Combination of visible and mid-infrared spectra for the prediction of chemical parameters of wines

Ilknur Sen [a], Burcu Ozturk [b], Figen Tokatli [a,*], Banu Ozen [a]

[a] Izmir Institute of Technology, Department of Food Engineering, 35430 Urla-Izmir, Turkey
[b] Namik Kemal University, Sarkoy Vocational School, Tekirdag, Turkey

ABSTRACT

Rapid and environmentally friendly methods for the prediction of chemical compositions have been an interest in the wine industry. The objective of the study was to show the potentials of combined use of visible and mid-infrared (MIR) spectroscopies to improve the prediction of various chemical compounds of wine as opposed to using mid-infrared range only. Wine samples of twelve grape varieties from two harvest years were analyzed. The chemical composition of wine samples was related to MIR and visible spectra using orthogonal partial least square (OPLS) regression technique. The prediction abilities were tested with crossvalidation and independent validation sets. The coefficient of determination of validation (R^2_{val}) for anthocyanin compounds of red wines were between 0.76 and 0.90, and that for total phenol content was 0.90. Range of R^2_{val} for glycerol, glycerol/ethanol ratio, malic acid, o-coumaric acid and °Brix were between 0.77 and 0.96. The spectral ranges that played significant roles in the predictions were also determined. The validations with independent data sets showed that the combination of visible and MIR ranges with multivariate methods improved the prediction of anthocyanin compounds and total phenols; produced comparable results for the rest of the parameters as MIR. This is the first study in the literature that shows the practical use of visible spectra along MIR. The combined use of these spectral ranges with multivariate models can be applied for the rapid, on-line determination of quality parameters and chemical profiles of wines.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Polyphenol content of grapes and wines are basically grouped as flavonoids (flavonols, flavan-3-ols, anthocyanins) and non-flavonoids (hydroxybenzoic acids, hydroxycinnamic acids, stilbenes). Wine quality highly depends on phenolic compositions as they improve wine taste and color and also have positive health effects [1]. The color of red wine is a major concern of wine industry since it strongly affects the consumer demands. Anthocyanin content is the main reason for the color of red wine and depends on the grape variety, degree of grape ripeness, soil and climatic conditions, and it undergoes a progressive change from production to consumption of any wine due to polymerization, copigmentation and oxidation reactions [2,3]. In terms of white wine, flavonols are the color pigments giving the yellow color of white wines [4]. Chromatographic methods are commonly used to determine the phenolic and anthocyanin profiles of wines. However, due to their

laborious and expensive work, spectroscopic methods have been offered as the rapid and reliable alternatives including near and mid-infrared spectroscopies.

The wine production sector requires rapid, reliable and simultaneous determination of analytes from the beginning of the grape ripening to the end of fermentation for monitoring and screening purposes. Online determination of wine composition enables the control of production process and ensures optimal product quality. The use of vibrational spectroscopy in wine sector started with the near infrared reflectance (NIR) spectroscopy. Recent studies have also focused on Fourier transform infrared spectroscopy (FTIR) technology due to advantages in the analysis of some of the wine constituents [5,6]. This technique has been employed in monitoring of grape ripening, fermentation and quantification of some chemical parameters such as organic acids, sugars or ethanol [7,8]. UV–Visible spectroscopy (UV–Vis) is another cost- and time-saving technique that is used in routine analysis such as determination of total phenols and color of wines. There are reports that showed the combined use of visible and NIR spectroscopies in classification of wines and in prediction of certain phenolic compounds [9]. UV–Vis range along with NIR region

* Corresponding author.
E-mail addresses: ilknursen@iyte.edu.tr (I. Sen), bozturk@nku.edu.tr (B. Ozturk), figentokatli@iyte.edu.tr (F. Tokatli), banuozen@iyte.edu.tr (B. Ozen).

was shown as a good predictor of certain phenolic compounds in red wines [10]. Borras, Ferre, Boque, Mestres, Acena and Busto [11] reviewed the fusion of data from different instruments for food authentication. In recent studies, UV–Vis and NIR spectroscopies were used to predict the quality parameters of grapes and real time total sugar content, alcohol content and pH values of rice wines [12,13].

Implementation of spectroscopic techniques in process monitoring and authentication has always been of interest, as they allow the collection of wide range spectra of large number of samples. Therefore, application of chemometric tools becomes necessary to extract the valuable information in the spectral data. The outcomes of the multivariate analysis can be used in classification of samples and determination of chemical and biological variables. Here, the aim of the study was to highlight the potentials of visible and mid-infrared spectra, and multivariate analysis using orthogonal partial least squares regression (OPLS), for the rapid quantification of some chemical compounds in red and white wines. So far, there are no examples of combined use of visible and MIR spectroscopies to predict quality parameters such as individual anthocyanins, phenolic compounds and glycerol. The spectral regions that played important role in the quantification of chemical compounds of red and white wines were also identified.

2. Materials and methods

2.1. Chemicals

In the chromatographic determination of phenolic profiles, HPLC grade standards and solvents were used: malvidin-3-*O*-glucoside, (+)-catechin hydrate, quercetin, quercetin-3-rutinoside, quercetin-3-glucoside, quercetin-3-galactoside, acetonitrile and methanol (Sigma-Aldrich, Steinheim, Germany), (-)-epicatechin, caffeic acid, ferulic acid, gallic acid, kaempferol, myricetin, o-coumaric acid, p-coumaric acid, vanillic acid (Fluka, Steinheim, Germany), *trans*-resveratrol and procyanidin B_1 (Extrasynthese, Genay, France). Other analytical reagents used in HPLC analysis were $NH_4H_2PO_4$, H_3PO_4 (Merck, Darmstadt, Germany), ethanol and glycerol (Sigma-Aldrich, Steinheim, Germany). Deionized water was obtained with Sartorius Arium 611 VF system (Sartorius AG, Gottingen, Germany).

2.2. Wine samples

The samples cover a wide range of wine producing regions, cultivars, and vintages representing varietal and year to year variations. A total number of 45 red and 27 white wine samples belonging to 2006 and 2007 vintages were collected. Monovarietal wine samples were two-year old commercial wines and were produced from twelve different grape varieties. Among these grape varieties, seven were characteristic to Turkey (Boğazkere, Öküzgözü, Papazkarası, Kalecik Karası as red grapes and Emir, Narince, Sultaniye as white grapes). Other five (Merlot, Cabernet Sauvignon, Syrah, Chardonnay and Muscat) were commercially valuable grapes. The grape varieties were grown in different regions of the eastern, western and central Anatolia. The wine bottles were stored at 4 °C and were immediately analyzed after purchasing.

