Determination of Chemical Composition of Commercial Honey by Near-Infrared Spectroscopy

P. Y. Qiu, H. B. Ding, Y. K. Tang, and R. J. Xu*

Department of Zoology, The University of Hong Kong, Pokfulam Road, Hong Kong

The feasibility of using near-infrared spectroscopy to determine chemical composition of commercial honey was examined. The influences of various sample presentation methods and regression models on the performance of calibration equations were also studied. Transmittance spectra with 1 mm optical path length produced the best calibration for all constituents examined. The regression model of modified partial least squares (mPLS) was selected for the calibration of all honey constituents except moisture, for which the optimal calibration was developed with PLS. Validation of the established calibration equations with independent samples showed that the spectroscopic technique could accurately determine the contents of moisture, fructose, glucose, sucrose, and maltose with squared correlation coefficients (R^2) of 1.0, 0.97, 0.91, 0.86, and 0.93 between the predicted values and the reference values. The prediction accuracy for free acid, lactone, and hydroxymethylfurfural (HMF) contents in honey was poor and unreliable. The study indicates that near-infrared spectroscopy can be used for rapid determination of major components in commercial honey.

Keywords: *Honey; NIR; fructose; glucose; sucrose; maltose*

INTRODUCTION

Honey, a natural sweet substance produced by honey bees, has a wide range of applications in the food industry. It can be processed for direct consumption or be used as an ingredient of various processed food products. Because of its superior nutritional value and unique flavor, natural bee honey is preferred by consumers. The price of bee honey is much greater than other sweetening commodities, such as refined cane sugar and corn syrup. Adulteration of bee honey with cheaper sweetening materials has been reported in the literature (Cienfuegos et al., 1997; Tien and Shau, 1997; Gonzalez et al., 1998). For this reason, various countries set strict standards for commercial honey, including specific physical properties and chemical compositions. To enforce such standards, chemical composition analysis and physical property assessment are routinely performed in commercial trading of honey. However, most techniques available for honey composition analysis and physical property determination are time-consuming and require considerable sample preparation and analytical skills.

Near-infrared (NIR) spectroscopy is a relatively new analytical technique, which is based on the electromagnetic absorption of organic compounds. The technique was initially developed in the 1960s by Norris and associates (Butler and Norris, 1960; Massie and Norris, 1965). The application of this technique was limited during the early stage of development due to the complexity of the NIR spectrum. Facilitated by more recent improvements in computing power and mathematical modeling techniques, NIR spectroscopy has now been applied to a wide range of industries, from food and feed composition analyses to petrochemical process control and noninvasive medical analysis. In the food and agricultural industries, NIR spectroscopy has been used to analyze chemical composition and physical properties of grains (Williams et al., 1984; Graybosch et al., 1995; Williams, 1996), meat (Ellekjaer et al., 1994; Cozzolino et al., 1996), and dairy products (Rodriguez-Otero et al., 1997) and to detect adulteration of various food ingredients (Downey, 1996). Apparently, in the literature there is no report of any application of the NIR spectroscopic technique to honey composition analysis and physical property assessment. In the present study, we evaluated the feasibility of using NIR spectroscopy to determine the chemical composition of commercial bee honey samples. The effects of different mathematical models, spectral scanning modes, sample presentation methods, and spectral data pretreatments on the performance of the NIR technique were also examined.

EXPERIMENTAL PROCEDURES

Samples. A total of 74 brands of bee honey produced in 11 countries was purchased from supermarkets in four different cities (Beijing, Hong Kong, Los Angeles, and Melbourne). All samples were stored in air-tight jars at 4 °C. Prior to chemical and spectroscopic analyses, honey samples were incubated in a water bath at about 50 °C until all the sugar crystals were melted. The samples were then mixed at room temperature before subsamples were taken for analyses.

Spectroscopic Analysis. Prior to spectroscopic analysis, all samples were left on the top of the laboratory bench to reach equilibrium room temperature. The spectroscopic analysis was then performed using a visible/near-infrared scanning spectrophotometer (model 6500, NIRSystem, Inc., Silver Spring, MD). Samples were scanned with the spectrophotometer in a quartz cuvette with 1, 2, 4, or 10 mm of optical path length. Reflectance and transmittance spectra between 400 and 2500 nm were recorded in 2 nm steps as log (1/R) and log (1/T) respectively, where R represents reflected energy and T

* Address correspondence to this author. Tel: 852-2859 7014. Fax: 852-2559 9114. E-mail: xuruojun@hkucc.hku.hk.

