

Calibration models for routine analysis of cider by mid-infrared spectroscopy

Anna Picinelli Lobo*, Belén Suárez Valles, Norman Fernández Tascón, Roberto Rodríguez Madrera, Ovidio Fernández García

Área de Tecnología de los Alimentos, SERIDA, Apdo. 13, 33300-Villaviciosa (Asturias), Spain

Abstract

Partial least square regression was used to develop different calibration data sets for multi-parameter routine analysis of ciders. Parameters included were those related with the monitoring of fermentation process and cider quality: specific gravity, total and volatile acidities, alcoholic proof, pH and fructose. Calibration performances were evaluated by means of the prediction residual sum of squares (PRESS), the root mean squared prediction error of cross-validation (RMSECV) and the residual predictive deviation (RPD) values (ratio of the standard deviation of the population to the standard error of cross-validation). Validation of the models was assessed in terms of accuracy and precision. Mean recoveries of the predicted results compared to the reference values were close to 100%, with repeatability and reproducibility similar to those of the reference methods.

© 2006 Swiss Society of Food Science and Technology. Published by Elsevier Ltd. All rights reserved.

1. Introduction

From the oenological point of view, monitoring of the cider-making process is needed to prevent alterations. Nowadays, it is important to have fast and reliable analytical responses to satisfy the increasing demand from cider makers of determining more parameters in a greater number of samples.

The potential of Fourier transform infrared (FTIR) spectroscopy has long been recognized in different fields, both research (Coimbra, Gonçalves, Barros, & Delgado, 2002; Edelman, Diewok, Schuster, & Lendl, 2001; Palma & Barroso, 2002) and routine analysis (Duarte, Barros, Almeida, Spraul, & Gil, 2004; Dubernet & Dubernet, 2000; Dubernet et al., 2000; Kupina & Shrikhande, 2003). Infrared frequency signals are processed through a series of mathematical procedures and are transformed into concentrations of the different components in the sample. Calibration process involves multi-variate statistical procedures such as principal component analysis, principal component regression and partial least squares (PLS) regression (Martens and Næs, 1996). The last approach

has been proved to be the most appropriate to extract information from complex spectra, containing overlapping absorption peaks. The choice of wavelengths for calibration model building is critical to ensure both the prediction ability and the robustness of these models. Traditionally, PLS analysis involved the selection of all wavelengths contained in a particular region of the spectrum, known to contain useful information. However, it has been shown that reducing variables by eliminating those with great noise or irrelevant information improved the ability of the prediction models. The selection of the most informative variables is often performed through interactive trial and error processes, taking the root mean square prediction error of cross-validation (RMSECV) as a measure of performance of the model (Garrido Frenich et al., 1995; Osborne, Jordan, & Künnemeyer, 1997).

Furthermore, the successful application of FTIR in the routine analysis implies the development of reliable and robust calibration models, which enable the detection of poorly predicted samples. Evaluation of commercial calibration models for different parameters in wine has recently been published (Kupina & Shrikhande, 2003; Nieuwoudt, Prior, Pretorius, Manley, & Bauer, 2004). Conclusions inferred from these works showed the

*Corresponding author.

E-mail address: apicinelli@serida.org (A. Picinelli Lobo).

necessity for developing suitable and powerful tools for quality control.

This paper was focused on the developing and validation of prediction models for the routine analysis of cider.

2. Material and methods

2.1. Cider samples

The calibration data sets used to build up the prediction models for cider are summarized in Table 1. These samples correspond to ciders exclusively made from cider-apple pressing, analysed in different stages of the making process, from the end of the fermentation to several months in bottle. Additionally, a validation set of cider samples was analysed by reference methods to verify the suitability of the calibration models (Table 1).

2.2. Spectroscopic measurements

Infrared analyses were performed with an Avatar 360 spectrometer (Thermo Nicolet, USA) equipped with a Globar source, a deuterated triglycine sulphate detector with KBr beamsplitter, temperature control of the sample cell by Peltier effect and automatic sampler. Spectroscopic measurements were made from 3300 to 950 cm^{-1} , using a nominal accumulating of 30 scans and a resolution of 4 cm^{-1} .

2.3. Software

The spectrometer is controlled by means of the Nicolet's OMNIC software to collect and analyse FTIR spectra. Additionally, specific software for wine analysis, through a series of commercial calibration models, and calibration developing (CETIM, v. 2003.1) were purchased from TDI (Barcelona, Spain).

