



UNIVERSITAT ROVIRA I VIRGILI

Treball de Fi de Grau

7-6-2019

METODOLOGIA I VALIDACIÓ PER NIR

David Sabaté Garcia



Dra. Ariadna Campos Carrasco
Dra. Maria Besora Bonet
Grau de Química

Contenido

1	RESUM	3
2	ABSTRACT	3
3	OBJECTIU	4
4	INTRODUCCIÓ	4
4.1	DESCOBRIMENT DE LA LLUM INFRAROJA.....	4
4.2	L'ESPECTRE ELECTROMAGNÈTIC	5
4.3	INTERACCIÓ DE LA RADIACIÓ I LA MATERIA	5
4.3.1	QUIN TIPUS DE MOLECULES ABSORBEIX LA LLUM INFRAROJA?	5
4.3.2	SEPARACIÓ DE LES ZONES ESPECTRALS	6
4.4	PARTICULARITATS SOBRE EL NIR	6
4.5	METODOLOGIA AMB EL NIR.....	6
4.5.1	CALIBRACIÓ UNIVARIANT.....	6
4.5.2	CALIBRACIÓ MULTIVARIANT	7
4.5.3	VALIDACIÓ DEL MÈTODE	9
4.6	EMULSOGEN HCO	10
4.7	PER QUÈ UTILITZAR EL NIR?	10
5	PROCEDIMENT EXPERIMENTAL	11
5.1	DIAGRAMA DE GANTT	11
5.2	RECOPILACIÓ DE MOSTRES	11
5.3	REVISIÓ DELS VALORS DELS ANÀLISIS DE REFERÈNCIA.....	11
5.4	RECOPILACIÓ DE DADES:.....	11
5.4.1	MÈTODES DE REFERÈNCIA UTILITZATS	11
5.5	ANALITZAR PER NIR	16
5.6	TRACTAMENT DELS ESPECTRES.....	17
5.7	DESENVOLUPAMENT DEL MÈTODE	17
5.7.1	PART EXPERIMENTAL	17
5.7.2	PART ESTADÍSTICA	19
5.8	VALIDACIÓ DEL MÈTODE	23
5.8.1	ROBUSTESA DEL MÈTODE	24
6	RESULTATS.....	25
6.1	MÈTODES	25
6.1.1	MÈTODE DEL ÍNDEX DE SAPONIFICACIÓ	25
6.1.2	MÈTODE DEL ÍNDEX D'HIDROXILACIÓ	29
6.1.3	MÈTODE DEL PUNT DE TERBOLESA	33
6.2	ROBUSTESA DE YOUTDEN-STEINER.....	37

6.2.1	ROBUSTESA EN L'ÍNDIX DE SAPONIFICACIÓ	38
6.2.2	ROBUSTESA EN L'ÍNDIX DE HIDROXILACIÓ	40
6.2.3	ROBUSTESA EN EL PUNT DE TERBOLESA.....	41
7	CONCLUSIONS.....	44
8	CONCLUSIONS.....	45
9	BIBLIOGRAFIA.....	46
	ANNEX.....	47

1 RESUM

Aquest treball es basa en el desenvolupament i posterior validació de tres mètodes que determinin per al producte Emulsogen HCO 040 i 060 el seu Índex de saponificació, el seu Índex d'hidroxilació i el seu punt de terbolesa, fent servir un model multivariant a partir de l'absorbància en freqüències d'infraroig proper (NIR). Per realitzar tot això, es farà una petita introducció als fonament teòrics de l'espectroscopia, fins a com es realitza un model multivariant, passant pels mètodes de referència utilitzats i tot el que cal saber per escollir el millor mètode desenvolupat i avaluar si presenta una certa validesa.

2 ABSTRACT

This project studies the development and validation of three methods for the determination of Saponification value, hydroxyl value and turbidity point in Emulsogen HCO 040 and 060 using multivariate model by near infrared spectroscopy (NIR). The project will start with an introduction of the basic of spectroscopy and how to develop a multivariate model, Also, it will be explained the reference methods that it has been used and to conclude it will explain how we choose the best method and the evaluation of their reliability.

3 OBJECTIU

Desenvolupament i validació d'un mètode quantitatiu via espectroscòpia d'infraroig proper (NIR), que permeti la determinació del Índex d'hidroxiació, de saponificació i el punt de terbolesa en mostres industrials en anàlisis de rutina.

4 INTRODUCCIÓ

Clarian Ibèrica Producciones S.A és una empresa que aposta per les noves metodologies analítiques menys invasives i destructives del producte i que presenten un avantatge sobre la resta en quant es refereix al temps d'anàlisi. La branca estructural de Clariant de BU i ICS, ha començat recentment utilitzar una d'aquestes tècniques. La tècnica que mesura l'absorbència de les mostres a freqüències d'infraroig proper (NIR). Seguidament parlarem de que es tracta aquesta tècnica i de com funciona i com es desenvolupa un mètode a partir d'aquesta.

4.1 DESCOBRIMENT DE LA LLUM INFRAROJA

L'any 1800 l'astrònom Friedrich Wilhelm Herschel (Figura - 1), va analitzar l'espectre de la llum del sol. Herschel va crear l'espectre dirigint la llum solar a través d'un prisma vítric, per tant, la llum es va dividir en diferents colors.

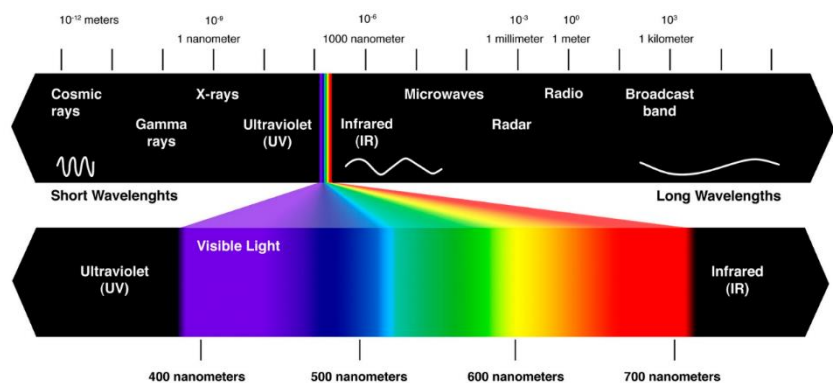


Figura - 1 Autoretrat de Friedrich Wilhelm Herschel (esquerra)¹ i freqüències espectrals (dreta) Font: <https://www.thegreenage.co.uk/tech/infrared-heating-panels/>

Ell va mesurar la capacitat calorífica de cada color utilitzant termòmetres amb una bombeta carbonitzada. Quan va mesurar la temperatura just per sota de la part vermella de l'espectre, va notar algun tipus de radiació invisible. Per la seva sorpresa, va trobar que en l'àrea propera a la part vermella (una àrea aparentment sense llum solar), tenia la major capacitat calorífica de tots. Herschel va concloure que hi havia d'haver un tipus diferent de llum per sota de la porció vermella, el qual no és visible per l'ull humà. Aquest tipus de llum va començar a coneixes com a llum "infraroja" (per sota del vermell, veure Figura - 1).¹

Després, Herschel va col·locar un recipient ple d'aigua entre el prisma i el termòmetre i va observar que la temperatura mesurada era menor que al mesurada sense aigua. Conseqüentment l'aigua havia d'absorbir parcialment la radiació. A més, Herschel va poder provar que segons com es rotava el prisma (depenent del rang espectral), la diferència en les temperatures mesurades per cada color, variaven. Aquest va ser el començament de l'espectroscopia infraroja.¹

L'espectroscopia infraroja mesura la llum infraroja que és absorbida per una substància. Aquesta absorció depèn de la longitud d'ona de la llum.

4.2 L'ESPECTRE ELECTROMAGNÈTIC

La llum visible i la llum infraroja són dos tipus de radiació electromagnètica, però amb diferents longituds d'ona o freqüències. En general, la radiació electromagnètica està definida per la longitud d'ona λ o la freqüència lineal ν . La longitud d'ona és la distància entre dos màxims en una ona sinusoïdal (Figura - 2).

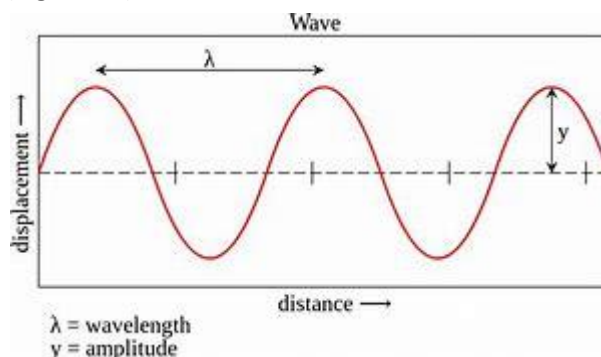


Figura - 2 Característiques de l'ona electromagnètica¹

La freqüència és el nombre de longituds d'ona per unitat de temps (Equació - 1). Des de que totes les ones electromagnètiques viatgen a la velocitat de la llum, la freqüència correspon a una longitud d'ona donada que es pot calcular.

$$\nu = c/\lambda$$

Equació - 1. Freqüència.¹

Segons la llei de la radiació de Planck, la freqüència és una radiació electromagnètica proporcional a la seva energia (Equació - 2).

$$E = h \cdot \nu$$

Equació - 2. Llei de Planck¹

En espectroscopia infraroja el nombre d'ona s'utilitza per descriure la radiació electromagnètica. El nombre d'ona és el nombre de longituds d'ona per unitat de distància.

4.3 INTERACCIÓ DE LA RADIACIÓ I LA MATÈRIA

L'espectroscopia IR està basada en l'absorció de la llum infraroja de la substància que s'ha de mesurar. Aquesta absorció excita les vibracions i rotacions moleculars, les quals tenen freqüències que són semblants a aquestes que estan dintre de la zona infraroja de l'espectre electromagnètic.

4.3.1 QUIN TIPUS DE MOLECULES ABSORBEIX LA LLUM INFRAROJA?

La llum infraroja sol es pot absorbir per molècules que durant la vibració presentin canvis en el moment dipolar dels grups específics dels àtoms implicats. Com més gran sigui el canvi en el moment dipolar, major serà la banda d'absorció corresponent en el IR.

Vibracions que no presenten canvis en el moment dipolar no poden ser excitades per absorció de la llum infraroja i són inactives en IR. Degut a això les molècules diatòmiques homonuclears, per exemple H_2 o O_2 , no tenen cap espectre IR.

En la dispersió de Raman, durant la vibració, la polaritzabilitat de l'enllaç canvia. Això significa que les vibracions inactives en IR són actives en Raman si la polaritzabilitat canvia. Els espectres de IR i Raman són complementaris entre ells.

4.3.2 SEPARACIÓ DE LES ZONES ESPECTRALS

L'infraroig mitjà, o MIR, és la zona espectral de 4000 a 400 cm^{-1} nombres d'ona. En aquesta zona, les vibracions fonamentals són típicament excitades. En contrast, en l'infraroig proper, o NIR, la zona espectral, la qual cobreix la zona de 12500 a 4000 cm^{-1} nombres d'ona, sobretons i vibracions de combinació són excitades. L'infraroig llunyà, o FIR, la zona espectral cobreix de 400 a 5 cm^{-1} nombres d'ona. Aquesta zona cobreix les freqüències vibracionals de la part estructural de molècules llargues com vibracions fonamentals que inclouen àtoms pesants, (per exemple, compostos organometàl·lics. També es pot observar en el FIR bandes rotacionals pures de gasos.

4.4 PARTICULARITATS SOBRE EL NIR

L'espectre NIR s'origina a partir de l'energia irradiada de l'energia mecànica associada al moviment dels àtoms units per enllaços químics en una molècula.

Normalment per entendre millor l'espectroscopia vibracional, es fa ús del model harmònic però aquest model presenta moltes restriccions cap a l'espectroscopia NIR perquè no permet transicions on Δv és major a 1. Transicions amb un $\Delta v=2$ o major són prohibides pel model harmònic/quàntic i, per tant, molts dels fenòmens observats en la regió del NIR, les bandes de sobretons, no haurien d'existir. També, les vibracions en el model harmònic són independents i les seves combinacions no haurien d'existir sota les restriccions imposades pel model. Per altre banda les bandes de sobretons i combinació existeixen.²

El model que explica aquestes transicions prohibides seria el model anarmònic/quàntic, encara que també imposa algunes restriccions en els possibles estats energètics de les molècules. De totes formes, aquest model, prediu l'aparició de transicions amb $\Delta v=2$ o major i l'existència de bandes de combinació entre vibracions. Aquests dos tipus de bandes són les bandes més comunes en la regió espectral del NIR.

La complexitat de l'espectre que s'obté, prohibeix el seu ús directe, com es fa amb el MIR, on hi ha bandes d'absorció molt netes i fàcilment distingibles i atribuïdes per exemple a la presència de certs grups funcionals. Per altre banda l'espectroscopia NIR s'està utilitzant de forma molt ampla per identificar matèries primes, per exemple, en l'indústria farmacèutica, tenint en compte que s'observen gràfiques de correlació on estan localitzades les bandes d'absorció dels enllaços de O-H, C-H, N-H i S-H de diferents compostos i es poden utilitzar com a una primera aproximació per un anàlisi qualitatiu.²

El NIR no és una tècnica absoluta sinó relativa, per tant, s'utilitzen mostres en les que es sap el contingut de l'analit d'estudi, determinat a partir d'un mètode ben acceptat, designat com a mètode de referència, i es mesura.³ A partir d'aquí es calcula la relació matemàtica entre l'absorbància de la mostra i el seu valor de referència. Més tard, el model matemàtic calculat es pot aplicar per l'anàlisi d'espectres de mostres desconegudes per tal de estimar el contingut de un o múltiples analits de forma simultània en anàlisis de rutina.

4.5 METODOLOGIA AMB EL NIR

En el cas d'una avaluació quantitativa de l'espectre NIR, el valor mesurat és típicament un espectre d'absorció o d'emissió, i el valor a determinar és la concentració de l'analit. Hi ha dos mètodes per posar a punt un model de calibració per aquests sistemes: la calibració univariant i la calibració multivariant.

4.5.1 CALIBRACIÓ UNIVARIANT

En el cas de la calibració univariant, els valors d'absorció del pic amb màxima absorbància són representats gràficament enfront la concentració de l'analit. De la recta de calibratge que és

forma es determina la seva equació de la recta que permet el càlcul de la concentració de l'analit de mostres desconegudes.

En molts casos la calibració univariant porta a una insuficient capacitat predictiva, ja que el mètode presenta diverses limitacions:

La concentració de l'analit sol està correlacionat a un sol punt de l'espectre, per tant, alhora d'avaluar noves , mostres univariants, no es poden reconèixer ni punts discrepants ni altres interferències. En altres paraules: a partir de l'altura o l'àrea del pic, un no pot assumir que l'estructura de l'espectre mesurat es apropiat per al conjunt que es farà servir en la calibració.

La variació estadística de la senyal, com el detector de soroll, és directament incorporat al conjunt de calibració. La incertesa resultant s'ha de minimitzar amb la mesura de múltiples mostra i com a resultat s'ha de fer una mitjana dels resultats obtinguts.

Una calibració satisfactòria de sistemes amb múltiples components necessita una separació suficient del pic amb màxima absorbència. Encara que en molts casos això no és possible, especialment en l'espectroscopia NIR.

En l'anàlisi de sistemes amb múltiples components, una addició dels valors de l'absorbència de tots els analits en les freqüències mesurades és assumida (Llei de Lambert-Beer). Representar els valors de l'absorbència contra la concentració porta a una funció de la calibració lineal. En molts casos, això no és vàlid en els sistemes reals. Forces intramoleculars o efectes de la temperatura poden provocar la distorsió dels respectius enllaços de l'analit. Per altre banda, hi ha varis tècniques pel qual la llei de Lambert-Beer no és vàlida, com la mesura de la reflectància difusa, una tècnica que s'utilitza molt sovint en l'espectroscopia IR .

