Available online at www.sciencedirect.com

SCIENCE DIRECT®

Food Chemistry 99 (2006) 615–621

Food Chemistry

www.elsevier.com/locate/foodchem

Analytical, Nutritional and Clinical Methods

# Prediction of organic acids and other quality parameters of wine vinegar by near-infrared spectroscopy. A feasibility study

M.J. Sáiz-Abajo, J.M. González-Sáiz, C. Pizarro *

*Department of Chemistry, University of La Rioja, c/ Madre de Dios 51, 26006 Logroño (La Rioja), Spain*

Received 7 April 2004; received in revised form 30 November 2004; accepted 14 August 2005

**Abstract**

A method for quality control of wine vinegar has been developed based on near-infrared spectroscopy that can be applied to monitor processes or in quality control in vinegar industries. In vinegar production, stages such as on-line fermentation monitoring, dilution process and quality control of the final product are relevant and necessary. The aim of this study was to achieve and simplify these manufacturing stages through the determination of 14 parameters from a spectral measurement. Total acids, non-volatile and volatile acids, organic acids, L-proline, solids, ash and chloride in vinegar were determined by near-infrared spectroscopy. Different data pre-processing methods such as orthogonal signal correction, multiplicative scatter correction or column centering were used and the final multivariate methods were evaluated and compared. The calibration models were validated and prediction capacity studied in order to examine their practical applicability to the vinegar industry.

*Keywords:* Calibration; Wine vinegar; Quality control; Predictor selection; Partial least squares; Characterisation

## 1. Introduction

Food quality and safety has always been very important. Ever since food became a trade object, there has always been fraud, usually the wilful manipulation of a product for financial gain, despite possible health risks to consumers (Anklam & Battaglia, 2001). As a result of this constant search for continuous assessment of food quality, many methods based on near-infrared spectroscopy (NIRS) have evolved rapidly over the past decade making NIR spectroscopy an appropriate technique for food analysis (Blanco & Villarroya, 2002; Xiccato, Trocino, Tulli, & Tibaldi, 2004).

One of the positive features of near-infrared spectroscopy is that it is a non-destructive analytical technique, it has the ability to record spectra of solid and liquid samples with no prior manipulation (Xiccato et al., 2004), it is not time-consuming and, in recent years, near-infrared spectroscopy has gained wide acceptance in different fields due to its advantages over other analytical techniques. NIR spectroscopy used in connection with multivariate calibration techniques affords important industrial applications that are spreading rapidly (Büning-Pfaue, 2003; Garrido-Varo, 1997; Larrechi & Callao, 2003; Shenk & Westerhaus, 1995).

However, working with NIR data involves some knowledge of the system and some of the parameters that affect NIR measurements. The complexity of NIR spectra makes direct interpretation impossible and prompts the need for multivariate approaches in order to obtain all the chemical information contained in the spectral variables. Moreover, the existence of optical interferences such as light scatter in NIR data always demands indirect multivariate calibration. Other parameters such as temperature or turbidity should also be taken into account when calibrating a NIR spectrophotometer. Multicollinearity in NIR data creates lack or robustness and loss of accuracy problems that, when added to those mentioned previously, makes the calibration process a complex and delicate step that

---

* Corresponding author. Tel.: +34 941299626; fax: +34 941299621.
*E-mail address:* consuelo.pizarro@dq.unirioja.es (C. Pizarro).

0308-8146/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodchem.2005.08.006

should be prepared and validated before being used in everyday operations in the industry.

The dilution process is an important stage in vinegar production. From the vinegar just fermented and until it is bottled, the product undergoes a dilution process in order to adjust the acetic grade to commercial specifications. Spanish regulations establish that wine vinegars should have a minimum total acidity of 60 g $L^{-1}$ expressed as acetic acid (Royal Decree 2070/1993 November 26th). Dilution is economically important to vinegar manufacturers because an error in this stage of the process may cause either an economic loss or a problem for the manufacturer if the final product does not comply with legal specifications.

Any kind of industrial process is complex due to its multicomponent nature. Thus, even with extensive automation, controlling industrial processes is a very difficult task. Industry is rapidly moving from post-line production control to on-line production control in order to avoid, or at least reduce, non-conformities in the production process and enable remote process monitoring (Coffey, Cooley, & Walker, 1999). Models for monitoring dynamic processes are usually constructed by withdrawing samples from the system of interest and analysing them using reference methods; the results from the different analytes in such samples and the spectral information obtained from them are used for constructing calibration models. These were the dynamics adopted to develop the analytical method described here.

