Quantitative and qualitative measurement of pear firmness based on near infrared spectroscopy and chemometrics

Fu Xiaping, Ying Yibin, Zhou Ying, Lu Huishan, Xu Huirong

(*College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou* 310029, China)

Abstract: Firmness is one of the most important characteristics to estimate fruit maturity and quality. The potential of near-infrared (NIR) diffuse reflectance spectroscopy as a nondestructive way for pear firmness evaluation of three varieties (‘Cuiguan’, ‘Xueqing’ and ‘Xizilv’) was studied, both quantitatively and qualitatively. NIR models were established using partial least square (PLS) methods in the spectral range of 800 to 2500 nm. For quantitative analysis, the correlation coefficient r increased with more varieties involved in the model. Best results were obtained in the model for all three varieties: r_{cal} was 0.934, root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP) were 2.06 N and 3.14 N, respectively. For qualitative analysis, the overall accuracies of discriminant PLS models for classifying pears into three firmness levels: low, medium and high firmness level were not so good, percentage of samples correctly classified ranged from 70.63% to 81.25% for calibration and from 56.25% to 74.38% for validation. The results indicate that NIR spectroscopy together with PLS chemometrics method is feasible for quantitative analysis of pear firmness, however, the classification accuracy is too low to put into practical application.

Keywords: Firmness, pear, near infrared, partial least square, quantitative measurement, qualitative measurement

DOI: 10.3965/j.issn.1934-6344.2008.01.069-074

Citation: Fu Xiaping, Ying Yibin, Zhou Ying, Lu Huishan, Xu Huirong. Quantitative and qualitative measurement of pear firmness based on near infrared spectroscopy and chemometrics. Int J Agric & Biol Eng. 2008;1(1): 69－74.

1 Introduction

Firmness is one of the most important internal quality indices of fruits including pears. As pears ripen, the puncture test value (the ability of pear flesh to resist compressive force) declines, which indicates the fruit maturing and softening. Even though many techniques have been employed for firmness measurements, such as penetration, finite element modal analysis, laser air-puff method, and so on[1-5], few of these techniques are accurate and fast enough. Hand-held penetrometer measurement, a commonly used destructive method for firmness detection, varies greatly with operator’s skill and carefulness. Therefore, development of nondestructive techniques for measurement of fruit firmness is beneficial to both planters and sellers.

Measurement of near infrared (NIR) properties of fruits has been one of the most successful nondestructive techniques for internal quality assessment such as sugar content, acidity, dry-matter[6-9]. NIR spectra can be measured by reflectance, transmittance, or interactance modes[10] . Diffuse reflectance measurements are easier to obtain than transmittance measurements. Partial least square (PLS) regression is a commonly used method in chemometrics for modeling and well suited for problems with multi-collinear predictor and response variables (such as quality attributes and NIR spectra)[11]. NIR diffuse reflectance together with PLS regression has been proved to be a promising technique for quantitative analysis of fruit internal quality. McGlone et al.[12] developed multivariate models for predicting kiwifruit firmness from NIR measurements using a narrow spectral range from 800 nm to 1100 nm. Lu and Ariana[6,13] studied the potential of NIR reflectance spectroscopy for firmness detection of sweet cherries and apples in the region of 800-1700 nm and 900-1500 nm, respectively. Their study on cherries indicated that no single wavelength was strongly correlated with the firmness. Dijk, et al.[14] developed practical applicable models which were capable to describe and to predict the firmness of tomatoes during storage using near infrared spectral data in the region of 1100−2500 nm. Zude et al.[15] used VIS/NIR spectrometer to predict apple fruit

Received date: 2008-07-14 Accepted date: 2008-08-06

Authors: Fu Xiaping, PhD; **Ying Yibin**, ASABE Member Engineer, Professor; **Zhou Ying**, Graduate student; **Lu Hushan**, PhD, **Xu Huirong**, PhD Candidate, Associate professor, College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, 310029, China.

