



Determination of ethanol and specific gravity in gasoline by distillation curves and multivariate analysis

Helga G. Aleme^a, Letícia M. Costa^b, Paulo J.S. Barbeira^{a,*}

^a Laboratório de Ensaios de Combustíveis, Departamento de Química, ICEx, UFMG, Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, Minas Gerais, Brazil

^b Grupo de Espectrometria Atômica e Preparo de Amostras, Departamento de Química, ICEx, UFMG, Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, Minas Gerais, Brazil

ARTICLE INFO

Article history:

Received 25 November 2008

Received in revised form 17 February 2009

Accepted 18 February 2009

Available online 4 March 2009

Keywords:

Ethanol

Gasoline

Partial least-squares regression

Distillation curve

Specific gravity

ABSTRACT

The partial least-squares regression method (PLS) was employed to predict the amount of ethanol and specific gravity in automotive gasoline using distillation curves (ASTM-D86). Additionally, a comparison was made between regression coefficients of all the algorithms, after selecting the number of latent variables. The low values obtained for RMSEC and RMSEP, associated with high accuracy when compared to the standard methodologies (NBR-13992, ASTM-D4052 and D1298) showed that PLS was efficient to determine the ethanol content and specific gravity in gasoline, since the model contains samples of different gasoline compositions, thus reflecting the variety of fuel in the Brazilian market. In addition, the proposed method is low cost, time reducing and easy to implement, as it utilizes the results of a routine assay carried out to evaluate the quality of automotive fuel.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

With the increase in energy demands throughout the latest centuries, fossil fuels are becoming limited. As a result, there is a major interest in developing alternative energy resources. Brazil was the first country in the world to change tetra-ethyl lead for ethanol as an antiknocking additive in automotive gasoline. This occurred after the introduction of Programa Nacional do Álcool (PROALCOOL) in 1975, which was created to stimulate the production and consumption of alcohol in the domestic and external markets. The addition of alcohol to gasoline has had some benefits such as the reduction of harmful emissions and the lessening of dependence on petroleum [1].

Since 1979 different amounts of ethanol, varying from 19 to 27% (v/v), have been added to Brazilian gasoline and, nowadays, this percentage is approximately 24–26% (v/v) of anhydrous ethanol [2]. According to ANP (Brazilian regulatory authority), around 23 billion liters of ethanol–gasoline mixture were commercialized [3] in 2008, and about 30% of these samples violated Brazilian legislation [4]. In Brazil, the determination of ethanol content in gasoline is carried out according to NBR 13992 [5], based on an extraction with solvent, whereby an aqueous NaCl solution is added to the sample. The alcoholic content, % (v/v), in the sample is determined by the

difference between final and initial volumes of water, after phase separation.

The specific gravity of gasolines commercialized in gas stations has no limits defined by ANP [2], with the exception of the gasolines used for the first filling and motor homologation, but it is an important parameter to evaluate in order to guarantee the good functioning of the engine. This assay can be performed by standardized methods such as American Society for Testing and Materials (ASTM) D1298 and D4052. ASTM-D1298 [6] describes the use of glass hydrometers, while ASTM-D4052 [7] explains the digital density meter. In the latter, a small volume of liquid sample is introduced into an oscillating U-tube, excited into oscillation by a Piezo element, and the oscillating frequency of the U-tube is directly related to the specific gravity of the filled sample.

The addition of ethanol to gasoline causes some changes in its volatility, vapor pressure and enthalpy of vaporization. In general, the increase of ethanol concentration causes a reduction of the vapor pressure and an increase in the enthalpy of vaporization [8].

Multivariate calibration associated to infrared and gas chromatography techniques has been used to predict several automotive fuel properties as specific gravity [9–11], viscosity [9,11–13], distillation fractions [9,11,13,14], octane number [14,15], cetane number [9,11,14–16], sulfur content [17,18], oxygen (MTBE) [15,19], aromatics [12,15,20], flash point [12,13], benzene [15], motor octane number (MON) [10,21] and research octane number (RON) [21].

Eight parameters (flash point, aromatics, freezing point, viscosity and distillation—initial, 10% (v/v), 90% (v/v) and final boiling

* Corresponding author. Tel.: +55 3134095767; fax: +55 3134096650.

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

point) of kerosene were predicted by Gómez-Carracedo et al. [13] applying vapor-phase generation, Fourier transform mid-infrared (FT-MIR) spectra and partial least-squares regression (PLS). The corrected average prediction errors are low and the joint test for the slope and intercept ensures that no bias is being introduced. Further, the precision values (repeatability and reproducibility) are lower than those from the corresponding official methods, the unique relevant exception being viscosity.

Oliveira et al. [10] evaluated MON and specific gravity for gasoline and alcoholic content, and specific gravity for hydrated ethanol by PLS and FT-NIR (Fourier transform near infrared). In the determination of alcoholic content and specific gravity, the method proposed herein – compared to the correlation spectra method for choosing the adequate spectral region – leads to PLS models with lower root mean square error of prediction (RMSEP) values. For specific gravity, the RMSEP values obtained were 0.0026 and 0.00019, for gasoline and hydrated ethanol respectively, and 0.0399 for alcoholic content.

