

**LÍNEA BASE ENERGÉTICA PARA LA UNIDAD U-200 DE LA GRB: ANÁLISIS Y
REGRESIÓN UTILIZANDO ESTADÍSTICA MULTIVARIADA**

RICARDO IVÁN RUBIO LIZCANO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICO QUIMICAS
ESCUELA DE INGENIERIA QUÍMICA
MAESTRÍA EN INGENIERÍA QUÍMICA
BUCARAMANGA**

2017

**LÍNEA BASE ENERGÉTICA PARA LA UNIDAD U-200 DE LA GRB: ANÁLISIS Y
REGRESIÓN UTILIZANDO ESTADÍSTICA MULTIVARIADA**

RICARDO IVÁN RUBIO LIZCANO

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

Director:

GIOVANNI MORALES MEDINA

Ph.D en Ingeniería Química

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO QUÍMICAS
ESCUELA DE INGENIERÍA QUÍMICA
MAESTRÍA EN INGENIERÍA QUÍMICA
BUCARAMANGA**

2017

CONTENIDO

	Pág.
INTRODUCCIÓN.....	16
1. OBJETIVOS.....	18
1.1 OBJETIVO GENERAL.....	18
1.2 OBJETIVOS ESPECÍFICOS.....	18
2. MARCO TEÓRICO.....	19
2.1 CONSUMO ENERGÉTICO Y LÍNEAS BASE.....	19
2.2 UNIDAD DE DESTILACIÓN PRIMARIA U-200.....	22
2.3 ANÁLISIS POR COMPONENTES PRINCIPALES (PCA) Y MÍNIMOS CUADRADOS PARCIALES (PLS).....	24
2.4 ALCANCES.....	26
3. METODOLOGÍA.....	27
3.1 CONSTRUCCIÓN DE UNA BASE DE DATOS CON LOS HISTÓRICOS DE PROCESO.....	27
3.2 ANÁLISIS DE LA BASE DE DATOS.....	27
3.3 IDENTIFICACIÓN DE ESCENARIOS DE CONSUMO ENERGÉTICO POR MEDIO DEL ANÁLISIS DE LOS COMPONENTES PRINCIPALES (PCA).....	28
3.4 DESARROLLO DE LA LÍNEA BASE ENERGÉTICA: REGRESIONES RLM, PCR Y PLS.....	29
4. ANÁLISIS DE RESULTADOS.....	31
4.1 DATOS HISTÓRICOS DE LA U-200.....	31
4.2 ESQUEMAS OPERATIVOS.....	37

4.3 ANÁLISIS PCA EN CADA CLASE	45
4.4 RLM PARA LA PREDICCIÓN DEL CONSUMO DE ENERGÍA EN CADA CLASE	53
4.4.1 Clase 0.....	53
4.4.2 Clase 1	56
4.4.3 Clase 2.....	58
4.4.4 Clase 3.....	59
4.4.5 Clase 4.....	60
4.5 PCR Y PLS PARA LA PREDICCIÓN DEL CONSUMO DE ENERGÍA EN CADA CLASE	65
4.5.2 Clase 1.....	69
4.5.3 Clase 2.....	72
4.5.4 Clase 3.....	75
4.5.5 Clase 4.....	78
5. MEJORAS EN EL PROCESO BASADAS EN LAS LÍNEAS BASE OBTENIDAS CON LAS REGRESIONES MULTIVARIADAS	83
6. IMPACTOS DEL TRABAJO DE APLICACIÓN	85
7. CONCLUSIONES	86
REFERENCIAS BIBLIOGRAFICAS.....	88
BIBLIOGRAFÍA.....	92
ANEXOS	96

LISTA DE FIGURAS

	Pág.
Figura 1. Diagrama de proceso resumido para la U-200.	23
Figura 2. Carga, Jet y Fondo de la T201 para la ventana de operación entre enero de 2013 y agosto de 2016. Los círculos muestran periodos con baja carga a la torre.	33
Figura 3. Carga, Jet y Fondo de la T204 para la ventana de operación entre enero de 2013 y agosto de 2016. Los círculos muestran periodos con alteraciones en la carga a la torre.....	34
Figura 4. Energía requerida por la U-200 por flujo de carga de crudo procesada versus carga procesada en la ventana de operación.....	36
Figura 5. Energía requerida por la U-200 por flujo de carga de crudo procesada versus carga procesada por año en la ventana de operación.	37
Figura 6. Análisis PCA a la base de datos histórica con pretratamiento de autoescalado. (a) varianza explicada. (b) Gráfico de influencia.....	38
Figura 7. Gráfica de los <i>scores</i> para los dos primeros componentes principales derivados del PCA. Las agrupaciones fueron obtenidas por el método <i>k-mean</i> utilizando la distancia euclidiana. (d) <i>Scores</i> para los dos primeros PC con la elipse de Hotelling T^2 al 95% de confianza.....	40
Figura 8. Gráfica de barras para los <i>loadings</i> del tercer componente principal (PC3). En el círculo de (a) se encierran en su orden de izquierda a derecha la destilación T10, T30, T50, T70 y T90.	41
Figura 9. Gráfica de la presión del flujo de entrada al H-201 versus la temperatura del flujo de entrada al mismo horno para las diferentes clases en la ventana operativa.	42

Figura 10. Gráfica de la presión del flujo de entrada al H-201 versus la temperatura del flujo de entrada al mismo horno para las diferentes clases en la ventana operativa.	42
Figura 11. Gráfica temperatura de chimenea H-201 versus temperatura de chimenea H-202 versus reflujo medio T-205 para las diferentes clases en la ventana operativa.	43
Figura 12. Representación de las clases en el gráfico de consumo energético específico versus el crudo proceso en la U-200 en la ventana de operación.	44
Figura 13. Resultados PCA para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (30%) y PC2 (16 Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) temperatura de entrada H-202 versus API versus BSW.	46
Figura 14. Resultados PCA para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (14%) y PC2 (12%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) flujo acpm T-201 versus flujo jet T-204 versus gasóleo T-205. ...	48
Figura 15. Resultados PCA para la clase 2. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (15%) y PC2 (12%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) temperatura de entrada H-201 versus temperatura de entrada H-202 versus flujo gasóleo T-205.	50
Figura 16. Resultados PCA para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (22%) y PC2 (11%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) flujo fondo T-201 versus flujo fondo T-204 versus temperatura flujo de entrada H-202.	53
Figura 17 Resultados PCA para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (28%) y PC2 (15%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95%	

de confianza; (d) poder calorífico combustible (COM) versus temperatura de chim. H-202 versus temperatura flujo de entrada H-201.....	54
Figura 18. Regresión RLM para la clase 0. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.....	55
Figura 19. Regresión RLM para la clase 1. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.....	57
Figura 20. Regresión RLM para la clase 2. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.....	59
Figura 21. Regresión RLM para la clase 3. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.....	62
Figura 22. Regresión RLM para la clase 4. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.....	63
Figura 23. Resultados PCR para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).....	67
Figura 24. Resultados PLS para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).....	68
Figura 25. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 0 (5 VL).....	69
Figura 26. Resultados PCR para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).....	70
Figura 27. Resultados PLS para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).....	71

Figura 28. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 1 (6 VL).....	72
Figura 29 Resultados PCR para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).....	73
Figura 30. Resultados PLS para la clase 2. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).....	74
Figura 31. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 2 (4 VL).....	75
Figura 32. Resultados PCR para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).....	76
Figura 33. Resultados PLS para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).....	77
Figura 34. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 3 (4 VL).....	78
Figura 35. Resultados PCR para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).....	79
Figura 36. Resultados PLS para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).....	80

Figura 37. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 4 (4 VL).....81

LISTA DE TABLAS

	Pág.
Tabla 1 Relación entre las clases y algunas de las variables operacionales.....	44
Tabla 2. ANOVA para la RLM de los datos de la clase 0 (Tomada de Unscrambler).	56
Tabla 3. ANOVA para la RLM de los datos de la clase 1 (Tomada de Unscrambler).	58
Tabla 4. ANOVA para la RLM de los datos de la clase 2 (Tomada de Unscrambler).	61
Tabla 5. ANOVA para la RLM de los datos de la clase 3 (Tomada de Unscrambler).	64
Tabla 6. ANOVA para la RLM de los datos de la clase 4 (Tomada de Unscrambler).	65
Tabla 7. Coeficientes regresión PLS para las variables autoescaladas de la clase 0 (5 VL).....	69
Tabla 8 Coeficientes regresión PLS para las variables autoescaladas de la clase 1 (6 VL).....	71
Tabla 9. Coeficientes regresión PLS para las variables autoescaladas de la clase 2 (4 VL).....	74
Tabla 10. Coeficientes regresión PLS para las variables autoescaladas de la clase 3 (4 VL).....	77
Tabla 11. Coeficientes regresión PLS para las variables autoescaladas de la clase 4 (4 VL).....	80

LISTA DE ANEXOS

	Pág.
Anexo A. Variables	96
Anexo B Estadísticos de los datos históricos.....	99

RESUMEN

TÍTULO: LÍNEA BASE ENERGÉTICA PARA LA UNIDAD U-200 DE LA GRB: ANÁLISIS Y REGRESIÓN UTILIZANDO ESTADÍSTICA MULTIVARIADA*.

AUTOR: RICARDO IVÁN RUBIO LIZCANO**

PALABRAS CLAVES: Línea base energética, unidad de destilación de crudos, análisis por componentes principales (PCA), regresión con estadística multivariada.

CONTENIDO: La norma ISO 50001 propone un sistema de gestión de energía (SGE) que establece una revisión energética de los procesos industriales y la construcción de un patrón de comparación denominado línea base energética. La implementación y el seguimiento de esta norma en refinerías, como la Gerencia Refinería de Barrancabermeja (GRB), conlleva a la disminución de sus consumos de energía. Actualmente la estadística multivariada es una buena opción para la obtención de líneas base energéticas en procesos complejos con una gran cantidad de variables. El presente trabajo expone los resultados de un protocolo para el desarrollo de la línea base energética para la unidad U-200 de la GRB, utilizando información histórica de proceso y las propiedades de los crudos tratados y los productos obtenidos. La metodología envuelve el análisis de esta información utilizando métodos de estadística multivariada como el análisis por componentes principales (PCA), regresión lineal múltiple (RLM), regresión por componentes principales (PCR) y regresión por mínimos cuadrados parciales (PLS). El análisis realizado conduce a descartar una sola línea base para la U-200 teniendo en cuenta que se identificaron 5 clases o esquemas de operación por el método de agrupación *k-mean*, corroborados con los reportes de planta. La regresión RLM aplicada sobre las 5 clases reproduce los datos de consumo energético específico de la unidad con el mayor desempeño, seguido por las regresiones PLS y PCR, por lo cual se seleccionan los modelos obtenidos con la regresión lineal múltiple (RLM) para las 5 clases.

* Trabajo de aplicación

** Facultad de Ingenierías Físicoquímica. Escuela de Ingeniería Química. Maestría en Ingeniería Química. Director: Giovanni Morales

SUMMARY

TITLE: ENERGY BASELINE FOR UNIT U-200 GRB: ANALYSIS AND REGRESSION USING MULTIVARIATE STATISTICS*

AUTHOR: RICARDO IVÁN RUBIO LIZCANO **

KEYWORDS: Energy baseline, crude distillation unit, principal component analysis (PCA), multivariate statistics regression.

CONTENTS: ISO 50001 proposes an energy management system (SGE) that establishes an energy review of industrial processes and the construction of a comparison pattern called energy baseline. The implementation and monitoring of this standard in refineries, such as the Barrancabermeja Refinery (GRB), leads to the reduction of its energy consumption. Multivariate statistics are currently a good choice for obtaining energy base lines in complex processes with a large number of variables. The present work presents the results of a protocol for the development of the energy baseline for unit U-200 GRB, using historical process information and the properties of the crude treated and the products obtained. The methodology involves the analysis of this information using multivariate statistical methods such as Principal Component Analysis (PCA), Multiple Linear Regression (RLM), Principal Component Regression (PCR) and Partial Least Squares (PLS) regression. The analysis performed leads to discarding a single baseline for the U-200 considering that 5 classes or operation schemes were identified by the k-mean grouping method, corroborated with the plant reports. The RLM regression applied on the 5 classes reproduces the specific energy consumption data of the unit with the highest performance, followed by the PLS and PCR regressions, whereby the models obtained with the multiple linear regression (RLM) for the 5 classes are selected.

* Thesis

** Faculty of Physico-Chemical Engineering. School of Chemical Engineering. Magister in Chemical Engineering Director: Giovanni Morales

INTRODUCCIÓN

Las organizaciones han usado desde hace algunos años un amplio rango de procedimientos técnicos para monitorizar y controlar su consumo de energía, especialmente aquellos cuyos costos son críticos para su actividad. El desarrollo de un sistema de gestión de energía (SGE) representa un enfoque proactivo, sistemático y lógico para abordar los problemas de la eficiencia energética en lugar de las actividades reactivas y poco sistemáticas [1]. La norma ISO 50001 propone un SGE que establece una revisión energética de los procesos industriales y la construcción de un patrón de comparación denominado línea base energética. La implementación y el seguimiento de esta norma en refinerías, como la Gerencia Refinería de Barrancabermeja (GRB), conllevarían a la disminución de sus consumos de energía, logrando disminuir el índice de intensidad de energía (IIE¹). Valores bajos en el IIE representan una disminución en la generación de gases de efecto invernadero (potencial de reducción de 49.4 ton/año de CO₂), ahorros en servicios industriales (se estima que la GRB invirtió \$148 millones de dólares en 2014 por concepto de gas combustible) y una mejor posición en el estudio de *benchmarking* realizado por SOLOMON [2].

El Departamento de Refinación de la GRB procesa alrededor de 240 kilobarriles por día (kbpd) en 5 unidades de destilación, de las cuales la U-200 es la unidad que cuenta con la mayor capacidad (83 kbpd) y por lo tanto es la que presenta los mayores consumos energéticos. Una disminución en los gastos de operación de la U-200 en lo referente a los consumos de gas combustible, vapor y energía eléctrica dependiendo del esquema operativo según la carga a la unidad repercutiría en un aumento en el margen de refinación y una mejora del IIE de la

¹ IIE de una refinería o un proceso se obtiene de la relación entre el consumo real de energía de la refinería y el consumo de energía estándar según una ecuación empírica establecida por la empresa SOLOMON.

GRB; según estimaciones de la GRB, la disminución en consumo de energía en la U-200 puede conllevar a una reducción de hasta 90 gigaBTU/año (GBTU/año), para un potencial de disminución del gasto de \$283 mil dólares anuales y una reducción de 1.43 ton/año de emisiones de CO₂ [2,3].

La variación en el IIE de la U-200 depende de la característica de la carga de crudo y los productos de la unidad, de las condiciones operacionales, de integridad operacional y de la planeación tanto en carga como en productos. Estas variables y condiciones pueden ser analizadas utilizando datos históricos en una ventana operacional de mínimo 1 año para conseguir la suficiente cantidad de estados estacionarios de la unidad. El tratamiento estadístico de estos datos históricos puede reportar un modelo de IIE referente a las principales variables que afectan el consumo energético del proceso [4-6]. Este modelo de IIE es representado a través de una línea base energética que puede ser acoplada a los programas de operación de la unidad para restringir las variables de proceso a regiones de consumo reducido de energía. Asimismo, una línea base corresponde a una herramienta que ayuda a los ingenieros de consola en el soporte de decisiones operativas.

El presente trabajo de aplicación final de maestría expone los resultados de un protocolo para el desarrollo de la línea base energética para la unidad U-200 de la GRB, utilizando información histórica de proceso y las propiedades de los crudos tratados y los productos obtenidos. La metodología envuelve el análisis de esta información utilizando métodos de estadística multivariada como el análisis por componentes principales (PCA), regresión lineal múltiple (RLM) y los mínimos cuadrados parciales (PLS).

1. OBJETIVOS

1.1 OBJETIVO GENERAL

Proponer una línea base energética para la unidad de destilación primaria U-200 a partir de datos históricos de proceso y técnicas de estadística multivariada.

1.2 OBJETIVOS ESPECÍFICOS

1. Construir una base de datos con los históricos de proceso e identificar las variables principales que afectan el desempeño energético de la unidad U-200.
2. Identificar patrones con los datos históricos para el proceso operando en estado estacionario por medio del análisis por componentes principales.
3. Identificar mejoras de proceso basadas en la regresión multivariada de la línea base energética considerando los escenarios operativos en estado estacionario para la U-200.

2. MARCO TEÓRICO

2.1 CONSUMO ENERGÉTICO Y LÍNEAS BASE

Los sistemas de gestión integral de la energía (SGIE) utilizan estándares que tienden a disminuir los consumos de las diferentes clases de energía, minimizando la inversión económica y al mismo tiempo minimizando los impactos sobre el medioambiente [22]. Entre los estándares internacionales para los sistemas de gestión de energía se encuentra la norma ISO 50001: 2011, aprobada por la Organización Internacional de Normalización (International Organization for Standardization, 2011).

La norma ISO 50001 es aplicable a cualquier organización que busque mejorar sus indicadores de consumo energético a través del ciclo de mejora continua planear-hacer-verificar-actuar [23]. El estándar considera el análisis de datos de consumos históricos, con lo cual, un equipo de trabajo –aprobado por la alta dirección– determina unos índices de consumo energético –relacionados con energía eléctrica, agua, vapor, combustibles– en la organización para una posterior evaluación basada en requerimientos de orden legal. Estos índices, que pueden ser diferentes dependiendo de la empresa, son utilizados para comparar los desempeños energéticos con diferentes empresas del sector y establecer planes indicadores de desempeño energético de control y mejora [25]. Entre los indicadores utilizados en diferentes empresas se tienen:

- Emisiones de CO₂, medidos en tonelada por año.
- Consumo de energía por empleado, medido en %.
- Consumo de energía en áreas específicas, medido en %.
- Valor de energía consumida, medida en Julios por pesos m.c.

Otros indicadores, más elaborados y dirigidos hacia una empresa en particular, son presentados por May y colaboradores (2015) [26]. Los indicadores de desempeño son utilizados en conjunto con datos históricos de consumo energético de la organización para definir la denominada línea base de energía; usualmente, la línea base es obtenida por un procedimiento de ajuste de curvas por medio del algoritmo de mínimos cuadrados. Esta línea base identifica las variables principales que influyen en el consumo energético y su impacto en la reducción del consumo energético. Asimismo, la línea base corresponde a una herramienta para la predicción de consumos futuros, así como un insumo para minimizar los usos actuales de la energía en la organización [27]. La evaluación por medio de la línea base puede indicar las áreas de consumo significativo, así como oportunidades de mejora en cuanto al desempeño energético en la organización, con lo cual es posible establecer planes de acción y de control del uso de la energía [23, 28].

El seguimiento de la norma ISO 50001 ha reportado ahorros en los consumos energéticos industriales del orden de un 40% [29] y un 50%. Estos ahorros en consumos energéticos puede disminuir en gran manera los costos en muchas organizaciones –e.g. en la industria del cemento los costos por consumo de energía se acercan al 20% de los costos totales [28]; asimismo, se estima que la refinería de Barrancabermeja invirtió \$148 millones de dólares en 2014 por concepto de gas combustible [2]–. Las publicaciones técnicas mundiales reportan 305 casos –entre 1976 y 2014– en donde la aplicación de un SGIE redujo los costos en diferentes empresas y organizaciones como compañías de procesamiento eléctrico, condominios residenciales, hospitales, edificios comerciales, empresas de construcción, universidades e instituciones educativas, centros recreativos y comerciales, industrias de producción de pulpa y papel, industrias petroleras y petroquímicas, empresas automotoras y compañías de celulares y telecomunicaciones, entre otras [26].

Debido al incremento de los costos de la energía y a la disminución de los recursos energéticos, las industrias han tenido la necesidad de explorar nuevas tecnologías y procedimientos que ayuden a supervisar y controlar su consumo. En particular, en la GRB se cuenta con un sistema de gestión de energía (SGE) para analizar y reducir los consumos energéticos en sus unidades de proceso. Asimismo, el SGE planea implementar los lineamientos descritos en la norma ISO 50001, de indicadores de desempeño energético, en la GRB; un indicador de energía es, por ejemplo, la cantidad de energía consumida por unidad. La implementación de esta norma incluye el desarrollo de líneas base energéticas para cada unidad; los demás elementos requeridos para la implementación de la norma en la GRB ya se han aplicado a través del SGE.

La línea base corresponde a un estándar de comparación propia del proceso con la cual se puede evaluar una eficiencia energética con respecto a algún procedimiento o cambio en el mismo; también se puede evaluar la eficiencia en el trabajo de los operarios y/o entrenamiento de los mismos. Con esto es posible detectar desviaciones en el consumo de las áreas de uso significativo y evaluar los ahorros que se pueden conseguir; es importante mencionar que las líneas de base se enfocan principalmente en las áreas de mayor consumo de energía. Si el consumo energético de un proceso está por encima de su línea de base puede ser que haya algún problema que esté haciendo que las instalaciones consumidoras utilicen más energía de la necesaria. Por el contrario, si el consumo de un área de uso está por debajo de su línea de base, se deberá probablemente a una mejora introducida [7].

Actualmente la estadística multivariada es una opción para la obtención de líneas base energéticas en procesos complejos con una gran cantidad de variables, como los ocurridos en la industria petrolera. Velázquez *et al.* (2012) realizaron un estudio que describe el desarrollo de un sistema de gestión energética para una planta de reformado catalítico de naftas, caracterizando el desempeño energético

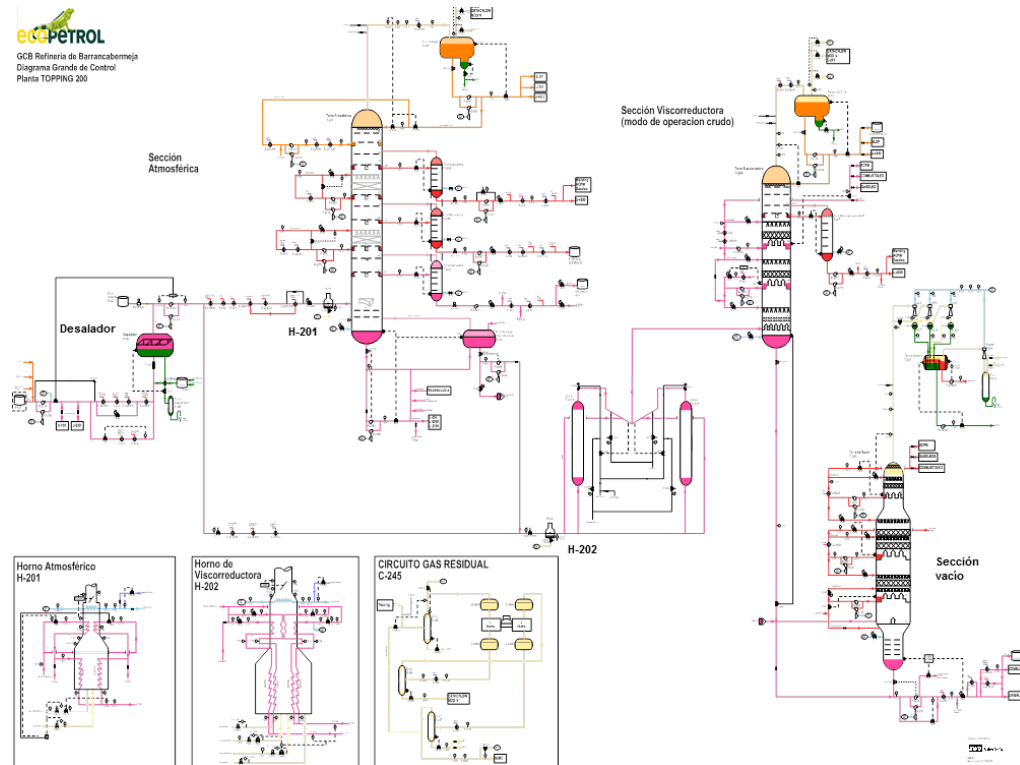
de la planta, identificando las variables clave de mayor influencia y representando mediante un modelo matemático la dependencia de los indicadores de desempeño, por medio de la técnica de estadística multivariable denominada análisis de función discriminante. Los modelos predictivos propuestos por *Velázquez et al.* fueron empleados para calcular potenciales de ahorro debido a mejoras en equipos y procesos [4].

De otro lado, *Biscarri et al.* (2012) abordaron la optimización del consumo de energía en un proceso de destilación de nafta, asegurando una tasa de producción adecuada y cumpliendo la calidad requerida del producto. Asimismo, obtuvieron el mínimo el costo de operación a través de la técnica de estadística multivariada denominada análisis canónico de discriminante [6]. Otro estudio es realizado por *Monedero et al.* (2012) para una unidad de reformado catalítico utilizando una técnica de minería de datos y su combinación con redes neuronales, para la predicción e interpolación de posibles puntos de operación. Los resultados de *Monedero et al.* fueron aplicados como parte de un software piloto para un sistema de control tipo SCADA, obteniendo disminuciones de consumo de energía en promedio del 7% [5]. De acuerdo a la revisión bibliográfica realizada este grupo de investigadores es el que tiene estudios más actuales y enfocados a la industria de gas y petróleo para la aplicación de análisis multivariado en la obtención de líneas base energéticas.

2.2 UNIDAD DE DESTILACIÓN PRIMARIA U-200

La unidad U-200 consta de tres secciones: una zona atmosférica, una zona viscorreductora que opera normalmente en modo operación de crudos y una sección de vacío. El siguiente es el diagrama general del proceso de refinación de la U-200:

Figura 1. Diagrama de proceso resumido para la U-200.



El crudo llega de los campos a través de oleoductos y se carga a un tanque de almacenamiento. Luego, en la U-200, el crudo es impulsado por unas bombas hacia un tren de intercambiadores de calor, los cuales suministran energía para aumentar la temperatura de la carga que entrará al desalador. En este equipo, el crudo es lavado con agua y químicos con el propósito de remover las impurezas que pueden causar daños, fallas y afectaciones, en equipos posteriores. Una vez sale de la unidad de desalado, el crudo pasa por un segundo tren de intercambiadores de calor, los cuales suministran energía al crudo para que pueda alcanzar la temperatura necesaria para entrar al horno H-201 de la sección atmosférica. En este horno, el crudo desalado es calentado hasta su zona de dos fases para efectuar su separación en la torre atmosférica T-201 y de esta manera recuperar productos livianos valiosos.

El procesamiento del crudo en la zona viscorreductora inicia con la llegada del crudo a un tambor de carga. Luego, mediante unas bombas, es impulsado hacia los serpentines del horno de viscorreductora H-202 en donde se le incrementa la temperatura hasta vaporizarlo parcialmente para entrar a la torre fraccionadora por la línea de transferencia. En modo de operación crudo, la torre Fraccionadora T-204 de la viscorreductora trabaja como una torre destiladora normal, donde se separan: gases, nafta, jet, ACPM y crudo reducido que pasa por vasos comunicantes. Este último sale por el fondo de la torre y es enviado a la zona flash de la torre de vacío T-205. La parte vaporizada asciende por la torre y a medida que se va enfriando produce corrientes laterales líquidas que se recuperan en los colectores de salida de la torre (gasóleos de vacío), y el corte más pesado va a los platos del fondo (fondos de vacío). Cuando la T-204 opera en modo viscorreductora se independizan varios sistemas para que esta unidad cargue crudo reducido, fondos de vacío de las otras unidades de destilación o un mezcla de ellos, se realiza la viscorreducción de esta corriente con el H-202 y se obtienen los siguientes productos: nafta de viscorreductora, disolventes y fondos de viscorreductora[8].

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-Principal Component*) o variables latentes. Cada componente principal contiene información de diferente relevancia, así 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 se reducirá al menor número de variables, manteniendo la mayor parte de la varianza de la información posible [9,10].

La proyección sobre los nuevos ejes de variación no colineales (componentes principales) genera los denominados *scores* (puntuaciones), mientras que los cosenos de los ángulos respecto las antiguas variables generan los denominados *loadings* (cargas). Con los *scores* y los *loadings* de los componentes principales es posible, principalmente a través de gráficos, la identificación de posibles agrupaciones, tendencias y patrones relevantes en los datos analizados [10-13]. 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 [10-13]. 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 [10-15].

Otro de los procedimientos de estadística multivariada comúnmente utilizados para la calibración de propiedades en datos multivariados 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) [13, 15]. 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. [10, 15]

2.4 ALCANCES

Este trabajo de aplicación tiene como alcance la obtención de un modelo adecuado para la línea base energética de la U-200 de la GRB aplicando metodologías de estadística multivariada para el análisis y la regresión de los respectivos datos históricos entre los años 2013 a 2015.

3. METODOLOGÍA

3.1 CONSTRUCCIÓN DE UNA BASE DE DATOS CON LOS HISTÓRICOS DE PROCESO

Los datos de los diferentes sensores disponibles en la U-200 en los últimos 3 años fueron recolectados del sistema de información de la GRB denominado PI; esta unidad durante los últimos 4 años ha presentado el menor número de eventos operativos como ampliaciones, reparaciones mayores, paradas de planta y eventos importantes a nivel de seguridad de procesos. Asimismo durante estos 4 años la unidad U-200 fue programada para que operara con diferentes crudos mezcla y considerando el escenario de viscorreducción, mostrando estado estacionario durante gran parte del tiempo. La información que incompleta en la base de datos fue imputada o adjudicada según la tendencia que presenten los datos globales [15]. Con la base de datos completa los consumos energéticos relacionados con gas combustible, vapor y energía eléctrica en los equipos de la unidad fueron estimados.

