

Near infrared spectroscopy, cluster and multivariate analysis – contributions to wine analysis

Wolfgang Guggenbichler¹, Christian W. Huck^{1*}, Armin Kobler², Michael Popp³ and Günther K. Bonn¹

¹Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University, Innsrain 52a, 6020-Innsbruck, Austria. ²Land- und Forstwirtschaftliches Versuchszentrum Laimburg, Pfatten, 39040-Auer, Italy. ³Bionorica AG - The Phytoneering Company, Kerschensteinerstrasse 11-15, 92308-Neumarkt, Oberpfalz, Germany. *e-mail: Christian.W.Huck@uibk.ac.at

Received 20 December 2005, accepted 22 March 2006.

Abstract

In this work we investigated different vine grapes (genus *Vitis*), grape varieties and bottled red wines using fiber-optics based near infrared transmittance and transmission spectroscopy. Recording of spectra in the transmission mode allowed distinguishing between different vine grapes. 57 bottles of *Cabernet Sauvignon*, *Lagrein* and *Chianti Classico* were analysed. Principal component analysis (PCA) over a wavenumber range from 4,500 to 10,000 cm⁻¹ including data pretreatment allowed assigning each wine to a separate cluster with a Q-value of 0.72. Partial least square regression (PLSR) additionally allowed determining the mixing ratio of grape varieties in one bottle of red wine. Therefore this model can help the consumer check the wine for its quality and thus prevent him from being misled. In order to achieve a fast and simple quantitative analysis in the transmittance mode of the carbohydrate, total acid, tartaric acid, malic acid, polyphenol content and pH value in Weißburgunder, Chardonnay, Ruländer, Silvaner, Müller Thurgau, Gewürztraminer, Sauvignon, Lagrein, Grossvernatsch, Blauburgunder, Cabernet and Merlot grapes a NIRS method was established.

Key words: Wine, near infrared spectroscopy, cluster analysis, multivariate analysis.

Introduction

Since the discovery of grape vines in Europe, Japan and Northern America about 100 million years ago¹ and the first production of wine in Asia about 8000 B.C., the quality of wine has always been a demanding task for mankind. Today, winegrowers and analysts mainly strive to protect the consumer from misleading information and wrong product names. A differentiation between different types of wines is normally carried out by sensory analysis. Tasting is subject to a variety of factors which can distort final results. Therefore, several analytical techniques for the identification have been developed.

Volatile ingredients can be analyzed by headspace-micro solid phase extraction (HMSPE) coupled to gas chromatography (GC)²⁻⁶. This technique allows identifying different grape sorts⁵. For the identification of the grape cultivation area, inductively coupled plasma mass spectrometry (ICP-MS)^{5, 7}, atomic absorption spectroscopy (AAS)⁸ and isotope dilution analysis in combination with AAS⁹ can be used. These methods use element pattern recognition in order to establish possible correlations between soil and geographical region, respectively. Alternatively, pyrolysis-mass spectrometry (MS) in combination with chemometric data analysis allows to identify the geographic origin of a wine¹⁰. Amino acid composition of a grape is not only typical of a certain vine variety, but also of climatic conditions and fertilization¹¹. Analysis of the amino acid pattern by ion exchange chromatography (IEC) or reversed-phase (RP) high-performance liquid chromatography (HPLC)¹¹, allows to classify wines according to their vine composition and geographical origin. Analysis and chemometric evaluation of polyphenols, mainly flavonoids and anthocyanins, by HPLC enable us to

differentiate between different grape mixing ratios¹²⁻¹⁴. Furthermore, with the help of mid infrared spectroscopy (MIRS) we are able to distinguish wines based on their geographical origin¹⁵, vines composition¹⁶ or age group¹⁷. A rapid method for the identification of the geographical origin and the age group of a wine without time-consuming sample preparation by UV-VIS spectroscopy was described by Heredia *et al.*¹⁸. In a second stage, this established method was also employed to obtain wines with standard defined colors.

Multivariate statistical analysis (MSA) of the investigated samples is used for establishing relations between the used vine, geographical origin, vintage, etc.¹⁹⁻²⁶. MSA comprises qualitative methods (e.g. principal component analysis (PCA)) or quantitative models (e.g. multiple linear regression (MLR))²⁷⁻²⁹. Arozarena *et al.*²⁰ found by MSA correlations for 20 analytical parameters such as color, pH, polyphenols, dry extract, alcohol content, volatile compounds or enological properties. These data are used to verify the geographical origin.

Near infrared reflectance spectroscopy (NIRS) is a fast and user-friendly alternative that is rugged enough for consistent batch-to-batch quality control^{30, 31} supported by sophisticated statistical software. NIRS offers the benefits of improved sample handling, reduced costs, quick analysis and the possibility to measure also chemical and physical properties. It has been successfully used for the analysis of carbohydrates in fruit juices over years³²⁻³⁴. Previous published works concentrated on the analysis of the ethanol³⁵ content as well on the determination of different parameters in Spanish³⁶ and Australian³⁷ wines. Cozzolino *et al.*³⁸ described the use of NIR for the discrimination

of Riesling and Chardonnay commercial wine varieties. In this work we present a NIR method for the cluster analysis of red wines according to their grape composition. Furthermore, a method for quantitative analysis of polyphenol content³⁹ and pH value beside the carbohydrate, total acid, tartaric acid and malic acid in Weißburgunder, Chardonnay, Ruländer, Silvaner, Müller Thurgau, Gewürztraminer, Sauvignon, Lagrein, Grossvernatsch, Blauburgunder, Cabernet and Merlot grapes during the wine production is presented. In a winery the carbohydrate content is used as a measure for the quality⁴⁰, while acids influence the taste and polyphenols the color³⁹. The presented method allows controlling the quality already at a very early stage during wine production and gives qualitative information on the final product by the determination of several ingredients. Compared to time-consuming standard procedures used in a winery to determine the content of the mentioned components, NIR-spectroscopic analysis can be carried out within a few seconds. Additionally, no expensive chemicals are needed. On the one hand this NIRS method is a helpful tool for a winery to optimize their wine production, on the other hand it protects the consumer from trickery.

Materials and Methods

Materials and reagents: Folin-Ciocalteu reagent, D(+)-glucose monohydrate (synthesis grade) and sodium carbonate (analytical reagent grade) were purchased from Merck (Darmstadt, Germany), D(-)-fructose (min. 99%) from Sigma (St. Louis, USA) and gallic acid-1-hydrate (reagent grade) from Riedel-de Haën (Hannover, Germany). For the identification of red wines *Cabernet Sauvignon*, *Lagrein*, *Chianti Classico*, *Don Miguel silver* and *Don Miguel red* were used. *Don Miguel* wines were a gift from Vinorica (Consell, Mallorca, Spain), all others from the producer Laimburg (Auer, Italy).

