Classification of Microbial Defects in Milk Using a Dynamic Headspace Gas Chromatograph and Computer-Aided Data Processing. 2. Artificial Neural Networks, Partial Least-Squares Regression Analysis, and Principal Component Regression Analysis

Yasumi Horimoto,* Karoline Lee, and Shuryo Nakai

Department of Food Science, The University of British Columbia, 6650 Northwest Marine Drive, Vancouver, BC, V6T 1Z4 Canada

Objective, yet cost-effective evaluation of flavor is difficult in quality control of milk. Inexpensive gas chromatographs in conjunction with computer models make it feasible to construct an objective flavor evaluation system for routine quality control purposes. The purpose of this study was to classify milk with microbial off-flavors using a low-cost headspace gas chromatograph and computer-aided data processing. Principal component similarity (PCS) analysis was discussed in part 1. In part 2, artificial neural networks (ANN), partial least-squares regression (PLS) analysis, and principal component regression (PCR) analysis are examined. UHT milk was inoculated with various bacteria (*Pseudomonas fragi, Pseudomonas fluorescens, Lactococcus lactis, Enterobactor aerogenes,* and *Bacillus subtilis*) and a mixed culture (*P. fragi*:*E. aerogenes*:*L. lactis* = 1:1:1) to approximately 4.0 $\log_{10}$ CFU mL^{-1}. ANN were able to make better predictions than PLS and PCR. The prediction ability of PLS was better than PCR. The performance of each method depended on the content of training and testing of data, i.e., more data resulted in better predictive ability.

Keywords: *Dynamic headspace gas chromatography; artificial neural network; partial least-squares regression analysis; principal component regression analysis; off-flavor*

INTRODUCTION

Objective evaluation of flavor quality has been one of the most difficult problems in quality control of milk. Recent progress in instrumental analysis and computer-aided data processing now make objective evaluation practicable.

Enormous amounts of data produced by automated instrumental analysis make efficient data-processing techniques an absolute necessity in modern food analysis. Several different multivariate analyses have been used. Especially in flavor research, multivariate analysis is an essential tool for processing numerous peaks obtained from GC patterns to classify samples. Aishima and Nakai (1991) reviewed chemometric techniques in flavor research as well as methods for multivariate analysis. Forina et al. (1987) reviewed theory and application of chemometrics in food chemistry. Applications of pattern recognition for quality control can be found in papers by Page (1986), Jeon (1991), and Resurreccion (1988).

There are two categories of classification in multivariate analysis techniques: supervised and unsupervised, depending on whether the sample grouping is known in advance.

Unsupervised methods do not require information for classification. These methods cluster individual samples on the basis of similarity among their data (Aishima and Nakai, 1991). The most popular unsupervised method is principal component analysis (PCA). PCA is a technique to reduce dimensionality of the data. It computes a few linear combinations of the original variables which can be used to summarize the data with minimal loss of information. PCA has been applied for whisky (Headley and Hardy, 1989), soy sauce (Aishima, 1979), wines (Heymann and Noble, 1987), and sugar cane (Cadet et al., 1991).

Supervised learning methods assume that the user has information about the groups prior to application of the algorithms. Multiple regression analysis (MRA), linear discriminant analysis (LDA), principal component regression analysis (PCR), partial least-squares regression analysis (PLS), and artificial neural networks (ANN) have been used for supervised pattern recognition (Aishima and Nakai, 1991).

PCR is a combination of principal component analysis and linear regression analysis. It provides the possibility of relating blocks of variables and allows an unknown pattern to be classified and predicted (Aishima and Nakai, 1991). PCR was applied for the prediction of shelf life for pasteurized milk with a standard error of estimate of 1.3 days within the anticipated shelf life of 21 days (Vallejo-Cordoba and Nakai, 1994).

PLS is one of several multivariate calibration techniques. Banks et al. (1992) applied PLS to gas chromatographic data from cheddar cheese. An excellent correlation was found between gas chromatographic data and sensory scores of cheese samples with various ages. PLS gave more reliable predictions than PCR. Martens et al. (1983) applied PLS for determining relationships between different dependent variables and sensory descriptive independent variables of cauliflower. Arteaga et al. (1994) applied PLS to fourth-derivative ultraviolet spectrums to determine the composition of protein mixtures. They found a good correlation between measured and predicted protein composition. The standard errors of prediction for 16 test samples were 13.4, 5.5, and 11.9% for α_{s1}-, β-, and κ-casein, respectively, and the correlation coefficients between measured and predicted composition were 0.91, 0.99, and 0.94 for the three proteins.

