Classification of Argentinean Sauvignon Blanc Wines by UV Spectroscopy and Chemometric Methods

Silvana Mariela Azcarate, Miguel Ángel Cantarelli, Roberto Gerardo Pellerano, Eduardo Jorge Marchevsky, and José Manuel Camiña

Abstract: Argentina is an important worldwide wine producer. In this country, there are several recognizable provinces that produce Sauvignon blanc wines: Neuquén, Río Negro, Mendoza, and San Juan. The analysis of the provenance of these white wines is complex and requires the use of expensive and time-consuming techniques. For this reason, this work discusses the determination of the provenance of Argentinean Sauvignon blanc wines by the use of UV spectroscopy and chemometric methods, such as principal component analysis (PCA), cluster analysis (CA), linear discriminant analysis (LDA), and partial least square discriminant analysis (PLS-DA). The proposed method requires low-cost equipment and short-time analysis in comparison with other techniques. The results are in very good agreement with results based on the geographical origin of Sauvignon blanc wines.

Keywords: chemometric classification, Sauvignon blanc, white wine

Practical Application: This manuscript describes a method to determine the geographical origin of Sauvignon wines from Argentina. The main advantage of this method is the use of nonexpensive techniques, such as UV-Vis spectroscopy.

Introduction

Wine is a product widely consumed of economic importance in many countries around the world (Bentlin and others 2012). Argentina is the 9th wine exporting country and the 5th wine producer after Italy, France, Spain, and the United States, according to the annual statistical report on world vitiviniculture made by the Intl. Organization of Vine and Wine (OIV) for 2011 (Intl. Organization of Vine and Wine 2012). Argentinean wines represent more than 5% of global wine production, and are characterized for their product quality (Arévalo and Córdoba 2008; Intl. Organization of Vine and Wine 2012). Wine has been extensively investigated because of frauds, including adulteration, false declaration of age, and geographical origin (Almeida and Vasconcelos 2003; Taylor and others 2003). In Argentina, the diversity of viticultural areas generates the need for establishing the provenance of wines and their properties (Fabani and others 2010; Granato and others 2011).

The wine growing areas in Argentina are divided into 3 well-differentiated regions based on their ecological characteristics and soil diversity:

- The northwest region, which includes the provinces of Salta, Catamarca, La Rioja, and the northwest of Tucumán.
- The central west region, which includes the provinces of Mendoza and San Juan and represents 91.19% of the vineyard surface of the country.
- The south region, which includes the provinces of La Pampa, Neuquén, Río Negro, and Chubut (Arévalo and Córdoba 2008).

According to the data provided by the *Inst. Nacional de Vitivinicultura* (INV) (Argentinean Vitiviniculture Inst.) the Sauvignon blanc variety is the 3rd white wine variety of greatest exportation after Chardonnay and Torrontes Riojano. Sauvignon blanc wines exported in 2009 correspond to 2.29% of the total red and white varietal wines exported and have continued a growth trend in 2010 and 2011 (Inst. Nacional de Vitivinicultura 2011). In Argentina, Sauvignon blanc varieties are mainly cultivated in Rio Negro (south region) and Mendoza and San Juan (central west region); however, in the northwest region production is fairly scarce (Inst. Nacional de Vitivinicultura 2012).

In Argentina, the Controlled Denomination of Origin (CDO) has been established to define the geographical origin, avoid frauds, and improve the quality of wines (Inst. Nacional de Vitivinicultura 2004). CDO can be studied using modern and sophisticated instrumentation, which is selective and reliable but requires well-trained analysts and involves high costs (De Villiers and others 2003; Wang and others 2003). These studies often involve lengthy analyses since they require sample pretreatment steps. However, now sophisticated techniques and direct measurements, combined with multivariate analysis, have demonstrated the ability to characterize wines by means of simple procedures (Arvanitoyannis and others 1999).