Near infrared reflectance spectroscopy (NIRS): NIR spectra were recorded with a scanning polarization interferometer NIR-spectrometer (Büchi, Uzwil, Switzerland) crosswise over a wavelength range from 4,000 to 10,000 cm⁻¹ in the transmission mode using a cuvette (silica glass, Suprasil, Hellma, Jena, Germany, optical thin layer 1 mm) and from 4,500 to 10,000 cm⁻¹ in the transfection mode using an optic glass fiber (silica glass, Infrasil, Bes Optics Inc., Warwick, Great Britain, length 2 m, optical thin layer 0.5, 1.0, 1.5, 2.0, 2.5, 3.0 mm). Ten scans were used for one average spectrum to equilibrate inhomogeneities. The instrument offers a resolution of 12 cm⁻¹, an absolute wavelength accuracy of ± 2 cm⁻¹ and a relative reproducibility of 0.5 cm⁻¹. Chemometrical software NirCal 3.0 and 4.0 (Büchi) was used for creating a model, i.e., selection of spectra and wavelengths, mathematical pretreatment and statistical analysis performing cluster analysis, principal component analysis (PCA) and partial least square regression (PLSR). Samples were thermo-stated in a water bath to 23°C (model PC/4, Julabo, Seelbach, Germany). Spectra were randomly divided into a so called learning-set (75%), i.e., calibration samples, and a test set (25%), i.e., samples for testing the calibration equation. The optimum number of factors used for the individual prediction was determined by cross-validation. Quality of cluster analysis was described in the Q-value calculated by the NirCal 4.0 software. The selection of the best quantitative regression model is based on the following calculated values: 1) BIAS, i.e., the average deviation between the predicted values

(y_n) and the actual values (x_n), in the calibration-set, should be close to zero.

$$Bias = \frac{1}{N} \sum (x_n - y_n)$$

2) PRESS, Predicted Residual Error Sum Square is the sum of the square of the deviation between predicted and reference values. The PRESS value of the validation set should be as small as possible and similar to that of the calibration set.

$$PRESS = \sum (x_n - y_n)^2$$

3) Standard error of estimation (SEE), i.e., the standard deviation of the differences between reference values and NIRS-results in the calibration set.

$$SEE = \sqrt{\frac{1}{N} \sum (x_n - y_n - Bias)^2}$$

4) Standard error of prediction (SEP), i.e., the counterpart for the test-set samples. SEE and SEP should be as small as possible.

$$SEP = \sqrt{\frac{1}{N} \sum (x_n - y_n - Bias)^2}$$

5) The correlation coefficient (R^2) should approach 1⁴¹.

Reference measurements

Carbohydrates: Refractometry³⁹ was used to determine the total carbohydrate content. Values are given in degree Öchsle. 1 degree Öchsle is equivalent to a density of 1.001 g/l. For the calculation of weight percent (gram carbohydrates per 100 gram must) corresponding to the Klosterneuburger Mostwaage (KMW), the following equation was used³⁹:

$$^{\circ}KMW = \frac{^{\circ}Öchsle}{4} - 3$$

pH: pH was determined using a Multilab 540 (WTW, Weinheim, Germany).

Acids: Total acid content was determined by alkaline titration³⁹, tartaric acid by colometric analysis according to the method of Rebelein⁴² and malic acid was quantified enzymatically^{43,44}.

Polyphenol content: For the determination of polyphenolic content the method according to Folin-Ciocalteu was used as a reference method^{45,46}. Gallic acid-1-hydrate was used as a reference standard in a concentration range from 0 to 4.93 µg/ml with equidistant steps. To 1 ml of each concentration 500 µl of the Folin-Ciocalteu reagent and 1.5 ml of a 20% sodium carbonate solution were added and the volume adjusted to 10 ml with water. Each concentration was determined fourfold. For the photometric determination of a double-beam spectrometer (model U-2000, Hitachi, Japan) and a silica glass cuvette (Suprasil, optical thin layer 10 mm) were used. Spectra were recorded at 750 nm.

Fifty grapes were crushed using a chopper and the resulting solution was centrifuged for 3 minutes at 6000 rpm. The supernatant was filled into a quartz cuvette (optical thin layer of 1 mm) and measured by NIRS in the transmission mode. For

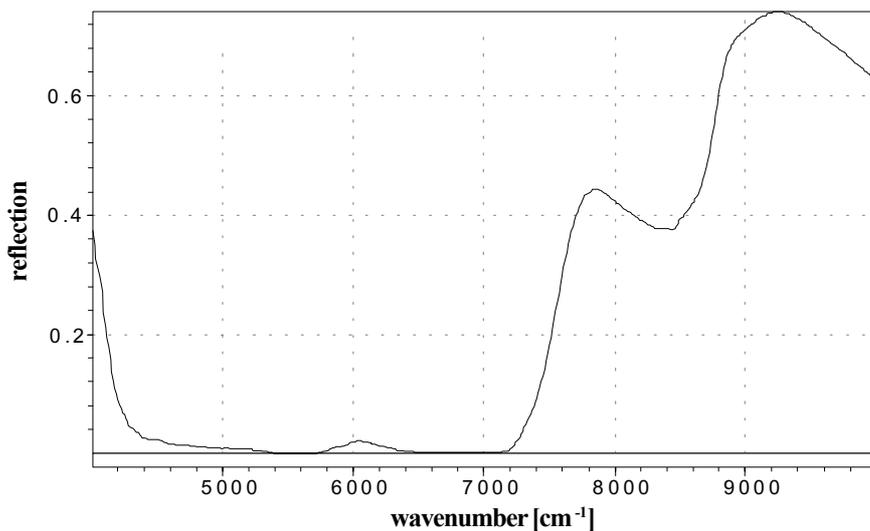


Figure 1. Transreflection spectrum of a red wine sample (*Lagrein*). Optical thin layer thickness, 3 mm; scans, 10; temperature, 23°C.

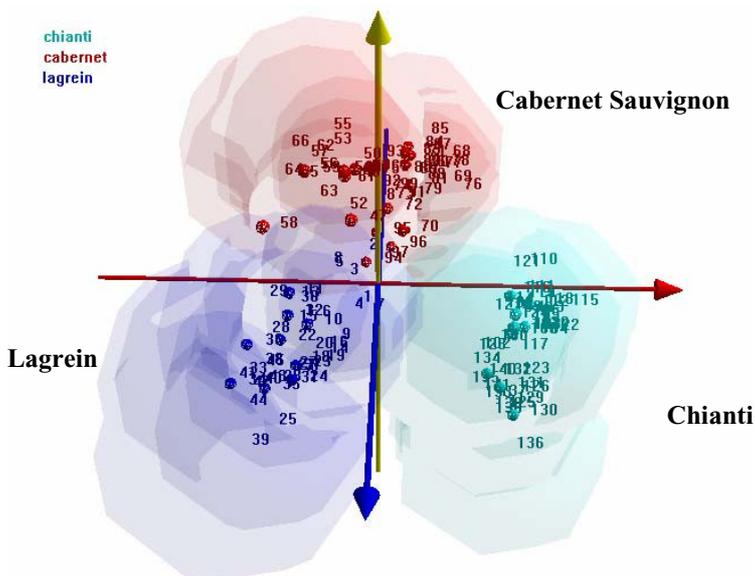


Figure 2. Factor-plot of 171 spectra of different wines (*Lagrein*, *Chianti Classico*, *Cabernet Sauvignon*). Conditions: normalization between 0 and 1, smoothing average 3 points, 1st derivative Savitzky-Golay 3 points; wavenumber range, 4500 - 10000 cm⁻¹; optical thin layer thickness, 3 mm; scans, 10; temperature, 23°C.

