JOURNAL OF
AGRICULTURAL AND FOOD CHEMISTRY

Development of a Rapid "Fingerprinting" System for Wine Authenticity by Mid-infrared Spectroscopy

Christopher J. Bevin,*,† Allison J. Fergusson,† Wade B. Perry,†
Leslie J. Janik,‡ and Daniel Cozzolino‡

Hardy Wine Company, Reynell Road, Reynella, SA 5161, Australia, and The Australian Wine
Research Institute, Post Office Box 197, Adelaide, Glen Osmond, SA 5064, Australia

This paper reports on the development of a rapid and simple method for red wine authenticity confirmation during transport and processing; namely, a wine "fingerprinting" system. When wine is transported between two sites, a sample is taken and a mid-infrared (MIR) spectrum is obtained. One hundred sixty-one ($n = 161$) samples of three main red wine varieties grown in Australia, Shiraz, Cabernet Sauvignon, and Merlot, were collected from six commercial wineries across Australia and scanned in transmission on two MIR spectrophotometers located at The Hardy Wine Company's main site at Reynella, South Australia (Foss WineScan FT 120) (926−5012 cm^{-1}). A similarity index (*SI*) method was used as a tool to classify wine samples on the basis of their spectral data. The results showed that high rates of classification were obtained when wine samples scanned in different instruments were analyzed. The *SI* has been proven to provide an acceptable measurement for authentication of red wine integrity during transportation. In five of the six winery data sets, the *SI* correctly classified 98% of the wines. It was also observed that less than 1% of wines were misclassified between the different wineries investigated. Further studies are needed in order to test the applicability of the *SI* in a commercial situation and to evaluate its potential as a rapid quality control tool for routine use to authenticate wine samples during transport.

KEYWORDS: Mid-infrared; MIR; similarity index; red wine; classification; fingerprinting

INTRODUCTION

Wine is routinely transported to bottling and packaging facilities and transported between wineries for blending purposes. Presently, there is no recognized method available to easily monitor wine authenticity before and after transportation. Refractive index (RI) is commonly used as an indicator of wine dilution; however, it is not recognized as a technique to authenticate wine samples (*1, 2*). Analytical control methods in an industrial environment, whether qualitative or quantitative, are essential in order to assess raw materials, products, and byproducts as well as to optimize the manufacturing process itself (*3*). Conventional chemical methods of wine analysis involve time-consuming, laborious and costly procedures (*4−6*). Therefore a robust, rapid, and inexpensive method for quality assurance purposes is needed in the wine industry to monitor that wine parameters conform to specification, in order to ensure the quality of the final product delivered to the consumer. Vibrational spectroscopy in the infrared region has been applied to rapid qualitative analyses for a wide variety of products (*7, 8*), and applications have recently been reported in wine (*9, 10*). It is well-known that visual examination of the infrared spectra

is subjective and often cannot discriminate between authentic and adulterated product (*11*). Therefore the application of multivariate data analysis techniques like principal component analysis (PCA) or discriminant analysis (DA) opens the possibility to unravel and interpret the spectral properties of the sample and allow a classification without the use of direct chemical compositional information (*11, 12*). One of the advantages of the spectroscopic technology is that it allows the assessment of chemical structures through analysis of the molecular bonds in the mid-infrared (MIR) and near-infrared (NIR) spectrum and also builds a characteristic spectrum that represents a "fingerprint" of the sample. The mid-infrared region typically contains the sharp fundamental rotational molecular vibrations, which provide important functional group information, while the broader overtone and combination vibrations are found in the near-infrared region. This opens the possibility of using spectra to determine complex attributes of organic structures in the sample, which are related to molecular chromophores, organoleptic scores, and sensory characteristics (*13*).

This study was established to develop a rapid and simple method for wine authenticity confirmation; namely, a wine "fingerprinting" system. When a wine is being dispatched for transportation to another processing site, a sample is taken and an infrared spectrum is obtained. When the wine arrives at its

* Corresponding author: fax +61 8 8392 2191; e-mail chris.bevin@hardywines.com.au.
† Hardy Wine Company.
‡ The Australian Wine Research Institute.

