Contents lists available at ScienceDirect

Food Control

Identification of coriander oil adulteration using a portable NIR spectrometer

Karine Cristine Kaufmann [a], Klicia Araujo Sampaio [a], Juan F. García-Martín [b], Douglas Fernandes Barbin [a],[*]

[a] Department of Food Engineering, University of Campinas (UNICAMP), Rua Monteiro Lobato, 80, Cidade Universitária, Campinas-SP, 13083-862, Brazil
[b] Departamento de Ingeniería Química, Facultad de Química, Universidad de Sevilla, C/ Professor García González, s/n, 41012, Sevilla, Spain

ARTICLE INFO

ABSTRACT

Coriander oil is a vegetable oil extracted from coriander seed that has about 70% of petroselinic acid, apart from anti-inflammatory and anti-aging properties, thus gaining the status of new food ingredient. Due to its properties and added value, it can become the target of adulteration as occurs with other edible vegetable oils of high market value. Therefore, the objective of this work was to identify the authenticity of coriander oil and adulteration with other commercial vegetable oils such as palm olein, canola oil and soybean oil. Principal component analysis (PCA) differentiated the matrices of pure oils using 3 principal components, which explained 87% of the variance. Linear discriminant analysis (LDA) and k-nearest neighbors algorithm (k-NN) were used to classify pure oil samples and adulterated coriander oils. Partial Least Squares (PLS) regression models presented coefficient of determination (R^2) of 0.98, 0.99 and 0.99, for coriander oil adulterated with palm olein soybean oil and canola oil, respectively. RPD was between 7.1 and 10, which indicates robust models that can be used for quality control during the processing of coriander oil.

1. Introduction

Coriander oil is extracted from seeds of *Coriandrum sativum* L., a plant cultivated in several regions worldwide, since its leaves are widely used in cooking and as a medicine. The seed contains around 28% of vegetable oil and approximately 0.9% of essential oils (volatile oils). These two oil fractions have important compounds such as petroselinic acid, found in the vegetable oil and linalol, which represents roughly 70% of these respective oil fractions (Sahib et al., 2013).

Petroselinic acid is the most abundant fatty acid present in coriander vegetable oil, and its importance is related to its chemical properties. Although this fatty acid has the same number of carbons and double bond number than oleic acid, what makes it unusual is the position of the double bond that has made it a positional isomer of oleic acid. The different position of the double bond within the carbon chain changes the physicochemical properties of petroselinic acid, which has a melting point at 30 °C at 1 atm, while oleic acid melting point is 14 °C at the same pressure (Delbeke et al., 2016). In addition, petroselinic acid has significant potential for the chemical industry, as oxidative cleavage leads to industrially interesting compounds such as lauric acid, a

commercial surfactant, and adipic acid, a precursor to Nylon 66 (Kleiman, 1990; Uitterhaegen et al., 2016).

Besides its chemical interests, fatty acids of coriander oil have attractive properties, such as anti-aging and anti-inflammatory activity, making them suitable for the cosmetic and functional food industry (Sahib et al., 2013; Uitterhaegen et al., 2016). Due to its functional properties, coriander oil has recently been labeled as a New Food Ingredient (NFI) in the context of Regulation (EC) No 258/97 of the European Parliament and up to 600 mg/day consumption by healthy adults as a dietary supplement has been considered safe (EFSA, 2013). Given the potential for application in the chemical, pharmaceutical and food industries, coriander oil is an important raw material that can be the target of adulteration to obtain a higher profit margin, as with other high-priced vegetable oils. Neem oil (Elzey, Pollard, & Fakayode, 2016), avocado (Rohman et al., 2016), sandalwood (Kuriakose, Thankappan, Joe, & Venkataraman, 2010), pumpkin seed oil (Van Hoed et al., 2017) and rosehip (De Santana et al., 2016) oils have been reported to be liable to adulteration with other lower-quality or cheaper oils.

Oil adulteration presents a global challenge for the production of natural oils. The development of food authenticity protocols capable of

---

determining the authenticity of natural products is of great interest for financial and safety reasons. Effective analytical methods for determination of adulteration include gas chromatography and spectroscopic techniques (Elzey et al., 2016). However, these techniques are expensive and require high investment for implementation.

Near infrared spectroscopy (NIRS) is a vibrational spectroscopy technique that uses electromagnetic radiation in the range of 780–2500 nm. Among other vibrational techniques (mid infrared spectroscopy, RAMAN spectroscopy), NIRS has the advantage of analyzing several components simultaneously, most of the times not requiring sample preparation, placing itself as a promising tool for rapid or online analysis without loss of fundamental chemical information (Armenta, Moros, Garrigues, & de la Guardia, 2010; García Martín, 2015; Wang, Sun, Zhang, & Liu, 2016). The use of NIRS to obtain information on pharmacochemical properties and adulteration data for edible vegetable oils is well studied due to the low cost required for the acquisition and implementation of this technique (Azizian, Mossoba, Reza, Kia, & Delmonte, 2015; Giovenzana, Beghi, Civelli, & Marai, 2015; Jinkui, Han, Tomohiro, Yelian, & C, 2014; Kuligowski, Carrión, Quintás, Garrigues, & Guardia, 2012; Laroussi-mezghani et al., 2015; Li et al., 2012; Man & Moh, 1998; Zeng et al., 2014).

Advances in electronic systems have provided innovations for the use of portable NIR spectrometers in the processing line. This improvement, in addition to making the equipment cheaper, when compared to traditional benchtop equipment available in the market, also improved speed of response, reduced energy costs and sample preparation time (Santos et al., 2021; Teye, Elliott, Sam-Amoah, & Mingle, 2019). Portable NIR spectrometers can also be employed to obtain edible vegetable oil quality parameters of palm oil adulterated with lard (Basri et al., 2017), adulteration of palm oil with Sudan dyes (Teye et al., 2019), identification and quantification of adulterations in extra virgin olive oils (Borghi et al., 2020), etc.