2.3. Mid-infrared analysis

The MIR spectra were collected in the mid-infrared region of 4000–650 cm^{-1} using a Perkin Elmer Spectrum 100 FTIR instrument (Perkin Elmer Inc., Wellesley, MA, USA) equipped with a horizontal attenuated total reflectance accessory (HATR) with ZnSe crystal and deuterated tri-glycine sulfate detector (DTGS). The transmittance spectra were recorded at 4 cm^{-1} spectral resolution and 0.50 $cm\,s^{-1}$ scan speed. A blank of air was acquired prior to samples. Each spectrum was averaged from 64 scans. The measurements were performed in triplicates.

2.4. Visible analysis

Transmittance scans in the visible range (400–700 nm) were recorded in 2 nm sampling intervals with a quartz cuvette of 1 mm path length for red and 10 mm path length for white wines. Spectrometric scans were acquired by using UV2450 model instrument (Shimadzu Inc., Kyoto, Japan). The measurements were repeated in triplicates. In the text, the results of visible and mid-infrared ranges will be given in terms of wavelength (nm), and wavenumber (cm^{-1}), respectively.

2.5. HPLC analysis

Agilent 1200 series HPLC (Agilent Technologies, CA, USA) system containing a G1322A degasser, G1311 quat pump, G1329 autosampler, G1316B column oven (set at 20 °C), and G1315D diode array detector was used in chromatographic analyses with a C18 column for separation (250 × 4.6 mm, 5 μm, ACE, Aberdeen, Scotland). After filtering through a 0.45 μm pore size syringe filter (Minisart RC, Sartorius AG, Gottingen, Germany), wine samples (20 μL) were injected into the HPLC system and phenolic compounds were quantified at 280, 320, 360 and 520 nm. The oven was set at 20 °C. The solvent flow rate was adjusted as 1 mL/min. A gradient mobile phase consisting of solvent A ($NH_4H_2PO_4$, 50 mM), solvent B (20% mobile phase A and 80% acetonitrile), and solvent C (H_3PO_4, 200 mM) was employed as described in Sen and Tokatli [14].

Phenolic compounds were identified by the retention times of pure standards and were quantified by the use of external standard method. The stock solutions were prepared by dissolving the pure standards in HPLC grade methanol. Peonidin-3-*O*-glucoside, petunidin-3-*O*-glucoside, delphinidin-3-*O*-glucoside and their derivatives, vitisin A and pinotin A were quantified in terms of malvidin-3-*O*-glucoside calibration curve; myricetin-3-*O*-glucoside and quercetin-3-*O*-glucuronide were quantified by using the quercetin-3-*O*-glucoside calibration curve. Identification of these phenolic compounds was performed according to the chromatograms of wine samples studied by Gómez-Alonso, García-Romero and Hermosín-Gutiérrez [15]. Data acquisition and peak processing were performed with Chemstation Rev. B.03.02 (Agilent Technologies, CA, USA). The repeatability of the HPLC method was evaluated by calculating the relative standard deviation of replicate measurements (RSD < 15%). The detection limits (LOD) were calculated as three times the standard deviation of signal of blank sample, prepared ten times according to the OIV method [16]. Recoveries were calculated based on the difference of spiked and un-spiked sample and by taking the ratio of this difference to the assigned value.

For the quantification of glycerol and ethanol, an HPLC method based on Castellari, Versari, Spinabelli, Galassi and Amati [17] was employed with HPX-87H column (300 × 7.8 mm, 9 μm, Bio-Rad Laboratories, Hercules, CA, USA). The mobile phase of 0.045 N H_2SO_4 with 6% acetonitrile (isocratic elution) was applied at a flow rate of 0.5 mL/min. The column oven was set at 45 °C and detection was performed with refractive index detector. All the measurements were performed in duplicates.

2.6. Total polyphenol analysis and soluble solid contents by refractive index (°Brix)

Total polyphenol content was determined according to the Folin Ciocalteu method modified by Arnous, Makris and Kefalas [18] as a micro scale protocol. All measurements were performed in triplicates and results were expressed as gallic acid equivalent (mg GAE L^{-1}). Soluble solid contents were measured by a refractometer (Mettler Toledo GmbH, Schwerzenbach, Switzerland).

2.7. Spectral data preprocessing and multivariate statistical analysis

In this study, phenolic compositions of the wine samples were related to the visible and mid-infrared spectra using orthogonal partial least-squares (OPLS) regression, which is used for the prediction of dependent variables Y with multivariate data set **X**. The technique differs from partial least squares (PLS) in the way of handling the variability in X-matrix. OPLS may lead better interpretation and predictions: PLS divides the variance of X-matrix as systematic and residual, whereas OPLS divides the systematic variance into two parts, the part that is correlated to **Y** (predicted) and the part that is uncorrelated to **Y** (orthogonal) [19]. The number of components in OPLS regression for a single variable y is given as $p_p + p_o$, where p_p and p_o are the number of components expressing information of **X** predictive to **Y**, and information of **X** orthogonal to **Y**, respectively. The multivariate analyses were carried out by using Simca-P 13 (Umetrics Inc., Umea, Sweden).

Prior to multivariate modeling, the mid-infrared and visible spectra were standardized by subtracting the averages and dividing them by the standard deviations. Standardization was followed by the second order derivative data pre-processing. This technique was applied with quadratic polynomial order. The aim of pre-processing was to remove the undesired systematic variations in **X** data, and eventually improve the predictive ability of the model. Such variations in the spectra can be due to baseline drift or wavelength regions of low information. The observations, i.e. the wine samples, were divided into two parts: calibration set to build the model and validation set to test the predictive ability of calibration model. For this purpose, approximately two thirds of data were randomly chosen and used for calibration. The rest of the samples were used in the external validation. The validation sets were chosen within the calibration ranges for all compounds. Regression results of validation models were given as regression coefficient R^2_{val}. So, 32 of the red wine samples and 19 of the white wine samples were employed in the calibration models. The OPLS models were also evaluated by R^2_{cal} as the regression coefficient of calibration model, the coefficient of leave-one-out cross-validation (R^2_{CV}), the root mean square error of calibration (RMSEC) and validation (RMSEP). In the regression analysis, whole spectra were used first and then certain ranges of spectra were selected. The selection of the significant variables in the spectra was performed according to the variable importance in the projection (VIP) as a feature of Simca-P software [20]. The variables with VIP greater than 1.0 were selected and then the models were rebuilt with them.

