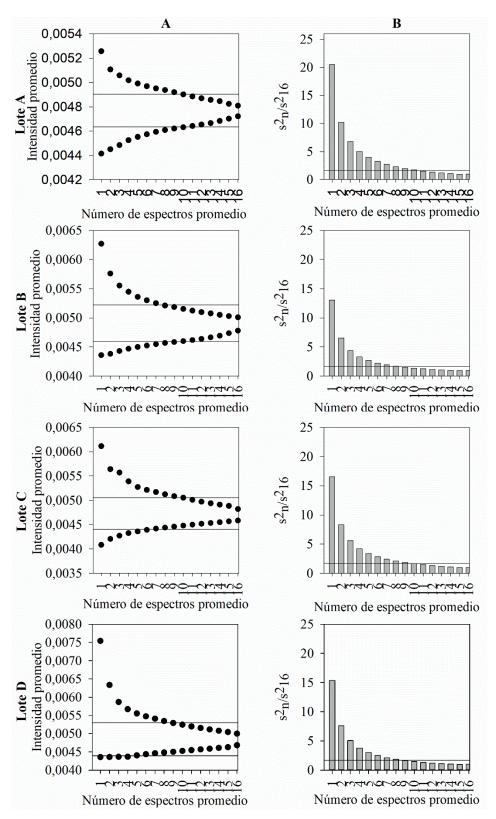


Figura VIII. 9. Resultados de la intensidad promedio tras la aplicación del análisis combinatorio para dosis unitarias de todos los lotes usando como pretratamiento BLC-N. (A) Representación del valor máximo y mínimo de la intensidad

del pico analítico; (B) Relación entre la varianza de todas las posibles combinaciones con n elementos y la varianza con las combinaciones de 16 elementos. El límite de control correspondo a $F_{(15,\infty,0.05)}$

Una reducción en el mapeo de 16 a 12 espectros supondría un tiempo de reducción de aproximadamente 3 min por comprimido. Para poder demostrar que, usando solo 12 espectros para determinar la distribución del API en el comprimido, se ha recalculado el AV promediando 12 espectros elegidos aleatoriamente entre los 16 adquiridos. La **Tabla VIII.**3 muestra los resultados. El resultado obtenido para el AV prácticamente no varía al disminuir en cuatro el número de espectros individuales adquiridos. En todos los casos, la prueba de UC sería conforme.

3.4. <u>Distribución del API en las mitades</u>

Los comprimidos estudiados presentan una marca en su cara superior que da la posibilidad de dividir el comprimido en sus dos mitades. Una de las ventajas de la espectroscopia Raman es que, gracias a su mapeo de puntos individuales, es posible adquirir espectros en cada una de sus mitades y evaluar la distribución del API en ambas partes. De este modo, a partir de los ocho espectros Raman adquiridos en cada una de las mitades, se puede evaluar la uniformidad del contenido para evaluar si la distribución del API en ambas mitades es homogénea.

Antes del cálculo de la uniformidad de contenido, se ha realizado una prueba t de datos pareados ($\alpha=0,05$) para comparar el promedio de la intensidad del pico analítico de ambas mitades para todos los comprimidos. Los resultados mostraron que, tras la realización de BLC-N, no hay diferencias significativas entre ambas mitades, con excepción de un único comprimido del lote C.

3.4.1. Uniformidad de las unidades de dosificación para mitades

La uniformidad de contenido se ha evaluado siguiendo el mismo procedimiento presentado para los comprimidos enteros. Por ello, cada mitad de cada comprimido se ha considerado como una unidad de dosificación propia.

Como se han evaluado diez comprimidos por lote, se dispone de un total de 20 mitades para evaluar. Sin embargo, la prueba de UC de la farmacopea requiere de diez unidades para la realización del cálculo del AV. Con el fin de reducir el riesgo a obtener un resultado no significativo de todo el conjunto de datos por la elección de una única posible combinación

de mitades, se ha calculado el resultado de una selección aleatoria de 1000 combinaciones de 10 mitades.

Los resultados de la UC que se muestran en la **Tabla VIII. 4** corresponden a los valores máximo y mínimo de AV calculados para las 1000 combinaciones y el porcentaje de combinaciones que pasan la prueba. Como puede verse, los resultados obtenidos se encuentran dentro del límite establecido. Sin embargo, cabe destacar que los valores de AV obtenidos son superiores que los vistos previamente al evaluar el comprimido entero.

Tabla VIII. 4. Resultados de AV usando el promedio de la intensidad del pico analítico de 8 y 6 espectros Raman individuales para la evaluación de las mitades. k = 2,4.

BLC-N								
Batch	8 espectros			6 e	spectro	S		
	Max	Min	%		Max	Min	%	
A	3,12	1,03	100		3,06	0,84	100	
В	7,01	2,37	100		8,49	2,52	100	
C	7,51	2,22	100		11,75	3,65	100	
D	9,94	3,52	100		12,43	1,85	100	

Para poder ver si se puede optimizar un poco el tiempo de análisis al realizar una normalización en los datos, dado que con ocho espectros individuales los lotes estudiados cumplen siempre con el criterio establecido, se realiza una optimización del número de puntos de mapeo.

3.4.2. Estudio del número mínimo de espectros en las mitades

La optimización del número mínimo de espectros que se deben registrar en una mitad del comprimido para obtener resultados de AV fiables se ha realizado mediante la aplicación del mismo algoritmo presentado anteriormente (**Figura VIII. 8**).

En este caso, para la realización del algoritmo, el valor máximo de medidas K=8. Después de obtener los resultados de los Z_n promedios de intensidad para cada mitad de comprimido, se ha calculado el valor máximo y mínimo y se ha realizado la representación. La **Figura VIII. 10** muestra los resultados obtenidos para las intensidades medias y la relación de varianzas. Se ha utilizado el mismo tipo de límites de control, $\bar{x}\pm 3s$ y de F límite que para los comprimidos enteros. En este caso, el valor de $F_{(7,\infty,0,05)}=2,01$ y se cumple para la relación de varianzas obtenidas a partir de 4 medidas. En el caso de utilizar el criterio sobre el valor medio $\bar{x}\pm 3s$, se cumple a partir de 6 medidas, por lo que se decide que el

número mínimo de puntos a considerar en el mapeo es de 6. Esta decisión se toma considerando que todo el sistema debe estar controlado, por lo que, se necesita que los promedios sean correctos y la varianza mínima. Se deben cumplir ambos criterios a la vez para concluir que el sistema de optimización está bajo control. Finalmente, se ha realizado el cálculo del valor de aceptación al promediar seis valores de intensidad Raman. Los resultados se muestran en la misma **Tabla VIII. 4**.

De las 1000 combinaciones aleatorias calculadas para las mitades, todas ellas proporcionan un resultado de UC dentro del límite establecido por la farmacopea al usar seis puntos de mapeo. Sin embargo, puede verse que el valor máximo de este parámetro llega hasta un valor de 12,43; válido, pero más próximo al límite (AV = 15) Este resultado coincide con el obtenido para la evaluación de la UC del comprimido entero, donde se establecía que era posible usar 12 espectros, que se podrían dividir en 6 espectros para cada una de las mitades del comprimido.

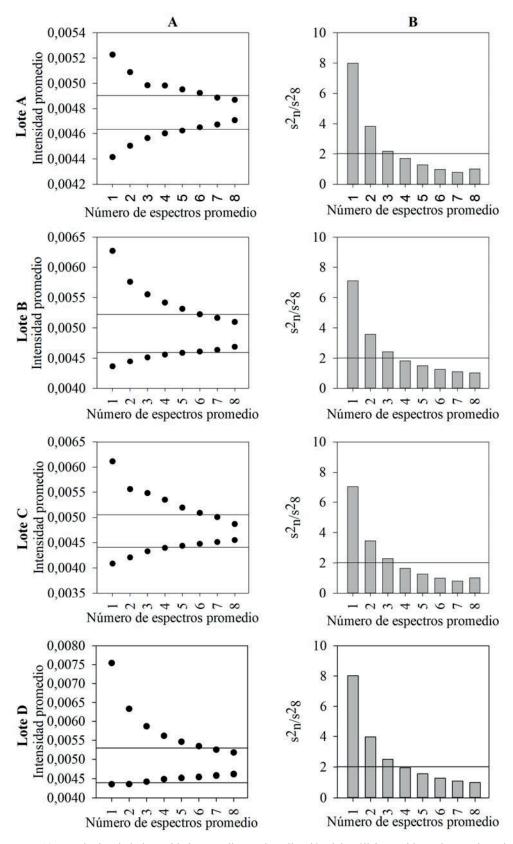


Figura VIII. 10. Resultados de la intensidad promedio tras la aplicación del análisis combinatorio para las mitades de todos los lotes usando como pretratamiento BLC-N. (A) Representación del valor máximo y mínimo de la intensidad del pico analítico; (B) Ratio entre la varianza de todas las posibles combinaciones. El límite de control correspondo a $F_{(7, \infty, 0.05)}$.

4. CONCLUSIONES

La imagen química de la distribución del API en un comprimido se puede obtener fácilmente mediante la aplicación de la espectroscopia Raman y la realización de una metódica de un mapeo en su superficie. Mediante la monitorización de la variabilidad de la intensidad de un pico analítico del API se puede estudiar la homogeneidad del API en el comprimido.

