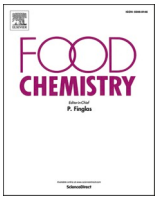
Food Chemistry 430 (2024) 137043

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Assessment of extra virgin olive oil quality by miniaturized near infrared instruments in a rapid and non-destructive procedure

Alejandra Arroyo-Cerezo [a,1], Xueping Yang [b,1], Ana M. Jiménez-Carvelo [a,*], Marina Pellegrino [b,c], Angela Felicita Savino [c], Paolo Berzaghi [b]

[a] Department of Analytical Chemistry, University of Granada, C/ Fuentenueva s/n, 18071 Granada, Spain
[b] Department of Animal Medicine, Production and Health, University of Padua, Via Dell'Università 16, 35020 Legnaro, Italy
[c] Laboratorio di Perugia –ICQRF-MASAF, Via della Madonna Alta 138c/d, 06128 Perugia, Italy

ARTICLE INFO

Keywords:
Extra virgin olive oil quality
Chemical parameter
Near-infrared spectroscopy
Low-cost instrument
Non-invasive method

ABSTRACT

Food fraud in olive oil is a major concern for consumers and authorities due to the health risks and economic impacts. Common frauds include blending with other cheaper non-olive oils, or misleading labelling. The main issue is that legislation and methods presently used in routine laboratories are not always up to date with current fraudulent practices, making detection difficult, so new analytical methods development is required.

This study focuses on developing an affordable and non-destructive analysis method based on NIR spectroscopy and chemometrics for EVOO quality assessment, specifically by monitoring 7 parameters of interest in EVOO measured by official methods and used to develop calibrations through NIR data. For this, two NIR low-cost portable instruments were employed, studied in-depth and compared with a NIR benchtop instrument. Calibration results enabled detection of atypical olive oils and excellent accuracy, especially for palmitic and oleic acid predictions, demonstrating the potential of the instruments.

1. Introduction

The detection of food fraud is undoubtedly one of the most worrying issues for consumers and authorities. From a public point of view, the concern is mainly linked to health and misleading. And it is no less important in the economic field, given that expert estimates of food crime expenditure situate this cost at around $40 million per year, taking into account everything from direct costs (victims and the judicial system) to intangible and market costs (Cox et al., 2020). Food fraud is becoming increasingly sophisticated, making it harder to detect. Unfortunately, the methods used in routine laboratories and the official food control methods outlined in legislation are not always keeping pace with current food fraud practises.

One of the most common sources of food fraud is the olive oil, according to the European Commission's Knowledge Centre for Food Fraud and Quality (KC-FFQ). In fact, olive oil is a target of food fraud worldwide for years (Yan et al., 2020). Several types of fraud have been described that affect not only the economic level, but also public health, such as the intoxication of more than 20,000 people in Spain in 1981 by the illegal sale of rapeseed oil as olive oil. Due to the high economic value of this product in its highest quality category (extra virgin olive oil, EVOO) for its characteristics and attributes, this product is one of the most common sources of food fraud in Europe, and some cases have been detected infringing the legislation that protects and differentiates EVOO from other edible oils. This type of fraud involves mainly the adulteration of EVOO with vegetable oils of lower quality or other vegetable oils, leading to mislabelling as to the commercial category and consumer deception (Lozano-Castellón et al., 2022). Therefore, ensuring the authenticity and quality of this product is essential.

Current olive oil legislation in Europe is based on the recent implementing regulation (Regulation (EU) 2022/2105), and on the repeals of the previous ones (Regulation (EU) 2022/2104). These regulations establish the limits of 8 quality characteristics (acidity, peroxide index, K232, K268 or K270, Delta-K, organoleptic evaluation, and fatty acid ethyl esters) as requirements to commercially classify an olive oil as EVOO, and of the other categories of edible oil produced from olives. Additionally, the composition ranges of 6 fatty acids and 6 sterols, as well as 7 other fatty acids and water content are defined as purity

* Corresponding author.
E-mail address: amariajc@ugr.es (A.M. Jiménez-Carvelo).
[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.foodchem.2023.137043
Received 27 February 2023; Received in revised form 25 July 2023; Accepted 26 July 2023
Available online 28 July 2023

characteristics. All these quality and purity criteria must be analyzed following the official methods and standards of the International Olive Council (IOC). This relies on the need to individually analyze the chemical parameters that determine the quality and purity characteristics by at least 8 different chemical methods using a targeted approach (one method for one analyte) to ensure that the indicated olive oil category is the correct one (García González et al., 2018). Moreover, the parameters to be evaluated and official methods of analysis have remained the same for more than 30 years and still continue in the new regulation.

The official methods often require the use of chemicals and are time-consuming. As a result, both official control laboratories and routine control laboratories in the olive oil industries face limitations in analyzing a large number of samples. This becomes particularly concerning in the case of official control, as it takes days or even weeks to obtain the results relating to the correct labelling of EVOO according to the legislation. The possibility of having a single multiparametric method for rapid screening of edible oils and detection of atypical would greatly benefit the food fraud police and officials routine control laboratories (García Martín, 2022). This would increase the efficiency of controls and optimize the overall food control process for olive oils.

Several studies have been published to date proposing various alternatives to traditional chemical methods for the determination of olive oil quality and authentication. A recent review by Zaroual and colleagues gathers a wide variety of analytical techniques that have been used for this purpose, from the most complex because of the need for a trained analyst, such as gas or liquid chromatography, to the simplest, such as spectroscopic techniques, as well as more innovative techniques such as electronic sensing (Zaroual et al., 2022). The goals of the studies covered in that review could be categorized into: (i) geographical authentication, (ii) variety authentication, (iii) adulteration detection, (iv) classification by olive oil type and (v) prediction of chemical parameters. The least abundant studies in the literature are those related to (iv) and (v), while efforts to develop methods for (i) geographical and (ii) variety authentication and (iii) adulteration detection (González-Pereira et al., 2021) are higher, probably because there are no recognized methods for the identification of these specific frauds (Conte et al., 2020).

