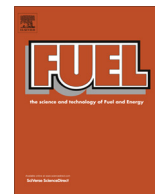


Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

Fuel

journal homepage: www.elsevier.com/locate/fuel

Detection and quantification of adulterants in gasoline using distillation curves and multivariate methods



Gisele Mendes, Paulo J.S. Barbeira *

Laboratório de Ensaios de Combustíveis, Departamento de Química, ICEx – Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, Minas Gerais, Brazil

HIGHLIGHTS

- PCA applied to distillation curves enables the detection of gasoline adulterations.
- The most important fractions for the discrimination were the 4–40%(v/v) interval.
- This interval was related to the increase of paraffin and isoparaffin content.
- PLS-DA enables the detection and quantification of the solvent used in adulterations.
- The method produced highly accurate results and is suitable for routine analysis.

ARTICLE INFO

Article history:

Received 4 February 2013
Received in revised form 15 April 2013
Accepted 24 April 2013
Available online 24 May 2013

Keywords:

Adulteration
Gasoline
Distillation
Multivariate analysis

ABSTRACT

This study has shown that the use of distillation curves combined with PCA (Principal Component Analysis) and PLS-DA (Partial Least Squares Discriminant Analysis) provides a model with enough sensitivity to discriminate adulterated and unadulterated gasoline samples, as well as, the determination of the solvent used in adulteration with minimum percentage of 97% accuracy. PLS-DA provided the prediction of adulterants with low RMSEC (Root Mean Square Error of Calibration) and low RMSEP (Root Mean Square Error of Prediction) when compared to other methods. The great advantage is the possibility to apply the results of the distillation curves to routine analysis (ASTM D86), therefore not requiring various assays, speeding up the analytical process. In addition to its feasibility this method can be quite useful in fuel quality monitoring and inspection procedures whilst having low cost and good reliability.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The end of government monopoly brought serious changes to retail and distribution of fuels in Brazil. One of the effects is the planned adulteration of fuel with the addition of controlled solvents, which have the objective of maintaining the product within the current specification [1]. This practice has occurred quite frequently despite the measures of Brazilian authorities through the ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis), which establishes technical specification through specific legislation regarding the minimum quality of fuels mainly guaranteeing the standardization of retail production [2].

The addition of solvents is one of the most common practices of adulteration of fuel due to the enormous difference in taxation between gasoline and solvents. The addition of illegal compounds to fuels can cause damaging and unpleasant issues to society such as environmental risk due to the emission of vapours and toxic

gases, i.e. CO and NO_x, less durability to the vehicles' engine, as well as, unfair market competition of fuel prices causing a great loss to the State in tax revenues [3,4]. The most commonly used solvents include: ethanol in excessive amounts, diesel, kerosene, refined petrochemicals, toluene, xylene, hexane, among others [5].

Several physicochemical properties are monitored to ensure the quality of Brazilian gasoline through the ANP [2], and these tests include specific mass measurements, distillation analysis, octane analysis, among others. However, the current specifications were chosen largely based on the good functioning of the engine, rather than identifying an illicit addition of solvents. Even though they are capable of indirectly identifying a fraud of any type of solvent in any proportion of addition, which can lead to adulterated gasoline to be in conformity with the physicochemical assays [5].

In recent years the ANP, after numerous cases reported by the media, created means to protect the consumer against the harmful consequences caused by adulteration. The Agency invested in a Marking Solvents Program, where mandatorily every solvent commercialised throughout the country had the addition of a chemical marker, which was designed to have no impact on the applicability

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

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

of the solvents when used individually. The presence of the marker in gasoline indicates that it has been adulterated with some kind of solvent, although it is a rather laborious process that requires laboratorial analysis, a tight logistic for the markers, and the monitoring process which becomes highly expensive [5]. It is very common to find gasoline in the Brazilian market with all the physical and chemical properties in accordance with the specifications of the ANP, but with the presence of traces of solvents [6].

A study carried out by LEC-UFGM (Laboratório de Ensaios de Combustíveis of UFGM) showed that in 2012 approximately 40% of the gasoline samples analysed were considered atypical [7]. Atypical samples are the ones that have a different profile than the majority of samples, though they are in accordance with the parameters established by the ANP. This can be attributed to careful and meticulous adulteration of fuel with the addition of solvents so that the final product stays in accordance with legal standards or a different origin from the one stated [8]. Therefore, new analytical methods must be developed to detect these adulterations. The methods are simple, fast and efficient in ensuring the quality and authenticity of commercial fuels and therefore highly recommended for routine analysis.

Currently the scientific literature has reported alternative methodologies in the detection of adulteration in gasoline. These methodologies use the employment of chemometric tools combined with several conventional techniques of gasoline analysis [1,3,5,9–17]. The large majority refer to studies using chromatographic [1,3,9] and spectroscopic [11–17] methods.

Wiedemann et al. [3] carried out a study to detect adulteration in gasoline samples using the results obtained through physicochemical properties and chromatographic data in the gas phase of samples combined with hierarchical clustering analysis (HCA).

Balabin et al. [13–16] used near infrared, different chemometric tools and artificial neural networks to predict different gasoline properties [13,14] and classify gasoline and gasoline fractions by source (refinery or process) and type [15,16]. The models have low errors and could be useful to detect anomalous gasoline samples.

In another study Monteiro et al. [17] used hydrogen nuclear magnetic resonance (^1H RMN) of retail gasoline and added solvents combined with principal component analysis (PCA) and hierarchical clustering analysis (HCA) to distinguish samples in conformity and not in conformity. The results indicated a tendency of non-complying samples clustering with the increase of the content of solvent added.

Recent studies have shown the great potential of distillation curves, or a few specific points, for the analysis of different parameters of petroleum products [8,5,18–39], as refinery origin [24,25,27] enthalpy of combustion [21], specific gravity [26,31], kinematic viscosity [31], octane numbers [30], cetane index [32], flash point [32], ethanol content [26] and biodiesel [39]. The addition of solvents changes the properties related to the gasoline volatility and also distillation temperatures can change significantly for them to be detected by this assay [5,8,18,23,28,37,38].

The assay is carried out according to the ASTM D86 [40], which describes the distillation at atmospheric pressure of various petroleum products. The aim of this assay is to determine volatility characteristics verifying if the light and heavy properties of the fuel being produced are adequate and also to detect the contamination of other products. The ANP establishes that Brazilian fuel must have maximum temperature values of 10%, 50% and 90% of the recovered volume, as is for the final boiling point and residual volume [40]. The great advantage of this tool is the possibility to use the results from routine analysis reducing the number of assays, as well as, eliminating the demand for sample pre treatment. Therefore, distillation curves have become quite a useful tool for quality control of automotive fuel.

