Original Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.94322

CC BY

Rapid determination of tea polyphenols content in Qingzhuan tea based on near infrared spectroscopy in conjunction with three different PLS algorithms

Shengpeng WANG[1], Panpan LIU[1], Lin FENG[1], Jing TENG[1], Fei YE[1], Anhui GUI[1], Xueping WANG[1], Lin ZHENG[1], Shiwei GAO[1]*, Pengcheng ZHENG[1]*

Abstract

Tea polyphenols are one of the most important ingredients in Qingzhuan tea. Usually, a chemical method is used to determine tea polyphenols content, but it was time-consuming and laborious. This paper attempted to use near infrared spectroscopy (NIRS) technology combined with three partial least squares methods to predict tea polyphenols content quickly and nondestructively. The partial least squares (PLS), synergy interval PLS (siPLS) and genetic algorithm based PLS (gaPLS) were used to establish prediction models, the performance of the final model was showed by root mean square error of prediction (RMSEP) and determination coefficient (Rp^2) in prediction set. The best spectral preprocessing method was multivariate scattering correction (MSC); the RMSEP and R_p^2 of PLS model were 0.145% and 0.8974, respectively; the siPLS model was established with four spectral regions (4377.6 cm^{-1}-4751.7 cm^{-1}, 4755.6 cm^{-1}-5129.7 cm^{-1}, 6262.7 cm^{-1}-6633.9 cm^{-1} and 7386 cm^{-1}-7756.3 cm^{-1}), whose RMSEP and R_p^2 were 0.0652% and 0.9235, respectively; the gaPLS model was established with 36 spectra dada points and showed the best performance (RMSEP=0.0624%, Rp^2=0.9769) compared with the PLS and si-PLS models. Therefore, the application of near infrared technology combined with the gaPLS method could predict tea polyphenols content in Qingzhuan tea more accurately and rapidly.

Keywords: Qingzhuan tea; tea polyphenols; near infrared spectroscopy; partial least squares; genetic algorithm.

Practical Application: Rapid determination of tea polyphenols content in Qingzhuan tea.

1 Introduction

Qingzhuan tea, a unique kind of dark tea, had more than 200 years' history (Xia, 2016), which was mainly produced in Chibi city, Hubei province of China. When Qingzhuan tea produced, fermentation was one of the most critical procedure. In the stages of fermentation, the raw tea materials have occurred a series of biochemical reactions, formed the unique quality characteristics of Qingzhuan tea soup with bright color and tasty thick mellow (Wang et al., 2020a). Qingzhuan tea soup contained many kinds of beneficial components, such as theabrownin, thearubigin, tea polysaccharide (Zhou et al., 2022a, b) and tea polyphenols. If the tea soup absorbed by human body, it can decompose fat (Yang et al., 2017), relax stomach (Chen et al., 2010), resist oxidation (Yan et al., 2018) and decrease blood sugar (Yamagata et al., 2015) and so on function effect. Therefore, Qingzhuan tea is very popular with consumers, and has become an indispensable necessity in daily life (Figure 1).

Tea polyphenols are of great interest owing to their beneficial medicinal properties (Al-Hatim et al., 2022). Increasing evidence shows that polyphenols substances found in tea can enhance general health. Recently, researchers has suggested that antioxidants found in polyphenol substances might play important roles in preventing cardiovascular disease (Nakachi et al., 2000), chronic gastritis (Setiawan et al., 2001), and some cancers (Jian et al., 2004; Fujiki et al., 2001; Inoue et al., 2001). Furthermore, polyphenol compounds are mainly responsible for the characteristic astringent and bitter tastes of brewed tea (Wang et al., 2020b). Therefore, the tea polyphenol content is an index for evaluating the quality of Qingzhuan tea. Generally, the Folin phenol chemical method (Zhou et al., 2008) is used to determine tea polyphenol contents, but this method is not only time-consuming and laborious, but also pollutes the peripheral environment. Therefore, developing a fast and nondestructive method to determine tea polyphenol contents is necessary.