Així, utilitzar en un anàlisis mètodes que fan servir calibracions univariants clàssiques poden portar resultats poc útils en sistemes en múltiples components. Un mètode per resoldre aquests tipus de problemes, és la utilització de mètodes de calibració multivariants, com la Regressió Lineal Múltiple (MLR), la Regressió per Components Principals (PCR), o la Regressió de mínims quadrats (PLS).⁴

4.5.2 CALIBRACIÓ MULTIVARIANT

La quimiometria és la utilització de tècniques matemàtiques i estadístiques per a extreure la informació rellevant provinent d'informació analítica.

El PCR i el PLS són considerats com les tècniques de calibració estàndards en l'espectroscopia NIR. El major avantatge que presenta aquestes tècniques és que evita problemes de col·linealitat permetent així treballar amb un número de variables que és major al número de mostres. Una comparació entre aquestes dues tècniques dona a lloc a resultats similars pel que fa a la capacitat de predicció, tampoc s'ha observat cap diferència significativa quan s'han utilitzat el número òptim de components principals (PCs). El PLS produeix bons models utilitzant un menor número de PCs que el PCR. Degut a això el PLS és més utilitzat en l'espectroscopia NIR que el PCR.

4.5.2.1 REGRESSIÓ PER MÍNIMS QUADRATS (PLS)

Per a dur a terme un PLS, la informació de l'espectre ha de ser comparat per els corresponents valors de la concentració dels analits. Els canvis que passen en les estructures de les dues fonts d'informació s'han de reconèixer i correlacionar en cada una.

Per aquest propòsit, un gran número de mostres s'ha de mesurar. Tota la informació ve representada en forma de matrius en les quals se'n generen els vectors propis.

Aquests vectors propis són nomenats factors o components principals. Aquests es poden utilitzar per la predicció de concentracions en comptes de l'espectre original, ja que en ells contenen tota la informació rellevant del sistema en estudi. L'avantatge d'aquesta descomposició és obvia: la informació analítica rellevant d'un conjunt enorme d'informació és reduïda en factors que es poden utilitzar en la calibració.

En el cas de la calibració PLS, els vectors propis són disposats en ordre descendent. El primer factor caracteritza els canvis principals de l'espectre i presenta la major importància en el model de calibració. Això presenta una important conseqüència per a l'avaluació de l'espectre: menys factors caracteritzen els major canvis en l'estructura de l'espectre, i en canvi més factors representa la part del soroll espectral (Figura - 3).⁴

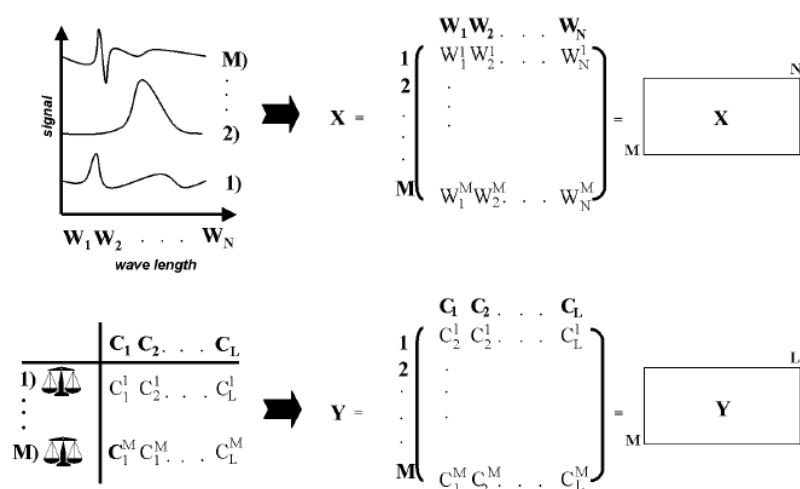


Figura - 3 Matrius generades en el model PLS

La selecció del número òptim de factors és de gran importància per la qualitat del model PLS. Amb massa pocs factors, les estructures espectrals no es reconeixen de forma suficient i la corresponent regressió de les dades no portarà mai cap a uns resultats dels anàlisis satisfactoris.

Això és descrit com a "underfitting". Si el número de factors és massa gran, això porta a la degradació dels anàlisis, ja que massa parts del soroll espectral són incorporats ("overfitting").⁴

En la regressió PLS, la matriu de la informació espectral X i la de la concentració Y són reduïdes cap a poc factors.

Generalment, el nombre de valors de l'absorbància mesurada excedeix enormement al número de components presents. Per tant, el sistema està "over-determined" i es pot correlacionar no solament amb un punt espectral (com el pic amb màxima absorbància en la calibració univariant), sinó amb tota l'estructura de la informació espectral.⁴

La informació emmagatzemada en el conjunt de calibració, d'aquesta forma, és molt més gran que la que s'obté en la calibració univariant. Així és possible detectar punts discrepants durant el anàlisis i es pot decidir si components desconeguts que no estan correlacionats amb el conjunt d'informació poden donar canvis en l'estructura de l'espectre.

En la regressió PLS el conjunt d'informació es descompon primerament en els components principals. Seguidament es du a terme un condicionament dels vectors dels scores provinents dels valors espectrals i dels valors del mètode de referència.

Aquest anàlisi multivariant presenta dos tipus d'algoritmes, el PLS1 i el PLS2:

L'algoritme PLS1 sol té en compte el valor de concentració d'un analit. La resta d'informació l'interpreta com a interferències, per tant, la matriu Y de la informació de les concentracions resulta ser un vector.

En l'algoritme PLS2 les concentracions de tots els components del sistema es tenen en compte per la calibració. Per la predicció de noves mostres, el model realitza un anàlisi simultani de totes les substàncies.

En contrast amb el PLS1, tota la informació de la matriu de concentracions s'ha de correlacionar amb la informació de la matriu espectral.

El PLS2 normalment obté resultats més pobres que la predicció amb el PLS1. Per aquesta raó és recomanable realitzar la calibració amb l'algoritme PLS1. Per a dur a terme un anàlisi d'un sistema amb múltiples components, aquests algoritme és aplicat de forma successiva en tots els components calibrats, obtenint el model desitjat amb tots els components tal i com s'obtindria amb el PLS2.⁴

4.5.3 VALIDACIÓ DEL MÈTODE

Per a realitzar la validació d'un mètode multivariant, hi ha dues formes de fer-ho, sempre segons el número de mostres que s'utilitzi per fer la calibració del mètode.

4.5.3.1 VALIDACIÓ INTERNA O CREUADA

Aquest tipus de validació es porta a terme extraient mostres individuals del conjunt de calibració. Utilitzant les mostres restants es crea un model quimiomètric i s'utilitza per analitzar les mostres prèviament separades. Una comparació dels resultats obtinguts amb la concentració real de la mostra demostra com de precís és el model alhora de predir les concentracions de les mostres. Extraient les mostres abans de crear el model quimiomètric es garanteix que la mostra sigui desconeguda i independent del model de calibració.

Per avaluar tot el conjunt de calibració, les mostres analitzades es retornen al conjunt de calibració i un segon grup s'extreu del conjunt. El procediment de remoure mostres, analitzar-les, i retornar-les al conjunt de calibració es continua de forma successiva fins que s'han analitzat totes les mostres.

Una comparació amb els resultats de cada anàlisi amb la informació original permet el càlcul de l'error de predicció de tot el sistema, RMSECV (Root Mean Square Error of Cross Validation). Aquesta és una mesura quantitativa de la diferència de la capacitat predictiva del model quimiomètric.

Per a la validació creuada és important remoure poques mostres del conjunt de calibració, ja que el model construït a partir de la resta de mostres ha de ser el més similar possible al model creat a partir del conjunt de mostres original. Per a un conjunt amb menys de 50 mostres és recomanable no moure més d'una mostra per la validació creuada.

4.5.3.2 VALIDACIÓ EXTERNA

En contrast amb la validació interna, totes les mostres s'utilitzen per a generar el model. Aquest model es manté constant per a futures validacions. El anàlisi de l'espectre ja no es remou del conjunt de calibració. Per a estimar l'error de predicció del model, altres mostres es mesuren i s'afegeixen en l'anomenada conjunt de proba. Per la validació solament les mostres del conjunt de proba són analitzades, per tant el conjunt de mostres per a la validació externa es divideix en conjunt de calibració i conjunt de proba. En contrast amb la validació creuada, no hi ha un intercanvi entre els dos conjunts de mostres.

La comparació amb els resultats obtinguts i la concentració real de les mostres s'utilitza per calcular el RMSEP (Root Mean Square Error of Prediction). Aquest valor representa la presició en la mesura quantitativa de la predicció en el model.

Bons models són caracteritzats per un valor de RMSEP baix. Les validacions internes i externes haurien de donar resultats comparables. Si no es dona el cas implica que hi ha massa poques mostres analitzades per establir un model fiable.

Una forma segura de saber quan utilitzar cada una de les validacions és comparant els resultats d'aquestes dues. Si els paràmetres estadístics són millor en la validació interna que en la externa, això significa que no hi ha mostres suficients per realitzar la validació externa.

Generalment és recomanable utilitzar la validació externa a partir de 50 mostres, sinó s'hauria d'utilitzar la validació interna, encara que no sigui tan precisa com la validació externa.

4.6 EMULSOGEN HCO

El producte que utilitzem per a desenvolupar el seu mètode en el NIR és resultat de l'esterificació d'una glicerina (alcohol) quan reacciona amb oli de Ricí (un àcid gras) i la posterior polimerització d'aquesta molècula a partir d'afegir òxid d'etilè que reacciona amb els hidroxils lliures de la cadena d'oli de Ricí.

4.7 PER QUÈ UTILITZAR EL NIR?

L'anàlisi via NIR utilitza tècniques de reflectància difusa com a una mesura no destructiva, el qual fa possible treballar amb la mostra directament sense pretractament. A més el seu anàlisi es realitza en menys d'un minut.

Partint de tot el que s'ha dit i comparant amb els mètodes de referència que es fan servir, és evident que la utilització de mètodes via NIR redueix enormement el consum de reactius que poden arribar a ser tòxics, el temps d'anàlisi, la interacció amb la mostra a analitzar i l'economia d'anàlisi de rutina.

5 PROCEDIMENT EXPERIMENTAL

5.1 DIAGRAMA DE GANTT

El diagrama de Gantt (Figura – 4) és una eina per planificar i programar feines durant un període determinat. Gràcies a una fàcil i còmode visualització de les accions previstes, permet realitzar el seguiment i control del progrés de cada una de les etapes d'un projecte i, a més, reproduïx gràficament les diferents feines, la seva durada i seqüència, a més del calendari general del projecte.

ACTIVITAT	03-abr	04-abr	05-abr	06-abr	07-abr	08-abr	09-abr	09-maig	10-maig	11-maig	12-maig	13-maig	14-maig	15-maig	16-maig	17-maig	18-maig	19-maig	20-maig
Recopil·lació mostres																			
Anàlisi mostres																			
Organització dades																			
Desenvolupament mètodes																			
Robustesa																			

Figura - 4 Diagrama de Gantt de la part experimental

5.2 RECOPIACIÓ DE MOSTRES

Cada matí, en arribar al laboratori, es revisa la safata que s'ha disposat per deixar aquells productes que s'han portat per analitzar de producció, i agafar del producte Emulsogen HCO 040 i 060.

Aquestes mostres seran provinents de producció, però per obtenir un conjunt de mostres representativa també es necessita mostres de producte final, els quals estan en un magatzem al laboratori de l'edifici social de l'empresa.

5.3 REVISIÓ DELS VALORS DELS ANÀLISIS DE REFERÈNCIA

Per a les mostres que es recol·lecten, revisar si presenten tots els valors dels anàlisi de referència els quals farem l'estudi.

Sabem que en les mostres de producció sol tenen com a especificació realitzar l'anàlisi del Índex de saponificació, per tant, caldrà que es realitzin els anàlisis del Índex d'hidroxilació i el punt de terbolesa per aquestes mostres. A més, si una de les mostres ha sigut anul·lada degut a que un dels altres anàlisis que requeria ha donat fora e l'especificació que demanava, doncs tampoc se li haurà fet l'anàlisi del Índex de saponificació i caldrà realitzar-lo també.

En el cas dels productes finals, en la seva fulla d'especificacions si que hi són tots els anàlisis que es vol estudiar, per tant, podem estar segurs que s'han realitzat.

5.4 RECOPIACIÓ DE DADES:

Un cop determinada l'absència dels valors dels anàlisis de referència, realitzar-les tal i com s'ha detallat anteriorment en l'apartat de de mètodes de referència que s'han utilitzat.

5.4.1 MÈTODES DE REFERÈNCIA UTILITZATS

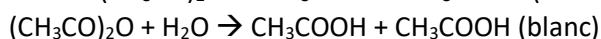
Per assegurar la qualitat del producte en tot moment, és necessari establir una paràmetres de qualitat que venen com a especificacions del producte en cada etapa del procés i per acabar en el producte final.

5.4.1.1 ANÀLISIS DE L'ÍNDEX D'HIDROXILACIÓ:

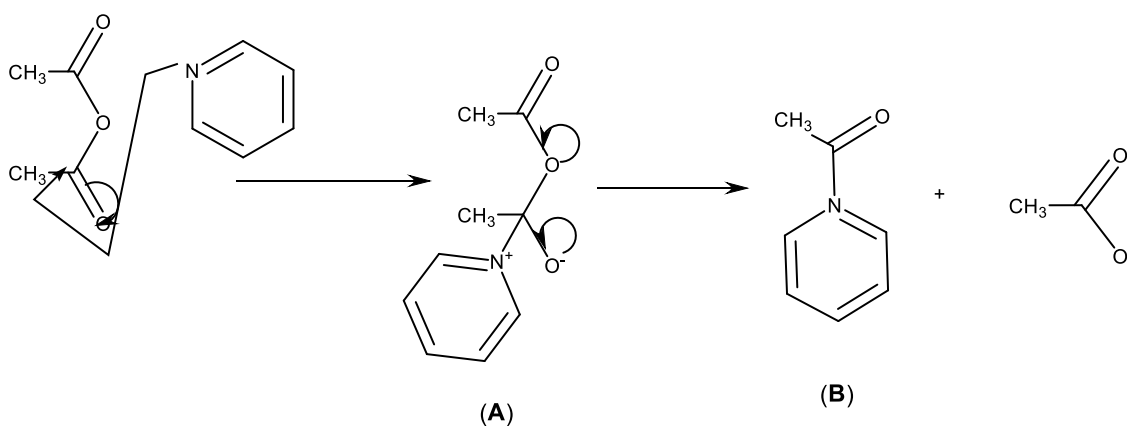
L'Índex de hidroxilació I(OH) es defineix com el número de mil·ligrams d'hidròxid potàssic equivalents al contingut en radicals hidroxil (OH) d'un gram de producte.