10.1021/jf9811368 CCC: $18.00
Published on Web 06/18/1999

Table 1. Chemical Composition of Honey Samples in Calibration and Validation Sets Determined by Reference Methods

	calibration set ($n = 50$)			validation set ($n = 24$)		
	mean	range	SD[a]	mean	range	SD[a]
moisture (%)	18.2	14.5−28.0	2.75	18.1	15.2−24.8	2.54
HMF (mg/kg)	85	2−731	161	76	0−543	141
fructose (%)	37.5	31.1−43.7	2.46	37.6	32.7−42.1	2.61
glucose (%)	30.6	26.2−33.8	1.70	30.8	28.4−32.6	0.98
sucrose (%)	1.62	0.00−12.60	2.08	1.23	0.13−3.44	0.82
maltose (%)	2.7	0.7−6.3	1.14	2.8	1.2−5.5	0.88
free acid (mequiv/kg)	20.3	9.1−51.0	8.10	22.9	9.4−149.4	27.50
lactone (mequiv/kg)	2.5	1.2−4.5	0.71	2.2	0.9−3.1	0.60

[a] Standard deviation.

represents transmitted energy. The honey samples were filled into the cuvette via a injection needle, and the cuvette was cleaned by rinsing under pressurized water beam and blown to dryness under a beam of compressed air.

Chemical Analysis. After spectroscopic analysis, the honey samples were analyzed for chemical compositions following the official methods of Association of Official Analytical Chemists (AOAC, 1990). Briefly, moisture content was measured using a refractometer (ATAGO 8436, Japan). The contents of free acid and lactone were determined by titrimetric method. Concentration of hydroxymethylfurfural (HMF) was determined by a spectrophotometric method. The contents of fructose, glucose, sucrose, and maltose were determined by the HPLC method with modified procedures described by Scott (1992). All analyses were performed in triplicate, and the mean values of the triplicate determination were used in calibration and validation studies.

Data Analysis. ISI software (ISI, 1992) was used for all data analysis. Fifty samples were randomly selected and formed a calibration set, and the remaining 24 samples formed a validation set. The reference values determined by chemical methods and the spectral data of both calibration and validation sets were analyzed to detect any outliers using a cutoff "T" value of 2.5 and a cutoff "H" value of 4.0 as described by Murray (1990). The number of outlying spectra and chemical values detected in the present study was between 2 and 4. Scatter correction of the spectral data was performed by standard normal variate transformation and de-trend method (Barnes et al., 1989). To magnify the fine structure and to improve resolution, the spectral data were further transformed with derivative operation (Stuart, 1996). Calibration equations were developed using modified partial least-squares (mPLS) regression analysis (Osborne et al., 1993), and the modification involved standardization of the residues after each iteration. Other calibration models, including principal component regression (PCR), partial least squares (PLS), and stepwise multiple linear regression (MLR) were also tested. The optimal calibration equations were chosen based on the highest squared correlation coefficient (R^2) and the lowest standard error of cross-validation (SECV). The performances of the established calibration equations were further validated using the samples in the validation set. The predicted values were correlated with the reference values, and the accuracy of prediction was assessed by the standard error of prediction (SEP), squared correlation coefficient (R^2), and bias.

RESULTS AND DISCUSSION

Results of chemical composition of honey samples in both calibration and validation sets are presented in Table 1. The average contents of various constituents were similar to those of American honey reported by White et al. (1962), with the exception of maltose content, which was much lower than the reported average value of 7.3% (White et al., 1962). This discrepancy may be due to differences in analytical methods. In the present study, maltose content in honey was determined by the HPLC method while in the earlier report maltose content in honey was calculated from reducing disaccharide content. The variation ranges and standard deviations of all constituents determined in the present study were generally greater than those of American honey reported by White et al. (1962), possibly representing diversity of samples used in the present study.

