SHIMADZU
Excellence in Science

Application News

No.A475

Spectrophotometric Analysis

Quantitative Analysis of Fat in Milk by UV-Vis-NIR Reflectance Spectroscopy and Multivariate Analysis

Milk is one of the most popular drinks among humans. In recent years, however, there has been a great increase in the sales of milk products with adjusted fat content. Milk fat content is typically measured using the Roese-Gottlieb or the Gerber method, but measurement by these methods is extremely time-consuming. We therefore investigated the use of the spectral reflectance method as a simpler quantitative method. By applying a combination of reflectance measurement using a screw-top glass tube in conjunction with multivariate analysis, we found that the fat content could be determined quite easily. As multivariate analysis methods, the multiple linear regression method, PLS method, and support vector regression method (SVR) were used to conduct a comparative analysis of the quantitative accuracy of these quantitative methods. The results indicated that the support vector regression method provided the best quantitative accuracy. The results are introduced in this paper.

Total Luminous Reflectance Measurement of Milk

Nine types of measurement samples with differing levels of fat content (3 types of high-fat milk, 3 types of medium-fat milk, 3 types of low-fat milk) were used. The fat content values displayed on the various milk cartons are shown in Table 1. To ensure that all of the samples were positioned identically, measurement was conducted with the integrating sphere mounted in a securing jig. Each sample was transferred to a screw-top glass tube which was then set in the integrating sphere as shown in Fig. 1, and the total light reflectance was measured twice for each sample by transferring the same sample to a different screw-top glass tube. Thus, a total of 18 data points were obtained (9 × 2 = 18). The disposable screw-top glass tubes were discarded after each use. In addition, a Spectralon® reflectance standard (U. S. Labsphere, Inc.) was used for the reflectance measurements.
The measurement results and measurement conditions are shown in Fig. 2 and Table 2, respectively. In Fig. 2, the red-colored trace lines correspond to high-fat milk, the black trace lines to medium-fat milk, and the blue trace lines to low-fat milk. The data clearly indicate that the lower the fat content, the lower the reflectance becomes overall.
An enlarged view of a portion of Fig. 2 is shown in Fig. 3 (next page). Looking at the blue-trace spectra of the low-fat milk, each of the two respective repeat measurement results overlap nearly perfectly, suggesting that exchanging the screw-top glass tube has little effect. The blue line spectra are clearly divided into 3 groups, corresponding to the descending order of fat content from 2.0, 1.0 and 0.2 (g/200 mL).

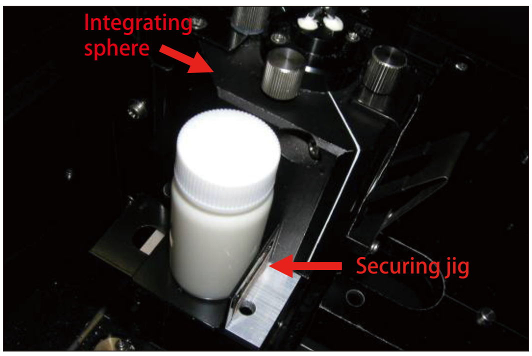

Fig. 1 Sample Set in Integrating Sphere

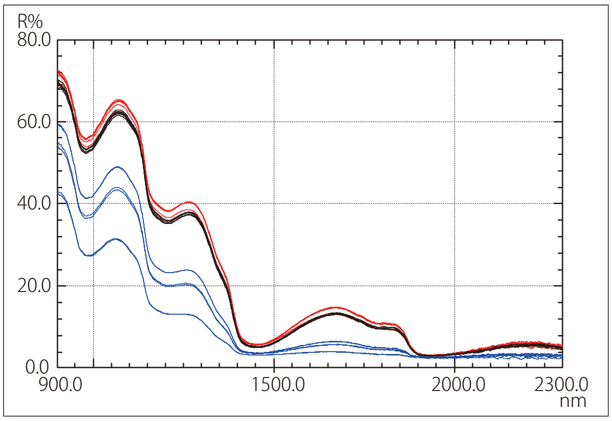

Fig. 2 Total Luminous Reflection Spectra of Samples
Red: High-Fat Milk, Black: Medium-Fat Milk, Blue: Low-Fat Milk

Table 1 Measurements of 9 Types of Milk Samples

Sample	Fat Content as Listed on Milk Carton (g/200 mL)
High-fat 1	9.3
High-fat 2	9.4
High-fat 3	9.5
Medium-fat 1	7.6
Medium-fat 2	7.8
Medium-fat 3	7.6
Low-fat 1	0.2
Low-fat 2	1.0
Low-fat 3	2.0

Table 2 Analytical Conditions

Instrument	: UV-3600 UV-visible-near-infrared spectrophotometer MPC-3100 Large sample compartment (with built-in integrating sphere)
Measurement wavelength range	: 900 nm – 2300 nm
Scan speed	: Medium
Sampling pitch	: 1.0 nm
Photometric value	: Reflectance
Slit width	: (12) nm

