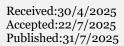# HALALSPHERE

International Islamic University Malaysia - INHART

Integration of spectroscopy and chemometric analysis for food authentication: A review

Katrul Nadia Basri

International Institute for Halal Research and Training (INHART), International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia.
Corresponding author: E-mail address: katrulnadia@iium.edu.my

Abstract

The high demand in the food sector necessitates an authentication process to verify products and prevent food fraud. In the food industry, applying spectroscopy techniques combined with chemometric analysis has become an efficient tool for authentication, thanks to its rapid results, cost-effectiveness, and reliability. Among the spectroscopy techniques reviewed in this paper are near-infrared (NIR), Fourier transform infrared (FTIR), ultraviolet (UV), and Raman spectroscopy. This paper reviews recent spectroscopy applications to food, such as oils, agricultural products, and beverages. This paper reviews various types of classification and regression algorithms in chemometric analysis to demonstrate the capabilities of these algorithms for use in food authentication processes. The chemometric model can accurately identify authentic samples with minimal risk of misclassification. This review presents a comparative synthesis of algorithms for various types of food samples, highlighting the performance of predictive algorithms. This review will provide a practical framework for researchers seeking to develop a robust predictive system for food authentication using spectroscopy and chemometrics. The review discusses the emerging trend of integrating spectroscopy into real-time authentication systems, particularly in the application of data fusion and deep learning techniques.

1. Introduction

Food authentication is a crucial issue for consumers due to religious motivations (Musfiroh *et al.*, 2025), economic reasons, and the assurance of the food's quality and safety. In Islam, followers are obliged to eat permissible and good food, per Islamic dietary laws, also known as halal, as stated in the *Qur'an*, verse 2:168. In addition to religious considerations, vegan consumers ensure that their food is not mixed with any animal-derived ingredients. Certification schemes on the Protected Designation of Origin (PDO) play a crucial role in assuring consumers of product authenticity to guarantee that the product is entirely produced, processed, and prepared in a specific geographical region (Candeias *et al.*, 2025; Mohammadi *et al.*, 2024; Özdemir *et al.*, 2024). Additionally, the authentication process ensures a trustworthy relationship between producers and consumers. The flow process of food authentication is as illustrated in Figure 1. After selecting the product, the sample will be prepared and analysed. The data obtained will be interpreted and verified by the regulatory body.

> *"O mankind, eat from whatever is on earth [that is] lawful (halal) and good (tayyib)"*
>
> *Qur'an* 2:168

Ensuring the authenticity of food products, such as verifying their geographic origin, species, and purity, is crucial in determining their halal status. Food authentication can help prevent food fraud, including adulteration, misbranding, or tampering. Adulteration involves the replacement of high-cost ingredients with lower-grade and cheaper substitutes. Adulteration cases in food products can lead to health problems for consumers (Aslam *et al.*, 2024; Dahimi *et al.*, 2014). The authentication process enables consumers to check for any forbidden substance, allergen or unsafe ingredients. Examples of analytical methods to detect cases of adulteration include chromatography, polymerase chain reaction (PCR) and spectroscopy techniques.

Food authentication may take some time, depending on the analytical method used. Thus, finding a rapid analysis method to save time and cost is crucial. In comparing chromatography or other molecular biology techniques with spectroscopy techniques, the latter method often requires more complicated procedures and extensive analysis. Besides that, some analytical methods, such as DNA extraction, require sample preparation, which might take some time to extract the result. The chromatography method does provide accurate results, but the instrument is expensive and humongous, and the technique is destructive to the sample.

Figure 1: The flow of the food authentication process.

Among the wide range of emerging technological solutions in food analysis and authentication, spectroscopic techniques have increasingly garnered substantial and noteworthy attention from researchers and industry. This is primarily due to their ability to deliver rapid, non-invasive, and non-destructive measurements while requiring minimal sample preparation. Various spectroscopic approaches, including but not limited to near-infrared (NIR), mid-infrared (MIR), Fourier-transform infrared (FTIR), Raman, and ultraviolet-visible (UV-Vis) spectroscopy, have found expanding and widespread applications across a broad spectrum of food matrices. Nevertheless, despite their many advantages, the spectral data produced by these techniques are often highly complex and frequently characterised by overlapping signals, which pose significant challenges in interpretation and therefore necessitate sophisticated and advanced chemometric analysis methods to extract meaningful information accurately.

It is precisely at this critical juncture that chemometric, which involves the comprehensive, and rigorous application of advanced mathematical modelling approaches combined with sophisticated statistical methodologies, specifically intended for the extraction, interpretation, and modelling of meaningful patterns, underlying structures, and complex interrelationships inherent within high-dimensional spectroscopic data, assumes an indispensable role within the overarching analytical framework that supports contemporary spectroscopic approaches to food authentication. Many machine learning algorithms have become powerful tools for handling nonlinearities and large datasets. Researchers are now equipped with enhanced capabilities for achieving significantly improved accuracy, reliability, and robustness in classifying food products, quantifying specific components or contaminants, and predicting authenticity markers or adulteration events across various food matrices and production systems.

Based on previous works, there is a lack of comprehensive reviews that explore how spectroscopy and chemometric analysis can be combined with analytical methods to authenticate food as the reference method. This review addresses these limitations by critically examining recent advancements in food authentication, focusing on matters related to halal status, including adulteration, product origin, and quality. It is also designed with the primary objective of thoroughly exploring and evaluating current scholarly studies and scientific investigations that specifically focus on integrating spectroscopy techniques and chemometric methods for authentication related to food safety and the detection of adulteration, with references to halal and religious compliance where applicable. The review aims to assess the application of chemometric methods in conjunction with spectroscopy. Furthermore, it highlights and discusses the emerging trends, methodological innovations, and thematic directions that characterise the current state of research in this area. It further identifies and discusses challenges and future spectroscopy and food authentication research directions.

2. Methodology

This comprehensive review was meticulously undertaken to thoroughly examine and analyse the existing research on utilising various spectroscopy techniques in conjunction with chemometric analysis for food authentication. A clearly defined and methodically structured approach was employed to guarantee that the literature coverage was extensive, representative, objective, and free from bias.

2.1 Literature search strategy

A literature search was conducted to identify relevant studies, involving queries of several primary and reputable scientific databases, including Scopus, Web of Science, PubMed, and ScienceDirect. The search was limited to peer-reviewed articles published within the time frame spanning from 2020 to 2025. A predefined set of keywords and Boolean operators effectively refined and focused the search results. A total of 97 publications were identified based on publication year and relevant keywords; however, only 48 were selected for review due to their close relevance to the focus of this paper, which includes adulteration, protection of origin, and quality control.

The keywords included:
- ("food authentication" OR "halal" OR "food fraud" OR "food adulteration") AND
- ("spectroscopy" OR "NIR" OR "FTIR" OR "Raman" OR "UV-Vis" OR "Raman") AND
- ("chemometric" OR "PCA" OR "PLS-DA" OR "SVM" OR "LDA" OR "machine learning")

In addition to the database searches, the reference lists of all initially selected articles were carefully reviewed and manually screened to identify any additional studies that may have been relevant but were not captured through the database search alone.

