DOI: 10.21577/0103-5053.20170073

http://dx.doi.org/10.21577/0103-5053.20170073

J. Braz. Chem. Soc., Vol. 28, No. 9, 1822-1829, 2017.
Printed in Brazil - ©2017 Sociedade Brasileira de Química
0103 - 5053 $6.00+0.00

Article

Prediction of Total Acid Number in Distillation Cuts of Crude Oil by ESI(–) FT-ICR MS Coupled with Chemometric Tools

Luciana A. Terra,[a] Paulo R. Filgueiras,[b] Rosana C. L. Pereira,[c] Alexandre O. Gomes,[c] Géssica A. Vasconcelos,[d] Lilian V. Tose,[b] Eustáquio V. R. Castro,[b] Boniek G. Vaz,[,d] Wanderson Romão[b,e] and Ronei J. Poppi[a]*

[a]*Instituto de Química, Universidade Estadual de Campinas, 13083-970 Campinas-SP, Brazil*

[b]*Laboratório de Petroleômica e Química Forense, Departamento de Química, Universidade Federal do Espírito Santo, 29075-910 Vitória-ES, Brazil*

[c]*CENPES/PETROBRAS, Av. Jequitiba 950, 21941-598 Rio de Janeiro-RJ, Brazil*

[d]*Instituto de Química, Universidade Federal de Goiás, 74001-970 Goiânia-GO, Brazil*

[e]*Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo, 29106-010 Vila Velha-ES, Brazil*

Competitive adaptive reweighted sampling-partial least squares (CARS-PLS) and negative-ion mode electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI(–) FT-ICR MS) data were adopted to assess the total acid number (TAN) of crude oil distillation cuts. Two crude oil samples and 24 derivatives with TAN ranging from 0.20 to 0.39 mg of KOH g^{-1} were investigated. The multivariate calibration PLS model was built with 18 calibration samples and tested with 8 validation samples. CARS-PLS reduced the number of variables from 1610 to only 4, allowing the identification of molecular formulas that are truly related to the TAN. The root mean square error of prediction (RMSEP) obtained was 0.01 mg of KOH g^{-1}, which is lower than the error when using all variables (0.03 mg of KOH g^{-1}). Finally, it was observed that the N and O_2 compound classes are the most important classes for providing a better correlation between ESI(–) FT-ICR mass spectra and TAN values.

Keywords: ESI(–) FT-ICR MS, TAN, distillation cuts, petroleomics, CARS-PLS

Introduction

The chemical characterization of petroleum using high-resolution mass spectrometry is referred to as petroleomics. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) provides a detailed view of petroleum composition, especially for polar compounds and heavy oil fractions. Owing to its ultra-high resolution and accuracy, FT-ICR MS routinely details the individual components (at the level of molecular formula assignment) of petroleum samples. The detailed elemental composition makes it possible to visualize the heteroatom classes (i.e., $C_cH_hN_nO_oS_s$ molecules with the same N, O, and S), degree of unsaturation from DBE (double-bond equivalent) pattern, and carbon number. Such detailed composition information reveals the differences among crude oil samples. These composition differences are correlated with the chemical and physical behavior of petroleum.[1-10]

Predicting the properties of crude oil and its derivatives based on molecular information has long been one of the main objectives of petroleomics. Owing to the extensive information provided by crude oil and its derivatives through high-resolution mass spectra data, the interpretation and correlation of molecular information with the physicochemical properties of petroleum is not an easy task. One strategy for finding a link between the complex and rich information obtained from the mass spectra and the properties of crude oil is to develop chemometric tools to access and correlate large amounts of data. Qian *et al.*[10] developed a univariate model to predict the total acid number (TAN) of crude oil samples based electrospray mass spectrometry (ESI MS) response. However, as the

*e-mail: boniek@ufg.br

acidic crude oil compounds ESI MS response showed a good correlation only for petroleum with high acidity, the model developed by Qian *et al.*[10] is only suitable for oil with TAN > 0.90 mg of KOH g^{-1}. In 2013, Vaz *et al.*[1] successfully employed chemometric tools such as partial least squares (PLS) and support vector machines (SVMs) to predict TAN of crude oil samples based on the normalized relative abundance of O_2 compounds detected via ESI(−) FT-ICR MS. Yeo *et al.*[11] used principal component analysis (PCA), hierarchical clustering analysis (HCA), and k-means clustering, to comparatively interpret the FT-ICR mass spectra data.

In 2014, Terra *et al.*[2] used negative-ion mode electrospray ionization, ESI(−), with FT-ICR MS coupled with PLS regression and variable selection methods to estimate the TAN of Brazilian crude oil samples. Lateefah *et al.*[12] described the acidic polar heteroatomic molecular class composition of three distillate fractions, showing different compositions of carboxylic acids and oxygen-sulfur species in distillates. Shi *et al.*[13] investigated the acid distribution in a type of Chinese crude oil and its fractions. Yang *et al.*[14] characterized the feed and products of high TAN crude oil which was subjected to five different temperatures between 300 and 500 °C, detecting low molecular weight acidic compounds which are responsible for refinery corrosion. Yingrong *et al.*[15] evaluated the petroleum cuts produced from two crude oils at three different boiling points, and they observed that the TAN decreased as the temperature increased. This behavior is due to the thermal decomposition of carboxylic acids, mainly molecules with a high carbon number. Dalmaschio *et al.*[3] characterized the polar compounds present in a true distillation point system, correlating their chemical composition (N, O, and O_2 classes) and DBE with the TAN and the corrosion process.