Table 1. Mean (Standard Deviation) of Chemical Data for Wines from Each Winery

winery	alcohol, % v/v	pH	TA[a] to pH 8.2, g L^{-1}	specific gravity (gcm^{-3})	N samples
A	13.67 (0.659)	3.41 (0.073)	6.20 (0.324)	0.9942 (0.00070)	32
B	13.64 (0.578)	3.49 (0.055)	6.18 (0.377)	0.9943 (0.00049)	27
C	14.04 (0.375)	3.51 (0.047)	6.02 (0.266)	0.9943 (0.00053)	26
D	14.42 (0.539)	3.42 (0.046)	6.87 (0.251)	0.9940 (0.00082)	26
E	14.27 (0.642)	3.41 (0.048)	6.53 (0.274)	0.9937 (0.00066)	20
F	13.94 (0.435)	3.37 (0.062)	6.51 (0.491)	0.9942 (0.00067)	30
overall	13.97 (0.617)	3.43 (0.075)	6.38 (0.431)	0.9941 (0.00068)	161
range	12.13–15.94	3.20–3.63	5.34–8.17	0.9921–0.9962	

[a] TA, titratable acidity.
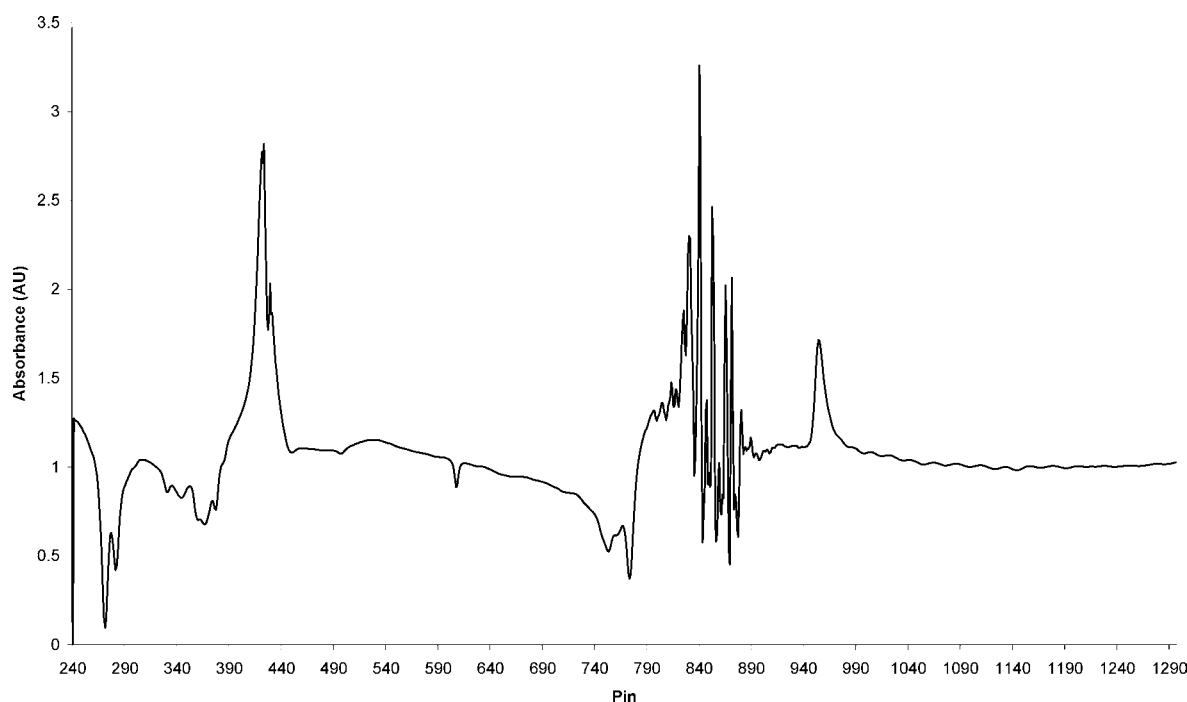
Figure 1. Example of a typical mid-infrared spectrum of wine.

final destination, a second sample from the same wine is taken and a new infrared spectrum is acquired on an instrument of spectral similarity to the first. The two spectra are then compared to confirm the authenticity of the wine or to observe changes that would have occurred during transport. The advantage of using infrared spectra is that they represent the combined absorbance of all components in the wine matrix, rather than simply the composition of a few chemical compounds. The two main criteria for the development of a system were (i) the system developed had to be easily understood by various operators in a winery from a diverse skill base and (ii) the system had to be able to operate using readily available equipment and inexpensive and commonly used software.