3. Results and discussion

The polyphenol composition, glycerol, ethanol, and soluble solid contents by refractive index (°Brix) values of red and white wine samples are reported in Table 1. The validation parameters of HPLC methods are given as Supplementary material. The most abundant anthocyanin in red wines was malvidin-3-O-glucoside and its derivatives which were followed by petunidin-3-O-glucoside and delphinidin-3-O-glucoside. In relation to the other

Table 1
Chemical compositions (mgL^{-1}) and soluble solid content values of red and white wines.

Parameters	Red wines			White wines		
	Min	Max	Median	Min	Max	Median
Malvidin-3-O-glucoside	0.54	66.7	21.6	–	–	–
Peonidin-3-O-glucoside	0.04	4.47	1.10	–	–	–
Petunidin-3-O-glucoside	0.08	7.63	1.86	–	–	–
Delphinidin-3-O-glucoside	0.06	5.20	1.37	–	–	–
Delphinidin-3-O-(6-acetyl)-glucoside	0.00	1.68	0.28	–	–	–
Petunidin-3-O-(6-acetyl)-glucoside	0.00	1.91	0.47	–	–	–
Peonidin-3-O-(6-acetyl)-glucoside	0.00	2.58	0.70	–	–	–
Malvidin-3-O-(6-acetyl)-glucoside	0.00	20.7	5.00	–	–	–
Delphinidin-3-O-(6-p-coumaroyl)-glucoside	0.00	1.27	0.25	–	–	–
Malvidin-3-O-(6-p-coumaroyl)-glucoside	0.06	11.4	2.45	–	–	–
Rutin	0.82	6.03	2.72	0.03	0.27	0.10
Quercetin	0.48	15.7	3.93	0.02	8.07	0.69
Myricetin	0.04	6.81	2.53	0.02	1.21	0.02
Kaempferol	0.03	1.29	0.64	0.03	0.03	0.03
Quercetin-3-O-glucoside	0.21	47.0	10.3	0.02	3.21	0.02
Quercetin-3-O-galactoside	0.43	12.5	4.92	0.06	2.79	0.06
Quercetin-3-O-glucuronide	0.65	26.4	9.88	0.00	7.97	0.28
Myricetin-3-O-glucoside	0.01	40.4	12.4	0.00	0.30	0.06
Caffeic acid	4.61	36.3	15.5	0.37	22.4	2.52
p-coumaric acid	1.64	11.1	5.35	0.08	13.6	1.22
Ferulic acid	0.64	3.02	1.19	0.06	1.43	0.69
Trans-resveratrol	0.19	1.52	0.48	0.00	0.78	0.10
Gallic acid	13.7	103.9	39.0	6.58	17.2	10.8
(+)-Catechin	12.4	153.6	38.8	1.61	20.1	4.32
Vanillic acid	4.08	11.0	6.74	0.06	1.59	0.72
(−)-Epicatechin	5.38	81.9	18.6	0.55	4.51	2.05
o-coumaric acid	0.57	3.09	1.34	0.11	2.88	0.44
Procyanidin B$_1$	4.99	99.2	39.1	1.88	9.08	2.99
Glycerol	1777	14,199	8650	2185	8134	5583
Ethanol (%)	11	14	13	10	14	13
Total polyphenol content	1269	3579	1984	178	417	275
°Brix	6.7	9.2	7.8	5.9	9.3	6.6
Malic acid	121	400	256	315	2503	1322

phenolic compounds, gallic acid, procyanidin B$_1$ and (+)-catechin were the most abundant in both red and white wine samples.

The mid-infrared and visible spectra of wine samples are shown in Fig. 1. The FTIR spectrum of wine was dominated by the water and ethanol absorption peaks. The O-H stretching vibrations can be observed in the ranges of 3600–3200 cm^{-1} and 1700–1565 cm^{-1} absorption bands. The peaks at 1636 cm^{-1} and 3040–2800 cm^{-1} were associated with C-O-C and C–H stretching vibration bonds. The vibrations of C–O, C–C, CH$_2$, CH$_3$, C–OH, C–H and C≡N bonds can be observed around 1565–965 cm^{-1}. The peaks between 900 and 670 cm^{-1} were attributed to the C-H out-of-plane bending vibrations [21]. In this study, the elimination of particular ranges was based on the values determined according to the VIP feature of Simca-P software. In case of mid-infrared spectra, the regions between 4000 and 3900 cm^{-1} were not included in the models since no contribution was seen in the models.

For the pre-processing of the data, several filtering algorithms such as wavelet compression spectra, orthogonal signal correction and the first and second order derivative were tried. The best R^2 values were established with the second order derivative technique. Elsewhere, it was also reported that the second order derivative of MIR data revealed sharper characteristic absorption bands than the first order derivative [21].

