Evaluation of Ultraviolet, Visible, and Near Infrared Spectroscopy for the Analysis of Wine Compounds

María J. MARTELO-VIDAL and Manuel VÁZQUEZ

Department of Analytical Chemistry, Faculty of Veterinary Science, University of Santiago de Compostela, Lugo, Spain

Abstract

Martelo-Vidal M.J., Vázquez M. (2014): **Evaluation of ultraviolet, visible, and near infrared spectroscopy for the analysis of wine compounds**. Czech J. Food Sci., **32**: 37–47.

Spectroscopy of UV-VIS-NIR combined with chemometric analyses was used as a non-destructive technique to build models for the quantitative characterisation of the main compounds of wine. The work in mixtures can give insight into how interferences affect the performance of calibrations in wines. Ethanol, glycerol, glucose, tartaric acid, malic acid, lactic acid, and acetic acid were evaluated as pure compounds and in mixtures. Different pre-treatments for the spectra and modelling strategies such as partial least squares (PLS) regression or Principal Component Regression (PCR) were evaluated. All pure compounds studied showed a good relationship between spectra and concentrations. However, interferences were observed in the mixtures and only good models for ethanol, tartaric acid, and malic acid were obtained. The best model was obtained in the NIR region for ethanol and in the UV region for tartaric acid and malic acid. The results indicate that NIR spectroscopy could be used as an alternative to conventional chemical methods for ethanol determination and UV spectroscopy for the determination of tartaric acid and malic acid.

Keywords: spectral analysis; glycerol; chemometrics; PLS; PCR; organic acids; ethanol

Food production demands a high level of quality. Satisfying this demand during and after production requires appropriate analytical tools for analysis. Desirable features of such tools include speed, ease-of-use, minimal or no sample preparation. These features are characteristic of a wide range of spectroscopic methods (Cozzolino *et al*. 2011d; Restaino *et al.* 2011). In the last years, wineries have invested in technology to improve the quality of wine. Fermentation monitoring is a growing need in the wine industry, which implies methods providing online information in order to assure the quality of the product at all stages of the process (Di Egidio *et al.* 2010).

The rapid determination of ethanol and reducing sugars (the sum of glucose and fructose which constitutes the bulk of residual sugars), along with volatile acidity, would have a direct technological impact on the production of wines because fermentation in this rich medium can be particularly erratic and difficult. At the final quality control and regulatory level, it is also of interest to know the contents of individual sugars in the final product (Garcia-Jares & Médina 1997). However, the methods of analysis are time consuming, laborious, costly, and inconvenient for online, rapid quality evaluation of wines. Therefore, it is necessary to develop a new and expeditious detection method for ethanol, sugars, and organic acids of wines (Liu *et al.* 2011; Rios-Corripio *et al.* 2012).

Near infrared (NIR) spectroscopy has become one of the most attractive and most frequently used methods of analysis, providing simultaneous, rapid, and non-destructive quantification of the major components in many agricultural products and plant materials (Urtubia *et al.* 2004; Cozzolino *et al.* 2006; Nicolai *et al.* 2007; Martelo-Vidal *et al.* 2013).

NIR spectroscopy has been applied to a wide array of applications from agri-food to pharmaceutical and petroleum industries. Infrared spectroscopy is a use-

ful method for the quality control and quantitative analysis of solid materials and liquids in chemical, pharmaceutical, and food industries (Kolomiets *et al.* 2010; Oliveri *et al.* 2011). In food-related applications, several reviews synthesised the current status of research. On the other hand, chemometrics builds a bridge between the methods and their application in chemistry, playing a very relevant role in spectroscopy. The combination of vibrational spectroscopy and chemometrics provides calibration models for specific complex-matrix analyses and classification and/or discrimination tools, as it is also suitable to handle dimensional overload, collinearity, spectral interferences and spectral noise on vibrational spectra, thus providing good data-acquisition and data-processing methods (Moros *et al.* 2010; Rohman *et al.* 2011).

Food products are mainly composed of water, carbohydrates, proteins, fats, and other constituents that are present at low concentrations, for example vitamins and minerals. All these compounds may contribute to the shape or the absorbance spectrum obtained in the UV/VIS and NIR region. Although the major compounds (water, carbohydrates, proteins, fats) dominate because constituents present at concentrations below approximately 0.1% (w/w) are difficult to detect in water-rich systems (Queji *et al.* 2010; Cozzolino *et al.* 2011d).

The major constituent of fruits, fruit juices and wine is water, thus the NIR spectrum is dominated by the water peaks, reducing prediction accuracy for constituents that are present in relatively low concentrations, particularly if their spectra overlap with that of water. The second most dominant analyte is sugar, expressed as total reducing sugar, or as individual sugars such as glucose (Nicolai *et al.* 2007; Walsh & Kawano 2009). Ethanol has a strong NIR absorbance signal in alcoholic beverages, usually the second only to water, but accuracy and robustness of calibrations can be limited by matrix variations, particularly variations in sugar concentrations (Cozzolino *et al.* 2011b). On the other hand, the determination of organic acids in biomaterials is usually performed by high-performance liquid chromatography (HPLC), but these methods are time-consuming and cost-intensive. In recent years, new applications involving the determination of other minor compounds (volatile compounds, elements and amino acids) in plant materials have been also reported (Cozzolino *et al.* 2011c).

Compared to traditional methods, multivariate data analysis combined with modern UV-VIS-NIR instrumental techniques (Urbano-Cuadrado *et al.* 2005; Riovanto *et al.* 2011) gives a new and a better insight into complex problems by measuring a great number of chemical compounds at once, thus enabling the fingerprinting of each sample. These methods are attractive due to their inherent features of versatility, flexibility, effectiveness, and richness of information (Cozzolino *et al.* 2011c).

Partial least squares (PLS) regression is a method for constructing predictive models when the factors are numerous and highly collinear. The general idea of PLS is an attempt to extract as much latent factor variation as possible while modelling several responses well (Tobias 1995). In PLS regression an orthogonal basis of latent variables is constructed one by one in such a way that they are oriented along the directions of maximal covariance between the spectral matrix and the response vector (Wold *et al.* 2001). Principal component regression (PCR) is a widely used regression model for data having a high degree of covariance in the independent or predictor variables, or where ill-conditioned matrices are present (Xiaobo *et al.* 2011).

