Selection of wavelength region for partial least squares Brix calibration of mango using the multiple linear regression method

Sirinnapa Saranwong,[a] Jinda Sornsrivichai[a] and Sumio Kawano[b]

[a]*Department of Biology, Faculty of Science, Chiang Mai University, Chiangmai 50002, Thailand*

[b]*National Food Research Institute, 2-1-12 Kannondai, Tsukuba 305-8642, Japan*

Introduction

In order to establish the calibration equation of near infrared (NIR) spectroscopy, multiple linear regression (MLR) was used during the early period[1–11] while partial least squares (PLS) regression has been widely applied recently because no selection of wavelength is needed.[9–13] However, even in a PLS calibration, the selection of a wavelength region is needed to make a good calibration, which is more complicated and time-consuming.[14,15] Therefore, in this work, the performance of MLR and PLS calibration equations were compared and the relationship between them was also investigated by using the Brix value of mango as a model.

Materials and methods

Materials

A total of 138 Philippine mango fruits (*Mangifera indica* cv. Caraboa), which were commercially available, were used as samples for this experiment. The fruits packed with corrugated fibreboard boxes were purchased at the wholesale market, transported to our laboratory at the National Food Research Institute (NFRI) and then kept in the cold storage room at a temperature of 5°C for the experiment the following day. Six hours before NIR spectra acquisition, the fruits were moved from the cold storage room to a room with a temperature of 25°C so that the sample temperature reached room temperature.

Spectral acquisition

NIR spectra of mango fruits were measured in the short wavelength region from 700 nm to 1100 nm. A commercially available NIR instrument (NIRSystems 6500) with a fibre-optic "Interactance Probe" was used to measure the NIR spectra of intact mango fruits in a similar way to our previous study.[4] The NIR spectra were obtained at the fruit shoulder by averaging 50 scans. A standard measurement of a teflon sphere (8 cm diameter) was performed on every six fruits.

Prior to NIR measurement of samples, control of sample temperature was performed by dipping each sample into a water bath controlled at 25°C. In order to prevent samples from getting wet, the surface of the water bath was covered with a polyethylene film and then the samples were dipped into the water, together with the film, for ten minutes.

Near Infrared Spectroscopy: Proceedings of the 10th International Conference
© IM Publications Open LLP 2002

Table 1. Characteristics of calibration and validation sample sets of mango used.

Item	Calibration set	Validation set
Number of sample	74	64
Range	10.80 – 18.25	11.15 – 16.65
Mean	14.44	14.22
Standard deviation	1.72	1.38
LE[a]	0.23	0.25
Unit	°Brix	°Brix

[a]LE is laboratory error calculated from standard deviation of error between duplicate measurements of each reference analysis

Chemical analysis

A portion of flesh, about 10 mm deep from the peel, which was illuminated by NIR, was sampled and analysed for Brix value with a digital refractometer (ATAGO, Model PR101).

Each constituent value used for the following calculation was given by averaging of duplicate measurements. Statistical characteristics of calibration and validation sets in both wavelength regions are shown in Table 1.

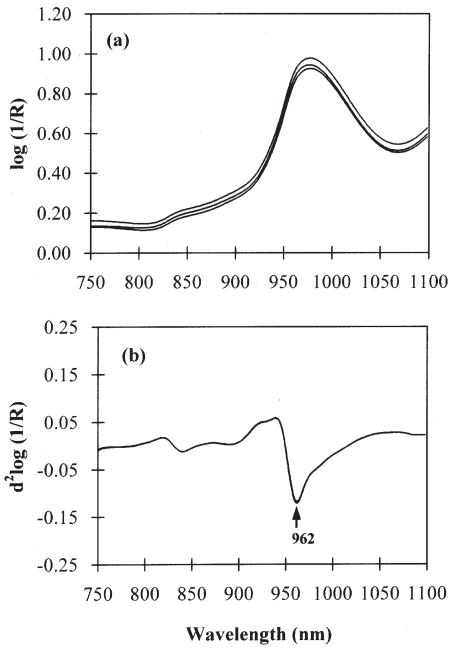

Figure 1. (a) Original spectra and (b) second derivative spectra of typical mango fruits having low, medium and high Brix values.

Data analysis

The Near Infrared Spectral Analysis Software (NSAS) program and the Unscrambler program were used for MLR and PLS regression, respectively.

Results and discussion

NIR spectra of intact mango

The original spectra and second derivative spectra of typical intact mangoes having low, medium and high Brix values are shown in Figure 1(a) and 1(b), respectively. Strong absorption bands due to water were observed at the wavelength of 962 nm.