reference analysis following the procedure by Folin–Ciocalteu^{45, 46}, 10 µl were filled into a 10 ml volumetric flask, to which 500 µl Folin-Ciocalteu and 1.5 ml sodium carbonate 20% solution was added. After adjusting the volume to 100 ml, absorption was measured after 2 hours at a wavelength of 750 nm. The linear regression equation from gallic acid-1-hydrate, $y=0.1228x - 0.0013$ (y , absorption; x , concentration in µg/ml), $R^2=0.99$ was used to calculate the polyphenol concentration.

Results and Discussion

As the grape has a big influence on the quality of a red wine, a method to distinguish between different varieties was established.

Identification of red wines

Principal component analysis (PCA): 57 bottles of *Cabernet*

Sauvignon, *Lagrein*, both harvested in 1997/1998 and *Chianti Classico* (1996) were investigated. Each *Cabernet Sauvignon* and *Lagrein* is produced by Laimburg (Auer, Bolzano, Italy) in Southern Tirol from only one single vine, *Chianti*, produced by Antinori in the area of Toscana, Italy, is a mixture of *Sangiovese* (75-90%), *Canaiole* (5-10%), *Malvasia* (2-5%), *Trebbiano* (2-5%) and other grapes like *Colorino*, *Cabernet-Sauvignon* and *Merlot* (< 10%). After thermo stating of the samples at 23°C, 171 spectra of 3 different red wines were recorded using an optical glass fiber in transfection mode over a wavelength range from 4,500 to 10,000 cm⁻¹ (Fig. 1) and divided randomly into a calibration (75%) and validation set (25%). Before cluster analysis all recorded spectra were normalized between 0 and 1, smoothed (average 3 points) to improve signal to noise ratio, and transformed to their first derivative (Savitzky-Golay 3 points) to eliminate shifts in the baseline. Five primary factors were necessary to reach the best calibration model. Three-dimensional display after principal component analysis (PCA) and calculation of Mahalanobis-distances in a factor plot (Fig. 2) showed that spectra of the same wine can be united within one cluster without any outlier spectrum. The maximal spectra residual for calibration was 8.16×10^{-5} and the resulting Q-value 0.72. Due to the fact that the investigated *Cabernet Sauvignon* and *Lagrein* were produced in the same winery (Laimburg, Auer, Italy) from grapes coming from the exactly same area in Southern Tirol following the same protocol, it can be assumed that clustering results from the different grape sorts. *Chianti Classico* from Antinori in Toscana is originating from a completely different geographical region with differences in soil. The calibration model proved to be robust/stable against any age group influences and differences in ripeness of the used grapes. This classification model

is a full spectrum method comprising the interaction of all organic compounds. We did not compare occurrence, amount or isotope proportions of elements, but the sum of all parameters are used for calculation.

To test the sensitivity of the system due to variations in vine composition, two Spanish wines *Don Miguel red* and *Don Miguel silver* were analyzed. *Don Miguel red* contains 70% *Tempranillo* (Ull de Liebre) and 30% *Cabernet Sauvignon*, *Don Miguel silver* 70% *Cabernet Sauvignon* and 30% *Tempranillo* (Ull de Liebre). A few milliliter of a 23°C thermo stated sample were measured in transmission mode using a quartz cuvette in a wavenumber range from 4,200 to 10,000 cm⁻¹. Transmission offers higher spectral reproducibility compared to transfection and allows the measurement below 4,500 cm⁻¹, where combination vibrations of CC + CH, CH + CH occur. Again 75% of the recorded spectra were used for calibration and 25% for validation. Data preparation was

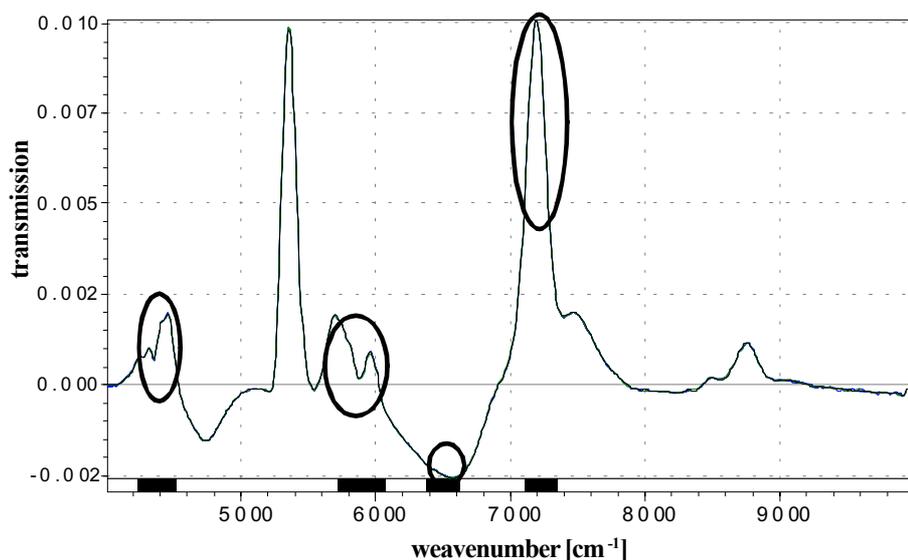


Figure 3. Pretreated spectrum of a *Don Miguel red* wine. Conditions: vector normalization, smoothing average 9 points, 1st derivative Savitzky-Golay 9 points; wavenumber range, 4,236-4,512 cm⁻¹, 5,724-6,072 cm⁻¹, 6,384-6,624 cm⁻¹ and 7,116-7,344cm⁻¹; optical thin layer thickness, 1 mm; scans, 10; temperature, 23°C.

