Shatang Mandarin Sugar Degree Detection Based on Near Infrared Spectrum

Kun Zhang[1,2,3]

[1]Huizhou Engineering Vocational College, Huizhou, Guangdong, 516001, China

[2]Key Laboratory of Key Technology for South Agricultural Machine and Equipment, South China Agricultural University, Ministry of Education, Guangzhou, 510642, China

[3]College of Engineering of South China Agricultural University, Guangzhou, 510642, China

Keywords: Shatang mandarin, Brix value, Near infrared spectrum, Partial least squares, BP artificial neural network.

Abstract: To measure Brix in Shatang mandarin, FieldSpec 3 porTable spectroradiometer made Inc. (ASD) was used to measure 200 Shatang mandarin's absorbance. The PLS model of Shatang mandarin was built. 8 different popular spectroscopy pretreatment approaches were used to find the suiTable optimal pretreatment methods and the optimal modeling band of the model of Shatang mandarin. Then the BP neural network model of Shatang mandarin was built. The results show the l correlation coefficient R of BP neural network model of Shatang mandarin based on the PLS mode calibration model of Shatang mandarin was 0.967. On validation the correlation coefficient R was 0.868. The mean absolute error was 0.56 Brix. The standard error of the prediction was 0.74.

1. Introduction

Sugar orange, also known as October orange. It is native to the Guangning and Sihui areas and is one of the featured fruits of Guangdong Province. The single fruit weighs 62-86 g, and its edible rate is 71%, containing 11% of soluble solids, 10.55 g of whole sugar per 100 ml, and 0.35 g of fruit acid. Brix is an important indicator for evaluating the quality of fruit. The traditional method of measuring sugar content directly measures the fruit juice with a sugar meter. This method of detection can cause damage to fruits, and is time consuming, inefficient, and prone to operational errors. Therefore, in order to increase the added value of fruits, a non-destructive testing technique is urgently needed to detect the internal quality of fruits. Near-infrared spectroscopy has the characteristics of high speed, no damage to samples, simple operation, good stability and high efficiency. It was first used to determine the moisture of agricultural by-products and has been extended to the detection of fruit quality in recent years. At present, many scholars have studied the estimation of the internal quality indicators of these fruits in the visible/near infrared band: apple, peach, pear, Jiangxi navel orange, seedless tangerine and citrus, etc., and established relevant prediction models.

2. Experimental materials and methods

2.1 Laboratory apparatus

Field Spec 3 porTable ground spectrometer of American ASD (analytical spectral device), its spectral sampling interval is 1.4nm@350-1050nm, 2nm@1000-2500nm. The measurement range is 350-2500 nm, the sampling time is 10 times/s, the spectral resolution is 3 nm@350-1000 nm, 10 nm@1000-2500 nm, and the front field angle is 25°. Wavelength accuracy: +/- 1 nm, wavelength repeatability: +/- 0.02 nm. Sensitivity adjustment: fully automatic. Japan's Aiyi company's PAL-1 digital display handheld sugar meter, measuring range: Brix 0.0 ~ 53.0%, its resolution is ± 0.2 Brix, automatic temperature compensation range is 10-60 ° C.

2.2 Experimental sample

A batch of fresh Sihui sugar oranges was purchased from the market, and 200 samples were randomly selected as experimental samples, 150 samples were used as model samples, and 50 samples were used as prediction samples. The branches and leaves were cut off, the skin of the sugar orange was cleaned, marked in turn, and allowed to stand at room temperature for 24 hours. The sugar orange sample is shown in Figure 1.

Fig.1 Shatang mandarin samples

2.3 Experimental methods

2.3.1. Spectral acquisition

At room temperature, after preheating the ground spectrometer for half an hour, place the granulated orange whole fruit on the blade holder and press it properly. The measurement site was selected at the equator of the sugar orange, and each sample was collected three times at 120° relative to the equator. When collecting spectra, try to avoid obvious scars and spots. The three spectra were averaged using the RS3 standard software included with the Field Spec 3 porTable ground spectrometer as the raw spectral data for the sample. The Original average spectroscopy of Shatang mandarin is shown in Figure 2.

