Food Analytical Methods (2024) 17:332–347
https://doi.org/10.1007/s12161-023-02568-4

RESEARCH

Comparison of Spectroscopic Techniques Using the Adulteration of Pumpkin Seed Oil as Example

Carolin Lörchner[1,2] · Carsten Fauhl-Hassek[1] · Marcus A. Glomb[2] · Vincent Baeten[3] · Juan A. Fernández Pierna[3] · Susanne Esslinger[1]

Received: 10 May 2023 / Accepted: 18 December 2023 / Published online: 5 January 2024

Abstract
The aim of the present study was to compare different spectroscopic techniques using the example of adulteration of pumpkin seed oil with rapeseed oil in combination with a multivariate regression method. A total of 124 pure seed oils and 96 adulterated samples (adulteration levels from 0.5 to 90.0% w/w) were analyzed using mid infrared, Raman, and ^{1}H-nuclear magnetic resonance spectroscopy. To build quantification models, partial least squares regression (PLS-R) was used. The regression performance parameters, latent variables, and the detection limits (in terms of root mean square error of PLS prediction) calculated when applying the different spectroscopic approaches were compared. For the studied example (pumpkin seed oil adulterated with refined rapeseed oil), the lowest detection limit (3.4% w/w) was obtained for ^{1}H-nuclear magnetic resonance spectroscopy. For the mid infrared and Raman spectroscopy, detection limits of 4.8% w/w and 9.2% w/w, respectively, were obtained, which might be used as screening methods.

Keywords Spectroscopy · Chemometrics · PLS-R · Edible oil · Adulteration

Introduction

The Food Fraud Network's report, published by the European (EU) Commission, annually indicates the most frequently listed product categories in the Administrative Assistance and Cooperation (AAC) system (European Commission 2023). This system collects requests from EU member states to share information about non-compliances and potential deliberate violations of EU legislation in order to tackle food fraud. In 2022, numerous requests were attributed to the product category "fats and oils" (5th place), with the highest number of non-compliance notifications due to "mislabeling" (European Commission 2023), which could be an indication of fraudulent practices.

Cold-pressed pumpkin seed oil is a variety of oil, which is highly susceptible to adulteration with other edible oils. Compared to other seed oils (such as refined rapeseed oil), this is a high-priced oil due to the costly cultivation of the crop, the manual harvesting of the seeds, the extraction of the oil (Wenzl et al. 2002), and nutritional benefits (Šamec et al. 2022). In contrast, refined rapeseed oil is produced in a large-scale industrial process (including mechanical harvesting of the seeds), which is associated with a high oil yield, making it an inexpensive product. Furthermore, pumpkin seed oil has specific organoleptic properties (strong inherent taste as well as odor and a dark green color), which may complicate identification of neutral edible oils as adulterants. For instance, due to the refinement process, the color of a refined rapeseed oil varies from saturated light yellow to light yellow and the taste is almost neutral. Thus, the color is easily concealed when mixed with other edible oils. Since such adulterations can hardly be determined through organoleptic investigation, robust, fast, and reliable analytical methods are required for the detection and prevention of instances of food fraud (Regulation (EU) 2017/625 2017).

✉ Susanne Esslinger
susanne.esslinger@bfr.bund.de

[1] German Federal Institute for Risk Assessment, Max-Dohrn-Str. 8-10, 10589 Berlin, Germany

[2] University of Halle-Wittenberg, Kurt-Mothes-Str. 2, 06120 Halle (Saale), Germany

[3] Quality and Authentication of Products Unit (QAP Unit), Knowledge and Valorization of Agricultural Products Department, Walloon Agricultural Research Centre, Chaussée de Namur, 24, 5030 Gembloux, Belgium

With respect to the detection of adulteration in the edible oil domain, several chromatographic approaches are well established, such as gas chromatography (GC) using a flame ionization detector (FID) (Official Method 2009/7th) or high-performance liquid chromatography (HPLC) in combination with evaporative light scattering (HPLC-ELSD) detector. These techniques are applied in particular to characterize the triacylglycerol and the fatty acid profiles (Salghi et al. 2014). If different types of olive oils cannot be distinguished on the basis of their fatty acid and triacylglyceride profile, the content of minor components (e.g., sterols) as identifying features can be determined by coupling HPLC and GC (HPLC-GC) (Grob et al. 1990). Such approaches are usually the basis of targeted methods, i.e., they are specific to certain molecules or profiles. One advantage of these targeted methods involves the ability to reliably identify and calculate low limits of detection (LOD). For instance, Wenzl et al. (2002) developed a method to discover adulteration of pumpkin seed oil using GC-FID, based on the quantification of the β-sitosterol content in pumpkin seed oil (Wenzl et al. 2002). The authors determined a LOD of 3.3 mg/L and showed that the addition of 1% by weight of corn oil to genuine pumpkin seed oil increased the β-sitosterol content by about 35%. However, a major disadvantage is that most of these chromatography-based techniques involve time-consuming sample preparation, including steps such as acylation, saponification, and various derivatizations. Therefore, rapid measurement and evaluation of the sample is not possible, which hinders detecting food fraud in an early stage. In addition, hazardous substances such as *n*-hexane, methanol, or different ethers are required during these sample preparations, which are hepatotoxic or even toxic to the central nervous system and, consequently, require experienced and trained personnel. Moreover, the analytical approaches described focus solely on the investigation of targeted, i.e., known substances or compositions (based for instance on the fatty acid profile), whereas unknown mixtures, whose presence or absence may provide more specific information, might be overlooked.

To circumvent these drawbacks, non-targeted analysis has been used more frequently in recent years to authenticate food products. This analytical approach uses, depending on the analytical method used, the entire spectrum/chromatogram of a specific food matrix in combination with chemometric evaluation for a comprehensive sample characterization. Accordingly, in order to capture the so-called chemical fingerprint and to avoid elimination of typical compounds of a sample, non-selective sample preparation is required (Esslinger et al. 2014; McGrath et al. 2018). Therefore, in contrast to targeted chromatographic methods, extensive sample preparation procedures such as esterification, transesterification, or fat extraction are not necessary (Castejón et al. 2014). In addition to being more time efficient, this procedure often provides the advantage of a reduced use of hazardous chemicals.

Thus, several studies are available focusing on non-targeted screening tools using mid infrared (MIR) (Fernández Pierna et al. 2016; Javidnia et al. 2013; McDowell et al. 2018), Raman (Berghian-Grosan and Magdas 2020, 2021; Jiménez-Carvelo et al. 2017; McDowell et al. 2018), near infrared (NIR) (Baeten et al. 2014), and nuclear magnetic resonance (NMR) (Alonso-Salces et al. 2022; McDowell et al. 2019) spectroscopy in combination with multivariate data analysis.

In recent years, spectroscopy-based approaches have been employed to develop multivariate regression models, such as partial least squares-regression (PLS-R) analysis or orthogonal projection on latent structures (OPLS) regression, which can be applied to detect and quantify specific adulterants in edible oil (Alonso-Salces et al. 2022; Balbino et al. 2022; Haughey et al. 2015; Jiménez-Carvelo et al. 2017; McDowell et al. 2019; McDowell et al. 2018; Pfister et al. 2018; Rohman et al. 2014). For this purpose, the regression models are trained with "adulterated" samples containing known concentrations of different adulterants.

Rohman et al. (2014) were able to detect adulterations of canola oil in extra virgin olive oil at concentrations of 1% *v/v* using MIR spectroscopy and a PLS-R analysis (Rohman et al. 2014). In contrast, Christopoulou et al. (2004) were able to detect only 5% of canola oil in olive oil using commonly applied targeted, chromatographic methods in combination with a considerably more time-consuming sample preparation (Christopoulou et al. 2004). In another study using near-infrared spectroscopy (NIR), colorimetry, and GC-MS, in combination with OPLS regression, Balbino et al. (2022) investigated the adulteration of pumpkin seed oil with refined sunflower oil (Balbino et al. 2022). The root mean square error of estimation (RMSEE) and of cross validation (RMSECV) values obtained ranged from 2.298 to 6.668 based on different sterol contents, but the estimated LODs were not calculated. The detection of mineral oil in sunflower oil was studied by Pfister et al. (2018) using NIR and MIR spectroscopy. The LOD of 0.12% w/w for NIR and 0.16% w/w for MIR were determined based on the calculation of three times the standard deviation of the predicted mineral oil content of the non-spiked samples (Haughey et al. 2015; Pfister et al. 2018). This example illustrates that very low LOD can be obtained using spectroscopy-based and regression models.

The adulteration of cold-pressed rapeseed oil with refined sunflower and rapeseed oils using MIR and Raman spectroscopy in combination with multivariate regression analysis (PLS-R) was described by McDowell et al. (2018) (McDowell et al. 2018). The authors obtained minimum detection levels (based on 2 × RMSE of prediction (RMSEP)) of 15% (Raman) and 9% (MIR) for adulteration

with sunflower oil. In a further study, McDowell et al. (2019) compared PLS-R results from low and high-field ^{1}H-NMR spectroscopy for the detection of adulterated cold-pressed rapeseed oil with other refined edible oils (McDowell et al. 2019). Based on PLS-R results, minimum detection levels (MDLs) for the adulteration with refined sunflower oil of 8% (400 MHz low field) and 12% (60 MHz) were determined. As a result, the adulteration of cold-pressed rapeseed oil with refined sunflower oil indicated that ^{1}H-NMR spectroscopy (400 MHz) yielded the lowest minimum detection level.

Alonso-Salces et al. (2022) determined RMSEP values between 0.32 and 3.4 (% vegetable oil) for adulteration of olive oil with different oil varieties based on ^{1}H-NMR spectroscopy (500 MHz) and statistical data analysis (Alonso-Salces et al. 2022). These publications highlight strong differences, not only regarding the approach to data evaluation, such as model development and optimization, but also in the assessment of the developed models.