Near-infrared (NIR) spectroscopy is a fast, accurate, and non-destructive technique that can be used to replace traditional chemical analysis methods. NIR spectroscopy is a powerful analytical tool widely applied in agricultural (Zhou et al., 2015; Xuemei & Jianshe, 2014), petrochemical (Wu et al., 2014), textile (Tavanaie et al., 2015), and pharmaceutical industries (Blanco & Peguero, 2010; Lee et al., 2011). NIR spectroscopy has been used in the tea industry, for example, to analyze caffeine (Esteban-Diez et al., 2004) and free amino acids, and determine the origins of tea varieties (Kachrimanis et al., 2003). Backward interval partial least squares (bi-PLS) and back propagation-artificial neural network (BP-ANN) algorithms, and other pattern recognition systems (Shi et al., 2011), such as the support vector machine, have been used to calibrate tea nutrition models (Chen et al., 2015). However, in terms of determining tea polyphenol contents in Qingzhuan tea, it has not been reported to use NIRS combined with different PLS algorithms to study Qingzhuan tea.

Received 29 July, 2022
Accepted 18 Sept., 2022
[1]*Institute of Fruit and Tea, Hubei Academy of Agricultural Sciences, Wuhan, PR China*
*Corresponding author: gsw0609@126.com; zpct15@163.com

Figure 1. Sample of Qingzhuan tea.

The partial least squares (PLS) method was a new type of multivariate statistical data analysis method (Lu, 2010), which combined three methods' advantages of principal component analysis, canonical correlation analysis and multivariate linear regression analysis. The main research was the relevance between multivariate regression modeling and multiple independent variables. The partial least squares regression method was more effective, especially when the internal variables are highly linear correlation.

The synergy interval partial least squares (siPLS) (Wang et al., 2012) was a wavelength optimization method. The full NIR spectrum was divided into 10-25 spectral regions, and then the 2, 3 and 4 spectral intervals were combined to establish prediction models by using PLS method. When the RMSECV value was the lowest, the siPLS prediction model was the best, and the selected spectral intervals were the best spectral interval combination. The siPLS method not only can reduce spectral noise effectively, but also enhance the prediction accuracy of models.

Genetic algorithm partial least squares (gaPLS) (Escobar et al., 2014) was an optimization method to simulate the evolutionary mechanism of biological species competition selection based on the theory of biological evolution. It realized the iterative optimization of the group by applying the genetic manipulation to the individual in the group in accordance with fitness function. GaPLS algorithm was applied to select the spectral data points for optimization models. When the calibration model was established and the root mean square error of cross-validation (RMSECV) was the lowest, it was meaning that the calibrating model had the best prediction ability.

In this research, after near infrared spectroscopy of Qingzhuan tea obtained and pretreated, the prediction models of tea polyphenols content were established by using PLS method, siPLS method and gaPLS method,respectively. And the final models were evaluated by the root mean square error of prediction (RMSEP) and the determination coefficient (Rp^2) in prediction set. The results of this study hoped to determine tea polyphenols content in Qingzhuan tea rapidly and nondestructively for evaluation the quality of Qingzhuan tea products.

2 Materials and methods

2.1 Sample preparation

The total number of 128 Qingzhuan tea samples, processed between April 2016 and September 2016, were obtained from Yangloudong tea limited liability company, Zhaoliqiao tea limited liability company and Dongzhuang tea company Ltd.. All samples were crushed over 200 mesh sieve for determination of tea polyphenols content. Then, the samples were randomly divided into two subsets according to 3:1 ratio, meaning 96 samples in calibration set and 32 samples in prediction set, which were used to establish the calibration models and prediction models, respectively.

2.2 Determination of tea polyphenols

The content of tea polyphenols was determined by the standard Folin phenol method (Zhou et al., 2008). After Qingzhuan tea samples grinded, tea polyphenols were distilled under the conditions of 70% methanol and 70 °C water bath. The -OH groups of tea polyphenols were oxidized by Folin phenol reagent, and showed blue color, then the absorbance was measured by a spectrophotometer at the wavelength of 765 nm. The standard curve was drawn by using gallic acid as the standard correction substance used to determine the content of tea polyphenols.

