doi: 10.17221/370/2014-CJFS

Application of Multivariate Regression Methods to Predict Sensory Quality of Red Wines

José Luis ALEIXANDRE-TUDÓ, Inmaculada ALVAREZ, Maria José GARCÍA, Victoria LIZAMA and José Luis ALEIXANDRE

Food Technology Department, Polytechnic University of Valencia, Valencia, Spain

Abstract

Aleixandre-Tudó J.L., Alvarez I., García M.J., Lizama V., Aleixandre J.L. (2015): **Application of multivariate regression methods to predict sensory quality of red wines**. Czech J. Food Sci., 33: 217–227.

Several multivariate methods including partial least squares (PLS) regression, principal component regression (PCR) or multiple linear regression (MLR) have been applied to predict wine quality, based on the definition of chemical and phenolic parameters of grapes and wines harvested at different ripening levels. Three different models including grape phenolic maturity parameters (grape), wine phenolic parameters (wine) and a combination of grape and wine phenolic parameters (grape + wine) were analysed for each of the wine sensory attributes. The grape parameter model has been presented as the best test to predict wine quality based on sensory scores. On the other hand, wine models showed lower accuracy. The combination of grape and wine parameters presented intermediate results showing sometimes good predictability. Moreover, PLS and PCR appeared as more accurate multivariate methods compared to MLR. Although MLR showed higher correlation coefficients, lower RPD values were observed, displaying thus its lower prediction accuracy. Multivariate calibration statistics appeared as a promising tool to predict wine sensory quality in an easy and inexpensive way.

Keywords: PLS regression; sensory attributes; phenolic parameters

Phenolic compounds are highly important for the overall wine quality. In red wine tannins and anthocyanins are the most important phenolic classes. Tannins contribute to the mouthfeel of wines but they also form more stable pigmented polymers in association with the anthocyanins. It has generally been accepted that the concentration of phenolic compounds in grapes increases through berry development. Tannins and hydroxycinnamic acids increase until veraison (Downey *et al*. 2003; De Freitas *et al*. 2000). Anthocyanins accumulate in the berries at veraison and increase their concentration during fruit ripening due to the sunlight effect. Some authors have reported a decline of these compounds late in berry development (Kennedy *et al*. 2002). A number of factors influence the decision on the optimum harvest date, influencing the future wine quality (Nogales-Bueno *et al*. 2014). At harvest time within a month a huge variation in phenolic concentration and composition takes place. As mentioned previously, the evolution of phenolic compounds during ripening highly affects their concentration in the berries and also influences their presence in future wines. Thus harvesting grapes at different ripening levels will lead to a bunch of different wines with a substantial variability in phenolic and organoleptic properties (Canals *et al*. 2005; Llaudy *et al*. 2008).

Bobal is an indigenous Spanish cultivar grown mainly in the Valencian Region. Bobal cultivar was selected, on the one hand, because it is the most frequently used in the Designation of Origin Utiel-Requena (Valencia, Spain), on the other hand, because despite being used traditionally in the production of rosé and red "doble pasta" wines and being sold mostly in bulk, during the last few years Bobal grapes have been used to produce high-quality aged wines, which shows the high potential of the cultivar (Méndez 2005; Sánchez 2008; Gómez-García Carpintero *et al*. 2011; Gómez Gallego *et al*. 2012).

Excellent oenological parameters were observed in the particular varietal wines, especially when wines were obtained from old vines. The wines were also

doi: 10.17221/370/2014-CJFS

identified as potential long-term aged wines. Moreover, a remarkable content of resveratrol in Bobal berry skins was also reported; results in accordance with those observed by Gómez-Gallego *et al*. (2012). The previous authors studied the phenolic composition and antioxidant activity of Bobal red wines and noticed a high antioxidant activity which might be correlated with the high phenolic concentration observed. The aromatic profile of Bobal red wines was also described by Gómez García-Carpintero *et al*. (2011). The characterisation of young red Bobal wines shows a complex chemical profile with a wealth of aromas. C_6 and benzene compounds were the most predominant free aromas while benzene compounds followed by C_{13} norisoprenoids were the most abundant in the glycosylated fraction. Moreover, olfactory descriptors of blackberries, raspberries, licorice, leather, pepper and sweet were used in order to describe Bobal wines. The authors also reported an increase in complexity and/or intensity when Cencibel grapes were used together with Bobal grapes for winemaking.

Partial least squares regression, principal component regression, and multiple linear regression are multivariate statistical techniques that have been applied in food research to obtain calibration models as an alternative to other statistical methods (Poveda *et al*. 2004). The PLS regression technique has recently been widely used in combination with chemometrics to predict several wine components (Cozzolino *et al*. 2004, 2008; Urbano-Cuadrado *et al*. 2004; Smith 2005; Tarantilis *et al*. 2008; Lorenzo *et al*. 2009). MLR has also been used for the prediction of wine tannin concentrations (Dambergs *et al*. 2012). Moreover, PCR has been mentioned as a possible solution against noise and correlations in the data, while this technique showed its ability to deal with nonlinear relationships between variables (Cozzolino *et al*. 2009).

As far as we know, no available literature exists with regard to the application of multivariate calibrations to assess the wine sensory quality. Quality measurements are usually defined by examining sensory panel scores. A disadvantage of sensory panels is that they require highly trained personnel who can be expensive to train and employ. Sensory panels are also impractical for use on a large scale. Therefore the aim of the present work was to assess the potential of PLS, PCR and MLR regression models to predict the wine sensory quality based on the definition of chemical and phenolic parameters of Bobal grapes and wines harvested at different ripening levels.

