ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Valorization of Onion Waste and By-Products: MCR-ALS Applied to Reveal the Compositional Profiles of Alcoholic Fermentations of Onion Juice Monitored by Near-Infrared Spectroscopy

José-María González-Sáiz, Isabel Esteban-Díez, Sofía Rodríguez-Tecedor, Consuelo Pizarro

Department of Chemistry, University of La Rioja, C/Madre de Dios 51, 26006 Logroño (La Rioja), Spain; telephone: +34-941-299634; fax: +34-941-299621; e-mail: josemaria.gonzalez@unirioja.es

Received 26 September 2007; revision received 18 January 2008; accepted 10 April 2008

Published online 21 April 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.21939

ABSTRACT: The overall purpose of the project, of which this study is a part, was to examine the feasibility of onion waste as a support-substrate for the profitable production of food-grade products. This study focused on the efficient production of ethanol from worthless onions by transforming the onion juice into onion liquor via alcoholic fermentation with the yeast *Saccharomyces cerevisiae*. The onion bioethanol produced could be later used as a favorable substrate for acetic fermentation to finally obtain onion vinegar. Near-infrared spectroscopy (NIRS), coupled with the multivariate curve resolution-alternating least squares (MCR-ALS) method, has been used to reveal the compositional and spectral profiles for both substrates and products of alcoholic fermentation runs, that is, total sugars, ethanol, and biomass concentration. The ambiguity associated with the ALS calculation was resolved by applying suitable inequality and equality constraints. The quality of the results provided by the NIR-based MCR-ALS methodology adopted was evaluated by several performance indicators, including the variance explained by the model, the lack of fit and the agreement between the MCR-ALS achieved solution and the results computed by applying previously validated PLS reference models. An additional fermentation run was employed to test the actual predictive ability of the ALS model developed. For all the components resolved in the fermentation system studied (i.e., total sugars, ethanol, and biomass), the final model obtained showed a high predictive ability and suitable accuracy and precision, both in calibration and external validation, confirmed by the very good agreement between the ALS responses and the reference values (the coefficient of determination was, in all cases, very close to 1, and the statistics confirmed that no significant difference was found between PLS reference models and the MCR-ALS methodology applied). Thus, the proven reliability of the MCR-ALS model presented in this study, based only on NIR measurements, makes it suitable for monitoring of the key species involved in the alcoholic fermentation of onion juice, allowing the process to be modeled and controlled in real time.

Biotechnol. Bioeng. 2008;101: 776–787.

KEYWORDS: onion waste; valorization; batch processes; alcoholic fermentation; MCR-ALS; NIR spectroscopy

Introduction

The food industry generates large volumes of wastes and by-products, both solids and liquid, resulting from processing and commercialization stages. The disposal and environmental problems relating to agro-food industry waste management are as heterogeneous and varied as the large variety of different waste materials produced by different sources. The EU waste management policy, which is set out in a series of Directives and Thematic Strategies, is ultimately focused on waste prevention and recycling, seeking to achieve a significant reduction in waste volumes, to decouple waste generation from economic growth and to shift to more sustainable consumption patterns (European Commission, 2006). Nevertheless, the overall volumes of waste streams are continuing to grow and, if current trends persist, the amount of waste produced in Europe in 2020 is expected to nearly double the amount produced today (European Environment Agency, 2005). Therefore, current European environmental regulations need to be complemented by alternative practices to minimize and valorize waste and thus achieve more efficient resource use.

Correspondence to: J.-M. González-Sáiz

Contract grant sponsor: Science and Technology
Contract grant number: AGL2003-09138-C04-03/ALI
Contract grant sponsor: Government of La Rioja
Contract grant number: ACPI-2004/02

In particular, waste handling and disposal represent a extremely delicate problem for the vegetable industry, owing to the perishable nature of their products and to the fact that vegetable business depends on a fragile balance between supply and demand. In the case of onion processing industry, over 450,000 tonnes of onion waste are generated annually in Europe, mainly from the UK, Holland, Italy, and Spain, in such a way that the worthless onions account for the 20% of the total production. This percentage varies depending on the harvest, but typically the surplus of onions generated is quite high. Moreover, the current quality levels demanded by customers and the search for products with high quality levels lead to even more worthless onions being discarded each year during selection and calibration stages (irregular shape, injured parts, non-commercial sizes). Total onion waste (which essentially comprises a mixture of worthless onions, damaged, whole onions and outer, dry leaves and fleshy tissues) is not particularly suitable as fodder due to their strong characteristic aroma and its disposal commonly involves landfill which results in very high economic costs and a harmful environmental impact owing to the rapid development of phytopathogenic agents (e.g., *sclerotium cepivorum*, white rot). Moreover, their removal by incineration either does not seem proper, since it would be rather expensive as a result of the high percentage of moisture in the onion waste. Thus, onion producers and processors, regulatory authorities and consumer groups are all interested in developing alternative means for the valorization of the onion waste to promote its profitable usage and its subsequent conversion into food-grade products.

Onion (*Allium cepa* L., Alliaceae) consumption has increased steadily (about 25%) over the past decade, reaching a global production of over 44 million tonnes, since onions have been promoted as a healthy, nutritious and versatile food, and also thanks to the popularity of ethnic cuisines. Sensory properties as well as health benefits of onions and many other *Allium* species (stressing their anti-cancer, anti-thrombosis and anti-cholesterol properties, and their antibacterial character) can be related to the broad variety of active compounds contained (Griffiths et al., 2002; Hertog and Hollman, 1996; Hollman and Arts, 2000; Keusgen, 2002).

The proportion of fermentable sugars contained in onions indicates that onion wastes could be successfully utilized as a favorable substrate in fermentation processes. In fact, several studies have reported the valorization of onion by-products as potential source of different valuable food ingredients (Aguilera et al., 2006), including several biological approaches dealing with ethanol, vinegar (Horiuchi et al., 1999, 2000a,b, 2004; Park et al., 1999), and lactic acid (Gardner et al., 2001; Roberts and Kidd, 2005) production from onions by fermentation. Vinegar produced from worthless onions has very interesting potential as a new functional condiment, owing to its particular physiological properties. Onion vinegar can be obtained from onion juice by a twofold fermentation process: the anaerobic transformation of fermentable sugars to ethanol (alcoholic fermentation) and the aerobic conversion of ethanol to acetic acid (acetic fermentation). Nevertheless, the vinegar production from onion waste poses some potential difficulties due to the relatively low concentration of sugars compared with other substrates commonly used in fermentation processes, and to the high concentration of antibacterial and antioxidant substances which might affect the fermentative capacity. Therefore, the control and optimization of ethanol and vinegar production from worthless onions require to carefully monitor all significant parameters involved in the respective fermentation process.

In spite of their reliability, the conventional analytical methods usually employed for off-line determination of both substrates and products during alcoholic fermentation (high performance liquid chromatography for sugars and gas chromatography for ethanol) may be fairly elaborate, costly and/or time-consuming, requiring samples preparation and even chemical manipulation, so that a significant time delay between sampling and analysis results is unavoidably introduced. Thus, simpler and faster methods, such as those based on spectroscopic techniques, capable of being implemented to provide real-time measurements, would provide a very attractive and useful alternative analytical tool for real-time monitoring and optimization of fermentation processes.

