

**Análisis estadístico y por simulación del desempeño operacional de la planta DEMEX de la
GRB**

Gabriel Antonio Morales Salazar

Trabajo de aplicación para optar al título de Magister en Ingeniería Química

Director:

Giovanni Morales Medina

Dr. En Ingeniería Química

Universidad Industrial de Santander

Facultad de Ingenierías fisicoquímicas

Escuela de Ingeniería Química

Maestría en Ingeniería Química

Bucaramanga

2017

Agradecimientos

En primera instancia agradezco a Dios por el regalo de la vida y por las opciones de crecimiento en los ámbitos profesional, personal y familiar.

Asimismo, a mi familia que brindó las bases para llegar al lugar en donde estoy. Me encuentro en deuda con mi familia, que sin lugar a duda son el motor que impulsa mi desarrollo y fija mi horizonte.

También, quiero expresar mis agradecimientos a ECOPETROL por brindar esta oportunidad de desarrollo académico; me siento orgulloso de laborar en la mejor empresa del país.

Al igual, expreso mi admiración y agradecimiento al cuerpo profesoral de la escuela de Ingeniería Química de la UIS. Mi especial respeto y agradecimiento a mi director de proyecto por su orientación metodológica y dedicación para con este trabajo de aplicación final.

Finalmente, mis agradecimientos a todas las personas que creyeron en mis capacidades y colaboraron en forma directa o indirecta en la construcción de este documento.

Contenido

	Pág.
Introducción	17
1. Objetivos	19
1.1 Objetivo General	19
1.2 Objetivos Específicos.....	20
2. Marco Teórico.....	20
2.1 Unidad de desasfaltado DEMEX	20
2.1.1 Sección de Extracción.....	21
2.1.2 Sección de Recuperación de DMO	22
2.1.3 Sección de Recuperación de Asfalto.....	22
2.1.4 Sección de Recuperación de Solvente	23
2.2 Estado del arte procesos de desmetalizado y desasfaltado	23
2.3 Análisis por componentes principales (PCA) y mínimos cuadrados parciales (PLS)	25
2.4 Alcances	27
3. Metodología	27
3.1 Recolección de una base de datos con los históricos de proceso.....	27
3.2 Identificación de escenarios operativos por análisis PCA y el método de agrupación <i>k-means</i>	28
3.3 Aplicación de regresiones múltiples lineales	28

3.4 Simulación de la unidad DEMEX.....	29
4. Análisis de resultados	30
4.1 Datos históricos de la unidad DEMEX.....	30
4.2 Escenarios de operación para la unidad DEMEX.....	32
4.3 Regresiones para el rendimiento de DMO en los escenarios operativos de la unidad DEMEX	47
4.3.1 Clase 0.....	48
4.3.2 Clase 1.....	67
4.3.3 Clase 2.....	80
4.3.4 Clase 3.....	96
4.4 Simulación del proceso DEMEX.....	110
5. Contribuciones del proyecto de aplicación.....	119
6. Conclusiones.....	120
7. Recomendaciones	122
Referencias Bibliográficas	123
Apéndice	126

Lista de Figuras

	Pág.
Figura 1. Diagrama de entrada-salidas unidad DEMEX.	21
Figura 2. Tendencia para algunas variables contenidas en la base de datos históricos para el proceso DEMEX; ventana operativa entre enero de 2014 y diciembre de 2016.	32
Figura 3. Análisis PCA a la base de datos históricos de DEMEX con pretratamiento de autoescalado. (a) varianza explicada acumulada del conjunto. (b) Gráfico de influencia para PC-7. (c) Residuales por muestra para PC-7; círculos negros corresponden a las muestras con estadísticos T^2 y F fuera de los respectivos límites (Figura b).	34
Figura 4. Gráfica de los scores para los dos primeros componentes principales (47%) derivados del PCA. Las agrupaciones fueron obtenidas por el método k-mean utilizando la distancia euclidiana.	36
Figura 5. Gráfica de los scores para los tres primeros componentes principales (57%) derivados del PCA, con cuatro agrupaciones ($k=4$).	38
Figura 6. Gráfica de los scores para los tres primeros componentes principales (57%) derivados del PCA.	38
Figura 7. Tendencias para la varianza interna total no ponderada en las clases (SS Interna clases) y la varianza no ponderada entre las clases (SS Entre clases) respecto al número de clases según resultados obtenidos con el programa de uso libre R. El valor de la segunda derivada denota que la cohesión y la separación adecuadas en la agrupación son alcanzadas con cuatro clases.	39

Figura 8. Gráficas de los coeficientes silueta para las clases 2 – 7, según resultados con el programa R.....	40
Figura 9. Gráficas de clases para la ventana de operación y la carga al proceso.....	41
Figura 10. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases.	43
Figura 11. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases (continuación).....	44
Figura 12. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases (continuación).....	45
Figura 13. Ilustración clases en el espacio de las variables Butano vs Carga vs Temperatura T_D2501A.	45
Figura 14. Ilustración clases en el espacio de las variables Butano vs Carga vs Temperatura T_D2501A (continuación).....	46
Figura 15. Variables de salida de la unidad DEMEX y su relación con las clases obtenidas.	47
Figura 16. Resultados PCA para la clase 0. (a) varianza explicada, 81.7% calibración, 68.7% validación. (b) Gráfica de los scores denotando dos ($k=2$) agrupaciones (subdivisión 1 en azul, subdivisión 2 en rojo). (c) Gráfica de los loadings.	49
Figura 17. Dispersión de los datos para la clase 0 resaltando las subdivisiones obtenidas con el método k-means (subdivisión 1 en azul, subdivisión 2 en rojo). (a) Gráfica Q vs T_2520A vs T_RectA. (b) Gráfica T_2520B vs T_2520A.	50
Figura 18. Regresión MLR para todos los datos de la clase 0.....	53
Figura 19. Regresión MLR para todos los datos de la clase 0 (continuación).....	54
Figura 20. Regresión MLR para todos los datos de la clase 0.....	55

Figura 21. Tendencia del rendimiento de DMO con la temperatura crítica del solvente (a) y con la temperatura del proceso de extracción (b). Fuente: Manual de operación de la planta DEMEX.

Capítulo 5..... 56

Figura 22. Valor predicho vs valor de referencia reportada por PLS para todos los datos de la clase 0..... 56

Figura 23. Regresión MLR para la subdivisión 1 de la clase 0. 57

Figura 24. Regresión MLR para la subdivisión 1 de la clase 0 (continuación). 58

Figura 25. Regresión MLR para la subdivisión 1 de la clase 0 (continuación). 59

Figura 26. Regresión MLR para la subdivisión 1 de la clase 0 (continuación). 60

Figura 27. Regresión PLS para la subdivisión 1 de la clase 0. 60

Figura 28. Regresión MLR para la subdivisión 2 de la clase 0. 62

Figura 29. Regresión MLR para la subdivisión 2 de la clase 0 (continuación). 63

Figura 30. Regresión MLR para la subdivisión 2 de la clase 0 (continuación). 64

Figura 31. Regresión MLR para la subdivisión 2 de la clase 0 (continuación). 65

Figura 32. Regresión PLS para la subdivisión 2 de la clase 0. 65

Figura 33. Regresión PLS para la subdivisión 2 de la clase 0 (continuación). 66

Figura 34. Resultados PCA para la clase 1. 68

Figura 35. Resultados PCA para la clase 1 (continuación). (a) varianza explicada, 85.1% calibración, 74.5% validación. (b) Gráfica de los scores denotando dos ($k=2$) agrupaciones. (c) Gráfica de los datos históricos considerando dos agrupaciones. (d) Gráfica de los datos históricos considerando tres agrupaciones. (e) Gráfica de los datos históricos considerando dos agrupaciones. (f) Gráfica de influencia considerando dos agrupaciones. 69

Figura 36. Resultados regresiones MLR con los datos de la clase 1. (a) Todos las 3243 muestras. (b) Subdivisión1, 2573 muestras. (c) Subdivisión2, 730 muestras. Datos en azul y en rojo corresponden a las etapas de calibración y validación, respectivamente.....	71
Figura 37. Resultados regresión MLR con los datos de la clase 1.	72
Figura 38. Resultados regresión PLS para la clase 1.....	74
Figura 39. Resultados regresión PLS para la clase 1 (continuación).....	74
Figura 40. Resultados regresión PLS para la subdivisión 1 de la clase 1. (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión. (d) Influencia de las variables independientes.	75
Figura 41. Resultados regresión PLS para la subdivisión 1 de la clase 1 (continuación). (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión. (d) Influencia de las variables independientes.	76
Figura 42. Resultados regresión PLS para la subdivisión 2 de la clase 1.....	76
Figura 43. Resultados regresión PLS para la subdivisión 2 de la clase 1. (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión.....	77
Figura 44. Resultados análisis PCA para los datos que conforman la clase 2.....	81
Figura 45. Subdivisión de las muestras de la clase 2 según el método k-mean (k=2).....	84
Figura 46. Resultados regresión MLR para las muestras de la clase 2.....	85
Figura 47. Resultados regresión MLR para las muestras de la clase 2 (continuación).	86
Figura 48. Resultados regresión MLR para las muestras de la clase 2.....	87
Figura 49. Resultados regresión PLS para las muestras de la clase 2.....	87
Figura 50. Resultados regresión PLS para las muestras de la clase 2 (continuación).	88

Figura 51. Resultados regresión MLR para las muestras de la subdivisión 1 para la clase 2.....	88
Figura 52. Resultados regresión MLR para las muestras de la subdivisión 1 para la clase 2.....	90
Figura 53. Resultados regresión PLS para las muestras de la subdivisión 1 para la clase 2.	91
Figura 54. Resultados regresión PLS para las muestras de la subdivisión 1 para la clase 2 (continuación).	92
Figura 55. Resultados regresión MLR para las muestras de la subdivisión 2 de la clase 2.....	92
Figura 56. Resultados regresión MLR para las muestras de la subdivisión 2 de la clase 2 (continuación).	93
Figura 57. Resultados regresión PLS para las muestras de la subdivisión 2 de la clase 2.	94
Figura 58. Resultados regresión PLS para las muestras de la subdivisión 2 de la clase 2 (continuación).	95
Figura 59. Regresión MLR para los datos de la clase 3.....	100
Figura 60. Regresión PLS para la clase 3.	101
Figura 61. Resultados análisis PCA con los datos de la clase 3.	102
Figura 62. Resultados método de agrupación k-mean con k=2 para la subdivisión de los datos de la clase 3.....	102
Figura 63. Variación de las temperaturas de los tambores T_D2501A-B para la subdivisión de los datos de la clase 3.....	103
Figura 64. Variación de las temperaturas de las torres rectificadoras T_RectA-B para la subdivisión de los datos de la clase 3.....	104
Figura 65. Regresión MLR para la subdivisión 1 de los datos de la clase 3.	105
Figura 66. Regresión MLR para la subdivisión 2 de la clase 3.	106
Figura 67. Regresión MLR para la subdivisión 2 de la clase 3 (continuación).	107

Figura 68. Regresión MLR para la subdivisión 2 de la clase 3 (continuación).....	108
Figura 69. Regresión PLS para la subdivisión 1 de los datos de la clase 3.	108
Figura 70. Regresión PLS para la subdivisión 2 de los datos de la clase 3.	108
Figura 71. Algunos assays codificados en la herramienta Petroleum Assays de HYSYS.	111
Figura 72. Ilustración del objeto Petroleum Feeder en la simulación realizada para el proceso DEMEX en HYSYS.	112
Figura 73. Propiedades reportadas por la simulación para la mezcla de alimentación a DEMEX.	113
Figura 74. Equipo de extracción con solvente representando a la sección respectiva de DEMEX.	114
Figura 75. Equipos de separación del solvente del DMO.....	114
Figura 76. Equipos para la separación del solvente del fondo DEMEX.	114
Figura 77. Equipos de recuperación del solvente.	115

Lista de Tablas

	Pág.
Tabla 1. Valores para las variables de operación según la clase o escenario operativo.	42
Tabla 2. Valores para el rendimiento a DMO y su contenido de carbón Conradson para los escenarios.....	43
Tabla 3. Resumen modelos ajustados para la clase 0.	66
Tabla 4. Resumen modelos ajustados para la clase1.	78
Tabla 5. Variables que definen la subdivisión en la clase 2.	84
Tabla 6. Resumen modelos ajustados para la clase2.	95
Tabla 7. Variables que definen la subdivisión en la clase 3.	103
Tabla 8. Resumen modelos ajustados para la clase3.	109
Tabla 9. Crudos componentes de la Mezcla Vasconia	110
Tabla 10. Composición volumétrica de los flujos de entrada a las unidades de destilación de la GRB.	112
Tabla 11. Composición volumétrica para los diferentes solventes probados en la simulación.	116
Tabla 12. Temperatura media para la operación de la extracción con cada mezcla de solvente.	116
Tabla 13. Resultados de las simulaciones del proceso DEMEX a la T_{media} , según el solvente.	118

Lista de Apéndices

	Pág.
Apéndice A. Procedimiento para la caracterización de un ASSAY en HYSYS	126

Resumen

Título: Análisis estadístico y por simulación del desempeño operacional de la planta DEMEX de la GRB*

Autor: Gabriel Antonio Morales Salazar**

Palabras Claves: Desasfaltado, Desemetalizado, Análisis Multivariado, DEMEX, PCA, PLS, *K-means*, Simulación.

Descripción:

En la GRB el proceso de desasfaltado denominado DEMEX presenta una oportunidad de mejora para incrementar sus rendimientos de aceite desmetalizado (DMO), el cual es una carga importante a las unidades FCC; un incremento de un 5% en el rendimiento de DMO presenta un potencial económico de *ca.* 31 MMUS\$ anual. Debido a este potencial el presente trabajo final de aplicación fue dedicado al análisis de la operación de la unidad DEMEX por medio de estadística multivariada y simulación de procesos.

El análisis del proceso DEMEX involucró la aplicación de análisis multivariado a una base de datos históricos en una ventana entre enero de 2014 y diciembre de 2016. Los métodos *k-means* y PCA fueron aplicados a los datos para obtener y analizar agrupaciones referentes a esquemas operativos de la unidad. Las regresiones MLR y PLS fueron aplicadas sobre estos esquemas para la calibración del rendimiento a DMO. Asimismo, una simulación en el programa HYSYS fue codificada para el análisis de diferentes mezclas de solventes sobre el desempeño de DMO y la evaporación del solvente durante la operación.

El análisis de los datos históricos reportó cuatro escenarios operativos representativos para DEMEX. Estos escenarios se distinguen unos de otros por sus valores de las variables temperatura, carga y butano en el solvente. La aplicación de MLR y PLS en los diferentes escenarios conllevó a la validación de regresiones con utilidad a nivel cuantitativo y semicuantitativo para el rendimiento de DMO. Por último, los resultados de las simulaciones del proceso en HYSYS sugieren que la mezcla actual de solvente utilizada en DEMEX exhibe valores adecuados para el rendimiento, el contenido de CCR y la pérdida de solvente, con lo cual, la operación actual de la unidad corresponde a un punto de beneficio para la refinería.

* Trabajo de Aplicación Final.

** Facultad de Ingenierías Físicoquímicas, Escuela de Ingeniería Química Maestría en Ingeniería Química, Director: Prof. Giovanni Morales Medina, I.Q., Dr.

Abstract

Title: Estadistical and simulation analysis of operational performance for DEMEX unit of GRB*

Author: Gabriel Antonio Morales Salazar**

Keywords: Deasphalting, Demetallization, Multivariate Analysis, DEMEX, PCA, PLS, *K-means*, Simulation.

Description

Deasphalting process (DEMEX) of GRB refinery shows an opportunity for improving its yield to demetallized oil (DMO), which is an important feed to FCC units; it is shown by GRB that an increase of 5% in DMO yield means an economic potential of *ca.* 31 MMUS\$ by year. Due to this economic potential, the present document related to a final application work was devoted to the analysis of performance of DEMEX unit by the use of multivariate statistics, as well as, process simulation.

DEMEX unit analysis took into account application of multivariate statistics to a historical process data collected from January 2014 to December 2016. Statistical methods called *k-means* and PCA were applied to historical data to clustering data into different operational states for the units. MLR and PLS regression considering the different operational states were calibrated for the corresponding DMO yields. Likewise, a simulation using HYSYS program was designed to predict effects of different solvent mixtures on DMO yields and solvent evaporation levels.

Analysis of historical process data showed four operational states for the DEMEX unit. These operational states can be differentiated by the different values of temperatures, feed and butane solvent content. Application of MLR and PLS regressions considering the operational states calibrated models with potential applications to predict DMO yields at quantitative and semiquantitative levels. Finally, simulations in HYSYS suggested that the current solvent composition utilized in DEMEX yielded the best performance results among the solvents considered in this work.

* Final Application Work.

** Faculty of Physicochemical Engineering, School of Chemical Engineering Master of Chemical Engineering, Director: Prof. Giovanni Morales Medina, I.Q., Dr

Introducción

Las reservas mundiales de crudos pesados y bitúmenes ascienden a *ca.* 6 trillones de barriles (Espeith, 2014). Las características de estos crudos (*i.e.* baja relación H/C, contenido alto de metales, estabilidad limitada, incremento en el residuo de carbón, tendencia a la formación de depósitos y corrosión en los equipos) afectan los desempeños de las operaciones industriales y el margen de refinación; en especial, el contenido de vanadio, níquel y el contenido de residuo carbonoso afectan la composición y la estructura de los catalizadores utilizados para la refinación y por lo tanto afectan los costos de capital (Magodemov *et al*, 2017). Debido a cambios en la calidad de los crudos medianos disponibles para cargar, la Gerencia Refinería de Barrancabermeja, GRB, ha experimentado variaciones en los desempeños de sus unidades de proceso. Particularmente, la planta DEMEX ha procesado fondos de vacío con diferentes características, a lo cual, su desempeño ha fluctuado debido al requerimiento de la modificación de las variables operacionales para la consecución de la mayor cantidad posible de aceite desmetalizado (DMO), con un contenido aceptable de carbón residual; sin embargo, debido a la variabilidad de los crudos y cuestiones operativas, la mezcla de solvente utilizada en DEMEX podría no ser la más adecuada.

La unidad DEMEX opera con una carga de fondos de vacío que puede ser ajustada entre 32 y 45 kbpd, la composición de estos fondos de vacío varía de acuerdo a la caracterización los crudos cargados en la refinería y esto se ve reflejado principalmente en un parámetro denominado penetración. Las principales variables operacionales que son manipuladas en la operación de la planta son la composición del solvente, la relación solvente carga y la temperatura de extracción.

El solvente está compuesto por una mezcla cuyos componentes principales son propano, isobutano y normal butano; el porcentaje de estos componentes varía y a esto se le llama composición del solvente. El rendimiento de DMO esperado en Demex es de *ca.* 55 % de la carga, con un contenido de carbón Conradson de 5% en peso (Grupo de capacitación Ecopetrol, 1978). A pesar de la consecución de la calidad en el DMO, la unidad DEMEX no ha sido operada en el máximo desempeño esperado de 55%, reportando un valor medio de 52% en los últimos 8 años (Ecopetrol S.A Base de datos PI, 2008-2016). La unidad está diseñada para procesar 45 kbpd de fondos de vacío y generar 24,75 kbpd de DMO; sin embargo, las condiciones de operación actuales permiten la consecución de 22 kbpd de DMO, con lo cual se tiene un potencial de mejora de 3 kbpd. Aunque la cantidad de mejora parece pequeña, según estimaciones de planeación de la refinería, un incremento de un 5 % en el rendimiento de DMO representa una ganancia adicional para la refinería de 31 MMUS\$ anual.

De otro lado, entre las herramientas que pueden ayudar en el análisis operacional se encuentran la simulación de procesos y el análisis estadístico de datos históricos. Una simulación de procesos desarrollada en el módulo de extracción del programa HYSYS puede suministrar información del comportamiento de diferentes mezclas de solvente al ser utilizados en una carga de fondos de vacío de la refinería de Barrancabermeja. Asimismo, los datos históricos de proceso contienen información sobre los desempeños posibles de la unidad, los escenarios operativos y las fallas ocurridas en la ventana de operación en análisis; esta información puede permitir orientar los diferentes cambios que requiere el proceso para alcanzar un desempeño sobresaliente según los escenarios de carga y características de la misma. Las dos herramientas, simulación y análisis estadístico, son complementarias ya que los resultados de simulación son validados por datos

históricos y la factibilidad de las proyecciones generadas con regresiones estadísticas pueden ser validadas y analizadas por medio de simulación de los nuevos escenarios operativos.

En el presente trabajo los datos de proceso son analizados para definir escenarios operativos que relacionen las variables de proceso, las propiedades de la carga y la composición del solvente utilizado; la ventana operativa considera desde enero de 2014 a diciembre de 2016. El análisis de datos históricos se efectúa por medio del cálculo de los componentes principales y por el procedimiento *k-means*. Asimismo, diferentes regresiones fueron propuestas en cada escenario operativo identificado utilizando los métodos MLR y PLS; la variable de respuesta en las regresiones fue el rendimiento de DMO. Complementariamente, un módulo fue desarrollado en el programa HYSYS para analizar el impacto de la mezcla de solventes sobre el rendimiento y la calidad del DMO obtenido en la unidad DEMEX. Por medio de las simulaciones, la relación entre la temperatura crítica de la mezcla de solvente y su comportamiento en el proceso de extracción fue verificada; también, las simulaciones condujeron a la identificación de las mezclas con mejor desempeño según rendimiento, contenido de carbón Conradson y evaporación del solvente.

1. Objetivos

1.1 Objetivo General

Analizar la superficie operacional para el proceso de desmetalizado de fondos de vacío a partir de la aplicación de la estadística multivariada a los datos históricos de proceso y simulaciones computacionales de diferentes solventes.

1.2 Objetivos Específicos

1. Definir diferentes escenarios operativos de la unidad DEMEX a partir de la aplicación de un análisis por componentes principales y una clasificación de los datos históricos.
2. Proponer y validar regresiones que permitan la predicción del rendimiento de DMO en los diferentes escenarios operativos identificados.
3. Desarrollar simulaciones computacionales al proceso de extracción utilizando el simulador comercial HYSYS para determinar los desempeños de diferentes mezclas de solventes.

2. Marco Teórico

2.1 Unidad de desasfaltado DEMEX

DEMEX es un proceso de separación de compuestos pesados y livianos de los fondos de vacío (desasfaltado y desmetalizado), utilizando una mezcla de propanos, butanos y pentanos como solvente, para obtener un extracto con cantidades bajas de asfáltenos y metales, denominado DMO y un fondo rico en contenido de asfalto (Figura 1) (Ahn *et al.*, 2015); los fondos de vacío se reciben desde las secciones de vacío de las unidades de fraccionamiento U2000, U200, U2100, U250 y U150. El proceso se divide en cuatro secciones: extracción, recuperación de DMO, recuperación de Asfalto y recuperación de solvente.

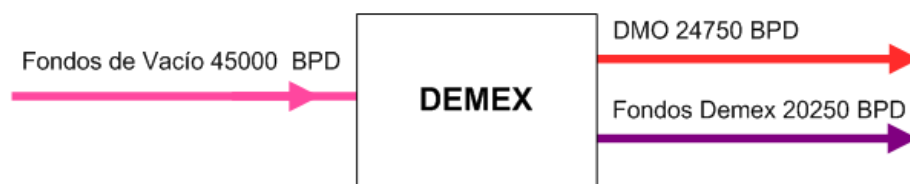


Figura 1. Diagrama de entrada-salidas unidad DEMEX.

2.1.1 Sección de Extracción. El propósito de la unidad Demex se logra en la sección de extracción. El flujo de solvente (propano-butanos-pentanos) se mezcla con el flujo de fondos de vacío para extraer selectivamente los componentes livianos de las resinas y los asfáltenos; los asfáltenos corresponden a los constituyentes pesados del petróleo ricos en metales, nitrógeno y azufre. El flujo de solvente de salida con componentes livianos –principalmente parafinas de alto peso– y una baja porción de resinas y asfáltenos corresponde a un extracto denominado DMO, del cual se pueden obtener productos de mayor valor, como lo es la gasolina (por craqueo catalítico). De otro lado, el flujo de salida –refinato o fondos DEMEX– corresponde a una opción para cargar a la unidad de viscorreducción o mezclar con otros flujos pesados y producir asfalto.

Antes de ingresar, la corriente de carga se mezcla con un flujo de solvente de predilución para alimentar los sedimentadores horizontales y las torres desasfaltadoras; la predilución mejora la extracción al disminuir la viscosidad de la carga y promover el contacto entre las fases. La selectividad y el coeficiente de distribución son los parámetros más importantes que influyen en la selección del solvente. El coeficiente de distribución (D) se define como la relación entre la concentración (X) de un compuesto particular en la fase del extracto (E) y la fase de asfalto (A), según:

$$D = \frac{X_E}{X_A}$$

Los solventes adecuados reportan coeficientes de distribución con un valor elevado. Esto significa que la concentración de extracto en la fase del extracto (X_E) es más grande que la concentración del soluto en la fase de asfalto (X_A) –eficiencia elevada de separación–. Asimismo, el proceso depende de otras variables como la relación solvente/carga (S_Q) y de las propiedades de la carga como API, penetración (PEN) y residuo de carbón (CCR).

2.1.2 Sección de Recuperación de DMO. La solución que sale en la cima de la sección de extracción, específicamente de los sedimentadores secundarios y las torres desasfaltadoras, contiene DMO y la mayoría del solvente. El propósito de esta sección es separar el solvente por medio de la elevación de la temperatura, reducción de presión y despojamiento con vapor. El extracto es calentado para elevar la temperatura y vaporizar parcialmente la mezcla. El vapor del flujo de extracto es separado en la torre T-2501, la cual posee dos rehervidores calentados por vapor. El arreglo interno de la torre y las condiciones de operación permiten una eficiencia de separación del solvente de aproximadamente 90%.

2.1.3 Sección de Recuperación de Asfalto. El asfalto que abandona el fondo de los sedimentadores primarios y las torres desasfaltadoras contiene residuos de solvente; el propósito de la sección de recuperación de asfalto es separar el solvente del asfalto. El asfalto precalentado en los Hornos H-2501A/B/C entra por la parte intermedia al tambor separador de asfalto D-2506. La mayoría del solvente de esta corriente es separada y sale por la cima en forma de vapor. El producto de fondo, contiene asfalto con solvente y pasa por un filtro atrapa-coque. El flujo ingresa a una torre despojadora (T-2522) para eliminar el solvente remanente por medio de un flujo de vapor de despojo.

2.1.4 Sección de Recuperación de Solvente. Los hidrocarburos presentes en la corriente de solvente causan daños en la etapa de compresión y disminuyen la extracción del DMO. El lavado se utiliza con el fin de retirar las trazas de DMO y asfalto remanentes en el solvente por medio de las torres de lavado de baja presión T-2504. El lavado se efectúa utilizando solvente contaminado con DMO que desciende por las torres de lavado. El DMO disuelve las trazas de hidrocarburos pesados presentes en el solvente. El solvente purificado que sale por la cima es enviado a un condensador para retirar el agua presente. El solvente separado es enviado por el compresor C-2501 para posteriormente condensarlo y recircularlo a la etapa de extracción.

2.2 Estado del arte procesos de desmetalizado y desasfaltado

El proceso de desasfaltado con solvente, denominado en la GRB como DEMEX, comparado con las tecnologías de rechazo de carbón y de adición de hidrogeno, ofrece bajos costos de instalación y operación y genera una amplia variedad de aceites desasfaltados (L.C., J.A.D, & J., 2012). Los productos del proceso son el aceite desmetalizado o desasfaltado (DMO) y el asfalto desaceitado, llamado fondos Demex. El DMO es usado como carga a las unidades de ruptura catalítica fluidizada, FCC, y su valor se expresa en términos de las propiedades residuo de carbón Conradson y contenido de metales; entre más alto sea el residuo de carbón Conradson y el contenido de metales, menor es la calidad y por lo tanto el valor del DMO. Un control deficiente en la calidad del DMO (*e.g.* aumento de residuo de carbón Conradson en valores superiores 5.5%) ocasiona una disminución en la actividad del catalizador de las unidades FCC, así como, fluctuaciones operacionales en la temperatura del regenerador disminuyendo su capacidad de procesamiento.

Los procesos de mejoramiento de residuos pesados del petróleo, pueden ser realizados por diferentes técnicas tales como: extracción, pirolisis, descomposición catalítica, hidroconversión catalítica y gasificación. En general estos procesos son realizados a altas temperaturas, incrementando su costo y limitando su aplicación (Lee, 2014). El proceso de desasfaltado utilizando extracción líquido-líquido con solventes parafínicos para remover la fracción rica en asfáltenos, ofrece ventajas tales como: bajo costo en sus instalaciones y flexibilidad en términos de control de calidad (Lee, 2014). Cao *et al.*, en 2009, reportó un análisis de sensibilidad para el proceso de desasfaltado en una refinería China. Este autor sugiere que una de las variables con mayor impacto el desempeño del proceso corresponde a la relación de solvente/carga, en concordancia con Fahai *et al.*, 2010; altas relaciones de solvente favorecen la calidad de DMO, pero incrementa el costo de operación de la unidad. En este último aspecto, Lee *et al.*, 2014, describen técnicas más eficientes de recuperación de solvente, evitando la pérdida del mismo y disminuyendo el costo energético implicado en este proceso. De otro lado, el rendimiento de DMO puede ser incrementado utilizando aditivos. Long, en 2011, reporta el uso de alquitrán en el proceso de desasfaltado, lo cual incrementa el flujo de DMO en 2%, sin afectar el contenido de carbón Conradson en este efluente. Otra mejora en la tecnología ha consistido en la adición de un proceso posterior de extracción para retirar la fracción de asfáltenos presente en el aceite desmetalizado (Ping-An *et al.*, 2012).

A nivel local, el instituto colombiano del petróleo ha desarrollado predicciones del rendimiento de DMO para la planta DEMEX aplicando la EOS de Peng-Robinson, junto con técnicas de análisis de datos (Leon *et al.*, 2009). Asimismo, Campos *et al.*, 1995, reportan solventes alternos para el proceso; Rojas *et al.*, 2002, analizaron el efecto de las corrientes intermedias en la carga de

alimentación a DEMEX. También, Nuncira et al., 2014, propusieron desarrollaron una metodología para la estimación de las propiedades del DMO resultante de la unidad DEMEX.

Según lo anterior, la literatura encontrada no reporta el uso combinado de regresiones de datos históricos y simulaciones, con el objetivo de proponer modelos de predicción de rendimientos de DMO y analizar la influencia de la composición del solvente en los rendimientos, el contenido de carbón residual y la evaporación del solvente.

2.3 Análisis por componentes principales (PCA) y mínimos cuadrados parciales (PLS)

Uno de los métodos más utilizados para la reducción de variables es la descomposición de los datos en componentes principales, mediante el análisis PCA por sus siglas en inglés (*Principal Component Analysis*); con el método PCA se concentra la variabilidad presente en la matriz de datos en unos pocos parámetros representativos, denominados componentes principales (PC); los PC corresponden a combinaciones lineales de las variables originales. Cada componente principal contiene información de diferente relevancia; los primeros componentes principales describen la fuente de variación más importante de los datos; por lo tanto ante una base de datos con muchas variables, el análisis PCA reducirá al menor número de variables, manteniendo la mayor parte de la varianza de la información posible (Berenton, 2002; Díaz & Morales, 2012).

En el PCA, la proyección sobre los nuevos ejes y los cosenos directores generan los denominados *scores* (puntuaciones) y *loadings* (cargas). Los *scores* representan a las muestras, mientras los *loadings* representan a las variables en el nuevo espacio de los PC. Con los *scores* y los *loadings* de los componentes principales se pueden identificar posibles agrupaciones, tendencias y patrones relevantes en los datos analizados (Berenton, 2002). Debido a la

descomposición propuesta en el análisis PCA, gran parte de la varianza total presentada en los datos iniciales es concentrada en los primeros componentes principales; aunque en varias aplicaciones no existe una clara relación con las variables físicas, los componentes principales presentan de forma concisa las principales tendencias del problema en análisis (Díaz & Morales, 2012). Los componentes principales que explican una buena parte de la varianza del sistema pueden ser utilizados en la calibración de un modelo de regresión PCR (por sus siglas en inglés *Principal Component Regression*) para la predicción de propiedades relacionadas con los datos analizados (Berenton, 2002).

Otro de los procedimientos de estadística multivariada comúnmente utilizados para la calibración de propiedades corresponde a la regresión por mínimos cuadrados parciales o PLS (*Partial Least Square*). Similar al análisis PCA, en el PLS se concentra la información contenida en las variables originales en un número menor de variables denominadas latentes; a diferencia del PCA, el PLS realiza la descomposición de los datos iniciales siguiendo la ruta de máxima covarianza con el vector de propiedades a correlacionar (variables dependientes) (Berenton, 2002; Díaz & Morales, 2012). La regresión se realiza sobre las nuevas variables, eliminando así los problemas de colinealidad que puedan tener los datos originales. En los modelos PLS, se establece una representación de la matriz X de datos iniciales en términos de los *scores* y los *loadings* de cada variable latente. El número de variables latentes necesarias para explicar la variación de la matriz X (dentro de un error máximo permitido) es una medida de la complejidad del modelo (Berenton, 2002).

2.4 Alcances

Este trabajo de aplicación tiene como alcance la proposición de modelos de predicción para el rendimiento de DMO, según los escenarios operativos identificados en los datos históricos de proceso de la unidad DEMEX. Asimismo, la operación de la unidad DEMEX será representada por una simulación desarrollada en el programa HYSYS, con lo cual se analizarán los desempeños en cuanto a rendimiento de DMO, carbón Conradson del DMO y evaporación del solvente según la composición especificada para el solvente.

3. Metodología

3.1 Recolección de una base de datos con los históricos de proceso

Esta etapa envuelve una recolección de datos desde el programa en línea de la refinería de Barranca, denominado PI. Asimismo, las muestras recolectadas son reducidas descartando datos atípicos con posibles errores, utilizando los rangos intercuartíles en cada variable; las muestras a descartar se encuentran fuera del límite superior $Q_3 + 1.5 * (Q_3 - Q_1)$ y límite inferior $Q_1 - 1.5 * (Q_3 - Q_1)$, donde Q_i corresponde al cuartil i (Berenton, 2002; Mendenhall & Sincich, 1996).

3.2 Identificación de escenarios operativos por análisis PCA y el método de agrupación *k-means*

Esta etapa aplica la metodología de los componentes principales (PCA) para el análisis de la base de datos (variables originales) pretratadas según el autoescalado (i.e. valor de la variable menos su media, divididos entre la desviación estándar). Con esta metodología se obtienen nuevas variables no correlacionadas (componentes principales), las cuales reportan valores característicos de *loadings* y *scores*. Con estos valores se determina la interrelación entre las variables de proceso, las composición del solvente y las propiedades de la carga en el espacio PC1 vs PC2 vs PC3. También, el método de agrupación *k-means* fue aplicado para detectar conjuntos de muestras con tendencias similares (*cluster* en inglés) denominados en este documento como clases. El método de agrupación es aplicado considerando desde 2 hasta 7 clases y sus resultados son visualizados en el espacio de los *scores*. La verificación del número adecuado de clases es realizada por medio de los conceptos de cohesión y separación (Berenton, 2002). Complementariamente, los *loadings* expresan la relación entre las clases y las variables que más influyen en su formación. Para fines aplicativos estas clases constituyen los escenarios operativos para la unidad DEMEX.

3.3 Aplicación de regresiones múltiples lineales

El desarrollo de regresiones para el rendimiento de DMO en cada escenario operativo es obtenida por medio de regresión lineal múltiple (MLR) y mínimos cuadrados parciales (PLS); estas regresiones son aplicadas a los datos de cada clase y a sus subdivisiones obtenidas por medio de nuevos PCA y *k-means* (k=2, dos subdivisiones). Por otra parte, las regresiones consideraron 170

entradas formadas por las variables originales y sus interacciones dobles. Para MLR, el método de selección *backward elimination* (eliminación hacia atrás) fue aplicado para construir las ecuaciones de regresión con las variables estadísticamente significativas ($p\text{-value}<0.05$), con la mayor contribución a la varianza del rendimiento de DMO. Asimismo, la regresión por mínimos cuadrados parciales (PLS) fue aplicada en cada escenario siguiendo la metodología de descomposición de acuerdo a la máxima covarianza para obtener las variables latentes (Berenton, 2002). Los estadísticos R² y RMSE fueron considerados para determinar las mejores regresiones en cada clase y sus subdivisiones.

Las regresiones son validadas aplicando la técnica de validación cruzada (*cross-validation*) durante la calibración de los modelos. Asimismo, el número de variables latentes que se utilizaron para las regresiones PLS fueron seleccionadas de acuerdo a la tendencia en el valor de RMSE obtenido en la etapa de validación cruzada. La aparición de una tendencia constante en la gráfica de RMSE indica un número de factores o variables latentes adecuado para la reproducción de la varianza del rendimiento a DMO. Los cálculos fueron realizados utilizando el programa libre R y el programa de uso comercial *TheUnscrambler X* licenciado para ECOPETROL y la UIS.

3.4 Simulación de la unidad DEMEX

Los principales tratamientos aplicados en la unidad de extracción con solvente, DEMEX, son codificadas en el programa comercial de simulación HYSYS. El paquete termodinámico a utilizar corresponde a una variación del modelo Peng-Robinson que incluye mezclas no ideales (Shi & Liang, 2011). El código de simulación se ejecuta considerando diferentes mezclas de solvente para determinar su efecto en el rendimiento de DMO, su contenido de carbón residual (CCR) y la

cantidad de solvente evaporado durante el proceso. El solvente actual utilizado en la unidad DEMEX es evaluado con base en los anteriores valores de desempeño.

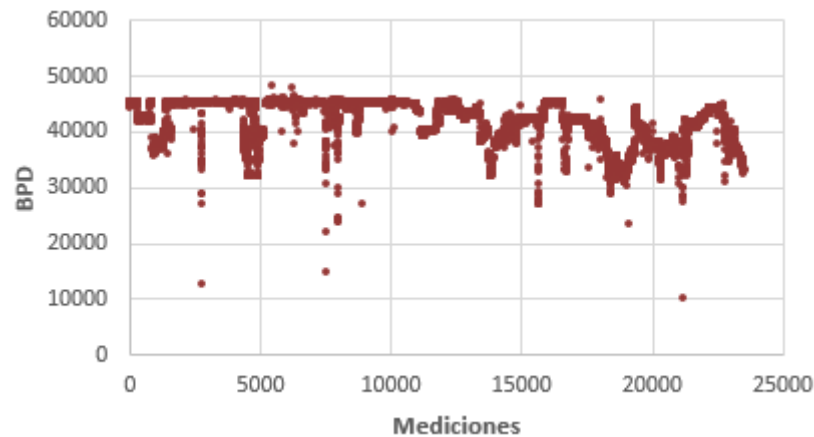
4. Análisis de resultados

4.1 Datos históricos de la unidad DEMEX

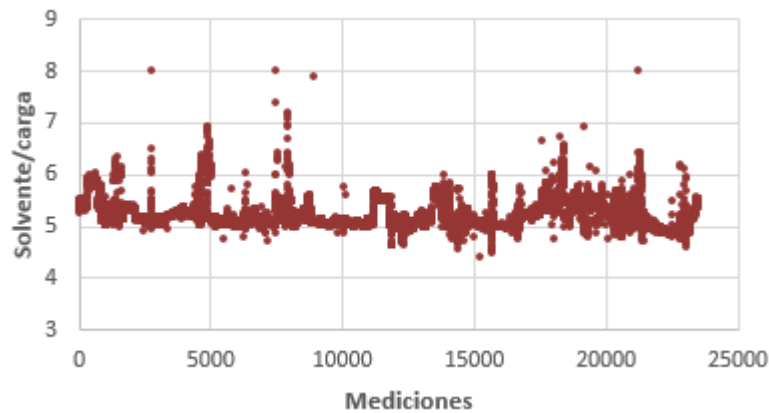
La base de datos históricos de la unidad DEMEX analizada en el presente trabajo incluye las propiedades físicas del crudo API, residuo de carbón por el método Conradson (CCR), vanadio y penetración, composición del solvente y variables operacionales; la lista completa de los parámetros y variables que conforman los datos históricos se encuentra en la Tabla 3.1. El tamaño de la base de datos recolectada fue de 26281x19 que corresponde al número de mediciones cada hora desde el primero de enero de 2014 hasta el 31 de diciembre de 2016 para las 19 variables disponibles de la unidad. El análisis de los datos históricos mostró que la variable relación solvente-carga (S_Q) y el rendimiento (%R) presentaban valores “Bad Input” y “I/O”, así como valores negativos; estas mediciones erróneas fueron eliminadas, con lo cual la base se redujo a 23533 muestras.

La Figura 3.1a presenta la variación de las cargas a DEMEX según la base de datos de históricos. De acuerdo a esta figura, las cargas a la unidad presentan zonas en estados dinámico y cuasi-estacionario en la ventana de operación analizada; asimismo, diferentes muestras atípicas (*e.g.* puntos con baja carga debido a inconvenientes operacionales y mantenimientos, Figura 3.1

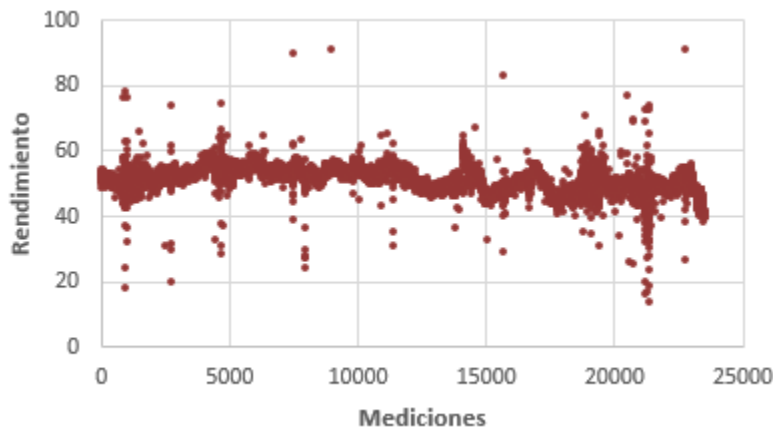
valores por debajo de 36 kbps) pueden ser encontradas en la base de datos; estas muestras atípicas pueden ser identificadas en las variables S_Q y %R de DMO (Figuras 3.1b y 3.1c, respectivamente). Para trabajar sólo con los estados cuasi-estacionarios, la base de datos fue sometida a una depuración inicial omitiendo las muestras atípicas según el rango intercuartíl de cada variable; de esta forma la base de datos fue reducida a 23113 muestras. Sin embargo, esta depuración no garantiza la eliminación completa de estados de operación dinámicos, por lo cual, diferentes tratamientos estadísticos multivariados reportarán muestras atípicas; asimismo, la identificación de agrupaciones o escenarios operativos puede presentar zonas de difícil delimitación en las diferentes variables; esto debido a la acumulación de muestras en estado dinámico en las fronteras de los escenarios operativos.