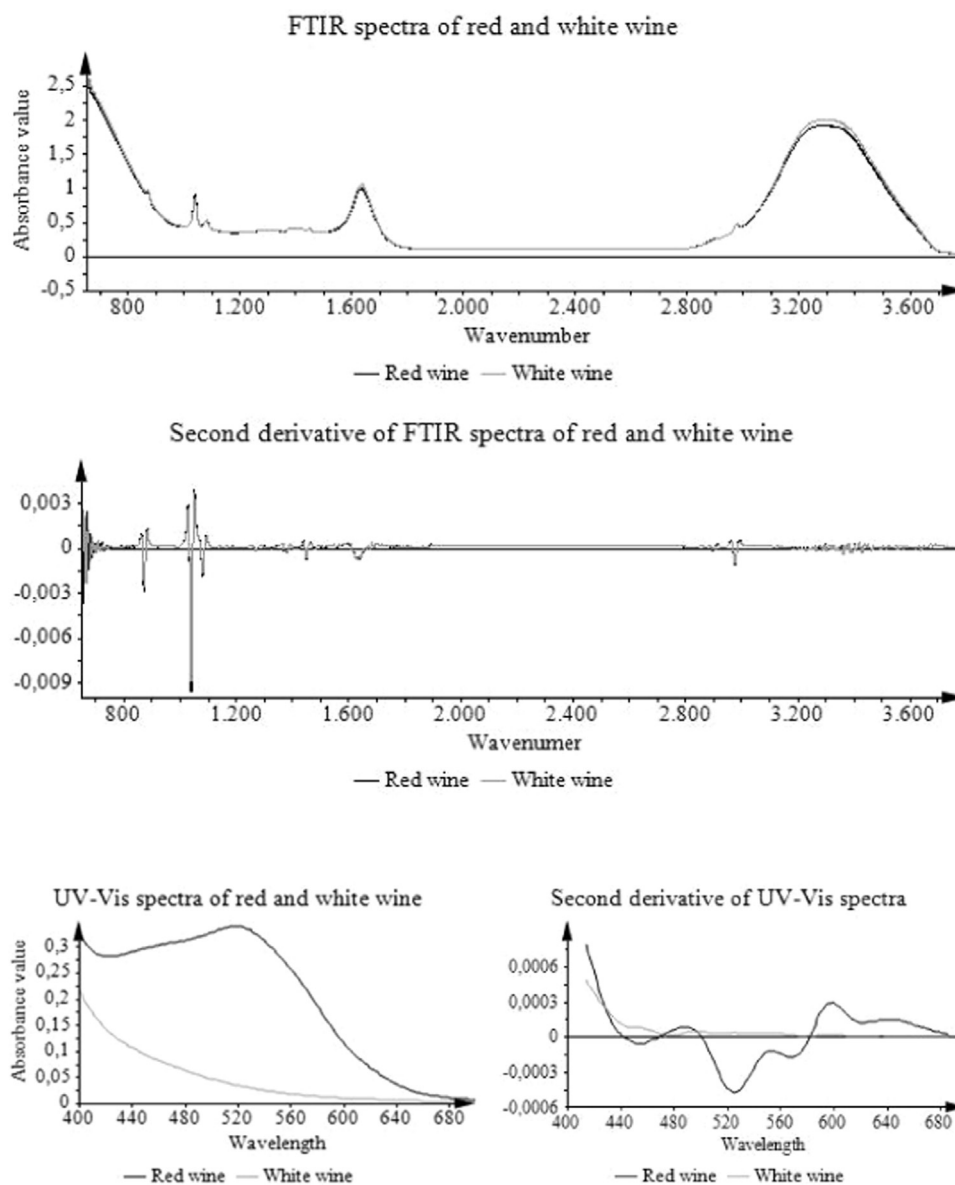The results of the regression models for the chemical

Fig. 1. FTIR and UV–Vis spectra of wine samples and the second derivative of spectra.

compounds with R^2_{val} less than 0.5 were not given. In building the regression models of some chemical compounds such as glycerol and °Brix, the use of all observations (red and white wines) provided satisfactory models. The regression models, on the other hand, failed to predict lower amounts of total phenolics in white wines. Therefore, only red wines were used in regression models for total polyphenol content. The o-coumaric acid content was modeled for white wines only. It should be noted that the multivariate regression models were developed individually for each parameter, not by using the collinear Y matrix.

OPLS models of red wines were established for individual anthocyanin compounds like malvidin-3-O-glucoside, peonidin-3-O-glucoside, petunidin-3-O-glucoside, delphinidin-3-O-glucoside, delphinidin-3-O-(6-acetyl)-glucoside, petunidin-3-O-(6-acetyl)-glucoside, peonidin-3-O-(6-acetyl)-glucoside, malvidin-3-O-(6-acetyl)-glucoside, malvidin-3-O-(6-p-coumaroyl)-glucoside and delphinidin-3-O-(6-p-coumaroyl)-glucoside. The RMSEC and RMSEP values were all lower than 0.73 except for the compounds with higher concentrations such as total phenol content, malvidin-3-O-glucoside and its acetate derivative (Table 3). The R^2_{cal} values

were greater than 0.87. All the proposed models indicated a good predictive ability ($R^2_{CV} > 0.5$) [22]. In general, the slopes and intercept of calibration models are in the vicinity of one and zero, respectively. In the OPLS analysis of anthocyanin compounds of red wines, the visible spectral ranges had higher VIP values than those from mid-infrared region, indicating that visible ranges were more significant than mid-infrared data in model building (Table 2). In the mid-infrared spectra, on the other hand, the most significant variable was 2260 cm^{-1} for the prediction of acetylated malvidin derivatives and malvidin-3-O-glucoside. Meanwhile, 1425 cm^{-1} was the most important variable with the highest VIP value for the prediction of delphinidin-3-O-(6-p-coumaroyl)-glucoside and peonidin-3-O-glucoside. In the models of delphinidin-3-O-glucoside, petunidin-3-O-(6-p-coumaroyl)-glucoside, and malvidin-3-O-(6-p-coumaroyl)-glucoside, 1283 cm^{-1} and 1286 cm^{-1} had the most impact, respectively. In the study of Fernandez and Agosin [21], it was reported that the peak around 1285 cm^{-1} indicated a characteristic of flavonoid-based tannins. The results of regression analysis for the calibration and validation sets are presented in Fig. 2 and Fig. 3.

Table 2
Visible and mid-infrared spectral ranges employed in OPLS models.

