Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 527 (2004) 81–88
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

Near infrared reflectance spectroscopy and multivariate analysis in enology
Determination or screening of fifteen parameters in different types of wines

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

[a] *Department of Analytical Chemistry, University of Córdoba, Campus of Rabanales, Annex Marie Curie Building, Córdoba E-14071, Spain*
[b] *LIEC, Agrifood R&D Laboratory, Manzanares E-13200, Spain*
[c] *NIR/MIR Spectroscopy Unit, Central Service for Research Support (SCAI), University of Córdoba, Campus of Rabanales, Ramón y Cajal Building, Córdoba E-14071, Spain*
[d] *Department of Computing and Numerical Analysis, University of Córdoba. Campus of Rabanales, Albert Einstein Building, Floor 3, Córdoba E-14071, Spain*

Received 18 May 2004; received in revised form 26 July 2004; accepted 26 July 2004
Available online 2 October 2004

Abstract

A study of the feasibility of near infrared reflectance spectroscopy (NIRS) for analytical monitoring in wineries is presented, in which equations for the determination or screening of the commonest enological parameters are proposed. The training and validation sets to develop NIR general equations were built with samples (180) from different apellation d'origine, different wine types, etc. By the calibration step (partial least squares regression and cross-validation were used for multivariate calibration), major components such as ethanol, volumic mass, total acidity, pH, glycerol, colour, tonality and total polyphenol index are accurate determined by the proposed equations as compared with the reference data obtained by the official and standard methods—determination coefficients (R^2) were higher than 0.800 (and higher than 0.900 most times) and standard error cross-validation (SECV) values were close to those of the reference methods. The proposed method also offers screening capability for components such as volatile acidity (R^2 = 0.481), organic acids (R^2 = 0.432 for malic acid, R^2 = 0.544 for tartaric acid, R^2 = 0.541 for gluconic acid)—with the exception of the accurate determination of lactic acid (0.860 and 0.35 g l^{-1} for R^2 and SECV, respectively)—reducing sugars (R^2 = 0.705) and total sulphur dioxide (R^2 = 0.615). In equations validation, the correlation between the reference and NIRS methods was tested, and slope and bias values statistically not different from 1 and 0, respectively, were obtained for most parameters.

Keywords: Near infrared spectroscopy; Wineries; Multivariate calibration

1. Introduction

The main objective of an analytical chemist is the development of methods able to extract physical, chemical and biological information from a target system. In agrifood industries, as in other areas, the new methodologies should be as close as possible to both time and operational requirements, and aims of clients regarding to product characterisation. Multiparameter approaches—that permit to determine more than one parameter in a single analysis—such as those based on chromatography, electrophoresis [1–3], or emission atomic spectrometry with multichannel detection [4,5], etc. have been widely employed.

In this sense, the use of various spectral regions and chemometrics is often aimed at obtaining multiparametric

* Corresponding author. Tel.: +34 95 7218581; fax: +34 95 7218581
E-mail address: qa1lucam@uco.es (M.D. Luque de Castro).

0003-2670/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.07.057

information, as is the case with approaches based on Fourier transform infrared [6,7]. Also Near Infrared Reflectance Spectroscopy (NIRS) is a very useful technique in this area as it enables multiparametric determination with minimum or null sample pretreatment [8,9]. The efficiency of NIRS, of particular significance in food analysis, has been tested regarding to some official and standard methods in wineries and other spirit industries. Thus, key parameters in brewing processes—namely: ethanol and sugar (fructose, glucose or residual sugars)—can be determined by both off-line and on-line approaches [10–12]. The possibility of multideterminationcompensates for both the time-consuming calibration step and the lower or similar accuracy of these methods as compared with the reference methods. The accuracy can be improved by both high reproducible reference data and homogeneous training and validation sets. The applicability of NIRS to wine parameters different from ethanol and sugars—namely, soluble solids, pH, colour and methanol—has also been studied [13].

A method for the determination of 15 enological parameters—alcoholic degree, volumic mass, total acidity, pH, volatile acidity, glycerol, total polyphenol index, reducing sugars, lactic, malic, tartaric and gluconic acids, colour, tonality, total sulphur dioxide and free sulphur dioxide—by NIRS has been developed. The objective was to check the applicability of this technique for either determination or screening of key parameters in a variety of both apellation d'origine and types of wines (red, rosé and white wines).