2.4. Mathematical tools

Calibration models were developed by PLS. The accuracy of the predictive ability of the model is assessed

by leave-ten-out cross-validation, by means of the RMSECV, which is obtained according to the expression:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (\hat{c}_i - c_i)^2}{n}}, \quad (1)$$

where \hat{c}_i is the predicted concentration and c_i the actual concentration, as determined by the reference method, and n the number of samples used in the calibration model. The number of factors to be included in each model was determined taking as criterion the minimum prediction residual sum of squares (PRESS) (Geladi & Kowalski, 1986).

When the assessment of the prediction performance is evaluated by means of an independent validation set, the expression (1) is substituted by the root mean-squared error of prediction (RMSEP), where n is the number of samples used in the validation set.

The residual predictive deviation (RPD) is defined as the ratio of the standard deviation of the population (S_d) to the standard deviation of prediction errors (S_e), and is used to evaluate the predictive ability of the calibration models.

Validation of the models against the reference methods were performed in terms of accuracy (bias and recovery) and precision (reproducibility and repeatability). Bias, which indicates systematic errors in the predictive values, is defined by the expression:

$$\text{BIAS} = \frac{1}{n} \sum_{i=1}^n (\hat{c}_i - c_i), \quad (2)$$

whereas recovery (%) was calculated as the ratio of the bias to the reference values:

$$\text{Recovery} = \frac{1}{n} \sum_{i=1}^n \frac{(\hat{c}_i - c_i)}{c_i} 100. \quad (3)$$

Repeatability was obtained as the standard deviation of the replicates obtained each day, whereas reproducibility was calculated by the expression $2\sqrt{2}S_R$, S_R being the standard deviation of the prediction errors for the n samples.

Table 1
Description of the calibration and validation data sets

Parameters	Calibration sets					Validation sets				
	Number of standards	Mean	SD	Maximum	Minimum	Number of standards	Mean	SD	Maximum	Minimum
Specific gravity (g/l)	80	0.9988	0.0315	1.0122	0.9962	47	0.99813	0.00136	1.00243	0.99645
Total acidity ^a	116	3.44	0.67	6.62	2.21	51	3.46	0.70	5.90	2.45
Volatile acidity ^b	147	1.47	0.91	4.76	0.12	44	1.13	0.63	2.76	0.20
pH	122	3.80	0.13	4.15	3.48	48	3.78	0.10	4.05	3.60
Alcoholic proof (% v/v)	105	6.04	0.83	8.24	3.46	48	6.31	0.38	7.06	4.17
Fructose (g/l)	51	1.64	2.69	13.9	<1.0	25	2.16	3.73	11.70	<1.0

^ag sulphuric acid/l.

^bg acetic acid/l.

Different mathematical pre-processing treatments are available: mean centred, which calculates the difference between the mean spectrum of the calibration set and each of the spectra of this set; reduced variance, which subtract from each spectrum the mean spectrum of the calibration set and then divides each point of the spectra by the typical deviation of the calibration set. First-order derivative is obtained as the difference between the absorbances at two consecutive wavelengths. Higher-order derivatives can be obtained by subsequently applying the operation to the derivatives. The use of derivatives allows to reduce correlation between variables and to better differentiate between overlapped peaks; however, it increases the signal-to-noise relationship.

2.5. Reference methods

Reference values for titratable acidity (TA, g sulphuric acid/l), volatile acidity (VA, g acetic acid/l) and pH were obtained according to the [EU Official Analytical Methods \(1998\)](#). Different internal validated methods were used for determining the alcoholic proof (% v/v) (near infrared spectroscopy), the specific gravity (SG, g/l) (electronic densitometry based on the oscillation of a U-tube density sensor) and fructose (g/l) (HPLC with refraction index detection, according to the technique described by [Blanco Gomis, Gutiérrez Álvarez, Mangas Alonso, and Noval Vallina \(1988\)](#)).

3. Results and discussion

3.1. Mathematical modelling

In [Fig. 1](#), a typical FTIR absorbance spectrum of cider is shown. Relevant spectroscopy areas are located between 1300 and 950 cm^{-1} , corresponding to the C–O stretching, 1470–1200 cm^{-1} (C–H bending), 1750–1680 cm^{-1} (C=O stretching) and 2990–2900 cm^{-1} (C–H stretching) ([Willard, Merritt, Dean, & Settle, 1990](#)). Carbon dioxide absorbs around 2350 cm^{-1} , whereas maximum transmittance for

water occurs around 1600 cm^{-1} , so that those regions should be avoided for calibration modelling.