Santos et al. [11] carried out a comparative study using infrared spectroscopy techniques (FTIR, FT-NIR and FT-Raman) associated with PLS and artificial neural network (ANN) along with diesel oil samples, for the determination of specific gravity, distillation temperatures at 50 and 85% (v/v) volume evaporated, viscosity and total sulfur content. In the specific gravity determination, considering the entire spectra region, the RMSEP values obtained in the range of 0.00071 and 0.0029 for other types of techniques led us to conclude that the best technique was Fourier transform infrared spectroscopy with conventional attenuated total reflectance (FTIR-ART1), since it provides a larger spectral range, better signal-to-noise ratios and lower RMSEP values.

Balabin et al. [22] used infrared (NIR) spectroscopy data and compared the performances of linear (MLR – multiple linear regression, PCR – principal component regression and PLS) and non-linear calibration techniques (polynomial and spline PLS, ANN) to predict six important gasoline properties: specific gravity and the temperatures of initial boiling point (IBP), final boiling point (FBP), and 10, 50 and 90% (v/v) volume evaporated points. The results showed that the non-linear methods proved their superiority over the linear ones, and neural networks turned out to be the most suitable methods for building a calibration model, providing low root mean square error of calibration with cross validation (RMSECV) values.

In order to simplify and speed up the analytical process for gasoline during inspection or production procedures, this research proposes to quantify the ethanol content in gasoline samples – which is the main nonconformity detected with the ANP regulations in Brazil – and the specific gravity, employing distillation curves (ASTM-D86), a routine assay that is done to evaluate the quality of gasoline associated with multivariate calibration, based on the PLS method.

ASTM-D86 [23] describes the method for distillation at atmospheric pressure of several petroleum products in order to determine volatility features, another critical problem of nonconformities in Brazil. This is carried out by checking if the light and heavy proportions of the fuel produced are appropriate to enable good performance during combustion, and to detect adulterations with other products. For Brazilian automotive gasoline, the ANP establishes maximum temperature values for 10, 50 and 90% of the volume recovered, beyond final boiling point and residue volume [2].

2. Experimental

2.1. Samples

2.1.1. Alcoholic content prediction

Two different sets of samples were used to build calibration and validation PLS models. In the first group, named ethanol doped

gasoline (EDG), different amounts of anhydrous ethanol were used to dope a gasoline in a final proportion of 15–30% (v/v). The second group was composed by samples collected in fuel stations by the ANP's Fuel Quality Control Program (FQCP-ANP), with different amounts of ethanol. As these samples have different origins, which directly affect the properties of the fuel [24–26], a representative set of samples from different refineries was selected and named Monitoring Program Group (MPG).

The gasoline, without ethanol, used to prepare the EDG samples was supplied by REGAP (Gabriel Passos Refinery) and doped with anhydrous ethanol (BR Distribuidora). The calibration matrix was constructed with 34 samples, whereas validation matrix with 14 samples.

The MPG group was built with a total of 135 samples of regular gasolines, collected in the east of Minas Gerais, Brazil. In this group, the addition of 19–31% (v/v) of ethanol has been a routine procedure performed by fuel distributors. For these sets of monitoring samples, 100 samples were used as calibration matrix and 35 samples were used as the validation matrix.

2.1.2. Specific gravity prediction

Another set of 135 samples was used in the determination of specific gravity. In this set, 35 samples were used for validation, aiming at determining the specific gravity of the gasoline, and the values were later compared with those obtained using the digital density meter (ASTM-D4052). The samples of this set came from five different refineries, all of them belong to FQCP-ANP, and have different quantities of ethanol, varying from 19 to 31% (v/v) and specific gravities in the range of 0.740–0.768 g mL⁻¹.

The gasoline samples were stored in appropriate polyethylene flasks, sealed and kept in cold storage (8–15 °C) until distillation analysis in order to avoid loss of volatile components [23].

2.2. Equipment and materials

Gasoline samples were distilled in a Herzog HDA 627 Automatic Distillation Analyzer, according to ASTM-D86 assay [23].

In the ethanol addition procedure, a graduated cylinder of 100 mL with 1 mL resolution, calibrated and covered according to NBR-13992 [5] was used. The 10% (w/v) NaCl solution was prepared by dissolving the salt (Reagen) in distilled water.

The specific gravity was determined by using an Anton Paar DMA 4500 digital density meter, according to ASTM-D4052 assay [7].

2.3. Experimental procedure

2.3.1. Distillation

For the distillation procedure, 100 mL of gasoline were transferred to a specific distillation flask equipped with a thermocouple sensor, and heated to keep the distillation ratio between 4 and 5 mL min⁻¹. The distilled and condensed steam was collected in a cooled test-tube (13–18 °C) and the recovered volume was measured with a digital volume sensor. Distillation curves (distillation temperature according to the recovered volume) were obtained after correcting temperature readings to atmospheric pressure of 101.3 kPa, and volume loss after measuring residue volume, according to ASTM-D86 [23].

2.3.2. Ethanol content

In the ethanol addition procedure, 50 mL of gasoline were added to a clean and dry graduated cylinder of 100 mL. Volume was completed with a 10% (w/v) NaCl solution, followed by a non-vigorous agitation and allowed to rest for 15 min, as described in NBR-13992 [5]. The final volume of the aqueous phase was read after this time and expressed in milliliters. The ethanol content is calculated

according to the equation below:

$$\text{ethanol content (v/v\%)} = [(A - 50) \times 2] + 1 \quad (1)$$

in which A is the final volume of the liquid phase given in milliliters. The addition of 1% (v/v) to the final result is necessary to compensate the volume contraction resulting from the water–ethanol mixture. When the difference $(A - 50)$ is less than 0.5 mL, the result must be written as $\leq 1\%$ (v/v) [5].

