Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 14, No. 2, 2025

Rapid and Non-destructive Detection of Sucrose Content in Tea by Near-Infrared Spectroscopy

Jinhui Luo [1], Li Luo [1,*], Xiangyang Yu [2,3], Yuchen Guo [3], Weibin Hong [3]

[1] School of Physics and Optoelectronic Engineering, Guangdong University of Technology, Guangzhou, Guangdong, 510006, China
[2] Nanchang Research Institute, Sun Yat-sen University, Nanchang, Jiangxi, 330096, China
[3] Guangzhou Guangxin Technology Co., Ltd., Guangzhou, Guangdong, 510300, China

*** Corresponding author:** Li Luo

Abstract: This study aims to establish a rapid and non-destructive method for determining sucrose content in green tea using near-infrared (NIR) spectroscopy. A total of 164 representative tea samples were collected from Yiling District, Wufeng County, and Yidu City in Hubei province as well as major tea-producing areas like Yunnan, Sichuan, Anhui, and Jiangxi. The research systematically verified and compared a variety of spectral preprocessing combinations (such as SG smoothing, SNV and D1) and characteristic wavelength selection algorithms (SPA, iPLS, CARS) to build a robust quantitative model. Among the tested models (PLSR, SVR and RFR), the Random Forest Regression (RFR) model demonstrated superior performance. Its key construction steps involved initial SG+SNV preprocessing, followed by characteristic wavelength screening using iPLS. During calibration, the model achieved a correlation coefficient (R^2c) of 0.9656, and a root mean square error (RMSEC) of 1.0316. For validation, the correlation coefficient (R^2p) reached 0.9043 and the root mean square error (RMSEP) was 1.7806. These results demonstrate that the model efficiently enables non-destructive sucrose testing in tea and provides a reliable technical solution for quality control.

Keywords: Green Tea; Sucrose Content; Near-infrared Spectroscopy; Rapid and Non-destructive Detection.

1. Introduction

As a traditional beverage with its profound cultural heritage and unique sensory characteristics, tea has long held a prominent and enduring position in both domestic and international markets. With increasing consumer demands for higher quality, particularly in terms of safety and stability, processing quality has become a key priority in the tea industry. In order to improve flavor of tea soup and reduce bitterness [1], a small amount of sucrose is often added during processing. However, the unethical addition of excessive sucrose for increased profit not only destroys the appearance and flavor of the final product but also interferes with objective sensory evaluation. Traditional methods for determining sucrose content in tea, such as high-performance liquid chromatography (HPLC) [1], are often time-consuming, labor-intensive, and require sophisticated equipment. This makes them inadequate to meet the demands of rapid, standardized, and automated production. Therefore, the development of rapid, accurate and efficient testing technologies is crucial for implementing effective quality control throughout the entire tea production chain.

Near-infrared spectroscopy (NIRS) is located in the electromagnetic spectrum just beyond the red end of the visible light region and within the shortwave region of the infrared spectrum, and was the first non-visible light region discovered in absorption spectroscopy. The American Society for Testing and Materials (ASTM) [2] provides a clear definition of this technology. The wavelength range spans 780 to 2526 nm (12820 to 3959 cm^{-1}). Since the mid-1980s, NIRS has rapidly evolved into a powerful tool for both qualitative and quantitative analysis, fueled by advances in chemometrics and experience gained from mid-infrared spectroscopy. NIRS is valued for its speed and efficiency, requiring minimal sample preparation and causing no sample damage. It enables simultaneous analysis of multiple components and can be applied directly to intact samples, such as whole fruits [3-5] and grain kernels [6-7]. Consequently, NIRS has been used successfully in many fields, including agriculture, food, pharmaceuticals, petrochemicals, tobacco, and textiles [8-14].

Many studies have demonstrated that NIRS is an effective non-destructive method for analyzing the chemical composition of tea. Schulz et al. [15] developed a quantitative model to predict the contents of polyphenols and caffeine by NIRS combined with partial least squares regression (PLS-R). The determination coefficient (R^2p) of both component prediction sets exceeds 0.85. Huang et al. [16] proposed a hybrid method combining ant colony optimization (ACO) with the interval partial least squares method (iPLS) to improve model performance. They demonstrated that the hybrid model can more effectively screen characteristic spectral intervals, resulting in a calibration model that is both more accurate and more stable. Luypaert et al. [17] expanded the application scope of NIRS by developing a NIRS-PLS model to determine the total antioxidant capacity of tea. The study demonstrated that NIRS is highly flexible in detecting plant bioactive compounds. Shi et al. [18] further successfully applied NIRS to the simultaneous detection of multiple components. Using this technology, the contents of water, free amino acids, caffeine, total catechins, polyphenols, and water extracts in Guizhou green tea were rapidly determined. The research team used the partial least squares (PLS) method combined with multivariate linear regression to construct a quantitative calibration model, and evaluated the quality of the model through cross-verification of correlation coefficients, root mean square error (RMSE), and predicted root mean square error (RMSEP). The results demonstrated that this method is easy to operate, highly accurate, and provides an effective approach for the rapid analysis of key