PLS was chosen for calibration. Selection of the wavenumber regions for calibration was performed taking into account both the chemical groups involved in each parameter and the higher correlations between the instrumental response and the corresponding reference values for each parameter, which are automatically calculated by the software. At this step, it is important to get the maximum explained variance with the lower number of variables and the minimum PRESS. Usually, this value decreases as the number of principal factors increases; however, a compromise should be maintained between the error value, which measures the prediction ability of the model, and the number of PC taken for calibration, to reach a calibration model both robust and accurate.

The influence of the number of predictor variables on the prediction ability of the different models is summarized in [Tables 2 and 3](#). Calibration models were built by trial and error, by combining different spectral regions and mathematical treatments in such a manner to reach the minimum PRESS value.

With respect to the parameters related to acidity (pH, TA and VA), it could be expected to have prominent absorbance peaks between 1180 and 1000 cm^{-1} (C–O stretching), 1470–1400 cm^{-1} (C–H bending) and 1750–1680 cm^{-1} (C=O stretching). In fact, there were two high-correlation regions between 1770 and 1170 cm^{-1} , and a third one at 2669–2391 cm^{-1} ; whereas for pH, five regions were found ([Table 2](#)). The spectral band located at 1750–1690 cm^{-1} had the highest correlation in all of these calibration models. This can be explained on the basis that VA (and therefore, TA) includes both acetic acid and esters belonging to the acetic series, which characteristic group (C=O) absorbs at that wavenumber.

In general, the application of mathematical pre-treatment improved the statistics of the models, especially that corresponding to pH, as evidenced by the increasing values of the ratio RPD ([Table 2](#), models M1-A to M1-C). This parameter has been associated with the ability of the calibration models for prediction. In this sense, [Nieuwoudt et al. \(2004\)](#) take as criterion those values for RPD higher than 5 indicate that the calibration equations are suitable for quantification. In our case, models for both TA and VA exhibited higher values of the RPD than those of the pH models. It can be explained because the corresponding calibration data sets for both TA and VA represented a greater variability in terms of concentration levels than that used for pH calibration. In any case, the MRSECV were always lower than 0.11 ([Table 2](#)).

In the case of the SG, six correlation regions were found between 1785 and 950 cm^{-1} . Sugars are di-functional compounds, with alcohol (primary and secondary) and C=O (aldehyde and ketone) groups. Thus, the correlation bands found were coincident with typical infrared bands for all these groups (1300–1000 cm^{-1}). The other correlations can be assigned to C–H bending (1470–1370 cm^{-1})

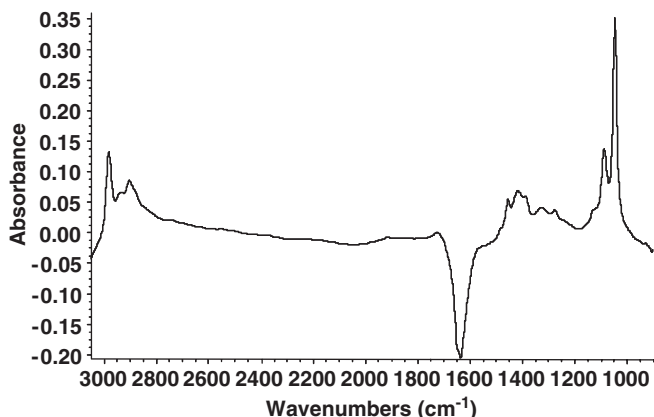


Fig. 1. Typical FTIR spectrum of cider.

Table 2
Mathematical modelling for calibration of pH, titratable and volatile acidity in ciders

Parameter	Selected wave numbers (cm ⁻¹)	A: No pre-treatment					B: Mean centred					C: Mean centred + first derivative				
		Model reference	PC	MRSECV	R ²	RPD	PC	MRSECV	R ²	RPD	PC	MRSECV	R ²	RPD		
pH	1014–960															
	1349–1095															
	1573–1519	M1	9	0.152	0.7017	1.403	10	0.032	0.9820	5.053	9	0.029	0.9856	5.905		
	1789–1712															
Titratable acidity ^a	1473–1176	M2	5	0.092	0.9933	8.686	7	0.077	0.9954	10.479	5	0.074	0.9958	10.930		
	1770–1693															
	1473–1176	M3	4	0.093	0.9932	8.574	7	0.077	0.9956	10.661	5	0.072	0.9956	10.633		
	1770–1693															
Volatile acidity ^b	2669–2391															
	1473–1176	M4	10	0.106	0.9978	15.225	12	0.067	0.9983	17.233	10	0.063	0.9987	19.511		
	1747–1686															
	1473–1176	M5	10	0.099	0.9976	14.403	12	0.070	0.9986	18.785	11	0.067	0.9989	21.511		
1747–1686																
2669–2522																

M1–M5: references for calibration models with different sets of wavenumbers; A, B, C: mathematical options for calibration modelling.

