JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Principal Component Analysis Applied to Fourier Transform Infrared Spectroscopy for the Design of Calibration Sets for Glycerol Prediction Models in Wine and for the Detection and Classification of Outlier Samples

HELENE H. NIEUWOUDT,*,†,‡ BERNARD A. PRIOR,† ISAK S. PRETORIUS,‡,|| MARENA MANLEY,§ AND FLORIAN F. BAUER‡

Department of Microbiology, Stellenbosch University, Private Bag X1, 7602 Matieland (Stellenbosch), South Africa, Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, 7602 Matieland (Stellenbosch), South Africa, Department of Food Science, Stellenbosch University, Private Bag X1, 7602 Matieland (Stellenbosch), South Africa, and The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, SA-5064, Australia

Principal component analysis (PCA) was used to identify the main sources of variation in the Fourier transform infrared (FT-IR) spectra of 329 wines of various styles. The FT-IR spectra were gathered using a specialized WineScan instrument. The main sources of variation included the reducing sugar and alcohol content of the samples, as well as the stage of fermentation and the maturation period of the wines. The implications of the variation between the different wine styles for the design of calibration models with accurate predictive abilities were investigated using glycerol calibration in wine as a model system. PCA enabled the identification and interpretation of samples that were poorly predicted by the calibration models, as well as the detection of individual samples in the sample set that had atypical spectra (i.e., outlier samples). The Soft Independent Modeling of Class Analogy (SIMCA) approach was used to establish a model for the classification of the outlier samples. A glycerol calibration for wine was developed (reducing sugar content < 30 g/L, alcohol > 8% v/v) with satisfactory predictive ability (SEP = 0.40 g/L). The RPD value (ratio of the standard deviation of the data to the standard error of prediction) was 5.6, indicating that the calibration is suitable for quantification purposes. A calibration for glycerol in special late harvest and noble late harvest wines (RS 31−147 g/L, alcohol > 11.6% v/v) with a prediction error SECV = 0.65 g/L, was also established. This study yielded an analytical strategy that combined the careful design of calibration sets with measures that facilitated the early detection and interpretation of poorly predicted samples and outlier samples in a sample set. The strategy provided a powerful means of quality control, which is necessary for the generation of accurate prediction data and therefore for the successful implementation of FT-IR in the routine analytical laboratory.

KEYWORDS: FT-IR spectroscopy; principal component analysis; outliers; glycerol; wine

INTRODUCTION

The potential of Fourier transform infrared spectroscopy (FT-IR) as a powerful analytical tool in enology has been recognized for many years. The technology is based on the measurement of the frequencies of the vibrations of chemical bonds in functional groups such as C−C, C−H, O−H, C═O and N−H, upon absorption of radiation in the mid infrared (IR) region (*1*). The IR region is usually defined as ranging from 4000 to 400 cm^{-1}, or in terms of nanometers, from 2500 to 2.5×10^4 nm (*1*). The measured frequencies are processed through a series of mathematical procedures (which include Fourier transformation) to an absorbance spectrum, which in turn is correlated to the actual concentrations of the relevant components in the sample matrix through a calibration process that involves multivariate statistical procedures such as principal component analysis (PCA), principal component regression (PCR) and partial least squares (PLS) regression (*2*, *3*). Recent improvements in IR instrumentation (*4*) and the development of versatile

* To whom correspondence should be addressed. Current address: Dr HH Nieuwoudt. Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, 7602 Matieland (Stellenbosch), South Africa. Tel.: +27 21 808 2569. Fax +27 21 808 3771. E-mail: hhn@sun.ac.za.

† Department of Microbiology, Stellenbosch University.

‡ Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University.

§ Department of Food Science, Stellenbosch University.

|| Present address: The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, SA, Australia.

10.1021/jf035431q CCC: $27.50
Published on Web 05/13/2004

Table 1. Wines Used for FT-IR Spectroscopy ($n = 329$)

style/description[a]	no.of wines
dry white, maximum sugar 4 g/L[a]	
Chenin blanc, Sauvignon blanc, Chardonnay, blends, maximum 2 years aging, unwooded[b] wines	23
wooded[b] wines	63
dry red, maximum sugar 4 g/L[a]	
Shiraz, Pinotage, Merlot, Cabernet Sauvignon, maximum 2 years aging	101
various blends, 3–6 years aging	27
single cultivars of Malbec, Pinot Noir, Petit Verdot	9
off-dry white, sugar 4.1–12 g/L[a]	
Chenin blanc, Sauvignon blanc, unwooded[b] andwooded[b] styles	17
sweet wines (Semillon, Gewurztraminer, Weisser Riesling and blends)	
special late harvest, maximum sugar content 50 g/L[a]	15
noble late harvest, minimum sugar content 50 g/L[a]	28
Blanc De Noir, maximum sugar content 30 g/L[a]	
Pinotage, red muscadel	5
low alcohol wines, alcohol less than 10% v/v, sugar content specified[a]	2
Chenin blanc and blend	
young wines[c]	39
total	329

[a] Specified according to the South African National Wine Show Association. [b] Unwooded is defined as "no noticeable wood character"; wooded as "noticeable wood character". [c] Young wines close to the end of fermentation, but not yet bottled.

and innovative software applications designed specifically for wine analysis (*13*) have optimized this technology. Currently, multicomponent analytical instruments with impressive performance data in terms of accuracy, precision, and speed of analysis are available. The evaluation of the use of one such an instrument, the WineScan FT120 (*13*) has recently received much attention (*5*−*8*).

Glycerol ($CH_2OH-CHOH-CH_2OH$) is quantitatively a major component of wine, and its determination at various stages of the winemaking process provides important information regarding issues that are directly or indirectly related to quality control (*9*−*11*). The application of FT-IR for quantifying glycerol in dry wine has been tested, and standard error of prediction (SEP) values of 0.49 g/L (*5*) and 1.32 g/L (*6*) have been reported. In addition, a ready-to-use calibration for glycerol in dry finished wine (SEP = 1.13 g/L) was recently made commercially available (*12*). On the basis of theoretical considerations (and assuming a normal probability distribution) it should be kept in mind that the prediction error of only ca. 68% of the samples in a sample set will lie within ± one times the SEP value, whereas the remaining ca. 32% of the samples will have prediction errors that can be expected to be larger than this interval. The evaluation of calibration models based on regression statistics alone, therefore, often presents an over optimistic view of the predictive abilities of a model and provides limited information in terms of: (i) sample types for which the model would not be suitable; (ii) identifying poorly predicted samples in the sample set; and (iii) detecting and interpreting extreme deviating samples, for example, the so-called "outlier samples" in the sample set, if present (*3*). These aspects are of particular importance in routine analysis of commercial wines, where wide variation in terms of style, vintage, cultivar, geographic origin, chemical composition, maturation periods, and process technologies are encountered. It is to be expected that this variation will be reflected in the spectral properties of the samples and that some of these sources of variation will have major implications for the accuracy of prediction of calibration models. For the successful application of FT-IR in the analytical laboratory, accurate prediction data are required. A strategy aimed both at the development of robust calibration models encompassing this variation, as well as implementing quality control measures that enable the early detection and interpretation (where possible) of poorly predicted samples and outlier samples, is therefore required.

