European Network for assuring food integrity using non-destructive spectral sensors (SensorFINT).

Protocol and guidelines for recommendations on NDSS combination for *in situ* analysis of food products

1.- Non-Destructive Spectral Sensors (NDSS)

Non-destructive spectral sensors (NDSS) refer to instruments capable of measuring the interaction between electromagnetic radiation and matter without causing any alteration or damage to the sample. In fact, these sensors do not physically or chemically alter the sample during measurement, preserving its integrity for further analysis. They operate across a range of wavelengths to obtain valuable information about the composition, structure, and properties of materials, including food products. Several NDSS are commonly used in the food industry and in the agronomic production for various applications, such as Near- and Mid-Infrared (NIR and MIR, respectively), Raman, Ultraviolet-Visible (UV-Vis), Fluorescence and Hyperspectral Imaging (HSI). All these sensors can be used to perform in situ analyses, which involve the real-time or on-site assessment of a sample within its natural environment or production setting, allowing for immediate and direct measurement, enabling timely decision-making and process monitoring and optimization.

Several aspects make in situ analysis a powerful tool to improve the efficiency of quality monitoring in general or specific properties in particular. Firstly, by analyzing the sample directly in its natural environment, there is no need for sample extraction or manipulation, reducing the risk of contamination or alteration. It also eliminates the need for transporting samples to a lab, which can be time-consuming and may lead to changes in sample properties over time. As a matter of fact, it can be performed without interrupting production processes, allowing for continuous monitoring without slowing down operations. These analyses provide immediate feedback on the quality, composition, and characteristics of food products, allowing for quick decision-making during processing or quality control. Therefore, they are particularly valuable in production sites of food industries, where rapid and accurate information about product quality, safety, and processing conditions is crucial for ensuring compliance with standards and meeting consumer expectations. Lastly, in situ analyses can be more cost-effective than traditional lab-based methods, as they may require fewer resources and specialized equipment and can cover a larger number of samples.

Generally speaking, there are many aspects to consider when a specific sensor is chosen in place of another one. Two of the most important parameters to assess are

accuracy and precision. Accuracy reflects the closeness of measurements to the true value, while precision indicates the consistency of repeated measurements. Both attributes are crucial for reliable data acquisition, and it is fundamental to verify the accuracy of the sensor through calibration against known standards and validate its performance in real-world conditions, as well as to evaluate the precision of the sensor by conducting multiple measurements on the same or similar samples (repeatability and reproducibility, respectively). Another aspect to consider is sampling rate, which refers to the frequency at which measurements are taken, while resolution pertains to the level of detail or granularity in the measurements. For dynamic processes or rapidly changing conditions, a higher sampling rate may be necessary to capture relevant information. Moreover, is important to consider the size and granularity of the areas being measured. Higher resolution may be required for small or heterogeneous samples. Going forward, portability refers to the ease with which a NDSS can be transported and operated in different field or production environments. User-friendly interfaces and intuitive operation are vital considerations. A key aspect is to assess whether the sensor can be easily transported and set up in diverse field conditions, also considering the level of training required for personnel to operate the sensor effectively. Lastly, it is important to make consideration about the cost of an instrument in terms of initial investment, maintenance expenses, and potential upgrades.

Food products encompass any substances intended for human consumption, including raw agricultural commodities, processed goods, and beverages. These can be of plant, animal, or microbial origin and undergo various stages of production, processing, and preparation. Food products also range from fresh fruits and vegetables to processed items like dairy products, grains, and packaged goods. Each category presents distinct spectral characteristics that must be considered when selecting sensors for in situ analysis. Understanding the optical properties of the food (e.g., transparency, scattering) is essential for selecting sensors that can effectively penetrate and interact with the sample. In the next section will be inspected the NDSS mainly used for in situ analysis of food products, listing their main properties, applications, and selection criteria in respect to other sensors.

