A review on the application of chromatographic methods, coupled to chemometrics, for food authentication

(Chromatography-chemometrics in food authentication)

M. Esteki\textsuperscript{a,*} (m.esteki@znu.ac.ir)
J. Simal-Gandara\textsuperscript{b,*} (jsimal@uvigo.es)
Z. Shahsavari\textsuperscript{a} (z.shahsavair@znu.ac.ir)
S. Zandbaaf\textsuperscript{a} (s.zandbaaf@znu.ac.ir)
E. Dashtaki\textsuperscript{a} (e.dashtaki@znu.ac.ir)
Yvan Vander Heyden\textsuperscript{c} (yvanvdh@vub.ac.be)

\textsuperscript{a} Department of Chemistry, University of Zanjan, Zanjan (Iran).
\textsuperscript{b} Nutrition and Bromatology Group, Department of Analytical and Food Chemistry, Faculty of Food Science and Technology, University of Vigo – Ourense Campus, Ourense (Spain).
\textsuperscript{c} Department of Analytical Chemistry, Applied Chemometrics and Molecular Modelling, Center for Pharmaceutical Research (CePhaR), Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium

Corresponding authors:
J. Simal-Gandara (Email: jsimal@uvigo.es) & M. Esteki (Email: m.esteki@znu.ac.ir)
Abstract

The increase of food adulteration, inducing losing a large amount of money as well as of the confidence of consumers, has become an urgent issue for producers, researchers, governments and consumers. Chromatographic methods, in combination with chemometrics, are usually developed and applied throughout the food chain to verify the nature or origin of food, with both targeted (metabolomics) and non-targeted (profiling) approaches. Their operation, together with their advantages and drawbacks, will be discussed in this review to show strategies to solve food authentication issues.

Keywords: food authentication; chemometrics; chromatography; gas chromatography; high performance liquid chromatography; profiling; metabolomics.
1.-Food fraud and authentication of food products

Fraud and adulteration detection, so-called food authentication, is a process which verifies the food compliance with its label description (Oliveri and Downey, 2012; Walker, 2017). This may include the geographical origin, production method, processing technologies and food composition. The declaration of specific quality attributes in expensive products is of particular interest, since these products are often target of fraudulent labeling. Rapid population growth results in raw material shortage and high prices, which forces some producers to bulk up their products with questionable fillers, a strategy that, in itself, introduces additional ingredients of unknown origin. Besides economic losses, food fraud represents a threat to human health, for instance prohibited additives may be toxic or contaminated with pathogens, or non-declared substitutes and production processes can cause health problems, such as allergic reactions (Gerbig et al., 2017). Evidence of provenance is an important topic when dealing with food quality and consumer protection, as well as is the compliance with national legislation, international standards, and guidelines (Danezis et al., 2016). Labelling and compositional regulations, which may differ from country to country, have a fundamental place in determining what scientific tests are appropriate for a particular issue. Consumers around the world are increasingly demanding reassurance that the origin and content of their food complies with the information written on the packaging labels. This pressure from consumer side intensifies the regulations imposed on the food supply worldwide.

Food adulteration is a growing challenge for food manufacturers and analysts because most adulterants are unknown, and difficult to recognize using the typical targeted screening methods. This industry urgently needs non-targeted screening methods for food samples to provide proof
of origin and prevent deliberately or accidentally undeclared additions to those samples (Esslinger et al., 2014). Since the 1980s, consumers, producers and regulatory bodies have recognized the authenticity of food products as an important quality criterion. Rapid, specific methods for detecting adulterations, verifying the quality, guaranteeing the geographical origin, and the type of production of food products are nowadays widely in demand (Rodríguez-Bermúdez et al., 2018). Thus, there is a necessity for accurate standardized food authentication. Consumers and producers similarly put a high value on accurate and defensible labeling, and providers are now proactively offering supplying clear labeling, traceability, and transparency. Reliable and robust analytical tools must be available throughout the food chain to verify its nature. Typically, there are several criteria for choosing a method, including its detection limits and sensitivity, sample preparation steps, feasibility and cost, and its throughput. Such tools should permit rapid, nondestructive, and inexpensive analysis. Various techniques are available for testing food authenticity, including chromatography (Cseháti et al., 2005; Cuadros-Rodríguez et al., 2016b), ultraviolet (Martelo-Vidal et al., 2013), near-infrared (Rohman, 2017; Wang and Yu, 2015), mid-infrared (Rodriguez-Saona and Allendorf, 2011), and Raman spectroscopy (Oroian and Ropciuc, 2017). All are routinely used to check both raw materials and finished food products with respect to production standards.

2. Adulteration of different food products

Food adulteration is not a contemporary phenomenon, but its history dates back to the beginning of the manufacturing of food. Special attention to food fraud issues is a rapidly growing field due to increasing public awareness concerning economically motivated adulteration, which might pose serious public health risks to humans. In this regard, non-compliance with food laws,
misleading the consumer and intentional fraud with prospect of financial gain are the key characteristics of food fraud. Nowadays, due to the increased globalization of food markets, many food products are also supplied from other countries. Therefore detection and trace the source of intentional and unintentional contamination would be really difficult especially in highly processed foods. In general, foods and food ingredients that are usually accompanied by food fraud are included edible oils, honey, milk and dairy products, fruit juices, wine and alcoholic beverages, meat products, spices, coffee, tea, grain-based foods, organic foods and some highly processed foods.

2.1 Edible oils and fats

Edible oils and fats, including vegetable oils, animal fats, salad and cooking oils, margarine and butter, are classified as the foods which are most frequently susceptible to adulteration (Hong et al., 2017). At present, adulteration in edible oils and fats has become the most important food safety issues because it has seriously impacted people's health. Replacing more expensive oils and fats with cheaper ones and mixing cold press oil with a refined one are two major ways of adulterations in edible oils and fats (Azadmard-Damirchi and Torbati, 2015).

Olive oil, especially, with regard to its high demand and profit potential, is a significant target for economically motivated adulteration. The chemical composition of the olive oils might differ due to the fruit cultivar, degree of fruit ripeness, environmental and storage conditions, geographical production area, and the extraction methods (Kalua et al., 2007). These properties not only affect the fatty acid composition of the oil, but also especially regulate the levels of diverse classes of minor compounds such as polyphenols, tocopherols, phytosterols and carotenoids. Virgin olive oil generally mixed with lower grades of olive oil such as refined or
pomace olive oil or other cheaper vegetable oils such as hazelnut, cotton, and sunflower or soybean oil. Geographical origin of olive oil production is another important choice factor for consumers (Benincasa et al., 2007; Menapace et al., 2011). Identification of the geographical origin of olive oil is an indispensable issue to ensure the high quality of the extra virgin olive oils which shows the specific characteristics of quality attributed to their geographical origin (Bendini et al., 2007).

2.2 Dairy products

Milk and dairy products with highly nutritive value make a major contribution to a well balanced diet for different consumer groups such as children and pregnant women (Jenkins and McGuire, 2006). Milk as one of the seven top-adulterated food products could be counterfeited by several ways (Borková and Snášelová, 2005). Mixing with different types of milk (Dias et al., 2009) or whey (Neelima et al., 2013), neutralizing to mask acidity of milk (Silva et al., 2015) and addition of melamine (Jawaid et al., 2013), salt or sugar (Nirwal et al., 2013) to mask extra water or high solid contents are some instances. Commercial ultra-high temperature milks (UHT) may be presented by addition of adulterants such as starch, chlorine, formalin, hydrogen peroxide, urine, etc (Souza et al., 2011). Excessive addition of water into milk leading decrease of nutritional composition (Mabrook and Petty, 2003), and addition of non-milk fat into dairy products (Kamal and Karoui, 2015; Nascimento et al., 2017) are other counterfeiting methods which generally used. The authenticity of dairy products could also be attributed to their geographical origin (Osorio et al., 2015) and processing technology (Schmidt and Mayer, 2018). Labeling of conventional milk as a product from organic farming is another aspect related to the authenticity of milk and dairy products which should be considered (Erich et al., 2015). Therefore,
authentication of dairy products is a major concern in order to carry out the monitoring (Sardina et al., 2015) to ensure that dairy products are correctly labeled in terms of actually processed for consumption (Di Domenico et al., 2017). Traceability of dairy products is also an important issue indicating the presence of undesirable compounds such as antimicrobials, mycotoxins, organochlorine pesticide, antibiotic residues and heavy metals in order to protect consumers from harmful contamination (Motarjemi et al., 2014). Although authenticity issues of dairy products mostly include detecting of the adulteration, traceability and safety, other factors such as processing conditions and packaging could also be considered.

2.3 Honey

Honey is a rich conventional natural resource with a specific flavour, odour, taste, nutritive value and therapeutic characteristics. Its quality is mainly determined by its sensorial, chemical, physical and microbiological characteristics depends on its chemical composition, floral origin, production method, thermal treatment, climatic conditions of the region and the conditions of manipulating and packaging (Alvarez-Suarez et al., 2010; Turhan et al., 2008).

Honey has the potential to be used in a variety of applications in the food industry. It can be used by direct consumption or incorporated as an ingredient in various processed food products for its sweetness, smell, color, texture, flavor, caramelisation and viscosity (Azeredo et al., 2003; Kahraman et al., 2010). Because of its nutritional value and unique flavor of honey, and considering its production cost the price of natural honey is much higher than that of other sweeteners, such as beet sugar or refined cane sugar. Therefore, honey is always the main target of food adulteration. Changing the composition of honey by adding other sweet ingredients, such as high fructose corn syrup, glucose syrup or saccharose syrups in any part of the production or
processing can be an attractive way to achieve financial benefits. Adulteration of honey by adding cheap sweet ingredients has been reported in the literature (Downey et al., 2003; Kelly et al., 2004; Li et al., 2012). It is often difficult to detect this kind of adulteration, due to the fact that the sugar compounds of these cheap syrups are sometimes close to honey. Moreover, adulteration of honey by feeding bees with artificial sources such as sugar or syrup (Kast and Roetschi, 2017) and incorrect information about the honey’s geographic and botanical origin lead to serious problems for honey producers and consumers (Bougrini et al., 2016). Therefore, investigation of the authenticity of honey is especially important for commercial and health reasons.