In this study, the WineScan FT120 instrument was used to generate the FT-IR spectra of a large number of wines of various styles, and PCA was used as a tool to identify the main sources of variation between the wines. The implications of this variation for the accuracy of prediction, and hence the design of calibration sets, were evaluated using glycerol calibration in wine as a model system. The PCA results were also used to establish a classification model for the early detection and interpretation of outlier samples in the sample set by using the Soft Independent Modeling of Class Analogy (SIMCA) approach.

MATERIALS AND METHODS

Wine Samples. The sample set consisted of bottled commercial South African red wines and white wines ($n = 290$), as well as young wines (both red wines and white wines, $n = 39$) that were close to the end of fermentation, but not yet bottled (**Table 1**). Collectively, the wines in the sample set represented more than 13 different cultivars and 22 different wine styles, as well as wide variation in terms of process technologies, geographic origin, vintage, and maturation periods. The majority of the commercial wines ($n = 263$) were of vintages from 1998 to 2002. A subset of red wines ($n = 27$) were of vintages older than 1998 and contained samples that had undergone three or more years maturation at the time of analysis.

FT-IR Spectral Measurements. Commercial wine samples were scanned upon reception without any further sample preparation. The young wines that appeared turbid were filtered with a Filtration Unit (type 79500, Foss Electric, Denmark) using filter paper circles graded at 20−25 μm and with diameter 185 mm (Schleicher & Schuell, reference number 10312714), to avoid disturbances in the optical path length of the cuvette. A WineScan FT120 instrument (Foss Electric, Denmark) that employs a Michelson interferometer was used to generate the FT-IR spectra. Because the WineScan FT120 is a specialized instrument designed specifically to generate quantitative data against the background matrix of wine, the number of scans generated per sample, the selection of wavenumbers, and the processing of the spectra have been pre-selected by the manufacturer and are not accessible to change by the user. Samples (7 mL) were pumped through the CaF_2-lined cuvette (optical path length 37 μm), which is housed in the heater unit of the instrument. The temperature of the samples are brought to

exactly 40 °C before analysis. Samples were scanned from 5011 to 929 cm^{-1} at 4 cm^{-1} intervals (i.e, 1056 data points per spectrum), which includes a small section of the near-IR region. The frequencies of the IR beam transmitted by a sample were recorded at the detector and used to generate an interferogram that is calculated from a total of 20 scans before being processed by Fourier transformation to generate a single beam transmittance spectrum. Background absorbance in the wine sample (which includes the absorbance of water) is corrected through the use of a Foss Zero Liquid S-6060 (*13*), which is scanned prior to the wine sample. The single beam transmittance spectrum of the zero liquid is stored on the computer of the instrument, and the ratio of the sample spectrum to the zero liquid spectrum, at each recorded data point, is used to generate the final transmittance spectrum. Absolute repeatability of the spectral measurements is calculated as the standard deviation of the transmittance, at each recorded data point, of the 20 replicate scans of the same sample (*13*). The transmittance spectra were finally converted into linearized absorbance spectra through a series of mathematical procedures. An aqueous solution of glycerol (10 g/L, analytical grade, BDH) was scanned under the same conditions as the wine samples.

Multivariate Data Analysis. *Principal Component Analysis* (*PCA*). FT-IR spectra were exported to the Unscrambler Software (version 6.11, Camo ASA, Trondheim, Norway) for the purpose of PCA. Duplicate spectra were averaged. The complete data set, defined by the variables in the columns (in this study, 1056 wavenumbers) and the samples in the rows, was autoscaled through mean centering by column. PCA models the maximum directions of variation in a data set by projecting the objects (in this study, the FT-IR spectra) as a swarm of points in a space defined by principal components (PC's). Each PC is a linear function of a number of original variables, resulting in a reduction of the original number of variables. PC's describe, in decreasing order, the most variation among the objects, and because they are calculated to be orthogonal to one another, each PC can be interpreted independently. This permits an overview of the data structure by revealing relationships between the objects as well as the detection of deviating objects. To find these sources of variation, the original data matrix, defined by X(*n*,*m*), is decomposed into the object space, the variable space, and the error matrix. The latter represents the variation not explained by the extracted PC's and is dependent on the problem definition. The algorithm describing this decomposition is presented as

$$X(n,m) = T(n,k)P(k,m)^{T} + E(n,m)$$

where X is the independent variable matrix, T is the scores matrix, P is the loadings matrix, E is the error matrix, n is the number of objects, m is the number of variables, and k is the number of PC's used (*2*, *3*).

Partial Least Squares Regression1 (*PLS1*). PLS1 is a bilinear regression modeling method where the original x variables are projected onto a smaller number of PLS components (*3*). These components, also referred to as "latent variables" or "factors", are calculated according to the same mathematical procedures as PC's, but the data in the Y-matrix are incorporated in the calculation. The regression establishes the relationship between the X-matrix and the Y-matrix (in this study, the reference data for glycerol), with the objective to predict the y variables by using the most relevant PLS components. The relationship between the y and x variables can be described by the polynomial

$$y = b_0 + b_1x_1 + b_2x_2 + b_nx_n$$

where y is the dependent variable, b_0-b_n are the regression coefficients (b_0 is the intercept) and x_1-x_n represent the absorbance at the selected wavenumbers (see section Wavenumber selection).

SIMCA. Detection and classification of outlier samples were done using the SIMCA application of the Unscrambler Software (version 6.11, Camo ASA, Trondheim, Norway). Two training sets, "red wine" ($n = 30$) and "white wine" ($n = 30$), were modeled using separate PCA models. The samples used in the training sets were randomly selected. The test set consisted of the spectral outliers ($n = 6$). Class membership was defined at a significance level of 5%.

Wavenumber Selection. With the WineScan FT120 instrument, a maximum of 15 "filters" (wavenumbers or small groups of wavenumbers) can be defined for calibration purposes. The wavenumbers at which the correlation between the measured absorbance and the corresponding reference values for glycerol (as determined with the enzymatic method, see Reference Methods) was the highest were selected by using the Advanced Performance Software Module version 2.1.0, which is an extension of the basic software of the WineScan FT120 instrument. To exclude noise being introduced into the spectral data, only three regions, 964−1543 cm^{-1}, 1716−2732 cm^{-1}, and 2434−2970 cm^{-1} are made available for wavenumber selection. Possible overfitting of the calibration models (which introduces noise and uninformative variation into the calibration) was evaluated by deselecting the filters, explaining a very low percentage of variation in a sample set in a stepwise manner, using the standard error of cross validation (SECV) values as guide. Here the objective was to find the lowest number of filters and the lowest possible SECV value for a particular calibration model.

