ASSESSMENT OF MATURITY OF LOIRE VALLEY WINE GRAPES BY MID-INFRARED SPECTROSCOPY

D. PICQUE[1*], Pascale LIEBEN[1], Ph. CHRÉTIEN[2],
J. BÉGUIN[3] and L. GUÉRIN[3]

1: Unité Mixte de Recherche 782 Génie et Microbiologie des Procédés Alimentaires
(AgroParisTech-INRA), BP 1, 78850 Thiverval-Grignon, France
2 : Institut Français de la Vigne et du Vin (IFV, Anjou Unit), 42 rue Georges Morel, BP 60057
49070 Beaucouzé cedex, France
3 : Institut Français de la Vigne et du Vin
(IFV, Touraine Unit), 46 Avenue Gustave Eiffel, BP 39537, 37095 Tours cedex 2, France

Abstract

Aim : The objective of the present study was to assess the ripening of grapes collected at different stages of maturation between the "véraison" and harvest periods from mid-infrared spectroscopy (MIRS) analysis.

Methods and results : Grape berries of Cabernet Franc collected in two locations of the Loire Valley region (Touraine and Anjou), from 28 vine plots and during two vintages (2005 and 2006) were analysed. With principal component analysis (PCA) of spectral data of grape musts, different levels of ripening were described during the three to four weeks before harvest. A separation according to origin (Touraine or Anjou) was also observed and confirmed by the results of partial least squares (PLS) discriminant analysis (88 % of correct classification). Similar evolutions and geographical discriminations were obtained for specific physicochemical parameters. By PLS regression, good predictions of titratable acidity and sugar concentration from berry spectral data were obtained, with root mean square error of prediction (RMSEP) of 0.53 g/L for titratable acidity (expressed in H_2SO_4) and 8 g/L for sugar concentration. Moreover, when the data from only one of the regions were considered, the predictions of titratable acidity and sugar concentration were improved and those of real acidity (pH) and maturity index were then satisfactory. The RMSEP values for samples from Touraine and Anjou were reduced, respectively, to 0.05 and 0.02 units for pH, 0.4 and 0.12 g/L for titratable acidity (expressed in H_2SO_4), 6.6 and 3.2 g/L for sugar concentration, and 5 and 2.2 units for maturity index.

Conclusion : Spectroscopic and classic chemical analyses of grape berries yielded highly similar results. The evolution of berries from "véraison" to harvest can be characterized according to both time course and region. The samples showed similar PCA results for chemical and IR spectra parameters. PLS regression between chemical and spectral data showed that Fourier transform IR is a good method to predict acidity and sugar concentration throughout ripening. And the results for these parameters, as well as for pH, maturity index and anthocyanin concentration, are improved if the regressions are calculated from sample sets restricted to a single growing region. Consequently, a calibration model is required for each grape geographical origin.

Significance and impact of the study : The potential of MIRS was demonstrated for the quantification of the main indicators of maturity during berry ripening. Furthermore, these spectra can be used to estimate grape maturity in particular in reference to a spectral database established over several years of study. The association infrared spectroscopy, chemometric methods and database will help to monitor ripening and to determine the optimum harvest date.

Key words : ripening, infrared spectroscopy, chemical methods, correlations

Résumé

Objectifs : L'objectif de ce travail est d'évaluer la maturation de baies de raisin collectées entre la véraison et la récolte par spectroscopie moyen infra rouge (MIRS).

Méthodes et résultats : Les baies de raisin ont été collectées dans deux régions de la vallée de la Loire (Touraine et Anjou) sur 28 parcelles au cours de deux millésimes (2005 et 2006). À partir d'une analyse en composantes principales (ACP) des spectres MIR des moûts issus des baies de raisin, différents niveaux de maturation sont mis en évidence durant les trois à quatre semaines précédant la récolte. Une séparation selon l'origine géographique est également observée et est confirmée par les résultats de l'analyse discriminante par les moindres carrés partiels (PLS) avec 88% de réponses correctes. Une évolution en fonction du temps et une discrimination géographique des baies sont également obtenues à partir des données d'analyses physico-chimiques des moûts. À partir des données spectrales, de bonnes prédictions des concentrations en sucres et de l'acidité titrable sont obtenues par régression PLS. Cependant, quand les données sont traitées séparément en fonction de l'origine géographique des échantillons (Touraine et Anjou), les prédictions des teneurs en sucres et de l'acidité titrable sont améliorées et celles du pH et de l'indice de maturité sont alors satisfaisantes.

Conclusion : Les analyses des données spectrales MIR et physico-chimiques conduisent à des résultats similaires. Les baies de raisin peuvent être caractérisées entre la véraison et la récolte en fonction du degré d'avancement de la maturation et en fonction de l'origine géographique des échantillons. Les régressions PLS établies entre les données spectrales et physico-chimiques montrent que la spectroscopie moyen infrarouge est une bonne méthode pour prédire les évolutions des teneurs en sucre et de l'acidité titrable des baies au cours de la maturation. Ces résultats, ainsi que ceux de prédiction du pH, de l'indice de maturité et de la concentration en anthocyanes, sont améliorés si les régressions sont établies région par région. En conséquence, une calibration est recommandée pour chaque groupe de baies de raisin en fonction de l'origine géographique.

Impact de l'étude : Le potentiel du MIRS a été démontré pour la quantification des principaux indicateurs de maturité des baies de raisin. De plus, ces spectres peuvent être utilisés pour estimer le niveau de maturité des baies en particulier en référence à une base de données spectrales établie sur plusieurs millésimes. L'association MIRS, chimiométrie et base de données pourra aider à suivre la maturation et à déterminer la date optimale de récolte.

Mots clés : maturation, baies de raisin, spectroscopie moyen infra rouge, méthodes chimiques, corrélations

*Corresponding author : picque@grignon.inra.fr

INTRODUCTION

Reliable maturity indices and methods to objectively evaluate grape quality are sought in order to monitor ripening, to determine the optimum harvest date and to adapt winemaking practices to the desired wine styles. To assess grape maturity, various types of analyses are carried out : i) the concentration of total soluble solids (TSS, which is mainly a measure of sugars) and the acidity. This maturity is called the technological maturity. ii) the total phenolic compounds (TPC) and the total anthocyanins by spectrophotometric methods. This maturity is called the phenolic maturity. Nevertheless, these parameters are admitted to be insufficient to assess grape quality. Therefore, other analysis, such as aromatic profile by gas chromatography (Sanchez Palomo *et al.*, 2007), phenolic compounds by liquid chromatography (Pena-Neira *et al.*, 2004), and sensory analysis (Rousseau and Delteil 2003 ; Le Moigne *et al.*, 2008) are used and present a real potential. Unfortunately, these analyses are time consuming and are generally too expensive to be considered by winemakers.

