Study of the Spectroscopic Data Analysis Pretreatments for Enhancing Performance of NIR Calibration Model for Determining the Brix Value of Japanese Pear

Karunrat Sakulnarmrat[1]* and Piyamart Jannok[2]

[1]Faculty of Agriculture and Technology, Rajamangala University of Technology Isan, Surin campus, Muang, Surin, 32000

[2]Faculty of Engineering and Architecture, Rajamangala University of Technology Isan, Nakorn Ratchasima, 30000

*Corresponding author. E-mail address: karunrat@hotmail.com

Received: 31 May 2016; Accepted: 30 June 2016

Abstract

Spectral pretreatment is one of essential procedures to improve data quality prior to establishment of calibration model and removal of physical interferences from spectra, which could enhance the model accuracy. An evaluation of Japanese Pear (*Pyrus pyrifolia*) sweetness using Near Infrared (NIR) spectroscopy technique with short wavelength (700 ~ 1100 nm) region and interactance mode was investigated. Several multivariate calibration techniques were compared and validated by establishing figures of merit. Various pretreatments with truncation were also employed in order to construct the best model of sweetness assessment. This included an individual and combination pretreatments. The best calibration was obtained through combination of multiplicative scatter correction (MSC) and second derivative (MSC+2D). It was characterized as follows: (F = 3), R^2 = 0.77, SEC = 0.56°Brix, SEP = 0.64°Brix and bias of -0.10°Brix. The results suggest potential application of NIR technique by food industries to assess pear sweetness that is non-destructive, fast and flexible.

Keywords: Pear, Near infrared spectroscopy, spectroscopic data analysis, spectral pretreatment

Introduction

Pear (*Pyrus pyrifolia*) is a fruit tree native to coastal and mildly temperate regions of western Europe, east regions of north Africa and across Asia, including China, Taiwan, Japan, and Korea. Currently cultivars derived from *Pyrus pyrifolia* are grown commercially throughout East Asia, Australia, New Zealand, and the United States (particularly Californai). Depeding on the production site the fruit is called various names as Asian pear, Chinese pear, Korean pear and Japanese pear (Wikipedia, 2014). Cultivar of *P. x bretschneideri* and *P. ussuriensis* are called Nashi pear (Wikipedia, 2014).

The large, fragrant and juicy pear fruit has a high water content which makes it-unsuitable for baking, including making a pie. Pears are utilized predominantly as a fresh fruit. The sweetness of fresh fruit is typically an important quality factor affecting customer preferences. Therefore, it is essential for the fruit industry that a fast and reliable method to evaluate the sweetness of a raw fruit is available. Near Infrared (NIR) spectroscopy represents a modern technique recognized as fast, non-destructive, low-cost and reliable method to measure the sugar content of fruit, preferred to a conventional measurement using a refractometer, which is fast but destructive.

The NIR spectroscopy technique is increasingly applied to characterize various fruit quality indexes, in particular the sweetness of an intact peach fruit (Kawano, Watanabe, & Iwanoto, 1992), mango (Saranwong, Sornsrivichai, & Kawano, 2003a), apple (Shi, Ji, Zhu, Tu, & Qing, 2008; Angra, Dimri, & Kapur, 2009; Fan, Zha, Dub, & Gao, 2009; Paz, Sanchez, Perez-Marin, Guerrero, & Garrido-Varo, 2009; Jannok, Kamitani, & Kawano, 2014), pear (Sirisomboon, Tanaka, Fujita, & Kojima, 2007; Jannok et al., 2014) and persimmon (Jannok et al., 2014). Due to interference by other factors (e.g. light scatter, sample position, sample presentation, temperature)

scattering effects occur and, therefore, the applied methods require constant improvement, calibration and validation. Scattering effects were monitored by employing different pretreatments.

Spectral pretreatments are essential tools to improve data quality prior to establishing the model. It consists in removing physical information/effects from the spectra in order to enhance the repeatability/reproducibility of the method, model robustness as well as accuracy. However, there still remain some effects, which require thorough examination in order to determine whether they need to be eliminated without affecting the validity of the method. The second derivative pretreatment is also common and is useful for removing the baseline shift, resolving broad peaks and reducing temperature effects of fruit spectra (Kawano, Abe, & Iwamoto, 1995; Saranwong, Sornsrivichai, & Kawano, 2003b).