Recently, artificial neural networks (ANN) have become the focus of interest in many disciplines includ-

* To whom correspondence should be addressed. Telephone: (604) 822-3404. FAX: (604) 822-3959. E-mail: yasumih@unixg.ubc.ca.

S0021-8561(96)00660-7 CCC: $14.00 © 1997 American Chemical Society

ing food science. The ANN are computer techniques which simulate the massive parallel structure of the brain (Eberhart and Dobbins, 1990). There are two principal types of network architecture: feed-forward and feedback (Lawrence, 1991). The most popular method is by example and repetition, also called back-propagation network (BPN). BPN has been extensively studied (Eberhart and Dobbins, 1990; Jansson, 1991; Lawrence, 1991; Wythoff, 1993).

The BPN technique is one example of supervised learning in feed-forward networks, in which the learning rule is a mathematical equation known as the Δ rule, or the related least mean squares rule, which minimizes errors between the known values and the network responses (Lawrence, 1991). The BPN is usually built from three layers: input, hidden, and output. The first layer, called the input layer, takes the input values of a pattern. The last layer, called the output layer, produces the pattern outputs. The layers between are called the hidden layers. Each layer has neurons, which are also called processing elements, units, or cells. The strength of a connection between two neurons is called the weight, which determines the magnitude of effect which one neuron can have on the other (Eberhart and Dobbins, 1990). The weighted signals are summed to form a net value. Usually they are simply added together. Total input is run through the activation function, which specifies what the neuron is to do with the signals after the weights have had their effect (Eberhart and Dobbins, 1990). The transfer function is then applied to the activation values to produce output. In a training sequence, the output of the network is compared to known values and errors are back-propagated to the hidden and input layers to adjust the weights and minimize the error. This is repeated many times until the errors between the output and known values are minimized.

A neural network produced better simulation of experimental foam capacity of food proteins than did PCR (Arteaga and Nakai, 1993). Horimoto et al. (1995) also reported better prediction ability of a neural network than PCR for wheat quality for breadmaking.

MATERIALS AND METHODS

Data sampling, data manipulation, and dynamic headspace gas chromatographic analysis were conducted as described in the preceding paper (Horimoto et al., 1996). Group I consisted of milk inoculated with *P. fragi* and *P. fluorescens*. Group II consisted of milk inoculated with *L. lactis*, *E. aerogenes, B. subtilis,* and a mixed culture (*P. fragi*:*L. lactis*:*E. aerogenes* = 1:1:1). Three supervised multivariate analyses were applied: artificial neural networks (ANN), partial least-squares regression analysis (PLS), and principal component regression analysis (PCR). To estimate the true predictive ability of each method, cross-validation was used (Borggard and Thodberg, 1992). The data set is divided into two groups: training and testing data. The model is fitted to the training data set. Predictions are calculated by fitting the model to the testing data set.

For the experiments with *P. fragi* and *P. fluorescens*, 30 samples were divided into training data (24 samples) and testing data (six samples). For the experiments with *L. lactis*, *E. aerogenes*, *B. subtilis,* and a mixed culture, 104 samples were divided into training data (89 samples) and testing data (15 samples). First, random numbers were generated for the samples with Lotus 123 software (version 3.0, Lotus Development Corp., Cambridge, MA). Then samples were arranged in ascending order according to the generated random numbers. The testing data were picked up from the top. The remaining data were used as the training data. After a model was calculated using training data, each class of samples in the testing data was predicted. The statistical parameters of coefficient of determination (r^2) and standard error of prediction (SEP) for the known and predicted values were employed to estimate predictive ability of each method. This procedure was repeated five times.

