

RESEARCH

Open Access



Rapid monitoring of flavonoid content in sweet tea (*Lithocarpus litseifolius* (Hance) Chun) leaves using NIR spectroscopy

Zhaoxia Tian^{1,2}, Zifeng Tan¹, Yanjie Li¹ and Zhiling Yang^{1*}

Abstract

Background: Sweet tea, which functions as tea, sugar and medicine, was listed as a new food resource in 2017. Flavonoids are the main medicinal components in sweet tea and have significant pharmacological activities. Therefore, the quality of sweet tea is related to the content of flavonoids. Flavonoid content in plants is normally determined by time-consuming and expensive chemical analyses. The aim of this study was to develop a methodology to measure three constituents of flavonoids, namely, total flavonoids, phloridin and trilobatin, in sweet tea leaves using near-infrared spectroscopy (NIR).

Results: In this study, we demonstrated that the combination of principal component analysis (PCA) and NIR spectroscopy can distinguish sweet tea from different locations. In addition, different spectral preprocessing methods are used to establish partial least squares (PLS) models between spectral information and the content of the three constituents. The best total flavonoid prediction model was obtained with NIR spectra preprocessed with Savitzky–Golay combined with second derivatives (SG + D2) ($R_p^2 = 0.893$, and $RMSEP = 0.131$). For trilobatin, the model with the best performance was developed with raw NIR spectra ($R_p^2 = 0.902$, and $RMSEP = 2.993$), and for phloridin, the best model was obtained with NIR spectra preprocessed with standard normal variate (SNV) ($R_p^2 = 0.818$, and $RMSEP = 1.085$). The coefficients of determination for all calibration sets, validation sets and prediction sets of the best PLS models were higher than 0.967, 0.858 and 0.818, respectively.

Conclusions: The conclusion indicated that NIR spectroscopy has the ability to determine the flavonoid content of sweet tea quickly and conveniently.

Keywords: Sweet tea, Flavonoids content, Near-infrared (NIR) spectroscopy, Partial least squares (PLS) model, Model calibration

Background

Lithocarpus litseifolius (Hance) Chun, which has abundant wild resources in the mountainous region of southern China, is commonly named as “sweet tea” [36]. Sweet tea has been listed as a new food resource by the National Health and Family Planning Commission of China since

2017 [9]. Its leaves, which function as tea, sugar and medicine, have been used as a traditional herbal medicine and tea, and it has the potential to become a functional food and food additive [26]. In sweet tea, the most abundant compounds are polyphenols, with flavonoids are the primary compounds in it [18]. Dihydrochalcones are the main flavonoids in sweet tea entailing trilobatin, phloridin, and other compounds. Flavonoids are the main medicinal ingredients in the sweet tea, which has extensive pharmacological activities, such as anti-diabetes effects, antitumor activity, anticancer activity,

*Correspondence: zlyang0002@126.com

¹ Research Institute of Subtropical Forestry, Chinese Academy of Forestry, No. 73, Daqiao Road, Fuyang, Hangzhou 311400, Zhejiang Province, China
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

antimicrobial activity, anti-inflammatory activity, hepatoprotective activity, antioxidant effects and so on [15, 26, 33].

The principal method to determine the flavonoids content in sweet tea is high-performance liquid chromatography [16]. Although chemical methods can accurately detect the composition and content of substances in sweet tea, their steps are time-consuming, complex and expensive. Thus, these characteristics make the process difficult to popularize and cannot be applied to the rapid detection of flavonoids in sweet tea. Therefore, it is important to establish a cheap, simple and reliable analytical approach for detecting flavonoids content in sweet tea leaves rapidly.

NIR spectroscopy, which uses reflectance in a wavelength range of 700–2500 nm, provides a rapid, convenient, and economical way to determine plant ingredients [29]. NIR spectroscopy is a sensing technology that is largely absorbed by O–H, N–H, R–OH, COOH and C–H bonds, which are the major constituents of organic compounds [32]. The principle of NIR spectroscopy is that the ratio and combination of energy absorbed are different due to the stretching and vibration of different bonds between samples [22]. Therefore, the combinations of peaks and troughs in the NIR spectrum vary from sample to sample because NIR radiation interacts differently with different bonds [23]. Thus, the characteristic information of hydrogen-containing groups in organic molecules can be obtained by scanning the near-infrared spectra of samples [25]. It has been widely used in agriculture, [12, 17, 19], petrochemicals [1, 14], food [28] and pharmaceuticals [6]. The use of NIR spectroscopy in plant leaf tissue analysis started in the mid-1970s [20], and there are an increasing number of papers on NIR spectroscopy application on leaf tissues [4, 13]. Many studies using NIR spectroscopy combined with multivariate calibration methods have analyzed the content of polyphenols and other ingredients in tea [5, 24].

These studies reveal that NIR spectroscopy has the potential to determine compositions in tea. Therefore, in this study, the feasibility of rapidly measuring the flavonoids in sweet tea using NIR spectroscopy was studied. At the same time, to improve the accuracy of the model, different preprocessing methods in combination with PLS were used and compared to determine the most desirable calibration model. NIR spectroscopy is expected to provide a rapid method to measure the flavonoids in sweet tea.

Methods

Sweet tea leaf sample preparation

Leaves of sweet tea were collected from different locations, covering major provinces in southern China,

including Hunan, Zhejiang, Guangxi, Sichuan, Jiangxi and Fujian. These areas vary in location, soil characteristics, etc. A total of 213 samples were collected to expand the range of variation in the data. The details of different locations of sweet tea leaf samples are outlined in Table 1. Sweet tea leaves were air-dried at the collection sites. Experimental data were derived from three independent repeats for each sample.

