Chapter from the book *Latest Research into Quality Control*
Downloaded from: http://www.intechopen.com/books/latest-research-into-quality-control

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
1. Introduction

Fraud can be generally defined as “the intentional deception made for personal gain or to damage another individual”. In particular, food fraud consists in the deliberate misdescription in order to deceive the consumers about the real nature of the product or of any of its ingredients. It results in the mismatch between what a food product is and what it is claimed to be. Food fraud is a broad term that also involve criminal acts such as tax-avoidance and smuggling. In the following we will mainly discuss about economically motivated adulteration and mislabelling i.e. food fraud issues falling in one of the following categories: 1) the substitution of an ingredient with a cheaper alternative (e.g., substitution of ethanol with methanol in wine or proteins with melamine in milk powders), 2) misdescription of the real nature of the product or one of its ingredients (e.g. counterfeiting, conventional products that are sold as added value products such as organic, fair trade, biodynamic), 3) incorrect quantitative ingredient declaration and 4) implementation of non-acceptable process practices such as irradiation, heating or freezing (e.g. thawed fish sold as fresh).

Food fraud can be implemented in any step of the food chain but it is mainly a food industry issue. In criminology, there are 3 elements of fraud opportunity (the crime triangle): victim, fraudster and guardian [1]. The typical set-up is that where final consumers play the role of victims, food industry (but in general food producers, processors, traders or retailers) plays the role of the fraudsters and governmental control authorities, non-governmental and certification organizations play the role of the guardian. However, food producers can also be victim or guardian. They are victims when, for example, their products are counterfeited or simulated and guardian when they implements QA systems for the assessment of the authenticity of the raw materials. Nowadays food fraud represents a major problem that costs the EU food industry and governments hundreds of millions of euros every year. But the
problem that food fraud poses is not merely an economic one: It is also a problem of public health because the adulteration can pose toxicological and hygienic risks to purchasers and consumers. In 2008, for example, Chinese milk was adulterated with melamine, a hazardous chemical, to increase milk nitrogen content causing 900 infants to be hospitalized with six deaths. Several other such examples can be given.

Since no one likes to be swindled, neither producers, traders, importers, retailers, and consumers, fraud prevention and detection is an important issue. Nowadays, authenticity of ingredients or products is mainly warranted by paper trailing. Analytical tests which can help to confirm the authenticity of ingredients/products compose a very useful complementary approach to paper trailing.

2. Fingerprinting approach: generalities and tools

Traditional strategies for the food fraud control have relied on the determination of the amount of a marker compound or compounds and the comparison of the obtained values for the test material with those established for the genuine material. The presence of an undesired adulterant can be uncovered by checking for its presence in the food material whereas the compliance of the food composition with the established legislative standards or with the amount of an ingredient as declared on the label can be simply proved by measuring the target compound or compounds. However, some aspects of food authenticity such as the geographical origin, the farming management systems (organic, free range..), or the application of some specific processes cannot be dealt with those traditional approaches. No single marker exists for the unequivocal authentication of an organic egg or a Dutch specialty cheese. Furthermore, based on conventional target analyses, an adulteration can be detected only if the adulterant is known beforehand and explicitly searched for by the analyst. Traditional quality control strategies are not designed to look for a near infinite number of potential contaminants so that new adulterants will not be unveiled until their presence in food is first acknowledged. For those reasons a more holistic approach is needed that is based on the measurement and the evaluation of several compounds at once, i.e. a fingerprinting approach. Moreover, in industrial and laboratory settings, there is always the need of implement screening methods that are able to reliably identify, in large numbers of samples, those that are potentially non-compliant before more detailed and accurate analysis with confirmatory methods are performed. A fingerprinting approach may, in many cases, provide rapid and high-throughput analyses well suited for screening purposes.

Fingerprint refers to the characteristic spectrum or image of a test material which can be related to its properties and thus to its authenticity in the same way as a human fingerprint is specific of a certain person and unequivocally identify him/her. The term thus recalls a comprehensive description of a test material that is carried out in a non-selective (or untargeted) way. Fingerprints can be generated through many analytical techniques. They can be obtained from chromatograms, spectroscopic measurements, spectral measurements or any other specific signal of complete spectra (Figure 1).
In a mass spectrum (Figure 1a), a collection of m/z and relative signal intensities is a chemical fingerprint of the material. Similarly, a chromatographic profile (Figure 1b) is a fingerprint of a more or less broad class of constituents of the material. The specific distribution of the restriction fragments of a selected DNA sequence on the electrophoretic gel (Figure 1c) is a genetic fingerprint of the test material. A NIR spectrum (Figure 1d) is a representation of the interaction of a test material with the infrared radiation whereas a thermogram (Figure 1e) is a representation of its interaction with thermal energy. (a) and (b) can be referred to as chemical fingerprints. They may be composed of as many groups of compounds as possible or alternatively of a specific group of compounds which requires higher level of purification and a selective extraction from the sample. (d) and (e) can be referred to as physical fingerprints even though chemical information can be obtained as well.

A fingerprinting approach implies that the whole information contained in the fingerprint (or a selected part of it) is used to infer about the properties of the system under study. To do that, a special statistical tool is necessary, i.e. chemometrics. Chemometrics can be defined as the science of extracting chemically relevant information from multivariate data by using statistical techniques to reduce the dimensionality of the dataset. It offers a tool to graphically summarise the analytical data to reveal relationships between samples and to detect characteristic patterns that can be used to identify a certain material. As a first step, an exploratory analysis is carried out in order to investigate the natural relations between the samples. This is carried out by so called unsupervised pattern recognition techniques because they do not require any prior knowledge of the properties of the samples. Examples of such techniques...
are: Hierarchical cluster analysis (HCA), cluster analysis (CA) and principal component analysis (PCA). PCA is the most widespread of those explorative tools. In a PCA model the original variables are transformed in new uncorrelated variables that arise from the linear combination of the original variables: the principal components (PCs). A number of PCs are extracted in sequence with each principal component accounting for the maximum of the residual variance in the data. The PCs extraction stops when most of the variance in the original data (typically around 90%) is explained. The new set of PCs define therefore a new space where the contribution of each original variables to each PC can be easily represented and the relationships between the original samples highlighted (Figure 2).

**Figure 2.** Plot of the first two dimensions in a typical PCA model. In a scores plot (panel a), samples are plotted in the space defined by the PCs. Similarities and differences between samples can be highlighted (in this case, two well separated groups of samples, red and green are apparent). In a loading plot (panel b) each variable (analytical response) is plotted on the new space defined by PCs. When score plot and loading plots are superimposed, information on the effect of the variable on samples properties can be obtained. When a variable is plotted close to a sample, this indicates that this variable shows relatively high concentration in this particular sample compared to the other samples. In the figure, variables 1, 4, 12 are higher in the samples of the red group and variables 6, 7 and 11 are higher in the green group.

As a further step, multivariate methods are applied to either classify a certain product or quantify a certain property of the product. In the **classification models**, information about the class membership of the samples to a certain group (class or category) is used to classify new unknown samples in one of the known classes on the basis of its pattern of measurements. Classification models are useful, for instance, for the authentication of organic products or of geographical origin. Supervised pattern recognition techniques are used such: k nearest neighbours (kNN), soft independent modelling of class analogy (SIMCA), partial least square regression discriminant analysis (PLS-DA), linear discriminant analysis (LDA), support vector machine (SVM) and artificial neural network (ANN). Classification models may either build a delimiter between the classes so that they always assign a new object to the class to which it most probably belongs (suitable for limited and defined number of possible classes) or build a model for each class studied and then evaluate the fitting of new objects to each model (suitable for unlimited number of classes). In the **regression models**, a functional relationship is established between some quantitative sample property, the dependent variable, and a multivariate independent variables such as a raw chromatogram or a IR spectrum and the model is used to predict the property of interest in the unknown sample. Regression model are used, for instance, to quantify the level of adulteration in a food sample. For the build-
ing of regression models, multivariate regression methods such as principal component regression (PCR) and partial-least square regression (PLSR) are used.