This study reports a rapid and safe method for performing final dilution stage in vinegar production. This is a general objective within the new research trends included in the PAT initiative (process analytical technology) aimed at improving commercial products and cutting costs by controlling production processes on-line in the manufacturing process.

To achieve this goal, 14 chemico-physical parameters in wine vinegar were determined by the Spanish official methods (AOAC, 1995, Chap. 43) or by other methods developed and validated in the laboratory (Sáiz-Abajo, González-Sáiz & Pizarro, ACA 2005) to calibrate the NIR instrument for simultaneous monitoring of all components in the final product of a vinegar producer.

Some of these parameters are part of the analyses encompassed within the technical regulations governing the safety of food products applied in the vinegar industry to test product quality, and others are important for assessing the genuineness of the final product. Spanish legislation (Royal Decree 2070/1993 November 26th) establishes certain levels for some of these physico-chemical parameters for wine vinegar. The total acids content expressed as acetic grade must be at least of 6° (60 g $L^{-1}$), as stated above. Solids must have values larger than 1.30 g $L^{-1}$ of vinegar and per acetic grade. Ash content must be between 1 and 5 g $L^{-1}$ and chloride concentration must be under 1 g $L^{-1}$ expressed as sodium chloride. On the other hand, L-proline and organic acids are indicative of the genuineness of wine vinegar (Sáiz-Abajo, González-Sáiz & Pizarro, JAFC 2004) and the levels of these compounds are significant for the vinegar quality.

## 2. Materials and methods

### 2.1. Instrumentation and software

NIR spectra were collected for the 1100–2500 nm range using a FOSS NIRSystem 5000 liquid analyser spectrometer (Foss NIRSystems, Silver Spring, MD20904, USA) equipped with a flow cell. The instrument was controlled by a compatible PC using Vision 2.22 (Foss NIRSystems, Silver Spring, MD20904, USA) for data acquisition.

For high-performance liquid chromatography measurements, a modular apparatus comprising a complete HP 1100 Series System with a vacuum degasser, a quaternary pump, an autosampler, a thermostatic column compartment and a diode array detector were used. Separation was performed on an analytical Zorbax SB-C18 (i.d. $4.6 \times 250$ mm with 5 μm particle size). Detection was performed using a diode array detector. The detection wavelength was set at 210 nm and the slit-width was set at 2 nm.

For regulatory analyses (*Reglamento Técnico Sanitario*-Technical Health Regulations), an oven (P Selecta), a furnace and a pH electrode (Crison, micropH 2002, Barcelona Spain) and a $Cl^{-}$ Ion-selective electrode (Crison, Alella, Barcelona, Spain) were used.

Chemometric analysis was performed using the following software packages: Unscrambler v7.8, MATLAB 6.1 and V-PARVUS package (version 2003).

### 2.2. Reagents

For organic acids and L-proline analyses, L(+)-tartaric, L(−)-malic, lactic, acetic, succinic, citric acid and L-proline of analytical-reagent grade were supplied by Merck (Darmstadt, Germany). For the preparation of the mobile phase, potassium di-hydrogen phosphate of analytical-reagent grade from Panreac (Barcelona, Spain), methanol of gradient grade and phosphoric acid 85% from Merck were used. Water purified using a Milli-Q system académic A10 (Millipore S.A., Molsheim, France) was used. Standard and calibration solutions of the compounds were prepared in Milli-Q water. All the solutions were stored at 4 °C in the refrigerator.

For the regulatory analyses, sodium hydroxide from Prolabo (Fontenay, France), hydrogene peroxide 33% (110 vol) and sodium nitrate from Panreac (Barcelona, Spain), phosphoric acid 85% from Merck, sodium chloride from Sigma–Aldrich Co.(Gillingham, Dorset, UK) were used.

