

Determination of octane numbers in gasoline by distillation curves and partial least squares regression

Gisele Mendes, Helga G. Aleme, Paulo J.S. Barbeira *

Laboratório de Ensaios de Combustíveis, Departamento de Química, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Av. Presidente Antônio Carlos 6627, CEP: 31270-901 Belo Horizonte, Minas Gerais, Brazil

ARTICLE INFO

Article history:

Received 10 December 2010

Received in revised form 20 January 2012

Accepted 28 January 2012

Available online 20 February 2012

Keywords:

Octane numbers

Gasoline

Distillation

Partial least squares

ABSTRACT

In this work, distillation curves (ASTM D86) associated with multivariate calibration (PLS-Partial Least Squares) were used to predict MON (motor octane number) and RON (research octane number) parameters of automotive fuel *RMSEC* getting values (Root Mean Square Error of Calibration) and *RMSEP* (Root Mean Square Error of Prediction) smaller than the literature. *RMSEC* values obtained were 0.051 and 0.078, and *RMSEP* values were 0.063 and 0.085 for MON and RON, respectively. These low values along with high accuracy, when compared to standard test methods (ASTM D2700 and ASTM D2999), indicated that the PLS models were efficient to predict MON and RON values and can be used as an alternative to control quality for automotive gasoline.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Gasoline is a complex mixture of hundreds of volatile and combustible compounds derived from petroleum, with 4–12 carbon atoms and boiling points in the range of 30–220 °C. In Brazil, gasoline is marketed in gas stations with $25 \pm 1\%$ (v/v) ethanol [1]. According to the National Agency of Petroleum, Natural Gas and Biofuels (ANP), approximately 35 billion L of gasoline were marketed in 2011 [2] and the quality of this fuel is guaranteed by several tests and specifications established by this agency [2]. Gasoline's octane number below the level stipulated by legislation has been one of the most frequent causes of non-conformity, according to ANP's Fuel Quality Monitoring Program (PMQC) [2], and reached 15.3% of the samples analyzed in 2011.

Octane number is one of the main parameters used in quality control of gasoline and provides information about the resistance to auto ignition. This phenomenon occurs when the temperature of the fuel–air mixture under the effect of compression, leading to sufficiently increased self-detonation of the mixture without the help of a spark [3,4]. Fuel self-detonation generates pressure pulses in the engine cylinder and causes an increase in fuel consumption, loss of engine power and, at worst, may even damage the engine [4]. Octane number is mainly affected by the presence of aromatic hydrocarbons, isoparaffins, olefins and additives such as ethanol and esters [5]. The chemical structure of hydrocarbons in gasoline has great influence on detonation. Paraffins that have

many ramifications, olefins and aromatics (benzene, toluene, xylenes) are very resistant to self-detonation, and conversely the long chains paraffins with less ramifications and olefins, with more than four carbon atoms, are more susceptible to the phenomena of self detonation [6].

Different methods are used to determine the octane number of gasoline by simulating engine specific requirements. The most important methods are: Motor Octane Number (MON) [7], Research Octane Number (RON) [8] and Antiknock Index (AKI) [9]. For Brazilian automotive fuel, ANP determines a minimum MON value of 82.0 [2]. The RON parameter is not determined in Brazilian specifications for automotive fuel intended for dealers, and it is only included in standard fuel specifications used for the approval of engines [10]. The arithmetic average of RON and MON values defines the antiknock index (AKI), according to ASTM D4814 [9], and the established minimum value is 87.0 [2].

Octane number is measured in special engines (CFR engines – Cooperative Fuel Research), single cylinder engines with variable compression ratio, equipped with the required tools and set up on a fixed base [11]. TSF's blends (toluene standardization fuel) with several MON and RON known values for the calibration system [2,8] were used in both tests. To determine the MON value, according to ASTM D2700 [7], the engine operates in controlled conditions: rotation 900 ± 9 rpm, lube oil temperature at 58 ± 8 °C and cooling liquid at 100 ± 2 °C. To determine the RON value, according to ASTM D2699 [8], the controlled conditions are: 600 ± 6 rpm, 58 ± 8 °C and 100 ± 2 °C, respectively [10].

Despite the wide use of engine tests, there is a general consensus toward some inherent problems in these tests: the standard

* Corresponding author. Tel./fax: +55 31 3409 5767.

E-mail address: barbeira@ufmg.br (P.J.S. Barbeira).

engines used are extremely expensive, large amount of iso-octane and *n*-heptane standards as well as of samples are employed (approximately 500 mL per test), the engines are noisy, produce exhaust gases and require regular cleaning and maintenance, analyses are slow (approximately 30 min per sample) and automation is impracticable [10]. Due to these reasons, several alternative methods such as dielectric spectroscopy and FT NIR spectra associated with chemometric techniques have been tested to determine MON and RON [12–14]. Fuel automatic analyzers, based on mid-infrared spectroscopy associated with multivariate calibration are also an available commercial alternative [15].