MS 20121596 Submitted 11/20/2012, Accepted 1/2/2013. *Authors Azcarate, Cantarelli, and Camiña are with Facultad de Ciencias Exactas y Naturales, Univ. Nacional de La Pampa, Av. Uruguay 151 (6300) Santa Rosa, La Pampa, Argentina also authors Azcarate, Cantarelli, and Camiña are with Inst. de las Ciencias de la Tierra y Ambientales de La Pampa (INCITAP), Av. Uruguay 151 (6300) Santa Rosa, La Pampa, Argentina. Author Pellerano is with Facultad de Ciencias Exactas, Naturales y Agrimensura, Univ. Nacional del Nordeste, Av. del Libertador 5460 (3400) Corrientes, Argentina. Author Marchevsky is with Inst. de Química de San Luis (INQUISAL), Chacabuco y Pedernera (5700), San Luis, Argentina. Direct inquiries to author Camiña (E-mail: jcaminia@gmail.com, jcaminia@exactas.unlpam.edu.ar).*

doi: 10.1111/1750-3841.12060

Journal of Food Science
A Publication of the Institute of Food Technologists

Table 1–Identification of Sauvignon Blanc wine samples.

Trademark	Region	Year	% vol/vol ethanol	pH
Latitud 33 (L)	Mendoza	2011	13.6	3.3
Hereford (H)	Mendoza	2010	13.0	3.0
Humberto Canale (HC)	Río Negro	2010	13.7	3.2
Alamos (A)	Mendoza	2010	13.2	3.3
Portillo (P)	Mendoza	2011	13.5	3.2
Finca Natalina (FN)	San Juan	2011	13.0	3.2
Marcus (M)	Río Negro	2011	13.5	3.1

In the last years, the use of chemometric methods to classify products according to their geographical origin, variety or other properties has been preferred by researchers (Martens 1983; Forina and others 1987). In particular, numerous applications of multivariate methods for wine classification have been reported based on sensorial properties and chemical variables (Green and others 2011) along with the use of different statistical methods (Granato and others 2012). Principal component analysis (PCA) has been widely used in the analysis of wines. PCA offers information for other methods, for example, cluster analysis (CA) and linear discriminant analysis (LDA). LDA and soft independent modeling of class analogy (SIMCA) allow the classification of wines into defined categories. In the same way, artificial neural networks (ANNs) and partial least squares (PLS) regression have been used for classification purposes (Saurina 2010). Some of these statistical procedures, such as PCA and SIMCA, point out the possibility to discriminate 1 particular denomination of origin from others, using UV-Vis spectroscopy (Urbano and others 2005). Acevedo and others (2007) have carried out the classification of red and white Spanish wines using UV-Vis spectrophotometric analysis and support vector machines to improve the classification. On the other hand, in more recent research, the study of fingerprints and terroir of Argentinean wines was carried out by the use of inorganic, organic, and stable isotope analyses in combination with chemometric methods (Di Paola-Naranjo and others 2011).

In this paper, an inexpensive method for the classification of Argentinean Sauvignon blanc wines of different origins is proposed by using modern chemometric techniques and UV-Vis spectroscopy.

Materials and Methods

Wine samples

Commercial Sauvignon blanc wine samples from different wine-producing provinces of Argentina (Mendoza, San Juan, and Rio Negro) were obtained from local supermarkets and wine stores. A total of 7 Sauvignon blanc wine trademarked samples (5 individual bottles of each one, declared as monovarietal wines on their labels) were analyzed: 4 samples from Mendoza, 2 from Río Negro, and 1 from San Juan; in all cases wines were declared from 89% to 100% of Sauvignon blanc. It should be noted that in Argentina, wines which are considered as monovarietal contain more than 85% of the corresponding declared variety, compliance with Argentinean laws (Inst. Nacional de Vitivinicultura 2004). All wines belonged to 2 vintages (2010/2011). All samples were immediately analyzed after being opened and within 1 mo after purchase. The alcoholic content ranged from 12% to 13.7% vol/vol ethanol (Table 1).

Sample preparation

All reagents were of analytical grade. Ultrapure water (Millipore UV Synergy System, Billerica, Mass., U.S.A.), with a resistivity of 18.2 MΩ cm, was used to prepare all solutions. For the method development, each white wine sample was subjected to the following treatment: 2 mL wine and 10 mL of Clark and Lubs's buffer, pH 10.2 (Meites 1982), were diluted to 25 mL. Clark and Lubs's buffer was prepared with a 0.1 mol/L KCl and H_3BO_3 (Merck, Darmstadt, Germany) solution and a 0.1 mol/L NaOH solution. All dilutions were prepared and tested immediately after the wine bottles were opened, in order to prevent oxidation reactions.

