

**DETERMINACIÓN DE LA MASA MOLECULAR PROMEDIO DE FONDOS DE  
VACÍO DE CRUDOS COLOMBIANOS UTILIZANDO ESPECTROMETRÍA DE  
MASAS LDI TOF/TOF**

**KAREN JULIETTE VEGA RAMIREZ**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE QUÍMICA  
BUCARAMANGA**

**2014**

**DETERMINACIÓN DE LA MASA MOLECULAR PROMEDIO DE FONDOS DE  
VACÍO DE CRUDOS COLOMBIANOS UTILIZANDO ESPECTROMETRÍA DE  
MASAS LDI TOF/TOF**

**KAREN JULIETTE VEGA RAMIREZ**

**Trabajo de grado para optar al título de  
Química**

**Director**

**ENRIQUE MEJÍA OSPINO**

**Químico, Ph.D.**

**Codirectora**

**DIANA CATALINA PALACIO**

**Física, M.Sc.**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER**

**FACULTAD DE CIENCIAS**

**ESCUELA DE QUÍMICA**

**BUCARAMANGA**

**2014**

## DEDICATORIA

*Con todo mi amor para:*

*Mi madre Yaneth Ramírez, mi Padre Enrique Vega, mis hermanos Harvey, Katherine, Yari y Jhon, mi esposo Andrés Castillo y mis hijas Karoll Andrea y María Silvana, por ser la inspiración en mi vida con sus valiosas virtudes de alegría, unión, perseverancia en la fé y profundo amor de entrega; además, por permitirme encontrar en ustedes la fuerza y el soporte para llevar a término mi formación profesional y hacer un sueño realidad.*

*¡Los amo con todas las fuerzas de mi corazón!*

*Karen Juliette Vega Ramírez*

## **AGRADECIMIENTOS**

Mis sinceros agradecimientos al Doctor Enrique Mejía por su confianza y sus valiosos aportes para llevar a cabo este trabajo en el grupo de investigación LEAM.

A Catalina Palacio, por su disposición y entera colaboración, además por el compartir de sus valiosos conocimientos.

A Freddy Ramírez Pbro. por su hermosa y valiosa amistad con la que aprendí a reconocer en mi vida la luz y el amor que fortalecen el principio de cualquier proyecto a emprender.

A todos mis familiares, amigos y cercanos por su compañía, sus consejos y toda la ayuda brindada para hacer de mi formación profesional no sólo un proceso de aprendizaje, sino también, un encuentro en el sentir humano y en la esperanza de realizar los sueños apoyados en la unión.

A todos mis compañeros de estudio que de alguna manera me apoyaron y aconsejaron durante mi estancia en la Universidad.

## CONTENIDO

	pág.
INTRODUCCIÓN .....	14
1. CONSIDERACIONES TEÓRICAS .....	16
1.1 ASPECTOS GENERALES DE LOS CRUDOS PESADOS Y SUS FRACCIONES .....	16
1.1.1 Fraccionamiento por Destilación Molecular. ....	17
1.2 CLASIFICACIÓN Y CARACTERIZACIÓN .....	18
1.2.1 Composición Química.....	18
1.2.2 Análisis Composicional. ....	19
1.2.3 Propiedades Fisicoquímicas .....	20
1.3 ESPECTROMETRIA DE MASAS .....	22
1.3.1. Fuentes de ionización. ....	23
1.3.1.1 Desorción/Ionización por Láser.....	24
1.3.2Analizadores. ....	26
1.3.2.1 Analizador de tiempo de Vuelo. ....	27
1.3.2.2Analizador Tiempo de Vuelo Reflector.....	29
1.4 QUIMIOMETRÍA .....	31
1.4.1 Calibración multivariante.....	32
1.4.2 Clasificación de los métodos de calibración multivariable.....	32
1.4.3 Construcción de modelos.....	34
1.4.4 Pre-tratamiento de datos.....	35
1.4.5 Análisis por Componentes Principales (PCA) .....	37
1.4.6 Regresión por Mínimos Cuadrados Parciales (PLS-R).....	41
1.4.6.1 Factores a incluir en el modelo .....	44
1.4.7 Validación del modelo PLS .....	46
1.4.7.1 Validación Cruzada.....	46

1.4.8 Errores de Calibración y predicción .....	48
2. PARTE EXPERIMENTAL .....	50
2.1 SELECCIÓN DE MUESTRAS.....	51
2.1.1 Caracterización de las muestras.....	52
2.2 INSTRUMENTACIÓN .....	54
2.3 DETERMINACIÓN DE PARÁMETROS INSTRUMENTALES .....	55
2.4 ADQUISICIÓN Y ANÁLISIS DE ESPECTROS .....	57
2.5 PROCESAMIENTO DE DATOS ESPECTRALES .....	57
2.5.1 Pre-tratamiento de datos.....	57
2.5.2 Determinación del peso molecular promedio por LDI-TOF/TOF.....	58
2.5.3 Análisis por componentes Principales (PCA).....	60
2.5.4 Regresión por Mínimos Cuadrados Parciales (PLS-R).....	60
3. RESULTADOS Y ANALISIS.....	62
3.1 DETERMINACIÓN DEL PESO MOLECULAR PROMEDIO POR ESPECTROMETRÍA DE MASAS LDI-TOF/TOF .....	62
3.2 ANALISIS POR COMPONENTES PRINCIPALES (PCA).....	68
3.3 DESARROLLO DE MODELO DE REGRESIÓN PLS.....	73
3.3.1 Análisis de muestras excluidas.....	77
3.3.2 Validación cruzada del modelo PLS-R.....	80
4. CONCLUSIONES.....	83
5. RECOMENDACIONES .....	86
CITAS BIBLIOGRAFICAS.....	87
BIBLIOGRAFIA.....	92
ANEXOS.....	96

## LISTA DE FIGURAS

	pág.
<b>Figura 1.</b> Esquema de la obtención de fondos de vacío .....	17
<b>Figura 2.</b> Esquema general de los espectrómetros de masas .....	22
<b>Figura 3.</b> Esquema del principio de operación de un analizador TOF Reflectrón. 29	
<b>Figura 4.</b> Esquema de un analizador TOF-TOF .....	30
<b>Figura 5.</b> Esquema del método para dos componentes principales.....	40
<b>Figura 6.</b> Descripción gráfica del modelo PLS-R.....	42
<b>Figura 7.</b> Gráfico del residual frente al leverage.....	45
<b>Figura 8.</b> Esquema general de la validación cruzada.....	47
<b>Figura 9.</b> Esquema Metodológico experimental. ....	50
<b>Figura 10.</b> Diagrama experimental del fraccionamiento por Destilación Molecular de un fondo de vacío. ....	51
<b>Figura 11.</b> Destilador Molecular KD-6-1S (Chem. Tech. Services Inc. Co).....	52
<b>Figura 12.</b> Imagen fotográfica del Espectrómetro de Masas.....	54
<b>Figura 13.</b> Espectro de la fracción R2R obtenido a intensidad láser de (a) 60% y (b) 90% .....	56
<b>Figura 14.</b> Espectros de Masas finales de los fondos de vacío VRT, VRG y VRR	59
<b>Figura 15.</b> Espectros de Masas de C2R, R2R y VRR .....	62
<b>Figura 16.</b> Gráfica de Mnde los Condesados de Rubiales con su repetibilidad espectral .....	64
<b>Figura 17.</b> Gráficas del peso molecular promedio para fondos de vacío y sus fracciones de condensados y residuos. ....	66
<b>Figura 18.</b> Gráfica de Mn, Mw y °API de las muestras R .....	67
<b>Figura 19.</b> Gráfico de puntuaciones (Scores) PC2 vs PC1 .....	69
<b>Figura 20.</b> Gráfico de Influencia con varianza residual X vs influencia .....	71
<b>Figura 21.</b> Gráfico de influencia vs. Muestra .....	73

<b>Figura 22.</b> Varianza Explicada para la calibración en el modelo PLS .....	74
<b>Figura 23.</b> Error calculado en función del número de componentes para el modelo .....	75
<b>Figura 24.</b> Coeficientes de regresión para el cuarto factor del modelo PLS de gravedad API .....	76
<b>Figura 25.</b> Gráfico de Influencia del modelo PLS-R .....	77
<b>Figura 26.</b> Varianza Residual en Y para cada muestra PLS-R .....	78
<b>Figura 27.</b> Curva de valores de referencia contra valores predichos en la calibración.....	79
<b>Figura 28.</b> Curva de Validación de valores de referencia contra valores de predicción .....	80

## LISTA DE TABLAS

	pág.
<b>Tabla 1.</b> Clasificación de crudos según su densidad y gravedad API .....	20
<b>Tabla 2.</b> Principales características de los métodos de ionización utilizados en MS .....	24
<b>Tabla 3.</b> Tipos de Láseres más usados para LDI .....	25
<b>Tabla 4.</b> Principales características de los analizadores .....	27
<b>Tabla 5.</b> Caracterización fisicoquímica de los fondos de vacío R, T y G, y sus fracciones .....	53
<b>Tabla 6.</b> Datos peso molecular promedio para C2R, R2R y VRR .....	63
<b>Tabla 7.</b> Valores de $M_n$ y $M_w$ para los fondos de vacío y sus fracciones.....	65
<b>Tabla 8.</b> Varianza explicada por 10 componentes principales de las 20 muestras	68
<b>Tabla 9.</b> Parámetros del modelo seleccionado para la predicción de Gravedad API en fondos de vacío y sus fracciones de DM .....	79
<b>Tabla 10.</b> Validación cruzada del modelo de predicción de la gravedad API .....	81

## RESUMEN

**TITULO:** DETERMINACIÓN DE LA MASA MOLECULAR PROMEDIO DE FONDOS DE VACÍO DE CRUDOS COLOMBIANOS UTILIZANDO ESPECTROMETRÍA DE MASAS LDI TOF/TOF\*

**AUTOR**  
KAREN JULIETTE VEGA RAMÍREZ\*\*

**PALABRAS CLAVES**  
FONDOS DE VACÍO, ESPECTROMETRÍA DE MASAS LDI TOF/TOF, PCA, PLS-R

### DESCRIPCIÓN:

En este trabajo se determinó el peso molecular promedio de fondos de vacío y las fracciones de condensados y residuos, encontrándose un aumento gradual en el peso para estas muestras, en el siguiente orden: residuos, fondos de vacío y condensados, respectivamente; también, se desarrolló una metodología para obtener un modelo matemático que permite determinar la densidad API de fondos de vacío y sus fracciones de la Unidad de Destilación Molecular (UDM), a partir de la correlación del espectro de masas y la caracterización por el método de referencia (suministrada por el Instituto Colombiano de Petróleo – ICP), mediante el análisis por componentes principales (PCA) y regresión por mínimos cuadrados parciales (PLS-R).

Con el análisis por componentes principales se identificaron tres posibles agrupaciones de las muestras como: Livianas, medianas y pesadas atendiendo a las semejanzas y/o diferencias entre ellas de su densidad API. Se determinó la región espectral de 497 a 762 m/z como la de mayor importancia para el desarrollo del modelo.

La validación cruzada del modelo de regresión PLS mostró resultados satisfactorios para la predicción de la densidad API con error estándar de calibración (RMSEC) de 0,477 y error estándar de validación (RMSECV) de 0,919. El coeficiente de correlación ( $R^2$ ) entre los valores de referencia y los predichos por el modelo fue superior al 0,97; sin embargo, el modelo desarrollado no es viable para la predicción de la densidad API en muestras extrapesadas.

Se demostró que mediante la espectrometría de masas LDI-TOF/TOF se puede determinar el peso molecular promedio de fracciones pesadas de crudos Colombianos, con resultados más precisos que otras técnicas analíticas, y combinada con métodos de calibración multivariable permite el desarrollo de un método alternativo más fácil, rápido y económico para determinar la densidad API de estas muestras.

---

\*Trabajo de Grado

\*\*Facultad de Ciencias. Escuela de Química. Director: Enrique Mejía Ospino. Codirectora: Diana Catalina Palacio.

## ABSTRACT

**TITLE:** DETERMINATION THE AVERAGE MOLECULAR WEIGHT OF VACUUM| FUND OF COLOMBIAN CRUDE OIL USING LDI-TOF/TOF MASS SPECTROMETRY\*

**AUTHOR**  
KAREN JULIETTE VEGA RAMÍREZ\*\*

**KEYWORDS**  
LDI TOF/TOF MASS SPECTROMETRY, VACUUM RESIDUES, PCA, PLS-R

### DESCRIPTION

The average molecular weight of vacuum fund, condensates fractions, and residue fractions was determined by LDI mass spectrometry. A gradual increase in the molecular weight of these samples was identified in the following order: residues, vacuum funds and condensates, respectively. A methodology was also developed to obtain a mathematical model to determine the API gravity of vacuum funds and fractions obtained in molecular distillation units (UDM). The methodology started with the correlation of the mass spectrum and characterization by the method of reference (supplied by the Colombian Petroleum Institute - ICP), the approach also used principal component analysis (PCA) and partial least squares regression (PLS-R).

Light, medium and heavy clusters were found using the principal component analysis following similarities and/or differences in API density. The spectral region covering 497-762 m/z was the most important to develop of the model presented in this work.

Cross-validation of PLS regression model showed satisfactory results for the prediction of API gravity with standard error of calibration (RMSEC) of 0.477 and standard error of validation (RMSECV) of 0.919. The correlation coefficient ( $R^2$ ) between the reference values and those predicted by the model was higher than 0.97; however, the developed model is not feasible to predict the API gravity in extra heavy samples.

It was shown that by using mass spectrometry LDI-TOF / TOF it is possible to determine the average molecular weight of heavy fractions of Colombian crude oil with more accurate results than other analytical techniques. Mass spectrometry combined with methods of multivariate calibration allows the development of a simpler, quicker, and economic approach to determine the API gravity of these samples.

---

\* Work Degree

\*\*Sciences Faculty. Chemistry School. Directress: Enrique Mejía Ospino. Codirectress: Diana Catalina Palacio

## INTRODUCCIÓN

En Colombia el porcentaje correspondiente a la producción de crudos pesados es del 45%, pero debido a su alta viscosidad éste petróleo necesita de estimulación térmica y química para su extracción. Por ello, para mejorar el bajo porcentaje de aprovechamiento de este crudo, las empresas petroleras buscan identificar mejores tecnologías enfocadas en la extracción de crudos pesados (hidrocarburos más viscosos y densos que poseen metales), para obtener mayor beneficio y rendimiento del crudo explotado, y por ende, optimizar los métodos de extracción y producción.

La industria petrolera mundial, también debe trabajar en la búsqueda de soluciones para contrarrestar la degradación natural del crudo pesado por oxidación, debido al alto contenido de sal y sustancias venenosas como el azufre, para finalmente aportar en la optimización de los tratamientos de refinación.<sup>[1]</sup>

Debido a la gran cantidad de crudo pesado encontrado en el país, también se obtiene un alto porcentaje de fondos de vacío y se hace necesario conocer detalladamente su estructura molecular y propiedades fisicoquímicas. Es imprescindible como primer paso, caracterizarlos con pruebas preliminares como: densidad, viscosidad y análisis SARA, entre otros procedimientos regidos por la ASTM (American Society Test and Materials).<sup>[2]</sup>

Estos análisis de crudos son costosos y requieren elevado consumo de solventes tóxicos con prolongado tiempo de análisis, en consecuencia, actualmente se han implementado diversos métodos analíticos e instrumentales para el estudio de la composición química de estas fracciones, y que permiten la predicción de

propiedades a partir de parámetros obtenidos por técnicas de espectrometría y/o espectroscopía, como la espectrometría de masas (MS: Mass Spectrometry) y la Resonancia Magnética Nuclear (RMN). De esta manera, se obtienen resultados importantes para minimizar la complejidad de su tratamiento en las diferentes unidades de procesamiento, implementadas en las refinerías Colombianas. [3]

La espectrometría de masas ha mostrado ser una técnica válida para determinar la masa molecular promedio de los residuos de destilación molecular de fondos de vacío. [4] Sin embargo, también se ha demostrado que deben llevarse a cabo experimentos sistemáticos que permitan determinar las condiciones óptimas de adquisición de los espectros, evitando de esta manera la formación de clusters o fragmentos que lleven a malas interpretaciones del peso molecular promedio de las muestras. Teniendo en cuenta la alta complejidad de los datos obtenidos por ésta técnica es apropiado hacer uso del análisis multivariable, el cual permite disminuir la dimensionalidad de los datos y avanzar en la correlación de los mismos para predecir propiedades fisicoquímicas de las muestras.

En este trabajo se plantea una metodología para obtener un modelo matemático que permite predecir una propiedad fisicoquímica (Densidad API) de los fondos de vacío y sus fracciones de la Unidad de Destilación Molecular (Condensados y Residuos), a partir del tratamiento de los datos espectrales adquiridos por Espectrometría de Masas con ionización/desorción láser (LDI) acoplado a analizadores de tiempo de vuelo (TOF/TOF),<sup>[5]</sup> y mediante Análisis por Componentes Principales (PCA) y Regresión por mínimos cuadrados parciales (PLS-R), como técnicas quimiométricas.

## 1. CONSIDERACIONES TEÓRICAS

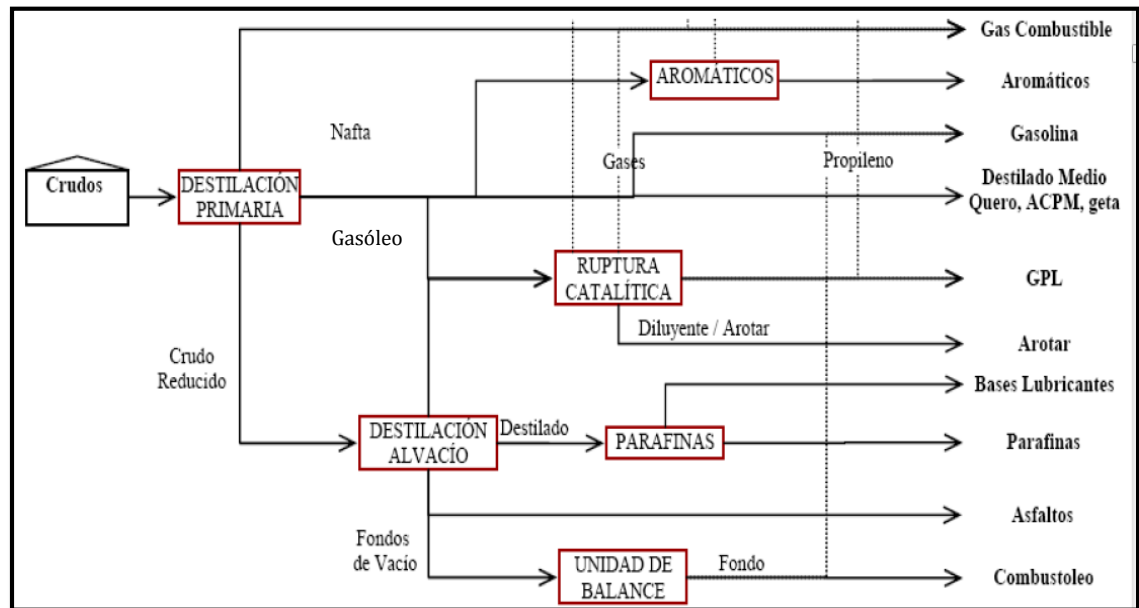
### 1.1 ASPECTOS GENERALES DE LOS CRUDOS PESADOS Y SUS FRACCIONES

El petróleo crudo es un recurso natural no renovable que aporta el mayor porcentaje del total de la energía que se consume en el mundo. Consiste principalmente en una mezcla de hidrocarburos líquidos, sólidos y gaseosos, con pequeñas cantidades de metales y de compuestos orgánicos de azufre, nitrógeno y oxígeno, también, incluye impurezas como agua, sales y sedimentos.

En la refinación del crudo se obtienen gran variedad de productos, entre ellos, se encuentran fracciones pesadas que provienen de la destilación al vacío del petróleo, estas son los fondos de vacío (FV) y están compuestos principalmente por resinas y asfáltenos, algunos hidrocarburos lineales de alto peso molecular y tal vez por hidrocarburos saturados de cadenas cíclicas;<sup>[6]</sup> estos compuestos pueden ser convertidos a través de procesos de hidrocrqueo o por reducción de viscosidad, en fracciones con menor punto de ebullición. En especial, el hidrocrqueo de esta fracción pesada da lugar a la obtención de volúmenes adicionales para combustibles de motor de alta calidad.

El término “pesado” hace relación a materiales con puntos de ebullición por encima de 345°C, incluyendo los residuos de destilación (Figura 1) que varían dependiendo de las temperaturas a las cuales fueron obtenidos.<sup>[7]</sup>

**Figura 1.** Esquema de la obtención de fondos de vacío



Fuente: Moreno González, D.H. 2011

Los fondos de vacío como fracción pesada, por mucho tiempo ha sido calificada como subproducto de bajo valor comercial debido a que concentra moléculas muy complejas y gran parte de las impurezas del crudo. No obstante, a causa de la disminución gradual de crudo liviano, las recientes investigaciones se orientan en el aprovechamiento de esta fracción para la producción de combustibles ligeros y otros materiales de mayor valor agregado como carbones y fibras activadas, grafito, compositos, entre otros. [8]

**1.1.1 Fraccionamiento por Destilación Molecular.** El fraccionamiento por destilación molecular consiste en la separación no convencional de mezclas líquidas homogéneas que contienen sustancias termo-sensibles de alto peso molecular y baja volatilidad; está basado en la evaporación de los componentes

más volátiles, su posterior condensación y enfriamiento. La evaporación ocurre a bajas presiones, en donde el principio de la separación se basa en la disminución de la temperatura de ebullición a causa del vacío, durante el proceso se generan destilados que contienen las moléculas que escaparon del evaporador y fueron atrapadas en el condensador, y residuos que corresponden a la fracción que no se evapora.<sup>[9]</sup>

## 1.2 CLASIFICACIÓN Y CARACTERIZACIÓN

La composición química de la materia prima como indicador del comportamiento en su refinación y representada en términos de tipos de compuestos, permite determinar la naturaleza de las reacciones en dicho proceso. Por lo tanto, el análisis composicional del crudo es una herramienta importante en el conocimiento de estos productos de refinación.

