J Food Sci Technol (March–April 2010) 47(2):207–213
DOI: 10.1007/s13197-010-0033-1

ORIGINAL ARTICLE

Non-destructive prediction of quality of intact apple using near infrared spectroscopy

Jha S. N. · Garg Ruchi

Revised: 24 August 2009 / Accepted: 27 August 2009

Abstract Potential of near infrared spectroscopy (NIRS) in the wavelength range of 900–1700 nm for determination of sweetness (total soluble solids, TSS); sourness (acidity) and their ratio for 5 cultivars of apple was studied. Partial least square and multiple linear regression (MLR) employing pre-processing techniques were carried out. MLR models were found to be the best for prediction after treating the spectral data with multiple scatter correction technique. The multiple correlation coefficients for calibration and validation were found to be 0.887, 0.745 °Brix for TSS, 0.890, 0.752 % for acidity and 0.893, 0.751 for acidity/TSS ratio, respectively. The standard errors of calibration, prediction, biases and differences in them were low, which indicated that NIRS has potential to predict internal quality of apple non-destructively.

Keywords Apple · Sourness · Sweetness · Acidity/TSS ratio · NIR spectroscopy

Jha S. N. · Garg R.
Agricultural Structures and Environmental Control Division, Central Institute of Post-Harvest Engineering and Technology, Ludhiana - 141 004, India

Jha S. N. (✉)
E-mail: snjha_ciphet@yahoo.co.in

Introduction

Quality evaluation of agricultural products has been a subject of interest for researchers. The choice of the best technique to measure fruit quality is difficult, and even more complicated is the definition of internal fruit quality, which changes during pre- and post-harvest operations. The sugars and organic acids content as also the texture and flavour represent fruits' internal quality. Traditionally, taste of a fruit in terms of sweetness or sourness is determined by destructive laboratory methods, which render the tested fruit unusable. Currently, non-destructive techniques for quality evaluation have gained momentum (Iwamoto et al. 1995, Jha and Matsuoka 2004). These techniques, particularly for fruits and vegetables, are rapid and easy to use. Many physical characteristics of fruits and vegetables have been determined non-destructively using various techniques (Lesage and Destain 1996, Nussinovitch et al. 1996, Kato 1997, Jha et al. 2005, 2006, 2007). Near infrared spectroscopy (NIRS) is an important technique amongst them.

Physiological indices and firmness of apple have been determined using NIRS (Schmilovitch et al. 2000). Aneshansley et al. (1997) identified wavelengths of 540, 750, 970 and 1030 nm are useful to distinguish between sound and damaged tissues in several apple cultivars. A NIR system, which combined a charged coupled devices (CCD) spectrophotometric camera and bifurcated fiber optics for determining total soluble solids (TSS) in apples, was reported by Bellon and Sevila (1993).

Most of the reported literature on determination of quality of apple using NIRS has used either TSS or acidity separately for a single variety of apple (Lammertyn et al. 1998, Ventura et al. 1998, Lu and Ariana 2002, Saranwong et al. 2004). Predictions of TSS in majority of cases have not been found so satisfactory. The taste of apple however, may also be determined using either acidity or acidity/TSS ratio as has been done for tomato juice (Jha and Matsuoka, 2004). Thus, the objective of this study was to investigate the potential of NIRS for prediction of TSS, titratable acidity and acidity/TSS ratio for commonly used 5 cultivars of apple to minimize the effect of variety during prediction.

Materials and methods

Freshly harvested apples (*Malus domestica* Borkh) variety 'Golden Delicious', 'Red Delicious', 'Ambri' and 2 unknown varieties were procured from local fruit market in the year 2006. The samples were brought to laboratory and screened manually to discard the damaged ones. Sound apples were wiped with muslin cloth to remove dirt and kept for 28 days at 32±0.5°C and 65±7% RH for accelerated changes in biochemical quality parameters. Three samples from each variety were randomly chosen bi-weekly for experimentation. Altogether 118 apples were used in the experiment.

