Nondestructive Measurement of Sugar Content in Chestnuts Using Near-infrared Spectroscopy

Jie Liu[1], Xiaoyu Li[1*], Peiwu Li[2], Wei Wang[1], Jun Zhang[1,3], Wei Zhou[1], Zhu Zhou[1]

1. College of Engineering, Huazhong Agricultural University, Wuhan, P.R.China
2. Oil Crops Research Institute, Chinese Academy of Agricultural Science, Wuhan, P.R.China
3. Key Laboratory of Highway Construction Technology and Equipment of Ministry of Education, Chang'an University, Xi'an, P.R.China
{ Jie Liu, Xiaoyu Li, Peiwu Li, Wei Wang, Jun Zhang, Wei Zhou, Zhu Zhou}lixiaoyu@mail.hzau.edu.cn

Abstract. The chestnut (Castanea) is an important fruit in Europe and Asia. As a highly variable fruit, its quality is graded according to nutrition components, especially according to the sugar content, which are traditionally measured by using chemical methods. However, the traditional methods are time-consuming, laborious, and expensive. Here, we analyzed the sugar content of intact and peeled chestnuts by near-infrared spectroscopy. The spectra of intact and peeled chestnut samples were collected in the wavelength range from 833 nm to 2500 nm. The Sample Set Partitioning based on joint X–Y distances was used when the calibration and validation subsets were partitioned. The predictive models for intact and peeled chestnut samples respectively, were developed using partial least squares (PLS) regression based on the original spectra and the spectra derived from different pretreatments. The models developed from the spectra of peeled samples gave accurate predictions. The correlation coefficient (R^2) of the optimized model for calibration set and validation set were 0.90 and 0.86. Although the models established on the spectra of intact samples did not perform excellently, they were still qualified to measure sugar content of the chestnut kernel. The correlation coefficient (R^2) of optimized model for calibration set and validation set were 0.84 and 0.59. These results suggested that NIR spectroscopy could be used as a fast and accurate alternative method for the nondestructive evaluation of sugar content in chestnuts during orchard and postharvest processes.

Keywords: Near-infrared Spectroscopy; sugar; nondestructive measurement; PLS; chestnut.

1. Introduction

The chestnut is popular in Asia and Europe for its delicious flavor and abundant nutrition. Every year, more than 1,1Mt tons chestnuts were produced all over the world[1, 2]. However, the chestnut is a highly variable species. The nutrition components, flavor and taste are varied according the variety, origin area, cultivation and storage condition[3-5]. Therefore, evaluating the nutrition components accurately, especially the sugar content, plays an important role in determination of the price, the time for marketing and preservation, the processing methods and the product quality control[6-8]. Recently, the sugar content in chestnuts are determined by using chemical approach, which is destructive, time-consuming, laborious and expensive[4]. Therefore, developing a rapid nondestructive means to analyze the sugar content of chestnuts, especially for intact chestnuts, is critical to the chestnut industry.

Near-infrared spectroscopy (NIR) is a fast, simple non-destructive, cost-efficient and environment-friendly measurement method that requires minimal sample preparation. It has been successfully used in the quantitative and qualitative analysis of horticultural crops. Because the near-infrared region related to vibration and combination overtones of the fundamental O–H, C–H and N–H bonds, which are the primary functional groups of sugar (glucose, fructose, sucrose, etc), this technology is powerful in quantitative analysis of this component[9, 10]. Numerous papers have demonstrated the usefulness of NIR in measurement of sugar content of cereal crops, vegetables and fruits[11-14]. Additionally, NIR spectra, with penetration more than 9 mm, can provide adequate information of the core part or the area covered by peel. Tian et al used this technology to predict the firmness of watermelon[15], Han et al applied it to determine the brown heart of pears[16], and Mexis et al obtained the internal properties in packages of walnuts by this means[17]. Recent research also showed that the NIR was employed for detection of moldy chestnuts[18] and for quantitative analysis of targets with peel such as oranges[19], avocados[20], peanuts[21] and intact asparaguses[22]. However, although it is beneficial and meaningful to apply NIR spectroscopy to the nut industry, NIR has not been popularly used for evaluating the quality of nuts, such as chestnuts, walnuts, and hazelnuts. In this investigation, we exploited the feasibility of NIR spectroscopy for evaluating the sugar content of intact and peeled chestnuts to provide an effective method for the chestnut industry. Also, the efficiency and influence of different preprocessing methods on calibration models were compared to identify the optimized model, as well as to provide reference for further research.