Al estudiar comprimidos divisibles, se ha podido evaluar si existían diferencias entre todas sus caras mediante la adquisición de espectros Raman en todas ellas. En el caso de la cara anterior y posterior, éstas no presentan diferencias significativas. En cambio, al estudiar la cara interior, se han visto diferencias en esta cara. Estas diferencias se atribuyen a su superficie debido a una pérdida de homogeneidad tras la división de los comprimidos.

Por otro lado, se ha demostrado posibilidad de calcular la UC directamente con el espectro Raman sin realizar una calibración previa. Este procedimiento es posible si se identifica un pico relacionado con el API, así se puede relacionar la concentración del analito directamente con la intensidad de la banda. Para poder calcular la UC, ha sido necesario preprocesar los espectros Raman. En este caso, se ha necesitado corregir la línea base y aplicar una normalización para ajustar la línea base del pico analítico bajo estudio.

Además, se ha establecido una estrategia para optimizar el tiempo de análisis de los comprimidos mediante Raman. Tras la aplicación de esta estrategia, se ha visto que es posible reducir de 16 a 12 espectros Raman adquiridos. Esta disminución implica una reducción del tiempo de análisis de la prueba de uniformidad.

A pesar de que los resultados y las conclusiones se refieren solamente a un producto farmacéutico con una determinada proporción de API, el procedimiento descrito ha demostrado la utilidad de la técnica Raman para este propósito.

5. BIBLIOGRAFÍA

- [1] Chapter 2.9.40. Uniformity of dosage units, Eur. Phamacopoeia 8.0. (2013) 357–359.
- [2] C. Berg, A. Ekedahl, *Dosages involving splitting tablets: common but unnecessary?*,
 J. Pharm. Heal. Serv. Res. 1 (2010) 137–141. https://doi.org/10.1111/j.1759-8893.2010.00020.x.

- [3] R.H. Richey, C. Hughes, J. V Craig, U.U. Shah, J.L. Ford, C.E. Barker, M. Peak, A.J. Nunn, M.A. Turner, *A systematic review of the use of dosage form manipulation to obtain required doses to inform use of manipulation in paediatric practice*, Int. J. Pharm. 518[1] R.H (2017) 155–166. https://doi.org/10.1016/j.ijpharm.2016.12.032.
- [4] D.A. Volpe, A. Gupta, A.B. Ciavarella, P.J. Faustino, V.A. Sayeed, M.A. Khan, *Comparison of the stability of split and intact gabapentin tablets*, Int. J. Pharm. 350 (2008) 65—69. https://doi.org/10.1016/j.ijpharm.2007.08.041.
- [5] J.M. Rosenberg, J.P. Nathan, F. Plakogiannis, Weight Variability of Pharmacist-Dispensed Split Tablets, J. Am. Pharm. Assoc. 42 (2002) 200–205. https://doi.org/10.1331/108658002763508498.
- [6] R.B. Shah, J.S. Collier, V.A. Sayeed, A. Bryant, M.J. Habib, M.A. Khan, *Tablet Splitting of a Narrow Therapeutic Index Drug: A Case with Levothyroxine Sodium*, AAPS PharmSciTech. 11 (2010) 1359–1367. https://doi.org/10.1208/s12249-010-9515-8.
- [7] N. Zhao, A. Zidan, M. Tawakkul, V.A. Sayeed, M. Khan, *Tablet splitting: Product quality assessment of metoprolol succinate extended release tablets*, Int. J. Pharm. 401 (**2010**) 25–31. https://doi.org/10.1016/j.ijpharm.2010.09.004.
- [8] A.B. Ciavarella, M.A. Khan, A. Gupta, P.J. Faustino, Dose uniformity of scored and unscored tablets: Application of the FDA Tablet Scoring Guidance for Industry, PDA J. Pharm. Sci. Technol. 70 (2016) 523–532. https://doi.org/10.5731/pdajpst.2016.006411.
- [9] I. Elliott, M. Mayxay, S. Yeuichaixong, S.J. Lee, P.N. Newton, *The practice and clinical implications of tablet splitting in international health*, Trop. Med. Int. Heal. 19 (2014) 754–760. https://doi.org/10.1111/tmi.12309.
- [10] J.K. Eserian, M. Lombardo, J.R. Chagas, J.C.F. Galduróz, Unmet Quality Needs in Oral Drug Delivery: Contrasts of Drug Content and Uniformity on Distinct Approaches for Achieving Individual Dosing, AAPS PharmSciTech. 20 (2019) 332. https://doi.org/10.1208/s12249-019-1546-1.
- [11] U.S. Department of Health and Human Services. Food and Drug Administration.,

- Guidance for Industry- Tablet Scoring: Nomenclature, Labeling, and Data for Evaluation, (2013).
- [12] M. U.S. Pharmacopoeial Convention: Rockville, 905 Uniformity of Dosage Units. United States Pharmacopoeia 39., (2016).
- [13] M. U.S. Pharmacopoeial Convention: Rockville, 1216 Tablet Friability, (2016).
- [14] U.S. Department of Health and Human Services. Food and Drug Administration., Guidance for Industry- Dissolution Testing of Immediate Release Solid Oral Dosage Forms, (1997).
- [15] M.I. Walash, F. Ibrahim, S.A. El Abass, A validated HPLC method for the determination of dimethyl-4,4'-dimethoxy-5,6,5',6'-dimethylene dioxybiphenyl-2,2'-dicarboxylate (DDB) with fluorescence detection in raw material and pill form: application to an in vitro dissolution test and a content unif, Luminescence. 29 (2014) 914–919. https://doi.org/10.1002/bio.2641.
- [16] T.E. Mürdter, D.U. Schulze, M. Schwab, S. Klumpp, *Determination of content uniformity of busulfan capsules by liquid chromatography–mass spectrometry*, Eur. J. Hosp. Pharm. Sci. Pract. 19 (2012) 443 LP 446. https://doi.org/10.1136/ejhpharm-2011-000047.
- [17] A.N. Zaid, R.J. Al-Ramahi, A.A. Ghoush, A. Qaddumi, Y.A. Zaaror, Weight and content uniformity of lorazepam half-tablets: A study of correlation of a low drug content product, Saudi Pharm. J. 21 (2013) 71–75. https://doi.org/10.1016/j.jsps.2011.12.009.
- [18] S.J. Mazivila, A.C. Olivieri, *Chemometrics coupled to vibrational spectroscopy and spectroscopic imaging for the analysis of solid-phase pharmaceutical products: A brief review on non-destructive analytical methods*, TrAC Trends Anal. Chem. 108 (2018) 74–87. https://doi.org/10.1016/j.trac.2018.08.013.
- [19] A. Farkas, B. Nagy, G. Marosi, Quantitative evaluation of drug distribution in tablets of various structures via Raman mapping, Period. Polytech. Chem. Eng. 62
 (2018) 1–7. https://doi.org/10.3311/PPch.10143.
- [20] H. Eksi-Kocak, S. Ilbasmis Tamer, S. Yilmaz, M. Eryilmaz, I.H. Boyaci, U. Tamer,

- Quantification and spatial distribution of salicylic acid in film tablets using FT-Raman mapping with multivariate curve resolution, Asian J. Pharm. Sci. 13 (2018) 155–162. https://doi.org/10.1016/j.ajps.2017.07.010.
- [21] K.C. Gordon, C.M. McGoverin, *Raman mapping of pharmaceuticals*, Int. J. Pharm. 417 (**2011**) 151–162. https://doi.org/10.1016/j.ijpharm.2010.12.030.
- [22] A. V Ewing, S.G. Kazarian, *Recent advances in the applications of vibrational spectroscopic imaging and mapping to pharmaceutical formulations*, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 197 (**2018**) 10–29. https://doi.org/10.1016/j.saa.2017.12.055.
- [23] L. Liu, H. Qu, Recent advancement of chemical imaging in pharmaceutical quality control: From final product testing to industrial utilization, J. Innov. Opt. Health Sci. 13 (2019) 1930014. https://doi.org/10.1142/S1793545819300143.
- [24] J. Arruabarrena, J. Coello, S. Maspoch, *Raman spectroscopy as a complementary tool to assess the content uniformity of dosage units in break-scored warfarin tablets*, Int. J. Pharm. 465 (**2014**) 299–305. https://doi.org/10.1016/j.ijpharm.2014.01.027.
- [25] D.A. Gómez, J. Coello, S. Maspoch, *Raman spectroscopy for the analytical quality control of low-dose break-scored tablets.*, J. Pharm. Biomed. Anal. 124 (**2016**) 207–15. https://doi.org/10.1016/j.jpba.2016.02.055.
- [26] W. Wang, H. Zhang, Y. Yuan, Y. Guo, S. He, *Research Progress of Raman Spectroscopy in Drug Analysis*, AAPS PharmSciTech. 19 (**2018**) 2921–2928. https://doi.org/10.1208/s12249-018-1135-8.
- [27] T. Firkala, A. Farkas, B. Vajna, I. Farkas, G. Marosi, *Investigation of drug distribution in tablets using surface enhanced Raman chemical imaging*, J. Pharm. Biomed. Anal. 76 (2013) 145–151. https://doi.org/10.1016/j.jpba.2012.12.017.
- [28] M.N. Slipchenko, H. Chen, D.R. Ely, Y. Jung, M.T. Carvajal, J.-X. Cheng, Vibrational imaging of tablets by epi-detected stimulated Raman scattering microscopy., Analyst. 135 (2010) 2613–2619. https://doi.org/10.1039/c0an00252f.
- [29] P.-Y. Sacré, P. Lebrun, P.-F. Chavez, C. De Bleye, L. Netchacovitch, E. Rozet, R.

