

Desarrollo de un Modelo Predictivo para la Determinación de Estabilidad en Crudos
Colombianos a partir de FTIR y PLS

Alexander Jair Rey Niño

Trabajo de Grado para Optar el título de Químico

Director

Enrique Mejía Ospino

Dr. En Ciencias Químicas

Codirector

Diana Catalina Palacio

Dra. En Ciencias Naturales (Física)

Universidad Industrial de Santander

Facultad de Ciencias

Escuela de Química

Bucaramanga

2020

Dedicatoria

“Educa a los niños y no será necesario castigar a los hombres”

- Pitágoras

A mi familia, cuyo apoyo y motivación constante, me demostró por medio de sus actos el valor de la perseverancia, a afrontar las pequeñas dificultades del día a día con paciencia, a entender y buscar una vida llena de plenitud.

“La gratitud es la memoria del corazón”

-Lao Tse

A mis amigos, por su constante apoyo, por su tiempo, consejos, amistad, por esas charlas que me llenaban de ánimo.

Agradecimientos

A mis Padres, Jair Rey Duarte y Luz Helena Niño, los regalos más bellos de Dios en esta vida, su constante sacrificio, apoyo, preocupación y sobre todo su amor, me ayudaron en los momentos más difíciles.

A mis amigos Eduardo y Eliana, quienes no solo con palabras, sino también con actos me demostraron su amistad y cariño.

Al profesor Enrique Mejía Ospino, por brindarme la oportunidad de realizar mi proyecto de grado a su lado, por ser el ejemplo más humano como persona y como académico, por el respeto, enseñanzas y apoyo brindado en esta etapa, y la motivación constante de no tener límites para lograr grandes avances en la ciencia.

Al profesor Jader Guerrero, por sus valiosas enseñanzas, por el acompañamiento científico y fuerte contribución en el proyecto, y sobre todo por esas jornadas de trabajo tan preciadas y provechosas.

Al profesor Rafael Cabanzo, por su disposición a colaborar ante cualquier duda o necesidad en el laboratorio.

A mis compañeros del LEAM, en especial a Ximena Ramírez, quien me guio y colaboro en todas las etapas de este proyecto, a Wilmar Contreras y Brayan Martínez, con quienes compartí la mayor parte del tiempo, por las buenas charlas, explicaciones y chistes, también a Ximena Calderón, por su disposición y colaboración.

A la Universidad Industrial de Santander (UIS) por adoptarme en su alma máter y al Instituto Colombiano de Petróleo ICP por brindarme la posibilidad de trabajar con ellos.

Contenido

	Pág.
Introducción	15
1. Planteamiento del Problema	18
2. Justificación	20
3. Consideraciones Teóricas	22
3.1 El Petróleo.....	22
3.2 Clasificación del Petróleo Crudo	23
3.2.1 Gravedad API.....	23
3.2.2 Viscosidad Cinemática.....	25
3.2.3 Contenido de Azufre	25
3.3 Análisis Composicional	26
3.3.1 Fraccionamiento.....	27
3.3.1.1 Fraccionamiento PIONA.	27
3.3.1.2 Análisis SARA (basado en su polaridad).....	28
4. Espectroscopia Infrarrojo.....	30
4.1 Introducción y Concepto de Espectroscopia Infrarroja	30

4.2 Principio de Mediciones Espectrales y Transformada de Fourier 31

4.3 Interferómetro de Michelson..... 32

4.4 Reluctancia Total Atenuada..... 34

4.5 Bandas de Absorción FT-IR/ATR en MIR..... 36

5. Estabilidad del Petróleo Crudo 38

5.1 Definición de Asfaltenos..... 39

5.2 Interacciones Moleculares y Modelos de Agregación 40

5.3 Estructura de Asfaltenos 44

6. Estudios de Estabilidad de Crudos de Petróleo..... 46

6.1 Índice de Inestabilidad Coloidal (CII) 47

6.2 Índice de Estabilidad Coloidal (CSI) 48

6.3 Índice de Estabilidad..... 49

6.4 Gráfico de Stankiewics (SP) 49

6.5 Criterio de estabilidad de Sepúlveda (QQA: análisis cualitativo – cuantitativo) 50

6.6 Parámetro Heithaus (parámetro P)..... 51

6.7 Modelo de Compatibilidad de Aceite (OCM) 52

6.8 Índice de Clase de Solubilidad de Asfaltenos (ASCI) 53

7. Quimiometría 54

7.1 ¿Qué es Quimiometría?..... 54

7.2 Métodos para la construcción de modelos predictivos 55

7.3 Regresión por Mínimos Cuadrados Parciales (PLSR)..... 56

7.3.1 Aspectos Matemáticos 58

7.3.2 SIMPLS Algoritmo para PLS 62

7.4 Validación Cruzada de Monte Carlo..... 63

8. Construcción de Modelos de Calibración (Aplicado a IR-MIR) 64

8.1 Evaluación y Optimización de Modelos de Calibración (Estadísticas) 65

8.1.1 Error estándar de calibración 66

8.1.2 Error estándar de calibración SEC 66

8.1.3 Error estándar de validación cruzada SECV 66

9. Metodología 67

9.1 Reactivos y Muestras 68

9.2 Espectroscopía Infrarroja 68

9.2.1 Adquisición Espectral 69

9.2.2 Análisis Espectral..... 69

9.3 Índice de Clase de solubilidad Asfaltenos (ASCI) 71

9.4 Desarrollo del modelo PLS-R..... 74

9.4.1 Preprocesamiento espectral..... 74

9.4.2 PLS – R..... 75

10. Análisis y Resultados 77

11. Comparación Índice de Clase de Solubilidad de Asfaltenos (ASCI) con otros métodos de Estabilidad.....	85
12. Conclusiones	88
Referencias Bibliográficas	89
Apéndices.....	95

Lista de Tablas

	Pág.
Tabla 1. Composición en general porcentual del petróleo crudo por elementos químicos.	22
Tabla 2. Especificaciones generales de clasificación de crudos según la gravedad API.....	24
Tabla 3. Frecuencias de absorción Infrarrojo	37
Tabla 4. Resumen estadístico de la fracción ASCI, API y SARA para las muestras de crudo colombiano.....	73
Tabla 5. Descripción del rendimiento del modelo	81
Tabla 6. Comparación de los datos obtenidos con diferentes métodos que evalúan la estabilidad y precipitación de asfaltenos en dos aspectos: composición y calidad.....	86
Tabla 7. Comparación valores propiedad ASCI e IN, utilizando la ecuación 29, para la Conversión Matemática de las Propiedades en Mención.....	86

Lista de Figuras

	Pág.
Figura 1. Crude oil fractionation based on PIONA analysis.	27
Figura 2. Esquema de la determinación SARA según la norma ASTM D2007.....	29
Figura 3. Esquema del Espectro de radiación electromagnética.....	30
Figura 4. Vista superior, imagen esquemática de Interferómetro de Michelson.	33
Figura 5. Procesos ópticos que tienen lugar cuando un haz infrarrojo en un cristal de alto índice de refracción, encuentra una muestra con menor índice de refracción.....	34
Figura 6. Ejemplo de cristal ATR con dos puntos calientes en su superficie superior.....	36
Figura 7. Estructura multinivel de asfaltenos de la molécula a la fase de depósito.....	39
Figura 8. Macroestructura de asfaltos.	41
Figura 9. Modelo Yen-Mullins. Este modelo muestra las estructuras moleculares y coloidales dominantes para los asfaltenos en solventes de laboratorio y aceites crudos.	42
Figura 10. Representación Esquemática de un ensamblaje supramolecular compuesto por especies huésped en un agregado de asfaltenos representativos.	43
Figura 11. (a) modelo continental o Isla, (b) modelo archipiélago.....	45
Figura 12. Criterio de estabilidad de Stankiewics.	49
Figura 13. Ejemplo de clasificación de zonas de estabilidad con el criterio de Sepúlveda (QQA: análisis cualitativo-cuantitativo).	50
Figura 14. Representación esquemática de ASCI, indicando cada una de las mezclas de solventes en sus mediciones.	53

Figura 15. Esquema fundamental de la Quimiometría 55

Figura 16. PLS como método de regresión lineal múltiple para la predicción de una propiedad y a partir de las variables X_1, \dots, X_m , aplicando coeficientes de regresión b_1, \dots, b_m (datos centrados en la media). 57

Figura 17. Matrices en PLS. 60

Figura 18. Descripción General de la calibración Multivariante. 65

Figura 19. Descripción Metodología (primera parte) FT-IR/ATR en la región MIR..... 69

Figura 20. Espectros FTIR-ATR de las 82 muestras de petróleo crudo. La flecha muestra la señal predictiva utilizada en el modelo de calibración (1900–400 cm^{-1}) desarrollada en Matlab R2019a. 70

Figura 21. Espectro FT-IR/ATR muestra de crudo colombiano, bandas características..... 70

Figura 22. Descripción Metodología (Parte 2) Índice de Clase de solubilidad de Asfaltenos (ASCI) 71

Figura 23. Paso a paso Metodología "Índice de Clase de Solubilidad de Asfaltenos" (ASCI). ... 73

Figura 24. Correlación de dos vectores b consecutivos contra el número de factores PLS 78

Figura 25. Errores estándar de 1000 particiones para cada uno de los primeros 20 LV utilizando el vector \bar{b} , en el plano SECV-SEC. 79

Figura 26. Métricas de rendimiento del modelo para predecir el valor del índice ASCI vs el número de variables latentes (Lv) 80

Figura 27. Vector de coeficientes de regresión para 1000 particiones, correspondientes a la undécima variable latente (sombreado en gris claro). 82

Figura 28. Correlación entre lo observado y lo predicho (valores ASCI). 83

Figura 29. Residuales para el valor ASCI de predicción usando FT-IR/ATR en la región media del infrarrojo. 84

Lista de Apéndices

	Pág.
Apéndice A. Información suplementaria propiedades ASCI y SARA, para cada una de las 82 muestras de crudos colombianos.	95

Resumen

Título: Desarrollo de un Modelo Predictivo para la Determinación de Estabilidad en Crudos Colombianos a partir de FTIR y PLS*

Autor: Alexander Jair Rey Niño**

Palabras Claves: Asfaltenos, Quimiometría, ASCI, Validación Cruzada

Descripción:

Se desarrolló un modelo para predecir el índice de clase de solubilidad de asfaltenos (ASCI), a partir de datos obtenidos por espectroscopia infrarroja con transformada de Fourier acoplada al uso de reflectancia total atenuada en la región del infrarrojo medio (FT-IR/ATR) combinada con la regresión de mínimos cuadrados parciales (PLS-R) y utilizando el método de Validación Cruzada de Monte Carlo (MCCV).

La determinación del inicio de precipitación de los asfaltenos para cada una de las 82 muestras de crudos colombianos, se determinó utilizando el índice de clase de solubilidad de asfaltenos (ASCI), la evaluación de este índice se realizó en las diferentes soluciones de Heptols (*n*- heptano/ tolueno), generando una base de datos del índice en mención. Los espectros (FT-IR/ATR) se registraron en la región del infrarrojo medio ($4000 - 400 \text{ cm}^{-1}$), para cada una de las 82 muestras de petróleo crudo.

Con la información recopilada se desarrolló un modelo de predicción mediante el algoritmo SIMPLS, que incluía los criterios para elegir el número adecuado de variables latentes (LVs) con la ayuda de la validación cruzada de Monte Carlo (MCCV). El modelo desarrollado se clasifica como un modelo robusto estadísticamente, presentando un error estándar de validación cruzada (SECV) de 1.42; permitiendo predecir el inicio de precipitación (ASCI), en una banda de confianza del 95% y con una alta correlación entre lo predicho y medido SECV y SEC alrededor de la unidad. El modelo obtenido permite superar los problemas de tiempo de respuesta y subjetividad en la determinación de la estabilidad de crudos.

* Trabajo de Grado

** Facultad de Ciencias. Escuela de Química. Director: Enrique Mejía Ospino. Dr. En Ciencias Químicas. Codirector: Diana Catalina Palacio. Dra. En Ciencias Naturales (Física)

Abstract

Title: Development of a Predictive Model for the Determination of Stability in Colombian Crude from FTIR y PLS*

Author: Alexander Jair Rey Niño**

Keywords: Asphaltenes, Chemometry, ASCI, Cross Validation

Description:

A model was developed to predict the asphaltene solubility class index (ASCI), from data obtained by Fourier transform infrared spectroscopy coupled to the use of attenuated total reflectance in the mid – infrared region (FT-IR/ATR). Combined with Partial Least Squares Regression (PLS-R) and using the Monte Carlo Validation (MCCV) method.

The determination of the onset of precipitation of asphaltenes for each of the 82 Colombian crude samples was determined using the asphaltene solubility class index (ASCI), the evaluation of the index was carried out in the different Heptols solutions (n-heptane/ toluene), generating a database of the index in question. Spectra (FT-IR/ ATR) were recorded in the mid-infrared region ($4000 - 400 \text{ cm}^{-1}$), for each of the 82 crude oil samples.

Whit the information collected, a prediction model was developed using the SIMPLS algorithm, which included the criteria to choose the appropriate number of latent variables (LVs) whit the help of Monte Carlo cross – validation (MCCV). The developed model is classified as statically robust model, presenting a standard error of cross validation (SECV) of 1.42; allowing to predict the onset of precipitation (ASCI), in a 95% confidence band and whit a high correlation between the predicted and measured SECV and SEC around the unit. The model obtained allows to overcome the problems of response time and subjectivity in determining the stability of crude oil.

* Bachelor Thesis

** Science Faculty. School of Chemistry. Director: Enrique Mejía Ospino. Dr. in Chemical Sciences. Co-director: Diana Catalina Palacio. Dra. In Natural Sciences (Physics)

Introducción

La industria del petróleo en Colombia ha cobrado una gran importancia, gracias a la alta demanda del petróleo, en los diversos procesos de uso y consumo de energía. Con ello ha aumentado la preocupación sobre su desempeño a futuro. Factores como la dinámica del mercado de combustibles, los cambios en las características de los mismos, las modificaciones por avances tecnológicos entre otros; han llevado a la industria del petróleo en Colombia, a desarrollar proyectos de mejoramiento y tratamiento, para adecuarse a las necesidades¹.

ECOPETROL S.A ha planteado tres estrategias para mantener el desarrollo y la autosuficiencia petrolera del país: (1) incrementar el factor de recobro en campos maduros, (2) aprovechar campos descubiertos no desarrollados y, (3) desarrollar los crudos pesados². El aprovechamiento de las reservas de crudos pesados, trae consigo la necesidad de enfrentar una serie de retos en las diferentes etapas operacionales de producción, transporte y refinación; considerando muy importante la caracterización química de los crudos para la selección de condiciones operacionales y control de calidad de las etapas involucradas³.

La caracterización de un crudo o fracción, consiste en el análisis de una serie mínima de propiedades que permiten un conocimiento básico de la calidad del crudo. La densidad, contenido de azufre, gravedad API^o, porcentaje de saturados, aromáticos, resinas y asfaltenos (análisis SARA), distribución de puntos de ebullición y análisis elemental (porcentaje en peso de: carbono, hidrogeno, oxigeno, nitrógeno y azufre) son algunas de las principales propiedades de calidad de un crudo o fracción.

De acuerdo con el valor de las propiedades de un crudo, este puede clasificarse por la propiedad de gravedad API, según la definición mundial, la gravedad API del crudo pesado varía de 10 a 22. 3°, mientras que el petróleo crudo extrapesado y el betún exhiben valores inferiores a 10°, también existen los crudos medianos (22, 3° a 31. 1°), ligero (31. 1° a 39°) y superligero (> 39°). Otra clasificación común es la viscosidad dinámica, el crudo pesado y el crudo extrapesado alcanzan valores de viscosidad entre 100 y 10.000 cSt, mientras que la viscosidad del betún es siempre superior a 10.000 cSt⁴. Diversos estudios indican que la viscosidad del crudo pesado, extrapesado o betún, depende de la concentración y tipo de asfaltenos presentes en el crudo y no de su composición química⁵.

La precipitación de asfaltenos es causal de inestabilidad de los crudos de petróleo, este fenómeno de precipitación se observa cuando se realizan mezclas (mezcla entre crudos pesado y crudos livianos o mezclas crudo pesado con nafta). el fenómeno de precipitación de asfaltenos, afecta negativamente a una variedad de procesos del petróleo, ya que exhibe algunos impactos indeseables en la calidad del crudo, como lo son, el aumento en la viscosidad de los crudos y estabilidad de las emulsiones, una disminución de los rendimientos de destilado y separación de fases inestables⁶.

Los problemas de inestabilidad ocasionados por precipitación o aglomeración de asfaltenos, debido a cambios de temperatura y/o presión, o cuando se presentan precipitaciones debido a mezclas, no pueden predecirse con exactitud con base en la caracterización que actualmente se realiza a un crudo. Por lo tanto, es necesaria la evaluación de parámetros de estabilidad y compatibilidad con la finalidad de generar recomendaciones para la minimización de problemas de producción.

Debido al actual panorama, Este trabajo de grado, desarrolló un modelo de predicción con parámetros de estabilidad de crudos colombianos, llevando a cabo el estudio de 82 crudos; caracterizados por espectroscopia infrarroja y estabilidad por clase de solubilidad de la fracción más polar y pesada presente en el petróleo crudo “Asfaltenos”. Como objetivo principal se tiene fortalecer y avanzar en el entendimiento y comportamiento de los asfaltenos en la estabilidad de los crudos y aportar de esa manera a la industria del petróleo.

1. Planteamiento del Problema

A medida que la producción de crudos livianos disminuye en el país, como los crudos provenientes de los campos de Caño Limón, Cusiana y Cupiagua quienes llevan su mismo nombre y el aumento de la producción de crudos pesados ($API < 10$), se ha hecho necesario realizar mezclas de crudos pesados con crudos livianos o diluyendo mezcla de pesados con fracciones como nafta, para poder transportar esta clase de crudos de un punto a otro y así incrementar su producción.

Este proceso de mezcla ha puesto en evidencia que aquellos crudos que tienen una caracterización similar (densidad, análisis SARA, contenido de azufre, True Boiling Point, número de ácido), y que se pensaría que no tienen problemas de estabilidad, al mezclarse presentan inestabilidad, observándose la aparición de sólidos (asociados a asfaltenos o ceras). Esta precipitación de compuestos en hidrocarburos o mezclas es un fenómeno que se presenta por la inestabilidad del sistema, ocasionada por un cambio en su equilibrio de fases; se manifiesta por las alteraciones ocurridas en la presión, la temperatura y la misma composición del crudo, ocasionando precipitación y aglomeración de asfaltenos, lo anterior resulta ser un problema grave ya que puede retardar el proceso de recuperación, producción, y transporte de petróleo.

Actualmente se cuenta con métodos de evaluación de estabilidad como el índice de inestabilidad coloidal (CII)⁷, índice de estabilidad coloidal (CSI)⁸, índice de estabilidad (SI)⁹, diagrama de Stankiewics (SP)¹⁰, análisis cualitativo-cuantitativo (QQA)¹¹, los cuales están directamente relacionados con el grado de confiabilidad del sistema, en función del porcentaje de saturados, aromáticos, resinas y asfaltenos (análisis SARA) presentes en el crudo. Estos parámetros mencionados anteriormente no están relacionados directamente con la “calidad” de los

asfaltenos (estructuras), la cual representa un parámetro más significativo para la evaluación de la precipitación en el crudo de petróleo.

Algunos autores han desarrollado métodos para evaluar esta “calidad” de los asfaltenos. Entre ellos el método propuesto por Wiehe¹² (IN-SBN): el número de insolubilidad (IN), siendo este una calificación de los asfaltenos, y el número de mezcla de solubilidad (SBN), que es una calificación del poder de solvencia del petróleo crudo en el que se encuentran estos asfaltenos. Al mismo tiempo TOTAL¹³ desarrolló un índice, llamado “Índice de clase de solubilidad de asfaltenos” (ASCI), cuyo propósito principal, es poder evaluar la calidad de los asfaltenos y utilizar este índice para estudios comparativos de crudos; presentando beneficiosamente una ventaja en la simplicidad de medición en laboratorios o incluso en campo.