2.2 Inclusion and exclusion criteria

The selection of inclusion and exclusion criteria used in this study is carefully defined in Table 1.

2.2.1 Inclusion criteria

Studies specifically involved using spectroscopy-based techniques, combined with chemometric methods, to authenticate food products. Research applying this combined spectroscopy and chemometric approach to a diverse range of food matrices, including but not limited to dairy products, edible oils, spices, and various processed food items. Investigations that utilised multivariate statistical models or machine learning algorithms for tasks such as classification of food types, quantification of components, or detection of adulteration within the food samples. Review articles were included as data sources because they provide background information and current updates on the topic.

Table 1: The inclusion and exclusion criteria for the literature search strategy

Criteria	Inclusion	Exclusion
Sample studied	Food	Non-food
Food authentication	Adulteration, origin, and quality control	Use not for food authentication
Method	Spectroscopy method (NIR, RAMAN, FTIR, UV)	Conventional methods, such as DNA or chromatography
Chemometric	Classification and regression	Dimension reduction

2.2.2 Exclusion criteria

Studies that did not include any chemometric analysis in their methodology were excluded from consideration, as well as research focusing on applications of spectroscopy techniques unrelated to food authentication or involving non-food-related subjects. Additionally, conference abstracts, book chapters, and other forms of literature that were not peer-reviewed were excluded from the study to maintain a high standard of scientific rigour and reliability. This review did not include a formal quality assessment of the included studies, such as evaluations based on sample size or methodological rigour. Future research should incorporate a standardised quality appraisal to enhance the reliability and validity of the conclusions.

3. Results and discussion

3.1 Application of spectroscopy

3.1.1 Overview of spectroscopy method

Spectroscopy has been widely used to analyse the composition, adulteration, and quality of food authentication products. It is an analytical method that analyses the interaction between electromagnetic radiation and the sample studied. The different wavelengths of electromagnetic radiation emitted during the measurement will determine the method of spectroscopic measurement used. The advantages of spectroscopic measurement compared to other analytical tools include less labour-intensive and time-consuming methods and direct detection. This method can also identify and detect low levels of composition in analytes with minimal or no sample preparation. (de Freitas Oliveira *et al.*, 2025). The spectroscopic techniques commonly employed in food authentication include near-infrared (NIR), Fourier Transform Infrared (FTIR), ultraviolet (UV), and Raman spectroscopy. Each spectroscopic technique presents unique analytical strengths for detecting and characterising food composition and potential adulteration. The sample's physicochemical properties largely determine the method under investigation.

FTIR spectroscopy is a method that provides molecular fingerprinting of the sample by measuring the mid-infrared (IR) absorption band corresponding to a specific molecular bond and functional group. The wavelength range (refer to Table 2) of FTIR utilises a longer wavelength; thus, it will be more advantageous because more absorption bands can be investigated. The wavelength emission used in FTIR is less detrimental to the sample; thus, multiple analyses can be run on the same sample without introducing sample degradation or altering absorption. Several FTIR spectroscopy techniques include transmission, attenuated total reflectance, and photoacoustic spectroscopy. (Pasieczna-Patkowska *et al.*, 2025). The method of spectroscopy used is dependent on the sample studied and the objective of the analysis.

Given the lower molecular specificity of NIR, its application in detecting subtle adulterants such as lard may require enhanced chemometric modelling. By contrast, FTIR, with its ability to capture fundamental vibrations described above, provides more evident spectral fingerprints for such compounds. There are three modes of sampling in NIR spectroscopy: transmission, reflectance and transflectance, depending on the sample type. One of the advantages of NIR spectroscopy is its rapid measurement capability, but it requires a calibration method to be deployed in the food authentication process. UV spectroscopy is a method that measures the absorption of light by a sample in the UV range, from 200 nm to 400 nm. The electronic transition also occurs in UV spectroscopy, and it is primarily applied to samples containing molecules with functional groups that include conjugated double bonds, aromatic rings, or lone pairs of electrons. Compared to NIR, UV light has a lower penetration depth, which limits its applicability primarily to aqueous samples, liquids, and soft solids.

Raman spectroscopy is a method of measuring the scattering of laser light that occurs due to the excitation of molecular vibration, unlike the previous spectroscopy method based on absorption. There are two types of scattering: elastic (Rayleigh) and inelastic (Raman). In Raman scattering, the molecule is excited from its ground state to a higher state and then relaxes to another state. A minor portion of the scattered photons exchanges energy with the target molecule. The method of Raman spectroscopy depends on the change of the polarised molecule at the frequency at which scattered radiation occurs. This spectroscopy method is suitable for investigating molecular bonds, such as C-C, C=C, C-CH, and C-CH$_3$ (Kolašinac *et al.*, 2025). Raman spectroscopy also offers rapid analysis with high sensitivity, making it a practical analytical tool for food authentication (Li *et al.*, 2024). Table 2 shows the wavelength range for each spectroscopic method, along with the use and type of sample studied.

3.1.2 Example of spectroscopy application

NIR spectroscopy is suitable for rapid quality evaluation due to its portability for detecting adulteration and identifying food origin. The application of NIR spectroscopy in detecting the quality index of ginger to identify its origin is exemplified. (Chen *et al.*, 2024). In NIR spectra, the peak at 1205 nm was attributed to the second overtone of the C-H bond in ginger powder, and the peak near 1460 nm corresponded to a signal from C-H stretching, the second overtone of O-H, and the radical features of gingerols. UV-Visible (UV-Vis) spectroscopy is applied to determine sugar adulteration in Mediterranean honey and investigate the molecular information of a sugar and honey mixture. The absorptions in the range of 250 nm to 350 nm are related to glucose, fructose, amino acids, proteins, and phenolic compounds, as shown by Dimakopoulou-Papazoglou *et al.* (023). Additionally, UV spectra were measured and analysed to differentiate between species of Javanese turmeric

Table 2: The details of the FTIR, NIR, UV AND Raman spectroscopy methods

Technique	Wavelength	Used for	Sample Types	Sensitivity	Portability	Drawback
FTIR	2500–25,000 nm	Functional group, adulterant	Powders, solids, liquids	High due to a sharp and well-defined peak	Less portable due to the bulky instrument	Strong water absorption requires sample contact
NIR	780–2500 nm	Moisture, fat, protein, sugar	Grains, dairy, liquids	Moderate due to broad and overlapping peaks	Yes, because handheld instruments are available	Require calibration and are sensitive to sample texture
UV	200–400 nm	Coloured compounds, polyphenols	Clear liquids, beverages, and extracts	High due to a sharp peak	Yes, because miniature equipment is available	Limited sample and cannot identify structural components
Raman	400–1000 nm	Pigments, lipid or adulterant	Aqueous samples, liquids, soft solids	Highly sensitive using surface-enhanced Raman spectroscopy	Yes, if it is integrated with fibre probes	Fluorescence interference, and sometimes, the signal intensity is weak

Table 3: Examples of recent research using the spectroscopy method for food authentication

