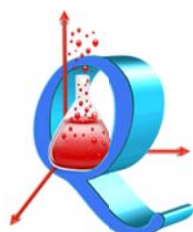


Màster Universitari Ciència i Tecnologia Químiques

Iniciació a la recerca i treball fi de màster

Avaluació d'un espectrofotòmetre NIR en miniatura pel seu ús en processos industrials i controls de qualitat



Daniel Moyano Girbau

Director: Dr. Manel Alcalà

Departament de Química, Facultat de Ciències

3 de Setembre de 2012

Memòria presentada per superar el mòdul d'Iniciació a la Recerca i treball fi de màster corresponent al Màster Universitari Ciència i tecnologia químiques.

Alumne

Daniel Moyano Girbau

Vist i plau

Dr. Manel Alcalà Bernàrdez

Bellaterra, 3 de Setembre de 2012

Agraiments

Índex

Abstract	i
1.- Introducció	1
1.1.- Espectroscòpia d'Infraroig proper	1
1.2.- Quimiometria	5
1.2.1.- Disseny d'Experiments	6
1.2.2.- Tractaments Espectrals	6
1.2.3.- Anàlisi Qualitativa i Quantitativa	7
2.- Objectius	10
3.- Part Experimental	10
4.- Resultats i Discussió	13
4.1.- Anàlisis dels Nivells de Senyal i Soroll	13
4.2.- Identificació de Matèries Primeres	16
4.3.- Quantificació d'API en una Formulació Farmacèutica	19
5.- Conclusions	29
6.- Referències	30

Abstract

Avui en dia, per assegurar la qualitat dels productes, s'han de realitzar moltes determinacions al llarg dels processos de manufactura, des de matèries primes fins als productes finals, passant pels intermedis. Amb l'avanç de la tecnologia, la instrumentació cada cop es torna més reduïda, la qual té prestacions similars als equips convencionals i sol ser més econòmica que els estàndards. Fer servir en la indústria instrumentació portàtil permet reduir els costos globals de producció sense afectar la qualitat dels resultats obtinguts.

El treball que es presenta avalua la idoneïtat d'un nou espectrofotòmetre NIR en miniatura per ser emprat com a eina d'anàlisi de productes farmacèutics. L'avaluació es presenta en tres tipus d'assaigs: (1) anàlisi i comparació dels nivells de senyal i soroll de la resposta instrumental, (2) la identificació de matèries primes mitjançant la construcció de biblioteques espectrals i (3) la quantificació d'ingredient actiu en una formulació farmacèutica mitjançant models de calibratge multivariables. Els resultats obtinguts amb aquest nou instrument són comparables amb els obtinguts fent servir aparells convencionals, demostrant així que aquesta nova tecnologia suposa un clar avenç en el control global de processos farmacèutics.

1.- Introducció

El major interès del Grup de Quimiometria Aplicada es centra en la espectroscòpia molecular, emprant tècniques com l'ultraviolat/visible (UV/Vis), l'infraroig amb transformada de Fourier (FTIR), l'infraroig proper (NIR), el dicroisme circular, la fluorescència molecular, l'espectrometria de mobilitat iònica (IMS) i l'espectroscòpia Raman. Els temes d'estudi del grup són l'aplicació d'aquestes tècniques al desenvolupament de mètodes ràpids per al control analític en molts camps industrials (tèxtil, petroquímic, alimentari, etc.) amb especial atenció a l'anàlisi farmacèutica amb NIR aplicant tècniques de calibratge multivariable [1]. En el treball que es presenta l'Espectroscòpia a l'Infraroig Proper (NIRS) s'emprarà en diferents anàlisis de productes farmacèutics conforme als objectius que es cerquen.

1.1.- Espectroscòpia a l'Infraroig Proper

L'infraroig proper va ser la primera regió no visible descoberta en la primera dècada del segle XIX per William Herschel. No obstant això, abans de la Segona Guerra Mundial la regió NIR no era considerada útil per l'espectroscòpia perquè es va observar que les bandes es solapen i són difícils d'interpretar [2], per aquest motiu la regió NIR va ser considerada de poc potencial fins a principis de la dècada de 1960. Amb el desenvolupament de l'electrònica i la informàtica, i una relativa simplicitat dels espectrofotòmetres, es van desenvolupar nombroses aplicacions basades en espectres NIR durant els darrers 30 anys.

En la regió NIR es detecten les vibracions d'estirament i compressió dels enllaços. Mentre que a l'infraroig mitjà (MIR) els fotons tenen energies inferiors corresponents a les vibracions moleculars fonamentals, el NIR dona informació sobre els primers, segons i tercers sobretons moleculars i combinacions d'aquests que tenen lloc a la regió MIR (veure *Figura 1*). Per absorbir a la regió infraroja (IR) les molècules han de presentar un dipol elèctric i la fluctuació del moment dipolar provoca l'absorció d'un fotó incident. L'origen de l'espectre NIR és la vibració inharmònica, quan l'oscil·lació entre la compressió i l'estirament d'un enllaç covalent entre àtoms és asimètric (és més fàcil estirar l'enllaç que comprimir-lo). La vibració inharmònica permet a les molècules absorbir fotons en el rang NIR, el que resulta ser els "sobretons" de la freqüència de vibració fonamental. En la

Figura 2 es pot veure l'oscil·lador inharmònic i les transicions que s'observen. Els sobretons o harmònics són una sèrie d'absorpcions múltiples (no necessàriament nombres enters) de la freqüència fonamental de vibració; les bandes de combinació tenen lloc quan dues o més freqüències fonamentals comparteixen la mateixa energia. Els grups que contenen hidrogen són els més inharmònics i les molècules biològiques són riques en enllaços $-CH$, $-OH$, $-SH$, i $-NH$ [3].

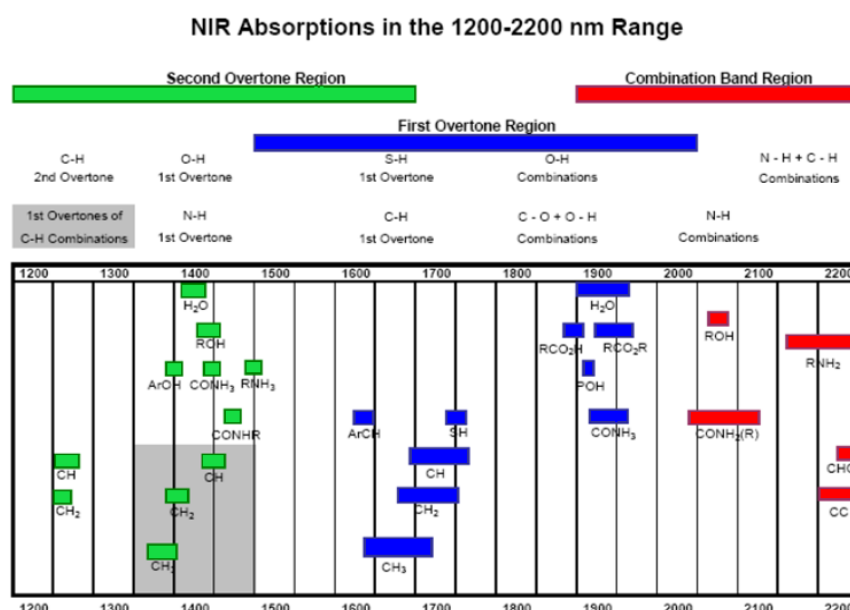


Figura 1.- Assignació de bandes en la regió espectral NIR dels grups que contenen hidrogen.

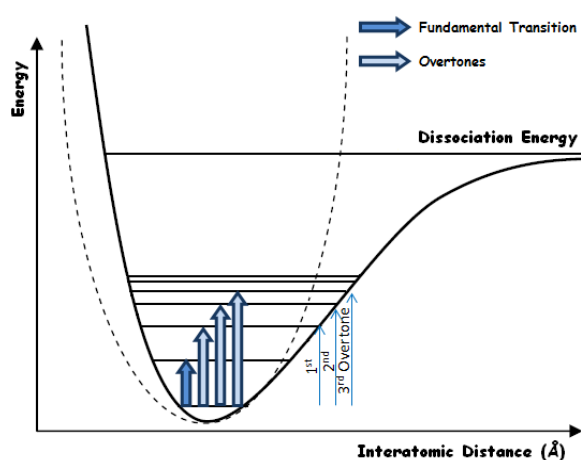


Figura 2.- Energia potencial d'un oscil·lador harmònic (línia discontinua) i la d'un inharmònic (línia continua). També es pot observar la transició entre diferents nivells energètics.

Tots els mètodes espectroscòpics, a excepció de l'espectrometria de masses, es poden classificar en funció de la energia involucrada durant la mesura, per tant les regions espectrals es poden definir en funció de la longitud d'ona. La regió infraroja (observeu la *Figura 3*) és aquella que va des de 780 fins 10^6 nm i la principal causa d'absorció són els moviments vibratoris i rotacionals (els rotacionals es troben, generalment, superposats a les bandes de vibració). La regió IR es pot dividir en tres parts: regió NIR (780 – 2.500 nm o $12.500 - 4.000 \text{ cm}^{-1}$), regió MIR (2.500 – 25.000 nm o $4.000 - 400 \text{ cm}^{-1}$) i regió FIR (25.000 – 10^6 nm) [4].

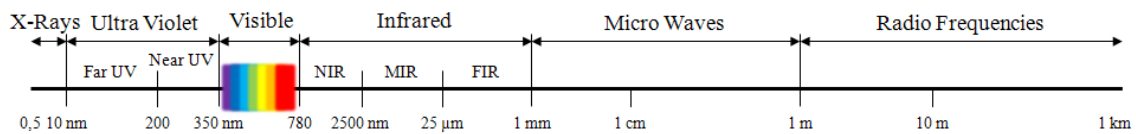


Figura 3.- Espectre electromagnètic amb les regions espectrals d'interès.