**1.2.1 Composición Química.** Los fondos de vacío como fracción pesada del petróleo tienen un punto de ebullición por encima de aproximadamente 540°C y como consecuencia del proceso de destilación en el que se eliminan los hidrocarburos ligeros, está constituido por moléculas bastante complejas formadas por varios anillos aromáticos con ramificaciones de cadenas alifáticas, cuyos pesos moleculares varían en un amplio rango masas.<sup>[10]</sup>

- **Crudos Parafínico:** Los compuestos principales de este tipo de crudo es la parafina. En el proceso de refinación dan lugar a la obtención de mayor cantidad de nafta y lubricantes. Son fluidos y de color claro.

- **Crudos Nafténicos:** Sus componentes primordiales son los hidrocarburos aromáticos y los naftenos. Este tipo de hidrocarburo genera mayor cantidad de residuo. Es de color oscuro y apariencia muy viscosa.
- **Crudos Mixtos:** Son aquellos que presentan una mezcla de los anteriores compuestos.

Por otra parte, esta fracción también contiene impurezas que incluyen azufre (hasta en un 6% en peso), nitrógeno (0.1-2% en peso), oxígeno (0.005-1.5% en peso), y trazas de metales como níquel, vanadio, hierro, calcio y magnesio, en forma de óxidos o sales o en estructuras organometálicas tipo porfirina.<sup>[11]</sup>

**1.2.2 Análisis Composicional.** Siendo el petróleo una mezcla tan compleja, debido al número elevado de componentes, conocer su composición química es un trabajo tedioso;<sup>[12]</sup> sin embargo, en las refinerías se utilizan diversas normas ASTM para determinar la calidad de los productos.

Basados en la naturaleza del crudo existen varias maneras de expresar su composición, como: PIONA (parafinas, isoparafinas, olefinas, naftenos y aromáticos), PONA (parafinas, olefinas, naftenos y aromáticos) y PNA (parafinas, naftenos y aromáticos) para cargas livianas con rangos de ebullición más estrechos, en cambio, para fracciones pesadas y residuos que contienen mayor contenido de asfáltenos y resinas, es útil el análisis SARA (ASTM D2007, D4124 y D2549).<sup>[13]</sup>

### 1.2.3 Propiedades Físicoquímicas

**1.2.3.1 Densidad API.** Una de las características que define al petróleo, es su gravedad. La densidad aumenta con el incremento de hidrocarburos y productos pesados (resinas y asfáltenos), y disminuye con la temperatura. La densidad de los crudos se expresa comúnmente en términos de densidad API (parámetro internacional establecido por el Instituto Americano de Petróleo), con unidades °API y se relaciona con el peso específico mediante la *ecuación (1)* en su correspondiente cálculo.<sup>[14]</sup> La gravedad específica del petróleo es su densidad con relación al agua a temperatura de 60° Fahrenheit.

$$^{\circ}API = \frac{141.5}{\text{densidad relativa estándar}_{\frac{60^{\circ}F}{60^{\circ}F}}} - 131.5 \quad (1)$$

Como se relaciona en la tabla 2, normalmente el crudo se puede clasificar en base a su densidad API desde extra pesado hasta super liviano. <sup>[15]</sup>

**Tabla 1.** Clasificación de crudos según su densidad y gravedad API

Tipo de Crudo	Densidad (g/cm <sup>3</sup> )	Gravedad API (°API)
Extra pesado	>1.0	<10
Pesado	1.0 – 0.92	10.0 – 22.3
Mediano	0.92 – 0.87	22.3 – 31.1
Liviano	0.87 – 0.83	31.1 – 39
Super liviano	<0.83	>39

Fuente. Wauquier, J. P. 2004

El petróleo liviano tiene una gravedad API de 22°API o más, también denominado 'petróleo convencional', éste tipo de crudo es el más codiciado en la industria petroquímica, y el de mayor valor económico, debido a su menor viscosidad y alta tendencia a emulsificarse, con lo que disminuyen sus costos de extracción y refinación comparados con los crudos, que son densos y viscosos con una gravedad API baja. Es generalmente aceptado que el límite máximo de los petróleos pesados sea 22° API. Es poco probable que su extracción se da en forma natural en un pozo petrolífero, por lo general se necesita de altos procesos térmicos para que pueda fluir.

El crudo extrapesado tiene una gravedad API de menos de 10°API. Los crudos pesados y extrapesados son una importante reserva mundial y el agotamiento de crudos livianos los convierte en una fuente obligada de hidrocarburos.

**1.2.3.2 Contenido de Azufre.** Se puede caracterizar un crudo en términos de los niveles de azufre (entre dulce y agrio). Crudo 'dulce' es definido comúnmente como el petróleo con un contenido de azufre menor de 0,5%, en tanto que un crudo 'agrio' posee niveles de 0,5% o más. <sup>[16]</sup> Los crudos pesados y extra-pesados presentan dos características composicionales importantes y vinculadas entre sí:

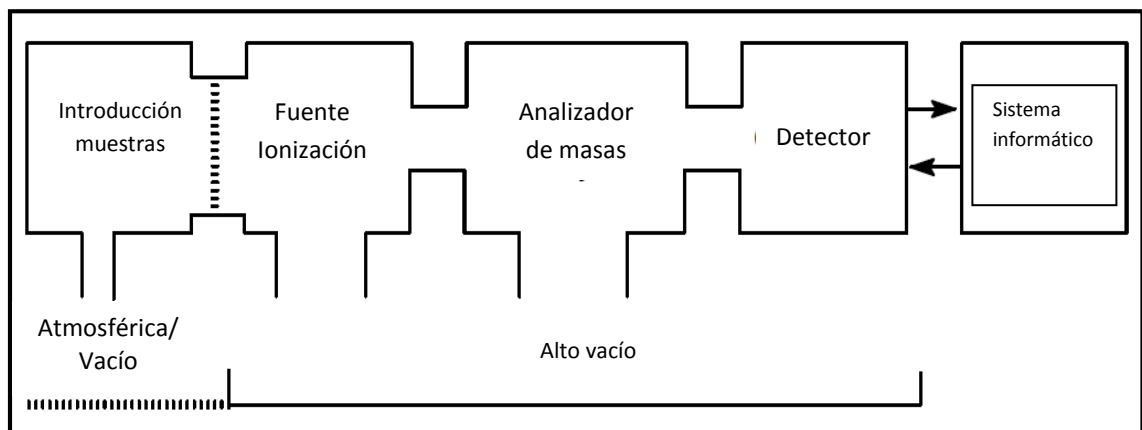
- Un alto residuo cuando se destilan por encima de 360°C.
- Una cantidad importante de contaminantes en procesos de refinación de gasolinas como Azufre, Vanadio y Níquel, además por la presencia elevada de Asfáltenos.

### 1.3 ESPECTROMETRIA DE MASAS

La espectrometría de masas es una poderosa técnica analítica que permite identificar compuestos desconocidos y algunas propiedades químicas de las moléculas. Se puede obtener información característica como el peso molecular del analito, con la detección de cantidades realmente pequeñas del orden de nanogramos (ng).<sup>[17]</sup>

En un espectrómetro de masas un haz de iones es separado de acuerdo a la relación masa/carga ( $m/z$ ), permitiendo con esta medida obtener la masa del ión. En general todo espectrómetro de masas está compuesto por unos elementos esenciales como se muestran en la figura 2.

**Figura 2.** Esquema general de los espectrómetros de masas



*Fuente: J. Gross, Mass Spectrometry, 2004.*

Los elementos que definen el tipo de espectrómetro y las capacidades del sistema son *lafuente de ionización*, donde son producidos los iones al suministrarle energía a la muestra que se estudia, y, *elanalizador* (un campo eléctrico, un campo magnético, de tiempo de vuelo, trampa de iones, ICR, etc.), donde los iones son separados de acuerdo a la relación masa/carga ( $m/z$ ).

La energía suministrada puede ser suficiente no solamente para ionizar la molécula y obtener el llamado ion molecular (la molécula intacta pero con una carga positiva o negativa), sino también para provocar la fragmentación de la molécula dando lugar a iones fragmentos. Los iones producidos son característicos de un determinado compuesto químico y por tanto es posible su identificación a partir del espectro de masas.<sup>[18]</sup>

**1.3.1. Fuentes de ionización.**La ionización de un analito es una etapa crucial en el análisis de cualquier compuesto por Espectrometría de Masas (MS: Mass Spectrometry). Existe una amplia variedad de técnicas como se observa en la tabla 3, pero ninguna de ellas es universal. La elección del tipo de ionización viene determinada por la naturaleza de la muestra y la clase de información que se desea obtener. El aspecto de los espectros de masas obtenidos para diferentes especies moleculares depende en gran medida del sistema de ionización utilizado.<sup>[19]</sup>

**Tabla 2.** Principales características de los métodos de ionización utilizados en MS

Técnica de Ionización	Naturaleza del analito	Introducción de la muestra	Rango de Masas	Breve descripción
<b>Impacto Electrónico (EI)</b>	Volátiles; térmicamente estables	Cromatografía de gases	<1000 Da	Método duro: principalmente fragmentación de iones
<b>Ionización Química (CI)</b>	Volátiles; térmicamente estables	Cromatografía de gases	<1000 Da	Método suave; ión molecular
<b>Bombardeo de átomos rápidos (FAB)</b>	Componentes órgano-metálicos	Cromatografía de líquidos; entrada directa	<5000 Da para un rango óptimo	Método suave, requiere matriz
<b>Ionización/Desorción láser asistida por matriz (MALDI)</b>	Biomoléculas (proteínas, DNA, entre otros)	La muestra co-cristalizada en una matriz	Puede ir más allá de los 500.000 Da	Método muy suave
<b>Ionización por Electrospray (ESI)</b>	Componentes orgánicos e inorgánicos	Muestra en solución	Desde muy bajas a muy altas	Método muy suave

Fuente: Palacio Lozano, D. C. 2012

**1.3.1.1 Desorción/Ionización por Láser.** La Desorción/Ionización por Láser (LDI: Laser Desorption/Ionization) es una técnica de ionización suave introducida a finales de los 60's y mucho antes de la técnica Bombardeo de átomos rápidos (FAB: FastAtomBombardment). LDI es un proceso de que usa por láser, de forma continua o en pulsos (para un láser Nd:YAG 335-226 nm la duración del pulso láser es aproximadamente de 4-7 ns). El haz de radiación incide en un área pequeña de la muestra problema a fin de desorber moléculas, para producir la vaporización (*desorción*; aumento de la energía cinética o transnacional) y la

ionización (energía de excitación electrónica que conduce al despojo de electrones) del analito, en la que sucede captura electrónica ( $M^-$ ).

- Longitud de onda del láser. La longitud de onda de láser es un parámetro importante en MALDI (Matrix-Assisted Laser Desorption/Ionization).<sup>[20]</sup> La longitud de onda que se usa con mayor frecuencia es la de 337 nm de un láser de nitrógeno, pero también han sido empleados los armónicos del láser de Nd:YAG y más recientemente láseres infrarrojos tales como Er:YAG (IR-MALDI).<sup>[21]</sup> En la tabla 4 se muestran algunos de los láseres más usados con su correspondiente longitud de onda y energía de los fotones emitidos.

**Tabla 3.** Tipos de Láseres más usados para LDI

Rango espectral	Longitud de onda	Energía del Fotón	Tipo de Laser
UV	193 nm	6.4 eV	ArFExcimer laser
UV	248 nm	5.0 eV	KrFExcimer laser
UV	266 nm	4.7 eV	Cuarto armónico Nd:YAG laser
UV	308 nm	3.8 eV	XeClExcimer laser
UV	337 nm	3.7 eV	Nitrogen laser*
UV	355 nm	3.5 eV	Tercer armónico Nd:YAG laser*
IR	1.06 $\mu\text{m}$	1.2 eV	Nd:YAG laser**
IR	2.94 $\mu\text{m}$	0.4 eV	Er:YAG laser**
IR	1.7–2.5 $\mu\text{m}$	0.7–0.5 eV	Opticalparametricoscillator (OPO) laser
IR	10.6 $\mu\text{m}$	0.1 eV	CO <sub>2</sub> laser

\* Láseres UV de uso más frecuente; \*\* Láseres IR de uso más frecuente

Fuente: J. Gross, *Mass Spectrometry*, 2004.

Cada pulso láser genera un pequeño plasma que se expande en vacío, del cual se extraen iones (positivos o negativos) del analito que son acelerados por un potencial de 10-30 kV y controlados por medio de campos electromagnéticos. Debido a diferencias de velocidades que resultan de sus diferencias en masas los iones llegan al detector en diferentes tiempos. [22]

Mientras que sales orgánicas de baja masa y moléculas orgánicas que absorben la luz son fácilmente accesibles en LDI, se necesita un gran esfuerzo para obtener espectros de masa útil de Biomoléculas, [23] en particular, cuando la masa del analito excede 2000 u. Cuando se combina con un analizador de masas TOF, ambos métodos son capaces de producir espectros de masas de las proteínas de aproximadamente 100.000 u de peso molecular, y, actualmente LDI y MALDI representan herramientas analíticas importantes en las ciencias de la vida moderna [24, 25] el análisis de biopolímeros [26,27] y moléculas orgánicas complejas con peso molecular entre unos pocos miles y cientos de miles de Da (uma).

**1.3.2 Analizadores.** El analizador de masas constituye el corazón de un espectrómetro de masas, su funcionamiento y desempeño dependen en gran medida de su diseño. El analizador tiene dos misiones fundamentales: separar los iones en función de su relación  $m/z$  y enfocar los iones separados hacia un determinado punto. El movimiento de las partículas cargadas permite distinguir unas de otras en función de la energía cinética y velocidad de cada ión. [28]

Los diferentes analizadores relacionados en la tabla 5 aprovechan alguna o varias de estas propiedades para separar los iones de distintas masas.

**Tabla 4.** Principales características de los analizadores

Analizador de masas	Rango de masas	Resolución	Ventajas	Desventajas
<b>Sector magnético (B)</b>	1-15.000 m/z	$10^2$ - $10^5$	Alta resolución	Baja sensibilidad Costoso
<b>Cuadrupolar (Q)</b>	1-5.000 m/z	$10^3$ - $10^4$	Fácil de usar Poco costoso Alta sensibilidad	Baja resolución Bajo rango de masas
<b>Trampa de iones (IT)</b>	1-5.000 m/z	$10^3$ - $10^4$	Poco costoso Alta sensibilidad Tandem MS (MS <sup>n</sup> )	Bajo rango de masas
<b>Tiempo de vuelo (TOF)</b>	ilimitado	$10^3$ - $10^4$	Alto rango de masas Diseño simple	Baja resolución
<b>Resonancia ciclotrónica de iones (ICR)</b>	Superior a 70 KDa	$10^4$ - $10^6$	Alta resolución y rango de masas	Alto costo Alto conocimiento técnico

Fuente: Palacio Lozano, D. C. 2012

**1.3.2.1 Analizador de tiempo de Vuelo.** Es uno de los analizadores más sencillos, el primero fue construido y dado a conocer en 1946 por W. E. Stephens,<sup>[29]</sup> pero actualmente tiene una gran presencia en los laboratorios. El principio de operación del analizador de tiempo de vuelo (TOF: time-of-flight) involucra la medición del tiempo requerido por un ion para viajar desde la fuente de iones hasta el detector localizado a 1-2 m de la fuente. Todos los iones reciben la misma energía cinética durante la aceleración instantánea, pero debido a que tienen diferentes valores de m/z, se separan en grupos de acuerdo a su

velocidad a medida que recorren la región de deriva entre la fuente de iones y el detector. Los iones chocan secuencialmente en el detector a medida que aumenta su masa. Los de baja  $m/z$  llegan al detector antes que aquellos con  $m/z$  alta, debido a que, entre más  $m/z$  tengan los iones tendrán una velocidad menor.

- Ventajas y limitaciones.

El analizador TOF es el analizador de masas más rápido, puede adaptarse a los métodos de ionización por pulsos (MALDI), tiene alta transmisión de iones y el rango de masas más grande de todos los analizadores de masas, por lo que son de amplia utilización para el análisis de sustancias de muy elevado peso molecular. <sup>[30]</sup>

Algunas limitantes son que requieren exclusivamente de un método de ionización por pulsos y que la selectividad de los iones puede limitarse en algunos experimentos, además, el principal problema en este tipo de analizador es que el tiempo que tardan los iones en alcanzar el detector se encuentra en el entorno de unos pocos microsegundos, lo que exige sistemas de detección muy rápido.

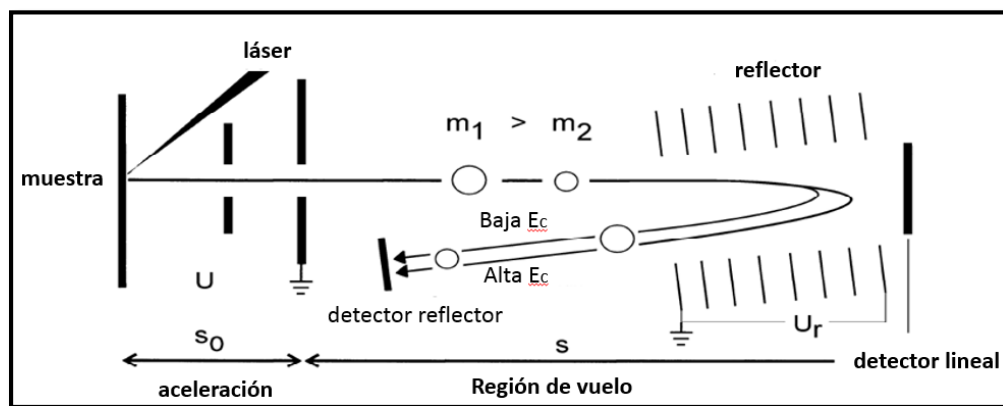
La baja resolución que presenta este analizador se debe a la dispersión que sufren en las velocidades que alcanzan los iones tras la aceleración, por no ser éstos generados en la fuente en un punto geométrico sino en un volumen de espacio.<sup>[31]</sup> El efecto de dispersión de velocidades tiende a compensarse en un punto, llamado *punto focal primario*, y si el detector se ubica en este punto, la anchura del paquete de iones será la menor posible y el analizador alcanzará su máxima resolución; sin embargo, en la gran mayoría de los equipos, la distancia

del punto focal primario es del orden de 10 cm; muy corta para proporcionar un tiempo de vuelo suficiente. [32]

**1.3.2.2 Analizador Tiempo de Vuelo Reflector.** El reflector o reflectrón fue ideado por Mamyrin en 1994. [33] Generalmente, los instrumentos reflectores también están equipados con un detector detrás del reflector que permite el funcionamiento en modo lineal con sólo apagar el voltaje de éste. La siguiente es una breve explicación cualitativa de este tipo de analizador TOF.

Un simple reflector consiste en un campo eléctrico de retardo ubicado detrás de la región de deriva libre de campo libre y opuesto a la fuente de iones. En la práctica, un reflector está compuesto por una serie de electrodos en forma de anillo en el que se va aumentando el potencial. El voltaje de la reflexión  $U_r$  se establece alrededor de 1,5 a 1,10 veces el voltaje de aceleración  $U$  con el fin de asegurar que todos los iones se reflejen dentro de la porción homogénea del campo eléctrico del dispositivo. [34] Un ejemplo de la representación esquemática de este tipo de analizador TOF se muestra en la figura 3.

**Figura 3.** Esquema del principio de operación de un analizador TOF Reflectorón.

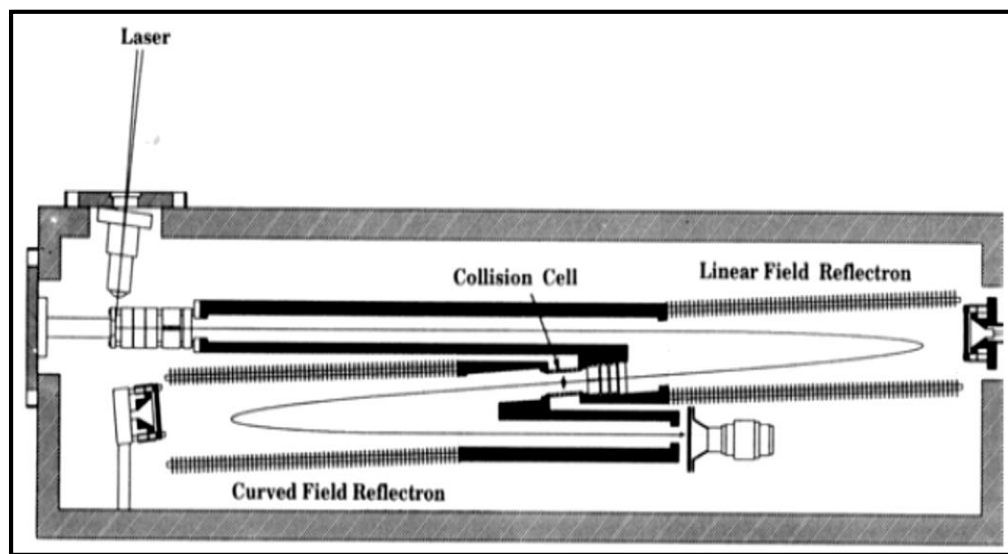


Fuente: J. Gross, *Mass Spectrometry*, 2011.

