Rapid and Non-destructive Determination of Oil Content of Peanut (*Arachis hypogaea L.*) Using Hyperspectral Imaging Analysis

Huali Jin[1] · Yongsheng Ma[2] · Linlin Li[1] · Jun-Hu Cheng[3]

Abstract Peanut (*Arachis hypogaea L.*) is rich in some important oils such as the high content of polyunsaturated fatty acids. The potential of hyperspectral imaging technique in the spectral range I (400–1000 nm) and II (1000–2500 nm) coupled with chemometrics analysis for predicting oil content in different peanut cultivars was investigated in this study. Hyperspectral images were obtained and the corresponding spectral data was extracted. Quantitative calibration models were established between pre-processing spectral data and the reference measured oil content by partial least squares regression (PLSR) analysis. By comparing the model performances based on different spectral pre-processing methods, the raw-PLSR models using full wavelengths presented better results with the determination coefficient (R^2_p) of 0.696 and 0.923, and root mean square errors by prediction (RMSEP) of 0.416 % and 0.208 %, respectively. In addition, six optimal wavelengths in the spectral range II were selected based on the regression coefficients of the established raw-PLSR model. The simplified PLSR model established only using identified optimal wavelengths also showed good performance with R^2_p of 0.934, and RMSEP of 0.197 %. The results demonstrated that hyperspectral imaging technique is a promising tool for rapid and non-destructive determination of oil content in peanut.

Introduction

Peanut (*Arachis hypogaea L.*) is one of the main commercial crops, which is universally cultivated especially in Asian countries. It has a high nutritional value, containing plentiful fat, protein, vitamins, sugar, and essential unsaturated fatty acid as well as the functional ingredients such as resveratrol and plant sterols. Because the nutritional contents of different peanut varieties are not exactly the same, it is used for oil production, peanut butter, candies, roasted peanuts and snack foods, extenders in meat product formulations, soups and desserts, and also for direct human consumption based on its dominant components (Rustom et al. 1996).

Peanut is rich in oil content (38–60 %) and relatively low in ash and carbohydrate compared with other oilseed crops (Shewfelt and Young 1977). Fatty acids, particularly oleic and linoleic, are the main unsaturated fatty acid in peanut oil (approximately 80 % of the total oil content), which have a large influence on the stability and nutritional quality of peanut oil (Holley and Hammons 1968; Sanders 1980; Brown et al. 1975). In addition, peanut is an important economical crop in the world trade of agricultural products, and about two thirds of peanut crops are crushed for oil. The value of oil content is an important criterion of monetary assessment in the trade of peanut, and the raw material price also depends on its oil content. Thus, selection of good quality varieties of peanut with high oil content is critical.

✉ Huali Jin
  jinhuali66@163.com

✉ Jun-Hu Cheng
  chengjunhu1229@163.com

[1] College of Food Science and Technology, Henan University of Technology, Zhengzhou 450000, China

[2] Food Department, Henan University of Animal Husbandry and Economy, Zhengzhou 450045, China

[3] College of Light Industry and Food Sciences, South China University of Technology, Guangzhou 510641, China

In order to evaluate the germplasm and screen the peanut varieties with different levels of oil content, the commonly used analytical methods for determining the oil content including the classical Soxhlet extraction, the American Oil Chemists' Society (AOCS) method (Firestone 1993), binary polar organic solvent extraction based on the Soxhlet extraction method (Folch et al. 1957; Bligh and Dyer 1959), and supercritical fluid extraction (SFE) method (King and O'Farrel 1997) have been developed. These above-mentioned methods provide an effective and accurate measurement of oil content for peanut quality evaluation and inspection. Even some of them have been used as excellent standards in oil crops. However, these methods are laborious, time consuming, highly empirical, generally work intensive, and require the use of large amounts of chemical solvents and analytical reagents that might be hazardous and harmful to analysts and the lab environment.

Looking for a rapid and non-destructive method for estimating and determining the oil content in different peanut varieties is very necessary. Near-infrared (NIR) spectroscopy, middle infrared (MIR) spectroscopy, and Raman spectroscopy as the rapid, non-destructive, and chemical-free techniques have been widely developed for measurement and evaluation of chemical information in the food industry. All of these three kinds of spectroscopic techniques have their unique advantages for the determination of the chemical components such as protein, oil, carbohydrates, amino acids, fiber components, and many other parameters in foods and food products (Kandala and Sundaram 2014; Sundaram et al. 2010; Tillman et al. 2006; Phan-Thien et al. 2011; Wang et al. 2013). However, they also have their own disadvantages. Concerning the NIR spectroscopy, for quantitative analyses, NIR spectroscopy is not independent of the disadvantages arising from the reference method applied for calibration, which requires a set of samples at least 20–50 with known concentrations (Cheng et al. 2013). As to Raman spectroscopy, the inherently weak effect of Raman scattering, the strong interfering of biological fluorescence background signals, the traditionally high instrumental costs, and the heat generated by the laser may affect or alter the measurement effectiveness (Cheng et al. 2013). More importantly, these spectroscopic techniques cannot provide the spatial information of the tested object.