Therefore, the aim of the presented study was to evaluate and compare the potential of three spectroscopic methods (MIR, Raman, and ^{1}H-NMR spectroscopy) as screening tools using the results of multivariate, quantitative regression analysis to investigate the adulteration of cold-pressed pumpkin seed oil with refined rapeseed oil.

Materials and Methods

Sample Collection and Preparation

For this study, 44 rapeseed and 80 pumpkin seed oil samples, purchased from the German retail market from 2017 to 2019, were randomly analyzed within each seed edible oil group. Variability among each seed oil group was covered as best as possible by purchasing oils from different manufacturers, batch numbers, production processes, etc. (Table S1 in the Supplementary material).

The following procedure describes the preparation/selection of the different sample types:

(a) Pooling of pumpkin seed oil samples

In order to cover the highest variability of the pumpkin seed oil class in a sample, the 80 investigated pumpkin seed oil samples were randomly divided into four groups of 20 seed oil samples each, resulting in four pool samples. For each pool sample, 3 ± 0.02 g of the respective pumpkin seed oil samples was weighed into a 100 mL amber DURAN® bottle (Schott AG, Mainz, Germany).

(b) Rapeseed oil samples as adulterants

To identify a representative subset of the rapeseed oil class for the adulteration of the pooled pumpkin seed oil samples, principal component analysis (PCA) of the MIR data of the 44 individual rapeseed oil samples was performed. To ensure high variability within the adulterants (rapeseed oil group), four rapeseed oil samples were selected as adulterants based on the first two principal components (furthest apart from each other in the PCA scores plot) (results not shown).

(c) Adulterated samples

Each pumpkin seed oil pool sample (pool 1-pool 4) was spiked with each rapeseed oil sample (rapeseed oil 1-4), respectively. Pool samples 1 and 2 were adulterated at 19 different concentrations ranging from 0.5 to 90% w/w, and pool samples 3 and 4 were adulterated at five different concentrations ranging from 1 to 10% w/w under gravimetric control (Genius ME254S, Sartorius AG, Göttingen, Germany) (Table S2/S3 in the Supplementary material). The samples were homogenized for 3 min with a Vortex Mixer (Grant Instruments, Cambridge, UK).

The aliquots of the pure rapeseed and pumpkin seed oil samples and pool samples as well as the 96 different adulterated oil samples were filled in 1.2 mL cryogenic tubes (neoLab Migge GmbH, Heidelberg, Germany) and were finally stored under completely dark conditions at −18 °C excluding any headspace volume in order to better preserve the oil samples from oxidation.

Prior to spectroscopic analysis, the samples were tempered to 21 °C and homogenized with an overhead shaker (Reax 2, Heidolph Instruments, Schwabach, Germany) for 30 s at 40 rounds per minute.

MIR Spectroscopy

The mid infrared spectra were recorded on a Vertex 70v Fourier transform spectrometer (Bruker Corporation, Ettlingen, Germany), which was equipped with the standard air-cooled source, a single attenuated total reflectance (ATR) diamond crystal, a wideband IR beamsplitter, and a room temperature deuterated lanthanum α-alanine-doped triglycine sulfate (DLaTGS) detector. For each measurement, 1 µL of sample was transferred onto the crystal surface. Spectra for each sample were recorded in triplicate at room temperature in the absorbance mode from 3996 to 550 cm^{-1} with a spectral resolution of 4 cm^{-1} (data spacing of 1.928 cm^{-1}, Happ-Genzel apodization) by accumulating 32 scans. A background spectrum (laboratory air) was recorded immediately before each sample measurement and inspected visually in order to exclude any signals from solvent (from cleaning) or sample residues. The performance (spectral resolution, signal-to-noise-ratio, and wavenumber accuracy) of the spectrometer

was inspected every 2 months using a polystyrene standard. Spectra visualization was performed with the OPUS 6.5 software package (Bruker, Waltham, USA). The triplicate spectra were averaged and the mean spectrum was used for chemometric analysis.

Raman Spectroscopy

The Raman spectra were acquired on a Vertex 70v Fourier transform spectrometer equipped with RAM II module (Bruker Optics, Ettlingen, Germany) with a spectral range of 4000–50 cm^{-1} and a spectral resolution of 4 cm^{-1} by collecting 128 scans. The Raman module was equipped with a Nd:YAG laser source (yttrium aluminum garnet crystal doped with triply ionized neodymium) (1064 nm), a CaF_2 beamsplitter, and a liquid nitrogen cooled germanium diode detector. Each sample was placed in a glass cuvette (10 mm) and then measured in duplicate at room temperature. The performance of the spectrometer was checked every month using a polystyrene and a naphthalene standard. Visualization of spectra was performed using the OPUS 6.5 software package (Bruker Optics, Ettlingen, Germany). The duplicate spectra were averaged and the mean spectrum was used for chemometric analysis.

^{1}H-NMR Spectroscopy

All chemicals used for the ^{1}H-NMR analysis are presented in Table S4 in the Supplementary material. The measurements were performed with a 400 MHz spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany) applying the instrument specifications in Table S5 in the Supplementary material.

For data collection, 140 ± 2 mg of oil was weighed into a sample tube (Safe-Lock tube, 2 mL, Eppendorf AG, Hamburg, Germany). Using a positive displacement pipette (Microman M1000, 100 µL, Gilson Inc., Middleton, USA), 700 µL of deuterated chloroform ($CDCl_3$, with 0.03% tetramethylsilane (TMS) as internal standard) was added to the same 2 mL sample tube. The weight of the sample and the added $CDCl_3$ were recorded on an analytical balance (Genius ME254S, Sartorius AG, Göttingen, Germany). The weight of the $CDCl_3$ was converted to volume (ρ = 1.5 g/cm^3). The total volume of each diluted oil sample was 700 ± 20 µL. The sample tube was sealed and the solution homogenized for 10 s on a Vortex Mixer (Grant Instruments, Cambridge, UK). From this homogenized solution, 600 µL was transferred to a 5 mm NMR tube (507-PP-7, Wilmad-LabGlass, Sigma Aldrich, St. Louis, USA) using a piston-stroke pipette. During the measurement, the pumpkin seed oil and adulterated samples were positioned in the autosampler holder (BACS-120, Bruker BioSpin GmbH, Rheinstetten, Germany) under dark conditions at room temperature to avoid oxidation processes.

All ^{1}H-NMR measurements were performed without rotation in an automated mode. After sample transfer to the magnet by the autosampler and an equilibration time of 5 min for the temperature, the following optimization steps of the NMR parameters (Godelmann et al. 2013) for each sample were carried out: (i) locking, (ii) automatic tuning and matching, (iii) shimming, and (iv) pulse calibration. Sample acquisition and processing were performed within a single experiment. The instrument settings are presented in Table S6 in the Supplementary material. The collected free induction decay (FID) was automatically processed in TopSpin (Bruker BioSpin GmbH, Rheinstetten, Germany), which included the multiplication with an exponential line broadening (LB) function of 0.3 Hz, Fourier transformation, phase and baseline (polynomial) correction. The chemical shift axis of the spectra was referenced to the signal of TMS at δ 0.00 ppm. The obtained spectra were used for multivariate data analysis.

Quality Control Sample

A refined rapeseed oil purchased from the local market (Germany) was used as quality control (QC) sample for all analysis. The QC sample was analyzed on each measurement day using the three selected spectroscopy approaches (at the beginning, the middle, and the end of each batch) and was filled and stored in the same way as described in "Sample Collection and Preparation".

In addition to the QC sample, two NMR tubes (5 mm) containing methanol-d_4 and sucrose standard solution (Bruker GmbH, Rheinstetten, Germany) were analyzed for defined parameters (absolute temperature of methanol-d_4; water suppression: length of 90 °C pulse, half-width of TMS signal, and signal-to-noise ratio) before starting the NMR measurements each working day and the corresponding assessment limits are listed in Table S7 in the Supplementary material. Whenever the half-width of the standard signal of TMS (~ 0.5 Hz, maximum 0.7 Hz at LB = 0) was >0.5 Hz, the measurements were repeated.

Multivariate Data Analysis

The pre-processing (only bucketing) of ^{1}H-NMR spectra was performed using Mnova version 12.0 (Mestrelab Research, S.L., USA). Multivariate data analysis was carried out using PLS toolbox version 7.0.3 (Eigenvector Research, Wenatchee, WA, USA) together with Matlab version 7.11.0584 R2010b (The MathWorks Inc., Natick, MA, USA).

Data Reduction and Optimal Pre-processing

Regions of the averaged spectra of MIR measurements that did not contain relevant spectral information were excluded (4000–3040 cm^{-1} and 1625–1490 cm^{-1} baseline area, 2790–1790 cm^{-1} absorption of diamond crystal). After this reduction process, each MIR spectrum consisted of 678 data points. The same procedure was performed for the Raman spectra, where the regions 4000–3100 cm^{-1}, 2600–1800 cm^{-1}, and 600–50 cm^{-1} were not included in the data analysis, resulting in 3111 data points per spectrum.

The generated Fourier transformed, phase- and baseline-corrected ^{1}H-NMR spectra were further processed. Bucketing was performed within 0.50–10.02 ppm using a bucket width of 0.04 ppm. The regions of the residues of the undeuterated chloroform signal (7.22–7.34 ppm) and without relevant information (9.02–10.02 ppm) were eliminated, leaving 203 buckets for multivariate statistical analysis.

For each spectroscopic method, PLS-R models were built, and the optimal pre-processing was identified based on the following parameters: the root mean square error of calibration (RMSEC), cross validation (RMSECV), and prediction (RMSEP) for external validation, determination coefficient (R^2) of calibration, R^2 of cross validation, R^2 of prediction as well as the number of latent variables (LV). To select the optimal number of LV, the values for RMSEC and RMESCV were plotted against the number of LVs and selecting the first minimum. Based on these parameters, different pre-processing steps and combinations (Table S8–S10 in the Supplementary material) were tested for each method, and thus, the optimal approach determined was applied: (i) MIR spectroscopy - standard normal variate (SNV), first Savitzky-Golay derivative (filter width 15; polynomial 2), Savitzky-Golay smoothing and mean center, (ii) Raman spectroscopy - SNV, first Savitzky-Golay derivative (filter width 15; polynomial 2) and mean center, (iii) ^{1}H-NMR spectroscopy - normalization (α-signals of glycerol 3.9–4.56 ppm) (Fauhl-Hassek et al. 2000) and Pareto scaling (Table S11 in the Supplementary material).

