Keeping-Quality Assessment of Pasteurized Milk by Multivariate Analysis of Dynamic Headspace Gas Chromatographic Data. 1. Shelf-Life Prediction by Principal Component Regression

Belinda Vallejo-Cordoba† and Shuryo Nakai*

Department of Food Science, Faculty of Agricultural Sciences, 6650 North West Marine Drive, Vancouver, British Columbia, Canada V6T 1Z4

An objective and rapid analytical system for milk shelf-life prediction was established. Volatiles as determined by dynamic headspace capillary gas chromatography (DH-GC), psychrotrophic bacterial counts, and sensory evaluation were monitored during refrigerated storage of pasteurized milk. High correlation between flavor-related shelf life and volatiles indicated that a good proportion of the total variation was explained by the mathematical model obtained by principal component regression (PCR). The resulting standard error of the estimate of less than 2 days was an excellent approximation in shelf-life studies. Using DH-GC, results are available in 20 h (18 h of preliminary incubation + 2 h of detection and data processing). Although other rapid instrumental techniques have been proposed, they are indirectly related to flavor-related shelf life. Milk quality deterioration and consequently the termination of shelf life are caused by the appearance of off-flavors mainly determined by a composite effect of spoilage volatiles. Therefore, multivariate interpretation of headspace gas chromatograhic data and flavor-related shelf life as derived by PCR offers a direct and accurate approach to milk shelf-life prediction.

INTRODUCTION

The period between processing/packaging and the time when milk becomes unacceptable to consumers is called the "shelf life" of milk, and it reflects its keeping quality. To allow consumers to assess the age of the product at the time of purchase, a date is placed on the container indicating the last day the product may be offered for sale. However, this date is a determined number of days, generally 14, which does not consider the initial quality of the milk and does not always reflect milk shelf life.

The keeping quality of milk has traditionally been assessed by bacterial counts and sensory evaluation. However, because the deterioration in flavor of milk is due to a combined effect of compounds produced by bacteria and/or lipid oxidation generally present at low concentration, sensory evaluation remains the most useful means of assessing milk quality. Standard microbiological methods, including the Moseley keeping-quality test, have proven to be of limited value for shelf-life prediction purposes (Bishop and White, 1986). Furthermore, the time needed to obtain results renders the Moseley test impractical for monitoring the quality of milk (Bishop et al., 1984). There is a trend toward the development of techniques measuring chemical changes produced by bacteria rather than their number (Bishop and White, 1986). At the forefront of these techniques is electrical impedance. Yet these methods do not always measure spoilage compounds which are the direct cause of the appearance of off-flavors and the end of shelf life. Therefore, there is need for an objective and rapid test that would provide a direct measurement of milk flavor quality, thus resulting in reliable estimates of milk shelf life within a period that allows for effective corrective measures.

The objective of this research was to develop an analytical system for shelf-life prediction of pasteurized milk. The system proposed was an integration of tools needed to associate chemical data to sensory responses: volatile detection by dynamic headspace capillary gas chromatography (DH-GC) analysis and flavor scoring correlated by multivariate statistics. An optimized method for DH-GC as described by Vallejo-Cordoba and Nakai (1993) that had shown to provide the sensitivity and repeatability desired for detecting volatiles that developed during refrigerated storage of pasteurized milk was used.

MATERIALS AND METHODS

Sampling. Samples were commercially pasteurized and homogenized whole milk, bottled in 4-L polyethylene jugs. Four jugs were collected from a local dairy plant the day of processing on six different occasions during a period of 6 months. Milk from each jug was aseptically dispensed into 17 sterile 250-mL Delong culture flasks provided with stainless steel caps and labeled as being subsamples of one jug. Flasks containing milk were immediately refrigerated at 4 °C, until the end of shelf life. During refrigerated storage, one milk flask was removed for analysis from each sample (milk from the same 4-L jug) at predetermined intervals. Sampling was carried out three times a week in the initial stage of storage and daily when there was a decline in flavor quality as determined by sensory judges. Milk (25 mL) was aseptically transferred to a sterile culture flask for preliminary incubation (24 ± 1 °C for 18 h). The remainder of the milk in the flask was immediately subjected to sensory tests. Milk aliquots were aseptically removed before and after preliminary incubation for determination of psychrotrophic bacterial counts. Sampling for bacterial counts was once a week. After preliminary incubation, samples were immediately frozen at −40 °C in 40-mL amber glass vials provided with Teflon-lined caps until used for volatile analysis.