In a case of interactance mode in sample evaluation, the particular variances come from temperature change, relative humidity, and NIR instrument system response (Sakulnarmrat et al., 2015). Thus, a slight change in ambient temperature could easily influence fruit sample temperature, resulting in spectral interference (Kawano et al., 1995).

In order to minimize/eliminate the scattering and sample size effects, spectral pretreatments such multiplicative scatter correction (MSC) are generally used. MSC correction is achieved by regressing a measure spectrum against a reference spectrum and then correcting the measured spectrum using the slope and intercept of the linear fit. MSC has proven to be effective to eliminate baseline offsets and multiplicative effect (Osborne, Fearn, & Hindle, 1993).

The calibration set that can be developed to evaluate a calibration model should provide all the likely sources of variance. Evaluating the Brix value of pear using NIR spectroscopy has been reported by previous studies and different results had been obtained. Jannok et al. (2014) reported the SEP of 0.40°Brix and bias of 0.01°Brix, while Ying, Liu, and Fu (2006) showed SEP of 0.32°Brix and correlation coefficient (R) of 0.87. Machado et al. (2012) characterized pear quality using Vis/NIR spectroscopy in transmittance mode and the result was satisfactory for evaluation of soluble solid content (SEP 0.25 ~ 0.34°Brix with R of 0.92 ~ 0.97 for three different cultivars). Japanese (Tsai, Chen, Hsieh, & Sheng, 2007) and Chinese pears (Ying & Liu, 2008) were evaluated for soluble solid content and the authors reported SEP of 0.40 ~ 0.62°Brix and R ranging from 0.60 ~ 0.85. However, the authors did not pay much attention to the effects influenced by spectral pretreatments which are known as one important step involved in calibration. Therefore, within the present study, the performance of individual and combinations of pretreatments was examined in order to establish the most suitable model for evaluation of the sweetness of pear.

Materials and methods

Sample

The pear sample (*Pyrus*, Rosaceae), (a total of 150 fruits) was obtained from Nishimuta store during the harvesting time in August 2014 (Kagoshima, Japan). Samples were stored at 4 °C before the experiment taking place the following day. On the day before the experiment (conducted in Food and Biosystems Science Laboratory, Faculty of Agriculture, Kagoshima University), samples were gently cleaned by wiping with a tissue the area designated for spectrum scanning. Then the samples were labeled and stored in the controlled temperature room at 25 °C overnight.

Spectral acquisition

The interactance mode of NIR instrument Model 6500 (Foss NIRSystems, Laurel, MD, USA) was

used to measure the NIR spectra (400 ~ 1100 nm) of pear samples in the equatorial position. The light could penetrate about 7 ~ 10 mm deep under the skin (Kawano et al., 1992; Saranwong et al., 2003a; Jannok et al., 2014). The interactance probe was composed of a concentric outer ring illuminator and an inner ring detector. Before the spectral acquisition, the sample temperature was maintained at 25 °C for at least 30 min using a water bath (T22LAS, Low Temperature Thermostatic Shaking Water Bath, Thomas Kagaku CO., Ltd., Japan), covered with a thin polyethylene film. Sample temperature was also controlled by covering it with 5 mm-thick foam sheet. A reference measurement was conducted with a white ceramic plate every time after examining five fruit. The NIR measuring conditions were as follows, (1) there were 50 scans per sample at one point, (2) the measuring position of each sample was the equator of the fruit, and (3) during the measurement, each sample was covered with an opaque-cylindrical cover.

Data analysis

The relationship between NIR spectra and Brix value was analyzed using the Unscrambler program (version 9.8, CAMO, Japan) for partial least squares (PLS) regression. Scattering effects were monitored by employing different spectral pretreatments, such as second derivative (2D), multiplicative scatter correction (MSC), standard normal variate (SNV), detrending (De) and normalization (Nor), and wavelength truncation (wavelength selection) as well as a combination of those pretreatments in different order (e.g. 2D+MSC, MSC+2D, 2D+Nor, Nor+2D) with wavelength truncation. The original NIR spectra and pretreated spectra were used to construct calibration models. The best proper calibration model was considered by the low SECV or SEP including low factor, high R^2 and bias should close to zero.

Reference analysis

After the NIR measurement, a segment of the fruit was immediately cut out (approximately 7 ~ 10 mm deep from the peel) in the NIR measuring area. The Brix value of finger-squeezed juice from each sample was measured by a digital refractometer (Model PR101, ATAGO, Tokyo, Japan). The measurement was carried out and two comparable and/or equal values were selected for data analysis (Jannok et al., 2014).