PLS analysis was used as a method to extract the latent variables (LV) of the original spectral data. Thus, LV could reduce the dimensionality and compress the original spectral data and explain the variance of the original spectral data related to the chemical constituents. The regression coefficients obtained by PLS analysis are helpful to find which variables were relevant and important for the prediction of Y-variables and to obtain the calibration models (Chen *et al.* 2006; Liu *et al.* 2011).

Wine is a complex system where great interferences can be present. The work in mixtures can give an insight into how interferences affect the performance of calibrations. Therefore the aim of this paper was to assess the application of UV-VIS-NIR spectroscopy in the analysis of the main compounds of wine (ethanol, glycerol, glucose, tartaric acid, malic acid, acetic acid, and lactic acid). Models were obtained using an *in-vitro* approach where the spectra of aqueous solutions and mixtures of the cited compounds were obtained and evaluated. Interferences and interactions in the spectra of the main compounds of wine were determined by UV/VIS/NIR spectroscopy.

MATERIAL AND METHODS

Samples. The *in-vitro* solutions of pure compounds and mixtures of ethanol, glycerol, glucose, tartaric acid, malic acid, acetic acid, and lactic acid were

prepared and the spectra were measured. The calibration was done with pure compounds. Seven *in vitro* solutions were prepared for each compound. The ranges of studies for pure solutions were (in g/l): ethanol 50–160, glycerol 0–12, glucose 0–10, tartaric acid 0–10, malic acid 0–5, lactic acid 0–5, and acetic acid 0–5. These *in vitro* solutions were prepared from standard solutions. All standard solutions and samples were prepared with distilled water in aseptic plastic tubes using micropipettes and HPLC quality reagents.

Mixtures of the wine compounds were also prepared. The ranges of studies for mixtures are shown in Table 1. An experimental design with 152 samples of mixtures was performed following a Central Composite Design (CCD). CCD has good design properties, little collinearity, rotatable, orthogonal blocks, insensitive to outliers and missing data. Each factor was studied at five levels. The region of operability must be greater than the region of interest to accommodate axial runs. 152 experiments combining the seven variables were performed following the design of Table 1. To validate the selected models, 60 new samples were measured.

Spectral measurements. Samples were scanned in transmittance mode in UV, VIS, and NIR regions (190–2500 nm) using a V-670 spectrophotometer. Spectral data were collected using the Spectra Manager™ II software (both Jasco Inc., Tokyo, Japan). Samples were scanned in a quartz cell with 1 mm path length and equilibrated at 33°C (Cozzolino *et al.* 2007) for 10 min before scanning. Spectral data were stored as transmittance (T) at 2 nm intervals. The samples were scanned in duplicate to obtain a total of 304 spectra of *in-vitro* mixture samples and 14 spectra for each component evaluated in pure solutions.

Spectral data pre-treatments. The spectra were exported from Spectra Manager™ II software (Jasco Inc., Tokyo, Japan) to Unscrambler software (version X 10.2; CAMO ASA, Oslo, Norway) for the chemometric analysis. Prior to the calibration, the spectral data were pre-processed for optimal performance (Liu *et al.* 2011). The UV/VIS and NIR spectra were transformed using different mathematical pretreatments to remove and minimise the unwanted spectral contribution (Di Egidio *et al.* 2010) and to reduce undesirable systematic noise, such as baseline variation, light scattering and enhance the contribution of the chemical composition (Chen *et al.* 2011). The pre-treatments applied were: normalisation + smoothing; normalisation + 1st derivative; normalisation + 2nd derivative; normalisation + SNV; normalisation + detrending; normalisation + baseline; normalisation + MSC; normalisation + deresolve; normalisation + noise; detrending; baseline; baseline + smoothing; baseline + 1st derivative; baseline + SNV; baseline + MSC; baseline + deresolve; baseline + noise; smoothing + SNV + detrending; normalisation + 1st derivative + baseline; normalisation + 2nd derivative + baseline; detrending + SNV + smoothing; smoothing + SNV + baseline; baseline + 1st derivative + normalisation; smoothing + SNV + baseline; baseline + SNV + smoothing; 1st derivative + MSC + SNV; 2nd derivative + MSC + SNV; MSC + SNV + 1st derivative; MSC + SNV + 2nd derivative; SNV + detrending + 1st derivative; SNV + detrending + 2nd derivative; 1st derivative + detrending + SNV; 2nd derivative + detrending + SNV.

Quantitative calibrations were developed to predict the studied compounds. For comparative purposes, two model techniques were applied: partial least-square (PLS) and principal component regression (PCR) (Urbano-Cuadrado *et al.* 2004; González-Caballero *et al.* 2010).

In development of PLS and PCR models, full cross-validation for pure *in vitro* solutions and random cross-validation with 50 segments and 6–7 for each segment for mixture *in-vitro* solutions were performed. Cross-validation was used to validate the quality and to prevent the overfitting of the calibration model (Ribeiro *et al.* 2010; Chen *et al.* 2011).

The following spectral regions and groups of peaks were tested for calibration purposes: the whole

Table 1. Levels of the CCD experimental design for 152 mixtures of wine compounds

Compounds	Levels (g/l)				
Ethanol	67.47	80	100	120	132.53
Glycerol	4.12	6	9	12	14
Glucose	0.039	1.7	4.35	7	8.66
Tartaric acid	0.01	1.2	3.1	5	6.2
Malic acid	0.03	0.5	1.25	2	2.5
Lactic acid	0.08	0.3	0.65	1	1.2
Acetic acid	0.08	0.3	0.65	1	1.2

UV/VIS/NIR spectral range (190–2500 nm); region A (spectral ranges covering different regions of higher absorption: 2261–2257, 1879–1870, 1392–1376, 856–850, 348–332, 237–224, and 202–190 nm); peak group B (spectral peaks with higher absorption): 2257, 1870, 1390, 335, 223, and 202 nm; peak group with higher absorption for each compound (2257, 1871, 1388, 856, 348, 334, 223, 209, and 202 nm for tartaric acid; 2293, 2257, 2240, 1892, 1874, 1743, 1691,1416, 1391, 1373, 1144, 250, 236, and 202 for glucose; 2261, 1876, 1392, 332, 332, 252, 232, and 208 nm for malic acid; 2257, 1870, 1388, 348, 334, 236, 220, 208, and 202 nm for lactic acid; 2257, 1872, 1687, 1390, 239, 224, 210, and 202 nm for glycerol; 2257, 1871, 1690, 1391, 239, 224, 210, and 202 nm for acetic acid and 2261, 1889, 1666, 1395, 300, 228, and 202 nm for ethanol), NIR region (780–2500 nm); VIS region (400–800 nm) and UV region (190–400 nm).