NIR Spectra acquisition: Transmittance spectra in wavelength range of 900–1700 nm of 118 apples, were acquired using a portable NIR-spectrometer (model EPP 2000-InGaAs, 2.25 nm resolution StellarNet Inc., USA) connected to 30 W halogen lamp and sample holder with 400 micron optical fibre cable and spectra wiz software (version 3.3). Dark and reference spectra for a standard supplied with the equipment were taken for 50 scan in 100 ms integration time. Apple fruit was then placed on sample holder arbitrarily from girth side in stable position by hand. The probe (sensor) was fixed in the centre of the base of sample holder and transmittance spectra were acquired at an interval of 2.25 nm for the wavelength range of 900–1700 nm.

Determination of TSS and acidity: Immediately after recording the spectra, juice of the whole apple was extracted using domestic juicer at ambient room temperature (28–30°C). Juice was filtered through a new piece of muslin cloth every time. The TSS (Jha and Matsuoka 2004) of the filtered juice was measured thrice using a hand held digital refractometer (Pal-1, Atago, range 0–53°Brix, least count 0.2°Brix, Japan) and acidity was determined using standard titration method (AOAC 1990). Mean values were then used for NIR calibration and validation.

NIR spectral data were imported to MS-excel software from spectra wiz software, and then to Unscrambler (CAMO AS, Trondheim, Norway, version 8.0.5) software for multivariate analyses with TSS, acidity and acidity/TSS ratio (Jha and Matsuoka 2004). Spectral data were plotted to inspect the nature of the spectra. Four curves visually odd amongst the group were identified as outlier and were deleted. Altogether 114 samples were used for calibration and validation for prediction of TSS, acidity and acidity/TSS ratio. The ranges, means and standard deviations of these parameters are presented in Table 1.

Two methods of regression, partial least square regression (PLS) and multiple linear regression (MLR) with the option of full cross validation available in the software were performed on the whole original spectra to develop the NIR model for predicting TSS, acidity and acidity/TSS ratio of apples non-destructively. In order to search for a small and simple model, the whole range of spectra was divided into small groups of 35 wavelengths continuously at an interval of 2.25 nm and both PLS and MLR were performed on each group. The best models based on the standard error of calibration (SEC), multiple correlations coefficients (R) and standard error of prediction (SEP) were selected (Jha et al. 2005, 2006).

In order to improve the predictability of selected models, pre-processing of spectra such as smoothing (second order), full multiplicative scatter correction (MSC) and second order derivative (Savitzky-Golay, by averaging one point to the left and one point to the right and fitting a second order polynomial) of selected range of wavelengths were performed. A few outlier samples were also identified with the help of software and removed for further improvements in the models. To minimize the number of wavelengths further, the effect of individual wavelength by eliminating them from the model of best performing group of wavelengths on the root mean square error was investigated (Kawano et al. 1995). Scatter plots between measured and predicted parameters were plotted to know the actual predictability using NIRS non-destructively.

Results and discussion

Typical spectral curves in the wavelength of 900–1700 nm for all five varieties (2 curves for each variety of apples) did not show any varietal differences in peaks and depressions (Fig. 1). The peaks and depressions in spectra show

Table 1 Statistical details of samples used in calibration and validation of different models

	Mean
TSS, °Brix	12.98±1.98
Acidity, %	0.104±0.041
Acidity/TSS ratio	0.008±0.003

Nr. of samples: 114, TSS: Total soluble solids

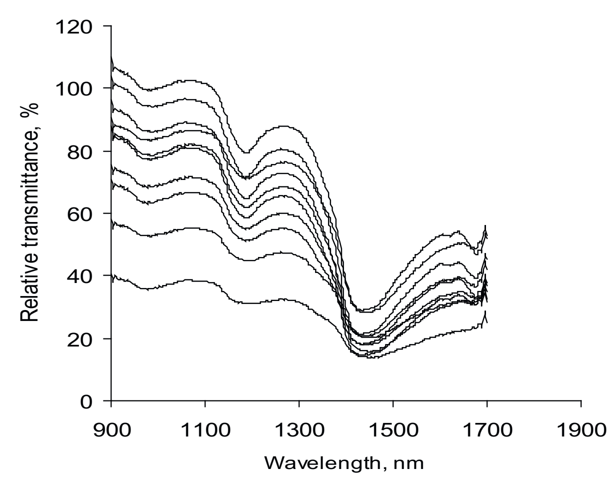

Fig. 1 Typical NIR spectra of different varieties of apple in wavelength range of 900–1700 nm

the strong and the weak transmittance characteristics of the apples, respectively within the range of the study. The relative values in other regions of spectra, however, differed from sample to sample.