Sensory Evaluation. Five judges of eight initial trainees were selected by using sequential sampling (Amerine et al., 1965). The judges were four females and one male between 26 and 38 years of age. The method used for training and testing was scoring with the ADSA score card for milk on a 10-point scale according to the scoring guide (American Dairy Science Association, 1987). Training was carried out by tasting milk of a wide range of days left to sell-by date, including some milks past this date. The

† Health Protection Branch, Health and Welfare Canada, 3155 Willingdon Green, Burnaby, BC, Canada V5G 4P2.

0021-8561/94/1442-0989$04.50/0

panelists were provided with the standard definitions or descriptions of the off-flavors encountered in milk to familiarize them with the flavor defects most likely to be encountered during storage (American Dairy Science Association, 1987). After discussion, the panelists agreed on the recognition of specific milk off-flavors in the samples used for training, and they were asked to identify them on the score cards. For sensory testing, samples (15 mL) at a temperature of 15–20 °C were served in 250-mL Styrofoam cups. Shelf life was determined by the panelists through sensory testing as described by Bishop et al. (1984). Accordingly, shelf life was ended whenever a score of 5 or lower was recorded by three of the judges, and the day before was considered the end of shelf life.

Psychrotrophic Bacterial Counts. Psychrotrophic bacterial counts (PBC) were determined by preparing spread plates on plate count agar (PCA, Difco Laboratories, Detroit, MI). PBC were carried out once a week, before and after preliminary incubation of the milk. Plates were incubated at 5 ± 1 °C for 10–15 days before colony forming units per milliliter (CFU/mL) were recorded. Serial dilutions of milk were made with 0.1% (w/v) Bactopeptone (Difco) in distilled deionized water.

Milk Volatile Detection by Dynamic Headspace Gas Chromatographic Analysis. Milk volatiles developed during refrigerated storage were determined under optimized conditions as described Vallejo-Cordoba and Nakai (1993) until the end of shelf life was reached as detected by sensory testing. Milk (9.1 mL) provided with 10 mg of tetradecanol (Sigma, St. Louis, MO) to diminish foaming and 10 μL of a 2.5 ppm solution of 4-methyl-2-pentanone to serve as an internal standard was purged for 12.5 min at 44 °C, and volatiles were absorbed onto a Tenax TA trap (60/80 mesh, Tekmar, Cincinnati, OH) and chromatographed on a DB-624 (30 m, 0.32 mm i.d., 1.8-μm film thickness, J&W Scientific, Rancho Cordova, CA) capillary column.

Data Manipulation. Gas chromatographic data were processed by a data processor (Chromatopac CR3A, Shimadzu Corp., Kyoto, Japan) and acquired by a personal computer provided with an interface (Chromatopac Data Archive Utility, version 2.1, Shimadzu Scientific Instruments, Inc., Columbia, MD). ASCII files were created and imported into a Lotus spreadsheet (version 2.1, Lotus Development Corp., Cambridge, MA) for elimination of unwanted information in the chromatographic reports. At this stage, an automated peak recognition procedure was applied by using a program written in BASIC for a personal computer (Vallejo-Cordoba, 1992). Peak matching or recognition was based on retention time ranges established after visual inspection of typical chromatograms. The automated peak recognition program added peak areas when there was more than one peak matching a certain retention time range. Chromatograms were divided in 47 retention time ranges which included all peaks in the chromatograms. Areas of 47 retention time ranges including the internal standard for every chromatogram report file were imported to a Lotus spreadsheet for data matrix construction. Ratios of each of the 46 peak areas (corresponding to a retention time range) to the area of the internal standard (standardized peaks) were used for statistical analysis. Finally, a spreadsheet consisting of gas chromatographic data and sensory evaluation based shelf-life data was imported to a Systat statistical software package (version 5.0, Systat, Inc., Evanston, IL, 1990). The data were checked for the assumption of multivariate normality by generating a normal probability plot using the NORM option of the PPLOT procedure of Sygraph (version 5.0, Systat, 1990). A log transformation of the data improved the fit.