Title	Sample	Method of detection	Performance / Accuracy	Relevance to the halal sector
Authentication of Mentha arvensis essential oil using attenuated total reflectance-Fourier transform infrared spectrophotometry coupled with chemometric (Jayasekher et al., 2024a)	Mentha arvensis essential oil	Attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectrophotometry. The wavenumber range of 500-4000 cm-1 was used.	The ATR-FTIR used to authenticate Mentha arvensis Essential oils produced a prediction model of 100% accuracy.	Adulteration in essential oil may undermine transparency and consumer trust.
Moringa oleifera seed oils: Physico-chemical characterisation and their authentication using FTIR spectroscopy and chemometrics (Irnawati et al., 2024)	Moringa oleifera seed oils	FTIR spectroscopy method that utilised the mid-infrared region between 4000-6000 cm^{-1}.	The coefficient determination result obtained was 0.9971 for quantifying the presence of oil adulterants.	Identifying oil adulteration is essential to avoid quality degradation and mislabeling.
Portable NIR spectrometer and chemometric tools for predicting quality attributes and adulteration levels in butteroil (da Silva Medeiros et al., 2023a)	Butteroil	NIR spectroscopy using DLP NIRscan ranges from 900 to 1700.	The sensitivity in detecting adulterants obtained was more than 91.5%.	Adulteration of non-halal ingredients such as lard is not permissible in Islamic dietary law.
Simple analytical method using ultraviolet spectral dataset and chemometrics for the authentication of Indonesian ground roasted coffee with different botanical and geographical indications (Suhandy et al., 2023)	coffee	A UV-visible spectrometer in the range 190 to 399 nm was used for the measurement.	The prediction model for sample authentication achieved 100% accuracy.	Geographical origin may affect halal claims, particularly about contaminants.
Adulteration detection of multi-species vegetable oils in camellia oil using Raman spectroscopy: Comparison of chemometric and deep learning methods (Wang et al., 2025)	Camellia oil	Raman spectroscopy	A Raman spectrometer with a 785 nm laser was used to measure the oil samples. The deep learning algorithms deployed have achieved an R^2 of 0.999 and an RMSE of 0.9.	Adulteration detection to ensure product quality and safety for consumer consumption.

(Temulawak), which is used in medicine and supplements (Kusumadewi *et al.*, 2022). Raman spectra were used to analyse the chemical properties of cocoa butter, cupuassu butter, passion fruit oil, and sacha inchi oil and their blend with cocoa butter used in chocolate manufacturing (Balcázar-Zumaeta *et al.*, 2025). All these methods of spectroscopic measurement discussed provide minimal to no sample preparation, are non-destructive, and produce rapid results. Table 3 reviews a few previous experimental works conducted by researchers and food analysts in the food industry using different spectroscopy methods.

NIR spectroscopy and chemometrics have been used to differentiate among edible oils and oilseeds (Li *et al.*, 2020). NIR spectroscopy has also been used to discriminate between the spectra of extra-virgin olive oil and virgin olive oil, and the result obtained using Partial Least Squares (PLS) density modelling has a correct classification rate of 82.35% (Garrido-Cuevas *et al.*, 2024; Rodionova *et al.*, 2024). Isdhiyanti *et al.* (2024) have utilised fingerprints unique in the NIR spectral range to monitor the quality change of cooking oil. The resulting NIR spectral display peaks, originating from C-H, CH2, and CH3 bonds, appear at wavelengths of 1730 nm and 2350 nm due to stretching and deformation. The peroxide number and free fatty acid are the chemical contents analysed using NIR spectroscopy (Isdhiyanti *et al.*, 2024).

The application of spectroscopy in halal food authentication primarily focuses on detecting non-halal ingredients, such as pork and its derivatives, and verifying authenticity to prevent false claims regarding origin or composition. The review work done by Claude Mona Airin *et. al.* discussed the application of FTIR to determine non-halal ingredients in meat products. Based on the review, the limit of detection of the FTIR method to detect non-halal meat in the range 1 % to 5 % (w/w) (*Airin et al.*, 2025). Besides halal concerns in meat products, the substitute for pork fat, known as lard, has been adulterated with butter to reduce production costs. The use of NIR spectroscopy to detect the presence of lard in butter oil, as tabulated in Table 3, demonstrates the method's capability to differentiate between halal and non-halal ingredients in a mixture. The method of Raman spectroscopy has been utilised to detect lard in the butter, and the result of R^2 is 0.99 (Taylan *et al.*, 2020). This result indicates that the Raman method is more suitable than NIR for detecting lard in butter, particularly due to the lipid compounds listed in Table 2.

The assurance of product authenticity represents a fundamental aspect of food safety management, extending across all supply chain segments from raw material procurement to final consumer delivery. In the context of halal compliance, this requirement becomes even more critical, as it necessitates stringent oversight by producers, suppliers, retailers, and regulators to ensure that the integrity of halal status is preserved throughout the production and distribution processes. A few studies have been summarised in Table 3 to highlight the application of various spectroscopy methods in the context of halal authentication.

Jiajun Zan *et al.*'s current work combines the methods of NIR and Raman spectroscopy and performs spectral analysis to detect tea seed oil adulteration. (Zan *et al.*, 2024). The feature layer fusion strategy was used to enhance the comprehensiveness of information usage. The accuracy obtained after the analysis reached more than 90% for calibration and prediction datasets. Motahari and coworkers characterise edible and non-edible oils using spectroscopic

methods, including Raman, FTIR, and UV-Vis spectroscopy. (Motahari *et al.*, 2023). The research employed a variety of spectroscopic methods to conduct a thorough investigation of the molecular fingerprint of the oil sample. The broad UV-Vis absorption peak observed around 370 nm signifies $\pi \rightarrow \pi*$ transitions within the carbon double bonds of the oil. A 270–300 nm peak indicates $n \rightarrow \pi*$ transitions from carbonyl groups (C=O). FTIR spectra exhibit peaks at specific wavelengths corresponding to molecular bonds, including C=O, C–O, -CH2, H–O–H, and C–OH. Raman peaks occur due to the stretching of molecular bonds, including C–H, CH2, CH3, and N–H.

UV-Vis and FTIR spectroscopy have been deployed by Cagri Cavdaroglu *et. al.* to detect adulteration using spirit vinegar and acetic acid. Although the method of data fusion from both methods did not improve classification results, the percentage of correct classification and sensitivity is more than 90% (Cavdaroglu & Ozen, 2022). A more accurate interpretation of spectra and correlation with the reference method requires chemometric analysis to extract information due to the low sensitivity and weak intensity of spectral data. (Motahari *et al.*, 2023). Specific spectra lack corresponding reference values, rendering them unlabelled data. In such instances, clustering algorithms can partition the spectra into subsets based on feature similarity or distance metrics. The multi-source data fusion technique was employed to integrate visible/NIR spectroscopy with machine vision data to predict the soluble solids content of oranges. The convolutional neural network model exhibited a 36.4% improvement in prediction RMSE. (Sun *et al.*, 2025). Data from three spectroscopic methods, NIR, mid-IR, and Raman, were combined to classify different milk types. The performance of prediction accuracy has increased from 85.71% to 95% (Mohammadi *et al.*, 2024).

