

DESARROLLO DE UN MODELO SIMCA PARA LA CLASIFICACIÓN DE KEROSINAS MEDIANTE EL EMPLEO DE LA ESPECTROSCOPÍA INFRARROJA

Ángel Dago Morales*, Alberto Cavado Osorio, Reinaldo Fernández Fernández y Ernesto Linchenat Dennes

Centro de Investigaciones del Petróleo, Washington 169 esquina a Churrucá, Municipio Cerro, Ciudad de la Habana, Cuba

Recebido em 16/10/07; aceito em 29/11/07; publicado na web em 26/8/08

DEVELOPMENT OF A SIMCA MODEL FOR CLASSIFICATION OF KEROSENE BY INFRARED SPECTROSCOPY. In the petroleum refining industry, the use of crude from several origins is frequent. This leads to a product of variable chemical composition during refining, hindering quality control. Therefore, it is important to develop classification models that help to better characterize those products. The objective of this study is to develop a SIMCA recognition pattern to classify kerosene using infrared spectroscopy data. The model permits to differentiate two kerosene groups with different chemical compositions, which was corroborated by mass spectrometry.

Keywords: SIMCA model; kerosene classification; infrared spectroscopy.

INTRODUCCIÓN

En el control de la calidad de los destilados medios del petróleo se emplean procedimientos de ensayo que han sido oficialmente establecidos como procedimientos de referencia por diferentes organizaciones internacionales; como por ejemplo; la American Society for Testing Materials (ASTM), la ISO (Internacional Standard Organization) y el IP (The Institute of Petroleum). Es conocido que estos métodos de referencia aunque son exactos emplean un tiempo de medición considerable y requieren, por lo general, de apreciable cantidad de muestra.

Los métodos quimiométricos han tenido gran aplicación en la industria de refinación del petróleo; por ejemplo, los métodos de calibración multivariada para el desarrollo de procedimientos alternativos para la predicción simultánea de propiedades físico químicas en derivados como gasolinas, kerosinas, turbocombustibles y diesel.¹⁻⁹ La exactitud de estos procedimientos alternativos es comparable con la exactitud de los métodos de referencia; presentan además las ventajas de utilizar pequeñas cantidades de muestras y de que el tiempo invertido en el análisis es pequeño. Por otra parte los métodos quimiométricos de clasificación y reconocimiento de patrones han sido ampliamente utilizados en la detección de adulteraciones de combustibles, así como en el control de calidad de los procesos productivos.¹⁰⁻¹³

En nuestro laboratorio se han desarrollado modelos de regresión multivariados para la predicción simultánea de propiedades físico químicas de kerosinas mediante el empleo de la espectroscopía infrarroja de rango medio (FT-MIR) y el método de mínimos cuadrados parciales (PLS). El establecimiento de estas metodologías analíticas alternativas no culmina con el desarrollo y posterior validación del modelo; es necesario establecer un procedimiento de control de su validez con respecto al tiempo, principalmente en la industria de refinación del petróleo, donde ocurren con frecuencia variaciones de la materia prima y de los propios procesos productivos, que originan cambios en la composición química de los productos terminados; y por ende, problemas con la predicción de los modelos.¹⁴⁻¹⁶ Para minimizar estas dificultades relacionadas con la detección de situaciones anómalas originadas por muestras atípicas, o no contempladas en el

proceso de calibración, es necesario desarrollar modelos de clasificación que permitan identificar las muestras que se encuentran fuera de especificación y definir además cuando es necesario realizar una actualización de las calibraciones, y de esta manera lograr modelos más robustos.^{16,17}

El método SIMCA de reconocimiento de patrones (*Soft Independent Modeling of Class Analogy*) desde su introducción por Svante Wold en 1976¹⁸ ha sido uno de los métodos de clasificación más utilizados en el control de calidad en las industrias alimenticia¹⁹ y farmacéutica.²⁰ En la industria del petróleo se ha utilizado para la detección de adulteraciones en gasolinas.^{10,13} Es un método supervisado de reconocimiento de patrones que se basa en el principio de analogía entre las muestras que pertenecen a una misma clase, y emplea para el cálculo de las distancias los *scores* determinados mediante análisis por componentes principales (PCA). El método SIMCA calcula un modelo PCA para cada clase o categoría presente en el sistema objeto de estudio, posteriormente integra cada una de las clases y calcula sus límites o fronteras con una probabilidad dada, comúnmente del 95%.²¹