The building of a classification (or regression) model comprises four essential steps: 1) Selection of a training set, which consist of objects of known class membership (or known quantitative values of a sample property) for which variables are measured. (2) Variable selection. Variables that contain information for the aimed classification are retained, whereas those encoding noise and/or with no discriminating power are eliminated. (3) Building of a model using the training set. A mathematical model is derived between the selected variables measured on the training set and their known categories (or quantitative values of the sample property). (4) Validation of the model. The model is validated in order to evaluate the reliability of the classification achieved either using an independent test set of samples (external validation) or the training set (cross validation).

The fingerprinting techniques are gaining more and more popularity over the past years thanks to advancements in the analytical instruments that are able to generate enormous amount of data at once and the application of chemometrics techniques. Herewith, fingerprinting techniques are classified in five broad categories according to the kind of fingerprint that can be obtained: Mass spectrometry (MS) fingerprinting, chromatographic fingerprinting, electrophoretic fingerprinting, spectroscopic fingerprinting, and other fingerprinting. This classification is shown in Table 1.

<table>
<thead>
<tr>
<th>MS fingerprinting</th>
<th>Chromatographic fingerprinting</th>
<th>Electrophoretic fingerprinting</th>
<th>Spectroscopic fingerprinting</th>
<th>Other fingerprinting</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTR-MS</td>
<td>LC (HPLC, LC-MS..)</td>
<td>CE (CZE, CIF…)</td>
<td>NMR</td>
<td>DSC</td>
</tr>
<tr>
<td>ICP-MS</td>
<td>GC (GC-FID, GC-MS..)</td>
<td>Gel electrophoresis (isoelectric focusing, DNA electrophoresis…)</td>
<td>Fluorescence spectroscopy</td>
<td>Electronic nose</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR-MS</td>
<td>IR (NIR, MIR, FTIR)</td>
<td>Microarray technologies, reverse PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct infusion and ambient MS (ESI, MALDI-TOF, DART…)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 1. Classification of fingerprinting techniques. For abbreviations, see text.*

3. MS fingerprinting

MS is a powerful analytical technique that measures the mass-to-charge ratio of ions. The samples are first ionised, the ions are separated and their relative abundance assessed based
on the intensity of the ions flux. MS produces therefore a mass spectrum representing the fingerprint of the sample components (Figure 1a). A number of different MS set-ups are possible based on the ionisation technique and the mass analyser used. MS can be used alone or they can be coupled with separation techniques. In this section, the stand-alone MS techniques will be dealt with in details whereas the application of MS as coupled with separation techniques will be dealt with in the next sections.

Stand-alone MS fingerprinting techniques that proved to be very useful for the fraud control and prevention are: proton transfer reaction MS (PTR-MS), inductively coupled MS (ICP-MS), isotope ratio mass spectrometry (IRMS), and direct infusion MS techniques.

Proton Transfer Reaction Mass Spectrometry (PTR-MS)

PTR-MS is a relatively new technique that is rapidly gaining popularity in the food analysis. PTR-MS allows quantitative on-line monitoring of volatile organic compounds. The volatile compounds are softly ionized by means of hydroxonium ions that are generated in an external ion source operating in pure water vapour. Only the volatile compounds that have a higher affinity for the ions are protonated and then accelerated by an electric field to the reaction chamber where they are separated and quantified. Because of this soft chemical ionisation the fragmentation of the parent compounds is limited and the interpretation of the spectra are much easier. Other major advantages of this technique are the great sensitivity with detection limits as low as few part per trillion, volume (pptv) and the possibility to monitor the food samples in real time, without any work up procedure. As a result, a fingerprint of all the volatile compounds comprised in a well definite mass range is obtained. The main disadvantage of this technique is that compounds are characterized only via their masses which is insufficient for their unequivocal identification.

PTR-MS has been extensively used in several aspects of food fraud control. It proved, for example, very successful for the geographical authentication of foods. The EU has long recognized the importance of differentiating food products on a regional basis. The normative framework introduced by the EU comprises the EU Regulations 509/2006 and 510/2006 and the EU Regulation 1898/2006. The EU Regulation introduced three geographical indications to a food product: protected designation of origin (PDO), protected geographical indication (PGI) and traditional specialities (TSG). In a study of 2008, the geographical origin (country) of butter samples was successfully predicted in 88% of the cases based on PTR-MS fingerprint and PLS-DA [2]. Recently, volatile fingerprint generated by PTR-MS has been used to discriminate between the Boeren Leidse specialty cumin cheeses with EU PDO from other 29 cumin cheese manufactured in the Netherlands [3]. The volatile fingerprint coupled with a PLS-DA model allowed the correct classification of 96% of the traditional boeren leidse cheese samples and 100% of the other commercial cheese samples. Another typical added value that is protected by the EU regulations is represented by the monovarietal extra virgin olive oil (EVOO), i.e. oil that is produced out of just one variety of olive trees. Frauds can be committed by mixing the more valuable monovarietal virgin olive oil with cheaper oils or by mixing different monovarietal olive oils. Volatile fingerprint of virgin olive oil obtained by PTR-MS and subjected to PLS-DA proved successful in discriminating among 5 different monovarietal EVOO from Spain with 100% sensitivity (% of objects of the modelled class
correctly accepted by the model) and specificity (% of object, extraneous from the modelled class, correctly refused by the model) close to 100% [4].

In conclusion, PTR-MS is a rapid and low cost analytical technique that can be also fully automated and implemented on-line. Recently, the coupling of the time of flight (TOF, see below in this section) mass analyser to PTR-MS instruments has generated PTR-TOF-MS which is characterized by a high sensitivity with limits of detection down to few pptv and a high time resolution. The technique has been recently applied to discriminate among PDO labelled hams from Spain and Italy [5].

**Inductively coupled plasma mass spectrometry (ICP-MS)**

ICP-MS provides quantitative measurements of a wide range of metals and non-metals (inorganic elements) at trace and ultratrace concentration level (ppt). In this technique, the sample (even solid or liquid) is decomposed to neutral elements in a high-temperature argon plasma and the single elements are separated based on their mass/charge ratio and analysed. The great advantage of this technique compared to others (e.g. atomic spectroscopy) is that more than one element can be analysed at once so that a multi-elemental fingerprint is obtained in a very fast and sensitive way. The multi-elemental composition of animal and vegetal tissues can provide valuable information on the characteristics of the soil where a crop has been cultivated and on plants composition of the animal diet. The multi-element fingerprint is thus a valuable marker of the geographical origin of food. For instance, the authenticity of Tropea red onion, an onion Italian variety that achieved the PGI certification by the European Union as “Cipolla Rossa di Tropea Calabria” can be proved by means of multi-elemental analysis by ICP-MS and multivariate statistics [6]. All the statistical models applied (LDA, stepwise LDA, SIMCA, ANN), allowed a success rate of prediction >90% for the genuine samples. Moreover, the availability of nutrients from the soil strictly depends on the fertilization strategies and the pest and weed control management systems. In organic farming synthetic fertilizers are not permitted and the pest and weed control is based exclusively on natural products. It has been thus proposed that the multi-elemental fingerprint might be a marker for organically cultivated crops as compared to conventionally cultivated ones. Laursen et al. managed to discriminate between organic and conventional wheat, barley and faba beans (but not potatoes) based on the profile of 25 elements measured by ICP-MS [7].

**Isotope ratio mass spectrometry (IRMS)**

IRMS is a technique that can measure the ratio of the stable isotopes of the constituents of a biological material. Light elements like carbon, nitrogen, hydrogen, oxygen and sulphur stable isotopes ratios are most frequently assessed with this technique. Those ratios vary according to specific food production factors and geo-climatic conditions. Carbon stable isotope ratio depends, for example, from the plant composition of ruminant diets and can then be used to authenticate feeding regime or the farming management system (organic, free-range). Nitrogen stable isotope ratio is on the other hand depending on the type of fertilizers used in agriculture and is thus much useful for the authentication of farming practices for vegetal products and crops. Oxygen isotope ratio is instead highly dependent on the
distance from the ocean and the altitude above sea level and could then be used for the authentication of the geographical origin of a food product.

