

Short communication

Feasibility of quality process control of a naphtha fractioning unit based on near-infrared spectroscopic prediction of physical and chemical properties of medium naphtha streams

Marcio V. Reboucas^{a,*}, Edson C. Santos^a, Fernanda S.V. Vieira^{a,b}

^a *Braskem S.A., Unidade de Insumos Básicos, Laboratório, Rua Eteno 1561, Complexo Petroquímico de Camacari, 42810-000 Camacari/Bahia, Brazil*

^b *Universidade Federal da Bahia, Instituto de Química, Campus Universitário de Ondina, Ondina, 40170-290 Salvador/Bahia, Brazil*

Received 25 March 2006; accepted 11 September 2006

Available online 25 October 2006

Abstract

Near-infrared spectroscopy multivariate calibration models were used to predict physical and chemical properties of medium naphtha samples providing the means of appropriate process control of naphtha fractioning units. Relative density, distillation curve (initial boiling point, 5%, 10%, 30%, 50%, 70%, 90%, 95% evaporated and final boiling point) and chemical composition (PINA) were estimated from FT-NIR spectra in the 1660–1800 and 2200–2500 nm regions. Principal component analysis was used for exploratory data analysis and cross-validated one-block partial least-squares models were employed for calibration. Root mean square errors of prediction, based on 4–8 PLS factors, were close to the measured repeatability values and below the limits defined by ASTM methods. The present work has demonstrated the feasibility of performing several analyses required for quality control of a naphtha fractioning unit based on a single NIR spectrum, providing good means for fast process control. © 2006 Elsevier B.V. All rights reserved.

Keywords: NIR; PLS; Relative density; Distillation curve; PINA; Naphtha

1. Introduction

Naphtha is a light fraction of petroleum enriched in paraffinic, isoparaffins (alkanes) and naftenic (cyclo-alkanes), and also containing lower levels of aromatic and olefinic (alkenes) compounds, usually with a distillation curve ranging from 20 to 220 °C. It is used as feedstock for petrochemical industries to produce a large number of basic petrochemicals such as ethylene, propylene, benzene, *p*-xylene, etc. Naftenic-rich naphtha is mainly used to feed catalytic reforming units which aims at converting C6–C8 naftenic compounds into benzene, toluene and xylenes. To obtain the maximum efficiency in the reforming catalytic reaction the crude naphtha must be previously fractioned to obtain a stream called medium naphtha. In such a process, liquefied petroleum gas (LPG), light naphtha and heavy naphtha area also produced. Several analytical parameters, namely distillation curve, relative

density and chemical composition, are provided to ensure the process control of the naphtha fractioning unit to produce the tightly-specified medium naphtha.

Near-infrared (NIR) spectroscopy has proved to be quite suitable for such applications, providing the basis for predicting several chemical and physical properties from a single sample's spectrum [1–16]. By means of powerful multivariate techniques like partial least-squares regression (PLS), each property of interest can be regressed on the NIR absorption spectra, resulting in a calibration model that can be used to predict that property's value for a given test sample. In the present work, we report on successful multivariate modelling achieved for 14 analytical parameters.

2. Experimental

2.1. Samples

Forty-one medium naphtha samples were collected over a period of 2 months from the naphtha fractioning unit. Samples of heavy and light naphtha were collected once.

* Corresponding author. Tel.: +55 71 3632 5072; fax: +55 71 3632 5072.

E-mail address: marcio.reboucas@braskem.com.br (M.V. Reboucas).

2.2. Experimental design

A 10-mixture simplex-lattice design simulating variations in the medium naphtha profile due to “out of control” process situations was chosen, to improve model robustness [17,18]. Components of the light and heavy naphthas are the most likely sources of medium naphtha contamination. A conventional three-component design would lead to mixtures with too low medium naphtha amounts or even no medium naphtha at all. A different pseudocomponent design was chosen instead, subjected to the following concentration constraints: medium naphtha: 70–100% (v/v); heavy naphtha: 0–30% (v/v); light naphtha: 0–30% (v/v) (Fig. 1). The central point mixture was prepared in duplicate.