Most of these proposed methods have some disadvantages as they are not easily transferable to the industry, especially when we talk about producers in the olive oil industry. There is still a lack of low-cost tools which in turn allow quick and easy screening of olive oils without the need for a trained analyst, especially for those olive oil industries that can implement a portable, fast, cost-effective, non-destructive, and simple method at the production site. In this line, spectroscopic techniques are the clear candidates to offer this type of rapid, non-destructive and low-cost solution. And within these, near-infrared (NIR) spectroscopy is the one that stands out the most, especially for quantitative analysis over compound identification (García Martín, 2022). In this sense, some reviews (Zaroual et al., 2022; García Martín, 2022) were focused on the calibration models developed for the prediction of chemical parameters using NIR data for EVOO quality assessment. For olive oil quality characteristics, namely acidity, peroxides value, K232 and K270, collected NIR spectra allowed obtaining good performance metrics for prediction (Manley & Eberle, 2006; Inarejos-García et al., 2013; Willenberg et al., 2019). Fatty acid contents were also used as reference to develop calibration models with acceptably low prediction errors (Özdemir et al., 2018). But scarce studies can be found on the feasibility of low-cost instruments for this purpose, to make it more affordable to EVOO producers. Garrido-Varo and colleagues (Garrido-Varo et al., 2017) achieved good quantification performance metrics for the most important parameters with the use of low-cost portable instruments, although fatty acids were not included in this study. Fatty acids, especially oleic and linoleic, provide information on the possible adulteration of EVOO with poorer quality vegetable oils.

The chemical information provided by a NIR spectrum is a combination of the chemical constituents found in the system to be analyzed, which are observed in the spectrum in the form of bands or peaks from fundamental vibrations of a chemical bond, which may be overlapped or indicate the presence of several different compounds in the same band. Therefore, it is visually difficult to work with these techniques by identifying individual bands, and the use of supervised chemometric tools is necessary for both qualitative and quantitative analysis, given the volume of data obtained with this technique (Jiménez-Carvelo et al., 2019). The potential offered by chemometric methods using the non-targeted approach in spectroscopic techniques such as NIR is remarkable (Karunathilaka et al., 2016). Through this approach, the full non-specific signal is used as an instrumental fingerprint providing relevant chemical information to characterize the material (Mialon et al., 2023). Advantages of non-targeted over targeted approaches for food authentication purposes have already been highlighted in literature (Sarkar et al., 2022; Hassoun et al., 2023). Although integration into official methods is progressing with publications as ASTM standards and USP guidance (ASTM, 2017; ASTM, 2018; Pharmacopoeia, 2019), further development is still needed.

In this line, the aim of the present work was to develop an affordable, low-cost, and ready-to-use screening method, based on NIR spectroscopy coupled with chemometrics for the rapid and non-invasive control of EVOO from a non-targeted approach, using NIR spectra of olive oils as characteristic signals of each sample as an instrumental fingerprint. The handling of two low-cost portable instruments was also studied in depth, to stablish the optimal acquisition conditions, and compared with spectral data acquired by a benchtop NIR instrument to evaluate the quality of the results.

2. Materials and methods

2.1. Olive oil samples

A total of 195 olive oils (132 from 2021 and 63 from 2022 harvest) samples were analyzed for this study. The origin of olive oil samples was Italian and they were provided by the laboratory of the Central Inspectorate of Quality Protection and Fraud Repression of Agri-food Products (ICQRF), from the Ministry of Agricultural Food and Forestry Policies (Perugia, Italy).

Twenty-five samples corresponding to the 2022 harvest were received later, so that measurements were taken on different days. This set (VAL) was used to assess the applicability of the model to new independent samples. All of the previous samples (n. 170) were used for calibration development.

2.2. Portable NIR instruments

Two low-cost portable instruments were used to collect the NIR spectra of the samples in transmission mode (NIR-M-T1 and NIR-M-T11, Innospectra Corp., Taiwan). Both instruments had an integrated halogen tungsten lamp and the detector consists of a single 1 mm InGaAs element. Dimensions of both devices were similar: 92×76×41 and 96×48×38 mm respectively for NIR-M-T1 (NIT1) and NIR-M-T11 (NIT2) and weighed approximately 100 g each. The main difference between the two instruments was the collected wavelength range: 900–1700 nm for NIT1 and 1350–2150 nm for NIT2.

Both devices are based on the Texas Instrument DLP NIRScan Nano Evaluation Module (DLPNIRNANOEVM, Texas Instruments (TI), Dallas, United States) with a single InGaAs detector and digital micromirror device that can be optimized for number of pixels (equivalent to 128 pixel with no overlapping, or 256 pixels with overlapping) and exposure time in the range of 0.635–60.960 ms. The instruments were connected via USB cable to a laptop computer and controlled using ISC Winform v3.77 software (Innospectra corp. TW). As scanning settings are programmable, preliminary tests were performed looking for the best scanning time and resolution.

2.2.1. NIT1 configuration

The chosen configuration for this instrument consisted of 0.635 ms of exposure time, and 7.03 nm pattern width. Each spectrum was obtained averaging 20 scans with a digital resolution of 300 points in the 950–1650 nm range. The total measurement time was around 7.6 s per spectrum collected. The first and the last 50 nm in the scanning range were not considered because of greater spectral noise.

2.2.2. NIT2 configuration

The chosen configuration for the upper-wave range instrument was divided into two sections to amplify the captured absorbance intensity. First section, consisted of the 1350–1600 nm range with a digital resolution of 104 points, was performed with 7.03 nm pattern width and 0.635 ms of exposure time, while the second section, with a resolution of 176 points in the 1602–2150 nm range, had 15.22 nm pattern width and exposure time of 2.54 ms. The total digital resolution was 280 points. Each spectrum was obtained averaging 10 scans and the total measurement time was around 8.6 s.

2.2.3. NIR spectra acquisition

Each olive oil sample was transferred into two vials (VWR n. 548-0042) with a diameter of 8 mm. Each vial was scanned in duplicate on each instrument within 5 s.

To perform scans, the instruments were first warmed up to a system temperature of approximately 40 °C. Then a reference scan was taken with an empty vial before starting the analysis. This reference scan was repeated every 20–25 min during prolonged scanning sessions. All single scans were exported from the software as a single file in CSV format.

In order to develop the fastest possible method for EVOO control, all samples were measured at room temperature (25 ± 2 °C) and then heated at 50 °C, to be compared in this study with laboratory instruments (Vanstone et al., 2018) and to check the necessity of heating the EVOOs before collecting their NIR spectra, as although recommended, this temperature can cause oxidation in olive oils and be prejudicial to quality analysis.

2.3. Reference data: NIR spectra and chemical parameters

The spectra obtained with NIT1 and NIT2 were compared with spectra obtained by an FT-NIR MPA II spectrometer (Bruker Corporation, Billerica, Massachusetts) (FT-NIR) in the 12,500–4000 cm^{-1} wavelength range with 8 cm^{-1} of resolution (converted to 875–2530 nm range) to compare the performances of the two portable low-cost against a laboratory benchtop NIR instrument. FT-NIR was located at ICQRF laboratory and NIR data were not available for the entire VAL set.

