

Feasibility of Near-Infrared Spectroscopy to Detect and to Quantify Adulterants in Cow Milk

Sumaporn KASEMSUMRAN,*† Warunee THANAPASE,* and Artaya KIATSOONTHON**

*Kasetsart Agricultural and Agro-Industrial Product Improvement Institute, Kasetsart University, 50 Pahonyothin, Jatujak, Bangkok 10900, Thailand

**Bureau of Quality Control of Livestock, Department of Livestock Development, Ministry of Agriculture and Cooperatives, Pathumthani 12000, Thailand

Cow milk adulteration involves the dilution of milk with a less-expensive component, such as water or whey. Near-infrared spectroscopy (NIRS) was employed to detect the adulterations of milk, non-destructively. Two adulteration types of cow milk with water and whey were prepared, respectively. NIR spectra of milk adulterations and natural milk samples in the region of 1100–2500 nm were collected. The classification of milk adulterations and natural milk were conducted by using discriminant partial least squares (DPLS) and soft independent modelling of class analogy (SIMCA) methods. PLS calibration models for the determination of water and whey contents in milk adulteration were also developed, individually. Comparisons of the classification methods, wavelength regions and data pretreatments were investigated, and are reported in this study. This study showed that NIR spectroscopy can be used to detect water or whey adulterants and their contents in milk samples.

(Received March 31, 2007; Accepted May 30, 2007; Published July 10, 2007)

Food quality and safety are of growing interest to consumers and to regulators. Cow milk adulteration is one of the most common types of sophisticated food fraud. Adulteration is simply achieved by adding common adulterants (water or whey) into natural milk. Thus, the food quality of milk is considerably reduced by adulteration. The detection of adulterants in milk is not straightforward, and may require more than one method for verification. Currently, various physical techniques, using differences in the freezing point and specific gravity, are carried out for detecting milk adulteration. However, these methods are time consuming and may not yield accurate detection results. As for quantitative analysis, the use of mid-infrared (MIR) spectroscopy has widely spread in the determination of common constituents in milk.¹ Unfortunately, this method involves destructive techniques for milk samples as well as high operation cost. Presently, near-infrared (NIR) spectroscopy in combination with chemometrics is widely employed in food analysis, including milk examination,^{2–9} because it involves a fast, non-destructive and inexpensive method, as well as easy operation. Especially, the NIR method has been studied in the quantitative determination of many constituents in milk,^{2–8} because it can be employed to determine several components in milk at once. Regarding milk adulteration, Jha and Matsuoka⁹ have used short-wavelength of NIR in the region of 700–1124.8 nm to determine such adulterants as vegetable oil, urea, NaOH and shampoo components, in which samples were prepared by mixing those components and milk powder into water as synthetic milk.

The present study has two proposes. One is to investigate the feasibility of using a NIR technique, while aiming at the detection and quantification of water and whey as adulterants in

natural milk. The other is to demonstrate the potential of a chemometric method in combination with a NIR technique for the classification milk adulteration from natural milk and for the determination of adulterant contents in milk samples.

Experimental

Samples

Milk samples from cows were purchased from Kasetsart University dairy center, Bangkok, Thailand. Whey (Fonterra, Newzealand) was donated from a dairy-product company in Thailand. Two sample sets of milk adulteration were prepared. The first adulteration sample set was produced by mixing water and milk at a ratio of 1.00–97.00% (v/v); the other set was prepared by adding whey into milk at a ratio of 2.15–48.40% (w/v). Totally, 90 adulterations of natural milk mixed with water ($n = 50$) or whey ($n = 40$) were prepared. Table 1 summarizes the distribution of the concentrations of the adulterants in the thus-prepared samples.