(a) Carga de entrada a DEMEX



(b) Relación solvente-carga



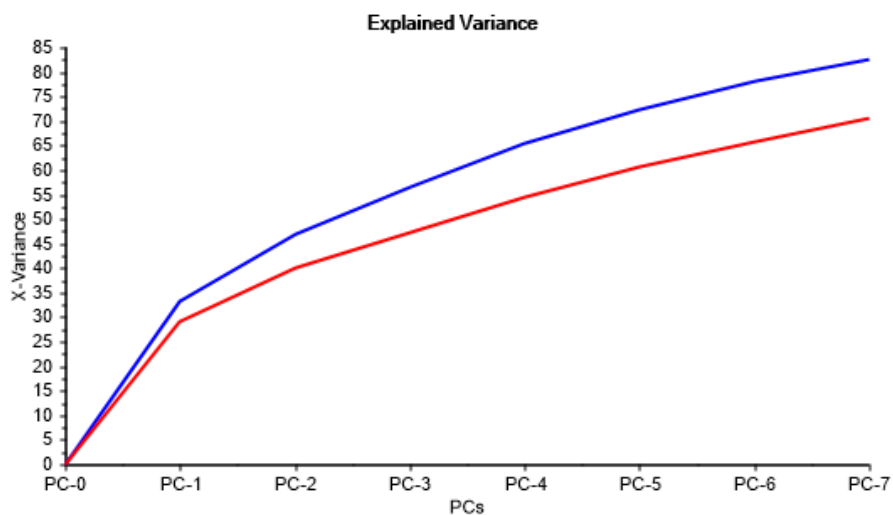
(c) Rendimiento de DMO

Figura 2. Tendencia para algunas variables contenidas en la base de datos históricos para el proceso DEMEX; ventana operativa entre enero de 2014 y diciembre de 2016.

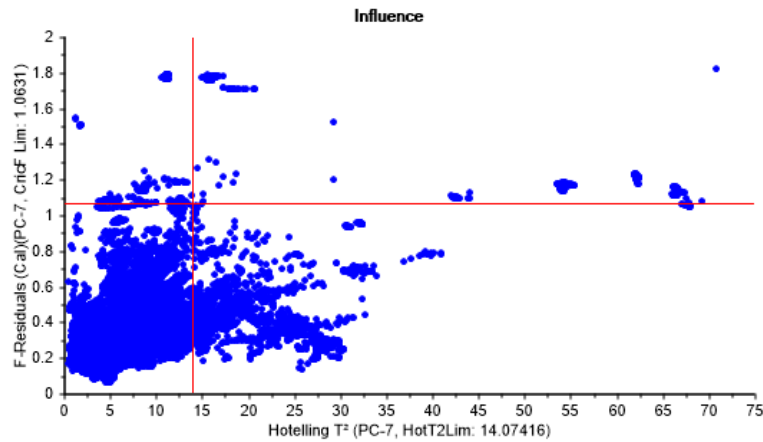
4.2 Escenarios de operación para la unidad DEMEX

El conjunto de datos independientes (*i.e.* sin las variables rendimiento de DMO, ni el valor de su CCR) fue sometido a una reducción de variables por medio del análisis por componentes principales. El análisis PCA fue aplicado a la matriz de datos autoescalados para obtener 7

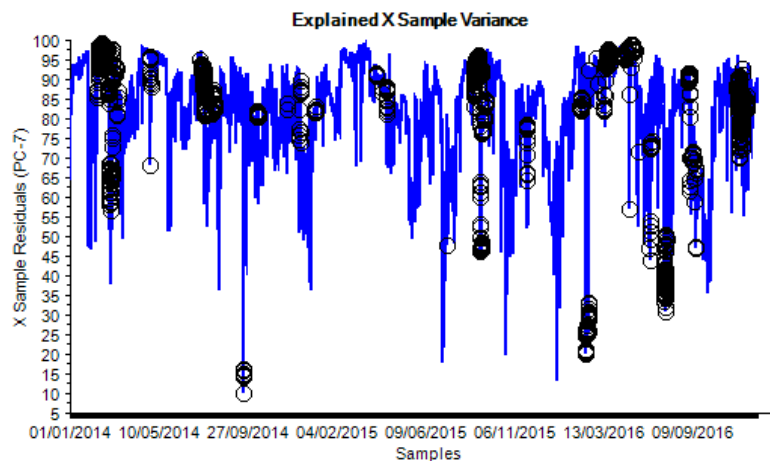
componentes principales (PC). Según los resultados del PCA, los 7 componentes principales reproducen 82.7% de la varianza total de los datos históricos (Figura 3.2a), con una concordancia entre los procedimientos de calibración (línea azul) y validación (línea roja); con lo anterior, la varianza de las 17 variables originales es concentrada en sólo 7 variables nuevas no correlacionadas entre sí. Según el gráfico de influencia (Figura 3.2b), una porción importante de muestras pertenecen a una población diferente –con T^2 superior a 14.07 para PC-7–; a pesar de utilizar 7 PC y obtener una varianza explicada para el conjunto de 82.7% en calibración, los valores del estadístico de Hotelling para una gran cantidad de muestras, sugieren la búsqueda de subconjuntos que representen diferentes esquemas operativos en los datos históricos de la unidad DEMEX. La Figura 3.2c exhibe que la varianza explicada para diferentes muestras es inferior a la del conjunto para PC-7; los puntos con círculos negros en esta figura corresponden a las muestras con $T^2 > 14.07$ y $F > 1.06$ para PC-7 (Figura 3.2b). Asimismo, el gráfico de influencia presenta puntos atípicos (primer cuadrante) que podrían ser omitidos; sin embargo, no son eliminados del conjunto ya que pueden pertenecer a un esquema operativo en particular.



(a) Varianza acumulada explicada del conjunto (PC-7: 82.7% calibración, 70.6% validación).



(b) Gráfico de Influencia.

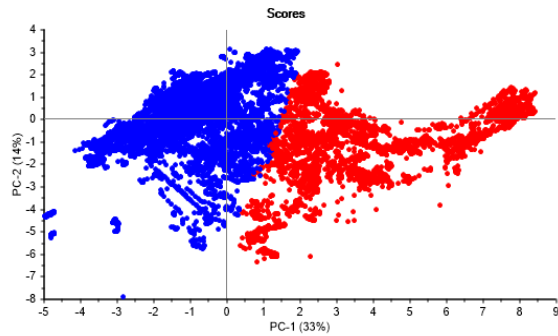


(c) Varianza explicada por muestra.

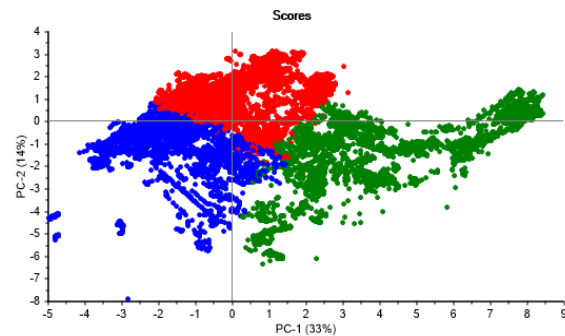
Figura 3. Análisis PCA a la base de datos históricos de DEMEX con pretratamiento de autoescalado. (a) varianza explicada acumulada del conjunto. (b) Gráfico de influencia para PC-7. (c) Residuales por muestra para PC-7; círculos negros corresponden a las muestras con estadísticos T^2 y F fuera de los respectivos límites (Figura b).

En este sentido, el método de agrupamiento denominado *k-means* fue aplicado sobre el conjunto de valores independientes autoescalados, considerando la distancia euclidiana para clasificar los datos en diferentes grupos. En la Figura 3.3, la gráfica de los *scores* para los dos primeros PC ilustra con colores las diferentes agrupaciones (hasta 5 clases) aplicadas por medio del método *k-*

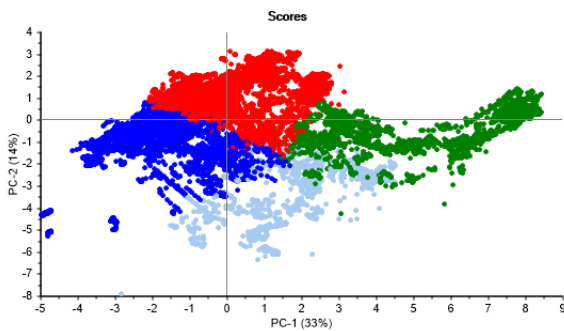
means. Un análisis de los valores de los PC revela que las clase 0 y 1 se forman debido a los valores positivos y negativos del primer componente principal (33% varianza) (Figuras 3.3a y 3.4). Las variables más influyentes en cada componente pueden ser analizadas en la gráfica de los *loadings*, Figura 3.5; con la definición de $k=2$, los datos se dividen en clase 0 y clase 1 caracterizada por diferentes valores de: carga (clase 0 mayor a 40 kbps y clase 1 menor a 40 kbps), isobutano en el solvente (clase 0 mayor a *ca.* 30% y clase 1 menor a *ca.* 30%) y temperaturas T_{2520} y T_{R2520} (clase 0 menor a *ca.* 160°F y clase 0 menor a *ca.* 160°F). Al aumentar el número de agrupaciones, las divisiones se presentan principalmente por los valores del eje del componente principal 2; Por otra parte, la clase 2 se caracteriza por los valores positivos del segundo PC (14% varianza) (Figura 3.4), mientras la clase 3 corresponde a aquellas muestras que poseen valores de PC-2 por debajo de -2; las clases 2 y 3 se encuentran conformadas principalmente por valores bajos para las temperaturas T_{2520} y T_{R2520} (valores menores a 200 °F) y por valores elevados de las propiedades CCR y V (mayores que 20 y 230, respectivamente).



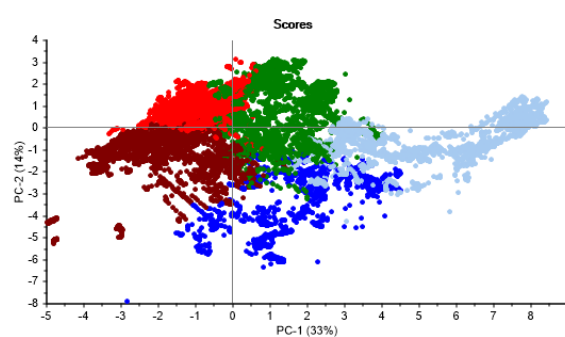
(a) Dos agrupaciones.



(b) Tres agrupaciones.



(c) Cuatro agrupaciones.



(d) Cinco agrupaciones.

Figura 4. Gráfica de los *scores* para los dos primeros componentes principales (47%) derivados del PCA. Las agrupaciones fueron obtenidas por el método *k-mean* utilizando la distancia euclidiana.

Es importante mencionar que los límites entre las agrupaciones se tornan más difusos a medida que aumenta el número de clases. De este modo, para las clases 5 y 6 no se visualizan rangos operacionales delimitados en forma simple. A nivel cuantitativo, la definición del número de clases se logra por medio de las tendencias de la varianza interna total no ponderada en las clases y la varianza no ponderada entre las clases (Rousseeuw, 1987 y referencias). La Figura 3.6 muestra las tendencias mencionadas respecto al número de clases, lo cual sugiere que cuatro clases agrupan adecuadamente las varianzas presentadas por los datos de la unidad DEMEX; con cuatro clases se alcanza una cohesión interna o similaridad suficiente en los datos de las clases (varianza interna

total), con una separación apreciable entre las diferentes clases (varianza no ponderada entre las clases), visualizadas en el punto de mayor reducción según la segunda derivada (Figura 3.6).

También, la relación entre la cohesión interna de los datos en las clases y la separación entre las agrupaciones indica el mejor número de clases para los datos; esta relación es denominada como coeficiente silueta (*silhouette coefficient* en inglés), el cual relaciona para cada dato la cohesión y la separación alcanzada con las clases, según $1-a/b$, con a como el valor promedio de la distancia al punto en su clase y b como la mínima distancia entre el punto y las otras clases (Rousseeuw, 1987). En las Figuras 3.7 se presentan en forma gráfica los coeficientes silueta para los datos de las clases 2 – 7 obtenidos por medio del programa de uso libre R; según los valores presentados en estas figuras, el mayor valor para el coeficiente silueta promedio (0.37) se alcanza con un número de clases igual a cuatro (Figura 3.7c). Este número de clases coincide con el criterio del punto de mayor cambio en la cohesión y la separación de las agrupaciones. Por lo tanto, los análisis posteriores serán aplicados considerando las cuatro clases ya presentadas; la población de datos en cada clase puede ser comparada en forma cualitativa en la Figura 3.8.

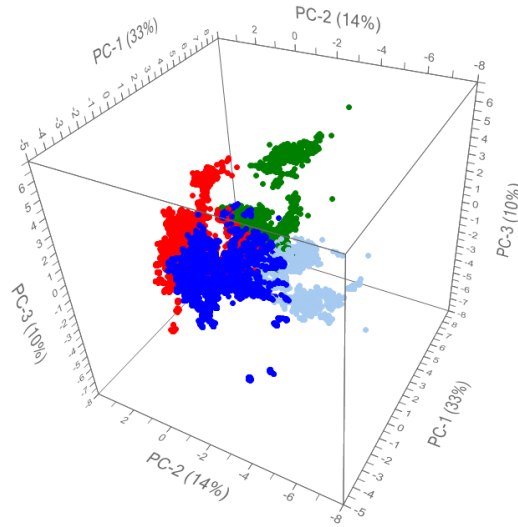


Figura 5. Gráfica de los *scores* para los tres primeros componentes principales (57%) derivados del PCA, con cuatro agrupaciones ($k=4$).

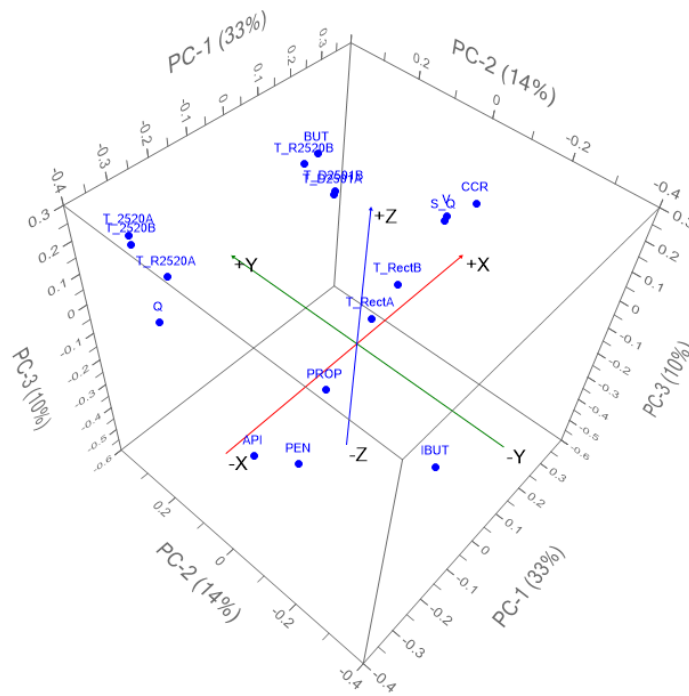


Figura 6. Gráfica de los *scores* para los tres primeros componentes principales (57%) derivados del PCA.

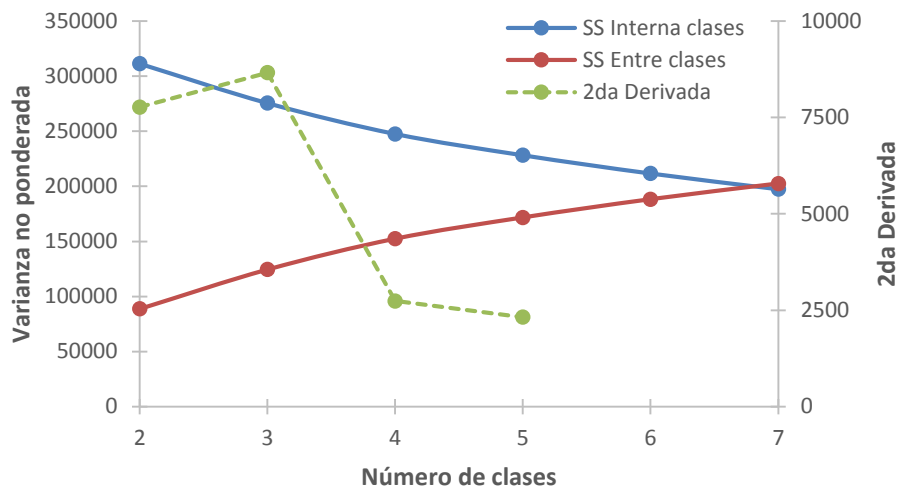
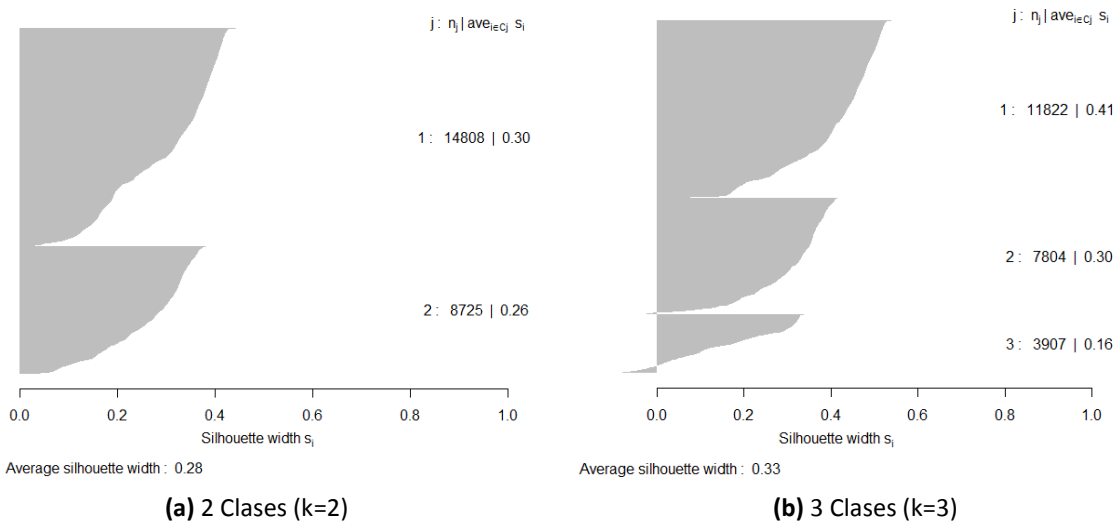
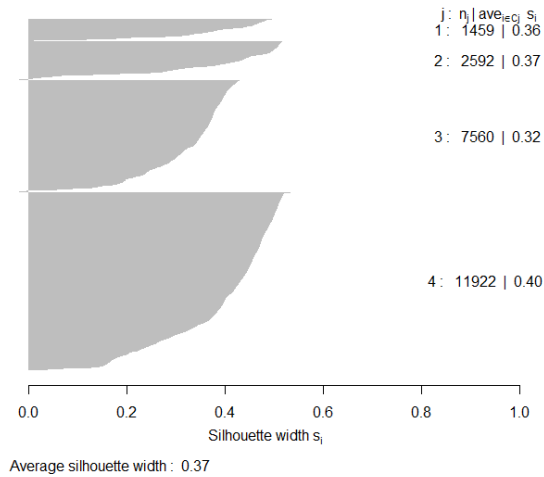


Figura 7. Tendencias para la varianza interna total no ponderada en las clases (SS Interna clases) y la varianza no ponderada entre las clases (SS Entre clases) respecto al número de clases según resultados obtenidos con el programa de uso libre R. El valor de la segunda derivada denota que la cohesión y la separación adecuadas en la agrupación son alcanzadas con cuatro clases.

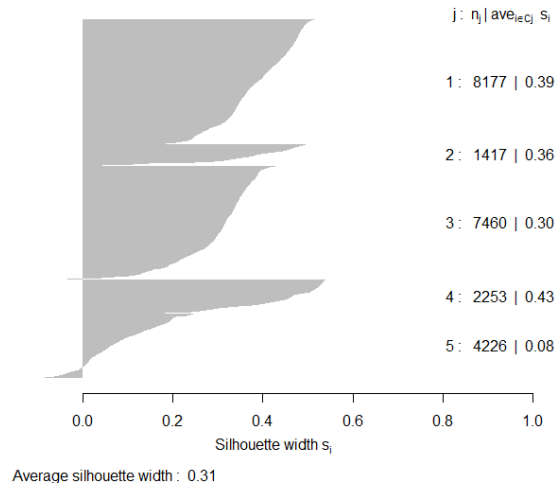


(a) 2 Clases (k=2)

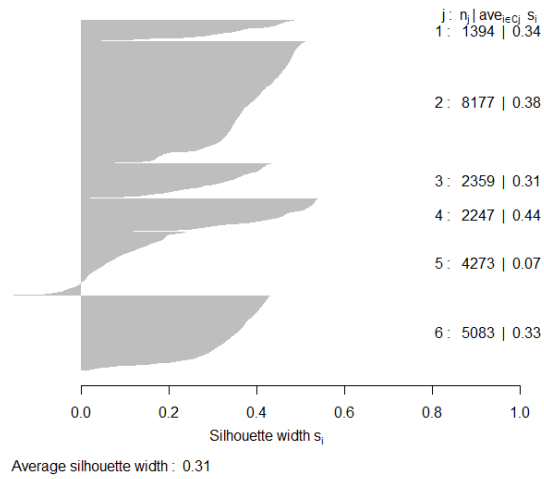
(b) 3 Clases (k=3)



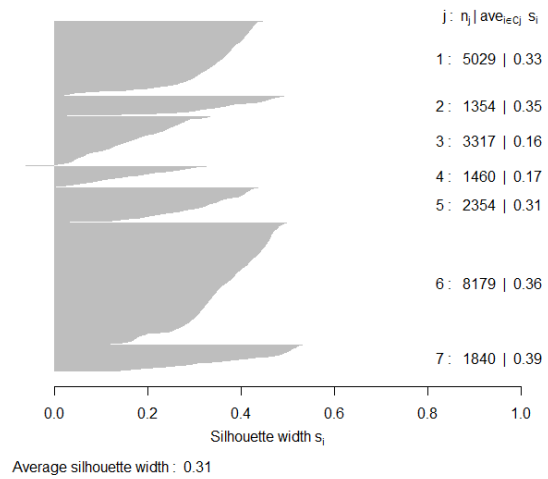
(c) 4 Clases (k=4)



(d) 5 Clases (k=5)



(e) 6 Clases (k=6)



(f) 7 Clases (k=7)

Figura 8. Gráficas de los coeficientes silueta para las clases 2 – 7, según resultados con el programa R.

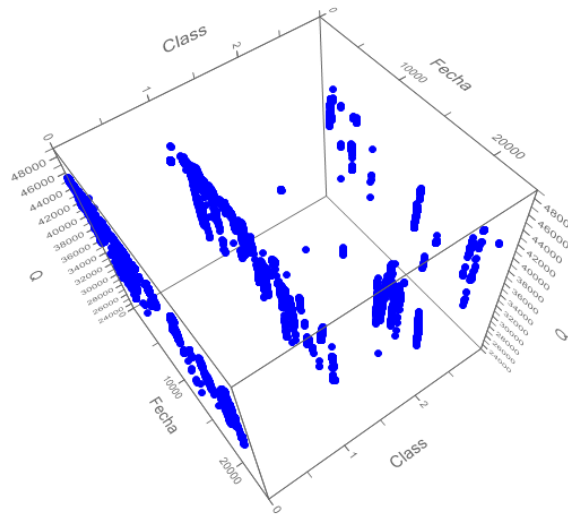


Figura 9. Gráficas de clases para la ventana de operación y la carga al proceso.

A nivel práctico, las clases representan diferentes escenarios operativos para la unidad; *i.e.* las clases sugieren que los datos históricos no presentan datos distribuidos aleatoriamente sino, en patrones definiendo condiciones operacionales. Las Tablas 3.1 y 3.2 presentan los diferentes intervalos en los que las variables definen la formación de las clases para la unidad DEMEX; estos intervalos consideran una inclusión del 95% de los datos para cada variable en cada clase. Los valores de los límites que definen cada clase se encuentran soportados en las Figuras 3.9-3.10; en estas figuras es posible inferir que las diferentes clases presentan límites según las variables de proceso y las propiedades de la carga. Según los valores de la Tabla 3.1, las clases 0 y 1 se distinguen de las clases 2 y 3 por sus mayores valores de T_{2520} , T_{R2520} y Q (Figuras 3.9 y 3.10); las clase 0 y 2 se diferencian de las clases 1 y 3, respectivamente, debido principalmente a los menores valores de BUT y T_{D2501} (Figuras 3.9f y 3.10). Las Figuras 3.10 resume las diferencias en la operación de la unidad DEMEX de acuerdo a tres variables de mayor influencia en las clases; según estas figuras, la clase 0 está caracterizada por valores altos de Q y valores bajos de T_{D2501} y BUT (Figura 3.10b); la clase 1 está caracterizada por valores altos de Q ,

T_D2501 y BUT (Figura 3.10c); la clase 2 se encuentra definida por valores bajos de Q y valores altos de T_D2501 y BUT (Figura 3.10d), mientras que la clase 3 se encuentra definida por valores bajos de Q, T_D2501 y BUT (Figura 3.10e). Por otra parte, las clases obtenidas por *k-means* no presentan tendencias definidas para el rendimiento a DMO y su contenido de carbón Conradson (Figuras 3.11). Según la Figura 3.11 no se presenta relación aparente entre los escenarios operacionales y los rendimientos a DMO y su contenido de CCR. En la siguiente sección diferentes regresiones serán calibradas para %R en cada escenario operativo.

Tabla 1.

Valores para las variables de operación según la clase o escenario operativo.

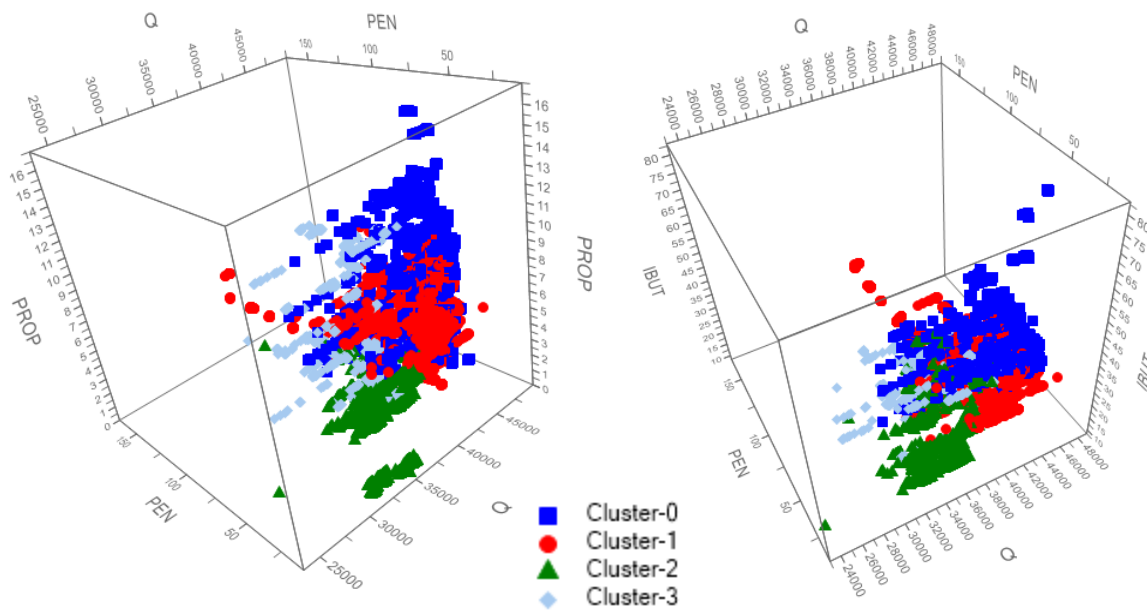
	T_D250 1A, °F	T_D250 1B, °F	T_2520 A, °F	T_2520 B, °F	T_Rect A, °F	T_Rect B, °F	T_R252 0A, °F	T_R252 0B, °F
Clase 0	[198; 210]	[198; 209]	[200; 230]	[105; 215]	[224; 240]	[220; 249]	[198; 255]	[190; 300]
Clase 1	[203; 218]	[204; 219]	[200; 236]	[90; 224]	[228; 242]	[223; 249]	[200; 253]	[201; 300]
Clase 2	[208; 221]	[207; 225]	[100; 190]	[85; 175]	[228; 241]	[223; 249]	[110; 225]	[175; 292]
Clase 3	[198; 212]	[197; 215]	[100; 150]	[85; 173]	[230; 243]	[201; 251]	[115; 215]	[94; 258]

	API	CCR, %	V, %	Q, kbps	PEN, mm	PROP, %	IBUT, %	BUT, %	S_Q
Clase 0	[4.4; 8.0]	[18.5; 24]	[156; 333]	[36.1; 45]	[29; 70]	[4.0; 13]	[25; 48]	[21; 60]	[4.7; 6.9]
Clase 1	[5.0; 7.1]	[18.8; 23]	[197; 348]	[37; 45.5]	[17; 97]	[3.9; 8.9]	[23; 40]	[49; 69]	[4.4; 6.2]
Clase 2	[3.0; 6.0]	[21.5; 27]	[249; 350]	[31.1; 38.8]	[15; 49]	[0.0; 5.0]	[18; 41]	[54; 74]	[4.7; 7.0]
Clase 3	[4.5; 7.0]	[20; 23]	[229; 333]	[32.8; 38.2]	[34; 71]	[4.0; 11]	[18; 42]	[48; 61]	[4.8; 6.4]

Tabla 2.

Valores para el rendimiento a DMO y su contenido de carbón Conradson para los escenarios.

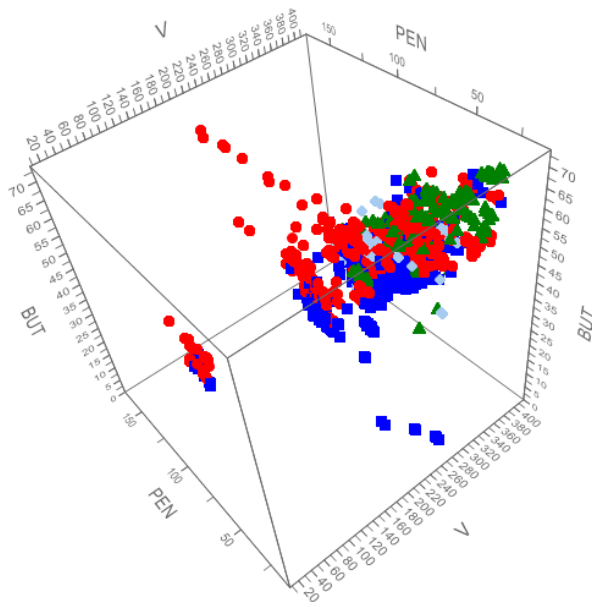
	%R	CCR_DMO, %
Clase 0	[39; 67]	[4.5; 6.6]
Clase 1	[42; 58]	[4.4; 6.8]
Clase 2	[40; 53]	[5.1; 6.8]
Clase 3	[41; 55]	[4.8; 6.2]



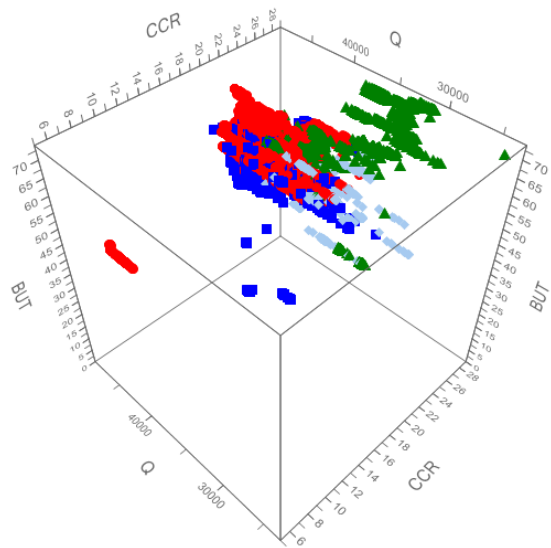
(a) Propano vs Penetración vs Carga

(b) *i*-Butano vs Penetración vs Carga

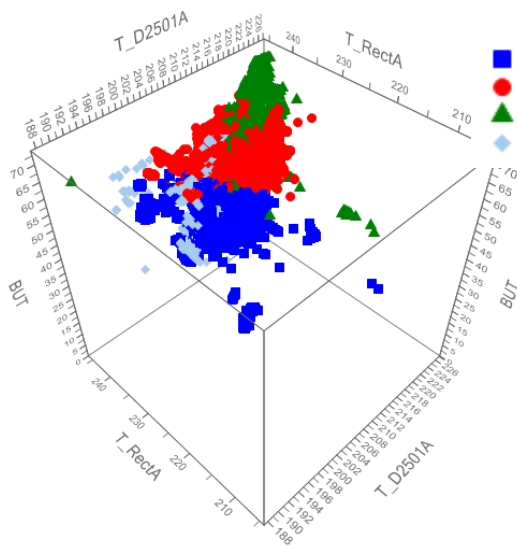
Figura 10. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases.



(c) Butano vs Penetración vs Vanadio

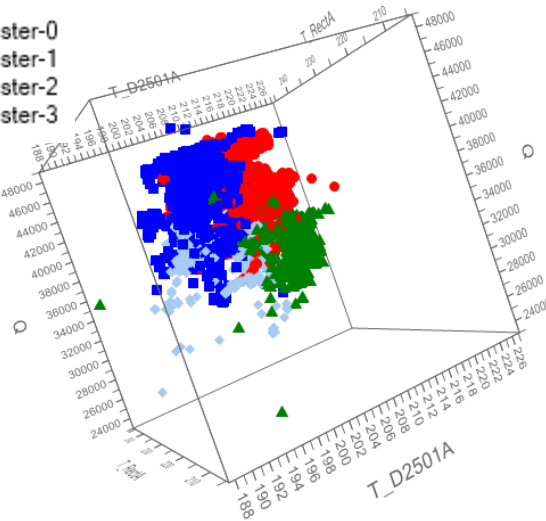


(d) Butano vs Carga vs Residuo de carbón



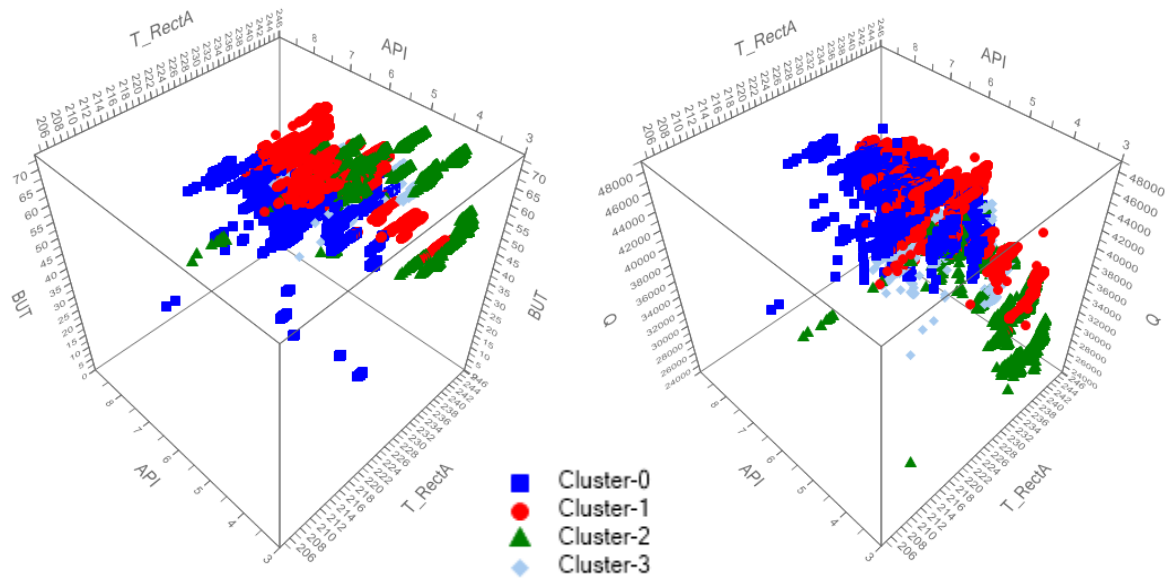
(e) Butano vs T_D2501 vs T_Rect

■ Cluster-0
● Cluster-1
▲ Cluster-2
◆ Cluster-3



(f) Butano vs T_D2501 vs T_Rect

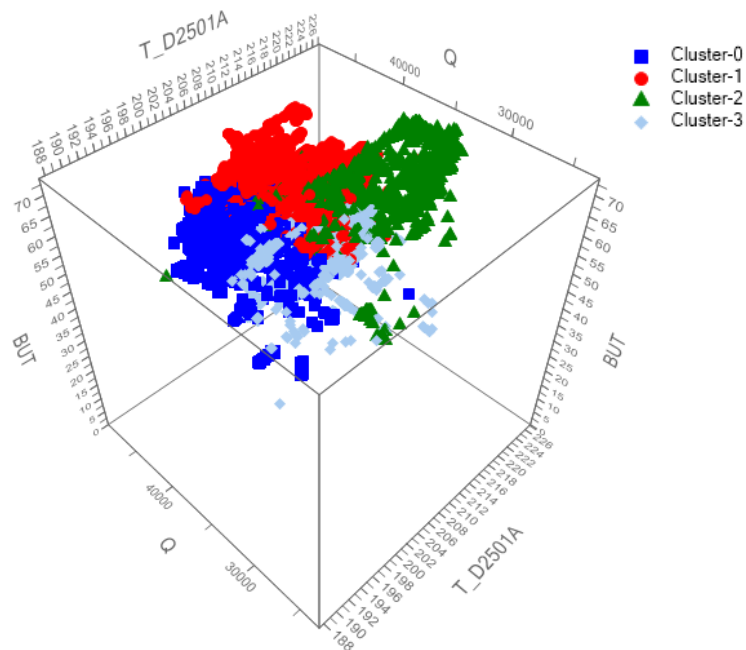
Figura 11. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases (continuación).



(g) Butano vs API vs T_Rect

(h) Carga vs API vs T_Rect

Figura 12. Gráficas de variables operacionales y propiedades de la carga para las diferentes clases (continuación).



(a) Todas las clases

Figura 13. Ilustración clases en el espacio de las variables Butano vs Carga vs Temperatura T_D2501A.

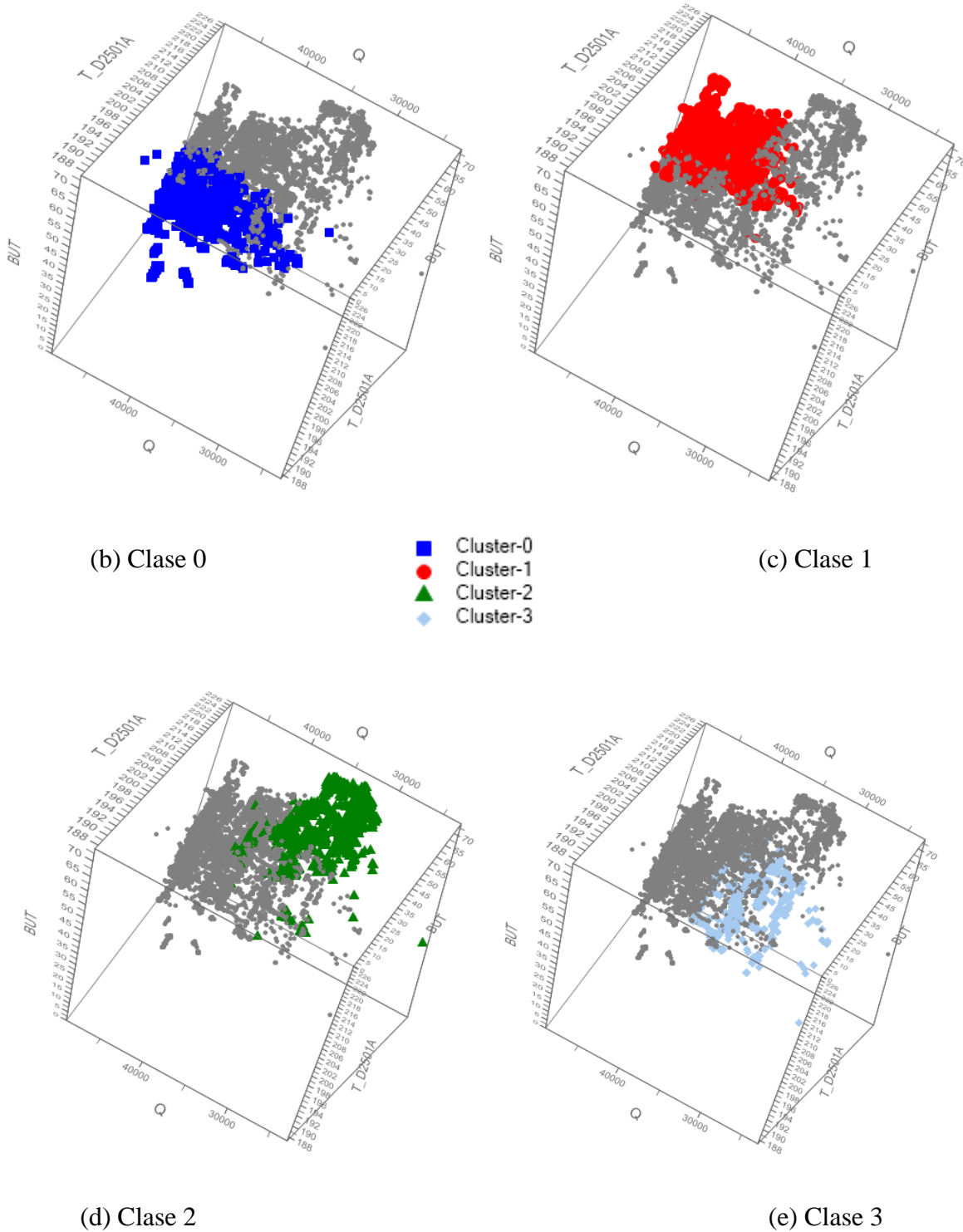


Figura 14. Ilustración clases en el espacio de las variables Butano vs Carga vs Temperatura T_D2501A (continuación).