A number of qualitative chemometric techniques such as soft independent modeling of class analogies (SIMCA) (14), Mahalanobis distance (15), and artificial neural networks (ANN) (16) have been shown to be able to classify foods on the basis of spectral data. However, these advanced qualitative methods can be difficult to understand and difficult to apply under industrial conditions. A simple approach for comparing two spectra is the "similarity index" method, as described by Coene et al. (17).

This paper details the application of the similarity index (SI) combined with mid-infrared (MIR) spectra in order to develop a rapid and inexpensive authenticity tool for monitoring wines during transportation and processing for the wine industry.

MATERIALS AND METHODS

Wine Samples. One hundred sixty-one ($n = 161$) wine samples of the three main red wine grape varieties grown in Australia, Shiraz ($n = 65$), Cabernet Sauvignon ($n = 58$), and Merlot ($n = 38$), were collected from different wineries across Australia belonging to the Hardy Wine Company. The samples used in this study were of the same vintage (2005), had completed malolactic fermentation, and had not been blended with any other variety or wine from another region. As such, these samples could not be considered commercial wines but do represent diversity in wine styles and qualities related to a specific region or winery. **Table 1** details the main chemical characteristics as measured by the MIR instrument for the wine samples from each winery.

Instrumentation. Wine samples were scanned in transmission on two similar MIR spectrophotometers (Foss WineScan FT 120; Foss, Hillerød, Denmark) using the range between wavenumbers 926 and 5012 cm^{-1} (instrument manufacturer "pin numbers" 240−1299). The WineScan FT 120 is capable of the simultaneous prediction of numerous wine components in approximately 40 s, a significant advantage over traditional methods of analysis. The spectrophotometers were configured to acquire two replicate spectra for every sample. Two Hardy Wine Company instruments, located at Reynella (South Australia), were used in the study. The instrument manufacturer manual (18) recommends that pin numbers 250−400, 445−589, and 631−770 be used for calibration development. This excludes the water absorption peaks in the regions between pin numbers 400 and 445 and 770 and 940 (1543−1717 and 2971−3627 cm^{-1}) as well as carbon dioxide, which has a fundamental absorption peak between pin nymbers 590 and 630 pin