2.3 Spectra collection

NIR spectra were obtained in the reflectance mode using a Thermo Antaris II Fourier transform (FT) NIR spectrometer (Nicolet AntarisII, Thermofisher Scientific, U.S.A.) coupled with an InGaAs detector, a quartz halogen lamp, and an integrating sphere accessory. The samples were placed in a sample cup (ø 30 mm) specifically designed for this application. For each sample, the Qingzhuan tea (10 g) was placed into the sample cup according to the procedure specified by the manufacturer. The spectral data were obtained from 10,000 cm^{-1} to 4,000 cm^{-1} at 3.857 cm^{-1} intervals while rotating the sample cup 360° such that the entire sample was analyzed. This analysis resulted in 1,557 variables. During spectral collection, the sample cup fitted with intermediate samples was rotated steadily at a constant speed. Duplicates of each sample were scanned three times. The average spectrum of each sample was employed in following analysis.

2.4 Spectral preprocessing and model establishment

To reduce spectral noise, each spectrum was pretreated with first derivative (FD), second derivative (SD), smoothing (Sm) and multiple scattering correction (MSC) by using TQ Analyst 9.4.45 software, OPUS 7.0 software and Matlab7.0 software for selecting the best pretreatment method. After comparing the results of models, the smoothing was as the best preprocessing method. The prediction ability of NIR models were evaluated by determination coefficient of cross validation (Rc^2), determination coefficient of prediction (Rp^2), root mean square error of cross validation (RMSECV), root mean square error of prediction (RMSEP) and factors.

The RMSECV is calculated as follows (Equation 1):

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i^{'} - y_i\right)^2}{n}} \quad (1)$$

where n is the number of samples in the calibration set, y_i is the true value for sample i, and $y_i^{'}$ is the theoretical value for sample i predicted from the calibration set.

For the prediction set, the root mean square error of prediction (RMSEP) is calculated as follows (Equation 2):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_i^{'}\right)^2}{n}} \quad (2)$$

where n is the number of samples in the prediction set, y_i is the true value for sample i, and $y_i^{'}$ is the predicted value for sample i.

R^2 was calculated as follows (Equation 3):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{'} - y_i)^2}{\sum_{i=1}^{n}(y_i^{i} - \bar{y})^2} \quad (3)$$

where $\bar{y}$ is the mean true value, y_i is the true value for sample i, and $y_i^{'}$ is the predicted value for sample i.

3 Results and discussion

3.1 The contents of tea polyphenols

The content of tea polyphenols was determined by using the standard method, and the results were shown in table 1.

As seen from Table 1, the values of the range, mean and S. D. in calibration set were from 3.25% to 11.04%, 6.52% and 2.36, respectively; the values of the range, mean, and S.D. in prediction set were from 3.46% to 10.32%, 5.77% and 2.10, respectively. Thus, it was evident that the range of tea polyphenols content in prediction set was included in the range of samples in calibration set, which laid the foundation for prediction tea polyphenols content more accurately.

3.2 Screening of spectral preprocessing methods

Generally speaking, there was more noise information from 9800 to 10000cm^{-1} and from 4000 to 4200 cm^{-1} in NIR spectra, so these two wavenumber intervals should be discarded in modeling. There are -OH absorption peaks of free water near 6900cm^{-1} and 5100cm^{-1}, which may affect the prediction accuracy of the tea polyphenols content models. Besides, it was necessary to try to use a variety of pretreatment methods to pretreat the sample spectrum, and get a better pretreatment method.

It can be seen from table 2 that the processing results obtained by different spectral pretreatment methods were quite different. The results without spectral preprocessing were the worst (R_c^2=0.6021, RMSECV=0.248%); the best results were obtained by using the MSC preprocessing method(R_c^2=0.8421, RMSECV=0.125%), whose RMSECV was reduced by 49.6% compared with the worst non spectral pretreatment method. It was thus clear that denoising the original spectrum can improve the prediction effect of the tea polyphenol contents models.

3.3 Results of PLS model

Partial least squares method (PLS) is an effective linear modeling method, and the number of factor is very important. The NIRS-PLS prediction model of tea polyphenols content was established by using full spectra except the both spectral ranges of 4000-4200 $cm^{-1}$and 9800-10000 cm^{-1}, the R_c^2 and RMSECV were 0.9145 and 0.0832%, respectively, when 10 factors were included in calibration model. The robustness of the calibration model was tested by all the prediction set samples, and the results were shown in Figure 2a. As seen from Figure 2a, the R_p^2 and RMSEP were 0.8974 and 0.145%, respectively. So, the model established by partial least squares method had good prediction ability and overfitting phenomenon was not appeared.