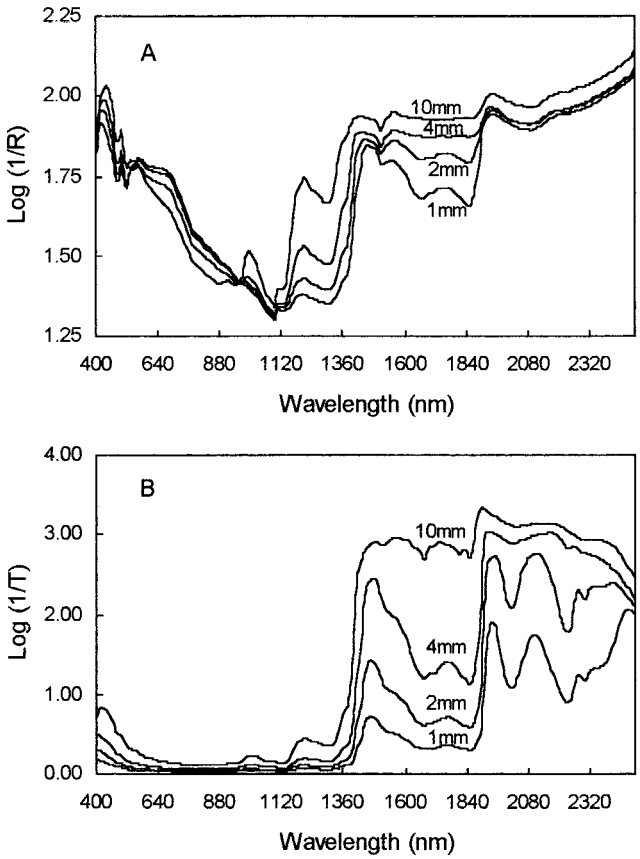

Figure 1. Average reflectance (panel A) and transmittance (panel B) spectra of honey samples presented to the spectrophotometer in a quartz cuvette with 1, 2, 4, or 10 mm optical path length.

Figure 1 shows the average reflectance and transmittance spectra of honey samples scanned with the spectrophotometer in a quartz cuvette with an optical path length of 1, 2, 4, or 10 mm. Compared with the reflectance spectra, spectra collected in transmittance mode had sharper peaks and better resolution. Calibration tests with mPLS regression model showed that transmittance spectra performed 30−70% better than reflectance spectra as assessed by SECV.

The optical path length of the cuvette had a marked impact on the transmittance spectra (Figure 1). Increase

Table 2. Statistics of the Optimal Calibration Equations Established with Transmittance Spectra with 1 mm Optical Path Length and Regression Model of Modified PLS

	spectral range	math treatment[a]	scatter correction[b]	SECV[c]	R^2	term[d]
moisture	1100−2500	1,5,5,1	no	0.08	1.0	7
HMF	400−2500	2,10,10,1	yes	60	0.88	8
fructose	400−2500	2,8,6,1	yes	0.57	0.94	8
glucose	400−2500	2,5,5,1	yes	0.52	0.90	4
sucrose	1100−2500	2,5,5,1	yes	0.28	0.91	8
maltose	400−2500	2,5,5,1	yes	0.31	0.92	6
free acid	1100−2500	1,10,10,1	yes	3.51	0.75	8
lactone	400−1100	2,5,5,1	yes	0.44	0.42	2

[a] Math treatment: the first figure represents the derivative order number (i.e., 0 indicates no derivative operation, 1 means first derivative, and so on); the second figure represents gap, the number of data points over which derivation was computed; the third figure represents the number of data points for spectral smoothing; and the fourth figure represents the data points for second spectral smoothing. [b] Scatter correction was performed by standard normal variate correction and de-trend. [c] SECV, standard error of cross-validation. [d] Term, the number of PLS factors used in the regression equations.

of optical path length from 2 to 10 mm led to a marked reduction of energy transmission, resulting in saturated spectra in the region of 1300−2500 nm. Calibration tests with modified PLS model showed that transmittance spectra of 1 mm optical path length produced the lowest SECV for all honey constituents examined when compared with the spectra of 2, 4, or 10 mm optical path length.

Calibration tests with different regression models showed that PLS performed better than MLR and PCR. Calibration with MLR uses selected wavelengths only while PCR captures as much of the variation in the whole spectral range as possible (Osborne et al., 1993). However, both MLR and PCR do not take account of reference values when selecting or constructing spectral components; whereas in the PLS regression model, information about reference values is involved in constructing spectral components (Osborne et al., 1993). This may explain why PLS performed better than MLR and PCR models. Further comparison showed that modified PLS performed better than PLS for all the honey constituents studied except moisture content. It has been stated in the ISI manual that modification involving standardization of the residues after each iteration can improve stability and accuracy of the PLS calibration.

