Rapid quantitative detection of mineral oil contamination in vegetable oil by near-infrared spectroscopy

Mantong Zhao (赵曼彤)¹, Dawei Zhang (张大伟)^{1,2,*}, Lulu Zheng (郑璐璐)¹, Otto Condliffe³, and Yi Kang (康 祎)¹

¹Engineering Research Centre of Optical Instrument and System, Ministry of Education, Shanghai Key Laboratory of

Modern Optical System, University of Shanghai for Science and Technology, Shanghai 200093, China

²Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092, China

³Sino-England International College, University of Shanghai for Science and Technology, Shanghai 200093, China *Corresponding author: dwzhang@usst.edu.cn

Received November 13, 2019; accepted December 11, 2019; posted online March 18, 2020

This study provides a rapid method for quantification of mineral oil in rapeseed oil using near-infrared spectroscopy. The data were processed by direct orthogonal signal correction (DOSC), successive projections algorithm (SPA), partial least squares, and principal component regression (PCR). Good correlation coefficients (R) of 0.998 and root-mean-squared error (RMSE) of 0.005 were obtained, and the DOSC-SPA-PCR model was identified as the optimal method. A satisfactory accuracy with R and RMSE of prediction by DOSC-SPA-PCR of 0.990 and 0.006, was obtained. The results demonstrate that the proposed methodology is a promising method for the rapid quantitative detection of mineral oil in vegetable oil.

Keywords: near-infrared spectroscopy; mineral oil; oil quality control; DOSC-SPA-PCR. doi: 10.3788/COL202018.043001.

Foods are often contaminated by mineral oils, which have negative effects on human health. Mineral oils are complex mixtures of hydrocarbons derived from petroleum. They are absorbed by the body, stored in various organs, and may cause microgranulomas in several tissues, which are extremely harmful to the human body. In 2008, a research on mineral oil in food reviewing the toxicity and occurrence of this kind of contamination was developed by the European Food Safety Authority (EFSA)^[1,2]. Although data on mineral oil contamination are available only for a limited number of food categories, a significant contribution to the human daily intake can be ascribed to vegetable oils, which have been reported to be frequently contaminated with amounts of mineral oils around 0.02–0.08 g \cdot kg^{-1]3}. In 2008, nearly 100,000 tons of imported Ukrainian sunflower seed oil was contaminated with more than 1000 mg \cdot kg⁻¹ of mineral oil in the European Union. Even though the source of the contamination was never officially confirmed, it is probable that mineral oil was added as a fraud^[4]. The occurrence of</sup> mineral oils in foodstuffs is related to different sources: environmental contamination [5,6], food-grade mineral oils widely used for different purposes in many processed foods^[7], and food contact materials^[8]. Along with the above mentioned contamination sources, some unscrupulous traders illegally adulterate vegetable oils in order to reduce production costs. Thus, in order to protect consumers from the hazards of mineral oil, the control of mineral oil in vegetable oils is of major importance, and food testing laboratories must ensure testing speed in order to provide a timely warning to the producers and relevant authorities in the case of a crisis. For these reasons, laboratories should be provided with rapid, sensitive, and

reliable analytical and cost effective methods for the determination of mineral oil.

A large number of analytical methods for the determination of mineral oils in edible oils have been proposed over the years, such as liquid–liquid–gas chromatography $(GC)^{[9]}$, GC-flame ionization detection $(FID)^{[10]}$, and liquid chromatography (LC-GC-FID)^[11]. These methods not only require advanced laboratory facilities but also involve elaborate processing steps and consume a great deal of time $\underline{12}$, for instance, the removal of olefins is necessary before subjection of the sample to LC-GC-FID. Mineral oil was studied using Raman spectroscopy by de Jong *et al.*^[13], but the detection limit of mineral oil is about 0.25% (=2500 mg/kg), which is much higher than the maximum limit of 50 mg/kg that is defined by the European Union for mineral oil in edible oil. Thus, some rapid, convenient, and accurate quantitative detection methods are needed. With the combination of modern chemometrics and instrumentation, near-infrared (NIR) spectroscopy is widely applied for rapid, accurate, lowcost, and non-destructive analysis in multiple industries. NIR spectroscopy has been revealed as promising for classification and quantitative determination of various materials, including oil quantity^[14], olive oil^[15,16]</sup>, energy</sup> drinks^[17], soybean oil^[18], and biodiesel/diesel blends^[19,20].