Los iones penetran en el reflectrón hasta que alcanzan energía cinética cero y luego son expulsados del reflector en la dirección opuesta. La energía cinética de los iones que salen del reflectrón permanece inalterada, sin embargo, los iones que tienen más energía cinética volarán más profundamente en el campo de desaceleración, y así pasarán más tiempo en el reflectrón en comparación con los iones menos energéticos.<sup>[35]</sup>

Los analizadores TOF/TOF (figura 4) proporcionan un mejor enfoque de los iones y por tanto, mayor resolución y precisión. Estos analizadores tienen tres secciones: sección de tiempo de vuelo lineal, la región de colisión que sirve como fuente de iones y un TOF reflectrón; esta última sección introduce una corrección en la energía de dispersión de los iones que salen de la fuente de ionización y que tienen el mismo valor  $m/z$ .

**Figura 4.** Esquema de un analizador TOF-TOF



Fuente: Cotter. *Time-of-flight mass spectrometry*, 1997

## 1.4 QUIMIOMETRÍA

En 1971, el científico sueco Svante Wold acuñó el término "Kemometri", en inglés equivalente a "Quimiometría". Específicamente la quimiometría trata de todos los procesos que transforman datos complejos y señales analíticas en información útil para el analista. Esta disciplina utiliza métodos estadísticos, matemáticos y de lógica formal para lograr sus fines, situándose en un campo interdisciplinar, pero sin duda alguna, totalmente dependiente de los problemas químicos que se puedan satisfacer.<sup>[36]</sup>

Dado que, los métodos espectroscópicos proporcionan datos analíticos sobre muchos componentes de una única muestra, la calibración multivariable se emplea en el análisis de sistemas químicos complejos como aguas residuales, formulaciones de medicamentos, aceites y combustibles, caracterización de materiales industriales, especialmente en la producción de hidrocarburos,<sup>[37]</sup> siendo un método de calibración efectivo, pues, la información contenida en las variables de una señal puede ser reducida a un número menor de variables sin que haya pérdida de información relevante.

En comparación con los métodos de calibración convencional, como la determinación simultánea de varias variables aplicando métodos ASTM, la técnica de calibración multivariable presenta ventajas como la sencillez, rapidez y bajo costo en el análisis.

Por lo anterior, se considera importante el desarrollo y aplicación de los nuevos métodos de calibración multivariable en la quimiometría, para lograr así, extraer la

máxima información química del análisis de muestras a partir de señales o respuestas instrumentales con escasa selectividad.<sup>[38]</sup>

**1.4.1 Calibración multivariante.** La calibración multivariante permite predecir una propiedad utilizando respuestas instrumentales no selectivas con la relación entre una o más variables de respuesta “Y” y una matriz de variables predictorias “X”, es el caso del espectro de una muestra; esta relación se cumple según la ecuación (5).

$$Y = G(X) \quad (5)$$

La variable “Y” puede ser un parámetro cualitativo o cuantitativo que representa una propiedad de interés en el sistema, y la matriz “X” contiene información relevante de la muestra determinada por un método de análisis químico. La función “G(X)” puede ser difícil de determinar por el número de variables en la matriz “X” y por la incertidumbre asociada a la determinación de los parámetros “Y” y “X”. La forma de la función “G(X)” depende del método de regresión empleado, por lo tanto, puede existir más de una posibilidad de ajuste de los datos diferenciándose básicamente en la complejidad de la función y en sus parámetros estadísticos.

**1.4.2 Clasificación de los métodos de calibración multivariable.** Según Martens & Neas,<sup>[39]</sup> se proponen sistemas de clasificación basados en los métodos de calibración y conceptos fundamentales no tan matemáticos, así:

**1.4.2.1 Según la relación entre las variables dependiente e independiente.**

- Calibración Lineal. Descrita por un modelo lineal en los parámetros, como se describe en la *ecuación (6)*.

$$y = b_0 + \sum_{k=1}^k b_k X_k \quad (6)$$

Los parámetros a determinar son  $b_0$  y  $b_k$ , y las variables independientes  $X_k$ .

- Calibración no-lineal. Descrita por un modelo no lineal en los parámetros, según la *ecuación (7)*.

$$y = \alpha x^\beta \quad (7)$$

Los parámetros a determinar son  $\alpha$  y  $\beta$ .

#### **1.4.2.2 Según la forma de encontrar la relación entre las variables, y pueden ser:**

- Métodos directos. En esta calibración los parámetros se calculan directamente a partir de la señal de cada uno de los analitos de forma individual.
- Métodos indirectos. Los parámetros de la calibración se calculan a partir de las señales analíticas de mezclas de los componentes.

**1.4.2.3 Sistemas Rígidos-Flexibles.** En los primeros es necesario tener información de todas las especies presentes que pueden contribuir a la señal, mientras que en los flexibles únicamente es necesario tener información de los analitos que se desean cuantificar.

#### **1.4.2.4 Según cuál sea la variable dependiente y cuál la independiente**

- Calibración clásica. Sigue un criterio directamente relacionado con la ley de Beer-Lambert. La señal analítica actúa como variable dependiente de la concentración, que es la variable independiente.
- Calibración inversa. Se utiliza la concentración como variable dependiente y la señal analítica como independiente.

**1.4.3 Construcción de modelos.** El principal objetivo de los métodos de calibración multivariable es construir modelos de calibración para predecir el valor de la propiedad analizada para nuevas muestras, a las que se ha registrado la señal analítica. Para obtener modelos válidos y de los que depende la fiabilidad de las predicciones, se deben construir bajo el desarrollo de los siguientes parámetros:

**1.4.3.1 Selección de las muestras de calibración.** Se tienen en cuenta muestras altamente representativas del sistema en estudio, considerando las variaciones químicas y físicas, y que permitan aplicar el modelo de predicción a muestras futuras.

**1.4.3.2 Caracterización de muestras de Calibración.** Esta caracterización se debe ejecutar mediante métodos de referencia establecidos para este fin, que sean altamente confiables y estadísticamente evaluados.

**1.4.3.3 Toma de espectros.** Se deben seleccionar las condiciones óptimas para la adquisición espectral y realizar pre-tratamiento a los datos, para minimizar las contribuciones presentes pero no deseadas en la señal analítica que a su vez disminuyen la reproducibilidad del sistema.

**1.4.3.4 Cálculo del modelo matemático.** Enseguida del pre-tratamiento óptimo de la señal, se emplean técnicas de regresión sobre los datos espectrales.

**1.4.3.5 Selección del modelo matemático.** Establece la relación más simple posible entre la propiedad a determinar y la señal analítica. Para esto, en muchos casos se requiere de un estudio complejo de la matriz de datos analíticos obtenidos, ya que no siempre es necesaria la utilización de toda la información registrada.

Las últimas etapas corresponden a la validación e implementación del modelo para el análisis de muestras desconocidas, realizando controles frecuentes para evaluar su desempeño.

**1.4.4 Pre-tratamiento de datos.** De una manera genérica se encuentran además de la contribución del analito a la señal, efectos no deseados como lo es el ruido; éste puede ser un ruido aleatorio o un ruido estructurado consecuencia de

interferencias químicas y/o físicas. Además, no siempre el instrumento responde linealmente a los cambios de concentración del analito por lo que la calidad de la señal puede ser deficiente. Existen varios pre-tratamientos que se pueden aplicar antes de un procesamiento para la mejora de la calidad de datos.<sup>[40]</sup>

**1.4.4.1 Corrección línea base.** Por medio de este tratamiento se pretende corregir determinadas tendencias en la línea base que aporta ruido espectral a la señal. Determinado por el efecto que se desee corregir, hay varios tipos de corrección como el Offset y la corrección lineal.

**1.4.4.2 Normalización.** Se utiliza para obtener todos los datos en aproximadamente la misma escala, o una distribución más equitativa de las variaciones y los valores medios. Existen varios tipos, los siguientes son los más aplicados.

- Normalización media: Consiste en dividir cada fila de una matriz de datos por su media, neutralizando así la influencia de los factores ocultos. *Ecuación (8)*.

$$\hat{X}_i = X_i / \bar{x}(\mathbf{g})$$

- Normalización Máxima: Se considera una alternativa a la normalización media clásica. Divide cada fila por su valor máximo absoluto en lugar de la media. *Ecuación (9)*.

$$\hat{X}_i = \frac{X_i}{x_{max}} \quad (9)$$

- Normalización por rangos: En esta transformación se normaliza un espectro  $X_i$  calculando el área bajo la curva del espectro. *Ecuación (10)*.

$$\hat{X}_i = X_i / (x_{max} - x_{min}) \quad (10)$$

**1.4.4.3 Suavizado espectral.** En ciertos casos en el que la relación señal/ruido es pequeña, suele no ser suficiente la disminución del ruido de alta frecuencia, por lo que se hace necesario aplicar sobre los espectros diferentes algoritmos que minimicen este efecto. Los métodos más utilizados son los basados en filtros de SavitzkyGolay<sup>[41]</sup> y transformadas de Fourier.<sup>[42]</sup>

**1.4.5 Análisis por Componentes Principales (PCA).** El análisis por componentes principales (PCA: Principal Component Analysis) es una técnica de descomposición multivariante, que permite reducir el número de variables perdiendo la menor cantidad de información posible y reduciendo la contribución del ruido. El PCA construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos denominados componentes principales (PCs) el cual explica la máxima variabilidad de las muestras; el segundo PC se escoge de forma que sea ortogonal al primero mostrando la máxima variabilidad de las muestras una vez restada la explicada por el primer PC.

Esta técnica está definida por la norma ASTM E131 como un procedimiento matemático para resolver conjuntos de datos hallando componentes o vectores

ortogonales entre sí, llamados componentes principales (PC), cuya combinación lineal se aproxima a los datos originales de acuerdo al grado de exactitud deseado. Tales componentes modelan la variación estadísticamente significativa en el conjunto de datos así como el error aleatorio presente, de manera que se puedan detectar fenómenos que a primera vista no parecen evidentes. [43]

Para aplicar el PCA, es necesaria una matriz de datos donde cada fila corresponde a la descripción completa de una medida, y cada columna a una variable indicadora concreta de cada experimento. En términos algebraicos, las filas hacen referencia a vectores pertenecientes a un espacio vectorial de dimensión  $p$ , en base canónica, en la que cada una de sus coordenadas se asocia a la respectiva variable descriptiva. El algoritmo PCA supone por lo tanto realizar un cambio de base, desde la original canónica a otra generada por los vectores propios resultantes de la matriz de covarianzas de los datos. Existen diferentes algoritmos de cálculo para obtener las matrices  $T$  y  $P$ , uno de ellos es el NIPALS (Nonlinear Iterative Partial Least Squares) utilizado por el software TheUnscrambler.

A partir de la matriz  $X$  construida de la información espectral obtenida para  $m$  muestras medidas a  $k$  relación masa/carga, se realiza una descomposición por componentes principales según la ecuación (11), que proporciona una aproximación a la matriz  $X$  como un producto de dos matrices,  $T$  (matriz de puntuaciones) y  $P$  (matriz de cargas), más una matriz  $E$  de residuales de  $X$ .

$$X = TP^T + E \quad (11)$$

La matriz T contiene información pertinente a las relaciones entre muestras y está construida por m filas, que corresponden al número de muestras u objetos, y A columnas, que corresponden al número de componentes principales más dominantes. La matriz P explica la relación existente entre variables originales y está constituida por A filas y k columnas, y E representa el ruido o la información no relevante.

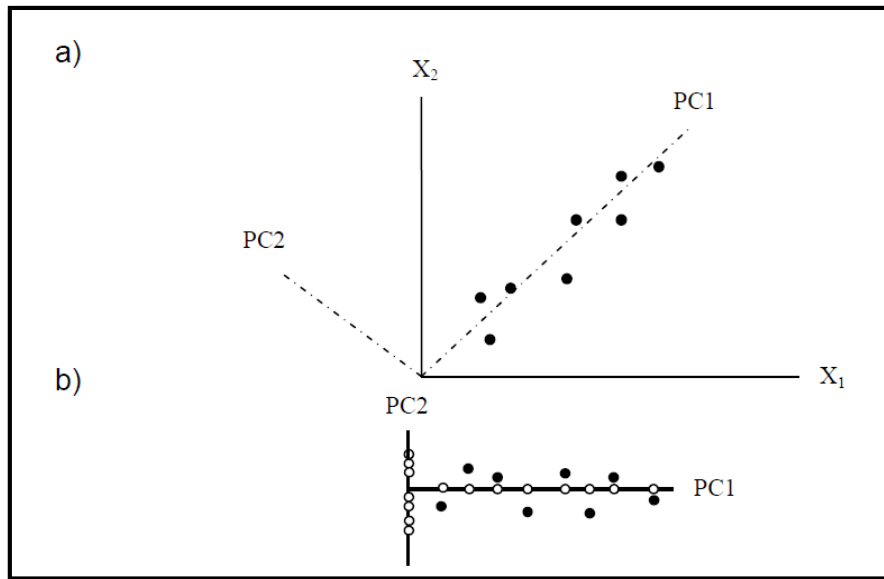
El producto  $TP^T$  se puede representar como la suma de A términos, según *ecuación (12)*, que corresponde a cada una de las columnas y filas de las matrices T y P respectivamente, cada uno de dichos términos se denominan factor o componente principal.

$$X = t_1p_1^T + t_2p_2^T + t_ap_a^T + E \quad (12)$$

Para tener una reproducción más precisa de la matriz X se necesitan k vectores de cargas y puntuaciones, sin embargo es posible representar la matriz original con un número  $a < k$  de vectores.

En la figura 5 se describe el método para dos variables y por lo tanto dos componentes principales. En la figura 5a se muestra los PC mediante líneas punteadas y formando ángulos rectos entre ellas. La figura 5b muestra los puntos referidos a estos dos nuevos ejes y también la proyección de los puntos sobre PC1 y PC2, en este caso se observa como  $Z_1$  recoge la mayor parte de la variación, permitiendo trabajar en una dimensión.

**Figura 5.** Esquema del método para dos componentes principales



*Fuente: Miller, N. J. Estadística y Quimiometría para química analítica. 2002*

**5a)** Ilustración de dos componentes principales PC1 y PC2, para dos variables  $X_1$  y  $X_2$ . **5b)** Puntos referidos a los ejes de los PC, los puntos negros indican datos y los puntos blancos sus proyecciones sobre los ejes.

En el PCA la gráfica de pesos (loadings) y la de puntuaciones (scores), permiten definir los PC; la primera contiene los coeficientes de las transformaciones lineales de las variables originales que luego, darán lugar a ellos mismos, y geoméricamente son los cosenos de los ángulos que forman con los antiguos ejes, y la segunda, revela las coordenadas de las muestras en estos nuevos ejes.<sup>[44]</sup>La gráfica de puntuaciones permite interpretar las diferencias y similitudes de las muestras, ya que, cuanto más cerca se encuentren éstas, más similares

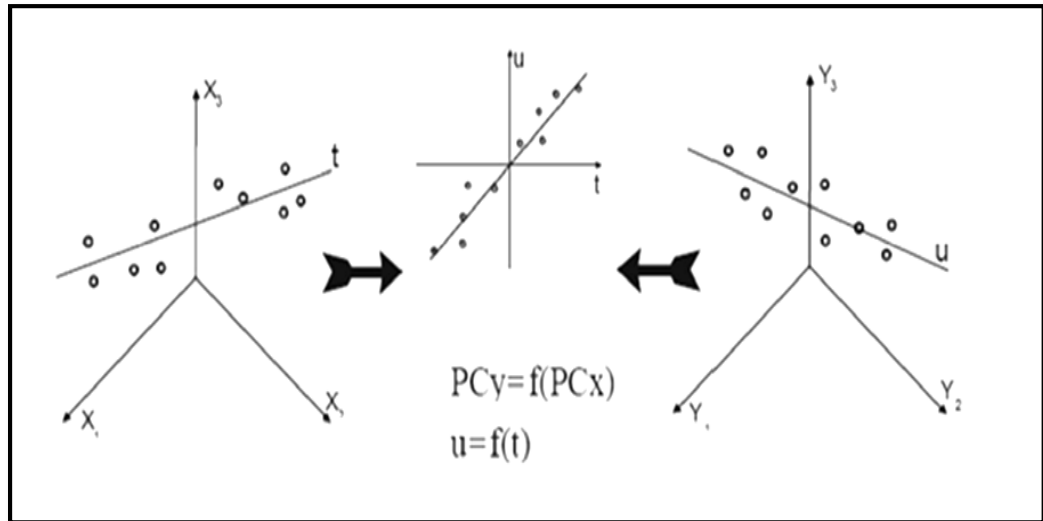
son con respecto a los dos componentes en cuestión. Por el contrario, las que se encuentren muy separadas son diferentes unas de otras.

En resumen, el conjunto de datos  $X$  que estaba descrito por variables correlacionadas, en este momento está definido por un nuevo número de variables, llamadas componentes principales, que son variables no correlacionadas entre sí, en un nuevo sistema de ejes ortogonales. Por todo lo anterior, el Análisis de Componentes principales es en primer lugar un método matemático para la reducción de datos y no supone que los datos tengan ninguna distribución concreta. Se ha evidenciado como el PCA se puede utilizar para reducir la dimensionalidad de un conjunto de datos y cómo puede, por tanto, revelar conglomerados.

**1.4.6 Regresión por Mínimos Cuadrados Parciales (PLS-R).** Como método de calibración multivariante la regresión por mínimos cuadrados parciales (PLS-R, Partial Least-Squares Regression), es utilizado para la reducción de variables, generaliza y combina características del Análisis por Componentes Principales y Análisis de Regresión Múltiple. Además, para estudiar la estructura de covarianza entre los espacios correspondientes a dos matrices  $X$  y  $Y$ , para predecir un conjunto de variables dependientes a partir de uno grande de variables independientes.

El algoritmo PLS utiliza la información contenida en la matriz de datos ( $X$ , Datos espectroscópicos) y en la matriz de la propiedad a determinar ( $Y$ , propiedad), obteniéndose unas variables auxiliares llamadas variables latentes (VL), una descripción gráfica se presenta en la figura 6.

**Figura 6.** Descripción gráfica del modelo PLS-R



Fuente: Miller, N. J. *Estadística y Quimiometría para química analítica*. 2002

Estas nuevas variables (VL) en la matriz X y la matriz Y se pueden representar como un producto de matrices, como se muestra en las ecuaciones (13) y (14).

$$X = TP^T + E = \sum t_a p_a^T + E \quad (13)$$

$$Y = UQ^T + F = \sum u_a q_a^T + F \quad (14)$$

Donde, para X y Y, las matrices de puntuación (scores) son T y U respectivamente, y las matrices de carga (loadings) son P y Q., y los residuos son E y F.

Las variables originales X y Y se pueden relacionar mediante los scores de cada una de las nuevas variables latentes como lo muestra la *ecuación (15)*.

$$u_a = b_a t_a (15)$$

Donde,  $b_a$  es el coeficiente de regresión para cada variable latente obtenido a través de la *ecuación (16)*.

$$b_a = \frac{u_a^T t_a}{t_a^T t_a} (16)$$

Los coeficientes  $b_a$  hallados para cada componente son agrupados en una matriz diagonal B, que contiene los coeficientes de regresión de los scores U y T de las matrices Y y X respectivamente. De tal forma que la matriz Y se puede calcular mediante la *ecuación (17)*.

$$Y = TBQ^T + F (17)$$

Resumiendo el procedimiento para realizar la descomposición de la matriz de las respuestas consta de ciertos pasos como son: calcular los loadings y scores, calcular la matriz residual y calcular una nueva matriz. Finalmente es importante determinar el tamaño óptimo del modelo, decidiendo el número de factores o componentes principales (PC).

#### 1.4.6.1 Factores a incluir en el modelo

- **Detección de outliers.** Gracias a la capacidad que tienen los métodos multivariantes de detectar la observación inconsistente con el resto de los datos, la localización de los outliers en esta etapa de establecimiento del modelo, se puede realizar con la información de la influencia de los objetos en el conjunto de calibración (leverage) y de los residuales, tanto en la propiedad de interés como en la respuesta instrumental.

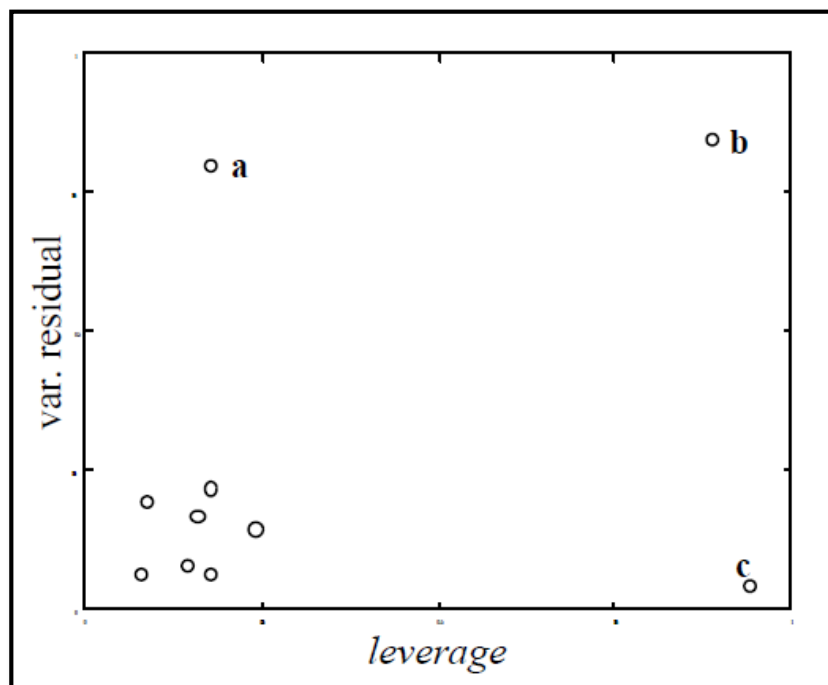
Es importante excluir estas muestras discrepantes en el modelo para minimizar la degradación de la capacidad predictiva del mismo.