Instrumentation

UV spectroscopy measurements were performed using an Ocean Optics Model CHEMUSB4 UV-Vis spectrophotometer, with a linear Charge-Coupled Device (CCD) array detector (Ocean Optics, Duiven, the Netherlands). The pH measurements were performed with a Horiba F42 pH meter (Tokyo, Japan). The absorbance spectra were collected from 5 different bottles of each sample, with a working range from 200 to 500 nm.

Data analysis

In order to sort the wine samples according to their brand and geographical origin, PCA, CA, discriminant analysis (DA), and PLS-DA were used as multivariate tools. These methods allowed verification of the contribution of each variable to the model and its capacity to discriminate 1 category from another. UV-Vis spectral data were used as response matrix, without prior signal pretreatment, and all data were autoscaled for every variable. The Unscramble 6.11 software (CAMO-ASA, Trondheim, Norway) was used for the PCA and PLS-DA modeling, while CA and LDA were calculated using Infostat software (Córdoba, Argentina).

Results and Discussion

Optimal pH buffer

Prior to obtaining the multivariate models, a study of spectral behavior of wines was performed at 6 different pH values: 1, 3, 5, 7, 10.2, and 11.4. This study assessed the pH value with the best spectral condition, for further use in the multivariate models. Figure 1 shows the results obtained from the different samples at 6 pH values. For the selection of optimal pH, the different spectral curves were analyzed, which showed a peak between 300 and 400 nm at pH 10.2 and 11.4; this peak did not appear in the curves corresponding to lower pH values and, therefore, could be used for classification. Since the spectral curve at pH 11.4 is quite

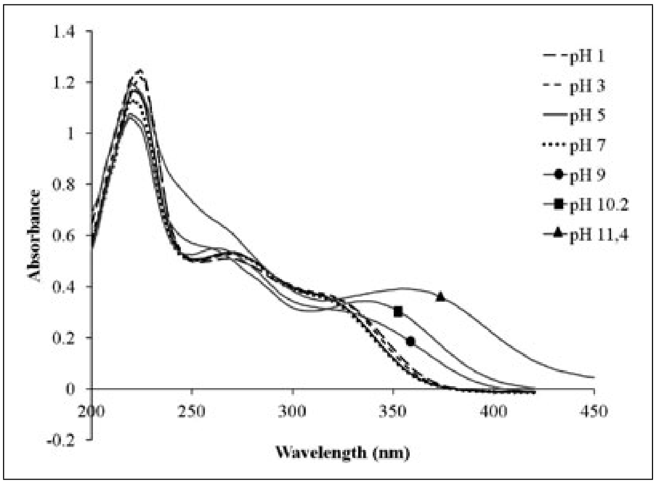

Figure 1–Spectral curves of a wine sample at 6 different pH values.

different from the rest of the curves, we preferred to select pH 10.2 for the analysis.

Principal component analysis

PCA is an unsupervised technique that reduces the dimensionality of responses matrix in a few new variables (principal components), which concentrate the maximum variability and information from the system under study (Forina and others 1986). PCA can be useful to detect groupings, outliers and structures in the data.

PCA was applied to the matrix formed by the UV-Vis spectra corresponding to the white wine samples. The PCA model was built using 190 variables, corresponding to the absorbance values, from 300 to 350 nm. Using the selected variables, the model was obtained using only 2 principal components, which explain 99.95% of the original information. Figure 2 shows the classification obtained by PCA, through the scores plot of samples in the bi-dimensional space formed by the 1st and 2nd PCs. As it can be observed, there are 3 incipient groups corresponding to the 3 Argentinean provinces studied.

Classification was possible due the presence of different compounds in the wines, which depend of the region (Urbano and others 2006). UV-Vis spectrum represents the information based in the composition of absorbent species, as polyphenolic compounds (Cheynier 2005) phenolic, benzoic, and hydroxycinnamic acids and stilbenes, flavanols, and anthocyanins. The composition of these compounds produces the color and taste characteristics of wines (Saurina 2010).