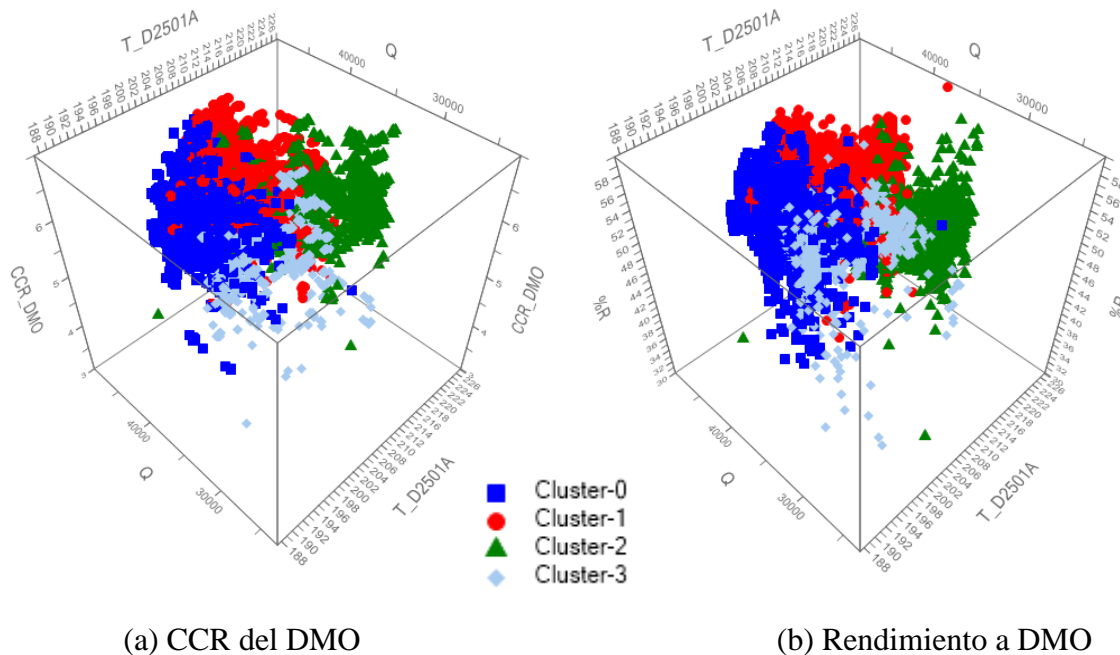
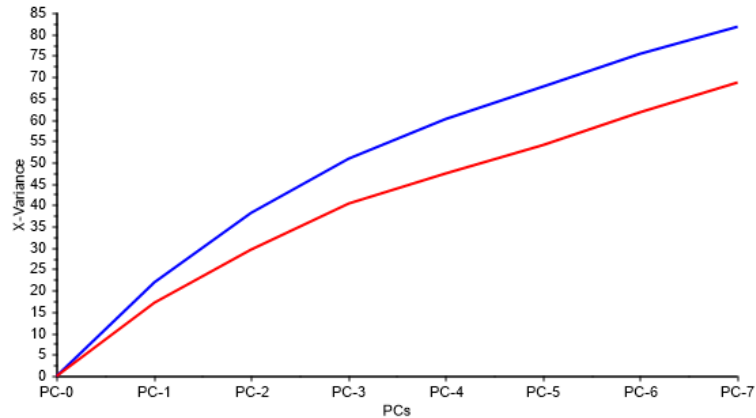


Figura 15. Variables de salida de la unidad DEMEX y su relación con las clases obtenidas.

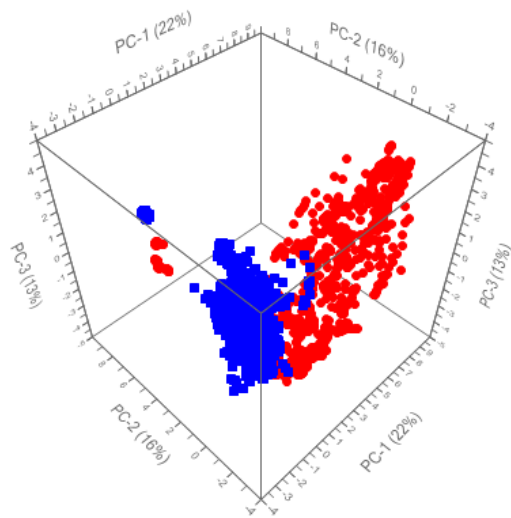
4.3 Regresiones para el rendimiento de DMO en los escenarios operativos de la unidad DEMEX

Con las clases definidas como los escenarios operacionales representativos para la unidad DEMEX, el rendimiento reportado para DMO en cada clase fue modelado utilizando regresiones lineales (MLR) y mínimos cuadrados parciales (PLS) según los algoritmos especificados en el programa comercial *The Unscrambler X*. Las regresiones incluyen efectos cuadráticos de las 17 variables independientes, con lo cual, las entradas para cada modelo correspondieron a 170. Los ajustes de los mejores modelos lineales por MLR fueron obtenidas por medio del procedimiento *backward elimination*, utilizando el criterio de información de Akaike según su codificación en el paquete de uso libre R.

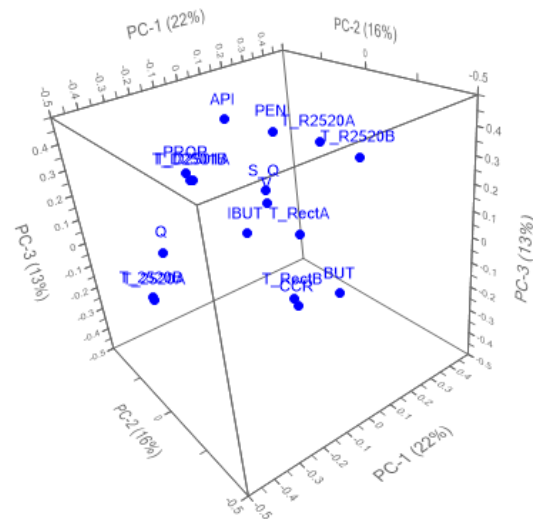
4.3.1 Clase 0 La clase 0 conformada por 7547 muestras fue analizada por medio de los métodos PCA y el método de agrupamiento *k-mean* ($k=2$); los resultados son mostrados en las Figuras 3.11a-c. Según la Figura 3.11, la varianza de los datos es explicada al 81.7% utilizando 7 componentes principales (PC) –68.7% en validación–; asimismo, los resultados del método *k-mean* son mostrados sobre la gráfica de los scores para los tres primeros PC (41%) (Figura 3.11b), denotando la formación de dos subdivisiones (*i.e.* subdivisión 1 con valores de PC1 negativos en color azul y subdivisión 2 con PC1 positivos en color rojo). Complementariamente, la gráfica de los *loading* para los tres primeros PC denota a las variables Q, T_2520 y T_2520 como las influyentes para la formación de la subdivisión 1 (puntos azules, Figura 3.11b), mientras que las variables T_R2520 y T_Rect como las influyentes para la formación de la subdivisión 2 (puntos rojos, Figura 3.11b). Según la dispersión de los datos en las anteriores variables, la formación de los grupos presenta el valor limitante en 205°F en T_2520 (Figuras 3.12); las muestras de la subdivisión 2 denotan una situación operacional en donde las torres de extracción de platos (T-2520A y B) se encuentran fuera de servicio (temperaturas bajas denotados por puntos rojos, Figura 3.12b); por lo general, la torre de extracción de T-2520 apoya al sistema para alcanzar flujos superiores a 40 kbpd (Figura 3.12a). Las otras variables no presentan cambios para la subdivisión de los datos de la clase 0 (Tabla 3.1).



(a) Varianza acumulada explicada hasta 7 componentes principales (calibración 81.7%, validación 68.7%)

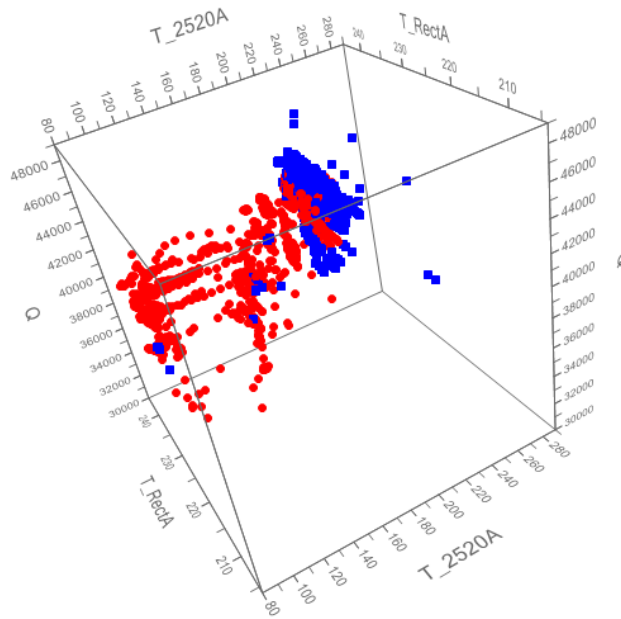


(b) Scores

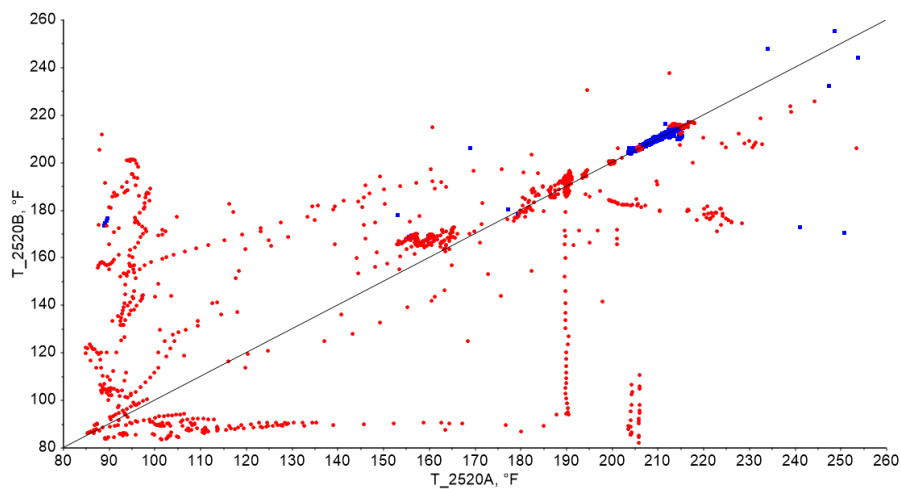


(c) Loadings

Figura 16. Resultados PCA para la clase 0. (a) varianza explicada, 81.7% calibración, 68.7% validación. (b) Gráfica de los *scores* denotando dos ($k=2$) agrupaciones (subdivisión 1 en azul, subdivisión 2 en rojo). (c) Gráfica de los *loadings*.



(a) Gráfica Q vs T_2520A vs T_RectA



(b) Gráfica T_2520B vs T_2520A

Figura 17. Dispersión de los datos para la clase 0 resaltando las subdivisiones obtenidas con el método *k-means* (subdivisión 1 en azul, subdivisión 2 en rojo). (a) Gráfica Q vs T_2520A vs T_RectA. (b) Gráfica T_2520B vs T_2520A.

Las regresiones MLR y PLS fueron aplicadas a todos los datos de la clase 0 y por aparte a los datos de cada subdivisión; el anterior procedimiento fue aplicado para determinar si la subdivisión

de los datos de la clase 0 conlleva a la obtención de regresiones con mejores desempeños para la predicción del rendimiento de DMO (%R). Los resultados para la regresión MLR con todos los datos de la clase 0 son presentados en las Figuras 3.13a-c; según estas figuras, la regresión MLR presenta una utilidad cualitativa, ya que reproduce *ca.* 80% de la varianza del rendimiento (%R, Figura 3.13a), utilizando 48 entradas estadísticamente significativas al 95% de confianza (Figura 3.13b); la expresión matemática de la regresión puede ser establecida utilizando los coeficientes mostrados en la Figura 3.13c (columna “B-coefficient”). Asimismo, según los valores absolutos de los coeficientes obtenidos (Figura 3.13c), la composición del solvente tienen una influencia en la predicción de *ca.* 39% –con 21% de influencia del BUT e igual influencia del IBUT y del PROP (9%)–, mientras que la influencia de las condiciones operativas llega a 44.5% –T_D2501, 16%, T_Rect, 14%, T_R2520, 10%, T_2520, 4.5%– y el resto de influencia en las propiedades de la carga (11%, incluyendo el flujo de alimentación) y de la relación solvente a carga, S_Q (5.5%). Respecto al signo de los coeficientes de las entradas, un aumento en la composición del *n*-butano (BUT) en el solvente, junto con una elevación en la temperatura de la rectificación (T_Rect) conducen a la elevación del rendimiento de DMO (%R); el aumento en el rendimiento se debe principalmente al aumento en la temperatura crítica de la mezcla de solvente (Figura 3.14a) ocasionada por el incremento en la composición de butano –*i.e.* el butano presenta mayor temperatura crítica (153.2°C) que el propano (96°C) y el *i*-butano (134.7°C)–. En este sentido, el aumento en la composición del *i*-butano (IBUT) o del propano (PROP) conducen a un incremento menor que el obtenido por el BUT, lo cual coincide con los reportes de la literatura abierta (Ahn *et al.*, 2015; Campos *et al.*, 1995; Fahai *et al.*, 2009). Por otra parte, un aumento en las temperaturas T_D2501 y T_R2520 conduce a la disminución del rendimiento de DMO (%R), lo cual coincide con los datos operacionales especificados por la tecnología DEMEX (Figura 3.14b). Según la

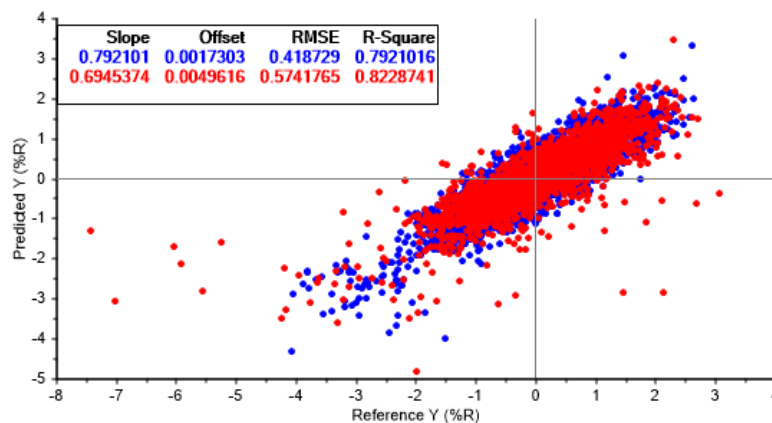
Figura 3.14b, la capacidad extractora del solvente disminuye con la temperatura base (ca. 205°F dependiendo de la temperatura crítica del solvente, Figura 3.14a), hasta encontrar su menor valor en la condición crítica del solvente.

La regresión PLS con todos los datos de la clase 0 reporta menor desempeño que la obtenida con la regresión MLR (Figura 3.15). La varianza del %R predicha con la regresión PLS alcanza sólo el 72%, con lo cual no es posible una aplicación práctica para la misma. Esta regresión se encuentra formada por 6 factores resultantes de la combinación lineal de 25 entradas; aunque contiene menor número de entradas que la regresión MLR, el desempeño de la regresión PLS no soporta una aplicación a nivel cualitativo (Tabla 3.2).

Para la subdivisión 1, conformada por 5479 muestras, la regresión MLR presenta un mejor desempeño que el obtenido con todos los datos de la clase 0 (Figura 3.16, Tabla 3.2), aunque con un mayor número de entradas; la regresión MLR alcanzó una reproducción del 95% de la varianza (95.1% en validación), con una desviación de 0.208 (RMSE=0.212 en validación). El desempeño alcanzado por la regresión MLR en la subdivisión 1 soporta su aplicación a nivel cuantitativo. La Figura 3.16b valida la suposición de residuos normalizados con varianza constante, mientras que la Figura 3.16e presenta la tabla ANOVA (copiada directamente de *The Unscrambler X*) valida la significancia estadística para la regresión y sus entradas. Asimismo, la Figura 3.16d muestra que las variables de mayor influencia en el rendimiento de DMO para la subdivisión 1 de la clase 0 son T_Rect (15.5%), PROP (12.9%) y T_2520 (11.5%); las temperaturas de operación de las secciones de extracción impactan el %R en un 44.2% (T_D2501, T_R2520, T_2520 y T_Rect), mientras que la composición del solvente y las propiedades de la carga influyen en un 29.6% y un 10.5%, respectivamente (Figura 3.16d); la mayor influencia de las temperaturas en el proceso de

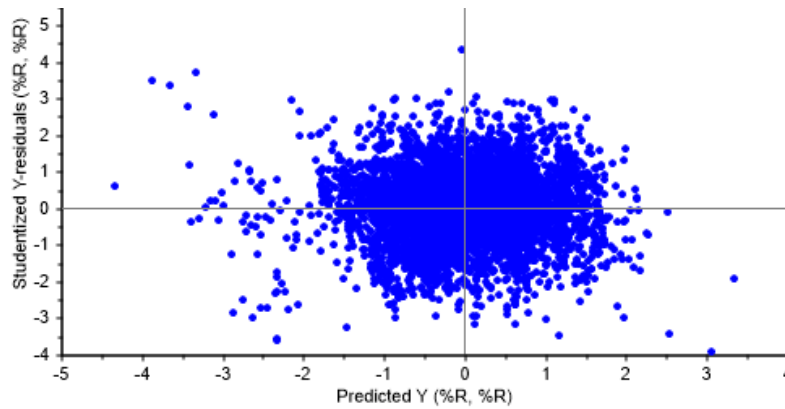
extracción se debe a que la cantidad de solvente utilizada se encuentra por encima de la necesaria (influencia de S_Q de 6.1%).

De otro lado, la regresión PLS para la subdivisión 1 reporta un desempeño bajo (ca. 80% varianza de rendimiento), con lo cual, sus posibilidades de aplicación se reducen al nivel cualitativo (Figura 3.17, Tabla 3.2); la forma matemática de la regresión PLS es mostrada en (1). Según los coeficientes de (1), la carga y sus propiedades afectan un 36.8% el rendimiento obtenido en el proceso (5.8% la carga Q, 16% el residuo de carbón CCR, 12% la penetración PEN y 3% el API), mientras que las variables operacionales afectan un 35.5% y la composición del solvente utilizado un 25%. Asimismo, el %R en este esquema operativo disminuye con el aumento en CCR y PEN en el flujo cargado a la unidad. También, el aumento en las temperaturas T_2520A y B, T_RectB y T_R2520A disminuyen el %R del proceso.



(a) Gráfica valor predicho vs valor de referencia.

Figura 18. Regresión MLR para todos los datos de la clase 0.



(b) Residuales estudentizados.

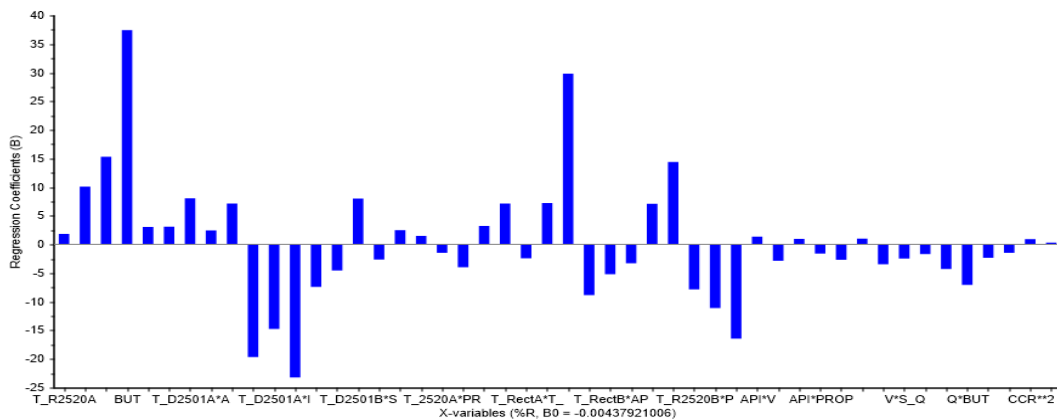
Anova Table							
Multiple Correlation: 0.8900009 (cal) 0.8203433 (val)							
R-Square: 0.7921016 (cal) 0.8228741 (val)							
	SS	df	MS	F ratio	p value	B-coefficient	STDerr
Summary							
Model	3390.91	48	70.64407	399.0221	0		
Error	889.995	5027	0.177042				
Adjusted Total	4280.91	5075	0.843529				
Variables							
Intercept	0.09663	1	0.096632	0.545812	0.460068196	-0.0043792100	0.0059275
T_R2520A	13.5225	1	13.52254	76.38079	3.14090777E-18	1.89467418	0.2167916
T_R2520B	25.2543	1	25.25435	142.6446	1.92450157E-32	10.1149168	0.8469048
PROP	33.9825	1	33.98255	191.9446	7.27273903E-43	15.3293743	1.1064617
BUT	176.365	1	176.3650	996.1608	0	37.4376869	1.1861627
S_Q	4.17591	1	4.175915	23.58704	1.23010238E-06	3.08685613	0.6355937
T_D2501A*T_R25	4.12638	1	4.126382	23.30715	1.42178158E-06	3.12263155	0.6468091
T_D2501A*API	32.1137	1	32.11372	181.3894	1.21257159E-40	8.09384251	0.6009643
T_D2501A*V	5.66703	1	5.667032	32.00944	1.61917004E-08	2.48477983	0.4391863
T_D2501A*Q	120.273	1	120.2735	679.3424	0	7.16701031	0.2749753
T_D2501A*PROP	66.5222	1	66.52220	375.7391	0	-19.6408539	1.0132513
T_D2501A*IBUT	39.5725	1	39.57256	223.5181	0	-14.7328434	0.9854398
T_D2501A*BUT	95.3549	1	95.35492	538.5936	0	-23.1806259	0.9988368
T_D2501B*T_252	36.5180	1	36.51807	206.2663	1.40129846E-45	-7.40304947	0.5154617
T_D2501B*T_R25	11.5422	1	11.54220	65.19432	8.42241441E-16	-4.53085613	0.5611453
T_D2501B*S_Q	30.5464	1	30.54649	172.5371	8.93244394E-39	8.04651833	0.6125854
T_2520A*T_R252	137.269	1	137.2694	775.3467	0	-2.58414221	0.0928043
T_2520A*API	40.0175	1	40.01756	226.0328	0	2.53509688	0.1686199
T_2520A*V	13.2767	1	13.27672	74.99137	6.28229572E-18	1.51572657	0.1750310
T_2520A*PROP	30.3211	1	30.32110	171.2637	1.65910767E-38	-1.42345917	0.1087707
T_2520A*BUT	27.9629	1	27.96291	157.9440	1.08977173E-35	-3.93797708	0.3133441
T_2520A*S_Q	50.8140	1	50.81408	287.0149	0	3.28464794	0.1938815
T_2520B*T_RectA	36.3253	1	36.32538	205.1780	1.40129846E-45	7.17750788	0.5010812
T_RectA*T_RectB	14.0331	1	14.03316	79.26374	7.46183811E-19	-2.39868355	0.2694236
T_RectA*PROP	10.5343	1	10.53438	59.50183	1.46354537E-14	7.26011658	0.9411923
T_RectA*IBUT	98.2588	1	98.25882	554.9950	0	29.8573151	1.2673780
T_RectA*S_Q	65.2760	1	65.27601	368.7007	0	-8.81078625	0.4588572

(c) Imagen tabla ANOVA

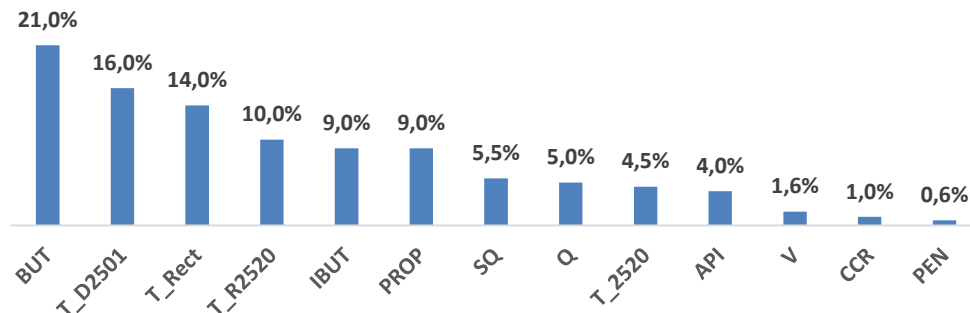
Figura 19. Regresión MLR para todos los datos de la clase 0 (continuación).

T_RectB*API	21.8553	1	21.85539	123.4468	2.37526072E-28	-5.17358446	0.4656414
T_RectB*V	11.2236	1	11.22360	63.39485	2.07527583E-15	-3.27048182	0.4107567
T_RectB*PROP	14.6961	1	14.69615	83.00867	1.15592649E-19	7.13444567	0.7830662
T_RectB*BUT	32.1076	1	32.10765	181.3548	1.2330866E-40	14.4399147	1.0722597
T_R2520B*PROP	42.7277	1	42.72772	241.3402	0	-7.82095575	0.5034368
T_R2520B*IBUT	34.8885	1	34.88859	197.0620	6.16571324E-44	-11.0861444	0.7897307
T_R2520B*BUT	43.7522	1	43.75222	247.1268	0	-16.4215641	1.0446108
API*V	16.8437	1	16.84378	95.13950	2.79346173E-22	1.38370979	0.1418615
API*Q	23.7933	1	23.79331	134.3934	1.0975066E-30	-2.8397665	0.2449591
API*PEN	21.6373	1	21.63737	122.2153	4.35336771E-28	0.983158231	0.0889324
API*PROP	22.0248	1	22.02480	124.4037	1.48354998E-28	-1.56449986	0.1402680
API*S_Q	20.0189	1	20.01895	113.0741	3.93237794E-26	-2.63404965	0.2477093
CCR*PEN	22.7218	1	22.72189	128.3408	2.14214026E-29	1.03978348	0.0917826
CCR*BUT	78.3735	1	78.37351	442.6816	0	-3.40620923	0.1618920
V*S_Q	20.1674	1	20.16746	113.9125	2.6005604E-26	-2.41455388	0.2262304
Q*PEN	32.5315	1	32.53153	183.7490	3.86015688E-41	-1.63279295	0.1204532
Q*IBUT	19.0755	1	19.07555	107.7449	5.45857579E-25	-4.26175404	0.4105726
Q*BUT	42.3641	1	42.36419	239.2869	0	-7.0148406	0.4534801
Q*S_Q	50.0361	1	50.03610	282.6198	0	-2.28884959	0.1361494
T_R2520A**2	8.07030	1	8.070303	45.58409	1.62768601E-11	-1.42286766	0.2107452
CCR**2	37.8734	1	37.87347	213.9227	0	0.96680522	0.0661013
V**2	23.2876	1	23.28767	131.5364	4.4593975E-30	0.345437318	0.0301193

(c) Imagen tabla ANOVA (continuación).



(d) Coeficientes MLR.



(e) Influencia de las variables en la predicción del %R.

Figura 20. Regresión MLR para todos los datos de la clase 0.

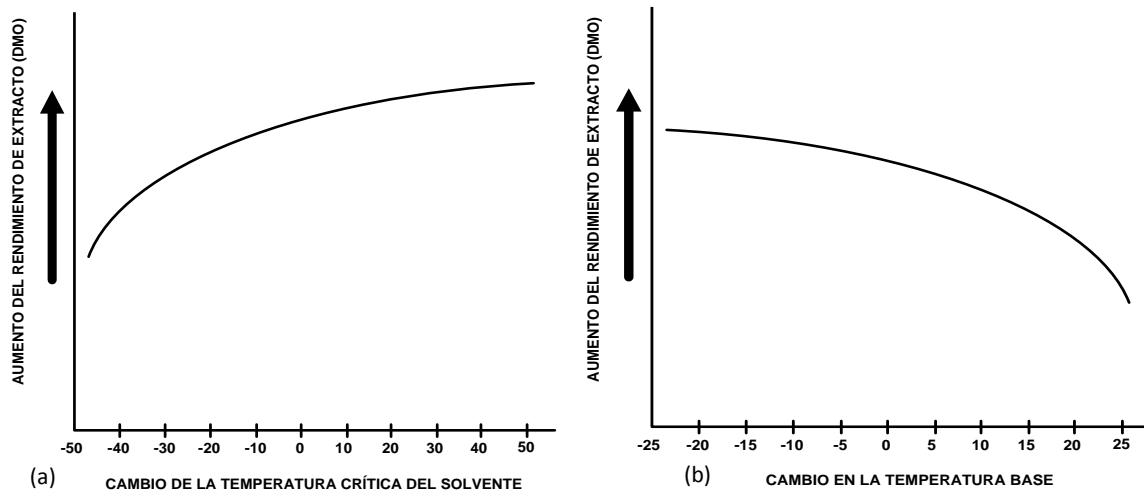


Figura 21. Tendencia del rendimiento de DMO con la temperatura crítica del solvente (a) y con la temperatura del proceso de extracción (b). Fuente: Manual de operación de la planta DEMEX. Capítulo 5.

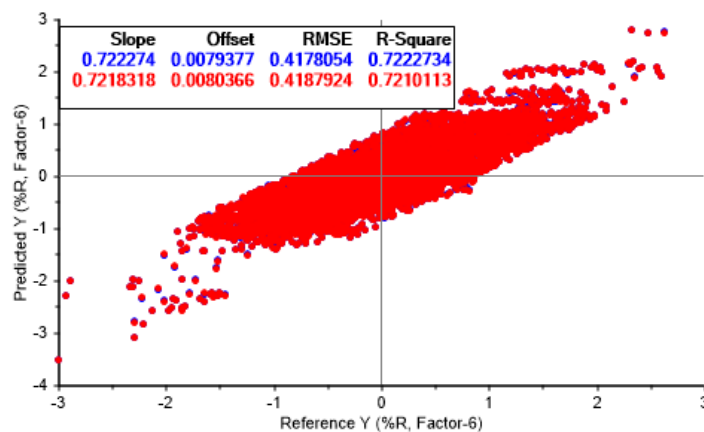
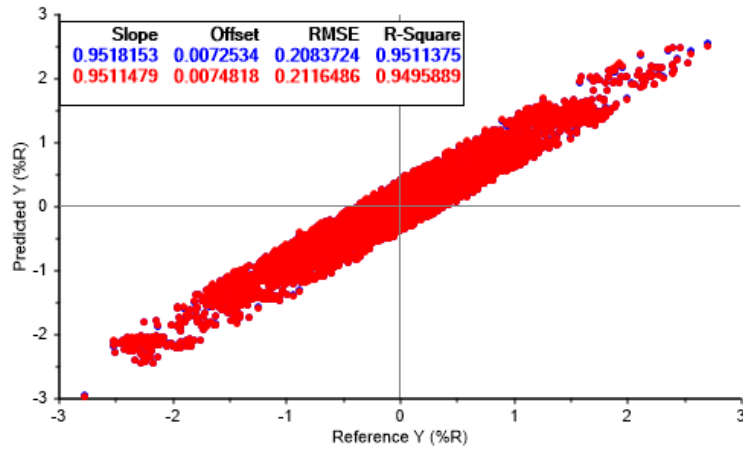
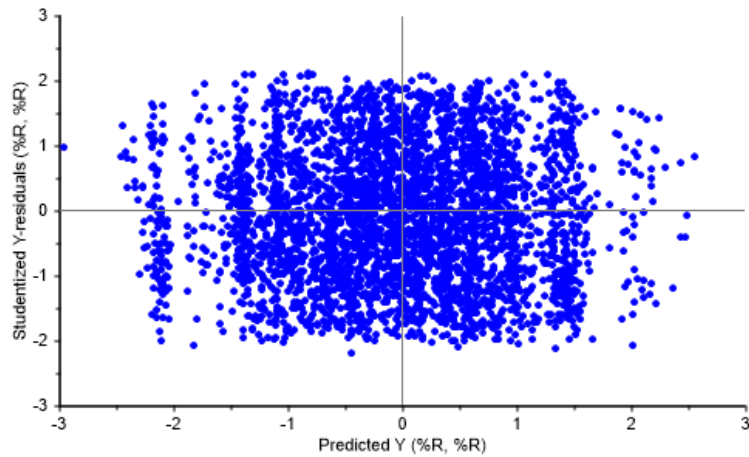


Figura 22. Valor predicho vs valor de referencia reportada por PLS para todos los datos de la clase 0.

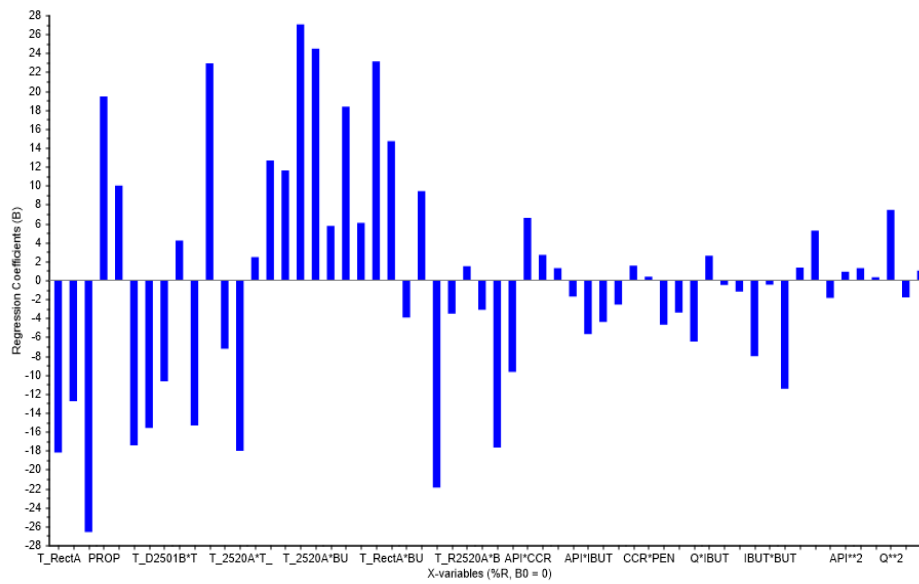


(a) Valor predicho vs valor de referencia.

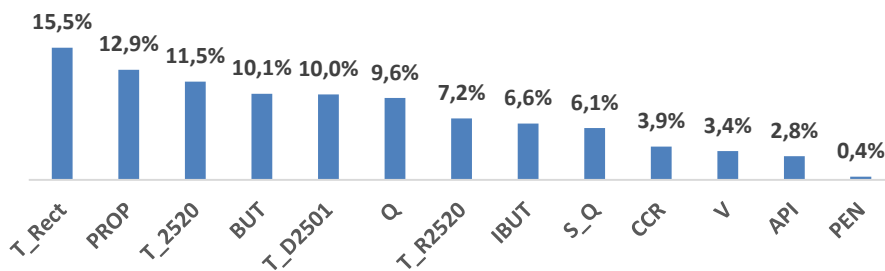
Figura 23. Regresión MLR para la subdivisión 1 de la clase 0.



(b) Residuos de la regresión.



(c) Coeficientes de la regresión MLR.



(d) Influencia variables independientes en la regresión MLR.

Figura 24. Regresión MLR para la subdivisión 1 de la clase 0 (continuación).

Anova Table							
Multiple Correlation: 0.975322 (cal) 0.9745319 (val)							
R-Square: 0.9511375 (cal) 0.9495889 (val)							
	SS	df	MS	F ratio	p value	B-coeficie	STDerr
Summary							
Model	3195.40674	58	55.0932198	1249.32495	0		
Error	163.472733	3707	0.04409838				
Total	3358.87939	3765	0.89213264				
Variables							
Intercept	0	0				0	0
T_RectA	41.7841721	1	41.7841721	947.512268	0	-18.1740131	0.5904164
CCR	12.8698921	1	12.8698921	291.846161	0	-12.7596836	0.7469007
Q	79.9723892	1	79.9723892	1813.49841	0	-26.5993862	0.6246159
PROP	24.5233898	1	24.5233898	556.103638	0	19.4274139	0.8238290
T_D2501A*T_Rect	12.2937269	1	12.2937269	278.778748	0	10.0189323	0.6000557
T_D2501A*PROP	25.3146553	1	25.3146553	574.046936	0	-17.4273167	0.7273724
T_D2501A*S_Q	8.70520782	1	8.70520782	197.404144	9.9492191E-44	-15.5949755	1.1099580
T_D2501B*T_R25	17.3405113	1	17.3405113	393.222534	0	-10.6673689	0.5379452
T_D2501B*CCR	1.54888523	1	1.54888523	35.123436	3.37696182E-09	4.22692013	0.7132233
T_D2501B*BUT	12.7230139	1	12.7230139	288.512848	0	-15.314785	0.9016299
T_D2501B*S_Q	17.4637699	1	17.4637699	396.017426	0	22.9342937	1.1524663
T_2520A*T_2520	3.6664319	1	3.6664319	83.1429672	1.22170105E-19	-7.22602367	0.7924768
T_2520A*T_RectB	20.3413639	1	20.3413639	461.272308	0	-18.0198669	0.8390210
T_2520A*T_R252	0.5232687	1	0.5232687	11.8659039	0.000578053237	2.47854304	0.7195252
T_2520A*V	29.9924545	1	29.9924545	680.125305	0	12.6838551	0.4863588
T_2520A*Q	19.0706596	1	19.0706596	432.459015	0	11.6227531	0.5589031
T_2520A*PROP	55.8553123	1	55.8553123	1266.60547	0	27.0646095	0.7604682
T_2520A*BUT	19.772789	1	19.772789	448.379791	0	24.4802914	1.1560951
T_2520B*CCR	2.20072412	1	2.20072412	49.9050446	1.91865035E-12	5.75897551	0.8152166
T_RectA*T_R2520	21.8196526	1	21.8196526	494.803925	0	18.3730984	0.8259729
T_RectA*PROP	1.65967178	1	1.65967178	37.635601	9.42306677E-10	6.07919073	0.9909377
T_RectA*IBUT	17.3082123	1	17.3082123	392.489288	0	23.1553516	1.1687927
T_RectA*BUT	7.73008966	1	7.73008966	175.287567	3.95307138E-39	14.7121124	1.1112185
T_RectB*V	5.08980227	1	5.08980227	115.419037	1.55983994E-26	-3.92955661	0.3657672

Anova Table							
T_RectB*Q	20.7579422	1	20.7579422	470.717529	0	9.42715645	0.4345107
T_RectB*IBUT	18.3386688	1	18.3386688	415.857483	0	-21.891592	1.0735075
T_R2520A*V	3.36727524	1	3.36727524	76.3580933	3.52299987E-18	-3.52182865	0.4030330
T_R2520A*BUT	2.89468431	1	2.89468431	65.6413879	7.2717355E-16	1.49364448	0.1843564
T_R2520B*CCR	2.65432024	1	2.65432024	60.1909294	1.10507493E-14	-3.11189222	0.4011058
T_R2520B*PROP	28.0114422	1	28.0114422	635.208862	0	-17.654335	0.7004757
T_R2520B*BUT	4.68911743	1	4.68911743	106.331604	1.33680221E-24	-9.65595245	0.9364054
API*CCR	38.7566032	1	38.7566032	878.865845	0	6.60280085	0.2227239
API*V	22.274025	1	22.274025	505.098969	0	2.70951438	0.1205599
API*PEN	19.6969681	1	19.6969681	446.65921	0	1.2943393	0.0612434
API*PROP	3.996351	1	3.996351	90.6236725	3.03304492E-21	-1.70795166	0.1794133
API*IBUT	16.4478016	1	16.4478016	372.980103	0	-5.68687105	0.2944628
API*BUT	9.77243614	1	9.77243614	221.605484	0	-4.37900019	0.2941607
API*S_Q	5.19562674	1	5.19562674	117.819031	4.82574565E-27	-2.56114507	0.2359535
CCR*Q	2.09561825	1	2.09561825	47.5214195	6.36754885E-12	1.5734843	0.2282536
CCR*PEN	1.20600963	1	1.20600963	27.3482761	1.79310646E-07	0.392885983	0.0751279
V*Q	18.3039646	1	18.3039646	415.071533	0	-4.6841836	0.2299177
V*S_Q	6.46573877	1	6.46573877	146.620468	3.97747101E-33	-3.405756	0.2812653
Q*PROP	21.5308456	1	21.5308456	488.245819	0	-6.45304823	0.2920422
Q*IBUT	1.76948595	1	1.76948595	40.1259232	2.6662908E-10	2.61279917	0.4124710
PEN*PROP	4.67790985	1	4.67790985	106.078903	1.51323145E-24	-0.47519168	0.0461375
PEN*BUT	14.3080893	1	14.3080893	324.45993	0	-1.18178928	0.0656084
PROP*S_Q	28.696146	1	28.696146	650.729431	0	-8.00941086	0.3139788
IBUT*BUT	7.1563673	1	7.1563673	162.282135	2.05999412E-36	-0.46029558	0.0361327

(e) Tabla ANOVA.

Figura 25. Regresión MLR para la subdivisión 1 de la clase 0 (continuación).

BUT*S_Q	28.7179623	1	28.7179623	651.225525	0	-11.4338484	0.4480502
T_2520B**2	0.28101053	1	0.28101053	6.37240267	0.0116322711	1.35695744	0.5375447
T_RectA**2	7.45437956	1	7.45437956	169.039474	7.96010171E-38	5.27605534	0.4058030
T_R2520B**2	0.74923557	1	0.74923557	16.9904213	3.83910956E-05	-1.8600173	0.4512476
API**2	2.8674984	1	2.8674984	65.0250854	9.88789099E-16	0.926759005	0.1149281
CCR**2	5.05913687	1	5.05913687	114.723938	2.19141738E-26	1.28716588	0.1201731
V**2	12.0468521	1	12.0468521	273.182526	0	0.348269522	0.0210711
Q**2	18.3225574	1	18.3225574	415.491791	0	7.4460001	0.3652933
PROP**2	36.5571136	1	36.5571136	828.98999	0	-1.78915632	0.0621403
IBUT**2	3.6423223	1	3.6423223	82.5955811	1.60176921E-19	1.02863407	0.1131833

(e) Tabla ANOVA (continuación).

Figura 26. Regresión MLR para la subdivisión 1 de la clase 0 (continuación).

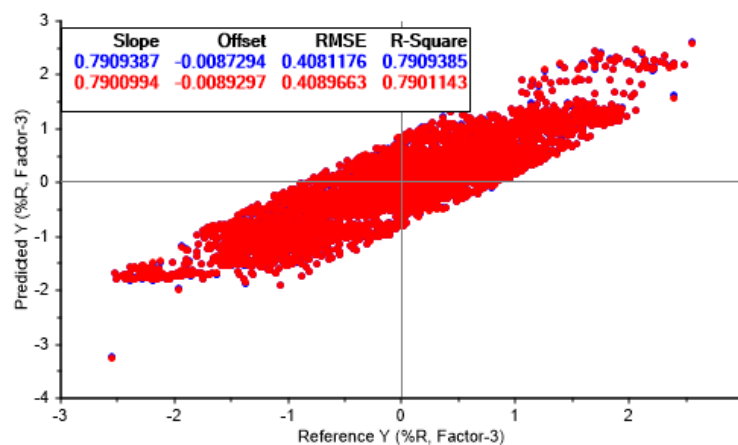
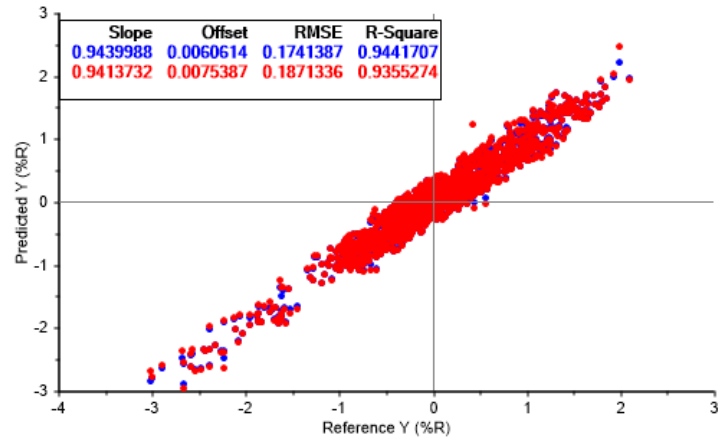


Figura 27. Regresión PLS para la subdivisión 1 de la clase 0.