| Compound	Visible spectra (nm)	Mid-infrared spectra (cm^{-1})
Refractive index (red+white wines)	414–422, 440–468, 504–686	850–1223, 1268–1539, 1600–1788, 1804–1813, 1853–3024, 3794–3800
Glycerol (red+white wines)	414–426, 440–468, 502–686	662–664, 858–1142, 1181–1546, 1597–3022, 3799
Glycerol/ethanol (Red+white wines)	414–426, 438–470, 488–490, 502–686	859–2270, 2423–2498, 2566–2631, 2704–3021, 3345–3714
Malic acid (red+white wines)	434–472, 500–688	896–3716
Anthocyanin compounds (red wines)	414–434, 470–688	799–1198, 1245–1704, 1881–2131, 2256–2309, 2406–2572, 2700–3052, 3277–3309, 3575–3726
Total phenol content (red wines)	414–426, 440–474, 494–500, 564–584, 600–602, 626–686	750–2052, 2205–2309, 2379–2608, 2663–3451, 3535–3897
o-coumaric acid (white wines)	438–450, 526–538, 612–620, 638–646	872–1162, 1219–1289, 1343–1732, 1765–1839, 1877–2319, 2397–3480, 3516–3578, 3638–3881
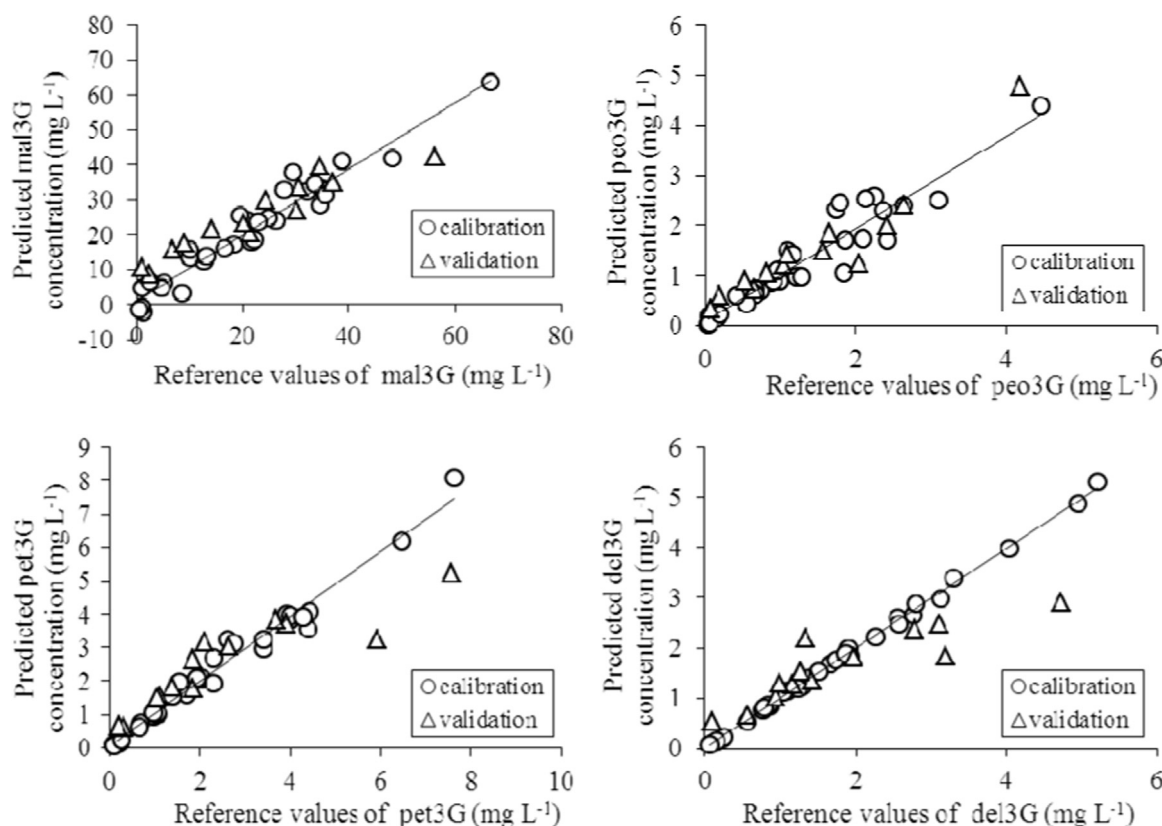
Fig. 2. Regression plots of calibration and validation sets for the OPLS models of red wines. mal3g: malvidin-3-O-glucoside, peo3g: peonidin-3-O-glucoside, pet3g: petunidin-3-O-glucoside, del3g: delphinidin-3-O-glucoside.

For the case of white wines, a powerful model could be established for o-coumaric acid value through visible and mid-infrared spectra. The most significant variable in the o-coumaric acid model was from the mid-infrared spectra (2542 cm^{-1}). OPLS models were also constructed using only mid-infrared data to compare the prediction performance of two spectrometric sets (mid-infrared data and combined visible and mid-infrared data). The results of the models are given in terms of correlation coefficients of calibration and validation data sets and leave-one-out cross-validation coefficients (Table 3). It was observed that the use of visible spectra together with mid-infrared spectra improved the prediction of anthocyanin compounds and total polyphenol content (higher R^2_{CV} values). On the other hand, °Brix, glycerol, glycerol/ethanol ratio and o-coumaric acid models revealed similar prediction performances for the two data sets. Anthocyanin compounds are absorbed in the visible range, whereas some of the phenolic compounds give the highest absorbance in the UV range [23]. This explains higher prediction capability of the method for

anthocyanins; however, not for other compounds like flavan-3-ols (catechin, epicatechin), flavonols (quercetin, myricetin and kaempferol), and most of phenolic acids (gallic acid, ferulic acid). The models with combined spectra and mid-infrared spectra resulted in similar results for glycerol, glycerol/ethanol, malic acid, and °Brix, since their absorbance values in the visible region had very low contribution.

Total phenol content of red wines was predicted with lower number of significant variables in the visible range than the malvidin compounds. Among the significant variables, 1307 cm^{-1} had the most significant impact on the model with the highest VIP value. The spectral range of 2000–750 cm^{-1} was included in the prediction of total phenols and anthocyanins. Similarly, it was reported that absorption band between 1800–900 cm^{-1} was useful in the prediction of total phenolic content, flavonoid content and antioxidant capacity of Moscatel dessert wines [24].