3.2 ANÁLISIS DE LA BASE DE DATOS

Los valores promedio, las desviaciones estándar y las varianzas fueron calculados para cada una de las variables más representativas de cada equipo. También se incluyeron los consumos energéticos en este análisis, lo que permitió determinar las tendencias de consumo, su magnitud y posibles errores en las mediciones o cálculos de los mismos. La detección de datos atípicos fue efectuada por medio del gráfico *boxplot* (caja y bigotes) [12,14]; los valores atípicos, que pueden corresponder a posibles errores de medición o diferentes tendencias de consumo

(diferentes escenarios), fueron identificados en el gráfico *boxplot* al ubicarse por fuera de los límites: límite superior $Q_3+1.5*(Q_3-Q_1)$, límite inferior $Q_1-1.5*(Q_3-Q_1)$, donde Q_i corresponde al cuartil i [14, 15].

3.3 IDENTIFICACIÓN DE ESCENARIOS DE CONSUMO ENERGÉTICO POR MEDIO DEL ANÁLISIS DE LOS COMPONENTES PRINCIPALES (PCA)

En esta etapa se aplicó la metodología de los componentes principales (PCA) para el análisis de la base de datos (variables originales). Con esta metodología se obtuvieron nuevas variables no correlacionadas (componentes principales), las cuales reportaron valores característicos de *loadings* y *scores*. Con estos valores se determinó la interrelación entre las variables de proceso y los consumos energéticos por medio de las gráficas de dispersión (PC1 vs PC2, PC2 vs PC3, etc.). El PCA se aplicó a las variables transformadas (autoescalado) para alcanzar una distribución normal aceptable y una distribución bivariada con distribución homogénea en la varianza (homocedasticidad) [10, 15]. Asimismo, con el PCA se identificaron y agruparon por medio de estadísticos [10] los diferentes escenarios que identifican tendencias de consumos energéticos; se aplicó el test estadístico de Hotelling (T^2) para establecer posibles datos atípicos [15, 16]. El estadístico T^2 corresponde a una generalización multidimensional del estadístico F utilizado en las pruebas de validación de hipótesis nula [17, 18]. Con este estadístico se construye un elipsoide que limita las observaciones para un nivel significancia definido ($\alpha=5\%$); las observaciones ubicadas fuera del elipsoide pueden ser consideradas como atípicas o representativas de una distribución normal multivariada diferente [18]. El elipsoide descrito por el estadístico T^2 se construirá sobre el espacio multidimensional no correlacionado descrito por los *scores* de los componentes principales. Los diferentes agrupamientos o clases fueron establecidos por el método *k-mean* y su ubicación sobre la gráfica de los *scores* del PCA.

3.4 DESARROLLO DE LA LÍNEA BASE ENERGÉTICA: REGRESIONES RLM, PCR Y PLS

El desarrollo de la línea base energética (relación entre el índice de energía y las variables de proceso) para cada escenario de consumo energético se obtuvo por medio de una regresión lineal múltiple (RLM) con las variables iniciales autoescaladas y por otro lado, utilizando como variables independientes los componentes principales del PCA (regresión PCR) [10]; para RLM, el método de selección *backward elimination* (eliminación hacia atrás) fue seleccionado para construir las ecuaciones de regresión con las variables estadísticamente significativas ($p\text{-value} < 0.05$), con la mayor contribución a la varianza del consumo de energía específico. Asimismo, la regresión por mínimos cuadrados parciales (PLS) fue aplicada en cada escenario de consumo energético siguiendo la metodología de descomposición de acuerdo a la máxima covarianza para obtener las variables latentes [10, 15]. Con esto, las regresiones RLM, PCR y PLS generaron un modelo matemático por cada escenario identificado en la actividad anterior.

Las regresiones RLM, PCR y PLS (en los datos originales y en los autoescalados) serán calibradas utilizando el 70% de los casos de la base de datos, mientras que la validación de los modelos fue realizada con el porcentaje restante; también, se analizaron los estadísticos generados por la validación cruzada (*cross-validation*) en la calibración de los modelos. Los desempeños para cada modelo en cada una de estas etapas fueron determinados por medio de los estadísticos del error estándar de predicción, el coeficiente de determinación $(R^2)^2$ y la raíz del error

² Para calibración $R^2 = 1 - \text{RSS}/\text{SST}$, donde RSS corresponde a la suma de los cuadrados de los residuos de calibración y SST representa la varianza total del conjunto. Para validación R^2 corresponde a $Q^2 = 1 - \text{PRESS}/\text{SST}$, donde PRESS es la suma de los cuadrados de los residuos de predicción según el conjunto externo (30%) y SST representa la varianza total del conjunto.

cuadrático medio (RMSE)³. Asimismo, el número de componentes principales y variables latentes que se utilizaron para las regresiones PCR y PLS, respectivamente, fueron seleccionadas de acuerdo a la tendencia en el valor de PRESS⁴ obtenido en la etapa de validación cruzada (PRESS-CV). Un incremento en la tendencia de PRESS-CV define el número de componentes principales o variables latentes óptimas para los modelos; i.e. el valor de PRESS-CV que presenta incremento en la tendencia hace referencia a un componente principal o variable latente con un aporte insignificante a la explicación de la varianza del modelo [10].

Los cálculos fueron realizados utilizando *TheUnscrambler X* licenciado para ECOPETROL y la UIS.

³ Para calibración $RMSE = \sqrt{RSS/N}$, donde N es el número total de muestras de calibración. Para validación $RMSE = \sqrt{PRESS/N}$ donde N es el número de muestras de la validación externa.

4. ANÁLISIS DE RESULTADOS

4.1 DATOS HISTÓRICOS DE LA U-200

La base de datos históricos de la U-200 analizada en el presente trabajo incluye las propiedades físicas del crudo API, BSW, azufre y temperaturas TBP al 10, 30, 50, 70 y 90% de evaporado, así como variables de proceso como carga a los hornos H-201 y H-202, temperaturas de los flujos y flujos de productos de las torres T-201, T-204 y T-205; la lista completa de los parámetros y variables que conforman los datos históricos se encuentra en el anexo A. El tamaño de la base de datos históricos fue de 30057x55 que corresponde al número de mediciones cada hora desde el primero de enero de 2013 hasta el 31 de agosto de 2016 para las 55 variables y parámetros disponibles de la U-200. Durante la depuración de la base fue posible identificar que los sensores para los flujos de jet y de fondo de la T-204 presentaron medidas incompletas para periodos de más de dos meses, por lo cual sus valores fueron adjudicados –imputados– por medio de regresiones desarrolladas utilizando las variables de la base de datos con mayor correlación respecto a estos flujos, según,

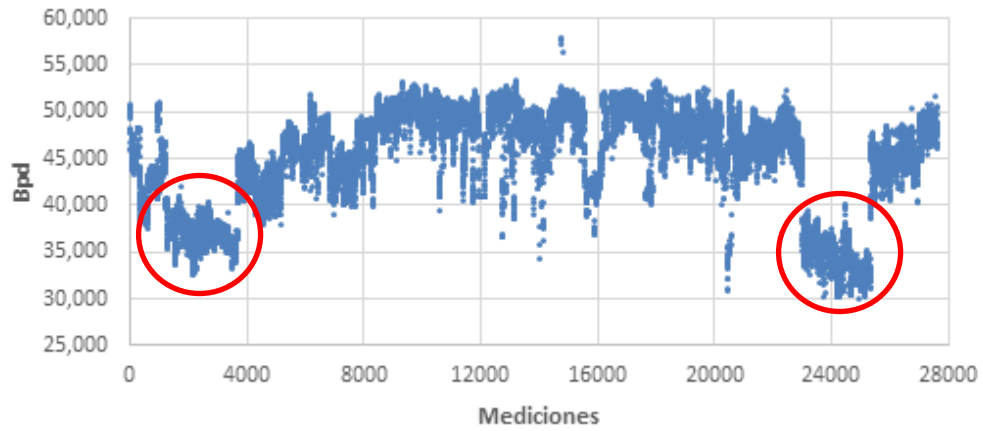
$$Jet_{T204} = 0.0452 * Carga_{H202} + 0.3225 * RC_{T204} - 0.0802 * RM_{T204} - 0.0425 * RI_{T204} \quad (1)$$

$$Fondo_{T204} = 0.638 * Carga_{H202} - 13.6358 * TE_{H202} - 0.074 * RI_{T204} \quad (2)$$

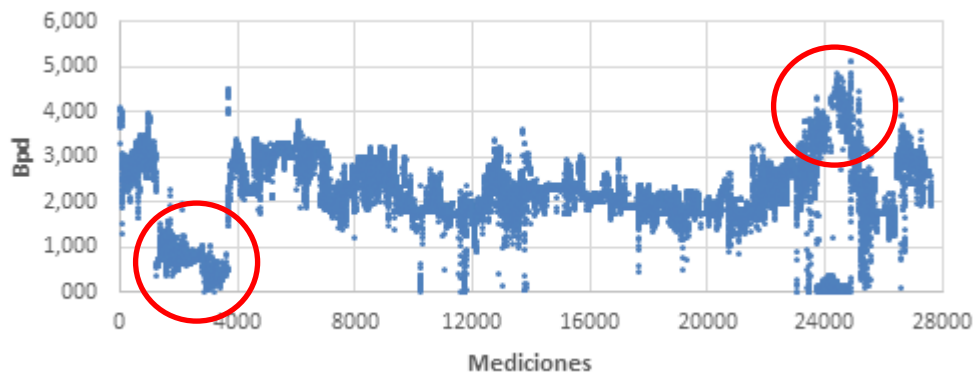
Donde, $Carga_{H202}$ y TE_{H202} corresponden al flujo y a la temperatura del crudo cargado al horno H202, mientras que RC_{T204} , RM_{T204} y RI_{T204} representan en su orden los reflujos de cima, medio e inferior de la torre T204.

Asimismo, diferentes mediciones fueron eliminadas de la base de datos debido al incumplimiento en el balance de materia en las torres de destilación por sensores sin calibración o por eventualidades operacionales. Después de la depuración, la base de datos quedó conformada por 27634 mediciones para la misma ventana de operación. Las Figuras 2 y 3 presentan la variación de las cargas y de algunos productos de las torres T201 y T204 para la base de datos de históricos. Según estas figuras, las cargas a las torres exhiben valores que representan operación continua con estados cuasi-estacionarios durante la ventana de operación analizada. La U-200 muestra dos periodos prolongados de baja carga de crudo (entre 33 y 38 kbpd) referentes a febrero-agosto de 2013 y a enero-mayo de 2016, durante los cuales la T-204 operó como unidad de viscorreducción; durante estos periodos la operación de la T201 presentó fluctuaciones acentuadas en los productos livianos debido a ajustes en el tren de precaliente y a errores en los medidores durante ese modo de operación, que fueron corroborados con problemas en esos instrumentos de acuerdo a los informes de turno respectivos de la unidad para ese periodo.

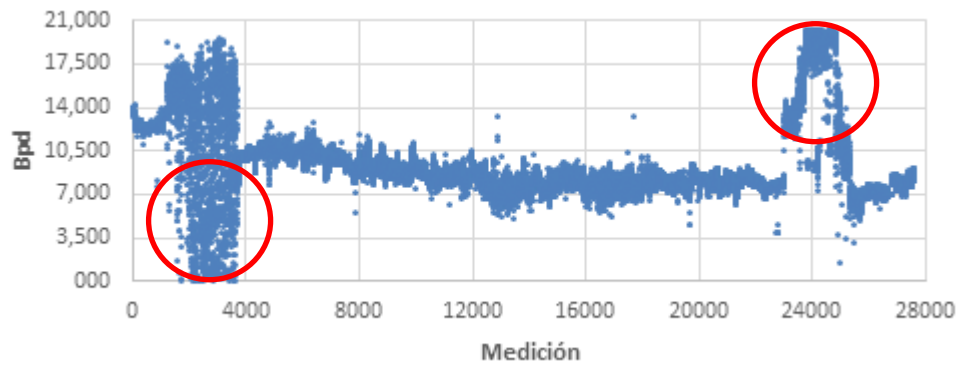
Figura 2. Carga, Jet y Fondo de la T201 para la ventana de operación entre enero de 2013 y agosto de 2016. Los círculos muestran periodos con baja carga a la torre.



(a) Carga a T201

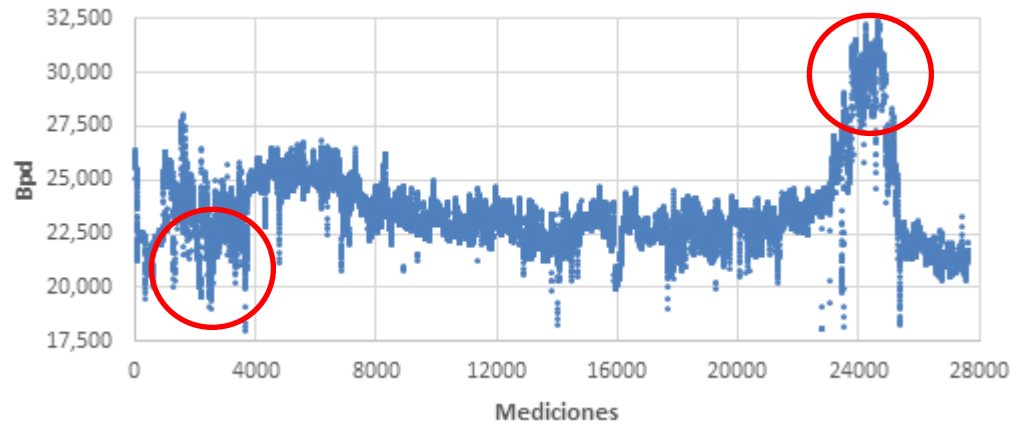


(b) Jet de T201

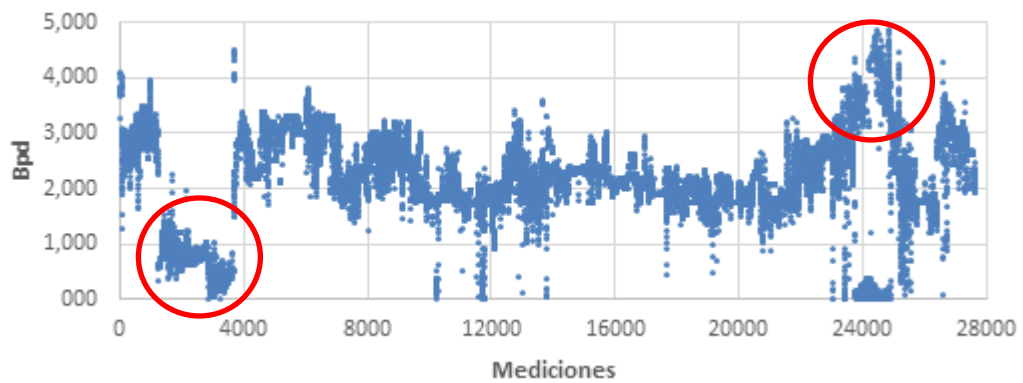


(c) Fondo de T201

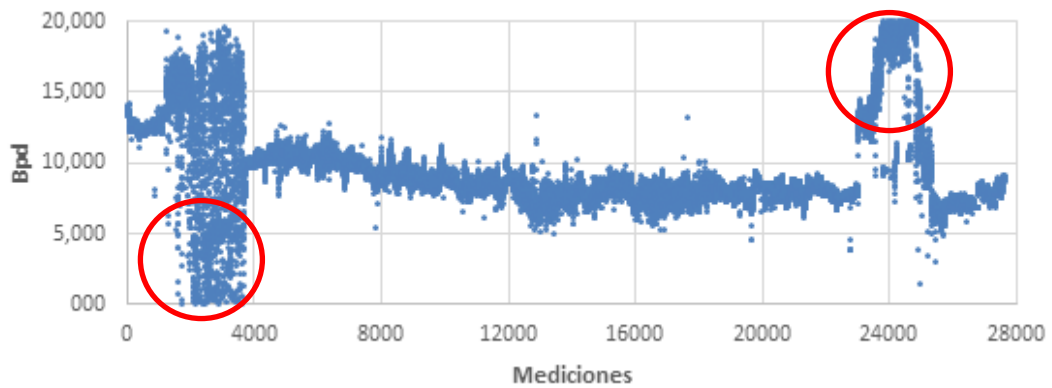
Figura 3. Carga, Jet y Fondo de la T204 para la ventana de operación entre enero de 2013 y agosto de 2016. Los círculos muestran periodos con alteraciones en la carga a la torre.



(a) Carga a T204



(b) Jet de T204

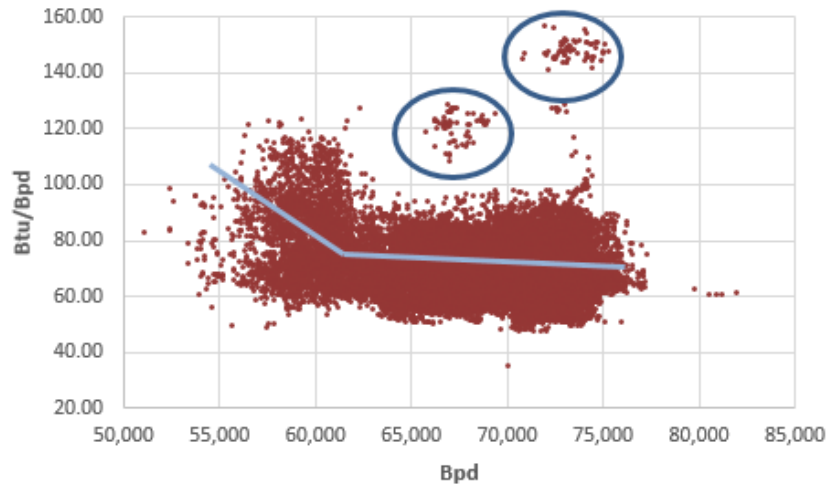


(c) Fondo de T204

De otro lado, la torre T204 durante los periodos de operación en modo viscorreductora presentó una carga baja (entre 18 y 23 kbpd) referente a febrero-agosto de 2013 y otro periodo de carga alta (entre 27 y 33 kbpd) referente a enero-mayo de 2016. Durante el periodo de carga baja la operación de la T204 presentó fluctuaciones importantes en los productos de fondo debido a la manipulación de los reciclos. También, para el periodo de carga alta, la T204 presenta rendimientos altos de nafta de viscorreductora y de disolventes, así como rendimientos altos en los productos de fondo como consecuencia de su operación en modo viscorreductora.

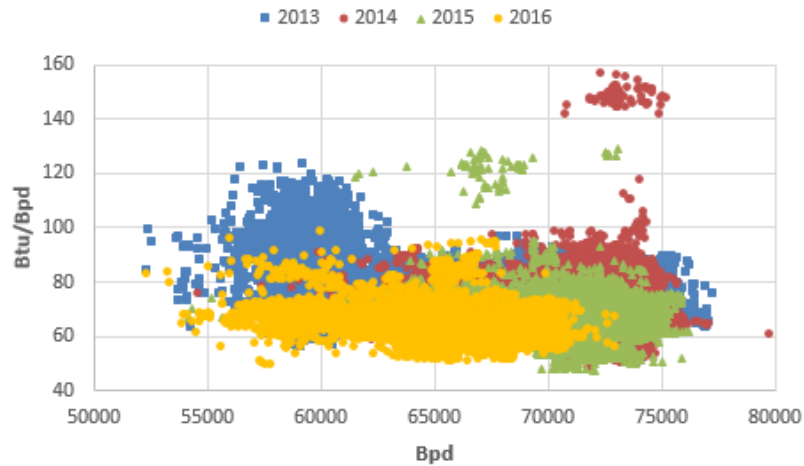
Respecto al consumo energético, la Figura 4 presenta la energía requerida por flujo de alimentación total a la U-200 –consumo energético específico– vs el flujo total de alimentación a la U-200 para la ventana de operación analizada. Según esta figura, los consumos energéticos específicos tienden a presentar valores mayores con flujos menores de alimentación (correlación de -0.18), sin embargo, esta tendencia es dispersa exhibiendo mayor desviación a mayores flujos de alimentación; asimismo, se presentan zonas con tendencia contraria (círculos azules). Con esto, la tendencia de consumo energético específico en la U-200 en la ventana de operación puede representarse de forma aproximada con dos pendientes, la primera con valor negativo hasta ca. 62 kbpd y la segunda, también de valor negativo pero con menor valor absoluto, hasta ca. 77 kbpd.

Figura 4. Energía requerida por la U-200 por flujo de carga de crudo procesada versus carga procesada en la ventana de operación.



Visualizando el consumo energético específico por año (Figura 5), se puede inferir que la U-200 durante el año 2013 reportó los mayores consumos energéticos para la operación con cargas inferiores a ca. 62 kbpd. Durante este año, la T-204 operó varios meses en modo de viscorreductora cargando principalmente fondos de vacío. La tendencia en el año 2013 conduce a la obtención de la correlación negativa de mayor valor absoluto en el consumo energético específico para el procesamiento de flujos de alimentación bajos. Para los años 2014, 2015 y 2016, la tendencia en el consumo de energía específica es también decreciente, con coeficientes de correlación de -0.13, -0.13 y -0.26, con consumos promedio de 72.47, 68.71 y 64.80 Btu/bpd y con desviaciones estándar de 10.58, 8.36 y 7.25 Btu/bpd, respectivamente. Salvo algunas excepciones en la operación para los años 2014 y 2015, el requerimiento energético específico de procesamiento se mantiene entre 55 y 90 Btu/bpd a cargas de procesamiento superiores a 62 kbpd.

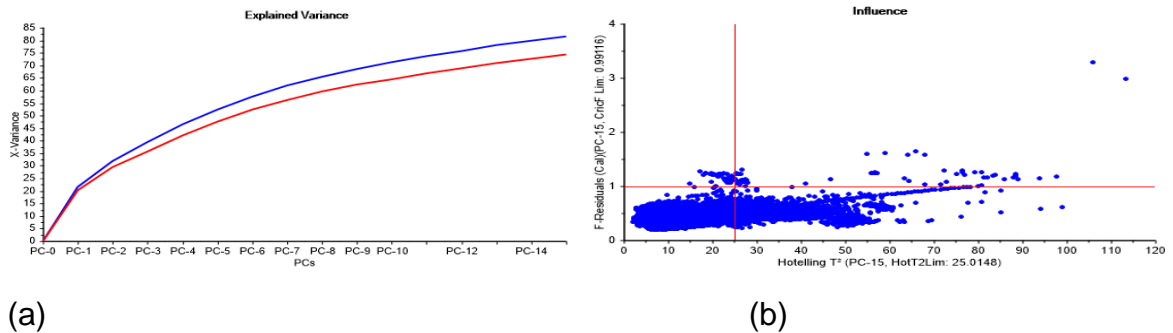
Figura 5. Energía requerida por la U-200 por flujo de carga de crudo procesada versus carga procesada por año en la ventana de operación.



4.2 ESQUEMAS OPERATIVOS

Los datos históricos de proceso de la U-200 fueron pretratados utilizando el procedimiento de autoescalado con los valores de la media y la desviación estándar según la variable en turno. Los datos autoescalados fueron clasificados según el método *k-mean* utilizando la distancia euclidiana y definiendo las clases 2, 3, 4 y 5. El análisis PCA fue aplicado a la matriz de datos autoescalados para obtener 15 componentes principales a partir de las variables de proceso iniciales. Según los resultados del PCA, los 10 componentes principales reproducen un 80% de la varianza total de los datos históricos (Figura 6a), con una excelente concordancia entre los procedimientos de calibración (línea azul) y validación (línea roja). Sin embargo, el gráfico de influencia presentado en la Figura 6b muestra que más del 50% de los datos pueden no ser representativos del conjunto global –puntos con valor de T^2 superior a 25 para una significancia estadística del 95%–, con lo cual, la división de los datos en subconjuntos describiría mejor la varianza de los datos.

Figura 6. Análisis PCA a la base de datos histórica con pretratamiento de autoescalado. (a) varianza explicada. (b) Gráfico de influencia.

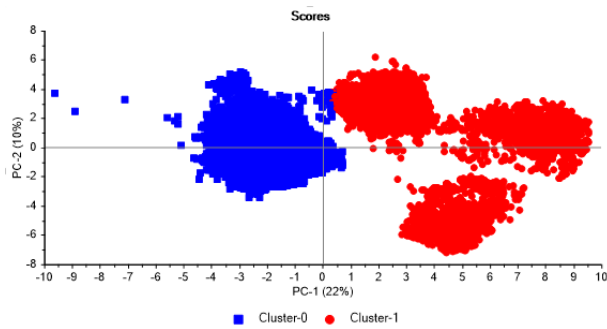


Según el gráfico de los *scores* para los dos primeros componentes principales (Figura 7), los cuales reproducen la mayor proporción de la varianza total de los datos históricos (32%), los valores exhiben discontinuidades que pueden reflejar diferentes esquemas operativos. En la Figura 7 se presentan sobre la gráfica de los *scores* los resultados del método de clasificación *k-mean* para determinar visualmente la ubicación de las diferentes agrupaciones o clases definidas. Según las Figuras 7a-7d, la ubicación de las agrupaciones coincide con las diferentes discontinuidades exhibidas por los *scores*. La Figura 7d muestra que la clase 2 corresponde a un subconjunto de la clase 1, en donde los valores del componente principal número 3 (PC3) son mayoritariamente positivos (Figura 7e). Los valores de los *loadings* muestra que este PC3 se encuentra influenciado principalmente por la destilación del crudo procesado en la U-200 (Figura 8). Debido a que el componente PC3 reproduce ca. 7% de la varianza total de los datos, las cinco agrupaciones muestran diferentes características operativas que representan distintos esquemas operativos energéticos para la U-200. La Figura 8b valida la influencia de las temperaturas para 10, 30, 50, 70 y 90% de evaporado del crudo procesado en la aparición de las clases o esquemas operativos; según esta figura, la clase 2 hace referencia a mayores temperaturas de evaporado, mientras que las clases 0 y 4 contienen muestras con las menores temperaturas de evaporado; las clases 1 y 3 presentan muestras con temperaturas intermedias de evaporado. De otro lado, un número de agrupaciones mayor que 5 por el método *k-mean* no

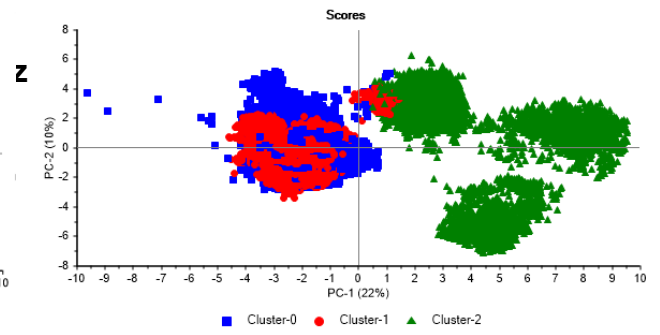
representa comportamientos importantes de las variables independientes del sistema.

Asimismo, las clases 0, 1 y 2 (Figuras 7d y 7e) se forman principalmente por los valores de las temperaturas y presiones del crudo en los hornos H-201 y H-202 y, por los valores de los flujos de fondo de la T-201 y de nafta de la T-204. Según la Figura 9, la clase 0 se encuentra constituida por aquellas operaciones en donde las temperaturas y presiones de los flujos de entrada a los hornos exhibieron sus mayores valores, mientras que las clases 1 y 2 reportan los menores valores para estas variables de operación. Respecto a los flujos, la clase 0 presenta los menores valores para el fondo T-201 y la nafta de la T-204, mientras que las clases 1 y 2 exhiben los mayores valores para estos flujos (Figura 10). Operativamente, esta clasificación se explica por el comportamiento de los consumos de gas de los hornos, los cuales se afectan por variaciones en temperatura y presión del crudo que ingresa y por la cantidad y calidad de los productos obtenidos; cabe resaltar que determinar cuáles variables operativas influyen en la clasificación permite realizar un mejor análisis energético de la unidad.

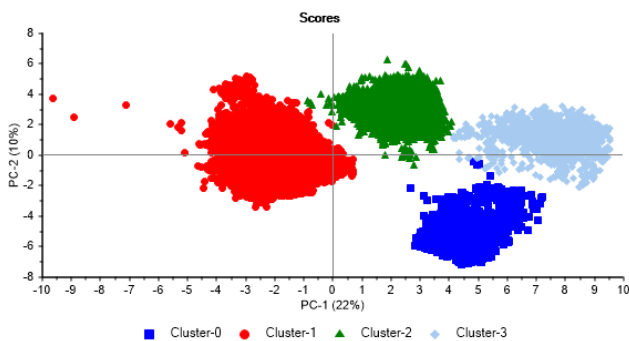
Figura 7. Gráfica de los scores para los dos primeros componentes principales derivados del PCA. Las agrupaciones fueron obtenidas por el método *k-mean* utilizando la distancia euclidiana. (d) Scores para los dos primeros PC con la elipse de Hotelling T^2 al 95% de confianza.



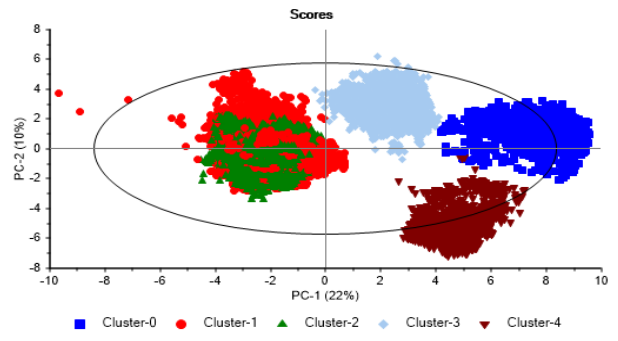
(a) Dos clases.