For this study, important chemical parameters for defining EVOO quality were determined to develop the models, namely: acidity, peroxides, K232, K268 and fatty acids (palmitic, palmitoleic, stearic, oleic, linoleic, linolenic and eicosenoic). Reference data were provided by ICQRF laboratory of Perugia (Italy) and all analyses were performed according to the official methods specified in the regulation (Regulation (EU) 2022/2105).

A statistical analysis of wet chemistry results was performed. Moreover, the methods were validated by ICQRF laboratory, and the standard errors of laboratory (SEL) methods had been determined with 10 replicates of the same sample. Maximum R^2 obtainable in the subsequently development of calibration was calculated with the following equation. Since if the wet chemistry data had errors, these would be carried over to the next step when developing prediction models.

$$R^2_{max} = \frac{SD^2 - SEL^2}{SD^2}$$

Where SD: standard deviation of data; SEL: standard error of laboratory.

2.4. Multivariate data analysis

Spectra from all single CSV files were imported and linearly interpolated every 2 nm by an R script using RStudio (RStudio version 2022.2.2.485, PBC, Boston, MA). All spectra were averaged by samples (4 scans), by scanning temperature and by instrument, defining four different sets of data (two instruments at two temperatures).

Replicate scans for each sample within each portable instrument were used to calculate repeatability of the instruments and to compare the quality of the obtained spectra between the two devices (NIT1 and NIT2) using the root mean squared (RMS) (Xue et al., 2014; De la Roza-Delgado et al., 2017) calculated according to the following equation:

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} (y_{im} - y_{ik})^2}{n}}$$

Where y_{im} = absorbance value of scan m of one sample at a wavelength *i*; y_{ik} = absorbance value of scan k of the sample at wavelength *i*; n = number of wavelengths.

Different combinations of spectra pre-processing methods were tested and the following was selected and applied to the data: Savitzky-Golay first derivative (9 filter window and 2nd polynomial order) and standard normal variation (SNV).

Wet chemistry data were evaluated for the detection of outliers, as laboratory reference data may have a great impact on the development of predicting models. For this purpose, a matrix composed of 170 rows (samples) and 11 columns (chemical parameters) was autoscaled to perform a principal component analysis (PCA) and the samples with large Q-residuals were removed.

Partial Least Square (PLS) regression method was used to develop calibration models for all the chemical parameters. The calibration dataset was randomly split into train and test sets in an 80:20 ratio and a 4 groups cross-validation was used to optimize the models. Optimal number of PLS components was selected on basis of the minimum root mean squared error (RMSE) balanced with the minimum possible value of the beta coefficient to avoid overfitting the model (Stotltzfus, 2011). All calibration developments were performed under Python 3.9, using the NumPy package (Harris et al., 2020).

Lastly, the VAL set (see section 2.1) was used to test the predictive capacity of the developed models, comparing prediction and reference values by applying models to samples not used for calibration (García Martín, 2022). For this purpose, it was calculated the standard error of prediction (SEP), Bias, and GH distance (Williams et al., 2017). The latter was calculated as the Mahalanobis distance divided by the number of PLS components (Garrido-Varo et al., 2019).

Note that in this study the calibration set is considered the set used to develop the models, randomly split into training and test sets. VAL set was the set of independent samples because they were sampled and analyzed at a different time.

Spectral repeatability was translated into predictive repeatability. For this purpose, the non-averaged spectra of the VAL set (4 spectra per sample) were introduced into the developed models. An analysis of variance (ANOVA) of the predicted parameters was performed and the repeatability for each data set was calculated as the square root of the residual variance divided by the degrees of freedom.

3. Results and discussion

The spectral range obtained by each of the three instruments (NIT1, NIT2 and FT-NIR) is shown in Fig. 1. Note that FT-NIR spectra were trimmed over 2250 nm, as spectra were saturated and no longer provided useful information, according to García Martín (García Martín, 2022) can be discarded without losing important information in olive oil samples. As expected, with the bench spectrometer (FT-NIR) spectra show sharper peaks due to the higher resolution than portable

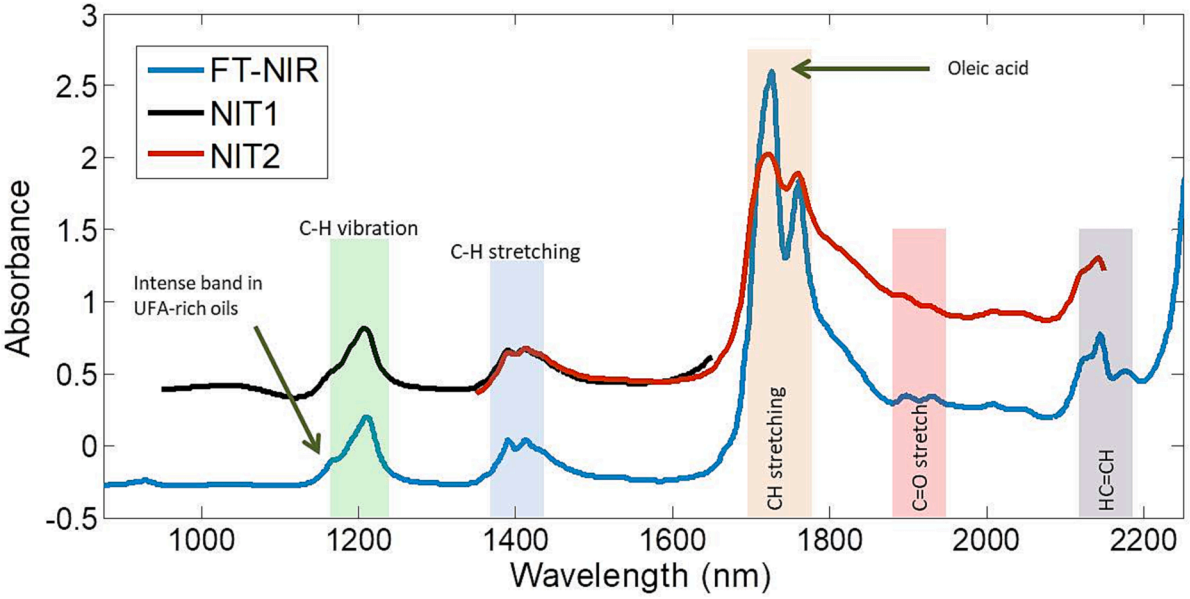


Fig. 1. EVOO NIR spectra obtained with the two portable instruments (NIT1 and NIT2) and bench instrument (FT-NIR).

instruments, NIT1 and NIT2.