- Klinkenberg, B. Streel, P. Hubert, E. Ziemons, *A new criterion to assess distributional homogeneity in hyperspectral images of solid pharmaceutical dosage forms*, Anal. Chim. Acta. 818 (**2014**) 7–14. https://doi.org/10.1016/j.aca.2014.02.014.
- [30] J.I. Jérez Rozo, A. Zarow, B. Zhou, R. Pinal, Z. Iqbal, R.J. Romañach, Complementary Near‐Infrared and Raman Chemical Imaging of Pharmaceutical Thin Films, J. Pharm. Sci. 100 (2011) 4888–4895. https://doi.org/10.1002/jps.22653.
- [31] H. Hisada, A. Okayama, T. Hoshino, J. Carriere, T. Koide, Y. Yamamoto, T. Fukami, *Determining the Distribution of Active Pharmaceutical Ingredients in Combination Tablets Using Near IR and Low-Frequency Raman Spectroscopy Imaging*, Chem. Pharm. Bull. 68 (2020) 155–160. https://doi.org/10.1248/cpb.c19-00791.
- [32] H. Rebiere, C. Ghyselinck, L. Lempereur, C. Brenier, *Investigation of the composition of anabolic tablets using near infrared spectroscopy and Raman chemical imaging*, Drug Test. Anal. 8 (2016) 370–377. https://doi.org/10.1002/dta.1843.
- [33] H. Rebiere, M. Martin, C. Ghyselinck, P.-A. Bonnet, C. Brenier, *Raman chemical imaging for spectroscopic screening and direct quantification of falsified drugs*, J. Pharm. Biomed. Anal. 148 (2018) 316–323. https://doi.org/10.1016/j.jpba.2017.10.005.
- [34] L.M. Kandpal, B.-K. Cho, J. Tewari, N. Gopinathan, *Raman spectral imaging technique for API detection in pharmaceutical microtablets*, Sensors Actuators B Chem. 260 (2018) 213–222. https://doi.org/10.1016/j.snb.2017.12.178.
- [35] M.I. Sanhueza, M.F. Meléndrez, C. von Plessing, C.Y. Bastidas, J.Y. Neira, J. Freer, R. del P. Castillo, *Raman microimaging as an analytical technique for simultaneous quantification and localization of active principles in pharmaceutical solid dosage forms*, J. Raman Spectrosc. 51 (2020) 649–659. https://doi.org/10.1002/jrs.5833.
- [36] B. Vajna, A. Farkas, H. Pataki, Z. Zsigmond, T. Igricz, G. Marosi, *Testing the performance of pure spectrum resolution from Raman hyperspectral images of*

- differently manufactured pharmaceutical tablets, Anal. Chim. Acta. 712 (2012) 45–55. https://doi.org/10.1016/j.aca.2011.10.065.
- [37] B. Vajna, I. Farkas, A. Farkas, H. Pataki, Z. Nagy, J. Madarász, G. Marosi, Characterization of drug-cyclodextrin formulations using Raman mapping and multivariate curve resolution, J. Pharm. Biomed. Anal. 56 (2011) 38–44. https://doi.org/10.1016/j.jpba.2011.05.005.
- [38] X. Chen, K. Stoneburner, M. Ladika, T.-C. Kuo, T.H. Kalantar, *High-Throughput Raman Spectroscopy Screening of Excipients for the Stabilization of Amorphous Drugs*, Appl. Spectrosc. 69 (2015) 1271–1280. https://doi.org/10.1366/15-07914.
- [39] S. Šašić, *Chemical imaging of pharmaceutical granules by Raman global illumination and near-infrared mapping platforms*, Anal. Chim. Acta. 611 (**2008**) 73–79. https://doi.org/10.1016/j.aca.2008.01.063.
- [40] B. Li, Y. Casamayou-Boucau, A. Calvet, A.G. Ryder, *Chemometric approaches to low-content quantification (LCQ) in solid-state mixtures using Raman mapping spectroscopy*, Anal. Methods. 9 (2017) 6293–6301. https://doi.org/10.1039/C7AY01778B.
- [41] M. Windbergs, M. Haaser, C.M. McGoverin, K.C. Gordon, P. Kleinebudde, C.J. Strachan, *Investigating the relationship between drug distribution in solid lipid matrices and dissolution behaviour using Raman spectroscopy and mapping.*, J. Pharm. Sci. 99 (2010) 1464–1475. https://doi.org/10.1002/jps.21894.
- [42] J.. Rindone, Evaluation of tablet-splitting in patients taking lisinopril for hypertension., J. Clin. Outcomes Manag. (2002) 22–24.
- [43] K.M.S. Fahelelbom, M.M.M. Al-Tabakha, N.A.M. Eissa, J. Javadi, Evaluation of Certain Pharmaceutical Quality Attributes of Lisinopril Split Tablets, Sci. Pharm. 84 (2016) 646–653. https://doi.org/10.3390/scipharm84040646.
- [44] D.A. Gómez, J. Coello, S. Maspoch, *The influence of particle size on the intensity and reproducibility of Raman spectra of compacted samples*, Vib. Spectrosc. 100 (2019) 48–56. https://doi.org/10.1016/j.vibspec.2018.10.011.
- [45] H. A. Laitinen, W. E. Harris, Chapter 27: Sampling, in: Chem. Anal. An Adv. Text

ESTUDIO DE LA DISTRIBUCIÓN DEL CONTENIDO DE API EN COMPRIMIDOS UTILIZANDO ESPECTROSCOPIA RAMAN

- Ref., 1975: pp. 565-582.
- [46] LibreTexts, 6.4. Normal Approximation to the Binomial Distribution, in: Stat. Probab., **2022**.
- [47] M.L. Scherholz, B. Wan, G. McGeorge, *A Rational Analysis of Uniformity Risk for Agglomerated Drug Substance Using NIR Chemical Imaging*, AAPS PharmSciTech. 18 (2017) 432–440. https://doi.org/10.1208/s12249-016-0523-1.
- [48] S. Šašić, W. Yu, L. Zhang, *Monitoring of API particle size during solid dosage form manufacturing process by chemical imaging and particle sizing*, Anal. Methods. 3 (2011) 568–574. https://doi.org/10.1039/C0AY00562B.

IX. <u>DETERMINACIÓN DEL CONTENIDO DE DOS</u> <u>APIS EN COMPRIMIDOS MEDIANTE NIR</u>

1. INTRODUCCIÓN

La espectroscopia NIR es una técnica analítica muy utilizada en la industria farmacéutica y que puede proporcionar información tanto química como física de los materiales analizados. Algunas de las principales aplicaciones están relacionadas con el análisis de los materiales de partida, la monitorización de distintas etapas de un proceso o la determinación del contenido de API en productos terminados [1–3]. Muchas de estas aplicaciones se realizan usando equipos NIR de sobremesa, que suelen ser costosos y de gran volumen.

Los espectrómetros NIR miniaturizados son una buena alternativa a los equipos de sobremesa dada su portabilidad y costes [4–6]. El pequeño tamaño de estos instrumentos y su portabilidad los convierten en una gran herramienta para la realización y monitorización de análisis *on-line*. Además, dadas sus características, su utilización para el control de calidad de productos farmacéuticos presenta un gran atractivo.

Uno de los espectrómetros NIR miniaturizados más conocidos es el MicroNIR de VIAVI Solutions, presentado en el Capítulo III. Este instrumento portátil se ha usado sobre todo para aplicaciones cuantitatives relacionadas con la industria alimentaria [7–9]. Sin embargo, para la industria farmacéutica, solo se han encontrado unas pocas aplicaciones relacionadas con métodos de identificación [10–12].

El objetivo principal de este trabajo es evaluar el uso del instrumento MicroNIR para la cuantificación de dos APIs en comprimidos comercializados. A pesar de que el uso de la técnica NIR para la determinación de APIs en productos farmacéuticos ha sido muy estudiada, la mayoría de estos procesos se basan en la cuantificación de un solo API. La complejidad de cuantificar 2 APIs está relacionada con la diferencia en la proporción en masa que puede representar cada uno. Una gran diferencia en masa para ambos APIs puede suponer que la contribución espectral del mayoritario influya en la determinación del minoritario. La **Tabla IX. 1** presenta un resumen de algunos de los trabajos donde se han cuantificado más de un API [13–21].

Tabla IX. 1. Resumen de métodos NIR para la caracterización de productos acabados con más de un API.

