AGRICULTURAL AND FOOD CHEMISTRY

On-line Multisensor Monitoring of Yogurt and Filmjölk Fermentations on Production Scale

MARIÁN NAVRÁTIL, CHRISTIAN CIMANDER,[†] AND CARL-FREDRIK MANDENIUS*

Division of Biotechnology, IFM, Linköping University, S-58183 Linköping, Sweden

Near-infrared (NIR) spectrometry and electronic nose (EN) data were used for on-line monitoring of yogurt and filmjölk (a Swedish yogurt-like sour milk) fermentations under industrial conditions. The NIR and EN signals were selected by evaluation of principal component analysis loading vectors and further analyzed by studying the variability of the selected principal components. First principal components for the NIR and the EN signals were used for on-line generation of a process trajectory plot visualizing the actual state of fermentation. The NIR signals were also used to set up empirical partial least-squares (PLS) models for prediction of the cultures' pH and titratable acidity (expressed as Thorner degrees, °T). By using five or six PLS factors the models yielded acceptable predictions that could be further improved by increasing the number of reliable and precise calibration data. The presented results demonstrate that the fusion of the NIR and EN signals has a potential for rapid on-line monitoring and assessment of process quality of yogurt fermentation.

KEYWORDS: Bioprocess monitoring; electronic nose; near-infrared spectroscopy; principal component analysis; partial least-squares (PLS) regression; process trajectory plot

INTRODUCTION

One of the main goals of industrial fermentation processes is to achieve as high consistency and reproducibility as possible. This is particularly important in the food industry, where unchanged high product quality is expected by the consumers. A way to suppress batch-to-batch variations would be a control system able to monitor key process variables according to prespecified trajectories. Unfortunately, there is a lack of rapid, reliable, and robust monitoring techniques applicable under industrial conditions. However, near-infrared (NIR) spectroscopy and electronic noses (EN) were recently used to predict analytes in microbial cultures via multivariate calibration models, thereby making on-line trajectory control of important process variables possible (1-3).

In the dairy industry, NIR spectroscopy is being widely used for quality control of raw materials, intermediates, and final products (4). Rapid analysis of milk for protein, fat, lactose, and total solids by IR spectroscopy is employed as a standard method by the Association of Official Analytical Chemists (5). NIR spectroscopy is established as a very powerful method for qualitative and quantitative analysis of dairy products of all kinds (6–8). Also, electronic noses, an array of chemical gas sensors, have been used in a vast number of applications, including product quality classification in the food industry (9, 10).

It has previously been shown that fusion of NIR spectroscopy and EN sensors with bioreactor probes can be successfully used for laboratory monitoring of yogurt fermentation (11). This paper describes how this multisensor system can be applied to realtime production monitoring in a dairy under industrial conditions. Yogurt fermentation and fermentation of a typical Swedish yogurt-like sour milk ("filmjölk") were studied. By following the process trajectories and predicting basic culture parameters, deviations in the process could easily be detected, which is important for qualified and timely decision-making during manufacture and for automatic control.

MATERIALS AND METHODS

Yogurt Fermentations. Fermentations were carried out in a 1000 L stainless steel tank in a dairy for regular production (Wapnö Mejeri AB, Halmstad, Sweden). Pasteurized (15 s at 75 °C) milk (fat content = 3%) was filled into the tank and cooled to 43 or 23.5 °C for yogurt or filmjölk production. Filmjölk is a typical Swedish yogurt-like sour milk fermented with mesophilic lactic acid bacteria. The most significant difference is in lactose content; lactose during filmjölk fermentation (unlike yogurt) is not fully hydrolyzed. The inocula were prepared from commercial starter cultures of lactic acid bacteria (employing *Lactococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* for yogurt fermentation and *Lactococcus lactis*, *Lactococcus cremoris*, *Lactococcus diacetylactis*, and *Leuconostoc cremoris* for filmjölk fermentation). After inoculation, the broth was thoroughly mixed for a short period, whereafter anaerobic, unstirred fermentations were carried out for 5-7 h for yogurt or for 20-22 h for filmjölk.