En el NIR, la feblesa dels sobretons permet a l'investigador mesurar mostres intactes i sense diluir, mitjançant diferents modes de registre que es mostren a la *Figura 4*. Hi ha tres modes de registre en NIRS: transmitància, reflectància i transfectància.

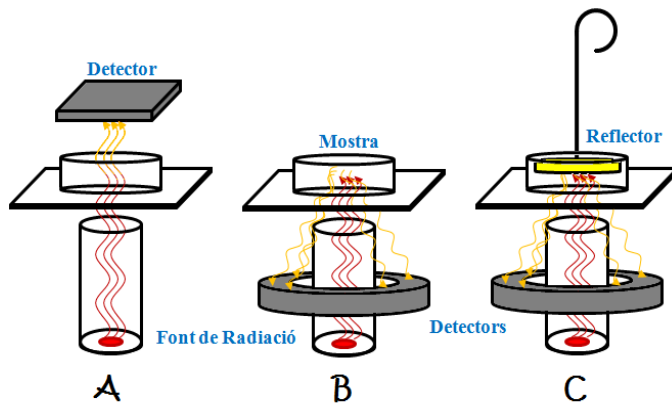


Figura 4.- Diferents modes de registre en tècniques NIR; (A) mode de transmitància, (B) mode de reflectància i (C) mode de transfectància.

En la primera forma de registre la radiació absorbida segueix la llei de Lambert-Beer, essent la transmitància la fracció de radiació transmesa per la mostra. En la configuració de reflectància, la mostra reflexa la radiació i la retorna de dues formes diferents: especular o difusament. La reflectància especular és descrita per les lleis de Fresnel i no proporciona

cap informació de la composició de la mostra; la reflectància difosa té lloc en totes direccions a conseqüència dels processos d'absorció i dispersió. En el mode de transflectància la radiació travessa la mostra però es reflexada per un reflector, creuant un altre cop la mostra i es dirigeix cap al detector situat al voltant de la font de radiació com en el mode de reflectància [5].

En els darrers anys, l'espectroscòpia NIR ha estat emprada cada cop més com una eina analítica en diversos camps i ha substituït els mètodes tradicionals. La principal raó d'aquesta estesa acceptació són la rapidesa amb la que les mostres es poden caracteritzar sense manipular-les i la flexibilitat dels equips NIR. No obstant, algunes característiques de la tècnica restringeixen aplicacions més àmplies o impedeixen usos particulars. Els principals avantatges i inconvenients de l'espectroscòpia NIR com a eina analítica es mostren a continuació [6].

Avantatges:

- + Tècnica no invasiva i no destructiva
- + Requereix mínima o cap preparació de la mostra
- + Mesures i obtenció de resultats ràpids
- + Sense necessitat de reactius o materials per preparar les mostres
- + Un únic espectre permet la determinació simultània de diversos analits
- + És possible determinar paràmetres físics com densitat, viscositat o mida de partícula
- + La instrumentació és adient per ser emprada en plantes de producció
- + Les fibres òptiques proporcionen sensors robustos i no invasius per anàlisis de control de processos *at-line*, *on-line* i *in-line*
- + Els resultats són comparables en precisió als d'altres tècniques analítiques

Inconvenients

- Les bandes són molt àmplies i solapades amb el què s'han de fer servir tècniques quimiomètriques per modelar les dades de les quals extreure informació rellevant
- El procés de calibratge per fer anàlisis qualitatives o quantitatives és complex, però es poden obtenir mètodes analítics robustos per ser utilitzats inclòs durant anys

- Cal introduir als models de calibratge tota aquella variabilitat que hom es pugui trobar durant l'anàlisi rutinària
- El paràmetre objectiu s'ha de determinar prèviament emprant mètodes de referència
- Aplicable als components majoritaris però en certes aplicacions es pot reduir el límit de quantificació fins a les ppm
- La construcció de models NIR requereix una inversió substancial que es compensa per la utilització d'aquests durant anys sense recalibratge.

1.2.- Quimiometria

El terme quimiometria va ser encunyat fa moltes dècades per descriure una nova manera d'analitzar dades químiques, en les quals es combinen tant conceptes químics com estadístics. Des de llavors, la quimiometria s'ha convertit en un camp tècnic i està creixent ràpidament en popularitat dins d'una àmplia gamma de disciplines químiques. Definicions d'aquest terme n'hi ha moltes [7-9], tant mateix, hi ha tres conceptes que apareixen constantment en les aplicacions, essent aquests: (1) modelatge empíric, (2) modelatge multivariant i (3) dades químiques. Tenint present aquests conceptes es pot proposar una definició per quimiometria: l'aplicació de mètodes de modelatge empíric i multivariable a les dades químiques.

La part de modelatge empíric indica una èmfasis en les dades que es generen i no pas en la teoria del modelatge d'aquestes. Respecte a l'element multivariable, aquest indica que per la construcció del model es fa servir més d'una variable obtinguda amb l'analitzador. La combinació dels termes de modelatge empíric i multivariant fa de la quimiometria una eina molt poderosa al poder modelar sistemes que són complexos i desconeguts. El modelatge s'ha de realitzar procurant de no sobreajustar els models o aquests el descriuran de forma errònia [10].

D'eines quimiomètriques n'hi ha moltes i molt diverses. A continuació es citaran i s'explicaran aquelles eines emprades en aquest treball, essent el disseny d'experiments, els pretractaments espectrals i alguna tècnica qualitativa i quantitativa.

1.2.1.- Disseny d'Experiments

El disseny d'experiments (DoE) és el conjunt de mètodes emprats per planejar la feina experimental per tal d'obtenir la màxima informació rellevant, de la millor qualitat i amb el menor nombre d'experiments. Els resultats d'un experiment són funció d'una o més variables (anomenades factors) que l'experimentador pot controlar. Amb la variació de les condicions experimentals es pot emprar el disseny experimental amb la finalitat de conèixer com influeixen els factors al resultat i quina combinació proporciona els millors resultats. En aquest treball s'ha aplicat un disseny Factorial Complet per estudiar la relació que existeix entre la configuració de l'equip en miniatura i la qualitat dels espectres obtinguts.

1.2.2.- Pretractaments Espectrals

Els pretractaments són correccions matemàtiques aplicades a les dades espectrals per corregir el soroll spectral o augmentar la resolució spectral. Hi ha molts pretractament, però en aquest treball es mencionaran els fets servir.

- ✓ Promig d'espectres: calcula l'espectre promig a partir dels replicats disminuint el soroll spectral aleatori
- ✓ Suavitzat: fent servir filtres basats en càlculs polinòmics el soroll spectral es pot minimitzar aplicant aquest pretractament. Els tipus de suavitzats que s'han testat són *Moving Average*, *Median Filter* i *Gaussian Filter*
- ✓ Variable Normal Estàndard (*Standard Normal Variate*, SNV): corregeix l'efecte de dispersió de la radiació produït, per exemple, per la diferència en la mida de partícula. Cada absorbància es resta al valor d'absorbància mitjà de l'espectre i es divideix per la desviació estàndard, centrant l'espectre al valor zero i amb una desviació estàndard igual a u
- ✓ Derivades: suprimeix la variació en la línia base i incrementa la diferència entre dos bandes àmplies i solapades. Una primera derivada elimina el desplaçament de la línia base mentre que la segona derivada treu els termes que depenen de la longitud d'ona [11]. A la *Figura 5* hi ha un exemple dels efectes de les derivades. El tipus de derivades provades són *Gap-Segment* i *Norris Gap*

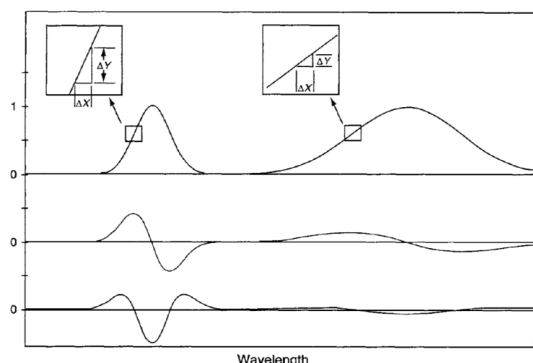


Figura 5.- Bandes gaussianes en absorbància (a dalt) i les seves corresponents primera (al mig) i segona (a baix) derivada [12].

1.2.3.- Anàlisi Qualitativa i Quantitativa

La tècnica NIR s'il·lustra en totes les farmacopees per ser emprada tant en anàlisis qualitatives com quantitatives i és una de les tècniques amb les millors perspectives en l'anàlisi de productes farmacèutics al ser ràpida, no destructiva i no necessitar preparació de la mostra. Les anàlisis qualitatives es realitzen mitjançant la construcció de biblioteques espectrals mentre que les anàlisis quantitatives consisteixen en el càlcul de models de calibratge multivariables.

En les biblioteques espectrals es realitzen identifikacions de productes comparant els espectres registrats amb els espectres de totes les substàncies presents en la biblioteca. Els passos a seguir per a la seva construcció són: (1) registre dels espectres NIR, (2) tria d'espectres, (3) tria del mètode de reconeixement de pautes, pretractament dels espectres i construcció de la biblioteca amb validació interna, (4) construcció de sub-biblioteques si s'escau i, finalment, (5) validació externa [13].

