We are IntechOpen,
the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Multivariate Modeling in Quality Control of Viscosity in Fuel: An Application in Oil Industry

Leandro Valim de Freitas[1,2], Fernando Augusto Silva Marins[2],
Ana Paula Barbosa Rodrigues de Freitas[3],
Messias Borges Silva[2,3] and Carla Cristina Almeida Loures[2,3]
[1]Petróleo Brasileiro SA (PETROBRAS)
[2]São Paulo State University (UNESP)
[3]University of São Paulo (USP)
Brazil

1. Introduction

This chapter aims to present an alternative to quality control of the viscosity of two important fuels in the international scenery - aviation kerosene and diesel oil – by statistical multivariate modeling (Pasadakis et al., 2006).

Viscosity is one of the most important properties of fuels; it influences the circulation and the fuel injection in the operation of injection engines. Engines efficiency in the combustion process depends on this property.

Out of specification values can decrease the fuel volatilization, thus implying, in an incomplete combustion (Pontes et al., 2010). This physicochemical property can vary significantly with the modification of the cast during the processing of crude in a refinery (Figure 1), maintaining the same conditions of production control, which compromises the quality standards. This leads to the need to determine the viscosity or provide it as often as possible in lieu of performing the traditional point analysis in the laboratory that can take long time.

According to Dave et al. (2003), the use of field instruments in conjunction with statistical multivariate techniques to determine, in real-time, properties of the products is one way to optimize the operations of oil refining.

Each refinery has at least one primary distillation tower, where the components of crude oil are separated into different sections using different boiling points, and different arrangements of unit conversion. In general, the refining margin increases with the complexity of the refinery. Decisions about how to operate and monitor a refinery and how to build the units, are factors that provide competitive advantages to oil companies.

The hydrotreatment is a catalytic process that removes large amounts of sulfur and nitrogen from the distillation fractions (Fernández et al., 1995). The fluidized catalytic cracking (FCC) is a process widely used in a petroleum refining industry. It consists in cracking large molecules into smaller ones by high temperatures. Thus, heavy oils are converted into products with

higher added value (IEA, n.d.). In addition, some refineries have used the coking unit to maximize the refining margin in the conversion of waste from distillation towers.
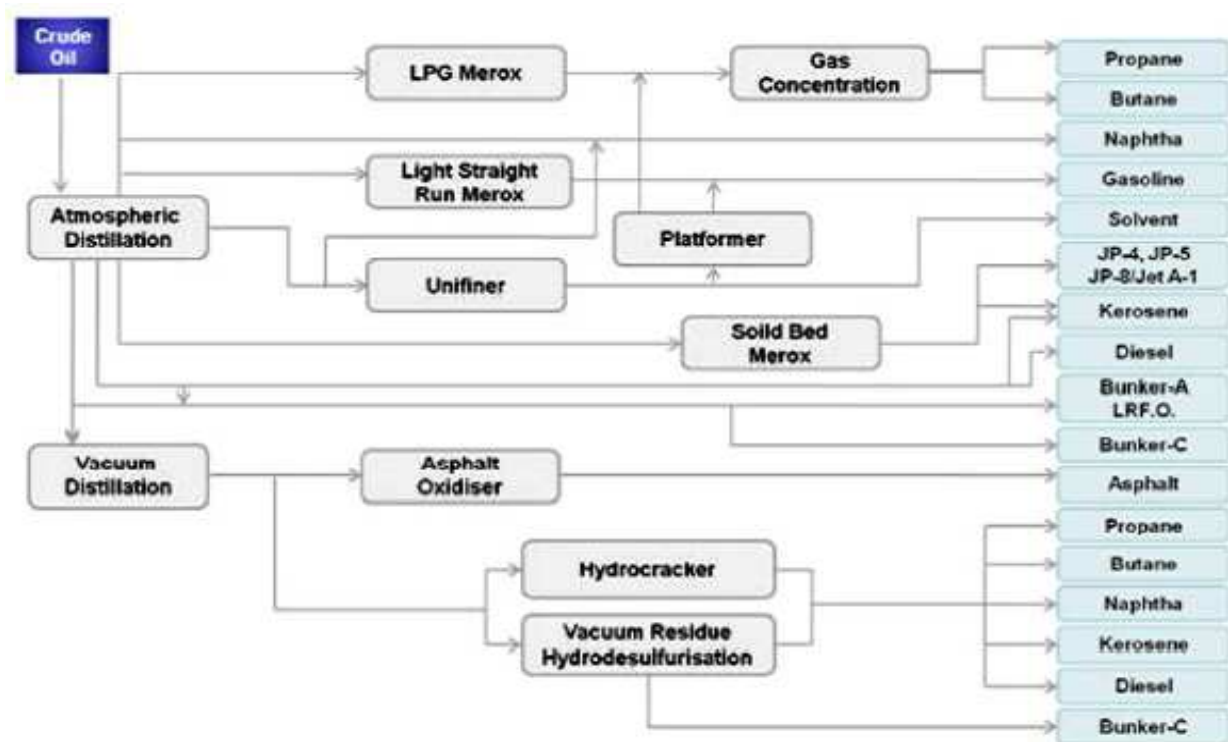

Fig. 1. Simplified process flow diagram of a refinery (Park et al., 2010)

Important products of refineries are the fuels for several kinds of existing engines. This way, fuel for jets is derived from petroleum and it is suitable for power generation by combustion in gas turbine engines for aircraft (see Fig.2). Jet fuel is produced by fractionation of petroleum by distillation at atmospheric pressure, with boiling range between 150 and 300 ° C, followed by finishes and treatments that aim primarily eliminate the undesirable effects of sulfur compounds, nitrogen and oxygen.

The viscosity of kerosene is limited to a maximum value to obtain a minimum loss of pressure in the flow at low temperature, as well as to allow using adequate spray nozzles for the fuel in order to improve the conditions of combustion. The viscosity property can significantly affect the lubricity of the fuel property, and, consequently, the life of the aircraft fuel pump.

The diesel fuel is a derivative of petroleum used in internal combustion engines compression to move motor vehicles (Fig. 3). It can also be used in marine engines and as a fuel for home heating. It is composed mainly of paraffinic hydrocarbons, and it is not desirable to the presence of olefins and aromatics. Its normal boiling range is 100 to 390 ° C, while the number of carbon atoms should be located between six and eighteen atoms.

The chemical composition of diesel oil directly affects its performance and is related to the type of oil used and with the adopted processes for their production in refineries.

Overall this product is composed of one or more cuts from the distillation of petroleum, and it can be added to other current refining processes, for example, the product obtained from catalytic cracking called light oil recycled.
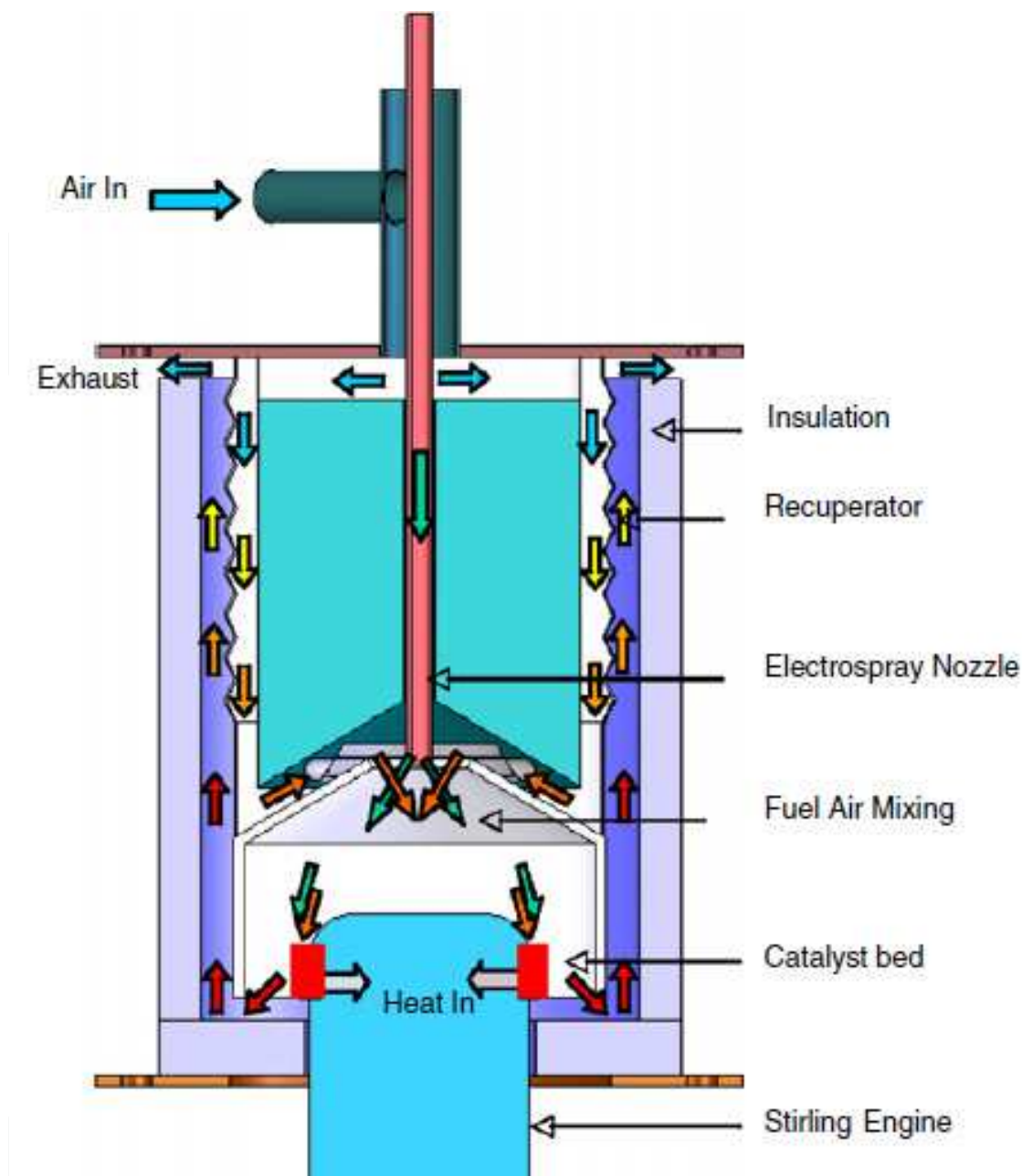
Fig. 2. Combustor and engine - schema (Gomez et al., 2007).