2.3. Instrument and apparatus

The near-infrared transmission spectra were recorded in the 1000–2500 nm region on a Fourier-Transform NIR Spectrometer MB 160 Series (Bomem/Hartmann & Braun, Canada) interfaced to an IBM microcomputer. A CaF₂ sample cell, model S1-4, with 0.5 mm pathlength (International Crystal Labs, NJ, USA) and an InAs detector under nitrogen purge were used. Instrument parameters such as noise level and wavelength accuracy were verified before performing the analyses according to the manufacturer’s instructions, as already mentioned elsewhere [15,19,20].

Reference results of each modelled property were obtained for each sample, using the ASTM standard methods D-4052 and D-86 for relative density and distillation parameters, respectively [21]. The chromatographic data, paraffins, isoparaffins, olefins, naftenics, and aromatics (PIONA) were obtained from a gas chromatograph with flame ionization detector and 40 m × 0.1 mm × 0.20 μm PONA column (J&W Scientific), using ASTM D-5134 as reference method.

2.4. Data set preprocessing

A background spectrum of air was recorded before each analysis and then subtracted from the spectrum of each sample. All FT-NIR spectra were collected as averages of 20 scans at 4 cm⁻¹ intervals corresponding to 3113 points. Raw FT-NIR spectra were fed into the PLSIQ software [22] together with the

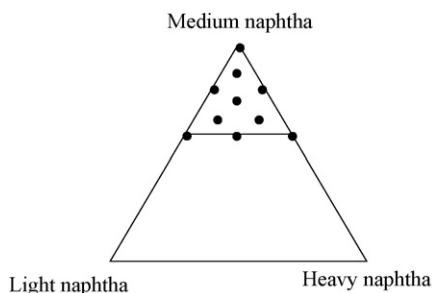


Fig. 1. Simplex lattice design used for specifying the mixtures of light, medium and heavy naphtha. A 70% (v/v) lower bound is imposed on the medium naphtha content.

sample property values. As described below, only a fraction of the entire spectrum (228–740 points) was selected for the property calibration. Each spectrum was normalised by division with the average absorbance value over the whole spectrum. First- and second-order derivatives were also tried, being used only for FBP calibration, since they did not improve model performance for the other properties. Prior to regression, the wavelength-variables were mean-centered and the property-variables were auto-scaled (mean-centered and also scaled to unit variance) [23,24]. Fifty-three samples were analysed, including medium naphtha samples and designed mixtures, for the training set. Fifteen additional samples were analysed to carry on the models validation.

2.5. Statistical procedure

A first exploratory analysis using PCA was performed on the auto-scaled data, to search for outlying points in the calibration set, possibly resulting from contamination, sampling errors, or extreme composition differences between the designed mixtures and actual naphtha samples. PCA score, Studentized Residual versus Sample Leverage, and Spectral Residual versus Concentration Residual plots were used to identify and exclude outlying points. The detected outliers were left out of the calibration set. Calibration was accomplished with the partial least-squares technique. The leave-one-out cross-validation method was used to estimate the optimal number of principal components, in which one sample was randomly left out of each calibration step [23,24]. Root mean square error of prediction (RMSEP) and Residuals plots were used for model evaluation. The samples excluded from the calibration set in the model refining process were added to the external validation set, a test set consisting of samples not included in the calibration set. After successive refinement steps, the external validation set was used to assess the final calibration model. The predicted values were then compared to the measured values by means of a paired *t*-test [25]. The average deviation between the NIR prediction and the reference method result was compared to the reproducibility value of the reference method.

3. Results and discussion

3.1. Multivariate calibration

After several calibration trials, the 1660–1800 and 2200–2500 nm spectral ranges were selected for calibration. A 1600–1800 nm region is usually assigned to first overtones of methyl C–H, methylene C–H and C–H aromatic stretching vibrations while 2200–2500 nm results of combination bands [26–28]. Wavelengths below 1660 nm and between 1800 and 2200 nm were excluded from the model because they do not carry relevant information, as it can be observed from Fig. 2, and could only add to noise. In addition, it has already noticed that they are particularly sensitive to temperature changes and instrumental variations [15,20]. The results of the best calibration model for each property are summarised in Table 1.

