Available online at www.sciencedirect.com
SCIENCE DIRECT®
Talanta 66 (2005) 218–224
Talanta
www.elsevier.com/locate/talanta

Comparison and joint use of near infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters

M. Urbano Cuadrado[a], M.D. Luque de Castro[a,*], P.M. Pérez Juan[b], M.A. Gómez-Nieto[c]

[a] Department of Analytical Chemistry, University of Córdoba, Campus of Rabanales, Annex C3, E-14071 Córdoba, Spain
[b] LIEC, Agrifood and R&D Laboratory, E-13200 Manzanares, Spain
[c] Department of Computing and Numerical Analysis, University of Córdoba, Campus of Rabanales, Building C2, Floor 3, E-14071 Córdoba, Spain

Received 20 July 2004; accepted 17 November 2004
Available online 23 December 2004

Abstract

A study of the statistic characteristics of the multidetermination of several enological parameters – namely, alcoholic degree, volumic mass, total acidity, glycerol, total polyphenol index, lactic acid and total sulphur dioxide – depending on the spectroscopic zone employed, was carried out. The two techniques used were near infrared spectroscopy (NIRS) and Fourier transform mid infrared spectroscopy (FT-MIRS). The combination of these two regions (sum of their spectra) was also studied. NIRS yielded better results, but the use of both zones improved the determination of glycerol and total sulphur dioxide. The training and validation sets used for developing general equations were built with samples from different apellation d'origine, different wine types, etc. Partial least squares regression was used for multivariate calibration, using systematic cross validation in the calibration stage and external validation in the testing stage. Sample preparation was not required.

Keywords: Near infrared spectroscopy; Attenuated total reflection mid infrared; Wineries; Multivariate calibration

1. Introduction

One of the last trends in analytical chemistry is aimed at shortening the time required for a given analysis and subsequent availability of the pursued information in a short time. This goal is so far achieved by different ways: namely, the reduction and simplification of the steps involved in sample preparation [1,2], the recent advances in automation and instrumentation [3–5], the development of qualitative and screening methodologies [6,7], the higher use of chemometrics thanks to advances in computers [8,9], etc.

In the food area, and specifically in the wine industry, the above commented ways have also been used in order to obtain fast information about the production process. Several methods based on flow injection (FI) have been developed in order to automate and reduce the complexity of some official methods in wine laboratories. Thus, parameters as ethanol, glycerol, total and free sulphur dioxide, etc. [10,11] can be determined by FI methods.

Chemometrics in wine production and laboratory is widely referenced in the literature from several points of view. Thus, supervised and non-supervised pattern recognition techniques have been used to distinguish different varieties, geographical areas, elaboration processes, etc. Variables used for differencing can be classified into non-spectroscopic and spectroscopic variables [12–14]. Concerning the latter, the spectra obtained, generally in few seconds, by spectroscopic techniques namely, near infrared reflectance spectroscopy (NIRS), Fourier transform mid infrared spectroscopy (FT-MIRS), Ultraviolet–visible spectroscopy (UV–vis), nuclear magnetic resonance (NMR), etc. constitute large data sets which contain hidden information. Multivariate regression, such as principal component regression (PCR) and partial

---

[*] Corresponding author. Tel.: +34 957218615; fax: +34 957218615.
E-mail address: qa1lucam@uco.es (M.D. Luque de Castro).

0039-9140/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2004.11.011

least squares regression (PLSR) are standards in chemometrics, which has been used for developing equations for the determination of quantitative parameters in wine and other food industries using the data provided by the spectroscopic techniques above cited. Approaches based on NIRS are the most used for the determination of different analytes; namely, ethanol, glycerol, sugars, [15,16] etc. FT-IR applications for wine analyses have also been developed [17,18].

Here, the applicability of both NIRS and FT-MIRS techniques to the determination of several enological parameters – alcoholic degree, volumic mass, total acidity, glycerol, total polyphenol index, lactic acid and total sulphur dioxide – is studied in wines from both several appellations d'origine and different grape varieties. The aim of this study is the development of equations independent of the type of wine. In addition, the joint use of spectral regions of shorter wavelength for improving the statistic parameters of the calibrations is discussed for the first time.

2. Experimental

2.1. Samples and sample preparation

Different wines – including red, rosé and white wines; young and aged wines; wines from different appellation d'origine ("La Mancha", "Valdepeñas", "Jumilla", "Navarra", "Alicante" and "Madrid") and grape varieties ("Cencibel", "Cabernet Sauvignon", "Cencibel-Cabernet Sauvignon", "Merlot", "Garnacha Tintorera" and "Syrah") – were used in the present study. Thus, the number of samples employed in the calibration and validation steps was 180. The samples were used as such, because filtering, dilution, preconcentration, interferents removal, etc. were not required.