Por lo anteriormente expuesto se planteó como objetivo de esta investigación, generar un parámetro robusto y simple, que calificara de mejor manera la solubilidad de los asfaltenos y diese proporción a una idea determinante de la estabilidad e inicio de precipitación de los asfaltenos en las diferentes muestras de petróleo crudo. Adicional a esto, la implementación de datos espectroscópicos obtenidos por FT-IR/ATR en la región media del infrarrojo (MIR), permite desarrollar un modelo estadístico, utilizando regresión de mínimos cuadrados parciales (PLS), junto con el método de validación cruzada de Monte Carlo, que nos permitió determinar de manera adecuada las variables latentes para el desarrollo del modelo.

2. Justificación

La inestabilidad de los crudos es uno de los problemas más graves que se presenta en la industria del petróleo, en algunos casos se asocia con la precipitación y aglomeración de la fracción de asfaltenos presentes en los crudos. Constituye un problema importante durante todas las etapas de producción del petróleo; lo que implica una gran inversión económica para su solución.

La metodología de índice de clase de solubilidad de asfaltenos (ASCI), representa una opción confiable y sencilla para la determinación de la estabilidad e inicio de la precipitación de los asfaltenos en las diferentes muestras de crudos colombianos en esta investigación, ya que nos revela una verdadera calificación de la solubilidad que presentan los asfaltenos. El parámetro ASCI puede generar una propiedad robusta para el desarrollo de un modelo predictivo y/o quimiométrico.

Debido a los gastos elevados, procesos y análisis complejos utilizados en la clasificación y calificación del comportamiento de los crudos, el desarrollo de métodos de análisis precisos y rápidos se ha convertido en una necesidad urgente para el control de la calidad de todas las etapas de refinación y producción del petróleo; a su vez la utilización de las diversas técnicas analíticas entre ellas la espectroscopia infrarroja, han contribuido en análisis estructural y funcional de diversos crudos, reemplazando métodos tradicionales, ya que esta técnica requiere un tratamiento mínimo de la muestra y ofrece una repetibilidad aceptable para la mayoría de muestras.

Ahora, la combinación de las técnicas analíticas y medición de propiedades fisicoquímicas del petróleo crudo o de sus fracciones, se pueden introducir al análisis estadístico quimiométrico

aplicado a datos espectroscópicos, generando modelos novedosos que permiten evaluar y predecir diferentes propiedades fisicoquímicas o concentraciones de muestras complejas.

En este campo el Laboratorio de Espectroscopia Atómica y Molecular (LEAM) ha desarrollado importantes investigaciones para predecir parámetros y propiedades de los crudos en Colombia. Enrique Mejía-Ospino y col¹⁴ desarrollaron un método para el uso de análisis multivariado en la determinación SARA de crudos por Espectroscopia NIR; Enrique Mejía-Ospino y col¹⁵ desarrollaron una metodología basada en espectroscopia MIR-ATR y PLS-R para predicción del contenido de asfaltenos en residuos de vacío de crudos Colombianos. Este tipo de trabajos utilizan la quimiometría combinada con estadística y resultados de análisis químicos para extraer información útil de datos a través de métodos multivariados con el objetivo de la determinación cuantitativa y cualitativa de las propiedades de los crudos. Estas investigaciones hacen parte de un grupo de trabajos que se realizaron en el Laboratorio de Espectroscopia Atómica y Molecular LEAM.

Teniendo en cuenta los problemas actuales de inestabilidad de crudos y la experiencia en el área de caracterización de mezclas complejas, se propuso realizar un análisis de la estabilidad de 82 crudos colombianos por espectroscopia infrarroja con transformada de Fourier acoplada al uso de Reflectancia Total Atenuada FT-IR/ATR en la región del infrarrojo medio (MIR), combinada con la regresión de mínimos cuadrados parciales (PLS-R) y utilizando el método de validación cruzada de Monte Carlo. Como propiedad a determinar o predecir por medio del modelo quimiométrico, se recolectaron los valores del parámetro o propiedad “Índice de Clase de Solubilidad de Asfaltenos” (ASCI), el cual genera una calificación de la solubilidad de asfaltenos y de la estabilidad de cada uno de los 82 crudos colombianos.

3. Consideraciones Teóricas

3.1 El Petróleo

La palabra Petróleo, deriva del latín *petra* y *oleum*, significa literalmente aceite de roca y se refiere a los hidrocarburos (parafínicos, nafténicos o aromáticos), que se producen ampliamente en las rocas sedimentarias en forma de gases, líquidos, semisólidos o sólidos¹⁶. El petróleo está compuesto principalmente por hidrocarburos, a su vez, también contiene pequeñas trazas de compuestos orgánicos, diversos heteroátomos como azufre, nitrógeno, oxígeno y metales (principalmente Níquel y Vanadio); también contiene agua, materia inorgánica (sales y azufre elemental), (ver Tabla 1)^{17 18}.

Tabla 1.

Composición en general porcentual del petróleo crudo por elementos químicos.

Elemento	Concentración
Carbono (C)	83.0 - 87.0
Hidrogeno (H)	10.0 – 14.0
Nitrógeno (N)	0.1 – 2.0
Oxígeno (O)	0.05 – 1.5
Azufre (S)	0.05 – 6.0

Nota. Adaptado de Del libro Speight, J. G. The Chemistry and Technology of Petroleum. The Chemistry and Technology of Petroleum (2014). doi:10.1201/b16559.

3.2 Clasificación del Petróleo Crudo

El petróleo crudo varía dramáticamente en sus propiedades dependiendo de su origen¹⁹. El petróleo pesado, el petróleo extra pesado, la arena de petróleo, las arenas bituminosas, el esquisto bituminoso y el betún son recursos petroleros poco convencionales. Todos los crudos tienen características fisicoquímicas diferentes y muy variables. Por lo tanto, el petróleo puede variar en composición de un campo petrolero a otro, de un pozo a otro en el mismo campo, e incluso de un nivel a otro en el mismo pozo²⁰.

Para obtener la información necesaria para la descripción y clasificación de los diferentes crudos, se utilizan comúnmente esquemas analíticos, uno de ellos es la realización de un ensayo de inspección²⁰, estos ensayos generalmente implican la determinación de varias propiedades fisicoquímicas de los aceites crudos como la gravedad API°, el contenido de azufre, la viscosidad cinemática, el análisis SARA. Las propiedades anteriormente mencionadas son las tendencias generales de mayor importancia en la industria del petróleo²¹.

3.2.1 Gravedad API

La gravedad API° es una propiedad fundamental en el proceso de producción inicial para estimar los parámetros relacionados con la producción, el refinamiento, el almacenamiento y la medición del petróleo, precios y toxicidad de los crudos²². El petróleo crudo se clasifica como “ligero”, “medio” y “pesado” (aunque el rango de grados API° no se ha estandarizado), haciendo referencia a la densidad relativa del petróleo en función de los valores de gravedad API°.

Debido a la presencia de hidrocarburos que se pueden separar por destilación. Los aceites crudos ligeros con un alto grado de gravedad API° son adecuados en la industria petrolera. Sin embargo, los crudos ligeros se consideran aceites con carácter toxico e inflamable. Por otro lado, los crudos pesados contienen altas concentraciones de azufre y diferentes metales (como el vanadio y niquel), que no pueden ser purificados por métodos convencionales. Por ejemplo, los aceites pesados se consideraron aquellos materiales de tipo petróleo que tienen una gravedad API° algo menor que 20° API²⁰, y un crudo ligero está en el orden de 31.1 a 39° API. (Ver Tabla 2).

Tabla 2.

Especificaciones generales de clasificación de crudos según la gravedad API

Aceite Crudo	Densidad (g/cm ³)	Gravedad API°
Extra pesado	>1.0	10
Pesado	1.0 – 0.92	10.0 – 22.3
Mediano	0.92 – 0.83	22.3 – 31.1
Ligero	0.87 – 0.83	31.1 – 39
Superligero	< 0.83	>39

Nota. Adaptado Del libro Speight, J. G. The Chemistry and Technology of Petroleum. The Chemistry and Technology of Petroleum (2014). doi:10.1201/b16559.

La gravedad específica del petróleo está influenciada por la composición química del mismo; pero, la correlación cuantitativa es difícil de establecer. Sin embargo, generalmente se reconoce, que el aumento de las cantidades de compuestos aromáticos da como resultado un aumento de la densidad, mientras que un aumento de los compuestos saturados, da como resultado, la disminución de la densidad. De hecho, también es posible reconocer ciertas tendencias preferidas entre la densidad del petróleo y una u otras propiedades físicas. Por ejemplo, existe una correlación aproximada entre la gravedad API° y, el contenido de azufre y viscosidad¹⁶.

3.2.2 Viscosidad Cinemática

La viscosidad en un fluido es una medida de la resistencia interna de un fluido al flujo. Da una noción del grosor del fluido ²³. La viscosidad cinemática es una medida del flujo resistivo de un fluido, solo bajo la acción de la gravedad. La viscosidad cinemática se determina experimentalmente por la norma ASTM D445 ²⁴; Según este método, la viscosidad se determina a diferentes temperaturas, generalmente 298.15 K (77 F°) y 373.15 K (212 °F); midiendo el tiempo que tarda un volumen fijo de aceite en fluir bajo gravedad a través de un viscosímetro calibrado. La viscosidad cinemática se calcula multiplicando el tiempo por la constante de calibración del viscosímetro.

3.2.3 Contenido de Azufre

El azufre elemental es un componente común de los sedimentos y, si está presente en la roca del depósito, se disuelve en el petróleo crudo y reacciona lentamente con él para producir varios compuestos de azufre y/o sulfuro de hidrogeno, que pueden reaccionar aún más con ciertos componentes del petróleo. Estas reacciones son probablemente muy parecidas a las que se producen en el lecho fuente y se supone que son gran parte responsables del contenido de azufre del petróleo. Estas reacciones van acompañadas de un oscurecimiento del aceite y un aumento significativo en la gravedad específica y la viscosidad ²⁰.

El contenido de azufre de un crudo es un factor importante, ya que se trasladaría a los productos de refino, que están sometidos a fuertes restricciones de calidad respecto a su contenido. Será necesario eliminarlo de forma adecuada mediante una serie de procesos que encarecerán la

operación, por esta razón, los crudos de bajo contenido de azufre necesitaran unos procesos más suaves y normalmente su valoración en el mercado será superior a los de alto azufre.

El contenido en azufre depende del origen y antigüedad del crudo; los hay con contenido muy bajo, con valores del 0.1% en peso, hasta con valores superiores al 6% en peso. Ambos extremos son excepciones, normalmente el intervalo se encuentra entre el 0.2% (sweet crude oil) y 4% (sour crude oil) en peso.

3.3 Análisis Composicional

La composición del petróleo obtenido del pozo es variable y depende no solo de la composición original del petróleo in situ, sino también de la forma de producción y la etapa alcanzada en la vida del pozo o deposito. Esta variación puede ser tanto en peso molecular como en los tipos de moléculas presentes en el petróleo²⁵. Los hidrocarburos presentes en el petróleo crudo se clasifican en:

Parafinas: hidrocarburos saturados con cadenas lineales o ramificadas, pero sin ninguna estructura de anillos.

Cicloparafinas (naftenos): hidrocarburos saturados que contienen uno o más anillos, cada uno de los cuales puede tener una o más cadenas laterales de parafina (más correctamente conocido como hidrocarburos alicíclicos).

“Las parafinas y los naftenos juntos se llaman saturados”

Aromáticos: hidrocarburos que contienen uno o más núcleos aromáticos como los sistemas de anillos de benceno, naftaleno y fenantreno que pueden ser unidos con anillos de naftaleno (sustituidos) o cadenas laterales de parafina.

“Los hidrocarburos aromáticos altamente ramificados y de alto peso molecular son conocidos como moléculas de asfaltenos, que normalmente contienen entre el 4 y 20%”

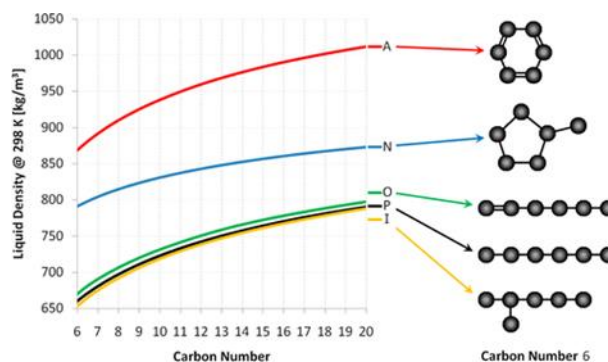
3.3.1 Fraccionamiento

Es posible obtener una visión general detallada de la composición del aceite que se puede utilizar para las predicciones del proceso. Por lo tanto, los métodos de fraccionamiento también juegan un papel importante, junto con los métodos de prueba física, ya que nos permiten evaluar los aceites pesados y los diversos procesos de recuperación.

3.3.1.1 Fraccionamiento PIONA. La caracterización basada en el tipo de hidrocarburos se conoce mejor como análisis de n-parafina, isoparafinas, olefina, nafteno y aromático (PIONA)²³²⁶. Porque las olefinas no suelen estar presentes en el petróleo crudo; este análisis para el petróleo crudo se conoce como análisis PNA. Esta técnica fue desarrollada para estudiar las propiedades termodinámicas de la mezcla de hidrocarburos.

Figura 1.

Crude oil fractionation based on PIONA analysis.



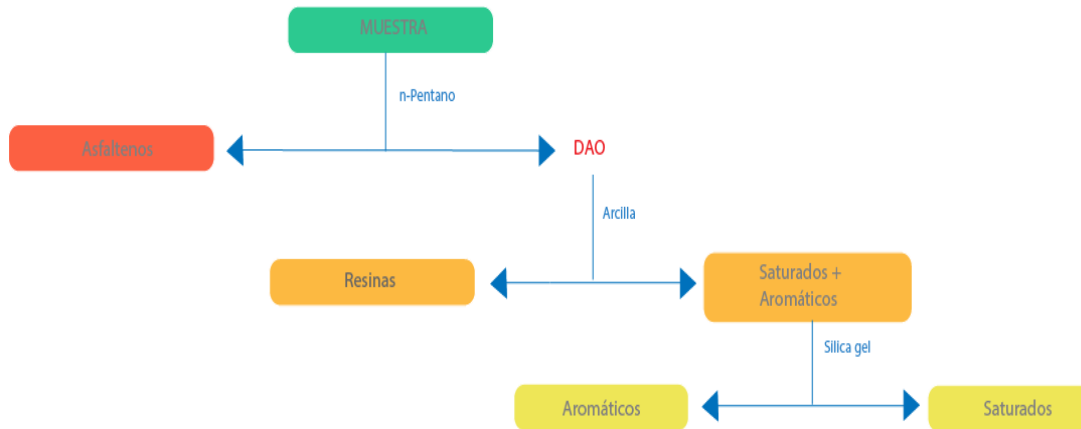
Nota: Ilustración adaptada de Hay, G., Loria, H. & Satyro, M. A. Thermodynamic modeling and process simulation through PIONA characterization. Energy and Fuels 27, 3578–3584 (2013).

En 2013 Hay y Col²⁷, desarrollaron una nueva técnica de Caracterización PIONA, basada en la estructura molecular de los grupos componentes. En este método, se supone que un grupo con el mismo número de carbono tiene un valor promedio de propiedades físicas con diferentes propiedades termodinámicas. Esta suposición introduce algún error, como en el caso de 1-buteno, 2-buteno e iso-buteno. Incluso si estos tres compuestos tienen los mismos cuatro carbonos, sus propiedades físicas varían significativamente. La *Figura 1*, muestra cómo cambia la densidad en la estructura molecular de PIONA versus diferentes números de carbono a 298 K. Además, muestra los diferentes grupos de estructura molecular de PIONA con seis átomos de carbono²⁸. El análisis PIONA se realiza mediante GC capilar, que utiliza una columna capilar unida no polar. Las desventajas de este método son su largo tiempo de análisis y vulnerabilidad al error humano (Excelencia de Shimadzu en ciencia 2017).

3.3.1.2 Análisis SARA (basado en su polaridad). El petróleo crudo se puede fraccionar en cuatro componentes: saturados, aromáticos, resinas y asfaltenos (análisis SARA). Una caracterización de composición confiable de las fracciones de petróleo crudo, de gran importancia para la optimización y evaluación del desempeño de los productos, los procesos de refinación. En general, el desasfaltado de la muestra, es el primer paso en la separación SARA de los aceites crudos. La separación es de clase SAR, se logra mediante HPLC, utilizando columnas de fase unida, o una combinación de sílice, y un alcanos como fase móvil ²⁹. La ASTM ofrece varios procedimientos para el análisis SARA: la norma ASTM D2007³⁰ (Ver *Figura 2*) y la ASTM D4124³¹, junto con las normas ASTM D2549³² que usa la cromatografía líquida para la separación de las fracciones de saturados y aromáticos.

Figura 2.

Esquema de la determinación SARA según la norma ASTM D2007.



La fracción Saturada consiste en hidrocarburos saturados, no polares, que incluyen hidrocarburos saturados lineales, ramificados y cíclicos. Los compuestos Aromáticos contienen uno o más anillos aromáticos que son más polarizables³³. Las Resinas y los Asfaltenos, ambas fracciones tienen sustituyentes polares, pero tienen una diferencia muy particular, cuya definición está contemplada en términos de solubilidad con ciertos solventes aromáticos. Las resinas son miscibles con *n*- alcanos, como el pentano y el *n*- heptano. Por otra parte, los asfaltenos son la fracción que es insoluble en *n*- alcanos, como *n*- heptano, pero soluble en solventes aromáticos, como tolueno³³.

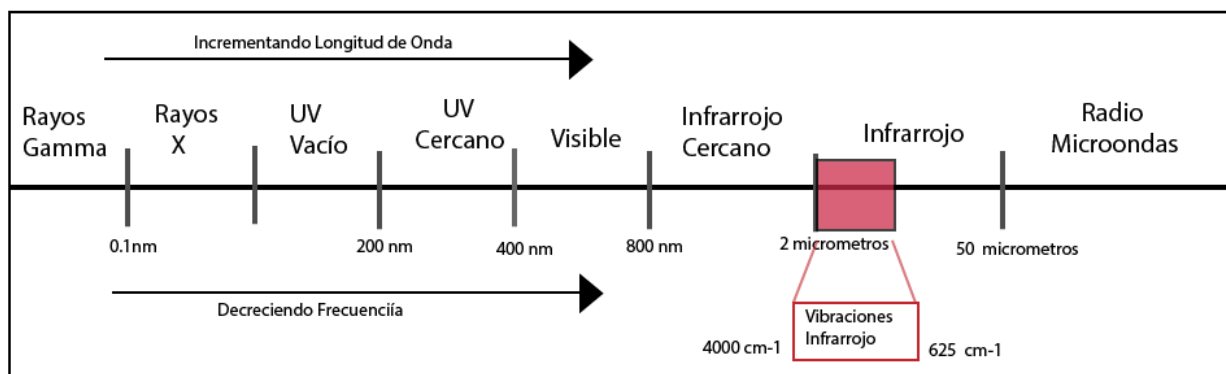
4. Espectroscopia Infrarrojo

4.1 Introducción y Concepto de Espectroscopia Infrarroja

La región infrarroja ocupa solo una pequeña porción del espectro de radiación electromagnética, situada entre la región ultravioleta más energética y la región de radiofrecuencia de baja energía³⁴.

Figura 3.

Esquema del Espectro de radiación electromagnética.



Nota: Ilustración tomada de Thompson, J. M. *Infrared Spectroscopy*. (2018).

Los espectros infrarrojos de los compuestos orgánicos generalmente consisten en una variedad de bandas de absorción, todas con diferentes intensidades y formas, y todas ocurren en longitudes de onda usualmente entre 4000 – 625 cm⁻¹, tanto las intensidades como las posiciones de las bandas de absorción son importantes para deducir las características estructurales de los compuestos orgánicos. Casi todos los compuestos orgánicos tienen su propio y único espectro infrarrojo, considerándose así mismo, como la huella digital de la molécula. Los enlaces químicos

no son entidades rígidas e inflexibles, pero son capaces de experimentar una serie de vibraciones diferentes. Las vibraciones de enlace pueden ser de varios tipos generales, que consisten en vibraciones de estiramiento simétricas y asimétricas, así como los modos de flexión dentro y fuera del plano, todas estas se conocen como vibraciones fundamentales. Además de los modos de estiramiento y flexión, los espectros de los compuestos orgánicos a menudo muestran ciertas vibraciones no fundamentales a niveles de energía superiores a las vibraciones fundamentales. Clasificándose como bandas de sobretonos, combinación o diferencia³⁴.