The objective of this study was to develop a rapid and accurate method based on NIR spectroscopy integrated with chemometrics for rapid quantitative detection of mineral oil contamination in vegetable oil. The NIR spectra of mineral oil/vegetable oil blends were acquired in the spectral range of 910–2150 nm. The original spectrum is preprocessed by direct orthogonal signal correction (DOSC). The successive projections algorithm (SPA) was employed for selecting effective wavelengths in the calibration set after the DOSC preprocessing. Based on the effective wavelengths, prediction models were set up using two different modeling approaches—partial least squares (PLS) and principal component regression (PCR). After the models were established, the prediction set was then analyzed in order to estimate the actual predictive capability of the established models.

The vegetable oil sample in this study was rapeseed oil. Two brands (B1 and B2) of rapeseed oil were used during the period from February to May 2019 within the framework of a monitoring program for the control of mineral oil in food products. The mineral paraffin oil "Paraffin viscous PH Eur, BP, USP" [CAS 8012-95-1] was from Merck (Darmstadt, Germany). B1 brand rapeseed oil samples were adulterated with mineral oil at eight different concentrations: 1, 0.5, 0.2, 0.1, 0.08, 0.05, 0.03, and $0.01 \text{ g} \cdot \text{kg}^{-1}$, respectively. The samples were divided into a calibration set and a prediction set. The calibration set was used to establish models for quantitative detection of mineral oil contamination in vegetable oil, and the prediction set was used as an external test set for validating the actual prediction ability of the models. We took 25 samples for each concentration in the B1 brand. Fifteen samples were for calibration and ten for prediction. The other brand (B2) of rapeseed oil sample was adulterated with $0.8 \text{ g} \cdot \text{kg}^{-1}$ and $0.1 \text{ g} \cdot \text{kg}^{-1}$ mineral oil for prediction, and we obtained 10 samples of $0.8 \text{ g} \cdot \text{kg}^{-1}$ and $0.1 \text{ g} \cdot \text{kg}^{-1}$, respectively. That is to say, we also validated the model with a new concentration (concentration not present in the calibration set) from another brand of rapeseed oil. The sample information in the calibration and prediction sets is shown in Table 1.

NIR spectrometer (Zeiss MCS 611NIR 2.2 WR) was used for spectral acquisition between 910 nm and 2150 nm. The NIR spectrum of each sample was obtained by taking the average of 33 scans. The spectra of rapeseed oil/mineral oil blends were obtained in the range from 910 to 2150 nm with 4 nm spectral resolution, and the average spectrum was used in modeling the data. The time required to achieve a spectral measurement was 30 s.

NIR spectra are often influenced by instrumental variation and measurement conditions, such as background noise and baseline shift. Therefore, the spectra are often pretreated before calibration. A relatively new preprocessing technique, DOSC, was introduced by Westerhuis *et al.*^[21].

 Table 1. Sample Information in Calibration and Prediction Sets

Set	Brand	Concentration $(\mathbf{g} \cdot \mathbf{kg}^{-1})$							
Calibration	B1	1 0	.5	0.2	0.1	0.08	0.05	0.03	0.01
Prediction	B1	1 0	.5	0.2	0.1	0.08	0.05	0.03	0.01
	B2		0).8			0	.1	



Fig. 1. Transmittance spectra of rapeseed oil (B1 brand) and mineral oil.

This technique has better ability to subtract the irrelevant information compared to the previous methods. The DOSC was performed using MATLAB (R2018a, the MathWorks Inc., USA). The spectrum after being pretreated (X^{DOSC}) was used to do the following analysis.