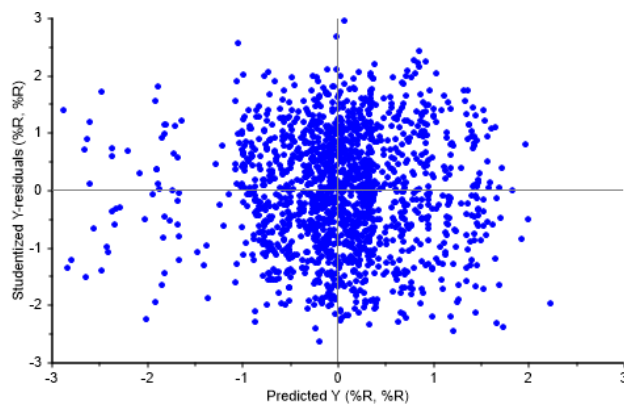
$$\begin{aligned} \%R = & -0.08517 - 0.1915 * T_{2520A} - 0.2441 * T_{2520B} + 0.1724 * T_{RectA} - 0.0835 * T_{RectB} - \\ & 0.1752 * T_{R2520A} + 0.0453 * T_{R2520B} - 0.0721 * API - 0.4109 * CCR + 0.15 * Q - 0.3071 * \\ & PEN - 0.251 * PROP - 0.1882 * iBUT + 0.2037 * BUT \end{aligned} \quad (1)$$

En lo referente a la subdivisión 2, conformada por 2078 muestras, la Figura 3.18 presenta un desempeño alto de la regresión MLR con una reproducción de la varianza de %R de 94.4% (93.6% en validación) y un RMSE de 0.174 (0.187 en validación); debido a los valores de sus estadísticos (Tabla 3.2), la regresión MLR obtenida puede ser aplicada para el análisis cuantitativo del proceso

DEMEX; sin embargo, el uso de esta regresión puede estar restringida por la cantidad de términos de entrada requeridos (88 entradas). La Figura 3.18b soporta el cumplimiento de los supuestos de distribución normal y homocedasticidad de la distribución de los residuales para la regresión MLR. Los valores de los coeficientes para la regresión MLR son mostrados en la Tabla ANOVA de la Figura 3.18d; según esta figura, tanto la regresión como sus entradas presentan significancia superior al 95% de confianza; asimismo, la Figura 3.18c presenta una comparación entre los coeficientes para las diferentes entradas de esta regresión, mientras que la Figura 3.17e denota la influencia de cada variable en la reproducción del valor del rendimiento de DMO (%R). Según la Figura 3.17e, la composición del solvente influye en *ca.* 53% el valor predicho para %R; la principal influencia del solvente en esta subdivisión se encuentra en el contenido de IBUT (25.1%), mientras que el contenido de BUT influye en *ca.* 18.7 en el valor de %R. También, respecto a las condiciones de operación, la temperatura de rectificación (T_Rect) exhibe una influencia de 9.4% en el valor predicho de %R, mientras que el flujo de alimentación presenta una influencia de 10.4%. Por otra parte, la ecuación (2) muestra la regresión PLS (Figura 3.18) obtenida para la subdivisión 2; la regresión PLS exhibe un desempeño menor que la correspondiente MLR, ($R^2=0.907$ en calibración y $R^2=0.906$ en validación) (Tabla 3.2); sin embargo, la regresión PLS puede ser utilizada para análisis semicuantitativos del desempeño de la unidad DEMEX, ya que requiere sólo de 7 entradas. Según los coeficientes de (2), la composición combinada del PROP y el BUT en el solvente influye en *ca.* 20% en el rendimiento obtenido en la unidad, mientras que la carga y sus propiedades influyen *ca.* 48% en el valor del rendimiento obtenido; las propiedades de la carga que afectan el rendimiento corresponden a CCR, 10%, PEN, 14%, Q, 17% y S_Q, 7%. Probablemente, el mejor desempeño de la regresión MLR se deba al término IBUT en el modelo.

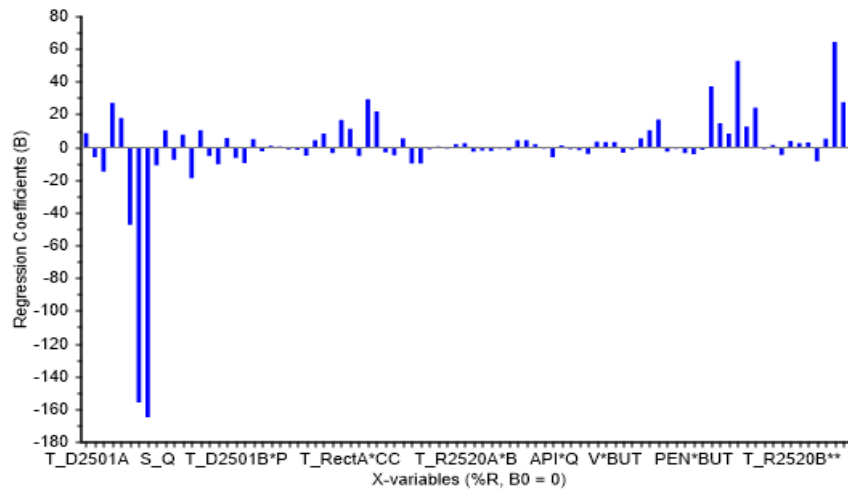


(a) Valor predicho vs valor de referencia.



(b) Residuales regresión MLR.

Figura 28. Regresión MLR para la subdivisión 2 de la clase 0.



(c) Coeficientes de la regresión MLR.

Anova Table							
Multiple Correlation: 0.971708 (cal) 0.9672839 (val)							
R-Square: 0.9441707 (cal) 0.9355274 (val)							
	SS	df	MS	F ratio	p value	B-coeficie	STDerr
Summary							
Model	801.118225	88	9.10361671	283.284	0		
Error	47.336235	1473	0.03213593				
Total	848.454468	1561	0.54353267				
Variables							
Intercept	0	0				0	0
T_D2501A	1.58042836	1	1.58042836	49.1772	3.55130725E-12	8.45720673	1.2059931
API	0.26927554	1	0.26927554	8.37937	0.00385094318	-6.14219427	2.1218650
CCR	3.19091105	1	3.19091105	99.2950	1.11508985E-22	-14.9890938	1.5042210
Q	2.91771245	1	2.91771245	90.7982	6.28605567E-21	26.9881725	2.8322701
PEN	1.2069006	1	1.2069006	37.5561	1.13799903E-09	17.7452297	2.8956167
PROP	0.53376948	1	0.53376948	16.6083	4.83980111E-05	-47.4535255	11.644070
IBUT	1.86045635	1	1.86045635	57.8893	4.91986517E-14	-155.956818	20.497692
BUT	1.28552687	1	1.28552687	39.9988	3.36214917E-10	-165.033707	26.094493
S_Q	0.75287425	1	0.75287425	23.4280	1.43300474E-06	-10.9924335	2.2710430
T_D2501A*T_252	0.17093528	1	0.17093528	5.31902	0.0212320369	10.4165688	4.5165681
T_D2501A*T_Rect	1.32384598	1	1.32384598	41.1945	1.85354745E-10	-7.67776108	1.1962305
T_D2501A*CCR	1.47913015	1	1.47913015	46.0255	1.68549792E-11	7.66223383	1.129421
T_D2501A*Q	2.35108328	1	2.35108328	73.1592	2.95200478E-17	-18.8690796	2.2060544
T_D2501A*PEN	0.40972888	1	0.40972888	12.7497	0.000367496512	10.2634077	2.8743541
T_D2501A*BUT	0.16513280	1	0.16513280	5.13854	0.0235446598	-5.43921471	2.3994736
T_D2501B*T_252	0.16867084	1	0.16867084	5.24856	0.0221055727	-10.3566599	4.5206360
T_D2501B*Q	0.43637034	1	0.43637034	13.5784	0.000237081345	5.73055077	1.5551445
T_D2501B*PEN	0.17774763	1	0.17774763	5.53095	0.0188140012	-6.4483037	2.7418613
T_D2501B*PROP	0.79256916	1	0.79256916	24.6635	7.62210277E-07	-9.70627213	1.9544497
T_2520A*T_RectB	5.69747829	1	5.69747829	177.292	2.80217626E-38	5.02209759	0.3771725
T_2520A*T_R252	12.1757374	1	12.1757374	378.889	0	-2.48348451	0.1275866
T_2520A*V	0.80062472	1	0.80062472	24.9144	6.70617396E-07	0.985732198	0.1974844
T_2520A*PEN	0.51422244	1	0.51422244	16.0006	6.64631734E-05	0.473066807	0.1182643
T_2520A*PROP	5.18767452	1	5.18767452	161.435	3.57885311E-35	-1.43164873	0.1126774
T_2520A*BUT	0.32321047	1	0.32321047	10.0573	0.00154866045	-1.51892447	0.4789538
T_2520B*T_RectB	2.31141233	1	2.31141233	71.9262	5.35425606E-17	-5.23570204	0.6173495

(d) Tabla ANOVA.

Figura 29. Regresión MLR para la subdivisión 2 de la clase 0 (continuación).

	Anova Table						
T_2520B*Q	1.91610563	1	1.91610563	59.6260	2.1054244E-14	4.27548265	0.5536907
T_RectA*T_RectB	2.15424323	1	2.15424323	67.0341	5.71941157E-16	8.3996973	1.0259248
T_RectA*T_R2520	0.93877297	1	0.93877297	29.2116	7.55863638E-08	-3.62914109	0.6714684
T_RectA*API	2.64875865	1	2.64875865	82.4223	3.43223344E-19	16.758297	1.8458966
T_RectA*CCR	1.00058055	1	1.00058055	31.1340	2.86167587E-08	11.3216915	2.0290544
T_RectA*V	0.43910574	1	0.43910574	13.6643	0.000226580174	-5.50446272	1.4890893
T_RectA*IBUT	1.35003662	1	1.35003662	42.0110	1.23484903E-10	29.3159294	4.5229535
T_RectA*BUT	0.76356142	1	0.76356142	23.7608	1.20872971E-06	21.7699509	4.4660782
T_RectB*API	0.21342967	1	0.21342967	6.64137	0.0100602498	-3.0761857	1.1936682
T_RectB*CCR	0.33509647	1	0.33509647	10.4275	0.00126887183	-4.86834049	1.5076142
T_RectB*V	0.54422134	1	0.54422134	16.9353	4.0816336E-05	5.46921301	1.3290076
T_RectB*Q	3.66607189	1	3.66607189	114.081	1.06093681E-25	-9.91141605	0.9279567
T_RectB*PEN	3.79512143	1	3.79512143	118.091	1.62837931E-26	-10.0522614	0.9250272
T_R2520A*T_R25	0.29033067	1	0.29033067	9.03468	0.00269383797	-1.05383682	0.3506039
T_R2520A*CCR	0.14719580	1	0.14719580	4.58033	0.032504715	0.413424522	0.1931734
T_R2520A*V	0.71441268	1	0.71441268	22.2301	2.64724804E-06	-0.96149432	0.2039272
T_R2520A*Q	1.63398421	1	1.63398421	50.8458	1.5603224E-12	1.88864446	0.2648637
T_R2520A*BUT	0.49215951	1	0.49215951	15.3152	9.51305992E-05	2.48747849	0.6356205
T_R2520A*S_Q	2.23666763	1	2.23666763	69.5989	1.65020613E-16	-2.68282056	0.3215809
T_R2520B*API	0.83831685	1	0.83831685	26.0871	3.68996723E-07	-2.03404188	0.3982418
T_R2520B*CCR	2.88171768	1	2.88171768	89.6751	1.0731817E-20	-2.33897781	0.2469960
T_R2520B*Q	0.14027331	1	0.14027331	4.36500	0.036855597	-0.97465497	0.4665073
T_R2520B*PEN	0.90415579	1	0.90415579	28.1346	1.30427949E-07	-1.89980805	0.3581700
T_R2520B*BUT	0.76645469	1	0.76645469	23.8501	1.1547753E-06	4.35697603	0.8921529
T_R2520B*S_Q	3.98696208	1	3.98696208	124.065	1.00749979E-27	4.26238441	0.3826718
API*CCR	2.91731477	1	2.91731477	90.7802	6.34023681E-21	1.70420873	0.1788657
API*V	0.42671936	1	0.42671936	13.2780	0.000277847721	-0.97238385	0.2668516
API*Q	3.80599976	1	3.80599976	118.436	1.38613323E-26	-6.17638588	0.5675334
API*PEN	0.94480514	1	0.94480514	29.3985	6.87669939E-08	1.0475179	0.1931961
API*S_Q	0.15235826	1	0.15235826	4.74111	0.0296082273	-1.11604786	0.5125574
CCR*S_Q	1.22611094	1	1.22611094	38.1520	8.4490942E-10	-1.71628523	0.2778628
V*Q	1.10267413	1	1.10267413	34.3109	5.78416159E-09	-3.9854331	0.6803919
V*PROP	2.74688959	1	2.74688959	85.4788	7.94984804E-20	3.27702641	0.3544464
V*IBUT	0.38402000	1	0.38402000	11.9493	0.000562212022	3.22595859	0.9332255
V*BUT	0.25982257	1	0.25982257	8.08496	0.00452500768	3.24973488	1.1429015

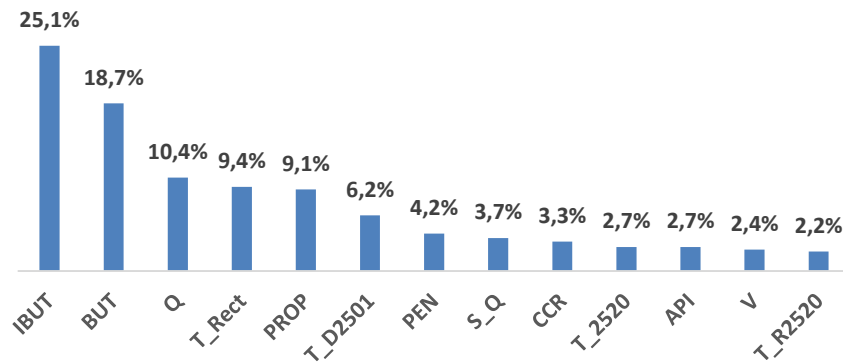
	Anova Table						
V*S_Q	1.29407525	1	1.29407525	40.2667	2.94193336E-10	-3.29453802	0.5191836
Q*PEN	0.28192886	1	0.28192886	8.77297	0.00310625136	-1.33124149	0.4494521
Q*PROP	0.32702261	1	0.32702261	10.1765	0.00145230535	5.44821548	1.7078626
Q*IBUT	0.2447851	1	0.2447851	7.61734	0.00585260103	10.3652439	3.7555871
Q*BUT	0.53619176	1	0.53619176	16.6853	4.64952864E-05	16.9631767	4.1527891
Q*S_Q	1.09705973	1	1.09705973	34.1375	6.31030161E-09	-2.65553856	0.4545024
PEN*PROP	0.15525506	1	0.15525506	4.83139	0.0281012077	-0.87905448	0.3999260
PEN*IBUT	0.40679946	1	0.40679946	12.6591	0.00038558268	-3.57997036	1.0061831
PEN*BUT	0.30621430	1	0.30621430	9.52909	0.00206012139	-4.36648941	1.4145119
PEN*S_Q	0.54870867	1	0.54870867	17.0743	3.79676785E-05	-1.53307295	0.3710145
PROP*IBUT	0.82598090	1	0.82598090	25.7010	4.49125025E-07	37.3006859	7.3576817
PROP*BUT	0.25637236	1	0.25637236	7.97699	0.00480136042	14.6253309	5.1782846
PROP*S_Q	2.32190871	1	2.32190871	72.2555	4.56676689E-17	8.35809803	0.9832677
IBUT*BUT	0.83157473	1	0.83157473	25.8745	4.11168003E-07	52.7399864	10.368213
IBUT*S_Q	1.2367053	1	1.2367053	38.4845	7.15642212E-10	12.651598	2.0393991
BUT*S_Q	2.44682264	1	2.44682264	76.1368	7.02579171E-18	24.1996136	2.7733907
T_2520A**2	2.29174352	1	2.29174352	71.3110	7.20809421E-17	-1.16702032	0.1381973
T_2520B**2	2.31285667	1	2.31285667	71.9696	5.24317398E-17	1.3580004	0.1600756

(d) Tabla ANOVA (continuación).

Figura 30. Regresión MLR para la subdivisión 2 de la clase 0 (continuación).

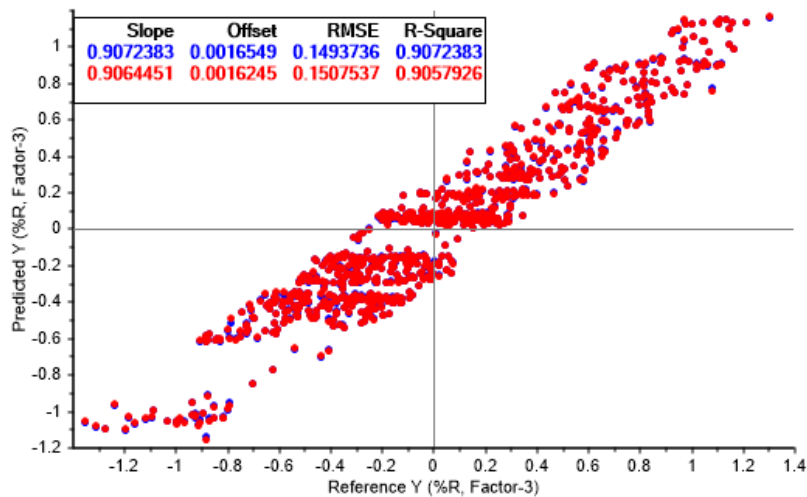
T_RectA**2	0.42694068	1	0.42694068	13.2859	0.000276689098	-4.73398399	1.2987623
T_R2520B**2	2.43981934	1	2.43981934	75.9210	7.79507878E-18	3.83094621	0.4396680
CCR**2	1.69966722	1	1.69966722	52.8905	5.70622054E-13	2.52847242	0.3476716
V**2	1.8851608	1	1.8851608	58.6626	3.37086451E-14	2.91851354	0.3810489
Q**2	4.86422682	1	4.86422682	151.367	3.4887765E-33	-8.49530602	0.6904991
PROP**2	0.20463034	1	0.20463034	6.36709	0.0117300935	5.18547916	2.0550308
IBUT**2	1.17155445	1	1.17155445	36.4531	1.97601602E-09	64.3739014	10.662089
BUT**2	0.40065577	1	0.40065577	12.4662	0.000427133869	27.4554329	7.7760705
S_Q**2	0.77828913	1	0.77828913	24.2185	9.56662689E-07	3.01383495	0.6124149

(d) Tabla ANOVA (continuación).



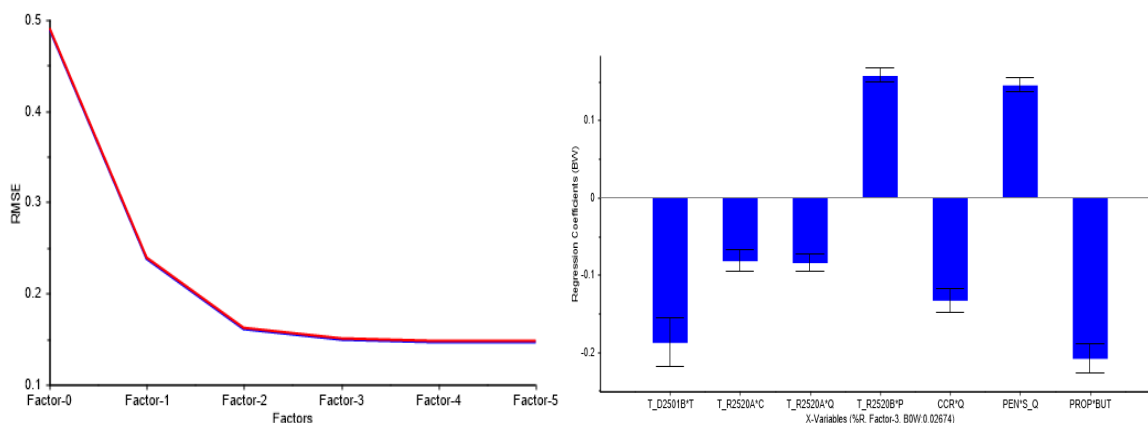
(e) Influencia variables independientes en la regresión MLR.

Figura 31. Regresión MLR para la subdivisión 2 de la clase 0 (continuación).



(a) Valor predicho vs valor de referencia

Figura 32. Regresión PLS para la subdivisión 2 de la clase 0.



(b) Disminución RMSE según factores.

(c) Coeficientes de la regresión PLS.

Figura 33. Regresión PLS para la subdivisión 2 de la clase 0 (continuación).

$$\begin{aligned} \%R = & 0.02674 - 0.1881 * T_{D2501B} * T_{RectB} - 0.0821 * T_{R2520A} * CCR - 0.0845 * T_{R2520A} * Q \\ & + 0.1575 * T_{R2520B} * PEN - 0.1334 * CCR * Q + 0.1452 * PEN * S_Q - 0.2088 \\ & * PROP * BUT \end{aligned}$$

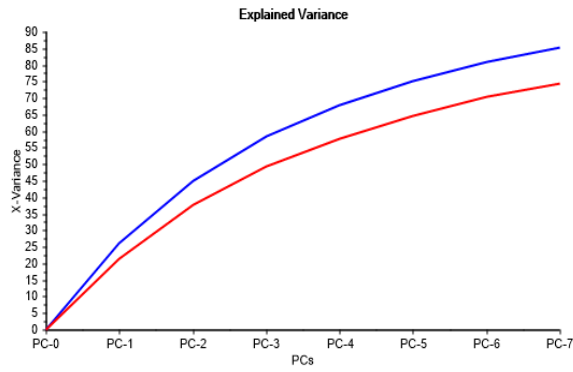
Tabla 3.

Resumen modelos ajustados para la clase 0.

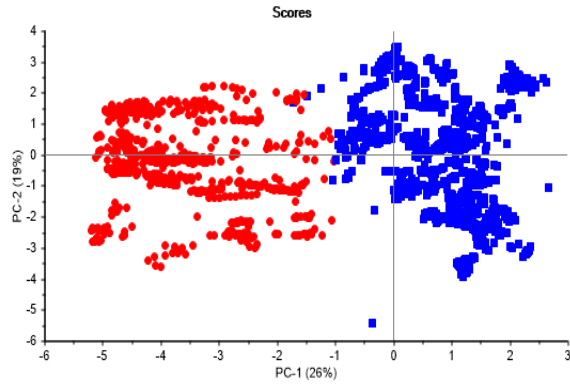
	MODELO	NÚMERO DE ENTRADAS	R2	RMSE
Todos los datos	MLR	48	0.792 (0.823)	0.419 (0.574)
	PLS	6 Factores resultantes de la combinación lineal de 25 entradas.	0.722 (0.721)	0.418 (0.420)
Subdivisión1	MLR	58	0.951 (0.950)	0.208 (0.212)
	PLS	3 Factores resultantes de la combinación lineal de 13 entradas.	0.791 (0.790)	0.408 (0.409)
Subdivisión2	MLR	88	0.944 (0.936)	0.174 (0.187)

MODELO	NÚMERO DE ENTRADAS	R2	RMSE
PLS	3 Factores resultantes de la combinación lineal de 7 entradas.	0.907 (0.906)	0.149 (0.151)

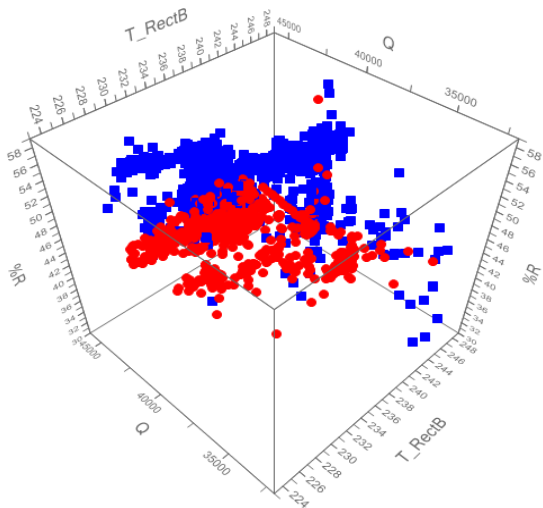
4.3.2 Clase 1 Para la clase1 conformada por 3273 muestras, la aplicación de un nuevo PCA reporta una varianza explicada acumulada de 85.1% en calibración (74.5% validación) con sólo 7 PC (Figura 3.20a). La aplicación del método *k-means* con dos (Figura 3.20b) y tres agrupaciones dentro de la clase1 reportó grupos definidos en las gráficas de las variables de proceso Q vs T_rectB vs %R (Figuras 3.20c-d); el método *k-means* con k=3 fue considerado debido a que las subdivisiones delimitan en un valor de PC-1 diferente de cero (Figura 3.20b); sin embargo, del análisis de las variables originales con las subdivisiones para k=3 se tiene que presentan una relación no constante para los respectivos límites de la clase1 (Figura 3.20d). Con lo anterior, las regresiones de predicción para las tres subdivisiones involucrarían dos ecuaciones para la descripción de los límites; en busca de simplicidad, las subdivisiones con k=3 son descartadas y se considera k=2 como las subdivisiones representativas de la clase 1 (Figuras 3.20b-c). De acuerdo con k=2, la subdivisión1 está formada por 2543 muestras, mientras que la subdivisión2 contiene 730 muestras. Según los resultados obtenidos, la subdivisión de la clase 1 en dos agrupaciones se presenta debido al caudal de crudo a tratar; el límite entre las dos agrupaciones se establece en un valor de caudal de *ca.* 40 kpbd (Figuras 3.20c-e). De otro lado, los resultados del PCA muestran puntos atípicos en las agrupaciones de la clase 1 que influyen en el escenario operativo; estos puntos atípicos –referidos a posibles condiciones operativas de mantenimiento y/o estados dinámicos de la planta– son ilustrados en las Figuras 3.20e-f y se identifican por sus respectivos valores del estadístico T².



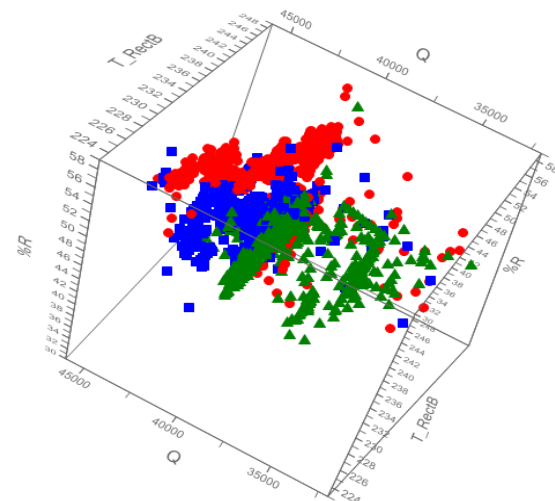
(a) Varianza explicada



(b) Ilustración k=2 en los scores para PC1 y PC2.

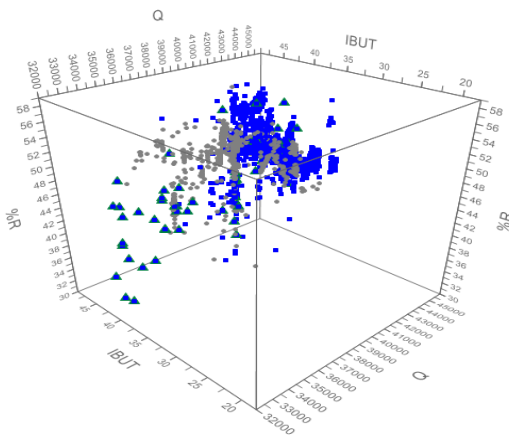


(c) Ilustración k=2 en las variables Q vs T_RectB vs %R.

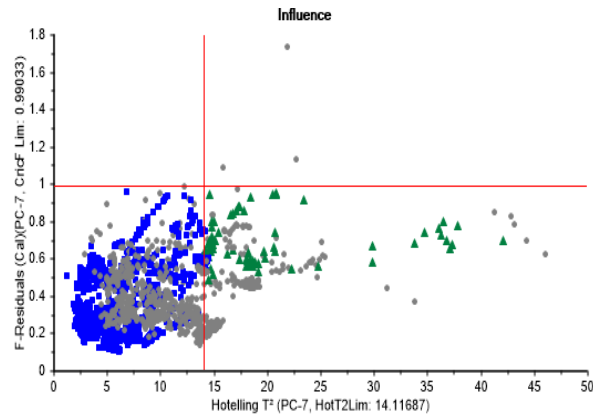


(d) Ilustración k=3 en las variables Q vs T_RectB vs %R.

Figura 34. Resultados PCA para la clase 1.



(e) Ilustración puntos atípicos (triángulos) con $k=2$ en las variables Q vs T_RectB vs %R.

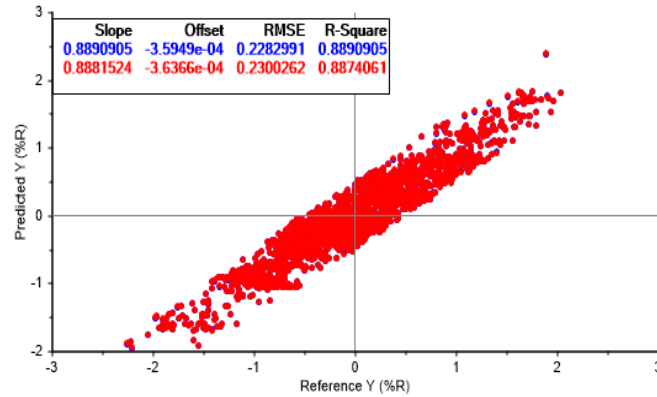


(f) Gráfica de influencia ilustrando $k=2$ agrupaciones y puntos atípicos (triángulos verdes).

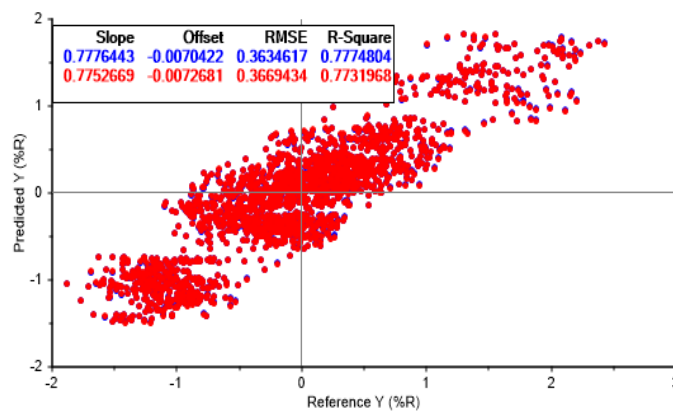
Figura 35. Resultados PCA para la clase 1 (continuación). (a) varianza explicada, 85.1% calibración, 74.5% validación. (b) Gráfica de los scores denotando dos ($k=2$) agrupaciones. (c) Gráfica de los datos históricos considerando dos agrupaciones. (d) Gráfica de los datos históricos considerando tres agrupaciones. (e) Gráfica de los datos históricos considerando dos agrupaciones. (f) Gráfica de influencia considerando dos agrupaciones.

Respecto a las regresiones, la Figura 3.21 presenta las gráficas de rendimiento predicho versus el valor para todos los datos (a), los datos del grupo 1 (b) y los datos del grupo 2 (c) de la clase 1 utilizando MLR. Según esta figura, la subdivisión de la clase 1 desmejora el desempeño de las regresiones; para todos los datos de esta clase, el coeficiente de determinación (R^2) reporta el valor de 0.889 (0.887 en validación), mientras que para los datos de las subdivisiones 1 y 2 de esta clase, R^2 desciende a los valores de 0.777 y 0.873, respectivamente. (Tabla 3.3) Con estos valores para el estadístico R^2 , la regresión para la clase presenta aplicación a nivel semicuantitativo, mientras que las regresiones para las subdivisiones presentan utilidad sólo a nivel cualitativo. Adicionalmente, la regresión con todos los datos requiere una menor cantidad de entradas (Figura 3.22a) con significancia estadística (95%) y cumple los supuestos de normalidad (Figura 3.22b) y homoscedasticidad (Figura 3.22c).

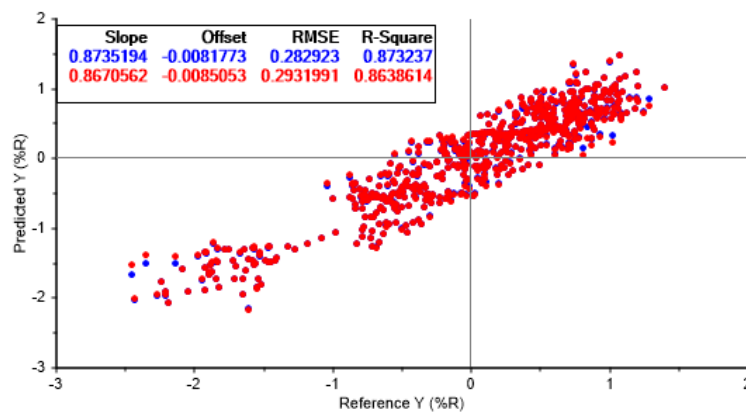
Según los coeficientes de la regresión para todos los datos de la clase 1 (Figuras 3.22a, d), la composición del solvente influye en *ca.* 21% sobre la predicción del rendimiento de DMO (%R) –10.7% BUT, 2.1% IBUT, 8.5% PROP–, mientras que la influencia de las propiedades de la carga se aproxima a 24.7% –24% API y 0.2% V– y en lo referente a las condiciones operativas la influencia llega a *ca.* 49.1% –T_D2501 23.4%, T_2520 4.4%, T_Rect 20%, T_R 1.9%–. Asimismo, las variables que más influyen son API, T_D2501, T_Rect, BUT y PROP, con una incidencia de 83.7% en conjunto sobre el valor de %R. También, la influencia del residuo de carbón (CCR) en el modo operativo referente a la clase 1 resulta nulo, con lo cual, el CCR del DMO resultante se mantiene entre unos límites cercanos.



(a) Todos los datos.



(b) Subdivisión1

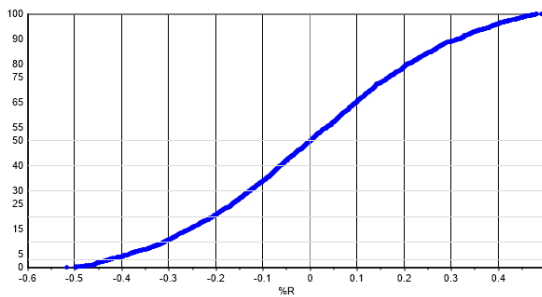


(c) Subdivisión2

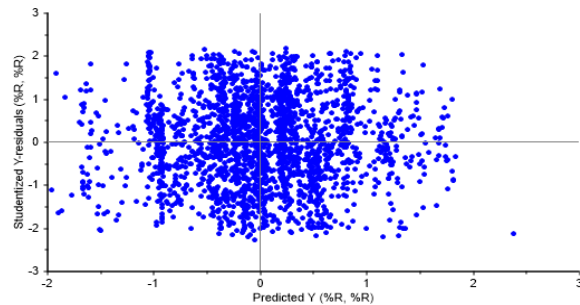
Figura 36. Resultados regresiones MLR con los datos de la clase 1. (a) Todos las 3243 muestras. (b) Subdivisión1, 2573 muestras. (c) Subdivisión2, 730 muestras. Datos en azul y en rojo corresponden a las etapas de calibración y validación, respectivamente.

Anova Table					
Multiple Correlation: 0.9429159 (cal) 0.9420227 (val)					
R-Square: 0.8890905 (cal) 0.8874061 (val)					
	MS	F ratio	p value	B-coefficients	STDerr
Summary					
Model	59.0427055	1124.29443	0		
Error	0.0525153391				
Adjusted Total	0.470145017				
Variables					
Intercept	3.2385695	61.669014	6.2434087E-15	-0.0437336601	0.0055690
T_D2501A	0.552461684	10.5200405	0.00119819853	0.29169929	0.0899345
T_D2501B	2.79691172	53.258812	4.03735042E-13	-0.650737166	0.0891681
T_RectB	1.9035964	36.2484093	2.02367989E-09	-0.072051473	0.0119673
API	28.6687965	545.914124	0	0.245119333	0.0104909
V	8.79629898	167.499817	5.3006417E-37	-0.109503008	0.0084609
PROP	16.0982056	306.542999	0	-0.171965107	0.0098218
T_D2501B*Q	22.938839	436.802795	0	0.176095262	0.0084256
T_2520A*PROP	1.741992	33.1711006	9.59812585E-09	-0.0883391425	0.0153381
T_2520A*IBUT	0.645202816	12.2859936	0.000465319463	0.0534102544	0.0152377
T_2520B*IBUT	1.58592606	30.1992588	4.34016023E-08	0.068661429	0.0124943
T_RectA*API	82.0670776	1562.72998	0	0.472028464	0.0119406
T_RectB*API	42.9051781	817.001648	0	-0.236065477	0.0082588
T_RectB*BUT	19.9497852	379.88501	0	0.173520058	0.0089027
T_R2520A*Q	1.90589797	36.2922478	1.97937466E-09	-0.0620839372	0.0103055
API*BUT	30.3318539	577.581299	0	-0.214914918	0.0089425
PROP*BUT	30.8364067	587.188843	0	0.242759719	0.0100181

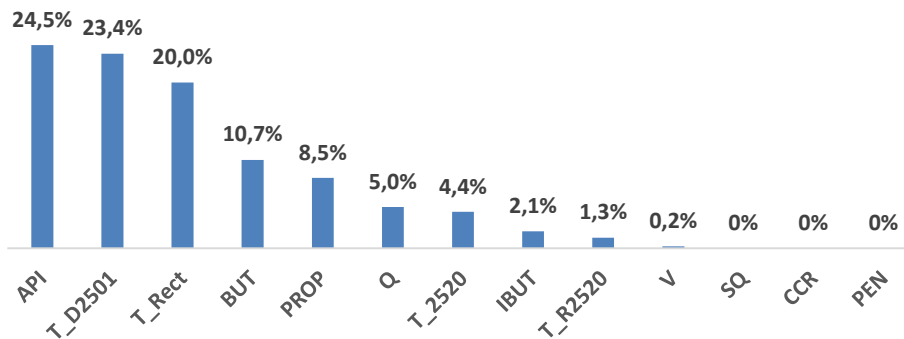
(a) Tabla ANOVA.



(b) Distribución de los residuos.



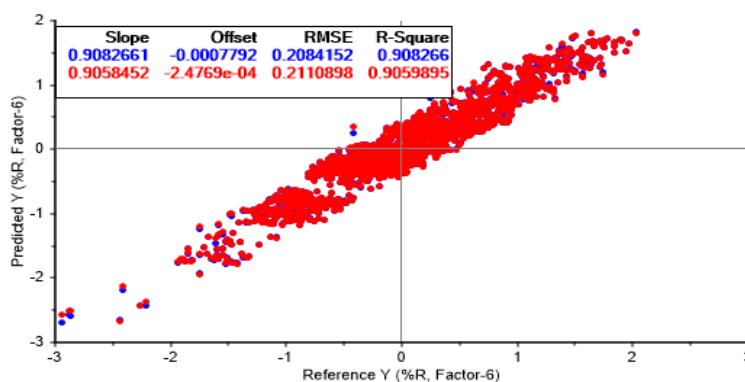
(c) Residuales estudentizados.



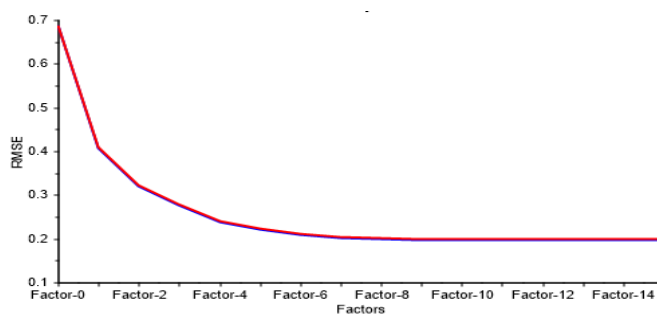
d) Influencia de las variables independientes.

Figura 37. Resultados regresión MLR con los datos de la clase 1.

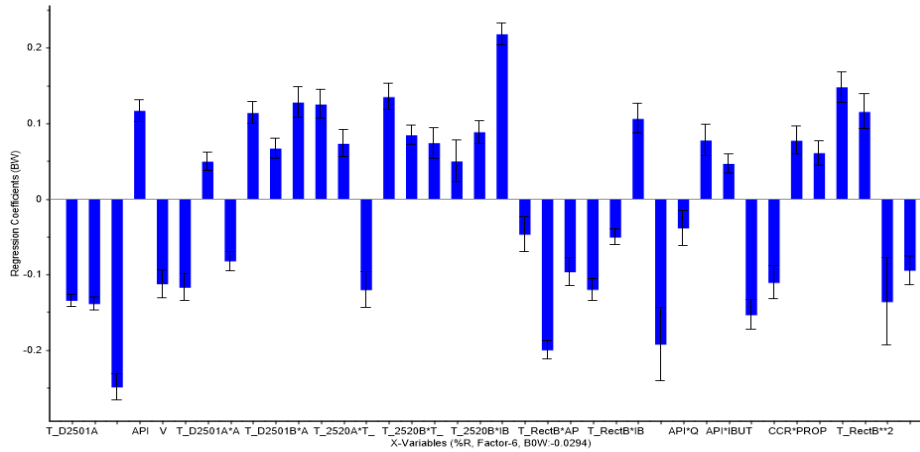
Por otro lado, las regresiones PLS aplicadas a la clase 1 y sus subdivisiones presentan un desempeño superior (Figuras 3.12-14, Tabla 3.3) que el reportado por las regresiones MLR (Figura 3.21); sin embargo, el desempeño mostrado por la regresión PLS para la subdivisión 2 reporta mejoras respecto a la regresión PLS para todos los datos; lo anterior puede deberse a la menor cantidad de datos en la subdivisión 2 respecto al total. Las ecuaciones (3-5) presentan las regresiones PLS para la clase 1 y sus subdivisiones, respectivamente.



(a) Valores predichos vs valores de referencia.

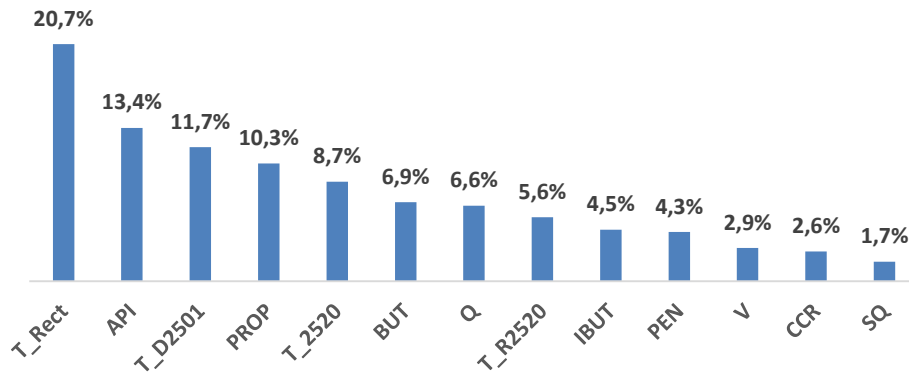


(b) Variación del RMSE según el número de factores.



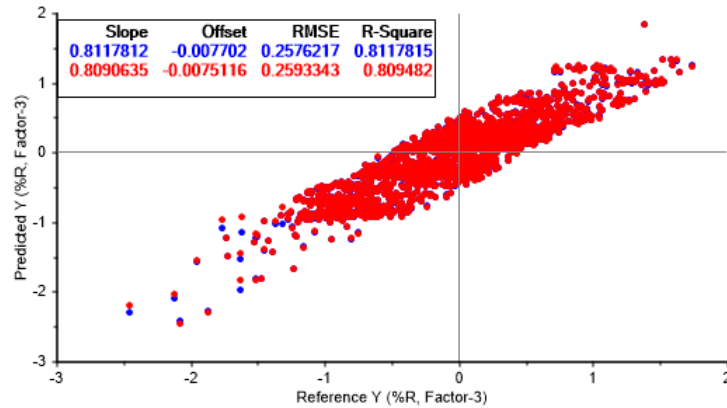
(c) Coeficientes obtenidos por la regresión.

