



Vibrational Spectroscopy Coupled with Chemometrics for Olive Oil Authentication: Recent Advances and Future Perspectives

Huri İlyasoğlu¹

Received: 7 March 2026 / Accepted: 13 June 2026 / Published online: 26 June 2026
© The Author(s) 2026

Abstract

Food authentication has become a critical issue for consumers, producers, and regulatory authorities, particularly for high-value products such as olive oil, which is highly susceptible to adulteration. In recent years, vibrational spectroscopy techniques, including mid-infrared (MIR), near-infrared (NIR), and Raman spectroscopy, have attracted significant attention as rapid and non-destructive analytical tools for assessing the authenticity of olive oil. When combined with multivariate chemometric approaches, these techniques enable the extraction of meaningful spectral fingerprints and the development of predictive models for adulteration detection. This review provides a comprehensive evaluation of studies conducted over the past decade, with a focus on the application of vibrational spectroscopy coupled with chemometrics for the authentication of olive oil. The reviewed studies demonstrate that combining vibrational spectroscopy with chemometric methods offers advantages, including minimal sample preparation, rapid analysis, environmental sustainability, and potential suitability for routine quality control and official food inspection. The current limitations and future research directions are also highlighted to facilitate the extensive implementation of these techniques in olive oil authentication frameworks.

Keywords Adulteration · Authentication · Chemometrics · Extra virgin olive oil · Vibrational spectroscopy

Introduction

Olive oil, a key component of the Mediterranean diet, is widely recognised for its diverse health benefits, attributable to its unique chemical composition. Olive oil is characterised by its high concentration of monounsaturated fatty acids, predominantly oleic acid. Furthermore, it contains minor components such as sterols, polyphenols, hydrocarbons, terpenes, chlorophyll and carotenoid pigments, and volatile compounds, which account for only 1–2% (Gagour et al. 2024). The research evidence, when considered as a whole, suggests that olive oil has a beneficial effect on several chronic diseases, including cardiovascular disease, breast cancer, and type 2 diabetes (Foscolou et al. 2018).

Olive oil is one of the foodstuffs most prone to adulteration due to its high price. Consequently, the authenticity of olive oil is of paramount importance (Kakouri et al. 2021). Authentication can be defined as the process of controlling

fraudulent practices, including adulteration, mislabelling, and mislabelling of origin (Bajoub et al. 2018). Various analytical techniques have been developed to determine the authenticity of olive oil. The classical analytical techniques include chemical analysis, liquid chromatography, and gas chromatography. Chromatographic methods enable the identification and quantification of individual constituents, including fatty acids, sterols, and minor compounds, thereby offering strong confirmatory evidence for authenticity assessment. However, these methods are both time-consuming and destructive (Zaroual et al. 2022). Therefore, the development of rapid and non-destructive methods for determining the authenticity of olive oil is imperative. Spectroscopic methods, with a particular emphasis on vibrational spectroscopy, have been shown to facilitate rapid and non-destructive procedures requiring minimal sample preparation. These methodologies enable the representation of a wide range of constituents in the oil through their respective spectra (Ozen et al. 2024).

The interpretation of spectral data often requires subsequent data analysis due to the complex nature of such data. In this context, multivariate analysis is a valuable tool for data analysis. In the olive oil field, chemometrics can be

✉ Huri İlyasoğlu
huriilyasoglu@yahoo.com; hilyasoglu@gumushane.edu.tr

¹ Department of Nutrition and Dietetics, Gümüşhane University, Gümüşhane, Türkiye

useful for identifying cultivar markers, facilitating the differentiation of samples by cultivar and geographical origin, guaranteeing the authenticity of olive oil, and/or detecting adulteration with other oils (Gómez-Caravaca et al. 2016; Rodionova and Pomerantsev 2020).

Rather than targeting individual compounds, vibrational spectroscopic methods rely on chemical fingerprints combined with chemometric analysis to enable classification and prediction (Kakouri et al. 2021). In this context, vibrational spectroscopy is more appropriately considered as a screening tool or complementary approach, rather than a direct replacement for classical analytical techniques. The effectiveness of spectroscopic approaches is strongly linked to the application of chemometric methods, which facilitate the extraction of meaningful information from complex, high-dimensional datasets. Recent studies in complex food systems have demonstrated the potential of chemometrics-assisted untargeted fingerprinting approaches, including colloidal and volatile profiling, to enable rapid and green authentication strategies. Such approaches illustrate how complementary analytical techniques can be integrated within a unified chemometric framework to enhance discrimination, traceability, and quality assessment (Zappi et al. 2022, 2023; Giordani et al. 2024).

Numerous literature reviews have explored the application of spectroscopic techniques in the analysis of olive oils. The present review aimed to provide an overview of vibrational spectroscopic techniques (MIR, NIR, and Raman spectroscopy) published between 2015 and 2025, with a particular focus on their application to olive oil authentication.

Vibrational Spectroscopy

Vibrational spectroscopy is a technique that measures the vibrational energy in a compound. Each chemical bond corresponds to a specific vibrational energy, thereby generating a distinctive fingerprint. This characteristic fingerprint can enable the determination of the structure of an unknown compound by comparing it with the fingerprints of known compounds (Campanella et al. 2021).

Vibrational spectroscopy involves the analysis of the specific spectral range within the electromagnetic spectrum. This range extends from 0.76 to 1000 μm or from 13,000 to 10 cm^{-1} , which corresponds to the frequency of molecular vibrations (Wang et al. 2016).

Vibrational spectroscopic techniques include IR spectroscopy and Raman spectroscopy. IR spectroscopy is the most used experimental technique in vibrational spectroscopy, relying on measuring the absorption of infrared radiation by matter at different wavelengths. The technique enables investigation of the molecular structure and the interactions of functional groups within molecules, as well as their

interactions with their environment. The IR spectrum of each sample is unique, thus serving as a fingerprint to identify functional groups or whole compounds and to provide insight into the evolution of the structure (Ditler and Lubert 2022).

The infrared component of electromagnetic radiation can be categorised into three distinct regions: NIR, which extends from 13,000 to 4000 cm^{-1} ; MIR, which ranges from 4000 to 400 cm^{-1} ; and the far-infrared, which covers the spectral range from 400 to 10 cm^{-1} (Wang et al. 2016).

MIR spectroscopy is a valuable tool in analysing chemical functional groups, providing insight into their fundamental vibrations. These functional groups exhibit absorption within the MIR region. The MIR region of the electromagnetic spectrum can typically be subdivided into two distinct regions: the functional group region (4000 to 1500 cm^{-1}) and the fingerprint region (1500 to 400 cm^{-1}). Most of the pertinent information utilised for interpreting MIR spectra is obtained from the functional group region (Mousa et al. 2022).

NIR spectroscopy records molecular combination and overtone vibrations. The vibration signals are analogous, and the spectrum is composed of numerous broad and overlapping bands. The detailed structural elucidation of the analysed sample is limited by overtone and combination bands (Li et al. 2020).

Raman spectroscopy relies on the inelastic scattering of light across the near-infrared, visible, and near-ultraviolet regions. Light excites the molecule to a higher energy state, and as it relaxes back to the ground state, information about vibrational modes is conveyed via the resulting frequency shift (Ditler and Lubert 2022).

Methods

This study was designed as a structured narrative review focusing on the application of vibrational spectroscopic techniques and chemometric approaches for olive oil authentication and adulteration detection.

The Scopus database was systematically searched for articles published in English between 2015 and 2025. The search was conducted using keywords appearing in the title, abstract, and keyword fields (last searched in September 2025). The search strategy combined terms related to olive oil authentication and adulteration with vibrational spectroscopic techniques and chemometric approaches. A representative search string included (“olive oil” AND “adulteration” AND (“FTIR” OR “NIR” OR “Raman” OR “vibrational spectroscopy” OR “infrared”) AND (“chemometrics” OR “machine learning”)) as well as (“olive oil” AND “authentication” AND (“FTIR” OR “NIR” OR “Raman” OR “vibrational spectroscopy” OR “infrared”) AND (“chemometrics”

OR “machine learning”). These terms were applied to the “title”, “abstract”, and “keywords” search fields to capture studies employing both classical chemometric techniques and data-driven analytical approaches.

Study selection was performed manually by the author through title screening, abstract evaluation, and full-text assessment. Studies were included if they (i) applied vibrational spectroscopy to olive oil authentication, (ii) employed chemometric or machine learning methods, and (iii) were published in peer-reviewed journals in English. Studies were excluded if they (i) did not involve olive oil authentication, (ii) lacked vibrational spectroscopy or chemometric analysis, and (iii) were conference abstracts, review papers, or duplicate records.

The initial search retrieved 193 records. After removal of duplicates and screening based on title, abstract, and full-text evaluation, 46 studies were ultimately included in the review.

Studies employing multiple analytical techniques were included when vibrational spectroscopy (FTIR, Raman, or NIR) was part of the analytical approach. The inclusion of these studies aimed to provide broader insights into applications and complementary information relevant to vibrational spectroscopic approaches for olive oil authentication.

Digital tools were used exclusively for non-scientific, supportive purposes. Elicit (free version) was employed to extract and organise structured information from the selected publications, including reported methodologies and key findings. All extracted information was subsequently verified, interpreted, and critically assessed by the author. ChatGPT (Academic GPT, free version) and DeepL (free version) were used solely for language refinement, grammatical correction, and improvement of clarity and readability.

All stages of literature screening, study selection, synthesis of findings, interpretation, and conclusions were conducted exclusively by the author. All AI-assisted edits were reviewed and verified by the author, who retained full responsibility for the scientific content and final version of the manuscript.

MIR Spectroscopy Applications in Detection of Olive Oil Adulteration

Over the past decade, several studies have explored the use of MIR spectroscopy coupled with chemometrics for the discrimination and quantification of EVOO adulteration (Table 1).

Sun et al. (2015) studied the detection and quantification of adulteration in EVOO with edible oils, including camellia, soybean, sunflower, and corn oils. The analysis of the FTIR data was performed using chemometrics, with principal component analysis (PCA), locally linear embedding (LLE), and supervised locally linear embedding (SLLE)

being utilised for dimensionality reduction. SLLE provided satisfactory results in identifying and quantifying adulteration of EVOO. For the PLS models, the R^2 value for the validation model ranged from 0.9757 to 0.9988. The root mean square error (RMSE) for the test set was between 0.0172 and 0.0951. The findings revealed that applying the first derivative, followed by smoothing, mean centering, and SLLE, achieved the best results in both the classification and quantification models.

A method using FTIR spectroscopy was developed to detect adulteration of EVOO with peanut oil (Vasconcelos et al. 2015). Correct classification rates of 92.3% for the calibration set and 88.3% for the cross-validation set were obtained using LDA. The PCR and PLS-R enabled the quantification of adulteration levels, with satisfactory R^2 values and low RMSE. The PLS-R model, when applied to the wavenumbers at 3007, 2922, 2853, 1754, 1160, and 1117 cm^{-1} , was found to be a reliable method for quantifying and discriminating peanut oil in EVOO. The study successfully detected adulteration at levels as low as 0.5% v/v.

Rohman et al. (2017) developed a methodology to detect adulteration with grape seed oil (GSO), soybean oil (SO), and walnut oil (WO). DA, PCR, and PLS were applied using the combined spectral regions of 3018–3002 and 1200–1000 cm^{-1} for GSO and SO, and 3029–2954 and 1125–667 cm^{-1} for WO. The use of optimised wavelengths by the DA provided classification of EVOO and adulterated samples with a 100% accuracy rate. Furthermore, PLS at optimised wavelengths demonstrated good predictive performance for adulterant levels ($R^2 > 0.99$, RMSEC: 0.101–0.658%).

