

**Modelo quimiométrico para la curva de destilación del crudo mezcla vasconia de la
Gerencia Refinería Barrancabermeja (GRB)**

Rosangela Pacheco Acosta

**Trabajo de Aplicación presentado para optar por el título de Master en Ingeniería Química
Modalidad Profundización**

Director

Giovanni Morales Medina

Dr. En Ingeniería Química

Universidad Industrial de Santander

Facultad de Ingenierías Fisicoquímicas

Escuela de Ingeniería Química

Maestría en Ingeniería Química

Bucaramanga

2019

Contenido

	Pág.
Introducción	13
1. Objetivos	15
1.1 Objetivo General	15
1.2 Objetivos Específicos.....	15
2. Marco Teórico.....	16
2.1 Naturaleza química de los crudos	16
2.1.1 Clasificación de los crudos	17
2.2 Curvas de destilación de crudos.....	18
2.2.1 Destilación TBP (<i>True Boiling Point</i>)	18
2.2.2 Destilación simulada o SimDis	19
2.3 Crudos recibidos en la GRB	20
2.4 Espectroscopia del infrarrojo cercano (NIR)	24
2.5 Análisis quimiométricos	26
2.6 Estado del arte.....	27
3. Metodología	30
3.1 Recolección de muestras crudo vasconia.....	30
3.2 Medición de espectros FT-NIR y de propiedades.....	30
3.3 Análisis y pretratamiento de espectros NIR.....	31

3.4 Relaciones entre las propiedades y la composición	32
3.5 Análisis por componentes principales	32
3.6 Desarrollo de regresiones PLS.....	33
4. Análisis de resultados	34
4.1 Espectros FT-NIR mezclas Vasconia	34
4.2 Propiedades de las mezclas Vasconia	36
4.3 Análisis PCA.....	46
4.4 Regresión lineal múltiple	50
4.5 Regresiones PLS	51
4.5.1 Temperaturas de evaporados.....	51
4.5.2 Rendimientos acumulados	61
4.5.3 Composiciones aromática, nafténica y parafínica	64
5. Conclusiones.....	75
6. Recomendaciones	77
Referencias Bibliográficas	78
Apéndices.....	82

Lista de Figuras

	Pág.
<i>Figura 1.</i> Distribución de oleoductos, poliductos y gasoductos en Colombia. Adaptado de: ECOPETROL S.A.	20
<i>Figura 2.</i> Porcentaje en volumen de crudos cargados diariamente en GRB	23
<i>Figura 3.</i> Tendencia de gravedad API en crudo mezcla Vasconia durante el año 2017 y enero del 2018.....	23
<i>Figura 4.</i> Tendencia de acidez en el crudo mezcla Vasconia durante el año 2017 y enero del 2018.....	24
<i>Figura 5.</i> Diagrama del espectro electromagnético. Adaptado de: Atlas of Science (2016).....	25
<i>Figura 6.</i> Espectros FT-NIR medidos de las muestras de crudo mezcla Vasconia.....	36
<i>Figura 7.</i> Espectros FT-NIR con corrección por línea base.	36
<i>Figura 8.</i> Pretratamientos efectuados a los espectros FT-NIR (rango 6500–3500 cm ⁻¹) muestras de crudo mezcla Vasconia. Iz: primera derivada, der: segunda derivada.	37
<i>Figura 9.</i> Curvas SimDis medidas a las muestras de crudo mezcla Vasconia.	38
<i>Figura 10.</i> Diagramas <i>boxplot</i> temperaturas de evaporación muestras de crudo mezcla Vasconia. IBP y FBP hacen referencia al 0% y al 100% de evaporados, respectivamente	39
<i>Figura 11.</i> Matriz de dispersión propiedades crudo mezcla Vasconia.	39
<i>Figura 12.</i> Figura de dispersión para K_w vs VGC de las muestras de crudo mezcla Vasconia.	41

<i>Figura 13.</i> Gráficos de dispersión para K_w vs VGC de las muestras de crudo mezcla Vasconia. Agrupación 0 en color fucsia, agrupación 1 en color verde claro y agrupación 2 en azul.	43
<i>Figura 14.</i> Gráficos <i>boxplot</i> de las fracciones de las mezclas Vasconia. Fuente: Autora.	44
<i>Figura 15.</i> Gráficos de los <i>scores</i> y los <i>loading</i> resultantes del PCA aplicado a la matriz de espectros pretratados de las mezclas Vasconia. Fuente: Autora.	48
<i>Figura 16.</i> Identifiación de las agrupaciones basadas en los espectros FT-NIR obtenidas con $kmean=3$ sobre el plano de los <i>scores</i>	49
<i>Figura 17.</i> Identifiación de las agrupaciones obtenidas con $kmean=3$ sobre el espacio de las propiedades. Fuente: Autora.	50
<i>Figura 18.</i> Gráficos de paridad de los resultados de las regresiones RLM en las diferentes temperaturas de evaporado.	51
<i>Figura 19.</i> Gráficos de paridad de los resultados de las regresiones PLS en diferentes temperaturas de evaporado.	56
<i>Figura 20.</i> Residuales obtenidos de las regresiones PLS en diferentes temperaturas de evaporado.	57
<i>Figura 21.</i> Coeficientes de regresión PLS en el ajuste de las diferentes temperaturas de evaporado.	59
<i>Figura 22.</i> Gráficas de paridad regresión PLS de los rendimientos acumulados.	67
<i>Figura 23.</i> Residuales regresión PLS de los rendimientos acumulados.	69
<i>Figura 24.</i> Contribuciones números de onda regresión PLS de los rendimientos acumulados ...	70
<i>Figura 25.</i> Gráficos de paridad regresión PLS en el ajuste de las fracciones volumétricas de naftenos, aromáticos y parafinas.	71

<i>Figura 26.</i> Residuales regresión PLS en el ajuste de las fracciones volumétricas de naftenos, aromáticos y parafinas.	72
<i>Figura 27.</i> Coeficientes regresión PLS en el ajuste de las fracciones volumétricas de naftenos, aromáticos y parafinas. Fuente: Autora.	73

Lista de Tablas

	Pág.
Tabla 1. <i>Crudos que se procesan diariamente en la GRB. Fuente: Ecopetrol.....</i>	21
Tabla 2. <i>Crudos que componen la Mezcla Vasconia. Fuente: Ecopetrol.....</i>	22
Tabla 3. <i>Frecuencias de absorción de algunos grupos funcionales en hidrocarburos.</i>	25
Tabla 4. <i>Estadísticos univaridos para las propiedades de las mezclas Vasconia.</i>	40
Tabla 5. <i>Rangos de propiedades según la componente aportante en las mezclas Vasconia.....</i>	44
Tabla 6. <i>Modelos de regresión PLS finales para las temperaturas de evaporado de las muestras de la mezcla Vasconia.</i>	59
Tabla 7. <i>Modelos de regresión PLS finales para los rendimientos acumulados volumétricos de la mezcla Vasconia a diferentes temperaturas.....</i>	63
Tabla 8. <i>Modelos de regresión PLS finales para las composiciones de las muestras de la mezcla Vasconia.....</i>	73

Lista de Apéndices

	Pág.
Apéndice A. Curvas SimDis mezcla Vasconia. Resultados, regresiones y residuales	82

Resumen

Título: Modelo quimiométrico para la curva de destilación del crudo mezcla vasconia de la Gerencia Refinería Barrancabermeja (GRB)*

Autora: Rosangela Pacheco Acosta**

Palabras claves: Crudos, destilación simulada, SimDis, PLS, quimiometría.

La GRB recibe diferentes tipos de crudos mezclados, provenientes de un amplio número de pozos y estaciones; la mezcla denominada Vasconia corresponde *ca.* 45% de la carga actual. Debido a situaciones operacionales, las proporciones de la mezcla fluctúan continuamente, con lo cual, las condiciones operacionales de los procesos deben ser ajustadas –según la curva de puntos de ebullición SimDis, ASTM D7169– para mantener los rendimientos en los productos de refinación y un margen de ganancia favorable. Sin embargo, la medición de la curva SimDis puede alrededor de 1 día, incluyendo los procesos de gestión interna en el laboratorio. Por otro lado, los espectros FT-NIR son medidos en cuestión de minutos y proporcionan información composicional que puede ser utilizada en el desarrollo de modelos quimiométricos para la SimDis. En el presente documento se exponen los principales resultados obtenidos del desarrollo de modelos de regresión PLS para la curva SimDis, basados en el espectro NIR. Según los resultados, el análisis PCA conduce a la identificación de los números de onda 3849 y 3843 cm^{-1} como indicadores de las fluctuaciones en composición de la mezcla Vasconia; asimismo, los números de onda 4335 y 4324 cm^{-1} caracterizan las temperaturas de la curva SimDis. Los modelos de regresión PLS se encuentran definidos por expresiones simples en términos de las absorbancias pretratadas en los números de onda 4335 y 4324 cm^{-1} . Los modelos para las temperaturas reportan valores de RMSE% de validación entre 4,1 y 8,29%, mientras que los modelos para los rendimientos acumulados reportan valores de RMSE% entre 4,89 y 20,58%. Por su parte, las regresiones PLS para las fracciones de aromáticos, naftenos y parafinas presentan valores de RMSE% de validación inferiores a 2,2%. Los anteriores desempeños de las regresiones obtenidas en el presente trabajo son comparables con ajustes reportados en la literatura abierta.

* Proyecto de grado

** Facultad de Ingenierías Fisicoquímicas, Escuela de Ingeniería Química, Maestría en Ingeniería Química. Director Giovanni Morales Medina

Abstract

Title: Chemometric model for distillation curve from vasconia crude oil of Gerencia Refinería Barrancabermeja (GRB) *

Author: Rosangela Pacheco Acosta **

Keywords: Crude oils, simulated distillation, SimDis, PLS, chemometrics.

GRB processes different class of blending crude oils from several petroleum wells and ports; blending crude called Vasconia corresponds to *ca.* 45% of the current load. Due to operation, composition of blending crudes change continuously and therefore, operational variables of the units have to be adjusted, accordingly to SimDis curve (ASTM D7169), in order to set product yield properly and to maintain net margin. However, experimental measurement of SimDis can last for one day, mainly due to sample queuing in the laboratory. Conversely, FT-NIR spectra are measured rapidly, giving information about composition that can be inputs for developing chemometric models for SimDis. Herein, principal results about regression models for SimDis based on NIR are exposed. According to the results, application of PCA leads to identification of wavenumbers 3849 y 3843 cm^{-1} as indicators for the composition of Vasconia blending; likewise, wavenumbers 3849 y 3843 cm^{-1} indicates fluctuations in SimDis temperatures. PLS regression models are defined in simple terms of the wavenumbers 4335 y 4324 cm^{-1} . Regressions for temperatures and cumulated yields reported RMSE% in the intervals 4,1– 8,29% and 4,89–20,58%, respectively. In addition, regressions for aromatic, naphthene and paraffin fractions yielded RMSE% values lower than 2,2%. The performances for the regressions developed in this work are comparable with corresponding models from literature.

* Proyecto de grado

** Facultad de Ingenierías Fisicoquímicas, Escuela de Ingeniería Química, Maestría en Ingeniería Química. Director Giovanni Morales Medina

Introducción

La Gerencia de la Refinería de Barrancabermeja (GRB) recibe diferentes tipos de crudos, puros y mezclados, provenientes de un amplio número de pozos y estaciones (ca. 122 corrientes de crudo). En el caso de las mezclas, debido a situaciones operacionales internas y externas al negocio (e.g. incremento en crudos pesados, mantenimiento de equipos y atentados en oleoductos), sus proporciones y por tanto sus propiedades fluctúan continuamente, con lo cual, las condiciones operacionales de los procesos deben ser ajustadas para mantener los rendimientos en los productos de refinación y un margen de ganancia favorable. Por lo anterior, una predicción con baja incertidumbre para las propiedades de los crudos y sus mezclas constituye una meta para la GRB; esta meta se encuentra en el Marco Estratégico 2020 de ECOPETROL S.A., consignada como “...establecer una dieta de crudos con características acertadas que conlleven al cubrimiento de las necesidades internas del país y los requerimientos en sus exportaciones”.

Una de las propiedades importantes de los crudos y sus mezclas corresponde a la curva de destilación o curva de puntos de ebullición denominada como TBP (*true boiling point*) según la norma ASTM D2892¹. La obtención de esta curva a nivel de laboratorio puede tardar varios días, por lo cual la norma ASTM D7169² describe un procedimiento que permite la predicción de la curva de ebullición a partir de la dilución de los diferentes componentes en un acolumna cromatográfica; esta curva es denominada como SimDis y su determinación tarda alrededor de 1

¹ ASTM D2892. (2017). Standard Test Method for Distillation of Crude Petroleum (15-Theoretical Plate Column).

² ASTM D7169. (2016). Standard Test Method for Boiling Point Distribution of Samples with Residues Such as Crude Oils and Atmospheric and Vacuum Residues by High Temperature Gas Chromatography.

día, que incluye la preparación del equipo, la ejecución de la prueba, la limpieza del equipo y los procesos de gestión interna en el laboratorio. Esta curva TBP o SimDis es utilizada continuamente por los ingenieros de refinación para determinar rendimientos de productos y sustentar decisiones operacionales. A pesar de la importancia de esta propiedad, los costos asociados restringen la determinación experimental de la TBP para la totalidad de los crudos recibidos en la GRB. Con esto, los ingenieros en diversas situaciones cuentan sólo con medidas de calidad globales de los crudos (*i.e.* °API, acidez, azufre, sal y BSW), las cuales son insuficientes para la predicción adecuada de los desempeños operacionales.

Como alternativa para obtener una curva TBP o SimDis representativa de una mezcla de crudos con proporciones desconocidas, se cuenta con modelos de predicción quimiométricos basados en espectros infrarrojos (NIR) tomados en operación. A partir de estos espectros se han propuesto modelos de regresión para la densidad API, el contenido de azufre, la acidez y el análisis SARA (saturados, aromáticos, resinas, asfaltenos), entre otros. Debido a esta aplicación, la GRB invirtió en tecnología para la medición en línea de espectros infrarrojos; estos sensores cuentan con conexión a un ordenador para la codificación de modelos de predicción. La puesta en marcha de este sistema, actualmente en desuso, con un modelo de predicción quimiométrico permitirá apoyar las decisiones operativas acordes con la variabilidad de las cargas.

El presente documento expone los principales resultados obtenidos de la calibración y validación de modelos quimiométricos PLS basados en espectros NIR para la predicción de la curva SimDis de muestras del crudo mezcla Vasconia. Esta mezcla reporta la mayor variabilidad en sus especificaciones y adicionalmente constituye la mayor proporción de carga a los procesos de destilación (*ca.* 45% de la carga o dieta actual) a la GRB.

1. Objetivos

1.1 Objetivo General

Validar modelos PLS de regresión de la curva de destilación SimDis del Crudo Mezcla Vasconia, recibido en la Gerencia Refinería de Barrancabermeja (GRB), a partir de los correspondientes espectros FT-NIR.

1.2 Objetivos Específicos

- Construir y depurar una base de datos de espectros FT-NIR, curvas de destilación SimDis y propiedades globales representativas del crudo Mezcla Vasconia.
- Relacionar las tendencias en las propiedades, las curvas de destilación SimDis y los espectros FT-NIR, por medio de estimaciones de las composiciones del crudo mezcla Vasconia y el análisis por componentes principales (PCA).
- Calibrar modelos de mínimos cuadrados parciales (PLS) dirigidos a la predicción de la curva de destilación SimDis del crudo Mezcla Vasconia, considerando los espectros FT-NIR.

2. Marco Teórico

2.1 Naturaleza química de los crudos

El petróleo crudo no es un elemento o un compuesto uniforme; se trata de una mezcla compleja que posee una cantidad significativa de componentes químicos llamados hidrocarburos. Los hidrocarburos son los compuestos orgánicos más importantes y reciben este nombre debido a que el carbono y el hidrógeno son sus constituyentes mayoritarios, representando porcentajes aproximados del 85% y el 12%, respectivamente. De igual manera, los crudos contienen proporciones menores de azufre (0.05 – 6.0%), nitrógeno (0.1 – 2.0%), oxígeno (0.05 – 1.5%) y metales como Ni, V, Na, Ca y Cu (0.001 – 0.1%) (Wauquier, 1995). Por ser una mezcla compleja de hidrocarburos, el petróleo crudo presenta una distribución muy grande de pesos moleculares y por lo tanto, tiene un amplio rango de temperaturas de ebullición. Este fenómeno físico es el más usado en la industria del petróleo para llevar a cabo el proceso de destilación (Navarro, 2005).

El petróleo crudo está formado por tres grandes familias de hidrocarburos, identificados como parafínicos, naftenos y aromáticos, cuya distribución en el crudo define su clasificación, sus propiedades fisicoquímicas, su calidad, los rendimientos de cada uno de los productos, la calidad de éstos productos y algunas de las condiciones de operación de los procesos de refinación. Adicionalmente, la distribución porcentual de estos hidrocarburos en el crudo, define su precio y las condiciones de transporte por oleoductos (UOP, 1978).

2.1.1 Clasificación de los crudos Existen diferentes métodos para clasificar los crudos, cada uno de ellos proporciona información valiosa, para su producción, comercialización, transporte y refinación, por lo que se requiere correlacionar éstos métodos para obtener el mayor provecho y conocimiento del crudo. Los crudos se pueden clasificar según (Riazi, 2005),

- **°API:** crudos livianos (>30 °API), crudos intermedios (20-30 °API), crudos pesados (10-20 °API) y crudos extrapesados (<10 °API).
- **Factor de Caracterización K_w o UOP:** los crudos parafínicos tienen un factor $K_w > 12.2$, los crudos intermedios entre 11.5-12.2 y los pesados donde predominan las estructuras nafténicas y aromáticos con $K_w < 11.5$.
- **Contenido de Azufre:** crudos dulces ($\%p/p \text{ S} < 0.5$), crudos medios ($0.5 \leq \%p/p \text{ S} \leq 1.5$) y crudos agrios ($\%p/p \text{ S} > 1.5$).

Estas propiedades, proporcionan información del tipo de hidrocarburo predominante en un crudo y/o en sus fracciones (Wauquier, 1995; Navarro, 2005). El análisis de los crudos y sus fracciones por tipo de componentes, especialmente las fracciones pesadas, se centra en cuatro grandes familias (SARA):

- **Saturados:** corresponde a la suma de los hidrocarburos parafínicos y cicloparafínicos.
- **Aromáticos:** son todos aquellos hidrocarburos que en su estructura tiene por lo menos un anillo bencénico.
- **Resinas:** son hidrocarburos de naturaleza aromática, cuya estructura tiene entre 4-8 anillos bencénicos, con la presencia de heteroátomos, especialmente N, S, O y pequeñas cantidades de metales.
- **Asfaltenos:** son las estructuras y los hidrocarburos más complejos y de mayor peso molecular presentes en un crudo.

2.2 Curvas de destilación de crudos

Existen dos técnicas utilizadas en el laboratorio de calidad de GRB para determinar la destilación de los crudos recibidos, estas son la curva TBP y la destilación simulada (SimDis).

2.2.1 Destilación TBP (*True Boiling Point*) La destilación TBP permite dar una imagen casi exacta de la composición de un crudo mediante la medida de la temperatura de ebullición de los constituyentes que lo componen, a través de la norma técnica ASTM D2892, en un tiempo aproximado de 6 días incluyendo la limpieza del equipo. Este método permite evaluar el petróleo crudo para predecir los rendimientos de las fracciones de diversos rangos de ebullición y, por lo tanto, es valioso en las discusiones técnicas de naturaleza comercial, estrategias operacionales y de optimización del proceso. De igual forma, las fracciones producidas pueden ser analizadas como productos o combinarse para llevar a cabo estudios analíticos, ingeniería y evaluaciones de calidad del producto; para llevar a cabo dicho análisis se requieren aproximadamente 6 días adicionales (Riazi, 2005).

Este método de prueba cubre el procedimiento para la destilación de petróleo crudo estabilizado a una temperatura de corte final de 400°C de Temperatura Atmosférica Equivalente (AET) empleando una columna de fraccionamiento que tiene una eficacia de 14 a 18 placas teóricas operadas a una relación de reflujo de 5:1. La destilación TBP detalla los procedimientos para la producción de gas licuado, fracciones de destilado y residuos de calidad estandarizada sobre los cuales se pueden obtener datos analíticos y determinar los rendimientos de las fracciones anteriores por masa y volumen. A partir de la información obtenida, se puede producir un gráfico de la temperatura vs % de masa o volumen destilado acumulado.

2.2.2 Destilación simulada o SimDis La destilación simulada (SimDis) establece la distribución del punto de ebullición, los intervalos del punto de corte de los crudos y los residuos mediante cromatografía de gases a alta temperatura, a través de la norma técnica ASTM D7169. Esta prueba se usa para determinar la distribución del punto de ebullición a una temperatura de 720 °C y demora aproximadamente 6 horas, incluyendo la calibración del equipo. La cantidad de residuo o recuperación de muestra se evalúa por medio de análisis matemáticos y, a diferencia del método TBP, los productos no son caracterizados (Riazi, 2005).

El principio operativo en el que se basa la destilación simulada utiliza columnas capilares con películas delgadas, lo que da como resultado la separación incompleta de C4-C8 en presencia de grandes cantidades de disulfuro de carbono, produciendo una distribución de punto de ebullición poco confiable correspondiente a este intervalo de elución. Por tal motivo es necesario un análisis por cromatografía de gases (GC) de mayor resolución para obtener una descripción más precisa de la curva del punto de ebullición.

La determinación de la distribución del punto de ebullición de los crudos y los residuos de vacío, así como otras fracciones de petróleo, proporciona información importante para el funcionamiento de la refinería, como el rendimiento porcentual en masa de los productos. Este método de prueba puede proporcionar información útil que puede ayudar a establecer condiciones operacionales en la refinería. El conocimiento de la cantidad de residuos producidos es importante para determinar la economía del proceso de refinación.

2.3 Crudos recibidos en la GRB

Los crudos colombianos son movidos por medio de oleoductos desde diferentes puntos de recolección hasta las refinerías de Barrancabermeja y Cartagena. En la Figura 1.1 se ilustra la red de oleoductos y poliductos actual en el territorio nacional. Según esta figura, una gran proporción de los oleoductos que transportan los crudos pasan por la estación de bombeo Vasconia; en esta estación se almacenan y mezclan los crudos para luego continuar su recorrido a la GRB.