Producto	APIs	mg API/comp	% API (m/m)	Ref.	
Comprimidos	Amlodipino	14	-	13	
	Valsartán	80	-	13	
Proceso de	API 1	-	86	14	
homogeneización	API 2	-	4	14	
Comprimidos	Bisoprolol	2,5 – 10	1,8 – 5,9	15	
	Hidroclorotiazida	6,25 - 25	3,7 – 14,7		
Comprimidos	Irbesartán	300	-	1.0	
	Hidroclorotiazida	25	-	16	
Comprimidos	Valerato de estradiol	2	1,6	17	
	Acetato de medroxiprogesterona	10	8,0		
Mezclas para	Atorvastatina cálcica	-	7,58	10	
compresión	Besilato de amlodipino	-	9,25	18	
Proceso de	API 1	-	22,7		
homogeneización	API 2	-	16,5	10	
	API 3	-	0,98	19	
	API 4	-	0,57		
Gránulos	Paracetamol		6,50		
	Ácido ascórbico		2,50		
	Hidrobromuro de dextrometorfano		0,20	20	
	Cafeína		0,15		
	Clorfeniramina maleato		0,04		
Comprimidos	Paracetamol	250	26,3		
	Ácido acetilsalicílico	250	26,3	21	
	Cafeína	65	6,8		

En este caso, el producto farmacéutico seleccionado para la evaluación del desempeño del instrumento miniaturizado presenta una gran diferencia en la proporción de sus dos APIs. El API mayoritario, Paracetamol (PCT), representa aproximadamente un 81% m/m del comprimido mientras que el API minoritario, Tramadol (TRA), representa solo un 9% m/m.

Esta gran diferencia en la proporción de ambos APIs dificulta la cuantificación del minoritario, siendo un reto su determinación. Como se ve en la tabla, existe un estudio donde también se encuentra una desproporción similar en el contenido de dos APIs, Sin embargo, el objetivo principal del estudio era la monitorización del proceso de mezclado y no la cuantificación de los APIs [14].

El API minoritario, TRA, es un derivado de la morfina que forma parte del grupo de opioides y se considera uno de los analgésicos más seguros del mercado. Sin embargo, después de su comercialización, se han reportado un incremento en el riesgo a obtener convulsiones al tomar esta sustancia [22,23]. En cambio, el API mayoritario, PCT, es considerado un analgésico para el tratamiento de dolores y no se ha reportado ningún problema con su uso. La combinación de ambos APIs se usa como analgésico para el tratamiento del dolor moderado.

En este trabajo, se evaluará la capacidad de un instrumento miniaturizado para cuantificar el contenido de ambos APIs aún y su gran diferenciación en el porcentaje de peso del comprimido. Los resultados se compararán con los obtenidos con un instrumento de sobremesa convencional en términos de exactitud y precisión.

2. MATERIALES Y MÉTODOS

2.1. Muestras

Los comprimidos de Tramadol/Paracetamol (TRA/PCT) estudiados son comercializados por Tarbis Farma en estuches de 20 o 60 unidades. Este producto contiene dos APIs en su formulación en una proporción estimada del 81,3% m/m y 9,3% m/m para el PCT y TRA, respectivamente. El porcentaje restante está compuesto por los excipientes de la formulación.

El trabajo se ha realizado con comprimidos de tres lotes comerciales distintos (A, B y C) adquiridos en farmacias cercanas a la UAB. Para cada lote, han sido analizados un total de 10 comprimidos, tanto por NIR como por HPLC, siguiendo el método descrito en el Capítulo III (Sección 3).

2.2. Comprimidos preparados en el laboratorio

El conjunto de comprimidos preparados en el laboratorio se preparó sobre- y subdosificando comprimidos comerciales. Para ello, se molturaron comprimidos y se les añadió una cierta cantidad de PCT, TRA y almidón siguiendo un diseño establecido para minimizar la correlación entre los componentes [24,25]. De esta manera, las muestras preparadas en el laboratorio también contienen la variabilidad asociada al proceso de fabricación. Todas las mezclas preparadas se homogeneizaron en una Turbula T2C antes de realizar su compresión a 140 MPa en una prensa (Perkin Elmer 15.000).

El diseño experimental usado se presenta en la **Figura IX. 1**. Se han preparado comprimidos a nueve niveles de concentración de PCT. En cada nivel, se ha variado la concentración de TRA para obtener 3 comprimidos por cada nivel de PCT, con el objetivo de reducir la colinealidad entre APIs y excipientes. Se prepararon 27 comprimidos variando la concentración de TRA entre 6 y 34 % m/m, y la de PCT entre 65 y 90 % m/m. Además, se prepararon 6 comprimidos extra entre los límites establecidos del diseño para usarlos en el conjunto de validación.

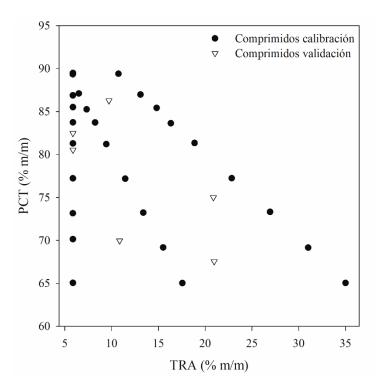


Figura IX. 1. Diseño experimental para la preparación de los comprimidos en el laboratorio.

2.3. Adquisición de espectros

Los espectros de los comprimidos del estudio han sido registrados usando el instrumento NIR de sobremesa (FOSS) y el instrumento NIR portátil (MicroNIR).

Los parámetros de registro para el MicroNIR han sido: tiempo de adquisición de 10 ms, usando un total de 100 barridos en el intervalo de longitudes de onda de 908 – 1676 nm. En el caso del FOSS, cada espectro fue el promedio de 32 barridos en el intervalo de longitudes de onda de 1100 – 2498 nm. En ambos casos, se ha utilizado el modo reflectancia. Los comprimidos se han colocado en la ventana correspondiente y se han obtenido tres espectros para cada comprimido, rotando su posición. Finalmente, para realizar la construcción del modelo cuantitativo, se ha usado el promedio de los tres registros.

2.4. Análisis de los datos

Los modelos cuantitativos de regresión se han construido con el algoritmo PLS. Durante el desarrollo del modelo, se han ido probando distintos pretratamientos e intervalos de longitudes de onda, considerando las regiones de absorción de los dos APIs. Los mejores modelos se han escogido considerando distintos parámetros estadísticos como el RMSECV, RMSEP y la varianza explicada.

Los modelos PLS se construyeron mediante una validación cruzada aleatoria, usando 20 segmentos (2 muestras por segmento). El análisis de datos multivariable se realizó con el software The Unscrambler X 10.5 (CAMO, Noruega).

Además, se realizó una prueba t pareada entre los resultados obtenidos con HPLC y las predicciones NIR para poder garantizar que no existen diferencias significativas entre ambos resultados a un nivel de significación $\alpha = 0.05$.

3. RESULTADOS Y DISCUSIÓN

El objetivo principal de este trabajo es comprobar el potencial de un instrumento miniaturizado frente a los tradicionales equipos de sobremesa. Una de las peculiaridades del instrumento miniaturizado estudiado (MicroNIR), es que cubre un intervalo de longitudes de onda de longitudes de onda menor que el instrumento de mesa (**Figura IX. 2**. Considerando esta diferencia, la comparación de los resultados se ha realizado en dos etapas. En la primera, usando el intervalo de longitudes de onda común para ambos instrumentos

(1100 – 1676 nm), para observar las posibles diferencias en los resultados debido al cambio de instrumento. En la segunda etapa, se optimizarán los modelos independientemente. Es decir, se buscará el mejor intervalo de longitudes de onda y pretratamiento individual. El objetivo final en ambas etapas es el mismo: predecir la concentración de PCT y TRA en comprimidos comerciales.

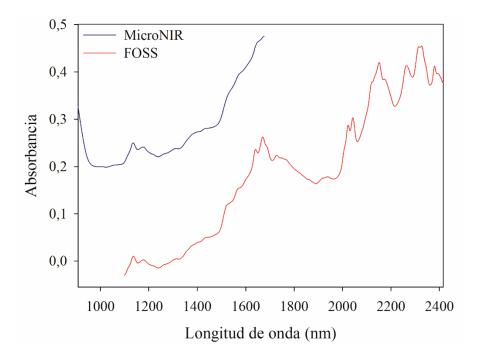


Figura IX. 2. Comparación intervalo de longitudes de onda de ambos instrumentos NIR.

Uno de los problemas en el desarrollo de los modelos PLS es la gran proporción en masa de uno de los APIs, que representa aproximadamente el 81% m/m de todo el comprimido. La **Figura IX. 3** presenta el espectro puro de ambos APIs junto con el espectro de un comprimido comercial para ambos equipos. Como puede observarse, el espectro del producto comercial es prácticamente el espectro del PCT. Una de las diferencias espectrales más clara entre el espectro de ambos APIs es el desplazamiento de la banda de 1200 nm, que se atribuye al segundo armónico del enlace C-H, y a la banda de 1550 nm presente en el espectro del PCT correspondiente al armónico del enlace N-H. Además, en la región de bandas de combinación del enlace C-H (2200 – 2400 nm) se pueden ver diferencias en la absorbancia de ambos APIs.

Para poder cuantificar la similitud entre los espectros presentados en la Figura IX. 3, se ha calculado su coeficiente de correlación. En la Tabla IX. 2;Error! No se encuentra el

origen de la referencia. pueden verse los resultados obtenidos para los espectros en absorbancia y tras la aplicación de pretratamientos de la señal. Como se observa, la correlación entre el espectro de PCT y el comprimido es muy alta, tanto en absorbancia como tras la eliminación del efecto de dispersión de la luz (*scattering*). Al aplicar una segunda derivada, las diferencias entre el espectro de TRA y PCT aumentan. Considerando estas correlaciones, será muy importante la selección de un pretratamiento e intervalo de longitudes de onda que permitan magnificar las diferencias entre APIs para la obtención de los modelos individuales.