The stable isotope fingerprint has been successfully used for the authentication of geographical origin and the farming practice. Normally the data are measured for many different elements and analysed with multivariate statistics. As an example, Fontina PDO cheese can be discriminated with good success from other cheeses based on stable isotopes $^{2}\text{H}/^{1}\text{H}$, $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$ and $^{34}\text{S}/^{32}\text{S}$ and PCA analysis [8]. However, frequently the isotope ratios (or a selection of them) are combined with other markers (elements) to improve the accuracy of the classification models. IRMS is often combined with ICP-MS for simultaneous elemental analysis. Stable isotope analysis combined with multi-elemental analysis has proven ideally suited to determine geographical origin of foods. The most accurate measurements of the isotope ratios is obtained by dual inlet (DI) IRMS. However, the purchasing and operating costs of a DI-IRMS instrument coupled with the time-consuming sample preparation are major disadvantages for the diffusion of this technique. The introduction of continuous flow (CF) IRMS instruments offers on-line, rapid and automated sample preparation, greater cost-effectiveness and easier interfacing with other preparation techniques.

**Direct infusion mass spectrometry**

Direct infusion MS techniques are based on the direct injection of the sample in the ion source without or with small sample pre-treatment. This allows for rapid analysis suited for high-throughput screenings. Electrospray ionisation (ESI), matrix assisted laser desorption ionization (MALDI) and direct analysis in real time (DART) are typical ionization techniques used for direct infusion MS. They are coupled with a variety of mass analysers, e.g. time of flight (TOF), Fourier Transform Ion Cyclotron Resonance (FT-ICR), single quadrupole (Q) and ion trap (IT) in many different set-ups.

ESI is a typical soft ionisation technique that is particularly suited for the determination of the molecular mass of large molecules (proteins, peptides, polysaccharides, triglycerides), because the ionisation does not bring about the fragmentation of the molecule. The liquid in which the analyte is contained is dispersed by electrospray to a fine aerosol. The droplets shrink as the solvent evaporates till solvated ions desorb from their surface. ESI-MS has proven to be very helpful in the authentication of vegetable oils. Lipid composition of vegetable oils depends on their botanical origin and the way they are processed. Fatty acids (FAs) and/or triglycerides (TGs) profile can thus help authenticate the type of oil, its origin, its quality grade and potential adulteration. Direct infusion ESI-MS has been for example used to predict the olive oil quality according to European Union marketing standards based on fatty acids and LDA analysis [9]. In the same research, the percentage of either EVOO and VOO in binary mixture with other lower grade oils was predicted with 5–11% average prediction errors by using PLS and multilinear regression (MLR). Samples were 1:50 diluted in an alkaline 85:15 (v/v) propanol/methanol mixture and directly infused into the MS system. Triglycerides analysis has some advantage over the analysis of the fatty acids profile for authentication or fraud control. Indeed, different oils can have specific TG fingerprint despite showing the same fatty acids composition. The triglyceride profiles, obtained using Q-TOF-ESI-MS was used to predict adulteration of olive oils with other vegetable oils.
The adulteration with hazelnut oil was predicted at a level of 10% v/v [10]. This adulteration is difficult to detect at levels below 20% by conventional methods due to the compositional similarity between the two oils. Similarly, PCA and HCA methodologies, applied to the ESI(+)‐MS data were able to readily detect adulteration of EVOO with ordinary olive oils, at levels as low as 1% w/w [11]. Mono-, di- and triglycerides together with vitamins and antioxidants were detected and quantified with this method. Direct infusion ESI-MS has been used to authenticate other food commodities. The chemical fingerprint generated by direct infusion ESI-Q-TOF-MS in the negative mode can be used to discriminate between genuine whisky from Scotland and US, from counterfeited whisky produced in Brazil [12] and between alembic (the most valuable) and industrial cachaças (Brazilian sugarcane spirit) as well as the fraudulent addition of sucrose to the spirit [13]. Finally, direct-infusion ESI-QqQ-TOF-MS and atmospheric pressure photoionization (APPI)-QqQ-TOF-MS have been used for Iberian ham typification. APPI is a soft ionization technique based on a photoionisation mechanism. Five types of Iberian hams were successfully classified. Applying a PLS-DA model [14].

MALDI is another soft ionisation technique that proved very useful in the analysis of macromolecules, especially proteins. In MALDI the molecules are desorbed from the support matrix and ionised by means of a UV laser beam in a complex process mediated by the support matrix itself. MALDI is mainly coupled with a time-of-flight (TOF) mass analyser which separate the ions based on their flying time to the detector, which on turns depends on their m/z ratio. An example of application of MALDI-TOF-MS for authentication issues is represented by the fast method developed by Wang et al. for the fingerprinting of honey proteins [15]. The mass spectra were used to build up a database library to be used for authentication purpose. The protein fingerprint was thus successfully used to authenticate the geographical origin of commercial honeys produced in the US and other countries. In a similar fashion, peptide fingerprinting obtained by MALDI-TOF has been converted in a biological bar code for the authentication of high quality Campania white wines [16].

An innovative technique for food fingerprinting is represented by the direct analysis in real time (DART)-MS. DART is an ambient ionisation technique i.e. in which ions are formed outside the mass spectrometer without sample preparation or separation. The samples, either gaseous, liquid or solid are ionised in open air under ambient conditions. This means that organic compounds can be directly, and in real time, determined without time-consuming analytical protocols and thus with high sample throughput. DART coupled with TOF-MS has been used to obtain the fingerprint of the triglycerides from olive oil [17]. This method, coupled with LDA allowed the discrimination between EVOO, olive oil and olive oil pomace and the detection of hazelnut oil in EVOO at 6% v/v. DART-TOF-MS with solid phase micro extraction (SPME) pre-concentration of the analytes has been also reported to allow discrimination between trappist and non-trappist beers based on volatiles and phenolic compounds [18]. A combination of DART-TOF-MS and chemometrics was used for animal fat (lard and beef tallow) authentication [19]. TGs and polar compounds were extracted and analyzed. Mass spectral records were processed by PCA and stepwise LDA. The LDA model developed using TAG data enabled the classification of lard and beef tallow samples but also detection of admixed lard and tallow at adulteration levels of 5 and 10% w/w.
Additional ambient ionisation techniques have been recently proposed for authentication and fraud control by product fingerprinting. For instance, easy sonic spray ionisation (EASI)-MS fingerprinting of fatty acids and phenolic compounds have been used for the authentication of olive oil geographical origin [20].

4. Chromatographic fingerprinting

Chromatographic techniques aim at resolving complex mixtures in well separated compounds. Based on the detection system, each single compounds generates a signal that can be used for the qualitative and quantitative analysis of the mixture. The graphical representation of such signal as a function of time is referred to as a chromatogram and can be thought of as the fingerprint of one or more classes of compounds occurring in the sample. Different strategies are available to obtain multivariate data matrices from chromatographic analyses (Figure 3). The fingerprint can be composed by the set of concentrations of the separated compounds based on an identification/calibration/quantification procedure as depicted in the path (a) of Figure 3. Alternatively, the fingerprint can be represented by the set of peak areas/heights (b). In this case the identification of each single peak is not necessary. Finally, it can be represented by the whole chromatogram that is handled as a continuous signal (c). In this case, the multivariate dataset is composed by as many variables as the sampling points the chromatogram is made up of (each data point of the chromatogram represents an individual variable). However, the application of chemometrics on raw chromatographic data requires specific data pre-processing techniques. In fact, problems related to the peak alignment or baseline shifts are particularly critical when a raw chromatogram is used as a data set.