### 2.3. Samples

Six genuine wine vinegar samples were taken directly from different fermentors of a vinegar producer in order to join a representative amount of samples. The samples were collected once the two-fold fermentation process had concluded. Several dilute samples, simulating those usually taken in the industry, were obtained in the laboratory to cover the whole range from the just fermented product (about 115–120 g $L^{-1}$ of total acidity expressed as acetic acid) to the commercial product (60 g $L^{-1}$ expressed as acetic acid). The range was also extended below the lower limit to study and detect possible fraudulent dilutions. The final amount of samples comprised 54 vinegar samples. The experimental parameters and the ranges studied are shown in Table 1.

Apart from the samples used in the calibration and validation processes, 10 commercial white wine vinegar samples collected from several supermarkets in northem Spain were used to test a posteriori the ability of the models performed to evaluate the different quality parameters of the vinegar.

Table 1
Range of the experimental parameters

| Parameter | Unit | Range |
|---|---|---|
| Total acids | g of acetic acid in 100 mL vinegar | 1.00–11.91 |
| Non-volatile acids | g of acetic acid in 100 mL vinegar | 0.01–0.18 |
| Volatile acids | g of acetic acid in 100 mL vinegar | 0.99–11.64 |
| Chloride | g $L^{-1}$ of NaCl | 0.177–0.533 |
| Solids | g $L^{-1}$ | 1.30–17.56 |
| Ash | g $L^{-1}$ | 0.16–2.08 |
| L-proline | g $L^{-1}$ | 0.20–2.89 |
| L(+)-tartaric acid | g $L^{-1}$ | 0.12–1.47 |
| L(−)-malic acid | g $L^{-1}$ | 0.05–0.62 |
| Lactic acid | g $L^{-1}$ | 0.25–0.91 |
| Acetic acid | g $L^{-1}$ | 9.87–118.39 |
| Citric acid | g $L^{-1}$ | 0.09–1.18 |
| Succinic acid | g $L^{-1}$ | 0.24–0.80 |
| D-malic acid | g $L^{-1}$ | 0.003–0.043 |

In order to avoid any optical interferences or scattering effects in the NIR measurements that can affect the subsequent calibration process, all the samples were filtered beforehand.

### 2.4. NIR spectra collection

The spectra were recorded at 43.0 ± 0.1 °C. The samples were heated to this temperature before being introduced in the NIRSystem to take NIR spectra. Each spectrum consisted in an average of 32 scans at 2 nm intervals within the wavelength range 1100–2500 nm. Four replicates of each sample were taken and the mean value was calculated. Fig. 1 shows the original NIR spectra of the samples collected at different dilution stages. Changes in several parts of the spectrum can provide information about the dilution process.

The spectral data were taken from the Vision software and treated using Unscrambler. Two segments of the spectrum were removed: the first, from 1880 to 2080 nm, due to the saturation of the spectrum caused by the strong combination band of O–H from water (1950 nm); and the second, from 2300 to 2500 nm, because of the low signal/noise ratio value. The mean values were exported for further mathematical pre-processing and calibration step.

### 2.5. Reference analyses

Regulatory analyses (total acids, volatile and non-volatile acids, solids and ash content) were analysed in accordance with the Official Methods of Analysis for vinegars (AOAC, 1995, Chap. 43).

Chloride was determined by direct potentiometric determination with a $Cl^-$ ion-selective electrode as a practical application of Nernst equation. It is based on the use of

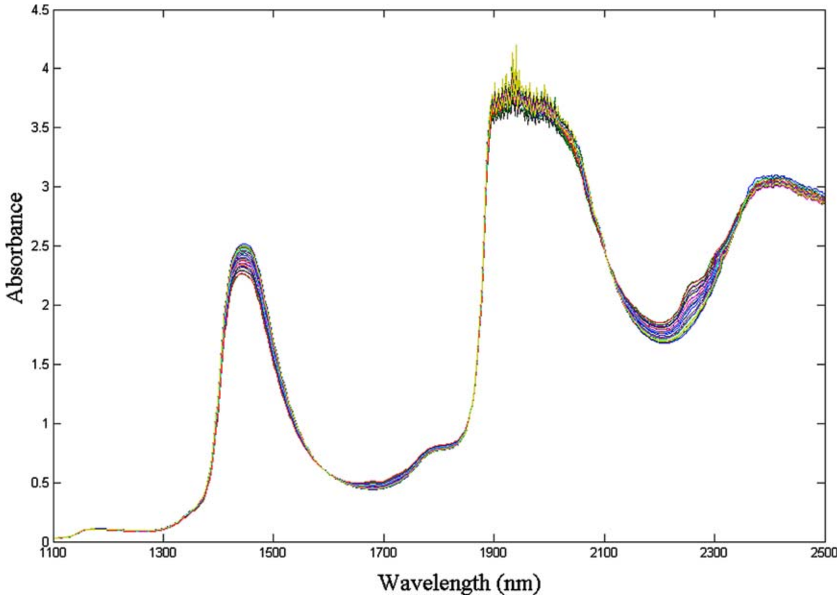

Fig. 1. Original NIR spectra of the vinegar samples.