Experimental Setup. The in-situ NIR probe was attached to a steel holder top-mounted on the fermentation tank and immersed into milk broth. The emission from the broth was sampled directly from the headspace of the tank and passed through a separation trap before entering the electronic nose at a flow rate of 80 mL/min using a built-in membrane pump. The NIR spectrometer and EN output signals were

^{*} Author to whom correspondence should be addressed (e-mail cfm@ifm.liu.se; telephone +46-13-288967; fax +46-13-122587). [†] Present address: Novozymes Biopharma, St:Lars väg 47, P.O. Box

¹ Present address: Novozymes Biopharma, St:Lars väg 47, P.O. Box 965, S-22009 Lund, Sweden.



Figure 1. Experimental setup for on-line monitoring of yogurt and filmjölk fermentation.

connected via a MATLAB (MathWorks Inc., Natick, MA) bridge (using Javalink/Gateway and Java runtime environment) to a G2 expert system (Gensym Inc., Boston, MA), allowing versatile and flexible data acquisition and processing (**Figure 1**). The system ran continuously for 6 weeks to acquire a sufficient amount of data for calibration and validation of models.

NIR spectra in the range between 400 and 2500 nm (steps of 2 nm) were acquired using a NIR spectrometer (model 6500, FOSS NIRSystems Inc., Silver Spring, MD). Thirty-two co-added reflectance scans were taken every 5 min and referenced to 32 co-added reflectance scans of air.

The EN was equipped with 10 MOSFET sensors (metal oxide semiconductor field-effect transistors with catalytic gates of palladium, iridium, or platinum operated at 140 or 170 °C), 19 MOS sensors (semiconductor metal oxide sensors of Figaro and FIS type; SnO₂ sensors operated at 400 °C), and an infrared CO₂ sensor. As reference gas process air was used. Gas samples were injected for 30 s every 15 min. Besides the sensor response, four additional signal parameters were calculated during the first 15 s of the sampling phase (on-derivative, on-integral) and during the first 15 s of the recovery phase (off-derivative, off-integral). Thus, in total 150 (30 sensors × 5 signal parameters) signal parameters were obtained from the EN at each measurement cycle.

Data Preprocessing. Forward selection was used to find a suitable subset of signals with high linear correlation to a process variable in order to reduce the number of noncontributing signals. The selected signals were scaled to unit variance by dividing them with their standard deviations and centered by subtracting their averages (*12*). This procedure was carried out using MATLAB software with a PLS_Toolbox (Eigenvector Research, Inc., Manson, WA).

Because of different sampling intervals of NIR (5 min) and EN (15 min) signals, trajectory plots were updated every time a NIR signal was acquired using the latest EN signal value.

Computational Methods. Principal component analysis (PCA) was used to analyze the sensor responses of NIR and EN and to calculate the score values of their first principal components (NIR-PC1 and EN-PC1). These were subsequently used to construct trajectory plots. PCA was performed by SIMCA-P software (Umetrics AB, Umeå, Sweden), and for real-time score calculations the PCA function in MATLAB /PLS_Toolbox was used.

A partial least-squares (PLS) approach was used for all calibrations and predictions. PLS is a regression between the spectra of the calibration samples (X matrix) and the off-line analyzed data of these calibration samples (Y matrix). The PLS procedure has been widely discussed elsewhere (see, e.g., ref 12). The key steps of the method are data selection, application of the PLS algorithm, and determination of the number of factors to be used in the model. PLS models were calculated with SIMCA-P software (Umetrics AB). The X matrix contained spectral data in the range between 700 and 1900 nm (based on forward selection by inspection of the loading profiles for the first PLS factors; data not shown).

Reference Data. Reference data for pH were measured off-line by a standard portable pH meter. Titratable acidity of the samples was expressed in Thorner degrees, $^{\circ}T(13)$. The value is obtained by titrating 100 mL of the product and 200 mL of distilled water with 0.1 M NaOH using phenolphthalein as indicator. Normal milk gives values of 15, which corresponds to a lactic acid content of 0.135%.



Figure 2. Visible and NIR reflectance spectra for yogurt (a) and filmjölk (b) fermentation; *t* is fermentation time.