There is clearly a need to develop new methods to evaluate ripening of berries. The measurement of mechanical properties of berries has made it possible to discriminate the « veraison » period, but no evolution of different rheological parameters was found between the "veraison" and the harvest periods (Robin *et al.*, 1997, Grotte *et al.*, 2001). More recently, it has been shown that the texture indices depended on grape variety and harvest season (Letaief *et al.*, 2008). Optical systems have been the object of particular attention these last years. Colour or reflectance measures were used by Celotti *et al.* (2005, 2007) to assess grape phenolic quality, and fluorimetry was established to estimate flavonol and sugar content in white grape berries (Kolb *et al.*, 2003, 2006). However, the most important work was dedicated to infrared spectroscopy. Infrared spectroscopy has many advantages because it is a high speed and easy to use analytical technique. It quantifies the energy absorbed by molecular bonds and provides spectral data, principally on acids, carbohydrates, alcohols and phenolic compounds. In recent years, combined with multivariate data analysis, infrared spectroscopy has been developed for rapid quantitative analysis of the main compounds in several products such as wine (Patz *et al.*, 2004 ; Moriera *et al.*, 2004). Consequently, several studies were devoted to the evaluation of important characteristics of wine quality such as tannins (Fernandez *et al.*, 2007), anthocyanins (Soriano *et al.*, 2007) or colour (Versari *et al.*, 2007).

Infrared spectroscopy for grape analysis has mostly focussed on near-infrared (NIR) measurements. For must samples collected during the harvest period, Manley *et al.* (2001) found a strong correlation between NIR spectroscopic data and TSS, but the predictions of free amino nitrogen, malic acid, lactic acid and ethyl carbamate were not satisfactory. Dambergs *et al.* (2003) and Gishen *et al.* (2005) confirmed that NIR spectroscopy to measure TSS, acidity and colour (expressed as total anthocyanins) can be considered as a good predictor of red wine composition and quality. Arana *et al.* (2003) showed that NIR reflectance technology is suitable to determine TSS content directly in the grape berries. Guggenbichler *et al.* (2006) established a fast quantitative analysis of carbohydrates, total acidity, tartaric acid, malic acid, polyphenol content and pH of different grapes varieties by NIR. These results were confirmed by Casiraghi *et al.* (2007) who analysed samples collected during the last period of ripening just before harvest. Local PLS regression (Dambergs *et al.*, 2006) and artificial neural networks (Janik *et al.*, 2007) have been proposed as means of improving calibration performance mainly for the prediction of anthocyanin concentrations.

Herrera *et al.* (2003) made a first evaluation of a NIR spectrometer equipped with an optical fibre probe directly applied onto the berries. The Brix degree was estimated with an error of 1.06. Using two types of infrared sensors working by transmission or diffuse reflectance respectively and allowing a direct measure on the berries or on the grape bunches, Desseigne *et al.* (2003) obtained a good correlation between infrared signals and TSS in grapes (coefficient of determination $R^2 = 0.90$).

Only a few studies have been devoted to the use of Fourier transform mid-infrared spectroscopy for the assessment and monitoring of grape quality. With an equipment working in near and mid infrared (5,000 cm^{-1} – 1,000 cm^{-1}), the measures of sugars, tartaric acid, malic acid, acetic acid, acidity, TPC and colour were validated according to the OIV (Office International de la Vigne et du Vin) methods for must samples (Dubernet *et al.*, 2001, Rousseau *et al.*, 2002). From a great number of South African grape musts analysed by mid-infrared spectroscopy, the standard errors of prediction were 0.38°Brix for TSS, 0.050 units for pH and 0.49 g/L for titratable acidity after adjustments of the global calibration algorithms (Swanepoel *et al.*, 2007).

The objective of the present study was to assess the ripening of grape berries between the « véraison » and harvest periods by mid-infrared spectroscopy. Grape berries of Cabernet Franc from 28 vine plots located in two regions (Touraine and Anjou, France) were collected during two vintages (2005 and 2006) in order to compose our sample. Musts from berries analysed by infrared spectroscopy and by classic chemical methods were discriminated by principal component analysis (PCA) and classified by partial least squares (PLS) analysis, according to ripening and to geographic origin.

Correlations between infrared spectra and classic chemical parameters were determined by PLS regression.

MATERIALS AND METHODS

1. Loire Valley : study sites

Berries of Cabernet Franc were collected during two vintages (2005 and 2006) from 28 parcels located in two different regions of the Loire Valley, France. The parcels were chosen to represent the pedo-climatic conditions met in the Loire Valley :

- 14 parcels in the Tours region (Touraine) : 5 in the Bourgueil appellation area, 4 in the Saint Nicolas de Bourgueil appellation area and 5 in the Chinon appellation area.

- 14 parcels in the Angers region (Anjou) : 10 in the Anjou appellation area and 4 in the Saumur appellation area.

2. Grapes

Grape sampling was carried out in each parcel over a 3 or 4-week period each year. A total of 211 samples was carried out during the study. Samples were collected at one-week interval and the last sampling date of each vintage corresponded to harvest. Two batches of two-hundred berries, with pedicels, were randomly picked per sample. Sampling was performed according to the ITV (French Technical Institute for Viticulture and Oenology) method (Cayla *et al.*, 2002) in order to limit the effects of grape heterogeneity.

3. Physicochemical analysis :

The first batch of two-hundred berries of each sample was crushed with a kitchen blender (Moulinex, Vivacio). The must was then filtered through a glass wool to make the following analyses. The analytical methods used are described in the collection of the international methods of analyses of wines and musts (OIV, 2009) : sugar concentration (g/L) was estimated with a refractometer with temperature correction (Fabre, Mesurelec, accuracy 3.3 g/L) ; titratable acidity (expressed in g/L eq. H_2SO_4) was determined by the titration method (accuracy 0.14 g/L in the range 4 - 4.5 g/L) and real acidity (pH) was determined with a pH meter (WTW, pH 196, accuracy 0.05). Maturity index was calculated by dividing the sugar concentration by the titratable acidity.

The second batch of two-hundred berries was also crushed with a blender, and maceration was performed to determine the total phenolic compounds (TPC, accuracy 0.8 g/kg) and the anthocyanin concentration (accuracy 0.06 g/kg) following the ITV France method (Cayla *et*

al., 2002). These analyses and accuracy estimations were performed at the Laboratory of Touraine (Tours, France).