2.2. Apparatus and methods

The instrument employed for NIR spectra collection was a Foss-NIRSystems 6500 System II spectrophotometer (Foss-NIRSystems Inc., Silver Spring, MD, USA) equipped with a transport module and capable of making measurements at 2 nm resolution in the spectral range covering 400–2500 nm. The samples were analysed by folded transmission using a ring cup with a 0.1 mm pathlength. A diffuse reflecting gold surface placed at the bottom of the cup reflected the radiation back through the sample to the reflectance detector. The spectra were collected using WinISI software 1.50 (Infrasoft International, Port Matilda, PA, USA). Before recording the spectra, the samples were thermostated at 24 °C. The reflectance spectra (log 1/*R*) were collected in duplicate.

The instrument employed for MIR spectra collection was an FT-MIR Nicolet Magna-IR550 Serie II (Nicolet Instrument Corp., Madison, Wisconsin, USA), capable of making measurements at 4 cm^{-1} resolution in the spectral range covering 4000–400 cm^{-1}. The instrument was furnished with an infrared attenuated total reflection (ATR) solid, liquid and mellow sample cell with a zinc selenide crystal (Spectra Tech., Stamford, CT, USA) for Nicolet Spectrometers. The ATR cell consists of the following: base accessory unit containing the mirrors to direct the IR beam to and from the sampling plate; a sampling plate containing the ZnSe ATR crystal; precision single or dual readout controller.

Table 1
Enological parameters and reference methods

Parameter	Reference method
Alcoholic degree	Distillation and aerometry
Volumic mass	Aerometry
Total acidity	Titration with NaOH up to pH = 7.0
Glycerol	Enzymatic reaction
Total polyphenol index	Folin–Ciocalteu reagent in alkaline medium
Lactic acid	High pressure liquid chromatography
Free sulphur dioxide	Iodometry in acid medium

Other characteristics were a transmission range between 20,000 and 650 cm^{-1}, a refractive index (at 1000 cm^{-1}) of 2.4, density of 5.27 $\mathrm{g\,cm}^{-3}$ and volume of 3.5 ml. Before recording the spectra, the samples were thermostated at 24 °C. The reflectance spectra (log 1/*R*) were collected in duplicate.

On the other hand, the samples were also analysed in duplicate by the reference methods, and standard error laboratory (SEL) was estimated from the duplicates. The reference methods used are shown in Table 1.

2.3. Chemometric software for data processing and statistical techniques used

The Unscrambler 7.8 (Camo Process AS, Oslo, Norway) was used for data processing. The chemometric procedure consists of the following steps.

2.3.1. Principal components analysis (PCA) for visualisation of spectral outliers

PCA [19] was required for the reduction of the number of variables showing co-linearity, thus representing the samples in a new, reduced *n*-dimensional space. Once the samples were in the new space defined by principal components, the leverage value – a measure of how far an object is compared to the majority – was computed. Sample spectra with leverage higher than 0.5 were considered outliers. The outliers were examined in order to know if either they provided any useful information or they must be removed.

2.3.2. Spectra preprocessing

Different treatments were applied to the spectra in the calibration and validation steps, namely, data centring, scatter correction, first derivative, and some of them jointly.

2.3.3. Calibration step: cross validation

In this step, partial least square regression (PLSR) [19,20] was used for developing the equations. The number of calibration groups and maximum number of PLS factors were set at 5 and 16, respectively. The latter is based on the following rule: one PLS factor per 10 samples of the training set plus 2. On the other hand, a study of possible outliers in the prediction of the cross validation was carried out taking into account the statistic *t* (Student's test) parameter, which was set at 2.50. Statistic parameters as standard error cross validation (SECV) and the determination coefficient (R^2) were employed.

2.3.4. Validation step

The validation set was introduced in the model, thus statistic parameters as standard error prediction (SEP) and r^2 were obtained.

The procedure above commented was applied to the spectral data from the NIR and the MIR zones and a combination of both.

3. Results and discussion

3.1. Reference data

Table 2 shows information about reference data. Thus, the range, mean, standard deviation (S.D.) and number of samples for the calibration and validation sets are summarised in the table, in addition to the standard error laboratory (SEL). As can be seen, the range of reference values encompassed the characteristic values for a high diversity of wines.