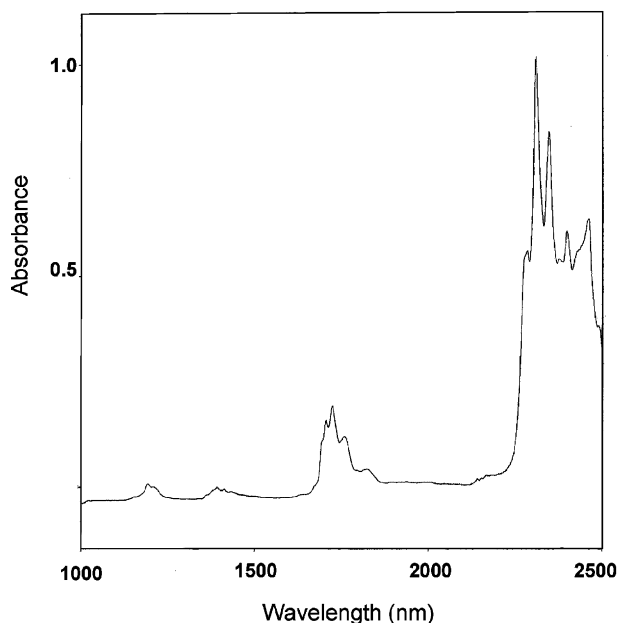


Fig. 2. Typical NIR spectrum of medium naphtha sample (0.5 mm pathlength, Bomem/Hartmann & Braun equipment).

3.2. Model precision

A medium naphtha sample was analysed in five replicates to provide an estimate of pure experimental errors of the reference method, i.e. the standard deviation under repeatability conditions. Table 1 brings this experimental error estimate (RMSEP target) and ASTM repeatability value for comparison with the RMSEP obtained for each regression model.

The results confirm the excellent performance of the PLS models in predicting PIONA composition (except for Olefins), relative density and distillation properties of medium naphtha

samples, especially when compared to ASTM values. Since ASTM values usually represent maxima limits, the RMSEP target provides a more realistic value. Even if compared to such values, the PLS prediction errors are still quite good, usually in the same order of magnitude.

The Olefins model showed poorer performance, clearly related to the very low concentration levels present in the samples and insufficient property variation for NIR calibration. It is worthy mentioning this property is not needed for process control and an approximation value obtained from normalisation of NIR PINA prediction is usually suitable for this purpose.

3.3. Model accuracy

The best calibration model for each property was submitted to a final assessment. From twelve to 30 samples were used to test model accuracy. The NIR spectrum of each sample was acquired, submitted to the same preprocessing used for the samples belonging to the calibration set, and then fed into the PLS models for property prediction. The performance of each model was evaluated by comparison of the predicted and measured values through paired *t*-tests. None of the observed *t* values was higher than the respective critical *t* value. Therefore, the null hypothesis that there was no significant difference between the predicted and measured values could not be rejected, and the models were accepted.

The average percentage difference between the NIR prediction and reference value was also compared to the relative reproducibility value of the reference method. The values ranged from -3.2 to 2.0% and since they were found within acceptable values there is no evidence of bias. The percentage difference for each property is shown in box-plots in Fig. 3.

Table 1
Validation results of PLS calibration models

Property	<i>N</i>	Range	Typical value	r^2	r^2 target	RMSEP	RMSEP target	1σ (ASTM)
Paraffins (wt%)	7	26–32	28.2	0.9550	0.9994	0.24	0.03	0.33
Isoparaffins (wt%)	7	26.8–30.2	27.9	0.9325	0.9017	0.17	0.27	0.04
Olefins (wt%)	12	0.30–0.64	0.5	0.2385	0.7380	0.07	0.10	0.20
Naftenics (wt%)	4	32–38	36.0	0.9590	0.9461	0.26	0.34	0.15
Aromatics (wt%)	8	5.8–8.8	7.4	0.9366	0.9712	0.14	0.12	0.01
Relative density 20/4 °C	7	0.71–0.74	0.7303	0.9840	0.9999	0.0004	0.00005	0.0001
Initial boiling point (°C)	7	56–92	84.3	0.9743	0.9975	1.04	0.54	3.50
5% evaporated (°C)	5	78–100	93.4	0.9647	0.9994	0.62	0.15	2.08
10% evaporated (°C)	5	83–101	95.4	0.9736	0.9991	0.43	0.15	2.20
30% evaporated (°C)	7	96–109	101.9	0.9724	0.9998	0.35	0.04	1.68
50% evaporated (°C)	8	106–113	109.1	0.9014	0.9986	0.48	0.13	1.75
70% evaporated (°C)	7	112–136	117.6	0.9718	0.9814	0.65	0.52	1.91
90% evaporated (°C)	7	122–148	129.3	0.9696	0.9961	0.80	0.28	1.74
95% evaporated (°C)	6	128–154	134.8	0.9449	0.9884	1.09	0.54	2.29
Final boiling point (°C)	6	138–176	145.0	0.8351	0.9956	3.13	0.47	3.50