Chemical measurements

Measurement of total flavonoids

The total flavonoid content was measured by a spectrophotometric method [35]. The principle of this method was that under certain conditions, flavonoids react with aluminum to form a red complex, and the color intensity was proportional to the content of flavonoids. Rutin was used as a standard for colorimetry. The absorbance of the standard was measured, and a standard curve was generated. The standard curve was used to calculate the total flavonoid content of sweet tea leaf samples. Briefly, the dried leaves were first ground into powder with a mill. Then, 0.5–1 g of sweet tea leaf powder was weighed and placed into a 250 mL conical flask. A 30 mL extraction solution was used to extract total flavonoids from the powder under supersonic (300 W, Kedao Ultrasonic Instrument Co., Ltd., Shanghai, China) conditions for 1 h. Then, the sample was filtered into a 50 mL volumetric flask while it was hot, and the extraction solution was

Table 1 Locations of sweet tea leaf samples

| Location | Georeference | Sample number |
|---------------------|------------------------------------|---------------|
| Lushan, Sichuan | 103°52′–103°11′E 30°01′–30°49′N | 23 |
| Jinyunshan, Sichuan | 106°17′–106°24′E 29°41′–29°52′N | 16 |
| Pidu, Sichuan | 103°42′–104°02′E 30°43′–30°52′N | 4 |
| Pengzhou, Sichuan | 103°10′–103°40′E 30°54′–31°26′N | 9 |
| Xupu, Hunan | 110°15′–111°01′E 27°19′–28°17′N | 58 |
| Jiangshan, Zhejiang | 118°22′–118°48′E 28°15′–28°53′N | 55 |
| Yiyang, Jiangxi | 117°24′–117°61′E 27°99′–28°73′N | 5 |
| Wuyuan, Jiangxi | 117°21′–118°12′E 29°01′–29°34′N | 14 |
| Bama, Guangxi | 107°25′–108°09′E 29°01′–29°34′N | 14 |
| Pinghe, Fujian | 117°21′–118°12′E 23°53′–24°14′N | 15 |
| Total | | 213 |

used to make the volume constant. Next, a 1 mL sample was placed in a 10 mL test tube, 2 mL of 0.1 mol/L $AlCl_3$ solution was added, and 3 mL of 1 mol/L potassium acetate solution was added. A 70% methanol solution was fixed to the mark. The solution was shaken well at room temperature for 30 min and used to detect the absorbance value at 420 nm using a spectrophotometer (723 N, Jingke Instrument Co., Ltd., Shanghai, China). Rutin was used as the standard for a calibration curve. The total flavonoid content was calculated using the following equation based on the calibration curve:

$$A = 28.63C + 0.0023, R^2 = 1.$$

where A is the absorbance and C is the total flavonoid content in mg/mL.

Determination of phloridin and trilobatin

Phloridin and trilobatin were determined by high-performance liquid chromatography (HPLC) [16, 31]. First, 0.50 g powder was accurately weighed and mixed with 20 mL 80% methanol solution and ultrasonically extracted for 30 min. After centrifuging at 8000 r/min for 10 min, the supernatant was poured out in a 50 mL volumetric flask, and the process was repeated once. Then, the volume was diluted to 50 mL with 80% methanol solution and passed through a 0.22 μ m organic membrane for liquid chromatographic determination. Second, a standard curve was prepared with phloridin and trilobatin standard products and then operated according to the chromatographic conditions to find the peak area and calculate the relevant content according to the standard curve equation. The chromatographic detection conditions were as follows: Waters Atlantis T3 column (5 μ m, 250 mm \times 4.6 mm); mobile phase: phase A was methanol, phase B was water (V/V = 52/48); the flow rate was 1.0 mL/min; the injection volume was 2 μ L; the detection wavelength was 285 nm; and the column temperature was 30 $^{\circ}C$.

NIR spectra collection

The NIR spectra of sweet tea leaves were collected with a handheld fiber optic contact probe from a field-based spectrometer (LF-2500, Spectral evolution, USA). The range of the spectrometer was 1100–2500 nm, and the whole operation was performed at room temperature (approximately 25 $^{\circ}C$). Before the leaf sample spectral measurements, a baseline correction was performed with a standard white panel of known absorbance on the absorbance sample port. The powder of sweet tea leaf samples was placed on the sample port closely, and then the absorbance spectrum of each sample was obtained. Three replicates were taken for each sample, and the

average was considered. To exclude issues arising from sample mixing, devices were wiped with paper towels after each sample was collected.

Data analysis

Model evaluation

All experimental data was imported into the Unscrambler X software version 10.4 (64-bit) from CAMO (Computer Aided Modeling, Trondheim, Norway). Partial least squares analysis, which is universally applied for calibration in present chemometrics analysis, was performed for the establishment of the flavonoid prediction model. Before the construction of a model, we have scientifically divided a total of 213 samples. The sample set was randomly divided into a calibration set for developing the model and a prediction set to evaluate the built model at a ratio of 2:1 [24]. All the sweet tea leaf samples were sorted from small to large according to chemical content values, and one sample for every three samples is selected as the prediction set samples to test the robustness of the model. In addition, full cross-validation was used to verify the accuracy and stability of the model and the highest number of latent variables was fixed at 15. The main parameters of the PLS models evaluation were as follows: the root mean square error of calibration set ($RMSE_C$), validation set ($RMSE_V$), and prediction set ($RMSE_P$) and the coefficient of determination of calibration set (R^2_C), validation set (R^2_V), and prediction set (R^2_P). Models with low RMSE and high R^2 values were considered to be an excellent model [11]. These parameters are presented in the following equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \bar{y}_i - y_i}{n}} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

where, in Eq. (1) and (2), \bar{y}_i and y_i are the calibration, validation and prediction set values of the samples, respectively, and n is the number of samples.