Kinematic viscosity of this product is an important property in terms of its effect on power systems and in fuel injection. Both high and low viscosities are undesirable since they can cause, among others, problems of fuel atomization. The formation of large and small droplets (low viscosity), can lead to a poor distribution of fuel and compromise the mixture air - fuel resulting in an incomplete combustion followed by loss of power and greater fuel consumption.

In Section 2, it will be presented a method for acquiring and construct the database. In Section 3, theoretical foundations of the statistical multivariate methods used. In Section 4, are presented the method application and the results for a real process of production. Finally, in Section 5, the conclusions reached are discussed.
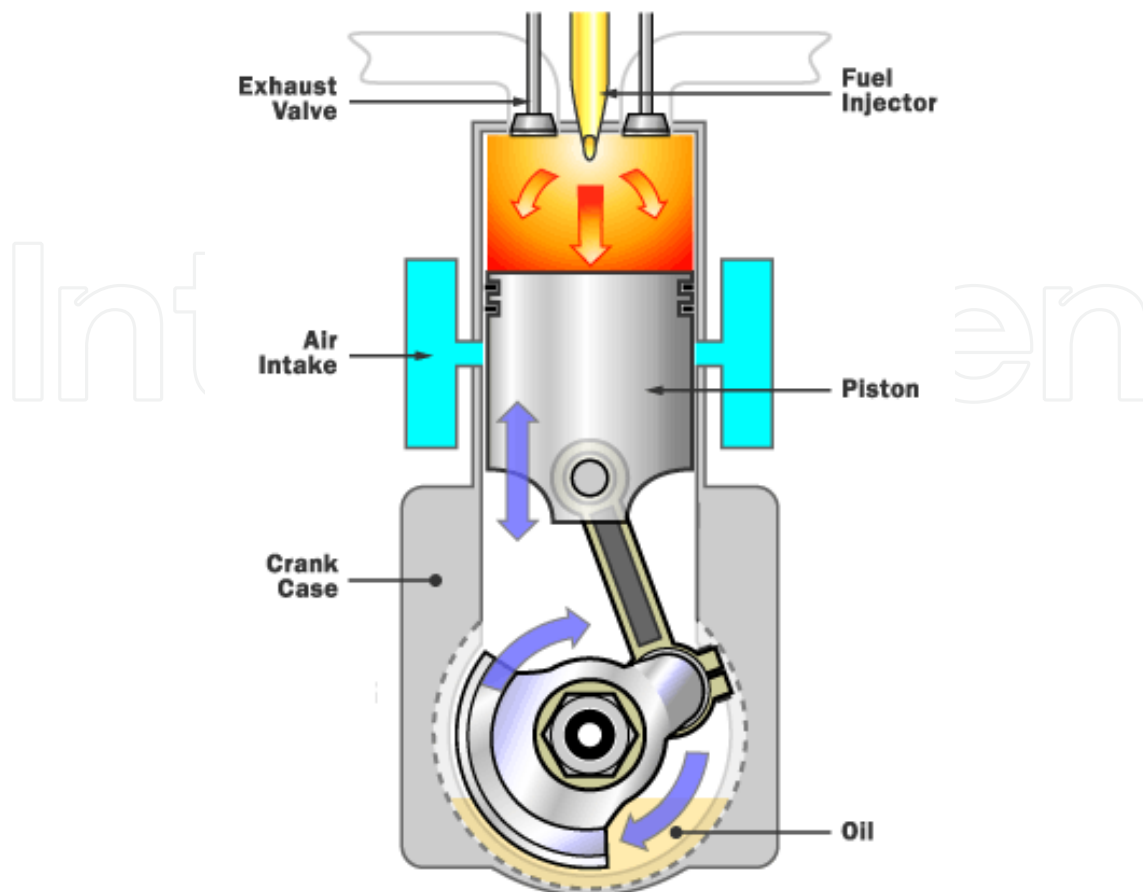
Fig. 3. Diesel engine (Marshall, 2002)

2. Acquisition data base

The multivariate analysis, using infrared signals, allow manipulate data absorption, called spectra, associated with more than one frequency or wavelength at the same time. In the oil industry, their applications are associated with the prediction of the quality of distillates such as naphtha, gasoline, diesel and kerosene (Kim, Cho, Park, 2000).

The chemical bonds of the type carbon-hydrogen (CH), oxygen-hydrogen (OH) and nitrogen-hydrogen (NH) are responsible for the absorption of infrared radiation, but are not very intense and overlap, creating broad spectral bands, that are correlated and difficult to interpret. However, a multivariate approach has proven quite adequate for modeling physical and chemical properties from samples of the input variables, it is known as absorbed infrared radiation (Behzadian et al., 2010).

The polychromatic radiation emitted by the source has a wavelength selected by a Michelson interferometer. The beam splitter has a refractive index such that approximately half of the radiation is directed to the fixed mirror and the other half is reflected, reaching the movable mirror and is therefore reflected by them. The optical path differences occur due to movement of the movable mirror that promotes wave interference. Nowadays, the instrumentation has introduced an improved Michaelson interferometer, to develop the system "Wishbone", as illustrated in Figure 4. In this system, instead of having a movable mirror and a fixed mirror, both mirrors cubic move, tied to a single support.
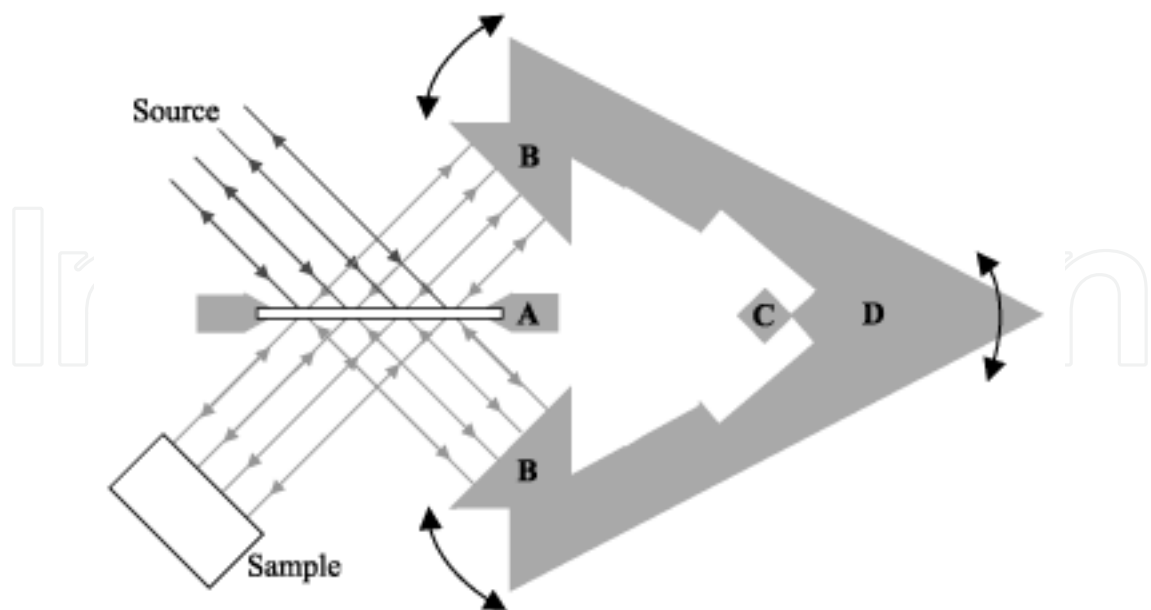
Fig. 4. "Wishbone" interferometric system employed in modern NIR spectrophotometers based on Fourier Transform. A, beam splitter; B, corner cubic mirrors; C, anchor, and D, "wishbone" (Pasquini, 2003)