N is the number of PLS factors in the model. Range presents the range of application of each model. The typical analytical results of medium naphtha characterization obtained at the Braskem naphtha fractioning unit are also showed. The correlation between measured (*Y*) vs. predicted (*X*) values is given by the coefficient of determination r^2 . r^2 target is calculated based on the variance of the reference method and the variance of the data base [r^2 target = $1 - (\text{variance of the reference method}/\text{variance of the data base})$]. RMSEP stands for root mean square error of prediction. RMSEP target is the experimental standard deviation of the reference method obtained under repeatability conditions. The ASTM repeatability is showed in the last column for comparison, except chromatographic results which are repeatability based on 10 replicates of a real sample.

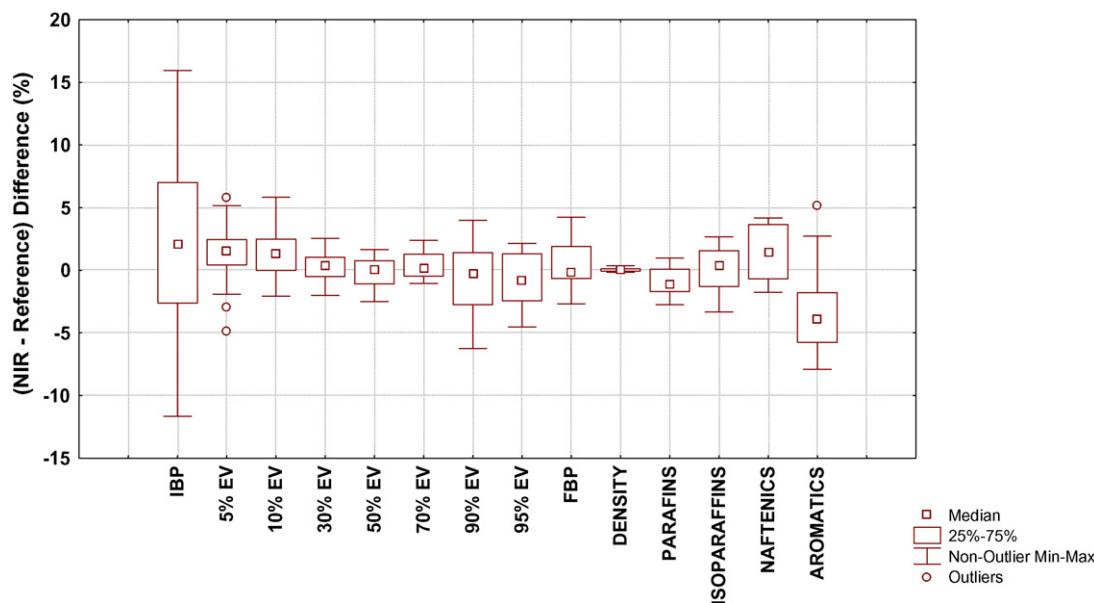


Fig. 3. Box-plots on the percentage differences obtained between NIR prediction and reference method result for the validation set. Olefins was excluded from the graph for best scaling.

4. Conclusions

It has been shown that near-infrared spectroscopy can be used as a reliable method for estimating different properties of medium naphtha samples produced at the Braskem petrochemical plant. PLS calibration models provided very good prediction results for the PINA composition, relative density and distillation parameters. Predicted values were always close to the values determined with reference methods, and the required ASTM correlation for model validation was successfully achieved. The method is already in routine use in Braskem laboratory, and the replacement of chromatographic, distillation and density methods by a single NIR spectrum acquisition has significantly reduced the time to report the results providing better process control. For example, the fast chromatographic analysis is carried out in 45 min, while less than 10 min is required for the whole procedure of NIR prediction.

Finally, it is worthy emphasizing the RMSEP and coefficient of determination of the calibration models, widely used by authors for model evaluation, should be handled with great care. In some cases, the best results for these parameters lead to an over-refined model, whose behaviour in routine analyses is not acceptable because of high occurrence of outliers or high deviation from reference values. So, a compromise between model robustness and good r^2 and RMSEP should be pursued and the final evaluation with an external validation set is the most appropriate way to accomplish it.