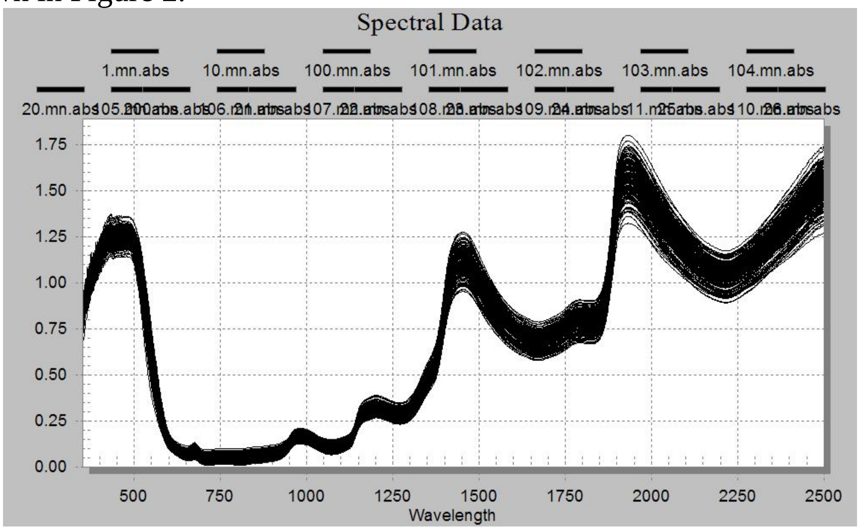

Fig.2 Original average spect roscopy of Shatang mandarin

2.3.2. Brix measurement

At room temperature, the granulated orange samples were placed in a beater in the order of the labels, and a portion of the broken samples were placed on a quantitative filter paper, and the filter paper was placed on a clean beaker. After each sugar orange sample is finished, clean the proofer and dry it with a clean towel. After a certain amount of juice in the beaker, the sugar content was measured. First, the sugar meter is zero-corrected with purified water, and then 2 to 3 drops are used to drop the droplets into the measurement area of the PAL-1 digital display handheld sugar meter. Press the Start button, wait for the reading to stabilize, record the reading of the sugar meter, and then press the Start button again. After the reading is sTable, read the data again, and average the two data read as the sugar value of the sugar orange. The 150 sugar model values of 150 modeling and 50 prediction samples are shown in Table 1.

Table 1. Brix value of 200 sugar oranges

Sample No.	Maximum value(Brix)	Minimum value(Brix)	Average value(Brix)	Standard deviation
150	15.5	10.3	12.8	1.03
50	15.8	10.5	13.5	1.39

3. Data processing and analysis

3.1 Spectral pretreatment

Commonly used spectral data preprocessing methods include data smoothing, baseline correction, vector normalization (VN), multiplication scattering correction (MSC),Standard normal variate transformation (SNV), first derivative (FD), and second derivative (SD).In this study, the spectral pre-processing of the modeled samples was followed by PLS modeling to obtain the best spectral pre-processing method. The results after pretreatment are shown in Table 2.

Table 2. PLS modeling indicators values after different pre-processing of modeling samples

Different pretreatment methods	PC	R	RMSEC
None	10	0.958764	0.268438
Moving average method (3)	10	0.957537	0.272318
Savitzky-Golay	10	0.957052	0.273834
Baseline	9	0.955024	0.280079
VN	9	0.957363	0.272864
MSC	9	0.959778	0.265187
SNV	9	0.959648	0.265605
FD	6	0.945482	0.307608
SD	6	0.941966	0.317086

It can be seen from the Table that after various pretreatments, the multi-scatter correction pretreatment has the best effect, the correlation coefficient reaches 0.959778, and the RMSEC decreases to 0.265187.