Hyperspectral imaging (HSI) is an emerging and innovative technique, which has advantages, that it is rapid, noninvasive, reliable for quality inspection, but compared to NIR spectroscopy, it also integrates traditional spectroscopy and digital imaging into one system, making it possible for providing both spectral and spatial information of a certain object simultaneously (Zheng et al. 2006). The outstanding advantages of HSI make it possible to accomplish the evaluation of external features such as size, appearance, color, defects, and so on, and estimation of internal properties that moisture, protein, fat, carbohydrates, etc., concurrently.

Up to now, HSI has been successfully used for classification of maize kernel hardness (Williams et al. 2009),

discrimination of maize, oat and groat kernels (Zhang et al. 2012; Serranti et al. 2013), fusarium infection in wheat (Bauriegel et al. 2011; Shahin and Symons 2012), and prediction of the composition of maize kernels (Weinstock et al. 2006; Hurburgh et al. 2004). However, using hyperspectral imaging technique for determination of oil content in peanut has not been reported to date.

Therefore, this study was aimed to investigate the feasibility of using HSI for predicting oil content in peanut. The specific objectives of this article were to (1) acquire hyperspectral images of peanut in the spectral range I (400–1000 nm) and spectral range II (1000–2500 nm); (2) extract the corresponding spectral information from identified regions of interests (ROIs) within the acquired hyperspectral images; (3) preprocess the original spectral information using several different spectral pretreatment methods and then establish the calibration model between the processed spectral information and the reference measured oil content; (4) verify the prediction ability of the calibration model; (5) choose the optimal wavelengths that carried the most relevant and useful information related to oil content prediction; and (6) build the new calibration model based on the selected optimal wavelengths and verify it.

Material and Methods

Sample Preparation and Oil Content Measurement

Five varieties (Huayu, Luohanguo, Zhonghua, Dabaisha, Xiaobaisha, which are expressed as HY, LHG, ZH, DBS, and XBS peanuts were purchased from a local seed market in Zhengzhou, China. After the shells were removed, 30 kernels with uniform size of each variety were selected for testing. Each peanut kernel was scanned by HSI, and the reference oil content of all samples were determined by using nuclear magnetic resonance (NMR), according to the protocol of the Official Methods and Recommended Practices of the American Oil Chemists' Society (1993). All experiments were conducted in triplicate. All data were expressed as mean ± standard deviation. The least significant difference procedure was carried out to test for difference between means (significance was defined at $P < 0.05$). Table 1 shows the relevant statistics of oil content for the samples. The whole dataset (150 samples) was divided into two groups, one group for building the calibration model consisting of 100 samples (calibration set) and another group used for validation consisting of 50 samples (prediction set).

Hyperspectral Imaging System

Two HSIs (HSI-1, the spectral range of 400–1000 nm; HSI-2, the spectral range of 1000–2500 nm) were assembled to acquire hyperspectral images of peanuts in reflectance mode. HSI-1 system consists of an imaging spectrograph (Inspector V10E,

Table 1 Reference oil content (%) of peanut kernels measured by traditional method

Source	Calibration set				Prediction set			
	Samples	Min	Max	Mean ± SD	Samples	Min	Max	Mean ± SD
HY	20	48.31	49.76	48.86 ± 0.37^b	10	48.36	49.42	48.77 ± 0.30^b
LHG	20	47.83	48.29	48.07 ± 0.15^a	10	47.83	48.25	48.05 ± 0.13^a
ZH	20	47.72	48.36	48.05 ± 0.23^a	10	47.75	48.26	48.03 ± 0.23^a
DBS	20	48.46	49.69	48.98 ± 0.30^b	10	48.53	49.18	48.96 ± 0.22^a
XBS	20	49.38	50.61	49.99 ± 0.34^b	10	49.62	50.34	49.99 ± 0.24^a
Total	100	47.72	50.61	48.79 ± 0.77^c	50	47.75	50.34	48.76 ± 0.76^c

[a], [b], [c] indicate the statistical significance ($p < 0.05$)