1.1 Near-Infrared Spectroscopy

Among the above-mentioned techniques, the most used are certainly the ones who relies on vibrational spectroscopy, due to being rapid, non-invasive, able to predict many parameters at the same time and requiring very low or even absent sample preparation. In particular, NIR spectroscopy is widely used for rapid analysis of various components in food, including moisture content, fat content, protein content, and more. This analytical technique involves high-energy vibrational spectroscopy within the wavelength range of 800 to 2500 nanometers (12500-4000 cm^{-1}), providing physio-chemical information on the sample. Spectral data acquired within this range are subsequently employed for both qualitative and quantitative analysis of various foods and food products in conjunction with chemometric methods. The interaction of NIR radiation with the sample can occur through different means, i.e. transmittance, diffuse reflectance and transflectance. These terms refer to specific geometric configurations of the probing radiation beam, sample, and detection system, which are utilized to gather analytical spectral information about the sample.

Different modes of NIR spectroscopy measurements are employed depending on the type of sample. Commonly, transmittance mode is used for liquid and semi-liquid food samples, as the NIR light passes through the sample allowing to measure the transmitted light on the other side. However, it can be also used to analyze thin solid samples or gases. This acquisition mode is particularly useful for assessing parameters as moisture content, sugar content, and other soluble constituents, and finds applications in samples like oil, milk, juices, sauces, honey, soft drinks, alcoholic beverages (wines, beer, spirits and liqueurs) [1-7] and so on, which are contained in quartz or glass cuvettes of various sizes. On the other hand, diffuse reflectance is used for solid samples, and it is particularly suitable for samples where it is not feasible or practical to use transmittance due to the sample's physical form, such as grains, cereals, flours, seeds, nuts, tea, coffee, herbs, fruits, vegetables, powders (milk, whey proteins, cocoa, spice blends, etc.) [8-17] and many more. It provides valuable information about the composition and properties of the samples. NIR transflectance, also known as diffuse transmittance, is used for food samples that are semi-solid, such as gels and emulsions, or have some level of opacity, transparency, or translucency. This acquisition mode involves NIR radiation onto the surface of the sample and collecting the radiation that is both transmitted through the sample and reflected back from the surface, allowing the analysis of both surface and internal properties of the sample. It is used to analyse semi-solid sauces (such as pesto),

yogurts, soft cheeses, processed meats (patés and sausages), nut butters, jellies, jams, mayonnaise-based dressings [18-23] and other similar products.

Among the greatest advantages of NIR, there is the possibility to use fiber optics, which allow the remote placement of the spectrometer from the sample. This is especially valuable in situations where direct contact with the sample is impractical or not possible. Furthermore, they can be designed for specific applications, allowing handling specific sample types, shapes, and sizes, and can be coupled with various types of sampling probes, including reflectance, transflectance, and transmission probes. This flexibility enables the analysis of different types of samples, such as liquids, solids, and powders. Overall, the flexibility and adaptability of fiber optic probes allows the real-time monitoring of processes, making them invaluable in food industries where timely interventions or adjustments are critical.

Lastly, the reason that makes NIR spectroscopy the most used NDSS for in situ analysis the wide range of portable instruments present in the market, as their construction is easier in respect to other spectroscopic techniques. This leads to the development of portable spectrometers designed for a specific application/kind of sample. Their use enables on-site analysis of food products directly in the production or processing environment, eliminates the need to transport samples to a central laboratory, saving time and resources [24].

1.2 Mid-Infrared Spectroscopy

MIR spectroscopy operates in the wavelength range of 2.5 to 25 micrometers (4000 to 400 cm^{-1}) in the electromagnetic spectrum, covering a broader range of wavelengths than NIR and making it suitable for applications like the identification of functional groups and chemical bonds in food products. It relies on the characteristic atomic vibration patterns within a molecule, leading the identification and quantification of various components in food products, such as carbohydrates, lipids, proteins, and additives [25-28]. The applications of MIR for food analysis often overlaps with the NIR ones regarding the kind of food products, therefore the choice of one rather than the other is not always clear. The absorption bands in MIR spectra are well-defined and correspond to fundamental vibrational modes of molecules rather than overtones and combination bands in NIR, making interpretation and identification of compounds more straightforward. However, MIR radiation is heavily affected by water absorption, making

it not suitable for samples with high water content. Generally, MIR is better for analyzing functional groups, while NIR is effective for quantifying organic compounds and moisture content.