![Figure 3. Schematic representation of the strategies to obtain multivariate dataset from a chromatogram for chemometric analysis.](image)
**Liquid chromatography**

In liquid chromatography, the mixture components are separated as they pass through a column based on their selective partition behaviour between a stationary phase (column material) and a mobile liquid phase. Depending on the type of stationary phase, compounds can be separated based on their size, charge, molecular mass, hydrophobicity etc. The most popular LC technique is high performance liquid chromatography (HPLC) that is a straightforward, robust and reproducible technique. HPLC has been used for the analysis of a wide range of food compounds such as vitamins, proteins, carbohydrates, TGs, additives, secondary plant metabolites. A typical example of the application of HPLC based fingerprint for fraud control is the authentication of the organic eggs by means of the carotenoids profile [21]. Carotenoids are a group of fat-soluble pigments that occur in the egg yolk in concentrations of about 10 mg kg\(^{-1}\). In animals carotenoids are entirely of dietary origin. Since the feeding regime of organic hens is clearly different from that of conventionally reared hens, the carotenoids fingerprint in eggs were used to discriminate between production systems. The carotenoids profile was determined by HPLC with UV detection, the single carotenoids quantified and the resulting concentrations used to build up a classification model by kNN. Almost all the conventional eggs and all the organic ones were correctly classified in external validation. The robustness of the method has been recently improved by testing eggs produced in several EU and non-EU countries. An example of HPLC fingerprint based on raw chromatographic data is in [22]. The authors applied PLS to the full TGs chromatogram profiles of vegetable oils to predict the % of olive oil in the mixtures with errors not exceeding 10%. Liquid chromatography can also be coupled with MS which allows higher resolution and higher sensitivity for metabolites occurring in relatively low amount. As an example, an untargeted method for proteins analysis based on LC-QTOF-MS has been developed which allowed to detect the fraudulent addition of cheaper vegetal proteins (from soy and pea) to skimmed milk powders based on the different peptides profile [23].

**Gas Chromatography**

In gas chromatography, the mixture is first vaporised in a heated chamber and then the mixture components are separated as they travel through the column transported by the flow of an inert gas (helium, nitrogen or hydrogen) based on their selective interaction with the column material. GC is a very popular separation technique mainly used for the analysis of volatile compounds. However a wide spectrum of compounds can be rendered volatile by proper derivatisation and thus analysed by GC. The analysis of fatty acids and triglycerides is usually carried out by GC with flame ionisation detector (FID) previous derivatisation in fatty acids methyl esters (FAME) and TG trimethylethers, respectively. FID is a general detector capable of high sensitivity and robustness. Fatty acids composition of animal tissues and animal products strongly depends on the feeding regime. FAs composition of fish muscle fat is affected by animal diet/feeding, the geographical area of catch and the marine conditions and is thus different between farmed and wild fishes. The discrimination between wild and farmed Atlantic salmon (Salmo salar L.) and Wild Turbot (Psetta maxima) has been reported based on FAs analysis and chemometrics [24-25]. The fatty acid fingerprint obtained by GC-FID followed by PLS-DA analysis has been also reported for the authentica-
tion of organic eggs and of organic feeds [26-27]. In the last case 90% of the analysed samples were correctly classified in their proper group in external validation. GC-FID can also be used for the TG profiling. TG fingerprinting by GC-FID has been for example reported for the authentication of three fat classes (animal fats, fish oils, recycled cooking oils) [28]. The TGs fingerprint was subjected to multivariate analysis (PLS-DA) and allowed the correct classification of 96% of the fat samples.

GC coupled with MS represents the method of choice for the analysis of volatile compounds because of its high reproducibility. On the other hand, GC-MS analysis requires careful sample cleaning and is quite expensive and time-consuming. The volatile fingerprint of coffee obtained by GC-TOF-MS after SPME has been reported for the geographical authentication of coffee [29]. SPME preconcentration of volatiles followed by GC-MS analysis coupled with PCA analysis allowed the detection of adulteration of ground roasted coffee with roasted barley [30]. The adulteration is detectable at level of 1% w/w in mixtures of dark roasted barley and coffees. Metabolomics studies can be also fruitfully performed by GC-MS. The fingerprint of a large range of metabolites obtained by GC-MS has been used to discriminate between mechanical separated meat (MSM) from hand-deboned meat [31]. MSM could be detected in raw meat mixtures down to a level of 10%.

5. Electrophoretic fingerprinting

Electrophoretic techniques are able to separate a complex mixture under a spatially uniform electric field, based on electrophoretic mobility of its components that depends, in turn, from their hydrodynamic properties and charge. Positively charged molecules move towards the anode and negatively charged molecules towards the cathode at a different rate based mainly on their mass to charge ratio. Smaller molecules move faster than larger ones.

Gel electrophoresis

In a gel electrophoresis, a gel is used as a medium for the movement of the charged particles under the applied electric field. Agar and polyacrylamide are typical medium used in gel electrophoresis. Proteins and nucleic acid fragments are usually separated by gel electrophoresis. Gel electrophoresis is of major importance for the genomic fingerprint of a sample material. Genomic fingerprints are obtained when properly amplified targeted or untargeted DNA or RNA fragments are separated by electrophoresis thus providing patterns that can be associated to sample properties (specie, variety and the like). Unlike the fingerprints discussed in the previous (and the next) sections, DNA fingerprint shows somehow different characteristics. The single features of the fingerprint are not quantitative variables (physical or chemical variables allowed to take on quantitative values, e.g. area of a peak in a chromatogram, signal intensity for a m/z or absorbance at a fixed wavelength in a IR spectra) but rather categorical variables, i.e. electrophoretic bands that can be either present or absent (see Figure 1 (c)). The sample identification is thus mainly carried out by checking for the presence (or absence) of one or more target bands. Multivariate analysis of the DNA fragments patterns is rarely performed.
Genomic fingerprinting mainly relies on polymerase chain reaction (PCR) based techniques. PCR is based on the amplification of a target DNA sequence by means of a thermostable DNA polymerase. The process consists of several cycles where the DNA molecule is denatured, specific primers (small DNA sequences) anneal to the target DNA sequence and the DNA polymerase synthesizes a new DNA fragment delimited by the two primers. In each cycle the number of DNA molecules increases exponentially. A PCR-derived fingerprint can be obtained in different ways. In PCR-RFLP (restriction fragment length polymorphism) the amplified region is digested with an endonuclease and the resulting DNA fragments are separated by electrophoresis and properly visualised. The pattern of fragment represents a fingerprint of the DNA sequence that has been amplified. In multiplex PCR, two or more DNA fragments are simultaneously amplified by means of different target primer pairs, separated by electrophoresis and visualised. In RAPD (random amplification of polymorphic DNA) random DNA fragments are amplified by means of arbitrarily created primers. After separation, the DNA fragments will give rise specific patterns on the gel. Finally, in single-strand conformation polymorphism (SSCP), DNA sequences are amplified, denatured and the resulting single strand DNA molecules separated by electrophoresis based on their specific secondary structures.

PCR-based fingerprinting techniques have been widely used for species identification. For instance, the identification of ten species of salmon genus in a wide range of commercial products can be accomplished by PCR-RFLP based on the amplification of a specific region of the mitochondrial cytochrome b gene followed by polyacrylamide gel electrophoresis (PAGE) [32]. Similarly, PCR-RFLP has been used to identify 15 species of gadoid fishes based on the amplification of a small region of the cytochrome b gene and three restriction enzymes [33]. Gadidae family is one of the most commercially important fish family comprising species as Atlantic cod (Gadus morhua), the pollack (Pollachius pollachius) and the haddock (Melanogrammus aeglenus). Duplex PCR targeting the cytochrome b gene can be used to detect cow milk in buffalo mozzarella at a level of 1%. Buffalo mozzarella is labelled with PDO and can be produced only with pure water buffalo milk (Bubalus bubalis) [34]. In an original approach, the multiplex PCR fingerprint of the 16S and 23S rDNA genes of the lactic bacteria naturally occurring in milk has been used to discriminate the geographical origin of PDO mozzarella cheese [35]. The PCR fingerprint was subjected to cluster analysis (neighbour-joining algorithm) which allowed a fair discrimination of the samples.

Genomic fingerprinting shows a unique potential for the species or variety authentication in food products. The introduction of PCR has notably increased the potential of this approach. However, compared to other fingerprinting techniques, genomic fingerprinting is relatively time-consuming and labour-intensive. Its applicability to fraud issues other than genetic identification is limited. Furthermore, food processing may degrade the DNA molecule and lower its recovery thus negatively affecting the results of an analysis when applied to heavily processed foods.