Spectral Imaging Ltd., Oulu, Finland) covering the spectral range of 308–1105 nm with 1.58 nm increments, a high-performance camera (DL-604 M, Andor Technology PLC, Belfast, Northern Ireland) with the effective resolution of 1004×1002 pixels, a camera lens (OLE23, Schneider Electric SA, RueilMalmaison, France), an illumination unit composed of two 150 W halogen lamps (2900-ER, Illumination Technologies Inc., New York, USA), a stepping motor (IRCP0076-1COMB, Isuzu Optics Corp., Taiwan, China), and a mobile platform and a computer system executing imaging data acquisition software (Spectral Image, Isuzu Optics Corp., Taiwan, China), which was used to control the exposure time, motor speed, combining mode, wavelength range, and image acquisition. HSI-2 system comprises a line-scan spectrograph (Specim V25E, Spectral Imaging Ltd., Oulu, Finland) spanning the spectral range of 1000–2500 nm with 6.3 nm increments, a high-performance 320×256 CCD camera (XC403, Xenics Infrared Solutions, Leuven, Belgium), a camera lens (OLES30, Xenics Infrared Solutions, Leuven, Belgium), an illumination unit including two halogen lamps (3900-ER, Illumination Technologies Inc., NY, USA), a conveyor belt controlled by a stepper motor (IRCP0076-1COMB, Isuzu Optics Corp., Taiwan, China), data acquisition software (Spectral Image software, Isuzu Optics Corp., Taiwan, China), and a computer.

Image Acquisition and Correction

Each peanut kernel was placed on the mobile platform and scanned line by line with the adjusted speed and exposure time to obtain the hyperspectral images, which has three dimensions (x, y, λ), where x and y are the spatial dimensions and λ is the number of wavebands. Thus, the original images were created and stored. In order to eliminate the effects of illumination and geometry and detector sensitivity, the raw acquired hyperspectral images (I_0) should be corrected using two reference standards: a white one (W) to set-up the maximum reflectance (~99 %) condition, which was obtained for a white calibration tile under the same condition of the raw image, and a black one (B) to define the no reflectance (~0 %) condition, which was acquired by turning off the light source and

completely covering the lens with its black cap. The calibrated image (I) was then calculated by the following equation:

$$I = \frac{I_0 - B}{W - B} \qquad (1)$$

All the corrected images were used to extract the spectral data in the subsequent image analysis.

Spectral Extraction

In order to extract spectral data from each peanut kernel in the corrected image, the Environment for Visualizing Images
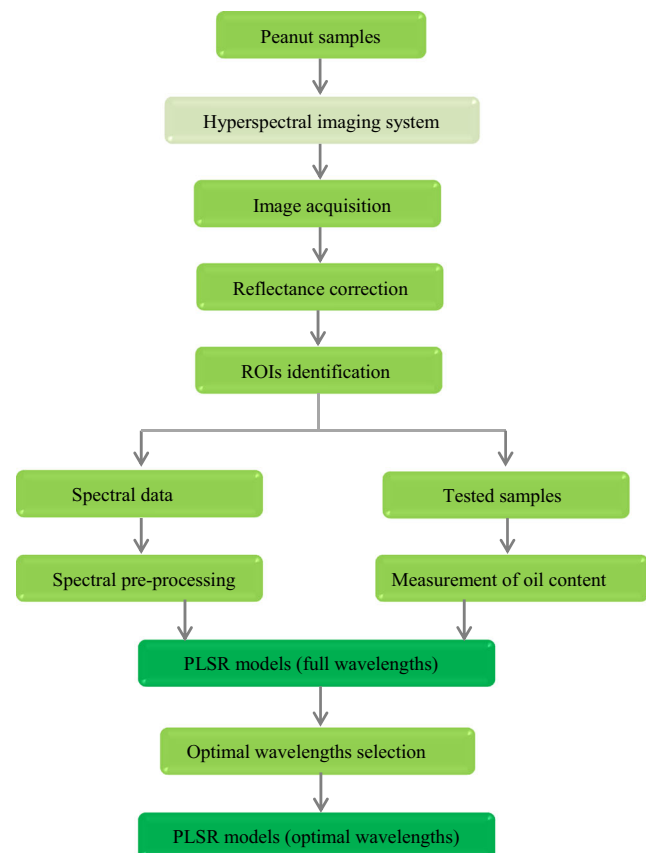