3.4 Results of siPLS model

Table 3 were the results of calibration models established by using siPLS algorithm. As seen from Table 3, when the spectra were divided into 16 sub-regions and the factor number was 7, the results of the calibration model were the best, and the RMSECV value was the least, 0.0624%. The selected spectral regions were [2 3 7 10], and the corresponding spectral wavelengths were 4377.6 cm^{-1}-4751.7 cm^{-1}, 4755.6 cm^{-1}-5129.7 cm^{-1}, 6262.7 cm^{-1}-

Table 1. Classification and tea polyphenols contents of all samples.

Set	Range (%)	Mean (%)	S. D.
Calibration set	3.25-11.04	6.52	2.36
Prediction set	3.46-10.32	5.77	2.10

S.D.: standard deviation.

Table 2. Comparison of spectral pretreatment methods.

No.	Pretreatment methods	R_c^2	RMSECV (%)
1	None	0.6021	0.248
2	FD	0.7041	0.215
3	SD	0.7253	0.194
4	MSC	0.8421	0.125
5	Sm	0.7312	0.168

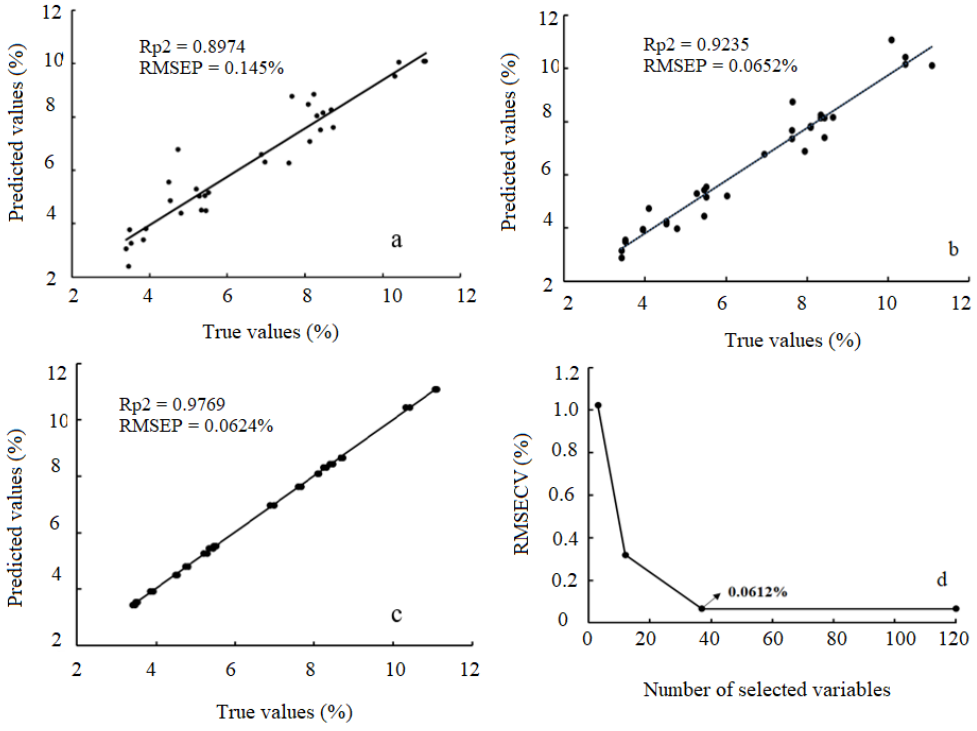

Figure 2. Three PLS kinds scatter plots of tea polyphenols content. Note: (a) True values versus NIR predicted values by PLS in prediction set; (b) True values versus NIR predicted values by siPLS in prediction set; (c) True values versus NIR predicted values by gaPLS in prediction set; (d) The best number of selected variables by gaPLS algorithm.

6633.9 cm^{-1} and 7386 cm^{-1}-7756.3 cm^{-1}, respectively (Figure 3). The determination coefficient (Rc^2) in the calibration set model was 0.9457. The determination coefficient (Rp^2) and RMSEP were 0.9235 and 0.0652%, respectively (Figure 2b). The prediction results were better than PLS model results.