- **Leverage.** Es una medida de la influencia de una muestra en relación al modelo. Muestras con un elevado valor de leverage están muy alejadas del centro del modelo, por lo que tendrán una influencia muy alta sobre el mismo.
- **Residuales en la propiedad.** En la etapa de establecimiento del modelo se dispone del valor de la propiedad de interés determinado por el método de referencia. Los residuales en esta propiedad comparan el valor predicho por el modelo multivariante  $\hat{c}$  con el valor considerado verdadero,  $c$ , que proporciona el método de referencia. *Ecuación (19).*

$$f = c - \hat{c} \quad (19)$$

En varias ocasiones la detección de outliers se realiza combinando las herramientas descritas, un buen ejemplo es la figura 7 en el que se representa el residual (propiedad) frente al leverage de las muestras. [45]

**Figura 7.** Gráfico del residual frente al leverage.



Fuente: Miller, N. J. *Estadística y Quimiometría para química analítica*. 2002

**(a)** Objetos con una varianza residual elevada se consideran outliers, **(b)** si además tienen un leverage alto son outliers peligrosos para el modelo, debido a que tienen mucha influencia sobre él. Las muestras con un leverage alto **(c)** son muestras influyentes y no necesariamente outliers.

- **Estadístico  $T^2$  de Hotelling.** Originalmente propuesto por Hotelling; se calcula como la suma de los cuadrados de los scores y mide la variación de cada muestra dentro del modelo PCA. El gráfico  $T^2$  monitoriza la distancia de una nueva medida al valor de referencia en el espacio reducido de los factores PC. Permite detectar si la variación incluida en los componentes principales considerados es más grande que la que le correspondería si solo influyeran variaciones aleatorias, su interpretación gráfica es como la de cualquier gráfico univariante, donde las muestras fuera de control aparecen más allá de la línea de control por poseer un valor de  $T^2$  superior al límite.

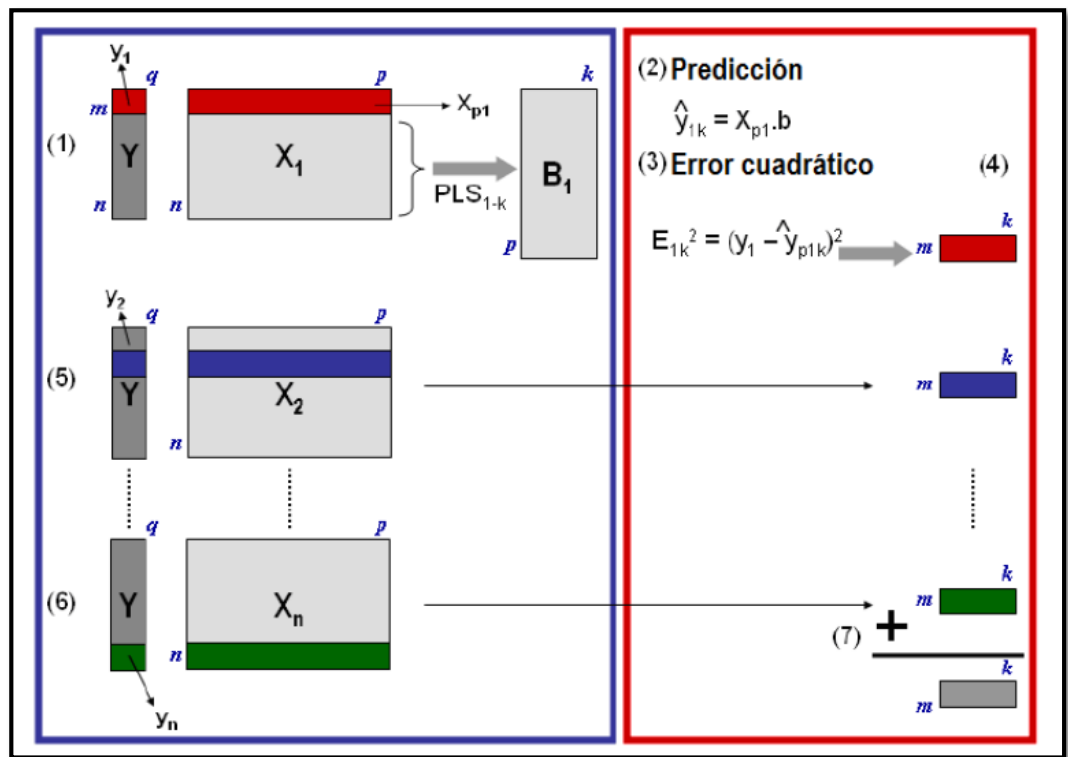
**1.4.7 Validación del modelo PLS.** Debido a que el método de calibración PLS no se apoya directamente en un modelo teórico y puede incorporar viabilidad de los datos no necesariamente relacionada con la propiedad de interés, es necesario validarlo cuantitativa y cualitativamente.

La validación se puede realizar con un análisis de un grupo de muestras independiente al utilizado en la calibración, comprobando que no existe un error sistemático entre las predicciones que realizan el modelo y los valores proporcionados por el método de referencia. Pero, en caso de no disponer de un conjunto independiente de muestras se puede utilizar el error obtenido por validación cruzada, *RMSECV*, dando una estimación de la exactitud del modelo y del error que se comete al predecir muestras futuras. <sup>[46]</sup>

**1.4.7.1 Validación Cruzada.** El fundamento de este proceso de validación es ir construyendo sub-modelos con diferentes secciones de los datos no utilizadas en cada uno de ellos. En la validación cruzada total, frecuentemente aplicada, cada sub-modelo está definido por la totalidad de las medidas excepto una, hasta que

todas queden en algún momento fuera. Así pues, se construyen tantos submodelos como medidas existen. Este procedimiento aproximado es extensamente aceptado y efectivo principalmente en aplicaciones en las que el número de medidas no es excesivamente elevado, en la figura 8 se presenta su esquema general.

**Figura 8.** Esquema general de la validación cruzada



A la izquierda de la anterior figura, modelos PLS con entre 1 y  $k$  variables y sus correspondientes coeficientes. A la derecha, la predicción de la muestra que no hacía parte de la calibración, el cálculo de los residuos cuadrados.

**1.4.8 Errores de Calibración y predicción.** Mediante el análisis de los errores en la regresión, se determina la validez del proceso de calibración y de predicción, y estos son:

**1.4.8.1 Raíz cuadrada del error cuadrado medio de calibración (RMSEC** por sus siglas en inglés): describe el grado de concordancia entre las propiedades estimadas por el modelo de calibración para las muestras de calibración y los valores aceptados como verdaderos para las muestras de calibración utilizados para obtener los parámetros del modelo. De acuerdo a la *ecuación (20)*:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-m-1}} \quad (20)$$

**1.4.8.2 Raíz cuadrada del error cuadrado medio de predicción (RMSEP** por sus siglas en inglés): El modelo de regresión se calcula a partir de los datos del conjunto de calibración, y el modelo calculado, se pone a prueba con los datos del conjunto de validación que no han sido utilizados para calcular la ecuación de calibración del modelo. El número de muestras de validación, p, debe ser grande, de manera que la estimación del error de predicción refleje con precisión todas las fuentes de variabilidad en el método de calibración. De acuerdo a la *ecuación (21)*:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^p (y_i - \hat{y}_i)^2}{p}} \quad (21)$$

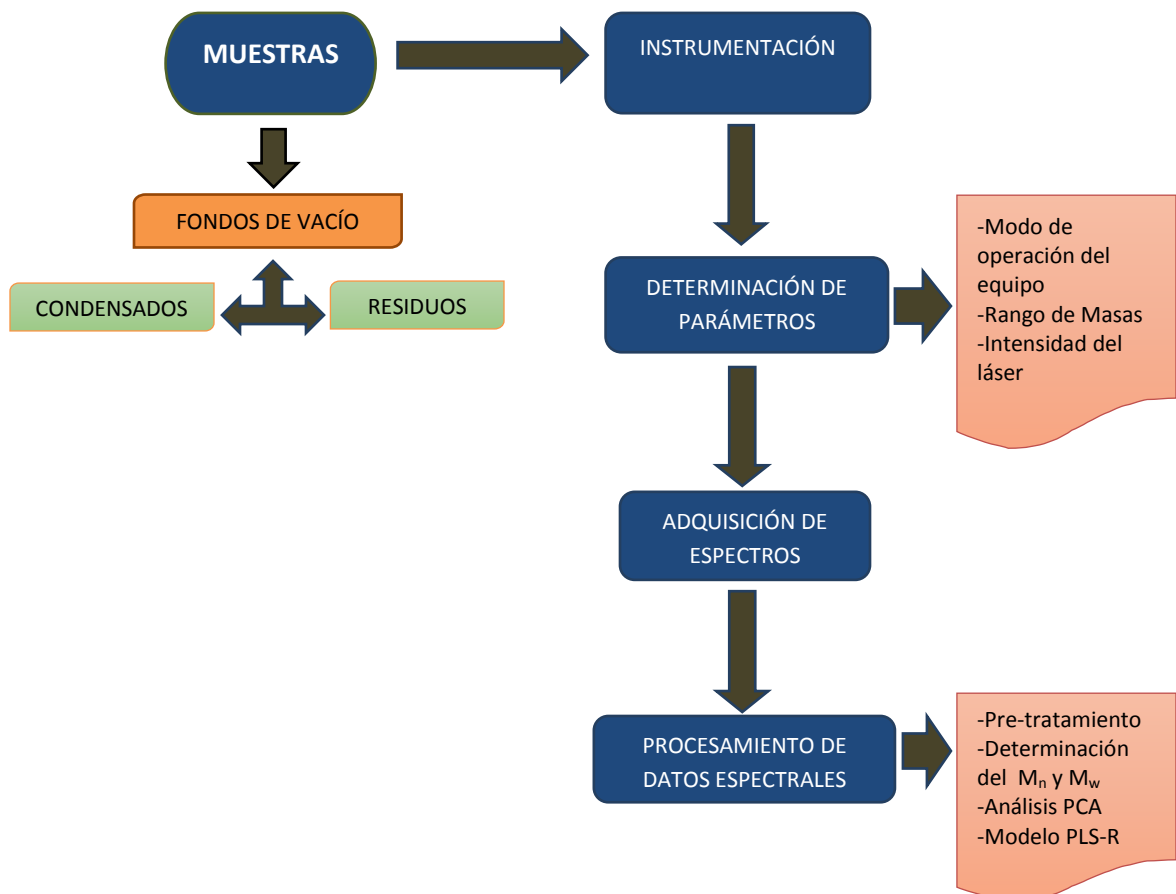
**1.4.8.3 Raíz cuadrada del error cuadrado medio de calibración cruzada** (RMSECV por sus siglas en inglés): Consiste en la utilización de los datos del conjunto de calibración para validar el modelo. El conjunto de calibración inicial se divide en diferentes grupos de cancelación. Cada vez que se calcula el modelo se deja un grupo de cancelación fuera y se calcula la ecuación de la regresión con los demás datos. Una vez realizada la calibración se pone a prueba el modelo con los datos del grupo de cancelación excluido. De acuerdo a la *ecuación (22)*:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^C (\hat{y}_i - y_{CV,i})^2}{n_C}} \quad (22)$$

## 2. PARTE EXPERIMENTAL

En el marco del convenio de cooperación tecnológica Universidad Industrial de Santander (UIS) -Instituto Colombia del Petróleo (ICP), el Laboratorio de Espectroscopia Atómica y Molecular (LEAM) de la UISha venido desarrollando una serie de investigaciones utilizando diferentes técnicas analíticas y quimiometría, enfocadas en el estudio y caracterización de los Crudos Colombianos y sus fracciones. En la figura 9 se muestra la estructura metodológica que se siguió en el presente trabajo.

**Figura 9.** Esquema Metodológico experimental.

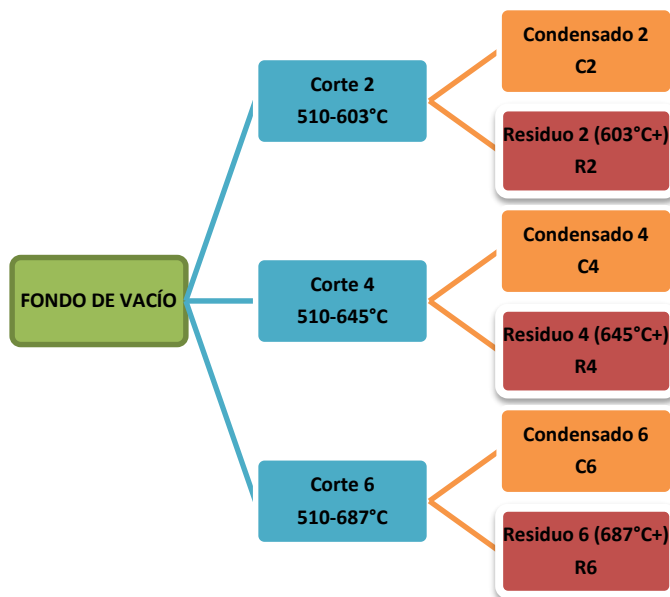


## 2.1 SELECCIÓN DE MUESTRAS

En esta investigación se estudiaron 20 muestras conformadas por fondos de vacío y sus fracciones de condensados y residuos obtenidas en la Unidad de Destilación Molecular (UDM) de crudos Colombianos. Las muestras y sus datos de caracterización estándar fueron suministradas por el Instituto Colombiano del Petróleo (ICP).

En la figura 10 se muestra un esquema general de las fracciones obtenidas del fondo de vacío sometido a cortes con puntos de ebullición iniciales y finales entre 510°C y 687°C. De cada corte se obtiene una fracción liviana llamada condensado y una fracción pesada como residuo de la destilación.

**Figura 10.** Diagrama experimental del fraccionamiento por Destilación Molecular de un fondo de vacío.



**2.1.1 Caracterización de las muestras.**El fraccionamiento de los tres fondos de vacío (VRT,VRG y VRR)se llevó a cabo en la unidad de destilación molecular KD-6-1S (figura11) de Chem. Tech. Services, Inc. Co.Cada una de las fracciones fue caracterizada siguiendo una metodología estándar así: análisis SARA (ASTM D2007), gravedad API (ASTM D70), residuo de carbón micro –MCR (ASTM D4530), contenido de azufre (ASTM D5453), metales níquel y vanadio (ASTM D5708) y nitrógeno básico (ASTM D5291); en la tabla 6 se relacionan los correspondientes datos.

**Figura 11.** Destilador Molecular KD-6-1S (Chem. Tech. Services Inc. Co)



*Fuente: Moreno González, D. H. 2011*

**Tabla 5.** Caracterización fisicoquímica de los fondos de vacío R, T y G, y sus fracciones

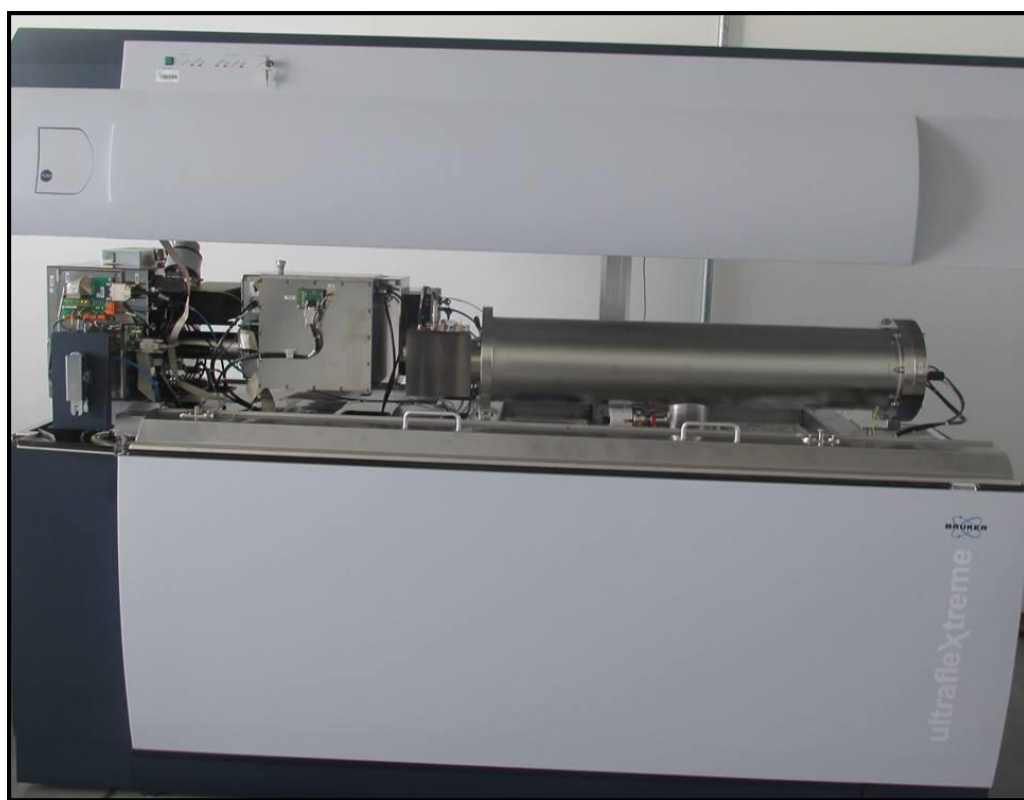
MUESTRA	PROPIEDAD				
	Azufre	Carbono	Hidrógeno	MCR	API
<b>C2G</b>	1.41	85.71	11.76	4.6	15.5
<b>C4G</b>	1.39	85.83	11.89	5.3	15.7
<b>C6G</b>	1.41	86.03	12.16	7.2	15.1
<b>VRG</b>	1.77	85.98	10.88	18.3	9.9
<b>R2G</b>	1.91	86.22	10.73	23.6	7.6
<b>R4G</b>	2	85.5	10.72	27.1	6.0
<b>C2T</b>	1.47	85.93	12.43	5.7	19.3
<b>C4T</b>	1.54	85.7	12.16	5.5	19.2
<b>C6T</b>	1.53	86.1	12.52	8.3	17.7
<b>VRT</b>	2.12	85.35	11.61	18.3	11.6
<b>R2T</b>	2.23	85.88	11.21	19.4	11.9
<b>R4T</b>	2.47	86.1	11.25	24.7	9.2
<b>R6T</b>	2.73	85.47	10.15	29.5	5.7
<b>C2R</b>	1.49	86.96	11.04	10.4	10.2
<b>C4R</b>	1.52	86.76	11.61	8.7	10.7
<b>C6R</b>	1.45	86.52	11.81	9.0	10.5
<b>VRR</b>	2.11	87.34	9.52	35.1	2.0
<b>R2R</b>	2.15	86.52	9.02	38.0	0.4
<b>R4R</b>	2.23	86.98	9.04	41.0	-0.5
<b>R6R</b>	2.29	86.96	8.6	44.5	-2.1

Los tres fondos de vacío que en adelante tendrán el nombre teórico R, G y T, se consideran prototipos y modelo de las cargas de refinería de Ecopetrol y como se ilustra en la Tabla 5 presentan una escala que abarca de 2 a 11,6 °API.

## 2.2 INSTRUMENTACIÓN

Se utilizó el espectrómetro de masas TOF/TOF, ultrafleXtreme (Bruker Daltonik GmbH, Bremen, Germany), como el ilustrado en la figura 12.

**Figura 12.** Imagen fotográfica del Espectrómetro de Masas



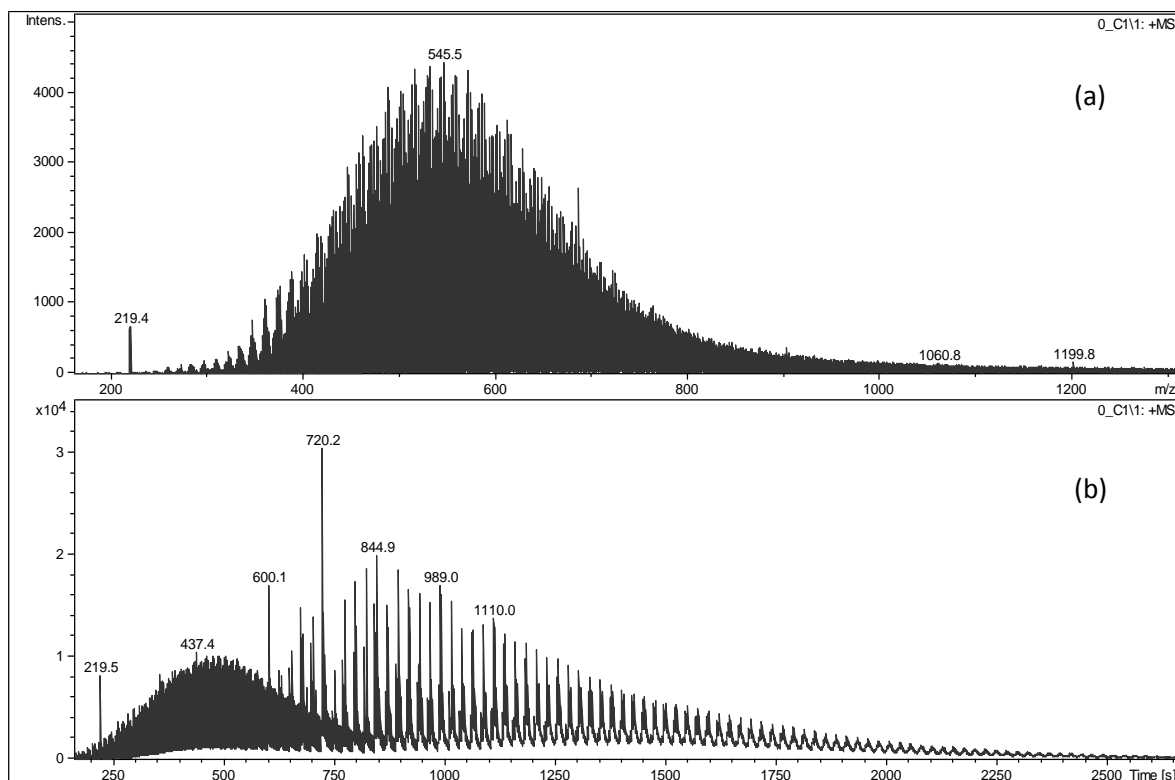
## 2.3 DETERMINACIÓN DE PARÁMETROS INSTRUMENTALES

A continuación se describe la metodología usada para la determinación de los mejores parámetros instrumentales para el análisis por Espectrometría de Masas de las muestras de Fondos de Vacío y sus fracciones en el Laboratorio de Espectroscopía Atómica y Molecular de la UIS.