NIR acquisition

The NIR spectra in the region of 1100–2500 nm were collected for a mixture of samples and natural milk samples. The NIR measurements were performed using the reflectance

Table 1 Distribution of the contents of adulterants in milk samples

Adulterant	Min	Max	Mean
Water, % (v/v)	1.00	97.00	49.00
Whey, % (g/v)	2.15	48.40	25.25

† To whom correspondence should be addressed.
E-mail: aapspk@ku.ac.th

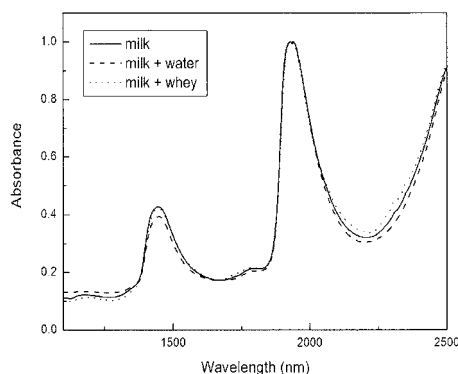


Fig. 1 Three mean NIR spectra in the region of 1100–2500 nm of natural milk (solid line), milk adulterated by water (dash line), and milk adulterated by whey (dotted line).

mode in the region of 1100–2500 nm by means of an InfraAlyzer 500 NIR reflectance analyzer (BRAN+LUEBBE, Germany) equipped with a PbS detector. The temperature of the sample was kept constant by using an IUCHI Thermal ROBO TR2 temperature controller.

Spectral analysis

Sesame (Ver. 3.1: BRAN+LUEBBE, Germany) was used for spectral data collection, in which they were converted into JCAMP files for Unscrambler (Ver. 9.6: CAMO AS, Trondheim, Norway). Discriminant PLS (DPLS), SIMCA and PLS regression were calculated by using Unscrambler.

In the present study, comparisons of using two classification methods, discriminant partial least squares (DPLS) and soft independent modeling of class analogy (SIMCA), were investigated to classify natural milk, milk adulterated by whey and milk adulterated by water, with various wavelength ranges and pretreatment methods being used. The DPLS method is an alternative method for principle component analysis (PCA) in discrimination analysis. It optimizes the fitting and prediction to {0/1}, coded membership, indicating variables in the development of latent variables.^{10,11} SIMCA classification is based on making a PCA model for each class in the training set. Unknown samples are then compared to class models and assigned to classes, according to their analogy to training samples.¹⁰ Several groups have evaluated the potential of these two classification methods with various samples.^{11–14}

For the present study, the DPLS analysis involved giving values of -1.0 , 0.0 and 1.0 to the NIR spectra of milk adulterated by whey, natural milk, and that adulterated by water, respectively. Because there are three classes, PLS2 was used with one response variable coding for each class. The PLS model was built, and was then cross-validated. Samples with the predicted result were classified as follows: milk adulterated by whey ≤ -0.5 , $-0.5 <$ natural milk < 0.5 , and milk adulterated by water ≥ 0.5 , respectively. Furthermore, individual PLS calibration models for the quantitative determination of adulterant contents in milk were developed, respectively.

The NIR spectra were subjected to a multiplicative scatter correction (MSC) and 2nd derivative (9-point Savitsky-Golay filter) before developing PLS models for the classification and quantitative determination. Four wavelength regions, *i.e.* the whole region, 1100–1850, 2048–2500 nm and the combination of 1100–1850 and 2048–2500 nm wavelength regions were employed for developing the classification and PLS models. The NIR spectra of each adulterant, *i.e.* that with

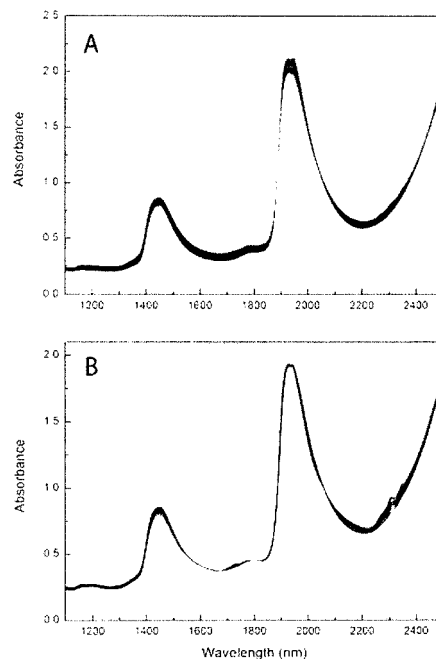


Fig. 2 NIR spectra of adulteration milks in the region of 1100–2500 nm; A, milk adulterated by water; B, milk adulterated by whey.

water, or whey, were used for both calibration and validation. PLS calibration models for quantification of the adulterant contents in natural milk were built separately. Full cross validation was used to validate and find the optimum number of PLS factors for the models used for the classification and quantitative determinations.