A hand-held FTIR instrument was employed to develop a portable screening strategy for identifying adulterated olive oils (Pan et al. 2018). The study involved canola, peanut, corn, soybean, and sunflower oils. The SIMCA model correctly identified different types of vegetable oils and detected adulteration in olive oils containing at least 10% adulterant. Adulterant oil content in the range of 5–45% was estimated using a PLS model, with a reported error of 3%.

Nurwahidah et al. (2019) employed FTIR spectroscopy in conjunction with chemometric methods (DA, PLS, and PCR) to detect adulteration of EVOO with rambutan seed fat (RSF). DA successfully distinguished between unadulterated and adulterated EVOO without making any misclassifications. The best quantitative model for predicting RSF levels in EVOO was attained by utilising normal FTIR spectra at designated wavenumber regions (1446.8–1409.7 cm^{-1} and 2368.6–1769.9 cm^{-1}) in conjunction with PCR.

A methodology was proposed using FTIR spectroscopy and PLS algorithms to identify and quantify adulterants in EVOO (Filoda et al. 2019). The study demonstrated the capability of the FTIR-PLS method in identifying the types and levels of adulteration. Using siPLS with the spectrum

Table 1 Application of MIR spectroscopy for detecting adulterants in olive oil or olive oil classification (cultivar and geographical origin)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Camellia, corn, soybean, and sunflower oils	EVOO ($n=40$) Edible oils ($n=20$) Adulterated samples ($n=240$) Laboratory-prepared mixtures, EVOO, and edible oils	4000–400 cm^{-1}	PCA LLE SLLE PLS	Random train/test split; cross-validation	SLLE performed best in the classification and quantification tasks R^2 : 0.9757–0.9998 and RMSE:0.0172–0.0951	Sun et al. (2015)
Peanut oil	EVOO ($n=5$) Peanut oil ($n=5$) Adulterated samples ($n=207$) Laboratory-prepared mixtures, EVOO, and peanut oil	4000–500 cm^{-1}	PCA LDA PCR PLS-R	Cross-validation; independent test set	LDA provided 88.3% correct classification in cross-validation PLS-R model had the highest R^2 value (0.977) and lowest RMSEP (1.05%)	Vasconcelos et al. (2015)
Grape seed, soybean, and walnut oils	EVOO ($n=20$) Adulterated samples ($n=75$) Laboratory-prepared mixtures	4000–650 cm^{-1}	DA PCR PLS	Cross-validation; independent test set	DA model classified EVOO and adulterated samples with 100% accuracy PLS models had a higher R^2 (0.991–0.998) and a lower RMSEC (0.101–0.658%)	Rohman et al. (2017)
Canola, peanut, corn, soybean, and sunflower oils	EVOO ($n=25$) Vegetable oils ($n=59$) Adulterated samples ($n=45$) Laboratory-prepared mixtures and vegetable oils	4000–650 cm^{-1}	PCA SIMCA PLS	Independent test set	SIMCA model can detect adulteration with at least 10% adulterant content R^2 : 0.9912–0.9943 and SEP: 1.01–1.26%	Pan et al. (2018)
Rambutan seed fat	Adulterated samples ($n=40$) Laboratory-prepared mixtures	4000–650 cm^{-1}	DA PLS PCR	Independent test set	DA model successfully classified EVOO and adulterated EVOO samples PCR yielded the best quantification model (R^2 :0.9915 and RMSEP: 3.68%)	Nurwahidah et al. (2019)
Soybean, canola, sunflower, and corn oils	Adulterated samples ($n=68$) Laboratory-prepared mixtures and commercial EVOOs	3200–650 cm^{-1}	PLS	Cross-validation; external validation	siPLS32 model exhibited the best results (R^2 : 0.9991 and RMSECV: 0.86%)	Filoda et al. (2019)
Hazelnut, canola, and safflower oils	EVOO ($n=6$) Vegetable oils ($n=8$) Adulterated samples ($n=108$) Laboratory-prepared mixtures and EVOOs	4000–625 cm^{-1}	PCA SIMCA PLS-DA	Cross-validation; external validation set	The developed models yielded good predictive performance (> 92%)	Ordoudi et al. (2022)

Table 1 (continued)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Morocco Picholine EVOO adulteration with Arbequina VOO	Pure olive oil samples ($n=15$) Adulterated samples ($n=60$) Laboratory-prepared mixtures and pure olive oil samples	4000–600 cm^{-1}	PCA PCR PLS-R	Cross-validation; independent test set	Calibration models resulted in a high R^2 value (> 0.995) and a low RMSE (< 2.17)	Hirri et al. (2023)
Refined olive oil, pomace olive oil, sunflower, soybean, and maize oils	Pure oil samples ($n=53$) Adulterated samples ($n=600$) Laboratory-prepared mixtures	4000–650 cm^{-1}	PLS-DA SVM RF BPNN	Train/test split	SVM model achieved the highest accuracy (99.1%)	Jiao et al. (2025)
Old olive oil	Pure olive oil samples ($n=10$) Adulterated samples ($n=45$) Laboratory-prepared mixtures	4000–600 cm^{-1}	PLS-DA PCR PLS-R	Cross-validation; independent test set	PLS-DA model achieved 100% correct classification PLS-R model showed better performance (R^2 : 0.999 and RMSEP: 0.2401)	Hirri et al. (2015)
Old olive oil	Fresh olive oil samples ($n=20$) Old olive oil samples ($n=5$) Adulterated samples ($n=100$) Laboratory-prepared mixtures and fresh olive oils	4000–650 cm^{-1}	OPLS-DA PLS-R	Cross-validation; external validation set	OPLS-DA models achieved higher correct classification rates (83–93%) in validation sets	Uncu and Ozen (2019)
Old olive oil	Adulterated samples ($n=50$) Laboratory-prepared mixtures	4000–650 cm^{-1}	PLS-R	Cross-validation	PLS-R model yielded a satisfactory R^2 value (> 0.94) and a low RMSECV (< 1.6)	Laouni et al. (2023a)
Origin/cultivar						
Geographical origin of Moroccan VOO	Four regions ($n=49$)	4000–600 cm^{-1}	PCA PLS-DA	Cross-validation; external validation	PLS-DA model achieved 100% accuracy	Benlmaalam et al. (2015)
Commercial-grade olive oils from Morocco	Four grades ($n=70$)	4000–600 cm^{-1}	PLS-DA	Cross-validation; external validation	PLS-DA model achieved 100% accuracy	Hirri et al. (2016)
Greek cultivars	Koroneiki ($n=41$) Megaritiki ($n=30$) Amfissis ($n=19$)	4000–650 cm^{-1}	PCA LDA QDA	Cross-validation	LDA and QDA models yielded a correct classification rate of 96.1%	Revelou et al. (2021)
Maltese EVOO	Maltese EVOO ($n=30$) Non-Maltese EVOO ($n=35$)	4000–630 cm^{-1}	PCA LDA PLS-DA SVM ANN	Cross-validation	Supervised models discriminated between Maltese and non-Maltese EVOOs	Lia et al. (2021)

Table 1 (continued)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Geographic origin of Italian EVOO	Six regions ($n = 72$)	3890–650 cm^{-1}	PCA SIMCA SVM-R PCR		SVM-R discriminated Toscana, Lazio, Puglia and Calabria samples	Scatigno and Festa (2022)
Geographic origin of Moroccan VOO	Six regions ($n = 360$)	4000–600 cm^{-1}	PCA LDA PLS-DA	Cross-validation; external validation	PCA-LDA and PLS-DA models achieved satisfactory performance (CCR > 84%)	Laouni et al. (2023b)

partitioned into 32 regions, the calibration model provided the best results when 12 latent variables were applied along the spectral ranges 3042–2806, 2726–2648, and 1225–1068 cm^{-1} . The calibration model performed best when pre-processing was applied using the second derivative and mean centering.

Ordoudi et al. (2022) developed a method to detect ternary blends of adulterants in EVOO at a standard adulteration level of 20% v/v. The second-derivative FTIR spectra were examined for outliers during multivariate calibration using PLS-DA and SIMCA. The stepwise external validation of the models showed 100% sensitivity in detecting adulteration, such as 5% canola and 15% hazelnut oils. The developed models achieved good predictive performance (> 92%).

The application of FTIR spectroscopy combined with PLS and PCR was examined for detecting and quantifying adulteration in Moroccan Picholine EVOO blended with Arbequina VOO (Hirri et al. 2023). The PLS and PCR calibration models exhibited high R^2 values and low RMSE values, indicating satisfactory performance in quantifying adulteration.

Jiao et al. (2025) investigated the detection of adulteration in EVOO with various types of vegetable oils. Four classification algorithms (PLS-DA, SVM, RF, and BPNN) were employed in the study. The application of FTIR spectroscopy coupled with an SVM achieved the highest accuracy of 99.1%, accompanied by a detection limit of 2.5%, indicating its suitability for high-precision detection. The models based on SNV + first derivative + SG smoothing or SNV + second derivative + SG smoothing pre-processing exhibited the best classification performance.

A methodology was developed using FTIR spectroscopy to detect EVOO adulteration with lower-quality aged olive oils (Hirri et al. 2015). The frequencies within the fingerprint region, particularly those occurring at 3100–600 cm^{-1} , were applied for both identifying and quantifying. PLS-DA model achieved a classification rate of 100% according to the oils' respective groups. PLS-R model provided better results ($R^2 = 0.999$, RMSEP = 0.2401) than the PCR model ($R^2 = 0.997$, RMSEP = 0.4865) for quantitative analysis of lower-quality olive oil in EVOO, thereby underscoring its capability in such applications.

Uncu and Ozen (2019) investigated the detection of adulteration in fresh olive oils with aged olive oils. Various pre-processing techniques were applied, including the first and second derivatives, and a combination of wavelength denoising methods and the first derivative. The OPLS-DA models successfully distinguished adulterated from non-adulterated oils, achieving classification accuracies above 80%.

A study was conducted using FTIR spectroscopy to identify and quantify adulteration of fresh olive oils with aged and commercial oils (Laouni et al. 2023a). The PLS-R model showed good performance in predicting adulteration

levels, as evidenced by satisfactory coefficients of determination and low error values for both calibration and cross-validation. The study evaluated the performance of different pre-processing techniques and found that applying the second derivative to FTIR enhanced model performance.

The reviewed studies consistently demonstrate that FTIR in conjunction with chemometric techniques provides accurate classification and robust predictive capability for detecting adulteration in EVOO. Most studies report very high coefficients of determination ($R^2 > 0.97$) and low prediction errors, particularly when PLS-based models are used. However, a more critical evaluation reveals that these performance metrics are strongly influenced by experimental design, variable selection, spectral pre-processing, and validation strategy. In many studies, the reported high predictive performance is obtained using laboratory-prepared binary mixtures with well-defined adulteration ranges. Such systems are characterised by simplified experimental conditions, including controlled sample preparation, known adulterant identity and concentration, single-adulterant formulations, and relatively homogeneous fresh samples. Although these experimental models are valuable for proof-of-concept investigations, they may not adequately represent the chemical complexity and variability encountered in real commercial fraud scenarios. In practice, adulteration may involve multiple adulterants, lower-quality oils with similar compositional profiles, processing-induced variability, ageing effects, and sophisticated blending strategies that mimic authentic EVOO. Accordingly, studies investigating more complex systems, such as ternary mixtures and compositionally similar adulterants, generally report lower model performance. Furthermore, studies employing optimised wavelength selection generally achieve greater robustness than full-spectrum models, suggesting that targeted spectral regions, especially within the fingerprint and C–H stretching regions, contain the most discriminatory information. Several studies also demonstrate that spectral pre-processing substantially influences model accuracy and robustness. Appropriate pre-processing techniques can improve signal discrimination and reduce irrelevant spectral variation. Nevertheless, the literature does not support a universally optimal pre-processing workflow, as pre-processing performance often depends on instrumental conditions and modelling strategy. This variability highlights the importance of systematic pre-processing optimisation to ensure analytical transferability across different datasets and laboratory environments. Another critical factor is the validation approach. The validation strategies exhibit considerable variability, with important implications for the interpretation of reported model performance. A substantial proportion of studies rely on internal validation techniques. While these approaches are valuable for model optimisation, they cannot provide robust evidence of generalizability. Notably, the highest

reported performance metrics are primarily associated with internal validation methods, suggesting potential optimism bias. A few studies employ external validation, which represents the most rigorous assessment of model transferability.