Figura 1. Distribución de oleoductos, poliductos y gasoductos en Colombia. Adaptado de: ECOPETROL S.A.

Entre la variedad del crudo que ingresa a la GRB se encuentran los presentados en la Tabla 1.1. Según los datos de esta tabla, la mezcla Vasconia constituye *ca.* 45% del total de la carga a

la GRB (Figura 1.2); inclusive, este porcentaje de aporte de la mezcla Vasconia puede aumentar cuando las situaciones de orden público afectan la disponibilidad de otros crudos. La mezcla Vasconia está compuesta por diferentes crudos nacionales, cuyos porcentajes en mezcla cambian diariamente según la producción de los respectivos campos.

Tabla 1.

Crudos que se procesan diariamente en la GRB. Fuente: Ecopetrol.

CRUDO	RECIBO	DISPONIBLE PARA CARGA
	KBPD	KBPD
M. VASCONIA	94.28	94.28
ELC LCT	40.21	33.21
VELÁSQUEZ	29.43	15.43
CAÑO LIMON	20.32	20.32
CASABE	16.13	16.13
CUSIANA	15.05	15.05
ISLA 6	14.97	0.00
CASTILLA	8.00	0.00
PROVINCIA	6.29	6.29
GALÁN	3.58	3.58
ELC HCT	2.60	2.60
PAYOA	2.28	2.28
TOTAL	253.13	209.15

La mezcla Vasconia tiene tres grandes componentes o mezclas intermedias: el crudo que llega por el Oleoducto de Colombia (OCL), el que llega por el Oleoducto del Alto Magdalena (OAM) y el recibido desde los campos Ocelote y Palagua, denominado como Vasconia 1 (VS1). La Tabla 1.2 muestra los crudos iniciales constituyentes de las mezclas intermedias.

Debido a la cantidad de crudos que constituyen la mezcla Vasconia y al cambio continuo en su composición, es posible observar la variabilidad diaria de sus propiedades básicas. La Figura 1.3 muestra como la gravedad API presenta valores tan bajos como 19.7 y tan altos como 32.8 °API, así como una gran dispersión alrededor del valor promedio 23.8°API. Por otra parte, la

Figura 1.4 ilustra el parámetro de acidez, el cual presenta un valor máximo de 2.63 mgKOH/g, un valor mínimo de 0.09 mgKOH/g y una alta dispersión alrededor de 0.66 mgKOH/g.

La operación de las unidades de destilación de crudos se ve afectada por la variación constante en las características del crudo procesado. Para el caso de la GRB, como se mencionó anteriormente, el crudo mezcla Vasconia es el de mayor porcentaje en la dieta de crudos y la variación de esta mezcla implica problemas operacionales en las unidades, debido a que los equipos no pueden operar de una manera estable ni obtener rendimientos máximos. Como se mostró en el estado del arte, una operación basada en predicciones en línea evita problemas y mejora la producción, razón por la cual la implementación de un sistema que permita evaluar los crudos teniendo en cuenta sus rendimientos es de suma importancia.

Tabla 2.

Crudos que componen la Mezcla Vasconia. Fuente: Ecopetrol.

OCL	Apiay
	Barquereña
	Estación Jaguar
	Ocelote
	Rubiales
	San Francisco
	Santiago
OAM	Andalucía
	Balcón
	Dina Cretaceos
	Dina
	Dina Terciarios 04
	Estación Gavan
	Galán
	Guando
	Matachin
	Rio Ceibas
	Rubiales
	South Blend K-002
	SanFrancisco
	Santa Clara
	Santiago
	Tello
	Toqui-Toqui
VS1	Yaguara
	Ocelote
	Palagua

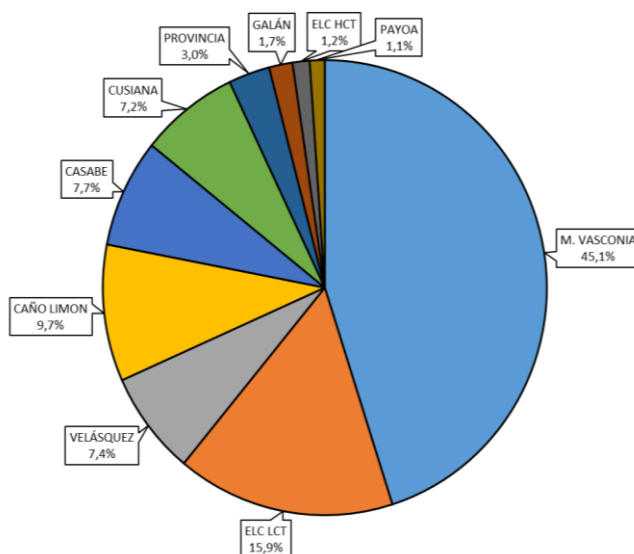


Figura 2. Porcentaje en volumen de crudos cargados diariamente en GRB

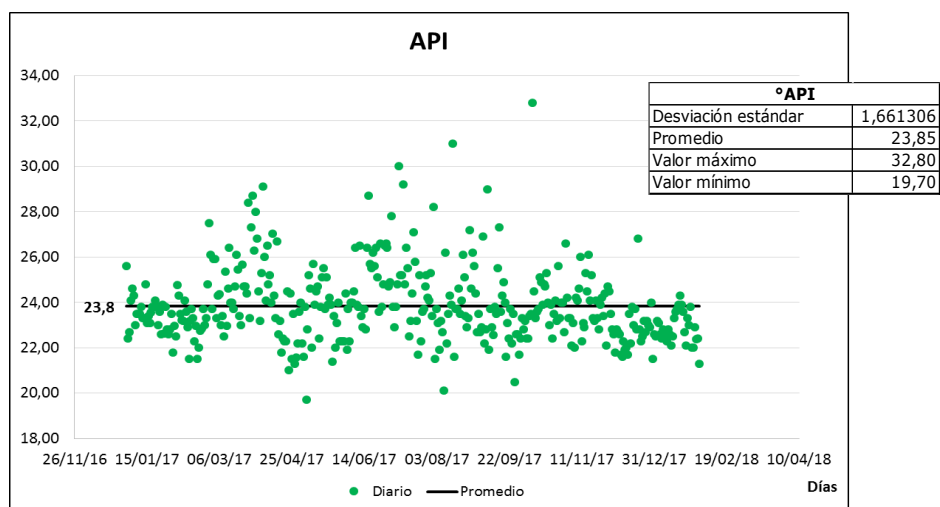


Figura 3. Tendencia de gravedad API en crudo mezcla Vasconia durante el año 2017 y enero del 2018.

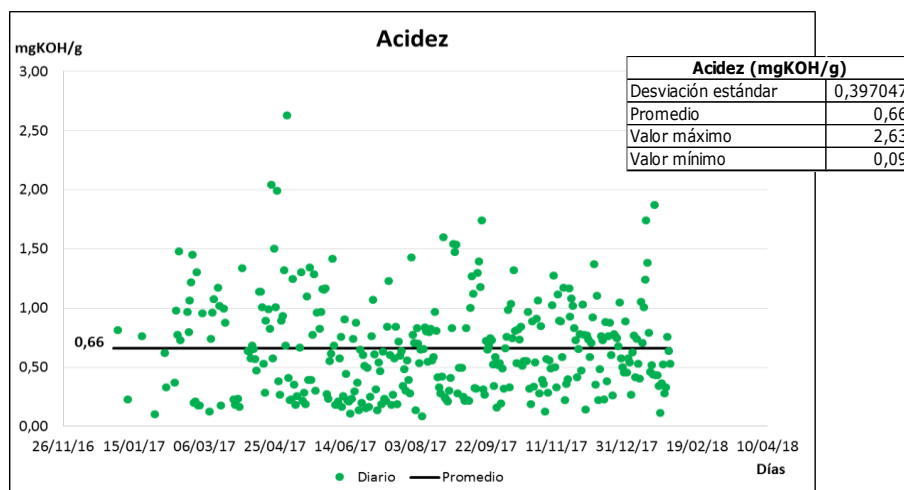


Figura 4. Tendencia de acidez en el crudo mezcla Vasconia durante el año 2017 y enero del 2018.

2.4 Espectroscopia del infrarrojo cercano (NIR)

El infrarrojo (IR, *InfraRed*) es la región del espectro electromagnético comprendida entre los 760 y los 10^6 nm ($12800 - 10\text{cm}^{-1}$, en número de onda) entre el visible y las microondas en el espectro electromagnético, como se muestra en la Figura 1.5. La radiación infrarroja no es lo suficientemente energética como para provocar transiciones electrónicas como sucede en el ultravioleta – visibles (UV – Vis) o los rayos X, por lo que las bandas que aparecen en esta zona corresponden principalmente a cambios en los estados vibracionales y rotacionales de la molécula.

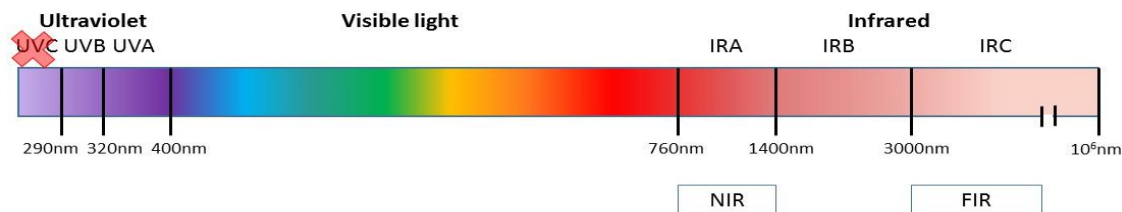


Figura 5. Diagrama del espectro electromagnético. Adaptado de: Atlas of Science (2016).

En la identificación cualitativa de grupos funcionales orgánicos e inorgánicos, la presencia de bandas de combinación y sobretonos en la región NIR limita la técnica. A pesar de esto, se han desarrollado tablas de asignación de bandas de absorción a grupos funcionales en la región NIR, que orientan el análisis cualitativo. La Tabla 1.3 presenta algunas asignaciones de bandas de absorción a grupos funcionales típicamente presentes en hidrocarburos. En un espectro NIR típico para un crudo, las principales bandas se encuentran en el rango electromagnético entre 6000 y 3600 cm^{-1} , en donde se obtienen espectros de absorción de muy buena calidad para los hidrocarburos debido a la combinación de las bandas vibracionales de los enlaces C-H (Pasquini & Bueno, 2007).

Tabla 3.

Frecuencias de absorción de algunos grupos funcionales en hidrocarburos.

Grupo funcional	Frecuencia (cm^{-1})
Acetilenos	9800 – 9430
	6580 – 6400
Alcoholes	7140 – 7010
Aldehídos	4640 – 4520
Alifáticos	8000
Aromáticos	4525
	4445
Formaldehído	4775 – 4630

Nota. Adaptado de Patnaik (2004).

Los espectros deben ser pretratados para corregir desviaciones tales como la línea base (desviación de la línea horizontal para el menor valor del espectro ocasionada por fluctuaciones de la señal eléctrica, inestabilidad de la fuente de luz, cambios en la temperatura) y errores asociados a la operación del equipo.

2.5 Análisis quimiométricos

La información analítica contenida en los espectros NIR, es poco selectiva y muy influenciada por un gran número de factores físicos, químicos e instrumentales, por lo cual, el análisis de estos espectros se suele realizar utilizando los métodos de la quimiometría. Uno de los métodos utilizados es el de los componentes principales (PCA, *Principal Component Analysis*), con el cual, la variabilidad presente en los datos originales es concentrada en un número reducido de variables denominadas componentes principales (PC). Cada componente principal contiene información de diferente relevancia, con lo cual es posible determinar las fuentes de variabilidad en el conjunto de datos iniciales, principalmente en forma gráfica a través de los *scores* y los *loadings*. Con esto es posible realizar un análisis entre las semejanzas y diferencias entre diferentes muestras (Aske *et al.*, 2001).

Por otra parte, la información espectroscópica puede correlacionarse con magnitudes físicas o químicas de las muestras, conduciendo a modelos de regresión. El fundamento de cualquier calibración es establecer una relación entre la señal instrumental y la propiedad a determinar, y posteriormente predecir esta propiedad a partir de la señal medida en una muestra desconocida. La regresión por mínimos cuadrados parciales (PLS, *partial least square*) concentra la información relevante en nuevas variables, de forma similar al PCA, maximizando la matriz de

varianza - covarianza. La regresión se aplica sobre las nuevas variables, eliminando así los problemas de colinealidad que puedan tener los datos originales. En los modelos PLS, se establece una representación de la matriz X en término de estas variables latentes, según:

$$X = TP^T + E$$

Donde, T representa *scores* y P^T *loadings*. La matriz X queda descompuesta en un número de variables latentes, cada una caracterizada por un vector T y un vector P^T . De esta forma, es posible representar una matriz X (dentro de un cierto error) por una matriz T con un número menor de columnas. Si se incluyen todas las variables latentes, el error es cero ($E=0$). El número de variables latentes necesarias para explicar la matriz X (dentro de un error máximo permitido) es una medida de la complejidad del modelo (Blanco *et al.*, 2001).

2.6 Estado del arte

Considerando la necesidad de caracterización y predicción de las propiedades de los crudos, la literatura reporta diversos trabajos en donde se exponen tecnologías basadas en mediciones con espectroscopia NIR, viscosímetros en línea y otras. Pasquini & Bueno, 2007, aplican espectros NIR en el rango de 5000 a 3900 cm^{-1} , en la predicción de la gravedad API y la curva TBP para 79 muestras de crudos mezcla con un rango de gravedad API de 23.7 – 30.4° y 43 muestras de crudos sin mezclar con gravedad API extendida desde 13.2 hasta 49.6°, por medio de regresión de mínimos cuadrados parciales (*PLS*). Estos autores obtienen un valor de error cuadrático medio de predicción (*RMSEP*) y coeficiente de determinación (r^2) para la curva de destilación TBP de 1.13 V% y 0.688, respectivamente (para la gravedad API de 0.24 y 0.923, respectivamente). Por su parte, Falla *et al.*, 2006, proponen una metodología para la estimación de la curva de

destilación para 40 muestras de crudos con grados API entre 31 y 36.4, basada en espectroscopia NIR (5600 y 6000 cm^{-1}) y redes neuronales artificiales. Falla *et al.* emplean una red de alimentación de tres capas (entrada, oculta y salida) en donde las variables de entrada son el espectro de absorción y el grado API; y la variable de salida es el porcentaje en peso evaporado a diferentes temperaturas. Estos autores presentan valores de *RMSEP* que varían desde 0.66 hasta 3.5 % y coeficiente r^2 igual a 0.9931. Asimismo, Argirov *et al.*, 2012, utilizan una base de datos con propiedades para 117 crudos (*i.e.* densidad, azufre, viscosidad, metales, carbono Conradson, contenido de asfalto y cortes TBP) y el método de distribución de dos parámetros de Riazi para el desarrollo de un modelo para la predicción de la curva TBP. Hidajat & Chong, 2000, correlacionan los espectros NIR en el rango de 6250 a 3800 cm^{-1} de 110 muestras de crudo con sus curvas de destilación TBP y su densidad por medio de 14 modelos de regresión PLS, obteniendo coeficientes de correlación superiores a 0.98. También, Chu *et al.*, 2011, proponen un modelo para la identificación rápida del tipo de crudo, denominado método del coeficiente de correlación de la ventana móvil, basados en la espectroscopia NIR, en la región comprendida entre 4000 y 6050 cm^{-1} , y una base de datos de 265 crudos a través de la regresión PLS; según sus resultados el coeficiente de determinación de cada ventana móvil analizada supera 0.9940. Por último, Meléndez *et al.*, 2012, desarrollan 8 modelos quimiométricos para predecir el análisis SARA en 50 muestras representativas de crudos colombianos mediante espectroscopia infrarroja por transformada de Fourier (ATR-FTIR) usando regresión PLS; según sus resultados, para cada componente del análisis los errores estándares de la predicción (SEP) variaron de 0.4 a 1.9 en crudos livianos y de 1.4 a 3.7 en crudos pesados; para todos los casos el coeficiente de determinación (r^2) fue superior a 0.95.

De igual manera, la aplicación del NIR se extiende al análisis y la regresión de cortes de crudo como diésel, jet, gasolina, gasóleo y asfalto. Algunos trabajos reportados en la literatura son: (i) Zhi-Na *et al.*, 2012) utilizan espectros NIR y regresión PLS en el desarrollo de un modelo para la verificación de la calidad del diésel, por medio de la estimación de la tasa de cambio de la viscosidad cinemática a 100°C y el punto de inflamación. (ii) Balabin *et al.*, 2010, clasifican las gasolinas según su fuente (destilación y proceso) con base en espectros NIR; (iii) Brouillette *et al.*, 2016, desarrollaron 22 regresiones para propiedades de los cortes diésel, gasolina y jet con valores de error comparables con la reproducibilidad del método ASTM correspondiente; (iv) Laxalde *et al.*, 2011, desarrollaron modelos predictivos a partir de espectros NIR para el contenido de compuestos saturados, aromáticos, resinas y asfaltos (SARA) en fracciones pesados del petróleo; (v) Blanco *et al.*, 2001, reporta un método para la determinación de varias propiedades fisicoquímicas y la composición química de asfaltos usando NIR y regresión PLS.

En trabajos realizados en ECOPETROL S.A., no publicados en literatura abierta, se han desarrollado modelos de espectroscopia NIR combinada con técnicas como el análisis de componentes principales (PCA) y la regresión PLS (ver Echeverría, 2016). Aunque estos modelos predicen la gravedad API y el azufre en mezclas de crudos, los desarrollos no han considerado todavía la regresión de la curva TBP obtenida mediante la norma ASTM D2892 y/o por SimDis con la ASTM D7169.

Según la revisión de literatura, diferentes modelos han sido propuestos para la predicción de las propiedades para el crudo y sus fracciones de destilación, a partir de los espectros FT-NIR. Los desempeños de las regresiones son variados dependiendo de las características de los crudos, medidas en términos de sus propiedades globales como °API. Los mejores desempeños en el

ajuste de las curvas TBP o SimDis han sido reportados para regresiones que consideran crudos livianos.

3. Metodología

3.1 Recolección de muestras crudo vasconia

En esta primera etapa, 100 muestras de crudo mezcla Vasconia fueron recolectadas en un periodo de seis meses. Los tiempos de recolección de las muestras fueron aleatorizados para evitar variaciones diferentes a los contenidos de crudos en la mezcla. Las muestras fueron enviadas al laboratorio de analítica de la GRB para las mediciones de propiedades y espectros.

3.2 Medición de espectros FT-NIR y de propiedades

En el laboratorio, cada muestra recolectada fue analizada por espectroscopia FT-NIR, en la región espectral que comprende 9000 y 3500 cm^{-1} , con una resolución de 2 cm^{-1} . El espectro fue medido utilizando del equipo para análisis de muestras de crudos fuera de línea *ABB* modelo *MB3000*. El espectro FT-NIR consistió de 32 escaneos, tomando aire como el espectro de referencia.

Por su parte, la curva SimDis fue obtenida siguiendo el protocolo establecido en la norma ASTM D7169. Según esta norma, como primer paso se realiza una calibración del cromatógrafo

utilizando una mezcla de patrones de parafinas diluidas en CS₂. Las muestras de crudo y el gasóleo de referencia fueron también diluidos en el disulfuro de carbono. El CS₂ fue inyectado al cromatógrafo, con lo cual fue definida la línea base de la muestra de calibración; después se inyectaron el gasóleo y la muestra de mezcla Vasconia a analizar. Los detalles de operación del cromatógrafo, así como, los procedimientos numéricos para el cálculo de la curva SimDis son detallados en la norma ASTM D7169.

Asimismo, las propiedades °API, contenido de azufre, contenido de agua, acidez y sal fueron medidas para las muestras recolectadas; estas propiedades fueron determinadas según los protocolos establecidos en las respectivas normas ASTM. Es importante mencionar que el laboratorio de analítica permanentemente verifica la reproducibilidad de las mediciones instrumentales y analíticas, por medio de los protocolos internos establecidos en la GRB.

3.3 Análisis y pretratamiento de espectros NIR

Los espectros NIR fueron pretratados con el fin de eliminar la variabilidad en las mediciones del equipo, incrementar la relación señal/ruido y compensar las variaciones en la intensidad de la fuente de emisión IR (Rinnan, 2014; Bi *et al.*, 2015). Los pretratamientos aplicados fueron corrección por línea base, suavizado según el procedimiento de Savitzky-Golay, y primera y segunda derivadas por medio de la aproximación de Savitzky-Golay. Los pretratamientos fueron aplicados según las funciones internas codificadas en el programa *The Unscrambler X* versión 10.3, licenciado para Ecopetrol GRB.

3.4 Relaciones entre las propiedades y la composición

Las propiedades globales determinadas para las muestras fueron analizadas con el fin de determinar tendencias relacionadas con la composición predominante en la mezcla Vasconia. Este análisis fue guiado utilizando los protocolos del método *kmean* según la distancia euclidiana. La predicción en los contenidos de aromáticos, naftenos y parafinas fue realizada según las regresiones expuestas por Riazi & Daubert, 1986, y Riazi, 2005.

3.5 Análisis por componentes principales

Los espectros FT-NIR pretratados fueron dispuestos en una matriz en donde las filas representaban las muestras de mezcla Vasconia, mientras que las columnas representaban los aportes para cada número de onda. A esta matriz se le aplicó el análisis por componentes principales (PCA, *principal component analysis*), obteniendo los respectivos *scores* y *loadings*. Con los *scores* y los *loadings* de los componentes principales con mayor varianza se procedió a la detección de patrones en los espectros relacionados con las composiciones y las propiedades de las muestras de las mezclas Vasconia. La búsqueda de patrones o agrupaciones fue definido según el método *kmean* según la distancia euclidiana. La información contenida en los *scores* y los *loadings* fue obtenida principalmente por inspección visual de las respectivas gráficas de dispersión de los componentes principales, tanto en sus valores *scores* como en sus valores *loadings*. El análisis PCA fue aplicado según los protocolos programados en *The Unscrambler X* versión 10.3.