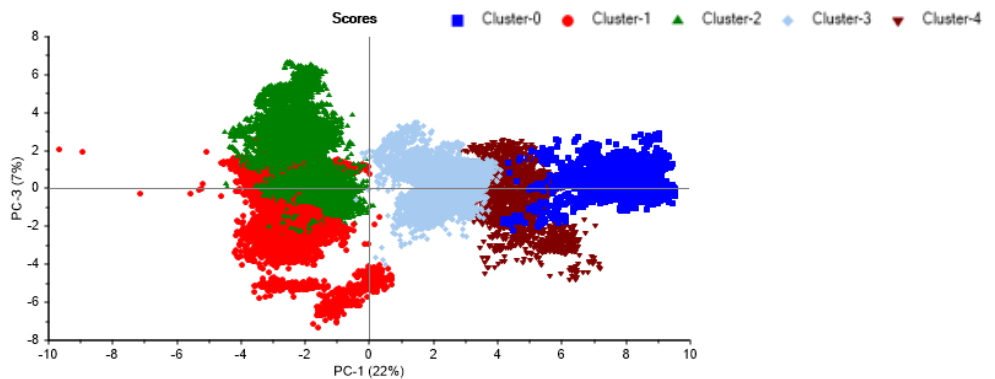
(b) Tres agrupaciones.



(c) Cuatro agrupaciones.



(d) Cinco agrupaciones.

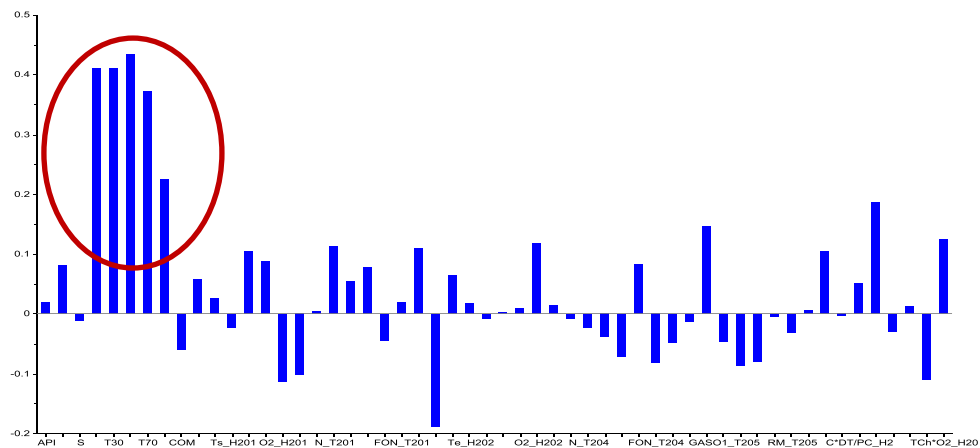


(e) Cinco agrupaciones en scores para PC1 versus PC3.

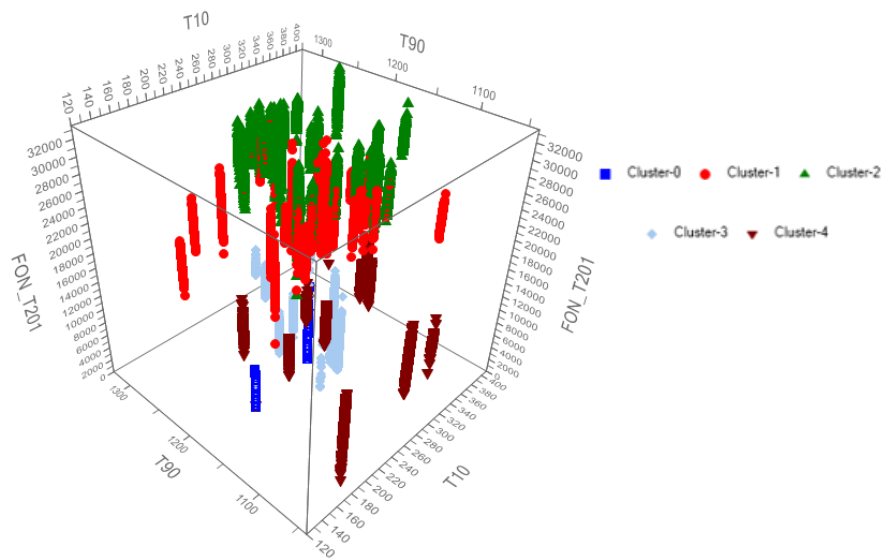
De otro lado, las clases 3 y 4 representan operaciones con valores bajos y valores elevados, respectivamente, para las temperaturas de chimenea de los hornos y

para el reflujo medio de la torre T-205 (Figura 11). Las temperaturas de chimenea están asociadas a la eficiencia de la combustión, mientras que el reflujo medio de la T-205 afecta el perfil de temperatura de la unidad de vacío. Las variables responsables de las clases definen esquemas operativos reportados durante la ventana de operación contenida en los datos históricos.

Figura 8. Gráfica de barras para los *loadings* del tercer componente principal (PC3). En el círculo de (a) se encierran en su orden de izquierda a derecha la destilación T10, T30, T50, T70 y T90.



(a)



(b)

Figura 9. Gráfica de la presión del flujo de entrada al H-201 versus la temperatura del flujo de entrada al mismo horno para las diferentes clases en la ventana operativa.

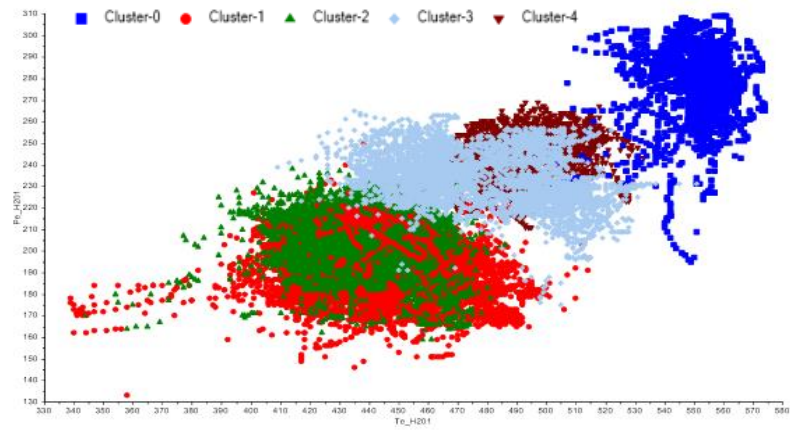


Figura 10. Gráfica de la presión del flujo de entrada al H-201 versus la temperatura del flujo de entrada al mismo horno para las diferentes clases en la ventana operativa.

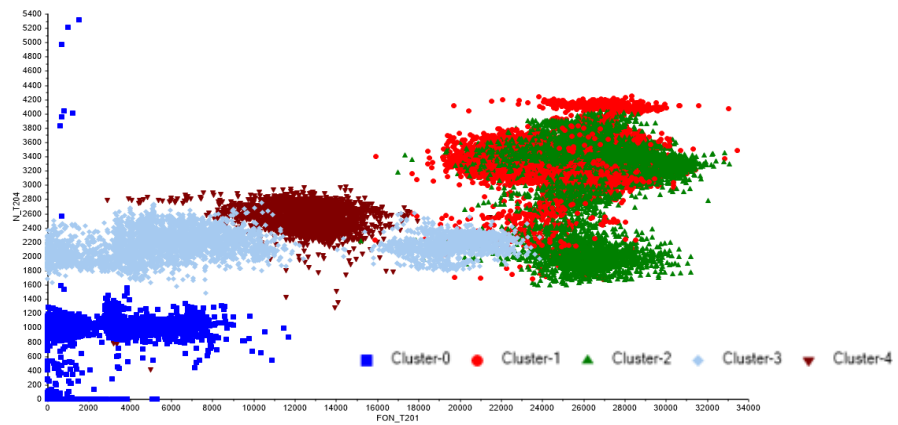
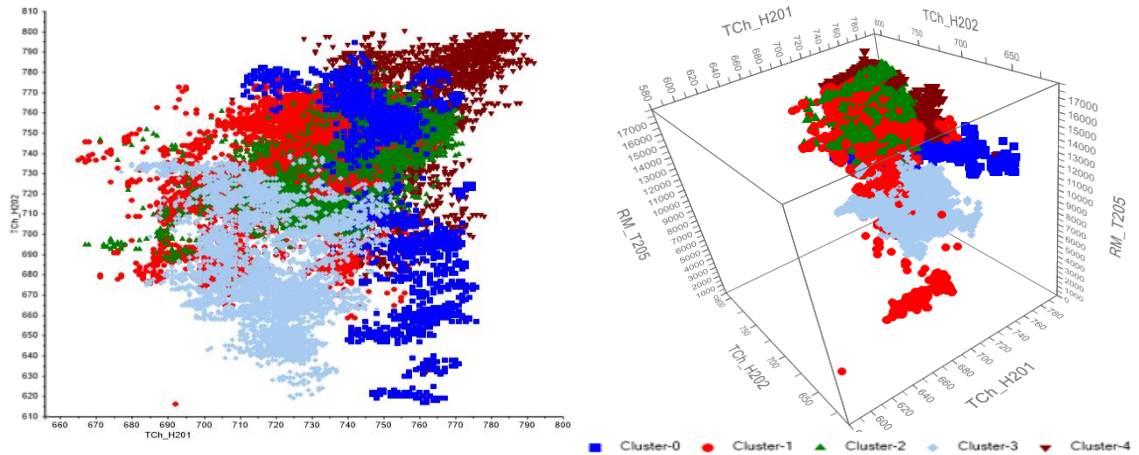
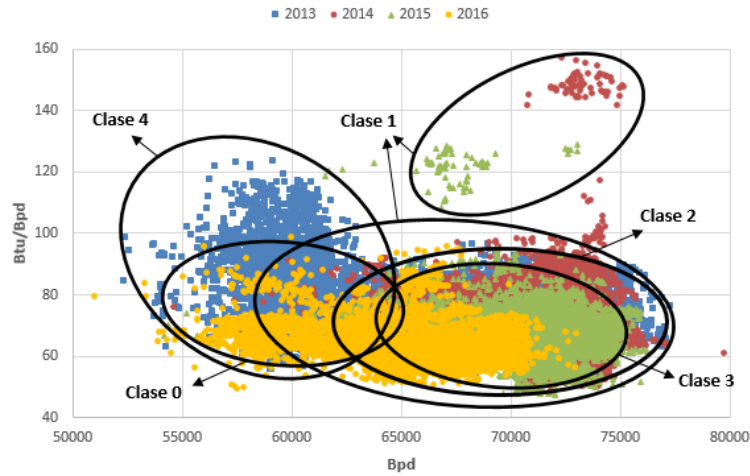


Figura 11. Gráfica temperatura de chimenea H-201 versus temperatura de chimenea H-202 versus reflujo medio T-205 para las diferentes clases en la ventana operativa.



Por otra parte, las clases obtenidas por el método *k-mean* están conformadas por datos de operación de los diferentes años, en concordancia con la Figura 12. Según esta figura, las clases 0 y 4 están caracterizadas por baja carga de procesamiento –asociadas a los periodos de operación en modo viscorreductora–, mientras que las clases 1, 2 y 3 presentan las mayores cargas de procesamiento de la U-200. También, la Figura 12 muestra que los mayores consumos de energía específica se presentan en la clase 4, excepto por algunos puntos de la clase 1; estos puntos de consumo elevado de energía específica en la clase 1 pueden deberse a salidas de servicio de intercambiadores del sistema de precaliente en los hornos.

Figura 12. Representación de las clases en el gráfico de consumo energético específico versus el crudo proceso en la U-200 en la ventana de operación.



Respecto al objetivo general del presente documento, el análisis anterior conduce a descartar una línea base para la U-200 por medio de una regresión sencilla del consumo energético con la carga de crudo a la unidad [19, 20]. Las 5 clases reportadas por el método *k-mean* y las tendencias y los valores para las variables representativas de cada clase deben ser consideradas para el desarrollo de regresiones que conduzcan a establecer líneas base para la U-200 con aplicaciones útiles. La siguiente tabla resume las condiciones operativas que definen cada clase para la U-200; la lista completa de los estadísticos para cada variable se encuentra en el anexo B.

Tabla 1 Relación entre las clases y algunas de las variables operacionales.

	Carga kbpd	Pe_H201 psi	Te_H201 °F	N_T204 kbpd	Fon_T201 kbpd	TCh_H201 °F	RM_T205 kbpd
Clase 0	55 – 68	240 – 310	520 – 580	<1.5	<10	710 – 775	6.5 – 13.5
Clase 1	59 – 76	160 – 230	340 – 500	1.7 – 4.3	20 – 32	670 – 770	8.0 – 17.0
Clase 2	63 – 77	170 – 240	340 – 480	1.7 – 4.0	20 – 32	665 – 770	8.0 – 17.5
Clase 3	60 – 74	200 – 265	430 – 530	1.6 – 2.4	<24	690 – 750	5.0 – 12.0
Clase 4	55 – 66	210 – 270	470 – 520	1.6 – 3.0	6 – 17	730 – 790	12.0 – 16.5

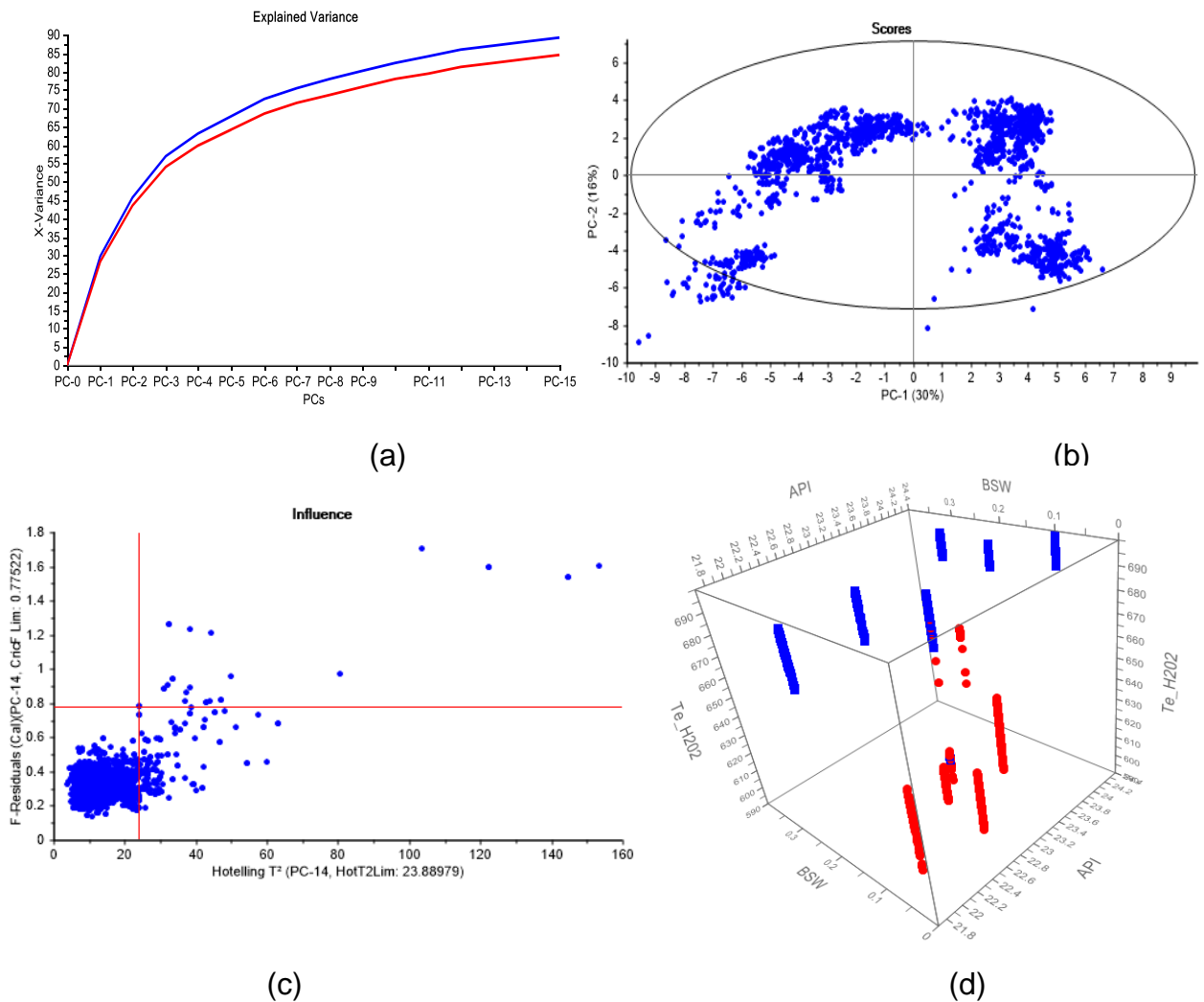
4.3 ANÁLISIS PCA EN CADA CLASE

El PCA define los componentes que representan la mayor varianza según la clase analizada. Para la clase 0, el PCA reproduce con 15 componentes principales ca. 88% de la varianza de las variables independientes (84% en validación), alcanzando una explicación del 46% con los dos primeros componentes principales (Figura 13a y b). También, una gran proporción de muestras (93%) presentan valores bajos de T^2 (Figura 13c, 23.8 al 95% confianza) con lo cual es posible suponer que todas las muestras son representativas de la clase 0. El gráfico de los scores para esta clase exhibe dos tendencias para esta clase relacionadas con operaciones opuestas en el horno H-202; los puntos con PC1 positivo representan la operación del H-202 con valores altos de temperatura (670 – 700°F), mientras que los puntos con PC1 negativo representan la operación contraria (600 – 650°F). Lo anterior conduce también a operaciones opuestas en la torre T-204; *i.e.* operaciones que conducen a flujos altos de livianos con flujos bajos de pesados en T-204 y viceversa. Una explicación a esta clase de operación se encuentra en las propiedades del crudo alimentado; en las fechas relacionadas con la clase 0, en la mayoría de las muestras, el H-202 opera a temperaturas bajas –inferiores a 650 °F–, cuando el crudo procesado presenta valores de BSW por debajo de 0.1 como se muestra en la Figura 13d. Desde la operación, esta clase se relaciona principalmente con el esquema viscorreductora; los valores bajos de BSW reflejan crudos de mejor calidad y por lo tanto habilitan que los fondos producidos a partir de estos crudos sean procesados en el H-202 a una temperatura menor.

Para la clase 1, la figura 14a muestra que el PCA reproduce ca. 85% de la varianza de las variables independientes (15 componentes principales), con una contribución del 26% con los dos primeros componentes principales (Figura 14b). De otro lado, el análisis de los *scores* y los *loadings* conduce a concluir que la clase 1 presenta dos casos, el primero muestra operaciones con baja producción

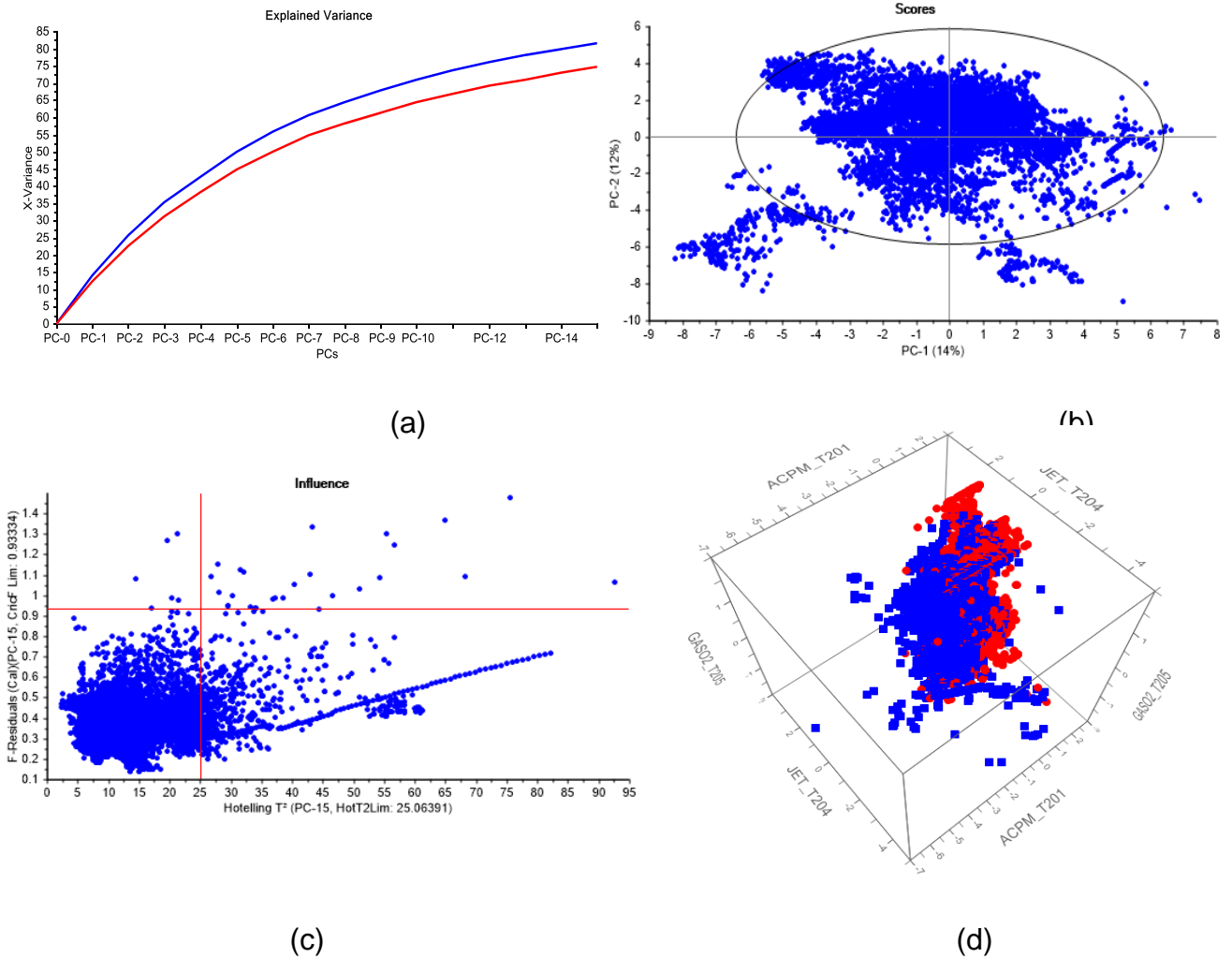
de acpm en la T-201 y baja producción de Jet en la T-204 (Figura 14d, puntos de color azul), mientras que para el segundo caso los flujos presentan la tendencia contraria (Figura 14d, puntos de color rojo).

Figura 13. Resultados PCA para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (30%) y PC2 (16% Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) temperatura de entrada H-202 versus API versus BSW.



Los puntos operativos del caso 1 presentan flujos de acpm en la T-201 mayores a 9600 bpd y flujos de jet en la T-204 superiores a 2450 bpd, mientras que los puntos del caso 2 presentan flujos de acpm en la T-201 menores a 9000 bpd y flujos de jet en la T-204 inferiores a 2200 bpd. Operativamente cuando se requiere maximizar la producción de acpm se desvía el Jet de la T-204 como una corriente adicional de acpm, por lo cual se disminuyen las restricciones de calidad y se pueden aumentar los flujos. De otro lado, de los valores de T^2 (Figura 14c, 25.1 al 95% confianza) es posible inferir que la mayoría de los puntos (95%) son representativos de la clase 1.

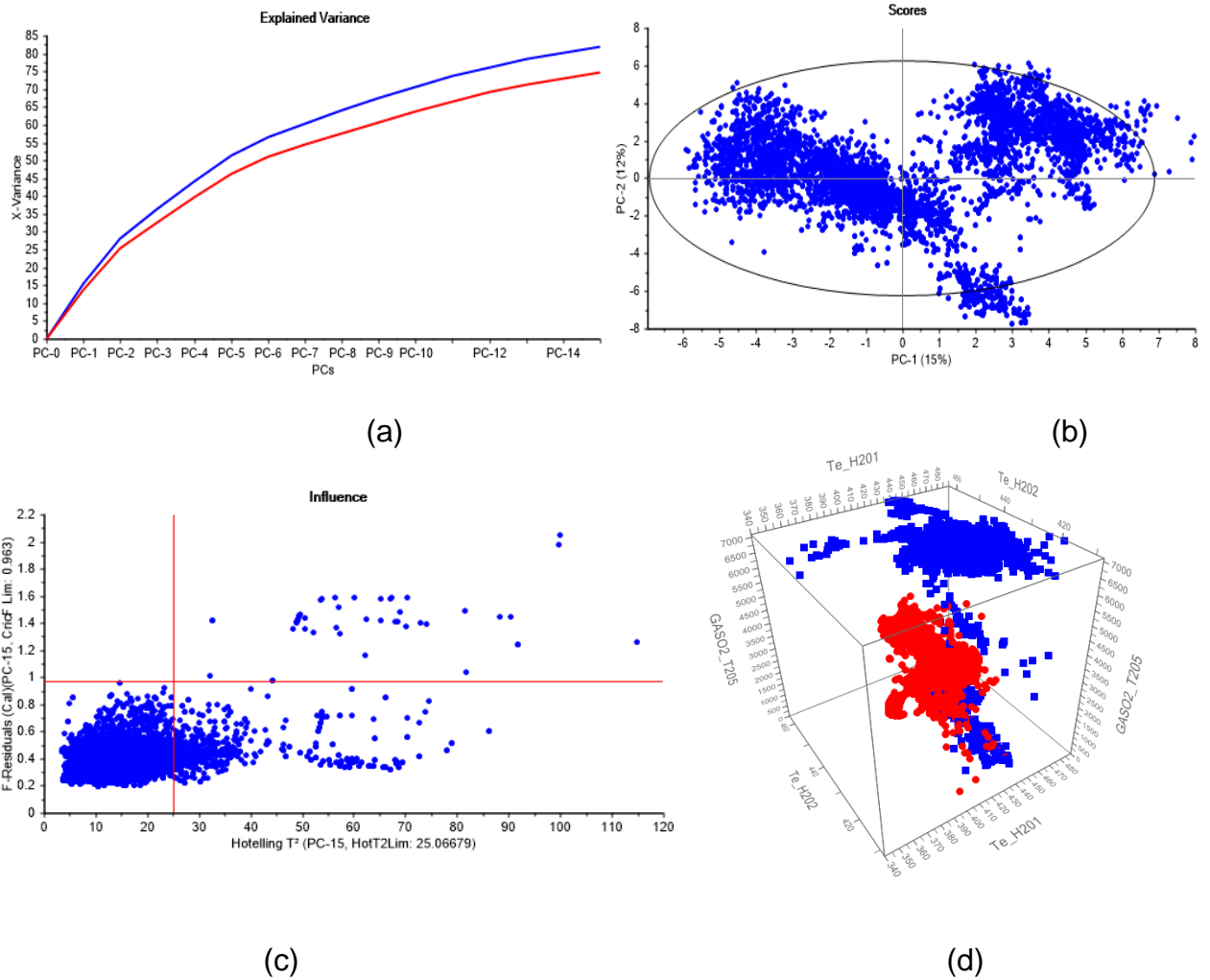
Figura 14. Resultados PCA para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (14%) y PC2 (12%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) flujo acpm T-201 versus flujo jet T-204 versus gasóleo T-205.



De otro lado, el PCA para la clase 2 muestra una explicación de la varianza con 15 componentes principales de ca. 85% (78% en validación) y con PC1 y PC2 se alcanza 27% de la varianza (Figuras 3.14a y 3.14b). En la clase 2 se tienen puntos de operación caracterizados por valores bajos en la temperatura del flujo de crudo al H-201, valores bajos de la temperatura del flujo de crudo al H-202 y valores

bajos de flujo de gasóleo en la T-205 (Figura 15d, puntos rojos) –caracterizados por valores positivos de PC1 (Figura 15b)–; los demás puntos de operación que constituyen la clase 2 presentan valores elevados de la temperatura de entrada al H-201 y valores altos de flujo de gasóleo de la T-205 (Figura 3.14d, puntos azules. Esta relación se explica principalmente las condiciones de temperatura de entrada y salida de los hornos; asimismo, determinan los flujos de crudo reducido de las T-201 y 204 los cuales son la carga de la T-205 y afectan el volumen de los productos de esta unidad, especialmente el gasóleo. De otro lado, de los valores los valores de T^2 (Figura 15c, 25.1 al 95% confianza) es posible establecer que ca. 87% de las muestras son representativas de la clase 2.

Figura 15. Resultados PCA para la clase 2. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (15%) y PC2 (12%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) temperatura de entrada H-201 versus temperatura de entrada H-202 versus flujo gasóleo T-205.