An interferogram is obtained as a result of a graph of the signal intensity received by the detector versus the difference in optical path traveled by the beams. Then, like the Fourier Transform translates the recurring phenomenon in a series of sines and cosines (see Fig. 5), it is possible to transform the interferogram in a spectrum transmittance. The amount of radiation absorbed is determined by using the co-logarithm of the transmittance spectrum.
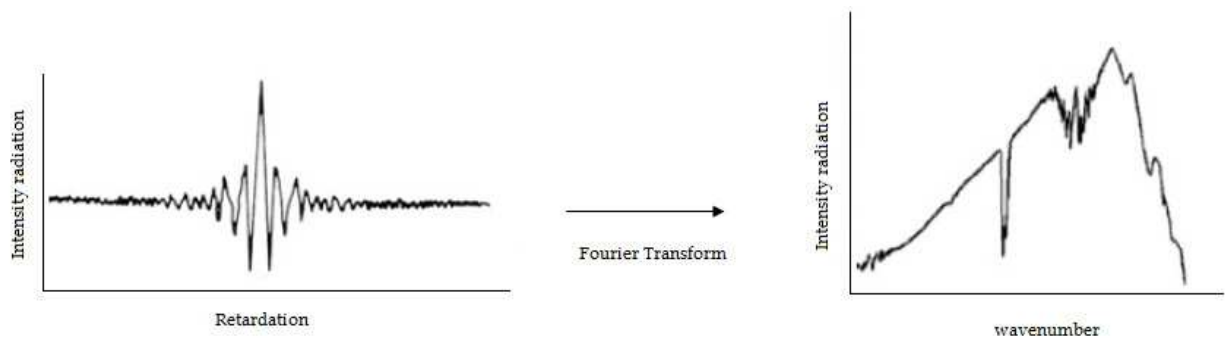
Fig. 5. Interferogram of radiation containing several wavelengths (adapted from Smith, 2011)

3. Multivariate analysis

The characterization of the mathematical models more adequate was performed by using the multivariate technique with partial least squares regression (PLS). This is an analysis technique where the original matrix of data is represented by factors or latent variables. Only the portion of the spectral data that correlates with the property assessed is included in this representation.

The first factor, calculated by a statistical program The Unscrambler®, has the highest correlation of spectral data with respect the property of interest. The residual spectrum,

which is the original spectrum minus the proportion represented by the first factor, then the same statistical program evaluates it. Thus, the second factor has the highest correlation with the residual spectrum property. This procedure is replicated until each one of the important information, which has a correlation with the property under study, was represented by the factors or latent variables. Observe that some caution is needed to determine the appropriate number of factors, because an insufficient number of them will not include all the necessary spectral information and too many of them will add noise (see Fig. 6).
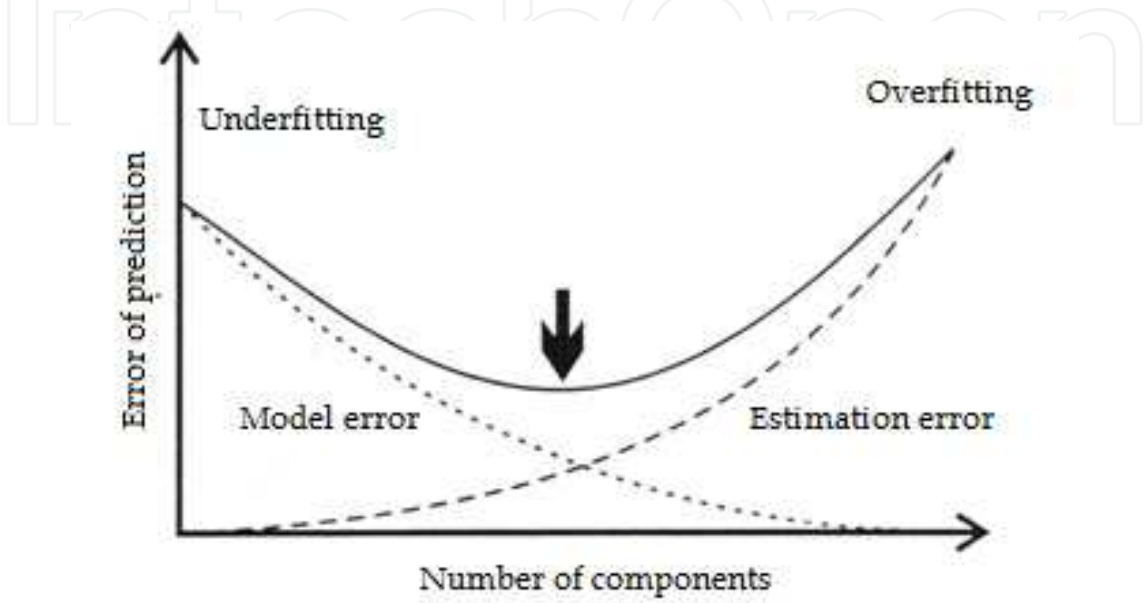
Fig. 6. Optimized number of components (Naes et al., 2002)

This regression model has the advantage of using the entire spectrum, be quick and offer a stable result. In addition, the regression using partial least squares, that use a number of factors less than the principal component regression (PCR) method, is more resistant to noise and in presence of weaker correlations.

The decomposition of the original variables on principal components can be represented by Equation (1), where for the k principal components, t is the value called score, that indicates the differences or similarities between samples, γ is the parameter that relates the original variables with the latent variable, it is called loading and represents how much a variable contributes to a major component and it considers the variation of the data:

$$t_{ik} = \upsilon_{k1}X_{i1} + \upsilon_{k2}X_{i2} + ... + \upsilon_{kq}X_{kq} \tag{1}$$

So the principal components are related to the concentration, or property of interest, according to Equation (2), where h is the number of principal components used. The Fig. 7 represents the Equation (2) in matrix form:

$$Y = b_0 + b_1 t_{i1} + b_2 t_{i2} + ... + b_h t_{ih} + e_1 \tag{2}$$

The Fig. 8 illustrates the absorbance of three wavelengths. Observe that the first principal component PC1 is a linear combination of the absorbance values that representing the maximum variance between samples. The projection of the sample point on the axis of PC1 is the score of PC1 (see Fig. 9).

Fig. 7. Matrix representation of mathematical model

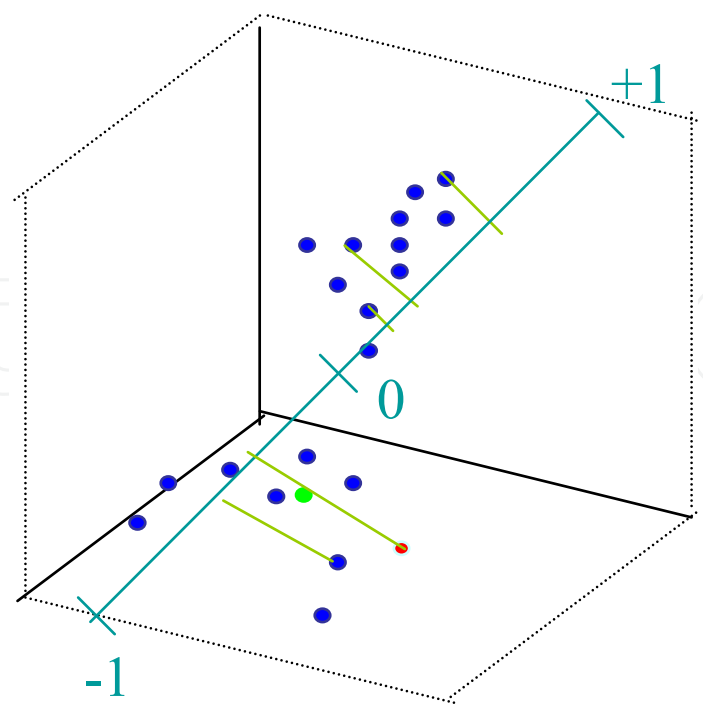

Fig. 8. Samples measured with three variables

Fig. 9. First principal component in 3D

The PC2 must be orthogonal to PC1 and is positioned to capture the maximum residual variance. When all the data variability can not be explained by only one major component (red and green samples), a second PC is needed and so on. The score for the PC2 is obtained by projection, in a manner analogous to the previous situation (Fig. 10).
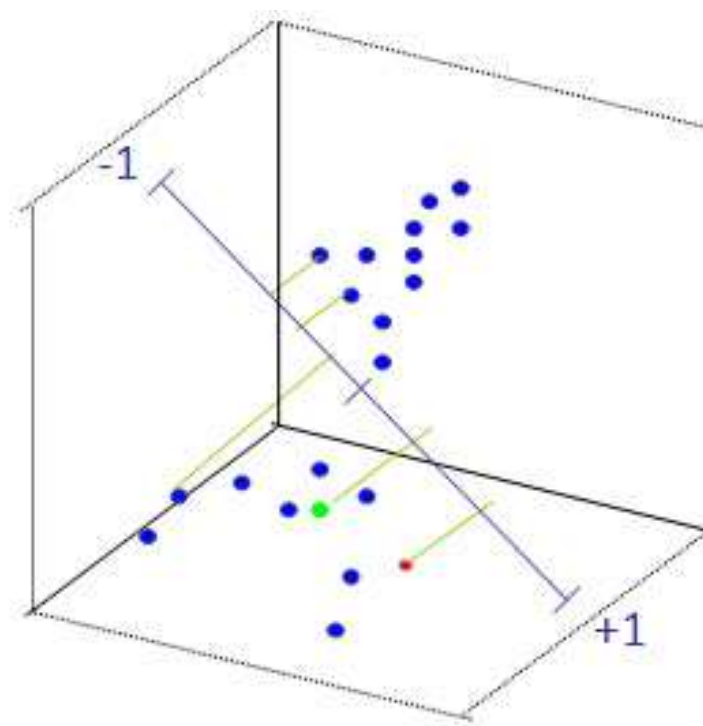

Fig. 10 . Second principal component in 3D

To obtain the remaining principal components, the procedure is the same. It be noted that for a set of 100 different wavelengths, for example, is not necessary 100 principal components to represent the data variability.