2.4 Beverages

Fluid foods, such as alcohol, fruit juices and other beverages are especially prone to simply being adulterated since they can be easily mixed with a number of cheaper liquids. This cause not only producing low quality products, but in some cases also creates toxic compounds that may lead to serious health problems.

Fruit juices

Fruit juice production is a significant and rapidly growing demand in the beverage industry. High-quality fruit juices and nectars are another target for adulteration by simple dilution with water or juices obtained from cheaper fruits, addition of sugar syrups, acids and colorant agents (Ogrinc et al., 2003).

Alcoholic beverages

The economic value of wine and the growing of market globalization made the wine authentication and traceability important tasks worldwide. Variety, provenance, production year,
process of vinification and quality ratings, which are the most important features used to specify, describe and pricing of wines, can be affected by the total composition of small molecules (Cuadros-Inostroza et al., 2010). In the past decades, the only adopted method for characterization and discrimination of wines was based on sensory evaluation performed by a panel of experts (Ballester et al., 2008). It is also obvious that the modern wine industry needs both fast and reliable process quality control analytical methods that allow rapid and efficient analysis to assure the quality of the final product to the consumer. The volatile aromas produced by wine yeast during fermentation include higher alcohols, esters and volatile fatty acids are particularly useful for characterizing wines cultivars (Garde-Cerd et al., 2008).

Wine adulteration, especially regarding varieties and original regions, has been extensively investigated, because wine is a widely distributed product which could be easily adulterated. Adulteration of wine or other alcoholic beverages is usually accomplished by dilution with water, addition of alcohol, dyes and aromas, and mixing with lower quality beverages (Stanziani, 2009). Beside this type of adulteration, mislabeling about composition, verities or geographical origin could be carried out about the wine (Ohtsubo et al., 2008).

Coffee

In the past two decades, It has been noted that the global consumption of coffee is growing (Roberts, 2016), stimulated by advances made in terms of product quality. However, the adulteration of roasted coffee is a common strategy used to reduce costs (Toci et al., 2016). Coffee fraud frequently practiced by using the low quality beans (geographical origin and defective beans), or by addition of other substances (coffee husks or parchments, corn and barley, cereals or caramelized sugar, wheat middling, soybean and rye) to coffee blends in order
to achieve less expensive product (de Moura Ribeiro et al., 2017; Nogueira and Do Lago, 2009; Pauli et al., 2014). The coffee consumed worldwide primarily comes from two cultivated species: *Coffea arabica* (arabica coffee) and *Coffea canephora* (robusta coffee), which account for 75% and 25% of the world market, respectively. Arabica coffee is the most appreciated by consumers because of producing a high quality beverage with intense aroma, low bitterness and good flavor, low caffeine content and also by the absence of off-flavors, such as smoky notes. This makes the price of arabica coffee usually 20–25% higher than that of robusta. Therefore, detection of coffee adulteration (such as illegal admixture of cheaper robust coffee) and its quantification is crucial to guarantee coffee quality, minimize unfair trade and raise consumer confidence.

### Tea

Tea (*Camellia sinensis* L.) is one of the most popular beverages in a wide range of countries in the world. In general, green and oolong tea are consumed mostly in Asian countries such as India, China, Japan and Thailand, while black tea is more popular in Western countries (Sharangi, 2009).

Adulteration of tea usually involves addition of colorants, substitution with other plants such as cashew husk, mixing with tea from other geographical origins or lower quality grades. In this way, appropriate analytical methods are important tools for monitoring adulteration activities and preventing incorrect labeling. Metabolite profiles for various teas collected from the most famous tea producing regions may be used to assess the quality or tea discriminate distinct tea products from different locations.
2.5 Meat and meat products

Recently, interest in meat authenticity and traceability has increased. The demand of precise information about origin, accurate labeling and ingredients of meat products is growing rapidly due to increased awareness of consumers about their health. Four main categories including meat species, production processes, processing treatment, geographic origin and non-meat ingredient addition (additives and water) should be considered in authentication of meat and meat products (Ballin, 2010). Moreover, accurate and unambiguous labeling of meat products is important to reassure consumers, protect regional designations and ensure fair competition. There is a strong preference and willingness to pay a higher price for particularly protected geographical indication (PGI) meat products which almost are organic products from selected breeds produced in a particular area (Deselniciu et al., 2013; Pla et al., 2007).

Fish and seafood fraud usually involves a deliberate increase in product weight and the use of prohibited additives in production. Addition of excess water to frozen product (overglazing) or partial substitution by cheaper meats are some examples (Moore et al., 2012).

3. Chemometrics

Chemometrics is a chemical discipline that uses statistical and mathematical methods to achieve objective data evaluation by extracting meaningful information from related and unrelated collections of chemical data. Chemometrics and specifically multivariate analysis, find several applications in quality control, quantitative and qualitative determination of chemical parameters
for assessing the food-products authenticity (Yu et al., 2018). Chemometrics provides powerful tools in calibration analysis of spectroscopic and chromatographic data, applied in targeted and non-targeted approaches to identify various food fraud situations or to verificate their geographic or biological origin (Beale et al., 2017; Martínez Bueno et al., 2018; Popping et al., 2017). The most common multivariate methods and principles for food authentication can be classified into three categories: exploratory data analysis; data description and visualization, discrimination and classification; and regression and prediction.

Classification methods, based on multivariate data analysis, could be supervised or unsupervised (Matera et al., 2014). In unsupervised methods, the purpose is to identify clusters or relationships between samples, without any prior knowledge of classes or groups. In contrast, supervised methods require information on class membership and a training stage to build up a proper mathematical model (Moncayo et al., 2015). Subsequently, unknown objects of a test or validation set can be predicted. The most commonly used unsupervised methods in food authentication analysis are principal component analysis (PCA) and hierarchical clustering (HC) (Granato et al., 2018). Many algorithms can be used to perform supervised methods, such as linear discriminant analysis (LDA) (D’Archivio and Maggi, 2017; Ma et al., 2016; Yudthavorasit et al., 2014), partial least squares (PLS) regression (Lenhardt et al., 2015; Sampaio et al., 2018), as linear methods, and artificial neural network (ANN) (da Silva et al., 2015; Gonzalez-Fernandez et al., 2018), as a non-linear classification method.

3.1.- Multivariate classification for qualitative analysis
In qualitative analysis, the sample properties one likes to relate to the corresponding chromatographic signal have discrete values that represent food product quality or identity, e.g. “pure” or “blended”, “fresh” or “non-fresh agent” (Callao and Ruisánchez, 2018). Multivariate classification methods are used in classification of samples in order to solve the selectivity and interference problems. Multivariate classification methods, also known as pattern recognition methods, are subdivided into two major categories: “supervised” and “un-supervised” learning algorithms, which will be explained in more detail in following sections (Cuadros-Rodríguez et al., 2016a).

3.1.1. Pattern recognition

Nowadays, rapid growth of the capabilities of modern analytical instruments provides a large amount of data involve a wide range of factors (features) which need efficient multivariate analysis methods for extraction of meaningful information. Classification, recognition, and grouping of patterns are important issues in a variety of different scientific fields such as food science and technology. A pattern could be achieved from various kinds of analytical methods, such as chromatography and spectroscopy. The recognition/classification of patterns may include supervised or/and unsupervised pattern-recognition (Palacios-Morillo et al., 2016). In supervised classification methods, members of each class are predefined because the information used to group samples into a subset is known, whereas, in unsupervised pattern recognition methods, the problem is to find similarities and differences between samples, without applying any additional prior information about them (Cen et al., 2016).
While rapid and affordable computing allows faster processing of large datasets, this is also facilitated by the use of sophisticated and diverse methods for data analysis and classification. At the same time, because of the access to large databases and precise performance requirements (accuracy, speed, and cost), the demand for automatic pattern recognition systems increases (Berrueta et al., 2007). Creating a pattern-recognition system involves three steps, i.e. obtaining data sets and their pre-processing, data representation, and ultimately decision making. In this process, problems to solve are the selection of proper sources of data production, the pre-processing, the design of the representation and the building of a decision model (Assis et al., 2018). In general, it has been shown that a problem of accurate and sufficiently limited recognition will lead to a simple model representation, a simple conclusion and a simple decision strategy. Learning from a collection of known samples that make up the training/calibration set, is an important feature of most pattern-recognition methods (Brereton, 2015).

3.1.1-Unsupervised pattern recognition

Unsupervised methods, also known as exploratory data analysis methods, do not need any prior knowledge about the class structure of the data, but instead may produce the grouping, i.e. clustering, themselves (Berkhin, 2006). Exploratory data analysis techniques are often quite helpful in elucidating the complex nature of multivariate relationships by using mapping and display techniques for understanding the structure of complex multivariate data sets (Bro et al., 2002). Principal component analysis is the most widely used unsupervised pattern recognition technique in chemometrics. PCA is often the first step in data analysis to identify or verify patterns in measured data (Berrueta et al., 2007). PCA converts the measured primary variables into new independent variables, which are linear combinations of the original variables and
called principal components. PCA successively provides a set of orthogonal axes indicating the
direction of the largest (remaining) variance in the data. The first principal component (PC1)
accounts for the maximum of the total variance, the second (PC2) is orthogonal to the first and
lies in the direction of the largest remaining variation, and so on, until the total variance is
accounted for. Each principal component contains different sources of information, since each
defines a direction of distribution in the data. The projections of the points from the original data
space on PCs are called the scores of the objects. The scores are weighted sum of the original
variables; these weights are called loadings. Row mode eigenvectors (loadings) form an
orthonormal set of basis vectors that span the space of the original data set.