Explorative Data Analysis

Principal component analysis (PCA) was performed to reduce dimensionality by calculating principal components (PCs) (Wold et al. 1987) and to visualize the possible grouping of pure and adulterated samples, applying the singular value decomposition (SVD) algorithm. For the detection of potential outliers, Student's t distribution, Hotelling's T^2 probability distribution in combination with the Q-statistics, was used (Hotelling 1992; Joe Qin 2003).

Model Building, Optimization, and Validation Using PLS

Detection of adulterants was performed using PLS-R (Eriksson et al. 2006). This involves calculating various statistical parameters such as RMSEs and regression coefficients, which are used to assess the quality of the model (Medina et al. 2019; Riedl et al. 2015; Uncu and Ozen 2015). For this study, the RMSEC, RMSECV, and RMSEP as representation of the spectral differences between the predictions of calibration/validation steps (Liu et al. 2017; Riedl et al. 2015) as well as the determination coefficient (R^2) (Uncu and Ozen 2015) were used. Furthermore, the RMSEP of the system challenge was applied to estimate minimum detection level as shown by Downey and Kelly (Downey and Kelly 2004).

To perform a comparison of PLS-R model results, the following procedure was selected (Fig. 1) in accordance to recommendations of Riedl et al. (2015) (Riedl et al. 2015). The whole data set (pure pumpkin seed oil samples; rapeseed oil samples - except the four rapeseed oil samples used for adulteration; pool 1 and pool 3 samples, adulterated with rapeseed oil samples 3 and 4) was divided into training (66%) and test set (34%) using the Kennard-Stone (KS) algorithm (Kennard and Stone 1969) to avoid the bias that results from manual data splitting. The training set was used to build and optimize the PLS-R model. An internal cross validation was applied with ten data splits on the training set using the venetian blind algorithm (Ballabio and Consonni 2013). The test set was used for external validation to validate the PLS-R model.

Finally, the data of pool 2 and pool 4 adulterated with rapeseed oil samples 1 and 2 were included in the optimized model as independent extra test sets (so called system challenge according to Riedl et al. 2015) for prediction.

Results and Discussion

Visualization and Explorative Data Analysis

To identify differences in the composition of the two types of seed oil varieties, the respective spectral characteristics of the oil samples for each method were firstly described and an exploratory data analysis was performed. A detailed explanation of the respective fatty acid composition of the two seed oil varieties has already been described by several research groups (Baeten et al. 1998; Baeten et al. 1996; Berghian-Grosan and Magdas 2020; Castejón et al. 2014; Guillén and Cabo 1999, 2000) and therefore was not the focus of this study.

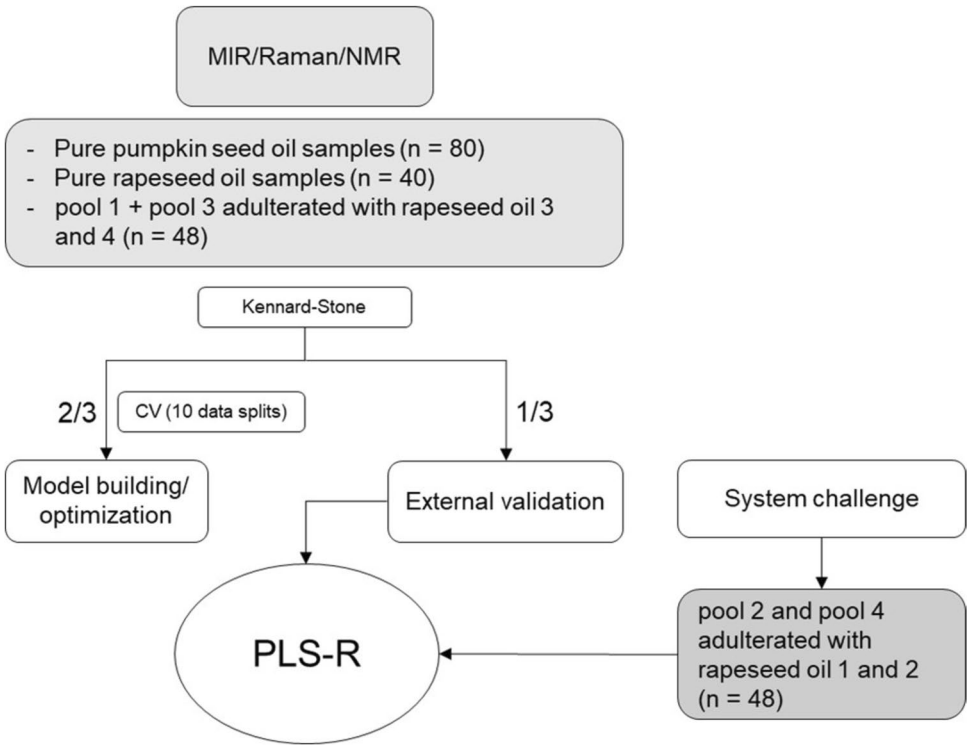

Fig. 1 Scheme of PLS-R model building/optimization, internal and external validation, and system challenge with pure seed oil and adulterated samples. CV: internal cross validation (10 data splits on training set)

MIR Spectroscopy

Figure 2a illustrates the superimposed, raw MIR (after Fourier transformation) spectra of ten cold-pressed pumpkin seed oil samples (m = 3, gray lines) and ten refined rapeseed oil samples (m = 3, black lines). The highlighted areas (Fig. 2a, regions 1 and 2) represent wavenumber regions that exhibit the most noticeable differences in absorbance between the seed oil varieties. The most prominent difference is that the spectra of the pumpkin seed oil samples exhibit higher absorbances in the range between 1050 and 880 cm^{-1} than the rapeseed oil samples. This area is associated with the content of saturated fatty acids (Sherazi et al. 2009). Pumpkin seed oil generally contains 16-19% and rapeseed oil 5-10% saturated fatty acids (Potočnik et al. 2016; Stevenson et al. 2007), which is in good agreement with the superimposed spectra.

To visualize and identify differences and similarities in the data set, a PCA was performed using the pure oil sample set and the training set. Two main groupings can be recognized in Fig. 2b, which are described in particular along PC1 (65% explained variance)—one group builds the data points of the pure rapeseed oil samples (black diamonds) in quadrants 3 and 4. The other cluster consists of the data points of pure pumpkin seed oil samples (gray squares) and adulterated samples (green and yellow triangles), with no clear separation between these to sample sets. According to the loading plot (Fig. 2c), the absorbance differences of the two seed oil varieties along PC1 are described by the range of 1050–880 cm^{-1}, which seems to correlate with the information from the line plot and thus because of the different proportion of saturated fatty acids in seed varieties. Other areas that showed differences in the two oil types in the loading plot of PC1, but barely in the visualization of the spectral data, were the bands around 2924 cm^{-1} and 2854 cm^{-1} as well as around 1746 cm^{-1}. According to literature, these bands are associated with the asymmetric and symmetric stretching vibration of the aliphatic -CH_2 functional group (different fatty acid composition of the seed oil samples and therefore –CH_2 chain lengths) and the functional group of the ester compound C=O (stretching vibration) (Guillén and Cabo 1999, 2000; Vlachos and Arvanitoyannis 2008).

Raman Spectroscopy

The raw Raman (after Fourier transformation) spectral data were not suitable for showing spectral differences (results not shown), because of background noise. Therefore, the pre-processed spectral data are presented in Fig. 3a. The areas with the clearest differences in the spectra are depicted as magnifications.

The region between 3100 and 2800 cm^{-1} reveals differences and reflects characteristic scattered bands of symmetric and antisymmetric (ν(C-H)) vibrations of the terminal chains of methyl (CH_3) and methylene (CH_2) groups of aliphatic molecules (Baeten et al. 1998). The two seed oils' different fatty acid compositions, and thus the different -CH_2 chain lengths in the oil mixtures, are probably responsible