2. Experimental

2.1. Samples and sample preparation

Different wines—including red, rosé and white wines (98, 12 and 70 samples, respectively); young and aged wines (the proportion was 70 and 30% for young and aged wines, respectively); wines from different apellation d'origine (80 from "La Mancha", 25 from "Valdepeñas", 17 from "Alicante", 15 from "Jumilla", 18 from "Navarra" and 25 from "Madrid") and grape varieties (27 from "Cencibel", 30 from "Cabernet Sauvignon", 15 from "Cencibel-Cabernet Sauvignon", 20 from "Merlot", 15 from "Garnacha", 30 from "Bobal" and 43 from "Arien")—were used in the present study. Thus, the number of samples employed in the calibration and validation steps was 180 for all the parameters, but for organic acids (155 samples), colour (98 samples) and tonality (98 samples). The last two parameters were obtained only for red wines. These samples were used as such, because filtering, dilution, interferents removal, etc. were not required.

2.2. Apparatus and methods

The instrument employed for spectra collection was a Foss-NIR Systems 6500 System II spectrophotometer (Foss-NIR Systems Inc., Silver Spring, MD, USA) equipped with a transport module. 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 (log1/*R*) spectra were collected in duplicate.

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
pH	Potentiometry
Volatile acidity	Distillation, vapour dragging and titration with NaOH
Glycerol	Enzymatic reaction
Total polyphenol index	Folin–Ciocalteu reagent in alkaline medium
Reducing sugars	Reduction of Cu^{2+} in boiling alkaline medium
Lactic acid	High pressure liquid chromatography
Malic acid	High pressure liquid chromatography
Tartaric acid	High pressure liquid chromatography
Gluconic acid	Enzymatic kits
Colour	Absorbance sum at 420, 520 and 620 nm
Tonality	Absorbance ratio at 420 and 520 nm
Total sulphur dioxide	Hydrolysis with NaOH and iodometry in acid medium
Free sulphur dioxide	Iodometry in acid medium

On the other hand, samples were analysed in duplicate by the reference methods shown in Table 1 and standard error laboratory (SEL) was estimated from these duplicates using the following equation:

$$\mathrm{SEL} = \sqrt{\frac{\sum_{i=1}^{n}(y_{i1} - y_{i2})^2}{n}} \quad (1)$$

where n is the number of samples and y_{i1} and y_{i2} are values obtained for the replicates 1 and 2, respectively, of sample i.

2.3. Chemometric software used for data processing and statistical techniques used

WinISI software 1.50 (Infrasoft International, Port Matilda, PA, USA) was also used for data processing. The chemometric procedure consists of the following steps.

2.3.1. Root mean square (RMS) calculus

The RMS [14] was used for the study of similarity between spectra corresponding to aliquots of the same sample. The following equation was used:

$$\mathrm{RMS}\,(j) = 10^6 \times \sqrt{\frac{\sum_{i=1}^{n}(y_{ij} - \overline{y_i})}{n}} \quad (2)$$

where n is the number of wavelengths, y_{ij} is the log(1/R) for the sub-sample j at λ_i and $\overline{y}_i$ is the log(1/R) for the averaged

spectrum of a sample at λ_i. The factor 10^6 is introduced in the calculation of RMS for avoiding to work with too low values. As two spectra were collected per sample, their RMS values were equal. Thus, a unique RMS value was considered, then compared with the $RMS_{cut\ off}$, which was calculated from the individual RMS values of the set of samples using the MEAN and STD parameters and Eqs. (3) and (4).

$$\text{MEAN} = \sqrt{\frac{\sum_{j=1}^{N}(RMS_j)^2}{N}} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{N}(y_{ij} - \overline{y_i})^2}{nN}} \quad (3)$$

$$\text{STD} = \sqrt{\frac{\sum_{j=1}^{N}(RMS_j)^2}{N-1}} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{N}(y_{ij} - \overline{y_i})^2}{n(N-1)}} \quad (4)$$

where N and n are the number of samples and the number of wavelengths, respectively.