Artificial Neural Networks (ANN). The neural network software program "Brainmaker" (California Scientific Software, Nevada, CA) was used. A three-layer neural network was used to predict classes using the back-propagation algorithm. A sigmoid function was used as a transfer function because the sigmoid function is particularly useful for a nonlinear relationship (Lawrence and Peterson, 1992). As input neurons for networks, 24 variables for *P. fragi* and *P. fluorescens* and 38 variables for *L. lactis, E. aerogenes, B. subtilis,* and the mixed culture were used. The number of output neurons was one, which represents each group. Since output values were groups, each sample was expressed with a two-digit number, the first digit indicating bacterial species and the second storage time. Arbitrary ranges for each class were used. As the first digits, 1, 2, and 3, were assigned to negative control, milk inoculated with *P. fragi*, and milk inoculated with *P. fluorescens*, respectively. The second digits, 1, 2, 3, 4, and 5, were assigned to storage days 0, 2, 4, 6, 8, and 10, respectively. In the case of *L. lactis, E. aerogenes, B. subtilis,* and a mixed culture, 1, 2, 3, 4, and 5 as the first digit were assigned to negative control, milk inoculated with *L. lactis*, *E. aerogenes*, *B. subtilis,* and a mixed culture, respectively. Storage times 0, 4, 8, 12, and 24 h were assigned 1, 2, 3, 4, and 5 as the second digit, respectively.

The number of hidden neurons is an important factor for the effectiveness of a network. Network performance may vary with the number of the hidden neurons (Lawrence and Peterson, 1992). With too many neurons, a network may not learn but instead memorize patterns, or it may train and run more slowly. On the other hand, without enough hidden neurons, a network may not be trainable (Lawrence and Peterson, 1992). Therefore the number of hidden neurons was diminished starting from a default number, which is the average of the number of input neurons and the output neurons (Lawrence and Peterson, 1992). The default parameters were used for the learning rate (1.00) and momentum factor (0.9).

Partial Least-Squares Regression Analysis (PLS) and Principal Component Regression Analysis (PCR). PLS and PCR were performed using the commercial software "PLSplus Version 2.1" and add-on software to the spectroscopic/chromatographic software system "LabCalc" (Galactic Industries Co., Salem, NH). The optimum number of factors for PCR and PLS was determined using cross-validation procedures as described by Martens and Naes (1989). The input and output variables were the same as those used for ANN.

RESULTS AND DISCUSSION

Artificial neural networks (ANN), partial least-squares regression analysis (PLS), and principal component regression analysis (PCR) were applied to the gas chromatographic data. Before being compared, each method was optimized regarding the number of hidden neurons of ANN and the number of factors of PLS and PCR.

Effects of Inoculating Milk with *P. fragi* and *P. fluorescens*. The number of hidden neurons was optimized for the effectiveness of the ANN. With too many hidden neurons, a network may not learn but instead memorize patterns, or it may train and run too slowly. On the other hand, without enough hidden neurons, a network may not be trained (Lawrence and Peterson, 1992). Thus the number of hidden neurons was decreased starting from the default number. The default number was the sum of the numbers of input and output divided by 2 (Lawrence and Peterson, 1992). ANN with too few hidden neurons cannot be completely trained. Therefore, the training time was set for one-half hour at maximum.

Figure 1 shows results using all data for training ANN. The standard error of prediction (SEP) varied depending on the content of testing data. The pattern

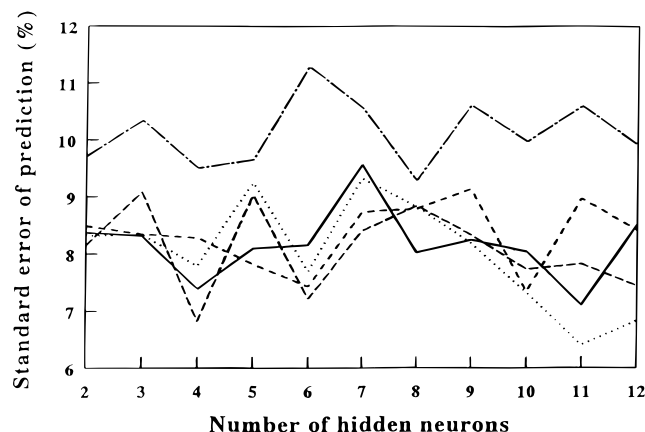

Figure 1. Standard error of prediction (%) by artificial neural networks (ANN) using different hidden neurons for *P. fragi* and *P. fluorescens*: (—) trial 1; (- - -) trial 2; (— — —) trial 3; (···) trail 4; (— · —) trial 5. Training data include testing data.