RESULTS AND DISCUSSION

Data Preprocessing and Sensor Selection for Trajectories. Trajectory plotting has been demonstrated as a convenient way to monitor process variables in real time (14). Two sources of experimental data have been used to generate process trajectory plots: near-infrared spectroscopy data and electronic nose signals. NIR spectra were obtained with the fiber-optic immersion probe providing 1050 signals every measurement cycle. Representative NIR spectra of yogurt and filmjölk fermentation are presented in Figure 2. Spectra of yogurt are very similar to those of filmjölk, which is considered an advantage during simultaneous data processing. The explanation of this similarity lies in the origin of NIR absorption bands, containing mainly overtones and combination bands of the fundamental vibration bands in the mid-infrared range (15). Unlike ultraviolet-visible and infrared applications, the development of NIR applications is almost totally dependent on statistics and chemometrics because of the low intensity of the broad overlapping absorption bands (16). Major differences in the spectra arise from changes in highly absorbing components (water, light-scattering solids, etc.), whereas information on chemical compounds (e.g., lactose, galactose, and lactate) must be deconvoluted from minor changes in the spectra by means of chemometric routines (6).

The electronic nose analyzed gases from the headspace of the tank, giving 150 signals at every measurement. Typical time courses of five randomly selected sensor responses for yogurt and filmjölk fermentation are given in **Figure 3**. Of course, development of volatile aroma compounds is fairly dependent on the fermentation temperature. This fact explains the relatively different sensor profiles acquired during yogurt and filmjölk fermentations.

To generate the process trajectory plots, a calibration data set is required. The calibration was performed on two yogurt and four filmjölk fermentations over a period of 2 weeks. Because the instruments were running continuously for more



Figure 3. Relative response of five selected gas sensors during yogurt (a) and filmjölk (b) fermentation: (■) PtPd140; (●) Pt175; (▲) 2181; (▼) SP53; (♦) SP42.

than a month (including the validation period), both EN and NIR signals must be corrected due to baseline drift during the long-term operation. This was done using the procedure described above.

To decrease the large amount of processed data and reduce the impact of noncorrelating signals, preselection of the NIR and EN signals is of great importance. In general, the diffuse reflectance geometry mode in NIR spectroscopy usually offers less information beyond 1900 nm than transmittance mode measurements. Hence, the spectral information in reflectance measurements as we used here is mostly restricted to below 2000 nm (17). The signals were selected from inspection of the PCA loading plots, which show the importance of the Xvariables (i.e., the NIR and EN signals) in the approximation of the X matrix. The loading vectors of the NIR spectra for the first principal component (PC) (77.6 and 95.5% explained variance for yogurt and filmjölk, respectively) are plotted in Figure 4. The spectral information modeled by the first PC can be inferred from the corresponding loadings for the wavenumbers between 800 and 1800 nm, which were chosen for further calculations, showing the highest contribution to the first PC.

The first PC loadings of the EN response signals are given in **Figure 5**. Because the loading vectors for yogurt and filmjölk indicate the importance of the signals, the 16 most contributing sensor response signals were chosen for the PCA.

Calibration and Validation of the Process Trajectory Plot. Both NIRS and EN signals were used separately for PCA, and scores of the first principal components were used to make a trajectory plot for each fermentation batch.

The produced models were consequently evaluated on external validation data sets (two and four batches for yogurt



Figure 4. First principal component loadings for the whole set of NIRS signals for yogurt (a) and filmjölk (b) fermentation.



Figure 5. First principal component loadings for the whole set of EN response signals for yogurt (a) and filmjölk (b) fermentation: (*) selected sensors.

and filmjölk, respectively). **Figure 6** shows the trajectory plots generated.



Figure 6. Trajectory plots generated from the first principal components of preselected NIRS and EN signals: (a) yogurt fermentation (\Box and \bigcirc , calibration datasets; \blacktriangle and \lor , validation datasets); (b) filmjölk fermentation (\Box , \bigcirc , \triangle , and \bigtriangledown , calibration datasets; solid triangles pointing sideways, \blacklozenge , and \blacklozenge , validation datasets).

Although the trajectory plots for yogurt and filmjölk do look quite different, they both describe the same changes that milk undergoes during the fermentation. In particular, the milk protein coaggulation region is noticeable. During this stage of the fermentation, chemical and rheological properties change rapidly, which results in a significant alteration of the NIR signals due to the formation of light-scattering colloid components (11). These changes are mostly observed in the middle of the trajectory plot (the diagonal region for yogurt and the horizontal zone for filmjölk), where the dataset points are slightly distant from each other.