Un dels mètodes emprats en la construcció de biblioteques espectrals és l'ús del coeficient de correlació (r). Aquest es defineix com el cosinus de l'angle entre el vector de l'espectre de la mostra i el vector de l'espectre promig del producte a ser inclòs en la biblioteca. La fórmula matemàtica és:

$$r_{jk} = \frac{\sum_{i=1}^p (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^p (x_{ij} - \bar{x}_j)^2} \sqrt{\sum_{i=1}^p (x_{ik} - \bar{x}_k)^2}} \quad \text{Equació 1}$$

On p és el nombre de longituds d'ona; els subíndex k i j denoten la mostra i la referència respectivament; x_i és el valor mesurat a la longitud d'ona i ; $\bar{x}_j$ és l'espectre promig del producte de referència j ; i $\bar{x}_k$ és l'espectre promig de la mostra.

Si el valor r entre els espectres de la substància desconeguda i el del producte contingut en la biblioteca supera un valor llindar establert (normalment s'estableix un llindar de 0,97), llavors els dos espectres comparats es consideren idèntics i la mostra s'identifica amb el producte de referència. Teòricament, el r de dos espectres de la mateixa mostra hauria de ser la unitat, però degut al soroll aleatori associat a les mesures experimentals fan que aquest valor no sigui 1 exactament. Aquest paràmetre té l'avantatge de ser independent de la mida de la biblioteca i de canvis en la concentració, el que permet identifications correctes en biblioteques amb pocs espectres. D'altra banda, r es calcula en els espectres en segona derivada, fent que les mostres del mateix producte però amb mida de partícula diferent tinguin la mateixa correlació [14].

En els models de calibratge multivariants calculats en les anàlisis qualitatives es pretén calcular la quantitat d'una substància present en la mostra a partir dels espectres d'aquesta. El procediment general per calcular-los es pot veure en la Figura 6.

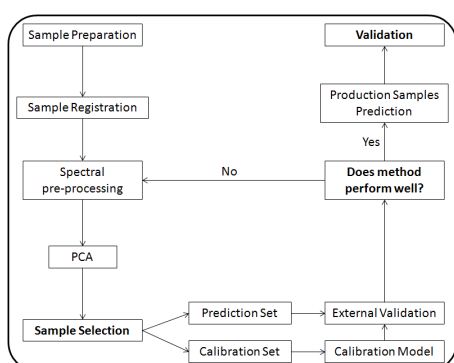


Figura 6.- Procediment general per la construcció de models de calibratge [15]. Hom també es pot referir a la validació externa com a predicció.

El terme PCA correspon a l'Anàlisi en Components Principals. Aquest és un mètode d'anàlisi de dades per reduir la dimensió de dades complexes. Els models lineals multivariants per PCA es desenvolupen emprant vectors ortogonals que s'anomenen Components Principals (PC). Una de les fites més significants del PCA és l'eliminació dels PC associats al soroll, reduint així la dimensió del sistema i minimitzant els efectes de l'error en la mesura. La tasca del PCA és la descomposició de la matriu de dades espectrals en diversos termes segons l'equació següent:

$$X = t_1 p_1^T + t_2 p_2^T + \dots + t_k p_k^T + \varepsilon \quad \text{Equació 2}$$

X correspon a la matriu original de dades, en el cas del NIR els espectres de totes les mostres; t_n és un vector anomenat *score* d'una fila i tantes columnes com mostres hi ha en el sistema, p_n^T és un vector anomenat *loading*, el qual està transposat i d'aquí el superíndex T , d'una columna i tantes files com longituds d'ona. El producte $t \cdot p$ s'anomena factor i se'n poden calcular tants com fonts de variació hi hagi en el sistema. Finalment, ε és la matriu residual que conté la informació que els altres factors no expliquen i que s'associa a l'error experimental [16]. Els factors amb més informació són els primers mentre que els últims no n'aporten gaire.

La representació de dos *scores* genera un gràfic en dues dimensions on cada mostra es representa per un punt. Aquest gràfic permet veure la similitud o diferència de les mostres segons aquestes estiguin agrupades o disperses.

El model de calibratge es realitza emprant l'algoritme de Mínim Quadrats Parcial (Partial Least Squares o PLS). Aquest es presenta sovint com la major tècnica de regressió per sistemes de dades multivariades. Una característica important del PLS és que té en compte que els espectres i les concentracions estimades contenen errors, a diferència d'altres algorismes que suposen que les concentracions estan lliures d'error. Aquest algorisme descompon les matrius segons es mostra en les Equacions 3 i 4.

$$X = T \cdot P + E \quad \text{Equació 3}$$

$$Y = U \cdot Q + F \quad \text{Equació 4}$$

X correspon a la matriu de dades espectrals mentre que Y és la matriu de concentració, T i U són els *scores*, P i Q són els *loadings* per les respectives matrius, E i F són les matrius de la variabilitat no explicada.

Com es pot observar, l'element comú en totes dues equacions són els *scores*. La descomposició d'ambdues matrius es realitza simultàniament i s'estableixen relacions lineals entre les dues matrius. D'aquesta manera s'obtenen factors PLS en X que millor prediran el valor en c tenint en compte els dos conjunts de variables [17].

Un cop calculat el model de calibratge, aquest es pot aplicar a les mostres que no s'hagin emprat en el càlcul del model per tal de validar-lo. El darrer pas de l'anàlisi qualitativa és aplicar el model a mostres de producció per verificar que la concentració que es prediu no difereix a la real.

2.- Objectius

La majoria dels treballs de camp en anàlisis químiques es realitza prenent la mostra i transportant-la a un instrument situat a distància. La demora entre la presa de mostra i la recepció dels resultats no només limita el nombre de mostres que es poden analitzar en un cert període sinó que a més pot afectar greument les oportunitats d'influir en el resultat d'una operació del procés. A més, la preparació de les mostres no permet l'obtenció de resultats en temps real. Una solució a aquests problemes és el desenvolupament d'una instrumentació de camp portàtil capaç de realitzar anàlisis *at-line* o *in-line* [18].

Hi ha molts exemples en la literatura d'instrumentació portàtil, i els espectrofotòmetres NIR no són pas una excepció. Aquest estudi té l'objectiu d'avaluar la idoneïtat d'un nou espectrofotòmetre NIR en miniatura per ser utilitzat com a eina en anàlisis qualitatives i quantitatives de fàrmacs. Aquesta avaluació s'ha dut a terme realitzant tres tipus de proves:

- ✓ Anàlisi i comparació dels nivells de senyal i soroll de la resposta instrumental
- ✓ Identificació de matèries primes
- ✓ Quantificació d'ingredient actiu en una formulació farmacèutica

Els resultats obtinguts permetran determinar si és factible utilitzar aquest nou espectrofotòmetre en diferents camps d'estudi i les limitacions de l'aparell.

3.- Part Experimental

Espectroscòpia NIR

La companyia nord-americana *JDS Uniphase Corporation* ha desenvolupat un espectrofotòmetre NIR ultra compacte, lleuger i de baix cost. Aquest instrument té completament integrats les fonts de llum, els components òptics i electrònics així com el software bàsic d'operació i s'ofereix tot en un paquet [19]. Aquest nou espectrofotòmetre es mostra en la *Figura 7* i a la *Taula 1* hi ha algunes de les especificacions de l'equip.

Taula 1.- Paràmetres i especificacions de l'espectrofotòmetre NIR en miniatura.

Paràmetre	Especificacions
Font d'il·luminació	Dos làmpades al buit de tungstè integrades
Tipus de detector	Arrenglament de 128 fotodíodes d'InGaAs
Element dispersiu	Filtre Variable Lineal
Rang de longituds d'ona	Des de 1.110 a 2.136 nm
Resolució	8 nm aprox.
Dimensions (diàmetre × alçada)	45 × 42 mm
Pes	< 60 g
Alimentació requerida	Alimentació per USB (< 500 mA, 5 V)



Figura 7.- Imatge de l'espectrofotòmetre NIR en miniatura desenvolupat per JDSU.

El mode de registre emprat en totes les anàlisis va ser per reflectància, obtenint espectres en aquesta forma de registre. Per al tractament de les dades l'espectre s'ha de transformar en absorbància aplicant següent la fórmula:

$$A_{\lambda} = \log \left(\frac{1}{R_{\lambda}} \right) = -\log(R_{\lambda}) \quad \text{Equació 5}$$

On A_{λ} és l'absorbància a la longitud d'ona λ determinada i R_{λ} és el valor de reflectància a la λ especificada.

El procediment que es segueix per registrar qualsevol mostra és el que es s'observa a la *Figura 8*. Aquest consisteix en definir els paràmetres de registre, registrar el 100% emprant una referència de ceràmica blanca del 99% de reflectància inclosa amb l'equip i finalment registrar la mostra per triplicat. Un cop es tenen les dades s'aplica l'*Equació 5* i es realitza la mitjana dels replicats per tal d'obtenir l'espectre promig.

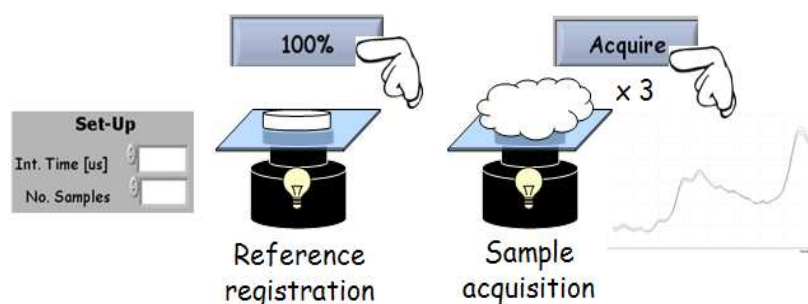


Figura 8.- Procediment experimental general de registre de les mostres emprant el NIR en miniatura. La distància que s'ha de deixar entre la superfície de la mostra i la finestra del detector és de 3 mm.