Respecto a la clase 3, el análisis PCA reporta una explicación de la varianza de las variables independientes de ca. 85% con 15 componentes principales (80% validación), alcanzando 33% con los dos primeros PC, según las Figuras 16a y 16b; asimismo, los valores de T^2 (Figura 16c, 25.1 al 95% confianza) sustentan la

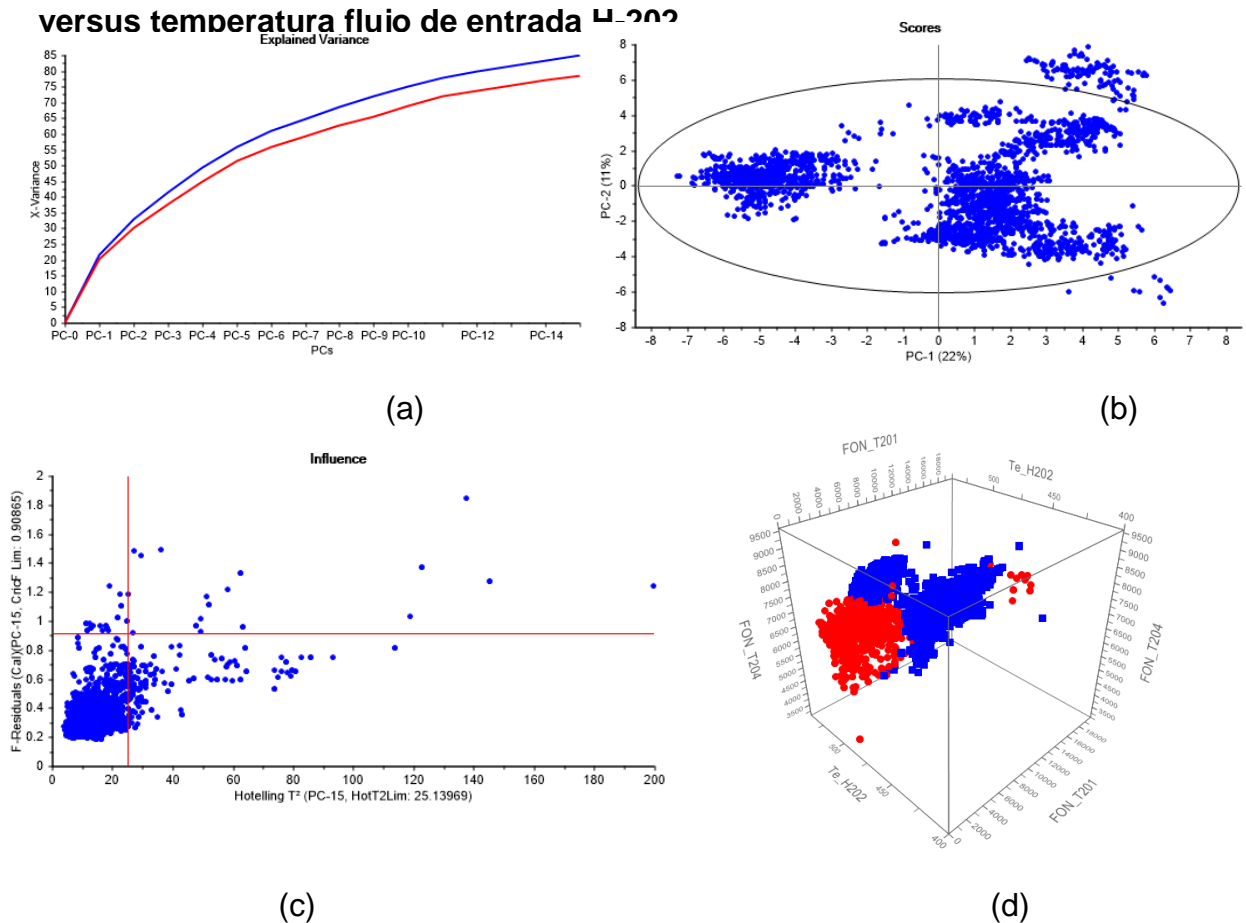
representatividad de ca. 97% de las muestras de la clase 3 (Figura 16c). Una condición operacional representativa de la clase 3 se encuentra en que presenta muestras con flujos bajos en los fondos de las T-201 –inferiores a 3000 bpd– y T-204 –inferiores a 8000 bpd– causadas por valores elevados de temperatura en los flujos de entrada a los hornos –superiores a 470°F– (valores negativos de PC1 en la Figura 16b y puntos de color rojo en la Figura 16d). De otro lado, la clase 3 también presenta muestras con condición operativa de flujos altos en los fondos de las T-201 –superiores a 3000 bpd– y T-204 –superiores a 8000 bpd– causadas en parte por valores bajos de temperatura en los flujos de entrada a los hornos –inferiores a 470°F– (valores positivos de PC1 en la Figura 16b y puntos de color azul en la Figura 16d). Operacionalmente, cuando la temperatura de entrada a los hornos es alta el sistema de precaliente está operando de forma adecuada, lo cual es aprovechado para aumentar la severidad y la temperatura de salida de los hornos y generar un aumento en la producción de livianos y disminuir los fondos en la T-201 y T-204.

Por último, el análisis PCA aplicado a la clase 4 conduce a la explicación del 90% de la varianza contenida en las variables independientes utilizando 15 componentes principales (84% validación), reportando una explicación del 43% de la varianza con los dos primeros PC (Figuras 17a y 17b). La clase 4 exhibe una separación marcada en dos condiciones operativas reflejada en los valores del primer componente principal (Figura 17b); la condición con PC1 positivo representa muestras con valores de poder calorífico del gas combustible a los hornos (COM) inferiores a 800 BTU/ft³, los cuales ocasionan valores de temperaturas inferiores a 760°F en los gases de chimenea de los hornos (Figura 17d, puntos rojos); de otro lado, la condición operativa con PC1 negativo representa valores altos de COM, lo cual genera temperaturas superiores a 760°F en los gases de chimenea de los hornos (Figura 17d, puntos azules). Esta condición se observa durante la operación de la T-204 como viscorreductora; los menores valores de poder calorífico aumentan el volumen de gas requerido a los

hornos, por lo cual se genera un aumento en el flujo de gases de chimenea y una disminución en su temperatura. Aunque se tienen las condiciones operativas en nubes dispersas en la Figura 3.16b de los *scores*, los valores de T^2 (Figura 14c, 25.3 al 95% confianza) sustentan la representatividad de ca. 97% de las muestras de la clase 4 (Figura 17c).

Asimismo, según las Figuras 6b, 13c, 14c, 15c, 16c y 17c, la división en clases con el método *k-mean* alcanza una mayor explicación de la varianza de las variables independientes incluyendo un mayor número de muestras que el conjunto total de valores de la base de datos histórica de la U-200 (Figura 6). Con lo anterior, un mejor ajuste de las regresiones multivariadas para el consumo energético específico puede ser obtenido dividiendo el conjunto de datos según las 5 clases o esquemas operativos (Figura 7d y 12).

Figura 16. Resultados PCA para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (22%) y PC2 (11%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) flujo fondo T-201 versus flujo fondo T-204 versus temperatura fluio de entrada



4.4 RLM PARA LA PREDICCIÓN DEL CONSUMO DE ENERGÍA EN CADA CLASE

4.4.1 Clase 0. El desempeño de la RLM obtenida con el método *backward elimination* (BE) para la clase 0 se muestra en la Figura 18. Según esta figura, el ajuste por RLM reproduce con un coeficiente de determinación corregido de 0.9764 (0.9764 validación) la varianza del consumo energético específico

autoescalado con un error medio (RMSE) de 0.154 (0.154 validación). La distribución de los residuales y la forma de la gráfica de probabilidad (Figuras 18b-c) cumplen los supuestos de varianza constante y distribución normal respectivamente. De otro lado, la tabla ANOVA (Tabla 2) valida la relación matemática de la regresión y la significancia estadística de la relación entre las variables independientes y el consumo energético específico (*p-value*) al 95% confianza. Esta tabla también presenta el valor de los coeficientes para la regresión (*B-coefficients*), que en total corresponden a 42 variables, de las cuales el poder calorífico del combustible (COM) es la de mayor contribución con 10.4% del total del valor del consumo energético específico.

Figura 17 Resultados PCA para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) scores para PC1 (28%) y PC2 (15%) con la elipse de Hotelling T^2 al 95% de confianza. (c) Gráfico de influencia para las muestras 95% de confianza; (d) poder calorífico combustible (COM) versus temperatura de chim. H-202 versus temperatura flujo de entrada H-201.

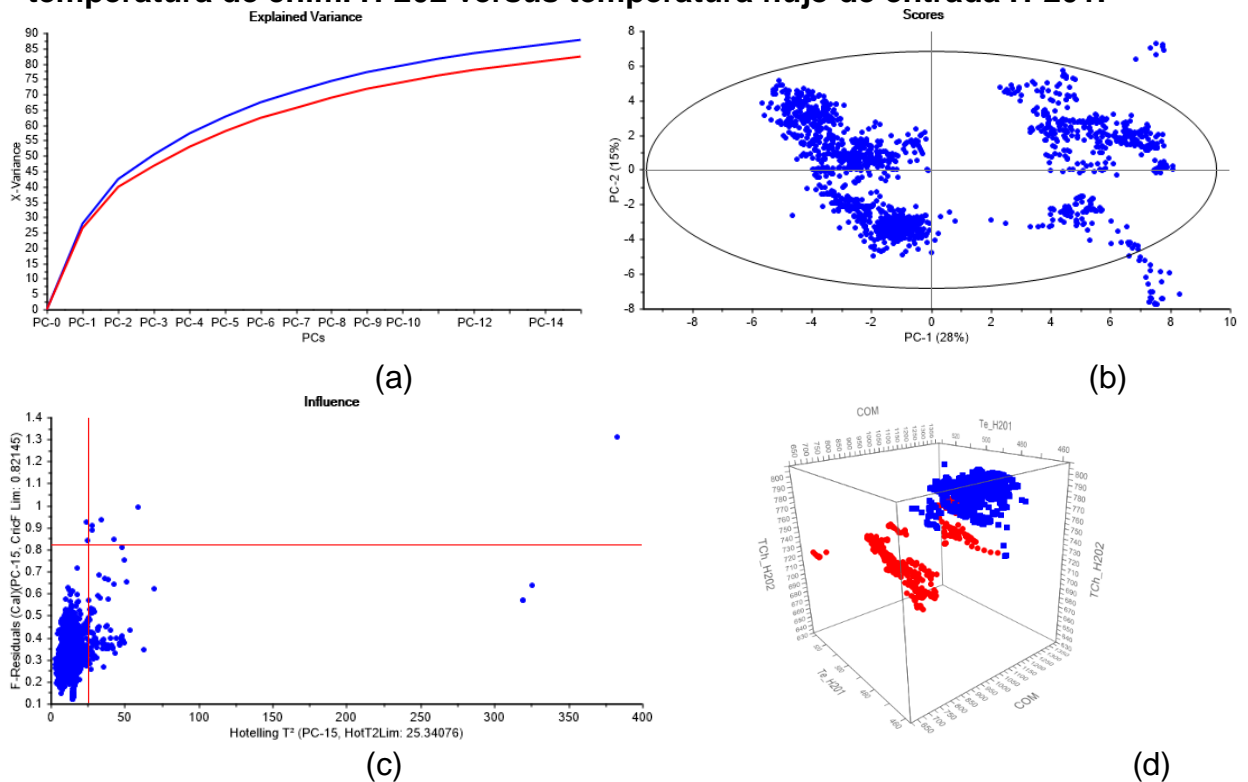


Figura 18. Regresión RLM para la clase 0. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.

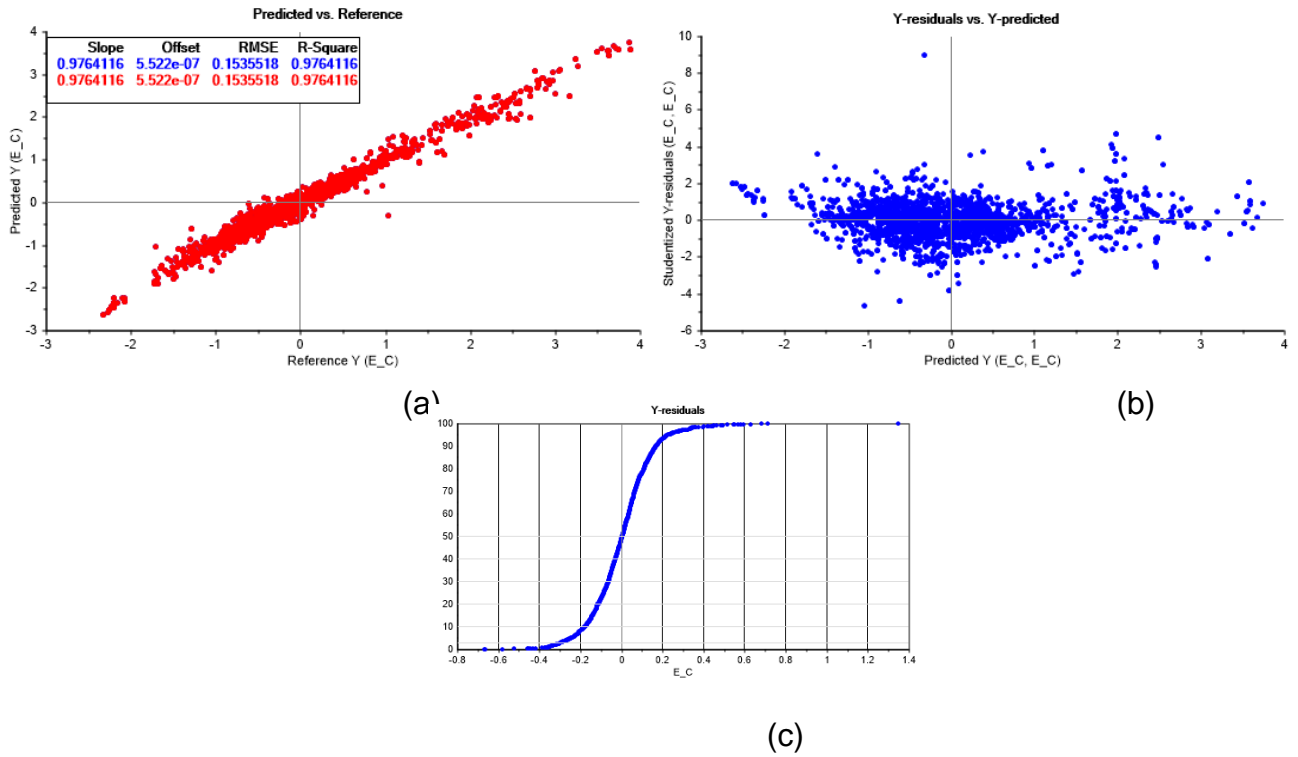


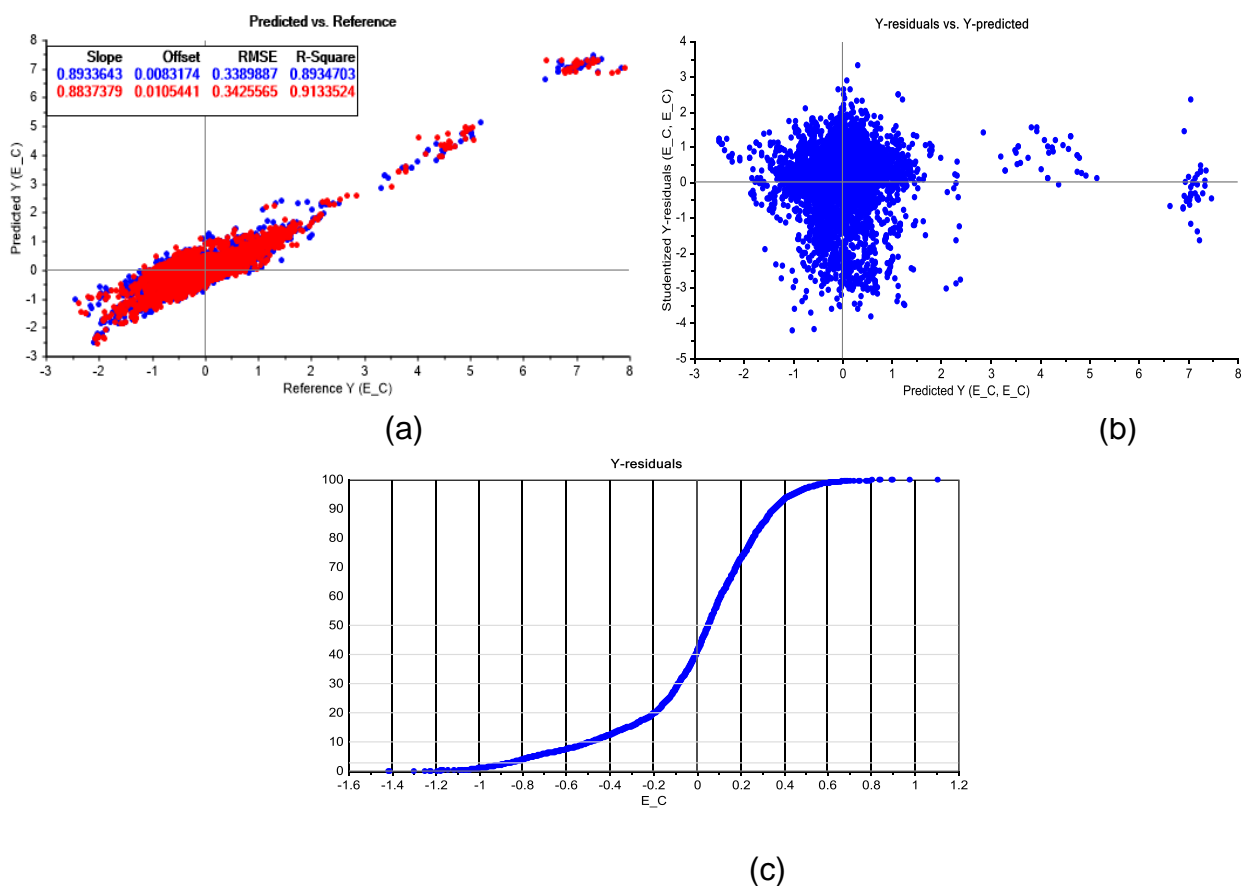
Tabla 2. ANOVA para la RLM de los datos de la clase 0 (Tomada de Unscrambler).

Anova Table							
Multiple Correlation: 0.9881355 (cal) 0.9881355 (val)							
R-Square: 0.9764116 (cal) 0.9764116 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	2255.51099	42	53.7026443	2236.24927	0		
Error	54.4891396	2269	0.0240146052				
Total	2310.00024	2311	0.999567389				
Variables							
Intercept	0	0				0	0
BSW	4.35060692	1	4.35060692	181.164185	8.69900863E-40	0.0941135138	0.0069922
S	6.11839676	1	6.11839676	254.778107	0	-0.122249916	0.0076589
T90	2.26435566	1	2.26435566	94.2906189	7.21212889E-22	-0.0790209249	0.0081378
COM	96.3556519	1	96.3556519	4012.37598	0	0.749862373	0.0118380
G_H201	98.0953064	1	98.0953064	4084.80884	0	0.509215891	0.0079673
Te_H201	0.137921423	1	0.137921423	5.74323225	0.0166327916	-0.0179866124	0.0075053
Ts_H201	0.587130547	1	0.587130547	24.4488621	8.18961382E-07	0.0254399832	0.0051450
Pe_H201	0.357835382	1	0.357835382	14.9006653	0.000116489529	0.330924332	0.0857286
TCh_H201	13.6731596	1	13.6731596	569.36792	0	-0.179464325	0.0075211
G_T201	0.329685539	1	0.329685539	13.728487	0.000216276967	0.0251365136	0.0067841
JET_T201	1.1242274	1	1.1242274	46.8143425	1.00037079E-11	-0.0557641983	0.0081501
ACPM_T201	0.589131176	1	0.589131176	24.5322056	7.84655072E-07	-0.0348169878	0.0070294
FON_T201	20.0280876	1	20.0280876	833.995422	0	-0.312394828	0.0108173
RC_T201	0.865159869	1	0.865159869	36.0263901	2.26006858E-09	-0.0414454974	0.0069050
RM_T201	1.66268826	1	1.66268826	69.2365189	1.48763748E-16	0.0577441268	0.0069396
Te_H202	3.22323179	1	3.22323179	134.219299	3.39945819E-30	-0.588538826	0.0508004
Ts_H202	2.84247255	1	2.84247255	118.364708	6.57843281E-27	0.130608305	0.0120049
Pe_H202	0.567411602	1	0.567411602	23.6277218	1.24889482E-06	0.404456466	0.0832071
TCh_H202	0.493825227	1	0.493825227	20.5635719	6.06859567E-06	0.0953723118	0.0210316
O2_H202	1.07124472	1	1.07124472	44.6080284	3.01755843E-11	0.185596123	0.0277883
TfI_T204	0.824013472	1	0.824013472	34.3129616	5.37524691E-09	-0.0396631733	0.0067710
G_T204	0.753052294	1	0.753052294	31.3579903	2.40515341E-08	0.0770224854	0.0137544
ACPM_T204	3.89294004	1	3.89294004	162.107132	6.43910608E-36	0.0718057975	0.0056397
GASO_T204	4.61965942	1	4.61965942	192.368698	4.78123036E-42	0.127209514	0.0091717
FON_T204	11.3466358	1	11.3466358	472.488525	0	-0.127392396	0.0058606
RC_T204	3.41128445	1	3.41128445	142.050461	8.26119445E-32	-0.118513457	0.0099436
RM_T204	0.349184513	1	0.349184513	14.5405197	0.000140835604	-0.0232270118	0.0060912
RI_T204	0.414122939	1	0.414122939	17.2446365	3.40776132E-05	-0.0324546546	0.0078153
GASO1_T205	0.263063937	1	0.263063937	10.9543133	0.000948471308	-0.0543571897	0.0164234
GASO2_T205	0.984441996	1	0.984441996	40.9934807	1.85033849E-10	0.0398350358	0.0062216
FON_T205	0.099502705	1	0.099502705	4.1434207	0.0419126004	-0.0120506361	0.0059201
RC_T205	3.09436917	1	3.09436917	128.853622	4.37474535E-29	-0.128924146	0.0113575
RM_T205	0.165663674	1	0.165663674	6.89845181	0.0086850673	-0.022764815	0.0086673
RI2_T205	0.81154108	1	0.81154108	33.7936172	6.9919337E-09	-0.0390729606	0.0067213
DP_H201	0.172131404	1	0.172131404	7.16773844	0.00747609418	-0.226808384	0.0847165
DP_H202	1.61451161	1	1.61451161	67.2300491	3.99485831E-16	-0.602463901	0.0734766
C*DT/COM_H201	0.103822984	1	0.103822984	4.32332468	0.0377054699	-0.00869149622	0.0041800
C*DP/COM_H201	8.96945763	1	8.96945763	373.50061	0	-0.384873241	0.0199146
C*DT/COM_H202	1.11119068	1	1.11119068	46.2713127	1.31247877E-11	-0.544486761	0.0800444
C*DP/COM_H202	0.356927872	1	0.356927872	14.8629169	0.000118828524	0.279864818	0.0725932
TCh*O2_H201	1.76869476	1	1.76869476	73.6507339	1.7004428E-17	0.0660794154	0.0076997
TCh*O2_H202	0.494542032	1	0.494542032	20.5933933	5.97563576E-06	-0.123034582	0.0271121

4.4.2 Clase 1. La Figura 19 presenta los resultados de la regresión RLM obtenida con el método BE aplicada a las muestras de la clase 1. Según la Figura 19a el ajuste por RLM presenta un R2 corregido de 0.8934 (0.9133 validación externa) con un error medio (RMSE) de 0.334 (0.343 validación). La distribución de los residuales y la forma de la gráfica de probabilidad (Figuras 19b y 19c) cumplen los

supuesto para RLM. Asimismo, la tabla ANOVA (Tabla 3.3) valida la relación matemática de la regresión y la significancia estadística de la relación entre las variables independientes y el consumo energético específico ($p\text{-value}<0.05$) al 95% confianza. En esta tabla se presentan los coeficientes para la regresión (*B-coefficients*) para las 38 variables significativas, de las cuales el oxígeno al H-201 (O2_H202) y la variable auxiliar TCh*O2_H201 presentan la mayor contribución (18% cada una) en la estimación del valor del consumo energético específico. Lo anterior corrobora los resultados del PCA aplicado al conjunto de datos total (sección 3.2).

Figura 19. Regresión RLM para la clase 1. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.



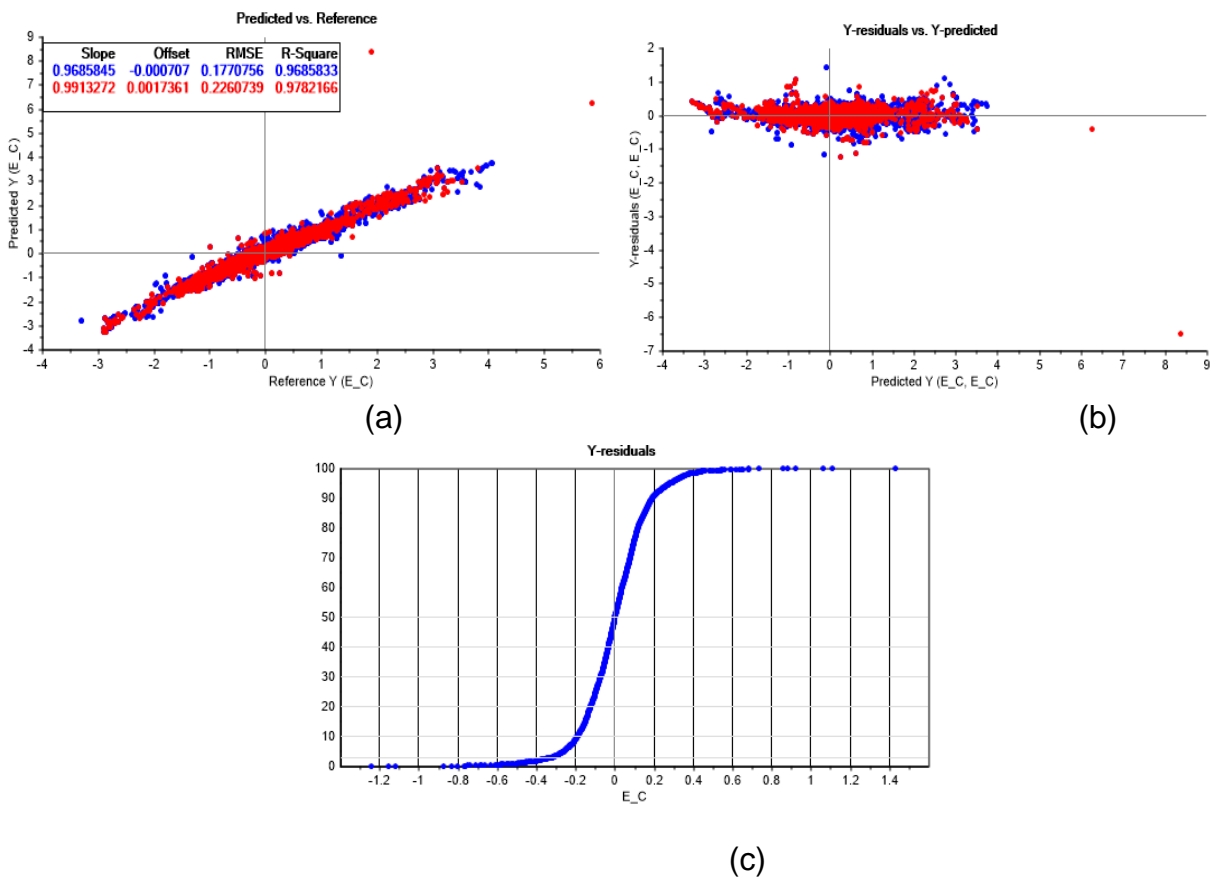
4.4.3 Clase 2. Según las Figuras 20a, 20b y 20c la regresión RLM resultante del método BE aplicado a las muestras de la clase 2 cumple los supuestos de normalidad y varianza constante, con un R2 corregido de 0.9685 (0.9782 validación externa) y un RMSE de 0.177 (0.2660 validación externa). Según la Tabla 4, los resultados de RLM son estadísticamente significativos al 95% confianza. La regresión requiere de 48 variables autoescaladas, de las cuales, las variables TCh*O2_H202, O2_H202, TCh*O2_H201 y O2_H201 presentan la mayor contribución (10.5, 9.6, 9.3 y 8.8%, respectivamente).

Tabla 3. ANOVA para la RLM de los datos de la clase 1 (Tomada de Unscrambler).