Statistical Analysis. The Systat statistical procedures (version 5.0, Systat, 1990) used were principal component analysis (FACTOR) and stepwise multiple regression (MGLH). In multivariate analysis, each of the 46 log transformation of standardized peak areas (referred as peaks) from a chromatogram was considered a variable for a given milk subsample corresponding to a shelf life determined by sensory evaluation. First, principal component analysis (PCA) was used to reduce the original 46 correlated variables (standardized peaks) to a set of 30 uncorrelated principal components (PC). Principal component regression (PCR) was used for building a shelf-life prediction model. The dependent variable was the log transformation of shelf life and the independent variables were the first 30 principal components. Forward stepwise selection with an alpha-to-enter and alpha-to-remove of 0.05 and a minimum tolerance for entry into the model of 0.01 was used to build the model. The model for shelf-life prediction was expressed by

$$Y_i = B_1 + B_2X_{2i} + B_3X_{3i} + ... + B_pX_{pi} + E_i \quad (1)$$

where $i = 1...n$ (n is the number of cases), B_1 is a constant, $B_2...B_p$ represent partial regression coefficients of p predictors, E_i is the residual term, Y_i represents the response log (shelf life) in the ith trial, and $X_{2i}...X_{pi}$ represent the p predictors, principal components, in the ith trial.

To assess the adequacy of the regression model for making predictions, graphic analysis of the residuals, the differences between the observed and predicted values of the dependent variable, was performed by using Sygraph (version 5.0, Systat, 1990). A normal probability plot was drawn for assessing normal distribution of the errors. Plots of the Studentized residuals and residuals against estimated values were generated to assess constant variance and independence of the errors. The final step in the model building was validation of the selected regression model. Cross-validation, used for this purpose, consists of splitting the data into two sets. The first, called the model-building data set, was used to develop the model, and the second set, called the validation of prediction data set, was used to evaluate the predictive ability of the selected model. Generally, the number of cases in the model-building set should be 6–10 times the number of variables (Neter et al., 1990). In this research, cases were 134 milk samples tested and variables were 14 principal components resulting from stepwise PCR. Thus, 21 cases were held for a validation data set, leaving 113 cases for building the model.

RESULTS AND DISCUSSION

Volatile isolation by DH-GC was used as part of an integrated system for milk shelf-life prediction. The analytical scheme consisted of the integration of the tools needed to permit the association of sensory responses to chemical data, purge and trap capillary gas chromatography or DH-GC, and flavor scoring correlated by multivariate statistics on a personal computer. Since the shelf life is ultimately determined by its sensory acceptability, flavor scoring has been the standard by which shelf-life end points have been determined (Kahn and Firstenberg-Eden, 1987). Therefore, calibration of an alternate method, such as purge and trap GC, requires the input of a trained sensory panel. However, once the mathematical models are established, the objective evaluation of unknown samples can be accomplished without the need for human input.

Preliminary shelf-life experiments with and without preliminary incubation of the milk indicated that enhancement of volatile production by preliminary incubation was required for shelf-life prediction early in its storage. Therefore, preliminary incubation of the milk for 18 h at 24 ± 1 °C was included before purge and trap gas chromatography. Plots of log psychrotrophic bacterial counts (PBC) before and after preliminary incubation for four different milk containers (4-L jugs) as a function of storage are presented in Figure 1. Enhancement of psychrotrophic bacterial activity, much earlier in storage, after samples were subjected to preliminary incubation was clearly shown. Psychrotrophic bacterial populations for samples without preliminary incubation on days 1 and 3 were counts of less than 10. Counts of between 10^6 and 10^8 CFU/mL were obtained after only 1 day for samples with preliminary incubation.