The band at 1180–1200 nm corresponds to the C—H vibration of CH_2 and CH_3 groups, but this range is not covered by NIT2 instrument. Spectral band at 1165 nm usually appears in vegetable oils with high unsaturated fatty acid content and is observed in FT-NIR spectrum, but practically unobservable in NIT1, probably due to the lower spectral resolution. Similarly, the double band between 1890 and 1945 nm, corresponding to C═O stretching, is detectable in FT-NIR spectrum but not in NIT2. The highest absorbance peak (1725 nm) characteristic of oleic acid, specifically triolein (García-González et al., 2013), is not covered by the NIT1 spectral range, which will influence the calibration performances with data from this instrument. Finally, the double bands at 1380–1420 and 1700–1770 nm corresponding to C—H stretching are present in the NIT2 and FT-NIR spectral ranges (Özdemir et al., 2018; Borghi et al., 2020). Pre-processed spectra can be seen in Fig. S1 of supplementary material.

3.1. Handling of portable instruments

The spectral repeatability of both portable instruments was calculated by means of the RMS value, the lower the better the repeatability. When raw data were used to calculate it, results were: 7.13×10^{-4}, 1.56×10^{-3}, 2.10×10^{-3}, 2.80×10^{-3} u.A. respectively for olive oils at room temperature and heated using NIT1 and NIT2. Note that this is spectra repeatability, i.e., this was calculated with the duplicate scans measured from one vial per sample. In addition, the repeatability per sample was also calculated, since two vials were measured for each sample in this study, and the results were: 8.54×10^{-3}, 1.02×10^{-2}, 1.04×10^{-2}, 1.67×10^{-2} u.A. These results show the sampling effect, and as expected, spectral repeatability by sample is poorer than repeatability by vials for a factor of 10. It is noteworthy that the repeatability of NIT1 is greater than that of NIT2, as is the case for samples measured at room temperature better than for heated samples. These results would indicate a priori that measurements performed at room temperature with the NIT1 spectrometer were more repeatable than with NIT2 or after heating the samples.

3.2. Wet chemistry data

The detection of outliers through the development of a PCA resulted in 6 samples with high values of Q residuals and Hotelling T^2. However, only three samples with high values of Q-residuals were removed from the dataset. High values of Q-residuals indicate those samples that are not well explained by the model, while high values of Hotelling T^2 correspond to those samples that show deviations. Therefore, in order to have maximum variability when running the model, the three samples with high Hotelling T^2 were not removed. In addition, the variability of these samples, and therefore being detected as outliers, coincides with their chemical values differing from those what would be expected for EVOO. In fact, these results agree with the information provided by ICQRF that, although all three samples were labelled as EVOO, the chemical parameters were very different from those that EVOO should contain. Fig. 2 shows graphically the differences between the values of the chemical parameters analyzed, representing the ranges of the EVOO samples (n = 164, in blue), those suspected non-EVOO (n = 3, in red) and the range allowed for EVOO by legislation for each parameter (in green) (Regulation (EU) 2022/2104). For better visualization, the 11 parameters were divided into two groups: high ranges (peroxides, and palmitic, oleic, and linoleic acids) and low ranges (acidity, K232, K268, and palmitoleic, stearic, linolenic and ecosanoic acids). Note that from now on, authentic EVOO samples will be named *typical EVOO*, and the suspected non-EVOO samples will be called *atypical EVOO*, following the Pharmacopoeia guidelines (Pharmacopoeia, 2019).

Note the difference in the K232 and K268 values, which in the atypical EVOO samples were above the permitted value of 2.50 and 0.22, respectively, for EVOO (Regulation (EU) 2022/2104), this could be caused by a bad oxidation state on these three olive oils. Remarkable was the difference found in oleic and linoleic acid, which were also outside the range allowed by legislation and the expected range for EVOO found in the literature (Özdemir et al., 2018) in the three atypical EVOO samples. Further, looking at these lower oleic and higher linoleic values, it could be concluded that these three samples may had been adulterated with other non-olive vegetable oils, as these values were different from typical EVOO (Aykas et al., 2020; Borghi et al., 2020). In addition, it is worth noting the high values in the peroxide index of some typical EVOO samples, above the permitted value, and this could be due to oxidation processes on those samples.

Table 1 shows the descriptive statistics of the calibration set after outliers removal, and VAL set, including: number of samples (N), mean, standard deviation, minimum and maximum values, the standard error of laboratory (SEL) of the chemical analysis methods, the maximum R^2 value and the Pearson's correlation coefficients between the chemical

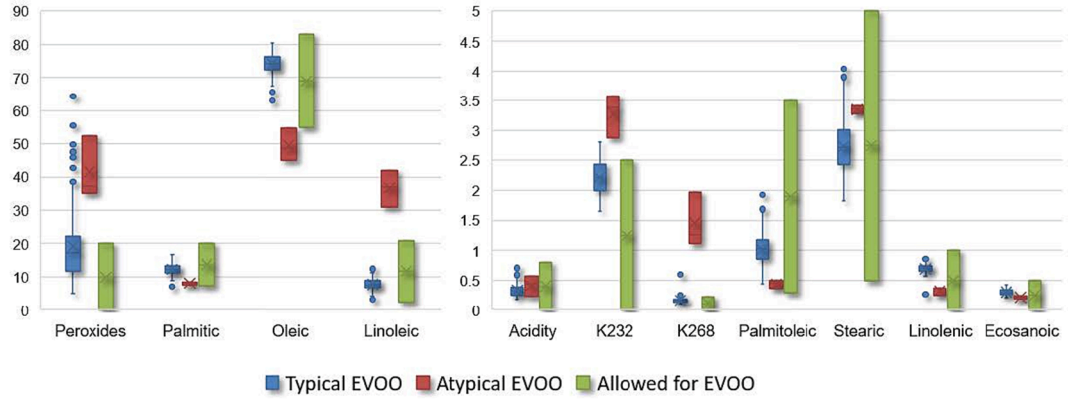


Fig. 2. Wet chemical data of typical EVOO (authentic) samples and atypical EVOO (suspicious) samples with specified ranges allowed by legislation for EVOO. *Note that for better visualization, the 11 parameters were divided into two groups: high ranges on the left and low ranges on the right.*

Table 1
Descriptive statistics of the wet chemistry data of EVOO samples from the calibration (CAL) set and VAL set (N = 25).