MATERIAL AND METHODS

Wine samples. Grapes of *Vitis vinifera* cv. Bobal were harvested in 2009 from 40 years old vines at the Coloraos experimental vineyard located in Requena (Valencia, Spain). Grapes were carefully harvested into 15-kg boxes and transported to an experimental wine production centre. The grapes were destemmed and crushed, and the must was homogenised and distributed to 50-l stainless steel tanks. Sulphur dioxide as potassium metabisulfite was added prior to carry out the different vinification practices. Alcoholic fermentation was induced by inoculation with *Saccharomyces cerevisiae* strain CT007 (Agrovin, Alcázar de San Juan, Spain). To ensure the development of malolactic fermentation selected *Oenococcus oeni* strain OE 104 (Agrovin) lactic acid bacteria were inoculated at the end of alcoholic fermentation.

Grapes were harvested at eight different ripening levels. The first harvest was carried out on 15th September while the last harvest took place on 15th October. Grapes were harvested every 4 days approximately. Wines were elaborated in triplicate, thus 24 different wines corresponding to different ripeness levels were obtained. Triplicate analyses of the wines were performed after malolactic fermentation (Figure 1).

Grape and wine analysis. Samples were analysed according to Glories and Augustin (1993), in order to determine the total potential anthocyanins (ApH1), the potential in extractable anthocyanins (ApH3.2) and the phenolic richness of grapes (A280). Complementary indexes such as anthocyanin extractability index (EA% = [(ApH1 – ApH3.2)/ApH1] × 100), skin tannin levels (dpell = (ApH3.2 × 40)/1000), relative proportion of skin tannins (dpell% = (dpell/A280) × 100), seed tannin levels (dTpep = A280 – dpell) and the relative proportion of seed tannins (Mp% = [(A280 – dpell)/A280] × 100) were also determined.

Colour density (CD) and hue (Blouin 1992), Folin-Ciocalteu index (Singleton & Rossi 1965), anthocyanins (Ribereau-Gayon & Stonestreet 1965) and tannins (Saint-Cricq De Gaulejac *et al*. 1998) were determined by spectrophotometric methods. Polyvinylpolipyrrolidone (PVPP) and astringency were estimated by the method reported by Llaudy

doi: 10.17221/370/2014-CJFS

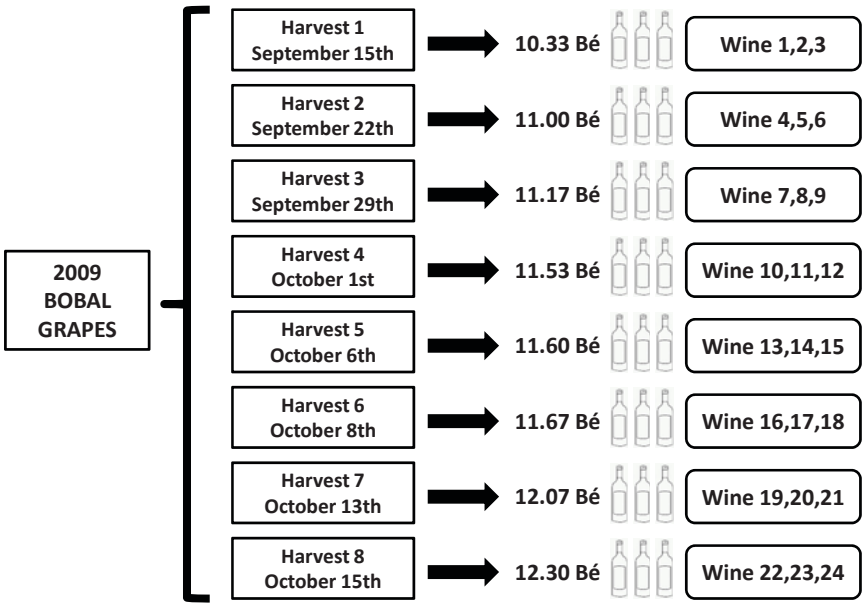

Figure 1. Experimental design of 2009 wines

et al. (2004). Hydrochloridric acid (HCl), ethanol (EtOH), and polymerisation indexes (Glories 1984) were also determined. Each analysis was performed in triplicate.

Sensory analysis. A sensory panel composed of 9 expert trainers with considerable experience in sensory analysis evaluated 24 wines corresponding to 2009 wine samples during 6 sessions using a complete block design. Wine samples were stored at 4°C and brought at room temperature before the tasting sessions. 20 ml of the wine sample were presented for the detection of odour, aroma, and taste. Assessments took place in a standard sensory analysis chamber, equipped with separate booths. Wines were presented in coded standard wine tasting glasses and covered with a watch glass. The sensory profile was determined using six descriptors (colour, aroma intensity and aroma quality, taste intensity and taste quality and overall quality). The panellists used a 10-point scale to rate the intensity of each attribute.

Statistical analysis. Partial least squares (PLS) regression is a biased multilinear regression based on latent variables that aims to obtain a linear model between a set of predictor variables (grape and wine chemical and phenolic parameters or *X* variables) and a set of response variables (sensory attributes or *Y* variables). PLS searches the directions in the predictor space with the maximum variance but avoiding those that are not correlated with the responses to achieve the highest prediction capacity.

Multiple linear regression (MLR) is based on ordinary least squares regression and is used with explanatory or predictive purposes. It correlates information between predictor (*X* variables) and response variables (*Y* variables). The method assumes the *X* variables as linearly independent variables, i.e. no linear relationship exists between *X* variables. Collinearity problems can arise if variables are not linearly independent. The ability to vary independently of each other is an important requirement for predictive variables in this method.