3.6 Desarrollo de regresiones PLS

En esta etapa diferentes modelos de regresión PLS basados en los espectros NIR fueron calibrados y validados en el ajuste de las curvas de destilación SimDis medidas para la mezcla Vasconia. Las temperaturas para los evaporados al 5%, 10%, 20%, 30%, 40%, 50%, 70% y UBP (punto de ebullición último) fueron utilizadas en el ajuste. Asimismo, los rendimientos a las temperaturas 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 y 550 °C fueron utilizadas en un ajuste posterior; estos rendimientos fueron determinados por medio de interpolación con trazadores cúbicos que preservan la tendencia (*pchip*, *piecewise cubic hermitian interpolating polynomial*) según el código de la función *pchip* del programa de uso libre R versión 3.4.3 (<https://www.r-project.org>). También, las fracciones de aromáticos, naftenos y parafinas fueron ajustadas por medio de regresiones PLS posteriores.

La validación de las diferentes regresiones fue aplicada según el protocolo LOO (*leave-one-out*) enmarcado en la validación cruzada (Li *et al.*, 2002). Los estadísticos coeficiente determinación múltiple (R^2), raíz del error cuadrático medio (RMSE) y porcentaje de RMSE (RMSE%) fueron utilizados en la determinación del desempeño de cada ajuste realizado. Estos estadísticos se definen según,

$$R^2 = 1 - \frac{\sum (y_i - \tilde{y}_i)^2}{\sum (y_i - \bar{y})^2}, \quad RMSE = \sqrt{\frac{\sum (y_i - \tilde{y}_i)^2}{n}}, \quad RMSE \% = \frac{RMSE}{\bar{y}} * 100$$

Donde, y_i , $\tilde{y}_i$, $\bar{y}$, n corresponden al valor de la propiedad experimental para la muestra i -ésima, el valor obtenido con el modelo de regresión para la propiedad, el promedio muestral de la propiedad y el número de muestras utilizadas, respectivamente.

Las regresiones PLS fueron depuradas mediante un proceso repetitivo de eliminación de números de onda con contribución nula, hasta lograr la reducción al mínimo número de variables independientes en los modelos finales. Las regresiones PLS fueron desarrolladas según los códigos del programa *The Unscrambler X* versión 10.3.

4. Análisis de resultados

4.1 Espectros FT-NIR mezclas Vasconia

Las muestras de crudo de la mezcla Vasconia recolectadas en diferentes tiempos fueron enviadas al laboratorio para la medición de sus espectros FT-NIR y sus propiedades; en total, los espectros de 100 muestras de crudo fueron medidos, así como sus propiedades gravedad API, azufre, acidez, agua y sal. Los espectros en el rango $9000\text{--}3500\text{ cm}^{-1}$ son presentados en la Figura 3.1; de esta figura se deduce el requerimiento de corrección de la línea base para cada espectro medido. La línea base de los espectros de las muestras de mezcla Vasconia exhibe dos regiones con características diferentes; la primera comprendida entre 6500 y 3500 cm^{-1} contiene un desfase constante de la línea base (*offset*), mientras que la segunda región presenta con un incremento monótonico no lineal de la línea base, denominado pendiente de absorción (*absorption slope*), a partir de 6500 cm^{-1} (Pasquini & Bueno, 2007; Laxalde *et al.*, 2014). Según Mullins, 1990, la aparición de una pendiente de absorción en la mencionada región es característica en crudos con contenidos apreciables de asfaltenos, los cuales exhiben bandas de

transición electrónica $\pi - \pi^*$ y $n - \pi^*$ residuales que, aunque pertenecientes al espectro UV-vis, modifican la parte inicial del espectro IR cercano. Por otra parte, Aske *et al.*, 2001, adjudican la aparición de *offset* en la línea base a la agregación de asfaltenos en la muestra medida. En efecto, los contenidos apreciables de asfaltenos en la mezcla Vasconia han ocasionado ensuciamiento en líneas y en intercambiadores con eventos de mantenimiento, debido a una posible problemática de incompatibilidad en la mezcla formada. Las líneas bases de los espectros fueron corregidas, con lo cual se generaron las tendencias mostradas en la Figura 3.2.

Por otra parte, a pesar de que el espectro FT-NIR presenta bandas con gran cantidad de solapamiento, esta región contiene vibraciones de importancia (Pantoja *et al.*, 2018). Según la Figura 3.1, la contribución espectroscópica de las fracciones de los crudos presenta regiones de poca relevancia, relacionadas con segundos sobretonos de las bandas de estiramiento C–H en las regiones $10000 - 6500 \text{ cm}^{-1}$ y $5300 - 4800 \text{ cm}^{-1}$; las bandas de C–H para esta región proviene de los aromáticos, solapamientos del primer sobretono C–H y combinaciones el estiramiento del enlace C=C (Hannisdal *et al.*, 2005; Abbas *et al.*, 2012). Estas contribuciones exhiben valores inferiores a 10% de la intensidad de la señal (Figura 3.1). En contraposición, en las regiones $5900 - 5665 \text{ cm}^{-1}$ y $4500 - 4000 \text{ cm}^{-1}$ se exhiben las mayores contribuciones del espectro; la primera región manifiesta la contribución de los primeros sobretonos de las bandas de estiramiento del grupo C–H, mientras que en la segunda región, combinaciones de bandas para el mismo grupo funcional son visualizadas (Pantoja *et al.*, 2018; Pabón & Filho, 2019).

De otro lado, la Figura 3.3 exhibe los pretratamientos adicionales aplicados a los espectros; estos pretratamientos consistieron en la primera y la segunda derivada. El efecto de las derivadas acentúa la contribución de las regiones discutidas anteriormente, lo cual facilita la obtención de modelos de predicción de propiedades a partir del espectro FT-NIR.

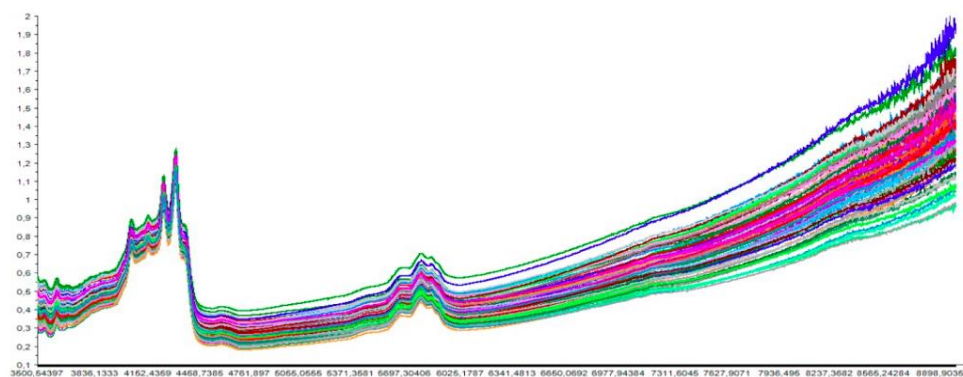


Figura 6. Espectros FT-NIR medidos de las muestras de crudo mezcla Vasconia.

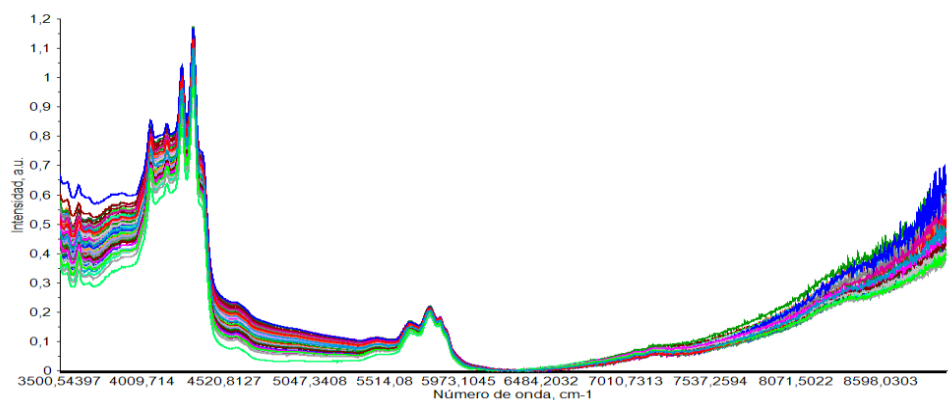


Figura 7. Espectros FT-NIR con corrección por línea base.

4.2 Propiedades de las mezclas Vasconia

Respecto a las propiedades macroscópicas, las Figuras 3.4 y 3.5 muestran las tendencias para las 100 curvas de destilación SimDis –representadas como temperatura al 0, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95 y 100% de evaporados–, mientras que la Figura 3.6 relaciona la matriz de dispersión para los resultados de Acidez, Agua, °API, Azufre, Sal y temperatura de evaporados al 50% (T50). Según la Figura 3.4, el punto inicial de ebullición (IBP) para todas las muestras, excepto una, presenta un valor alrededor de los 30 °C; la muestra 7 inicia la curva SimDis en 70

°C. Asimismo, todas las muestras presentan valores de temperatura hasta T70, mientras que el número de muestras con resultados para mayores valores de evaporados disminuye; para T80, T90, T95 y FBP el número de muestras con valores corresponde a 76, 38, 9 y 2, respectivamente (parte final Figura 3.4). Esta disminución manifiesta el carácter variable de las proporciones de los crudos constituyentes en la mezclas Vasconia (Tabla 1.2). El rango de variación para cada temperatura de evaporados puede ser visualizada en los gráficos *boxplot* de la Figura 3.5; según esta figura, el rango intercuartíl de la curva aumenta con el aumento en el porcentaje de evaporados; esto sugiere que las mayores variaciones en la curva SimDis provienen de las fluctuaciones en las proporciones de los crudos pesados constituyentes de la mezcla Vasconia.

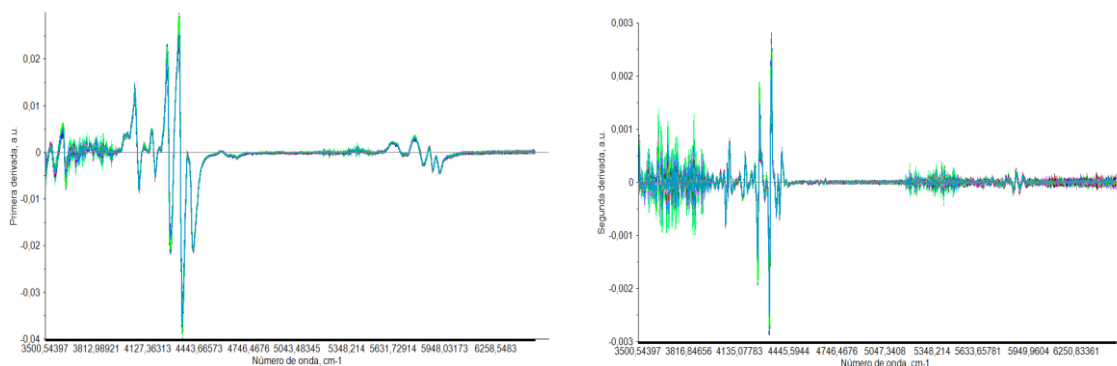


Figura 8. Pretratamientos efectuados a los espectros FT-NIR (rango 6500–3500 cm^{-1}) muestras de crudo mezcla Vasconia. Iz: primera derivada, der: segunda derivada.

Esta variabilidad en la composición de la mezcla Vasconia también es expresada en los valores de las propiedades de la Figura 3.6. Según esta figura, los rangos de variación de las propiedades de la mezcla Vasconia es amplio; la acidez, el agua, los °API, el azufre, la Sal y la T50 varían en los intervalos [0,1 – 1,5] mg/g, [0 – 0,55] %, [20 – 28], [6000 – 13000] ppm, [4 – 17] PTB (*pounds of salt per thousand barrels of crude*) y [340 – 440] °C, respectivamente. Los estadísticos univariados para estas propiedades son mostrados en la Tabla 3.1. De la Figura 3.6

es posible visualizar tendencias en el comportamiento de las propiedades; *i.e.* la acidez, el contenido de agua, el azufre y el contenido de sal tienen un comportamiento inversamente proporcional con el API, con lo cual, la mezcla Vasconia exhibe una característica de un crudo intermedio agrio (Riazi, 2005).

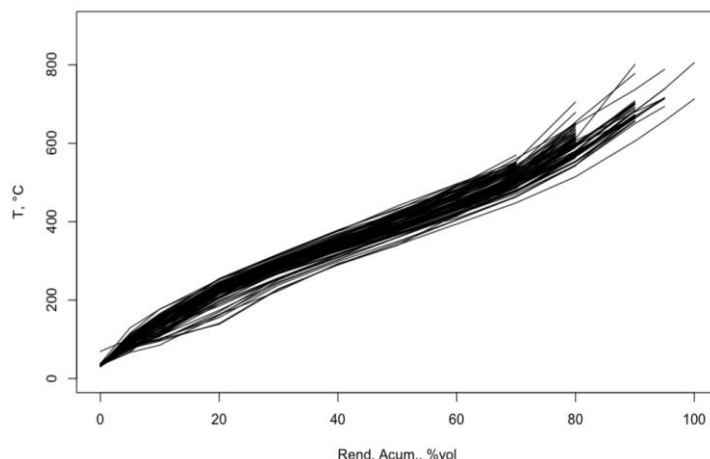


Figura 9. Curvas SimDis medidas a las muestras de crudo mezcla Vasconia.

El factor UOP o K_w^3 y el constante VGC^4 (*viscosity gravity constant*) determinan una clasificación de los crudos a partir de su propiedades globales (Riazi, 2005). Valores de $K_w < 11$ o $VGC < 0.75$ indican crudos con características aromáticas, mientras $K_w > 13$ o $VGC > 0.95$ son representativos de crudos parafínicos; los crudos nafténicos exhiben K_w o VGC con valores intermedios. Según la Tabla 3.1 y la Figura 3.7, los valores exhibidos de K_w y VGC por la mezcla Vasconia son característicos para un crudo con preeminencia nafténica. Con esto, a pesar de las fluctuaciones, la mezcla Vasconia se encuentra constituida por componentes similares

$K_w = \frac{(1.8 * VABP)^{1/3}}{SG}$, donde VABP es la temperatura media de ebullición en K y SG a la gravedad específica.
 $VGC = (10 * SG - 1.0752 * \log_{10}(v_{38} - 38)) / (10 - \log_{10}(v_{38} - 38))$ Donde v_{38} está en SUS.

durante el tiempo de muestreo. De hecho, uno de los mayores componentes de la mezcla Vasconia es el crudo Rubiales con característica pesada y nafténica.

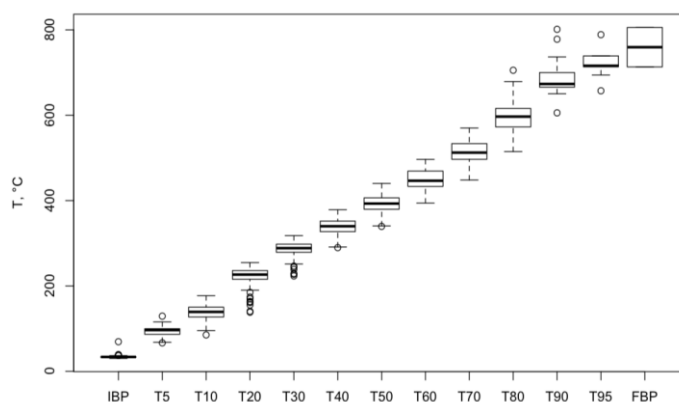


Figura 10. Diagramas *boxplot* temperaturas de evaporación muestras de crudo mezcla Vasconia. IBP y FBP hacen referencia al 0% y al 100% de evaporados, respectivamente

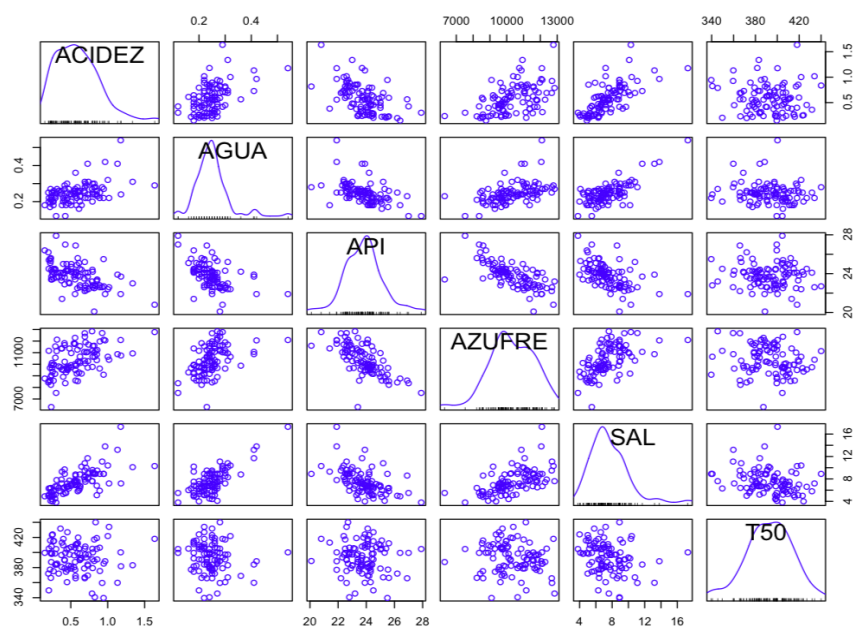


Figura 11. Matriz de dispersión propiedades crudo mezcla Vasconia.

Tabla 4.

Estadísticos univaridos para las propiedades de las mezclas Vasconia.

Propiedad	Mínimo	Máximo	Media	Q1	Q3	Des. Est.
Acidez, mg/g	0,15	1,64	0,58	0,34	0,75	0,28
Agua, %masa	0,12	0,54	0,25	0,22	0,27	0,06
°API	20,10	27,90	23,81	22,88	24,40	1,26
Azufre, ppm	6307	12844	10355	9552	11281	1255
Sal, PTB	3,80	17,30	7,51	6,23	8,90	2,16
T50, °C	339,6	440,1	392,8	380,0	406,4	20
K_w	11,27	12,04	11,68	11,60	11,79	0,15
VGC	0,84	0,90	0,87	0,86	0,88	0,01
M	263,4	388,4	336,5	320,4	357,4	26,9
RI	1,052	1,065	1,058	1,057	1,059	0,002
x_a , vol	0,15	0,32	0,23	0,21	0,25	0,03
x_p , vol	0,54	0,61	0,20	0,56	0,58	0,02
x_n , vol	0,13	0,23	0,20	0,19	0,21	0,01

El cálculo de las fracciones volumétricas para los componentes aromático, nafténico y parafínico fue efectuado por medio de las relaciones descritas en Riazi & Daubert, 1986; estas relaciones se encuentran en términos del peso molecular (M), del intercepto de refractividad (RI, *refractivity intercept*) y del VGC⁵. La Tabla 3.1 también expone los valores obtenidos para M, RI, VGC y las fracciones (x_a , x_n , x_p); en esta tabla se verifican los valores de peso molecular requeridos en las regresiones de las fracciones volumétricas (M>200). Según la Tabla 1, la mayor contribución al volumen de la mezcla Vasconia se encuentra en la fracción parafínica; esta contribución se debe al menor valor de su densidad comparada con la correspondiente para las fracciones aromática y nafténica (Riazi, 2005).

5 Para peso molecular superior a 200 (M>200),
 $x_a = -8,66 + 7,37 \cdot RI + 1,255 \cdot VGC$; $x_p = -9 + 12,53 \cdot RI - 4,228 \cdot VGC$; $x_n = 1 - x_a - x_p$

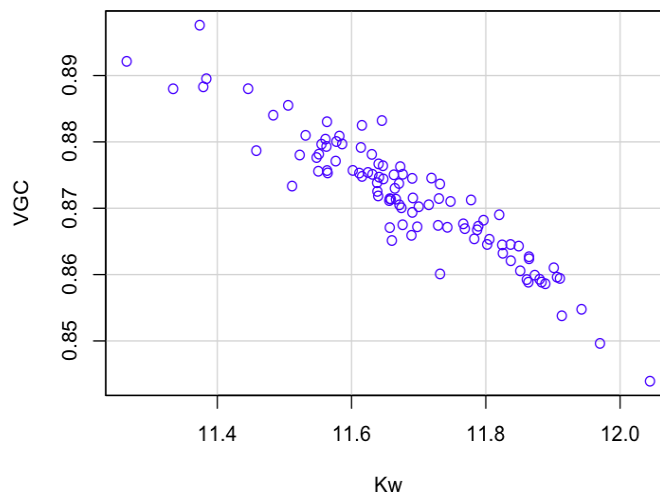
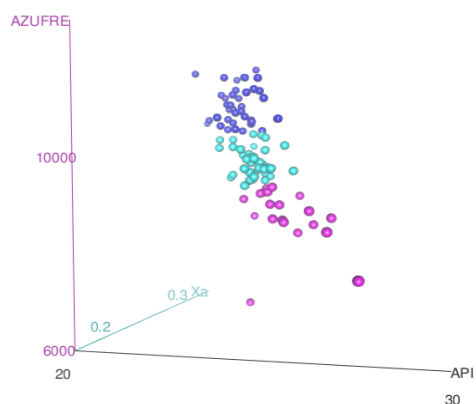
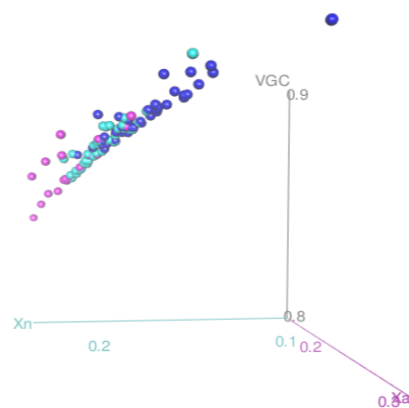


Figura 12. Figura de dispersión para K_w vs VGC de las muestras de crudo mezcla Vasconia.