^ag sulfuric acid/l.

^bg acetic acid/l.

and C=O stretching in aldehydes and ketones (1735–1681 cm⁻¹).

By taking all of these correlation regions, the optimal number of PC was the highest, together with the value of the standard error of calibration. When first derivative and/or mean centred were applied, the prediction ability of the model improved, with only four PC being necessary. After that, slight improvements in the RMSECV values were obtained by taking into account only two regions, those related with C–O stretch; the selected model was that including two spectral regions and mean centred as pre-processing treatment (model M7-B in Table 3).

Prediction modelling for fructose determination can be considered as described for SG. Independently of the spectroscopic signals chosen or the mathematical processing applied, the RPD values obtained were always higher than those found for the SG calibration models (Table 3). As explained for TA and VA, this fact should be related to the greater variability of the samples used as calibration set for fructose. From the theoretical point of view, the best prediction model was obtained by only taking two regions with mean centred and first derivative (model referred to as M9-C in Table 3).

For alcoholic proof, up to five correlation regions were found between 1619 and 950 cm⁻¹, and between 2990 and 2960 cm⁻¹ (Table 3). These areas include the aforementioned interactions. By combining the lowest values for RMSECV and number of factors, the optimal model was considered to be that referred to as M10-B.

As said before, both PRESS and RMSECV decrease as the number of PC increases, until a local minimum is reached. This minimum establishes a compromise between the prediction ability of the model and its capacity for estimating the analytical property of unknown samples, that is, accuracy and robustness. As shown in Fig. 2, this trend was observed for all the parameters studied, except VA, which presented a steady decrease of these error values with the number of factors. The selection of the optimal number of PC was automatically performed by increasing the number of factors until differences between actual and previous PRESS value become small compared to some previously established measurement error.

3.2. Method validation: accuracy, repeatability and reproducibility

Selected calibration models were checked by means of an independent set of samples, including the whole range of concentrations (Table 1). From this set, the following parameters were obtained and compared: bias, recovery, RMSPE, repeatability and reproducibility. In Table 4, the validation results for the optimal prediction models are summarized. As seen, the RMSPE were similar to those obtained in the cross-validation process or slightly higher, meaning that the models are robust and suitable to predict unknown samples. Bias and mean recovery values demonstrated that no systematic errors exist compared to the

Table 3
Mathematical modelling for calibration of specific gravity, alcoholic proof and fructose in ciders

Parameter	A: No pre-treatment				B: Mean centred				C: Mean centred + first derivative					
	Selected wave numbers (cm ⁻¹)	PC	MRSECV	R ²	RPD	PC	MRSECV	R ²	RPD	PC	MRSECV	R ²	RPD	
Specific gravity (g/l)	1033–950													
	1203–1052													
	1265–1214													
	1373–1322	M6	15	0.042	0.1638	1.002	4	0.0006	0.9850	5.068	4	0.0007	0.9849	10.076
	1473–1419													
	1735–1681													
Fructose (g/l)	1033–950	M7	12	0.055	0.1216	1.001	4	0.0006	0.9950	10.024	4	0.0007	0.9749	4.488
	1203–1052													
	1018–956	M8	8	0.222	0.9990	22.195	9	0.179	0.9995	30.487	4	0.241	0.9972	13.424
	1268–1060													
	1373–1334													
	1481–1461	M9	8	0.283	0.9979	15.458	10	0.229	0.9994	28.352	5	0.206	0.9988	43.193
Alcoholic proof. (%av/v)	991–952													
	1141–1037													
	1373–1338													
	1482–1465	M10	3	0.125	0.9898	11.044	5	0.088	0.9959	11.050	7	0.079	0.9970	11.619
	1619–1592													
	2990–2962													
Alcoholic proof. (%av/v)	991–952													
	1141–1037	M11	6	0.094	0.9958	9.962	8	0.081	0.9969	9.696	6	0.080	0.9970	12.627
2990–2962														

M6–M11: references for calibration models with different sets of wavenumbers; A, B, C: mathematical options for calibration modelling.