Model assessment and predictive capability was evaluated by the following indices: determination coefficient (r^2) (Davrieux *et al.* 2010; Chen *et al.* 2011) and root mean square error of prediction (RMSEP) (Castritius *et al.* 2010; Liu *et al.* 2009). Generally, a good model should have high determination coefficient and low RMSEP. The r^2 should be close to the value 1 (Chen *et al.* 2011). Thus, r^2 values higher than 0.90 indicate excellent precision, values between 0.70 and 0.90 mean good precision and on the other hand, values lower than 0.70 indicate that the equation can be used only for screening purposes, which enable distinction between low, medium and high values for the measured parameter (Urbano-Cuadrado *et al.* 2004; Novales *et al.* 2009).

RESULTS AND DISCUSSION

Raw UV/VIS/NIR spectra. Any water present in the samples dominates spectra of natural products. For this reason, quantitative analysis often relies on minor changes in spectra (García-Jares & Médina 1997). Therefore the determination of the studied compounds in aqueous solutions had to be based on very small differences.

Figure 1 shows the spectra of ethanol solutions. It was observed that the transmittance in several zones changes with the ethanol concentration. The plot shows the principal amplified regions of ethanol absorptions in 1600–1900 nm region related with O-H combinations (Xiaobo *et al.* 2010) and C-H stretch first overtones (Liu *et al.* 2008; Xiaobo *et al.* 2010) and 2200–2300 nm related with C-H vibrations (Osborne *et al.* 1993; Dambergs *et al.* 2002; Cozzolino *et al.* 2003).

The spectra of other compounds in aqueous solutions had minor variations in the transmittance. In the case of glucose, the principal variations were in the UV region and in 1200 nm and 1450 nm (NIR region). However, these variations were very slight. In the case of glycerol and organic acids, the principal variations were in the UV region. Water overlapped the absorption of the constituents of solutions and made the visualisation of the compound variations difficult.

Figure 2 shows the UV/VIS/NIR raw spectra of mixtures with all the main wine compounds. In the NIR region, the spectrum was dominated by an absorption band at around 1200 nm region related with sugars (González-Caballero *et al.* 2010). Water-related absorption bands were also found at around 950 nm and 1460 nm, which are related to the third overtone of O-H, as it is usually the case for fruits and vegetables and their juices, with 70–80% of water (Murra 1987; McGlone & Kawano 1998; Williams 2001). Spectral variations at 990 nm were produced by the O-H stretch second overtones from sugars and organic acids. The absorption bands at 1450 and 1950 nm were related to the first overtone of the O-H stretch (Liu *et al.* 2008) and a combination