PCA, by reducing the dimensions of the data, allows them to be visualized, while maintaining as
much information as possible from the original data (Elmqvist and Fekete, 2010). It means the
PCA model must capture a large fraction of the variance in the data set, say 70% or more, in the
first few principal components. The basic idea behind PCA as an unsupervised learning
 technique is that data vectors representing objects in a high-dimensional space can be efficiently
 projected into a low-dimensional space and visualized graphically as scatter plots of PC scores.
Objects (samples) that are similar tend to cluster in the score plots, while different objects tend to
be separated.

Other methods of recognizing patterns such as cluster analysis (CA) that give insight into the
structure of a data set can also be used for initial evaluation of the contents of information in data
matrices. CA attempts also finding sample groupings or clusters within data, using criteria
developed from the data itself. In other words, in CA, samples are grouped according to
similarities, regardless of class membership information (Newby and Tucker, 2004). Cluster analysis is based on the principle that distances between pairs of points (i.e., samples) in the measurement space are inversely related to their degree of similarity. The basis of this approach is based on the idea that the similarity is inversely related to the distance between the samples. Therefore, CA determines the distances between all pairs of points in the data set using a defined metric parameter, such as Euclidean distance or Manhattan distance. Correlation is another criterion used to express similarity. High correlation indicates a high degree of similarity. Similar objects, then sequentially are merged according to a given clustering algorithm. Although several different types of clustering algorithms exist, e.g., Hierarchical clustering, K-means and FCV (Rezaee, 2010), by far the most popular is hierarchical clustering (Zhao and Karypis, 2005). The result of this technique is a dendrogram, a visual representation of the relationships between the samples in the dataset. Interpretation of the results is intuitive, which is the major reason for the popularity of these methods.

3.1.2.-Supervised pattern recognition

In “supervised” methods (also known as discriminant analysis), the classes and their attributes are known, so qualitative information is added to a classification multivariate analytical data. In general, supervised techniques are used to calibrate or classify collections by using the information about the class membership of the samples to build a classification model. Then these models classify new unknown samples in one or several of the known classes based on its measured pattern. Supervised pattern recognition methods use the following common strategy for all learning algorithms (Berrueta et al., 2007):
Selection of a training set (calibration set) and a test set, which includes members of the known class membership. The calibration set is used for the adjustment of the parameters of the model, while the test set is used for measuring the predictive ability of the model. However, such division is only possible when the data set is large enough.

Variable selection, which is a crucial problem in building a multivariate analysis model. The purpose of variable selection or feature selection is to eliminate irrelevant variables encoding the noise and/or without discriminating power to enhance the generalization performance of the learning algorithm.

Model construction by using the training set. A mathematical model is made by a set of selected variables from the samples that belong to the training set (model input) and their known categories (model output).

Verification and validation of the model involves running a model using input parameters measured for an independent test set of samples, in order to assess the quality and the reliability of the classification model.

The most popular techniques for the classification of food products include linear discriminant analysis (LDA) (D’Archivio et al., 2016), k-nearest neighbor (k-NN) (Sikorska et al., 2005), soft independent modeling of class analogy (SIMCA) (Urbano et al., 2006), unequal dispersed classes (UNEQ) (Oliveri and Downey, 2012), artificial neural network (ANN) (Dębska and Guzowska-Świder, 2011), wavelet neural network (WNN) (Zhang et al., 2014), support vector machine (SVM) (Du and Sun, 2005), partial least squares–discriminant analysis (PLS-DA) (Souto et al., 2010) and orthogonal projections to latent structures–discriminant analysis (OPLS-DA) (Sales et al., 2017). These methods can be subdivided into two groups; first classification
methods also called discrimination methods which focus on discriminating between classes, such as LDA, kNN, PLS-DA and ANN; and those oriented towards modeling classes, such as SIMCA and UNEQ. Most discriminating methods are based on using all categories for model building, while disjoint class-modeling methods create a separate model for each category. One of the problems associated with discriminating methods is the risk of the classification of samples that do not belong to any of the categories. They obligates are awarded to one of the considered classes. Class-modeling methods are specifically designed to describe a single class at a time, which means that decision-making on adaptation is not affected by samples outside the class. However, recent empirical experiments show that SIMCA is not as powerful in classification as previously believed since the directions defined by the significant components are based on the largest variation, which may be different from the direction of separating the classes (Brereton, 2015).

The nature of pattern recognition methods can also typically be classified using terms as parametric/non-parametric, deterministic/probabilistic or linear/non-linear methods (Cooper et al., 2015). A parametric technique assumes that the obtained data can be described with a probability density function, which determines its distribution. In most cases, it is assumed that the data is normally distributed with a known mean and variance. In the parametric techniques, statistical parameters of the normal distribution of samples are used in the decision rules. Parametric methods of pattern recognition assume that probability density functions are known or can be estimated. Parametric techniques, such as LDA, PLS-DA, SIMCA and UNEQ, use the statistical parameters of samples distribution in extracting the decision function. In non-parametric methods such as kNN and ANN, the statistics of the models are not based on
distribution assumption, which makes the estimation of the probabilities of correct classification more difficult.

3.2.-Multivariate calibration for quantitative analysis

Recently, multivariate calibration has been widely used in different food and chemical industries without complex separation and pre-treatment processes prior to their analysis (De Carvalho et al., 2015; Miaw et al., 2018). The aim of multivariate calibration methods is to determine relationships between a response variable and several independent variables. In these methods, a mathematical model is developed that allows quantitative predictions of desired properties, such as adulteration percent, from a set of input variables, such as chromatographic or spectroscopic measurements at different time points or wavelength. The constructed model should be able to recognize and capture important features of the observed data in order to provide an effective framework, accurate and precise enough to be able to properly predict the output variable for future samples. Multivariate calibration methods are a collection of linear and non-linear mathematical techniques that can be applied to chemical data analysis when more than one measurement is acquired for each sample. Multiple linear regression (MLR) (Liu et al., 2014), principal component regression (PCR) (Sahin and Demir, 2016), partial least squares (PLS) regressions (Lim et al., 2016), uninformative variable elimination PLS (UVE-PLS) (Liu et al., 2014), partial robust M-regression (PRM) (Fu et al., 2017), uninformative-variable-elimination genetic-algorithm PLS (UVE-GA-PLS) (Yuan et al., 2016) and orthogonal projections to latent structures (O-PLS) (Masoum et al., 2014) are such calibration methods based on statistical linear models which can be developed by using linear transfer functions to connect input and output variables. Multiple Linear Regression (MLR) is one of the most common techniques for
calibration and regression in chemistry; however, collinearity of variables may lead to unstable regression coefficients and could affect the accuracy of the model prediction. Stepwise-MLR is a multiple linear regression in which the variables are selected during the steps. This process starts by selecting the input variable that has the highest correlation with the output. Then, for a given selected variable $x_i$, its regression coefficient is determined and, if the coefficient is statistically significant (based on a t-test), the variable is retained (forward step). In the following step, the next variable is selected according to its highest partial correlation with the dependent variable.

Since the addition of new variables in the model can reduce the importance of previously selected variables, the significance of regression terms is checked after each new addition. If the terms are not significant, they are removed from the model (backward step). The inclusion and exclusion steps are repeated until no more improvement is achieved by adding new variables.

PCR solves the collinearity problem by generating new linearly independent variables and by reducing the noise through the elimination of the less important principal components. However, PCR is a two-step method and thereby has the risk that useful (predictive) information may end up in discarded principal components and that some noise will remain in the components used for regression (Dormann et al., 2012). However, the correlation between the response and the coefficients generated by the PCR with could be inappropriate because the selected PCs components are derived only based on capturing the highest variation in the data matrix of x-variables without considering the output values.

A distinct advantage in PLS techniques is that the latent variables are created based on the maximum covariance between the input data matrix $X$ and the response vector $y$. PLS is a widely used multivariate statistical technique, which can be implemented to develop a regression model.
to progress the prediction of chemical parameters. However, the interpretation of the regression coefficients of the model may be complicated due to the variability in $X$ that is orthogonal to $y$ (Xie et al., 2016). O-PLS is another multivariate calibration method with as main objective improving the interpretation of PLS models and reduce model complexity. O-PLS provides a way to remove variation from an input data set $X$ not correlated to the response $y$; in other words it removes variability in $X$ that is orthogonal to $y$ (Féraud et al., 2017). This can be performed by subtracting PLS components, orthogonal to $y$, from the $X$ data. Removing non-correlated variation in data prior to modelling is not only interesting from a predictive point of view, but also the ability to interpret the results of the models also improves. UVE (Uninformative Variable Elimination)-PLS has been developed in order to increase the predictive ability of the PLS model. Here, uninformative variables which do not contribute to the model, i.e. variables comparable to noise, are eliminated from the $X$-matrix. UVE-PLS can also be applied in order to remove uninformative samples from the calibration set before model construction. To obtain a robust and efficient PLS approach, Hoffmann et al. (2015) proposed partial robust M-regression (PRM). The main idea of PRM is to introduce a weight coefficient for each data object, which is inversely proportional to the distance between the point and the centre of the data cloud. Outliers are often objects located far from the centre, so their effect on the model can be limited by a small weight coefficient. One of the main advantages of the PRM technique is that no outlier detection is required prior to model building. Support vector machines (SVMs), Artificial neural network (ANN) and random forest (RF) (Segal, 2004) are relatively new pattern recognition methods well developed for multi-class problems.