Multivariate calibration associated with infrared analysis may be used to determine physical or chemical characteristics of different materials. The spectra obtained for a set of reference calibration samples and their concentration values, or of the characteristic of interest, are correlated in order to obtain a multivariate calibration model. The result of this model is used to analyze spectra of unknown samples in order to provide an estimate of the component concentration or value of the characteristic in the unknown sample. Multiple Linear Regression (MLR), principal components regression (PCR), and PLS are some examples of multivariate mathematic techniques commonly employed to develop a calibration model. In addition, statistical tests are used to detect outliers during the development of a calibration model [16].

There are several methods described in the literature that apply multivariate calibration associated with different techniques, such as infrared spectroscopy and gas chromatography, to predict various properties of automotive fuels, as specific gravity [17,18], distillation fractions [17], octane number [19], aromatics [20], MON and RON [21,22].

Flumignan et al. [22] evaluated several physico-chemical parameters of gasoline, amongst these MON, RON and AKI, using the chromatographic profile obtained by a flame ionization detector (GC–FID) associated with PLS. In determining MON and RON the standard error calibration values (*SEC*) were 0.6 and 0.8 and standard error validation values (*SEV*) were 0.7 and 1.0, respectively.

Kelly et al. [23] determined MON, RON and AKI associating PLS multivariate calibration with near infrared spectrometry (NIR) in the interval 660–1215 nm. The *RMSEC* values (Root Mean Square Error of Calibration) 0.383, 0.355 were obtained for RON and MON, respectively.

Cooper et al. [21] used PLS regression analysis to build regression models which correlate the Raman of spectra of fuels with the experimentally determined values for MON, RON, and AKI. The *SEV* values were 0.415, 0.535, and 0.410 for MON, RON and AKI, respectively.

Oliveira et al. [18] carried out a comparative study of two methods to choose the spectral region, combining FTNIR absorption measures with PLS to determine MON as well as other parameters. The method is based on spectral distribution of standard deviation of concentration (*sc/c*) that provided the lowest *RMSEP* value of 0.33, leading to more reliable calibration models.

With the purpose of simplifying the analytical process of gasoline by reducing the number of tests to evaluate quality, as well as costs, this work describes the use of distillation curves, obtained according to ASTM D86 [24], and PLS multivariate calibration in the determination of MON and RON parameters, comparing results with those obtained by an automatic analyzer based on mid-infrared spectroscopy. The use of manual distillations and later identification of the components in each distillation fraction, resorting to infrared spectroscopy and gas chromatography coupled to mass spectrometry, aims at explaining the relationship between hydrocarbons and main fractions in the determination of MON and RON.

The distillation test was carried out according to ASTM D86 [24], which describes distillation at atmospheric pressure of petro-

leum products. The aim of this test is to determine volatility features by checking if the light and heavy proportions of fuel produced are appropriate, and seeking good performance in combustion and detection of contamination with other products. For Brazilian automotive fuel, ANP establishes maximum temperature values for 10%, 50% and 90% of recuperated volume, as well as final boiling point and waste volume [2].

2. Experimental

2.1. Samples

The samples (300 regular gasolines) were collected in fuel stations of the state of Minas Gerais (Brazil), and were produced by five different refineries, with an alcoholic content of 19–33% (v/v) (after the refinery produces the gasoline, ethanol is added by the distributors, which pass the mixture to the fuel stations). From these samples, 150 were used to predict MON and 150, to predict RON. The choice of these samples was made randomly from the larger set one, and subsequently submitted to the PLS regression. The 150 samples used in each model were considered a representative number, enough to acquire 1/3 of the samples to the validation set [16]. The gasoline samples were stored in appropriate polyethylene bottles, sealed and refrigerated (8–15 °C) until physico-chemical analyses were performed, to prevent the loss of volatile components [24].

2.2. Materials and equipment

The gasoline samples were distilled in Herzog HDA 627 automatic distillers, according to ASTM D86 [24].

MON and RON values, paraffin, olefin, and aromatic were obtained using a Petrospec GS1000 automatic analyzer based on mid-infrared spectroscopy associated with multivariate calibration methods as PLS, PCR (Principal Components Regression) and MLR (Multiple Linear Regression), according to ASTM E1655 [16]. The equipment's database is composed of samples whose octane number values were obtained using engine tests [15] and periodically updated with the introduction of new samples. An identical procedure was carried out with the concentration values of different hydrocarbons obtained using chromatography. The accuracy of the values obtained with a commercial spectrometer is usually evaluated by ANP's interlaboratory testing program, which has more than twenty participants and holds three annual rounds [2].

2.3. Experimental procedures

2.3.1. Automatic distillation

For the distillation test, 100 mL of previously cooled gasoline were transferred to a specific distillation flask coupled with a sensor and heated in order to maintain the distillation rate between 4 and 5 mL min^{−1}, according to ASTM-D86 [24]. The distilled steam was condensed and collected in a cooled beaker and the distillation curves (distillation temperature depending on recuperated volume), at 1% (v/v) intervals, were obtained after correcting atmospheric pressure temperature readings to 760 mm Hg and considering volume loss, according to ASTM-D86 [24].

