Original Article
Food Science and Technology

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

DOI: https://doi.org/10.1590/fst.104422

Quickly determination of sesame lignans in sesame oil using a portable near-infrared spectrometer

Wen-Cong JIAO[1], Ya-Ke LI[1], Mian JIA[1], Dong-Min WANG[1]*, Kun QI[2], Xue-De WANG[1]*

Abstract

Sesame oil is one of the most commonly used oils in life. It contains a special antioxidant substance called sesame lignans, which has high nutritional value and pharmacological activity. Chromatographic methods are accurate and reliable but not suitable for quality control due to time-consuming. Near-infrared (NIR) quantitative models for quickly determining sesamin, sesamolin and sesame lignans in sesame oil were built using a portable NIR spectrometer combined partial least squares(PLS) algorithm, and the optimal PLS models were developed by comparing the performance of the models with different spectral pretreatment methods and bands selection, the correlation coefficients (R^2) were R_C^2 = 0.98, 0.99, 0.99, R_P^2 = 0.99, 0.97, 0.94, and the root mean square error (RMSEP) was 2.69 μg/mL, 3.73 μg/mL, and 7.96 μg/mL, respectively. The acceptable results demonstrated that portable NIR spectrometer could be used for monitoring the contents of sesamin, sesamolin, and sesame lignans in sesame oil during the production process to carry out quality control.

Keywords: portable NIR; sesame oil; sesame lignans; pretreatment methods.

Practical Application: The three quantitative models can be used to quickly detect sesamin, sesamolin, and sesame lignans in sesame oil by portable Near-Infrared spectrometer.

1 Introduction

Sesame is one of the oldest oil crops known to mankind. Sesame oil is rich in nutrients, contains sesame lignans, vitamin E, and other substances, has a strong aroma and long shelf life, and is loved by consumers (Das et al., 2019; Shen et al., 2020, Nikzad et al., 2021). Sesame lignans are a general term for a class of compounds formed by the paired oxidation of ρ-hydroxyphenyl propane, whose structural features include 3,4 -methylenedioxyphenyl. Sesame lignans mainly consist of sesamin, sesamolin, sesamol, and asarinin, among which sesamin and sesamolin are the most abundant, with sesamin content of 0.4%-0.8% in sesame oil and sesamolin content of 0.2%-0.4% in sesame oil (Andargie et al., 2021). In recent years, sesame lignans have received worldwide attention for their super antioxidant effects and significant health benefits, such as promoting ethanol metabolism or liver detoxification, regulating blood lipids, anti-cancer properties, etc. For example, Majdalawieh et al. reported the in vitro and in vivo anticancer activity of sesamol from sesame lignans in several tumor cell lines and animal models, demonstrating potent anticancer properties of sesamol in vitro and in vivo (Majdalawieh & Mansour, 2019). Oikawa et al. reported that sesamin and its related lignans have inhibitory effects on the intestinal bacteria L-tryptophan indole-Lyase, and drugs based on this machine can be made for the treatment of chronic kidney disease (Oikawa et al., 2022). Aslam et al. reported that the increase of sesame lignans had a lipid-lowering effect on rats (Aslam et al., 2021). As a result, many sesame oil products have started to label the sesame lignans content, using the level of content to indicate the high quality of sesame oil in order to attract consumers.

The current methods for the analytical determination of lignans in sesame oil include gas chromatography (GC) (Schwertner & Rios, 2012; Tashiro), gas chromatography-mass spectrometry (GC-MS) (Qadir et al., 2018), HPLC (Mikropoulou et al., 2019; Moazzami et al., 2010; Shi et al., 2018), and thin layer chromatography (TLC) (Dar et al., 2015). Chromatography methods have high accuracy, but there are some problems such as long detection time, complex operation, and harmful reagents, among which the HPLC method is the most commonly used. These problems make it difficult for them to measure expediently, quickly and nondestructively, and it is difficult to popularize the characteristics of complex operations. Thus, a straightforward, portable, quick, and nondestructive approach must be developed to determine the amount of lignans in sesame seeds, and modern spectroscopic techniques correspond with these advantages.

Past developments have seen the utilization of molecular spectroscopy in modern spectral analysis technology has been comprehensively developed. Although the vibration of each atom in a molecule is very complicated, information can be gained by analyzing the infrared spectrum of compounds to many reflect the information about the molecular structure used for the determination of the molecular structure of compounds and the identification of unknown materials and mixtures. It can also be inferred to determine the content of components

Received 05 Oct., 2022
Accepted 07 Dec., 2022
[1]*College of Food Science and Engineering, Henan University of Technology, Zhengzhou, China*
[2]*Henan Anyang Mantianxue Protein Co. Ltd., Henan Anyang, China*
*Corresponding author: wdongmin@126.com; wangxuede1962@126.com

in the mixture from the intensity of characteristic absorption peaks. In addition, it can also determine the bond length and angle of the molecule, thereby inferring the 3D configuration of the molecule and judging the strength of the chemical bond. Among them, the application of portable NIR spectroscopy in the quantitative detection of substances is very broad due to its maturity, convenience, and low cost, and a large number of research reports have been accumulated. NIR spectroscopy is widely used in various fields of food, such as cereals and tubers (Aykas et al., 2020; Chadalavada et al., 2022; Su et al., 2020; Tilahun et al., 2020), animal foods (Diaz-Olivares et al., 2020; Lin et al., 2010; Rahim & Ghazali, 2012), legumes and their products (Natcha & Panmanas, 2014; Plans et al., 2014; Szigedi et al., 2013), vegetables and fruits (Malvandi et al., 2022; Minas et al., 2021; Sirisomboon, 2018), and beverages (Wang et al., 2022) etc. Zhenzhen Xia (Xia et al., 2020) and others have verified the feasibility of NIR spectroscopy technology for the prediction of sesamin and sesamolin in sesame seeds. However, the NIR spectrometers they use are all large instruments used in laboratories, which are expensive and cannot meet the demand of the public for simple, portable, and cheap use.