**2.3.1 Modo de operación del equipo.** El equipo fue usado en modo reflectrón positivo, debido a que en este modo se cuenta con la mayor resolución y sensibilidad hacia masas de alto peso molecular. La fuente de ionización es un láser de tipo SmartBeam-II de Granate de Aluminio e Ytrio dopado con Neodimio (Nd:YAG a 355 nm) y con anchos de pulso del orden de nanosegundos.

**2.3.2 Rango de Masas.** El rango de masas que se utilizó para la toma de espectros fue de 180 a 5000 Da con precisión de 1ppm, pues en este rango se observaron las distribuciones de masas de las muestras.

**2.3.3 Intensidad del láser.** Se escogió una intensidad laser equivalente al 60% de su energía total ya que a mayor intensidad se observa una distribución adicional que se extiende hasta 2500 Da con picos espaciados entre sí 24 Da (ver figura 13(b)) y que podría corresponder a la formación de clusters debido a la ablación láser. <sup>[47]</sup> Cada adquisición corresponde a la suma de 500 disparos láser que fueron adquiridos de forma aleatoria en el spot para cada muestra.



**Figura 13.** Espectro de la fracción R2R obtenido a intensidad láser de (a) 60% y (b) 90%

**2.3.4 Preparación de muestras.** Las muestras fueron preparadas por el método de la gota seca. En este método, la muestra se disuelve en cloroformo a una concentración de 1 mg/ml para los destilados y 0,2 mg/ml para los fondos de vacío y los residuos. Se coloca 1 µl de la solución en un spot sobre una placa de acero inoxidable (target) para su subsecuente evaporación al aire.

## **2.4 ADQUISICIÓN Y ANÁLISIS DE ESPECTROS**

La adquisición de espectros se realizó mediante el software de análisis FlexControl, que permite la variación de diferentes parámetros como: intensidad láser, rango de masas, método de adquisición de los espectros, voltaje de extracción de las placas, entre otros.

Bajo las condiciones experimentales establecidas se obtuvieron espectros con un rango de masas de 250-950 m/z.

## **2.5 PROCESAMIENTO DE DATOS ESPECTRALES**

Se realizó un pre-tratamiento a los espectros obtenidos utilizando las herramientas del software de procesamiento FlexAnalysis 3.3.

**2.5.1 Pre-tratamiento de datos.** Inicialmente, con el fin de detectar influencias de posibles errores instrumentales y humanos durante el manejo de muestras y posterior adquisición de espectros de masas, se realizó una exploración visual a los mismos.

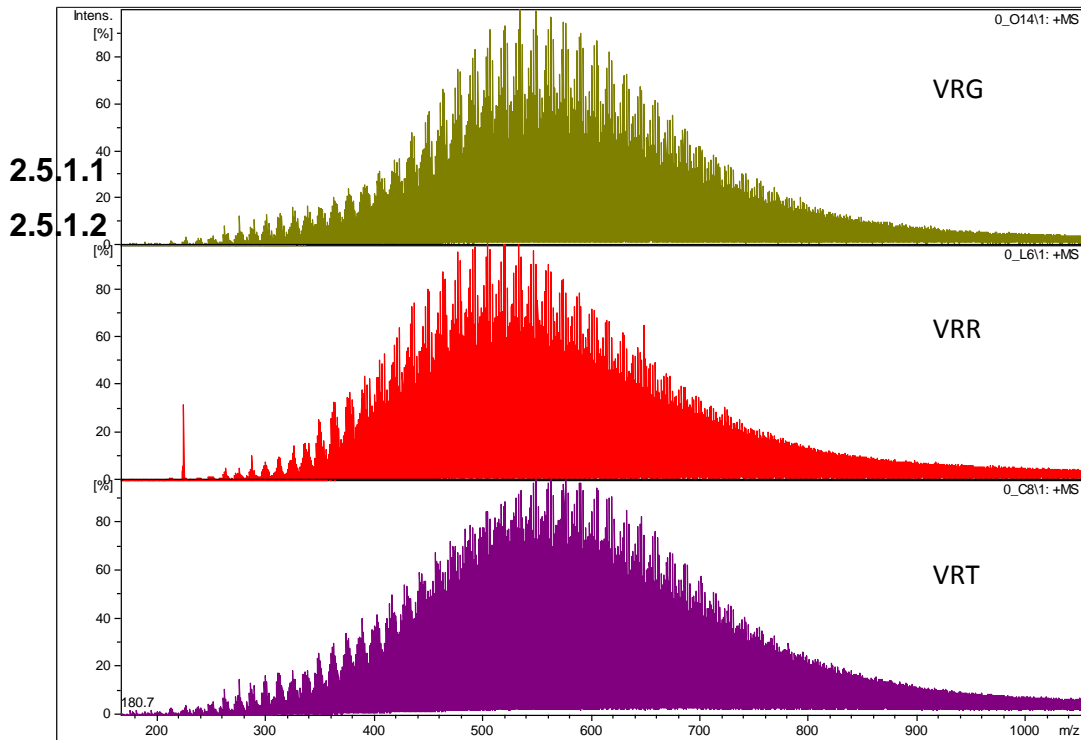
El pre-tratamiento de los datos se realizó con el software FlexAnalysis, con el fin de obtener la máxima y más útil información de los datos, minimizando las contribuciones presentes en la señal analítica que reducen la reproducibilidad e inducen comportamientos del sistema que darían lugar a estimaciones erróneas de los parámetros esperados.

Teniendo en cuenta que esta técnica analítica requiere calibración antes de la toma de espectros con celebrantes que cubran el rango de masas de las muestras se utilizó una mezcla de péptidos (patrón estándar) como calibrante, el cual cubre un rango de 700 a 3000 Da. Esta mezcla fue depositada en un spot del target lo más cercano posible a la muestra a analizar.

A continuación, se realizó una calibración interna del espectro de masas con la del calibrante más próximo según su ubicación en el target. Finalmente, se implementó un método de trabajo definiendo las características de los parámetros de suavizado espectral -para reducir el ruido aleatorio que existe en la señal analítica-, y de la corrección de la línea base -para intentar corregir ciertas tendencias en ésta que aporta ruido espectral a la señal-. También, se especificó el valor de la señal ruido (S/N) de aproximadamente 1.8, el ancho de pico (0.002 m/z), la altura del 80% y el número máximo de picos (rango entre 2000 y 3000), todo lo anterior para obtener una mejor identificación de los picos.

**2.5.2 Determinación del peso molecular promedio por LDI-TOF/TOF.** Luego del procesamiento (calibración, suavizado espectral, corrección de la línea base y finalmente determinación del número de picos en el rango de masas evaluado) de los espectros de masas de los fondos de vacío (figura 14) y sus fracciones, se calculó el *peso molecular promedio en número* ( $\overline{Mn}$ ) y el *peso molecular promedio en peso* ( $\overline{Mw}$ ) con las expresiones dadas según las ecuaciones 23 y 24, respectivamente.<sup>[48]</sup>

**Figura 14.** Espectros de Masas finales de los fondos de vacío VRT, VRG y VRR



**2.5.2.1 Ecuaciones para determinar el Peso Molecular.** Con la *ecuación (23)* se determina  $\overline{M}_n$  donde  $I_i$  es la intensidad relativa de la masa  $m_i$ , esta ecuación es afectada principalmente por las moléculas de bajos pesos moleculares en el espectro.

$$\overline{M}_n = \frac{\sum I_i M_i}{\sum I_i} \quad (23)$$

Con la ecuación(24) se calcula  $\overline{M_w}$  que es la medida aritmética de los pesos moleculares de todas las moléculas. Es afectado principalmente por las moléculas de altos pesos moleculares, debido a que no solamente depende del número de especies (moléculas o agregados) de cada peso, como  $\overline{M_n}$ , sino que depende del peso de cada una de las especies.

$$\overline{M_w} = \frac{\sum I_i M_i^2}{\sum I_i M_i} \text{(24)}$$

Donde,  $M_i$  y  $I_i$  representan los pesos moleculares de los componentes y sus intensidades de señal respectivamente (suponiendo una relación lineal entre el número de iones y la intensidad de la señal) de las especies detectadas. <sup>[49]</sup>

**2.5.3 Análisis por componentes Principales (PCA).** Se utilizó el software Unscrambler versión 10.3 para el análisis por PCA. En este análisis se escogió una matriz de 99.999 datos que corresponde a un rango de masas de 270 – 900 m/z para cada muestra. Debido a que el software Unscrambler tiene diferentes tipos de normalización para la corrección de los datos espectrales; se realizó un análisis exploratorio por el cual se identificó la Normalización Media (Mean Normalization) como la más apropiada, ya que logra la mejor agrupación de las muestras con alta varianza explicada.

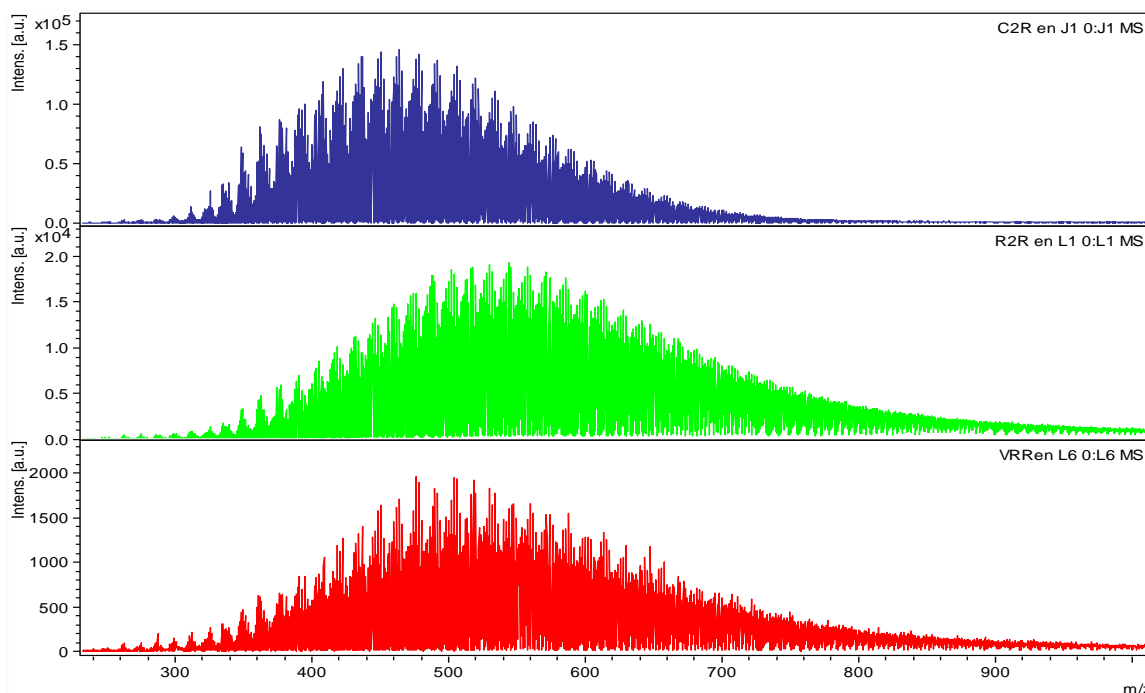
**2.5.4 Regresión por Mínimos Cuadrados Parciales (PLS-R).** Mediante la Regresión PLS se evaluó el potencial de la técnica Espectrometría de Masas LDI-TOF/TOF para predecir la gravedad API a muestras de fondos de vacío y sus fracciones de la Unidad de Destilación Molecular.

Para el desarrollo de este modelo se realizó la respectiva calibración y validación, se evaluaron los parámetros estadísticos RMSEC y RMSECV y la varianza explicada para identificar el número óptimo de Variables Latentes (VL).

### 3. RESULTADOS Y ANALISIS

#### 3.1 DETERMINACIÓN DEL PESO MOLECULAR PROMEDIO POR ESPECTROMETRÍA DE MASAS LDI-TOF/TOF

La inspección visual de los espectros de masas de los fondos de vacío y las fracciones de la Unidad de Destilación Molecular, permitió diferenciar la fracción de condensados de los demás productos analizados, observándose un desplazamiento del pico más intenso del espectro color azul (figura 15) en esta fracción, comparada con el fondo de vacío y la fracción residuos. Esto, probablemente se debe a que en los condensados aumenta la concentración de hidrocarburos livianos (pesos moleculares cada vez menores), por la disminución de la concentración de asfáltenos. Los espectros correspondientes a las demás muestras se presentan en el Anexo A.



**Figura 15.** Espectros de Masas de C2R, R2R y VRR

En la tabla 6 se relacionan los pesos moleculares para las muestras del primer corte de destilación con repetibilidad espectral, y en la figura 16 se puede observar una gráfica de la relación entre  $\overline{Mn}$  de los condensados R de cada corte con su respectiva repetibilidad espectral.

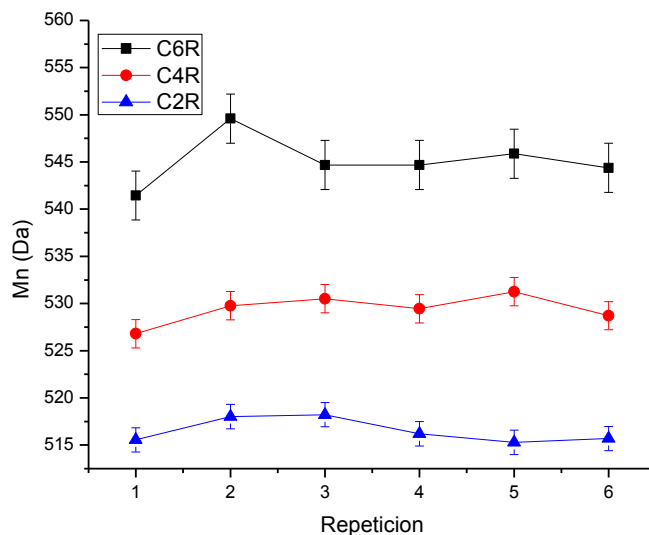
**Tabla 6.**Datos peso molecular promedio para C2R, R2R y VRR

Ensayo	C2R		R2R		VRR	
	$\overline{Mn}$	$\overline{Mw}$	$\overline{Mn}$	$\overline{Mw}$	$\overline{Mn}$	$\overline{Mw}$
1	515,549	541,591	658,904	714,760	625,099	676,527
2	518,007	545,327	680,196	750,318	650,396	711,766
3	518,220	545,490	679,614	750,689	667,252	736,084
4	516,192	542,231	680,484	749,676	655,274	718,511
5	515,290	540,991	677,359	747,006	646,642	707,418
6	515,689	541,659	681,725	753,314	649,452	712,491
<b>Media</b>	516,491	542,881	676,380	744,294	649,019	710,466
<b>Error Estándar</b>	0,528	0,815	3,544	5,964	5,632	7,933
<b>Desviación Estándar Relativa</b>	0,250	0,368	1,283	1,963	2,125	2,735

En este caso para las muestras R, el cálculo de la desviación estándar relativa (RSD) presenta valores menores al 3%, lo cual demuestra la repetibilidad y fiabilidad de los resultados.

De acuerdo a los datos reportados (ver Anexo B) e ilustrados en la figura 16, se observa un incremento en los valores de  $\overline{Mn}$  en el orden C2R, C4R y C6R. En general, este resultado se evidenció también para las muestras G y T; esto se puede explicar por el incremento de la aromaticidad y la relación C/H en las muestras con la temperatura en el proceso de Destilación Molecular.<sup>[50]</sup> Por lo tanto, se observa que el peso molecular medio de los destilados aumenta con la temperatura de destilación.

**Figura 16.** Gráfica de  $\overline{Mn}$  de los Condesados de Rubiales con su repetibilidad espectral



De los resultados obtenidos y relacionados en la tabla 7 y la gráfica 17 se puede concluir que, los residuos presentan un mayor peso molecular ( $\overline{Mn}$ ) que los fondos

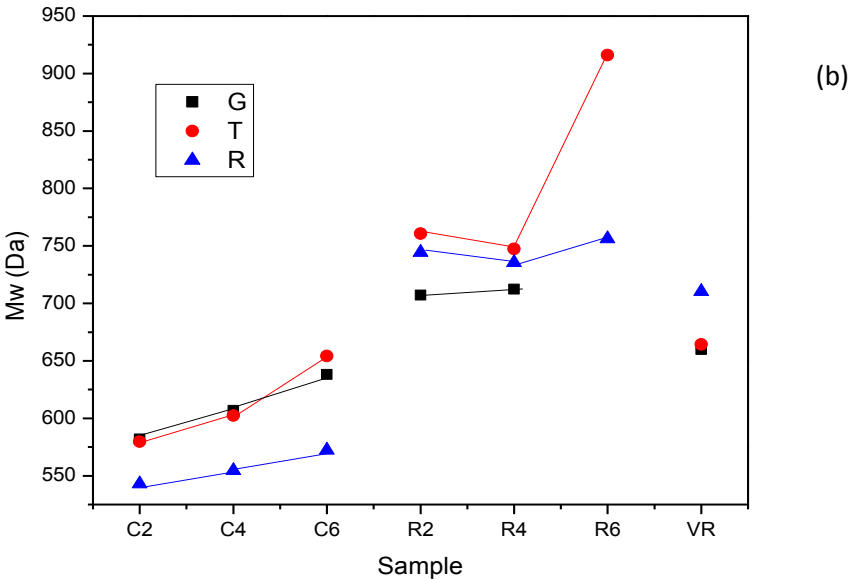
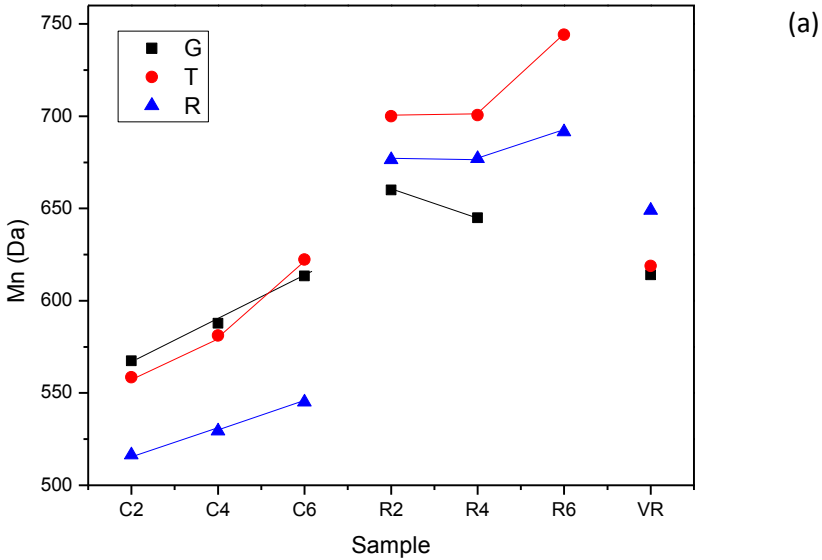
de vacío, esto debido a que tras el proceso de destilación molecular estas fracciones tienen un alto contenido de resinas y asfaltenos quienes tienen una alta complejidad composicional y altos puntos de ebullición. Adicionalmente, de estas fracciones fueron removidas las especies con bajos puntos de ebullición lo que contribuye al aumento de su peso molecular medio.<sup>[51]</sup>

- ✓ El peso molecular promedio estimado de las fracciones de condensados de los tres fondos de vacío oscila entre 542,88 y 654,20 Da.
- ✓ Los residuos considerados como fracción pesada de los fondos de vacío, tienen un rango de peso molecular más alto que los condensados y los propios fondos, que puede variar entre 707,18 y 756,16 Da.
- ✓ Los fondos de vacío tienen un peso molecular promedio que oscila entre 659,80 y 710,46Da.