3.5 Results of gaPLS model

As seen from Figure 2d, with the selected data points gradually increased, RMSECV value rapidly decreased from 1.0223% to 0.0612%, when 36 data points were selected (Table 4) (the ratio of accounting for all the 1557 spectral data points was 2.4%), and the lowest RMSECV value was 0.0612%. Here, determination coefficient (Rc^2) was 0.9821 and the best number of PLS factors was 10 in calibration set. When the performance of gaPLS calibration model was evaluated by the samples in prediction set, the root mean square error of prediction set (RMSEP) was 0.0624% and determination coefficient (Rp^2) is 0.9769 in prediction set (Figure 2c). The predicted results were showed in Table 5.

From the distribution of data points (Table 4), we can see that there were 32 spectral data points distributed in the long wave region of the near infrared spectrum, and the ratio of accounting for all the 36 spectral data points was 88.89%; there were 4 spectral data points distributed in the short wave region of the near infrared spectrum, and the ratio of accounting for all the 36 spectral data points was 11.11%. This was because the near infrared spectra of Qingzhuan tea samples were mainly distributed in the long wavelength region, whose spectral information was abundant and represented the most spectral information of Qingzhuan tea samples; On the contrary, short-wave near-infrared spectral region of Qingzhuan tea samples had less spectral information. Therefore, when the spectral data points were optimized by genetic algorithm, the spectral data points were mainly located in the long wave of near infrared spectrum region.

Table 3. Results of siPLS calibration model selected different spectral regions.

Number of intervals	PLS factors	Selected intervals	RMSECV (%)
10	10	[1 3]	0.0644
11	10	[1 3 6 10]	0.0650
12	9	[5 6 7 11]	0.0645
13	9	[6 7 10 11]	0.0649
14	10	[1 6 7 10]	0.0647
15	10	[1 6 9]	0.0643
16	**7**	**[2 3 7 10]**	**0.0624**
17	10	[1 6 9 11]	0.0644
18	9	[3 9 12 16]	0.0637
19	8	[8 10 16 18]	0.0643
20	8	[3 9 13 14]	0.0625
21	9	[3 9 12 20]	0.0635
22	8	[9 12 18]	0.0638
23	10	[1 2 4 6]	0.0638
24	10	[1 4 8 13]	0.0639
25	9	[4 8 16 21]	0.0628

Table 4. The 36 spectra data points selected by genetic algorithm.

No.	Data points (cm^{-1})	No.	Data points (cm^{-1})	No.	Data points (cm^{-1})
1	4065.21	13	6741.92	25	8631.82
2	4335.19	14	6954.05	26	8670.39
3	4389.19	15	6981.05	27	8693.53
4	4551.18	16	7409.17	28	8724.38
5	5654.26	17	7482.45	29	8789.95
6	5820.11	18	7679.15	30	8932.66
7	5839.40	19	7729.29	31	8994.37
8	5908.82	20	8072.56	32	9052.22
9	5947.39	21	8103.42	33	9102.36
10	5997.53	22	8176.70	34	9144.79
11	6032.24	23	8427.40	35	9191.07
12	6055.39	24	8616.39	36	9503.48

Table 5. The results in prediction set samples by the best gaPLS model.

No.	Predicted values (%)	True values (%)	No.	Predicted values (%)	True values (%)	No.	Predicted values (%)	True values (%)
1	8.08	8.11	2	10.37	10.32	3	7.03	7.00
4	7.64	7.63	5	9.8	9.88	6	4.8	4.80
7	9.45	9.47	8	9.35	9.43	9	3.64	3.57
10	6.24	6.22	11	8.48	8.54	12	3.97	3.85
13	5.5	5.53	14	8.24	8.37	15	4.57	4.50
16	3.69	3.74	17	6.02	6.11	18	4.91	4.88
19	6.74	6.83	20	7.05	7.13	21	4.34	4.25
22	5.81	5.87	23	8.54	8.48	24	4.58	4.57
25	6.51	6.56	26	5.08	5.03	27	5.88	5.85
28	6.18	6.21	29	10.25	10.19	30	4.5	4.50
31	6.65	6.62	32	9.97	9.88			