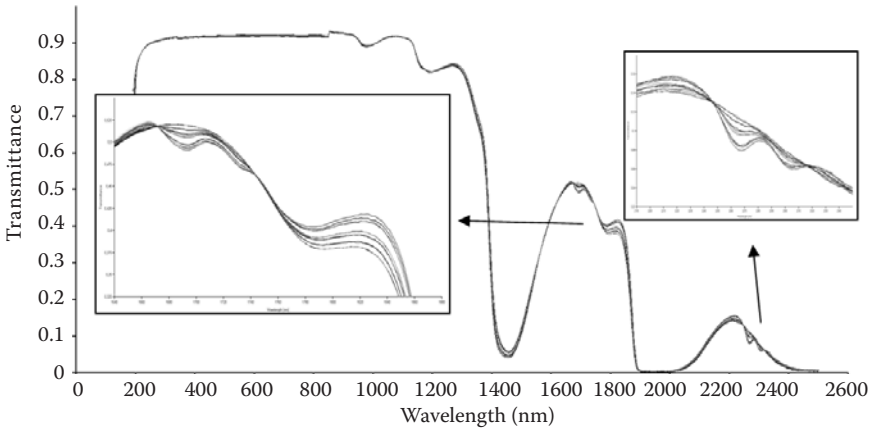

Figure 1. Spectra of pure ethanol solutions

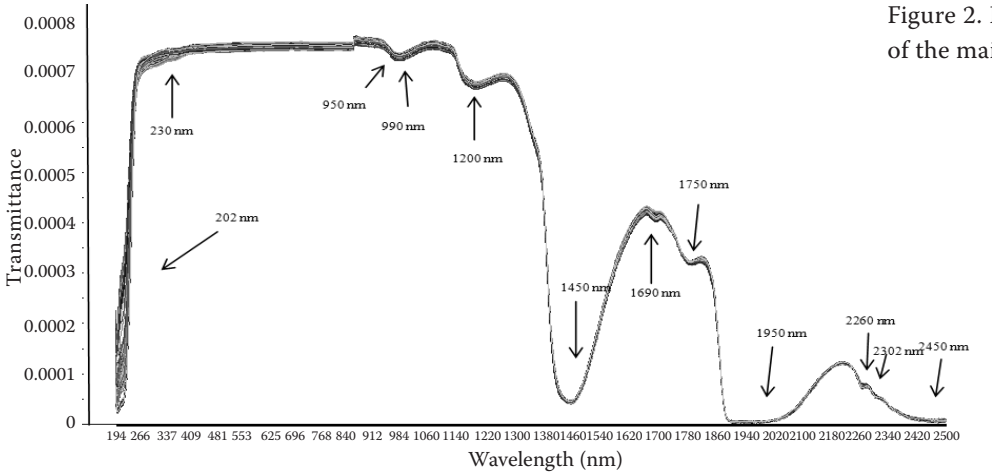

Figure 2. Raw spectra of mixtures of the main wine compounds

of stretch and deformation combination of O-H group in water, glucose and ethanol (Da Costa Filho 2009; Ferrari *et al.* 2011; Liu *et al.* 2011). Absorptions were observed at 1690 nm related with either C-H_3 stretch first overtone and 1750 nm, related to C-H_2, C-H stretch first overtones in glucose and ethanol (Liu *et al.* 2008; Ferrari *et al.* 2011). Absorption at 2260 nm was likely related with C-H combinations and O-H stretch overtones, the latter for glucose. Absorption at 2302 nm was mainly related with C-H combination vibrations (CH_3 and CH_2) of ethanol, carbohydrates and organic acids. In the UV region, 202 and 230 nm were the peaks with the highest responses. These are relative to carboxyl groups of organic acids (Shen *et al.* 2010; Cozzolino *et al.* 2011a; Ferrari *et al.* 2011).

Development of multivariate calibration models for pure wine compounds. UV-VIS and NIR spectral data were correlated with glucose, glycerol, ethanol or organic acid (tartaric, malic, lactic, and acetic acids) concentrations in pure solutions using PLS regression and PCR. The calibrations and cross-validation statistics for pure compounds of wine are shown in Table 2. The calibrations were performed for raw and several pre-treated data (smoothing; smoothing and 2nd derivative; smoothing, 2nd derivative and SNV) and models evaluated by the values of r^2 and RMSEP.

The PLS model for predicting glucose showed the best results. The pre-treatment used in this case was smoothing and 2nd derivative. The results obtained showed the value of r^2 0.999 and RMSEP 0.125 g/l. The model used two principal components (2 PC) that explain 100% of the variation of samples.

For glycerol concentration, the best model was PLS, showing the values for r^2 0.999 and for RMSEP 0.112 g/l. The model with 2 PC explains 100% of the variation of samples. Ethanol model was PLS with pre-treatment of smoothing and 2nd derivative. The model used 2 PC and explains 100% of the variation of data. The value of r^2 was 0.999 and that of RMSEP 1.719 g/l.

For tartaric acid, the best model was PLS with pre-treatment of smoothing, 2nd derivative and SNV. The results showed a high level of precision with r^2 of 0.999 and RMSEP of 0.066 g/l. The model used 3 PC that explain 100% of the variation. For acetic acid concentration, the PLS model used 2 PC without pre-treatment. The value of r^2 was 0.999 and RMSEP was 0.059 g/l. This model explains 100% of the variation of samples. For malic acid, the best model used 2 PC that explain 100% of the variation and raw data. The best model was obtained by PLS. The value of r^2 was 0.998 and RMSEP 0.0916 g/l. Finally, the best model for lactic acid was PLS with r^2 of 0.998 and RMSEP of 0.093 g/l.

New samples were measured to validate the models confirming the good fitting. The results showed that each component of the wine could be modelled using the UV-VIS-NIR spectra and the models can be used for prediction. However, the mixtures of these compounds, like in the wine, can imply interferences in the spectra and the models could not be adequate. Therefore, the mixtures of the wine compounds were studied and the results are shown in the following section.

Development of multivariate calibration models for mixtures of wine compounds. The models developed with all compounds in the whole UV/VIS/NIR region correlated well with the spectra of tartaric acid, malic acid, and ethanol in PLS regression models and PCR models. The better prediction results for calibration sets generated by the best models are shown in Table 3. The predictive models for the

Table 2. Calibration and cross-validation statistics for determination of pure compounds of wine by UV-VIS-NIR transmittance (whole spectra) performed on solutions containing only one compound each

	Pre-treatments	Model	r^2	RMSE (g/l)
Tartaric acid	raw	PLS	0.997	0.198
		PCR	0.997	0.207
	smoothing	PLS	0.997	0.201
		PCR	0.996	0.210
	smoothing + 2nd derivative	PLS	0.995	0.250
		PCR	0.995	0.253
	smoothing + 2nd derivative + SNV	PLS	0.999	0.066
		PCR	0.999	0.104
Glucose	raw	PLS	0.987	0.542
		PCR	0.987	0.546
	smoothing	PLS	0.987	0.540
		PCR	0.987	0.543
	smoothing + 2nd derivative	PLS	0.999	0.125
		PCR	0.986	0.560
	smoothing + 2nd derivative + SNV	PLS	0.997	0.231
		PCR	0.996	0.298
Malic acid	raw	PLS	0.998	0.092
		PCR	0.998	0.098
	smoothing	PLS	0.998	0.092
		PCR	0.998	0.098
	smoothing + 2nd derivative	PLS	0.997	0.105
		PCR	0.997	0.098
	smoothing + 2nd derivative + SNV	PLS	0.994	0.151
		PCR	0.994	0.153
Lactic acid	raw	PLS	0.998	0.093
		PCR	0.996	0.125
	smoothing	PLS	0.998	0.094
		PCR	0.996	0.125
	smoothing + 2nd derivative	PLS	0.997	0.103
		PCR	0.997	0.104
	smoothing + 2nd derivative + SNV	PLS	0.996	0.126
		PCR	0.995	0.135
Glycerol	raw	PLS	0.989	0.427
		PCR	0.989	0.428
	smoothing	PLS	0.989	0.425
		PCR	0.989	0.425
	smoothing + 2nd derivative	PLS	0.999	0.112
		PCR	0.996	0.226
	smoothing + 2nd derivative + SNV	PLS	0.999	0.136
		PCR	0.998	0.166
Acetic acid	raw	PLS	0.999	0.059
		PCR	0.998	0.079
	smoothing	PLS	0.999	0.060
		PCR	0.998	0.079
	smoothing + 2nd derivative	PLS	0.992	0.182
		PCR	0.992	0.184
	smoothing + 2nd derivative + SNV	PLS	0.998	0.088
		PCR	0.998	0.097
Ethanol	raw	PLS	0.994	4.265
		PCR	0.994	4.314
	smoothing	PLS	0.994	4.263
		PCR	0.993	4.312
	smoothing + 2nd derivative	PLS	0.999	1.719
		PCR	0.999	1.973
	smoothing + 2nd derivative + SNV	PLS	0.998	2.586
		PCR	0.997	2.911

other compounds (glucose, glycerol, acetic acid, and lactic acid) showed lower r^2 and RMSEP. Therefore, these results show that there are important interactions and similar responses in the spectra of these compounds that did not allow a good identification of the spectral response due to each compound.