Anova Table							
Multiple Correlation: 0.9452374 (cal) 0.9448305 (val)							
R-Square: 0.8934703 (cal) 0.9133524 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	4159.81445	37	112.427422	969.948364	0		
Error	494.127411	4263	0.115910724				
Total	4653.94189	4300	1.08231211				
Variables							
Intercept	0	0				0	0
API	27.3726845	1	27.3726845	236.154144	0	-0.131170973	0.0085357
BSW	23.3798313	1	23.3798313	201.705948	9.80908925E-45	0.143492997	0.0101034
S	0.823139489	1	0.823139489	7.10151434	0.0077309045	-0.0226531494	0.0085006
T10	18.8682594	1	18.8682594	162.782639	1.29418438E-36	0.125716045	0.0098534
T50	16.0573578	1	16.0573578	138.531952	1.70456155E-31	-0.213041469	0.0181004
T70	22.3339577	1	22.3339577	192.682495	6.99247934E-43	0.207361162	0.0149384
T90	13.7795162	1	13.7795162	118.880623	2.544134E-27	0.0870148167	0.0079806
COM	304.271973	1	304.271973	2625.04956	0	0.819227397	0.0159895
G_H201	197.99231	1	197.99231	1708.14819	0	0.59404999	0.0143734
Ts_H201	3.03841496	1	3.03841496	26.2134171	3.19171363E-07	0.0433880799	0.0084743
TCh_H201	1.0407927	1	1.0407927	8.97942543	0.00274609402	0.0610949807	0.0203883
O2_H201	2.20745707	1	2.20745707	19.0447655	1.30710041E-05	1.4003042	0.3208740
G_T201	2.79355621	1	2.79355621	24.1008644	9.4837668E-07	-0.0543588623	0.0110727
N_T201	0.665912271	1	0.665912271	5.74504709	0.0165780149	-0.0213789772	0.0089194
JET_T201	0.604212701	1	0.604212701	5.2127409	0.0224709567	-0.016026929	0.0070196
ACPM_T201	1.04049039	1	1.04049039	8.97664165	0.00275027123	-0.0338308848	0.0112916
FON_T201	25.6193447	1	25.6193447	221.027191	0	-0.159456819	0.0107255
RM_T201	7.55736446	1	7.55736446	65.1999512	8.72677929E-16	0.0608988516	0.0075419
RI_T201	2.55083537	1	2.55083537	22.0068684	2.80184486E-06	-0.0378466286	0.0080676
Te_H202	1.69748259	1	1.69748259	14.6447735	0.000131653738	0.0769605413	0.0201106
TCh_H202	1.65132654	1	1.65132654	14.2465668	0.000162533703	0.0485445186	0.0128613
TfI_T204	3.64439058	1	3.64439058	31.4413776	2.18551772E-08	0.057434503	0.0102428
G_T204	1.81972957	1	1.81972957	15.6993933	7.54619905E-05	0.0263653342	0.0066541
FON_T204	24.1271305	1	24.1271305	208.152908	0	-0.225480005	0.0156284
RC_T204	8.7415514	1	8.7415514	75.416214	5.3476325E-18	-0.082691215	0.0095219
RM_T204	2.79768252	1	2.79768252	24.136591	9.31041768E-07	-0.0642230064	0.0130723
RI_T204	11.4065638	1	11.4065638	98.408371	6.02362891E-23	-0.113878816	0.0114796
GASO1_T205	9.30050087	1	9.30050087	80.2384567	4.86303101E-19	-0.0990084857	0.0110530
GASO2_T205	1.24265909	1	1.24265909	10.7208223	0.00106782361	-0.0269640442	0.0082351
FON_T205	1.93113804	1	1.93113804	16.660574	4.55190275E-05	0.0228940658	0.0056089
RI2_T205	4.62674332	1	4.62674332	39.9164429	2.92199626E-10	-0.0709406808	0.0112284
DP_H201	5.88307762	1	5.88307762	50.7551956	1.22269989E-12	0.146432191	0.0205539
C*DT/PC_H201	15.0001287	1	15.0001287	129.411026	1.46546868E-29	-0.258892983	0.0227580
C*DP/PC_H201	7.51508713	1	7.51508713	64.8349228	1.04735857E-15	-0.309237391	0.0384049
C*DT/PC_H202	28.6219482	1	28.6219482	246.931046	0	0.414200544	0.0263586
TCh*O2_H201	2.21627569	1	2.21627569	19.1208191	1.25625866E-05	-1.38980579	0.3178343
TCh*O2_H202	1.15989137	1	1.15989137	10.0068007	0.0015705945	-0.0249251816	0.0078793

4.4.4 Clase 3. De las Figuras 3.20a - 3.20c se sigue que la regresión obtenida con el método BE utilizando las muestras de la clase 3 cumple los supuestos de la RLM, con un R2 corregido de 0.980 (0.9792 validación cruzada) y un RMSE de 0.141 (0.144 validación cruzada). Según la tabla ANOVA (Tabla 5), los resultados de RLM son estadísticamente significativos al 95% confianza. La regresión requiere de 50 variables autoescaladas, de las cuales, las variables temperatura de evaporado T50, T30 y T70 presentan la mayor contribución (22.2, 13.2 y 10.1%, respectivamente).

Figura 20. Regresión RLM para la clase 2. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.



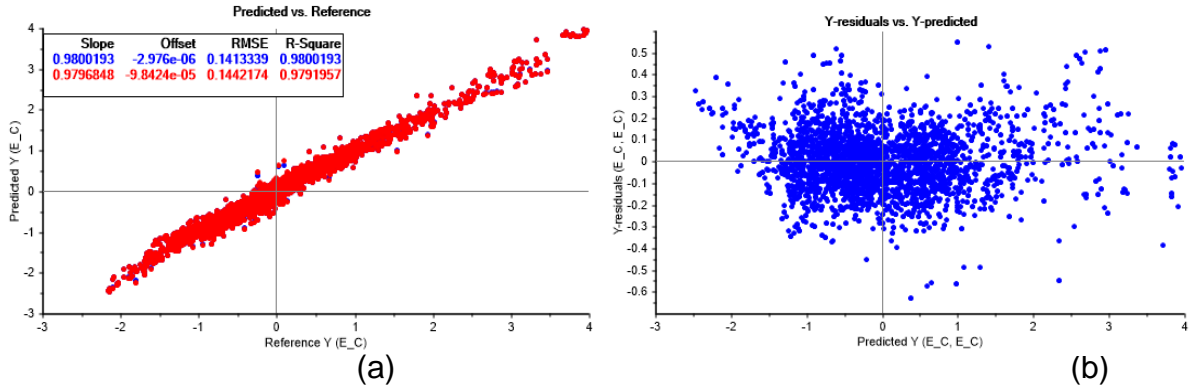
4.4.5 Clase 4. La regresión aplicada a las muestras de la clase 4 (RLM obtenida con el método BE) cumple los supuestos de la RLM (Figuras 3.21a - 3.21c); los estadísticos de la regresión son R^2 corregido de 0.9286 (0.9260 validación cruzada) y un RMSE de 0.267 (0.272 validación cruzada). La tabla ANOVA (Tabla 3.6) confirma la significancia estadística de los resultados de RLM (95% confianza). Según esta tabla, la regresión requiere de 28 variables autoescaladas, de las cuales, las variables C*DP/COM_H201, fondos T204 (Fon_T204) y C*DT/COM_H201 presentan la mayor contribución (12.9, 10.2 y 10.1%, respectivamente).

Las regresiones RLM aplicadas a cada clase reportaron desempeños con valores de R^2 y RMSE suficientes para su aplicación como línea base energética [21] de la unidad U-200. Estas regresiones para las 5 clases pueden ser aplicadas según los límites de cada variable (Anexo B), sin embargo, el número de variables independientes es elevado y puede limitar su aplicación práctica.

Tabla 4. ANOVA para la RLM de los datos de la clase 2 (Tomada de Unscrambler).

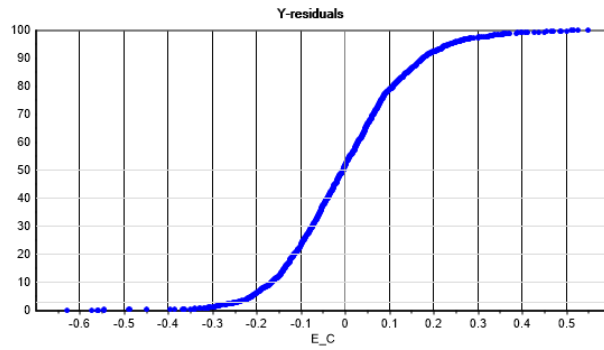
Anova Table							
Multiple Correlation: 0.9841666 (cal) 0.9751034 (val)							
R-Square: 0.9685833 (cal) 0.9782166 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	5026.87793	47	106.954849	3380.18018	0		
Error	163.049988	5153	0.031641759				
Total	5189.92773	5200	0.998063028				
Variables							
Intercept	0	0				0	0
API	0.444898188	1	0.444898188	14.06071	0.000178943956	-0.0210573971	0.0056156
BSW	0.329044372	1	0.329044372	10.399045	0.00126867264	0.0162130296	0.0050276
S	0.798318982	1	0.798318982	25.2299423	5.26031386E-07	-0.0274400227	0.0054629
T10	0.247180879	1	0.247180879	7.81188011	0.00520958845	-0.0282678455	0.0101138
T30	0.495734721	1	0.495734721	15.6673117	7.65346776E-05	0.132650733	0.0335129
T50	0.445600301	1	0.445600301	14.0828924	0.000176851594	-0.0999457166	0.0266328
T90	5.33502817	1	5.33502817	168.607117	5.84245387E-38	0.0657360926	0.0050625
COM	74.094429	1	74.094429	2341.65625	0	0.666045129	0.0137639
G_H201	121.299736	1	121.299736	3833.53955	0	0.400291681	0.0064651
Te_H201	0.629554391	1	0.629554391	19.8966103	8.34840102E-06	-0.0804852545	0.0180437
Ts_H201	0.967031181	1	0.967031181	30.5623474	3.39224151E-08	0.0710286275	0.0128481
TCh_H201	0.239441484	1	0.239441484	7.56725693	0.00596457999	0.0486719459	0.0176933
O2_H201	1.09942949	1	1.09942949	34.7461891	3.9949084E-09	0.792623997	0.1344663
TfI_T201	0.289840788	1	0.289840788	9.16008186	0.00248573767	-0.0102846092	0.0033981
N_T201	0.881984711	1	0.881984711	27.8741474	1.34781374E-07	-0.0191348922	0.0036243
ACPM_T201	0.188989162	1	0.188989162	5.97279215	0.0145615637	-0.0137175182	0.0056128
GASO_T201	4.37379646	1	4.37379646	138.228928	1.63599532E-31	-0.063633807	0.0054123
FON_T201	17.8140697	1	17.8140697	562.991882	0	-0.107544012	0.0045324
RM_T201	1.92381096	1	1.92381096	60.7997437	7.59329801E-15	0.0261748899	0.0033568
RI_T201	2.07630992	1	2.07630992	65.6194305	6.7710125E-16	0.0415870547	0.0051338
Te_H202	5.53424644	1	5.53424644	174.904221	2.72702771E-39	-0.2621108	0.0198191
Ts_H202	2.60648274	1	2.60648274	82.3749771	1.57150306E-19	0.0731071234	0.0080549
Pe_H202	0.412523538	1	0.412523538	13.0373106	0.000308250514	0.0882323235	0.0244362
TCh_H202	2.74079704	1	2.74079704	86.6196136	1.90100564E-20	0.143557461	0.0154265
O2_H202	1.18133616	1	1.18133616	37.3348389	1.06798048E-09	0.820554912	0.1342921
TfI_T204	0.327473044	1	0.327473044	10.3494234	0.00130316219	0.0139596881	0.0043392
N_T204	1.75456142	1	1.75456142	55.4510002	1.1170642E-13	-0.0450579599	0.0060508
JET_T204	3.53656888	1	3.53656888	111.76915	7.37432413E-26	-0.0472267233	0.0044671
ACPM_T204	1.21366048	1	1.21366048	38.3562431	6.35101249E-10	-0.0290877968	0.0046966
GASO_T204	1.80094481	1	1.80094481	56.9167099	5.34340062E-14	-0.0310167	0.0041112
FON_T204	30.1459808	1	30.1459808	952.728394	0	-0.150220588	0.0048668
RM_T204	1.99989867	1	1.99989867	63.2043915	2.27198377E-15	-0.0414155461	0.0052094
RI_T204	0.231956035	1	0.231956035	7.33069038	0.00680106552	-0.0159177128	0.0058790
GASO1_T205	23.1792927	1	23.1792927	732.553284	0	-0.108199544	0.0039976
GASO2_T205	1.95350289	1	1.95350289	61.737999	4.74131901E-15	-0.0526071005	0.0066952
FON_T205	1.25276327	1	1.25276327	39.5920715	3.3882952E-10	-0.0196951404	0.0031300
RC_T205	0.217042908	1	0.217042908	6.85943699	0.00884326361	-0.0150805907	0.0057580
RM_T205	4.61765432	1	4.61765432	145.935364	3.74930168E-33	-0.0499216691	0.0041324
RI2_T205	1.06106532	1	1.06106532	33.5337029	7.41869366E-09	0.027656056	0.0047758
DP_H201	0.237301871	1	0.237301871	7.49973297	0.0061920248	0.0528855808	0.0193114
DP_H202	4.17472887	1	4.17472887	131.937393	3.58744212E-30	-0.342880577	0.0298510
C*DT/PC_H201	1.78133047	1	1.78133047	56.2977409	7.29518036E-14	-0.315482438	0.0420464
C*DP/PC_H201	0.748442292	1	0.748442292	23.6539898	1.18736921E-06	-0.220963791	0.0454327
C*DT/PC_H202	1.51305342	1	1.51305342	47.818325	5.24499888E-12	-0.429273039	0.0620778
C*DP/PC_H202	4.32404661	1	4.32404661	136.656738	3.53748921E-31	0.739804685	0.0632851
TCh*O2_H201	1.02530789	1	1.02530789	32.403656	1.32176794E-08	-0.753348231	0.1323423
TCh*O2_H202	1.29903865	1	1.29903865	41.0547028	1.61203245E-10	-0.8942765	0.1395694

Figura 21. Regresión RLM para la clase 3. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.



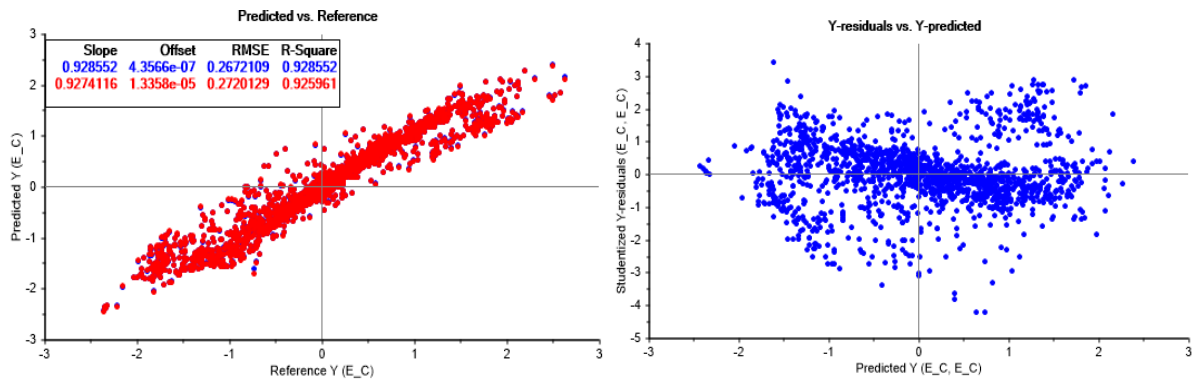
(a)

(b)



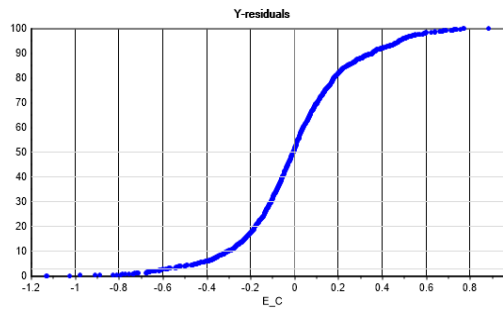
(c)

Figura 22. Regresión RLM para la clase 4. (a) Desempeño de la regresión; (b) residuales versus valor predicho. (c) Distribución de los residuales.



(a)

(b)



(c)

Tabla 5. ANOVA para la RLM de los datos de la clase 3 (Tomada de Unscrambler).

Anova Table							
Multiple Correlation: 0.9899592 (cal) 0.9895433 (val)							
R-Square: 0.9800193 (cal) 0.9791957 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	3592.7561	49	73.3215485	3621.56812	0		
Error	73.2493134	3618	0.0202458017				
Total	3666.00537	3667	0.999728739				
Variables							
Intercept	0	0				0	0
API	0.978893161	1	0.978893161	48.3510666	4.20990066E-12	-0.053210821	0.0076523
BSW	0.337259829	1	0.337259829	16.6582928	4.57207789E-05	0.0210368298	0.0051542
S	4.92836952	1	4.92836952	243.429764	0	0.1289173662	0.0082627
T10	12.6087723	1	12.6087723	622.785034	0	0.262076914	0.0105017
T30	12.6965351	1	12.6965351	627.126526	0	-1.68818033	0.0674126
T50	11.5820217	1	11.5820217	572.079529	0	2.83095479	0.1183599
T70	11.1233616	1	11.1233616	549.425232	0	-1.2930094	0.0551629
T90	4.98337746	1	4.98337746	246.148682	0	-0.12713711	0.0081035
COM	121.974388	1	121.974388	6024.67285	0	0.891337872	0.0114835
G_H201	111.39006	1	111.39006	5501.87451	0	0.830679297	0.0111989
Ts_H201	0.0893515646	1	0.0893515646	4.41332579	0.0357282795	0.00821781252	0.0039117
Pe_H201	0.178411901	1	0.178411901	8.81228638	0.00301171863	-0.171692669	0.0578372
TCh_H201	1.00494003	1	1.00494003	49.6369247	2.2048073E-12	-0.038743306	0.0054991
O2_H201	3.74385691	1	3.74385691	184.919907	4.11701489E-41	0.0592646897	0.0043581
TfI_H201	4.08477974	1	4.08477974	201.759537	1.40129846E-44	0.0835159421	0.0058796
G_T201	2.02889419	1	2.02889419	100.212746	2.74252984E-23	-0.0440187752	0.0043972
N_T201	0.399120957	1	0.399120957	19.7138672	9.26323355E-06	0.0331349261	0.0074627
JET_T201	0.955615699	1	0.955615699	47.2008247	7.51190429E-12	-0.0227908492	0.0033173
ACPM_T201	2.61029553	1	2.61029553	128.930237	2.19527907E-29	-0.0644441769	0.0056755
GASO_T201	0.750977218	1	0.750977218	37.0930023	1.24406363E-09	0.0333663784	0.0054785
FON_T201	1.7509836	1	1.7509836	86.4862823	2.36798787E-20	-0.0472768545	0.0050836
RC_T201	1.36807799	1	1.36807799	67.5733566	2.79778874E-16	-0.0384168029	0.0046734
RM_T201	0.903833091	1	0.903833091	44.6430359	2.72693378E-11	0.0187376663	0.0028043
RI_T201	0.546053112	1	0.546053112	26.9712391	2.17893117E-07	0.0196089447	0.0037757
Te_H202	0.402812511	1	0.402812511	19.8962288	8.42469944E-06	-0.116362497	0.0260872
Ts_H202	0.831087947	1	0.831087947	41.0499496	1.67499292E-10	-0.0400861353	0.0062565
Pe_H202	2.41011429	1	2.41011429	119.042786	2.71559935E-27	0.668792844	0.0612971
TCh_H202	22.9953117	1	22.9953117	1135.80652	0	0.391853839	0.0116271
O2_H202	3.08028436	1	3.08028436	152.144089	2.85115604E-34	0.161040232	0.0130559
TfI_H202	1.18427718	1	1.18427718	58.4951973	2.59551461E-14	-0.0340272561	0.0044490
G_T204	0.861907423	1	0.861907423	42.5720596	7.75872561E-11	-0.0340601318	0.0052201
N_T204	3.41476035	1	3.41476035	168.666	9.96485983E-38	-0.0971236303	0.0074784
JET_T204	0.671760619	1	0.671760619	33.1802406	9.10068643E-09	-0.0260721091	0.0045262
GASO_T204	0.094499588	1	0.094499588	4.66761732	0.0308019929	-0.00855334848	0.0039590
FON_T204	15.0219793	1	15.0219793	741.982239	0	-0.149153456	0.0054756
RC_T204	0.433508098	1	0.433508098	21.4122295	3.83386214E-06	-0.0389190912	0.0084106
RI_T204	1.53432059	1	1.53432059	75.784996	4.72663906E-18	-0.039737545	0.0045646
GASO1_T205	15.9214172	1	15.9214172	786.406372	0	-0.212761447	0.0075869
GASO2_T205	0.0823357403	1	0.0823357403	4.06680965	0.0438075438	0.00736143254	0.0036503
FON_T205	0.502848387	1	0.502848387	24.8371639	6.53066934E-07	-0.0189817362	0.0038087
RC_T205	0.414259255	1	0.414259255	20.4614143	6.27985082E-06	0.0241163597	0.0053314
RI2_T205	6.36838007	1	6.36838007	314.552704	0	-0.0849790722	0.0047914
DP_H201	0.431103259	1	0.431103259	21.2934208	4.07748212E-06	0.26374352	0.0571556
DP_H202	2.28508973	1	2.28508973	112.867462	5.54875085E-26	-0.693039417	0.0652339
C*DT/PC_H201	0.0794064999	1	0.0794064999	3.92211866	0.0477304198	0.00615756586	0.0031092
C*DP/PC_H201	15.2873802	1	15.2873802	755.089783	0	-0.343808949	0.0125117
C*DT/PC_H202	0.273344457	1	0.273344457	13.5014	0.000241839356	0.124027625	0.0337542
C*DP/PC_H202	0.205489457	1	0.205489457	10.1497765	0.00145546102	0.106065191	0.0332923
TCh*O2_H202	7.12325525	1	7.12325525	351.838074	0	-0.255304009	0.0136108

Tabla 6. ANOVA para la RLM de los datos de la clase 4 (Tomada de Unscrambler).

Anova Table							
Multiple Correlation: 0.963614 (ca) 0.9622698 (val)							
R-Square: 0.928552 (ca) 0.925961 (val)							
	SS	df	MS	F ratio	p value	B-coefficients	STDerr
Summary							
Model	1429.97119	28	51.0704002	702.259033	0		
Error	110.02993	1513	0.0727230236				
Total	1540.0011	1541	0.999351799				
Variables							
Intercept	0	0				0	0
API	1.47478688	1	1.47478688	20.2800827	7.20237631E-06	-0.0538895875	0.0119665
T10	0.491844296	1	0.491844296	6.76327562	0.00939594116	0.158620447	0.0609930
T30	0.907447994	1	0.907447994	12.4781694	0.000424103113	-0.29542166	0.0836309
T70	0.781410515	1	0.781410515	10.7450047	0.00106933364	0.126778275	0.0386759
COM	7.2389555	1	7.2389555	99.5414124	9.51964461E-23	0.399213493	0.0400132
G_H201	32.7354126	1	32.7354126	450.13797	0	0.252248198	0.0118892
Te_H201	1.78241754	1	1.78241754	24.5096741	8.22147285E-07	-0.303575099	0.0613193
Ts_H201	1.96915758	1	1.96915758	27.0774918	2.22254414E-07	0.197504178	0.0379552
TCh_H201	0.579310477	1	0.579310477	7.96597338	0.00482883584	0.0466724113	0.0165363
O2_H201	1.67067742	1	1.67067742	22.9731579	1.80424627E-06	0.0522648059	0.0109043
JET_T201	0.767360449	1	0.767360449	10.5518312	0.00118618878	-0.0378666855	0.0116571
FON_T201	9.24546719	1	9.24546719	127.132713	2.27041357E-28	-0.154951617	0.0137425
RC_T201	5.64699793	1	5.64699793	77.6508102	3.30385304E-18	-0.107038885	0.0121469
TfI_T204	0.954972386	1	0.954972386	13.1316395	0.000299956737	-0.0476429164	0.0131473
G_T204	0.348926157	1	0.348926157	4.7980113	0.0286447555	-0.0414751992	0.0189346
JET_T204	2.46211171	1	2.46211171	33.8560257	7.23067961E-09	-0.0857177675	0.0147316
ACPM_T204	0.768665552	1	0.768665552	10.5697975	0.00117479544	0.0616844334	0.0189732
FON_T204	372.479706	1	372.479706	5121.89453	0	-0.575427711	0.0080403
RM_T204	1.93121183	1	1.93121183	26.5557518	2.89770952E-07	0.0873345211	0.0169475
RI_T204	1.03272343	1	1.03272343	14.2007771	0.000170611936	-0.0545431636	0.0144738
GASO2_T205	0.592857003	1	0.592857003	8.1522665	0.00435940782	-0.0359459631	0.0125895
FON_T205	2.48975992	1	2.48975992	34.2362289	5.97294969E-09	-0.0479585156	0.0081963
DP_H201	2.21384001	1	2.21384001	30.4420948	4.04043483E-08	-0.182234868	0.0330289
DP_H202	3.33670545	1	3.33670545	45.8823242	1.79229132E-11	0.126321048	0.0186488
C*DT/PC_H201	2.32623434	1	2.32623434	31.9876347	1.85209981E-08	-0.568990231	0.1006036
C*DP/PC_H201	2.21463418	1	2.21463418	30.4530029	4.01822469E-08	0.72503376	0.1313842
C*DT/PC_H202	4.05102444	1	4.05102444	55.7048225	1.41421515E-13	0.337883085	0.0452709
C*DP/PC_H202	4.51570749	1	4.51570749	62.0945053	6.20932622E-15	-0.450790316	0.0572068

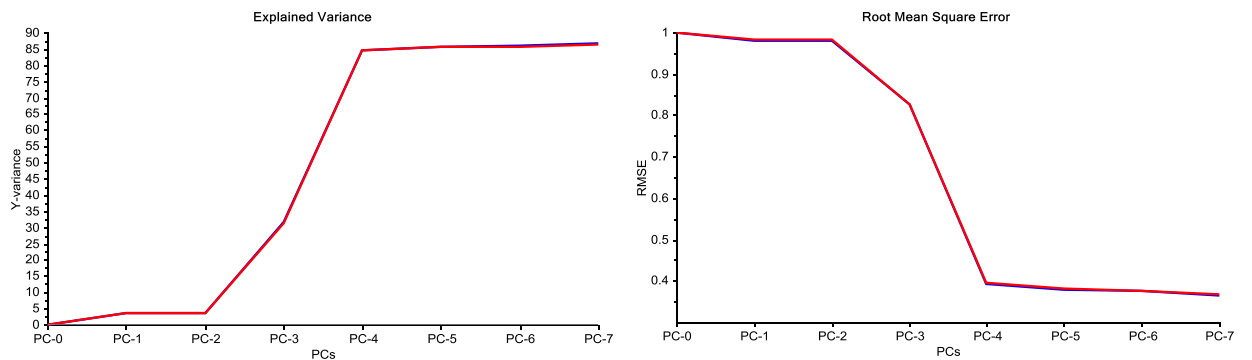
4.5 PCR Y PLS PARA LA PREDICCIÓN DEL CONSUMO DE ENERGÍA EN CADA CLASE

3.5.1 Clase 0. Los resultados de las regresiones PCR y PLS aplicadas a las muestras de la clase 0 son mostrados en las Figuras 23 y 24, respectivamente. Según estas figuras, el PCR reproduce ca. 86% de la varianza del consumo de energía específica (Figura 23a) con un número óptimo de componentes principales de 5, según lo recomendado por la variación del RMSE (Figura 23b). La Figura 23c resume el desempeño del PCR para la predicción del consumo de energía específica en la clase 0; según esta figura, el PCR reporta un desempeño

en la reproducción de la energía específica que puede ser utilizado para aplicaciones cualitativas.

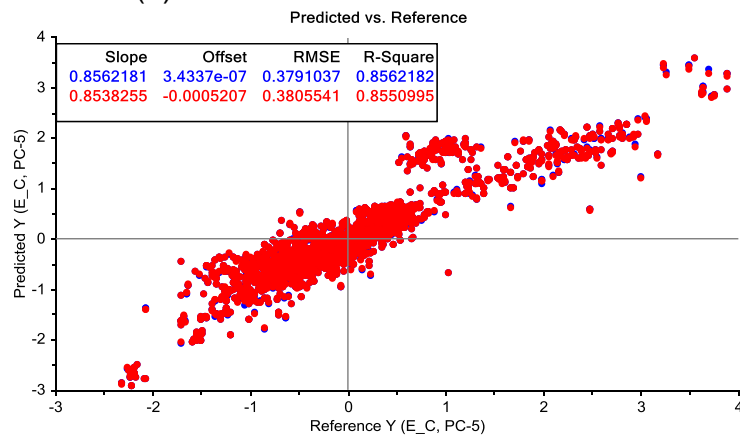
De otro lado, la regresión PLS reproduce *ca.* 96% de la varianza contenida en la energía específica consumida por la U-200 operando según la clase 0 (Figura 24); según los valores de RMSE (Figura 3.23b), el número óptimo de variables latentes corresponde a 5. Los valores de calibración y validación cruzada concuerdan con un error inferior al 0.1%, con lo cual, la regresión PLS puede ser utilizada en aplicaciones cuantitativas con un error de ± 0.4 en la energía específica autoescalada al 95% confianza. Los coeficientes de la regresión PLS se presentan en la Tabla 3.7; según estos coeficientes, el poder calorífico del combustible (COM) afecta *ca.* 18% del valor del consumo de energía específica de la U-200 (Figura 25), coincidiendo con lo obtenido por RLM (sección 3.4.1). A pesar de que el coeficiente R2 en PLS es inferior al reportado por RLM (0.9764), el número de variables requeridas por la regresión PLS es menor que el requerido por RLM (42 variables).

Figura 23. Resultados PCR para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).



(a)

(b)



(c)

Figura 24. Resultados PLS para la clase 0. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).