4.-Chromatographic fingerprinting
A fingerprint can be defined as a characteristic profile reflecting the complex chemical composition of the analyzed sample. It can be obtained by spectroscopic or chromatographic techniques. Fingerprinting methods are provided by a variety of analytical techniques. They produce analytical signals containing information about the food composition in a non-selective way with as primary purpose authenticating or identifying food products (Pérez-Castaño et al., 2015). The results of numerous investigations demonstrate that fingerprint techniques, as non-targeted strategies, could provide powerful and effective tools for comprehensive food control (Riedl et al., 2015). Chemometrics, as a multivariate data analysis tool, is often coupled to fingerprints to assess the quality and to authenticate food and beverage products. Additionally, a comprehensive chemical fingerprint analysis can also detect non-label contamination compounds and unauthorized additives, or using of prohibited technological processes (Esslinger et al., 2014).

There are several analytical techniques available for fingerprint development including spectroscopic (e.g. NMR, MIR and NIR), chromatographic (GC and HPLC) and mass spectrometric methods. Among the various techniques, chromatography is found to be more informative in fingerprint analysis. Gas chromatography (GC) and high-performance liquid chromatography (HPLC) are the most commonly modes of chromatography. In the past decade GC-FID has been used successfully as a well-established, robust, cheap, and fast procedure to develop fingerprints. However, the main drawback of GC is that derivatization is frequently required (Raynie, 2018). In contrast, various compounds including both polar and non-polar in food products, can be analyzed by HPLC without derivatization. Currently, tremendous gains in efficiency, selectivity, throughput, speed and resolution are possible by shifting from HPLC to
UHPLC due to the reduced particle size of the stationary phase, while the sample size and mobile phases are also significantly reduced. The information obtained from chromatographic methods include peak intensity (area or height) and their positions (retention time) and, if MS is coupled, peak m/z. However, one may also use the entire profile instead of distinct peak information (Viaene and Heyden, 2018). Comprehensive two-dimensional (2D) chromatography is also an innovative technique that recently has been developed (Cortes et al., 2009). Since its presentation it has been used in a variety of research fields and has gained a lot of attention, mainly because of the presence of an additional chromatographic dimension, which increases the potential of fingerprinting and subsequently of authentication (Tranchida et al., 2004).

The vast and remarkable capabilities of chromatographic techniques, beside their high sensitivity, reproducibility, and robustness, enable them to obtain analytical signals with the highest information content. Therefore, these methods form systematic approaches to produce convenient and reliable fingerprints for food characterization, as well as for food authentication.

5.-Chromatographic analysis of food samples

Chromatographic techniques, equipped with various detectors, could be used to identify chemical adulteration markers, including fatty acids, oligosaccharides and tocopherols (Fanali et al., 2016).

Most methods are developed for carbohydrates, carotenoids, amino acids, phenolic or other organic compounds employing high performance liquid chromatography or gas chromatography (Herrero et al., 2017). The determination of triacylglycerols (TAGs), sterols and fatty acids provides the possibility to detection of oil adulteration and to characterize the blend’s
composition (Wernig et al., 2018). Characterization and differentiation of different vegetable oils can be achieved using the contents of fatty acid methyl esters (FAME), obtained through transesterification of vegetable oils determined by gas chromatography (Brodnjak-Vončina et al., 2005).

The field of metabolomics, focusing on global analysis of numerous low molecular compounds (metabolites), either targeted or non-targeted, in a biological sample, has recently drawn the attention of researchers in different areas (Naz et al., 2014) including food authentication (Riedl et al., 2015). Substantial advances in analytical techniques have made metabolomics rapidly grown, which facilitates the analysis of various metabolites with diverse physicochemical properties.

Triacylglycerols (TG) can be found in various vegetable oils; therefore, over the past decades extensive attention has been paid to their analysis. Papers discussing the chromatographic analysis of TGs in edible oils can be found in several papers (Indelicato et al., 2017). The methods include gas chromatography (Ruiz-Samblás et al., 2015), liquid chromatography (LC) (La Nasa et al., 2013), and supercritical fluid chromatography (SFC) (Lesellier et al., 2014). Each of these techniques has advantages and drawbacks and the choice depends on the objective of the analysis. The thermal instability of polyunsaturated triglycerides, associated with the temperature rise, is an important problem in gas chromatography. Therefore, long chain TGs cannot be separated individually by GC and SFC. In this context, GC and SFC of TGs seem to be complementary to HPLC (Sandra et al., 2010). Non-aqueous reversed-phase HPLC has been widely used for the analysis of TGs (Cajka and Fiehn, 2014). One particularly interesting feature
of HPLC is its ability to separate TGs directly without derivatization. However, due to insolubility, higher molecular weight TGs cannot be analysed by HPLC.

Typically, GC is useful for analysing non-polar and semi-polar, volatile and semi-volatile chemicals. Without chemical derivatization, GC is often used for the analysis of sterols, oils, fatty acids, aroma components and off-flavours (Ben Brahim et al., 2018; Grasso et al., 2016; Henna Lu and Tan, 2009). HPLC can be useful for separating all types of organic chemicals regardless their polarity or volatility. Nevertheless, because of the advantages of GC, HPLC has been mostly used for the analysis of polar, thermolabile, and/or non-volatile chemicals. Chemical derivatization of polar chemicals, such as amino acids, hydroxyl (poly) carboxylic acids, fatty acids, phenolic compounds, sugars and vitamins is also performed to permit their analysis by GC (Hammi et al., 2018). Only the non-volatile compounds, such as inorganic salts, proteins, polysaccharides, nucleic acids, and other large molecular weight organics, are outside the realm of GC (Lehotay and Hajšlová, 2002).

An estimation of chromatographic techniques used in food applications can be made fairly easily using PubMed, a free literature-search database provided by the US National Institutes of Health (Lehotay and Hajšlová, 2002). PubMed covers the main analytical and application journals, but it is designed for the biomedical researcher. However, it serves the purpose of this review to display trends. Fig. 1 gives the number of publications in the PubMed database in relation to the main food application category, chromatographic technique, and year of the publication. Searches were limited by the terms, “chromatography” AND “food” AND “adulteration”. The
publication rate in this field has been increased significantly from 1980 to 2017, which indicates
the increased attention of scientists to the importance of food authentication.

5.1.-Gas Chromatography

Advances in analytical chemistry and instrumentation in recent years may increase the ease with
which these chromatographic techniques can be applied in industrial food authentication.
Development and assessment of methods for the detection of adulteration of olive oil samples
using GC focussed on the comparison of their fatty acid composition and chemometric analysis
(Gamazo-Vázquez et al., 2003; Ollivier et al., 2003). Yang et al. (2013) proposed a GC-MS
method for the detection and identification of extra virgin olive oil adulteration with four types of
oils, including corn, peanut, rapeseed, and sunflower oils. The fatty acid content besides the
ratios of linoleic/linolenic acid and oleic/linoleic acid, the total saturated fatty acids (SFAs),
polyunsaturated fatty acids (PUFAs), monounsaturated fatty acids (MUFAs), MUFAs/PUFAs)
were determined. Univariate analysis demonstrated that higher contents of eicosanoic acid,
docosanoic acid, tetracosanoic acid, and SFAs were the peculiarities of peanut adulteration
while, higher levels of linolenic acid, 11-eicosanoic acid, erucic acid, and nervonic acid were the
characteristics of rapeseed adulteration. PLS-LDA was also used in this work for the detection of
adulteration with a 1% detection limit and 90% prediction ability.

Mansor et al. (2011) investigated the application of GC with surface acoustic wave detector (GC-
SAW system), combined with chemometrics, to analyze the presence of lard in virgin coconut
oil. Different lard and virgin coconut oil mixtures ranging from 1% to 50% (v/v) were subjected
to a fast GC-SAW system. In the chromatograms, ten peaks were identified as the adulterant
Peaks. PLS regression was able to predict correctly the lard contents in virgin coconut oil, while LDA could classify virgin coconut oil and the samples adulterated with lard.

Jabeur et al. (2014) used the fatty-acids composition as an indicator of purity and the linolenic acid content as a parameter for the detection of extra virgin olive oil fraud with 5% of soybean oil. Their results showed that the sterols profile is almost crucial in clarifying the adulteration of olive oils with other cheaper ones, e.g. soybean, corn or sunflower oil. They used LDA as a tool for the fast detection of extra-virgin olive oil adulteration.

Esteki et al. (2017) solved the detection of the adulteration of almond powder samples with apricot kernel by GC fatty-acid fingerprinting combined with chemometric methods, including PCA, PCA-LDA, PLS and least square support vector machine (LS-SVM). Different ratios of almond and apricot kernel samples were mixed to give desired proportions from 10 to 90% w/w. PCA was used as exploratory data analysis and PCA-LDA for the classification of almonds, apricot kernels, and their mixtures. PLS and LS-SVM were used as regression methods for the determination of the adulteration ratios in almond. The root mean square error (RMSE) and the coefficient of determination ($R^2$) for the validation data set obtained for LS-SVM were 2.3 and 0.995, respectively, which indicate the feasibility of the method for testing almond adulteration.

The fatty-acid composition of olive oil samples obtained by flash gas chromatography electronic nose were subjected to chemometric methods to determine their authenticity and geographical origin (Melucci et al., 2016). Similar analyses of Italian and Argentinian olive oils enabled their classification based on the triacylglycerol components (Ollivier et al., 2003). Gutiérrez et al
(2009) used GC to determine triacylglycerol profiles in milk and non-milk fat. In order to detect and quantify non-milk fat in milk fat, the triacylglycerol profiles were subjected to LDA. Raw milk fat from the central region of Mexico and ultra-pasteurized milk fat from three factories, as well as pork fat lard, bovine tallow, fish oil, peanut-corn-oil, olive oil and soy oil were analyzed. The samples of raw milk fat were adulterated with non-milk fats in different proportions (0-20%). The first function obtained from the LDA allowed the correct classification of 94.4% of the adulterated samples. The triacylglycerol profiles of the ultra-pasteurized milk fats were evaluated with the LDA model, demonstrating that one factory added non-milk fat to its products.