Corresponding author: Ying Yibin, College of Biosystems Engineering and Food Science, Zhejiang University, 268 Kaixuan Road, Hangzhou, 310029, China; phone: 86-571-86971885; Fax: 86-571-86971885;
Email: ybying@zju.edu.cn

flesh firmness on tree and in shelf life. Their results showed the potential of non-destructive spectral measurements for predicting accepted fruit parameters enabling the determination of optimum harvest date and fruit quality in shelf life.

Alternatively, NIR spectral data of a sample can be treated as a signature, allowing samples to be grouped on the basis of their spectral similarities[16], namely, qualitative analysis. NIR spectroscopy for qualitative analysis is often used in production control, for example, to detect adulteration of food product[17,18]. Besides, it is sometimes used to discriminate food with different varieties or blends[19,20]. In addition, NIR spectra can also be used to classify fruit according to the internal quality levels. For example, Park et al.[21] reported on principal component regression (PCR) and Mahalanobis Distance (MD) analysis for soluble solid content (SSC) and firmness evaluation of two variety apples and investigated two regions of 800-1100 nm and 400-1800 nm when conducting MD analysis to classify apples according to SSC and firmness values. For SSC, the accuracies of PCR models were high and the MD classifier also performed well. However, both results of quantitative PCR models and qualitative MD classifier for firmness were much worse than those of SSC.

As defined by ASTM (American Society for Testing Materials), NIR spectral region is from 780 nm to 2526 nm. Studies that have been reported on using NIR techniques for measuring firmness mostly used part of the NIR region. The objective of this research was to evaluate the firmness of pears both quantitatively and qualitatively by NIR diffuse reflectance spectroscopy and chemometrics. More specific objectives were: 1) to measure the diffuse reflectance of pears over the spectral region between 800 nm to 2500 nm; 2) to develop quantitative model for firmness prediction of pears by PLS regression method; 3) to develop qualitative model for firmness classification of pears by discriminant PLS method.

2 Materials and methods

2.1 Materials

Three varieties of pears, ‘Xizilv’, ‘Cuiguan’, and ‘Xueqing’ (n=160 of each variety), were used for the experiments of this study. These fruits were hand-picked from Zhejiang University’s orchard in July and August, 2005. According to the experience and expertise of the orchard manager, the commercial harvest dates for ‘Xizilv’, ‘Cuiguan’, and ‘Xueqing’ were around July 20th, July 25th, and August 10th, respectively. In order to cover a wide range of firmness, pears of each variety were successively picked at three times (Table 1). The selected samples were without surface defects such as bruises, injuries, or wormholes, etc. Samples were placed in room condition (25℃) for 24 hours to equilibrate to room environment after they were picked. Every sample was signed, and morphological properties (including diameter, height, and weight) of each sample were measured and recorded before spectral collection. The same samples were used for classifying pear varieties based on NIR spectroscopy before[22].

Table 1 Picked time and samples of different varieties of pears

Variety	Picked time	Number of samples
‘Xizilv’	July 10th	40
	July 20th	80
	July 30th	40
‘Cuiguan’	July 15th	40
	July 25th	80
	August 4th	40
‘Xueqing’	July 31st	40
	August 10th	80
	August 20th	40

2.2 Spectral measurement

An FT-NIR spectrometer (Thermo Electron Corp., USA), equipped with a bifurcated optical fiber probe, was used to acquire diffuse reflectance spectra in a wavelength range of 800～2630 nm (InGaAs detector)[22]. However, the signal to noise ratio above 2500 nm was very weak. The light source of the spectrometer was a 50 W quartz halogen lamp. The optical fiber probe was enclosed in a 16 mm diameter stainless steel cylindrical tube. Both light source beams and receptor beams were enclosed in it randomly.