components in tea. Liu et al. [19] investigated black tea, exploring the application of NIRS for rapid non-destructive quantification of Theaflavins (TFs), Thearubigins (TRs) and Theabrownins (TBs). To achieve this, a total of 240 samples (180 for calibration and 60 for prediction) were analyzed in the study, and the best band range and pretreatment method were determined using OPUS 7.0 software. The prediction model exhibited high accuracy, indicating its potential for the quantitative analysis of theanine. Wang and Lu [20] studied Pu'er tea, exploring a method that combines NIRS with partial least squares regression (PLSR) for rapid determination of tea polysaccharide content, employing standard normal variable (SNV) pretreatment. The cross-verification root mean square error of the developed model was 0.0822, the predicted root mean square error was 0.1264, and the predicted set correlation coefficient reached 0.8217, indicating that the model performs well in quality control and active ingredient analysis. Similarly, Shen et al. [21] investigated white tea, using NIRS to achieve rapid quantitative analysis of total soluble sugar. Their partial least squares (PLS) model yielded a correlation coefficient (R) of 0.963, a root mean square error of calibration (RMSEC) of 0.363, a root mean square error of prediction (RMSEP) of 0.349, and an average relative error of the validation set of 3.11%. The results confirm that the method has the advantages of high precision, fast and non-destructive nature in this specific application scenario.

In summary, NIRS has emerged as a reliable tool for tea composition analysis, offering the advantage of non-contact and non-destructive analysis. Compared to traditional chemical methods, NIRS presents significant benefits for the quantitative analysis of tea ingredients. However, there are relatively few studies on the detection of sucrose content in tea. To address this gap, this study aims to develop a method for the rapid and non-destructive detection of sucrose content in tea leaves based on NIRS.

2. Material and Methods

2.1. Samples and Instrumentation

In this study, 164 intact green tea samples, representing diverse varieties, were collected from various regions across China, including Yiling District, Wufeng County, and Yidu City in Hubei Province, as well as other major tea-producing provinces such as Yunnan, Sichuan, Anhui and Jiangxi. Spectral data was collected using the GX-1-S1325-S200 portable near-infrared spectrometer produced by Guangzhou Guangxin Technology Co., Ltd, which works in diffuse reflection mode with a wavelength range of 1350-2550 nm.

2.2. Data Acquisition and Preprocessing

The experiment began with powering on the spectrometer, followed by a 30-minute warm-up period to stabilize its detection performance. A white reference calibration was then performed using the instrument's internal calibration tile. With the calibration complete, five replicate measurements were taken for each tea sample. After each measurement, the sample was returned to its original bag to maintain consistency. Each of these five portions was placed in a separate quartz dish, about two-thirds full. Then, a near-infrared spectrum was acquired for each replicate in diffuse reflectance mode. The resulting five spectra for each tea sample were then averaged to create one representative spectrum used in all subsequent analyses. The experimental setup for NIR data collection is depicted in Figure 1. The reference sucrose content was measured according to Method 3 of the Chinese National Standard GB 5009.8-2023. The measured sucrose content across all samples ranged from 1.81% to 29.1%.

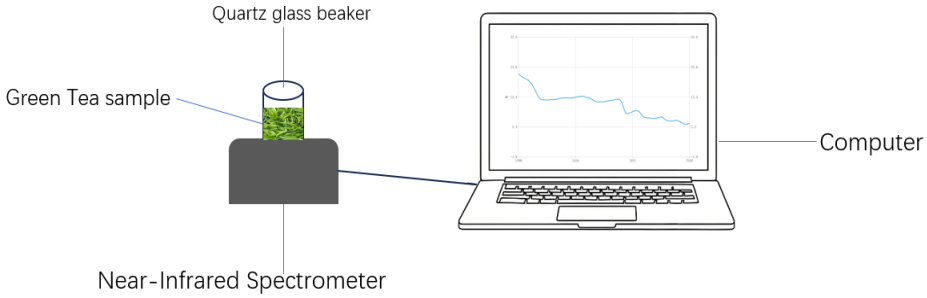

Figure 1. Schematic diagram of the experimental setup for spectral acquisition from tea samples.