Taking into account Eqs. (3) and (4), the relationship between the values of MEAN and STD is defined by the following equation:

$$\text{STD} = \sqrt{\frac{N}{N-1}} \times \text{MEAN} \quad (5)$$

Eq. (5) can be transformed into Eq. (6) as 2 replicates per sample were obtained in this work:

$$\text{STD} = \sqrt{2} \times \text{MEAN} \quad (6)$$

The expression for calculating STD is a variance of the error that shows a χ^2 distribution. For this reason, Eq. (7) is used:

$$\text{STD}_{\text{limit}} = 1.036 \times \sqrt{\frac{\sum_{k=1}^{k=m}\text{STD}_k^2}{m}} = 1.036 \times \sqrt{\overline{\text{STD}}^2} \quad (7)$$

where m is the number of samples.

Finally, $RMS_{cut\ off}$ is obtained using $\text{STD}_{\text{limit}}$ and Eq. (5). For samples with RMS lower than $RMS_{cut\ off}$ an average of the two spectra was obtained; for samples with RMS higher than $RMS_{cut\ off}$ a third spectrum was obtained, and the two more similar were used for recalculation of the RMS. In this way, RMS values lower than the $RMS_{cut\ off}$ were also obtained in all instances.

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

PCA was required for the reduction of the number of variables showing co-linearity, thus representing the samples in a new, reduced p-dimensional space ($p < n$). Once the samples were in the new space defined by principal components, the H Mahalanobis distance was computed. According to Shenk and Westerhaus [15] criterion, sample spectra with $H > 3.0$ were considered outliers. Because of the existence of clusters of samples, PCA and H were computed for each cluster. These outliers were examined in order to know if either they provided any useful information or they must be removed.

2.3.3. Selection of the calibration and validation sets

Once the outliers without useful information had been removed, the calibration and validation sets were defined. Both sets were independent; thus, the validation set was only used for testing the equations. The percent of samples in each set were 85 and 15% for the training and validation sets, respectively. The selection of the validation set was carried out by calculating both PCA and subsequent H distance. The criterion used was samples more separated between them with the highest number of neighbours ($H <$ 0.6). This strategy was used for white, red and rosé wines separately.

2.3.4. Spectra preprocessing

Different treatments were applied to the spectra in the calibration and validation steps, namely scatter correction, derivatives (Saviztky–Golay filter with different number of points per gap and polynomial order was used with variable number of points).

2.3.5. Calibration step: cross-validation

In this step, partial least square regression (PLSR) [16] was used for developing the equations. The number of calibration groups and maximum number of PLS factors were set at 4 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 in cross-validation (SECV) and determination coefficient (R^2) were obtained.

2.3.6. Validation step

In order to validate the equations, the model was applied to spectra from the validation set, thus statistic parameters as standard error prediction (SEP), determination coefficient in validation (r^2), slope and bias were obtained. The correlation between the reference and NIR methods was studied using the criteria proposed by the Office International de la Vigne et du Vin (OIV) [17].

3. Results and discussion

3.1. Reference data

Table 2 shows information about reference data. Thus, range, mean, standard deviation (S.D.) and number of samples for the calibration and validation sets are summarised in the table, in addition to standard error laboratory (SEL). As can be seen, the range of reference values encompasses the characteristic values for a high diversity of wine. An exception is reducing sugars, whose values were only within the interval for dry and semi-dry wines but not for sweet wines.