	Parameter	Acidity	Peroxides	K232	K268	Palmitic	Palmitoleic	Stearic	Oleic	Linoleic	Linolenic	Eicosenoic
	Units	%	mEq O_2/kg	N/A	N/A	%TFA	%TFA	%TFA	%TFA	%TFA	%TFA	%TFA
CAL	**N**	167	166	167	167	141	141	141	141	141	141	141
	Mean	0.34	19.58	2.23	0.19	12.30	1.01	2.73	73.49	8.40	0.69	0.30
	SD	0.10	11.12	0.31	0.18	1.66	0.28	0.49	5.17	4.59	0.09	0.05
	Min	0.18	5.00	1.65	0.10	7.09	0.36	0.08	45.07	3.31	0.25	0.20
	Max	0.70	64.45	3.57	1.98	16.79	1.96	4.04	80.57	41.90	0.86	0.42
	SEL	0.03	2.4	0.06	0.01	0.48	0.07	0.12	0.98	0.24	0.06	0.06
	$\mathbf{R^2_{max}}$	0.917	0.958	0.962	0.997	0.917	0.938	0.940	0.964	0.997	0.552	−0.648
VAL	**Mean**	0.28	10.16	1.89	0.13	12.83	1.09	2.55	74.06	7.74	0.69	0.27
	SD	0.10	2.93	0.20	0.01	1.28	0.25	0.43	3.13	1.87	0.04	0.03
	Min	0.08	5.50	1.52	0.10	10.09	0.77	2.03	62.65	4.88	0.62	0.22
	Max	0.45	15.20	2.24	0.16	17.00	2.08	3.64	78.74	13.99	0.76	0.32

N: number of samples; SD: standard deviation; Min: minimum; Max: maximum; SEL: standard error of laboratory; N/A: not applicable; TFA: total fatty acids.

parameters. The data corresponding to the measured fatty acids were not available for the whole calibration set of samples, this is the reason for the difference in N value between the parameters.

The maximum R^2 value calculated is quite high and acceptable in most cases (greater than 0.9), especially for K268 and linoleic acid. However, maximum R^2 values for linolenic and eicosenoic acids showed some possible limitations, because reference SELs were large in relation to the variability of the calibration data. Also, it must be noted that in the data corresponding to the VAL set, some ranges were outside the range covered by the calibration data. This was the case of the minimum value of acidity, and maximum values of palmitic and palmitoleic acids. This will be a factor to take into account when evaluating prediction performances by using the VAL set.

3.3. Predictive models

After studying the handling and usability of the data obtained with NIT1 and NIT2 and the wet chemistry reference data, calibrations were developed for 7 parameters with the 5 data sets (NIT1 and NIT2 at room and high temperature and FT-NIR at high temperature, following their protocol for this type of analysis). The VAL set was used to evaluate the predictive capability, except for the calibrations developed with FT-NIR data, since spectral data for all 25 samples were not available. Table 2 shows the results obtained for the calibration and VAL sets. In this table, "_RT" refers to olive oils measured at room temperature and "_50" after heating.

As expected, best calibration performance metrics were obtained for FT-NIR data, as spectral range was larger, and resolution was greater than NIT1 and NIT2. However, RMSE values obtained FT-NIR were in most cases not much lower than the portable instruments, and for palmitic and oleic acids, calibrations with NIT1 at room temperature had slightly lower RMSE than using FT-NIR data. It should be noted that all calibration errors were lower than 1%, with the only exception of the peroxides index. This index is a parameter with large fluctuations (see Table 1) and is not particularly stable over time, as it in fact determines the deterioration of olive oil with time. Even so, the calibration R^2 were high, especially for the data obtained with NIT1 at high temperature.

Data acquired at room temperature with both portable instruments showed better calibration and prediction performances (errors and R^2) for K232 and the four fatty acids calibrations. This agrees with the results from Azizian et al. (Azizian et al., 2007), who found that the increase in edible oil temperature above 40 °C caused classification and quantification errors of the models to be increasingly larger, this could be explained because such temperature could provoke the onset of fatty acids oxidation in olive oils. In addition, better results were also observed with the NIT1 data for fatty acids, both in calibration and in independent set prediction, compared to NIT2. This may be because measurements performed with NIT1 were more repeatable than with NIT2 as discussed in section 3.1 above. In the case of acidity and peroxides, NIT1 data achieved better calibration performances, but prediction performances were better with NIT2, including K232, which could be related to a possible loss of relevant information in the 1652–2150 nm spectral range not covered by NIT1.

Beside the R^2, a useful criterion for estimating the prediction accuracy of a model is the proximity of the SEP to the standard error of laboratory (SEL). The criteria proposed by (Shenk & Westerhaus, 1996) state an excellent accuracy when SEP/SEL value is lower than 1.5; good accuracy for SEP/SEL values < 3; medium accuracy for SEP/SEL values < 4; and low accuracy for SEP/SEL values between 4 and 5. According to this, the predictions performed here for acidity, K232 and fatty acids had good accuracy (excellent for palmitic and oleic, as SEP/SEL was below 2 for NIT1), and low accuracy for peroxides (García Martín, 2022).

Table 2
Calibration (CAL) and prediction (VAL) set metrics of developed PLS models.