Differentiation of monovarietal Sicilian olive oil samples from the same Italian region (southwestern Sicily) obtained from four monovarietal cultivars was accomplished using LDA analysis of 10 GC peaks with classification accuracy of 95% (Mannina et al., 2003). This study showed that GC fingerprinting in conjunction with multivariate data analysis achieved a better clear separation between cultivars than NMR multivariate data analysis. Chemometric methods have also been successfully applied to GC data for the authentication of several other food products and beverages, such as coffee (Toledo et al., 2014) and fruit juices ( Cuevas et al., 2017).

Reid et al. (2004) combined GC and chemometrics to differentiate between apple-juice samples on the basis of apple variety and applied heat treatment. The chromatographic data were subjected to PLS regression and PCA-LDA. PLS gave 92.5% correct classifications of the apple juice samples, according to both variety and heat treatment, while using PCA-LDA, 87.5% and
80% of the samples were correctly classified according to apple variety and heat treatment, respectively.

Winterová et al. (2008) determined different concentrations of volatile compounds, fatty acids and stable isotope ratios to assess the authenticity of fruit spirits by GC combined with LDA. The results show that especially the isotope ratios can be used for discriminating fruit spirits and others spirits, i.e. those made from maize, cane sugar, beet sugar, grain, potato, or synthetic alcohol. It was also possible to distinguish the respective spirits derived from one type of fruit, such as sweet cherry brandy, sour cherry brandy, apple brandy, apricot brandy, pear brandy, or plum brandy.

The gas chromatographic determination of volatile constituents can also be used for checking the authenticity of food products containing volatile organic compounds. González-Arjona et al. (2006) carried out the volatile compound analysis of 52 commercialized whiskeys by GC-MS after liquid-liquid extraction with dichloromethane. Multivariate data analysis, including LDA, k-nearest neighbors (KNN), SIMCA, procrustes discriminant analysis (PDA), and artificial neural network techniques, involving multilayer perceptrons (MLP) and probabilistic neural networks (PNN) were used for classification. The constructed models were validated by considering the number of false positives (FPs) and false negatives (FNs) of each class associated to the prediction set. Artificial neural networks showed the best results because of their intrinsic nonlinear properties. Both MLP and PNN, achieved 100% selectivity and 100% specificity for categories. KNN as a nonparametric method provides also reasonable results. PDA produced 100% sensitivity and 100% specificity by selecting nine principal components for
class modeling. LDA showed a lower classification performance, because of building linear borders between classes that does not apply well in this situation.

Lin et al. (2013) developed an analytical method to discriminate oolong tea varieties and adulterated samples based on volatile compounds. Oolong tea samples of five varieties, including Tieguanyin, Maoxie, Jinguanyin, Benshan, and Huangjingui, were analyzed by headspace solid phase microextraction (HS-SPME) coupled to GC–MS. The chromatographic fingerprint of the major volatile compounds varied significantly with variety, indicating that the aromatic profile could play a significant discriminating role for tea cultivars with very close origin. PCA was applied on the chromatographic fingerprints and showed that the feature of variety dominated over the other features, such as producing region and quality. By using stepwise linear discriminant analysis (S-LDA), 18 volatile compounds with the best discriminating potential were selected, and 4 discriminant functions enabled to discriminate five oolong varieties with a sensitivity of 100%.

Lorenzo et al. (2002) applied direct coupling of a headspace sampler with a mass spectrometer (HS-MS) for the rapid detection olive oil adulterations. Samples of olive oils were mixed with different concentration ranges of sunflower and olive-pomace oil. Volatile compound patterns in the original and mixed samples were obtained. LDA was sufficient to differentiate the adulterated from the non-adulterated oils and to discriminate the type of adulteration. The classification accuracy was 100% for the training set and the prediction set. Additional applications of GC in food authentication are represented in Table 1.
5.2.-High Performance Liquid Chromatography

The number of new applications of HPLC in food authenticity in recent years is low. Nevertheless some applications have been reported for the technique (Table 2). Detection of the adulteration of ovine and caprine cheese and milk with bovine milk at levels as low as 2% v/v has been accomplished using HPLC analysis of the whey protein β-lactoglobulin (Chen et al., 2004). Discrimination of bovine, ovine and caprine milks using reversed-phased HPLC was also performed (Veloso et al., 2002). Additional applications of HPLC in detection of milk adulteration are given in Table 2.

Wines from different denominated areas of origin in the Canary Islands (Spain) were correctly classified using HPLC analysis determining 15 polyphenol concentrations combined with PCA and LDA (Rodríguez-Delgado et al., 2002). Spanish table wines were correctly discriminated using PCA, cluster analysis and LDA of the RP-HPLC data obtained for eight biogenic amine compounds in the samples (Roberto Romero et al., 2002). HPLC with UV detection combined with PCA and LDA was also used for the geographical classification of some Australian wines, with 75%-100% accuracy (Bellomarino et al., 2009).

The triglyceride and tocopherol composition of green and roasted coffee beans in coffee samples, based on RP-HPLC analysis, was combined with PCA and LDA to discriminate different varieties (González et al., 2001). In another report, HPLC–UV–Vis was used to detect adulterations in roasted and ground coffee, using PCA (Domingues et al., 2014b). Mass spectrometry using an atmospheric pressure chemical ionization interface (APCI-MS) with HPLC was used for the triacylglycerol evaluation to characterize olive oil adulteration by...
soybean oil, using PCA (Fasciotti and Pereira Netto, 2010). Triglyceride composition determined by HPLC analysis was used to classify Spanish olive oil varieties applying PCA and DA (Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004)(Aranda et al., 2004). Discrimination of extra virgin olive oils was performed based on diacylglycerols, triacylglycerols and sterols composition of samples. The data obtained with HPLC–MS, and PCA and LDA models were well suited to classify the samples. HPLC analysis of marker proteins was investigated for durum wheat and wheat cultivars commonly used as adulterants in durum wheat (Bonetti et al., 2004). Using this approach, detection of adulteration of durum wheat cultivars in high-quality pasta with wheat cultivars is possible even at levels as low as 5% w/w.

Pauli et al (2014) determined carbohydrates as the main nutrients of grains as chemical biomarkers to evaluate the quality of coffee. They used soybeans and wheat as sources of
adulteration. High performance anion-exchange chromatography with pulsed amperometric
detection was used to characterize the pure roasted coffee bean and the adulteration profiles
based on total carbohydrates composition. The proposed PCA and LDA models were accurate
and reliable to recognize and predict of different mixture proportions and to distinguish genuine
coffee. Pure roasted coffee contains high levels of galactose and mannose, which, glucose and
fructose are considered chemical markers for wheat and soybean adulteration, respectively.

Tavares et al. (2016) investigated the effectiveness of tocopherols as markers for coffee
adulteration with maize, based on their ability to detect adulterations with coffee husks, the main
residue of coffee pre-processing. Roasted husks, cleaned roasted husks, and roasted maize were
used as adulterants in this study. Extracted lipids were analyzed by normal-phase HPLC (with
florescence detection), and the tocopherol amounts found were analyzed by PCA, LDA and
SIMCA. Based on the tocopherol profiles maize adulterant is detectable at levels above 10%. For
heavy adulterations, discrimination of husks and cleaned husks is also possible.

Vaclavik et al. (2012) explored the feasibility of HPLC–MS technique employing a hybrid triple
quadrupole/linear ion trap (QqQ/LIT) mass analyzer, for metabolomic-based authentication of
fruit juices (orange, grapefruit, apple, and their mixtures) representing different price categories.
Comprehensive metabolomic fingerprinting of several fruit juice types, prepared from expensive
(oranger) or relatively low-priced (apple, grapefruit) fruits was done. Following the automated
data mining and pre-treatment step, the suitability of the data for authentication, i.e.,
classification of fruit juices and adulteration detection, was assessed with PCA and LDA. The
LDA classification model, constructed and validated employing a highly variable sample set,
was able to reliably detect 15% addition of apple or grapefruit juice to orange juice. Comprehensive, non-targeted HPLC–MS metabolomic fingerprinting coupled to chemometric methods was then formed to be a suitable tool for fruit-juice authenticity testing.

Jabeur et al. (2014) used linolenic acid content as a parameter for the detection of extra-virgin olive oil (EVOO) fraud with 5% of soybean oil. For this purpose, analyses of fatty acid, sterol, and triacylglycerol profiles were performed using gas and liquid chromatography. The objective of their research was the detection of Chemlali EVOO adulteration with soybean, corn, and sunflower oils. Application of LDA, after feature selection, was sufficient to differentiate Chemlali EVOO and all adulterated extra-virgin olive oils. The success was 100% in classification and close to 100% in prediction.

Sabir et al. (2017) combined HPLC fingerprint analysis with chemometrics to discriminate red and white rice bran, grown in Indonesia. Cycloartenol ferulate, cyclobranol ferulate, campesterol ferulate and β-sitosterol ferulate are the main components for characterizing the rice bran. PCA and LDA analysis were successfully used for discriminating the two classes with good predictive ability.

Agozzino et al. (2010) used a large number of certified samples (84) of Sicilian olive oils arising from eight cultivars. HPLC/MS analysis was performed without sample preparation or chemical derivatization. LDA has been used for the olive oil cultivar traceability. In addition, changes in the composition of olive oils glyceride components leads to easy discrimination between fresh and one-year-old oils samples.
6.-Conclusions

Over the past decade, due to recent crises and food industry scandals that have seriously affected consumer confidence, the quality of food products and production methods has increased significantly. For this reason, several rapid analytical techniques have been tried as methods for empirically assessing the authenticity and the detection of adulteration. The purpose of this article is to review the most recent applications of chromatographic methods coupled to chemometrics in food identity validation and food authentication. In this way, more than seventy scientific publications were thoroughly examined, focusing on targeted and non-targeted analysis for authentication, including supervised and un-supervised pattern recognition methods.