Figura 38. Resultados regresión PLS para la clase 1.

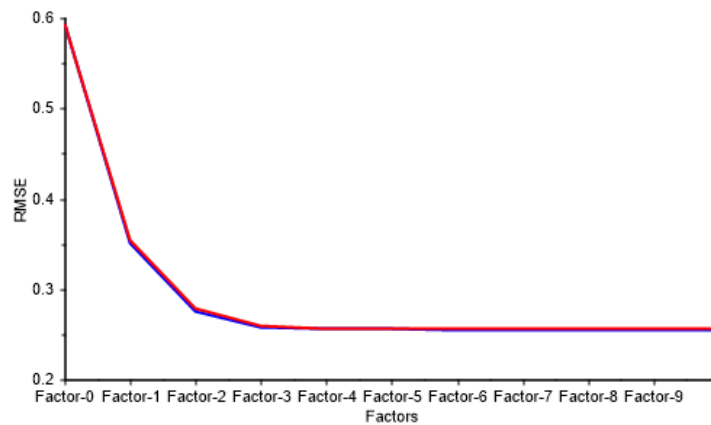


(d) Influencia variables independientes.

Figura 39. Resultados regresión PLS para la clase 1 (continuación).

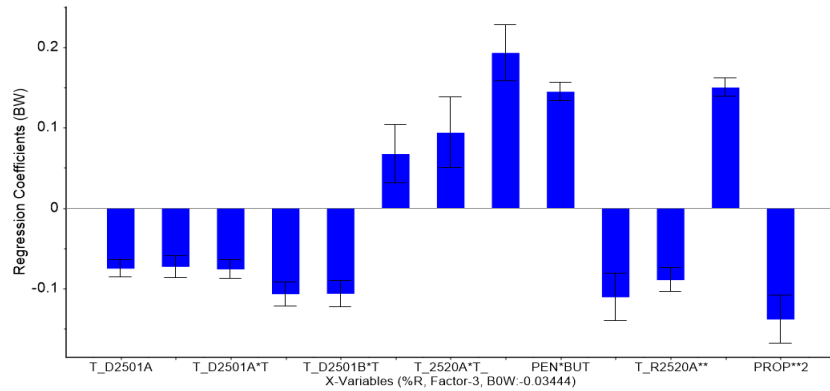


(a) Valores predichos vs valores de referencia

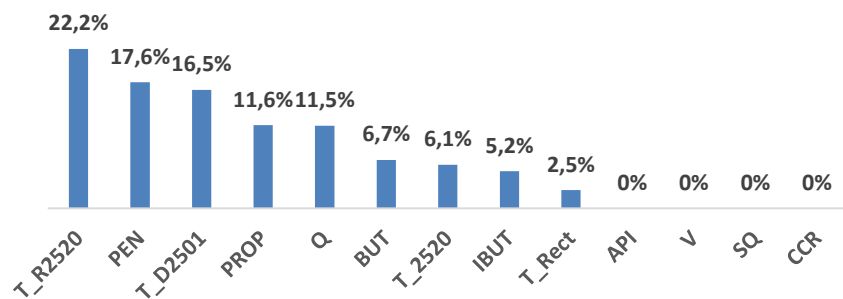


(b) Variación del RMSE según el número de factores

Figura 40. Resultados regresión PLS para la subdivisión 1 de la clase 1. (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión. (d) Influencia de las variables independientes.

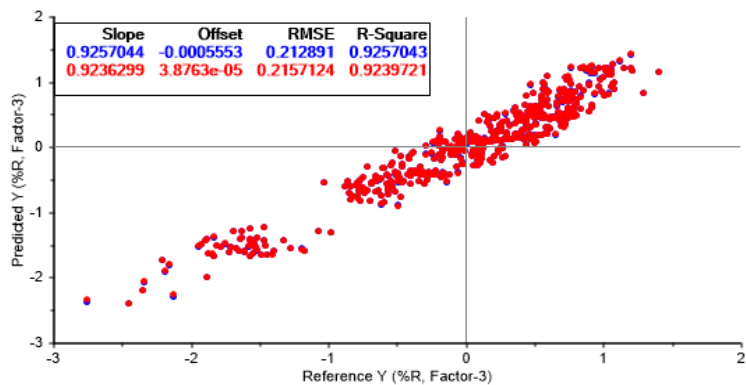


(c) Coeficientes obtenidos por la regresión.



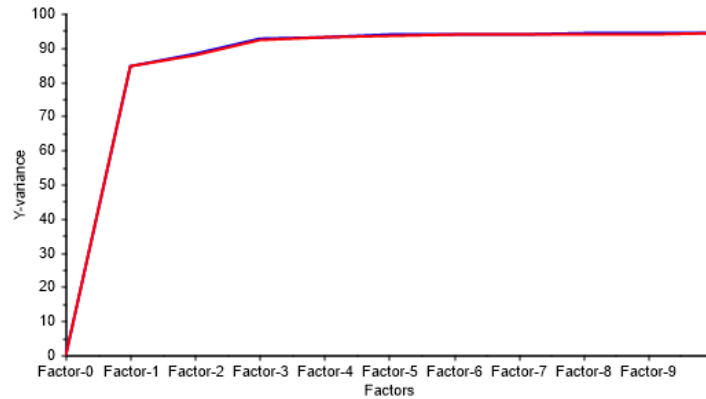
(d) Influencia variables independientes.

Figura 41. Resultados regresión PLS para la subdivisión 1 de la clase 1 (continuación). (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión. (d) Influencia de las variables independientes.

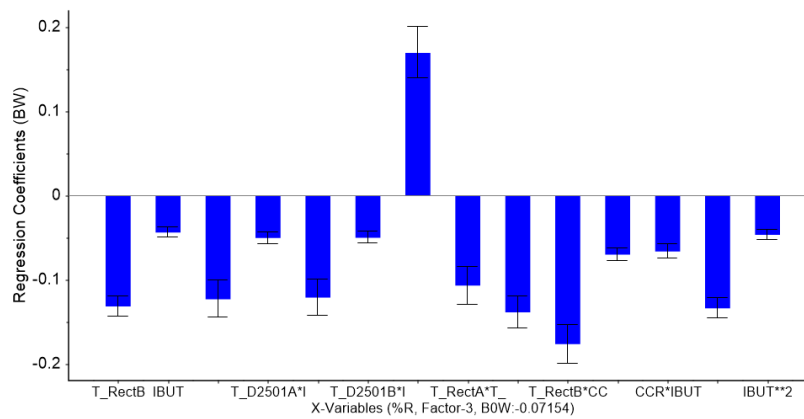


(a) Valores predichos vs valores de referencia.

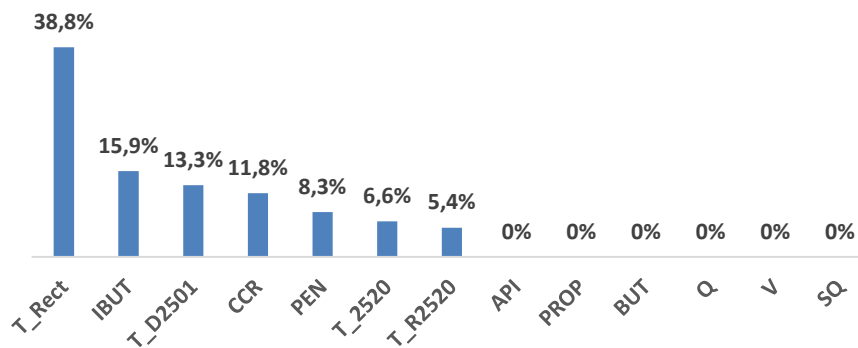
Figura 42. Resultados regresión PLS para la subdivisión 2 de la clase 1.



(b) Variación del RMSE según el número de factores.



(c) Coeficientes obtenidos por la regresión.



(d) Influencia variables independientes.

Figura 43. Resultados regresión PLS para la subdivisión 2 de la clase 1. (a) Valores predichos vs valores de referencia. (b) Variación del RMSE según el número de factores. (c) Coeficientes obtenidos por la regresión.

Tabla 4.

Resumen modelos ajustados para la clase1.

	Modelo	Número de entradas	R2	RMSE
Todos los datos	RLM	16	0.889 (0.887)	0.228 (0.230)
	PLS	6 Factores resultantes de la combinación lineal de 38 entradas.	0.908 (0.906)	0.208 (0.211)
Subdivisión1	RLM	17	0.778 (0.773)	0.363 (0.367)
	PLS	3 Factores resultantes de la combinación lineal de 13 entradas.	0.812 (0.809)	0.257 (0.259)
Subdivisión2	RLM	16	0.873	0.283
	PLS	3 Factores resultantes de la combinación lineal de 14 entradas.	0.926 (0.924)	0.213 (0.216)

$$\begin{aligned}
\%R = & -0.02939563 - 0.1349153 * T_{D2501A} - 0.1392738 * T_{D2501B} - 0.2494535 * T_{RectB} \\
& + 0.1163927 * API - 0.1129091 * V - 0.1176338 * PROP + 0.04913263 \\
& * T_{D2501A} * API - 0.08283015 * T_{D2501A} * BUT + 0.1136433 * T_{D2501B} \\
& * T_{RectA} + 0.06658479 * T_{D2501B} * API + 0.1274585 * T_{D2501B} * Q \\
& + 0.1249363 * T_{D2501B} * PROP + 0.07294101 * T_{2520A} * T_{RectA} - 0.1207292 \\
& * T_{2520A} * PROP + 0.134677 * T_{2520A} * IBUT + 0.0841281 * T_{2520B} * T_{RectA} \\
& + 0.07362676 * T_{2520B} * T_{R2520B} + 0.04960961 * T_{2520B} * PROP \\
& + 0.08815237 * T_{2520B} * IBUT + 0.2178191 * T_{RectA} * API - 0.04744632 \\
& * T_{RectA} * S_Q - 0.200455 * T_{RectB} * API - 0.09727894 * T_{RectB} * V \\
& - 0.1205831 * T_{RectB} * PEN - 0.0511715 * T_{RectB} * IBUT + 0.1058134 \\
& * T_{RectB} * BUT - 0.1930533 * T_{R2520A} * Q - 0.03922813 * API * Q \\
& + 0.07730959 * API * PEN + 0.04621422 * API * IBUT - 0.1539221 * API \\
& * BUT - 0.1112718 * CCR * PEN + 0.07702775 * CCR * PROP \\
& + 0.06044926 * Q * S_Q + 0.1478131 * PROP * BUT + 0.1150301 * T_{RectB}^{(3)2} \\
& - 0.1367152 * T_{R2520A}^2 - 0.0951406 * PROP^2
\end{aligned}$$

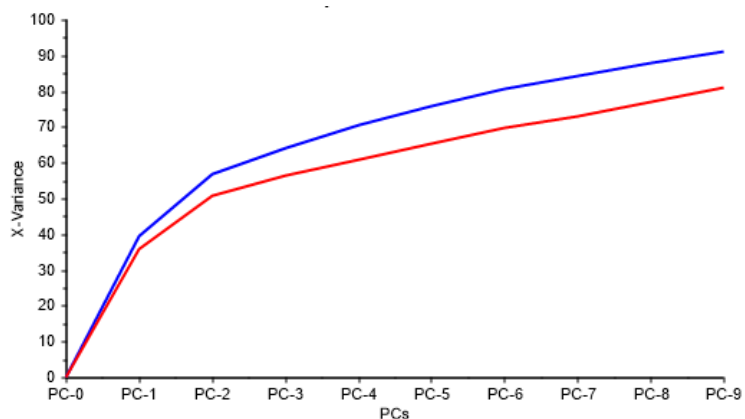
$$\begin{aligned}
\%R = & -0.03444 - 0.07532434 * T_{D2501A} - 0.0731402 * T_{D2501B} - 0.07628634 * T_{D2501A} \\
& * T_{D2501B} - 0.1072249 * T_{D2501A} * T_{R2520A} - 0.1067742 * T_{D2501B} * T_{R2520A} \\
& + 0.06727139 * T_{2520A} * T_{RectA} + 0.09394754 * T_{2520A} * T_{R2520B} \\
& + 0.1931411 * T_{R2520B} * Q + 0.1449376 * PEN * BUT - 0.1109933 * PROP \\
& * iBUT - 0.08964583 * T_{R2520A}^2 + 0.1500444 * PEN^2 - 0.1386758 * PROP^2
\end{aligned}$$

$$\begin{aligned}
\%R = & -0.07154 - 0.131335 * T_{RectB} - 0.043664 * iBUT - 0.122686 * T_{D2501A} * T_{RectB} \\
& - 0.05005 * T_{D2501A} * iBUT - 0.12076 * T_{D2501B} * T_{RectB} - 0.04973 * T_{D2501B} \\
& * iBUT + 0.16944 * T_{2520A} * PEN - 0.106618 * T_{RectA} * T_{RectB} - 0.1383 \\
& * T_{RectB} * T_{R2520A} - 0.17599 * T_{RectB} * CCR - 0.0699 * T_{RectB} * iBUT \\
& - 0.06612 * CCR * iBUT - 0.13349 * T_{RectB}^2 - 0.04628 * iBUT^2
\end{aligned} \tag{5}$$

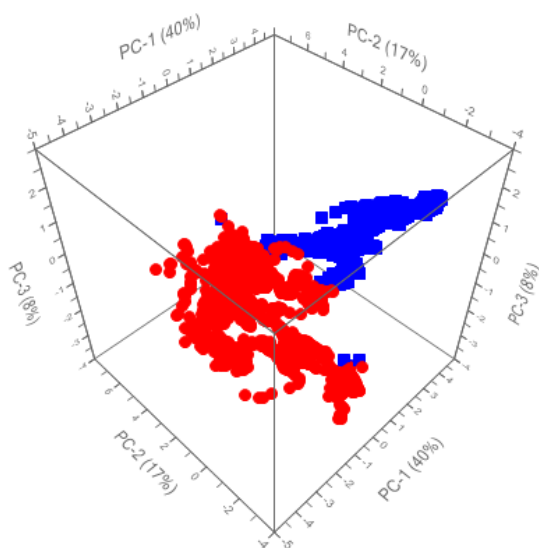
Según los desempeños de las diferentes regresiones (Tabla 3.3), el rendimiento de DMO en el esquema operativo de la clase 1 (Tabla 3.1) es representado semicuantitativamente por la regresión PLS (ecuación 3). Según esta regresión, la temperatura de la rectificación de a mezcla DMO y solvente influye 20.7% en el valor del rendimiento (Figura 3.22d); en conjunto, las temperaturas de operación influyen en un 48.7%, seguidas de las propiedades de la carga con una influencia de 23.1% y la composición del solvente con una incidencia de 21.7%. La mayor influencia de las temperaturas de operación radica en los altos flujos de solvente utilizados para la extracción; un ligero aumento en las temperaturas de operación modifica el rendimiento según lo expuesto en la Figura 3.14a. Por otra parte, la regresión PLS para la subdivisión 1 (ecuación 4), aunque con mejor desempeño que la respectiva MLR, presenta un desempeño que soporta su utilidad sólo a nivel cualitativo; según la ecuación (4), las variables operacionales presentan la mayor influencia

(47.3%), sin embargo, el solvente ocupa un segundo lugar con una incidencia de 23.5%; respecto a las propiedades, la única propiedad influyente en esta regresión corresponde a la penetración (17.6%). Para la subdivisión 2, la regresión PLS (ecuación 5) presenta una aplicación a nivel semicuantitativo (Tabla 3.3). La incidencia de las condiciones operativas, del solvente y de las propiedades de la carga alcanzan 64.1%, 15.9% y 20.1%, respectivamente.

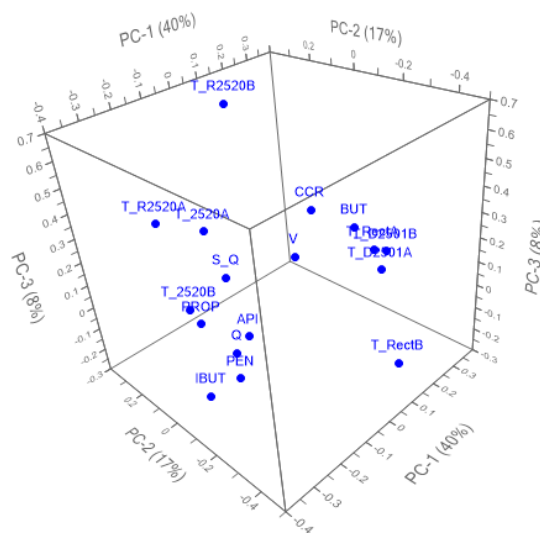
4.3.3 Clase 2 El análisis PCA aplicado a las 2557 muestras que conforman la clase 2 puede ser consultado en las Figuras 3.25a-c; la etapa de calibración del PCA (Figura 3.25a) reportó una reproducción en la varianza de las variables independientes de *ca.* 92% por medio de 9 componentes principales (82% en validación). La gráfica de los *scores* en las coordenadas de los 3 primeros PC (Figuras 3.25b), muestra una subdivisión marcada de los datos sobre el eje de PC-1 (40% varianza), con lo cual, el valor de los *loadings* para PC-1 refleja la influencia de las variables en la subdivisión de los datos (Figura 3.25c). Según el valor de los *loadings* para PC-1, las variables IBUT, PEN, Q y API son las principales responsables de la aparición de la subdivisión 1 (puntos rojos en Figura 3.25b-c), mientras que las variables CCR, BUT y T_D2501 son las responsables de la aparición de la subdivisión 2 (puntos azules en Figura 3.25b-c). Asimismo, un análisis de las tendencias en las anteriores variables permite definir los rangos para cada subdivisión; las Figuras 3.26a-b ilustran los límites para algunas variables, mientras que la Tabla 3.4 presenta los valores para la formación de las subdivisiones. Es importante mencionar que los valores de las variables para la formación de las subdivisiones en algunos casos se presenta difuso (Figura 3.26b), por lo cual las fronteras fueron definidas como un promedio entre los respectivos datos de las subdivisiones. Las demás variables continúan según la clase 2 (Tabla 3.1).



(a) Varianza explicada vs número de componentes principales.



(b) Gráfica de los *scores*.



(c) Gráfica de los *loadings*.

Figura 44. Resultados análisis PCA para los datos que conforman la clase 2.

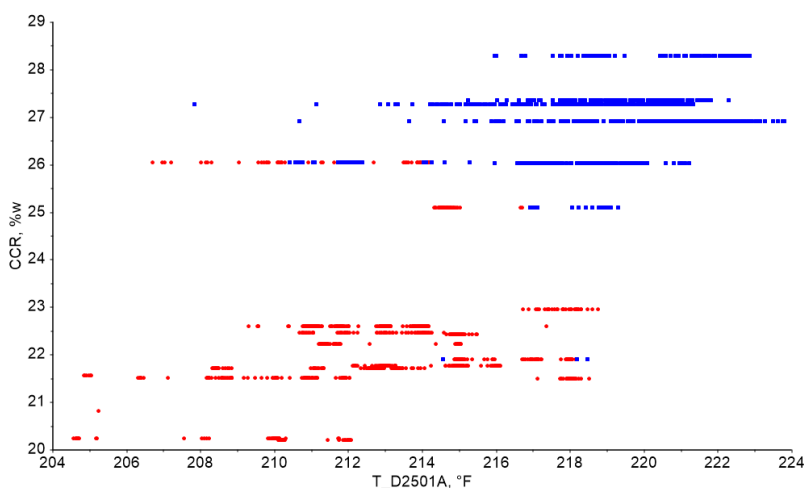
Los rangos totales de las variables y los rangos de las variables en las subdivisiones fueron tomados para ajustar regresiones MLR y PLS dirigidas a la predicción del rendimiento de DMO (%R). Respecto a las muestras totales de la clase 2, las Figuras 3.27 y 3.28 presentan los desempeños para las regresiones MLR y PLS, respectivamente. Según estas figuras, la regresión MLR obtenida presenta mejor desempeño en el ajuste del %R que la regresión PLS. La regresión

MLR reproduce un 93.6% de la varianza (93.1% en validación), utilizando 49 entradas (Figura 3.26c-d, Tabla 3.5) y cumpliendo los supuestos de normalidad, homoscedasticidad (Figura 3.26b) y significancia estadística (Figura 3.26d); la influencia de las condiciones operativas en la regresión alcanza un 52.2%, mientras que el solvente y las propiedades de la carga alcanzan 25.6% y 16.7%, respectivamente (Figura 3.27e). De otro lado, aunque la regresión PLS presenta menor desempeño que la MLR, reproduciendo el 90% de la varianza, su aplicación se encuentra favorecida debido a la menor cantidad de entradas (6 entradas) (Figura 3.28c, Tabla 3.5). La significancia estadística de los coeficientes de la regresión PLS es comprobada por los bajos valores de los intervalos de confianza respectivos (Figura 3.28c); la influencia de las condiciones operativas, según PLS, alcanza un 39.8%, mientras que la influencia del solvente y las propiedades de la carga llegan a 40.9% y 10%, respectivamente (Figura 3.28d).

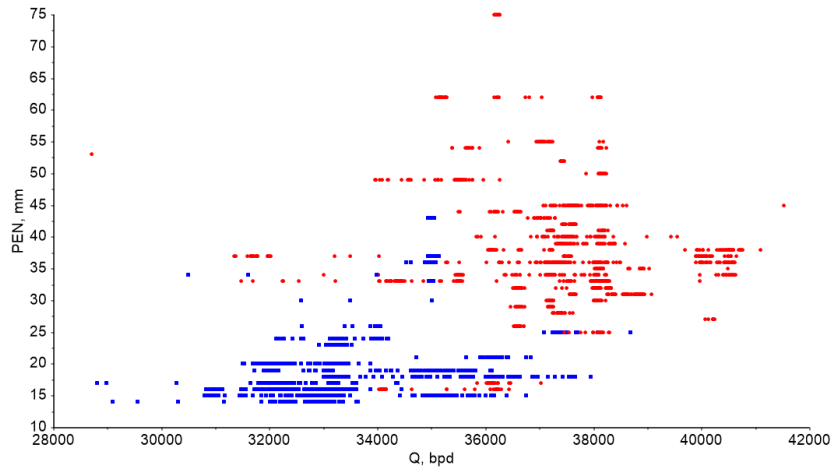
Por otra parte, la subdivisión de la clase 2 mejora el desempeño de los modelos de regresión MLR y PLS. En las dos subdivisiones los desempeños de las regresiones MLR mejoran los obtenidos por las regresiones PLS (Figuras 3.29, 3.30, 3.31, 3.32; Tabla 3.5); la reproducción de las varianzas con las regresiones MLR alcanza valores de 97.3% (96.3% en validación) y 94.4% (94.3% en validación) para las subdivisiones 1 y 2, respectivamente, con lo cual la aplicación de estos modelos se encuentra en el plano cuantitativo; sin embargo, la aplicación de la MLR para la subdivisión 1 puede estar restringida por la cantidad de entradas requeridas (45, Tabla 3.5). Los modelos MLR para las subdivisiones cumplen los supuestos correspondientes (Figura 3.29c-d y 3.31b-c). Respecto a la influencia de las variables en los modelos MLR, en la subdivisión 1, el solvente alcanza una incidencia del 50.3%, mientras que las condiciones de operación un 26.3% y las propiedades de la carga un 21.7% (Figura 3.29e). En la subdivisión 2, la influencia del solvente y las condiciones de operación disminuyen a 39.2% y 20%, respectivamente, mientras que la

incidencia de las propiedades de la carga aumenta a 30.8% (Figura 3.31e). La incidencia del solvente en esta clase es mayor que en las anteriores debido a la menor carga tratada en la unidad (Tabla 3.1).

Por su parte, las regresiones PLS en las subdivisiones presentan igual desempeño (*i.e.* reproducción de la varianza del rendimiento de DMO en 92% y RMSE de 0.2) con un número de entradas bajo (3 para la subdivisión 1 y 4 para la subdivisión 2); la aplicación de estas regresiones presenta una utilidad a nivel semicuantitativo. La regresión PLS de la subdivisión 1 presenta una influencia del IBUT, del vanadio (V) y de la temperatura de rectificación T_R2520 de 41.7%, 24.7% y 16.8%, respectivamente (Figura 3.30d), mientras que la subdivisión 2 muestra una incidencia de las propiedades de la carga de 36.9%, del solvente de 37.9% y de la temperatura de extracción T_2520 de 12% (Figura 3.32d). Las regresiones PLS coinciden en que la mayor influencia corresponde a la del solvente, seguida por las propiedades de la carga a procesar.



(a) Gráfica de CCR vs T_D2501A.



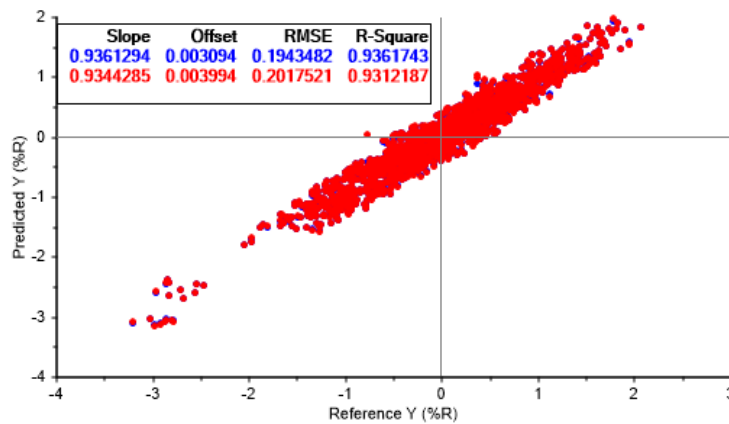
(b) Gráfica de PEN vs Q.

Figura 45. Subdivisión de las muestras de la clase 2 según el método *k-mean* ($k=2$).

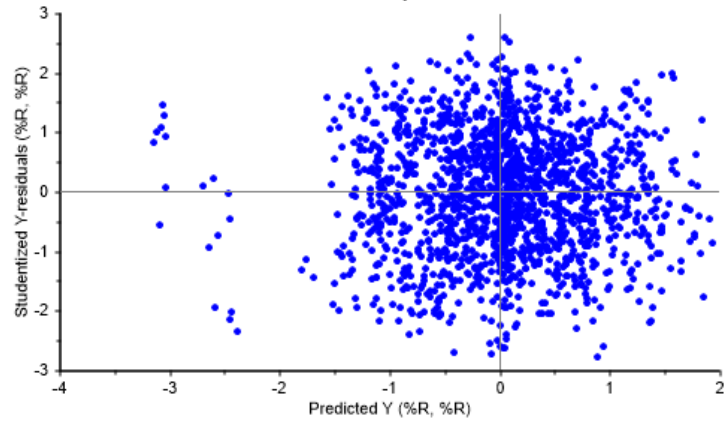
Tabla 5.

Variables que definen la subdivisión en la clase 2.

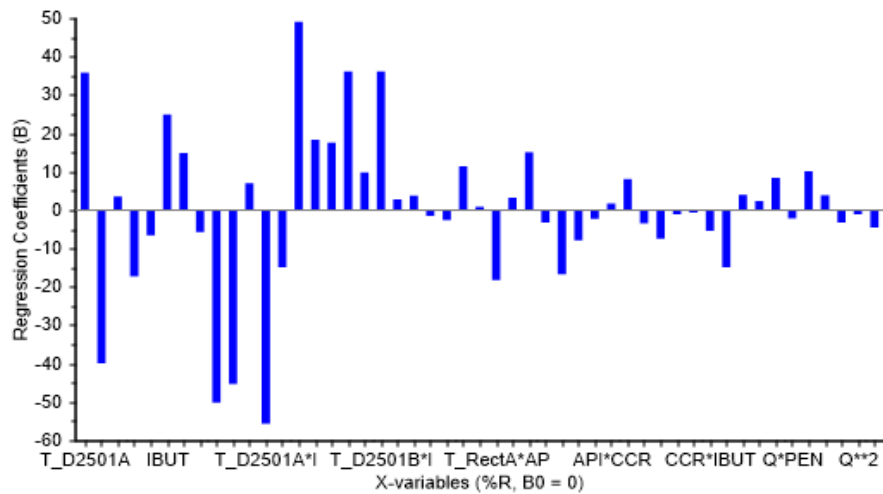
	T_D250 1A, °F	T_D2501 B, °F	Q, kbps	BUT, %v	IBUT, %v	API	PEN	CCR, %w
Subdivisión 1	[208, 217]	[207, 219]	[31.1, 36]	[54, 65]	[18, 29]	[3, 5]	[14, 36]	[21.5, 25]
Subdivisión 2	[217, 221]	[219, 225]	[36, 38.8]	[65, 74]	[29, 41]	[5, 6]	[36, 49]	[25, 27]



(a) Valor predicho vs valor de referencia.



(b) Residuales estudentizados.



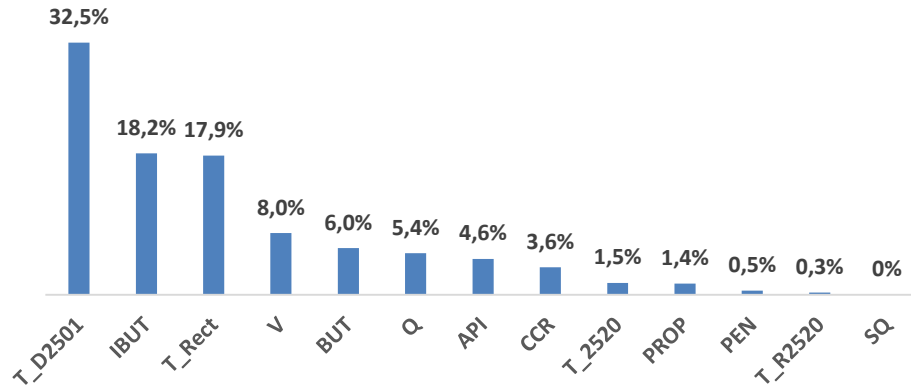
(c) Coeficientes obtenidos en la regresión.

Figura 46. Resultados regresión MLR para las muestras de la clase 2.

Anova Table							
Multiple Correlation: 0.9675657 (cal) 0.9650112 (val)							
R-Square: 0.9361743 (cal) 0.9312187 (val)							
	SS	df	MS	F ratio	p value	B-coeficien	STDerr
Summary							
Model	932.11352	49	19.022724	488.9583	0		
Error	63.531196	1633	0.0389045				
Total	995.64471	1682	0.5919409				
Variables							
Intercept	0	0				0	0
T_D2501A	13.199887	1	13.199887	339.2776	0	35.7908554	1.9430968
T_D2501B	10.498572	1	10.498572	269.8510	0	-39.8868752	2.4281077
T_2520B	2.9950993	1	2.9950993	76.98585	4.24655998E-18	3.58463454	0.4085445
Q	12.220121	1	12.220121	314.1068	0	-17.2315845	0.9722692
PROP	2.5801274	1	2.5801274	66.31942	7.53444171E-16	-6.60413694	0.8109530
IBUT	4.7345314	1	4.7345314	121.6976	2.41769439E-27	24.9409256	2.260849
BUT	1.5860816	1	1.5860816	40.76867	2.22794588E-10	14.8842297	2.3311114
T_D2501A*T_252	4.6462569	1	4.6462569	119.4231	7.03199547E-27	-5.68310452	0.5200454
T_D2501A*T_Rect	13.287906	1	13.287906	341.5409	0	-50.1029778	2.7110786
T_D2501A*V	9.9592008	1	9.9592008	255.9808	0	-45.3070679	2.8317978
T_D2501A*PROP	2.3728358	1	2.3728358	60.99129	1.01766765E-14	7.00235176	0.8966232
T_D2501A*IBUT	16.721805	1	16.721805	429.8152	0	-55.6759605	2.6855104
T_D2501B*T_Rect	4.5921058	1	4.5921058	118.0352	1.3500849E-26	-14.8512554	1.3669645
T_D2501B*T_Rect	10.053177	1	10.053177	258.4012	0	49.0621452	3.0521035
T_D2501B*API	6.934618	1	6.934618	178.2472	1.15043283E-38	18.3288708	1.3728539
T_D2501B*CCR	4.7667856	1	4.7667856	122.5245	1.64052793E-27	17.5536785	1.5858298
T_D2501B*V	6.1905212	1	6.1905212	159.1160	7.05164482E-35	36.0896339	2.8610498
T_D2501B*Q	9.2505445	1	9.2505445	237.7772	0	9.83165646	0.6375900
T_D2501B*IBUT	9.9684028	1	9.9684028	256.2284	0	36.1612816	2.2590723
T_2520A*IBUT	4.2904992	1	4.2904992	110.2806	5.22135644E-25	2.77595091	0.2643395
T_2520A*BUT	3.9052810	1	3.9052810	100.3774	5.71159782E-23	3.8144443	0.3807265
T_2520B*IBUT	1.5872892	1	1.5872892	40.79956	2.19385066E-10	-1.44957995	0.2269417
T_2520B*BUT	3.6401276	1	3.6401276	93.56569	1.46928754E-21	-2.54041505	0.2626312
T_RectA*T_RectB	3.6944761	1	3.6944761	94.96288	7.53860826E-22	11.423213	1.1722259
T_RectA*T_R2520	1.4087141	1	1.4087141	36.20951	2.18316076E-09	0.920766056	0.1530163
T_RectA*API	8.5675964	1	8.5675964	220.2220	0	-18.1740646	1.2246772
Anova Table							
T_RectA*PEN	1.5217176	1	1.5217176	39.11398	5.09352338E-10	3.27380657	0.5234642
T_RectA*IBUT	3.8038601	1	3.8038601	97.77430	1.97224377E-22	15.1327953	1.5304064
T_RectB*V	0.7908999	1	0.7908999	20.32919	6.98436224E-06	-3.21342754	0.7127028
T_RectB*IBUT	3.3413419	1	3.3413419	85.88684	5.82022955E-20	-16.7108803	1.8031672
T_RectB*BUT	0.4034379	1	0.4034379	10.37002	0.00130596396	-7.82314348	2.4293570
T_R2520B*IBUT	1.5169255	1	1.5169255	38.99094	5.41687251E-10	-2.24705219	0.3598580
API*CCR	0.9863774	1	0.9863774	25.35383	5.30030434E-07	1.69486976	0.3366003
API*Q	15.582440	1	15.582440	400.5283	0	8.07302094	0.4033847
API*IBUT	2.9522047	1	2.9522047	75.88332	7.23868732E-18	-3.39502239	0.3897350
API*BUT	9.6538276	1	9.6538276	248.1416	0	-7.40841627	0.4703006
CCR*PEN	0.7057874	1	0.7057874	18.14151	2.16770368E-05	-1.02470756	0.2405820
CCR*PROP	0.3080734	1	0.3080734	7.918685	0.00495132897	-0.60817259	0.2161226
CCR*IBUT	4.9973750	1	4.9973750	128.4517	1.02417921E-28	-5.38756323	0.4753596
CCR*BUT	4.9095702	1	4.9095702	126.1945	2.94166371E-28	-14.8784342	1.3244541
V*Q	11.075243	1	11.075243	284.6772	0	4.00540781	0.2373943
V*IBUT	3.4829671	1	3.4829671	89.52635	1.01520394E-20	2.33862257	0.2471636
V*BUT	5.2986288	1	5.2986288	136.1958	2.77486501E-30	8.37352753	0.7175075
Q*PEN	1.4820632	1	1.4820632	38.09455	8.48419279E-10	-2.17233872	0.3519622
Q*IBUT	27.371397	1	27.371397	703.5524	0	10.1889668	0.3841332
T_D2501A**2	0.8385692	1	0.8385692	21.55466	3.71393003E-06	3.89244318	0.8384007
CCR**2	0.8633641	1	0.8633641	22.19176	2.67641121E-06	-3.22402096	0.6843872
Q**2	0.2374148	1	0.2374148	6.102475	0.0136007825	-1.09824836	0.4445775
IBUT**2	8.4756650	1	8.4756650	217.8563	0	-4.490376	0.3042269

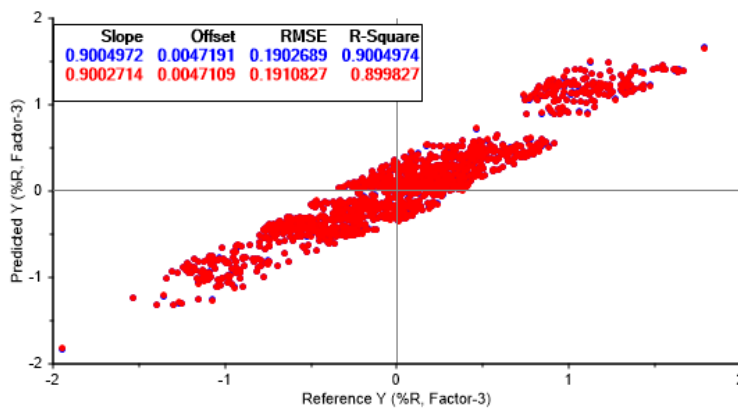
(d) Tabla ANOVA.

Figura 47. Resultados regresión MLR para las muestras de la clase 2 (continuación).

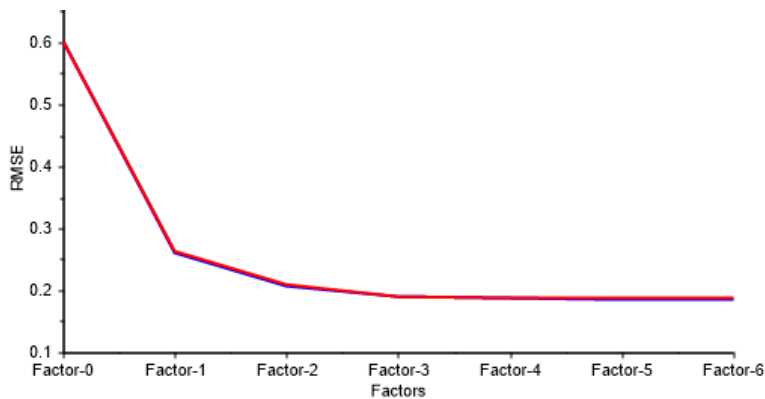


(e) Influencia de las variables independientes en la predicción.

Figura 48. Resultados regresión MLR para las muestras de la clase 2.

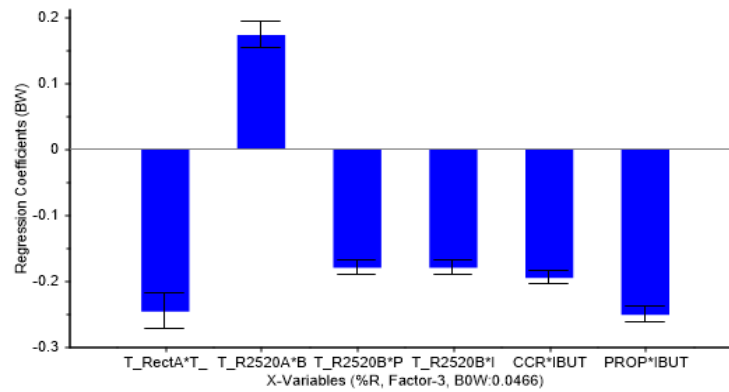


(a) Valor predicho vs valor de referencia.

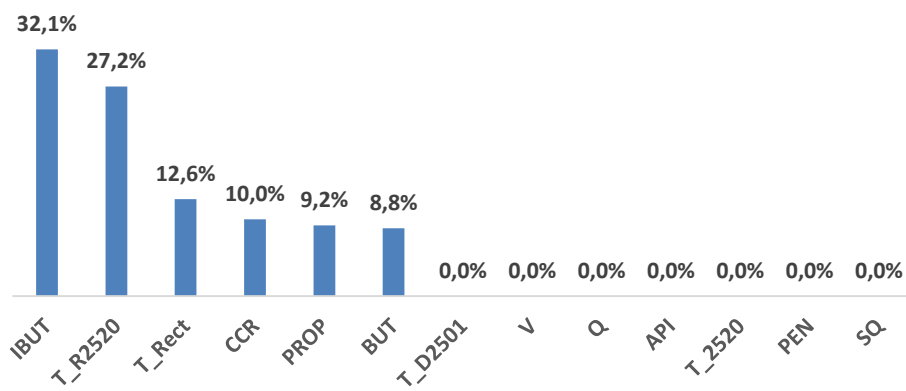


(b) Variación RMSE con el número de factores de la regresión.

Figura 49. Resultados regresión PLS para las muestras de la clase 2.

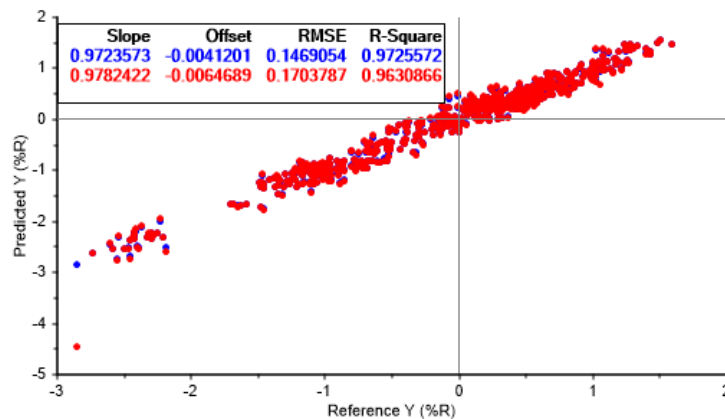


(c) Coeficientes de la regresión.



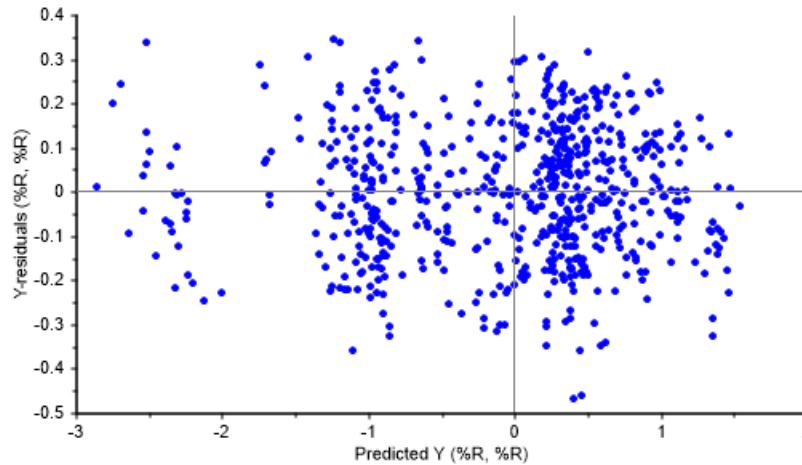
(d) Influencia de las variables independientes en el %R.

Figura 50. Resultados regresión PLS para las muestras de la clase 2 (continuación).



(a) Valor predicho vs valor de referencia.

Figura 51. Resultados regresión MLR para las muestras de la subdivisión 1 para la clase 2.



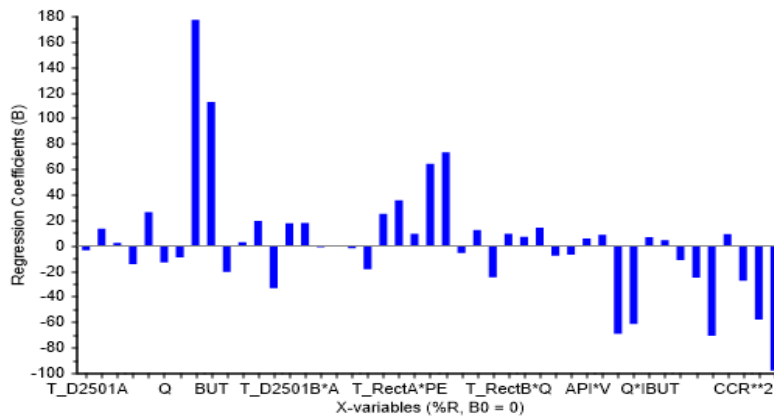
(b) Residuales de la regresión.