Preprocessing method

The raw spectra not only contain effective information related to target substance content but are also covered by some useless information, that is, the noise and baseline that occurred in the detector and background or other unwanted changes due to the difference in physical structure between samples [34]. Therefore, it is absolutely vital that preprocessing should be done for a good calibration model in spectroscopy analysis. Generally,

preprocessing steps on the raw spectra are given below: (1) First, remove noise and baseline; (2) Second, correct the difference between samples. Moving average filters (MA) and Savitzky–Golay (SG) are two methods commonly used for reducing noise. Multiplicative scatter correction (MSC), standard normal variate (SNV) and differential coefficient (first and second derivatives, D1 and D2) are used to correct sample differences. Derivative operations are usually applied in combination with SG. In the present study, to compare different preprocessing methods, six preprocessing methods on PLS were studied, including MA, SG, SNV, MA + MSC, SG + D1 and SG + D2.

Results

Variation in the flavonoid content of sweet tea leaf samples among different locations

HPLC revealed that the contents of flavonoids in sweet tea leaves from different locations were different. As shown in Fig. 1, sweet tea leaves of Bama, Guangxi, had the maximum levels of total flavonoids and phloridin, 8.12 ± 0.62 and 13.99 ± 1.06 g/100 g, respectively (all values represent means \pm sem). The maximum level of trilobatin in sweet tea leaves was Xupu, Hunan, with an average of 25.02 ± 1.05 g/100 g.

The correlation coefficients of flavonoid content between different locations are shown in Table 2. It can be used to evaluate the relationship between locations and ingredients. We can see that the content of total flavonoid had a negative correlation with the location

Table 2 Correlation Matrix of constituents between different locations

| | Location | Total flavonoids | Phloridin | Trilobatin |
|------------------|----------|------------------|-----------|------------|
| Location | 1.000 | | | |
| Total flavonoids | -0.094 | 1.000 | | |
| Phloridin | 0.159* | -0.089 | 1.000 | |
| Trilobatin | 0.272** | 0.316** | -0.219** | 1.000 |

*P < 0.05; **P < 0.01

($r = -0.094$), phloridin has a significant positive correlation with it ($r = 0.159$), and trilobatin has a very significant positive correlation with the location ($r = 0.272$). In addition, trilobatin and total flavonoid showed a very significant positive correlation ($r = 0.316$) and a very significant negative correlation with phloridin ($r = -0.219$).

NIR Spectra

Raw spectra

The raw NIR spectra of sweet tea leaf samples are shown in Fig. 2. As shown, it was obviously hard to observe the difference between different samples in the original spectra. The spectra had two important peaks. The predominant peaks were related to O–H stretching and bending combinations observed between 1400–1500 nm and 1900–2000 nm [19, 21]. The small plate absorption peaks appearing at 1200 nm, 1650–1800 nm, 2100–2200 nm and 2300 nm were assigned to C–H stretching vibrations [24]. To visualize differences between different locations, principal component analysis was applied to examine the

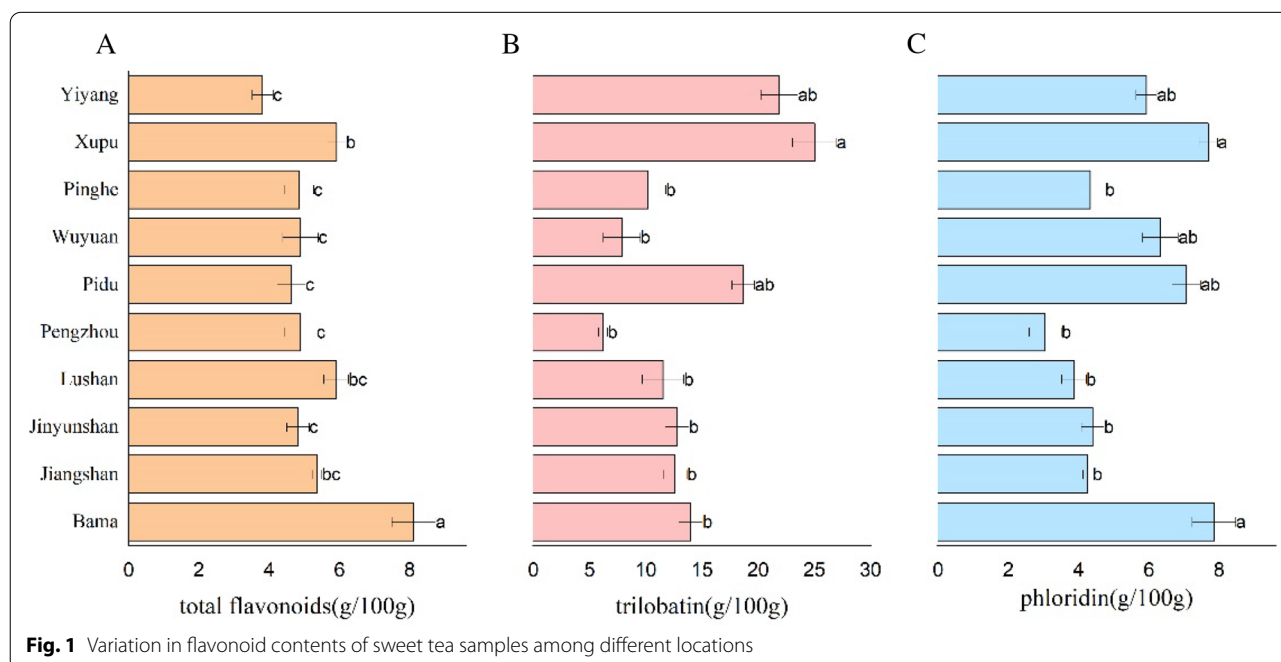
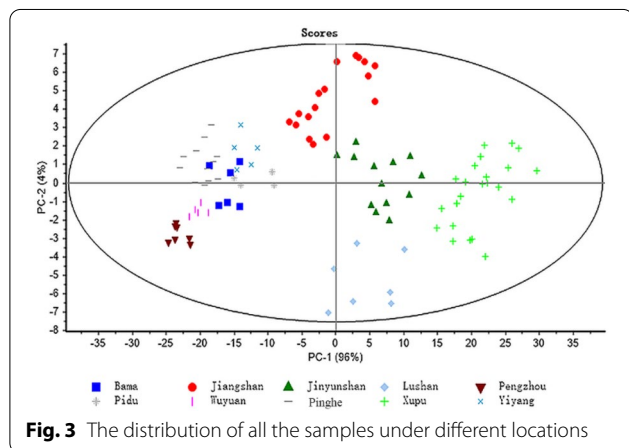
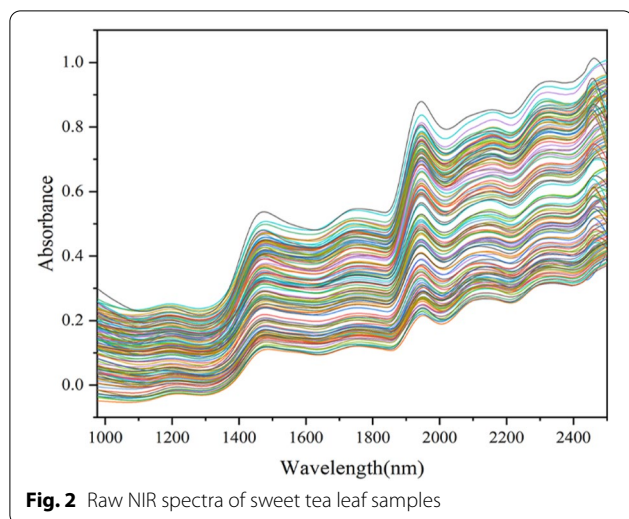