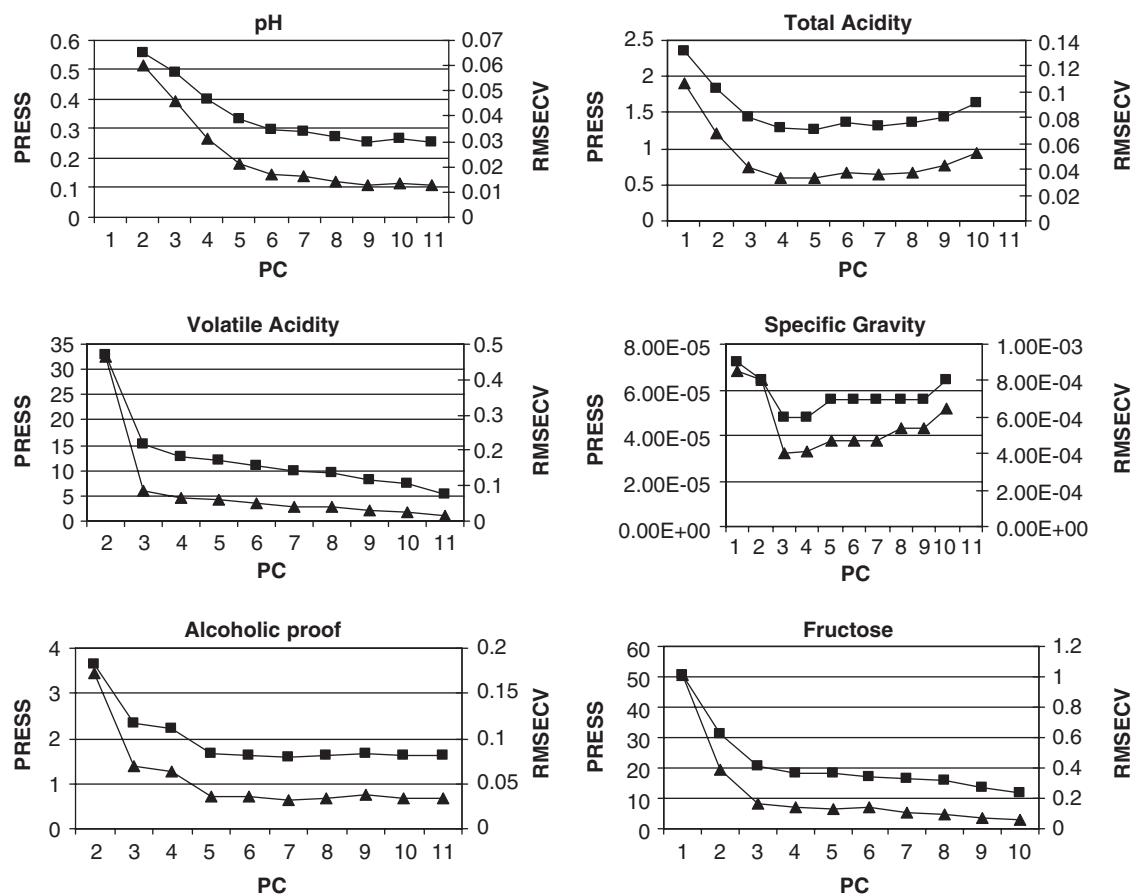


Fig. 2. Plot of the statistical PRESS and RMSECV against the number of components taken for calibration (Legends: ▲, PRESS; ■, RMSECV).

Table 4
Validation of the FTIR technique against official methods by means of an independent set of samples

Parameter	Model reference ^a	RMSEP	Bias	Recovery (%)	Repeatability		Reproducibility	
					FTIR	Reference method	FTIR	Reference method
pH	M1-C	0.04	-0.01	99.8	0.02	0.02	0.12	0.10
Total acidity (g sulphuric/l)	M2-C	0.090	-0.03	100.9	0.03	0.05	0.24	0.20
Volatile acidity (g acetic/l)	M4-C	0.075	0.01	100.2	0.05	0.04	0.21	0.08
Specific gravity	M7-B	0.00109	0.00021	100.02	0.00010	0.00010	0.00305	0.00037
Fructose (g/l)	M9-C	0.44	-0.10	96.9	0.06	0.08	1.23	0.18
Alcoholic proof (% v/v)	M10-B	0.085	-0.02	99.7	0.04	0.10	0.24	0.19

^aSee Tables 2 and 3 for references for the selected calibration models.

reference methods. Likewise, the prediction models were as repetitive and reproducible as the reference methods, especially those for TA and pH. Even though the results for VA, alcoholic proof and SG obtained by FTIR were slightly less reproducible than the reference methods, these can be considered as suitable for routine purposes. The calibration model for fructose presented the highest bias error and the lowest value for recovery; however, this is normal, considering the usual concentration levels found in natural ciders and the quantification limit of our reference method (1 g/l).