New models were developed with different regions of spectra. In order to find the different regions, calibration solutions in whole spectra were used for determining the regression coefficients. The regression coefficients of PLS model in whole spectra and pre-treatments of normalization, 2nd derivative and baseline correction were evaluated (Figures 3 and 4). Ranges with high absolute regression coefficient values and also the peaks of maximum wavelengths were selected. The main criteria for selection were that the wavelength should have a high absolute regression coefficient value and should be at specific peaks and

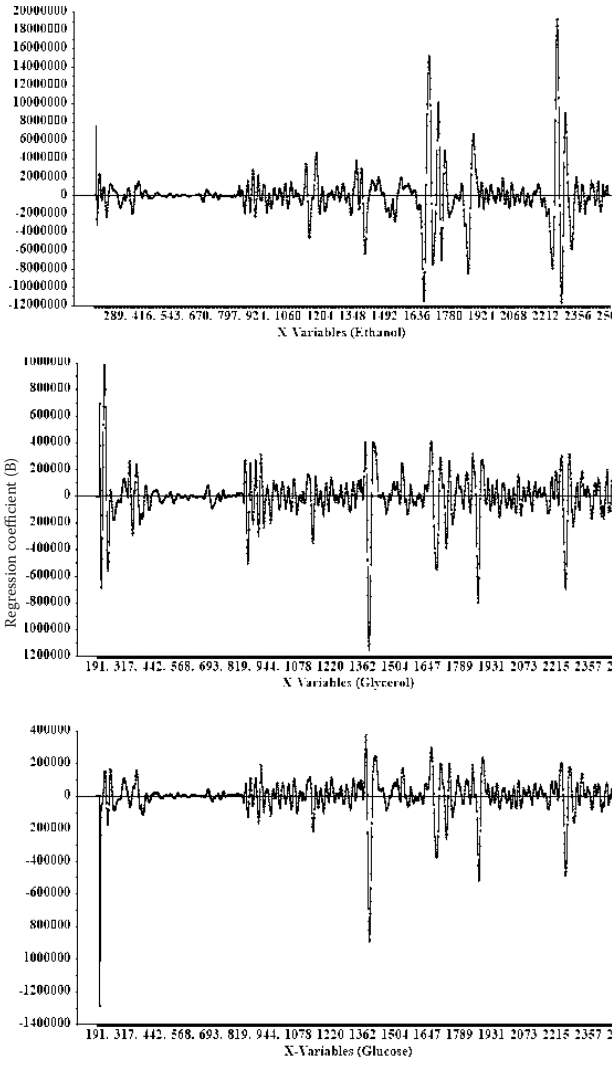

Figure 3. Regression coefficients for ethanol, glycerol, and glucose

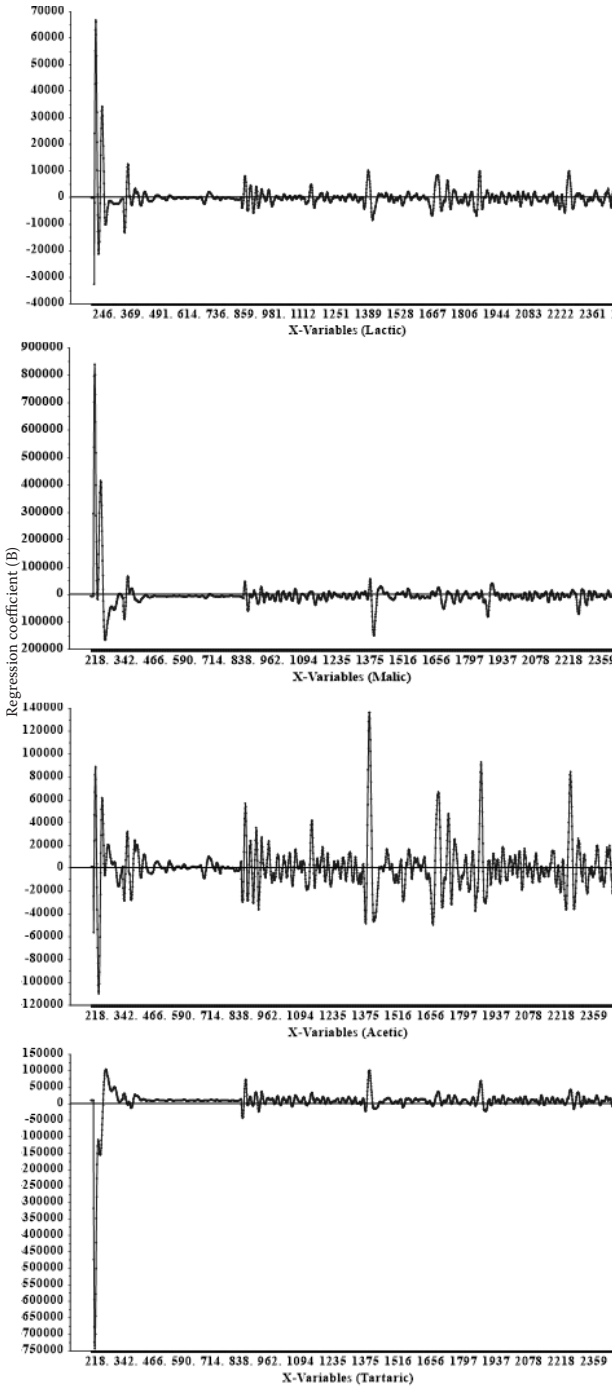

Figure 4. Regression coefficients for lactic, acetic, tartaric, and malic acid

valleys or the regression coefficient curve (Liu *et al.* 2011). It was assumed that the wavelength with a high absolute regression coefficient value could represent useful information on wavelength bands at the peaks and valleys. Therefore, ranges of wavelengths were selected as region A and group of peaks called B. These were selected by comparison of regression coefficients of each solution compounds (ethanol, glycerol, glucose, tartaric, malic, acetic, and lactic acid concentration). The common regions were selected, firstly by zones (region A), secondly by concrete wavelengths (group of peaks B). Concrete wavelengths for each compound were also studied. Table 3 shows the values of r^2 and RMSEP obtained for all models.