Cuando un compuesto orgánico se irradia por infrarrojos en este rango, todos los enlaces en las moléculas cuyas vibraciones van acompañadas de cambios de momento dipolar absorberán el infrarrojo a frecuencias que coinciden con sus propias frecuencias vibratorias naturales, estas absorciones dan como resultado el aumento en la amplitud del movimiento vibratorio que es detectado y registrado por el instrumento infrarrojo. El resultado es un registro de absorciones de enlaces versus longitudes de onda o un espectro infrarrojo³⁴. Las absorciones de enlaces están acompañadas por una serie de rotaciones de enlaces, debido a esto, las absorbancias aparecen como bandas anchas a agudas en lugar de líneas.

4.2 Principio de Mediciones Espectrales y Transformada de Fourier

La fuente IR emite rayos IR, y una porción del rayo a través de la ranura se convierte en una luz paralela mediante espejos, $I(t)$ que van al prisma. Por el momento, la luz tiene una imagen de la onda plana de los campos eléctricos y magnéticos en función del tiempo t , y la intensidad I es proporcional al campo eléctrico al cuadrado³⁵.

El prisma es un elemento óptico que funciona como el corazón del espectrómetro: el ángulo de refracción depende de la longitud de onda de la luz. De esta manera, la luz paralela recta se convierte en una luz “dispersa”. Si la luz es visible, aparece un patón de arco iris en la pantalla. Como la luz IR es invisible, se debe usar un detector para tener un espectro en lugar del patrón visible.

El espectro aparece como un gráfico desarrollado por dos ejes de longitud de onda k , e intensidad S . Dado que la longitud de onda está directamente relacionada con un parámetro más útil, la frecuencia angular ω , el espectro se denota como $S(\omega)$. De esta manera, el elemento óptico trabaja para transformar la función del dominio del tiempo $I(t)$, en el dominio de frecuencia $S(\omega)$, que corresponde a la transformación de Fourier en matemáticas. En otras palabras, el papel del elemento óptico es que la transformada de Fourier, transforma la luz en un espectro rápidamente. La transformada de Fourier está representada por las ecuaciones (1) y (2)³⁵:

$$S(\omega) = \int_{-\infty}^{\infty} I(t)e^{-i\omega t} dt \quad \text{Ecuación (1)}$$

$$I(t) = \frac{1}{2} \int_{-\infty}^{\infty} S(\omega)e^{i\omega t} d\omega \quad \text{Ecuación (2)}$$

4.3 Interferómetro de Michelson

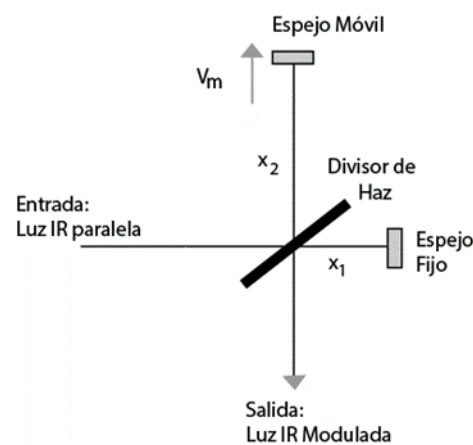
El interferómetro consta de un divisor de haz y dos espejos planos. El divisor de haz permite que la mitad de la luz pase a través de él, y la mitad restante se refleje en la superficie. Uno de los espejos se coloca en posición fija con una distancia de x_1 al punto de división, mientras que el otro espejo está en un riel para cambiar la distancia x_2 que se llama el “espejo móvil”. El espejo móvil se mueve muy suavemente con una velocidad muy constante controlada electrónicamente³⁵. (Ver *Figura 4*).

Cuando el espejo móvil se desplaza desde la posición inicial de $x_2 = x_1$, con la velocidad constante de v_m , la posición del espejo móvil $x_2(t)$ se representa como:

$$x_2(t) = x_1 + v_m t \quad \text{Ecuación (3)}$$

Figura 4.

Vista superior, imagen esquemática de Interferómetro de Michelson.



Nota: Adaptado de Hasegawa, T. Quantitative infrared spectroscopy for understanding of a condensed matter. Quantitative Infrared Spectroscopy for Understanding of a Condensed Matter (2017). doi:10.1007/978-4-431-56493-5

Un rayo IR paralelo se dirige al interferómetro, la mitad de la luz va al espejo fijo, y la luz reflejada se refleja a medias en el divisor, para salir. De manera similar, la mitad restante se refleja para ir al espejo en movimiento, y la luz reflejada pasa a través del divisor para salir como luz IR modulada. Como resultado, las dos líneas de luz se superponen en la ruta de salida, lo que genera interferencia. La superposición de las dos ondas de oscilación del campo eléctrico $E(t)$, puede expresarse simplemente considerando tres hechos: (1) las amplitudes A , de las dos ondas son

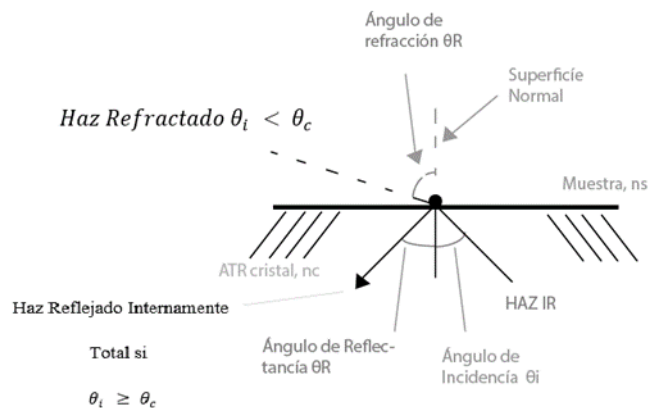
iguales entre sí, ya que el divisor de haz divide la luz para la transmisión y la reflexión a medias, (2) la frecuencia angular ω , es común a las dos ondas y (3) un viaje de ida y vuelta entre el divisor del haz y un espejo influye en el cambio de fase³⁵.

4.4 Reluctancia Total Atenuada

La Reflectancia Total Atenuada (ATR) se basa en la reflectancia interna, la *Figura 5* muestra un haz de luz que viaja a través de un cristal de alto índice de refracción n_c , que encuentra un límite con una muestra con menor índice de refracción n_s . Las líneas punteadas en la figura son la superficie normal, y el ángulo que forma el haz entrante con la superficie normal, es el ángulo de incidencia θ_i .

Figura 5.

Procesos ópticos que tienen lugar cuando un haz infrarrojo en un cristal de alto índice de refracción, encuentra una muestra con menor índice de refracción.



Nota: Adaptado de Smith, B. C. Fundamentals of fourier transform infrared spectroscopy, second edition. Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition (2011).

Lo que sucede al rayo infrarrojo cuando alcanza el límite de los dos medios depende de θ_i . Si el ángulo de incidencia es pequeño, parte del haz se reflejará en la superficie interna del cristal, y parte del haz se refractará fuera del cristal hacia la muestra, dando el haz refractado. El ángulo de refracción denotado por θ_R , es el ángulo que forma el haz refractado con la superficie normal. Las leyes de refracción dicen que como θ_i aumenta, θ_R aumenta. En algún ángulo de incidencia, el ángulo de refracción se convertirá en 90° y el haz infrarrojo ya no dejará el cristal, sino que permanecerá dentro de él. En este caso, toda la luz se refleja en la superficie interna del cristal, de ahí el término de reflectancia total interna³⁶.

El ángulo mínimo de incidencia en el que la reflectancia interna total ocurre en un material, se denomina ángulo crítico θ_C . La reflexión interna total, tendrá lugar en θ_C y todos los ángulos de incidencia mayores que θ_C . El ángulo crítico depende de los índices de refracción del cristal y la muestra, como se muestra en la ecuación (4):

$$\theta_C = \sin^{-1}(n_s/n_c) \quad \text{Ecuación (4)}$$

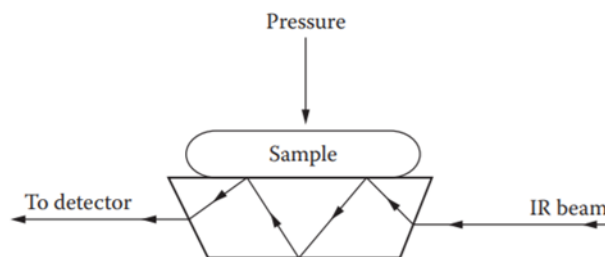
En el punto de reflectancia interna, los rayos infrarrojos entrantes y salientes ocupan el mismo volumen. Bajo las condiciones correctas incluyendo $\theta_i \geq \theta_C$, estos dos haces, sufren interferencia constructiva. Esto significa que la amplitud infrarroja en el punto de reflectancia interna será mayor que la amplitud en ambos lados³⁶. Como ya se está en la superficie del cristal, la amplitud mejorada no tiene más que ir hacia arriba, lo que da como resultado, que la luz infrarroja se adhiera al espacio sobre la superficie del cristal, que se llama onda evanescente o también denominados “puntos calientes”.

Para tomar el espectro de una muestra, se pone en contacto con el punto caliente, parte del haz es absorbido por la muestra y luego se enfoca en el detector (Ver *Figura 6*). El espectro de

fondo se obtiene del cristal limpio y seco. Dado que las absorbancias de la muestra atenúan la intensidad del haz infrarrojo totalmente reflejado, se utiliza el término Reflectancia Total Atenuada (ATR).

Figura 6.

Ejemplo de cristal ATR con dos puntos calientes en su superficie superior.



Nota: Tomado de Smith, B. C. Fundamentals of fourier transform infrared spectroscopy, second edition. Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition (2011)

4.5 Bandas de Absorción FT-IR/ATR en MIR

La espectroscopia de absorción infrarroja es una de las técnicas más utilizadas para examinar la constitución química de los materiales orgánicos. De hecho, los estudios de los espectros infrarrojos de una variedad de compuestos han demostrado una correlación sistemática con la estructura. Por lo tanto, no es sorprendente que se haya hecho intentos para estimar las entidades estructurales presentes en el petróleo y otros materiales carbonosos mediante espectroscopia infrarroja. Cuando se prepara y escanea adecuadamente con instrumentos adecuados, los espectros infrarrojos de las fracciones de petróleo muestran varias bandas bien definidas¹⁶ (Ver Tabla 3).

Tabla 3.

Frecuencias de absorción Infrarrojo

Frecuencias de Absorción Infrarrojo		
Enlace	Tipo de compuesto	Frecuencia, Longitud de Onda (cm^{-1})
O – H	Alcoholes y fenoles Monoméricos	3640 – 3160 Estiramiento
O – H	Alcoholes y fenoles unidos por puentes de hidrogeno	3600 – 3200 Estiramiento
N – H	Aminas, Amidas	3560 – 3320 Estiramiento; sin enlaces hidrogeno
N – H	Aminas	3400 – 3100 Estiramiento; con enlaces de hidrogeno
C – H	Alquinos	3333 – 3267 Estiramiento
C – H	Anillos aromáticos	3100 – 3000 Estiramiento
C – H	Alquenos	3080 – 3020 Estiramiento
O – H	Ácidos Carboxílicos	3000 – 2500 Estiramiento
C – H	Alcanos	2960 – 2850 Estiramiento
C ≡ N	Nitrilos	2260 – 2220 Estiramiento
C ≡ C	Alquinos	2260 – 2100 Estiramiento
C – H	Sustitución del anillo de fenilo	2000 – 1600
C = O	Aldehídos, Cetonas, Acido carboxílico, Esteres	1760 – 1670 Estiramiento
C = C	Alquenos	1680 – 1640 Estiramiento
NO ₂	Nitrocompuestos	1660 – 1500 Estiramiento Asimétrico
N – H	Aminas	1650 – 1580 Flexión
C = C	Anillos aromáticos	1600, 1500 Estiramiento
C – H	Alcanos	1470 – 1350 Tijera y Flexión
NO ₂	Nitrocompuestos	1660 – 1500 Estiramiento asimétrico
C – H	CH ₃ deformación	1380 Isopropil, t-butil
C – N	Aminas	1340 – 1020 Estiramiento
C – O	Alcoholes, Éteres, Ácidos carboxílicos, Esteres	1260 – 1000 Estiramiento
C – H	Alquenos	1000 – 675 Flexión
C – H	Sustitución del anillo fenilo	870 – 675 Flexión
C – H	Alquinos	700 – 610 Flexión
C – Cl	Compuestos Clorados	600 – 800
C – Br	Compuestos Bromados	500 – 600
C – I	Compuestos Yodados	500

Nota. Adaptado de Speight, J. G. The Chemistry and Technology of Petroleum. Chem. Technol.

Pet. (2014) doi:10.1201/b16559.

La espectroscopia Infrarroja también se ha utilizado ampliamente en la aclaración de las entidades estructurales presentes en los asfaltenos y resinas. Por ejemplo, las bandas de aproximadamente 3030 cm^{-1} se han asignado al estiramiento aromático carbono-hidrogeno y han demostrado que esta banda puede cambiar a valores tan altos como 3052 cm^{-1} , con un numero de

anillo decreciente. Además de las funciones carbonilo, los espectros infrarrojos de los asfaltenos muestran una fuerte banda de carbono-hidrogeno a 2940 cm^{-1} , un pico característico de los grupos CH_2 a 1470 cm^{-1} y picos para los grupos metilo a 1380 cm^{-1} .

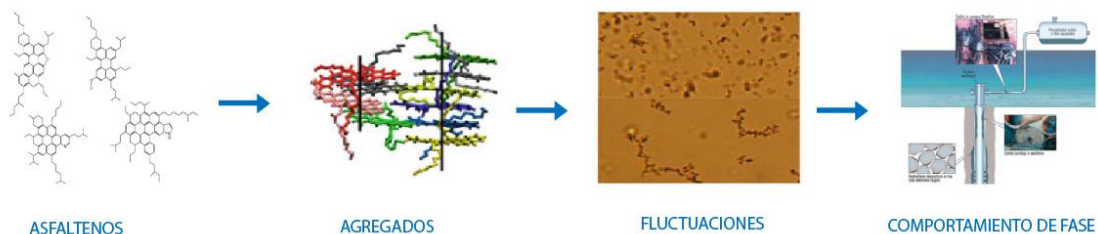
Se cree que el grupo de los picos de aproximadamente 760 , 814 , y 870 cm^{-1} se debe a estructuras de anillo aromático sustituido. La banda de anillo aromático a 1600 cm^{-1} es preeminente, mientras que la absorción amplia y fuerte entre 1330 y 1110 cm^{-1} se asigna a especies que contienen oxígeno. Los picos a 1300 y 1135 cm^{-1} , se han atribuido a enlaces azufre-oxígeno, pero un pico característico a 1040 cm^{-1} no se ha diagnosticado adecuadamente. La banda a 725 cm^{-1} , generalmente atribuida a cadenas de metileno que tienen cuatro o más unidades, a menudo está presente, pero varía en intensidad dependiendo de la fuente de petróleo. Se ha sugerido que la banda a 740 cm^{-1} surge de alquilbencenos, la banda aproximadamente a 1600 cm^{-1} se ha asignado a enlaces carbono-carbono aromático. Además, la aromaticidad (banda de 1600 cm^{-1}) de los asfaltenos es significativamente mayor que la de la fracción de malteno, pero, por espectroscopia infrarroja, aumenta solo ligeramente con el tamaño molecular¹⁶.

5. Estabilidad del Petróleo Crudo

Cuando se hace alusión a la estabilidad de crudos, se hace referencia a la propiedad de un crudo para mantener los asfaltenos estables y prevenir la posible precipitación o aglomeración de los mismos.

Figura 7.

Estructura multinivel de asfaltenos de la molécula a la fase de depósito.



Nota: Tomado y modificado de Xu, C. & Shi, Q. Structure and Modeling of Complex Petroleum Mixtures.

La estructura de los asfaltenos es uno de los factores más importantes en la determinación de la estabilidad del crudo, los asfaltenos provenientes de crudos inestables se caracterizan por una alta aromaticidad, bajo contenido de hidrogeno y alta condensación de anillos aromáticos. Los asfaltenos, la fracción más polar en el petróleo crudo, son críticos para todos los aspectos de utilización del petróleo. Las fuertes interacciones entre los asfaltenos conducen a varios niveles de agregación; agregación que es responsable de una variedad de problemas de transporte³⁷.

5.1 Definición de Asfaltenos

En la industria del petróleo, los asfaltenos se definen operacionalmente como la parte del petróleo crudo que es insoluble en alcanos con bajo contenido de carbono (pentano o n-heptano) y solubles en solventes fuertes (tolueno)³⁸. La definición de asfaltenos indica inherentemente que son solo una clase de solubilidad que tiene composiciones químicas y estructuras moleculares indefinidas³⁹.

Los asfaltenos son una mezcla polidispersa de la fracción de petróleo crudo más pesada. Es una fracción compleja, constituida por hidrocarburos policíclicos (HAP), con cadenas laterales de alquilo, pero también heteroátomos (N, S, O), y metales (Ni, V), que los convierte en la fracción de petróleo más polar y pesada. Los asfaltenos han atraído la atención de la industria y ciencia del petróleo, ya que son responsables de muchos problemas durante la producción, transporte y refinación del petróleo, a causa de la fuerte tendencia a agregarse, debido a la presencia de heteroátomos en la estructura aromática de los asfaltenos, estas especies están involucradas en una amplia variedad de interacciones físicas y químicas, que conducen a la asociación intermolecular en los asfaltenos y la formación de agregados. Si la condición es termodinámicamente inestable, la agregación continuará y se producirá fluctuación de los asfaltenos³⁹.

5.2 Interacciones Moleculares y Modelos de Agregación

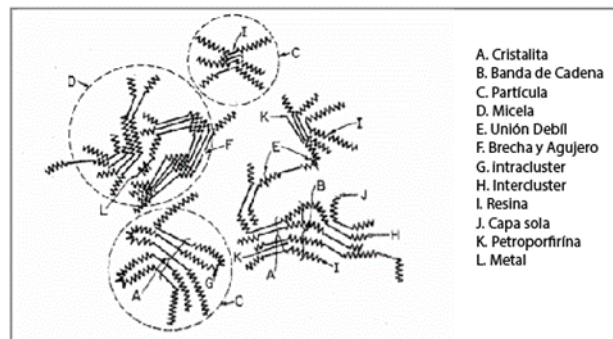
El estudio y análisis sobre los asfaltenos es fenomenológico. La complejidad de los asfaltenos se manifiesta en dos aspectos: (1) las numerosas especies involucradas en el sistema, (2) las diferentes formas y tamaños de ensamblaje molecular. La estructura y la agregación siguen siendo el tema a estudiar con mayor profundidad; considerándose un problema enorme la estructura molecular y composición de los asfaltenos, al igual, que su naturaleza intrínseca de agregación de los mismos.

Una variedad de estudios ha revelado que los agregados de asfaltenos y resinas se originan principalmente a la interacción de hidrogeno, uniones ácido-base, enlaces de coordinación e interacciones de apilamiento $\pi - \pi$. Se han desarrollado una variedad de modelos con el firme propósito de dar una mejor explicación a este proceso de agregación e inestabilidad³⁹.

Basado principalmente en datos de caracterización de asfaltos sólidos por mediciones de Difracción de Rayos X (DRX), el profesor The Fu Yen y col⁴⁰ propusieron previamente una jerarquía de estructuras de asfaltos para dar cuenta de muchas de sus características conocidas en ese momento, denominando una estructura jerárquica de los asfaltos como el “Modelo de Yen”⁴¹.

Figura 8.

Macroestructura de asfaltos.



Nota: Adaptada de Dickie, J. P. & Yen, T. F. Macrostructures of the Asphaltic Fractions by Various Instrumental Methods. Anal. Chem. 39, 1847–1852 (1967).