Evaluation of the Performance of Calibration Sets. The statistical indicators for evaluating the performance of the calibration models were calculated using the Advanced Performance Software Module provided with the WineScan FT120 instrument and included bias, SECV, and SEP. Cross validation was automatically done by the software and involved keeping out successive groups of samples from the calibration set (10% of the total number of calibration samples at a time), and using these subsets for prediction, until all samples have been kept out once. The selection strategy for cross validation purposes have been set by the manufacturer and are not accessible to change by the user of the instrument. Bias gives an indication of a systematic error in the predictive values (*3*), and it was calculated as the average of the residuals (residuals being the difference between the reference values and the predicted values). The accuracy of the predictive ability of the calibration model, relative to the reference data, was expressed as SECV when based on the calibration samples and using cross validation as explained before, and as SEP when based on independent validation sets. The calculations of these indicators are standard statistical procedures and several authors describe these procedures (*2*, *3*, *13*). The RPD value (ratio of the standard deviation of the data to the standard error of prediction) was used to evaluate the predictive ability of the calibration models (*19*, *20*). It has been proposed that an RPD value less than 3 indicates that the calibration model is unsuitable for quantification, a value between 3 and 5 is suitable for screening, while a value greater than 5 is suitable for quantification (*20*).

Reference Methods. *Glycerol Determinations*. Glycerol was assayed with the Boehringer Mannheim test kit (catalog number 148270). The total assay volume was scaled down to 100 μL, and duplicate determinations for each sample were carried out in microtiter plates (Sterilin, catalog number 612F96), and readings were taken at 340 nm, using a μQuant spectrophotometer (Bio-Tek Instruments, Winooski, VT). The accuracy of the reference method was expressed as the standard error of laboratory (SEL) and calculated as

$$\mathrm{SEL} = \sqrt{\frac{\sum (y_1 - y_2)^2}{2n}}$$

where y_1 and y_2 are the results of duplicate determinations and n is the number of samples.

Routine Wine Analysis. Reference data for alcohol, residual sugar (RS), titratable acidity (TA) and volatile acidity (VA) for the commercial wines were those officially approved by the South African Wine and Spirit Board upon certification of the wines. Corresponding data for the young wines were obtained by using the WineScan FT120 instrument and the commercial calibrations for these components (*13*).

RESULTS AND DISCUSSION

Analysis of FT-IR Spectra. A FT-IR spectrum of a wine provides the collective absorbance of all the IR-active components present in the sample. The contribution of the absorbance of water to the FT-IR spectra was significant in two wavenumber

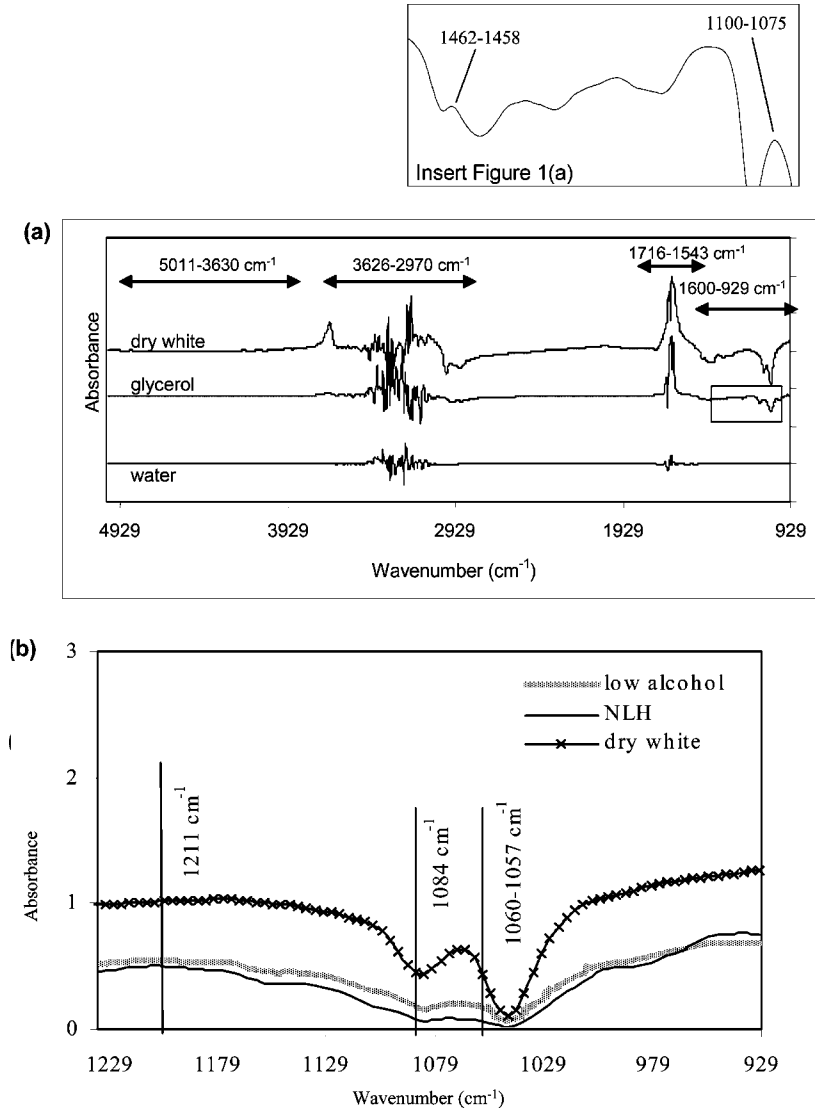

Figure 1. (a) FT-IR spectra of water, glycerol (10 g/L) and a dry white wine in the region 5011–929 cm^{-1}. Spectra were offset for clarity. The absorbance for all spectra recorded ranged from zero to a maximum value of 2.5. (b) Spectral variation between dry white, low alcohol, and noble late harvest (NLH) wines in the region 1229–929 cm^{-1}. Vertical lines indicate the wavenumbers, where collectively, more than 85% of the variation in the glycerol content of the samples was explained.

regions, 3626−2970 cm^{-1} and 1716−1543 cm^{-1}, respectively (**Figure 1a**). These regions were typically broad and covered several hundreds of wavenumbers. The absolute repeatability of the FT-IR spectra was poor in these regions (data not shown), and these areas are known to contribute considerable noise in the spectra. The FT-IR spectrum of the aqueous glycerol solution showed prominent absorbance peaks in the 1600−929 cm^{-1} region. Some characteristic features of the FT-IR spectrum of glycerol that were observed (see insert **Figure 1a**) included the peak at 1462−1458 cm^{-1} (representing the H−C bend) and the peak at 1100−1075 cm^{-1} (representing the C−O stretch) (*14*). The major infrared band associated with the −OH group is located at 3350 ± 50 cm^{-1} (*14*). In a wine matrix, where water is abundant, the absorbance of the −OH group would not be expected to be very useful for the purpose of developing a calibration for glycerol. Visual inspection of the FT-IR spectrum of a dry white wine indicated that the region from 5011 to 3630 cm^{-1} showed little variation in absorbance, whereas prominent peaks were present in the region from 1600 to 929 cm^{-1} (**Figure 1a**). The information in the latter region, referred to as the "fingerprint" area, is particularly useful in molecular absorption spectroscopy, because many different IR bands, including those corresponding to the vibrations of the C−O, C−C, C−H, and C−N bonds occur in this region (*1*).

Distinct variation between the FT-IR spectra of wines of various styles was observed in the region from 1229 to 929 cm^{-1}, as illustrated by the spectra of the dry white, low alcohol, and noble late harvest wines (**Figure 1b**). Collectively, more than 85% of the variation in the glycerol content of the wine samples could be correlated to the absorbance at the filters shown in **Figure 1b**. The absorbance at filters 1060−1057 cm^{-1} and 1084 cm^{-1} corresponded to the C−O stretch, respectively (*14*).

