

Quantitative Analysis of Individual Sugars and Acids in Orange Juices by Near-Infrared Spectroscopy of Dry Extract

Weijie Li, Pierre Goovaerts,[†] and Marc Meurens*

Unité de Biochimie de la Nutrition, Faculté des Sciences Agronomiques, Université Catholique de Louvain, Place Croix du Sud 2/8, Louvain-la-Neuve B1348, Belgium

The combination of near-infrared spectroscopy (near-IR) and multivariate calibration for determination of glucose, fructose, sucrose, and citric and malic acids in orange juices was investigated. The concentrations of these components analyzed by enzymatic assays were considered as references relative to near-IR spectroscopy. Dry extract spectra of 218 orange juice samples were recorded in transmission mode between 1100 and 2500 nm. The original near-IR spectral data could be improved by mathematical pretreatments such as derivative transformations or multiplicative signal correction. Stepwise multiple linear regression (SMLR) and partial least-squares regression (PLSR) were used to create calibration models relating chemical reference values to spectral data. The prediction ability of calibration models is acceptable in comparison with the reference methods. The calibration and validation results provided by PLS-1 calibration models are slightly better than those obtained with SMLR calibration models.

Keywords: *Sugars; acids; orange juices; near-infrared spectroscopy; multivariate calibration*

INTRODUCTION

The orange juice market alone in Europe and the United States was estimated to be in excess of £4 billion (Robertson, 1988). Thus, it becomes increasingly important to have adequate analytical techniques for quantitative analysis of major components in orange juices. The conventional laboratory analysis of fruit juices such as wet chemistry testing is generally laborious and time consuming. In contrast, near-infrared (near-IR) spectroscopy has proved to be a rapid and accurate technique for a wide range of analytical applications and has emerged as an interesting alternative to wet chemistry in analysis and quality control of agricultural and food products (Osborne and Fearn, 1986; Barton, 1987). Near-IR spectroscopy has been used for determining individual sugars in dry fruit model systems by Giangiacomo et al. (1981). Lanza and Li (1984) have performed fruit juice analysis using near-IR transmittance, yet they determined only total sugars. In this paper, the performance of near-IR spectroscopy for determining individual sugars and acids is investigated.

In the analysis of aqueous solutions, such as fruit juices, overlap of the vibrational bands of water with those of the solutes is inevitable, resulting in broad contours that cannot usually be deconvoluted into their constituents (Fischer et al., 1994). In this case, near-IR spectroscopy can be made more sensitive and more accurate by placing the liquid sample on a fiberglass support and eliminating water until dry extract is obtained. This method has been developed by Meurens et al. (1982, 1990) to analyze aqueous solutions. Studies on the analysis of synthetic solutions of sugars and of a limited number of fruit juice samples by near-IR reflectance of dry extract on fiberglass filters have been

published (Alfaro et al., 1990, 1991). The present aim is to report the analytical results of near-IR spectroscopy obtained by transmittance of dry extract on a large collection of orange juice samples.

For each particular component, near-IR spectroscopy requires calibration of the instrument by a recognized laboratory method known as a reference method. In this study, individual sugars and acids in orange juices were determined by enzymatic methods. The stepwise multiple linear regression (SMLR) and the partial least squares regression (PLSR) algorithms can be used to establish mathematical relations (i.e., calibration models) between chemical reference values and spectral data for each component being measured. The SMLR calibration models are built up using a stepwise procedure that selects wavelengths as long as their contribution to the model is statistically significant (Mark, 1992). In contrast, the PLSR algorithm is a full-spectrum method that provides extensive possibilities for extracting information, both with respect to the variables and to the objects, to make the calibration models (Martens and Naes, 1989). In this paper, both SMLR and PLSR methods were applied to orange juice analysis by near-IR spectroscopy and their results were compared.

Mathematical pretreatments of near-IR spectra should enhance the qualitative interpretation of spectra and the prediction ability of calibration models. Derivative transformations could partially compensate for baseline offset between samples and reduce instrument drift effects (Norris, 1982). Another property of these transformations is to invert the spectrum so that the peaks become narrow valleys (Shenk et al., 1992). In near-IR measurement, light scattering and sample thickness have multiplicative effects whereas chemical absorption has additive effects. The multiplicative signal correction (MSC) is a kind of standardization of spectral data that allows the separation of additive and multiplicative effects from each other (Martens and Naes, 1989). In this paper, effects of mathematical pretreatments of the spectral data of dry extracts of orange juices on the calibration results were also investigated.

* Author to whom correspondence should be addressed.

[†] Unité de Biométrie, Faculté des Sciences Agronomiques, Université Catholique de Louvain, Place Croix du Sud 2/16, Louvain-la-Neuve B1348, Belgium.

MATERIALS AND METHODS

Samples. Orange juices, including single strength juices and pulp washes (120 samples) and concentrates (98 samples), were collected from different countries in Europe, Africa, and America. These samples had been previously analyzed by Schutzgemeinschaft der Fruchtsaftindustrie within the frame of the FLAIR research project of the Commission of the European Communities on new technology and laboratory methodology for glucose, fructose, sucrose, and citric and malic acids by enzymatic methods (Boehringer, Mannheim, Germany). These results were used as chemical reference values in contrast with near-IR spectral data for the multivariate calibration.