Fig. 1 Variation in flavonoid contents of sweet tea samples among different locations



flavonoids of sweet tea from different locations (Fig. 3). After removing the outliers, the result showed that the differences between leaf samples collected in different locations could be reflected by NIR spectroscopy combined with PCA.

Spectral preprocessing

Preprocessing methods were applied to the spectra, including MA, SG, MA combined with MSC, SNV, SG combined with first derivative spectra (SG + D1), and SG combined with second derivative spectra (SG + D2). The spectra after preprocessing are shown in Fig. 4.

Establishment of the PLS model

After the calibration set and the prediction set were established, PLS models of the three constituents were constructed with the raw and preprocessed NIR spectra. The evaluated parameters of different preprocessing PLS models are shown in Table 3 to select

the model with the best performance. Models with low root mean square error (RMSE) and high coefficient of determination (R^2) values were considered to be excellent models [11]. Figure 5 showed the prediction results of the PLS models according to which the developed models could predict total flavonoid, trilobatin and phloridin contents. (Notes: MA, Moving average filters; SG, Savitzky–Golay smoothing with 9 points; MSC, MA combined with multiplicative scatter correction; SNV, Standard normal variate; SG + D1, SG combined with first derivative spectra; SG + D2, SG combined with second derivative spectra; RMSEC, RMSEV and RMSEP, the root mean square error of calibration set, validation set, and prediction set; R^2_C , R^2_V and R^2_P , the coefficient of determination of calibration set, validation set and prediction set.)

PLS model for total flavonoids

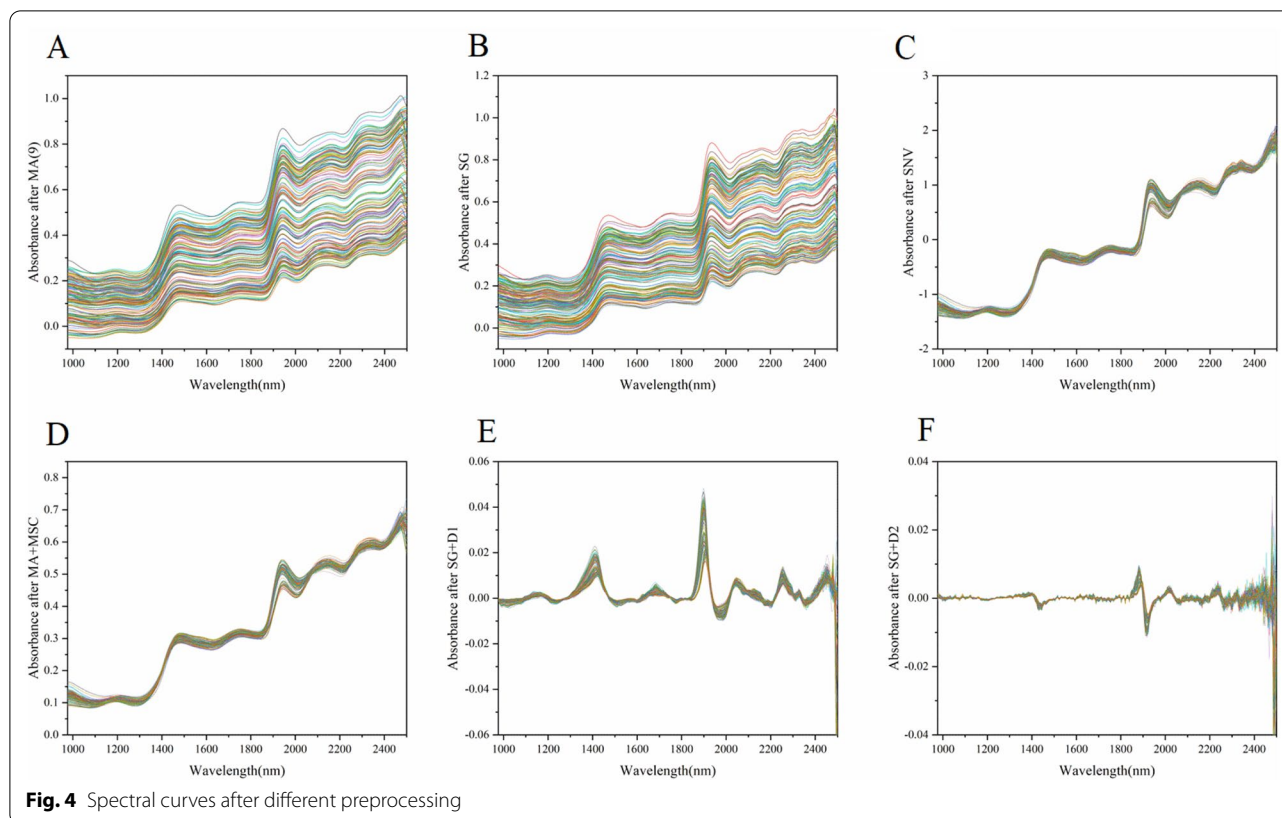
As seen from Table 3, the best result of the total flavonoid calibration model was developed with NIR spectra preprocessed with SG + D2. The combination of both, which makes the model more robust and reliable, can play a good role in reducing noise and correcting sample differences [17]. This model, which is shown in Fig. 5A, had the highest R^2_C (0.998) and lowest $RMSEC$ (0.069). For the validation set, the best performance was also obtained with SG + D2, with an R^2_V of 0.931 and an $RMSE_V$ of 0.409. In addition, the model was tested using the prediction set, and it performed best with a higher R^2_P of 0.893 and a lower $RMSE_P$ of 0.131.