Fig. 1 The key steps of the experiment by using hyperspectral imaging technique
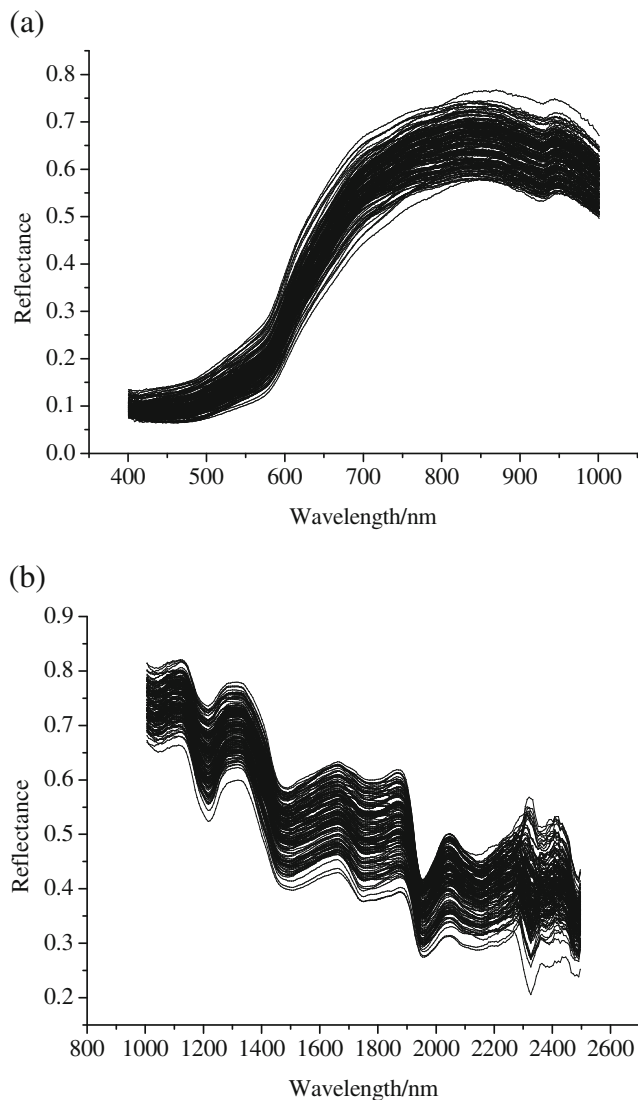
(a)

(b)

Fig. 2 Spectral data of the examined peanut samples in two different wavelength ranges. **a** 400–1000 nm. **b** 1000–2500 nm

software (ENVI v4.8, ITT Visual Information Solutions, Boulder, CO, USA) was used to identify the ROIs with a circle shape (324 pixels), and then the spectral data within ROIs for samples were extracted and averaged at each wavelength to obtain one mean spectrum representing the ROI. All of the

extracted spectral data from peanut kernels were then arranged in a matrix where the rows of this matrix represent the number of samples and the columns represent the number of variables (381 and 236 wavelengths).

Chemometrics Analysis

Before developing a calibration model, spectral pre-processing was executed in order to correct spectral data. By pre-processing procedure, some random noises resulting from instrumental effects or variable physical sample properties can be removed (Rinnan et al. 2009). In this study, five pretreatment methods of smoothing, normalize, multiplicative scatter correction (MSC), baseline, and standard normal variate (SNV) were used to pre-process the raw spectral data. The calibration models were established using a multivariate data analysis method. Among them, SNV, MSC and normalization are usually used to eliminate the effects of the spectral variability resulting from the light scattering (Tavallaie et al. 2011). Spectra smoothing and baseline are applied to smooth spectra by elimination of baseline shifts and resolution of o-verlapped peaks (Amigo et al. 2015). Partial least squares regression (PLSR) is a reliable and effective multivariate chemometrics method, it has prominent advantages that solve multicollinearity problems and allowing variables more than the samples, has been widely used for developing a mathematical model (Abdi 2010; Wu et al. 2012; He et al. 2013). In this study, quantitative calibration models were established between pre-processed spectral data of calibration set and the reference oil content by PLSR analysis, and the performance of these models were evaluated and compared. In this part, all the spectral pre-process and multivariate data analysis were carried out by means of The Unscrambler 9.7 (CAMO Software AS, Norway).

Optimal Wavelengths Selection

The hyperspectral images suffered from the problem of high dimensionality and multi-collinearity among contiguous wavelength bands lead to slowing down the computation

Table 2 Results of PLSR models based on different spectral pretreatment in spectral range I

PLSR model	No. of varieties	PCs	Calibration		Cross-Validation		Prediction	
			R^2	RMSEC (%)	R^2_{CV}	RMSECV (%)	R^2_P	RMSECP (%)
Raw-PLSR	381	8	0.729	0.400	0.605	0.485	0.696	0.416
Smoothing-PLSR	381	8	0.701	0.419	0.605	0.486	0.646	0.448
Normalize-PLSR	381	7	0.710	0.412	0.600	0.489	0.649	0.446
MSC-PLSR	381	3	0.616	0.475	0.592	0.494	0.538	0.512
Baseline-PLSR	381	7	0.706	0.415	0.599	0.490	0.641	0.451
SNV-PLSR	381	3	0.626	0.469	0.602	0.488	0.536	0.513