Principal component regression (PCR) is a method which relates the variation in a response variable (sensory attributes or *Y* variables) to the variations of several predictors (grape and wine parameters or *X* variables). PCR is a two-step method. At first, a principal component analysis (PCA) is performed in the *X* variables. The principal components are then used as predictors in a MLR.

The statistics used to describe and compare the model performance include the root mean square error of cross validation $(RMSECV)_{CV}$, the root mean square error of prediction $(RMSEP)_{CV}$, the coefficient of determination (r^2_{val}) between measured and predicted values (obtained by full cross validation), the residual predictive deviation (RPD) and the coefficient of variation (*CV*). The RMSECV is a measure of the average difference between the values determined by the sensory panellists and those predicted by the model in cross validation and it is expressed in the same units as the sensory analysis. The RMSEP is a measure of the average difference between values predicted by the model and values determined by the sensory analysis during independent

doi: 10.17221/370/2014-CJFS

testing of the model. The residual predictive deviation (RPD = standard deviation (SD)/RMSEP) is a useful statistic that is often applied to evaluate how well a calibration model can predict. If the RMSEP is large compared with the range of composition (as SD), relatively small RPD values result and the calibration model is considered not to be robust. The higher the RPD, the greater the probability of the model to predict accurately in samples outside the calibration set. Calculations were performed using the Unscrambler v. 9.2 program (CAMO, AS, Trondheim, Norway).

The analysis of correlation between grape and wine parameters was also performed with the objective of investigating if there is a positive or negative correlation between the grape phenolic parameters studied and the phenolic measurements observed in the wines made thereafter. The sensory analysis data was analysed by ANOVA. The analysis of variance (ANOVA) was applied in order to evaluate whether there exist significant differences between the samples as well as to select the variable that most influences the differences between them. *LSD* test was used to separate the means ($P < 0.01$) when the ANOVA test was significant. For the statistical processing of the data the Statgraphics Plus v. 5.1 softwarewas used.

RESULTS AND DISCUSSION

Correlation analysis of grape and wine phenolic parameters. When evaluating the grape quality we should not only keep in mind the level of phenolic compounds in the grapes since also the extractability of the phenolic material should be considered. Important factors to bear in mind are for example the seasonal and cultivar differences. Positive correlations between grapes and wines have been generally found in the literature for anthocyanins and colour density but, on the other hand, parameters such as total phenolics or tannin concentrations have shown poorer correlations (Du Toit & Visage 2011; van der Merwe 2012). The correlation analysis of the parameters analysed in grapes and wines has been performed and discussed (Table 1).

Interestingly, the potential extractable anthocyanins and total extractable anthocyanins showed very strong positive correlations with HCl, EtOH, PVPP, and polymerisation indexes showing correlation coefficients of 0.98, 0.84, 0.83, and 0.91, respectively. Moreover, lower positive correlations were observed for the anthocyanin concentration (mg/l), Folin index, CD, and hue. On the other hand, negative strong correlations were found between the potential and total extractable anthocyanins with the tannin concentration (g/l) and gelatin index. Surprisingly, the same behaviour as that reported previously has been observed for the parameters dpell and dpell%. This fact could be explained by the anthocyanin location in the skins (Kennedy *et al*. 2006a) and the dpell and dpell% measurements since both of them measure compounds located in the berry skins.

When the parameters evaluating the phenolic material (dTpep and Mp%) present in the seeds (tannins) were correlated, only a positive strong correlation was observed for the gelatin index. On the other hand, strong negative correlations were obtained in all the remaining parameters. Interestingly, a weak positive correlation was observed between the parameter that evaluates the tannin concentration in the seeds and the total tannin concentration (g/l) ((Saint-Cricq

Table 1. Correlation analysis between grape and wine phenolic parameters. Grape parameters determined by the method proposed by Glories and Augustin (1993) and phenolic parameters determined in the obtained wines

	ApH1	ApH3.2	EA%	A280	dpell	dpell%	dTpep	Mp%
Anthocyanins (mg/l)	0.55	0.556	–0.48	–0.62	0.56	0.70	–0.74	–0,70
Tannins (g/l)	–0.66	–0.57	0.15	–0.21	–0.57	–0.44	0.31	0.44
Folin index	0.51	0.51	–0.42	–0.60	0.51	0.64	–0.69	–0.64
Colour density	0.61	0.53	–0.22	–0.42	0.53	0.61	–0.62	–0.62
Hue	0.63	0.60	–0.34	–0.31	0.60	0.63	–0.60	–0.63
HCl index	0.98	0.96	–0.62	–0.11	0.96	0.87	–0.76	–0.87
EtOH index	0.84	0.88	–0.76	–0.44	0.88	0.91	–0.89	–0.91
Gelatin index	–0.91	–0.96	0.81	0.25	–0.96	–0.91	0.84	0.90
PVPP index	0.83	0.86	–0.71	–0.57	0.86	0.94	–0.94	–0.94
Polymerisation index	0.91	0.89	–0.60	–0.45	0.89	0.93	–0.89	–0.93

doi: 10.17221/370/2014-CJFS

Table 2. Statistical values of grape and wine phenolic parameters measured in 2009 wines

	Mean	Min.	Max.	SD	*CV*
Grape phenolic parameters					
ApH1	596.64	501.00	694.77	67.86	11.37
ApH3.2	326.74	244.26	395.44	47.27	14.47
EA%	45.33	42.95	51.22	2.62	5.78
A280	24.85	22.13	26.98	1.38	5.55
dpell	13.07	9.77	15.82	1.89	14.47
dpell%	52.82	36.21	61.97	8.35	15.81
dTpep	11.77	9.30	17.20	2.57	21.85
Mp%	47.16	38.01	63.70	8.34	17.69
Wine phenolic parameters					
Anthocyanins (mg/l)	605.89	480.07	697.14	74.02	12.22
Tannins (g/l)	2.35	2.02	2.70	0.24	10.09
Folin index	47.53	41.51	52.16	3.56	7.49
Colour density	14.64	10.95	20.88	3.34	22.80
Hue	40.81	33.87	43.45	3.43	8.40
HCl index	25.20	20.04	32.01	3.96	15.70
EtOH index	14.19	10.23	15.96	1.94	13.69
Gelatin Index	54.58	49.91	61.76	3.72	6.81
PVPP index	35.83	23.49	42.55	6.66	18.59
Polymerisation index	22.65	18.68	25.37	2.48	10.95