Models for the complete UV-VIS-NIR range were well correlated with ethanol and tartaric acid. The same was obtained for models of region A and group of peaks B. Like in the previous models, the models for the other organic acids, glucose, and glycerol had very low r^2 values. It can be concluded that these models were not good predictors of wine compounds.

Models were also obtained for concrete wavelengths for each compound. The values of r^2 and RMSE obtained were well correlated with ethanol, tartaric acid, and malic acid. For the other compounds the predictive models were badly fitted.

Models were also evaluated for the classical distribution of spectral regions in UV, VIS, and NIR regions. The results in VIS failed to provide good results to predict the wine compounds. In the UV region, the models PLS and PCR gave good correlations for tartaric acid and malic acid (Table 3). PLS regression and PCR models for the other compounds were badly fitted. The PLS regression and PCR models obtained in the NIR region were well fitted for ethanol and tartaric acid.

Sixty new unknown samples were measured with better models for ethanol, tartaric acid and malic acid. The results of prediction for tartaric acid, malic acid, and ethanol indicated that these models can be used to predict the concentration of these three wine compounds. However, the good quantitative prediction in model solutions does not imply necessarily good results in real wines.

The results showed that PLS and PCR models provided good results to predict the concentrations of each main wine compound in individual solutions using the UV/VIS/NIR spectra. The models of calibration can be used for quantitative determination of these compounds. However, in mixtures like wine, these compounds show interferences that do not allow good fits for many of them.

In mixtures, the UV/VIS/NIR models provided good predictions for ethanol, tartaric acid, and malic acid concentrations although the models were better in different regions. The optimal pre-treatments in whole spectra were normalisation + 2nd derivative + baseline to PLS model of tartaric acid, 2nd derivative + MSC + SNV to PLS model of malic acid, smoothing + SNV + detrending to PLS model of ethanol

Table 3. Calibration and cross-validation statistics for determination of tartaric acid, malic acid and ethanol in mixtures by UV-VIS-NIR transmittance. Calibrations were performed on the 152 mixtures studied and predictions on additional new 60 mixtures

Compounds	Spectral range	Pre-treatments	Model	r^2	RMSEP (g/l)
Tartaric acid	UV/VIS/NIR	normalisation + 2nd derivative + baseline	PLS	0.969	0.314
		1st derivative + MSC + SNV	PCR	0.969	0.313
	region A	normalisation + 2nd derivative + baseline	PLS	0.885	0.603
	region B		PCR	0.878	0.622
	concrete wavelength		PLS	0.957	0.362
	NIR	smoothing + SNV + detrending	PLS	0.740	0.906
			PCR	0.743	0.902
	UV	normalisation + 1st derivative + baseline	PLS	0.985	0.221
		SNV + detrending + 2nd derivative	PCR	0.982	0.239
	VIS	detrending + SNV + smoothing	PLS	0.034	1.748
		baseline + SNV + smoothing	PCR	0.005	1.775
Malic acid	UV/VIS/NIR	2nd derivative + MSC+ SNV	PLS	0.845	0.277
		1st derivative + MSC + SNV	PCR	0.849	0.273
	region A	normalisation + 2nd derivative + baseline	PLS	0.207	0.625
	region B		PLS	0.249	0.608
	concrete wavelength		PLS	0.743	0.356
	NIR	smoothing + SNV + detrending	PLS	0.101	0.666
			PCR	0.107	0.666
	UV	2nd derivative + MSC + SNV	PLS	0.926	0.191
		SNV + detrending + 2nd derivative	PCR	0.900	0.222
	VIS	normalisation + 1st derivative + baseline	PLS	0.033	0.691
		baseline + 2nd derivative + normalisation	PCR	0.007	0.61
Ethanol	UV/VIS/NIR	smoothing + SNV + detrending	PLS	0.983	2.453
		1st derivative + MSC + SNV	PCR	0.977	2.853
	region A	normalisation + 2nd derivative + baseline	PLS	0.824	7.849
	region B		PLS	0.776	8.865
	concrete wavelength		PLS	0.950	4.197
	NIR	normalisation + 1st derivative + baseline	PLS	0.989	1.949
		smoothing + SNV + detrending	PCR	0.991	1.782
	UV	baseline + 2nd derivative + normalisation	PLS	0.185	16.910
			PCR	0.145	17.320
	VIS	normalisation + 1st derivative + baseline	PLS	0.289	15.790
		1st derivative + detrending + SNV	PCR	0.008	18.650

and 1st derivative + MSC + SNV to PCR model of ethanol, tartaric, and malic acids.

Ethanol was better determined in the NIR region and tartaric and malic acids were better determined in the UV region. In the NIR region, tartaric acid gives also good results, using both region A and group of peaks B. The ethanol model obtained also good results in region A and group of peaks B, but it was better predicted using the complete NIR region. In the UV region, the results for ethanol were lower than when other spectral regions were used.

The model for malic acid obtained good results in the NIR region but it was better in the UV region. The best pre-treatment for malic acid models was normalisation + 2nd derivative + baseline in region A and group peaks B. For PLS and PCR models of tartaric acid in the NIR region the best pre-treatment was smoothing + SNV + detrending.

For PLS model of ethanol in the NIR region was normalisation + 1st derivative + baseline and for PCR model was smoothing + SNV + detrending. In the UV region, for the PLS models the pre-treatments

were normalisation + 1st derivative + baseline for tartaric acid and 1st derivative + MSC + SNV for malic acid. For PCR models in the UV region it was SNV + detrending + 2nd derivative.