Calibration and validation by MLR

The correlation plot which were correlation coefficients between $d^2\log(1/R)$ and Brix value plotted against wavelength is shown in Figure 2. Large negative peaks were observed at 848, 906 and 990 nm. The wavelength of 906 nm was selected manually as the first wavelength in a similar way to our previous study.[4] The calibration

© IM Publications Open LLP 2002

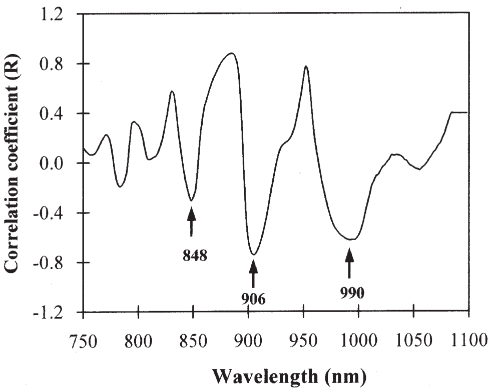

Figure 2. Correlation plots for selecting the first calibration wavelength of Brix value.

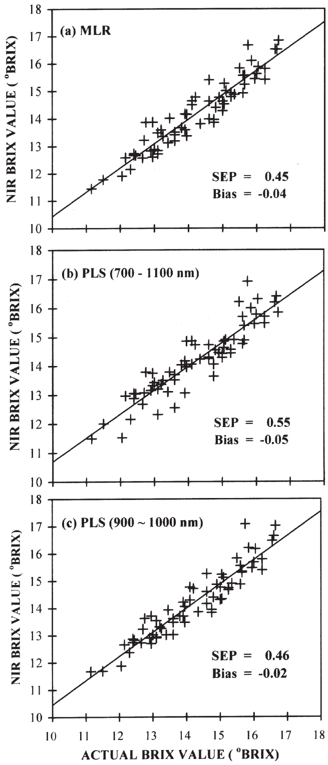

Figure 3. Scatter plot of Brix value of mango fruits. The Brix values are predicted by (a) MLR, (b) PLS (700–1100nm), and (c) PLS (900–1000 nm).

and validation results are shown in Table 2. The scatter plot of Brix values of mango fruits predicted by MLR developed calibration equation are shown in Figure 3(a).

Calibration and validation by PLS

The PLS calibration and validation results using different wavelength regions are shown in Table 3. The *SEP* was smallest when the wavelength region from 900 nm to 1000 nm was se-

Table 2. Calibration and validation results obtained from MLR regression regions for determining Brix values of mango fruits.

Selected Wavelengths (nm)				*R*	*SEC*	*SEP*	Bias
λ_1	λ_2	λ_3	λ_4		(°Brix)	(°Brix)	(°Brix)
906^M	—	—	—	–0.74	1.17	1.20	0.27
906^M	996^M	—	—	0.91	0.73	0.81	0.13
906^M	996^M	978^M	—	0.91	0.73	0.77	0.15
906^M	996^M	978^M	942^C	0.96	0.49	0.45	–0.04

^{M} Selected manually
^{C}: Selected by computer
R: Multiple correlation coefficients
SEC: Standard error of calibration,
SEP: Bias-corrected standard error of prediction
Bias: The average of difference between actual value and NIR value

© IM Publications Open LLP 2002

Table 3. Calibration and validation results of PLS calibration on various wavelength regions for determining Brix values of mango fruits.

Wavelength region (nm)	Factors	*R*	*SEC* (°Brix)	*SEP* (°Brix)	Bias (°Brix)
700 – 1100	9	0.97	0.43	0.55	–0.05
800 – 1100	4	0.95	0.52	0.53	–0.02
900 – 1100	5	0.96	0.48	0.49	0.01
1000 – 1100	3	0.92	0.68	0.77	0.23
900 – 1000	5	0.96	0.47	0.46	–0.02
900 – 950	3	0.93	0.62	0.72	0.08

R: multiple correlation coefficients
SEC: standard error of calibration
SEP: bias-corrected standard error of prediction
Bias: the average of difference between actual value and NIR value

lected. The scatter plots of Brix values of mango fruits predicted by PLS developed calibration equation are shown in Figure 3.

Relationship between MLR and PLS calibrations

It was found that the best wavelength region for PLS calibration from 900 nm to 1000 nm was identical to the wavelength region which involved the wavelengths selected by MLR.