Results and Discussion

The characteristics of pears are shown in Table 1. Calibration set was used for calibration and validation set was used for validation/prediction representing from the different sample groups. The average Brix value employed in the calibration and validation set were comparable in both sets of fruit samples.

(1) NIR spectra

The original spectra of pear are presented in Figure 1, where only one peak of water band was prominent around the wavelength of 976 nm. However, the spectra shifted downward to minus y-axis region, which could be explained as a consequence of a difference in absorbance between reference and samples by which the former penetrated lower light (Terdwongworakul, Nakawajana, Teerachaichayut, & Janhiran, 2012). There is agreement between various studies that reported similarity of original spectra obtained from different kinds of fruit such as mango (Saranwong et al., 2003a; Saranwong et al., 2003b), peach (Kawano et al., 1992) and also apple, pear and persimmon (Jannok et al., 2014).

(2) Spectroscopic data analysis pretreatment

To achieve accurate model accuracy, the knowledge of interference from factors (e.g. light scatter, sample position, sample presentation, temperature) is required in order to efficiently eliminate such interference using appropriate spectral pretreatments. Within this study, the interactance mode was performed, and various spectral

pretreatments, single and combined, were compared to achieve the best model. As food matrix is complex (e.g. carbohydrate, protein, sugar, fiber), it comprises various components/characteristics which will make spectral difference. Thus, it is necessary to employ appropriate spectral pretreatments to minimize such physical effects in order to model the chemical effects more accurately. Generally, different spectral pretreatments and their combinations are considered when analyzing spectral data.

Table 1 Characteristics of calibration and validation sets of pear.

Item	Calibration set	Validation set
Number of samples	75	73
Range (°Brix)	8.25 - 13.1	8.30 - 12.7
Mean (°Brix)	10.6	10.5
Standard deviation (°Brix)	1.17	1.11

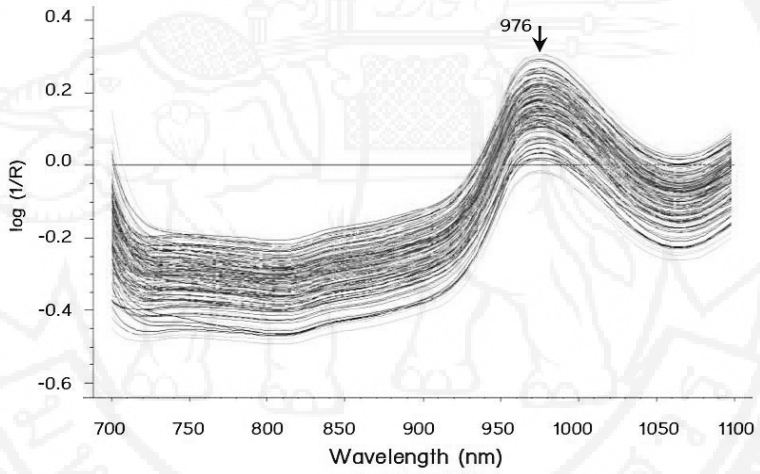

Figure 1 Original spectra of pear samples.

Thus, prior knowledge related to effects of food matrix is required in order to apply appropriate spectral pretreatments, not only random trials. Derivatives are among the most common spectral pretreatments applied to spectral data. Derivatives are mainly used to enhance resolution and eliminate constant and linear baseline drift between samples. Second derivatives are more common in practice than higher-order ones. Spectral derivatives can be calculated by obtaining the differences between two consecutive points, or by smoothing, specified gap distance or Savitzky-Golay Polynomial fitting (Osborne et al., 1993). Some common disadvantages of applying derivatives are noise increase and difficulty in spectral interpretation.

In this study, the analysis was conducted using a single spectral pretreatment such as 2D, MSC, SNV, De and Nor pretreatments, and a combination of spectral pretreatments with different orders. Multiplicative scatter correction (MSC) followed by 2D was successful to remove the scattering effect (Table 2). From the results obtained in this study, it was found that the order of spectral pretreatment provides different results (better results). At the

wavelength of 976 nm of the MSC, Figure 2 shows a strong positive peak of water band. After further 2D applied, water band at 972 nm showed clearly negative peak. Moreover, the wave-length truncation is the simplest and most effective spectral pretreatment tool applied in NIR spectroscopy if the expected region is already known. Wavelength truncation in the region with the exception of the interest compound provides better conditions in terms of removal of noise or unrelated variables, simplifying baseline variation and more robust models with fewer components. The proper truncation of this study was 800 ~ 1058 nm, which was similar to a study of Park, Abbott, Lee, Choi, & Choi at 800 ~ 1068 nm (Park et al., 2003).