Results and Discussion

Figure 1 shows the normalization of three mean NIR spectra in the region of 1100–2500 nm of natural milk and milk adulterated by water or whey. They illustrate two broad bands around 1450 and 1900 nm, respectively. The band of 1450 nm is mainly due to the combination of OH symmetric and antisymmetric stretching modes of water.¹⁵ The other band is assigned to the combination mode of the OH stretching and deformation vibrations of water.¹⁵ It can be seen that the intensity of the water bands increase with an increase in the content of water added into milk samples (Fig. 2A). On the other hand, the intensity of the band in the 2300–2400 nm wavelength region changes with a change in whey content in samples (Fig. 2B). This region is rather rich in protein bands, and may contain overtone and combination bands for lactose as well.^{16,17} Nevertheless, those mean NIR spectra are very similar to each other, and it is not easy to classify the spectra. Therefore, the DPLS and SIMCA methods were applied in this study to classify between natural and adulterated milk samples.

The statistical results for the DPLS regressions and the SIMCA are reported in Tables 2 and 3, respectively. These classification models were developed using the different spectral regions and pretreatment spectra. The classification performances of these two methods were compared. The results show that the DPLS method yields better classification results than those using the SIMCA method. The percentage of correct classifications for the validation set (%CP) is an important factor that express the classification performance in this study.

Table 2 Statistical results of DPLS analysis for classifying adulterants in milk samples

Pretreatment	Wavelength/nm							
	Whole region		1100 – 1850		2048 – 2500		1100 – 1850, 2048 – 2500	
	Factor	%CP	Factor	%CP	Factor	%CP	Factor	%CP
Original	9	97.96	8	98.98	8	100.00	8	97.96
MSC	7	97.96	7	98.98	5	93.88	5	88.78
MSC + 2nd-derivative	7	100.00	6	98.98	5	97.96	5	98.98

%CP: %correct classification for validation set.

Table 3 Statistical results of SIMCA analysis for classifying adulterants in milk samples

Pretreatment	Wavelength/nm							
	Whole region		1100 – 1850		2048 – 2500		1100 – 1850, 2048 – 2500	
	MD	%CP	MD	%CP	MD	%CP	MD	%CP
Original	622/3552	81.63	336/1231	73.47	595/8743	84.69	980/8743	82.65
MSC	318/859	84.69	297/1773	83.67	240/506	85.71	450/1648	84.69
MSC + 2nd-derivative	133/347	86.73	122/278	83.67	91/599	75.51	96/550	84.69

MD: model distance. Former value, the distance between model for natural milk and model for milk adulterated by water; latter value, the distance between model for natural milk and model for milk adulterated by whey.

The %CP value obtained from the DPLS method was over 97.00%; on the other hand, the best result of %CP obtained from the SIMCA method was only 86.73%. Nevertheless, the classification performance of SIMCA increased when it was applied to classify the milk samples adding adulterant contents from about of 11%.

In the DPLS results, the best DPLS classification model for natural milk, milk adulterated by water and milk adulterated by whey was developed using the MSC and 2nd-derivative spectra in the whole region of 1100 – 2500 nm with a PLS factor of 7. It gives the best classification performance of 100.00% correct classification of the prediction set (%CP) (Table 2). Therefore, it seems that the MSC can reduce the scattering effect in milk samples,¹⁸ and the 2nd-derivative pretreatment can enhance any spectral differences in the whole region where bands due to any differences in the milk constituents appear. Figure 3 is a plot between the given actual value (*X*-axis) for each class and the DPLS predicted value (*Y*-axis) for validation samples. It is clear from the results in Table 2 and Fig. 3 that the DPLS method is a very powerful technique to classify adulterants in natural milk, such as water or whey.