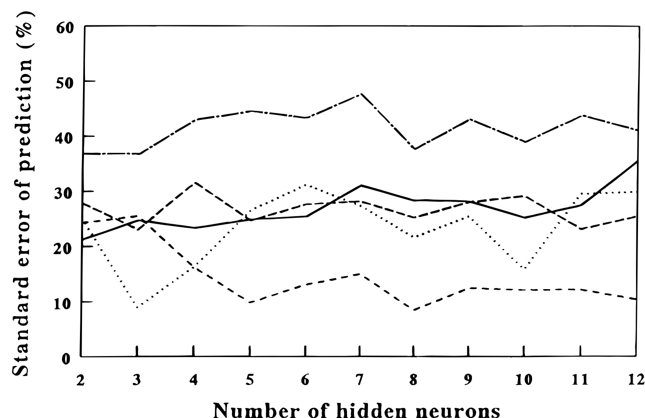

Figure 2. Standard error of prediction (%) by artificial neural networks (ANN) using different hidden neurons for *P. fragi* and *P. fluorescens*: (—) trial 1; (- - -) trial 2; (— — —) trial 3; (···) trail 4; (— · —) trial 5. Training data do not include testing data.

of trial 5 was different from those of the other trials. The difference of SEP within each trial was small (2~3%). There was no trend with different numbers of hidden neurons. Generally, only a small difference in SEP was found with various numbers of hidden neurons. This indicates that the number of hidden neurons did not influence the system. The difference was presumed to come from another parameter, random weight for each connection. The strength of a connection between two neurons is called the weight. It determines the amount of effect that one neuron can have on the other (Lawrence and Peterson, 1992). This weight is randomly assigned. The first assigned weight might influence a network system.

Figure 2 shows prediction ability of cross-validated ANN. Compared with Figure 1, the SEP was much greater. The SEP in trial 5 was consistently greater than that of other trials, while the SEP in trial 2 was smaller. This indicates that the SEP varied depending on the training and testing data. SEP was not greatly different within each trial. These results suggest that the prediction ability is dependent on the content of training and testing data. There was no trend with different numbers of hidden neurons, except trial 2, where the SEP became smaller with increasing hidden neurons. Finally, the number of hidden neurons which had the smallest SEP in the ANN in each trial was used for the comparison with PCR and PLS.

Table 1 presents a comparison of the prediction ability of ANN, PLS, and PCR. A model for each method was

Table 1. Comparison of Prediction Ability of Artificial Neural Networks (ANN), Partial Least-Squares Regression Analysis (PLS), and Principal Component Regression Analysis (PCR) for *P. fragi* and *P. fluorescens* (Training Data Include Testing Data)[a]

	SEP (%)[b]			r^2 [c]		
trial	ANN	PLS	PCR	ANN	PLS	PCR
1	7.1	21.9	23.9	0.98	0.73	0.67
2	7.3	17.7	19.8	0.99	0.88	0.96
3	6.8	9.0	26.1	0.98	0.98	0.75
4	6.4	16.8	15.4	0.98	0.83	0.89
5	9.3	19.5	32.0	0.96	0.78	0.43
mean	$7.4^{**A1,B}$	17.0^{**A1}	23.4^{**B}	0.98^{*A2}	0.84^{*}	0.74^{*A2}
SD[d]	1.1	4.9	6.3	0.01	0.10	0.21
CV (%)[e]	15.3	28.7	26.9	1.1	11.4	28.0

[a] *, significant ($P < 0.05$); **, significant ($P < 0.01$) in one-way ANOVA test. Superscript A1, A2, significant ($P < 0.01$); superscript B, significant ($P < 0.05$) in a Turkey HSD test. [b] Standard error of prediction was divided by range of experimental values in testing. [c] Coefficient of determination. [d] Standard deviation. [e] Coefficient of variation.