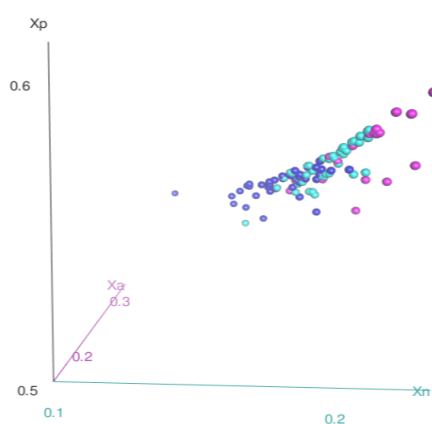
La contribución de cada fracción a las propiedades de las mezclas de crudo Vasconia fue estimada por medio de la aplicación del método de agrupación *kmean*. Los valores de las propiedades y parámetros definidos en la Tabla 3.1 fueron tomados como datos de entrada para el algoritmo, considerando la distancia euclidiana como métrica de valoración. Los resultados del algoritmo *kmean* definiendo 3 agrupaciones –que representan a las fracciones parafínica, nafténica y aromática– son mostrados gráficamente en las gráficas de dispersión en el espacio de propiedades (Figura 3.8). las muestras referentes a las agrupaciones 0, 1 y 2 son detalladas en colores fucsia, verde claro y azul, respectivamente. Por cada agrupación se tienen 20, 39 y 41 muestras, respectivamente.



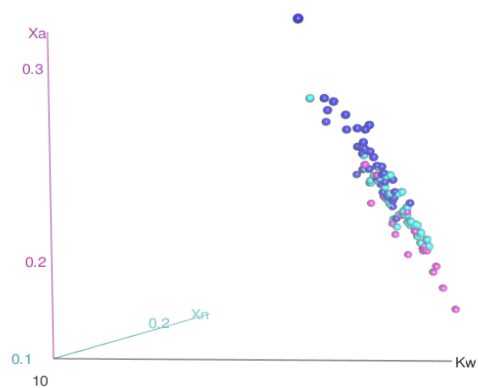
(a) x_a vs API vs Azufre



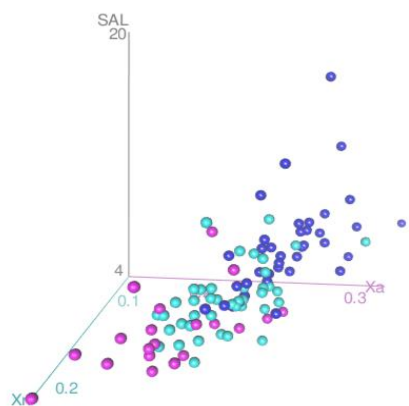
(b) x_a vs x_n vs VGC



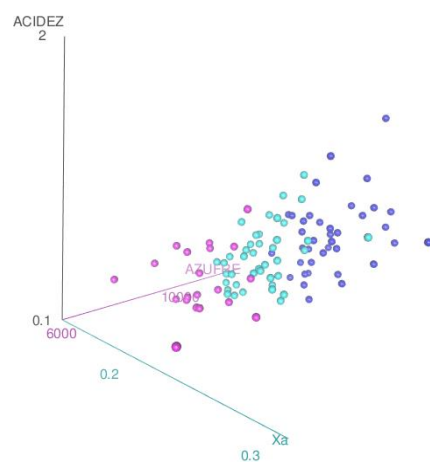
(c) x_a vs x_n vs x_p



(d) x_a vs x_n vs K_w



(e) x_a vs x_n vs SAL



(f) x_a vs Acidez vs Azufre

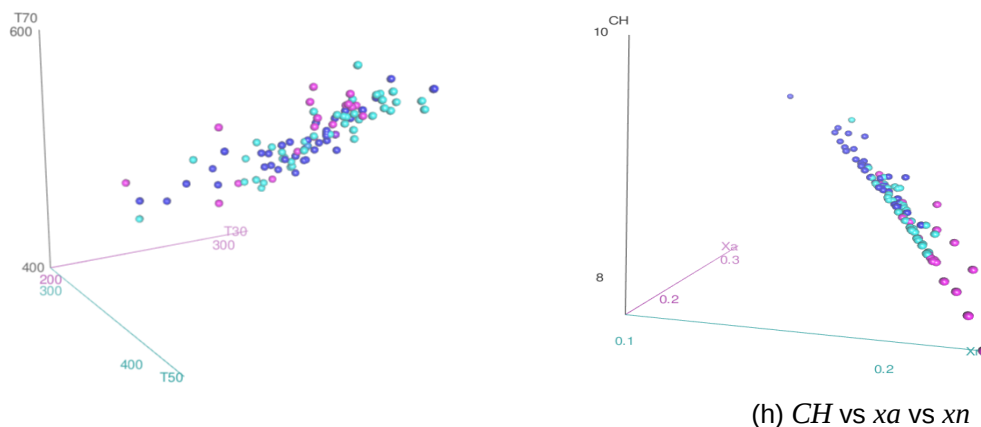


Figura 13. Gráficos de dispersión para K_w vs VGC de las muestras de crudo mezcla Vasconia. Agrupación 0 en color fucsia, agrupación 1 en color verde claro y agrupación 2 en azul.

Según la Figura 3.8(a), las tres agrupaciones aparecen con límites definidos en el eje de la propiedad Azufre. Las agrupaciones 0, 1 y 2 poseen valores de azufre entre [6307 – 9300], [9300 – 10800] y [10800 – 12844] ppm, respectivamente. Por su parte, la Figura 3.8(b) presenta la distribución según el parámetro VGC y las fracciones aromática y nafténica; según esta figura, la agrupación 0 enmarca muestras con características nafteno-parafínicas, mientras que las agrupaciones 1 y 2 enmarcan muestras con características nafténicas y nafteno-aromáticas, respectivamente, coincidiendo con los intervalos en propiedades reportados en literatura (Pabón & Filho, 2019; Riazi, 2005). Estas características son verificadas en el espacio de las fracciones y en el eje del parámetro K_w , Figuras 3.8(c) y 3.8(d), respectivamente. La diferencias de medias entre las diferentes agrupaciones en el espacio de las fracciones son visualizadas en la Figura 3.9. Asimismo, las Figuras 3.8(e)-(h) ilustran las tendencias de cada agrupación según las propiedades Sal, Acidez, contenido de carbono a hidrógeno y temperaturas. En estas propiedades, excepto para las temperaturas de evaporado, se aprecia una separación entre las agrupaciones.

De acuerdo al análisis realizado en esta sección, las diferentes muestras de mezcla Vasconia presentan una componente nafténica base (valores de K_w y VGC, Tabla 3.1), con fluctuaciones en los aportes de los crudos constituyentes con composición parafínica y aromática (Figura 3.9). Estas fluctuaciones influyen drásticamente en las propiedades exhibidas por las muestras de la mezcla Vasconia. La Tabla 3.2 define los rangos de fluctuación de las propiedades según los aportes de las fracciones y las agrupaciones obtenidas por el método *kmean*.

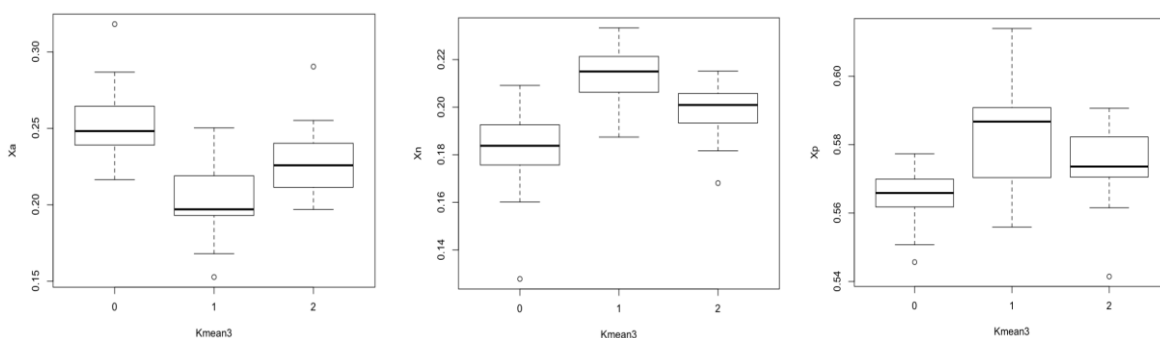


Figura 14. Gráficos *boxplot* de las fracciones de las mezclas Vasconia. Fuente: Autora.

Tabla 5.

Rangos de propiedades según la componente aportante en las mezclas Vasconia

Componente	Propiedad	Media	Desv. Estandar	Intervalo, 95%
Parafínica, $0,20 \leq x_a \leq 0,29$ $0,17 \leq x_n \leq 0,22$ $0,54 \leq x_p \leq 0,60$ Crudo nafteno- parafínico	Acidez, mg/g	0,38	0,21	$\pm 0,42$
	Agua, %masa	0,20	0,04	$\pm 0,08$
	API	25,3	1,18	$\pm 2,36$
	Azufre, ppm	8628	699	± 1388
	Sal, PTB	5,92	1,62	$\pm 3,24$
	T50, °C	392	20	± 40

Componente	Propiedad	Media	Desv. Estandar	Intervalo, 95%
Nafténica, 0,15≤xa≤0,25 0,19≤xn≤0,23 0,55≤xp≤0,61 Crudo Nafténico	Acidez, mg/g	0,55	0,24	± 0,48
	Agua, %masa	0,24	0,04	± 0,08
	API	23,9	0,70	± 1,40
	Azufre, ppm	9982	382	± 764
	Sal, PTB	6,91	1,48	± 2,96
Aromática, 0,22≤xa≤0,32 0,13≤xn≤0,21 0,55≤xp≤0,58 Crudo nafteno- aromático	T50, °C	397	22	± 44
	Acidez, mg/g	0,71	0,30	± 0,60
	Agua, %masa	0,28	0,07	± 0,14
	API	22,9	0,99	± 1,98
	Azufre, ppm	11631	564	± 1128
	Sal, PTB	8,95	2,17	± 4,34
	T50, °C	389	18	± 36

A nivel operativo, el conocimiento de las diferentes propiedades que puede presentar la mezcla Vasconia es de utilidad en la preparación de las condiciones operativas de los equipos como el desalador y las torres de destilación. La cantidad de campo eléctrico aplicado se encuentra en relación con la cantidad de sal en la mezcla; asimismo, la cantidad de azufre y la acidez del crudo influye en la inyección de inhibidores en las torres de destilación para evitar corrosión en los ductos de cima (UOP, 1978; Riazi, 2005).

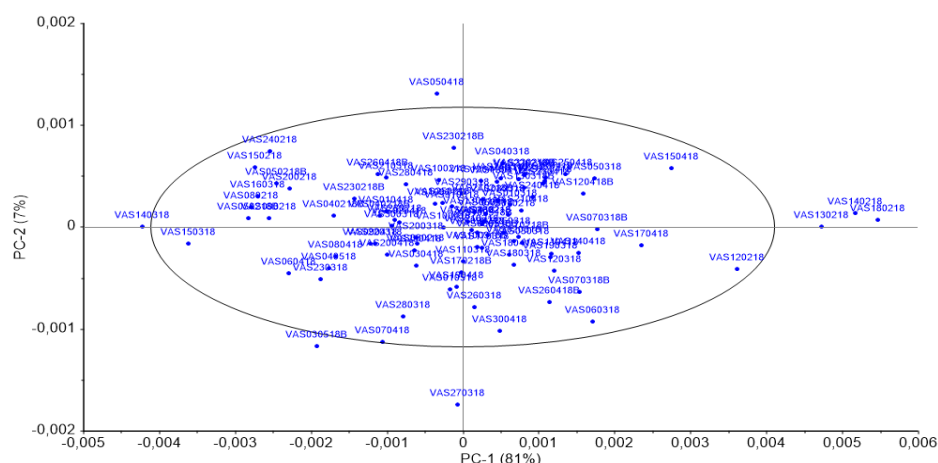
4.3 Análisis PCA

Los espectros pretratados fueron utilizados en la conformación de una matriz de análisis por componentes principales (PCA). Los resultados de los *scores* y los *loading*, representando a las muestras y las contribuciones en los números de onda, son presentados en la Figura 3.10. Según la Figura 3.10(a), la reproducción de la varianza lograda con el componente PC1 asciende a 81%, mientras que con el componente PC2 se alcanza un 7%; la contribución más importante en los *scores* la presentan las muestras que se encuentran por fuera de la elipse de Hotelling (*i.e.* VAS140218, VAS180218, VAS130218, VAS270318, VAS050418, VAS140318, VAS030618). Las muestras por fuera de la elipse, exhibiendo valores de PC1 negativos –VAS270318, VAS050418, VAS140318, VAS030618– se clasifican como resultantes de la fluctuación debido a la fracción aromática (Tabla 3.2), mientras las que exhiben valores de PC1 positivos –VAS140218, VAS180218, VAS130218– se clasifican como derivadas de fluctuaciones en la fracción parafínica, según la Tabla 3.2.

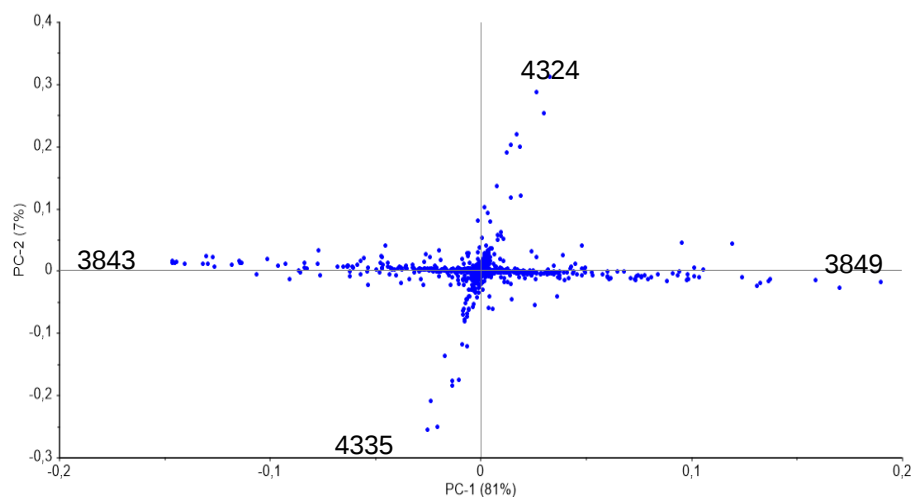
La Figura 3.10(b) ilustra los resultados de los PC1 y PC2 para los *loading*. De esta figura se sigue que los *loading* que representan a los números de onda 3843 y 3849 cm^{-1} poseen la mayor contribución para el PC1, en valor absoluto, mientras que los *loading* representando a los números de onda 4324 y 4335 cm^{-1} muestran la mayor contribución para el PC2. Una superposición entre las gráficas de los *scores* –Figura 3.10(a)– y los *loading* –Figura 3.10(b)– revela que las absorbancias pretratadas en los números de onda 3843 y 3849 cm^{-1} indican las fluctuaciones en composición de la mezcla Vasconia; *i.e.* espectros con valores elevados en la segunda derivada evaluada en el número de onda 3843 cm^{-1} indican un aumento en el componente aromático de la mezcla, mientras que valores extremos negativos en la segunda

derivada evaluada en 3849 cm^{-1} indican un aumento en el componente parafínico de la mezcla. Con esto, por medio del valor de la segunda derivada en los números de onda mencionados es posible inferir un intervalo en las propiedades (Tablas 3.2) para la mezcla en turno.

De otro lado, un análisis en las tendencias de las propiedades muestra que los números de onda 4324 y 4335 cm^{-1} son característicos de los valores para las temperaturas de evaporado. Las muestras VAS050418 y VAS270318 exhiben los valores extremos para las temperaturas de evaporado o curva SimDis (Figura 3.4); VAS050418 con el mayor valor de PC2 presenta el menor valor de temperaturas de evaporados, mientras que VAS270318 con el menor valor de PC2 presenta el mayor valor entre las curvas SimDis. Con lo anterior, las segundas derivadas en los números de onda 4324 y 4335 cm^{-1} sugiere la cantidad de compuestos volátiles contenidos en el crudo mezcla Vasconia.



(a) Scores



(b) Loadings

Figura 15. Gráficos de los *scores* y los *loading* resultantes del PCA aplicado a la matriz de espectros pretratados de las mezclas Vasconia. Fuente: Autora.

Los resultados de la aplicación del método *kmean* utilizando la matriz de espectros FT-NIR pretratada son mostrados en la Figura 3.11. Según el análisis anterior, las agrupaciones obtenidas utilizando las absorbancias pretradas del espectro podrían representar el aumento en cada componente de la mezcla Vasconia. Sin embargo, la Figura 3.12, que presenta las agrupaciones en los ejes de las propiedades, indica una carencia en la identificación de las componentes

fluctuantes en las mezclas. En esta figura, las muestras pertenecientes a cada agrupación se organizan sin ninguna tendencia aparente. Esta carencia en un patrón en las propiedades basada en la matriz de espectros FT-NIR manifiesta la variación en las composiciones de los crudos que conforman la mezcla Vasconia. Considerando la naturaleza nafténica de la mezcla, las variaciones se presentan en los crudos minoritarios; estas variaciones no afectan en gran medida las tendencias en las propiedades globales mostradas por las mezclas Vasconia, pero si las tendencias en sus espectros. Por lo anterior, las agrupaciones basadas en la aplicación del método *kmean* sobre el espectro FT-NIR no conduce a la generación de tendencias en las propiedades de las mezclas. Con esto, las agrupaciones considerando las propiedades macroscópicas son el mejor indicador de la contribución de cada componente en la mezcla Vasconia.

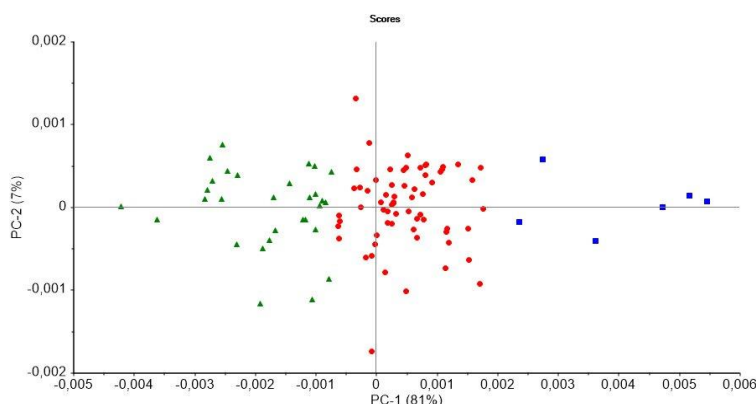


Figura 16. Identificación de las agrupaciones basadas en los espectros FT-NIR obtenidas con *kmean*=3 sobre el plano de los *scores*.

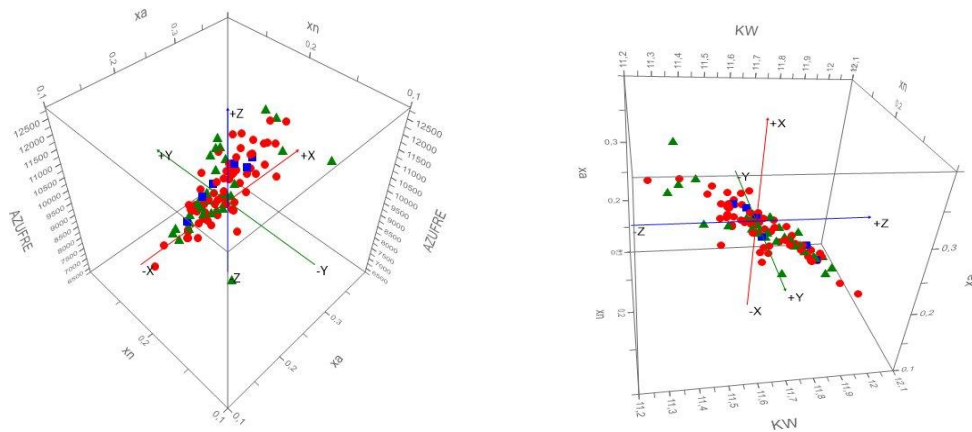


Figura 17. Identificación de las agrupaciones obtenidas con $kmean=3$ sobre el espacio de las propiedades. Fuente: Autora.

4.4 Regresión lineal múltiple

Los espectros pretratados fueron utilizados en la predicción de las temperaturas de evaporados por medio de la aplicación de la regresión lineal múltiple (RLM). Los resultados de la aplicación de la RLM son mostrados en la Figura 3.13. En esta figura se aprecia que el mejor desempeño en las RLM corresponde a la regresión para T80 con una reproducción en la varianza de 59%; asimismo, la etapa de validación cruzada resultó fallida en todos los casos, debido a valores cercanos a cero en la matriz de pretratamiento. Con lo anterior, el uso de las RLM no son una opción para la predicción de las temperaturas de evaporados.

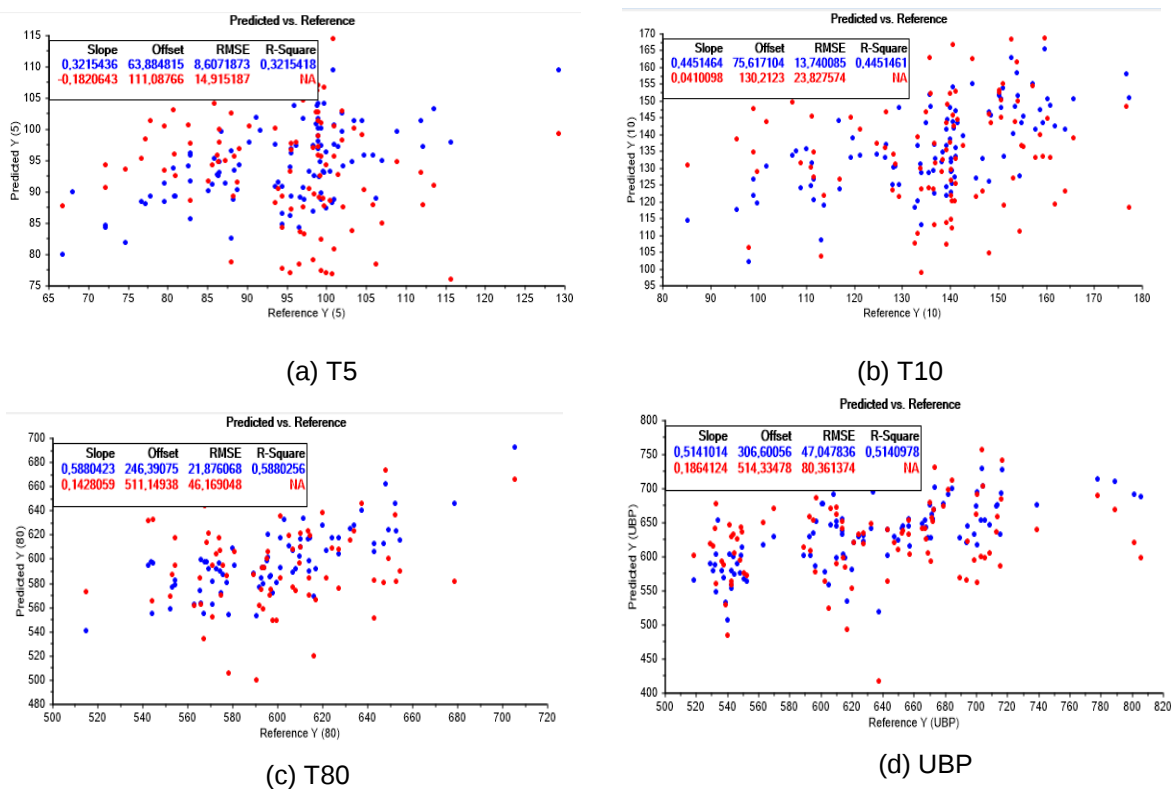


Figura 18. Gráficos de paridad de los resultados de las regresiones RLM en las diferentes temperaturas de evaporado.