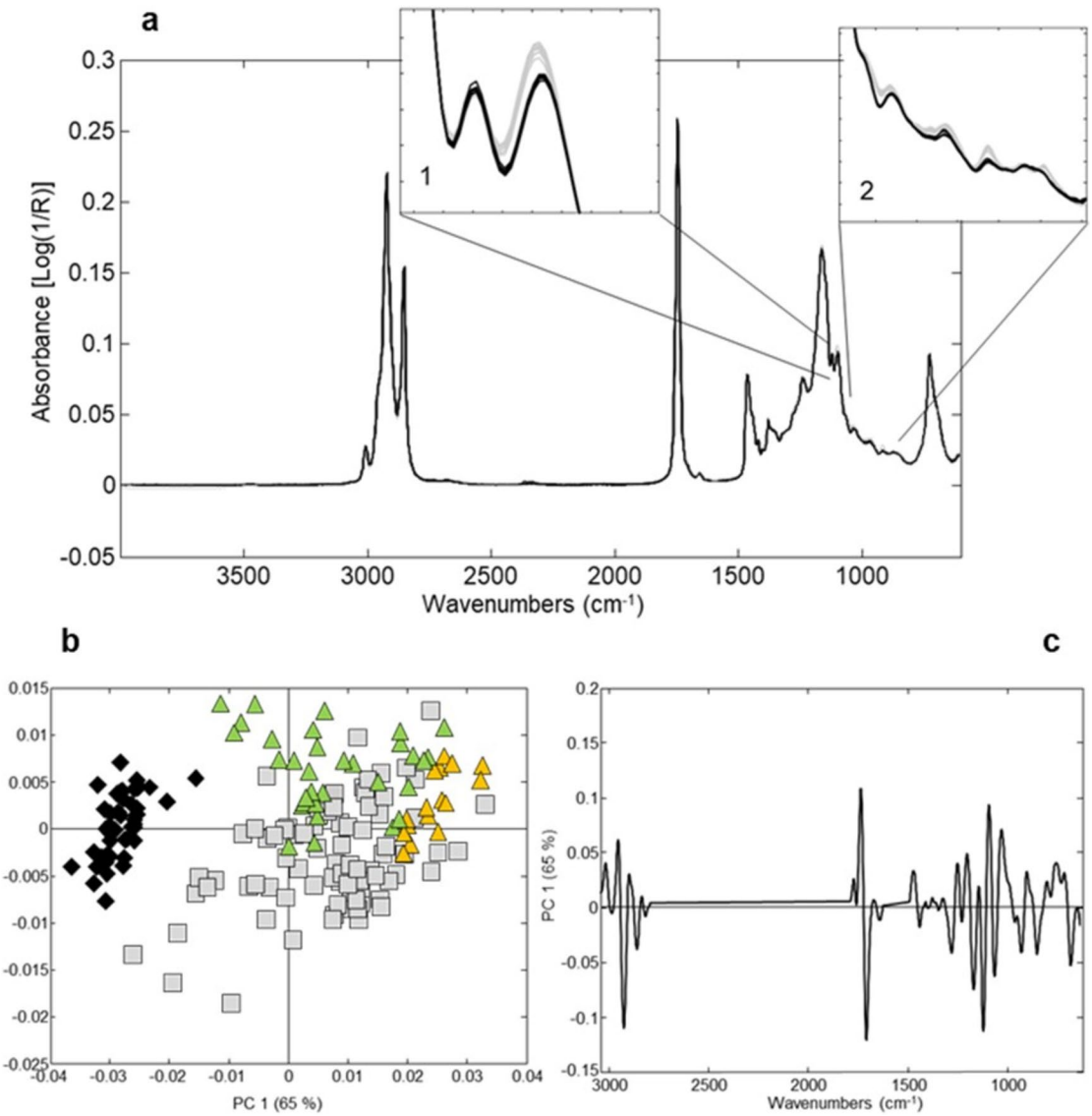

Fig. 2 **a** Line plot of MIR measurements, pumpkin seed oil samples (n = 10, gray lines), rapeseed oil samples (n = 10, black lines)—region 1/2: around 1050–880 cm^{-1}. **b** PCA score plot (PC1/PC2) of pure oil samples (rapeseed oil samples, n = 44, black diamonds; pumpkin seed oil samples, n = 80, light gray squares) and training set (n = 48, yellow triangles < 10% adulteration and green triangles > 10% adulteration) after pre-processing. **c** Loading plot of PC1

for these variations. Vibrations of the olefin groups, i.e., the saturated fatty acids, are characteristic at a Raman shift around 1660 cm^{-1} (region 2 in Fig. 3a) (Baeten et al. 1998). This also correlates with the content of saturated fatty acids in the two varieties of seed oils.

In Fig. 3a in region 3 at 1530 cm^{-1}, the spectra of the pumpkin seed oil samples (gray lines) show significantly higher scattering intensities compared to the spectra of the rapeseed oil samples (black lines). Baeten et al. (2001) demonstrated, depending on the refining process, that pigments (chlorophylls and carotenoids) were removed in edible oils that have a high intensity in this wavelength region (Baeten et al. 2001). Hence, the less extensive a refining process is, the more of these compounds remain in the edible oil and exhibit high intensities in this region of the spectrum. And in the presented study, exclusively refined rapeseed oil samples were investigated.

The apparent differences in the spectra are also described in the scores plot or by the loadings plot of the PCA. The loading plot (Fig. 3c) reveals that the grouping of the two seed oil types along PC1 is especially due to Raman scattering intensity differences in the range of 3100–2800 cm^{-1}, around 1660 cm^{-1} as well as 1050–880 cm^{-1}. Two main groupings can be observed in the PCA score plot (Fig. 3b) of the first components, discriminating along PC1 with 55% of the variance explained. This represents the seed

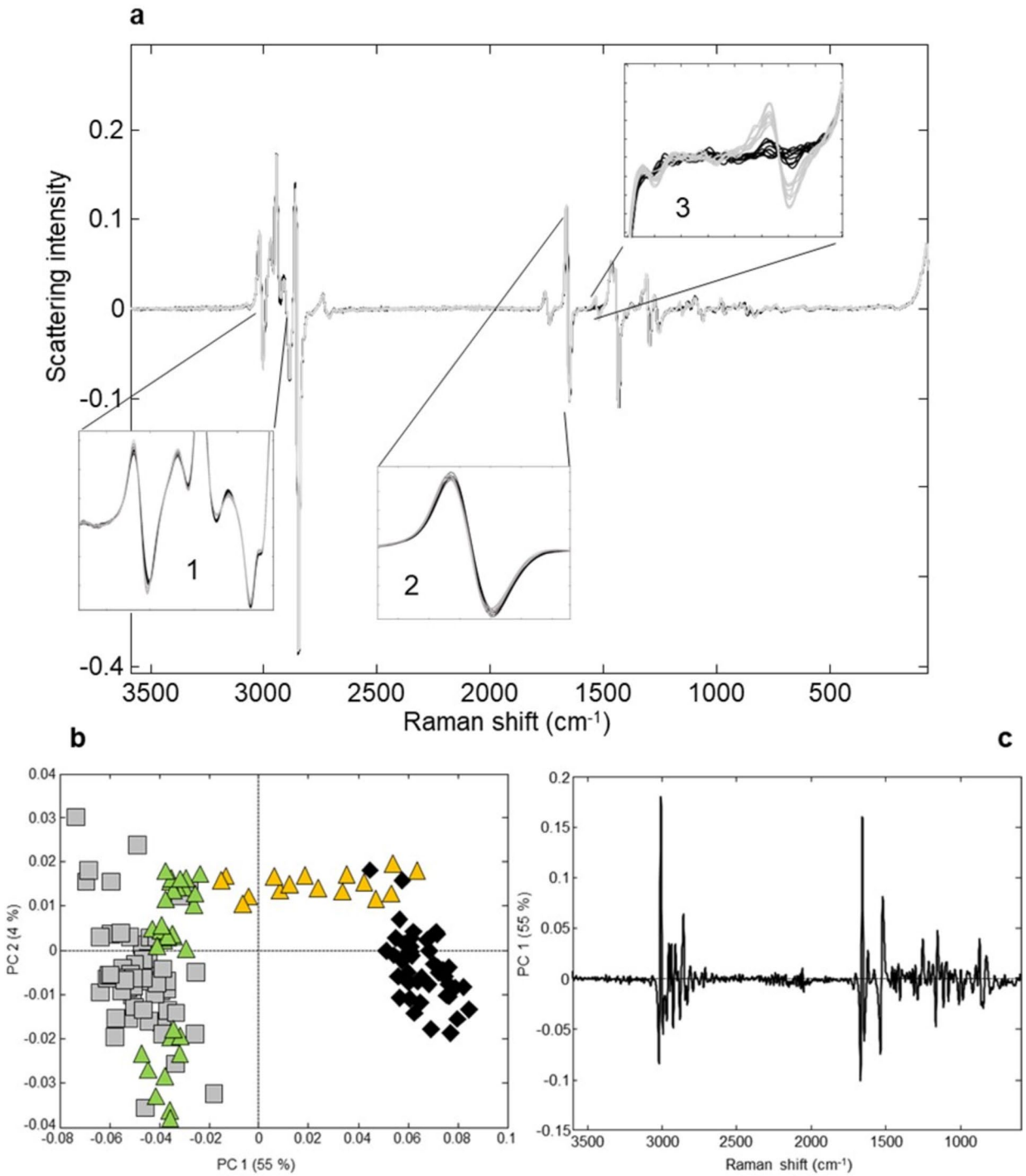

Fig. 3 **a** Line plot of pre-processed Raman spectra, pumpkin seed oil samples (n = 10, gray lines), rapeseed oil samples (n = 10, black lines) - region 1: 3100 cm–2800 cm^{-1}; region 2: 1660 cm^{-1}; region 3: 1530 cm^{-1}. **b** PCA score plot (PC1/PC2) of pure oil samples (rapeseed oil samples. n = 44, black diamonds; pumpkin seed oil samples, n = 80, light gray squares) and training set (n = 48, yellow triangles > 10% adulteration and green triangles < 10% adulteration) after pre-processing. **c** Loading plot of PC1

oil differentiation. The data points of the pure rapeseed oil samples (black diamonds) are located in quadrants 3 and 4, whereas the data points of the pure pumpkin seed oil samples (gray squares) are sited in quadrants 1 and 2. The data points from a degree of adulteration below 10% group along PC2 in the pumpkin seed oil group. The PC2 describes the variation within a seed oil group, which might be explained by different chemical compositions (due to phenotype,

harvest year, etc.) (Stevenson et al. 2007; Szydłowska-Czerniak et al. 2010).