Anova Table							
Multiple Correlation: 0.9861858 (cal) 0.9815021 (val)							
R-Square: 0.9725572 (cal) 0.9630866 (val)							
	SS	df	MS	F ratio	p value	B-coeficie	STDerr
Summary							
Model	511.682892	45	11.3707314	491.17395	0		
Error	14.3299179	619	0.02315011				
Total	526.012817	664	0.79218798				
Variables							
Intercept	0	0				0	0
T_D2501A	0.08674211	1	0.08674211	3.7463333	0.0533785522	-3.6067454	1.8634265
T_D2501B	1.23721313	1	1.23721313	53.43713	8.28184906E-13	13.3195238	1.8220787
T_2520B	0.19415718	1	0.19415718	8.3868875	0.00391277019	2.27909756	0.7869778
T_R2520A	0.45847511	1	0.45847511	19.804628	1.01727255E-05	-14.602153	3.2812066
CCR	0.63570815	1	0.63570815	27.459341	2.20318441E-07	26.370491	5.0323767
Q	0.98002499	1	0.98002499	42.333416	1.58936364E-10	-13.168047	2.0238564
PROP	0.80894196	1	0.80894196	34.941631	5.61493341E-09	-9.3103838	1.5750559
IBUT	0.51263582	1	0.51263582	22.142063	3.12582961E-06	177.342117	37.687976
BUT	0.24447904	1	0.24447904	10.559620	0.00121847307	112.800896	34.712699
T_D2501A*T_Rect	1.55433691	1	1.55433691	67.143234	1.45305061E-15	-20.624378	2.5169799
T_D2501A*T_R25	0.15786878	1	0.15786878	6.8193073	0.00923634041	2.80398083	1.0737543
T_D2501A*BUT	1.52640641	1	1.52640641	65.934646	2.52755627E-15	19.4656544	2.3972437
T_D2501B*T_Rect	2.29098582	1	2.29098582	98.951416	9.97291335E-22	-33.235183	3.3410816
T_D2501B*API	1.14513993	1	1.14513993	49.464965	5.36256456E-12	17.4892941	2.4867000
T_D2501B*V	2.90480185	1	2.90480185	125.47494	1.20180871E-26	17.8192883	1.5907863
T_2520A*IBUT	0.16969674	1	0.16969674	7.3303017	0.00696737692	-1.1010540	0.4066752
T_2520A*BUT	0.17399568	1	0.17399568	7.5160217	0.00629194267	0.42128005	0.1536656
T_2520B*BUT	0.19586882	1	0.19586882	8.4608240	0.00375893572	-1.9461848	0.6690795
T_RectA*API	0.67438101	1	0.67438101	29.130386	9.65069944E-08	-18.451629	3.4187047
T_RectA*CCR	1.24468553	1	1.24468553	53.764633	7.10399784E-13	24.9365864	3.4008593
T_RectA*PEN	3.53795838	1	3.53795838	152.82353	1.57117527E-31	35.6564789	2.8843193
T_RectA*PROP	0.81385105	1	0.81385105	35.153610	5.0652802E-09	9.26582336	1.5627843
T_RectA*IBUT	1.52613056	1	1.52613056	65.917366	2.54766568E-15	64.3669205	7.9279851
T_RectA*BUT	1.63202548	1	1.63202548	70.492286	3.15193877E-16	73.3211517	8.7328996

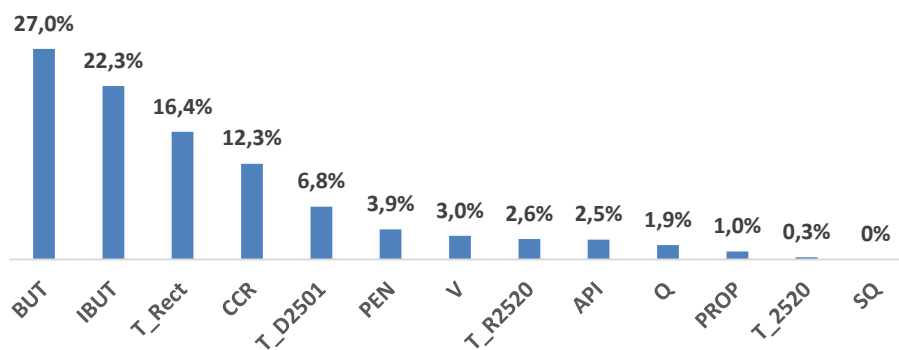
(c) Tabla ANOVA

Anova Table							
T_RectB*API	0.16509905	1	0.16509905	7.1317782	0.00777193857	-5.8606224	2.1945467
T_RectB*CCR	1.55075312	1	1.55075312	66.984771	1.56233999E-15	12.1640835	1.4862482
T_RectB*V	3.18600368	1	3.18600368	137.61872	7.73971902E-29	-24.743681	2.1092376
T_RectB*Q	0.49262833	1	0.49262833	21.280181	4.82451651E-06	9.26226425	2.0078418
T_R2520A*IBUT	0.30352515	1	0.30352515	13.111292	0.000317553728	7.01951361	1.9385824
T_R2520A*BUT	0.54955244	1	0.54955244	23.738857	1.40323959E-06	14.0796232	2.8897557
T_R2520B*IBUT	0.22659046	1	0.22659046	9.7878150	0.00183926756	-7.8162870	2.4983749
T_R2520B*BUT	0.14653754	1	0.14653754	6.3298482	0.0121241547	-7.0382556	2.7974889
API*V	1.42664099	1	1.42664099	61.621749	1.83927462E-14	5.63607931	0.7179762
CCR*V	0.37864899	1	0.37864899	16.354446	5.91671269E-05	8.57626057	2.1207039
CCR*IBUT	1.42535686	1	1.42535686	61.566864	1.88649755E-14	-69.038139	8.7986402
CCR*BUT	0.84560525	1	0.84560525	36.525558	2.60342059E-09	-61.491596	10.174599
Q*IBUT	0.26739057	1	0.26739057	11.550336	0.000720862648	6.67366505	1.9636635
Q*BUT	0.09059913	1	0.09059913	3.9135835	0.0483403057	4.42000723	2.2342701
PEN*IBUT	2.73605418	1	2.73605418	118.18426	2.5900288E-25	-11.347893	1.0438436
PEN*BUT	3.78495598	1	3.78495598	163.49321	2.18432417E-33	-25.226810	1.9729336
IBUT*BUT	0.49829858	1	0.49829858	21.522829	4.26907491E-06	-70.692596	15.237870
T_D2501A**2	0.34512180	1	0.34512180	14.906849	0.000124827726	9.03916264	2.3411824
CCR**2	2.90034628	1	2.90034628	125.28309	1.30241536E-26	-27.566265	2.4628143
IBUT**2	0.75610417	1	0.75610417	32.658336	1.70859416E-08	-58.104438	10.167455
BUT**2	0.40802678	1	0.40802678	17.623672	3.08705348E-05	-98.012245	23.347057

(c) Tabla ANOVA (continuación)

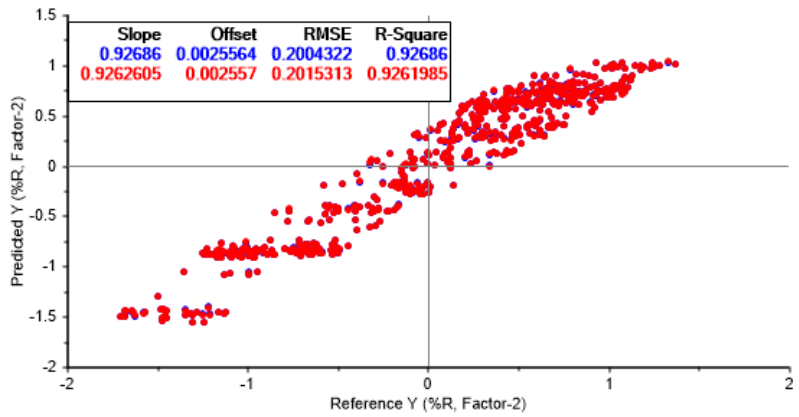


(d) Coeficientes de la regresión.

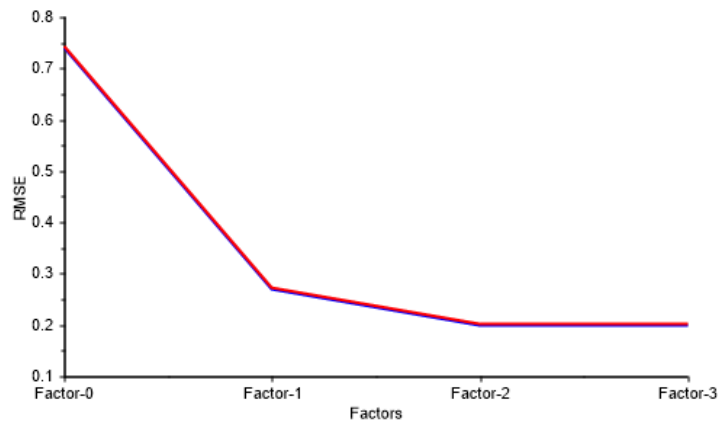


(e) Influencia de las variables independientes en la predicción.

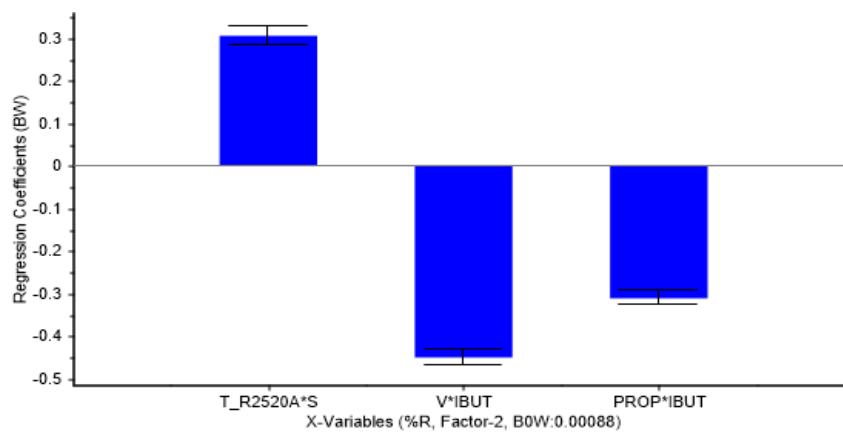
Figura 52. Resultados regresión MLR para las muestras de la subdivisión 1 para la clase 2.



(a) Valor predicho vs valor de referencia.

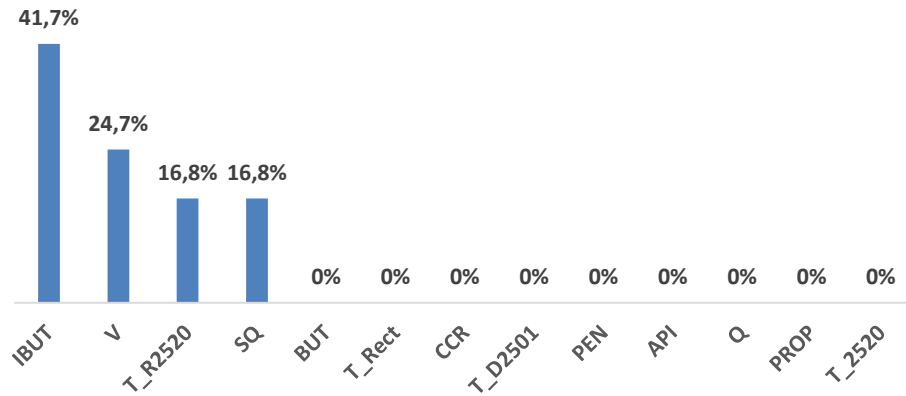


(b) Variación RMSE con los factores en la regresión.



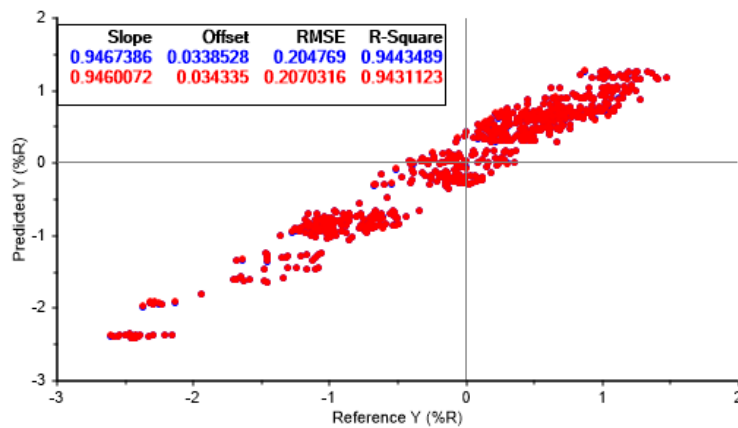
(c) Coeficientes de la regresión.

Figura 53. Resultados regresión PLS para las muestras de la subdivisión 1 para la clase 2.

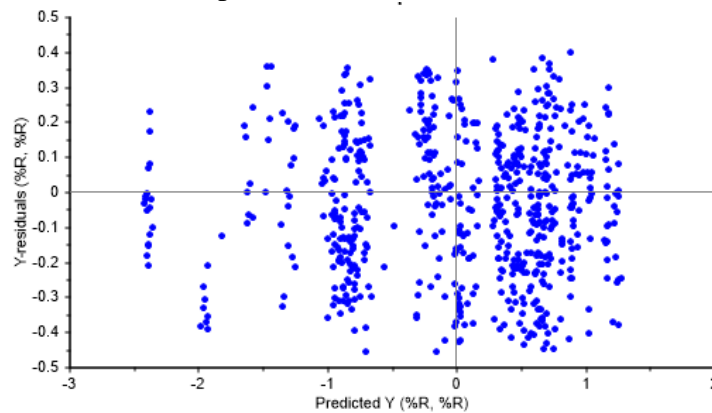


(d) Influencia de las variables independientes en la predicción.

Figura 54. Resultados regresión PLS para las muestras de la subdivisión 1 para la clase 2 (continuación).



(a) Valores predichos vs valores de referencia.

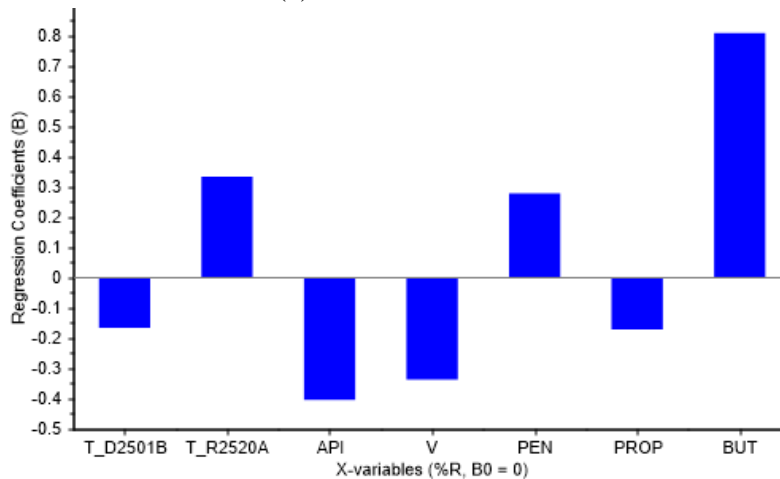


(b) Residuales de la regresión.

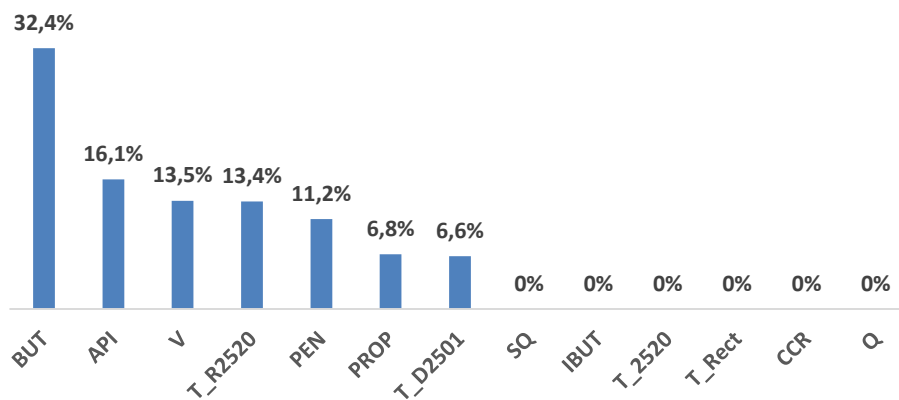
Figura 55. Resultados regresión MLR para las muestras de la subdivisión 2 de la clase 2.

Anova Table							
Multiple Correlation: 0.9726848 (cal) 0.9720753 (val)							
R-Square: 0.9443489 (cal) 0.9431123 (val)							
	SS	df	MS	F ratio	p value	B-coeficien	STDerr
Summary							
Model	457.646	7	65.3780	1542.177	0		
Error	26.8773	634	0.04239				
Total	484.523	641	0.75588				
Variables							
Intercept	0	0				0	0
T_D2501B	8.61456	1	8.61456	203.2052	3.39341239E-40	-0.164511904	0.0115406
T_R2520A	17.5457	1	17.5457	413.8797	0	0.334692955	0.0164516
API	26.1115	1	26.1115	615.9350	0	-0.403427213	0.0162554
V	54.5770	1	54.5770	1287.396	0	-0.336600721	0.0093812
PEN	13.2586	1	13.2586	312.7522	0	0.27960968	0.0158107
PROP	11.9491	1	11.9491	281.8643	0	-0.170115232	0.0101326
BUT	115.553	1	115.553	2725.739	0	0.810400963	0.0155223

(c) Tabla ANOVA.

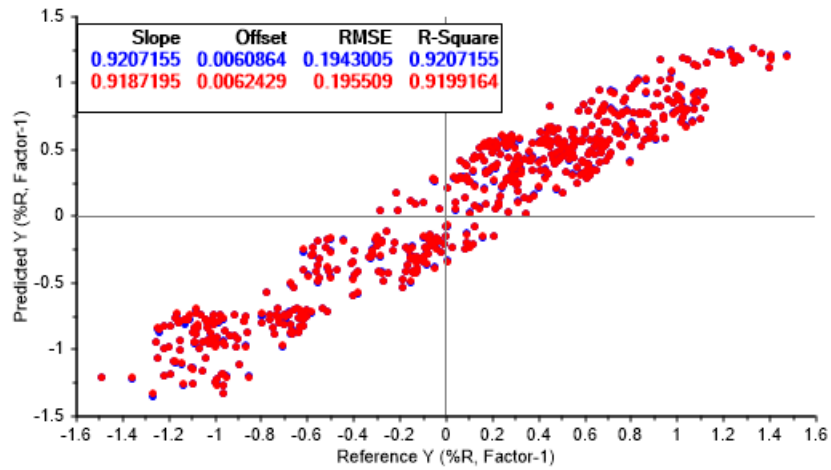


(d) Coeficientes de la regresión.

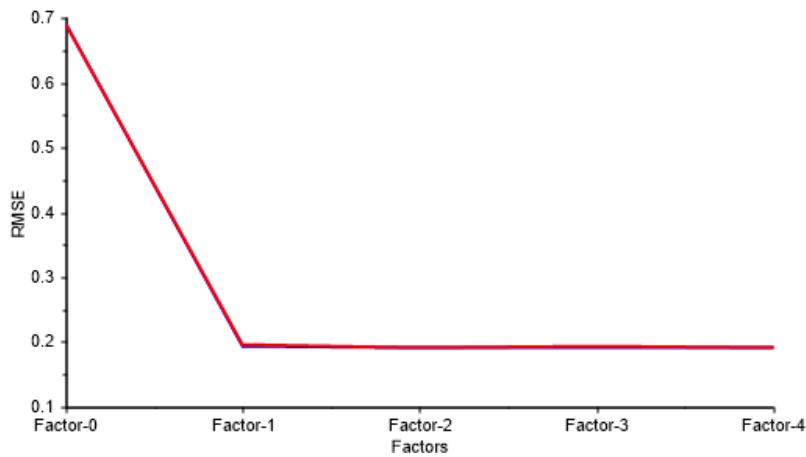


(e) Influencia de las variables independientes en la predicción.

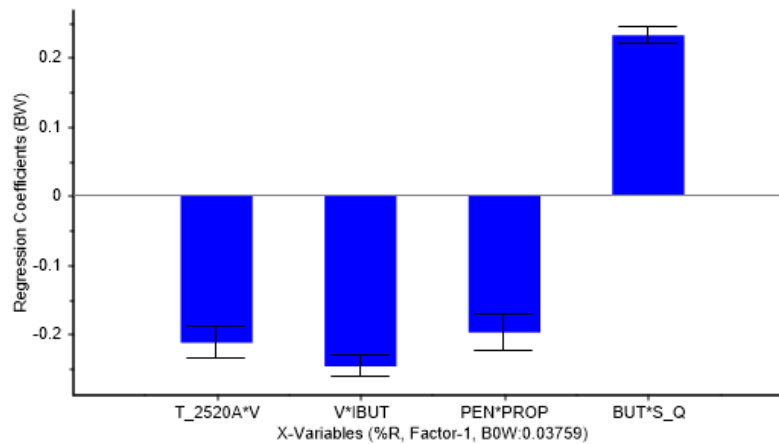
Figura 56. Resultados regresión MLR para las muestras de la subdivisión 2 de la clase 2 (continuación).



(a) Valores predichos vs valores de referencia.

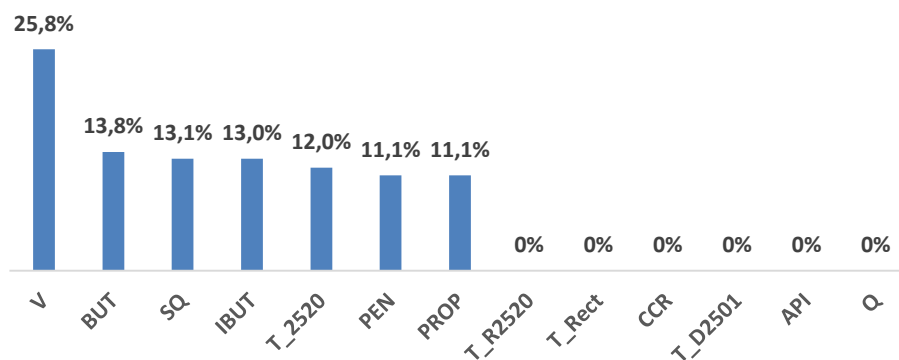


(b) RMSE vs número de factores.



(c) Coeficientes de la regresión PLS.

Figura 57. Resultados regresión PLS para las muestras de la subdivisión 2 de la clase 2.



(d) Influencia de las variables independientes en la predicción.

Figura 58. Resultados regresión PLS para las muestras de la subdivisión 2 de la clase 2 (continuación).

Tabla 6.

Resumen modelos ajustados para la clase2.

	MODELO	NÚMERO DE ENTRADAS	R2	RMSE
Todos los datos	MLR	49	0.936 (0.931)	0.194 (0.202)
	PLS	3 Factores resultantes de la combinación lineal de 6 entradas.	0.900 (0.900)	0.190 (0.191)
Subdivisión1	MLR	45	0.973 (0.963)	0.147 (0.170)
	PLS	2 Factores resultantes de la combinación lineal de 3 entradas.	0.927 (0.926)	0.200 (0.202)
Subdivisión2	MLR	7	0.944 (0.943)	0.205 (0.207)
	PLS	1 Factor resultante de la combinación lineal de 4 entradas.	0.921 (0.920)	0.194 (0.196)

Resumiendo lo obtenido para la clase 2, las regresiones MLR de las subdivisiones presentan un mejor desempeño que la regresión MLR para la clase completa; estas regresiones presentan utilidad a nivel cuantitativo, aunque su aplicación puede estar restringida por el número de entradas de cada modelo (exceptuando la regresión para la subdivisión 2). Por su parte, las regresiones PLS

obtenidas presentan una estructura con pocas entradas (ecuaciones 6-8) y con aplicación a nivel semicuantitativo.

$$\begin{aligned} \%R = & 0.0466 - 0.2464 * T_{RectA} * T_{RectB} + 0.1725 * T_{R2520A} * BUT - 0.1800 * T_{R2520B} \\ & * PROP - 0.1800 * T_{R2520B} * iBUT - 0.1954 * CCR * iBUT - 0.2513 * PROP \\ & * iBUT \end{aligned} \quad (6)$$

$$\%R = 0.0009 - 0.3069 * T_{R2520A} * S_Q - 0.4492 * V * iBUT - 0.3098 * iBUT * PROP \quad (7)$$

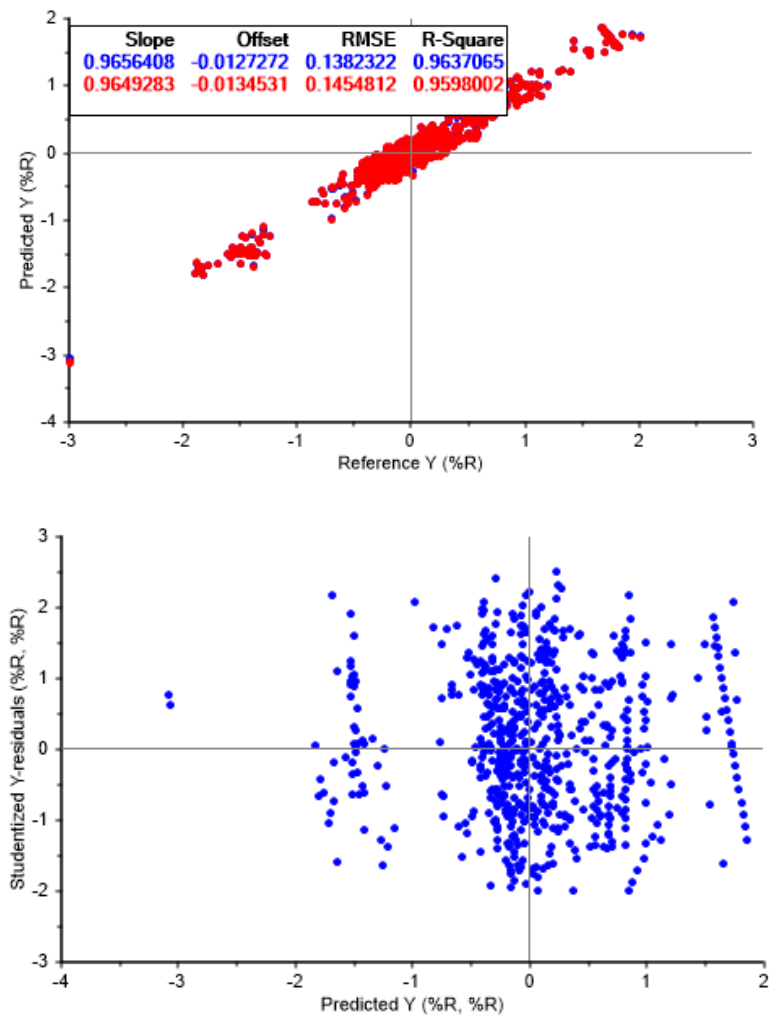
$$\begin{aligned} \%R = & 0.03759 - 0.2124 * T_{2520A} * V - 0.2462 * V * BUT - 0.1978 * PEN * PROP \\ & + 0.2320 * S_Q * iBUT \end{aligned} \quad (8)$$

4.3.4 Clase 3 El ajuste del %R por MLR para la clase 3, considerando efectos dobles de las 17 variables independientes, conduce a un modelo con 31 entradas exhibiendo una reproducción de la varianza del 96.4% y un RMSE de 0.138 (Figura 3.33a). La distribución de los residuos permite verificar los supuestos de la MLR (Figura 3.33b); aunque es posible apreciar un patrón lineal tenue en algunos residuales, la prueba de Kolmogorov-Smirnov aplicada según el código R reportó un valor p inferior a 0.05, confirmando una distribución normal, aleatoria, con media cero y con varianza constante de los residuos (Mendenhall & Sincich, 1996). En este mismo sentido, la tabla ANOVA presenta la verificación de la significancia estadística de la regresión (Figura 3.33c, *Model p value*) y la de sus variables de entrada (Figura 3.33c, *p value*). También, la Figura 3.33d muestra la relación entre los coeficientes de las entradas en la MLR obtenida, mientras que la Figura 3.33e presenta el porcentaje de contribución de cada grupo de las variables independientes a la predicción del %R. Según la Figura 3.33e, las variables de proceso exhiben la mayor

contribución (45.6%) para la predicción del rendimiento de DMO (%R); *i.e.* T_D2501, T_R2520, T_Rect y T_2520 aportan 26.9%, 11.7%, 7.0% y 0.9%, respectivamente. En segundo lugar de aporte a la predicción de DMO se encuentran las variables relacionadas con la composición del solvente (26% como contribución total, distribuido en 12.7% de BUT, 10.7% de IBUT y 2.6% de PROP). Por último, las propiedades de la carga aportan un total de 26.6% (PEN 11%, API 10.4%, V 2.7% y CCR 2.5%). Una influencia casi nula la presenta la relación solvente a carga (SQ) y una relación nula la exhibe el flujo de carga a la unidad (Q).

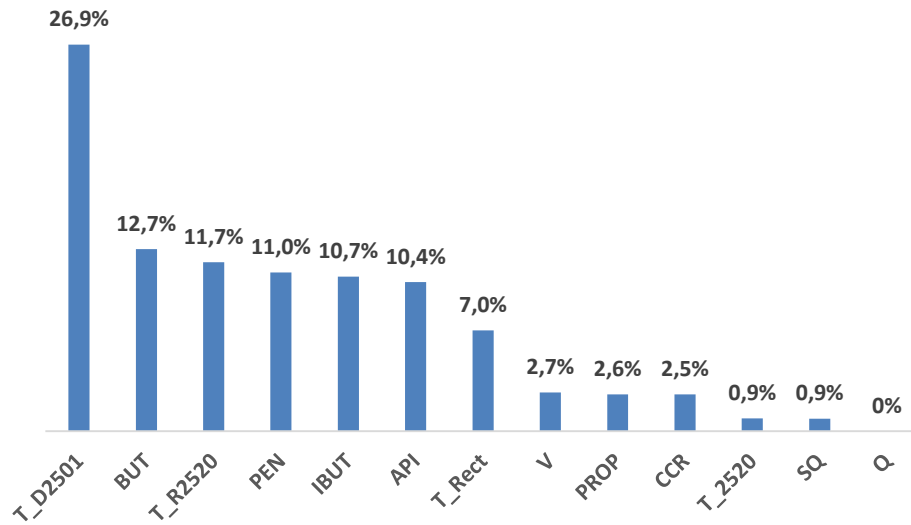
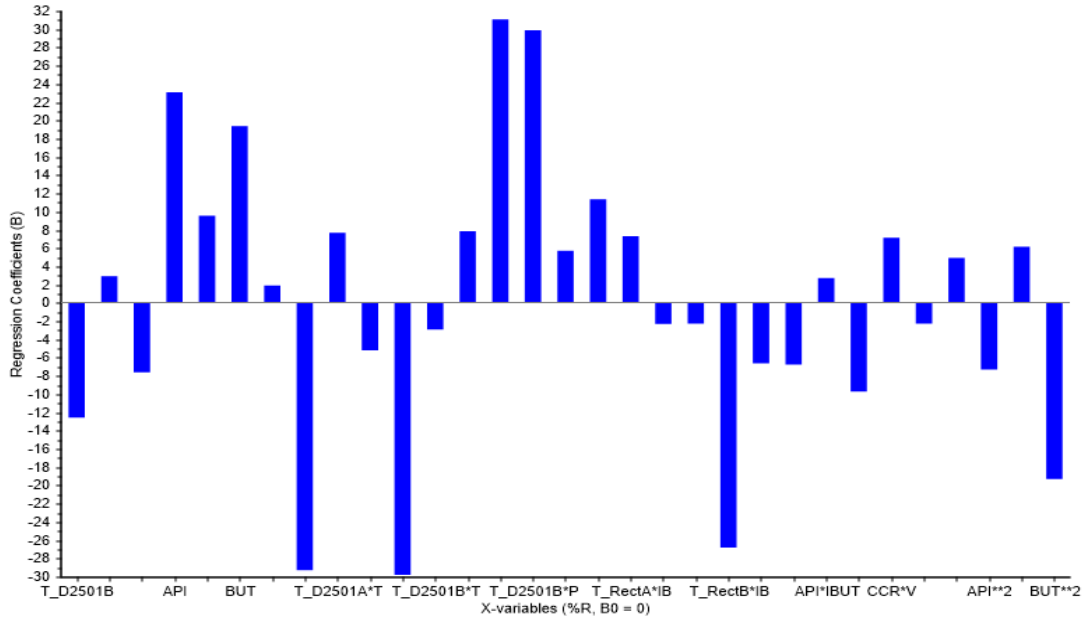
Por otra parte, la aplicación del método PLS presenta una regresión con sólo 12 entradas, producto de la combinación lineal de 4 factores, pero con una reproducción de la varianza de sólo 77% y un RMSE de 0.330 (Figura 3.34). Con esto, la mejora en las regresiones es buscada por medio de la aplicación de un PCA y una agrupación por *k-means*. Para esta clase, la subdivisión establecida por el método de agrupación *k-means* con $k=2$ conduce a los resultados mostrados en las gráficas tridimensionales de los *scores* y los *loading* para los tres primeros componentes principales (Figura 3.35). La comparación entre las Figuras 3.35a y 3.35b sugiere que las variables con mayor influencia en la formación de la subdivisión en este grupo corresponden a PROP (PC1=0.414), T_D2501A (PC1=-0.439), T_D2501B (PC1=-0.436) y T_RectB (PC1=-0.435). Según la Figura 3.36 en donde se presenta la dispersión de los datos de la clase 3 en las variables T_D2501A, T_D2501B, T_RectB y PROP, la subdivisión en esta clase es establecida en los valores límite de 208, 235 y 7.9 para estas variables, respectivamente –desde luego también se presentan zonas con datos atípicos–. La Tabla 3.6 presenta los límites para cada variable en la subdivisión de la clase 3. Según esta tabla, las variables que definen la formación de la subdivisión corresponden a T_D2501A, T_D2501B, T_RectB y PROP; La Figura 3.37 muestra que la relación entre T_D2501A y T_D2501B es lineal con una pendiente de *ca.* 1°F, mientras que la Figura 3.38

exhibe una relación no lineal entre T_RectA y T_RectB ($R=0.602$); lo anterior se debe a problemas operativos con los serpentines ubicados en el tambo de extracción D2501A, lo cual crea diferencias significativas en las temperaturas de las torres de rectificación del solvente (T_RectA y T_RectB).



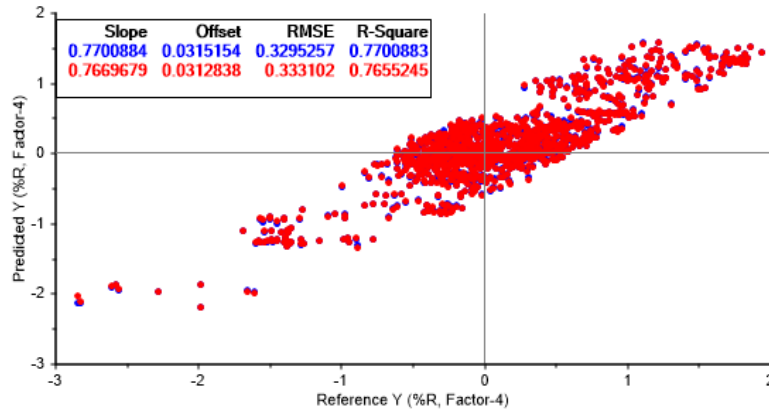
Anova Table							
Multiple Correlation: 0.9819059 (cal) 0.979948 (val)							
R-Square: 0.9637065 (cal) 0.9598002 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	332.7717	31	10.73457	534.9875	0		
Error	12.42028	619	0.020065				
Total	345.1920	650	0.531064				
Variables							
Intercept	0	0				0	0
T_D2501B	8.689617	1	8.689617	433.0852	0	-12.5591259	0.6034936
T_2520A	0.742184	1	0.742184	36.98907	2.08008188E-09	2.93733931	0.4829668
T_R2520B	2.507016	1	2.507016	124.9441	1.50128365E-26	-7.56811714	0.6770643
API	7.835206	1	7.835206	390.4934	0	23.1112347	1.1695432
IBUT	0.923974	1	0.923974	46.04875	2.70333218E-11	9.57660866	1.4112463
BUT	7.055322	1	7.055322	351.6232	0	19.3938484	1.0342491
S_Q	0.592894	1	0.592894	29.54864	7.8536516E-08	1.93590176	0.3561349
T_D2501A*T_R25	1.284985	1	1.284985	64.04131	6.02996995E-15	-29.2578506	3.6560516
T_D2501A*T_R25	2.587357	1	2.587357	128.9481	2.81440649E-27	7.70617485	0.6786273
T_D2501A*V	2.900058	1	2.900058	144.5332	4.54362082E-30	-5.1816411	0.4310060
T_D2501A*PEN	1.428999	1	1.428999	71.22003	2.26379534E-16	-29.790863	3.5300591
T_D2501B*T_252	0.749241	1	0.749241	37.34077	1.75467285E-09	-2.88895059	0.4727684
T_D2501B*T_Rect	5.316837	1	5.316837	264.9829	0	7.89369726	0.4849216
T_D2501B*T_R25	1.843535	1	1.843535	91.87866	2.20381223E-20	31.1094303	3.2455232
T_D2501B*PEN	1.440784	1	1.440784	71.80743	1.7335813E-16	29.9175167	3.5305378
T_D2501B*PROP	4.285000	1	4.285000	213.5566	9.11124262E-42	5.75185633	0.3935967
T_D2501B*BUT	2.566133	1	2.566133	127.8906	4.3754328E-27	11.385726	1.0067957
T_RectA*IBUT	0.715393	1	0.715393	35.65373	3.97309252E-09	7.34015894	1.2292859
T_RectA*S_Q	0.606204	1	0.606204	30.21195	5.66693288E-08	-2.26715302	0.4124691
T_RectB*T_R2520	0.139720	1	0.139720	6.963379	0.00852885656	-2.24799156	0.8518922
T_RectB*IBUT	8.134676	1	8.134676	405.4201	0	-26.7809124	1.3300645
API*CCR	12.73664	1	12.73664	634.7748	0	-6.6112113	0.2624044
API*PROP	5.527219	1	5.527219	275.4670	0	-6.73911572	0.4060393
API*IBUT	0.704739	1	0.704739	35.12256	5.14227194E-09	2.73870707	0.4621175

Anova Table							
API*BUT	4.813898	1	4.813898	239.9148	0	-9.67483425	0.6246187
CCR*V	11.73448	1	11.73448	584.8274	0	7.17839146	0.2968337
V*BUT	0.987042	1	0.987042	49.19223	6.09948767E-12	-2.24556875	0.3201681
T_D2501A**2	1.747803	1	1.747803	87.10870	1.81132301E-19	4.94706631	0.5300501
API**2	15.07549	1	15.07549	751.3338	0	-7.27680969	0.2654755
IBUT**2	6.423926	1	6.423926	320.1543	0	6.18262291	0.3455358
BUT**2	26.93616	1	26.93616	1342.442	0	-19.271286	0.5259721



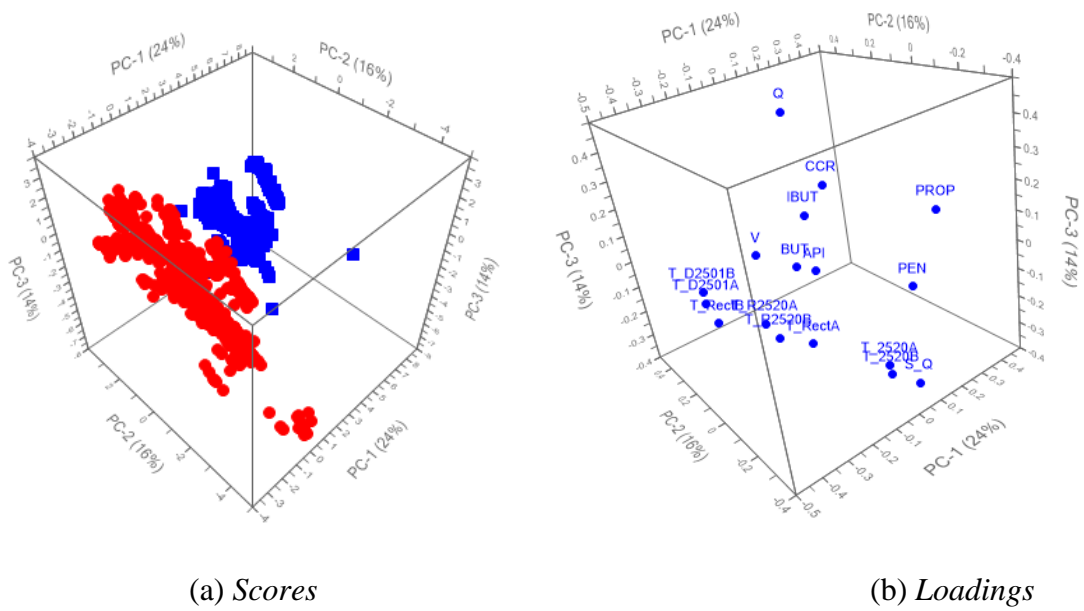
(a) Regresión MLR

Figura 59. Regresión MLR para los datos de la clase 3.



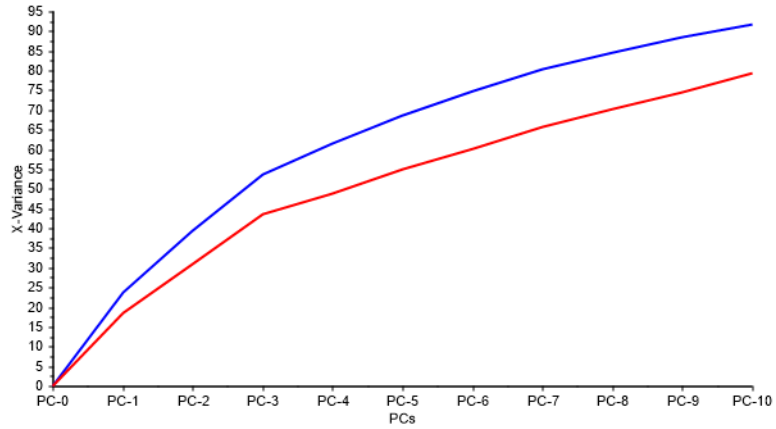
(b) Regresión PLS

Figura 60. Regresión PLS para la clase 3.