Table 2. Comparison of Prediction Ability of Artificial Neural Networks (ANN), Partial Least-Squares Regression Analysis (PLS), and Principal Component Regression Analysis (PCR) for *P. fragi* and *P. fluorescens* (Training Data Do Not Include Testing Data)

	SEP (%)[a]			r^2 [b]		
trial	ANN	PLS	PCR	ANN	PLS	PCR
1	21.2	20.3	26.0	0.75	0.78	0.70
2	8.5	28.9	35.7	0.96	0.89	0.81
3	23.1	43.7	45.3	0.93	0.61	0.52
4	9.0	20.2	21.8	0.95	0.85	0.79
5	36.8	13.5	37.5	0.41	0.30	0.13
mean	21.7	29.5	33.3	0.80	0.69	0.59
SD[c]	10.0	10.0	9.4	0.23	0.24	0.28
CV (%)[d]	46.0	33.8	28.2	29.3	35.1	47.7

[a] Standard error of prediction was divided by the range of experimental values in testing. [b] Coefficient of determination. [c] Standard deviation. [d] Coefficient of variation.

trained with all data. The number of factors for PLS and PCR were optimized automatically while running the software (PLSplus, 1992). PLS used three factors and PCR used two factors for the optimum model.

The mean standard errors of prediction (SEP) were 7.4%, 17.0%, and 23.4% for ANN, PLS, and PCR, respectively. The SEP among methods was significantly different ($P < 0.01$) in the one-way ANOVA test. To determine which of the methods is significantly different from each other, a Turkey HSD test was applied. The result is also shown in Table 1 using a letter (A and B). "A" or "B" indicates a significant difference at $P < 0.01$ or $P < 0.05$, respectively. The SEP of the ANN was significantly different from those of PLS ($P < 0.05$) and PCR ($P < 0.01$). The difference between PLS and PCR was not significant. Coefficients of variation for the SEP were 15.3%, 28.7%, and 26.9% for the ANN, PLS, and PCR, respectively. The ANN consistently gave better prediction ability.

The mean coefficients of determination (r^2) were 0.98, 0.84, and 0.74 for the ANN, PLS, and PCR, respectively. There were significant differences among the methods ($P < 0.05$). From a Turkey HSD test, the r^2 of the ANN was significantly different from that of PCR ($P < 0.01$). The coefficient of variation for r^2 of the ANN was much smaller than those of both PLS and PCR.

Table 2 shows the prediction ability of each method using cross-validation. The mean SEP were 21.7%, 29.5%, and 33.3% for the ANN, PLS, and PCR, respectively. The mean r^2 between actual and predicted values were 0.80, 0.69, and 0.59 for the ANN, PLS, and PCR, respectively. The SEP and r^2 among each method were not significantly different from those of the one-

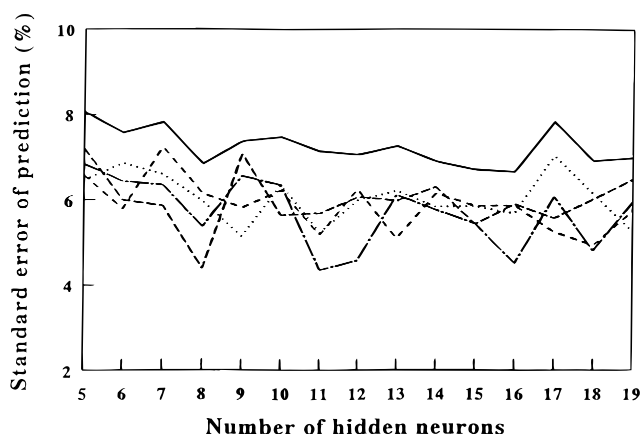

Figure 3. Standard error of prediction (%) by artificial neural networks (ANN) using different hidden neurons for *L. lactis, E. aerogenes, B. subtilis,* and a mixed culture: (—) trial 1; (- - -) trial 2; (— — —) trial 3; (···) trail 4; (— · —) trial 5. Training data include testing data.

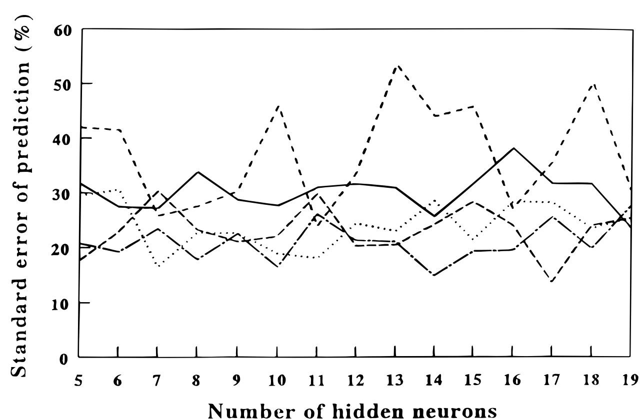

Figure 4. Standard error of prediction (%) by artificial neural networks (ANN) using different hidden neurons for *L. lactis, E. aerogenes, B. subtilis,* and a mixed culture: (—) trial 1; (- - -) trial 2; (— — —) trial 3; (···) trail 4; (— · —) trial 5. Training data do not include testing data.