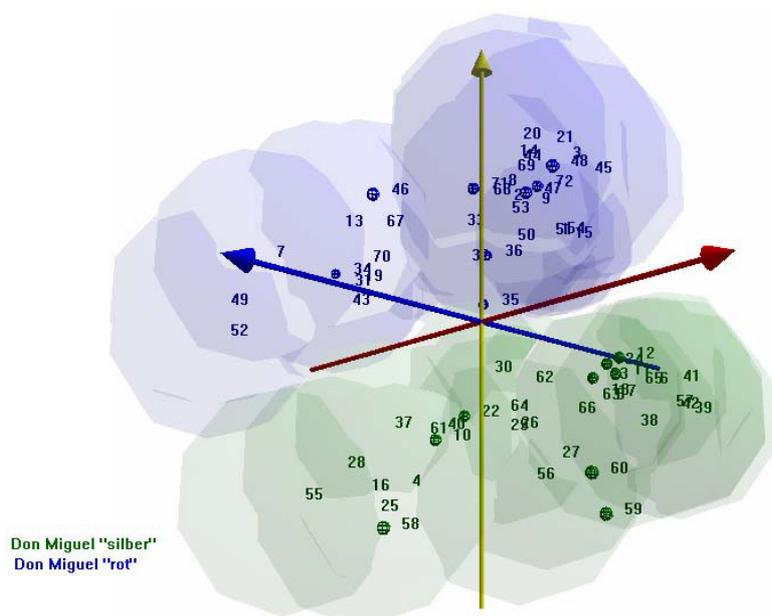


Figure 4. Factor-plot of 72 spectra of *Don Miguel red* and *silver* wines: Conditions: vector normalization, smoothing average 9 points, 1st derivative Savitzky-Golay 9 points; wavenumber range, 4,236 - 4,512cm⁻¹, 5,724 - 6,072 cm⁻¹, 6,384 - 6,624 cm⁻¹, 7,116 - 7,344 cm⁻¹; optical thin layer thickness, 1 mm; scans, 10; temperature, 23°C.

optimized to minimize baseline shifts, to optimize signal to noise ratio and to point out the differences between the two wines. Vector normalization, smoothing average 9 points and calculation 1st derivative Savitzky-Golay 9 points allowed to find differences in the spectra in the wavenumber ranges 4,236 - 4,512cm⁻¹, 5,726 - 6,072 cm⁻¹, 6,384 - 6,624 cm⁻¹, 7,116 - 7,344 cm⁻¹ (Fig. 3). 72 spectra of 24 bottles were used for calibration. After principal component analysis and calculation of Mahalanobis distances two uniform clusters with a Q-value of 0.75 were achieved, which is depicted in Fig. 4. Testing this model with 3 mixtures blends of 50% *Don Miguel red* and 50% *Don Miguel silver* resulted in factors in between the two original clusters. Although ethanol is present in a wine more than 10% (v/v), this confirms that the model is able to determine changes in the vine composition independent of alcohol

content, as both wines contain exactly the same amount of ethanol.

Partial least square regression (PLSR):

Calculation of the recorded spectra in the partial least square regression model using three factors allowed to distinguish between *Don Miguel red* and *silver* and to determine the mixing ratio of those. For data pretreatment vector normalization, smoothing average 9 points and calculation of the 1st derivative (Savitzky-Golay 9 points) in the wavenumber range 4,236 - 4,512 cm⁻¹, 5,724 - 6,072 cm⁻¹, 6,384 - 6,624 cm⁻¹ and 7,116 - 7,344 cm⁻¹ were carried out. The PRESS function showed the necessity to apply 2 factors for the calculation of the model. Recording of spectra using an optical thin layer of 1 mm (10 scans) resulted in the following equations: calibration (75% of spectra) was $f(x)=0.9189x+0.0405$ with a correlation coefficient of $R^2=0.95$, equation for validation (25%) was $f(x)=0.9261x+0.0311$ with $R^2=0.95$ (Fig. 5). The robustness of the system is high, which is demonstrated in the similarity of the results for SEE and SEP: 0.14 and 0.16.

Quantitative analysis of grape variety ingredients:

NIRS methods, which allow determining the carbohydrate, total acid, tartaric acid, malic acid, pH in grape variety and the polyphenol content in grapes were established. First method can control the quality already at a very early stage during the wine production and allows designing the quality of the final product. Grapes of 12 different vines (Weißburgunder, Chardonnay, Ruländer, Silvaner, Müller Thurgau, Gewürztraminer, Sauvignon, Lagrein, Grossvernatsch, Blauburgunder, Cabernet, Merlot) were harvested in autumn 2000 and squeezed. The obtained grape variety was thermo-stated at 23°C and analyzed quantitatively by NIRS in the transfection mode using an optical thin layer

thickness of 1 mm. Higher thickness was proved to be inappropriate due to total absorption between 4,500 and 5,300 cm⁻¹ and 6,564 - 7,164 cm⁻¹. In order to establish a calibration model 252 spectra of samples with lower and upper concentration as a reference, given in Table 1, were recorded. 76 % of all spectra were randomly used for calibration, 24% for validation. Data preparation was carried out in order to minimize technical influences, which mainly cause a drift in baseline. Quantitative analysis was carried out by partial least square regression (PLSR).

Carbohydrates. Data preparation comprised normalization between 0 and 1 and following calculation of the first derivative using a wavenumber range from 4,500 to 7548 cm⁻¹. The PRESS function showed that 3 factors were needed for the calculation of the model. Calculation with 3 factors resulted in a good conformity

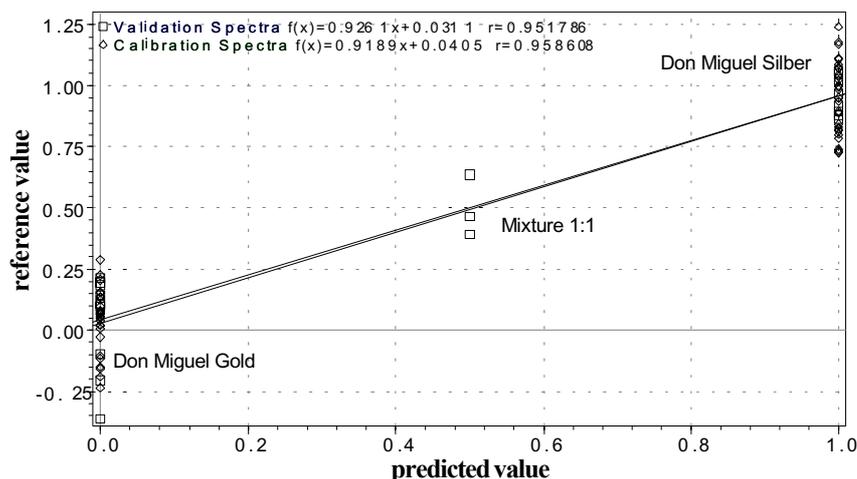


Figure 5. Predicted (NIRS) versus true property for the determination of *Don Miguel red* and *silver* mixing ratios. Conditions: vector normalization, smoothing average 9 points, calculation of 1st derivative Savitzky-Golay 9 points; wavenumber range, 4,236 - 4,512 cm⁻¹, 5,724 - 6,072 cm⁻¹, 6,384 - 6,624 cm⁻¹, 7,116 - 7,344 cm⁻¹; optical thin layer thickness, 1 mm; scans, 10; temperature, 23°C.