Prior to near-IR spectroscopy, the concentrates were diluted to 11.18° Brix (w/w) with distilled water, whereas the single strength juices and pulp washes were homogenized by a Kontes homogenizer. Then each orange juice sample (0.6 mL) was pipetted and deposited on a fiberglass disk (Millipore AP40047) inside a centralizing device and was dried for 4 min in a special dryer, named DESIR (Dry extract system for infrared), unit. This DESIR unit, designed by NIRSystems (Perstorp Analytical Co.), is a thermostat-controlled fan oven operating at temperatures between 45 and 50 °C (Meurens and Alfavo, 1990). Immediately after drying, the fiberglass filter was placed between two glass windows in a sample cup for near-IR scanning.

Near-IR Spectroscopy. All near-IR measurements were made with a Pacific Scientific spectrometer Model 6250 in combination with an IBM computer Model PS/2 using the NSAS software version 3.30 of NIRSystems. This spectrometer has a single-beam scanning monochromator that provides a linear scan over the 1100–2500 nm range at intervals of 2 nm. The absorbances of orange juice samples were recorded in the transmission mode as $\log(1/T)$ (T = transmittance).

The spectrum of each orange juice sample was calculated as the average of six spectra, which were obtained using two fiberglass disks and, for each disk, three positions (0°, 120°, and 240°) of the sample cup inside the spectrometer. The measurement of each spectrum was separately scanned 10 times from 1100 to 2500 nm. The reference spectrum was obtained by scanning a blank fiberglass disk.

Database Analysis. The acquisition of near-IR spectra for a set of calibration samples (called the calibration set) allows one to establish calibration models relating spectral data to chemical reference values for each component being analyzed. Such calibration models can be checked using a set of validation samples (called the validation set) that are not part of the calibration set but are similar to those with respect to constituent ranges, physical state, origin, etc. The check proceeds as follows: (1) spectral data are collected and interpolated into a calibration model to predict concentration values and (2) these estimates are then compared with chemical reference values in terms of residual variance and correlation coefficient.

In this study, the 218 orange juice samples were split into two groups of 150 and 68 samples, respectively. The first group of 150 samples was used to establish the calibration models for determination of glucose, fructose, sucrose, and citric and malic acids. The validity of these models was assessed using the remaining 68 samples. Note that the samples were randomly assigned to either the calibration set or the validation set, except for the samples with maximum and minimum values for each component that were put into the calibration set so as to span the complete range of component concentrations during the calibration modeling. Tables 1–3 show for each sample set the statistics of component concentrations measured by enzymatic methods.

The SMLR algorithm is included in the software package NSAS 3.30, whereas the PLSR algorithm is applied by the software package Unscrambler 5.5 of CAMO A/S (Trondheim, Norway). The PLSR method in Unscrambler allows efficient detection of outliers that are not representative of the calibration and validation sets (Unscrambler, 1994). Outliers were detected by looking at residuals provided by a calibration

Table 1. Statistics of Component Concentrations Measured by Enzymatic Assays for the Total Sample Set (218 Samples) (Unit = g/L)

	glucose	fructose	sucrose	citric acid	malic acid
min value	4.10	4.70	6.50	0.80	0.28
max value	39.30	44.00	95.20	23.40	3.41
average	20.61	22.45	37.28	9.78	1.55
SD ^a	5.05	5.39	12.97	4.68	0.59

^a SD = standard deviation.

Table 2. Statistics of Component Concentrations Measured by Enzymatic Assays for the Calibration Set (150 Samples) (Unit = g/L)

	glucose	fructose	sucrose	citric acid	malic acid
min value	4.10	4.70	6.50	0.80	0.28
max value	39.30	44.00	95.20	23.40	3.41
average	20.59	22.40	36.32	9.79	1.53
SD	5.50	5.82	13.54	4.91	0.60

Table 3. Statistics of Component Concentrations Measured by Enzymatic Assays for the Validation Set (68 Samples) (Unit = g/L)

	glucose	fructose	sucrose	citric acid	malic acid
min value	10.60	11.60	9.50	3.50	0.47
max value	29.30	34.50	71.70	23.20	3.10
average	20.68	22.55	39.36	9.74	1.59
SD	3.91	4.34	11.45	4.15	0.57

Table 4. Number of Outliers Detected by Multivariate Calibration in Quantitative Analysis of Orange Juices

	number of outliers for				
	glucose	fructose	sucrose	citric acid	malic acid
calibration	9	6	9	6	7
validation	3	5	4	7	8

model and identifying points that are both suspect and influential on the score plot of the first two principal components. The largest outliers had been identified after preliminary calibration, and they were eliminated first. The calibration modeling was repeated until all outliers had been removed. The same outliers were discarded for SMLR algorithm to be compared with the PLSR in the orange juice analysis. Table 4 shows the number of outliers detected by the calibration model for the quantification of individual sugars and acids in the orange juices.

Statistics. The performance of a calibration model to predict component concentrations based on spectral data can be assessed using the prediction error and the correlation coefficient between estimates and chemical reference values. The three following statistics (Martens and Naes, 1989) were used in this paper:

1. Root mean square error of prediction (RMSEP) is computed as the square root of the average squared difference between chemical reference and predicted values

$$\text{RMSEP} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (1)$$

where n is the number of validation samples, \hat{y}_i is the near-IR predicted value and y_i is the chemical reference value provided by the reference method.

2. Standard error of prediction (SEP) is the square root of the prediction variance and is computed as

$$\text{SEP} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\hat{y}_i - y_i - \hat{\text{Bias}})^2} \quad (2)$$

where Bias is estimated by

$$\hat{\text{Bias}} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)$$

3. Correlation coefficient between predicted and chemical reference values

$$r = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{[\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 (y_i - \bar{y})^2]^{1/2}} \quad (3)$$

where $\bar{\hat{y}}$ = average value of \hat{y}_i , \bar{y} = average value of y_i .