3.2 Chemometric

Spectroscopic techniques, such as near-infrared (NIR), mid-infrared (MIR), Raman, and UV-Vis spectroscopy, generate detailed spectral fingerprints that reflect food samples' molecular composition and structure. However, these raw spectra are often complex and challenging to interpret due to overlapping peaks and background noise. Chemometric algorithms address this challenge by modelling the relationship between spectral features (independent variables) and reference values such as chemical concentrations or physical properties (dependent variables).

Chemometrics is a branch of science that relates measurements made on a chemical system or process to the system's state by applying mathematical or statistical methods (Kusumadewi *et al.*, 2022). Chemometrics is also a multivariate data analysis technique applied in spectroscopy. Currently, most researchers focus on data mining numerous spectral datasets. Chemometric algorithms implemented in spectroscopy analysis extract large amounts of hidden information from the spectral data. The recent advancements in the chemometric field have raised hope for overcoming the challenges in food authentication. The combination of chemometric techniques has helped extend the application of spectroscopy in many areas, such as halal authentication (Musfiroh *et al.*, 2025).

In this paper, the review of chemometric techniques is classified into classification and regression analysis. Classification analysis, also known as pattern recognition, uses information about the class membership of the samples and classifies new unknown samples in one of the known classes based on their measurement pattern. Regression is a

quantitative analysis that involves correlating the input variables (spectra) with the continuous value of the numerical output (target). The performance measures for classification and regression analysis differ. In classification, the algorithm's performance was evaluated based on accuracy, sensitivity, precision or specificity. Error-based matrices were used in regression analysis, including root mean square error (RMSE), mean absolute error (MAE), and mean squared error (MSE). R2, computed in regression, indicates how well the independent variables explain the variability of the target variable, with values closer to 1 showing better model fit.

3.2.1 Classification analysis

3.2.1.1 Principal component analysis (PCA)

One of the most employed techniques for classification analysis is principal component analysis (PCA). PCA is a technique used to classify two or more unknown groups. PCA is an unsupervised pattern recognition method with the simplest eigenvalue-based multivariate analysis. The PCA algorithm transforms the data and creates a new space based on the principal components, which can be used for the dimension reduction step. (Aslam *et al.*, 2024; Mohammadi *et al.*, 2024). The PCA algorithm decomposes the spectral data using either eigenvectors and eigenvalues or singular value decomposition (SVD). The eigenvectors and eigenvalues were calculated from the covariance matrix of the spectra, indicating the variables' variance relative to one another. The decomposition output produces principal components used as input to train the model. The optimal number of principal components depends on the model's performance. Rui Zhu *et al.*, thoroughly explained the details of the PCA method's process in their work on compressing spectral data (Zhu *et al.*, 2024).

Joanna Banas *et al.* carried out an investigation using PCA to characterise selected herb honeys, such as raspberry, lemon balm, rose, mint, black currant, instant coffee, pine, hawthorn, and nettle. (Banaś & Banaś, 2024). The first three principal components resulting from PCA are used to examine the natural pattern among samples and their clustering. The total variance in the data of the three PC was 97.6% variation. The first PC shows that the signals from all the fluorescent compounds in the herb honeys contribute nearly equally. These results suggest extracting and training the pro-health properties based on the relationship between individual polyphenols, pigments, vitamin B3, and vitamin B9. The classification of Lampung robusta coffee based on fermentation using PCA algorithms has a cumulative percentage of variance of 98% based on two principal components. (Yulia & Suhandy, 2020).

3.2.1.2 Linear discriminant analysis (LDA)

The linear discriminant analysis (LDA) algorithm distinguishes samples based on class differences. The computation of the scatter matrix, both within and between classes, based on distance, is used to calculate eigenvectors and eigenvalues. The objective of the computation in LDA is to maximise variance between classes and minimise variance within classes. LDA, combined with the dimension reduction method, has demonstrated improvements in classification methods. LDA was used to discriminate peanut kernel samples based on their origin from different cities (Zhu *et al.*, 2024). Combining preprocessing techniques and an ensemble method with LDA has achieved 100% classification accuracy. The results have shown that the correct combination of chemometric processes applied to the dataset can produce a robust and reliable model. The combination of PCA and LDA is also used to solve the problem of high-dimensional data. The LDA algorithm was combined with PCA to differentiate various adulterants in Mentha arvensis essential oil, where the principal components derived from PCA were used to train the LDA model. The PCA-LDA model achieved 100% accuracy, indicating a highly effective performance. (Jayasekher *et al.*, 2024b). Besides that, LDA algorithms have been used as spectral dimension reduction to extract information. The extracted features or the linear discriminant vector were utilised as the input for the classification model. (Wu *et al.*, 2024).

3.2.1.3 K-nearest neighbour (KNN)

The k-nearest neighbour (KNN) algorithm calculates the K nearest neighbours using distance metrics such as Euclidean, Mahalanobis, or cosine distance to perform classification analysis. (Suyal & Goyal, 2022). The category with the most occurrences among the *K* closest data points determines the classification of a new data point. The application of KNN with different preprocessing methods achieved 95.2% accuracy for the test Raman spectra dataset of Pericarpium citri reticulatae (PCR) (Yang *et al.*, 2025).

NIR spectroscopy and the KNN algorithm were utilised to determine the quality of mei tea by classifying its grade (Wu *et al.*, 2024). The sample size and the parameter, K, set will determine the accuracy of the classification model. Wu and his coworker set the value of the parameter K to 1, 3, 5, 7, and 9. The highest accuracy was achieved with K values of 1 and 3, yielding an accuracy of 44% (Wu *et al.*, 2024). KNN has also been employed in a classification method to categorise peanut kernels using NIR spectra (Zhu *et al.*, 2024). The combination of maximum LDA to project the dataset and KNN enables improvement in classification accuracy from 62.96% to 70.99%. Based on previous works, it has been proven that selecting the optimal value of K is crucial for achieving the maximum accuracy.

3.2.1.4 Support vector machine (SVM)

The fundamental principle of the support vector machine (SVM) is that it finds the optimal hyperplane by dividing the dataset into different classes and maximising the margin of separation between these classes. The hyperplane was defined based on Equation 1, where w is the adjustable weight vector, x is the input vector, and b is the bias. (Mohammadi *et al.*, 2024; Yang *et al.*, 2025).

$$w^T x + b = 0 \qquad \text{Equation 1}$$

SVM algorithms were utilised to determine the optimal qualitative analysis model using the sample prediction accuracy to identify the adulterant in Tartary buckwheat samples mixed with whole wheat flour, oat flour, soybean flour, barley flour, and sorghum flour using NIR spectroscopy, and the accuracy reported was 100% (Wu *et al.*, 2024).