3.2 Selection of optimal modeling band

The field spectrum of the sucrose orange sample scanned by the FieldSpec 3 spectrometer has an absorbance of 2150 wavelength points. If 2150 absorbance values are modeled directly, the amount of computation is too large. Therefore, it is necessary to determine the optimal modeling spectral band after selecting the optimal spectral preprocessing method, which not only can reduce the calculation amount, simplify the model, but also can eliminate the irrelevant or nonlinear variables, and obtain strong prediction ability and good robustness. Correction model.

In this study, the interval-taking method is combined with the correlation coefficient method to select the optimal band. First, the points were taken at 780-1100 nm and 1100-2500 nm according to the resolution of the spectrometer. After the treatment, a total of 248 wavelength data were obtained. Then, according to the correlation coefficient method, the correlation coefficient is obtained for the absorbance matrix and the sugar matrix at each wavelength of the sample, and 248 correlation coefficients are obtained, as shown in Fig. 3:

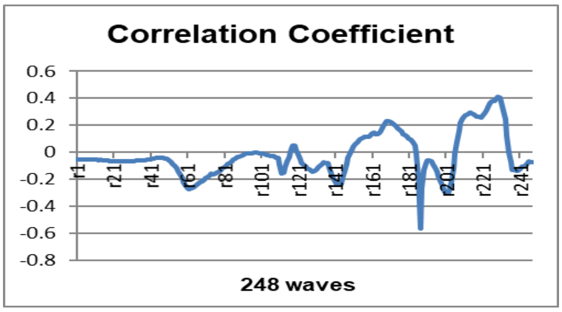

Fig.3 Correlation coefficients of different wavelengths and sugar degree

According to the method of correlation coefficient method, the wavelength is selected in the near red spectral range of 780 to 2500 nm, and the absolute value of the correlation coefficient is in descending order. There is no specific formula for the selection of the wavelength threshold due to the correlation coefficient method. Therefore, the absorption spectrum data modeling of the first 100, 120, and 150 wavelengths was selected in this study. Using the partial least squares modeling, the PLS model parameters of different wavelengths are compared to select the optimal modeling wavelength. The model parameters are shown in Table 3.

Table 3. Preferred PLS model parameters after wavelength

No. of waves	PC	R	RMSEC
100	10	0.961708	0.258872
120	9	0.957483	0.272487
150	7	0.949822	0.295439

From the above Table, it can be seen that when the deviation is the smallest, under the different principal components, when the preferred number of wavelengths is 100, the prediction correlation coefficient of the model is the best, 0.961708, and the corrected root mean square error is 0.258872.

3.3 BP neural network model

Based on the optimal pretreatment method based on the above PLS model and the selection of the optimal band, a BP neural network model of sugar orange sugar is established. The input layer is the absorbance spectrum data of the granulated orange with the best pretreatment and the optimal band; the hidden layer adopts the logig type transfer function; the training function selects the trainscg function. The output layer is the sugar value of a single neuron, ie, sugar orange. The pureline linear transfer function is used. The learning function uses the leanngdm function, and the number of hidden layer neurons is 10. After the model was established, the absorbance spectrum values of 50 samples of sugar-free oranges that were not involved in the training were input into the BP neural network model, and the value of the skin sugar value was predicted. The robustness of the model was analyzed by comparing with the measured values. The regression model of the predicted and measured values in the obtained BP neural network prediction model is shown in Fig. 4.