Generally, methods of vibrational spectroscopy are used in combination with chemometrics. These mathematical and statistical methods are important for obtaining good results, since they facilitate the extraction of information from the NIR spectra that most of the times are large and overlay bands (Wang et al., 2016). NIRS and chemometrics proved to be a powerful tool for identifying adulteration. Generally, traditional grading methods were employed to detect whether the target oil was adulterated with other oils. Chemometrics such as partial least squares (PLS) calibration, have been successfully applied in olive oil tampering detection (Zhang et al., 2017), while supervised soft independent modeling of class analogy (SIMCA) and linear discriminant analysis (LDA) methods were used (Sinelli, Cerretani, Egidio, Bendini, & Casiraghi, 2010) to classify olive oil samples by denomination of origin.

Considering the economic importance of coriander oil and its attractiveness to adulteration, this study aims to identify coriander oil adulteration using a low-cost, portable NIR spectrometer, in tandem with chemometric methods. Several cheaper vegetable oils were tested as adulterants, and the results could provide a potential fast and low-cost method for inspection agents and producers.

2. Material and method

2.1. Oil samples

Coriander seeds were acquired in two different batches and cold pressed for oil extraction in a continuous press. Solid particles were removed by press filtering (2.0 kgf/cm^2) the oil obtained. Three different samples of canola oil and three samples of soybean oil used as adulterants were acquired from the local market in Campinas, Brazil, while two different samples of palm olein were obtained from an oil processing industry (Limeira-SP, Brazil).

2.2. Sample preparation

Coriander oil was adulterated in twelve different concentrations (1%, 3%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% (wt.)), in triplicate, of other types of oils. Hence, it was prepared 108 samples (12 concentrations x 3 samples x 3 repetitions) of coriander oil adulterated with canola oil, 108 samples of coriander oil adulterated with soybean oil, and 72 samples (12 concentrations x 2 samples x 3 repetitions) of coriander oil adulterated with palm olein. Also, pure samples of coriander oil (0% concentration of adulteration) and pure adulterants (100% concentration) were used in the classification and prediction models.

The samples were split into calibration and validation sets for the regression models, with different level of adulteration for each dataset, to guarantee an external validation. Pure samples were used in the calibration set. Hence, the calibration set consisted of pure coriander oil (0% adulteration), 3%, 10%, 30%, 50%, 70%, 90% adulteration, and pure adulterant (100%), totaling 59 samples. The validation dataset was comprised of samples with 1%, 5%, 20%, 40%, 60%, 80% of adulterant, totaling 54 samples.

2.3. Determination of fatty acid (FA) composition

Pure oils were characterized for their fatty acid (FA) compositions, expressed as relative percentage of total oil determined by gas chromatography (GC) following the AOCS official method Ce 1–62 (AOCS, 2009). Conversion of FA to fatty acid methyl esters (FAME) was performed based on the method described by Hartman and Lago (1973). Gas chromatography was performed in a (GC) Clarus 600 (PerkinElmer, USA) with a flame ionization detector (FID) and a DB-WAX capillary column (30m length, 0.25 mm internal diameter, 0.25 μm film thickness) (Agilent Technologies, USA). Helium was used as carrier gas (1.78 mL/min), FID temperature of 250 °C, with 1 μL injection at 250 °C, column temperature ramp from 50 °C to 250 °C at 10 °C/min. The total FA results were verified by comparing the retention times of each analysis with the retention peaks of an external standard (FAME mix C8 – C24; Sigma-Aldrich, USA), using Total Chrom software (version 6.3.2, PerkinElmer, USA).

2.4. NIR spectra acquisition

Near infrared spectra of the samples at 25 °C were obtained using a portable NIR equipment (TIDA-00554 DLP NIRScan Nano, Texas Instruments, Dallas, TE, USA) by diffuse reflectance measurements in the range of 900–1700 nm. Oil samples were scanned using a cuvette (2 mm optical path), where transflectance was measured with a *Spectralon* disk in the back to avoid light scattering.

2.5. Chemometric analysis

NIR spectra were pre-processed with smoothing and 1st Savitzky-Golay derivative (15 points). Principal component analysis (PCA) was applied as an exploratory and unsupervised method to investigate the differences in spectra of pure oil samples. This technique consists of linear transformation of the original information, in this case the NIR spectra, which are highly correlated with each other into a smaller set of non-correlated variables, called principal components (PCs) (Brasil et al., 2021). The scores of PCs provide information about the differences among samples, while the PC loadings provide information about the most important wavelengths related to the given set of samples.

2.5.1. Classification of adulterated coriander oil

Linear discriminat analysis (LDA) and k-Nearest Neighbors (k-NN) were used for classification of coriander oil samples adulterated with soybean, canola and palm oil with concentrations of 3%, 10% and 70%, pure adulterant and pure coriander oil. Two different set of samples

were prepared for calibration (60% of samples) and validation (40% of samples).

2.5.2. *Prediction of coriander oil adulteration*

Partial least squares regression (PLSR) was used for prediction of the concentration of adulterant (canola oil, palm olein or soybean oil) in coriander oil. Calibration models included samples of pure coriander oil (0% adulteration), and samples adulterated with the following concentrations: 3%, 10%, 30%, 50%, 70%, 90%, 100% (pure adulterant). Diiferent concentrations were prepared for the validation model, to provide an external set of samples, adulterated in the concentrations of 1%, 3%, 5%, 20%, 40%, 60%, 80% of adulterant. Chemometric analyses were performed using PLS Toolbox software from Matlab R2007b (Matlab Inc., Natick, MA, USA), and the LDA models were constructed using MINITAB Release 14 (Minitab Ltd., Coventry, United Kingdom).

Variable selection was performed using Interval partial least squares (iPLS), forward stepwise, and the PCA loadings with larger absolute value, and identification of informative region. iPLS is a variable selection method that fragments the spectrum into smaller information intervals and for each interval a PLS regression is performed using the parameter of interest. This method can significantly improve the prediction of PLS models (Magwaza et al., 2012). Forward stepwise is a variable selection method that eliminates collinearity or highly correlated variables, analyzing the combination of all variables step by step (Chen & Xie, 2014). PCA loadings were used to select the most important variables for discrimination between coriander oil and adulterant oil. The informative region was selected as the region of the spectra with peaks and differences among samples.

