Available online at www.sciencedirect.com
SCIENCE DIRECT®
Food Chemistry 97 (2006) 166–175
Food Chemistry
www.elsevier.com/locate/foodchem

Analytical, Nutritional and Clinical Methods

Ultraviolet–visible spectroscopy and pattern recognition methods for differentiation and classification of wines

Manuel Urbano [a], María D. Luque de Castro [a,*], Pedro M. Pérez [b], Juan García-Olmo [c], Miguel A. Gómez-Nieto [d]

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

Received 6 September 2004; received in revised form 6 April 2005; accepted 2 May 2005

Abstract

The feasibility of developing a qualitative model based on cheap and simple instrumentation to differentiate and classify wines from the apellation d'origine "La Mancha" (Spain) has been studied. The criteria for discrimination were origin, grape variety and ageing process. Ultraviolet–visible spectroscopy was used for the development of a inexpensive and simple screening approach. Once spectra were collected, a data exploratory analysis was carried out in order to both show trends hidden in the data matrix from the sample spectra and study the characteristics of the models thus developed. Principal components analysis and soft independent modelling of class analogy were used for the exploratory analysis and the development of classification models, respectively. The ultraviolet region has for the first time been used for the discrimination of types of wines. This region is of great importance for differentiation of wines according to both origin within a same apellation d'origine and ageing process. The sum of false positives and false negatives – the criteria used for evaluating errors – were not higher than 25%, and the classification of wines according to the origin zone yielded the best values (10%). Sample preparation was not necessary.

Keywords: Ultraviolet–visible spectroscopy; Wine classification; Pattern recognition methods

1. Introduction

Quality control of foods is often based on modern and sophisticated instruments that involve high costs and require well-trained analysts (De Villiers, Alberts, Lynen, Crouch, & Sandra, 2003; Encinar, Sliwka-Kaszynska, Polatajko, Vacchina, & Szpunar, 2003; Flurer, 2003; Wang, Geil, Kolling, & Padua, 2003). However, the results from these instruments are often not obtained in time despite being supported on modern techniques. This is owing to the necessity of steps, previous to measurement, focused on improving the accuracy and precision of the results. Quantification of components either involved in a fraudulent wine elaboration or those that constitute a key family in wine from a given apellation d'origine is, most times, a long and expensive step (Cullere, Aznar, Cacho, & Ferreira, 2003; Legin et al., 2003; Lopez, Aznar, Cacho, & Ferreira, 2002). A present trend in analytical chemistry is the development of methodologies able to provide "fitness for purpose" results, which take into account aspects related with the importance of time against accuracy achieved. These aims are often supported on qualitative aspects in

* Corresponding author. Tel./fax: +34 95 75 18 61 5.
E-mail address: qa1lucam@uco.es (M.D. Luque de Castro).

0308-8146/$ - see front matter
doi:10.1016/j.foodchem.2005.05.001

contrast to quantitative results (Trullols, Ruisánchez, & Rius, 2004).

Chemometric techniques allow construction of models to characterise target samples within previously defined and validated groups. Multivariate analysis is a powerful tool that permits to extract qualitative and quantitative chemical information from large data sets. Thus, qualitative techniques as cluster analysis (CA), principal components analysis (PCA), linear discriminant linear analysis (LDA), *k*-nearest neighbours (*k*-NN), soft independent modelling of class analogy (SIMCA), etc., are aimed at improving the quality of food products (Kos, Lohninger, & Krska, 2003; Tura, Prenzler, Bedgood, Antolovich, & Robards, 2004; Yu & MacGregor, 2003).

Wine industry has employed the above commented tools with several objectives; namely: wine classification based on either grape variety or climate factors; assessment of the authenticity of wine; study of different brownings; etc. The datasets used were of different nature, namely: the concentration of phenolic compounds (De la Presa & Noble, 1995; García-Parrilla, González, Heredia, & Troncoso, 1997), the composition of volatile compounds (García-Jares, García-Martín, Marino, & Torrijos, 1995; Guth, 1997), amino acids content (Etiévant, Schlich, & Bouvier, 1998), concentration of metal ions (Baxter, Crews, Dennis, Goodall, & Anderson, 1997), isotopic determination (Martin, Guillou, & Martin, 1998), etc.

Spectroscopy has been widely used for the development of classification models. Thus, near infrared spectroscopy (NIRS), Fourier transform infrared (FT-IR), mass spectrometry (MS), etc., have been used for the differentiation and classification of samples in various areas (Downey, McIntyre, & Davies, 2002; Pérez-Pavón et al., 2003; Reeves & Zapf, 1998). Although quantitative methods based on these techniques have been developed in the enological area, there are few approaches regarding to their qualitative use. Visible and NIR regions have been used to discriminate between white wines of two varietal origins (Cozzolino, Smyth, & Gishen, 2003). The visible region has also been used in an indirect way for certification of roséwines from the apellation d'origine "Rioja" (Spain) (Meléndez, Sánchez, Íñiguez, Sarabia, & Ortiz, 2001). The authors employed six parameters obtained from the absorption spectrum in the visible range for classification. The ultraviolet zone has not been employed for differentiation and classification of wines.

The aim of this work was to study the use of both the ultraviolet and visible zones in order to obtain spectra for classification of wines based on criteria more restrictive than those above commented. Criteria as ageing process and origin within an apellation d'origin have been studied. Varietal discrimination has also been evaluated. Thus, the approach here reported is aimed at studying the feasibility of developing cheap and simple models – using only the ultraviolet and visible zones – for the differentiation and classification of wines without the help of trained specialists.