3.2. Sample spectra

Spectra in the whole measurement range provided by the NIR instrument; that is, 400–2500 nm were obtained. Nevertheless, the zone of the MIR spectra used was only the 800–3000 cm^{-1} region; thus, the 400–800 and 3000–4000 cm^{-1} zones were not employed in this study owing to the high irreproducibility because of the high absorbance values. NIR and MIR spectra are shown in Fig. 1A and B, respectively. The highest differences in the NIR spectra were within 400–1100 and 1910–1960 nm, while the MIR spectra presented the largest differences in the 1000–1500 and 2300–2400 cm^{-1} regions. The combination of the NIR and MIR spectra is shown in Fig. 1C.

In the NIR zone, the absorption bands with the highest intensity – namely, 1820–2030 and 1370–1600 nm – are second and first harmonics, respectively, of the MIR absorption band corresponding to the strength vibration of the −OH group. The fundamental vibration band corresponding to this group is at 3300–3500 cm^{-1}, which was not used owing to the high irreproducibility, as commented before.

3.3. Spectral outliers

PCA was applied to the three data matrixes corresponding to the sample spectra for NIR, MIR and their combination. The leverage values were calculated in order to detect spectral outliers. Four and three spectra were outliers for NIR and MIR, respectively. These spectral outliers were not common for the two regions; which can be due to operational errors. With respect to the combination of the two zones, only two spectra were computed as outliers. The samples plotted in the bi-dimensional space formed by the first two principal components for the NIR region are shown in Fig. 2. The outliers are within circles. Two groups can also be distinguished: white wines are in the left side and red wines in the right. Rosé wines (a small group) are located between the other two groups, next to white wines.

3.4. Equations development

3.4.1. Influence of the spectra preprocessing

Accuracy and precision, based on SECV and R^2 values, were similar, independent of the mathematical preprocessing employed.

3.4.2. Equations accuracy

Table 3 shows the statistic results obtained in the calibration and external validation stages (SECV and R^2). Table 4 summarises the criteria proposed by Shenk and Westerhaus [21] to compare statistic results from calibrations and the subsequent validations. NIRS provided the best statistic results for alcoholic degree and volumic mass, being the SECV values very close to those of the standard methods – namely 0.19% (v/v) and 0.33 kg/l for the determination of alcoholic degree and volumic mass, respectively. The R^2 values for the correlation between the reference and NIRS methods were 0.986 and 0.980, respectively, for the determination of these

Table 2
Reference data

Parameter	Calibration set				Validation set				SEL
	N	Range	Mean	S.D.	*N*	Range	Mean	S.D.	
Alcoholic degree (% v/v)	150	9.58–15.15	12.14	1.24	25	10.13–14.96	12.26	1.37	0.19
Volumic mass (kg/l)	150	989.5–999.3	992.9	2.1	25	990.4–999.4	993.5	2.4	0.4
Total acidity (meq/l)	150	3.55–8.72	5.42	0.92	25	4.15–8.69	5.70	1.09	0.35
Glycerol (g/l)	150	1.95–12.38	6.29	2.47	25	2.57–14.56	6.74	2.88	0.43
Total polyphenol index	150	5.0–131.0	35.3	25.4	25	6.0–92.0	36.6	25.8	3.2
Lactic acid (g/l)	130	0.06–5.32	1.36	1.10	20	0.22–5.09	1.33	1.09	0.22
Free sulphur dioxide (mg/l)	150	8.0–24.0	16.45	4.7	25	8.0–59.0	20.9	10.8	1.4