In this context, near-infrared spectroscopy (NIRS) has demonstrated great potential for cost-effective real-time bioprocess monitoring (Blanco et al., 2004; Ferreira et al., 2005; Garrido-Vidal et al., 2004; Tosi et al., 2003; Vaidyanathan et al., 2001). NIRS is a non-invasive and rapid method, which allows the simultaneous determination of multiple constituents from every single analysis, even without any sample pre-treatment. Nevertheless, it should be borne in mind that the actual applicability of NIRS to on-line process control and monitoring depends crucially on chemometrics, which supplies the suitable tools for gathering and extracting significant information from highly overlapped and noisy NIR spectra, and to improve method specificity and sensitivity. Partial least squares (PLS) regression can be considered, without question, to be the most popular and widely applied regression technique used in NIRS multivariate analysis. However, in order to develop reliable and robust PLS calibration models based on NIR measurement for predicting the diverse significant parameters involved in a fermentation process, the corresponding calibration reference values must be assessed beforehand using traditional analysis methods (i.e., PLS regression is not applicable if the response variables associated with the spectral data are not available).

There are, however, alternative chemometric methodologies exempted from this applicability limitation, such as the so-called multivariate curve resolution methods (MCR), which do not precise any reference analytical information to extract from spectral data relevant information about chemical changes occurring during complex evolving processes (Jiang et al., 2004). Multivariate self-modeling

curve resolution describes a family of different techniques and algorithms, of which alternating least squares (ALS) is the most widely applied, the aim being to estimate pure component spectra and composition profiles from data matrices of unresolved mixture spectra recorded from evolutionary systems when no prior or little information is available about the nature and composition of these mixtures. MCR methodology has been successfully applied to solve a broad variety of specific challenges in process analysis (e.g., chemical reactions, industrial processes, chromatographic elutions, or environmental data), also including approaches for monitoring fermentation processes (Blanco et al., 2006).

This study focuses on the potential to convert worthless onions and by-products discarded as wastes into added-value products such as onion vinegar and, more specifically, on the first essential step of the fermentation system involved in onion vinegar production: the effective production of ethanol from juice extracted from worthless onions via alcoholic fermentation. In order to find a suitable methodology to monitor this process in real time, and given that the combination of NIR spectroscopy with MCR-ALS methodology appeared to have great potential for cost-effective bioprocess monitoring, the objective of the present study is precisely to examine the feasibility of implementing a NIR-ALS-based approach for predicting the compositional profiles of key species in alcoholic fermentations from onion juice, which would enable the process to be controlled in real time to ensure efficient ethanol yield. Thus, the main strength of the monitoring approach proposed here has to do precisely with unifying under the same methodology the recording of NIR spectra during alcoholic fermentation nearly in real-time and the application of MCR-ALS to the analysis of monitored fermentation runs. Thus, a final and reliable model (based only on NIR measurements) was developed which would allow for a straightforward future determination of both substrates and products in alcoholic fermentation processes and contribute to the practical valorization of onion wastes by transforming them into onion vinegar.

Materials and Methods

Multivariate Curve Resolution

We will only provide an introduction to the MCR-ALS method here; a more detailed description of the algorithm can be found in literature (De Juan and Tauler, 2003; Tauler, 1995).

In MCR, a bilinear decomposition of an experimental spectroscopic data matrix **D** (with I mixture spectra in rows measured at J wavelengths) is generally performed using the following model equation:

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \quad (1)$$

where **C** is an $I \times K$ matrix describing the variation of the K chemical species contributing to the signal in the I different observations of the data matrix (concentration profiles), and $\mathbf{S}^{\mathrm{T}}$ is a $K \times J$ matrix describing how the response of these K components considered changes with respect to the J different wavelengths of the data matrix (pure spectra profiles). The dimension relating to the number of spectra recorded often coincides, for an evolving system, with the dimension related to the temporal variation. **E** is the residual matrix which contains the variance of the data unexplained by $\mathbf{C}\mathbf{S}^{\mathrm{T}}$ (i.e., due to no analyte) and is assumed to be independent and have constant variance.

Thus, according to this bilinear model, resolving matrix **D** implies firstly determining the number of analytes (K) responsible for the observed data variance (i.e., the analytes whose concentrations change with time and modify the registered signal), then obtaining the composition profiles of these species (**C**), and, finally, finding the pure spectra profiles of these components ($\mathbf{S}^{\mathrm{T}}$).

Nevertheless, the solution to this bilinear decomposition for both concentration and spectral profiles is not unique if no additional information is available, since it is subject to rotational and intensity (scale) ambiguities (Tauler et al., 1995). Fortunately, such ambiguities can be avoided, thus considerably limiting the range of possible solutions, by applying suitable constraints based on the particular nature and available prior knowledge about the target system, to model the shape of the concentration and/or spectral profiles. Two types of constraints are commonly used in MCR methods: inequality (natural) and equality constraints. Inequality constraints include:

(a) *Non-negativity constraints*: the concentration profiles of the chemical components must be positive to have physical meaning, and/or the pure component spectra must be non-negative.
(b) *Unimodality constraints*: the concentration profiles have unimodal peaks or cumulative shapes.
(c) *Closure or mass balance constraints*: the sum of all concentrations remains constant throughout the process.

Equality constraints refer to the possibility of fix known values in the concentration profiles or in the spectra, for example, pure spectra of known compounds or selectivity/local rank information.

MCR-ALS methodology is based on the iterative optimization of the resolved concentration profiles and spectra subject to selected constraints, hence, to achieve a result with chemical significance that fits satisfactorily the experimental behavior observed. This optimization is carried out for a previously determined number of significant components in the system (chemical rank) and using initial estimates of either concentration or spectral profiles. The initial starting guess of **C** or $\mathbf{S}^{\mathrm{T}}$ can be obtained by a number of exploratory procedures, including evolving

factor analysis (EFA) (Maeder, 1987), needle search (Gemperline, 1986) or simple to use interactive self-modeling mixture analysis (SIMPLISMA) (Windig and Guilmet, 1991) derived methods. The ALS optimization algorithm consists of repeatedly alternating the following two steps:

(a) From the initial or intermediate estimate of concentration profiles, the spectra profiles are obtained from the least squares resolution:

$$\mathbf{S}^{\mathrm{T}} = \mathrm{pinv}(\mathbf{C}) \times \mathbf{D} \tag{2}$$

(b) From the initial estimation of pure spectra or from the spectral profiles of the previous step, a new estimate of the concentration profiles is calculated by least squares:

$$\mathbf{C} = \mathbf{D} \times \mathrm{pinv}(\mathbf{S}^{\mathrm{T}}) \tag{3}$$

Each application of these two ALS steps produces an improved estimate of the constrained concentration and spectral profiles, so that a simple iterative refinement procedure is carried out until no further improvement in the estimates is found (i.e., the difference between the residual of one iteration and the next is less than a prefixed convergence value) or a maximum number of iterations is reached.

The performance of the optimization procedure can be evaluated by several diagnostic parameters, such as the percent of lack of fit (LOF), the percent of variance explained (EV) and the standard deviation of residuals respect experimental data (σ):

$$\begin{aligned} \mathrm{LOF}\,(\%) &= \sqrt{\frac{\sum_i \sum_j (d_{ij} - \hat{d}_{ij})^2}{\sum_i \sum_j d_{ij}^2}} \times 100 \\ \mathrm{EV} &= \frac{\sum_i \sum_j \hat{d}_{ij}^2}{\sum_i \sum_j d_{ij}^2} \times 100 \\ \sigma &= \sqrt{\frac{\sum_i \sum_j (d_{ij} - \hat{d}_{ij})^2}{n_{\mathrm{rows}} \times n_{\mathrm{columns}}}} \end{aligned} \tag{4}$$

where d_{ij} represents the elements in experimental data matrix $\mathbf{D}$, and $\hat{d}_{ij}$ the corresponding values computed by using the MCR-ALS model according to Equation (1).