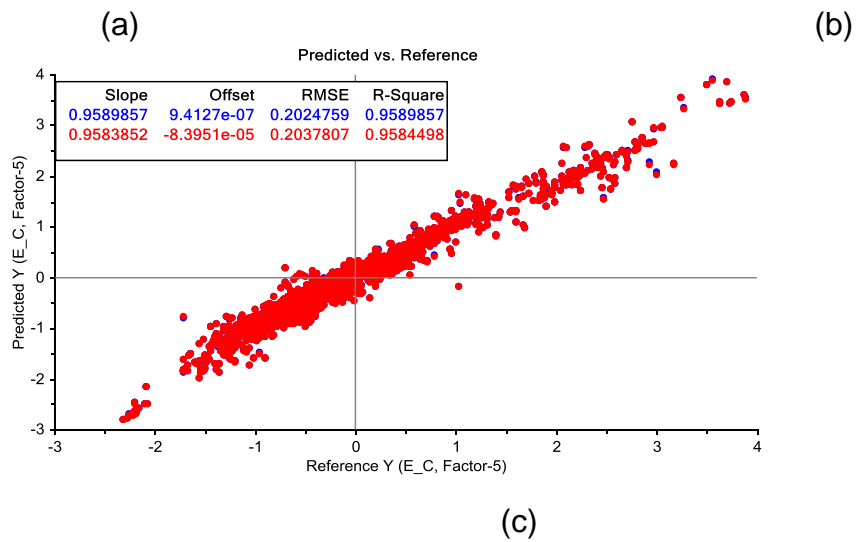
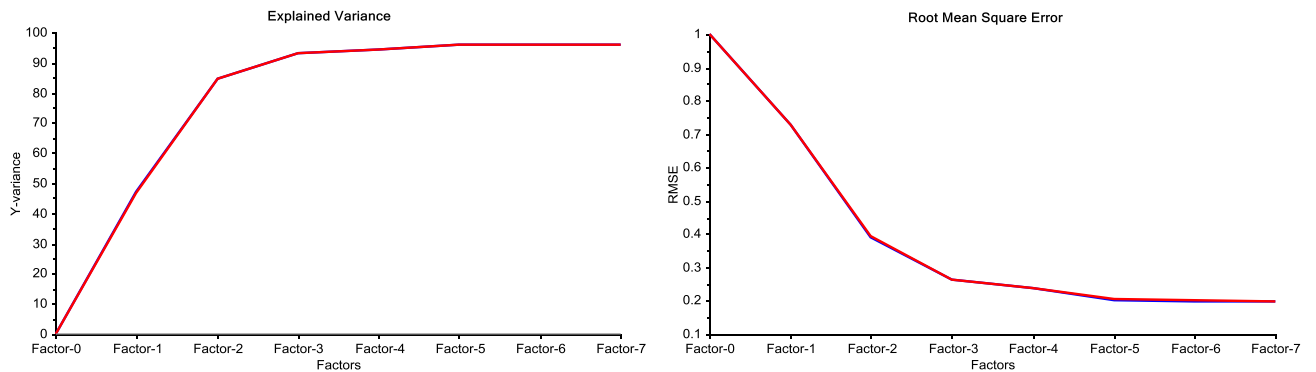
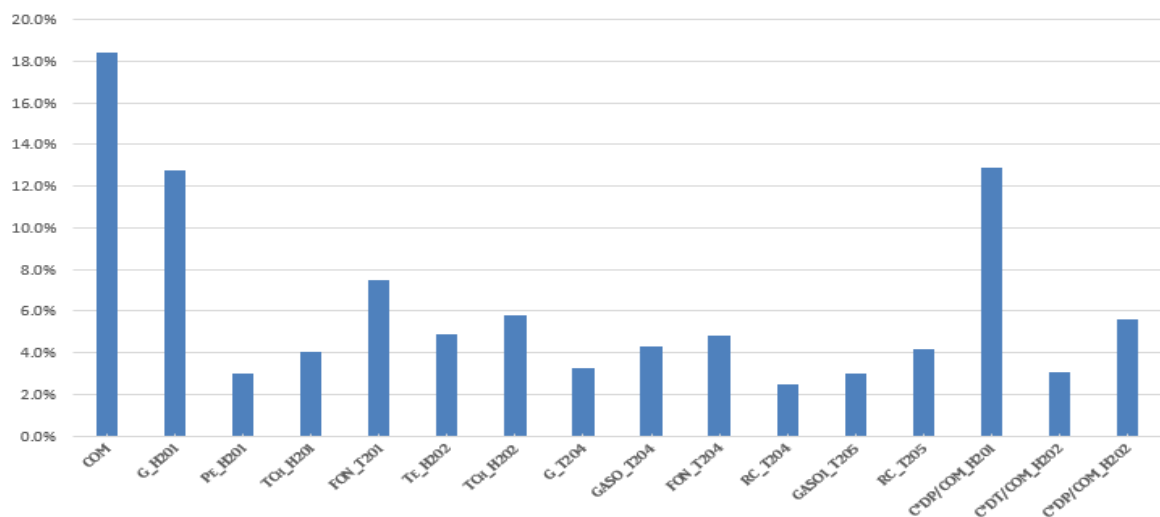


Tabla 7. Coeficientes regresión PLS para las variables autoescaladas de la clase 0 (5 VL).

COM	0.65053	GASO_T204	0.15246
G_H201	0.45182	FON_T204	-
Pe_H201	0.10559	RC_T204	-
TCh_H201	-0.14194	GASO1_T205	-
FON_T201	-0.26366	RC_T205	-
Te_H202	-0.17326	C*DP/COM_H201	-
TCh_H202	0.20525	C*DT/COM_H202	-
G_T204	0.11593	C*DP/COM_H202	-

Figura 25. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 0 (5 VL).



4.5.2 Clase 1. Las Figuras 26 y 27 presentan los resultados de las regresiones PCR y PLS aplicadas a las muestras de la clase 1. Según estas figuras, el PCR reproduce ca. 74% de la varianza del consumo de energía específica (Figura 26a) con un número óptimo de componentes principales de 6 (Figura 23b). Por su parte, PLS reproduce la energía específica con un R2 de 0.7352 y RMSE de 0.3737 (Figura 27). El desempeño de PLS es similar al obtenido por RLM, sin

embargo PLS reporta un menor número de variables, 26, que RLM, 38. Al igual que PCR, PLS requiere de 6 variables latentes (VL) para el óptimo de la regresión. Los coeficientes de la regresión PLS para 6 VL son mostrados en la Tabla 8; con estos coeficientes es posible inferir que el poder calorífico del combustible (COM) aporta ca. 21% de la varianza de la energía específica requerida por la U-200 (Figura 28). El desempeño de la RLM es ligeramente superior al de PLS debido a que incorpora mayor peso a las variables relacionadas con el consumo de oxígeno en los hornos (sección 3.4.2), lo cual impacta en mayor grado el desempeño energético de los mismos.

Figura 26. Resultados PCR para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).

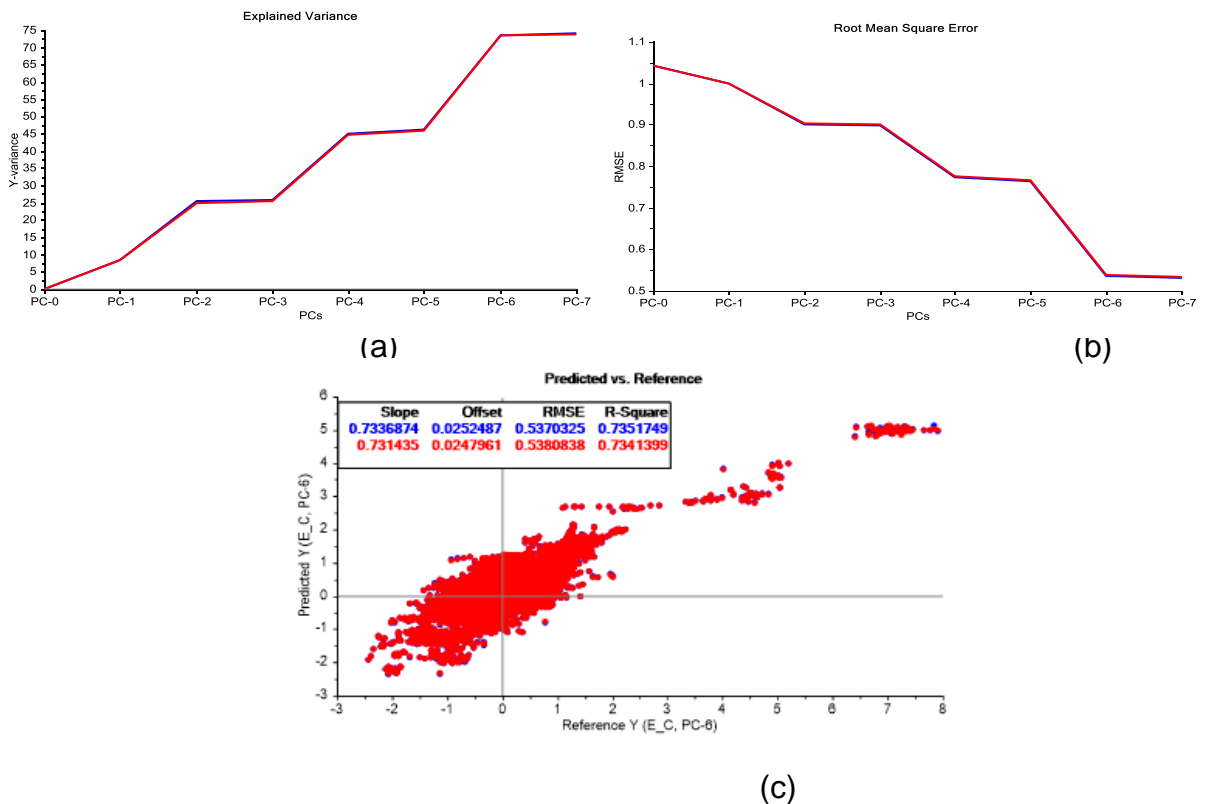


Figura 27. Resultados PLS para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).

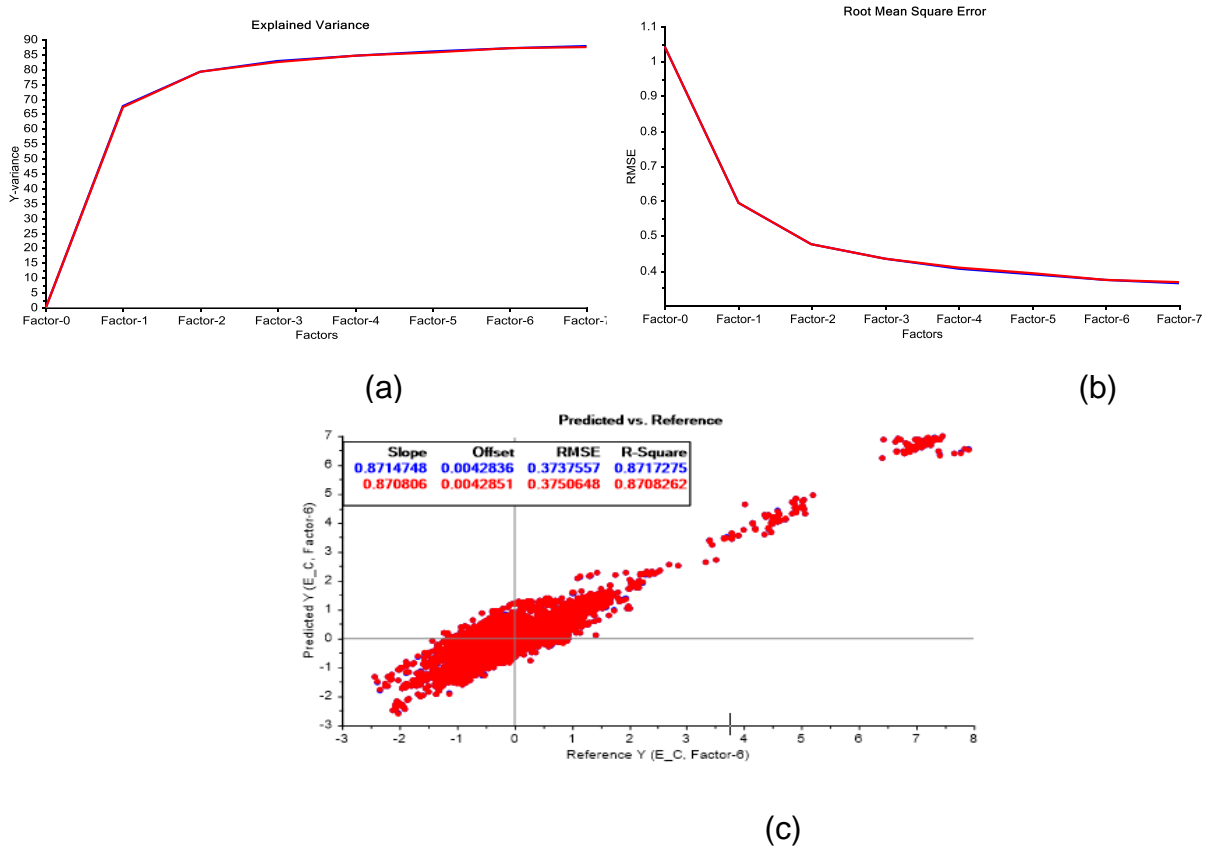
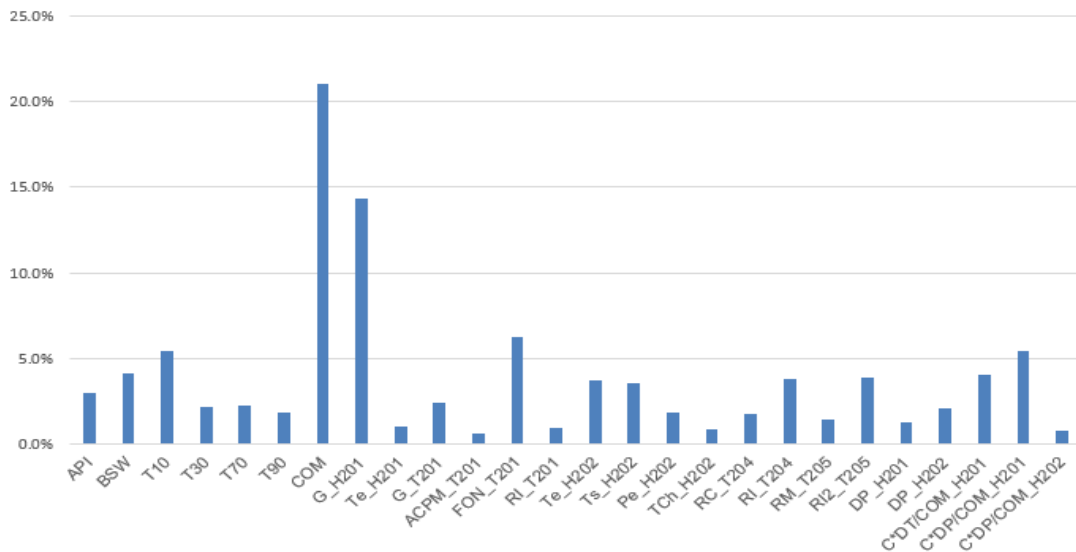


Tabla 8 Coeficientes regresión PLS para las variables autoescaladas de la clase 1 (6 VL).

API	-0.09489	Te_H202	-0.11745
BSW	0.13210	Ts_H202	0.11389
T10	0.17449	Pe_H202	0.05780
T30	-0.07024	TCh_H202	0.02667
T70	0.07076	RC_T204	-0.05518
T90	0.05933	RI_T204	-0.12107
COM	0.67293	RM_T205	0.04666
G_H201	0.45823	RI2_T205	-0.12481

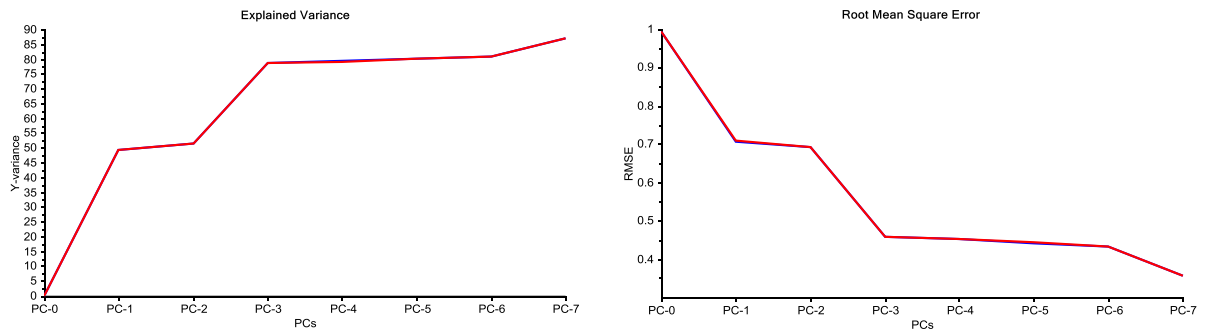
Te_H201	-0.03353	DP_H201	0.04164
G_T201	-0.07674	DP_H202	0.06695
ACPM_T201	0.01921	C*DT/PC_H201	-0.12937
FON_T201	-0.19945	C*DP/PC_H201	-0.17454
RI_T201	0.03065	C*DP/PC_H202	-0.02448

Figura 28. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 1 (6 VL).



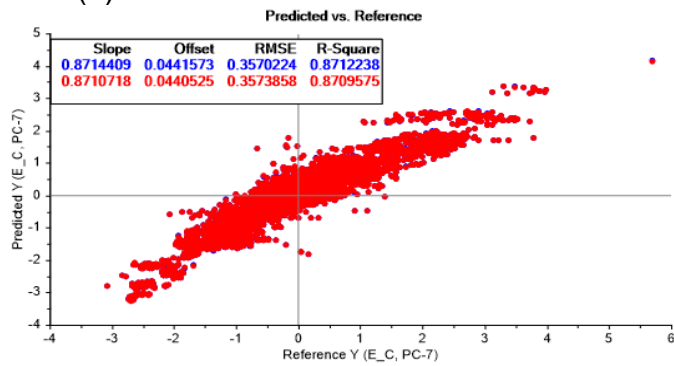
4.5.3 Clase 2. Los resultados de las regresiones PCR y PLS aplicadas a las muestras de la clase 2 muestran que la regresión PCR reproduce ca. 87% de la varianza del consumo de energía específica con 7 PC (Figura 3.28), mientras que la regresión PLS reporta un R2 de 0.932 y RMSE de 0.2598 utilizando sólo 4 VL (Figura 3.29). A pesar de que el desempeño de PLS es inferior al obtenido con RLM (R2=0.9685), la regresión PLS requiere 20 variables menos que RLM. Al igual que para la clase 1, los coeficientes de PLS (Tabla 3.9) dan mayor peso al COM (Figura 3.30), mientras que los coeficientes de RLM concentran ca. 40% de la varianza en las variables relacionadas con el consumo de oxígeno en los hornos (sección 3.4.3).

Figura 29 Resultados PCR para la clase 1. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).



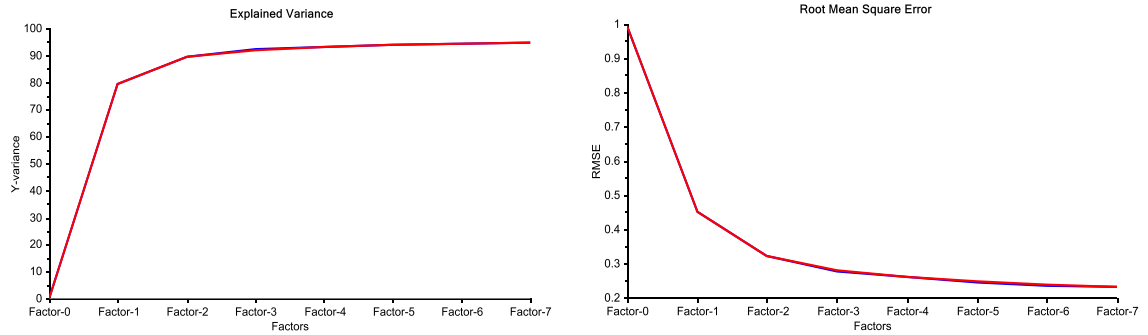
(a)

(b)



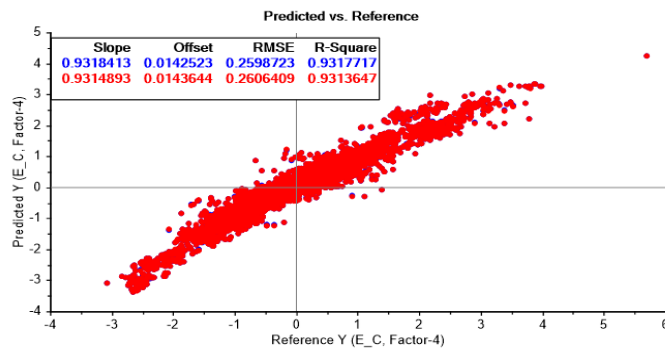
(c)

Figura 30. Resultados PLS para la clase 2. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).



(a)

(b)



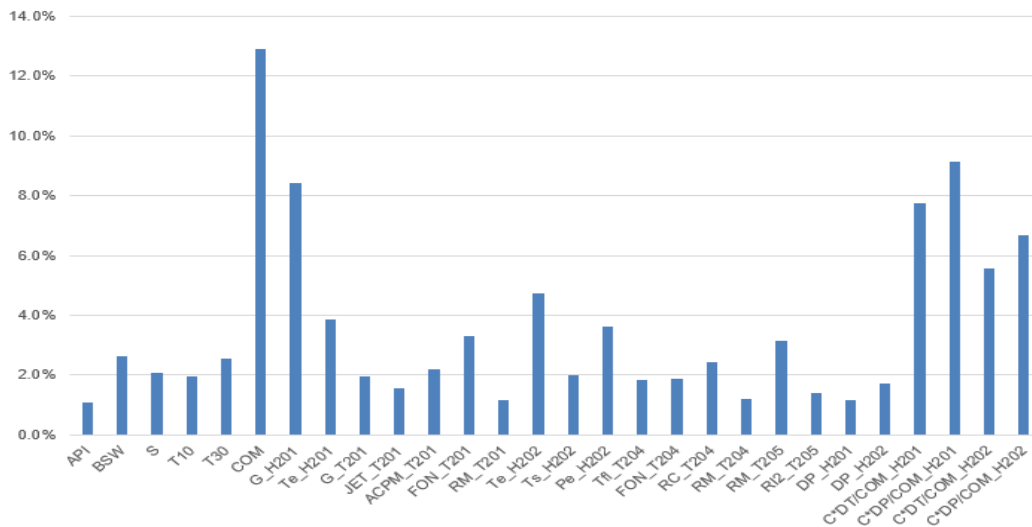
(c)

Tabla 9. Coeficientes regresión PLS para las variables autoescaladas de la clase 2 (4 VL).

API	0.02544754	Ts_H202	0.04710171
BSW	-0.06164817	Pe_H202	0.08519831
S	-0.04843931	Tfl_T204	0.04311051
T10	0.04556998	FON_T204	-
T30	0.05952024	RC_T204	0.05675523
COM	0.3032535	RM_T204	-
G_H201	0.1974708	RM_T205	-
Te_H201	-0.09109749	RI2_T205	0.03338612

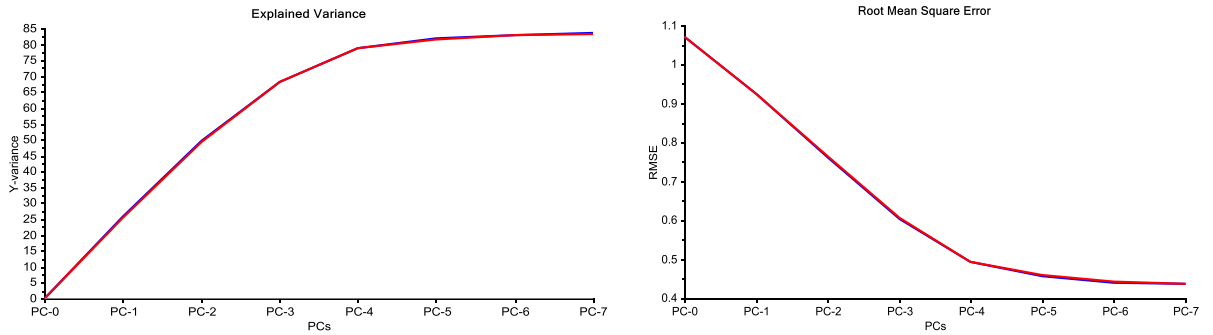
G_T201	-0.04601879	DP_H201	-
JET_T201	-0.03691234	DP_H202	0.04071357
ACPM_T201	0.05190605	C*DT/COM_H201	-0.1821778
FON_T201	-0.07799181	C*DP/COM_H201	-0.2148114
RM_T201	0.02702068	C*DT/COM	-0.1307534
Te_H202	-0.1109204	C*DP/COM	-0.1568936

Figura 31. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 2 (4 VL).



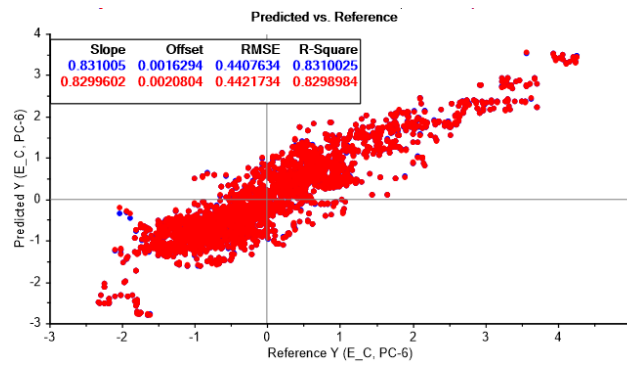
4.5.4 Clase 3. Los resultados de las regresiones PCR y PLS en la clase 3 reproducen los datos con R2 de 0.8310 y 0.9263, respectivamente (Figuras 3.31 y 3.32). Estos desempeños son bajos en comparación con RLM aplicado a la misma clase (sección 3.4.4); RLM reporta R2 de 0.98 y RMSE de 0.141. Asimismo, según la Tabla 3.10 y la Figura 3.33, la variable que tiene mayor peso en la regresión PLS corresponde al COM. RLM reporta un mayor desempeño que PLS debido a que considera mayor peso en las variables que identifican la clase 3 –i.e. temperaturas de evaporado del crudo tratado– (sección 3.2); sin embargo, PLS reporta una regresión que requiere 30 variables menos que RLM.

Figura 32. Resultados PCR para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).



(a)

(b)



(c)

Figura 33. Resultados PLS para la clase 3. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).

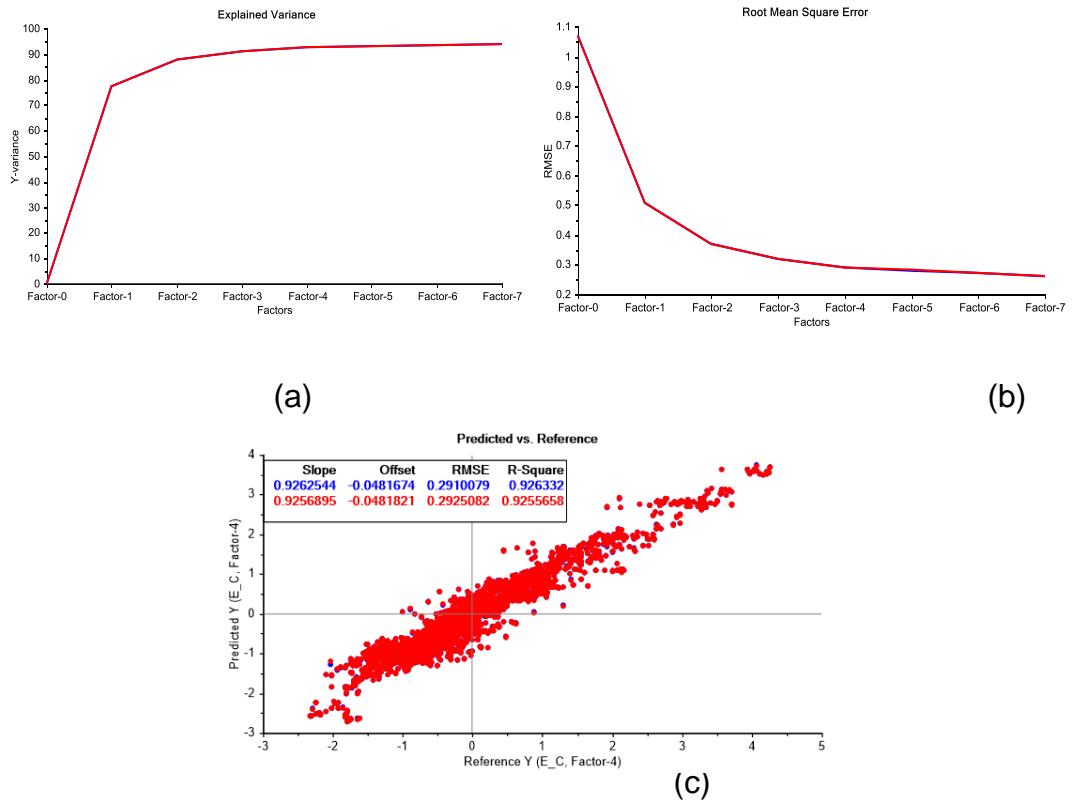
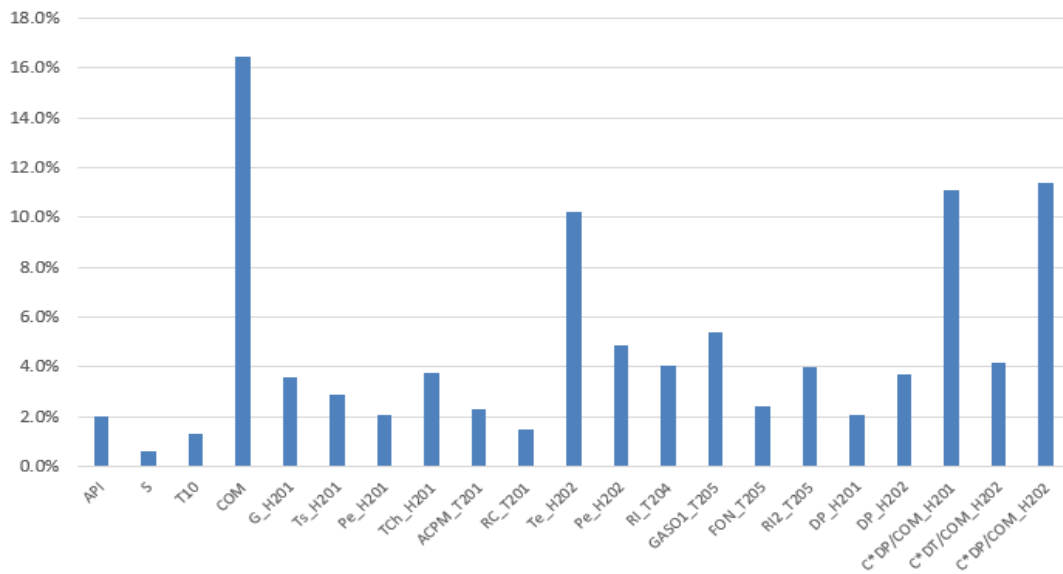


Tabla 10. Coeficientes regresión PLS para las variables autoescaladas de la clase 3 (4 VL).