2.3.2. Manual distillation

A manual distillation system adapted to the specifications established by ASTM-D86 [24] was used for the analysis of distillation fractions of automotive fuel. The samples of gasoline were prepared by adding 25% (v/v) ethanol to gasoline originally from REGAP refinery. This was done due to the fact that most of the set samples were originally from this refinery. The specifications

used for the manual method were the same as those used for the automatic method; distilled fractions were collected every 10 mL. Five distillations of this gasoline were made, and each distillation fraction was mixed in order to increase the representativeness of the results. The solutions of the different percentages of recovered volume were cooled (from 8 to 15 °C) and later submitted analysis in a commercial infrared spectrometer, then to gas chromatography associated with mass spectrometry to determine the composition of the distilled fractions.

2.3.3. Gas chromatography–mass spectrometry (GC–MS)

The composition of the distilled fractions of gasoline samples was obtained using an Electron-Ionization Mass Spectrometry (EI-MS). The analyses were performed in a GC–MS Shimadzu, model GC-17A/QP-5050A using a fused capillary column (50 m × 0.2 mm × 0.5 µm, PONA50, HP), with poly(methylsiloxane) as the stationary phase and helium as the carrier gas at a constant flow rate of 0.1 mL min⁻¹. Sample aliquots of 1.0 µL were injected in split mode (1:16) without solvent delay. The analyses were carried out in the following conditions: initial temperature 34 °C for 8 min at 2 °C per minute until reaching 60 °C, 3 °C per minute up to 185 °C and 10 °C up to 250 °C for 2 min. Injector and detector temperatures were 230 and 250 °C, respectively. The mass spectrometer worked on ionization mode 70 eV operating in scan mode (*m/z* 45–350). The presence of different types of compounds in the samples was discovered using Total Ion Chromatogram (TIC) together with library information (Wiley Class 5000, 6th edition). Compounds with less than 90% similarity of mass spectra were discarded.

2.3.4. Octane number (MON and RON)

For the determination of MON and RON values, a small volume of gasoline (approximately 20 mL) was transferred to a clean, dry flask coupled to the equipment that pumps the sample into the reading cell and was submitted to infrared radiation.

To estimate MON and RON values with the infrared spectrometer, a calibration model was constructed using a series of samples from different origins, correlating the absorption spectra of the samples and the MON and RON values obtained applying the engine method according to ASTM D2700 [7] and ASTM D2699 [8], and using Multiple Linear Regression (MLR) according to ASTM E1655 [16].

2.3.5. Construction of models

The calibration models were developed using PLS algorithm and estimates were carried out in Minitab Release (version 14 for Windows) and SOLO (version 2007–2008 for Windows) softwares. For each model, the distillation curves of the calibration set used to build the PLS model were correlated with their MON and RON values separately.

Autoscaling was the pre-process used in order to assign the same importance to all the variables. The calibration model was developed using cross-validation “leave-one-out”, whereby the best number of factors (latent variables) in the calibration model is determined by the *PRESS* value (Prediction Error Sum of Squares).

MON and RON values of the prediction set samples were always obtained with the best number of factors [25,26], as this is a critical parameter in the calibration model. Hence, a strict test for the predictive significance of each PLS component is necessary, and this test is stopped when components start to be non-significant [27]. The number of latent variables with lowest *PRESS* values was selected for the construction of models [28].

Through PLS the original matrix data *X* (distillation curves) and *Y* (MON or RON values) were separated into scores and weights vectors, plus a residue matrix of non-modeled data. The product

of these vectors originate the designated latent variables which are capable of representing the samples at less dimensions and describe the direction of maximum variance. The scores are the coordinates of the samples to the new axes system. Its graph provides the possibility to analyze through similarities, grouping and outliers. The weights correspond to the variables present in the data set and its graph provides the possibility to identify the important variables. The higher the weight the more important the variable is [29–31].

2.3.6. Evaluation of accuracy

Different procedures were used to measure accuracy. The first one was calculating *RMSEP* [28] and the estimated values were obtained from two sets of external validation. Each of these sets was built using 50 samples with MON and RON values varying from 81.6 to 83.2 and 97.4 to 101.4, respectively, and were independent from the samples used for the calibration set. In addition, the predicted values for the validation sets samples were compared using *t* test, with the values obtained following the reference methods (ASTM D2700 and D2699). Another procedure that used was calculating *RPD* (Residual Prediction Deviation), which is defined as the ratio between standard deviation and prediction error. A good calibration model must have an *RPD* value higher than three, according to Zhang et al. [32]. Test *t* was also used to compare the MON and RON results of seven samples from ANP Interlaboratory Programs [2], which includes the participation of the 23 laboratories of the PMQC with the proposed method. Therefore, the distillation curves of these samples were applied to the calibration set of the proposed method and the predicted values were used in test *t*.