Results of Quantitative Analysis

The multiple linear regression method of multivariate analysis, the PLS method, and support vector regression method (SVR) were applied to the acquired data to quantify the fat content. The first and second samples of the high-fat, medium-fat, and low-fat samples of Table 1 were used as standard samples to generate a calibration model. The fat content values displayed on the milk cartons were taken as the true fat content values of the standard samples. As for the support vector regression method, this can be considered as a type of quantitative method of kernel multivariate analysis which can also be applied to non-linear data. In addition, the number 3 samples of the respective fat content samples of Table 1 were used as verification samples to confirm the prediction accuracy of the calibration models. The fat content prediction results for the verification samples using each of the calibration models are shown in Table 3. The values in parentheses represent the error of the predicted value with respect to the fat content displayed on the milk carton. By comparing the amount of error among the various techniques, it was determined that the best results were obtained using the support vector regression method.

It is thought that the relationship between the reflectance spectrum and concentration is not a linear relationship (proportional relationship). Compared to the multiple linear regression method and PLS method, which demonstrate their effectiveness with linear data, the support vector regression method may provide better results because it can also handle non-linear data effectively.

Calculations associated with the PLS method and the support vector regression method were conducted using The Unscramber® 1) multivariate analysis software of CAMO Software company. In addition, regression analysis calculations for the multiple linear regression method were conducted using the "Regression Analysis" feature the Microsoft Excel® 2) spreadsheet software.

Note:
Calculations using the multiple linear regression method were conducted with respect to four wavelengths, 1100 nm, 1200 nm, 1500 nm and 1800 nm. As for the PLS method, all the data from 1100 to 1500 nm were used and mean centering was conducted to conduct the calculations. As for the support vector regression method, all the data from 1100 nm to 1500 nm were used, and the parameters C = 1 and ε = 0.08 were calculated using a linear kernel function.

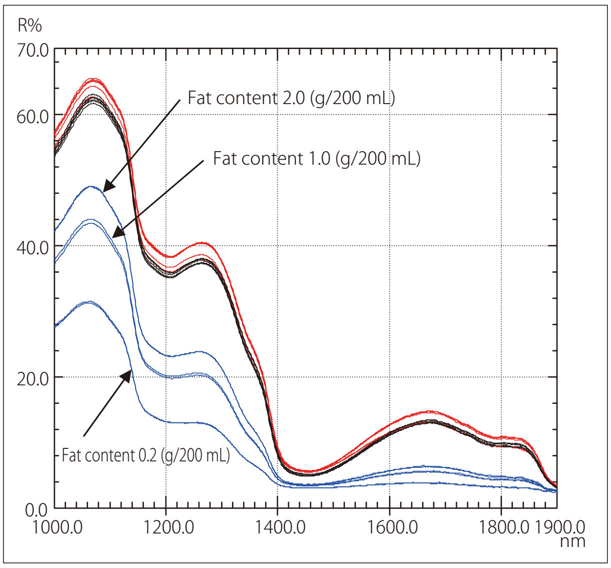

Fig. 3 Expanded Spectra of Fig. 2
Red: High-Fat Milk, Black: Medium-Fat Milk, Blue: Low-Fat Milk

Table 3 Concentration of Fat for Samples and Results Calculated by Multiple Linear Regression, PLS Regression and Support Vector Regression

Sample	Fat Content Listed on Milk Carton (g/200 mL)	Predicted Results by Multiple Linear Regression Method	Predicted Results by PLS Method	Predicted Results by Support Vector Regression Method
High fat 3 (1st)	9.5	8.87 (6.6 %)	9.57 (0.7%)	9.72 (2.3 %)
High fat 3 (2nd)	9.5	8.89 (6.4 %)	9.67 (1.6 %)	9.47 (0.3 %)
Medium fat 3 (1st)	7.6	8.04 (5.8 %)	8.40 (10.5 %)	7.88 (3.7 %)
Medium fat 3 (2nd)	7.6	7.59 (0.1 %)	7.71 (1.4 %)	7.31 (3.8 %)
Low fat 3 (1st)	2.0	2.24 (12.0 %)	1.78 (11.0 %)	2.01 (0.5 %)
Low fat 3 (2nd)	2.0	2.36 (18.0 %)	1.72 (14.0 %)	1.86 (7.0 %)

Conclusion

We conducted quantitation of fat content in various milk products by applying multivariate analysis to reflectance data.

The result of a comparison of three types of multivariate analysis including the multiple linear regression method, the PLS method, and the support vector regression method indicated that the support vector regression method offered the best quantitative accuracy. In addition, while spectral transmission measurement of liquid samples typically involves time-consuming washing of the cell after each measurement, the current method which combines the use of the spectral reflectance method with disposable screw-top glass tubes eliminates the need for troublesome cleaning, while also saving time.

The results obtained here suggest that this method is effective for the determination of fat content in highly turbid samples such as milk.

1) The Unscrambler is a trademark or registered trademark of CAMO Software.
2) Excel is a trademark or registered trademark of Microsoft Corporation.

First Edition: Nov. 2013

SHIMADZU
Shimadzu Corporation
www.shimadzu.com/an/