Competitive adaptive reweighted sampling-partial least squares (CARS-PLS)[16] is a method for variables selection in PLS multivariate regression that presents a high power to readjustment to new conditions, because the regression coefficients change as the variables receive new weights. The CARS algorithm selects subsets of variables using Monte Carlo sampling with fixed ratio variables (80-90% of the calibration set) to establish a calibration model. After that, based on the regression coefficients, two process are applied to variable selection: exponentially decreasing function (EDF) and adaptive reweighted sampling (ARS). EDF is used to remove small absolute regression coefficients, which is fast in first stage and in the second stage the decreasing is very slow performing a refined selection. Further, ARS eliminates variables in a competitive way, variables with higher weights (most important for the calibration model) are chosen more often instead of variables with small weights. Finally, cross validation method is applied to ensure that the variables with lowest root mean square error of cross validation (RMSECV) are the optimal subset with chemical meaning.

The CARS-PLS method has been used in spectroscopic data treatment, although the authors mention the possibility of its application in the omics sciences.[16] Here, in this paper, we use the rich and detailed information provided by distillation cuts ESI FT-ICR MS analysis with efficient multivariate methods, CARS-PLS model, to build an accurate and robust method for predicting the TAN of the low acidity distillation cuts. The main objective is to obtain a satisfactory set of molecules that explain the correlation between ESI(−) FT-ICR MS response and low-TAN distillation cuts using CARS-PLS in order to reduce the number of variables.

Experimental

Samples

Twenty-four distillation cuts obtained from two petroleum (petroleum 1 (P1) and petroleum 2 (P2)) offshore crude oil samples (Table 1) were used in this method. The TAN of each distillation cut was determined using the ASTM D664-09 standard method.[17]

Table 1. Crude oils and derivatives with their respective TAN values (mg of KOH g^{-1})

Fraction	Temperature / °C	P1	P2
Petroleum	25	0.23	0.29
7	200-225	0.20	0.24
8	225-250	0.21	0.25
9	250-275	0.22	0.23
10	275-300	0.22	0.27
11	300-325	0.25	0.28
12	325-350	0.30	0.39
13	350-375	0.32	0.34
15	400-425	0.22	0.21
16	425-450	0.21	0.22
17	450-475	0.24	0.24
18	475-500	0.29	0.28
19	500-525	0.26	0.27

TAN: total acid number.

ESI(−) FT-ICR MS

Approximately 5 mg of the sample were dissolved in 5 mL of toluene to obtain a solution of 1 mg mL^{-1} petroleum/derivate. Then, 0.5 mL of this solution was diluted in 0.5 mL of methanol containing 1% ammonium

hydroxide to run the analysis in negative mode. The resulting solution was injected into the 7 T LTQ FT Ultra system (Thermo Scientific, Bremen, Germany)[1,18,19] using a syringe pump (Harvard). The parameters used were as follows: ESI voltage, 3.10 kV; tube lens voltage, −100 V; capillary voltage, −39 V; and flow, 1 μL min^{-1}. The MS data were processed, and the elemental compositions of the compounds were determined by measuring the *m/z* values. For each elemental composition, $C_cH_hN_nO_oS_s$, the heteroatom class, the type, and the carbon number, CN, were tabulated to generate the class diagrams and the DBE *versus* intensity histogram.

For the DBE *versus* intensity histogram, DBE is the number of rings added to the number of double bonds in each molecular structure and can be deduced according to equation 1:[8]

$$DBE = c - h/2 + n/2 + 1 \quad (1)$$

where c, h, and n are the numbers of carbon, hydrogen, and nitrogen atoms, respectively, in the molecular formula.

Formula assignment

In general, in petroleum analyses via ESI FT-ICR MS, around 15,000 compounds are detected and identified. Molecular attribution for large ion sets is added using Composer® software (Sierra Analytics, Pasadena, CA, USA). The compound's elemental composition is determined by measuring the *m/z* values with a mass accuracy of < 1 ppm and resolving power of $m/Dm_{50\%} \cong 450{,}000$, where $Dm_{50\%}$ is the full peak width at half-maximum peak height for *m/z* = 400.