Fruit were placed centrally and steadily upon the fruit holder by hand, with stem-calyx axis horizontal. There is a rubber grommet between fruit sample and fiber probe, acted as both a light seal against surface reflections and a flexible support to accommodate different sample sizes. Diffuse reflectance spectra were collected from three locations around the equator of each pear, about 120° apart. These locations were marked for subsequent Magness-Taylor(MT) firmness measurements. The signals accumulated over 64 repetitive scans were averaged and then transformed to absorption, log(1/*R*), to create one spectrum per measurement with resolution of 2 cm^{-1}. The reflectance spectrum of each pear was obtained by averaging the three spectra. Before fruit spectra acquisition, a reference spectrum was collected from a white Teflon cylinder. Spectra data acquisition and storage were achieved with a computer by running specially developed software OMNIC v6.1 (Thermo Electron corp., Madison, Wisc, USA).

2.3 Firmness measurement

MT firmness of each pear was measured by puncture test using a standard 6 mm MT probe mounted in an Instron5543 universal testing machine (Instron Corp., USA) with a loading rate of 20 mm/min, which was interfaced to a computer to obtain continuous force-deformation curves. Only the maximum force (N) was used for analysis, with which the pear skin can be penetrated. The MT firmness was measured at the corresponding locations where NIR diffuse reflectance spectra were acquired on each pear. Therefore, three maximum forces were acquired and averaged for each pear.

2.4 Data analysis

TQ Analyst v6.2.1 software (Thermo Electron corp., Madison, Wisc, USA) was used for data analysis and model establishment. Partial least squares (PLS) regression method was employed to develop relationships between spectroscopic measurements and puncture tests for both quantitative and qualitative analysis. Models were established for each single variety, every two varieties, and for all three varieties in the region of 800-2500 nm. Samples used in each model were randomly divided into two equally set for calibration and validation, respectively. For quantitative analysis, the algorithm of PLS to predict firmness is: $firmness = \beta_0 + \sum \beta_i S(\lambda_i)$, where λ_i are wavelengths, i is from 800 to 2500, β_0 is constant coefficient, β_i are regression coefficients, and S is a discrete function representing spectral values at each wavelength of the diffuse reflectance measurements. The correlation coefficient r is calculated between measured MT firmness and firmness predicted from NIR spectra.

For qualitative analysis, each sample in the data set is assigned a dummy variable as reference value, which is an arbitrary number if the sample belongs to a particular firmness level or if it does not – in this case samples of high firmness level were assigned a numeric value of 1, medium level 2 and low level 3. Pears harvested on three different dates were pulled together for each variety before modeling. For each variety, samples were nearly equally divided into three firmness levels according to their firmness. PLS models were then developed by regression of the spectral data against the assigned reference value (dummy variable).

The optimum number of calibration factors for each model was selected on the basis of root mean square error of cross-validation (RMSECV). In quantitative PLS models, calculated statistics include root mean square error of calibration (RMSEC), the correlation coefficient of calibration (r_{cal}), and root mean square error of prediction (RMSEP). In qualitative PLS models, percentages of samples correctly classified were calculated for both calibration and validation sets.

3 Results and discussion

3.1 Morphological properties and absorbance spectra

Table 2 shows the morphological properties (including maximum diameter, height and weight) and the MT firmness of the samples used in this study. The average size of ‘Xueqing’ pears is the largest, and the average weight is much heavier than those of the other two varieties. For ‘Cuiguan’ and ‘Xizilv’, the differences of their average sizes and weights are very small. In the case of MT firmness, the ‘Xizilv’ pears were more firm than ‘Cuiguan’ and ‘Xueqing’ pears. Specifically, the MT firmness of ‘Xizilv’ ranged from 20.80 N to 49.54 N, while the MT firmness of ‘Cuiguan’ was from 19.29 N to 36.69 N and ‘Xueqing’ from 15.00 N to 35.86 N. As shown in the Table 2, the variability of firmness for ‘Xizilv’ was more than those for ‘Cuiguan’ and ‘Xueqing’. Figure 1 shows the firmness distribution of the three varieties. For ‘Xizilv’ pears, more than 90% of the samples were included in the range between 25 N and 45 N. Very few samples were very firm (＞45 N) or very soft (＜25 N). For ‘Cuiguan’ and ‘Xueqing’ pears, the firmness of most samples were between 20 N and 35 N.