The evidence indicates that FTIR spectroscopy is a powerful screening tool for olive oil authentication; however, greater emphasis on complex sample sets, the diversity of real market samples, external validation, and standardised performance metrics is required to enable routine implementation in official control laboratories.

MIR Spectroscopy Applications in Geographical and Botanical Origin Authentication of Olive Oil

MIR spectroscopy in conjunction with chemometrics has also been examined for geographical and botanical authentication (Table 1).

MIR spectroscopy combined with chemometric approaches was employed to differentiate VOOs from four distinct geographical regions in Morocco (Benlamma et al. 2015). The full-spectrum data acquired ($4000\text{--}600\text{ cm}^{-1}$) was subjected to PCA and PLS-DA analysis. In the validation set, the PLS-DA algorithm achieved 100% correct classification according to the geographical origin.

Hirri et al. (2016) studied the classification of four commercial-grade virgin olive oils. The study included olive oils, representing different commercial grades (extra virgin, virgin, ordinary virgin, and lampante), collected from Central Morocco. The PLS-DA model demonstrated high classification accuracy in both calibration and external validation, achieving 100% accuracy in classifying VOO grades. Furthermore, they developed a PLS-DA model using physico-chemical properties, but it was unable to distinguish between the different grades.

Revelou et al. (2021) developed chemometric models using FTIR spectra for the purpose of differentiating Greek olive cultivars (Koroneiki, Megaritiki, and Amfissis). The FTIR spectra were subjected to forward stepwise variable selection to identify discriminative features. The spectral regions selected corresponded primarily to C=O stretching vibrations of triglyceride ester groups and to C–H bending vibrations related to aliphatic CH_2 chains and double bonds. LDA and QDA achieved a classification accuracy of 96.1%, demonstrating their potential as powerful tools for the authentication of Greek olive oils.

A study was conducted to evaluate the capability of FTIR spectroscopy for authenticating Maltese EVOO (Lia et al. 2021). The application of supervised classification methods, including PLS-DA, ANN, LDA, and SVM, demonstrated the capability for discriminating Maltese and non-Maltese

EVOOs. Applying Savitsky-Golay first- and second-order derivative pre-processing enhanced the performance of multivariate models for discrimination and prediction of Maltese EVOOs.

Scatigno and Festa (2022) used FTIR spectroscopy to analyse Italian EVOO samples from 6 regions, with a particular focus on geographical traceability. The machine learning algorithms used were SIMCA, SVM-R, and PCR. The regions of Toscana, Lazio, Puglia, and Calabria exhibited strong reproducibility with respect to geo-traceability, unlike Sicily and Umbria.

The use of FTIR spectroscopy for predicting the geographical origin of Moroccan virgin olive oils was investigated using samples collected from six regions across two consecutive harvest seasons (Laouni et al. 2023b). The PCA-LDA and PLS-DA methods successfully classified olive oils according to geographic provenance, achieving classification rates above 84.0%. The study highlighted the importance of pre-processing techniques in enhancing the accuracy of classification models.

The reviewed studies demonstrate the potential of FTIR spectroscopy combined with chemometric methods for the authentication of olive oils according to their geographical and botanical origins. Supervised methods such as PLS-DA, LDA, and QDA have frequently achieved classification accuracies exceeding 90%, and in several cases, near-perfect discrimination. However, a more critical evaluation indicates that these performance metrics are strongly influenced by sample design and experimental conditions. Studies reporting near-perfect classification accuracy frequently rely on relatively small sample sets ($n < 100$) or on oils originating from geographically distinct regions, where compositional differences are more pronounced, and discrimination is inherently less challenging. Most studies focus on simplified experimental settings involving single-factor authentication, such as geographical, varietal, or quality classification. These experimental designs typically involve regional classifications with well-defined geographic boundaries or varietal classifications based on clearly distinguishable botanical characteristics. However, the chemical similarity of olive oils produced in neighbouring regions or from closely related cultivars represents a substantial analytical challenge. In such cases, overlapping spectral features may reduce the discriminatory capability of MIR-based models, leading to decreased classification performance under more realistic authentication conditions. The variability observed among Italian regional studies and the moderate accuracies reported in certain multi-regional Moroccan investigations illustrate the increasing complexity of real-world traceability scenarios. In addition, many studies rely on single or two-harvest datasets, which do not fully capture the effects of seasonal variability, climatic conditions, and agricultural practices on olive oil composition. As a result, the generalizability

of these models across different production years remains uncertain. Although reporting of pre-processing procedures is often limited, several studies demonstrate that spectral pre-processing substantially influences classification performance and model robustness. The available evidence suggests that optimal pre-processing strategies depend on the chemometric method employed. Accordingly, the absence of standardised pre-processing workflows remains an important limitation affecting method transferability and inter-study reproducibility. Another important consideration is the validation strategy. While high classification accuracy is frequently reported, a substantial proportion of studies rely primarily on internal validation approaches, which may lead to optimistic estimates of model performance. Consequently, external validation is critical for developing reliable models applicable to real-world traceability.

Although FTIR-based chemometric models show capability for the geographical and botanical authentication of olive oil, ensuring reproducibility and enabling their implementation in official traceability systems require greater emphasis on large-scale, multi-harvest datasets and standardised validation protocols.

NIR Spectroscopy Applications in Detection of Olive Oil Adulteration

The use of NIR spectroscopy combined with chemometrics has been investigated for detecting adulteration in EVOO (Table 2).

A rapid methodology was developed using FT-NIR spectroscopy and PLS calibration models to authenticate EVOO and detect adulterants (Azizian et al. 2015). A PLS1 calibration model was constructed to predict the FT-NIR index based on the relative intensities of selected carbonyl overtone absorption bands. Furthermore, four PLS1 calibration models were developed using FT-NIR spectral data to discriminate among four distinct adulterant groups, each characterised by a specific fatty acid profile. The first group of adulterants comprised oils high in linoleic acid, namely soybean, sunflower, corn, and canola oils. The second group comprised oils high in oleic acid, including hazelnut, high-oleic acid sunflower, and peanut oils. The third and fourth groups of adulterants were palm olein and refined olive oil, respectively.

Karunathilaka et al. (2016) developed a rapid tool to evaluate the authenticity of EVOO retail products using NIR spectroscopy. The study validated these analytical methods using reference samples, spiked samples, and commercial products to assess their capability in identifying authentic EVOO. Visual examination of FT-NIR spectra revealed distinct spectral features that differentiated EVOO from adulterant oils. These differences were observed in bands

Table 2 Application of NIR spectroscopy for detecting adulterants in olive oil or olive oil classification (cultivar and geographical origin)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Vegetable oils	NR	4500–8000 cm^{-1}	PLS	Cross-validation	Four PLS calibration models were developed (R^2 : 97.6–99.99, RMSECV: 0.9–3.7%)	Azizian et al. (2015)
Authenticity of EVOO retail products	Reference EVOO samples ($n=62$) Edible oils ($n=10$) Blends ($n=20$) EVOO retail products ($n=88$) Laboratory-prepared mixtures and commercial samples	4000–12500 cm^{-1}	PCA SIMCA	Cross-validation; external validation	The SIMCA model demonstrated 61% of retail products were not classified as EVOO	Karunathilaka et al. (2016)
Corn, sunflower, soybean and canola oils	EVOO ($n=47$) Vegetable oils ($n=20$) Adulterated samples ($n=40$) Laboratory-prepared mixtures and EVOOs	4000–12000 cm^{-1}	PCA SIMCA	Independent test set	The detection limit was established, ranging from 10 to 20%, without an unadulterated sample for reference	Vanstone et al. (2018)
Peanut, sunflower, soybean, sesame, and maize oils	Adulterated samples ($n=100$) Laboratory-prepared mixtures	4000–10000 cm^{-1}	BOSS-PLS CARS-PLS IVIR-PLS MCUVE-PLS	Cross-validation; independent validation set	BOSS-PLS model outperformed other models (R^2 : 0.9992 and RMSEP: 1.48889%)	Jiang and Chen (2019)
ROO and POO	EVOO ($n=80$) ROO ($n=40$) POO ($n=10$) Commercial olive oil samples	908–1676 nm	PCA OckNN OCSVM SIMCA	Repeated random cross-validation; external validation set	The SIMCA model demonstrated higher accuracy	Yan et al. (2019)
Soybean, sunflower, corn, and canola oils	Adulterated samples ($n=228$) Laboratory-prepared mixtures and commercial olive oil samples	908–1676 nm	PCA SIMCA PLS	Train/test split; cross-validation; external test set	SIMCA model classified correctly EVOO samples PLS model for transmittance mode yielded a high R^2 value (0.99) and a low RMSEP (0.17–1.40%)	Borghini et al., (2020)
Cottonseed, canola, corn, soybean, and sunflower oils	EVOO ($n=10$) Edible oils ($n=5$) Adulterated samples ($n=75$) Laboratory-prepared mixtures and EVOOs	908–1676 nm	SVM PLS-DA PLS	Train/test split; cross-validation; external validation	SVM demonstrated better accuracy than PLS-DA R^2 : 0.90–0.99 and RMSEP: 2.06–4.36%	Folli et al. (2022)

Table 2 (continued)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Refined hazelnut, sunflower, linseed, and corn oils	EVOO ($n=30$) Adulterated samples ($n=480$) Laboratory-prepared mixtures	800.62–2499.64 nm	PCA/RDA LDA CART kNN SVM RF	Cross-validation	RDA factors improved the results obtained from PCA factors in both calibration and validation	Sánchez-Rodríguez et al. (2022)
Pomace, soybean, corn, and sunflower oils	EVOO ($n=40$) Adulterated samples ($n=144$) Laboratory-prepared mixtures and EVOOs	350–750 nm	PCA LDA SVM LR	Cross-validation; external validation	LDA, SVM, and LR models achieved highly accurate predictions ($>97\%$)	Nanou et al. (2025)
Origin/cultivar						
Varietal origin of Tunisian VOO	Tunisian origin ($n=184$) Foreign countries ($n=63$)	4500–10000 cm^{-1}	PCA SIMCA	Independent test set	SIMCA model achieved high correct classification rates for Chemlali Sfax, Cheuti and Queslatti oils (89.6–98.50%)	Laroussi-Mezghani et al. (2015)
Geographical origin of Cretan EVOO	Istria ($n=23$) Kvarner ($n=10$) Dalmatia ($n=16$) Others ($n=11$)	400–2500 nm	PCA PLS-DA	Cross-validation; external validation set	PLS-DA model demonstrated a high potential for geographical origin discrimination	Peršurić et al. (2018)
Geographical origin of Argentinean EVOO	RDO-EVOOs ($n=52$) Unknown RDO-EVOOs ($n=8$)	1100–2500 nm	PCA PLS-DA	Train/test split; cross-validation; external test set	PLS-DA model verified RDO samples with a 100% correct classification rate	Jiménez-Carvelo et al. (2019)

NR not reported

near 5280 and 5180 cm^{-1} , which are attributed to the second overtone of the C=O stretching vibration. The SIMCA analysis indicated potential quality issues or adulteration in approximately 61% of the commercial EVOO products.