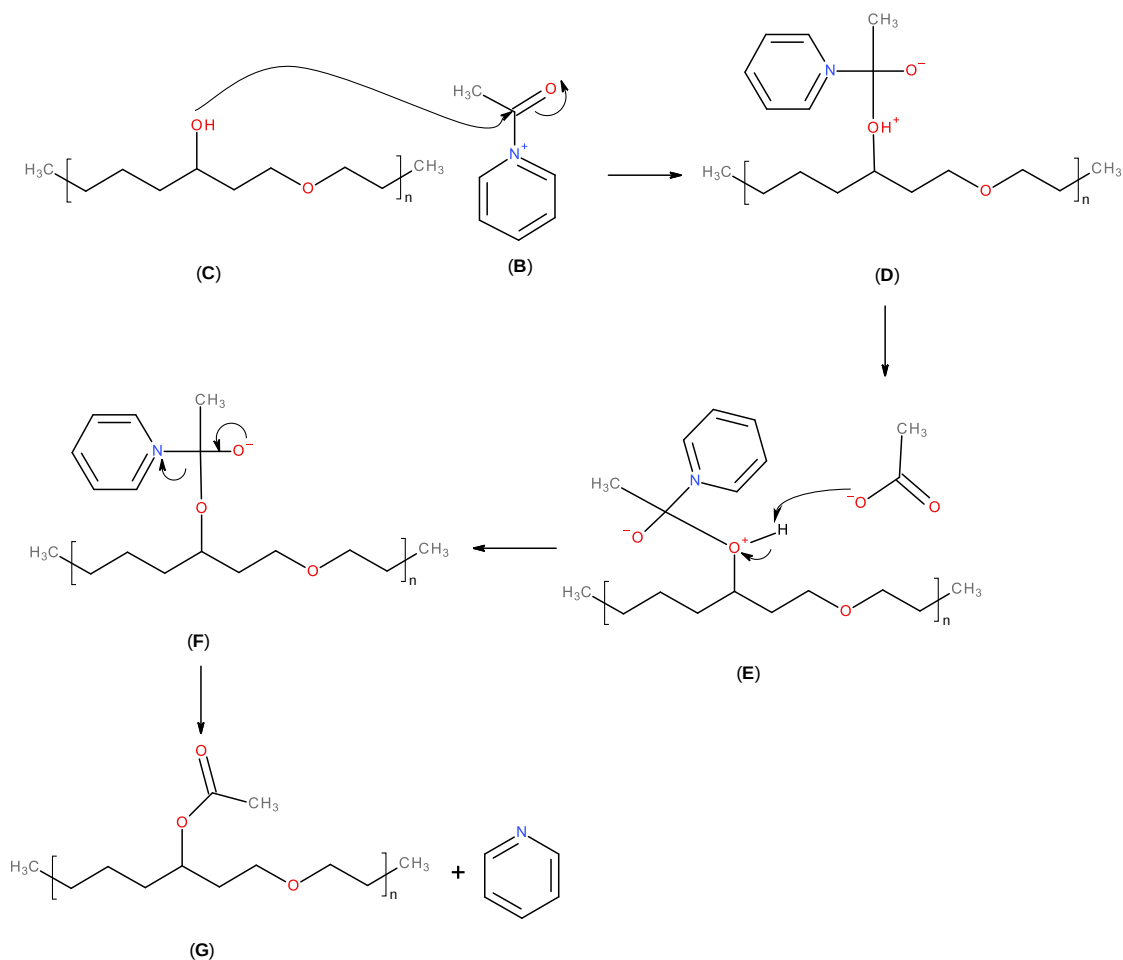
El mètode és aplicable a la determinació de les funcions hidroxil en derivats etoxilats, propoxilats o barreges d'alcohols grassos primaris o secundaris, de alquilfenols i d'àcids grassos. Permet la determinació del Índex de OH per a valors compresos, aproximadament, entre 10 i 1000 mgKOH/g.⁵

La fonamentació d'aquest mètode recau en l'esterificació (Esquema - 1), amb un excés d'anhídrid ftàlic (o acètic) en piridina, dels grups -OH primaris o secundaris (no terciaris) presents en la mostra. A continuació s'hidrolitza l'excés d'aquest anhídrid per adició d'aigua i es valora l'àcid format amb una dissolució patró de NaOH.⁵



La piridina, que és un dels reactius que s'utilitza, és de molta importància per a que la reacció es dugui a terme, ja que activa el complex que ha de reaccionar amb l'anhídrid (espècie B, Esquema - 1).





Esquema - 1 Mecanisme d'activació del complex de la piridina i reacció d'esterificació de l'alcohol i l'àcid carboxílic⁶

El mecanisme de reacció consisteix en l'addició de la piridina a l'anhídrid (A), trencament de l'anhídrid (B), atac del grup hidroxil al carboni carbonílic (C), formació del complex intermedi amb la piridina (D), el sintó de l'anhídrid ataca el protó del complex per formar àcid acètic (E), formació de l'èster (F) i recuperació de la piridina amb la formació de l'èster (G).

Per a efectuar l'anàlisi, prèviament, s'ha de preparar les dissolucions d'anhídrid ftàlic o acètic amb piridina i guardar-los en una nit en un recipient de color topazi.

El mètode operatiu es tracta de preparar per triplicat en matrassos Erlenmeyer de 250 mL, perfectament secs.

Seguidament es pesa una quantitat de producte que està tabulada segons el resultat del Índex d'hidroxilació que s'espera.

Després s'ha d'afegir, en el matràs Erlenmeyer, mitjançant una pipeta, 25 mL de la mescla preparada d'acetilació amb anhídrid ftàlic (o bé 50 mL de la mescla amb anhídrid acètic), i uns trossets de porcellana porosa.

Posar en reflux uns 35 minuts i deixar refredar seguidament uns 20 minuts. Un cop passat aquest temps, sense desconnectar el condensador dels matrassos, afegir 25 mL d'aigua destil·lada (quan s'utilitzi l'anhídrid acètic amb 10 mL d'aigua destil·lada és suficient). Agitar bé els Erlenmeyers i netejar els condensadors amb 25mL piridina per si han quedat traces.

Realitzar un assaig del blanc per duplicat amb el mateix mètode descrit anteriorment però ometent la mostra del assaig.

El blanc es realitza per saber quina quantitat total hi ha d'anhídrid i així saber la seva diferència amb el consumit quan es realitzi la valoració de la mostra.

Per finalitzar l'anàlisi, es valora la mescla dels matrassos amb una solució de NaOH 1M en presència de fenolftaleïna, o si la mescla és massa fosca, es pot utilitzar el potenciògraf.

El consum del blanc serà aproximadament de 36 mL de NaOH 1M i la diferència de volums entre el blanc i la mostra serà d'aproximadament de 8 mL.

Adicionalment es determina l'Índex d'acidesa (IA) o basicitat (IB) de la mostra del producte dissolent uns 10 g del producte en 100 mL d'IPA (isopropanol) mitjançant el potenciògraf, valorant amb HCl 0.1M si la mostra és alcalina i amb NaOH 0.1M si la mostra és àcida.

Càlcul:

El càlcul a realitzar bàsicament és un factor de conversió entre el consum real de NaOH 1M cap a mg de KOH ja que inicialment aquest anàlisi es feia amb potassa en comptes de NaOH.⁸

De forma simplificada és la següent (Equació - 3 Càlcul de l'Índex d'hidroxilacióEquació - 3):

$$I(OH) = \frac{(Vb - Va) * N * f * 56.11}{P} (+IA)(-IB)$$

Equació - 3 Càlcul de l'Índex d'hidroxilació⁵

On:

Vb: mL de NaOH consumits en l'assaig del blanc

Va: mL de NaOH consumits per la mostra

N: normalitat de la solució de NaOH = 1

f: factor de la solució de NaOH

P: pes de la mostra (g)

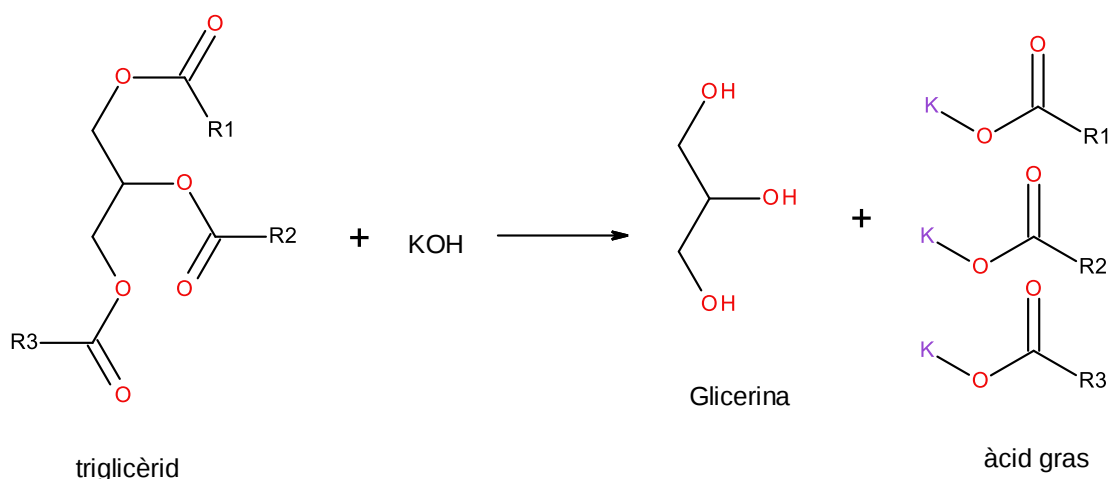
IA/IB: Índex d'acidesa/basicitat del producte

5.4.1.2 ANÀLISIS DE L'ÍNDEX DE SAPONIFICACIÓ:

Aquest procediment s'utilitza per a la determinació del Índex de saponificació de tensioactius; olis i greixos vegetals i productes similars. S'obté el contingut d'èster i anhídrids i/o àcids lliures que hi pugui haver a la mostra.⁷

El límit de quantificació és de 1 a 500 mgKOH/g i cobreix a èsters i anhídrids.

L'índex de saponificació (unitat: mgKOH/g) es defineix com el número de mil·ligrams de KOH necessaris per saponificar un gram de producte. Es determina, després de saponificar la mostra del producte amb un excés de solució de KOH, per valoració amb HCL del KOH present en excés.⁷



Esquema - 2 Reacció de saponificació del triglicèrid del producte

La seva reacció es basa en la reacció d'un triglicèrid amb KOH obtenint la glicerina i l'àcid gras que ve en forma de sal (Esquema - 2).

La seva valoració es farà amb la base conjugada d'aquesta sal utilitzant HCl 0.5M.

El seu mètode operatiu es basa en fer l'anàlisi per triplicat pesant en un Erlenmeyer net i sec, una quantitat de mostra, la qual està tabulat segons l'Índex de saponificació esperat, amb una precisió de 0.01g.

A cada Erlenmeyer, afegir 25 mL de solució metanòlica de KOH 0.5N amb una bureta d'unitat intercanviable i 50 mL de iso-butanol. Afegir uns trossets de porcellana porosa per a que l'ebullició sigui uniforme.

Col·locar a reflux durant dues hores.

Posteriorment, refredar, afegir unes gotes de solució de fenolftaleïna i valorar amb la solució de HCl 0.5N fins que la solució passa de color lila a incolora.

També s'ha de realitzar un assaig en el blanc, seguint el mateix procediment però ometent la mostra.

En aquest cas el blanc es realitzà per saber la quantitat total de KOH que s'afegeix en cada mostra i així poder determinar la seva diferència quan es valori amb HCl 0.5M en les mostres.

Valorar en calent quan el producte solidifiqui al refredar-se.

Quan el color de la mostra no permeti veure el punt final de la valoració, es pot efectuar la valoració amb el titroprocessador. En aquest cas el blanc ha de valorar-se pel titroprocessador.

L'índex de saponificació es determina mitjançant la següent fórmula (Equació - 4):

$$Ind. Sapo = \frac{(V0 - V1) * N * f * 56.11}{p}$$

Equació - 4 Càlcul de l'Índex de saponificació⁷

On:

V0 = mL de la solució de HCl consumits en el blanc.

V1 = mL de la solució de HCl consumits en la valoració de la mostra.

F = factor de la solució de HCl utilitzada.

N = concentració de la solució de HCl utilitzada = 0.5N

p = pes, en g, de la mostra utilitzada.

5.4.1.3 ANÀLISI DEL PUNT DE TERBOLESA (TP):

Aquest mètode té com a objectiu fixar les condicions de mesura de la temperatura d'enterboliment dels agents de superfície no iònics que es deriven del òxid d'etilè (OE) per condensació d'aquest sobre un compost lipòfil sense grups oxipropilènics.⁸

El TP es determina normalment en derivats de OE de alquil(aril)alcohols, àcids grassos i amines grasses.⁸

El TP depèn del número i amplada de la distribució estadística dels grups OE i es veu influenciat per electròlits i altres compostos orgànics.

El seu fonament es basa en que les dissolucions aquoses dels agents iònics de superfície no iònics derivats del òxid d'etilè es fan heterogènies, a causa de la formació de dos fases líquides, quan s'eleva la seva temperatura.

Aquesta particularitat es presenta a una temperatura més elevada quan hi ha un major número de molècules de OE condensades sobre la base lipòfila del agent de superfície.

La mescla de les dos fases es torna un altre cop homogènia quan la temperatura baixa. Aquesta temperatura, on la mescla torna a ser homogènia, es nomena de forma convencional, temperatura de clarificació o punt d'enterboliment.

Per a realitzar aquest anàlisi s'ha de preparar prèviament una solució al 10% de clorur sòdic en aigua.

A continuació, es pesen 0.5 grams del producte i s'afegeixen 49.5 ml (52.4 g, considerant la densitat de la solució de NaCl 1.059 g/mL) de la solució al 10% de NaCl.

5.5 ANALITZAR PER NIR

Un cop realitzats tots els anàlisis que faltaven, passem la mostra per un espectre d'infraroig proper (NIR).

Els passos a seguir són els següents:

En la pantalla del equip s'ha d'especificar el producte que s'està analitzant, l'operació on prové la mostra, ja sigui de producció o de producte final, el número de mostra i el nom de l'analista que està realitzant l'anàlisi. D'aquesta forma es pot tenir un seguiment del procés i totes aquestes dades queden enregistrades en el nom de l'espectre que s'obté.

Seguidament mitjançant una pipeta Pasteur introduïm mostra en una cubeta de vidre fins omplir-la unes tres quartes parts. Tapem la cubeta amb un tap de plàstic i introduïm la cubeta tapada en la cavitat que hi ha en l'equip.

Prèviament haurem configurat l'equip amb les següents condicions:

1. Temperatura de mesura: 70°C
2. Mesures per temps: 16 scans
3. Resolució: 8 cm⁻¹
4. Mesures per mostra: 1

L'anàlisi dura menys d'un minut, (depèn de la temperatura en que estigui la mostra i el temps que es triga en condicionar-se a la temperatura de treball).

5.6 TRACTAMENT DELS ESPECTRES

Després de haver punxat en l'equip del NIR totes les mostres, es transfereix els espectres en una carpeta per facilitar la seva cerca. A més, es reparteix els espectres en tres carpetes diferents, segons el mètode que es vulgui desenvolupar, i s'han anomenat de la següent forma:

SAPO: per al desenvolupament del mètode per a la determinació del Índex de saponificació.

OH: per al desenvolupament del mètode per a la determinació del Índex d'hidroxilació.

TP: per al desenvolupament del mètode per a la determinació del punt de terbolesa.

Un cop tenim tots els espectres, copiem els espectres en una fulla d'Excel i assignem a cada espectre el seu valor del mètode de referència que utilitzem.

Tenint en compte que hi havia mostres de producció que no les teníem físicament i les faltaven alguns dels anàlisis, es va separar en tres fulles excels els espectres segons el tipus de mètode que es volia desenvolupar, determinació del Índex de saponificació, determinació del Índex d'hidroxilació o determinació del punt de terbolesa.

5.7 DESENVOLUPAMENT DEL MÈTODE

Per a desenvolupar el mètode, utilitzarem el software OPUS que ens l'ha proporcionat la mateixa casa que gestiona l'equip del NIR, Bruker.

5.7.1 PART EXPERIMENTAL

Els passos a seguir són els següents:

En la pestanya de components del software indicar el component del qual es desenvoluparà el mètode i així en permetrà afegir la seva corresponent columna en la pestanya d'espectres (Figura - 4).