Chemometric analysis

After molecular formula attribution using Composer®, the data were arranged as a matrix using MATLAB® software. The dataset was mean centered and normalized between 0 and 1, by column, to build the PLS calibration model.[20-22] Further, k-fold cross validation by venetian blinds was adopted to optimize the choice of a suitable number of variables for the PLS model components. To determine the *m/z* values or molecular formulas related to the TAN, a variable selection method was used, e.g., CARS-PLS. This method considers the calibration set variables starting from a random sampling. The withdrawal of an important variable to explain the TAN increases the prediction error of the model. Therefore, the CARS-PLS model tries to identify these key variables of the model. However, some less important variables remain to be selected. The most important variables are usually selected by running the CARS-PLS algorithm several times and retaining the variables that are selected more frequently.[16] All models were developed using libPLS software version 1.95 under MATLAB® R2014a.[20]

Figures of merit

The PLS model evaluations were based on root mean square error of prediction (RMSEP), limit of detection (LOD), limit of quantification (LOQ) and the inverse of the analytical sensitivity (γ^{1}).[23] The RMSEP is calculated from equation 2:

$$RMSEP = \sqrt{\sum_{i=1}^{n-1} (\hat{y}_i - y_i)^2 / np} \quad (2)$$

where np is the number of validation samples, y_i and $\hat{y}_i$ are the measured and the predicted values of the concentration for each oil, respectively.

The LOD (equation 3) is the smaller amount of the analyte which could be differentiate in a confident way of zero or noise and LOQ (equation 4) is the smaller concentration in a sample that can determine with accuracy.

$$LOD = 3.3\delta x\|b\| \quad (3)$$

$$LOQ = 10\delta x\|b\| \quad (4)$$

where δx is instrumental noise and $\|b\|$ is the norm of the coefficient regression.

The equation 5 presents the analytical sensibility (γ) that is a ratio between the norm of the coefficient regressions ($\|b\|$) and instrumental noise (δx). This value indicates the instrument capability to differentiate among small differences in analytical concentration. The inverse of the analytical sensibility (γ^{-1}) is extensively used where values close to zero means high sensibility, desirable to analytical analyses when working with quantification.

$$\gamma = 1 / (\|b\|\delta x) \quad (5)$$

Regarding the decimal cases, we changed the results by placing 2 decimal cases, as follow in Table 2 and on the Results and Discussion section.

Results and Discussion

ESI(–) FT-ICR MS

In view of the large number of signals detected in ESI(–) FT-ICR mass spectra of both samples P1 and P2

(Figure S1, Supplementary Information), and of their respective distillation cuts (with boiling point varying from 200 to 525 °C, Figures S2 and S3, Supplementary Information), we will describe, for illustration purposes, only the ESI(−) mass spectra of three representative distillation cut samples for each crude oil. Figure 1 shows the ESI(−) FT-ICR mass spectra for distillation cuts (200-225 °C, 350-375 °C, and 500-525 °C) of crude oils P1 and P2. Note that when the distillation cut temperature increases (200→525 °C), the number of signals detected increases as does the average molecular weight distribution, M_w, that shifts to higher *m/z* values (M_w = 169→350 Da) (Figure 1). Polar species are detected as deprotonated species, $[M - H]^-$ ions, in which M corresponds primarily to pyrroles and carboxylic acids (e.g., naphtenic acids) and their analogs.

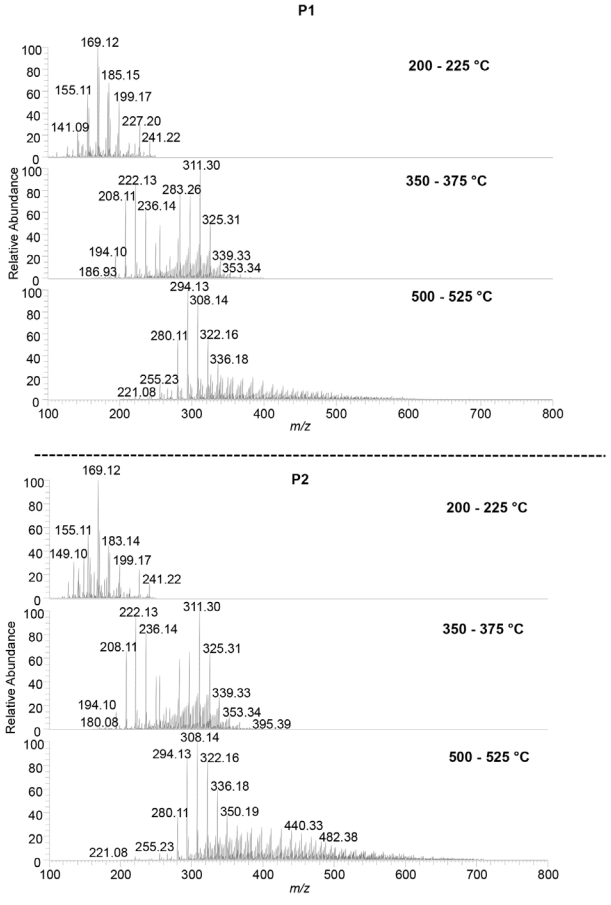

Figure 1. ESI(−) FT-ICR mass spectra of P1 and P2 of the distillate cuts.