Raw spectral data are often affected by noise, light scattering, and baseline drift, stemming from inherent sample variability, instrument factors and environmental fluctuations. These issues can compromise the prediction accuracy of the calibration model [22]. Therefore, spectral pretreatment must be carried out before model development to offset these adverse effects and improve the stability of the model. This research investigates and compares a variety of pretreatment technologies and their combinations, including Savitzky-Golay (SG) smoothing method, standard normal variable (SNV) conversion and first-order derivative (D1) processing. SG smoothing reduces random noise by local polynomial fitting of spectral data; while SNV correction can effectively eliminate baseline offset and scattering effects related to variations in light path and particle size. D1 treatment helps to eliminate baseline drift, separate overlapping absorption peaks, and improve the overall spectral resolution.

2.3. Feature Wavelength Extraction

Characteristic wavelength selection is used to reduce the dimensionality of spectral data. This step not only improves model stability and accuracy but also reduces the computing workload. In order to build a sucrose prediction model, various feature selection methods were compared to determine the optimal scheme. The methods evaluated included recursive projection algorithm (SPA), interval partial least squares (iPLS), and competitive adaptive resampling (CARS). As a forward variable screening method, SPA identifies the variable set containing the least redundant information through vector projection analysis [23], effectively reducing collinearity, singularity and instability among spectral variables. By reducing collinearity and streamlining the number of variables required for modeling, SPA simplifies the model structure, accelerates the development process and improves efficiency [24-27]. The

iPLS method is a wavelength interval selection method proposed by Lars Norgaard in 2000 [28]. The algorithm initially divides the complete near-infrared spectrum into multiple equal intervals, and then constructs a PLS regression model for each interval. The interval with the lowest cross-validation root mean square error (RMSECV) is selected as the optimal interval, which is then refined by iteratively expanding or reducing the interval to optimize the final spectral region. Zhu Xiaobo et al. have applied iPLS to select spectral regions in the near-infrared spectral analysis of agricultural products [29]. CARS, based on adaptive weight sampling, iteratively prioritizes wavelengths with large absolute regression coefficients in the PLS model, gradually eliminates features with lower weights, and finally selects the best variable subset, i.e., the model with the smallest RMSECV at each step [30-32].

2.4. Data Division and Model Evaluation Methods

The 164 samples were divided into a calibration set and a validation set at a 4:1 ratio, using the SPXY algorithm [33], resulting in 131 samples for the calibration set and 33 samples for the validation set. Three regression models, including Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), and Random Forest Regression (RFR), were then developed. To optimize these models, a cross-validated grid search was employed to find the best hyperparameters for each. Model performance was evaluated using the coefficient of determination (R^2) and root mean square error (RMSE), calculated for both the calibration (R^2c and RMSEC) and validation (R^2p and RMSEP) sets. Higher R^2 value closer to 1 indicate stronger correlation and better agreement between actual and predicted sucrose content, while lower RMSEC and RMSEP values closer to 0 signify greater model accuracy. The relevant calculation formulas are as follows:

$$R_C^2 = 1 - \frac{\sum_{t=1}^{n_c}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n_c}(y_t - \bar{y}_c)^2} \quad (1)$$

$$R_p^2 = 1 - \frac{\sum_{t=1}^{n_p}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n_p}(y_t - \bar{y}_p)^2} \quad (2)$$

$$\text{RMSEC} = \sqrt{\frac{1}{n_c}\sum_{t=1}^{n_c}(y_t - \hat{y}_t)^2} \quad (3)$$

$$\text{RMSEP} = \sqrt{\frac{1}{n_p}\sum_{t=1}^{n_p}(y_t - \hat{y}_t)^2} \quad (4)$$

In equations (1) to (4), n_c and n_p represent the number of samples in the calibration and validation set, respectively; $\hat{y}_t$ is the predicted value for the *t*-th sample; while y_t is the corresponding actual measured value. $\bar{y}_c$ and $\bar{y}_p$ denote the means of the actual values in the calibration and validation sets, respectively. All data processing in this study was performed using PyCharm 2024.1.4.

3. Results and Discussion

3.1. Spectral Data Acquisition and Wavelength Selection

Near-infrared spectra were acquired for all 164 green tea samples, following the procedure outlined in Section 2.2. The raw reflectance spectra are shown in Figure 2. Visual inspection of the original spectral curves revealed significant noise at both the beginning and end of the spectral range. To minimize the impact of this noise on model performance, the noisy regions were excluded, and only spectral data between 1400–2450 nm was retained for subsequent processing and analysis.