A NIR spectroscopy-based method was developed to detect the adulteration of EVOO with various edible oils, with PCA and SIMCA employed for data analysis and classification (Vanstone et al. 2018). PCA plot demonstrated that adulteration was detectable at a low level (2.7%) if an authentic sample was included as a control. However, in the absence of unadulterated control samples, the detection limits were found to depend on the types of EVOO and adulterant. PCA with SIMCA was used to detect adulteration, with established detection limits of 20% for corn and sunflower oils, 15% for soybean oil, and 10% for canola oil.

Jiang and Chen (2019) employed the FT-NIR spectroscopy to determine adulteration content in EVOO. The bootstrapping soft shrinkage (BOSS) algorithm was utilised for the variable selection during PLS modelling to select informative spectral wavenumbers. The BOSS-PLS model demonstrated superior predictive performance relative to other models, including competitive adaptive reweighted sampling PLS (CARS-PLS), Monte Carlo uninformative variable elimination PLS (MCUVE-PLS), and iteratively retaining informative variables PLS (IRIV-PLS). Using the BOSS-PLS model yielded both a high R^2 value (0.9922) and a low RMSEP (1.4889% v/v).

The use of handheld NIR spectroscopy was investigated to differentiate EVOO from other olive oil grades, including refined olive oil (ROO) and pomace olive oil (POO) (Yan et al. 2019). The study involved one-class classification models with three different classifiers: OCKNN, OCSVM, and SIMCA. SIMCA classification models accurately differentiated EVOO from ROO (94%) and POO (100%), indicating the method's effectiveness.

Borghi et al. (2020) developed a method using a portable near-infrared spectrometer to identify and quantify adulterations in EVOO with vegetable oils. Two distinct spectral acquisition modes, namely reflectance and transmittance, were rigorously evaluated. Quantification of adulterants in EVOO was performed using PLS models, which demonstrated superior performance for the transmittance mode compared with the reflectance mode. Using the SIMCA model, EVOO samples and their mixtures were assigned correctly to the target EVOO class.

Folli et al. (2022) investigated the classification and identification of adulterants in EVOO using a portable NIR spectrometer. The study involved cottonseed, canola, corn, soybean, and sunflower oils as adulterants. They compared the performance of SVM and PLS-DA models for classification, with SVM showing better accuracy. The PLS models achieved good performance in quantifying adulterants in EVOO (RMSEP: 2.06–4.36%).

The application of PCA and redundancy analysis (RDA) for reducing high-dimensional near-infrared spectral data to enhance olive oil authentication was examined using various machine learning algorithms, including LDA, CART, kNN, SVM, and RF (Sánchez-Rodríguez et al. 2022). PCA and RDA factors were incorporated as explanatory variables in the developed models to evaluate oil authenticity using qualitative or quantitative approaches. RDA factors improved the results in both regression and classification tasks compared to PCA factors. RDA showed accuracy approaching 100% for certain algorithms and types of vegetable oils.

Nanou et al. (2025) examined the use of UV-Vis NIR spectroscopy to detect adulteration of EVOO with cheaper vegetable oils, including pomace, soybean, corn, and sunflower oils. The study analysed EVOO samples collected from four Greek regions recognised for producing premium olive oils: Achaia, Kalamata, Crete, and Lesvos. LDA, SVM, and LR were employed. The prediction accuracy of the three models ranged from 97.5 to 99.8%.

The reviewed studies demonstrate that NIR, coupled with chemometric methods, presents a rapid and practical approach for detecting adulteration in EVOO. Laboratory-grade FT-NIR systems and portable/handheld devices have shown accurate classification and quantitative prediction capabilities, frequently yielding high R^2 values and low prediction errors. However, a more critical assessment reveals that these performance metrics are highly sensitive to the sample design, experimental conditions, and validation strategies. Most studies rely on simplified experimental settings involving laboratory-prepared binary mixtures with controlled adulterant concentrations. While such models are useful for proof-of-concept investigations, they may simplify the discrimination task and lead to overestimation of predictive performance compared with real commercial samples. In contrast, authentication of commercial EVOO presents substantially greater analytical complexity due to natural compositional variability, multi-adulterant interference, and the chemical similarity between EVOO and lower-quality olive oils. Accordingly, studies evaluating commercial samples report generally lower performance under these more realistic conditions. Reported detection limits vary considerably depending on the type of adulterant, the concentration range, and the availability of authentic control samples. Studies incorporating authentic control samples tend to achieve lower detection limits. Furthermore, classification performance is affected by chemical similarity between EVOO and adulterants, especially in cases involving oils with comparable fatty acid profiles. Another important consideration is the spectral acquisition mode (i.e. reflectance vs. transmittance) and pre-processing strategies. The existing literature suggests that transmittance measurements and optimized pre-processing approaches may improve predictive capability and reduce irrelevant spectral variation.

However, these conditions are not consistently standardised across studies, complicating direct comparison. Although appropriate pre-processing frequently improves model robustness, systematic evaluation of pre-processing effects on model performance and transferability remains limited. Furthermore, validation strategies exhibit considerable heterogeneity. Several studies rely exclusively on internal validation approaches, which may lead to optimistic performance estimates. Some studies employ more robust external validation frameworks, which provide stronger evidence of model transferability. Notably, the highest reported performance metrics are frequently associated with internal or hybrid validation strategies. These findings highlight the importance of validation design in interpreting model performance.

Although NIR spectroscopy demonstrates considerable potential as a rapid screening tool for olive oil authentication, its routine implementation in official quality control frameworks requires broader validation across multi-source datasets, real-market samples, and standardised performance metrics.

NIR Spectroscopy Applications in Geographical and Botanical Origin Authentication of Olive Oil

Some studies have investigated the use of NIR spectroscopy in conjunction with chemometrics for the varietal and geographical authentication of EVOO (Table 2).

Laroussi-Mezghani et al. (2015) used chemometric methods to classify the varietal origins of Tunisian VOOs based on NIR spectra. The study encompassed six Tunisian varieties (Chemlali Sfax, Chetoui, Chemchali, Oueslati, Zarrazi, and Zalmati), Algeria, Morocco, and French VOOs. The varietal origin of olive oils was predicted using the SIMCA model as a supervised classification method, with Chemlali Sfax, Chetoui, and Oueslati varieties correctly classified at rates of 89.55%, 92.53%, and 98.50%, respectively. The study revealed that Tunisian VOOs can be distinguished from those of other origins, such as Algerian and Moroccan, using this method.

The determination of the geographical origin of Croatian EVOO using NIR spectra and employing PCA and PLS-DA was investigated (Peršurić et al. 2018). NIR spectra between 1000 and 2222 nm were found to be indicative of chemometric analysis. In the prediction set, the PLS-DA model constructed from NIR significant variables correctly classified all samples, except one.

Jiménez-Carvelo et al. (2019) investigated the use of NIR spectroscopy to authenticate the geographical origin of Argentinean EVOOs. The study encompassed EVOO samples from six Argentinean RDOs and unknown RDOs.

The PLS-DA model demonstrated excellent performance in determining the RDO of EVOO samples, achieving 100% classification accuracy.

The reviewed studies suggest that combining NIR spectroscopy with chemometric methods has considerable potential for distinguishing EVOO according to varietal and geographical origins. The reported classification rates exceed 90%. However, a critical evaluation indicates that reported performance is often closely linked to experimental design. Studies often rely on well-defined, geographically distinct, or well-characterised sample sets, which may simplify classification tasks relative to real-world conditions. Binary geographical classification generally represents a relatively simplified authentication scenario, whereas multi-varietal and multi-regional authentication involve substantially greater analytical complexity. These challenges increase in cases involving overlapping geographic origins or sample groups that exhibit similar chemical signatures. The chemical similarity between oils from neighbouring regions or closely related cultivars represents a significant challenge, as overlapping chemical characteristics may reduce discriminatory capability, potentially limiting model generalisability. Moreover, olive oil composition is strongly influenced by environmental factors and seasonal variability; reliance on single-harvest datasets may restrict model generalisability. Additionally, studies based on internal validation may yield overly optimistic performance estimates.

Consequently, NIR spectroscopy shows considerable promise for olive oil authentication; however, greater emphasis on multi-season datasets, chemically similar sample groups, and standardised performance metrics is required to ensure reproducibility and enable adoption within official traceability systems.

Raman Spectroscopy Applications in Detection of Olive Oil Adulteration

Several studies have evaluated the use of Raman spectroscopy in conjunction with chemometrics for detecting adulteration in EVOO (Table 3).

Yildiz Tiriyaki and Ayvaz (2017) investigated the use of portable Raman spectroscopy to quantify soybean oil adulteration in EVOO, with PLS-R applied to predict the adulteration levels. The PLS-R showed good performance in predicting the level of adulteration, as evidenced by a low standard error of prediction (SEP) of 1.34% and a high correlation coefficient of 0.99.

The application of Raman spectroscopy for detecting adulteration of Cretan EVOO with sunflower oil was investigated (Philippidis et al. 2017). Higher intensities were observed in the Raman bands at 1265 cm^{-1} and 1657 cm^{-1} for sunflower oil, corresponding to the stretching

Table 3 Application of Raman spectroscopy for detecting adulterants in olive oil or olive oil classification (cultivar and geographical origin)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Soybean oil	Adulterated samples ($n=38$) Laboratory-prepared mixtures	250–3200 cm^{-1}	PLS-R	Cross-validation; independent test set	PLS-R model yielded a high R^2 value (0.99) and a low SEP (1.34)	Yildiz Tiryaki and Ayvaz (2017)
Sunflower oil	Adulterated samples ($n=14$) Laboratory-prepared mixtures	100–3300 cm^{-1}	PLS-R	Train/test split; cross-validation	R^2 : 0.985 and RMSECV: 3.5%	Philippidis et al. (2017)
Waste cooking oils	Adulterated samples ($n=96$) Laboratory-prepared mixtures	100–3500 cm^{-1}	iPLS SiPLS	SPXY-based train/test split; cross-validation	SiPLS model achieved better results (R^2 : 0.982 and RMSEP: 0.485)	Li et al. (2018)
Soybean oil	EVOO ($n=8$) Blend of SO and EVOO ($n=4$) Vegetable oils ($n=4$) Adulterated samples ($n=19$) Laboratory-prepared mixtures and commercial samples	400–2300 cm^{-1}	PCA PLS iPLS	Cross-validation; external validation	iPLS achieved the best results (R^2 : 0.9940 and RMSEP: 1.9823%)	Barros et al. (2021)
Sunflower, pomace olive, and refined olive oils	Adulterated samples ($n=63$) Laboratory-prepared mixtures	400–2300 cm^{-1} (handheld) 350–2000 cm^{-1} (SORS system)	PCA PLS-R	Bootstrap resampling validation	R^2 : 0.90–1.00 and RMSEP: 1.4–10.2	Varnasseri et al. (2021)
Synthetic olive oil (sunflower seed oil blended with β -carotene)	EVOO ($n=4$) Sunflower seed oil ($n=1$) Synthetic OO ($n=9$)	800–3050 cm^{-1}	SVR	Train/test split	Linear kernel function achieved the best results (R^2 : 0.9868 and RMSE: 0.0653)	Fang et al. (2022)
Sunflower seed oil	Blended samples ($n=9$) Laboratory-prepared mixtures	200–3600 cm^{-1}	PLS-R SVM ANN RFR	Random train/test split	RFR model showed the best performance (R^2 : 0.9799 and RMSE: 0.0447)	Chen et al. (2025)
Origin/cultivar						
Varietal and geographical origin of Andalusian EVOO	Hojiblanca ($n=142$) Picual ($n=128$) Frantoio ($n=19$) Gordal ($n=18$) Mixtures of varieties ($n=105$)	100–3100 cm^{-1}	PCA LDA	Cross-validation	LDA achieved a high correct classification rate (≥ 84)	Sánchez-López et al. (2016)

Table 3 (continued)

Adulterant	Number of sample datasets	Spectroscopic information	Chemometrics	Validation approach	Main findings	Reference
Varietal origin of Greek EVOO	Koroneiki ($n=154$) Kolovi ($n=41$) Adramytiani ($n=8$) Commercial samples ($n=30$)	1000–1700 cm^{-1}	SVM RF	Cross-validation; external validation	RF combined with SVM achieved accuracy score of 94%	Drakopoulou et al. (2022)

vibrations of unsaturated fatty acids. The Raman region from 1000 to 1800 cm^{-1} was used for building the PLS model. The PLS model exhibited R^2 (0.985) and RMSECV (3.5%) values.