Table 1. Upper and lower limit of reference values used for calibration.

Concentration	Carbohydrates [KMW]	Total acid [g/l]	Tartaric acid [g/l]	Malic acid [g/l]	pH
Minimum	13.1	5.0	3.1	2.9	3.09
Maximum	19.8	11.0	6.7	7.0	3.74

Table 2. Calibration results for the determination of the carbohydrate, total acid, tartaric acid, malic acid content and pH.

Parameter	Unit	Regression Equation Calibration		Regression Equation Validation		R ² Calibration		R ² Validation	
		a	b	a	b	a	b	a	b
Carbohydrates	KMW	f(x)=0.993x+0.1159	f(x)=0.9839x+0.9918	f(x)=0.9973x+0.0343	f(x)=0.9692x+0.4539	0.99	0.99	0.99	0.99
Total acids	g/l	f(x)=0.8616x+1.0340	f(x)=0.8616x+1.0340	f(x)=0.6104x+0.1159	f(x)=0.6104x+3.0884	0.92	0.92	0.85	0.85
Tartaric acid	g/l	f(x)=0.8308x+0.8633	f(x)=0.8219x+0.9089	f(x)=0.5936x+2.1505	f(x)=0.6941x+1.6032	0.91	0.90	0.87	0.84
Malic acid	g/l	f(x)=0.7926x+0.8903	f(x)=0.7274x+0.12762	f(x)=0.6139x+1.8469	f(x)=0.4460x+2.7197	0.89	0.85	0.80	0.71
pH		f(x)=0.6753x+1.0868	f(x)=0.5398x+1.5406	f(x)=0.6476x+1.1613	f(x)=0.4802x+1.7244	0.82	0.73	0.89	0.73

Note. a Single analysis b Simultaneous analysis

Table 3. Prediction results for the determination of the carbohydrate, total acid, tartaric acid, malic acid content and pH.

Parameter	Unit	SEE		SEP		BIAS	
		a	b	a	b	a	b
Carbohydrates	KMW	0.13	0.21	0.11	0.19	2.30×10^{-15}	3.33×10^{-16}
Total acids	g/l	0.60	0.43	0.61	0.53	7.17×10^{-15}	-1.08×10^{-14}
Tartaric acid	g/l	0.40	0.41	0.54	0.55	-1.08×10^{-14}	-3.43×10^{-15}
Malic acid	g/l	0.43	0.49	0.55	0.65	-4.25×10^{-15}	-2.44×10^{-15}
pH		0.07	0.09	0.06	0.09	-1.26×10^{-15}	-7.15×10^{-15}

Note. a Single analysis b Simultaneous analysis.

between SEE and SEP (Fig. 6f). Linear regression between true and predicted values resulted in a value for the correlation coefficient of $R^2=0.99$ for calibration and $R^2=0.99$ for validation (Table 2, Fig. 6a). Results for SEE and SEP: 0.13° KMW and SEP: 0.13° KMW and 0.11° KMW, the BIAS value is 2.30×10^{-15} (Table 3).

Total acids. PLSR in a concentration range between 5 and 11 g/l included normalization between 0 and 1, full multiplicative scatter correction (MSC) and calculation of the 1st derivative

(Taylor 3 points) between 4,500 and 7,548 cm⁻¹. Three factors were necessary to obtain a minimum for PRESS and an agreement between SEE (0.60 g/l) and SEP (0.61 g/l) of nearly 100% (Table 3, Fig. 6f). The highly linear model (Fig. 6b) allows determining the total acid content with a prediction error of 0.61 g/l (Table 3).

Tartaric acid. After normalization, performing of the 1st derivative over between 4,500 and 7,308 cm⁻¹, four factors (Fig. 6f) were used for creation of the highly linear model depicted in Fig. 6c with $R^2=0.91$ for calibration and $R^2=0.87$ for validation (Table 2). Despite the small concentration range between 3.1 and 6.7 g/l used for