PCA Modeling. (*i*) *PCA Modeling of the Complete Data Matrix.* In the explorative stages of PCA, the complete data matrix, which included all the samples and all the wavenumbers, was modeled. The score plot of PC1 versus PC2 showed a distinct clustering of the samples that was related to wine style, as well as individual deviating samples and groups of deviating samples (**Figure 2a**). Group A consisted exclusively of two unwooded dry white wine styles ($n = 23$). The cluster located near the origin of the model (group B) consisted of the dry red, dry white, off-dry white, and young unbottled wines. Most of the young wines located at the extreme periphery of group B.

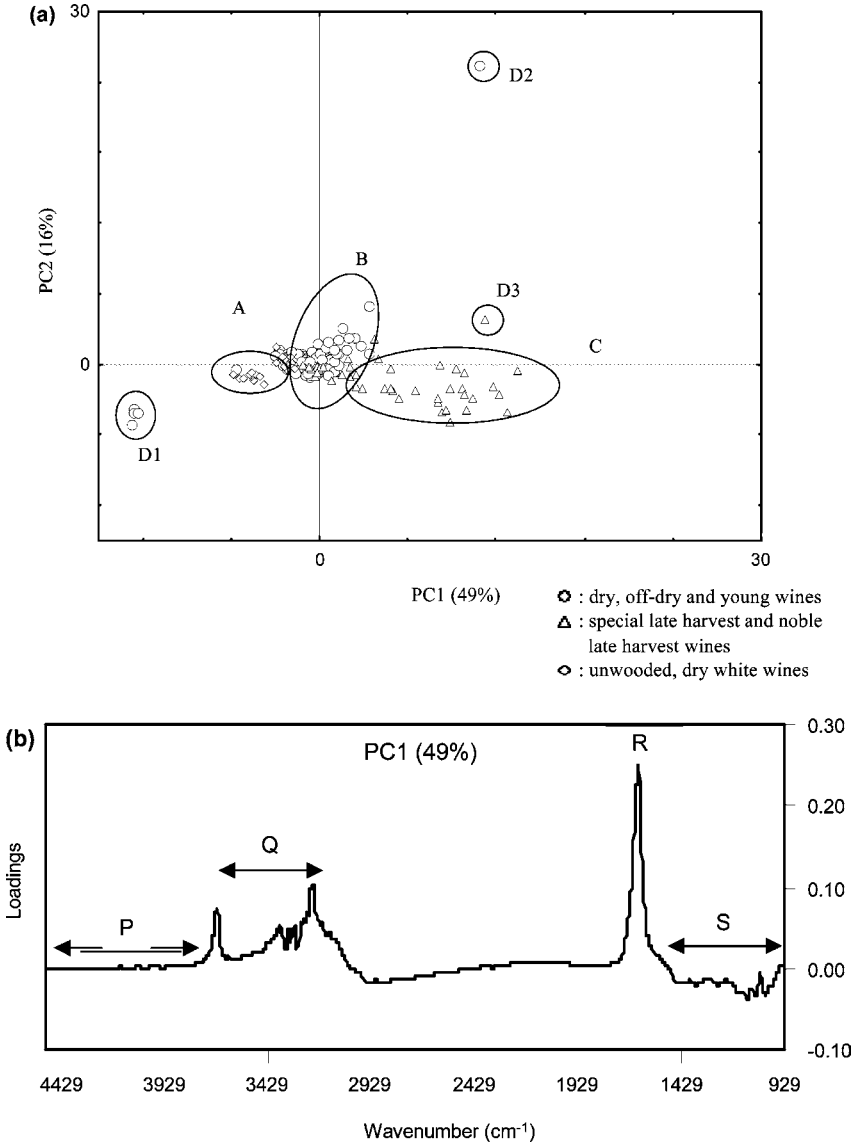

Figure 2. (a) PCA score plot, PC1 versus PC2, of the FT-IR spectra of the total set of wines ($n = 329$) and all the wavenumbers included. The model was centered, and the axes cross each other at the origin. (b) PC1 loadings plot. See text (PCA Modeling of Wine Samples) for the assignment of the symbols.

Group C located toward the positive end of PC1 as well as toward the negative end of PC2 and consisted exclusively of the special late harvest and noble late harvest wines. Six samples (groups D1, D2, and D3) were located far away from the other samples and clearly had spectra deviating considerably from the rest. An analysis of the leverage of these deviating spectra showed a significant influence on the PCA model (results not shown), and these samples were therefore considered to be true outliers. PC1 explained 49% of the variation in the sample set, which was not unexpected, because all the wavenumbers were included in the modeling and background noise due to the absorbance of water would be present. The PC1 loadings plot showed high loadings for wavenumbers where water is known to absorb (regions S and R, respectively, **Figure 2b**). Loadings were also observed in region S (1500−1000 cm^{-1}), which point to the contribution of the absorbance of several of the group wavenumbers of, among other components, alcohol, glycerol, sugars, and organic acids, which are known to absorb in this region, to PC1 (*1*). Region P showed very low loadings and confirmed the earlier interpretation that this region provided very little useful information and could contribute to noise in the spectra.

(ii) PCA Using Selected Wavenumbers. To evaluate the contribution of the major chemical components (other than water) to the variation between the different wine styles, subsequent PCA was done with the two wavenumber regions where water absorbs strongly (3626−2970 cm^{-1} and 1716−1543 cm^{-1}), and the region showing little useful information (5011−3630 cm^{-1}) was deselected. The outlier samples ($n = 6$) were also deselected from the original data matrix. Where appropriate, the interpretation of the score plots was based on the concentration ranges of the major chemical components of the wine samples (**Table 2**). PC1 (explaining 96% of the variation) seemed to distinguish between samples based on sugar content (**Figure 3a**). Samples in group A, which consisted almost exclusively of noble late harvest wines (RS levels ranging from 82 to 147 g/L; average 130.2 g/L, **Table 2**), appeared as a highly diverse and scattered group. Samples in group B consisted of special late harvest wines (RS levels ranging from 31 to 47 g/L, average 43.1 g/L). Group C consisted of commercial dry red and white wines, off-dry white wines, and the young wines (RS levels collectively ranging from 0.5 to 13 g/L). Some of the young wines and some of the older red wines that have undergone more than three years of maturation located

Table 2. Component Range of Calibration Samples[a]

style	glycerol (g/L)	alcohol (% v/v)	residual sugar (g/L)	volatile acidity (g/L)[b]	titratable acidity (g/L)[c]
dry white	6.81 ± 0.82	12.47 ± 0.71	2.34 ± 1.10	0.36 ± 0.13	6.01 ± 0.59
off-dry white	6.58 ± 0.77	12.48 ± 0.88	5.94 ± 1.99	0.28 ± 0.10	6.05 ± 0.46
low alcohol	3.47 ± 0.53	7.28 ± 0.31	2.04 ± 0.05	0.21 ± 0.06	5.43 ± 0.31
dry red	10.61 ± 1.08	12.93 ± 0.73	1.75 ± 0.69	0.54 ± 0.14	5.88 ± 0.36
SLH[e]	6.61 ± 0.87	11.62 ± 0.49	43.14 ± 8.12	0.33 ± 0.12	5.64 ± 0.50
NLH[f]	15.02 ± 3.82	12.93 ± 1.35	130.19 ± 24.95	0.87 ± 0.27	7.13 ± 0.67
Blanc De Noir	5.60 ± 0.46	12.39 ± 0.80	14.90 ± 9.79	na[d]	5.86 ± 0.95
young red wines	9.86 ± 0.93	11.01 ± 0.58	4.93 ± 1.17	0.40 ± 0.13	4.77 ± 0.53
young white wines	6.43 ± 0.57	11.08 ± 1.30	4.70 ± 1.34	0.37 ± 0.17	5.61 ± 0.56

[a] Values given are average ± standard deviation. [b] Expressed as g/L acetic acid. [c] Expressed as g/L tartaric acid. [d] Not available. [e] Special late harvest wines. [f] Noble late harvest wines.