Therefore, in this study, a portable NIR spectrometer was used to establish PLS models for the determination of sesamin, sesamolin, and sesame lignans content in sesame oil. Although NIR spectroscopy is convenient and quick, due to the poor sensitivity of the spectrum, stoichiometry methods are needed to optimize the model after the spectral acquisition, such as spectra pretreatment and variable selection. Consequently, in order to improve the PLS models, the first derivative (1st), second derivative (2nd), and standard normal variable transformation (SNV) spectral pretreatment methods were performed in this study. Based on the spectral information of sesame oil and the structure of sesame lignans, the spectral variables were selected. The results obtained by the quantitative model were compared with those obtained by HPLC to evaluate the established model detection method.

2 Materials and methods

2.1 Reagents and samples

Sesamin (purity > 98%), sesamolin (purity > 98%), sesamol (purity > 98%) and asarinin (purity > 98%) were purchased from Macklin Reagent Network. Collect 60 kinds of sesame oil (purchased from the market respectively, including three kinds of sesame oil with different processes of cold pressing, hot pressing, and small grinding, and homemade sesame oil with different microwave power and time). Methanol (chromatographic grade, purchased from Honeywell Trading (Shanghai) Co.)

2.2 HPLC analysis

The standard curves of sesamin, sesamolin, sesamol, and asarinin were first established in HPLC. Appropriate amounts of each of the four standards were taken and prepared with chromatographic grade methanol to form 224 μg/mL of sesamin, 85 μg/mL of sesamin, 237 μg/mL of sesamol, and 104 μg/mL of asarinin. Then the master mixes were diluted separately and subjected to liquid chromatographic detection, and the results were plotted as standard curves for sesamin, sesamolin, sesamol, and asarinin.

Detection conditions for HPLC:

a) Chromatographic column: Sunfire C18 reversed-phase column (250 × 4.6 mm, 5 μm);

b) Injection volume: 20 μL;

c) Mobile phase: methanol: water = 70:30 (V/V);

d) Flow rate: 0.8 mL/min;

e) Column temperature: 30 °C;

f) UV detector wavelength: UV detector, detection wavelength 287 nm.

Extraction of sesame lignans from the sample: 0.2 g of oil sample was accurately weighed in a 10 mL white plastic centrifuge tube and extracted with chromatographic methanol three times. First, add 4 mL of chromatographic methanol, vortex for 5 min, sonicate for 5 min, freeze for 5 min, centrifuge at 4000 r/min for 5 min, and transfer the supernatant to a 10 mL volumetric flask. Repeat the above steps with 3 mL of chromatographic methanol and 2 mL of chromatographic methanol. Finally, the volumetric flask was filled with chromatographic methanol and the filter membrane was pumped into the liquid flask for machine determination. The extracted samples were detected by HPLC, and the contents of sesamin, sesamolin, sesamol, and asarinin were calculated from the standard curve for each sample, and then the four substances were summed up to obtain the content of sesame lignans.

2.3 NIR spectroscopy detection

The AvaSpec NIR256-1.7 TEC spectrometer, tungsten halogen light source, color matching tube frame, and two optical fibers were used (Figure 1). The signal source interface is a standard SMA905 interface to connect the optical probe, and the light source signal is calibrated by the spherical mirror. Spectra were collected in a colorimetric dish by the transmission principle.

The required sesame oil samples were taken out of the refrigerator and put on the experimental bench until the temperature reached room temperature of 15 °C, at which point the samples were taken. The oil sample was added to a quartz cuvette with an optical length of 10 mm and a capacity of 10 mL. In order to avoid stray light interfering with the measurement, the sample to be measured was put into place and then covered with a sample cover before its spectrum was collected. When the measurement of a sample was finished, the sample was dumped out and 3 rinses were performed with the next sample to be measured. The NIR spectrometer used an integration time of 0.470 ms, and each spectrum was obtained by averaging 100 repeated measurements. Each spectrum was in the wavelength range of 1100-1700 nm and consisted of 203 points.

2.4 NIR spectra pretreatment and modeling

In this study, we used PLS to construct NIR models of sesamin, sesamolin, and sesame lignans contents in sesame oil, and we used the model to predict the sesamin, sesamolin, and sesame lignans contents in the validation set and compare them with the actual values determined by HPLC. The calibration set and