R values for calibration and validation of TSS in PLS regression were found to be 0.562 and 0.454, respectively for the wavelength range of 1136.25–1212.75 nm, whereas these values were 0.749 and 0.457 in case of MLR (Table 2) for the same range of wavelengths, which indicated that MLR is better for prediction of TSS. There is however a large difference in R values of calibration and validation and thus results may be unstable during prediction. Ventura et al. (1998) obtained higher R values for prediction of TSS as compared to this study. Lower values here are mainly due to 5 cultivars of apples which are distinctly different in characteristics were used in this study. The present model is thus applicable to wide varieties of apple whereas it was only for one variety (cv. Jonagold) in case of Lammertyn et al. (1998).

Effect of various pre-processing techniques for enhancement of predictability of the selected model showed that the MSC treatment yielded highest R values for calibration and validation, 0.889 and 0.745, respectively (Table 3), which shows that after treatment, TSS can be predicted with reasonable accuracy. If no treatment is given even then we can predict the TSS but stability in prediction is at stake because of large difference in R values of calibration and prediction. Similar views have also been reported by Saranwong et al. (2004). Scatter plots of models after MSC treatments for the wavelength range of 1136.25–1212.75 nm are shown in Fig. 2, which indicate that slope of the curve is near to 45° and thus prediction is near to measured values.

Acidity can be predicted in the wavelength range of 900–976.5 nm using PLS as well as MLR. R values in case of PLS were found to be 0.736 for calibration and 0.66 for validation (Table 4). So, prediction is stable but may not be so accurate as in MLR, where R values for calibration and validation were found to be 0.853 and 0.481, respectively (Table 4). In this case large gap indicates that prediction may not be as stable as PLS. The predictability of selected MLR model was further enhanced after applying MSC and removing some outlier samples. The lowest SEC, SEP and highest R values for calibration and validation of MSC treated spectra were found to be 0.016, 0.890, –0.001 and 0.024, 0.752, –0.001, respectively (Table 3). Negative biases indicate that predicted values may be lower than the

Table 2 SEC, SEP, R and biases of PLS and MLR models for different range of NIR wavelengths for prediction of TSS

Wavelength range, nm	Calibration			Validation		
	SEC	R	Bias	SEP	R	Bias
PLS model						
900–976.5	1.716	0.499	–0.000	1.905	0.328	–0.04
978.75–1055.25	1.980	0.0133	0.000	2.026	–0.551	0.000
1057.5–1134	1.980	0.018	0.000	2.013	–0.411	–0.004
1136.25–1212.75	1.638	0.562	–0.000	1.783	0.454	–0.004
1215–1291.5	1.859	0.344	–0.000	1.954	0.216	0.0152
1293.75–1370.25	1.980	0.023	0.000	2.019	–0.482	–0.008
1372.5–1449	1.821	0.393	0.000	1.888	0.314	0.003
1451.25–1527.75	1.976	0.069	0.000	2.028	–0.256	0.0036
1530–1606.75	1.935	0.214	0.000	1.991	0.081	–0.005
1608.75–1685.25	1.980	0.025	0.000	2.012	–0.374	–0.0008
MLR model						
900–976.5	1.46	0.676	–0.000	2.101	0.291	–0.084
978.75–1055.25	1.602	0.588	0.000	2.319	0.108	–0.023
1057.5–1134	1.560	0.603	–0.000	2.766	–0.054	–0.185
1136.25–1212.75	1.313	0.749	0.000	1.924	0.457	0.013
1215–1291.5	1.577	0.605	0.000	2.493	0.148	0.0381
1293.75–1370.25	1.549	0.623	–0.000	2.419	0.083	–0.106
1372.5–1449	1.457	0.677	0.000	2.282	0.219	–0.10
1451.25–1527.75	1.583	0.601	0.000	2.271	0.173	0.009
1530–1606.75	1.331	0.740	–0.000	2.167	0.350	–0.09
1608.75–1685.25	1.517	0.643	–0.000	2.634	0.073	–0.131