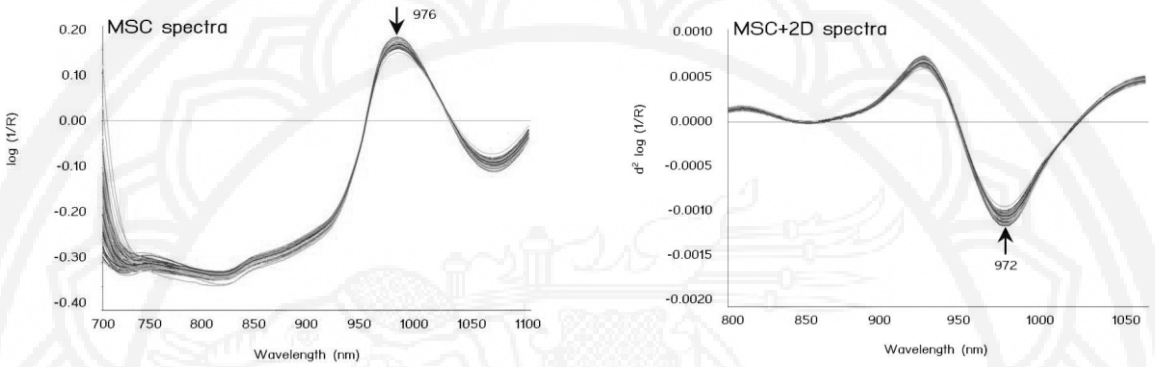

Figure 2 Pear spectra after pretreated by the multiplicative scatter correction (MSC) and the multiplicative scatter correction followed by second derivative (MSC+2D).

(3) Calibration models

The PLS regression model was established using the calibration set and validated using the validation set. The calibration models were optimized by the exclusion of two outliers, which were visibly isolated from the group (the outliers were separated from the target line by more than 3 x SEP in a tentative PLS calibration model and thus were removed (Jannok et al., 2014)). Consequently, calibration and validation sets were subjected to data sorting. Within this study, various spectral pretreatments were employed in order to evaluate the best PLS model in order to remove the effects from the instrument or environment (Table 2). PLS regression based on MSC + 2D-20pt (MSC+2D) (calibration model N) with the wavelength truncation of 800~1058 nm were developed and show the best results. The best proper calibration model was considered by the low SECV or SEP including low factor, high R^2 and bias should close to zero (Table 2).

Table 2 Spectral pretreatments of the calibration models for determining the Brix values of intact pears.

Calibration Model	Spectral Pretreatment	Wavelength (nm)	F	R^2	SEC (°Brix)	SECV (°Brix)	Bias (°Brix)
A	Original	700-1098	7	0.79	0.57	0.61	0.01
B	SNV	700-1098	6	0.79	0.54	0.61	0.01
C	De	700-1098	4	0.75	0.59	0.64	0.01
D	MSC	700-1098	6	0.79	0.53	0.60	-0.01
E	Nor	700-1098	1	0.04	1.15	1.17	0.01
F	2D-5pt	710-1088	5	0.77	0.56	0.65	0.01
G	2D-10pt	720-1078	4	0.75	0.59	0.65	0.00
H	2D-15pt	730-1068	4	0.76	0.57	0.63	0.00
I	2D-20pt	740-1058	4	0.77	0.56	0.61	0.00
J	2D-20pt	750-1058	4	0.77	0.56	0.61	0.01
K	2D-20pt	800-1058	3	0.73	0.61	0.66	0.01
L	SNV + 2D-20pt	800-1058	3	0.76	0.57	0.61	0.00
M	De + 2D-20pt	800-1058	3	0.75	0.59	0.64	0.01
N	**MSC + 2D-20pt (MSC+2D)**	**800-1058**	**3**	**0.77**	**0.56**	**0.60**	**0.00**
O	Nor + 2D-20pt	800-1058	1	0.04	1.15	1.17	0.01
P	2D-20pt + SNV	800-1058	3	0.71	0.63	0.67	0.00
Q	2D-20pt + De	800-1058	4	0.76	0.57	0.63	0.01
R	2D-20pt + MSC	800-1058	3	0.71	0.63	0.67	0.01
S	2D-20pt + Nor	800-1058	3	0.75	0.58	0.63	0.00

SNV: standard normalize variate; De: detrending; MSC: multiplicative scatter correction; Nor: normalization at 960 nm; 2D: second derivative.