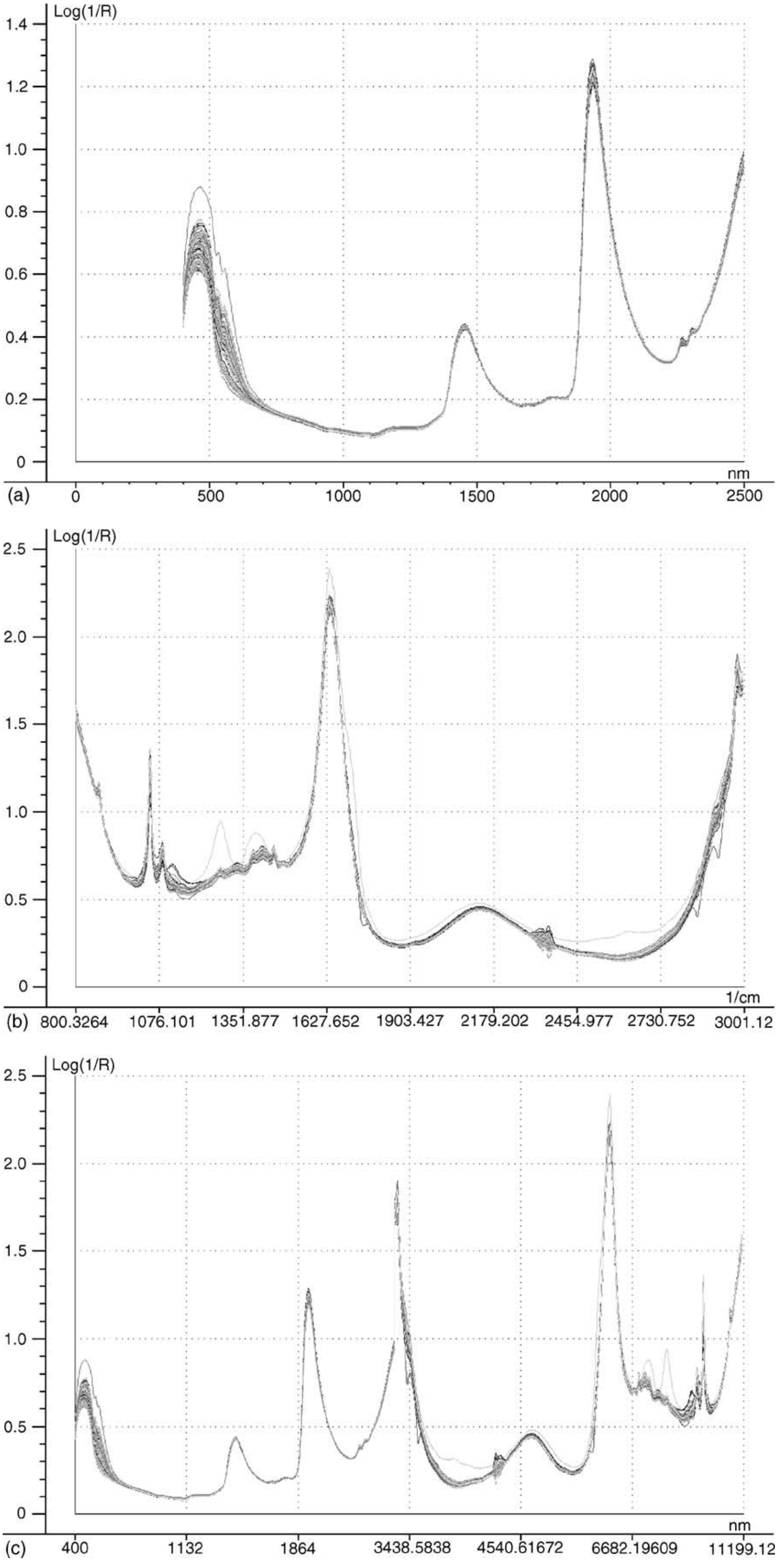

Fig. 1. (A) NIR spectra, (B) MIR spectra, and (C) combination of NIR and MIR spectra.

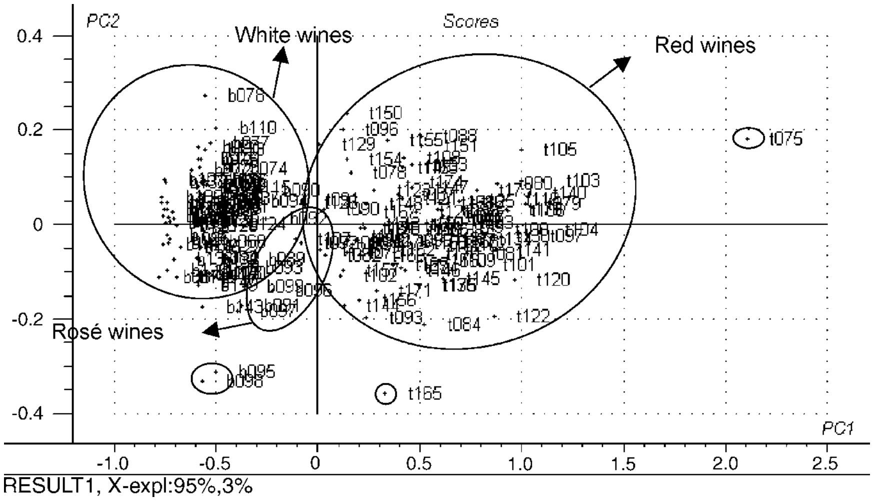

Fig. 2. Samples in the space determined by the first two principal components for NIR region.