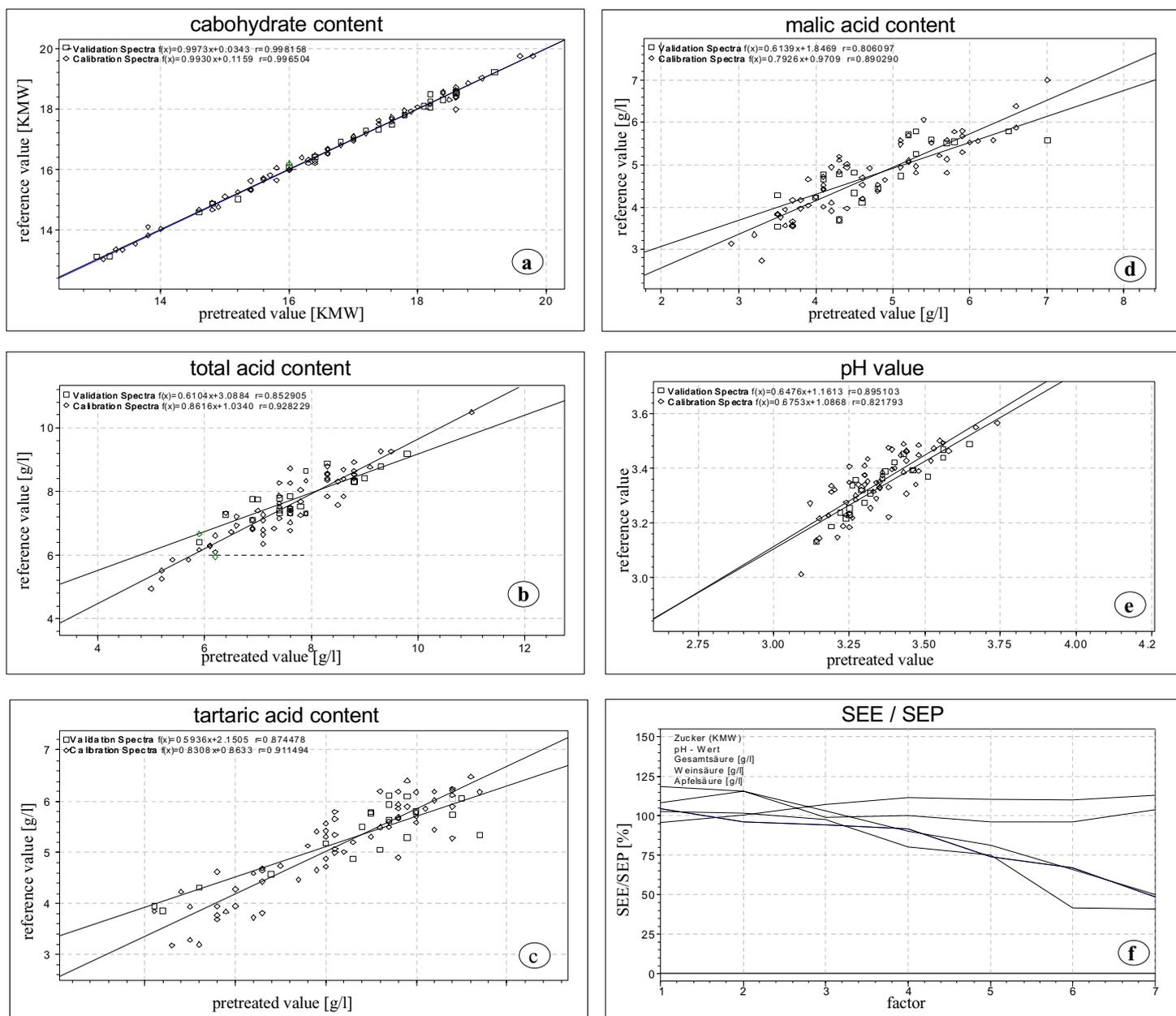


Figure 6. Predicted (NIRS) versus true property for the determination of (a) carbohydrate, (b) total acid, (c) tartaric acid, (d) malic acid content, (e) pH, (f) SEE vs. SEP. Conditions: normalization (e) between 0 and 1, 2nd derivative Taylor 3 points; scans, 7; optical thin layer thickness, 0.5 mm; temperature, 23°C.

calibration (Table 1) this system allows to determine the tartaric acid content in grape variety with an absolute error of estimation of 0.40 g/l and prediction of 0.54 g/l. (Table 3, Fig. 6f).

Malic acid. Malic acid often shows 2-5 times higher values compared to tartaric acid. Calibration between 2.9 and 7.0 g/l (Table 1) after normalization between 0 and 1 and calculation of a second smoothed derivative was carried out using three factors, SEE and SEP showing acceptable agreement (Fig. 6f). Absolute values for SEE, and BIAS were 0.43 g/l and -4.25×10^{-15} (Table 3). Straight line for calibration showed a linearity of $R^2=0.89$ (Table 2) and allowed a prediction of the malic acid content with an absolute error of 0.55 g/l (Table 3).

pH. Normalization and calculation of the smoothed 2nd derivative between 4,500 and 7,308 cm^{-1} showed an optimum for BIAS at five factors (Table 3). Despite the narrow calibration range of pH 3.09-3.74 the calibration equation shows a R^2 of 0.82

(Table 2, Fig. 6e).

In order to enable the determination of these parameters with only one single measurement, simultaneous analysis of the carbohydrate, total acid, tartaric acid and malic acid content and pH was achieved by performing normalization (between 0 and 1) and calculating its 2nd derivative (Taylor 3 points). Four factors over a wavenumber range from 4,500 to 7,308 cm^{-1} showed 73-100% agreement between SEE and SEP. Linear regression showed high linearity for each investigated parameter with slightly lower values for R^2 (Table 2). Compared to the above-described single analysis this method allows a quantitative analysis of all parameters at once within a few seconds. Values for SEP are slightly increased (Table 3).

Polyphenols mainly influence taste, sensory properties and color of a wine. Therefore, a rapid method to analyze its quantity is important. The method according to Folin-Ciocalteu was used as

Table 4. Calibration and prediction results for the determination of gallic acid and polyphenol content.

Parameter	Unit	Regression Equation Calibration	Regression Equation Validation	R ² Calibration	R ² Validation	SEE ¹	SEP ¹	BIAS ¹
Gallic acid	mg/ml	$f(x)=0.9637x+0.1274$	$f(x)=0.9632x+0.0885$	0.98	0.96	0.45	0.46	8.37×10^{-15}
Polyphenols	mg/ml	$f(x)=0.9506x+0.0881$	$f(x)=0.9900+0.0497$	0.97	0.96	0.45	0.46	7.86×10^{-16}

¹ Calibration Set.

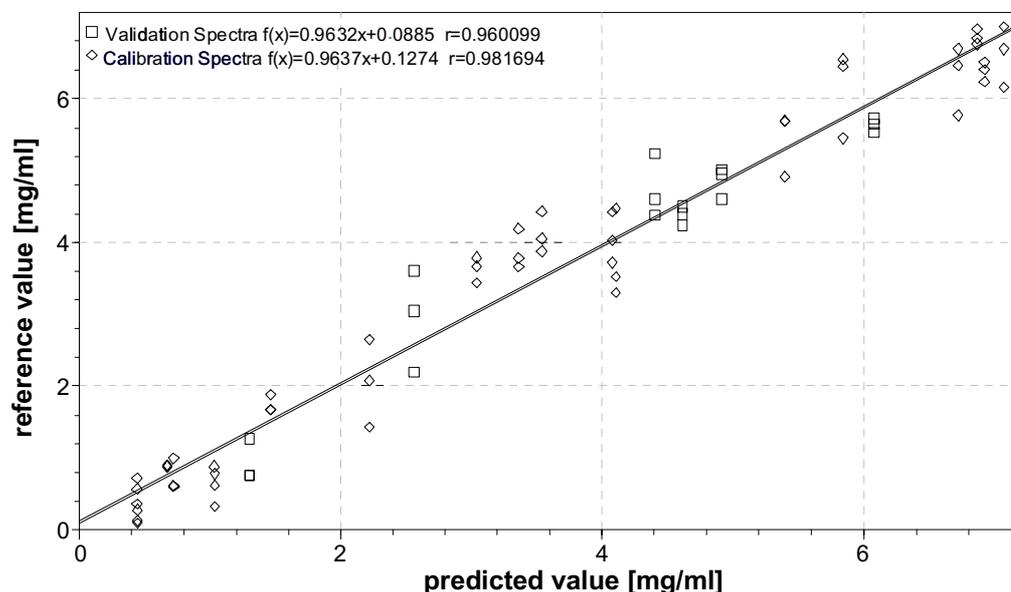


Figure 7. Predicted (NIRS) versus true property for the determination of gallic acid content. Conditions: normalization between 0 and 1, 1st derivative Savitzky-Golay; scans, 10; optical thin layer thickness, 0.5 mm; temperature, 23°C.