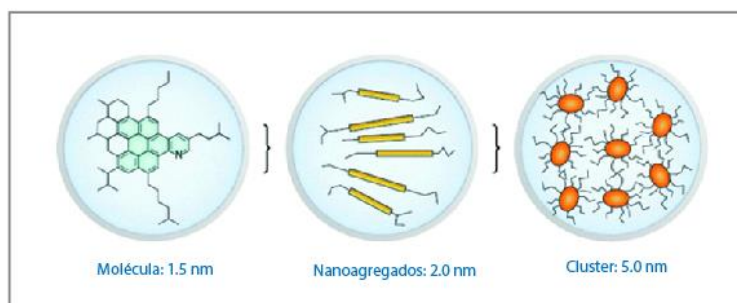
Como se observa en la *figura 8*. Este modelo muestra el mecanismo de agregación de los asfaltos, desde el estado de la molécula hasta el estado del racimo: la lámina aromática individual, se puede apilar para formar partículas elementales, y estas partículas pueden asociar micelas: las micelas de asfaltos pueden agruparse en agregados, cuando la concentración es suficientemente alta⁴¹. La asociación molecular de las hojas aromáticas apiladas en solución es impulsada por la interacción $\pi - \pi$ y los enlaces de hidrógeno. Los contaminantes, como el metal, pueden ser una ayuda para la asociación de micelas. A pesar de que las resinas son similares a las

fracciones de asfaltenos correspondientes en tamaño, tienen el menor grado de condensación. Por lo tanto, las resinas y otros compuestos aromáticos presentes en el aceite solo pueden asociarse con los asfaltenos.

Basándose en los preceptos del “Modelo de Yen”, Mullins presentó un nuevo modelo precedido de enormes estudios⁴². En este modelo, la arquitectura molecular predominante de asfaltenos consiste en un único sistema de anillo de hidrocarburo aromático policíclico de tamaño moderado (HAP) con sustituyentes alcanos periféricos⁴³. A su vez, expuso a los asfaltenos existentes en el petróleo crudo como tres estructuras jerárquicas distintas: (1) Moléculas, (2) Nanoagregados, (3) Racimos (Ver figura 9).

Figura 9.

Modelo Yen-Mullins. Este modelo muestra las estructuras moleculares y coloidales dominantes para los asfaltenos en solventes de laboratorio y aceites crudos.



Nota: Tomado y modificado de Mullins, O. C. et al. *Advances in Asphaltene Science and the Yen – Mullins Model*. (2012) doi:10.1021/ef300185p.

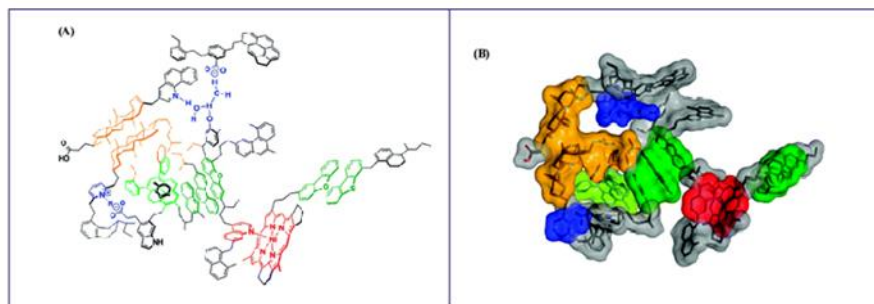
Las moléculas de asfaltenos pueden formar nanoagregados con una sencilla y desordenada pila de aromáticos policíclico (HAP). Las dimensiones de estos nanoagregados son de 2 nm, con agregación aproximadamente de 6. El exterior del nanoagregado está dominado por los

sustituyentes alcano. No necesitan resinas para formarse o estar suspendidas de manera estable. Los nanoagregados pueden formar grupos de nanoagregados de asfaltenos y los números de agregación se estima aproximadamente en ocho nanoagregados⁴².

Tanto el modelo de Yen como el modelo de Yen – Mullins, enfatizaron en la formación de grupos de moléculas mediante el apilamiento $\pi - \pi$ de anillos aromáticos. Gray⁴⁴ demostró que el apilamiento $\pi - \pi$ es solo una de las interacciones y propuso un nuevo modelo de ensamblaje supramolecular, para la agregación de asfaltenos.

Figura 10.

Representación Esquemática de un ensamblaje supramolecular compuesto por especies huésped en un agregado de asfaltenos representativos.



Nota: Las asociaciones entre moléculas están codificadas por colores en (A) la representación molecular y (B) la versión que llena el espacio: interacciones ácido-base y enlaces de hidrogeno (azul), complejo de coordinación metálica (rojo), una bolsa hidrofóbica (naranja), apilamiento $\pi - \pi$ (cara a cara verde oscuro; dentro de un caltrato que contiene tolueno, verde claro).

Las fuerzas impulsoras para este modelo, comprenden interacción ácida – base, enlaces hidrogeno, coordinación de metales, bolsas hidrofóbicas y apilamiento $\pi - \pi$. La implicación de

este ensamblaje supramolecular de moléculas complejas es que los nanoagregados dispersos de asfaltenos en solución son estructuras porosas⁴⁴.

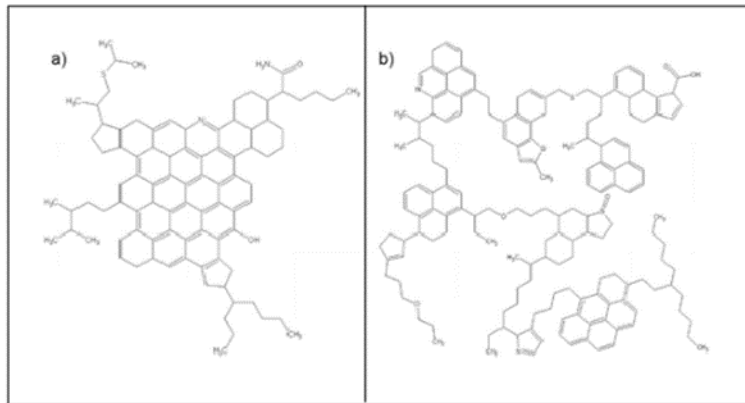
5.3 Estructura de Asfaltenos

La evolución de la visión molecular de la estructura de asfaltenos está llena de controversia y debates no concluyentes⁴⁵. Algo que, si es claro hoy en día, es el consenso de que los asfaltenos son monoméricos. Las técnicas como la despolarización de fluorescencia resuelta en el tiempo (TRFD) y la espectrometría de masas (MS), convergen en que los pesos moleculares de los monómeros de asfaltenos están entre 250 y 1200 g/mol, con un promedio de aproximadamente 750 g/mol⁴⁵.

Aún existen desacuerdos sobre la estructura molecular de los monómeros de asfaltenos, existen dos modelos, en cuanto a la naturaleza del núcleo aromático: (1) La arquitectura clásica de Isla, representa asfaltenos que consisten en un núcleo aromático, con cadenas laterales periféricas de alquilo⁴⁶, (2) También está el modelo Archipiélago, que a diferencia del modelo Isla, sugiere que los asfaltenos contienen más de un núcleo aromático, unido por puentes alquilo⁴⁷. Determinar los motivos estructurales, Islas vs Archipiélago, es una tarea extremadamente difícil y ha sido objeto del debate continuo. (Ver *figura 11*⁴⁸).

Figura 11.

(a) modelo continental o Isla, (b) modelo archipiélago



Nota: Tomado de Durand, E. et al. Effect of chemical composition on asphaltenes aggregation. Energy and Fuels 24, 1051–1062 (2010)

Los estudios realizados a muestras de asfaltenos mediante fotoionización de presión atmosférica de iones positivos por espectrometría de masas/ resonancia de ion ciclotrón de transformada de Fourier, arrojo una luz sobre la abundancia de motivos estructuras Islas y Archipiélagos. Estos resultados demostraron que los dos motivos estructurales coexisten en asfaltenos de petróleo. A diferencia de los motivos Islas que son fácilmente accesibles. Se requiere purificación de asfaltenos para detectar y caracterizar especies de Archipiélago por espectrometría de masas⁴⁵.

6. Estudios de Estabilidad de Crudos de Petróleo

Ser capaz de evaluar las propiedades de solubilidad de los asfaltenos es esencial para la evaluación del riesgo de precipitación de los mismos en campos petroleros¹³. La polaridad de la molécula de asfaltenos es causada a nivel local por electrones en átomos tales como azufre, oxígeno y nitrógeno. Los electrones pueden unirse cuando los metales están presentes en el catalizador, es decir, Ni, Co, Mo etc. La precipitación de los asfaltenos no solo depende de la composición del petróleo crudo, sino también de las propiedades fisicoquímicas de su fase dispersa y continua, así como de la temperatura y presión. Además, la solubilidad de la fracción de asfaltenos, también depende de la naturaleza de la fracción de maltenos en los aceites crudos, es decir, la concentración y la composición.

Aunque se han alcanzado progresos sobre el mecanismo real de los asfaltenos: (1) aglomeración, (2) Floculación y, (3) precipitación, aun no se ha entendido completamente. El mecanismo de desestabilización se mantiene principalmente en dos aspectos: (1) los asfaltenos se consideran suspensiones coloidales y, (2) los asfaltenos se describen como mezcla líquido-líquido. La estimación de parámetros para los modelos predictivos de la estabilidad en crudos, requiere de la determinación de datos experimentales que suelen ser muy costosos y cuya medición lleva largo tiempo. En este contexto, para la realización de pruebas preliminares de detección de estabilidad en el petróleo, generalmente se emplea principalmente el análisis SARA^{49 8}.

El riesgo de precipitación de los asfaltenos en el petróleo está fuertemente relacionado con cada uno de los porcentajes de cada una de las fracciones del análisis SARA (saturados, aromáticos, resinas y asfaltenos). Existe una base de indicación de estabilización de los asfaltenos,

respaldada por la presencia de resinas adsorbidas en la superficie; siendo la relación resina/asfaltenos un índice muy utilizado en la valoración de la estabilidad. La ruptura del equilibrio de fuerzas de atracción entre moléculas de resinas adsorbidas y las partículas de asfaltenos es la razón media de la floculación “desestabilización” de los asfaltenos coloidales en el aceite crudo; lo que indica riesgo de precipitación de asfaltenos, a medida que disminuyen las resinas. Por lo tanto, es de vital importancia determinar el inicio de la precipitación de asfaltenos, durante la producción de petróleo. En este sentido han surgido varias técnicas en la literatura, cada una con sus pros y contras. Sin embargo, desde el punto de vista práctico, cada quien debe decidir qué método sirve y puede llegar a ser el adecuado ⁹.

En el transcurso del tiempo se han propuesto diferentes métodos, índices, modelos y caracterizaciones muy finas mediante sucesivos experimentos de precipitación de los asfaltenos en el petróleo crudo, teniendo en cuenta parámetros de solubilidad. Dentro de los índices más utilizados se encuentra el índice de inestabilidad coloidal (CII)⁷, índice de clase de solubilidad (CSI)⁵⁰, índice de estabilidad¹⁰, diagrama de Stankiewics (SP)¹¹. Criterio de estabilidad de Sepúlveda (QQA)¹¹. Parámetro de Heithaus (parámetro P)⁵¹. Equivalencia de Tolueno (TE)⁷. Modelo de Compatibilidad (OCM)⁵².

6.1 Índice de Inestabilidad Coloidal (CII)

Según este índice, el petróleo se considera como un sistema coloidal compuesto por fracciones SARA. La estabilidad de los asfaltenos es una función de los saturados, aromáticos y resinas. Se expresa matemáticamente como la suma de asfaltenos y saturados (floculantes), dividida por la

suma de compuestos aromáticos y resinas (peptizadores). Los floculantes causan desestabilización de asfaltenos, mientras que los peptizadores estabilizan la fracción de asfaltenos⁷:

$$CII = \frac{(\text{Asfaltenos en \% w})+(\text{Saturados en \% w})}{(\text{Resinas en \% w})+(\text{Aromaticos en \% w})} \quad \text{Ecuación (5)}$$

Cuanto menor es el valor de CII, mayor es la estabilidad de los asfaltenos en los aceites crudos. Sin embargo, esta estabilidad debe considerarse como relativa al evaluar si un petróleo crudo es más estable que otro. Según Asomaning y Watkinson¹⁰, si $CII > 0.9$, la fracción de asfaltenos tiende a ser inestable dentro del petróleo crudo. Cuando $0.7 \leq CII \leq 0.9$, la estabilidad de los asfaltenos es incierta, mientras que si $CII \leq 0.7$, la fracción de asfaltenos es estable.

6.2 Índice de Estabilidad Coloidal (CSI)

Según este índice, la fracción de asfaltenos de los crudos inestables es más polar en comparación con la de los crudos estables. En fraccionamiento SARA y la polaridad se tienen en cuenta como se muestra en la ecuación⁵⁰:

$$CII = \frac{(\epsilon^{Asf})(\text{Asfaltenos en \% w})+(\epsilon^{Sat})(\text{Saturados en \% w})}{(\epsilon^{Res})(\text{Resinas en \% w})+(\epsilon^{Aro})(\text{Aromaticos en \% w})} \quad \text{Ecuación (6)}$$

Donde existen valores estándar de constantes dieléctricas para todos los crudos. Si $CSI > 0.95$, el petróleo crudo es inestable y los asfaltenos precipitaran. Si $CSI < 0.95$, el petróleo crudo es estable y la precipitación de asfaltenos es improbable.

6.3 Índice de Estabilidad

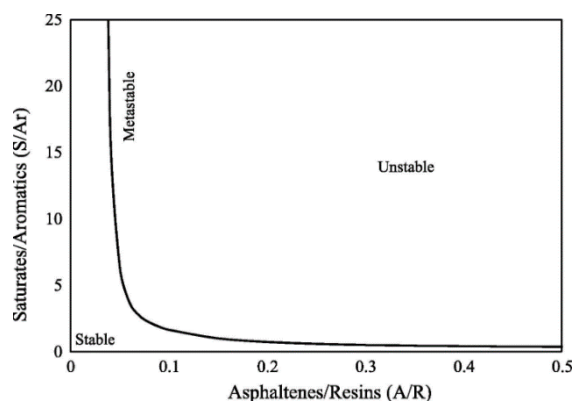
La relación asfaltenos/resinas (A/R) se usa ampliamente, porque ambas fracciones son pesadas, no volátiles y se pueden cuantificar con precisión. Las resinas se consideran peptizadores naturales de asfaltenos que las mantienen en solución. Asomaning y Watkinson¹⁰ informaron que si la relación A/R es inferior a 0.35, entonces el petróleo crudo será estable⁵³.

6.4 Gráfico de Stankiewics (SP)

Este método tiene en cuenta las fracciones SARA. Si se usa conjuntamente junto con CII, se obtienen resultados más confiables. En este método, las relaciones de saturados/ aromáticos se representa frente a asfaltenos/ resinas, como se muestra en la *figura 12*. Para evaluar la tendencia de precipitación de los asfaltenos. Dos regiones están claramente marcadas como estables o inestables^{11 53}.

Figura 12.

Criterio de estabilidad de Stankiewics.



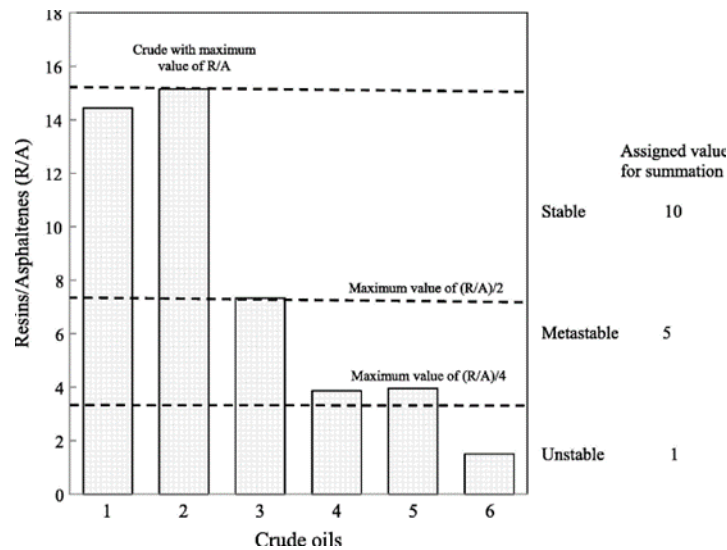
Nota: Tomado de Guzmán, R., Ancheyta, J., Trejo, F. & Rodríguez, S. Methods for determining asphaltene stability in crude oils. Fuel 188, 530–543 (2017).

6.5 Criterio de estabilidad de Sepúlveda (QQA: análisis cualitativo – cuantitativo)

Sepúlveda y colaboradores, llevaron a cabo un análisis cualitativo – cuantitativo, utilizando una serie de relaciones del fraccionamiento SARA: R/ A, S/ A, Ar/ A, Ar/ (S/A), (S/ Ar /A), (R/A) / (S/ Ar). Cada relación es dividida en tres zonas con un valor específico como se muestra en la *Figura 13*. Para R/A^{11 53}.

Figura 13.

Ejemplo de clasificación de zonas de estabilidad con el criterio de Sepúlveda (QQA: análisis cualitativo-cuantitativo).



Nota: Tomado Guzmán, R., Ancheyta, J., Trejo, F. & Rodríguez, S. Methods for determining asphaltene stability in crude oils. Fuel 188, 530–543 (2017).

Los valores más altos para los crudos en cada relación serán tomados en cuenta para futuros cálculos. Cada línea discontinua entre la región estable, metaestable e inestable se obtiene de la siguiente manera: (a) Para la zona estable y metaestable, el valor máximo es dividido por dos, (b)

Para la zona metaestable e inestable, el máximo del valor se divide por cuatro. Luego, los valores se asignan a diferentes zonas, es decir, zona estable = 10, zona metaestable = 5 e inestable = 1. Suma de valores determinados para cada relación se consigue. En el caso de la relación $Ar/(S/A)$, la zona inferior adquiere un valor de 10,5 para la zona media y 1 para la zona superior.

Siguiendo este procedimiento para todas las relaciones, se realiza una suma para cada muestra de petróleo. Si la suma es superior a 30, entonces el petróleo crudo es estable. Para valores entre 15 y 30, el petróleo crudo se considera metaestable, y si la suma es inferior a 15, el petróleo crudo es inestable.

6.6 Parámetro Heithaus (parámetro P)

Este método fue desarrollado por Heithaus⁵¹, para predecir la estabilidad de los crudos pesados y los asfaltos. Una solución de petróleo crudo o betún en un disolvente aromático (tolueno o α – metil naftaleno) se titula con disolvente parafínico (n – heptano, iso-octano o n- hexadecano) hasta que se produce la precipitación de asfaltenos⁵¹. En este punto se obtiene una relación de floculación (FR) teniendo en cuenta la fracción volumétrica del disolvente aromático en relación con el volumen total del disolvente (V_S) y el volumen del titulador (V_T) de acuerdo con la ecuación (7):

$$FR = \frac{V_S}{V_S + V_T} \quad \text{Ecuación (7)}$$

La concentración del petróleo o betún en un solvente aromático (C) está dada por la masa de petróleo crudo dividida por el volumen total de solvente aromático y titulante de la siguiente manera:

$$C = \frac{W_A}{V_S + V_T} \quad \text{Ecuación (8)}$$

Donde V_S es el volumen del solvente, V_T el volumen del titulador y W_A la masa de aceite en la muestra.

6.7 Modelo de Compatibilidad de Aceite (OCM)

En este modelo de compatibilidad, se requieren dos parámetros: (1) número de insolubilidad (I_N), número de mezcla de solubilidad (S_{BN})^{53 7}. Las soluciones contienen n – heptano y tolueno en diferentes proporciones, que se agregan al petróleo crudo para determinar el inicio de la precipitación de asfaltenos. Se necesitan dos pruebas: (1) prueba de dilución de heptano y (2) prueba de equivalencia de tolueno. Al aplicar esta prueba, la mezcla de n – heptano/ tolueno se llama “líquido de prueba” y para cada relación de volumen de aceite (V_{aceite}), a líquido prueba (V_{TL}), se determina el volumen mínimo de tolueno para mantener los asfaltenos en solución⁷. A partir de estos valores, se obtiene la siguiente ecuación lineal (9), donde el coeficiente lineal representa el número de insolubilidad (I_N):

$$S_{BN} = I_N \left[1 + \frac{100}{H_D} \right] = I_N \left[1 + \frac{V_{HD}}{V_{oil}} \right]_{V_T=0} \quad \text{Ecuación (9)}$$

La intersección del eje x se evalúa directamente, determinando el volumen máximo de n – heptano que causa la precipitación de asfaltenos, es decir, la llamada prueba de dilución de heptano ($S_{BN} = I_N$). La relación S_{BN} / I_N se utiliza para predecir la estabilidad del aceite puro mediante el cual un alto valor corresponderá a los aceites crudos más estables.