Li et al. (2018) investigated the detection and quantification of olive oil adulteration with waste cooking oil, collecting spectra of the samples using a portable Raman spectroscopy system. The quantitative analysis was conducted using the interval partial least squares (iPLS) and synergy interval partial least squares (SiPLS) methodologies. The spectral data processed with SNV proved to be optimal for SiPLS, thereby improving the model's accuracy. The wavelength ranges identified as optimal for the were 1159 cm^{-1} –1190 cm^{-1} , 1264 cm^{-1} –1295 cm^{-1} , 1433 cm^{-1} –1464 cm^{-1} , and 1627 cm^{-1} –658 cm^{-1} , respectively. The RMSEC and RMSEV values demonstrated good performance of calibration and validation models.

The application of portable Raman spectroscopy, employing PCA, PLS, and iPLS, was investigated for the identification and quantification of soybean oil adulteration in EVOO (Barros et al. 2021). The PLS model was built using the spectral region from 1165 to 1732 cm^{-1} , while the iPLS model was developed based on selected variables within the 1163–1732 cm^{-1} range. The iPLS method achieved the best analytical results.

Varnasseri et al. (2021) examined the detection of adulteration in EVOO with sunflower oil, pomace olive oil, and refined olive oil using portable conventional Raman and spatially offset Raman spectroscopy (SORS). The experimental setup employed two handheld conventional Raman spectrometers and a portable SORS instrument, with Raman spectroscopy performed in vial mode. Bootstrapping was used to validate the PLS-R models. The PCA plots demonstrated discernible trends associated with the level of adulteration across all three mixtures. The capability of Raman spectroscopy and SORS in detecting adulteration in EVOO was demonstrated, with low prediction errors and limits of detection. SORS could analyse samples through their original container with similar accuracy, making it a non-destructive method.

Vegetable oils are blended with β -carotene to imitate EVOO. Fang et al. (2022) developed a methodology using a laser confocal Raman technique to analyse and compare molecular differences between EVOO and sunflower seed oil. The methodology relied on theoretical Raman spectra of β -carotene, oleic acid, and linoleic acid, computed using density functional theory (DFT). The SVR was employed for quantitative analysis of β -carotene in synthetic olive oils. The identification of specific Raman bands (1161, 1265, 1526, and 1658 cm^{-1}) has been shown to serve as key markers for distinguishing between oleic acids, linoleic acids, and β -carotene ratios. A combination of the laser confocal Raman technique and theoretical Raman spectra based on

DFT proved to be an appropriate methodology for identifying low-cost imitation of olive oil.

A model was developed using Raman spectroscopy combined with regression techniques to quantify olive oil content in binary blends with sunflower oil (Chen et al. 2025). CARS was used for the purpose of selecting feature bands associated with β -carotene and unsaturated fatty acids. The Raman peaks at approximately 1265 and 1658 cm^{-1} were attributed to unsaturated fatty acids, while those at around 1156 and 1526 cm^{-1} corresponded to key feature bands of β -carotene. The performance of four regression models, namely SVM, PLS-R, RFR, and ANN, was evaluated. The RFR model demonstrated the best performance in predicting olive oil content, with RMSE of 0.0447 and R^2 of 0.9799. The study emphasised the importance of Raman feature peaks associated with β -carotene and unsaturated fatty acids in the prediction model.

The reviewed studies demonstrate that Raman spectroscopy, in conjunction with chemometric methods, shows considerable potential for detecting and quantifying adulteration in EVOO. The high coefficients of determination and low prediction errors indicate that Raman spectra can capture meaningful compositional differences, particularly those associated with unsaturated fatty acid content and β -carotene content. Nevertheless, a more critical analysis reveals that these performance metrics are dependent on sample design, experimental conditions, pre-processing techniques, and validation strategies. Many studies rely on laboratory-prepared binary mixtures with controlled adulterant concentrations, which may not adequately reflect the compositional complexity and variability of commercial olive oils. Most investigations represent simplified experimental settings under idealized analytical conditions that differ substantially from real-world commercial fraud scenarios. The limited number of studies evaluating more realistic authentication conditions suggests that multi-adulterant systems increase analytical complexity and reduce predictive robustness. The capability of Raman-based models is strongly influenced by the type of adulterant and its compositional similarity to EVOO. The available evidence indicates that detection limits depend on chemical similarity, with olive oil derivatives being the most difficult to detect. Furthermore, a notable trend in recent studies is the shift toward advanced variable selection methods (e.g. iPLS, SiPLS) and machine learning algorithms (SVM, RFR, ANN), which frequently outperform traditional PLS regression. However, direct comparisons among algorithms are not consistently conducted under standardised conditions, making it difficult to determine whether performance gains reflect genuine methodological improvement or dataset-specific optimisation. Several studies demonstrate that spectral pre-processing substantially influences model robustness and predictive capability. The available evidence suggests that optimal pre-processing strategies are highly

dependent on the chemometric method employed, whereas suboptimal pre-processing may reduce classification stability and transferability. Although individual studies report substantial improvements when appropriate pre-processing is applied, the absence of standardised pre-processing workflows remains a major limitation for inter-study reproducibility and method transferability. Another critical factor is the validation approach. The validation strategies are primarily based on internal validation methods. While these methods are valuable for assessing model robustness, they may lead to optimistic performance estimates, particularly when applied to small or homogeneous datasets. Notably, the consistently high predictive accuracies reported across many internally validated studies suggest potential optimism bias. These findings emphasise that reported performance metrics should be interpreted with caution. Additionally, the effects of fluorescence interference and instrumental variability, recognised challenges in Raman spectroscopy, may limit the transferability of models across laboratories or devices.

Consequently, Raman spectroscopy exhibits potential for a rapid screening tool for olive oil adulteration. However, broader validation with diverse samples, standardised pre-processing, and standardised performance metrics are required before routine implementation in official quality control systems.

Raman Spectroscopy Applications in Geographical and Botanical Origin Authentication of Olive Oil

Some studies have evaluated the use of Raman spectroscopy in conjunction with chemometrics for the varietal and geographical authentication of EVOO (Table 3).

Sánchez-López et al. (2016) investigated the authentication of Andalusian EVOOs using Raman spectroscopy. The study involved monovarietal EVOO from Frantoio, Hojiblanca, Picual, and Gordal cultivars, and mixtures of different varieties. LDA was used for the classification of samples based on harvest year, olive variety, geographical origin, and PDO. The LDA achieved good classification performance for qualitative variables: 94.3% for harvest year, 84.0% for olive variety, 89.0% for geographical origin, and 86.6% for PDO.

Drakopoulou et al. (2022) conducted a study for the authentication of Greek EVOOs. The study used SVM and RF to analyse data for variety authentication. The model-based feature selection process, which employed an RF algorithm in conjunction with an SVM for classification, achieved good classification performance. The selected features were primarily linked to the spectral regions at 1040–1087 cm^{-1} , 1122–1275 cm^{-1} , and the band at 1655 cm^{-1} . A total of 49 out of the 52 test samples were

accurately classified, thereby attaining an accuracy rate of 94%.

The reviewed studies suggest that Raman spectroscopy, combined with chemometric methods, has substantial potential for discriminating EVOOs according to varietal and geographical origins. Reported classification accuracy (> 85%) indicates that Raman spectral fingerprints capture compositional differences related primarily to fatty acid profiles and minor constituents. However, a critical evaluation highlights that model performance is strongly influenced by the sample design and experimental conditions. Studies based on monovarietal oils, well-defined geographical regions, or regulated designations (e.g. PDO) may facilitate classification by reducing intra-class variability and increasing inter-class compositional differences. In contrast, real-world authentication scenarios involving overlapping geographical regions or closely related cultivars present substantially greater analytical challenges due to increased compositional overlap. The available evidence suggests that compositional similarity is a major determinant of classification complexity, as geographically proximate regions and genetically related cultivars often exhibit highly similar chemical profiles, thereby reducing discriminatory capability.

Consequently, Raman spectroscopy shows considerable promise for the varietal and geographical authentication of EVOO. Nevertheless, larger and more diverse datasets with standardised performance metrics are required to establish its reliable application for routine traceability and regulatory contexts.

Applications of Multiple Vibrational Spectroscopic Techniques for Olive Oil Adulteration Detection and Geographical and Botanical Authentication

In some studies, two or three vibrational spectroscopy techniques have been employed to evaluate EVOO authenticity (Table 4).

Wójcicki et al. (2015) developed chemometric models using NIR and MIR spectroscopy to identify adulterants, including refined and mildly deodorised olive oils. The PLS and PCR algorithms were employed. The optimal model was achieved using both MSC and first-derivative pre-processing techniques in the region of 1500–900 cm^{-1} for the PLS and PCR models, based on the MIR data. The PLS model developed from the NIR data yielded the best results when employing MSC pre-processing of spectra in the region of 6150–4500 cm^{-1} . The optimal PCR model was obtained in the same region using both MSC and first-derivative pre-processing techniques. The study revealed that four models provided R^2 values above 0.96 and RMSEP values below 4.0.

The detection of soybean adulteration in EVOO was examined using FTIR and NIR spectroscopy (Meng et al. 2023). FTIR and Vis–NIR spectroscopy combined with PLS-DA demonstrated superior performance, achieving 100% classification accuracy. This finding indicated the capability of these spectroscopic techniques for identifying adulterated olive oils. Furthermore, the performance of excitation-emission matrix fluorescence spectroscopy was evaluated, revealing that FTIR and vis–NIR spectroscopy were more suitable for the rapid identification of adulterated olive oils. This finding was associated with the fact that these methods could provide a more extensive array of information and were comparatively uncomplicated to operate and construct models.

Adulteration of EVOO with edible oils was evaluated using FTIR and Raman spectroscopic techniques (Jiménez-Carvelo et al. 2017). Different classification algorithms (kNN, PLS-DA, OCPLS, SIMCA, and SVM-C) and a quantification algorithm (PLS-R) were applied using the pre-processed data. According to the PCA, the FTIR regions, 3100–2700 cm^{-1} , 1800–1600 cm^{-1} , and 1205–1080 cm^{-1} , were selected, and the Raman region, 950–650 cm^{-1} , was selected. For classification, the best results were achieved from PLS-DA models. The PLS-based Raman model yielded higher R^2 (0.93) and RMSEV (34.2) values than the FTIR-based model (0.86 and 17.6, respectively). Normally, the findings indicated that FTIR fingerprints provided more discriminative power than Raman fingerprints in both classification and regression tasks.