The regression coefficient (*K*) plot of the best PLS calibration is shown in Figure 4. Four negative peaks were observed at 908, 942, 978 and 1000 nm, which corresponded to the wavelengths selected for the MLR calibration equation as shown in Table 2. The MLR calibration equation can be written as follows:

$$\text{Brix value} = 13.61 - 1245.02\, d^2\log(1/R_{906}) - 329.01 d^2\log(1/R_{996}) - 1052.89 d^2\log(1/R_{978}) - 824.89\, d^2\log(1/R_{942}) \quad (1)$$

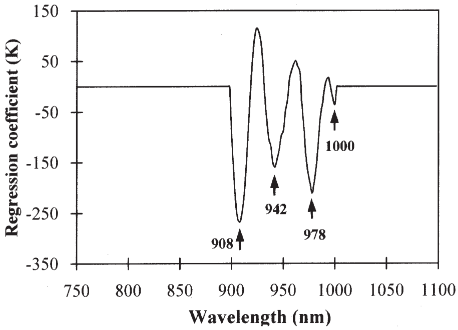

Figure 4. Regression coefficient plot for PLS calibrations using the wavelength region from 900 nm to 1000 nm.

In addition, it was found that there is also high similarity in the order of the regression coefficients. The order for MLR and PLS calibration equations is:

$K_{906} < K_{978} < K_{942} < K_{996}$ for MLR
and $K_{908} < K_{978} < K_{942} < K_{1000}$ for PLS

where, $K_{\lambda i}$ is the regression coefficient at the wavelength λ_i.

These results indicated that even if the procedures for getting the calibration equation are different, the MLR and PLS calibration equations to predict Brix value were based on the same important wavelengths.

© IM Publications Open LLP 2002

Conclusion

It was concluded that NIR spectroscopy in the short wavelength region is sufficiently accurate to determine Brix value of mango fruit nondestructively. Both MLR and PLS could be used in making calibration equations with similar predictive efficiency. It was found that the selection of wavelength region is needed in PLS calibration in the same way as in MLR calibration. The wavelength region selected by MLR could be used as a good navigator for the selection of the wavelength region in the PLS calibration.

Acknowledgement

This research was supported by the Royal Golden Jubilee PhD Programme under the Thailand Research Fund (TRF).

References

1. G.S. Birth, G.G. Dull, W.T. Renfroe and S.J. Kays, *J. Amer. Soc. Hort. Sci.* **110,** 297 (1985).
2. G.G. Dull, G.S. Birth, D.A. Smittle and R.G. Leffler, *J. Food Sci.* **54,** 393 (1989).
3. S. Kawano, H. Watanabe and M. Iwamoto, in *The Proceedings of the 2nd International Near Infrared Spectroscopy Conference,* Ed by M. Iwamoto and S. Kawano. Korin Publishing, Japan, p. 343 (1990).
4. S. Kawano, H. Watanabe and M. Iwamoto, *J. Japan. Soc. Hort. Sci.* **61(2),** 445 (1992).
5. S. Kawano, H. Abe and M. Iwamoto, *J. Near Infrared Spectrosc.* **3,** 211 (1995).
6. S. Kawano, T. Fujiwara and M. Iwamoto, *J. Japan. Soc. Hort. Sci.* **62(2),** 465 (1993).
7. T. Temma, M. Chiba, K. Hanamatsu and F. Shinoki, in *Near Infrared Spectroscopy: The Future Waves*, Ed by A.M.C. Davies and P. Williams. NIR Publications, Chichester, UK, p. 638 (1996).
8. R.K. Cho, M.R. Sohn and Y.K. Kwon, *J. Near Infrared Spectrosc.* **6,** A75 (1998).
9. J. Guthrie and K. Walsh, *Aust. J. Exp. Agr.* **37,** 253 (1997).
10. Z. Schmilovitch, A. Mizrach, A. Hoffman, H. Egozi and Y. Fuchs, *Postharv. Biol. Technol.* **19,** 245 (2000).
11. P. Dardenne, G. Sinnaeve and V. Baeten, *J. Near Infrared Spectrosc.* **8,** 229 (2000).
12. V.A. McGlone and S. Kawano, *Postharv. Biol. Technol.* **13,** 131 (1998).
13. K. Miyamoto, M. Kawauchi and T. Fukuda, *J. Near Infrared Spectrosc.* **6,** 267 (1998).
14. S.D. Osborne, R. Kunnemeyer and R.B. Jordan, *J. Near Infrared Spectrosc.* **7,** 9 (1999).
15. P. Carlini, R. Massantini and F. Mencarelli, *J. Agric. Food Chem.* **48,** 5236 (2000).

© IM Publications Open LLP 2002