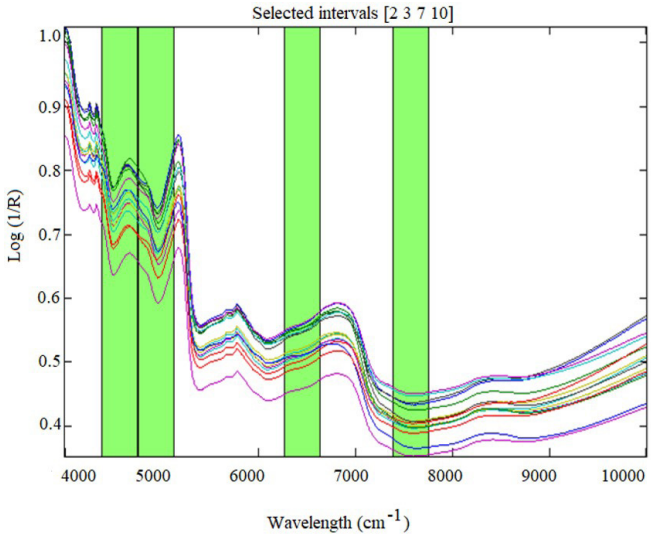

Figure 3. Optimal spectral region selected by siPLS with wavelength 4377.6 cm^{-1}-4751.7 cm^{-1}, 4755.6 cm^{-1}-5129.7 cm^{-1}, 6262.7 cm^{-1}-6633.9 cm^{-1} and 7386 cm^{-1}-7756.3 cm^{-1}.

3.6 Discussion

By comparing the prediction results of the PLS, siPLS and gaPLS models, we can see that, gaPLS model was the best, followed by the siPLS model. The worst performance was PLS model. This may be due to that, the full wavelength was used to establish the PLS prediction model, because of the noise information in the spectra, it would affect the prediction effect of the model, resulting in poor prediction ability. Based on PLS algorithm, siPLS method has optimized the spectral intervals which were closely related to tea polyphenols. When these characteristic spectral intervals were used for modeling, it not only can eliminate some spectral noise information, reduce the modeling computation, but also effectively improve the prediction ability of the model. Therefore, the results of siPLS model were better than the results of the PLS model. GaPLS showed its incomparable superiority. By applying genetic algorithm, the most useful spectral data points were extracted for modeling, further eliminating more noise information, so the performance of gaPLS model was the best.

4 Conclusion

In this paper, near infrared spectroscopy (NIRS) together with three different partial least squares (PLS) were applied to predict tea polyphenols content in Qingzhuan tea rapidly and nondestructively. Among the three PLS methods, the worst was PLS model, followed by siPLS model, and the best was gaPLS model, which was established by using 36 spectra data points. The optimal calibration model was achieved with $Rp^2 = 0.9769$ and RMSEP = 0.0624% in prediction set. The experimental results will

provide a reference for rapid detection of the inner components in Qingzhuan tea products.

Acknowledgements

This work was supported by the National tea industry technology system project (CARS-19), National Natural Science Foundation of China (31902081), Hubei agricultural science and technology innovation center project (2021-620-000-001-024), Key research and development projects of Hubei Province (2022BBA0068).

References

Al-Hatim, R. R., Al-Alnabi, D. I. B., Al-Younis, Z. K., Al-Shawi, S. G., Singh, K., Abdelbasset, W. K., & Mustafa, Y. F. (2022). Extraction of tea polyphenols based on orthogonal test method and its application in food preservation. *Food Science and Technology*, 42, e70321. http://dx.doi.org/10.1590/fst.70321.

Blanco, M., & Peguero, A. (2010). Analysis of pharmaceuticals by NIR spectroscopy without a reference method. *Trends in Analytical Chemistry*, 29(10), 1127-1136. http://dx.doi.org/10.1016/j.trac.2010.07.007.

Chen, Q., Zhang, D., Pan, W., Ouyang, Q., Li, H., Urmila, K., & Zhao, J. (2015). Recent developments of green analytical techniques in analysis of tea's quality and nutrition. *Trends in Food Science & Technology*, 43(1), 63-82. http://dx.doi.org/10.1016/j.tifs.2015.01.009.

Chen, Y. Q., Zhang, W., Ni, D. J., Cheng, Q., Liu, X. H., & Gan, D. P. (2010). Study on the hypolipidemic effect and antioxidative activity of Hubei Qingzhuan tea. *Chaye Kexue*, 30(2), 124-128.