**Tabla 7.** Valores de  $\overline{M}_n$  y  $\overline{M}_w$  para los fondos de vacío y sus fracciones

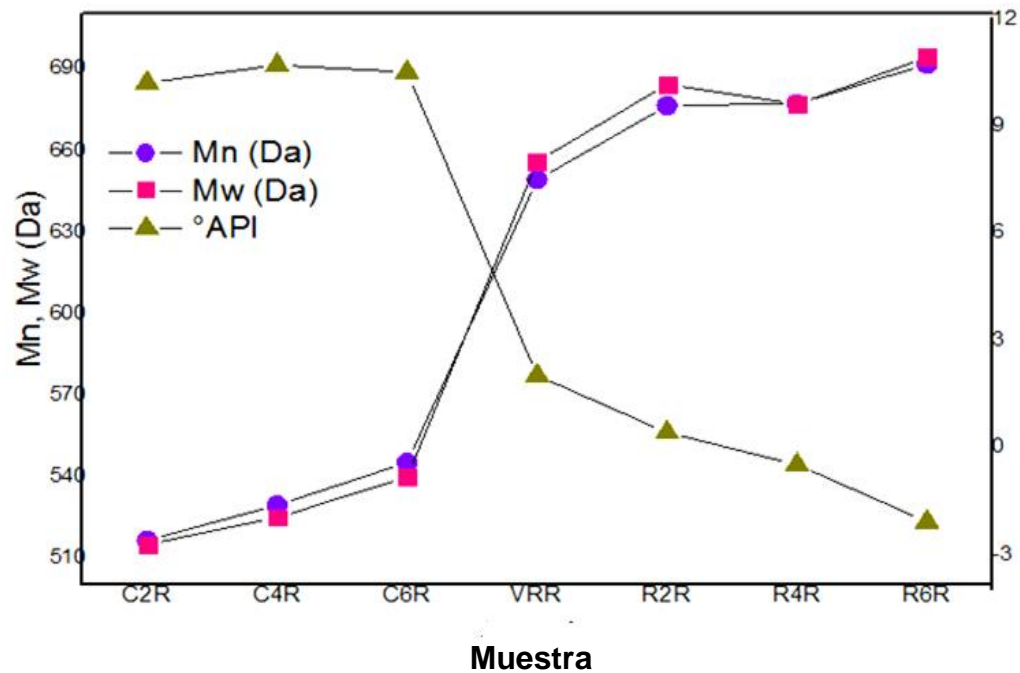
Muestra	$\overline{M}_n$	$\overline{M}_w$	Muestra	$\overline{M}_n$	$\overline{M}_w$	Muestra	$\overline{M}_n$	$\overline{M}_w$
<b>C2G</b>	567,47	581,84	<b>C2T</b>	558,53	579,71	<b>C2R</b>	516,49	542,88
<b>C4G</b>	587,79	606,81	<b>C4T</b>	581,08	602,43	<b>C4R</b>	529,41	554,59
<b>C6G</b>	613,43	638,10	<b>C6T</b>	622,32	654,20	<b>C6R</b>	545,11	572,25
<b>R2G</b>	660,00	707,18	<b>R2T</b>	700,00	760,66	<b>R2R</b>	676,38	744,29
<b>R4G</b>	644,89	712,43	<b>R4T</b>	700,63	747,42	<b>R4R</b>	677,08	735,52
--	--	--	<b>R6T</b>	744,20	915,87	<b>R6R</b>	691,59	756,16
<b>VRG</b>	614,07	659,80	<b>VRT</b>	618,79	664,31	<b>VRR</b>	649,02	710,46

**Figura 17.** Gráficas del peso molecular promedio para fondos de vacío y sus fracciones de condensados y residuos.



Es importante tener en cuenta el grado de correlación de las propiedades fisicoquímicas de los fondos de vacío y sus fracciones, para esto, se realizó una gráfica comparativa del peso molecular medio de las muestras y su correspondiente °API (figura 18). Como se observa en la gráfica los destilados presentan un menor peso molecular promedio y a su vez una °API mayor, igualmente las muestras más pesadas tienen una mayor °API. Esto permite evidenciar una correlación importante entre la °API y los espectros de masas obtenidos para los fondos de vacío y sus fracciones, por lo que podría encontrarse un modelo matemático de regresión.

**Figura 18.** Gráfica de  $\overline{M}_n$ ,  $\overline{M}_w$  y °API de las muestras R



### 3.2 ANALISIS POR COMPONENTES PRINCIPALES (PCA)

Se realizó una descomposición por componentes principales al conjunto de 20 espectros de muestras entre, fondos de vacío, destilados y residuos de diferentes crudos Colombianos, corregidos con pre-tratamiento espectral.

En este análisis se definió el número de componentes principales (PC) a utilizar en el modelo de calibración, con el fin de disminuir la dimensionalidad de los datos lo más posible. Debido a que no existe una regla definida sobre el número de PC que se debe utilizar, ésta se determinó a través de la varianza explicada acumulada (tabla 9).

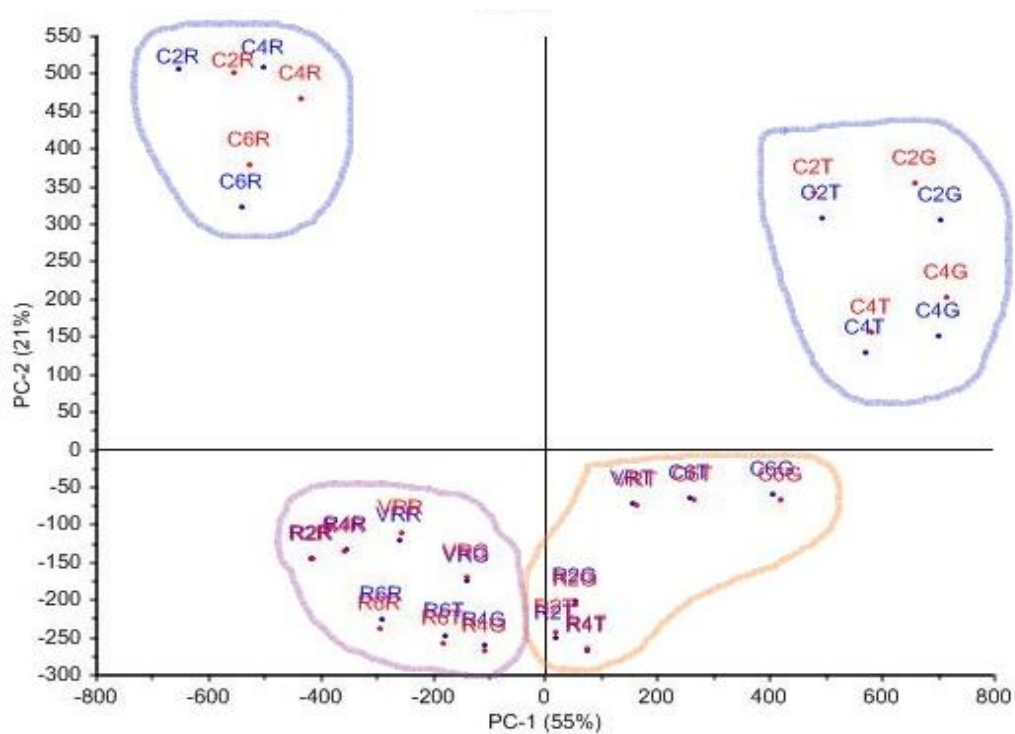
**Tabla 8.** Varianza explicada por 10 componentes principales de las 20 muestras

<b>Componente Principal</b>	<b>Varianza explicada %</b>	<b>Varianza explicada acumulada %</b>
<b>PC 1</b>	54,65875	54,65875
<b>PC 2</b>	21,47241	76,13116
<b>PC 3</b>	6,27665	82,40781
<b>PC 4</b>	5,60246	88,01027
<b>PC 5</b>	4,0639	92,07417
<b>PC 6</b>	2,92229	94,99646
<b>PC 7</b>	1,63911	96,63557
<b>PC 8</b>	0,87178	97,50735
<b>PC 9</b>	0,62771	98,13506
<b>PC 10</b>	0,50417	98,63923

Teniendo en cuenta el porcentaje de varianza de los datos espectrales que explica los primeros 10 componentes principales, se encontró que a partir del PC8 el aumento en la varianza explicada no es considerable, por lo que 7 es el número óptimo de componentes a utilizar en los modelos de calibración.

La representación de los dos primeros componentes principales se muestra en el gráfico de puntuaciones en la figura 19. En el primer cuadrante aparecen las muestras que son explicadas de una forma positiva tanto por el PC1 como por el PC2. En el segundo aparecen las que son explicadas de una forma negativa por el PC1 pero positiva por el PC2. En el tercer cuadrante los de forma negativa para los dos componentes. Y para el cuarto cuadrante aparecen las que son explicadas de una forma positiva para el PC1, pero negativa para el PC2.

**Figura 19.** Gráfico de puntuaciones (Scores) PC2 vs PC1



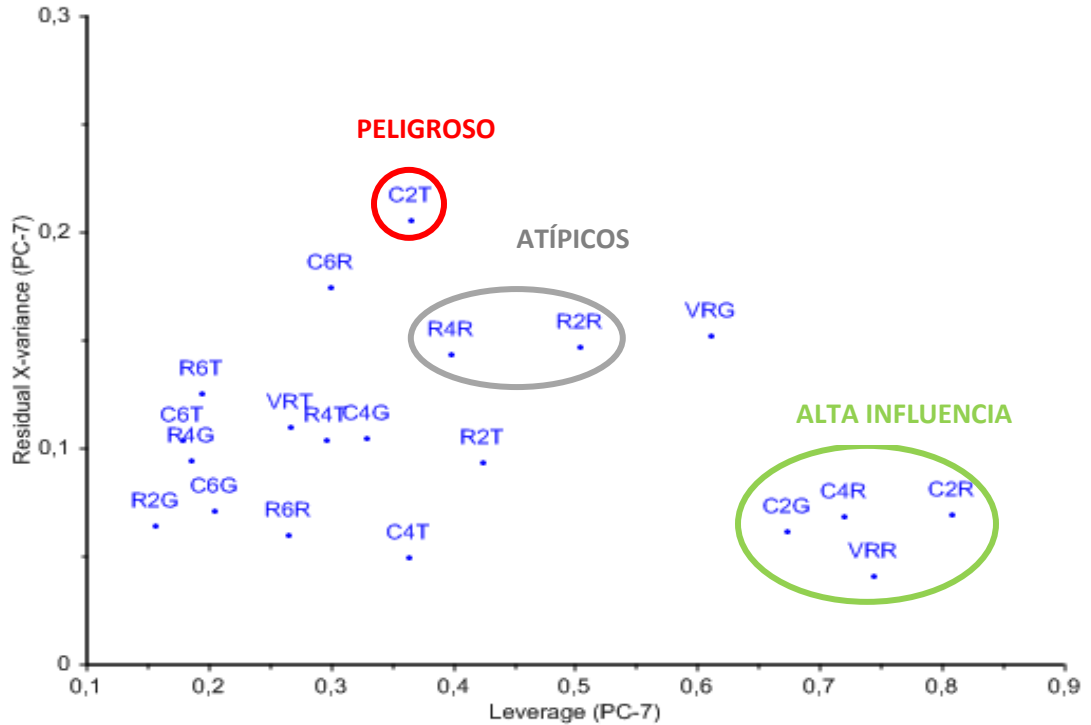
La gráfica de puntuaciones permite identificar tres grupos de muestras para los fondos y sus fracciones como: **Livianos** (azul), **medianos**(anaranjado) y **pesados**(morado), atendiendo a las semejanzas y/o diferencias fisicoquímicas de cada agrupación (gravedad API).

Se observa una muy buena agrupación para los condensados, explicando de forma positiva al PC1 los Condensados T y G, fracciones un poco más livianas que los Condensados R, que explican muy bien el PC2.

Las muestras de los Residuos y fondos de vacío se agrupan en los dos últimos cuadrantes debido a las similitudes fisicoquímicas que estos dos presentan, y se evidencia que las muestras R se caracterizan por su distribución en el gráfico de scores como las más pesadas del conjunto total de muestras y en relación a las que provienen de los crudos T y G.

El gráfico de Influencia (figura 20) permite clasificar de manera detallada e identificar con mayor precisión aquellas muestras que de alguna manera podrían distorsionar el modelo, como: valores atípicos (outliers), de alta influencia y posibles valores peligrosos (dangerous outliers) de las muestras analizadas.

**Figura 20.**Gráfico de Influencia con varianza residual X vs influencia



Las muestras ubicadas en la parte superior izquierda de esta gráfica presentan *altos valores de varianza residual* y pueden ser catalogadas como posibles outliers. Las muestras ubicadas en la parte inferior derecha de la gráfica son catalogadas como muestras con alta influencia para los modelos, denominadas *muestras con alto leverage* (apalancamiento), lo que significa que estas muestras atraerán el modelo para que las describa mejor; *las muestras con valores altos de leverage y varianza residual* se definen como dangerous outliers, lo que significa que no serán descritas por el modelo que describe la mayoría de las muestras, éstas distorsionan el modelo con el fin de estar mejor descritas.

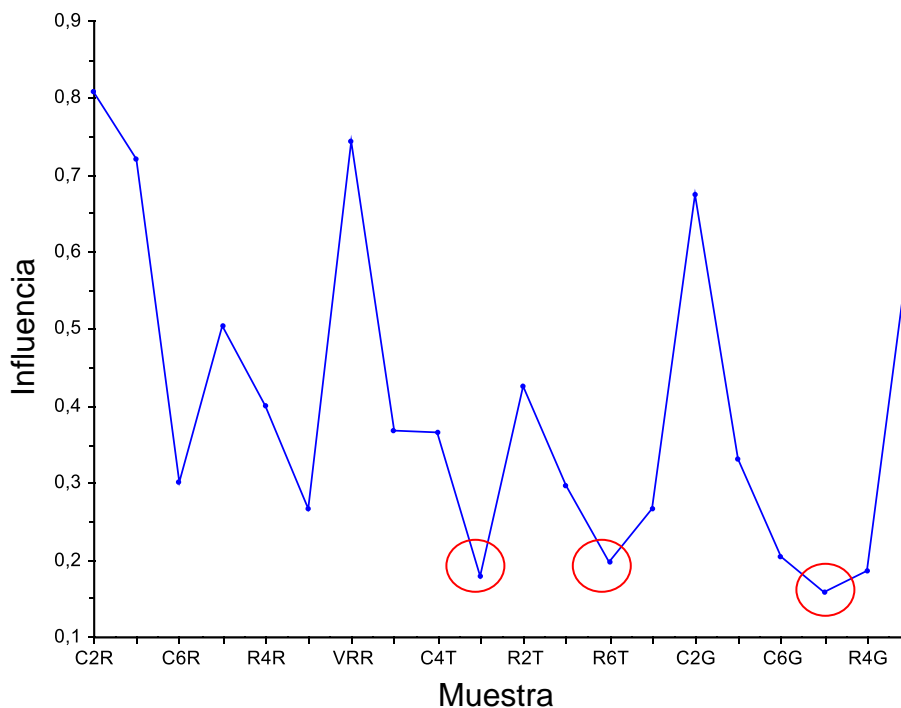
Teniendo en cuenta lo anterior y con la información suministrada en la gráfica de Influencia se observa que la muestra C2T presenta alta varianza residual o Q-residual y alto influencia, clasificándose como un valor potencialmente peligroso, por esto se considera su exclusión para el desarrollo del modelo PLS; revisando las propiedades fisicoquímicas de esta muestra con respecto a las demás, se encuentra que C2T presenta el valor más alto de gravedad API.

Las fracciones de residuos R (R2R, R4R) se pueden catalogar como posibles outliers ya que presentan valores altos de varianza residual, son muestras que por sus bajos valores de gravedad API pueden en definitiva no ajustarse al modelo.

Las muestras que presentan alta influencia como las C2R, VRR, C4R y C2G, atraerán el modelo para que las describa o modele mejor, en estas muestras se aprecia un rango de distribución de los valores de la gravedad API que va desde 2 hasta 15,5 °API, podría esperarse que alguna de estas muestras no sea modelada satisfactoriamente, teniendo en cuenta que su °API puede estar muy distante del valor medio (9,78) de las muestras.

La gráfica de influencia (figura 21), permitió apreciar aún más el carácter negativo que representan las muestras que presentan valores más bajos valores de influencia (C6T, R6T, R2G), considerándose como posibles outlier en el desarrollo del modelo.

**Figura 21.** Gráfico de influencia vs. Muestra



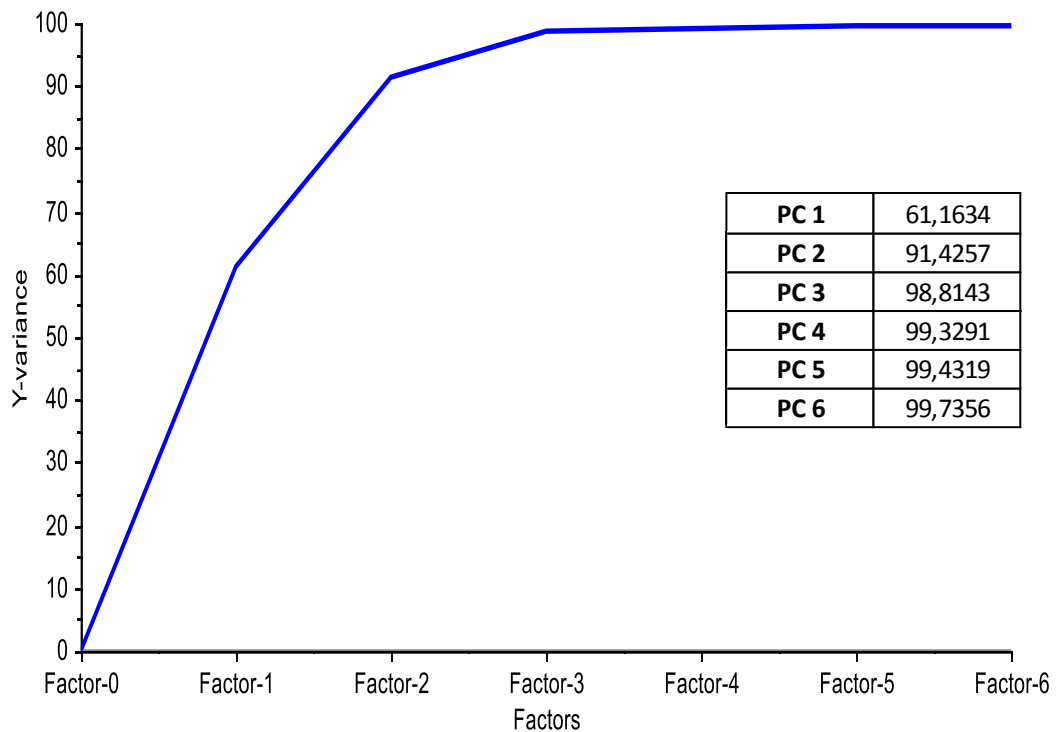
### 3.3 DESARROLLO DE MODELO DE REGRESIÓN PLS

Mediante el software TheUnscrambler® versión 10.3<sup>[52]</sup> licenciado por la UIS y el algoritmo de Regresión por mínimos cuadrados parciales (PLS-R), se analizaron los datos espectroscópicos para predecir la gravedad API en muestras de fondos de vacío y sus fracciones de condensados y residuos de la Unidad de Destilación Molecular. El desarrollo de este modelo de predicción abarcó dos grandes etapas: calibración y validación, procedimientos que se detallan a continuación.

**3.3.1 Calibración del modelo PLS-R.** Debido al número de muestras disponibles (veinte) se utilizó el conjunto total de muestras para construir el modelo de calibración.

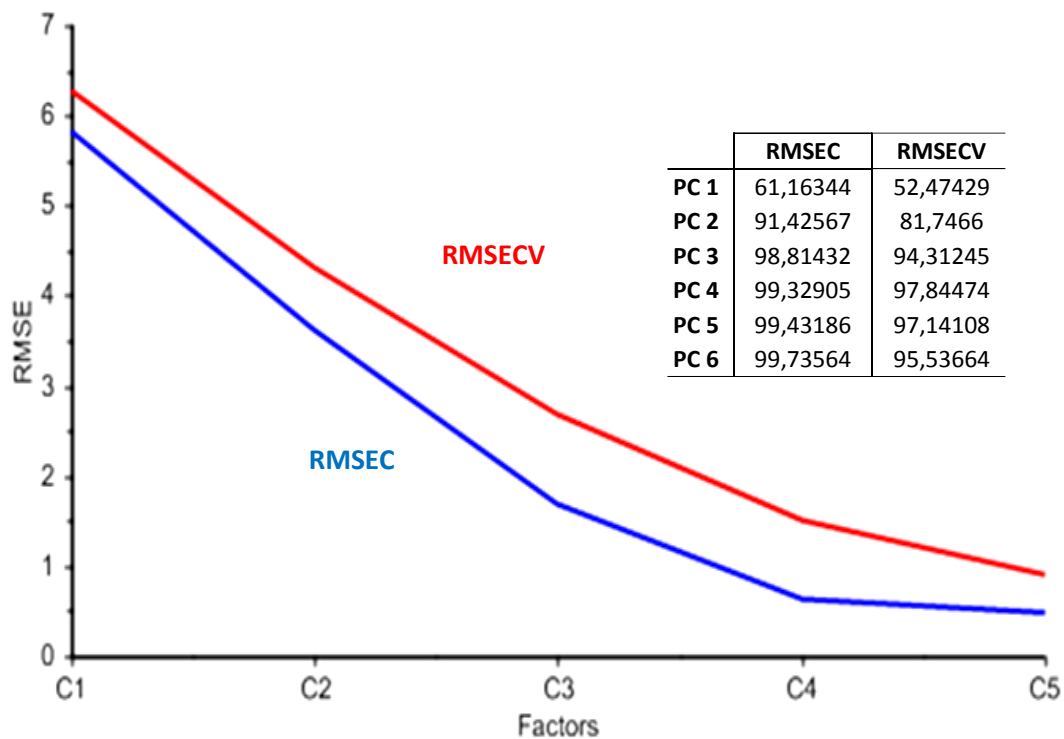
El modelo de predicción de la gravedad API explica más del 99% en la variabilidad de los datos de referencia (variable Y) a partir de los espectros de masas LDI-TOF/TOF, con un mínimo de 4 componentes principales (figura 22).