Instrument	CAL				VAL				
	Parameter	PLS f	RMSE	R^2_{cal}	SEP	Bias	GH_{av}	R^2_{pred}	Acc
NIT1_RT	Acidity	5	0.07	0.295	0.09	0.01	1.96	0.208	3.0
	Peroxides	4	3.94	0.845	11.88	11.37	6.01	0.289	5.0
	K232	7	0.19	0.369	0.15	0.00	2.62	0.477	2.5
	Palmitic	13	0.47	0.907	0.66	0.19	3.84	0.815	1.4
	Palmitoleic	10	0.15	0.754	0.14	−0.03	3.96	0.664	2.0
	Oleic	9	0.92	0.922	1.46	0.88	2.97	0.856	1.5
	Linoleic	8	0.48	0.923	0.57	−0.17	7.39	0.921	2.4
NIT1_50	Acidity	5	0.06	0.414	0.08	0.02	2.11	0.305	2.7
	Peroxides	8	3.47	0.880	13.00	12.30	3.19	0.321	5.4
	K232	7	0.20	0.304	0.17	0.04	2.57	0.376	2.8
	Palmitic	10	0.92	0.647	0.85	−0.13	4.35	0.636	1.8
	Palmitoleic	10	0.17	0.688	0.16	−0.05	3.74	0.646	2.3
	Oleic	8	1.78	0.704	1.88	−1.20	2.62	0.780	1.9
	Linoleic	8	0.82	0.774	1.25	1.05	3.78	0.898	5.2
NIT2_RT	Acidity	3	0.08	0.133	0.08	−0.03	2.16	0.328	2.7
	Peroxides	7	4.70	0.781	5.96	5.10	2.94	0.551	2.5
	K232	5	0.16	0.557	0.17	−0.10	1.70	0.528	2.8
	Palmitic	12	0.64	0.832	0.96	−0.09	2.47	0.502	2.0
	Palmitoleic	7	0.15	0.743	0.24	0.09	2.23	0.292	3.4
	Oleic	8	0.98	0.911	2.35	−0.89	2.26	0.584	2.4
	Linoleic	8	0.50	0.915	1.29	0.76	2.80	0.724	5.4
NIT2_50	Acidity	3	0.07	0.304	0.09	0.03	1.46	0.298	3.0
	Peroxides	8	3.86	0.852	12.31	11.70	1.88	0.142	5.1
	K232	5	0.19	0.335	0.18	0.04	1.98	0.193	3.0
	Palmitic	11	0.72	0.786	0.85	−0.30	2.66	0.639	1.8
	Palmitoleic	9	0.19	0.616	0.13	0.01	1.71	0.721	1.9
	Oleic	11	0.92	0.922	1.11	−0.16	2.44	0.873	1.1
	Linoleic	8	0.79	0.792	0.65	−0.39	2.46	0.920	2.7
FT-NIR	Acidity	6	0.08	0.463					
	Peroxides	11	2.84	0.895					
	K232	6	0.18	0.603					
	Palmitic	8	0.50	0.927			N/A		
	Palmitoleic	6	0.12	0.856					
	Oleic	7	0.98	0.976					
	Linoleic	9	0.23	0.999					

PLS f: number of PLS factors used for calibration; RMSE: root mean squared error; SEP: standard error of prediction; Acc: accuracy calculated as SEP/SEL; N/A: not available data.

The GH parameter was developed to identify spectra not well represented by the calibration data set. It should be noted that GH threshold in this type of application is usually set at 3, so a value above 3 would indicate an outlier within the developed model (Garrido-Varo et al., 2019). Table 2 shows the average GH values obtained for the 25 samples of the VAL set. Values below 3 were observed for all predictions with NIT2, however, with NIT1 this only occurred for acidity, K232 and oleic acid. These results could be interpreted in two ways. Having no outliers in the NIT2 VAL set, one could conclude that in the VAL set all samples were typical EVOO. On the other hand, with NIT1 some samples in the VAL set having large (>3) GH values, could raise the suspicion of not being EVOO. However, GH distance is not the appropriate parameter to be used for discriminating being EVOO, but it is useful in the maintenance and updating calibrations to find those samples not well represented in the calibration set, in order to remove them or update the calibration with new samples inclusion.

A more suitable and easy way for the detection of atypical EVOO, could be focused on the predicted value for each chemical parameter and not directly on the spectral data as in the GH parameter. The three atypical EVOO samples (section 3.2) scored Hotelling T^2 values higher than 3 in all cases (FT-NIR, NIT1 and NIT2 at both temperatures) when performing a PCA with the predicted chemical data (all calibration and VAL set samples). The performed prediction also showed that these three samples were out of the allowed range for EVOO (see supplementary material, Figs. S2 and S3), especially in K232 and oleic and linoleic acids, although the acidity value was within the allowed range, being one of the official parameters to be analyzed. Therefore, with the developed PCAs, atypical EVOO could be detected when the Hotelling T^2 value is greater than 3. These results show the capability of NIR as a multiparametric tool to be used in olive oil quality control according to J.F. García Marín review (García Martín, 2022) as a rapid and non-invasive screening, through the detection and identification of suspicious olive oils.

It should be noted in the PCA results (Fig. S2) that some samples scored too large Q residuals values, especially sample "529″. This could be explained by the fact that it was the unique unfiltered olive oil, so the physical state of the edible oil should be taken into account, or a representative content of such samples should be included in the model. This did not hold true for the PCA developed using the reference chemical data, but it occurred when using the chemical data predicted from the NIR spectra.

3.4. Repeatability of predictions

The calculated spectral repeatability (see section 3.1) was transferred to predictive repeatability. The aim of this predictive repeatability study was to check whether it is necessary to measure the same sample in duplicate vials and also double scan or whether similar results could be obtained without duplicates, making faster analysis. Note that predictive repeatability is a required evaluation for any NIR method development (Williams et al., 2017). Table 3 shows the results obtained for predictive repeatability. The difference between repeatability per vial and per sample in NIT1 is remarkable. This shows that the effect of performing 2 scans of the same vial does not affect the results to a great extent. However, the fact of measuring two vials per sample due to the sampling and/or vial differences has a much greater effect. To the best of our knowledge, the repeatability per vial had not yet been reported, although this study showed it could influence the acquired NIR signal,

Table 3
Predictive repeatability by ANOVA analysis for oleic and linoleic acid predictions of EVOO measured at room temperature and heated with NIT1 and NIT2.

	Repeatability by scan *Df = 50*		Repeatability by vial *Df = 25*	
Instrument	**Oleic**	**Linoleic**	**Oleic**	**Linoleic**
NIT1_RT	0.32	0.11	0.71	0.32
NIT1_50	0.34	0.15	0.66	0.31
NIT2_RT	1.52	0.86	1.55	0.95
NIT2_50	1.7	0.76	1.48	0.83

Df = degrees of freedom.

and thus the results. Therefore, with NIT1 it is advisable to measure the same sample in two different vials, considering the short analysis time required compared to a more accurate prediction. Predictive repeatability of NIT2 was poorer than NIT1, as it could be expected from spectral repeatability (see section 3.1), and was similar between vials and samples. To obtain more precise predictions when using NIT2, it would be advisable to scan in two vials, both in duplicate.

These findings were in concurrence with the spectral repeatability assessment, as they substantiated that the outcomes derived from NIT1 exhibited superior spectral repeatability compared to NIT2, and moreover predictive repeatability calculated here was also higher. There was no significant difference between analyzing the olive oils at room or at higher temperature, so for a faster and more affordable analysis they could be measured at room temperature avoiding the heating steps.