The variability of the spectrum can be compressed into a less number of principal components without significant loss of analytical information. After this compression, the scores are considered as independent variables in the regression to obtain the dependent variable y (concentration or physicochemical properties).

The PLS determines the principal components that are the best ones with respect to the variable y and that explain as best as possible the variable x (see Fig. 11).
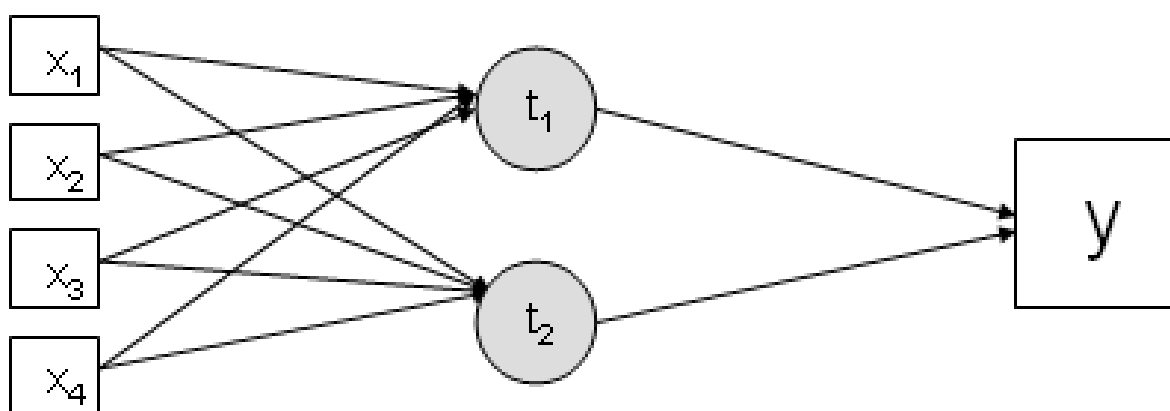

Fig. 11. Data compression in principal components

4. Proposed method, results and analysis

The first step is collect samples of kerosene and diesel oil in a reasonable period of time, to obtain a data set that best reflect all possible operational variability, as changes in the cast of oil and operating conditions of the units.

In the second stage experiments are performed to characterize, on a laboratory scale, the product, aiming to determining, from the samples, the real kinematic viscosity to be modeled. The samples were also subjected to infrared radiation.

In the third step the mathematical models were developed using The Unscrambler® and Excel® softwares, associating the information to absorbed infrared radiation with the physicochemical property. In the end, the model is implemented on an industrial scale for forecasting the viscosity in real time, providing to the production area, high power decision-making, and enabling increase the profitability of the blending process.

For each oil studied was developed a mathematical model with 800 input variables. To help determine the number of latent variables and minimize the residual variance was used the full cross-validation method, which is a mathematical algorithm able to gradually reduce the number of samples. In the sequence, a model constructed from the remaining samples is tested by comparing it with the true values of the samples excluded. The models are developed using The Unscrambler® program. Several forms of preprocessing were

evaluated to obtain the minimum value of RMSECV (Root Mean Square Error of Cross Validation).
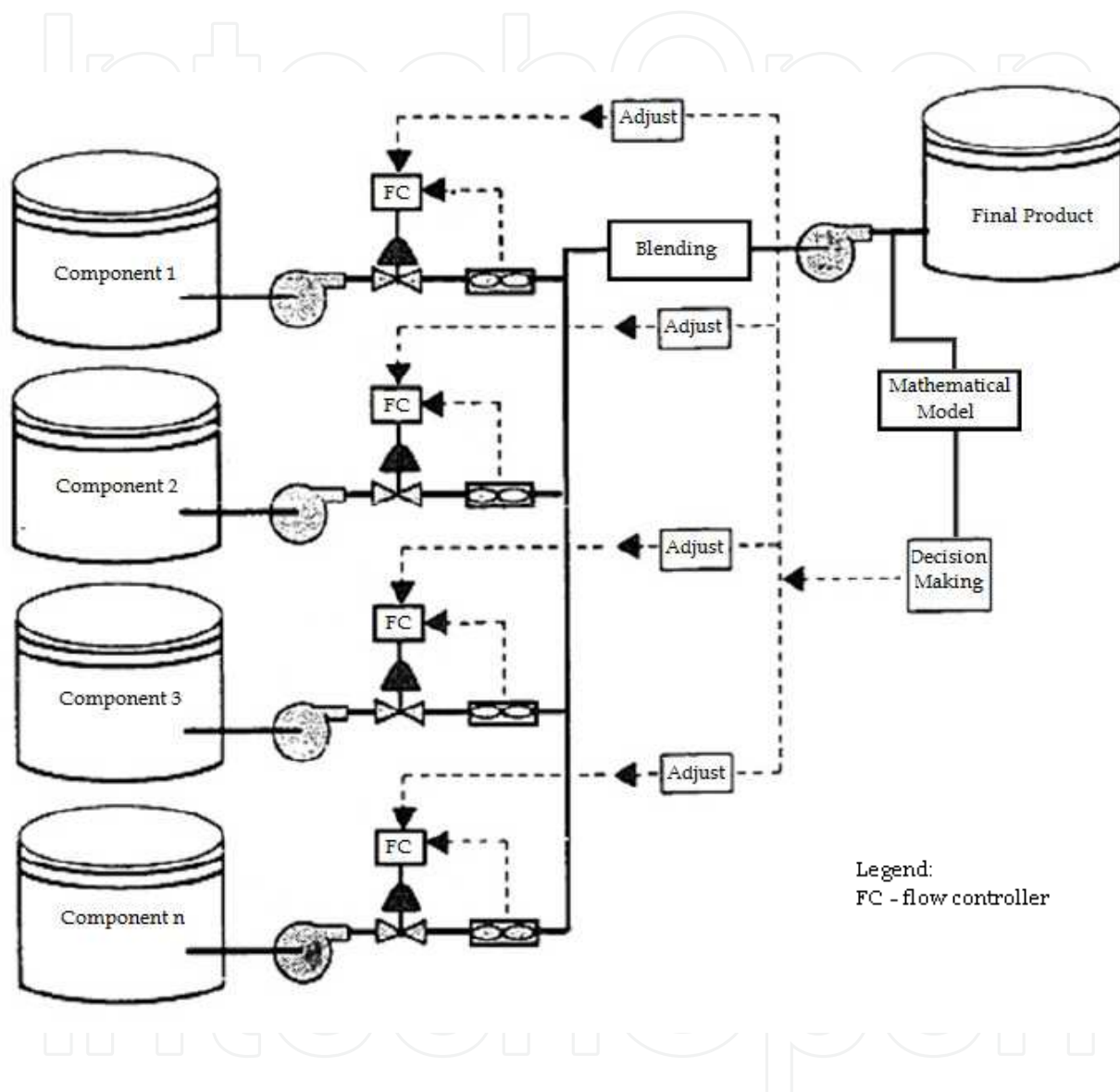

Fig. 12. Representation of online determination of viscosity by mathematical modeling (adapted from Early Jr, 1990)