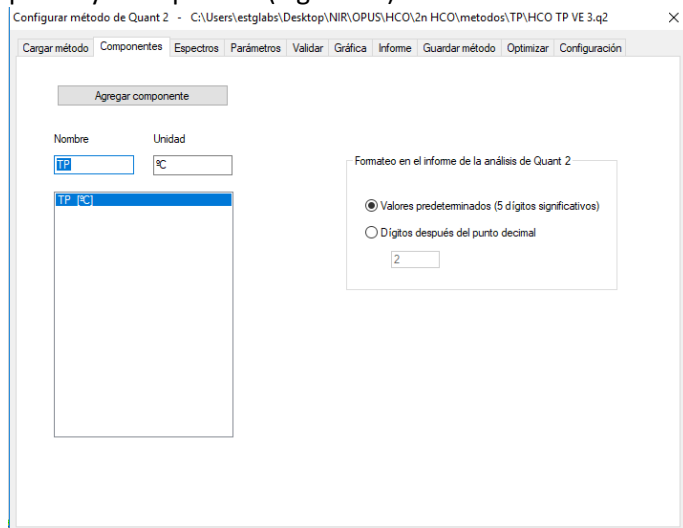


Figura - 4 Assignació de components

Agregar els espectres que s'han d'utilitzar per a realitzar la calibració juntament amb els valors provinents del mètode de referència de cada espectre (Figura - 5).

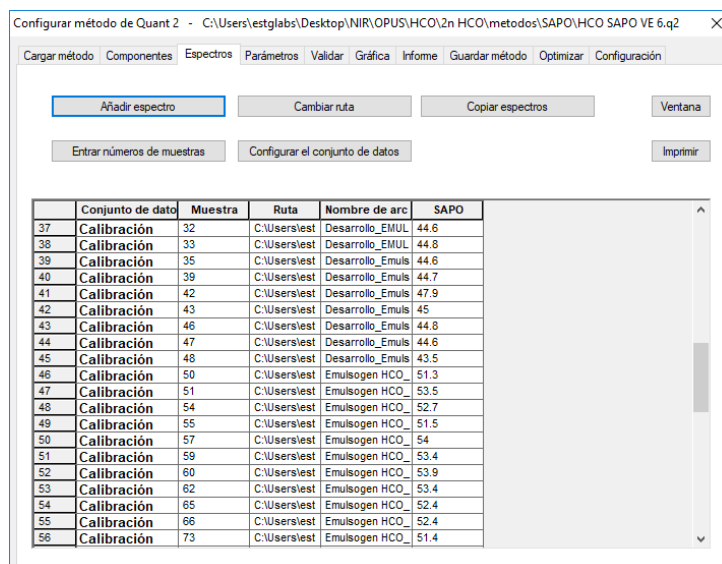


Figura - 5 Pestanya per agregar els espectres i els respectius valors de referència

Un cop seleccionats els espectres per a realitzar la calibració del mètode realitzem la configuració del paràmetres per optimitzar i seleccionar la zona espectral de la qual es vol calibrar el mètode (Figura - 6).

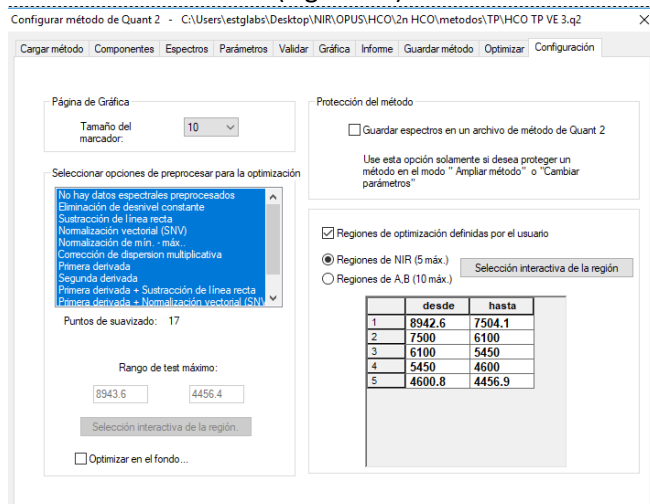


Figura - 6 Pestanya de configuració de l'òptimització de l'espectre

Per començar es realitzarà una validació creuada per tenir els paràmetres estadístics resultants com a referència i després es realitzaria la validació externa.

Quan s'ha seleccionat tots el paràmetres de qualitat per desenvolupar el mètode es clica l'opció d'optimitzar el mètode i es selecciona el mètode que presenti un valor de RMSECV o RMSEP amb el valor més baix.

Per realitzar la validació externa prèviament factoritzem el conjunt d'espectres i seleccionem a l'atzar un 20% del conjunt total d'espectres que s'utilitzen i especifiquem que el aquest conjunt d'espectres s'utilitzi com a prova del mètode calibrat.

A partir d'aquí obtenim les gràfiques corresponents a la predicció del valor vs el valor real provinent de mètode de referència i la diferència del valor predit amb el valor real. Amb aquestes gràfiques el software identifica els possibles punts discrepant a partir de si el seu valor calculat per a la distància de Mahalanobis és major al establert coma

referència. Si és aquest cas per tal d'obtenir millors paràmetres estadístics s'haurà d'eliminar de la calibració aquests espectres.

Es realitza els punts 6 i 7 tants cops com faci falta procurant que el subconjunt del 20% d'espectres per a realitzar la validació externa sigui a l'atzar i diferents cada cop.

Per acabar escollir el mètode que presenti millors paràmetres estadístics.

5.7.2 PART ESTADÍSTICA

5.7.2.1 PARÀMETRES ESTADÍSTICS

Alhora d'escollir el mètode òptim s'ha de tenir en compte una sèrie de paràmetres estadístics que ens ajudaran a comprendre si el mètode que s'ha desenvolupat és òptim per a predir la concentració de l'analit desitjat.

5.7.2.1.1 Bias

El Bias és la mitjana sistemàtica de la diferència entre els valors reals de concentració i els valors predits. És calculat fent la mitjana de totes les diferències de les mostres dintre del conjunt de mostres (Equació - 5).

$$\text{bias} = \frac{\sum_{i=1}^M Y_i^{\text{meas}} - Y_i^{\text{pred}}}{M}$$

Equació - 5 Equació del Bias⁴

On:

Y_i^{meas} : és el valor de concentració real

Y_i^{pred} : és el valor predit pel model

M: és el número de mostres analitzades

5.7.2.1.2 DISTÀNCIA DE MAHALANOBIS (MD)

Durant la factorització, la mesura espectral es descompon en diferents factors i un residu espectral. Això s'aplica en les mostres de calibració com en les mostres d'anàlisi. Si es realitza l'anàlisi d'una mostra desconeguda amb l'algoritme PLS, s'haurà de revisar si es pot analitzar l'espectre amb confiança amb el mètode. Això és possible amb el càlcul de la distància de Mahalanobis, on es revisa com de bé l'espectre de l'analit s'acomoda a l'espectre del conjunt de calibració.

La distància de Mahalanobis (Equació - 6) es defineix com la diferència de l'espectre mesurat de l'analit amb la diferència dels valors de tots els espectres del conjunt de calibració. Com més gran sigui aquesta diferència, més gran serà el valor de la distància de Mahalanobis. Hi ha diverses raons perquè això passi. Factors externs, com la contaminació de les mostres o canvis de la temperatura, que porten a la distorsió de la simetria del pic, el qual resulta en un augment del valor del MD. De la mateixa forma aquest valor augmenta si s'analitzen mostres que presenten una concentració fora del interval de concentracions. El MD és una mesura quantitativa per la credibilitat de l'anàlisi, perquè tant els punts discrepants com les mostres amb valors de referència fora de l'interval de concentracions vàlid són detectats.

Durant la regressió PLS, el MD de tots els espectres de calibració són calculats. Dels resultats el valor màxim tolerable és determinat, tal que l'espectre pot ser analitzat de manera segura amb

la funció de calibració donat. Mostres que estan per sobre d'aquest valor determinat, no poden ser determinats amb credibilitat. Aquestes mostres són potencials punts discrepants.

$$h_i = t_i^T (X^T \cdot X)^{-1} \cdot t_i$$

Equació - 6 Equació de la distància de Mahalanobis⁴

On:

h_i : és la distància de Mahalanobis

t_i^T : és el vector propi de la matriu T

t_i : és el vector propi de la matriu original

X^T : és la matriu T

X: és la matriu original

5.7.2.1.3 COEFICIENT DE DETERMINACIÓ (R^2)

El coeficient de determinació (Equació - 7) dóna el percentatge de varianza present en els valors dels components que són reproduïts en la predicció. Això és normalment nomenat com a la varianza explicada. Com més alt sigui el valor del coeficient, millor és la correlació entre els valors de la concentració i els valors espectrals. Hi ha diverses raons per a que hi hagi una mala correlació: seleccionar paràmetres del model inapropiats porta a un baix coeficient de determinació ($R^2 \ll 90\%$), o per exemple a una insuficient presició dels valors de referència, o l'existència de punts discrepant en el conjunt de calibració.

$$R^2 = \left[1 - \frac{SSE}{\sum_{i=1}^M (Y_i - Y_m)^2} \right] \cdot 100$$

Equació - 7 Equació del coeficient de determinació⁴

On:

SSE (Sum of Square Errors): és la suma dels valors residuals.

Y_i : és el valor predit pel model

Y_m : és la mitjana dels valors predits

Segons el valor de R^2 que s'obté es pot aplicar el mètode desenvolupat en diferents àmbits (Taula - 1).

Taula - 1 Classificació dels valors del coeficient de determinació segons la seva aplicació⁹

R ²	Clasificació	Interpretació
< 0,25	Very poor	Not usable
0,26 – 0,49	Poor	Poor correlation
0,50 – 0,64	Fair	Rough screening applications
0,66 – 0,81	Reasonable	Fair, Ok for screening applications
0,83 – 0,90	Good	Use with caution (QC)
0,92 – 0,96	Very good	Use with most application including some QA
0,98 – 0,998	Excellent	Use with any application
> 0,998	Superior	As good as reference

5.7.2.1.4 RMSECV (ROOT Mean Square Error of Cross Validation)

Aquest valor és la mesura quantitativa de la precisió amb que són predites les mostres durant la validació. El RMSECV (Equació - 8) és comparable amb el RMSEP de la validació externa.

$$\text{RMSECV} = \sqrt{\frac{1}{M} \cdot \sum_{i=1}^M (Y_i^{\text{meas}} - Y_i^{\text{pred}})^2} = \sqrt{\frac{1}{M} \cdot \text{PRESS}}$$

Equació - 8 Equació del RMSECV⁴

On:

M: és el número de mostres

Y_i^{meas} : és el valor de concentració original

Y_i^{pred} : és el valor predit

5.7.2.2 RMSEP (Root Mean Square Error of Prediction)

Aquest valor (Equació - 9) és una mesura quantitativa de la precisió del anàlisi del conjunt de prova. És comparable al RMSECV de la validació interna.

$$\text{RMSEP} = \sqrt{\frac{1}{M} \cdot \sum_{i=1}^M (Y_i^{\text{meas}} - Y_i^{\text{pred}})^2}$$

Equació - 9 Equació del RMSEP⁴

On:

M: és el número de mostres

Y_i^{meas} : és el valor de concentració original

Y_i^{pred} : és el valor predit

5.7.2.2.1 RPD (Residual Prediction Deviation)

El valor de RPD (Equació - 10) és el quocient de la desviació estàndard dels valors de referència (SD) (Equació - 11) i de la correcció del bias de la diferència d'error en la predicció de la validació (SEP_{bias}) (Equació - 12).

$$RPD = \frac{SD}{SEP_{bias}}$$

Equació - 10 Equació del RPD⁴

On:

$$SD = \sqrt{\frac{1}{M-1} \cdot \sum_{i=1}^M (Y_i^{meas} - Y_m)^2}$$

Equació - 11 Equació del SD⁴

On:

$$SEP_{bias} = \sqrt{\frac{1}{M-1} \cdot \sum_{i=1}^M (Y_i^{meas} - Y_i^{pred} - bias)^2}$$

Equació - 12 Equació del SEP_{bias} ⁴

On:

M: és el número de mostres

Y_i^{meas} : és el valor de concentració original

Y_i^{pred} : és el valor predit

Es una mesura qualitativa de l'avaluació dels resultats de la validació. Com més petit sigui el valor del SEP_{bias} , comparat amb la variança dels valors de referència, millor és el model, també com major sigui el valor de RPD, millor és la calibració.

Per avaluar la qualitat de la validació, calcular el RPD és molt més significatiu que sol mirar l'error de predicció. Per exemple, validar la calibració amb un interval relativament petit, porta a un error de predicció petit. El model reconeix sol una petita variació dels valors de referència i predirà els valors dels analits amb una variança petita (i aparentment un error petit), encara que el mateix model no sigui bo. Per evitar una interpretació massa positiva dels resultats, el valor del RPD és calculat: d'aquesta manera es pot observar que la desviació estàndard dels valors de referència són relativament petits (valors de SD petits) i de fet en una menor deterioració de l'error de predicció (augmenten els valors del SEP_{bias}) pot esdevenir a pitjors valors del RPD. Per tant, errors de predicció petits poden donar mals models de calibració.⁴

Mirant el valor del RPD presenta un altre avantatge: s'ha explicat que les mostres de calibració s'haurien de posicionar al llarg de tot l'interval de calibració de forma homogènia. Mètodes que presenten una gran proporció de mostres en el centre però poca proporció al final de l'interval, són menys robustes. Això ho té en compte el valor de RPD. Una distribució més homogènia dels valors de referència dona a lloc a uns valors de desviació estàndard (SD) majors concentrats en una posició. El RPD resultant és major i el model de calibració és categoritzat millor.⁴

Segons el valor de RPD obtingut, el mètode desenvolupat es pot aplicar en un àmbit (Taula - 2).

Taula - 2 Classificació dels valors del RPD segons la seva aplicació⁹

RPD	Clasificació	Aplicació
< 1,0	Very poor	Not recommended
1,0 – 2,4	Poor	Not recommended
2,5 – 2,9	Fair	Rough screening
3,0 – 3,9	Reasonable	Screening
4,0 – 5,9	Good	QC
6,0 – 7,9	Very good	QA
8,0 – 10,0	Excellent	Any application
> 10,0	Superior	As good as reference

5.8 VALIDACIÓ DEL MÈTODE

D'acord amb l'Organització Internacional de l'Estandardització, una validació del mètode es defineix com “ la confirmació mitjançant una examinació i aportació d'evidències objectives, pel qual els requeriments per a un ús específic o aplicació són complertes.

En un mètode univariant hi ha un seguit de passos per validar el mètode desenvolupat, que són el cel·lular una sèrie de paràmetres:

- Exactitud
- Precisió
- Especificitat
- Límit de detecció
- Límit de quantificació
- Linealitat
- Reproductibilitat
- Robustesa

En el cas del mètode multivariant ja presenta un seguit de validacions que serien la interna i externa i no està establert que s'hagin de realitzar els mateixos passos que en un mètode univariant.

Tots aquests passos es poden realitzar en el mètode del NIR però molts d'ells, sinó bé la majoria no fa falta aplicar-los. Tenint en compte que la tècnica del NIR és una tècnica relativa que relaciona els valors d'altres mètodes de referència en els seus respectius espectres, l'exactitud, la precisió, l'especificitat, el límit de detecció i quantificació, i la linealitat seran les mateixes que la dels mètodes de referència utilitzats.

En el cas de la reproductibilitat, la tècnica del NIR és poc invasiva i no destructiva de la mostra, el qual no requereix un pretractament previ, per tant no té sentit avaluar la seva reproductibilitat.

El que si que es pot avaluar és la robustesa del mètode per veure si factors externs i interns de l'equip del NIR, poden afectar a l'estructura de l'espectre de la mostra analitzada.

Per tant avaluarem la Robustesa segons Yodur-Stenner.