En caso de que no esté disponible la información externa, las categorías o clases se pueden definir a partir de una evaluación de las agrupaciones mediante el análisis exploratorio de datos (reconocimiento de patrones no supervisado) que brindan métodos como: el Análisis Jerárquico de Grupos (*Hierarchical Cluster Analysis, HCA*) y el Análisis por Componentes Principales (*Principal Components Analysis, PCA*). Una vez definidas las categorías, se construye y refina un modelo para el conjunto de muestras de entrenamiento conocidas *training set*; y después éste se utiliza para predecir las clases de nuevas muestras desconocidas *test set*. Una muestra desconocida o muestra problema puede asignarse a una categoría, a más de una, o a ninguna; y se calcula la probabilidad de cada asignación.

El objetivo de nuestro trabajo es desarrollar un modelo SIMCA de reconocimiento de patrones para clasificar los cortes de producción de kerosinas.

MATERIALES Y MÉTODOS

Se utilizaron 60 muestras de kerosinas provenientes de una refinería de la Unión Cuba Petróleo (CUPET). El análisis preliminar de componentes principales evidenció que las muestras se dividen en dos grupos diferentes, las cuales se identificaron como C y V.

*e-mail: adago@ceinpet.cupet.cu

destaca que se logra la correcta clasificación de todas las muestras y que no se presentan falsos negativos. Se obtuvo una distancia entre las clases de 3.80, la selectividad del modelo se considera adecuada cuando el valor de este parámetro es mayor de 3.^{21,26}

Tabla 1. Resumen de los resultados del cálculo del modelo SIMCA

Clases	Nº de muestras	Nº de factores	% de varianza explicada	Clasificados correctamente	Clasificados incorrectamente	No clasificados
C	26	5	93.4	26	-	-
V	14	5	88.4	14	-	-

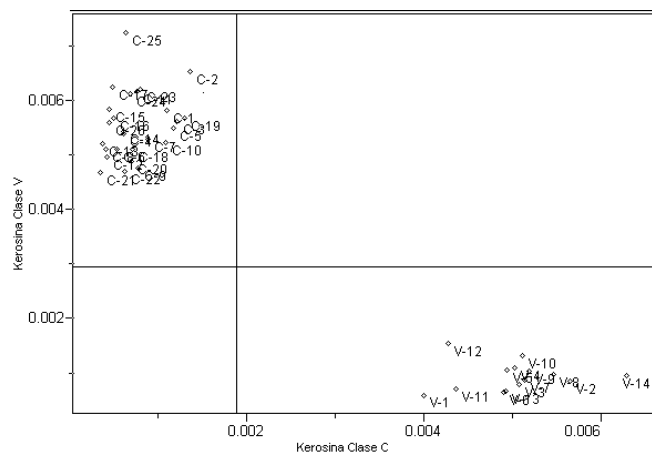


Figura 3. Modelo SIMCA: Distancias de las muestras del conjunto de calibración con respecto a las clases C y V (con líneas continuas se destacan los límites de las clases para un nivel de significación del 5%)

Validación del modelo SIMCA

El grupo de validación se conformó con 13 muestras de la clase C, 7 de la clase V y 10 muestras de turbocombustibles (T1-T10). Las muestras de turbocombustibles son ajenas al modelo desarrollado, y fueron incluidas en el grupo de validación para verificar la selectividad del mismo; es decir, para comprobar si el modelo es capaz de reconocerlas como que no pertenecen a ninguna de las clases.

En la Figura 4 se presenta el diagrama de Coomans, en el cual se reflejan las distancias de las muestras del conjunto de validación con respecto a las clases definidas por el modelo. Se observa como las muestras de validación de las clases C y V clasifican correctamente en sus respectivas clases y las muestras de turbocombustibles caen fuera de los límites que definen las fronteras de las clases. En la Figura 5 se reportan los estadígrafos Q residual y T² de Hotelling calculados para la clase V: se observa como las muestras de validación de la clase V caen dentro de los límites de confianza del 5% y las muestras pertenecientes a la clase C y las muestras de turbocombustibles caen fuera de los límites establecidos.