2. Materials and methods

2.1. Fruit samples

185 chestnut smaples, from MaCheng area (Hubei, China) were used in this study, and the individual weight of intact chestnuts ranged from 9.26 g to 27.74 g. Considering the chestnut growing property, the samples were representative enough for different growing conditions, maturation and weather at harvest time. The samples were stored under standard refrigeration conditions (0 °C~4 °C, 80%~90% relative humidity, SB/T 10192-1993) for three to four months before experiment tests. The samples were exposed to temperature 26 °C for 2 hours before experimental tests to allow for temperature equilibrium. The experiment was executed in 10 days with same time interval in order to provide models with wide applicability as the change process of chestnut fruit during storage is perplexing.

2.2. Spectra collection and reference value determination

NIR spectra were acquired by using a VECTOR33 NIR spectrometer (Brucker Optics, Ettlingen, Germany) equipped with fiber optic accessories and were stored on a personal computer connected via an AQP card to the spectrometer. The instrument control, parameter setup, spectral acquisition and storage were carried out with the OPUS software (v. 4.0 Brucker Optics, Ettlingen, Germany). The environment temperature was kept at 26 °C and each sample was scanned 64 times, from which an average spectrum was calculated. Spectra were collected in reflectance mode in the range of 833 nm~2500 nm at 1.25×106 nm and stored as the absorbance mode. To stabilize the light sources, the spectrometer was warmed up for a period of 1 hour prior to measurement. Before the sample measurement, the spectrum of standard background was obtained as the basic reference. Firstly, every sample was scanned with an intact peel. To avoid differences caused by optical path and angle of incidence, the chestnut samples were scanned with their flat side facing the probe and covering the probe completely. Secondly, it was dissected parallel to the flat side and the hemispheric part was set into the sample container to be scanned. For the same reason, the probe was completely covered by the section.

After spectra acquisition, the kernel of the sample was taken out for reference value determination according to the Chinese national standard methods. The sugar content was determined by using an enzymatic method and expressed as the percentage of the fresh weight (GB T 5009.7—2003, SAC).

2.3. Data analysis

The sample set partitioning based on joint X–Y distances (SPXY) was performed to partition the calibration and validation subsets at the rate of 3 to 1. Because of its con-

sideration of the variability in both x and y information, the SPXY was cited as a method that can cover the multidimensional space more effectively in comparison with the randomly sampling (RS) or partitioning schemes based only on x- or y- space alone (such as the Kennard–Stone (KS) algorithm or concentration gradient sampling)[23]. Consequently, this method can be potentially used to improve the performance of the predictive models.

Before calibrating the models, four preprocessing methods, namely multiplicative scatter correction (MSC), standard normal variate transformation (SNV), first derivation (FD) and second derivation (SD) were applied to the spectra. These methods can effectively reduce or remove shifts of baseline and superposed peaks, or interferences due to scatter, particle size, and light distance changes[10, 24]. However, they also result in amplification of noises or loss of information. Therefore, the contribution of a preprocessing method to the predictive models depends on the characteristics of the spectra.

The partial least square (PLS) regression method was used to establish the predictive models for sugar content. The optimal latent variables were determined by the lowest root mean square error of cross validation (RMSECV) among the calibration set and the models were validated with the independent validation set. The performance of models were evaluated by the correlation coefficient (R^2) between predicted values and reference values of the parameters, the root mean square error of calibration (RMSEC) and the root mean and square error of prediction (RMSEP). The higher the R^2 is, the better the model performed, while the lower the RMSEC and RMSEP are, the better the model is. All the data analyses were carried out with the aid of Matlab software (V.7.6, Mathworks, Natick, USA) and the PLS models were developed by using the PLS toolbox (V. 5.20, Eigenvector Research, Inc., Wenatchee, USA).

3. Result and discussion

3.1. Spectra configuration, reference values, and subsets partition

Raw spectra of intact and peeled samples are presented in figure 1. It was found that the absorbance spectra of intact samples were scatter over a wide range, while those of the peeled samples fell in a narrow range. The more prominent diversity of peel surface possibly contributed to these differences. However, the two kinds of spectra show similar features over the entire spectra range except for the bands between 833 nm~1053 nm and 2000 nm~2040 nm. The former band mainly reflected the information of color or texture about the sample surface and the later band possibly revealed the features of peel component. Generally, from both the spectra of intact samples and the spectra of peeled samples, characteristic absorption bands associated with sugar can be observed. For example, the dominant peak at around 1450 nm was influenced by the first –OH overtone. The weak bands at around 1780 nm were mainly attributed to the first sugar-related overtone and the prominent peak around 1160 nm was also associated with sugar[25, 26]. Figure 2 shows the raw spectra of a randomly selected