2.3.7. Evaluation of repeatability and reproducibility

The repeatability and reproducibility evaluation of the methods was performed in accordance with ISO 5725-2 [33]. Ten gasoline samples were used for this evaluation and physico-chemical assays, with seven replicates for each sample were performed by three different analysts, producing 21 results per sample [33].

3. Results and discussion

A preliminary study of the distillation curve showed the need to use exclusively the 4–93% (v/v) interval for PLS models, due to the low reproducibility of the results outside this interval. ASTM-D86 [24] determines that the distillation rate should occur between 4 and 5 mL min⁻¹ and, therefore, initial adjustments of the heating resistance were required. Initially, there is a vigorous boiling in the distillation curve that causes fluctuations in temperature values from the initial point up to 4% (v/v) and low reproducibility of the measurements. However, pyrolysis of larger molecules occurs in the final stage of the distillation from 94% to 98% (v/v), thus reducing boiling temperature [24] and therefore the samples finish distillation at different percentages.

3.1. Prediction of MON

The calibration matrix was built with 100 samples, and the validation matrix with 50 samples. The subdivision of these sets was done randomly and carried out before using PLS. Being a natural calibration, the samples of this set contained MON values in the 81.6–83.2 range. Although limited, this range reflects the variability of the gasolines commercialized in the sampling area, during a period of at least one year.

The number of latent variables was defined by cross validation leave-one-out, because it provided the lowest prediction error for the cross-validation set (*PRESS*). Eight latent variables were esteemed the appropriate number, considering there is no significant

difference at 95% in PRESS values using 8–10 latent variables, according to test *F*.

Three first latent variables presented 90% of the total variance, 38.0% for the first variable, 47.0% for the second and 5.0% for the third variable. Unlike chemometric techniques like PCA (Principal Components Analysis), where principal components are built in decreasing order of the number of variances they describe [34] in PLS the first latent variable does not necessarily have more explained variance than the second, etc. One possible reason for this to happen may be a perfect linear between data projections [34], which in this case correspond to distillation curves and MON values.

Fig. 1A indicates that in the first latent variable the most important fractions for the determination of MON include the 60–93% (v/v) interval and, for the second latent variable, fractions from 40% to 60% (v/v) due to the existence of hydrocarbons with high MON values.

Fig. 2 shows MON values obtained by commercial spectrometer for the different fractions of gasoline containing 25% (v/v) ethanol. MON values reduce significantly in the 10–40% (v/v) fractions, thus coinciding with the interval of higher negative weights of the first latent variable (Fig. 1A). In the following interval, from 40% to 80% (v/v), MON values remain almost unchanged and match with the highest positive weights in the first and second latent variables. Then, there is an increase of the MON value between 80% and 90% (v/v) that matches with the heaviest positive weights in the first latent variable.

Fig. 3 shows the composition of distilled fractions obtained by the commercial spectrometer. In the 10–40% (v/v) interval the

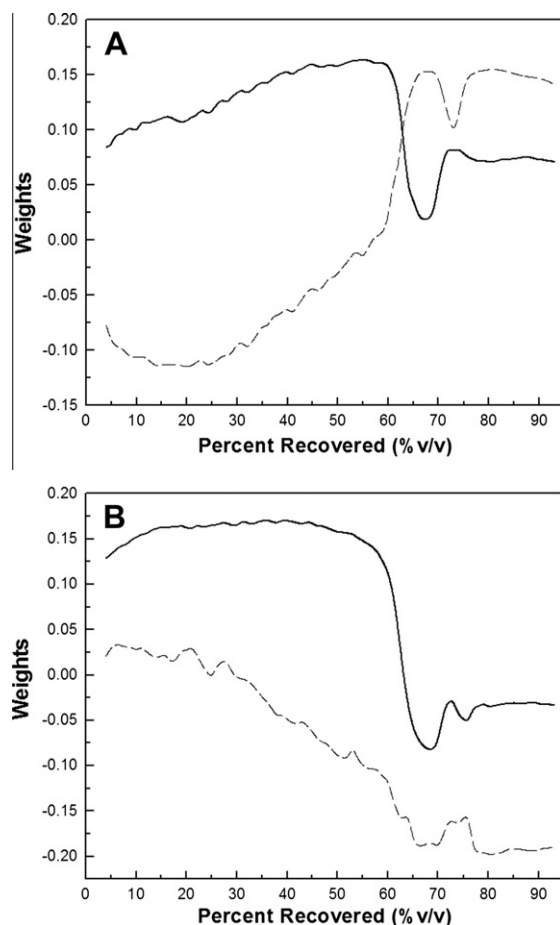


Fig. 1. Weight graphs obtained in the determination of octane numbers. (A) MON: LV1 (---) with 38.0% and LV2 (—) with 47.0% of explained variance. (B) RON: LV1 (—) with 51.0% and LV3 (---) with 28.2% of explained variance.