Robustness of classification models were assessed by accuracy (Acc, Eq. (1)), sensibility (SEN, Eq. (2)), and selectivity (SEL, Eq. (3)).

$$\text{Acc} = \frac{TP + TN}{TP + FN + FP + TN} \times 100 \tag{1}$$

$$\text{SEN} = \frac{TP}{TP + FN} \times 100 \tag{2}$$

$$\text{SEL} = \frac{TN}{FP + TN} \times 100 \tag{3}$$

Where: TP true positives; TN true negatives; FP false positives; FN false negatives.

Robustness of prediction models was assessed by root mean squared errors for calibration (RMSEC, Eq. (4)), cross-validation (RMSECV) and prediction (RMSEP, Eq. (5)), and the ratio of performance to deviation (RPD, Eq. (6)) (García Martín, 2015; Kaufmann et al., 2019).

$$\text{RMSEC} = \frac{\sqrt{\sum_{i=1}^{n} (y_{\text{calibration}} - y_{\text{measured}})^2}}{n} \tag{4}$$

$$\text{RMSECV, RMSEP} = \frac{\sqrt{\sum_{i=1}^{k} (y_{\text{predicted}} - y_{\text{measured}})^2}}{k} \tag{5}$$

$$\text{RPD} \approx \sqrt[2]{\frac{1}{1 - r^2}} \tag{6}$$

Where: $y_{predicted}$ is the value of the property of interest predicted by the model; $y_{measured}$ is the value of the measured property of interest for the sample by the reference method (in this case GC); n is the number of samples used in the calibration set and k is the number of samples used in the external validation set.

In order to obtain adequate PLS models, the balance of the relationship between the bias and the variance explained by the model must be studied. With the increase in the number of latent variables (LV), the complexity of the model increases, so the best prediction model is the one that has the least prediction error. The ideal model should have bias equal to zero, models with values less than zero are under-adjusted and models with bias greater than zero are over-adjusted (Kalivas & Palmer, 2014; Mas, Rubio, Valverde-Som, Sarabia, & Ortiz, 2020).

3. Results and discussion

3.1. *Characterization of pure oils*

All samples investigated presented fatty acid (FA) profiles (Table 1) similar to those reported in the literature (Chowdhury, Banu, Khan, & Latif, 1970; Farhoosh, Esmaeilzadeh Kenari, & Poorazrang, 2009; Sampaio et al., 2017). Petroselinic acid (68–84%) and linolenic acid (13–16%) were the predominant FA for coriander oil (Sahib et al., 2013). As illustrated in Table 1, the contents of petroselinic and oleic acids in coriander oil are noticeably different from the other oils. The greatest differences between these oils are in petroselinic acid (higher in coriander oil), lineloeic acid (higher in soybean oil) and oleic acid (higher in canola oil and palm olein). The major fatty acid component of canola oil and palm olein is oleic acid; it is also noticeable that coriander oil and palm olein have high palmitic acid content.

3.2. *NIR spectra of oil samples and exploratory analysis*

NIR spectra for different types of oils were acquired (Fig. 1A). It was possible to verify that the samples presented peaks in the region at 1200 nm and 1400 nm, similar to other oil samples previously reported (Basri et al., 2017; Teye et al., 2019). Bands of the NIR spectroscopy method are wide, which causes overlap, so it is necessary to use multivariate analysis to extract the information contained in the spectrum. Spectra were pre-treated with Savitzky-Golay smoothing filter (15 points) and Savitzky-Golay 1st derivative 2nd order (7 windows, 15 points) (Fig. 1B). Principal component analysis was applied to investigate the differences in spectral information of samples (Fig. 1C and D).

Using two principal components (PC), it is possible to observe that the matrices of pure oils could be separated in the scores plot (Fig. 1C). PC1 accounts for 86.88% of variance, discriminating palm olein and coriander oil (both on the positive side of PC1) from canola and soybean oils (both on the negative side of PC1, and also separated from each other). PC2 is responsible for 7.21% of variance, and is more important for discriminating coriander oil (in the positive side of PC2) from the other oils (near zero (0) or in the negative side of PC2).

The differences observed in the PCA could be related to the fatty acid composition (Table 1), since the loadings from PC1 and PC2 (Fig. 1D) present peaks and valleys related to C–H bonds. It is possible to observe a peak in the region of 1185 nm, responsible for the HC=CH stretch in the second overtone (Jinkui et al., 2014). Stretching overtone of C–H at 1225 nm (Teye et al., 2019) is related to oil compositions that have a greater number of double bonds, observed in PC2, which separates coriander oil from all other oils.

Table 1
Fatty acid composition (% wt.) of the assayed oils.