The preprocessing that provided the best result was the first derivative of the second-degree polynomial proposed by Savitzky-Golay (Galvão et al., 2007), that highlight the differences between samples, contributing to the model can be used to explain the variance between them. The Fig. 13 and Fig. 14 show the original spectra for jet fuel and diesel, and the Fig. 15 and Fig. 16 show the same data after preprocessing by the above derivative.
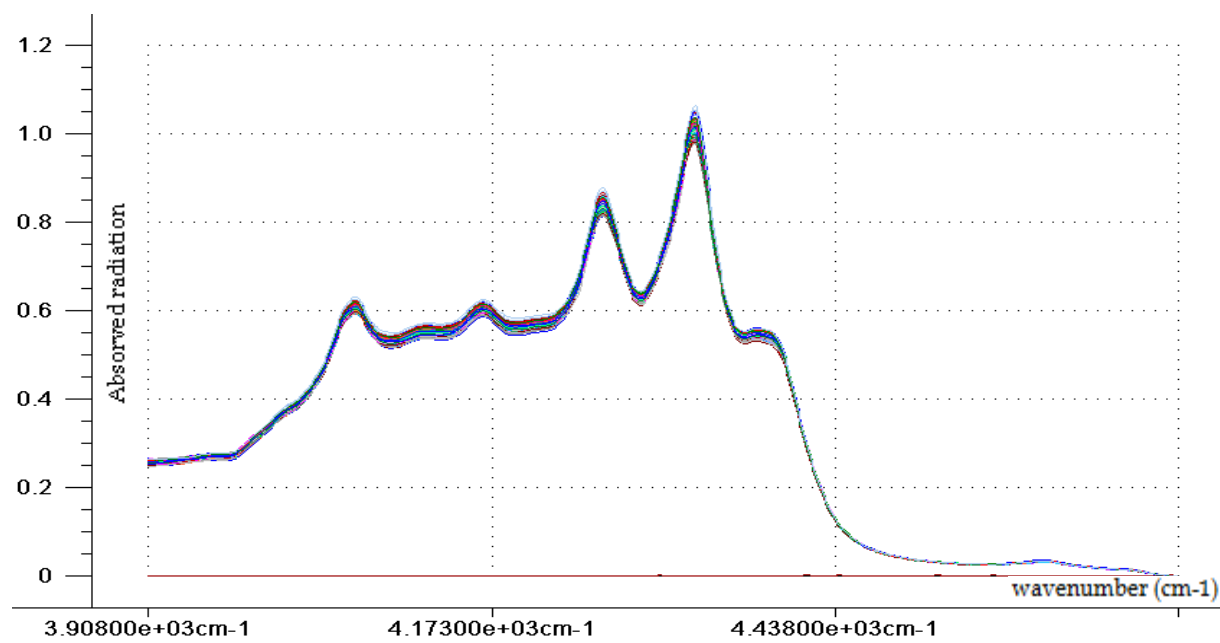
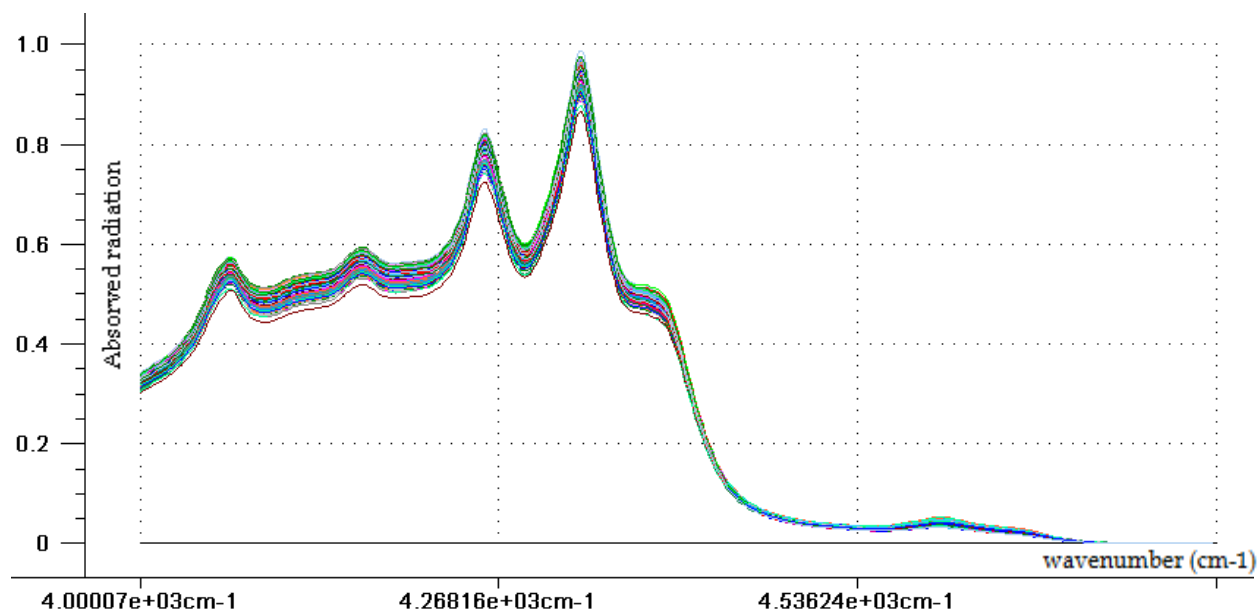Fig. 13. Set of spectra of kerosene used in modeling

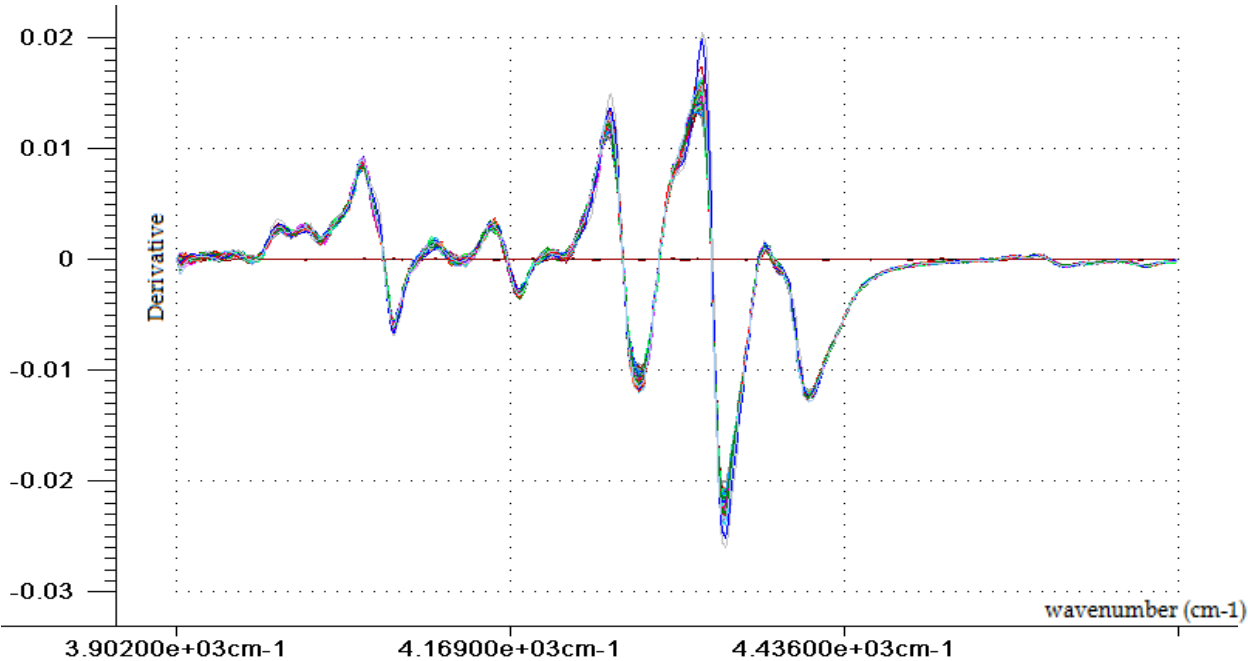Fig. 14. Set of spectra of diesel used in modeling

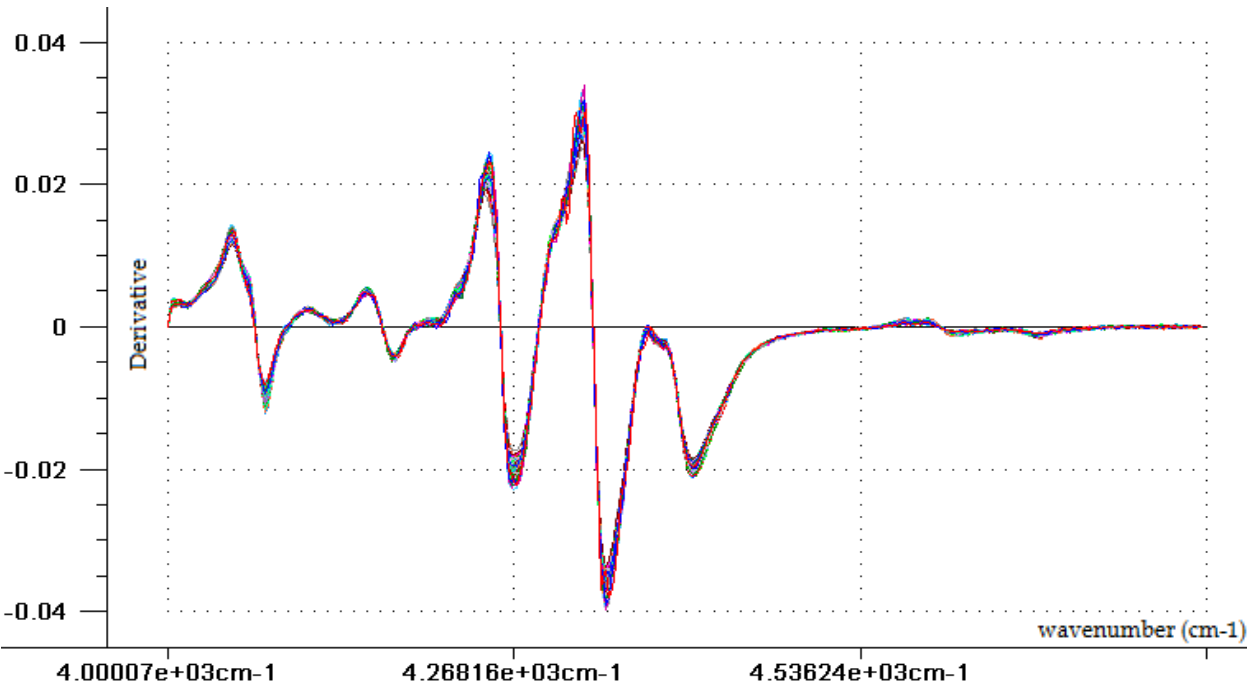Fig. 15. Derivative (Savitzky-Golay) spectra of kerosene