Medium and Strains

Red onions from the *Figueres* variety were used as raw material for alcoholic fermentation. The product had an initial pH of 5.09 and 9.17°Brix. Onion juice for fermentation was obtained from worthless onions as follows. Onions were cut and the roots and stalks were separated. Onions were then triturated in a vertical cutter with a dicing grid of 10 mm × 10 mm. The pulp was then pressed with a manual crusher. The extract was adjusted to pH < 4.6 by adding citric acid (0.10%), packed in heat-sealed pouches of 4 L and pasteurized at 100°C during 8 min.

A commercial *Saccharomyces cerevisiae* strain, trade mark Uvaferm-CM, was used to inoculate the onion juice in the alcoholic fermentation.

Batch Fermentation Runs

Two separate fermentation runs were conducted in a 20 L bioreactor furnished with a double jacket through which water was circulated for thermostating. Processes were run without aeration (in order to minimize potential losses of ethanol by evaporation) and agitation was controlled at 300 rpm during the fermentation.

The working volume was 10 L. The lyophilized strains were added to approximately 800 mL of onion juice and poured directly into the fermentor at 30°C. No other carbohydrates or nutrients were added to the original juice extracted from worthless onions.

Temperature was carefully controlled during the process and was set to 30°C. Both fermentation processes were monitored during around 40 h to ensure the completion of alcoholic fermentation.

NIR Measurements

NIR spectra were on-line recorded on a near-infrared spectrophotometer NIRSystems 5000 (Foss NIRSystems, Raamsdonksveer, The Netherlands) equipped with a liquid analyzer module, using a 2 mm flow cell. The instrument was controlled by a compatible PC, and Vision v. 2.22 software package was used for data acquisition. Spectra were collected directly from uncentrifuged samples, at regular time intervals (45 min). Each spectrum was obtained from 32 scans performed at 2 nm intervals over the wavelength range 1,100–2,500 nm. When NIRS is applied to liquid samples, temperature control is essential to assess the reproducibility of spectra. The heater of the liquid analyzer module was set to maintain the temperature of the sample to be measured constant at 43°C. The software was programmed to maintain the flow cell inside the liquid analyzer module for 105 s before scanning to allow the sample to reach the desired temperature. In order to avoid potential interferences, automated clean cycles were performed between individual spectra collection (sample cell was cleaned with diluted sodium hypochlorite chlorine (10%)). Figure 1 shows the evolution of the NIR absorbance spectra during the alcoholic fermentation of onion juice (in particular, the variation of the spectra corresponding to the fermentation batch used for the MCR-ALS model development is displayed).

Data Processing

Two segments of the whole wavelength range 1,100–2,500 nm were removed in all the spectra: first, the region from 1,880 to 2,080 nm due to the saturation of the signal

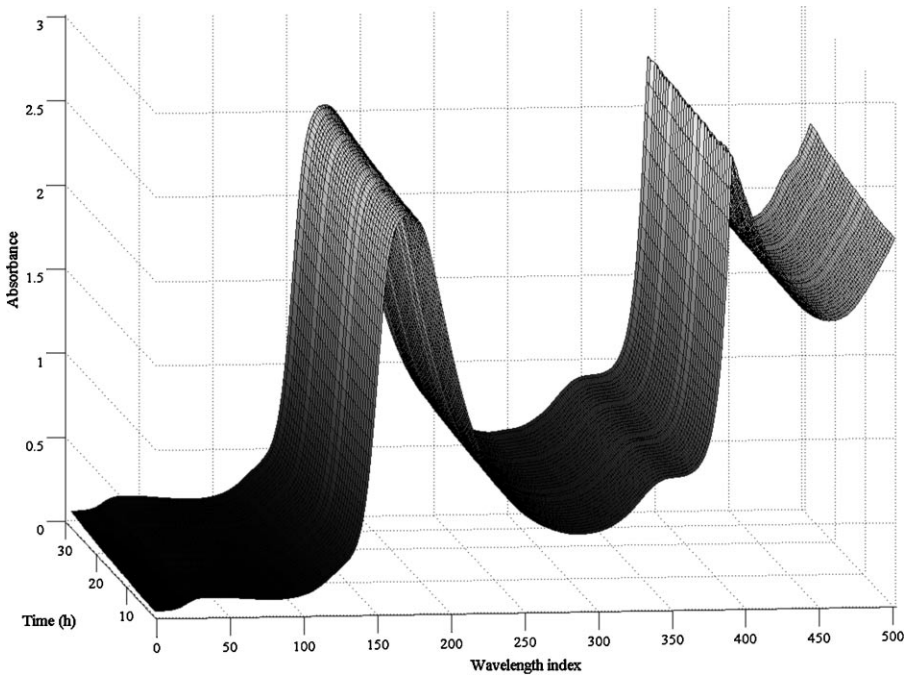

Figure 1. Evolution of the absorbance NIR spectra over time along the alcoholic fermentation batch used for calibration purposes.

caused by the strong combination band of O–H bonds from water (1,950 nm), and second, the zone from 2,300 to 2,500 nm because of its considerably low signal/noise ratio.

A total of two fermentation runs were completed and monitored by NIRS in the present study. The first monitored fermentation run (calibration run) was used to develop the MCR-ALS model best fitting the experimental system, by incorporating suitable inequality and equality constraints. Then, an additional fermentation run (a validation run conducted under similar conditions) served to evaluate the actual reliability and applicability of the ALS model previously obtained (which only relied on NIR measurements) to monitoring and predicting the concentration of both substrates and products of the alcoholic fermentation of onion wastes.

Chemometric analysis of experimental data was performed using various functions and toolboxes implemented in MATLAB® technical computing language, version 7.0 (Mathworks, Natick, MA). Specifically, all calculations relating to MCR-ALS were carried out using the software GUIPRO (Gemperline and Cash, 2003), which runs under a MATLAB® 6.0 computer environment and higher.

Each individual experimental NIR spectrum was corrected prior to any calculation by offset correction (i.e., by subtracting its lowest absorbance value from the absorbances at the rest of wavelengths) in order to remove constant background differences and any vertical shift in the spectra (Fig. 2).

The chemical rank (rank in absence of experimental noise), which indicated the number of significant components in the system studied, was estimated by applying singular value decomposition (SVD) (Malinowski, 1991).

The initial starting estimates of profiles were selected using the needle search method as introduced in the GUIPRO package.

Aside from the parameters already mentioned for testing the performance of the ALS model finally obtained, and taking into account that reference concentration of total sugars, ethanol and biomass could be determined by using previously validated PLS regression models (González-Sáiz et al., 2007), the quality of the results provided by the MCR-ALS approach in the concentration domain was evaluated via the coefficient of determination (R^2) between the ALS provided and the PLS reference values.