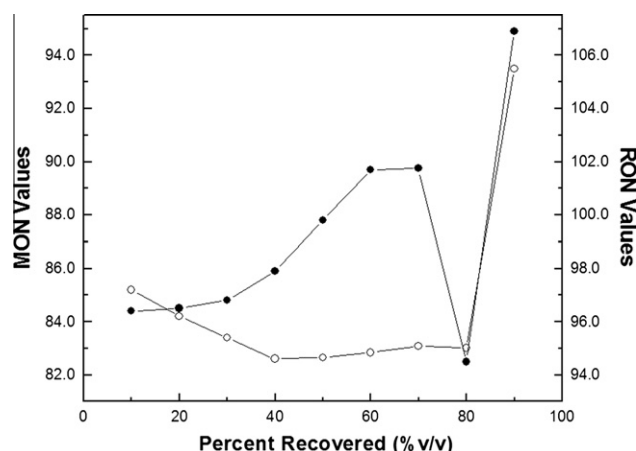


Fig. 2. Octane numbers for different distillation fractions for automotive gasoline containing 25% (v/v) ethanol. (○) MON; (●) RON.

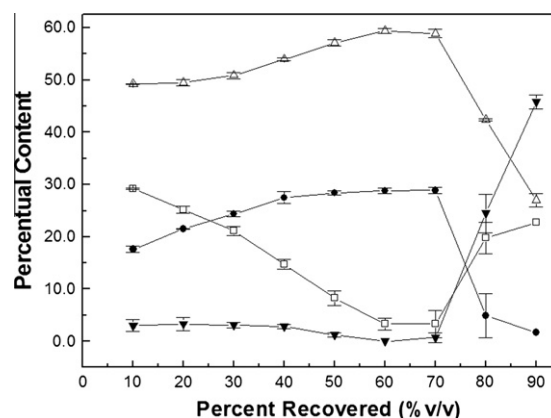


Fig. 3. Percentage of different hydrocarbons in distilled fractions of automotive gasoline with 25% (v/v) ethanol. (□) Olefins, (▼) aromatics, (●) ethanol and (Δ) paraffins.

content of paraffins (with low MON values) is higher than olefins and ethanol (with high MON values) [6,35–37], thus indicating that paraffins have more influence on MON values in this interval. The results obtained by GC/MS showed that even fraction 40% (v/v) becomes more concentrated in linear paraffins (as hexane) and branched paraffins (as 3-methylpentane and 3,3,4-trimethylpentane) whose MON values are very low [35,36]. This justifies the negative weights of these variables.

Fig. 3 indicates that in the interval from 40% to 60% (v/v) the content of paraffin is still higher than other components, whereas ethanol remains steady. Up to the 60% (v/v) fraction there are significant amounts of hexane (even in small proportions) and, at the same time, there is an increase of cyclic paraffins with five to six atoms of carbon (as methylhexane and methylcyclopentane) that show high MON values [35,36]. This interval shows high positive weights due to cyclic paraffins that compensate the low MON values of the linear paraffins, thus maintaining these values low and almost unchanged. In the 60–70% (v/v) interval, the content of ethanol and mainly of paraffins is still high, but with no significant variations. The high positive weights of this interval are due to cyclic paraffins that, although in small amounts, remain until the 80% (v/v) fraction.

It was expected that the MON values of the 80% (v/v) fraction would be higher due to the presence of aromatic hydrocarbons. However, in this fraction there are also saturated hydrocarbons with eight to nine carbon atoms (as 2,5-dimethyloctane and

3-methylnonane), which have very low MON values [38]. From this fraction on, the reduction of levels of paraffins parallels a significant increase of aromatic contents (as 1,3-dimethylbenzene, 1-ethyl-3-methylbenzene, 1,2,4-trimethylbenzene and 1,3-diethylbenzene) and cyclic olefins (as cycloheptatriene) with MON values higher than those in other hydrocarbons [6,35,36].

The rotation system of the engine in the MON determination test requires severe operating conditions, as high temperature and higher rotation of the engine [10]. The areas with greater positive contribution in this model present components such as cyclical paraffins, ethanol, aromatic and olefins that tolerate these conditions longer due to their high MON values. This fact is related to the stability of the compounds in these fractions [39], which prevent gasoline from spontaneous ignition.

It is common to use the regression coefficients to explain the most important variables in predictions using PLS. However, in this work, these coefficients were not used for this purpose since they were not sufficient to explain the importance of the variables in the models, due to its random distribution.

The results obtained from the proposed method, compared with the results from the commercial spectrometer using test *t*, with 95% confidence, indicated that the calculated *t* values (t_{cal}) were lower than the *t* values (t_{tab}) (Table 1). This shows there is no significant difference between the PLS model based on distillation curves and the commercial infrared spectrometer [40]. Similarly, test *t* also indicated that there is no significant difference between the interlaboratory program and the proposed method, since calculated *t* values (t_{cal})^{*} were lower than the *t* values in Table 1 (t_{tab})^{*}, thus proving the accuracy of the proposed method.