The OPLS technique was shown to predict the °Brix, malic acid, glycerol and glycerol/ethanol ratio and malic acid parameters of all
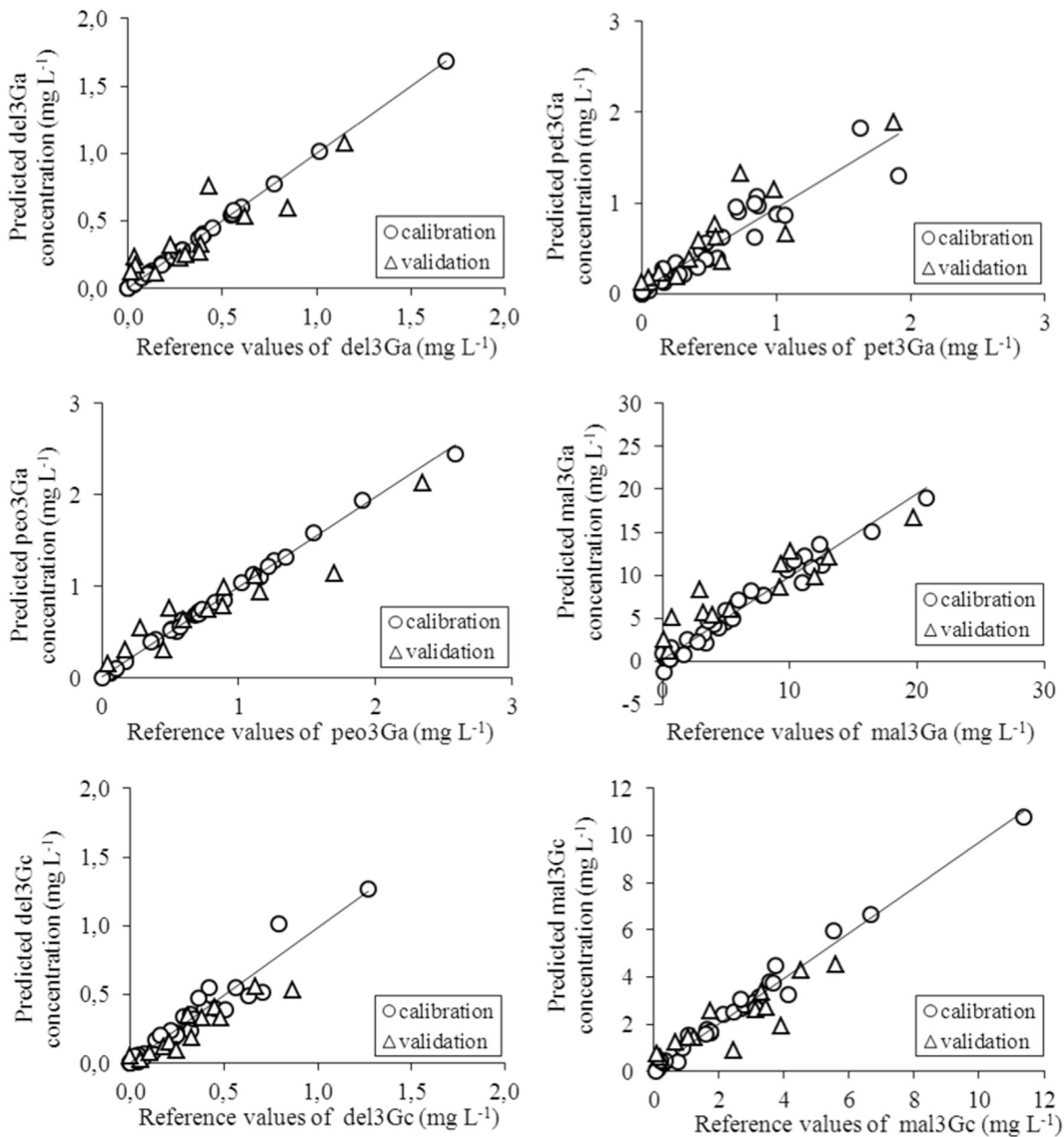
Fig. 3. Regression plots of calibration and validation sets for the OPLS models of red wines. del3ga: delphinidin-3-O-(6-acetyl)-glucoside, pet3ga: petunidin-3-O-(6-acetyl)-glucoside, peo3ga: peonidin-3-O-(6-acetyl)-glucoside, mal3ga: malvidin-3-O-(6-acetyl)-glucoside, del3gc: delphinidin-3-O-(6-p-coumaroyl)-glucoside, mal3gc: malvidin-3-O-(6-p-coumaroyl)-glucoside.

Table 3
Results of calibration and validation models of visible and mid-infrared and only mid-infrared data.

Parameter	Vis+MIR data							MIR data		
	PC[a]	R^2_{cal}	R^2_{CV}	R^2_{val}	Calibration equation	RMSEC	RMSEP	R^2_{cal}	R^2_{CV}	R^2_{val}
°Brix (red+white wines)	1+1	0.96	0.95	0.96	y=1.00x−0	0.17	0.16	0.96	0.95	0.96
Glycerol (red+white wines)	1+3	0.95	0.89	0.83	y=1.00x−0	545.8	740.7	0.94	0.89	0.84
Glycerol/ethanol (red+white wines)	1+6	0.99	0.82	0.89	y=1.00x−0	8.92	49.4	0.99	0.84	0.89
Malic acid (red+white wines)	1+3	0.88	0.76	0.86	y=0.88x+80.00	250.4	240.5	0.90	0.76	0.84
Total polyphenol content (red wines)	1+3	0.98	0.83	0.90	y=0.99x+15.00	72.6	249.1	0.98	0.75	0.87
Malvidin-3-O-glucoside	1+2	0.94	0.73	0.90	y=1.00x−0	3.94	6.87	0.93	0.53	0.71
Peonidin-3-O-glucoside	1+2	0.89	0.68	0.89	y=0.98x+0.04	0.35	0.38	0.99	0.56	0.73
Petunidin-3-O-glucoside	1+3	0.98	0.80	0.80	y=1.00x−0	0.32	1.10	0.99	0.67	0.64
Delphinidin-3-O-glucoside	1+4	0.99	0.78	0.79	y=1.00x−0	0.07	0.73	0.99	0.71	0.38
Delphinidin-3-O-(6-acetyl)-glucoside	1+7	0.99	0.67	0.78	y=1.00x+0	0.00	0.15	0.33	0.27	0.78
Petunidin-3-O-(6-acetyl)-glucoside	1+2	0.87	0.63	0.80	y=0.97x+0.02	0.18	0.24	0.90	0.29	0.59
Peonidin-3-O-(6-acetyl)-glucoside	1+5	0.99	0.66	0.90	y=1.02x−0.01	0.04	0.22	0.47	0.31	0.48
Malvidin-3-O-(6-acetyl)-glucoside	1+2	0.96	0.63	0.86	y=1.00x−0	1.11	2.59	0.48	0.41	0.64
Delphinidin-3-O-(6-p-coumaroyl)-glucoside	1+2	0.93	0.62	0.87	y=0.94x+0.02	0.08	0.12	0.89	0.45	0.81
Malvidin-3-O-(6-p-coumaroyl)-glucoside	1+2	0.98	0.70	0.76	y=1.02x−0.03	0.35	0.65	0.99	0.69	0.67
o-coumaric acid (white wines)	1+2	0.99	0.63	0.77	y=1.07x−0.03	0.10	0.33	0.99	0.63	0.72