PLSR model	No. of varieties	PCs	Calibration		Cross-Validation		Prediction	
			R^2	RMSEC (%)	R^2_{CV}	RMSECV (%)	R^2_P	RMSECP (%)
Raw-PLSR	236	4	0.904	0.237	0.893	0.254	0.923	0.208
Smoothing-PLSR	236	4	0.904	0.237	0.893	0.254	0.923	0.208
Normalize-PLSR	236	3	0.800	0.343	0.783	0.361	0.851	0.289
MSC-PLSR	236	7	0.752	0.382	0.609	0.484	0.697	0.414
Baseline-PLSR	236	9	0.769	0.368	0.625	0.474	0.636	0. 453
SNV-PLSR	236	3	0.626	0.469	0.602	0.488	0.536	0.513

Table 3 Results of PLSR models based on different spectral pretreatment in spectral range II

speed of the calibration process and influencing the modeling accuracy. Therefore, several optimal wavelengths that carry the most sensitive and useful information corresponding to the oil content of peanut should be found. However, there is no common method available to select the optimal wavelengths, although a number of approaches have been proposed, including successive projection algorithm (SPA) (Wu et al. 2010), principal component analysis (PCA) (Lorente et al. 2012), regression coefficients based on the PLSR model (ElMasry et al. 2013; Iqbal et al. 2013; He et al. 2013), and genetic algorithm (Lin et al. 2012). In this study, regression coefficients based on the PLSR analysis was used to select the informative and feature wavelengths from the full spectral range. The weighted regression coefficients analysis, also called β-coefficients method, use the regression coefficients resulting from the calibration model (PLSR) to choose the effective wavelengths. The wavelengths with the highest absolute values of β-coefficients are selected as the feature wavelengths, whereas the ones with the lowest absolute values of β-coefficients are completely removed due to their little contribution in prediction (ElMasry et al. 2013). Based on the selected optimal wavelengths, the original PLSR models using the full wavelengths both in the spectral range of 400–1000 nm and 1000–2500 nm were simplified for further analysis.

Modeling Establishment and Evaluation

The PLSR models were built with the calibration set under full cross validation by using leave-one-out cross-validation method. In this method, one test sample was excluded from the original sample set, and used as validation data, then the PLSR model was calculated using the remaining calibration samples. The established model was then used to predict the result of the excluded sample. This procedure was repeated until each sample was used as the validation data. Besides validating the PLSR models by leave-one-out cross-validation, the predictive capacity of the established PLSR models was rather verified in the prediction set by comparing the actual measured oil content values with predictive oil content values of the prediction set. Parameters that evaluate the robustness and

predictive accuracy of a PLSR model consist of determination coefficients of calibration (R^2_C), cross-validation (R^2_{CV}), and prediction (R^2_P) and root mean square error of calibration (RMSEC), cross-validation (RMSECV), and prediction (RMSEP). Generally, a good model should have high R^2_C, R^2_{CV}, and R^2_P and low RMSEC, RMSECV, and RMSEP, as well as a slight difference between them. The key steps of the experimental procedure are showed in Fig. 1

Results and Discussion

Spectral Features Analysis

The full wavelength scanned from HSI-1 and HSI-2 ranges from 328 to 1115 nm and 916 to 2533 nm, respectively; however, in order to avoid high signal-to-noise ratio, only a spectral range of 400–1000 nm (spectral range I) and a spectral range of 1000–2500 nm (spectral range II) were used for further analysis. As we have known, the most dominant absorption bands in the visible and near-infrared region are ascribable to the strong overtone and combination absorptions of hydrogen containing bonds O-H (found in water), C-H (usually from fats and oil), and N-H (found in

Fig. 3 Selection of optimal wavelengths corresponding to the large values of regression coefficients of the raw-PLSR model in the spectral range II

Table 4 PLSR models established using different important wavelengths combinations in spectral range II