Gasoline is the second most consumed fuel in Brazil while diesel is the most consumed [2], approximately 39.7 billion litres in 2012. Hence, there is a great demand from society for high quality gasoline, which requires the development of methods for the detection of possible contaminants in fuel to support monitoring and surveillance programs.

In this work, distillation curves, a routine procedure in fuel analysis, combined with chemometric techniques of classification PCA and PLS-DA were employed to identify the adulteration of gasoline with solvents sold at fuel stations. Furthermore, multivariate calibration models were built using PLS in the prediction of the solvents added to automotive gasoline.

2. Experimental

2.1. Material

This study used 150 samples from common gasoline C collected at fuelling stations located in the eastern region of the State of Minas Gerais, Brazil, through the Fuel Quality Monitoring Program (PMQC-ANP). The origin of the samples was determined from the data provided in the invoices from the fuel stations at the moment of collecting the samples, since they come from different refineries from other states. The samples were stored in polyethylene flasks, sealed, and kept at 8 to 15 °C temperature range.

Four groups of doped gasoline samples were prepared using gasoline A (mixture of hydrocarbons without ethanol from the Gabriel Passos Refinery – REGAP – Petrobras) and 5 to 40% (v/v) of each one of the solvents, in 1% (v/v) steps, more the 45 and 50% (v/v) ones. The solvents used were kerosene, turpentine, thinner and rubber solvent (a mixture of aliphatic and aromatic hydrocarbons from C6 to C8). The ethanol concentration was kept at 25% (v/v) in all samples maintaining the same content of gasoline sold at fuel stations. Petrobras provided all solvents as well as the gasoline, except thinner, which was acquired through retail (Dissolminas 3500). These samples were used to identify the types of solvents used in automotive gasoline adulteration through statistical tools i.e. PCA and PLS-DA and also the PLS models built for quantification.

All samples were previously analysed through several physicochemical parameters established by the ANP [2] such as distillation temperatures equivalent to 10%, 50% and 90% of the recovered volume, final boiling point, residue volume (ASTM D86) [40], specific gravity (ASTM D4052) [41], octane numbers (MON and IAD) (correlated with ASTM D2699 and D2700) [42,43], contents of benzene (%v/v) (ASTM D6277) [44], anhydrous ethanol (NBR 13992) [45] and hydrocarbons (saturated, olefins and aromatics, correlated with ASTM D1319) [46]. An automatic distiller was used for the analysis (Herzog HDMA 627), a densimeter (Anton Paar 4500) and an automatic gasoline analyser (Petrospec GS1000) based in infrared spectrometry and multivariate analysis. From these results the samples were classified as in-conformity (in agreement with the specification) and nonconformity (not in agreement to at least one of the specifications).

Besides the aforementioned assays the analysis of the markers of solvents was also carried out. The detection of the solvents' markers carried out by the ANP was done by submitting the samples to chromatographic analysis. The marker is a product developed exclusively for the detection of adulterations and due to legal issues it is not possible to publish information on the marker neither standards required for the analysis nor the methodology used. However, it is possible to state that the analysis takes 20 min, besides having a high cost it can only be carried out by ANP authorized labs.

After obtaining the results of analysis using the markers mentioned above, the samples were classified into adulterated (having a positive result in the presence of the solvent marker) and unadulterated (having a negative result in the presence of the solvent marker).

2.2. Distillation

For the distillation test, 100 mL of previously cooled gasoline was 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 [40]. The distilled vapour 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 mmHg and considering volume loss, according to ASTM D86 [40].

2.3. Chemometric analysis

The distillation curves of 302 samples – 150 samples of retail gasoline, 75 adulterated and 75 non-adulterated, and 152 samples of tampered gasoline with four different solvents – were arranged into data matrixes where each line of the matrix constitutes a sample and the rows the variables.

The first model was built with 150 retail samples, from different origins, employing PCA and PLS-DA to distinguish the adulterated samples from the unadulterated samples. Later two PCA and PLS-DA models were built using 40 adulterated samples and 40 unadulterated samples (from the first model and only from REGAP refinery) and 152 samples tampered with solvents. The origin of the fuel was an important factor in this stage of the process since samples from different refineries present distinct compositions. Therefore, a doped sample from REGAP could present similarities to another sample from a different origin, which would lead to an erroneous classification.

After the PCA analysis, which had the aim to extract maximum data from all the set visually, models were built using the multivariate method for the PLS-DA classification in order to carry out a statistical data analysis.

In the first PLS-DA model 100 samples were used for the validation set. The second model used 138 samples in the calibration set and 54 in the validation set. Both sub sets (calibration and validation) were selected using the *Kennard-Stone* algorithm [47]. To validate the PLS-DA model the “leave-one-out” cross validation was used and also the external validation. The criterion used to select the number of latent variables was based on the lowest *RMSECV* (Root Mean Square Error Cross Validation) value. The sensitivity and specificity parameters were determined for each one of the classes, for both calibration and validation in the evaluation of the models' quality.

PLS was employed in the prediction of content of solvents added to samples doped with solvents in concentrations of 5 to 50% (v/v). In this work in order to determine the content of solvents of the gasoline samples, four PLS multivariate calibration models, and one for each solvent was built using the same samples as the aforementioned PCA model. In each case, 1/3 of the samples of the validation set were used and the selection of these samples was carried out also using the *Kennard-Stone* algorithm. The number of latent variables used to build each model was determined through cross-validation.

Auto scaling was the preprocessing method used for all models. The models were evaluated by accessing the *RMSEC* (Root mean square error of calibration) value, the correlation between the real and predicted values for the concentrations (*R*), *RMSEP* (Root Mean

Square Error of Prediction), *LV* (Latent Variable) and *EV%* (explained variance percentage).

The calculations were performed using the Minitab Release 14 Windows version) and Matlab (version 7.9) softwares.

3. Results and discussion

3.1. Segregation of adulterated and unadulterated samples

The distillation curves of adulterated and unadulterated samples can present similarities so that a simple visual comparison would not be able to distinguish them, hence the use of chemometric methods are necessary. In this study the interval of the distillation curve ranged from 4 to 93% (v/v), due to low reproducibility outside this range.

The distillation curves of 150 samples from PMQC-ANP (75 adulterated and 75 unadulterated) were obtained and put into an **X** matrix data with the samples displayed in lines and the equivalent temperatures of the recuperated percentages in columns. The distillation curves of these samples were used for the PCA chemometric treatment and the preprocessing used for the treatment of data was auto scaling.