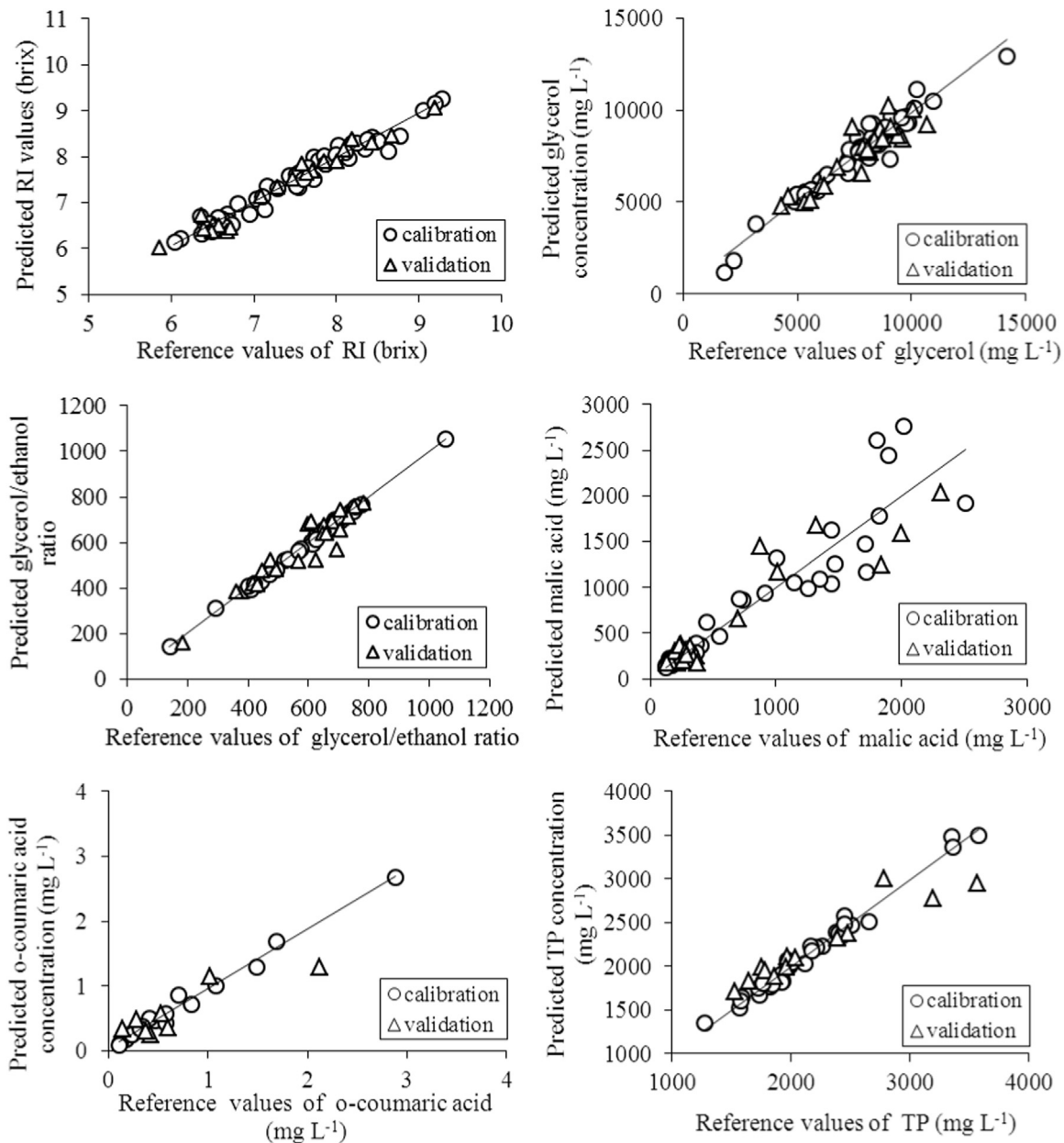[a] PC: Number of principle components in OPLS models (pp+po).

Fig. 4. Regression plots of calibration and validation sets for the OPLS models of white and red wines. Regression plots of RI (soluble solid contents by refractive index, °Brix); glycerol (mg L^{-1}); glycerol/ethanol; malic acid (mg L^{-1}) for red and white wines; o-coumaric acid (mg L^{-1}) for white wines, TP (total polyphenol content) for red wines.

wine samples (red and white) well. In the prediction of °Brix, 2803 cm^{-1} had the highest VIP value, whereas 1195 cm^{-1}, 996 cm^{-1} and 926 cm^{-1} were the most significant variables for malic acid, glycerol and glycerol/ethanol ratio. The correlation between reference amounts determined with HPLC and predicted values for the calibration and validation sets are presented in Fig. 4.

In literature, the use of FTIR for total polyphenol and total flavonol contents were reported with similar leave-one-out cross-validation or validation R^2 values to this study, however, with rather narrower range of wine samples (wines of limited varieties /red wine only) or vintages [24]. Prediction of individual phenolics by UV–Vis and NIR were studied, and only the results of calibration sets were reported [10]. Their results indicated that *trans*-resveratrol and malvidin 3-O-glucoside (oenin) could be predicted successfully. The crossvalidation results for individual anthocyanins are higher in the present study. In another work, malvidin compounds were used to predict anthocyanin content of red wines

with FTIR [25] and their validation results (0.84–0.88) are similar to R^2 values presented in Table 3. The wine samples used here covered a wide range of factors such as different geographic regions, grape varieties and harvest years. Successful OPLS models could be established for anthocyanin compounds and total polyphenol content in red wines, for o-coumaric acid in white wines and glycerol, glycerol/ethanol ratio, malic acid and °Brix in general, with mid-infrared region or combined spectra. It should be noted that the use of whole spectrum between 200 and 700 nm instead of only visible range can provide better predictions of other phenolic substances and chemical parameters in wine.