Fig. 16. Derivative (Savitzky-Golay) spectra of diesel

The method of the derivative of Savitzky-Golay smooths the spectrum by polynomial mobile. The derivative of the values of the absorbance as a function of wavenumber is calculated with the polynomial equation cited.

After the derivative of the spectrum, was determined the optimal number of latent variables for each product. In this way, for kerosene, were adopted four latent variables. Above this number, the explained variance decreases due to the incorporation of noise in the model, as shown in Table 1 and Fig. 17.

Latent Variable	Cumulative Variance Explained (%)
1	8.20
2	28.46
3	55.96
4	75.67
5	78.71
6	83.61
7	85.50
8	87.48
9	87.60
10	86.00
11	84.90
12	84.72

Table 1. Cumulative explained variance (%) versus latent variables (kerosene)
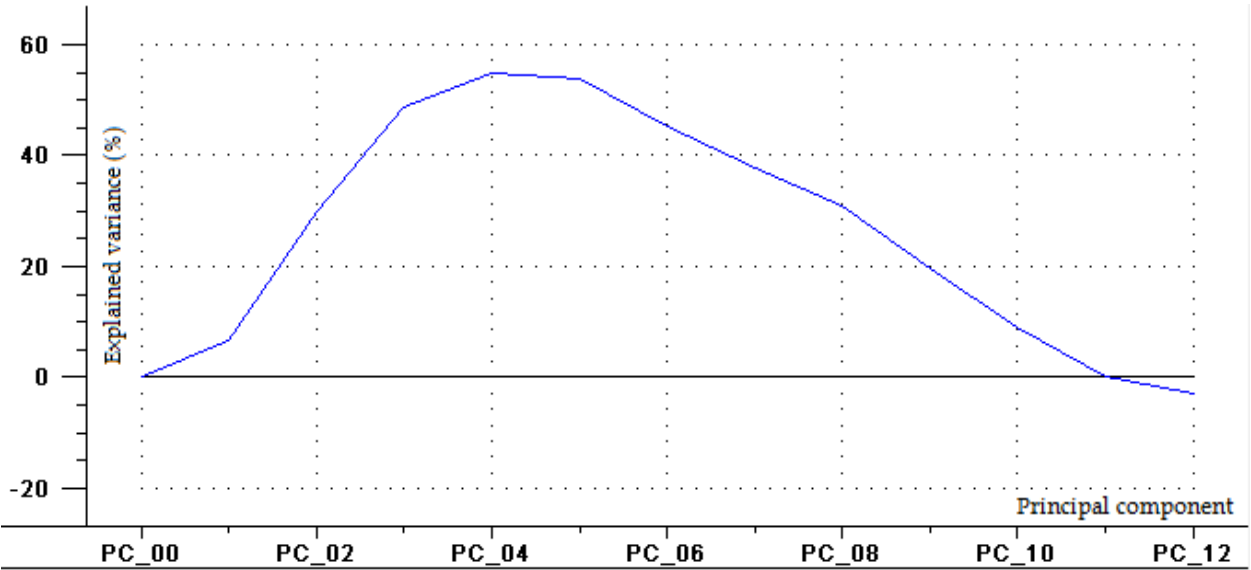

Fig. 17. Explained variance (%) versus latent variables (kerosene).

For the diesel, using the same procedure, were adopted eight latent variables, because above that number, there is no significant gain for explanation of the variance, as shown in Table 2 and Fig. 18.

Latent Variable	Cumulative Variance Explained (%)
1	8.20
2	28.46
3	55.96
4	75.67
5	78.71
6	83.61
7	85.50
8	87.48
9	87.60
10	86.00
11	84.90
12	84.72

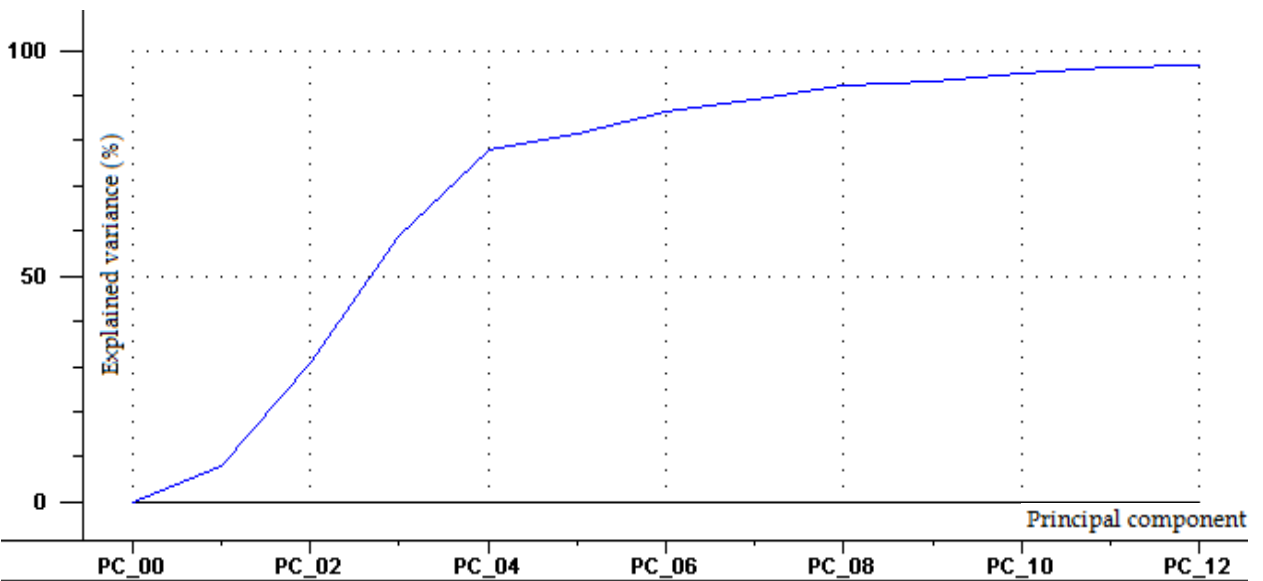Table 2. Cumulative explained variance (%) versus latent variables (diesel)

Fig. 18. Explained variance (%) versus latent variables (kerosene).