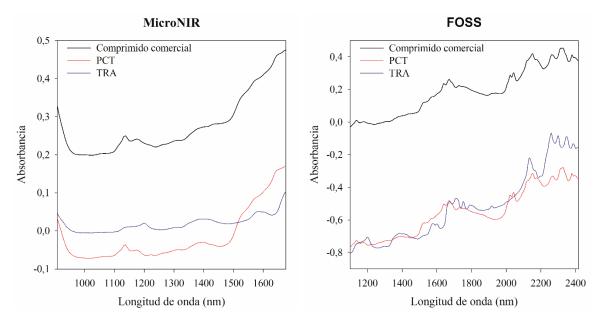


Figura IX. 3. Espectros NIR de ambos APIs y de un comprimido comercial para ambos instrumentos.

Tabla IX. 2. Coeficientes de correlación entre los dos APIs y entre cada uno de los APIs y el comprimido comercial utilizando ambos instrumentos en el intervalo de longitudes de onda común (1100 – 1646 nm) en absorbancia (Abs) y tras la aplicación de pretratamientos.

	FOSS				MicroNIR		
	Abs	Abs SNV SNV + 2D			SNV	SNV + 2D	
PCT-TRA	0,88	0,88	0,10	0,88	0,87	0,12	
PCT-comprimido	0,99	0,99	0,98	0,99	0,99	0,93	
TRA-comprimido	0,89	0,89	0,11	0,91	0,99	0,30	

3.1. <u>Modelos PLS construidos usando el mismo intervalo de longitudes de</u> onda (1100 – 1646 nm)

3.1.1. Construcción de los modelos

La comparación del desempeño de ambos instrumentos se ha realizado usando el mismo intervalo de longitudes de onda y pretratamiento. Se han probado distintas combinaciones de pretratamientos hasta obtener el mejor modelo según los criterios comentados en la Sección 2.4.

Se ha seleccionado el uso de SNV para reducir el efecto del *scattering* y una segunda derivada (algoritmo Savitzky-Golay, segundo orden, 15 puntos de suavizado) para aumentar la diferenciación de los espectros de los productos del comprimido. La **Figura IX. 4**presenta los espectros obtenidos tras la aplicación de estos pretratamientos para los comprimidos comerciales y los preparados en el laboratorio.

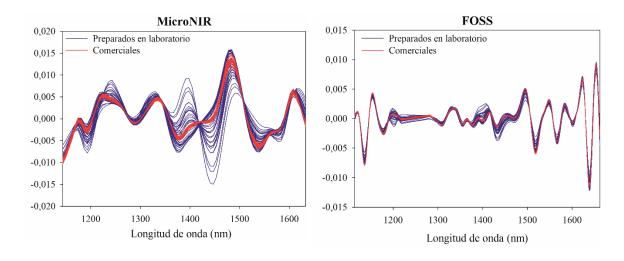


Figura IX. 4. Espectros NIR los comprimidos comerciales y preparados en el laboratorio tras la aplicación de una SNV + 2D en el intervalo de longitudes de onda común.

El conjunto de calibración está formado por los comprimidos preparados en el laboratorio más 8 comprimidos comerciales del lote B. En la presentación del diseño experimental realizado, se puede ver que la concentración de TRA se encuentra en el intervalo de 6 – 34 %m/m. Sin embargo, durante el desarrollo de los modelos, se observó que, al incluir un intervalo de concentración tan amplio, se perdía la linealidad, por lo que se utilizaron únicamente 23 de los comprimidos preparados en el laboratorio para conseguir un intervalo de la concentración de TRA de 6 a 18 %m/m.

La **Tabla IX.** 3 presenta un resumen de los parámetros NIR de los dos modelos desarrollados y la **Figura IX.** 5 las rectas de verificación de la exactitud del modelo de calibración, en las que, para cada muestra, se representa el valor calculado utilizando CV frente al valor de referencia.

El número de factores escogido para el modelo es el que proporciona el mínimo valor de RMSECV cuando se representa frente al número de factores. Para evaluar la veracidad (*trueness*) de los modelos construidos, se ha estudiado si la región de confianza conjunta de la pendiente, b, y la ordenada en el origen, a, de las rectas de la **Figura IX. 5** contienen los puntos 1 y 0, respectivamente. Esta región de confianza conjunta es una región elíptica definida por el conjunto de valores que cumplen la **Ecuación IX-1.**

$$\sum_{i=1}^{N} (Y_i - a - bX_i)^2 - \hat{S}\left(1 + \frac{F}{N-2}\right) = 0$$
 (IX-1)

Donde X_i e Y_i son los pares de los N valores de la recta y \hat{S} la suma de mínimos cuadrados. F es el valor crítico de la distribución con 2 y (N-2) grados de libertad para el nivel de confianza (1- α).[26,27]. Para ambos analitos, en todos los casos, ambas regiones incluían los dos valores, por lo que se confirma la veracidad del modelo PLS para un nivel de confianza 1- α .

Tabla IX. 3. Resumen de los parámetros NIR de los dos modelos desarrollados usando el mismo intervalo de longitudes de onda y pretratamiento.

	FC	OSS	Micr	oNIR
	PCT	TRA	PCT	TRA
Calibración				
Nº de muestras	31	31	31	31
Pretratamiento	SNV+2D	SNV+2D	SNV+2D	SNV+2D
Intervalo de	1100-1676	1100-1676	1100-1676	1100-1676
longitudes de onda				
(nm)				
Nº factores PLS	2	3	2	5
RMSEC (% m/m)	0,95	0,31	1,01	0,40
Ordenada de origen	$0,85 \pm 2,00$	0.07 ± 0.16	$1,09 \pm 2,11$	$0,05 \pm 0,14$
Pendiente	$0,98 \pm 0,03$	$0,99 \pm 0,02$	$0,98 \pm 0,03$	$0,99 \pm 0,01$
\mathbb{R}^2	0,98	0,99	0,98	0,99
Validación				
Nº de muestras	24	24	24	24
Ordenada de origen	$1,44 \pm 2,20$	$0,17 \pm 0,20$	$0,58 \pm 0,33$	$0,07 \pm 0,23$
Pendiente	$0,98 \pm 0,03$	$0,98 \pm 0,02$	$0,99 \pm 0,02$	$0,99 \pm 0,02$
\mathbb{R}^2	0,98	0,99	0,98	0,99
RMSEP (% m/m)	1,31	0,56	1,38	0,65
Bias (%m/m)	-0,67	0,02	-0,39	-0,04

Los modelos construidos para el PCT muestran resultados prácticamente iguales, ya que son necesarios el mismo número de factores para ambos instrumentos. Sin embargo, para los modelos del TRA el número de factores necesario es mayor para el instrumento miniaturizado. La necesidad de emplear un número mayor de factores puede estar relacionada con la diferente resolución de los espectros registrados en cada instrumento y con la menor proporción de TRA en la formulación. Como se observa en la **Figura IX. 4.**, el número de bandas que aparecen en el espectro registrado con el instrumento portátil es menor que en el espectro registrado con el instrumento de sobremesa. La resolución espectral del instrumento MicroNir es de 6 nm, mientras que la del instrumento FOSS es de 2 nm.

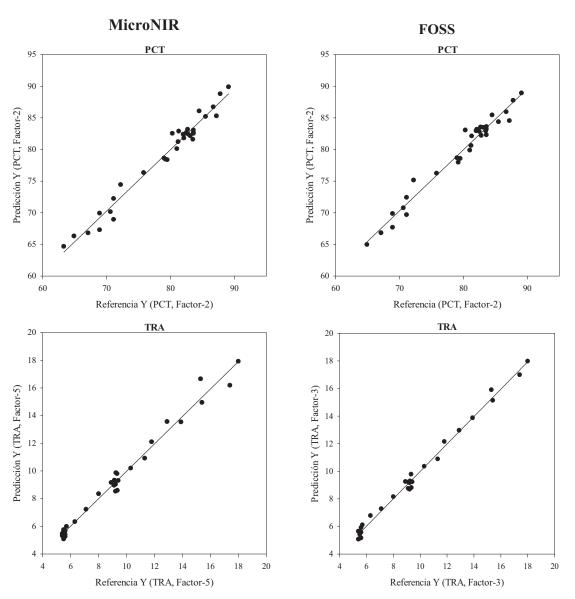


Figura IX. 5. Rectas de verificación de la exactitud de los modelos de calibración. Obtenidas para ambos instrumentos y analitos a partir de los mejores modelos construidos en el intervalo común de longitudes de onda.

3.1.2. Validación de los modelos

La validación de los modelos construidos se realizó calculando el contenido de diez comprimidos adicionales de los lotes A y C que no se habían incluido en la construcción de los modelos y 4 comprimidos preparados en el laboratorio. Los resultados para ambos instrumentos y APIs se muestran en la **Tabla IX. 4**. El contenido expresado es el promedio de los diez comprimidos analizados de cada lote.

Tabla IX. 4. Resultados medios de la validación externa (valor ± desviación estándar) de los modelos PLS desarrollados usando mismo intervalo de longitudes de onda y pretratamiento.