Table 4 summarizes the statistical results for predicting the concentrations of water contained in milk samples. It can be seen in Table 4 that good prediction results were obtained mostly from models built by using the MSC pretreatment spectra. Due to the power of MSC, one can reduce the scattering effects that occur in milk samples.¹⁸ The best prediction result is obtained for water adulterated in natural milk, when the model is developed by using the MSC spectra over the whole region of 1100 – 2500 nm. Its statistical results are the lowest value of the root mean square error of prediction (RMSEP) 2.159% (v/v) with a PLS factor of 4. This region contains bands arising from the combination of OH symmetric and antisymmetric stretching modes of water.¹⁵ It is noted that the error limit of $\pm 2.159\%$ (v/v) is slightly higher than the minimum water content contained in a milk sample set. Therefore, when milk is adulterated by adding a very small volume of water within the error limitation value, it is not easy

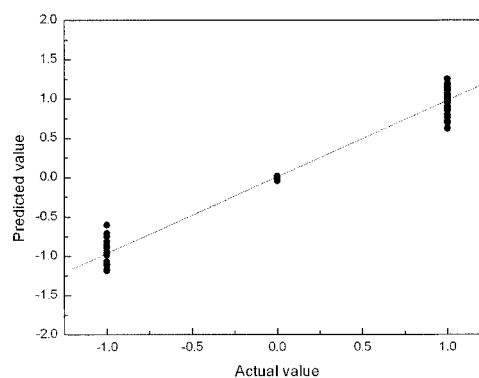


Fig. 3 DPLS predicted vs. actual value for classification of natural milk and milk adulterations.

to determine the content of water in such a milk sample. Perhaps it is due to the fact that typical cow milk normally contains water as a main component. Thus, the small contamination of water in milk is not significantly appear in the NIR spectrum.

As for the prediction result of whey contents in adulteration of milk samples, results are given in Table 5. The best calibration model for milk adulteration by mixing whey yields the prediction result with a RMSEP value of 0.244% (g/v) by a PLS factor of 4. This model was built using the MSC pretreated spectra of the combination regions of 1100 – 1850 and 2048 – 2500 nm. It may be because these regions contain useful information on whey.^{16,17}

Conclusions

The results obtained in this study have demonstrated that NIR spectroscopy is highly feasible to detect such adulterants as water and whey, and to determine their contents in milk

Table 4 Prediction results of PLS calibration models for determining water contents in milk samples

Pretreatment	Wavelength/nm							
	Whole region		1100 – 1850		2048 – 2500		1100 – 1850, 2048 – 2500	
	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)
Original	0.997	2.311 (3)	0.992	3.702 (3)	0.997	2.263 (4)	0.997	2.227 (3)
MSC	0.997	2.159 (4)	0.997	2.377 (3)	0.997	2.173 (4)	0.997	2.219 (3)
MSC + 2nd-derivative	0.996	2.468 (3)	0.997	2.195 (4)	0.996	2.358 (1)	0.997	2.242 (3)

The number in parenthesis is the PLS factor number. R, Correlation coefficient; RMSEP, root mean square error of prediction.

Table 5 Prediction results of PLS calibration models for determining whey contents in milk samples

Pretreatment	Wavelength/nm							
	Whole region		1100 – 1850		2048 – 2500		1100 – 1850, 2048 – 2500	
	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)	R	RMSEP, % (g/v)
Original	0.999	0.264 (4)	0.999	0.701 (2)	0.999	0.490 (5)	0.999	0.348 (3)
MSC	0.999	0.251 (4)	0.999	0.457 (5)	0.999	0.310 (5)	0.999	0.244 (4)
MSC + 2nd-derivative	0.999	0.371 (7)	0.999	0.334 (5)	0.998	0.802 (3)	0.999	0.474 (3)

The number in parenthesis is the PLS factor number. R, Correlation coefficient; RMSEP, root mean square error of prediction.