^{1}H-NMR Spectroscopy

Figure 4a inset 1 shows ten overlaid ^{1}H-NMR spectra per pure seed oil variety. Three main spectral differences can be identified and are highlighted in the figure. In the range of 2.74–2.76 ppm, for the rapeseed oil samples (black lines), an overlap of two triplets can be identified and for the pumpkin seed oil group (gray lines) one triplet is exhibited. In this region, the signals are associated with specific bis-allylic protons (=HC-CH_2-CH=), which are also defined as polyunsaturated fatty acids. Depending on the content and composition of the alkyl groups in the acyl group in an oil, the signals lead to different shapes in the spectrum. Oil varieties

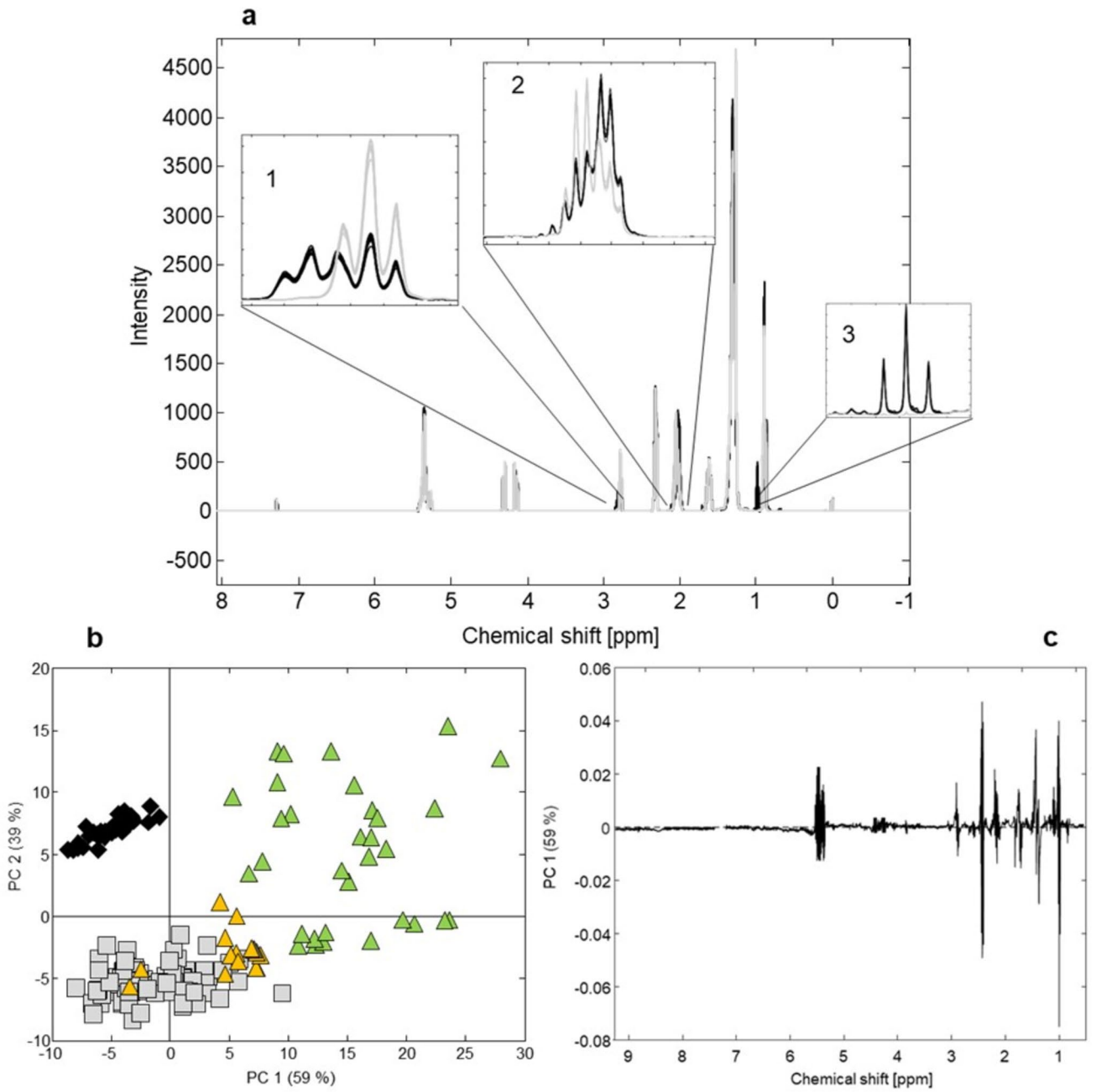

Fig. 4 **a** Line plot of ^{1}H-NMR measurements, pumpkin seed oil samples (n = 10, gray lines), rapeseed oil samples (n = 10, black lines) - region 1: 2.74–2.76 ppm; region 2: 1.99–2.01 ppm; region 3: 0.98 ppm. **b** PCA score plot (PC1/PC2) of pure oil samples (rapeseed oil samples, n = 44, black diamonds; pumpkin seed oil samples, n = 80, light gray squares) and training set (n = 48, yellow triangles < 10% adulteration and green triangles > 10% adulteration) after preprocessing. **c** Loading plot of PC1

with a high content of linoleic and linolenic acids show an overlap of two triplets. If only linoleic acyl groups and no linolenic acyl groups are present in an oil sample, the spectrum exhibits a triplet (Guillén and Ruiz 2003b).

Another difference in the oil varieties can be seen in inset 2 (1.99–2.01 ppm). Here, the shape of the signals caused by the different contents of acyl groups varies in particular. The pattern of signals differs between the two varieties. This spectral region represents also the amount and distribution of unsaturated fatty acids (specific allylic protons) (Guillén and Ruiz 2003b). In particular, the pattern for the pumpkin seed oil group at 1.99–2.01 ppm indicates (Fig. 4a, inset 2) that only small amounts of oil acyl groups are present in the samples. In contrast, the distribution of the peaks in the signal of the rapeseed oil samples implies high contents of oleaginous allyl protons in combination with smaller proportions of linoleic and linolenic acid acyl groups (Guillén and Ruiz 2003b). This correlates with the content of these fatty acids in rapeseed and pumpkin seed oil samples. While the highest content of acyl groups in rapeseed is the oleic acid groups (63%), the linoleic acid groups are dominant in pumpkin seed oils (54%) (Belitz et al. 2009; Guillén and Ruiz 2003b).

Additional variation as displayed in the line plot is present around 0.98 ppm. In this region, all rapeseed oil samples (black lines) show a triplet, whereas the pumpkin seed oil samples (gray lines) exhibit no intensity at all. This region is associated with the terminal methyl protons of *n*-3 polyunsaturated fatty acids (called Ω-3 including, e.g., linolenic fatty acid) (Castejón et al. 2014; Guillén and Ruiz 2003a). According to literature, the content of these fatty acids is 9% for rapeseed oils and only 0.5% for pumpkin seed oils (Guillén and Ruiz 2003b; Rezig et al. 2012; Sakhno 2010; Stevenson et al. 2007) explaining the different behavior in the spectra.

To consider the impact of these regions on the multivariate data analysis, a PCA was calculated. The score plot of the first two principal components is depicted in Fig. 4b. Two main groupings can be observed, differing along PC2 with 39% of the variance explained, representing the variety of seed oil samples. Along PC1 (59% of the explained variance), the variance is described within a seed oil group. The data points with a degree of adulteration of less than 10% (green triangles) scatter along PC1 and PC2 in the first quadrant. The loading plots indicate that the differences within the seed oil cluster (Fig. 4c) mostly result from the signal around 0.98 ppm, while the grouping along PC2 (results not shown) is due to differences in intensity in the range at 2.74–2.76 ppm (region 1) and 1.99–2.01 ppm (region 2).

Comparison of the Performances of the Three Spectroscopy-Based Techniques

Assessment Based on Partial Least Squares-Regression (PLS-R) Results

A common method for quantifying adulteration using spectroscopic techniques is the multivariate PLS-R. The aim of PLS is to establish the functional relationship between an independently measurable variable (e.g., wavenumbers) and a dependent target variable (e.g., levels of adulteration of refined rapeseed oil in a sample of cold-pressed pumpkin seed oil). A PLS-R was performed to estimate the MDL (2 × RMSEP) for quantifying refined rapeseed oil in cold-pressed pumpkin seed oil for each analytical technique using an independent extra test set (system challenge), according to the description of Downey and Kelly (2004). In the present study, besides the calculation of the MDL, additional focus was on the assessment and comparison of other model parameters such as RMSE and R^2 values, generated from the three spectroscopic methods. The PLS-R quantification results are summarized in Table 1.

The RMSE values (a measure of the spectral differences between the predictions of calibration/validation steps) and determination coefficients (R^2_{Pred}) were used to assess the quality of the model (Medina et al. 2019; Riedl et al. 2015; Uncu and Ozen 2015). The lower and more comparable RMSEC, RMSECV, and RMSEP values are, the more reliable is the model (Liu et al. 2017; Riedl et al. 2015). R^2_{Pred} indicates how accurately the model predicts new samples (separate independent test sets for system challenge: pool 2

Table 1 PLS-R results of MIR, Raman, and ^{1}H-NMR spectroscopy to detect refined rapeseed oil in pumpkin seed oil

Spectroscopic method	Optimal number of LV	R^2_{Pred}	RMSEC	RMSECV	RMSEP (system challenge)	Estimated MDL [% w/w]
MIR	4	0.993	2.196	3.122	2.384	4.8
Raman	3	0.983	3.202	4.465	4.591	9.2
^{1}H-NMR	3	0.998	1.148	1.326	1.722	3.4

LV optimal number of latent variables, R^2_{Pred} determination coefficient for prediction, *RMSEC* root mean square error of calibration, *RMSECV* root mean square error of cross validation, *RMSEP* root mean square error of prediction for system challenge, *MDL* estimated minimum detection level = according to Downey and Kelly (2 × RMSEP) (Downey & Kelly, 2004)

and pool 4, adulterated with rapeseed oil samples 1 and 2, see Fig. 1). In general, a R^2_{Pred} value above 0.9 also indicates high prediction ability (Uncu and Ozen 2015).

As reported in Table 1, the RMSEP and (accordingly) MDL values for Raman spectra are higher compared to the results of the MIR and ^{1}H-NMR analysis. Only for adulteration levels higher than 9.2% w/w refined rapeseed oil could be detected in pumpkin seed oil. In addition, the R^2_{Pred} value for Raman spectroscopy indicates a lower predictive ability compared to calculations based on MIR and ^{1}H-NMR spectral data. The lowest estimated MDL is at 3.4% w/w using ^{1}H-NMR spectroscopy. The highest R^2_{Pred} value was also determined for ^{1}H-NMR spectroscopy.