Figure 1 presents the transmittance^[22] curves in the range of 910-2150 nm, collected from the mineral oil and rapeseed oil (B1 brand). The two average curves were obtained using MATLAB 2018a and Origin 8.0. Mineral oils are mixtures of liquid hydrocarbons derived from petroleum. In detail, the C-H first overtone in hydrocarbon appears at about 1800 nm. It can be seen from Fig. 1 that there is an absorption peak in the transmission curve of mineral oil at 1814 nm. Rapeseed oils are made up of esters. The C = O keys second overtone in esters appears at about 1910–1950 nm. Figure 1 shows that rapeseed oil has an absorption peak at 1930 nm. Moreover, it can be seen from the figure that the transmittance curves of mineral oil and rapeseed oil are quite different, preliminarily indicating that pure rapeseed and mineral oils can be discriminated.

Principal component analysis (PCA) is a multivariate projection method, widely used in chemometrics to compress high-dimensional data into a lower-dimensional space with a minimal loss of information^[23]. It is designed to extract and display essential information from a data set and forms new variables called principal components (PCs). In the study, PCA was conducted on the fullspectrum data. The PCA was performed using the Statistical Product and Service Solutions (SPSS). All of the spectra of the calibration set sample were evaluated by PCA. Figure <u>2</u> presents the PC2 × PC1 score plot resulting from the application of PCA. It is noteworthy that the first PC represents 90.47% of the data variance, while the second PC represents 7.51%. As can be seen, there is no overlap between different concentrations of mineral oils,



Fig. 2. PC2 \times PC1 score plot for the overall set of 120 rapes eed–mineral oil samples.

indicating that the NIR spectrum conveys appropriate information for the classification task. This observation shows that the NIR spectra together with PCA can realize the classification of different concentrations of mineral oil in rapeseed oil.

Spectral data are known to possess variable (wavelength) dimensions with redundancy among contiguous variables^[24]. Elimination of irrelevant variables can predigest calibration modeling and improve results in terms of accuracy and robustness^[25]. The SPA was originally proposed by Araújo et al. in the context of multivariate calibration $\frac{[26,27]}{2}$. In this study, SPA was applied to select the effective wavelengths that have the greatest contribution to the quantitative detection of mineral oil. SPA was carried out in a MATLAB environment (R2018a, the MathWorks Inc., USA). According to previous studies, effective wavelengths might be equally or more efficient than full wavelengths because they carry the most important information relevant to determination. In this study, 36 variables were selected as the effective wavelengths from the full spectra (910–2150 nm) for the quantitative detection of mineral oil contamination in vegetable oil, as shown in Table 2.

Once the effective wavelengths were selected, the spectral data sets were then reduced to a matrix with a dimension of $s \times t$, where s is the number of samples (s = 120), and the number of variables t is reduced from 311 to 36 wavelengths (the number of effective wavelengths). In this study, the models were established using two calibration algorithms, $PLS^{\underline{[28,29]}}$ and $PCR^{\underline{[30]}}$, based on the effective wavelengths. In the PCR and PLS models, root-mean-square error (RMSE) and correlation coefficient (R) are the two most important parameters. RMSE is closer to zero, whereas R is closer to one, implying the model's better prediction ability. Table <u>3</u> presents the results of different calibration models. The best performance was obtained by DOSC-SPA-PCR, which provides

		0	v
No.	Wavelength (nm)	No.	Wavelength (nm)
1	910	19	1906
2	930	20	1922
3	1066	21	1930
4	1614	22	1970
5	1674	23	2026
6	1690	24	2030
7	1706	25	2050
8	1710	26	2066
9	1734	27	2078
10	1746	28	2090
11	1750	29	2098
12	1754	30	2106
13	1782	31	2110
14	1790	32	2114
15	1814	33	2122
16	1850	34	2126
17	1866	35	2134
18	1898	36	2150