SEC: Standard error of calibration, SEP: Standard error of prediction, R: Multiple correlation coefficient. PLS: Partial least square; MLR: Multiple linear regression Number of elements in each case = 114

Table 3 Effect of data treatment on MLR, NIR modelling for the best range of wavelength selected

Parameters	Wavelength range, nm	Data treatment	Nr of elements	Calibration			Validation		
				SEC	R	Bias	SEC	R	Bias
TSS	1136.25–1212.75	No treatment	109	1.087	0.807	–0.000	1.630	0.565	0.034
		Smoothing	109	1.097	0.803	–0.001	1.630	0.574	0.024
		MSC	101	0.829	0.887	0.035	1.258	0.745	–0.00
		2nd order deriv	109	1.146	0.783	0.00	0.033	0.53	0.033
Acidity	900–976.5	No treatment	109	0.077	0.88	–0.000	0.025	0.718	0.000
		Smoothing	109	0.017	0.876	–0.000	0.025	0.728	–0.000
		MSC	106	0.016	0.890	–0.001	0.024	0.752	–0.001
		2nd order deriv	109	0.017	0.876	0.000	0.025	0.725	–0.000
	1215–1291.5	No treatment	108	0.018	0.851	0.000	0.028	0.627	–0.000
		Smoothed	108	0.018	0.845	0.000	0.028	0.629	–0.001
		MSC	106	0.017	0. 854	–0.00	0.027	0.646	–0.001
		2nd order deriv	108	0.018	0.850	0.000	0.027	0.648	0.000
Acidity/TSS ratio	900–976.5	No treatment	105	0.001	0.893	–0.000	0.002	0.751	–0.000
		Smoothing	108	0.001	0.875	–0.000	0.002	0.721	–0.000
		MSC	106	0.001	0.880	–0.000	0.002	0.712	–0.000
		2nd order deriv	109	0.001	0.876	–0.000	0.002	0.726	–0.000

NIR: Near infrared, MSC: Multiple scatter correction. SEC, SEP, R, MLR,TSS: See Table 2

Table 4 SEC, SEP, R and biases of PLS and MLR models for different range of NIR wavelengths for prediction of acidity

Wavelength range, nm	Calibration			Validation		
	SEC	R	Bias	SEP	R	Bias
PLS model						
900–976.5	0.028	0.736	.000	0.031	0.66	–0.000
978.75–1055.25	0.030	0.665	0.000	0.033	0.57	–0.000
1057.5–1134	0.030	0.642	0.000	0.035	0.536	0.000
1136.25–1212.75	0.032	0.633	0.000	0.034	0.550	–0.000
1215–1291.5	0.316	0.633	0.000	0.033	0.575	–0.000
1293.75–1370.25	0.032	0.615	0.000	0.035	0.520	0.000
1372.5–1449	0.032	0.619	–0.000	0.034	0.547	–0.000
1451.25–1527.75	0.032	0.618	0.000	0.034	0.561	–0.000
1530–1606.75	0.032	0.613	–0.000	0.035	0.517	0.000
1608.75–1685.25	0.032	0.611	0.000	0.034	0.560	–0.000
MLR model						
900–976.5	0.021	0.853	–0.000	0.039	0.481	–0.003
978.75–1055.25	0.025	0.787	0.000	0.040	0.418	–0.001
1057.5–1134	0.024	0.815	0.000	0.038	0.491	–0.001
1136.25–1212.75	0.025	0.781	0.000	0.048	0.477	0.003
1215–1291.5	0.024	0.813	–0.000	0.037	0.512	–0.001
1293.75–1370.25	0.027	0.750	–0.000	0.042	0.379	–0.001
1372.5–1449	0.025	0.783	0.000	0.042	0.363	–0.002
1451.25–1527.75	0.029	0.711	0.000	0.058	0.337	0.003
1530–1606.75	0.025	0.784	–0.000	0.040	0.416	–0.001
1608.75–1685.25	0.024	0.811	–0.000	0.0353	0.565	–0.001