RESULTS AND DISCUSSION

Mathematical Pretreatments of Near-IR Spectra. The objective of mathematical pretreatments of original spectra is 2-fold: (1) make all major interferences vary as independently as possible of each other in the samples and (2) reduce the influence of light scattering variation with a potential improvement of calibration results.

Derivative transformations of spectra aim to glean additional information from the original spectra and to avoid somewhat band interference and overlap when several components are measured simultaneously. Figure 1 shows the spectrum of a dry extract of orange juice before and after derivative transformations. The original spectrum $\log(1/T)$ has broad overlapping absorption bands. The first and second derivative transformations sharpen these absorption bands and so facilitate their separation. Either the first or second derivative can be used, yet the correlation plot is more easily interpreted for the second derivative results because the spectrum has minimum values at the wavelengths at which the $\log(1/T)$ spectrum has maximum values. A shortcoming of the second derivative transformation is that two false valleys are generated in the positive ordinate scale for every band in a negative direction (Hruschka, 1992).

Multiplicative Signal Correction (MSC) of spectra allows the separation of the multiplicative effects due to light scattering and sample thickness from the additive effects due to the chemical absorption. Figure 2a shows the original spectra for some orange juice samples. Multiplicative and additive effects lead to the baseline shifts of the original spectra. In this study, the light absorbances within the wavelength interval 1900–2000 nm (the main water absorption peak) are independent of the component being measured, and so they provide no relevant information. Consequently, the wavelengths from 1900 to 2000 nm were chosen as base variables for MSC standardization, and the original spectra were corrected until all spectra have similar absorption scatter level (Figure 2b).

Selection of Mathematical Pretreatments. Different calibration models, which were established on the basis of different transformed spectra of orange juices, were appraised by prediction error. For example, Figure 3 shows the estimated prediction error provided by different calibration models to determine glucose. Mathematical pretreatments tend to reduce the mean square error relative to original spectral data, better results being obtained for the first and second derivative transformations. These results confirm the utility of pretreatments prior to calibration. The selection of the mathematical pretreatments usually depends on final regression results. In the example above, the first derivative was chosen because it yields the lowest prediction error.

Calibration and Prediction. In this section, we discuss and compare SMLR and PLSR performances to determine the concentrations of individual sugars and acids in orange juices.

SMLR algorithm has been the cornerstone of near-IR spectroscopy for more than 30 years and is based on the forward calibration model,

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F} \quad (4)$$

where \mathbf{Y} is the component concentration matrix, \mathbf{X} is the absorbance data matrix, and \mathbf{F} is the residual

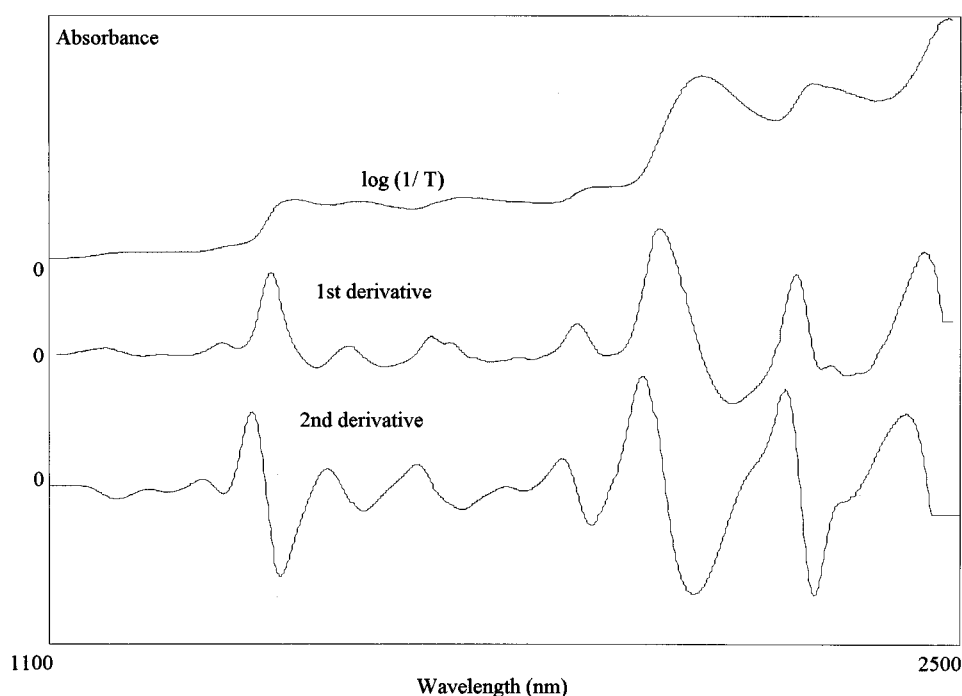


Figure 1. Spectra of a dry extract of orange juice sample before and after derivative transformations.