SD – standard deviation; Min. – minimum; Max. – maximum; *CV* – coefficient of variation

De Gaulejac *et al*. 1998). Moreover, the anthocyanin extractability index (EA%) showed a strong positive correlation with the gelatin index, which indicates that this index is measuring more tannin than the anthocyanin fraction. Finally, a relatively strong negative correlation (–0.6) was detected when A280 and Folin index were correlated. This result suggests the important role that winemaking plays in the phenolic composition and concentration and the fact that the Folin index does not measure the total phenolic pool as the A280 value does (all the phenolic compounds show absorbance features at 280 nm). The presence of strong negative and positive correlations between the analysed parameters in grapes and in the corresponding wines points out the suitability of the data set to be subjected to multivariate calibrations.

Grape and wine phenolic data set. The analysed wines showed high coefficients of variation (> 10%) in the majority of the determined parameters (Table 2). The high *CV* was observed for dpell% (15.81%), dTpep (21.85%), Mp% (17.69%), colour density (22.80%), HCl index (15.70%) and PVPP index (18.59%) and the high values of the coefficient of variation mean that there is variability in the data set. Due to the small data set (24 samples) and the fact that wines were made within the same variety, the variability in the determined parameters was considered suitable for developing multivariate calibrations.

Sensory analysis. The analysis of variance (ANOVA) was applied to evaluate whether there exist significant differences between the samples (Table 3). Generally speaking, higher values were obtained for those wines that were elaborated from more mature grapes (wines 5–8, elaborated from grapes harvested between 11.6 and 12.3 Bé). Although the only parameters which showed statistical differences were taste quality and overall quality, the wines cited previously (wines 5–8) had better ratings.

Multivariate data analysis. Statistics for the PLS, PCR and MLR calibration models developed for wine and grape phenolics in 2009 wines are shown in Tables 4–6, respectively. Three different models were built for each sensory attribute. The first one (named as grape) includes only the grape related phenolic parameters such as ApH1, ApH3.2, EA%, A280, dpell, dpell%, dTpep and Mp%. Grape parameters were included in the model as *X* variables while the different rated sensory attributes were used as reference methods

doi: 10.17221/370/2014-CJFS

Table 3 Mean values and standard deviations of the sensory parameters obtained in the sensorial analysis with a numeric scale 0–10

Ripening level	Colour	Aroma		Taste		Overall quality
		intensity	quality	intensity	quality	
1	8.14 ± 0.69	6.86 ± 0.90	5.71 ± 1.38	6,00 ± 1.15	5.57 ± 1.27[a]	5.86 ± 1.35[a]
2	8.14 ± 0.69	6.29 ± 0.76	5.71 ± 1.11	6,29 ± 1.11	5.57 ± 0.98[a]	5.71 ± 1.25[a]
3	8.14 ± 0.69	6.29 ± 0.95	5.29 ± 1.60	6,57 ± 1.13	5.36 ± 1.35[a]	5.59 ± 1.38[a]
4	8.14 ± 0.69	6.57 ± 1.40	5.43 ± 1.27	6,43 ± 1.27	5.43 ± 1.51[a]	5.71 ± 1.38[a]
5	8.14 ± 0.69	6.43 ± 1.27	6.86 ± 1.07	7,14 ± 1.07	7.00 ± 1.29[b]	7.19 ± 0.95[b]
6	8.29 ± 0.76	6.43 ± 1.40	6.57 ± 1.62	7,00 ± 0.82	6.53 ± 1.13[b]	6.69 ± 1.11[b]
7	8.29 ± 0.76	6.29 ± 1.38	6.71 ± 1.70	7,00 ± 1.15	6.71 ± 1.38[b]	6.71 ± 1.38[b]
8	8.29 ± 0.76	6.14 ± 1.35	6.00 ± 1.15	7,14 ± 0.69	6.57 ± 1.40[b]	6.67 ± 0.98[b]

Different letters within the same column are used to compare grape maturity influence

or *Y* variables. The second model (named as wine) was built considering only the phenolic parameters determined in wine [anthocyanins (mg/l), tannins (g/l), Folin index, colour density, hue, HCl, EtOH, gelatin, PVPP, and polymerisation indexes]. Finally, a model built including grape and wine parameters (named as grape + wine) was also tested. In both models the rated sensory attributes were used as target or *Y* variables.

PLS calibration models for each of the sensory evaluated attributes are documented in Table 4. Generally speaking, the higher coefficients of determination in cross-validation (R^2_{val}) were obtained for the grape parameters models. The tested sensory attributes, with the exception of colour, showed higher regression coefficients in the models where only grape phenolic measurements were included. On the other hand, models where wine parameters were considered (wine and grape + wine models) presented lower correlation coefficients. Only for the attribute "Taste intensity" the wine model exhibited the same coefficient of correlation as that obtained with the grape model (0.88), in this specific case the combined model was slightly lower (grape + wine) (R^2_{val} = 0.86). Moreover, in four out of the six attributes the grape + wine models showed higher coefficients than the wine model. Only when analysing overall quality, the wine model showed a higher correlation coefficient than the combined model (grape + wine). The statistical descriptors $RMSECV_{CV}$, $RMSEP_{CV}$, and RPD showed almost the same trend as that observed for correlation coefficients.