Escobar, M. S., Kaneko, H., & Funatsu, K. (2014). Flour concentration prediction using GAPLS and GAWLS focused on data sampling issues and applicability domain. *Chemometrics and Intelligent Laboratory Systems*, 137, 33-46. http://dx.doi.org/10.1016/j.chemolab.2014.06.005.

Esteban-Diez, I., Gonzalez-saiz, J. M., & Pizarro, C. (2004). An evaluation of orthogonal signal correction methods for the characterisation of arabica and robusta coffee varieties by NIRS. *Analytica Chimica Acta*, 514(1), 57-67. http://dx.doi.org/10.1016/j.aca.2004.03.022.

Fujiki, H., Suganuma, M., Okabe, S., Sueoka, E., Sueoka, N., Fujimoto, N., Goto, Y., Matsuyama, S., Imai, K., & Nakachi, K. (2001). Cancer prevention with green tea and monitoring by a new biomarker, hnRNP B1. *Mutation Research. Fundamental and Molecular Mechanisms of Mutagenesis*, 480-481, 299-304, 480-481. http://dx.doi.org/10.1016/S0027-5107(01)00189-0. PMid:11506822.

Inoue, M., Tajima, K., Mizutani, M., Iwata, H., Iwase, T., Miura, S., Hirose, K., Hamajima, N., & Tominaga, S. (2001). Regular consumption of green tea and the risk of breast cancer recurrence: follow-up study from the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC), Japan. *Cancer Letters*, 167(2), 175-182. http://dx.doi.org/10.1016/S0304-3835(01)00486-4. PMid:11369139.

Jian, L., Xie, L. P., Lee, A. H., & Binns, C. W. (2004). Protective effect of green tea against prostate cancer: a case-control study in southeast China. *International Journal of Cancer*, 108(1), 130-135. http://dx.doi.org/10.1002/ijc.11550. PMid:14618627.

Kachrimanis, K., Karamyan, V., & Malamataris, S. (2003). Artificial neural networks (ANNs) and modeling of powder flow. *International Journal of Pharmaceutics*, 250(1), 13-23. http://dx.doi.org/10.1016/S0378-5173(02)00528-8. PMid:12480269.

Lee, M. J., Seo, D. Y., Lee, H. E., Wang, I. C., Kim, W. S., Jeong, M. Y., & Choi, G. J. (2011). In line NIR quantification of film thickness on pharmaceutical pellets during a fluid bed coating process. *International Journal of Pharmaceutics*, 403(1-2), 66-72. http://dx.doi.org/10.1016/j.ijpharm.2010.10.022. PMid:21035529.

Lu, W. Z. (2010). *Modern near infrared spectroscopy analytical technology* (2nd ed.). Beijing: Sinopec Press.

Nakachi, K., Matsuyama, S., Miyake, S., Suganuma, M., & Imai, K. (2000). Preventive effects of drinking green tea on cancer and cardiovascular disease: epidemiological evidence for multiple targeting prevention. *BioFactors*, 13(1-4), 49-54. http://dx.doi.org/10.1002/biof.5520130109. PMid:11237198.

Setiawan, V. W., Zhang, Z. F., Yu, G. P., Lu, Q. Y., Li, Y. L., Lu, M. L., Wang, M. R., Guo, C. H., Yu, S. Z., Kurtz, R. C., & Hsieh, C. C. (2001). Protective effect of green tea on the risks of chronic gastritis and stomach cancer. *International Journal of Cancer*, 92(4), 600-604. http://dx.doi.org/10.1002/ijc.1231. PMid:11304697.

Shi, J. Y., Zou, X. B., Zhao, J. W., & Mao, H. P. (2011). Selection of wavelength for strawberry NIR spectroscopy based on BiPLS combined with SAA. *Hongwai Yu Haomibo Xuebao*, 30(5), 458-462. http://dx.doi.org/10.3724/SP.J.1010.2011.00458.

Tavanaie, M. A., Esmaeilian, N., & Mojtahedi, M. R. M. (2015). Olive hue visible-near infrared camouflage properties of high speed melt spun poly (ethylene terephthalate) multifilament yarn. *Dyes and Pigments*, 114, 267-272. http://dx.doi.org/10.1016/j.dyepig.2014.11.019.