way ANOVA test ($P > 0.05$). The coefficient of variation of the ANN was much larger than that of PLS and PCR. In ANN, trial 2 gave the smallest SEP. The difference between the smallest and largest SEP for the ANN was about 28%. Even though the ANN generally had smaller SEP than PLS and PCR, the content of training and testing data have great influence on prediction ability in the ANN.

Generally, the ANN was able to make better predictions than PCR and PLS. This indicates that the relationship between dependent and independent values may be nonlinear. The ANN was also faster and easier to use than PCR and PLS. The result using all data was better than cross-validated results. This suggests that a larger data set increases prediction ability for the ANN.

Effects of Inoculating Milk with *L. lactis, E. aerogenes, B. subtilis,* and a Mixed Culture. Figure 3 presents the performance of the ANN with differing numbers of hidden neurons using all data for training. The SEP tended to be smaller with more hidden neurons. Small variations of the SEP were observed within each trial (2~3%).

Figure 4 shows the performance of the cross-validated ANN. Except in trial 2, there was a small difference for SEP with different numbers of hidden neurons. Compared with Figure 3, the SEP was much greater. These trends were similar to those for milk inoculated with *P. fragi* and *P. fluorescens*.

The ANN with the smallest SEP was used for the comparisons with PCR and PLS. The number of factors for PLS and PCR were 5 and 6, respectively. Table 3 shows the comparisons of the three methods. A model was trained with all data. The mean SEP were 5.2%, 21.1%, and 25.1% for ANN, PLS, and PCR, respectively. There were significant differences ($P < 0.01$) among the three methods. In the Turkey HSD test, the SEP of the ANN was significantly different from those of PLS and PCR ($P < 0.01$). The difference between PLS and PCR was significant ($P < 0.05$). Coefficients of variation for the SEP were 16.8%, 14.3%, and 9.5% for the ANN, PLS, and PCR, respectively. These values suggest that prediction ability of the ANN varied depending on testing data, even though the ANN consistently gave smaller SEP than PLS and PCR. PCR consistently had poorer prediction ability than the ANN and PLS. The mean r^2 values were 0.98, 0.71, and 0.56 for the ANN, PLS, and PCR, respectively. There were significant differences ($P < 0.01$) among the three methods. The r^2 of the ANN was significantly different from those of PLS and PCR ($P < 0.01$). The difference between PLS and PCR was also significant ($P < 0.05$).

Table 4 shows the cross-validation results of ANN, PLS, and PCR. The mean SEP were 18.8%, 26.9%, and 26.2% for the ANN, PLS, and PCR, respectively. They were significantly different ($P < 0.01$) from those of the one-way ANOVA. From a Turkey HSD test, the ANN had a significantly smaller SEP than PLS ($P < 0.01$) and PCR ($P < 0.05$). No significant difference was found between PLS and PCR. The mean r^2 values were 0.73, 0.50, and 0.53 for the ANN, PLS, and PCR, respectively. They were significantly different ($P < 0.05$). The differences between the ANN and PLS and the ANN and PCR were significant ($P < 0.05$) in the Turkey HSD test. The difference between PLS and PCR was not significant. ANN gave the best predictive ability. The values for the SEP and r^2 of PCR for the cross-validated

Table 3. Comparison of Prediction Ability of Artificial Neural Networks (ANN), Partial Least-Squares Regression Analysis (PLS), and Principal Component Regression Analysis (PCR) for *L. lactis, E. aerogenes, B. subtilis,* and a Mixed Culture (Training Data Include Testing Data)[a]

trial	SEP (%)[b]			r^2 [c]		
	ANN	PLS	PCR	ANN	PLS	PCR
1	6.7	22.8	27.6	0.97	0.73	0.45
2	5.1	24.5	26.9	0.98	0.63	0.60
3	4.9	22.3	24.0	0.98	0.63	0.60
4	5.1	17.5	25.4	0.98	0.74	0.45
5	4.4	18.4	21.6	0.98	0.82	0.70
mean	5.2**A1,A2	21.1**A1,B1	25.1**A2,B1	0.98**A3,A4	0.71**A3,B2	0.56**A4,B2
SD[d]	0.9	3.0	2.4	0.004	0.08	0.11
CV (%)[e]	16.8	14.3	9.5	0.46	11.4	19.4