		PCT		TRA		
		% m/m	% m/m mg/comp		mg/comp	
HPLC	Α	$82,1 \pm 0,8$	$329,1 \pm 3,0$	$9,1 \pm 0,1$	$36,4 \pm 0,5$	
	C	$82,3 \pm 0,8$	$327,2 \pm 2,7$	$9,1 \pm 0,1$	$36,4 \pm 0,5$	
FOSS	A	$82,8 \pm 0,4$	$331,5 \pm 1,2$	$9,1 \pm 0,3$	$36,6 \pm 1,1$	
	C	$83,1 \pm 0,4$	$330,4 \pm 2,3$	$9,1 \pm 0,2$	$36,0 \pm 0,7$	
MicroNIR	A	$82,6 \pm 0,5$	$330,7 \pm 1,7$	$8,8 \pm 0,4$	$35,3 \pm 1,5$	
	C	$82,7 \pm 0,5$	$328,6 \pm 2,5$	$9,1 \pm 0,3$	$36,3 \pm 1,2$	

Para poder comparar los resultados obtenidos en la validación con el método de referencia, se ha realizado una prueba t de datos pareados entre los resultados, a un nivel de significación del 0,05. En el caso del instrumento FOSS, no se han encontrado diferencias significativas entre los resultados. Sin embargo, en el caso del instrumento MicroNIR, se encuentran diferencias significativas en la determinación del contenido de TRA del lote A. Se intentará construir un modelo individual para cada analito para obtener una correcta determinación del contenido de TRA con el instrumento miniaturizado.

3.2. <u>Modelos individuales construidos para ambos APIs y</u> espectrofotómetros

3.2.1. Construcción de los modelos

Los modelos individuales para cada analito e instrumento se han desarrollado siguiendo el mismo esquema presentado anteriormente. Se ensayaron distintas combinaciones de intervalos de longitudes de onda y pretratamientos, considerando siempre las bandas de absorción de ambos analitos. Los parámetros NIR de los modelos finales obtenidos se muestran en la **Tabla IX**. 5.

Como puede verse en la **Tabla IX. 5**, en el caso de los modelos óptimos para el instrumento FOSS no mejoran los parámetros estadísticos obtenidos previamente; tanto los

modelos individuales como los construidos en el intervalo de longitudes de onda común para ambos instrumentos, muestran resultados similares para la determinación de PCT y TRA. Lo mismo sucede para la determinación de PCT con el instrumento MicroNIR. Donde sí se observa una pequeña mejora es en el modelo PLS para cuantificar el contenido de TRA con el instrumento miniaturizado. Estas pequeñas mejoras se aprecian mejor durante la validación de los modelos, sobre todo en el análisis de los residuales obtenidos.

Tabla IX. 5. Resumen de los parámetros NIR de los dos modelos individuales desarrollados para ambos instrumentos y analitos.

	FC	OSS	MicroNIR		
	PCT	TRA	PCT	TRA	
Calibración					
Nº de muestras	31	31	31	31	
Pretratamiento	SNV + 2D	2D	SNV + 2D	2D	
Intervalo de longitudes de onda (nm)	1100-2406	1100-1780	982-1366	1087-1329	
Nº factores PLS	2	3	3	5	
RMSEC (% m/m)	0,93	0,30	1,01	0,33	
Ordenada de origen	$1,58 \pm 2,07$	0.06 ± 0.15	$2,02 \pm 2,33$	$0,07 \pm 0,16$	
Pendiente	$0,98 \pm 0,03$	$0,99 \pm 0,02$	$0,97 \pm 0,03$	$0,99 \pm 0,02$	
\mathbb{R}^2	0,99	0,99	0,99	0,99	
Validación					
Nº de muestras	24	24	24	24	
Ordenada de origen	$1,77 \pm 2,37$	$0,06 \pm 0,20$	$2,37 \pm 2,70$	$0,14 \pm 0,22$	
Pendiente	$0,98 \pm 0,03$	$0,99 \pm 0,02$	$0,97 \pm 0,03$	$0,98 \pm 0,02$	
\mathbb{R}^2	0,99	0,99	0,97	0,99	
RMSEP (% m/m)	1,32	0,45	1,22	0,58	
Bias (%m/m)	0,10	-0,05	-0,03	0,11	

En la **Figura IX. 6** se muestran las rectas de verificación de la exactitud del calibrado realizado con ambos instrumentos. Todas ellas cumplen con el criterio veracidad de la región de confianza conjunta de la pendiente y de la ordenada en el origen.

3.2.2. Validación de los modelos

La **Tabla IX.** 6 presenta los resultados de la validación externa de los diez comprimidos de los lotes A y C.

Tabla IX. 6. Resultados de la validación externa de los modelos óptimos desarrollados.

		P	PCT TRA			
		% m/m	mg/comp	% m/m mg/comp		
IIDI C	A	$82,1 \pm 0,8$	$329,1 \pm 3,0$	$9,1 \pm 0,1$	$36,4 \pm 0,5$	
HPLC	C	$82,3 \pm 0.8$	$327,2 \pm 2,7$	9.1 ± 0.1	36.4 ± 0.5	

		P	CT	TRA		
		% m/m mg/comp		% m/m	mg/comp	
FOSS	A	$82,3 \pm 0,6$	$329,1 \pm 1,6$	$9,3 \pm 0,3$	$36,8 \pm 1,1$	
	\mathbf{C}	$82,0 \pm 0,6$	$326,2 \pm 2,9$	$9,2 \pm 0,2$	$36,4 \pm 0,8$	
MicroNIR	A	$82,7 \pm 0,9$	$329,4 \pm 2,3$	$9,0 \pm 0,4$	$36,0 \pm 1,5$	
	C	$82,8 \pm 0,6$	$327,5 \pm 2,6$	$9,0 \pm 0,3$	$35,9 \pm 1,2$	

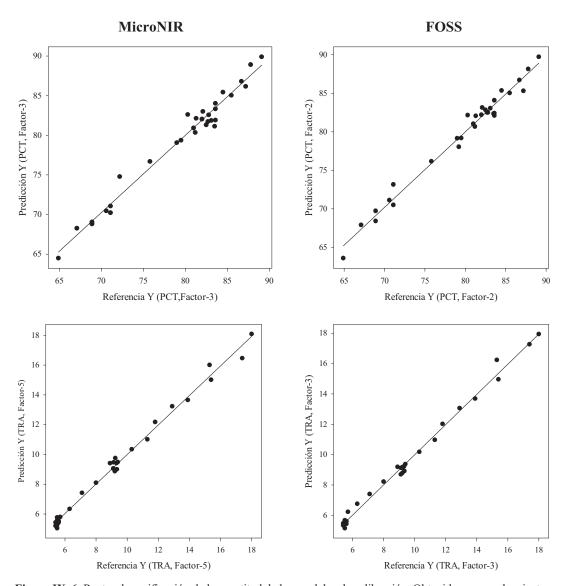


Figura IX. 6. Rectas de verificación de la exactitud de los modelos de calibración. Obtenidas para ambos instrumentos y analitos a partir de los mejores modelos individuales construidos.

Al utilizar un modelo optimizado individualmente para cada API, no se observan diferencias significativas entre los resultados HPLC y las predicciones NIR para ninguno de los dos lotes analizados.

3.2.2.1. Repetibilidad y precisión intermèdia

Durante la validación de los modelos individuales, se evaluó la repetibilidad y la precisión intermedia. La repetibilidad se estimó registrando el mismo comprimido diez veces consecutivas, cambiando cada vez la posición de éste en la ventana de medida de ambos equipos. La precisión intermedia se estimó registrando el espectro de un mismo comprimido durante cinco días consecutivos. Para la evaluación, se ha calculado el valor de la desviación estándar relativa (RSD%). Los resultados se muestran en la **Tabla IX. 7**.

Tabla IX. 7. Repetibilidad y p	recisión intermedia,	expresadas como %RS	O, para los modelos PLS individuales.
---------------------------------------	----------------------	---------------------	---------------------------------------

	Repetibilidad (n = 10)	Precisión intermedia (n = 5)
PCT (FOSS)	0,18	0,22
PCT (MicroNIR)	0,35	0,67
TRA (FOSS)	0,99	1,11
TRA (MicroNIR)	0,54	1,75

Como es de esperar, los valores de estos dos parámetros son mejores para los modelos del analito con mayor concentración (PCT). Los resultados obtenidos para ambos APIs se encuentran dentro de los límites establecidos por la EMA y ICH, \leq 2%, por lo que para ambos instrumentos se cumplen los criterios de repetibilidad y precisión intermedia.[28,29]

En la evaluación de la precisión intermedia, los resultados obtenidos con el instrumento de sobremesa son mejores que los obtenidos con el instrumento miniaturizado. Esta diferencia en los resultados está relacionada con la posición del comprimido en la celda de medida, ya que el equipo de sobremesa permite fijarlo siempre en la misma posición, mientras que pare el instrumento miniaturizado no se dispone de ningún accesorio que permita mantener los comprimidos en una posición fija respecto al instrumento.

3.2.2.2. Uniformidad de dosis unitarias (UDU)

La uniformidad de dosis unitarias se ha evaluado siguiendo la monografía específica de la Farmacopea Europea (EP).[30] En la monografía 2.9.40. se establecen las pautas en las que es necesario o no la realización de esta prueba y el método con el cual debe evaluarse: uniformidad de contenido (*Content uniformity*, UC, según EP) o variación de masa (*Mass variation*, según EP). Para el analito con mayor concentración, PCT, no sería necesario la realización de la prueba UC, pero se ha realizado como un ejemplo, dado que también se indica en la monografía que esta prueba puede realizarse en todos los casos.