small quantities of both sample and standard and the generation of a standard curve by serial addition.

L-proline and organic acids were determined by HPLC with DAD detector. All the standard solutions and vinegar samples were passed through a 0.7 μm glass microfibre GF/F supplied by Whatman (Whatman International Ltd., Maidstone, England) prior to injection into the HPLC System.

The column used was Zorbax SB-C18 with a stable bond packaging suitable for working at low pH values. The operating temperature was 25 °C.

The mobile phase was 0.009 M potassium di-hydrogen phosphate (adjusted to pH 2.06 with phosphoric acid)–methanol (92:8 v/v) at a flow-rate of $0.64\,\mathrm{mL\,min^{-1}}$ at 25 °C and a working pressure of 90 bar ($1\,\mathrm{bar} = 10^5\,\mathrm{Pa}$). Detection was performed by measuring UV absorption at 210 nm.

This chromatographic method was developed and optimised in the laboratory using experimental design and desirability functions as a multicriteria-decision-making method for the separation, determination and quantification of these compounds.

### 2.6. Chemometrics and data analysis

The matrix had 54 objects, 499 prediction variables (near-infrared absorbance values) and 14 response variables. The initial set of 54 objects was divided into two sets: the first was the training set used to construct the multivariate models; and the second the external test set, used to test the models and obtain RMSEP (root mean squared error in prediction) values expressing the prediction capacity of the computed models. The objects assigned to the external test set were randomly selected covering the experimental domain from the initial set of 54 samples and accounted for 18.5% of the total number of samples.

Three pre-processing methods (centering, multiplicative scattering correction (MSC) and orthogonal signal correction (OSC)) were applied to the data. MSC was used to estimate the error produced by scatter effects in the samples (Geladi, Macdougall, & Martens, 1985). OSC was applied to avoid all the effects and interferences in the response that presented zero correlation with the reference value $y$ (Blanco, Coello, Montoliu, & Romero, 2001; Svensson, Kourti, & MacGregor, 2002; Wold, Antic, Lindgren, & Öhman, 1998). In the present study, the original method proposed by Wold (Wold et al., 1998) and implemented in the *osccal* function of the PLS Matlab toolbox was used.

Recently, a number of theoretical approaches have addressed the selection of predictor strategies. This is necessary when there are a large number of variables, often highly correlated together with a moderate number of samples. SELECT was used as a predictor selection method. SELECT is a V-Parvus program that generates a set of decorrelated variables on the basis of their correlation coefficients with a response $y$. SELECT searches for the variable with the maximum correlation coefficient (correlation weight) with a response variable. This variable is selected and decorrelated from the other variables. SELECT then searches among the other variables with the maximum correlation coefficient until a specified number of variables are selected (Forina, Lantieri, Armanino, & Cerrato-Oliveros, 2003).

Calibrations were developed using partial least squares (PLS) (Næs, Isaksson, Fearn, & Davies, 2002, Chap. 5) with external test set to estimate prediction error. The quality of the results provided in the calibrations was compared using the root mean square error (RMSE) of the residuals obtained with the PLS model, defined as

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}, \tag{1}$$

where $y_i$ is the reference value, $\hat{y}_i$ the calculated value and $n$ the total number of samples. The RMSE is expressed as a percentage (in both calibration and prediction) taking into account the response range in its calculation

$$\mathrm{RMSE}(\%) = \frac{\mathrm{RMSE}}{(y_{i,\max} - y_{i,\min})} \times 100. \tag{2}$$

RMSE is termed root mean square error in calibration (RMSEC) for the calibration set and root mean square error in prediction (RMSEP) for the external prediction set.