API	-0.04818198	Pe_H202	0.1161894
S	0.01441705	RI_T204	-0.09715626
T10	0.03073446	GASO1_T205	-0.1281131
COM	0.3918348	FON_T205	-0.05767988
G_H201	0.08599742	RI2_T205	-0.09511658
Ts_H201	0.06826679	DP_H201	0.04911313
Pe_H201	0.04993093	DP_H202	0.08767294

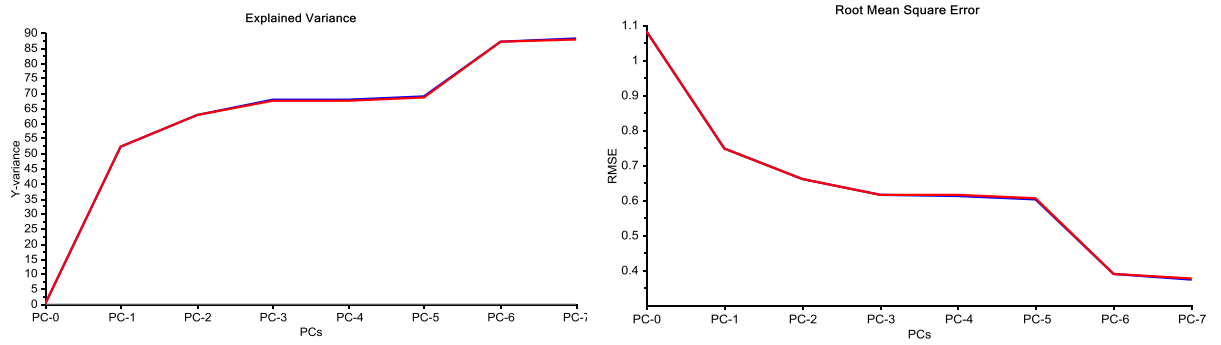
TCh_H201	0.09015111	C*DP/COM_H201	-0.2650075
ACPM_T201	-0.0554354	C*DT/COM_H202	-0.09919567
RC_T201	-0.03610331	C*DP/COM_H202	-0.2716183
Te_H202	-0.2432301		

Figura 34. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 3 (4 VL).



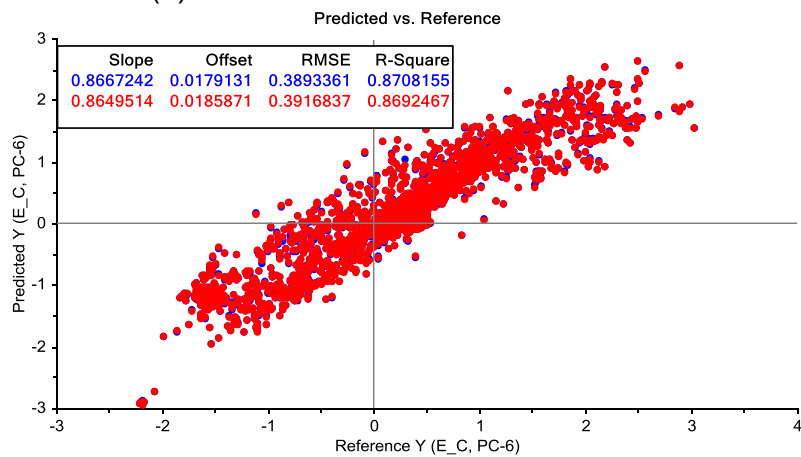
4.5.5 Clase 4. Las Figuras 35 y 36, presentan los resultados de las regresiones PCR y PLS, respectivamente, aplicadas a las muestras de la clase 4 son mostrados en. Según estas figuras, el PCR reproduce ca. 87% de la varianza del consumo de energía específica (Figura 3.34a) con un número óptimo de PC de 6 (Figura 3.34b), mientras que la regresión PLS requiere 4 VL para reproducir ca. 90% de la varianza. Estos desempeños son inferiores al obtenido por RLM que reproduce ca. 93% de la varianza de la energía específica de la clase 4. A pesar de esto, PLS reporta una regresión que requiere la mitad de variables que la regresión obtenida por RLM (sección 3.4.5). El desempeño de PLS está cercano a RLM en la clase 4 ya que las dos confieren un peso importante a la variable de flujo de fondos de la T204 (Fon_T204).

Figura 35. Resultados PCR para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por los PC (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PCR (calibración en azul y validación en rojo).



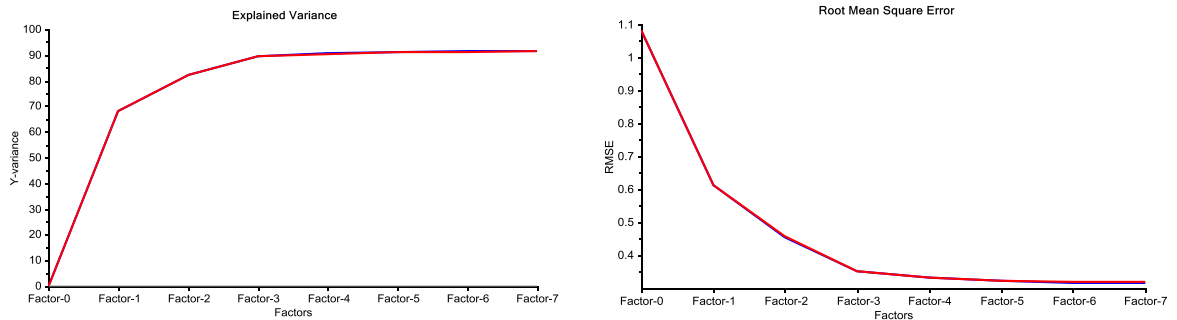
(a)

(b)



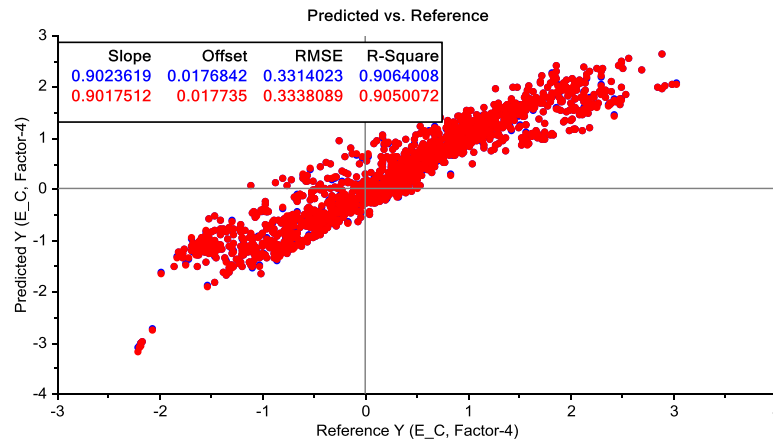
(c)

Figura 36. Resultados PLS para la clase 4. (a) Varianza explicada (calibración en azul y validación en rojo); (b) Variación RMSE por las VL (calibración en azul y validación en rojo). (c) Gráfico de desempeño del PLS (calibración en azul y validación en rojo).



(a)

(b)



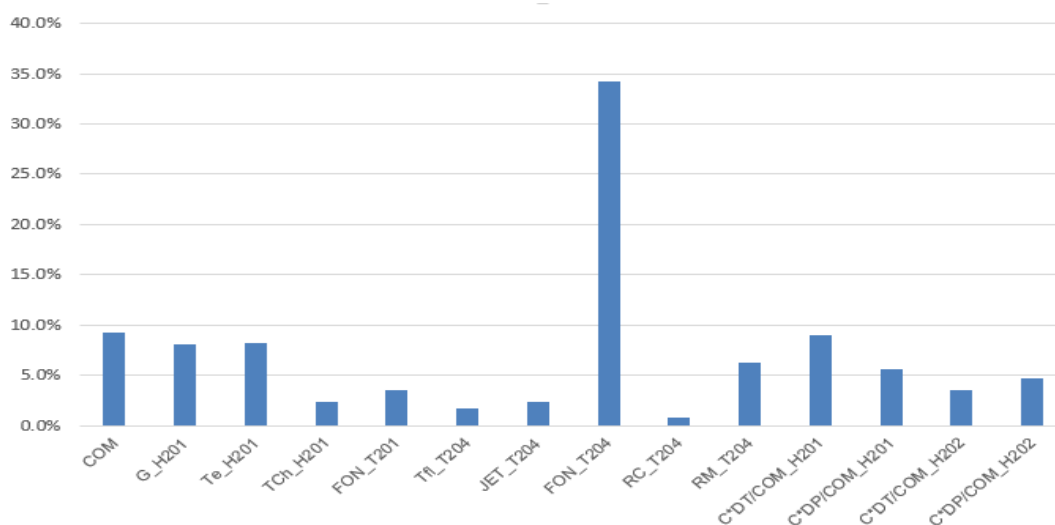
(c)

Tabla 11. Coeficientes regresión PLS para las variables autoescaladas de la clase 4 (4 VL).

COM	0.1758552	FON_T204	-0.6541597
G_H201	0.1547836	RC_T204	0.01690111
Te_H201	-0.157007	RM_T204	0.1210062
TCh_H201	0.04477188	C*DT/COM_H201	-0.1721013
FON_T201	-0.0686041	C*DP/COM_H201	-0.1068911

Tfl_T204	-0.03329964	C*DT/COM_H202	-0.06764134
JET_T204	-0.04644935	C*DP/COM_H202	-0.08938543

Figura 37. Porcentaje de contribución de cada variable en la regresión PLS en la predicción de la energía específica autoescalada de la clase 4 (4 VL).



Según los resultados de PCR y PLS, esta última regresión presenta mejores desempeños para la predicción del consumo energético específico de la U-200 para las clases definidas por k-mean, excepto para las clases 1 y 4 que contienen datos importantes de operación de la unidad en alternativa viscorreductora. RLM reporta mayores desempeños que PLS, pero con un mayor número de variables. El análisis de la reducción de variables en las regresiones RLM para las 5 clases es un tópico recomendado para la continuación del presente trabajo de aplicación. En caso de que no exista ninguna restricción en la disponibilidad de la información de las variables operacionales necesarias para cada modelo y teniendo en cuenta los resultados obtenidos para el cálculo de consumo de energía, se seleccionan los modelos obtenidos con la regresión lineal múltiple (RLM) para las 5 clases. De otro lado, en caso de falla de algún instrumento o no disponibilidad de algún análisis de laboratorio se seleccionan como segunda opción los modelos

obtenidos con la regresión por mínimos cuadrados parciales (PLS) para las 5 clases debido a la menor cantidad de variables necesarias para la predicción.

5. MEJORAS EN EL PROCESO BASADAS EN LAS LÍNEAS BASE OBTENIDAS CON LAS REGRESIONES MULTIVARIADAS

Teniendo en cuenta las variables de mayor contribución de cada uno de los modelos (RLM) de las 5 clases analizadas, las siguientes mejoras operativas fueron identificadas:

- Oportunidad con la instalación de un medidor de poder calorífico. Debido a que la variable poder calorífico del combustible (COM) se presenta en todas las líneas base, se encuentra en evaluación la consecución e instalación de un medidor de poder calorífico para la actualización en línea; actualmente, el valor del COM es actualizado cada 2 a 3 días.
- Seguimiento operacional a variables clave. Actualmente, un control operacional de las variables de mayor impacto en el consumo de energía es efectuado en planta con base en los valores de operación deseados; con esto, se espera mejorar las condiciones de calidad y producción requeridas con un bajo costo energético asociado. Entre las variables clave se encuentran:
 - Clase 0 a la 5: COM, G_H-201
 - Clase 0: PE_H-201, FON_T-201, TE_H-202
 - Clase 0 y 3: PE_H-202, DP_H-202
 - Clase 1 y 2: O2_H-201, TCh*O2_H201
 - Clase 2: O2_H-202
 - Clase 2 y 3: TCh*O2_H202
 - Clase 3: T30, T50, T90
 - Clase 4: FON_T-204, TE_H-201
- Análisis del assay del crudo. La frecuencia muestreo para los análisis de laboratorio para destilación y capacidad calorífica del crudo procesado en la U-

200 es recomendada para la estimación adecuada de los consumos energéticos. Estas mediciones pueden ser obtenidas en línea por medio de un equipo de espectroscopia NIR.

- Seguimiento del desempeño de los hornos. Para cada horno se ha implementado un análisis operacional respectivo para posibles ajustes; por ejemplo, para modificar la temperatura de chimenea del H-202 este cuenta con un venteo que se puede abrir y cerrar y que modifica de forma importante esta variable y por lo tanto la operación del horno.

6. IMPACTOS DEL TRABAJO DE APLICACIÓN

La implementación de las mejoras obtenidas a través de la línea base energética puede llegar a tener un impacto económico por reducciones en consumos de energía hasta del 7% del total de una unidad de proceso [5]; para el caso de la U-200 se estima un ahorro de consumo del 3% el cual corresponde a \$283 mil dólares anuales [2]. También se tendría un impacto ambiental por reducción de gases de efecto invernadero de 1,43 ton/año de CO₂ [3]. De igual forma se aplicaría esta misma metodología de construcción de línea base para otras plantas de proceso e inclusive se podría integrar dentro del control avanzado de la unidad U-200 para una optimización energética en línea.

Asimismo, con los resultados de este trabajo se confirmó los supuestos de mayor consumo energético de la unidad U-200, cuando esta se encuentra operando en modo viscorreductora; *i.e.* como el observado en el año 2013 cuando la unidad operó durante varios meses en este modo (Figura 3.1). Este mayor consumo era presumido por los ingenieros de proceso porque la carga de la T-204 es un producto mucho más pesado (crudo reducido y fondos de vacío), pero no se había evaluado cuantitativamente como se en el presente trabajo.

También, los resultados del presente trabajo muestran el impacto de la salida de servicio de uno o más intercambiadores en el sistema de precaliente, disminuyendo la temperatura del crudo y requiriendo mayores consumos de gas y de energía como se observa en los subgrupos de consumos del 2014 y 2015 que se encuentran en la región de altas cargas de la U-200 de la Figura 3.4. Este aumento en consumo energético es coherente con el aumento de flujo que se observa cuando se realiza el procedimiento operativo de desvío (*By-pass*) de un intercambiador.

7. CONCLUSIONES

- El protocolo seguido en el presente trabajo utilizando análisis por componentes principales y el método *k-mean* para identificar diferentes escenarios operativos concuerda con los reportes de planta y permite dilucidar 5 clases o esquemas operativos diferentes de la unidad de destilación de crudos U-200.
- La regresión RLM aplicada sobre las 5 clases obtenidas por el análisis PCA y el método *k-mean* genera las correspondientes líneas base. Los desempeños reportados por las diferentes regresiones RLM son suficientes para la aplicación cuantitativa de las líneas base en los estudios energéticos de la U-200 en el respectivo esquema operativo.
- La regresión RLM aplicada sobre las 5 clases reproduce los datos de consumo energético específico de la unidad con mayor desempeño que el mostrado por las regresiones PCR y PLS; PLS reporta mejor desempeño que PCR. De otro lado, PLS reporta un menor número de variables para las respectivas líneas base y un desempeño suficiente para aplicaciones cuantitativas para las clases 0, 2 y 3.
- En caso de que no exista ninguna restricción en la disponibilidad de la información de las variables operacionales, las regresiones RLM deben ser implementadas –según la clase o esquema operativo– para el cálculo del consumo de energía específica. De otro lado, en caso de falla de alguna medida o no disponibilidad de algún análisis de laboratorio, las regresiones obtenidas por PLS –según la clase o esquema operativo– corresponden a una opción adecuada para la aproximación de los consumos energéticos específicos.

- Según la concordancia entre las tendencias de las variables y las clases obtenidas utilizando PCA y el método *k-mean* se recomienda extender el protocolo a otras unidades de la refinería de Barrancabermeja.
- El análisis de la reducción de variables en las regresiones RLM para las 5 clases dilucidadas en la operación de la U-200 es un tópico recomendado para la continuación del presente trabajo de aplicación.

REFERENCIAS BIBLIOGRAFICAS

[1] ISO 50001:2011. Energy management systems-requirements with guidance for use; 2011.

[2] ECOPETROL S.A. Informe de estudio de benchmarking de SOLOMON para la GRB. Coordinación de Economía y Gestión, GRB, 2015.

[3] ECOPETROL S.A. Informe de GEI (Gases de efecto invernadero) de la GRB. Coordinación de Energía y Pérdidas, GRB, 2015.

[4] Velázquez, D., González-Falcón, R., Pérez-Lombard, L., Gallego, L. M., Monedero, I., & Biscarri, F. (2013). Development of an energy management system for a naphta reforming plant: a data mining approach. *Energy Conversion and Management*, 67, 217–225.

[5] Monedero, I., Biscarri, F., León, C., Guerrero, J. I., González, R., & Pérez-Lombard, L. (2012). Decision system based on neural networks to optimize the energy efficiency of a petrochemical plant. *Expert Systems with Applications*, 39(10), 9860-9867.

[6] Biscarri Félix et al. (2012). A decision support system for consumption optimization in a naphtha reforming plant.

[7] García, Carlos. Colombia Tendencias a largo plazo en el sector de energía. Octubre 2014. UPME. http://www.upme.gov.co/SeccionDemanda/UPME-WEC_CARTAGENA-OCT%202104_ENERGY%20TR3LEMMA.pdf

- [8] ECOPETROL S.A. Manual de operación U-200. Barrancabermeja: ECOPETROL S.A. 2010
- [9] Terrádez Gurrea, M. Análisis de componentes principales. Universidad abierta de Cataluña. http://www.uoc.edu/in3/emath/docs/Componentes_principales.pdf
- [10] Brereton, R.G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*. 1st edition. USA: Wiley, 2002.
- [11] González Carrasco C. Control Estadístico Multivariable mediante análisis de componentes principales (PCA) aplicado al área de destilación Planta I Methanex Chile. Limited. Trabajo de grado Ingeniero Civil Químico. Universidad de Magallanes.
- [12] Kannan, K.S. and Manoj, K. 2015. Outlier Detection in Multivariate Data. *Applied Mathematical Sciences*, Vol. 9, No 47, 2317 – 2324.
- [13] Jobson, J.D. 1991. *Applied Multivariate Data Analysis. Volume 1: Regression and Experimental Design*. Springer Science. 1st edition. USA.
- [14] Chatterjee, S. & Hadi, A.S. 1988. *Sensitivity Analysis in Linear Regression*. John Wiley & Sons, USA.
- [15] Hair, J.f.; Black, W.C.; Babin, B.J.; Anderson, R.E. 2010. *Multivariate Data Analysis*. Prentice Hall, 7th edition, USA.
- [16] Maesschalck, R.D.; Jouan-Rimbaud, D.; Massart, D.L. 2000. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems*, 50, 1–18.

- [17] Worley, B., Halouska, S., Powers, R. 2013. Utilities for Quantifying Separation in PCA/PLS-DA Scores Plots. *Anal Biochem.*, 433(2): 102–104.
- [18] Kourti, M. and MacGregor, J.F. 1995. Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and Intelligent Laboratory Systems*, 28, 3–21.
- [19] Puerto Vargas, S.D., Rodríguez Duran, E.V., Maradei García, M.P., Escalante Hernández, H. 2015. Línea base energética de la sección atmosférica y la sección de vacío de la unidad U-250 de refinación de crudo en la refinería de ECOPETROL S.A. Proyecto de grado, escuela de Ingeniería Química, UIS.
- [20] Salazar Cachopo, M.S., Morales Medina, G. 2015. Análisis de datos de consumo energético de la unidad de refinación de crudos U-200 del complejo industrial Barrancabermeja: regresión lineal múltiple. Proyecto de grado, escuela de Ingeniería Química, UIS.
- [21] Hileman, M.J. 2001. Best practices determine refinery pacesetter performance. *Oil & Gas Journal*, 99 (41), 44-47.
- [22] Borroto, A.E., 2006. Gestión y economía energética. Cienfuegos, Cuba: Editorial Universo Sur, ISBN 959-257-114-7.
- [23] Correa, J., Borroto, A.E., Alpha, M., González, R., Curbelo, M., Díaz, A.M. 2014. Diseño y aplicación de un procedimiento para la planificación energética según la NC-ISO 50001:2011. *Ingeniería Energética* Vol. XXXV, 38-47.
- [24] Borroto, A.E., 2006. Gestión y economía energética. Cienfuegos, Cuba: Editorial Universo Sur, ISBN 959-257-114-7.

[25] Javied, T., Rackowa, T., Franke, J. 2015. Implementing energy management system to increase energy efficiency in manufacturing companies. *Procedia CIRP* 26, 156 – 161.

[26] May, G., Barletta, I., Stahl, B., Taisch, M. 2015. Energy management in production: A novel method to develop key performance indicators for improving energy efficiency. *Applied Energy* 149, 46–61.

[27] García, C. 2014. Colombia Tendencias a largo plazo en el sector de energía. UPME. http://www.upme.gov.co/SeccionDemanda/UPME-WEC_CARTAGENA_OCT%202104_ENERGY%20TR3LEMMMA.pdf.

[28] Castrillón, R. del P., González, A.J., Quispe, E.C. 2013. Mejoramiento de la eficiencia energética en la industria del cemento por proceso húmedo a través de la implementación del sistema de gestión integral de la energía. *Dyna*, 177, 115-123.

[29] Campos, J.C., Quispe, E.C., Vidal, J.R., Lora, E. 2008. EL MGIE, un modelo de gestión energética para el sector productivo nacional. *El hombre y la máquina*, pp. No 30,18-31.

BIBLIOGRAFÍA

BISCARRI Félix et al. A decision support system for consumption optimization in a naphtha reforming plant. 2012

BORROTO, A.E., Gestión y economía energética. Cienfuegos, Cuba: Editorial Universo Sur, 2006. ISBN 959-257-114-7.

BORROTO, A.E., Gestión y economía energética. Cienfuegos, Cuba: Editorial Universo Sur, ISBN 959-257-114-7. 2006.

BRRERETON, R.G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*. 1st edition. USA: Wiley, 2002.

CAMPOS, J.C., QUISPE, E.C., VIDAL, J.R., LORA, E. EL MGIE, un modelo de gestión energética para el sector productivo nacional. *El hombre y la máquina*, pp. No 30,18-31. 2008.

CASTRILLÓN, R. DEL P., GONZÁLEZ, A.J., QUISPE, E.C. Mejoramiento de la eficiencia energética en la industria del cemento por proceso húmedo a través de la implementación del sistema de gestión integral de la energía. *Dyna*, 177, 115-123. 2013.

CHATTERJEE, S. & HADI, A.S. *Sensitivity Analysis in Linear Regression*. John Wiley & Sons, USA. 1988.

CORREA, J., BORROTO, A.E., ALPHA, M., GONZÁLEZ, R., CURBELO, M., DÍAZ, A.M. Diseño y aplicación de un procedimiento para la planificación

energética según la NC-ISO 50001:2011. *Ingeniería Energética* Vol. XXXV, 38-47. 2014.

ECOPETROL S.A. Informe de estudio de benchmarking de SOLOMON para la GRB. Coordinación de Economía y Gestión, GRB, 2015.

ECOPETROL S.A. Informe de GEI (Gases de efecto invernadero) de la GRB. Coordinación de Energía y Pérdidas, GRB, 2015.

ECOPETROL S.A. Manual de operación U-200. Barrancabermeja: ECOPETROL S.A. 2010

GARCÍA, C. Colombia Tendencias a largo plazo en el sector de energía. UPME 2014.. [en línea] disponible en: http://www.upme.gov.co/SeccionDemanda/UPME-WEC_CARTAGENA-OCT%202104_ENERGY%20TR3LEMMMA.pdf.

GARCÍA, Carlos. Colombia Tendencias a largo plazo en el sector de energía. Octubre 2014. UPME. [en línea] disponible en http://www.upme.gov.co/SeccionDemanda/UPME-WEC_CARTAGENA-OCT%202104_ENERGY%20TR3LEMMMA.pdf

GONZÁLEZ CARRASCO C. Control Estadístico Multivariable mediante análisis de componentes principales (PCA) aplicado al área de destilación Planta I Methanex Chile. Limited. Trabajo de grado Ingeniero Civil Químico. Universidad de Magallanes.

HAIR, J.F.; BLACK, W.C.; BABIN, B.J.; ANDERSON, R.E *Multivariate Data Analysis*. Prentice Hall, 7th edition, USA. . 2010.

HILEMAN, M.J. Best practices determine refinery pacesetter performance. *Oil & Gas Journal*, 99 (41), 44-47. 2001.

ISO 50001:2011. Energy management systems-requirements with guidance for use; 2011.

JAVIED, T., RACKOWA, T., FRANKE, J. Implementing energy management system to increase energy efficiency in manufacturing companies. *Procedia CIRP* 26, 156 – 161. 2015.

JOBSON, J.D. *Applied Multivariate Data Analysis. Volume 1: Regression and Experimental Design*. Springer Science. 1st edition. USA. 1991.

KANNAN, K.S. and MANOJ, K. Outlier Detection in Multivariate Data. *Applied Mathematical Sciences*, Vol. 9, No 47, 2317 – 2324. 2015.

KOURTI, M. AND MACGREGOR, J.F. Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and Intelligent Laboratory Systems*, 28, 3–21. 1995.

MAESSCHALCK, R.D.; JOUAN-RIMBAUD, D.; MASSART, D.L. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems*, 50, 1–18. 2000.

MAY, G., BARLETTA, I., STAHL, B., TAISCH, M. Energy management in production: A novel method to develop key performance indicators for improving energy efficiency. *Applied Energy* 149, 46–61. 2015.

MONEDERO, I., BISCARRI, F., LEÓN, C., GUERRERO, J. I., GONZÁLEZ, R., & PÉREZ-LOMBARD, L. Decision system based on neural networks to optimize the

energy efficiency of a petrochemical plant. *Expert Systems with Applications*, 39(10), 9860-9867. 2012

PUERTO VARGAS, S.D., RODRÍGUEZ DURAN, E.V., MARADEI GARCÍA, M.P., ESCALANTE HERNÁNDEZ, H. Línea base energética de la sección atmosférica y la sección de vacío de la unidad U-250 de refinación de crudo en la refinería de ECOPETROL S.A. Proyecto de grado, escuela de Ingeniería Química, UIS. 2015.

SALAZAR CACHOPO, M.S., MORALES MEDINA, G. Análisis de datos de consumo energético de la unidad de refinación de crudos U-200 del complejo industrial Barrancabermeja: regresión lineal múltiple. Proyecto de grado, escuela de Ingeniería Química, UIS. 2015.

TERRÁDEZ GURREA, M. Análisis de componentes principales. Universidad abierta de Cataluña. [en línea] disponible en http://www.uoc.edu/in3/emath/docs/Componentes_principales.pdf

VELÁZQUEZ, D., GONZÁLEZ-FALCÓN, R., PÉREZ-LOMBARD, L., GALLEGO, L. M., MONEDERO, I., & BISCARRI, F. Development of an energy management system for a naphta reforming plant: a data mining approach. *Energy Conversion and Management*, 67, 217–225. 2013

WORLEY, B., HALOUSKA, S., POWERS, R. Utilities for Quantifying Separation in PCA/PLS-DA Scores Plots. *Anal Biochem.*, 433(2): 102–104. 2013.

ANEXOS

Anexo A. Variables

- API: Gravedad API.
- BSW: Propiedad BSW.
- S: Azufre, %.
- T10: Temperatura al 10% evaporado, °F.
- T30: Temperatura al 30% evaporado, °F.
- T50: Temperatura al 50% evaporado, °F.
- T70: Temperatura al 70% evaporado, °F.
- T90: Temperatura al 90% evaporado, °F.
- COM: Poder calorífico gas combustible, BTU/ft³.
- G_H201: Flujo de gas combustible al H201.
- Te_H201: Temperatura entrada crudo al H201, °F .
- Ts_H201: Temperatura salida crudo del H201, °F.
- Pe_H201: Presión de entrada del crudo al H201, psi.
- TCh_H201: Temperatura de chimenea gases H201, °F.
- O2_H201: Oxígeno en el flujo de aire al H201, %.
- Tfl_T201: Temperatura zona flash T201, °F.
- G_T201: Flujo gas de cima T201, CFSD.
- N_T201: Flujo de nafta T201, bpd.
- JET_T201: Flujo de jet T201, bpd.
- ACPM_T201: Flujo de acpm T201, bpd.
- GASO_T201: Flujo de gasóleo T201, bpd.
- FON_T201: Flujo de fondo T201, bpd.
- RC_T201: Flujo de reciclo de cima T201, bpd.
- RM_T201: Flujo de reciclo medio T201, bpd.
- RI_T201: Flujo de reciclo inferior T201, bpd.