The outliers were eliminated observing the graphs of scores of the first two principal components and their influence (residual variance in Y versus leverage). The first two principal components captured the largest variability between the data and both for the samples of diesel and kerosene they are statistically close, as represented in the Fig. 19 and Fig.20.
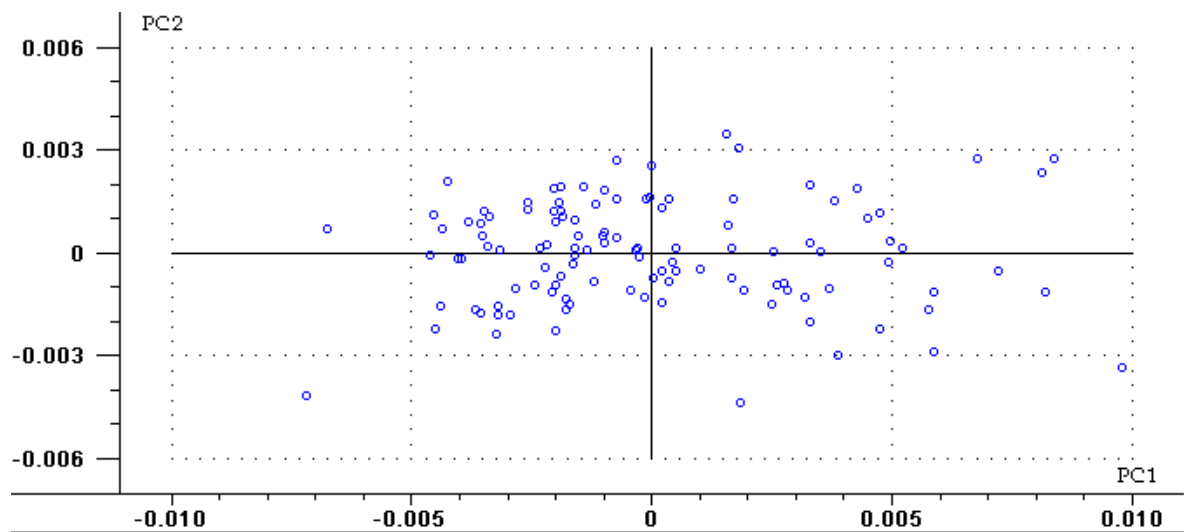
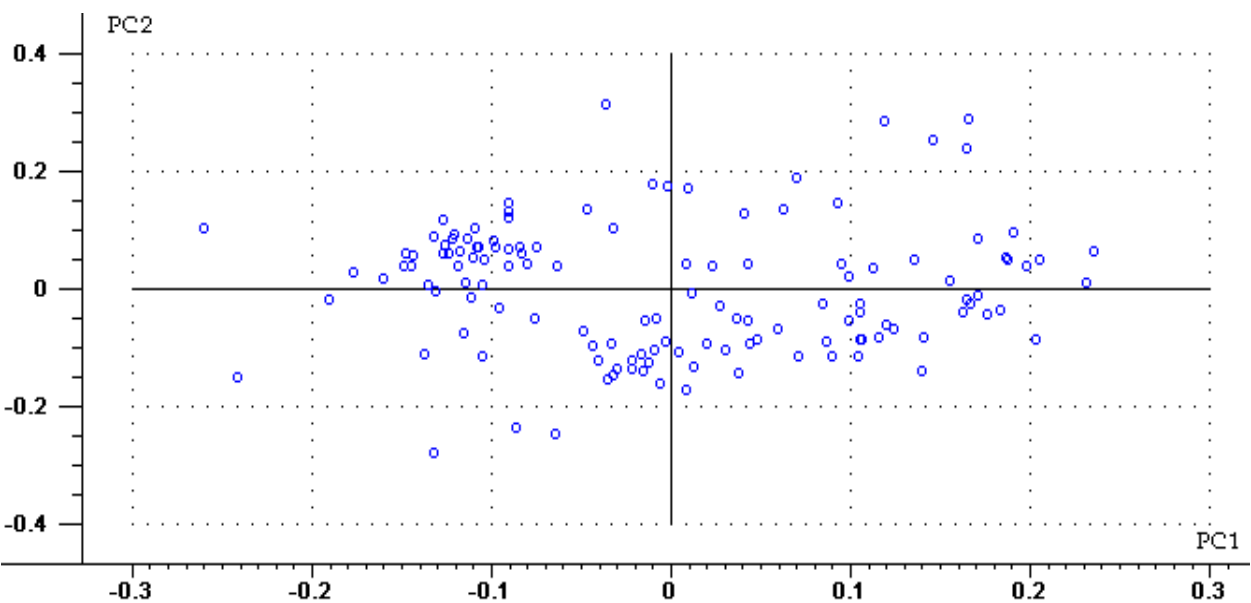Fig. 19. Scores of the first two latent variables for kerosene

Fig. 20. Scores of the first two latent variables for diesel

Among the statistical tools used to detect outliers stands out the Student residues technique versus leverage. The leverage can be interpreted as the distance from the centroid of a sample data set. High values of leverage, means that the sample is located far from the mean and has a major influence in the model. The Student residues can be interpreted as the difference between the actual values and the values predicted by the model. For both products (jet fuel and diesel) there were no evident outliers, and it is also observed that the samples that have great influence have low residue (Fig. 16 and Fig. 17).
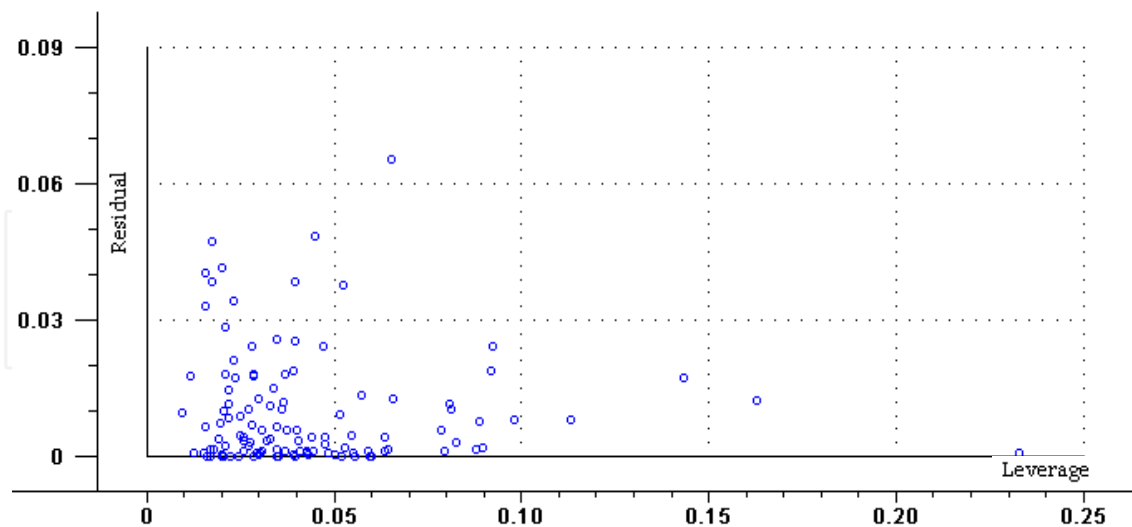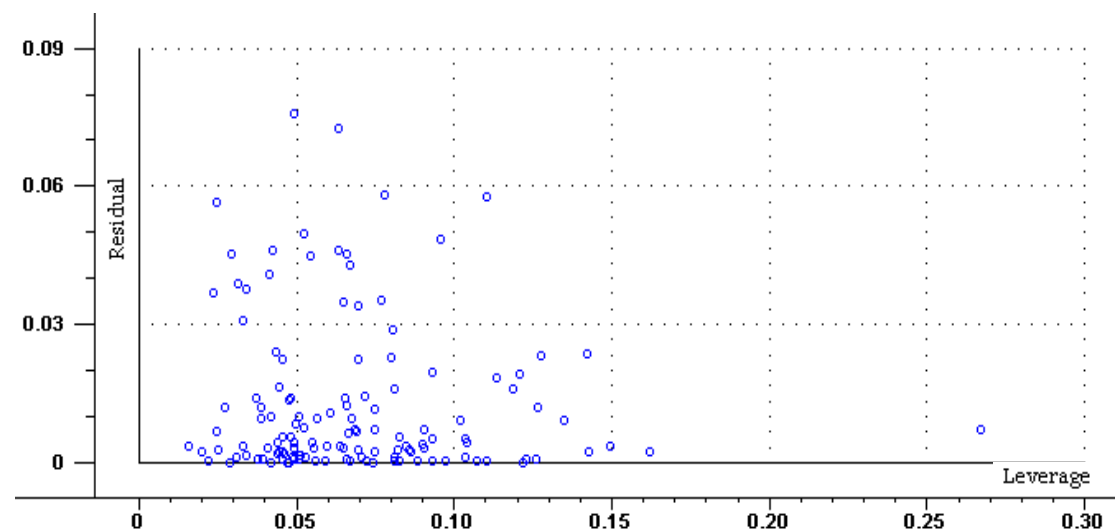
Fig. 21. Residual versus leverage for kerosene

Fig. 22. Residual versus leverage for diesel

The correlation coefficient (r) indicates the degree of correlation between the estimated values and those obtained experimentally. In developing the model, the objectives were to minimize the RMSECV and maximize the coefficient of multiple determination (R^2) or the correlation coefficient (r). Graphs in Fig. 23 and Fig. 24 show the actual values versus predicted values along with the correlation coefficients.

The results of the models are statistically equivalent to those ones from laboratory methods. The Table 3 summarizes the results of modeling.