3.2. Pretreatment Methods

The original near-infrared spectrum is highly susceptible to interference from various factors, including the sample's physical properties (e.g., particle size, surface scattering), instrument characteristics, and environmental conditions. These factors introduce noise, baseline drift and scattering effects, which can compromise the accuracy and robustness of quantitative models [34]. Therefore, preprocessing the raw spectral data is crucial to mitigate these adverse effects. The resulting spectra after applying various pretreatment techniques are illustrated in Figures 3(a), (b), (c), and (d).

3.3. Feature Wavelength Selection

Near-infrared spectral data typically contains hundreds to thousands of wavelength variables, leading to significant data redundancy and increased susceptibility to noise interference. Employing all variables in modeling increases the risk of overfittings and reduces model stability. Therefore, variable selection methods are necessary to eliminate redundant and irrelevant wavelengths, thereby improving model prediction accuracy, robustness, and computational efficiency. In this

study, characteristic wavelength selection was performed on the preprocessed spectral data using the SPA, iPLS, and CARS algorithms. The selected wavelengths were subsequently used for model development. The results are summarized in Table 1.

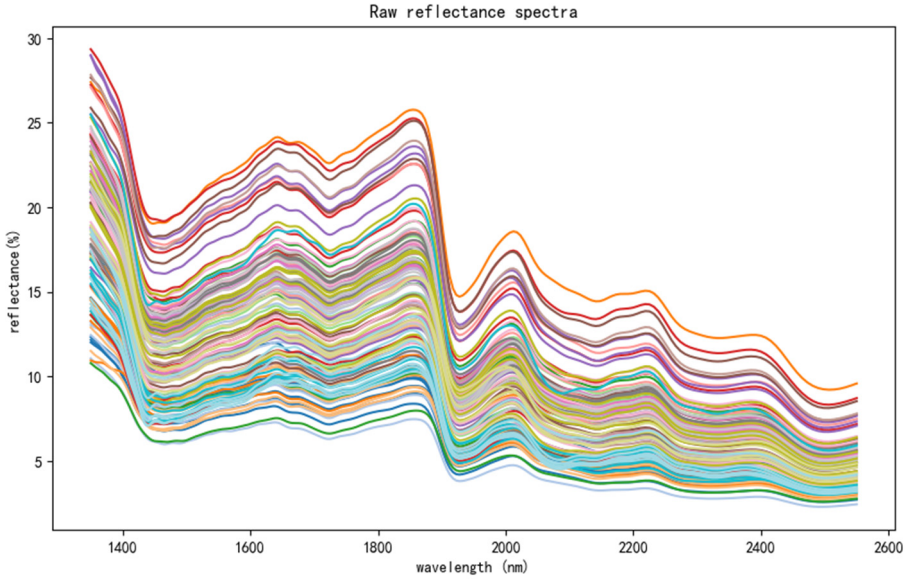

Figure 2. Near-infrared (NIR) reflection spectra of green tea samples

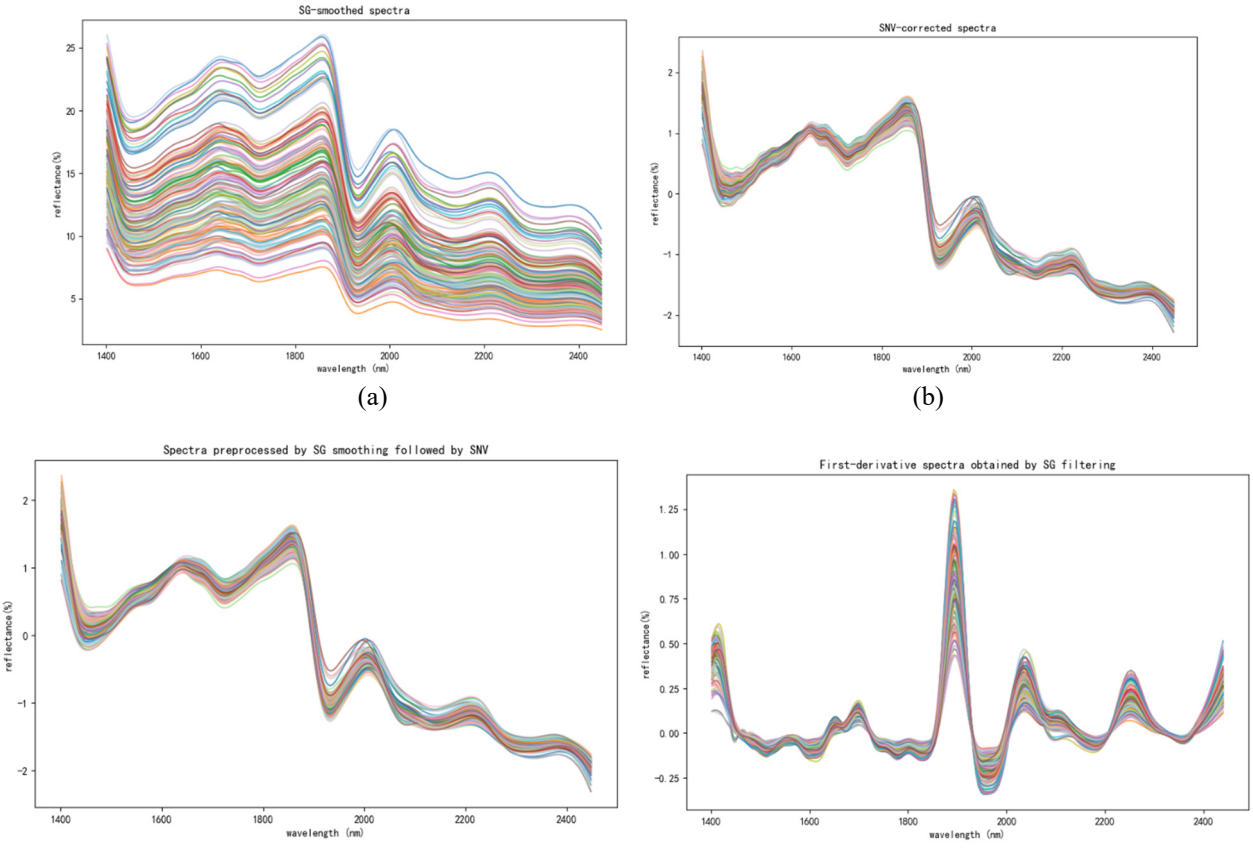

Figure 3. Spectra processed by different pretreatment methods.

3.4. Model Optimization

To develop a precise prediction model, the raw spectra were first preprocessed, followed by the application of feature selection algorithms: SPA, iPLS, and CARS. Three models–PLSR, SVR, and RFR–were then developed to predict sucrose content in green tea. Their performances are summarized in Table 2.