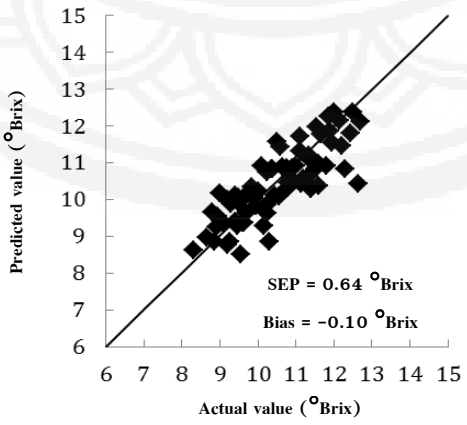

Figure 3 Scatter plot for predicting the Brix values of the validation set of pears by using the developed calibration model N.

The scatter plots as shown in Figure 3 presented that the calibration model N validates/predicts values which are located close to the target line with lower SEP and bias. This equation provides 3 factors for predicting Brix value of pear with R^2 of 0.77, SEC of 0.56 °Brix, SEP of 0.64 °Brix and bias of -0.10°Brix. In addition, the regression coefficient plots in Figure 4 present the weight at each wavelength of the calibration model N. The regression coefficient plot of calibration model N which was obtained from the MSC + 2D-20pt (MSC+2D), provided smooth and clear wavelength. The wavelengths were 918 and 998 nm which were related to sugar band (Kawano et al., 1992; Jannok et al., 2014) and the second overtone of OH stretching (Porteous, Muir, & Wastie, 1981), respectively.

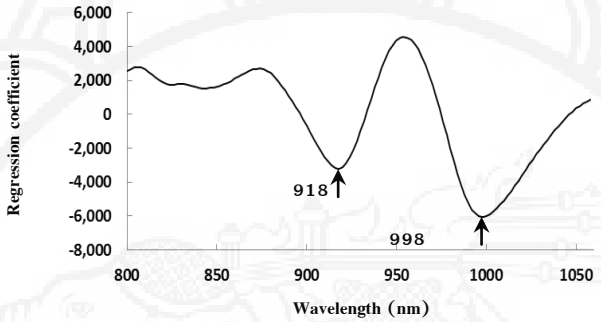

Figure 4 Regression coefficient of calibration model N.

The accuracy of NIR spectroscopy in quantitative analysis is influenced by various parameters such as instrumentation, sampling, measurement, calibration etc. (Osborne et al., 1993). This indicates that the spectral pretreatment is one of the essential procedures to improve data quality prior to establishing the calibration model by eliminating unwanted variables.

Conclusions

The performance of individual and combination of spectral pretreatments was studied for determination of the Brix value of commercially available Japanese pear. The combination of multiplicative scatter correction and second derivative (MSC+2D) was found to be more advantageous in a comparison to those others. However, selection of appropriate spectral pretreatments should be done carefully. It is important to recognize that various pretreatments exhibited different effectiveness in eliminating unwanted variations, depending on the objectives of the spectral analysis and sources of baseline variations. Both individual and combined pretreatments provide a difference regarding physical variations commonly interfering with chemical responses that occur in NIR spectroscopy. The order of combined pretreatments also affected the results. The results showed difference in removing various scattering effects due to interfering factors obtained by application of those pretreatments. The application of pretreatments should follow careful selection of factors relevant to the objectives of specific analyses in order to develop the most appropriate model.

Acknowledgements

The authors are gratefully acknowledging Rajamangala University of Technology Isan (RMUTI), Thailand and Kagoshima University, Japan for the supports. The authors also thank Mr T. Shinomiya for

his assistance in the experiments and Dr. I. Konczak for improving the use of English in the manuscript.

References

Angra, S. K., Dimri, A. K., & Kapur, P. (2009). Nondestructive Brix Evaluation of Apples of Different Origin Using Near Infrared (NIR) Filter Based Reflectance Spectroscopy. *Instrum Sci Technol*, *37*(2), 241-253.