Important wavelengths (nm)	No.	PCs	Calibration		Cross-Validation	
			R^2_C	RMSEC (%)	R^2_{CV}	RMSECV (%)
1127, 1216, 1477, 1738, 1953, 2073, 2143, 2319	8	3	0.898	0.245	0.889	0.258
1127, 1216, 1477, 1738, 1953, 2073, 2319	7	4	0.911	0.228	0.903	0.240
1216, 1477, 1738, 1953, 2073, 2319	6	3	0.909	0.232	0.902	0.242
1127, 1216, 1738, 1953, 2073, 2319	6	4	0.912	0.227	0.904	0.239

protein). The spectral reflectance curves of rested peanut samples exhibited a similar trend in spectral range I and II (Fig. 2), but there were nuances between them. The distinctions were possibly ascribed to the quality attributes, uneven surface structure, and the unfixed scatter on the surface of samples (Wu et al. 2012). To be specific, the absorption peak at 925 nm in Fig. 2a was probably related to the third overtone of C-H stretches (Wold et al. 2001). In Fig. 2b, however, the main absorption peaks were observed at around 1216, 1738, and 2330 nm; these absorption peaks were presumably attributed to the contribute of first and second overtone of C-H stretches of methyl or methylene group in fat components (Holman and Edmondson 1956; Kays et al. 1998; Petisco et al. 2010).

Calibration of Oil Content Using Full Spectral Range

Based on all the spectral data of samples and their corresponding measured oil content values, the PLSR calibration models using the whole spectral wavelengths were established in the spectral range I and II. And before this, the spectral data were preprocessed using different spectral treatment methods. The performances of PLSR models are presented in Tables 2 and 3. In spectral range I, the statistical parameters of PLSR models in Table 2 suggested that PLSR models based on different spectral pretreatment methods exhibited poor model performances. It can be concluded that the used spectral pretreatment methods were not helpful to improve the PLSR calibration capability. It was also indicated that using the raw spectral data for modeling was considered as to be good, which means that the applications of these pretreatment methods possibly removed some important information about oil content in peanut. As Table 3 shown, in spectral range II, raw-PLSR, and Smoothing-PLSR showed the same high performance, which were obviously superior to other models. Therefore, the better choice was to establish the PLSR

prediction model for oil content using raw spectra or spectra based on smoothing process. In this study, raw spectra were used for further analysis. Raw-PLSR model presented almost similar good performance in both calibration and cross-validation with R^2_C, R^2_{CV} values of 0.904, 0.893 and RMSEC, RMSECV of 0.237 %, 0.254 %, respectively. R^2_P value and RMSEP were 0.923 and 0.208 %, which implied that raw-PLSR with whole wavelengths for predicting oil content of peanut showed excellent predictive accuracy. In brief, the established PLSR models with low prediction errors in this work confirmed the suitability of hyperspectral imaging technique for oil content prediction in peanut, and the better results would be achieved using spectral range II for analysis.

Modeling of Oil Content by PLSR Using Selected Optimal Wavelengths

In order to avoid the large dimensionality of the acquired hyperspectral images, minimize inessential information among contiguous bands, and reduce the time of computation, it is of necessity to select optimal wavelengths with effective and useful information for predicting oil content of peanut. The weighted regression coefficients resulting from the best PLSR models were considered as an indication of the possible optimal wavelengths that could explain most of the variation. In this study, the optimal wavelengths were selected based on the regression coefficients of the established raw-PLSR model in spectral range II; the wavelengths with large regression coefficients values (regardless of the sign) could be chosen. By means of this approach, eight individual wavelengths (1127, 1216, 1477, 1738, 1953, 2073, 2143, and 2319 nm) for oil content prediction in the spectral range II were selected preliminarily (shown in Fig. 3), then the selected wavelengths were compared by using different combinations among them to establish a PLSR model, as shown in Table 4. Finally, six wavelengths (1127, 1216, 1738, 1953, 2073, and 2319 nm) in

Table 5 The performance of the optimized model during calibration, cross-validation, and prediction

Model	No. of varieties	PCs	Calibration		Cross-Validation		Prediction	
			R^2	RMSEC (%)	R^2_{CV}	RMSECV (%)	R^2_P	RMSECP (%)
Optimized -PLSR	6	4	0.912	0.227	0.904	0.239	0.934	0.197

OC = 47.75% OC = 48.03% OC = 48.98% OC = 49.51% OC = 50.23%
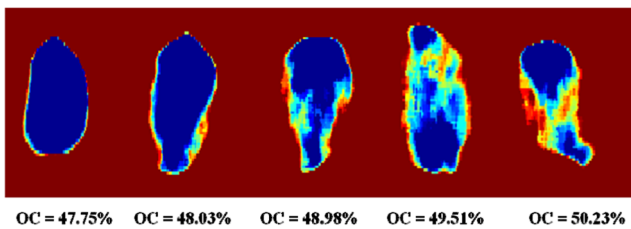
Fig. 4 Some examples of visualized distribution maps of oil content in peanut kernels