Table 3
Results obtained with the proposed equations

Parameter	Spectrum	Calibration equation				Validation equation	
		N	Mean	SECV	R^2	R^2	SEP
Alcoholic degree (% v/v)	NIR	140	12.07	0.19	0.986	0.978	0.24
	MIR	140	12.11	0.23	0.972	0.961	0.29
	NIR–MIR	138	12.07	0.20	0.980	0.953	0.35
Volumic mass (kg/l)	NIR	141	992.8	0.33	0.980	0.917	0.54
	MIR	139	992.7	0.42	0.961	0.912	0.60
	NIR–MIR	142	992.8	0.40	0.967	0.901	0.63
Total acidity (meq/l)	NIR	139	5.41	0.38	0.845	0.812	0.48
	MIR	140	5.52	0.45	0.837	0.795	0.54
	NIR–MIR	139	5.47	0.43	0.841	0.814	0.49
Glycerol (g/l)	NIR	139	6.27	0.59	0.936	0.845	0.72
	MIR	140	6.33	0.55	0.940	0.813	0.68
	NIR–MIR	140	6.33	0.50	0.954	0.926	0.57
Total polyphenol index	NIR	138	32.31	4.50	0.975	0.919	6.70
	MIR	139	32.10	4.70	0.937	0.892	7.13
	NIR–MIR	141	32.05	4.83	0.914	0.890	7.24
Lactic acid (g/l)	NIR	125	1.38	0.35	0.860	0.814	0.41
	MIR	123	1.36	0.42	0.825	0.790	0.55
	NIR–MIR	122	1.35	0.37	0.871	0.811	0.52
Total sulphur dioxide (mg/l)	NIR	141	53.2	21.5	0.615	0.569	23.5
	MIR	140	52.3	24.3	0.573	0.520	27.0
	NIR–MIR	138	52.7	19.0	0.765	0.670	22.7

parameters. The use of the MIR region did not improve the results, as did the combination of both zones.

Total acidity was determined with good precision (the SECV and R^2 values were 0.38 meq/l and 0.845, respectively) by NIRS. The determination of this parameter by FT-MIR showed statistic results similar to NIRS (0.45 meq/l and 0.837), but not surpassed the efficiency of the NIR region. The determination of polyphenol total index also yielded similar

Table 4
Criteria for the evaluation of the results

R^2		SEP	
$R^2 \geq 0.90$	Excellent precision	SEP = 1–1.5 SEL	Excellent precision
$R^2 = 0.70–0.89$	Good precision	SEP = 2–3 SEL	Good precision
$R^2 = 0.50–0.69$	Good separation between low, medium, and high values	SEP = 4 SEL	Medium precision
$R^2 = 0.30–0.49$	Correct separation between low and high values	SEP = 5 SEL	Low precision
$R^2 = 0.05–0.29$	It is better than no analysing		

values by the two procedures proposed, as can be seen in Table 3.

Glycerol and total sulphur dioxide were determined by the combination of MIR and NIR regions in a more accurate way than by either NIR or MIR separately. It is worth emphasising the improvement achieved for the determination of total sulphur dioxide – taking into account the R^2 values, from 0.615 by NIR to 0.765 – by the combination of the two zones studied. In this way, this parameter can thus be quantitatively determined, according to the criteria showed in Table 4.

The applicability of NIRS to the most representative organic acids, namely, malic, tartaric and gluconic was limited to screening methodologies owing to the low concentration of these compounds. For this reason, these parameters are not shown in the tables. Only the determination of lactic acid, with values of R^2 higher than 0.800 and SECV values acceptable – according to the criteria shown in Table 4 – with respect to SEL value in Table 2.

3.4.3. *External validation*

The equations were tested with the validation set (samples not used for the development of the equations). Almost all the SEP values in the external validation were within the limit value – SEP = 1.5 × SECV. Only the validation of volumic mass yielded an SEP value slightly upper the limit; thus, the equations developed were robust.

3.4.4. *Comparison between the spectral zones employed*

The determinations carried out using the NIR region yielded generally better statistic results than those using the MIR region. This behaviour is due to the low signal/noise ratio in the MIR spectra caused by both instrumental and technical characteristics and the high intensity of the absorption bands corresponding to the strength vibration of the –OH group. In spite of the best efficiency of NIRS, drastic differences were not found.