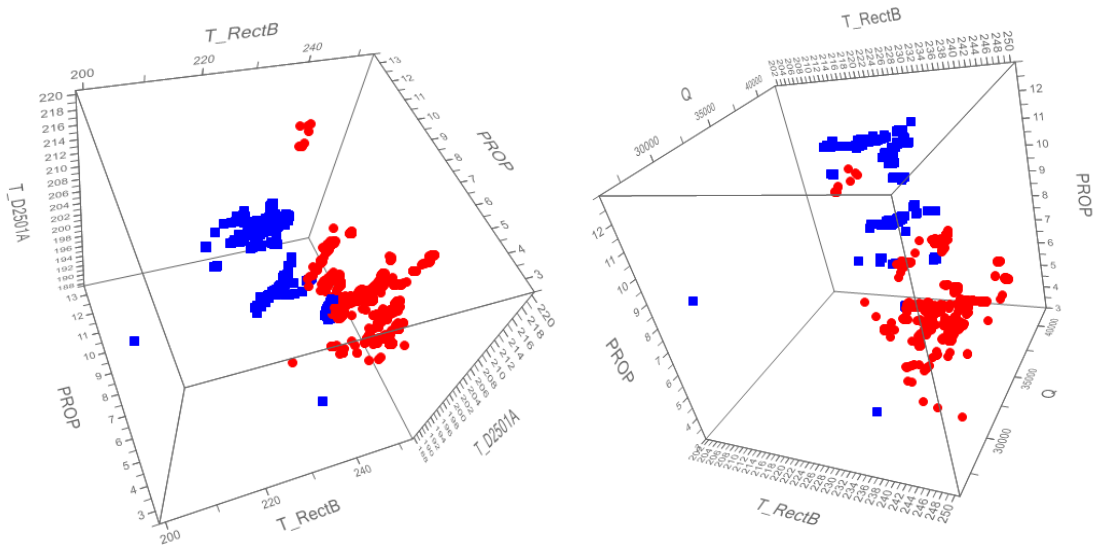
(a) Scores

(b) Loadings



(c) Reproducción de la varianza según el PCA.

Figura 61. Resultados análisis PCA con los datos de la clase 3.



(a) T_RectB vs PROP vs T_D2501A

(b) T_RectB vs PROP vs Q

Figura 62. Resultados método de agrupación k-mean con $k=2$ para la subdivisión de los datos de la clase 3.

Tabla 7.

Variables que definen la subdivisión en la clase 3.

	T_D2501A	T_D2501B	T_RectB	PROP, %v
Subdivisión 1	[198, 208]	[197, 208]	[201, 232]	[4, 7.9]
Subdivisión 2	[208, 212]	[208, 215]	[232, 251]	[7.9, 11]

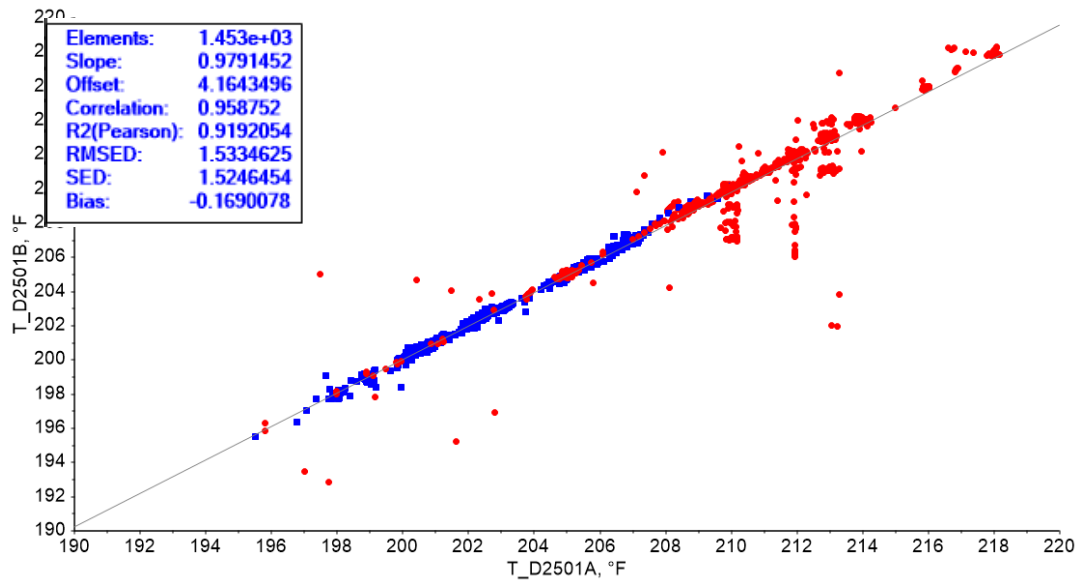


Figura 63. Variación de las temperaturas de los tambores T_D2501A-B para la subdivisión de los datos de la clase 3.

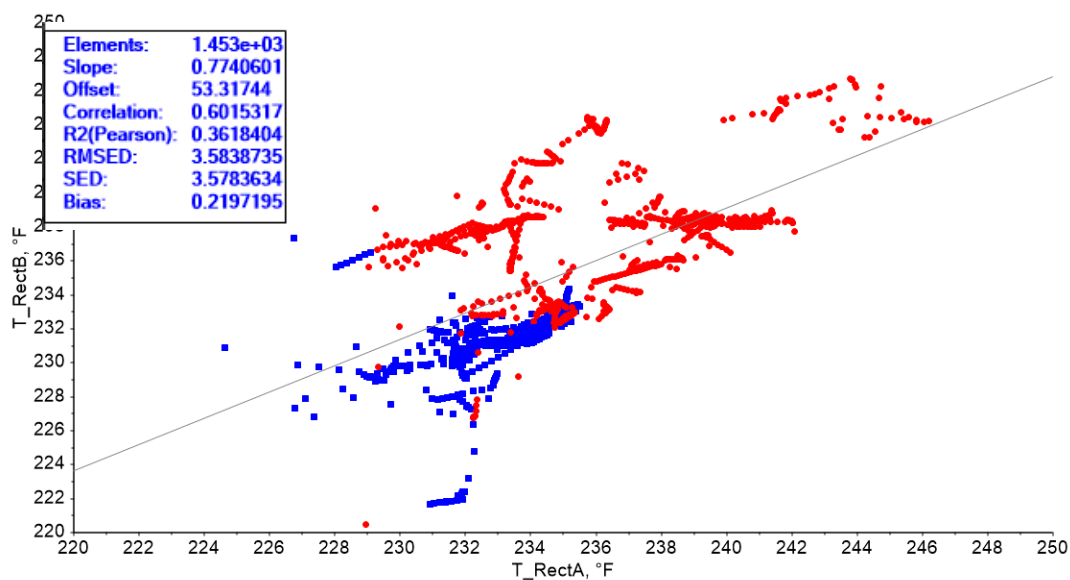


Figura 64. Variación de las temperaturas de las torres rectificadoras T_RectA-B para la subdivisión de los datos de la clase 3.

La aplicación de regresiones en la subdivisión de los datos según los resultados *k-means* con $k=2$, proporciona los desempeños mostrados en las Figuras 3.39-3.42 y en la Tabla 3.7. La regresión MLR sobre los datos de la subdivisión 1 reporta un desempeño bajo con una reproducción pobre de la varianza (67.6%) y un RMSE de 0.264 (Figura 3.39); aunque las entradas requeridas para este modelo corresponden a sólo 3 su aplicación es nula para la representación del %R (Tabla 3.7). Por otra parte, la regresión MLR sobre los datos de la subdivisión 2 presenta un desempeño suficiente para su aplicación a nivel cuantitativo, con un coeficiente de determinación de 0.954 en calibración (R2 de 0.948 en validación) y un RMSE de 0.144 (0.153 en validación) (Figura 3.40). Esta regresión requiere de sólo 22 entradas, por lo que se facilita su aplicación práctica; la Figura 3.40c presenta el análisis de varianza para esta regresión –tomado directamente de *The Unscrambler X*–, en donde se pueden detallar los coeficientes de la regresión. Según los coeficientes de la regresión MLR para la subdivisión 2 (Figura 3.40c-d), las entradas de mayor

influencia corresponden a T_RectB*PROP (25.7%) y T_D2501B*PROP (25.3%), con lo cual la variable PROP –porcentaje de propano en el solvente– presenta una incidencia de un 30% en el valor del %R (Figura 3.40e); el solvente influye en un 37% del valor del %R, mientras que las variables operativas llegan a una incidencia del 53.5% y las propiedades de la carga en sólo un 4.1% (Figura 3.40e).

En cuanto a las regresiones PLS, las Figuras 3.41-42 presentan las tendencias obtenidas para las subdivisiones respectivas. Según los datos de desempeño (Tabla 3.7), la regresión PLS para la subdivisión 1 representa sólo el 66.8% de la varianza del %R, mientras que para la subdivisión 2, la regresión PLS reproduce el 92.6% del %R. La ecuación 9 presenta la regresión obtenida para el modelo PLS en la subdivisión 2; según los coeficientes de esta ecuación, la T_RectB se encuentra relacionada con *ca.* 36% del valor de %R, mientras que V reporta una influencia de *ca.* 29%. Es importante mencionar que a pesar de la no inclusión de las interacciones dobles en el análisis por componentes principales, el PCA pudo identificar a la variable T_RectB (componente principal PC-1, 24% varianza, Figura 3.35) como una de las de mayor aporte tanto en la regresión MLR como en la PLS.

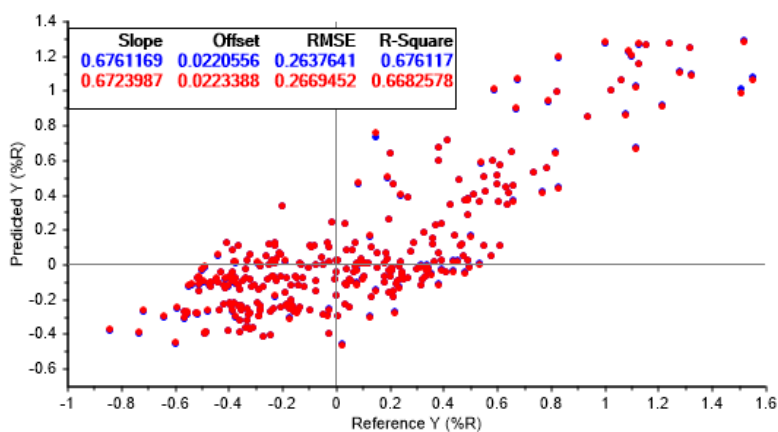
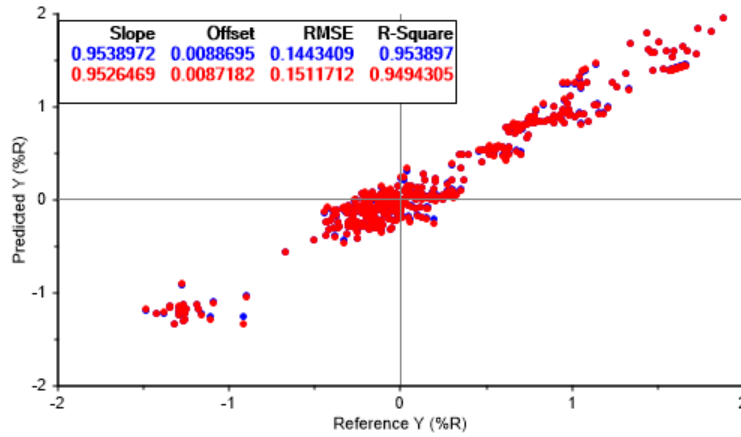
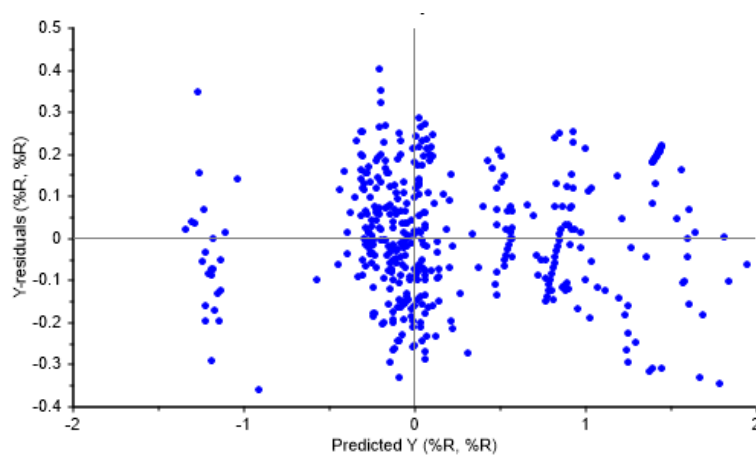


Figura 65. Regresión MLR para la subdivisión 1 de los datos de la clase 3.



(a) Valor predicho vs valor de referencia.

Figura 66. Regresión MLR para la subdivisión 2 de la clase 3.

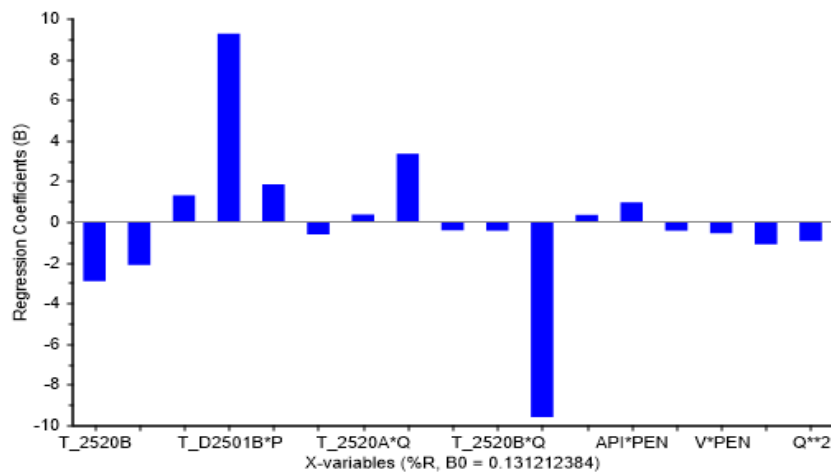


(b) Residuales estudentizados.

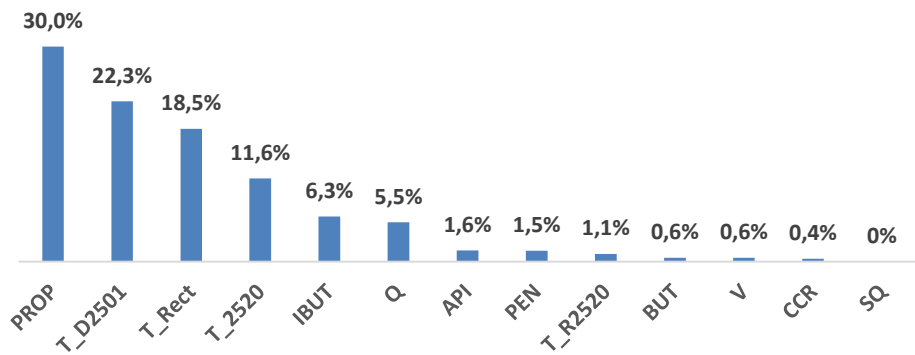
Anova Table							
Multiple Correlation: 0.9766765 (cal) 0.974393 (val)							
R-Square: 0.953897 (cal) 0.9494305 (val)							
	SS	df	MS	F ratio	p value	B-coeficien	STDerr
Summary							
Model	176.740	17	10.39647	477.100	0		
Error	8.54205	392	0.021790				
Adjusted Total	185.282	409	0.453012				
Variables							
Intercept	3.28420	1	3.284208	150.713	1.54413886E-29	0.131212384	0.01068
T_2520B	0.18385	1	0.183858	8.43730	0.0038842163	-2.9114933	1.00233
T_D2501A*IBUT	0.35278	1	0.352783	16.1894	6.87873471E-05	-2.11280465	0.52510
T_D2501B*Q	0.17644	1	0.176446	8.09722	0.00466604205	1.30429304	0.45836
T_D2501B*PROP	1.46774	1	1.467740	67.3551	3.30109365E-15	9.26028156	1.12833
T_D2501B*IBUT	0.27138	1	0.271380	12.4538	0.000466627593	1.85057151	0.52438
T_2520A*API	9.25475	1	9.254753	424.704	0	-0.60819846	0.02951
T_2520A*Q	3.81018	1	3.810189	174.852	2.8969656E-33	0.369479805	0.02794
T_2520B*T_RectB	0.24739	1	0.247398	11.3532	0.000827869982	3.36856627	0.99973
T_2520B*T_R252	1.75576	1	1.755761	80.5727	1.18076864E-17	-0.40700259	0.04534
T_2520B*Q	0.09278	1	0.092783	4.25786	0.0397268347	-0.43189901	0.20930
T_RectB*PROP	1.51551	1	1.515512	69.5473	1.2813132E-15	-9.58366966	1.14918
T_R2520B*BUT	2.41321	1	2.413218	110.743	5.62143669E-23	0.36025852	0.03423
API*PEN	7.98705	1	7.987058	366.530	0	0.967313349	0.05052
CCR*Q	2.72032	1	2.720321	124.837	2.38111625E-25	-0.42505809	0.03804
V*PEN	3.63579	1	3.635794	166.848	4.77975105E-32	-0.56038469	0.04338
T_D2501B**2	1.37948	1	1.379482	63.3053	1.9213704E-14	-1.09839582	0.13805
Q**2	0.11693	1	0.116935	5.36626	0.0210448243	-0.92977702	0.40136

(c) Tabla ANOVA.

Figura 67. Regresión MLR para la subdivisión 2 de la clase 3 (continuación).



(d) Coeficientes obtenidos en la regresión.



(e) Influencia de las variables independientes en la predicción.

Figura 68. Regresión MLR para la subdivisión 2 de la clase 3 (continuación).

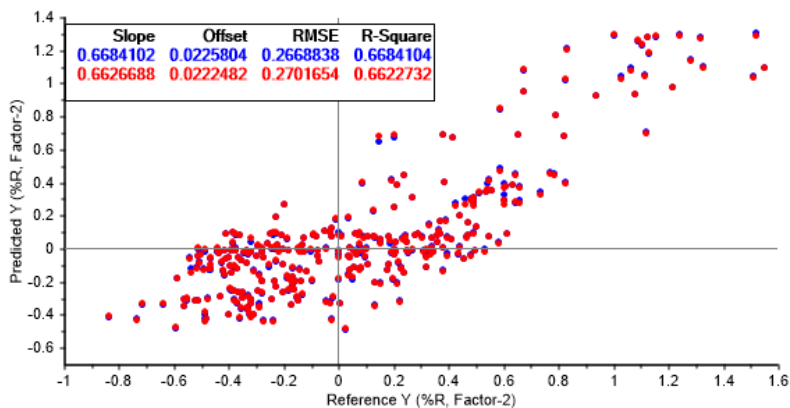


Figura 69. Regresión PLS para la subdivisión 1 de los datos de la clase 3.

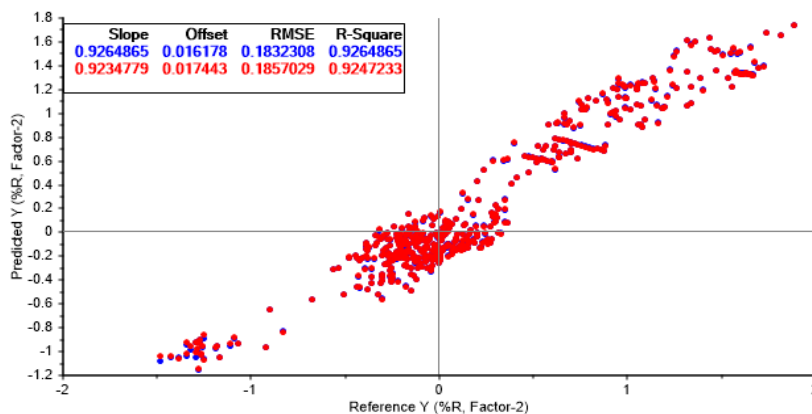


Figura 70. Regresión PLS para la subdivisión 2 de los datos de la clase 3.

$$\begin{aligned} \%R = & 0.2074 - 0.1674 * T_{D2501A} * T_{RectB} - 0.1622 * T_{D2501B} * T_{RectB} - 0.2127 * T_{2520A} * V \\ & - 0.1637 * T_{2520B} * V - 0.1440 * T_{RectB} * CCR + 0.2341 * T_{R2520B} * PEN \end{aligned} \quad (9)$$

Tabla 8.

Resumen modelos ajustados para la clase3.

	MODELO	NÚMERO DE ENTRADAS	R2	RMSE
Todos los datos	MLR	31	0.964 (0.960)	0.138 (0.145)
	PLS	4 Factores resultantes de la combinación lineal de 12 entradas.	0.770 (0.766)	0.330 (0.333)
Subdivisión1	MLR	3	0.676 (0.668)	0.264 (0.267)
	PLS	2 Factores resultantes de la combinación lineal de 7 entradas.	0.670 (0.662)	0.267 (0.270)
Subdivisión2	MLR	17	0.954 (0.950)	0.147 (0.151)
	PLS	2 Factores resultantes de la combinación lineal de 6 entradas.	0.926 (0.925)	0.183 (0.186)

Resumiendo esta sección, la metodología de subdivisión de los escenarios operativos identificados, propuesta en el presente documento, condujo a la mejor en la predicción de los rendimientos de DMO para las clases 0 y 2; la aplicación de los métodos MLR y PLS en los diferentes escenarios operativos y sus subdivisiones conllevó a la proposición y validación de regresiones con utilidad a nivel cuantitativo y semicuantitativo para la predicción de los rendimientos de DMO. Con esto, los modelos validados constituyen un apoyo a la evaluación y definición de valores operacionales tendientes a la mejora de los rendimientos de la unidad DEMEX

4.4 Simulación del proceso DEMEX

Los assays de las corrientes de crudo de carga a la refinería GRB fueron utilizados para codificar las propiedades de los fondos de vacío para la carga a la unidad DEMEX; los assays fueron suministrados por el Departamento de Planeación de la GRB. Las principales corrientes que componen la carga a la refinería corresponden a los crudos Cupiagua, HCT, Mezcla Norte, Mezcla Vasconia, Galán, Payoa, Provincia, LCT, Casabe, Teca y Rubiales. Cada crudo está compuesto por flujos de otros crudos primarios, por lo cual, para la simulación se crearon mezcladores para la representación de cada corriente principal; los crudos que componen la Mezcla Vasconia, junto con su composición volumétrica, son mostrados en la Tabla 3.8. Asimismo, la Figura 3.43 presenta algunos de los crudos codificados en la simulación con el programa HYSYS.

Tabla 9.

Crudos componentes de la Mezcla Vasconia

Crudo	Composición Vol.
OAM	0,4125
OCL	0,1125
VS1	0,0875
RUBIALES	0,3000
CUPIAGUA/FLOREÑA	0,0875

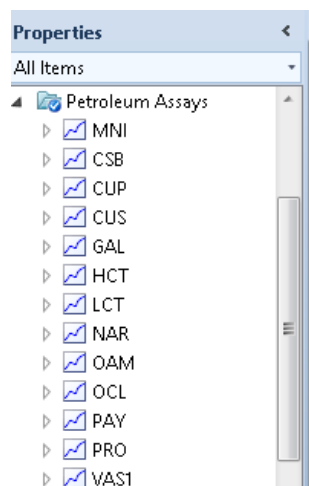


Figura 71. Algunos assays codificados en la herramienta *Petroleum Assays* de HYSYS.

Los flujos de crudos de alimentación a las unidades de destilación fueron establecidos en la simulación según la Tabla 3.9. Estos flujos fueron especificados en el objeto *Petroleum Feeder* (Figura 3.44), del cual se obtienen los correspondientes flujos de fondos de vacío por medio de la definición del corte 800°F+. Las corrientes de salida de los *Feeder* son mezcladas para representar el flujo de fondos de vacío carga a la unidad DEMEX (Corriente Fondos, Figura 3.44). De otro lado, el paquete termodinámico utilizado para la simulación del proceso DEMEX fue el denominado PSRV, el cual es una mejora de la ecuación de estado PR que incluye mezclas no ideales (Shi & Liang, 2011). Esta ecuación de estado ha sido reportada en simulaciones de fases en equilibrio de productos del petróleo en condiciones subcríticas y supercríticas (Zhao *et al.*, 2006).

Tabla 10.

Composición volumétrica de los flujos de entrada a las unidades de destilación de la GRB.

CRUDO	U200	U250	U2000	U2100
HCT	-	-	-	0.0300
Mezcla Norte	-	-	-	0.3500
Mezcla Vasconia	0.7800	0.4200	0.4300	0.3100
Galán	-	0.0400	-	0.0400
Payoa	-	-	-	0.0200
Provincia	0.0500	0.0300	0.0300	0.0200
LCT	-	0.2500	0.4400	-
Casabe	0.1700	0.1000	0.1000	0.0700
Nare	-	0.1600	-	0.1600

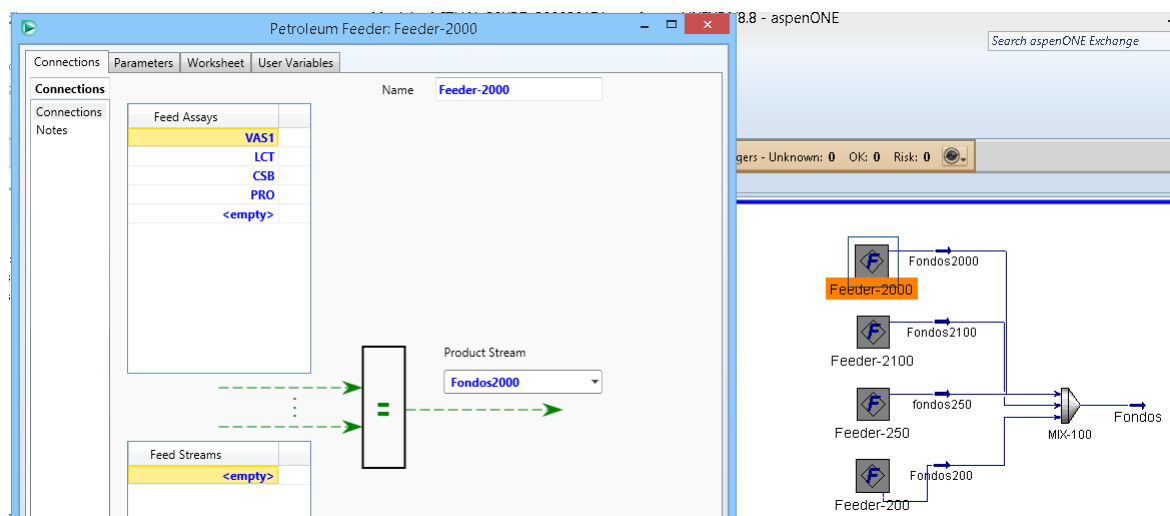
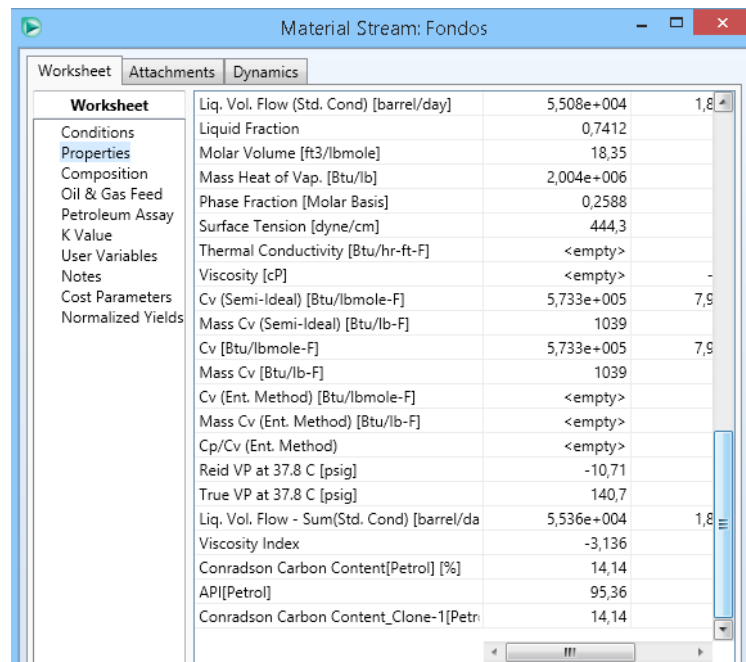


Figura 72. Ilustración del objeto *Petroleum Feeder* en la simulación realizada para el proceso DEMEX en HYSYS.

La mezcla de fondos obtenida con el anterior procedimiento reporta un contenido de carbón residual de 14.4 (Figura 3.45), muy cercano al contenido de carbón de diseño de la unidad DEMEX; la gravedad API (“API[Petrol]”, Figura 3.45) de la mezcla es de 95.36, que a condiciones estándar corresponde a 9.5, valor aproximado al utilizado para el diseño de la unidad DEMEX (API=10). De otro lado, la simulación del proceso DEMEX involucró las cuatro etapas principales:

extracción, separación del solvente del DMO, separación del solvente del asfalto y recuperación del solvente separado de las corrientes de proceso.

El módulo de extracción (Figura 3.46) fue simulado utilizando datos de diseño de la unidad y parámetros heurísticos de la operación, así como, intervalos de operación según la sección anterior. Asimismo, para la separación de solvente contenido en el DMO se adicionaron los equipos de calentamiento (E2503), la torre flash T2501 y la torre despojadora con vapor T2502 (Figura 3.47). También, para la etapa de separación del solvente contenido en el asfalto, se adicionaron los hornos H2501, el tambor separador y la torre despojadora con vapor T2522 (Figura 3.48). Asimismo, para la simulación de la recuperación del solvente se incluyen condensadores de solvente (E2505, E2507), sub enfriadores de solvente (E2505), un compresor (C2501), tambores acumuladores de solvente (D2504, D2505) y una bomba de recuperación de solvente de baja presión (P2502) (Figura 3.49).



Worksheet	Attachments	Dynamics
Worksheet	Liq. Vol. Flow (Std. Cond) [barrel/day]	5,508e+004 1,8
Conditions	Liquid Fraction	0,7412
Properties	Molar Volume [ft ³ /lbmole]	18,35
Composition	Mass Heat of Vap. [Btu/lb]	2,004e+006
Oil & Gas Feed	Phase Fraction [Molar Basis]	0,2588
Petroleum Assay	Surface Tension [dyne/cm]	444,3
K Value	Thermal Conductivity [Btu/hr-ft-F]	<empty>
User Variables	Viscosity [cP]	<empty>
Notes	Cv (Semi-Ideal) [Btu/lbmole-F]	5,733e+005 7,9
Cost Parameters	Mass Cv (Semi-Ideal) [Btu/lb-F]	1039
Normalized Yields	Cv [Btu/lbmole-F]	5,733e+005 7,9
	Mass Cv [Btu/lb-F]	1039
	Cv (Ent. Method) [Btu/lbmole-F]	<empty>
	Mass Cv (Ent. Method) [Btu/lb-F]	<empty>
	Cp/Cv (Ent. Method)	<empty>
	Reid VP at 37.8 C [psig]	-10,71
	True VP at 37.8 C [psig]	140,7
	Liq. Vol. Flow - Sum(Std. Cond) [barrel/day]	5,536e+004 1,8
	Viscosity Index	-3,136
	Conradson Carbon Content[Petrol] [%]	14,14
	API[Petrol]	95,36
	Conradson Carbon Content_Clone-1[Petr	14,14

Figura 73. Propiedades reportadas por la simulación para la mezcla de alimentación a DEMEX.

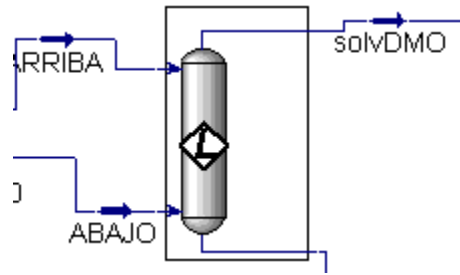


Figura 74. Equipo de extracción con solvente representando a la sección respectiva de DEMEX.

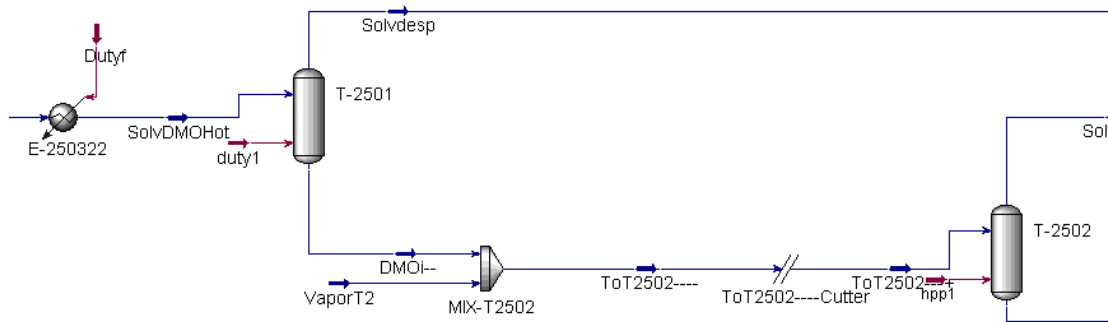


Figura 75. Equipos de separación del solvente del DMO.

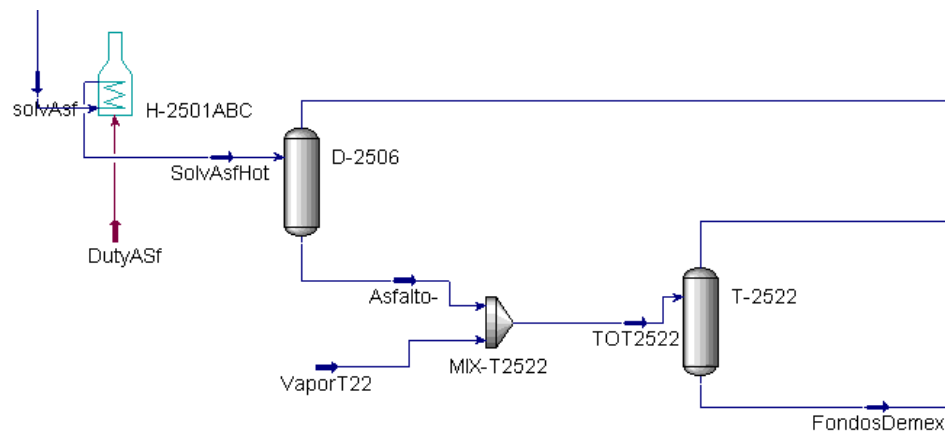


Figura 76. Equipos para la separación del solvente del fondo DEMEX.

Por otra parte, la corriente de solvente fue definida según el esquema actual de operación en la GRB (Columna “Actual”, Tabla 3.10); asimismo, diferentes composiciones referenciadas en la

literatura (Carrillo *et al.*, 1996; Cao *et al.*, 2010) fueron establecidas para determinar el impacto del cambio en la composición del solvente sobre el rendimiento de la unidad DEMEX (Tabla 3.10). La temperatura crítica para cada mezcla (Tabla 3.11) fue determinada para establecer el intervalo de temperatura de extracción. La Tabla 3.11 presenta los valores para las temperaturas de máxima extracción (T_{mex}), media (T_{media}), máxima (T_{max}) y crítica ($T_{critica}$) según la mezcla definida (Tabla 3.10); por recomendación consignada en el manual del fabricante de la tecnología DEMEX, la T_{max} debe ubicarse 50°F por debajo de la $T_{critica}$. El intervalo de operación debe fijarse entre la temperatura máxima T_{max} y la temperatura donde se encuentra la mayor capacidad de extracción de la mezcla de solvente; según el manual de operación de DEMEX, la T_{mex} se encuentra *ca.* 60°F por debajo de la T_{max} . Tomando en consideración los anteriores criterios, el intervalo de operación para la temperatura de la mezcla actual de solvente corresponde a $[180-240]^{\circ}\text{F}$; este intervalo coincide con los reportes de operación de la base de datos histórica (sección anterior).

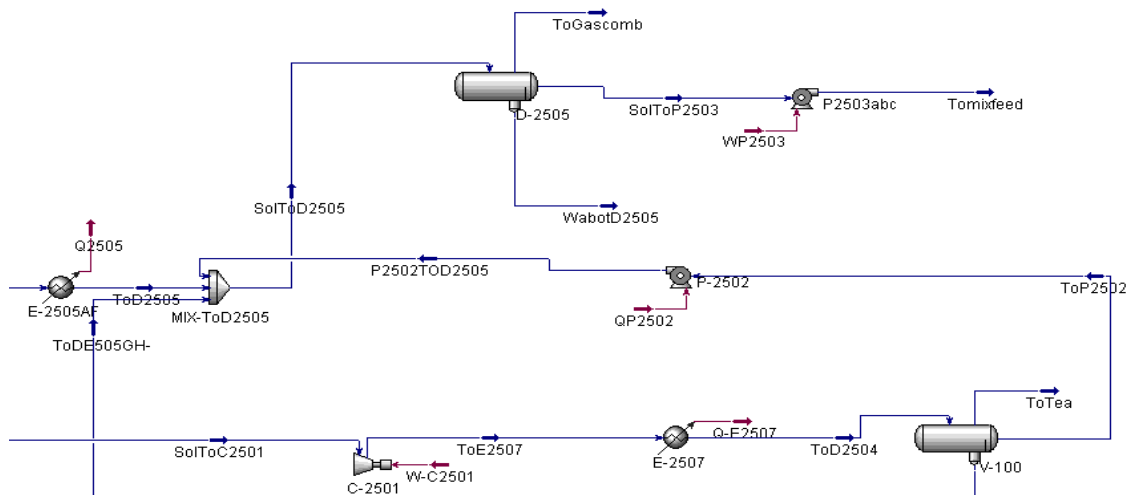


Figura 77. Equipos de recuperación del solvente.

Tabla 11.

Composición volumétrica para los diferentes solventes probados en la simulación.

Componente	Mezclas para Simulación										
	Actual	Propano	Propileno	UOP	Demex	n- Butano	i- Butano	I	II	III	Diseño
Etano	0,28	6,50									
Etileno	0,02										
Propano	3,58	85,1	37,20	0,10	6,90			1,57	11,59		40
Propileno	0,14		62,20	0,10	6,60			0,56	0,56	0,56	
iso-Butano	37,41	3,20	0,60	25,50	31,30	12,20	86,50	6,22		1,69	
n-Butano	57,93	5,20		8,70	43,40	80,40	13,50	87,21	85,3	85,30	60
1-Buteno	0,13							0,42	0,42	0,42	
Isobuteno	0,02							0,44	0,44	0,44	
butileno	0,19			65,60	11,80	7,40					
Trans-2-buteno	0,14							0,95	0,95	0,95	
Cis-2-buteno	0,06							0,74	0,74	0,74	
C5+	0,10							1,89		9,90	

Tabla 12.

Temperatura media para la operación de la extracción con cada mezcla de solvente.

Mezcla	T _{mex} , °F	T _{media} , °F	T _{max} , °F	T _{crítica} , °F
Actual	180	210	240	290
Propano	96	126	156	206
Propileno	91	121	151	201
UOP	181	211	241	291
Demex	171	201	231	281
n-Butano	191	221	251	301
i-Butano	169	199	229	279
I	193	223	253	303
II	184	214	244	294
III	203	233	263	313
Diseño	156	186	216	266

Las simulaciones consideraron el valor de la relación volumétrica solvente a carga de 5.3; este valor de la relación solvente a carga es una de las más utilizadas en la operación de la unidad

DEMEX; para la mezcla actual de solvente, una relación de 5.3 equivale a *ca.* 34.5 kgmol/h de solvente por cada kgmol/h de fondo de vacío. Con esta relación de solvente a carga, diferentes simulaciones fueron efectuadas para determinar las características de la extracción en el valor de la T_{media} ; la Tabla 3.12 muestra los resultados obtenidos con las simulaciones para cada uno de los solventes especificados (Tabla 3.10). Según la Tabla 3.12, los resultados obtenidos por las simulaciones sugieren que la mezcla actual de solvente utilizada en DEMEX (columna “Actual”) exhibe rendimientos inferiores a los nombrados como UOP, n-Butano, I y III. Sin embargo, el solvente “Actual” presenta menor contenido de CCR y menor pérdida de solvente, con lo cual, la operación actual de la unidad corresponde a un punto de beneficio para la refinería, al generar flujos considerables de DMO que pueden ser llevados a las unidades de craqueo catalítico. Este solvente, inclusive es mejor que el especificado por los fabricantes de la tecnología DEMEX (columna “Diseño”, Tabla 3.12). Los resultados obtenidos con el solvente “Actual” son similares al solvente II (Cao *et al.*, 2010), el cual utiliza menor cantidad de componentes, pero genera mayor cantidad de evaporados.

De igual forma, los resultados concuerdan con la literatura en que el solvente denominado DEMEX tiene un mayor rendimiento y un menor residuo de carbón a condición de temperatura de extracción en el punto medio de la curva de rendimiento vs temperatura (Figura 3.14). Sin embargo, la operación en esta temperatura reporta una mayor pérdida de solvente y por lo tanto la necesidad de una mayor restitución del mismo (Vera *et al.*, 1995).

También, las mezclas denominadas como Propano y Propileno presentan composición apreciable de propano y propileno (Tabla 3.10), los cuales a pesar de generar un extracto de excelente calidad (bajo CCR, Tabla 3.12) no cumplen con los rendimientos esperados para la planta Demex; además, la temperatura de operación requerida está por debajo del intervalo

operacional posible en las instalaciones de la planta. Por su parte, las mezclas denominadas UOP y III reportan rendimientos superiores a las demás mezclas, sin embargo, no cumplen con la calidad necesaria (alto CCR, Tabla 3.12) para el suministro de carga a las unidades de craqueo catalítico fluidizado; adicionalmente, estas mezclas presentan pérdidas de solvente elevadas, con costos apreciables y riesgos para la integridad de las instalaciones.

Tabla 13.

Resultados de las simulaciones del proceso DEMEX a la T_{media} , según el solvente.

Componente	Mezclas utilizadas en las simulaciones										
	Actual	Propano	Propileno	UOP	Demex	n- Butano	i- Butano	I	II	III	Diseño
T_{media} , °F	210	126	121	211	201	221	199	223	214	233	186
DMO, BPD	12900	4741	4495	19560	12950	15240	11760	14560	13160	16040	9770
%Rendimiento	37.9	13.94	13.22	57.5	38.0	44.8	34.58	42.81	38.69	47.16	28.0
%Carbón residuo (CCR)	5.45	2.46	2.46	7.32	5.29	6.01	5.24	5.93	5.42	6.55	4.33
Pérdida de solvente, BPD	89	31	26	326	107	127	72	107	97	147	60

Con lo anterior, los resultados de las simulaciones en HYSYS concuerdan con diferentes tendencias exhibidas en literatura y reportadas por la unidad DEMEX. Esta simulación corresponde a una herramienta de ayuda para la definición de una mezcla de solvente adecuada según la carga y las condiciones operacionales definidas en la unidad; una validación de mayor rigurosidad es requerida para ajustar diferentes detalles y validar la simulación en una mayor superficie operacional. Asimismo, sus resultados pueden complementar los análisis realizados utilizando los datos históricos y la estadística (presentados en la sección anterior), lo cual corresponde a un futuro trabajo en la línea de la extracción con solvente de fondos de vacío.

5. Contribuciones del proyecto de aplicación

El desarrollo del presente trabajo de aplicación final ha contribuido en la definición de cuatro escenarios operativos representativos del rendimiento para la unidad DEMEX de la GRB. Para cada escenario operativo diferentes regresiones fueron validadas para la predicción del rendimiento de DMO con aplicación a nivel cuantitativo y semicuantitativo. Es importante mencionar que no se contaba con regresiones basadas en datos históricos de proceso industrial para la predicción del rendimiento de DMO en la unidad. Estas regresiones constituyen un apoyo a la operación diaria de la unidad DEMEX, al suministrar un punto de comparación a los ingenieros de los cuartos de control en caso de eventualidades operacionales que requieran la definición de las condiciones operacionales en un tiempo limitado.