| Fatty Acid	Nomenclature	Oil			
		Coriander	Soybean	Canola	Palm olein
Palmitic Acid	C16:0	3.7	11.13	4.75	24.73
Palmitoleic Acid	C16:1	0.07	–	0.26	–
Stearic acid	C18:0	0	3.49	2.53	3.04
Petroselinic Acid	C18:1 (6)ω12	79.79	–	–	–
Oleic Acid (Omega 9)	C18:1 (9)ω9	0.62	26.07	62.82	47.27
Linoleic Acid (Omega 6)	C18:2	14.87	52,89	18.51	14.63
Linolenic Acid (Omega 3)	C18:3	–	4.55	7.77	–
Araquic Acid	C20:0	–	–	0.95	–
Eicosanoic Acid	C20:1	–	–	0.78	–
Eicosadiene acid	C20:2	–	–	1.08	–
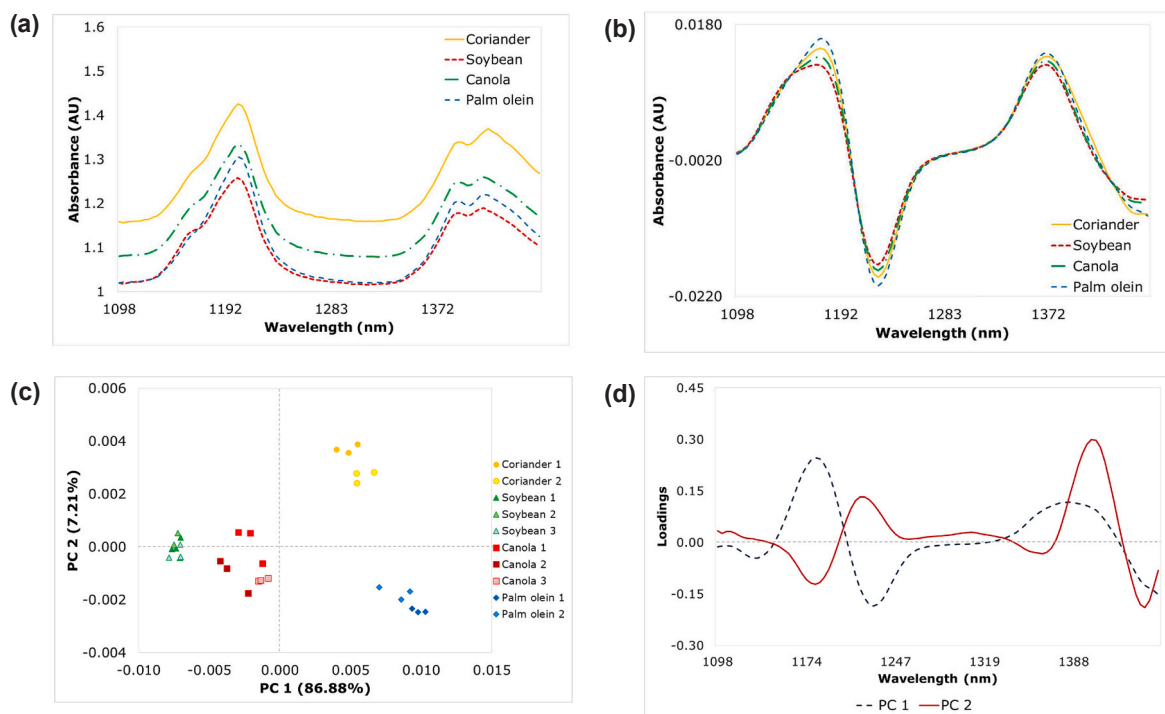
Fig. 1. (a) Raw spectra of vegetable oil samples; (b) Pre-processed spectra of pure vegetable oil samples (1st derivative plus symbol smoothing); (c) PCA scores of pure vegetable oil samples; (d) PCA loadings of pure vegetable oils samples.

3.3. Classification models

3.3.1. Variable selection

In this study, we compared five different inputs for prediction, i. e, full spectra, using the informative region, and applying variable selection (iPLS, forward stepwise and PCA loadings. The wavelengths selected for each method are presented in Table 2, and further prediction models using the selected wavelengths from each method are reported.

Variable selection are used to remove uninformative variables and noise, providing more robust models with less influence from external variation, mainly in investigations of adulteration that have raw spectra very similar to each other (Heise, Damm, Lampen, Davies, & Mcintyre, 2005). Variable selection methods were applied to verify the improvement of the classification obtained by the models. These results can be seen in Table 3 and Table 4.

3.3.2. Supervised classification models

Supervised classification methods LDA and k-NN were developed using coriander oil adulterated with canola oil, soybean oil and palm olein, (Table 3).

The k-NN model was evaluated for different variable selections using 3 nearest neighbors: informative region (1100–1450 nm), forward stepwise and PCA loadings. For LDA models, the same variable selections were used except for informative region, because for the analysis of

Table 2
Variables selected from samples spectra using different methods.

Method selection	Variables (nm)
Full spectrum	900–1700
Informative region	1100–1450
Stepwise (15 variables)	1102; 117; 1132; 1151; 1174; 1215; 1254; 1270; 1283; 1340; 1375; 1378; 1413; 1447; 1500
PCA Loadings (8 variables)	1131; 1181; 1221; 1228; 1365; 1388; 1402,1443
iPLS (10 variables)	1208; 1215; 1389; 1523; 1571; 1624; 1634; 1648; 1677; 1692

LDA it is necessary to eliminate highly correlated variables and the number of variables cannot be higher than the number of samples.

LDA models using variables selected by stepwise showed better results for both calibration and validation than models with variables selected in the PCA loadings. The LDA-stepwise showed 99% accuracy of calibration and validation in the models used to classify coriander oil adulterated with palm olein, compared to 91% accuracy obtained by k-NN in the validation set. For samples of coriander oil adulterated with soybean oil, the best results were obtained for LDA using stepwise variable selection (93% accuracy) and k-NN using informative region (95% accuracy). For the models using canola oil as an adulteration source, accuracy was slightly lower when compared to classification models for samples adulterated with soybean oil and palm olein. Better results were obtained using stepwise as a variable selection method, reaching accuracy higher than 90% for both LDA and KNN models. Previous applications of LDA for classification of oil samples using NIR spectra have been previously reported (Sinelli et al., 2010; Yang, Irudayaraj, & Paradkar, 2005; Zhou, Liu, Li, & Chen, 2015), with accuracy around 70%, depending on the complexity and differences among the samples evaluated.

It was observed that the lowest accuracy for classification was obtained using canola oil as adulterants. It was observed that canola oil has a higher content of oleic acid (Table 1), which is the positional isomer of petroselinic acid, found in larger amount in coriander oil. As a result, the similarity between the major fatty acids present in canola oil and in coriander oil could be responsible for the models' lower accuracy.

The identification of samples of coriander oil adulterated with soybean oil, canola and palm oil can be explained mainly by the difference in the composition of major fatty acids. For soybean oil the major compound is linoleic acid with two double bonds and for canola oil and palm olein it is oleic acid with a double bond. Although petroselinic acid also has a double bond, the position of the double bond in the chain is different, and this variation influenced the results. Regarding palm olein has the high content of palmitic acid, that is saturated, may explain the better results obtained.