Table 2
Reference data

Parameter	Calibration set				Validation set				SEL
	N	Range	Mean	Standard deviation	*N*	Range	Mean	Standard deviation	
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 ($kg\,l^{-1}$)	150	989.5–999.3	992.9	2.1	25	990.4–999.4	993.5	2.4	0.4
Total acidity ($meq\,l^{-1}$)	150	3.55–8.72	5.42	0.92	25	4.15–8.69	5.70	1.09	0.35
pH (pH units)	150	3.26–4.04	3.65	0.15	25	3.30–4.03	3.63	0.17	0.02
Volatile acidity ($g\,l^{-1}$)	150	0.14–0.87	0.42	0.15	25	0.19–0.82	0.45	0.20	0.08
Glycerol ($g\,l^{-1}$)	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
Reducing sugars ($g\,l^{-1}$)	150	0.65–9.78	2.19	1.24	25	0.85–14.3	2.86	3.39	0.12
Lactic acid ($g\,l^{-1}$)	130	0.06–5.32	1.36	1.10	20	0.22–5.09	1.33	1.09	0.22
Malic acid ($g\,l^{-1}$)	130	0.03–1.83	0.77	0.49	20	0.19–1.79	0.80	0.50	0.10
Tartaric acid ($g\,l^{-1}$)	130	1.54–4.64	2.59	0.44	20	1.76–4.20	2.63	0.55	0.25
Gluconic acid ($g\,l^{-1}$)	130	0.06–1.80	0.63	0.48	20	0.06–1.85	0.73	0.65	0.16
Colour (only red wines)	80	3.80–21.40	10.59	3.77	14	8.10 -16.10	11.25	2.77	1.48
Tonality (only red wines)	80	0.440–0.950	0.627	0.120	14	0.430–0.840	0.602	0.104	0.045
Total sulphur dioxide ($mg\,l^{-1}$)	150	16.0–149.0	59.9	35.4	25	19.0–204.0	69.3	43.0	13.3
Free sulphur dioxide ($mg\,l^{-1}$)	150	8.0–24.0	16.45	4.7	25	8.0–59.0	20.9	10.8	1.4

3.2. Spectral similarity

After calculating individual RMS values for each sample, the $RMS_{cut\ off}$ value obtained was 3200. Fig. 1 shows the evolution of this statistic parameter versus the sample number (no sample identifier). Four samples had an RMS value higher than the limit (samples that did not fulfil the spectral similarity control); thus, a third spectrum was collected for each outlier and the RMS values were recalculated using the two closer spectra and the remaining spectrum was deleted. This means that a spectrum of the three collected per sample was anomalous owing to operational errors. In this way, the four samples had an RMS value lower than the upper limit.

3.3. Spectral outliers

After studying and controlling the spectral similarity within samples, similarity between samples was studied with the aim of detecting spectral outliers regarding sample population. For this, the averaged spectrum per sample was considered (see Fig. 2). Significant dif-

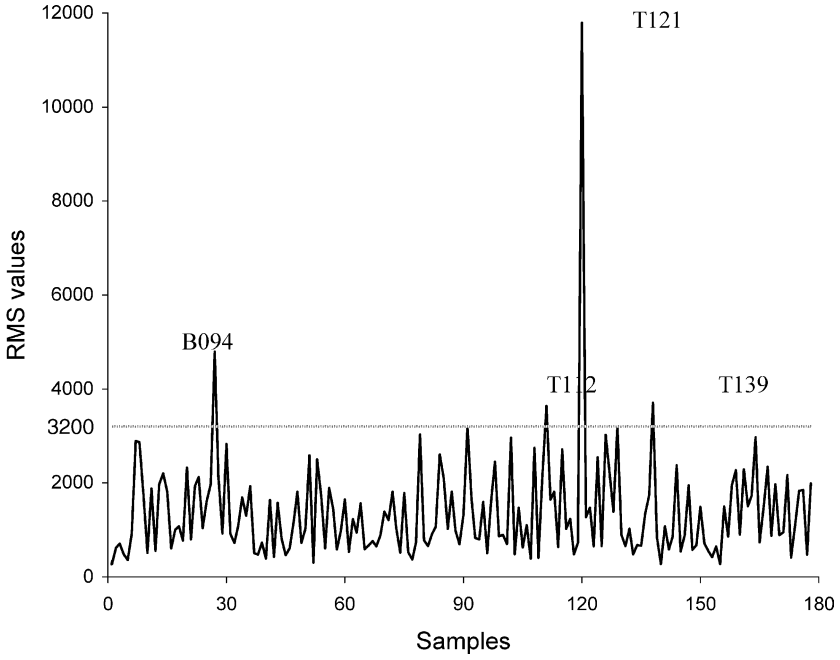

Fig. 1. RMS values vs. sample number.