Asimismo, la simulación desarrollada corresponde a una herramienta para el análisis de las condiciones operacionales y su influencia según la mezcla de solvente utilizada en la unidad DEMEX. Los resultados de la simulación mostraron que la mezcla actual de solvente utilizada conlleva a uno de los mejores rendimientos según diferentes solventes utilizados en pruebas piloto y en literatura. Con esta simulación es posible profundizar en los resultados obtenidos con el análisis estadístico de los cuatro escenarios operativos detectados.

Por último, los resultados derivados del desarrollo del presente documento, corresponden a un apoyo para la evaluación y la definición de las entradas requeridas en el esquema de control avanzado DMC de la unidad DEMEX.

6. Conclusiones

El análisis de los datos históricos utilizando los métodos PCA y *k-means* condujo a la definición de cuatro escenarios operativos representativos para la unidad DEMEX durante la ventana de operación desde enero de 2014 hasta diciembre de 2016; este número de escenarios operativos fue validado considerando la cohesión y la separación entre las agrupaciones, así como, por el cálculo del coeficiente silueta (*silhouette coefficient*). Los escenarios operativos se distinguen unos de otros principalmente por los valores de las variables T_2520, T_R2520, T_D2501, Q y BUT; con esto, el escenario operativo denominado como clase 0 presenta valores altos de Q y valores bajos de temperaturas y BUT, mientras el escenario operativo clase 1 exhibe valores altos en las anteriores variables. Por su parte, la clase 2 se encuentra definida por valores bajos de Q y valores altos de temperaturas y BUT, mientras que la clase 3 se encuentra definida por valores bajos en las mencionadas variables. Por otra parte, no se presenta una relación aparente entre los escenarios operacionales y los rendimientos a DMO y su contenido de CCR.

La metodología de subdivisión de los escenarios operativos identificados, propuesta en el presente documento, condujo a la predicción de los rendimientos de DMO; la aplicación de los métodos MLR y PLS en los diferentes escenarios operativos y sus subdivisiones conllevó a la proposición y validación de regresiones con utilidad a nivel cuantitativo y semicuantitativo para la predicción de los rendimientos de DMO. Para la clase 0, las regresiones MLR en las dos subdivisiones reportaron, en promedio, predicciones de la varianza de *ca.* 95% y valores de RMSE de *ca.* 0.2; sin embargo, el número de entradas requeridas para estas regresiones es elevada, con

lo cual su aplicación puede estar restringida en situaciones prácticas. Respecto a la clase 1, las regresiones con todos los datos reportaron mejores desempeños que los obtenidos por las regresiones en las dos subdivisiones del escenario operativo; la regresión PLS con todos los datos presenta un desempeño a nivel semicuantitativo con una reproducción de la varianza de 91% y un RMSE de 0.23. Por su parte, en la clase 2, las regresiones para las subdivisiones reportan mejores desempeños que las regresiones aplicadas a todos los datos del escenario operativo; las regresiones MLR de las subdivisiones reportan, en promedio, reproducción de la varianza de *ca.* 96% y RMSE de *ca.* 0.18, soportando su aplicación a nivel cuantitativo. En lo referente a la clase 3, la regresión MLR con todos los datos presentó utilidad a nivel cuantitativo al reproducir una varianza de 96.4% y un RMSE de 0.138. Con estos resultados, los modelos validados constituyen un apoyo a la evaluación y definición de valores operacionales tendientes a la mejora de los rendimientos de la unidad DEMEX.

De otro lado, los resultados de las simulaciones del proceso DEMEX en HYSYS concuerdan con diferentes tendencias exhibidas en literatura y reportadas por los datos históricos. Esta simulación corresponde a una herramienta de ayuda para la definición de una mezcla de solvente adecuada según la carga y las condiciones operacionales definidas en la unidad; los resultados obtenidos por las simulaciones sugieren que la mezcla actual de solvente utilizada en DEMEX exhibe valores adecuados para el rendimiento, el contenido de CCR y la pérdida de solvente, con lo cual, la operación actual de la unidad corresponde a un punto de beneficio para la refinería; la operación actual de DEMEX genera flujos considerables de DMO que pueden ser llevados a las unidades de craqueo catalítico. El solvente utilizado actualmente, inclusive es mejor que el especificado por los fabricantes de la tecnología DEMEX.

7. Recomendaciones

Efectuar validación de las regresiones propuestas en el presente documento utilizando datos de proceso por fuera de la ventana de operación considerada.

Aplicar una validación de mayor rigurosidad a los resultados de las simulaciones del proceso DEMEX, con lo cual es posible el ajuste de diferentes detalles y la extensión del análisis en una mayor superficie operacional.

Referencias Bibliográficas

- Ahn, S., Shin, S., Im, S.I., Lee, K.B., Nho, N.S. 2015. Solvent recovery in solvent deasphalting process for economical vacuum residue upgrading. *Korean J. Chem. Eng.* DOI: 10.1007/s11814-015-0146-3.
- Bradley Worley, S.H. 2012. Utilities for quantifying separation in PCA/PLS-DA scores plots. *Analytical Biochemistry*, 433(2):102-104.[17,18].
- Brereton, R. G. 2002. *Chemometrics: Data analysis for the laboratory and Chemical Plant. 1st edition.* USA: John Wiley & Sons, Ltd.
- Campos, M. M., Martinez, R. D., Pantoja, E. F., Rueda, L. P. 1995. *SELECCION DE SOLVENTES ALTERNOS PARA EL PROCESO DEMEX.* Bucaramanga: TESIS (INGENIERA QUIMICA) - UIS. ESCUELA DE INGENIERIA QUIMICA.
- Cao, F., Jiang, D., Li, W., Du, P., Yang, G., Ying, W. 2010. Process analysis of the extract unit of vacuum residue through mixed C4 solvent for deasphalting. *Chemical Engineering and Processing*, 49, 91–96.
- Cuadras, C. M. 2007. *NUEVOS MÉTODOS DE ANÁLISIS.* Barcelona, España: CMC Editions.
- Díaz, L. G., & Morales, M. A. 2012. *Estadística multivariada Tercera edición.* Bogotá, Colombia: Editorial Universidad Nacional de Colombia.
- Ecopetrol S.A Base de datos PI. (365 de Enero-Diciembre de 2014-2016). Datos de variables planta Demex. Barrancabermeja, Santander, Colombia.
- Fahai, C., Dan, J., Wudong, L., Ping'an, D., Guizhong, Y., & Weiyong, Y. 2010. Process analysis of the extract unit of vacuum residue through mixed. *Chemical Engineering and Processing* 49, 91–96.

- Grupo de excelencia operacional Ecopetrol S.A. 2007. *Manual de la Planta Demex ECOGCB-DMX-MDPU-CA02*. Barrancabermeja: Ecopetrol S.A.
- Long, J. B. S. 2011. Novel Solvent Deasphalting Process by Vacuum Residue Blending. *Industrial & Engineering Chemistry Research*, 11259–11269.
- Lee, J. M. 2014. Separation of solvent and deasphalted oil for solvent deasphalting process. *Fuel Processing Technology* 119, 204-210.
- L.C., C., J.A.D, M., & J., A. 2012. Combined process schemes for upgrading of heavy petroleum. *Fuel*, 2-18.
- Leon, A. Y., Grosso, J. L., Parra, M. J. 2009. *Estimacion de rendimiento del desasfaltado de fondos de vacio basado en la ecuacion de estado de Peng- Robinson y la tecnica de quimiometria*. Bucaramanga: CONVENIO UIS - ICP 002 DE 2006 TESIS (MAGISTER EN INGENIERIA QUIMICA) - UIS. ESCUELA DE INGENIERIA QUIMICA.
- Magodemov, R.N., Pripakhaylo, A.V., Maryitania. 2017. Solvent demetallization of heavy petroleum feedstock usingsupercritical carbon dioxide with modifiers. *J. of Supercritical Fluids* 119, 150–158.
- Mendenhall, W. and Sincich, T. 1996. *A Second Course in Statistics. Regression Analysis*. Prentice Hall, Fifth Edition, USA.
- Montaña, D. D. 2016. Beneficio economico de incremento en rendimiento de DMO. Barrancabermeja, Santander, Colombia.
- Nuncira, S. H. 2014. *Desarrollo de correlaciones para la extensión de curvas de propiedades de “assays” de crudos Colombianos*. Barrancabermeja: Universidad de los Andes - Ecopetrol.
- Ping-An, D. b., Man-Nian, R. b., Dan, J. a., Zhen-Hui, Z. b., & Wei-Yong, Y. A. 2012. An experimental study on optimization utilization of deasphalted oil. *Fuel Processing Technology* 99, 64–68.
- Rojas, E. C., Alvarez, M. D., Russi, J. A., Afanasieva, N. N., Grosso, J. L., Navarro, U. D. 2002. *ESTUDIO DEL EFECTO DE LA ADICION DE CORRIENTES INTERMEDIAS AL*

FONDO DE VACIO EN EL PROCESO DEMEX. Bucaramanga: TESIS (INGENIERO QUIMICO) - UIS. ESCUELA DE INGENIERIA QUIMICA.

Rousseeuw, P.J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.

Shi, L. and Liang, X. 2011. An Object-Oriented Approach for Structure Design of Property Calculation Programs Using Equations of State. *Ind. Eng. Chem. Res.*, 50, 6404–6412.

Speight, J.G. 2006. *The Chemistry and Technology of Petroleum*, fourth ed., CRCPress, Boca Raton, Florida.

Terrádez, M. 2017. *Universitat Oberta de catalunya*. Recuperado el 24 de 11 de 2016, de Universidad abierta de cataluña: www.uoc.edu/in3/emath/docs/Componentes_principales.pdf

Zhao, S., Wang, R., Lin, S. 2006. High-pressure phase behavior and equilibria for Chinese petroleum residua and light hydrocarbon systems. *Pet. Sci. Technol.* 24, 297–318.

Apéndice

Apéndice A. Procedimiento para la caracterización de un ASSAY en HYSYS

TIPO DE CARACTERIZACIÓN

El tipo de caracterización se establece con base en las necesidades originadas a partir de su uso. En el Instituto Colombiano del Petróleo por ejemplo, se manejan tres tipos de caracterización de crudos que dependiendo de la profundidad de estudio se clasifican en:

Caracterización tipo 1. Es el más básico y con menos número de pruebas que se realizan al crudo. En este tipo de pruebas no se realiza caracterización a las fracciones del crudo.

Caracterización tipo 2. Se realiza al crudo como también a algunos de sus derivados. Para este análisis las muestras que tengan un contenido de agua superior al 0.5% requieren deshidratación del crudo.

Caracterización tipo 3. En este ensayo se hacen pruebas para las fracciones del crudo y la información es más detallada y precisa.

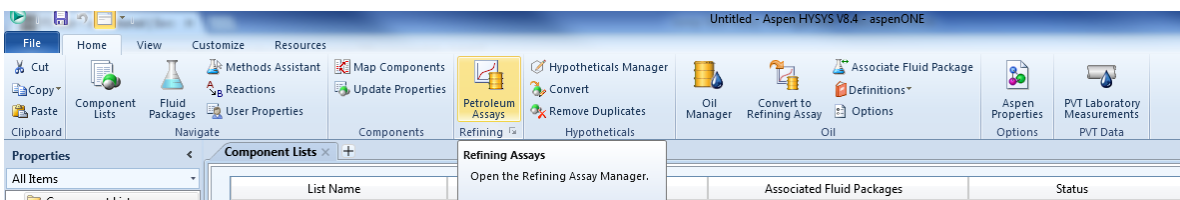
CONFIGURACIÓN DE LOS ENSAYOS

Cuando se configura un ensayo en Aspen HYSYS Petroleum Refining, primero hay que seleccionar y caracterizar los ensayos. Esto se hace desde el ambiente de Propiedades de HYSYS y se usa para poder evaluar las propiedades de los crudos antes de que sean usados en el ambiente de simulación. Si un usuario no ha definido el paquete de fluidos o la lista de componentes al

momento de crear el ensayo, HYSYS va a crear de forma automática una lista de componentes hipotéticos con Peng-Robinson como paquete de fluidos para ser usados en la simulación.

Agregar un nuevo ensayo

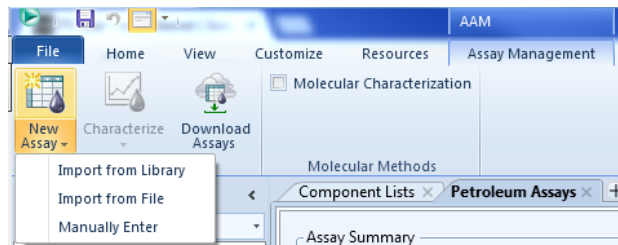
1. Comience con un caso nuevo de HYSYS haciendo click en File | New | Case
2. Para importar uno de los ensayos de petróleo, se debe utilizar el botón de “Petroleum Assays” disponible en la cintilla en el ambiente de propiedades.



Al hacer click en esta opción se abre la sección de HYSYS Petroleum Refining Assay Summary donde se pueden ver y editar los ensayos con los que se va a trabajar en la simulación. Al abrir esta página, se habilita el ribbon de Assay Management. En este nuevo menú, se encuentra el botón de New Assay. Haciendo click en esta opción permite las siguientes opciones:

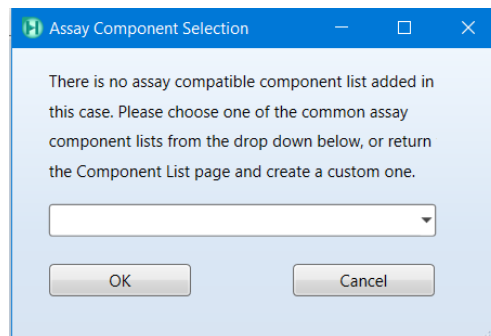
- Import from the Library – Selecciona un ensayo disponible en la librería de ensayos.
- Import from File – Importa un archivo soportado de Excel, CSV, o .AFAM que contenga información del ensayo.
- Manually enter – Introduce datos de forma manual en la tabla del ensayo.

3. Haga click en el botón de New Assay. Por defecto, e elige la opción para importar desde la librería de ensayos.

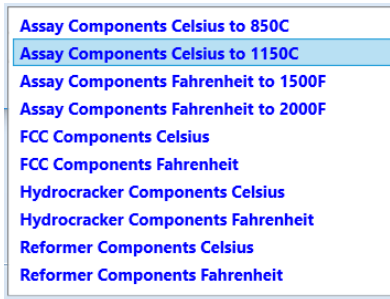


Como no existe un paquete termodinámico definido con una lista de componentes, HYSYS da la opción de elegir a partir de una lista predeterminada de componentes hipotéticos que se van a utilizar para representar el crudo.

Los componentes utilizados, a diferencia de aquellos obtenidos por medio del Oil Manager no tienen propiedades fijas, simplemente sirven para poder dividir componentes por puntos de ebullición. Y sus propiedades, tales como % de azufre, propiedades frías, contaminantes, etc., van a variar dependiendo de los crudos caracterizados y las mezclas hechas.



4. De las opciones disponibles, elija Assay Components Celsius to 1150 C para que se utilice la plantilla para componentes más pesados en grados Celsius.



5. Haga click en OK

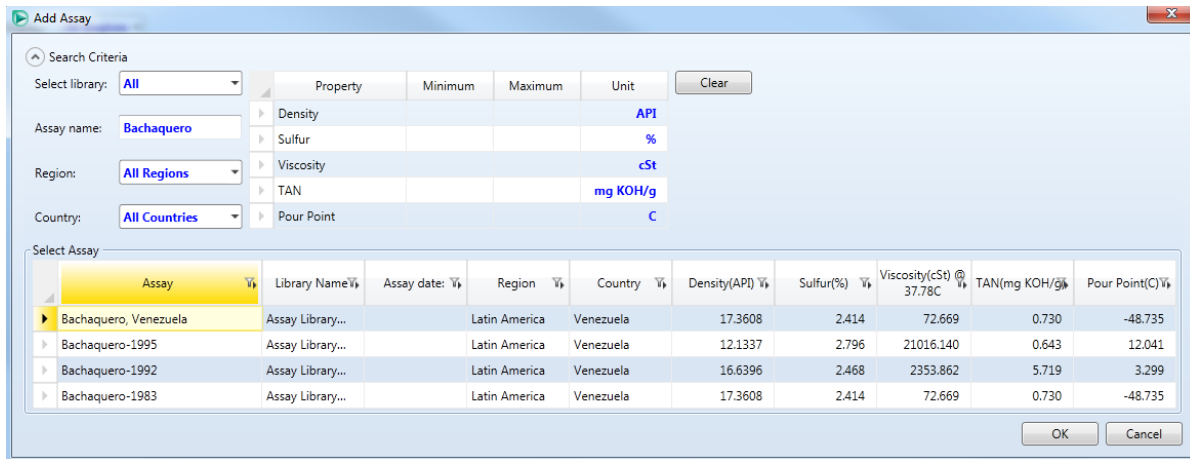
Una vez que la librería de ensayos se ha abierto, existe una larga lista de ensayos para seleccionar, y una gran variedad de opciones para filtrar los resultados. En este ejemplo, se va a importar un ensayo de la librería llamado Bachaquero. Este es un crudo pesado de Venezuela.

6. Para encontrar este crudo en específico, escriba “Bachaquero” dentro de la sección de “Assay Name”.

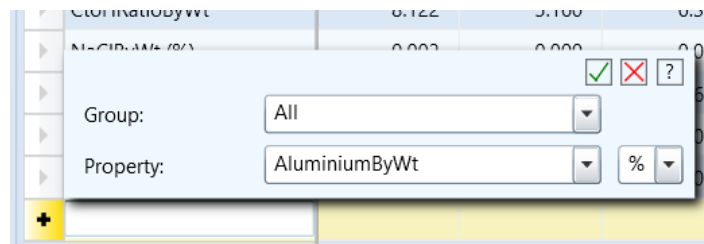
7. Seleccione el crudo más pesado de la lista de ensayos (o haga doble click).

8. Para finalizar la selección, haga click en OK en el botón de la parte inferior derecha.

Después de que el ensayo haya sido importado en la simulación, se pueden verificar las propiedades del ensayo en la sección de Input Assay y Conventional Results desde el Panel de Navegación.



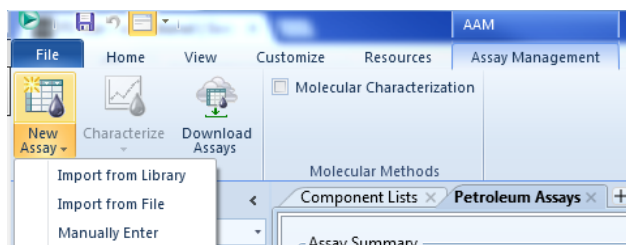
Se pueden adicionar más columnas para revisar propiedades globales del crudo, o propiedades para cortes a temperaturas específicas. De igual manera, se pueden adicionar filas de propiedades para evaluar datos adicionales.



CREACIÓN DE ENSAYOS DE FORMA MANUAL

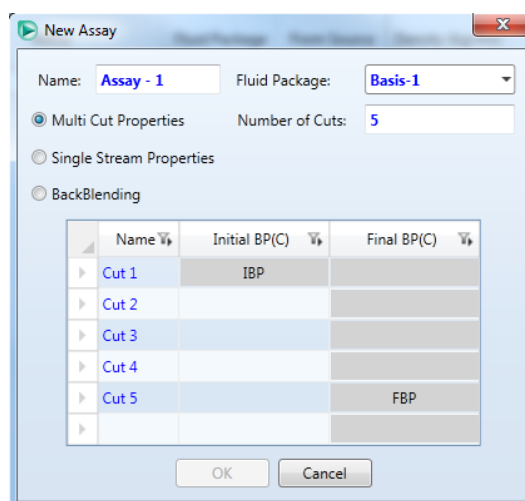
Ya que existe una gran variedad de formas para introducir la información del ensayo en un modelo de refinería, Aspen HYSYS Petroleum Refining permite a los usuarios introducir la información de forma manual al crear un nuevo ensayo. Al elegir esta opción se va a abrir una nueva ventana con varias opciones para darle formato a la tabla del ensayo y así poder introducir la información del ensayo según se tenga disponible.

1. Haga click en la flecha del botón para adicionar nuevos ensayos y elija Manually Enter



Existen tres opciones para introducir la información de los ensayos:

- Multi Cut Properties – Se usa cuando se tiene información específica de los cortes en el ensayo, como el rendimiento para cada corte. Se permite definir el número de cortes necesarios y la información de los rangos de punto de ebullición.



- Single-stream Properties – Se usa para evaluar un ensayo a partir de la información de la curva de destilado y las propiedades de bulbo para posteriormente crear una sola corriente que represente este ensayo en el modelo.

- Backblending – Se utiliza cuando los usuarios tienen información de las corrientes de productos y quieren conocer los datos del crudo que se alimenta a partir de información de componentes puros, curvas de destilación y propiedades de los productos en específico.

2. Borre las filas para introducir los datos de 5% y de 95%
3. Elija el tipo de datos de entrada como D86 de tipo volumétrico
4. Introduzca los siguientes datos para el crudo a caracterizar:

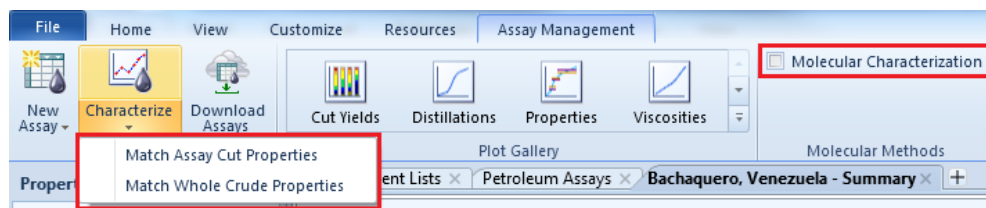
% Volumen Líquido	TBP Temp, F
0	35
10	147
30	281
50	468
70	662
90	985
100	1392

5. Haga click en OK
6. Guarde el archivo

CARACTERIZACIÓN DE LOS ENSAYOS

El Assay Management permite caracterizar un crudo de distintas formas utilizando el botón de Characterize. Estas opciones permiten relacionar las propiedades de corte del ensayo y las propiedades completas del crudo, así como llevar a cabo una caracterización molecular del ensayo.

Por defecto, las propiedades de bulto definidas por el usuario van a tener precedencia sobre las propiedades definidas para cada corte. (AspenTech, 2017)



Las tablas mostradas dentro de “Input Assay” pueden ser personalizadas para agregar propiedades generales del crudo y datos para cada corte en específico, así como más cortes, para que ayuden con la caracterización.

Input Summary		Pure Component	Distillation Data				
	Whole Crude	Cut1	Cut2	Cut3	Cut4	Cut5	
Initial Temperature (C)	IBP	IBP	8.8900	40.0000	96.4706	152.9412	
Final Temperature (C)	FBP	8.8900	40.0000	96.4706	152.9412	209.4118	
VanadiumByWt (%)	0.032	0.000	0.000	0.000	0.000	0.000	
NickelByWt (%)	0.005	0.000	0.000	0.000	0.000	0.000	
ParaffinsByVol (%)	6.245	96.809	85.629	64.451	40.681	24.008	
NaphthenesByVol (%)	50.569	3.191	14.288	32.534	47.158	56.543	
AromByVol (%)	43.186	0.000	0.083	3.015	12.162	19.449	
AromByWt (%)	41.721	0.000	0.162	4.065	14.478	21.540	
PourPoint (C)	-48.735	-273.150	-270.639	-243.425	-147.370	-99.814	
AnilinePoint (C)	78.297	35.986	34.097	34.719	39.263	45.317	
FreezePoint (C)	-29.192	-120.882	-99.576	-78.166	-72.222	-62.843	
API Gravity	11.012	100.478	284.360	102.121	24.733	19.005	
Group:	All		0.05	0.04	0.03	0.02	
Property:	AluminiumByWt	%	84.23	71.62	65.71	67.71	
				-47.80	3.41	43.81	

1. En el nuevo ensayo creado, haga click en el botón de Characterize Assay en la parte inferior derecha.

Observe que algunas propiedades van a ser estimadas y otras se van a quedar en blanco, ya que no existen valores de bulto para utilizarlos en la estimación (por ejemplo, % de contaminantes o % de azufre).

CARACTERIZACIÓN MOLECULAR

Al seleccionar la opción de caracterización molecular, se abre una lista de moléculas de hidrocarburos en el crudo y se muestran tres columnas para editar: porcentaje peso (Wt %) y los coeficientes alfa y beta. Adicionalmente, también se encuentran disponibles coeficientes de viscosidad. (AspenTech, 2017)

Dichas variables pueden ser modificadas antes llevar a cabo la caracterización molecular y una vez seleccionadas se puede hacer click en el botón de “Run Evaluation” para caracterizar el ensayo utilizando esta información. (AspenTech, 2017)

El usuario puede cambiar el método de caracterización quitando la selección de Caracterización Molecular en la cintilla y usar otro de los métodos disponibles para la caracterización. (AspenTech, 2017)

Profile name: Latin America_Venezuela_Bachaquero Select Profile Save Profile

Molecular Specification Enable regression Auto Setup

Molecular Class	Wt(%)	Carbon Number	CentroidTb (C)	ρ (kg/m ³)	Segment Name	Segment Structure	Function	β	α
Hydrocarbons molecules									
Paraffins	8.02	10	238.48	712.63	Methylene	-CH ₂ -	Gamma	7.45063	1
					One-branch methylene		Gamma	1.31755	1
					Two-branch methylene		Gamma	0.0142907	1
Naphthenes	25.24	18	348.78	853.36	Naphthenic side ring		Gamma	1.03348	1
					Side chain carbon number	-C _n H _{2n+1}	Gamma	10.7362	1
					Mole fraction of 6CR			0.818645	
Aromatics	53.85	26	516.64	1028.54	Aromatic side ring		Gamma	4.91126	1
					Naphthenic side ring		Gamma	4.21859	1
					Side chain carbon number	-C _n H _{2n+1}	Gamma	5.55045	1

Viscosity

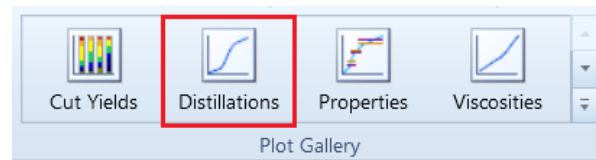
k0	k1	k2	k3
0	-10	0	0

REPORTE DE LA INFORMACIÓN DEL ENSAYO

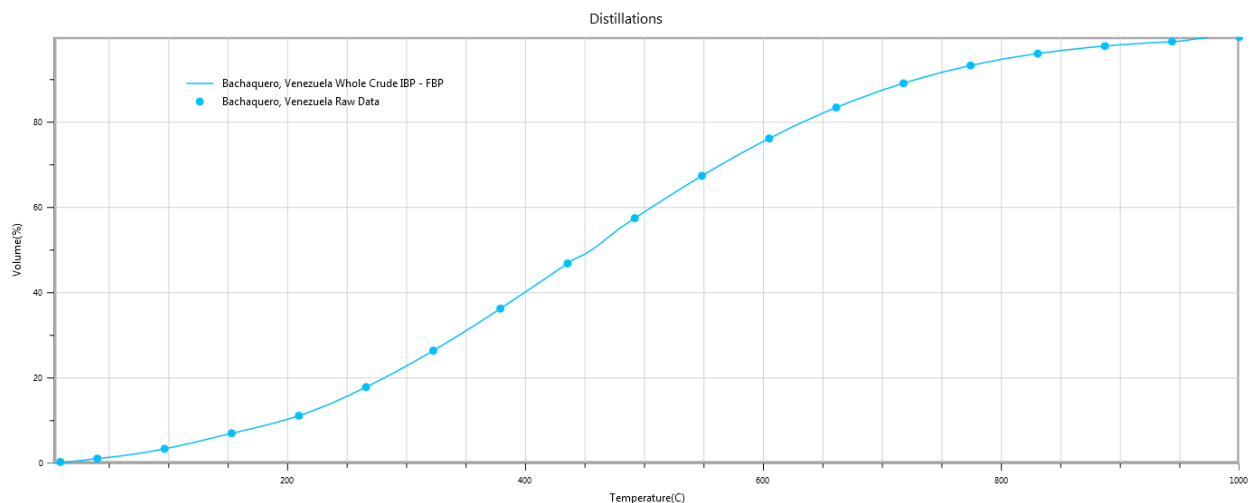
Dentro del Assay Management en HYSYS se puede graficar información de los ensayos. Esta herramienta hace más fácil visualizar las características de una corriente de crudo y permite al

usuario comparar distintos ensayos. Las propiedades de los ensayos que pueden ser graficadas se muestran dentro de la pestaña de Assay Management en la sección de Plot Gallery.

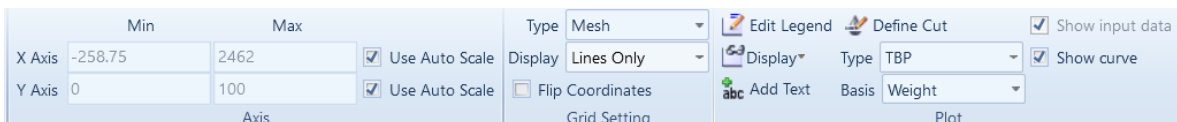
1. Haga click en el crudo Bachaquero desde el Panel de Navegación.
2. Haga click en el ícono de Distillations en la sección de gráficos en la cinta.



En este ejemplo, se muestra la curva de destilado para el crudo de Bachaquero.



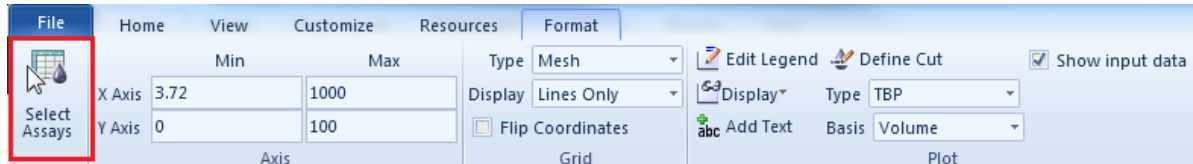
Para poder personalizar la gráfica, la cinta cuenta con opciones para cambiar las coordenadas de la gráfica, el tipo, la leyenda, el texto, etc.



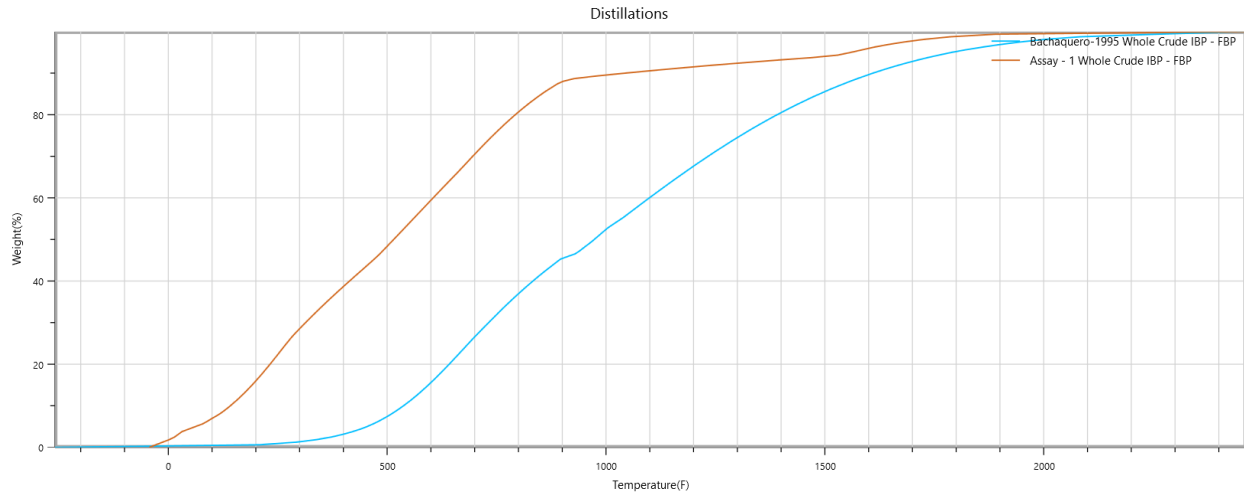
COMPARACIÓN DE VARIOS ENSAYOS

Para comparar datos de varios ensayos, primero se debe agregar un crudo nuevo siguiendo el procedimiento mencionado anteriormente.

1. Haga click en el botón de “Select Assays” en la cinta.



Después de agregar el ensayo adicional a la gráfica, se pueden comparar las propiedades más fácilmente.

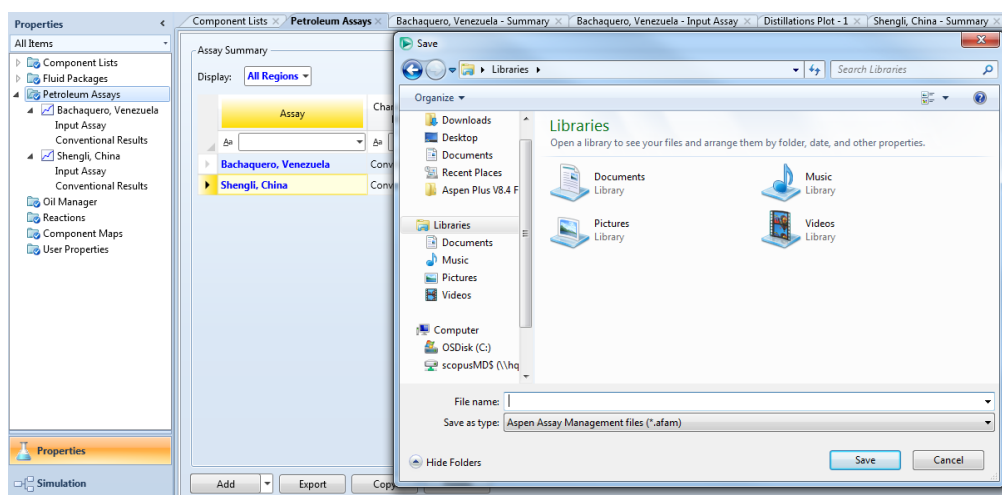


De igual manera, pueden verificarse otras propiedades de los ensayos, así como las distribuciones de los cortes, de acuerdo con las opciones de gráficos disponibles. Estas opciones también permiten hacer la comparativa entre distintos ensayos de la simulación.

EXPORTANDO INFORMACIÓN DEL ENSAYO

Los ensayos pueden ser exportados desde un caso de HYSYS Petroleum Refining mientras uno se encuentre en el ambiente de propiedades. Dentro de la sección de Petroleum Assays en el Panel de Navegación se accede al resumen de los ensayos.

Haga click en el botón de Export para guardar el ensayo como un archivo AFAM, el cual puede ser importado en otro caso de Aspen HYSYS Petroleum Refining o en Aspen PIMS y con esto, tener consistencia entre los modelos de planeación e ingeniería.



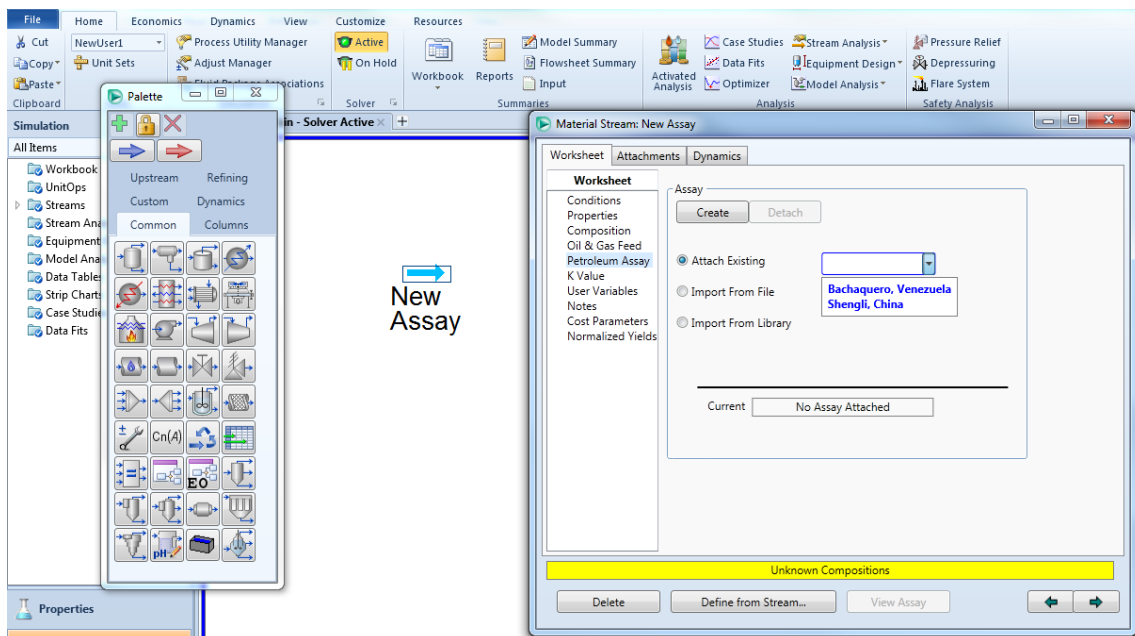
Ensayos en el ambiente de simulación

El objetivo principal de importar y crear ensayos en Aspen HYSYS Petroleum Refining es poder utilizarlos para evaluar el comportamiento de varios crudos dentro de un modelo de refinación. Existe un número de opciones para poder usar estos ensayos y sus resultados con Aspen Assay Management. Para empezar a trabajar con estos ensayos, hay que ir al ambiente de simulación.

1. Haga click en Simulation en la parte inferior izquierda de la ventana.

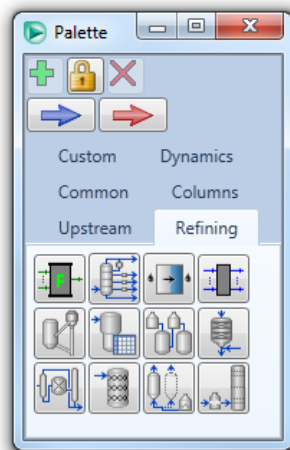
Cuando un ensayo ha sido creado y caracterizado en el ambiente de propiedades, inmediatamente puede ser utilizado en el ambiente de simulación. Esto se puede hacer agregando el ensayo a una corriente o llamándolo por medio de un bloque llamado Petroleum Feeder. Si el ensayo va a ser tratado como una sola corriente, se puede anexar a una corriente en la simulación. Primero se debe crear una corriente de material en el diagrama de flujo de procesos utilizando la paleta de objetos o cualquier otro de los métodos disponibles. (AspenTech, 2017)

2. Haga click en F11 para adicionar una nueva corriente.
3. Haga click en la opción de Petroleum Assay en la pestaña de Worksheet
4. Una vez que se abra la página para definir los ensayos, seleccione la opción de Attach Existing para mostrar una lista de los ensayos que ya se hayan caracterizado dentro del ambiente de propiedades.



5. Seleccione cualquiera de los dos ensayos caracterizados
6. Una vez, que el ensayo se seleccione, HYSYS va a definir la composición y las propiedades de la corriente de tal manera que el usuario solo debe definir las condiciones de la corriente para que esta se encuentre totalmente definida.

Si más de un ensayo se va a usar o si el ensayo se va a mezclar con otra corriente, se puede agregar el ensayo a través del bloque de Petroleum Feeder. En el Feeder se pueden seleccionar las corrientes de alimentación y los ensayos, así como las especificaciones de mezclado, y los paquetes de fluidos. Con esto, Aspen HYSYS va a calcular una corriente de producto que puede ser usada en la simulación.



En algunos casos, puede ser necesario crear un nuevo ensayo en una corriente. Para hacerlo, agregue la corriente a la simulación y proceda directamente a la sección de Petroleum Assay en la vista de propiedades de la corriente. El usuario puede agregar un ensayo existente, importar desde un archivo externo, importar desde la librería, o introducir los datos de forma manual directamente en la corriente.

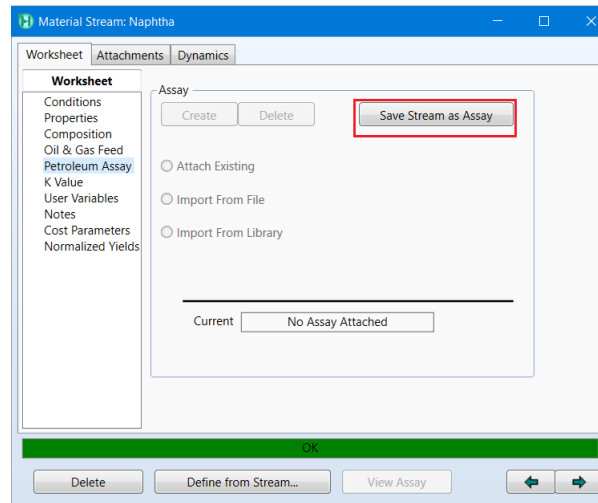
Una vez que el ensayo haya sido creado, se puede hacer click en el botón de View Assay en la parte inferior de la corriente para ver la información detallada del ensayo.

The screenshot displays the Aspen Plus software interface. On the left, a 'New Assay' button is visible. The main window shows the 'Material Stream: New Assay' dialog box with the 'Import From Library' option selected and 'Isthmus, Mexico' entered in the search field. Below the dialog, a yellow bar indicates 'Unknown Temperature'. On the right, the 'Input Assay' table is displayed, showing various properties for 'Whole Crude' and its cuts (Cut1, Cut2, Cut3).

	Whole Crude	Cut1	Cut2	Cut3
Initial Temperature (C)	IBP	IBP	8.8900	40.0000
Final Temperature (C)	FBP	8.8900	40.0000	96.4706
CutYieldByVol (%)	100.00	0.41	1.14	6.51
StdLiquidDensity (kg/m...	853.6915	281.5050	539.6696	706.6192
SulfurByWt (%)	1.326	0.020	0.006	0.014
nCSinsolublesByWt (%)	2.395	0.000	0.000	0.000
ConradsonCarbonByWt...	3.948	1.285	1.056	0.941
KinematicViscosity (cSt)...	4.935	0.399	0.374	0.492
KinematicViscosity (cSt)...	3.557	0.354	0.329	0.378
KinematicViscosity (cSt)...	1.594	0.291	0.264	0.291
TotalAcidNumber (mg...	0.103	0.001	0.002	0.017
VanadiumByWt (%)	0.004	0.000	0.000	0.000
NickelByWt (%)	0.001	0.000	0.000	0.000
MercaptanSulfurByWt (...)	0.004	0.013	0.009	0.006
ParaffinsByVol (%)	24.281	99.096	92.218	68.652
NaphthenesByVol (%)	28.968	0.000	6.007	22.425

En algunos casos, se puede requerir evaluar las propiedades de una corriente, posiblemente de los cortes de una columna o de los productos después de su procesamiento. Para ello, Assay Manager cuenta con la opción de exportar los datos de una corriente como ensayo para su evaluación de la misma manera que como cuando se caracteriza un ensayo. Para este caso, vamos a usar un ejemplo existente. Para ello hay que ir a la carpeta de ejemplos y abrir el archivo llamado “Assays and Crude Distillation unit with Preheat.hsc”.

1. Haga doble click en la corriente Naphtha
2. Vaya a la sección de Petroleum Assays
3. En esa ventana haga click en el botón de Save Stream as Assay



4. Vaya al ambiente de propiedades y vea cómo se genera un ensayo llamado Assay from Naphtha, que trae las propiedades del crudo para ser evaluadas dentro del Assay Manager.