2.3.3. Specific gravity

In this assay, approximately 0.7 mL of the gasoline previously cooled down to 10 °C is introduced, with a dry and clean syringe, into an oscillating sample tube, and the change in oscillation frequency caused by the change in the mass of the tube is used in conjunction with calibration data to determine the specific gravity of the sample [7]. After temperature stabilization of the measurement cell at 20 °C the specific gravity is noted.

2.3.4. Calibration multivariate tool

The aim of PLS is to find a small number of relevant factors that are predictive for Y and utilize X efficiently. The method effectively achieves a canonical decomposition of X in a set of orthogonal factors which are used for fitting Y [27]. In this method, each latent variable of the X matrix is modified for its covariance and the vector of matrix Y was maximized. In PLS, matrixes X and Y were decomposed in lesser matrixes according to the equations below:

$$X = TP^T + E \quad (2)$$

$$y = UQ^T + F \quad (3)$$

where X and Y are matrixes that will be decomposed, T and U are score matrixes with perpendicular rows (T and U are n -rows and d -columns matrixes, respectively), P is the matrix of d -rows and p -columns of loading matrix X , and E is the error (residue). Q is the matrix of d -rows and m -columns of the loading matrix of Y , F is the error (residue) for the Y matrix [28].

All calculations were performed with Minitab Release software (version 14 for Windows).

2.3.5. Evaluation of repeatability and reproducibility

The evaluation of repeatability and reproducibility of the methods were carried out according to ISO-5725-2 [29]. Therefore, 10 samples of gasoline were used and, for each sample, assays were done by three different analysts with seven replicates for each one [30], producing a total of 21 results.

3. Results and discussion

In order for the distillation to occur within a rate of 4–5 mL min⁻¹ [23], the resistance needs to be significantly heated. Initially, this can lead to a vigorous ebullition – causing oscillation in temperature values – which raises from the initial point to approximately 4% (v/v) and produces low reproducible measurements. Another issue that should be emphasized is that not all the distillation curves have equivalent temperatures in 94–98% (v/v) range due the pyrolysis phenomenon, in which the larger molecules are broken causing the decrease of boiling temperature [23]. Therefore, the interval 4–93% (v/v) (T4 to T93) was used to construct the PLS models.

The addition of ethanol leads to an essential change in the shape of the distillation curve: a plateau appears at about 70–77 °C. Fig. 1 shows that the distillation curves from the mixtures of gasoline–ethanol – up to a concentration of 40% (v/v) of ethanol – have lower boiling temperatures than the distillation curves of the gasoline without the addition of ethanol. This occurs due to the formation of an azeotropic mixture of hydrocarbons–ethanol

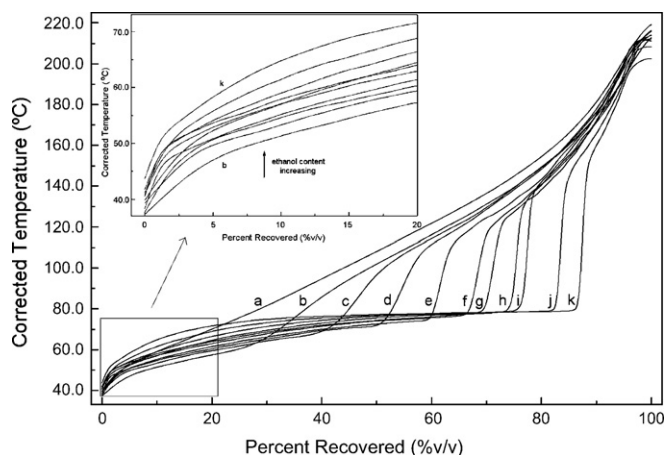


Fig. 1. Distillation curves (ASTM-D86) obtained for a gasoline doped with different ethanol contents. (a) 0% (v/v); (b) 5% (v/v); (c) 10% (v/v); (d) 15% (v/v); (e) 20% (v/v); (f) 25% (v/v); (g) 30% (v/v); (h) 35% (v/v); (i) 40% (v/v); (j) 50% (v/v); (k) 60% (v/v).

that have lower boiling temperatures than the original hydrocarbons [31], since mixtures of ethanol with hydrocarbons exhibit large deviations from ideal mixing. Consequently, Raoult's law cannot be used. It has also been observed that, the more ethanol added, the higher the plateau formed in the distillation curve. As from the concentration of 50% (v/v) of ethanol the plateaus become closer to each other and the initial temperatures rise. Fig. 1 also shows that at the beginning of the distillation process the influence of the azeotropic mixture is stronger than that at the end, because the amount of ethanol in the system is higher at the beginning. Most of the ethanol boiled in the 75–80 °C range and did not form an azeotropic mixture at the end of the distillation process [32]. Another fact to be considered is the dispersed structure of ethanol–gasoline mixture, demonstrated by Balabin et al. [33] using the method of correlation spectroscopy of scattered light. The results showed that the colloidal structures are microspheres, whose size depends on the concentration of ethanol.

Due to the fact that the samples are from different refineries, and therefore have different chemical composition, some normalizations were used in an attempt to minimize the differences in behavior of the distillation curves, and to improve the RMSEC values. For this, each distillation point was divided by the temperature reading at different evaporated volumes, from 5% (v/v) (T5) to the final boiling point (FBP), out of a total of 19 normalizations.