**Figura 22.** Varianza Explicada para la calibración en el modelo PLS



Esta selección del número óptimo de componentes principales (PC) o factores, requirió además evaluar los parámetros estadísticos RMSEC y RMSECV, errores estándar de calibración y validación, respectivamente (figura 23); donde, el primero mide la desviación estándar de los residuales obtenidos por la diferencia entre los valores observados y los predichos por el modelo para las muestras de calibración, y el segundo es un algoritmo iterativo que relaciona las muestras dentro de un grupo de calibración para desarrollar el modelo de predicción y posteriormente evaluarlo sobre las muestras restantes. [53] Se observa como disminuyen RMSEC y RMSECV a medida que aumenta el número de factores, a partir del PC 4 estos valores permanecen constantes.

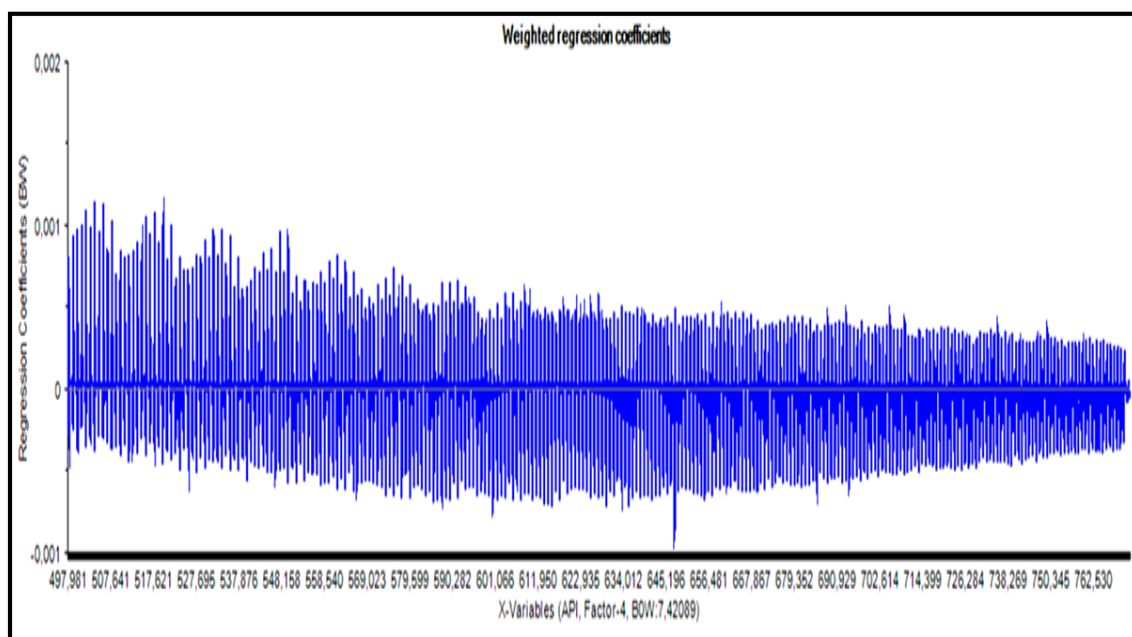
**Figura 23.** Error calculado en función del número de componentes para el modelo



El modelo inicialmente contaba con 47898 variables x, del cual se eliminaron las variables x de menor importancia (con coeficientes de regresión cercanos a cero) aumentando así, el coeficiente de correlación ( $R^2$ ) y disminuyendo la raíz cuadrada media del error de calibración *RMSEC*, hasta que permanecieron estables con valores menores a uno; este procedimiento se llevó a cabo de forma sistemática para la variable latente óptima (factor 4).

En la figura 24 se observan los coeficientes de regresión para el número óptimo de componentes (factor 4) que explica la mayor variabilidad en los datos. Se observa que no hay regiones parcializadas en el rango espectral (497,981 – 762,530), por el contrario, existen datos de m/z con intensidades que contribuyen tanto positiva como negativamente al componente principal.

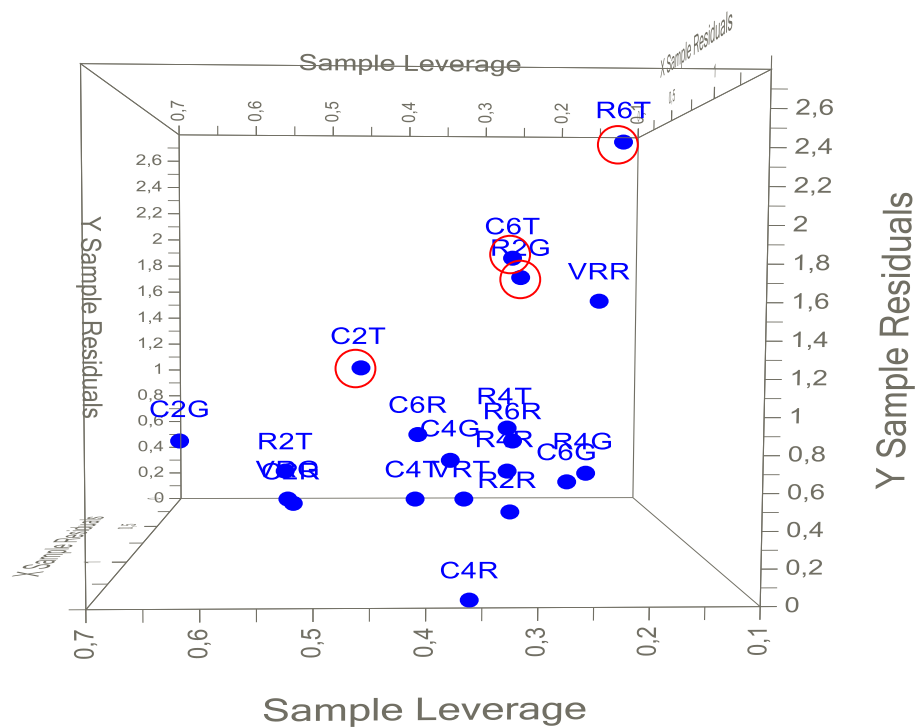
**Figura 24.** Coeficientes de regresión para el cuarto factor del modelo PLS de gravedad API



**3.3.1 Análisis de muestras excluidas.** En el desarrollo de este modelo, además de la reducción de variables, se excluyeron varias muestras teniendo en cuenta los resultados del análisis por componentes principales.

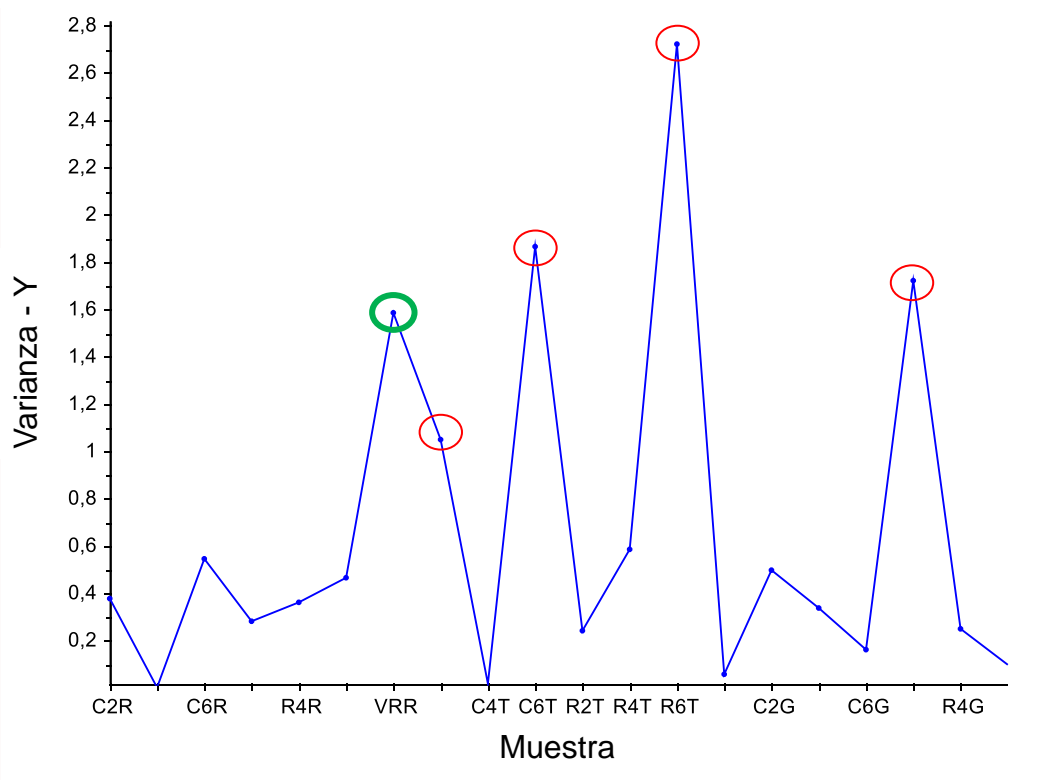
Se eliminaron del modelo seis muestras de las cuales C6T, R2T, R2G y C2T se consideraron posibles outliers, según el PCA. Además, se confirmó el carácter negativo de estos condensados y residuos para el modelo, teniendo en cuenta la gráfica de Influencia de la Regresión PLS (figura 25).

**Figura 25.** Gráfico de Influencia del modelo PLS-R



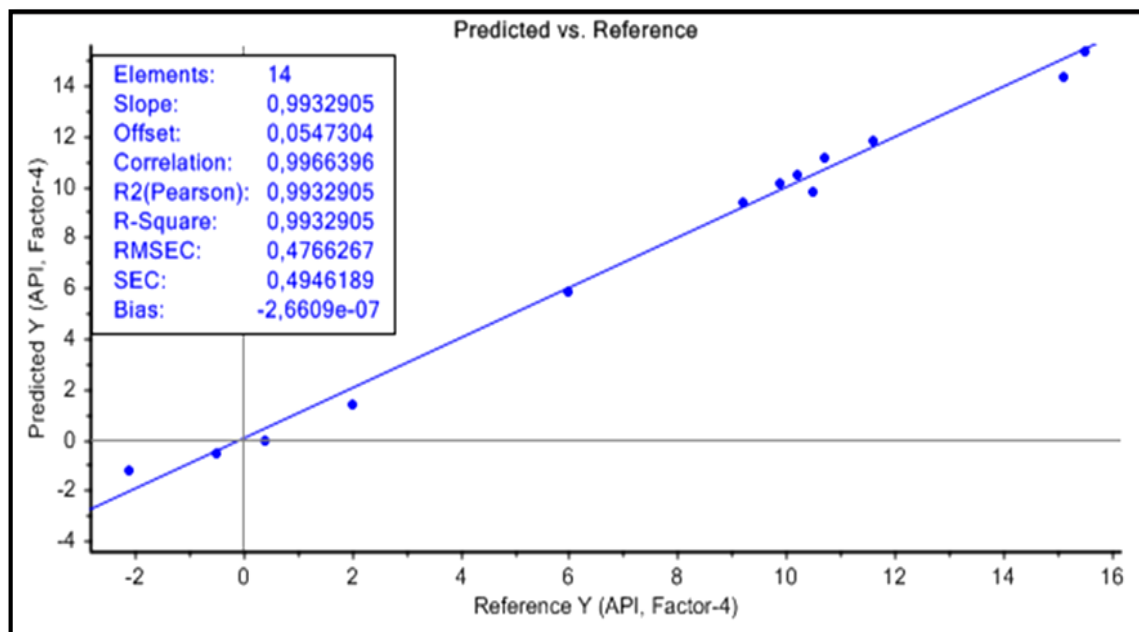
Como se observa en la gráfica de Varianza residual para cada muestra (figura 26), las fracciones R6T, C6T, R2G y C2T presentan los mayores valores de varianza en Y, ratificando que no explicaban el modelo. También, se identificó como muestra atípica el fondo de vacío VRR (verde), la cual según el PCA presentaba alta influencia en el desarrollo del modelo.

**Figura 26.** Varianza Residual en Y para cada muestra PLS-R



En la curva de valores de referencia contra valores predichos (figura 27) se identifican los parámetros estadísticos para esta etapa de calibración (azul).

**Figura 27.** Curva de valores de referencia contra valores predichos en la calibración



Los parámetros estadísticos evaluados en el modelo PLS-R para predecir la gravedad API con 4 componentes principales en muestras de fondos de vacío, condensados y residuos de la UDM, se relacionan en la tabla 9.

**Tabla 9.** Parámetros del modelo seleccionado para la predicción de Gravedad API en fondos de vacío y sus fracciones de DM

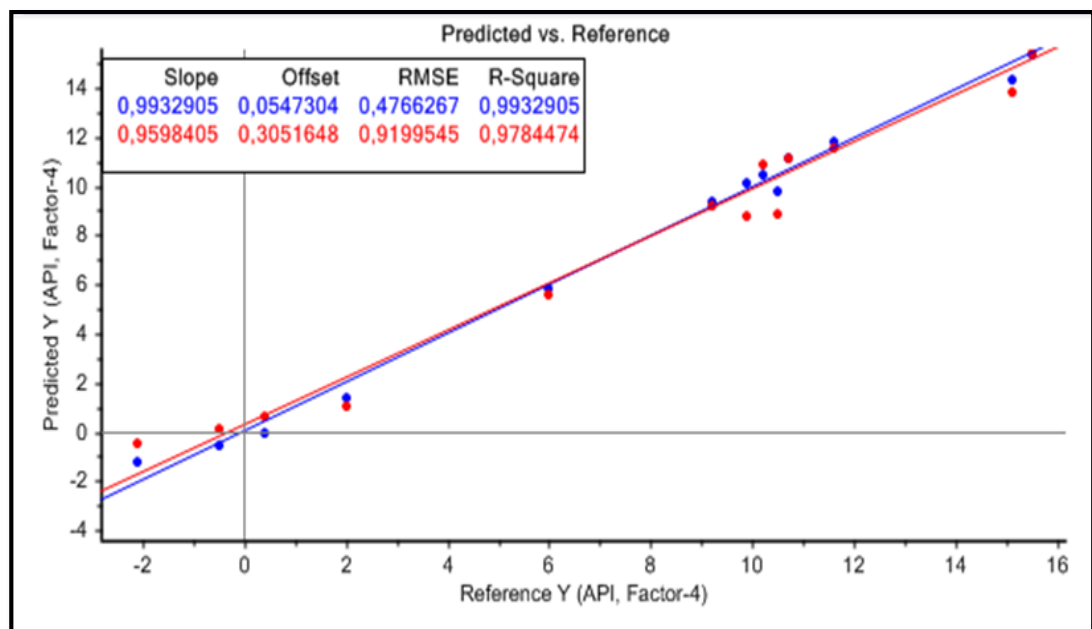
Rango espectral (m/z)	Muestras excluidas	PC	Varianza Explicada (%)	RMSEC(%)	R <sup>2</sup>
497,981 – 762,530	6	4	99,329	0,477	0,99

**3.3.2 Validación cruzada del modelo PLS-R.** Se utilizó la validación cruzada o *cross-validation* para obtener predicciones independientes utilizando todos los espectros disponibles en el conjunto de muestras de la calibración.<sup>[54-55]</sup>

En este tipo de validación se obtiene un error similar al RMSEP, que se denomina raíz cuadrada del error medio de validación cruzada (*Root-Mean-Square Error of Cross-Validation RMSECV*) descrito por la ecuación (22).

El coeficiente de correlación ( $R^2$ ) calculado entre los valores de referencia y de predicción fue mayor a 0,97 y el error estándar de calibración para el conjunto de muestras de validación (SECV) fue 0,945 (figura 28).

**Figura 28.** Curva de Validación de valores de referencia contra valores de predicción



La tabla 10 muestra la aproximación satisfactoria entre los valores de referencia determinados por la técnica de Espectrometría de Masas LDI-TOF/TOF, con los de predicción por el modelo planteado; sin embargo, tres muestras (R2R, R6R y VRR) de las 14 utilizadas en la etapa de validación, presentan error relativo mayor al 30% en la determinación de la gravedad API.

**Tabla 10.** Validación cruzada del modelo de predicción de la gravedad API

Muestra	Referencia (°API)	Predicho (°API)	Residual	Error Relativo
C2R	10,2	10,49	-0,29	2,84
C4R	10,7	11,17	-0,46	4,35
C6R	10,5	9,82	0,68	6,52
R2R	0,4	-0,04	0,44	110,02
R4R	-0,5	-0,50	0	0,78
R6R	-2,1	-1,18	-0,92	43,78
VRR	2	1,39	0,61	30,63
R4T	9,2	9,41	-0,21	2,33
VRT	11,6	11,82	-0,22	1,87
C2G	15,5	15,32	0,18	1,19
C4G	15,7	16,19	-0,49	3,12
C6G	15,1	14,36	0,77	4,93
R4G	6	5,85	0,15	2,53
VRG	9,9	10,13	-0,23	2,29

Las muestras que presentan estos valores elevados de error relativo, se caracterizan por tener los menores valores de gravedad API en el conjunto de las muestras analizadas. Lo que indicaría que este modelo no se ajusta de manera adecuada para muestra pesadas.

Finalmente, la validación cruzada en el modelo PLS-R permitió identificar la calidad del modelo como favorable para la predicción de la gravedad API, ya que, el coeficiente de correlación entre los valores de referencia y de predicción fue superior a 0,99 para la curva de calibración y superior a 0,97 para la curva de validación. Además, los valores de corte de las curvas (offset) presentaron valores con tendencia a cero.

#### 4. CONCLUSIONES

- La espectrometría de masas LDI-TOF/TOF combinada con técnicas de calibración multivariable como PCA y PLS-R, permitió clasificar y analizar fondos de vacío y sus fracciones de destilación molecular (condensados y residuos) para el desarrollo de modelos de predicción de propiedades fisicoquímicas, como alternativa al uso de técnicas comúnmente utilizadas (normas ASTM) y de esta manera, evitar el pre -tratamiento de las muestras y reducir el tiempo de análisis.
- A partir de la información de los espectros de masas LDI-TOF/TOF se determinó el peso molecular promedio de los fondos de vacío y sus fracciones de DM, encontrándose un aumento gradual en el peso para estas muestras, en el siguiente orden: residuos, fondos de vacío y condensados, respectivamente. Además, se identificó una correlación importante entre la gravedad API y peso molecular promedio de las muestras, donde los condensados presentan menor peso molecular promedio y a su vez un mayor °API.
- El análisis por componentes principales (PCA) permitió reducir la dimensionalidad de los datos con el uso de 7 componentes principales, que explicaron aproximadamente el 99% de la variabilidad de los datos. También, se clasificaron las muestras por su gravedad API como pesadas, medianas y livianas, y se determinó la región espectral m/z (497-762) de mayor importancia para el desarrollo del modelo.

- Las gráficas estudiadas en el PCA fueron de gran utilidad, ya que brindaron información detallada para la identificación de muestras atípicas y consideradas como posibles outlier o de gran influencia para el desarrollo y optimización del modelo PLS-R.
  
- Mediante la Regresión PLS se evaluó el potencial de la técnica espectrometría de masas LDI-TOF/TOF para predecir la gravedad API a muestras de fondos de vacío y sus fracciones de la Unidad de Destilación Molecular, con una variabilidad de los datos de referencia mayor al 99% para un mínimo de cuatro componentes principales.
  
- Se obtuvo valores de RMSEC y RMSECV de 0,477 y 0,919, respectivamente y se demostró que la reducción de las variables predictorias (intensidades espectrales), contribuye a la calidad del modelo en términos de disminución de los errores de calibración y en el incremento del coeficiente de Regresión; ya que si se toma todo el espectro, se pueden incluir regiones con baja relación señal/ruido.
  
- El modelo se ajustó por la eliminación de varias muestras consideradas como posibles outliers, que presentaron los más bajos valores de leverage en el PCA y los más elevados en la varianza residual en Y para el mismo PLS.
  
- Por medio de la validación cruzada se evaluó la calidad del modelo como satisfactorio, debido los valores de referencia y de predicción ( $R^2$ ) de 0,99 y 0,97 para la curva de calibración y validación, respectivamente. Sin embargo,

para muestras con valores de °API igual o menores a 2, el modelo implementado no es viable para su predicción.

## 5. RECOMENDACIONES

- ✓ Es importante establecer metodologías para la construcción y validación de los modelos de predicción de propiedades fisicoquímicas de fracciones pesadas de crudos, ampliando la cantidad y rango de calibración de muestras. Además, alimentar el modelo con nuevas muestras, para evaluar su reproducibilidad en el tiempo, y optimizar estas metodologías para controlar la calidad de procesos en la refinería del petróleo.
  
- ✓ Para la posterior aplicación del modelo desarrollado, se deben emplear muestras de la misma naturaleza y espectros adquiridos bajo las mismas condiciones y parámetros instrumentales, descritos en este trabajo.
  
- ✓ Teniendo en cuenta que el modelo desarrollado en este trabajo, es satisfactorio para predecir la gravedad API con espectros de masas LDI-TOF/TOF en muestras de fondos de vacío y sus fracciones de condensados, es recomendable ajustar la metodología con varias y nuevas muestras de la misma naturaleza para evaluar su desempeño.