Table 2 Morphological properties and firmness of pears

properties	variety	maximum	minimum	average	Standard deviation
MT-firmness /N	‘Xizilv’	49.54	20.80	33.85	5.58
	‘Cuiguan’	36.69	19.29	26.39	3.12
	‘Xueqing’	35.86	15.00	25.05	4.00
Diameter/mm	‘Xizilv’	86	64	73.99	4.42
	‘Cuiguan’	83	66	73.35	3.46
	‘Xueqing’	92	75	82.97	3.83
Height/mm	‘Xizilv’	79	53	64.79	4.66
	‘Cuiguan’	79	59	67.13	3.69
	‘Xueqing’	91	63	73.54	4.48
Weight/g	‘Xizilv’	326.57	137.44	208.04	36.23
	‘Cuiguan’	308.40	155.46	212.69	28.82
	‘Xueqing’	420.15	206.70	294.74	41.90

Figure 2 shows spectra of three typical pear samples of different varieties in the wavelength range of 800-2500 nm. For the shape of the curves, no obvious differences were detected from a visual observation of spectra among the three pear varieties. Three varieties have absorption bands around 970 nm, 1190 nm, 1450 nm, 1790 nm and 1940 nm. Absorption band at 1940 nm is related to the O-H stretching vibration and deformation vibration of water. Absorption bands at 1450 nm and 970 nm are related to the O-H first and second overtones of water, respectively. And bands at 1790 nm and 1190 nm are related to the C-H first and second overtones. As fruit ripens, the components like water content, sugar content may change, which can induce the firmness changes. Therefore, NIR spectra might be indirectly related to fruit firmness.

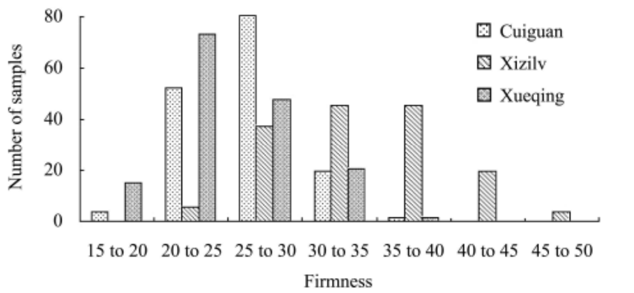

Figure 1 Distribution of firmness of three pear varieties

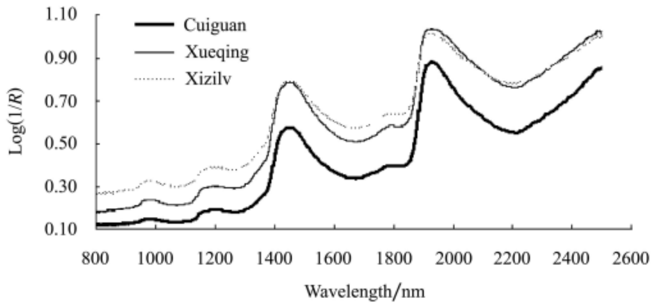

Figure 2 NIR spectra of pear samples of three varieties

3.2 Quantitative models for firmness prediction

The component information contained thousands of spectral data points can usually be condensed into a few factors. It is important to use enough factors to adequately describe all the variation in the data. However, if too many factors are used, the performance of the method may decrease because of overfitting. Figure 3 shows the *RMSECV* with different number of factors of PLS quantitative models for firmness prediction. For each model, the *RMSECV* decreased with an increasing number of factors until it exceeded a certain number. The optimum number of calibration factors was selected based on the minimum *RMSECV*. Table 3 shows the results of PLS quantitative models for firmness prediction based on the selected optimum number of factors. As shown in the Table 3, the correlation coefficients of calibration model (r_{cal}) increased with more varieties involved in the model ranged from 0.895 to 0.934. The r_{cal} of the model for all three varieties is higher than that of the models for two varieties, and the r_{cal} of the model for two varieties are higher than those of the models for each single variety. In the model for all three varieties, the r_{cal} to predict firmness was 0.934; *RMSEC* and *RMSEP* were 2.06 N and 3.14 N, respectively. Figure 4 shows the plots of measured (actual) and predicted (calculated) firmness of calibration and validation sets in the model for all three varieties.