The last tool of this integrated system was data manipulation in a personal computer. Data acquisition from the GC followed by automatic peak recognition in a personal computer was a unique feature of this study. Other researchers have recognized the difficulties asso-

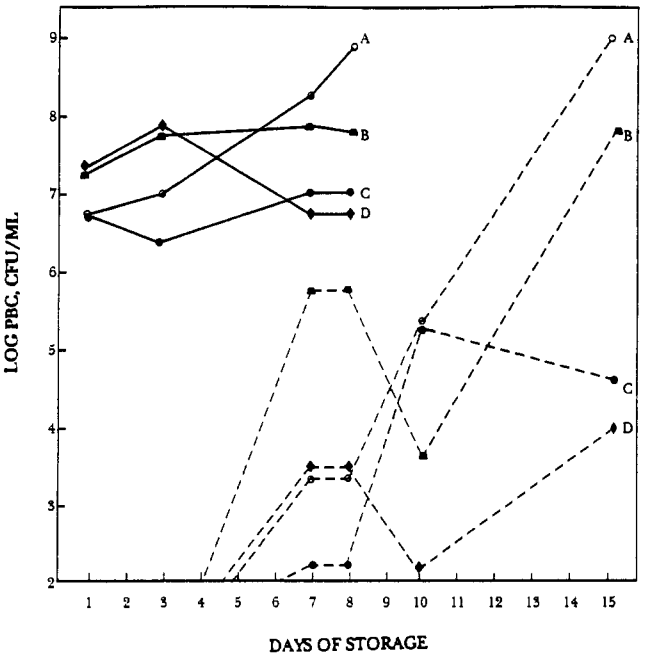

Figure 1. Psychrotrophic bacterial counts with (solid lines) and without (dotted lines) preliminary incubation during refrigerated storage for four different milk containers from the same production line. Missing values for 1 and 3 days for nonincubated samples correspond to counts of less than 200 CFU/mL.

ciated with matching up corresponding peaks in different gas chromatograms (Dravnieks et al., 1973; Liardon et al., 1984). Furthermore, the large amount of GC data in flavor studies is subject to human error due to fatigue during peak matching. The simplified approach of computerized peak recognition resulted in considerable savings in time and effort.

Principal Component Regression. Principal component analysis transformed the original 46 variables into fewer uncorrelated principal components. Cumulative percent of total variance explained by up to PC_{30} was 96% (Table 1). There is no universally accepted method for indicating how many principal components to retain for subsequent analysis. The decision is largely arbitrary. In this study, 30 principal components were retained for further analysis. Since principal component regression (PCR) requires a fewer number of samples than multiple linear regression (MLR) and is immune to multicollinearity, PCR was used for building a milk shelf-life prediction model. The regression model for milk shelf-life prediction was expressed by

$$\log(\text{shelf life}) = 1.069 - 0.178PC_1 - 0.115PC_2 - 0.123PC_3 + 0.075PC_4 - 0.065PC_6 + 0.031PC_7 + 0.068PC_9 + 0.033PC_{10} - 0.059PC_{12} + 0.037PC_{16} - 0.032PC_{17} + 0.048PC_{18} - 0.035PC_{22} - 0.028PC_{27} \quad (2)$$

Regression statistics for the milk shelf-life prediction model based on gas chromatographic data are summarized in Table 2. The F value was highly significant ($P < 0.001$), meaning that shelf life was dependent on the principal components. Values of greater than 0.8 for the adjusted R^2 (adjusted coefficient of multiple determination) implied that a good proportion of the total variation in shelf life was explained by the model. The final step in the model-building process was the validation of the selected regression model. Similar regression statistics for the model-building set for cross-validation were obtained. Residuals of less than 2 days in shelf life were observed across the validation data set, thus confirming the strength of the model for predictive purposes.