From this exploratory analysis, it can be observed that the 3 provinces are clearly separated, probably due to differences in the proportions of these families of compounds between different geographical origins within the same variety. In accord with Urbano and others (2006), this discrimination cannot be visual since the differences are found in the ultraviolet region; however, UV spectra could be used for the differentiation and classification of wines.

Cluster analysis

Similar to the PCA model, CA is used to classify objects into groups characterized by the values of a set of variables. CA is therefore an alternative to PCA for describing the structure of a data table (Massart and others 1997). CA is an unsupervised method used to examine classificatory structure of data. CA involves techniques that produce a classification from unclassified data, which allows obtaining groups based on their similarities (Camiña and others 2008). CA was carried out using the same data matrix used for PCA (30 samples and 70 variables). Figure 3 shows the dendrogram obtained from samples of the CA. The Ward's method, which calculates the distances between objects of a cluster, was used as amalgamation criterion, while the Euclidean distance was used as association criterion. Once again, the classification was successful and, as a result, 3 groups were obtained in concordance to the PCA model. The top group corresponded to the M group (4 trademarks), the middle group to the RN group (2 trademarks), while the bottom group corresponded to the SJ group (1 trademark). Dot line represents the border from which the classification is significant in the cluster.

Linear discriminant analysis

Discrimination models were developed using the LDA technique. LDA is a supervised learning method in which a classification model is constructed using the data of the precategorized objects into known categories (the training data set), and the calculation algorithm is trained to discriminate the objects (here the wine samples) into the given categories (types) (Sharma 1996; Massart and others 1997; Otto 1999). The goal is to find the allocation rule that gives the highest percentage of correct classification. The LDA classification procedure maximizes the variances between categories, and minimizes the variances within categories (Adams 1995). The power of classification can be calculated as the number of samples correctly predicted in conjunction to the group population: classification power = (number of correctly classified individuals/sample population) × 100 (Sperková and Suchánek 2005). The LDA model was constructed using the original variables used in PCA and CA. For LDA, samples were organized into categories according to wine trademark and its geographical origin. Twenty-five random samples were used to build the model in the training set, and 10 samples were used as the prediction set. Table 2 shows the results of the LDA model, evidencing good fitting, for which a discriminant function that correctly classifies 100% of the samples analyzed by geographical origin (for training and prediction sets) was obtained.

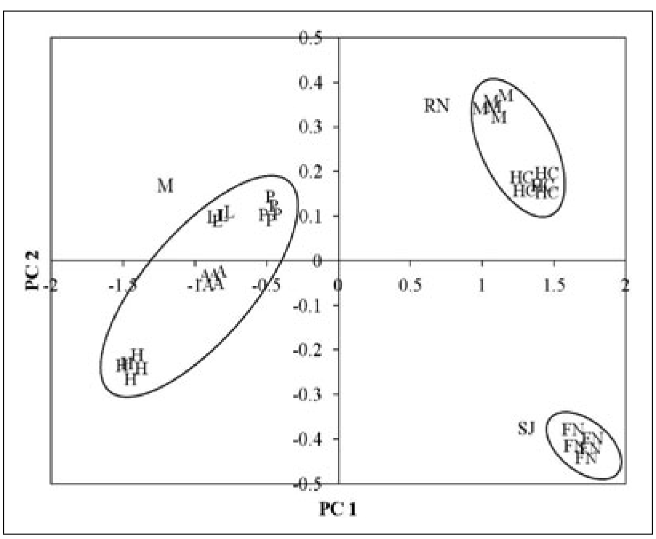

Figure 2–Score plot for the classification of Sauvignon blanc wines from Mendoza (M), Río Negro (RN), and San Juan (SJ).

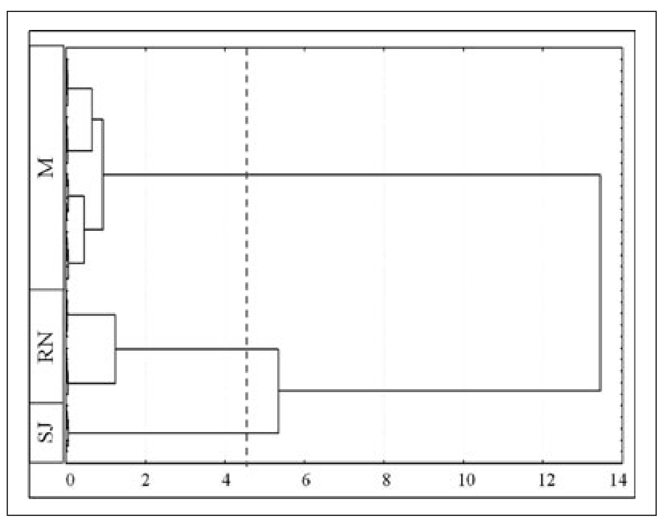

Figure 3–Dendogram of Sauvignon blanc wines obtained by CA.