4. Conclusions

Six reliable and suitable calibration models were optimized for the routine analysis of different parameters in cider. Selection of the most adequate spectral regions and mathematical treatments was performed using as criterion the value of the prediction residual squares sum, which measures the prediction ability of the model. The external validation showed that the optimized models were accurate, precise and robust, therefore, suitable for routine analysis purposes.

Acknowledgements

The authors are indebted to the Principado de Asturias for financial support.

References

- Blanco Gomis, D., Gutiérrez Álvarez, M. D., Mangas Alonso, J. J., & Noval Vallina, A. (1988). Determination of sugars and alcohols in apple juice and cider by HPLC. *Chromatographia*, *25*, 701–706.
- Coimbra, M. A., Gonçalves, F., Barros, A. S., & Delgado, I. (2002). Fourier transform infrared spectroscopy and chemometric analysis of white wine polysaccharide extracts. *Journal of Agricultural and Food Chemistry*, *50*, 3405–3411.
- Duarte, I. F., Barros, A., Almeida, C., Spraul, M., & Gil, A. M. (2004). Multivariate analysis of NMR and FTIR as a potential tool for the quality control of beer. *Journal of Agricultural and Food Chemistry*, *52*, 1031–1038.
- Dubernet, M., & Dubernet, M. (2000). Utilisation de l'analyse infrarouge à transformée de Fourier pour l'analyse œnologique de routine. *Revue Française d'Œnologie*, *181*, 10–13.
- Dubernet, M., Dubernet, M., Dubernet, V., Coulomb, S., Lerch, M., & Traineau, I. (2000). Analyse objective de la qualité des vendanges par spectrométrie infrarouge à transformée de Fourier (FTIR) et réseaux de neurones. *Revue Française d'Œnologie*, *185*, 18–21.
- Edelman, A., Diewok, J., Schuster, K. C., & Lendl, B. (2001). Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts. *Journal of Agricultural and Food Chemistry*, *49*, 1139–1145.
- European Union Official Methods of Analysis. (1998). Wines, juices and grape musts analyses. In , *MAPA*, Vol. II (pp. 487–729). Madrid: Ministerio de Agricultura, Pesca y Alimentación.
- Garrido Frenich, A., Jouan-Rimbaud, D., Massart, D. L., Kuttatharmakul, S., Martínez Galera, M., & Martínez Vidal, J. L. (1995). Wavelength selection method for multicomponent spectrophotometric determination using partial least squares. *Analyst*, *120*, 2787–2792.
- Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. *Analytica Chimica Acta*, *185*, 1–17.
- Kupina, S. A., & Shrikhande, A. J. (2003). Evaluation of a Fourier transform infrared instrument for rapid quality-control wine analyses. *American Journal of Enology and Viticulture*, *54*, 131–134.
- Martens, H., Næs, T., 1996. In H. Martens, T. Næs (Eds.), *Multivariate calibration*. Chichester: Wiley.
- Nieuwoudt, H. H., Prior, B. A., Pretorius, I. S., Manley, M., & Bauer, F. F. (2004). 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. *Journal of Agricultural and Food Chemistry*, *52*, 3726–3735.
- Osborne, S. O., Jordan, R. B., & Künnemeyer, R. (1997). Method of wavelength selection for partial least squares. *Analyst*, *122*, 1531–1537.
- Palma, M., & Barroso, C. G. (2002). Application of FT-IR spectroscopy to the characterisation and classification of wines, brandies and other distilled drinks. *Talanta*, *58*, 265–271.
- Willard, H. H., Merritt, L. L., Jr., Dean, J. A., & Settle, F. A., Jr. (1990). Infrared spectrometry. In H. H. Willard, L. L. Merritt Jr., J. A. Dean, & F. A. Settle Jr. (Eds.), *Instrumental methods of analysis* (pp. 287–320). Belmont, CA: Wadsworth Publishing Company.