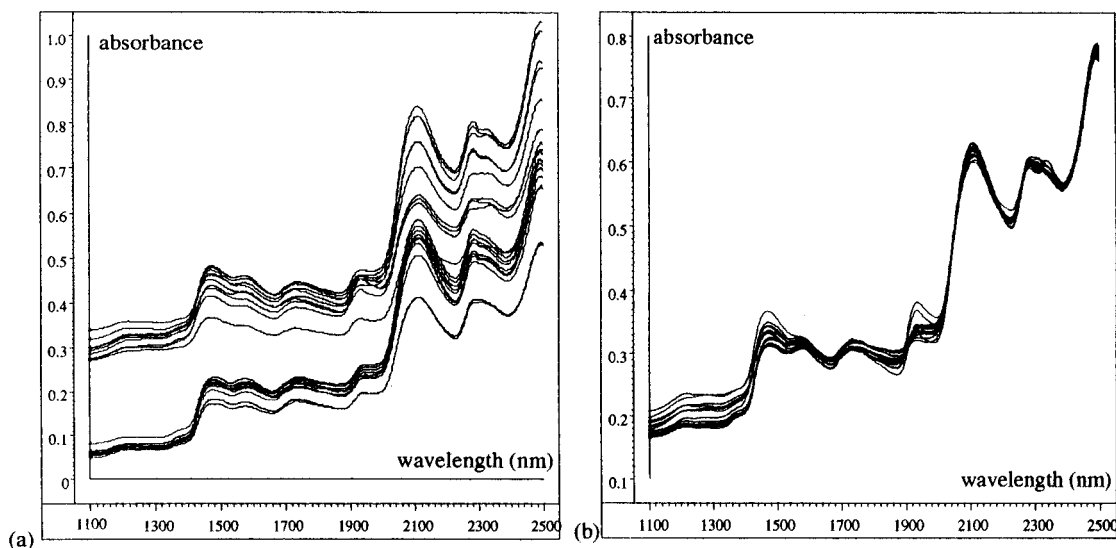


Figure 2. Comparison between the original and MSC standardized spectra of orange juices: (a) original spectra and (b) MSC standardized spectra.

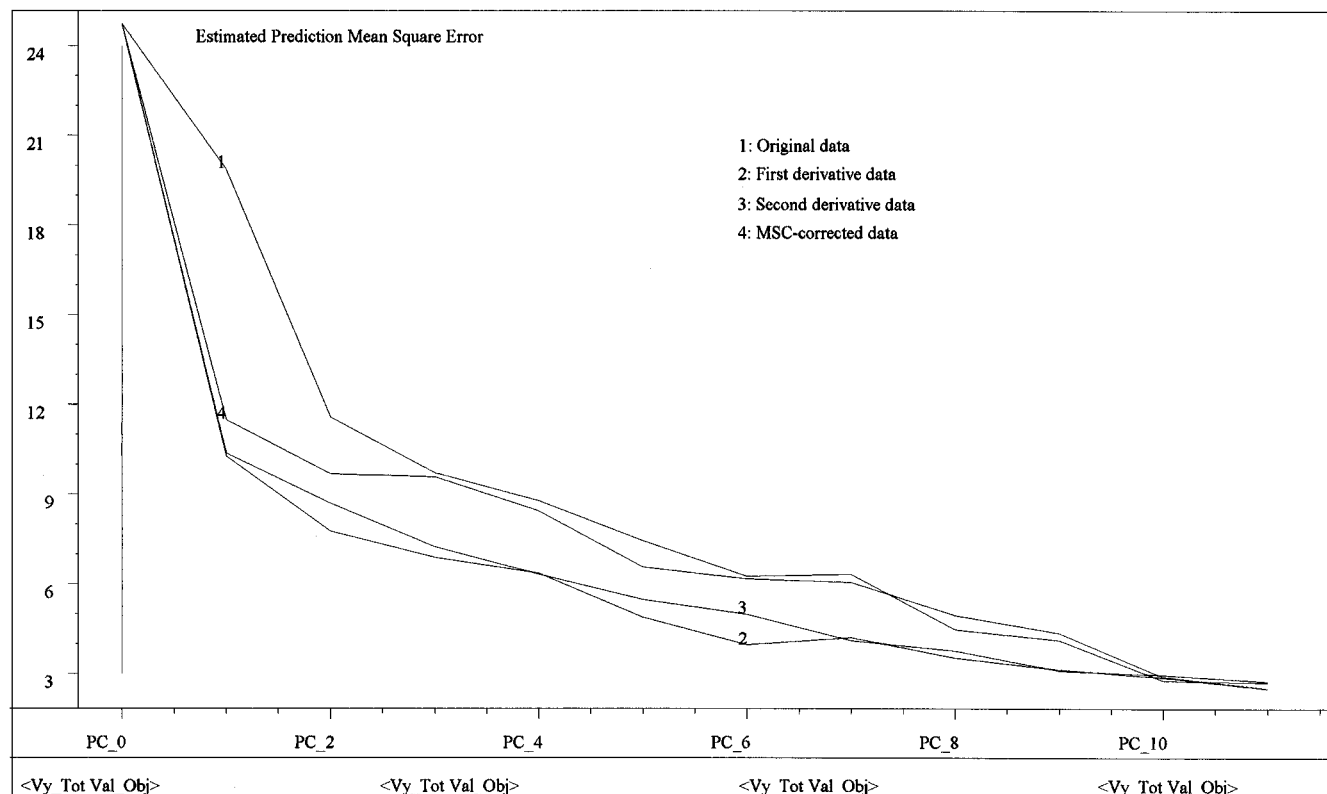


Figure 3. Impact of mathematical pretreatments of the original near-IR spectral data on the prediction error of glucose concentration in orange juices.

matrix. A calibration model in SMLR is built up by adding wavelengths one at a time, each wavelength being chosen so that the resulting model yields the smallest residual sum of squares possible. The starting point for such a stepwise procedure is usually the wavelength that gives the best one-term model. In this study four terms have been included in calibration models for the individual sugars and acids of orange juice samples so as to obtain a good compromise between thoroughness and efficiency.

The prediction ability of SMLR calibration models is assessed using the statistics (1–3), and their values are listed in Table 5. For each sugar and acid, results relate to the mathematical pretreatment that provides the best reestimation scores, see Table 5 (right column). The

large values of the correlation coefficients indicate that the values predicted by the calibration models based on spectral data fit well the chemical reference values.