The results obtained with PCA showed that the three first principal components explained 86.3% of the original information. Fig. 1 shows the scores graph obtained for the first two principal components (PC1 e PC2), which displays the formation of two groups, one formed by a more dense group of unadulterated samples and the other by a more disperse group of samples of various behaviours classified as adulterated by the solvent marker assay. PC1 described 63.6% of the total variance, PC2 15.7% and PC3 7.0%.

Fig. 2 shows the loadings graph of the variables (percentage of recuperated volume) associated with the first principal component. In PC1 the most important fractions for the separation of both sets of adulterated samples and unadulterated samples correspond to 4 to 40% (v/v). This is related to the increase of paraffin and isoparaffin content in adulterated gasoline that forms azeotropes with ethanol, causing alterations in the initial range of the distillation curves (Fig. 6).

Ethanol when mixed with gasoline forms minimum azeotropes with hydrocarbons, which have lower boiling point than both components separately [48]. These mixtures exhibit high deviations from ideal mixtures defined by Raoult's Law. During distillation a mixture of gasoline (hydrocarbons) and ethanol, boils and distills constantly until it reaches its azeotropic temperature, which

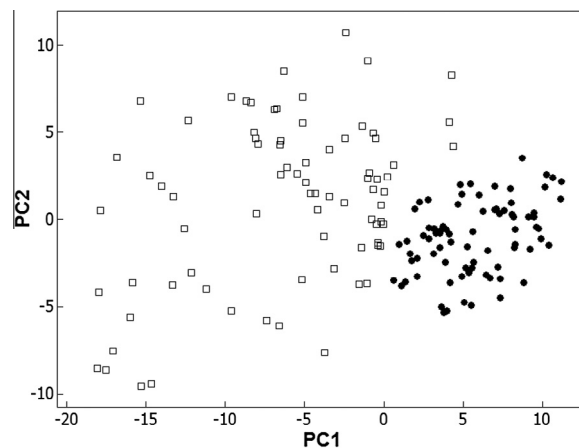


Fig. 1. Score graph for the first two principal components obtained in the discrimination of gasoline C samples from PMQC-ANP. (□) adulterated samples; (●) unadulterated samples.

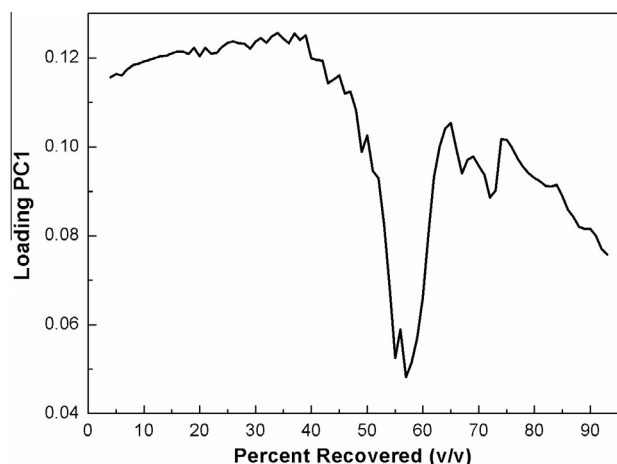


Fig. 2. Loading graph for PC1 obtained in the discrimination of adulterated and unadulterated gasoline C samples from PMQC-ANP.

depends on the mixtures' chemical composition. At this stage the temperature remains constant until all the azeotropic mixture is distilled so then the rest of the hydrocarbon mixture is free from ethanol, and "boils" at its respective individual boiling points continuously. This phenomenon creates a region of transition between the azeotropic mixture and the lasting hydrocarbon mixture easily identified by a sudden rise of temperature during distillation [49].

According to Pasquini and Scafi [50], until the 40% (v/v) fraction of gasoline C distillation, there is an increase in content of paraffins and isoparaffins, the fraction of 60% (v/v) at its maximum. The addition of aliphatic solvents, the most used type in adulterations in Brazil [6], produce the dilution of olefins and aromatics due to the increase of paraffins and isoparaffins contents in the mixtures [51]. Taking this into consideration the paraffins and isoparaffins are the main hydrocarbons responsible for the segregation studied. The increase of these hydrocarbons in adulterated samples to the fraction of 40% (v/v), produces a rapid formation of azeotropes among the hydrocarbons and ethanol causing the latter to produce azeotropes in smaller fractions compared to unadulterated gasoline, therefore changing the profile of the distillation (Fig. 6).

The results obtained from the physico-chemical parameters assays according to the ANP [2] for adulterated samples and unadulterated samples confirm this tendency. According to Table 1 it is possible to observe that adulterated samples present a greater dispersion of RON values, final boiling point and temperatures equivalent to 10% and 90% of the recovered volume. In addition, it is

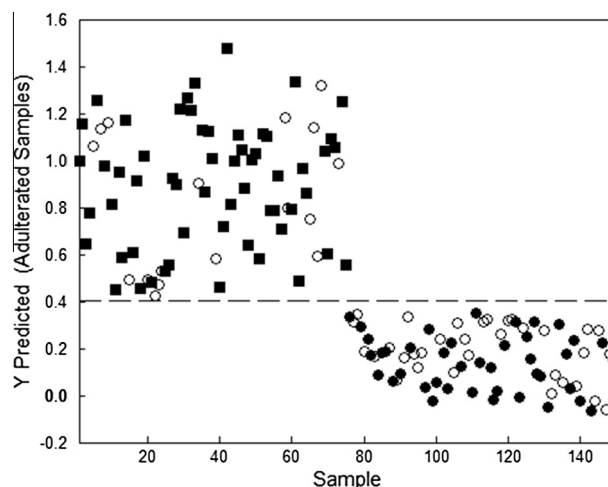


Fig. 3. PLS-DA classification of gasoline C samples from PMQC-ANP. (■) adulterated samples; (●) unadulterated samples; (○) respective validation samples.

possible to observe that adulterated samples present a tendency to raise the values of the temperature equivalent to 10% and the content of saturated hydrocarbons, in relation to unadulterated samples, followed by smaller olefin content. In the Brazilian South-east region the content of olefins in gasoline is high once the production process uses catalytic cracking to convert complex molecules into simpler compounds (light hydrocarbons) producing large quantities of these components. Therefore, the addition of solvents causes modifications in the original volatility characteristics of gasoline C [52].

The PCA analysis had the objective of exploring important variables in the segregations once in PLS-DA the results obtained using loadings provided graphs with distinct behaviours. This is because in PLS-DA the simultaneous regression amongst matrixes X and Y occurs and in PCA only the X matrix is decomposed. After performing PCA, PLS-DA was employed in the same sample set because it provides a statistical data analysis.

To carry out the PLS-DA model three latent variables were used and auto scaling as pre processing which yielded an accumulated variance of 85.0% in X a 76.6% in Y . Furthermore, the model did not present anomalous samples and the results obtained for both their classes (adulterated samples and unadulterated samples from PMQC) are shown in Fig. 3.