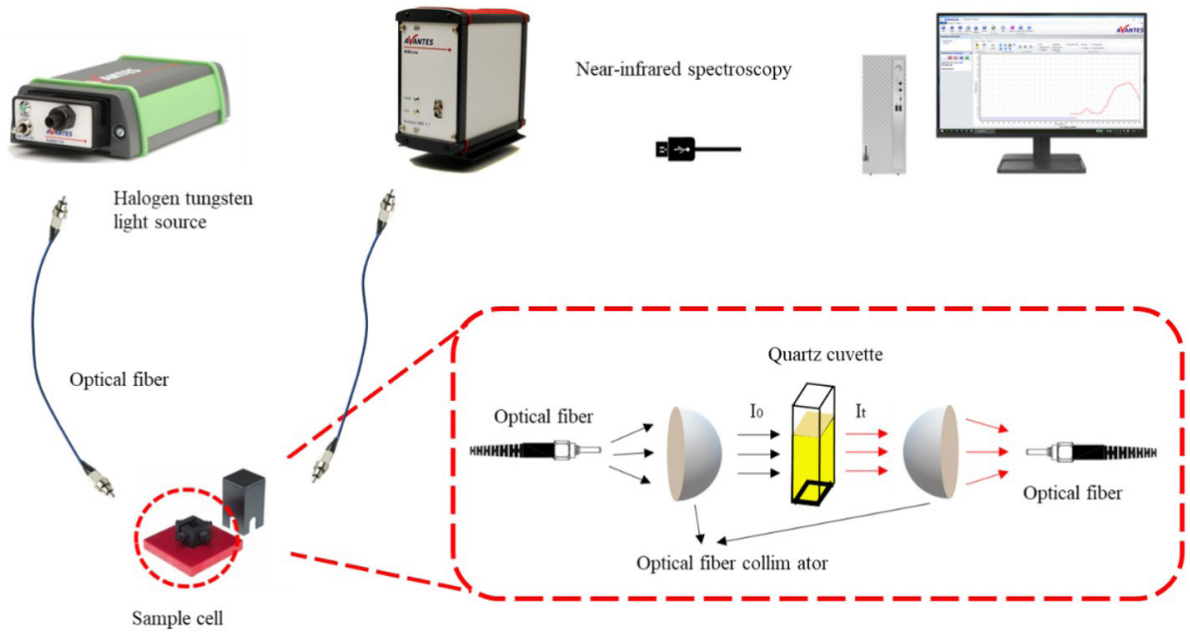

Figure 1. Portable NIR spectrometer.

validation set were created by randomly dividing the dataset's samples. The spectra also contain unimportant information and noise, such as sample background, astigmatism, and electrical noise, except as the chemical data of the samples themselves. The spectrum of the 1st and 2nd derivatives spectral analysis is commonly used in pretreatment methods. The 1st shows the rate of change of the spectrum, the 2nd shows the change of the whole spectrum rate, which can effectively eliminate the interference of the baseline and another background, to distinguish the overlapping peaks, and improve the sensitivity and resolution, but can introduce noise at the same time, reducing the signal-to-noise ratio. In addition, the 2nd has more baseline noise than the 1st. For example, for a discrete spectrum x_k, the first and second derivative spectra at wavelength k with a different width g are calculated according to the following equation, respectively (Equations 1, 2 and 3). The SNV is mainly used to eliminate the influence of sample surface scattering and optical path variation on NIR spectra. SNV is the raw spectrum after subtracting the average of the spectrum, then dividing by the standard deviation of the data, essentially normalizing the raw spectral data standard. The difference from the standardized algorithm is that the SNV algorithm processes each spectrum and computes the following formula for spectra requiring SNV transformation. Therefore, the 1st, 2nd, and SNV spectra pretreatment methods were selected to be combined with the original spectra of samples, respectively, to establish PLS models for comparison, and to select the optimal pretreatment method.

$$x_{k,1st} = \frac{x_{k+g} - x_{k-g}}{g} \quad (1)$$

$$x_{k,2nd} = \frac{x_{k+g} - 2x_k + x_{k-g}}{g^2} \quad (2)$$

$$x_{SNV} = \frac{x - \bar{x}}{\sqrt{\frac{\sum_{k=1}^{m}(x_k - \bar{x})^2}{(m-1)}}} \quad (3)$$

Where $\bar{x} = \frac{\sum_{k=1}^{m} x_k}{m}$, m is the number of wavelength points, k = 1, 2, ..., m.

The performance and stability of the prediction model were evaluated by correlation coefficient (R^2), root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP) and the ratio of prediction to deviation (RPD). R^2 is used to evaluate the correlation between the predicted value of the sample and the true value, and the closer the value is to 1, the better the prediction, given the same concentration range. RMSECV and RMSEP are used to evaluate the predictive power of the model, and the smaller their values, the stronger the predictive power of the model. The RPD is the ratio of the standard deviation (SD) of the validation set to the RMSEP. The smaller the SD is, the wider and more uniform the dispersion of the validation set samples will be, and the larger the RPD value will be. According to the same concentration range assumption, the model's prediction accuracy increases with RPD size, and it is typically thought that the RPD should be more than 2 for the forecast results to be regarded as acceptable.

3 Results and analysis

3.1 Division of calibration and validation sets

From the HPLC analysis, it is clear that the homemade ones have a larger range of composition than the market-purchased samples, expanding the range of detection and

increasing the applicability (Table 1). From the analysis of the raw spectrogram (Figure 2), sample 27 was found to be turbid or even granular during the assay, which eventually resulted in a large absorbance and the following spectral results occurred. It was also found that the color difference between cold pressed, hot pressed, and water extraction sesame oil had no significant effect on the spectrum. Subsequently, 60 samples were randomly divided into 50 calibration groups and 10 validation groups.

The concentration content of these three components was different, so the samples for calibration and validation sets were selected differently. The coverage of sesamin, sesamolin, and sesame lignans distribution in the sesame oil calibration set samples was larger than the variation range of the validation set, indicating that the constructed PLS model for the calibration set samples could be better applied to the validation set samples (Table 2).