The PLS data pre-processing was carried out in two different ways, as reported by Geladi and Kowalsky [34]: autoscaling and centering. In the first one, each data was subtracted from the average of its column and divided by the standard deviation. In the latter, each data was subtracted from the average value of its corresponding column. The cross validation was carried out using the leave-one-out method.

3.1. Alcoholic content prediction

The PLS model was built from calibration data. The alcoholic content of the prediction sample set was obtained using the best number of factors. The selection of components in PLS is very important to reach a good prediction [34,35]. With numerous and correlated X -variables there is a substantial risk of “over-fitting”, thus creating a well fitting model with little or no predictive power. Hence, a strict test of the predictive significance of each PLS component is necessary, stopping when components start to be non-significant [36].

The number of latent variables with higher Q^2 values (coefficient of linearity for prediction using cross validation) were chosen

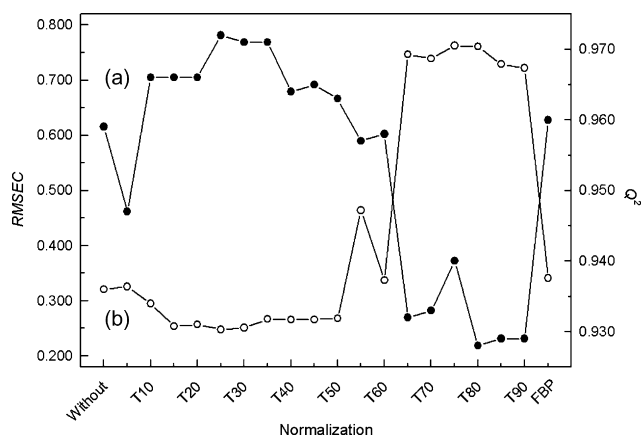


Fig. 2. RMSEC and Q^2 values obtained for models with different normalizations for the determination of alcoholic content in gasoline with the MPG set. (a) Q^2 ; (b) RMSEC.

to build the PLS model [37], and it was calculated by PRESS values, according to Wakeling and Morris [38]. For the EDG set, it was observed that only three variables were used to build the model, and for the MPG set, 4–10 latent variables were used, depending on the type of normalization applied. Fig. 2 (curve a) shows that the higher values were obtained for the normalization with temperature reading at 25% evaporated volume (T25).

The weight graphs of sets EDG and MPG showed a different behavior of the variables in the distillation curve. This difference may be related to the fact that the EDG set is composed of samples of only one gasoline doped with different concentrations of ethanol, unlike the MPG set composed of samples from various refineries. In the EDG set, three latent variables were responsible for 81.9% of the total variance explained by the model (Table 1), out of which 76.4% were from the first latent variable.

Analysis of Fig. 3A shows that three areas of distilled volume, % (v/v), were formed in the EDG set. The first area (I) is related to the initial part of the distillation curve, between 4 and 55% (v/v). The second area (II) is related to the interval between 56 and 63% (v/v), and a third area (III) between 64 and 93% (v/v). It was also observed that these areas have similar weights in latent variable 1. In the MPG set without normalization (Fig. 3B), three latent variables explained 80.8% of the model variance, 31.1% corresponds to the first latent variable and 45.6% to the second latent variable. The three areas have different weights for the two first latent variables, and show a distinct behavior from the EDG set. This behavior is due to the different origin of the samples, thus indicating that gasoline composition has a significant impact on the weight of variables. A

Table 1

RMSEC, RMSEP and Q^2 values, among other parameters, obtained for several models built with different sets (EDG and MPG), in the determination of alcoholic content in gasoline as from the distillation curves.

Parameter	EDG	MPG without normalization	MPG normalized by 25% (v/v) point
Q^2	0.992	0.959	0.972
Number of latent variables	3	9	10
Explained variance LV1 (%)	76.4	31.1	42.9
Explained variance LV2 (%)	3.6	45.6	21.7
Explained variance LV3 (%)	1.9	4.1	11.7
RMSEC	0.227	0.321	0.248
RMSEP	0.304	0.691	0.538
RMSEP autoscaled	0.437	0.665	0.652
RMSEP mean centered	0.422	0.641	0.534
Test $t(t_{calc})$	1.0	1.7	1.8
Test $t(t_{tab})$	2.1	2.0	2.0