spectral range II were identified as the optimal wavelengths. Accordingly, the optimized model was redesigned only using the selected six optimal wavelengths, and the predictive capacity of the model was verified in the prediction set by comparing the actual measured oil content values with predictive oil content values of prediction set, whose robustness and reliability were also assessed by the determination coefficients and root mean square errors. As illustrated in Table 5, the optimized model exhibited the R^2_C, R^2_{CV}, and R^2_P values of 0.912, 0.904, and 0.934 with RMSEC, RMSECV, and RMSEP of 0.227 %, 0.239 %, and 0.197 %, respectively. Compared with the original PLSR model, the optimized PLSR model displayed better performance, the robustness and reliability of the model was proved to be good simultaneously. It can be concluded that using six optimal wavelengths to replace the full wavelengths in spectral range II for predicting the oil content of peanut is possible and better.

In addition, Fig. 4 illustrates some examples of distribution maps of different oil contents in peanuts. It is very obvious to observe the changes of oil content (from small to big) in peanut kernels. The obtained distribution maps were generated by image processing method using pseudo color. In the distribution maps, the pixels holding similar spectral characteristics usually show the same predicted values, which are then visualized in a similar color in the image. Different colors in the final distribution maps represent different chemical indicator values in the image in proportion to the spectral differences of the corresponding pixels. Therefore, according to the distribution maps of oil content, it is useful to rapidly and non-destructively estimate and evaluate the oil content in different peanut kernels.

Conclusions

The feasibility of hyperspectral imaging technique in the spectral range of 400–1000 nm and 1000–2500 nm for predicting oil content in peanut was investigated. Through spectral pre-treatment, quantitative calibration models using full spectral wavelengths were established. By comparing the performance of models undergoing different spectral pre-processing, it was concluded that raw-PLSR models presented better results. Raw-PLSR model in the spectral range II exhibited good

performance with the high values (R^2_P = 0.923). The HSI-2 was more reliable in prediction of oil content of peanut. Besides, six optimal wavelengths were selected using the regression coefficients of the established raw-PLSR model in the spectral range II and the optimized calibration model was established. The optimized calibration model showed satisfactory capability with R^2_C, R^2_P of 0.912 and 0.934, which presented the equivalent performance with the original raw-PLSR model. On the whole, the above results indicated that the established optimized PLSR model had accurate predictive capability and good model robustness, and the hyperspectral imaging technique in tandem with chemometrics analysis has the potential to predict oil content of peanut in a rapid and non-destructive manner. To the best of our knowledge, it is the first time to use the hyperspectral imaging technique for predicting oil content in peanut, the varieties of peanut adopted are limited, and extending this technique for more diversified peanut samples would be more meaningful.

Compliance with Ethical Standards

Conflict of Interest Huali Jin declares that she has no conflict of interest. Yongsheng Ma declares that he has no conflict of interest. Linlin Li declares that she has no conflict of interest. Jun-Hu Cheng declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Not applicable.

References

Abdi H (2010) Partial least squares regression and projection on latent structure regression (PLS regression). Wiley interdisciplinary reviews. Comput Stat 2:97–106

Amigo JM, Babamoradi H, Elcoroaristizabal S (2015) Hyperspectral image analysis. A tutorial. Anal Chim Acta 896(10):34–51

Bauriegel E, Giebel A, Geyer M, Schmidt U, Herppich WB (2011) Early detection of Fusarium infection in wheat using hyper-spectral imaging. Comput Electron Agric 75:304–312

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Physiol 37:911–917

Brown DF, Cater CM, Mattil KF et al (1975) Effect of variety, growing location and their interaction on the fatty acid composition of peanuts. J Food Sci 40(5):1055–1060

Cheng JH, Dai Q, Sun DW, Zeng XA, Liu D, Pu HB (2013) Applications of non-destructive spectroscopic techniques for fish quality and safety evaluation and inspection. Trends Food Sci Technol 34(1):18–31

ElMasry G, Sun DW, Allen P (2013) Chemical-free assessment and mapping of major constituents in beef using hyperspectral imaging. J Food Eng 117:235–246

Firestone D (1993) Official methods and recommended practices of the American oil chemists' society, vol 4th. AOCS Press, Champaign

Folch J, Lees M, Sloane Stanley GH (1957) A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem 226:497–509

He HJ, Wu D, Sun DW (2013) Non-destructive and rapid analysis of moisture distribution in farmed Atlantic salmon (Salmo salar) fillets using visible and near-infrared hyperspectral imaging. Innovative Food Sci Emerg Technol 18:237–245

Holley KT, Hammons RO (1968) Strain and seasonal effects on peanut characteristics. Res. Bull. 32 Ga agric. Exp. Stn 27