PLS model for trilobatin

As shown in Table 3, similar to total flavonoids, the best result of the trilobatin calibration model was developed with raw NIR spectra preprocessed with SG + D2. This model had the highest R^2_C (0.993) and lowest values of $RMSEC$ (0.832). For the validation set, the model also had a higher R^2_V of 0.829 and an $RMSE_V$ of 4.102. However, a lower performance was observed for the prediction set with lower R^2 values. The best calibration, validation and prediction PLS models of trilobatin shown in Fig. 5B were developed with raw NIR spectra with R^2_C (0.967), $RMSEC$ (1.654), R^2_V (0.872), $RMSE_V$ (3.283), R^2_P of 0.902 and $RMSE_P$ of 2.993.

PLS model for phloridin

Phloridin content was predicted by the PLS algorithm using different spectral preprocessing methods. According to the calibration set and prediction set of PLS models, the SNV spectral model had the best performance with the highest R^2_C (0.987) and lowest values of $RMSEC$ (0.263). For the validation set, the best performance was also obtained with SNV with an R^2_V of 0.858 and an



RMSE_V of 0.893. In addition, the model was tested using the prediction set, and it performed best with a higher R^2_P of 0.818 and a lower RMSE_P of 1.085. The best performance of the phloridin PLS model is presented in Fig. 5C.

Discussion

The variation in the flavonoid content of sweet tea leaf samples among different locations was different. The results are consistent with some previous studies [30, 33]. In our research, the content of trilobatin in sweet tea from Hunan was the highest, which is consistent with Yang's results. In addition, correlation coefficients of flavonoid contents between different locations indicate that there is a certain correlation between the content of the three components in the same leaf. The leaf with higher trilobatin has higher total flavonoids and lower phloridin. In the study of Wei, he found that trilobatin made a large contribution to the tender leaves, and Yang found that the total flavonoid content was highest in the tender leaves, declining as the leaves aged. From previous studies [27, 30, 33], we can conclude that this may be related to the freshness of the leaves, and further research is needed in follow-up experiments.

The results of spectrum preprocessing show that after pretreatment with SG and MA, the spectra did not change much from the original spectrum. The SNV preprocessing method is as good as MA + MSC, which is consistent with the study of Chen [3]. Furthermore, SG + D1 and SG + D2 were considered the best. The results are consistent with Nascimento's study, which used various preprocessing methods for the spectra, and the first derivative of Savitsky–Golay with five smoothing points was regarded as the best result [19].

Modeling with spectra after different pretreatments, the best PLS model of total flavonoids was developed with NIR spectra preprocessed with SG + D2. Many studies have also reported that the best preprocessing method for NIR spectroscopy is derivatization with SG. For example, Farhadi established a prediction model for starch, reducing sugar, and moisture content in potatoes [8], Zhang measured leaf water content using VIS/NIR spectroscopy [34] and Nascimento developed PLS models for soluble solids content and firmness determination in low-chilling peach [19]. The best PLS model of trilobatin was developed with raw NIR spectra. The results of Hassan's research also show that non preprocessing was

Table 3 The results of the three constituents by PLS models with different preprocessing methods