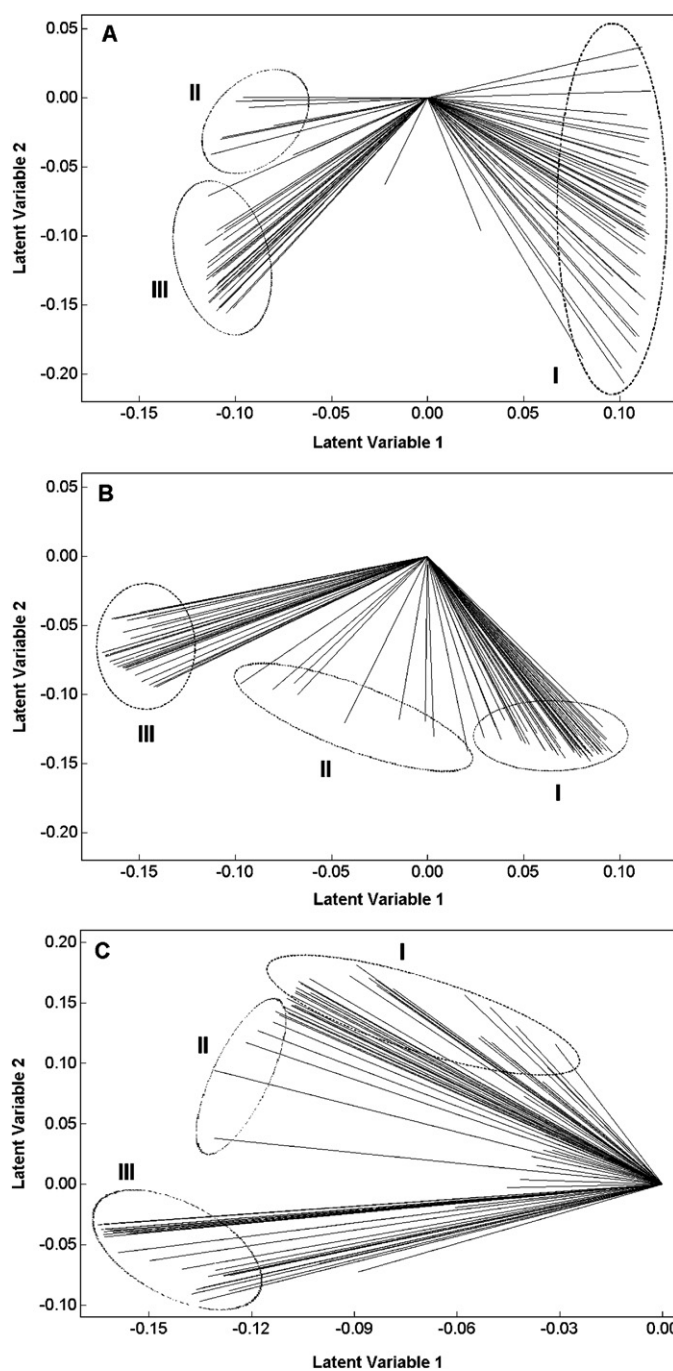


Fig. 3. Weight graphics obtained for the determination of alcoholic content in gasoline with EDG (A), MPG (B) and MPG25 (C) sets, highlighting three different areas of the distillation curves: (I) 4–55% (v/v); (II) 56–63% (v/v); (III) 64–93% (v/v).

previous study [26] confirmed that the initial part of the distillation curve, between 34 and 45% (v/v), had greater weight in the model for prediction of gasoline origin with distillation curves.

For the MPG set normalized by the temperature reading at 25% volume evaporated, MPG25 (Fig. 3C) three latent variables explained just 76.3% of the model variance, with 42.9% for the first one. The use of this normalization provided a weight increase of the first latent variable in area C, which corresponds to the distillation of ethanol. Thus, the variables of area C have greater weight in the explanation of the model, mainly in the 77–93% (v/v) range, since the first latent variable is the most important in determining the alcoholic content.

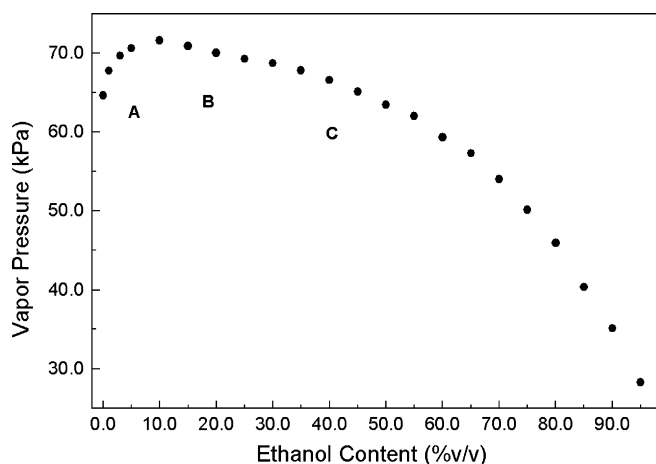


Fig. 4. Behavior of vapor pressure according to the volume of ethanol added to gasoline. Figure adapted from Brand et al. [39].

After building the calibration models, it was observed that the samples of sets EDG and MPG had separated according to the different alcoholic content in the two first latent variables. In the EDG set, the data set was separated into two groups with alcoholic content ranging from 15 to 23 and 24 to 30% (v/v), whereas for the MPG set the separations ranged from 19 to 24 and 25 to 31% (v/v), with or without normalizations. This separation is not only related to the formation of azeotropes mentioned above, but also to the vapor pressure variance in ethanol–gasoline mixtures, an important characteristic which is indicative of the behavior of fuel in different operating conditions in spark-ignition engines, and is also directly related to its volatility.

Fuels with high vapor pressure present a high rate of emission of volatile compounds [8], and the addition of different concentrations of ethanol causes alterations in vapor pressure, since the pressure of gasoline is approximately three times higher than that of ethanol, as shown in research done by Brand et al. [39]. In this research, vapor pressure measurements were carried out in several ethanol–gasoline mixtures, demonstrating that the vapor pressure of the gasoline was initially raised by the addition of up to 4% (v/v) of ethanol, but at an ethanol concentration of 15% (v/v) the vapor pressure dropped, as can be observed in Fig. 4 (reproduced from Brand et al. [39] results). Vapor pressure reduced approximately $0.159 \text{ kPa } ^\circ\text{C}^{-1}$ in the range of 15 and 25% (v/v) of ethanol, whereas in the range of 25 and 30% (v/v) it decreased slightly less at a rate of $0.110 \text{ kPa } ^\circ\text{C}^{-1}$. These first two ranges coincide, to an extent, with the ranges observed in the separation of sets EDG and MPG, demonstrating a close relationship between vapor pressure and the separation of alcoholic content.