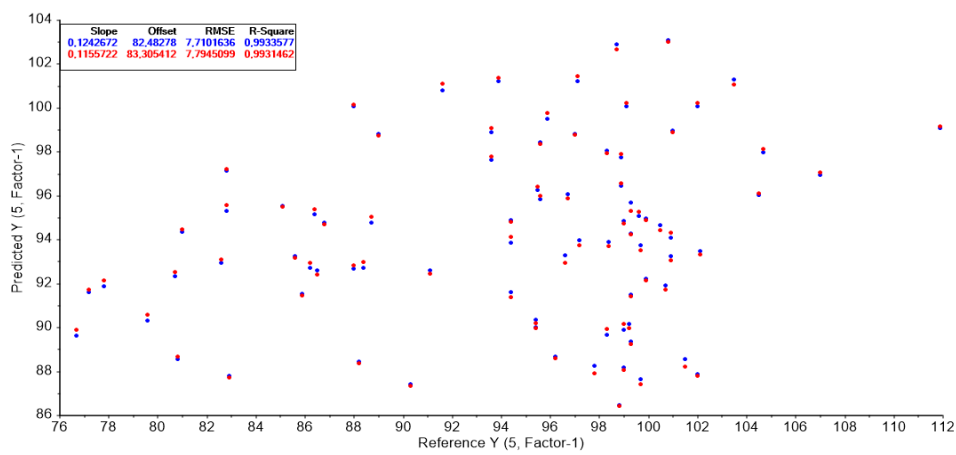
4.5 Regresiones PLS

4.5.1 Temperaturas de evaporados La regresión PLS fue aplicada en el ajuste de las temperaturas de evaporados utilizando la matriz de espectros FT-NIR pretratada. Las regresiones obtenidas con el algoritmo PLS fueron depuradas mediante un proceso repetitivo de eliminación de números de onda con contribución nula, con lo cual el número de variables independientes en los modelos fueron disminuídas favoreciendo la representatividad y la practicidad del modelo final. Las gráficas de paridad, de residuales y de contribución de los números de onda finales para los modelos de regresión depurados son mostradas en las Figuras 3.14, 3.15 y 3.16,

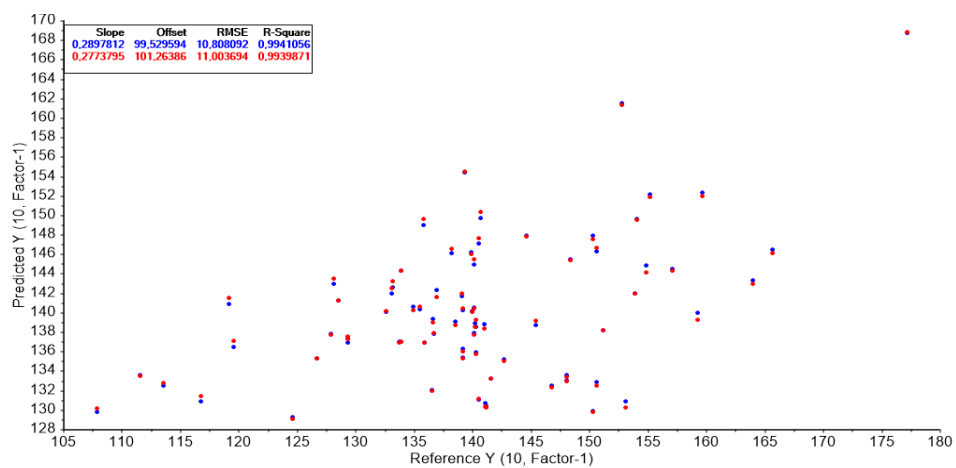
respectivamente. Según la Figura 3.14, en las gráficas de paridad la dispersión entre las predicciones de los modelos y los valores experimentales aumenta con el porcentaje de evaporados; este aumento se verifica con el aumento del RMSE en los respectivos modelos. El aumento en la incertidumbre de los modelos concuerda con el aumento en el rango de temperaturas para los evaporados de las diferentes mezclas Vasconia mostrado en las Figuras 3.4 y 3.5.

Por su parte, la Figura 3.15 presenta los residuales de los modelos de regresión depurados. En los residuales para T5, T10, T20 y T60 no se aprecian tendencias. Para las otras temperaturas de evaporado, incluyendo la última (UBP), las gráficas muestran aparentemente una tendencia decreciente con el aumento en la temperatura, sugiriendo la inclusión de nuevas variables independientes; sin embargo, la dispersión de los residuales sugiere que los sesgos mostrados son debidos a la carencia de muestras, con lo cual la inclusión de nuevos espectros conllevaría al cumplimiento de distribución aleatoria. Con esto, los modelos de regresión PLS obtenidos para las temperaturas de evaporados basados en los espectros FT-NIR representan la variabilidad expuesta por los respectivos datos experimentales.

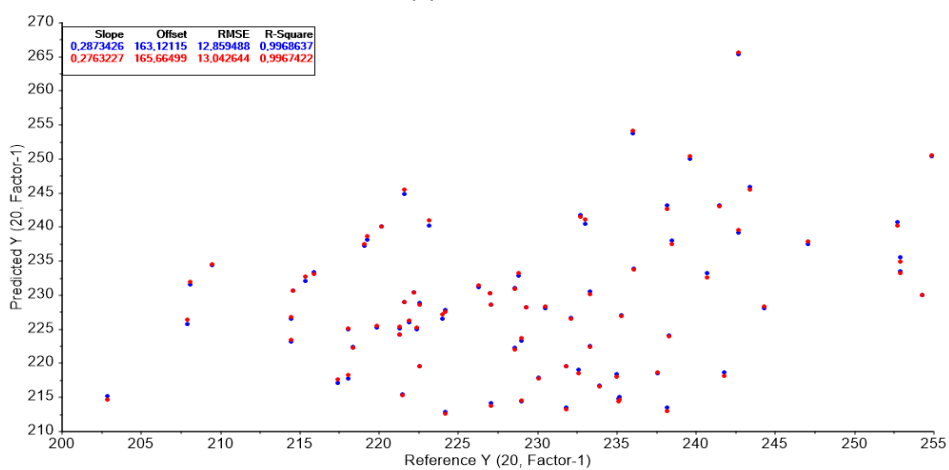
La Figura 3.16 expone los números de onda finales de los modelos de regresión. En esta Figura se identifican las dos variables independientes constitutivas de los modelos de predicción; *i.e.* las absorbancias pretratadas evaluadas en los números de onda 4324 y 4335 cm^{-1} definen las temperaturas de evaporados en los modelos PLS. Estos números de onda fueron identificados en el análisis PCA por denotar los valores extremos de las temperaturas de evaporados –Figura 3.10(b)–. Asimismo, todos los modelos coinciden en utilizar un solo factor o variable latente en la construcción de las regresiones. La Tabla 3.3 resume los modelos de regresión obtenidos para las temperaturas de evaporados.



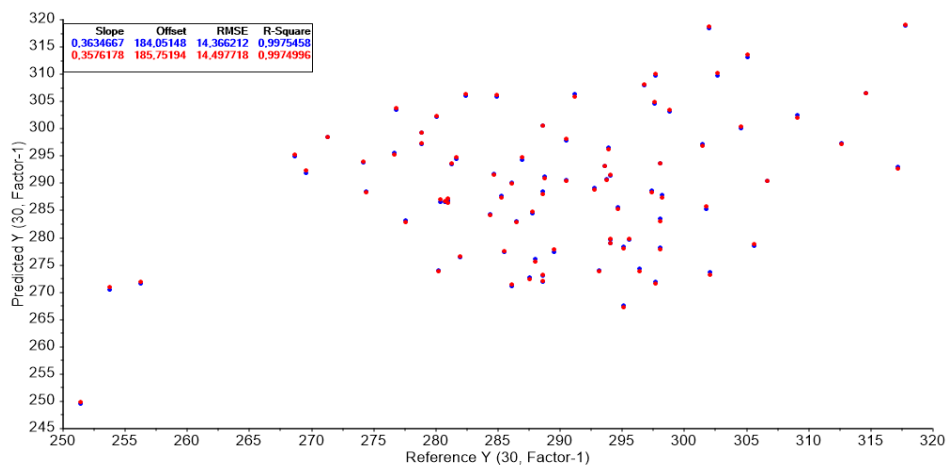
(a) T5



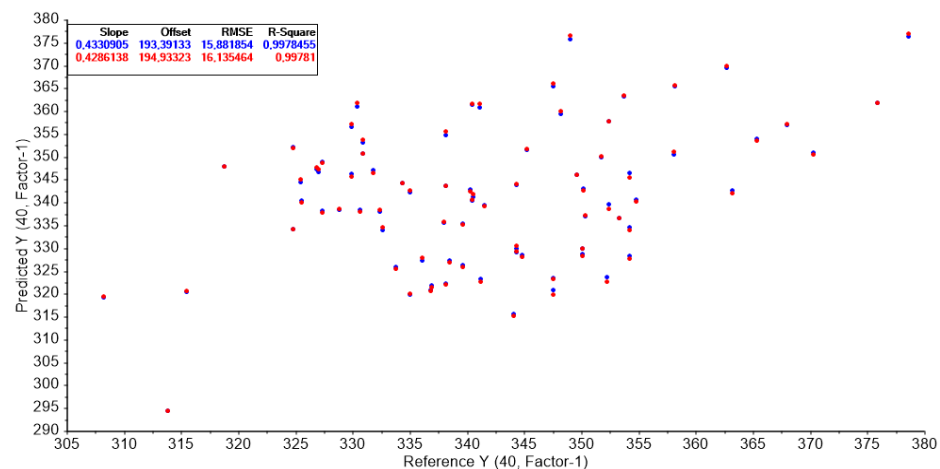
(b) T10



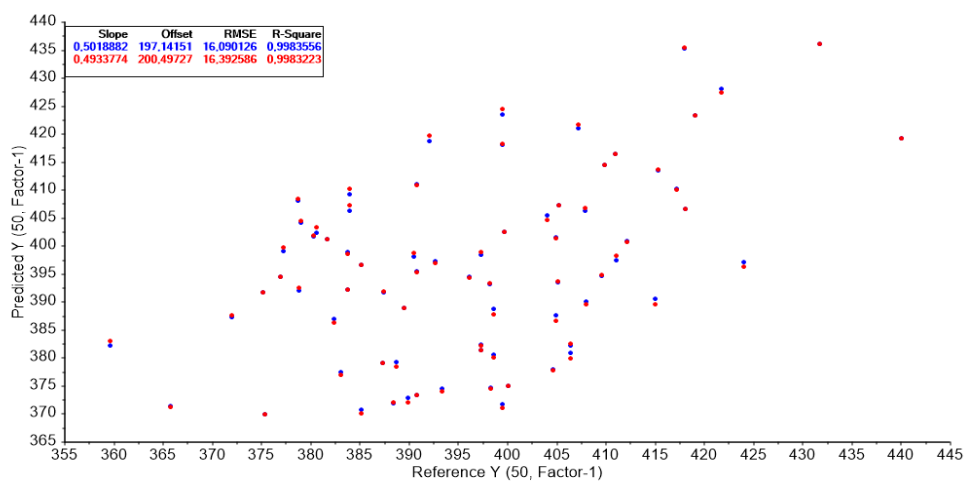
(c) T20



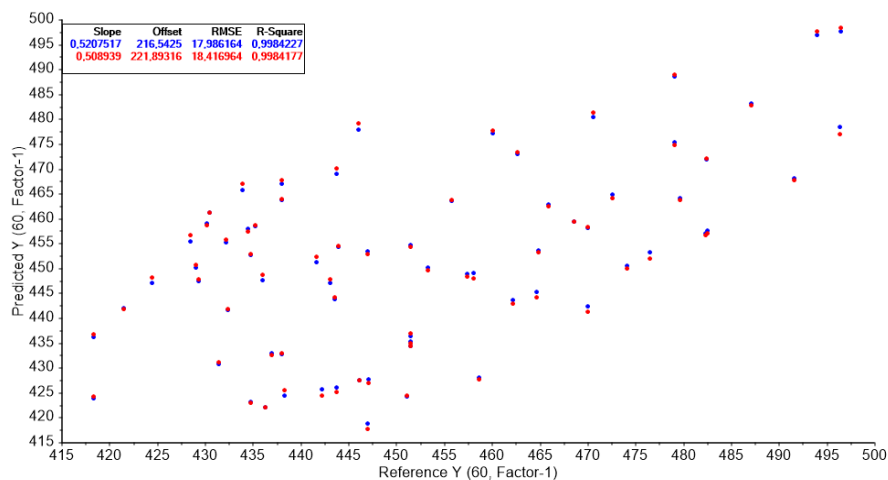
(d) T30



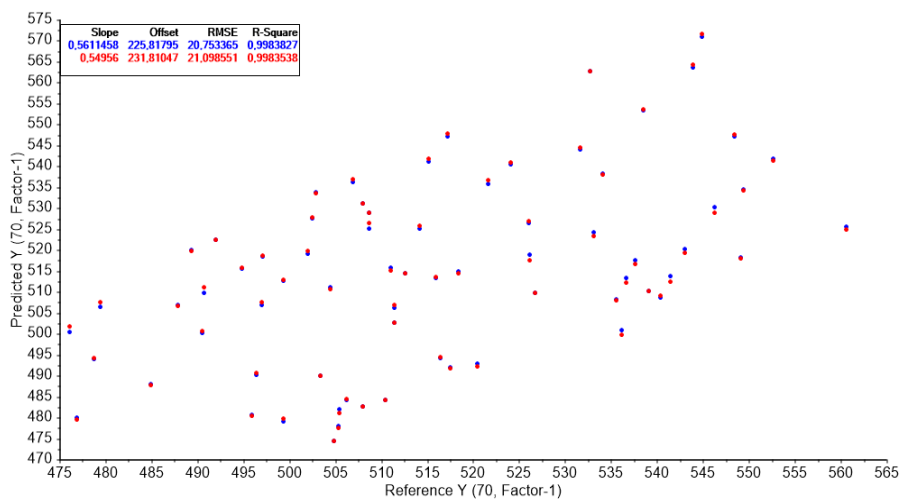
(e) T40



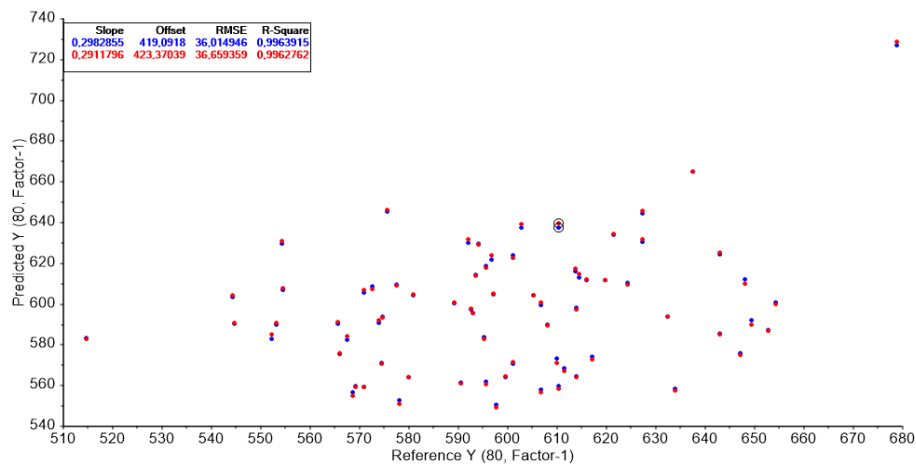
(f) T50



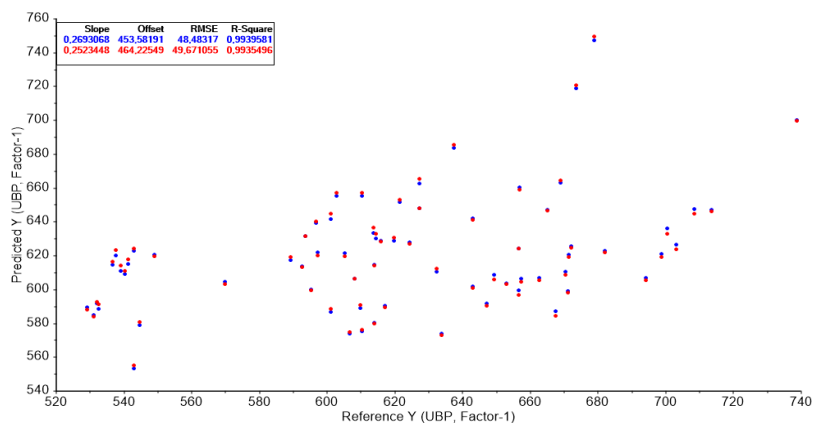
(g) T60



(h) T70

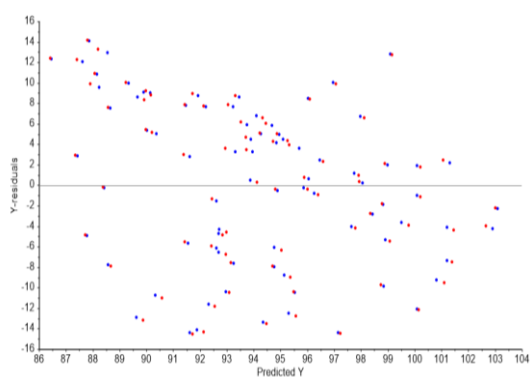


(i) T80

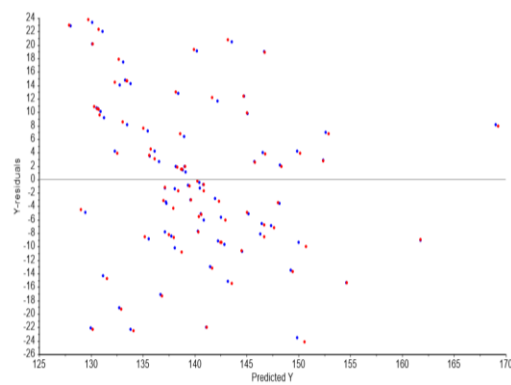


(j) UBP

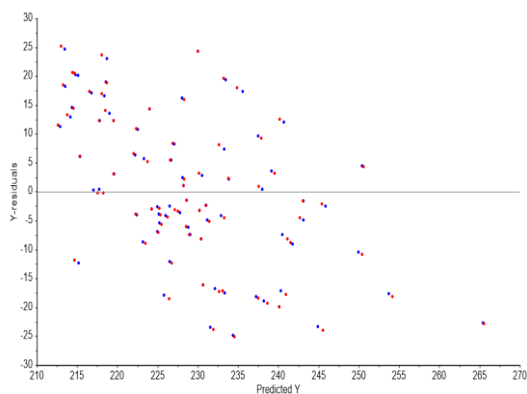
Figura 19. Gráficos de paridad de los resultados de las regresiones PLS en diferentes temperaturas de evaporado.