The MIR spectrometers acquisition modes, transmission and diffuse reflectance, works in the same way as already described in section 1.1 for NIR spectroscopy. The use of fiber optics in MIR is possible, it is less straightforward than in the NIR range, as common optical fibers made of silica glass have high absorption in MIR spectral range, which limits their transmission efficiency and attenuates the signal significantly as it travels through the fiber, reducing the signal-to-noise ratio. To overcome the limitations of standard silica fibers, specialized MIR fibers made of materials like chalcogenide glass or fluoride glass have been developed, but are less common and can be more expensive. MIR portable spectrometers are present in the market, although to a lesser extent than NIR ones. However, they are used to assess the quality of raw materials, ingredients, and finished food products by analyzing parameters like fat content, protein content, and other constituents [29-31].

1.3 Raman Spectroscopy

Raman spectroscopy is another technique which exploit vibrational spectroscopy to get detailed information about chemical composition, structural conformation, and identification of compounds in food samples [32-33]. Unlike MIR and NIR spectroscopies, which rely on absorption, Raman spectroscopy is based on inelastic scattering. When monochromatic light interacts with a sample, most of the photons are elastically scattered at the same energy as the incident light (Rayleigh scattering). However, a small fraction of the incident photons undergo inelastic scattering, resulting in a shift in energy. This is known as Raman scattering, and the difference in energy between the incident and scattered photons is called Raman shift. It corresponds to the energy associated with vibrational or rotational transitions within the sample. In Raman scattering, two main spectral lines are observed: Stokes lines and Anti-Stokes lines. Stokes lines are shifted to lower energy (longer wavelength) compared to the incident light, while Anti-Stokes lines are shifted to higher energy (shorter wavelength). The obtained spectrum contains peaks at specific Raman shifts corresponding to different vibrational modes in the sample, i.e. provides information about specific chemical bonds and functional groups. This method is sensitive to potential interference such as

fluorescence, a concern less prevalent in IR spectroscopy. As NIR and MIR, also in Raman spectroscopy it is possible to present the sample in different states. However, it is less common to find applications on non-fluorescent liquid samples, (clear solutions, extracts, or other transparent liquid-based food products).

Raman offers distinct advantages, including exceptional specificity with well-separated peaks. Raman and IR spectroscopies are complementary; in fact, Raman primarily measures the changes in polarizability of a molecule due to molecular vibrations, making it particularly sensitive to non-polar vibrations and, therefore, useful for compounds with symmetric structures. IR spectroscopy, on the other hand, measures the absorption of energy by the molecule as it undergoes vibrational transitions. It is sensitive to polar vibrations, that result in a change in dipole moment. Notably, Raman signals associated with water are relatively low; consequently, this approach proves valuable in analyzing liquid samples without the complications arising from water-induced signal masking—a challenge frequently encountered in IR spectroscopy, especially in MIR. Lastly, it has to be pointed out that different parameters have to be tuned for each specific application, such as the wavelength and the power of the incident laser, as well as the acquisition time and the number of accumulations on the sample. The availability of portable Raman spectrometers has been increasing over the years, driven by advancements in technology and a growing demand for Raman spectroscopy in various fields, including food analysis [34-36].

1.4 Ultraviolet-Visible Spectroscopy

UV-Vis spectroscopy covers the wavelength range from about 200 to 800 nanometers. The UV region (200-400 nm) corresponds to higher-energy transitions, while the visible region (400-800 nm) corresponds to lower-energy transitions. UV-Vis spectroscopy provides information about the electron transitions and energy levels of atoms and molecules, making it a valuable analytical technique in the food industry for assessing the quality, composition, and characteristics of various food products. It is commonly used to analyze the color of food products, as measure the absorbance or transmittance of light at specific wavelengths, providing information about the hue, chroma, and intensity of color [37]. Other common application are the identification and quantification of food additives (colorants, preservatives and antioxidants) [38-40], the study of natural pigments in foods (anthocyanins, carotenoids, chlorophylls, and

betalains) [41-44], the monitoring of the Maillard reactions progress [45] and the analysis of phenolic compounds, vitamins, food oxidation [46-48]. The wide range of application of these techniques in addition to its relatively low costs makes it suitable to be chosen for food analysis, using both benchtop and portable instruments. Many food samples are analyzed in a liquid state. This is suitable for clear liquids, juices, extracts, and other liquid-based food products [49]. The sample is typically diluted in a suitable solvent if necessary, and measurements are taken using cuvettes. Extraction or dissolution in a solvent is a common practice when specific compounds like pigments, polyphenols, or other UV-absorbing substances are studied. For some food samples with a transparent or translucent nature, direct measurements can be taken in the solid state, such as clear fruit juices, honey, or transparent gels [50]. In general, however, UV-Vis spectroscopy requires more sampling preparation in respect to vibrational spectroscopy techniques, on average.