Malavi et al. (2023) evaluated the capability of FTIR and Raman spectroscopy for the detection of adulteration with cheaper edible oils. To classify and quantify olive oil adulteration, PLS-DA and PLS regression models were established. The FTIR-based PLS-DA model achieved a better correct classification rate (99.8%) than the Raman-based model (96.6%). The FTIR model demonstrated proficiency for the discrimination of EVOO adulterated with safflower, soybean, canola, and sesame oils, achieving a correct classification rate (CCR) exceeding 90%. In contrast, the Raman model exhibited accuracy for the discrimination of canola, sunflower, and sesame oils, with a CCR exceeding 90%. The overall PLS regression model based on Raman spectra yielded a higher prediction R^2 (0.92) and a lower RMSEP (1.96%) than the FTIR-based model ($R^2 = 0.77$; RMSEP = 3.26%), demonstrating its superior performance.

Mendes et al. (2015) examined the adulteration of EVOO with soybean oil using three vibrational spectroscopy methods. PLS regression was applied to the full raw spectra. The PLS models of three spectroscopic methods achieved high R^2 values (≥ 0.98) for both calibration and validation sets. The RMSECV and RMSEP of the PLS model based on MIR data (4.89 and 14.71) were higher than those obtained from the NIR model (1.76 and 1.71) and the Raman model (1.57

Table 4 Application of vibrational spectroscopy for detecting adulterants in olive oil or olive oil classification (cultivar and geographical origin)

Adulterant	Number of sample datasets	Technique	Chemometrics	Validation approach	Main finding	References
Mild deodorized and refined olive oils	Adulterated samples ($n = 60$) Laboratory-prepared mixtures	NIR (4000–12000 cm^{-1}) MIR (4000–500 cm^{-1})	PCA PLS PCR	Train/test split, cross-validation	PLS and PCR models achieved satisfactory results ($R^2 > 0.96$ and RMSEP < 4.0)	Wójcicki et al. (2015)
Soybean oil	Pure olive oil samples ($n = 30$) Adulterated samples ($n = 150$) Laboratory-prepared mixtures and pure olive oil samples	NIR (400–2500 nm) MIR (4000–400 cm^{-1})	PCA PLS-DA	Train/test split; cross-validation; independent test set	PLS-DA achieved a total recognition rate of 100% for both NIR and MIR data	Meng et al. (2023)
Hazelnut, peanut, canola, safflower, sunflower, corn, and soybean oils	Olive oil ($n = 67$) Edible oils ($n = 79$) Blended samples ($n = 30$) Laboratory-prepared mixtures and olive oils	FTIR (4000–550 cm^{-1}) Raman (200–3200 cm^{-1})	PLS-DA OCPLS kNN SVM-C SIMCA PLS-R	KS- based train/test split; independent test set	PLS-DA achieved better classification results PLS model based on Raman data yielded higher R^2 and RMSEP values than the model based on FTIR data	Jiménez-Carvelo et al. (2017)
Safflower, corn, soybean, canola, sunflower, and sesame oils	EVOO ($n = 13$) Vegetable oils ($n = 6$) Adulterated samples ($n = 42$) Laboratory-prepared mixtures, EVOOs, and vegetable oils	FTIR (4000–400 cm^{-1}) Raman (500–2000 cm^{-1})	PCA PLS-DA PLS-R	Repeated train/test split; cross-validation	PLS-DA achieved a high classification rate (> 96%) PLS-R model developed from Raman data achieved better results (R^2 : 0.92 and RMSEP: 1.96%)	Malavi et al. (2023)
Soybean oil	Vegetable oils ($n = 8$) Adulterated samples ($n = 60$) Laboratory-prepared mixtures	NIR (4000–12000 cm^{-1}) MIR (4000–350 cm^{-1}) Raman (50–3500 cm^{-1})	PLS-R	Cross-validation; independent test set	PLS models yielded a high R^2 value (> 0.98) PLS model based on MIR data exhibited a higher RMSEP	Mendes et al. (2015)
Origin/cultivar						
Varietal origin of French olive oils	Aglandau ($n = 63$) Caillolletier ($n = 29$) Olivière ($n = 30$) Picholine ($n = 35$) Salonenque ($n = 39$) Tanche ($n = 34$)	NIR (4500–10000 cm^{-1}) MIR (4000–600 cm^{-1})	PCA PLS-DA	Stratified train/test split, cross-validation	High-level models using a majority vote with a control chart decision rule achieved the highest AUC values (0.97–1.00)	Maléchaux et al. (2020)
Geographical origin of VOO	Apulia, Italy ($n = 57$) Tuscany, Italy ($n = 13$) Foreign countries ($n = 27$)	NIR (4000–12500 cm^{-1}) MIR (4000–651 cm^{-1})	PCA PLS-DA	KS-based train/test split; cross-validation	PLS-DA achieved high sensitivity and specificity values (> 0.90) in the test set for both NIR and MIR data	De Angelis et al. (2025)

and 1.61). The MIR technique showed a greater dispersion. This finding can be associated with two situations. Firstly, the skeletal vibration bands of lipid structures are mostly located in the 800–1300 cm^{-1} region. This is a relatively minor segment of the entire MIR spectrum, which extends from 400 to 4000 cm^{-1} . Secondly, the heterogeneity of the oil drop distribution on the KBr window surface can also have an influence on the observed results.

Maléchaux et al. (2020) examined the varietal discrimination of French EVOOs, employing NIR and MIR spectroscopy combined with the PLS-DA algorithm. The study encompassed EVOO samples from the Aglandau, Cailletier, Olivière, Picholine, Salonenque, and Tanche cultivars, collected over three harvest seasons. The NIR spectra between 6100 and 4500 cm^{-1} and the FTIR spectra between 1800 and 700 cm^{-1} were selected for data analysis. The study revealed that the PLS-DA models, whether based on individual MIR or NIR datasets, exhibited similar performance for most cultivars. Aglandau was better predicted using NIR data, while MIR data yielded better predictions for Olivière. PLS-DA models were developed using three fusion strategies: low-level, mid-level, and high-level. The high-level models provided the best overall results among fusion strategies. The low- and mid-level fusion approaches typically did not yield better results than individual models, suggesting that integrating NIR and MIR data does not always improve the performance of models.

De Angelis et al. (2025) used the NIR and IR spectroscopy to classify VOO samples according to their geographical origins. They collected authentic samples from Italy (Apulia and Tuscany) and foreign countries (Morocco, Jordan, Greece, Tunisia, and Spain). The PLS-DA yielded high sensitivity and specificity values, thereby indicating the method's reliability in differentiating geographical origins. The values of specificity and non-error rate confirmed the adequate performance of the models. In the context of NIR spectra, the combination of SNV, 1D, and GLSW provided the highest predictive performance with respect to sensitivity, specificity, and classification error. Conversely, in IR spectra, the most favourable results were achieved through using SNV + GLSW.

The reviewed studies demonstrate the analytical potential of vibrational spectroscopy techniques, whether applied individually or comparatively, in detecting olive oil adulteration and origin discrimination. The high R^2 values, low prediction errors, and near-perfect classification accuracies observed in the studies highlight the sensitivity of these approaches under controlled experimental conditions. However, a critical assessment reveals that these results are frequently influenced by differences in experimental design, pre-processing strategies, and validation protocols, complicating direct comparison between techniques. Many studies rely on relatively small sample sets and laboratory-prepared

mixtures or otherwise well-defined sample groups, which may simplify analytical tasks and lead to overestimation of model robustness. Most investigations focus on simplified binary adulteration systems, whereas only a limited number address realistic authentication challenges involving complex commercial fraud scenarios. This discrepancy represents a major limitation in translating laboratory-based findings into practical, real-world authentication applications. The available evidence also indicates that spectral pre-processing strongly influences authentication model performance. Although an optimal pre-processing strategy appears to be technique-specific, the literature lacks comprehensive guidelines for ensuring analytical robustness and transferability across different datasets and instrumental platforms. Furthermore, objective assessment of methodological superiority is often complicated by differences in pre-processing strategies, selected spectral regions, and validation protocols when making direct comparisons among spectroscopic techniques. Comparative studies do not always employ standardized validation frameworks, making it difficult to determine whether observed differences in performance reflect true methodological advantages or dataset-specific effects. In addition, most reviewed studies rely primarily on internal validation approaches, while externally validated models remain comparatively limited. Notably, the highest reported classification performances are frequently observed in internally validated models, highlighting the importance of rigorous external validation for establishing reliable model generalisability. Additionally, variability in instrumental configuration and spectral acquisition modes presents a further challenge for model transferability and inter-laboratory reproducibility. Consequently, harmonised validation protocols, multi-season and multi-source datasets, and standardised performance metrics are essential for developing robust, generalisable, and practically implementable spectroscopic authentication models.

A comparative evaluation of MIR, NIR, and Raman spectroscopy reveals that, although all three techniques demonstrate strong potential for olive oil authentication, their analytical performance and practical applicability differ substantially depending on the context of use.

MIR spectroscopy is widely reported to provide the highest classification accuracy and strongest predictive performance, primarily due to the presence of well-defined fundamental vibrational bands that are directly related to key chemical functional groups. However, MIR measurements are generally more sensitive to sample preparation, and laboratory-based FTIR systems may limit portability for in-field applications.

NIR spectroscopy offers significant advantages in terms of minimal sample preparation and portability, particularly with the increasing availability of handheld devices. These characteristics make NIR highly suitable for rapid screening

and on-site quality control. Nevertheless, NIR spectra consist mainly of broad overtone and combination bands, which are less chemically specific and can reduce interpretability. As a result, NIR-based models are often more dependent on advanced chemometric techniques and may exhibit reduced robustness when applied to chemically similar samples or complex matrices.

Raman spectroscopy provides complementary information to both MIR and NIR, with particular sensitivity to unsaturated fatty acids and minor constituents such as carotenoids. Its ability to analyse samples through transparent containers makes it advantageous for non-destructive and in situ analysis. However, practical limitations such as fluorescence interference and instrumental variability can affect spectral quality and model transferability.

Based on the reviewed evidence, technique selection should depend on both the authentication task and application context. MIR spectroscopy is most suitable for laboratory-based adulteration detection due to its high chemical specificity, whereas NIR is better suited for rapid, on-site screening because of its portability. Raman spectroscopy provides complementary information and is advantageous for non-destructive analysis and detecting subtle compositional differences. For more complex tasks, such as geographical or varietal authentication, combining techniques through data fusion may improve robustness and discriminative performance.

Authentication Tasks

The reviewed literature encompasses several distinct analytical objectives, including adulteration detection, geographical origin and cultivar discrimination, grade classification, and identification of blends with lower-quality oils. These tasks differ fundamentally in terms of chemical similarity, class overlap, and application context and should therefore not be treated as analytically equivalent.

Adulteration detection involving chemically distinct vegetable oils typically represents a lower-complexity classification problem due to pronounced differences in fatty acid composition and spectral profiles. As a result, this task generally entails lower classification complexity and may yield higher reported accuracy, particularly under controlled experimental conditions.

Geographical and botanical origin authentication involves oils with high chemical similarity influenced by shared cultivars, environmental conditions, and processing practices, resulting in greater class overlap and reduced spectral separability. Therefore, these tasks are inherently more challenging and sensitive to subtle compositional variations.

Grade classification and discrimination of lower-quality blends introduce intermediate levels of complexity,

depending on the extent of processing differences and compositional overlap between oil categories. In such cases, spectral differences may be more subtle and require more sophisticated chemometric modelling to achieve reliable discrimination.

Consequently, a structured framework that accounts for chemical similarity, class overlap, and practical complexity is essential for a meaningful interpretation of spectroscopic authentication studies in olive oil analysis.

Limitation of Study

Although this review provides a comprehensive overview of MIR, NIR, and Raman spectroscopy combined with chemometric techniques for olive oil authentication, several limitations should be acknowledged.

Firstly, the review is inherently dependent on the methodological quality and reporting standards of the included studies. A significant proportion of the reviewed studies report high classification accuracies; however, variations in sample size, experimental design, pre-processing strategies, and validation protocols limit direct comparability across studies. The absence of standardised reporting criteria for spectroscopic authentication studies complicates objective performance assessment across techniques.