sample with and without peel. The difference of reflectance between the two kinds of spectra is obvious, indicating the influence of the peel on the spectrum.
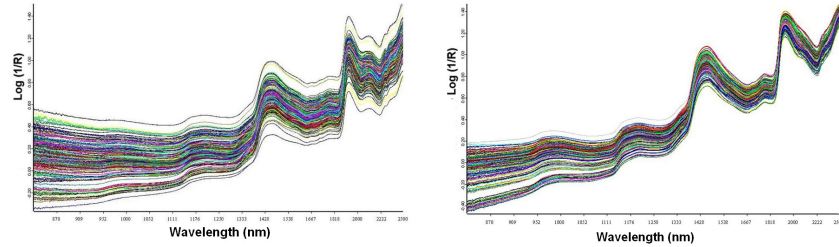

Fig. 1. Raw spectra of chestnut samples (Left: the spectra of intact samples; right: the spectra of peeled samples.)
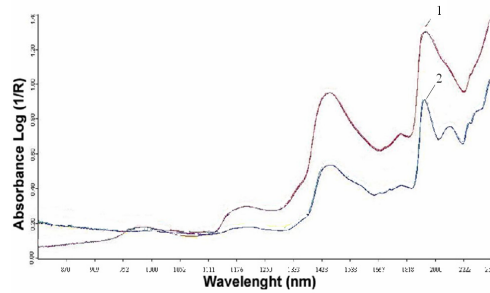

Fig. 2. Different spectra of a randomly selected sample acquired with and without peel (1: the raw spectrum of the peeled sample; 2: the raw spectrum of the intact sample).

Table 1. Statistic descriptions of subsets partition and reference sugar content (SN: sample number; SD: standard deviation; Max: maximum; Min: minimum.)

Subset	SN	Mean	SD	Max	Min
Intact Calibration	139	9.96	1.79	14.55	6.57
Intact Validation	46	9.32	1.38	12.92	6.58
Peeled Calibration	139	9.86	1.82	14.55	6.57
Peeled Validation	46	9.60	1.37	13.37	6.62
Total	185	9.80	1.72	14.55	6.57

The details of the reference data were shown in Table 1. The sugar content of samples was scattered from 6.6% to 14.6%. It is a wide range since approximately 10 % of the fresh weight is sugar normally. According to the SPXY algorithm, three quarters of the samples were chosen to be the calibration set while the rest served as independent validation. From the obvious difference between the spectra acquired in

intact and peeled condition, even of a same sample, it is clear that the subset structure for intact and that for peeled chestnuts are different. The statistic descriptions of subsets were also summarized in Table 1.

3.2 Calibration and validation of prediction models for intact sample

In the entire spectral range, the prediction models based on the raw and pretreated spectra of intact samples were established by PLS regression, respectively. Independent validation sets were used to evaluate the performance of models. Table 2 presents detailed information about the models. It can be found that, the FD model had the highest R^2 value of calibration and validation and the lowest RMSEC and RMSEP (0.84, 0.59, 1.12 and 1.32, respectively). Although the SD model had great parameters for calibration, it appeared as an over-fitting one because of the poor result for validation, which indicated the effect of derivation preprocessing in extracting the useful information in spectra and the result of this method to noise amplification along with the order increasing. The SNV was not a good preprocessing method for intact samples since the model performed worse than the one based on the raw spectra. While the MSC model had same parameters with the model for peeled samples, its latent variables was 9.

Table 2. Statistics parameters of PLS models for sugar content in intact samples (NP: no preprocessing; MSC: multiplicative scatter correction; SNV: standard normal variate transformation; FD: first derivation; SD: second derivation; R^2: correlation coefficient; RMSEC: root mean square error of calibration; RMSEP: root mean square error of prediction).

Preprocessing method	Calibration		Validation		Factors
	R^2	RMSEC	R^2	RMSEP	
NP	0.676	1.287	0.583	1.531	8
MSC	0.773	1.107	0.597	1.330	9
SNV	0.669	1.300	0.554	1.400	7
FD	0.838	1.116	0.589	1.320	4
SD	0.967	0.445	0.516	1.383	4

3.3 Calibration and validation of prediction models for peeled sample

The PLS models for peeled samples were calculated in the same way and the results were presented in Table 3. Similar with the results of models for peeled samples, the FD model gave the best performance with 0.91, 0.87, 0.76 and 0.74 as the R^2 of calibration, R^2 of validation, RMSEC and RMSEP, respectively. The SD model would also be over-fitting because of the unacceptable difference of R^2 of calibration and R^2 of validation while it had less factor number. Compared with the model based on raw spectra, the MSC and SNV models shown their capability of obtain the relevant in-

formation in spectra. However, these two models were not reliable enough because the differences between the R^2 of calibration and R^2 of validation were unacceptable.