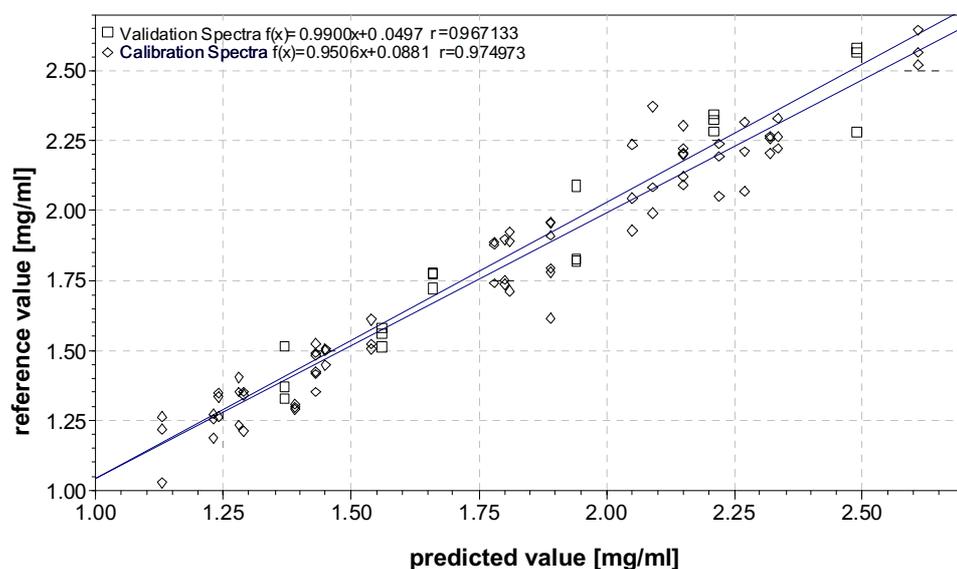


Figure 8. Predicted (NIRS) versus true property for the determination of polyphenol content. Conditions: normalization by Closure, 1st derivative Taylor; scans, 10; optical thin layer thickness, 0.5 mm; temperature, 23°C.

a reference method (see Materials and Methods). Gallic acid-1-hydrate was used as reference standard in a concentration range from 0 to 4.93 $\mu\text{g/ml}$ with equidistant steps. Twenty four gallic acid-1-hydrate solutions in a concentration range between 0.442 and 7.08 mg/ml were measured in the transmission mode threefold

and in random order by NIRS. Evaluation using PLSR was achieved by dividing 72-recorded spectra randomly into a calibration (54 spectra) and validation (18 spectra) set. Data pretreatment comprised normalization between 0 and 1 and calculation of the 1st derivative (Savitzky-Golay) between 4,008-7,512 cm^{-1} . Using three factors, the PRESS function showed a minimum and a good agreement between SEE (0.45 mg/ml) and SEP (0.46 mg/ml) (Table 4). Linear regression (Fig. 7) between predicted and true

Abbreviations: GC, gas chromatography; HPLC, high-performance liquid chromatography; ICP, inductively coupled plasma; MS, mass spectrometry; MSA, multivariate statistical analysis; PCA, principal component analysis; PLS, partial least square; PRESS, predicted residual error sum square; RP, reversed-phase; SEE, standard error of estimation; SEP, standard error of prediction.

values allowed to predict the gallic acid-1-hydrate concentration between 0 and 7 mg/ml with $R^2=0.98$ (Table 4).

In order to determine the total polyphenol concentration 30 must samples were measured in the transmission mode threefold and in random order. 90 spectra were divided into 72 calibration and 18 validation spectra. Normalization and performing of the 1st derivative allowed minimizing shifts in the baseline. 4 factors were necessary to obtain a minimum for the PRESS function and to get a maximum agreement of SEE and SEP (Table 4). Linear regression (Fig. 8) allowed correlating true and predicted values with a R^2 of 0.97 (Table 4). Compared to the traditionally used Folin - Ciocalteu method in a winery, which is very time-consuming and expensive due to the usage of different chemicals, the NIRS method is very simple, precise and incomparably fast.

Compared with the results obtained by other authors our established methods yielded R^2 , SEE and SEP values close to the values in the literature or even better³⁵⁻³⁷.

Finally, the established NIRS method possesses the advantage compared to laborious traditional analytical techniques of short analysis time and simultaneous determination.

References

- Johnson, H. 2002. Johnsons Weingeschichte. Hallwag Verlag, Munich.
- De la Calle Garcia, D., Reichenbacher, M., Danzer, K., Hurlbeck, C., Bartzsch, C. and Feller, K.-H. 1998. Analysis of wine bouquet components using headspace solid-phase microextraction-capillary gas chromatography. *J. High Resol. Chromatogr.* **21**(7):373-377.
- De la Calle Garcia, D., Reichenbacher, M. and Danzer, K. 1998. Classification of wines by means of multivariate data analysis using the SPME/CGC-chromatograms of volatile aroma compounds. *Vitis* **37**(4):181-188.
- Marengo, E., Aceto, M. and Maurino, V. 2002. Classification of Nebbiolo-based wines from Piedmont (Italy) by means of solid-phase microextraction-gas chromatography-mass spectrometry of volatile compounds. *J. Chromatogr. A* **943**(1):123-137.
- Danzer, K., De la Calle Garcia, D., Thiel, G. and Reichenbacher, M. 1999. Classification of wine samples according to origin and grape varieties on the basis of inorganic and organic trace analyses. *American Lab.* **31**(20): 26, 28, 30, 32-34.
- Weber, J., Beeg, M., Bartzsch, C., Feller, K.-H., De la Calle Garcia, D., Reichenbacher, M. and Danzer, K. 1999. Improvement of the chemometric variety characterization of wines by improving the detection limit for aroma compounds. *J. High Resol. Chromatogr.* **22**(6):322-326.
- Baxter, M.J., Crews, H.M., Dennis, M.J., Goodall, I. and Anderson, D. 1997. The determination of the authenticity of wine from its trace element composition. *Food Chemistry* **60**(3):443-450.
- Gomez-Plaza, E., Gil-Munoz, R. and Martinez-Cutillas, A. 2000. Multivariate classification of wines from seven clones of Monastrell grapes. *J. Sci. Food and Agric.* **80**(4):497-501.
- Day, M.P., Zhang, B. and Martin, G. 1995. The use of trace element data to complement stable isotope methods in the characterization of grape musts. *J. Sci. Food and Agric.* **67**(1):113-123.
- Montanarella, L., Bassani, M.R. and Breas, O. 1995. Chemometric classification of some European wines using pyrolysis mass spectrometry. *Rapid Commun. Mass Spectrom.* **9**:1589-1593.
- Soufleros, E.H., Boulompasi, E., Tsachopoulos, C. and Bilideris, C.G. 2002. Primary amino acid profiles of Greek white wines and their use in classification according to variety, origin and vintage. *Food Chemistry* **80**(2):261-263.
- Eder, R., Wendelin, S. and Barna, J. 1994. Classification of red wine cultivars by anthocyanin analysis. Part 1. Application of multivariate statistical methods for differentiation of grape samples. *Mitteilungen Klosterneuburg* **44**(6):201-212.
- Berente, B., De la Calle Garcia, D., Reichenbacher, M. and Danzer, K. 2000. Method development for the determination of anthocyanins in red wines by high-performance liquid chromatography and classification of German red wines by means of multivariate statistical methods. *J. Chromatogr. A* **871**:95-103.
- Pena-Neira, A., Fernandez, T., Garcia-Vallejo, C., Estrella, I. and Suarez, J.A. 2000. A survey of phenolic compounds in Spanish wines of different geographical origin. *Europ. Food Res. Technol.* **210**(6):445-448.
- Picque, D., Cattenoz, T. and Corrieu, G. 2001. Classification of red wines analyzed by middle infrared spectroscopy of dry extracts according to their geographical origin. *J. Intern. Des Sci. De La Vigne Et Du Vin* **35**(3):165-170.
- Edelmann, A., Diewok, J., Schuster, C.K. and Lendl, B. 2001. Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts. *J. Sci. Food Agric.* **49**:1139-1145.
- Palma, M. and Barroso, C.G. 2002. Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta* **58**(2):265-271.
- Heredia, F.J., Troncoso, A.M. and Guzman-Chozas, M. 1997. Multivariate characterization of aging status in red wines based on chromatic parameters. *Food Chemistry* **60**(1):103-108.
- Kallithraka, S., Arvanitoyannis, I.S., Kefalas, P., El-Zajouli, A., Soufleros, E. and Psarra, E. 2001. Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin. *Food Chemistry* **73**(4):501-514.
- Arozarena, I., Casp, A., Marin, R. and Navarro, M. 2000. Multivariate differentiation of Spanish red wines according to region and variety. *J. Sci. Food Agric.* **80**(13):1909-1917.
- Perez-Magarino, S. and Gonzalez-SanJose, M.L. 2002. Physico-chemical parameters justifying the vintage qualifications in wines from Spanish Protected Designation of Origin. *Europ. Food Res. Technol.* **214**(5):444-448.
- Etievant, P., Schlich, P., Bouvier, J.C., Symonds, P. and Bertrand, A. 1988. Varietal and geographic classification of French red wines in terms of elements, amino acids and aromatic alcohols. *J. Sci. Food Agric.* **45**(1): 25-41.
- Perez-Magarino, S., Revilla, I., Izcarra, E. and Gonzalez-San Jose, M.L. 1999. Discrimination of D.O.C. wine vintages as a function of chemical composition. *Quimica Analitica* **18**(1):133-136.
- Sivertsen, K., Holen, B., Nicolaysen, F. and Risvik, E. 1999. Classification of French red wines according to their geographical origin by the use of multivariate analyses. *J. Sci. Food Agric.* **79**:107-115.
- Cosmos, E., Heberger, K. and Simon-Sarkadi, L. 2002. Principal component analysis of biogenic amines and polyphenols in Hungarian wines. *J. Agric. Food Chem.* **50**(13):3768-3774.
- Rebollo, S., Pena, R.M., Latorre, M.J., Garcia, S., Botana, A.M. and Herrero, C. 2000. Characterization of Galician (NW Spain) Ribeira Sacra wines using pattern recognition analysis. *Anal. Chim. Acta* **417**(2):211-220.
- Arvanitoyannis, I.S., Katsota, M.N., Psarra, E.P., Soufleros, E.H. and Kallithraka, S. 1999. Application of quality control methods for assessing wine authenticity: Use of multivariate analysis (chemometrics). *Trends Food Sci. Technol.* **10**:321-326.
- Gonzalez, G. and Pena-Mendez, E.M. 2000. Multivariate data analysis in classification of must and wine from chemical measurements. *Europ. Food Res. Technol.* **212**:100-107.
- Martens, M. and Naes, T. 1989. Assessment, validation and choice of calibration method. In *Multivariate Calibration*, Wiley, New York, pp. 237-266.
- Grummisch, U. 1998. Qualitative and quantitative near infrared spectroscopy applications in the pharmaceutical industry. *Pharm. Ind.* **60**:1002-1006.