Table 2–Results of the classification ability of the LDA model for different wines according to their geographical origin.

Training	Correct%	Mendoza	Río Negro	San Juan
Mendoza	100	15	0	0
Río Negro	100	0	7	0
San Juan	100	0	0	3
Total	100	15	7	3
Prediction	**Correct%**	**Mendoza**	**Río Negro**	**San Juan**
Mendoza	100	5	0	0
Río Negro	100	0	3	0
San Juan	100	0	0	2
Total	100	5	3	2

Table 3–Results of the classification of the PLS-DA model according to geographical origin.

Training	Correct%	Mendoza	Río Negro	San Juan
Mendoza	100	16	0	0
Río Negro	100	0	6	0
San Juan	100	0	0	3
Total	100	16	6	3
Prediction	**Correct%**	**Mendoza**	**Río Negro**	**San Juan**
Mendoza	100	4	0	0
Río Negro	100	0	4	0
San Juan	100	0	0	2
Total	100	4	4	2

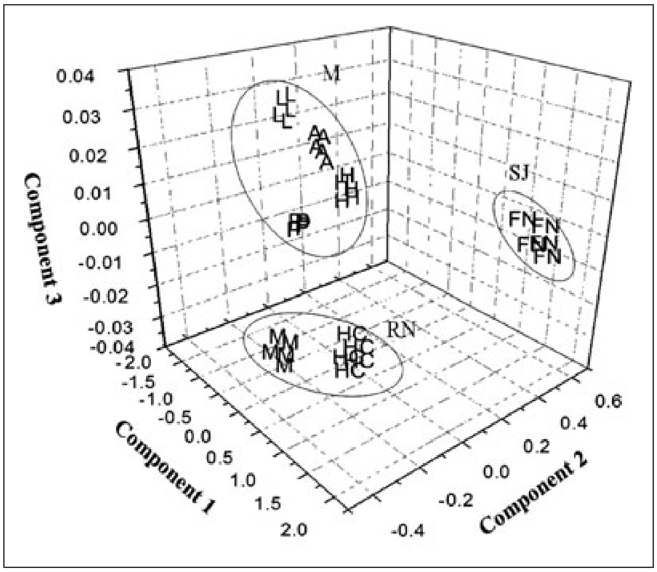

Figure 4–PLS-DA 3D score plot for 35 samples of white wines, from different geographical origin: Mendoza (M), Río Negro (RN), and San Juan (SJ).

Partial least square discriminant analysis

PLS-DA is a variant of partial least-squares regression (PLS). In this technique, a dummy variable is assigned to each sample in the calibration set as a reference value (Mendoza wines = 1, Río Negro wines = 2, and San Juan wines = 3). The PLS-DA model was developed using the same variables than for PCA, which served as a variable preselecting tool (Hernandez and others 2005; Louw and others 2009). The PLS-DA model was obtained using 3 principal components, which explained a 99.8% of information from the original spectral variables. Figure 4 shows the score plot of the 1st 3 PCs of the PLS-DA model. It is similar to the PCA score plot; however, separation of wines according to geographical origin seems to be more obvious. This might be explained by the fact that the PLS-DA algorithm maximizes the variance between groups rather than within the group (Kemsley 1996).

With these 3 PLS components for the model, the observed-predicted classification plot for white wines was obtained, showing an r^2 coefficient = 0.995, abscissa = 0.089, and slope = 0.990, which suggest a good fit of the model (ideal values of $r^2 = 1$, abscissa = 0, and slope = 1). Then the percentage of discrimination was obtained according to the geographical origin (Table 3), obtaining the same results as LDA.