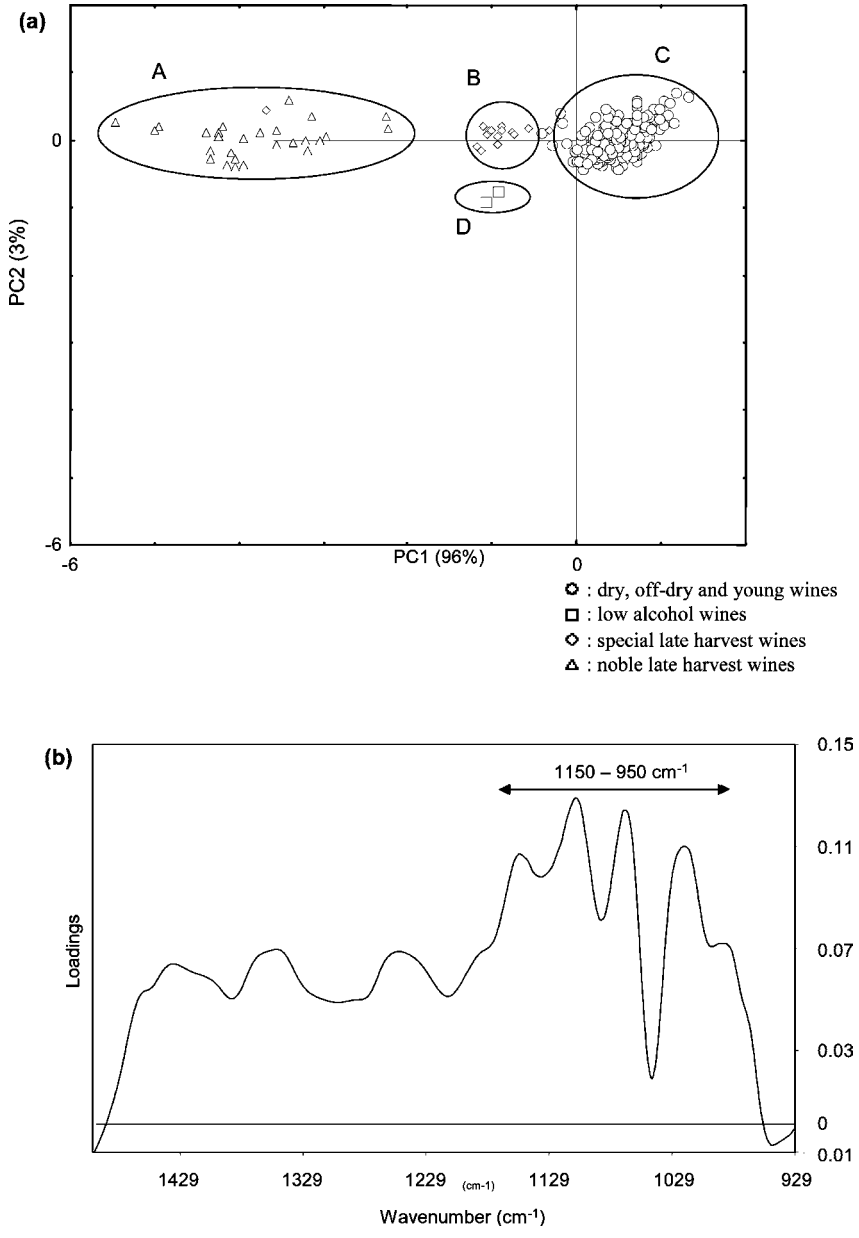

Figure 3. (a) PCA score plot, PC1 versus PC2, of the wine spectra, with the outliers ($n = 6$) removed and with the wavenumber regions 5011–2970 and 1716–1543 cm^{-1} deselected. The model was centered, and the axes cross each other at the origin. (b) PC1 loadings plot. See text (PCA Modeling of Wine Samples) for the interpretation of the symbols.

on the extreme periphery of group C (**Figure 3a**). The low alcohol wines (less than 8% v/v alcohol) located far away from the rest of the samples and PC2, explaining 3% of the remaining variation, seemed to distinguish between samples based on the alcohol content. These results confirmed the earlier observation that both the RS and alcohol levels were major sources of variation between the different styles and also suggested that the late harvest wines would have to be treated as a separate

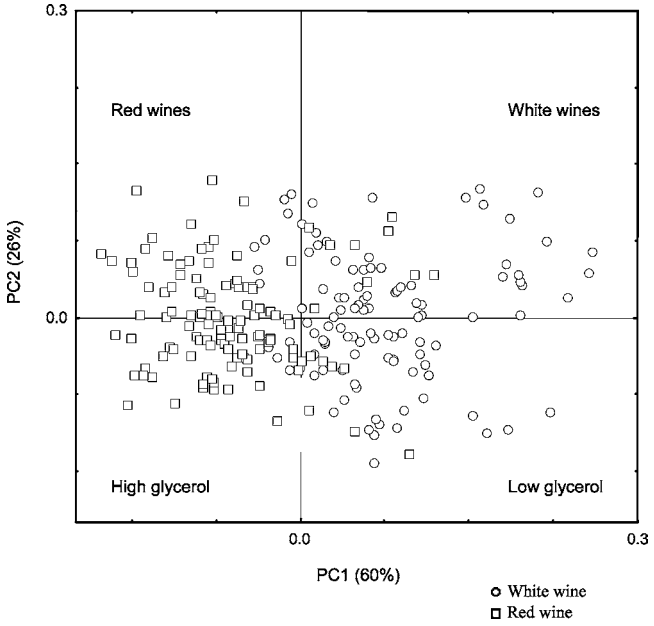

Figure 4. PCA score plot, PC1 versus PC2, of the FT-IR spectra of dry wines ($n = 284$; RS levels < 30 g/L). The 15 most pertinent wavenumbers, explaining more than 98% of the variation in the glycerol content of the samples, were used for the modeling.

group in the design of sample sets for glycerol calibration. The PC1 loadings plot (**Figure 3b**), showed particularly high loadings for several wavenumbers in the region from 1150 to 950 cm^{-1}, which confirmed the interpretation of the score plot in **Figure 3a**, but also pointed to the need to extract the most pertinent wavenumbers related to the variation in the glycerol content for the purpose of calibration. The separation of the samples based on the sugar content was not surprising, because the largest variation in the chemical composition between the different wine styles was seen in the RS levels, both in terms of the range in concentrations and the standard deviation within each style (**Table 2**).