La prueba de UC se basa en el cálculo de un valor de aceptación (*Acceptance* value, AV, según EP). En el caso del TRA, este valor se obtiene con la **Ecuación IX-2**, mientras que para el PCT se utiliza su forma simplificada (**Ecuación IX-3**).

$$AV = M - \bar{X} + k \cdot s$$
 (IX-2)

$$AV = k \cdot s$$
 (IX-3)

El valor de aceptación depende de \bar{X} , que representa el promedio de los valores individuales $(x_1, x_2, ..., x_n)$ expresados como tanto por ciento del valor nominal. M representa el valor de referencia, s la desviación estándar de la muestra y k la constante de aceptabilidad, que depende del número de muestras a analizar (n). En este caso, k vale 2,4 dado que el valor de n = 10.

El objetivo principal de esta prueba es demostrar la consistencia de cada dosis unitaria. Para cada lote, todas sus unidades deben tener un contenido de principio activo dentro de un rango establecido próximo a su valor nominal.

El criterio establecido para poder determinar si la prueba de uniformidad de contenido es aceptada, se basa en el número de muestras a analizar. Como en este trabajo se han analizado diez comprimidos para cada lote, el límite establecido es $AV \le 15$.

Los resultados de los AV obtenidos se presentan en la **Tabla IX. 8**. Como puede observarse, todos los valores se encuentran dentro del límite establecido por la EP, por lo que, se cumple la uniformidad de contenido.

	HPLC		FOS	SS	Mic	Micro	
	\overline{X}	AV	\overline{X}	AV	\overline{X}	AV	
PCT							
A	101.2	2.3	101.2	1.2	101.3	1.7	
C	100.7	2.0	100.4	2.1	100.8	1.9	
TRA							
A	97.1	4.9	98.2	7.6	95.9	12.2	
C	97.0	4.8	97.1	6.7	95.7	10.7	

Tabla IX. 8. Valores AV obtenidos para los modelos PLS individuales.

4. CONCLUSIONES

El objetivo principal de este trabajo era demostrar las posibilidades de un nuevo instrumento NIR miniaturizado para el análisis cuantitativo en comprimidos farmacéuticos.

Los comprimidos farmacéuticos analizados contienen dos APIs en una proporción de masa 9:1, dificultando así la cuantificación del API minoritario. Con los resultados presentados, se puede concluir que es posible cuantificar ambos APIs con este nuevo instrumento miniaturizado.

Las características del instrumente miniaturizado, como el cubrir solo el intervalo de longitudes de onda de los armónicos, no es un factor limitante, dado que hay suficiente información química para la cuantificación. La menor resolución del instrumento miniaturizado, en comparación al instrumento de sobremesa, hace que sea crítica la selección de un pretratamiento correcto y un intervalo de longitudes de onda adecuado.

La precisión de los resultados para ambos instrumentos se encuentra dentro del límite establecido. La diferencia entre ambos instrumentos es menor que la diferencia entre APIs. Como era de esperar, dada la mayor proporción de PCT, su %RSD es menor que para TRA.

En los resultados del TRA, los valores de aceptación de la uniformidad de contenido son mayores que para el PCT, ya que su proporción en el comprimido es mucho menor y la desviación estándar de su concentración mayor. Además, estos valores también son mayores para el instrumento miniaturizado. Sin embargo, todos los valores calculados a partir de los modelos NIR cumplen con el criterio.

Los resultados muestran que los instrumentos miniaturizados pueden utilizarse para la cuantificación de APIs en comprimidos, aún en situaciones en las que la relación de concentraciones sea muy desproporcionada, y para el cálculo de los parámetros exigidos por las autoridades reguladoras, siendo así una alternativa más económica a los instrumentos de mesa.

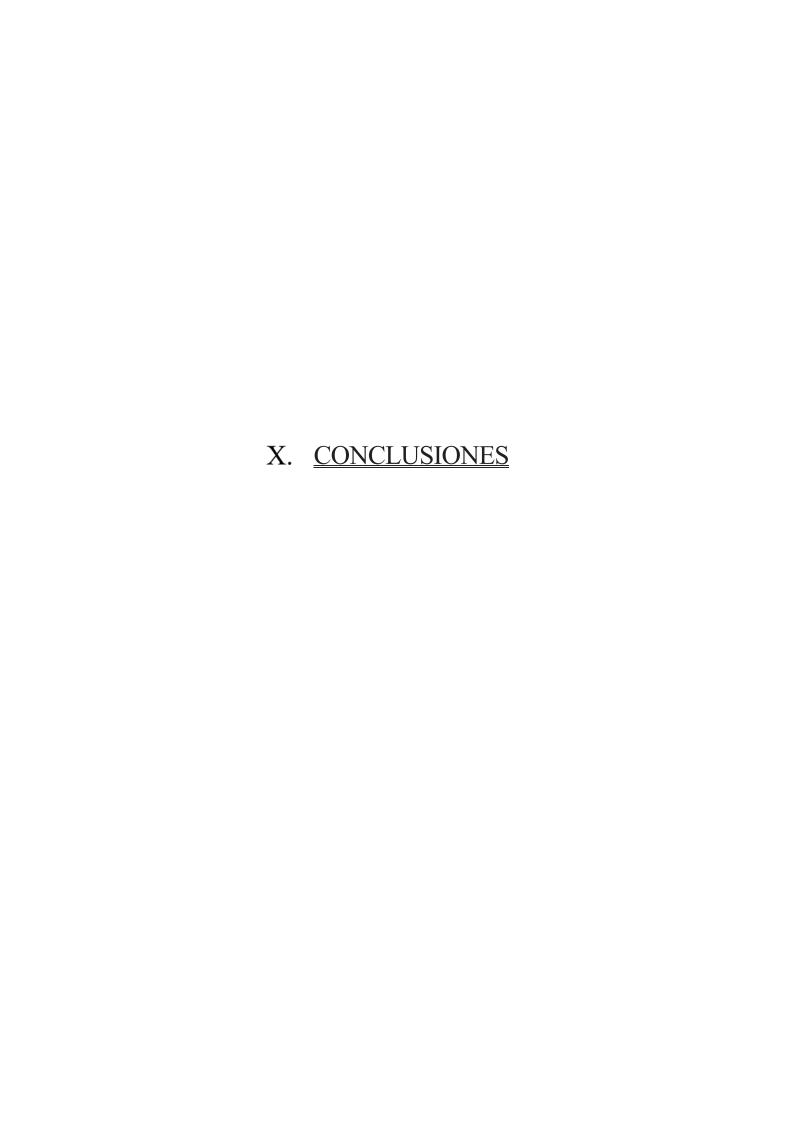
5. BIBLIOGRAFÍA

- [1] J. Luypaert, D.L. Massart, Y. Vander Heyden, *Near-infrared spectroscopy applications in pharmaceutical analysis*, Talanta. 72 (**2007**) 865–883. https://doi.org/10.1016/j.talanta.2006.12.023.
- [2] E.W. Ciurczak, B. Igne, Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy, **2014**.
- [3] C. Pasquini, Near infrared spectroscopy: A mature analytical technique with new perspectives A review, Anal. Chim. Acta. 1026 (2018) 8–36.

- https://doi.org/10.1016/j.aca.2018.04.004.
- [4] K.B. Beć, J. Grabska, C.W. Huck, *Principles and Applications of Miniaturized Near-Infrared (NIR) Spectrometers*, Chem. A Eur. J. 27 (2021) 1514–1532. https://doi.org/10.1002/chem.202002838.
- [5] R.A. Crocombe, *Portable Spectroscopy*, Appl. Spectrosc. 72 (2018) 1701–1751.
 https://doi.org/10.1177/0003702818809719.
- [6] H. Yan, H.W. Siesler, Hand-held near-infrared spectrometers: State-of-the-art instrumentation and practical applications, NIR News. 29 (2018) 8–12. https://doi.org/10.1177/0960336018796391.
- [7] R.M. Correia, F. Tosato, E. Domingos, R.R.T. Rodrigues, L.F.M. Aquino, P.R. Filgueiras, V. Lacerda, W. Romão, *Portable near infrared spectroscopy applied to quality control of Brazilian coffee*, Talanta. 176 (2018) 59–68. https://doi.org/10.1016/j.talanta.2017.08.009.
- [8] S. Modroño, A. Soldado, A. Martínez-Fernández, B. de la Roza-Delgado, *Handheld NIRS* sensors for routine compound feed quality control: Real time analysis and field monitoring, Talanta. 162 (2017) 597–603. https://doi.org/10.1016/j.talanta.2016.10.075.
- [9] C. Malegori, E.J. Nascimento Marques, S.T. de Freitas, M.F. Pimentel, C. Pasquini, E. Casiraghi, *Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the evaluation of acerola fruit quality, using PLS and SVM regression algorithms*, Talanta. 165 (2017) 112–116. https://doi.org/10.1016/j.talanta.2016.12.035.
- [10] M. Alcalà, M. Blanco, D. Moyano, N. Broad, N. O'Brien, D. Friedrich, F. Pfeifer, H. Siesler, Qualitative and quantitative pharmaceutical analysis with a novel handheld miniature near-infrared spectrometer, J. Near Infrared Spectrosc. 21 (2013) 445. https://doi.org/10.1255/jnirs.1084.
- [11] L. Sun, C. Hsiung, C.G. Pederson, P. Zou, V. Smith, M. Von Gunten, N.A. O'Brien, Pharmaceutical Raw Material Identification Using Miniature Near-Infrared (MicroNIR) Spectroscopy and Supervised Pattern Recognition Using Support Vector Machine, Appl. Spectrosc. 70 (2016) 816–825. https://doi.org/10.1177/0003702816638281.
- [12] R.M. Correia, E. Domingos, F. Tosato, N.A. dos Santos, J. de A. Leite, M. da Silva, M.C.A. Marcelo, R.S. Ortiz, P.R. Filgueiras, W. Romão, *Portable near infrared spectroscopy applied to abuse drugs and medicine analyses*, Anal. Methods. 10 (2018) 593–603. https://doi.org/10.1039/C7AY02998E.
- [13] T. Casian, A. Reznek, A.L. Vonica-Gligor, J. Van Renterghem, T. De Beer, I. Tomuţă,