Wang, S. P., Gong, Z. M., Zheng, P. C., Liu, P. P., Teng, J., Gao, S. W., & Gui, A. H. (2020a). Selection of characteristic near infrared spectra and establishment of prediction model for Qingzhuan tea quality. *Shipin Kexue*, 41(18), 283-287.

Wang, S. P., Teng, J., Zheng, P. C., Liu, P. P., Gong, Z. M., Gao, S. W., Gui, A. H., Ye, F., Wang, X. P., & Zheng, L. (2020b). Optimizing processing pressure of Qingzhuan tea and development of GCG models for near infrared spectroscopy detection. *Nongye Gongcheng Xuebao*, 36(8), 271-277.

Wang, X. F., Bao, Y. F., Liu, G. L., Li, G., & Lin, L. (2012). Study on the best analysis spectral section of NIR to detect alcohol concentration based on SiPLS. *Procedia Engineering*, 29, 2285-2290. http://dx.doi.org/10.1016/j.proeng.2012.01.302.

Wu, C.-F., Wu, T., Hashmonay, R. A., Chang, S.-Y., Wu, Y.-S., Chao, C.-P., Hsu, C.-P., Chase, M. J., & Kagann, R. H. (2014). Measurement of fugitive volatile organic compound emissions from a petrochemical tank farm using open-path Fourier transform infrared spectrometry. *Atmospheric Environment*, 82, 335-342. http://dx.doi.org/10.1016/j.atmosenv.2013.10.036.

Xia, T. (2016). *Tea processing* (1st ed). Beijing: China Agriculture Press.

Xuemei, L., & Jianshe, L. (2014). Using short wave visible-near infrared reflectance spectroscopy to predict soil properties and content. *Spectroscopy Letters*, 47(10), 729-739. http://dx.doi.org/10.1080/00387010.2013.840315.

Yamagata, K., Tagami, M., & Yamori, Y. (2015). Dietary polyphenols regulate endothelial function and prevent cardiovascular disease. *Nutrition*, 31(1), 28-37. http://dx.doi.org/10.1016/j.nut.2014.04.011. PMid:25466651.

Yan, S. S., Shao, H. J., Zhou, Z. H., Wang, Q., Zhao, L., & Yang, X. (2018). Non-extractable polyphenols of green tea and their antioxidant, anti-α-glucosidase capacity, and release during in vitro digestion. *Journal of Functional Foods*, 42, 129-136. http://dx.doi.org/10.1016/j.jff.2018.01.006.

Yang, X. H., Huang, M. J., Qin, C. Q., Lv, B. Y., Mao, Q. L., & Liu, Z. H. (2017). Structural characterization and evaluation of the antioxidant activities of polysaccharides extracted from Qingzhuan brick tea. *International Journal of Biological Macromolecules*, 101, 768-775. http://dx.doi.org/10.1016/j.ijbiomac.2017.03.189. PMid:28373046.

Zhou, H. F., Wang, Z. Y., Ma, H. M., Wang, S. Y., Xie, W. J., Chen, Y., Xie, C., Guo, A. R., Wang, C., & Zheng, M. (2022a). Characterisation and antioxidant activity of polysaccharide iron (III) complex in Qingzhuan Dark Tea. *Food Science and Technology*, 42, e119421. http://dx.doi.org/10.1590/fst.119421.

Zhou, H., Wang, S., Wang, Z., Xie, W., Wang, C., & Zheng, M. (2022b). Prepparation, characterization and antioxidant activity of polysaccharides selenides from Qingzhuan Dark Tea. *Food Science and Technology*, 42, e108421. http://dx.doi.org/10.1590/fst.108421.

Zhou, W. L., Xu, J. F., & Xu, L. (2008). *Determination of total polyphenols and catechins content in tea.* Beijing: China Light Industry Press.

Zhou, X. F., Yang, Z. L., Haughey, S. A., Galvin-King, P., Han, L. J., & Elliott, C. T. (2015). Classification the geographical origin of corn distillers dried grains with solubles by near infrared reflectance spectroscopy combined with chemometrics: a feasibility study. *Food Chemistry*, 189, 13-18. http://dx.doi.org/10.1016/j.foodchem.2014.09.104. PMid:26190595.