## CITAS BIBLIOGRAFICAS

- [1] Vargas Jiménez, C. A. Nuevos aportes a la estimación del potencial de hidrocarburos en Colombia. *Rev. Acad. Colomb. Cienc.* **2009**, 33 (126): 17-43.
- [2] Speight, J. G. Petroleum Analysis, En *The Chemistry and Technology of Petroleum*; Taylor and Francis Group, 4th ed.; U.S.A.**2007**; pp 241-243.
- [3] Amorocho, E.; Oliveros G. Situación Colombiana, En *Apuntes sobre Energía y Recursos Energéticos*; Universidad Autónoma de Bucaramanga (UNAB); Bucaramanga, Colombia.**2000**; pp 38-40.
- [4] Calderón Jiménez, N. M. Implementación de un método para la determinación de la masa molecular promedio de las fracciones pesadas del proceso de destilación molecular de fondos de vacío utilizando Espectrometría de Masas MALDI TOF/TOF. Tesis. Universidad Industrial de Santander, **2013**.
- [5] León, A-Y.; Parra, M. J. Determinación del peso molecular del residuo de vacío y sus fracciones SARA. *CT&F - Ciencia, Tecnología y Futuro.* **2010**,4(2): 101-112.
- [6] Sánchez, M. Desmetalización de los Residuos del petróleo. *Revista del Instituto Mexicano del Petróleo.* **1974**. pp 256-280.
- [7] Moreno González, D. H.; Estudio sobre la composición y la estructura de los productos de destilación molecular de fondos de vacío de petróleos colombianos por medio de RMN 1H y 13C. Tesis. Universidad Nacional de Colombia, **2011**.
- [8] Cabrera, E. R., Franco, J. F., Mondragón, F. & Fernández, J.J., Conversión de fondos de vacío de petróleo a semicoque. *Rev. Energética.***2007**, (37):39-50.
- [9] Zuñiga Liñan L. et al. Correlation for predicting the molecular weight of Brazilian petroleum residues and cuts: An application for the simulation of a molecular distillation process. *Journal of Petroleum Science and Engineering.***2011**, 78: 78-85.

- [10] Flórez Rodríguez, Cesar A. Construcción de modelos predictivos de la composición elemental de fondos de vacío de crudos Colombianos a partir de la Espectroscopía de Fluorescencia en estado estacionario y métodos quimiométricos. Tesis. Universidad Industrial de Santander, **2007**.
- [11] Vargas Patiño, G. E. Predicción del análisis SARA de fondos de vacío Colombianos utilizando espectroscopía fotoacústica en el infrarrojo medio y métodos quimiométricos. Tesis. Universidad Industrial de Santander. **2011**.
- [12] Muhammad Imran Ahmad. Molecular components-based representation of petroleum fractions. *Chemical Engineering Research and Design*. **2011**, 89(4): 410-420.
- [13] Speight, J. G. *The chemistry and technology of petroleum*. 4a ed.; Laramie. **2006**; Chemical Industry Series; p 916.
- [14] Himmelblau, D.M. *Principios básicos y cálculos en ingeniería química*; Pearson Educación, 6ª Ed.; México. **2002**; p 24.
- [15] Roldán, J. El Petróleo, *En Fuentes de Energía*; Paraninfo S.A.; Madrid, **2008**; pp 133,134.
- [16] Wauquier, J. P. Evaluación de los crudos de petróleo, En *El Refino del Petróleo*; Fundación Repsol YPF; Madrid. **2004**; pp 317-323.
- [17] El-Aneed, et al. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews*. **2009**, 44(3): 210-230.
- [18] Gross, J. H. Matrix-Assisted Laser Desorption/Ionization, En *Mass Spectrometry*; Springe. **2011**; pp 411-414.
- [19] Palacio Lozano, D. C. Diseño y montaje de un analizador de tiempo de vuelo lineal para el estudio de la ionización multifotónica en moléculas orgánicas simples. Tesis. Universidad Industrial de Santander. **2012**.
- [20] Knochenmuss, R. Ion formation mechanisms in UV-MALDI. *Journal The Royal Society of Chemistr*. **2006**, (131); 966-986.
- [21] Niu, S., Zhang, W. & Chait, B. T. Direct comparison of infrared and ultraviolet wavelength matrix-assisted laser desorption/ ionization mass spectrometry of proteins. *Journal of the American Society for Mass Spectrometry*. **1998**, 9(1): 1-6.

- [22] Plascencia Villa, G. Espectrometría de Masas. Universidad Nacional Autónoma de México, Cuernavaca, Morelos. **2003**. pp 16,17, 22, 23.
- [23] Wilkins, C.L.; Weil, D.A.; Yang, C.L.C.; High Mass Analysis by Laser Desorption Fourier Transform Mass Spectrometry. *Anal.Chem.***1985**,57: 520-524.
- [24] Hillenkamp, F.; Peter-Katalinic, J. *MALDI-MS. A Practical Guide to Instrumentation, Methods and Applications*; Wiley-VCH: Weinheim. **2007**. pp 65-69.
- [25] Cotter, R.J. Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research; *American Chem. Soc.*; Washington, DC. **1997**. pp 89-99
- [26] Siuzdak, G. *The Expanding Role of Mass Spectrometry in Biotechnology*; 2a ed.; MCC Press: San Diego. **2006**. pp 121-136
- [27] Montaudo, G.; Lattimer, R.P. *Mass Spectrometry of Polymers*, CRC Press: Boca Raton. **2001**. pp 110-125.
- [28] Martín, M. C.; Ballesteros, M. Espectrometría de masas y análisis de biomarcadores. *Sección Departamental de Química Analítica*. Universidad Complutense, Madrid, **2004**; pp 113-165.
- [29] Stephens, W.E. A Pulsed Mass Spectrometer With Time Dispersion. *Phys.Rev.* **1946**, 69: 691.
- [30] Vestal, M.L. Modern MALDI Time-of Flight Mass Spectrometry. *J. Mass Spectrom.* **2009**, 44: 303-317.
- [31] Guilhaus, M. Principles and Instrumentation in Time-of-Flight Mass Spectrometry. Physical and Instrumental Concepts. *J. Mass Spectrom.* **1995**, 30, 1519-1532.
- [32] Ioanoviciu, D. Ion-Optical Solutions in Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1995**, 9:985-997.
- [33] Mamyrin, B.A. Laser Assisted Reflectron Time-of-Flight Mass Spectrometry. *Int. J. Mass Spectrom. Ion Proc.* **1994**,131: 1-19.
- [34] Gross, J. H. Time-of-Flight Instruments, En *Mass Spectrometry*; Springer, **2011**; pp 121,122.

- [35] Vestal, M.; Hayden, K. High Performance MALDI-TOF Mass Spectrometry for Proteomics. *Int. J. Mass Spectrom.* **2007**, 268: 83-92.
- [36] Miller J.N.; Miller J. C. Análisis Multivariante, En *Estadística y Quimiometría para Química Analítica*; Pearson Educación S.A., 4a ed.; Madrid, **2002**; pp 221-239.
- [37] Meléndez Correa, L. V. Predicción de propiedades fisicoquímicas de productos destilados del Delayed Coking de fondos de vacío a partir de parámetros estructurales determinados por espectroscopía infrarroja FTIR-ATR y métodos quimiométricos. Tesis. Universidad Industrial de Santander, **2014**.
- [38] Rubinson, K. A.; Rubinson, J. F. Análisis Instrumental, Prentice Hall Hispanoamericana S.A. Madrid; Editorial Pearson Educación, **2001**.
- [39] Martens, H.; Naes, T. Multivariate calibration. New York; Jhon Wiley & Sons, **1989**. p 348.
- [40] Skoog. D. A.; Holler F. J.; Nieman T. A. *Principios de Análisis Instrumental*; 5ª ed.; Madrid, **2001**; pp 103-117.
- [41] Savitsky, S. & Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **1964**, (36): 1627-1639.
- [42] Geladi, P. Notes on the History and Nature of Partial Least Squares (PLS) modeling. *Journal of Chemometrics*, **1988**; (2): 231-245.
- [43] Miller, J. N.; Miller, J. C. *Estadística y Quimiometría para química analítica*. 4ª ed., España; Prentice Hall, **2002**. p 595.
- [44] Ballesteros, S. L. y colaboradores. Determinación de la Composición de Fondos de Vacío de petróleos Colombianos por Espectroscopia de Fluorescencia Inducida por Láser (LIF). *Revista Colombiana de Física-RCF*. **2011**, 3(3): 649-652
- [45] Abbas, O., et al. Assessing petroleum oils biodegradation by chemometric analysis of spectroscopic data, **2008**. (12): 1-15.
- [46] Thomas, E. V. A primer on multivariate calibration. *Anal. Chem.* **1994**, 66: 791-795.

- [47] Pereira T. M. C. et al. FT-ICR MS analysis of asphaltenes: Asphaltenes go in, fullerenes come out. *Fuel*, **2004**, 131: 50-56.
- [48] Kuangnan Qian. Desorption and Ionization of Heavy Petroleum Molecules and Measurement of Molecular Weight Distributions. *Energy & Fuels*, **2007**, 21: 1042-1047
- [49] Gilmore, I. S.; Seah, M. P. Ion detection efficiency in sims: dependencies on energy, mass and composition for microchannel plates used in mass spectrometry. *Int. J. Mass. Spectrom*, **2000**, 202 (3): 217-229
- [50] Orrego Ruíz, J. A. Fraccionamiento por destilación molecular y caracterización de fondos de vacío de crudos Colombianos. Tesis. Universidad Industrial de Santander, **2014**.
- [51] Navarro, L. et. al. Separación y Caracterización de Resinas y Asfáltenos provenientes del crudo Castilla. Evaluación de su interacción Molecular. *Ciencia, Tecnología y Futuro-CT&F*, **2004**, 2 (5): 57-59.
- [52] Unscrambler User Manual, CAMO A/S, Trondheim, **1998**.
- [53] Klaus, H. Altgelt. Mieczyslaw M. Boduszynski. Properties of Heavy Petroleum Fractions, En *Composition and analysis of heavy petroleum fractions*; Marcel Dekker, Inc; New York, **1994**; pp 90-101.
- [54] Wold, S. Cross-validatory estimation of the number of components in factor and principal component models, *Technometrics*. **1978**,20: 397.
- [55] Stone, M. J. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*. **1974**, B36: 111-147.

## BIBLIOGRAFIA

ABBAS, O., REBUFA, C., DUPUY, N., & PERMANYER, A. K. Assessing petroleum oils biodegradation by chemometric analysis of spectroscopic data. En: Talanta. 2008. Vol.75, pp 857-871.

BALLESTEROS, S. L. y colaboradores. Determinación de la Composición de Fondos de Vacío de petróleos Colombianos por Espectroscopia de Fluorescencia Inducida por Láser (LIF). En: Revista Colombiana de Física- RCF. 2011. Vol. 3, No. 3, pp 649-652.

CABRERA, E. R., FRANCO, J. F., MONDRAGÓN, F., & FERNÁNDEZ, J.J. Conversión de fondos de vacío de petróleo a semicoque. En: Rev. Energética. 2007. Vol. 37, pp 39-50.

COTTER, R.J. Time-of-Flight Mass Spectrometry: Instrumentation and Applications in Biological Research. En: American Chem. Soc., Washington, DC. 1997. pp 89-99.

EL-ANEED, et al. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and commonly used mass analyzers. En: Applied Spectroscopy Reviews. 2009. Vol. 3, No. 44, pp 210-230.

FLÓREZ RODRÍGUEZ, Cesar A. Construcción de modelos predictivos de la composición elemental de fondos de vacío de crudos Colombianos a partir de la Espectroscopía de Fluorescencia en estado estacionario y métodos quimiométricos. Tesis. Universidad Industrial de Santander, 2007.

GELADI, P. Notes on the History and Nature of Partial Least Squares (PLS) modeling. En: Journal of Chemometrics. 1988. Vol. 2, pp 231-245.

GILMORE, I. S. & SEAH, M. P. Ion detection efficiency in sims: dependencies on energy, mass and composition for microchannel plates used in mass spectrometry. En: Int. J. Mass. Spectrom. 2000. Vol. 3 No. 202, pp 217-229.

GROSS, J. H. Time-of-Flight Instruments. En: Mass Spectrometry. Springer, 2011. pp 121,122.

GUILHAUS, M. Principles and Instrumentation in Time-of-Flight Mass Spectrometry. Physical and Instrumental Concepts. En: J. Mass Spectrom. 1995. Vol.30, pp 1519-1532.

IOANOVICIU, D. Ion-Optical Solutions in Time-of-Flight Mass Spectrometry. Rapid Commun. En: Mass Spectrom. 1995. Vol. 9, pp 985-997.

KLAUS, H., & ALTGELT, Mieczyslaw. Properties of Heavy Petroleum Fractions, En: Composition and analysis of heavy petroleum fractions. Marcel Dekker. Inc; New York. 1994. pp 90-101.

KNOCHENMUSS, R. Ion formation mechanisms in UV-MALDI. En: Journal The Royal Society of Chemistr. 2006. Vol.131, pp 966-986.

KUANGNAN, Qian. Desorption and Ionization of Heavy Petroleum Molecules and Measurement of Molecular Weight Distributions. En: Energy & Fuels. 2007. Vol. 21, pp 1042-1047.

LEÓN, A-Y., & PARRA, M. J. Determinación del peso molecular del residuo de vacío y sus fracciones SARA. En: CT&F - Ciencia, Tecnología y Futuro. 2010. Vol. 2, No. 4, pp 101-112.

MAMYRIN, B.A. Laser Assisted Reflectron Time-of-Flight Mass Spectrometry. En: Int. J. Mass Spectrom. Ion Proc. 1994. Vol. 131, pp 1-19.

MARTENS, H., & NAES, T. Multivariate calibration. New York; Jhon Wiley & Sons. 1989. p 348.

MARTÍN, M. C., & BALLESTEROS, M. Espectrometría de masas y análisis de biomarcadores. Sección Departamental de Química Analítica. Universidad Complutense, Madrid. 2004. pp 113-165.

MILLER, J. N.; MILLER, J. C. Estadística y Quimiometría para química analítica. 4ª ed., España; Prentice Hall. 2002. p 595.

MORENO GONZÁLEZ, D. H. Estudio sobre la composición y la estructura de los productos de destilación molecular de fondos de vacío de petróleos colombianos por medio de RMN 1H y 13C. Tesis. Universidad Nacional de Colombia, 2011.

MUHAMMAD IMRAN, Ahmad. Molecular components-based representation of petroleum fractions. En: Chemical Engineering Research and Design. 2011. Vol. 4 No. 89, pp 410-420.

NAVARRO, L. et. al. Separación y Caracterización de Resinas y Asfáltenos provenientes del crudo Castilla. Evaluación de su interacción Molecular. En: Ciencia, Tecnología y Futuro-CT&F. 2004. Vol. 5, No. 2, pp 57-59.

NIU, S., ZHANG, W. & CHAIT, B. T. Direct comparison of infrared and ultraviolet wavelength matrix-assisted laser desorption/ ionization mass spectrometry of proteins. En: Journal of the American Society for Mass Spectrometry. 1998. Vol. 1 No. 9, pp 1-6.

ORREGO RUÍZ, J. A. Fraccionamiento por destilación molecular y caracterización de fondos de vacío de crudos Colombianos. Tesis. Universidad Industrial de Santander, 2014.

PEREIRA T. M. C. et al. FT-ICR MS analysis of asphaltenes: Asphaltenes go in, fullerenes come out. En: Fuel. 2004 Vol. 131, pp 50-56.

ROLDÁN, J. El Petróleo, En: Fuentes de Energía; Paraninfo S.A.; Madrid. 2008. pp 133,134.

RUBINSON, K. A.; Rubinson, J. F. Análisis Instrumental, Prentice Hall Hispanoamericana S.A. Madrid; Editorial Pearson Educación. 2001.p 457

SÁNCHEZ, M. Desmetalización de los Residuos del petróleo. En: Revista del Instituto Mexicano del Petróleo. 1974. pp 256-280.

SAVITSKY, S. & GOLAY. Smoothing and differentiation of data by simplified least squares procedures. En: Analytical Chemistry. 1964. Vol. 36, pp 1627-1639.

SPEIGHT, J. G. The chemistry and technology of petroleum. 4a ed.; Laramie.; Chemical Industry Series. 2006. p 916.

SPEIGHT, J. G. Petroleum Analysis, En: The Chemistry and Technology of Petroleum. Taylor and Francis Group, 4th ed.; U.S.A. 2007. pp 241-243.

STONE, M. J. Cross-validatory choice and assessment of statistical predictions. En: Journal of the Royal Statistical Society. 1974 Vol. B36, pp 111-147.

VARGAS JIMÉNEZ, C. A. Nuevos aportes a la estimación del potencial de hidrocarburos en Colombia. En: Rev. Acad. Colomb. Cienc. 2009. Vol. 126, No. 33, pp 17-43.

VESTAL, M.; HAYDEN, K. High Performance MALDI-TOF Mass Spectrometry for Proteomics. En: Int. J. Mass Spectrom. 2007. Vol. 268, pp 83-92.

WAUQUIER, J. P. Evaluación de los crudos de petróleo. En: El Refino del Petróleo; Fundación Repsol YPF. Madrid. 2004. pp 317-323.

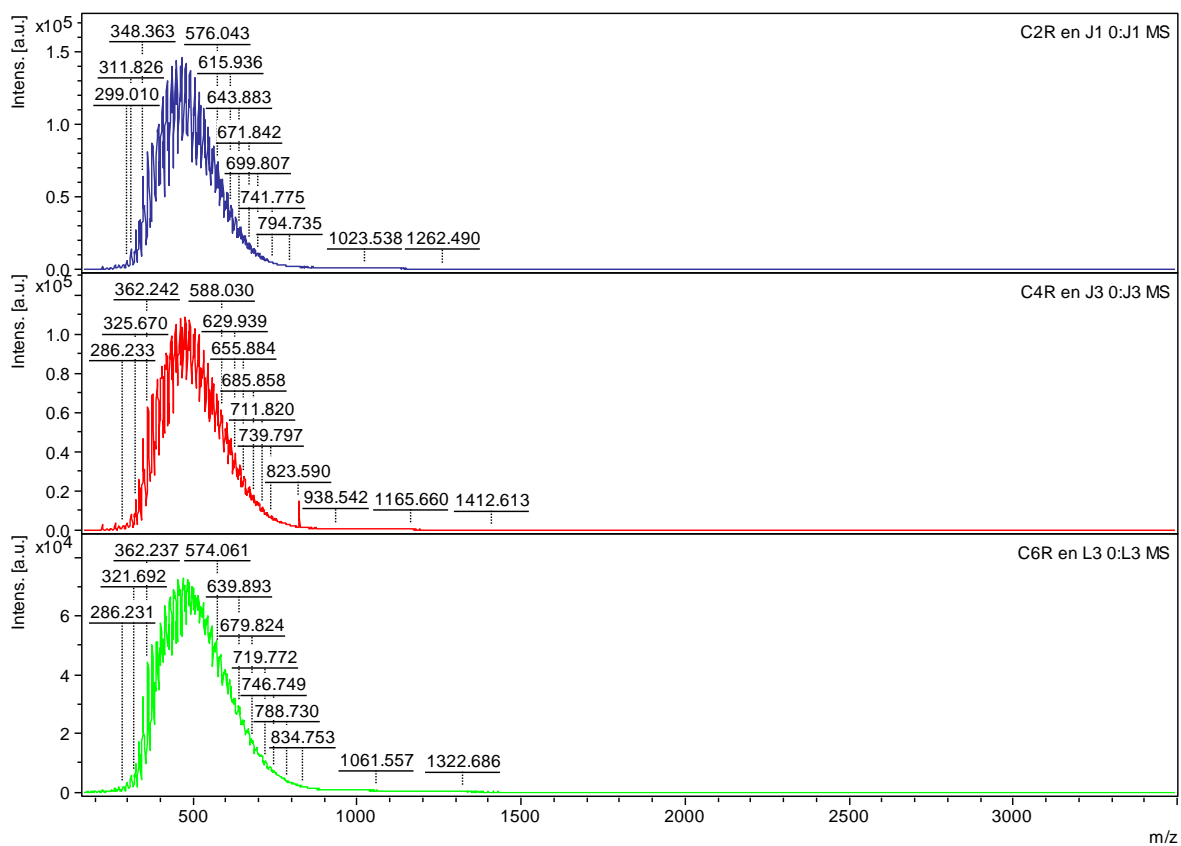
WILKINS, C.L.,WEIL, D.A., &YANG, C.L.C.; High Mass Analysis by Laser Desorption Fourier Transform Mass Spectrometry.En: Anal.Chem.1985.Vol. 57, pp 520-524.

WOLD, S. Cross-validatory estimation of the number of components in factor and principal component models. En: Technometrics. 1978. Vol. 20, p 397.

ZUÑIGA LIÑAN, L. et al. Correlation for predicting the molecular weight of Brazilian petroleum residues and cuts: An application for the simulation of a molecular distillation process. En: Journal of Petroleum Science and Engineering. 2011. Vol. 78, pp 78-85.

# ANEXOS

## ANEXO A. Espectros de masas de la fracción condensados del fondo de vacío R



**ANEXO B.** Tabla de datos de repetibilidad espectral para la fracción de condensados del fondo de vacío R

Repetibilidad	C2R		C4R		C6R	
	$\bar{M}_n$	$\bar{M}_w$	$\bar{M}_n$	$\bar{M}_w$	$\bar{M}_n$	$\bar{M}_w$
<b>1</b>	515,549	541,591	526,803	551,701	541,450	569,072
<b>2</b>	518,007	545,327	529,767	555,071	549,595	578,140
<b>3</b>	518,220	545,490	530,506	555,894	544,680	573,976
<b>4</b>	516,192	542,231	529,445	554,602	544,685	572,938
<b>5</b>	515,290	540,991	531,251	556,489	545,874	574,630
<b>6</b>	515,689	541,659	528,710	553,770	544,380	572,251
<b>Media</b>	516,491	542,881	529,414	554,588	545,111	573,501
<b>Desviación Estándar</b>	1,293	1,997	1,550	1,706	2,645	2,984
<b>Error Estándar</b>	0,528	0,815	0,633	0,697	1,080	1,218
<b>Desviación Estándar Relativa</b>	0,250	0,368	0,293	0,308	0,485	0,520