4. Fourier transform infrared spectroscopy analysis

Fourier Transform Mid-Infrared Spectroscopy (FT-MIR) analyses of must samples were carried out in transmission mode. The spectra of 1,060 data points in the range between 5,012 cm^{-1} and 926 cm^{-1} were collected using a FOSS WineScan system. Before statistical analysis, the first derivatives of spectra were calculated using the Savitzky-Golay method (4 data points for each side). This treatment was performed to take into account the eventual drift of spectrometric signal.

5. Statistical analysis

Statistical analyses were carried out using STATISTICA software (StatSoft, Paris, France). The Principal Component Analysis (PCA) was applied to the infrared data in order to evaluate their evolution throughout ripening. PCA helps to obtain an overview of the information in the data set by replacing the original variables by a few new variables called Principal Components (PC). PC contain almost all the information and are orthogonal among themselves.

The Partial Least Squares (PLS) regression was used to establish regression models between the physicochemical and the spectral data. PLS is a data compression technique that provides good relationships in case of collinearity, as in spectral data, maximizing the covariance between successive PLS scores and the dummy variables. One of the most important issues in PLS regression is to determine the size of the model to describe the significant source of variation, but without over-fitting.

Therefore, the samples, with their physicochemical values and spectral measurements, were split in two sets. The first set (168 samples), designated as the "calibration set", was used for the establishment of the PLS statistical models. The establishment of the PLS model consisted of modelling and validation procedures by a leave-one-out cross-validation method. The second set (43 samples), designated as the « prediction set », was randomly selected to validate the calibration. PLS regression was also applied to Tours and Angers data separately. In that case, the number of samples by set is shown in table 3.

The number of factors was determined by using the minimum Prediction Residual Error Sum of Squares (PRESS) value.

The accuracy of the regression models was expressed as the coefficient of determination (R^2). According to the R^2 value, the model provides good predictions if R^2 is

higher than 0.81 and approximate quantitative predictions if R^2 is between 0.66 and 0.81 (Karoui *et al.*, 2006).

To compare between the results obtained from the physicochemical and MIRS methods, the values of the root mean square error of cross-validation (RMSECV) and prediction (RMSEP) for calibration and prediction set, respectively, were considered (Bertrand, 2006).

The Ratio of Prediction to Deviation (RPD) was used as a broad indicator of the performance of the calibration models. RPD was calculated by dividing the Standard Deviation (SD) of the data set by the RMSEP. It has been proposed that a ratio higher than 2 indicates a good calibration (Mouazen *et al.*, 2005 ; Karoui *et al.*, 2006).

Classifications according to ripening (coded H-1, H-2, H-3, H-4 according to the week of sampling before harvest) or origin (Touraine or Anjou, coded 1 or 2, respectively) were carried out by partial least squares discriminant analysis (PLS-DA). As previously described, the samples were divided in two sets. The number of samples per set is shown in table 2 and 3.

RESULTS

1. Infrared spectra of grape must

Figure 1 shows the mean and standard deviation values of absorbance calculated from the 221 samples analysed during the study. The noisy and non informative parts of the spectra were excluded from our analyses : the regions from 1,582 cm^{-1} to 1,692 cm^{-1} and from 2,971 cm^{-1} to 3,627 cm^{-1} contained strong water absorption bands ; the regions from 1,700 cm^{-1} to 2,700 cm^{-1} and above 3,700 cm^{-1} were eliminated because the variations of absorption, characterized by the standard deviation, were very low ; and the region from 2,700 cm^{-1} to 3,000 cm^{-1}, assigned to methylene and methyl vibrations (Coates, 2000), was eliminated because it contained very little useful information for the characterization of wine grapes. The wavelength region used for all data manipulation was 950 cm^{-1} to 1,550 cm^{-1}, corresponding to the region called fingerprint. This region provided the most useful information (Versari *et al.*, 2006).

Figure 2 shows the infrared spectra of berry juice from four weeks before harvest (coded H-4) to harvest (coded H). The spectra show changes of absorption at different frequencies that could be attributed to functional groups. The region between 1,000 cm^{-1} and 1,250 cm^{-1} corresponds to the frequency of C-O stretching mode and more precisely to the antisymmetric stretching of C-C-O group. Its frequencies of vibration make it possible to differentiate the primary (1,075 cm^{-1} – 1,000 cm^{-1}) and secondary (1,150 cm^{-1} – 1,075 cm^{-1}) hydroxyls functions

(Lin-Vien *et al.*, 1991). These regions are found in carbohydrates and the increase of absorption between H-4 and H can be attributed to the increase of sugar concentration.

Between 1,200 cm^{-1} and 1,550 cm^{-1}, the spectra present several weak modifications of absorption. According to Coates (2000), the aromatic ring stretch bands and the aromatic combination bands are located in this region. The aromatic ring absorption C = C-C is found around 1,450 cm^{-1} and 1,510 cm^{-1} (Fernandez *et al.*, 2007). A strong contribution of OH deformation vibration can be found in the region from 1260 cm^{-1} to 1,410 cm^{-1}, and a C-O valence vibration can be found between 1,040 cm^{-1} and 1,150 cm^{-1} overlapping with the aromatic fingerprint bands between 950 cm^{-1} and 1,225 cm^{-1} (Edelmann *et al.*, 2001).

2. Discrimination of must samples analysed by infrared spectroscopy

PCA was applied to the spectra obtained from must in order to visualize the occurrence of grouping among the samples measured at different levels of ripening. The PCA results indicate that the first two principal components (PC1 and PC2) explain 97 % of the total variance with 96 % attributed to PC1 (figure 3). For a better reading, the PCA factorial map defined by PC1 and PC2 is presented in separate maps for Touraine (figure 3A) and Anjou samples (figure 3B), but the new coordinates of the spectra were calculated from the same PCA. A continuous evolution from samples collected 4 or 3 weeks before harvest to samples collected at maturity is visible according to axis 1. For the same time period before harvest, the position of the sample spectra varied according to the parcel and the vintage. The most important wavenumbers associated to PC1 were found in the region from 1,000 cm^{-1} to 1,120 cm^{-1}, assigned to carbohydrates.

Table 1 shows the PLS-DA results applied to predict the ripening for all the samples from Tours and Angers. Poor classifications were achieved for the three samples sets. For instance, the percentages of correctly classified samples were included between 38 and 75 % for the calibration set. Similar classification levels were achieved for prediction sets. A discriminant analysis applied separately to the data from Tours and Angers did not improve the level of the results.