The different PLS-R results obtained could be explained by the different excitation/vibration principles and resulting oscillations/signals of the three molecular spectroscopic techniques. This means that when molecules are excited in different ways, depending on the type of instrument, certain vibrations from molecules can be detected differently, resulting in a difference of the information provided by the spectra. In MIR spectroscopy, molecular vibrations are excited (by, e.g., tungsten source) as a function of vibrating masses and bond strength and molecules are IR active only if they have a high dipole moment (Skoog and Leary 1996). Conjugated double bonds (e.g., linoleic acid), on the other hand, show a low dipole moment and are therefore IR inactive.

However, because Raman spectroscopy is based on a different type of excitation (by monochromatic laser beam), these bonds are Raman active and therefore can be identified in the Raman spectrum (Skoog and Leary 1996). Thus, the two analytical techniques have a complementary detection capability. In addition to the determination of the fatty acid composition, pigments or volatile components can be identified by Raman spectroscopy (see " Raman Spectroscopy") (not by MIR spectroscopy). Nevertheless, in the present study, lower detection levels of adulteration were obtained for Raman spectroscopy. Possibly, this additional information has less impact on the model results.

^{1}H-NMR spectroscopy involves the excitation of hydrogen nuclei. This allows the deduction of the electronic structure of a molecule and its functional groups, rendering this analytical technique very sensitive (Webb 2006). Depending on the shape of a signal (of a singlet, triplet, etc.) (see "^{1}H-NMR Spectroscopy") in the edible oil spectrum, individual alkyl groups in the acyl group can also be identified. In the area of *n*-3-unsaturated fatty acids, the loading plot of PC1 shows strong differences between the two oil types (see "^{1}H-NMR Spectroscopy"), which could affect the model results and consequently would explain the lower detection limits of rapeseed oil in pumpkin seed oil for ^{1}H-NMR spectroscopy. In summary, depending on the variety and production process, different chemical compounds are predominant in the edible oil. Depending on the spectroscopic technique, this results in different bands/signals in the spectrum and possibly in a different weighting of the PLS-R results.

^{1}H-NMR spectroscopy exhibits the best results (regarding PLS-R parameters) for the example reported in this study, due to its excitation principle-based sensitivity for certain molecules, which are relevant to the model, as mentioned above in "Visualization and Explorative Data Analysis."

There are no studies describing the detection of adulteration of pumpkin seed oil with refined rapeseed oil using spectroscopic techniques. However, in order to be able to classify and assess the PLS-R results, the following section measures the results with approaches described in the literature (calculations, assessment parameters, model generation) for comparable authenticity problems and spectroscopic techniques.

Balbino et al. (2022) investigated the adulteration of pumpkin seed oil with refined sunflower oil (Balbino et al. 2022). The authors used OPLS regression in combination with GC-FID and NIR spectroscopy. Differences in sterol contents between the two types of oil were observed (C-H and O-H third overtones and ArC-H and C-H first overtones). Pumpkin seed oils are rich in Δ7-sterols (except Δ7-stigmasterol), whereas refined sunflower oil has a high content of Δ5-sterols (except campestenol). The RMSEP values were not evaluated, but the authors calculated the RMSEE and RMSECV, with values of 6.470 and 6.668 determined for the model based on the β-sitosterol content and 2.298 and 2.792 based on the Δ7,22,25-stigmasterol content. These RMSECV values are within the range of those determined for the present example and indicate that the quality of the generated models is comparable. In the present study, slight differences in sterol composition (e.g., bands around 1440 cm^{-1} and 1350 cm^{-1}) (Baeten et al. 2001) between the cold-pressed pumpkin seed oil and the refined rapeseed oil samples were detected by Raman spectroscopy, but these bands took a subordinate role in the multivariate data analysis. However, by using MIR spectroscopy in our study, no specific bands for the sterol molecules could be identified from the spectra, possibly related to the fact that these molecules were not absorbed or overlapped by other bands. Raman and MIR spectroscopy can be used to obtain sample structural information (e.g., fatty acid composition), while NIR spectroscopy can be used to determine broad bands (Eliaerts et al. 2020). In another study, the adulteration of pumpkin seed oil with sunflower oil was investigated using Raman spectroscopy (portable device) (Becze and Simedru 2020). The areas of the spectral bands from the Raman spectrum were used to quantify the adulteration, and a prediction equation was developed using PLS with four band areas included. These bands were assigned to vibration bands of *cis* (C=C) and *cis* (=C-H) of unsaturated fatty acids as well as scissoring vibrations and twisting vibrations of methylene. The highest band areas in pumpkin seed oil were

assigned to the vibrations of the methylene groups while in sunflower oil the highest band area could be assigned to the vibrations of unsaturated fatty acid group. Also, in the present study, these ranges showed differences between the oil types. Nevertheless, the Raman region between 3100 and 2800 cm^{-1}, which indicated the biggest difference between the pumpkin seed oil group and the rapeseed oil group in the current study, was not included in the multivariate data analysis in the study of Becze et al. (2020) (Becze and Simedru 2020).

Further comparisons cannot be carried out due to the different model design. Pfister et al. (2018) investigated the adulteration of sunflower oil with mineral oil by NIR and MIR (Pfister et al. 2018). The authors determined the LOD values, which are based on the determination of the detection limit using the blank method. This means that in a certain number of samples that were measured the analyte they were looking for was not present (the unspiked natural sunflower oils). Using a previously generated calibration function (using multivariate PLS regression from samples spiked with mineral oil), the content of the blank samples was determined. These calculated contents were plotted in the form of a distribution diagram. From the Gaussian normal distribution fitted to the data, the detection limit was estimated using the 3s limits (two-sided). Using this approach, it was possible to calculate LOD values of 0.12% for NIR and 0.16% for MIR. This example illustrates that very low LOD can be calculated using spectroscopy-based and regression models. The authors did not use an independent test set for the calculation, and the comparison of the methods on the basis of other parameters such as the RMSE values was not the focus. It is less possible to compare the two studies because the calculations of the LODs or MDLs are based on different mathematical calculations. Moreover, the aim in the present study was to include not only the MDLs but also other parameters such as RMSE values from the individual spectroscopic techniques in the assessment of the quality of the models.

McDowell et al. (2018) described the detection of refined sunflower and rapeseed oil in cold-pressed rapeseed oil by MIR and Raman spectroscopy (McDowell et al. 2018). The loading plots of the PCA of the spectra of both techniques show a similar picture as in the present study. Differences between sunflower and rapeseed oil are particularly evident in the fatty acid distribution, additionally in the areas revealing the specific bands for the pigments. Using MIR spectroscopy, better R^2_{Pred} and the MDL values (R^2_{Pred}: 0.99, MDL: 9% w/w) were calculated compared to analysis by Raman spectroscopy (R^2_{Pred}: 0.96, MDL: 15% w/w), which is in the same range to the present study. In a further study, McDowell et al. (2019) calculated the MDLs for ^{1}H-NMR spectroscopy for the identical authenticity questions and sample sets. For this and in contrast to the present study, three specific regions in the spectrum were used (0.52–3 ppm, 3.9–4.56 ppm, and 4.94–5.8 ppm), resulting in 3217 data points. Normalization to the glycerol signal (3.9–4.56 ppm) was subsequently performed (the same for the present study). For the detection of sunflower oil in rapeseed oil, the MDL of 8% (w/w) and an R^2_{Pred} of 0.99 were determined using PLS-R (McDowell et al. 2019). Accordingly, the lowest MDL and highest RMSE values were also obtained for ^{1}H-NMR analysis.

Alonso-Salces et al. (2022) determined the content of adulterant in olive oil samples, adulterated with various vegetable oils (sunflower oil, hazelnut oil, etc.), by ^{1}H-NMR spectroscopy and PLS-R analysis (Alonso-Salces et al. 2022). The authors carried out the same sample preparation as described in the present study. Normalization to the glycerol signal was also performed. However, Alonso-Salces et al. used a bin width of 0.02 ppm (bucketing) and autoscaling or centering as pre-processing steps. Depending on the adulterant and the degree of adulteration, RMSEP values between 0.32 and 3.4 (% vegetable oil) were determined. Furthermore, detection limits between 2 and 5% (depending on the type of adulteration) were calculated (in the range of the present study), but the evaluation of these limits was not described in the publication.

In summary, the values generated in the present study (RMSEP, RMSECV, MDL) are basically comparable or in a similar range as described in the literature. However, the studies described above illustrate that there are very different approaches for the quantitative determination of adulterants in edible oils. These procedures differ in spectrum evaluation (e.g., selection of specific, spectral regions or bucketing), in model building, optimization and validation, type of model (e.g., PLS-R, OPLS), or in the calculation of detection levels (LOD, MDL). A unified approach does not yet exist, which complicates the comparison of results with other techniques or authenticity questions.

Applicability for Laboratories (as Screening Methods)

With regard to the quantification results and the discussion with existing literature, it is not possible to provide a uniform conclusion concerning which spectroscopy-based method is the most suitable, since this depends strongly on the authenticity question to be investigated. Therefore, this study also considers general as well as environmental aspects, which are shown in Table 2.

A prerequisite for the application of the three spectroscopic techniques is the need for trained personnel who are knowledgeable about the techniques (operation, setting of parameters, etc.), to be able, for instance, to detect any solvent residues (from cleaning procedure).