2. Materials and methods

2.1. Samples

Different samples – red and white wines; young and aged wines; wines from different zones within the "La Mancha" apellation d'origine ("Quintanar de la Orden", "Fuente de Pedro Naharro", "Mota del Cuervo", "Corral de Almaguer" and "Villacañas") and grape varieties ("Cencibel" and "Cabernet Sauvignon") – were used in the study. Thus, the number of samples employed in the calibration and validation steps was 120.

2.2. Apparatus and procedure

The instrument employed for spectra collection was an Agilent 8453E UV–vis spectroscopy system (Agilent Technologies, Waldbronn, Germany). The spectra were collected using a UV–vis Chem Station Rev. A.06.03 (Hewlett–Packard, USA). The absorbance spectra were recorded in duplicate. The working range was 300–800 nm.

The instrument was equipped with a quartz cell with a pathlength of 0.1 cm. This short pathlength enabled to obtain absorbance values within the appropriate range regarding to accuracy and precision specified by the spectrophotometer characteristics manual.

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

Duplicate spectra from two aliquots of each sample were averaged. The Unscrambler 7.5 (Camo Process AS, Oslo, Norway) was used for data processing. PCA and SIMCA (Esbensen, 2002; Vandeginsten et al., 1998; Wold, 1976) were the multivariate pattern recognition methods.

2.3.1. Principal components analysis (PCA)

Principal components analysis has been extensively used for visualisation of hidden trends in a data matrix M consisting of n objects defined by m variables. In this work, the objects were the sample spectra and the variables were the wavelengths. This can also be seen as an m-dimensional space, in which each wavelength defines one dimension. In order to show trends or different data structures hidden in the M matrix, a new c-dimensional space can also be built from the original m-dimensional. The new dimensions, principal components – PCs – are built taking into account the maximum variance of data and the requirements about an orthogonal space. The number of PCs are much lower than the number of original variables, mainly in spectral analysis, due to the linear combination of the original variables in order to form the PCs thus removing co-linearity between variables. Objects plotted in the new space – score plot – often show trends that, in spite of having to be interpreted and explained, constitute a first step in subsequent modelling for samples classification.

2.3.2. Soft independent modelling of class analogy (SIMCA)

SIMCA, which is a supervised pattern recognition technique in contrast to the not supervised PCA, is considered a key chemometric approach for classification. This technique enables to classify the samples into an already existing group, assigning new objects to the class to which they show the largest similarity. SIMCA is strongly based on PCA, because each class is defined by an independent PCA, taking into account the optimal number of PCs for each class, which is endowed with a specific data structure.

The SIMCA classification process consists of two stages, namely: the training stage, in which the individual models of the data classes are developed, and the testing stage, in which new samples (not used in the training stage) are classified within the established class models to evaluate their efficiency.

The choice of the SIMCA technique in contrast to other supervised pattern recognition techniques is based on the modelling properties of SIMCA, which provides approaches more versatile than those obtained using discriminant techniques. The classification of a sample in one or several classes, or in none of them, is possible with SIMCA, while discriminant techniques only permits to classify a sample in a unique class.

3. Results and discussion

3.1. Exploratory analysis

3.1.1. Visualising trends or groups in white wines

PCA was applied to the matrix formed by the UV–vis spectra corresponding to samples of white wines. The maximum number of PCs was set at 10. The first two components explained almost all the data variance.

The first criterion was to differentiate wines from different zones within the same apellation d'origine, namely: "Quintanar de la Orden" (identified by A), "Fuente de Pedro Naharro" (B), "Mota del Cuervo" (C) and "Corral de Almaguer" (D). Fig. 1(a) shows the plot of samples in the bi-dimensional space formed by the first two PCs. As can be seen, there are four incipient groups corresponding to the four zones. Thus, UV–vis spectra can be used for the differentiation of wines as a function of wine origin within the same apellation d'origine.

First derivative is often used as a mathematical pre-processing for UV–vis spectra in order to enlarge the differences between them. This treatment was carried out with the aim of improving the discrimination above achieved. Fig. 1(b) shows the more significant differentiation between groups. The clusters corresponding to C and D were very close, but the difference between them was clearer after re-scaling. Discrimination was possible because of the different concentration depending on zones of families of compounds that absorb UV–vis radiation.

In order to find out these compounds, loadings for the first two PCs were plotted as shown in Fig. 2. Loading vectors can be considered as the bridge between the initial variables space and the PCs space. As commented before, each PC is a linear combination of the entire initial variables; that is, of all the wavelengths. The coefficients of this combination are called loadings, and each wavelength has a loading. The higher the loadings the higher the influence of the corresponding wavelengths in the explanation of the data variance. Fig. 2 shows that the key wavelengths for the discrimination of groups were in the range 300–400 nm; that is, the ultraviolet region. Compounds present in wine that absorb in this region are esters from hydroxicynamic acids (Flanzy, 2000). Thus, two conclusions can be obtained from this exploratory analysis. The first is that there are differences between the proportion of these families between different zones within an apellation d'origine. The second is that this discrimination cannot be visual because the differences lead in the ultraviolet region.

3.1.2. Visualising trends or groups in red wines

The first criterion used for the differentiation was also origin zone. Three zones were taken into account, namely: "Fuente de Pedro Naharro" (J), "Quintanar de la Orden" (K) and "Villacañas" (L). The major differences were found between wines from "Quintanar de la Orden" and those from "Villacañas". Wines from "Fuente de Pedro Naharro" were located between the other two zones. The differentiation was worse than that achieved for white wines.