As summarized in Table 2, the Random Forest Regression (RFR) model exhibited the best predictive performance. This model was developed using preprocessed raw spectra, with a combination of SG smoothing and SNV, followed by feature wavelength selection using iPLS. On the calibration set, this model achieved a coefficient of determination (R^2c) of 0.9656 and a root mean square error of calibration (RMSEC) of 1.0316. For the independent validation set, it reached an R^2p of 0.9043 and an RMSEP of 1.7806. This model was then used to predict the sucrose content of the validation set

samples. A visual comparison of the predicted and actual measured values is presented in Figure 4, illustrating the model's predictive capability.

Table 1. Characteristic wavelengths selected by different variable selection methods.

Wavelength Selection	Pretreatment	Feature Wavelength (nm)
	Raw	1437.18, 1593.06, 1715.89, 2080.45, 2141.11, 2212.05, 2259.69, 2369.03, 2448.00
	SG	1440.00, 1460.04, 1575.97, 1665.30, 1887.92, 1927.56, 1990.24
SPA	SNV	1401.53, 1451.38, 1661.54, 1711.89, 2074.58, 2205.41, 2316.72
	SG+SNV	1401.53, 1508.01, 1607.00, 1650.34, 1711.89, 1990.24, 2034.35, 2134.88, 2198.80, 2252.76
	SG+D1	1404.21, 1451.38, 1932.63, 2006.56
	Raw	1401.53~1440.00, 1692.17~1748.57, 1963.64~2448.00
	SG	1401.53~1440.00, 1635.63~1813.33, 2045.68~2225.45
iPLS	SNV	1692.17~1748.57, 1963.64~2225.45
	SG+SNV	1692.17~1748.57, 1963.64~2225.45
	SG+D1	1404.21~1440.00, 1486.64~1632.00, 1692.17~1813.33, 2134.88~2225.45
	Raw	1401.53, 1434.38, 1442.83, 1457.14, 1495.72, 1552.64, 1632.00, 1657.79, 1907.53, 1995.65
	SG	1401.53, 1431.58, 1454.26, 1553.19, 1639.29, 1642.95, 1669.09, 1826.87, 1912.50, 1995.65
CARS	SNV	1401.53, 1477.67, 1530.00, 1657.79, 1715.89, 1979.51, 2080.45, 2092.31, 2110.34, 2198.80
	SG+SNV	1406.90, 1498.78, 1552.64, 1669.09, 1688.28, 1711.89, 1791.22, 1748.57, 1963.64, 2080.45
	SG+D1	1483.64, 1465.87, 1610.53, 1728.00, 2006.56, 2068.73, 2166.37, 2212.05, 2239.02, 2392.18