Results and Discussion

Observations on NIR Spectra

Figure 1 shows the spectra of the samples collected at different times throughout the fermentation process. The dominant feature observed in the raw spectra was the water absorption band at approximately 1,450 nm, which was related to the first overtone band of the –OH stretching mode. This intense, broad band masks significant information in the spectra corresponding to much weaker and overlapped bands attributable to the hydroxyl bond of other main constituent present in the culture medium (ethanol and carbohydrates). Scattering effects due to biomass accumulation in the medium during alcoholic fermentation

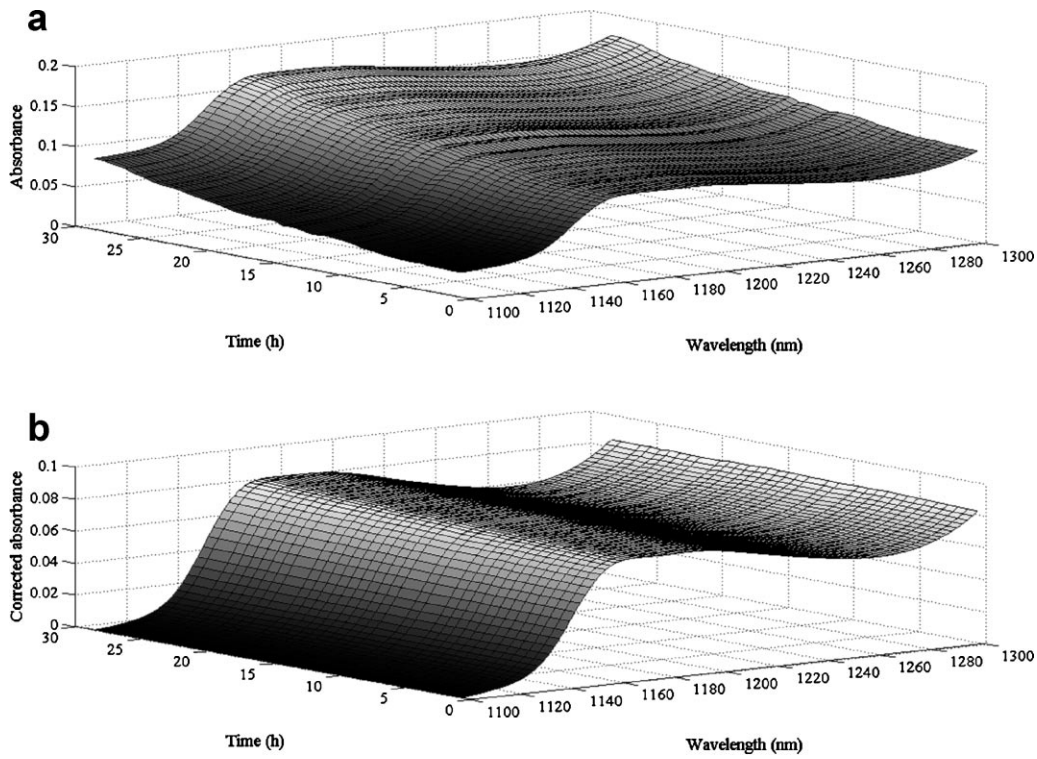

Figure 2. Absorbance mesh detail in the 1,100–1,300 nm region (**a**) before and (**b**) after applying offset correction on calibration NIR spectra.

gave rise to a slight baseline shift in the raw absorbance spectra, which especially affects to low-absorbance regions such as that ranging from 1,100 to 1,300 nm (Fig. 2a). Changes in the spectra with time at 1,680–1,700 nm (Fig. 3a) and 2,200–2,300 nm (Fig. 3b) regions (after applying offset correction to eliminate any vertical shift in the spectra) can be ascribed, respectively, to the first overtone for the C–H bonds and to the combination bands for the –OH group in ethanol structure, in such a way that spectral absorbance increases in these regions as the fermentation process progresses and ethanol is produced.

Number of Significant Components

As regards the rank of the studied system, and taking into account that the singular values associated with the significant components are expected to be much greater than other noisy contributions to the variance of the data, the number of relevant factors to be considered was estimated by simply inspecting the eigenvalues computed by SVD. The eigenvector plot of normalized eigenvalues corresponding to the first five components for the 1,100–1,880 nm/2,080–2,300 nm absorbance regions is presented in Figure 4. The first eigenvector represented the variation of sugar content in the culture medium, in such a way that the main sugars contained in the onion juice (glucose, fructose and sucrose), all consumed as substrates in alcoholic fermentation, appeared grouped together as one single, global component. As far as the second and third eigenvectors are concerned, they described clearly the change with time of the concentrations of ethanol and biomass, respectively. In contrast, the erratic trends exhibited by higher order eigenvector (e.g., those corresponding to factors 4 and 5) indicated that they had no chemical meaning. Thus, the rank analysis of the NIR data from the calibration run allowed us to conclude that there were only three significant components ($K = 3$) in the system, instead of five which was the number of main species involved in the fermentation process (glucose, fructose, sucrose, ethanol, and biomass). The reason for this rank deficiency is that the concentration profiles of the absorbing sugars are not independent since total sugar concentration (the sum of the glucose, fructose, and sucrose concentrations) is actually the essential substrate of alcoholic fermentation.

Selection of ALS Constraints

After determining the number of relevant factors for the alcoholic fermentation process studied, and as the first step to commence MCR-ALS iteration, the initial estimates of spectral profiles were computed based on the needle search method. Then, The P-ALS algorithm implemented in GUIPRO was applied to absorbance spectra recorded

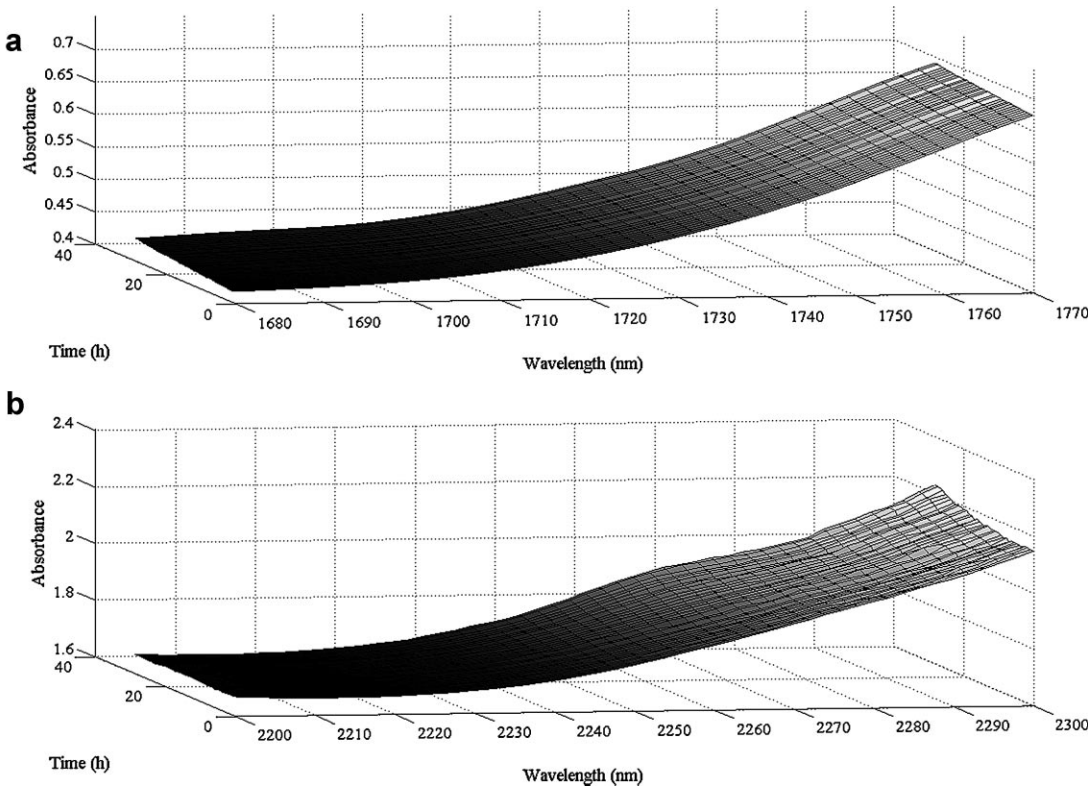

Figure 3. Absorbance mesh detail (**a**) in the 1,680–1,770 nm, and (**b**) in the 2,200–2,300 nm regions after applying offset correction on NIR spectra.