Un altre equip utilitzat per comparar els resultats obtinguts ha estat l'equip de sobretaula disponible al laboratori, un espectrofotòmetre FOSS acoblat a un mòdul RCA (*Rapid Container Analyzer*). Amb aquest equip es registren 32 *scans* per mostra entre 1.100 i 2.500 nm amb una resolució de 2 nm.

Anàlisi i comparació dels nivells de senyal i soroll de la resposta instrumental

Aquest estudi té la finalitat de trobar els millors paràmetres de registre, essent el nombre de *scans* a realitzar i el temps de integració. Aquesta anàlisi es va realitzar emprant una referència blanca amb un 99% de reflectància, dels espectres de la qual es van poder calcular el valor promig de l'espectre, el valor promig de la desviació estàndard i el temps d'adquisició. El disseny d'experiments plantejat es va plantejar amb el programa informàtic MODDE 8.0 (Unmetrics). Del mateix programa es va fer servir una de les opcions per al càlcul de superfícies de resposta.

Identificació de matèries primes

En aquesta anàlisi es van registrar diferents matèries primeres entre ingredients farmacèutics actius (API) i excipients. El registre es va fer a través d'un tros de plàstic puig que els productes farmacèutics es solen emmagatzemar i transportar en bosses de plàstic tancades i és interessant identificar-los a través d'aquest material per facilitar el registre i no alterar en cap cas el producte. Els programes informàtics que s'han fet servir són

l'Unscrambler 9.8 (Camo) per tractar les dades espectrals i l'Excel (Microsoft) per a la construcció d'una biblioteca espectral.

Quantificació d'ingredient actiu en una formulació farmacèutica

En el darrer estudi realitzat es va aplicar un disseny d'experiments per tal de preparar 24 mostres sintètiques, les qual amplien el rang de concentració del API en un $\pm 20\%$ del valor nominal de la formulació. Aquestes es van preparar al laboratori pesant amb balança analítica diferents quantitats de placebo i d'ingredient actiu i barrejant-les per obtenir les mostres. Aquestes es van registrar a través d'una cubeta de quars. Els models de calibratge per quantificar el ingredient actiu es van calcular emprant el programa informàtic Unscrambler 9.8.

4.- Resultats i Discussió

4.1.- Anàlisi i Comparació dels nivells de senyal i soroll de la resposta instrumental

Quan es fa servir l'espectrofotòmetre NIR en miniatura s'han de definir dos paràmetres abans del registre. Aquests són el temps d'integració (temps en μs que l'espectrofotòmetre triga en registrar un espectre) i el nombre de *scans* (nombre d'espectres per cada mostra). Per trobar quina és la combinació dels paràmetres amb millors resultats espectrals s'ha realitzat un DoE amb un disseny 3^2 Factorial Complet, el qual genera 9 experiments a realitzar. Aquest tipus de disseny s'aplica quan hom vol estudiar un nombre petit de factors (important que el nombre de factors sigui petit perquè el nombre d'experiments a realitzar creix exponencialment) i permet escollir els experiments més adequats per treure el màxim d'informació útil. Per a la anàlisi es van donar tres nivells als dos factors d'estudi; per al temps d'integració van ser 100, 550 i 1000 μs i per al nombre de *scans* 1.000, 10.500 i 20.000.

Combinant aquestes possibilitats es van fer 9 mesures a la referència blanca del 99% de reflectància emprant-la com a blanc i mostra, obtenint espectres que es poden considerar com els espectres del fons o soroll espectral. Amb aquests espectres es poden calcular dos paràmetres: el valor promig de l'espectre i el valor promig de la desviació estàndard (SD). El primer (Valor Promig) consisteix en sumar els valors d'absorbància de l'espectre promig i dividir pel nombre de longituds d'ones, obtenint un valor d'absorbància promig de tot l'espectre. El segon (Promig SD) consisteix en calcular la SD del triplicat per cada longitud

d'ona, i de tots els valors de SD calcular-ne el promig. A part d'aquestes dues respostes hi ha una tercera que també es considera, el Temps d'Adquisició i consisteix en el temps que triga en aparèixer l'espectre en la pantalla de l'ordinador un cop començada l'adquisició. A la *Taula 2* es poden observar els 9 experiments realitzats, quines són les condicions de registre i els valors numèrics de les respostes estudiades.

Taula 2.- Valors dels factors i les respostes en l'anàlisi realitzada. Els diferents experiments es van fer en ordre aleatori.

Exp. #	Factors		Respostes		
	Temps d'Integració (µs)	Nombre de Scans	Valor Promig	Promig SD	Temps d'Adquisició (s)
1	100	1.000	- 0,0038	0,0003	1,5
2	550	1.000	0,0003	0,0002	2
3	1.000	1.000	- 0,0002	0,0002	3
4	100	10.500	0,0015	0,0003	9
5	550	10.500	- 0,0020	0,0008	15
6	1.000	10.500	0,0001	0,0006	19
7	100	20.000	0,0004	0,0002	18
8	550	20.000	- 0,0008	0,0004	27
9	1.000	20.000	- 0,0023	0,0007	35

En totes tres respostes es desitja un valor el més proper a zero. En el cas del valor promig, que aquest valor sigui pròxim a zero significa que no hi ha desplaçament de l'espectre registrat. Un valor de promig SD baix indica que hi ha poca diferència entre els replicats d'una mostra i que els espectres són propers entre ells. Finalment, un temps d'adquisició baix permet fer més anàlisis.

Observant els valors de la taula hom pot observar que els millors resultats s'obtenen en els experiments 2, 3 i 7, tot i que el darrer té un temps d'adquisició més alt. Per tal de trobar les condicions òptimes es fa un estudi de les superfícies de resposta i amb l'opció *Optimizer* del programa emprat es poden trobar els valors exactes per als dos factors. A la *Figura 9* es pot observar el resultat de les superfícies de resposta.

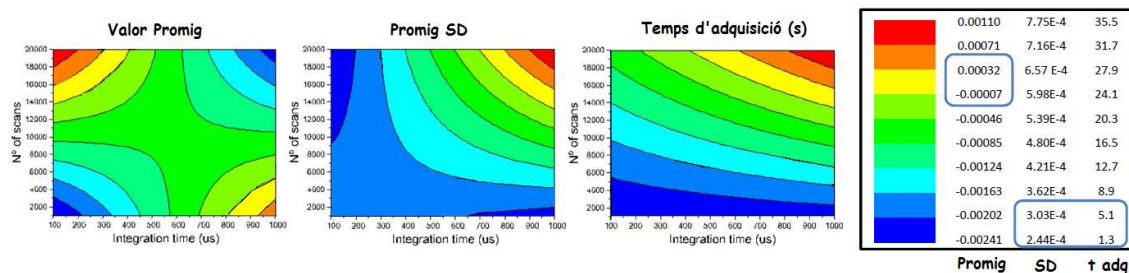


Figura 9.- Superfícies de resposta per les tres variables en estudi. A la dreta hi ha la llegenda pels tres gràfics.

Per al valor promig, la combinació de factors amb resultats propers a zero són les zones de color groc, corresponents a temps d'integració alts i baix nombre de *scans* o bé temps d'integració baix i nombre de *scans* alt. Respecte al valor promig, la combinació amb respostes properes a zero és aquella amb àrea de color blava, corresponents a les mateixes condicions que per al valor promig. Pel què fa al temps d'integració, es pot dir en general que a major nombre de *scans* major és el temps d'adquisició, amb el que les condicions més favorables són aquelles amb aquest factor a nivell baix.

El programa proposa com a condicions òptimes un temps d'integració de 865 μs i 1.000 *scans*. Aquests valors en les superfícies de respostes cauen en les zones que s'han considerat que donen els resultats desitjats. Aquestes són les condicions considerades òptimes i seran els valors que s'introduiran en la resta d'espectres que es registraran (el desenvolupament de la biblioteca espectral i el càlcul del model de calibratge per a API).

La part final d'aquesta anàlisi va consistir en la comparació dels resultats obtinguts amb l'espectrofotòmetre JDSU i l'equip disponible al laboratori, un equip FOSS amb el mòdul RCA. La *Taula 3* mostra els resultats obtinguts al analitzar la mateixa referència blanca del 99% de reflectància.

Taula 3.- Comparació de resultats entre el nou equip i el disponible al laboratori.

		NIR JDSU	NIR FOSS-RCA
Factors	Temps d'Integració (μ s)	865	---
	Nombre de Scans	1.000	32
Respostes	Valor Promig	- 0,0002	0,00004
	Promig SD	0,0002	0,00003
	Temps d'Adquisició (s)	2,8	20

Es pot dir que l'aparell del laboratori NIR de FOSS té un soroll espectral inferior puig que el valor promig i el promig SD són inferiors als del equip JDSU. Tot i això, si hom es fixa en la magnitud dels paràmetres es pot dir que tots dos equips són exactes i precisos. Un detall important és que mentre que el NIR JDSU realitza més *scans* que el FOSS, el primer adquireix les dades amb major rapidesa que no pas el segon. Una possible explicació d'aquest fet pot ser la diferència en el rang espectral i que mentre que l'equip JDSU adquireix dades cada 8 nm (obtenint 128 valors d'absorbància), l'aparell FOSS pren dades cada 2 nm (i genera 700 variables per espectre).

4.2.- Identificació de matèries primes

Per moltes indústries, analitzar els productes entrants és imprescindible per assegurar la identitat i la qualitat d'ells. El procés del control d'identificació significa que s'han de prendre decisions basades principalment en conclusions estadístiques [20]. En les farmacopees, la identificació es defineix com la confirmació d'una certa identitat química. Emprant la tècnica NIR, la identificació es basa en la comparació de les dades espectrals d'una substància amb les dades espectrals de moltes altres mostres i lots diferents d'aquestes presents en la biblioteca espectral [21].