SEC, SEP, R, PLS, MLR: See Table 2, 3. Number of elements in each case = 114

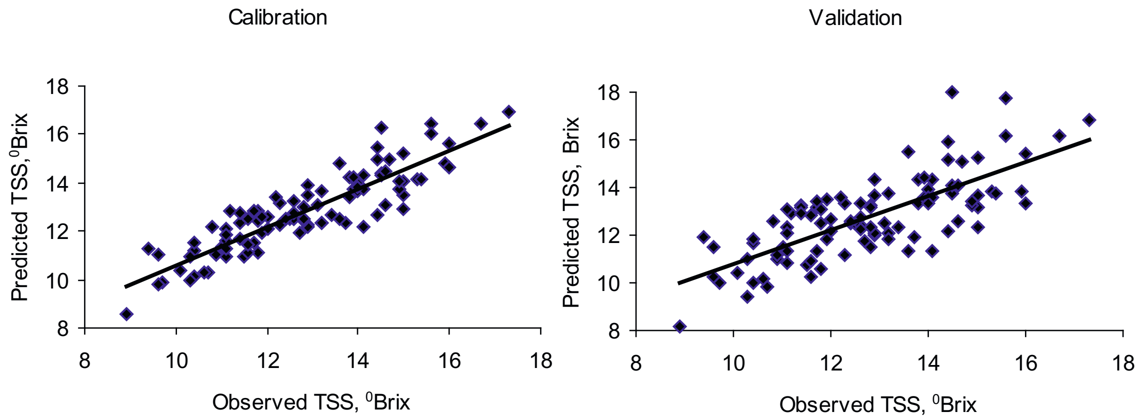

Fig. 2 Observed and predicted TSS in wavelength range of 1136.25–1212.75 nm after MSC treatment of calibration and validation set of samples

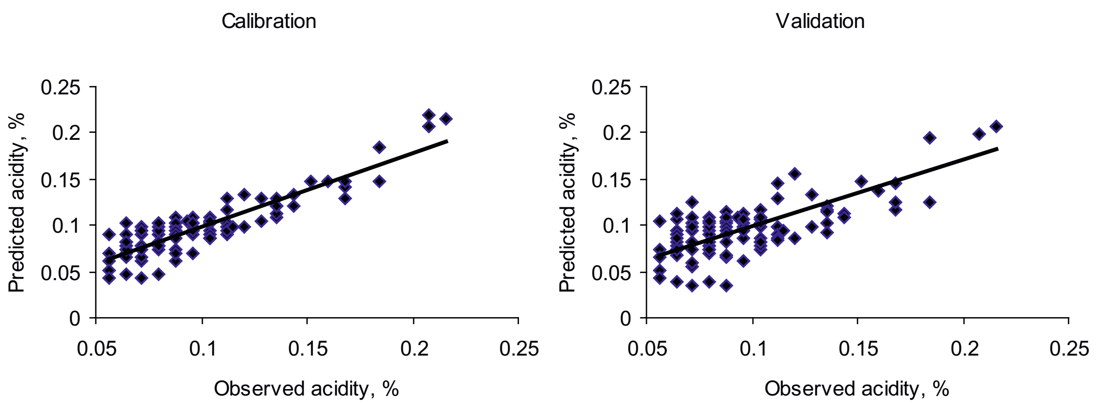

Fig. 3 Observed and predicted acidity of apple in wavelength range of 900–976.5 nm after MSC treatment of spectra of calibration and validation set of samples

actual values, but the lower SEC, SEP, higher R values and lower differences in them indicate that prediction of acidity is much better and stable. These parameters without giving any treatments to data were also closer to the values of MSC treated data. Thus, one may choose spectra of wavelength range 900–976.5 nm without treatment or with MSC treatment depending upon the accuracy requirement for prediction of acidity. Prediction of acidity is better than that of TSS in case of apple and the similar trends have been obtained by Lammertyn et al. 1998. Scatter plots (Fig. 3) drawn from the selected model indicate that slope of the curve is near to ideal 45° and thus proves that predicted values may be near to the measured one.