Alternating conditional expectations (ACE) was used to detect a possible non-linearity within the experimental data. ACE is a non-linear regression method that can be applied when the ratio between the number of objects and the number of variables is very high (at least 10) (Massart et al., 1997, Chap. 11).

## 3. Results and discussion

The calibration of the NIR instrument consists on finding a relationship between the spectral variables and the reference values, so that, the resulting mathematical equations can be used to predict the different compounds in future new vinegar samples from their NIR measurements.

In a first step, partial least squares (PLS) regression was used to establish several models to quantify the fourteen components. The selected pre-processing methods were used, evaluated and compared in order to eliminate scatter effects and ensure that the spectra contained as little irrelevant information as possible.

### 3.1. Prediction of the physico-chemical parameters by partial least squares regression

Calibrations for individual responses were performed using three different pre-treatment combinations: centering, MSC-centering and centering-OSC. The results are shown in Table 2. These pre-processing methods were chosen according to the type of data obtained. Multicollinearity in NIR data should be taken into account, prompting

Table 2
Calibration and prediction errors for the 14 responses using different pre-processing methods and PLS

| Pre-treatment | Centering | | MSC-centering | | Centering-OSC | |
|---|---|---|---|---|---|---|
| Compound | RMSEC | RMSEP | RMSEC | RMSEP | RMSEC | RMSEP |
| Total acids (°) | 0.35(5) | 0.58(5) | 0.31(5) | 0.55(5) | 0.29(1.4) | 0.54(1.4) |
| Non-volatile acids (°) | 0.82(5) | 1.14(5) | 1.03(4) | 1.05(4) | 0.79(1.3) | 1.15(1.3) |
| Volatile acids (°) | 0.36(5) | 0.64(5) | 0.32(5) | 0.60(5) | 0.25(1.5) | 0.59(1.5) |
| Chloride (NaCl, g L$^{-1}$) | 0.27(5) | 0.36(5) | 0.31(5) | 0.35(5) | 0.24(1.4)<br>0.21(1.5) | 0.35(1.4)<br>0.27(1.5) |
| Solids (g L$^{-1}$) | 0.02(1) | 0.01(1) | 0.13(1) | 0.12(1) | 0.002(1.1) | 0.004(1.1) |
| Ash (g L$^{-1}$) | 0.29(5) | 0.43(5) | 0.37(4) | 0.54(4) | 0.24(1.4)<br>0.21(1.5) | 0.43(1.4)<br>0.39(1.5) |
| L-proline | 4.18(5) | 6.73(5) | 4.46(4) | 7.33(4) | 3.54(1.4) | 6.32(1.4) |
| L(+)-tartaric acid | 0.03(1) | 0.03(1) | 0.13(1) | 0.13(1) | 0.01(1.1) | 0.02(1.1) |
| L(−)-malic acid | 0.05(1) | 0.05(1) | 0.13(1) | 0.14(1) | 0.02(1.1)<br>0.02(1.2) | 0.05(1.1)<br>0.05(1.2) |
| Lactic acid | 0.43(5) | 0.66(5) | 0.44(5) | 0.65(5) | 0.37(1.3) | 0.60(1.3) |
| Acetic acid | 0.52(5) | 0.61(5) | 0.52(5) | 0.68(5) | 0.46(1.4)<br>0.40(2.2) | 0.56(1.4)<br>0.51(2.2) |
| Citric acid | 0.03(1) | 0.03(1) | 0.13(1) | 0.13(1) | 0.02(1.1)<br>0.03(1.3) | 0.03(1.1)<br>0.02(1.2) |
| Succinic acid | 0.50(5) | 0.71(5) | 0.52(5) | 0.68(5) | 0.44(1.3) | 0.63(1.3) |
| D-malic acid | 0.02(1) | 0.02(1) | 0.13(1) | 0.12(1) | 0.004(1.1) | 0.008(1.1) |

the need for data compression. The pre-processing methods studied here are the most commonly used when working with NIR data.

In Table 2, the RMSEC (%) and RMSEP (%) values are shown together with a selected number of the optimal number of components used in the PLS model. In each case, the selected components gave the minimal RMSEC and RMSEP values. When there was no agreement, selection was performed according to the minimal RMSEP value in order to avoid overfitting problems. The same pattern was followed for each of the three pre-treatment methods and for each of the 14 responses. In the case of OSC pre-treatment, the numbers in brackets represents the number of OSC components and the number of latent variables in the PLS model.