4. Conclusion

Quantification of phenolic compounds and some quality variables in monovarietal red and white wines using combined visible and mid-infrared spectra was investigated for the first time in

literature and shown to be a promising method. The OPLS technique was used in the regression models. Sampling included wines of different grape varieties, geographic regions and vintages to cover wide ranges of predicted variables. The regression coefficient of determination of both crossvalidation and external validation showed that the analysis of wine samples with these methods could improve the predictions of individual compounds such as total polyphenol, anthocyanin compounds, o-coumaric acid, glycerol, and glycerol/ethanol ratio. The combination of visible and mid-infrared spectra especially provided better predictions of anthocyanin compounds compared to using only mid-infrared data. To increase the accuracy and robustness of these prediction models and to employ them in commercial applications, larger sample sets can be used in future studies. Inclusion of UV spectra in data analysis can increase the predictability of all phenolic compounds as well. Determination of wine phenolic compounds with such practical and reliable techniques makes it possible to monitor the changes of these compounds at various stages of production process and particularly aging.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2016.08.057.

References

[1] R. Rodríguez-Montealegre, R. Romero Peces, J.L. Chacón Vozmediano, J. Martínez Gascueña, E. García Romero, Phenolic compounds in skins and seeds of ten grape Vitis vinifera varieties grown in a warm climate, J. Food Compos. Anal. 19 (6–7) (2006) 687–693.
[2] L. Laghi, A. Versari, G.P. Parpinello, D.Y. Nakaji, R.B. Boulton, FTIR spectroscopy and direct orthogonal signal correction preprocessing applied to selected phenolic compounds in red wines, Food Anal. Method 4 (4) (2011) 619–625.
[3] A. Soriano, P.M. Pérez-Juan, A. Vicario, J.M. González, M.S. Pérez-Coello, Determination of anthocyanins in red wine using a newly developed method based on Fourier transform infrared spectroscopy, Food Chem. 104 (3) (2007) 1295–1303.
[4] N. Castillo-Munoz, S. Gomez-Alonso, E. Garcia-Romero, I. Hermosin-Gutierrez, Flavonol profiles of Vitis vinifera white grape cultivars, J. Food Compos. Anal.
23 (7) (2010) 699–705.
[5] A. de Villiers, P. Alberts, A.G. Tredoux, H.H. Nieuwoudt, Analytical techniques for wine analysis: an African perspective; a review, Anal. Chim. Acta 730 (2012) 2–23.
[6] M. Friedel, C.D. Patz, H. Dietrich, Comparison of different measurement techniques and variable selection methods for FT-MIR in wine analysis, Food Chem. 141 (4) (2013) 4200–4207.
[7] S. Fragoso, L. Acena, J. Guasch, O. Busto, M. Mestres, Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening, J. Agr. Food Chem. 59 (6) (2011) 2175–2183.
[8] C.D. Patz, A. Blieke, R. Ristow, H. Dietrich, Application of FT-MIR spectrometry in wine analysis, Anal. Chim. Acta 513 (1) (2004) 81–89.
[9] D. Cozzolino, H.E. Smyth, M. Gishen, Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins, J. Agr. Food Chem. 51 (2003) 7703–7708.
[10] M.J. Martelo-Vidal, M. Vazquez, Determination of polyphenolic compounds of red wines by UV–VIS–NIR spectroscopy and chemometrics tools, Food Chem. 158 (2014) 28–34.
[11] E. Borras, J. Ferre, R. Boque, M. Mestres, L. Acena, O. Busto, Data fusion methodologies for food and beverage authentication and quality assessment - a review, Anal. Chim. Acta 891 (2015) 1–14.
[12] Q. Ouyang, J. Zhao, W. Pan, Q. Chen, Real-time monitoring of process parameters in rice wine fermentation by a portable spectral analytical system combined with multivariate analysis, Food Chem. 190 (2016) 135–141.
[13] J.U. Porep, A. Mattes, M.S. Pour Nikfardjam, D.R. Kammerer, R. Carle, Implementation of an on-line near infrared/visible (NIR/VIS) spectrometer for rapid quality assessment of grapes upon receival at wineries,, Aust. J. Grape Wine Res. 21 (1) (2015) 69–79.
[14] I. Sen, F. Tokatli, Authenticity of wines made with economically important grape varieties grown in Anatolia by their phenolic profiles, Food Control 46 (2014) 446–454.
[15] S. Gómez-Alonso, E. García-Romero, I. Hermosín-Gutiérrez, HPLC analysis of diverse grape and wine phenolics using direct injection and multidetection by DAD and fluorescence, J. Food Compos. Anal. 20 (7) (2007) 618–626.
[16] OIV, Compendium of International Methods of Wine and Must Analysis Volume 1, International Organisation of Vine and Wine, 2013.
[17] M. Castellari, A. Versari, U. Spinabelli, S. Galassi, A. Amati, An Improved Hplc Method for the Analysis of Organic Acids, Carbohydrates, and Alcohols in Grape Musts and Wines, J. Liq. Chromatogr. Relat. Technol. 23 (13) (2000) 2047–2056.
[18] A. Arnous, D.P. Makris, P. Kefalas, Effect of principal polyphenolic components in relation to antioxidant characteristics of aged red wines, J. Agr. Food Chem. 49 (2001) 5736–5742.
[19] R.C. Pinto, J. Trygg, J. Gottfries, Advantages of orthogonal inspection in chemometrics, J. Chemom. 26 (6) (2012) 231–235.
[20] L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, Multi- and Megavariate Data Analysis:Principals and Applications, Umetrics, Sweden, 2001.
[21] K. Fernandez, E. Agosin, Quantitative analysis of red wine tannins using fourier-transform mid-infrared spectrometry, J. Agr. Food Chem. 55 (2007) 7294–7300.
[22] W. Saeys, A.M. Mouazen, H. Ramon, Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy, Biosyst. Eng. 91 (4) (2005) 393–402.
[23] E.H. Anouar, J. Gierschner, J.-L. Duroux, P. Trouillas, UV/visible spectra of natural polyphenols: a time-dependent density functional theory study, Food Chem. 131 (1) (2012) 79–89.
[24] S.D. Silva, R.P. Feliciano, L.V. Boas, M.R. Bronze, Application of FTIR-ATR to Moscatel dessert wines for prediction of total phenolic and flavonoid contents and antioxidant capacity, Food Chem. 150 (2014) 489–493.
[25] M. Romera-Fernandez, L.A. Berrueta, S. Garmon-Lobato, B. Gallo, F. Vicente, J.M. Moreda, Feasibility study of FT-MIR spectroscopy and PLS-R for the fast determination of anthocyanins in wine, Talanta 88 (2012) 303–310.