Wavelength range (900–976.5 nm) and regression method (MLR) for prediction of acidity/TSS ratio was found to be the same as that of acidity (Table 5). But after applying data treatment techniques not much improvement was found in R values. Thus model without data treatment whose SEC, SEP and R values found to be 0.001, 0.002, 0.893 for calibration and 0.751 for validation, respectively (Table 3) may be used for prediction of acidity/TSS ratio. Intercepts of the scatter plots (Fig. 4) were minimal and slopes were near to 45°, which shows that prediction is closer to actual values. Acidity/TSS ratio for apple has not been predicted using NIRS. The predictability is, however, in line of that of acidity and TSS reported in literature (Lammertyn et al. 1998, Ventura et al. 1998). The comparison of statistical results indicated that quality of apple could better be judged using acidity rather than TSS. The taste of apple however could be judged with almost same accuracy as acidity, using acidity/TSS ratio, which include both acidity and TSS. The multiple correlation coefficients in all cases are little lower than some of the reported values because of inclusion of five varieties to nullify the effect of varieties during commercial grading of apple.

Conclusion

Calibration models for different groups of wavelengths in the range of 900–1700 nm for prediction of TSS, acidity and acidity/TSS ratio of apples using PLS and MLR regression methods with respect to transmittance and its second-order derivatives, smoothing and MSC were developed and tested employing cross validation. MLR predicts the TSS fairly in

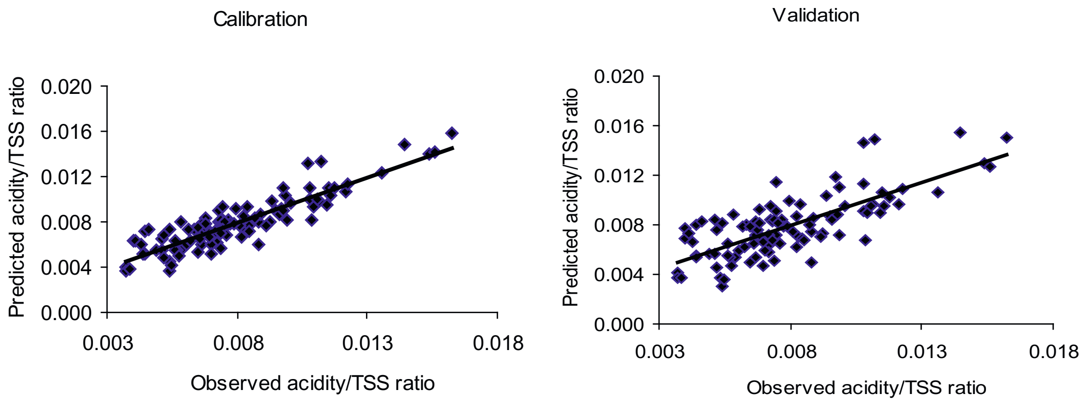

Fig. 4 Observed and predicted acidity/TSS ratio of apple in wavelength range of 900–976.5 nm for calibration and validation set of samples; TSS = total soluble solids

Table 5 SEC, SEP, R and biases of PLS and MLR models for different range of NIR wavelengths for prediction of acidity/TSS ratio

Wave lengths range, nm	Calibration			Validation		
	SEC	R	Bias	SEP	R	Bias
PLS model						
900–976.5	0.002	0.763	0.000	0.002	0.674	–0.00
978.75–1055.25	0.002	0.690	0.000	0.002	0.612	–0.000
1057.5–1134	0.002	0.651	0.000	0.003	0.536	–0.000
1136.25–1212.75	0.002	0.696	0.000	0.002	0.637	0.000
1215–1291.5	0.002	0.674	0.000	0.002	0.622	–0.000
1293.75–1370.25	0.002	0.655	0.000	0.002	0.576	0.000
1372.5–1449	0.002	0.637	–0.000	0.002	0.564	0.002
1451.25–1527.75	0.002	0.631	0.000	0.002	0.543	–0.000
1530–1606.75	0.002	0.626	0.000	0.002	0.551	0.000
1608.75–1685.25	0.002	0.632	0.000	0.002	0.585	–0.000
MLR model						
900–976.5	0.002	0.848	–0.000	0.003	0.549	–0.000
978.75–1055.25	0.002	0.787	0.000	0.003	0.428	–0.000
1057.5–1134	0.002	0.797	0.000	0.003	0.459	0.000
1136.25–1212.75	0.002	0.806	0.000	0.003	0.489	0.000
1215–1291.5	0.002	0.814	–0.000	0.003	0.473	–0.000
1293.75–1370.25	0.002	0.781	–0.000	0.003	0.466	0.000
1372.5–1449	0.002	0.823	–0.000	0.003	0.582	–0.000
1451.25–1527.75	0.002	0.706	–0.000	0.003	0.278	0.0002
1530–1606.75	0.002	0.775	–0.000	0.003	0.471	0.000
1608.75–1685.25	0.002	0.783	–0.000	0.003	0.500	0.000