As has been stated by other authors, RPD values higher than three can be used for screening purposes while an RPD higher than five is considered for quality control (Poveda *et al*. 2004). On the contrary, other researchers reported that RPD values below 1.5 are considered unusable, from 1.5 to 2 are suitable for rough screening, those between 2 and 2.5 are suitable for quantitative predictions, while RPD values from 2.5 to 3 or above are considered excellent for prediction purposes (Saeys *et al*. 2005; Davey *et al*. 2009). In this study all the grape models, unless when aroma quality was examined, showed RPD values higher than two, which clearly showed the potential of using grape phenolic parameters when predicting the sensory quality of red wines. Interestingly, the sensory attribute "Taste intensity" showed RPD values higher than 2 in all the models considered in the study, with RPD values of 2.44, 2.76, and 2.59 for the grape, wine and grape + wine phenolic models, respectively. The fact that most of the analysed parameters measure individual phenolic concentrations might explain the high prediction accuracy observed for this sensory attribute. In addition, grape + wine models presented RPD values higher than 1.5 for the attributes "Taste quality", "Aroma intensity" together with the mentioned "Taste intensity", being therefore considered suitable for rough prediction. Moreover, wine models always showed RPD values lower than 1.5 (except "Taste intensity"), being therefore considered unusable for prediction purposes. Finally, where the attribute "Colour" was included as *Y* variable very poor statistics were observed, which indicates that with the available sample set it is not possible to predict the wine colour.

Despite the results observed, some authors argue that in a small range of values the accuracy of the method should be evaluated by coefficient of

doi: 10.17221/370/2014-CJFS

Table 4. Partial least squares (PLS) and principal component regression (PCR) calibration statistics for sensory attributes including different phenolic parameters in 2009 wines

		PLS					PCR				
		R^2_{val}	$RMSECV_{CV}$	$RMSEP_{CV}$	RPD	CV	R^2_{val}	$RMSECV_{CV}$	$RMSEP_{CV}$	RPD	*CV*
Overall quality	grape	0.85	0.22	0.26	2.25	4.15	0.84	0.22	0.26	2.25	4.15
	wine	0.69	0.28	0.41	1.43	6.54	0.44	0.41	0.46	1.27	7.34
	grape+wine	0.65	0.30	0.44	1.33	7.02	0.59	0.33	0.42	1.40	6.70
Taste quality	grape	0.86	0.23	0.27	2.37	4.43	0.86	0.23	0.27	2.37	4.43
	wine	0.66	0.3	0.48	1.33	7.88	0.46	0.44	0.49	1.31	8.04
	grape+wine	0.82	0.18	0.41	1.56	6.73	0.48	0.44	0.47	1.36	7.71
Taste intensity	grape	0.88	0.14	0.17	2.44	2.54	0.88	0.14	0.17	2.44	2.54
	wine	0.88	0.13	0.15	2.76	2.24	0.87	0.14	0.15	2.76	2.24
	grape+wine	0.83	0.15	0.16	2.59	2.39	0.91	0.11	0.14	2.96	2.09
Aroma quality	grape	0.76	0.27	0.31	1.86	5.14	0.77	0.26	0.31	1.86	5.14
	wine	0.24	0.46	0.53	1.09	8.78	0.22	0.48	0.53	1.09	8.78
	grape+wine	0.58	0.27	0.51	1.13	8.45	0.49	0.32	0.47	1.23	7.79
Aroma intensity	grape	0.80	0.09	0.10	2.12	1.56	0.79	0.09	0.10	2.12	1.56
	wine	0.44	0.14	0.16	1.33	2.50	0.44	0.14	0.16	1.33	2.50
	grape+wine	0.68	0.10	0.12	1.77	1.87	0.68	0.10	0.12	1.77	1.87
Colour	grape	0.15	0.12	0.14	0.53	1.71	0.15	0.12	0.14	0.53	1.71
	wine	0.16	0.12	0.14	0.53	1.71	0.14	0.12	0.14	0.53	1.71
	grape+wine	0.14	0.12	0.14	0.53	1.71	0.02	0.13	0.14	0.53	1.71

$n = 24$ – number of samples; R^2_{val} – coefficient of determination in cross-validation; $RMSECV_{CV}$ – root mean standard error in cross-validation; $RMSEP_{CV}$ – root mean standard error of prediction, RPD – residual predictive deviation; *CV* – coefficient of variation

variation (Roberts *et al*. 2004), although the size and interpretation of the *CV* depends partly on the source of data used. Values between 5 and 20% were considered adequate for prediction purposes (Cozzolino *et al*. 2008). As can be seen in Table 4, coefficients of variation always lower than 10% were determined, meaning good prediction ability in most of the models. In this case a better interpretation of the results is made if we consider that the lower the *CV*, the better the prediction ability of the models.

The principal component regression analysis was also performed with the aim of evaluating another different multivariate statistical method. Table 5 shows the calibration statistics for the sensory attributes considered in the study. As can be seen in Table 4, more or less the same trends as those reported for the PLS analysis were identified. It seems that grape models had better prediction accuracy compared with wine or grape + wine models. RPD values higher than 2.5 were found again for the attributes "Overall quality, Taste quality, Taste intensity, and Aroma Intensity". Again, the taste intensity scores were accurately predicted even when wine or grape + wine parameters were included as *X* variables (RPD values of 2.44, 2.76, and 2.96 for grape, wine and grape + wine models, respectively, were observed). Moreover, as can be expected, models including grape and wine parameters showed better statistics than those built only including wine phenolic parameters. Finally, PCR was not able to accurately predict the wine colour again.