3.2 Modeling and optimization

Spectral pretreatment

Before establishing PLS models of sesamin, sesamolin, and sesamin lignans in sesame oil, the spectrum of sesame oil was pretreated with different combinations of 1st, 2nd, and SNV. These pretreatment methods can reduce the color difference obtained by the three processes and the problem of sesame oil 27 sample detection by different means. R^2, RMSECV, RMSEP, and RPD were used to assess the models, and the results of raw spectra were added to facilitate comparison (Table 3).

For sesamin, R^2 of the model generated by the raw spectrum was $R_C^2 = 0.97$, $R_P^2 = 0.99$, RMSECV, RMSEP, and RPD were 13.08, 2.43, and 8.54, respectively. The whole data set became worse after SNV treatment of the raw spectrum, and the results of the other two combinations did not improve greatly, probably because the special signal of sesamin was less obvious after SNV treatment, so its model is not suitable for SNV treatment.

Table 1. Concentration of sesame lignans.

samples	sesamin / (μg /mL)	sesamolin / (μg /mL)	sesame lignans / (μg /mL)
Purchased from market	114.48-140.37	34.40-67.05	154.31-217.95
Homemade samples	107.34-178.73	11.41-87.20	143.28-271.37

After pretreatment with 1st and 2nd, the results of the 1st were worse, but the results of the 2nd were better, where R^2 did not change significantly, RMSECV decreased significantly from 13.08 to 7.83, while RMSEP and RPD only changed slightly. Therefore, the 2nd pretreatment method should be selected for the sesamin PLS model.

The PLS model result for the raw spectra of sesamolin was $R_C^2 = 0.93$, $R_P^2 = 0.95$, RMSECV, RMSEP, and RPD were 9.36, 5.23, and 4.49, respectively. Compared with the raw spectrum, the PLS model results after pretreatment were optimized to a certain extent. In particular, the four evaluation indexes of the results after SNV pretreatment were the best among the six model results. R_C^2 increased from 0.93 to 0.99, R_P^2 increased from 0.95 to 0.97, and RMSEP decreased from 5.23 to 3.59. RPD increased from 4.49 to 6.55. After the other four kinds of pretreatment, the relevant indicators of the correction set were significantly improved, but the RPD value and relevant indicators of the validation set were decreased to a certain extent, which did not improve the predictive ability of the model. And their results were all similar, indicating that 1st and 2nd pretreatment did not increase the sesamolin signal in the spectrum. Thus, the optimal spectral pretreatment method in the sesamolin PLS model is the SNV pretreatment method.

For the raw spectra of sesame lignans, PLS model results were $R_C^2 = 0.98$, $R_P^2 = 0.92$, RMSECV, RMSEP, and RPD were 12.94, 8.04, and 4.06, respectively. The comparison of the results after the five pretreatments revealed that their R_C^2 was

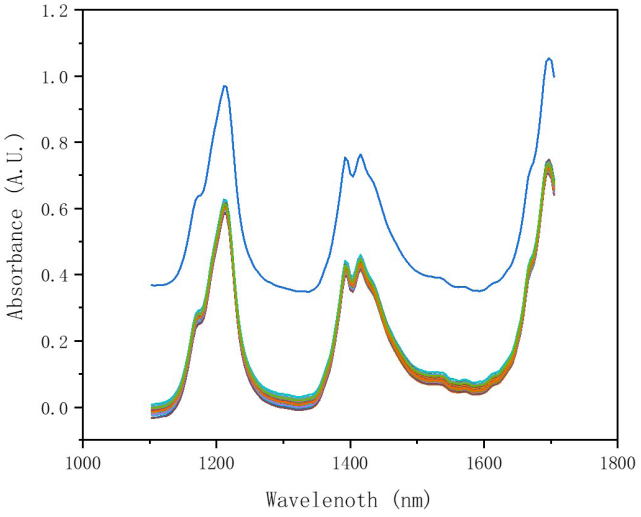

Figure 2. Raw NIR spectra.

Table 2. Distribution of sesamin, sesamolin and sesamin lignans in calibration and validation sets.

Samples	sesamin /(μg /mL)				sesamolin /(μg/mL)				sesame lignans /(μg/mL)			
	Max	Min	Average	SD	Max	Min	Average	SD	Max	Min	Average	SD
Totality	178.73	107.34	133.12	18.98	87.20	11.41	54.01	21.93	271.37	143.28	193.76	37.71
Calibration	178.73	107.34	131.68	18.49	87.20	11.41	54.75	21.78	271.37	143.28	193.94	38.94
Validation	167.09	108.26	140.34	20.75	79.34	16.41	50.30	23.51	249.40	155.51	192.84	32.66

Table 3. Results of different pretreatment methods.

Lignans	Methods	Calibration Set		Validation Set		RPD
		R^2	RMSECV	R^2	RMSEP	
Sesamin	Raw	0.97	13.08	0.99	2.43	8.54
	SNV	0.85	10.58	0.82	8.66	2.40
	1st	0.91	8.84	0.89	6.78	3.06
	1st+SNV	0.91	9.25	0.89	6.55	3.17
	2nd	0.98	7.83	0.98	3.36	6.17
	2nd+SNV	0.97	7.90	0.98	3.34	6.21
Sesamolin	Raw	0.93	9.36	0.95	5.23	4.49
	SNV	0.99	10.18	0.97	3.59	6.55
	1st	0.99	7.49	0.93	5.89	3.99
	1st+SNV	0.99	6.53	0.92	6.40	3.67
	2nd	0.98	7.78	0.94	5.90	3.98
	2nd+SNV	0.98	7.85	0.93	6.16	3.82
Sesame lignans	Raw	0.98	12.94	0.92	8.04	4.06
	SNV	0.99	9.58	0.94	8.36	3.91
	1st	0.99	7.21	0.94	7.96	4.10
	1st+SNV	0.99	6.67	0.94	7.96	4.10
	2nd	0.99	8.42	0.89	13.70	2.38
	2nd+SNV	0.99	8.94	0.91	12.10	2.70