[a] *, significant ($P < 0.05$); **, significant ($P < 0.01$) in one-way ANOVA test. Superscript A1, A2, A3, A4, significant ($P < 0.01$); superscript B1, B2, significant ($P < 0.05$) in a Turkey HSD test. [b] Standard error of prediction was divided by range of experimental values in testing. [c] Coefficient of determination. [d] Standard deviation. [e] Coefficient of variation.

Table 4. Comparison of Prediction Ability of Artificial Neural Networks (ANN), Partial Least-Squares Regression Analysis (PLS), and Principal Component Regression Analysis (PCR) for *L. lactis, E. aerogenes, B. subtilis,* and a Mixed Culture (Training Data Do Not Include Testing Data)[a]

trial	SEP (%)[b]			r^2 [c]		
	ANN	PLS	PCR	ANN	PLS	PCR
1	24.8	29.9	29.1	0.58	0.37	0.50
2	24.2	26.3	27.9	0.59	0.61	0.51
3	14.3	26.1	25.6	0.84	0.52	0.48
4	15.7	26.7	26.2	0.81	0.43	0.46
5	15.0	25.5	22.1	0.83	0.59	0.70
mean	18.8**A,B1	26.9**A	26.2**B1	0.73*B2,B3	0.50*B2	0.53*B3
SD[d]	5.2	1.7	2.7	0.13	0.10	0.10
CV (%)[e]	27.8	6.3	10.2	18.2	20.4	18.3

[a] *, significant ($P < 0.05$); **, significant ($P < 0.01$) in one-way ANOVA test. Superscript A, significant ($P < 0.01$); superscript B1, B2, B3, significant ($P < 0.05$) in a Turkey HSD test. [b] Standard error of prediction was divided by range of experimental values in testing. [c] Coefficient of determination. [d] Standard deviation. [e] Coefficient of variation.

data were better than those for PLS. However, the differences were not significant ($P > 0.05$). Therefore, there was no difference of predictive ability between PLS and PCR ($P > 0.05$).

From the above two experiments, it is concluded that the ANN gave the best prediction ability among three supervised methods. The prediction ability of PLS was better than PCR. However, the performance of each method was dependent on the content of training and testing data; the more data, the better the prediction ability. Each method gave better predictive ability when trained with all data.

CONCLUSIONS

Three supervised multivariate analyses, artificial neural networks (ANN), partial least-squares regression analysis (PLS), and principal component regression analysis (PCR), were applied to peak areas in chromatograms of UHT-sterilized milk samples with different bacterial species and storage times. The statistical parameters of coefficient of determination (r^2) and standard error of prediction (SEP) were used to estimate predictive ability of each method. ANN gave the best mean r^2 and SEP among the supervised methods. The coefficient of variation of SEP of the ANN was much larger than that of PLS and PCR. Even though the ANN generally had a smaller SEP and larger r^2 than PLS and PCR, the content of training and testing data had great influence on the predictive ability of the ANN. Generally, the ANN was able to make better predictions than PLS and PCR. This indicates that the relationship between dependent and independent variables may be nonlinear. The ANN was also faster and easier to use than PLS and PCR. The results from using all data were better than cross-validated results. This suggests that a larger data set increases predictive ability for the ANN. Other parameters such as learning rate, momentum factor, noise addition, weight control, and number of hidden neurons can be optimized to improve ANN predictive ability.

LITERATURE CITED

Aishima, T. Classification of soy sauce on principal components in GC profiles. *Agric. Biol. Chem.* **1979**, *43*, 1905–1910.

Aishima, T.; Nakai, S. Chemometrics in flavour research. *Food Rev. Int.* **1991**, *7*, 33–101.

Arteaga G. E.; Nakai, S. Predicting protein functionality with artificial neural networks: foaming and emulsifying properties. *J. Food Sci.* **1993**, *58*, 1152–1156.