The dashed line in Fig. 3 indicates the threshold value between the predicted values. Values above this limit indicate that the

Table 1

Mean values and respective standard deviations of physicochemical parameters and their specifications according to ANP.

Method	Parameter	Specification	Adulterated samples	Unadulterated samples
ASTM D4052	Specific gravity	No specification	0.7560 ± 0.0040	0.7549 ± 0.0035
ASTM D6277	Benzene (% v/v)	1.0 (max.)	0.5 ± 0.2	0.7 ± 0.1
NBR 13992	Ethanol (% v/v)	25 ± 1	24.6 ± 1.6	24.0 ± 2.3
ASTM D86	10% recovered (°C)	65.0 (max.)	59.3 ± 3.9	53.6 ± 1.5
	50% recovered (°C)	80.0 (max.)	72.8 ± 0.8	72.3 ± 0.6
	90% recovered (°C)	190.0 (max.)	170.0 ± 10.2	166.8 ± 4.8
	Final boiling point (°C)	220.0 (max.)	211.3 ± 8.6	209.3 ± 5.9
	Residue (mL)	2.0 (max.)	1.1 ± 0.2	1.1 ± 0.2
Automatic	MON	82.0 (min.)	81.6 ± 1.1	82.5 ± 1.2
	RON	No specification	95.2 ± 2.6	97.2 ± 1.0
Gasoline Analyser	AKI	87.0 (min.)	88.4 ± 1.5	89.8 ± 0.7
	Aromatics (% v/v)	38.0 (max.)	15.5 ± 4.9	17.6 ± 4.4
	Olefins (% v/v)	45.0 (max.)	13.8 ± 5.9	22.9 ± 4.8
	Saturated hydrocarbons (% v/v)	No specification	40.2 ± 9.9	30.2 ± 8.6

Table 2
Sensitivity and specificity for the PLS-DA model (3VL).

Parameter	Class 1 Adulterated samples	Class 2 Unadulterated samples
Sensitivity (calibration)	1.000	1.000
Specificity (calibration)	1.000	1.000
Sensitivity (prediction)	1.000	1.000
Specificity (prediction)	1.000	1.000
Classification error (calibration)	0	0
Classification error (prediction)	0	0

sample belongs to the modeled class and predicted values, below this threshold indicate that the sample does not belong to the modeled class. This line is calculated by the algorithm and is estimated from the calibration samples and the prediction values by PLS-DA. If a sample belonging to a certain class is present at the bottom line, it has been classified incorrectly.

In Fig. 3 the graphs indicate both classes classified by the PLS-DA for both calibration and validation sets. Table 2 shows the sensitivity and specificity parameters by the PLS-DA model. The results indicate 100% accuracy in the prediction of adulterated and unadulterated samples from PMQC-ANP.

These results confirmed that the use of distillation curves associated with the PLS-DA were able to discriminate adulterated and unadulterated samples of the monitoring program with 100% classification both for calibration and validation.

3.2. Models using adulterated and unadulterated samples and doped with solvents

This model used 40 adulterated and 40 unadulterated samples collected by PMQC-ANP, and more 152 samples doped with different solvents. Initially the set of data from the distillation curves of the unadulterated gasoline samples and adulterated samples with the four solvents in concentrations of 5 to 50% (v/v) were submitted to principal component analysis.

The results obtained by the PCA analysis showed the formation of four groups of adulterated samples, with kerosene, turpentine, thinner and rubber solvent, as well as, another two sets of adulterated and unadulterated samples (Fig. 4). A few samples from the PMQC-ANP grouped toward one or more sets due to the possible mixture of the solvents used. Turpentine and kerosene presented in their composition heavy aliphatic hydrocarbons (C13 to C15) and (C9 to C16), respectively, showing certain similarities compared to the other groups (thinner and rubber solvent), given that turpentine is one of the kerosene fractions. In rubber solvent there is the presence of light aliphatic hydrocarbons (C6 to C8) with saturated content in quantities larger than the aforementioned solvents and also the presence of olefins. And in thinner there is a high content of aromatic hydrocarbons as well as other components such as acetates and alcohols [6].

The first four principal components explained 95.4% of the original information. PC1 described 75.7%, PC3 13.9, PC2 3.5% and PC1 2.2%. Fig. 4 shows PC1 separated doped samples with thinner and rubber solvent from unadulterated samples. Another separation in PC1 occurred between samples doped with rubber solvent and unadulterated samples. While PC2 distinguished samples doped with kerosene from samples doped with turpentine. The loadings graph (Fig. 5) indicates that the essential region of the distillation curve for the separation of samples ranges from the intervals of 4 to 35% (v/v), the aforementioned model likewise, 50 to 65% (v/v). For PC2 the intervals of 4 to 35% (v/v) and 70 to 90% (v/v) were the most important.

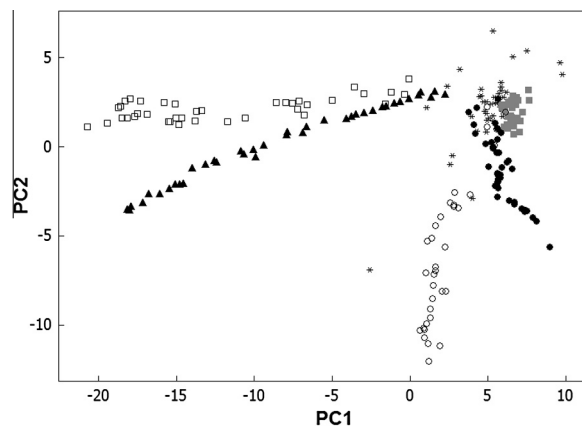


Fig. 4. Score plot for PC1 and PC2 obtained in the discrimination of gasoline C samples doped with different solvents. (□) kerosene; (▲) turpentine; (○) thinner; (●) rubber solvent; and also (*) adulterated samples and (■) unadulterated samples from PMQC-ANP.