Acknowledgements

The authors thank Braskem S.A. for supplying technical resources required for this work and Prof. Celio Pasquini (UNICAMP) for final manuscript revision.

References

- [1] S.J. Swarin, C.A. Drumm, *Spectroscopy* 7 (7) (1992) 42–49.
- [2] G. Buttner, *Process Control Quality* 9 (1997) 197–203.
- [3] D.E. Honigs, T.B. Hirschfeld, G.M. Hieftje, *Anal. Chem.* 57 (2) (1985) 443–445.
- [4] J.J. Kelly, C.H. Barlow, T.M. Jinguji, J.B. Callis, *Anal. Chem.* 61 (4) (1989) 313–320.
- [5] Z. Chen, X. Feng, *Hydrocarbon Process.* 71 (1) (1992) 94–96.
- [6] H. Chung, M. Ku, J. Lee, *Vib. Spectrosc.* 20 (2) (1999) 155–163.
- [7] N. Asker, S. Kokot, *Appl. Spectrosc.* 45 (7) (1991) 1153–1157.
- [8] A. Espinosa, M. Sanchez, S. Osta, C. Boniface, J. Gil, A. Martens, B. Descales, D. Lambert, M. Valleur, *Oil Gas J.* 92 (42) (1994) 49–56.
- [9] J.J. Kelly, J.B. Callis, *Anal. Chem.* 62 (14) (1990) 1444–1451.
- [10] F. Parisi, L. Nogueiras, H. Prieto, *Anal. Chim. Acta* 238 (1) (1990) 95–100.
- [11] G.E. Fodor, K.B. Kohl, R.L. Mason, *Anal. Chem.* 68 (1) (1996) 23–30.
- [12] M. Ku, H. Chung, *Appl. Spectrosc.* 53 (5) (1999) 557–564.
- [13] C.J. deBakker, P.M. Fredericks, *Appl. Spectrosc.* 49 (12) (1995) 1766–1771.
- [14] G.E. Fodor, K.B. Kohl, *Energy Fuels* 7 (5) (1993) 598–601.
- [15] M.V. Reboucas, B. de Barros Neto, *J. Near Infrared Spectrosc.* 9 (2001) 263–273.
- [16] F.S.V. Vieira, *Calibração Multivariada Aplicada à Predição da Composição da Nafta (COPENE) e Otimização do seu Processo de Fracionamento*, Academic Dissertation, Chemistry Institute, Federal University of Bahia (UFBA), Bahia, Brazil, 2001.
- [17] B. Barros Neto, I.S. Scarminio, R.E. Bruns, *Planejamento e Otimização de Experimentos*, second ed., Editora da Unicamp, Campinas, Brazil, 1996.
- [18] G.E.P. Box, W. Hunter, J.S. Hunter, *Statistics for Experimenters*, Wiley, New York, 1978.
- [19] J.R. Ferraro, L.J. Basile, *Fourier Transform Infrared Spectroscopy—Applications to Chemical Systems*, second ed., Academic Press, New York, 1979, pp. 218.
- [20] U. Hoffmann, N. Zanier-Szydowski, *J. Near Infrared Spectrosc.* 7 (1991) 33–45.
- [21] *Annual Book of ASTM Standards* (vols. 05.01, 05.02). American Society for Testing and Materials, Philadelphia, PA.

- [22] PLSplus/IQ, Add on application for Grams/32, Galactic Industries Corporation.
- [23] K. Beebe, R.J. Pell, M.B. Seasholtz, *Chemometrics: A Practical Guide*, first ed., John Wiley & Sons, Inc., 1998.
- [24] H. Martens, T. Naes, *Multivariate Calibration*, Wiley, New York, 1989.
- [25] J.C. Miller, J.N. Miller, *Statistics for Analytical Chemistry*, third ed., Ellis Horwood PTR Prentice Hall, London, 1993.
- [26] R.F. Goddu, D.A. Delker, *Anal. Chem.* 32 (1) (1960) 140–141.
- [27] J.J. Workman Jr., *Appl. Spectrosc. Rev.* 31 (3) (1996) 251–320.
- [28] L.G. Weyer, *Appl. Spectrosc. Rev.* 21 (1–2) (1985) 1–43.