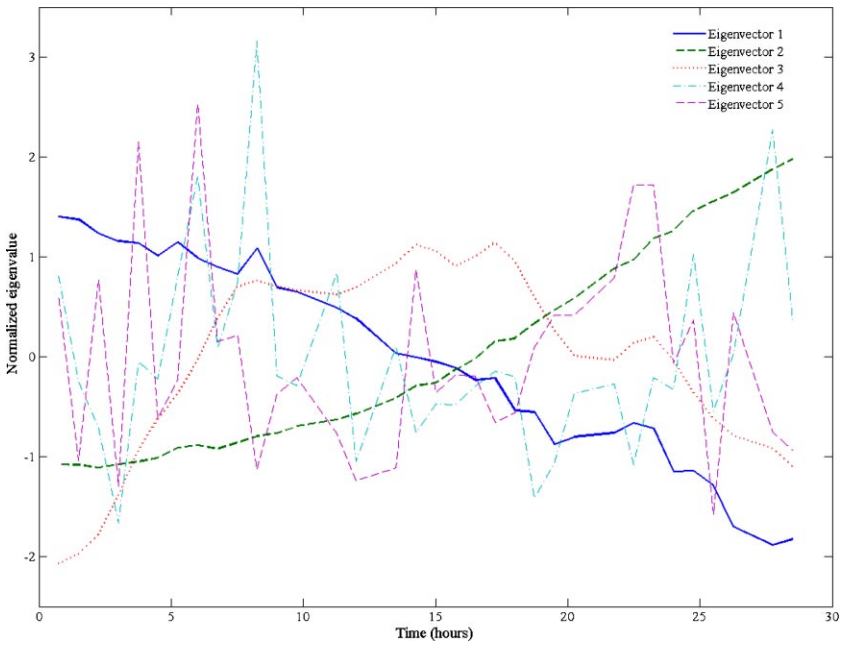

Figure 4. Eigenvector plot corresponding to the first five factors for the spectral range examined in the absorbance spectral mode. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

during alcoholic fermentation constraining all the concentration profiles to be non-negative and unimodal, that is, the temporal evolution of all the components (total sugars, ethanol and biomass) during fermentation process should be fitted to a growth/decay profile exhibiting a single maximum value. The spectra of all the components were forced to be non-negative, as corresponds to NIR spectroscopy properties. Moreover, the effect of incorporating equality constraints for all components was additionally explored as an attempt to resolve the problem of rotational and intensity ambiguities. Thus, the first spectrum recorded immediately before inoculation of the yeasts was assigned to the component total sugars. The last sample of the calibration batch, which was taken once the fermentation process was completed and all sugars were consumed from the medium, was properly filtered to remove the accumulated biomass, in such a way that the corresponding spectrum recorded from the resulting filtered sample was used as reference spectrum for the component ethanol. Finally, the difference spectrum obtained from subtracting the reference spectrum for ethanol from the last spectrum recorded during the batch was assigned to biomass.

Resolution Results

The application of the ALS procedure to the calibration fermentation run, using the above-mentioned constraints, allowed the resolution of both the concentration profiles (Fig. 5) and pure spectra (Fig. 6) associated to each significant component in the target system. The recovered concentration and spectral profiles satisfactorily reproduced the experimental data: 99.9959% of the variance associated with the experimental NIR data was successfully explained by the product $\mathbf{CS}^{\mathrm{T}}$. Additional figures of merit that were used to evaluate the fit quality were the percentage of lack of fit, which had a value of 0.0993%, and the standard deviation of residuals respect experimental data, which was equal to 0.0012. These extremely low LOF and σ values provided a global measure of the residual "noise" and confirmed that most variability in the experimental spectra was captured by the product $\mathbf{CS}^{\mathrm{T}}$.

It should be noted that the concentration profiles resolved by ALS method were not directly obtained as true concentration but rather as profile changes in concentration, and had the same units as the intensities of the data. However, the conversion of the recovered profiles to real scale for the process can be suitably performed by taking into account the maximum and minimum concentration values predicted by the respective PLS reference model for each constituent, so that the concentration profiles represented in Figure 5 are effectively on real scale. Figure 7 shows graphically the results provided by the ALS model comparing the concentration values computed by MCR-ALS for each component with the corresponding values predicted by PLS regression, in such a way that it can be used as an additional tool to evaluate the quality of the model developed in the present study. A linear regression of the ALS computed values for each modeled component versus PLS reference values was applied in such a way that Figure 7 also includes the slope, intercept and determination coefficient (R^2 value) between the NIR-ALS computed

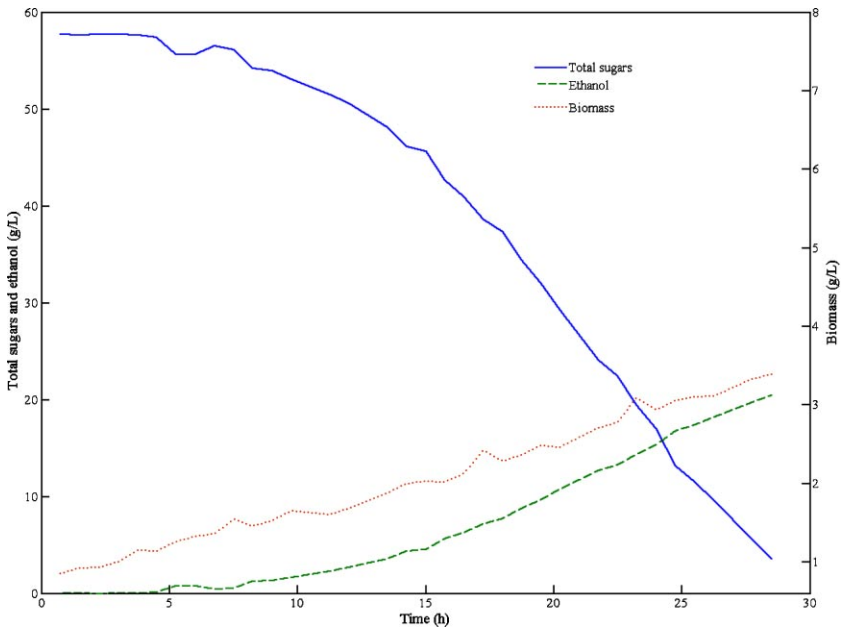

Figure 5. Resolved concentration profiles for total sugars, ethanol and biomass, extracted by the MCR-ALS procedure (all concentrations on a real scale). [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