PLSR algorithm belongs to the class of bilinear regression methods that could approximate complicated multivariate input data by a few bilinear PLSR factors **T** or principal components (PCs), which account for most of the variance of all the variables. The full bilinear calibration model can be written as

$$\begin{aligned} \mathbf{X} &= \mathbf{TP} + \mathbf{E} \\ \mathbf{Y} &= \mathbf{TQ} + \mathbf{F} \end{aligned} \quad (5)$$

where the loading matrix **P** represents the regression

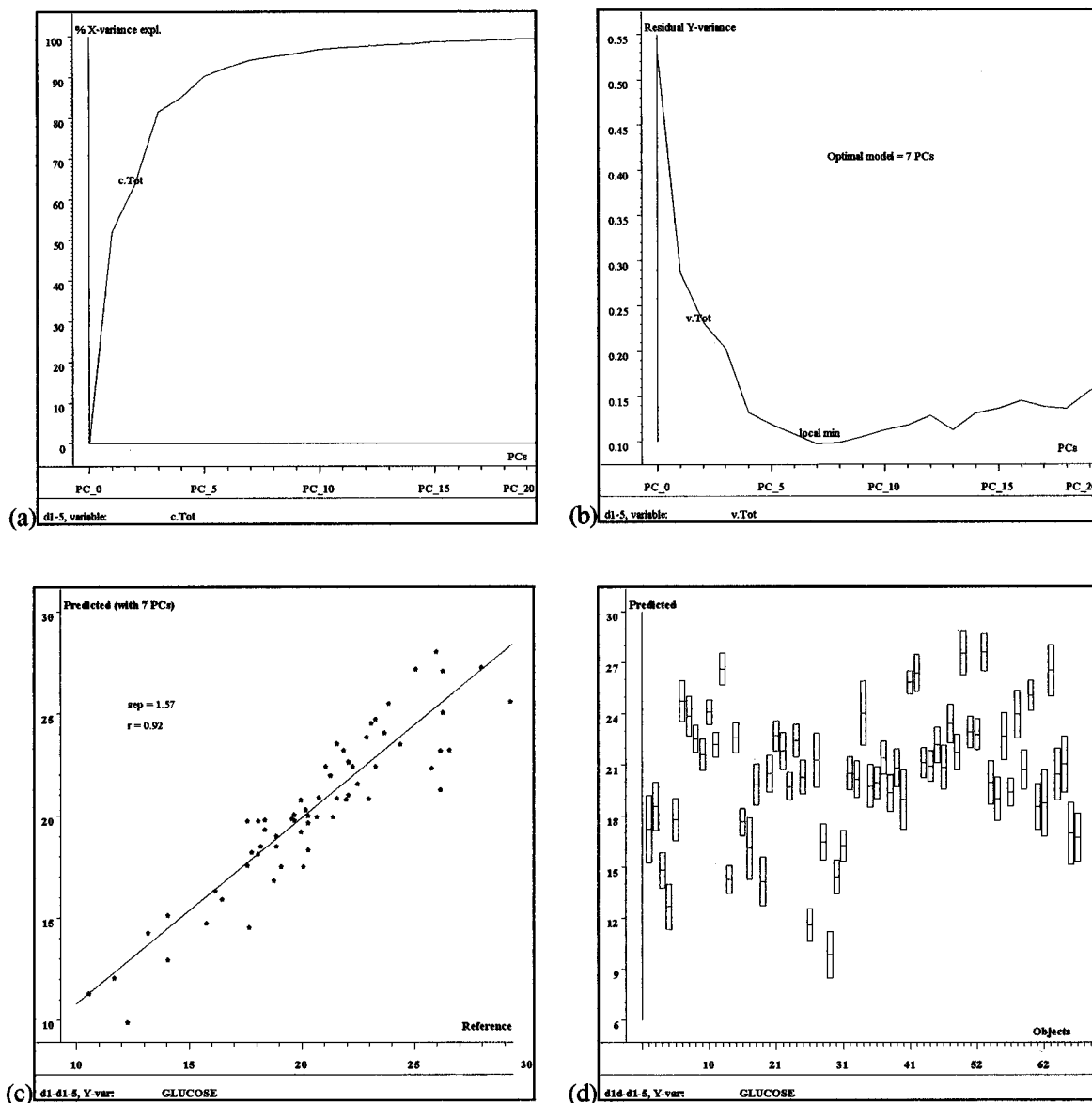


Figure 4. Calibration and prediction results provided by the PLS-1 algorithm for determination of glucose in orange juices: (a) calibration variance, (b) validation variance, (c) predicted vs reference values for validation set, and (d) predicted values with deviations for prediction set.

Table 5. Prediction Results of SMLR Calibration Models for the Orange Juice Analysis

	RMSEP (g/L)	SEP (g/L)	<i>r</i>	wavelength (nm)	method
glucose	1.80	1.72	0.88	2240/2096/2306/2226	D1OD ^a
fructose	1.93	1.82	0.90	1882/2278/2176/2300	MSC
sucrose	3.42	3.06	0.95	2272/2194/2110/1844	D1OD
citric acid	1.18	0.77	0.95	1734/2342/2316/1894	D2OD ^b
malic acid	0.26	0.21	0.82	2274/1936/2294/1462	D1OD

^a First derivative. ^b Second derivative.

coefficients of **X** on **T** in the same way as **Q** represents the regression coefficients of **Y** on **T**. The residual matrices **E** and **F** represent the variation in **X** and **Y** that are not explained by the bilinear structure, respectively.