Classification models using k-NN with 3 neighbors (k) were applied

Table 3
Results for supervised classification methods with different variable selection methods for coriander oil adulterated with palm olein, canola oil, and soybean oil.

Calibration															
	LDA						k-NN								
	Stepwise			PCA Loadings			Informative region			Stepwise			PCA Loadings		
	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean
Selectivity	0.99	0.97	0.98	0.92	0.95	0.96	0.99	0.83	0.97	0.92	0.94	0.91	0.92	0.89	0.90
Sensitivity	0.98	0.93	0.91	0.73	0.83	0.81	0.98	0.52	0.90	0.75	0.81	0.55	0.73	0.63	0.55
Accuracy	0.99	0.96	0.97	0.89	0.92	0.93	0.99	0.77	0.96	0.88	0.92	0.85	0.88	0.84	0.84
Validation															
	Stepwise			PCA Loadings			Informative region			Stepwise			PCA Loadings		
	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean	Palm	Canola	Soybean
Selectivity	0.99	0.94	0.95	0.93	0.89	0.96	0.98	0.92	0.97	0.94	0.94	0.91	0.86	0.89	0.91
Sensitivity	0.98	0.78	0.83	0.75	0.68	0.84	0.93	0.64	0.86	0.76	0.77	0.63	0.61	0.64	0.67
Accuracy	0.99	0.90	0.93	0.88	0.84	0.93	0.97	0.88	0.95	0.91	0.91	0.86	0.80	0.84	0.86

Palm is coriander oil adulterated with palm olein; Canola is coriander oil adulterated with canola oil; Soybean is coriander oil adulterated with soybean oil.

to FTIR spectra to classify olive oils, (Jiménez-Carvelo, Osorio, Koidis, González-Casado, & Cuadros-Rodríguez, 2017), and FT-NIR spectra to evaluate oil storage time (He, Jiang, & Chen, 2020). Although they presented good classification results, the performance of k-NN was inferior to models with LDA and SVM. Results presented in this work (Table 3) show that k-NN models can present high accuracy depending on the variables selection method and the complexity of the samples analysed.

3.4. PLS-regression

Partial least squares regression models were built to quantify the percentage of adulterant (canola oil, soybean oil or palm olein) in coriander oil. Calibration models were built using the full spectrum (900–1700 nm), informative region (1100–1450 nm), forward stepwise and iPLS for each method of variable selection. Previously, smoothing filter treatments and 1st Savitzky-Golay derivative were used (Table 4).

The PLS models with the lowest RMSECV and RMSEP and the highest R^2 were the models for prediction of adulteration of coriander oil with palm olein using the spectra between 1100 and 1450 nm pre-processed with smoothing and 1st derivative, with R^2 of 0.98 and RMSEP of 4.12% (wt.) using only 2 latent variables (LV). Notwithstanding, for the full spectrum it was possible to reach R^2 of 0.99 using 6 LV, and the RMSEP presented a higher value (4.26% wt.), indicating that the use of the full spectrum can increase the complexity of the model without bringing results that justify the increase in the number of wavelengths. However, when the number of wavelengths used to build the PLS model was reduced with stepwise and iPLS, the models presented worse values of RMSEP and R^2 in addition to increasing the number of LV.

The PLS models for quantification of soybean oil added to coriander oil showed better results using smoothing and 1st derivative, reaching RMSEP of 5.4% (wt.) and R^2 of 0.98 using 3 LV. Although the PLS models for the prediction of the amount of palm olein and soybean oil have similar R^2 for all pre-processing methods investigated, a higher number of LV was needed for palm olein, with a higher RMSEP, in comparison to models for prediction of coriander oil adulterated with soybean oil. Prediction of coriander oil adulterated with canola oil showed higher values of RMSEP, with lower R^2 and RPD, which could be related to the similarity between the chains of major fatty acids in canola oil and coriander oil when compared to the other adulterants.

These results show that the models provide good results to identify the percentages of different adulterants. The use of the selection of variables from the informative region, eliminating wavelengths that do not contain useful information about the samples, are widely used for regression models to identify the adulteration of edible vegetable oils (Heise et al., 2005). Literature data using a portable NIR spectrometer and PLS to predict palm oil adulteration (Basri et al., 2018), and quantification of copaiba oil adulteration (de Oliveira Moreira, de Lira

Machado, de Almeida, & Braga, 2018) presented $R^2 = 0.99$, similar to the best results presented in this work.

PLS regression models concentrate the most important spectral data for determining the predicted data (Elzey et al., 2016). Therefore, the number of latent variables (LV) was chosen according to the reduction of the BIAS and the RMSEP (Table 4). When more latent variables are used, BIAS and RMSEP increase, while if a lower LV number is used, the models can exclude important information. Good predictive models have low BIAS and small errors, thus increasing the accuracy of the prediction model implies minimizing the RMSEP.

RPD (Table 4) is an important quality parameter for PLS models using NIR spectra. Calibration models that show values between 2.5 and 5 can be used as a screening, whereas RPD values in the range between 5 and 10 range are robust models that can be used for quality control (Du et al., 2012; García Martín, López Barrera, Torres García, Zhang, & Álvarez Mateos, 2019). All regression models presented have RPD above 2.5 with values up to 7.07, indicating that they can be used for sample screening and quality analysis.

4. Conclusion

The results obtained demonstrate that a portable NIR spectrometer can identify pure coriander oil among other samples of soybean oil, canola oil and palm olein. Supervised classification models LDA and k-NN can be used to assess the authenticity of coriander oil and the choice of the most appropriate method depends on the type of oil used as adulterant. LDA-stepwise presented better results to identify coriander oil adulterated with palm olein, canola oil and soybean oil.

For the quantification of the adulterants in coriander oil, the PLS models with Savitzky-Golay smoothing and 1st derivative presented the best results for all types of adulterants investigated. As for variable selection, the informative region (1100–1450 nm) showed better results, compared to models using the full spectra or few wavelengths selected by stepwise and iPLS methods. The results obtained indicate that portable NIR spectrometers are promising tools for identification of coriander oil adulteration, and could be used for fast authentication and quality control in the oil processing industry.