**Capillary electrophoresis**

Capillary electrophoresis (CE) is the electrophoretic technique that shows a notable potential for food fraud detection based on product fingerprint. CE is a family of separation techni-
Techniques that separate charged analytes based on their electrophoretic mobility: capillary zone electrophoresis (CZE), capillary isoelectric focusing (CIF), capillary gel electrophoresis, capillary electrochromatography. An electric field is applied to the ends of a capillary column. The ions migrate through the column in the same direction pulled by the electrosomotic flow and are separated based on their electrophoretic mobility. The signal that is generated when the mixture components are detected as they are eluted from the column is referred to as a capillary electropherogram. Multivariate dataset can be obtained from electropherogram in the same way as depicted in Figure 3. However, only strategy (a) has been used so far for authentication and fraud control purposes. CE are capable of rapid, low cost and high resolution analysis with little consumption of mobile phase. Main disadvantages of the technique are the low reproducibility (compared to other separation techniques) and low sensitivity that makes CE not suitable for the analysis of compounds occurring in trace amounts. CE represents a good alternative for the multiple detection of inorganic and organic acids. Many fruits and vegetables is rich in organic acids occurring in varying quantities in different fruits types, giving each fruit a unique organic acid profile. These profiles can be thus used to authenticate a vegetable product or identify the addition of another fruit type. For example the organic acids content measured by CE and LDA has been used to classify Spanish white wines [36].

6. Spectroscopic fingerprinting

Spectroscopy is the study of the interaction between a material and radiated energy. The graphical representation of such interaction is what is referred to as a spectrum i.e. a plot of the response of interest as a function of the wavelength or the frequency of the radiation used (see Figure 1d). Such a spectrum is by its very nature a fingerprint of the target material and contains information that are multivariate in nature. The extraction of the chemically relevant information from such a fingerprint requires the application of multivariate statistical techniques. The whole spectrum is used (or part of it) to obtain a multivariate dataset for further chemometric analysis in the same way as described for raw chromatograms (see Figure 3c). Spectral fingerprinting can be used either to classify and discriminate between samples or to quantify a certain compounds. According to the nature of the radiating energy (infrared, visible, ultraviolet, x-rays) and the nature of the interaction between energy and matter (absorbance, emission, scattering, resonance) different kind of spectra can be obtained. In the following we will mainly focus on nuclear magnetic resonance (NMR), fluorescence spectra and infrared (IR) spectra.

NMR

NMR spectra are generated by the absorption of radiofrequency radiation by atomic nuclei with non-zero spin in a strong magnetic field. Such absorption is affected by the surroundings of the atomic nucleus so that precise information about the molecular structure of a sample can be obtained. The atomic nuclei with non-zero spin that are most frequently used in NMR are \(^1\)H, \(^{13}\)C even though \(^{15}\)N, \(^{17}\)O, \(^{19}\)F and \(^{31}\)P can be also employed.
Generally, NMR is superior to other spectroscopic technique because of the much richer and more detailed information that can be gathered from the NMR spectra, at least with high resolution instruments that use frequencies above 100 MHz. Those information can be used for the simultaneous quantitative determination of a number of compounds without any prior separation. Furthermore the NMR spectrum can be considered a molecular fingerprint of the test material and subjected to multivariate analysis. The main disadvantage of this technique is the elevated costs of the instruments and the running costs. Nowadays, low resolution NMR instruments are available that use frequency ranging from 10 to 40 MHz. Those instruments are much cheaper and easy to use but do not provide the same detailed information as the high resolution instruments. NMR instruments are also capable of good accuracy but the sensitivity is lower compared to MS.

A recent study on the quality control of cola beverages using NMR is exemplar of the potential application of this technique for food authentication and fraud control [37]. $^1$H NMR spectroscopy was used to discriminate with high success between premium and discount cola brands. This is important in the light of possible counterfeiting. The whole NMR spectra were used in combination with PCA. In addition, the information contained in specific regions of the NMR spectra combined with multivariate calibration (PLS) allowed the quantification of several cola ingredients (caffeine, aspartame, acesulfame-K, and benzoate) which concentration must comply with regulatory limits. NMR has been also used for the authentication of the geographical origin of olive oils. $^1$H NMR spectra of the bulk olive oil, its corresponding unsaponifiable fraction, and a subfractions of the unsaponifiable fraction (alcohol, sterol, hydrocarbon, and tocopherol fractions) were used to classify olive oils according to their origin [38]. The unsaponifiable fraction had to be extracted to avoid the signal to be masked by that from the TGs in the bulk oil. The adulteration of virgin olive oil with a wide range of seed oils can be detected at level as low as 5% by means of combined $^1$H and $^{31}$P NMR spectra and discriminant analysis provided that the virgin olive oil are fresh (as reflected by their high 1,2-diglycerides to total diglycerides ratio) [39]. In this case the multivariate analysis was performed on 13 compositional parameters derived from the spectra rather than on the whole NMR spectral fingerprint. $^{13}$C NMR spectra have been used for the authentication of fish and fish products. Discrimination between farmed and wild salmon is possible based on the NMR spectra of the muscle lipids and neural networks (PNN) and support vector machines (SVM) multivariate analysis [40]. Using the peak intensities of 12 selected chemical shifts an excellent discrimination is obtained by using PNN and SVM (98.5 and 100.0%, respectively). The authentication of different gadoid species was also achieved based on the NMR spectra of muscle lipids and Bayesian belief networks (BBN) with successful classification of 100% [41]. However, $^1$H NMR spectroscopy can also provide useful information for the authentication of wild fish. In Figure 4, the PLS-DA scores plot for the $^1$H NMR data measured in the authors’ group on frozen, smoked and canned salmons both wild and farmed is presented. The score plot shows a clear separation of the two groups in distinct regions of the three dimensional plot. The results of the classification model (leave 5 out internal validation) were extremely positive with 100% of the wild samples (29 samples) and almost 100% of the farmed samples (60 out of 62 samples) correctly classified.
Fluorescence spectroscopy

Fluorescence spectra are normally obtained by exciting the test material. With radiations at a fixed wavelength and recording the intensity of the emitted radiation over a range of wavelengths. However, a 3D spectrum can be obtained by recording the emission spectra at different excitation wavelengths: the so called fluorescence excitation-emission matrix (EEM) which provides more information about the fluorescent compounds occurring in the sample. In synchronous fluorescence scan (SFS) the whole fluorescence landscape can also be achieved by scanning the excitation and the emission wavelengths simultaneously keeping a fixed wavelength interval (the so-called offset). Products that contains natural (or added) fluorophores are suitable for Fluorescence spectroscopy. Tryptophan, tyrosine and phenylalanine residues are fluorophores typically present in a variety of foods along with vitamin A, riboflavin (vit B2), NAD, NADH and compounds originating from Maillard reaction/ lipid oxidation. Fluorescence spectra provide information about the amount of those compounds and on the way the fluorophore environment interacts with them. Traditionally, the fluorescence spectra have been treated by means of univariate approaches i.e. taking advantage of one specific wavelength or the derived fluorescence peak features. Nowadays the multivariate information contained in the fluorescence spectra is processed by chemometric techniques. When EEMs or SFS spectra are produced, decomposition methods such as two-way PCA, TUCKER and parallel factorial analysis (PARAFAC) are necessary to extract information from such a multi-way dataset (the data can be arranged in a cube instead of a
matrix as in standard multivariate data sets). In the right-angle fluorescence spectroscopy
the incidence angle between the excitation and the emission radiation is 90°. Only liquids
and diluted solutions can be analysed and an attenuation of the signal intensity at high ab‐
sorbance (> 0.1) is observed. To overcome this problem the front-face fluorescence (FFF) has
been developed where only the surface of the material is analysed and the incidence angle is
around 56° to minimise the artefacts from the excitation radiation reflected or scattered by
the sample. Solid and powdered samples as well as bulk liquids can be analysed by FFF.