Two PLS algorithms can be distinguished: PLS-1 handling only one **Y**-variable at a time, and PLS-2 handling several **Y**-variables simultaneously. PLS-2 regression is useful for a preliminary overview in exploitative data analysis. When nonlinearity, rather than random noise, is the major problem, a better approach would consist of performing PLS-1 on each

Y-variable separately (Martens and Naes, 1989). It is the reason why only the PLS-1 calibration models are used hereafter.

Similar to SMLR calibration the prediction ability of PLS-1 calibration models with the best mathematical pretreatments are summarized using three statistics listed in Table 6. Comparison of Tables 5 and 6 shows that PLS-1 regression generally provides better prediction of individual sugar and acid concentrations in orange juices.

An important parameter of the PLS-1 algorithm is the number of PCs that are included in the calibration model: using too few PCs can leave important near-IR structure unmodeled whereas using too many PCs draws too much measurement noise from the **X** and **Y** data into the calibration model. The optimal number of PCs can be chosen as the number associated with the first local minimum in the plot of the validation variance as shown in Figures 4b and 5b. Higher order PCs give a rise in the residual variance and hence are considered as irrelevant.

Figure 4 shows calibration and prediction results for the determination of glucose in orange juices. Glucose

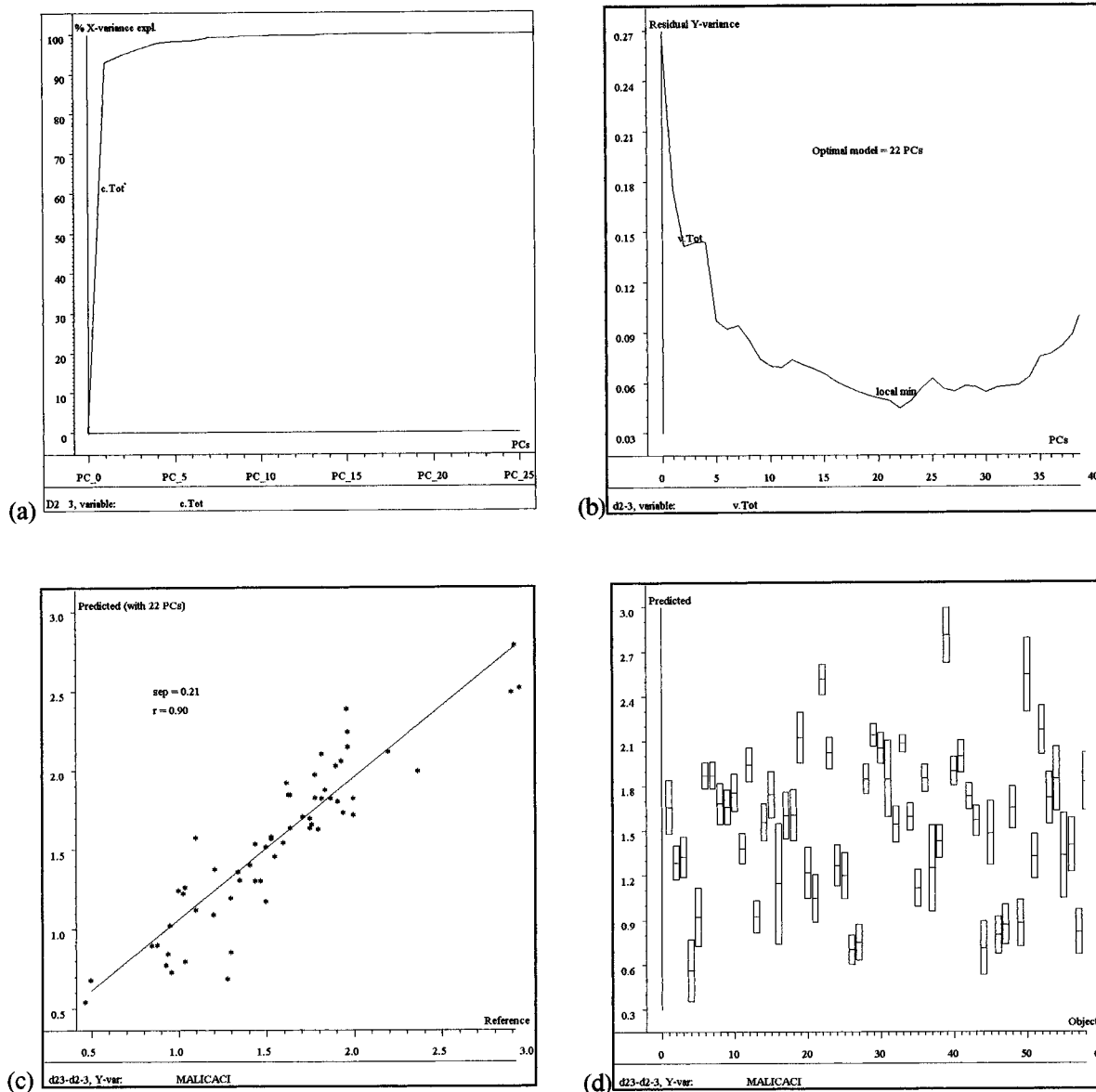


Figure 5. Calibration and prediction results provided by the PLS-1 algorithm for determination of malic acid in orange juices: (a) calibration variance, (b) validation variance, (c) predicted vs reference values for validation set, and (d) predicted values with deviations for validation set.