In order to detect food fraud at an early stage and therefore to be able to act quickly before food products reach

Table 2 Comparison of the three investigated spectroscopic techniques taking into account general aspects as well as environmental aspects, particularly with regard to the analyzed authentication issue

	Spectroscopic method		
	MIR spectroscopy	Raman spectroscopy	^{1}H-NMR spectroscopy
General aspects			
Principle	Molecular vibrations	Light inelastic scattering	Spin transition
Determination of content of specific components	✓	✓	✓
Structure elucidation	✓	✓	✓
For investigated example			
Sample preparation**	✗	✗	✓
Use of hazardous chemicals during sample preparation	✗	✗	✓
Use of hazardous chemicals during measurement	✓	✓	✗
Time of analysis	~ 2 min	~ 2 min	~ 20 min
Automatization of analysis (time saving)	✗	✗	✓

✗: no; ✓: yes; **excluding homogenization

the market, fast and simple analytical applications are necessary. Compared to NMR spectroscopy, MIR and Raman spectroscopy offer significant advantages. Sample preparation does not need to be performed. Moreover, a single MIR and Raman spectroscopic measurement takes approximately 2 min. In addition, there are handheld devices that can be taken on site (during the sampling) to perform the measurements onsite. With an adapted cloud-based solution, the evaluation could be performed in minutes. However, one disadvantage (when a large number of samples need to be measured) is that the instruments are not equipped with an autosampler and therefore each sample must be manually placed on the ATR crystal or in the glass cuvette. In NMR spectroscopy, on the other hand, autosamplers are available, which saves a considerable amount of time in the laboratory. Furthermore, NMR spectral data provide information about individual fatty acids (Guillén and Ruiz 2003a, 2003b), while MIR and Raman data do not. It is also possible to identify and determine from an obtained spectrum compounds (phenols, aldehydes) that play a significant role in sensory properties of an oil or indicate a progressing oxidation of the oil.

Nevertheless, with regard to environmental aspects, the advantages for Raman and MIR outweigh those for NMR spectroscopy. Sample preparation for NMR analysis requires in this particular example deuterated chloroform, which is hepatotoxic and likely carcinogenic. The other two techniques preclude the use of hazardous chemicals during the sample preparation. After a measurement by MIR and Raman spectroscopy, only the ATR crystal and the glass cuvette need to be cleaned with ethanol, which is much less hazardous than deuterated chloroform.

For the specific example, based on the PLS-R results (2 × RMSEP) and considering the described environmental aspects, MIR spectroscopy is recommended for the routine and official control to get a fast first inside about the authenticity of the respective edible oil sample. The samples could be measured directly after receipt in the laboratory without further sample preparation within a short time, evaluated and, if adulteration is suspected, analyzed for clarification using accredited, chromatographic methods according to AOAS (Official Method, 2009/7th). These techniques would help food inspectors detain suspect edible oil samples early, before it is available to consumers.

Conclusions

The results of this study demonstrate that it is possible to detect below 10% refined rapeseed oil as an adulterant in pumpkin seed oil using spectroscopic methods and chemometric evaluation. The lowest contents of adulterant could be detected by ^{1}H-NMR analysis (MDL 3.4% w/w). For MIR and Raman spectroscopy, minimum detection limits of 4.8% w/w and 9.2% w/w were obtained. MIR and Raman spectroscopy are more efficient as screening tools than ^{1}H-NMR analysis, since these two methods do not require sample preparation (only homogenization). This also applies for the commonly employed chromatographic techniques, which demand significantly more laboratory effort and therefore are more time-consuming. To challenge the obtained performance of the models, a larger independent extra test set should be analyzed.

There is a large variety of literature describing the quantification of adulteration in edible oils, but approaches differ in terms of the number of samples included for model calculation and development and basis of calculation for MDL as well as the performance parameters used to assess the mathematical models. Therefore, to harmonize spectroscopy-based methods in combination with multivariate data

analysis, the future focus should be on developing a uniform approach regarding model development, evaluation strategies, and calculation of minimum detection limits with the aim to better compare results generated by different spectroscopic techniques.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12161-023-02568-4.

Acknowledgements The authors thank Mr. Quentin Arnould of Walloon Agricultural Research Centre (CRA-W) for support with the MIR and Raman analysis and Mr. Felix Berger from German Federal Institute for Risk Assessment (BfR) for support with ^{1}H-NMR analysis.

Code Availability Not applicable.

Author Contributions Carolin Lörchner: investigation, formal analysis, writing - original draft; Carsten Fauhl-Hassek: project administration, supervision, writing - review and editing; Marcus A. Glomb: supervision, writing - review and editing; Vincent Baeten: supervision, writing - review and editing; Juan A. Fernández Pierna: formal analysis, writing - review and editing; Susanne Esslinger: conceptualization, writing - review and editing.

Funding Open Access funding enabled and organized by Projekt DEAL. The project is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE) under the innovation support programme.

Data Availability Not applicable.

Declarations

Conflict of Interest Carolin Lörchner declares that she has no conflict of interest. Carsten Fauhl-Hassek declares that he has no conflict of interest. Marcus A. Glomb declares that he has no conflict of interest. Vincent Baeten declares that he has no conflict of interest. Juan A. Fernández Pierna declares that he has no conflict of interest. Susanne Esslinger declares that she has no conflict of interest.

References

Alonso-Salces RM, Berrueta LÁ, Quintanilla-Casas B, Vichi S, Tres A, Collado MI, Asensio-Regalado C, Viacava GE, Poliero AA, Valli E, Bendini A, Gallina Toschi T, Martínez-Rivas JM, Moreda W, Gallo B (2022) Stepwise strategy based on 1H-NMR fingerprinting in combination with chemometrics to determine the content of vegetable oils in olive oil mixtures. Food Chem 366:130588. https://doi.org/10.1016/j.foodchem.2021.130588

Baeten V, Dardenne P, Aparicio R (2001) Interpretation of Fourier transform Raman spectra of the unsaponifiable matter in a selection of edible oils. J Agric Food Chem 49(11):5098–5107

Baeten V, Hourant P, Morales MT, Aparicio R (1998) Oil and fat classification by FT-Raman spectroscopy. J Agric Food Chem 46(7):2638–2646. https://doi.org/10.1021/jf9707851

Baeten V, Meurens M, Morales MT, Aparicio R (1996) Detection of virgin olive oil adulteration by Fourier transform Raman spectroscopy. J Agric Food Chem 44(8):2225–2230

Baeten V, Vermeulen P, Fernández Pierna JA, Dardenne P (2014) From targeted to untargeted detection of contaminants and foreign bodies in food and feed using NIR spectroscopy. NewFood 17(3):15–23. http://www.newfoodmagazine.com/14117/supplements/quality-control-supplement-2014/?utm_medium=email&utm;_term=&utm;_c%20Control%20supplement:&utm_campaign=Resend%20of%20NF%20-%20Issue%20#3%202014&utm_source=Email+marketing

Ballabio D, Consonni V (2013) Classification tools in chemistry. Part 1: linear models. Pls-DA. Anal Methods 5(16):3790. https://doi.org/10.1039/C3AY40582F

Balbino S, Vincek D, Trtanj I, Egređija D, Gajdoš-Kljusurić J, Kraljić K, Obranović M, Škevin D (2022) Assessment of pumpkin seed oil adulteration supported by multivariate analysis: comparison of GC-MS, colourimetry and NIR spectroscopy data. Foods 11(6):835. https://doi.org/10.3390/foods11060835

Becze A, Simedru D (2020) Rapid detection of walnut and pumpkin oil adulteration using Raman spectroscopy and partial least square methodology. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 48(3):1426–1438. https://doi.org/10.15835/nbha48312024

Belitz H-D, Grosch W, Schieberle P (2009) Food chemistry. In: *SpringerLink Bücher*, 4th edn. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-540-69934-7

Berghian-Grosan C, Magdas DA (2020) Raman spectroscopy and machine-learning for edible oils evaluation. Talanta 218:121176. https://doi.org/10.1016/j.talanta.2020.121176

Berghian-Grosan C, Magdas DA (2021) Novel insights into the vegetable oils discrimination revealed by Raman spectroscopic studies. J Mol Struct 1246:131211. https://doi.org/10.1016/j.molstruc.2021.131211

Castejón D, Mateos-Aparicio I, Molero MD, Cambero MI, Herrera A (2014) Evaluation and optimization of the analysis of fatty acid types in edible oils by 1H-NMR. Food Anal Methods 7(6):1285–1297. https://doi.org/10.1007/s12161-013-9747-9

Christopoulou E, Lazaraki M, Komaitis M, Kaselimis K (2004) Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. Food Chem 84(3):463–474. https://doi.org/10.1016/S0308-8146(03)00273-5

Downey G, Kelly JD (2004) Detection and quantification of apple adulteration in diluted and sulfited strawberry and raspberry purées using visible and near-infrared spectroscopy. J Agric Food Chem 52(2):204–209. https://doi.org/10.1021/jf035019a

Eliaerts J, Meert N, Dardenne P, Baeten V, Pierna J-AF, van Durme F, Wael K de, Samyn N (2020) Comparison of spectroscopic techniques combined with chemometrics for cocaine powder analysis. J Anal Toxicol 44(8):851–860. https://doi.org/10.1093/jat/bkaa101

Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikström, C, Wold S (2006) *Multi- and megavariate data analysis, part I: basic principles and applications* (2nd ed., Part 1). Umetrics AB