Table 3. Statistics parameters of PLS models for sugar content in peeled samples (NP: no preprocessing; MSC: multiplicative scatter correction; SNV: standard normal variate transformation; FD: first derivation; SD: second derivation; R^2: correlation coefficient; RMSEC: root mean square error of calibration; RMSEP: root mean square error of prediction.)

Preprocessing method	Calibration		Validation		Factors
	R^2	RMSEC	R^2	RMSEP	
NP	0.505	1.558	0.625	1.112	7
MSC	0.773	1.107	0.597	1.330	3
SNV	0.699	1.278	0.809	0.810	6
FD	0.907	0.761	0.865	0.739	4
SD	0.802	1.084	0.653	1.097	2

It can be seen that, the models based on the spectra of peeled samples performed better than those based on the spectra of intact samples. The influence of peel in spectra acquisition might be the reason. However, their spectral characteristics were similar because the effects of different preprocessing methods followed the same order: the first deviation was the most powerful; the second deviation was the least effective, and the SNV and the MSC performed medium and their results were near to that of the non-preprocessing. The poor performances of MSC and SNV models both for intact and peeled samples shown that the influence of peel to predictive models came more from its obstruction to penetration instead of its surface diversity.

According to Shenk and Westerhaus[27], an R^2 value of more than 0.9 indicates excellent quantitative information, a value of between 0.70 and 0.89 means good quantitative information, and a value of between 0.50 and 0.69 is qualified for good separation of samples into high, medium and low groups. Therefore, the FD models based on the spectra of samples peeled and intact, were reliable to test the sugar content of chestnuts.

4. Conclusion

The present work addressed the assessment of chestnut quality by near-infrared spectroscopy. The presented results demonstrated the potential of NIR spectroscopy in evaluating the sugar content of intact and peeled chestnuts. The correlation coefficients of the optimized models for sugar were 0.90 for peeled samples and 0.83 for intact samples. These results proved that NIR can be used as a nondestructive and rapid means for determining the sugar content of chestnuts with reduced time and labor compared with traditional methods. Further work would be needed to develop more accurate and robust models by selecting sensitive bands, intensifying spectra energy or combining chemometrics methods. Portable or online devices could be developed

based on these results to offer information for decisions on harvest, market or process, which will benefit the whole chestnut industry.

References

1. Food and Agriculture Organization of the United Nations, http://faostat.fao.org/site/336/DesktopDefault.aspx?Page ID=336
2. Ozcimen, D., and Ersoy-Mericboyu, A.: A study on the Carbonization of Grapeseed and Chestnut Shell. Fuel Process. Technol. 89, 1041--1046 (2008)
3. Feng Y., Qin L., and Li F., Technique of Chestnut Cultivation China Agricultural Aniversity Press, BeiJing (2007)
4. Pereira-Lorenzo, S., Ramos-Cabrer, A.M., Diaz-Hernandez, M.B., Ciordia-Ara, M. and D. Rios-Mesa:Chemical Composition of Chestnut Cultivars from Spain. Sci. Horticulturae. 107, 306--314 (2006)
5. Freinkel S., American Chestnut: The Life, Death, and Rebirth of a Perfect Tree. University of California Press, Berkeley (2009)
6. Pena-Mendez, E.M., Hernandez-Suarez, M., Diaz-Romero, C., and Rodriguez-Rodriguez, E.: Characterization of Various Chestnut Cultivars by Means of Chemometrics Approach. Food Chem. 107, 537--544 (2008)
7. Gao H., and Chang X., The Storage and Process of Chestnuts. JinDun Publishing House, BeiJing (2004)
8. Sacchetti, G., Pittia, P., Mastrocola, D. and Pinnavaia, G.G.: Stability and Quality of Traditional and Innovative Chestnut Based Products. In: III International Chestnut Congress, pp. 63--69. ISHS Press, Belgium (2004)
9. Lu W.Z., Yuan H.F., and Xu G.T.: Morden Near Infrared Spectroscopy Analytical Technology. China Petrochemical Press, Beijing (2007)
10. Williams P., and Norris K. eds., Near-Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemist, (2002)
11. Huang, H.B., Yu, H.Y., Xu, H.R. and Ying, Y.B.: Near Infrared Spectroscopy for on/in-line Monitoring of Quality in Foods and Beverages: A Review. J. Food Eng. 87, 303--313 (2008)
12 Yan T., Zhao L., and Han D., Basement and Application of Near Infrared Spectroscopy Analysis. China Light Industry Press, Beijing (2005)
13. Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I. and Lammertyn, J.: Nondestructive Measurement of Fruit and Vegetable Quality by Means of NIR Spectroscopy: A Review. Postharvest Biol. Technol. 46, 99--118 (2007)
14. Ying, Y.B., Liu, Y.D. and Fu, X.P.: Sugar Content Prediction of Apple using Near-infrared Spectroscopy Treated by Wavelet Transform. Spectrosc. Spect. Anal. 26, 63--66 (2006)
15. Tian, H.Q., Ying, Y.B., Lu, H.S., Xu, H.R., Xie, L.J., Fu, X.P. and Yu, H.Y.: Study on Predicting Firmness of Watermelon by Vis/NIR diffuse Transmittance Technique. Spectrosc. Spec. AnaL. 27, 1113--1117 (2007)