The prediction accuracy was close to similar studies in wine, beer and spirit drinks. Urtubia *et al.* (2004) applied NIR to predict malic acid (0.985 for r^2 and RMSE 0.56 g/l), ethanol (r^2 0.99 and RMSE 1.04 g/l), acetic acid (r^2 0.988 and RMSE 0.42 g/l), glycerol (r^2 0.988 and RMSE 0.81 g/l), and glucose (r^2 0.994 and RMSE 1.84 g/l) in wines. Liu *et al.* (2011) used NIR to determine acetic acid (r^2 0.999 and RMSE 0.603 g/l) and tartaric acid (r^2 0.995 and RMSE 0.246 g/l) in fruit vinegars. Di Egidio *et al.* (2010) used NIR in must during fermentations to predict glucose (r^2 0.990 and RMSE 1.32 g/l), glycerol (r^2 0.990 and RMSE 0.49 g/l) and ethanol (r^2 0.990 and RMSE 2.04 g/l). Castritius *et al.* (2010) used NIR in beers to determine ethanol (r^2 0.997 and RMSE 3.19 g/l). Kolomiets *et al.* (2010) also used NIR in alcoholic beverages to determine ethanol (r^2 0.984 and RMSE 0.22 g/l). Urbano-Cuadrado *et al.* (2004, 2005) used NIR in wine fermentations to determine ethanol (r^2 0.986 and RMSE 0.436 g/l), glycerol (r^2 0.936 and RMSE 0.77 g/l), lactic acid (r^2 0.860 and RMSE 0.59 g/l), malic acid (r^2 0.452 and RMSE 0.38 g/l) and tartaric acid (r^2 0.541 and 0.47 g/l).

CONCLUSION

All pure compounds studied showed a good relationship between spectra and concentrations, giving good models for prediction. The results show that the use of PLS and PCR models to quantify *in vitro* solutions of each wine compound separately is feasible. However, in mixtures, interferences were revealed that allowed to obtain good models only for ethanol, tartaric and malic acid concentrations. Therefore, ethanol can be determined with PLS model and normalisation + 1st derivative + baseline pre-treatment in the NIR region. Tartaric acid can be determined with PLS model and normalisation + 1st derivative + baseline correction pre-treatment in the UV region. Malic acid can be determined with PCR model and SNV + detrending + 2nd derivative pre-treatment in the UV region. The results give insight into how the interferences can affect the performance of calibrations in real wines. Further studies are needed to validate the models in real wines.

References

Castritius S., Kron A., Schäfer T., Rädle M., Harms D. (2010): Determination of alcohol and extract concentration in beer samples using a combined method of near-infrared (NIR) spectroscopy and refractometry. Journal of Agricultural and Food Chemistry, **58**: 12634–12641.

Chen L., Xue X., Ye Z., Zhou J., Chen F., Zhao J. (2011): Determination of Chinese honey adulterated with high fructose corn syrup by near infrared spectroscopy. Food Chemistry, **128**: 1110–1114.

Chen J.Y., Zhang H., Matsunaga R. (2006): Rapid determination for the main organic acid composition of raw Japonese apricot fruit juices using near-infrared spectroscopy. Journal of Agricultural and Food Chemistry, **54**: 9652–9657.

Cozzolino D., Liu L., Cynkar W.U., Dambergs R.G., Janik L., Colby C.B., Gishen M. (2007): Effect of temperature variation on the visible and near infrared spectra of wine and the consequences on the partial least square calibrations developed to measure chemical composition. Analytica Chimica Acta, **588**: 224–230.

Cozzolino D., Cynkar W., Janik L., Dambergs R.G., Gishen M. (2006): Analysis of grape and wine by near infrared spectroscopy – a review. Journal of Near Infrared Spectroscopy, **14**: 279–289.

Cozzolino D., Cynkar W.U., Shah N., Smith P. (2011a): Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zeland? Food Chemistry, **126**: 673–678.

Cozzolino D., Cynkar W.U., Shah N., Smith P. (2011b): Feasibility study on the use of attenuated total reflectance mid-infrared for analysis of compositional parameters in wine. Food Research International, **44**: 181–186.

Cozzolino D., Cynkar W.U., Shah N., Smith P. (2011c): Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality. Food Research International, **44**: 1888–1896.

Cozzolino D., Cynkar W.U., Shah N., Smith P. (2011d): Technical solutions for analysis of grape juice, must, and wine: the role of infrared spectroscopy and chemometrics. Analytical and Bioanalytical Chemistry, **401**: 1475–1484.

Cozzolino D., Smyth H.E., Gishen M. (2003): Feasibility study on the use of visible and near infrared spectroscopy together with chemometrics to discriminate between commercial white wine of different varietal origins. Journal of Agricultural and Food Chemistry, **51**: 7703–7708.

Da Costa Filho P.A. (2009): Rapid determination of sucrose in chocolate mass using near infrared spectroscopy. Analytica Chimica Acta, **631**: 206–211.

Dambergs R.G., Kambouris A., Francis I.L., Gishen M. (2002): Rapid analysis of methanol in grape derived distillation products using near infrared transmission spectroscopy. Journal of Agriculture and Food Chemistry, **50**: 3079–3084.

Davrieux F., Allal F., Piombo G., Kelly B., Okulo J.B., Thiam M., Diallo O.B., Bouvet J.M. (2010): Near infrared spectroscopy for high-throughtput characterization of shea tree (*Vitellaria paradoxa*) nut fat profiles. Journal of Agricultural and Food Chemistry, **58**: 7811–7819.

Di Egidio V., Sinelli N., Giovanelli G., Moles A., Casiraghi E. (2010): NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. European Food Research and Technology, **230**: 947–955.

Ferrari E., Foca G., Vignali M., Tassi L., Ulrici A. (2011): Adulteration of the anthocyanin content of red wines: Perspectives for authentication by Fourier Transform-Near InfraRed and ^{1}H NMR spectroscopies. Analytica Chimica Acta, **701**: 139–151.

Garcia-Jares C.M., Médina B. (1997): Application of multivariate calibration to the simultaneous routine determination of ethanol, glycerol, fructose, glucose and total residual sugars in botrytized-grape sweet wines by means of near-infrared reflectance spectroscopy. Fresenius' Journal of Analytical Chemistry, **357**: 86–91.

González-Caballero V., Sánchez M.T., López M.I., Pérez-Marín D. (2010): First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. Journal of Food Engineering, **101**: 158–165.