As can be seen, all of the plots of the same kind exhibit the same pattern. It would be possible to set boundaries or limits for the trajectory plots, which define the area that represents acceptable process conditions. A deviation of the plot outside the area would indicate an unwanted process state.

Calibration of the pH and °**T Prediction Models.** Another usage of the response data was to set up models for the prediction of the two basic process parameters in yogurt fermentation: pH value and titratable acidity (or Thorner degree, °T). The prediction of pH and °T values is based on regression models, which might cover changes also in other process variables detectable by NIR (i.e., lactose/lactic acid concentration, degree of protein coagulation, etc.). PLS was used to analyze the NIR data. One data set from the yogurt fermentation was applied as a calibration set (35 samples) to predict the pH and °T value of another data set (34 samples). A separate PLS model was developed for filmjölk fermentation, regressing spectral data with actual pH and °T values measured off-line during the process. The sample set was divided into a calibration

 Table 1. Optimization of Number of PLS Factors for pH and °T

 Prediction Model for Yogurt Fermentation Samples

no. of factors	pH prediction model				°T prediction model			
	<i>R</i> ²	SEC	Q 2	SEP	<i>R</i> ²	SEC	Q 2	SEP
1	0.948	0.26	0.938	0.30	0.966	7.2	0.977	6.7
2	0.958	0.24	0.923	0.29	0.973	6.6	0.964	9.2
3	0.969	0.21	0.942	0.30	0.981	5.5	0.953	12.3
4	0.972	0.20	0.950	0.28	0.991	3.8	0.983	7.2
5	0.990	0.12	0.959	0.35	0.995	2.8	0.976	10.0
6	0.996	0.08	0.973	0.17	0.999	1.5	0.978	6.6
7	0.998	0.06	0.966	0.28	0.999	1.5	0.978	7.0
8	0.999	0.04	0.970	0.26	0.999	1.4	0.979	6.8
9	0.999	0.04	0.968	0.27	0.999	1.1	0.978	6.8

set (246 samples from two fermentations) and a validation set (244 samples from another two fermentations).

Selection of the spectral data chose the wavelength range of 700–1900 nm, because the inspection of the loading vectors (described above) showed spectral response across this range to be the most correlating.

Due to the limited access of reference data compared to the large number of on-line data, a fitting and interpolation procedure of the experimental data was applied to increase the number of reference values for model development. The procedure was carried out according to a sigmoidal Boltzmann function:

$$y = \frac{A_1 - A_2}{1 + e^{(x - x_0)/dx}} + A_2$$

In the calibration step, the values of the constants $(A_1, A_2, x_0, and dx)$ were estimated by nonlinear curve fitting based on the Levenberg–Marquardt algorithm. These were used to interpolate pH and °T values corresponding to on-line sampled spectral absorbance values.

Cross-validation was used to test the predictive significance of each PLS factor in order to select the optimal number of the factors to be used in the PLS models and avoid overfitting, which results in a well-fitting model with little or no predictive power. Nine parallel PLS models were compared where one to nine PLS factors were used. Multiple correlation coefficient (R^2), standard error of correlation (SEC), cross-validated R^2 (Q^2), and standard error of prediction (SEP) were calculated for each model (Table 1). The predictive ability of the models was evaluated on the basis of a high Q^2 and a low SEP value, giving six PLS factors for the pH and °T models (yogurt fermentation). The same procedure was repeated for the filmjölk fermentation (data not shown). For pH, a five-factor model gave an R^2 of 0.986 and a Q^2 of 0.961. The same number of factors was chosen for a °T model with $R^2 = 0.989$ and $Q^2 = 0.977$. Other characteristics of the models selected for the evaluation are presented in Table 2, listing the concentration range, the number of samples used in the calibration, and basic statistical parameters.

Validation of the Prediction Models. It is demonstrated that analytical NIR and/or EN data in combination with chemometric evaluation have potential to become a useful tool in monitoring of industrial yogurt manufacture. However, successful longterm operation requires a sufficient amount of reliable calibration data.