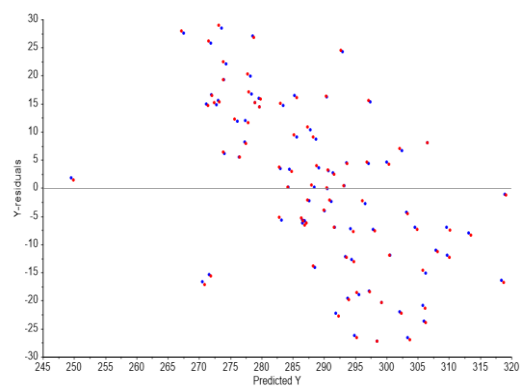
(a) T5



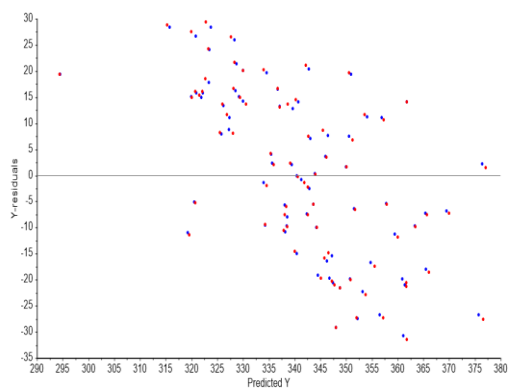
(b) T10



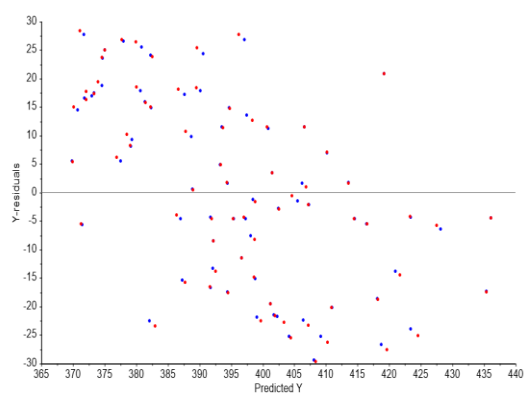
(c) T20



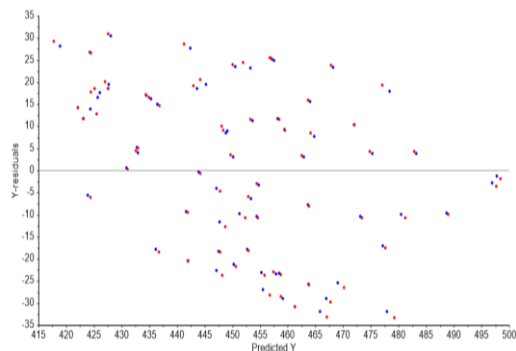
(d) T30



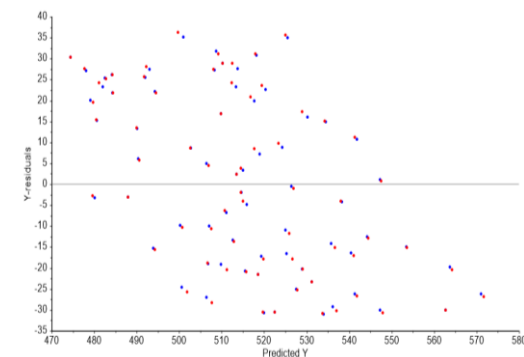
(e) T40



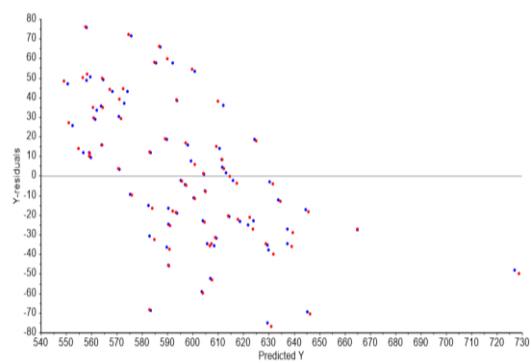
(f) T50



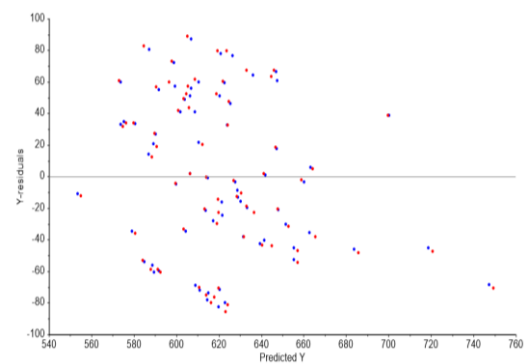
(g) T60



(h) T70

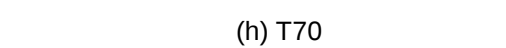
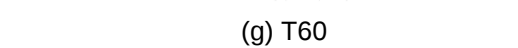
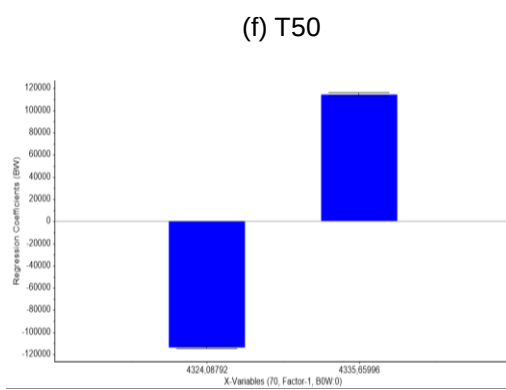
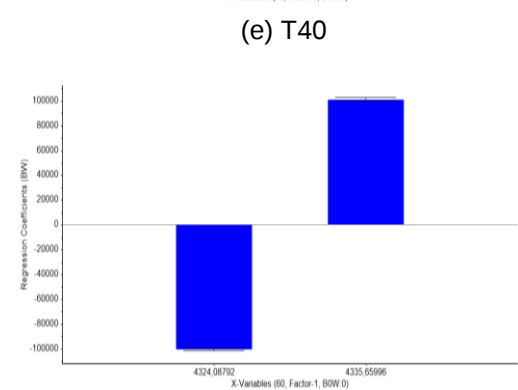
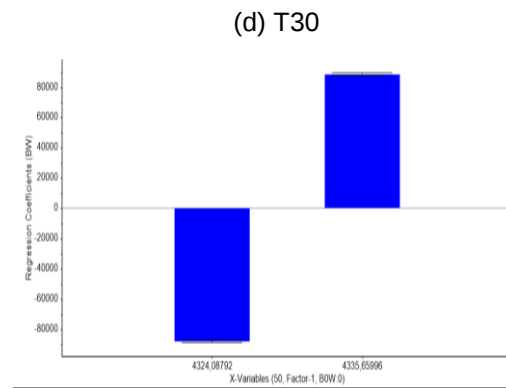
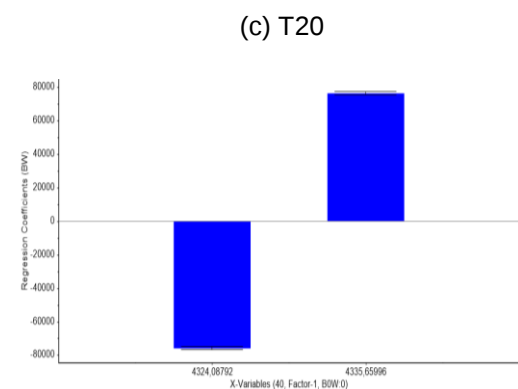
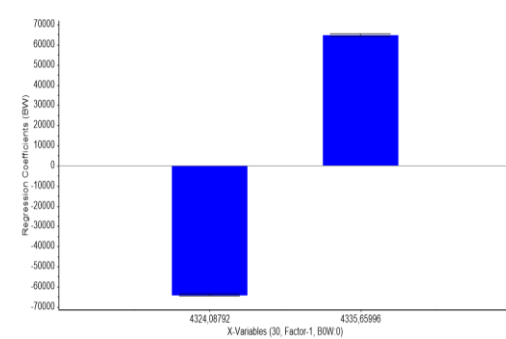
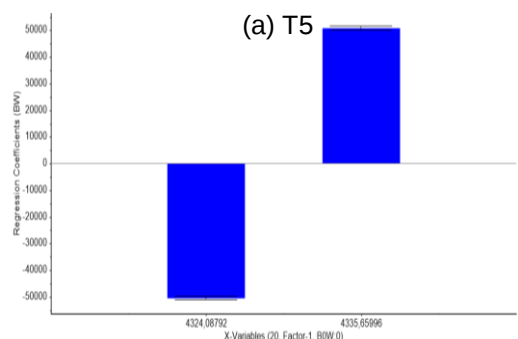
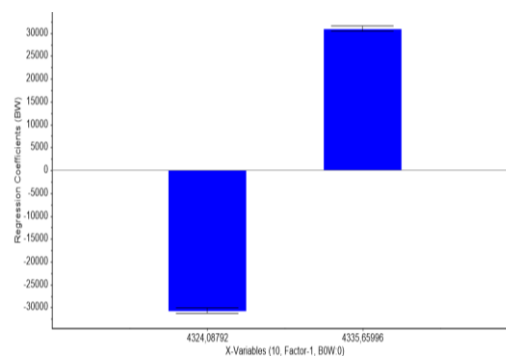
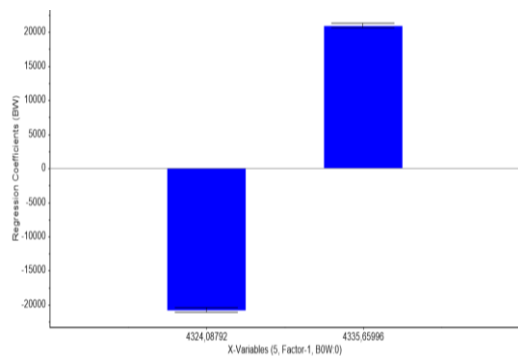


(i) T80



(j) UBP

Figura 20. Residuales obtenidos de las regresiones PLS en diferentes temperaturas de evaporado.



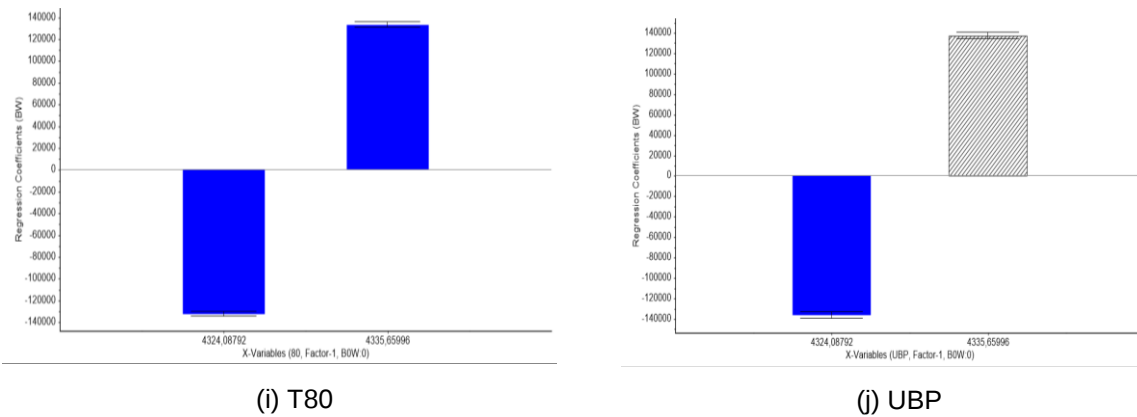


Figura 21. Coeficientes de regresión PLS en el ajuste de las diferentes temperaturas de evaporado

Tabla 6.

Modelos de regresión PLS finales para las temperaturas de evaporado de las muestras de la mezcla Vasconia.

REGRESIÓN	R^2_{val}	RMSE _{val} , °C	RMSE % _v al*	Intervalo 95%
$T5 = -20811,98 A_{4324} + 20909,68 A_{4335}$	0,993	7,80	8,28	± 14,60
$T10 = -30765,27 A_{4324} + 30909,1 A_{4335}$	0,994	11,00	8,07	± 22,00
$T20 = -50546,17 A_{4324} + 50805,96 A_{4335}$	0,997	13,04	5,89	± 26,08
$T30 = -64343,58 A_{4324} + 64740,55 A_{4335}$	0,997	14,50	5,08	± 29,00
$T40 = -75944,95 A_{4324} + 76396,73 A_{4335}$	0,998	16,14	4,76	± 32,28
$T50 = -87986,12 A_{4324} + 88516,18 A_{4335}$	0,998	16,40	4,17	± 32,80
$T60 = -100404 A_{4324} + 101045,9 A_{4335}$	0,998	18,42	4,10	± 36,84
$T70 = -113766,1 A_{4324} + 114385,7 A_{4335}$	0,998	21,10	4,11	± 42,20
$T80 = -132455,1 A_{4324} + 133252,6 A_{4335}$	0,994	36,66	6,13	± 73,32
$UBP = -136274 A_{4324} + 136896,1 A_{4335}$	0,993	49,67	7,87	± 99,34

* $\bar{T}5 = 94^\circ C$; $\bar{T}10 = 136^\circ C$; $\bar{T}20 = 221$; $\bar{T}30 = 285^\circ C$; $\bar{T}40 = 339^\circ C$; $\bar{T}50 = 393^\circ C$; $\bar{T}60 = 449^\circ C$; $\bar{T}70 = 514^\circ C$
 $\bar{T}80 = 598^\circ C$; $\bar{UBP} = 631^\circ C$

Según la Tabla 3.3, la incertidumbre para las temperaturas a partir de T20 supera los 50°C, lo cual disminuye la capacidad de predicción de las regresiones comparada con los datos

experimentales. Un intervalo de 50°C implica una diferencia de 7% en rendimiento volumétrico en promedio; asimismo un intervalo 74°C, mostrado por la regresión de T60, significa una incertidumbre promedio de 11% de rendimiento en la predicción. La regresión para UBP exhibe la mayor incertidumbre con un tamaño de intervalo de 200°C, con lo cual se tiene aproximadamente 22% de incertidumbre en rendimiento. El tamaño de los intervalos de incertidumbre muestra que la aplicación de las regresiones, a partir de T60, puede conducir a desviaciones importantes con consecuencias económicas para los procesos –ya sea por subestimación o sobreestimación de los flujos de mezcla o sus productos de destilación–. Por esto, las predicciones realizadas con las regresiones T60 en adelante corresponden a estimaciones semicuantitativas de la temperatura de evaporados; a pesar de estas incertidumbres, Pantoja *et al.*, 2009, recomiendan este nivel de estimación semicuantitativa en aplicaciones en línea para la caracterización de crudos en operaciones de refinación.

Los desempeños de las regresiones obtenidas en el presente trabajo son comparables con ajustes reportados en la literatura abierta. Duarte *et al.*, 2017, basados en el espectro RMN ^1H de crudos brasileiros, reportaron valores de ajuste PLS de la curva SimDis con RMSE% inferiores a 15%; los valores de RMSE% obtenidos por estos autores resultan, en algunos casos, inferiores al 5% especialmente entre T30 y T80. Lo encontrado por Duarte *et al.* coincide con lo obtenido en el presente trabajo.

Asimismo, Fakayode *et al.*, 2014, propusieron un modelo de regresión PLS para la curva SimDis de productos del petróleo, basada en descriptores de la actividad estructural, tales como densidad, punto de fusión, índices de refracción y tiempos de retención en la prueba de cromatografía de gases, entre otros. Estos autores aplicaron la regresión sobre mezclas de componentes orgánicos definidos, reportando errores RMSE% del orden de 10%. Es importante

mencionar que los resultados del presente trabajo se encuentran en el orden de magnitud de Fakayode *et al.*, con la diferencia de que estos últimos consideraron mezclas discontinuas de compuestos para simular crudos reales.

También, Moradi *et al.*, 2011, entrenaron y validaron una red neuronal utilizando como entradas el peso molecular del crudo y la gravedad específica del mismo para la predicción de la curva de destilación TBP en peso y la distribución de otras propiedades. Estos autores definieron una red de 12 neuronas internas con función de activación tansig, utilizando 579, 124 y 178 datos de crudos para las etapas de entrenamiento, validación y prueba, respectivamente; con esta red, Moradi *et al.*, reportaron un error promedio de 4,04% en la predicción de la curva de destilación de crudos livianos (SG promedio de 0,85). El error promedio obtenido en el presente trabajo corresponde al 8.4%, con la diferencia en que la mezcla Vasconia se caracteriza por ser un crudo intermedio con SG promedio de 0,92. La dificultad de trabajar con crudos intermedios corresponde a la obtención de un residuo no destilable, que debe ser tratado por destilación al vacío; esto puede generar una incertidumbre en el valor de los puntos de destilación SimDis (Espinosa-Peña *et al.*, 2004), lo cual justifica el mayor error relativo promedio en el presente trabajo.

Por otra parte, las predicciones pueden requerir el rendimiento en lugar de la temperatura; *i.e.* la predicción de los rendimientos acumulados a una temperatura definida, lo cual es presentado en la siguiente sección.

4.5.2 Rendimientos acumulados Los rendimientos acumulados volumétricos en las temperaturas a partir de 50°C, con incrementos de 50°C hasta 550°C –definidos como X50, X100, X150, X200, X250, X300, X350, X400, X450, X500 y X550– fueron calculados

utilizando interpolación por trazadores cúbicos sobre las curvas SimDis respectivas (Figura 3.4). Estos rendimientos acumulados a las temperaturas de referencia fueron ajustados utilizando regresión PLS sobre la matriz de espectros pretratados; asimismo, las regresiones obtenidas fueron depuradas por eliminación de variables independientes con baja contribución. Los resultados de las regresiones depuradas en cuanto a las gráficas de paridad, las gráficas de residuales y las gráficas de contribuciones de los números de onda son presentados en las Figuras 3.17, 3.18 y 3.19, respectivamente. Las gráficas de paridad (Figura 3.17) presentan coeficientes de determinación múltiple cercanos a la unidad, tanto en calibración como en validación; asimismo, los valores de RMSE presentan valores iniciales de 0,22% en rendimiento volumétrico, incrementando con el aumento en la temperatura; el valor máximo se presenta en 400°C con un RMSE de 4,3%.

Por otra parte, la Figura 3.18 exhibe las distribuciones de los residuales versus los valores de rendimiento volumétrico ajustados. Con esta figura es posible sugerir una distribución sin ninguna clase de tendencia, con lo cual el modelo es de utilidad en la predicción de los rendimientos. Asimismo, la Figura 3.19 ilustra los valores de los números de onda finales para los modelos PLS depurados; según esta figura –como se esperaba según lo obtenido en el numeral 3.5.1– los números de onda 4324 y 4335 cm^{-1} definen los valores para los rendimientos acumulados a cualquier temperatura.

La Tabla 3.4 resume los resultados obtenidos con la regresión PLS en el ajuste de los rendimientos acumulados a diferentes temperaturas. Según esta tabla, la incertidumbre en la predicción aumenta apreciablemente a partir de los 350°C (X350), con un tamaño de intervalo de 8,7% de rendimiento volumétrico, acorde con la incertidumbre obtenida en la sección anterior para la regresión de T60. Sin embargo, para las regresiones hasta X550 el tamaño del intervalo es

de ca. 16%, lo cual mejora la estimación realizada para T70, T80 y UBP de la sección anterior. A pesar de esta mejora, los resultados de la aplicación de las regresiones a partir de X400, deben ser asumidas como aproximaciones semicuantitativas y por consiguiente la aplicación debe considerar el análisis de escenarios de menor y mayor rendimiento acumulado.

Tabla 7.

Modelos de regresión PLS finales para los rendimientos acumulados volumétricos de la mezcla Vasconia a diferentes temperaturas.

REGRESIÓN	R ² _{val}	RMSE _{val} , %vol	RMSE %v al*	Interval o 95%
$X50 = -234,03 A_{4324} + 234,71 A_{4335}$	0,959	0,22	20,58	± 0,44
$X100 = -1185,38 A_{4324} + 1189,31 A_{4335}$	0,981	0,76	12,67	± 1,52
$X150 = -2432,58 A_{4324} + 2445,07 A_{4335}$	0,988	1,22	10,30	± 2,44
$X200 = -3702,06 A_{4324} + 3723,98 A_{4335}$	0,995	1,23	7,09	± 2,46
$X250 = -5211,05 A_{4324} + 5243,43 A_{4335}$	0,995	1,60	6,65	± 3,20
$X300 = -7169,95 A_{4324} + 7212,48 A_{4335}$	0,997	1,87	5,73	± 3,74
$X350 = -9389,02 A_{4324} + 9442,42 A_{4335}$	0,997	2,17	5,15	± 4,34
$X400 = -11512,75 A_{4324} + 11578,26 A_{4335}$	0,993	4,35	8,46	± 8,70
$X450 = -13439,29 A_{4324} + 13520,08 A_{4335}$	0,997	3,23	5,36	± 6,46
$X500 = -15201,88 A_{4324} + 15309,22 A_{4335}$	0,998	3,46	5,09	± 6,92
$X550 = -16490,56 A_{4324} + 16583,08 A_{4335}$	0,998	3,65	4,89	± 7,30

* $\bar{X}50 = 1,07$; $\bar{X}100 = 6,00$; $\bar{X}150 = 11,84$; $\bar{X}200 = 17,46$; $\bar{X}250 = 24,04$; $\bar{X}300 = 32,63$; $\bar{X}350 = 42,17$; $\bar{X}400 = 51,44$

$\bar{X}450 = 60,28$; $\bar{X}500 = 67,95$; $\bar{X}550 = 74,59$

Respecto a la comparación con reportes de literatura, Falla *et al.*, 2006, desarrollaron modelos de RNA para la predicción de la curva SimDis de crudos brasileiros con °API entre 31,1 y 36,4. Estos autores reportan valores de RMSE de predicción entre 0,66 y 3,5% en peso, utilizando una red con 6 neuronas internas. En el presente trabajo, la mezcla Vasconia reporta valores de °API

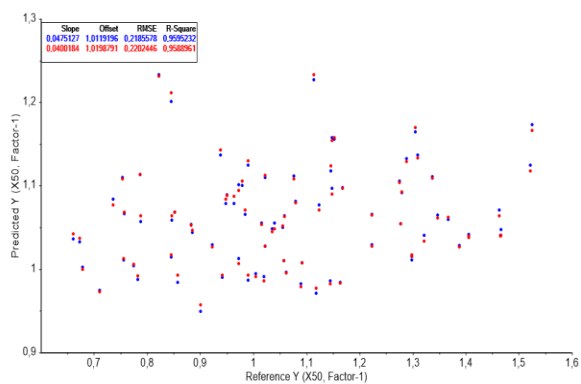
entre 20,1 y 27,9 (Tabla 3.1); a pesar de esto, las regresiones PLS para el rendimiento acumulado de la mezcla Vasconia obtuvo desempeños de RMSE entre 0,22 y 4,35%, similares a los reportados por Falla *et al.* Una mejora de la regresión PLS aplicada en el presente documento corresponde al cumplimiento en la monotonicidad creciente de la curva SimDis –siempre con un rendimiento acumulado creciente–; Falla *et al.* reportan el incumplimiento de este requisito por parte de las RNA aplicadas.

Asimismo, Hidajat & Chong, 2000, aplicaron regresiones PLS basadas en espectros NIR para el ajuste de propiedades de crudos con SG entre 0,80 y 0,90. Los valores de RMSE reportados por estos autores se encuentran en el rango 0,66–1,25 %vol, utilizando hasta 10 variables latentes en sus modelos. Los valores de RMSE reportados en el presente trabajo se encuentran en promedio, por encima del rango reportado por Hidajat & Chong; sin embargo, el número de variables latentes y la cantidad de números de onda utilizados en el presente documento son menores que los obtenidos por Hidajat & Chong. Con esto, el valor máximo de RMSE de 4,35% obtenido en el presente trabajo es justificado con base en su simplicidad y facilidad de aplicación.