Fluorescence spectra are a promising tool to verify the egg freshness. Albumen samples stor‐
ed for 1,2,3 and 4 weeks can be discriminated by means of the Maillard reaction products
fluorescence and factorial discriminant analysis (FDA) with high success rate [42]. Vitamin
A fluorescence together with FDA allows discrimination among egg yolk samples stored for
different times [43]. Similarly, the freshness of fish can be predicted based on NADH and
tryptophan fluorescence spectra. NADH fluorescence spectra can be also considered as a
promising tool for the discrimination between frozen-thawed fish and fresh fish. The NADH
emission spectra show a typical maximum at 455 nm in fresh fish and at 379 nm in frozen-
thawed fish. The multivariate analysis (FDA) of the NADH spectrum allowed the correct
classification of 100% of the analysed samples [44]. The authenticity of edible oils has also
been extensively investigated by fluorescence spectroscopy. Chlorophyll and vitamin E are
important fluorophores in olive oils and contribute greatly to oil colour and stability during
storage. Refining processes decreases the content of chlorophyll and vitamin E with a corre‐
sponding change in the fluorescence spectrum. However, the discrimination between virgin
and refined olive oil is mainly based on the fluorescence of lipid oxidation products (more
abundant in the less resistant refined oils). A fast screening method has been also developed
to detect adulteration of EVOO with olive-pomace oil [45]. It is based on the EEMs and it is
able to detect adulteration at a level of 5%. Similarly, the discrimination between olive oils
according to their overall acidity are also possible with fluorescence spectroscopy [46]. In
this case, the maximum differentiation between the oils was obtained in the region 429-545
nm of the spectrum and allowed 100% correct classification of lampante olive oil (acidity
>3.3%, not edible) from virgin olive oil with lower acidity (<3.3%). Finally, SFS with multiple
regression analysis has been reported to for the detection of adulteration of EVOO with
olive oil to a level as low as 8.4% when a 80 nm wavelength interval is used [47].

The great advantage of fluorescence spectroscopy is the rapidity, the limited costs and the
non-destructive nature of the analysis. The sensitivity is also much greater than that of other
techniques because the fluorescence signal has in principle no background.

**Infrared spectroscopy**

Infrared spectra are produced by measuring the intensity of the absorbance of infrared light
by a sample as function of the wavelength. The absorption of infrared light is ascribed to
transitions in the vibrational energies of the molecules contained in a sample. Each func‐
tional group of a molecule shows characteristic IR absorption at specific frequency ranges re‐
gardless of the interaction of the functional group with the rest of the molecule. However,
interaction between atoms within a molecule may sometimes affect the position of charac‐
teristic bands in a IR spectra depending on the surroundings of the functional group. IR
spectra can thus provide qualitative information about the nature of the functional group present in a food sample and quantitative information on their amount. When the effect of all the functional groups is taken together, the whole spectrum represents a molecular fingerprint that can be used to verify the nature of the sample. The IR region of the electromagnetic spectrum can be divided in 3 portion: The far IR (FIR, 400-10 cm\(^{-1}\)) has the lower energy and induces rotational transitions in the molecules. The mid IR (MIR, 4000-400 cm\(^{-1}\)) induces fundamental vibrational transitions in the molecules. The near IR region (NIR, 14000-4000 cm\(^{-1}\)) also induces transitions in the vibrational energies of the molecules. However, the transitions of the vibrational energy induced by the NIR portion of the spectrum are more complex than those induced by the MIR region. Overtones (transitions from the fundamental vibrational level over two or higher energy levels) and combination modes (arising from the interaction of two or more vibrations taking place simultaneously in different functional groups) give rise to very complex bands in the NIR spectrum that can give more complex structural information than MIR. On the other hand, NIR spectra are less selective than MIR spectra because of the superposition of different overtones and combination bands. A raw spectrum contains background information and noise beside valuable information. To remove those interferences as well as those coming from scattering, to normalise the effect of particle size and light distance, pre-processing methods such as smoothing, derivative, standard normal variate transformation (SNC), multiplicative scatter correction (MSC) or wavelet transforms (WT) are required. Recently, the introduction of the Fourier transform technique in IR (FTIR) has further increased the application range of the IR spectroscopy in the food field. In such a case the spectrum is obtained by mean of an interferogram in which multiple frequencies are measured simultaneously. The resulting interferogram is then deconvoluted using proper algorithms in order to have the original spectrum. The advantages of that technology is a faster analysis and a higher throughput and a better alignment of spectrum obtained by repetitive scans.

IR spectroscopy have been successfully applied to detect adulteration of juices, purees and syrups with cheaper juice concentrates. Adulteration of orange juice with orange pulpwash, grapefruit juice or synthetic sugars/acids mixture can be detected at a level as low as 50 g/kg by NIR [48]. Similarly, the adulteration of strawberry or raspberry juice with apple juice can be detected at level> 10% by transmittance NIR coupled with PLS [49]. MIR spectra have been used to detect adulteration of pure pomegranate juice with grape juice (2%-14% v/v) [50] and to predict the percent fruit content in strawberry jam [51]. Adulteration of honey and maple syrup can also be detected by NIR and MIR spectroscopy. NIR and FTIR have been successfully applied for the detection and quantification of cane and beet sugars in maple syrup [52]. Attenuated total reflectance (ATR)-FTIR coupled with LDA and PLS was used to discriminate the type of adulterant in three different honey varieties. A success rate of prediction of 100% was achieved for honey samples adulterated with 7-25% w/w of simple (glucose, fructose, sucrose) and complex (beet and cane invert) sugars [53]. NIR and MIR have also been employed for the authentication of lard and fats. Lard adulteration can be detected in cake [54] and in chocolate [55]. The adulteration of shortening with lard can be detected at levels ranging from 0 to 100% and a standard error of calibration (SEC) of 1.75 by using the regions 1.117-1.097 cm\(^{-1}\) and 990-950 cm\(^{-1}\) of the NIR spectrum. Adulteration of
olive oils has also attracted much attention due to the economic value of the product. Adulteration of EVOO with palm oil can be detected by FTIR and PLS in concentration varying from 1.0 to 50.0% w/w [56]. The region 1500-1000 cm$^{-1}$ of the MIR spectra was used for the regression model. The adulteration of EVOO with sunflower, corn, soyabean and hazelnut oil can be detected at level as low as 5% by using FTIR and LDA [57]. In this latter case, the normalized absorbance of peaks and shoulders areas were used in the model as predictors. The standard of identity for butter require that no vegetal oil nor margarine is added to the product. The presence of margarine can be detected by NIR coupled with PLS in the range 0-100% with a standard error of calibration after validation (SECV) <1.2% [58]. IR spectroscopy has been also widely used to predict and control meat quality. The discrimination between fresh and frozen-thawed beef can be accomplished by IR spectroscopy due to modification of the myofibrillar proteins and the corresponding change in their water holding capacity [59].

Infrared spectroscopy is a well-established technique for fast, high-throughput and non-destructive analysis of food and other biological samples. The analysis can be easily implemented on-line, can be automated and does not requires trained personnel to be carried out. It is little expensive and environmental friendly since does not require solvents, chemicals and does not produce waste. However, even though the analysis per se (collection of the spectrum) is fast, post-processing (pre-processing of the spectra and model building) of the input data can be laborious and time-consuming. The calibration models are usually built against reference analytical methods so that the measurement errors accumulated and the total predictive error increases. Finally, the classification or regression models are theoretically valid only on the instruments with which the training and the calibration has been carried out. The transfer of a multivariate model to other instruments affects its precision and accuracy compared to the original ones.

A case study on NIRS and adulteration.

As an example of the potential of the NIR Spectroscopy for the detection of adulteration, the results of an investigation that has been carried out in the authors’ research group will be shown. NIR spectra were used to detect the presence of nitrogen replacers in milk powders. The compositional standards for milk powders require that the amount of milk proteins in milk solids-not-fat should be at least equal to 34% m/m, unless declared. The low prices of some nitrogen containing compounds make them attractive as potential adulterants to increase the level of apparent proteins in milk powders. The Kjeldahl method (official reference method for proteins content) measures the total amount of N irrespective of whether it comes from proteins or not. Expensive and time-consuming analytical methods such as enzyme-linked immunosorbent assay (ELISA), LC-MS/MS and GC-MS/MS are necessary for confirmatory analysis of melamine and its analogues in milk powders. To prove the potential of NIR spectroscopy for the detection of such adulteration, 33 skim milk powders were randomly adulterated with adulterants ammonium chloride, caprolactam, diammonium phosphate and polyvinylpyrrolidone (PVP) in order to produce an increment of 0.10, 0.50, 1.00 and 2.00% in the (apparent) proteins content of the milk powder. The samples were measured by NIR spectroscopy and the spectra subjected to PLS-DA analysis. A few milk
powders were randomly selected and adulterated by melamine, ammelide and urea to test the robustness of the predictive models.