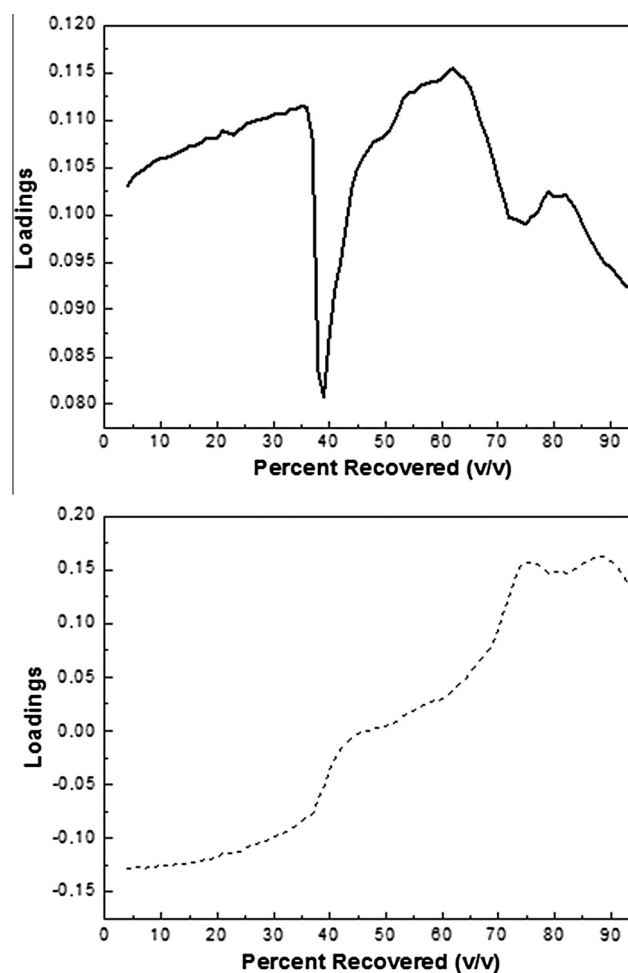


Fig. 5. Loading graphs for the first (—) and second (---) principal components obtained in the discrimination of gasoline C samples doped with different solvents.

The importance of these variances is related to the boiling temperature alterations caused by the addition of different solvents to gasoline previously discussed. Fig. 6 shows the variations produced by the additions in a gasoline C sample doped with different solvent in concentrations of 0, 10, 20, 30, 40 and 50% (v/v), separately.

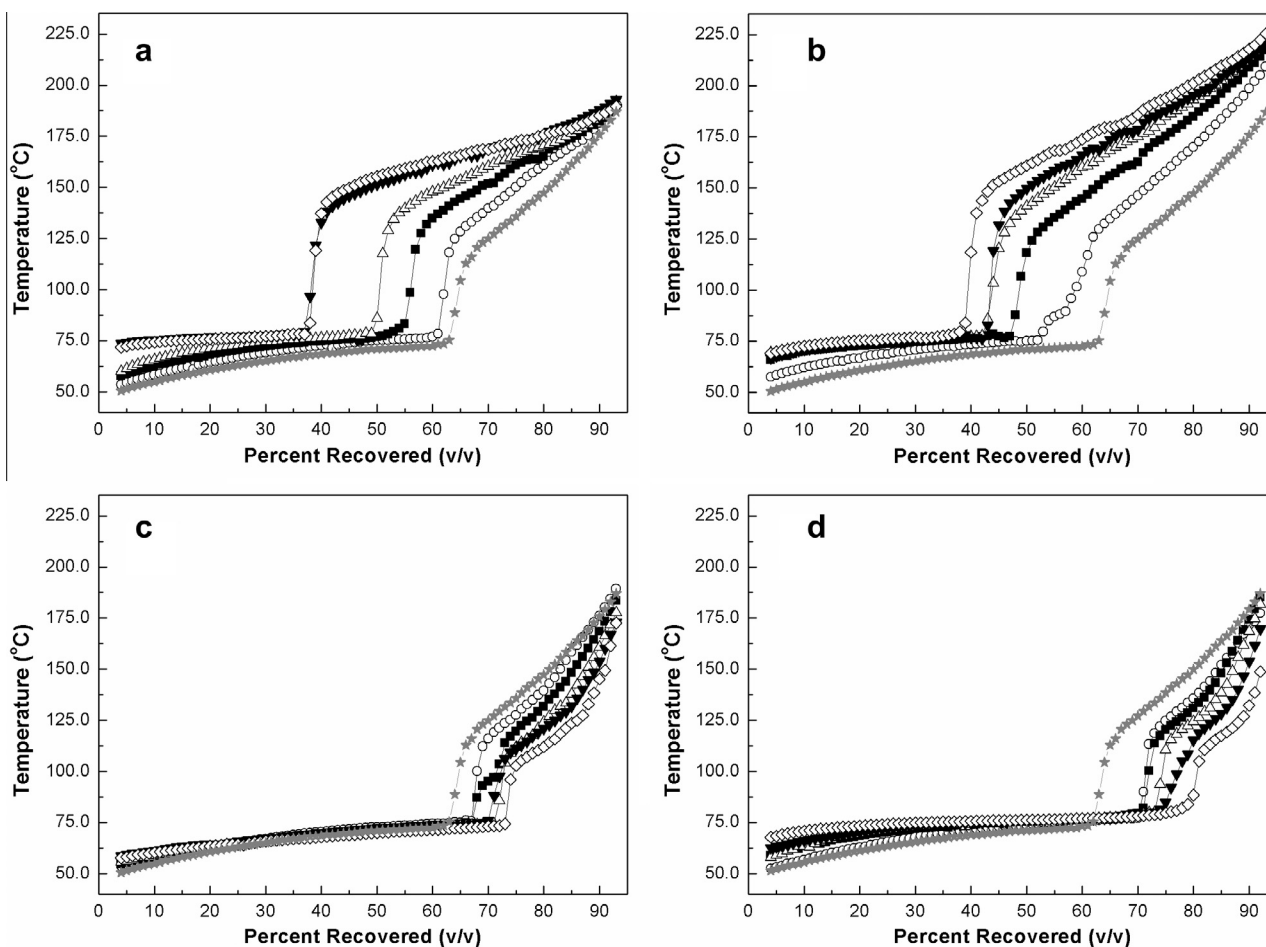


Fig. 6. Distillation curves obtained for different solvent contents in gasoline C samples. (★) 0% (v/v); (○) 10% (v/v); (■) 20% (v/v); (△) 30% (v/v); (▼) 40% (v/v); (◇) 50% (v/v). (a) turpentine; (b) kerosene; (c) rubber solvent and (d) thinner.

It was possible to observe that solvents such as turpentine and kerosene, when compared to others, lower the distillation temperature of gasoline C. The increase of the solvent content causes a sudden increase of temperatures in smaller fractions. The addition of turpentine produces severe variation in the distillation curve in fraction ranging from 4 to 40% (v/v) starting from this point onwards the increase of temperature for solvents with higher concentrations. Between 45 and 60% (v/v) abrupt increases in temperature occur in these solvents that take place in distinct intervals with rubber solvent and thinner, explaining the importance of the variables of the loadings graph (Fig. 5).

In the distinction between gasoline C and rubber solvent the fraction ranging from 4 to 40% (v/v) are responsible for the segregation, while the turpentine and kerosene solvents the distinction was made by PC2 in the fraction ranging from 65 to 90% (v/v). The distillation curves in Fig. 6 show that the boiling temperatures of kerosene are higher in the final fractions, possibly due to the presence of hydrocarbons C13 to C16.