On the PCA scores scatter plot, a difference between the position of Touraine and Anjou samples can also be noted according to axis 2 : the first set of samples (Tours) is principally located at the bottom of the PCA map, whereas the second one (Angers) is at the top. This separation according to origin was confirmed by the PLS-DA results. The models correctly predicted more than 90 % of the calibration samples (table 2). For the prediction

set, the level of correct classifications was similar (higher than 87 %) with only 5 misclassified samples.

3. Discrimination of must samples analysed by chemical methods

In order to evaluate the performance of infrared spectroscopy, the same types of discriminant analysis were applied to the chemical data usually measured in grape must. For PCA, the first three principal components explained 90 % of the total variance, with 52 %, 25 %, and 13 % for PC1, PC2 and PC3, respectively. The PCA factorial maps (figure 4) show an evolution of sample positions depending on the ripening time according to axis 1, as described above for PCA applied to spectral data. For a better reading, the PCA factorial map is presented in separate maps for Touraine (figure 4A) and Anjou samples (figure 4B), but the new coordinates of the spectra were calculated from the same PCA. The factor loadings indicate that maturity index and sugar concentration are positively correlated ($R^2 > 0.85$) and acidity is negatively correlated ($R^2 > 0.88$) with PC1. As previously described for the mid-infrared spectra, the position of the samples depended on the parcel and the year.

PLS-DA was performed to classify grapes according to ripening time using the chemical data (table 1). The percentages of correctly classified samples were very weak : between 17 and 55 % for the calibration set and between 20 and 80 % for the prediction set.

The factorial maps, formed by PC1 and PC2, clearly show a separation between the samples from Tours and Angers according to axis 2. The total phenolic compounds (TPC) variable shows a high level of correlation ($R^2 = 0.85$) with this axis. Figure 5 confirms the difference of TPC level between the two studied regions. At each step of ripening and for both years of the study, the concentrations of TPC were higher in the Angers area than in the Tours area. An average difference of 30 % is noted between the samples of the regions.

The PLS discrimination results according to the geographic origin were good, with 92 and 84 % of correctly identified samples, meaning that six and seven

Table 1 - Percentage of correct classification of must samples in calibration and prediction sets based on the Partial Least Squares discriminant analysis of spectral and chemical data according to the ripening of the berries.

Analysis	Sets	Number of samples	% of correct classification				
			H	-1	-2	-3	-4
Infrared data	calibration	168 (97T, 71A)	51	75	46	49	38
	prediction	43 (28T, 15A)	70	50	100	57	40
Chemical data	calibration	168 (97T, 71A)	35	55	37	50	17
	prediction	43 (28T, 15A)	40	50	80	44	20

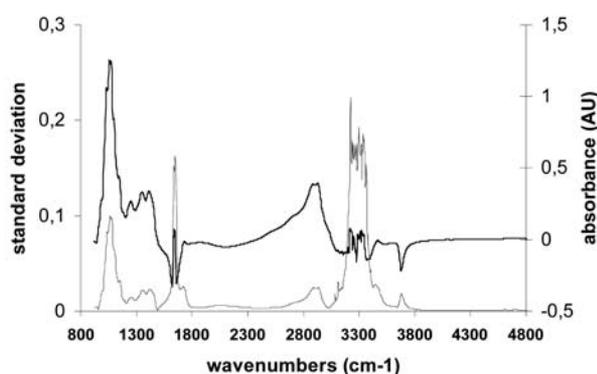T = Tours samples, A = Angers samples, H = harvest time, -1 to -4 = weeks before the harvest.

Figure 1 - Mean values of infrared absorbance (—) and standard deviation (—) calculated from the 221 grape samples analysed during the two-year study.
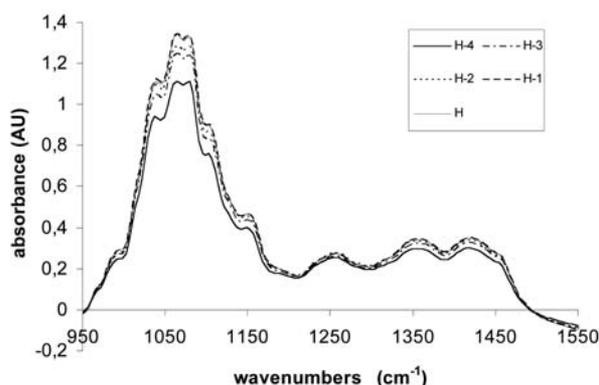
Figure 2 - Fourier Transform Mid-Infrared spectra of musts obtained from berries collected at harvest (H) and one to four weeks before harvest (H-1 to H-4).
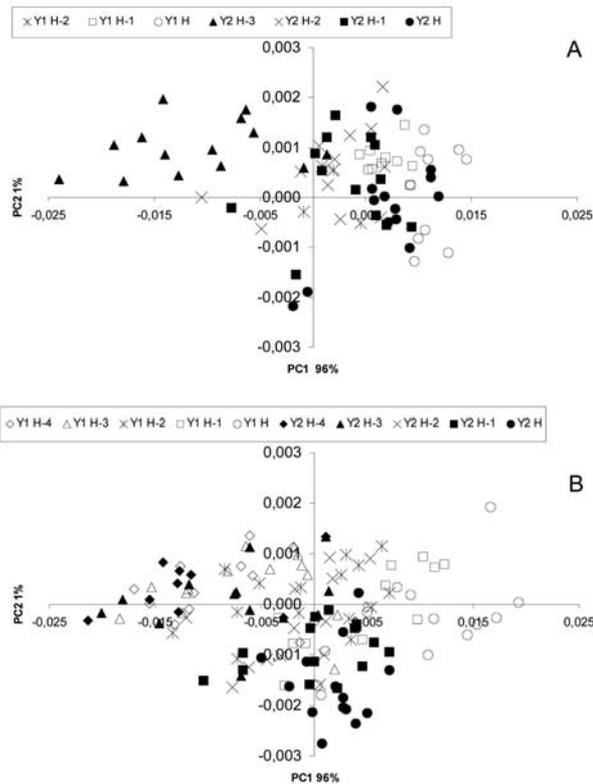
Figure 3 - Principal Components Analysis scores scatter plot (PC1 vs PC2) of the Fourier Transform Mid-Infrared spectra of must samples of berries collected for two seasons (Y1 = 2005, Y2 = 2006) during ripening. H-4 to H-1 = weeks before harvest, H = harvest, A = Anjou, B = Touraine.
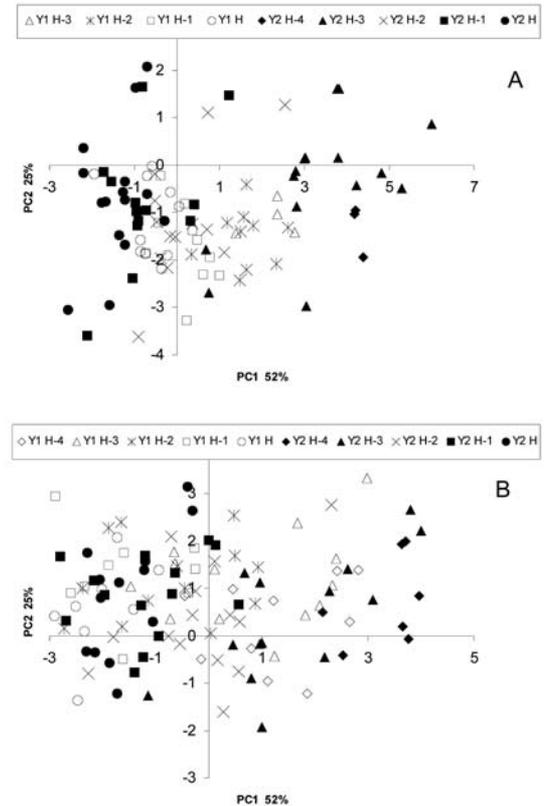
Figure 4 - Principal Components Analysis scores scatter plot (PC1 vs PC2) of the chemical data of must samples of berries collected for two seasons (Y1 = 2005, Y2 = 2006) during ripening. H-4 to H-1 = weeks before harvest, H = harvest, A = Anjou, B = Touraine.