Normally, UV-Vis spectroscopy is preferred for compounds with extended conjugated systems, like aromatic compounds, whereas vibrational spectroscopy is more versatile for a wide range of molecular vibrations. The choice also depends on the concentration levels of the compounds under investigation and the general complexity of the sample. Indeed, UV-Vis spectroscopy is generally more suitable for higher concentrations, whereas vibrational spectroscopy can be used for both low and high concentrations, and it may be more suitable for its ability to provide detailed information about chemical bonds and functional groups.

1.5 Fluorescence Spectroscopy

Fluorescence spectroscopy is used to study the interaction of electromagnetic radiation with fluorescent molecules. It is applicable to molecules, known as fluorophores, which have the ability to absorb and emit the radiation. Common natural fluorophores include certain aromatic compounds, dyes, and biomolecules like chlorophyll and fluorescent proteins. As instrument outputs, the excitation spectrum shows the intensity of emitted fluorescence as a function of excitation wavelength, whereas the emission spectrum displays the intensity of emitted fluorescence as a function of emission wavelength. Applications of Fluorescence spectroscopy for food analysis goes from assessment of quality and freshness of food products, detecting changes in the composition of food due to spoilage, degradation, or contamination [51-53], to detection

of contaminants (pesticides, heavy metals, and food dyes), determination of vitamins, minerals, proteins and lipids, especially for studying lipid oxidation, a major factor in food deterioration [54-57]. The use of Fluorescence spectroscopy over other NDSS sensors relies primarily on the presence or absence of fluorophores in the sample that are highly sensitive, such as aromatic compounds and certain biomolecules.

Food samples can be presented to a fluorescence spectrometer in various states, depending on the specific analysis and the nature of the food product, similar to the description present in section 1.4 for UV-Vis spectroscopy. Many food samples are analyzed in a liquid state. This is suitable for juices, sauces, extracts, and other liquid-based food products. The sample is typically diluted or extracted in a suitable solvent, and measurements are taken using cuvettes or specialized cells. Some food samples may need to be extracted or dissolved in a solvent to access the fluorescent compounds of interest. This is common for studying specific compounds like pigments, flavonoids, or vitamins. Other food products may exist as suspensions or emulsions, where solid particles or droplets are dispersed in a liquid matrix. Examples include salad dressings, milk, or mayonnaise. Some food samples, particularly those with solid or semi-solid textures, can be analyzed directly in their natural state. Examples include fruits, vegetables, meats, or baked goods. However, the sample may be sliced, homogenized, or prepared in a way that allows for fluorescence measurements.

1.6 Hyperspectral Imaging

HSI is an advanced technique that combines the power of imaging and spectroscopy to capture detailed information about a sample's spectral characteristics across a wide range of wavelengths. It provides a rich dataset with spectral information for each pixel in an image, allowing for comprehensive analysis and identification of materials based on their unique spectral signatures. HSI systems cover a broad range of wavelengths, typically spanning from the UV-Vis region to the NIR region, even if some instruments cover even MIR region. The power of this technique relies on the double information provided, that is spectral resolution and spatial resolution. Spectral resolution offers high spectral resolution, meaning it can distinguish between narrow spectral features, crucial for identifying subtle differences in materials based on their spectral signatures. Spatial resolution refers to the level of detail with which the spatial features of an object or scene are captured by the imaging system. It is a critical parameter as it

determines how finely the image can represent the spatial characteristics of the objects being observed. The data generated by HSI is typically organized in a data cube, where each pixel in the image contains a complete spectrum, allowing for simultaneous analysis of both spatial and spectral information. One of the primary strengths of HSI is its ability to identify and classify materials based on their spectral characteristics, providing at the same time information on their position in the sample. A great advantage of HSI in respect to other NDSS is that the measurements can be performed using airborne or satellite-based platforms. This allows for large-scale, remote sensing applications, such as environmental monitoring and agricultural assessments. Nevertheless, despite its power, HSI can generate large datasets that require advanced computational resources for analysis, and the interpretation of hyperspectral data can be complex and may require expertise in both imaging and spectroscopy. In fact, analyzing this kind of data requires sophisticated processing techniques and chemometric algorithms. This may involve techniques like spectral unmixing, classification and feature extraction to extract meaningful information from the data.