SEC, SEP, R, PLS, NIR, MLR: See Table 2, 3. Number of elements in each case = 114

the wavelength range of 1136.25–1212.75 nm without any data treatment. The prediction of acidity and acidity/TSS ratio was more accurate in the wavelength range 900-976.5 nm without treatment. The treatment of data with MSC improves the prediction in case of TSS whereas the improvement in prediction of acidity and acidity/TSS ratio was negligible.

References

Aneshansley DJ, Throop JA, Upchurch BL (1997) Reflectance spectra of surface defects on apples. Sensors for nondestructive testing of apples. In: Proc of Sensors for Non-destructive Testing. Int Conference; Orlando, Fla.; Feb 18–21, Ithaca, NY, Northeast Regional Agric Eng Service Cooperative Extension, p 143–160

AOAC (1990) Official methods of analysis. 14th edn, Association of Official Analytical Chemists, Washington DC

Bellon V, Sevila S (1993) Optimization of a nondestructive system for on-line infrared measurement of fruit internal quality. In: Proc 4th Int Symp Fruit, Nut and Vegetable Production Engineering, Zaragoza, Spain, p 317–325

Iwamoto M, Kawano S, Ozaki Y (1995) An overview of research and development of near infrared spectroscopy in Japan. J Near Infrared Spectroscopy 3:179–189

Jha SN, Matsuoka T (2004) Non-destructive determination of acid–brix ratio of tomato juice using near infrared spectroscopy. Int J Food Sci Technol 39:425–430

Jha SN, Chopra S, Kingsly ARP (2005) Determination of sweetness of intact mango using visual spectral analysis. Biosyst Eng 91:157–161

Jha SN, Kingsly ARP, Chopra S (2006) Non-destructive determination of firmness and yellowness of mango during growth and storage using visual spectroscopy. Biosyst Eng 94:397–402

Jha SN, Chopra S, Kingsly ARP (2007) Modelling of colour values for non-destructive evaluation of maturity of mango. J Food Eng 78:22–26

Kato K (1997) Electrical density sorting and estimation of soluble solids content of watermelon. J Agric Eng Res 67:161–170

Kawano S, Abe H, Iwamoto M (1995) Development of a calibration equation with temperature compensation for determining the brix value in intact peaches. J Near Infrared Spectroscopy 3:211–118

Lammertyn J, Nieolai B, Oomsk Smedt VD, Baerdemaeker JD (1998) Non-destructive measurement of acidity, soluble solids and firmness of Jonagold apples using NIR-spectroscopy. Trans ASAE 41:1089–1094

Lesage P, Destain MF (1996) Measurement of tomato firmness by using a non-destructive mechanical sensor. Postharvest Biol Technol 8:45–55

Lu R, Ariana D (2002) A Near Infrared sensing technique for measuring internal quality of apple fruit. Appl Eng Agric 18: 585–590

Nussinovitch A, Ward G, Mey-Tal E (1996) Gloss of fruits and vegetables. Lebensm Wiss Technol 29:184–186

Saranwong S, Sornsrivichai J, Kawano S (2004) Predictions of ripen stage quality of apple fruit from its harvest quality measured non-destructively by near infrared spectroscopy. Postharvest Biol Technol 31:137–145

Schmilovitch Z, Mizrach A, Hoffman A, Egozi H, Fuchs Y (2000) Determination of apple physiological indices by near infrared spectrometry. Postharvest Biol Technol 19:245–252

Ventura M, Jager A, Putter H, Roelofs FPMM (1998) Non-destructive determination of soluble solids in apple fruit by near infrared spectroscopy. Postharvest Biol Technol 14: 21–27