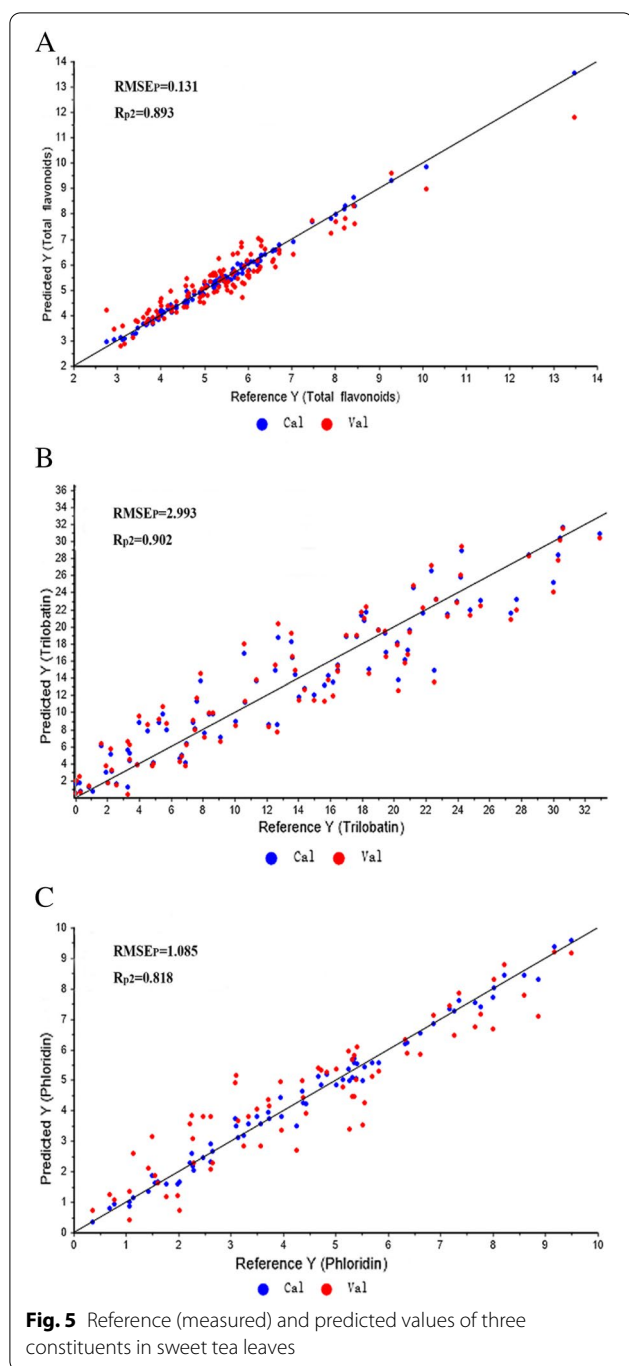
| Method | R ² _c | RMSEC | R ² _v | RMSEV | R ² _p | RMSEP |
|------------------|-----------------------------|--------------|-----------------------------|--------------|-----------------------------|--------------|
| Total flavonoids | | | | | | |
| None | 0.826 | 0.641 | 0.512 | 1.085 | 0.570 | 0.942 |
| SG | 0.751 | 0.768 | 0.502 | 1.095 | 0.569 | 0.944 |
| MA(9) | 0.708 | 0.831 | 0.474 | 1.126 | 0.434 | 1.082 |
| SNV | 0.746 | 0.775 | 0.408 | 1.194 | 0.499 | 1.017 |
| MA + MSC | 0.569 | 1.010 | 0.400 | 1.202 | 0.412 | 1.103 |
| SG + D(1) | 0.878 | 0.538 | 0.453 | 1.149 | 0.289 | 0.522 |
| SG + D(2) | 0.998 | 0.069 | 0.931 | 0.409 | 0.893 | 0.131 |
| Trilobatin | | | | | | |
| None | 0.967 | 1.654 | 0.872 | 3.283 | 0.902 | 2.993 |
| SG | 0.904 | 3.021 | 0.831 | 4.049 | 0.824 | 4.412 |
| MA(9) | 0.922 | 2.788 | 0.851 | 3.894 | 0.841 | 4.059 |
| SNV | 0.889 | 3.274 | 0.758 | 4.877 | 0.792 | 4.805 |
| MA + MSC | 0.837 | 4.271 | 0.728 | 5.564 | 0.793 | 4.789 |
| SG + D(1) | 0.953 | 2.083 | 0.871 | 3.482 | 0.194 | 3.370 |
| SG + D(2) | 0.993 | 0.832 | 0.829 | 4.102 | 0.250 | 2.457 |
| Phloridin | | | | | | |
| None | 0.948 | 0.589 | 0.752 | 1.306 | 0.684 | 1.378 |
| SG | 0.741 | 1.531 | 0.482 | 2.185 | 0.467 | 1.906 |
| MA(9) | 0.738 | 1.640 | 0.531 | 2.215 | 0.431 | 1.891 |
| SNV | 0.987 | 0.263 | 0.858 | 0.893 | 0.818 | 1.085 |
| MA + MSC | 0.776 | 1.519 | 0.559 | 2.151 | 0.234 | 2.077 |
| SG + D(1) | 0.942 | 0.686 | 0.750 | 1.436 | 0.139 | 2.256 |
| SG + D(2) | 0.692 | 1.464 | 0.432 | 2.007 | 0.127 | 2.368 |

confirmed to be an optimal PLS model to determine flavonoid contents in Chinese wild rice [10].

These results showed that the established PLS models had a good ability to predict the total flavonoid, trilobatin and phloridin contents in sweet tea with high R² and low RMSE values. It shows the correlations between the measured and predicted flavonoids in sweet tea leaves, and the scatter distribution for all constituents is concentrated around the diagonals. NIR spectroscopy has the potential for the rapid prediction of flavonoids in sweet tea. Similar studies for flavonoids have been reported for fresh *Ginkgo biloba* leaves [7] and propolis [2]. Fang's model of *Ginkgo biloba* leaves had a better performance, with an R²_c of 0.922, an RMSE_c of 3.153, a higher R²_p of 0.905 and a lower RMSE_p of 3.511. However, in Betances-Salcedo's study, the R² values for all prediction sets of the PLS models were between 0.600 and 0.860, and the RMSE_p values were between 3.800 and 119.600, which are much lower than those in our research. This may be related to the species, quantity, etc. of the sample.

Conclusion

The results of this study indicate that a combination of NIR spectroscopy and chemometrics could determine the flavonoid content in sweet tea leaves from different locations. The differences in sweet tea leaf spectra in different locations could be presented by PCA, which created a basis for the quantitative determination of the content of flavonoids by NIR spectroscopy. Furthermore, different spectral preprocessing methods are used in this research to establish the best PLS models between spectral information and the content of the three constituents. NIR spectra preprocessed with SG + D2, none preprocessing and SNV were confirmed to be optimal PLS models to determine total flavonoid, trilobatin and phloridin contents, respectively. The SG + D2 model was achieved with R²_p = 0.893 and RMSE_p = 0.131 for total flavonoid content prediction. For trilobatin content prediction, the none preprocessing model was achieved with R²_p = 0.902 and



RMSE_p = 2.993, and the SNV model for phloridin was obtained with R_p² = 0.818 and RMSE_p = 1.085.

NIR spectroscopy is expected to be a rapid, cheaper and convenient method to measure the content of flavonoid in sweet tea. In this study, the flavonoid content determined by NIR spectroscopy was in great agreement with the chemical analysis, indicating that NIR spectroscopy could potentially be used for the quick prediction of flavonoid contained in sweet tea.