0.99, and the smallest of the RMSECV was 6.67 after the 1st+SNV pretreatment, so the optimal result in the correction set was the 1st+SNV pretreatment method. Only SNV, 1st, and 1st+SNV pretreatment had the highest R_P^2 of 0.94 in the validation set; 1st and 1st+SNV pretreatment had the lowest RMSEP of 7.96, and both methods had the highest RPD value of 4.10. Only the RMSECV values decreased in the 1st+SNV pretreatment compared with the 1st pretreatment alone, while all other indices remained unchanged, indicating that SNV further optimized the model, so the 1st+SNV pretreatment method is the best choice for the sesame lignans PLS model pretreatment method.

Spectral band and model master factor number selection

When using a full spectrum to build a spectral model, a large amount of data and irrelevant information can affect and interfere with the model, resulting in slow processing speed, which leads to the model not being able to calculate predicted values more accurately and cannot meet the fast-paced portable inspection of sesame oil production. Therefore, the band regions associated with sesamin, sesamolin, and sesame lignans should be selected based on the sesame oil spectral information, and the model should be further optimized based on preprocessing to select the optimal band among available bands. Sesamin and sesamolin have similar chemical structural characteristics, and sesame lignans are the general term for this class of compounds, whose structural characteristics include 3, 4-methylene dioxophenyl. Their IR spectral information contains the stretching vibration characteristic peak of the carbon-oxygen carbon bond C-O-C 1050-1250 nm, the in-plane bending vibration peak of two adjacent hydrogens on the benzene ring 1000-1300 nm, and the vibration absorption peak of the benzene ring skeleton 1450-1600 nm. From the raw spectrogram (Figure 2), three peaks located around 1200, 1400, and 1700 nm can be seen, and from the basic chemical structure of sesame lignans, all these three peaks may be related to it. Accordingly, the spectra should be divided into three regions: 1100-1300 nm, 1300-1500 nm, and 1500-1700 nm, and the regions that are more suitable for the PLS model of the three related substances should be explored respectively. Based on the obtained optimal pretreatment method, the three regions are combined for modeling and compared with each other to select the optimal modeling region (Table 4). Meanwhile, the table also shows the optimal number of principal factors for each final PLS model. The selection of the quantity of principal factors of the model is directly related to the actual predictive power. Using too few master factors will not fully express the spectral information of the samples; using too many master factors will add noise and reduce the actual prediction ability of the model. As a result, one of the efficient ways to utilize spectral information and filter out noise is to determine the quantity of principal factors participating in model building reasonably. The method of selecting the optimal number of principal factors here is to determine the optimal number of factors in the model by RMSECV.

Table 4. Results for different spectral band regions.

Lignans	Spectral band /(nm)	Factor	Calibration Set		Validation Set		RPD
			R^2	RMSECV	R^2	RMSEP	
Sesamin	1100-1300	7	0.88	15.67	0.82	9.15	2.27
	1300-1500	10	0.97	12.51	0.93	5.24	3.96
	1500-1700	7	0.94	9.40	0.95	4.39	4.73
	1100-1500	8	0.96	14.64	0.93	5.22	3.97
	1100-1300+ 1500-1700	8	0.96	9.10	0.99	2.93	7.08
	1300-1700	12	0.98	7.79	0.99	2.69	7.71
	1100-1700	10	0.98	7.83	0.98	3.36	6.17
Sesamolin	1100-1300	11	0.98	10.45	0.86	8.41	2.80
	1300-1500	9	0.95	12.13	0.88	10.1	2.33
	1500-1700	22	0.99	7.09	0.95	4.97	4.73
	1100-1500	4	0.89	14.61	0.93	6.28	3.74
	1100-1300+ 1500-1700	27	0.99	5.71	0.96	4.76	4.94
	1300-1700	26	0.99	7.68	0.97	3.73	6.30
	1100-1700	26	0.99	10.18	0.97	3.59	6.55
Sesame lignan	1100-1300	17	0.99	11.94	0.95	6.83	4.78
	1300-1500	7	0.98	15.60	0.85	13.8	2.37
	1500-1700	8	0.99	7.38	0.94	8.85	3.69
	1100-1500	11	0.99	13.54	0.94	8.36	3.91
	1100-1300+ 1500-1700	4	0.99	7.82	0.95	7.89	4.14
	1300-1700	10	0.99	7.08	0.94	8.26	3.95
	1100-1700	10	0.99	6.67	0.94	7.96	4.10

The PLS models of sesamin in different regions showed the worst results at 1100-1300 nm, indicating that this region was not characterized for sesamin. All aspects of the result index are optimal at 1300-1700 nm, where the RMSEP decreases significantly and the RPD increases significantly in comparison with the full band. The results of the validation set are optimized at the same time as the calibration set optimization, which improves the prediction ability and reduces the used band. Therefore, the sesamin PLS model selects the optimal region between 1300 and 1700 nm, and its principal factor number is 12, $R_C^2 = 0.98$, $R_P^2 = 0.99$, RMSECV, RMSEP, and RPD are 7.79, 2.69, and 7.71, respectively.