Holman RT, Edmondson PR (1956) Near-infrared spectra of fatty acids and some related substances. Anal Chem 28(10):1533–1538

Hurburgh CR, Rippke GR, Cogdill RP (2004) Single-kernel maize analysis by near-infrared hyperspectral imaging. Trans Asae 47(1):311–320

Iqbal A, Sun DW, Allen P (2013) Prediction of moisture, color and pH in cooked, pre-sliced turkey hams by NIR hyperspectral imaging system. J Food Eng 117:42–51

Kandala CV, Sundaram J (2014) Nondestructive moisture content determination of three different market type in-shell peanuts using near infrared reflectance spectroscopy. Food Meas 8:132–141

Kays SE, Windham WR, Barton FE (1998) Prediction of total dietary fiber by near-infrared reflectance spectroscopy in high-fat- and high-sugar-containing cereal products. J Agric Food Chem 46(3):854–861

King J, O'Farrel WV (1997) SFE-new method to measure oil content, inform. 8:1047–1051

Lin P, Chen Y, He Y (2012) Identification of geographical origin of olive oil using visible and near-infrared spectroscopy technique combined with chemometrics. Food Bioprocess Technol 5(1):235–242

Lorente D, Aleixos N, Gómez-Sanchis J, Cubero S, García-Navarrete OL, Blasco J (2012) Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. Food Bioprocess Technol 5(4):1121–1142

Petisco C, Garcia-criado B, Vazquez-de-aldana BR et al (2010) Measurement of quality parameters in intact seeds of Brassica species using visible and near-infrared spectroscopy. Ind Crop Prod 32(2):139–146

Phan-Thien KY, Golic M, Wright GC, Lee NA (2011) Feasibility of estimating peanut essential minerals by near infrared reflectance spectroscopy. Sens & Instrumen Food Qual 5:43–49

Rinnan Å, van den Berg F, Engelsen SB (2009) Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal Chem 28(10):1201–1222

Rustom IYS, Lopez-Leiva MH, Nair BM (1996) Nutritional, sensory and physicochemical properties of peanut beverage sterilized under two different UHT conditions. Food Chem 56(1):45–53

Sanders TH (1980) Fatty acid composition of lipid classes in oils from peanuts differing in variety and maturity. J Am Oil Chem Soc 57(1):12–15

Serranti S, Cesare D, Marini F, Bonifazi G (2013) Classification of oat and groat kernels using NIR hyperspectral imaging. Talanta 103:276–284

Shahin MA, Symons SJ (2012) Detection of fusarium damage in Canadian wheat using visible/near-infrared hyperspectral imaging. Food Meas 6:3–11

Shewfelt AL, Young CT (1977) Storage stability of peanut-based foods: a review. J Food Sci 42(5):1148–1152

Sundaram J, Kandala CV, Ronald AH, Christopher LB, William RW (2010) Determination of in-shell peanut oil and fatty acid composition using near-infrared reflectance spectroscopy. J Am Oil Chem Soc 87:1103–1114

Tavallaie R, Talebpour Z, Azad J, Soudi MR (2011) Simultaneous determination of pyruvate and acetate levels in xanthan biopolymer by infrared spectroscopy: effect of spectral pre-processing for solid-state analysis. Food Chem 124(3):1124–1130

Tillman LB, Gorbet WD, Person G (2006) Predicting oleic and linoleic acid content of single peanut seeds using near-infrared reflectance spectroscopy. Crop Sci 46:2121–2126

Wang L, Wang Q, Liu HZ, Liu L, Du Y (2013) Determining the contents of protein and amino acids in peanuts using near-infrared reflectance spectroscopy. Soc Chem Ind 93:118–124

Weinstock BA, Janni J, Hagen L, Wright S (2006) Prediction of oil and oleic acid concentrations in individual corn (Zea mays l.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. Appl Spectrosc 60(1):9–16

Williams P, Geladi P, Fox G, Manley M (2009) Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. Anal Chim Acta 653:121–130

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 8(2):109–130

Wu D, He Y, Nie P, Cao F, Bao Y (2010) Hybrid variable selection in visible and near-infrared spectral analysis for non-invasive quality determination of grape juice. Anal Chim Acta 659(1):229–237

Wu D, Sun DW, He Y (2012) Application of long-wave near infrared hyperspectral imaging for measurement of colour distribution in salmon fillet. Innovative Food Sci Emerg Technol 16:361–372

Zhang XL, Liu F, He Y, Li XL (2012) Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. Sensors 12:17234–17246

Zheng CX, Sun DW, Zheng LY (2006) Recent developments and applications of image features for food quality evaluation and inspection—a review. Trends Food Sci Technol 17:642–655