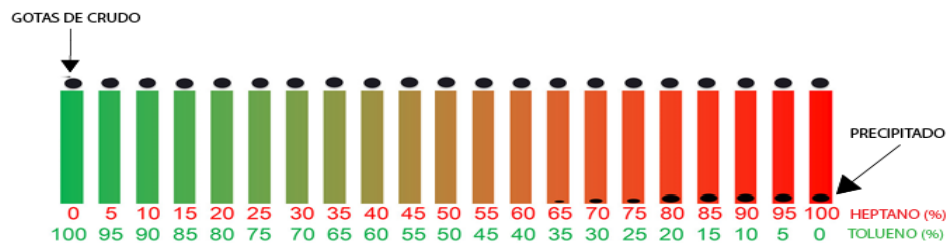
6.8 Índice de Clase de Solubilidad de Asfaltenos (ASCI)

Una de las características de la estabilidad de los asfaltenos, es su fuerte sensibilidad a la calidad del disolvente y su muy baja sensibilidad a la dilución. Si una cantidad dada de una fracción de asfaltenos flocula en una proporción dada de una mezcla de tolueno y *n* – heptano “Heptol”, estos asfaltenos aún flocularían incluso si se diluyen con 10 o 100 volúmenes de la misma mezcla de Heptol. Una modificación muy débil de la relación tolueno/ *n* – heptano puede modificar la solubilidad de una fracción de asfaltenos de una manera espectacular¹³.

Una descripción más apropiada de la capacidad de una fracción de asfaltenos para permanecer soluble debería ser su clase de solubilidad. Se puede llegar a caracterizar el comportamiento de una fracción de asfaltenos por un tipo de “Índice de clase de solubilidad de asfaltenos” (ASCI). Para medir este índice de una fracción de asfaltenos, se agregan unas gotas de

Figura 14.

Representación esquemática de ASCI, indicando cada una de las mezclas de solventes en sus mediciones.



Nota: Adaptado de Passade-Boupat, N., Gingras, J. P., Desplobins, C. & Zhou, H. Could the Asphaltene Solubility Class Index Be Used as the ‘wax Appearance Temperature’ of Asphaltenes? Illustration through the Study of the Polydispersity of PetroPhase 2017 Asphaltenes. Energy and Fuels 32, 2760–2768 (2018).

Un líquido que contiene asfaltenos a caracterizar a viales con diferentes mezclas de Heptol, como se describe en la *Figura 14*. La cantidad de petróleo crudo o fracción de asfaltenos debe ser insignificante en comparación con la cantidad de Heptol¹³.

Los viales se agitan manualmente y se almacenan a 45° C, durante aproximadamente 48h antes de la observación a simple vista. Se adopta una temperatura de 45° C para evitar interferencias con otros posibles depósitos, como las ceras, para evitar cualquier efecto sobre la solubilidad de los asfaltenos. Si la precipitación de asfaltenos se observa en el vial 0, en tolueno puro (100% tolueno), se obtiene un $ASCI=0$, en el caso de que los precipitados de “asfaltenos” solo se vean en el vial 20 en *n* – heptano puro (100% *n* – heptano), se obtiene un valor de $ASCI=20$ ¹³. Como se observa en la *Figura 14*. Se denota la aparición de un precipitado en la mezcla de Heptol (65 heptano/ 35 tolueno), se obtiene un $ASCI=13$.

7. Quimiometría

7.1 ¿Qué es Quimiometría?

El desarrollo de la disciplina de la quimiometría está fuertemente relacionado con el uso de las computadoras en el análisis químico, algunos grupos analíticos en la década de 1970 ya estaban trabajando con métodos estadísticos y matemáticos que se atribuyen hoy en día a los métodos quimiométricos⁵⁴. La notación de quimiometría, fue introducida en 1972 por el sueco Svante Wold y el estadounidense Bruce R. Kowalski; en 1974 la fundación de la sociedad internacional de

quimiometría condujo la primera descripción de esta disciplina. Una definición actual y real de la quimiometría es: “La disciplina química que utiliza métodos matemáticos y estadísticos, (a) para diseñar o seleccionar procedimientos y experimentos de medición óptimos, y (b) para proporcionar la máxima información química mediante el análisis de datos químicos.”

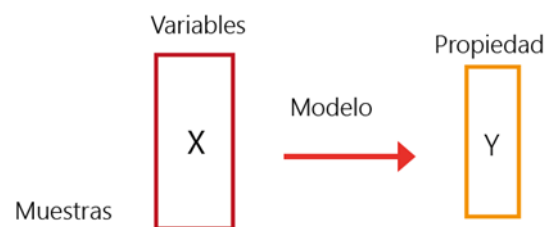
Para comprender la multidimensionalidad de los problemas analíticos, los datos analíticos se pueden organizar como una matriz X de n objetos (filas) y características definidas como p (columnas). Los objetos pueden ser muestras, moléculas, materiales, etc. Las características o variables típicas de esos objetos serán patrones elementales, espectros, características estructurales o propiedades físicas⁵⁴.

7.2 Métodos para la construcción de modelos predictivos

Una vez clasificadas las muestras, uno de los objetivos generales de la quimiometría “Construcción de modelos predictivos”, es posible gracias a varios métodos; el objetivo de estos métodos está en encontrar la forma como se relacionan dos matrices de datos, en este caso, una espectroscópica y una de propiedades.

Figura 15.

Esquema fundamental de la Quimiometría



El anterior esquema responde a la siguiente ecuación matricial

$$X * B = Y \quad \text{Ecuación (10)}$$

Donde la única incógnita es B o matriz regresora, pero es determinada como:

$$B = (X^t * X)^{-1} * X^t * Y \quad \text{Ecuación (11)}$$

Principalmente tres técnicas son usadas para esta finalidad y lo único que las diferencia entre sí son las variables que componen la matriz espectroscópica o X en la figura anterior; estas técnicas son: MLR (Multiple Linear Regression), PCR (Principal Components Regression), y PLSR (Partial Least Squares Regression).

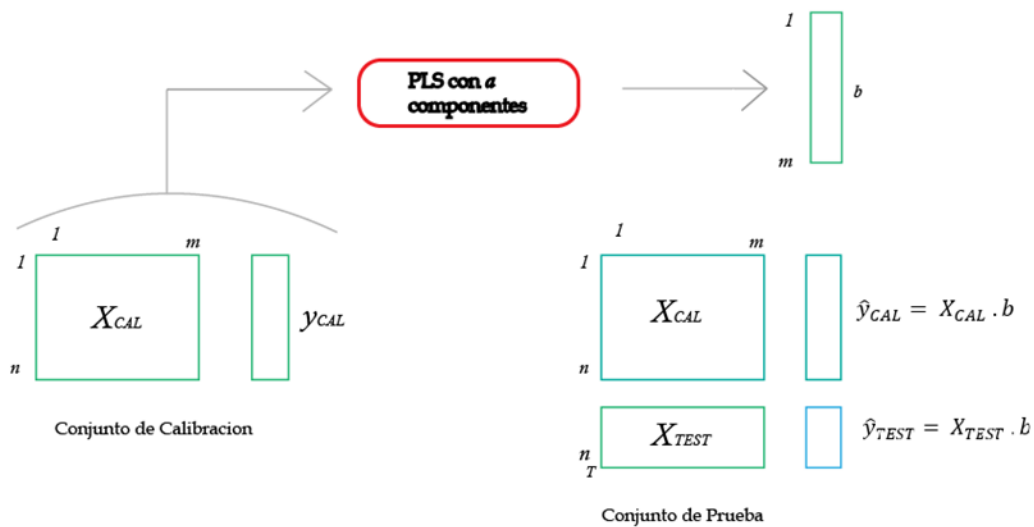
7.3 Regresión por Mínimos Cuadrados Parciales (PLSR)

PLSR ha demostrado ser un método muy versátil para el análisis de datos multivariados y el número de aplicaciones está aumentando constantemente en campos de investigación como la Bioinformática, el aprendizaje automático y la quimiometría, siendo la misma una proyección a estructuras latentes. PLS tiene en cuenta que en la mayoría de los casos las dos matrices tienen diferente dimensionalidad, por lo cual calcula otro tipo de variables conocidas como factores latentes. PLSR en su forma original no tiene implementación de selección de variables, ya que el enfoque del método es encontrar el subespacio lineal relevante de las variables explicativas, no las variables en sí, pero se han propuesto una gran cantidad de métodos para la selección de variables en PLSR⁵⁵. La selección de variables puede parecer innecesaria ya que la ponderación ascendente

y descendente de las variables es una propiedad inherente del estimador PLS. Se motiva la selección de variables para mejorar el rendimiento de estimación/predicción⁵⁶.

Figura 16.

PLS como método de regresión lineal múltiple para la predicción de una propiedad y a partir de las variables X_1, \dots, X_m , aplicando coeficientes de regresión b_1, \dots, b_m (datos centrados en la media).



Nota: A partir de un conjunto de calibración, el modelo PLS se crea y aplica a los datos de calibración y a los datos de prueba. Ilustración tomada y modificada de Brown, S. D. Book Reviews: Introduction to Multivariate Statistical Analysis in Chemometrics. Applied Spectroscopy vol. 64 (2010).

El modelo resultante predice una propiedad y , de las variables dependientes originales x_1 y x_m . El modelo lineal contiene coeficientes de regresión b_1, \dots, b_m (datos centrados en la media) y una intersección b_0 . A partir de un conjunto de calibración y a los datos de prueba. Durante el

desarrollo del modelo, se calcula un número relativamente pequeño de componentes PLS (variables latentes lineales intermedias) que se utilizan internamente para la regresión. El número de componentes PLS determina la complejidad del modelo y puede optimizarse para un alto rendimiento de predicción⁵⁶.

7.3.1 Aspectos Matemáticos

PLS a menudo se introduce y explica como un algoritmo numérico, maximiza una función objetivo bajo ciertas restricciones. La función objetivo es la covarianza entre los “Scores” x e y , y la restricción suele ser la ortogonalidad de los Scores. Se asumen datos multivariados de x e y dado por la matriz \mathbf{X} de dimensión $(n \times m)$, y la matriz \mathbf{Y} de tamaño $(n \times q)$. PLS se basa en la conjetura, que los datos en las filas se obtienen de los mismos n objetos, y que X contiene la información de m características (variables predictoras) e Y describe q propiedades (variables de respuesta). Para una notación más conveniente, suponemos que las columnas de X e Y , se han centrado.

$$Y = XB + E \quad \text{Ecuación (12)}$$

Entre las variables x e y , usando una matriz $m \times q$, matriz B de coeficientes de regresión, y una matriz de error E . La forma de esta relación es:

$$Y = XB + e \quad \text{Ecuación (13)}$$

Con coeficientes de regresión b y un término de error e_1 . En lugar de encontrar esta relación directamente, tanto X como Y se modelan mediante variables latentes lineales de acuerdo con los modelos de regresión:

$$X = TP^T + E_X \quad \text{Ecuación (14)}$$

Y.

$$Y = UQ^T + E_Y \quad \text{Ecuación (15)}$$

Con las matrices de error E_X y E_Y . Las matrices T y U (matrices de puntajes) así como las matrices P y Q (matrices de carga) tiene a columnas, donde $a \leq \min(m, q, n)$ siendo este el número de componentes PLS. Los Scores-x en T son combinaciones lineales de las variables x y pueden ser consideradas como buenos resúmenes de las variables x. Los Scores-y y U son combinaciones lineales de las variables-y, y pueden considerarse como buenos resúmenes de las variables- y.

A continuación, t_j , u_j , p_j y q_j denotan las columnas j_{th} de T, U, P, Q respectivamente ($j = 1, \dots, a$), además los Scores de x e y están conectados por la relación lineal interna.

$$u_j = d_j t_j + h_j \quad \text{Ecuación (16)}$$

Con h_j como los residuales y d_j los parámetros de regresión. Si, por ejemplo, la relación entre u_1 y t_1 es fuerte (si los elementos de h_j son pequeños), entonces el puntaje x del primer componente PLS es bueno para predecir Scores-y y finalmente para pronosticar datos y. Por lo general, se usa más de un componente PLS para modelar Y por X; el número óptimo de componentes PLS puede estimarse por CV. La relación entre los Scores luego se convierte:

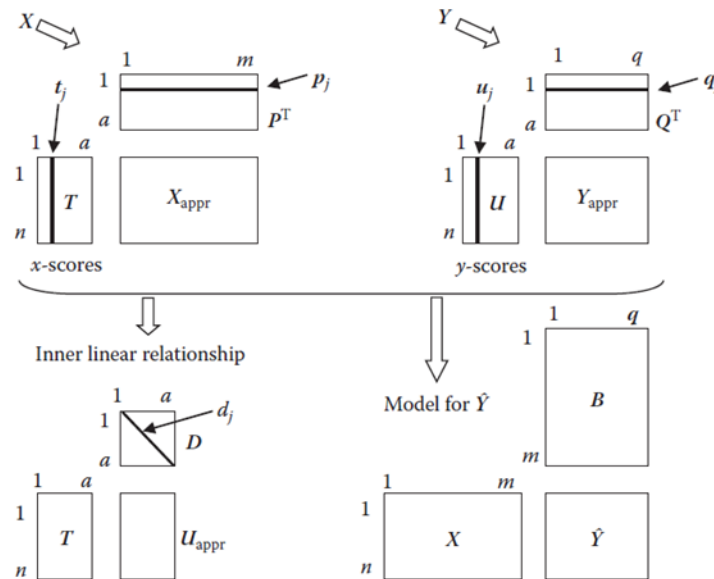
$$U = TD + H \quad \text{Ecuación (17)}$$

Donde D siendo una matriz diagonal con elementos d_1, d_2, \dots, d_a , y H la matriz residual con las columnas h_j . En la *Figura 17*, se ofrece una visión general esquemática de las relaciones

matriciales. La Ecuación (17) también motiva el nombre “PLS”, porque solo se usa información parcial de los datos x e y para la regresión.

Figura 17.

Matrices en PLS.



Nota: Tomada de Brown, S. D. Book Reviews: Introduction to Multivariate Statistical Analysis in Chemometrics. Applied Spectroscopy vol. 64 (2010).

Como se señaló anteriormente, la covarianza como criterio para las variables latentes combina una alta varianza de X , así como una alta correlación entre X e Y . De nuevo, depende de cómo se estime la covarianza ⁵⁶. En el caso clásico, la covarianza entre dos vectores de Scores t y u se estima mediante la covarianza de la muestra $\frac{t^T u}{n-1}$, pero también se puede usar estimadores robustos. Dado que el problema de maximización no sería único, se necesita una restricción en los vectores de Scores, que generalmente se toma como $\|t\| = \|u\| = \|1\|$ (longitud 1). Los vectores

de Scores resultan de la proyección de las matrices de datos X e Y en los vectores de carga. Sería lógico ahora usar los vectores de carga de las matrices P y Q de las ecuaciones. Sin embargo, por razones técnicas, se usan otros vectores de carga, digamos un vector w para las variables-x, $t = Xw$, y c para las variables-y; $u = Yc$.

El problema de maximización con restricciones por lo tanto es:

$$\text{cov}(X_w, Y_c) \rightarrow \max \quad \|t\| = \|X_w\| = 1 \quad Y \quad \|u\| = \|Y_c\| = 1 \quad \text{Ecuación (18)}$$

Donde “cov” es la covarianza de muestra. Las restricciones tienen que ser la longitud de los vectores de puntuación es igual a 1 o la longitud de los vectores de peso w y c, son iguales a 1. Las soluciones del problema de maximización son los primeros vectores de puntuación t_1 y u_1 para el espacio x e y, respectivamente. Para los siguientes vectores de puntuación, se maximiza el mismo criterio, pero se deben introducir restricciones adicionales. Por lo general, estas restricciones son la Ortogonalidad de los vectores de puntuación anterior:

$$t_j^T t_i = 0 \quad Y \quad u_j^T u_i = 0 \quad \text{para } 1 \leq j \leq l \leq a \quad \text{Ecuación (19)}$$

Los puntajes no correlacionados se obtienen por la mayoría de los otros algoritmos (Kernel, NIPALS, SIMPLS, O-PLS) y desde cada vector de puntuación adicional cubre una nueva variabilidad, esto podría ser preferible para fines de predicción. El primer componente PLS se encuentra de la siguiente manera: dado que tratamos con la varianza de la muestra, el problema de maximización, puede escribirse como maximización de:

$$t^T u = (Xw)^T Yc = w^T X^T Yc \rightarrow \max \quad \text{Ecuación (20)}$$

Bajo las mismas restricciones de los vectores de longitud 1. Las soluciones para w y c, se encuentran fácilmente mediante la descomposición de valores singulares (SVD) de $X^T Y$. De

acuerdo con eso, entre todas las direcciones posibles w y c , se alcanza el máximo de la ecuación (18) para los vectores $w = w_1$ y $c = c_1$ correspondientes al mayor valor singular de $X^T Y$.

7.3.2 SIMPLS Algoritmo para PLS

El nombre de este algoritmo propuesto por de Jong (1993) se origina a partir de la “implementación directa de una modificación estadísticamente inspirada del método PLS de acuerdo con un concepto simple”. Este algoritmo maximiza directamente el problema inicial de la Ecuación (18):

$$\text{cov}(X_w, Y_c) \rightarrow \max \quad \|t\| = \|X_w\| = 1 \quad Y \quad \|u\| = \|Y_c\| = 1 \quad \text{Ecuación (18)}$$

Bajo la restricción de ortogonalidad de los puntajes t para diferentes componentes. El primer componente PLS de SIMPLS es idéntico al resultado de NIPALS o el algoritmo de Kernel. En SIMPLS no hay deflación de los datos centrados en las matrices X y Y realizados, pero la deflación se lleva a cabo por la matriz de covarianza, o más precisamente, la matriz de productos cruzados $S = X^T Y$, entre datos- x y datos- y . El pseudocódigo para el algoritmo SIMPLS es el siguiente:

1. Se inicia con $S_0 = X^T Y$ y se repiten los pasos 2 a 6 para $j = 1, \dots, a$
2. Si $j = 1$, $S_j = S_0$; si $j > 1$, $S_j = S_{j-1} - P_{j-1}(P_{j-1}^T P_{j-1})^{-1} P_{j-1}^T S_{j-1}$
3. Estima w_j como el primer (left) vector singular de S_j
4. $w_j = w_j / \|w_j\|$
5. $t_j = X w_j$
6. $t_j = t_j / \|t_j\|$
7. $P_j = X_j^T t_j$
8. $P_j = [P_1, P_2, \dots, P_{j-1}]$

Los pesos resultantes w_j y Scores t_j se almacenan como columnas en las matrices W y T , respectivamente. Tenga en cuenta que la matriz W difiere ahora de los algoritmos anteriores, porque los pesos están directamente relacionados con X y no con las matrices deflactadas. El paso 2 explica la restricción de ortogonalidad de los Scores t_j a todos los vectores de Scores anteriores, porque la búsqueda se realiza en el componente ortogonal de S_{j-1} . El paso 3 maximiza directamente el problema inicial (Ecuación (18)), compárelo con la Ecuación (14). Las puntuaciones en el paso 4 se obtienen proyectando directamente con X en la dirección óptima, y las cargas en el paso 5 se obtienen por regresión OLS para el modelo (Ecuación (14)).

7.4 Validación Cruzada de Monte Carlo

Uno de los principales problemas del método de Validación Cruzada (CV), es que tiende a incluir componentes innecesarios en el modelo y hace que el modelo sea más grande de lo que debería ser. Por lo tanto, el modelo con el número de componentes determinado “*leave-one-out*” (CV), a menudo funciona bien en la calibración, pero presenta un potencial pobre en la predicción⁵⁷.

Para elegir correctamente la dimensión del modelo de calibración, se introduce un método sencillo y de gran efectividad llamado Validación cruzada de Monte Carlo (MCCV). La validación cruzada de Monte Carlo desarrollada en este documento es un método asintóticamente consistente para determinar el número de componentes de calibración. Puede evitar un modelo grande e innecesario y, por lo tanto, disminuye el riesgo de un ajuste excesivo para el modelo de calibración⁵⁸.