This technique is used for quality assessment of food products, as can identify defects, bruises, blemishes, foreign objects and other quality-related issues in fruits, vegetables, and processed food items [58-60]. Moreover, it can determine the freshness of perishable food products like seafood, meat, and fruits, being able to detect changes in color, texture, and biochemical composition that occur during spoilage [61-64]. Regarding food safety, HSI is employed to identify contaminants, such as foreign materials, mold, pathogens or chemical residues [65-67]. Furthermore, it can determine the maturity and ripeness of fruits and vegetables based on their biochemical composition to optimizing harvest times and managing post-harvest handling [68-69].

Deciding whether to use HSI or other NDSS for food analysis depends on several factors. First of all, it has to be determined if detailed spatial information about the samples is needed. HSI provides both spectral and spatial data, which can be crucial for certain applications. Then, it should be considered whether high spectral resolution is crucial for your analysis since certain applications may require fine spectral discrimination, while others may not. Moreover, it has to be kept in mind the amount of data that will be generated, as HSI produce large datasets that require specialized processing techniques and advanced chemometric tools for the data analysis and interpretation.

2 Good Practices in Data Acquisition

Accurate and reliable data acquisition is fundamental to the success of any on-site analysis using NDS. To achieve this goal, it is recommended that to carry out a few critical steps.

2.1 Sampling

When performing sampling on food products, it's essential to follow a structured approach to ensure representative and reliable results. First of all, it is needed to clearly define the purpose of the sampling, whether it's for quality control, compliance testing, authentication or other specific objectives. It is crucial to choose a sampling plan that is appropriate for the specific food product, taking into consideration factors like batch size, production process, and variability within the product. Sampling foodstuff — especially for authentication purposes such as PDO (Protected Designation of Origin), PGI (Protected Geographical Indication), etc. — entails two similar steps: ensuring representativeness concerning producers (taking into account the impact of their production on the market), year/season, variety/species, and all elements that may vary according to production protocols (e.g., raw materials and production practices), and subsampling from production batches/lots. Similarly, when assessing quality or traceability along the production chain, representativeness in sampling must be ensured both in terms of raw materials and the process settings over time, as well as how and how many (and how often) intermediate samples are taken. Finally, when obtaining an analytical sample from bulk samples, it is imperative to consider the fundamentals of sampling theory [70]. Ensure that the tools and equipment used for sampling are clean, sterilized, and appropriate for the type of food product being sampled is also important, as well as use a random selection process to choose sampling locations. This helps ensure that the samples are truly representative of the batch. Finally, even transportation and storage of samples should be planned and controlled to prevent spoilage or degradation.

2.2 Data acquisition

Many aspects should be considered to perform a reliable acquisition of the data. The NDSS needs to be properly calibrated according to the manufacturer's guidelines. This step is critical for accurate measurements. Obviously, appropriate settings and configurations should be selected, based on the specific requirements of the analysis (e.g., wavelength range, resolution). Even the position of the sensors needs to be chosen at the

correct distance and angle relative to the sample to ensure optimal data acquisition. If precision and accuracy are critical, consider taking multiple measurements of the same sample to ensure consistency. Environmental conditions, such as lighting, temperature, and humidity, can introduce variability in spectral measurements. Therefore, understanding and mitigating these factors is essential for obtaining consistent and accurate data. Implement control measures, such as shielding from ambient light, maintaining stable temperature conditions, and using appropriate calibration standards, is important to minimize the impact of environmental variability. For instance, humidity can heavily interfere with MIR and NIR measurements, as these techniques are highly sensible to water. On the same way, a non-homogeneous illumination can cause problems when performing HSI measurements. In general, temperature affects all analyses performed with the NDSS, so it should always be kept under control, or if this is not possible afterwards, some processing methods should be used to minimise this effect. In HSI, other key aspects to keep under control are the speed of bank movement (for line scan cameras), the positions of the lights (angle of the lights with respect to the lens axis) and the focus of the lens.