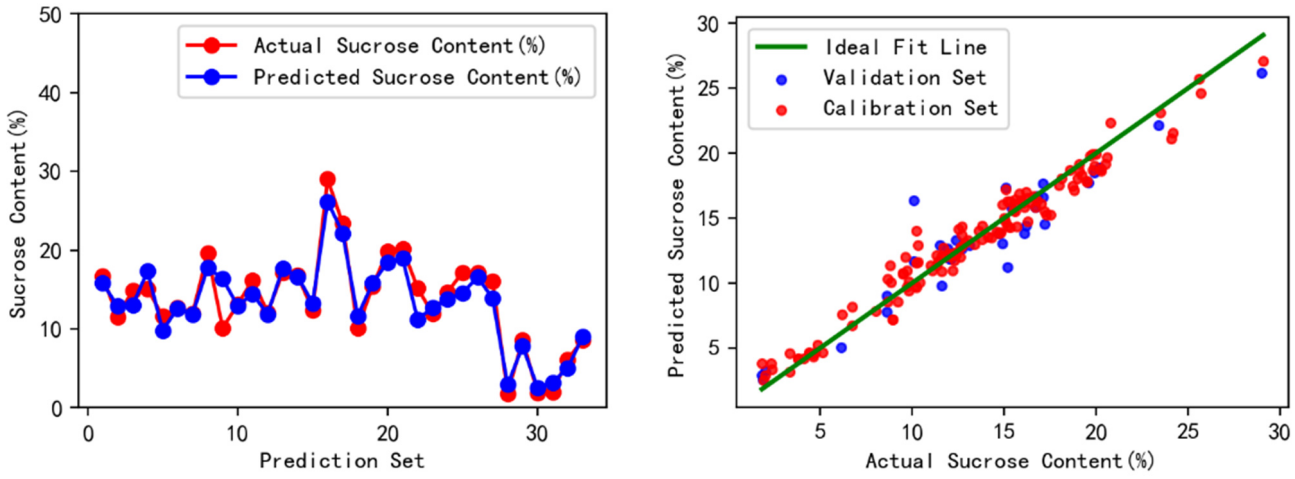

Figure 4. Prediction results of the random forest regression (RFR) model

4. Conclusion

Raw reflectance spectra exhibited significant noise interference, particularly at the beginning and end of the spectral range. To minimize the potential errors in subsequent model development, a spectral truncation method was applied, retaining data only within the 1400–2450 nm wavelength band. Beyond this truncation, the original spectra were subjected to various preprocessing combinations, including Savitzky-Golay (SG) smoothing, standard normal variate (SNV) transformation, and first-order derivative (D1) to reduce the impact of noise, baseline drift, and scattering on model accuracy. To further improve model accuracy, three regression models – PLSR, SVR, and RFR – were compared. All models are constructed after characteristic wavelength screening, which utilized three independent algorithms: SPA, iPLS and CARS. The random forest regression model exhibited the strongest prediction performance. The model adopts the original spectrum preprocessed by SG+SNV and is combined with iPLS for characteristic wavelength selection construction. This model achieved a correlation coefficient (R^2c) of 0.9656 and a root mean square error of calibration (RMSEC) of 1.0316 on the calibration set, and a correlation coefficient (R^2p) of 0.9043 and a root mean square error of validation (RMSEP) of 1.7806 on the validation dataset. These results demonstrate the model's capability for rapid and non-destructively determination of sucrose content in tea. In addition, the research makes two key contributions: first, a new method for the objective quantification of tea quality is presented; and second, the potential for advancing intelligent tea quality control, precision processing, and fair trade is highlighted. Future research will be committed to improve model universality and generalization by expanding the scope of sample types and geographical sources.

Table 2. Predictive performance of different models for sucrose content in green tea.