Secondly, a significant proportion of the available literature relies on laboratory-prepared binary mixtures and relatively controlled datasets. Whilst these experimental designs are appropriate for proof-of-concept demonstrations, they may not fully represent the compositional variability encountered in commercial markets.

Finally, due to the heterogeneity of datasets across studies, including differences in harvest year, cultivar diversity, geographical coverage, and adulterant types, it remains difficult to draw definitive conclusions regarding the universal superiority of one technique over another. Performance differences may reflect dataset structure rather than intrinsic analytical capability.

Future Perspectives

The future research directions proposed in this section, derived from the methodological limitations identified in the reviewed literature, highlight key pathways for the development of olive oil authentication using vibrational spectroscopy.

The reviewed studies rely on laboratory-prepared binary mixtures, which cannot adequately represent the compositional complexity and variability of real-market samples. Future research should therefore prioritise the use of commercial oils and multi-adulterant systems, enabling the

development of models that better reflect real-world conditions and reduce the risk of overestimating analytical performance.

The heterogeneity of validation strategies across studies further limits the comparability and reliability of reported performances. Therefore, there is a clear need for harmonised validation protocols. Particular emphasis should be placed on external validation using samples collected from different batches, origins, or harvest seasons to enhance the robustness and comparability of reported results.

The development of large-scale datasets remains essential, particularly for geographical and origin authentication. Future research should prioritise large-scale datasets covering multiple harvest years and geographical regions, capturing natural variability associated with climatic conditions, agricultural practices, and processing techniques. Such datasets are essential for developing robust and transferable chemometric models.

While the application of machine learning algorithms is increasing, greater emphasis should be placed on interpretability, allowing clearer chemical insight into the spectral features leading to discrimination rather than focusing solely on predictive accuracy.

Data fusion strategies combining FTIR, NIR, and Raman spectroscopy offer a promising approach to addressing limitations arising from chemical similarity among samples. By integrating complementary spectral information, multimodal approaches may improve classification robustness, especially in complex authentication tasks.

The expansion of portable NIR and Raman devices facilitates real-time field authentication. However, models developed under controlled laboratory conditions may not be directly transferable to field applications. Future research should focus on instrument standardisation, effective calibration transfer strategies, and inter-device reproducibility to support regulatory implementation.

For spectroscopic techniques to be adopted in official control systems, validation in accordance with international standards and collaborative inter-laboratory trials will be essential. Accordingly, future research should progress beyond laboratory-scale feasibility assessments toward comprehensive regulatory validation and official acceptance.

Conclusion

The authentication of olive oil remains a critical challenge due to the increasing prevalence of adulteration and the complexity of its chemical matrix. Vibrational spectroscopy, encompassing MIR, NIR, and Raman spectroscopy techniques, in combination with chemometrics methods, has emerged as a powerful untargeted approach, offering a molecular fingerprint of the entire food matrix. This

combination facilitates rapid, non-destructive analysis with minimal sample preparation.

Current evidence suggests that these methodologies are not only effective for detecting adulteration but also show potential for discriminating olive oils by varietal and geographical origins. The integration of chemometric and machine learning algorithms has been demonstrated to enhance predictive accuracy and robustness; however, challenges remain regarding standardisation and inter-laboratory reproducibility.

In the future, the continued development of portable instrumentation, multimodal data fusion strategies, and machine learning algorithms may facilitate the extensive integration of vibrational spectroscopy into routine quality control practices. Nevertheless, the implementation of regulatory frameworks necessitates the harmonisation of validation protocols, the demonstration of inter-laboratory reproducibility, and the alignment with international standards. In such circumstances, vibrational spectroscopic methods may serve as valuable supplementary tools for olive oil authentication, thereby contributing to consumer protection, product integrity, and the enhancement of traceability systems.

Abbreviations ANN: Artificial neural network; ATR: Attenuated total reflectance; BOSS-PLS: Bootstrapping soft shrinkage-partial least squares; BPNN: Back-propagation neural network; CARS: Competitive adaptive reweighted sampling; CARS-PLS: Competitive adaptive reweighted sampling-partial least squares; CART: Classification and regression trees; DA: Discriminant analysis; 1D: First derivative; EVOO: Extra virgin olive oil; FTIR: Fourier transform infrared; FT-MIR: Fourier transform mid-infrared; FT-NIR: Fourier transform near-infrared; GLSW: Generalized least squares weighting; IR: Infrared; IVIR-PLS: Iteratively retaining informative variables-partial least squares; kNN: K-nearest neighbours; KS: Kennard–Stone; LDA: Linear discriminant analysis; LLE: Locally linear embedding; LR: Logistic regression; MCV-PLS: Monte Carlo uninformative variable elimination-partial least squares; MSC: Multiplicative scatter correction; MIR: Mid-infrared; NIR: Near-infrared; OCSVM: One-class support vector machine; OCPLS: One-class partial least squares; OCkNN: One-class k-nearest neighbours; OO: Olive oil; OPLS-DA: Orthogonal partial least squares-discriminant analysis; PCA: Principal component analysis; PCR: Principal component regression; PLS: Partial least squares; PLS-DA: Partial least squares-discriminant analysis; PLS-R: Partial least squares regression; iPLS: Interval partial least squares; SiPLS: Synergy interval partial least squares; PDO: Protected designation of origin; POO: Pomace olive oil; QDA: Quadratic discriminant analysis; R²: Coefficient of determination; RDA: Redundancy analysis; RDO: Registered designation of origin; RF: Random forest; RFR: Random forest regression; RMSE: Root mean square error; RMSEC: Root mean square error of calibration; RMSEV: Root mean square error of validation; RMSECV: Root mean square error of cross-validation; RMSEP: Root mean square error of prediction; ROO: Refined olive oil; SIMCA: Soft independent modelling of class analogies; SG: Savitzky-Golay; SLLE: Supervised locally linear embedding; SPXY: Sample Set Partitioning Based on Joint X–Y Distance; SNV: Standard normal variate; SVM: Support vector machine; SVM-R: Support vector machine regression; SVM-C: Support vector machine classification; SVR: Support vector regression; VOO: Virgin olive oil; UV–Vis-NIR: Ultraviolet-visible-near-infrared; Vis-NIR: Visible-Near-infrared

Author Contribution Solely one author designed the study.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Data Availability No datasets were generated or analysed during the current study.

Declarations

Ethical Approval and Consent to Participate Not applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Azizian H, Mossoba MM, Fardin-Kia AR et al (2015) Novel, rapid identification, and quantification of adulterants in extra virgin olive oil using near-infrared spectroscopy and chemometrics. *Lipids* 50:705–718. <https://doi.org/10.1007/s11745-015-4038-4>
- Bajoub A, Bendini A, Fernández-Gutiérrez A, Carrasco-Pancorbo A (2018) Olive oil authentication: a comparative analysis of regulatory frameworks with especial emphasis on quality and authenticity indices, and recent analytical techniques developed for their assessment. A review. *Crit Rev Food Sci Nutr* 58:832–857
- Barros IHAS, Paixão LS, Nascimento MHC et al (2021) Use of portable Raman spectroscopy in the quality control of extra virgin olive oil and adulterated compound oils. *Vib Spectrosc* 116:103299. <https://doi.org/10.1016/j.vibspec.2021.103299>
- Benlmaalam S, Chigr M, Hirri A, et al (2015) Characterization and discrimination of virgin olive oils from different Moroccan geographical areas using MIR spectroscopy coupled to chemometric methods. *International Journal of Science and Research* 4:2443–2248. <https://doi.org/10.13140/RG.2.1.4760.9443>
- Borghi FT, Santos PC, Santos FD et al (2020) Quantification and classification of vegetable oils in extra virgin olive oil samples using a portable near-infrared spectrometer associated with chemometrics. *Microchem J* 159:105544. <https://doi.org/10.1016/j.microc.2020.105544>
- Campanella B, Palleschi V, Legnaioli S (2021) Introduction to vibrational spectroscopies. *Chemtexts* 7:5. <https://doi.org/10.1007/s40828-020-00129-4>
- Chen Y, Yang Z, Zeng S et al (2025) Quantitative analysis of β -carotene and unsaturated fatty acids in blended olive oil via Raman spectroscopy combined with model prediction. *Food Chem* 470:142621. <https://doi.org/10.1016/j.foodchem.2024.142621>
- De Angelis D, Totaro MP, Caponio F et al (2025) Could the quality characteristics of virgin olive oil influence its geographical classification using untargeted infrared spectroscopy and chemometrics? *Food Chem X* 28:102548. <https://doi.org/10.1016/j.fochx.2025.102548>
- Ditler E, Luber S (2022) Vibrational spectroscopy by means of first-principles molecular dynamics simulations. *WIREs Comput Mol Sci* 12:e1605. <https://doi.org/10.1002/wcms.1605>
- Drakopoulou S, Orfanakis E, Karagiannaki I et al (2022) Comparative evaluation of different targeted and untargeted analytical approaches to assess Greek extra virgin olive oil quality and authentication. *Molecules* 27:1350. <https://doi.org/10.3390/molecules27041350>
- Fang P, Wang H, Wan X (2022) Olive oil authentication based on quantitative β -carotene Raman spectra detection. *Food Chem* 397:133763. <https://doi.org/10.1016/j.foodchem.2022.133763>
- Filoda PF, Fetter LF, Fornasier F et al (2019) Fast methodology for identification of olive oil adulterated with a mix of different vegetable oils. *Food Anal Methods* 12:293–304. <https://doi.org/10.1007/s12161-018-1360-5>
- Folli GS, Santos LP, Santos FD et al (2022) Food analysis by portable NIR spectrometer. *Food Chemistry Advances* 1:100074. <https://doi.org/10.1016/j.focha.2022.100074>
- Foscolou A, Critselis E, Panagiotakos D (2018) Olive oil consumption and human health: a narrative review. *Maturitas* 118:60–66. <https://doi.org/10.1016/j.maturitas.2018.10.013>
- Gagour J, Hallouch O, Asbbane A et al (2024) A review of recent progresses on olive oil chemical profiling, extraction technology, shelf-life, and quality control. *Chem Biodivers* 21:e202301697. <https://doi.org/10.1002/cbdv.202301697>
- Giordani S, Kassouf N, Zappi A et al (2024) Rapid and green discrimination of bovine milk according to fat content, thermal treatment, brand and manufacturer via colloidal fingerprinting. *Food Chem* 440:138206. <https://doi.org/10.1016/j.foodchem.2023.138206>
- Gómez-Caravaca AM, Maggio RM, Cerretani L (2016) Chemometric applications to assess quality and critical parameters of virgin and extra-virgin olive oil. A review. *Anal Chim Acta* 913:1–21. <https://doi.org/10.1016/j.aca.2016.01.025>
- Hirri A, Gammouh M, Gorfti A, Kzaiber F, Bassbasi M, Souhassou S, Balouki A, Oussama A (2015) The use of Fourier transform mid infrared (FT-MIR) spectroscopy for detection and estimation of extra virgin olive oil adulteration with old olive oil. *Sky J Food Sci* 4(5):60–66
- Hirri A, Bassbasi M, Platikanov S et al (2016) FTIR spectroscopy and PLS-DA classification and prediction of four commercial grade virgin olive oils from Morocco. *Food Anal Methods* 9:974–981. <https://doi.org/10.1007/s12161-015-0255-y>
- Hirri A, Bouchafray H, Zaraybya L et al (2023) Development of fast analytical method for the detection and quantification of Moroccan picholine extra virgin olive oil adulteration using MIR spectroscopy and chemometrics tools. *Curr Chem Lett* 12:579–586. <https://doi.org/10.5267/j.ccl.2023.2.005>
- Jiang H, Chen Q (2019) Determination of adulteration content in extra virgin olive oil using FT-NIR spectroscopy combined with the BOSS-PLS algorithm. *Molecules* 24:2134. <https://doi.org/10.3390/molecules24112134>
- Jiao Z, Song L, Zhang Y et al (2025) A comparative study of fluorescence hyperspectral imaging and FTIR spectroscopy combined with chemometrics for the detection of extra virgin olive oil adulteration. *J Food Meas Charact* 19:1761–1776. <https://doi.org/10.1007/s11694-024-03069-6>
- Jiménez-Carvelo AM, Osorio MT, Koidis A et al (2017) Chemometric classification and quantification of olive oil in blends with any edible vegetable oils using FTIR-ATR and Raman spectroscopy. *LWT* 86:174–184. <https://doi.org/10.1016/j.lwt.2017.07.050>
- Jiménez-Carvelo AM, Lozano VA, Olivieri AC (2019) Comparative chemometric analysis of fluorescence and near infrared spectroscopies for authenticity confirmation and geographical origin