Conclusions

Even though the information contained in the UV-Vis spectra is unspecific and can limit the discrimination possibilities, they can provide similar power of classification comparable to the use of chemical composition, which present high-cost and time-consuming analysis. Thus, UV-Vis spectra present several advantages as simplicity, availability, and minimal sample treatment. In conclusion, UV-Vis spectroscopy and chemometric methods produced a correct classification and correlation between the Sauvignon blanc wine samples analyzed from 3 different geographical origins and 2 vintages. Even though the proposed method is qualitative, it avoids the need of a quantitative method that would require the use of standards, calibration and time-consuming analysis. However, some factors can limit the precision of the classification models, such as the number of samples and the similarities between some wines due to climate, soil, or other characteristics. Therefore, before these methods can be used with confidence by the wine industry, further studies are needed to improve the precision of the developed classification models.

Acknowledgments

The authors thank the Univ. Nacional de San Luis, Univ. Nacional de La Pampa, and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) for financial support. They also thank Prof. Roberto Olsina (Univ. Nacional de San Luis) for providing the Unscrambler 6.11 software; and the Facultad de Agronomía, Univ. Nacional de La Pampa, for providing the InfoStat software.

References

Acevedo FJ, Jimenez J, Maldonado S, Dominguez E. 2007. Classification of wines produced in specific regions by UV-visible spectroscopy combined with support vector machines. J Agric Food Chem 55:6842–9.

Adams MJ. 1995. Chemometrics in analytical spectroscopy. Analytical spectroscopy monographs. Cambridge: The Royal Society of Chemistry. p. 216.

Almeida CMR, Vasconcelos MTSD. 2003. Multielement composition of wines and their precursors including provenance soil and their potentialities as fingerprint of wine origin. J Agric Food Chem 51:4788–98.

Arévalo G, Córdoba C. 2008. Argentina, un gran viñedo. Buenos aires: Albatros. p. 16–33.

Arvanitoyannis IS, Katsota MN, Psarra EP, Soufleros EH, Kallithraka S. 1999. Application of quality control methods for assessing wine authenticity: use of multivariate analysis (chemometrics). Trends Food Sci Technol 10:321–36.

Bentlin FRS, dos Santos CMM, Flores EMM, Pozebon D. 2012. Lanthanides determination in red wine using ultrasound assisted extraction, flow injection, aerosol desolvation and ICP-MS. Anal Chim Acta 710:33–9.

Camiña JM, Cantarelli MA, Lozano VA, Boeris MS, Irimia ME, Gil RA, Marchevsky EJ. 2008. Chemometric tools for the characterization of honey produced in La Pampa, Argentina, from their elemental content, using inductively coupled plasma optical emission spectrometry (ICP-OES). J Apicult Res 47:102–7.

Cheynier V. 2005. Polyphenols in foods are more complex than often thought. Am J Clin Nutr 81:223S–9S.

De Villiers A, Alberts F, Lynen F, Crouch A, Sandra P. 2003. Evaluation of liquid chromatography and capillary electrophoresis for the elucidation of the artificial colorants brilliant blue and azorubine in red wines. Chromatographia 57:393–7.

Di Paola-Naranjo RD, Baroni MV, Podio NS, Rubinstein HR, Fabani MP, Badini RG, Inga M, Ostera HA, Cagnoni M, Gallegos E, Gautier E, Peral-García P, Hoogewerff XJ, Wunderlin D. 2011. Fingerprints for main varieties of Argentinean wines: terroir differentiation by inorganic, organic, and stable isotopic analyses coupled to chemometrics. J Agric Food Chem 59:7854–65.

Fabani MP, Arrúa RC, Vázquez F, Diaz MP, Baroni MV, Wunderlin DA. 2010. Evaluation of elemental profile coupled to chemometrics to assess the geographical origin of Argentinean wines. Food Chem 119:372–9.

Forina M, Armanino C, Castino M, Ubigli M. 1986. Multivariate data analysis as a discriminating method of the origin of wines. Vitis 25:189–201.

Forina M, Lanteri S, Armanino C. 1987. Chemometrics in Food Chemistry. In: Dewar MJS, Dunitz JD, Hafner K, Heilbronner E, Ito S, Lehn JM, Niedenzu K, Raymond KN, Rees CW, Wögtle F, Witting G, editors. Topics in current chemistry. Berlin: Springer-Verlag. p. 93–143.