Other criterion used to differentiate red wines was the grape variety used for obtaining the two groups of wine. The varieties employed were "Cabernet Sauvignon" (S) and "Cencibel" (T). The score plot is shown in Fig. 3(a). Two trends can be observed: T samples, which are the most abundant, provided a swarm of points that comprises a large zone and S samples were located in the below-right part of the plot. Loadings from the visible region were higher than those obtained from white wines. Obviously, this is due to the red colour.

The ageing process of wines was also studied. Aged (Y) and no-aged wines (Z) were considered. The former

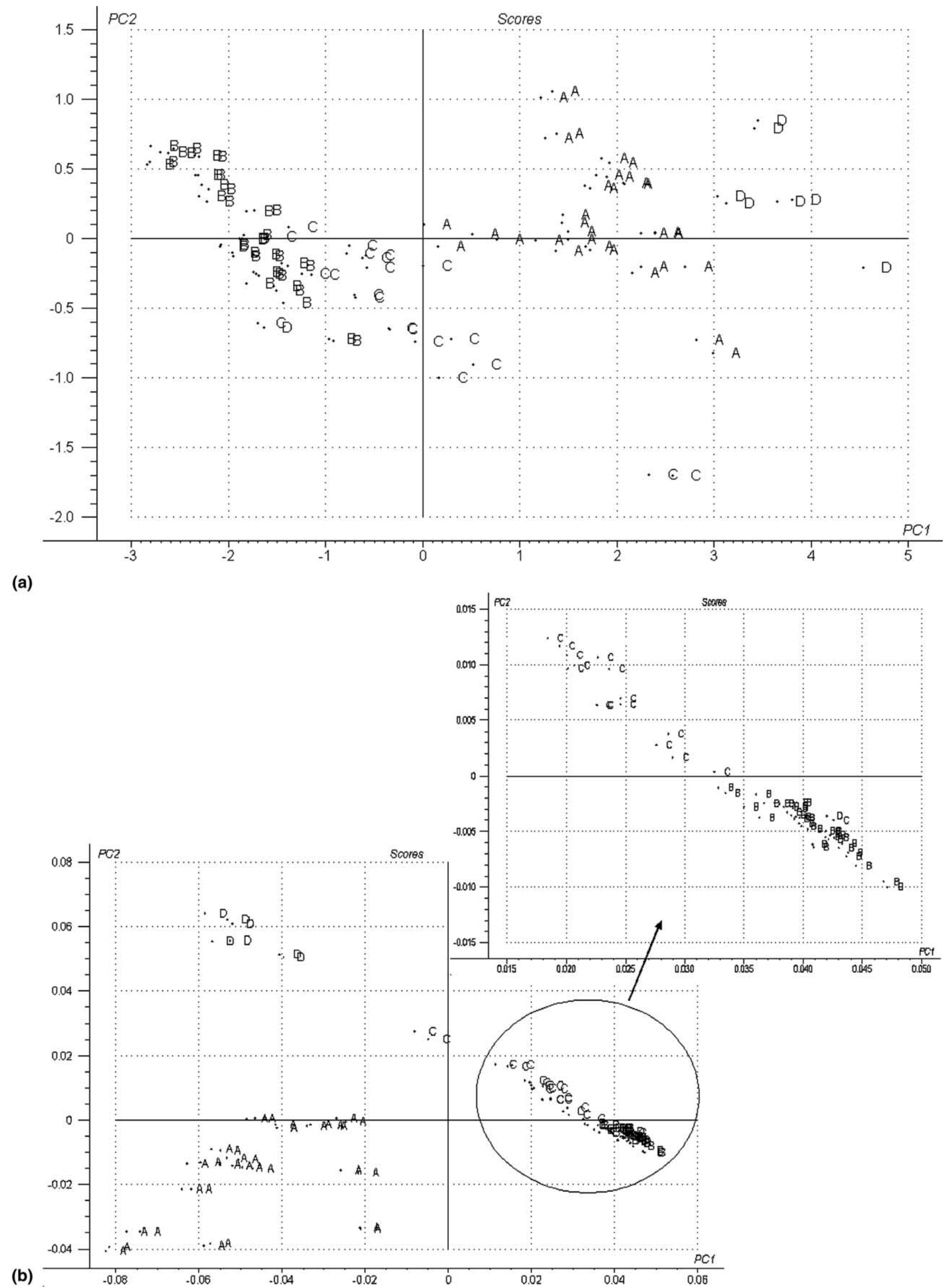


Fig. 1. Plot of white wines samples in the space formed by the first two PCs using: (a) original spectra data; (b) first derivative spectra data.