Finally, with the above-mentioned aim of investigating different multivariate methods the third analysis was also performed using multiple linear regression analysis. Table 6 shows the calibration statistics of the model using sensory attributes as reference values. Interestingly, higher correlation coefficients were observed. R^2_{val} were on average higher than those found for PLS and PCR analysis. Again, as has been observed, grape models showed high correlation coefficients, but in this case models built with a combination of grape and wine phenolic parameters showed higher R^2_{val}. Surprisingly, the increased R^2_{val} were not supported by the RPD values, while the

doi: 10.17221/370/2014-CJFS

Table 5. Principal component regression (PCR) calibration statistics for sensory attributes including different phenolic parameters in 2009 wines

		R^2_{val}	$RMSECV_{CV}$	$RMSEP_{CV}$	RPD	*CV*
Overall quality	grape	0.84	0.22	0.26	2.25	4.15
	wine	0.44	0.41	0.46	1.27	7.34
	grape+wine	0.59	0.33	0.42	1.40	6.70
Taste quality	grape	0.86	0.23	0.27	2.37	4.43
	wine	0.46	0.44	0.49	1.31	8.04
	grape+wine	0.48	0.44	0.47	1.36	7.71
Taste intensity	grape	0.88	0.14	0.17	2.44	2.54
	wine	0.87	0.14	0.15	2.76	2.24
	grape+wine	0.91	0.11	0.14	2.96	2.09
Aroma quality	grape	0.77	0.26	0.31	1.86	5.14
	wine	0.22	0.48	0.53	1.09	8.78
	grape+wine	0.49	0.32	0.47	1.23	7.79
Aroma intensity	grape	0.79	0.09	0.10	2.12	1.56
	wine	0.44	0.14	0.16	1.33	2.50
	grape+wine	0.68	0.10	0.12	1.77	1.87
Colour	grape	0.15	0.12	0.14	0.53	1.71
	wine	0.14	0.12	0.14	0.53	1.71
	grape+wine	0.02	0.13	0.14	0.53	1.71

n = 24 – number of samples; R^2_{val} – coefficient of determination in cross-validation; $RMSECV_{CV}$ – root mean standard error in cross-validation; $RMSEP_{CV}$ – root mean standard error of prediction, RPD – residual predictive deviation; *CV* – coefficient of variation

models showed lower accuracy than PLS and PCR analysis. Very accurate models were observed for the sensory attribute "Taste intensity", which supports the reasoning reported earlier (PLS analysis section). In addition, grape models showed RPD values lower than 2 when the sensory attributes "Overall quality (1.89) and Aroma intensity (1.93)" were tested. On the contrary, other sensory attributes such as "Taste quality, Taste intensity, and Aroma quality" presented the RPD values of 2, 2.07, and 5.25, respectively. Moreover, a large number of models built with the wine and grape + wine sets presented values lower than 1.5 and therefore they are not able to accurately predict samples. Based on RPD results, the multiple linear regression has shown a lower ability to build models able to accurately predict sensory attributes. Finally, no large differences were observed regarding *CV* when comparing different multivariate methods.

Additionally, throughout the study the wine models have shown a poorer ability to predict sensory ratings, being thus considered as unusable models in most of the cases. Looking for a possible explanation of the result observed we suggest here that the obtained low prediction ability could be explained by the well-known low accuracy/reliability of the analytical methods due to different issues (Herderich & Smith 2005).

For example anthocyanins determined by bisulphite bleaching show higher values than those analysed by HPLC. The fact that some polymeric pigments also react with sulphur dioxide has been presented as the reason why higher values are obtained (Versari *et al*. 2008). Hence, both anthocyanins and polymerisation index could be affected by this phenomenon. Colorimetric methods for the determination of tannins have been presented as a lack of selectivity methods and they are only appropriate for the analysis of purified condensed tannin samples or monomeric polyphenols (Peng *et al*. 2002), as opposed to wine or grape extracts. Polyvinylpolypirrolidone (PVPP) binds strongly with tannins and this has been exploited in numerous assays. A significant problem of such strongly binding polymers is their lack of selectivity for tannins and removal of other non-tannic phenols. Total tannin concentration (g/l), HCl index, and PVPP index are the parameters that

doi: 10.17221/370/2014-CJFS

Table 6. Multiple linear regression (MLR) calibration statistics for sensory attributes including different phenolic parameters in 2009 wines

		R^2_{val}	$RMSECV_{CV}$	$RMSEP_{CV}$	RPD	*CV*
Overall quality	grape	0.82	0.22	0.31	1.89	4.95
	wine	0.61	0.28	0.48	1.22	7.66
	grape+wine	0.87	0.10	0.37	1.58	5.90
Taste quality	grape	0.85	0.23	0.32	2.00	5.25
	wine	0.62	0.29	0.52	1.23	8.54
	grape+wine	0.86	0.12	0.41	1.56	6.73
Taste intensity	grape	0.87	0.14	0.20	2.07	2.99
	wine	0.90	0.08	0.16	2.59	2.39
	grape+wine	0.97	0.05	0.16	2.59	2.39
Aroma quality	grape	0.83	0.08	0.11	5.25	1.82
	wine	0.35	0.35	0.61	0.95	10.11
	grape+wine	0.74	0.16	0.49	1.18	8.12
Aroma intensity	grape	0.83	0.08	0.11	1.93	1.72
	wine	0.25	0.13	0.22	0.96	3.43
	grape+wine	0.93	0.05	0.14	1.52	2.18
Colour	grape	0.15	0.11	0.17	0.44	2.07
	wine	–	0.10	0.19	0.39	2.32
	grape+wine	–	0.08	0.29	0.26	3.54