Comparisons in Table 2 are biased because there is no agreement between the number of components considered in each case. However, as indicated previously, this number was selected from RMSEP optimal values and gave an idea of the degree of goodness of the regression model.

The data in the table enable the following conclusions to be drawn. The OSC pre-treatment method minimised the RMSEP value in all cases except for L-proline. Only in the case of L(−)-malic acid were the results with OSC and column centering methods the same and significant improvement was negligible.

The MSC-column centering procedure did not show any significant improvement in the regressions. It may therefore be concluded that the scatter effect is not the dominating source of variability within the data matrix and MSC pre-treatment could be removing some of the chemical information contained in the data matrix. Thus, in this system there is no need to include a scatter effect correction in the data if the samples have been filtered prior to the NIR measurement stage.

L-proline was the response with the highest number of calibration and prediction errors and a new combination of pre-treatment methods was tried in order to determine whether noticeable improvements were obtained. First derivative by Norris algorithm and column centering and first derivative followed by OSC were performed prior to PLS regression. However, in both cases calibration and prediction errors were higher than the results achieved with the previous pre-treatment strategies. Therefore, two alternative approaches were tested: a non-linear model to perform the calibration and a selection of predictors from the spectral data.

### 3.2. Alternating conditional expectations

ACE was used to determine whether a non-linear model was more appropriate to fit L-proline data. Different SPAN values were tested between 0.1 and 1 (linear model). ACE was performed over the first five principal components computed out of the data matrix of 499 NIR variables to achieve the ratio number of samples-number of predictors required by ACE algorithm. Table 3 shows the different RMSEC and RMSEP values (in

Table 3
Calibration and prediction errors for L-proline using ACE

| SPAN | RMSEC (%) | RMSEP (%) |
|---|---|---|
| 0.2 | 4.39 | 9.71 |
| 0.5 | 5.85 | 9.05 |
| 0.7 | 5.98 | 8.47 |
| 0.95 | 6.17 | 8.31 |

percentages) together with the SPAN value. The RMSEC values increase as the model gets closer to a linear model and the RMSEP values decrease as the SPAN value becomes larger. The results obtained showed that the non-linear model was not suitable for describing the behaviour of L-proline data. Moreover, ACE modelling produces overfitting as the SPAN value decreases; the calibration error decreases but, at the same time, the prediction error increases sharply.

### 3.3. SELECT and PLS regression

SELECT was used to obtain a decorrelated matrix of the predictor variables directly related to L-proline response. SELECT was performed over the whole matrix of the NIR variables using column centering as a pre-treatment method and selecting a maximum of 20 predictors. The program selected eight wavelengths from the total number of 499 near-infrared absorbances collected: 1144, 1148, 1270, 1278, 1526, 1534, 1608 and 2112 nm, which were supposedly related to L-proline values.

These selected bands correspond to the second overtone bands of the bond C–H (1144–1148, 1270–1278 nm, stretching mode), the first overtone of the bond O–H (1526 nm, stretching mode) and the first overtone of the bond N–H (1534 nm, stretching mode). The band 1608 nm, can correspond to the O–H stretching first overtone (intermolecular H-bond) and the last band (2112 nm) corresponds to aminoacid compounds (N–H stretching and C=O stretching) (Osborne, Fearn, & Hindle, 1993, Chap. 2).

PLS was performed over the eight predictors previously selected and once again performing column centering as pre-processing method. The results with four and five latent variables are shown in Table 4. This shows how the results with SELECT followed by PLS regression with five latent variables gave the best result for modelling L-proline response with a RMSEP (in percentage) of 4.81%.

Table 4
Calibration and prediction errors for L-proline using SELECT and PLS regression

| LVs | RMSEC (%) | RMSEP (%) |
|---|---|---|
| 4 | 9.20 | 12.78 |
| 5 | 3.57 | 4.81 |

### 3.4. Evaluation of the quality of commercial vinegar samples

A direct application of the calibration models developed before was done studying the prediction results for 10 new commercial white wine vinegar samples. OSC with the elimination of one orthogonal component was the pre-treatment used in the development of the regression models due to its ability to yield in general the best prediction results. From the values obtained it is possible to make some observations about the quality of the wine vinegar analysed.