Compared to the results of the sesamolin PLS model in different regions, R_C^2 reached 0.99 at 1500-1700 nm, 1100-1300+1500-1700 nm, 1300-1700 nm, and the full band. The RMSECV of the other three models was significantly lower than that of the full band, while their RMSEP and RPD were poor. R_P^2 is only highest at 1300-1700 nm and full band, so the 1300-1700 nm region modeling is closer to the results of full band modeling. In the 1300-1700 nm and full-band modeling results, the RMSECV of the 1300-1700 nm model was significantly improved, and the RPD value decreased slightly. Therefore, the 1300-1700 nm region should be selected for the sesamolin PLS model, and the principal factor number is 26.

The calibration set results of the PLS model in each region of sesame lignans showed that R_C^2 was similar, and there were 4 regions with RMSECV values around 7, among which the minimum value of the whole band was 6.67. The R_P^2 difference of the validation set results of the above four models were small, and the RMSEP value was only about 1100-1300+1500-1700 nm, with a small difference, as was the RPD value. However, the RMSECV of the full-band PLS model was significantly less than 1100-1300+1500-1700 nm, so the full-band region with a principal factor number of 10 was selected for the sesame lignans PLS model. This may be because sesame lignans contain a variety of substances, and the modeling of the whole band region is more complete than that of other combinations, so the obtained results are more accurate. Finally, the prediction plots and regression residuals of the determination of PLS models based on sesamin, sesamolin, and sesame lignans from the NIR spectra are plotted (Figure 3).

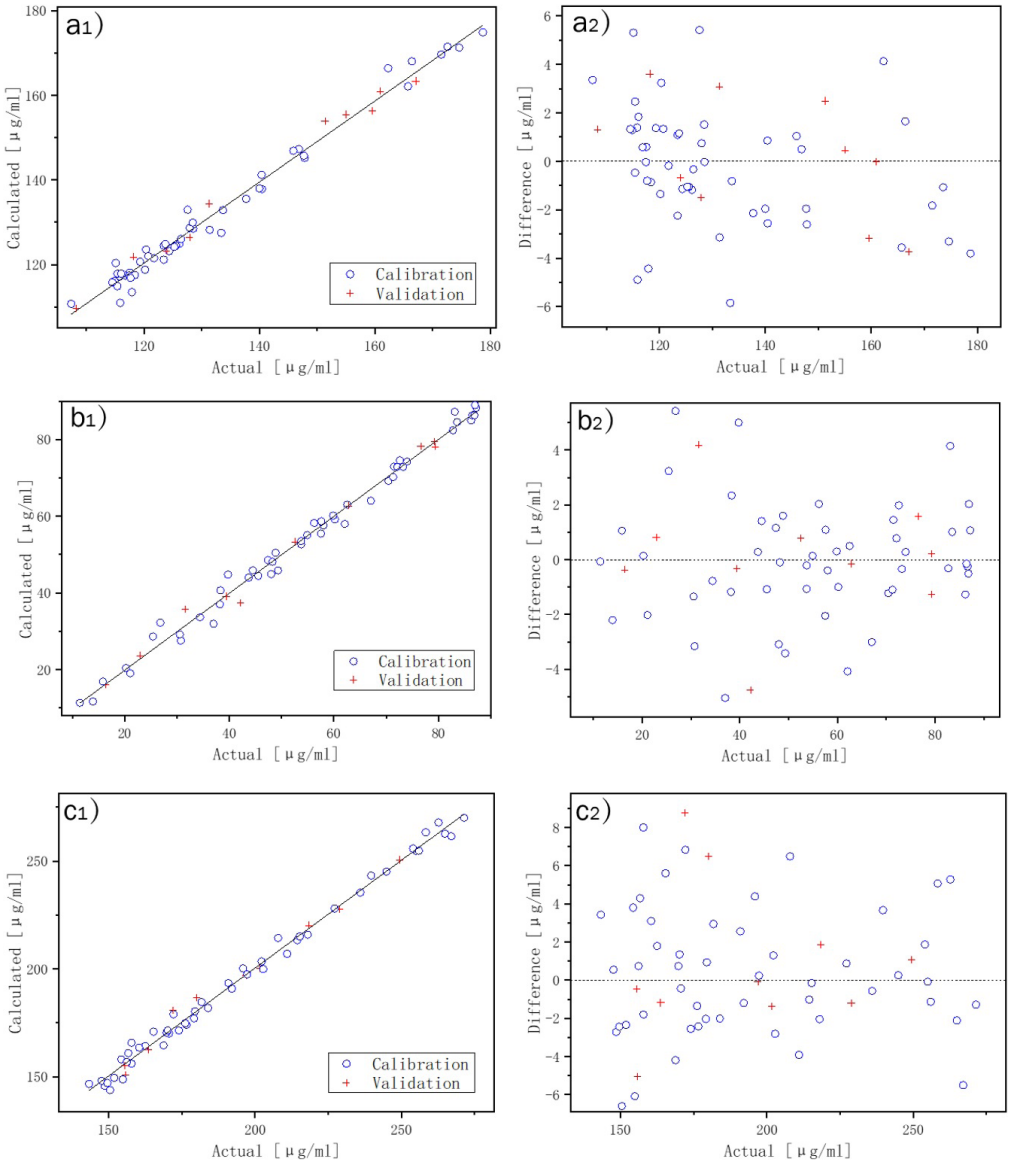

Figure 3. Predicted plots and regression residuals for sesamin (a), sesamolin (b), and sesame lignans (c) based on NIR spectra.