Fan, G., Zha, J., Dub, R., & Gao, L. (2009). Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *J Food Eng*, *93*(4), 416-420.

Jannok, P., Kamitani, Y., & Kawano, S. (2014). Development of a common calibration model for determining the Brix value of intact apple, pear and persimmon fruits by near infrared spectroscopy. *J Near infrared Spectrosc*, *22*(5), 367-373.

Kawano, S., Watanabe, H., & Iwanoto M. (1992). Determination of sugar content in intact peaches by near infrared spectroscopy with fiber optics in interactance mode. *J Japan Soc Hort Sci*, *61*(1), 445-451.

Kawano, S., Abe, H., & Iwamoto, M. (1995). Development of a calibration with temperature compensation for determining the Brix value in intact peaches. *J Near Infrared Spec*, *3*(4), 211-218. http://dx.doi.org/10.1255.jnirs.71

Machado, N. P., Fachinello, J. C., Galarca, S. P., Betemps, D. L., Pasa, M. S., & Schmitz, J. D. (2012). Pear quality characteristics by Vis/Nir spectroscopy. *Ann Braz Acad Sci*, *84*(3), 853-863.

Osborne, B. G., Fearn, T., & Hindle P. H. (1993). *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis* (2nd ed.). Longman Group: Harlow, England, UK.

Park, B., Abbott, J. A., Lee, K. J., Choi, C. H., & Choi, K. H. (2003). Near-infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of Delicious and Gala apples. *Trans Amer Soc Agr Eng*, *46*, 1721- 1731.

Paz, P., Sanchez, M-T., Perez-Marin, D., Guerrero, J-E., & Garrido-Varo, A. (2009). Evaluating NIR instruments for quantitative and qualitative assessment of intact apple quality. *J Sci Food Agric*, *89*(5), 781-790.

Porteous, R., Muir, A., & Wastie, R. (1981). The identification of diseases and defects in potato tubers from measurements of optical spectral reflectance. *J Agri Eng Res*, *26*(2), 151-160.

Sakulnarmrat, K., Shinomiya, T., Jannok, P., Kamitani, Y., & Kawano, S. (2015). A simple method of system response compensation of a near infrared instrument for determining Brix value of Peaches. In 6th Rajamangala University of Technology International Conference (RMUTIC), *Green Innovation for a Better Life*, *1-3 September 2015*, Nakhonratchasima, Thailand: Rajamangala University of Technology Isan.

Saranwong, S., Sornsrivichai, J., & Kawano, S. (2003b). On-tree evaluation of harvesting quality of mango fruit using a hand-held NIR instrument. *J Near Infrared Spectrosc*, *11*(4), 283-293.

Shi, B., Ji, B., Zhu, D., Tu, Z., & Qing, Z. (2008). Study on genetic algorithms-based NIR wavelength selection for determination of soluble solids content in fuji apples. *J Food Quality*, *31*(2), 232-249.

Sirisomboon, P., Tanaka, M., Fujita, S., & Kojima, T. (2007). Evaluation of pectin constituents of Japanese pear by near infrared spectroscopy. *J Food Eng*, *78*(2), 701-707.

Terdwongworakul, A., Nakawajana, N., Teerachaichayut, S., & Janhiran, A. (2012). Determination of translucent content in mangosteen by means of near infrared transmittance. *J Food Eng*, *109*(1), 114-119.

Tsai, C. Y., Chen, H. J., Hsieh, J. F., & Sheng C. T. (2007). Fabrication of a near infrared online inspection system for pear fruit. *Int Agric Eng J*, *16*, 57-70.

Wikipedia. (2014). Pear. Retrieved from http://en.wikipedia.org/wiki/Pyrus_pyrifolia

Ying, Y., Liu, Y., & Fu, X. (2006). Performance of FT-NIR instrument for Brix value measurement of intact pear fruit. *Int J Postharv Tech Inn*, *1*(2). http://dx.doi.org/10.1504/IJPTI.2006.011665

Ying, Y., & Liu, Y. (2008). Nondestructive measurement of internal quality in pear using genetic algorithms and FTNIR spectroscopy. *J Food Eng*, *84*(2), 206-213.

Saranwong, S., Sornsrivichai, J., & Kawano, S. (2003a). Performance of a portable near infrared instrument for Brix value determination of intact mango fruit. *J Near Infrared Spectrosc*, *11*, 175-181.