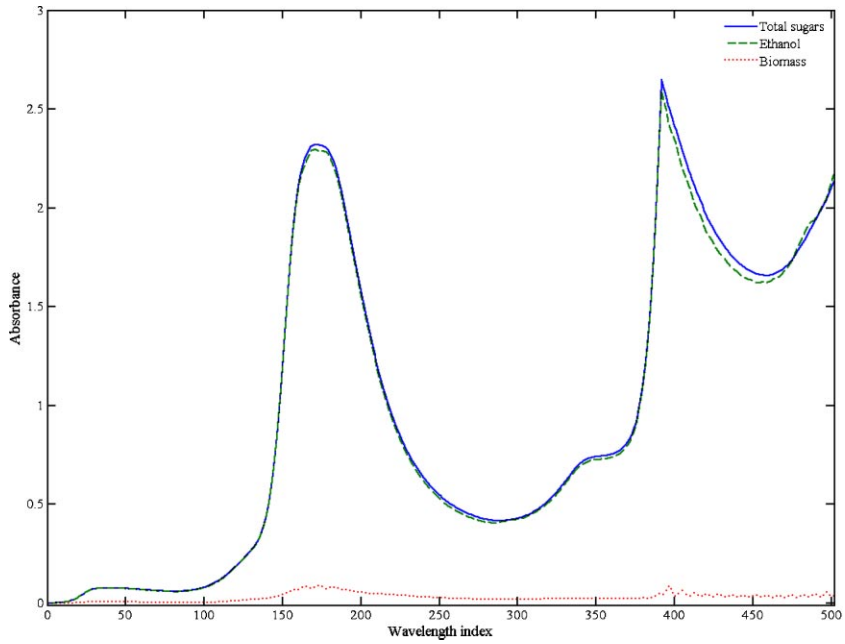

Figure 6. Pure component spectra for total sugars, ethanol, and biomass recovered from NIR absorbance spectra using MCR-ALS both with inequality and equality constraints. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

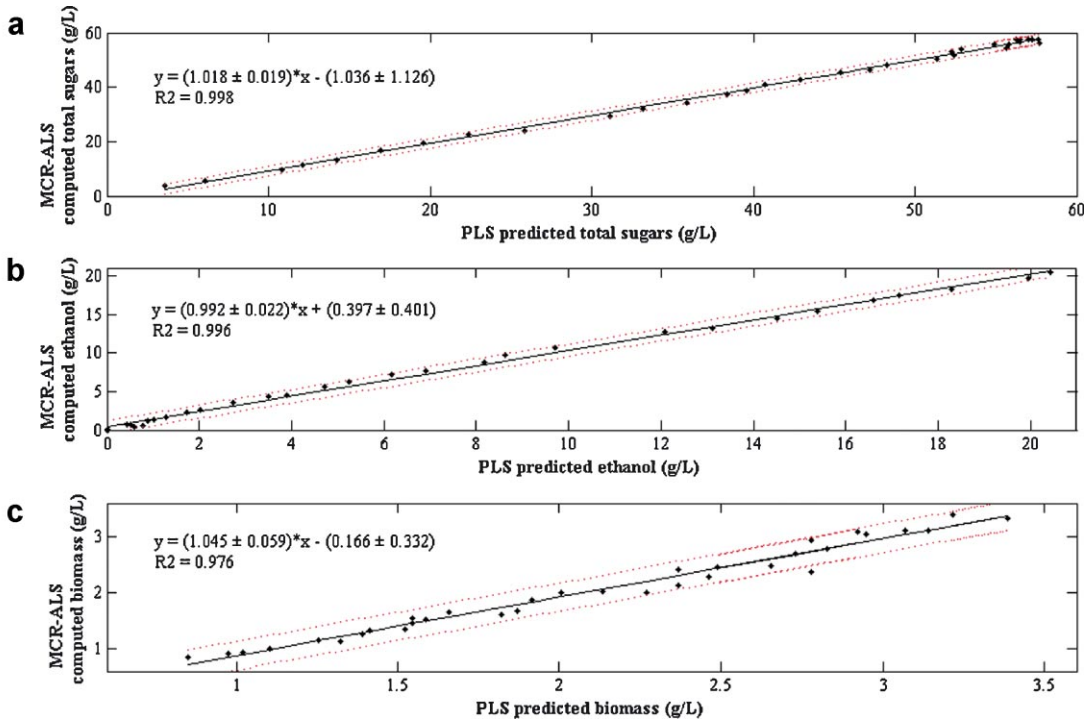

Figure 7. Correlation plot of MCR-ALS computed versus PLS reference values for each resolved component: (**a**) total sugars, (**b**) ethanol, and (**c**) biomass, when working with the alcoholic fermentation batch used for calibration. Figures of merit for the relationship between NIR computed values provided by the resulting ALS model and the corresponding reference values are shown for each modeled component, as well as the 95% confidence interval of the linear fitting. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

values and the respective values obtained by PLS calibration for each resolved species, which may be used as useful statistical measurements to better assess model performance and confirm the feasibility of using it to control fermentation in real time. Prediction intervals at the 95% confidence level were computed in all cases to properly estimate reliability and uncertainty. The very good agreement between the ALS and PLS values in all cases (the coefficient of determination, R^2, was 99.78%, 99.64%, and 97.58% for total sugars, ethanol and biomass, respectively) revealed that the inclusion of additional information in the form of the selected equality constraints significantly improved the reliability of the final ALS model to monitoring alcoholic fermentation, thus enabling accurate tracking of the total sugar consumption and ethanol and biomass production during the process.

The optimal pure spectra profiles resolved by MCR-ALS from the calibration run were then used to predict the concentration profiles for a new fermentation run monitored by NIRS (Fig. 8), performed under similar conditions, in order to confirm the fitness of the proposed methodology testing its performance when applied to an external validation run never employed in the ALS model development. The concentration profiles for this additional fermentation batch process were computed by least squares in accordance with the ALS model equation:

$$\mathbf{C}_{\text{new}} = \mathbf{D}_{\text{new}} \times \text{pinv}(\mathbf{S}^{\text{T}}) \quad (5)$$

where $\mathbf{D}_{\text{new}}$ was the experimental NIR data matrix corresponding to the validation run, $\mathbf{C}_{\text{new}}$ was the predicted concentration profiles for the new spectroscopic data, and pinv($\mathbf{S}^{\text{T}}$) was the pseudo-inverse of the pure spectra profiles previously optimized based on the calibration run. In Figure 9, the component profiles of alcoholic fermentation estimated using the proposed curve resolution model in the analysis of the batch NIR spectra used for validation are shown. The ALS model developed in this study proved to be highly reliable for predicting unknown fermentation samples, since, once the new concentration profiles had been resolved, the variance explained by the product $\mathbf{C}_{\text{new}}\mathbf{S}^{\text{T}}$ was 99.9998%, whereas the low values of both the percentage of lack of fit and the standard deviation of the residuals with regard to experimental signals (0.1354% and 0.0017%, respectively) again proved that variability in the NIR data on-line recorder during fermentation process was accurately modeled. Figure 10 shows graphically the results provided by the curve resolution approach for the test run, in such a way that they could be compared with respective results from the previously validated PLS models used as reference. As in the case of the calibration run, in order to gain further insight into the accuracy of the MCR-ALS model developed, a linear regression of the ALS predicted values for each fermentation component versus reference values for the external validation run was applied, and a number of regression statistics such as the slope, intercept, determination coefficient (R^2), and confidence limits for the slope and intercept at the 95% confidence level were calculated from each regression plot. These correlation plots in validation confirmed the reliability and accuracy of the proposed ALS model as it can be inferred from the corresponding R^2 values for resolved components, in all cases very close to 1 (99.38%, 99.70%, and 94.13% for total sugars, ethanol and biomass, respectively). Therefore, it may be argued that the model