It is possible to associate the total acid number as a function of the chemical profile by examining the ESI(−) FT-ICR mass spectra of each cut sample. The latter is illustrated in Figure 1; species with odd *m/z* numbers are carboxylic acids whereas molecules containing nitrogen species exhibit an even *m/z* number. The ESI(−) mass spectrum of the lightest derivative (fraction F7, 200-225 °C) mainly consisted of low-molecular-weight acid species because almost all the peaks detected are odd. For instance, signals with *m/z* 169.12340 correspond to the deprotonated form of 9-decenoic acid ($C_{10}H_{18}O_2$). As M_w shifts to higher values as a function of the distillation cut temperature, the proportion of molecules containing oxygen decreases, whereas the amount of nitrogen-containing compounds increases. The heaviest fraction (500-525 °C) contains predominantly non-basic nitrogen compounds because almost all the peaks have even *m/z* values (e.g., peak *m/z* 308.14473 corresponds to the deprotonated form of $C_{23}H_{19}N$). However, some highly aromatic acid species are detected in this fraction.

In general, different petroleum samples have significantly different chemical compositions. A way of displaying similarities or differences between the signal patterns of crude oil samples is to plot graphs, such as relative abundance *versus* different compound classes.[5,6,8] In the class profile diagrams, the total content of each compound group is presented in Figure 2; they were calculated by summing the abundance of each compound class and dividing by the total abundance of all species.

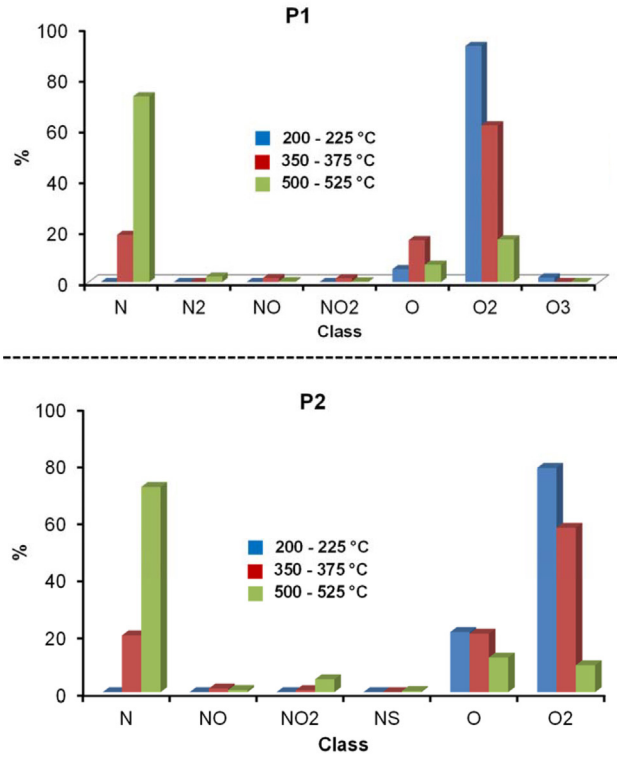

Figure 2. Class distribution diagram of three fractions obtained from P1 and P2 distillation.

Data obtained in petroleomic analyses is commonly organized as a histogram of heteroatom class ($N_nO_oS_s$) relative distribution. Figure 2 shows the class distribution

diagram for the three cuts of each sample. The histogram indicates that the majority of compounds belong to N and O_2 classes, which is in agreement with the profile obtained using ESI(–) mass spectra. Molecules detected in the lightest cut mostly have two oxygen atoms in their structure. As the temperature increases, the relative percentage of compounds containing oxygen gradually decreases, while the amount of nitrogen-containing species increases. The heaviest fraction, which is the most complex mixture, essentially consists of molecules containing one nitrogen in their structure. When analyzing the TAN values reported for crude oils and their respective distillation cuts, intermediary crude oils cuts (samples F11-F13 with boiling points of 300-375 °C) have higher TAN values (Table 1 and Figure 3).

DBE represents the number of rings plus double bonds of molecules that are within the same class.[5,6,8] DBE *versus* intensity histogram was plotted to confirm the TAN effect in a petroleum sample based on the chemical profile obtained via ESI(–) FT-ICR analyses. In order to simplify comparison, the abundance was scaled to the highest polar class species. Figure 4 shows the DBE distribution for N, O, and O_2 classes.