16. Han, D.H., Tu, R.L., Lu, C., Liu, X.X. and Wen, Z.H.: Nondestructive Detection of Brown core in the Chinese Pear 'Yali' by Transmission Visible-NIR Spectroscopy. Food Control. 17, 604--608 (2006)

17. Mexis, S.F., Badeka, A.V., Riganakos, K.A., Karakostas, K.X. and Kontominas, M.G.: Effect of Packaging and Storage Conditions on Quality of Shelled Walnuts. Food Control. 20, 743--751(2009)

18. Liu, J., Li, X.Y., Wang, W. Zhang, J., Zhou, Z., Xiao, W. and Zhou W.: Study of Discrimination of Moldy Chestnut Using Near- Infrared Spectroscopy. In: Annual International Meeting of American Society of Agricultural and Biological Engineers, pp. 095715. ASABE Press, Joseph (2009)

19. Tewari, J.C., Dixit, V., Cho, B.K., and Malik, K.A.: Determination of Origin and Sugars of Citrus Fruits using Genetic Algorithm, Correspondence Analysis and Partial Least Square Combined with Fiber Optic NIR Spectroscopy. Spectrochim. Acta. Part A. 71, 1119--1127 (2008)

20. Blakey, R.J., Bower, J.P. and Bertling, I.: Influence of Water and ABA Supply on the Ripening Pattern of Avocado (Persea americana Mill.) Fruit and the Prediction of Water Content using Near Infrared Spectroscopy. Postharvest Biol. Technol. 53, 72--76 (2009)

21. Sundaram, J. Kandala, C.V.K., Butts, C. L. and Windham W.R.: Application of NIR Reflectance Spectroscopy on Determination of Moisture Content of Peanuts: A Non Destructive Analysis Method. In: Annual International Meeting of American Society of Agricultural and Biological Engineers, pp. 095864. ASABE Press, Joseph (2009)

22. Flores-Rojas, K., Sanchez, M.T., Perez-Marin, D., Guerrero, J.E. and Garrido-Varo, A.: Quantitative Assessment of Intact Green Asparagus Quality by Near Infrared Spectroscopy. Postharvest Biol. Technol. 52, 300--306 (2009)

23. Galvao, R.K.H., Araujo, M.C.U., Jose, G.E., Pontes, M.J.C., Silva, E.C. and Saldanha, T.C.B.: A Method for Calibration and Validation Subset Partitioning," Talanta. 67, 736--740 (2005)

24. Isaksson, T. and Naes, T.: The Effect of Multiplicative Scatter Correction (Msc) and Linearity Improvement in Nir Spectroscopy," Appl. Spectrosc. 42, 1273--1284 (1988)

25. Osborne B.G., Fearn T., and Hindle P.H.: Food and Beverage Analysis. Longman, Essex (1993)

26. Walsh, K.B., Golic, M. and Greensill, C.V.: Sorting of Fruit using Near Infrared Spectroscopy: Application to a Range of Fruit and Vegetables for Soluble Solids and Dry Matter Content," J. Near Infrared Spectrosc. 12, 141--148 (2004)

27. Shenk, J.S. and Westerhaus, M.O.: Calibration the ISI Way. In:Davies, A.M.C., Williams, P.C. (Eds.), Near Infrared Spectroscopy: The Future Waves. NIR Publications, Chichester (1996)