Rapid Wine "Fingerprinting" by Mid-infrared Spectroscopy

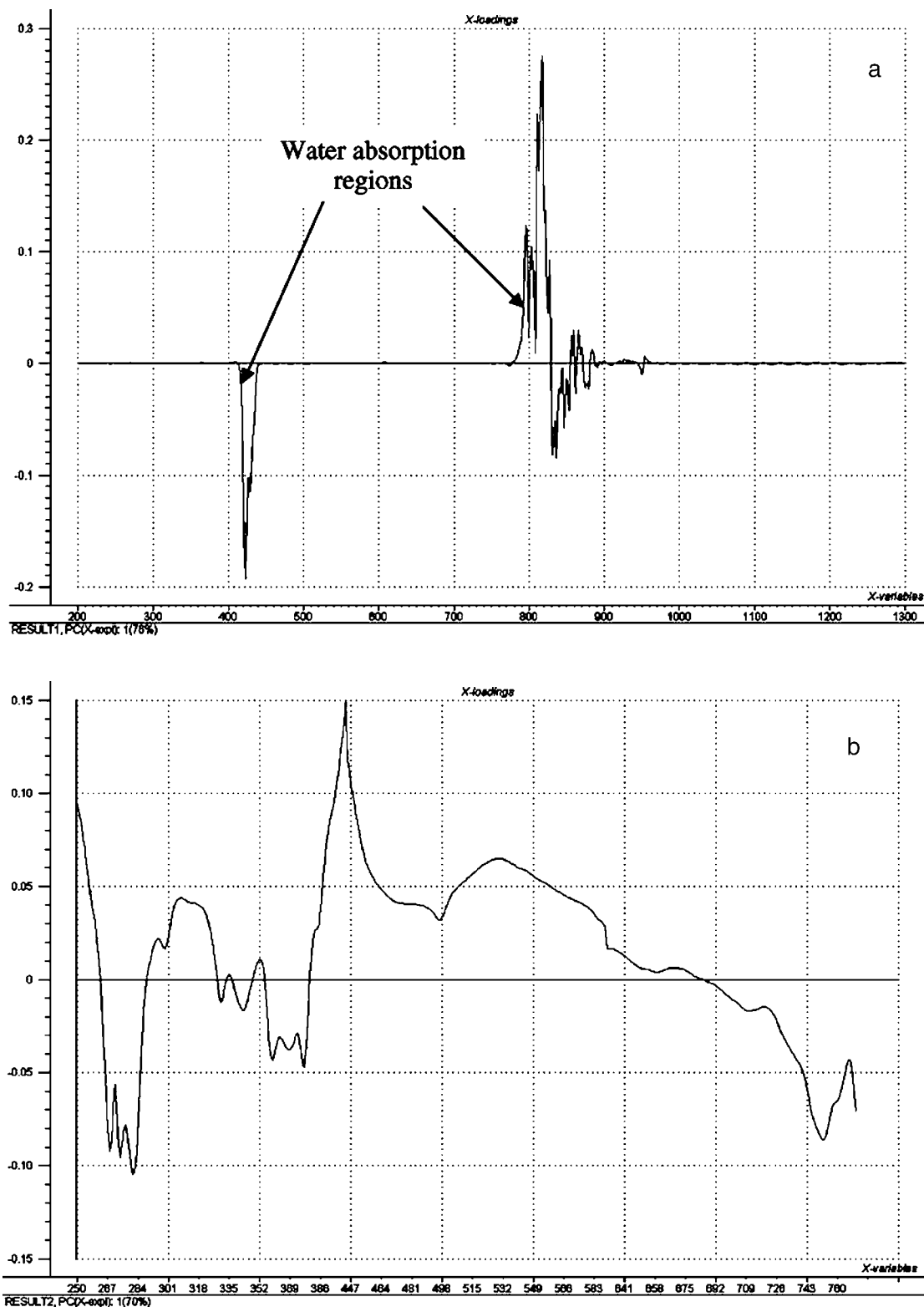*J. Agric. Food Chem.,* Vol. 54, No. 26, 2006   **9715**

Figure 2. (**a**) Eigenvectors of PC1 across all pin numbers, showing the influence of water peaks in the IR spectrum. (**b**) Eigenvectors of PC1 with instrument manufacturer-recommended pin numbers.

numbers (2276−2431 cm^{-1}). The effect of including all pins in the *SI* calculation was investigated; however, for the reported *SI* results only the manufacturer's recommended pins were used.

Data Processing and Statistical Analysis. Spectra were exported and reformatted in Microsoft Excel to enable basic data analysis, then imported into The Unscrambler software (version 9.2, Camo AS, Nedre Vollgate, Norway) for principal component analysis (PCA). To test between-instrument reproducibility in this study, duplicate spectra were