samples were misclassified in the calibration and prediction sets, respectively (table 2).

4. Correlation between infrared spectra and physicochemical analysis

PLS regression was applied in order to evaluate the relationship between the chemical values and the infrared spectral responses throughout ripening (table 3). The evaluation of the performance of the PLS models was mainly based on the coefficient of determination (R^2), the Root Mean Square Error of Cross-Validation (RMSECV) and Prediction (RMSEP). Three data sets were studied : one with the spectra of all samples (total data set), a second with the spectra of the Tours samples, and a third with the spectra of the Angers samples.

For the global data set (table 3), good correlations between the chemical values and the infrared spectra were only obtained for acidity and sugar concentration, with a Ratio of Prediction to Deviation (RPD) above 2. For pH and maturity index, the R^2 values were quite good (higher than 0.9), but those of RPD were lower than 2. For berry weight, TPC and anthocyanins, the performances of PLS

models were not satisfactory. TPC and anthocyanin concentration were very low and no significant variations were observed in time, except for the anthocyanins of the 2006 vintage, three and four weeks before the harvest (figures 5 and 6).

For Tours and Angers sets, the results are improved with the best results for Angers set. The predictions of sugar concentration and maturity index are very satisfying with a RPD included between 2.8 and 9.2. The predictions of pH and acidity are improved, if we considered RMSEP and RPD values. For example, the RMSEP of pH are 0.05 and 0.02 for Touraine and Anjou data respectively. The prediction of anthocyanins of Angers samples is slightly less satisfactory but can be considered as correct (RMSEP = 0.08, RPD = 2.1). However, the difference between the RMSECV and the RMSEP obtained respectively for the calibration and prediction sets is not high and allows us to think that the models are reasonably robust.

Results of PLS calibration and prediction for Angers samples are depicted in figure 7 with the regression curves. We can observe that measured and predicted values are quite close for the two data sets (calibration and
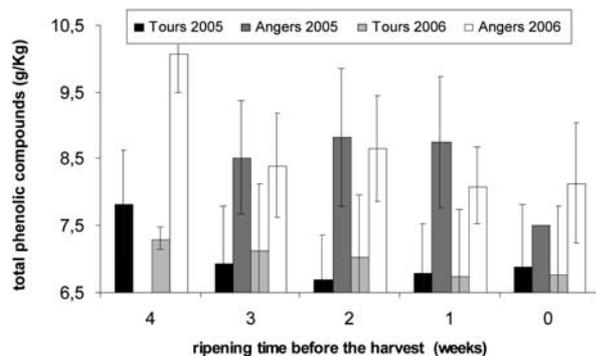
Figure 5 - Evolution of the mean values of total phenolic compounds in grape samples from Tours (14 parcels) and Angers areas (14 parcels) during the four weeks before harvest of the 2005 and 2006 vintages.
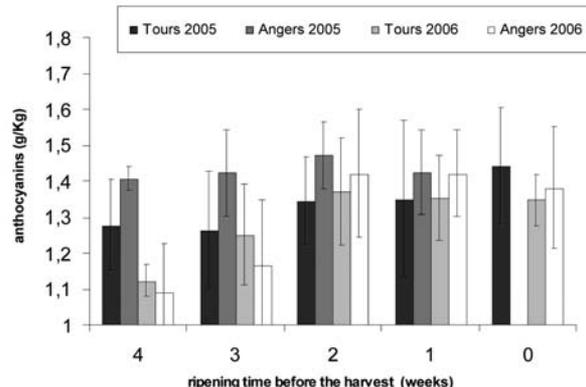

Figure 6 - Evolution of the mean values of anthocyanins in grape samples from Tours (14 parcels) and Angers areas (14 parcels) during the four weeks before harvest for the 2005 and 2006 vintages.

Table 2 - Percentage of correct classification of must samples in calibration and prediction sets based on the Partial Least Squares discriminant analysis of spectral and chemical data according to the geographical origin of the berries.

Analysis	Sets	Number of samples	% of correct classification (CC)		
			Touraine	Anjou	mean CC
Infrared data	calibration	168 (97T, 71A)	94	91	98
	prediction	43 (28T, 15A)	89	87	88
Chemical data	calibration	168 (97T, 71A)	94	88	92
	prediction	43 (28T, 15A)	86	80	84

T = Tours samples, A = Angers samples.

prediction), showing good accuracy and robustness of the models for pH, acidity, sugar concentration and maturity index. For anthocyanin and TPC predictions, the linear regression models are less accurate and a curvature is observed in the scatter plot.