5.8.1 ROBUSTESA DEL MÈTODE

Un estudi de la robustesa estudia l'alteració de factors important en el procediment analític. En els experiments per estudiar els efectes dels factors, dissenys de prova són utilitzats. Placket i Burman van desenvolupar aquests dissenys, els quals, sol són utilitzables quan les interaccions són negligibles o quan es considera un conjunt de factors com a dominant, perquè els seus dissenys no poden treballar amb alteracions dels factors.¹⁰

Youden i Steiner en 1975 van introduir un test de robustesa molt útil per a la química analítica. Aquest test examina en un sol laboratori, la susceptibilitat dels mètodes analítics contra petits canvis en els paràmetres del mètode utilitzant dissenys experimentals. Aquest test preliminar és menys costos que estudis col·laboratius i permet ajustar el mètode abans de realitzar l'estudi col·laboratiu.⁵

El test de robustesa de Youden es pot realitzar durant l'optimització de la tècnica utilitzada o durant la preparació de la mostra.

Altres mètodes aprovats es poden utilitzaren aquest punt. Però l'aproximació de Youden-Steiner, manté l'esforç i el temps requerits al mínim. Aquesta robustesa segueix un disseny factorial fraccionat, per tant, interaccions entre diferents factors no es poden detectar.

Per a fer l'avaluació de la robustesa del mètode, és necessari calcular la desviació estàndard dels mètodes de referència utilitzats. Per tal de dur a terme aquest càlcul es selecciona unes 7 variables i es determina de cada variable un límit superior i inferior d'aquesta.

Les 7 variables escollides són 3 que formen part dels paràmetres intern de l'equip del NIR (temperatura de mesura, resolució i número de scans per mesura de la mostra), 2 variables fan referència a la fusió de la mostra quan s'ha solidificat (temperatura i temps de fusió) i les dos últimes variables fan referència a la producció i envasament de la mostra, (presència de KOH i aigua en la mostra), durant la fabricació del producte s'utilitza KOH 0.5M com a medi de reacció i durant l'envasat del producte final pot haver traces d'aigua, en un principi una mostra de procés no ha de contenir cap traça d'aigua.

A partir d'aquí es crea un disseny factorial fraccionat amb un total de 8 experiments que presenten cada una d'aquestes 7 variables d'estudi.

Un cop preparats aquests 8 experiment, es mesuren pel NIR i amb el mètode multivariant creat, s'obtenen els valors predits d'aquests experiments per cada determinació desitjada. Seguidament amb aquests resultats es realitza l'avaluació de la seva robustesa enfront de cada variable d'estudi.

6 RESULTATS

En aquest apartat explicarem els resultats obtingut en seguir tots els punts descrits en l'apartat anterior del procés experimental d'haver analitzat i desenvolupat 3 mètodes per a determinar l'índex de saponificació, l'Índex d'hidroxilació i el punt de terbolesa de l'Emulsogen HCO 040 i 060, que és un producte resultant de fer reaccionar una glicerina amb oli de Ricí i posteriorment polimeritzar aquesta molècula amb òxid d'etilè que reacciona amb el grup hidroxil lliure de la part estructural de l'oli de Ricí.

Com s'ha explicat anteriorment per a desenvolupar els mètodes s'ha utilitzat l'algoritme PLS1, per tant, s'ha desenvolupat un mètode per separat de cada component on es volia determinar la concentració del analit, això significa que per a cada mètode hem repetit espectres que s'hagin utilitzat en les calibracions dels altres mètodes.

6.1 MÈTODES

Un cop recopilat i complimentat tots els valors de tots els espectres obtinguts, s'ha realitzat una optimització dels espectres (Figura - 7), eliminant els que estiguessin repetits o els que presentessin valors dels mètodes de referència equivocats i no es tinguessin les mostres físiques per a tornar a realitzar l'anàlisi.

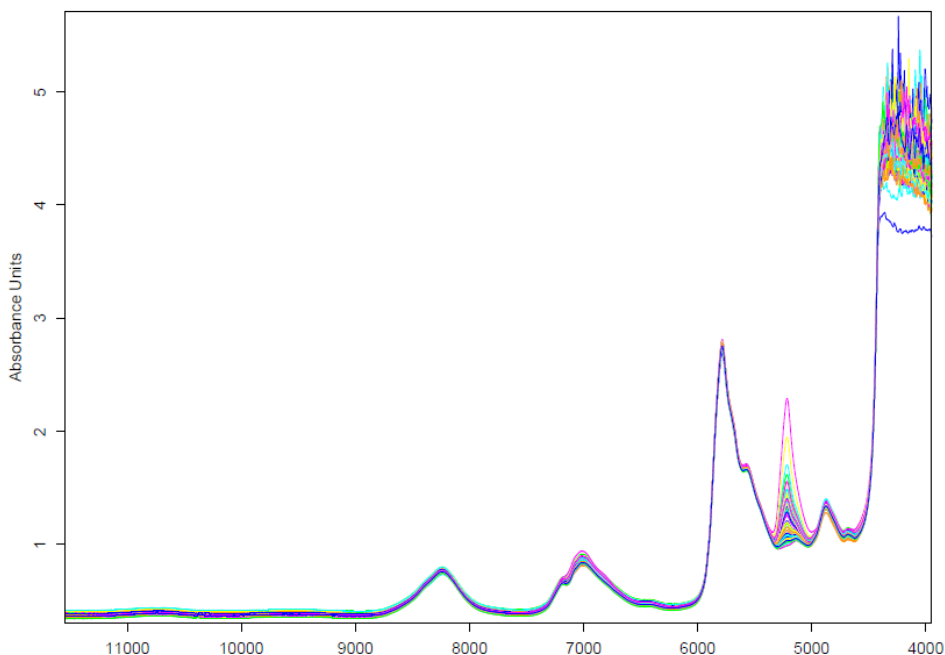


Figura - 7 Conjunt d'espectres analitzats i optimitzats

6.1.1 MÈTODE DEL ÍNDEX DE SAPONIFICACIÓ

Per a desenvolupar aquest mètode hem utilitzat un total de 90 mostres amb els seus respectius espectres.

La classificació d'aquestes mostres es pot observar en la següent (Taula -3):

Taula - 3 Classificació i número de les mostres finals i de procés per a la determinació de l'Índex de saponificació

FINAL	PROCÉS	TOTAL
51	39	90

Ahora de configurar els paràmetres d'optimització per la qualitat del mètode hem seleccionat tots els disponibles i hem tret del calibratge la zona espectral que presenta una absorbància total i la zona per sobre dels 9000 cm^{-1} ja que són zones on el soroll espectral presenta més influència.

Per començar s'ha realitzat una validació creuada per tal d'identificar els punts discrepants que el software reconeixia a partir del valor de la distància de Mahalanobis i els eliminaven del conjunt d'espectres.

Els resultat obtinguts duran la validació creuada (CV) es mostren en la (Taula - 4, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 4 Resultats del CV en el mètode per a la determinació de l'Índex de saponificació

VALIDACIÓ INTERNA									
anàlisis	R^2	RMSECV	RPD	Bias	Ranc	MD	Regions (cm^{-1})	Outliers	NOM
SAPO	93,12	0,959	3,82	-0,0353	7	0,54	8240-7504 6100-5448	040 ESD0031421 040 ESD0032698	HCO SAPO CV

Com es pot observar dos mostres s'han hagut d'excloure del conjunt d'espectres perquè el software els ha identificat com a punts discrepants i podien entorpir el conjunt de calibració.

Els dos punts discrepants trobats es tracten de dues mostres de producte finals del Emulsogen HCO 040.

Aclarir que no sempre l'eliminació de punts discrepants porta a un bon model de calibració, per això s'ha de fixar en els paràmetres estadístics resultants d'excloure del conjunt de calibració aquests punts discrepants detectats. En aquest cas eliminant-los s'ha obtingut millors resultats.

Les regions de l'espectre (Annex - 1) que el software ha seleccionat que presentaven un menor valor de RMSECV correspon a la zona espectral on hi ha presència de bandes del primer sobretó del metilè (CH_2), del metil (CH_3) i del primer sobretó d'alquens.

És curiós que les zones espectrals on el software ha detectat més canvis estructurals de l'espectre no corresponen a les bandes de combinació i del primer sobretó de l'èster del producte, ja que hauria de ser la zona espectral on s'hauria de veure més afectada la seva estructura, tenint en compte que l'Índex de saponificació es veu afectat pel número d'èsters formats.

Seguidament de forma aleatòria s'ha seleccionat un conjunt d'espectres que representessin el 20% del total d'espectres i s'ha realitzat la validació externa del mètode calibrat.

Els resultat obtinguts duran la validació externa (VE) es mostren en la (Taula - 5, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 5 Resultats del VE en el mètode per a la determinació de l'índex de saponificació

VALIDACIÓ EXTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
SAPO	95,65	0,845	5,34	-0,372	8	0,61	8972-5448		HCO SAPO VE 1
SAPO	95,72	0,813	5,23	-0,311	10	0,63	8972-5448	060 ESD0029965	HCO SAPO VE 2
SAPO	90,47	0,801	3,4	-0,243	8	0,55	8972-6100	060 ESD0029965	HCO SAPO VE 3
SAPO	92,12	0,744	3,65	-0,164	8	0,44	4604-4440	060 ESD0029965 040 E11 M2	HCO SAPO VE 4
SAPO	93,94	0,673	4,07	0,0376	7	0,41	4604-4440	060 ESD0029965 040 E11 M2 040 D14 M2	HCO SAPO VE 5
SAPO	98,01	0,595	7,68	0,23	5	0,45	8972-7504 6100-5448	060 ESD0029965 040 E11 M2 040 D14 M2	HCO SAPO VE 6
SAPO	97,49	0,649	6,43	-0,123	3	0,2	8972-7504 6100-5448 4604-4440	060 ESD0029965 040 E11 M2 040 D14 M2 040 E9 M2	HCO SAPO VE 7

Com es pot observar s'han realitzat diverses validació externes amb un conjunt de prova diferent i aleatori, per tant, també s'ha utilitzat un conjunt de calibració diferent en cada cas.

La regió espectral (Annex - 1) on el software ha detectat més variació estructural de l'espectre en el conjunt de prova és semblant al del CV però afegint la presència del segon sobretot d'alquens i del CH. També dir que no està la zona espectral de les bandes de combinació i sobretons de l'èster.

S'ha observat l'aparició de punts discrepant, els quals s'han eliminat per tal d'observar sense interferències com es desenvolupava cada conjunt de prova en cada mètode realitzat sense que hi haguessin espectres que poguessin interferir en aquests desenvolupament i poder obtenir d'aquesta manera els millors paràmetres estadístics possibles.

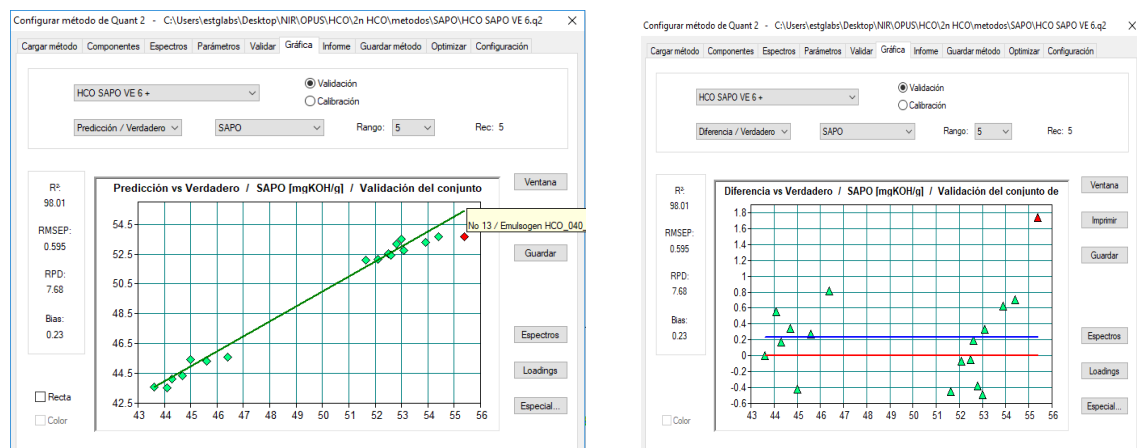


Figura - 8 - Gràfica de la comparació dels valors predits i reals en el mètode i de la seva diferència en la determinació de l'índex de saponificació

El mètode amb millors resultats és el que es pot observar que està nomenat com ha HCO SAPO VE 6. Hem escollit aquest mètode perquè es diferenciava de tots els altres sobretot en el número

de components principals que utilitzava que no era ni massa alt ni massa baix (Ranc=5). A més presenta la correlació més elevada ($R^2= 98.01$) i el valor de RPD més elevat (RPD= 7.68).

A més observant la gràfica de la diferència del valor predit amb el valor real (Figura - 8) podem observar que la línia vermella que marca el 0 i la línia blava que marca la mitjana de la diferència entre el valor predit i el real, els valors del conjunt de proba es distribueixen de forma homogènia per dalt i per baix d'aquestes dues línies, això ens indica que no hi ha una tendència ni factor de correcció en la predicció del model desenvolupat.

En aquest cas s'ha identificat un punt discrepant en el conjunt de proba, que si se l'eliminava s'obtenien pitjors resultats. Aquest és un exemple clar de que no sempre es bo eliminar els punts discrepants detectats sense una avaluació prèvia, ja que poden disminuir la capacitat de predicció del mètode.

Com s'ha explicat anteriorment, els valors obtinguts són aplicables en la majoria d'aplicacions analítiques de control de qualitat.

Per tal de ser més acurats en el moment d'escollir el millor mètode s'ha comprovat si únicament seleccionant el pic o pics que presenta una variació estructural de l'espectre com a zones espectrals per a la calibració del mètode obteníem millors resultats (Figura - 9).

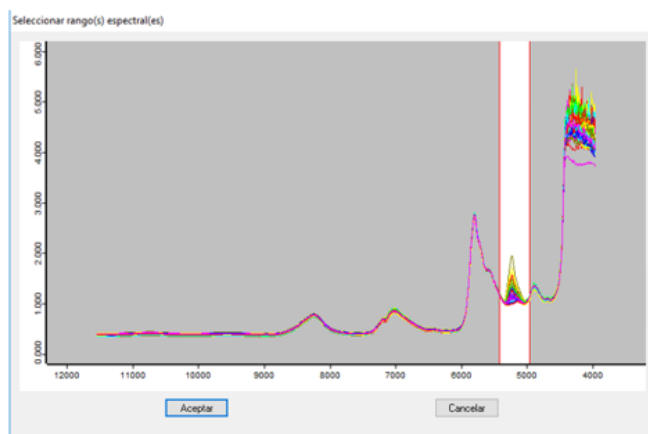


Figura - 9 Regió de l'espectre seleccionat

La zona espectral seleccionada correspon al pic on s'observava major variació estructural i correspon a la zona on hi ha presència de bandes del segon sobretó d'èsters i àcids carboxílics, i bandes de combinació de l'aigua.

Els resultat obtinguts durant la validació creuada (CV) optimitzant la zona espectral es mostren en la (Taula - 6, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 6 Resultats del CV optimitzant la zona espectral

VALIDACIÓ INTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
SAPO	91,81	1,03	3,49	-0,00196	8	0,45	5412-4952	040 ESD0031421 040 ESD0032698 060 ESD0029965	HCO SAPO CV OPT

Com es pot observar tant el valor de la correlació i el RPD resultants són inferiors a quan s'ha calibrat el mètode amb tota la zona espectral. A més el número de component principals és elevat i identifica un punt discrepant més que en l'anterior mètode.

En un primer cop d'ull aquest mètode és més restrictiu que l'anterior ja que presenta una distància de Mahalanobis inferior i identifica un major número de punts discrepant, però tot i així es realitza una validació externa, per tal d'observar quin comportament tenen mostres desconegudes, i ho realitzem de la mateixa forma que s'ha descrit anteriorment.