RMSEC values were calculated from the calibration models for each set of samples (Table 1), and showed that the EDG set presented the lowest value for the model whose data had neither been normalized nor preprocessed. The same did not occur with the MPG set, which presented the lowest RMSEC for the data normalized by the temperature reading at 25% volume evaporated. The EDG model was built with only one type of gasoline and, therefore, it presented lower RMSEC values than the MPG set. Test F [40] was carried out to determine the existence of a significant statistical difference between the RMSEC values obtained for the different normalizations. Using this test, it was determined with a confidence level of 95% that in the MPG set there was a significant difference between the models normalized by the temperature reading at 25% volume evaporated and those with no normalization. This demonstrates that this normalization was the most appropriate to determine the alcoholic content in automotive gasoline. In addition, for the EDG set the F test showed that all the RMSEC values are statistically

equivalent; indicating that in this instance there is no need to use normalizations.

The accuracy of each of the models may be verified by calculating the RMSEP [16]. According to Table 1, for the EDG set, the lowest RMSEP values were obtained for the models whose data had not been preprocessed nor normalized, and this value was lower than those for the MPG set. This is due to the higher linearity, that is, a higher Q^2 value of the analytical curves of set EDG over those of the MPG set.

The application of test F to the obtained RMSEP values demonstrated with a confidence level of 95% that, in general, preprocessing did not provide a significant improvement of the values in any of the sets studied. However, for the MPG set, the normalization by the temperature reading at 25% volume evaporated provided a significant decrease in the values found, and showed it is the best option.

An evaluation of the accuracy of the proposed method was done by comparing its results with those obtained through a different method such as NBR-13992, using test t (Eq. (4)) and analyzing the calculated t value (t_{calc}) [40]:

$$t_{calc} = \frac{\bar{d}}{\sqrt{\sum (d_i - \bar{d})^2 / n - 1}} \sqrt{n} \quad (4)$$

where \bar{d} is the average difference between methods, d_i the difference between methods for each sample, and n the number of data pairs.

The t values in the tables (t_{tab}) for the EDG and MPG sets of samples were lower than the calculated values (t_{calc}), with a confidence level of 95%, showing that there was no significant difference between the PLS models based on distillation curves and those of the NBR-13992 method. This demonstrates the accuracy of the proposed method.

To evaluate the repeatability and reproducibility of the method, the model chosen was one whose data was normalized by the temperature reading at 25% volume evaporated with the MPG set. The repeatability and reproducibility values obtained for NBR-13992 were 0.37 and 0.44% (v/v), respectively, whereas for the proposed method, the values were 0.54 and 0.57% (v/v), respectively. Test F , applied to these values, showed with a confidence level of 95% that the repeatability and reproducibility of both methods are equivalent and lower than the maximum limit established by NBR 13992 [5], 1 and 2% (v/v), respectively, thus demonstrating the effectiveness of the proposed method.

In addition to supplying the same results as the normalized method, the proposed method provides more decimal digits, i.e., it is more accurate, as the results obtained by NBR-13992 have no decimal digits due to the inexactness of the graduated cylinder. When the values obtained by the proposed method have no decimal digits, the repeatability and reproducibility values drop slightly (0.60 and 0.68, respectively), although they are still significantly equivalent to NBR-13992.

3.2. Prediction of specific gravity

The same normalizations and preprocesses mentioned above were applied to determine specific gravity, using distillation curves associated to multivariate calibration as well as cross validation (leave-one-out), in a set of samples from different refineries containing 100 samples for the calibration set and 35 for validation.

Several latent variables were used in this work, depending on the type of normalization used. Nine latent variables were employed to build the calibration curve in the non-normalized data, whereas 7–10 latent variables were used to build the calibration of data, depending on the type of normalization applied. As opposed to the

Table 2

RMSEC, RMSEP and Q^2 values, among other parameters, obtained in the determination of specific gravity in gasoline as from the distillation curves.

Parameter	Values
Q^2	0.841
Number of latent variables	9
Explained variance LV1 (%)	72.1
Explained variance LV2 (%)	5.3
Explained variance LV3 (%)	3.5
RMSEC	0.00063
RMSEP	0.00088
RMSEP autoscaled	0.00136
RMSEP mean centered	0.00088
Test t (t_{calc})	0.33
Test t (t_{tab})	2.0

determination of alcohol content, normalizations produced higher RMSEC values, as shown in Table 2 and Fig. 5. Test F , applied to these values showed, with a confidence level of 95%, that all the normalizations provided higher RMSEC values as well as statistically different, thus indicating that the calibration model should be built taking into consideration non-normalized data.

Compared to the values presented in other studies, the RMSEC value for this model (0.0006) was lower than the RMSECV value obtained by Balabin et al. [22] (0.0028) using PLS and FT-NIR. This comparison of values was feasible because the calibration model was built with the best amount of latent variables (associated to the highest Q^2 value), which was calculated using a cross-validation method, i.e., in this case the RMSEC value can be considered similar to the RMSECV value.

In the weight graph of non-normalized data, the first latent variable provided 72.1% of the model variance, whereas the second one presented 5.3%. Fig. 6 shows that the 8–53% (v/v) interval of the distillation curve presented the highest weight of the first latent variable, indicating that this variable is the most relevant to build the PLS model.

In the graph analysis of the three first latent variables, the same separation of samples was observed – according to the specific gravity – in two groups in the 0.740–0.750 and 0.751–0.768 g mL⁻¹ intervals. Although the addition of anhydrous alcohol to gasoline causes significant alterations in the distillation curve, the specific gravity of the mixture increases linearly in the 10 and 70% (v/v) interval.