All the samples fulfilled the specifications of the legal regulations about total acidity ranging from 6.05 to 6.25°. Solids content per liter and acetic grade ranged from 1.42 to 1.48 g. Ash content was larger than 1 g $L^{-1}$ in all cases; chloride concentration was from 0.32 to 0.34 g $L^{-1}$ expressed as sodium chloride and citric acid ranged from 0.58 to 0.61 g $L^{-1}$. Table 5 shows the complexity of the regression methods applied and the ranges of the parameters obtained in the analysis of the commercial samples. Regarding the organic acids, some of them are indicative of the origin and genuineness of the vinegar. For instance, values of tartaric acid between 0.45 and 1.15 g $L^{-1}$ can be considered normal (Sáiz-Abajo, González-Sáiz & Pizarro, JAFC, 2004) and values under 0.45 g $L^{-1}$ are indicative of an extra dilution process. In the case of the vinegar samples studied, the concentration of tartaric acid ranged from 0.73 to 0.76 g $L^{-1}$. The concentration of citric acid is below the maximum legal limit and the rest of parameters are within the normal ranges for wine vinegar.

It is important to note that the calibration developed and validated in this study is applicable only to the quality control and dilution process of white wine vinegar and it is potentially extendable to red wine vinegar with the corresponding validation process. When handling with vinegar from other origins (i.e., cider vinegar, malt vinegar, balsamic vinegar, Sherry vinegar), the calibration stage should be performed with samples representing as best as possible the future vinegar samples to be analysed.

Table 5
Ranges of concentration of the different parameters in the commercial vinegar samples

| Parameter | #PLS components | Range |
|---|---|---|
| Total acids (°) | 4 | 6.05–6.25 |
| Non-volatile acids (°) | 3 | 0.07–0.09 |
| Volatile acids (°) | 5 | 5.97–6.13 |
| Chloride (g $L^{-1}$ of NaCl) | 5 | 0.31–0.34 |
| Solids (g $L^{-1}$ per acetic grade) | 1 | 1.42–1.48 |
| Ash (g $L^{-1}$) | 5 | 1.06–1.11 |
| L-proline (g $L^{-1}$) | 4 | 1.02–1.63 |
| L(+)-tartaric acid (g $L^{-1}$) | 1 | 0.73–0.76 |
| L(−)-malic acid (g $L^{-1}$) | 1 | 0.31–0.32 |
| Lactic acid (g $L^{-1}$) | 3 | 0.55–0.60 |
| Acetic acid (g $L^{-1}$) | 4 | 56.46–60.71 |
| Citric acid (g $L^{-1}$) | 2 | 0.58–0.61 |
| Succinic acid (g $L^{-1}$) | 3 | 0.49–0.54 |
| D-malic acid (g $L^{-1}$) | 1 | 0.021–0.022 |

## 4. Conclusion

A method based on near-infrared spectroscopy has been developed for determining 14 parameters that influence the quality of wine vinegar.

OSC has proven to be the best pre-processing method for the spectral data in the calibration stage and the achieved prediction errors ranged from 0.008% for D-malic acid to 1.15% for non-volatile acids. Only L-proline presented slightly greater prediction errors that were reduced using SELECT as a predictor selection method.

This study demonstrates the suitability of NIR spectroscopy for monitoring different industrial processes such as dilution and quality control of the final product in the vinegar industry. NIR spectroscopy constitutes an important tool for assuring the genuineness of the products and for compliance with required national and foreign market regulations, which is nowadays a hugely competitive task.

## Acknowledgements

The authors thank the Spanish Government (Ministerio de Ciencia y Tecnología, INIA CAL01-053), the Local Government of La Rioja (Consejería de Educación, Juventud y Deportes for the research grant FPI-2001–2004) and the University of La Rioja for their financial support. We also thank Professor M. Forina for providing the latest version of V-Parvus and Daniel Gomez for the assistance in the experimental part.

## References

Anklam, E., & Battaglia, R. (2001). Food analysis and consumer protection. *Trends in Food Science and Technology, 12*(5–6), 197–202.

AOAC. (1995). *Official methods of analysis* (16th ed., pp. 10–11) 930.35.

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: a rapid-response analytical tool. *Trends in Analytical Chemistry, 21*(4), 240–250.