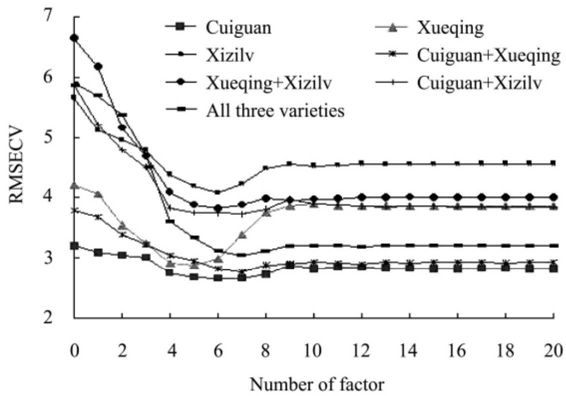

Figure 3 RMSECV with different number of factors of PLS quantitative models for firmness prediction

The reason why the results of the models for two or three varieties were better than those for single variety is that the more the varieties involved in the model, the wider the firmness range is. Firmness ranges of single varieties were shown in Table 2. From the Table 3, it can also be seen that the firmness range of 'Cuiguan'+'Xueqing', 'Xueqing'+'Xizilv', 'Cuiguan'+

Table 3 Results of PLS quantitative models for firmness prediction

Variety	Number of factors	$r_{cross-v}$	*RMSECV*/N	r_{cal}	*RMSEC*/N	*RMSEP*/N	*RPD*
'Cuiguan'	6	0.583	2.65	0.902	1.37	2.11	2.28
'Xueqing'	5	0.737	2.84	0.895	1.85	2.60	2.16
'Xizilv'	6	0.702	3.98	0.920	2.18	3.44	2.56
'Cuiguan'+'Xueqing'	7	0.715	2.67	0.912	1.50	2.42	2.43
'Xueqing'+'Xizilv'	6	0.827	3.72	0.928	2.44	3.13	2.68
'Cuiguan'+'Xizilv'	7	0.780	3.68	0.927	2.18	3.32	2.69
'Cuiguan'+'Xueqing'+'Xizilv'	7	0.845	3.10	0.934	2.06	3.14	2.83

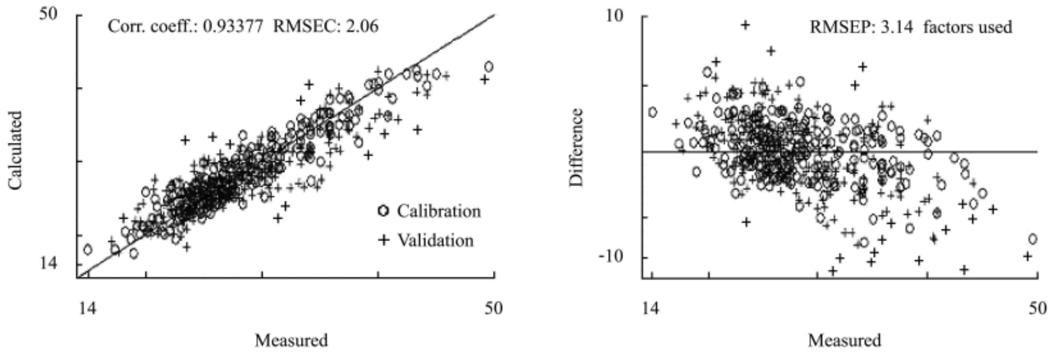

Figure 4 Quantitative analysis result of PLS model for all three varieties

‘Xizilv’, and ‘Cuiguan’+‘Xueqing’+‘Xizilv’ were 15.00～36.96 N, 15.00～49.54 N, 19.29～49.54 N, and 15.00～49.54 N, respectively. The adaptability and robustness of the statistical model could be improved when the distribution range of the firmness values increased. Harvesting the pears of each variety in three different dates was also for this reason.