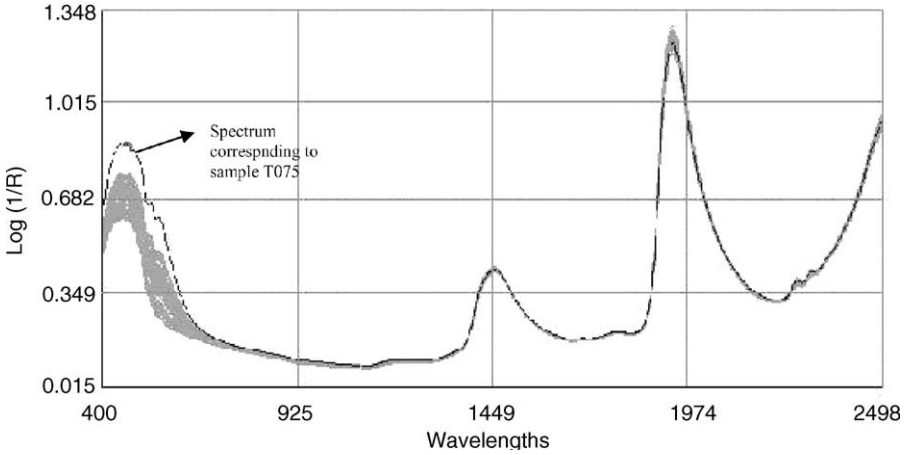

Fig. 2. NIR spectra.

ferences between spectra in the regions 400–1000 nm and 1800–2000 nm can be observed. The figure also shows that one spectrum differs significantly with the rest of spectra.

PCA was applied to the spectra. The samples plotted in the tri-dimensional space formed by the first three principal components are shown in Fig. 3. Two groups can be distinguished: white wine samples are placed in a down plane and red wine samples are in a top swarm. Rosé wines (a small group) are located between the other two groups, next to the plane corresponding to white wines. The spectrum far from the rest belongs to sample *T075* in Fig. 2.

Considering the above commented, PCA and *H*-Mahalanobis distance were computed for each of the two clear clusters in Fig. 3—white and rosé wines were considered jointly. Three spectra of red wines and one spectrum of white wine behaved as outliers. Ten and 12 principal components were used for *H* distance calculation. The criterion to fix the number of components was to obtain an increment of explained variance lower than 0.25%. On the other hand, the sum of explained variance for each model was close to 100%.

3.4. Equations development

3.4.1. Influence of spectra preprocessing

The results obtained, based on the statistic parameters described below, were similar independently of the mathematical preprocessing employed.

3.4.2. Equations calibration

The cross-validation procedure was used for equations calibration. The minimum value of SECV determined the number of PLS factors in each equation, thus avoiding ovefitting problems. The values of R^2 and SECV indicated the precision achieved in calibration. The analytical quality of the equations will be studied in the subsequent step of validation. The criteria proposed by Shenk and Westerhaus [15] based on the values of R^2 and SECV were employed in this section. Thus, R^2 values higher than 0.90 indicate excellent precision, as well as SECV values lower than 1.5 × SEL. R^2 values between 0.70 and 0.90 mean good precision, as do the SECV values between 2 × SEL and 3 × SEL. On the other hand, R^2 values lower than 0.70 indicate that the equation can only be used for screening purposes, which enable distinction between low, medium and high values for the measured

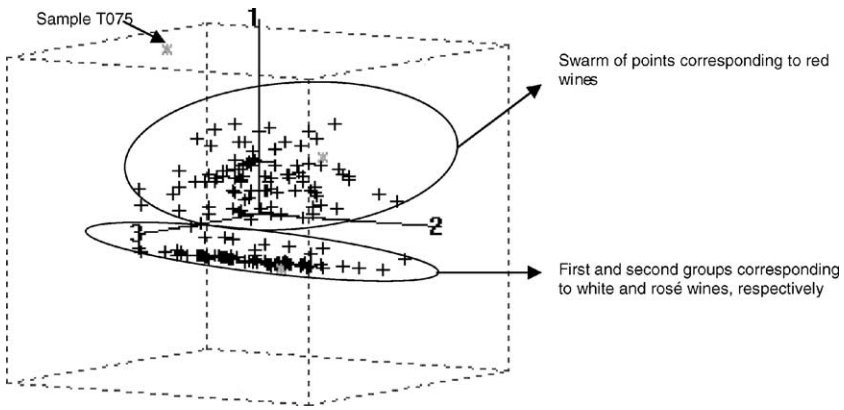

Fig. 3. Samples in the space determined by the first three principal components.