The authors believe that the use of chromatographic fingerprinting in food characterization and validation, including detection of adulteration, traceability and classification purposes, is perfectly appropriate. As expected, the determination of a variety of organic compounds was
mainly carried out using chromatographic separation techniques (HPLC or GC) along with
different detection methods. Multivariate qualitative methods are effective tools for dealing with
food authentication problems that cannot be solved with limited number of variables, because the
required and obtained response is complex in nature and furthermore no univariate signal acts as
an unambiguous marker. The review of resources indicates a rapid development of the latest
chromatographic techniques combined with chemometrics multivariate data analysis methods.

The results indicate that chemometric analysis of the chromatographic data offers reliability to
detect suspected fraud samples and appropriate way for classification of different food products.
These methods present special potential in routine daily analysis as well as in research about new
types of adulterants in a variety of food products such as edible oils, beverages, honey and dairy
products. This article is also a valuable resource for food scientists who want to use different
methods of chromatography along with chemometrics methods for food authentication.


doi:10.1016/j.meatsci.2010.06.001

doi:10.1016/j.jpba.2011.02.001


doi:10.1016/j.talanta.2009.08.001


doi:10.1021/jf034881f


analysis of the volatile composition of honeys from Brazilian stingless bees by static headspace GC–MS. Food Res. Int. 102, 536–543. doi:10.1016/j.foodres.2017.09.036


Food Control 14, 463–467. doi:10.1016/S0956-7135(02)00102-0


Jiao, J., Ding, N., Shi, T., Chai, X., Cong, P., Zhu, Z., 2011. Study of chromatographic...
doi:10.1080/00032711003783044
Fast-HPLC fingerprinting to discriminate olive oil from other edible vegetable oils by
multivariate classification methods. J. AOAC Int. 100, 345–350. doi:10.5740/jaoacint.16-0411
One input-class and two input-class classifications for differentiating olive oil from other
edible vegetable oils by use of the normal-phase liquid chromatography fingerprint of the
doi:10.1016/j.foodchem.2016.10.103
Fas
HPLC fingerprinting to discriminate olive oil from other edible vegetable oils by
multivariate classification methods. J. AOAC Int. 100, 345–350. doi:10.5740/jaoacint.16-0411
One input-class and two input-class classifications for differentiating olive oil from other
edible vegetable oils by use of the normal-phase liquid chromatography fingerprint of the
doi:10.1016/j.foodchem.2016.10.103
One input-class and two input-class classifications for differentiating olive oil from other
edible vegetable oils by use of the normal-phase liquid chromatography fingerprint of the
doi:10.1016/j.foodchem.2016.10.103
chromatography/flame ionization detector-based metabolite fingerprinting for
doi:10.1016/j.jbiosc.2015.03.005
blends for authentication of Asian palm civet coffee (Kopi Luwak) via metabolomics: A
Jumtee, K., Komura, H., Bamba, T., Fukusaki, E., 2011. Predication of Japanese green tea (Sen-
cha) ranking by volatile profiling using gas chromatography mass spectrometry and
in honey from different regions of Turkey. Food Chem. 123, 41–44.
doi:10.1016/j.foodchem.2010.03.123
Olive oil volatile compounds, flavour development and quality: A critical review. Food
Kamal, M., Karoui, R., 2015. Analytical methods coupled with chemometric tools for
determining the authenticity and detecting the adulteration of dairy products: A review.
52, 33–39. doi:10.1021/jf034985q
doi:10.5851/kosfa.2016.36.2.206
commercial cheeses from fatty acid profiles and phytosterol contents obtained by gc and
Kuš, P.M., van Ruth, S., 2015. Discrimination of Polish unifloral honeys using overall PTR-MS
and HPLC fingerprints combined with chemometrics. LWT - Food Sci. Technol. 62, 69–75.
doi:10.1016/j.lwt.2014.12.060


Olaoye, O.A., Ubbor, S.C., Lawrence, I.G., 2018. Assessment of the Effect of Different...


Spectrometry and Chemometrics. Foods 6, 13. doi:10.3390/foods6020013


Roberto Romero, Mercedes Sánchez-Viñas, *, Domingo Gázquez, and, Bagur, M.G., 2002. Characterization of Selected Spanish Table Wine Samples According to Their Biogenic Amine Content from Liquid Chromatographic Determination. doi:10.1021/JF025514R


Sharangi, A.B., 2009. Medicinal and therapeutic potentialities of tea (Camellia sinensis L.) - A


Fig. 1. Publications using GC/HPLC combined with chemometrics in food authentications, based on PubMed information.
Table 1. Applications of gas chromatography coupled with chemometrics in food authentication.

<table>
<thead>
<tr>
<th>Food</th>
<th>Purpose of analysis</th>
<th>Analytical Technique</th>
<th>Chemometrics methods</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Edible Oils</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coconut oil</td>
<td>Detection of lard adulteration in virgin coconut oil</td>
<td>GC-SAW</td>
<td>DA</td>
<td>(Mansor et al., 2011)</td>
</tr>
<tr>
<td>Olive oil</td>
<td>Classification of olive oil varieties</td>
<td>HT/GC-(IT) MS</td>
<td>PCA, HCA, PLS-DA</td>
<td>(Ruiz-Samblás et al., 2011)</td>
</tr>
<tr>
<td>Edible vegetable oil blends</td>
<td>Quantification of olive oil in oil blends</td>
<td>GC-(IT)MS</td>
<td>RF</td>
<td>(Ruiz-Samblás et al., 2014)</td>
</tr>
<tr>
<td>Extra-virgin olive oil</td>
<td>Adulteration detection (soybean, corn, sunflower oil)</td>
<td>GC-FID</td>
<td>LDA</td>
<td>(Jabeur et al., 2014)</td>
</tr>
<tr>
<td>Flaxseed oil</td>
<td>Adulteration detection</td>
<td>GC-MS</td>
<td>SVM</td>
<td>(Sun et al., 2015)</td>
</tr>
<tr>
<td>Olive oil</td>
<td>Distinction and detection of adulteration with cheap vegetable oils</td>
<td>GC-FID</td>
<td>LDA</td>
<td>(Jabeur et al., 2016)</td>
</tr>
<tr>
<td>Extra virgin olive oils</td>
<td>Discrimination of geographical origin of extra virgin olive oils</td>
<td>SPME/GC-MS</td>
<td>PCA, PLS-DA</td>
<td>(Melucci et al., 2016)</td>
</tr>
<tr>
<td>Vegetable oils</td>
<td>Classification of edible vegetable oils</td>
<td>FGC E-nose</td>
<td>SVM</td>
<td>(Li et al., 2016)</td>
</tr>
<tr>
<td>Extra virgin olive oil</td>
<td>Olive oil profiling</td>
<td>HS-GC-IMS</td>
<td>KNN, LDA</td>
<td>(Gerhardt et al., 2017)</td>
</tr>
<tr>
<td>Olive oil</td>
<td>Detection of olive oil adulteration with pomace olive oil</td>
<td>GC-FID</td>
<td>LDA</td>
<td>(Jabeur et al., 2017)</td>
</tr>
<tr>
<td>Dressing and olive oil blends</td>
<td>Quantification of extra virgin olive oil in dressing and edible oil blends</td>
<td>GC-FID</td>
<td>PLS-R</td>
<td>(Pérez-Castaño et al., 2018)</td>
</tr>
<tr>
<td><strong>Dairy products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>Detection of cow milk in donkey milk</td>
<td>HRGC-FID</td>
<td>LDA, ANN</td>
<td>(Cossignani et al., 2011)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Discrimination of commercial cheeses</td>
<td>GC-MS</td>
<td>PCA</td>
<td>(Kim et al., 2014)</td>
</tr>
<tr>
<td>Cheddar cheese</td>
<td>Classification and prediction of Cheddar cheese maturity</td>
<td>APCI-MS/GC-MS</td>
<td>PLS-DA</td>
<td>(Gan et al., 2016)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Detection for non-milk fat in dairy product</td>
<td>GC-FID</td>
<td>PCA</td>
<td>(Kim et al., 2016)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Authentication Cheeses</td>
<td>GC-FID</td>
<td>PCA</td>
<td>(Tociu et al., 2013)</td>
</tr>
</tbody>
</table>

58
<table>
<thead>
<tr>
<th>Food Product</th>
<th>Purpose of Analysis</th>
<th>Methodology</th>
<th>Identification Techniques</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk fat</td>
<td>Authentication of the organic status</td>
<td>GC-FID</td>
<td>MLR</td>
<td>(Rodríguez-Bermúdez et al., 2018)</td>
</tr>
</tbody>
</table>

**Honey**

<table>
<thead>
<tr>
<th>Honey</th>
<th>Geographical origin</th>
<th>GC-TOF-MS</th>
<th>LDA, DPLS, SVM</th>
<th>(Stanimirova et al., 2010)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honey</td>
<td>Differentiation of Manuka Honey from Kanuka Honey</td>
<td>HS-SPME-GC/MS</td>
<td>PCA</td>
<td>(Beitlich et al., 2014)</td>
</tr>
<tr>
<td>Honey</td>
<td>Classification of honey based on floral source</td>
<td>SPME-GC-MS</td>
<td>PCA, HCA</td>
<td>(Revell et al., 2014)</td>
</tr>
<tr>
<td>Honey</td>
<td>Geographical classification of Brazilian honey</td>
<td>HS-GC-MS</td>
<td>PCA</td>
<td>(de Lima Morais da Silva et al., 2017)</td>
</tr>
<tr>
<td>Honey</td>
<td>Geographical classification of bracatinga honeydew honey</td>
<td>GC-MS</td>
<td>CA</td>
<td>(Azevedo et al., 2017)</td>
</tr>
<tr>
<td>Honey</td>
<td>Classification of the botanical origin of Chinese honey</td>
<td>GC-MS</td>
<td>PCA</td>
<td>(Chen et al., 2017)</td>
</tr>
<tr>
<td>Honey</td>
<td>Assessment of the authenticity of honey</td>
<td>HS-GC-IMS</td>
<td>PCA-LDA, kNN</td>
<td>(Gerhardt et al., 2018)</td>
</tr>
</tbody>
</table>