(iii) PCA Based on the Glycerol Content of Wine Samples. To model the relationships between the wine samples on the basis of their glycerol content, PCA was done using only the 15 filters that collectively explained more than 98% of the variation in the glycerol content of the samples. The loadings for the 15 wavenumbers selected were high (loadings plot not shown). The separation of samples in the score plot of PC1 vs PC2 (**Figure 4**) could be interpreted as follows: (i) the samples clearly separated on the basis of their glycerol content, as would be expected, and samples with the highest glycerol levels located toward the negative end of PC1; (ii) the red and white wines partially separated in two groups, although with overlap between the two groups; (iii) the older red wines that have undergone more than three years of maturation as well as some of the young unbottled wines appeared on the "extreme" periphery of the clusters shown in **Figure 4** and seemed to be spectrally different from the rest of the samples. These results provided preliminary evidence that both the maturation period of the wines, as well as the stage of fermentation, should be taken into consideration in the design of calibration sets and that these sources of variation could have implications for the accuracy and robustness of prediction.

Design of Calibration Sets for Glycerol in Wine. The initial strategy used attempted to establish a single global calibration model for glycerol in all the wines (excluding the six outlier samples) in the sample set. For this purpose, the wine spectra were split in two equal sets, a calibration set and a validation set, respectively. The selection of samples was random and aimed at keeping the two sets balanced in term of chemical composition. A variable selection was performed on the calibration set ($n = 162$) and used to establish a calibration model. Samples that were poorly predicted included the low alcohol wine ($n = 1$, SEP > 2% v/v), several of the special late harvest and noble late harvest wines ($n = 10$, SEP > 0.8 g/L) and some of the young red wines ($n = 6$, SEP > 0.9 g/L). The poorly predicted samples ($n = 17$) were deselected from the calibration set and a new variable selection performed on the calibration set. The validation of this model provided a SEP value of 0.62 g/L, which was considered unsatisfactory. The customary approach of subdividing the original calibration set into subsets (and keeping the same internal subdivision between the calibration set and the validation set), in an attempt to improve the SEP, was not further pursued because this procedure resulted in subdivision of the special and late harvest wines in groups that were so small in number that it was not possible to develop a calibration model for these styles. Furthermore, the clustering of the wines according to the wine style (as observed with PCA) clearly suggested a subdivision according to wine style to establish calibration models with acceptable prediction accuracies. Using this approach, the strategy in the design of calibration sets was aimed at low SECV or SEP values, but at the same time keeping the number of calibrations as small as possible. In the exploratory stages of calibration a "cutoff" level of ca. 8% v/v alcohol was used to differentiate the "low alcohol" wines from the wines with higher alcohol levels and a RS level of ca. 30 g/L to differentiate between samples used for the "wine calibration" and a "sweet wine calibration", respectively. The former set (RS < 30 g/L) consisted of dry red, dry white, off-dry white, and Blanc De Noir wine styles, as well as the young wines. The "sweet wine calibration" set (RS 31−147 g/L) consisted of the special late harvest and noble late harvest styles (**Table 1**).

Glycerol Calibration for Wine (RS < 3.0 g/L, Alcohol > 11.6% v/v). No significant gain in terms of accuracy for glycerol prediction was obtained by separating the red and white wines into different calibration sets, whereas the inclusion of the young wines in the original calibration set resulted in an increase in the SECV (from 0.38 g/L to 0.52 g/L). On the basis of these results, the red and white wines were not separated for the final design of the calibration set, but the young unbottled wines were removed from the original sample set. Decisions regarding the design of calibration sets should, however, clearly be made in the context of a particular application. For the quantification of alcohol, for instance, where very high levels of accuracy are required, preliminary results indicated that the separation of white wines and red wines into different calibration sets improved the accuracy (data not shown). The final calibration set ($n = 135$) spanned the glycerol concentration range in the original sample set and also contained the "extreme" spectral members (samples that located toward the extreme periphery of the clusters observed by PCA modeling). Samples that were unusual in terms of their geographic origin, and some of the older red wines were also included in the calibration set.

Filters ($n = 15$) were selected that collectively explained more than 98% of the accumulated variation in the glycerol content of the calibration set. The SECV for the calibration set was 0.38 g/L (**Table 3**). An independent validation set ($n = 98$) was used to test the predictive accuracy of the calibration model. These samples were selected to span the glycerol concentration range, over which predictions in future samples had to be done.

Table 3. Regression Statistics for Glycerol Calibration and Validation in Dry Wine

(a) calibration set		(b) validation set	
number of factors	7	bias	0.02
number of filters	15	*r*	0.96
number of samples	135	number of samples	98
SECV[a] (g/L)	0.38	SEP[a] (g/L)	0.40
AR[a]	0.08	AR[a]	0.08
glycerol range (g/L)[b]	4.74–14.00	glycerol range (g/L)[b]	5.46–13.40
average glycerol (g/L)[b]	8.71	average glycerol (g/L)[b]	8.92
SEL[a] (g/L)	0.30		

[a] Abbreviations used: SECV, standard error of cross validation; SEP, standard error of prediction; AR, absolute repeatability; SEL, standard error of laboratory. [b] As determined by the reference method.

The correlation between the glycerol concentrations that were determined with the reference method and the values predicted by the glycerol calibration is shown in **Figure 5**. The SEP value (0.40 g/L, **Table 3**) was in agreement with the error for the reference method for glycerol, SEL = 0.30 g/L. The RPD value for the glycerol calibration model was 5.6 (SD of the data set = 2.22), indicating that the calibration was suitable for quantification purposes (*19*). SEP values of 1.32 g/L (*6*) and 1.13 g/L (*12*) have been reported for the quantification of glycerol in dry wine. In comparison, the glycerol prediction with the calibration established in this study showed an improvement in accuracy. The establishment of a glycerol calibration with SEP = 0.49 g/L has also been reported, but the predictive ability was found to be highly dependent on the sample set used for the validation (*5*).

The error of prediction for glycerol in the red wines that have undergone more than 3 years of aging was in excess of 0.6 g/L for some of the older wines, and these samples were clearly predicted less accurately by the model. This result was not surprising in view of the tendency of these samples to locate toward the extreme periphery of the red wine cluster in the PCA score plot (**Figure 4**). Complex changes occur in the chemistry of red wines during aging, particularly due to the polymerization and condensation of the tannins. Recent research on wines that have been subjected to different aging regimes showed changes in the chemical composition of tannins that were reflected in the FT-IR spectra of the samples (*16*).

Glycerol Determination in Low Alcohol Wines (8% v/v Alcohol). The initial calibration strategy that was aimed at developing a global calibration model for all the wine styles used in this study was not satisfactory, because the SEP value for the low alcohol wines was > 2% v/v. The sample number of this wine style was very small ($n = 2$), and future efforts will be directed toward the enlargement of the sample set. More work is also required to fully characterize these wines in terms of their spectral properties in the IR range. Very little information on the application of FT-IR for the analysis of low alcohol wines has been reported in the literature. In one study, wines with alcohol concentrations of 8.5% v/v were included in a calibration set designed for use with FT-IR, but no information was provided on the specific prediction error for these samples (*5*).