The SVM model achieved a sensitivity and specificity of 94.0% and 98.6% in authenticating tea samples based on their geographical origins. (Mohammadi *et al.*, 2024). The SVM algorithm's optimisation was tuned based on the parameters c (cost), epsilon, and gamma, in which optimisation can be performed using a genetic algorithm. (Mohammadi *et al.*, 2024). The output from the PCA, after dimension reduction, was used as input to an SVM using a radial basis function to

identify the presence of adulterant butteroil. (da Silva Medeiros *et al.*, 2023b). The SVM classification result has an accuracy of 98.46%, indicating that the model can identify pure samples. (da Silva Medeiros *et al.*, 2023b). SVM can solve nonlinear problems by employing a nonlinear kernel function to map input data into a high-dimensional feature space, wherein linear classification can be performed.

3.2.2 Regression analysis

3.2.2.1 Partial least squares (PLS)

Partial least squares (PLS) regression represents a multivariate regression method that uses the least squares algorithm to establish a relationship between variables. In PLS regression, the fundamental principle relies on latent variables, which are unobserved variables derived from the original predictors. The method decomposes the spectral data matrix, denoted as X, into a set of these latent variables or components. This decomposition process is carefully guided by the variation in the property values, represented as **y**, which are interesting in the analysis. The relationship between X and Y is governed by a functional relationship, f, as illustrated in Equation 2.

$$y = f(X)$$
<div align="right">Equation 2</div>

The co-variation between X and y is maximised to extract the variation in X directly correlated with y. Applying NIR can explain this by representing the spectra in the space of NIR wavelengths to show a linear correlation based on combinations of wavelengths, called factors, where the studied property is best described.

Balcázar-Zumaeta *et al.* employed PLS in the chemometric analysis, which was embedded with Raman spectroscopy, to quantify the concentrations of cupuassu butter, passion fruit oil, and sacha inchi oil blended with cocoa butter to make chocolate. (Balcázar-Zumaeta *et al.*, 2025). Three PLS models were developed for mixing cocoa butter with cupuassu butter, passion fruit oil, and sacha inchi oil, and the R^2 prediction obtained for each model was 0.749, 0.85, and 0.62. Although the results of some models are not above 0.80, with proper model tuning and optimisation, the model's performance can be improved.

3.2.2.2 Principal component regression (PCR)

Principal component regression (PCR) constitutes a multivariate regression methodology that integrates the frameworks of multiple linear regression (MLR) and principal component analysis (PCA) to address challenges inherent in high-dimensional spectral datasets. Spectroscopic data often exhibit substantial multicollinearity and dimensional redundancy, complicating conventional regression modelling. PCR circumvents these issues by applying PCA to the original set of correlated spectral variables, thereby projecting them onto a reduced-dimensional subspace defined by orthogonal principal components. These principal components, which are linear combinations of the original variables, retain the maximal variance structure of the data while eliminating inter-variable correlations. Subsequently, MLR is performed on this transformed set of uncorrelated components, facilitating more stable and interpretable regression modelling. This dimensionality reduction mitigates the adverse effects of multicollinearity and enhances model robustness and generalisation capacity in spectroscopic analysis. The number of principal components to be used in the PCR model is one of

the parameters that must be optimised by varying the number of principal components and computing the prediction error. (Biancolillo *et al.*, 2020). A PCR model was developed using the FTIR spectrum to detect adulteration in Mentha arvensis essential oil. (Jayasekher *et al.*, 2024b). The correlation coefficient (R^2) and RMSE for prediction were 0.9953 and 0.01687, respectively. The results demonstrate that the PCR prediction model is robust, as evidenced by its high correlation and low error rate when applied to the prediction dataset.

3.2.2.3 Support vector regression (SVR)

Support vector regression (SVR) is a supervised machine learning algorithm that seeks to construct a function capable of approximating the underlying relationship between spectral features and their corresponding target values while incorporating a specified margin of tolerance, commonly called the "ε-insensitive zone." This ε-insensitive loss function enables the model to tolerate minor deviations between predicted and actual values, thereby focusing on capturing the general trend of the data rather than precisely fitting every observation. A notable advantage of SVR is its robustness when dealing with relatively small datasets, as it minimises a regularised loss function that effectively balances the trade-off between predictive accuracy and model complexity. This balance is crucial in preventing overfitting, especially when data availability is limited, which is often the case in many applied spectroscopy and photonics applications.

Moreover, SVR demonstrates flexibility in handling both linear and nonlinear regression problems by applying kernel functions, which enable the mapping of input data into higher-dimensional feature spaces where complex patterns may be more readily identified. Like its counterpart, Support Vector Machines (SVM), which is widely used for classification tasks, SVR relies on carefully tuning several key hyperparameters to optimise its performance. These hyperparameters include the choice of kernel type, the regularisation parameter (which controls the trade-off between the flatness of the function and the allowed deviations), and the epsilon parameter that defines the width of the ε-insensitive margin. Hyperparameter optimisation is a critical step in model development, and methods such as grid search are commonly employed to explore various parameter combinations to identify the optimal configuration that yields the best generalisation performance.

The SVR algorithm has been applied in food authentication to quantify adulterants in coriander powder using FTIR spectroscopy. The R^2 obtained for the validation was 0.8289, and the RMSE was 0.1113 (Goyal *et al.*, 2025). SVR algorithms trained to detect mycotoxins in food samples have achieved an accuracy of 0.995, indicating that they can be a valuable tool for regulatory compliance and food safety monitoring (Tarcan & Küplülü, 2024; Usman *et al.*, 2025). The prediction model developed based on SVR algorithms to determine fructose, glucose, and sucrose concentrations in high-fructose syrup has yielded R^2 values of 0.8491 for fructose, 0.7812 for glucose, and 0.8871 for sucrose. (Erinawati *et al.*, 2025)Based on the nonlinear SVR model, the developed model offers a rapid and robust approach, making it suitable for routine monitoring.

3.2.2.4 Artificial neural network (ANN)

Artificial neural networks (ANNs) constitute a widely adopted machine learning technique in chemometric analysis. It has demonstrated considerable reliability and effectiveness, particularly in performing regression tasks such as adulterant

quantification, compositional analysis, and estimating food product quality. This modelling approach is especially advantageous in the context of spectroscopy. This field frequently generates complex, high-dimensional, and inherently noisy datasets, often presenting significant challenges for conventional analytical methods. ANN's capacity to model nonlinear data structures enables it to capture intricate patterns and interactions within the data that simpler, linear models may overlook. Nevertheless, to ensure the development of a robust and generalisable model, careful tuning of the network's hyperparameters is crucial to overcome the risk of overfitting, wherein the model excessively adapts to the noise in the training data rather than capturing the authentic underlying relationships.