Overall, the DBE value varied from 0 to 18, in which the nitrogen species presented the highest values. This is explained by the fact that nitrogen compounds are mainly present in fractions with high *m/z* values. Figure 4 shows

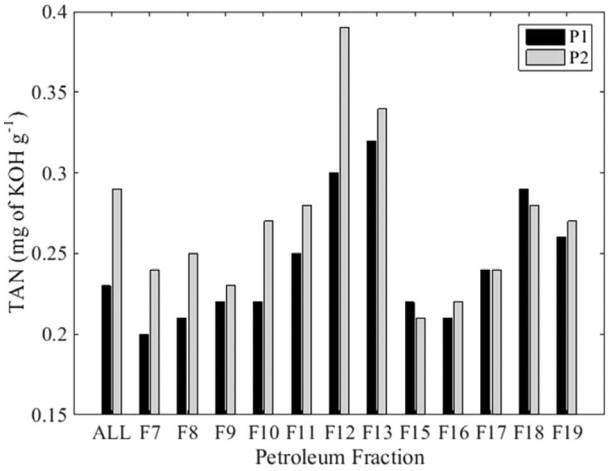

Figure 3. TAN values for crude oils P1 and P2 and their respective distillation cuts.

the DBE distribution of N class. The lightest cut barely has nitrogen compounds; therefore, it is not shown in this diagram. Most of the molecules in mid-fraction have a DBE of 9, which is represented by the carbazole species and their analogues. The heaviest distillation cut has a wider range of DBE, varying from 9 to 19. The most abundant species (DBE 15) is dibenzocarbazole and its analogues.

The DBE distribution of the O class is also shown in Figure 4. Most of the compounds have a DBE of 4, which corresponds to phenolic molecules. Its analogues can also be seen at DBE of 7 and 10. For the O_2 class, the majority

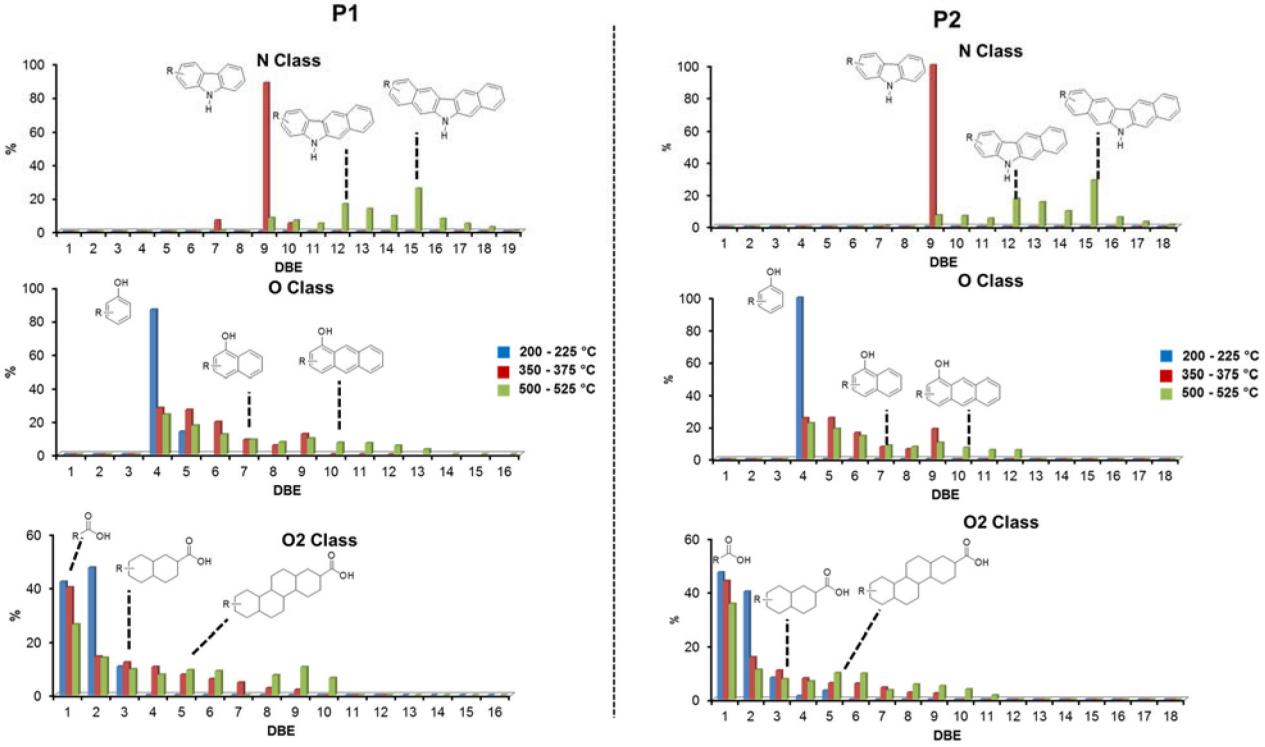

Figure 4. DBE *versus* relative intensity graph for (a) N class, (b) O class and (c) O_2 class.

of O_2 compounds have DBE 1 or 2, which are represented by aliphatic acids and acids with one ring. Species with DBE > 2 are analogues of naphthenic acids.

Prediction of total acid number

As reported by Qian *et al.*,[10] ESI signal is directly relative to the acid content in petroleum sample, that correlated to the KOH needed to neutralize the acid. However, good correlation between ESI MS signal with TAN values exists specifically for high-TAN crude oil samples (TAN > 0.9 mg of KOH g^{-1}). ESI MS response felled at approximately a constant value for low-TAN crudes. Here, we attempted to evaluate the CARS-PLS model to predict the TAN especially in low acidity distillation cuts samples and to identify the key variables (molecules) that explain to the model. Including the 24 cut fractions and the two crude oil samples, the TAN ranged from 0.20 to 0.39 mg of KOH g^{-1} (Table 1). Specifically, the petroleum samples P1 and P2 presented TAN values of 0.23 and 0.29 mg of KOH g^{-1}, respectively.