Els resultat obtinguts duran la validació externa (VE) optimitzant la zona espectral es mostren en la (Taula - 7, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R², bias, el ranc i la distància de Mahalanobis):

Taula - 7 Resultats del VE optimitzant la zona espectral

VALIDACIÓ EXTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
SAPO	97.33	0,652	6,12	-0,0101	9	0,59	5412-4952	040 D41 M2	HCO SAPO VE OPT 1
SAPO	97,13	0,675	5,91	0,028	9	0,63	5412-4952	040 D41 M2 040 E9 M2	HCO SAPO OPT 2

Havent realitzat els VE i eliminant els respectius punts discrepant que s'han detectat, es pot observar que s'obté pitjors paràmetres estadístics i sobretot el número de components principals que utilitza és massa elevat, per tant, aquest mètodes són massa sensibles al soroll espectral ja que gran part de la informació que utilitzen prové d'aquella zona espectral.

6.1.2 MÈTODE DEL ÍNDEX D'HIDROXILACIÓ

Per a desenvolupar aquest mètode s'ha seguit amb els mateixos passos que en el desenvolupament del Índex de saponificació, per tant, també s'ha començat escollint tota la zona espectral per tal de calibrar el mètode.

En aquest cas s'ha utilitzat 81 mostres per calibrar el mètode.

La classificació d'aquestes mostres es pot observar en la següent (Taula - 8):

Taula - 8 Classificació i número de les mostres finals i de procés per a la determinació de l'Índex d'hidroxilació

FINAL	PROCÉS	TOTAL
52	29	81

Els resultats obtinguts durant la validació creuada (CV) es mostren en la (Taula - 9 on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 9 Resultats del CV en el mètode per a la determinació de l'Índex d'hidroxilació

VALIDACIÓ INTERNA									
anàlisi	R^2	RMSECV	RPD	Bias	Ranc	MD	Regions (cm-1)	Outliers	NOM
OH	81,07	1,83	2,3	-0,0167	2	0,19	8944-7504 5452-4496	060 ESD0026274 060 E3 M2	HCO OH CV

En aquest cas el software ha identificat com a punt discrepant dues mostres del Emulsogen HCO 060, una de producte final i l'altre de producció, els quals els seus espectres podien entorpir el bon desenvolupament del mètode, per tant s'han eliminat del conjunt d'espectres per a realitzar la calibració.

Com s'ha explicat anteriorment no sempre és recomanable eliminar del conjunt de calibració tots els punt discrepant detectats, però en aquest cas també s'ha obtingut millors paràmetres estadístics, resultat d'eliminar els punts discrepant detectats.

La zona espectral (Annex - 1) que el software ha detectat més variació estructural de l'espectre correspon a la zona on hi ha presència de bandes del segon sobretó d'èsters i àcids carboxílics, de bandes de combinació de l'aigua, del segon sobretó del metil (CH_3), del segon sobretó del CH i de la banda de combinació de l'alcohol lliure.

En aquest cas el software si que ha seleccionat les zones espectrals d'interès com seria la del alcohol lliure, ja que en la determinació de l'índex d'hidroxilació, aquest valor es veu afectat pel número d'alcohols lliures que hi ha presents en el producte.

S'ha d'anar amb compte perquè també agafa la zona de la banda de combinació de l'aigua i pot afectar en la predicció del valor, perquè l'aigua també presenta enllaç OH.

Seguidament s'ha realitzat la validació externa tal i com s'ha descrit anteriorment.

Els resultats obtinguts durant la validació externa (VE) es mostren en la (Taula - 10, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 10 Resultats del VE en el mètode per a la determinació de l'Índex d'hidroxilació

VALIDACIÓ EXTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
OH	91,28	1,66	3,7	0,669	1	0,14	8944-8224 4976-4496		HCO OH VE 1
OH	92,84	1,55	3,93	0,475	3	0,39	5452-4496	060 D02 M1	HCO OH VE 2
OH	91,73	1,43	3,91	0,652	3	0,3	8944-6100 5452-4600	060 D02 M1	HCO OH VE 4
OH	92,69	1,23	4,05	-0,498	2	0,19	5452-4600	060 D02 M1	HCO OH VE 5
OH	94,93	1,06	4,63	-0,298	2	0,21	5452-4496	060 D02 M1 040 ESD0031513	HCO OH VE 6
OH	97,58	0,673	6,68	-0,18	5	0,39	4604-4548	060 D02 M1 040 ESD0031513 060 E02 M2	HCO OH VE 8
OH	98,1	0,597	7,69	-0,2	7	0,47	4604-4548	060 D02 M1 040 ESD0031513 060 E02 M2	HCO OH VE 9
OH	98,07	0,64	7,54	0,191	2	0,24	7500-6100 4604-4496	060 D02 M1 040 ESD0031513 060 E02 M2 060 D03 M2 060 D05 M1	HCO OH VE 9

Per a realitzar aquest mètode havíem de tenir en compte que les mostres que presentessin un percentatge d'aigua major al 0.5% es podia veure afectat a l'estructura de l'espectre en les zones que hi hagués la presència de les bandes de combinació i del sobretó de la vibració de l'enllaç O-H.

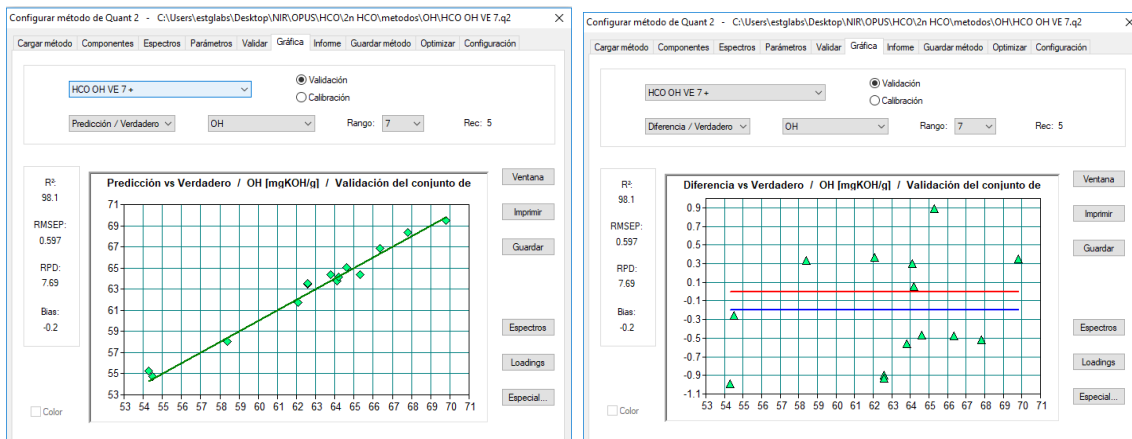


Figura - 10 Gràfica de la comparació dels valors predits i reals en el mètode i de la seva diferència en la determinació de l'Índex d'hidroxilació

En el mètode que el software ha seleccionat presenta la zona espectral on hi ha presència de la banda de combinació de l'alcohol lliure com ha zona on s'ha detectat més variació estructural de l'espectre. De tots els mètodes que es van calibrar s'ha escollit el que està nomenat com a HCO OH VE 8 perquè presenta el valor de correlació més elevat ($R^2 = 98.1$) i el valor de RPD també més elevat (RPD= 7.69), encara que el número de component principals que utilitza és força elevat (Ranc= 7), està compensat amb un valor de la distància de Mahalanobis més elevat que

la resta de mètodes, per tant, aquest mètode és més permissiu alhora d'identificar espectres com a punts discrepants.

Com s'ha explicat en l'anterior mètode, també en aquest podem observar una distribució homogènia de la diferència entre els valors predits i reals (Figura - 10) a dalt i baix de la línia blava, indicant que no hi ha cap tendència ni factor de correcció en els valors predits del model desenvolupat.

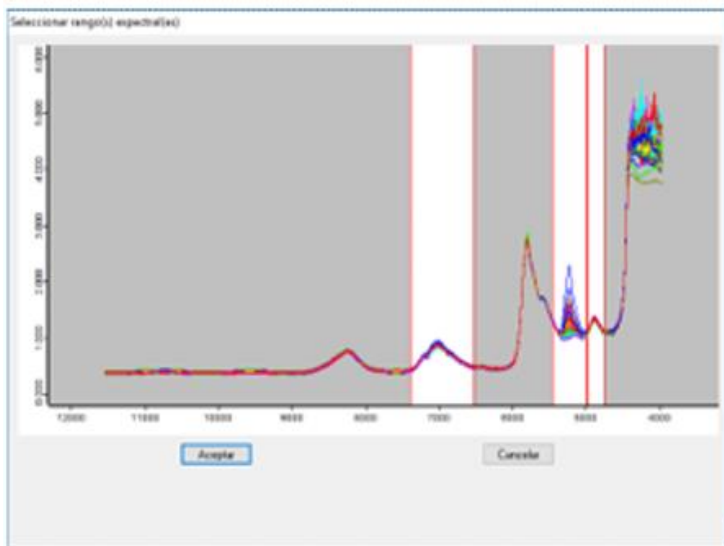


Figura - 11 Regió de l'espectre seleccionat

Tal i com s'ha realitzat anteriorment en el desenvolupament del mètode del Índex de saponificació, per ser més acurats alhora d'escollir el millor mètode també s'ha optimitzat la zona espectral utilitzada en la calibració del mètode (Figura - 11).

S'ha seleccionat les zones espectrals que presentaven major variació estructural en l'espectre, i corresponen a les zones on hi ha presència de bandes de combinació de l'alcohol lliure, de combinació de l'aigua, del segon sobretó d'èsters i àcids carboxílics i del primer sobretó de l'alcohol lliure.

Els resultats obtinguts durant la validació creuada (CV) optimitzant la zona espectral es mostren en la (Taula - 11, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 11 Resultats CV optimitzant la zona espectral

VALIDACIÓ INTERNA									
anàlisis	R^2	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
OH	79.29	1,92	2,2	0,0214	5	0,5	7388-6952 5452-4980 4968-4744	060 ESD0026274 060 E03 M2	HCO OH CV OPT

Com ha passat en el mètode del Índex de saponificació s'ha obtingut pitjors paràmetres estadístics, però en contrast s'utilitza un major nombre de components principals i el valor de la distància de Mahalanobis és major, tot i així, s'ha identificat els mateix nombre de punts discrepants.

Seguidament s'ha realitzat la validació externa.

Els resultat obtinguts duran la validació externa (VE) optimitzant la zona espectral es mostren en la (Taula - 12, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 12 Resultats VE optimitzant la zona espectral

VALIDACIÓ EXTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
OH	82,83	1,93	2,99	-1,14	3	0,32	5452-4980 4968-4744		HCO OH VE OPT 1
OH	85,29	2,16	2,79	0,777	3	0,28	7388-6520 5216-4980 4968-4856		HCO OH VE OPT 2

En aquest cas encara que no s'han identificat en cap cal cap punt discrepant en cap dels mètodes calibrat, els paràmetres estadístics resultants són molt pobres en comparació a quan s'ha utilitzat tota la zona espectral per a calibrar el mètode.

6.1.3 MÈTODE DEL PUNT DE TERBOLESA

En aquest cas s'ha utilitzat tota la zona espectral per calibrar el mètode i un total de 79 mostres.

La classificació d'aquestes mostres es pot observar en la següent (Taula - 13):

Taula - 13 Classificació i número de les mostres finals i de procés per a la determinació del punt de terbolesa

FINAL	PROCÉS	TOTAL
52	27	79

Els resultat obtinguts duran la validació creuada (CV) es mostren en la (Taula - 14, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 14 Resultats del CV en el mètode per a la determinació del punt de terbolesa

VALIDACIÓ INTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
TP	98,38	0,564	12,7	0,00788	7	0,63	8224-7504 5028-4600	060 E03 M2 040 ESD0026589 040 E11 M2 040 ESD0027665 060 D02 M1 040 E11 M2 040 ESD0031598 040 D18 M2 060 ESD0024302	HCO TP CV

En realitzar el CV dels espectres s'ha hagut d'eliminar un número elevat de punts discrepants que ha identificat el software per tal d'obtenir uns bons paràmetres estadístics.

Tot i així sense haver realitzat la validació externa amb un conjunt de proba de mostres desconegudes, s'ha obtingut uns valors de R^2 i RPD òptims per la utilització en qualsevol anàlisi de qualitat. Per tant, tot fa pensar que en la validació externa s'obtindrà uns millors resultats.

Una de les possibles raons per les quals en aquest model s'ha detectat molts punts discrepants, seria que la diferència dels valors reals als predits pel model de calibració difereixen molt, degut possiblement a que en aquest cas el mètode de referència utilitzat és molt subjectiu al criteri de l'analista que realitzi aquest anàlisi, ja que els valors reals utilitzats, molt d'ells, són dades recuperades, i en el laboratori hi ha 5 analistes que poden tenir un criteri diferent alhora de determinar el punt de terbolesa de la mostra.

Les zones espectrals (Annex - 1) que el software ha seleccionat que presentaven la major variació espectral, corresponen a les zones de l'espectre on hi ha presència de les bandes de combinació d'alquens i de combinació d'alcohols lliures.

Es pot dir que és curiós que no detectes el software variacions estructurals en l'espectre en la zona on hi ha presència de bandes del primer sobretó de l'èter, ja que la determinació del punt de terbolesa està estretament relacionada amb la relació molar d'òxid d'etilè que hi ha present al producte.

Els resultat obtinguts duran la validació externa (VE) es mostren en la (Taula 15, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 15 Resultats del VE en el mètode per a la determinació del punt de terbolesa

VALIDACIÓ EXTERNA									
anàlisis	R^2	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
TP	99,44	0,621	14	0,194	8	0,86	8944-7504 5452-4600		HCO TP VE 1
TP	99,37	0,648	12,9	0,139	7	0,81	8944-7504 5452-4600	040 E03 M2	HCO TP VE 2
TP	99,82	0,363	24	-0,0414	10	0,97	8944-7504 5452-4600	040 E03 M2	HCO TP VE 3
TP	99,81	0,369	23,5	-0,0657	10	0,97	8944-7504 5452-4600	040 E03 M2 060 D4 M1	HCO TP VE 4
TP	99,79	0,389	21,8	0,0293	10	0,94	8944-7504 6100-4600	040 E03 M2 060 D4 M1 060 E02 M2	HCO TP VE 5
TP	99,67	0,476	20,7	0,261	4	0,64	5452-4456	040 E03 M2 060 D4 M1 060 E02 M2	HCO TP VE 6

En el desenvolupament d'aquest mètode, com es preveia, s'han obtingut uns paràmetres estadístics molt bons en cada mètode realitzat i escollir quin és el millor pot resultar difícil.

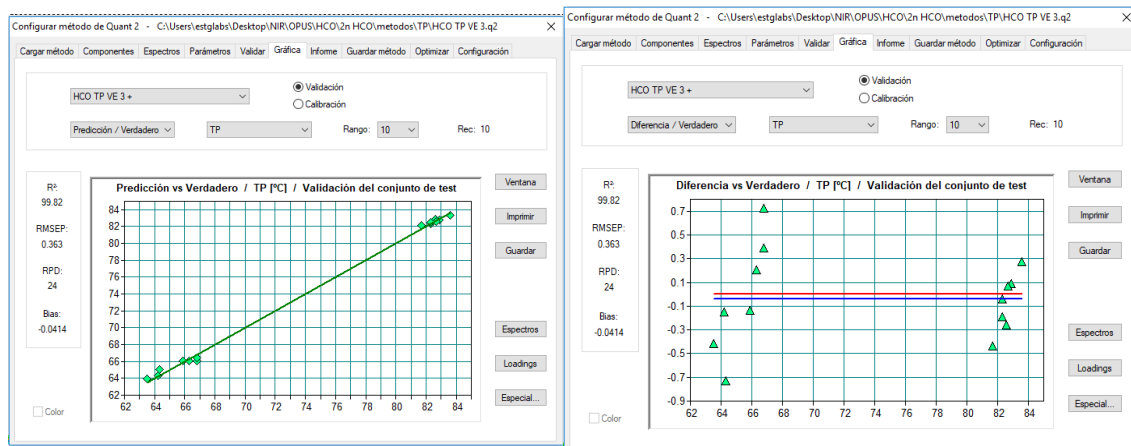


Figura - 12 Gràfica de la comparació dels valors predits i reals en el mètode i de la seva diferència en la determinació del punt de terbolesa

El mètode nomenat com HCO TP VE 3 és el que s'ha escollit com el millor mètode per aplicar en la predicció de la determinació del punt de terbolesa, degut a que presenta els millors paràmetres estadístics i encara que utilitza molts components principals (Ranc= 10) està compensat per un valor de la distància de Mahalanobis molt gran (MD=0.97) fent-lo molt permissiu alhora d'acceptar punts discrepants i a més sol s'ha necessitat eliminar un punt discrepant identificat per tal d'obtenir aquests resultats.