Table 3
Mean, minimum, maximum, SECV and R^2 in the calibration step

Parameter	N	PLS factors	Mean	Minimum	Maximum	SECV	R^2
Alcoholic degree (%, v/v)	140	7	12.07	9.69	15.18	0.19	0.986
Volumic mass ($kg\,l^{-1}$)	141	14	992.8	989.5	997.1	0.33	0.980
Total acidity ($meq\,l^{-1}$)	139	14	5.41	3.77	8.69	0.38	0.845
pH (pH units)	143	10	3.64	3.28	4.03	0.05	0.905
Volatile acidity ($g\,l^{-1}$)	143	10	0.40	0.17	0.82	0.09	0.481
Glycerol ($g\,l^{-1}$)	139	12	6.27	2.43	12.51	0.59	0.936
Total polyphenol index	138	11	32.31	6.17	83.21	4.50	0.975
Reducing sugars ($g\,l^{-1}$)	140	12	1.95	0.81	3.07	0.27	0.705
Lactic acid ($g\,l^{-1}$)	125	12	1.38	0.11	5.41	0.35	0.860
Malic acid ($g\,l^{-1}$)	128	7	0.73	0.08	1.72	0.34	0.452
Tartaric acid ($g\,l^{-1}$)	126	8	2.59	1.32	3.54	0.23	0.544
Gluconic acid ($g\,l^{-1}$)	128	8	0.64	0.04	1.75	0.31	0.541
Colour (only red wines)	75	7	10.29	4.18	22.8	1.52	0.820
Tonality (only red wines)	74	7	0.60	0.42	0.95	0.05	0.781
Total sulphur dioxide ($mg\,l^{-1}$)	141	12	53.2	17.0	138.2	21.5	0.615

parameter. If the R^2 value is lower than 0.50, the equation only discriminates high and low values.

Table 3 shows the results obtained in equations calibration. Thus, the number of samples used after outliers removal (using the Student's test), number of PLS factors, mean, minimum, maximum, SECV and R^2 are summarised in this table. The best results were achieved for the determination of alcoholic degree and volumic mass, being the SECV values very close to those of the standard methods—namely 0.19% (v/v) and $0.33\,kg\,l^{-1}$ 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 for the determination of these two parameters.

There are three parameters in wine related to acidity: total acidity, pH, and volatile acidity. The first two were determined with good precision (the SECV values were $0.38\,meq\,l^{-1}$ and 0.05 pH units, respectively). The model for the volatile acidity can be used only as screening methodology to distinguish between low and high values, according to the R^2 value (0.481). The SECV value was $0.09\,g\,l^{-1}$.

Glycerol and total polyphenol index (t.p.i.) showed both good R^2 values (0.936 and 0.975, respectively) and SECV values (0.59 and $4.50\,g\,l^{-1}$, respectively). Thus, a good precision was also achieved.

The results for reducing sugars (R^2 and SECV were 0.705 and $0.27\,g\,l^{-1}$, respectively) were in between the values corresponding to the applicability into either determination or screening.

The applicability of NIRS to organic acids—lactic, malic, tartaric and gluconic—was limited to screening methodologies due to the low concentration of these compounds; except for lactic acid, with values of 0.860 and $0.35\,g\,l^{-1}$ values for R^2 and SECV, respectively. Errors involved in the cross-validation for malic, tartaric and gluconic acids were high (SECV values of 0.34, 0.23 and $0.31\,g\,l^{-1}$ for malic, tartaric and gluconic acids, respectively).

With respect to colour and tonality parameters, the results obtained were acceptable. The standard methods for these parameters are based on absorbance measurements at established wavelengths in the visible range. The spectral region was 400–2500 nm; thus, the visible spectrum was also taken into account and, for this reason, good correlation between the reference and NIRS methods was achieved. The wavelength region longer than UV–vis added to the spectrum a specific noise, which affected to R^2 and SECV values (0.820 and 1.52 for colour, and 0.781 and 0.049 for tonality).