4 Conclusion

In this study, a portable NIR spectrometer combined with the PLS modeling method was used to establish the quantification of sesamin, sesamolin, and sesame lignans in sesame oil in the spectral range of 1100-1700 nm. For optimizing the PLS model, different spectral preprocessing methods (1st, 2nd, SNV, and spectral bands were used), as well as the choice of model master factor number. For the sesamin, sesamolin, and sesame lignans PLS models, the pretreatment conditions were 2nd, SNV, and 1st+SNV, respectively, and the spectral bands were 1300-1700 nm, 1300-1700 nm, and 1100-1700 nm, respectively, and the principal factor numbers were 12, 26, and 10, respectively, which gave better results. The concentrations of sesamin, sesamolin, and sesame lignans in the predicted samples were calculated using the three optimal PLS models, and the findings showed that they were reasonably similar to the values found in the HPLC analysis. As a result, the portable NIR spectrometer can realize the nondestructive, simple, and convenient detection of sesamin, sesamolin, and sesame lignans content in sesame oil, which is more convenient and cheaper than the large-scale NIR spectrometer in the laboratory and has a great application prospect.

Conflict of interest

The authors affirm that they have no known financial or interpersonal conflicts that would have appeared to have an impact on the research presented in this study.

Acknowledgements

This work was supported by the Key Project of Science and Technology of Henan Province (201300110600) and the project

number of "China Agricultural Research System funded by the Ministry of Agriculture" CARS-14, and the authors are thankful to the sponsors.

References

Andargie, M., Vinas, M., Rathgeb, A., Möller, E., & Karlovsky, P. (2021). Lignans of sesame (Sesamum indicum L.): a comprehensive review. *Molecules (Basel, Switzerland)*, 26(4), 883. http://dx.doi.org/10.3390/molecules26040883. PMid:33562414.

Aslam, M., Shabbir, M. A., Pasha, I., Shukat, R., Siddique, U., Manzoor, M. F., & Ayub, S. (2021). Protective effect of sesame (sesamum indicum) seed oil against hypercholesterolemic in sprague-dawley male rats. *Food Science and Technology (Campinas)*, 41(Suppl 2), 741-745. http://dx.doi.org/10.1590/fst.35320.

Aykas, D. P., Ball, C., Menevseoglu, A., & Rodriguez-Saona, L. E. (2020). In situ monitoring of sugar content in breakfast cereals using a novel FT-NIR spectrometer. *Applied Sciences (Basel, Switzerland)*, 10(24), 8774. http://dx.doi.org/10.3390/app10248774.

Chadalavada, K., Anbazhagan, K., Ndour, A., Choudhary, S., Palmer, W., Flynn, J. R., Mallayee, S., Pothu, S., Prasad, K. V. S. V., Varijakshapanikar, P., Jones, C. S., & Kholová, J. (2022). NIR instruments and prediction methods for rapid access to grain protein content in multiple cereals. *Sensors (Basel)*, 22(10), 3710. http://dx.doi.org/10.3390/s22103710. PMid:35632119.

Dar, A. A., Verma, N. K., & Arumugam, N. (2015). An updated method for isolation, purification and characterization of clinically important antioxidant lignans – Sesamin and sesamolin, from sesame oil. *Industrial Crops and Products*, 64, 201-208. http://dx.doi.org/10.1016/j.indcrop.2014.10.026.

Das, D., Datta, A. K., Kumbhakar, D. V., Ghosh, B., Pramanik, A., & Grupta, S. (2019). Heterosis, lipid profile and lignan glycosides assessment in raised inter-varietal promising hybrids of sesame (Sesamum indicum L.; Pedaliaceae). *Industrial Crops and Products*, 127, 164-171. http://dx.doi.org/10.1016/j.indcrop.2018.10.046.

Diaz-Olivares, J., Adriaens, I., Stevens, E., Saeys, W., & Aernouts, B. (2020). Online milk composition analysis with an on-farm near-infrared sensor. *Computers and Electronics in Agriculture*, 178, 105734. http://dx.doi.org/10.1016/j.compag.2020.105734.

Lin, H., Zhao, J. W., Chen, Q. S., Cai, J. R., & Zhou, P. (2010). Identification of egg freshness using near infrared spectroscopy and one class support vector machine algorithm. *Guang pu xue yu guang pu fen xi = Guang pu*, 30(4), 929-932. PMid:20545133.

Majdalawieh, A. F., & Mansour, Z. R. (2019). Sesamol, a major lignan in sesame seeds (Sesamum indicum): anti-cancer properties and mechanisms of action. *European Journal of Pharmacology*, 855, 75-89. http://dx.doi.org/10.1016/j.ejphar.2019.05.008. PMid:31063773.

Malvandi, A., Feng, H., & Kamruzzaman, M. (2022). Application of NIR spectroscopy and multivariate analysis for Non-destructive evaluation of apple moisture content during ultrasonic drying. *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, 269, 120733. http://dx.doi.org/10.1016/j.saa.2021.120733. PMid:34920303.

Mikropoulou, E. V., Petrakis, E. A., Argyropoulou, A., Mitakou, S., Halabalaki, M., & Skaltsounis, L. (2019). Quantification of bioactive lignans in sesame seeds using HPTLC densitometry: comparative evaluation by HPLC-PDA. *Food Chemistry*, 288, 1-7. http://dx.doi.org/10.1016/j.foodchem.2019.02.109. PMid:30902268.

Minas, I. S., Blanco-Cipollone, F., & Sterle, D. (2021). Accurate non-destructive prediction of peach fruit internal quality and physiological maturity with a single scan using near infrared spectroscopy. *Food Chemistry*, 335, 127626. http://dx.doi.org/10.1016/j.foodchem.2020.127626. PMid:32739812.