For each sample, the estimate of RMSEP values was carried out using different preprocesses, and indicated (with a confidence level of 95% according to test F) that the lowest value is also associated with the model containing non-preprocessed or self-scaled data.

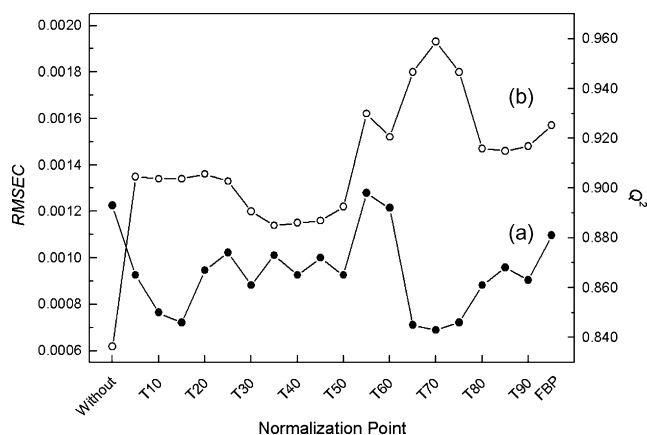


Fig. 5. RMSEC and Q^2 values obtained for models with different normalizations for the determination of specific gravity in gasoline. (a) Q^2 ; (b) RMSEC.

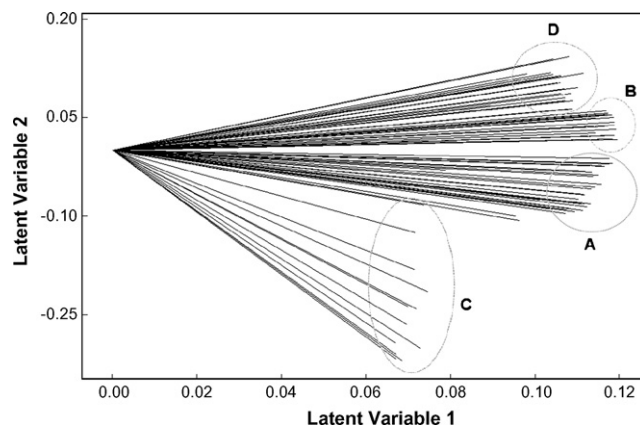


Fig. 6. Weight graphics obtained for the determination of specific gravity in gasoline, highlighting four different areas of the distillation curves: (A) 8–33% (v/v); (B) 34–53% (v/v); (C) 55–66% (v/v); (D) 70–93% (v/v).

Test t was also performed to check the accuracy of the proposed method compared to the ASTM-D4052 method. This test proved that the t_{calc} value was lower than the t_{tab} value, i.e., the accuracy of the proposed method is statistically similar to the accuracy of the ASTM-D4052 method. The analyses of the Q^2 values of the calibration models, as well as the results of tests t and F , indicated that the best method to determine specific gravity is the method with non-normalized distillation curves, unlike the result obtained in the determination of alcohol content.

Compared to the values presented in other studies, the RMSEP value for this model (0.0009) was lower than those obtained by Oliveira et al. [10] (0.0026 and 0.0050) using PLS and FT-NIR, thereby showing the great accuracy of the proposed method.

To evaluate the repeatability and reproducibility of the method, the model chosen included non-normalized and non-preprocessed data. The values obtained for repeatability and reproducibility for ASTM-D4052 were 0.00012 and 0.00020 g mL⁻¹, for ASTM-D1298 values were 0.00034 and 0.00040 g mL⁻¹, and for the method proposed in this work, values were 0.00065 and 0.00076 g mL⁻¹, respectively.

4. Conclusions

The distillation curves obtained by ASTM-D86, together with the PLS multivariate calibration method, managed to predict the alcoholic content in the 19–31% (v/v) range, with gasoline specific gravity in the 0.74–0.76 mg L⁻¹ range, regardless the refineries from which they came from, with low RMSEC and RMSEP values.

The weight graphs presented different behaviors of the distillation curve, of the EDG and MPG sets in the determination of the alcoholic content, and of the set of samples used in the determination of specific gravity.

For the determination of alcoholic content it was found that it is best to use a non-preprocessed, normalized by the temperature reading at 25% volume evaporated model with different origins and ethanol concentrations (as the MPG set). The repeatability and reproducibility values in the determination of the alcoholic content using the distillation curves and PLS were equivalent to those of the NBR-13992 method.

In the determination of specific gravity the best model was the one whose data had no preprocess nor normalization, with lower RMSEP values than those found in the literature. In addition to accuracy, the proposed method provides higher repeatability and reproducibility values and it is statistically different from the ASTM-D1298 and D4052 methods.

The use of distillation curves associated with chemometric techniques leads to high accuracy, demonstrating that it can be easily implemented in routine analyses, providing different parameters in only one assay.

Acknowledgements

The authors would like to thank CTPetro-FINEP and ANP for their financial support.