The database with the mass spectra contains 1610 *m/z* values, i.e., variables. Few of these variables are in fact related to the TAN of the samples. The PLS model was built using 18 calibration samples and 8 validation samples being selected by Kennard-Stone. The CARS-PLS algorithm was executed 200 times to obtain the frequency of choice of each variable. The RMSEP values of the different frequencies of selection are shown in Figure 5; it becomes possible to compare them with a selection frequency ranging from variable selected (VS) at least 1% (VS > 1) to 90% (VS > 90).

The RMSEP obtained by using all the variables for TAN determination is just above 0.03 mg of KOH g^{-1}. By using CARS for variable selection, it is observed that some variables have greater frequencies of selection, whereas others are selected in a few models only. Table 2 lists the number of variables selected by CARS and the respective calibration, cross-validation and prediction errors.

Table 2. Number of variables selected by CARS and the respective RMSEC, RMSECV and RMSEP values

	Number of variables	RMSEC / (mg of KOH g^{-1})	RMSECV / (mg of KOH g^{-1})	RMSEP / (mg of KOH g^{-1})
FULL	1610	0.04	0.04	0.03
VS0	69	0.02	0.03	0.02
VS1	22	0.02	0.03	0.02
VS10	11	0.02	0.02	0.02
VS20	10	0.02	0.02	0.02
VS30	7	0.02	0.02	0.02
VS40	4	0.02	0.02	0.01
VS50	4	0.02	0.02	0.01
VS60	4	0.02	0.02	0.01
VS70	4	0.02	0.02	0.01
VS80	4	0.02	0.02	0.01
VS90	3	0.02	0.02	0.01

CARS: competitive adaptive reweighted sampling; RMSEC: root mean square error of calibration; RMSECV: root mean square error of cross validation; RMSEP: root mean square error of prediction.

Therefore, when a PLS model was built only with variables that have at least 40% selection rate (VS ≥ 40), a minimum RMSEP value was reached, indicating that this set of variables best reflects the variation in the TAN content. These variables are presented in Table 3.

The results obtained for the predicted error from CARS-PLS, using 2 latent variables, was a RMSEP of 0.01 mg of KOH g^{-1} that explained 91.72% in X and 85.67% in Y, with reduction of the original data from 1610 variables

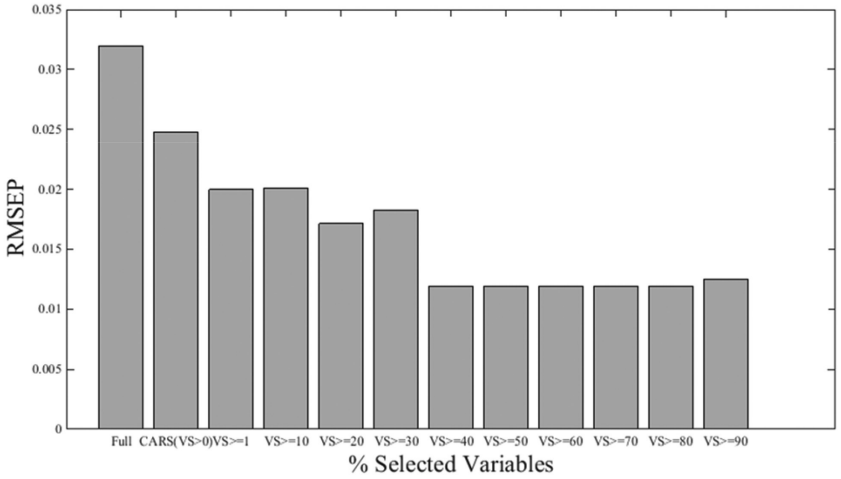

Figure 5. Histograms of RMSEP to the frequency of selected variable (VS) by CARS.

Table 3. Variables selected by CARS for the TAN using VS ≥ 1 and the VS ≥ 40 in bold

Formula	Exact mass	DBE
$C_{18}H_{36}O_2$	283.264254	1.0
$C_{19}H_{38}O_2$	297.279904	1.0
$C_{20}H_{40}O_2$	311.295554	1.0
$C_{22}H_{17}N$	294.128823	15.0
$C_{10}H_{18}O_2$	169.123415	2.0
$C_{10}H_{20}O_2$	171.139136	1.0
$C_{11}H_{16}O$	163.112856	4.0
$C_{11}H_{20}O_2$	183.139048	2.0
$C_{11}H_{22}O_2$	185.154702	1.0
$C_{12}H_{18}O$	177.128495	4.0
$C_{12}H_{24}O_2$	199.170354	1.0
$C_{13}H_{20}O$	191.144133	4.0
$C_{13}H_{26}O_2$	213.186145	1.0
$C_{14}H_{28}O_2$	227.201654	1.0
$C_{16}H_{17}N$	295.264436	9.0
$C_{17}H_{34}O_2$	269.248611	1.0
$C_{18}H_{15}N$	244.113157	12.0
$C_{18}H_{34}O_2$	281.248617	2.0
$C_{19}H_{36}O_2$	295.264436	2.0
$C_{21}H_{15}N$	280.113172	15.0
$C_{21}H_{42}O_2$	325.311405	1.0
$C_{23}H_{19}N$	308.144542	15.0