Arteaga, G. E.; Horimoto, Y.; Li-Chan, E.; Nakai, S. Partial least squares regression of fourth derivative of ultra violet absorbance spectra predicts composition of protein mixtures: application to bovine caseins. *J. Agric. Food Chem.* **1994**, *42*, 1938–1942.

Banks, J. M.; Brechany, E. Y.; Christie, W. W.; Hunter, E. A.; Muir, D. D. 1992. Volatile components in steam distillates of cheddar cheeses as indicator indices of cheese maturity, flavour and odour. *Food Res. Int.* **1992**, *25*, 365–373.

Borggard, C.; Thodberg, H. H. Optimal minimal neural interpretation of spectra. *Anal. Chem.* **1992**, *64*, 545–551.

Brainmaker software, 6th ed.; California Scientific Software: Nevada, CA, 1992.

Cadet, F.; Bertrand, D.; Pobert, P.; Maillot, J.; Dieudonne, J.; Rouch, C. Quantitative determination of sugar cane sucrose by multidimensional statistical analysis of their mid-infrared attenuated total reflectance spectra. *Anal. Spectrosc.* **1991**, *45*, 166–172.

Eberhart, R. C., Dobbins, R. W., Eds. Implementations. In *Neural Networks PC Tools: A Practical Guide*; Academic Press, Inc.: San Diego, CA, 1990.

Forina, M.; Lanteri, S.; Armanino, C. Chemometrics in food chemistry. *Top. Curr. Chem.* **1987**, *141*, 193–143.

Headley, L. M.; Hardy, J. K. Classification of whiskies by principal component analysis. *J. Food Sci.* **1989**, *4*, 1351–1358.

Heymann, H.; Noble, A. C. Descriptive analysis of commercial cabernet sauvignon wines from California. *Am. J. Enol. Vitic.* **1987**, *38*, 41–44.

Horimoto, Y.; Durance, T.; Nakai, S.; Lukow, O. M. Neural networks vs principal component regression for prediction of wheat flour loaf volume in baking tests. *J. Food Sci.* **1995**, *60*, 429–433.

Horimoto, Y.; Lee, K.; Nakai, S. Classification of microbial defects in milk using dynamic headspace gas chromatography and computer-aided data processing. 1. Principal component similarity (PCS) analysis. *J. Agric. Food Chem.*, submitted for publication.

Jansson, P. A. Neural networks: an overview. *Anal. Chem.* **1991**, *63*, 375A–362A.

Jeon, I. J. Pattern recognition techniques for food research and quality assurance. In *Instrumental Methods for Quality Assurance in Foods*; Fung, D. C., Richard, F. M., Eds.; Marcel Dekker Inc.: New York, 1991; pp 271–501.

Lawrence, J. *Introduction to Neural Networks*; Luedeking, S., Ed.; California Scientific Software: Nevada, CA, 1991.

Lawrence, J.; Peterson, A. *Brainmaker: User's Guide and Reference Manual*; Luedeking, S., Fredrickson, J., Eds.; California Scientific Software: Nevada, CA, 1992.

Lotus 123; Lotus Development Corp.: Cambridge, MA, 1990.

Martens, M.; Martens, H.; Wold, S. Preference of cauliflower related to sensory descriptive variables by partial least squares (PLS) regression. *J. Sci. Food Agric.* **1983**, *34*, 715–724.

Martins, H., Naes, T., Eds. *Multivariate Calibration*; John Wiley & Sons Ltd.: Chichester, U.K., 1989.

Page, S. W. Pattern recognition methods for the determination of food composition. *Food Technol.* **1986**, *40*, 11, 104–109.

PLS plus; Galactic Industries Corp.: Salem, NH, 1992.

Resurreccion, A. V. A. Applications of multivariate methods in food quality evaluation. *Food Technol.* **1988**, *42*, 11, 128–134.

Vallejo-Cordoba, B. and Nakai, S. Keeping-quality assessment of pasteurized milk by multivariate analysis of dynamic headspace gas chromatographic data. 1.Shelf life prediction by principal component regression. *J. Agric. Food Chem.* **1994**, *42*, 989–993.

Wythoff, B. J. Backpropagation neural networks: a tutorial. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 115–155.

Received for review May 22, 1996. Revised manuscript received December 4, 1996. Accepted December 13, 1996.⊗

JF960660G

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1997.