Los resultados de la validación fueron satisfactorios; el modelo desarrollado clasificó correctamente a las muestras del grupo de validación: todas las muestras fueron clasificadas en su clase respectiva, no se presentaron casos de falsos negativos, las 10 muestras de turbocombustibles fueron correctamente clasificadas como no pertenecientes a ninguna clase.

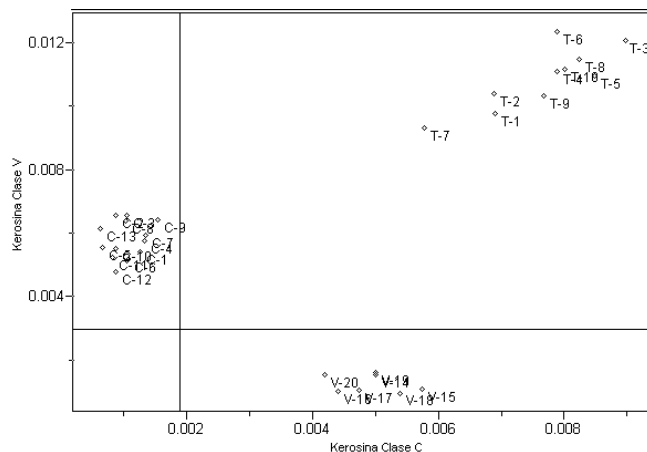


Figura 4. Validación del modelo SIMCA: Distancias de las muestras del grupo de validación con respecto a las clases C y V (con líneas continuas se destacan los límites de las clases para un nivel de significación del 5%)

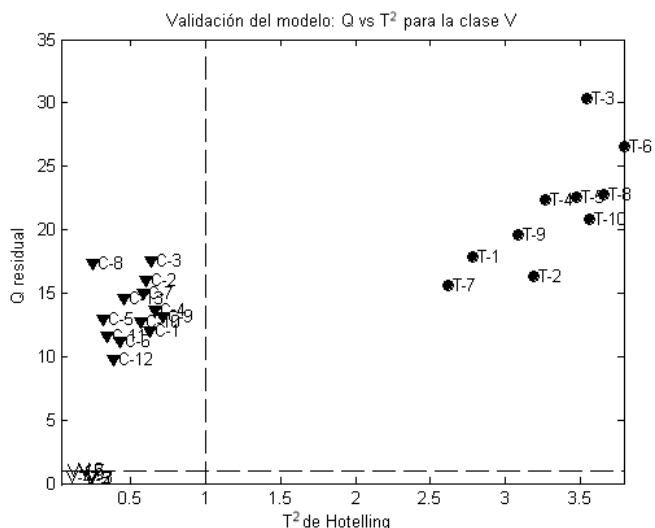


Figura 5. Validación del modelo SIMCA. Descriptores T² y residuales Q de la clase V (con líneas discontinuas se destacan los límites de la clase V para un nivel de significación del 5%)

CONCLUSIONES

Se demostró, mediante la aplicación de técnicas quimiométricas y cromatográficas, que en el proceso de producción se originan dos tipos de cortes de kerosinas, cuya diferencia fundamental está dada en la proporción de los componentes más pesados. Se desarrolló un modelo SIMCA de reconocimiento de patrones supervisado que permite la clasificación de las dos clases de kerosinas a partir de los datos de sus espectros infrarrojos. El modelo desarrollado posee una sensibilidad y selectividad adecuadas para el control de calidad del proceso productivo.

MATERIAL SUPLEMENTAR

A Figura 1S está disponibilizada gratuitamente em la pagina web <http://quimicanova.sbj.org.br> em la forma de un archivo .PDF.

AGRADECIMIENTOS

Los autores agradecen a la Dirección Técnica de CUPET el financiamiento de este trabajo.