averaged to provide a sample spectrum for each instrument. PCA was used to reduce the dimensionality of the data to a small number of components, to examine the possible grouping of samples, and to visualize the presence of outliers (*19*, *20*). The number of latent variables (principal components) used in order to develop the PCA models were selected by the PRESS (predicted residual error sum-of-squares) function. Full internal cross-validation (leave one out) was used when PCA models were developed. Two available methods were
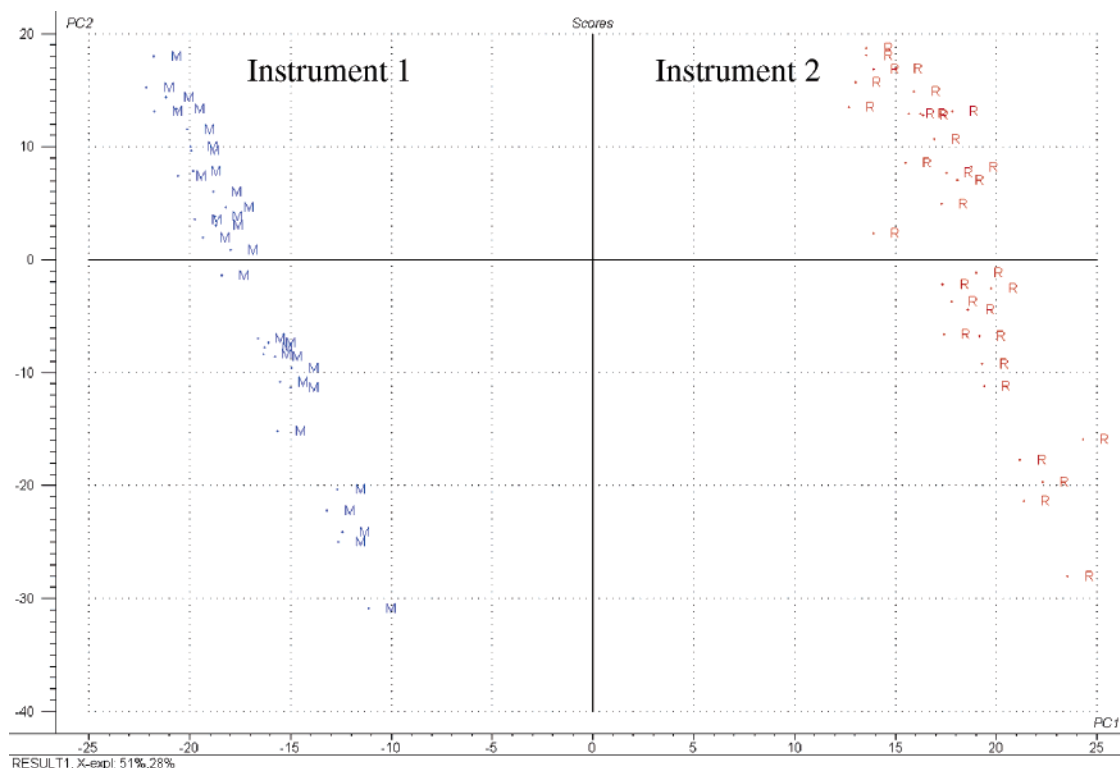
Figure 3. Score plot for the entire spectral region for 32 samples scanned on two WineScan instruments (instrument 1 = M; instrument 2 = R).

investigated as a tool to classify wines, namely, the similarity index (*SI*) method (*17*, *21*) and the polar qualification system (PQS) (*22*). In this study, only the results of applying *SI* as a classification method were reported.

Similarity Index. The similarity index was considered to be the most appropriate method to use, as it fulfils each of the criteria specified in this study. Indeed, it is specifically targeted at such an application, whereby only two spectra are being compared:

$$SI = \frac{1}{(1 - r^2)} \tag{1}$$

where r^2 is the coefficient of determination between the absorbance values from the two spectra at each wavelength. This can be easily determined by use of the RSQ function in Excel. The inverse relationship with r^2 means that *SI* is very sensitive to small changes in r^2, and *SI* can range in values from 1 (totally different spectra) → ∞ (identical spectra).

RESULTS AND DISCUSSION

Spectra. A typical MIR spectrum of wine is shown in **Figure 1**. It was observed that water and ethanol absorption peaks dominate the spectrum, with the C—O stretch for primary alcohols at pin 272 (1050 cm^{-1}) prominent. The contribution to the C—H stretch from pins 739 to 770 (2850−2960 cm^{-1}) from ethanol is also significant. The region from pins 438 to 456 (1690−1760 cm^{-1}) contains information relating to C=O stretching for aldehydes, carboxylic acids, and esters (*23*). It was noted that the manufacturer's recommendation of the cutoff point for one of the main water peaks might seriously impact on the quantitative measurement of key components in wine.

Principal Component Analysis. An objective of this study was to assess whether the variation in the wine spectra of the same wine measured on two instruments is significantly less than the variation between different wines measured on two instruments. If this is the case, then instrument noise might be

considered unimportant in this regard and the same wine should be able to be recognized, regardless of which instrument acquired the spectra.

The PCA of the wine samples analyzed confirmed that the differences between wine samples were small; therefore discrimination by rapid analytical techniques such as MIR would be difficult. In an industrial environment, therefore, subtle differences might be described only by sensory analysis done by an experienced winemaker rather than MIR methods. It is well-known that water accounts for approximately 85−90% of the wine sample matrix, representing the dominant component as shown in the eigenvector plot in **Figure 2a**. A number of chemical compounds including organic acids also have infrared absorption signatures overlapping with water MIR peaks. **Figure 2b** shows the major peaks in the PCA after removal of the spectral regions dominated by water. Peaks between pins 300 and 390 are largely due to -CH groups in the alcohol and acid components, while the peaks at pins 260−285 are largely due to -COH groups.