Blanco, M., Coello, J., Montoliu, I., & Romero, M. A. (2001). Orthogonal signal correction in near infrared calibration. *Analytica Chimica Acta, 434*(1), 125–132.

Büning-Pfaue, H. (2003). Analysis of water in food by near infrared spectroscopy. *Food Chemistry, 82*(1), 107–115.

Coffey, C., Cooley, B. E., & Walker, D. S. (1999). Real time quantitation of a chemical reaction by fiber optic near-infrared spectroscopy. *Analytica Chimica Acta, 395*(3), 335–341.

Forina, M., Lantieri, S., Armanino, C., & Cerrato-Oliveros, C. V-PARVUS 2003: an extendable package of programs for data explorative analysis, classification and regression analysis, Genoa, Italy: Dipartimento di Chimica e Tecnologie Farmaceutiche ed Alimentari.

Garrido-Varo, A. (1997). Current and future applications of NIRS technology in the feed industry. In Morand-Fehr P. (Ed.). *Feed manufacturing in Southern Europe: New challenges* (pp. 87–92). Zaragoza: CIHEAM-IAMZ. Cahiers Options Méditerranéennes; v. 26, South European Feed Manufacturers Conference, 1996/05/09–11, Reus (Spain).

Geladi, P., Macdougall, D., & Martens, H. (1985). Linearization and scatter-correction for NIR reflectance spectra of meat. *Applied Spectroscopy, 39*, 491–500.

Larrechi, M. S., & Callao, M. P. (2003). Strategy for introducing NIR spectroscopy and multivariate calibration techniques in industry. *Trends in Analytical Chemistry, 22*(9), 634–640.

Massart, D. L., Vandeginste, B. G. M., Buydens, L. M. C., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1997). Data handling in science and technology $20^{a}$. In B. G. M. Vandeginste & S. C. Rutan (Eds.). *Handbook of chemometrics and qualimetrics: Part A* (pp. 329). Sara Burgerhartstraat, Amsterdam: Elsevier Science.

MATLAB (2001). The Language of Technical Computing. Version 6.1.0.450 Release 12.1, May 18, 2001. The MathWorks, Natick, MA.

Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification* (p. 33). Chichester, UK: NIR Publications.

Osborne, B. G., Fearn, T., & Hindle, P. H. (1993). *Practical NIR Spectroscopy with applications in food and beverages analysis* (pp. 23–33) (2nd ed.). Harlow, UK: Longman Scientific & Technical.

Royal Decree No. 2070/1993 of November 26th. *Official State Gazette*, issue no. 293, dated 8th December 1993.

Sáiz-Abajo, M. J., González-Sáiz, J. M., & Pizarro, C. (2005). Multiobjective optimisation strategy based on desirability functions used for chromatographic separation and quantification of L-proline and organic acids in vinegar. *Analytica Chimica Acta, 528*, 63–76.

Sáiz-Abajo, M. J., González-Sáiz, J. M., & Pizarro, C. (2004). Classification of wine and alcohol vinegar samples based on Near-Infrared spectroscopy. Feasibilily study on the detection of adulterated vinegar samples. *Journal of Agricultural and Food Chemistry, 52*, 7711–7719.

Shenk, J. S., & Westerhaus, M. O. (1995). Analysis of agricultural and food products by near infrared reflectance spectroscopy. Monograph, NIR Systems, Inc., 12101 Tech Road, Silver Spring, MD 20904, PN IS-0119.

Svensson, O., Kourti, T., & MacGregor, J. F. (2002). A comparison of orthogonal signal correction algorithms and characteristics. *Journal of Chemometrics, 16*, 176–188.

The Unscrambler v7.6, Copyright 1986–2000 CAMO ASA. Nedre Vollgate 8, N-0158 OSLO, Norway. Phone: +47 22 39 63 00 Fax: +47 22 39 63 22. E-mail: camo@camo.no Support: support@camo.no, Web: http://www.camo.com.

Wold, S., Antic, H., Lindgren, F., & Öhman, J. (1998). Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems, 44*(1–2), 175–185.

Xiccato, G., Trocino, A., Tulli, F., & Tibaldi, E. (2004). Prediction of chemical composition and origin identification of european sea bass (*Dicentrarchus labrax L.*) by near infrared reflectance spectroscopy (NIRS). *Food Chemistry, 86*(2), 275–281.