La biblioteca construïda està formada per 23 substàncies diferents i s'inclouen dos lots per cada una (a excepció d'una única mostra que només es disposa d'un lot), amb el que la biblioteca estarà formada per 45 mostres. Un cop enregistrades totes les mostres es van aplicar a les dades espectrals diferents pretractaments. Aquestes correccions van ser diferents suavitzats per tal de corregir el soroll i després s'hi aplicà una segona derivada

que és el pretractament aplicat al construir una biblioteca. Es van provar dos suavitzats i dos tipus de derivades, i a aquell que es va considerar el millor es van fer uns canvis en els rangs d'aplicació del pretractament per veure si hi havia milloria. Aquella combinació de pretractaments que menys falsos positius tingués i que els valors del coeficient de correlació entre mostres iguals fossin els més alts va ser la que es va triar per la biblioteca.

A la *Taula 4* es poden observar els valors de r de cada mostra amb totes les altres. Els pretractaments aplicats són un suavitzat de tipus *Moving Average* (amb *segment size* de 3) i una segona derivada de tipus *Gap-Segment* (amb *gap* i *segment* de 3). Si s'analitza la *Taula 4*, el primer que s'observa és que només una meitat té valors. Això és degut al fet que el coeficient de correlació compara els espectres de dues mostres i és el mateix comparar una en front de l'altre com al revés. Si hom es fixa en la diagonal de la taula pot veure que tots els valors són la unitat. Això és degut a què l'espectre d'una mostra comparat amb el mateix espectre és completament igual, per tant els valors de r són 1 en tota la diagonal de la taula. Els valors que es troben emmarcats en la diagonal corresponen als r de la mateixa mostra però són de dos lots diferents. Degut a la variabilitat de cada fabricant entre lots diferent, els espectres dels dos lots de la mateixa substància no tenen perquè ser completament igual, però aquests no poden diferir gaire, motiu pel qual aquests r han de ser superiors al llindar establert.

Els valors que fins ara s'han esmentat es poden considerar correctes i és desitjable que siguin així sempre. Pel què fa a la resta de resultats, l'ideal és que els valors siguin inferiors al llindar establert, fet que confirma que són dues substàncies diferents. Malgrat això, en la biblioteca que s'ha construït hi ha dos parells de mostres amb falsos positius (quan una mostra A és identificada com A però també com B). Aquestes parelles són el mannitol amb el mannitol granulat i la sacarosa en pols amb la sacarosa en pols a la sílice. La diferència entre mannitols és la mida de partícula i, al aplicar la segona derivada, les diferències espectrals es redueixen, motiu pel qual es confonen i donen falsos positius. Pel què fa a les sacaroses, la diferència rau en que una és pura i l'altre té un baix contingut en sílice. Per poder resoldre aquests falsos positius caldria fer una sub-biblioteca basada en espai de PCA amb un poder discriminant més gran.

Taula 4.- Valors dels coeficients de correlació de les 45 mostres incloses a la biblioteca espectral. Les dades estan tractades amb un suavitzat de tipus *Moving Average* (amb *segment size* de 3) i una segona derivada de tipus *Gap-Segment* (amb *gap* i *segment* de 3).

[illegible]

Per finalitzar aquesta anàlisi es va realitzar una comparació entre la biblioteca calculada amb l'equip JDSU i el FOSS-RCA. Els espectres obtinguts amb el FOSS es van tractar amb una segona derivada de Savitzky-Golay amb 11 punts de finestra mòbil. La *Taula 5* mostra els valors de r per les mostres amb falsos positius en les biblioteques.

Taula 5.- Comparació de les identifikacions negatives entre l'equip JDSU i el FOSS.

	NIR JDSU	NIR FOSS-RCA
Mannitol - Mannitol granulat	0,99	1,00
Sacarosa pols - Sacarosa pols sílice	1,00	1,00
Diestearat glicerol - Monoestearat sorbitan	0,96	1,00

Tots dos equips identifiquen correctament totes les substàncies amb la diferència que en el JDSU hi ha dos falsos positius mentre que en l'equip FOSS n'hi ha tres. El fals positiu entre el diestearat de glicerol i el monoestearat de sorbitan rau en què totes dues substàncies comparteixen el mateix grup estearat i la causa que només es considerin iguals només amb el FOSS es deu a què aquest equip té una resolució més baixa, fet que fa que les bandes siguin més definides.

4.3.- Quantificació d'ingredient actiu en una formulació farmacèutica

En la quantificació de productes farmacèutics, les guies reguladores com la guia ICH, recomanen fer servir mostres ampliant el rang de concentració del API en un $\pm 20\%$ del valor nominal de la formulació (no hi ha un rang establert per als excipients) [22]. Hi ha diverses formes de realitzar l'ampliació del rang i en aquest estudi s'ha triat fer servir mostres de laboratori sintètiques preparades pesant quantitats conegudes de API i excipient per expandir el rang de concentració del API. A la *Taula 6* hi ha el percentatge en pes de cada component de la formulació farmacèutica que s'estudiarà.

Taula 6.- Components en la formulació farmacèutica i els corresponents percentatges en pes de cada un (expressats en % m/m).

Component	% m/m
API	9
Excipient 1	57,5
Excipient 2	32
Excipient 3	1
Excipient 4	0,5

Les mostres per a l'anàlisi es van preparar seguint un disseny binari, barrejant el API amb un placebo (mescla de tots els excipients). El problema d'aquest disseny és que hi ha una correlació molt alta entre la quantitat de API i de placebo; quan la concentració de API creix la de placebo disminueix. Per trencar aquesta correlació es van preparar 5 placebos variant un $\pm 5\%$ relatiu la proporció de cada excipient.

Les mostres es van preparar pesant els diferents productes purs amb una balança analítica, generant 24 mostres amb concentració de API creixent i de 10 g cada una. Les concentracions de cada component estan compreses aproximadament: entre 6,9 i 10,4% m/m per al API, entre 55,4 i 59,4% m/m per l'excipient 1, entre 30,4 i 33,6% m/m per l'excipient 2, entre 1,2 i 1,3% m/m per l'excipient 3 i, finalment, entre 0,48 i 0,53% m/m per l'excipient 4. Les mostres es van agitar manualment per tal d'homogeneïtzar el contingut en tot el contenidor. Es va enregistrar l'espectre de totes les mostres, així com el dels productes purs, col·locant els sòlids en una cubeta de quars. La *Figura 10* mostra els espectres de les substàncies purs així com la mostra amb la composició més propera a la de la formulació.

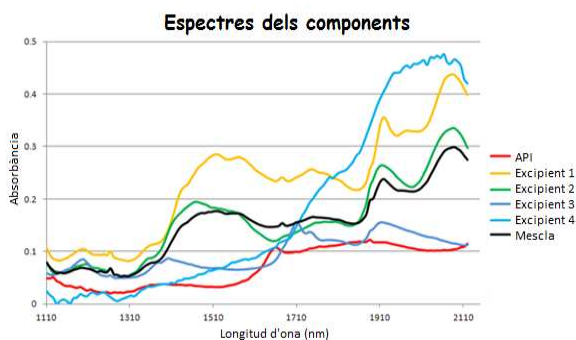


Figura 10.- Espectres dels components purs de la formulació farmacèutica i el de una mostra amb concentracions similars a les de la composició esmentada.

A la *Figura 11* es poden observar els espectres de les 24 mostres sintètiques preparades al laboratori un cop aplicats els pretractaments. Els pretractaments aplicats són un suavitzat de tipus *Gaussian Filter* (amb segment de 3) i una segona derivada de tipus *Gap-Segment* (amb gap i segment de 3).

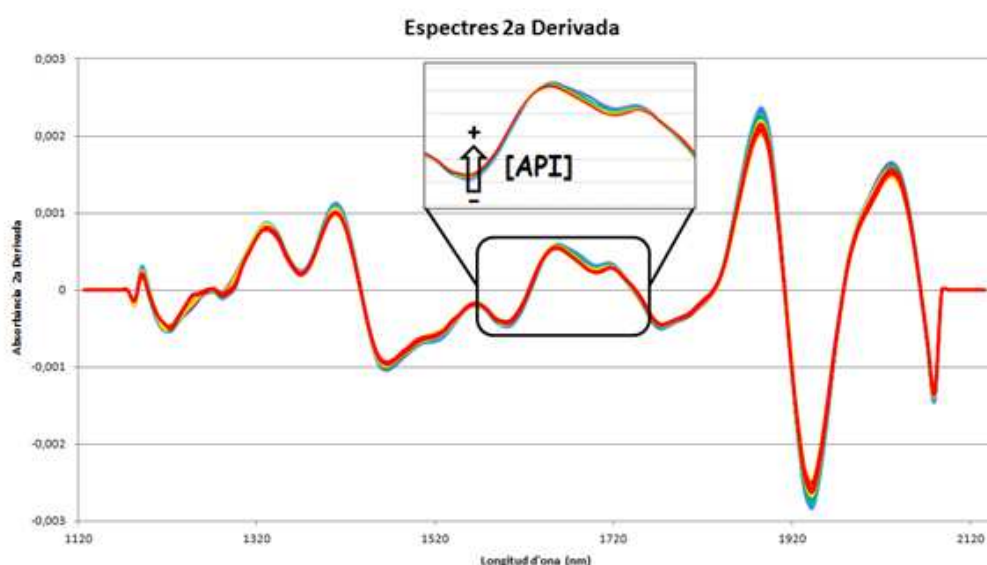


Figura 11.- Espectres suavitzats i derivats de les mostres en el model. Els colors freds indiquen concentracions baixes de API mentre que colors càlids concentracions altes. El rang ampliat és la zona on el API té una absortivitat diferent als altres components.