Granato D, Katayama F, Chizuko U, de Castro IA. 2011. Phenolic composition of South American red wines classified according to their antioxidant activity, retail price and sensory quality. Food Chem 129: 366–73.

Granato D, Katayama F, Chizuko U, de Castro IA. 2012. Characterization of red wines from South America based on sensory properties and antioxidant activity. J Sci Food Agric 92:526–33.

Green JA, Parr WV, Breitmeyer J, Valentin D, Sherlock R. 2011. Sensory and chemical characterisation of Sauvignon blanc wine: influence of source of origin. Food Res Intl 44:2788–97.

Hernandez OM, Fraga JMG, Jimenez AI, Jimenez F, Arias JJ. 2005. Characterization of honey from the Canary Islands: determination of mineral content by atomic absorption spectrophotometry. Food Chem 93:449–58.

Inst. Nacional de Vitivinicultura (INV). 2011. Exportaciones argentinas de productos vitivinícolas. Available from: http://www.inv.gov.ar/est_exportaciones.php. Accessed 2012 Sep 11.

Inst. Nacional de Vitivinicultura (INV). 2004. Ley de Denominación en Origen N° 25.163. Decreto reglamentario N° 57/2004, Argentina.

Inst. Instituto Nacional de Vitivinicultura (INV). 2012. Operativo de actualización del registro nacional de viñedos 2010/2011. Argentina: Departamento de estudios vitivinícolas – Inst. Nacional de Vitivinicultura. p. 50–1.

Intl. Organisation of Vine and Wine (OIV). 2012. Statistical report on world vitiviniculture 2012. Intergovernmental organisation. 18 rue d´Aguesseau 75008. Paris. Available from: http://www.oiv.int. Accessed 2012 Sep 11.

Kemsley EK. 1996. Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods. Chemom Intell Lab Syst 33:47–61.

Louw L, Roux K, Tredoux A, Tomic O, Naes T, Nieuwoudt HH, van Rensburg P. 2009. Characterization of selected South African young cultivar wines using FTMIR spectroscopy, gas chromatography, and multivariate data analysis. J Agric Food Chem 57:2623–32.

Martens H. 1983. A layman´s guide to multivariate data analysis. In: Martens H, Russwurm H Jr, editors. Food research and data analysis. London: Elsevier Applied Science. p. 473–92

Massart DL, Vandeginste BGM, Buydens LMC, De Jong S, Lewi PJ, Smeyers-Verbeke J. 1997. Handbook of chemometrics and qualimetrics. Vol. B. Amsterdam: Elsevier, p. 57–68; 207–42.

Meites L. 1982. Handbook of analytical chemistry. 1st ed. N.Y.: Mc Graw-Hill. p. 5–8.

Otto M. 1999. Chemometrics, statistics and computer. Application in analytical chemistry. 1st ed. Weinheim: Wiley-VCH. p. 41–7, 124–68.

Saurina J. 2010. Characterization of wines using compositional profiles and chemometrics. Trends Anal Chem 29:234–45.

Sharma S. 1996. Applied multivariate techniques. 1st ed. N.Y.: Wiley. p. 1–15, 59–66, 264–7, 287–93.

Sperková J, Suchánek M. 2005. Multivariate classification of wines from different Bohemian regions (Czech Republic). Food Chem 93:659–63.

Taylor VF, Longerich, HP, Greenough JD. 2003. Multielement analysis of Canadian wines by inductively coupled plasma mass spectrometry (ICP-MS) and multivariate statistics. J Agric Food Chem 51:856–60.

Urbano M, Luque de Castro MD, Pérez, PM, Gómez-Nieto MA. 2005. Study of spectral analytical data using fingerprint and scaled similarity measurements. Anal Bioanal Chem 381:953–63.

Urbano M, Luque de Castro MD, Pérez PM, García-Olmo J, Gómez-Nieto MA. 2006. Ultraviolet-visible spectroscopy and pattern recognition methods for differentiation and classification of wines. Food Chem 97:166–75.

Wang JF, Geil PH, Kolling DRJ, Padua GW. 2003. Analysis of zein by matrix-assisted desorption/ionization mass spectrometry. J Agric Food Chem 51:5849–54.