Property	Number of samples	Latent variables	RMSECV (cSt)	Correlation	Range (cSt)
Kerosene	115	4	0.09	0.889854	3.1 to 4.6
Diesel	131	8	0.11	0.959387	3.1 to 5.3

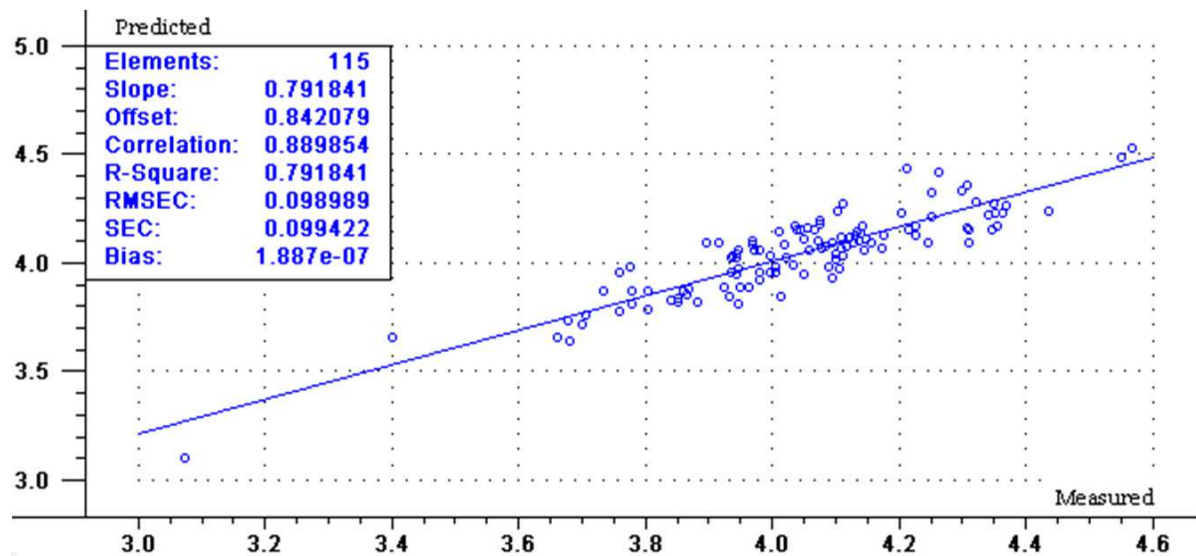Table 3. Summary of modeling results

Fig. 23. Comparison of the results provided by PLS regression model and the results obtained by the reference laboratory (kerosene)
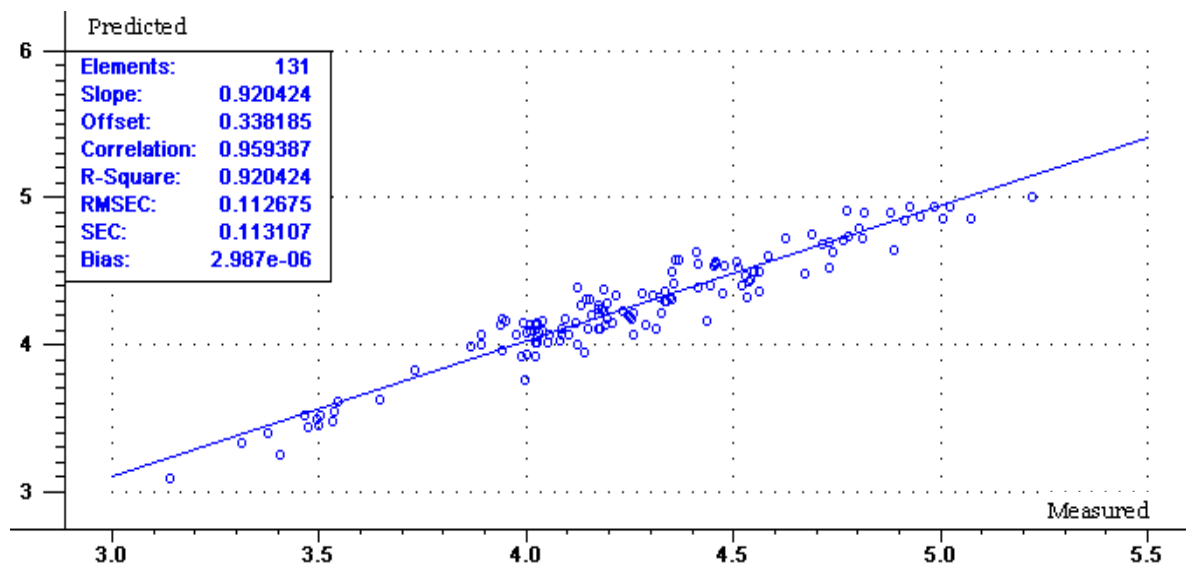

Fig. 24. Comparison of the results provided by PLS regression model and the results obtained by the reference laboratory (diesel)

5. Conclusions

It was possible modeling mathematically the kinematic viscosity of jet fuel and diesel oil by multivariate analysis. The values of the standard error of cross validation showed that the products meet the specifications.

It was possible to solve a real problem combining academia with industry. The results were used in an industrial plant at a refinery in Brazil and helped to speed up the decision-making in blending system, reduced process variability and increase the profitability of production.

There was a significant reduction of the analysis performed in the laboratory, because the proposed method is faster, more practical, does not generate chemical waste, minimize

reprocessing, reducing costs and energy costs. In addition, better quality fuel reduces the impact of burning on the environment.

This chapter opens up other possibilities for assistance, such as simultaneous determinations of several other parameters of oil products.

6. References

Behzadian, M.; Kazemzadeh, R. B.; Albadvi, A. & Aghdasi, M. (2010). PROMETHEE : A comprehensive literature review on methodologies and applications. *European Journal of Operational Reserarch*, Vol.200, (January 2009), pp.198-215, ISSN 0377-2217

Early Jr, P. L. (1990). In-line blending : using a PLC-based process controller. *ISA Transanctions*, Vol.29, No.2, (November 1990), pp. 57-62, ISSN 0019-0578

Fernández, I.; Martinez, M. T.; Benito, A. & Miranda, J. L. (1995). Application of petroleum processing technology to the upgrading of coal syncrude. *Fuel*, Vol.74, No.1, (January 1995), pp. 32-36, ISSN 0016-2361

Galvão, R. K. H.; Araújo, M. C. U.; Benito, A.; Silva, E. C.; José, G. E.; Soares, S. F. & Paiva, H. M. (2007). Cross-Validation for the selection of spectral variables using the sucessive projections algorithm. *Journal of the Brazilian Chemical Society*, Vol.18, No.8, (December 2007), pp. 1580-1584, ISSN 0103-5053

Gomez, A.; Berry, J. J.; Roychoudhury, S.; Coriton, B. & Huth, J. (2007). From jet fuel to electric power using a mesoscale, efficient stirking cycle. *Proceedings of the Combustion Institute*, Vol.31, (June 2006), pp. 3251-3259, ISSN 1540-7489

International Energy Agency (IEA). (n.d.). Refining Processes, In: *Reliance Industries Limited*, 2.10.2011, Available from:
< http://www.ril.com/html/business/refining_processes.html>

Kim, K.; Cho, I. & Park, J. (2000). Use of real-time NIR (near infrared) spectroscopy for the on-line optimization of a crude distillation unit, Proceedings of 12*th Annual Computer Security Incident Handling Conference*, pp. 1-4, Illinois, Chicago, USA, June 25-30, 2000

Marshall, B. (2002.). How diesel two-stroke engines work, In: *Howstuffworks*, 01.09.2011, Available from <http://auto.howstuffworks.com/diesel-two-stroke1.htm>

Naes, T.; Isaksson, T.; Fearn, T. & Davies, T. (2002). Multicollinearity and the need for data compression, In: *A user-friendly guide to multivariate calibration and classification*, NIR Publications, (Ed.), 19-24, ISBN 09528666, Chichester, England

Park, S.; Lee, S.; Jeong, S. J.; Song, Ho-Jun & Park, Jin-Won. (2010). Assessment of CO_2 emissions and its reduction potencial in the Korean petroleum refining industry using energy-environment models. *Energy*, Vol.35, No.6, (June 2010), pp. 2419-2429, ISSN 0360-5442

Pasadakis, N.; Sourligas, S. & Foteinopoulos, Ch. (2006). Prediction of distillation profile and cold properties of diesel fuels using mid-IR spectroscopy and neural networks. *Fuel*, Vol.85, No.7-8, (September 2005), pp. 1131-1137, ISSN 0016-2361

Pasquini, C. (2003). Near infrared spectroscopy : fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, Vol.14, No.2, (March 2003), pp. 198-219, ISSN 0103-5053

Smith, B. C. (2011). How an FTIR works, In: *Fundamentals of Fourier Transform Infrared Spectroscopy*, CRC Press, (Ed.), 19-54, Taylor & Francis, ISBN 978-142-0069-29-7, Boca Raton, United States

Fuel Injection in Automotive Engineering

Edited by Prof Kazimierz Lejda

The main topic of "Fuel injection in automotive engineering" book is fundamental process that determines the development of internal combustion engines and performances of automotive vehicles. The book collects original works focused on up-to-date issues relevant to improving injection phenomena per se and injection systems as the engine key components.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Leandro Valim de Freitas, Fernando Augusto Silva Marins, Ana Paula Barbosa Rodrigues de Freitas, Messias Borges Silva and Carla Cristina Almeida Loures (2012). Multivariate Modeling in Quality Control of Viscosity in Fuel: An Application in Oil Industry, Fuel Injection in Automotive Engineering, Prof Kazimierz Lejda (Ed.), ISBN: 978-953-51-0528-2, InTech, Available from: http://www.intechopen.com/books/fuel-injection-in-automotive-engineering/multivariate-modeling-in-quality-control-of-viscosity-in-fuel-an-application-in-oil-industry

INTECH

open science | open minds