Abbreviations

NIR spectroscopy: Near infrared spectroscopy; PCA: Principal component analysis; PLS: Partial least squares; MA: Moving average filters; SG: Savitzky–Golay smoothing with 9 points; MSC: MA combined with multiplicative scatter correction; SNV: Standard normal variate; SG + D1: SG combined with first derivative spectra; SG + D2: SG combined with second derivative spectra; RMSE: Root mean square error; R²: Decisive factor; HPLC: High-performance liquid chromatography.

Acknowledgements

The authors would like to appreciate the National Nature Science Foundation of China (32071785) for providing the financial support for this research.

Authors' contributions

ZX performed the experiments, analyzed the data, and wrote the manuscript. YJ designed the study. ZL collected the materials. YJ, ZL and ZF modified the article. All authors have read and approved the final manuscript.

Funding

This work was funded by the National Natural Science Foundation of China (32071785) and the Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (2021C02070-7).

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no potential conflicts of interest with respect to the research, author and publication of this article.

Author details

¹Research Institute of Subtropical Forestry, Chinese Academy of Forestry, No. 73, Daqiao Road, Fuyang, Hangzhou 311400, Zhejiang Province, China. ²College of Forestry, Nanjing Forestry University, Nanjing, People's Republic of China.

Received: 2 December 2021 Accepted: 18 March 2022

Published online: 02 April 2022

References

- Balabin RM, Safieva RZ. Gasoline classification by source and type based on near infrared (NIR) spectroscopy data. *Fuel*. 2008;87(7):1096–101. <https://doi.org/10.1016/j.fuel.2007.07.018>.
- Betances-Salcedo E, Revilla I, Vivar-Quintana AM, Gonzalez-Martin MI. Flavonoid and antioxidant capacity of propolis prediction using near infrared spectroscopy. *Sensors (Basel)*. 2017. <https://doi.org/10.3390/s17071647>.
- Chen Q, Zhao J, Liu M, Cai J, Liu J. Determination of total polyphenols content in green tea using FT-NIR spectroscopy and different PLS algorithms. *J Pharm Biomed Anal*. 2008;46(3):568–73. <https://doi.org/10.1016/j.jpba.2007.10.031>.
- Chen Y, Bin J, Zou C, Ding M. Discrimination of fresh tobacco leaves with different maturity levels by near-infrared (NIR) spectroscopy and deep learning. *J Anal Methods Chem*. 2021;2021:9912589. <https://doi.org/10.1155/2021/9912589>.
- Diniz PHGD, Pistonesi MF, Araújo MCU. Using iSPA-PLS and NIR spectroscopy for the determination of total polyphenols and moisture in