Table 1. Eigenvalues, Proportion of Variance Explained, and Cumulative Percent of Total Variance Explained in PCA Using Gas Chromatographic Data (n = 134)

principal component	eigenvalue	prop var explained	cum % of total var
PC_1	16.60	36.09	36.09
PC_2	3.59	7.81	43.91
PC_3	3.44	7.49	51.40
PC_4	2.45	5.32	56.72
PC_5	2.09	4.54	61.27
PC_6	1.81	3.94	65.22
PC_7	1.39	3.04	68.26
PC_8	1.19	2.59	70.86
PC_9	1.17	2.55	73.41
PC_{10}	0.98	2.15	75.56
PC_{11}	0.95	2.07	77.63
PC_{12}	0.87	1.89	79.53
PC_{13}	0.81	1.77	81.30
PC_{14}	0.69	1.51	82.81
PC_{15}	0.64	1.41	84.22
PC_{16}	0.60	1.31	85.54
PC_{17}	0.56	1.23	86.77
PC_{18}	0.52	1.14	87.91
PC_{19}	0.48	1.06	88.97
PC_{20}	0.47	1.04	90.01
PC_{21}	0.42	0.91	90.93
PC_{22}	0.36	0.79	91.73
PC_{23}	0.34	0.73	92.46
PC_{24}	0.32	0.71	93.18
PC_{25}	0.31	0.67	93.86
PC_{26}	0.27	0.60	94.46
PC_{27}	0.25	0.54	95.00
PC_{28}	0.23	0.50	95.50
PC_{29}	0.21	0.46	95.97
PC_{30}	0.21	0.46	96.43

Table 2. Regression Statistics for the Milk Shelf-Life Prediction Model Based on Gas Chromatographic Data (n = 134)[a]

	variable		
	R^2	R^2A	F value
14PC	0.83	0.81	42.11**

[a] Standard error of estimate = 0.134 (1.36 days). PC, principal components; R^2, coefficient of multiple determination; R^2A, adjusted coefficient of multiple determination. ** Highly significant ($P < 0.01$).

Volatiles produced during refrigerated storage of milk are the result of chemical changes mainly caused by bacterial activity; consequently, high correlations with flavor-related shelf-life were expected. A standard error of the estimate of less than 2 days in shelf life was obtained in this research. Several researchers have attempted to use volatile detection in pasteurized milk as an indication of keeping quality, yet failure of previous works may be attributed to attempts to correlate a single component to a sensory response. For example, Urbach and Milne (1988) could not find a good correlation between ethanol content in pasteurized milk and acceptability. Although simple correlations can sometimes be useful, it is widely recognized that flavor quality is related to a complex balance of several components. Therefore, the use of multivariate statistics for the correlation of most volatiles in the chromatographic profile to a sensory response and shelf life was the key feature of this study. Other workers have recognized the complexity of flavor and have used multivariate statistics to correlate chromatographic profiles to sensory responses in off-flavored milk; for example, Leland et al. (1987) used this approach to classify milk into oxidized flavor groups.

To assess the adequacy of the regression models for making predictions, graphic analysis of the residuals was performed. Plots of the residuals did not indicate vio-

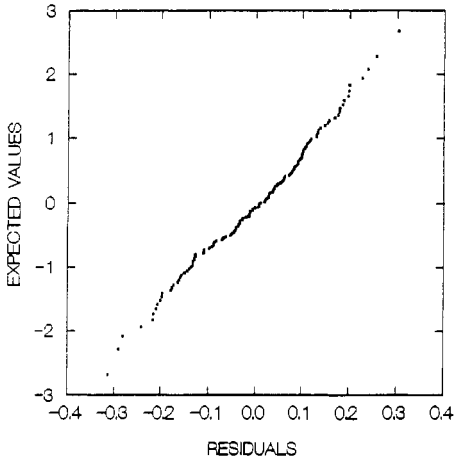

Figure 2. Normal probability plot of the residuals vs expected values for the milk shelf-life prediction regression model.