The reproducibility and repeatability values obtained by the proposed method (0.10 and 0.11) were compared with the maximum limits established by ASTM D2700 (0.2 and 0.9, respectively) using test *F* [40] and they indicated, with 95% confidence, that the repeatability and reproducibility values for the proposed method are lower than the maximum established by ASTM D2700. This confirms the accuracy of the proposed method.

Compared to the values described in literature, the *RMSEC* and *RMSEP* values of this model (0.051 and 0.078) were much lower than those obtained for other models using multivariate dielectric spectroscopy (1.65 and 0.68) [12], GC–FID (0.70 and 0.80) [22] and infrared mid-spectrometry (MIR) (0.31) [41], showing the high accuracy of the proposed method. The fact that *RMSEP* is slightly higher than the *RMSEC* value indicates that the model is not sensitive to random variations [23]. Moreover, the efficiency of the model is also shown by the *RPD* (5.97) value considered appropriate to obtain a good calibration model, indicating that residue was relatively low [32].

Table 1
RMSEC, *RMSEP* and Q^2 values, among other parameters used for the determination of MON and RON in automotive gasoline, based on distillation curves.

Parameter	MON	RON
Number of latent variables (LV)	8	10
Explained variance LV1 (%)	38.0	51.0
Explained variance LV2 (%)	47.0	4.2
Explained variance LV3 (%)	5.0	28.2
<i>RMSEC</i>	0.051	0.078
<i>RMSEP</i>	0.063	0.085
<i>t</i> Test (t_{cal}) (validation set)	2.4×10^{-4}	3.0×10^{-3}
<i>t</i> Test (t_{tab}) (validation set)	1.98	1.98
<i>t</i> Test (t_{cal}) [*] (interlaboratorial program)	1.63	1.81
<i>t</i> Test (t_{tab}) [*] (interlaboratorial program)	2.45	2.45
<i>RPD</i>	5.97	10.99
Repeatability (proposed method)	0.10	0.29
Reproducibility (proposed method)	0.11	0.40
Maximum repeatability (standard method)	0.20	0.20
Maximum reproducibility (standard method)	0.90	0.70

3.2. Prediction of RON

Another set of 150 samples was used to predict RON, out of which 100 were used for the calibration matrix and 50 for the validation matrix. The gasoline samples for this model were also chosen randomly to ensure that the sample set was well represented for modeling. In a similar manner to the procedure used for MON, the values were compared with those obtained using a commercial infrared spectrometer provided with a database from ASTM D2699. RON values for that set of data varied from 97.4 to 101.4.

Test *F* applied to *PRESS* values indicated that the model must be built using ten latent variables, as there is a significant difference in *PRESS* values when fewer variables are used. Three first variables explained 83.4% of model variance with 51.0% of explained variance for the first latent variable, 4.2% for the second and 28.2% for the third one. Fig. 1B indicates that in the first latent variable the most important fractions to determine RON were from 4% to 60% (v/v) and, for the third latent variable, fractions from 77% to 83% (v/v).

Fig. 2 shows that RON values increased until reaching the 60% (v/v) fraction, thus matching the interval with highest positive weights in the first latent variable (Fig. 1B). Three different behaviors may be observed in the next interval: from 60% to 70% (v/v), RON values remain unchanged, there is a sudden drop between 70% and 80% (v/v) and a significant increase in RON values in the 80–90% (v/v) interval. In the 60–80% (v/v) interval, RON value fraction variations match the highest negative weights in the first and third latent variables (Fig. 1B).

As in the case of MON, fraction compositions were also analyzed to explain the behavior of variables in the weight graphic. Fig. 3 shows that in the 10–60% (v/v) interval the content of paraffin increases gradually and has higher levels of paraffin than other components, indicating that they affect more the RON values. The results obtained through GC/MS showed that up to the 60% (v/v) fraction there is an increase in carbon chains complexity, that is, there is a decrease in linear (as hexane) and branched paraffins (as 3-methylpentane and 3,3,4-trimethylhexane) and a simultaneous increase of cyclic paraffins (as methylcyclopentane and 1,3-dimethylcyclopentane), which causes a gradual increase of RON values.

In the test to determine RON, the combustible mixture is not heated so the test starts at a lower temperature, compared to MON, and engine rotation is lower. Lighter hydrocarbons resist detonation for a longer period of time in an engine when operating conditions are mild, as opposed to severe operating conditions. In the distillation curve, hydrocarbons in the 4–60% (v/v) interval have those characteristics and thus explain the importance of that interval to determine RON.

Fig. 3 shows that the composition of fractions in the 60–70% (v/v) interval is similar and does not produce changes in RON values (Fig. 2). It was observed that in the 80% (v/v) fraction, branched paraffins with 8–9 carbon atoms (as 2,5-dimethyloctane and 3-methylnonane) have lower RON than MON values [38]. These hydrocarbons cause the sudden drop of RON values in that fraction and are enough to drastically reduce the value of this characteristic, even if the amount of aromatics and olefins were increased.