En aquest cas s'observa que en la gràfica de les diferències entre els valors reals amb els predits (Figura - 12), no hi ha una distribució homogènia, sinó que s'observa dues agrupacions als extrems, això es pot explicar pel que s'ha comentat abans, de que el mètode de referència utilitzat és molt subjectiu al criteri de l'analista, per tant, les diferències entre el valor predit i el real poden ser notòries.

A més en la gràfica dels valors predits enfront els reals es pot observar que hi ha dues agrupacions cosa que és clarament diferent en el que ha passat en els altres mètodes creats, podent fer més fàcil la creació de la recta. També com que la determinació presenta el rang d'especificació més gran fa més fàcil obtenir millors valors e calibració.

Tots aquests aspectes han fet que aquest mètode fos el més òptim.

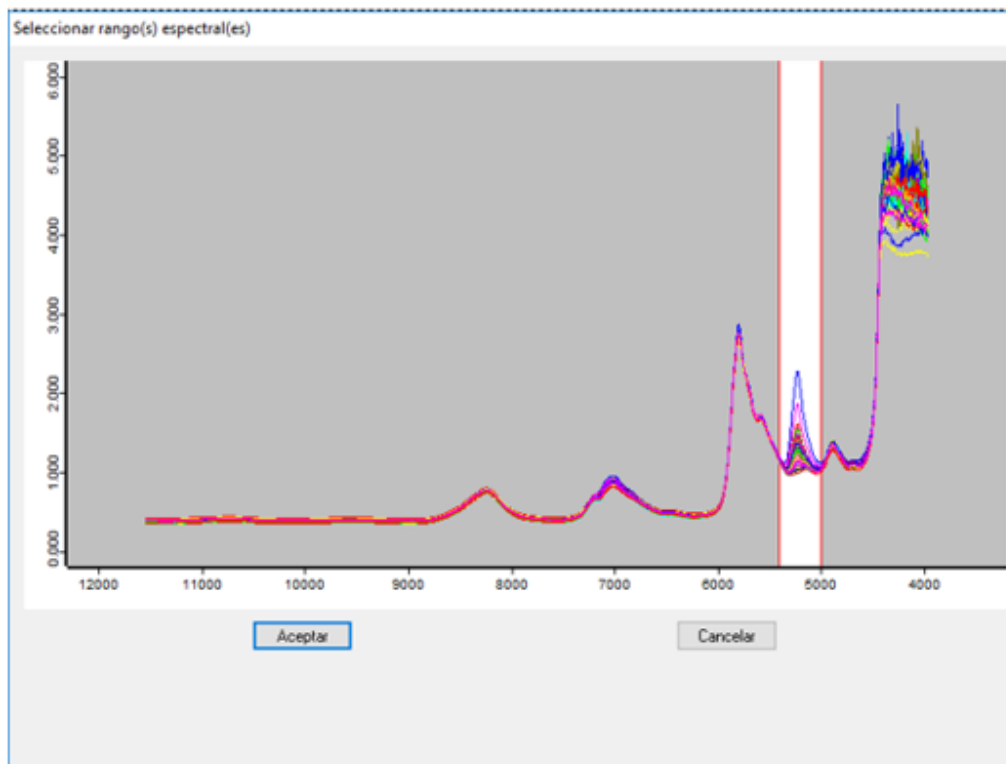


Figura - 13 Regió de l'espectre seleccionat

També, tal i com s'ha fet en els dos mètodes anteriors s'ha comprovat si els paràmetres estadístics dels mètodes calibrat, milloraven si es seleccionava les zones de l'espectre que presentessin canvis en la seva estructura en comptes d'utilitzar tot l'espectre.

La zona espectral seleccionada (Figura - 13 Regió de l'espectre seleccionat) és la del pic que s'observava major variació estructural en l'espectre i correspon a la zona on hi ha presència de bandes de combinació d'èsters i àcids carboxílics, i bandes de combinació de l'aigua.

Els resultat obtinguts durant la validació creuada (CV) optimitzant la zona espectral es mostren en la (Taula - 16, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 16 Resultats CV optimitzant la zona espectral

VALIDACIÓ INTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
TP	99,32	0,589	12,1	0,00226	7	0,69	5400-4992	060 E03 M2 040 ESD0026589 040 E11 M2 040 ESD0027665 060 D02 M1 040 E11 M2 040 ESD0031598 040 D18 M2 060 ESD0024302 040 E03 M2 040 D20 M2 060 ESD0026840 040 D19 M2	HCO TP CV

Com ha passat quan s'ha utilitzat tota la zona espectral per a realitzar la calibració, per ha obtenir bons resultats s'ha hagut d'eliminar un bon número de punts discrepants.

En aquest cas s'ha obtingut uns millors paràmetres estadístics pel que fa a la $R^2=99.32$ i la $MD=0.69$ en contrast al CV de tot l'espectre.

Els resultat obtinguts duran la validació externa (VE) optimitzant la zona espectral es mostren en la (Taula - 17, on es poden veure els diferents valors estadístics del mètode obtingut, com són el RPD, R^2 , bias, el ranc i la distància de Mahalanobis):

Taula - 17 Resultats VE optimitzant la zona espectral

VALIDACIÓ EXTERNA									
anàlisis	R ²	RMSECV	RPD	Bias	Ranc	MD	Regions (cm- 1)	Outliers	NOM
TP	99,64	0,464	19,5	-0,24	6	0,71	5400-4992		HCO TP OPT 1

Els resultats obtinguts són millors que en el CV però lleugerament pitjors que el mètode seleccionat quan s'ha utilitzat tota la zona espectral, sobretot en el $RPD=19.5$ que és molt més baix que el $RPD=24$ aconseguit anteriorment, encara que utilitza menys components principals que l'altre mètode, no és tan permissiu en quan acceptar o no un punt discrepant identificat ja que el seu valor de MD és menor ($MD=0.71$).

6.2 ROBUSTESA DE YOUNDEN-STEINER

S'ha realitzat el test de la robustesa en els dos tipus de producte els quals estem desenvolupant el mètode (Emulsogen HCO 040 i 060), i per a cada determinació que presenta.

En aquest estudi (Taula - 18) s'ha seleccionat 7 variables que s'ha considerat que podrien afectar l'espectre. A continuació es mostra una taula amb aquestes variables i el disseny factorial fraccionat amb els 8 experiments que s'han realitzat per a cada producte i anàlisis que es volia fer

Taula - 18 Disseny factorial fraccionat segons Youden-Steiner

HCO											
CONDICIÓ VARIABLE				ANÀLISIS							
Tipus	Clau	Valor alt X	Valor baix x	1	2	3	4	5	6	7	8
%H ₂ O	A,a	2%	0%	2%	2%	2%	2%	0%	0%	0%	0%
Temperatura mesura	B,b	75°C	60°C	75°C	75°C	75°C	60°C	75°C	60°C	60°C	60°C
Número de scanners	C,c	32	8	32	8	32	8	32	8	32	8
Resolució	D,d	16	8	16	16	8	8	8	8	16	16
Amb KOH 0,5M	E,e	SI	NO	SI	NO	SI	NO	NO	SI	NO	SI
Temps de fusió	F,f	1h30min	45min	1h30min	45min	45min	1h30min	1h30min	45min	45min	1h30min
Temperatura de fusió	G,g	105°C	80°C	105°C	80°C	80°C	105°C	80°C	105°C	105°C	80°C

L'arrel quadrada de 2 multiplicada per la desviació estàndard del mètode és el valor de referència que s'utilitza per determinar si el factor realment afecta o no a l'anàlisi. Si la diferència entre les mitjanes dels límits inferiors i superiors de cada factor per al valor predit en el mètode desenvolupat és superior a aquest valor, això voldrà dir que aquest factor afecta a l'anàlisi, si es dona el cas contrari, voldrà dir que el mètode és robust contra aquest factor.

6.2.1 ROBUSTESA EN L'ÍNDEX DE SAPONIFICACIÓ

Per començar s'ha determinat la desviació estàndard del mètode de la determinació de l'Índex de saponificació de referència Taula - 19) per als dos productes.

Taula - 19 Valors de realitzar el mètode de referència per a la determinació de l'Índex de saponificació en l'Emulsoaen HCO 040 i 060

SAPO	SAPO
Emulsogen HCO 040 E/13 M2	Emulsogen HCO 060 ESD0025449
54,46	45,08
54,71	45,13
53,6	45,17
56,22	45,21
53,95	44,33
53,23	45,83
55,6	44,27
54,61	45,18

Per al HCO 040 la seva desviació estàndard resultant ha sigut de 0,995242253 i per al HCO 060 la seva desviació estàndard ha sigut de 0,506303128.

Taula - 20 Valors de la predicció en el mètode desenvolupat per a la determinació de l'Índex de saponificació dels 8 experiments per a l'Emulsogen HCO 040 i 060

Experiment	SAPO	
	VALORS PREDITS PEL MÈTODE	
	Emulsogen HCO 040 E/18 M2	Emulsogen HCO 060 ESD0025449
1	48,375	38,583
2	49,526	43,435
3	53,597	46,582
4	54,809	47,806
5	53,711	45,252
6	45,249	53,935
7	48,28	39,481
8	47,695	39,719

Un cop s'ha preparat els 8 experiment es punxen en l'equip del NIR i es passen pel mètode desenvolupat i d'aquesta forma s'obté els valors predits per estudiar la robustesa (Taula - 20).

Taula - 21 Valoració de la robustesa del mètode per a la determinació de l'Índex de saponificació en l'Emulsogen HCO 040

Emulsogen HCO 040 E/18 M2					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	X-x (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	51,577	48,734	2,843	Sensible a variable
B	b	51,302	49,008	2,294	Sensible a variable
C	c	50,991	49,320	1,671	Sensible a variable
D	d	48,469	51,842	3,373	Sensible a variable
E	e	48,729	51,582	2,853	Sensible a variable
F	f	51,148	49,163	1,984	Sensible a variable
G	g	49,178	51,132	1,954	Sensible a variable

Com es pot observar per al Emulsogen HCO 040 el mètode desenvolupat no és gaire robust, ja que és sensible a tots els factors estudiats (Taula – 21).

Pel que fa a les variables B, C i D que fan referència a les variables internes de l'equip del NIR (Temperatura de mesura, número de scans i resolució), no porten cap problema perquè aquestes condicions estan fixades, el problema esdevé en les variables A, E, F i G (%H₂O, presència de KOH 0.5M, temps de fusió i temperatura de fusió), que caldrà anar amb compte alhora d'analitzar.

Taula - 22 Valoració de la robustesa del mètode per a la determinació de l'Índex de saponificació en l'Emulsogen HCO 060

Emulsogen HCO 060 ESD0025449					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	X-x (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	44,102	44,597	0,495	No sensible a variable
B	b	43,463	45,235	1,772	Sensible a variable
C	c	42,475	46,224	3,749	Sensible a variable
D	d	40,305	48,394	8,089	Sensible a variable
E	e	44,705	43,994	0,711	No sensible a variable
F	f	42,840	45,858	3,018	Sensible a variable
G	g	44,951	43,747	1,204	Sensible a variable

Pel que fa a l'Emulsogen HCO 060 és robust en el %H₂O i quantitat de KOH 0.5M que presenta la mostra i s'haurà d'anar amb compte alhora d'analitzar amb el temps i temperatura de fusió de la mostra, ja que és sensible (Taula - 22). Pel que fa a les variables internes de l'equip no fa falta que es vagi amb compte perquè sempre són fixes.

6.2.2 ROBUSTESA EN L'ÍNDEX DE HIDROXILACIÓ

Per començar s'ha determinat la desviació estàndard del mètode de la determinació de l'Índex de hidroxilació de referència per als dos productes (Taula - 23).
Taula - 23 Valors de realitzar el mètode de referència per a la determinació de l'Índex d'hidroxilació en l'Emulsogen HCO 040 i 060

OH	OH
Emulsogen HCO 040 E/14 M2	Emulsogen HCO 060 ESD0025630
67,82	62,25
67,9	62,35
67,47	61,86
67,47	61,74
67,53	61,69
65,67	62,13
67,35	61,42
67,02	
67,85	

Per al HCO 040 la seva desviació estàndard resultant ha sigut de 0,686053529 i per al HCO 060 la seva desviació estàndard ha sigut de 0,335857112.

Un cop s'ha preparat els 8 experiment es punxen en l'equip del NIR i es passen pel mètode desenvolupat i d'aquesta forma s'obté els valors predits (Taula - 24) per estudiar la robustesa.

Taula - 24 Valors de la predicció en el mètode desenvolupat per a la determinació de l'Índex d'hidroxilació dels 8 experiments per a l'Emulsogen HCO 040 i 060

Experiment	OH	
	VALORS PREDITS PEL MÈTODE	
	Emulsogen HCO 040 E/18 M2	Emulsogen HCO 060 ESD0025449
1	64,626	66,983
2	69,554	59,184
3	71,327	64,889
4	61,425	51,621
5	74,141	60,589
6	59,672	66,917
7	77,359	67,598
8	75,493	66,915

Taula - 25 Valoració de la robustesa del mètode per a la determinació de l'índex d'hidroxilació en l'Emulsogen HCO 040

Emulsogen HCO 040 E/18 M2					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	X-x) (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	66,733	71,666	4,933	Sensible a variable
B	b	69,912	68,487	1,425	Sensible a variable
C	c	71,863	66,536	5,327	Sensible a variable
D	d	71,758	66,641	5,117	Sensible a variable
E	e	67,780	70,620	2,840	Sensible a variable
F	f	68,921	69,478	0,557	No sensible a variable
G	g	65,771	72,629	6,858	Sensible a variable

Per al Emulsogen HCO 040, el mètode és robust per a la variable temps de fusió, però és sensible (Taula - 25 a la resta de variables, pel que fa a les variables internes de l'equip, no hi ha problemes perquè són fixes, però s'haurà d'anar en compte quan s'analitzi una mostra amb el %H₂O, la quantitat de KOH 0.5M present i la temperatura amb que es fon la mostra.

Taula - 26 Valoració de la robustesa del mètode per a la determinació de l'índex d'hidroxilació en l'Emulsogen HCO 060

Emulsogen HCO 060 ESD0025449					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	X-x) (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	60,669	65,505	4,836	Sensible a variable
B	b	62,911	63,263	0,352	No sensible a variable
C	c	65,015	61,159	3,856	Sensible a variable
D	d	65,170	61,004	4,166	Sensible a variable
E	e	66,426	59,748	6,678	Sensible a variable
F	f	61,527	64,647	3,120	Sensible a variable
G	g	63,280	62,894	0,386	No sensible a variable

En l'Emulsogen HCO 060, el mètode és robust pel que fa a la temperatura de mesura i la temperatura de fusió (Taula - 26). Com s'ha dit anteriorment les variables internes del equip de mesura, com que són fixes, no és necessari anar amb compte amb elles, en canvi s'ha de posar atenció alhora de mesurar una mostra amb el %H₂O, la quantitat de KOH 0.5M present a la mostra i el temps de fusió perquè el mètode és sensible a aquestes variables.

6.2.3 ROBUSTESA EN EL PUNT DE TERBOLESA

Per començar s'ha determinat la desviació estàndard del mètode de la determinació de l'índex de hidroxilació de referència per als dos productes (Taula - 27).