Credit author statement

Karine Cristine Kaufmann: Conceptualization, Methodology, Writing - Original Draft. Klicia Araujo Sampaio: Investigation, Writing – Review, Supervision. Juan F. García-Martín: Writing - Review & Editing. Douglas Fernandes Barbin: Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration.

Table 4
PLS regression models for coriander oil adulterated with canola oil, palm olein and soybean oil.

	900–1700 nm				1100–1450 nm				Stepwise				iPLS			
	Raw	SG	1D	SG+1D	Raw	SG.	1D	SG+1D	Raw	SG	1D	SG+1D	Raw	SG	1D	SG+1D
Coriander oil – canola oil																
LV	6.00	7.00	7.00	6.00	3.00	3.00	2.00	2.00	3.00	3.00	2	4	6	6.00	3.00	3.00
RMSECV	27.54	33.12	14.30	9.52	25.28	25.46	32.78	32.41	29.92	35.66	30.25	43.40	24.82	19.73	17.7	16.91
RMSEP	13.07	9.17	8.93	8.34	11.25	11.23	13.21	12.68	11.34	11.71	12.64	14.06	13.57	13.77	13.27	13.66
BIAS pred	4.94	−0.79	−0.24	0.35	−1.53	−1.48	−3.50	−3.59	−1.96	−1.193	−3.91	−1.96	6.37	6.97	6.40	6.70
R2pred	0.88	0.89	0.94	0.94	0.90	0.90	0.86	0.88	0.90	0.87	0.86	0.78	0.86	0.86	0.80	0.80
RPD	2.89	3.02	4.08	4.08	3.16	3.16	2.67	2.89	3.16	2.77	2.67	2.13	2.67	2.67	2.24	2.24
Coriander oil - palm olein																
LV	3.00	4.00	7.00	6.00	3.00	3.00	2.00	2.00	4.00	4.00	4.00	4.00	5.00	3.00	3.00	4.00
RMSECV	17.17	11.62	3.95	4.59	7.02	7.19	5.04	4.86	9.16	10.01	11.20	10.17	15.42	16.79	20.32	18.89
RMSEP	13.33	7.65	5.02	4.26	5.98	6.47	4.14	4.12	7.38	9.37	8.97	10.65	7.57	13.25	12.20	12.02
BIAS pred	−0.18	0.03	−0.10	0.33	−0.30	−0.44	−0.89	−0.85	0.61	0.15	−0.17	−0.66	0.77	−0.18	0.59	0.47
R2pred	0.81	0.95	0.99	0.99	0.97	0.96	0.98	0.98	0.94	0.92	0.92	0.91	0.95	0.80	0.81	0.82
RPD	2.29	4.47	10	10	5.77	5.00	7.07	7.07	4.08	3.54	3.54	3.34	4.47	2.24	2.29	3.54
Coriander oil - soybean oil																
LV	4.00	7.00	5.00	6.00	4.00	4.00	3.00	3.00	4.00	3.00	4.00	2.00	4.00	6.00	4.00	5.00
RMSECV	14.23	5.36	6.52	5.68	5.08	5.51	5.69	5.13	7.12	11.51	8.32	11.34	18.32	11.61	13.73	14.26
RMSEP	11.63	4.41	5.91	5.09	4.23	4.84	5.11	5.4	6.43	9.97	6.97	9.72	11.90	11.05	11.76	13.83
BIAS pred	−0.28	0.83	−0.72	0.90	−0.06	−0.07	−0.37	−0.37	−1.45	−0.73	−1.34	0.10	2.35	3.033	4.31	3.52
R2pred	0.86	0.98	0.97	0.98	0.98	0.98	0.98	0.98	0.96	0.90	0.95	0.90	0.88	0.89	0.88	0.83
RPD	2.67	7.07	5.77	7.07	7.07	7.07	7.07	7.07	5.00	3.16	4.47	3.16	2.89	3.02	2.89	2.43

Raw: raw spectra, SG: Savitzty-Golay smoothing; 1st Deriv.: 1st Derivative Savitzky-Golay; SG + 1st Deriv.: Savitzky-Golay Smoothing and 1st Derivative.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

Armenta, S., Moros, J., Garrigues, S., & de la Guardia, M. (2010). The use of near-infrared spectrometry in the olive oil industry. *Critical Reviews in Food Science and Nutrition, 50*(6), 567–582. https://doi.org/10.1080/10408390802606790

Azizian, H., Mossoba, M. M., Reza, A., Kia, F., & Delmonte, P. (2015). Novel , rapid identification , and quantification of adulterants in extra virgin olive oil using near - infrared spectroscopy and chemometrics. *Lipids, 50*(7), 705–718. https://doi.org/10.1007/s11745-015-4038-4

Basri, K. N., Hussain, M. N., Bakar, J., Sharif, Z., Khir, M. F. A., & Zoolfakar, A. S. (2017). Classification and quantification of palm oil adulteration via portable NIR spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 173*, 335–342. https://doi.org/10.1016/j.saa.2016.09.028

Basri, K. N., Laili, A. R., Tuhaime, N. A., Hussain, M. N., Bakar, J., Sharif, Z., et al. (2018). FT-NIR, MicroNIR and LED-MicroNIR for detection of adulteration in palm oil via PLS and LDA. *Analytical Methods, 10*(34), 4143–4151. https://doi.org/10.1039/c8ay01239c

Borghi, F. T., Santos, P. C., Santos, F. D., Nascimento, M. H. C., Corrêa, T., Cesconetto, M., et al. (2020). Quantification and classification of vegetable oils in extra virgin olive oil samples using a portable near-infrared spectrometer associated with chemometrics. *Microchemical Journal, 159*, 105544. https://doi.org/10.1016/j.microc.2020.105544

Chen, M., & Xie, L. (2014). A preliminary study of aflatoxin b1 detection in peanut oil by terahertz time-domain spectroscopy. *Transactions of the ASABE, 57*(6), 1793–1799. https://doi.org/10.13031/trans.57.10725