Kolomiets O.A., Lachenmeier D.W., Hoffmann U., Siesler W. (2010): Quantitative determination of quality parameters and authentication of vodka using near infrared spectroscopy. Journal of Near Infrared Spectroscopy, **18**: 59–67.

Liu L., Cozzolino D., Cynkar W.U., Dambergs R.D., Janik L., O'Neill B.K., Colby C.B., Gishen M. (2008): Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. Food Chemistry, **106**: 781–786.

Liu F., He Y., Wang L., Sun G. (2011): Detection of organic acids and pH of fruit vinegars using near-infrared spectroscopy and multivariate calibration. Food Bioprocess Technology, **4**: 1331–1340.

Liu Y., Sun X., Ouyang A. (2009): Nondestructive measurement of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN. LWT-Food Science and Technology, **43**: 602–607.

Martelo-Vidal M.J., Domínguez-Agis F., Vázquez M. (2013): Discrimination between subzones into Spanish designation of origin "Rías Baixas" wines using UV/VIS/NIR spectral analysis and chemometrics tools. Australian Journal of Grape and Wine Research, **19**: 62–67.

McGlone V.A., Kawano S. (1998): Firmness, dry matter and soluble solids assessment of postharvest kiwifruit by NIR spectroscopy. Postharvest Biology and Technology, **13**: 131–141.

Moros J., Garrigues S., De la Guardia M. (2010): Vibrational spectroscopy provides a green tool for multicomponent analysis. Trends in Analytical Chemistry, **29**: 578–591.

Murra I. (1987): The NIR spectra of homologous series of organic compounds. In: Hollow J., Kaffka K.J., Gonczy L.L. (eds): Near Infrared Diffuse Reflectance/Transmittance Spectroscopy. Proceedings International NIR/NIT Conference, Akademiai Kiado, Budapest: 13–18.

Nicolai B.M., Beullens K., Bobelyn E., Peirs A., Saeys W., Theron L.I. (2007): Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, **46**: 99–118.

Novales J.F., López M.I., Sánchez M.T., Morales J. (2009): Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines. Food Research International, **42**: 285–291.

Oliveri P., Di Egidio V., Woodcock T., Downey G. (2011): Application of class-modelling techniques to near infrared data for food authentication purposes. Food Chemistry, **125**: 1450–1456.

Osborne B.G., Fearn T., Hindle P.H. (1993): Practical NIR spectroscopy with applications in food and beverage analysis. Harlow, Longman Scientific and Technical, Essex.

Queji M.D., Wosiacki G., Cordeiro G.A., Peralta-Zamora P.G., Nagata N. (2010): Determination of simple sugars, malic acid and total phenolic compounds in Apple pomace by infrared spectroscopy and PLSR. International Journal of Food Science and Technology, **45**: 602–609.

Restaino E., Fassio A., Cozzolino D. (2011): Discrimination of meat patés according to the animal species by means of near infrared spectroscopy and chemometrics. CyTA – Journal of Food, **9**: 210–213.

Ribeiro J.S., Ferreira M.M.C., Salva T.J.G. (2010): Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. Talanta, **83**: 1352–1358.

Rios-Corripio M.A., Rojas-López M., Delgado-Macuil R. (2012): Analysis of adulteration in honey with standard sugar solutions and syrups using attenuated total reflectance-Fourier transform infrared spectroscopy and multivariate methods. CyTA – Journal of Food, **10**: 119–122.

Riovanto R., Cynkar W.U., Berzaghi P., Cozzolino D. (2011): Discrimination between Shiraz wines from different Australian regions: The role of spectroscopy and chemometrics. Journal of Agricultural and Food Chemistry, **59**: 10356–10360.

Rohman A., Che Man Y.B., Hashim P., Ismail A. (2011): FTIR spectroscopy combined with chemometrics for

analysis of lard adulteration in some vegetable oils. CyTA – Journal of Food, **9**: 96–101.

Shen F., Niu S., Yang D., Ying Y., Li B., Zhu G., Wu J. (2010): Determination of amino acids in Chinese rice wine by Fourier transform near-infrared spectroscopy. Journal of Agricultural and Food Chemistry, **58**: 9809–9816.

Tobias R.D. (1995): An introduction to partial least squares regression. In: Proceedings of the 12th Annual SAS Users. SAS Institute Inc., Cary.

Urbano-Cuadrado M., Luque de Castro M.D., Pérez-Juan P.M., García-Olmo J., Gómez-Nieto M.A. (2004): Near infrared reflectance spectroscopy and multivariate analysis in enology. Determination or screening of fifteen parameters in different types of wines. Analytica Chimica Acta, **527**: 81–88.

Urbano-Cuadrado M., Luque de Castro M.D., Pérez-Juan P.M., Gómez-Nieto M.A. (2005): Comparison and joint use of near infrared spectroscopy and Fourier transform mid infrared spectroscopy for the determination of wine parameters. Talanta, **66**: 218–224.

Urtubia A., Pérez-Correa J. R., Meurens M., Agosin E. (2004): Monitoring large scale wine fermentations with infrarred spectroscopy. Talanta, **64**: 778–784.

Walsh K.B., Kawano S. (2009): Near infrared spectroscopy. In: Zude M. (ed.): Optical Monitoring of Fresh and Processed Agricultural Products. CRC Press, Taylor and Francis Group, Boca Raton: 192–239.

Williams P.C. (2001): Implementation of near-infrared technology. In: Williams P.C., Norris K.H. (eds): Near-Infrared Technology in the Agricultural and Food Industries. 2nd Ed. AACC Inc., St. Paul: 145–169.

Wold S., Sjöstrom M., Eriksson L. (2001): PLS-Regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory System, **58**: 109–130.

Xiaobo Z., Jiewen Z, Povey M.J.W., Holmes M., Hanpin M. (2010): Variables selection methods in near-infrared spectroscopy. Analytica Chimica Acta, **667**: 14–32.

Received for publication April 5, 2013
Accepted after corrections June 21, 2013

Corresponding author.

Dr Manuel Vázquez, Universidade de Santiago de Compostela, Facultade de Veterinaria, Química Analítica, Nutrición e Bromatoloxía, Avda. Carballo Calero s/n., 27002-Lugo, Spain; E-mail: manuel.vazquez@usc.es