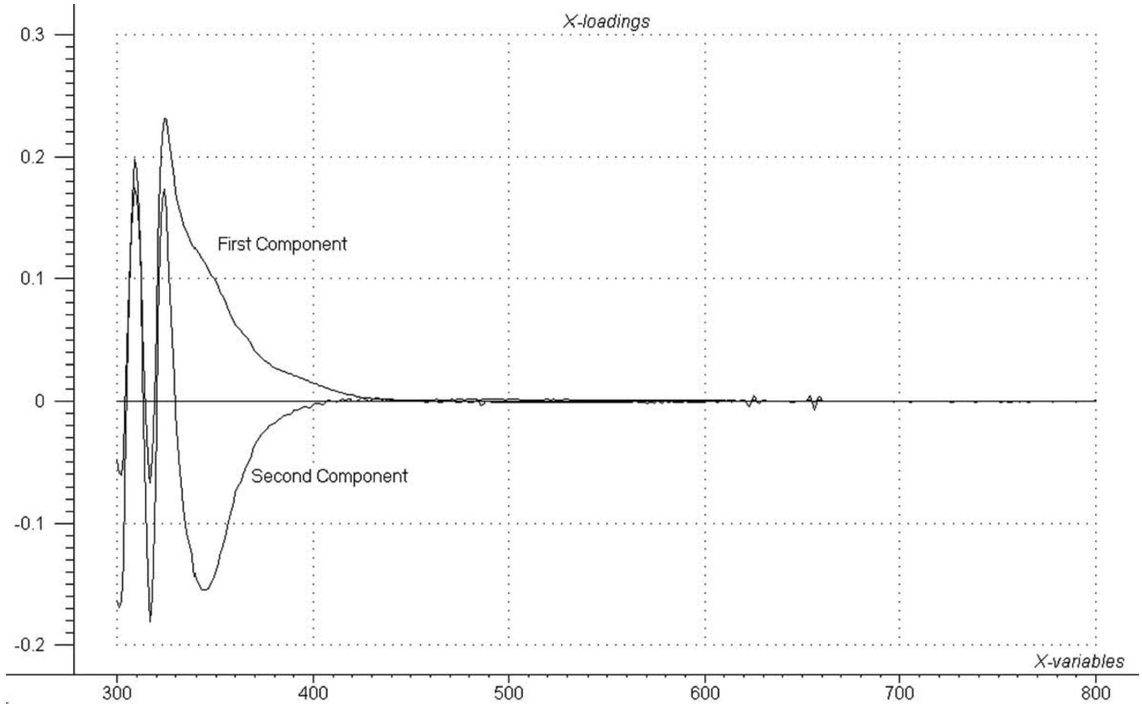


Fig. 2. Loadings plot for the two first components in white wines.

gathered in the right side and the latter in the left side of Fig. 3(b).

A higher dispersion was observed in the PCAs for red wines as compared with that obtained for white wines, which can be due to the higher complexity of the composition of red wines.

3.2. Development of classification models

3.2.1. Model for classifying white wines according to origin of the grape

A model for the classification of three types of wines according to zone of the grape was developed. The classes, following the nomenclature employed in Section 3.1.2, were named A ("Quintanar de la Orden"), B ("Fuente de Pedro Naharro") y C ("Mota del Cuervo"). Independent principal components analysis were carried out for each class. The first derivative spectra were employed. The prediction capacity of the model was studied by external validation using both samples of these classes that had not been used in the training stage, and samples from "Corral de Almaguer" (D), that has not been modelled because of the small number of samples. The significance level was set at 5%.

Coomans plots (Esbensen, 2002) were used for evaluating the results from the classification. These plots provide the orthogonal distance from all new objects to two selected classes at the same time. The critical cut-off class membership limits are also obtained from those plots. If an object belongs to a model (class) it should fall within the membership limit, which is on the left of the vertical line or below the horizontal line in Fig. 4(a).

As can be seen in Fig. 4(a), the validation samples corresponding to origin zones for grape A and B were all within the limit of each membership. Nevertheless, B samples were placed in the class C membership zone in the Cooman plot for the classes A and C shown in Fig. 4(b); thus, too many false positives were obtained in the classification table (the 50% of B samples were also assigned to class C). This is owing to the proximity of both groups – B and C – the major dispersion of the class B and, mainly, the existence of two C samples within the swarm of points corresponding to B zone, as shown in the PCA in Fig. 1(b). Better results were achieved when these two C samples were not used in the modelling of class C – that is, they were considered as outliers. Almost all the validation samples corresponding to zone C were classified correctly in class C, and all the validation samples corresponding to zone D were not classified in any class, as expected. The sum of percent of false positives and false negatives was 10% and the number of samples in the testing set was 32.

In order to explain the classification data, the concept of *Variable Discrimination Power* (Esbensen, 2002) was used; this gives information about the power of a variable to discriminate between any two models. A value close to 1 indicates no discrimination power at all, while a high value, e.g., greater than 3, indicates good discrimination for a given variable.