A PLS-DA model was first developed to predict the type of adulterant. The training set consisted of 80% genuine milk powders (26 samples) and 80% adulterated samples (19 samples from each adulterant) which were randomly selected. The remaining 20% of the samples, and those adulterated with the non-modelled adulterants melamine, ammelide and urea were used for external validation. The success rate of prediction was 100% in cross-validation and 78% for the external validation set (Table 2). Three out of 6 samples adulterated with melamine, ammelide and urea were correctly predicted as adulterated.

<table>
<thead>
<tr>
<th>Class item</th>
<th>Genuine powder</th>
<th>+ NH₄Cl</th>
<th>+ caprolactam</th>
<th>+ (NH₄)₂HPO₄</th>
<th>+ PVP</th>
<th>No match</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLS-DA model based on training set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genuine powder</td>
<td>26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ NH₄Cl</td>
<td>-</td>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ caprolactam</td>
<td>-</td>
<td>-</td>
<td>19</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ (NH₄)₂HPO₄</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ PVP</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19</td>
<td>-</td>
</tr>
<tr>
<td>External validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genuine powder</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ NH₄Cl</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>+ caprolactam</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+ (NH₄)₂HPO₄</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>+ PVP</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>External validation based on melamine, ammelide and urea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ melamine</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>+ ammelide</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+ urea</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Prediction results of PLS-DA model for the type of adulterant in milk powders

A PLS-DA model was then developed to discriminate generally between genuine and adulterated samples. The training set and the validation set were built in the same way as previously described. The PLS-DA scores plot is presented in Figure 4. The prediction results are reported in Table 3.

In external validation only one genuine sample was wrongly predicted as adulterated. One adulterated sample was incorrectly predicted as genuine (PVP added at its lowest concen-
tration) and 3 adulterated samples could not be classified. Samples adulterated with melamine, ammelide and urea were all correctly predicted as adulterated.

![Plot of PLS-DA scores for IR spectra of genuine and adulterated milk powders. Red = adulterated powders; green = genuine powders.](chart.png)

**Figure 5.** PLS-DA scores plot of IR spectra for genuine and adulterated milk powders. Red = adulterated powders; green = genuine powders.

<table>
<thead>
<tr>
<th>Class Item</th>
<th>Genuine powder</th>
<th>Adulterated powder</th>
<th>No match</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Internal validation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genuine powder</td>
<td>24</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Adulterated powder</td>
<td>0</td>
<td>68</td>
<td>8</td>
</tr>
<tr>
<td><strong>External validation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genuine powder</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Adulterated powder</td>
<td>1</td>
<td>22</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table 3.** Prediction results of PLS-DA model for genuine versus adulterated milk powders.

NIRS combined with chemometrics proved to be a promising tool for the cost-effective detection of adulteration of milk powders with a range of nitrogen replacers. The reliability of the classification models can be improved by the careful selection of those regions of the NIR-visible spectrum which showed the best discrimination power between genuine and adulterated samples.
7. Other fingerprintings

*Differential scanning calorimetry (DSC)*

Differential scanning calorimetry is a thermoanalytical technique that measures the amount of heat required to increase the temperature of a sample relative to a reference material. When the amount of heat absorbed/released is plotted as a function of temperature a thermal spectrum of the sample is obtained from which kinetic and thermodynamic information such as the heat capacity and the enthalpy of any phase transition (fusion, evaporation, glass transition etc) the sample undergoes in the temperature span of the experiments can be determined. Those physical properties depend in turn on compositional and structural properties of the samples.

DSC has been mainly applied for the authentication of fats/oils. The level of adulteration of canola oil with lard, beef tallow and chicken fat as well as that of virgin coconut oil with palm kernel oil (PKO) and soybean oil can be predicted by the analysis of the DSC thermogram and stepwise multilinear regression (SMLR) [60]. The melting profiles of cow, goat, sheep, camel, horse and water buffalo milk fat samples were also determined by DSC. Differences in the DSC profiles for the fat fraction of milk of different species can be attributed to their fatty acid/TAG composition. In general, the melting point of the fats decreases with decreasing chain length and increasing degree of unsaturation of the fatty acids in the milk. The thermograms were subjected to PCA analysis which showed a clear distinction of the four milk samples analysed. The loadings plots of the heat flow data showed that the data are most influenced between the temperatures of 13° and 24° C [61].

*Sensor technology*

Sensor technology is often referred to as electronic nose technology. In this technology the volatile compounds present in the headspace of a sample are detected by an array of semi-selective sensors. Normally each sensor is sensitive to all the volatile but each in a peculiar way. There are many sensors for EN instruments such as metal oxide sensors, conducting polymer sensors, quartz crystal membrane sensors, or mass spectrometers (MS-EN) available on the market. The sensor’s response is then transformed in a spectrum that represents a fingerprint of the volatile compounds of the tested sample. Electronic nose analysis are often cheaper and faster than GC analysis and the sample preparation is usually quite simple. The technique, however, does not have the same sensitivity of other techniques. Furthermore, single volatile compounds cannot be identified and the signal is sensitive to water vapour.

Electronic nose has been used mainly for the detection of adulteration in fats/oils. Rapid detection of pork and lard in food samples for halal authentication (compliance with Islamic dietary rules) are reported with electronic nose [62]. Under Jewish and Islamic dietary laws, foods containing porcine-based ingredients such as lard are strictly prohibited from consumption. Electronic nose and chemometric analysis was applied for the detection of adulteration of olive oil samples with sunflower and olive-pomace oil at levels as low as 5% [63]. Application are also reported for the authentication of the geographical origin of Italian wines [64], and emmenthal cheese [65].
Transcriptomics

Transcriptomics is a post-genomic technique that consists in the simultaneous measurements of all the transcripts (mRNA molecules) in a given organism, or of a specific subset of transcripts present in a particular cell type. Unlike genome, transcriptome varies according to the environmental conditions and represents the genes that are actively expressed in a certain cell at a certain time. Transcriptome is usually obtained by DNA microarray technologies and reverse PCR. The set of all the mRNA produced (and hence of the genes actively expressed) represents a fingerprint of that target cell/organism and can thus use for fraud detection and authentication purposes. At the present the potential of transcriptomics in this respect is still underexplored.

8. Fingerprinting options for other biomaterials

Biofuels are an important environmental-friendly alternative to fossil fuels. The term biodiesel refers primarily to FAME obtained after transesterification of triglycerides with methanol. The methyl esters can be produced from many different triglyceride sources, primarily rapeseed oil but also sunflower oil, soybean oil, palm oil, linseed oil, tallow, and used frying oil. Blends of biodiesel with conventional petrodiesel fuel represent a common utilization of biodiesel. The ability to predict retail biodiesel blend percent composition is important to detect adulterations. It has been reported that biodiesel from different sources can be discriminated by direct infusion ESI-MS and multivariate statistics [66]. In addition, the % of rapeseed or salmon biodiesel in petrodiesel in concentrations ranging from 0.5% to 10% can be predicted with good accuracy by applying a PLS model. Similarly, the feedstock source of blends of biodiesel and conventional diesel, as well as the % composition of the blend can be predicted applying a kNN and a PLS model respectively to the total ion current chromatograms from gas chromatography–quadrupole mass spectrometry (GC–qMS) using a polar column [67]. The precision of the prediction was between 4-5%. Furthermore, the addition of residual oil (non-transesterified residual vegetable oil) is one of the easiest ways of adulterating biofuels. Synchronous fluorescence combined to LDA can be used to discriminate between diesel oils, biodiesels and biodiesel adulterated with residual oil and the % of residual oil can be predicted with good accuracy when a PLS model is used [68].