Als espectres de les 24 mostres preparades al laboratori se'ls hi va aplicar diferents pretractaments. Aquests van ser el SNV, els suavitzats de tipus *Moving Average*, *Median Filter* i *Gaussian Filter* i les derivades de primer i segon ordre de tipus *Gap-Segment* i *Norris Gap*, provant-los per separat o combinant-los. Com ja s'ha esmentat, els calibratges per tècniques NIR són totalment empírics i és necessari calcular diversos models per tal de trobar aquell més adequat per les dades obtingudes. Tot seguit es llisten els paràmetres que determinen si un model és adequat a les dades o no.

- ✓ Nombre de factors: indica en quants termes s'ha descompost la matriu de dades espectrals per al càlcul del calibratge i és desitjable que sigui petit

- ✓ Percentatge de varianza Y explicada en calibratge (% Y-ECV): és el percentatge de similitud a cada factor entre les dades originals i les obtingudes al descompondre-les; es cerca que aquest valor superi el 95% en el factor triat
- ✓ Correlació NIR vs. Ref.: indica quan similars són les dades calculades amb el NIR i les reals o de referència
- ✓ Test t per mostres aparellades: és una prova estadística per comprovar si el residual de les mostres de predicció és estadísticament igual a zero
- ✓ RMSEC i RMSEP: corresponen a l'error quadràtic mitjà pel calibratge i la predicció respectivament; l'expressió matemàtica és la mateixa per tots dos i aquesta correspon a l'Equació 6. Aquest valor indica la diferència entre el valor calculat amb el model i el valor de referència, és un indicatiu de la magnitud dels residuals
- ✓ RSEC i RSEP: error estàndard relatiu pel calibratge i la predicció respectivament; l'expressió matemàtica és la mateixa per tots dos i correspon a l'Equació 7. Indica la similitud entre els valors reals i els calculats, com menor sigui menor serà l'error de la recta doncs més similars són els valors obtinguts per la regressió i els reals

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (Y_i^{NIR} - Y_i^{Ref})^2}{n}} \quad \text{Equació 6, [23]}$$

$$RSEC = 100 \cdot \sqrt{\frac{\sum_{i=1}^n (Y_i^{NIR} - Y_i^{Ref})^2}{\sum_{i=1}^n (Y_i^{Ref})^2}} \quad \text{Equació 7, [24]}$$

On Y_i^{NIR} és el valor calculat pel model de calibratge, Y_i^{Ref} és la concentració obtinguda per pesada. La diferència d'aquests dos valors és el que es coneix com a residual. Per últim, n és el nombre de mostres emprades en el calibratge.

A continuació es mostren alguns models calculats abans de trobar aquell que millor s'ajusta a les dades. La *Taula 7* mostra els resultats del calibratge i la *Taula 8* són els resultats de predicció aplicant el model calculat. En la segona taula, totes les columnes són valors obtinguts emprant les mostres sintètiques preparades al laboratori a excepció de la darrera columna, predicció de producció, els valors de la qual es van obtenir al aplicar el model de calibratge a dues mostres de producció.

Taula 7.- Paràmetres estadístics obtinguts per al calibratge de diversos models calculats aplicant diferents pretractaments.

Pretractaments espectrals	Nombre factors	% Variança Y explicada	RMSEC (%m/m)	Correlació NIR vs. Ref.
Suavitzat <i>Gaussian Filter</i> + SNV	1	74,624	0,553	0,864
	2	87,264	0,392	0,934
	3	96,109	0,216	0,980
	4	97,573	0,171	0,988
	5	98,504	0,134	0,992
Suavitzat <i>Gaussian Filter</i> + SNV + 2a Derivada <i>Gap-Segment</i>	1	79,877	0,492	0,894
	2	87,721	0,384	0,937
	3	98,129	0,150	0,991
	4	98,582	0,131	0,993
	5	99,223	0,097	0,996
1a Derivada <i>Gap-Segment</i>	1	86,531	0,403	0,930
	2	92,557	0,299	0,962
	3	96,928	0,192	0,985
	4	98,086	0,152	0,990
	5	99,307	0,091	0,997
Suavitzat <i>Gaussian Filter</i> + 1a Derivada <i>Gap-Segment</i>	1	86,512	0,403	0,930
	2	92,463	0,301	0,962
	3	96,729	0,198	0,984
	4	97,854	0,161	0,989
	5	98,973	0,111	0,995
2a Derivada <i>Gap-Segment</i>	1	84,693	0,429	0,920
	2	91,497	0,320	0,962
	3	95,046	0,244	0,975
	4	97,984	0,156	0,990
	5	98,418	0,138	0,992
Suavitzat <i>Gaussian Filter</i> + 2a Derivada <i>Gap-Segment</i>	1	84,595	0,431	0,920
	2	91,792	0,314	0,958
	3	95,007	0,245	0,975
	4	97,598	0,170	0,988
	5	98,347	0,141	0,992

Taula 8.- Paràmetres estadístics obtinguts en la validació externa dels corresponents models de calibratge.

Pretractaments espectrals	Nombre factors	Correlació NIR vs. Ref.	RMSEP (%m/m)	Test t mostres aparellades			Predicció Producció	
				t _{crit} = 2,365			Ref = 8,696 % m/m	
				Prom. Res	SD Res	t _{exp}	Pols 1	Pols 2
Suavitzat <i>Gaussian Filter</i> + SNV	1	0,631	0,799	-0,379	0,752	1,427	---	---
	2	0,700	0,705	-0,292	0,686	1,204	---	---
	3	0,858	0,578	-0,312	0,521	1,694	---	---
	4	0,833	0,588	-0,297	0,543	1,546	---	---
	5	0,846	0,580	-0,302	0,529	1,612	---	---
Suavitzat <i>Gaussian Filter</i> + SNV + 2a Derivada Gap-Segment	1	0,684	0,684	-0,194	0,701	0,782	---	---
	2	0,761	0,634	-0,239	0,627	1,077	---	---
	3	0,948	0,325	-0,145	0,310	1,324	---	---
	4	0,957	0,315	-0,156	0,292	1,511	---	---
	5	0,949	0,337	-0,168	0,312	1,518	---	---
1a Derivada Gap-Segment	1	0,817	0,537	0,141	0,554	0,718	---	---
	2	0,889	0,415	0,036	0,441	0,227	---	---
	3	0,952	0,288	0,085	0,294	0,815	---	---
	4	0,966	0,241	0,037	0,254	0,406	---	---
	5	0,939	0,317	-0,067	0,331	0,573	---	---
Suavitzat <i>Gaussian Filter</i> + 1a Derivada Gap-Segment	1	0,818	0,536	0,144	0,552	0,739	9,140	9,336
	2	0,888	0,417	0,037	0,443	0,238	8,413	8,610
	3	0,950	0,292	0,081	0,300	0,762	8,903	8,916
	4	0,964	0,245	0,029	0,260	0,316	9,495	9,565
	5	0,944	0,306	-0,070	0,318	0,618	9,399	9,541
2a Derivada Gap-Segment	1	0,822	0,538	0,171	0,546	0,886	9,689	9,904
	2	0,913	0,381	0,094	0,695	0,677	9,099	9,480
	3	0,958	0,321	0,178	0,286	1,764	8,027	8,205
	4	0,981	0,219	0,133	0,186	2,021	8,618	8,724
	5	0,983	0,222	0,147	0,178	2,342	8,569	8,732
Suavitzat <i>Gaussian Filter</i> + 2a Derivada Gap-Segment	1	0,825	0,535	0,172	0,542	0,895	9,734	9,955
	2	0,916	0,372	0,085	0,387	0,621	9,102	9,478
	3	0,958	0,312	0,162	0,285	1,612	8,113	8,308
	4	0,976	0,228	0,110	0,213	1,454	8,624	8,709
	5	0,984	0,204	0,122	0,175	1,962	8,697	8,854

Cada taula conté un determinat nombre de mostres. Les 24 mostres preparades es van dividir en dos grup, un per la construcció del model (amb un 70% de les mostres) i l'altre per comprovar la capacitat predictiva d'aquest (amb el 30% restant). Així, els valors de calibratge són de 17 mostres i els de predicció en són de 7 mostres.

Els millors resultats es van obtenir amb un suavitzat de tipus *Gaussian Filter* (amb *segment* de 3) i una segona derivada de tipus *Gap-Segment* (amb *gap* i *segment* de 3). Un cop tractades les dades es va realitzar un càlcul de PCA. Amb la representació dels dos primers *scores* s'obté el gràfic de la *Figura 12*. D'aquest gràfic de PCA es poden seleccionar els dos grups de mostra: el de calibratge (punts de color blau) i el grup de validació externa o predicció (punts de color vermell). És important que la selecció de mostres es realitzi tant al llarg del rang de concentració com en la distribució del PCA.