n = 24 – number of samples, R^2_{val} – coefficient of determination in cross-validation, $RMSECV_{CV}$ – root mean standard error in cross-validation, $RMSEP_{CV}$ – root mean standard error of prediction, RPD: residual predictive deviation, *CV* – coefficient of variation

could be affected by the above-mentioned lack of selectivity. An interesting aspect of the protein precipitation assays is that the interaction of proteins with tannins can be used to model astringency perception in humans (Kennedy *et al*. 2006b; Mercurio & Smith 2008). In the gelatin index tannins are racked by their propensity for precipitation by gelatin, but the variability in the composition and purity of gelatin might have caused problems with the reproducibility between studies (Llaudy *et al*. 2004). In general, all the protein precipitation essays are dependent on many variables including pH, isoelectric point, ionic strength, protein conformation, and temperature. Tannin structural features such as the proportion of catechin versus epicatechin, degree of galloylation, ratio of procyanidins to prodelphidins, polymer length and the presence of secondary or tertiary structures could also impact on binding kinetics with gelatin (Seddon & Downey 2008). As we have exposed before the determined parameters are dependent on many variables and their accuracy is not always good enough for analytical purposes and their intrinsic characteristics could influence the results obtained.

CONCLUSIONS

Considering the entire calibration statistics obtained by the three different multivariate methods, the determination of the phenolic maturity following the method developed by Glories and Augustine (1993) has been presented as the best test to predict wine sensory quality. Setting aside the effect that winemaking practices could have on the final wine composition, the determination of grape phenolic characteristics might accurately predict future wine sensory quality attributes. Models built with wine phenolic parameters have shown much lower prediction ability. Moreover, when grape and wine parameters are combined, intermediate results were obtained. The inclusion of some wine phenolic parameters might increase the model performance but further work is required to select those parameters that positively contribute to model accuracy. On the other hand, care should be taken into account when calibration statistics are analysed. As we reported previously, RPD and *CV* calibration statistics depend partly on the source of the data set.

doi: 10.17221/370/2014-CJFS

Even though good calibration statistics were obtained in the evaluated models, due to the small sample set and to the low variability present in the data set, since only one variety coming from the same location was used to develop the calibrations, the results observed in this study must be considered with caution. The authors consider the study as a preliminary screening of the potential of this technique to predict wine sensory quality attributes based on chemistry determinations.

References

Álvarez I., Aleixandre J.L., García M. J., Lizama V. (2005): Impact of prefermentative maceration on the phenolic and volatile compounds in Monastrell red wines. Analytical Chimical Acta, 563: 109–115.

Blouin J. (1992): Manuel pratique d'analyse des mouts et des vins. Blanquefort, Chambre d'Agriculture de la Girone.

Canals R., Llaudy M.C., Valls J., Canals J.M., Zamora F. (2005): Influence of ethanol concentration on the extraction of colour and phenolic compounds from the skin and seeds of Tempranillo grapes at different stages of ripening. Journal Agricultural and Food Chemistry, 53: 4010–4025.

Cozzolino D., Kwiatkowski M.J., Parker M., Cynkar W.U., Dambergs R.G., Gishen M., Herderich M.J. (2004): Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. Analytica Chimica Acta, 513: 73–80.

Cozzolino D., Kwiatkowski M.J., Dambergs R.G., Cynkar W.U., Janik L.J., Skouroumounis G., Gishen M. (2008): Analysis of elements in wine using near infrared spectroscopy and partial least squares regression. Talanta, 74: 711–716.

Cozzolino D., Cynkar W.U., Shah N., Dambergs R.G., Smith P.A. (2009): A brief introduction to multivariate methods in grape and wine analysis. International Journal of Wine Research, 1: 123–130.

Dambergs R.G., Mercurio M.D., Kassara S., Cozzolino, D., Smith P.A. (2012): Rapid measurement of methyl cellulose precipitable tannins using ultraviolet spectroscopy with chemometrics: application to red wine and interlaboratory calibration transfer. Applied Spectroscopy, 66: 565–664.

Davey M.V., Saeys W., Hof E., Ramon E., Swennen R.L., Keulemans J. (2009): Application of visible and near infrared reflectance spectroscopy (Vis/NIRS) to determine carotenoid contents in banana (*Musa* spp.) fruit pulp. Journal of Agricultural and Food Chemistry, 57: 1742–1751.

De Freitas V.A.P., Glories I., Monique A. (2000): Developmental changes of procyanidins in grapes of red *Vitis vinifera* varieties and their composition in respective wines. Australian Journal of Grape and Wine Research, 9: 15–27.

Downey M.O., Harvey J.S., Robinson S.P. (2003): Analysisi of tannin in seeds and skins or Shiraz grapes throughout berry development. Australian Journal of Grape and Wine Research, 9: 15–27.

Du Toit W.J., Visage M. (2011): Correlations between red grape and wine color and phenolic composition: Comparing the Glories, Illand and bovine serum albumin precipitation method. South African Journal of Enology and Viticulture, 33: 1–13.

Glories Y. (1984): La couleur des vins rouges. Mesure, origine et interpretation. Connaissance Vigne Vin, 18: 253–271.

Glories Y., Augustin M. (1993): Maturité phénolique du raisin, conséquences technologiques: application aux millésimes 1991 et 1992. Compte Rendu Colloque Journée Technique. Bordeaux, CIVB: 56–61.

Gómez García-Carpintero E., Sánchez-Palomo E., González-Viñas M.A. (2011): Aroma characterisation of red wines from cv. Bobal grape variety grown in La Mancha region. Food Research International, 44: 61–71.