As mentioned above, in the remaining interval from 80% to 90% (v/v) (Fig. 3) there is an increase of aromatics and olefins with very high RON values [35,36] as well as a reduction of branched paraffins with 8–9 carbon atoms (as 2,5-dimethyloctane and 3-methylnonane). The increase of aromatic hydrocarbons does not contribute as much to the model as paraffins do in fractions 4–60% (v/v) (Fig. 1B).

T test was also performed to verify the accuracy of the method compared to the commercial spectrometer. This test also showed that (t_{cal}) values were lower than those of (t_{tab}), with 95% confidence. This shows there is no significant difference between the

PLS model based on distillation curves and the commercial spectrometer. In addition, the test indicated that $(t_{cal})^*$ values, from interlaboratory programs, were lower than $(t_{tab})^*$ values, thus proving the accuracy of the proposed method (Table 1).

As in the model for the prediction of MON, the reproducibility and repeatability values obtained by the proposed method (0.29 and 0.40) were compared [40] with the maximum limits established by ASTM D2699 (0.2 and 0.7), respectively, using test *F*. This test indicated, with 95% confidence, that the reproducibility and repeatability values for the proposed method are lower than the maximum values established by ASTM D2699, hence the proposed method is more accurate.

Compared with the values described in literature, the *RMSEC* and *RMSEP* values of this model (0.063 and 0.085) were lower than those obtained [12] using dielectric spectroscopy (1.65 and 1.00), FT-MIR (0.43) [41] and GC-FID (0.80 and 1.00) [22]. This also shows the proposed method is highly accurate. For the same reasons discussed in the MON prediction model, the *RPD* value (10.99) also confirms the high predictive power of the model [32].

4. Conclusions

The PLS multivariate method applied to the distillation curves enabled the prediction of MON and RON values in gasoline samples from different refineries, in a range between 81.6 and 83.4 for MON, and 97.3–101.4 for RON, and obtained low *RMSEC* and *RMSEP* values.

Analyses of gasoline distillation fractions by means of GC/MS and infrared explained the relationship of the hydrocarbons in each fraction and the importance of the variables in each model. In the determination of MON, the presence of paraffins and ethanol is responsible for the great importance of variables in the second latent variable and of aromatics and olefins in the first latent variable. However, in the model to predict RON, the presence of paraffins is responsible for the importance of variables in the first latent variable and has more impact on the model than the third latent variable.

The use of distillation curves associated with chemometric techniques produced highly accurate results and demonstrated that the assay may be easily implemented in routine analyses. In addition to quality parameters established by ANP for the assay, other parameters, as alcoholic content, specific gravity [42] and steam pressure may also be predicted, to show the versatility of chemometric models based on distillation curves.