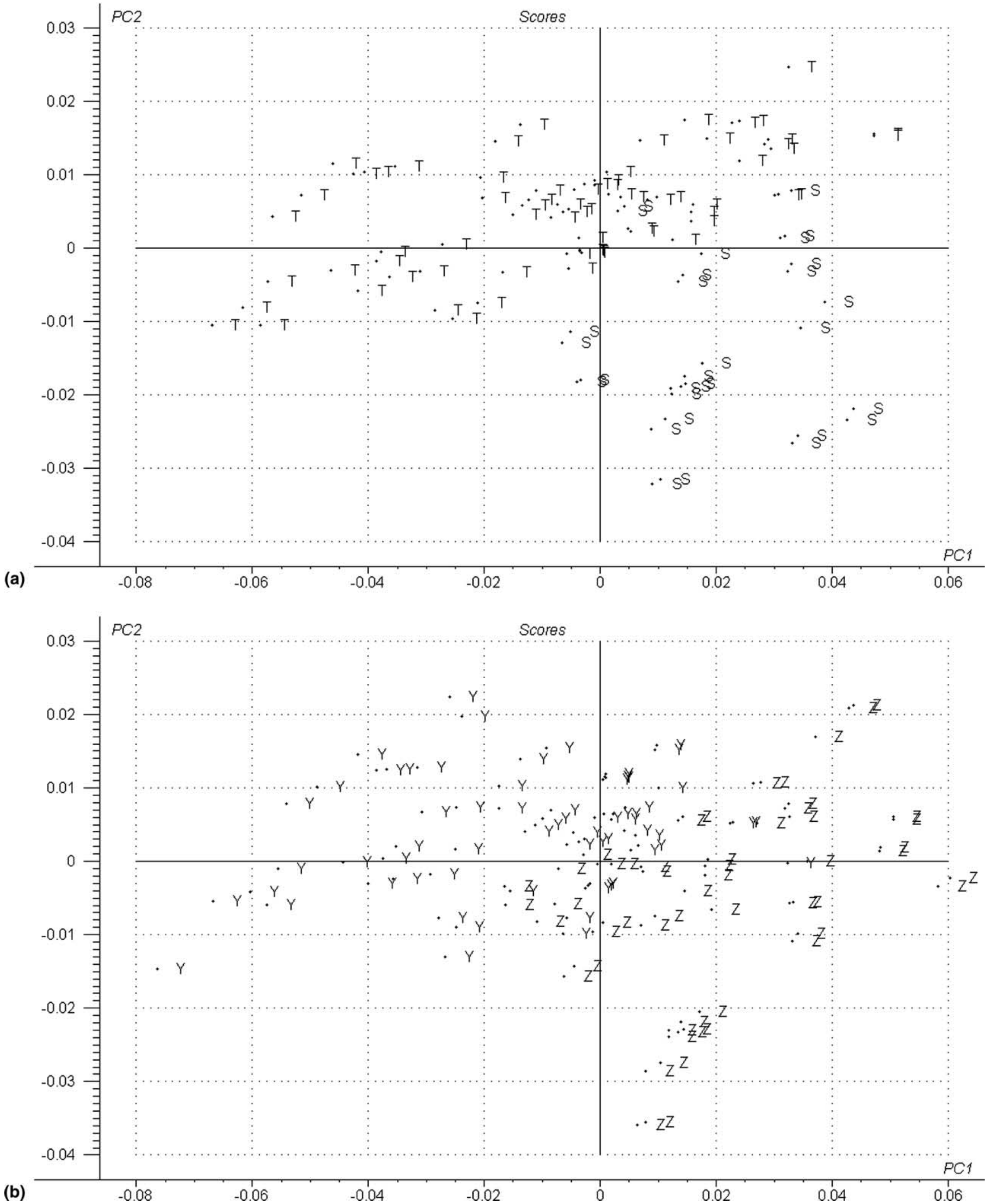
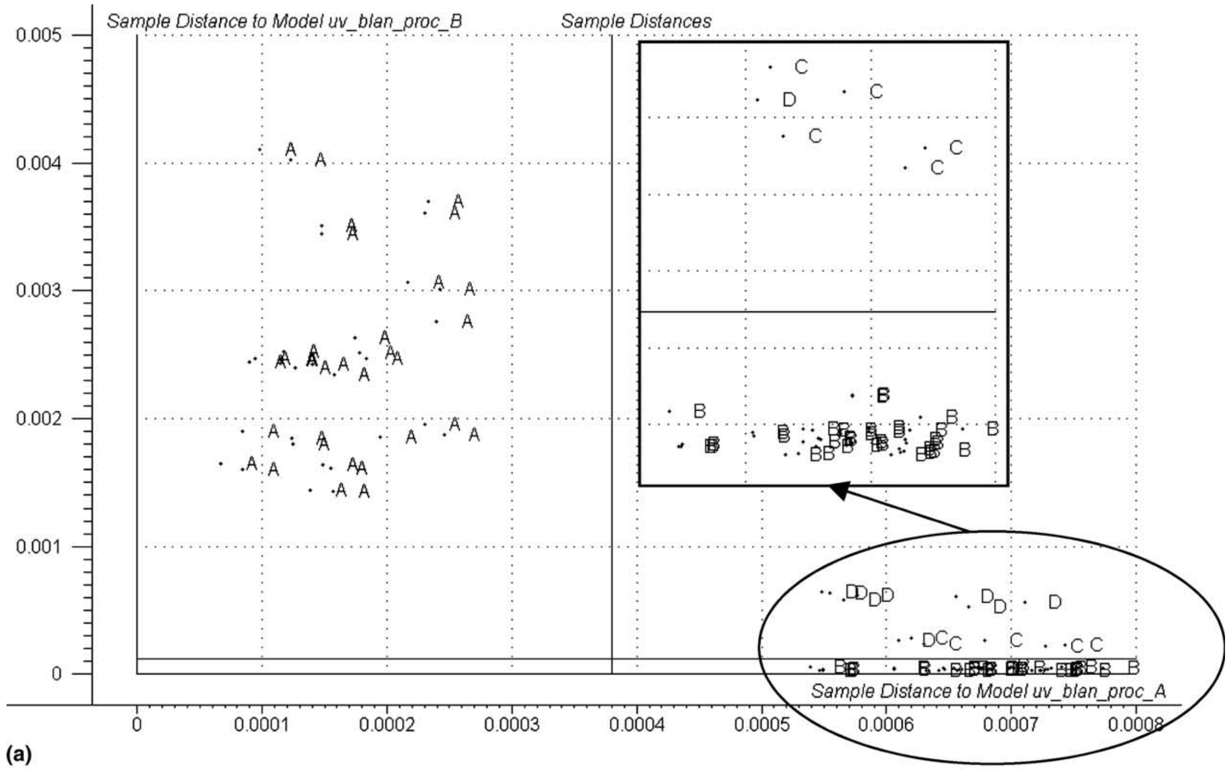


Fig. 3. (a) PCA for red wines according to the grape variety, (b) PCA for red wines according to the ageing process.