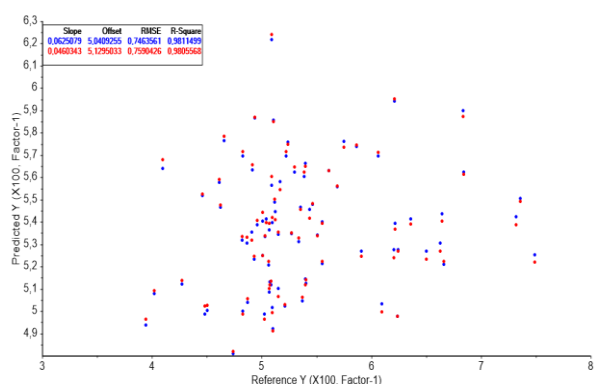
4.5.3 Composiciones aromática, nafténica y parafínica ~De igual forma, las composición de las diferentes mezclas Vazconia representadas por sus componentes aromático, nafténico y parafínico fueron ajustados, utilizando modelos de regresión PLS –como una especie de metamodelo de las ecuaciones de Riazi a partir del espectro NIR–. Al igual que para las temperaturas de evaporado, los modelos PLS fueron depurados, eliminando números de onda con contribuciones nulas. Las Figuras 3.20–3.22 presentan los resultados de los modelos PLS para los componentes, en las gráficas de paridad, de residuales y de contribución de las variables

independientes, respectivamente. Según la Figura 3.20, los gráficos de paridad presentan mayor dispersión para las composiciones parafínica y aromática, exhibiendo mayor valor de RMSE. Los residuales mostrados en la Figura 3.21 sugieren un sesgo para el modelo de la composición parafínica; sin embargo, al igual que para las temperaturas de evaporado, la dispersión de los residuales sugiere que una mayor cantidad de muestras conllevaría a la obtención de una ubicación aleatoria.

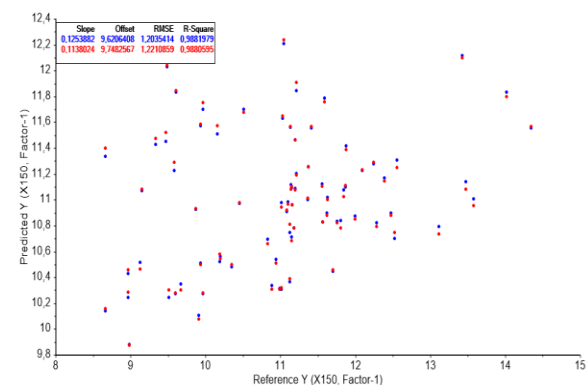
Por otro lado, las variables independientes de los modelos finales coinciden con los modelos para las temperaturas de evaporados (Tabla 3.3) y los rendimientos acumulados (Tabla 3.4), en que los aportes en los números de onda 4324 y 4335 cm^{-1} definen las curvas SimDis y las composiciones de las muestras de mezcla Vasconia (Tabla 3.5).



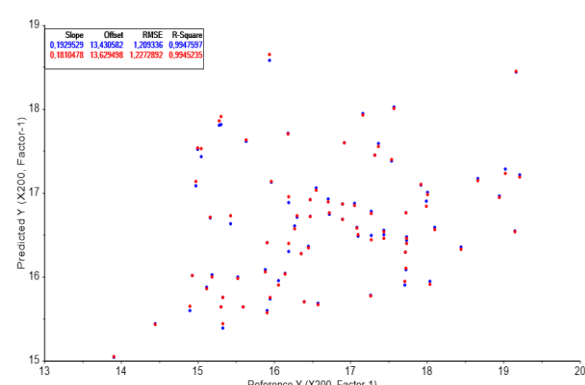
(a) X50



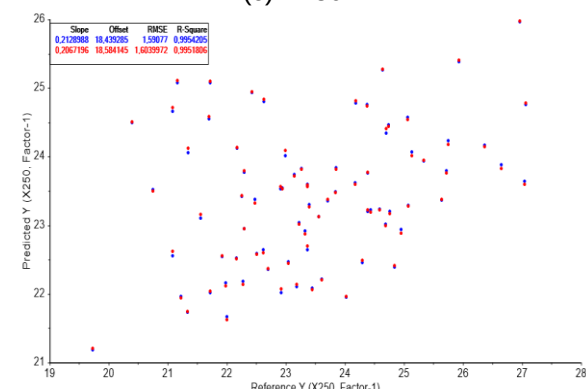
(b) X100



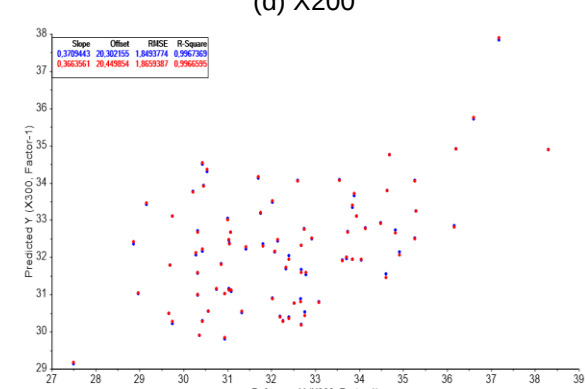
(c) X150



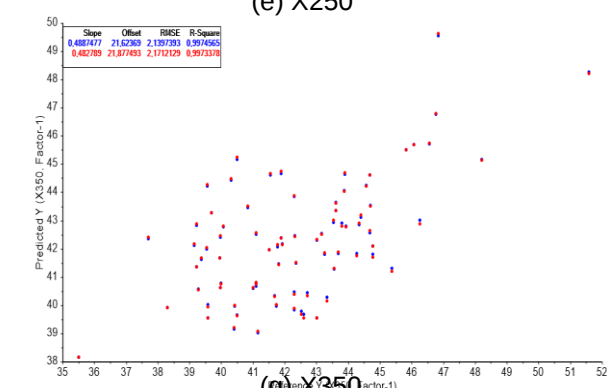
(d) X200



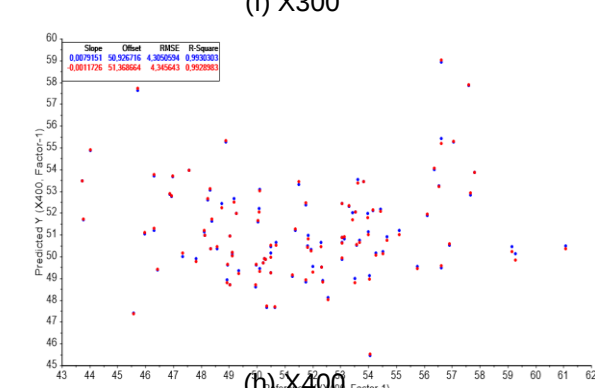
(e) X250



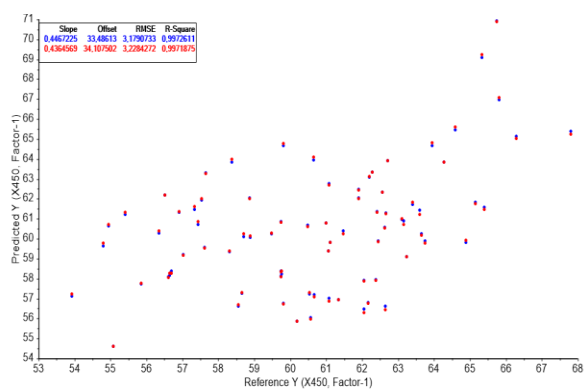
(f) X300



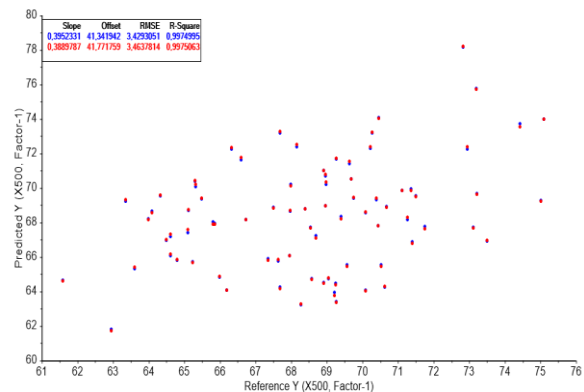
(g) X350



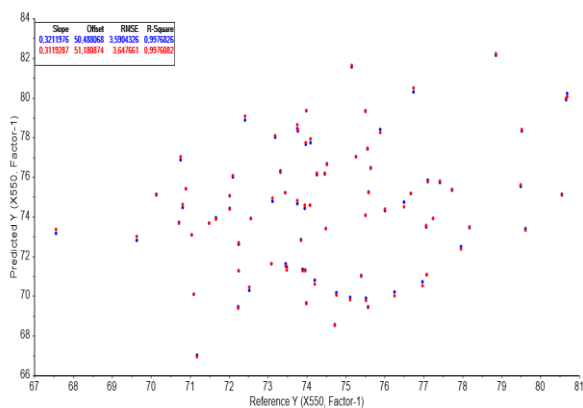
(h) X400



(i) X450

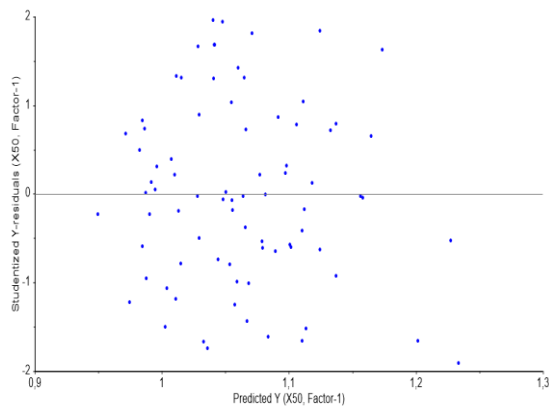


(j) X500

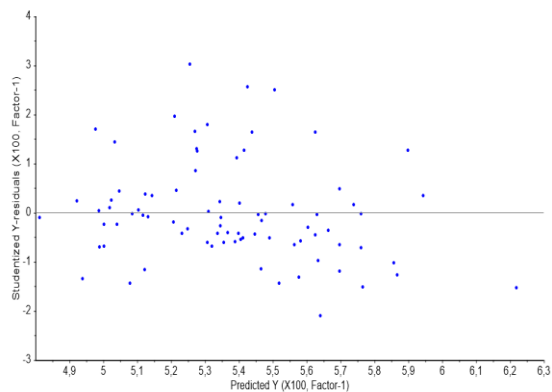


(k) X550

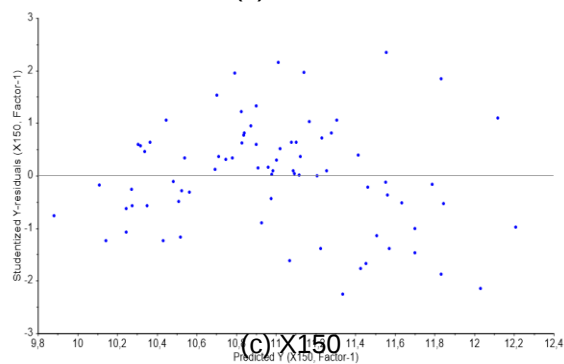
Figura 22. Gráficas de paridad regresión PLS de los rendimientos acumulados



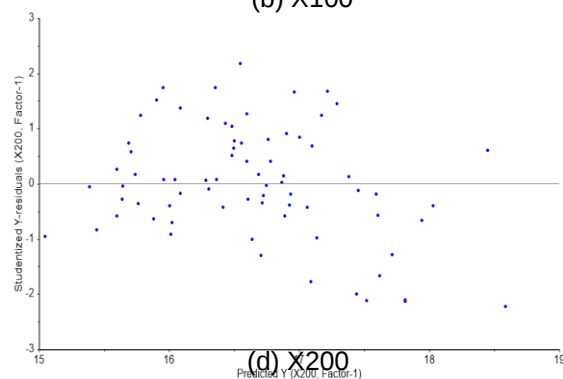
(a) X50



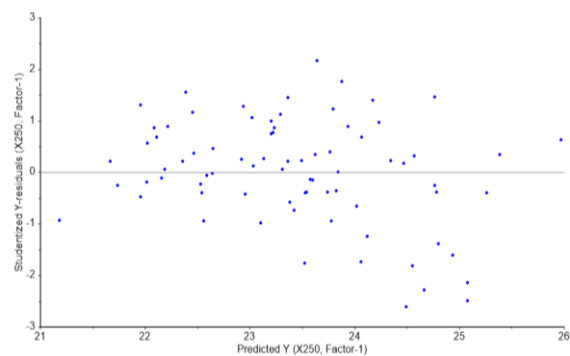
(b) X100



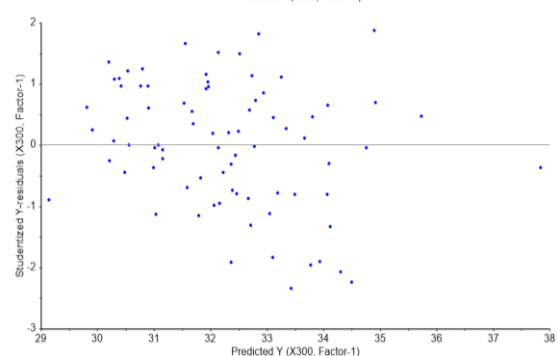
(c) X150



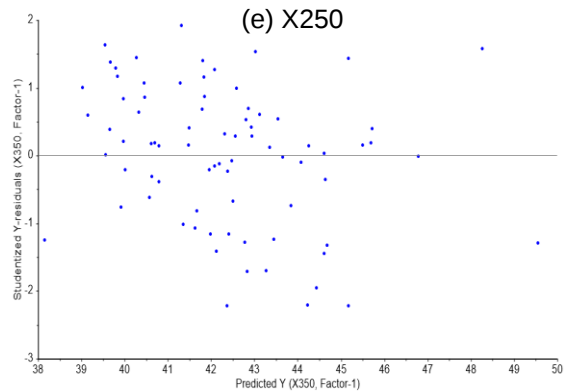
(d) X200



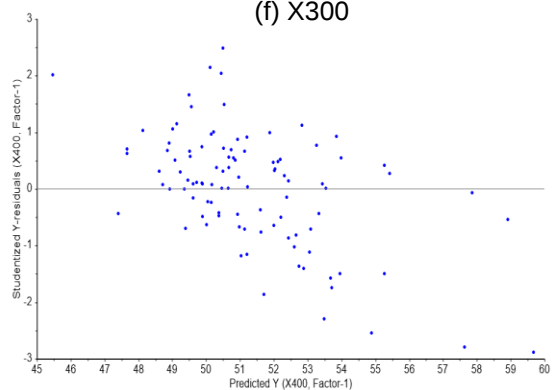
(e) X250



(f) X300



(g) X350



(h) X400

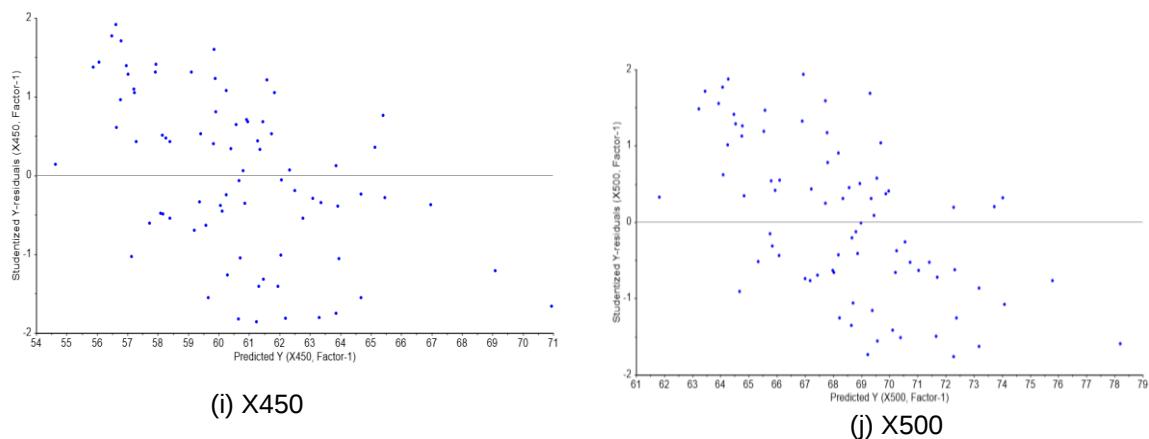
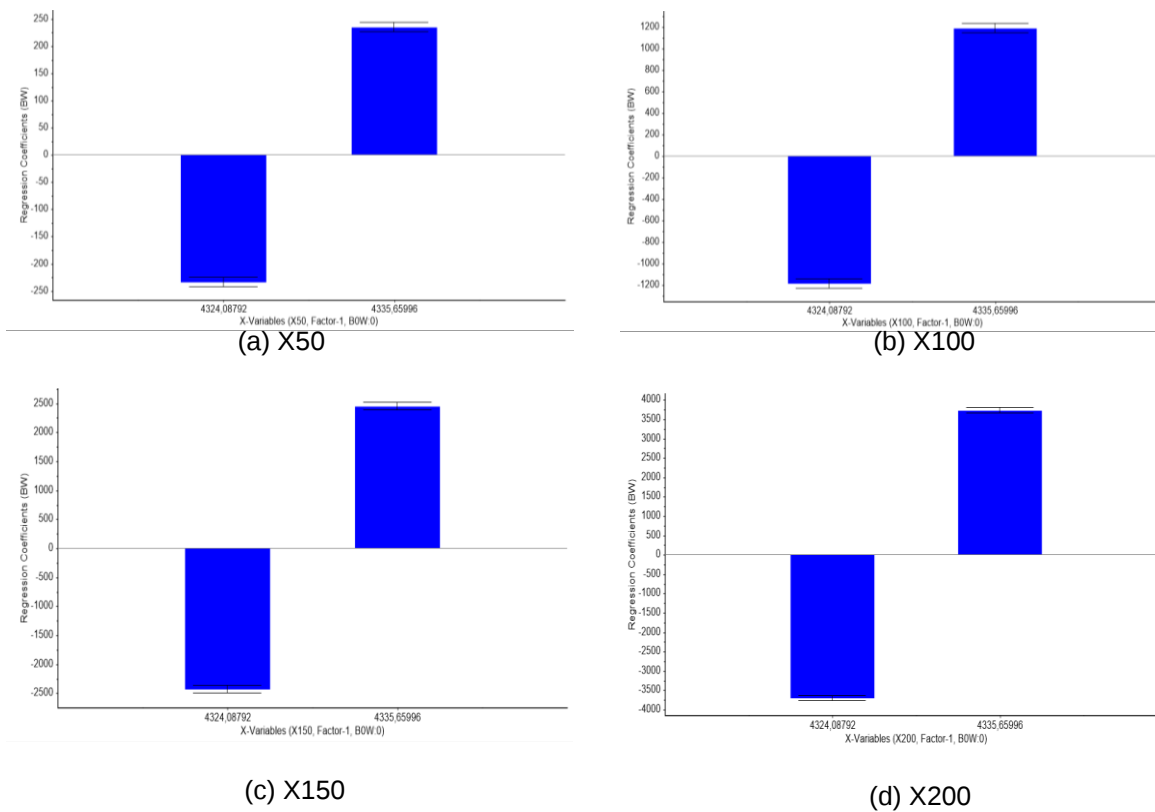
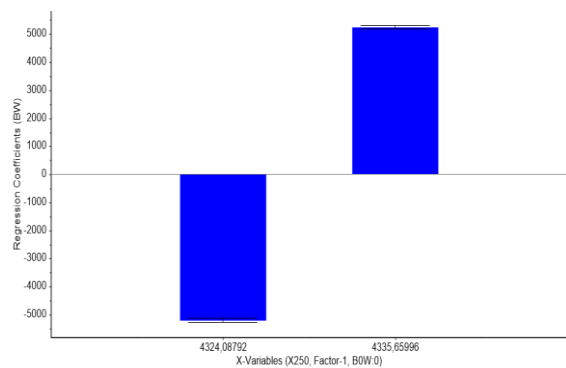
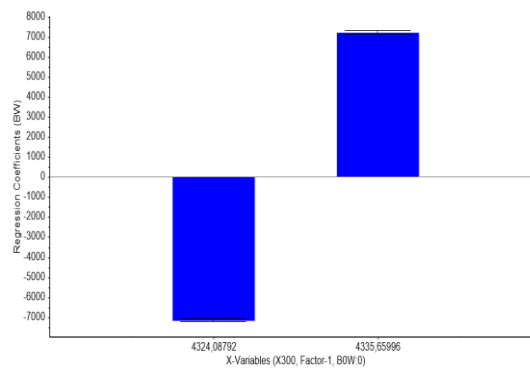