3.3 Qualitative models for firmness classification

According to the distribution of firmness of pears in each model, samples were divided into three classes: low, medium and high firmness level. If the samples were not divided equally into three classes, the phenomenon was that some samples had same firmness but belonged to different classes. The firmness thresholds of each model were selected artificially, which were integers as shown in Table 4. Although the sample numbers of each class (shown in brackets) were not the same for each model, the differences between each other were the minimum by the selected integer thresholds. In the PLS qualitative models, samples in the validation set were classified as low firmness level, medium firmness level and high firmness level if their values were 0.5～1.5, 1.5～2.5, and 2.5～3.5, respectively. Otherwise, if the absolute difference value between reference value and predicted value was larger than 0.5, the sample was misclassified. Table 4 also shows the results of PLS qualitative models for firmness classification. Accuracy was calculated as the percentage of samples correctly classified. As shown in the Table 4, the overall accuracies were a bit low for both calibration and validation of the models. For calibration, the accuracies ranged from 70.63% to 81.25%; and for validation, the accuracies ranged from 56.25% to 74.38%. The firmness of most misclassified samples was near the firmness thresholds of two adjacent levels. Therefore, the low accuracies of firmness classification could be attributed to the reason that the threshold firmness value of two adjacent levels was determined subjectively, however, the firmness of the samples in a modeling data set is a successive data array. For example, three ‘Cuiguan’ pears with firmness of 20.00 N, 24.98 N and 25.02 N were classified as low firmness level for the first two samples and medium firmness level for the last one according to the threshold value of 25 N. But the difference between the first two samples (4.98 N) was much larger than that of the second and third (0.04 N). The samples with firmness around the threshold value were easily to be misclassified.

Table 4 Results of PLS qualitative models for firmness classification

Variety	Firmness level			Accuracy	
	low	medium	high	Calibration	Validation
‘Cuiguan’	< 25 N (57)	25-28 N (63)	> 28 N (40)	72.50%	62.50%
‘Xueqing’	< 23 N (49)	23-26 N (53)	> 26 N (58)	73.75%	58.75%
‘Xizilv’	< 30 N (44)	30-36 N (66)	> 36 N (50)	81.25%	65.00%
‘Cuiguan’+‘Xueqin’g	< 24 N (99)	24-27 N (119)	> 27 N (102)	77.50%	56.25%
‘Xueqing’+‘Xizilv’	< 25 N (94)	25-32 N (120)	> 32 N (106)	79.38%	74.38%
‘Cuiguan’+‘Xizilv’	< 26 N (92)	26-32 N (122)	> 32 N (106)	70.63%	70.00%
‘Cuiguan’+‘Xueqing’+‘Xizilv’	< 25 N (152)	25-30 N (167)	> 30 N (161)	80.83%	64.58%

Note: The number in the brackets means the number of samples in the corresponding firmness level.

4 Conclusion

The results of PLS quantitative models, established for each single variety, two varieties, and all three varieties, indicate that NIR diffuse reflectance measurement is feasible for quantitative analysis of pear firmness of different varieties. The r_{cal} increased with more varieties involved in the model. Best results were obtained in the model for all three varieties, r_{cal} was 0.934, *RMSEC* and *RMSEP* were 2.06 N and 3.14 N, respectively. For qualitative analysis, discriminant PLS models were established. However, the overall accuracies are not so good. For calibration, the accuracies range from 70.63% to 81.25%; and for validation, the accuracies range from 56.25% to 74.38%, which are too low to put into practical application. Therefore, more research should be conducted to improve prediction accuracy of fruit firmness using NIR spectroscopy technique.

Acknowledgements

The authors gratefully acknowledge the financial support provided by National Natural Science Foundation of China（No. 30671197）and the Program for New Century Excellent Talents in University (No. NCET-04-0524).

[References]

[1] Duprat F, Grotte M G, Pietri E, et al. A multi-purpose firmness tester for fruit and vegetables. Computers and Electronics in Agric. 1995; 12: 211－223.