Taula - 27 Valors de realitzar el mètode de referència per a la determinació del punt de terbolesa en l'Emulsogen HCO 040 i 060

TP	TP
Emulsogen HCO 040 E/13 M1	Emulsogen HCO 060 ESD0025449
66,3	69,2
66,5	69,2
66,4	69,2
66,3	69,3
66,5	69,2
66,4	69,2
66,4	69,3
66,4	69,2

Per al HCO 040 la seva desviació estàndard resultant ha sigut de 0,075592895 i per al HCO 060 la seva desviació estàndard ha sigut de 0,046291005.

Taula - 28 Valors de la predicció en el mètode desenvolupat per a la determinació del punt de terbolesa dels 8 experiments per a l'Emulsogen HCO 040 i 060

Experiment	TP	
	VALORS PREDITS PEL MÈTODE	
	Emulsogen HCO 040 E/18 M2	Emulsogen HCO 060 ESD0025449
1	66,797	85,328
2	66,69	82,75
3	69,296	83,911
4	65,145	83,76
5	65,326	81,196
6	80,721	63,579
7	65,435	83,032
8	66,181	83,007

Un cop s'ha preparat els 8 experiment es punxen en l'equip del NIR i es passen pel mètode desenvolupat i d'aquesta forma s'obté els valors predits per estudiar la robustesa (Taula - 28).

Taula - 29 Valoració de la robustesa del mètode per a la determinació del punt de terbolesa en

Emulsogen HCO 040 E/18 M2					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	$\Delta(X-x)$ (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	66,982	69,416	2,434	Sensible a variable
B	b	67,027	69,371	2,343	Sensible a variable
C	c	66,714	69,684	2,971	Sensible a variable
D	d	66,276	70,122	3,846	Sensible a variable
E	e	70,749	65,649	5,100	Sensible a variable
F	f	65,862	70,536	4,673	Sensible a variable
G	g	69,525	66,873	2,651	Sensible a variable

Pel que es pot observar en la taula, l'Emulsogen HCO pel mètode de la determinació del punt de terbolesa és sensible a totes les variables estudiades (Taula - 29), per tant, exclouent aquelles variables que són internes de l'equip del NIR perquè són fixes, s'haurà d'anar amb compte amb les variables de %H₂O, quantitat de KOH 0.5M i, temperatura i temps de fusió, perquè el mètode és sensible a aquestes variables.

Taula - 30 Valoració de la robustesa del mètode per a la determinació del punt de terbolesa en l'Emulsogen HCO 060

Emulsogen HCO 060 ESD0025449					
CONDICIÓ VARIABLE		RESULTATS		DIFERÈNCIA	COMPARACIÓ
Valor alt X	Valor baix x	Mitjana X (mgKOH/g)	Mitjana x (mgKOH/g)	$\Delta(X-x)$ (mgKOH/g)	$\Delta < \sqrt{2} s$
A	a	83,937	77,704	6,234	Sensible a variable
B	b	83,296	78,345	4,952	Sensible a variable
C	c	83,367	78,274	5,093	Sensible a variable
D	d	83,529	78,112	5,418	Sensible a variable
E	e	78,956	82,685	3,728	Sensible a variable
F	f	62,654	78,318	15,664	Sensible a variable
G	g	78,925	82,716	3,791	Sensible a variable

Com en l'anterior cas, per a l'Emulsogen HCO 060, el mètode per a la determinació del punt de terbolesa no és gaire robust, ja que és sensible a totes les variables estudiades (Taula - 30). Com s'ha dit en el cas de l'Emulsogen HCO 040, s'haurà d'anar amb compte amb les variables de %H₂O, quantitat de KOH 0.5M i, temperatura i temps de fusió, perquè el mètode és sensible a aquestes variables i les altres variables al ser internes de l'equip del NIR i són fixes no seran cap inconvenient alhora de mesurar una mostra.

7 CONCLUSIONS

En aquest punt en que s'ha desenvolupat els tres mètodes i se'ls ha validat tant interna com externament, es pot concloure que s'ha desenvolupat tres bons mètodes amb l'algoritme PLS1 amb una bona capacitat predictiva de les determinacions de l'Índex de saponificació, l'Índex d'hidroxilació i el punt de terbolesa. Aquestes afirmacions venen amb la credibilitat que aporten els paràmetres estadístics que calcula el software OPUS i demostres el bon desenvolupament del conjunt de prova en la validació del mètodes desenvolupats.

De tots els mètodes que s'han creat el que presenta més dubtes seria l'escollit per al punt de terbolesa, ja que utilitza un número de components principals elevat. Una forma de solucionar això, és escollir un dels altres mètodes per a la determinació del punt de terbolesa amb paràmetres estadístics més baixos en el número de components principals però que presentin un bon número de RPD i de R^2 i veure amb un conjunt de prova independent al utilitzat quina capacitat de predicció presenta.

En el cas de la robustesa dels tres mètodes, poden concloure que no són gaire robustos enfront de les diferents variables estudiades, pel que fa a les variables internes de l'equip del NIR, com que són fixes, no són problema pels mètodes creats.

També cal afegir que en l'anàlisi de rutina de les diferents mostres que arribin al laboratori, les variables del temps i temperatura de fusió no són important perquè les mostres que arriben al laboratori ja venen foses.

S'haurà de posar important importància i anar amb precaució si la mostra presenta restes de KOH 0.5M, ja que s'utilitza com a medi de reacció i si hi ha un petit percentatge d'aigua en la mostra perquè aquestes dues variables són sensibles en els mètodes desenvolupats per a la determinació del Índex de saponificació i hidroxilació, i del punt de terbolesa.

Cal aclarir que les variables de $\%H_2O$ i KOH són factors que rarament es trobaran en les mostres tant de rutina com de producte final, ja que les variabilitats que es poden tenir un número elevat de mostres a l'hora d'elaborar els mètodes s'ha tingut presents. Però si que és cert que això ens pot servir de cara a mirar si en un altre regió de l'espectre que no estiguin presents els grups OH, es pot desenvolupar el mètode amb resultats semblants i que faci més robust a aquests paràmetres.

L'únic mètode que no presenta problemes en aquestes variables seria el del Índex de saponificació per al producte Emulsogen HCO 060, ja que no presenta aquesta sensibilitat vista en els altres mètodes.

8 CONCLUSIONS

At the end of the project it has been developed three methods and it has been validated internally and externally. It can be concluded that we have seen developed three good methods with the PLS1 algorithm with a good predict capacity to determine the saponification value, the hydroxylation value and the turbidity point. These results go with the reliability from the statistics parameters obtained using the software OPUS and it can be demonstrating the good development of the set test in the validation.

From all the methods created, the turbidity point method presents doubts because it has a great number of factors. One possible solution to solve this problem it can be chosen another method with a bit worse statistics parameter, but with good RPD and R^2 and see if with an independent test set verify its predicted capacity.

In the robustness validation, it can be concluding that three methods are not too robust versus studied factors. The internal factors from the NIR spectroscopy equipment they do not represent a problem, since they are fixed in the software.

On the other hand, the fusion temperature and time factors do not represent a problem, sine in the routine analysis the different samples are always in liquid state

We should pay attention to %H₂O and KOH factors because there are sensible in all three methods, but in other hands this factors its seen rarely in the routine samples because the different variabilities were taken into account during method creation. All of this factors are useful for detecting another spectral region where the OH bands are not present and develop a new method with similar results with more robustness in this factors.

The only method with robustness front almost all factors is the method to determine the saponification index in Emulsogen HCO 060.

9 BIBLIOGRAFIA

- ¹ BrukerOptics_IR-Tutorial; Studio für Mediendesign: B.Spieldenner Dipl.Ing. (FH), 2004
- ² Pasquini, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. *J. Braz. Chem. Soc.* **2003**, 14 (2), 198–219.
- ³ Kadenkin, A. Metrohm White Paper Near-Infrared Spectroscopy : Quantitative Analysis According to ASTM E1655 Metrohm White Paper. 1–7.
- ⁴ Bruker. A Practical Guide for the Method Development in the Analytical Chemistry. **2006**, 120.
- ⁵ AC-TH035 – Índice de Hidroxilo;MA035-V6.doc; Método interno QC/QA Clariant Tarragona, 2018
- ⁶ Fernandes, E.; Vicente, J.; Cruz, R.; Nele, M.; Nunes, É.; Costa-Félix, R. A Methodology Free of Pyridine for Measuring Hydroxyl Value in Polymers and Oils. *Polym. Test.* **2014**, 33 (February), 16–20.
- ⁷ AC-TH007 – Índice de saponificación; MA007-V4.doc; Método interno QC/QA Clariant Tarragona, 2018
- ⁸ AC-TH005 – Punto de turbidez; MA005-V3.doc; Método interno QC/QA Clariant Tarragona, 2000
- ⁹ Laboratorio De Calidad, instrucciones de uso del NIR – TANGO(BRUKER);No.TA-INCE-34, Laboratorio De Calidad,Clariant, Tarragona, **2019**.
- ¹⁰ Karageorgou, E.; Samanidou, V. Youden Test Application in Robustness Assays during Method Validation. *J. Chromatogr. A* **2014**, 1353, 131–139.

ANNEX

Annex - 1 Taula de freqüències del NIR

Overtone, combination	Wavelength/nm		Wavenumber/cm ⁻¹		Remarks
(1) Groups containing only C and H atoms					
1) -CH ₃ methyl					
combination	2275	2285	4400	4380	CH stret.+CH bend.
	1355	1365	7380	7330	2xCH stret.+CH bend.
	1010	1020	9900	9800	2xCH stret.+3xCH bend.
first overtone	1710	1730	5850	5780	first overtone of asym.stret.
	1770	1785	5650	5600	first overtone of sym.stret.
second overtone	1150	1165	8700	8580	second overtone of asym.stret.
	1190	1200	8400	8330	second overtone of sym.stret.
third overtone	870	885	11490	11300	third overtone of asym.stret.
	900	910	11110	10990	third overtone of sym.stret.
2) -CH ₂ methylene					
combination	2320	2330	4310	4290	CH stret.+CH bend.
	2305	2315	4340	4320	
	1410	1420	7090	7040	2xCH stret.+CH bend.
	1390	1400	7190	7140	
	1050	1060	9520	9430	
first overtone	1735	1750	5760	5710	first overtone of anti-sym.stret.
	1780	1795	5620	5570	first overtone of sym.stret.
second overtone	1170	1180	8550	8470	second overtone of anti-sym.stret.
	1200	1210	8330	8260	second overtone of sym.stret.
third overtone	885	895	11300	11170	third overtone of anti-sym.stret.
	910	920	10990	10870	third overtone of sym.stret.
3) -CH					
first overtone	1755	1775	5700	5630	
second overtone	1185	1195	8440	8370	
third overton	900	910	11110	10990	
4) C=C alkenes (vinyl, vinylidine, vinylene)					
combination	2340	2350	4270	4260	CH ₂ stret.+=CH ₂ bend.
	2185	2195	4580	4560	CH ₂ stret.+C=C stret.
	2135	2145	4680	4660	=CH stret.+C=C stret.
first overtone	1675	1695	5970	5900	
	1645	1660	6080	6020	vinyl group
second overtone	1130	1145	8850	8730	
	1110	1120	9010	8930	vinyl group
third overtone	860	870	11630	11490	
	840	850	11900	11760	vinyl group
5) -C≡CH alkynes (ethynyle)					
first overtone	1535	1545	6510	6470	
second overtone	1035	1045	9660	9570	
third overtone	780	790	12820	12660	
6) CH (aromatic)					
combination	1440	1450	6940	6900	2xCH stret.+CH bend.
	1410	1420	7090	7040	2xCH stret.+CH bend.
	1070	1085	9350	9220	2xCH stret.+2xC-C stret.
first overtone	1680	1690	5950	5920	
second overtone	1130	1140	8850	8770	
third overtone	850	860	11760	11630	
(2) O atom containing groups					
7) H ₂ O					
combination	1930	1940	5180	5150	OH stret.+OH bend.
	1375	1385	7270	7220	OH anti-sym. stret.+OH sym. stret.
first overtone	1450	1460	6900	6850	
second overtone	975	985	10260	10150	
third overtone	740	750	13510	13330	
8) free -OH alcohol					
combination	2060	2090	4850	4780	OH stret.+OH bend.
first overtone	1395	1425	7170	7020	
second overtone	2370	2390	4220	4180	second overtone of OH bend.
	940	955	10640	10470	second overtone of OH stret.
third overtone	730	745	13700	13420	

Overtone, combination	Wavelength/nm		Wavenumber/cm ⁻¹		Remarks
9) bound - OH alcohol					
first overtone	1435	1480	6970	6760	intermolecular hydrogen bond
	1500	1595	6670	6270	intramolecular hydrogen bond
second overtone	980	990	10200	10100	intermolecular hydrogen bond
	1035	1045	9660	9570	intramolecular hydrogen bond
10) COOH carboxylic acids, COOR esters					
second overtone	1890	1920	5290	5210	2x C=O stret. (carboxylic acids)
	1930	1950	5180	5130	2x C=O stret. (esters)
11) C=O ketones					
second overtone		1950		5130	
12) CHO aldehydes					
combination	2190	2210	4570	4520	CH stret.+C=O stret.
13) epoxides					
first overtone	1640	1650	6100	6060	first overtone of CH stret.
(3) N atom containing groups					
14) -NH₂ primary amines					
combination	1970	2010	5080	4980	NH stret.+NH bend.
first overtone	1520	1540	6580	6490	first overtone of NH ₂ sym.stret.
	1500	1520	6670	6580	first overtone of NH ₂ anti-sym.stret.
	1450	1480	6900	6760	ArNH ₂
second overtone	1020	1040	9800	9620	second overtone of NH ₂ sym.stret.
	1000	1020	10000	9800	second overtone of NH ₂ anti-sym.stret.
	980	1020	10200	9800	ArNH ₂
third overtone	800	820	12500	12200	third overtone of NH ₂ sym.stret.
	780	800	12820	12500	ArNH ₂
	770	790	12990	12660	third overtone of NH ₂ anti-sym.stret.
15) >NH secondary amines					
first overtone	1490	1545	6710	6470	
second overtone	1010	1040	9900	9620	
16) -CONH₂ primary amides					
combination	2140	2170	4670	4610	2x amide+amide
	2100	2130	4760	4690	NH stret.+amide
	2040	2060	4900	4850	NH stret.+amide
	1950	1970	5130	5080	NH stret.+amide
first overtone	1600	1620	6250	6170	intermolecular hydrogen bond
	1510	1530	6620	6540	intermolecular hydrogen bond
	1490	1510	6710	6620	first overtone of NH ₂ sym.stret.
	1440	1460	6840	6850	first overtone of NH ₂ anti-sym.stret.
second overtone	2020	2040	4950	4900	second overtone of amide
	1070	1090	9350	9170	intramolecular hydrogen bond
	1015	1035	9850	9660	intramolecular hydrogen bond
	1000	1020	10000	9800	second overtone of NH ₂ sym.stret.
	970	990	10310	10100	second overtone of NH ₂ anti-sym.stret.
17) -CONH- secondary amides					
combination	2150	2170	4650	4610	2 x amide+amide
	2100	2120	4760	4720	NH stret.+amide
	1990	2010	5030	4980	NH stret.+amide
first overtone	1530	1670	6540	5990	hydrogen bond
	1460	1510	6850	6620	free
second overtone	1910	1930	5240	5180	second overtone of amide
	1035	1120	9660	8930	hydrogen bond
	1000	1050	10000	9520	free
(4) S or P atom containing group					
18) -SH (thiols)					
first overtone	1735	1745	5760	5730	
19) P-OH (phosphorus acids, phosphonic acids, phosphinic acids)					
first overtone	1900	1910	5260	5240	first overtone of OH stret.
20) PH (phosphines, phosphorus acid esters)					
first overtone	1890	1900	5290	5260	