Gómez Gallego M.A., Gómez García-Carpintero E., Sánchez-Palomo E., González Viñas M.A., Hermosín-Gutierrez I. (2012): Effect of co-winemaking in phenolic composition, color and antioxidant capacity of young red wines from La Mancha region. European Food Research and Technology, 235: 155–167.

Herderich M.J., Smith P.A. (2005): Analysis of grape and wine tannins: Methods, applications and challenges. Australian Journal of Grape and Wine Research, 11: 205–214.

Hidalgo J. (2005): Tecnologia de elaboración y soluciones enológicas al aumento de pH en los vinos debido al potasio. II. Madrid, Encuentro Enológico de la Fundación para la Cultura del Vino.

Kennedy J.A. (2008): Grape and wine phenolics: observation and recent findings. Ciencia e Investigación Agraria, 35: 107–120.

Kennedy J.A., Matthews M.A., Waterhouse A.L. (2002): Effect of maturity and vine water status on grape skin and wine flavonoids. American Journal of Enology and Viticulture, 53: 268–274.

Kennedy J., Saucier C., Glories Y. (2006a): Grape and wine phenolics: History and perspective. American Journal of Enology and Viticulture, 57: 239–248.

Kennedy J.A., Ferrier J., Harbertson J.F., Peyrot Des Gachons C. (2006b): Analysis of tannins in red wine using multiple methods: correlation with astringency. American Journal of Enology and Viticulture, 57: 481–485.

Lorenzo C., Garde-Cerdan T., Pedroza M.A., Alonso G.L., Salinas M.R. (2009): Determination of fermentative vola-

doi: 10.17221/370/2014-CJFS

tile compounds in aged red wines by near infrared spectroscopy. Food Research International, 42: 1281–1286.

Llaudy M.C., Canals R., Canals J., Rosés N., Zamora F. (2004): New method for evaluating astringency in red wines. Journal of Agricultural and Food Chemistry, 52: 742–746.

Llaudy M.C., Canals R., Canals J.M., Zamora F. (2008): Influence of ripening stage and maceration length on the contribution of grape skins, seeds and stems to phenolic composition and astringency in wine-simulated macerations. European Food Research and Technology, 226: 337–344.

Méndez J.V. (2005): Estudio de la maduración fenólica y antociánica en uvas tintas de Bobal para diferentes condiciones agrológicas. [Doctoral Thesis.] Valencia, Universidad Politécnica de Valencia.

Mercurio M.D., Smith P.A. (2008): Tannin quantification in red grapes and wine: Comparison of polysaccharide- and protein-based tannin precipitation techniques and their ability to model wine astringency. Journal Agricultural and Food Chemistry, 56: 5528–5537.

Nogales-Bueno J., Hernández-Hierro J.M., Rodriguez-Pulido F.J., Heredia F.J. (2014): Determination of technological maturity of grapes and total phenolic compounds of grape skins in red and white cultivars during ripening by near infrared hyperspectral image: a preliminary approach. Food Chemistry, 152: 586–591.

Peng Z., Iland P.G., Oberholster A., Sefton M.A., Waters E.J. (2002): Analysis of pigmented polymers in red wine by reverse phase HPLC. Australian Journal of Grape and Wine Research, 8: 70–75.

Poveda J.M., Garcia A., Martin-Álvarez P.J., Cabezas L. (2004): Application of partial least square (PLS) regression to predict the ripening time of Manchego cheese. Food Chemistry, 84: 29–33.

Ribéreau-Gayon P., Stronestreet E. (1965): Le dossage des anthocyanes dans le vin rouge. Bulletine de la Société Chimique, 9: 119–142.

Roberts C.A., Stuth J., Flin P.C. (2004): Near Infrared Spectroscopy in Agriculture. Madison, ASA, CSSA, SSSA,

Saeys W., Mouazen A.M., Ramon H. (2005): Potential onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. Biosystem Engineering, 91: 393–402.

Saint-Cricq De Gaulejac N., Vivas N., Glories Y. (1998): Maturité phénolique: définition et contrôle. Revue Française d'Œnologie, 173: 22–25.

Sánchez N. (2008): Influencia de diferentes factores agrológicos y tecnológicos sobre la mejora de la calidad de los vinos tintos de Bobal. [Doctoral Thesis.] Valencia, Universidad Politécnica de Valencia.

Seddon T.J., Downey M.O. (2008): Comparison af analytical methods for the determination of condensed tannins in grape skin. Australian Journal of Grape and Wine Research, 14: 54–61.

Singleton V.L., Rossi J.A. (1965): Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. American Journal of Enology and Viticulture, 16: 144–158.

Smith P.A. (2005): Development of an assay to quantify tannins in grape and wine samples. Australian Wine Research Institute Technical Review, 154: 37–38.

Tarantilis P.A., Troianou V.E., Pappas C.S., Kotseridis Y.S., Polissiou L.G. (2008): Differentiation of Greek red wines on the basis of grape variety using attenuated total reflectance Fourier transform infrared spectroscopy. Food Chemistry, 111: 192–196.

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. Analyticia Chimica Acta, 527: 81–88.

Van der Merwe H. (2012): The color and phenolic content of Robertson red grape cultivars: Distribution, correlation with wines and analyses. [PhD Thesis.] Stellenbosch, Stellenbosch University.

Versari A., Boulton R.B., Parpinello G.P. (2008): A comparision of analytical methods for measuring the color components of red wines. Food Chemistry, 106: 397–402.

Received: 2014–07–10
Accepted after corrections: 2015–01–15

Corresponding author:

Prof Jose Luis Aleixandre, Universidad Politécnica de Valencia, Camino de Vera s/n, 46071 Valencia, Spain;
E-mail: jaleixan@tal.upv.es