[2] Dewulf W, Jancsok P, Nicolai B, et al. Determining the firmness of a pear using finite element modal analysis. J Agric Eng Research. 1999; 74: 217－224.

[3] McGlone V A, Ko S M W, Jordan R B. Noncontact fruit firmness measurement by the laser air-puff method. Trans ASAE. 1999; 42(5): 1391－1398.

[4] Peleg K. Development of a commercial fruit firmness sorter. J Agric Eng Research. 1999; 72: 231－238.

[5] Delong J M, Prange R K, Harrision P A, et al. Comparison of a new apple firmness penetrometer with three standard instruments. Postharvest Biol Tech. 2000; 19(3): 201－209.

[6] Lu R, Ariana D. Determination of firmness of intact apple fruit by near-infrared reflectance spectroscopy. Applied Eng in Agric. 2002; 18(5): 585－590.

[7] Clark C J, McGlone V A, Requejo C, et al. Dry matter determination in ‘Hass’ avocado by NIR spectroscopy. Postharvest Biol Tech. 2003; 29(3): 300－307.

[8] Ying Y, Liu Y, and Tao Y. Non-destructive quantification of soluble solids content and available acidity of apples using FT-NIR spectroscopy. Applied Optics. 2005; 44(25): 5224－5229.

[9] Liu Y, Ying Y. An optical system for measurement of internal pear quality using near-infrared spectroscopy. J Optical Eng. 2005; 47(7): 1－5.

[10] Schaare P N, Fraser D G. Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (Actinidia chinesis). Postharvest Biol Tech. 2000; 20: 175－184.

[11] Ramadan Z, Hopke P K, Johnson M J, Scow K M. Application of PLS and Back-Propagation Neural Networks for the estimation of soil properties. Chemometrics and Intelligent Lab Syst. 2004; 75(1): 23－30.

[12] McGlone V A, Kawano S. Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. Postharvest Biol Tech. 1998; 13: 131－141.

[13] Lu R. Predicting firmness and sugar content of sweet cherries using near-infrared diffuse reflectance spectroscopy. Trans ASAE. 2001; 44(5): 1265－1271.

[14] Dijk C V, Boeriu C, Stolle-Smits T, et al. The firmness of stored tomatoes (cv. Tradiro). 2. kinetic and near infrared models to describe pectin degrading enzymes and firmness loss. J Food Eng. 2006; 77(3): 585－593.

[15] Zude M, Herold B, Roger J, et al. Non-destructive tests on the prediction of apple fruit flesh firmness and soluble solids content on tree and in shelf life. J Food Eng. 2006; 77(2): 254－260.

[16] Kim J, Mowat A, Poole P, et al. Linear and non-linear pattern recognition models for classification of fruit from visible-near infrared spectra. Chemometrics and Intelligent Lab Syst. 2000; 51(2): 201－216.

[17] Ding H B, Xu R J. Near-Infrared Spectroscopic Technique for Detection of Beef Hamburger Adulteration. J Agric Food Chem. 2000; 48(6): 2193－2198.

[18] Downey G, Mcintyre P, Davies A N. Detecting and Quantifying Sunflower Oil Adulteration in Extra Virgin Olive Oils from the Eastern Mediterranean by Visible and Near-Infrared Spectroscopy. J Agric Food Chem. 2002; 50(20): 5520－5525.

[19] Cozzolino D, Smyth H E, Gishen M. Feasibility study on the use of visible and near-infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. J Agric Food Chem. 2003; 51(26): 7703－7708.

[20] Esteban-Diez I, Gonzalez-Saiz J M, Saenz-Gonzalez C, et al. Coffee varietal differentiation based on near infrared spectroscopy. Talanta. 2007; 71(1): 221－229.

[21] Park B, Abbott J A, Lee K J, et al.. Near-infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of Delicious and Gala apples. Trans ASAE. 2003; 46(6): 1721－1731.

[22] Fu X, Zhou Y, Ying Y, et al. Discrimination of pear varieties using three classification methods based on near-infrared spectroscopy. Trans ASABE. 2007; 50(4): 1355－1361.