CARS: competitive adaptive reweighted sampling; TAN: total acid number; VS: variable selected; DBE: double-bond equivalent.

to 4 variables. That result was compared with the results from PLS model with all variables using the accuracy test.[24] In this test, *p*-value of 0.08 was obtained being higher than the significance level of 0.05 and, therefore, it is possible to affirm, with 95% of confidence, that CARS-PLS presented the same accuracy than the model using all the variables. However, now the model presents the advantage of the reduced number of variables making possible to identify molecular formulas that are truly related to the TAN. CARS-PLS produces a simple, accurate, and robust model, as indicated by the regression curve for predicting the TAN (Figure 6) with relative prediction errors ranging from 5 to 17%.

It is important to mention that in this work, we obtained the prediction error using all the variables belonging to only the N and O_2 classes separately (632 and 306 variables, respectively), and the RMSEP obtained was 0.03 and 0.02 mg of KOH g^{-1}, respectively. These errors are much larger than those obtained with the CARS-PLS method. These results are another strong indicator of the ability of CARS-PLS to not only choose the compound class, but also assign the most important molecules into each class.

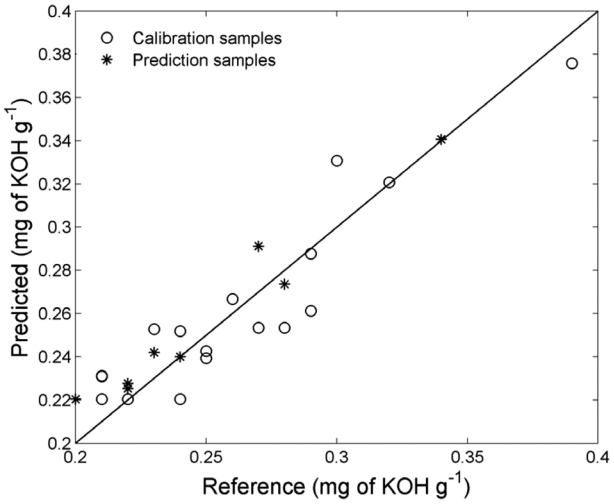

Figure 6. Regression curve using CARS-PLS model for TAN using 4 variables selected.

In Figure 6, it is possible to observe excellent concordance between the TAN estimated by CARS-PLS and the values measured by the reference methods for both the calibration and the prediction samples. This confirms that the model was properly developed, does not present overfitting, and can predict further samples with small errors. The figures-of-merit[23] obtained are as follows: determination coefficient (R^2), 0.95; limit of detection (LOD), 0.03 mg of KOH g^{-1}; limit of quantification (LOQ), 0.09 mg of KOH g^{-1}; the inverse of the analytical sensitivity (γ^{-1}), 0.01 mg of KOH g^{-1}. The TAN estimated values for the eight prediction samples are within the repeatability limit of the results obtained by the standard method ASTM D664-09 (Figure 7). The repeatability (r) was determined by the equation 6:

$$r = 0.044(X_{TAN} + 1) \quad (6)$$

where X_{TAN} is the average of the two test results.

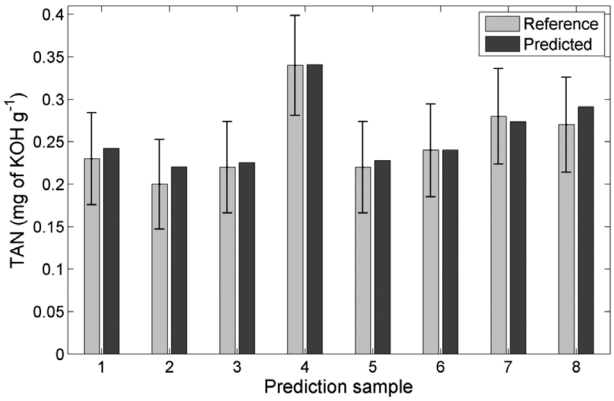

Figure 7. TAN values measured by standard method ASTM D664 and predicted by multivariate model for the eight prediction samples. Vertical bars represent the confidence intervals.

Conclusions

This work presented a chemically significant and effective application of the CARS-PLS strategy for reducing the number of variables of FT-ICR MS data for two crude oil samples and their respective derivatives over a wide range of distillation temperatures. The small number of variables, i.e., the reduction from 1610 to only 4 variables belonging to N and O_2 classes, facilitates the interpretation of variables that are truly important to predict and ensures better correlation between the ESI(−) FT-ICR mass spectra and low TAN values, presenting an adequate RMSEP of 0.01 mg of KOH g^{-1}.