El paso fundamental después de que los datos estén disponibles, es determinar la dimensión del número de componentes. Validación cruzada de Monte Carlo MCCV es un método simple y

efectivo: divide aleatoriamente las muestras en dos partes $S_c(i)$ (de tamaño de n_c) y $S_v(i)$ (de tamaño de n_c) y se repite el procedimiento para N veces ($i = 1, 2, \dots, N$). Se define el criterio repetido de MCCV ⁵⁸:

$$\text{MCCV}_{n_v}(k) = \frac{1}{N n_v} \sum_{i=1}^N \left\| y_{S_v(i)} - \hat{y}_{k * S_v(i)} \right\|^2 \quad \text{Ecuación (21)}$$

Mediante el método de Monte Carlo, la cantidad de complejidad de cálculo se puede reducir sustancialmente. Teóricamente, cuantas menos muestras se usen en la calibración del modelo, más tiempos de repetición son necesarios y $N = n^2$, en general, es suficiente para que MCCV, funcione también como ⁵⁸ CV_{n_v} “Validación Cruzada leave-one-out”. Cabe señalar que, después de que MCCV obtiene las variables, todas las muestras se utilizan para ajustar al modelo⁵⁹. Para adquirir $\hat{\beta}_{\alpha^*}$, donde tenemos en la ecuación (21) $\hat{y}_{k * S_v(i)}$, que es el vector de respuesta previsto, utilizando PLS (con k componentes). La calidad del modelo establecido se evalúa de acuerdo con su capacidad de predicción, ya que el papel principal del modelado de regresión es predecir muestras futuras desconocidas⁵⁹.

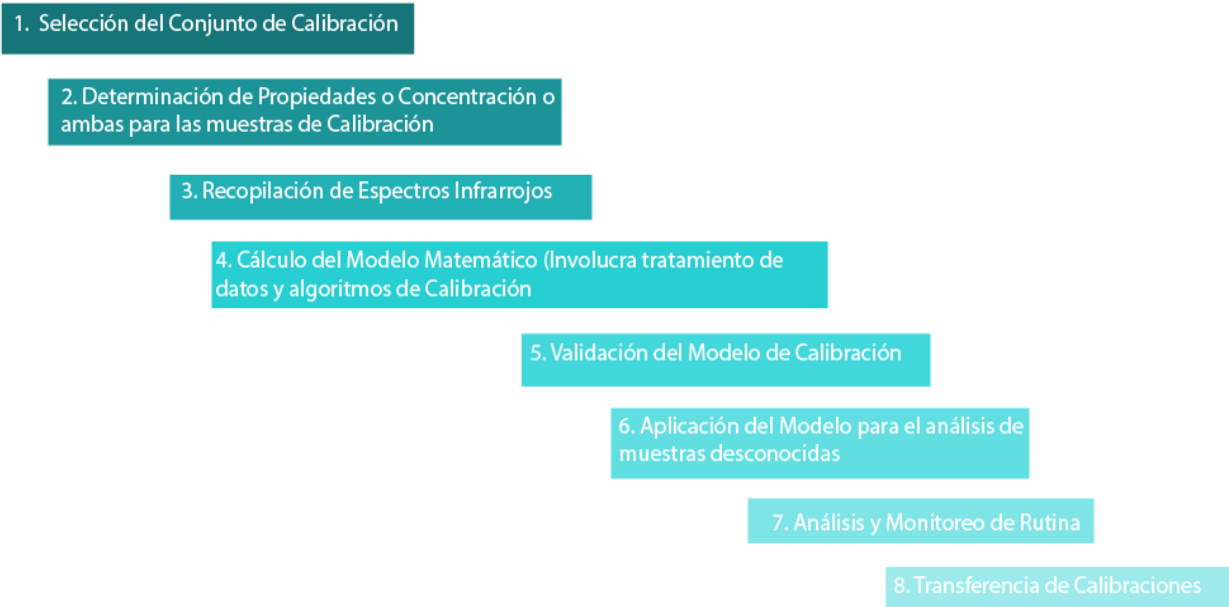
8. Construcción de Modelos de Calibración (Aplicado a IR-MIR)

El objetivo de los métodos de calibración Multivariable es construir modelos de calibración que sean capaces de predecir el valor de la propiedad estudiada con las absorbancias de un conjunto de muestras de referencia conocida en más de una longitud de onda o frecuencia. El modelo de calibración Multivariante resultante se aplica al análisis de espectros de muestras desconocidas

para proporcionar una estimación de la concentración de componentes o valores de propiedad para la muestra desconocida⁶⁰

Figura 18.

Descripción General de la calibración Multivariante.



Nota: ASTM E1655-05. Standard Practices for Infrared Multivariate Quantitative Analysis. ASTM Int. 05, 29 (2012).

8.1 Evaluación y Optimización de Modelos de Calibración (Estadísticas)

Se utilizan diversas estadísticas para evaluar y optimizar el rendimiento de los modelos de calibración Multivariable, estas estadísticas generalmente se aplican solo a datos del conjunto de calibración; no deben confundirse con las estadísticas que se utilizan para validar el modelo, que se calculan en función de un conjunto de validación analizado por separado⁶⁰.

8.1.1 Error estándar de calibración

Si \hat{y} son los valores estimados para las muestras de calibración, e y son los valores de referencias correspondientes, entonces e es el vector de error de calibración definido por:

$$e = \hat{y} - y \quad \text{Ecuación (22)}$$

Los errores de predicción incluyen contribuciones de errores en los valores de referencia para el conjunto de calibración, errores espectrales en los espectros del conjunto de calibración y errores del modelo (usando número incorrecto de variables).

8.1.2 Error estándar de calibración SEC

$$SEC = \sqrt{\frac{e^t e}{d}} = \sqrt{\frac{1}{n_{cv}} \sum_{i=1}^{n_t} (\hat{y}_i - y_i)^2} \quad \text{Ecuación (23)}$$

Como se muestra en la ecuación (20), d es el número de grados de libertad en el modelo de calibración, d es típicamente igual a $n - k$, donde n es el número de muestras de calibración, y k es el número de variables (variables latentes PLS). El error estándar de calibración se usa para estimar el acuerdo esperado entre los valores estimados, usando los modelos de calibración y los valores que se medirían con el método de referencia.

8.1.3 Error estándar de validación cruzada SECV

$$SECV = \sqrt{\frac{1}{n_{cv}} \sum_{i=1}^{n_{cv}} (\bar{y}_i - y_i)^2} \quad \text{Ecuación (24)}$$

En las ecuaciones (23) y (24), y_i junto con \hat{y}_i son los valores de referencia de y y estimados del conjunto de muestras de calibración (validación), respectivamente. n_t Y n_{CV} son la cantidad de muestras para entrenamiento y validación cruzada respectivamente.

Existen una variedad de pruebas estadísticas para evaluar los modelos de calibración, dentro de ellas algunas que destacan en este trabajo son: Valor P, para coeficientes lineales β_1 y cuadráticos β_2 , también la F_{tab} y F_{cal} .

9. Metodología

Algunas propiedades que describen la composición del petróleo crudo, como el contenido de asfaltenos, están estrechamente relacionadas con los grupos funcionales, cuyas señales aparecen en regiones infrarrojas particulares. Comprender estas relaciones fue motivación del presente trabajo. En esta investigación se desarrolla de una forma original, un modelo para predecir la estabilidad de 82 muestras de crudos colombianos. Este modelo se realizó por análisis quimiométrico multivariado, de la propiedad (ASCI) índice de clase de solubilidad de asfaltenos (propiedad- Y) y, las mediciones de los espectros infrarrojos por FT-IR/ATR (propiedad- X). En el desarrollo del modelo se utilizó la regresión parcial de mínimos cuadrados (PLS-R).

El presente trabajo muestra ventajas preponderantes para la caracterización del petróleo crudo respecto a los análisis convencionales, ya que permite evaluar y clasificar con la utilización de pequeñas cantidades de muestra, pretratamiento mínimo de la misma y procedimientos de análisis rápidos (2 a 3 minutos). Este trabajo de investigación tiene lugar bajo el convenio de

cooperación tecnológica Universidad Industrial de Santander –Instituto Colombiano del Petróleo UIS-ICP.

9.1 Reactivos y Muestras

Un conjunto de 82 muestras de aceites crudos, deshidratados y desalados, de diferentes áreas de exploración de Colombia, suministrados por el Instituto Colombiano de Petróleo (ICP), se utilizaron para desarrollar un modelo predictivo, para la determinación de estabilidad mediante la solubilidad de asfaltenos.

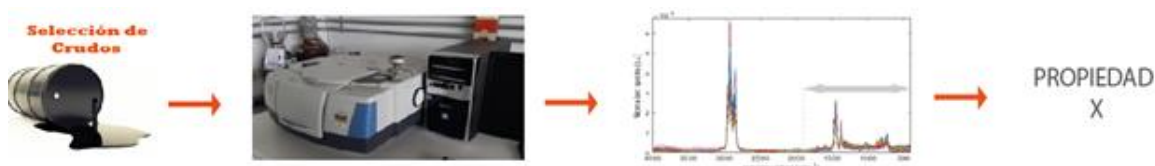
Se emplearon tolueno (grado analítico, J.T. Baker) y n-heptano (grado analítico, Merck) para preparar las soluciones de n-heptano/tolueno para evaluar la estabilidad de los crudos.

9.2 Espectroscopía Infrarroja

Las mediciones de FT-IR/ATR se realizaron para identificar y cuantificar grupos funcionales, los espectros de crudos se recolectaron y evaluaron en 7468 puntos de número de onda, igualmente espaciados, entre 4000 y 400 cm^{-1} , usando un espectrofotómetro Thermo Scientific Nicolet iS50 FTIR, equipado con un cristal de diamante con un medidor de interferencia Michelson y una unidad ATR.

Figura 19.

Descripción Metodología (primera parte) FT-IR/ATR en la región MIR.



9.2.1 Adquisición Espectral

La medición de Espectroscopia Infrarroja se realizó con una pequeña porción del petróleo crudo de cada una de las 82 muestras de crudos colombianos. Los parámetros promedio de medición se calcularon sobre la base de los espectros y mediante el uso de software con el trazado de línea base para el desarrollo de la relación señal/ruido, que fue calculada en el rango espectral de 2300 – 1800 cm^{-1} , cuyo valor más alto fue 11, correspondiente a los siguientes parámetros de adquisición de espectros: 128 escaneos, añadidos conjuntamente durante la adquisición de cada conjunto de datos, con una velocidad del espejo móvil del interferómetro de Michelson de 0.32 cm/s.

9.2.2 Análisis Espectral

Debido a que la espectroscopia Infrarroja tiene una sensibilidad relativamente baja, las diferencias de las intensidades en los espectros de cada una de las 82 muestras son muy pequeñas, lo cual se tornan difíciles de detectar (*Figura 20*). Además, las mismas muestras de aceites crudos generalmente también son bastantes similares con respecto a la estructura (grupos funcionales) los cuales se describen en la *Figura 21*.

En la *Figura 20*, podemos observar una flecha de color gris, la cual nos indica la selección de la señal predictiva, utilizada en el modelo de calibración ($1900 - 400 \text{ cm}^{-1}$). En el intervalo seleccionado, hay bandas de grupos funcionales representativos de asfaltenos, relacionados, por definición con la estabilidad del petróleo crudo (recuadro rojo *Figura 20*).

Figura 20.

Espectros FTIR-ATR de las 82 muestras de petróleo crudo. La flecha muestra la señal predictiva utilizada en el modelo de calibración ($1900-400 \text{ cm}^{-1}$) desarrollada en Matlab R2019a.

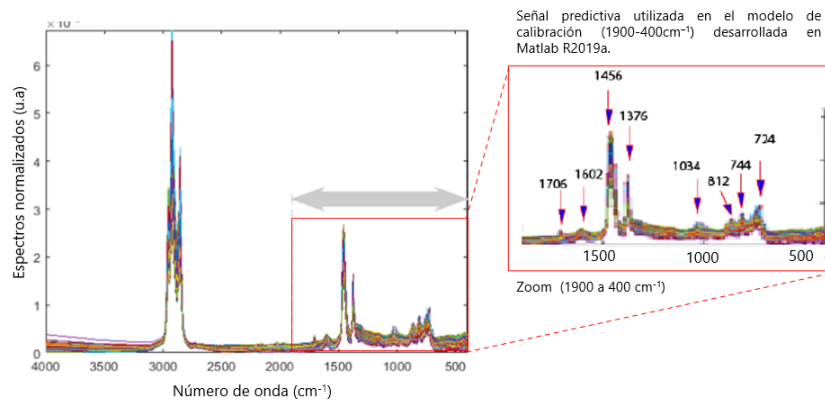
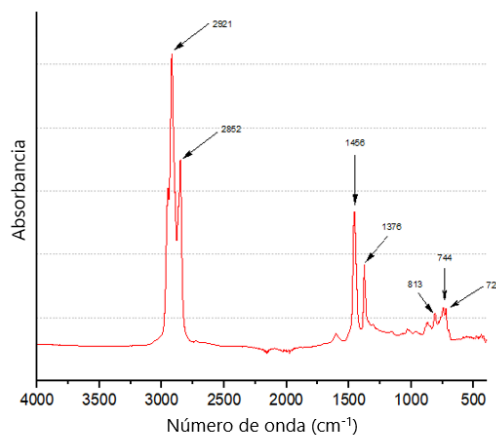


Figura 21.

Espectro FT-IR/ATR muestra de crudo colombiano, bandas características.

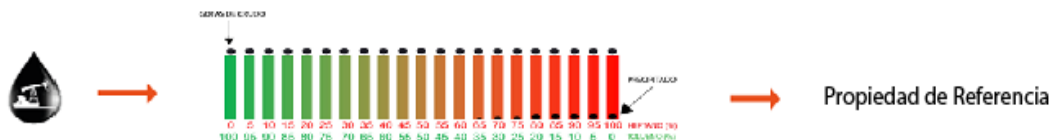


Los espectros de IR-FT/ATR de las 82 muestras de petróleo crudo, presentan las siguientes bandas de absorción relacionadas de la siguiente manera: se presentan bandas en 1456 y 1376 cm^{-1} que corresponden a la deformación simétrica de R_2CH_2 y RCH_3 respectivamente, en 722 cm^{-1} , se presenta una señal característica de una cadena de alquilo que consta de más de 4 grupos metileno, en 744 y 813 se observan bandas correspondientes a una serie de vibraciones de flexión $\text{C}_{\text{ar}} - \text{H}$. Adicional a las bandas de absorción ya descritas, también se presentan otras señales, que, aunque no participan en la señal predictiva del modelo de calibración, están presentes en las muestras del petróleo crudo como se muestra en la Figura 22: las regiones entre 2952 y 2921 cm^{-1} que corresponden al estiramiento asimétrico de R_2CH_2 .

9.3 Índice de Clase de solubilidad Asfaltenos (ASCI)

Figura 22.

Descripción Metodología (Parte 2) Índice de Clase de solubilidad de Asfaltenos (ASCI)



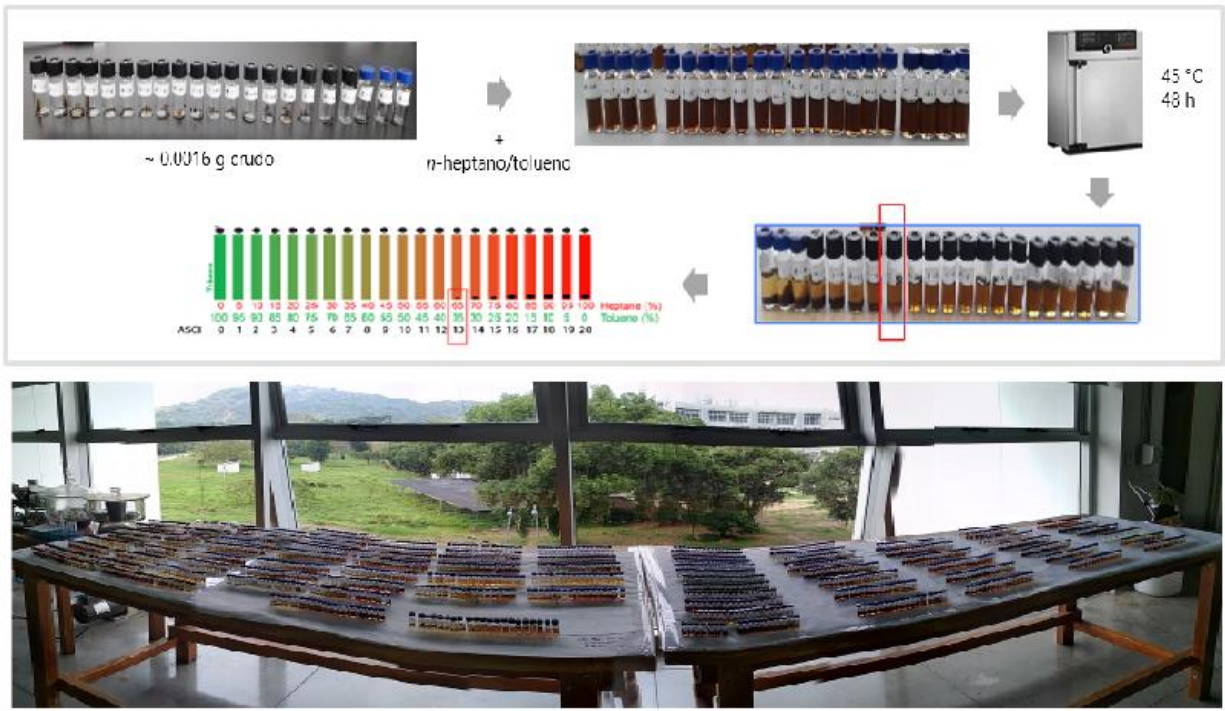
La estabilidad de los 82 crudos Colombianos se evaluó siguiendo el procedimiento propuesto por Passade – Boupat ¹³. Esta metodología denominada “índice de clase de solubilidad de asfaltenos” (ASCI), se basa en el inicio de precipitación de los asfaltenos en la muestra cruda con diferentes proporciones de soluciones *n*-heptano/tolueno. Para la preparación de cada uno de

los índices ASCI (total 21 soluciones de Heptol, para cada muestra de crudo) se tomaron 0.0016 g de la muestra crudo. En la imagen siguiente podemos observar el proceso general para determinar el índice ASCI.

La metodología ASCI puede tomar valores de 0 a 20; cuanto mayor sea el valor ASCI mayor será la estabilidad, los datos ASCI obtenidos (propiedad-Y) se muestran en su totalidad en el (Apéndice A). La Tabla 4, muestra un resumen estadístico de las mediciones ASCI y el rango de densidad API°, al igual que la fracción de saturados, aromáticos, resinas y asfaltenos (SARA) en porcentaje en peso (datos completos de 60 muestras en Apéndice A). Los datos de la fracción SARA presentada en este trabajo, sigue el método y los estándares de la norma (ASTM D2007). Esta información muestra que los crudos utilizados varían desde crudos ligeros (48° API) hasta crudos pesados (11° API), que también exhiben diferencias significativas en su distribución SARA. Como el valor ASCI es un entero, la media se aproximó al entero más cercano.

Figura 23.

Paso a paso Metodología "Índice de Clase de Solubilidad de Asfaltenos" (ASCI).



Nota: Passade-Boupat, N., Gingras, J. P., Desplobins, C. & Zhou, H. Could the Asphaltene Solubility Class Index Be Used as the ‘wax Appearance Temperature’ of Asphaltenes? Illustration through the Study of the Polydispersity of PetroPhase 2017 Asphaltenes. Energy and Fuels 32, 2760–2768 (2018)

Tabla 4.

Resumen estadístico de la fracción ASCI, API y SARA para las muestras de crudo colombiano

Propiedad	Cantidad de Muestras	Min	Max	Media	Mediana
ASCI	82	5	20	14	14
° API	65	11	48	25	24
Saturados (% w)	60	13	71	34	32
Aromáticos (% w)	60	8	51	31	30
Resinas (% w)	60	2	27	12	11
Asfaltenos (% w)	60	0	14	4	3

9.4 Desarrollo del modelo PLS-R

El análisis cuantitativo multivariado para predecir el valor ASCI se desarrolló utilizando la regresión de mínimos cuadrados parciales (PLSR), que se implementó empleando la modificación inspirada estadísticamente del algoritmo de mínimo cuadrado parcial (SIMPLS). La evaluación y determinación del número de componentes, se llevó a cabo por el método de validación cruzada de Monte Carlo (MCCV).