The PC score plot for the full spectrum (**Figure 3**) highlights that the two instruments are easily separated in the *x* direction (PC1, 51% variation). It was observed that by use of the manufacturer's recommendation to exclude the water absorption peaks in the 400−445 and 770−940 pin number ranges (1543−1717 and 2971−3627 cm^{-1}), the instrument effect along PC1 (85% variation) was virtually removed, as shown in **Figure 4**. Most of the variation, therefore, was now observed between wines rather than between the two instruments. Thus, in order to develop the fingerprinting system (reduce instrument variation), we removed the water region pin numbers (as suggested by the instrument manufacturer).

Similarity Index. Table 2 shows the results of the *SI* for a small subset of wine samples scanned on two MIR instruments for a single winery (winery A). In order to simulate the effect of transport from one winery to another, instrument 1 was designated as "dispatch site instrument" and instrument 2 was
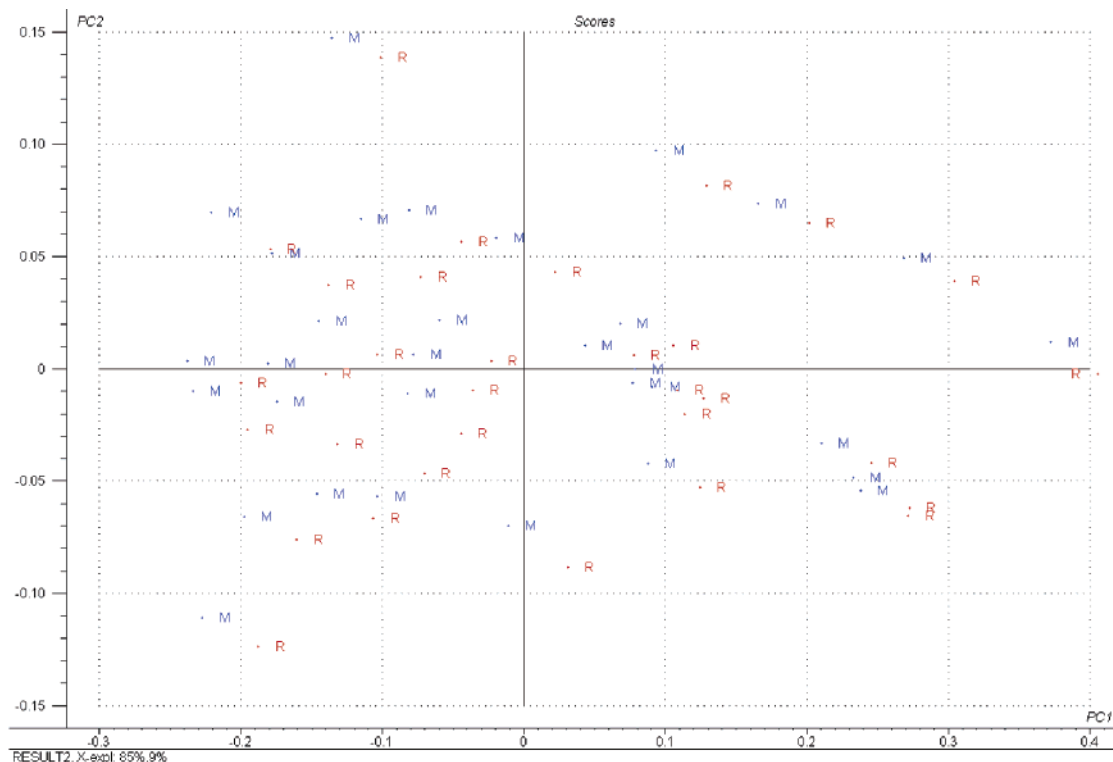
Figure 4. Score plot using Foss-recommended pin numbers for the same 32 samples scanned on two WineScan instruments (instrument 1 = M; instrument 2 = R).