The optimal calibration equations for determination of moisture, HMF, fructose, glucose, sucrose, maltose, free acid, and lactone contents in honey samples were constructed based on the lowest SECV and the highest R^2. The statistical parameters of the optimal calibration equations are presented in Table 2. The transmitance spectra with 1 mm optical path length were selected for all calibration equations. Modified PLS regression model was selected for all honey constituents examined except moisture content, of which optimal calibration was developed with PLS regression model. Scatter correction of the spectra generally improved the calibration performance for most of the constituents (Table 2).

Validation of the established calibration equations with independent samples showed that the spectroscopic technique could accurately predict the contents of major components in honey, such as moisture, fructose, and glucose (Table 3). Correlation between the predicted

Table 3. Means, Bias, Standard Error of Prediction (SEP), and Squared Correlation Coefficients (R^2) of Chemical Composition of Honey Samples in the Validation Set Predicted by Near-Infrared Spectroscopy

	mean	bias	SEP	R^2
moisture	18.1	0.0	0.16	1.00
HMF	34	44	110	0.66
fructose	37.5	0.1	0.42	0.97
glucose	30.7	0.1	0.34	0.91
sucrose	1.3	−0.1	0.34	0.86
maltose	2.9	−0.1	0.28	0.93
free acid	17.5	−0.1	4.39	0.49
lactone	2.3	−0.1	0.79	0.00

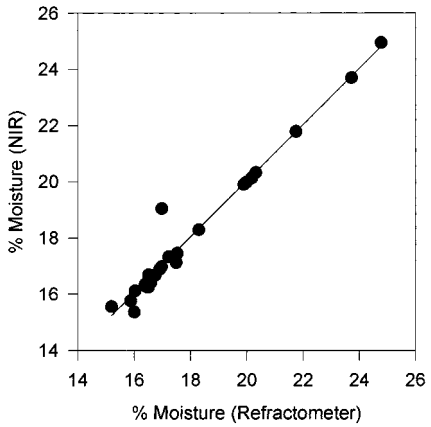

Figure 2. Scatter plot of moisture contents of the validation samples determined by refractometer and NIR spectroscopy, respectively.

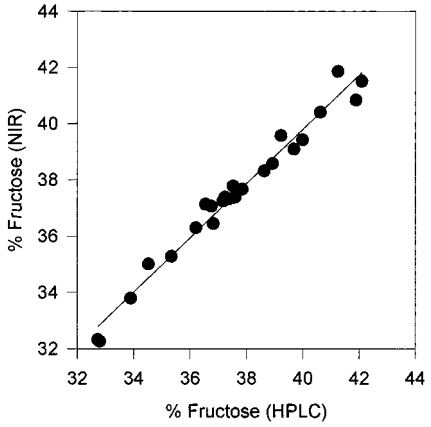

Figure 3. Scatter plot of fructose contents of the validation samples determined by HPLC and NIR spectroscopy, respectively.

values and the reference values produced squared correlation coefficients of 1.00, 0.97, and 0.91 for moisture, fructose, and glucose contents, respectively. Typical scatter plots for moisture and fructose contents are shown in Figures 2 and 3. The prediction performances of the spectroscopic technique for minor components of sucrose and maltose were also acceptable. The squared correlation coefficients between predicted values and the reference values for sucrose and maltose were 0.86 and 0.93, respectively. However, the prediction accuracy for trace components of HMF, free acid, and lactone was rather poor and unreliable.

Although fructose, glucose, sucrose, and maltose have similar structures and all contain C−H and O−H bonds,