The flexibility of ANN allows its application in classification and regression analyses, depending mainly on selecting the activation function employed within the network architecture. Structurally, an ANN is composed of three fundamental layers: the input layer, which serves to receive and introduce the raw data into the network; one or more hidden layers, where complex transformations and feature extractions are performed; and the output layer, which delivers the final predictive or classification outcomes. Due to its interconnected structure, an ANN typically requires a larger dataset than classical algorithms. This requirement arises from the need to iteratively adjust and optimise the numerous connection weights between neurons during the training process, allowing the network to effectively learn and model the complex relationships between the input features and the corresponding output variables. The research to predict adulteration in coriander powder utilised FTIR spectra to train an ANN with an input layer of 64 neurons, a hidden layer of 32 neurons, and an output layer of 1. R^2 of validation data for the ANN model was 0.9351, and the root mean square error (RMSE) was 0.0753 (Goyal *et al.*, 2025). The ANN model's high performance is due to its ability to calculate nonlinear relationships and complex relations between features. A convolutional neural network (CNN) was also utilised to detect sugar alcohol adulteration in coconut water using FTIR

spectroscopy (Thomas A. Teklemariam *et al.*, 2024). The accuracy obtained surpassed that of discriminant analysis and SVM. This demonstrates the capability of FTIR to detect alcohol in beverages, a critical requirement for ensuring compliance with Islamic dietary laws.

In addition, Table 4 shows that a few researchers utilised a regression algorithm to develop a predictive model for detecting adulteration or quality authenticity. Different types of regression algorithms, based on previous work, for example, support vector regression (SVR), PLS, and Gradient boosted regression (GBR), have achieved low error rates in the prediction model. As shown in Table 4, the support vector regression (SVR) application yields a low prediction error of less than 0.1 for nutrient prediction. However, SVR can also result in higher errors, with a maximum observed value of 3.78. These findings demonstrate that chemometric-based prediction models offer high accuracy and generally maintain a low risk of significant errors in quantifying food components, depending on the characteristics of the input spectra. The prediction model in chemometric analysis requires an optimisation process before being deployed to the instrument. More regression algorithms are available; thus, the chemometricians must try multiple algorithms before choosing the best model.

PLS, PCR and SVR are effective for small to medium-sized datasets due to their strong generalisation performance, but ANN require a large dataset to avoid model overfitting. In some applications, regression is combined with classification techniques to enhance food authentication. For instance, a classification model might first identify whether a food sample is halal or non-halal, while a regression model estimates the concentration of specific non-halal components. This hybrid approach increases the robustness and reliability of the authentication process. Overall, regression algorithms are indispensable in modern food spectroscopy, providing rapid, nondestructive, and accurate solutions for ensuring food authenticity, quality, and safety throughout various stages of the food supply chain.

Table 4*:* Different types of regression methods to assess food authenticity

Sample	Regression method	Root mean square error (RMSE)	Application	Industry implication
Potato plants (Abukmeil *et al.*, 2024)	SVR	< 0.1	Nutrient prediction (NPK) using Vis-NIR spectroscopy	Enhances precision agriculture through rapid nutrient assessment
Arabica coffee (Suhandy *et al.*, 2025)	PLS	2.01 %.	Detection of roasted soybean adulteration in Arabica coffee	Determine coffee quality control by identifying adulteration with cheaper ingredients
Honey (Boateng *et al.*, 2022)	GBR	2.183	Detection of syrup adulteration in honey	Ensures authenticity and protects consumer trust by detecting syrup adulteration
Extra virgin olive oil (Zaroual *et al.*, 2025)	SVR	3.78.	Adulteration quantification using emission spectra (excitation at 430 nm)	Aids in verifying purity and detecting economic fraud in premium oil products

Furthermore, regression models can be integrated into real-time monitoring systems, enabling continuous testing of food quality without the need for extensive sample preparation. This ability makes them ideal for high-throughput industrial applications. In more advanced systems, regression models may be combined with classification techniques to enhance food authentication's accuracy, such as distinguishing between halal and non-halal ingredients. Overall, regression algorithms play a crucial role in transforming spectral data into actionable insights, thereby ensuring the authenticity and safety of food.

4. Conclusion

Spectroscopy techniques such as FTIR, UV-Vis, NIR, and Raman have demonstrated strong potential in food authentication, with each method offering distinct advantages depending on the nature of the food matrix. For example, NIR and FTIR spectroscopy have efficiently quantified nutrients and detected adulterants due to their ability to capture complex molecular vibrations. With its high specificity, Raman spectroscopy has shown strengths in differentiating similar chemical compounds, making it suitable for origin verification and variety classification. When integrated with chemometric techniques, the predictive performance of these spectroscopic methods is significantly enhanced. For regression tasks such as nutrient quantification, SVR and PLS were among the most accurate models, with SVR achieving errors lower than 0.1 in some instances. However, model performance was also sensitive to the quality of spectral input, with occasional high errors indicating the need for robust preprocessing. For classification problems, methods such as PCA combined with LDA or classifiers like SVM have demonstrated high accuracy in distinguishing between authentic and adulterated samples and identifying the origin of food. The key contribution of this review lies in its focused examination of how spectroscopy and chemometric methods address not only general food authentication challenges but also specific religious requirements, such as halal verification.

However, despite these advancements, several critical challenges remain that limit the full deployment and standardisation of spectroscopy-based authentication systems. One primary concern is the lack of standardised protocols and calibration models across different instruments, which can affect the reproducibility and comparability of results. Data reproducibility and model transferability across different platforms and operators also remain challenging, especially for portable or handheld devices in field conditions. Methods, such as calibration transfer, should play a crucial role in ensuring the repeatability and reproducibility of data prediction across different instruments. Another significant issue is the limited availability of comprehensive spectral libraries and validated reference datasets, which are crucial for building robust predictive models. Future advancements in chemometric and spectroscopic instrumentation are expected to enhance these techniques' accuracy and usability and improve sensitivity and specificity. These techniques' usability promotes the evolution of artificial intelligence, real-time portable devices and transparent food authentication frameworks. In the context of halal food systems, these analytical tools support compliance with religious dietary laws by enabling the detection of non-halal adulterants such as pork derivatives without compromising the integrity of the sample. This technological approach not only reinforces trust and transparency within halal supply chains but also helps manufacturers and certifiers meet the increasing expectations of consumers for religiously permissible and ethically verified products. As advancements in portable devices and real-time spectral analysis continue, these methods are becoming increasingly accessible for on-site verification and ensuring halal integrity.

References

Abukmeil, R., Al-Mallahi, A., & Campelo, F. (2024). Using local support vector regression models to calibrate lab and field reflectance spectra for nutrient estimation in potato plants. Smart Agricultural Technology, 8. https://doi.org/10.1016/j.atech.2024.100492

Airin, C. M., Budikafa, M. J., Erwanto, Y., Sudjadi, S., Astuti, P., Sarmin, S., Windarsih, A., & Rohman, A. (2025). Fingerprinting using vibrational spectroscopy and chemometrics for determining the presence of non-halal meats in food products. In Applied Food Research (Vol. 5, Issue 1). Elsevier B.V. https://doi.org/10.1016/j.afres.2025.100982

Aslam, N., Fatima, R., Altemimi, A. B., Ahmad, T., Khalid, S., Hassan, S. A., & Aadil, R. M. (2024). Overview of industrial food fraud and authentication through chromatography techniques and its impact on public health. Food Chemistry, 460. https://doi.org/10.1016/j.foodchem.2024.140542