References

- [1] J.P.A. Magalhães, N. Kuperman, R.C. Machado, Proalcohol: Uma Avaliação Global, 1st ed., Publ. ASTEL, Rio de Janeiro, 1991.
- [2] Agência Nacional do Petróleo e Biocombustíveis, Portaria no. 309, 27/12/2001, <http://www.anp.gov.br>.
- [3] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Anuário Estatístico, 2007, <http://www.anp.gov.br>.
- [4] Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Boletim da Qualidade, Não-Conformidades no Brasil, December 2008, <http://www.anp.gov.br>.
- [5] Gasolina Automotiva-Determinação do Teor de Álcool Etilico Anidro Combustível (AEAC), NBR-13992, Associação Brasileira de Normas Técnicas, 1997.
- [6] Standard Test Method for Density, Relative Density (Specific Gravity), or API Gravity of Crude Petroleum and Liquid Petroleum Products by Hydrometer Method, ASTM-D1298, American Society for Testing and Materials, 2005.
- [7] Standard Test for Density and Relative Density of Liquids by Digital Density Meter, ASTM-D4052, American Society for Testing and Materials, 2002.
- [8] R. Cataluña, R. Silva, Quim. Nova 29 (2006) 580.
- [9] G.E. Fodor, R.A. Mason, S.A. Hutzler, Appl. Spectrosc. 53 (1999) 1292.
- [10] F.C.C. Oliveira, A.T.P.C. de Souza, J.A. Dias, S.C.L. Dias, J.C. Rubim, Quim. Nova 27 (2004) 218.
- [11] V.O. Santos Jr., F.C.C. Oliveira, D.G. Lima, A.C. Petry, E. Garcia, P.A.Z. Suarez, J.C. Rubim, Anal. Chim. Acta 547 (2005) 188.
- [12] V.K. Seinsche, Fuel Energy Abstr. 38 (1997) 11.
- [13] M.P. Gómez-Carracedo, J.M. Andrade, M. Calviño, E. Fernández, D. Prada, S. Munategui, Fuel 82 (2003) 1211.
- [14] D.L. Flumignan, F.O. Ferreira, A.G. Tininis, J.E. de Oliveira, Abstracts of Papers, 234th ACS National Meeting, Boston, MA, United States, August 19–23, 2007, FUEL-011.
- [15] B. Janis, Society of Automotive Engineers [Special Publication] SP-1461 (1999) 95.
- [16] J.J. Kelly, C.H. Barlow, T.M. Jinguji, J.B. Callis, Anal. Chem. 61 (1989) 313.
- [17] M.C. Breikreitz, I.M. Raimundo Jr., J.J.R. Rohwedder, C. Pasquini, H.A. Dantas Filho, G.E. Jose, M.C.U. Araujo, Analyst 128 (2003) 1204.
- [18] K.J. Johnson, R.E. Morris, S.L. Rose-Pehrsson, Energy Fuels 20 (2006) 727.
- [19] J.B. Cooper, K.L. Wise, W.T. Welch, R.R. Bledsoe, M.B. Sumner, Appl. Spectrosc. 50 (1996) 917.
- [20] J.B. Cooper, K.L. Wise, W.T. Welch, M.B. Sumner, B.K. Wilt, R.R. Bledsoe, Appl. Spectrosc. 51 (1997) 1613.
- [21] J.B. Cooper, K.L. Wise, J. Groves, W.T. Welch, Anal. Chem. 67 (1995) 4096.
- [22] R.M. Balabin, R.Z. Safieva, E.A. Lomakina, Chem. Intell. Lab. Syst. 88 (2007) 183.
- [23] Standard Test for Distillation of Petroleum Products at Atmospheric Pressure, ASTM-D86, American Society for Testing and Materials, 2007.
- [24] P.J.S. Barbeira, R.C.C. Pereira, C.N.C. Corgozinho, Energy Fuels 21 (2007) 2212.
- [25] R.M. Balabin, R.Z. Safieva, Fuel 87 (2008) 1096.
- [26] P.J.S. Barbeira, H.G. Aleme, L.M. Costa, Fuel 87 (2008) 3664.
- [27] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics. Part B, 1st ed., Elsevier Publ., Amsterdam, Lausanne, New York, Oxford, Shannon, Singapore, Tokyo, 1998.
- [28] M. Otto, Chemometrics: Statistics and Computer Application in Analytical Chemistry, 1st ed., Wiley-VCH Publ., Weinheim, New York, Chichester, Brisbane, Toronto, 1999.
- [29] ISO/IEC 5725-2, 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, 2002.
- [30] INMETRO DOQ-CGCRE-008, Orientação sobre Validação de Métodos de Ensaios Químicos, 2007.
- [31] R.M. Balabin, R.Z. Syunyaev, S.A. Karpov, Energy Fuels 21 (2007) 2460.
- [32] R. French, P. Malone, Fluid Phase Equilib. 228–229 (2005) 27.
- [33] R.M. Balabin, R.Z. Syunyaev, S.A. Karpov, Fuel 86 (2007) 323.
- [34] P. Geladi, B.R. Kowalsky, Anal. Chim. Acta 185 (1986) 1.
- [35] B. Hemmateenejad, A. Abbaspour, H. Maghamia, R. Miri, M.R. Panjehshahin, Anal. Chim. Acta 575 (2006) 290.
- [36] S. Wold, M. Sjöström, L. Eriksson, Chemometr. Intell. Lab. Syst. 58 (2001) 109.
- [37] Minitab-14, Statistical Software for Windows, Minitab Inc., State College, PA, 2003.
- [38] I.N. Wakeling, J.J. Morris, J. Chemometr. 7 (1993) 291.
- [39] J.I. Brand, J.A. Pumphrey, W.A. Scheller, Fuel 79 (2000) 1405.
- [40] D.C. Harris, Análise Química Quantitativa, 5th ed., LTC Editora, Rio de Janeiro, 2001.