Figura 23. Residuales regresión PLS de los rendimientos acumulados

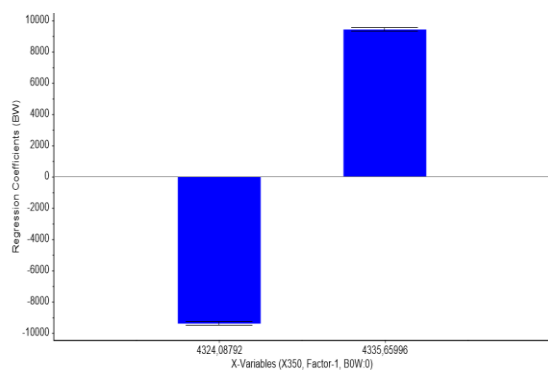




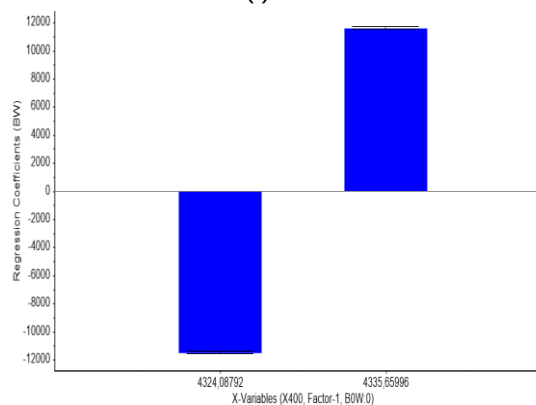
(e) X250



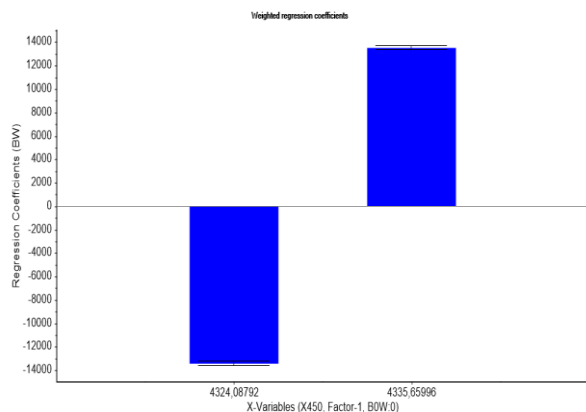
(f) X300



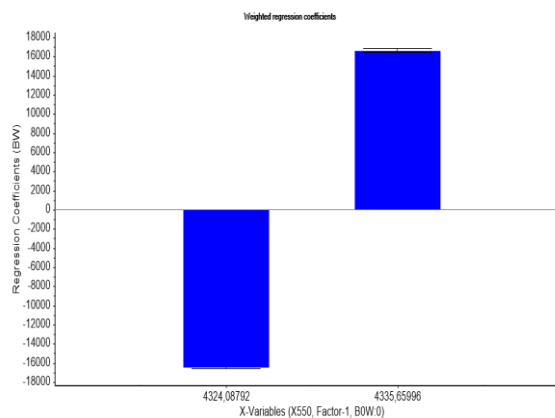
(g) X350



(h) X400

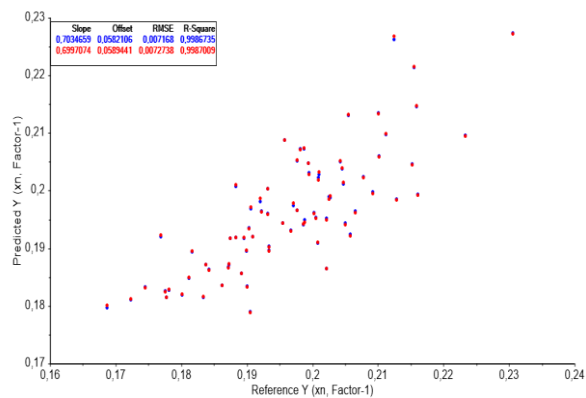


(i) X450

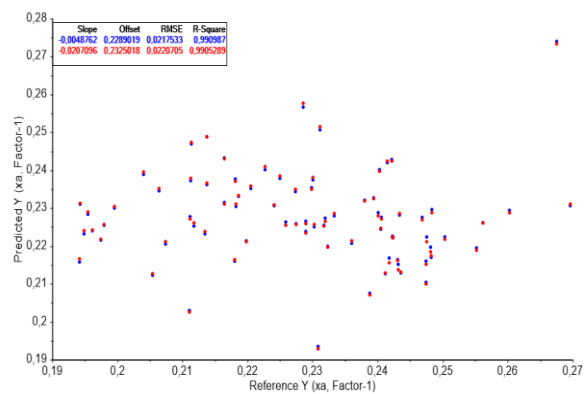


(j) X550

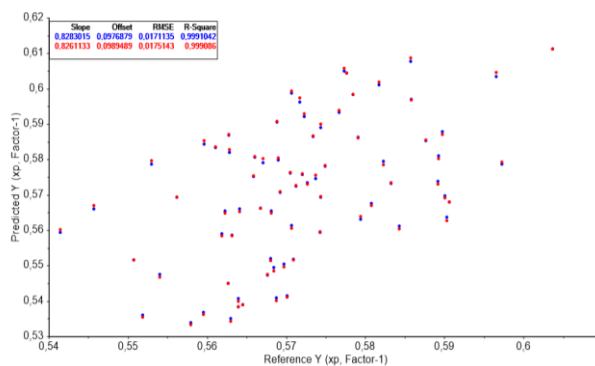
Figura 24. Contribuciones números de onda regresión PLS de los rendimientos acumulados



(a) Fracción nafténos, x_n

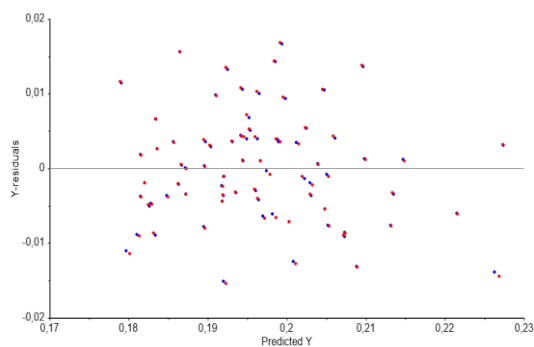


(b) Fracción aromáticos, x_a

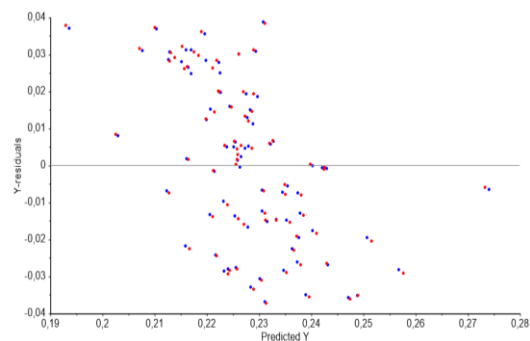


(c) Fracción parafinas, x_p

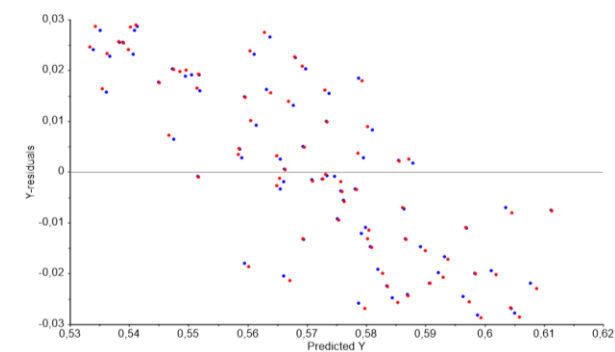
Figura 25. Gráficos de paridad regresión PLS en el ajuste de las fracciones volumétricas de nafténos, aromáticos y parafinas.



(a) Fracción nafténos, x_n



(b) Fracción aromáticos, x_a



(c) Fracción parafinas, x_p

Figura 26. Residuales regresión PLS en el ajuste de las fracciones volumétricas de naftenos, aromáticos y parafinas.

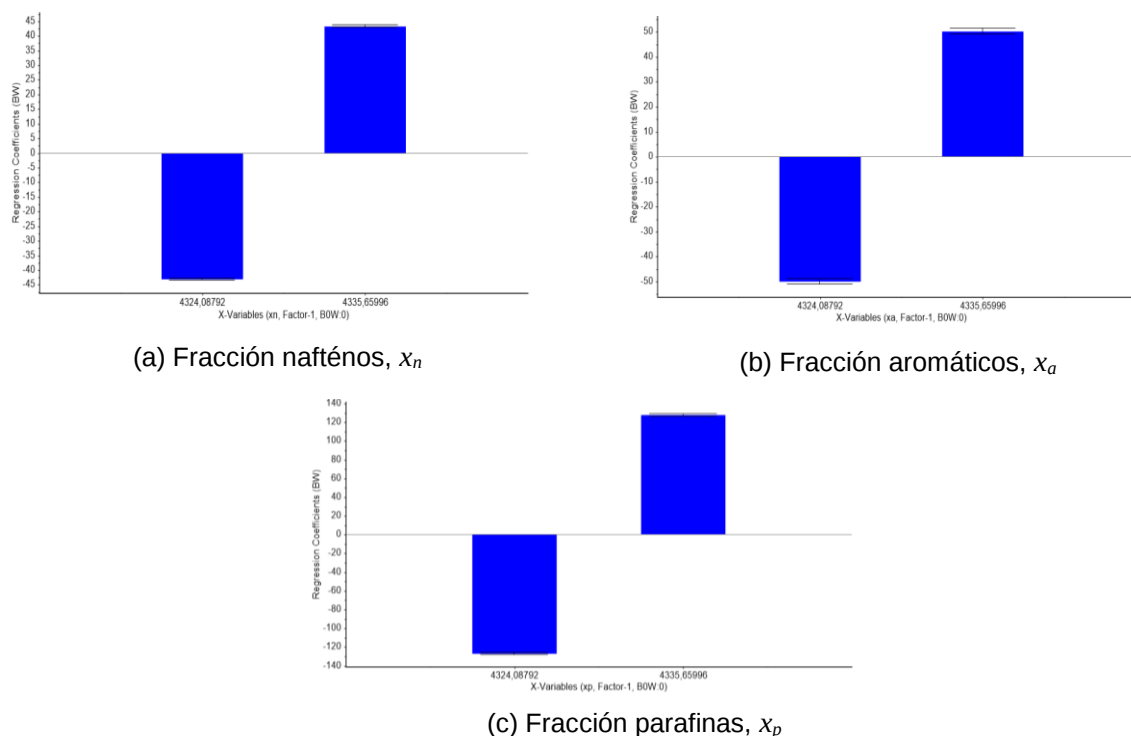


Figura 27. Coeficientes regresión PLS en el ajuste de las fracciones volumétricas de naftenos, aromáticos y parafinas. Fuente: Autora.

Tabla 8.

Modelos de regresión PLS finales para las composiciones de las muestras de la mezcla Vasconia.

REGRESIÓN	R^2_{val}	RMSE _{val}	RMSE % _{val}	Interval
		fracción	*	o 95%
$x_n = -43,1228 A_{4324} + 43,2829 A_{4335}$	0,999	0,007	3,60	$\pm 0,014$
$x_a = -49,9875 A_{4324} + 50,1701 A_{4335}$	0,991	0,022	3,60	$\pm 0,044$
$x_p = -127,0617 A_{4324} + 127,7303 A_{4335}$	0,999	0,018	7,80	$\pm 0,036$

* $\bar{x}_n = 0,196$; $\bar{x}_p = 0,573$; $\bar{x}_a = 0,232$

Comparando los desempeños de las regresiones del presente documento, Hannisdal *et al.*, 2005, aplicaron regresiones PLS en el ajuste de los contenidos de saturados, aromáticos, resinas

y asfaltenos (SARA) de 20 crudos de origen Europeo, Latinoamericano y Oriental; los modelos PLS consideraron espectros MIR. Hannisdal *et al.* reportaron valores de RMSE de predicción de 2,82%, 1,47%, 1,46% y 0,44% en peso para las fracciones SARA, respectivamente. En el presente trabajo se obtuvieron los valores de 0,7% en volumen para los naftenos y 1,8% en volumen para el RMSE de las parafinas, los cuales son comparables a los reportados por Hannisdal *et al.* para la fracción de saturados; considerando la densidad se tendría un menor valor para los RMSE respectivos en peso. De igual forma, los resultados de RMSE del ajuste PLS de aromáticos (2,2%) resultan comparables con los valores de Hannisdal *et al.* para las fracciones de aromáticos y resinas.

Por otra parte, Lexalde *et al.*, 2014, aplicaron un procedimiento de selección de variables sobre el espectro NIR, basado en algoritmos genéticos en regresiones PLS para los contenidos SARA de 230 muestras entre fondos de vacío y residuo atmosférico. Estos autores reportan valores de RMSE de predicción de 1,51%, 1,59%, 0,77% y 1,26% en peso para los componentes SARA. Estos valores resultan comparables con los RMSE de validación obtenidos en el presente trabajo (Tabla 3.5). Asimismo, los desempeños de las regresiones de la Tabla 3.5 resultan comparables con los reportes de otros autores como Aske *et al.*, 2001, Abbas *et al.*, 2012, Meléndez *et al.*, 2012 y Satya *et al.*, 2007. Por lo anterior, los modelos propuestos en el presente documento son de utilidad práctica para estimaciones de la fracción de aromáticos, naftenos y parafinas en la mezcla Vasconia.

5. Conclusiones

- Las propiedades medidas para las muestras indican que las proporciones de los crudos constituyentes de la mezcla Vasconia varían periódicamente. Los valores calculados para K_w y VGC indican que la mezcla Vasconia es un crudo con preeminencia nafténica. Con esto, a pesar de las fluctuaciones, la mezcla Vasconia se encuentra constituida por componentes similares, principalmente el crudo Rubiales, durante el tiempo de muestreo. Los resultados del algoritmo *kmean* definiendo 3 agrupaciones –que representan a las fracciones parafínica, nafténica y aromática– sugieren que 20 muestras presentan un aumento importante en su fracción parafínica, mientras que 41 muestras presentan un aumento importante en su fracción aromática. Las tres agrupaciones aparecen con límites definidos en el eje de la propiedad Azufre; los valores de azufre corresponden a [6307 – 9300], [9300 – 10800] y [10800 – 12844] ppm.
- Según el análisis PCA, los valores de PC1 para los *loading* en los números de onda 3843 y 3849 cm^{-1} indican las fluctuaciones en composición de la mezcla Vasconia; *i.e.* espectros con valores elevados en la segunda derivada evaluada en el número de onda 3843 cm^{-1} reflejan un aumento en el componente aromático, mientras que valores extremos negativos en la segunda derivada en 3849 cm^{-1} muestran un aumento en el componente parafínico de la mezcla Vasconia. Asimismo, los valores de PC2 para los *loading* en los números de onda 4324 y 4335 cm^{-1} son característicos de las temperaturas de evaporado; el *loading* PC2 en 4324 cm^{-1} representa el menor valor de temperaturas de evaporados,

mientras que *loading* PC2 en 4335 cm^{-1} representa el mayor valor entre las curvas SimDis. Con lo anterior, las segundas derivadas en estos números de onda sugieren la cantidad de compuestos volátiles contenidos en la mezcla Vasconia.

- El método kmeans considerando los espectros FT-NIR pretratados define agrupaciones sólo en el espacio de los componentes principales. La carencia en un patrón en las propiedades basada en la matriz de espectros FT-NIR manifiesta la variación en las composiciones de los crudos que conforman la mezcla Vasconia. Considerando la naturaleza nafténica de la mezcla, las variaciones se presentan en los crudos minoritarios; estas variaciones no afectan en gran medida las tendencias en las propiedades globales mostradas por las mezclas Vasconia, pero si las tendencias en sus espectros. Con esto, las agrupaciones considerando las propiedades macroscópicas corresponden a un indicador de la contribución de cada componente en la mezcla Vasconia.
- Las regresiones PLS finales para las temperaturas de evaporados, los rendimientos acumulados y las fracciones de los diferentes componentes se encuentran definidas por las absorbancias pretratadas en los números de onda 4324 y 4335 cm^{-1} . Los modelos PLS para las temperaturas de evaporados exhiben valores de RMSE% de validación entre 4,1 y 8,29%. Las regresiones PLS para los rendimientos acumulados reportan valores de RMSE% entre 4,89 y 20,58%. Por su parte, las regresiones PLS para las fracciones de aromáticos, naftenos y parafinas presentan valores de RMSE de validación inferiores a 2,2%. Los anteriores desempeños de las regresiones obtenidas en el presente trabajo son comparables con ajustes reportados en la literatura abierta. Una ventaja de los modelos PLS obtenidos en el presente documento corresponde a su simplicidad, lo cual los hace de utilidad en aplicaciones en línea.

6. Recomendaciones

- Continuar con el desarrollo de modelos de regresión para las propiedades de los crudos que recibe la GRB.

Referencias Bibliográficas

- Abbas, O., Rebufa, C., Dupuy, N., Permanyer, A., Kyster, J. 2012. PLS regression on spectroscopic data for the prediction of crude oil quality: API gravity and aliphatic/aromatic ratio. *Fuel*, 98, 5-14.
- Argirov, G., Ivanov, S. & Cholakov, G. 2012. Estimation of crude oil TBP from crude viscosity. *Fuel*, 97, 358–365.
- Aske, N., Kallevik, H., Sjöblom, J. 2001. Determination of saturate, aromatic, resin and asphaltenic (SARA) components in crude oils by means of infrared and near-infrared spectroscopy. *Energy Fuel*, 15, 1304-1312.
- Balabin, R. M., Safieva, R. Z. & Lomakina, E. I. 2010. Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. *Anal. Chim. Acta*, 671, 27–35.
- Bi, Y.M., Xie, Q., Peng, S.L., Tang, L., Hu, Y., Tan, J., Zhao, Y.H., Li, C.W. 2013. Dual stacked partial least squares for analysis of near-infrared spectra. *Anal. Chim. Acta*, 792, 19–27.
- Blanco, M., MasPOCH, S., Villarroya, I., Peralta, X., González, J. M., Torres, J. 2001. Determination of physico-chemical parameters for bitumens using near infrared spectroscopy. *Anal. Chim. Acta*, 434, 133–141.
- Brouillette, C., Smith, W., Shende, C., Gladding, Z., Farquharson, S., Morris, R.E., Cramer, J.A., Schmitgal, J. 2016. Analysis of Twenty-Two Performance Properties of Diesel, Gasoline, and Jet Fuels Using a Field-Portable Near-Infrared (NIR) Analyzer. *Applied Spectroscopy*, 70(5), 746–755.
- Chu, X., Xu, Y., Tian, S., Wang, J., Lu, W. 2011. Rapid identification and assay of crude oils based on moving-window correlation coefficient and near infrared spectral library. *Chemom. Intell. Lab. Syst.*, 107, 44–49.

- Duarte, L. M., Filgueiras, P. R., Dias, J. C. M., Oliveira, L. M. S. L., Castro, E. V. R., de Oliveira, M. A. L. 2017. Study of Distillation Temperature Curves from Brazilian Crude Oil by 1H Nuclear Magnetic Resonance Spectroscopy in Association with Partial Least Squares Regression. *Energy & Fuels*, 31, 3892–3897.
- Echeverria, O.D. 2016. *Análisis y correlación del contenido de biocombustibles, punto de chispa y T95 para el diésel utilizando espectroscopia infrarroja y métodos quimiométricos* (tesis de maestría). UIS, Barrancabermeja, Colombia.
- Espinosa-Peña, M., Figueroa-Gómez, Y., Jiménez-Cruz, F. 2004. Simulated Distillation Yield Curves in Heavy Crude Oils: A Comparison of Precision between ASTM D-5307 and ASTM D-2892 Physical Distillation. *Energy & Fuels*, 18(6), 1832–1840.
- Fakayode, S. O., Mitchell, B. S., & Pollard, D. A. 2014. Determination of boiling point of petrochemicals by gas chromatography–mass spectrometry and multivariate regression analysis of structural activity relationship. *Talanta*, 126, 151–156.
- Falla, F. S., Larini, C., Le Roux, G. A. C., Quina, F. H., Moro, L. F. L. y Nascimento, C. A. O. 2006. Characterization of crude petroleum by NIR. *Journal of Petroleum Science and Engineering*, 51, 127–137.
- Hannisdal, A., Hemmingsen, P. V., & Sjöblom, J. 2005. Group-Type Analysis of Heavy Crude Oils Using Vibrational Spectroscopy in Combination with Multivariate Analysis. *Industrial & Engineering Chemistry Research*, 44, 1349–1357.
- Hidajata, K. & Chong, S. M. 2000. Quality characterisation of crude oils by partial least square calibration of NIR spectral profiles. *J. Near Infrared Spectrosc.*, 8, 53–59.
- Li, B., Morris, J., Martin, E.B. 2002. Model selection for partial least squares regression. *Chem. Int. Lab. Syst.*, 64, 79-89.
- Lexalde, J., Ruckebusch, C., Devos, O, Caillol, N., Wahl, F., Duponchel, L. 2011. Characterisation of heavy oils using near-infrared spectroscopy: Optimisation of pre-processing methods and variable selection. *Anal. Chim. Acta*, 705, 227–234.
- Lexalde, J., Caillol, N., Wahl, F., Ruckebusch, C., Duponchel, L. 2014. Combining near and mid infrared spectroscopy for heavy oil characterisation. *Fuel*, 133, 310-316.

- Meléndez, L.V., Lache, A., Orrego-Ruiz, J.A., Pachón, Z. & Mejía-Ospino, E. 2012. Prediction of the SARA analysis of Colombian crude oils using ATR – FTIR spectroscopy and chemometric methods. *Journal of Petroleum Science and Engineering*, 91, 56–60.
- Moradi, G.R., Khoshmaram, A.A., Riazi, M.R. 2011. Estimation of properties distribution of C7+ by using artificial neural networks. *Journal of Petroleum Science and Engineering*, 76, 57-62.
- Mullins, O.C. 1990. Asphaltenes in crude oil: absorbers and/or scatterers in the near-infrared region? *Anal. Chem.*, 62, 508–514.
- Navarro, U. 2005. *La composición química y las propiedades del petróleo crudo y sus fracciones*. Barrancabermeja, Colombia.
- Pabón, R.E.C & Filho, C.R. De S. 2019. Crude oil spectral signatures and empirical models to derive API gravity. *Fuel*, 237, 1119-1131.
- Pantoja, P.A., Sotelo, F.F., Gatti, A.C., Ciriaco, M.F., Katata, A.C., Roux, G.A.C.L., Nascimento, C.A.O. 2009. Improving Accuracy of Refinery Optimization by the On-Line Characterization of Crude Oil. *10th International Symposium on Process Systems Engineering: Part A*, 339–344.
- Pantoja, P.A., López-Gejo, J., Nascimento, C.A.O., Roux, G.A.C.L. 2018. Application of Near-Infrared Spectroscopy to the Characterization of Petroleum. In: *Analytical Characterization Methods for Crude Oil and Related Products*. Wiley, 221-239.
- Pasquini, C. & Bueno, A.F. 2007. Characterization of petroleum using near-infrared spectroscopy: Quantitative modeling for the true boiling point curve and specific gravity. *Fuel*, 86, 1927–34.
- Riazi, M.R. 2005. *Characterization and Properties of Petroleum Fractions*. Philadelphia, USA: American Society for Testing and Materials.
- Riazi, M.R. & Daubert, T.E. 1986. Prediction of molecular-type analysis of petroleum fractions and coal liquids. *Ind. Eng. Chem. Process Des. Dev.*, 25, 1009-1015.

- Rinnan, A. 2014. Pre-processing in vibrational spectroscopy – when, why and how. *Anal. Methods*, 6, 7124–7129.
- Satya, S., Roehner, R. M., Deo, M. D., & Hanson, F. V. 2007. Estimation of Properties of Crude Oil Residual Fractions Using Chemometrics. *Energy & Fuels*, 21(2), 998–1005.
- UOP. 1978. *Train Your Refinery Operators: TYRO*. Illinois, USA: UOP Inc.
- Wauquier, J.P. 1995. *Petroleum Refining VI: Crude Oil, Petroleum Products, Process Flowsheets*. París, Francia: Editions Technip.
- Zhi-Na, X., Jiu-Xiang, W. y Jun, Q. 2012. Rapid Measurement of Diesel Engine Oil Quality By Near Infrared Spectroscopy (NIRS). *The Open Petroleum Engineering*, 2, 58–62.

Apéndices

Apéndice A. Curvas SimDis mezcla Vasconia. Resultados, regresiones y residuales

[illegible]