The discrimination values were much higher in the ultraviolet than in the visible region for classes A–B. For this reason, the classification of wines from "Quintanar de la Orden" (A) and wines from "Fuente de Pedro Naharro" (B) was based on the differences in the concentration of esters from hydroxicynamic acids, which absorb in the ultraviolet region. Distinction between wines from "Quintanar de la Orden" (A) and "Mota del Cuervo" (C) was explained by the same effect.



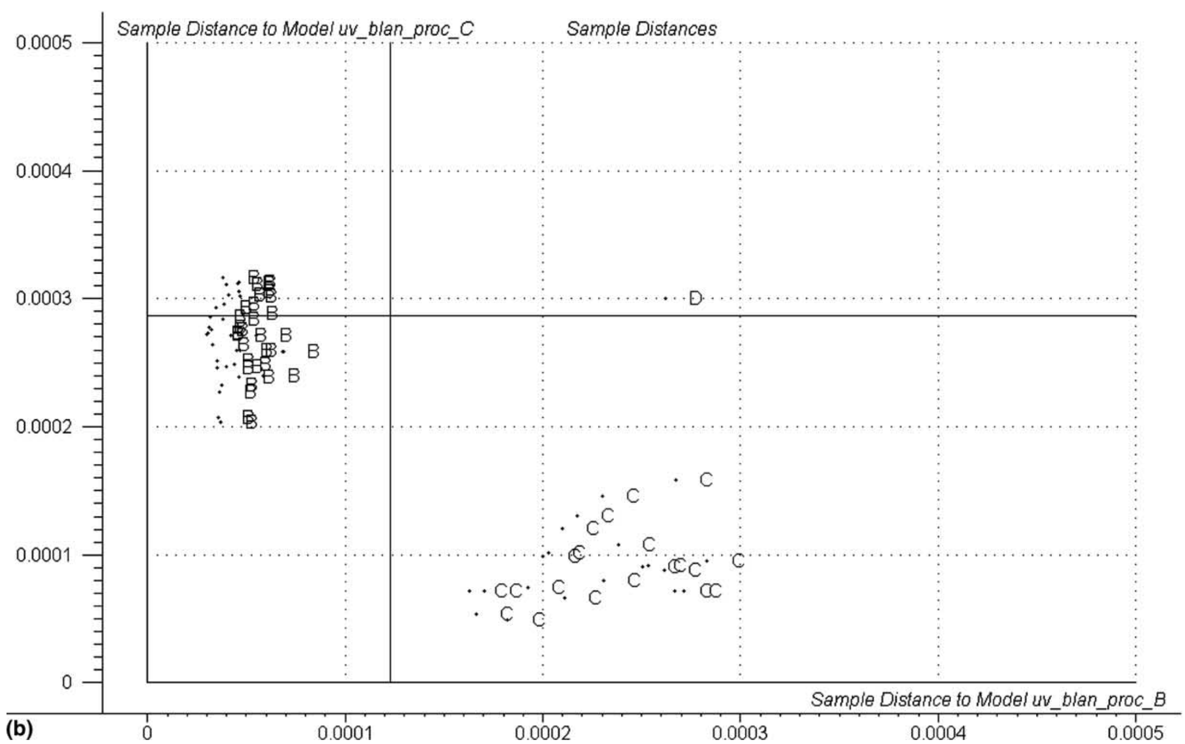


Fig. 4. (a) Cooman plot for the classes A and B, (b) Cooman plot for the classes B and C.

The discrimination values in for classes C–B were lower than those corresponding to classes A–B. This fact explains that classes B and C were close, as compared with the distance from A. The visible zone between 620 and 640 nm provided a high discrimination power for classes B and C. Phenolic compounds are responsible for the slightly green colour present in white wines; thus, the proportion of these compounds is different between B and C wines.

3.2.2. Model for classifying red wines according to origin of the grape

Classes corresponding to wines from "Quintanar de la Orden" (K) and wines from "Villacañas" (L) were developed using independent principal component analysis. External validation was carried out by samples that have not been used for models development.

The Distance vs. Leverage plot (Si vs. Hi plot) (Esbensen, 2002) was used in the evaluation of the external validation. It shows the limits used in the classification, both for the distance to the model (Si) and of the leverage (Hi). The objects within these limits have a high probability to belong to the class at the chosen significance level. They summarise the information contained in the model.

Fig. 5 shows the Si vs. Hi plot for the validation set of the model composed by classes K and L at 5% significance level. For the class K (Fig. 5(a)) almost all the validation objects were within the leverage limit. As can be seen in the figure, two samples were false negatives. They came from two zones placed in the border of zone K. Fig. 5(b) shows the classification achieved for class L. Four samples of the K group were plotted within the class L membership limit; thus, these samples were false positives. The classification of L samples was successful. The sum of false positives and false negatives was 20% and the number of samples for the test set was 30.

On the other hand, the zone between 500 and 560 nm enables the highest discrimination power for red wines origin, in contrast to the discrimination of the ultraviolet region for white wines. In this region, the absorbent compounds are anthocyans.

3.2.3. Model for classifying red wines according to grape variety

Classes corresponding to red wines from "Caubernet Sauvignon" grape (S) and red wine from "Cencibel" grape (T) were modelled. Class models were then externally validated. Figs. 5(c) and (d) show the Si vs. Hi plots of the validation set for the classes S and T, respectively. The error – sum of false positives and false negatives – was 25%. The test set was composed by 30 samples and the significance level was set at 5%. The larger data dispersion, as shown in the PCA in Fig. 3(a) under the variety criterion, meant a higher error.