Glycerol Calibration for Sweet Wines (RS 31−147 g/L). Due to the large spectral variation observed by PCA in the sweet wines (special late harvest and noble late harvest wines) and the relatively small sample size ($n = 43$, **Table 1**), the full sample set was used for calibration purposes. This produced a SECV of 0.65 g/L (**Table 4**) when evaluated by cross validation as described before. The correlation between the glycerol concentrations that were determined with the reference method and those predicted by the glycerol calibration for sweet wine is shown in **Figure 6**. In a recent study where near-infrared reflectance spectroscopy was used for the simultaneous determination of alcohol, glycerol, glucose, and fructose in botrytized sweet wines, the highest error in the prediction results was found for the estimation of glycerol (*17*).

Table 4. Regression Statistics for Glycerol Calibration in Sweet Wine

number of factors	9
number of filters	15
SECV[a] (g/L)	0.65
AR[a]	0.09
number of samples	43
glycerol range (g/L)[b]	4.74–14.00
average glycerol (g/L)[b]	8.71

[a] Abbreviations used: SECV, standard error of cross validation; AR, absolute repeatability. [b] As determined by the reference method.

Glycerol Determination in Young Wines. The young wines were treated as a separate validation set of the glycerol calibration established for the wines with RS < 30 g/L, and this resulted in an SEP of 0.85 g/L. A better fit of the data set was obtained by adjusting the intercept of the original calibration, and using this strategy, an SEP of 0.43 g/L was obtained. The average levels of the components listed in **Table 2** did not appear to be significantly different between the young wines and bottled, commercial dry red, or white wines. It is to be expected that the stage of the fermentation would have a major influence on the spectral properties of the samples. In this respect, the CO_2 levels as well as the stage of the malolactic fermentation have been shown to influence the accuracy of quantification of various components using FT-IR spectroscopy (*6*). This is, however, clearly a situation that should be evaluated for each specific sample set, and the results presented here merely serve the purpose of illustrating that the prediction accuracies for the young wines need monitoring and may, in some instances, require additional validation.

Interpretation and Classification of Outlier Samples. In the exploratory stages of PCA, all the wavenumbers were included in the modeling en lieu of selectivity, and six extreme outlier samples were identified. These samples (all commercial, bottled wines) were poorly described by the PCA model and did not appear to belong to any of the major groups shown in **Figure 2a**. The glycerol estimations for these samples, using the calibrations established in this study, were also poor, and the SEP values were in excess of 0.8 g/L. In this study, several strategies were used to interpret the outlier status of these samples, including (i) a comparison of the component ranges of the outlier samples to that of similar samples in the sample set; (ii) an examination of the statistics (mean, maximum, minimum, and SD) of absorbance over the entire wavenumber range; and (iii) visual inspection of the spectra. The component ranges of the outlier samples, as compared to that of similar types of wines, did not reveal any obvious unusual features. The outlier status of two of these samples could be ascribed to poor repeatability (as judged by a large SD between the absorbance at some wavenumbers of replicate scans), which could be an indication of poor sample quality or inhomogeneity in the sample. In the other instances, however, the spectra of the outlier samples were markedly different from that of similar wines and the atypical nature of the spectra was confirmed with repeated scanning of the same wine sample.

For the early detection and classification of the spectral outliers, the SIMCA application of the Unscrambler Software

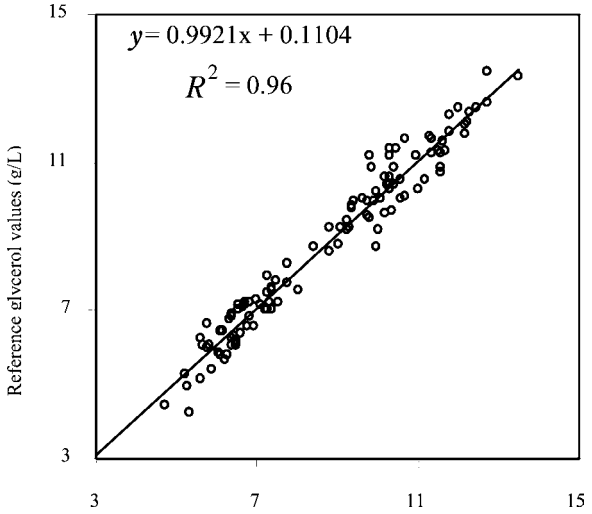

Figure 5. Regression plot of the comparison between the glycerol content of an independent validation set (n = 98; RS levels < 30 g/L) as determined with the reference method and predicted with the WineScan FT120 instrument.

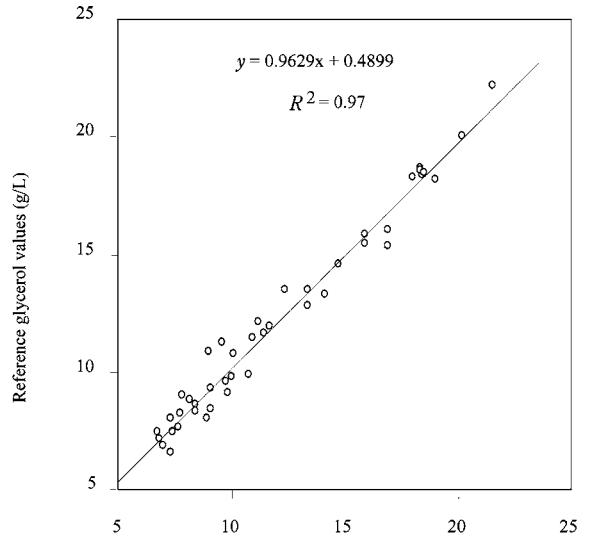

Figure 6. Regression plot of the comparison between the glycerol content of a set of noble late harvest and special late harvest wines (n = 43; RS levels ranging from 31 to 147 g/L), as determined with the reference method and predicted with the WineScan FT120 instrument.

was used to make two disjoint PCA class-models for "white wine" and "red wine". The class-membership of a test set containing the outliers (n = 6) was tested at a significance level of 5%. The classification results are graphically presented in **Figure 7**, where the area below the horizontal line delimits membership of the "white wine" model and the area to the left of the vertical line delimits membership of the "red wine" model. The area near the origin of the plot delimits samples showing membership to both models (*3*). Results showed a 100% nonmembership to both models for the outlier samples. The outlier samples were not investigated further, and future work will be aimed at enlarging the database of outlier spectra and to establish a discriminatory PCA calibration on the WineScan FT120 instrument, to provide a conformity test at the time of analysis, as well as a warning of suspected outlier samples. Such samples should then automatically be tested with appropriate reference methods and subjected to further investigation for the purposes of quality control. Recently, a quality assurance software module was made commercially available (*18*), and

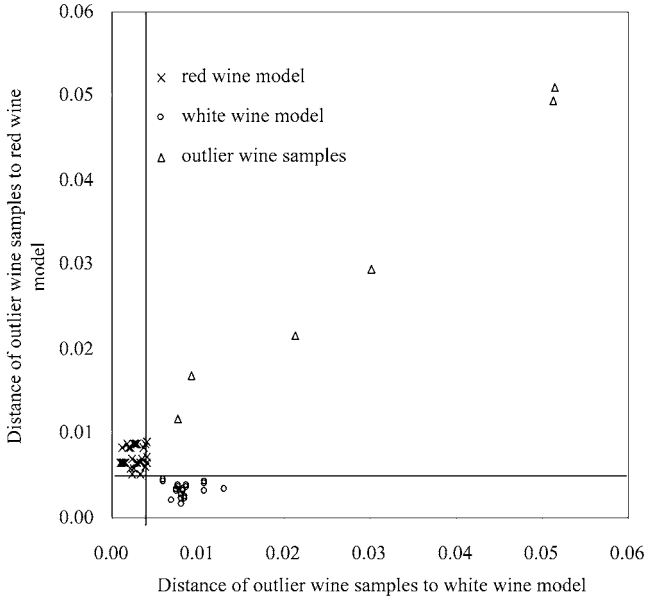

Figure 7. Coomans plot showing the distances of the outlier samples (n = 6) to the selected models for "red wine" (n = 30) and "white wine" (n = 30), respectively.

this application can be used as a basis from which to develop customized calibration models for the purpose of quality control.