Figures 7 and **8** present the correlation plots of the actual versus the predicted values for the pH and °T PLS models. The data from the calibration yogurt fermentation and from both validation fermentations are distributed evenly. The error in the

Table 2. Summary of Near-Infrared Calibration Results of pH and °T Models for Unmodified Yogurt and Filmjölk Samples

sample	parameter	Ν	range	regression model	SEC	SEP
yogurt	pН	35	4.3-6.5	PLS: six factors, 700–1900 nm	0.08	0.17
yogurt	°T	35	15-95	PLS: six factors, 700–1900 nm	1.5	6.6
filmjölk	pН	246	4.1-6.5	PLS: five factors, 700–1900 nm	0.13	0.22
filmjölk	°Τ	246	15-105	PLS: five factors, 700–1900 nm	5.1	7.7



Figure 7. Correlation between actual and NIR-predicted pH values for yogurt (a) and filmjölk (b) fermentation: (\Box) calibration dataset; (\bullet) validation dataset.

models can be attributed to variation in reference methods with the portable pH and °T devices, the precision of which in the complex yogurt media is low and subject to systematic errors. Major discrepancies occur in the midpart of the fermentation as shown in the correlation plot.

Correlation of the pH and °T models for the filmjölk fermentations is more scattered than those for yogurt. This can probably be attributed to distorted interpolation. Due to a limited number of off-line data from the process compared to the extensive access of on-line data, the pH and °T predictions were less accurate than would be possible with more off-line sampling. However, samples from the boundary zones (around initial and final pH and °T values) allow a precise prediction, where the corresponding points are located close to the correlation line. This part is also the decisive one from manufacturing control considerations.

For efficient surveillance and control of yogurt fermentation it is important to be able to monitor the change of pH and possibly the progression and balance of the mixed anaerobic lactic acid culture in relation to its utilization of lactose in the milk. These parameters are key indicators of the quality of the yogurt product and are decisive for its taste, aroma, and texture. The industrial pH measurement is mostly performed in a discontinuous manner because of signal drift and protein deposit (*18*). Determination of titratable acidity is laborious; therefore,



Figure 8. Correlation between actual and NIR-predicted °T values for yogurt (a) and filmjölk (b) fermentation: (□) calibration dataset; (●) validation dataset.

on-line prediction of these parameters using NIR spectroscopy is of interest. Consequently, it is attractive to predict these in order to accurately terminate the culture at a well-defined optimal and reproducible quality state.

In this paper exploitation of NIR spectroscopy for on-line pH and titratable acidity prediction was explored. However, there are numerous studies describing the applicability of NIR spectroscopy for the analysis of other physicochemical parameters of milk and fermented milk products (6-8, 19). Previous results from our laboratory also have shown that NIR spectroscopy and EN have potential to monitor the main chemical components (lactose, galactose, and lactate concentrations) and physical characteristics (viscosity) on-line (11). These findings can, as we have shown here, be implemented on a large scale with refined modeling.

In the present study it has been shown that despite the complexity of the yogurt and filmjölk matrices, changes in their key process variables can be captured from the NIR spectra and the EN signals. The information developed by one of the chemometrics methods applied exhibits relatively high robustness in performance over an extended process time and has the potential to provide rapid on-line monitoring and assessment of the process quality and state.

ABBREVIATIONS USED

NIR, near-infrared; EN, electronic nose; PC(A), principal component (analysis); PLS, partial least-squares; Q^2 , cross-validated multiple correlation coefficient; R^2 , multiple correlation coefficient; SEC, standard error of correlation; SEP, standard error of prediction; °T, Thörner degrees of titratable acidity of dairy products.

ACKNOWLEDGMENT

We thank Claus Fredriksen, Wapnö Mejeri AB, for valuable help and discussions.