Perfume counterfeiting is an illegal practice that causes huge economical loss to the perfume industry and pose potential health risk to consumers who might be exposed to harmful chemical from counterfeited products. Traditionally, the quality control analyses for perfume focuses on volatile and semi-volatile compounds and are performed by GC based techniques. Recently, more straightforward methods have been proposed for the fast authentication of premium perfumes and detection of adulteration/counterfeiting. A fingerprint of the polar compounds can be achieved by direct infusion EASI-MS (see section 3) [69]. The samples are sprayed onto a glass rod and directly exposed to the ionisation source of the MS system. An almost instantaneous, simple and reproducible fingerprint of the polar compounds in the product is obtained that allows a complete discrimination between au-
thentic and counterfeited products. A fast discrimination between authentic and counterfeit-
ed products can be also obtained by fingerprinting of the polar compounds by ESI-MS in the
positive ion mode and chemometric analysis previous extraction of few µL of the sample in
a 1:1 methanol/water solution [70]. Even more recently, a fast, simple and low-cost method
for the authentication of perfume based on a commercial electronic olfactory system (EOS)
equipped with thin film metal oxide semiconductors has been proposed [71]. The PCA anal‐
ysis of the R/R₀ values (resistance of the sensor in the presence of the volatile compounds
relative to that of the sensor balanced in air) generated by 6 sensors based on different metal
oxide semiconductors can unequivocally discriminate between authentic and counterfeited
perfumes. The prior removal of ethanol from the samples is necessary for the correct dis‐
crimination of the samples.

Essential oils are also widely employed for their fragrance in perfumery but also in cosmet‐
ics and household products. One of the most common fraud is the blend of valuable essen‐
tial oils with other less valuable alternatives. As an example, the valuable Rosewood
essential oil, obtained from the trees of Aniba rosaeodora Ducke and employed in fine per‐
fumery, can be blended with the far cheaper synthetic linalool, obtained by re-distillation of
Rosewood leaf oil. The ESI-MS fingerprint in the positive ions mode of the polar compounds
extracted by an acidified 1:1 water methanol solution can easily detect adulteration of Rose‐
wood oil with synthetic linalool at concentration as low as 10% v/v [72]. A PCA model is nec‐
essary to extract the relevant information from the fingerprint.

Direct infusion MS can be used to authenticate wood as well. Venturi easy ambient sonic spray
ionization (V-EASI)-MS fingerprint of a very simple methanolic extract of wood chips or
directly acquired from the freshly scratched wood surface may help to control the illegal
logging and trade of the noble Mahogany tropical wood and its falsification [73]. V-EASI-
MS is a novel ambient ionization MS technique characterized by sonic spray ionization and
a self-pumping system based on the Venturi effect. It allows the direct analysis of solid or
liquid samples. Ionization is assisted by compressed nitrogen and the apparatus is thus free
of electrical discharge, thermal interferences since no heating, voltage or radiation is used.
The introduction of fast and high-throughput analytical techniques for wood authentication
is especially valuable since the classical controls are based on time-consuming morphologi‐
cal evaluations. Recently an original and fast approach for the authentication of wood spe‐
cies has been proposed. It is based on the analysis of the volatile compounds measured by a
low-cost conductive polymer-based portable electronic nose formed by an array of only three
gas sensors and/or the elemental fingerprint measured by laser-induced breakdown spectrom‐
etry (LIBS) which performs a multielemental and direct analysis even in solid samples [74].

The assurance of quality of herbal supplements and medicines is a major concern for the
phytopharmaceutical and the food industry. The identification of the herbal drug and the
presence of adulterant is a mandatory test to ensure the quality, the efficacy and the safety
of a medical preparation or an herbal supplement. Among a variety of quality control meth‐
ods, chromatographic fingerprinting has gained more and more attention recently and have
been used to authenticate a large number of herbal products. They are accepted by many
international organization for the quality control of herbal medicine but are relatively time-
consuming. More recently, spectroscopic techniques have been explored to rapidly authenti-
cate herbal products. For instance, different species of *Echinacea*, e.g., *E. purpurea*, *E. angustifolia*, and *E. pallida* are used for commercial preparations to prevent or cure the common cold, flu, and several other diseases due to their nonspecific stimulating effect of the immune system. It is also well known that *Echinacea* preparation are commonly adulterated with roots of *Parthenium integrifolium* L. This adulteration can be detected by NIRS at a minimum of 10% of adulteration [75]. The method requires just the milling of the sample and can be carried out within 1 minute.

The list of potential applications of products fingerprinting is not limited to the cases discussed above and many other examples may be provided. In the authors’ research group for instance, PTR-MS and ICP-MS are used for the authentication of the geographical origin of flower bulbs. Counterfeiting of pharmaceuticals is another area where the application of fingerprinting techniques has proved of great help. The topic is so huge that we would address the interested readers to specialized publications.

9. Conclusions

Product fingerprinting combined with chemometrics represents a valuable tool for fraud detection and control for food products and other biomaterials. A fingerprinting approach is particularly useful:

- For the authentication of products for which target analyses based on specific markers are not available
- For the detection of adulteration based of yet unknown adulterants
- For a fast and high-throughput screening of the samples before more elaborated confirmatory analysis are applied.

At the same time a fingerprinting approach may substantiate nutritional, sensory or other product qualities.

LC, GC, and IR spectroscopy are already common instrumental platforms available in most QA laboratories and they will continue to provide valuable support for food fraud prevention. IR and other spectroscopic techniques have the great advantage of providing fast, high-throughput and non-destructive analyses with limited costs. They can be easily automated and adapted for *in-line or in-situ* analysis which makes these techniques well suited for implementation in the industrial setting. MS and NMR are not as common in QA laboratories, principally because of their high costs but they may become more important in routine QA testing because of their superior performances. However, even though the costs of MS and NMR instruments is still very high, the cost per sample can be very low if a high samples turnover can be achieved.

From an analytical point of view, a further improvement is expected in the future from the broader application of multi-dimensional separation techniques such as GCxGC or LCxLC which provide enhanced resolution and an higher number of peaks. Ultra-performance liq-
uid chromatography (UPLC) and CE may also provide rapid separation with limited costs. The application of novel (or still underexploited) ambient ionisation MS techniques that allow the rapid analysis of liquid and solid samples with little, if any, preparation will be particularly valuable in the future.

The applicability and the reliability of a fingerprint approach also depends from the correct and tailored usage of the relevant and appropriate chemometric tools. For the development of regression and/or classification models, special care should be devoted to ensure:

- The representativeness of the classes considered, in order to cover all the possible source of variability for the class at stake.
- A robust validation of the model (external validation to be preferred over internal validation).
- The use of the appropriate chemometric tools depending on the problem at stake. Whereas a few pattern recognition techniques are frequently used (PLS-DA, SIMCA, LDA...), some other such as classification and regression trees (CART), quadratic discriminant analysis (QDA) are still underexploited despite the good results that they can provide.

Another key aspect is represented by the validation of methods based on fingerprinting and chemometrics that is essential for their application in a commercial context. Whereas standards exist for the validation of regular analytical methods (see for instance, Commission Decision 2002/657/EC and ISO 17025), internationally accepted protocols for the validation of methods based on fingerprinting techniques and chemometric classification models are lacking at the present. Such protocols should indicate the performance characteristics that have to be checked and the criteria to be met in order to verify the compliance of the method with the performance characteristics.

Acknowledgements

We acknowledge Martin Alewijn, Grishja van der Veer and Pan Weijing for the work on NIR spectroscopy and milk powder and Angela Dura de Miguel, Arjen Lommen and Maikel Rozijn for the work on 1H NMR and fish. We also acknowledge the Dutch ministry of agriculture, economic affair and innovation for the funding of this book chapter.

Author details

Edoardo Capuano* and Saskia M. van Ruth1

*Address all correspondence to: edoardo.capuano@wur.nl

1 RIKILT – Institute of Food Safety, Wageningen University and Research Centre, The Netherlands
References


(EV with ON) as determined by direct infusion electrospray ionization mass spectrometry and chemometric approaches. DOI: 10.1002/rcm.4590.