Table 6. Prediction Results of PLS-1 Calibration Models for the Orange Juice Analysis

	RMSEP (g/L)	SEP (g/L)	<i>r</i>	PCs	method
glucose	1.57	1.57	0.92	7	D1OD
fructose	1.53	1.54	0.93	12	MSC
sucrose	3.20	3.02	0.96	9	D2OD
citric acid	0.92	0.91	0.98	14	MSC
malic acid	0.20	0.21	0.90	22	D2OD

is one of the major components in the orange juices, and its average concentration in the validation set is 20.68 g/L. According to the criterion in the previous paragraph, seven PCs have been chosen as the optimal number of PCs to retain in the final model for the prediction of glucose (Figure 4b); they explain 95% of the variance of spectral measurements (Figure 4a). The prediction ability of the optimal calibration model is illustrated in Figure 4c,d. Figure 4c shows a good correlation between predicted values and known chemical reference values. The reliability of the prediction is illustrated in Figure 4d that shows, for each of the validation samples, the predicted value and the corresponding deviation: actual value = predicted value \pm

deviation. Values of the deviation are computed by an empirical formula that accounts for the Y-residual variance in the validation set from the calibration step, X-residual variance in the prediction object and the average X-residual variance in the validation set (Unscrambler, 1994).

Figure 5 shows calibration and prediction results provided by PLS-1 algorithm for the determination of malic acid in orange juices. The average concentration of malic acid (1.59 g/L) in the validation set is smaller than for other components, see Table 4. Figure 5a indicates that a large proportion of X-variance in the calibration set is explained by the PLS-1 calibration model. However, the variance curve in Figure 5b shows that the validation residual variance reaches its first local minimum only after 22 PCs. Clearly, the explanation of original variance by the optimal model for malic acid needs more PCs extracted from the spectral data relative to other components (Table 6). Note also that more outliers were detected during the calibration modeling of malic acid than for the other components (Table 4). The deviations between chemical reference

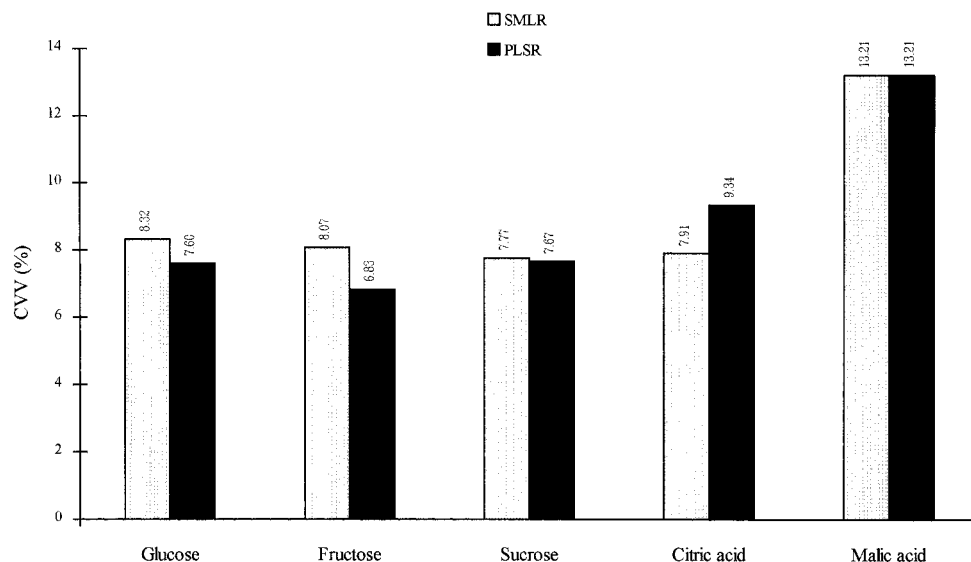


Figure 6. CVV values provided by SMLR and PLSR algorithms for the orange juice analysis.

and predicted values for malic acid shown in Figure 5d are relatively worse than for glucose in Figure 4d. This means that the predicted values for malic acid are less reliable than those of glucose, which is due to the small average concentration of malic acid relative to glucose.

Comparison of SMLR and PLSR. The SMLR is a traditional calibration method that calls for a wavelength selection procedure, whereas the PLSR is a predictive two-block regression method based on estimated latent variables rather than on the optical data directly. The SMLR approach may give better prediction results than the PLSR method, because the SMLR eliminates those X -variables that are not relevant. In addition, noninformative curvatures may contaminate the full spectrum calibration models such as PLS-1 models. Nevertheless, the response of a scanning NIR instrument at different wavelengths is usually highly collinear. A problem therefore arises in SMLR since each X -variable is assumed to bring unique information about Y . In the PLSR, using estimated latent variables, the collinearity in the X data is considered as a stabilizing advantage rather than as a problem (Bjørsvik and Martens, 1992). Thereby the PLSR method is expected to give more reliable predictions.

It is evident that the variety in the approaches to calibration stems from its empirical nature; none of the methods is theoretically more sound than the others, and they must be judged by what they achieve. Comparison of Tables 5 and 6 indicates that the prediction results of SMLR and PLS-1 calibration models are very close. The SMLR is still a workable alternative, although PLSR gives slightly better results than SMLR for the determination of individual sugars and acids in orange juices.

To account for the ranges of component concentrations in the validation set, the prediction results of SMLR and PLSR algorithms are compared using a coefficient of variation for validation defined as

$$CVV = \frac{SEP}{\bar{Y}} \times 100 \quad (6)$$

where \bar{Y} is the mean of chemical reference values and SEP is the standard error of prediction defined in (2). CVV values were computed for individual sugars and acids analyzed by both SMLR and PLSR algorithms. Figure 6 shows that PLSR gives the better CVV values

for three individual sugars, but for the determinations of the individual acids, SMLR is also a good calibration method.