REFERENCIAS

1. Fodor, G. E.; Mason, R. A.; Hutzler, S. A.; *Appl. Spectrosc.* **1999**, *53*, 1292.
2. Wentzell, P. D.; Andrews, D. T.; Walsh, J. M.; Cooley, J. M.; *Can. J. Chem.* **1999**, *77*, 391.
3. Andrade, J. M.; Garrigues, S.; de la Guardia, M.; Gómez-Carracedo, M.; Prada, D.; *Anal. Chim. Acta* **2003**, *482*, 115.
4. Vianney Santos, O. Jr.; Flavia Oliveira, C. C.; Daniella Lima, G.; Andrea Petry, C.; Garcia, E.; Paulo Suarez, A. Z.; Joel Rubim, C.; *Anal. Chim. Acta* **2005**, *547*, 188.
5. Côcco, L. C.; Yamamoto, C. I.; von Meien, O. F.; *Chem. Intel. Lab. Syst.* **2005**, *76*, 55.
6. Pasadakis, N.; Sourligas, S.; Foteinopoulos, Ch.; *Fuel* **2006**, *85*, 1131.
7. Brudzewski, K.; Kesik, A.; Kolodziejczyk, K.; Zborowska, U.; Ulaczyk, J.; *Fuel* **2006**, *85*, 553.
8. Dago-Morales, A.; Fernández, F. R.; Ruiz-Martínez, M. D.; Balmayor, M. M.; Laza, N. M.; Ross, J. E.; Otero, Z. M.; Simeón, A. L.; *Rev. CENIC, Ciencias Químicas* **2006**, *37*, 3.
9. Balabin, R. M.; Safieva, R. Z.; Lomakina, E. I.; *Chem. Intel. Lab. Syst.* **2007**, *88*, 183.
10. Santos de Oliveira, F.; Gomes Teixeira, L. S.; Ugulino Araujo, M. C.; Korn, M.; *Fuel* **2004**, *83*, 917.
11. Wiedemann, L. S. M.; d' Avila, L. A.; Azevedo, D. A.; *Fuel* **2005**, *84*, 467.
12. Pasadakis, N.; Kardamakis, A.; *Anal. Chim. Acta* **2006**, *578*, 250.
13. Flumignan, D. L.; Tininis, A. G.; Ferreira, F. O.; de Oliveira, J. E.; *Anal. Chim. Acta* **2007**, *595*, 128.
14. Andrade, J. M.; García, M. V.; López-Mahía, P.; Prada, D.; *Talanta* **1997**, *44*, 2167.
15. Swierenga, H.; de Weijer, A. P.; Van Wijk, R. J.; Buydens, L. M.; *Chemom. Intell. Lab. Syst.* **1999**, *49*, 1.
16. García-Mencía, M. V.; Andrade, J. M.; López-Mahía, P.; Prada, D.; *Fuel* **2000**, *79*, 1823.
17. Daszyłowski, M.; Kaczmarek, K.; Vander Heyden, Y.; Walczak, B.; *Chem. Intell. Lab. Syst.* **2007**, *85*, 203.
18. Wold, S.; *Pattern Recognition* **1976**, *8*, 127.
19. Forina, M.; Lanteri, S. In *Chemometrics. Mathematics and Statistics in Chemistry*; Kowalski, B. R., ed.; D. Reidel Publishing Company: Reidel, Dordrecht, 1984, p. 305.
20. Gabrielsson, J.; Lindberg, N. O.; Lundstedt, T.; *J. Chem.* **2002**, *16*, 141.
21. Ferreira, M. C.; Antunes, A. M.; Melgo, M. S.; Volpe, P. L. O.; *Quim. Nova* **1999**, *22*, 724.
22. *Konikrom Plus Software, version 2.3*, Konik-Tech Corporation, 2004.
23. *Nicolet's OMNIC Spectroscopy Software, version 5.2a*, Nicolet Instrument Corporation, Copyright (1992-2000).
24. *PLS_Toolbox 3.5 for use with MATLAB*, Eigenvector Research, Inc., 2004, version 3.5.
25. *MATLAB. The Language of Technical Computing*, The Math Works, 2003, version 6.5.
26. *Pirouette*, Infometrix Inc., Woodinville WA., 2003, version 3.11.