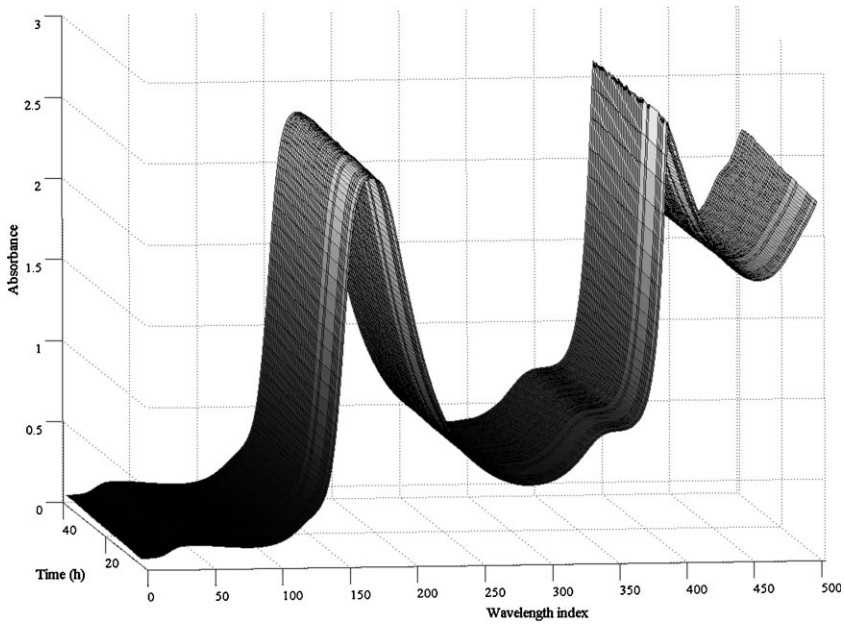

Figure 8. Evolution of the absorbance spectra over time along the alcoholic fermentation batch used for validation purposes.

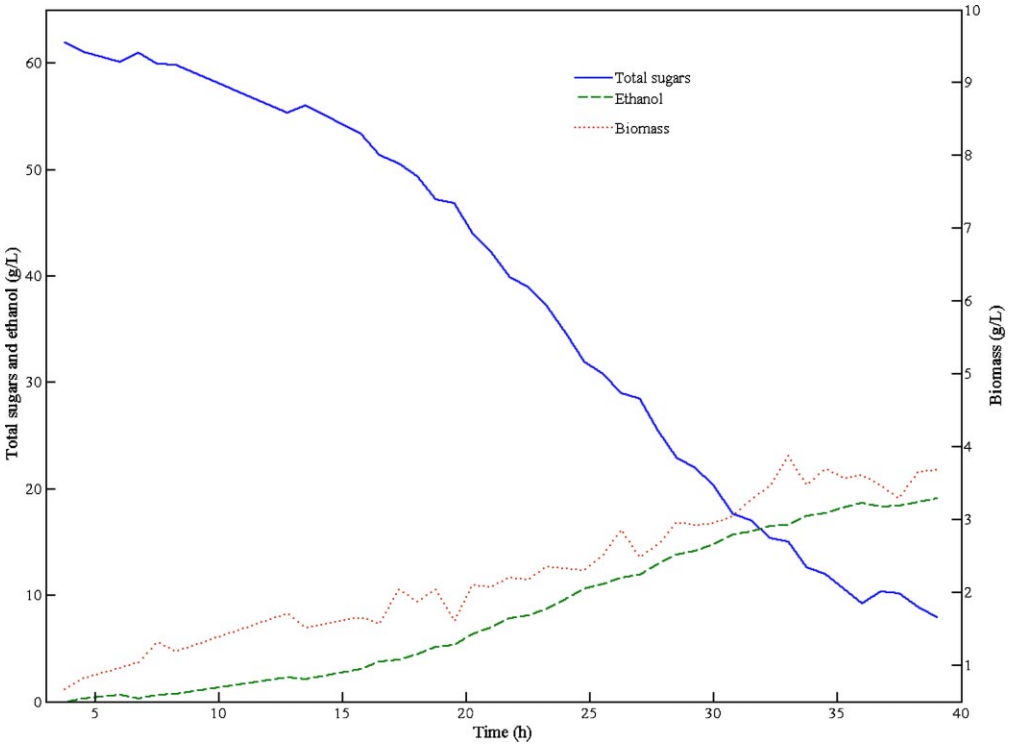

Figure 9. Concentration profiles predicted for total sugars, ethanol, and biomass by applying the MCR-ALS model to the fermentation batch used as validation run. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

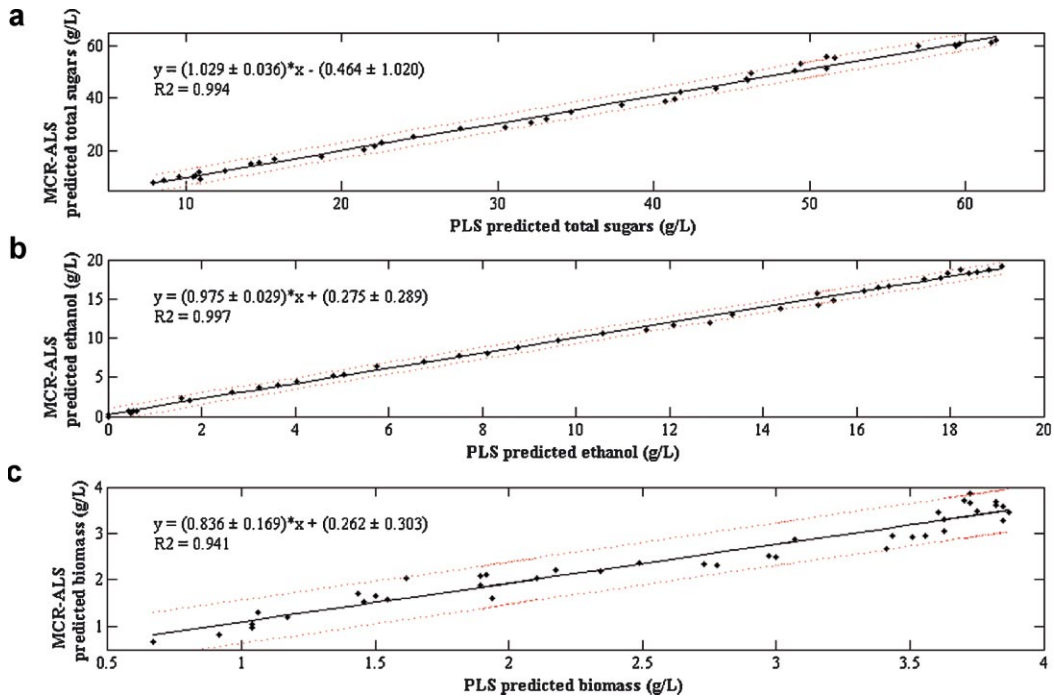

Figure 10. Correlation plot of MCR-ALS predicted versus PLS reference values for each resolved component: (**a**) total sugars, (**b**) ethanol, and (**c**) biomass, when working with the alcoholic fermentation batch used for external validation. Figures of merit for the relationship between NIR predicted values provided by the resulting ALS model and the corresponding reference values are shown for each modeled component, as well as the 95% confidence interval of the linear fitting. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

developed represents a very useful tool for simultaneously monitoring the evolution of the main species involved in the alcoholic fermentation of onion juice.