- Development, validation and comparison of near infrared and Raman spectroscopic methods for fast characterization of tablets with amlodipine and valsartan, Talanta. 167 (2017) 333–343. https://doi.org/10.1016/j.talanta.2017.01.092.
- [14] L. Martínez, A. Peinado, L. Liesum, *In-line quantification of two active ingredients in a batch blending process by near-infrared spectroscopy: Influence of physical presentation of the sample*, Int. J. Pharm. 451 (**2013**) 67–75. https://doi.org/10.1016/J.IJPHARM.2013.04.078.
- [15] A. Hertrampf, R.M. Sousa, J.C. Menezes, T. Herdling, *Semi-quantitative prediction of a multiple API solid dosage form with a combination of vibrational spectroscopy methods*, J. Pharm. Biomed. Anal. 124 (**2016**) 246–253. https://doi.org/10.1016/j.jpba.2016.03.003.
- [16] D. Pan, G. Crull, S. Yin, J. Grosso, *Low level drug product API form analysis Avalide tablet NIR quantitative method development and robustness challenges*, J. Pharm. Biomed. Anal. 89 (2014) 268–275. https://doi.org/10.1016/j.jpba.2013.11.011.
- [17] J. Arruabarrena, J. Coello, S. Maspoch, *Enhancing sensitivity and precision on nir reflectance determination of an api at low concentration: Application to an hormonal preparation*, J. Pharm. Biomed. Anal. 60 (**2012**) 59–64. https://doi.org/10.1016/j.jpba.2011.10.026.
- [18] A.L. Vonica-Gligor, T. Casian, A. Reznek, I. Tomuta, F. Gligor, *Simultaneous quantification of atorvastatin and amlodipine in powder blends for tableting by NIR spectroscopy and chemometry*, Farmacia. 63 (**2015**) 381–387.
- [19] Y. Lin, W. Li, J. Xu, P. Boulas, Development of a NIR-based blend uniformity method for a drug product containing multiple structurally similar actives by using the quality by design principles, Int. J. Pharm. 488 (2015) 120–126. https://doi.org/10.1016/j.ijpharm.2015.04.025.
- [20] M. Blanco, M. Alcalà, Simultaneous quantitation of five active principles in a pharmaceutical preparation: Development and validation of a near infrared spectroscopic method, Eur. J. Pharm. Sci. 27 (2006) 280–286. https://doi.org/10.1016/j.ejps.2005.10.008.
- [21] C. Pino-Torres, S. Maspoch, R. Castillo-Felices, M. Pérez-Rivera, M. Aranda-Bustos, C. Peña-Farfal, *Evaluation of NIR and Raman spectroscopies for the quality analytical control of a solid pharmaceutical formulation with three active ingredients.*, Microchem. J. 154 (2020) 104576. https://doi.org/10.1016/j.microc.2019.104576.
- [22] H. Potschka, E. Friderichs, W. Löscher, Anticonvulsant and proconvulsant effects of

- tramadol, its enantiomers and its M1 metabolite in the rat kindling model of epilepsy, Br. J. Pharmacol. 131 (2000) 203–212. https://doi.org/10.1038/sj.bjp.0703562.
- [23] V. Arjunan, R. Santhanam, M.K. Marchewka, S. Mohan, Comprehensive quantum chemical and spectroscopic (FTIR, FT-Raman, 1H, 13C NMR) investigations of O-desmethyltramadol hydrochloride an active metabolite in tramadol An analgesic drug, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 122 (2014) 315–330. https://doi.org/10.1016/j.saa.2013.11.040.
- [24] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, C. De La Pezuela, Strategies for constructing the calibration set in the determination of active principles in pharmaceuticals by near infrared diffuse reflectance spectrometry, Analyst. 122 (1997) 761–765. https://doi.org/10.1039/a700630f.
- [25] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, N. Pou, Influence of the procedure used to prepare the calibration sample set on the performance of near infrared spectroscopy in quantitative pharmaceutical analyses, Analyst. 126 (2001) 1129–1134. https://doi.org/10.1039/b102090k.
- [26] P. Munson, D. Rodbard, Statistical comparison of two assay methods (weighted linear regression with errors in x and y), in: Qual. Cntrol Clin. Endocrinol., **1979**: pp. 89–98.
- [27] W.A. Sadler, *Joint parameter confidence regions improve the power of parametric regression in method-comparison studies*, Accredit. Qual. Assur. 15 (**2010**) 547–554. https://doi.org/10.1007/s00769-010-0674-9.
- [28] ICH, ICH Q2 (R2) Validation of Analytical Procedures, (2022).
- [29] EMA, Guideline on the use of Near Infrared Spectroscopy (NIRS) by the Pharmaceutical Industry and the Data requirements for new Submissions and Variations, Eur. Med. Agency. 44 (2014) 1–28.
- [30] Chapter 2.9.40. Uniformity of dosage units, Eur. Phamacopoeia 8.0. (2013) 357–359.



1. CONCLUSIONES

Las distintas aplicaciones de las técnicas NIR y Raman presentadas en este trabajo ofrecen una alternativa para el análisis de identidad y cuantificación de los productos farmacéuticos.

Considerando los objetivos definidos, se concluye de forma general:

- Se ha desarrollado y validado un método de identificación para distintos materiales de partida recibidos en una industria farmacéutica mediante la técnica NIR. Concretamente un método para la identificación de los colorantes (Biblioteca Específica, 38 productos), y otro para el resto de los materiales (Biblioteca General, 398 productos). La estrategia empleada ha sido la construcción de una librería de espectros y el cálculo de la distancia en el espacio de los componentes principales para la separación y asignación de los grupos. La validación se ha realizado con un conjunto de espectros
- Se ha presentado un análisis de riesgo de la biblioteca desarrollada y validada con el fin de identificar los factores y las variables que tienen un mayor peso en la obtención de un resultado OOS. Tras la identificación, se ha calculado el riesgo asociado a cada producto incluido en la biblioteca de espectros. Se ha identificado un total de 13 colorantes y 78 producto con riesgo alto a obtener un fallo en su identificación. Para la Biblioteca Específica, la acción mitigadora ha sido una revalidación tras un periodo de tres meses, para incluir la mayorar variabilidad posible, y la optimización de los parámetros de clasificación. En el caso de la Biblioteca General, de los productos con riesgo alto, la mayoría son EXCs. Así, como acción mitigadora del riesgo, se ha dividido de la Biblioteca en dos, una para APIs y otra para EXCs. Tras la revalidación de las bibliotecas, se ha observado una disminución en el número de resultados OOS que indica la mejora de la robustez de estas y la efectividad del análisis de riesgos.
- Se ha realizado una comparación entre distintos métodos de identificación para discriminar entre sietes marcas comerciales del mismo producto farmacéutico, tras la obtención de los datos con dos instrumentos portátiles NIR y Raman. Todos los métodos han sido desarrollados y validados satisfactoriamente.
- Se ha desarrollo un método de cuantificación univariante para la determinación del contenido del API en muestras de gel mediante la técnica de adición estándar. Los

espectros han sido adquiridos mediante un equipo Raman portátil equipado con una fibra que estaba en contacto con la muestra de gel. Para el desarrollo experimental, se han optimizado las condiciones hidrodinámicas para asegurar los tiempos de adición del patrón y las condiciones de homogeneización de la muestra. Los resultados muestran la viabilidad de este procedimiento para la obtención del contenido del API (5% m/m) en las muestras.

- Se ha realizado una comparativa entre un instrumento de mesa y uno portátil NIR para la cuantificación de dos APIs en comprimidos farmacéuticos mediante un modelo PLS. Los modelos construidos han sido desarrollados y validados para este propósito. Los resultados muestran que el API en menor contenido necesita de un número mayor de factores en el instrumento portátil, pero, que los resultados son comparables a los obtenidos con el equipo de sobremesa.
- Se ha demostrado la viabilidad del uso de la técnica Raman para la determinación de la uniformidad de contenido del API en comprimidos. Se ha usado como señal analítica la intensidad de un pico característico del mismo. Con el fin de optimizar el procedimiento para una posible aplicación en rutina, se ha estudiado el número mínimo de puntos de mapeo necesarios para obtener una imagen química fiable de la distribución del API.

Con todo esto, se ha demostrado el potencial de estas técnicas para su implementación en la industria farmacéutica junto con la ayuda de la Quimiometria.