Chowdhury, K., Banu, L., Khan, S., & Latif, A. (1970). Studies on the fatty acid composition of edible oil. *Bangladesh Journal of Scientific & Industrial Research, 42*(3), 311–316. https://doi.org/10.3329/bjsir.v42i3.669

De Santana, F. B., Gontijo, L. C., Mitsutake, H., Mazivila, S. J., De Souza, L. M., & Borges Neto, W. (2016). Non-destructive fraud detection in rosehip oil by MIR spectroscopy and chemometrics. *Food Chemistry, 209*, 228–233. https://doi.org/10.1016/j.foodchem.2016.04.051

Delbeke, E. I. P., Everaert, J., Uitterhaegen, E., Verweire, S., Verlee, A., Talou, T., et al. (2016). Petroselinic acid purification and its use for the fermentation of new sophorolipids. *AMB Express, 6*, 28. https://doi.org/10.1186/s13568-016-0199-7

Du, R., Lai, K., Xiao, Z., Shen, Y., Wang, X., & Huang, Y. (2012). Evaluation of the quality of deep frying oils with fourier transform near-infrared and mid-infrared spectroscop. *Journal of Food Science, 77*(2), 261–266. https://doi.org/10.1111/j.1750-3841.2011.02551.x

Efsa, E. F. S. A. (2013). Scientific Opinion on the safety of " coriander seed oil " as a Novel Food. *EFSA Journal, 11*(10), 1–20. https://doi.org/10.2903/j.efsa.2013.3422

Elzey, B., Pollard, D., & Fakayode, S. O. (2016). Determination of adulterated neem and flaxseed oil compositions by FTIR spectroscopy and multivariate regression analysis. *Food Control, 68*, 303–309. https://doi.org/10.1016/j.foodcont.2016.04.008

Farhoosh, R., Esmaeilzadeh Kenari, R., & Poorazrang, H. (2009). Frying stability of canola oil blended with palm olein, olive, and corn oils. *JAOCS, Journal of the American Oil Chemists' Society, 86*(1), 71–76. https://doi.org/10.1007/s11746-008-1315-x

García Martín, J. F. (2015). Optical path length and wavelength selection using Vis/NIR spectroscopy for olive oil's free acidity determination. *International Journal of Food Science and Technology, 50*(6), 1461–1467. https://doi.org/10.1111/ijfs.12790

García Martín, J. F., López Barrera, M. del C., Torres García, M., Zhang, Q. A., & Álvarez Mateos, P. (2019). Determination of the acidity of waste cooking oils by near infrared spectroscopy. *Processes, 7*(5), 1–7. https://doi.org/10.3390/pr7050304

Giovenzana, V., Beghi, R., Civelli, R., & Marai, S. (2015). Postharvest characterization of olive oil fruits texture by NIR and. *Vis/NIR spectroscopy, 44*(2004), 61–66. https://doi.org/10.3303/CET1544011

Hartman, L., & Lago, R. C. A. (1973). Further observations concerning effects of unsaponifiable constituents on the properties of coffee seed oil. *Journal of the American Oil Chemists' Society, 50*(3), 99–100. https://doi.org/10.1007/BF02671111

Heise, H. M., Damm, U., Lampen, P., Davies, A. N., & Mcintyre, P. S. (2005). Spectral variable selection for partial least squares calibration applied to authentication and quantification of extra virgin olive oils using Fourier transform Raman spectroscopy. *Applied Spectroscopy, 59*(10), 1286–1294. https://doi.org/10.1366/000370205774430927

He, Y., Jiang, H., & Chen, Q. (2020). High-precision identification of the actual storage periods of edible oil by FT-NIR spectroscopy combined with chemometric methods. *Analytical Methods, 12*(29), 3722–3728. https://doi.org/10.1039/d0ay00779j

Jiménez-Carvelo, A. M., Osorio, M. T., Koidis, A., González-Casado, A., & Cuadros-Rodríguez, L. (2017). Chemometric classification and quantification of olive oil in blends with any edible vegetable oils using FTIR-ATR and Raman spectroscopy. *Lebensmittel-Wissenschaft und -Technologie- Food Science and Technology, 86*, 174–184. https://doi.org/10.1016/j.lwt.2017.07.050

Jinkui, M., Han, Z., Tomohiro, T., Yelian, M., & C, J. Y. (2014). Rapid determination of degradation of frying oil using near-infrared spectroscopy. *Food Science and Technology Research, 20*(2), 217–223. https://doi.org/10.3136/fstr.20.217

Kalivas, J. H., & Palmer, J. (2014). Characterizing multivariate calibration tradeoffs (bias, variance, selectivity, and sensitivity) to select model tuning parameters. *Journal of Chemometrics, 28*(5), 347–357. https://doi.org/10.1002/cem.2555

Kaufmann, K. C., Favero, F. D. F., de Vasconcelos, M. A. M., Godoy, H. T., Sampaio, K. A., & Barbin, D. F. (2019). Portable NIR spectrometer for prediction of palm oil acidity. *Journal of Food Science, 84*(3). https://doi.org/10.1111/1750-3841.14467

Kleiman, R. (1990). Chemistry of new industrial oilseed crops. p. 196-203. In J. Janick, & J. E. Simon (Eds.), *Advances in new crops*. Portland, OR: Timber Press.