Te_H202: Temperatura entrada crudo al H202, °F.
 Ts_H202: Temperatura salida crudo del H202, °F.
 Pe_H202: Presión de entrada del crudo al H202, psi.
 TCh_H202: Temperatura de chimenea gases H202, °F.
 O2_H202: Oxígeno en el flujo de aire al H202, %.
 Tfl_T204: Temperatura zona flash T202, °F.
 G_T204: Flujo gas de cima T202, CFSD.
 N_T204: Flujo de nafta T204, bpd.
 JET_T204: Flujo de jet T204, bpd.
 ACPM_T204: Flujo de acpm T204, bpd.
 GASO_T204: Flujo de gasóleo T204, bpd.
 FON_T204: Flujo de fondo T204, bpd.
 RC_T204: Flujo de reciclo de cima T204, bpd.
 RM_T204: Flujo de reciclo medio T204, bpd.
 RI_T204: Flujo de reciclo inferior T204, bpd.
 GASO1_T205: Flujo de gasóleo 1 T205, bpd.
 GASO2_T205: Flujo de gasóleo 2 T205, bpd.
 FON_T205: Flujo de fondo T205, bpd.
 RC_T205: Flujo de reciclo de cima T205, bpd.
 RM_T205: Flujo de reciclo medio T205, bpd.
 RI2_T205: Flujo de reciclo inferior T205, bpd.
 DP_H201: Caída de presión H201.
 DP_H202: Caída de presión H202.
 C*DT/COM_H201: Aproximación 1 a la relación Flujo de gas combustible H201 y capacidad calorífica del crudo.
 C*DP/COM_H201: Aproximación 2 a la relación Flujo de gas combustible H201 y capacidad calorífica del crudo.
 C*DT/COM_H202: Aproximación 1 a la relación Flujo de gas combustible H202 y capacidad calorífica del crudo.
 C*DP/COM_H202: Aproximación 2 a la relación Flujo de gas combustible H202

y capacidad calorífica del crudo.

Variable auxiliar contenido energético gases chimenea

TCh*O2_H201: H201.

Variable auxiliar contenido energético gases chimenea

TCh*O2_H202: H202.

Anexo B Estadísticos de los datos históricos

	E_C	API	BSW	S	T10	T30	T50	T70	T90	COM	G_H201	Te_H201	Ts_H201	Pe_H201	TCh_H201	O2_H201	Tfi_T201
Mean	71.82309	22.78735	0.1785384	1.085557	275.5361	548.8832	740.0837	956.5753	1185.618	982.5597	134.8356	462.2933	705.7915	214.9975	734.3244	3.393443	698.1176
Max	156.7	25.4	0.8	1.3	402.2	652.3	882.9	1262.3	1323.1	2317	162	574	757	309	792	10	816
Min	35.25	20.3	0.05	0.9	122.1	279.3	594.8	811.1	1045.1	621	55	339	531	133	587	1	509
Range	121.45	5.1	0.75	0.4	280.1	373	288.1	451.2001	278	1696	107	235	226	176	205	9	307
Std	10.61328	0.8813617	0.1246622	0.07567155	49.77393	51.19239	51.37012	97.74494	57.45919	147.1149	16.79404	37.3576	13.56111	29.0877	19.47153	1.056121	30.63515
Variance	112.6418	0.7767985	0.01554067	0.00572618	2477.444	2620.661	2638.89	9554.073	3301.559	21642.8	282.0397	1395.59	183.9037	846.0941	379.1404	1.115391	938.5125
Clase 0																	
Mean	67.9774	23.063	0.1348118	1.132627	276.0471	546.0289	731.9096	925.878	1191.981	987.1469	125.9362	550.7891	714.473	277.671	752.9301	3.793843	712.3542
Max	95.43121	24.4	0.3	1.2	284.3	549.7	735.5	929	1194.8	1597.2	146.4098	574.4599	728.6273	308.6473	775.9625	6.296724	721.0732
Min	51.57132	21.7	0.05	1	174.8	499.3	672.9	862.4	1134.5	623.8	95.94034	499.8265	693.5154	195.9061	711.3979	1.99863	684.2976
Range	43.8599	2.699999	0.25	0.2	109.5	50.40002	62.59998	66.59998	60.30005	973.4	50.46948	74.63339	35.11188	112.7413	64.56458	4.298095	36.77563
Std	7.071908	0.8581482	0.082341	0.08573584	28.58471	12.52607	12.33178	10.89104	9.838198	168.1993	9.413272	10.83416	4.360001	17.92165	12.5257	0.7552155	4.329659
Variance	50.01188	0.7364184	0.00678004	0.00735063	817.0856	156.9024	152.0728	118.6147	96.79014	28291.01	88.6097	117.3789	19.00961	321.1856	156.8932	0.5703505	18.74594
Clase 1																	
Mean	76.11983	22.74854	0.1902176	1.082034	247.4321	524.7985	719.4758	928.1696	1167.124	1003.046	136.2283	446.265	708.2485	191.4666	726.4909	3.723507	695.3653
Max	156.6957	25.4	0.8	1.3	352.5	618.9	829.5	1211.3	1293.8	2317.4	159.6623	514.1625	726.6567	249.5974	765.5476	10.18555	816
Min	35.25315	21.3	0.05	0.9	122.1	279.3	594.8	811.1	1072.4	706	66.83935	339.4745	553.8453	133.0502	587.2581	0.7532148	508.7197
Range	121.4425	4.1	0.75	0.4	230.4	339.6	234.7	400.2001	221.4	1611.4	92.82298	174.688	172.8114	116.5473	178.2896	9.432332	307.2803
Std	10.18715	0.7902364	0.1260572	0.06941746	50.0412	66.60442	45.86415	77.25297	44.40655	169.0134	12.5752	22.16785	14.04049	13.24556	15.85458	1.483865	28.41455
Variance	103.778	0.6244736	0.01589041	0.00481878	2504.122	4436.148	2103.521	5968.021	1971.942	28565.52	158.1356	491.4135	197.1354	175.4448	251.3678	2.201856	807.3863
Clase 2																	
Mean	70.52087	22.73186	0.2193673	1.097329	307.585	576.0594	770.5916	1019.271	1223.57	931.6475	144.003	440.7547	702.9918	200.0397	731.3027	3.111059	702.1486
Max	110.4189	25.4	0.6	1.3	396.5	647.3	882.9	1262.3	1305.9	1685.054	161.5244	489.3327	725.873	236.5835	771.9158	6.00793	816
Min	48.03662	21.4	0.1	0.9	234.4	512.7	682.8	858.6	1103.4	638.9	106.4013	343.6946	534.1895	158.694	666.6442	1.827751	640.2114
Range	62.38226	4	0.5	0.4	162.1	134.6	200.1	403.7001	202.5	1046.154	55.12305	145.6381	191.6836	77.88953	105.2717	4.180179	175.7886
Std	6.806979	0.9369388	0.1124801	0.07600105	39.18182	33.84859	50.25879	121.3082	55.30997	99.30087	8.425617	15.76125	12.88302	12.76332	15.00067	0.5087022	45.16307
Variance	46.33497	0.8778544	0.01265178	0.00577616	1535.215	1145.727	2525.946	14715.68	3059.193	9860.662	70.99104	248.4169	165.9723	162.9022	225.02	0.2587779	2039.703
Clase 3																	
Mean	63.06929	22.73852	0.1439187	1.088983	282.7892	552.1062	734.7538	933.0541	1190.452	1011.829	121.2164	471.829	707.5886	231.8284	716.1649	3.291124	705.902
Max	91.52386	24.5	0.5	1.2	330.6	618.9	833.3	1124.4	1323.1	1612.8	158.2392	527.7839	757.4553	263.9327	747.7633	7.003047	724.2467
Min	47.55764	21.8	0.05	1	239.2	499.3	672.9	862.4	1134.5	674.1	61.6201	412.8984	555.1974	175.9091	677.3892	0.6628814	664.0014
Range	43.96623	2.700001	0.45	0.2	91.40001	119.6	160.4	262	188.6	938.7001	96.61906	114.8855	202.2578	88.02365	70.37408	6.340166	60.2453
Std	7.172285	0.7120726	0.1214618	0.06538925	19.86567	22.95471	33.65229	58.80779	48.97599	149.9069	22.12171	22.40761	12.28907	12.009	11.50388	0.7060697	10.07006
Variance	51.44167	0.5070473	0.01475297	0.00427575	394.645	526.9187	1132.476	3458.356	2398.648	22472.09	489.3701	502.1011	151.0213	144.2162	132.3393	0.4985344	101.406
Clase 4																	
Mean	84.08171	22.30833	0.1992041	1.017796	247.2523	526.1974	700.555	885.4187	1103.979	1037.895	135.7789	490.9981	690.6693	242.7204	762.1921	3.00785	674.4636
Max	123.4202	24.2	0.8	1.1	349.3	587.3	773.1	1028.8	1194.8	1395.5	149.8513	532.6753	729.5411	269.3309	792.2	5.227216	716.1423
Min	55.34448	20.3	0.05	1	168.5	472.4	646.3	817.7	1045.1	621.3	102.4221	456.6592	641.3414	210.9499	707.9819	1.378508	627.1123

	E_C	API	BSW	S	T10	T30	T50	T70	T90	COM	G_H201	Te_H201	Ts_H201	Pe_H201	TCh_H201	O2_H201	Tfl_T201
Range	68.07573	3.900002	0.75	0.1	180.8	114.9	126.8	211.1	149.7001	774.2	47.42917	76.01605	88.19977	58.381	84.21814	3.848708	89.02997
Std	13.00408	0.9020126	0.1786147	0.03825565	59.58648	37.30139	34.74152	44.46185	54.31303	161.1653	7.325013	13.36309	10.42559	10.86529	13.33872	0.5784011	10.18278
Variance	169.1061	0.8136267	0.03190321	0.0014635	3550.549	1391.394	1206.973	1976.856	2949.905	25974.27	53.65581	178.5721	108.6929	118.0546	177.9215	0.3345479	103.689

	G_T201	N_T201	JET_T201	ACPM_T201	GASO_T201	FON_T201	RC_T201	RM_T201	RI_T201	Te_H202	Ts_H202	Pe_H202	TCh_H202	O2_H202	Tfl_T204	G_T204	N_T204
Mean	13506.52	6210.435	3829.599	11350.54	1163.628	19755.89	20170.34	21353.79	11554.64	472.3381	736.8741	332.3586	725.5978	3.50769	711.6619	8192.026	2781.854
Max	20612	9973	8492	15959	5044	33476	49665	24084	18470	700	849	422	801	10	744	49994	5312
Min	13	0	758	53	0	0	233	1509	0	229	534	147	612	1	507	0	0
Range	20599	9973	7734	15906	5044	33476	49432	22575	18470	471	315	275	189	9	237	49994	5312
Std	6728.041	940.2222	983.9852	1718.864	1015.59	9252.697	7743.212	1839.094	2431.186	91.8362	44.18342	32.01697	34.11649	0.8201955	7.967998	11561.23	825.1243
Variance	4.53E+07	884017.8	968227	2954493	1031423	8.56E+07	6.00E+07	3382269	5910663	8433.889	1952.175	1025.086	1163.935	0.6727206	63.48899	1.34E+08	680830.1

Class 0

Mean	3640.134	7800.014	6523.705	13260.83	1751.455	2828.557	26063.44	20796.86	12199.67	658.0967	829.4449	354.725	717.3831	3.859308	723.6241	35102.75	848.4665
Max	7142.439	9973.442	8492.229	15958.81	3538.31	11682.25	33656.14	21923.49	15747.63	699.6091	849.2128	420.1736	795.2826	10.2	744.4398	49994.3	5312.274
Min	153.9899	5009.431	2606.126	8875.866	402.5746	0	8174.772	1508.61	1536.578	598.8662	765.2042	146.9636	616.788	0.9662269	685.5443	12507.26	0
Range	6988.449	4964.012	5886.102	7082.944	3135.735	11682.25	25481.37	20414.88	14211.06	100.7429	84.00861	273.21	178.4946	9.233773	58.89551	37487.04	5312.274
Std	1266.258	863.1016	1109.792	1439.762	603.5419	2442.439	3054.393	2164.906	1310.474	34.62989	7.325779	52.61349	48.69713	1.113444	8.856478	6636.185	464.2616
Variance	1.60E+06	744944.4	1231637	2072915	364262.8	5.97E+06	9.33E+06	4686816	1717343	1199.229	53.66704	2768.179	2371.41	1.239757	78.43719	4.40E+07	215538.9

Class 1

Mean	17475.87	5643.958	3420.552	10500.44	987.4763	24783.44	17475.98	21951.15	10483.79	414.5483	714.8334	312.2052	731.8293	3.348211	707.7209	2994.552	3363.123
Max	20612.2	7726.339	4986.565	14389.08	5044.038	33476.2	49664.53	23980.28	16816.01	473.0757	728.6431	360.7282	779.0416	8.788181	722.8372	9513.309	4245.732
Min	3856.989	1.37E-07	1005.817	53.29693	0	17702.92	7789.151	12115.34	0	228.6494	533.8319	216.0466	611.801	0.6865893	507.45	38.0868	1679.6
Range	16755.21	7726.339	3980.749	14335.78	5044.038	15773.28	41875.38	11864.94	16816.01	244.4263	194.8112	144.6816	167.2407	8.101592	215.3872	9475.222	2566.132
Std	3505.168	629.655	393.9399	1555.283	1084.821	2137.038	7227.959	667.2548	2639.823	21.20666	5.373315	25.74056	23.5705	0.6921797	6.754736	564.666	299.565
Variance	1.23E+07	396465.4	155188.7	2418907	1176837	4.57E+06	5.22E+07	445229	6968663	449.7222	28.87252	662.5764	555.5686	0.4791128	45.62646	3.19E+05	89739.2

Class 2

Mean	14736.05	5988.66	3332.778	11579.96	524.1229	26113.01	25589.11	21843.86	10308.28	435.1442	716.6863	323.251	723.6806	3.456141	710.9371	2850.87	3112.323
Max	20612.2	7834.038	4781.628	14915.06	4619.852	32066.82	36582.07	22911.04	17695.17	463.8304	729.0062	369.8686	785.2528	7.646107	724.553	8719.825	4015.299
Min	3625.516	1335.466	1655.894	6274.132	0	17347.18	9630.19	15630.6	5743.688	408.001	702.4267	271.7994	686.8372	1.562624	683.1677	895.3015	1591.425
Range	16986.69	6498.572	3125.734	8640.926	4619.852	14719.63	26951.88	7280.438	11951.48	55.82941	26.57953	98.06918	98.41559	6.083483	41.38525	7824.524	2428.874
Std	5208.746	750.2652	356.8095	1638.628	421.9868	1890.823	7811.936	342.0336	1536.341	10.14519	3.673121	13.27829	12.90669	0.6658372	4.030019	448.3041	634.8953
Variance	2.71E+07	562897.9	127313	2685102	178072.9	3575213	6.10E+07	116987	2360344	102.9249	13.49182	176.3131	166.5826	0.4433392	16.24105	200976.6	403092.1

Class 3

Mean	3742.76	5971.309	3891.115	11645.87	1926.949	5387.464	16481.69	20325.11	11875.69	477.32	721.7541	345.9351	676.4469	3.468644	716.2796	2435.67	2170.202
Max	6778.809	7793.181	5598.107	14557.54	4408.184	19945.58	37599	21773.41	15335.49	536.9323	735.9069	373.1	737.5338	8.850112	726.4172	5983.771	2734.281
Min	459.5914	3239.747	757.9868	2657.832	522.8417	0	233.0354	13677.01	9688.632	403.6697	661.3571	242.5848	618.6665	0.9773397	657.748	0	1485.279
Range	6319.217	4553.434	4840.12	11899.71	3885.342	19945.58	37365.96	8096.396	5646.856	133.2626	74.5498	130.5152	118.8673	7.872772	68.66925	5983.771	1249.003
Std	965.4871	575.3662	497.0064	1315.064	641.2769	3014.447	6611.988	516.2742	983.6954	29.24219	4.613046	14.4616	27.06799	0.7934919	3.951307	1337.296	195.8161
Variance	9.32E+05	331046.3	247015.3	1729392	411236.1	9086889	4.37E+07	266539	967656.6	855.1055	21.2802	209.1379	732.6758	0.6296294	15.61283	1788360	38343.93

Class 4

Mean	18488.49	7358.818	3986.586	12908.58	1576.622	12562	17049.38	22536.91	15212.88	667.4109	836.1802	382.0503	763.0446	3.797878	711.661	30852.89	2550.568
Max	20612.2	9262.008	6970.613	14839.06	4228.024	21482.75	30388.86	23293.29	18470.45	694.0499	849.4066	405.844	800.979	8.690792	737.324	45316.81	2986.881
Min	3255.95	4963.681	2471.174	7587.011	624.9156	2900.354	8066.327	15904.62	5375.367	626.7521	805.7499	233.039	632.1634	1.617974	672.2801	14238.17	422.6971

	G_T201	N_T201	JET_T201	ACPM_T201	GASO_T201	FON_T201	RC_T201	RM_T201	RI_T201	Te_H202	Ts_H202	Pe_H202	TCh_H202	O2_H202	Tfl_T204	G_T204	N_T204
Range	17356.25	4298.327	4499.439	7252.051	3603.109	18582.4	22322.54	7388.665	13095.08	67.29779	43.65662	172.805	168.8156	7.072818	65.04388	31078.63	2564.184
Std	1095.911	568.7714	489.8983	1140.449	489.5011	1784.106	3243.29	717.968	1118.661	9.791236	2.919033	9.548728	29.60966	0.7344957	8.939549	3825.765	189.7101
Variance	1.20E+06	323500.8	240000.4	1300624	239611.3	3183034	1.05E+07	515478.1	1251403	95.86831	8.520752	91.1782	876.7317	0.5394839	79.91554	1.46E+07	35989.9

	JET_T204	ACPM_T204	GASO_T204	FON_T204	RC_T204	RM_T204	RI_T204	GASO1_T205	GASO2_T205	FON_T205	RC_T205	RM_T205	RI2_T205	DP_H201	DP_H202	C*DT/PC_H201
Mean	2150.419	5386.139	1016.24	9507.341	8373.333	15318.47	6203.495	2230.936	3653.573	8870.379	5339.867	13149.42	1116.942	199.6606	313.1801	11464.66
Max	5136	10455	3952	20297	16234	20211	11830	3696	7088	25981	11109	17918	6320	293	389	24700.9
Min	5	628	0	7	0	1233	0	0	0	0	0	0	0	125	121	-22479.9
Range	5131	9827	3952	20290	16234	18978	11830	3696	7088	25981	11109	17918	6320	168	268	47180.8
Std	789.3657	948.084	593.548	3021.697	1932.579	2045.059	2422.665	995.4213	2077.719	4255.959	4786.749	3359.102	1311.142	28.32894	29.86956	3469.913
Variance	623098.1	898863.3	352299.2	9130651	3734862	4182266	5869307	990863.5	4316917	1.81E+07	2.29E+07	1.13E+07	1719093	802.5289	892.1906	1.20E+07

Class 0																
Mean	2268.58	6208.67	1010.305	15443.54	10314.47	13372.53	6285.064	1822.282	4054.757	13512.91	8057.184	11412.38	1871.613	261.0194	318.2437	5796.358
Max	5135.587	10455.45	3951.908	20296.98	16120.87	16859.97	9812.35	3586.315	6010.404	25980.92	11109.45	15062.52	4560.187	292.5119	388.0797	16733
Min	5.381019	1382.01	18.66789	1508.209	1274.82	7321.033	0	1.675274	0	2212.74	3960.684	5537.921	0	181.7879	121.0542	-21065.91
Range	5130.206	9073.439	3933.24	18788.77	14846.05	9538.938	9812.35	3584.639	6010.404	23768.18	7148.769	9524.596	4560.187	110.724	267.0255	37798.91
Std	1480.582	1167.697	1208.745	3661.268	2843.526	1727.717	2409.523	975.0427	995.7697	2493.749	1956.705	1449.728	1554.226	17.7409	50.55907	1684.044
Variance	2192122	1363516	1461064	1.34E+07	8085640	2985006	5805800	950708.3	991557.3	6.22E+06	3.83E+06	2.10E+06	2415620	314.7394	2556.22	2.84E+06
	2268.58	6208.67	1010.305	15443.54	10314.47	13372.53	6285.064	1822.282	4054.757	13512.91	8057.184	11412.38	1871.613	261.0194	318.2437	5796.358

Class 1																
Mean	2513.123	5476.443	854.7321	9784.916	8493.974	14536.22	5486.125	2572.324	2850.698	8143.094	3578.203	13682.28	421.3392	176.7479	295.4989	12050.57
Max	4515.402	7066.478	3892.261	14261.17	15399.52	18476.2	10546.36	3696.191	7087.918	25911.95	11109.35	17409.16	3819.163	234.7776	342.5858	18517.42
Min	7.371612	627.6795	0	4542.335	5208.263	3907.059	1799.762	0	0	0	0	0	0	124.7044	201.9872	4022.032
Range	4508.031	6438.798	3892.261	9718.838	10191.26	14569.14	8746.602	3696.191	7087.918	25911.95	11109.35	17409.16	3819.163	110.0732	140.5986	14495.38
Std	544.5283	584.9364	426.988	1529.594	779.7685	1737.325	2676.397	941.4727	2123.644	4881.112	4884.026	3986.607	709.3566	13.05917	25.01522	1959.656
Variance	296511.1	342150.6	182318.8	2.34E+06	608038.9	3018297	7163100	886370.9	4509862	2.38E+07	2.39E+07	1.59E+07	503186.8	170.542	625.7614	3.84E+06
	2513.123	5476.443	854.7321	9784.916	8493.974	14536.22	5486.125	2572.324	2850.698	8143.094	3578.203	13682.28	421.3392	176.7479	295.4989	12050.57

Class 2																
Mean	2125.075	5486.466	1092.599	7924.311	8497.619	15927.8	7853.718	2856.245	4382.683	8230.26	9033.998	14674.74	803.6902	185.5111	306.4417	13864.49
Max	3946.84	6573.25	3951.887	13882.6	10841.17	19048.65	11830.48	3696.191	7087.918	25911.21	11109.14	17918.5	3014.847	223.6305	353.6852	20276.9
Min	131.6418	2941.332	213.2543	5000.817	6811.519	11471.23	4987.284	0	0	0	0	5544.296	0	147.0926	257.6409	3259.205
Range	3815.198	3631.917	3738.633	8881.781	4029.649	7577.423	6843.2	3696.191	7087.918	25911.21	11109.14	12374.2	3014.847	76.53789	96.04425	17017.7
Std	336.8412	459.4715	427.9915	886.611	559.4454	1282.977	1242	527.1908	3118.55	3636.229	3822.657	1536.25	800.9518	11.19928	13.7373	1777.667
Variance	113462	211114	183176.7	786079.1	312979.2	1646030	1542564	277930.1	9725351	1.32E+07	1.46E+07	2360064	641523.9	125.4239	188.7134	3160099

Class 3																
Mean	2082.343	5980.852	892.1991	7645.072	9078.159	14323.17	7356.574	954.0602	3733.973	10824.84	6998.029	8781.57	2335.496	216.8082	330.9844	10355.19
Max	3179.788	7735.142	2900.216	9580.168	16234.03	17468.31	10949.62	2744.41	6390.335	24232.45	11109.45	13204.76	6023.23	249.1522	359.3343	16735.53
Min	158.7727	1554.199	147.8412	3072.518	2005.932	1233.344	725.8421	0	0	4924.799	3276.293	5139.336	7.45E-20	162.7306	228.0418	-22323.78
Range	3021.016	6180.943	2752.375	6507.65	14228.1	16234.97	10223.78	2744.41	6390.335	19307.65	7833.159	8065.42	6023.23	86.42166	131.2924	39059.31
Std	451.4532	464.1901	311.007	668.6059	1615.343	1464.433	1399.116	533.3681	878.5079	1821.385	1835.804	987.5621	1489.616	11.9032	15.08159	4520.082
Variance	203810	215472.4	96725.35	447033.9	2609334	2144565	1957527	284481.5	771776.3	3.32E+06	3.37E+06	975278.8	2218957	141.6863	227.4543	2.04E+07
	2082.343	5980.852	892.1991	7645.072	9078.159	14323.17	7356.574	954.0602	3733.973	10824.84	6998.029	8781.57	2335.496	216.8082	330.9844	10355.19

Class 4																
Mean	723.1434	3976.356	1208.588	11393.36	4673.397	18329.68	4708.637	2094.419	3988.153	8871.628	1731.064	15233.52	1232.2	225.8913	355.8822	7288.409
Max	2267.226	7146.48	3774.583	19564.58	10137.77	20211.31	9416.767	3696.191	7087.918	25186.08	11100.46	17522.43	3579.672	253.2613	379.4353	10650.8

	JET_T204	ACPM_T204	GASO_T204	FON_T204	RC_T204	RM_T204	RI_T204	GASO1_T205	GASO2_T205	FON_T205	RC_T205	RM_T205	RI2_T205	DP_H201	DP_H202	C*DT/PC_H201
Min	12.2356	1200.63	11.06156	7.153816	0	2145.84	1593.025	0	936.6201	2171.704	0	11344.88	0	195.5055	196.3538	4481.837
Range	2254.99	5945.85	3763.521	19557.42	10137.77	18065.47	7823.742	3696.191	6151.298	23014.38	11100.46	6177.555	3579.672	57.7558	183.0815	6168.961
Std	285.9356	1041.889	574.62	5271.312	2732.434	1611.693	1534.49	623.4026	820.8955	1799.899	3733.919	810.5223	555.6281	10.79039	10.10763	1010.174
Variance	81759.2	1085533	330188.1	2.78E+07	7466194	2597555	2354658	388630.8	673869.4	3.24E+06	1.39E+07	656946.4	308722.5	116.4325	102.1641	1020451

	C*DP/PC_H201	C*DT/PC_H202	C*DP/PC_H202	TCh*O2_H201	TCh*O2_H202
Mean	9269.678	6449.838	7660.213	24.83322	25.48787
Max	16298	11629.85	17882.3	76.33	71.52
Min	2916.7	1789.11	1867.2	0	0
Range	13381.3	9840.739	16015.1	76.33	71.52
Std Deviation	1759.184	1553.681	1489.852	7.401023	5.87023
Variance	3094728	2413926	2219660	54.77514	34.4596

Clase 0

Mean	9351.367	5070.047	9395.687	28.54698	27.50124
Max	14842.15	9644.364	17828.68	46.74339	71.52412
Min	4348.872	1789.11	1867.224	14.81255	0
Range	10493.28	7855.254	15961.46	31.93084	71.52412
Std Deviation	1672.622	1723.23	2780.171	5.592339	7.757726
Variance	2797665	2969522	7729352	31.27425	60.18231
	9351.367	5070.047	9395.687	28.54698	27.50124

Clase 1

Mean	8145.369	7281.624	7190.381	27.01467	24.53054
Max	12339.5	11629.85	11412.19	76.32751	60.13092
Min	2916.745	2852.665	2910.471	5.097551	0
Range	9422.75	8777.182	8501.717	71.22997	60.13092
Std Deviation	1427.531	1212.417	1348.913	10.63464	5.286281
Variance	2037845	1469956	1819565	113.0956	27.94476
	8145.369	7281.624	7190.381	27.01467	24.53054

Clase 2

Mean	9833.919	6942.501	7569.917	22.73208	25.04062
Max	15988.25	9601.572	11812.12	43.87267	59.06598

	C*DP/PC_H201	C*DT/PC_H202	C*DP/PC_H202	TCh*O2_H201	TCh*O2_H202
Min	6089.095	3958.179	4087.501	0	11.21176
Range	9899.159	5643.394	7724.615	43.87267	47.85422
Std Deviation	1460.772	806.1874	1009.08	3.657093	5.040001
Variance	2133854	649938.1	1018242	13.37433	25.40161
Clase 3					
Mean	10077.22	5633.929	7652.672	23.55294	23.38664
Max	16297.99	8557.785	11227.17	48.37343	63.01131
Min	5581.865	3353.159	4414.672	4.649364	0
Range	10716.13	5204.626	6812.502	43.72406	63.01131
Std Deviation	1607.047	913.2361	1123.164	4.967506	5.744198
Variance	2582599	834000.1	1261498	24.67612	32.99582
	10077.22	5633.929	7652.672	23.55294	23.38664
Clase 4					
Mean	8306.065	3845.83	8094.486	22.9016	29.00914
Max	14652.15	7016.224	14974.27	38.62811	66.12581
Min	5000.284	1979.956	3299.354	10.15822	0
Range	9651.863	5036.268	11674.91	28.46988	66.12581
Std Deviation	1490.527	667.5256	1251.049	4.274816	5.949234
Variance	2221671	445590.4	1565123	18.27405	35.39338