samples. For the detection of adulterant types, DPLS analysis used a suitable wavelength region of 1100 – 2500 nm and pretreated spectra with MSC and 2nd-derivative methods; thus, an efficient DPLS model was obtained. The optimum DPLS classification model can correctly detect adulterant types with 100.00%CP for the validation set. Moreover, excellent PLS models for the determination of water or whey contents in natural milk were developed. Four wavelength regions and three pretreatment spectra methods using in PLS model development were evaluated. Finally, the most selective PLS calibration model for the quantitative determination of water content in milk built by using MSC pretreated spectra over the whole region of 1100 – 2500 nm yielded prediction results with the lowest error of 2.159% (v/v). For the quantitative analysis of whey in milk, PLS calibration developed using MSC pretreated spectra in the combination regions of 1100 – 1850 and 2048 – 2500 nm yielded excellent results with an error limit of $\pm 0.244\%$ (g/v). The results obtained in this study have revealed that NIR spectroscopy in combination with a chemometric method has a feasible potential to detect and to quantify water or whey in cow milk samples. For further study, we strongly believe that the model milk samples must be enlarged to cover the greatest variation of milk constituents for practical applications in the milk industry.

Acknowledgements

The authors would like to thank Prof. Yukihiro Ozaki (Kwansei Gakuin University, Japan) for valuable suggestions. S. K. gratefully thanks the Thailand Research Fund (TRF) and the Commission on Higher Education (CHE) for financial support.

References

1. R. Grappin and D. Lefier, "Milk Protein Definition and Standardization", **1995**, International Dairy Federation, Brussels, Belgium.
2. B. Czarnik-Matusiewicz, K. Murayama, R. Tsenkova, and Y. Ozaki, in Proceedings of the Ninth International Conference on Near Infrared Spectroscopy, ed. A. M. C. Davies and R. Giangiacomo, **2000**, NIR Publication, UK, 17.
3. H. Kamishikiryō-Yamashita, Y. Oritani, H. Takamura, and T. Matoba, *J. Food Sci.*, **1994**, *59*, 313.
4. J. Renata and S. Kvetoslava, *Czech J. Food Sci.*, **2003**, *21*, 123.
5. M. F. Laporte and P. Paquin, *J. Agric. Food Chem.*, **1999**, *47*, 2600.
6. R. Tsenkova, S. Atanassova, K. Itoh, Y. Ozaki, and K. Toyoda, *J. Anim. Sci.*, **2000**, *78*, 515.
7. S. Slobodan and Y. Ozaki, *Anal. Chem.*, **2001**, *73*, 64.
8. Y. A. Woo, Y. Terazawa, J. Y. Jie, F. Terada, and S. Kawano, *Appl. Spectrosc.*, **2002**, *56*, 599.
9. S. N. Jha and T. Matsuoka, *J. Food Sci. Tec.*, **2004**, *41*, 313.
10. H. Martens, "Multivariate Calibration", **1989**, John Wiley and Sons, New York, Chichester, Brisbane, Toronto.
11. B. K. Alsberg, D. B. Kell, and R. Goodacre, *Anal. Chem.*, **1998**, *70*, 4126.
12. J. McElhinney, G. Downey, and T. Fearn, *J. Near Infrared Spectrosc.*, **1999**, *7*, 145.
13. G. Downey, V. Fouratier, and J. D. Kelly, *J. Near Infrared Spectrosc.*, **2003**, *11*, 447.
14. Y. A. Woo, S. H. Chang, H. J. Kim, and J. Cho, in Proceedings of the Ninth International Conference on Near Infrared Spectroscopy, ed. A. M. C. Davies and R. Giangiacomo, **2000**, NIR Publication, UK, 523.
15. H. Maeda, Y. Ozaki, M. Tanaka, N. Hayashi, and T. Kojima, *J. Near Infrared Spectrosc.*, **1995**, *3*, 19.
16. S. Slobodan and Y. Ozaki, *Appl. Spectrosc.*, **2000**, *54*, 1327.
17. P. Whilliams and K. Norris, "Near-Infrared Technology in the Agricultural and Food Industries", **1990**, American Association of Cereal Chemists, Inc., St. Paul, Minnesota, USA.
18. C. J. Yu, L. Chie, T. Fuminori, and S. Kawano, *J. Near Infrared Spectrosc.*, **2002**, *10*, 301.