 Table 2. Effective Wavelengths Selected by SPA

Table 3. Results of Different Calibration Models

Spectral		Variable			
Set	Model	Number	R	RMSE	
X	PCR	311	0.991	0.009	
X	PLS	311	0.983	0.009	
X^{DOSC}	PCR	311	0.995	0.009	
X^{DOSC}	PLS	311	0.985	0.008	
X (SPA)	SPA-PCR	36	0.996	0.006	
X (SPA)	SPA-PLS	36	0.988	0.007	
$\begin{array}{c} X^{\text{DOSC}} \\ (\text{SPA}) \end{array}$	DOSC- SPA-PCR	36	0.998	0.005	
$\begin{array}{c} X^{\text{DOSC}} \\ (\text{SPA}) \end{array}$	DOSC- SPA-PLS	36	0.989	0.006	

a satisfactory calibration, RMSE 0.005, and R better than 0.99. DOSC and SPA made a significant contribution to the model, because most of the unrelated information in the original spectra has been removed. The results of DOSC-SPA-PLS and DOSC-SPA-PCR models are shown in Figs. <u>3(a)</u> and <u>3(b)</u>. The obtained DOSC-SPA-PLS and DOSC-SPA-PCR models are shown as follows:
$$\begin{split} Y_{\rm PLS} &= -0.19784X_{910\,\rm nm} - 0.00317X_{930\,\rm nm} + 0.20723X_{1066\,\rm nm} + 0.09309X_{1614\,\rm nm} + 0.21298X_{1674\,\rm nm} + 0.52350X_{1690\,\rm nm} \\ &+ 4.28272X_{1706\,\rm nm} - 2.75074X_{1710\,\rm nm} + 0.59575X_{1734\,\rm nm} + 1.34289X_{1746\,\rm nm} - 0.44838X_{1750\,\rm nm} - 10.30360X_{1754\,\rm nm} \\ &- 0.20254X_{1782\,\rm nm} + 0.48721X_{1790\,\rm nm} - 1.21490X_{1814\,\rm nm} - 0.44141X_{1850\,\rm nm} - 0.25346X_{1866\,\rm nm} + 0.49816X_{1898\,\rm nm} \\ &- 0.02357X_{1906\,\rm nm} - 0.13418X_{1922\,\rm nm} + 0.42399X_{1930\,\rm nm} - 0.10942X_{1970\,\rm nm} - 0.28962X_{2026\,\rm nm} - 0.53677X_{2030\,\rm nm} \\ &+ 0.04219X_{2050\,\rm nm} + 0.30724X_{2066\,\rm nm} - 0.03143X_{2078\,\rm nm} + 0.00698X_{2090\,\rm nm} - 0.31297X_{2098\,\rm nm} - 0.21812X_{2106\,\rm nm} \\ &+ 0.16126X_{2110\,\rm nm} + 0.44248X_{2114\,\rm nm} + 0.14626X_{2122\,\rm nm} + 0.07488X_{2126\,\rm nm} + 0.88256X_{2134\,\rm nm} + 0.26406X_{2150\,\rm nm} \\ &+ 0.63877, \end{split}$$

 $+ 0.47511X_{2050\,\mathrm{nm}} + 0.38152X_{2066\,\mathrm{nm}} - 0.72060X_{2078\,\mathrm{nm}} + 0.27836X_{2090\,\mathrm{nm}} - 0.35282X_{2098\,\mathrm{nm}} + 0.39210X_{2106\,\mathrm{nm}} - 0.08173X_{2110\,\mathrm{nm}} + 0.14682X_{2114\,\mathrm{nm}} - 0.45689X_{2122\,\mathrm{nm}} + 0.31962X_{2126\,\mathrm{nm}} + 0.73204X_{2134\,\mathrm{nm}} + 0.24402X_{2150\,\mathrm{nm}} + 0.73951,$

where X_{inm} is the transmittance at the wavelength of i nm, and Y_{PLS} and Y_{PCR} are the predicted values.

We used samples from the prediction set for the B1 brand for evaluating the prediction ability of the two models: DOSC-SPA-PLS and DOSC-SPA-PCR. There are 80 spectra for prediction. Figure 4 shows the predicted



Fig. 3. Calibration results of (a) DOSC-SPA-PCR and (b) DOSC-SPA-PLS. Fifteen samples per concentration were used for modeling, and there were 120 samples in total.