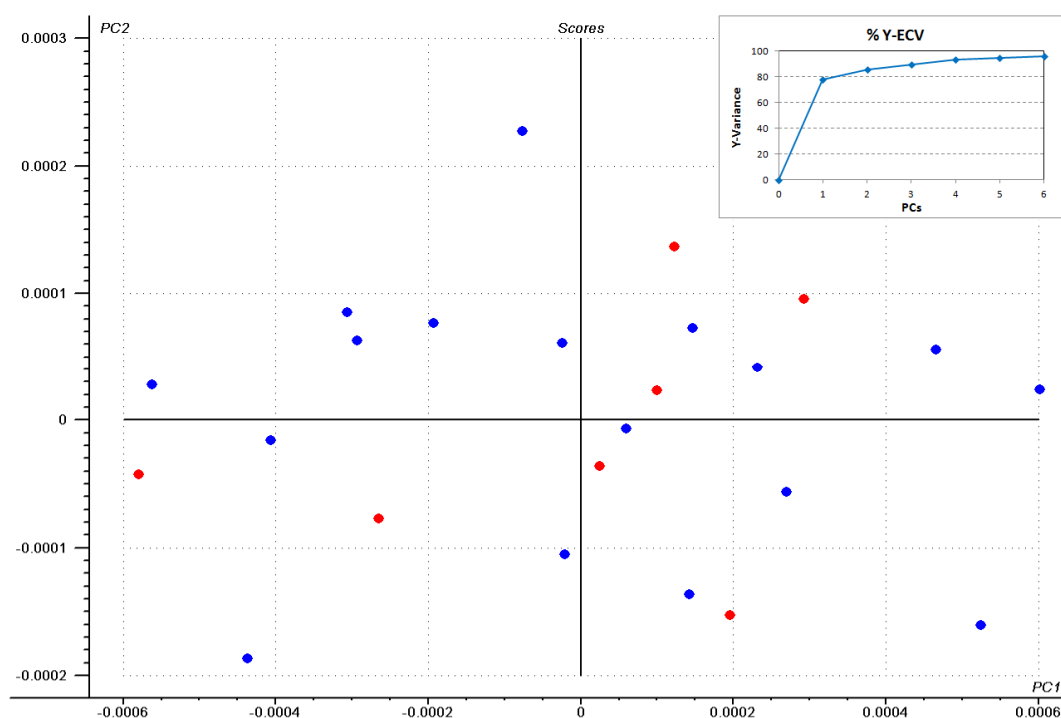


Figura 12.- Distribució de les mostres en l'espai PCA un cop aplicats els pretractaments esmentats. A la part superior dreta hi ha el gràfic que mostra la tendència de la variança explicada en funció dels factors.

Aquest nou espai de PCA conté la informació de totes les longituds d'ona i es representa en un gràfic de dues dimensions, podent observar visualment la variació de les

mostres. El primer PC (eix horitzontal) explica el 78,1% de la variabilitat de les mostres, essent la concentració del API el que la provoca. El segon PC (eix vertical) explica un 7,5% de la variabilitat, amb el què el 85,6% de la variació en els espectres es pot observar en aquesta figura.

Un cop s'ha decidit quines mostres van a calibratge i quines a validació es construeix el model de calibratge emprant l'algoritme PLS. Aquest algoritme calcula uns nous valors de concentració de API per cada mostra i la representació de les concentracions reals amb les noves genera una recta de calibratge tal i com es pot veure a la *Figura 13*.

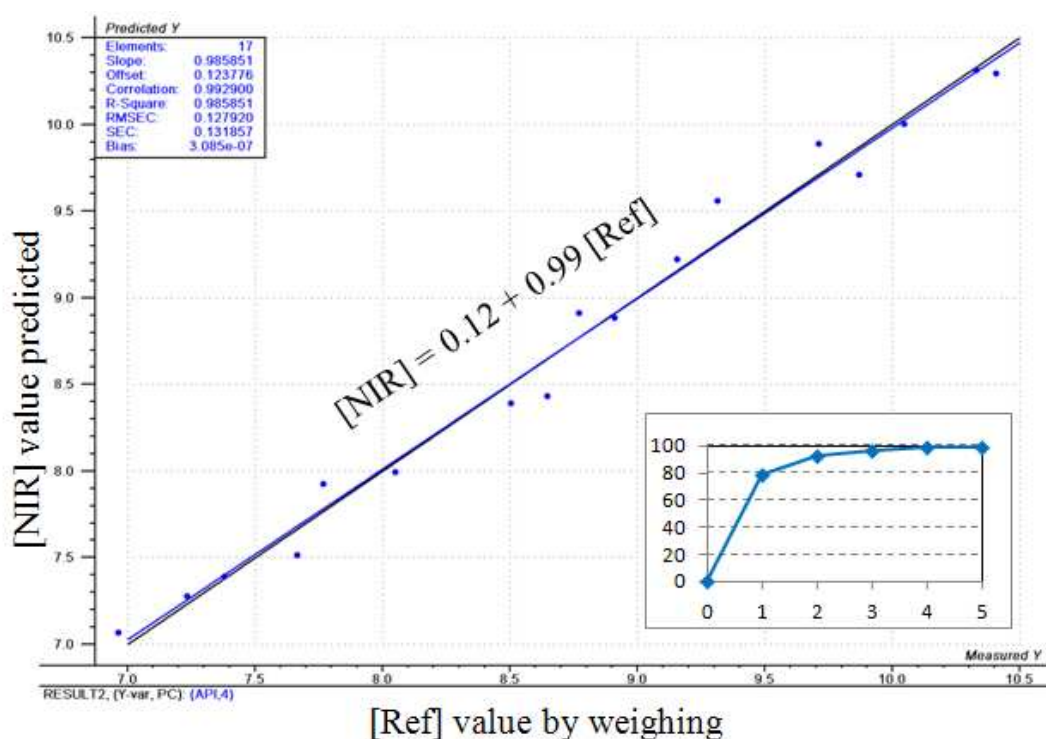


Figura 13.- Model de calibratge obtingut per PLS amb el set de calibratge. El gràfic de la variança explicada enfront els factors PLS es mostra a la part inferior dreta.

Els resultats que es mostren s'obtenen per 4 factors PLS i la variança explicada és del 98,6%. La recta de regressió obtinguda té un pendent de $0,99 \pm 0,06$ i una ordenada a l'origen de $0,1 \pm 0,6$; els intervals de confiança indiquen que la regressió està dintre de les especificacions puig que el valor 1,00 està contingut en la pendent i el 0,0 en l'origen. La

correlació entre la recta de regressió i la recta objectiu ($y = x$) és de 0,993. Els errors de la recta són un RMSEC de 0,128% m/m i un RSEC de 1,451%.

Amb el model de calibratge calculat, aquest es pot aplicar a les mostres que no s'han fet servir per comprovar la capacitat predictiva del model i per validar-lo. Els resultats obtinguts corresponen a la *Figura 14*.

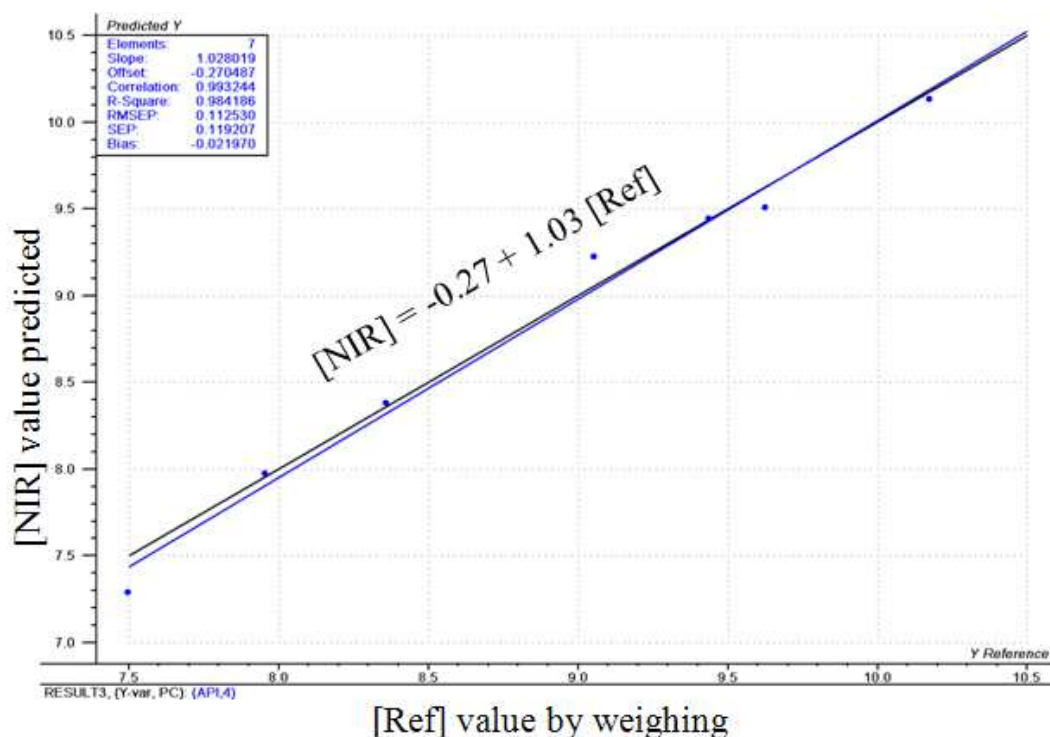


Figura 14.- Predicció aplicada al conjunt de mostres de predicció.

Els resultats de predicció, per al mateix nombre de factors PLS, són una pendent de $1,0 \pm 0,1$ i una ordenada a l'origen de $-0,3 \pm 1,2$. En aquesta ocasió el 1,0 també està comprès en el interval de la pendent i el 0,0 es troba dintre del definit per l'ordenada a l'origen. La correlació té un valor de 0,993 i els errors són un RMSEP de 0,113% m/m i un RSEP de 1,266%.

Els resultats obtinguts tant en el calibratge com en la predicció es troben dintre dels valors acceptables i es considera que aquest és el model que millor s'ajusta a les dades. El darrer pas per donar el model per bo és realitzar la predicció de mostres de producció i comprovar que la concentració de API d'aquestes mostres és propera a la descrita en la

formulació farmacèutica. En la *Taula 9* hi ha la predicció amb aquest model de dos lots diferents de mostres de producció en pols.