Thinner and rubber solvents ranging from 4 to 40% (v/v) present higher importance in the distinction among them. In this interval thinner addition cause the boiling temperature to lower, contrarily, the rubber solvent presents in its composition considerable concentration of ethanol, which form minimum azeotropes in this particular interval [25].

The results indicated that the adulterated samples with thinner and rubber solvent presented a tendency of grouping with unadulterated samples in lower concentrations of solvent. Fig. 4 shows that the adulteration with thinner (aromatic solvent) and with rubber solvent (light aliphatic solvent) seems more difficult to detect,

than the addition of kerosene and turpentine (heavy aliphatic solvents) using distillation curves for the detection.

As previously carried out the PLS-DA was also employed to same samples used in PCA except the adulterated samples from PMQC-ANP. Six latent variables were used to build the PLS-DA model as for the preprocessing auto scaling was employed. This yielded an accumulated variance of 98% in **X** and 75% in **Y**, not displaying any anomalous samples. The graphs of Fig. 7 indicate the samples fitting to each class they were classified into by PLS-DA.

Table 3 shows the sensitivity and sensibility parameters obtained by the PLS-DA model. The results reveal 97% accuracy in the prediction of doped samples with turpentine and rubber solvent, and for the rest of the doped samples and the unadulterated samples the percentage of accuracy was of 100%. The classification and validation errors for all the set of samples were quite small, lower than 0.06%. These results were better than the ones obtained by Pereira et al. [6] using LDA and infrared spectra with the efficiency of 93% in the identification of the type of solvent added to gasoline.

The individual models for each class with different concentrations and types of solvents were developed yielding good results before the results of PCA and PLS-DA were obtained, providing the classification of adulterated and unadulterated samples and the types of solvents used.

3.3. Quantification of solvents in gasoline samples

The **X** matrix (distillation curve) and the **y** vector (content of solvents) were used to build the multivariate calibration model,

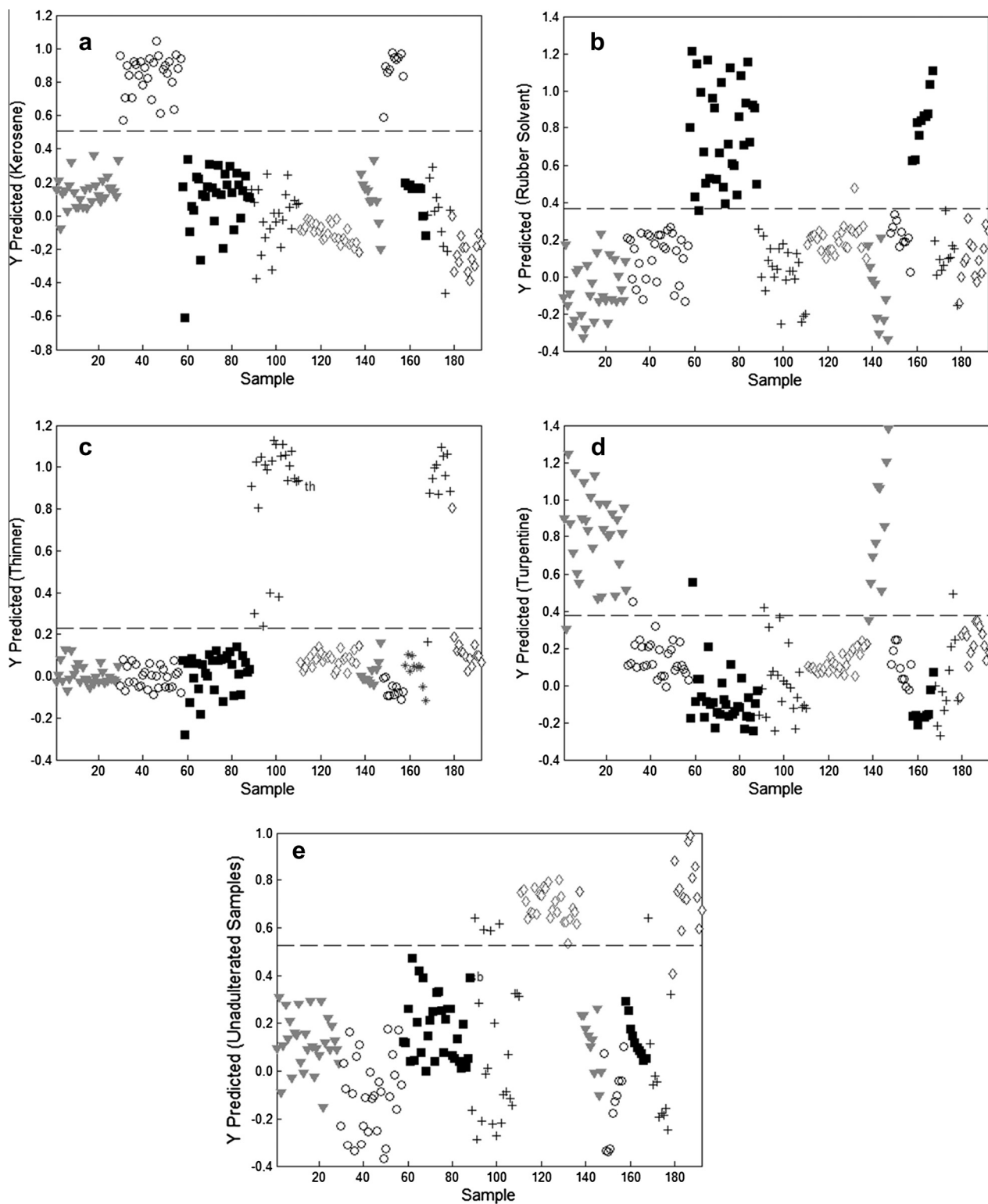


Fig. 7. PLS-DA classification of gasoline C samples doped with different solvents. (a) (○) kerosene; (b) (■) rubber solvent; (c) (+) thinner; (d) (▼) turpentine and (e) (◇) unadulterated samples.

with the number of latent variables accordingly defined by the lowest value of *PRESS* (Predictive Residual Error Sum of Squares) [53]. Table 4 shows the number of latent variables used in each model and the explained variance in *y*.

The autoscaling process was employed to build the chemometric models, it provided lower values of *RMSEC* (Root Mean Square

Error of Calibration) and *RMSEP* (Root Mean Square Error Validation), when compared to the ones obtained using original data and other tested pre processing methods. This is because in this pretreatment equal weight is given to the variables.