9.4.1 Preprocesamiento espectral

En términos generales, las muestras estudiadas presentaron un perfil infrarrojo con pequeñas diferencias en sus intensidades relativas. Al aplicar un enfoque manual para la selección de variables, el rango espectral inicial se redujo al intervalo (1900 – 400 cm^{-1}). Sistemáticamente, los rangos espectrales se descartaron si su exclusión no disminuía el error estándar de validación cruzada. En el intervalo seleccionado, hay bandas de grupos funcionales representativos de los asfaltenos, relacionados, por definición, con la estabilidad del petróleo crudo.

Esta región de los espectros, normalizada por el área bajo la curva espectral y centrada con respecto al promedio del conjunto de espectros de calibración, constituye la señal predictiva (ver región sombreada *Figura 20*). No se aplicaron corrección de señal de fondo, ni algoritmos de filtrado digital de los espectros FT-IR/ATR.

9.4.2 PLS – R

El análisis cuantitativo multivariado para predecir el valor ASCI se desarrolló utilizando la regresión de mínimos cuadrados parciales (PLS-R), que se implementó utilizando la modificación inspirada estadísticamente del algoritmo de mínimo cuadrado parcial (SIMPLS). El núcleo de la calibración para la regresión es definir la dimensión apropiada para el modelo y calcula un vector de coeficientes b (con el vector de transposición b^T) que permita estimar la variable de respuesta \hat{y} a partir del espectro (o señal predictiva) de la muestra X , de acuerdo con la ecuación (25). En este caso, \hat{y} corresponde al parámetro ASCI.

$$\hat{y} = b^T X \quad \text{Ecuación (25)}$$

La calibración multivariada se realizó con la validación cruzada de Monte Carlo (MCCV), para determinar la complejidad del modelo. Para un total de n muestras, aproximadamente n^2 particiones aleatorias, en dos conjuntos: uno para entrenamiento (60% de total de muestras) y otro para validar (40% de retención). Para una partición dada, el algoritmo SIMPLS genera un vector b para cada variable latente (LV) de la calibración modelo. En la práctica, los vectores b respectivos de cada uno de los factores PLS se promedian obteniendo el vector \bar{b} .

Un primer criterio para determinar el número de Lv (Variables Latentes) del modelo responde al comportamiento de las correlaciones r de dos vectores \bar{b} sucesivos (por ejemplo, los vectores \bar{b}_k y \bar{b}_{k+1} con magnitudes de $|\bar{b}_k|$ y $|\bar{b}_{k+1}|$, respectivamente) calculado por:

$$r = \frac{\bar{b}_k \cdot \bar{b}_{k+1}}{|\bar{b}_k| |\bar{b}_{k+1}|} \quad \text{Ecuación (26)}$$

En ese sentido, el número de Variables Latentes (Lv) corresponderá al punto donde r alcanza un valor aproximadamente constante. Es apropiado relacionar esta tendencia con algunas métricas de rendimiento del modelo de calibración a saber, la determinación entre los valores pronosticados y observados para el conjunto de entrenamiento R_C^2 y el conjunto de prueba R_V^2 ; el error estándar de calibración SEC y validación cruzada $SECV$, la desviación predictiva residual RPD o, de manera similar, la relación error de rango RER . Para la referencia y_i con los valores más altos y_{max} con los valores más bajos y_{min} , se calculó la desviación estándar $SD_y, SEC, SECV, RPD, RER$ mediante las siguientes ecuaciones (27), (28), (29) y (30).

$$SEC = \sqrt{\frac{1}{d} \sum_{i=1}^{n_t} (\hat{y}_i - y_i)^2} \quad \text{Ecuación (27)}$$

$$SECV = \sqrt{\frac{1}{n_{cv}} \sum_{i=1}^{n_{cv}} (\bar{y}_i - y_i)^2} \quad \text{Ecuación (28)}$$

$$RPD = \frac{SD_y}{SECV} \quad \text{Ecuación (29)}$$

$$RER = \frac{|y_{max} - y_{min}|}{SECV} \quad \text{Ecuación (30)}$$

En las ecuaciones (27) y (28), y_i junto con \hat{y}_i son los valores de referencia y estimados de la i -ésima muestra de calibración (validación), respectivamente. n_t y n_{cv} son la cantidad de muestras para entrenamiento y validación cruzada respectivamente. En la ecuación (27), $d = (n_t - k - 1)$ corresponde a los grados de libertad del modelo, donde k corresponde al número de Variables Latentes (Lv).

Para una fracción de n^2 y usando \bar{b} en la ecuación (25), las predicciones exhiben errores estándar distribuidos en segmentos de arco para cada uno de las variables latentes (Lvs), que se representan en el plano $SECV - SEC$. En general, a medida que aumenta el número de factores

PLS, la longitud de los arcos y la brecha entre ellos exhiben un comportamiento asintótico que podría usarse para definir la complejidad del modelo de regresión. Por otro lado, el límite de confianza a un nivel de 95% para el rendimiento del modelo se calculó en $\hat{y}_i \pm t \cdot SEC \cdot \sqrt{1 + h_{ii}}$, donde t es el valor t de estudent para d grados de libertad.

Para cada muestra de entrenamiento, con un valor estimado \hat{y}_i , h_{ii} es un escalar tomado en orden, desde la diagonal principal de la matriz $T \times T'$, donde las puntuaciones T son las coordenadas de la muestra en el espacio de componentes principales. La bondad de ajuste se determinó mediante una prueba F . Los códigos para el Preprocesamiento y el desarrollo del modelo se implementaron en Matlab R2018a (MathWorks Inc.).

10. Análisis y Resultados

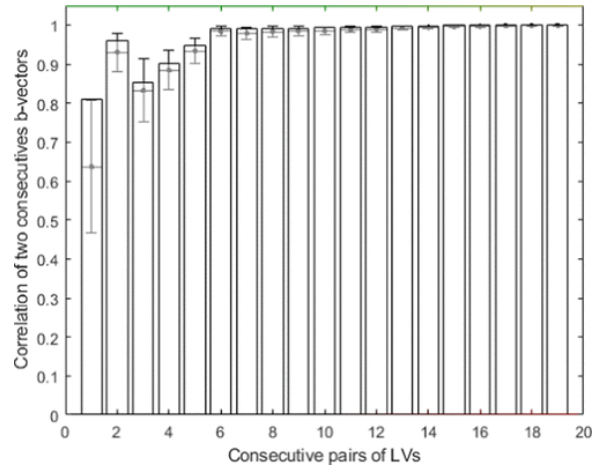
El modelo de calibración se desarrolló a partir de ($n = 82$) muestras de petróleo crudo divididas aleatoriamente en dos conjuntos: uno con 50 muestras para entrenamiento n_t y el otro con 32 muestras restantes n_{CV} para validación. Un total de 1000 particiones (o pruebas) generadas la misma cantidad de modelos para cada uno de los primeros 20 variables latentes (Lv).

En todos los casos, el conjunto de entrenamiento contenía las muestras con valores ASCII más altos y más bajos (ver Tabla 4). Por lo tanto, el modelo de regresión solo es válido para interpolar valores en ese rango. En la *Figura 24*, después de algunas oscilaciones, la correlación de dos de los vectores \bar{b} consecutivos prácticamente permanecen constantes, lo que sugiere que se

incorpora información nueva insignificante a partir del octavo (8) variable latente (Lv) y define, de esta manera, la dimensión del modelo de regresión.

Figura 24.

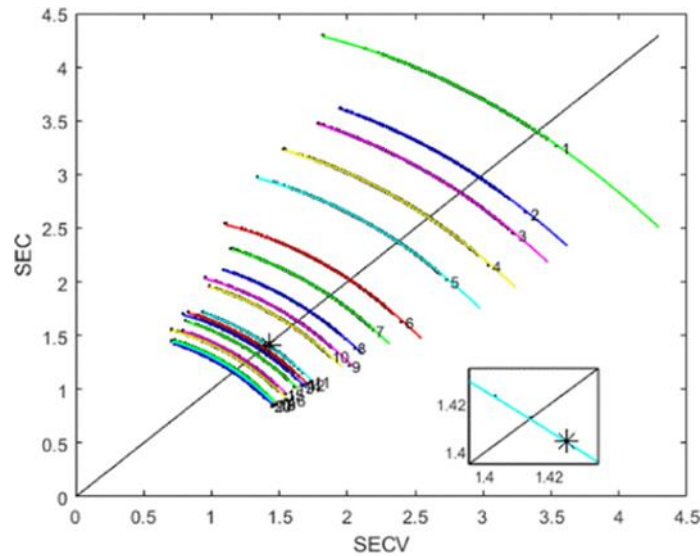
Correlación de dos vectores b consecutivos contra el número de factores PLS



Sin embargo, siguiendo la tendencia de la longitud de los arcos y la brecha entre ellos en el plano *SECV – SEC* representado en la *Figura 25*, una complejidad de once ($k = 11$) Lvs podrían ser aceptados. Este número de factores también está justificado para rastrear los promedios de las métricas de rendimiento.

Figura 25.

Errores estándar de 1000 particiones para cada uno de los primeros 20 LV utilizando el vector \bar{b} , en el plano $SECV$ - SEC .

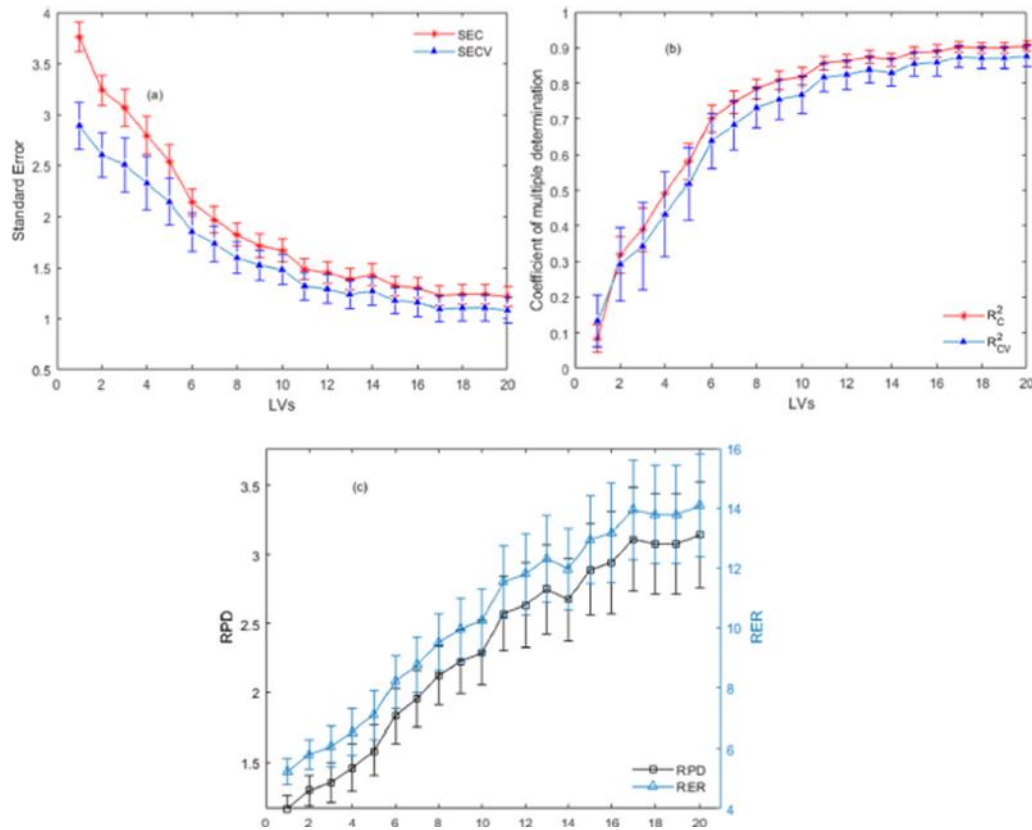


Nota: En el cuadro, un zoom a la partición (marcado con un asterisco) que muestra un optimismo moderado y cumple la condición de $SECV > SEC$

En la *Figura 26a*, es evidente que hay un pequeño salto descendente en los errores estándar ($SEC, SECV$) cerca del undécimo Lvs. Se puede ver en una situación análoga, pero con un salto hacia arriba, en los perfiles R_C^2, R_{CV}^2 (*Figura 26b*) y en el comportamiento del RPD, RER que se muestran en la *Figura 26c*, (con escala en los ejes izquierdo y derecho, respectivamente). Naturalmente, la información del RPD, RER , es similar, correspondiente a sus definiciones. Las barras verticales en la *Figura 26*, indican la desviación estándar del promedio de las 1000 particiones.

Figura 26.

Métricas de rendimiento del modelo para predecir el valor del índice ASCI vs el número de variables latentes (Lv)



Nota: (a) SECV y SEC. (b) Coeficientes de determinación para el conjunto de datos de entrenamiento (R_c^2) y validación (R_{CV}^2); (c) RPD y RER. Las barras verticales indican la desviación estándar del promedio de las 1000 particiones para cada Lv.

La calidad del modelo seleccionado se describió de acuerdo con los siguientes índices: R_C^2 , R_{CV}^2 , RER, RPD y SECV (Tabla 5), los valores informados para R_C^2 (0.82 ± 0.02) y R_{CV}^2 (0.82 ± 0.04) son típicos de los modelos de calibración multivariada y son consistentes con el comportamiento de los datos ASCI predichos versus los valores medidos (Figura 26). Por otro

lado, RPD (2.57 ± 0.3) está cerca del intervalo $[2, 2.5]$ y RER (11.51 ± 1.2) sugiere un modelo aceptable para predicciones cuantitativas del valor ASCI. Los valores de estos descriptores corresponden a promedios, para más de mil particiones (Ver *Tabla 5*).

Tabla 5.

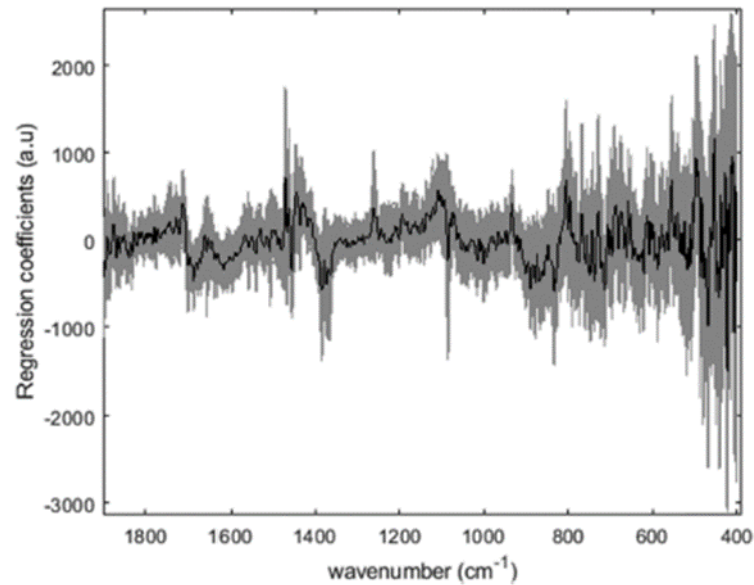
Descripción del rendimiento del modelo

Parámetro	Valor
<i>LVs</i>	11
R_C^2	0.85 ± 0.02
R_{CV}^2	0.82 ± 0.04
<i>RPD</i>	2.57 ± 0.3
<i>RER</i>	11.51 ± 1.2
<i>SECV</i>	1.32 ± 0.1
$\beta_1, p - valor$	0.06
$\beta_2, p - valor$	0.2
F_{cal}	17.3
F_{tab}	2.1

Los 1000 vectores b correspondientes a la undécima variable latente (Lvs), se muestran en la *Figura 27*, (gris claro), y su respectivo \bar{b} , aparece en negrita; para este vector \bar{b} , la partición con un asterisco (ver cuadro en la *Figura 25*) exhibe el siguiente rendimiento según métricas: $R_C^2 = 0.85, R_{CV}^2 = 0.83, SECV = 1.42, SEC = 1.40$. Naturalmente, estos valores están dentro de las estadísticas para todas las muestras (Ver *Tabla 5*) y mantienen una compensación adecuada en términos de optimismo (una medida de la degradación del rendimiento, para el mismo modelo, considerando diferentes conjuntos de capacitación y validación). Esta partición es la más cercana a la línea de identidad (optimismo moderado) en el plano $SECV - SEC$, que cumple con la condición que $SECV$ es ligeramente más grande que SEC y permite estimar la confianza del modelo de calibración.

Figura 27.

Vector de coeficientes de regresión para 1000 particiones, correspondientes a la undécima variable latente (sombreado en gris claro).

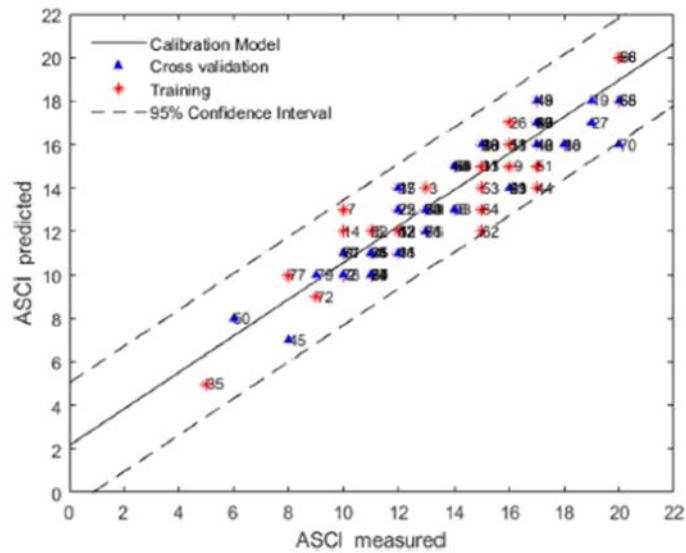


Nota: El vector b promedio aparece en negrita. Gráfica

En la *Figura 28*, el límite de la banda de confianza del 95% se muestra con una línea discontinua. Es ampliamente aceptado, y solo quedan un par de muestras de rango (uno de ellos es el conjunto de validación).

Figura 28.

Correlación entre lo observado y lo predicho (valores ASCI).

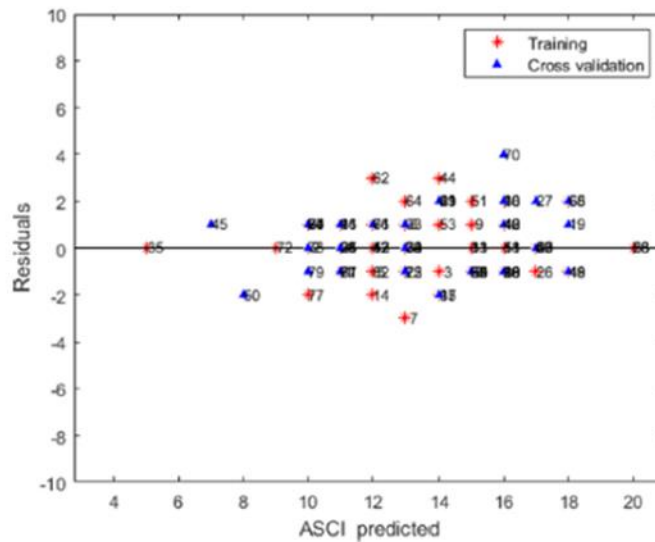


Nota: La predicción se realizó tomando el vector b y una partición (seleccionada entre 1000), correspondiente a la marcada con un asterisco en la Grafica 3

La inspección visual de la Figura 28, muestra un patrón aleatorio en los residuos en función de los valores de referencia. En particular, este hecho se corroboró para los residuos del conjunto de datos de entrenamiento mediante la aplicación de una prueba de 50000 permutaciones no paramétricas. Los valores p obtenidos para los coeficientes lineales (β_1) y cuadráticos (β_2) fueron, 0.06 y 0.2, superiores a 0.05, lo que evidencia la ausencia de tendencias de este orden (Ver Tabla 5).

Figura 29.

Residuales para el valor ASCI de predicción usando FT-IR/ATR en la región media del infrarrojo.