- ³¹Lüderwald, I. and Müller, M. 1993. In Günzler, H., Borsdorf, R., Danzer, K., Fresenius, W., Huber, W., Lüderwald, I., Tölg, G. and Wissler, H. (eds). *Analytiker Taschenbuch 11*. Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest, pp. 135-141.
- ³²Segtnan, V.H. and Isaksson, T. 1999. On-line versus at-line analysis of carbohydrates in fruit juice. *Proceedings of the ninth International Conference on Near Infrared Spectroscopy*, pp. 391-395.
- ³³Rodríguez-Saona, L.E., Fry, F.S., McLaughlin, M.A. and Calvey, E.M. 2001. Rapid analysis of sugars in fruit juices by FT-NIR spectroscopy. *Carbohydr. Res.* **336**:63-74.
- ³⁴Rambla, F.J., Garrigues, S. and de la Guardia, M. 1997. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Anal. Chim. Acta* **344**:41-53.
- ³⁵Urbano Cuadrado, M., Luque de Castro, M.D., Perez Juan, P.M. and Gomez-Nieto, M.A. 2005. Comparison and joint use of near infrared spectroscopy and Fourier transform and mid infrared spectroscopy for the determination of wine parameters. *Talanta* **66**:218-224.
- ³⁶Urbano Cuadrado, M., Luque de Castro, M.D., Perez Juan, P.M. and Gomez-Nieto, M.A. 2004. Near infrared reflectance spectroscopy and multivariate analysis in enology; Determination or screening of fifteen parameters in different types of wines. *Anal. Chim. Acta* **527**:81-88.
- ³⁷Gishen, M., Damberg, R.G., Kambouris, A., Kwiatkowski, M., Cynkar, W.U., Hoj, P.B. and Francis, I.L. 1999. Application of near infrared spectroscopy for quality assessment of grapes, wine and spirits. *Proceedings of the ninth International Conference on Near Infrared Spectroscopy*, pp. 917-920.
- ³⁸Cozzolino, D., Smyth, H.E. and 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. *J. Agric. Food Chem.* **51**:7703-7708.
- ³⁹Würidg, G. and Woller, R. 1989. *Chemie des Weines*. Verlag Eugen Ulmer, Stuttgart.
- ⁴⁰Bergner, K.G. 1984. Quality of wine from a chemical viewpoint. *Deutsche Apotheker Zeitung* **124**:2179-2183.
- ⁴¹Büchi, A.G. 1993. *Handbook of the NIR-spectrometer system*. Uzwil, Switzerland.
- ⁴²Almela, L., Lazaro, D., Lopez-Roca, J.M. and Fernandez-Lopez, J.A. 1993. Tartaric acid in frozen musts and wines. Optimization of Rebelein's method and validation by HPLC. *Food Chemistry* **47**(4):357-361.
- ⁴³Knichel, W. and Radler, F. 1982. Method for the enzymic determination of D(+)-malate. *Zeitschrift für Lebensmittel-Untersuchung und -Forschung* **174**(4):296-299.
- ⁴⁴Mayer, K. and Busch, I. 1963. An enzymic determination of malic acid in wine and grape juice. *Mitt. Gebiete Lebensm. Hyg.* **54**:60-65.
- ⁴⁵Singleton, V.L. and Rossi, J.A. 1965. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Viticult.* **16**(3):144-158.
- ⁴⁶Slinkard, K. and Singleton, V.L. 1977. Total phenol analysis: automation and comparison with manual methods. *Am. J. Enol. Viticult.* **28**:49-55.