Esslinger S, Riedl J, Fauhl-Hassek C (2014) Potential and limitations of non-targeted fingerprinting for authentication of food in official

control. Food Res Int 60:189–204. https://doi.org/10.1016/j.foodres.2013.10.015
European Commission. (2023) *2022 annual report* [alert and cooperation network]. https://food.ec.europa.eu/safety/acn/reports-and-publications_en#agri-food-fraud-network-ffn
Fauhl-Hassek C, Reniero F, Guillou C (2000) 1H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin. Magn Reson Chem 38(6):436–443. https://doi.org/10.1002/1097-458X(200006)38:6<436::AID-MRC672>3.0.CO;2-X
Fernández Pierna JA, Vincke D, Baeten V, Grelet C, Dehareng F, Dardenne P (2016) Use of a multivariate moving window PCA for the untargeted detection of contaminants in agro-food products, as exemplified by the detection of melamine levels in milk using vibrational spectroscopy. Chemom Intell Lab Syst 152:157–162. https://doi.org/10.1016/j.chemolab.2015.10.016
Godelmann R, Fang F, Humpfer E, Schütz B, Bansbach M, Schäfer H, Spraul M (2013) Targeted and nontargeted wine analysis by 1H NMR spectroscopy combined with multivariate statistical analysis. differentiation of important parameters: grape variety, geographical origin, year of vintage. J Agric Food Chem 61(23):5610–5619. https://doi.org/10.1021/jf400800d
Grob K, Lanfranchi M, Mariani C (1990) Evaluation of olive oils through the fatty alcohols, the sterols and their esters by coupled LC-GC. JAOCS, J Am Oil Chemists' Soc 67(10):626–634. https://doi.org/10.1007/BF02540412
Guillén MD, Cabo N (1999) Usefulness of the frequency data of the fourier transform infrared spectra to evaluate the degree of oxidation of edible oils. J Agric Food Chem 47(2):709–719. https://doi.org/10.1021/jf9808123
Guillén MD, Cabo N (2000) Some of the most significant changes in the Fourier transform infrared spectra of edible oils under oxidative conditions. J Sci Food Agric 80(14):2028–2036. https://doi.org/10.1002/1097-0010(200011)80:14<2028::AID-JSFA713>3.0.CO;2-4
Guillén MD, Ruiz A (2003a) 1h nuclear magnetic resonance as a fast tool for determining the composition of acyl chains in acylglycerol mixtures. Eur J Lipid Sci Technol 105(9):502–507. https://doi.org/10.1002/ejlt.200300799
Guillén MD, Ruiz A (2003b) Edible oils: discrimination by 1 H nuclear magnetic resonance. J Sci Food Agric 83(4):338–346. https://doi.org/10.1002/jsfa.1317
Haughey SA, Galvin-King P, Ho Y-C, Bell SE, Elliott CT (2015) The feasibility of using near infrared and Raman spectroscopic techniques to detect fraudulent adulteration of chili powders with Sudan dye. Food Control 48:75–83. https://doi.org/10.1016/j.foodcont.2014.03.047
Hotelling H (1992) The generalization of Student's ratio. In: *Breakthroughs in statistics*. Springer, pp 54–65. https://doi.org/10.1007/978-1-4612-0919-5_4
Javidnia K, Parish M, Karimi S, Hemmateenejad B (2013) Discrimination of edible oils and fats by combination of multivariate pattern recognition and FT-IR spectroscopy: a comparative study between different modeling methods. Spectrochim Acta A Mol Biomol Spectrosc 104(0):175–181
Jiménez-Carvelo AM, Osorio MT, 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 Sci Technol 86:174–184. https://doi.org/10.1016/j.lwt.2017.07.050
Joe Qin S (2003) Statistical process monitoring: basics and beyond. J Chemom 17(8-9):480–502. https://doi.org/10.1002/cem.800
Kennard RW, Stone LA (1969) sComputer aided design of experiments. Technometrics 11(1):137–148. https://doi.org/10.1080/00401706.1969.10490666
Liu W, Zhang B, Xin Z, Ren D, Yi L (2017) Gc-MS fingerprinting combined with chemometric methods reveals key bioactive components in Acori Tatarinowii Rhizoma. Int J Mol Sci 18(7):1342. https://doi.org/10.3390/ijms18071342
McDowell D, Defernez M, Kemsley EK, Elliott CT, Koidis A (2019) Low vs high field 1h Nmr spectroscopy for the detection of adulteration of cold pressed rapeseed oil with refined oils. Lebensmittel-Wissenschaft Und-Technologie-Food Sci Technol 111:490–499. https://doi.org/10.1016/j.lwt.2019.05.065
McDowell D, Osorio MT, Elliott CT, Koidis A (2018) Detection of refined sunflower and rapeseed oil addition in cold pressed rapeseed oil using mid infrared and raman spectroscopy. Eur J Lipid Sci Technol 120(7):1700472. https://doi.org/10.1002/ejlt.201700472
McGrath TF, Haughey SA, Patterson J, Fauhl-Hassek C, Donarski J, Alewijn M, van Ruth S, Elliott CT (2018) What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed? – Spectroscopy case study. Trends Food Sci Technol 76:38–55. https://doi.org/10.1016/j.tifs.2018.04.001
Medina S, Pereira JA, Silva P, Perestrelo R, Câmara JS (2019) Food fingerprints - a valuable tool to monitor food authenticity and safety. Food Chem 278:144–162. https://doi.org/10.1016/j.foodchem.2018.11.046
Official Method Ce 1-62 Surplus Fatty acid composition by gas chromatography (2009)
Pfister MK-H, Horn B, Riedl J, Esslinger S, Fauhl-Hassek C (2018) Vibrational spectroscopy in practice: detection of mineral oil in sunflower oil with near- and mid-infrared spectroscopy. NIR News 29(3):6–11. https://doi.org/10.1177/0960336018763196
Potočnik T, Ogrinc N, Potočnik D, Košir IJ (2016) Fatty acid composition and δ13C isotopic ratio characterisation of pumpkin seed oil. J Food Compos Anal 53:85–90. https://doi.org/10.1016/j.jfca.2016.09.005
Regulation (EU) No 2017/625 of the European Parliament and of the Council of 15 March 2017 on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products, amending Regulations (EC) No 999/2001, (EC) No 396/2005, (EC) No 1069/2009, (EC) No 1107/2009, (EU) No 1151/2012, (EU) No 652/2014, (EU) 2016/429 and (EU) 2016/2031 of the European Parliament and of the Council, Council Regulations (EC) No 1/2005 and (EC) No 1099/2009 and Council Directives 98/58/EC, 1999/74/EC, 2007/43/EC, 2008/119/EC and 2008/120/EC, and repealing Regulations (EC) No 854/2004 and (EC) No 882/2004 of the European Parliament and of the Council, Council Directives 89/608/EEC, 89/662/EEC, 90/425/EEC, 91/496/EEC, 96/23/EC, 96/93/EC and 97/78/EC and Council Decision 92/438/EEC (Official Controls Regulation) Text with EEA relevance., Official Journal of the European Union (2017) https://www.data.europa.eu/eli/reg/2017/625/oj. Accessed 28 Jan 2022
Rezig L, Chouaibi M, Msaada K, Hamdi S (2012) Chemical composition and profile characterisation of pumpkin (Cucurbita maxima) seed oil. Ind Crop Prod 37(1):82–87. https://doi.org/10.1016/j.indcrop.2011.12.004
Riedl J, Esslinger S, Fauhl-Hassek C (2015) Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. Anal Chim Acta 885:17–32. https://doi.org/10.1016/j.aca.2015.06.003
Rohman A, Che Man YB, Yusof FM (2014) The use of FTIR spectroscopy and chemometrics for rapid authentication of extra virgin olive oil. J Am Oil Chem Soc 91(2):207–213. https://doi.org/10.1007/s11746-013-2370-5
Sakhno LO (2010) Variability in the fatty acid composition of rapeseed oil: classical breeding and biotechnology. Cytol Genet 44(6):389–397. https://doi.org/10.3103/S0095452710060101

Salghi R, Armbruster W, Schwack W (2014) Detection of argan oil adulteration with vegetable oils by high-performance liquid chromatography-evaporative light scattering detection. Food Chem 153:387–392. https://doi.org/10.1016/j.foodchem.2013.12.084

Šamec D, Loizzo MR, Gortzi O, Çankaya İT, Tundis R, Suntar İ, Shirooie S, Zengin G, Devkota HP, Reboredo-Rodríguez P, Hassan STS, Manayi A, Kashani HRK, Nabavi SM (2022) The potential of pumpkin seed oil as a functional food-a comprehensive review of chemical composition, health benefits, and safety. Compr Rev Food Sci Food Saf 21(5):4422–4446. https://doi.org/10.1111/1541-4337.13013

Sherazi STH, Talpur MY, Mahesar SA, Kandhro AA, Arain S (2009) Main fatty acid classes in vegetable oils by SB-ATR-Fourier transform infrared (FTIR) spectroscopy. Talanta 80(2):600–606. https://doi.org/10.1016/j.talanta.2009.07.030

Skoog DA, Leary JJ (1996) Kernresonanzspektroskopie (NMR). In: *Instrumentelle Analytik*. Springer, Berlin, Heidelberg, pp 335–386. https://doi.org/10.1007/978-3-662-07916-4_14

Stevenson DG, Eller FJ, Wang L, Jane J-L, Wang T, Inglett GE (2007) Oil and tocopherol content and composition of pumpkin seed oil in 12 cultivars. J Agric Food Chem 55(10):4005–4013. https://doi.org/10.1021/jf0706979

Szydłowska-Czerniak A, Trokowski K, Karlovits G, Szłyk E (2010) Determination of antioxidant capacity, phenolic acids, and fatty acid composition of rapeseed varieties. J Agric Food Chem 58(13):7502–7509. https://doi.org/10.1021/jf100852x

Uncu O, Ozen B (2015) Prediction of various chemical parameters of olive oils with Fourier transform infrared spectroscopy. Lebensmittel-Wissenschaft Und-Technologie-Food Sci Technol 63(2):978–984. https://doi.org/10.1016/j.lwt.2015.05.002

Vlachos A, Arvanitoyannis IS (2008) A review of rice authenticity/adulteration methods and results. Crit Rev Food Sci Nutr 48(6):553–598

Webb GA (Ed.) (2006) Springer ebook collection / chemistry and materials science 2005-2008. Modern Magnetic Resonance: Part 1: applications in chemistry, biological and marine sciences, Part 2: applications in medical and pharmaceutical sciences, Part 3: applications in materials science and food science (Reprinted with corrections). Springer Science+Business Media B.V

Wenzl T, Prettner E, Schweiger K, Wagner FS (2002) An improved method to discover adulteration of Styrian pumpkin seed oil. J Biochem Biophys Methods, 53(1–3), 193–202. https://doi.org/10.1016/S0165-022X(02)00108-2

Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemom Intell Lab Syst 2(1-3):37–52. https://doi.org/10.1016/0169-7439(87)80084-9