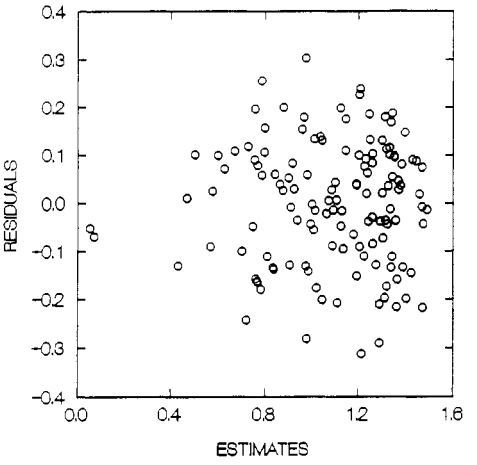

Figure 3. Relationship between residuals and estimates for the milk shelf-life prediction regression model.

lations of the assumptions in regression analysis. A normal probability plot showed the residuals fell approximately on a diagonal straight line, which indicated that the errors were normally distributed (Figure 2). A plot of the residuals against estimated values showed values were randomly scattered above and below the zero horizontal line. This indicated that the errors were independent (Figure 3). Plots of Studentized residuals against estimated values showed that residuals were arranged in a horizontal band within 3 units around zero, indicating that error had constant variance (data not shown).

Models relating psychrotrohic bacterial counts and shelf life, on the other hand, did not have predictive value. Very low correlations were produced by the linear and quadratic relationships of PBC to shelf life of pasteurized milk. Regression statistics for the best model based on psychrotrophic bacterial counts had a multiple R of 0.57, which is comparable to previously reported values (Phillips and Griffiths, 1985). It is widely recognized that initial PBC and total bacterial counts are of little use in predicting shelf life of dairy products (Blankenagel, 1976). According to Firstenberg-Eden and Eden (1984), shelf life is likely to correlate better with bacterial activity than with initial bacterial numbers. Thus, a technique that measures bacterial metabolic activity and, in fact, integrates the effects of number of organisms and their metabolic activity would be a most suitable tool for shelf-life prediction. Poor correlations of PBC to shelf life were reported to possibly be due to two limitations. First, spoilage is not always related to the number of organisms present. It is the type of bacteria rather than the actual numbers that determined the production of off-flavors and consequently the end of shelf life (Bishop and White, 1986). Second, it appears that poor flavor and keeping quality can also be attributed to the presence of microbial enzymes and metabolic products from organisms present before pasteurization (Patel and Blankenagel, 1972). Therefore, methods such as volatile detection by headspace gas chromatography are a better direct assessment of the quality of foods than the standard plate counts.

Volatile detection by DH-GC as well as other metabolism-based techniques includes very sensitive analytical instruments that can detect microbial components or metabolites, yet DH-GC, when coupled to MS, takes the analyst one step ahead by elucidating the chemical composition of these metabolites. Association of certain metabolites to specific contaminating microorganisms might give more insight into the nature of contamination, thus enabling production staff to identify and respond rapidly to problems in the processing plant. Since DH-GC is a relatively rapid method, the test may be used to monitor plant hygiene by assessing the degree of contamination of in-line samples. Furthermore, a significant improvement in the shelf life may be obtained by this routine monitoring, coupled with effective response in plant cleaning operations.

Conclusions. The analytical system established in this work has potential for application in the dairy industry for shelf-life prediction. Results confirmed previous findings that techniques based on metabolic activity correlate better with sensory quality and shelf life than bacterial counts. Volatile detection by dynamic headspace gas chromatography proved to be a useful analytical tool for milk shelf-life prediction since it is a direct measurement of the flavor quality of pasteurized milk. High correlation between flavor-related shelf life and volatiles detected by dynamic headspace gas chromatography indicated that a good proportion of the total variation in shelf life was explained by the model obtained by principal component regression. The end of shelf life was predicted with an accuracy of ±2 days, which is an excellent approximation in shelf-life studies. Although sensory evaluation was required for calibration of the instrumental method, once the mathematical models were established, the objective evaluation of unknown samples was accomplished without the need for human input.