The working range of the curves varied from 5 to 50% (v/v) and the minimum concentrations that the methodology can quantify

Table 3
Sensitivity and specificity for the PLS-DA model (6VL).

Parameter	Class 1 T	Class 2 K	Class 3 RS	Class 4 TH	Class 5 UN
Sensitivity (calibration)	0.966	1.000	0.968	1.000	1.000
Specificity (calibration)	0.972	1.000	0.991	1.000	0.966
Sensitivity (prediction)	0.966	1.000	0.968	0.955	0.963
Specificity (prediction)	0.963	1.000	0.991	0.991	0.964
Classification error (calibration)	0.031	0	0.020	0	0.018
Classification error (prediction)	0.061	0	0	0.041	0.011

K – kerosene, T – turpentine, RS – rubber solvent, TH – thinner, UN – unadulterated samples.

Table 4
Parameters of models PLS for quantification of solvents in gasoline C.

Solvent	RMSEC (% v/v)	RMSEP (% v/v)	NVL	R ²	VE (%)
Turpentine	0.53	1.1	3	0.99	99.8
Kerosene	0.66	0.66	3	0.99	99.9
Rubber solvent	0.85	1.4	3	0.99	99.0
Thinner	1.1	1.7	3	0.99	99.8

was 6% (v/v) – for thinner, turpentine and kerosene – and 7% (v/v) for rubber solvent.

The accuracy evaluation of the proposed method was carried out by calculating the *RMSEP*. Table 4 shows that the *RMSEP* values were lower than those reported in the literature [10] using infrared. From these results it can be stated that the combination of distillation curves and multivariate calibration provided excellent results in predicting the levels of solvent added to gasoline, since low values of *RMSEP* were observed compared with models obtained from spectrometric techniques.

To assess the adjustment of the data, the concentration of each solvent provided by the proposed method were compared with reference values. The four models showed a high correlation between real and predicted values, i.e., we obtained a large set of real and predicted values for model calibration (R greater than 0.99), the largest one obtained in predicting the levels of kerosene.

4. Conclusions

This study showed that the use of distillation curves associated with PCA and PLS-DA yielded the discrimination of adulterated and unadulterated samples from the PMQC-ANP, as well as, the distinction of samples doped with unadulterated thinner, rubber solvent, turpentine and kerosene. The lowest percentage of accuracy obtained by the PLS-DA method was 97% for all the models built.

The use of PLS provided the prediction of content of contaminants in the range 5 to 50% (v/v) with low *RMSEC* and *RMSEP* values when compared to other methods. The great advantage of this work in relation to others in the literature using other analytical techniques is that it is possible to apply the results of the distillation curves in routine analysis, with no need for further testing.

One can also conclude that the proposed methodology can be implemented to improve the analytical process, as it is possible to predict which samples show adulteration, and the marker test can be used only to reassure the results. This study will provide great benefits in assisting in the process of supervision and monitoring the quality of automotive gasoline combined with its low cost.

Acknowledgements

The authors would like to thank FINEP, CAPES and CNPq for the financial support.

References

- [1] Moreira LS, d'Avila LA, Azevedo DA. Automotive gasoline quality analysis by gas chromatography: study of adulteration. *Chromatographia* 2003;58:501–5.
- [2] <<http://www.anp.gov.br>> – Portaria Agência Nacional do Petróleo n° 309, de 27.12.2001 – Agência Nacional do Petróleo e Biocombustíveis (ANP). (accessed in March 2013).
- [3] Wiedemann LSM, d'Avila LA, Azevedo DA. Adulteration detection of Brazilian gasoline samples by statistical analysis. *Fuel* 2005;84:467–73.
- [4] Kalligeros S, Zannikos F, Stournas S, Lois E. Fuel adulteration issues in Greece. *Energy Fuel* 2003;28:15–26.
- [5] Takeshita EV, Rezende RVP, Souza SMAGU, Souza AAU. Influence of solvent addition on the physicochemical properties of Brazilian gasoline. *Fuel* 2008;87:2168–77.
- [6] Pereira RCC, Skrobot VL, Castro EVR, Fortes ICP. Determination of gasoline adulteration by principal components analysis – linear discriminant analysis applied to FTIR spectra. *Energy Fuel* 2006;20:1097–102.
- [7] <<http://www.anp.gov.br>> – Relatório Mensal, Programa de Monitoramento da Qualidade dos combustíveis – Agência Nacional do Petróleo e Biocombustíveis (ANP). pdf. (accessed in January 2013).
- [8] Barbeira PJS. Using statistical tools to detect gasoline adulteration. *Engenharia Térmica* 2002;1:48–50.
- [9] Wiedemann LSM, d'Avila LA, Azevedo DA. Brazilian gasoline quality: study of adulteration by statistical analysis and gas chromatography. *J Braz Chem Soc* 2005;16:139–46.
- [10] Teixeira LSG, Oliveira FS, Santos HC, Cordeiro PWL, Almeida SQ. Multivariate calibration in Fourier transform infrared spectrometry as a tool to detect adulterations in Brazilian gasoline. *Fuel* 2008;87:346–52.
- [11] Flumignan DL, Boralle N, Oliveira JE. Screening Brazilian commercial gasoline quality by hydrogen nuclear magnetic resonance spectroscopic fingerprintings and pattern – recognition multivariate chemometric analysis. *Talanta* 2010;82:99–105.
- [12] Al-Ghoutia MA, Al-Degsb YS, Amera M. Determination of motor gasoline adulteration using FTIR spectroscopy and multivariate calibration. *Talanta* 2008;76:1105–12.
- [13] Balabin RM, Safieva RZ, Lomakina EI. Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. *Chemometr Intell Lab Syst* 2007;88:183–8.
- [14] Balabin RM, Safieva RZ, Lomakina EI. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. *Chemometr Intell Lab Syst* 2008;93:58–62.
- [15] Balabin RM, Safieva RZ. Gasoline classification by source and type based on near infrared (NIR) spectroscopy data. *Fuel* 2008;87:1096–101.
- [16] Balabin RM, Safieva RZ, Lomakina EI. Gasoline classification using near infrared (NIR) spectroscopy data: comparison of multivariate techniques. *Anal Chim Acta* 2010;671:27–35.
- [17] Monteiro MR, Ambrozini ARP, Lião LM, Boffo EF, Tavares LA, Ferreira MMC, et al. Study of Brazilian gasoline quality using hydrogen nuclear magnetic resonance (¹H NMR) spectroscopy and chemometrics. *Energy Fuel* 2009;23:272–9.
- [18] Oliveira FS, Teixeira LSG, Araújo MCU, Korn M. Screening analysis to detect adulterations in Brazilian gasoline samples using distillation curves. *Fuel* 2004;83:917–23.
- [19] Bruno TJ. Improvements in the measurement of distillation curves. Part 1: A composition-explicit approach. *Ind Eng Chem Res* 2006;45:4371–80.
- [20] Bruno TJ, Smith BL. Improvements in the measurement of distillation curves – Part 2: Application to aerospace/aviation fuels RP-1 and S-8. *Ind Eng Chem Res* 2006;45:4381–8.
- [21] Bruno TJ, Smith BL. Enthalpy of combustion of fuels as a function of distillate cut: application of an advanced distillation curve method. *Energy Fuel* 2006;20:2109–16.
- [22] Bruno TJ, Smith BL. Advanced distillation curve measurement with a model predictive temperature controller. *Int J Thermophys* 2006;27:1419–34.
- [23] Bruno TJ. Improvements in the measurement of distillation curves: Part 3: Application to gasoline and gasoline + methanol mixtures. *Ind Eng Chem Res* 2007;46:297–309.
- [24] Barbeira PJS, Pereira RCC, Corgozinho CNC. Identification of gasoline origin by physical and chemical properties and multivariate analysis. *Energy Fuel* 2007;21:2212–5.
- [25] Barbeira PJS, Aleme HG, Costa LM. Determination of gasoline origin by distillation curves and multivariate analysis. *Fuel* 2008;87:3664–8.
- [26] Barbeira PJS, Aleme HG, Costa LM. Determination of ethanol and specific gravity in gasoline by distillation curves and multivariate analysis. *Talanta* 2009;78:1422–8.
- [27] Barbeira PJS, Aleme HG, Corgozinho CNC. Diesel oil discrimination by origin and type using physicochemical properties and multivariate analysis. *Fuel* 2010;89:3151–6.
- [28] Tanaka GT, Ferreira FD, Silva CEF, Flumignan DL, Oliveira JE. Chemometrics in fuel science: demonstration of the feasibility of chemometrics analyses