Model	Modeling Method	Variables	Calibration Set			Validation Set	
			R^2c	RMSEC		R^2p	RMSEP
	SG+SPA	7	0.6728	3.1828		0.7665	2.7820
	SNV+SPA	7	0.7664	2.6893		0.8318	2.3610
	SG+SNV+SPA	10	0.7770	2.6280		0.8205	2.4389
	SG+D1+SPA	4	0.5735	3.6340		0.7762	2.7238
	SG+iPLS	90	0.7624	2.7126		0.6772	3.2711
PLSR	SNV+iPLS	60	0.7880	2.5624		0.8410	2.2981
	SG+SNV+iPLS	60	0.7889	2.5569		0.8299	2.3744
	SG+D1+iPLS	104	0.7479	2.7939		0.7022	3.1418
	SG+CARS	10	0.7459	2.8049		0.7159	3.0684
	SNV+CARS	10	0.8109	2.4198		0.7742	2.7356
	SG+SNV+CARS	10	0.7966	2.5097		07930	2.6195
	SG+D1+CARS	10	0.7505	2.7793		0.7442	2.9118
	SG+SPA	7	0.6761	3.1669		0.7410	2.9300
	SNV+SPA	7	0.8091	2.4310		0.8045	2.5454
	SG+SNV+SPA	10	0.7948	2.5206		0.7573	2.8360
	SG+D1+SPA	4	0.6561	3.2633		0.8021	2.5611
	SG+iPLS	90	0.7570	2.7429		0.6485	3.4135
SVR	SNV+iPLS	60	0.7964	2.5108		0.8006	2.5706
	SG+SNV+iPLS	60	0.7952	2.5183		0.7979	2.5880
	SG+D1+iPLS	104	0.7999	2.4884		0.7784	2.7102
	SG+CARS	10	0.7305	2.8890		0.6966	3.1714
	SNV+CARS	10	0.8183	2.3718		0.8082	2.5211
	SG+SNV+CAR	10	0.8107	2.4212		0.8046	2.5452
	SG+D1+CARS	10	0.8006	2.4849		0.8199	2.4435
	SG+SPA	7	0.6687	3.2031		0.4956	4.0887
	SNV+SPA	7	0.9611	1.0978		0.9037	1.7862
	SG+SNV+SPA	10	0.9542	1.1903		0.8349	2.3395
	SG+D1+SPA	4	0.9009	1.7515		0.8520	2.2145
	SG+iPLS	90	0.9100	1.6698		0.5316	3.9401
RFR	SNV+iPLS	60	0.9660	1.0256		0.8980	1.8391
	SG+SNV+iPLS	60	0.9656	1.0316		0.9043	1.7806
	SG+D1+iPLS	104	0.9512	1.2292		0.7953	2.6047
	SG+CARS	10	0.7407	2.8336		0.5379	3.9136
	SNV+CARS	10	0.9656	1.0321		0.8991	1.8283
	SG+SNV+CAR	10	0.9675	1.0039		0.8647	2.1178
	SG+D1+CARS	10	0.9530	1.2067		0.8083	2.5208

References

[1] Chen Qi, Li Li, Wu Xueyuan, et al, High performance liquid chromatography-evaporative light detector for detection of carbohydrates doped in tea[J]. J. of Tea Business, 2015, 37(1):26-29.

[2] Shou-he Yan. Review: Evaluation of the composition and sensory properties of tea using near infrared spectroscopy and principal component analysis [J]. Journal of Near Infrared Spectroscopy, 2005, 13(6):313-325.

[3] Liu Yande, Ying Yibin. Determination of sugar content and effective acidity in honey peaches by near-infrared spectroscopy[J]. Transactions of the Chinese Society of Agricultural Engineering. 2004, 20(1): 189-192.

[4] Liu Yande, Gao Rongjie, Sun Xudong. Research progress on portable near-infrared detectors for internal quality of fruits[J]. Spectroscopy and Spectral Analysis. 2010, 30(10): 2874-2878.

[5] Wang Jiahua, Pan Lu, Li Pengfei, et al. Temperature compensation for near-infrared spectroscopy analysis models of apple sugar content[J]. Spectroscopy and Spectral Analysis. 2009, 29(6): 1517-1520.

[6] An Ling, Zhang Wujiu, Zhao Wushan. Study on early-generation screening technology for malting barley—Evaluating barley quality using near-infrared transmission spectroscopy[J]. Food and Fermentation Industries. 2001, 27(3): 32-33.

[7] Wu Jianguo, Shi Chunhai, Zhang Xiaoming, et al. Study on the analysis of three essential amino acid contents in rice using near-infrared reflectance spectroscopy[J]. Acta Agronomica Sinica, 2003, 29(5): 688-692.

[8] Sun Tong, Xu Huirong, Ying Yibin. Research progress on the application of near-infrared spectroscopy analysis technology in online non-destructive detection of agricultural product/food quality[J]. Spectroscopy and Spectral Analysis. 2009, 29(1): 122-126.

[9] Campbell M R, Mannis S R, Port H A, et al. Prediction of starch amylose content versus total grain amylose content in corn by near-infrared transmittance spectroscopy[J]. Cereal Chem. 1999,76(4): 552-557.