DISCUSSION

Infrared spectroscopy appears to be a very attractive method to characterize red grape berries. The PCA of the infrared spectra of must extracted from berries clearly shows the evolutions from "véraison" to harvest. The most important wavenumbers associated with these evolutions are attributed to the carbohydrates. Discrimination of the geographical origin is also possible between the two studied areas (Touraine and Anjou). The PCA performed with the chemical data shows very similar results with those performed with the infrared spectra, with an evolution depending on the ripening time, linked to sugar concentrations and acidity levels, and with discrimination between the two regions. This

discrimination between Tours and Angers, mainly attributed to total phenolic compounds by PCA on the chemical data, can be explained by the soil characteristics of the different parcels. These differences in soil structure can induce a modification of temperature and water availability in soils and also different vine development kinetics. Some researchers have already described this in previous works, and notably Brossaud *et al.* (1999).

PLS-DA applied spectral and chemical data confirmed the PCA conclusions. The samples collected during the same week present differences of coordinates in the map and do not form homogeneous groups. The results of PLS-DA discrimination are poor, ranging from 20 to 100 % of correct classification. This is not surprising because, for each parcel, different maturity levels can be recorded depending of the location.

A high level of geographical discrimination was reached with 88 and 84 % of well classified prediction samples for infrared and chemical data, respectively.
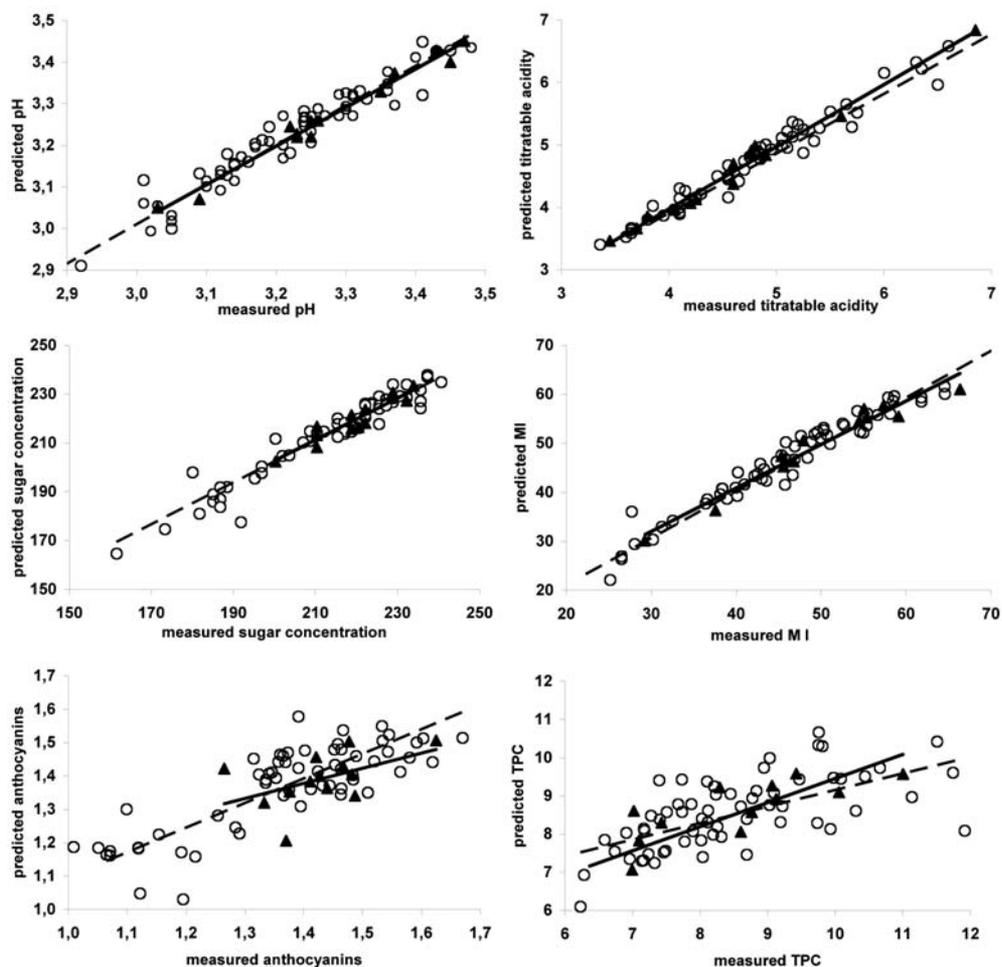
Figure 7 - Predicted versus measured values from Partial Least Squares regression models for pH, titratable acidity (g eq. H_2SO_4/L), sugar concentration (g/L), maturity index (MI), anthocyanins (g/Kg) and total phenolic compounds (TPC, g/Kg) for Angers grape samples.

○, — — — **data and regression curve for the calibration set;**

▲, —— **data and regression curve for the prediction set.**

Three of the five misclassified samples by infrared spectroscopy were also misclassified by chemical data.

The similarities of the results between spectral and chemical analyses suggest that the spectral data can be correlated with the chemical values in order to predict their evolutions during ripening. For the whole data set (two geographical origins and two vintages), the correlations are only satisfying for acidity and sugar predictions (RMSEP = 0.53 g/L and 8 g/L respectively). Previously published results, reviewed by Cozzolino *et al.*, 2006, showed RMSEP between 1° and 1.6° Brix for TSS measured by near infrared reflectance. More recently, by using visible/near infrared spectroscopy, the RMSEP was reduced to 0.75° (Casiraghi *et al.*, 2007) and 0.27°Brix

(Dambergs *et al.*, 2006). Using mid-infrared spectroscopy as in our study, the SEP values were 0.38°Brix for TSS and 0.49 g/L for titratable acidity (Swanepoel *et al.*, 2007). Our results (Angers and Tours) are not as satisfying as those obtained in this last study. But if we consider separately the predictions of each geographical region, the results are of the same order: RMSEP = 6.3 and 3.2 g/L for sugars (corresponding approximately to 0.58 and 0.30°Brix) and 0.4 and 0.12 g/L for titratable acidity in the Tours and Angers regions, respectively.

In fact, if the data sets are divided according to the geographical origin of the samples, a reduction of the prediction error is also observed for real acidity (pH), maturity index and anthocyanin concentration. For

Table 3 - Partial Least Squares regression results using the physicochemical values and Fourier Transform Mid-Infrared spectra of berries (n=number of samples).