Kuligowski, J., Carrión, D., Quintás, G., Garrigues, S., & Guardia, M. De (2012). Direct determination of polymerised triacylglycerides in deep-frying vegetable oil by near infrared spectroscopy using Partial Least Squares regression. *Food Chemistry, 131*(1), 353–359. https://doi.org/10.1016/j.foodchem.2011.07.139

Kuriakose, S., Thankappan, X., Joe, H., & Venkataraman, V. (2010). Detection and quantification of adulteration in sandalwood oil through near infrared spectroscopy. *Analyst, 135*(10), 2676–2681. https://doi.org/10.1039/C0AN00261E

Laroussi-mezghani, S., Vanloot, P., Molinet, J., Dupuy, N., Hammami, M., & Grati-kamoun, N. (2015). Authentication of Tunisian virgin olive oils by chemometric analysis of fatty acid compositions and NIR spectra . Comparison with Maghrebian and French virgin olive oils. *Food Chemistry, 173*, 122–132. https://doi.org/10.1016/j.foodchem.2014.10.002

Li, S., Zhu, X., Zhang, J., Li, G., Su, X., & Shan, Y. (2012). *Authentication of pure camellia oil by using near infrared spectroscopy and pattern recognition techniques, 77*(4), 374–380. https://doi.org/10.1111/j.1750-3841.2012.02622.x

Magwaza, L. S., Opara, U. L., Nieuwoudt, H., Cronje, P. J. R., Saeys, W., & Nicolaï, B. (2012). NIR spectroscopy applications for internal and external quality analysis of citrus fruit-A review. *Food and Bioprocess Technology, 5*(2), 425–444. https://doi.org/10.1007/s11947-011-0697-1

Man, Y. B. C., & Moh, M. H. (1998). Determination of free fatty acids in palm oil by near-infrared reflectance spectroscopy. *Journal of the American Oil Chemists' Society, 75*(5), 557–562. https://doi.org/10.1007/s11746-998-0065-0

Mas, C., Rubio, L., Valverde-Som, L., Sarabia, L. A., & Ortiz, M. C. (2020). Impact of the pretreatment of ATR-FTIR signals on the figures of merit when PLS is used. *Chemometrics and Intelligent Laboratory Systems, 201*, 104006. https://doi.org/10.1016/j.chemolab.2020.104006

de Oliveira Moreira, A. C., de Lira Machado, A. H., de Almeida, F. V., & Braga, J. W. B. (2018). Rapid purity determination of copaiba oils by a portable NIR spectrometer and PLSR. *Food Analytical Methods, 11*(7), 1867–1877. https://doi.org/10.1007/s12161-017-1079-8

Rohman, A., Windarsih, A., Riyanto, S., Sudjadi Ahmad, S. A. S., Rosman, A. S., & Yusoff, F. M. (2016). Fourier transform infrared spectroscopy combined with multivariate calibrations for the authentication of avocado oil. *International Journal of Food Properties, 19*(3), 680–687. https://doi.org/10.1080/10942912.2015.1039029

Sahib, N. G., Anwar, F., Gilani, A.-H. H., Hamid, A. A., Saari, N., & Alkharfy, K. M. (2013). Coriander (coriandrum sativum L.): A potential source of high-value components for functional foods and nutraceuticals- a review. *Phytotherapy Research, 27*(10), 1439–1456. https://doi.org/10.1002/ptr.4897

Sampaio, K. A., Ayala, J. V., Van Hoed, V., Monteiro, S., Ceriani, R., Verhé, R., et al. (2017). Impact of crude oil quality on the refining conditions and composition of nutraceuticals in refined palm oil. *Journal of Food Science, 82*(8), 1842–1850. https://doi.org/10.1111/1750-3841.13805

Santos, F. D., Santos, L. P., Cunha, P. H. P., Borghi, F. T., Romão, W., de Castro, E. V. R., et al. (2021). Discrimination of oils and fuels using a portable NIR spectrometer. *Fuel, 283*, 118854. https://doi.org/10.1016/j.fuel.2020.118854

Sinelli, N., Cerretani, L., Egidio, V. Di, Bendini, A., & Casiraghi, E. (2010). Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity. *Food Research International, 43*(1), 369–375. https://doi.org/10.1016/j.foodres.2009.10.008

Teye, E., Elliott, C., Sam-Amoah, L. K., & Mingle, C. (2019). Rapid and nondestructive fraud detection of palm oil adulteration with Sudan dyes using portable NIR spectroscopic techniques. *Food Additives & Contaminants Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 36*(11), 1589–1596. https://doi.org/10.1080/19440049.2019.1658905

Uitterhaegen, E., Sampaio, K. A., Delbeke, E. I. P., De Greyt, W., Cerny, M., Evon, P., et al. (2016). Characterization of French coriander oil as source of petroselinic acid. *Molecules, 21*(9), 1–13. https://doi.org/10.3390/molecules21091202

Van Hoed, V., Sampaio, K. A., Felkner, B., Bavec, F., Scippo, M. L., Brose, F., et al. (2017). Tocopherols and polyphenols in pumpkin seed oil are moderately affected by industrially relevant roasting conditions. *European Journal of Lipid Science and Technology, 119*(12), 1–9. https://doi.org/10.1002/ejlt.201700110

Wang, P., Sun, J., Zhang, T., & Liu, W. (2016). Vibrational spectroscopic approaches for the quality evaluation and authentication of virgin olive oil. *Applied Spectroscopy Reviews, 51*(10), 763–790. https://doi.org/10.1080/05704928.2016.1176034

Yang, H., Irudayaraj, J., & Paradkar, M. M. (2005). Discriminant analysis of edible oils and fats by FTIR, FT-NIR and FT-Raman spectroscopy. *Food Chemistry, 93*(1), 25–32. https://doi.org/10.1016/j.foodchem.2004.08.039

Zeng, L. L., Song, Z. Q., Zheng, X., Tu, B., Yin, C., He, D. P., et al. (2014). NIR Spectroscopy-based qualitative and quantitative detection of adulteration of peanut oil. *Applied Mechanics and Materials, 687–691*, 795–801. https://doi.org/10.4028/www.scientific.net/AMM.687-691.795.

Zhang, L., Huang, X., Li, P., Na, W., Jiang, J., Mao, J., et al. (2017). Multivariate adulteration detection for sesame oil. *Chemometrics and Intelligent Laboratory Systems, 161*, 147–150. https://doi.org/10.1016/j.chemolab.2016.11.009

Zhou, Y., Liu, T., Li, J., & Chen, Z. (2015). Rapid identification of edible oil and swill-cooked dirty oil by using near-infrared spectroscopy and sparse representation classification. *Analytical Methods, 7*(6), 2367–2372. https://doi.org/10.1039/c4ay02900c