[10] Lin Tao, Yu Haiyan, Ying Yibin. Research progress on the application of visible/near-infrared spectroscopy in the detection of liquid foods[J]. Spectroscopy and Spectral Analysis. 2008, 28(2): 285-290.

[11] Li Yanyan, Bian Baolin, Yang Jian. Quantitative analysis of major components in the extract of Gegan-Danshen herb-pair by near-infrared spectroscopy[J]. China Journal of Chinese Materia Medica. 2008, 33(17): 2194-2195.

[12] Xu Yupeng, Yuan Hongfu, Lu Wanzhen. Interpreting near-infrared spectra of petroleum products using correlation analysis[J]. Modern Scientific Instruments. 2002.

[13] Qian Ping, Sun Guoqin, Zhang Cunzhou. A new method for quantitative analysis of petroleum components based on near-infrared spectroscopy technology[J]. Spectroscopy and Spectral Analysis. 2008, 28(12): 2851-2854.

[14] Tao Shuai, Ma Xiang, Li Junhui, et al. Study on projection method based on principal component data of tobacco near-infrared spectra and its application in re-drying formula[J]. Spectroscopy and Spectral Analysis. 2009, 29(11): 2970-2974.

[15] SCHULZ H, ENGELHARDT U H, WEGENT A, et al. Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic substances in green tea leaves[J]. Journal of agricultural and food chemistry, 1999, 47(12):5064-5067.

[16] Huang Xiaowei, Zou Xiaobo, Zhao Jiewen, et al. Determination of anthocyanin content in scented tea by near-infrared spectroscopy combined with ant colony algorithm[J]. Journal of Jiangsu University: Natural Science Edition, 2014, 35(2): 165-170, 188.

[17] LUYPAERT J, ZHANG M H, MASSART D L. Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea, Camellia sinensis (L.) [J]. Analytica chemical acta, 2003, 478(2):303-312.

[18] Shi Yanmei, Wu Qing, Tan Gaohao, et al. Rapid determination of six major components in Qian tea by near-infrared spectroscopy[J]. Chemical Sciences, 2016, 34(4): 74-77.

[19] Liu Honglin. Determination of tea pigment content in Congou black tea based on near-infrared spectroscopy (NIRS)[J]. Journal of Food Safety & Quality, 2016, 7(12): 4874-4878.

[20] Wang Xi, Lu Hongmei. Rapid determination of tea polysaccharide content in Pu-erh tea by NIR spectroscopy combined with PLSR[J]. Spectroscopy and Spectral Analysis, 2017, 37(18): 5687-5692.

[21] Shen Shiyu, Sun Weijiang, Tang Qin, et al. Study on the rapid determination of total soluble sugar in white tea based on near-infrared spectroscopy[J]. Natural Product Research and Development, 2019, 31(1): 16-23.

[22] Zhuang Xingang. Applied study of near-infrared spectroscopy and incidence optical system of new spectral sensing node[D]. Jinan: Shangdong University, 2017.

[23] Araújo M C U, Saldanha T C B, Galváo R K H, et al. Chemometrics and Intelligent Laboratory Systems, 2001, 57 (2): 65.

[24] LIAO Y T, FAN Y X, CHENG F, et al. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(S1):379.

[25] LIU F, HE Y. Food Chemistry, 2009, 115(4):1430.

[26] Moreira E D T, Pontes M J C, Galvào R K H, et al. Talanta, 2009, 79(5):1264.

[27] LIN Z D, WANG Y B, WANG R J, et al, Chinese Journal of Luminescence, 2016, 37(11):1429.

[28] Norgaard L, Saudland A and Wagner J, et al. Interval partial least squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. Applied Spectroscopy, 2000, 54:413-419.

[29] Zou Xiaobo, Zhu Zeng, Zhao Jiewen. 2007. Wavelength region selection method for NIR spectra of agricultural products based on interval partial least squares[J]. Modern Scientific Instruments, (1): 86-88.

[30] CHEN Yu, QIU Zhijun, ZHANG Bin, Journal of Instrumental Analysis, 2021,40(12):1690.

[31] ZHANG Xuhui, ZHANG Kaixin, ZHANG Chao, et al. Journal of Xi'an University of science and Technology, 2020, 40(5): 760.

[32] TANG Haitao, MENG Xiangtian, SU Xunxin, et al. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(2):105.

[33] CHU X l. Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Applications. Beijing: Chemical Industry Press, 2011:31.

[34] GAO R Q, FAN S F, YAN YL, et al. Spectroscopy and Spectral Analysis, 2004, 24(12):76.