- of Argentinean extra virgin olive oils. *Food Control* 96:22–28. <https://doi.org/10.1016/j.foodcont.2018.08.024>
- Kakouri E, Revelou PK, Kanakis C et al (2021) Authentication of the botanical and geographical origin and detection of adulteration of olive oil using gas chromatography, infrared and Raman spectroscopy techniques: a review. *Foods* 10:1565. <https://doi.org/10.3390/foods10071565>
- Karunathilaka SR, Kia ARF, Srigley C et al (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. *J Food Sci* 81:C2390–C2397. <https://doi.org/10.1111/1750-3841.13432>
- Laouni A, El Orche A, Elhamdaoui O et al (2023) A preliminary study on the potential of FT-IR spectroscopy and chemometrics for tracing the geographical origin of Moroccan virgin olive oils. *J AOAC Int* 106:804–812. <https://doi.org/10.1093/jaoacint/qsac136>
- Laouni A, Orche AE, Kacemi ME et al (2023) Simultaneous concurrent assessment of extra virgin olive oil adulteration via Fourier transform mid-infrared and UV-visible spectroscopy combined with partial least squares regression. *Methods and Objects of Chemical Analysis* 18:160–169. <https://doi.org/10.17721/moca.2023.160-169>
- Laroussi-Mezghani S, Vanloot P, Molinet J et al (2015) Authentication of Tunisian virgin olive oils by chemometric analysis of fatty acid compositions and NIR spectra. Comparison with Maghrebian and French virgin olive oils. *Food Chem* 173:122–132. <https://doi.org/10.1016/j.foodchem.2014.10.002>
- Li Y, Fang T, Zhu S et al (2018) Detection of olive oil adulteration with waste cooking oil via Raman spectroscopy combined with iPLS and SiPLS. *Spectrochim Acta A Mol Biomol Spectrosc* 189:37–43. <https://doi.org/10.1016/j.saa.2017.06.049>
- Li X, Zhang L, Zhang Y et al (2020) Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils. *Trends Food Sci Technol* 101:172–181. <https://doi.org/10.1016/j.tifs.2020.05.002>
- Lia F, Mangion MZ, Farrugia C (2021) Application of Fourier transform mid-infrared attenuated total reflectance (FT-MIR-ATR) for the authentication of Maltese extra virgin olive oil. *LA RIVISTA ITALIANA DELLE SOSTANZE GRASSE-VOL XCVIII-GENNAIO/MARZO*
- Malavi D, Nikkha A, Raes K, Van Haute S (2023) Hyperspectral imaging and chemometrics for authentication of extra virgin olive oil: a comparative approach with FTIR, UV-VIS, Raman, and GC-MS. *Foods* 12:429. <https://doi.org/10.3390/foods12030429>
- Maléchaux A, Le Dréau Y, Artaud J, Dupuy N (2020) Control chart and data fusion for varietal origin discrimination: application to olive oil. *Talanta* 217:121115. <https://doi.org/10.1016/j.talanta.2020.121115>
- Mendes TO, da Rocha RA, Porto BLS et al (2015) Quantification of extra-virgin olive oil adulteration with soybean oil: a comparative study of NIR, MIR, and Raman spectroscopy associated with chemometric approaches. *Food Anal Methods* 8:2339–2346. <https://doi.org/10.1007/s12161-015-0121-y>
- Meng X, Yin C, Yuan L et al (2023) Rapid detection of adulteration of olive oil with soybean oil combined with chemometrics by Fourier transform infrared, visible-near-infrared and excitation-emission matrix fluorescence spectroscopy: a comparative study. *Food Chem* 405:134828. <https://doi.org/10.1016/j.foodchem.2022.134828>
- Mousa MAA, Wang Y, Antora SA et al (2022) An overview of recent advances and applications of FT-IR spectroscopy for quality, authenticity, and adulteration detection in edible oils. *Crit Rev Food Sci Nutr* 62:8009–8027. <https://doi.org/10.1080/10408398.2021.1922872>
- Nanou E, Bekogianni M, Stamatoukos T, Couris S (2025) Detection of adulteration of extra virgin olive oil via laser-induced breakdown spectroscopy and ultraviolet-visible-near-infrared absorption spectroscopy: a comparative study. *Foods* 14:321. <https://doi.org/10.3390/foods14020321>
- Nurwahidah AT, Rumiati, Riyanto S, et al (2019) Fourier transform infrared spectroscopy (FTIR) coupled with multivariate calibration and discriminant analysis for authentication of extra virgin olive oil from rambutan seed fat. *Food Res* 3:727–733. [https://doi.org/10.26656/fr.2017.3\(6\).182](https://doi.org/10.26656/fr.2017.3(6).182)
- Ordoudi SA, Özdikicierler O, Tsimidou MZ (2022) Detection of ternary mixtures of virgin olive oil with canola, hazelnut or safflower oils via non-targeted ATR-FTIR fingerprinting and chemometrics. *Food Control* 142:109240. <https://doi.org/10.1016/j.foodcont.2022.109240>
- Ozen B, Cavdaroglu C, Tokatli F (2024) Trends in authentication of edible oils using vibrational spectroscopic techniques. *Anal Methods* 16:4216–4233. <https://doi.org/10.1039/d4ay00562g>
- Pan M, Sun S, Zhou Q, Chen J (2018) A simple and portable screening method for adulterated olive oils using the hand-held FTIR spectrometer and chemometrics tools. *J Food Sci* 83:1605–1612. <https://doi.org/10.1111/1750-3841.14190>
- Peršurić Ž, Saftić L, Mašek T, Kraljević Pavelić S (2018) Comparison of triacylglycerol analysis by MALDI-TOF/MS, fatty acid analysis by GC-MS and non-selective analysis by NIRS in combination with chemometrics for determination of extra virgin olive oil geographical origin. A case study. *LWT* 95:326–332. <https://doi.org/10.1016/j.lwt.2018.04.072>
- Philippidis A, Poulakis E, Papadaki A, Velegrakis M (2017) Comparative study using Raman and visible spectroscopy of Cretan extra virgin olive oil adulteration with sunflower oil. *Anal Lett* 50:1182–1195. <https://doi.org/10.1080/00032719.2016.1208212>
- Revelou PK, Pappa C, Kakouri E et al (2021) Discrimination of botanical origin of olive oil from selected Greek cultivars by SPME-GC-MS and ATR-FTIR spectroscopy combined with chemometrics. *J Sci Food Agric* 101:2994–3002. <https://doi.org/10.1002/jsfa.10932>
- Rodionova OY, Pomerantsev AL (2020) Chemometric tools for food fraud detection: the role of target class in non-targeted analysis. *Food Chem* 317:126448. <https://doi.org/10.1016/j.foodchem.2020.126448>
- Rohman A, Che Man Y, Ismail A, Hashim P (2017) FTIR spectroscopy coupled with chemometrics of multivariate calibration and discriminant analysis for authentication of extra virgin olive oil. *Int J Food Prop* 20:S1173–S1181. <https://doi.org/10.1080/10942912.2017.1336718>
- Sánchez-López E, Sánchez-Rodríguez MI, Marinas A et al (2016) Chemometric study of Andalusian extra virgin olive oils Raman spectra: qualitative and quantitative information. *Talanta* 156–157:180–190. <https://doi.org/10.1016/j.talanta.2016.05.014>
- Sánchez-Rodríguez MI, Sánchez-López E, Marinas A et al (2022) Redundancy analysis to reduce the high-dimensional near-infrared spectral information to improve the authentication of olive oil. *J Chem Inf Model* 62:4620–4628. <https://doi.org/10.1021/acs.jcim.2c00964>
- Scatigno C, Festa G (2022) FTIR coupled with machine learning to unveil spectroscopic benchmarks in the Italian EVOO. *Int J Food Sci Technol* 57:4156–4162. <https://doi.org/10.1111/ijfs.15735>
- Sun X, Lin W, Li X et al (2015) Detection and quantification of extra virgin olive oil adulteration with edible oils by FT-IR spectroscopy and chemometrics. *Anal Methods* 7:3939–3945. <https://doi.org/10.1039/c5ay00472a>
- Uncu O, Ozen B (2019) A comparative study of mid-infrared, UV-visible and fluorescence spectroscopy in combination with chemometrics for the detection of adulteration of fresh olive oils with old

- olive oils. *Food Control* 105:209–218. <https://doi.org/10.1016/j.foodcont.2019.06.013>
- 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 Qual Saf* 2:189–198. <https://doi.org/10.1093/fqsafe/fyy018>
- Varnasseri M, Muhamadali H, Xu Y et al (2021) Portable through bottle SORS for the authentication of extra virgin olive oil. *Appl Sci* 11:8347. <https://doi.org/10.3390/app11188347>
- Vasconcelos M, Coelho L, Barros A, de Almeida JMMM (2015) Study of adulteration of extra virgin olive oil with peanut oil using FTIR spectroscopy and chemometrics. *Cogent Food Agric* 1:1018695. <https://doi.org/10.1080/23311932.2015.1018695>
- Wang P, Sun J, Zhang T, Liu W (2016) Vibrational spectroscopic approaches for the quality evaluation and authentication of virgin olive oil. *Appl Spectrosc Rev* 51:763–790
- Wójcicki K, Khmelinskii I, Sikorski M et al (2015) Spectroscopic techniques and chemometrics in analysis of blends of extra virgin with refined and mild deodorized olive oils. *Eur J Lipid Sci Technol* 117:92–102. <https://doi.org/10.1002/ejlt.201300402>
- Yan J, van Stuijvenberg L, van Ruth SM (2019) Handheld near-infrared spectroscopy for distinction of extra virgin olive oil from other olive oil grades substantiated by compositional data. *Eur J Lipid Sci Technol* 121:1900031. <https://doi.org/10.1002/ejlt.201900031>
- Yildiz Tiryaki G, Ayvaz H (2017) Quantification of soybean oil adulteration in extra virgin olive oil using portable raman spectroscopy. *Food Measure* 11:523–529. <https://doi.org/10.1007/s11694-016-9419-8/>
- Zappi A, Marassi V, Kassouf N et al (2022) A green analytical method combined with chemometrics for traceability of tomato sauce based on colloidal and volatile fingerprinting. *Molecules* 27:5507. <https://doi.org/10.3390/molecules27175507>
- Zappi A, Marassi V, Giordani S et al (2023) Extracting information and enhancing the quality of separation data: a review on chemometrics-assisted analysis of volatile, soluble and colloidal samples. *Chemosensors* 11:45. <https://doi.org/10.3390/chemosensors11010045>
- Zaroual H, Chénè C, El Hadrami EM, 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. *Crit Rev Food Sci Nutr* 62:4526–4549

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.