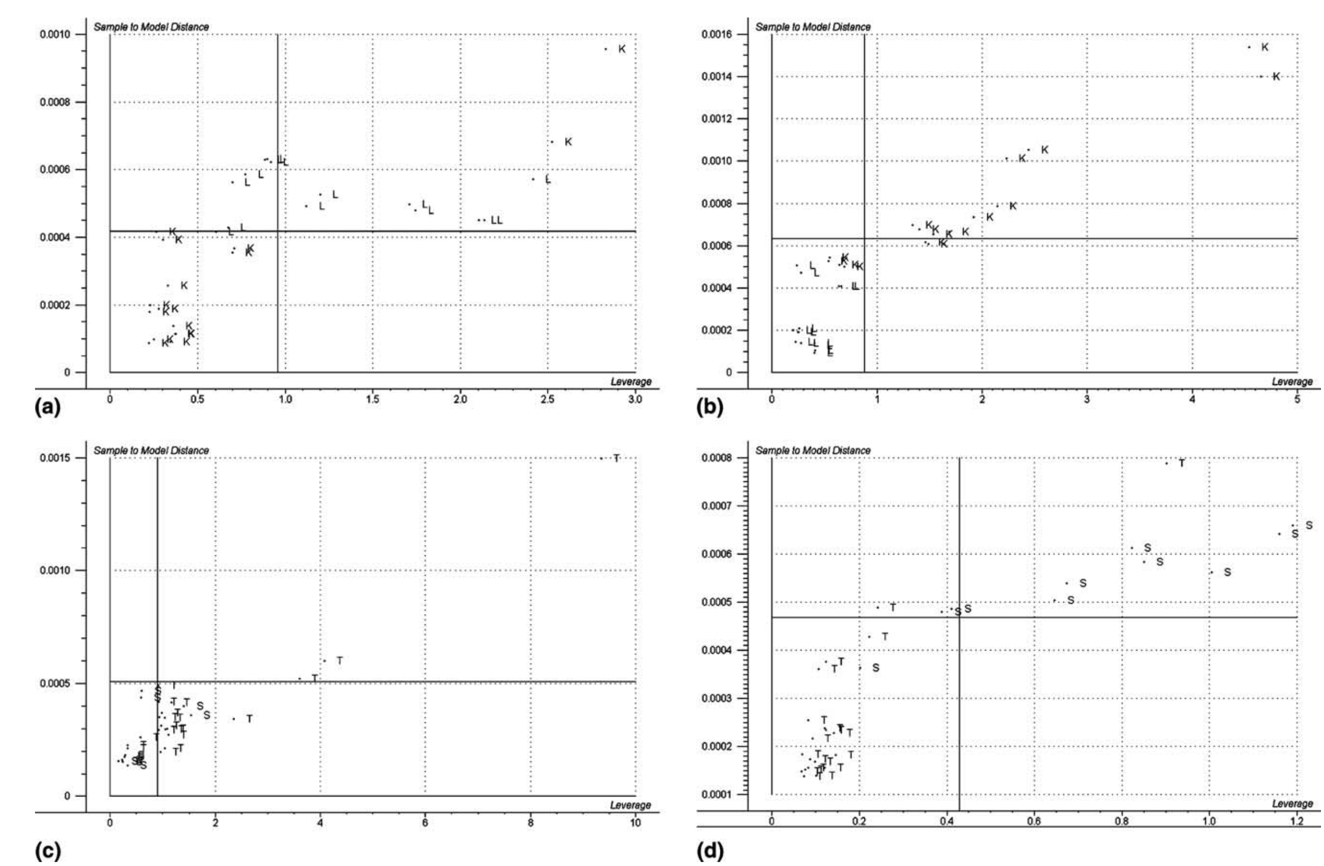


Fig. 5. (a) Si vs. Hi plot for the class K, (b) Si vs. Hi plot for the class L, (c) Si vs. Hi plot for the class S, (d) Si vs. Hi plot for the class T.

3.2.4. Model for classifying red wines according to the ageing process

In this case, the model developed was aimed at classifying red wines into aged wines and no-aged wines. Thus, two groups of wine were taken in order to model the classes, namely: Y class, non-aged wines, and Z class, aged wines. The sum of false positives and false negatives was 25%. The test set was composed by 30 samples and the significance level was set at 5%.

The discrimination power of the variables of the last two models was lower than that of the first model; this explains the higher number of false positives and false negatives. The last one shows the higher discrimination values in the ultraviolet region. This fact means that the quantitative and qualitative evolution of the phenolic fraction in the ageing process yields differences in ultraviolet absorbent compounds.

4. Conclusions

Differentiation and classification of wine under various criteria have been achieved in the approach presented in this work, using cheap and simple instrumentation. The error in the prediction depends on the type of wine to which the method is applied (namely, white or red wine).

The classification of the wines from different origins within the same apellation d'origine is around 90%. In addition, the classification of wines according to grape variety and ageing process is better than 75%. The model thus developed supplies a simple, inexpensive screening tool for wines from the apellation d'origine "La Mancha" from the data provided by an array diode UV–vis spectrophotometer. Other advantage of the model is the time required for analysis – about 10 min were enough for classification into a given class.

The ultraviolet region has been used for the first time for differentiation of wines. This region is the key for discrimination of wines according to the zone within the same apellation d'origine. In addition, the joint use of the ultraviolet and visible regions enables a better classification of wine according to the ageing process, as compared with the single use of one of the spectrum zones.

The phenolic compounds from the secondary metabolism of vegetables VER are excellent discriminators for wines. This characteristic is due to the variability of those compounds as a function of soils, grape variety, production conditions, etc.