Balcázar-Zumaeta, C. R., Maicelo, J. L., Pajuelo-Muñoz, A. J., Torrejón-Valqui, L., Muñóz-Astecker, L. D., Barrena, M., Cayo-Colca, I. S., & Castro-Alayo, E. M. (2025). PCA, PLS, and MCR-ALS applied to the analysis of miscibility and quantification of Cupuassu butter, Passion fruit oil and Sacha Inchi oil blended in Cocoa butter. Applied Food Research, 5(1). https://doi.org/10.1016/j.afres.2024.100678

Banaś, J., & Banaś, M. (2024). Combined application of fluorescence spectroscopy and principal component analysis in characterisation of selected honeybees. Molecules, 29(4). https://doi.org/10.3390/molecules29040749

Biancolillo, A., Marini, F., Ruckebusch, C., & Vitale, R. (2020). Chemometric strategies for spectroscopy-based food authentication. In Applied Sciences (Switzerland) (Vol. 10, Issue 18). MDPI AG. https://doi.org/10.3390/APP10186544

Boateng, A. A., Sumaila, S., Lartey, M., Oppong, M. B., Opuni, K. F. M., & Adutwum, L. A. (2022). Evaluation of chemometric classification and regression models for the detection of syrup adulteration in honey. LWT, 163. https://doi.org/10.1016/j.lwt.2022.113498

Candeias, D. N. C., Silva, K. M., Bezerra, L. P., de Sousa Fernandes, D. D., & Diniz, P. H. G. D. (2025). Hierarchical authentication of the geographical origin of instant coffee using digital image-based fingerprints and chemometrics. Food Chemistry, 144278. https://doi.org/10.1016/j.foodchem.2025.144278

Cavdaroglu, C., & Ozen, B. (2022). Detection of vinegar adulteration with spirit vinegar and acetic acid using UV–visible and Fourier transform infrared spectroscopy. Food Chemistry, 379. https://doi.org/10.1016/j.foodchem.2022.132150

Chen, R., Li, S., Cao, H., Xu, T., Bai, Y., Li, Z., Leng, X., & Huang, Y. (2024). Rapid quality evaluation and geographical origin recognition of ginger powder by portable NIRS in tandem with chemometrics. Food Chemistry, 438. https://doi.org/10.1016/j.foodchem.2023.137931

da Silva Medeiros, M. L., Brasil, Y. L., Cruz-Tirado, L. J. P., Lima, A. F., Godoy, H. T., & Barbin, D. F. (2023a). Portable NIR spectrometer and chemometric tools for predicting quality attributes and adulteration levels in butteroil. Food Control, 144. https://doi.org/10.1016/j.foodcont.2022.109349

da Silva Medeiros, M. L., Brasil, Y. L., Cruz-Tirado, L. J. P., Lima, A. F., Godoy, H. T., & Barbin, D. F. (2023b). Portable NIR spectrometer and chemometric tools for predicting quality attributes and adulteration levels in butteroil. Food Control, 144. https://doi.org/10.1016/j.foodcont.2022.109349

Dahimi, O., Hassan, M. S., Rahim, A. A., Abdulkarim, S. M., & A, S. M. (2014). Differentiation of Lard from Other Edible Fats by Gas Chromatography- Flame Ionisation Detector (GC-FID) and Chemometrics. Journal of Food and Pharmaceutical Sciences, 2, 27–31.

de Freitas Oliveira, T., de Jesus, J. C., Coqueiro, J. S., Silva, R. R., Ferrão, S. P. B., & Santos, L. S. (2025). Assessing the authenticity of the sport supplement creatine using spectroscopic techniques and multivariate analysis. Food Chemistry, 144282. https://doi.org/10.1016/j.foodchem.2025.144282

Dimakopoulou-Papazoglou, D., Ploskas, N., Serrano, S., Silva, C. S., Valdramidis, V., Koutsoumanis, K., & Katsanidis, E. (2023). Application of UV–Vis spectroscopy for the detection of adulteration in Mediterranean honeys. European Food Research and Technology, 249(12), 3043–3053. https://doi.org/10.1007/s00217-023-04347-1

Erinawati, D. A., Setyaningsih, W., & Palma, M. (2025). Spectroscopy and chemometric-based method for sugar profiling in high-fructose syrup. Applied Food Research, 5(1). https://doi.org/10.1016/j.afres.2025.100872

Goyal, R., Singha, P., & Singh, S. K. (2025). Machine learning-assisted Fourier transform infrared spectroscopy to predict adulteration in coriander powder. Food Chemistry, 477. https://doi.org/10.1016/j.foodchem.2025.143502

Irnawati, Rohman, A., Yamin, Fadzillah, N. A., Azmi, A. A., Nurlatifah, Windarsih, A., Susidarti, R. A., & Ruslin. (2024). Moringa oleifera seed oils: Physico-chemical characterisation and its authentication using FTIR spectroscopy and chemometrics. Case Studies in Chemical and Environmental Engineering, 10. https://doi.org/10.1016/j.cscee.2024.100994

Isdhiyanti, M. I. T., Budiastra, I. W., Mala, D. M., Mardjan, S., Hasrini, R. F., & Supriatna, D. (2024). Prediction of chemical contents of cooking oil using near-infrared spectroscopy. Food Research, 8(6), 147–152. https://doi.org/10.26656/fr.2017.8(6).520

Jayasekher, A., Panchariya, P. C., Maurelli, F., Prajapati, D., & Palit, A. K. (2024a). Authentication of Mentha arvensis essential oil using attenuated total reflectance-fourier transform infrared spectrophotometry coupled with chemometrics. Journal of Food Composition and Analysis, 135. https://doi.org/10.1016/j.jfca.2024.106576

Jayasekher, A., Panchariya, P. C., Maurelli, F., Prajapati, D., & Palit, A. K. (2024b). Authentication of Mentha arvensis essential oil using attenuated total reflectance-fourier transform infrared spectrophotometry coupled with chemometrics. Journal of Food Composition and Analysis, 135.

https://doi.org/10.1016/j.jfca.2024.106576

Kolašinac, S. M., Pećinar, I., Gajić, R., Mutavdžić, D., & Dajić Stevanović, Z. P. (2025). Raman spectroscopy in the characterisation of food carotenoids: Challenges and prospects. In Foods (Vol. 14, Issue 6). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/foods14060953

Kusumadewi, A. P., Martien, R., Pramono, S., Setyawan, A. A., & Rohman, A. (2022). Review on the application of chemometrics for the standardisation and authentication of Curcuma xanthorrhiza. In Food Research (Vol. 6, Issue 2, pp. 1–8). Rynnye Lyan Resources. https://doi.org/10.26656/fr.2017.6(2).044

Li, N., Hussain, N., Ding, Z., Qu, C., Li, Y., Chu, L., & Liu, H. (2024). Guidelines for Raman spectroscopy and imaging techniques in food safety analysis. Food Safety and Health, 2(2), 221–237. https://doi.org/10.1002/fsh3.12040

Li, X., Zhang, L., Zhang, Y., Wang, D., Wang, X., Yu, L., & Li, P. (2020). Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils.