DESARROLLO DE UN MODELO SIMCA PARA LA CLASIFICACIÓN DE KEROSINAS MEDIANTE EL EMPLEO DE LA ESPECTROSCOPÍA INFRARROJA

Ángel Dago Morales*, Alberto Cavado Osorio, Reinaldo Fernández Fernández y Ernesto Linchenat Dennes
Centro de Investigaciones del Petróleo, Washington 169 esquina a Churruca, Municipio Cerro, Ciudad de la Habana, Cuba

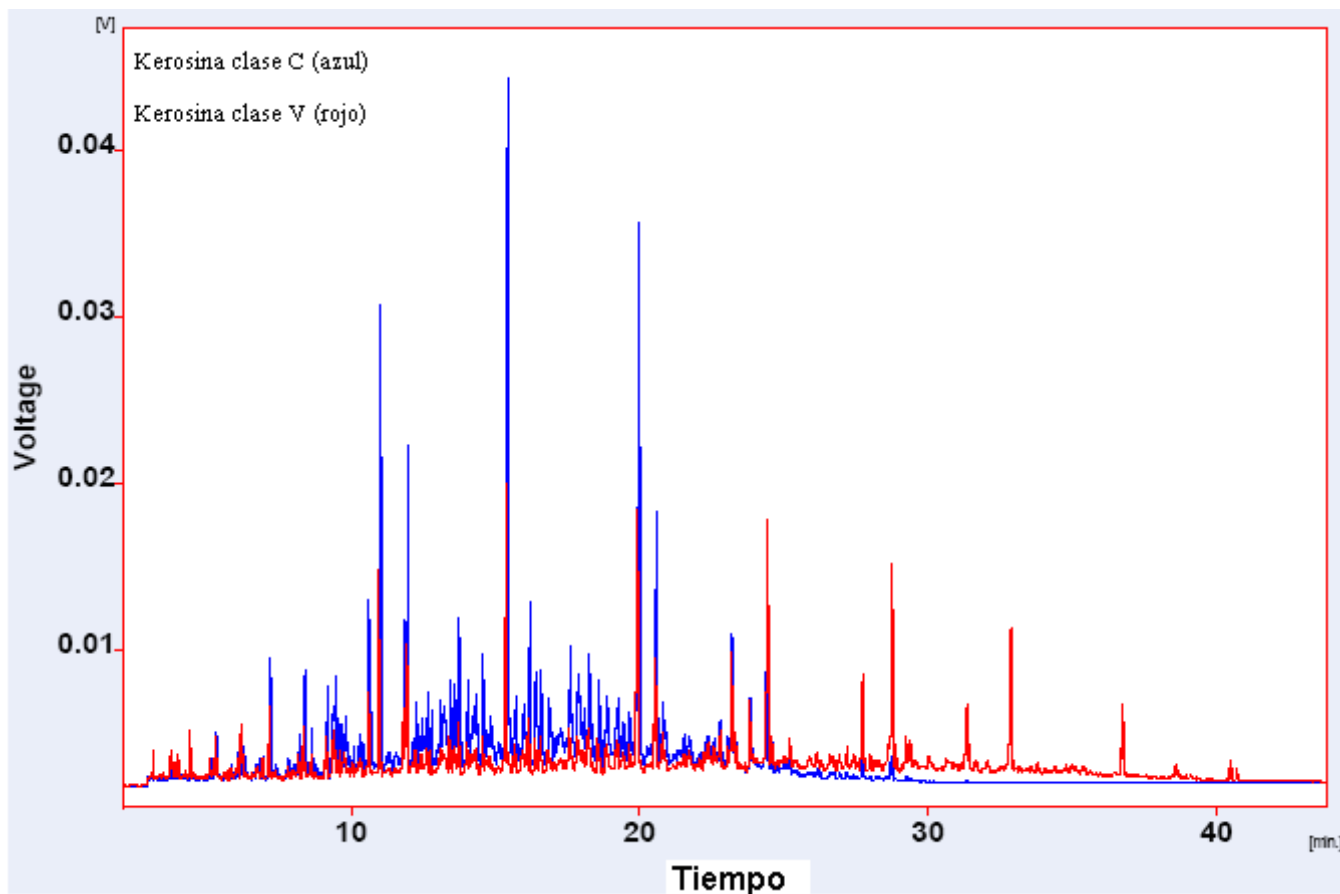


Figura 1S. Cromatogramas de iones totales de muestras de kerosinas de las clases C y V

*e-mail: adago@ceinpet.cupet.cu