Sulphur dioxide present in wine is divided in two fractions: the free fraction and the combined fraction or sulphur dioxide bonded to diverse organic components. The model was not appropriate for the determination of the free fraction as the sensitivity of NIRS is not high enough; so the values of this parameter are not shown in Table 4, which lists that the model yielded acceptable statistic values (R^2 and SECV were 0.615 and $21.5\,mg\,l^{-1}$, respectively) to distinguish low, medium and high values of total sulphur dioxide—that is, the sum of free and bonded sulphur dioxide. These results can be explained by both the highest concentration of combined sulphur dioxide and the suitability of NIRS for organic compounds.

3.4.3. Equations validation

The equations were tested with the validation set consisting of samples not used for calibration. Table 4 shows the number of samples used after removing outliers (from Student's test), mean, minimum, maximum, SEP, r^2, slope and bias. These statistic parameters were used for evaluating the analytical quality of the equations. The values of slope and bias parameters were useful for distinguishing systematic errors and studying the correlation between the reference and NIRS methods. Slope and bias values were evaluated for testing if they are statistically equal to 1 and 0, respectively. With this aim, the criteria proposed by the OIV [17] were used at a significance level of 0.5%. The range of non-significance is also shown in the slope and bias columns. Only the determination of volatile acidity (bias), and tartaric and gluconic

Table 4
Correlation between the reference and NIR methods in the validation step

Parameter	*N*	Mean	Minimum	Maximum	SEP	r^2	Slope	Bias
Alcoholic degree (%, v/v)	24	12.27	10.08	15.36	0.24	0.978	0.971 (0.969–1.031)	0.04 (−0.10–0.10)
Volumic mass ($kg\,l^{-1}$)	24	993.3	989.5	998.5	0.54	0.917	1.001 (0.973–1.027)	−0.09 (−0.30–0.30)
Total acidity ($meq\,l^{-1}$)	24	5.68	4.12	8.65	0.48	0.812	0.986 (0.789–1.211)	0.09 (−0.21–0.21)
pH (pH units)	24	3.62	3.22	3.91	0.07	0.819	0.989 (0.916–1.084)	−0.02 (−0.03–0.03)
Volatile acidity ($g\,l^{-1}$)	24	0.44	0.25	0.72	0.14	0.345	0.730 (0.582–1.418)	0.12 (−0.08–0.08)
Glycerol ($g\,l^{-1}$)	24	6.31	2.89	10.67	0.72	0.845	0.871 (0.710–1.290)	−0.29 (−0.32–0.32)
Total polyphenol index	24	32.14	5.54	67.14	6.70	0.919	0.918 (0.894–1.106)	−2.15 (−2.83–2.83)
Reducing sugars ($g\,l^{-1}$)	23	1.86	1.37	2.65	0.33	0.712	0.983 (0.929–1.071)	0.12 (−0.14–0.14)
Lactic acid ($g\,l^{-1}$)	19	1.40	0.03	2.87	0.41	0.814	0.941 (0.847–1.143)	−0.05 (−0.20–0.20)
Malic acid ($g\,l^{-1}$)	19	0.73	0.33	1.17	0.36	0.441	0.910 (0.682–1.318)	0.03 (−0.16–0.16)
Tartaric acid ($g\,l^{-1}$)	19	2.55	2.19	3.61	0.39	0.428	0.675 (0.754–1.246)	−0.09 (−0.17–0.17)
Gluconic acid ($g\,l^{-1}$)	19	0.72	0.07	1.82	0.38	0.498	0.701 (0.809–1.191)	0.13 (−0.15–0.15))
Colour (only red wines)	14	10.57	7.90	13.71	1.83	0.705	0.807 (0.725–1.275)	−0.91 (−1.07–1.07)
Tonality (only red wines)	14	0.65	0.45	0.83	0.06	0.729	0.95 (0.925–1.075)	0.06 (−0.08–0.08)
Total sulphur dioxide ($mg\,l^{-1}$)	23	63.16	23.24	112.11	23.5	0.569	0.845 (0.698–1.302)	1.62 (−8.24–8.24)

acids (slope) yielded values out of the limits. These values are bolded in Table 4.