- applied to physicochemical parameters to screen solvent tracers in Brazilian commercial gasoline. *J Chemometr* 2011;25:487–95.
- [29] Burger J, Lovestead TM, Windom BC, Bruno TJ. Characterization of renewable fuels and additives with the advanced distillation curve method. *Abstr Pap Am Chem Soc* 2011;242:145 [Fuel].
- [30] Barbeira PJS, Aleme HG, Mendes G. Determination of octane numbers in gasoline by distillation curves and partial least squares regression. *Fuel* 2012;97:131–6.
- [31] Barbeira PJS, Aleme HG, Assunção RA, Carvalho MMO. Determination of specific gravity and kinematic viscosity of diesel using distillation curves and multivariate calibration. *Fuel Process Technol* 2012;102:90–5.
- [32] Barbeira PJS, Aleme HG. Determination of flash point and cetane index in diesel using distillation curves and multivariate calibration. *Fuel* 2012;102:129–34.
- [33] Maheshwari AS, Chellani JG. Correlations for pour point and cloud point of middle and heavy distillates using density and distillation temperatures. *Fuel* 2012;98:55–60.
- [34] Burger JL, Bruno TJ. Application of the advanced distillation curve method to the variability of jet fuels. *Energy Fuel* 2012;26:3661–71.
- [35] Marinovic S, Bolanca T, Ukic S, Rukavina V, Jukic A. *Chem Tech Fuels Oil* 2012;48:67–74.
- [36] Albahri TA. Developing correlations for the properties of petroleum fuels and their fractions. *Fluid Phase Equilib* 2012;315:113–25.
- [37] Balakrishnan J, Balasubramanian V. A physicochemical study of the adulteration of motor gasoline with a mixture of aliphatic and aromatic hydrocarbons. *Pet Sci Technol* 2012;30:237–46.
- [38] Majhi A, Kukerti VS, Sharma YK, Khanna R, Datta A. The detection of kerosene as an adulterant in gasoline. *Pet Sci Technol* 2012;30:271–7.
- [39] Barbeira PJS, Aleme HG. Determination of biodiesel content in diesel using distillation curves and multivariate calibration. *Energy Fuel* 2012;26:5769–74.
- [40] ASTM – American Society for Testing and Materials – Standard test for distillation of petroleum products at atmospheric pressure. ASTM D86-11b; West Conshohocken PA; 2010.
- [41] ASTM – American Society for Testing and Materials – Standard test method for density, relative density, and API gravity of liquids by digital density meter. ASTM D4052-11; West Conshohocken PA; 2009.
- [42] ASTM – American Society for Testing and Materials – Standard test method for research octane number of spark-ignition engine fuel. ASTM D2699, West Conshohocken PA; 2009.
- [43] ASTM – American Society for Testing and Materials – Standard test method for motor octane number of spark-ignition engine fuel. ASTM D2700, West Conshohocken PA; 2009.
- [44] ASTM – American Society for Testing and Materials – Standard test method for determination of benzene in spark ignition engine fuels using mid infrared spectroscopy. ASTM D6277-07; West Conshohocken PA; 2009.
- [45] ABNT – Associação Brasileira de Norma Técnicas – Gasolina Automotiva – Determinação do teor de álcool etílico anidro combustível (AEAC). NBR-13992; Rio de Janeiro; 2008.
- [46] ASTM – American Society for Testing and Materials – Standard test method for hydrocarbon types in liquid petroleum products by fluorescent indicator adsorption. ASTM D1319-10; West Conshohocken PA; 2009.
- [47] Kennard RW, Stone LA. Computer aided design of experiments. *Technometrics* 1969;11:137–48.
- [48] Cataluña R, Silva R. Desenvolvimento de um equipamento para avaliação do efeito do etanol na pressão de vapor e entalpia de vaporização em gasolinas automotivas. *Quim Nova* 2006;29:580–5.
- [49] French R, Malone P. Phase equilibria of ethanol fuel blends. *Fluid Phase Equilib* 2005;228:27–40.
- [50] Pasquini C, Scafi SHF. Real-time monitoring of distillations by near-infrared spectroscopy. *Anal Chem* 2003;75:2270–5.
- [51] Lopes AR. Avaliação do efeito da adição de solventes sobre as propriedades da gasolina automotiva comercializada no estado do Paraná. Curitiba: Dissertação de Mestrado; 2004.
- [52] Flumignan DL, Anaia GC, Anaia GC, Tininis AG, de Ferreira FO. Screening Brazilian automotive gasoline quality through quantification of saturated hydrocarbons and anhydrous ethanol by gas chromatography and exploratory data analysis. *Chromatographia* 2007;65:617–23.
- [53] Hair Jr JF, Anderson RE, Tatham RL, Black WC. *Análise Multivariada de dados*. fifth ed. Porto Alegre: Bookman; 2005.