References

- [1] Wiedemann LSM, d'Avila LA, Azevedo DA. Adulteration detection of Brazilian gasoline samples by statistical analysis. *Fuel* 2005;84:467–73.
- [2] Agência Nacional do Petróleo e Biocombustíveis (ANP). <<http://www.anp.gov.br>> [accessed in November 2010].
- [3] Beires JS. Motores de explosão e diesel. 6th ed. São Paulo: LEP; 1964.
- [4] Pulkrabek WW. Engineering fundamental of internal combustion engine. 1st ed. Upper Saddle River: Prentice Hall; 1997.
- [5] Myers Jr ME, Stollsteimer J, Wims MA. Determination of gasoline octane numbers from chemical composition. *Anal Chem* 1975;47:2301–4.
- [6] Wauquier JP. El refinó del petróleo. Madrid: Diaz de Santos; 1995.
- [7] ASTM – American Society for Testing and Materials. Standard test method for motor octane number of spark-ignition engine fuel. West Conshohocken, (PA): ASTM D2700; 2009.
- [8] ASTM – American Society for Testing and Materials. Standard test method for research octane number of spark-ignition engine fuel. West Conshohocken, (PA): ASTM D2699; 2009.
- [9] ASTM – American Society for Testing and Materials. Standard specification for automotive spark-ignition engine fuel. West Conshohocken, (PA): ASTM D4814; 2010.
- [10] Santos JA. Informativo PETROBRAS – produtos PETROBRAS – gasolina automotiva. 5a ed. Betim: Refinaria Gabriel Passos; 2000.
- [11] Cruz JFM. Caracterização de gasolinas por espectroscopia FT – Raman. Thesis, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro; 2003.
- [12] Guan L, Feng XL, Li ZC, Lin GM. Determination of octane numbers for clean gasoline using dielectric spectroscopy. *Fuel* 2009;88:1453–9.
- [13] Tosello SA, Dupuyb N, Kisterb J. Contribution of external parameter orthogonalisation for calibration transfer in short waves – near infrared spectroscopy application to gasoline quality. *Anal Chim Acta* 2009;642:6–11.
- [14] Bao X, Dai L. Partial least squares with outlier detection in spectral analysis: a tool to predict gasoline properties. *Fuel* 2009;88:1216–20.
- [15] Petroleum Analyzer Company LP. Petrospec calibration update software windows version, operating manual version 1. USA; 1999.
- [16] ASTM – American Society for Testing and Materials. Standard practices for infrared multivariate quantitative analysis. West Conshohocken, (PA): ASTM E1655; 2005.
- [17] Fodor GE, Mason RA. A. Estimation of middle distillate fuel properties by FT-IR. *Appl Spectrosc* 1999;53:1292–8.
- [18] Oliveira FCC, de Souza ATPC, Dias JA, Dias SCL, Rubim JC. A escolha da faixa espectral no uso combinado de métodos espectroscópicos e quimiométricos. *Quim Nova* 2004;27:218–25.
- [19] Flumignan DL, Ferreira FO, Tininis AG, Oliveira JE. Abstracts of Papers. In: 234th ACS National Meeting, Boston, MA, Fuel-011, 2007. p. 19–23.
- [20] Cooper JB, Wise KL, Welch WT, Sumner MB, Wilt BK, Bledsoe RR. Comparison of near-IR, Raman, and mid-IR spectroscopies for the determination of BTEX in petroleum fuels. *Appl Spectrosc* 1997;51:1613–20.
- [21] Cooper JB, Groves J, Welch WT. Determination of octane numbers and Reid vapor pressure of commercial petroleum fuels using FT-Raman spectroscopy and partial least-squares regression analysis. *Anal Chem* 1995;67:4096–100.
- [22] Flumignan DL, Ferreira FO, Tininis AG, Oliveira JE. Multivariate calibrations in gas chromatographic profiles for prediction of several physicochemical parameters of Brazilian commercial gasoline. *Chemom Intell Lab Syst* 2008;92:53–60.
- [23] Kelly JJ, Barbow CH, Jinguji TM, Callis JV. Prediction of gasoline octane numbers from near-infrared spectral features in the range 660–1215 nm. *Anal Chem* 1989;61:313–20.
- [24] ASTM – American Society for Testing and Materials. Standard test for distillation of petroleum products at atmospheric pressure. West Conshohocken, (PA): ASTM-D86; 2010.
- [25] Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Anal Chim Acta* 1986;185:1–17.
- [26] Hemmateenejad B, Abbaspour A, Maghami H, Miri R, Panjehshahin MR. Partial least squares-based multivariate spectral calibration method for simultaneous determination of beta-carboline derivatives in peganum harmala seed extracts. *Anal Chim Acta* 2006;575:290–9.
- [27] Wold S, Sjöstrom M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 2001;58:109–30.
- [28] Minitab-14, Statistical software for windows. State College, (PA): Minitab Inc.; 2003.
- [29] Ni Y, Huang C, Kokot S. A Kinetic spectrophotometric method for determination of ternary mixtures of reducing sugars with the aid of artificial neural and multivariate calibration. *Anal Chim Acta* 2003;480:53–65.
- [30] Wold S, Sjostrom M. Chemometrics: present and futures success. *Chemometr Intell Lab Syst* 1998;44:3–14.
- [31] Li H, Liang Y, Xu Q, Cao D. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. *Anal Chim Acta* 2009;648:77–84.
- [32] Zhang XB, Fenga YC, Hua CQ. Feasibility and extension of universal quantitative models for moisture content determination in beta-lactam powder injections by near-infrared spectroscopy. *Anal Chim Acta* 2008;630:131–40.
- [33] ISO – International Organization for Standardization. Accuracy (trueness and precision) of measurement methods and results. Part 2. Basic method for the determination of repeatability and reproducibility of a standard measurement method. ISO 5725–2, Geneva; 2002.
- [34] Kramer R. Chemometric techniques for quantitative analysis. New York: Marcel Dekker; 1998.
- [35] Hobson GD, Pohl W. Modern petroleum technology. 5th ed. New York: John Wiley & Sons; 1984.
- [36] Satterfield CN. Heterogenous catalysis in practice. New York: McGraw-Hill; 1980.
- [37] Rasskazchikova TV, Kapustin VM, Karpov SA. Ethanol as high-octane additive to automotive gasolines. Production and use in Russia and abroad. *Chem Technol Fuels Oils* 2004;40:203–300.
- [38] Ghosh P, Hickey KJ, Jaffe SB. Development of a detailed gasoline composition-based octane model. *Ind Eng Res* 2006;45:337–45.
- [39] Solomons TW. Química orgânica. 9th ed. Rio de Janeiro: LTC; 2009.
- [40] Harris DC. Análise química quantitativa. 6th ed. Rio de Janeiro: LTC; 2005.
- [41] Andrade JM, Sánchez MS, Sarabia LA. Applicability of high-absorbance MIR spectroscopy in industrial quality control of reformed gasolines. *Chemometr Intell Lab Syst* 1999;46:41–5.
- [42] Aleme HG, Costa LM, Barbeira PJS. Determination of ethanol and specific gravity in gasoline by distillation curves and multivariate analysis. *Talanta* 2009;78:1422–8.