4. Conclusions

This study aimed to investigate the potential use of two low-cost portable NIR instruments for rapid and non-destructive quality assessment of extra virgin olive oil (EVOO). The chemical information obtained from the spectral data allowed the development of calibrations for 7 chemical parameters that define EVOO quality, which showed good predictive capabilities. The repeatability of the two instruments was evaluated, showing the short-wavelength instrument (NIT1) to have better spectral repeatability than the upper-wavelength instrument (NIT2). The study also examined the effect of temperature on the vegetable oils to be measured, and some quality metrics showed better results in the predictions developed with data from olive oils at room temperature than heated, consistent with the literature that found higher quantification errors with increasing vegetable oil temperature. However, when predictive repeatability study was carried out, no significant differences were found between room temperature and heated olive oils, so to save time and minimize oxidation of the vegetable oils, it would be advisable to measure at room temperature.

Finally, the calibrations developed proved to detect atypical EVOO and showed good accuracy for most parameters, but more non-EVOO samples would be needed to fully evaluate the performances. The study suggests the use of portable low-cost NIR instruments for rapid and non-destructive analysis to pre-screen EVOO lots directly at production site and retain only the suspect atypical ones for laborious and expensive official analysis. This pre-screening would enable official laboratories to target samples with a greater probability of being adulterated, increasing the effectiveness of controlling EVOO on a territory. Despite the fact that these portable NIR instrument have slightly lower accuracy than the more sophisticated laboratory ones, the accuracy achieved in this study by the portable instruments would enable even small producers to have access to an analytical tool at a limited cost, which would allow the monitoring and improvement of EVOO production even in small operations.

CRediT authorship contribution statement

Alejandra Arroyo-Cerezo: Conceptualization, Data curation, Formal analysis, Software, Writing – original draft. **Xueping Yang:** Conceptualization, Writing – original draft. **Ana M. Jiménez-Carvelo:** Methodology, Writing – review & editing, Supervision. **Marina Pellegrino:** Formal analysis, Methodology. **Angela Felicita Savino:** Conceptualization, Formal analysis, Validation. **Paolo Berzaghi:** Writing – review & editing, Validation, Project administration, Resources, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Paolo Berzaghi is administrator of GraiNit srl, spin-off of the University of Padua.

Data availability

The authors do not have permission to share data.

Acknowledgements

AAC gratefully acknowledges the Spanish Ministry of Universities for a pre-doctoral fellowship FPU (FPU20/04711, Formación del Profesorado Universitario) and SensorFINT Cost Action (CA-19145) for the Short Term Scientific Mission (STSM) grant awarded for a research stay at the University of Padua. In addition, **AMJC** acknowledges the Grant (RYC2021-031993-I) funded by MCIN/AEI/501100011033 and "European Union NextGenerationEU/PRTR", and **XY** is grateful to receive funding from the China Scholarship Council program (202006350063). Funding open access charge: University of Granada / CBUA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2023.137043.

References

ASTM. (2017). *E2617-17 Standard Practice for Validation of Empirically Derived Multivariate Calibrations*. ASTM International: West Conshohocken, PA, USA. https://doi.org/10.1520/E2617-17.

ASTM. (2018). *E1655-17 Standard Practices for Infrared Multivariate Quantitative Analysis*. ASTM International: West Conshohocken, PA, USA. https://doi.org/10.1520/E1655-17.

Aykas, D. P., Karaman, A. D., Keser, B., & Rodriguez-Saona, L. (2020). Non-targeted authentication approach for extra virgin olive oil. *Foods, 9*, 221. https://doi.org/10.3390/foods9020221

Azizian, H., Kramer, J. K., & Winsborough, S. (2007). Factors influencing the fatty acid determination in fats and oils using Fourier transform near-infrared spectroscopy. *European Journal of Lipid Science and Technology, 109*, 960–968. https://doi.org/10.1002/ejlt.200700062

Borghi, F.T., Santos, P.C., Santos, F.D., Nascimento, M.H., Correa, T., Cesconetto, M., Pires, A.A., Ribeiro, A.V.F.N., Lacerda Jr, V., Romao, W., & Filgueiras, P. R. (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.

Commission Delegated Regulation (EU) 2022/2104 of 29 July 2022 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards marketing standards for olive oil, and repealing Commission Regulation (EEC) No 2568/91 and Commission Implementing Regulation (EU) No 29/2012, L 284/1. http://data.europa.eu/eli/reg_del/2022/2104/oj.

Commission Implementing Regulation (EU) 2022/2105 of 29 July 2022 laying down rules on conformity checks of marketing standards for olive oil and methods of analysis of the characteristics of olive oil, L 284/23. http://data.europa.eu/eli/reg_impl/2022/2105/oj.

Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., Lacoste, F., Breneton, P., García-Gonzálex, D. L., Moreda, W., & Toschi, T. G. (2020). Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology, 105*, 483–493. https://doi.org/10.1016/j.tifs.2019.02.025

Cox, A., Wohlschlegel, A., Jack, L., & Smart, E. (2020). The cost of food crime. *Food Standard Agency (Research Project code: FS 301065)*. https://www.food.gov.uk/research/food-crime/the-cost-of-food-crime.

De la Roza-Delgado, B., Garrido-Varo, A., Soldado, A., Arrojo, A. G., Valdés, M. C., Maroto, F., & Pérez-Marín, D. (2017). Matching portable NIRS instruments for in situ

monitoring indicators of milk composition. *Food Control, 76*, 74–81. https://doi.org/10.1016/j.foodcont.2017.01.004

García González, D. L., Aparicio, R., & Aparicio-Ruiz, R. (2018). Olive oil. In *FoodIntegrity Handbook: A Guide to Food Authenticity Issues and Analytical Solutions* (pp. 335–357). France: Eurofins Analytics.

García Martín, J. F. (2022). Potential of near-infrared spectroscopy for the determination of olive oil quality. *Sensors, 22*, 2831. https://doi.org/10.3390/s22082831

García-González, D. L., Baeten, V., Pierna, J. A. F., & Tena, N. (2013). Infrared, Raman, and fluorescence spectroscopies: Methodologies and applications. In *Handbook of Olive Oil* (pp. 335–393). Boston, MA: Springer. https://doi.org/10.1007/978-1-4614-7777-8_10.