Acknowledgements

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

References

Baxter, J. M., Crews, M. E., Dennis, J., Goodall, I., & Anderson, D. (1997). The determination of the authenticity of wine from its trace elements composition. *Food Chemistry, 60*, 443–450.

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

Cullere, L., Aznar, M., Cacho, J., & Ferreira, V. (2003). Fast fractionation of complex organic extracts by normal-phase chromatography on a solid-phase extraction polymeric sorbent. Optimization of a method to fractionate wine flavor extracts. *Journal of Chromatographia A, 1017*, 17–26.

De la Presa, C., & Noble, A. C. (1995). Descriptive analysis of three white wines varieties from Penedes. *American Journal of Enological and Viticulture, 46*, 5–9.

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*, 393–397.

Downey, G., McIntyre, P., & Davies, A. N. (2002). Detecting and quantifying sunflower oil adulteration in extra virgin olive oils from the Eastern Mediterranean by visible and near infrared spectroscopy. *Journal of Agriculture and Food Chemistry, 50*, 5520–5525.

Encinar, J. R., Sliwka-Kaszynska, M., Polatajko, A., Vacchina, V., & Szpunar, J. (2003). Methodological advances for selenium speciation analysis in yeast. *Analytica Chimica Acta, 500*, 171–183.

Esbensen, K. H. (2002). *Multivariate data analysis – In practice*. Oslo: Camo Process AS.

Etiévant, P., Schlich, P., & Bouvier, J. C. (1998). Varietal and geographic classification of French red wines in terms of elements, amino acids and aromatic alcohols. *Journal of Science and Food Agriculture, 48*, 25–41.

Flanzy, C. (2000). *Enología: fundamentos científicos y tecnológicos (enology: scientific and technological fundamentals)*. Madrid: AMV-Mundi Prensa.

Flurer, C. L. (2003). Analysis of antibiotics by capillary electrophoresis. *Electrophoresis, 24*, 4116–4127.

García-Jares, C. M., García-Martín, M. S., Marino, N., & Torrijos, C. (1995). GC–MS identification of volatile components of Galician (northwestern Spain) white wines. Application to differentiate Rias Baixas wines from wines produced in nearby geographical regions. *Journal of Science and Food Agriculture, 69*, 175–184.

García-Parrilla, M. C., González, G. A., Heredia, F. J., & Troncoso, A. M. (1997). Differentiation of wine vinegars based on phenolic composition. *Journal of Agriculture and Food Chemistry, 45*, 3487–3492.

Guth, H. (1997). Quantification and sensory studies of character impact odorants of different white wine varieties. *Journal of Agriculture and Food Chemistry, 45*, 3027–3032.

Kos, G., Lohninger, H., & Krska, R. (2003). Development of a method for the determination of fusarium fungi on corn using mid-infra-red spectroscopy with attenuated total reflection and chemometrics. *Analytical Chemistry, 75*, 1211–1217.

Legin, A., Rudnitskaya, A., Lvova, L., Vlasov, Y., Di Natale, C., & D'Amico, A. (2003). Evaluation of Italian wine by an electronic tongue: recognition, quantitative analysis and correlation with human sensory perception. *Analytica Chimica Acta, 484*, 33–44.

Lopez, R., Aznar, M., Cacho, J., & Ferreira, V. (2002). Determination of minor and trace volatile compounds in wine by solid-phase extraction and gas chromatography with mass spectrometric detection. *Journal of Chromatographia A, 966*, 167–177.

Martin, G. J., Guillou, C., & Martin, M. L. (1998). Natural factors of isotope fractionation and the characterization of wines. *Journal of Agriculture and Food Chemistry, 36*, 316–322.

Meléndez, M. E., Sánchez, M. S., Íñiguez, M., Sarabia, L. A., & Ortiz, M. C. (2001). Pyschophysical parameters of colour and the chemometric characterisation of wines of certified denomination of origin 'Rioja'. *Analytica Chimica Acta, 446*, 159–169.

Pérez-Pavón, J. L., Del Nogal-Sánchez, M., García-Pinto, C., Fernández Laespada, M. E., Moreno-Cordero, B., & Guerrero-Peña, A. (2003). A method for the detection of hydrocarbon pollution in soils by headspace mass spectrometry and pattern recognition techniques. *Analytical Chemistry, 75*, 2034–2041.

Reeves, J. B., & Zapf, C. M. (1998). Mid-infrared diffuse reflectance spectroscopy for discriminant analysis of food Ingredients. *Journal of Agriculture and Food Chemistry, 46*, 3614–3622.

Trullols, E., Ruisánchez, I., & Rius, F. X. (2004). Validation of qualitative analytical methods. *Trends in Analytical Chemistry, 23*, 137–145.

Tura, D., Prenzler, P. D., Bedgood, D. R., Antolovich, M., & Robards, K. (2004). Varietal and processing effects on the volatile profile of Australian olive oils. *Food Chemistry, 84*, 341–349.

Vandeginsten, B. G. M., Massart, D. L., Buydens, S., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1998). *Handbook of chemometrics and qualimetrics: parts A and B*. Amsterdam: Elsevier.

Wang, J. F., Geil, P. H., Kolling, D. R. J., & Padua, G. W. (2003). Analysis of zein by matrix-assisted desorption/ionization mass spectrometry. *Journal of Agricultural and Food Chemistry, 51*, 5849–5854.

Wold, S. (1976). *Pattern Recognition, 8*, 127–139.

Yu, H. L., & MacGregor, J. F. (2003). Multivariate image analysis and regression for prediction of coating content and distribution in the production of snack foods. *Chemometrics and Intelligent Laboratory Systems, 67*, 125–144.