**Fruit juices**

<table>
<thead>
<tr>
<th>Juice Type</th>
<th>Purpose of Analysis</th>
<th>Methodology</th>
<th>Identification Techniques</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apple juices</td>
<td>Differentiation of apple juice samples</td>
<td>SPME-GC</td>
<td>PCA, PLS, LDA</td>
<td>(Reid et al., 2004)</td>
</tr>
<tr>
<td>Fruit spirits</td>
<td>Assessment of the authenticity of fruit spirits</td>
<td>GC</td>
<td>LDA</td>
<td>(Winterová et al., 2008)</td>
</tr>
<tr>
<td>Citrus fruits peels</td>
<td>Classification by fruit types</td>
<td>GC-MS</td>
<td>MCR</td>
<td>(Parastar et al., 2012)</td>
</tr>
<tr>
<td>Apple juice</td>
<td>Discrimination of juices by variety and geographical origin</td>
<td>GC-MS</td>
<td>PCA, LDA, SLDA, PLS-DA</td>
<td>(Guo et al., 2012)</td>
</tr>
<tr>
<td>Orange juices</td>
<td>Assessment of premium organic orange juices authenticity</td>
<td>HS-SPME-GC-MS</td>
<td>HCA, PLS-DA</td>
<td>(Cuevas et al., 2017)</td>
</tr>
</tbody>
</table>

**Alcoholic beverages**

<table>
<thead>
<tr>
<th>Beverage Type</th>
<th>Purpose of Analysis</th>
<th>Methodology</th>
<th>Identification Techniques</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcoholic beverages and fermented products</td>
<td>Varietal origin of wines and/or grapes</td>
<td>SBSE-TD-GC-MS</td>
<td>PCA, LDA</td>
<td>(Tredoux et al., 2008)</td>
</tr>
<tr>
<td>Alcoholic beverages and fermented products</td>
<td>Varietal origin of wines and/or grapes</td>
<td>GC-MS</td>
<td>PLS</td>
<td>(Skogerson et al., 2009)</td>
</tr>
<tr>
<td>Product Type</td>
<td>Process Description</td>
<td>Methodologies Used</td>
<td>Authors</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Beer</td>
<td>Classification by types, breweries and tastes</td>
<td>GC-FID, PCA, HCA</td>
<td>(Jiao et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Wines</td>
<td>Differentiation of wine brands</td>
<td>GC-MS, PCA, HCA, CTA, HCA, PCA</td>
<td>(Dall’Asta et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Wines</td>
<td>Authentication of geographical origin</td>
<td>GC-MS, PLS-DA, SIMCA, NPLSDA, NSIMCA</td>
<td>(Salvatore et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>White wine</td>
<td>Authentication of grape varieties</td>
<td>HS-SPME-GC-MS, PCA, CTA</td>
<td>(Springer et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>White wine</td>
<td>Authentication of grape varieties</td>
<td>HS-SPME-GC-MS, PCA, CTA</td>
<td>(Springer et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Cherry wines</td>
<td>Discrimination</td>
<td>GC-MS, PCA</td>
<td>(Xiao et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Rice wine</td>
<td>Classification of Chinese rice wine according to geographic origin and wine age</td>
<td>SBSE-TD-GC-MS, PCA, CA</td>
<td>(Xiao et al., 2015)</td>
<td></td>
</tr>
<tr>
<td>Wine</td>
<td>Differentiation of wines according to grape variety and geographical origin</td>
<td>PME-GC/MS, PCA, LDA</td>
<td>(Ziolkowska et al., 2016)</td>
<td></td>
</tr>
<tr>
<td>Coffee</td>
<td>Discrimination of coffee varieties</td>
<td>GC-MS, PLS-DA</td>
<td>(Hovell et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>Coffee</td>
<td>Modelling of sensory prediction</td>
<td>GC-FID, PLS</td>
<td>(Ribeiro et al., 2012)</td>
<td></td>
</tr>
<tr>
<td>Coffee</td>
<td>Detection and quantification of coffee adulteration</td>
<td>SPME-GC-MS, PLS</td>
<td>(Toledo et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Coffee</td>
<td>Differentiation of geographic origin of coffee beans</td>
<td>TD-GC-MS, PCA</td>
<td>(Thimmaraju and Yao, 2015)</td>
<td></td>
</tr>
<tr>
<td>Asian coffee</td>
<td>Authentication of Asian palm civet coffee (Kopi Luwak)</td>
<td>GC-MS, PCA</td>
<td>(Jumhawan et al., 2015)</td>
<td></td>
</tr>
<tr>
<td>Coffee beans</td>
<td>Authentication Asian palm civet coffee</td>
<td>GC/FID, GC/MS, PCA, OPLS-DA</td>
<td>(Jumhawan et al., 2016)</td>
<td></td>
</tr>
<tr>
<td>Tea</td>
<td>Classification of Japanese green tea by qualities</td>
<td>GC-(TOF)-MS, PCA, PLS</td>
<td>(Pongsuwan et al., 2007)</td>
<td></td>
</tr>
<tr>
<td>Green tea</td>
<td>Prediction of product quality</td>
<td>GC-MS, PLS</td>
<td>(Pongsuwan et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>Green tea</td>
<td>Classification of Japanese green tea by quality</td>
<td>GC-MS, PLS-DA, PLS</td>
<td>(Jumtee et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Tea</td>
<td>Geographical origin of tea</td>
<td>GC-MS, PCA, LDA</td>
<td>(Ye et al., 2012)</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Methodology</td>
<td>Technique</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Oolong teas</td>
<td>Discrimination of oolong tea</td>
<td>HS-SPME/GC-MS</td>
<td>PCA S-LDA (Lin et al., 2013)</td>
<td></td>
</tr>
<tr>
<td><strong>Meat and meat products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meat, fish and seafood</td>
<td>Detection of mechanically recovered meat in food products</td>
<td>GC–MS PCA OPLS-DA</td>
<td>OPLS-DA (Surowiec et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Fresh and frozen pork</td>
<td>Quality control</td>
<td>E-nose based on ultra-fast gas chromatography (UFGC) PCA ANN</td>
<td>(Górská-Horczyczak et al., 2017)</td>
<td></td>
</tr>
<tr>
<td><strong>Other food products</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balsamic vinegars</td>
<td>Classification of balsamic vinegars</td>
<td>GC-MS PCA SIMCA</td>
<td>SIMCA (Cocchi et al., 2004)</td>
<td></td>
</tr>
<tr>
<td>Balsamic vinegars</td>
<td>Prediction of sensory attributes</td>
<td>GC-FID NPLS unfold-PLS PCA CA</td>
<td>(Durante et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Rice brand wax</td>
<td>Prediction of product quality</td>
<td>GC-MS PCA SIMCA</td>
<td>PCA (Wang et al., 2007)</td>
<td></td>
</tr>
<tr>
<td>Balsamic vinegars</td>
<td>Evolution of the volatile compounds during ageing</td>
<td>GC-FID PCA</td>
<td>PCA (Cocchi et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>Apple</td>
<td>Classification of apple varieties</td>
<td>GC-MS PCA PLS-DA LDA SIMCA PCA</td>
<td>PLS-DA LDA SIMCA PCA (Aprea et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Tomato</td>
<td>Geographical Origin of tomatoes</td>
<td>GC-MS PCA SIMCA</td>
<td>PCA (Feudo et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Mushrooms</td>
<td>Discrimination of wild mushrooms species</td>
<td>GC-MS PCA SIMCA</td>
<td>PCA (Malheiro et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>Oxidized tallow</td>
<td>QC of oxidized tallow</td>
<td>SPME-GC-MS PCA</td>
<td>PCA (Song et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>Pineapple samples</td>
<td>Authentication of maturity stages</td>
<td>HS-SPME-GC-MS PCA</td>
<td>PCA (Steingass et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>White apricot and almond</td>
<td>Detection of adulteration</td>
<td>GC-MS PCA SIMCA</td>
<td>PCA (Tian et al., 2014)</td>
<td></td>
</tr>
<tr>
<td>Almond powder</td>
<td>Detection and quantification of apricot kernel in almond powder</td>
<td>GC-FID PCA PLS PCA-LDA SVM</td>
<td>(Esteki et al., 2017)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Applications of high performance liquid chromatography and chemometrics in food authentication.