Table 2. Similarity Index for a Subset of Wine Samples from Winery A[a]

sample, instrument 2	sample, instrument 1				
	C-01	C-02	C-03	C-04	C-05
C-01	(14 286)	3,226	2857	3846	2041
C-02		(25 000)	7692	6667	6667
C-03			(25 000)	**10 000**	6250
C-04				(25 000)	5263
C-05					(16 667)

[a] The results on the diagonal in parentheses are the *SI* for identical samples. Off-diagonal results are the *SI* between different samples. The result in boldface type is a false positive classification as identical wine.

Table 3. Summary of Similarity Index of Wine Samples for Each Winery[a]

winery	N samples	intrawinery comparisons	no. (%) matched correctly	incorrect classifications within each winery group
A	30	465	30 (100%)	6 (1.3%)
B	27	378	27 (100%)	8 (2.1%)
C	27	378	27 (100%)	6 (1.6%)
D	19	190	19 (100%)	8 (4.2%)
E	28	406	28 (100%)	7 (1.7%)
F	30	465	23 (77%)	8 (1.7%)

[a] Comparison of instrument 1 with instrument 2.

designated as "receival site instrument". An *SI* value of 10 000, corresponding to an r^2 value of 0.99990, was set as the minimum value for similarity, meaning that the wines were considered to be the same. Wine samples having values below this limit were considered different. It was observed that the *SI* successfully matched all identical wine samples. One sample, C-04, was also classified as matching sample C-03.

Table 3 summarizes the classification results for the whole set of red wine samples. In order to simplify the interpretation of the results obtained and exclude interwinery differences, samples within each winery group only were compared by the *SI* method. It was observed that wine samples belonging to winery F were 77% classified correctly. These wine samples were effervescing, with high levels of CO_2 and incomplete degassing prior to scanning, indicating malolactic fermentation had not been completed. Furthermore, PCA confirmed that there was poor spectral repeatability for wine samples from winery F. Therefore, caution must be taken when the classification results for this particular site are interpreted. Additionally, a small number of false positive samples, where two different

samples from the same site were incorrectly classified as matching (<5%), were observed in the classification.

It was not feasible to compare the spectrum of each sample from one instrument (instrument 1) to the entire wine set spectra acquired on the other instrument (instrument 2). However, a subset consisting of instrument 1 spectra from winery A were compared to instrument 2 spectra for wineries B− F. A total of 3930 permutations of wine spectra were compared. Only 21 samples from wineries B−F (0.5%) were incorrectly classified as belonging to winery A (*SI* > 10 000). These results suggested that the *SI* measurement can be used for an accurate comparison of wine samples between two very similar instruments.

The application of the similarity index to MIR spectra has been proven to provide an acceptable measurement of wine integrity across transportation. In five of the six winery data sets analyzed, the index correctly classified 98% of the wine samples. Additionally, it was found that less than 1% of wine samples were misclassified between the different regions investigated.

For quality assurance in the system, it is recommended that a reference wine such as cask wine also be analyzed with each

sample taken to ensure the instruments are comparable on wine known to be from the same origin. Further studies are needed in order to test practical adoption of the *SI* method to MIR spectra in commercial situations and to evaluate its potential as a rapid quality control tool in routine domestic and international wine transfer operations. The impact on *SI* from changes in dissolved oxygen, carbon dioxide, and sulfur dioxide levels requires investigation. Research into *SI* should also be extended to observe the effect of temperature and storage conditions for wine.

ABBREVIATIONS

MIR, mid-infrared; *SI*, similarity index; RI, refractive index; PCA, principal component analysis; DA, discriminant analysis; NIR, near-infrared; SIMCA, soft independent modeling of class analogy; ANN, artificial neural network; PRESS, predicted residual error sum-of-squares; PQS, polar qualification system; RSQ, *R*-squared.

ACKNOWLEDGMENT

LITERATURE CITED

(1) Somers, T. C.; Evans, M. E. Wine quality: correlations with color density and anthocyanin equilibrium in a group of young red wines. *J. Sci. Food Agric.* **1974**, *25*, 1369−1379.

(2) Somers, C. *The wine spectrum*; Winetitles: Adelaide, Australia, 1998; 136 pp.

(3) Bakeev, K. A. *Process Analytical Technology*; Blackwell Publishing: Oxford, U.K., 2005; 451 pp.