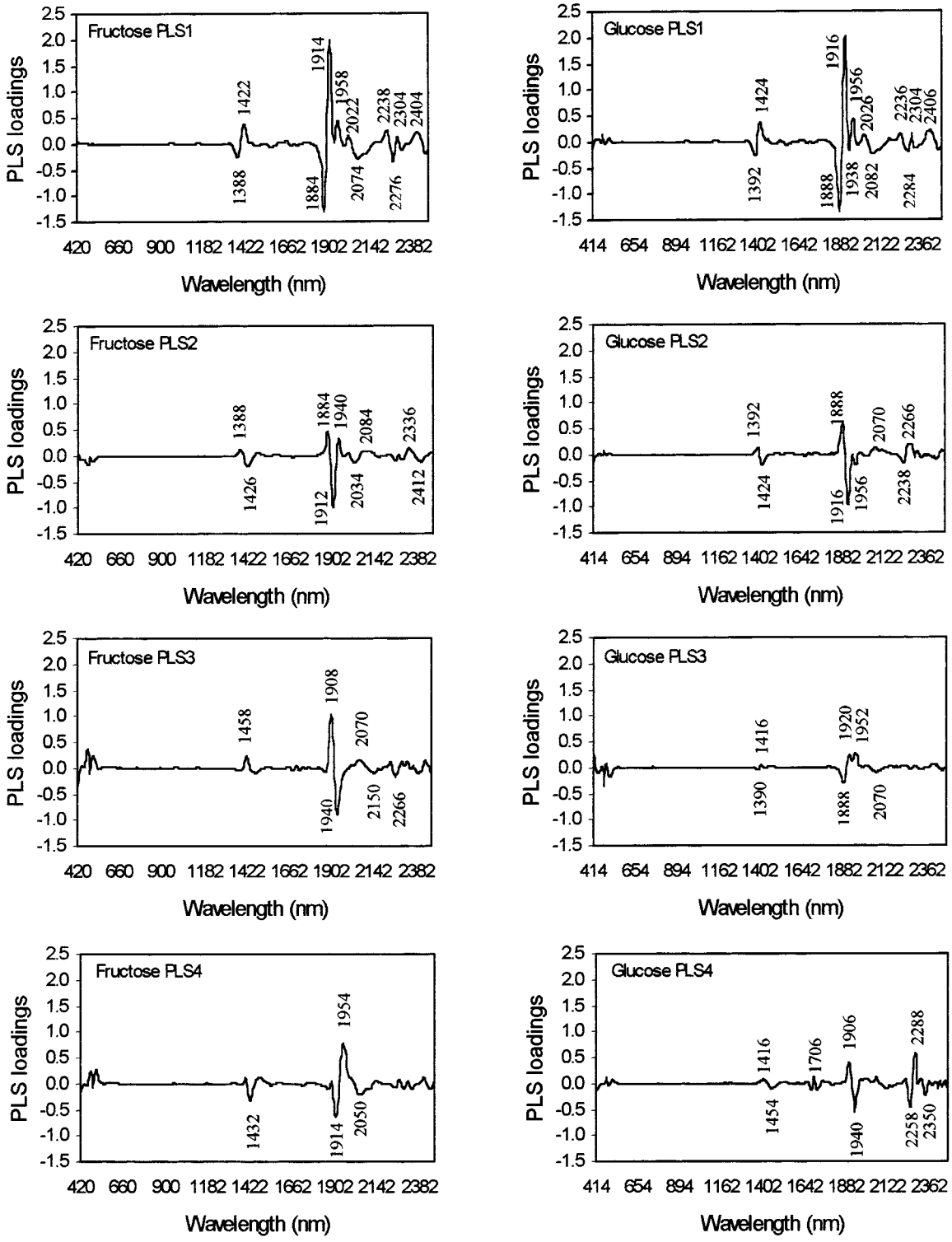

Figure 4. PLS loadings of the first four factors used in the regression calibrations of fructose and glucose contents in honey.

subtle differences exist in their NIR absorption spectra (Osborne et al., 1993). In the present study, major PLS loadings used in the calibration equations for individual sugars reflected the subtle differences (Figures 4 and 5), indicating the specificity of the calibration equations for individual sugars. The squared correlation coefficients between total sugar content and the contents of fructose, glucose, sucrose, and maltose determined by the reference method were 0.50, 0.03, 0.28, and 0.18, respectively. The squared correlation coefficients between the individual sugars determined by the reference method were below 0.3. In contrast, the squared correlation coefficients between the content of individual sugars determined by the reference method and the contents predicted by the NIR calibration equations were above 0.85 (Table 3), further indicating the specificity of the NIR calibration equations for individual sugars. Previous studies with mixtures of fructose, glucose, and sucrose in dry powders or in aqueous solutions have also shown that specific NIR calibration equations can be established for individual sugars (Giangiacomo et al., 1981; Giangiacomo and Dull, 1986).