The combination of the two spectral zones improved the determination for glycerol and total sulphur dioxide. This combination is particularly useful for the later as it provides a quantitative method and not a screening method as occurs when the spectral regions are used separately. The use of a larger dataset by combination of the two spectral zones (2100 variables) surpasses results obtained from the data of the MIR region.

4. Conclusions

A study of the applicability of the spectroscopic techniques in the near and mid infrared zones to the determination of wine parameters has been studied in this work. Thus, the spectroscopic methods developed enable multidetermination for alcoholic degree, volumic mass, total acidity, total polyphenol index, glycerol and total sulphur dioxide, surpassing the limitations of standard and reference methods regarding to time, reagent consumption, operational errors, etc.

The simplicity of the methods developed is similar for both spectral regions. The calibration stage, and development of the equations is the limiting step taking into account the time necessary for calibration. After this step, the determination of the parameters in a single analysis is carried out in 2 min. In addition, the equations were developed with the aim of covering a wide range of wines; for which wines from different appellation d'origine and grape varieties were used.

The study carried out shows that NIRS results were better than those obtained by FT-MIR due to the high signal/noise ratio of the latter. In addition, the combination of both spectral zones has been studied for the first time. The availability of a larger number of variables (wavelengths) than when used separately surpasses the noise of MIR spectra for two parameters; namely, glycerol and total sulphur dioxide. The improvements achieved are the key for the quantitative determination of the latter. The equations provided for each zone of the spectrum separately can only be used for screening.

Acknowledgement

The Comisión Interministerial de Ciencia y Tecnología (CICyT) is thanked for financial support (Project AGL2000-0321-P4-03).

References

[1] M. Valcárcel, M.D. Luque de Castro, Flow Injection Analysis: Principles and Applications, Ellis Horwood, Chichester, 1987.

[2] M.D. Luque de Castro, J.L. Luque García, Acceleration and Automation of Solid Sample Treatment, Elsevier, Amsterdam, 2002.

[3] L. Yang, V. Colombini, P. Maxwell, Z. Mester, R.E. Sturgeon, J. Chromatogr. A 1011 (2003) 135.

[4] M. Bonoli, M. Montanucci, T.G. Toschi, G. Lercker, J. Chromatogr. A 1011 (2003) 163.

[5] G. Alvarez Llamas, M.D.R. Fernández de la Campa, A. Sanz Medel, J. Anal. Atom. Spectrom. 18 (2003) 460.

[6] E. Trullols, I. Ruisánchez, F.X. Rius, Trends Anal. Chem. 23 (2004) 137.

[7] R. Muñoz-Olivas, Trends Anal. Chem. 23 (2004) 203.

[8] G. Kos, H. Lohninger, R. Krska, Anal. Chem. 75 (2003) 1211.

[9] J.L. Pérez-Pavón, M. Del Nogal-Sánchez, C. García-Pinto, M.E. Fernández Laespada, B. Moreno-Cordero, A. Guerrero-Peña, Anal. Chem. 75 (2003) 2034.

[10] E. Mataix, M.D. Luque de Castro, Analyst 123 (1998) 1547.

[11] E. Mataix, M.D. Luque de Castro, Talanta 51 (2000) 489.

[12] M.C. García-Parrilla, G.A. González, F.J. Heredia, A.M. Troncoso, J. Agric. Food Chem. 45 (1997) 3487.

[13] G.J. Martin, C. Guillou, M.L. Martin, J. Agric. Food Chem. 36 (1998) 316.

[14] M. Urbano-Cuadrado, M.D. Luque de Castro, P.M. Pérez-Juan, J. García-Olmo, M.A. Gómez-Nieto, Anal. Chim. Acta. 527 (2004) 81.

[15] C.M. García-Jares, B. Medina, Fresenius J. Anal. Chem. 357 (1997) 86.

[16] R. Eberl, J. Near Infrar. Spectrosc. 6 (1998) 133.

[17] A.K. Kupina, A. Shrikhande, J. Am. J. Enol. Viticult. 54 (2003) 131.

[18] M. Dubernet, Mathieu Dubernet, Rev. Franç. D'Oenolog. 81 (2000) 10.

[19] B.G.M. Vandeginsten, D.L. Massart, S. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam, 1998.

[20] K.H. Esbensen, Multivariate Data Analysis – in Practice, Camo Process AS, Oslo, 2002.

[21] J.S. Shenk, M.O. Westerhaus, Calibration the ISI way, in: En Near Infrared Spectroscopy: The Future Waves, NIR Publications, Chichester, 1996, p. 198.