<table>
<thead>
<tr>
<th>Food</th>
<th>Purpose of analysis</th>
<th>Analytical Technique</th>
<th>Chemometrics methods</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edible oils</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olive oil</td>
<td>classification of the commercial virgin olive oil samples</td>
<td>HPLC</td>
<td>PCA</td>
<td>(F. Aranda et al., 2004)</td>
</tr>
<tr>
<td>Olive oil</td>
<td>Characterization of olive oil adulteration by soybean oil</td>
<td>HPLC-APCI-MS-MS</td>
<td>PCA</td>
<td>(Fasciotti and Pereira Netto, 2010)</td>
</tr>
<tr>
<td>Olive oils</td>
<td>Determination of the cultivar and aging of Sicilian olive oils</td>
<td>HPLC/MS</td>
<td>LDA</td>
<td>(Agozzino et al., 2010)</td>
</tr>
<tr>
<td>Edible vegetable oils</td>
<td>Quantification of olive oil in blends with vegetable oils</td>
<td>HPLC-CAD</td>
<td>PLS, iPLS</td>
<td>(De La Mata-Espinosa et al., 2011)</td>
</tr>
<tr>
<td>Edible vegetable oils</td>
<td>Discrimination of olive oil from vegetable oils</td>
<td>HPLC-CAD</td>
<td>PCA</td>
<td>(Tomasi et al., 2011)</td>
</tr>
<tr>
<td>Edible vegetable oils</td>
<td>Discrimination of the virgin olive oils from edible vegetable oils</td>
<td>HPLC-UV</td>
<td>HCA, PCA</td>
<td>(Bagur-González et al., 2015)</td>
</tr>
<tr>
<td>Extra virgin olive oil</td>
<td>Adulteration detection (soybean, corn, sunflower oil)</td>
<td>HPLC</td>
<td>LDA</td>
<td>(Jabeur et al., 2014)</td>
</tr>
<tr>
<td>Extra virgin olive oil</td>
<td>Discrimination of the PDO Sabina olive oil</td>
<td>HPLC-DAD</td>
<td>PLS-DA, biPLS-GA</td>
<td>(Nescatelli et al., 2014)</td>
</tr>
<tr>
<td>Palm oil</td>
<td>Differentiation of geographical origin</td>
<td>HPLC-CAD</td>
<td>PCA</td>
<td>(Pérez-Castaño et al., 2015)</td>
</tr>
<tr>
<td>Edible vegetable oils</td>
<td>Discrimination of olive oil from edible oils</td>
<td>HPLC-UV</td>
<td>PLS-DA, SIMCA</td>
<td>(Jiménez-Cardel et al., 2017a)</td>
</tr>
<tr>
<td>Crude palm oil</td>
<td>Authentication of the geographical origin</td>
<td>HPLC-UV and HPLC-CAD</td>
<td>PLS-DA</td>
<td>(Obisesan et al., 2017)</td>
</tr>
<tr>
<td>vegetable oils</td>
<td>Discrimination of olive oil from edible oils</td>
<td>NPLC</td>
<td>kNN, PLS-DA, SVM</td>
<td>(Jiménez-Cardel et al., 2017b)</td>
</tr>
<tr>
<td>Dairy products</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk and cheese</td>
<td>Detection of adulteration of ovine, caprine and bovine milk</td>
<td>HPLC-DAD</td>
<td>PCA</td>
<td>(Rodríguez et al., 2010)</td>
</tr>
<tr>
<td>Infant Milk</td>
<td>Detection and differentiation of detergent powder adulteration in infant milk formula</td>
<td>LC-QTOF-MS</td>
<td>PCA, PLS</td>
<td>(Tay et al., 2013)</td>
</tr>
<tr>
<td>Milk Powder</td>
<td>Detection of adulteration of skim milk powder with foreign proteins</td>
<td>UHPLC</td>
<td>SIMCA</td>
<td>(Jablonski et al., 2014)</td>
</tr>
<tr>
<td>Fruit juices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apricot juices</td>
<td>Characterization of Italian commercial apricot juices</td>
<td>HPLC</td>
<td>PCA</td>
<td>(Versari et al., 2008)</td>
</tr>
<tr>
<td>Orange juice</td>
<td>Classification of fruit juice and</td>
<td>HPLC-MS</td>
<td>PCA</td>
<td>(Vaclavík et al., 2017)</td>
</tr>
<tr>
<td>Food Product</td>
<td>Adulteration Detection</td>
<td>Analytical Method</td>
<td>Multivariate Techniques</td>
<td>Ref.</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------</td>
<td>-------------------</td>
<td>-------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Pineapple, orange, grapefruit, apple juice</td>
<td>Rapid detection of the adulteration of fruit juices with cheaper alternatives</td>
<td>UPLC–QToF</td>
<td>LDA</td>
<td>(Jandrić et al., 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PCA</td>
<td>(Jandrić et al., 2014)</td>
</tr>
</tbody>
</table>

**Alcoholic beverages**

| Wine | Differentiation of organic and conventional wines | HPLC | PCA | (Tobolková et al., 2014) |
| Wine | Classification of southern Italy monovarietal wines | HPLC-DAD-MS | PCA | (Ragone et al., 2015) |
| Wine | Characterization of Polish wines | HPLC | PCA | (Socha et al., 2015) |
| Wine | Classification of Cabernet Sauvignon wines | HPLC | PCA | (Radovanovic et al., 2016) |
| White wine | Determination of polyphenols in white wines | HPLC-UV | PCA | (Larrauri et al., 2017) |
| Red wine | Classification of red wines from South America | HPLC | PCA | (Belmiro et al., 2017) |
| Wine | Authentication of vintage year | HPLC-DAD | PCA-LDA | (Liu et al., 2017) |
| Wine | Simultaneous analysis of sugars and organic acids in wine | HPLC | PCA | (Coelho et al., 2018) |

**Honey**

| Honey | Geographical Origin | HPLC-PAD | LDA | (Nozal et al., 2005) |
| Honey | Authentication and classification of floral origin | HPLC-DAD | SESCF | (Zhou et al., 2014) |
| Honey | Discrimination of Polish unifloral honeys | HPLC-DAD | PCA | (Kuš and van Ruth, 2015) |
| Honey | Identification of monofloral honeys | HPLC–ECD | DA | (Zhao et al., 2016) |
| Honey | Discrimination of botanical origins for Chinese honey | HPLC-FLD | PCA | (Chen et al., 2016) |
| Honey | Identification of botanical origin of Chinese unifloral honeys | RP-HPLC | CA | (Sun et al., 2017) |
| Honey | Discrimination of high altitude Indian honey | HPLC | CA | (Nayik et al., 2018) |

**Coffee**

<p>| Coffee | Detection of ground roasted coffee adulteration with roasted soybean and wheat | HPLC | PCA | (Pauli et al., 2014) |
| Coffee | Detection of roasted and ground coffee adulteration | HPLC | PCA | (Domingues et al., 2014a) |
| Green coffee | Differentiation of degrees of ripeness of Catuai and Tipica green coffee | HPLC | PCA | (Smrk et al., 2015) |</p>
<table>
<thead>
<tr>
<th>Product</th>
<th>Methodology</th>
<th>Analytical Technique(s)</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coffee</td>
<td>Characterization of the effects of different roasting conditions on coffee samples of different geographical origins</td>
<td>HPLC-DAD  PLS-DA SIMCA</td>
<td>(De Luca et al., 2016)</td>
</tr>
<tr>
<td>Arabica coffee</td>
<td>Detection of Arabica coffee adulteration with maize</td>
<td>HPLC-FLD PCA LDA SIMCA</td>
<td>(Tavares et al., 2016)</td>
</tr>
<tr>
<td>Coffee green beans</td>
<td>varietal classification of coffee</td>
<td>HPLC-DAD  PLS-DA SIMCA MCR-ALS</td>
<td>(de Luca et al., 2018)</td>
</tr>
<tr>
<td>Tea</td>
<td>Tea processing (variety)</td>
<td>HPLC-PCA SVM RF</td>
<td>(Zheng et al., 2009)</td>
</tr>
<tr>
<td>Green tea</td>
<td>Prediction of the total antioxidant capacity</td>
<td>HPLC-UV MLR OPLS</td>
<td>(Dumarey et al., 2010)</td>
</tr>
<tr>
<td>Green and white tea</td>
<td>Differentiation between tea categories</td>
<td>UPLC-DAD/(ESI)MS PCA</td>
<td>(Zhao et al., 2011)</td>
</tr>
<tr>
<td>Green tea</td>
<td>Evaluation of (dis)similarity</td>
<td>HPLC-UV CASE RF</td>
<td>(Alaerts et al., 2012)</td>
</tr>
<tr>
<td>Tea</td>
<td>Geographical origin of tea</td>
<td>UHPLC-MS PCA PLS-DA</td>
<td>(Fraser et al., 2013)</td>
</tr>
<tr>
<td>Green tea</td>
<td>Identification of Chinese Ziyang green tea from different plantations</td>
<td>HPLC-DAD SESCF</td>
<td>(He et al., 2015)</td>
</tr>
<tr>
<td>Meat and meat products</td>
<td>Assessment of the effect of different packaging materials on some quality indices of a Nigerian stick meat during Storage</td>
<td>HPLC-PCA</td>
<td>(Olaoye et al., 2018)</td>
</tr>
<tr>
<td>Stick meat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other food products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>Identification of Italian wheat cultivars and detection of durum wheat flour adulteration</td>
<td>HPLC-PCA</td>
<td>(Bonetti et al., 2004)</td>
</tr>
<tr>
<td>Vanillin</td>
<td>Discrimination of vanillin from different origins</td>
<td>HPLC-UV PCA</td>
<td>(van Nederkassel et al., 2006)</td>
</tr>
<tr>
<td>Raspberry</td>
<td>Identification of raspberries</td>
<td>HPLC-DAD SESCF</td>
<td>(Chen et al., 2009)</td>
</tr>
<tr>
<td>Kiwi and pomelo</td>
<td>Classification according to the fruit species and subspecies</td>
<td>HPLC-(ESI)MS PCA CA LDA</td>
<td>(Sârbu et al., 2012)</td>
</tr>
<tr>
<td>Tofu-type soybean</td>
<td>Differentiation of Tofu-type soybeans</td>
<td>HPLC-UV CASE</td>
<td>(Shen et al., 2012)</td>
</tr>
<tr>
<td>Honey bush tea</td>
<td>Differentiation between wild harvested and cultivated seedling plants</td>
<td>HPLC-DAD PCA</td>
<td>(Schulze et al., 2014)</td>
</tr>
<tr>
<td>Sugarcane</td>
<td>Discrimination of genotypes</td>
<td>HPLC-DAD PCA</td>
<td>(Leme et al., 2014)</td>
</tr>
<tr>
<td>Food Product</td>
<td>Study Title</td>
<td>Method(s)</td>
<td>Year</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>--------------------------------------------------------------</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>Pomegranate peel</td>
<td>Identification of Chinese pomegranate from different orchards</td>
<td>HPLC-UV</td>
<td>2014</td>
</tr>
<tr>
<td>Saffron components</td>
<td>Geographical characterization</td>
<td>MAE-HPLC</td>
<td></td>
</tr>
<tr>
<td>Red and the white rice</td>
<td>Discrimination of red and white rice bran from Indonesia</td>
<td>HPLC</td>
<td></td>
</tr>
</tbody>
</table>