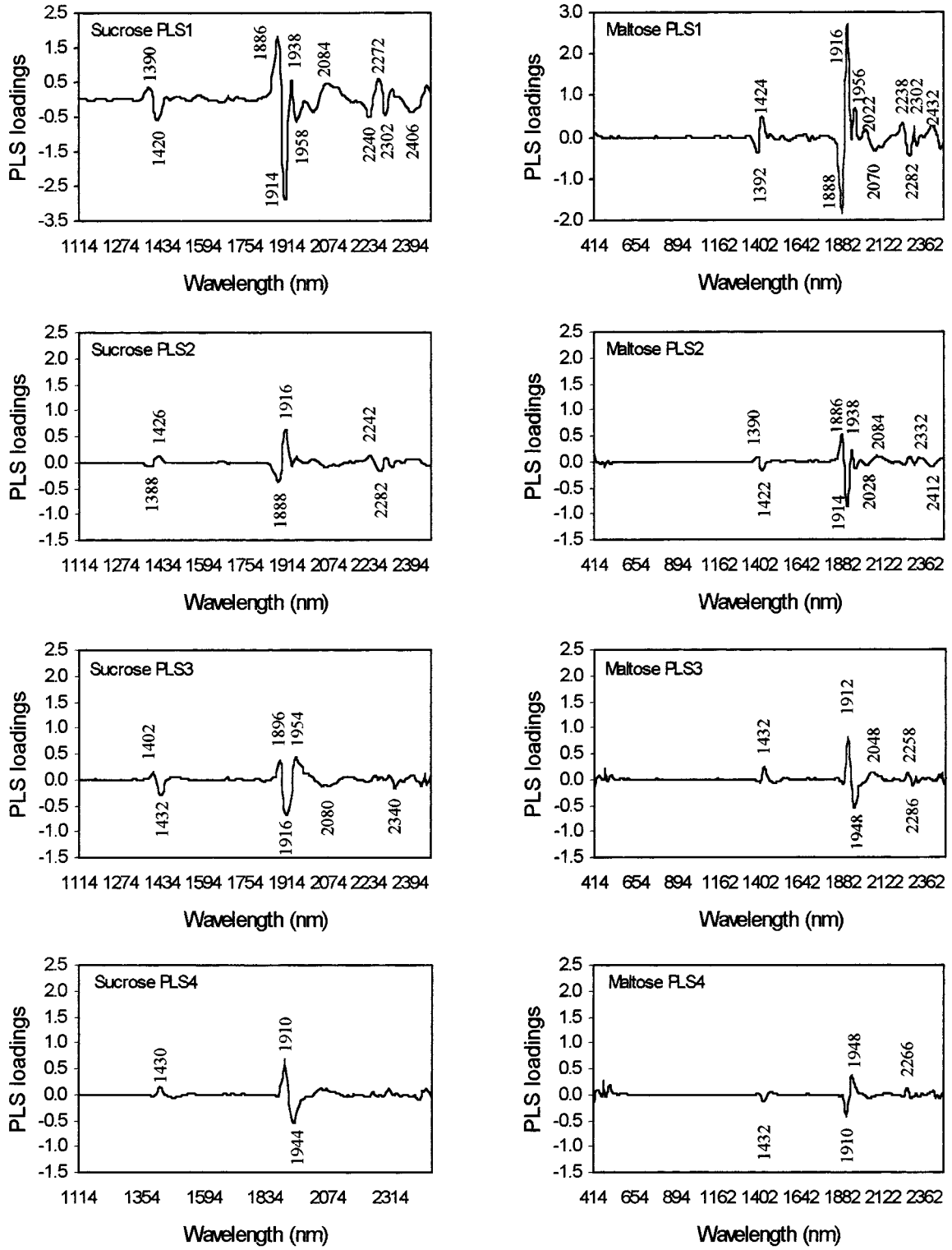

Figure 5. PLS loadings of the first four factors used in the regression calibrations of sucrose and maltose contents in honey.

In conclusion, the present study demonstrated that the NIR spectroscopy is feasible for rapid analysis of major components of commercial honey.

ABBREVIATIONS USED

AOAC, Association of Official Analytical Chemists; HMF, hydroxymethylfurfural; HPLC, high-pressure liquid chromatography; ISI, Infrasoft International; MLR, multiple linear regression; mPLS, modified partial least squares; NIR, near-infrared; PCR, principal component regression; PLS, partial least squares; R^2, squared correlation coefficient; SD, standard deviation; SECV, standard error of cross-validation; SEP, standard error of prediction.