Conclusions

The present work focused on the production of ethanol from the juice extracted from worthless onions via alcoholic fermentation by the yeast *S. cerevisiae* as an alternative strategy for adding value to onion waste, and on the search for a suitable methodology for monitoring such a process in real time.

The results reported in this study have shown that the ALS analysis of NIR spectra recorded during batch processes can be successfully applied to reveal the kinetic variables of alcoholic fermentation (i.e., total sugars, ethanol, and biomass) in spite of the complexity inherent in the study of biological systems. The goodness of the concentration profiles recovered for all the components modeled was evaluated by comparison with previously validated PLS reference models. The ALS model developed was also applied in a test fermentation run carried out under similar conditions in order to better validate its actual reliability and applicability to predicting the concentration of both substrates and products. Several equality and inequality constraints were successfully introduced to improve final resolution and to minimize the ambiguities related to MCR-ALS methodology.

The use of the simple and reliable monitorization strategy proposed in the present study (developed only from NIR measurements) has the potential of significantly reducing analytical time, efforts and costs of assessing key species involved in fermentation, allowing near-real-time determination fundamental for the control and global optimization of the process.

The authors thank the Ministry of Science and Technology (Project No. AGL2003-09138-C04-03/ALI) and the Government of La Rioja (Project No. ACPI-2004/02) for their financial support.

References

Aguilera Y, Mollá Lorente E, López Andreu FJ, Martín Cabrejas MA, Esteban Álvarez RM, Benítez V, González L. 2006. Estudio de la utilización de subproductos de diferentes variedades de cebolla como ingredientes funcionales. Alimentaria 372:112–113.

Blanco M, Peinado AC, Mas J. 2004. Analytical monitoring of alcoholic fermentation using NIR spectroscopy. Biotechnol Bioeng 88:536–542.

Blanco M, Peinado AC, Mas J. 2006. Monitoring alcoholic fermentation by joint use of soft and hard modelling methods. Anal Chim Acta 556:364–373.

De Juan A, Tauler R. 2003. Chemometrics applied to unravel multicomponent processes and mixtures. Revisiting latest trends in multivariate resolution. Anal Chim Acta 500:195–210.

European Commission. 2006. 2005 Environment Policy Review. Brussels: Commission of the European Communities.

European Environment Agency. 2005. The European environment—State and outlook 2005. Copenhagen: European Environment Agency.

Ferreira AP, Alves TP, Menezes JC. 2005. Monitoring complex media fermentations with near-infrared spectroscopy: Comparison of different variable selection methods. Biotechnol Bioeng 91:474–481.

Gardner NJ, Savard T, Obermeier P, Cladwell G, Champagne CP. 2001. Selection and characterization of mixed starter cultures for lactic acid fermentation of carrot, cabbage, beet and onion vegetable mixtures. Int J Food Microbiol 64:261–275.

Garrido-Vidal D, Esteban-Díez I, Pérez-del-Notario N, González-Sáiz JM, Pizarro C. 2004. On-line monitoring of kinetic and sensory parameters in acetic fermentation by near infrared spectroscopy. J Near Infrared Spectrosc 12:15–27.

Gemperline PJ. 1986. Target transformation factor analysis with linear inequality constraints applied to spectroscopic-chromatographic data. Anal Chem 58:2656–2663.

Gemperline PJ, Cash E. 2003. Advantages of soft vs. hard constraints in self-modeling curve resolution problems—Alternating least-squares with penalty functions (P-ALS). Anal Chem 75:4236–4243.

González-Sáiz JM, Pizarro C, Esteban-Díez I, Ramírez O, González-Navarro CJ, Sáiz-Abajo MJ, Itoiz R. 2007. Monitoring of alcoholic fermentation of onion juice by NIR spectroscopy: Valorization of worthless onions. J Agric Food Chem 55:2930–2936.

Griffiths G, Trueman L, Crowther T, Thomas B, Smith B. 2002. Onions—A global benefit to health. Phytother Res 16:603–615.

Hertog MGL, Hollman PCH. 1996. Potential health effects on the dietary flavonol quercetin. Eur J Clin Nutr 50:63–71.

Hollman PCH, Arts ICW. 2000. Flavonols, flavones and flavanols-nature, occurrence and dietary burden. J Sci Food Agric 80:1081–1093.

Horiuchi JI, Kanno T, Kobayashi M. 1999. New vinegar production from onions. J Biosci Bioeng 88:107–109.

Horiuchi JI, Kanno T, Kobayashi M. 2000a. Effective onion vinegar production by a two-step fermentation system. J Biosci Bioeng 90: 289–293.

Horiuchi JI, Yamauchi N, Osugi M, Kanno T, Kobayashi M, Kuriyama H. 2000b. Onion alcohol production by repeated batch process using a flocculating yeast. Bioresour Technol 75:153–156.

Horiuchi JI, Tada K, Kobayashi M, Kanno T, Ebie K. 2004. Biological approach for effective utilization of worthless onions—Vinegar production and composting. Resour Conserv Recy 40:97–109.

Jiang JH, Liang Y, Ozaki Y. 2004. Principles and methodologies in self-modeling curve resolution. Chemom Intell Lab Syst 71:1–12.

Keusgen M. 2002. Health and alliums. In: Rabinowitch HD, Currah L, editors. Advances in allium science. Wallingford: CAB International Publishing, p 357–378.

Maeder M. 1987. Evolving factor analysis for the resolution of overlapping chromatographic peaks. Anal Chem 59:527–530.

Malinowski ER. 1991. Factor analysis in chemistry. 2nd edn. New York: John Wiley & Sons.

Park YK, Jung ST, Kang SG, Park IB, Cheun KS, Kang SK. 1999. Production of a vinegar from onion. Korean J Appl Microbiol Biotechnol 27: 75–79.

Roberts JS, Kidd DR. 2005. Lactic acid fermentation of onions. Lebensm Wiss u Technol 38:185–190.

Tauler R. 1995. Multivariate curve resolution applied to second order data. Chemom Intell Lab Syst 30:133–146.

Tauler R, Smilde A, Kowalski BR. 1995. Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. J Chemom 9:31–58.

Tosi S, Rossi M, Tamburini E, Vaccari G, Amaretti A, Matteuzzi D. 2003. Assessment of in-line near-infrared spectroscopy for continuous monitoring of fermentation processes. Biotechnol Prog 19:1816–1821.

Vaidyanathan S, Arnold SA, Matheson L, Mohan P, McNeil B, Harvey LM. 2001. Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality. Biotechnol Bioeng 74:376–388.

Windig W, Guilmet J. 1991. Interactive self-modeling mixture analysis. Anal Chem 63:1425–1432.