Garrido-Varo, A., Garcia-Olmo, J., & Fearn, T. (2019). A note on Mahalanobis and related distance measures in WinISI and The Unscrambler. *Journal of Near Infrared Spectrocopy, 27*, 253–258. https://doi.org/10.1177/0967033519848296

Garrido-Varo, A., Sánchez, M. T., De la Haba, M. J., Torres, I., & Pérez-Marín, D. (2017). Fast, low-cost and non-destructive physico-chemical analysis of virgin olive oils using near-infrared reflectance spectroscopy. *Sensors, 17*, 2642. https://doi.org/10.3390/s17112642

González-Pereira, A., Otero, P., Fraga-Corral, M., Garcia-Oliveira, P., Carpena, M., Prieto, M. A., & Simal-Gandara, J. (2021). State-of-the-art of analytical techniques to determine food fraud in olive oils. *Foods, 10*, 484. https://doi.org/10.3390/foods10030484

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., & Oliphant, T. E. (2020). Array programming with NumPy. *Nature, 585*, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hassoun, A., Jagtap, S., Garcia-Garcia, G., Trollman, H., Pateiro, M., Lorenzo, J. M., Trif, M., Rusu, A. V., Aadil, R. M., Šimat, V., Cropotova, J., & Câmara, J. S. (2023). Food quality 4.0: From traditional approaches to digitalized automated analysis, 111216 *Journal of Food Engineering, 337*. https://doi.org/10.1016/j.jfoodeng.2022.111216.

Inarejos-García, A. M., Gómez-Alonso, S., Fregapane, G., & Salvador, M. D. (2013). Evaluation of minor components, sensory characteristics and quality of virgin olive oil by near infrared (NIR) spectroscopy. *Food Research International, 50*, 250–258. https://doi.org/10.1016/j.foodres.2012.10.029

Jiménez-Carvelo, A. M., González-Casado, A., Bagur-González, M. G., & Cuadros-Rodríguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity–A review. *Food Research International, 122*, 25–39. https://doi.org/10.1016/j.foodres.2019.03.063

Karunathilaka, S. R., Kia, A. R. F., Srigley, C., Chung, J. K., & Mossoba, M. M. (2016). Nontargeted, rapid screening of extra virgin olive oil products for authenticity using near-infrared spectroscopy in combination with conformity index and multivariate statistical analyses. *Journal of Food Science, 81*, C2390–C2397. https://doi.org/10.1111/1750-3841.13432

Lozano-Castellón, J., López-Yerena, A., Domínguez-López, I., Siscart-Serra, A., Fraga, N., Sámano, S., López-Sabater, C., Lamuela-Raventós, R. M., Vallverdú-Queralt, A., & Pérez, M. (2022). Extra virgin olive oil: A comprehensive review of efforts to ensure its authenticity, traceability, and safety. *Comprehensive Reviews in Food Science and Food Safety, 21*, 2639–2664. https://doi.org/10.1111/1541-4337.12949

Manley, M., & Eberle, K. (2006). Comparison of Fourier transform near infrared spectroscopy partial least square regression models for South African extra virgin olive oil using spectra collected on two spectrophotometers at different resolutions and path lengths. *Journal of Near Infrared Spectroscopy, 14*, 111–126. https://doi.org/10.1255/jnirs.597

Mialon, N., Roig, B., Capodanno, E., & Cadiere, A. (2023). Untargeted metabolomic approaches in food authenticity: A review that showcases biomarkers, 133856 *Food Chemistry, 398*. https://doi.org/10.1016/j.foodchem.2022.133856.

Özdemir, İ. S., Dağ, Ç., Özinanç, G., Suçsoran, Ö., Ertaş, E., & Bekiroğlu, S. (2018). Quantification of sterols and fatty acids of extra virgin olive oils by FT-NIR spectroscopy and multivariate statistical analyses. *LWT – Food Science and Technology, 91*, 125–132. https://doi.org/10.1016/j.lwt.2018.01.045

Pharmacopoeia, U. S. (2019). *Appendix XVIII: Guidance on Developing and Validating Non-targeted Methods for Adulteration Detection*. Rockville, MA, USA: US Pharmacopeial Convention.

Sarkar, T., Salauddin, M., Kirtonia, K., Pati, S., Rebezov, M., Khayrullin, M., … Lorenzo, J. M. (2022). A review on the commonly used methods for analysis of physical properties of food materials. *Applied Sciences, 12*, 2004. https://doi.org/10.3390/app12042004

Shenk, J., & Westerhaus, M. (1996). Calibration the ISI way. In *Near Infrared Spectroscopy: The Future Waves* (pp. 198-202), Eds. Davies, AMC and Williams. NIR Publications: Chichester, UK.

Stoltlzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine, 18*, 1099–1105. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Vanstone, N., Moore, A., Martos, P., & Neethirajan, S. (2018). Detection of the adulteration of extra virgin olive oil by near-infrared spectroscopy and chemometric techniques. *Food Quality and Safety, 2*, 189–198. https://doi.org/10.1093/fqsafe/fyy018

Willenberg, I., Matthäus, B., & Gertz, C. (2019). A new statistical approach to describe the quality of extra virgin olive oils using near infrared spectroscopy (NIR) and traditional analytical parameters. *European Journal of Lipid Science and Technology, 121*, 1800361. https://doi.org/10.1002/ejlt.201800361

Williams, P., Dardenne, P., & Flinn, P. (2017). Tutorial: Items to be included in a report on a near infrared spectroscopy project. *Journal of Near Infrared Spectroscopy, 25*, 85–90. https://doi.org/10.1177/0967033517702395

Xue, J., Yang, Z., Han, L., & Chen, L. (2014). Study of the influence of NIRS acquisition parameters on the spectral repeatability for on-line measurement of crop straw fuel properties. *Fuel, 117*, 1027–1083. https://doi.org/10.1016/j.fuel.2013.10.017

Yan, J., Erasmus, S. W., Toro, M. A., Huang, H., & van Ruth, S. M. (2020). Food fraud: Assessing fraud vulnerability in the extra virgin olive oil supply chain, 107081 *Food Control, 111*. https://doi.org/10.1016/j.foodcont.2019.107081.

Zaroual, H., Chénè, C., El Hadrami, E. M., & Karoui, R. (2022). Application of new emerging techniques in combination with classical methods for the determination of the quality and authenticity of olive oil: A review. *Critical Reviews in Food Science and Nutrition, 62*, 4526–4549. https://doi.org/10.1080/10408398.2021.1876624