Although almost all slope and bias values were within the non-significance range, this was wider for parameters with r^2 and SEP values, that only enable screening (namely, volatile acidity, tartaric and gluconic acids and total sulphur dioxide). The slopes for correlation were always lower than 1, but for volumic mass. This means that the NIRS values are systematically higher than those obtained by the reference methods, taking into account that the NIRS values correspond to abscissa axis in correlation plots.

On the other hand, almost all the SEP values in the external validation were within the limit value: SEP = 1.5 × SECV. Only the validation of volumic mass and tartaric acid yielded SEP values slightly upper the limit, thus, the equations developed were robust.

3.4.4. Comparison with the results obtained by other authors

The determination of alcoholic degree, volumic mass, pH and glycerol yielded R^2, SECV and SEP values close to the values in the literature [11–13]. The error obtained in the determination of sugars and colour was higher [11,13], but a variety of both apellation d'origine and types of wines were used for establishing the equations with the aim of obtaining a general approach. The determination of organic acids, volatile acidity, total acidity, t.p.i. and total sulphur dioxide had not been reported previously.

4. Conclusions

The applicability of NIRS to the evaluation of 16 enological parameters in wine has been studied in this work. The results have been compared with the values obtained by other authors, when available and quite similar values were obtained in spite of the fact the calibration and validation sets were more heterogeneous than those involved previous approaches [11–13]. Thus, the calibration and validation of the equations were carried out with a variety of both apellation d'origine and types of wines. The final equations were developed for the determination of 15 parameters.

Thus, the most remarkable aspects of this work are the evaluation of the applicability of NIRS to the quantitative analysis in a wide variety of wines and the high number of enological parameters, which can be determined.

Acknowledgement

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

References

[1] Q.J. Wang, H. Yu, H. Li, F. Ding, P.G. He, Y.Z. Fang, Food Chem. 83 (2003) 311.
[2] M. Bonoli, M. Montanucci, T.G. Toschi, G. Lercker, J. Chromatogr. A 1011 (2003) 163.
[4] G. Álvarez Llamas, M.D.R. Fernández de la Campa, A. Sanz Medel, J. Anal. At. Spectrom. 18 (2003) 460.
[5] F.M. Pennebaker, M.B. Denton, Appl. Spectrosc. 55 (2001) 504.
[6] A.K. Kupina, A. Shrikhande, J. Amer. J. Enol. Viticult. 54 (2003) 131.
[7] M. Dubernet, M. Dubernet, Revue Française d'Oenologie 81 (2000) 10.
[8] E. Albanell, G. Caja, X. Such, M. Rovai, A.A.K. Salama, R. Casals, J. AOAC Int. 86 (2003) 746.
[9] F.E. Barton, D.E. Akin, W.H. Morrison, A. Ulrich, D.D. Archibald, J. Agric. Food Chem. 50 (2002) 7576.
[10] Y. Li, C. Brown, J. Near Infrared Spectrosc. 7 (1999) 101.
[11] C.M. García-Jares, B. Medina, Fresen. J. Anal. Chem. 357 (1997) 86.
[12] R. Eberl, J. Near Infrared Spectrosc. 6 (1998) 133.
[13] M. Gishen, R.G. Dambergs, A. Kambouris, M. Kwiatkowski, W.U. Cynkar, P.B. Høj, I.L. Francis, Proceedings of the ninth International Conference on Near Infrared Spectroscopy, 1999, p. 917.
[14] J.S. Shenk, M.O. Westerhaus, Routine Operation, Calibration, Development and Network System Management Manual, NIR Systems Inc, Silver Spring, MD, USA, 1995.

[15] J.S. Shenk, M.O. Westerhaus, Calibration the ISI way. In Near Infrared Spectroscopy: the Future Waves, NIR Publications, Chichester, 1996, 198–202.
[16] 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.
[17] Resolution OENO 6/99. Validation Protocol for a Typical Analytical Method Compared to the OIV Reference Method. Office International de la Vigne te du Vin. http://www.oiv.int/Database/Images/Client/oeno699uk.doc.