Mohammadi, N., Esteki, M., & Simal-Gandara, J. (2024). Machine learning for authentication of black tea from narrow-geographic origins: Combination of PCA and PLS with LDA and SVM classifiers. LWT, 203. https://doi.org/10.1016/j.lwt.2024.116401

Mohammadi, S., Gowen, A., & O'Donnell, C. (2024). Vibrational spectroscopy data fusion for enhanced classification of different milk types. Heliyon, 10(16). https://doi.org/10.1016/j.heliyon.2024.e36385

Motahari, H., Mousavi, S. S., & Haghighi, P. (2023). Raman, FTIR, and UV–Vis spectroscopic investigation of some oils and their hierarchical agglomerative clustering (HAC). Food Analytical Methods, 16(7), 1237–1251. https://doi.org/10.1007/s12161-023-02481-w

Musfiroh, I., Maritha, V., Harlina, P. W., Muchtaridi, M., Bakar, N. K. A., Rohman, A., Dachriyanus, D., Ikram, N. K. K., & Windarsih, A. (2025). Chemometrics applications in omics studies (metabolomics, lipidomics, and proteomics) for halal authentication in food and pharmaceutical products. In Applied Food Research (Vol. 5, Issue 1). Elsevier B.V. https://doi.org/10.1016/j.afres.2025.100770

Özdemir, İ. S., Firat, E. Ö., Özturk, T., Zomp, G., & Arici, M. (2024). Geographical origin determination of the PDO hazelnut (cv. Giresun Tombul) by chemometric analysis of FT-NIR and Raman spectra acquired from shell and kernel. Journal of Food Science, 89(8), 4806–4822. https://doi.org/10.1111/1750-3841.17214

Pasieczna-Patkowska, S., Cichy, M., & Flieger, J. (2025). Application of fourier transform infrared (FTIR) spectroscopy in characterization of green synthesized nanoparticles. In Molecules (Vol. 30, Issue 3). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/molecules30030684

Rodionova, O. Y., Oliveri, P., Malegori, C., & Pomerantsev, A. L. (2024). Chemometrics as an efficient tool for food authentication: Golden pillars for building reliable models. In Trends in Food Science and Technology (Vol. 147). Elsevier

Ltd. https://doi.org/10.1016/j.tifs.2024.104429

Suhandy, D., Yulia, M., Munawar, A. A., & Kusumiyati, K. (2023). Simple analytical method using ultraviolet spectral dataset and chemometrics for the authentication of Indonesian specialty ground roasted coffee with different botanical and geographical indications. Data in Brief, 51. https://doi.org/10.1016/j.dib.2023.109820

Suhandy, D., Yulia, M., Widodo, S., Naito, H., & Al Riza, D. F. (2025). Fast authentication of Indonesian ground-roasted Arabica coffee adulterated with roasted soybean by portable LED-based fluorescence spectroscopy and chemometrics analysis. Food Chemistry, 479. https://doi.org/10.1016/j.foodchem.2025.143791

Sun, Z., Tian, H., Hu, D., Yang, J., Xie, L., Xu, H., & Ying, Y. (2025). Integrating deep learning and data fusion for enhanced oranges soluble solids content prediction using machine vision and Vis/NIR spectroscopy. Food Chemistry, 464. https://doi.org/10.1016/j.foodchem.2024.141488

Suyal, M., & Goyal, P. (2022). A Review on Analysis of K-nearest neighbor classification machine learning algorithms based on supervised learning. In International Journal of Engineering Trends and Technology (Vol. 70, Issue 7, pp. 43–48). Seventh Sense Research Group. https://doi.org/10.14445/22315381/IJETT-V70I7P205

Tarcan, B., & Küplülü, Ö. (2024). Rapid determination of chicken meat ratios in Beef Mixtures and Beef Sausages by Near Infrared Reflectance (NIR) spectroscopy. Ankara Universitesi Veteriner Fakultesi Dergisi, 71(3), 311–319. https://doi.org/10.33988/auvfd.1200920

Taylan, O., Cebi, N., Tahsin Yilmaz, M., Sagdic, O., & Bakhsh, A. A. (2020). Detection of lard in butter using Raman spectroscopy combined with chemometrics. Food Chemistry, 332. https://doi.org/10.1016/j.foodchem.2020.127344

Thomas A Teklemariam, Chou, F., Kumaravel, P., & Van Buskrik, J. (2024). ATR-FTIR spectroscopy and machine/deep learning models for detecting adulteration in coconut water with sugars, sugar alcohols, and artificial sweeteners. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 322. https://ssrn.com/abstract=4751204

Usman, A. G., Mati, S., Daud, H., Suleiman, A. A., Abba, S. I., Ahmad, H., & Radwan, T. (2025). Optimised SVR with nature-inspired algorithms for environmental modelling of mycotoxins in food virtual-water samples. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-99908-7

Wang, J., Qian, J., Xu, M., Ding, J., Yue, Z., Zhang, Y., Dai, H., Liu, X., & Pi, F. (2025). Adulteration detection of multi-species vegetable oils in camellia oil using Raman spectroscopy: Comparison of chemometrics and deep learning methods. Food Chemistry, 463. https://doi.org/10.1016/j.foodchem.2024.141314

Wu, B., Tang, W., Zhou, J., Jia, H., Shen, H., & Qi, Z. (2024). Near-infrared spectroscopy combined with fuzzy fast pseudoinverse linear discriminant analysis to discriminate mee tea grades. Heliyon, 10(5). https://doi.org/10.1016/j.heliyon.2024.e27732

Yang, Q., Xia, C., Cui, Y., Zhang, X., Zhang, T., Huang, Y., Liu, D., Wei, Y., Gong, Y., Fang, G., Wang, Z., Chen, X., & Hu, M. (2025). Fusion of terahertz spectroscopy and Raman spectroscopy combined with support vector machine to distinguish different pericarpium citri reticulatae species. Heliyon, 11(6). https://doi.org/10.1016/j.heliyon.2025.e43136

Yulia, M., & Suhandy, D. (2020). The potential application of UV-visible spectroscopy and chemometrics for discrimination of Lampung robusta coffee with different fermentations. IOP Conference Series: Earth and Environmental Science, 537(1). https://doi.org/10.1088/1755-1315/537/1/012035

Zan, J., Li, H., Cai, L., Wu, C., Fan, Z., & Sun, T. (2024). Detection of tea seed oil adulteration based on near-infrared and Raman spectra information fusion. LWT, 213. https://doi.org/10.1016/j.lwt.2024.117064

Zaroual, H., El Hadrami, E. M., Farah, A., Ez zoubi, Y., Chénè, C., & Karoui, R. (2025). Detection and quantification of extra virgin olive oil adulteration by other grades of olive oil using front-face fluorescence spectroscopy and different multivariate analysis techniques. Food Chemistry, 479. https://doi.org/10.1016/j.foodchem.2025.143736

Zhu, R., Wu, X., Wu, B., & Gao, J. (2024). High-accuracy classification and origin traceability of peanut kernels based on near-infrared (NIR) spectroscopy using Adaboost - Maximum uncertainty linear discriminant analysis. Current Research in Food Science, 8. https://doi.org/10.1016/j.crfs.2024.100766