En la Figura 29, es posible observar que la zona de la mayor variación del vector \bar{b} está entre 900 y 400 cm^{-1} , lo que corresponde a modos vibracionales de funcionalidades aromáticas con un alto grado de condensación en anillos aromáticos tales como anillos di-, tri-, tetra-, y hexa-aromáticos. En 744, 812 y 866 cm^{-1} , se observan señales correspondientes a bandas mono y diaromáticas. Estos modos vibratorios están asociados a los asfaltenos, que son responsables de la inestabilidad de los aceites crudos.

El valor crítico de F_{Tab} a un nivel de significancia del 95% es $F_{\text{Tab}} = (0.95; 10,39) = 2.08$. El valor calculado de $F_{\text{cal}} = 17.28$ es más alto que el valor crítico, y se puede concluir que la bondad de ajuste es estadísticamente significativa.

11. Comparación Índice de Clase de Solubilidad de Asfaltenos (ASCI) con otros métodos de Estabilidad

La comparación del índice de clase de solubilidad de asfaltenos (ASCI), se llevó a cabo calculando los valores de algunas pruebas de estabilidad de crudos o aceites de petróleo, las cuales permiten predecir y evaluar la precipitación de los asfaltenos, basándose en la composición del crudo; los métodos utilizados en esta comparación son: (1) índice de clase de insolubilidad (CII) ecuación (5), (2) índice de estabilidad (IE) se halla con la fracción (A/R), (3) número de insolubilidad (I_N); Este último método ya hace más referencia hacia la “calidad” de los asfaltenos, (Ver Tabla 6).

El número de insolubilidad I_N , se determinó utilizando la conversión matemática de ASCI, hacia I_N en determinaciones equivalentes, presentada a continuación en la ecuación (29):

$$ASCI = (100 - I_N) / 5 \quad \text{Ecuación (29)}$$

En la Tabla 6. Se observa un rendimiento de calidad en la predicción de la estabilidad por el índice de clase de solubilidad de asfaltenos (ASCI), mostrando un comportamiento similar en la predicción promedio de la estabilidad. También se denota que las muestras de crudo con un valor de $ASCI \leq 13$, son crudos generalmente inestables, a su vez aquellos crudos con $ASCI \geq 15$, presentan un buen comportamiento como crudos estables en la “predicción promedio”. En el valor de la propiedad $ASCI = 14$, se pudo observar un amplio grado de incertidumbre o concepto de “incierto” respecto a la estabilidad del crudo.

Tabla 6.

Comparación de los datos obtenidos con diferentes métodos que evalúan la estabilidad y precipitación de asfaltenos en dos aspectos: composición y calidad.

N° Muestra	ASCI	Estabilidad	CII	Estabilidad	IE	Estabilidad	IN	Estabilidad	Predicción Promedio
50	6	Inestable	1,13	Inestable	1,34	Inestable	75	Inestable	Inestable
45	8	Inestable	1,31	Inestable	1,13	Inestable	65	Inestable	Inestable
1	10	Inestable	0,89	Incierto	0,43	Inestable	55	Inestable	Inestable
4	11	Inestable	0,93	Inestable	0,61	Inestable	50	Inestable	Inestable
52	12	Inestable	1,38	Inestable	1,76	Inestable	45	Inestable	Inestable
61	13	Inestable	1,82	Inestable	0,39	Inestable	40	Inestable	Inestable
6	14	Incierto	0,82	Incierto	0,24	Estable	35	Estable	Incierto
54	14	Incierto	1,14	Inestable	0,6	Inestable	35	Estable	Inestable
13	15	Estable	0,89	Incierto	0,35	Estable	30	Estable	Estable
26	16	Estable	0,65	Estable	0,21	Estable	25	Estable	Estable
23	17	Estable	0,97	Inestable	0,09	Estable	20	Estable	Estable
10	18	Estable	0,76	Incierto	0,21	Estable	15	Estable	Estable
19	19	Estable	0,7	Estable	0,13	Estable	10	Estable	Estable
28	20	Estable	0,64	Estable	0,05	Estable	5	Estable	Estable

Verificando más a fondo la relación ASCII - I_N , se elaboró una comparación mostrada en la Tabla 7. Que muestra una similitud bastante robusta entre los dos métodos ya mencionados. Estos valores se determinaron, utilizando datos de número de insolubilidad (I_N) para 11 muestras de crudo problema, obtenidos por el Laboratorio de Resonancia Magnética Nuclear, vinculado al Laboratorio de Espectroscopía Atómica y Molecular de la Universidad Industrial de Santander. Se utilizó la ecuación (29) para realizar la conversión de datos I_N a datos ASCII.

Tabla 7.

Comparación valores propiedad ASCII e IN, utilizando la ecuación 29, para la Conversión Matemática de las Propiedades en Mención

Muestras	ASCI	$(100- I_N) / 5 = ASCII$	I_N
P6	10	10.7	46.33
P7	9	10.63	46.85
P25	13	12.32	38.36
P50	11	11.7	41.46

Muestras	ASCI	(100-IN) /5 = ASCI	I_N
P63	12	11.47	42.65
P77	13	10.49	37.51
P87	8	8.86	55.66
P105	10	8.68	52.77
P118	9	10.10	49.49
P137	11	10.36	48.17
P161	11	10.10	49.47

Este trabajo de investigación fue apoyado por el convenio de investigación N.º. 5222395, AC – 22 entre Ecopetrol S.A y la Universidad Industrial de Santander. Los resultados del proyecto fueron publicados en la revista Energy & Fuels, DOI: 10.1021/acs.energyfuels.9b02829. La cual es publicada por The American Chemical Society, cuya divulgación de artículos está centrada en todas las áreas de investigación de química de las fuentes de energía no nuclear. En este artículo se expresa también un agradecimiento a COLCIENCIAS bajo el proyecto N.º. FP44842-496-2016.

12. Conclusiones

La metodología desarrollada en este trabajo de investigación permitió obtener un modelo PLS, para predecir el índice de clase de solubilidad de asfaltenos, un método desarrollado recientemente para medir la estabilidad del petróleo crudo. El modelo se elaboró utilizando datos infrarrojos de las 82 muestras de crudos colombianos.

Los resultados obtenidos muestran una alta correlación entre la predicción y la medición con SECV y SEC, alrededor de la unidad, similar al error empírico. También se desarrolló una metodología precisa para encontrar el número de componentes del modelo, utilizando un método sencillo de validación cruzada denominada MCCV, cuya importancia y valor radican en evitar el ajuste o sobreajuste del modelo.

El índice de clase de solubilidad de asfaltenos (ASCI), resultó ser un parámetro robusto y sencillo para calificación de la solubilidad de los asfaltenos.

Finalmente, el procedimiento completo para determinar la estabilidad de los crudos a través del método ASCI, podría permitir superar los problemas de tiempo de respuesta en la determinación de estabilidad de crudos

Referencias Bibliográficas

- Alimohammadi, S., Zendejboudi, S. & James, L. A comprehensive review of asphaltene deposition in petroleum reservoirs: Theory, challenges, and tips. *Fuel* 252, 753–791 (2019).
- Alvarez-ram, F. & Ruiz-morales, Y. Island versus Archipelago Architecture for Asphaltenes : Polycyclic Aromatic Hydrocarbon Dimer Theoretical Studies. (2013) doi:10.1021/ef301522m.
- Ashoori, S., Sharifi, M., Masoumi, M. & Mohammad Salehi, M. The relationship between SARA fractions and crude oil stability. *Egypt. J. Pet.* 26, 209–213 (2017).
- Asomaning, S. & Watkinson, A. P. Petroleum stability and heteroatom species effects in fouling of heat exchangers by asphaltenes. *Heat Transf. Eng.* 21, 10–16 (2000).
- ASTM D2549. for Separation of Representative Aromatics and Nonaromatics Fractions of High-Boiling Oils by Elution Chromatography'. *ASTM Int.* West Conshohocken, PA 05, 5–10 (2002).
- ASTM E1655-05. Standard Practices for Infrared Multivariate Quantitative Analysis. *ASTM Int.* 05, 29 (2012).
- ASTMD4124. Standard Test Methods for Separation of Asphalt into Four Fractions. *ASTM Int.* 04, 1–7 (2009).
- Bartholdy, J. & Andersen, S. I. Changes in asphaltene stability during hydrotreating. *Energy and Fuels* 14, 52–55 (2000).

- Boduszynski, M. M. Composition of Heavy Petroleums. 1. Molecular Weight, Hydrogen Deficiency, and Heteroatom Concentration as a Function of Atmospheric Equivalent Boiling Point up to 1400 °F (760 °C). *Energy and Fuels* 1, 2–11 (1987).
- Boduszynski, M. M. Composition of Heavy Petroleums. 2. Molecular Characterization. *Energy and Fuels* 2, 597–613 (1988).
- Brown, S. D. Book Reviews: Introduction to Multivariate Statistical Analysis in Chemometrics. *Applied Spectroscopy* vol. 64 (2010).
- Chacón-Patiño, M. L., Rowland, S. M. & Rodgers, R. P. Advances in Asphaltene Petroleomics. Part 1: Asphaltenes Are Composed of Abundant Island and Archipelago Structural Motifs. *Energy and Fuels* 31, 13509–13518 (2017).
- Deposition, A. Asphaltene Deposition - Fundamentals, Prediction, Prevention, and Remediation. (2018).
- Dickie, J. P. & Yen, T. F. Macrostructures of the Asphaltic Fractions by Various Instrumental Methods. *Anal. Chem.* 39, 1847–1852 (1967).
- Drews, A. Standard Test Method for Characteristic Groups in Rubber Extender and Processing Oils and Other Petroleum-Derived Oils by the Clay-Gel Absorption Chromatographic Method. *Man. Hydrocarb. Anal.* 6th Ed. 11, 311-311–7 (2008).
- Drews, A. Standard Test Method for Kinematic Viscosity of Transparent and Opaque Liquids (the Calculation of Dynamic Viscosity). *Man. Hydrocarb. Anal.* 6th Ed. 126-126–8 (2008) doi:10.1520/mnl10842m.
- Durand, E. et al. Effect of chemical composition on asphaltenes aggregation. *Energy and Fuels* 24, 1051–1062 (2010).
- Ecopetrol, I. Muestreo y Caracterizacion de Crudos. (2018).

- Fan, T. & Buckley, J. S. Rapid and accurate SARA analysis of medium gravity crude oils. *Energy and Fuels* 16, 1571–1575 (2002).
- Filgueiras, P. R. et al. Determination of API gravity, kinematic viscosity and water content in petroleum by ATR-FTIR spectroscopy and multivariate calibration. *Fuel* 116, 123–130 (2014).
- García, C. A. et al. OPTIMIZACION DEL TRANSPORTE POR OLEODUCTO DE CRUDO PESADO CASTILLA. 8, 17–28 (2010).
- Gray, M. R., Tykwinski, R. R., Stryker, M. & Tan, X. Supramolecular Assembly Model for Aggregation of Petroleum Asphaltenes. 3125–3134 (2011) doi:10.1021/ef200654p.
- Grziwa. The chemistry and technology of petroleum. *Fuel Processing Technology* vol. 5 (1982).
- Guzmán, R., Ancheyta, J., Trejo, F. & Rodríguez, S. Methods for determining asphaltene stability in crude oils. *Fuel* 188, 530–543 (2017).
- Hasegawa, T. Quantitative infrared spectroscopy for understanding of a condensed matter. *Quantitative Infrared Spectroscopy for Understanding of a Condensed Matter* (2017). doi:10.1007/978-4-431-56493-5.
- Hay, G., Loria, H. & Satyro, M. A. Thermodynamic modeling and process simulation through PIONA characterization. *Energy and Fuels* 27, 3578–3584 (2013).
- Hay, G., Loria, H. & Satyro, M. A. Thermodynamic modeling and process simulation through PIONA characterization. *Energy and Fuels* 27, 3578–3584 (2013).
- Hoshyargar, V. & Ashrafizadeh, S. N. Optimization of flow parameters of heavy crude oil-in-water emulsions through pipelines. *Ind. Eng. Chem. Res.* 52, 1600–1611 (2013).
- Lundanes, E. & Greibrokk, T. Separation of fuels, heavy fractions, and crude oils into compound classes: A review. *J. High Resolut. Chromatogr.* 17, 197–202 (1994).

- McKenna, A. M. et al. Heavy petroleum composition. 2. Progression of the boduszynski model to the limit of distillation by ultrahigh-resolution FT-ICR mass spectrometry. *Energy and Fuels* 24, 2939–2946 (2010).
- Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. A review of variable selection methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* 118, 62–69 (2012).
- Meléndez, L. V., Lache, A., Orrego-Ruiz, J. A., Pachón, Z. & Mejía-Ospino, E. Prediction of the SARA analysis of Colombian crude oils using ATR-FTIR spectroscopy and chemometric methods. *J. Pet. Sci. Eng.* 90–91, 56–60 (2012).
- Mohammed, R. & Daubert, E. Characterization Parameters for Petroleum Fractions. *Am. Chem. Soc.* 26, 755–759 (1987).
- Moura, L. G. M., Santos, M. F. P., Zilio, E. L., Rolemberg, M. P. & Ramos, A. C. S. Evaluation of indices and of models applied to the prediction of the stability of crude oils. *J. Pet. Sci. Eng.* 74, 77–87 (2010).
- Mullins, O. C. & Sheu, E. Y. *Asphaltenes, Heavy oils, and Petroleomics.* (2007).
- Mullins, O. C. et al. *Advances in Asphaltene Science and the Yen – Mullins Model.* (2012) doi:10.1021/ef300185p.
- Mullins, O. C. The Modified Yen Model †. 2179–2207 (2010) doi:10.1021/ef900975e.
- Mullins, O. C. The modified yen model. *Energy and Fuels* 24, 2179–2207 (2010).
- Orrego-ruiz, J. A., Cabanzo, R. & Mejía-ospino, E. PLS models for determination of SARA analysis of Colombian vacuum residues and molecular distillation fractions using MIR-ATR Modelos PLS para determinación del Análisis SARA en residuos de vacío y fracciones de destilación molecular usando MIR-ATR PLS mo. *Rev. Ion* 27, 43–48 (2014).

- Passade-Boupat, N., Gingras, J. P., Desplobins, C. & Zhou, H. Could the Asphaltene Solubility Class Index Be Used as the 'wax Appearance Temperature' of Asphaltenes? Illustration through the Study of the Polydispersity of PetroPhase 2017 Asphaltenes. *Energy and Fuels* 32, 2760–2768 (2018).
- Pereira, V. J., Setaro, L. L. O., Costa, G. M. N. & Vieira De Melo, S. A. B. Evaluation and Improvement of Screening Methods Applied to Asphaltene Precipitation. *Energy and Fuels* 31, 3380–3391 (2017).
- Santos, R. . & Lo. *Brazilian Journal of Chemical Engineering - An overview of heavy oil properties and its recovery and transportation methods. Brazilian J. Chem. Eng.* 31, 571–590 (2014).
- Sepulveda Gaona, J. A., Bonilla Manrique, J. P. & Medina Majé, Y. Predicción de la Estabilidad de los Asfaltenos Mediante la Utilización del Análisis SARA para Petróleos Puros . *Stability Prediction for Asphaltenes Using SARA Analysis for Pure Petroleum . Ing. y Región* 7, (2010).
- Shadman, M. M., Saeedi Dehaghani, A. H. & Badizad, M. H. How much do you know about the methods for determining onset of asphaltene precipitation? *Petroleum* 3, 287–291 (2017).
- Smith, B. C. *Fundamentals of fourier transform infrared spectroscopy, second edition. Fundamentals of Fourier Transform Infrared Spectroscopy, Second Edition* (2011).
- SPEIGHT, J. G. *Handbook of Petroleum Product Analysis. second edition vol. 182* (2013).
- Speight, J. G. *The Chemistry and Technology of Petroleum. Chem. Technol. Pet.* (2014) doi:10.1201/b16559.
- Speight, J. G. *The Chemistry and Technology of Petroleum. The Chemistry and Technology of Petroleum* (2014). doi:10.1201/b16559.

Syunyaev, R. Z. & Likhatsky, V. V. Effects of temperature and pressure on the phase state of oils and asphaltene solutions observed using dielectric spectroscopy. *Energy and Fuels* 24, 2233–2239 (2010).

Thompson, J. M. *INFRARED SPECTROSCOPY*. (2018).

Tojima, M., Suhara, S., Imamura, M. & Furuta, A. Effect of heavy asphaltene on stability of residual oil. *Catal. Today* 43, 347–351 (1998).

UPME. *La Cadena del Petróleo*. (2013).

Vafaie-Sefti, M. & Mousavi-Dehghani, S. A. Application of association theory to the prediction of asphaltene deposition: Deposition due to natural depletion and miscible gas injection processes in petroleum reservoirs. *Fluid Phase Equilib.* 247, 182–189 (2006).

Wiehe, I. A. & Kennedy, R. J. Application of the oil compatibility model to refinery streams. *Energy and Fuels* 14, 60–63 (2000).

Wiehe, I. A. & Kennedy, R. J. Oil compatibility model and crude oil incompatibility. *Energy and Fuels* 14, 56–59 (2000).

Xu, C. & Shi, Q. *Structure and Modeling of Complex Petroleum Mixtures*.

Xu, C. & Shi, Q. *Structure and Modeling of Complex Petroleum Mixtures*. vol. 168 (2016).

Xu, Q. S. & Liang, Y. Z. Monte Carlo cross validation. *Chemom. Intell. Lab. Syst.* 56, 1–11 (2001).

Xu, Q. S., Liang, Y. Z. & Du, Y. P. Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *J. Chemom.* 18, 112–120 (2004).

Ziegel, E. R. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*. *Technometrics* vol. 43 (2001).

Apéndices

Apéndice A. Información suplementaria propiedades ASCI y SARA, para cada una de las 82 muestras de crudos colombianos.

Muestra	ASCI	API ^o	Saturados (% w)	Aromáticos (% w)	Resinas (% w)	Asfaltenos (% w)
1	10	25	31	29	11	5
2	10	30	36	32	12	3
3	13	24	34	40	18	3
4	11	13	30	26	17	10
5	11	27	31	26	10	5
6	14	25	34	34	10	3
7	10	23	33	30	10	6
8	14	23	18	25	16	3
9	16	24	22	8	2	0
10	18	24	31	33	11	2
11	16	24	30	29	12	2
12	12	24	31	26	12	4
13	15	24	32	29	11	4
14	10	47	27	9	2	0
15	12	23	34	28	12	4
16	12	23	34	31	15	3
17	12	48	24	8	2	1
18	15	23	32	30	13	3
19	19	25	31	30	16	2
20	15	25	31	29	12	2
21	11	24	31	29	12	4
22	12	24	32	28	13	4
23	17	33	33	28	6	1
24	11	13	22	37	15	14
25	11					
26	16	25.3	28	31	14	3
27	19	34.1	35	16	8	1
28	20		31	36	13	1
29	16		34	29	12	3
30	13	13	30	23	9	5
31	15	23	29	30	14	4
32	13	26	30	31	14	4
33	14	29	30	32	10	3
34	13	25	22	30	10	4
35	5	42	35	9	3	1
36	11					
37	12	31	33	26	9	2
38	18	40	33	28	8	2

Muestra	ASCI	API°	Saturados (% w)	Aromaicos (% w)	Resinas (% w)	Asfaltenos (% w)
39	15	25	24	37	14	3
40	17	16	34	33	27	1
41	16		13	41	21	15
42	17		23	41	14	2
43	12					
44	17					
45	8	14	47	35	8	9
46	17	12	29	45	24	3
47	17	12	29	51	19	2
48	17	14	35	43	19	3
49	17	11	24	49	26	3
50	6	15	41	37	8	10
51	17	21	44	38	15	3
52	12	20	47	35	6	10
53	15	13	25	51	18	7
54	14	19	48	38	9	5
55	20					
56	20	21	47	35	10	8
57	11	29	54	39	5	2
58	16	33	43	19	4	1
59	17	19	38	41	19	2
60	17					
61	13	27	62	30	6	2
62	15	39	71	25		1
63	16	39	67	28	3	2
64	15	30	43	42	7	9
65	14	25				
66	15	20	42	46	9	1
67	10					
68	20					
69	14					
70	20					
71	10					
72	9					
73	13					
74	11					
75	12					
76	13					
77	8					
78	10					
79	9					
80	11					
81	12					
82	11					