- commercial tea samples. *Anal Methods*. 2015;7(8):3379–84. <https://doi.org/10.1039/c4ay03099k>.
6. Dou Y, Sun Y, Ren Y, Ju P, Ren Y. Simultaneous non-destructive determination of two components of combined paracetamol and amantadine hydrochloride in tablets and powder by NIR spectroscopy and artificial neural networks. *J Pharm Biomed Anal*. 2005;37(3):543–9. <https://doi.org/10.1016/j.jpba.2004.11.017>.
 7. Fang LI, Shi JY, Zhang DT, Chen ZW. Determination of total flavonoid content distribution on ginkgo biloba leaf. *Sci Technol Food Ind*. 2015;36(9):270–2. <https://doi.org/10.13386/j.issn1002-0306.2015.09.050>.
 8. Farhadi R, Afkari-Sayyah AH, Jamshidi B, Mousapour Gorji A. Prediction of internal compositions change in potato during storage using visible/near-infrared (Vis/NIR) spectroscopy. *Int J Food Eng*. 2020. <https://doi.org/10.1515/ijfe-2019-0110>.
 9. Guo H, Fu MX, Zhao YX, Li H, Li HB, Wu DT, Gan RY. The chemical, structural, and biological properties of crude polysaccharides from sweet tea (*Lithocarpus litseifolius* (Hance) Chun) based on different extraction technologies. 2021. *Foods*. <https://doi.org/10.3390/foods10081779>.
 10. Hassan H, Fan M, Zhang T, Yang K. Prediction of total phenolics and flavonoids contents in Chinese wild rice (*Zizania latifolia*) using FT-NIR spectroscopy. *Am J Food Technol*. 2015;10(3):109–17. <https://doi.org/10.3923/ajft.2015>.
 11. Huang Y, Du G, Ma Y, Zhou J. Predicting heavy metals in dark sun-cured tobacco by near-infrared spectroscopy modeling based on the optimized variable selections. *Ind Crops Prod*. 2021. <https://doi.org/10.1016/j.indcrop.2021.114003>.
 12. Inostroza L, Lobos I, Acuña H, Vásquez C, Tapia G, Monzón G. NIR-Prediction of water-soluble carbohydrate in white clover and its genetic relationship with cold tolerance. *Chilean J Agric Res*. 2017;77(3):218–25. <https://doi.org/10.4067/s0718-58392017000300218>.
 13. Jin X, Shi C, Yu CY, Yamada T, Sacks EJ. Determination of leaf water content by visible and near-infrared spectrometry and multivariate calibration in miscanthus. *Front Plant Sci*. 2017;8:721. <https://doi.org/10.3389/fpls.2017.00721>.
 14. Kai-Xun HE, Cheng H, Qian F. A new sample-selection and modeling method based on near-infrared spectroscopy and its industrial application. *J Donghua Univ*. 2014;31(2):207–11.
 15. Li S, Zeng J, Tan J, Zhang J, Wu Q, Wang L, Wu X. Antioxidant and hepatoprotective effects of *Lithocarpus polystachyus* against carbon tetrachloride-induced injuries in rat. *Bangladesh J Pharmacol*. 2014. <https://doi.org/10.3329/bjpv.v8i4.16458>.
 16. Liu Y, Liu H-Y, Xia Y, Guo H, He X-Q, Li H, Gan R-Y. Screening and process optimization of ultrasound-assisted extraction of main antioxidants from sweet tea (*Lithocarpus litseifolius* [Hance] Chun). *Food Biosci*. 2021. <https://doi.org/10.1016/j.fbio.2021.101277>.
 17. Ma L, Liu D, Du C, Lin L, Zhu J, Huang X, Wu Z. Novel NIR modeling design and assignment in process quality control of Honeysuckle flower by QbD. *Spectrochim Acta A Mol Biomol Spectrosc*. 2020;242: 118740. <https://doi.org/10.1016/j.saa.2020.118740>.
 18. Meng Y, Ding L, Wang Y, Nie QT, Xing YY, Ren Q. Phytochemical identification of *Lithocarpus polystachyus* extracts by ultra-high-performance liquid chromatography-quadrupole time-of-flight-MS and their protein tyrosine phosphatase 1B and alpha-glucosidase activities. *Biomed Chromatogr*. 2020;34(1): e4705. <https://doi.org/10.1002/bmc.4705>.
 19. Nascimento PA, de Carvalho LC, Júnior LC, Pereira FM, de Almeida Teixeira GH. Robust PLS models for soluble solids content and firmness determination in low chilling peach using near-infrared spectroscopy (NIR). *Postharvest Biol Technol*. 2016;111:345–51. <https://doi.org/10.1016/j.postharvbio.2015.08.006>.
 20. Norris K, Barnes R, Moore J, Shenk J. Predicting forage quality by infrared reflectance spectroscopy. *J Anim Sci*. 1976;43:889–97.
 21. Osborne BG, Fearn T, Hindle PH. Practical NIR spectroscopy with applications in food and beverage analysis. Longman Scientific & Technical. 1993;153.
 22. Prananto JA, Minasny B, Weaver T. Near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues. *Advances in Agronomy*. 2020;164:1–49. <https://doi.org/10.1016/bs.agron.2020.06.001>.
 23. Richardson AD, Reeves JB, Gregoire IG. Special issue: plant speciation II multivariate analyses of visible/near infrared (VIS/NIR) absorbance spectra reveal underlying spectral differences among dried, ground conifer needle samples from different growth environments. *New Phytol*. 2004;161(1):291–301.
 24. Sanaeifar A, Huang X, Chen M, Zhao Z, Ji Y, Li X, Yu X. Nondestructive monitoring of polyphenols and caffeine during green tea processing using Vis-NIR spectroscopy. *Food Sci Nutr*. 2020;8(11):5860–74. <https://doi.org/10.1002/fsn3.1861>.
 25. Sandorfy C, Buchet R, Lachenal G. Principles of Molecular Vibrations for Near-Infrared Spectroscopy. In: Ozaki Y, McClure WF, Christy AA, editors. *Near-Infrared Spectroscopy in Food Science and Technology*. 2006;1–46.
 26. Shang A, Liu HY, Luo M, Xia Y, Yang X, Li HY, Gan RY. Sweet tea (*Lithocarpus polystachyus* rehd.) as a new natural source of bioactive dihydrochalcones with multiple health benefits. *Crit Rev Food Sci Nutr*. 2020. <https://doi.org/10.1080/10408398.2020.1830363>.
 27. Sheng-Hua LI, Xian-Jin WU, Niu YY. Determination of dihydrochalcone from *Lithocarpus polystarch* Rehd. *Food Sci Technol*. 2010;35(6):281–3. <https://doi.org/10.13684/j.cnki.spkj.2010.06.012>.
 28. Teye E, Huang X, Afoakwa N. Review on the potential use of near infrared spectroscopy (NIRS) for the measurement of chemical residues in food. *J Dairy Sci*. 2013;85(3):487–93.
 29. van Maarschalkerweerd M, Husted S. Recent developments in fast spectroscopy for plant mineral analysis. *Front Plant Sci*. 2015;6:169. <https://doi.org/10.3389/fpls.2015.00169>.
 30. Wei M, Tuo Y, Zhang Y, Deng Q, Shi C, Chen X, Zhang X. Evaluation of two parts of *Lithocarpus polystachyus* Rehd. from different chinese areas by multicomponent content determination and pattern recognition. *J Anal Methods Chem*. 2020;2020:8837526. <https://doi.org/10.1155/2020/8837526>.
 31. Wen C, Du M, Liu H. Assay of phlorizin from the water extract of *Lithocarpus litseifolius* by high performance liquid chromatography. *J Hunan Normal Univ Med Sci*. 2017;14(2):190–2.
 32. Xiong C, Liu C, Pan W, Ma F, Xiong C, Qi L, Zheng L. Non-destructive determination of total polyphenols content and classification of storage periods of Iron Buddha tea using multispectral imaging system. *Food Chem*. 2015;176:130–6. <https://doi.org/10.1016/j.foodchem.2014.12.057>.
 33. Yang J, Huang Y, Yang Z, Zhou C, Hu X. Identification and quantitative evaluation of major sweet ingredients in sweet tea (*Lithocarpus polystachyus* Rehd.) based upon location, harvesting time, leaf age. *J Chem Soc Pak*. 2018;40(1):158–64.
 34. Zhang Q, Li Q, Zhang G. Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. *Spectroscopy*. 2012;27:93–105. <https://doi.org/10.1155/2012/276795>.
 35. Zhong DL, Han SF, Ding M. Establishment of determination method for total flavonoids from tomato. *Food Sci*. 2009;30(22):272–4.
 36. Zhou CJ, Huang S, Liu JQ, Qiu SQ, Xie FY, Song HP, Lai XP. Sweet tea leaves extract improves leptin resistance in diet-induced obese rats. *J Ethnopharmacol*. 2013;145(1):386–92. <https://doi.org/10.1016/j.jep.2012.09.057>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