LITERATURE CITED

AOAC. *Official Methods of Analysis*, 15th ed.; Association of Official Analytical Chemists: Washington DC, 1990.

Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard normal variance transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl. Spectrosc.* **1989**, *43*, 772−777.

Butler, W. L.; Norris, K. H. The spectrophotometry of dense light-scatting material. *Arch. Biochem. Biophys.* **1960**, *87*, 31−40.

Cienfuegos, E.; Casar, I.; Morales, P. Carbon isotopic composition of Mexican honey. *J. Apic. Res.* **1997**, *36*, 169−179.

Cozzolino, D.; Murray, I.; Paterson, R.; Scaife, J. R. Visible and near infrared reflectance spectroscopy for the determination of moisture, fat and protein in chicken breast and thigh muscle. *J. Near Infrared Spectrosc.* **1996**, *4*, 213−223.

Downey, G. Authentication of food and food ingredients by near infrared spectroscopy. *J. Near Infrared Spectrosc.* **1996**, *4*, 47−61.

Ellekjaer, M. R.; Isaksson, T.; Solheim, R. Assessment of sensory quality of meat sausages using near infrared spectroscopy. *J. Food Sci.* **1994**, *59*, 456−464.

Giangiacomo, R.; Dull, G. G. Near infrared spectrophotometric determination of individual sugars in aqueous mixtures. *J. Food Sci.* **1986**, *51*, 679−683.

Giangiacomo, R.; Magee, J. B.; Birth, G. S.; Dull, G. G. Predicting concentrations of individual sugars in dry mixtures by near-infrared reflectance spectroscopy. *J. Food Sci.* **1981**, *46*, 531−534.

Gonzalez, M. I.; Marques, M. E.; Sanchez, S. J.; Gonzalez, R. B. Detection of honey adulteration with beet sugar using stable isotope methodology. *Food Chem.* **1998**, *61*, 281−286.

Graybosch, R. A.; Peterson, C. J.; Baenziger, P. S.; Shelton, D. R. Environmental modification of hard red winter wheat flour protein composition. *J. Cereal Sci.* **1995**, *22*, 45−51.

ISI. *NIRS 2: Software for Near Infrared Instruments*; Infrasoft International: 1992.

Massie, D. R.; Norris, K. H. Spectral reflectance and transmitance properties of grains in the visible and near-infrared. *Trans. Am. Soc. Agric. Eng.* **1965**, *8*, 598−600.

Murray, I. Aspects of the interpretation of NIR spectra. In *Analytical Application of Spectroscopy;* Creaser, C. S., Davies, A. M. C., Eds.; Royal Society of Chemistry: Cambridge, U.K., 1990.

Osborne, B. G.; Fearn, T.; Hindle, P. H. *Practical NIR Spectroscopy with Application in Food and Beverage Analysis;* Longman Scientific & Technical: Harlow, U.K., 1993.

Rodriguez-Otero, J. L.; Hermida, M.; Centeno, J. Analysis of dairy products by near-infrared spectroscopy: a review. *J. Agric. Food Chem.* **1997**, *45*, 2815−2819.

Scott, F. W. HPLC detrmination of carbohydrates in foods. In *Food Analysis by HPLC;* Nollet, L. M. L., Ed.; Marcel Dekker: New York, 1992; pp 259−274.

Stuart, B. *Modern Infrared Spectroscopy;* John Wiley & Sons: New York, 1996.

Tien, L. L.; Shau, M. O. A. Quality analysis of Longan honey in Taiwan market. *Food Sci. Taiwan* **1997**, *24*, 479−489.

White, J. W.; Riethof, M. L.; Subers, M. J.; Kushnir, I. *Composition of American honeys*; Technical Bulletin 1261; U.S. Department of Agriculture: Washington, DC, 1962.

Williams, P. C. Observations on the use, in prediction of functionality in cereals, of weights derived during development of partial least squares regression. *J. Near Infrared Spectrosc.* **1996**, *4*, 175−187.

Williams, P. C.; Preston, K. R.; Norris, K. H.; Starkey, P. M. Determination of amino acids in wheat and barley by near-infrared reflectance spectroscopy. *J. Food Sci.* **1984,** *49*, 17−20.

Received for review October 14, 1998. Revised manuscript received March 29, 1999. Accepted April 28, 1999. This study was partially supported by a research grant from the Committee of Research and Conference Grants of The University of Hong Kong.

JF9811368