CONCLUSION

Near-IR spectroscopy combined with multivariate calibration is an easy and rapid technique for the simultaneous determination of individual sugars and acids in orange juices. It is shown that the DESIR technique leads to easy sampling and reliable measurement in near-IR spectroscopy application for the aqueous solution analysis. Mathematical pretreatments of original spectral data prior to calibration improve the calibration and prediction results. Both SMLR and PLS-1 algorithms based on the near-IR spectral data allow for the quantitative analysis of component concentrations in orange juices. The limitation of the method is that minor components such as malic acid cannot be determined with an acceptable precision ($CVV \leq 10\%$). In the future, one should investigate different methods of spectral standardization to improve the sensitivity and precision of the near-IR spectroscopy analysis.

ABBREVIATIONS USED

CVV, coefficient of variation for validation; D1OD, first derivative; D2OD, second derivative; DESIR, dry extract system for infrared; MSC, multiplicative signal correction; near-IR, near infrared; PCs, principal components; PLSR, partial least squares regression; r , correlation coefficient; RMSEP, root mean square error of prediction; SEP, standard error of prediction; SMLR, stepwise multiple linear regression.

ACKNOWLEDGMENT

We are thankful to the Schutzgemeinschaft der Fruchtsaftindustrie for providing the fruit juice samples and the chemical reference data. We gratefully acknowledge Prof. J. Mayaudon for reading the manuscript and Michel Foulon for his technical support.

LITERATURE CITED

Alfaro, G.; Meurens, M.; Birth, G. S. Liquid Analysis by Dry-Extract Near-Infrared Reflectance on Fiberglass. *Appl. Spectrosc.* **1990**, *44*, 979–986.

- Alfaro, G.; Meurens, M.; Van Den Eynde, P. Application of DESIR and NIRS on Orange Juice. In *Proceeding of the Third International Conference on Near Infrared Spectroscopy*; Biston, R., Bartiaux-Thill, N., Eds.; Agricultural Research Center Publishing: Gembloux, Belgium, 1991.
- Barton, F. E. Analytical Application to Fibrous Foods and Commodities. In *Near-infrared Technology in the Agricultural and Food Industries*; Williams, P., Norris, K., Eds.; American Association of Cereal Chemists, Inc.: St. Paul, MN, 1987.
- Bjorsvik, H. R.; Martens, H. Data Analysis: Calibration of NIR Instruments by PLS Regression. In *Handbook of near-infrared analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992.
- Fischer, W. B.; Eysel, H. H.; Nielsen, O. F.; Bertie, J. E. Corrections to the Baseline Distortions in the OH-stretch Region of Aqueous Solutions. *Appl. Spectrosc.* **1994**, *48*, 107–112.
- Giangiaco, R.; Magee, J. B.; Birth, G. S.; Dull, G. G. Predicting Concentrations of Individual Sugars in Dry Mixtures by Near Infrared Reflectance Spectroscopy. *J. Food Sci.* **1981**, *46*, 531.
- Hruschka, W. R. Spectral Reconstruction. In *Handbook of near-infrared analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992.
- Lanza, E.; Li, B. W. Application for Near Infrared Spectroscopy for Predicting the Sugar Content of Fruit Juices. *J. Food Sci.* **1984**, *49*, 995–998.
- Mark, H. Data Analysis: Multilinear Regression and Principal Component Analysis. In *Handbook of near-infrared analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992.
- Martens, H.; Naes, T. Methods for Calibration; Assessment, Validation and Choice of Calibration Method; Pretreatment and Linearization. In *Multivariate calibration*; John Wiley & Sons Ltd.: Chichester, UK, 1989.
- Meurens, M. Analysis of Aqueous Liquids by Near Infrared Reflectance Spectroscopy. In *International Meeting of Chemical Engineering*; Achema 82: Frankfurt, Germany, 1982.
- Meurens, M.; Alfaro, G. La Spectroscopie Proche Infrarouge D'extrait Sec "DESIR": Nouvelle Technique de Contrôle de Qualité dans L'agro-alimentaire. *Belgian J. Food Chem. Biotechnol.* **1990**, *45*, 63–69.
- Norris, K. H. Extracting Information from Spectrophotometric Curves. Predicting Chemical Composition from Visible and Near Infrared Spectra. In *Food Research and Data Analysis*; Martens, H., Russwurm, H., Eds.; Applied Science Publ.: Oslo, Norway, 1982.
- Osborne, B. G.; Fearn, T. Applications of Near Infrared Spectroscopy in Food Analysis. In *Near infrared spectroscopy in food analysis*; Longman Scientific & Technical: Harlow, UK, 1986.
- Robertson, A. Chemical techniques in fruit juice authentication with special reference to orange juice. In *Topical Issues in Soft Drink and Fruit Juice Manufacture (Symposium)*; Campden Food & Drink Research Association: Chipping Campden, UK, 1988.
- Shenk, J. S.; Workman, J. J.; Westerhaus, M. O. Application of NIR Spectroscopy to Agricultural Products. In *Handbook of near-infrared analysis*; Burns, D. A., Ciurczak, E. W., Eds.; Marcel Dekker, Inc.: New York, 1992.
- Unscrambler *user's guide*. Version 5.5, Program package for multivariate calibration, Marketed by CAMO A/S, Trondheim, Norway, 1994, pp 45–47, 342.

Received for review February 2, 1995. Revised manuscript received November 28, 1995. Accepted April 29, 1996.[®] Funding for this study was provided by the FLAIR research project of the Commission of the European Communities.

JF9500750

[®] Abstract published in *Advance ACS Abstracts*, June 15, 1996.