ACKNOWLEDGMENT

The financial support of the Natural Sciences and Engineering Research Council of Canada and of the Consejo Nacional de Ciencia y Tecnologia (Mexico) is gratefully acknowledged. This study was partly presented at the 1991 National Meetings of the American Chemical Society, Agricultural and Food Chemistry Division, Atlanta, GA, April 14–16, and New York, NY, August 25–30.

LITERATURE CITED

American Dairy Science Association. Committee on evaluation of dairy products, score card sub-committee, Champaign, IL, 1987. Cited in: Bodyfelt, F. W.; Tobias, J.; Trout, G. M. In *The Sensory Evaluation of Dairy Products*; Van Nostrand Reinhold: New York, 1988.

Amerine, M. A.; Pangborn, R. M.; Roessler, E. B. *Principles of Sensory Evaluation of Food*, Academic Press: New York, 1965.

Bishop, J. R.; White, C. H. Assessment of dairy product quality and potential shelf-life—a review. *J. Food Prot.* **1986**, *49*, 739–753.

Bishop, J. R.; White, C. H.; Firstenberg-Eden, R. Rapid impedimetric method for determining the potential shelf-life of

pasteurized whole milk. *J. Food Prot.* **1984**, *47*, 471–475.
Blankenagel, G. An examination of methods to assess post-pasteurization contamination. *J. Milk Food Technol.* **1976**, *39*, 301–304.
Dravnieks, H. G.; Reilich, H. G.; Whitfield, J. Classification of corn odour by statistical analysis of gas chromatographic patterns of headspace volatiles. *J. Food Sci.* **1973**, *38*, 34–39.
Firstenberg-Eden, R.; Eden, G. *Impedance Microbiology*; Research Studies Press: Letchworth, Hertfordshire, England, 1984.
Kahn, P.; Firstenberg-Eden, R. Prediction of shelf-life of pasteurized milk and other fluid dairy products in 48 h. *J. Dairy Sci.* **1987**, *70*, 1544–1550.
Leland, J. V.; Reineccius, G. A.; Lahiff, M. Evaluation of copper-induced oxidized flavour in milk by discriminant analysis of capillary gas chromatographic profiles. *J. Dairy Sci.* **1987**, *70*, 524–533.
Liardon, R.; Ott, U.; Daget, M. Application of multivariate statistics for the classification of coffee headspace profiles. *Lebensmittel-Wissenschaft Technol.* **1984**, *17*, 32–38.
Neter, J.; Wasserman, W.; Kutner, M. H. *Applied Linear Statistical Models*, 3rd ed.; Irwin: Homewood, IL, 1990.
Patel, G. B.; Blankenagel, G. Bacterial counts of raw milk and flavour of the milk after pasteurization and storage. *J. Milk Food Technol.* **1972**, *35*, 203–206.
Phillips, J. D.; Griffiths, M. W. Bioluminescence and impedimetric methods for assessing shelf-life of pasteurized milk and cream. *Food Microbiol.* **1985**, *2*, 39–51.
Urbach, G.; Milne, T. The concentration of volatiles in pasteurized milk as a function of storage time and storage temperature-a possible indicator of keeping quality. *Aust. J. Dairy Technol.* **1988**, *43*, 53–58.
Vallejo-Cordoba, B. Shelf-life prediction and flavor quality assessment of pasteurized milk by multivariate analysis of dynamic headspace gas chromatographic data. Ph.D. Dissertation, University of British Columbia, 1992.
Vallejo-Cordoba, B.; Nakai, S. Using a simultaneous factor optimization approach for the detection of volatiles in milk by dynamic headspace gas chromatographic analysis. *J. Agric. Food Chem.* **1993**, *41*, 2378–2384.

Received for review September 1, 1993. Accepted January 6, 1994.[⊗]

[⊗] Abstract published in *Advance ACS Abstracts*, February 15, 1994.