	Weight of 200 berries (g)	pH	Acidity g eq. H_2SO_4/L	Sugars g/L	MI[a]	TPC[b] g/Kg	Anthocyanins g/Kg
Mean values	224	3.33	4.41	211	49.5	7.65	1.34
Standard deviation	30	0.16	1.1	17.8	13.1	1.02	0.17
Tours + Angers samples							
latent variables	4	11	6	1	10	7	2
R²	0.29	0.9	0.86	0.85	0.92	0.51	0.43
RMSECV (n=168)	23.5	0.09	0.54	7.8	9.0	0.8	0.11
RMSEP (n=43)	24.5	0.09	0.53	8	9.6	0.7	0.11
RPD	1.3	1.8	2.1	2.2	1.4	1.5	1.5
Tours samples							
latent variables	4	5	6	3	8	5	1
R²	0.31	0.84	0.92	0.85	0.92	0.25	0.17
RMSECV (n=97)	24	0.05	0.3	6.6	4.4	0.8	0.12
RMSEP (n=28)	22	0.05	0.4	6.3	5	0.8	0.12
RPD	1.4	3.2	2.8	2.8	2.6	1.3	1.4
Angers samples							
latent variables	6	3	6	2	4	4	3
R²	0.52	0.95	0.98	0.94	0.95	0.43	0.73
RMSECV (n=71)	20	0.03	0.15	3.3	2.3	0.61	0.08
RMSEP (n=15)	18	0.02	0.12	3.2	2.2	0.63	0.08
RPD	1.7	8.0	9.2	5.6	6.0	1.6	2.1

a: maturity index, b: Total Phenolic Compounds

example, the RMSEP values reported for pH decreased from 0.09 for the total sample set to 0.05 and 0.02 for the Tours and Angers sample sets, respectively. These values are similar to those published by Swanepoel *et al.* (2007) who concluded that it is important to use sample sets representative of geographical origins.

As reported by Cozzolino *et al.* (2006), the RMSEP values for anthocyanins (0.05 to 0.18 mg/g) increase with diverse sample sets in comparison with sample sets restricted to a single vine growing region. Our results show the same evolution with a decrease of 0.11 g/kg for the total sample set to 0.08 g/kg for the Angers sample set. However, the prediction of TPC and anthocyanins is not satisfactory. The small variations of concentrations, the sensitivity limit of spectroscopy and the performances of the PLS regression can partly explain these results. Still, the precision of the prediction, in particular for TPC, is close to values obtained from the spectrophotometric method (0.8 g/kg). To take into account the inaccuracy of the PLS regression due to high regression curvature or bias, an artificial neural network was proposed to improve the prediction of total anthocyanin concentration (Janik *et al.*, 2007).

The number of PLS factors incorporated into our models, between one and eight, is reasonable. It is the same for the values of RMSECV and RMSEP obtained respectively for the calibration and prediction set which are very close (Table 3). Therefore, the models are reasonably robust.

CONCLUSIONS

The evolution of berries from "véraison" to harvest can be characterized from their infrared spectra according to time course. Discrimination between samples from both studied regions was also observed. IR spectroscopic and classic chemical analyses of grape berries showed high similarities. PCA of chemical parameters showed similar results to those obtained for infrared spectra according to both time course and region.

PLS regression between chemical and spectral data showed that FTIR is a good method to predict acidity and sugar concentration throughout ripening. But the results are improved for these parameters and for the prediction of pH, maturity index and anthocyanin concentration if the regressions are established from sample sets restricted to a single growing region. Consequently, a calibration model is required for each geographical origin. Further

work with a more important number of samples and harvest seasons will be necessary to confirm these results.

Mid-infrared spectroscopy is a high speed, easy to use analytical technique that allows simultaneous determination of several chemical parameters. For oenological laboratories, it would be possible to provide quick answers to predict the maturity level and to establish a harvest date presenting the most optimal parameters.

BIBLIOGRAPHICAL REFERENCES

Arana I., Jaren C. and Arazuri S., 2005. Maturity, variety and origin determination in white grapes (*Vitis Vinifera* L.) using near infrared reflectance technology. *J. Near Infrared Spectrosc.* **13**, 349-357.

Bertrand D., 2006. Une démarche générale pour l'établissement d'applications analytiques. *In: Tec & Doc (Ed.), La spectroscopie infrarouge et ses applications analytiques*, pp. 459-467. Lavoisier Publishers, Paris, France.

Brossaud F., Cheynier V., Asselin C. and Moutounet M., 1999. Flavonoid compositional differences of grapes among site test plantings of cabernet franc. *Am. J. Enol. Vitic.*, **50** (3), 277-284.

Cayla L., Cottereau P. and Renard R., 2002. Estimation de la maturité polyphénolique des raisins rouges par la méthode ITV Standard. *Rev. Fr. Œnol.*, **193**, 10-16.

Casiraghi E., Sinelli N., Bodria L., Guidetti R., Beghi R. and Cabassi G., 2007. Evaluation of grape quality parameters by VIS/NIR and FT-NIR spectroscopy. *Ital. Food Beverage Technol.*, **50**, 5-10.

Celotti E. and Carcereri De Prati G., 2005. The phenolic quality of red grapes at delivery: Objective evaluation with colour measurements. *S. Afric. J. Enol. Vitic.*, **26** (2), 75-82.

Celotti E., Della Vedova T., Ferrarini R. and Martinand S., 2007. The use of reflectance for monitoring phenolic maturity curves in red grapes. *Ital. J. Food Sci.*, **19**, 9-100.

Coates J., 2000. Interpretation of infrared spectra, a practical approach. *In Encyclopaedia of analytical chemistry*. R. A. Meyers (Ed.), pp. 10815-10837. John Wiley & Sons Ltd, Chichester.

Cozzolino D., Dambergs R.G., Janik L., Cynker W.U. and Gishen M., 2006. Analysis of grapes and wine by near infrared spectroscopy. *J. Near Infrared Spectrosc.*, **14**, 279-289.

Dambergs R.G., Cozzolino D., Esler M.B., Cynkar W.U., Kambouris A., Francis I.L., Hoj I.L. and Gishen M., 2003. The use of near infrared spectroscopy for grape quality measurement. *Aust. Grapegrower Winemaker*, 473A, 69-76.

Dambergs R.G., Cozzolino D., Cynkar W.U., Janik L. and Gishen M., 2006. The determination of red grape quality parameters using LOCAL algorithm. *J. Near Infrared Spectrosc.*, **14**, 71-79.

Desseigne J.M., Payan J.C., Crochon M., Roger J.M., Ballester J.F., Boulet J.C., Mazollier J. and Toussaint C., 2003. *Spectroscopie proche infrarouge et appréciation de la qualité de la vendange*, pp 168-172, EUROVITI, Montpellier.