Fig. 4. Prediction results of (a) DOSC-SPA-PCR and (b) DOSC-SPA-PLS. Ten samples per concentration were used for prediction, and there were 80 samples in total.

results of the two models. Different RMSE of prediction (RMSEP) values can be calculated: DOSC-SPA-PLS 0.008 and DOSC-SPA-PCR 0.006, respectively, while the correlation coefficients (R_p) values are 0.985 and 0.990, respectively. After that, we used 0.8 g \cdot kg⁻¹ and 0.1 g \cdot kg⁻¹ from another brand of rapeseed oil (B2) to validate the ability of the models. There are 10 samples of 0.8 g \cdot kg⁻¹ and 0.1 g \cdot kg⁻¹ in the prediction set, respectively. By the standards of the oil manufacturer, for the concentration of 0.8 g \cdot kg⁻¹, the threshold can be 0.05. It means 0.8 ± 0.05 g \cdot kg⁻¹ can be considered as 0.8 g \cdot kg⁻¹. Likewise, 0.1 ± 0.05 g \cdot kg⁻¹ can be considered as 0.8 g \cdot kg⁻¹. Figure 5 presents the results of DOSC-SPA-PCR and DOSC-SPA-PLS for the concentration of 0.8 g \cdot kg⁻¹ and 0.1 g \cdot kg⁻¹ mineral oil for B2



Fig. 5. Prediction results of (a) DOSC-SPA-PCR and (b) DOSC-SPA-PLS for B2 rapeseed oil.

rapeseed oil. As can be seen from Fig. 5, in the DOSC-SPA-PCR model, the prediction results of 0.8 g \cdot kg⁻¹ and 0.1 g \cdot kg⁻¹ are with 90% and 100% accuracy, respectively, while in the DOSC-SPA-PLS model, the prediction results of 0.8 g \cdot kg⁻¹ and 0.1 g \cdot kg⁻¹ both are with a 90% accuracy. All of the results together demonstrate that DOSC-SPA-PCR could be considered as the better method for quantitative detection of mineral oil contamination in vegetable oil.

In conclusion, NIR spectroscopy exhibited good performance as a rapid, accurate, sensitive, reliable, and costeffective method for quantitative detection of mineral oil contamination in vegetable oil. Compared with previous studies, this study realized the rapid (about 30 s) quantitative detection of mineral oil contamination in vegetable oil with a limit of detection of 0.01 g \cdot kg⁻¹. The results of the transmittance curves and PCA indicate that the NIR analysis technique can realize the quantitative detection of mineral oil contamination in vegetable oil. Then, out of 311 wavelengths, only 36 wavelengths were selected by SPA as the effective wavelengths after DOSC preprocessing. After that, DOSC-SPA-PCR and DOSC-SPA-PLS analyses were performed in order to obtain a model suitable for the mineral oil rapid quantification. Relatively high R (0.998) and low RMSE (0.005) were obtained by the DOSC-SPA-PCR model compared with DOSC-SPA-PLS. We used two brands of rapeseed oil to validate the ability of the two models. A satisfactory accuracy with R and RMSE of prediction by the DOSC-SPA-PCR method of 0.990 and 0.006 for the B1 brand was obtained, and the accuracy of $0.8 \text{ g} \cdot \text{kg}^{-1}$ and $0.1 \text{ g} \cdot \text{kg}^{-1}$ for the B2 brand in the prediction set was 90% and 100%, respectively. The study indicates that the NIR spectroscopy combined with DOSC-SPA-PCR is appropriate for rapid quantitative detection of mineral oil contamination in vegetable oil with a limit of detection of $0.01 \text{ g} \cdot \text{kg}^{-1}$. The results may be useful for the development of a miniature spectrometer in mineral detection.

This work was partially supported by the National Natural Science Foundation of China (NSFC) (No. 61775140).

References

- S. Opinion, R. Efsagmouk, S. Seeds, and E. Panel, Food Sci. Nutr. 6, 38 (2010).
- A. Mukhopadhyay, A. Al-Haddad, M. Al-Otaibi, and M. Al-Senafy, *The European Food Safety Authority at Five* (Social Science Electronic Publishing, 2008).