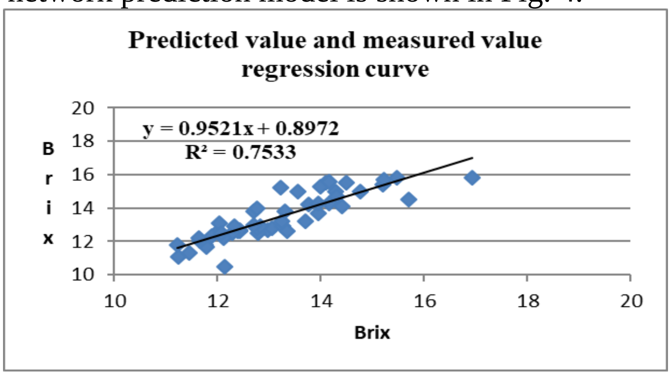

Fig.4 Regression curve of predicted value and measured value of BP neural network

It can be seen from Fig. 4 that for the BP neural network model of sugar orange sugar degree, after

50 sets of sugar orange absorbance data input model of the verification set, the sugar value prediction value of 50 sugar oranges is obtained, and 50 measured by the sugar meter. After the regression analysis is performed on the measured values, the regression equations of the predicted values and the measured values can be obtained.

In order to detect the sugar content of sugar orange, the absorbance spectrum of 200 sugar oranges was measured by FieldSpec3 portable ground spectrometer produced by American ASD Company, and the PLS model of sugar orange sugar was established. Eight common spectral preprocessing methods were compared, the optimal preprocessing method and the optimal modeling band were selected, and the BP neural network model was finally established. The experimental results show that the correlation coefficient R between the predicted value and the measured value of the BP artificial neural network model based on PLS is 0.967, and the correlation coefficient R between the predicted and actual values of the verification model is 0.868. The forecast standard error is 0.74.

4. Summary

In this study, near-infrared spectroscopy and partial least squares (PLS) were used to investigate the effects of eight common spectral pretreatment methods on the PLS quantitative analysis model of sugar orange sugar near-infrared spectrum. By comparing the spectral pretreatment methods, it is determined that the optimal common spectral preprocessing method is multivariate scatter correction. Determine the best modeling band 944 to 1013 nm by interval spacing method and correlation coefficient method (take one data for every three data), 1140 nm, 1150 nm, 1290 to 1330 nm (one data per ten data) 1410 to 1480 nm (per ten The data takes one data) 1630 to 2420 nm (one data per ten data). After optimal pretreatment and optimal band selection, the BP neural network sugar model of sugar orange sugar was established. The correlation coefficient R between the predicted and measured values of the model was 0.967. The correlation coefficient between the predicted and actual values of the model was R. 0.868. The standard error predicted is 0.74. The maximum error between the measured and predicted values is 1.9773 Brix, the minimum error is 0.026 Brix, and the average error is 0.265 Brix. The average absolute error is 0.56 Brix.

References

[1] Lu Wanzhen. Modern near-infrared spectroscopy, Beijing: China Petrochemical Press, 2007: 6-10.

[2] Clark C J, McGlone V A, Jordan R B. Deteetion of Brownheartin ‘Braeburn’ apple by transmission NIR Spectroscopy, Postharvest Biology and Technology, 2003(28): 87-96.

[3] Carlomagno G, Capozzo L, Attolico G, et al. Non-destructive grading of peaches by near-infrared spectrometry, Infrared physics and technology, 2004(46): 23-29.

[4] Hong Tiansheng, Qiao Jun, Ning Wang et al. Nondestructive testing of snow pear quality based on hyperspectral image technology, Transactions of the Chinese Society of Agricultural Engineering, 2007, 23(2): 151-154.

[5] Xia Junfang, Li Xiaowei, Li Peiwu et al. Nondestructive detection of citrus internal quality near-infrared spectroscopy based on wavelet denoising. Journal of Huazhong Agricultural University. 2007, 2(26): 20-123.

[6] Antihus Hernandez Gomez, Yong He, Annia Garcia Pereira. Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR–spectroscopy techniques. Journal of food engineering, 2006(77): 313-319.

[7] Liu Yande, Luo Ji, Ouyang Aiguo. Study on the sugar content of Jiangxi navel orange by visible light spectroscopy. Spectroscopy and spectral analysis, 2007, 3(27): 5 69-572.