(4) Francis, I. L.; Høj, P. B.; Dambergs, R. G.; Gishen, M.; de Barros Lopes, M. A.; Pretorius, I. S.; Godden, P. W.; Henschke, P. A.; Herderich, M. J.; Waters, E. J. Objective measures of grape quality—are they achievable? In *Proceedings 12th Australian Wine Industry Technical Conference*; Australian Wine Industry Technical Conference Inc. Australia, 2005, pp 85−90.

(5) Munck, L.; Norgaard, L.; Engelsen, S. B.; Bro, R.; Andersson, C. A. Chemometrics in food science: a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 31−60.

(6) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*; Wiley−VCH: Chichester, U.K., 1999; 314 pp.

(7) Kelly, J. F. D.; Downey, G. Detection of sugar adulterants in apple juice using Fourier transform infrared spectroscopy and chemometrics. *J. Agric. Food Chem.* **2005**, *53* (9), 3281−3286.

(8) Murray, I.; Kaffka, K. J.; Sergely, Z. S. PQS detects adulteration of fishmeal with meat and bone meal—a contribution to help eradicate "Mad Cow Disease". In *Near Infrared Spectroscopy: Proceedings of the 11th International Conference*; Davies, A. M. C., Garrido-Varo, A., Eds.; NIR Publications: West Sussex, U.K., 2004; pp 641−645.

(9) Gergely, S.; Farkas, K.; Forgács, A.; Salgó, A. Quantitative and qualitative differentiation of alcoholic beverages by near infrared spectroscopy. In *Near Infrared Spectroscopy: Proceedings of the 11th International Conference*; Davies, A.M.C., Garrido-Varo, A., Eds.; NIR Publications: West Sussex, U.K., 2004; pp 569−572.

(10) Nieuwoudt, H. H.; Prior, B. A.; Pretorius, I. S.; Manley, M.; Bauer, F. F. 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. *J. Agric. Food Chem.* **2004**, *52* (12), 3726−3735.

(11) Cordella, C.; Moussa, I.; Martel, A. C.; Sbirrazzuoli, N.; Lizzani-Cuvelier, L. Recent developments in food characterization and adulteration detection: technique-oriented perspectives. *J. Agric. Food Chem.* **2002**, *50* (7), 1751−1764.

(12) Downey, G. Qualitative analysis in the near infrared region. *Analyst* **1994**, *119*, 2367−2375.

(13) Cozzolino, D.; Smyth, H. E.; Gishen, M. Feasibility study on the use of visible and near infrared spectroscopy to discriminate between white wine of different varietal origin. *J. Agric. Food Chem.* **2003**, *51* (26), 7703−7708.

(14) Mark, H. Qualitative Analysis. In *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 2001; pp 233−238.

(15) Mark, H. L.; Tunnell, D. Qualitative near infrared reflectance analysis using Mahalanobis distances. *Anal. Chem.* **1985**, *57* (7), 1449−1456.

(16) Bertran, E.; Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Montoliu, I. Near infrared spectrometry and pattern recognition as screening methods for the authentication of virgin olive oils of very close geographical origins. *J. Near Infrared Spectrosc.* **2000**, *8*, 45−52.

(17) Coene, M. P. D.; Grinter, R.; Davies, A. M. C. The use of quadratic regression in qualitative near infrared and visible spectroscopic analysis. *J. Near Infrared Spectrosc.* **1996**, *4*, 153−161.

(18) Anon. *WineScan^TM FT 120 Type 77110 and 77310 Reference Manual;* Foss Electric; Hillerød, Denmark, 2002; 130 pp.

(19) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley and Sons: New York, 1989; 419 pp.

(20) Naes, T.; Isaksson, T.; Fearn, T.; Davies, A. M. C. *A user-friendly guide to multivariate calibration and classification*; NIR Publications: Chichester, U.K., 2000; 344 pp.

(21) Davies, A. M. C. Qualitative analysis: is it time to use the quadratic correlation approach? In *Near Infrared Spectroscopy: Proceedings of the 11th International Conference*; Davies, A. M. C., Garrido-Varo, A., Eds.; NIR Publications: West Sussex, U.K., 2004; pp 237−240.

(22) Kaffka, K. J.; Gyarmati, L. S. Investigating the polar qualification system. *J. Near Infrared Spectrosc.* **1998**, *6*, 191−200.

(23) Stuart, B. *Modern Infrared Spectroscopy*; John Wiley & Sons: Chichester, U.K., 1996; 180 pp.