From a spectroscopic perspective, wine is a challenging matrix both in terms of its chemical complexity, as well as in the variation introduced in the spectra by factors such as style, process technology, cultivars, and geographic origin. For the purposes of quality control in the analytical laboratory, the outlier samples are important and require special attention. Due to the atypical nature of the spectral properties of these samples, it is to be expected that the accuracy of prediction in these samples will be unsatisfactory. Furthermore, if the cause for the outlier status of these samples could be interpreted, the appropriate action could be taken, and the decisions on how to handle similar future samples would be more informed.

This study has shown that PCA provides a powerful tool to identify the major sources of variation in the FT-IR spectra of wine samples. The sources of variation identified in this study included the sugar and alcohol content of the samples, the stage of the fermentation process, and the maturation period of the wines. The implications of this variation for the accuracy of prediction of calibration models were evaluated using glycerol calibration as a model system and clearly showed that calibration sets have to be carefully selected in order to design calibration models that find a balance between robustness and accuracy of prediction. PCA of the FT-IR spectra also facilitated the early detection and classification of poorly predicted samples, as well as a small number of extreme outlier samples in the sample set. It is our opinion that the successful implementation of FT-IR for the routine analysis of wine requires an approach that combines the development of robust calibration models, as well as the implementation of quality control measures (such as PCA calibrations or SIMCA models) to enable the early detection and interpretation of poorly predicted samples and outlier samples. The latter aspect clearly also involves data management, specifically in terms of the interpretation of the reasons for the poor predictions or outlier status of deviating samples.

ABBREVIATIONS USED

SECV, standard error of cross validation; SEL, standard error of laboratory; SEP, standard error of prediction; AR, absolute

repeatability; RPD, ratio of standard deviation of data to standard error of prediction; PCA, principal component analysis; PLS, partial least squares; RS, reducing sugar.

LITERATURE CITED

(1) Smith, B. *Infrared Spectral Interpretation*: *A Systematic Approach*, 1st ed.; CRC Press LLC: Boca Raton, FL, 1999.

(2) Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Wold, S. *Introduction to Multi- and Megavariate Data Analysis using Projection Methods* (*PCA and PLS*), 1st ed.; Umetrics AB: Umeå, Sweden, 1999.

(3) Esbensen, K. H. *Multivariate Data Analysis - In Practise*, 4th ed.; Camo ASA: Oslo, Norway, 2000.

(4) Skoog, D. A.; Holler, F. J.; Nieman, T. A. *Principles of Instrumental Analysis*, 5th ed.; Harcourt Brace College Publishers: Philadelphia, PA, 1997.

(5) Patz, C.-D.; David, A.; Thente, K.; Kürbel, P.; Dietrich, H. Wine analysis with FTIR spectrometry. *Vitic. Enol. Sci.* **1999**, *54*, 80−87.

(6) Dubernet, M.; Dubernet, M. Utilisation de l'analyse infrarouge a transformee de Fourier pour l'analyse oenologique de routine. *Rev. Fr. Œnol.* **2000**, *181*, 10−13.

(7) Gishen, M.; Holdstock, M. Preliminary evaluation of the performance of the Foss WineScan FT120 instrument for the simultaneous determination of several wine analyses. *Aust. Grapegrower Winemaker* **2000**, *Ann. Technol. Issue*, 75−81.

(8) Kupina, S. A.; Shrikhande, A. J. Evaluation of a Fourier transform infrared instrument for rapid quality-control wine analyses. *Am. J. Enol. Vitic.* **2003**, *54*, 131−134.

(9) Würdig, G.; Woller, R. *Chemie des Weines*. Stuttgart: Ulmer, 1989.

(10) Ribéreau-Gayon, J., Glories, Y., Maujean, A. and Dubourdieu, D. *Handbook of Enology. The Microbiology of Wine and Vinifications.* 1st ed.; John Wiley and Sons: New York, Toronto, 1998; Vol. 1.

(11) Zoecklein, B. W.; Williams, J. M.; Duncan, S. E. The effect of sour rot on the composition of White Riesling (*Vitis Vinifera L.*) grapes. *Small Fruits Rev.* **2000**, *1*, 63−77.

(12) Application Note 191 Issue 2 GB, P/N 1025415, 2001. *WineScan Calibration Finished Wine-Glycerol*. Foss Electric, Denmark.

(13) WineScan FT120 Type 77110 and 77310 Reference Manual, Issue 4 GB Foss Electric, Denmark, 2001. http://www.foss.dk.

(14) Pavia, D. L.; Lampman, G. M.; Kriz, G. S. *Introduction to Spectroscopy*, 3rd ed.; Harcourt College Publishers: Philadelphia, PA, 2001.

(15) Coimbra, M. A.; Barros, A.; Barros, M.; Rutledge, D. N.; Delgadillo, I. Multivariate analysis of uronic acid and neutral sugars in whole pectic samples by FT-IR spectroscopy. *Carbohydr. Polym.* **1998**, *37*, 241−248.

(16) Edelmann, A.; Lendl, B. Toward the optical tongue: flow-trough sensing of tannin-protein interactions based on FTIR spectroscopy. *J. Am. Chem. Soc.* **2002**, *124*, 14741−14747.

(17) Garcia-Jares, C. M.; Médina, B. Application of multivariate calibration to the simultaneous routine determination of ethanol, glycerol, fructose, glucose, and total residual sugars in botrytized-grape sweet wines by means of near-infrared reflectance spectroscopy. *J. Anal. Chem.* **1997**, *357*, 86−91.

(18) *In Focus* Volume 2/18 2003. New Quality Assurance Module for the WineScan FT120. http://www.foss.dk.

(19) Pink, J.; Naczk, M., Pink, D. Evaluation of the quality of frozen minced red hake: use of Fourier transform infrared spectroscopy. *J. Agric. Food Chem.* **1998**, *46*, 3667−3672.

(20) Williams, P. Near-infrared technology implementation. Canadian Grain Commission, Grain Research Laboratory: Winnipeg, MB. 1995.

Received for review December 5, 2003. Revised manuscript received April 2, 2004. Accepted April 11, 2004. Financial support provided by the National Research Foundation (Core and THRIP programs) and the South African wine industry (Winetech).

JF035431Q