LITERATURE CITED

- Hall, J. W.; McNeil, B.; Rollins, M. J.; Draper, I.; Thompson, B. G.; Macaloney, G. Near-infrared spectroscopic determination of acetate, ammonium, biomass, and glycerol in an industrial *Escherichia coli* fermentation. *Appl. Spectrosc.* **1996**, *50*, 102– 108.
- (2) Bachinger, T.; Mandenius, C.-F. Physiologically motivated monitoring of fermentation processes by means of an electronic nose. *Eng. Life Sci.* 2001, *1*, 33–42.
- (3) Cimander, C.; Mandenius, C.-F. Online monitoring of a bioprocess based on a multi-analyser system and multivariate statistical process modeling. *J. Chem. Technol. Biotechnol.* 2002, 77, 1157–1168.
- (4) Yeung, K. S. Y.; Hoare, M.; Thornhill, N. F.; Williams, T.; Vaghjiani, J. D. Near-infrared spectroscopy for bioprocess monitoring and control. *Biotechnol. Bioeng*, **1999**, *63*, 684–693.
- (5) Kawano, S. Application to agricultural products and foodstuffs. In *Near Infrared Spectroscopy. Principles, Instruments and Applications*; Siesler, H. W., Ozaki, Y., Kawata, S., Heise, H. M., Eds.; Wiley-VCH Verlag: Weinheim, Germany, 2002; pp 269–288.
- (6) Laporte, M.-F.; Paquin, P. Near-infrared analysis of fat, protein, and casein in cow's milk. J. Agric. Food Chem. 1999, 47, 2600– 2605.
- (7) Šašić, S.; Ozaki, Y. Short-wave near-infrared spectroscopy of biological fluids. 1. Quantitative analysis of fat, protein, and lactose in raw milk by partial least-squares regression and band assignment. *Anal. Chem.* **2001**, *73*, 64–71.

- (8) Hermida, M.; Gonzales, J. M.; Sanchez, M.; Rodriguez-Otero, J. L. Moisture, solids-non-fat and fat analysis in butter by near infrared spectroscopy. *Int. Dairy J.* 2001, *11*, 93–98.
- (9) Mandenius, C.-F. Electronic noses for bioreactor monitoring. Adv. Biochem. Eng. Biotechnol. 1999, 66, 65–82.
- (10) Korel, F.; Balaban, M. Ö. Microbial and sensory assessment of milk with an electronic nose. J. Food Sci. 2002, 67, 758–764.
- (11) Cimander, C.; Carlsson, M.; Mandenius, C.-F. Sensor fusion for on-line monitoring of yoghurt fermentation. *J. Biotechnol.* 2002, 99, 237–248.
- (12) Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* 2001, 58, 109– 130.
- (13) Tamime, A. Y.; Robinson, R. K. Yoghurt: Science and Technology; Pergamon Press: Oxford, U.K., 1985.
- (14) Teppola, P.; Mujunen, S.-P.; Minkkinen, P.; Puijola, T.; Pursiheimo, P. Principal component analysis, contribution plots and feature weights in the monitoring of sequential process data from a paper machine's wet end. *Chemom. Intell. Lab. Syst.* **1998**, 44, 307–317.
- (15) Bokobza, L. Origin of Near-Infrared Absorption Bands. In *Near Infrared Spectroscopy. Principles, Instruments, Applications*; Siesler, H. W., Ozaki, Y., Kawata, S., Heise, H. M., Eds.; Wiley-VCH Verlag: Weinheim, Germany, 2002; pp 11–41.
- (16) Ciurczak, E. W. Principles of near-infrared spectroscopy. In Handbook of Near-Infrared Analysis; Burns, D. A., Ciurczak, E. W., Eds.; Dekker: New York, 2001; pp 7–18.
- (17) Vaidyanathan, S.; Arnold, A.; Matheson, L.; Mohan, P.; Macaloney, G.; McNeil, B.; Harvey, L. M. Critical evaluation of models developed for monitoring an industrial submerged bioprocess for antibiotic production using near-infrared spectroscopy. *Biotechnol. Prog.* **2000**, *16*, 1098–1105.
- (18) De Brabandere, A. G.; De Baerdemaeker, J. G. Effects of process conditions on the pH development during yoghurt fermentation. *J. Food Eng.* **1999**, *41*, 221–227.
- (19) Rodríguez-Otero, J. L.; Hermida, M. Analysis of fermented milk products by near-infrared reflectance spectroscopy. *J. AOAC Int.* **1996**, *79*, 817–821.

Received for review July 7, 2003. Revised manuscript received November 14, 2003. Accepted December 1, 2003. The project was supported by the Swedish Agency for Innovation Systems (VINNOVA, Contract 341-2001-03766).

JF0304876