Supplementary Information

Supplementary information is available free of charge at http://jbcs.sbq.org.br.

Acknowledgments

This research was generously funded by PETROBRAS/CENPES, FAPESP, CNPq, CAPES, and FINEP.

References

1. Vaz, B. G.; Abdelnur, P. V.; Rocha, W. F. C.; Gomes, A. O.; Pereira, R. C. L.; *Energy Fuel* **2013**, *27*, 1873.
2. Terra, L. A.; Filgueiras, P. R.; Tose, L. V.; Romão, W.; Souza, D. D.; Castro, E. V. R.; Oliveira, M. S. L.; Dias, J. C. M.; Poppi, R. J.; *Analyst* **2014**, *139*, 4908.
3. Dalmaschio, G. P.; Malacarne, M. M.; Almeida, V. M. D. L.; Pereira, T. M. C.; Gomes, A. O.; Castro, E. V. R.; Greco, S. J.; Vaz, B. G.; Romão, W.; *Fuel* **2014**, *115*, 190.
4. Hugheya, C. A.; Rodgers, R. P.; Marshall, A. G.; Qianc, K.; Robbins, W. K.; *Org. Geochem.* **2002**, *33*, 743.
5. Mapolelo, M. M.; Rodgers, R. P.; Blakney, G. T.; Yenc, A. T.; Asomaningc, S.; Marshall, A. G.; *Int. J. Mass Spectrom.* **2011**, *300*, 149.
6. Headley, J. V.; Peru, K. M.; Barrow, M. P.; *Mass Spectrom. Rev.* **2009**, *28*, 121.
7. Dias, H. P.; Dixini, P. V.; Almeida, L. C. P.; Vanini, G.; Castro, E. V. R.; Aquije, G. M. F.; Gomes, A. O.; Moura, R. R.; Junior, V. L.; Vaz, B. G.; Romão, W.; *Fuel* **2015**, *139*, 328.
8. McLafferty, F. W.; Turecek, F.; *Interpretation of Mass Spectra*; University Science Books: Sausalito, CA, 1993.
9. Zhigang, W.; Rodgers, R. P.; Marshall, A. G.; *Energy Fuels* **2004**, *18*, 1424.
10. Qian, K.; Edwards, K. K. E.; Dechert, G. J.; Jaffe, S. B.; Green, L. A.; Olmstead, W. N.; *Anal. Chem.* **2008**, *80*, 849.
11. Yeo, I.; Lee, J. W.; Kim, S.; *Bull. Korean Chem. Soc.* **2010**, *31*, 3151.
12. Lateefah, A.; Kim, S.; Rodgers, R. P.; Marshall, A. G.; *Energy Fuel* **2006**, *20*, 1664.
13. Shi, Q.; Zhao, S.; Xu, Z.; Chung, K. H.; Zhang, Y.; Xu, C.; *Energy Fuel* **2010**, *24*, 4005.
14. Yang, B.; Xu, C.; Zhao, S.; Hsu, C. S.; Chung, K. H.; Shi, Q.; *Sci. China: Chem.* **2013**, *56*, 848.
15. Yingrong, L.; Qundan, Z.; Wei, W.; Zelong, L.; Xinyi, Z.; Songbai, T.; *China Pet. Process. Petrochem. Technol.* **2014**, *16*, 8.
16. Li, H.; Liang, Y.; Xu, Q.; Cao, D.; *Anal. Chem. Acta* **2009**, *648*, 77.
17. ASTM D664-09, *Standard Test Method for Acid Number of Petroleum Products by Potentiometric Titration*, ASTM International, West Conshohocken, PA, 2006.
18. Benassi, M.; Berisha, A.; Romão, W.; Babayev, E.; Rompp, A.; Spengler, B.; *Rapid Commun. Mass Spectrom.* **2013**, *27*, 825.
19. Haddad, R.; Regiani, T.; Klitzke, C. F.; Sanvido, G. B.; Corilo, Y. E.; Augusti, D. V.; Pasa, V. M. D.; Pereira, R. C. C.; Romão, W.; Vaz, B. G.; Augusti, R.; Eberlin, M. N.; *Energy Fuels* **2012**, *26*, 3542.
20. Hoskuldsson, A.; *J. Chemom.* **1998**, *2*, 2112.
21. Jong, S.; *Chemom. Intell. Lab. Syst.* **1993**, *18*, 251.
22. Anderson, M.; *J. Chemom.* **2009**, *23*, 518.
23. Valderrama, P.; Braga, J. W. B.; Poppi, R. J.; *Quim. Nova* **2009**, *32*, 1278.
24. Van der Voet, H.; *Chemom. Intell. Lab. Syst.* **1994**, *25*, 313.

Submitted: January 3, 2017
Published online: April 20, 2017