Dubernet M., Dubernet V., Coulomb S., Lerch M. and Traineau I., 2001. Analyse objective de la qualité des vendanges par spectrométrie infra-rouge à transformé de Fourier et réseaux de neurones. *Bull. OIV*, **74**, 16-24.

Edelman A., Diewok J., Schuster K.C. and Lendl B., 2001. Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts. *J. Agric. Food Chem.*, **49**, 1139-1145.

Fernandez K., Labarca X., Bordeu E., Guesalagua A. and Agosin E., 2007. Comparative study of wine tannin classification using Fourier Transform mid-infrared spectroscopy and sensory analysis. *Applied Spectroscopy*, **61** (11), 1163-1167.

Gishen M., Dambergs R.G. and Cozzolino D., 2005. Grape and wine analysis enhancing the power of spectroscopy with chemometrics. A review of some applications in the Australian wine industry. *Aust. J. Grape Wine Res.*, **11**, 296-305.

Grotte M., Cadot Y., Poussier A., Loonis D., Piétri E., Duprat F. and Barbeau G., 2001. Détermination du degré de maturité des baies de raisin par mesures physiques: aspects rhéologiques. *J. Int. Sci. Vigne Vin*, **35** (2), 87-98.

Guggenbichler W., Huck C.W., Kobler A., Popp M. and Bonn G.K., 2006. Near infrared spectroscopy, cluster and multivariate analysis – contribution to wine analysis. *J. Food Agric. Env.*, **4** (2), 98-106.

Herrera J., Guesalaga A. and Agosin E., 2003. Shortwave-near infrared spectroscopy for non-destructive determination of maturity of wine grapes. *Meas. Sci. Technol.*, **14**, 689-697.

Janik L.J., Cozzolini D., Dambergs R., Cynkar W. and Gishen M., 2007. The prediction of total anthocyanin concentration in red-grape homogenates using visible-near-infrared spectroscopy and artificial neural networks. *Anal. Chim. Acta*, **594**, 107-118.

Karoui R., Mouazen A., Dufour E., Schoonheydt R. and Baerdemaeker J., 2006. A comparison and joint use of VIS-NIR and MIR spectroscopic methods for the determination of some chemical parameters in soft cheeses at external and central zones: A preliminary study. *Eur. Food Res. Technol.* **223**, 363-371.

Kolb C.A., Kopecky J., Riederer M. and Pfundel E.E., 2003. UV screening by phenolics in berries of grapevines (*Vitis vinifera*). *Functional Plant. Biology*, **30**, 1177-1186.

Kolb C.A., Wirth E., Kaiser W.M., Meister A., Riederer M. and Pfundel E.E., 2006. Noninvasive evaluation of the degree of ripeness in grape berries (*Vitis vinifera* L cv Bacchus and Silvaner) by chlorophyll fluorescence. *J. Agric. Food Chem.* **54**, 299-305.

Lemoigne M., Symonaux M. and Jourjon F., 2008. How to follow grape maturity for wine professionals with a seasonal judge training? *Food Quality and Preference*, **19** (8) 672-681.

Letaief H., Rolle L. and Gerbi V., 2008. Mechanical behavior of winegrapes under compression tests. *Am. J. Enol. Vitic.*, **59** (3), 323-329.

Lin-Vien D., Colthup N.B., Fateley W.G. and Grasselli J.G., 1991. *The handbook of infrared and Raman characteristic frequencies of organic molecules*, Academic Press, San Diego.

Manley M., Van Zyl A. and Wolf E.E.H., 2001. The evaluation of the applicability of Fourier transform near-infrared spectroscopy in the measurement of analytical parameters in must and wine. *S. Afric. J. Enol. Vitic.*, **22** (2), 93-100.

Moriera J.L. and Santos L., 2004. Spectroscopic interferences in Fourier transform infrared wine analysis. *Anal. Chem. Acta* **513**, 263-268.

Mouazen A., De Baerdemaeker M. J. and Ramon H., 2005. Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil Tillage Res.*, **80**, 171-183.

Organisation Internationale de la Vigne et du Vin,, 2009, *Recueil des méthodes internationales d'analyses des vins et des moûts,* volume I, O.I.V., Paris, France.

Patz C.D., Blieke A., Ristow R. and Dietrich H., 2004. Application of FT-MIR spectrometry in wine analysis. *Anal. Chem. Acta*, **513**, 81-89.

Pena-Neira A., Duenas M., Duarte A., Hernandez T., Estrella I. and Loyaola E., 2004. Effects of ripening stages of plant vegetative vigor on the phenolic composition of grapes (*Vitis vinifera* L.) cv. Cabernet-sauvignon in the Maipo Valley (Chile). *Vitis*, **43** (2), 51-57.

Robin J.P., Abbal P. and Salmon J.M., 1997. Fermeté et maturation du raisin. Définition et évolution de différents paramètres rhéologiques au cours de la maturité. *J. Int. Sci. Vigne Vin*, **31** (3), 127-138.

Rousseau J., Samirant M. and Granes D., 2002. Evaluation du fonctionnement d'un interféromètre à transformé de Fourier pendant les vendanges. *Rev. Fr. Œnol.*, **195**, 12-19.

Rousseau J. and Delteil D., 2003. Effets de différents facteurs sur le profil sensoriel des baies de raisin du cépage Chardonnay. *Bull. OIV*, **76** (873 – 874), 913-925.

Sanchez Palomo E., Diaz-Maroto M.C., Gonzales Vinas M.A., Soriano-Pérez A. and Pérez-Coello M.S., 2007. Aroma profile of wines from Albilllo and Muscat grape varieties at different stages of ripening. *Food Control*, **18**, 398-403.

Soriano A., Perez-Juan P.M., Vicario A., Gonzales J.M. and Perez-Coello M.S., 2007. Determination of anthocyanins in red wine using a newly developed method based on Fourier Transform infrared spectroscopy. *Food Chem.*, **104** (3), 1295-1303.

Swanepoel M., du Toit M. and Nieuwoudt H.H., 2007. Optimisation of the quantification of total soluble solids, pH and titratable acidity in South African grape must using Fourier transform mid-infrared spectroscopy. *S. Afr. J. Enol. Vitic.*, **28** (2), 140-149.

Versari A., Boulton R.B. and Parpinello G.P., 2007. Effect of spectral pre-processing methods on the evaluation of the color components of red wines using Fourier-Transform Infrared spectrometry. *Ital. J. Food Sci.*, **4** (18), 423-432.