Moazzami, A. A., Haese, S. L., & Kamal-Eldin, A. (2010). Lignan contents in sesame seeds and products. *European Journal of Lipid Science and Technology*, 109(10), 1022-1027. http://dx.doi.org/10.1002/ejlt.200700057.

Natcha, T., & Panmanas, S. (2014). Detection of adulteration of soy sauce by brine using near infrared spectroscopy. *Agriculture and Agricultural Science Procedia*, 2, 212-217. http://dx.doi.org/10.1016/j.aaspro.2014.11.030.

Nikzad, N., Ghavami, M., Seyedain-Ardabili, M., Akbari-Adergani, B., & Azizinezhad, R. (2021). Effect of deep frying process using sesame oil, canola and frying oil on the level of bioactive compounds in onion and potato and assessment of their antioxidant activity. *Food Science and Technology (Campinas)*, 41(3), 545-555. http://dx.doi.org/10.1590/fst.35819.

Oikawa, D., Yamashita, S., Takahashi, S., Waki, T., Kikuchi, K., Abe, T., Katayama, T., & Nakayama, T. (2022). (+)-Sesamin, a sesame lignan, is a potent inhibitor of gut bacterial tryptophan indole-lyase that is a key enzyme in chronic kidney disease pathogenesis. *Biochemical and Biophysical Research Communications*, 590, 158-162. http://dx.doi.org/10.1016/j.bbrc.2021.12.088. PMid:34974305.

Plans, M., Simó, J., Casañas, F., del Castillo, R. R., Rodriguez-Saona, L. E., & Sabaté, J. (2014). Estimating sensory properties of common beans (Phaseolus vulgaris L.) by near infrared spectroscopy. *Food Research International*, 56, 55-62. http://dx.doi.org/10.1016/j.foodres.2013.12.003.

Qadir, A., Ali, A., Arif, M., Al-Rohaimi, A. H., Singh, S. P., Ahmad, U., Khalid, M., & Kumar, A. (2018). Solvent extraction and GC-MS analysis of sesame seeds for determination of bioactive antioxidant fatty acid/fatty oil components. *Drug Research*, 68(6), 344-348. PMid:29232753.

Rahim, H. A., & Ghazali, R. (2012, March). The application of near-infrared spectroscopy for poultry meat grading. In *2012 IEEE 8th International Colloquium on Signal Processing and its Applications* (pp. 58-62). New York: IEEE. http://dx.doi.org/10.1109/CSPA.2012.6194691.

Schwertner, H. A., & Rios, D. C. (2012). Analysis of sesamin, asarinin, and sesamolin by HPLC with photodiode and fluorescent detection and by gc/ms: application to sesame oil and serum samples. *Journal of the American Oil Chemists' Society*, 89(11), 1943-1950. http://dx.doi.org/10.1007/s11746-012-2098-7.

Shen, W., Wang, H., Lu, H., Yu, K., Hu, G., Wei, X., & Wu, B. (2020). Determination of eight vitamin E in vegetable oils by gas chromatography-mass spectrometry and its application on authentication of sesame oil. *Se Pu*, 38(5), 595-599. [Chinese] [PMid:34213245.

Shi, L. K., Zheng, L., Xiang, Y. F., Liu, R. J., Chang, M., Jin, Q. Z., & Wang, X. (2018). A rapid method for simultaneous analysis of lignan and γ-tocopherol in sesame oil by using normal-phase liquid chromatography. *Journal of the American Oil Chemists' Society*, 95(1), 13-19. http://dx.doi.org/10.1002/aocs.12010.

Sirisomboon, P. (2018). NIR spectroscopy for quality evaluation of fruits and vegetables. *Materials Today: Proceedings*, 5(10), 22481-22486. http://dx.doi.org/10.1016/j.matpr.2018.06.619.

Su, W., Bakalis, S., & Sun, D. (2020). Chemometric determination of time series moisture in both potato and sweet potato tubers during hot air and microwave drying using near/mid-infrared (NIR/MIR) hyperspectral techniques. *Drying Technology*, 38(5-6), 806-823. http://dx.doi.org/10.1080/07373937.2019.1593192.

Szigedi, T., Fodor, M., Pérez-Marín, D., & Garrido-Varo, A. (2013). Fourier transform near-infrared spectroscopy to predict the gross energy content of food grade legumes. *Food Analytical Methods*, 6(4), 1205-1211. http://dx.doi.org/10.1007/s12161-012-9527-y.

Tilahun, S., An, H. S., Hwang, I. G., Choi, J. H., Baek, M. W., Choi, H. R., Park, D. S., & Jeong, C. S. (2020). Prediction of α-Solanine and α-Chaconine in potato tubers from hunter color values and VIS/NIR spectra. *Journal of Food Quality*, 2020, 1-9. http://dx.doi.org/10.1155/2020/8884219.

Wang, J. B., Wu, X. H., Zheng, J., & Wu, B. (2022). Rapid identification of green tea varieties based on FT-NIR spectroscopy and LDA/QR. *Food Science and Technology (Campinas)*, 42, e73022. http://dx.doi.org/10.1590/fst.73022.

Xia, Z., Yi, T., & Liu, Y. (2020). Rapid and nondestructive determination of sesamin and sesamolin in Chinese sesames by near-infrared spectroscopy coupling with chemometric method. *Spectrochimica Acta. Part A: Molecular and Biomolecular Spectroscopy*, 228, 117777. http://dx.doi.org/10.1016/j.saa.2019.117777. PMid:31727518.