- S. Moret, T. Populin, and L. S. Conte, Olives and Olive Oil in Health and Disease Prevention (Academic, 2010).
- P. A. Picouet, P. Gou, R. Hyypiö, and M. Castellari, J. Food Eng. 230, 18 (2018).
- S. Moret, T. Populin, L. S. Conte, K. Grob, and H. P. Neukom, Food Addit. Contam. 20, 417 (2003).
- H. P. Neukom, K. Grob, M. Biedermann, and A. Noti, Atmos. Environ. 36, 4839 (2002).
- 7. D. R. Tennant, Food Chem. Toxicol. 42, 481 (2004).
- 8. M. Biedermann and K. Grob, J. Chromatogr. A 1255, 76 (2012).
- M. Zoccali, L. Barp, M. Beccaria, D. Sciarrone, G. Purcaro, and L. Mondello, J. Sep. Sci. 39, 623 (2016).
- 10. M. Wrona, D. Pezo, and C. Nerin, Food Chem. 141, 3993 (2013).
- 11. M. Nestola and T. C. Schmidt, J. Chromatogr. A 1505, 69 (2017).
- N. C. A. Rashid, N. H. Ngajikin, A. I. Azmi, R. Arsat, S. Isaak, N. A. Cholan, and N. E. Azmi, Chin. Opt. Lett. **17**, 081701 (2019).
- J. de Jong, P. López, H. Mol, V. Baeten, J. A. F. Pierna, P. Vermeulen, U. Vincent, A. Boix, C. von Holst, and M. Tomaniova, Trac-Trend. Anal. Chem. 76, 203 (2015).
- L. Salguerochaparro, V. Baeten, J. A. Fernándezpierna, and F. Peñarodríguez, Food Chem. 139, 1121 (2013).
- M. I. Sánchez-Rodríguez, J. M. Caridad, E. Sánchez-López, A. Marinas, and F. J. Urbano, J. Sci. Food Agr. 99, 3417 (2019).
- J. Gromadzka and W. Wardencki, Pol. J. Food Nutr. Sci. 61, 89 (2011).
- A. Rácz, K. Héberger, and M. Fodor, Anal. Bioanal. Chem. 408, 6403 (2016).
- A. S. Luna, S. A. Da, J. S. Pinho, J. Ferré, and R. Boqué, Spectrochim. Acta A 100, 115 (2013).
- D. D. S. Fernandes, A. A. Gomes, G. B. D. Costa, G. W. B. D. Silva, and G. Véras, Talanta 87, 30 (2011).
- E. M. Paiva, J. J. R. Rohwedder, C. Pasquini, M. F. Pimentel, and C. F. Pereira, Fuel 160, 57 (2015).
- J. A. Westerhuis, S. D. Jong, and A. K. Smilde, Chemometr. Intell. Lab. 56, 13 (2001).
- H. Zhang, Z. Zhang, X. Zhao, X. Zhang, T. Zhang, C. Cao, and Y. Yu, Chin. Opt. Lett. 16, 103001 (2018).
- 23. J. Yang, J. Sun, L. Du, B. Chen, and W. Gong, Opt. Express 25, 3743 (2017).
- 24. M. Kamruzzaman, G. Elmasry, D. W. Sun, and P. Allen, J. Food Eng. **104**, 332 (2011).
- D. Jie, L. Xie, X. Fu, X. Rao, and Y. Ying, J. Food Eng. 118, 387 (2013).
- D. C. Gean Bezerra, D. D. S. Fernandes, A. A. Gomes, D. A. Valber Elias, and V. Germano, Food Chem. 196, 539 (2016).
- M. C. U. Araújo, T. C. B. Saldanha, R. K. H. Galvão, T. Yoneyama, H. C. Chame, and V. Visani, Chemometr. Intell. Lab. 57, 65 (2001).
- 28. Y. Ma, W. Gong, L. Wang, M. Zhang, and J. Yang, Opt. Express 24, 8170 (2016).
- 29. Y. Zheng, X. Zhu, Z. Wang, Z. Hou, F. Gao, R. Nie, X. Cui, J. She, and B. Peng, Chin. Opt. Lett. 8, 083001 (2017).
- M. Febrero-Bande, P. Galeano, and W. González-Manteiga, Int. Stat. Rev. 85, 61 (2015).