Taula 9.- Valors de concentració de API predits de dos mostres de producció.

Ref. = 8,696 % m/m	NIR (% m/m)	Residual (% m/m)
Pols de Producció 1	8,733	0,037
Pols de Producció 2	8,829	0,133

Observant els resultats d'aquesta taula es pot veure que el model de calibratge no només prediu correctament les mostres preparades al laboratori, sinó que a més prediu correctament la concentració de API en dues mostres de producció. Es pot dir aquesta afirmació puig que el valor dels residuals (diferència entre el valor NIR i el de Referència) és de la mateixa magnitud que en la validació externa.

De la mateixa forma que les anteriors anàlisis s'han comparat amb un altre equip, aquest estudi no ha estat una excepció. La *Taula 10* recull els resultats obtinguts amb els dos equips. El pretractament aplicat a les dades de l'espectrofotòmetre FOSS amb el mòdul RCA és una segona derivada de Savitzky-Golay i 11 punts de finestra mòbil.

Taula 10.- Comparació dels resultats quantitius obtinguts amb els dos instruments.

	Espectrofotòmetre JDSU		Espectrofotòmetre FOSS-RCA	
Calibratge	Factors	4	Factors	3
	% Y-ECV	98,6 %	% Y-ECV	99,1 %
	Pendent	0,99 ± 0,06	Pendent	0,99 ± 0,05
	Ordenada Origen	0,1 ± 0,6	Ordenada Origen	0,1 ± 0,4
	Correlació	0,993	Correlació	0,995
	RMSEC	0,128 % m/m	RMSEC	0,105 % m/m
	RSEC	1,451 %	RSEC	1,207 %
Predicció	Pendent	1,0 ± 0,1	Pendent	1,0 ± 0,1
	Ordenada Origen	- 0,3 ± 1,2	Ordenada Origen	0,0 ± 1,1
	Correlació	0,993	Correlació	0,992
	RMSEP	0,113 % m/m	RMSEP	0,122 % m/m
	RSEP	1,266 %	RSEP	1,371 %
	Lots Producció	8,781 % m/m	Lots Producció	8,575 % m/m

Com es pot observar en la taula, els resultats amb tots dos espectrofotòmetres són comparables. Tots dos models estan calculats amb pocs factors i superant en tots dos casos el 95% de la varianza explicada pel calibratge. Mentre que amb l'equip JDSU calen 4 factors PLS per tal d'explicar un 98,6% de Y-ECV, amb l'equip FOSS s'explica un 99,1% d'aquesta amb 3 factors. No hi ha diferències apreciable entre les rectes de calibratge dels dos models. La correlació també és de la mateixa magnitud i els errors en el calibratge són del mateix ordre de magnitud, però inferiors en el equip FOSS.

En el cas de la validació externa, tots dos equips donen paràmetres de la recta de regressió similars al igual que el valor de correlació. Els errors en ambdós casos són del mateix ordre de magnitud, amb la singularitat que ara els valors són més petits en l'equip JDSU. Respecte a la predicció de mostres de producció, tots dos espectrofotòmetres donen resultats propers al valor de referència.

5.- Conclusions

S'ha estudiat la funcionalitat d'un nou espectrofotòmetre NIR desenvolupat per la companyia JDSU i les conclusions que s'extreuen de l'estudi són:

- ✓ S'ha seguit una estratègia per tal de trobar les millors condicions de registre amb el NIR en miniatura. Els valors trobats són un temps d'integració de 865 μ s i 1.000 *scans*. Aquesta combinació de condicions donen nivells de senyal i soroll adequats.
- ✓ S'ha construït una biblioteca espectral compresa per 23 mostres diferents registrades a través de plàstic. El major nombre de positius entre diferents lots de la mateixa mostra i menor nombre de falsos positius es va obtenir aplicant un suavitzat *Gaussian Filter* (amb segment de 3) i una segona derivada de tipus *Gap-Segment* (amb gap i segment de 3). Tot i això, hi ha dos parelles de substàncies que són falsos positius: mannitol amb mannitol granulat i sacarosa en pols amb sacarosa en pols a la sílice.

- ✓ El calibratge per una formulació farmacèutica, a la qual s'ha ampliat el rang de concentració de API en un 20% (entre 6,9 i 10,4% m/m), es va obtenir aplicant sobre les dades un suavitzat de tipus *Moving Average* (segment de 3) i una segona derivada *Gap-Segment* (gap i segment de 3). Els valors de RSEC i RSEP són 1,45% i 1,27% respectivament.
- ✓ Els resultats obtinguts són comparables als obtinguts amb un equip estàndard. Els nivells de senyal i soroll són majors en l'equip JDSU, però són valors acceptables. En quant a la biblioteca espectral, l'equip JDSU té una capacitat de classificació similar a la del FOSS. A nivell quantitatiu, els dos espectrofotòmetres donen bons models, però el FOSS té resultats similars al JDSU amb un factor menys, el qual fa el model més senzill.

6.- Referències

[1] About the Applied Chemometric Research Group, URL <http://grupsderecerca.uab.cat/chemometrics/> (visitada el 25 de Maig de 2012).

[2] Roggo, Y.; Chalus, P.; Maurer, L.; et al.; *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*. J. Pharm. Biomed. Anal. Vol. 44 (2007) 683-700.

[3] Bosco, G.L.; *James L. Waters Symposium 2009 on near-infrared spectroscopy*. Trends Anal. Chem. Vol. 29, No. 3 (2010) 197-208.

[4] Dufour, É. In *Infrared Spectroscopy for Food Quality and Control*, 1st Ed.; Sun, D.; Elsevier (2009) pp 3-27. ISBN: 978-0-12-374136-3.

[5] Alcalà, M. Ph.D. thesis, Universitat Autònoma de Barcelona, Bellaterra (BCN), 2006.

[6] Blanco, M.; Villarroya, I; *NIR spectroscopy: a rapid-response analytical tool*. Trends Anal. Chem. Vol. 21, No. 4 (2002) 240-250.

[7] Mongay, C. *Quimiometria*; Universitat de València (2005) p 19. ISBN 978-84-370-8644-6.

[8] About the North American Chapter of the International Chemometrics Society (NAmICS), URL <http://www.namics.nysaes.cornell.edu/> (visitada el 12 de Juny de 2012).

[9] About the International Union of Pure and Applied Chemistry (IUPAC), URL http://www.iupac.org/nc/home/projects/project-db/project-details.html?tx_wfqbe_pi1%5bproject_nr%5d=2008-002-1-500 (visitada el 12 de Juny de 2012).

[10] Miller, C.E. In *Chemometrics in Process Analytical Chemistry*, 1st Ed.; Bakeev, K.A.; Blackwell Publishing (2005) pp 226-227. ISBN: 978-1-4051-2103-3.

[11] Fearn, T. *NIRnews* (Vol. 19, No 7), pp 16-17.

[12] Mark, H.; Workman, J.Jr. *Chemometrics in Spectroscopy*; Elsevier (2007) pp 51-55 and 340. ISBN: 978-0-12-374024-3.

[13] Blanco, M.; Romero, M.A.; *Near-infrared libraries in the pharmaceutical industry: a solution for identity confirmation*. *Analyst* (Cambridge, U. K.) Vol. 126 (2001) 2212-2217.

[14] Blanco, M.; Coello, J.; Iturriaga, H.; et al.; *Near-infrared spectroscopy in the pharmaceutical industry*. *Analyst* (Cambridge, U. K.) Vol. 123 (1998) 135R-150R.

[15] Alvarenga, L.; Ferreira, D.; Aktekruse, D.; et al.; *Tablet identification using near-infrared spectroscopy (NIRS) for pharmaceutical quality control*. *J. Pharm. Biomed. Anal.* Vol. 48 (2008) 62-69.

[16] Gemperline, P.J. In *Practical Guide to Chemometrics*, 2nd Ed.; Gemperline, P.J.; Taylor & Francis (2006) pp 70-71. ISBN 978-1-57444-783-5.

[17] Brereton, R.G. *Chemometrics. Data Analysis for the Laboratory and Chemical Plant*; John Wiley & Sons (2003) pp 298-300. ISBN 0-471-48977-8.

[18] Lysaght, M.J.; Van Zee, J.A.; Callis, J.B.; *Laptop chemistry: A fiber-optic, field-portable, near-infrared spectrometer*. *Rev. Sci. Instrum.* Vol. 62 (1991) 507-515.

[19] About JDS Uniphase Corporation (JDSU), URL www.jdsu.com/go/mironir (visitada el 21 de Maig de 2012).

[20] Molt, K.; *How safe are NIR-library search systems? Information-theoretical and practical aspects*. *Fresenius' J. Anal. Chem.* Vol. 359 (1997) 67-73.

[21] *Guideline on the use of Near Infrared Spectroscopy (NIRS) by the pharmaceutical industry and the data requirements for new submissions and variations*. European Medicines Agency (EMA), 2012.

[22] Blanco, M.; Peguero, A.; *Analysis of pharmaceuticals by NIR spectroscopy without a reference method*. *Trends Anal. Chem.* Vol. 29, No. 10 (2010) 1127-1136.

[23] Blanco, M.; Cueva-Mestanza, R.; Peguero, A.; *NIR analysis of pharmaceutical samples without reference data: Improving the calibration*. Talanta Vol. 85 (2011) 2218-2225.

[24] Blanco, M.; Bautista, M.; Alcalà, M.; *API Determination by NIR Spectroscopy across Pharmaceutical Production Process*. AAPS PharmSciTech Vol. 9, No. 4 (2008) 1130-1135.