In general, a good practice involves the regular assessment of spectral data quality using metrics like signal-to-noise ratio, outlier detection, and spectral resolution. This is possible simply displaying the data and look at them, a practice as banal as it is rarely done, or performing an exploratory data analysis using simple chemometric tools like Principal Component Analysis (PCA), especially for outlier detection.

2.3 Preprocessing

Another aspect to be considered is that data collected from NDSS often contain other undesirable information, such as noise and background data, in addition to the relevant information. It is imperative to eliminate this irrelevant information. Therefore, preprocessing techniques are employed to construct reliable, precise, and consistent models. In contemporary practice, a variety of preprocessing techniques have been employed to enhance the robustness of models. These techniques encompass processes such as smoothing, centering, multiplicative scatter correction (MSC), normalization, wavelet transforms, orthogonal signal correction, standard normal variate transformation (SNV), straight-line subtraction, first and second derivative, and direct orthogonal signal correction, among others. Each data set obtained from these techniques is processed

according to its distinct characteristics. For instance, Raman spectra are often normalized and corrected for the baseline, whereas NIR spectra are usually pre-processed with techniques that eliminate scattering, such as SNV, MSC or peak decomposition techniques as derivatives. However, scattering effects are not always negative, as they provide information about the physical structure of the sample, such as granularity. Thus, users must decide, depending on the application developed, when it is appropriate to apply scatter correction, and when its use is not appropriate because we eliminate information that could be of interest in the modelling of the data. Therefore, there is no single recipe as to which signal pre-treatments should be applied.

2.4 Modeling

Choosing the proper modeling approach depends on many factors, such as the goal of the analysis, (classification, regression, exploratory analysis, clustering, feature selection, etc.), the number of samples available for analysis, the presence of outliers and noise in the data and the computational resources available (some methods may be computationally intensive and may not be suitable for large datasets). As a given example, when a food authentication study is performed, it is advisable to use a class modelling approach rather than a class discrimination. In fact, class discrimination aims to identify differences among samples from different categories. Consequently, every sample is assigned exclusively to one of the modelled classes. This implies that all classes must be characterized by representative sampling. Additionally, for this approach to be effective, each modelled class should have samples with similar characteristics. In certain authentication tasks, like determining if a food product matches the indicated PDO, a two-class discrimination scenario is often used: Class 1 represents the PDO category, and Class 2 includes all other samples. However, this can be problematic as Class 2 may lack common characteristics. Non-authentic PDO samples are typically diverse and may not share a distinct region in the feature space. Additionally, obtaining a representative sample from the entire non-PDO category is often impractical.

Class modelling methods, on the other hand, are designed to capture the similarities among samples within the same category. They focus on the unique characteristics of each class rather than on the differences between classes. Class modelling may be applied to one or several classes, with each category being modelled independently, regardless of the presence of others. In summary, discrimination is

suitable when it is clear which category or categories need to be distinguished, such as distinguishing between fresh and frozen fish [70].

Another possible issue can occur when dealing with a large process monitoring datasets. In this case, if the goal is to use the process data to predict in real time the quality parameters of the final product, it is advisable to select an algorithm that is computationally efficient and does not requires long times to build prediction models, such as PLS (Partial Least Square regression) or Response-Oriented Sequential Alternation (ROSA) [71], a relatively recent multiblock regression method that makes its computational efficiency one of its strengths. In general, when dealing with a limited number of samples, it is often advisable to lean towards simpler, linear models. This is because with a small dataset, there may not be enough independent samples to effectively optimize a large number of parameters, which can lead to overfitting. Furthermore, since the use of different spectroscopic techniques leads to obtain large datasets, the use of variable selection methods, such as PCA or CovSel [72], can help to get rid of irrelevant information and keep just the variables relevant for the specific study.

2.5 Validation

One of the first aspect to consider when performing data analysis and modeling is certainly the fundamental step of performing calibration and validation with two independent datasets. Calibration involves establishing a relationship between the measurements obtained from the NDSS and the “true” properties or composition of the sample. This step is crucial for converting raw spectral data into meaningful information. On the other hand, validation ensures that the calibration model performs accurately and reliably under different conditions or with new samples. Basically, it involves testing the model with independent validation data to assess its predictive capability. Validation necessitates that the raw data collected from new or upcoming samples (referred to as validation samples) are subject to the same sources of variability as the calibration data, and that the analytical measurements maintain the same level of quality. Hence, one must address sampling considerations, as discussed in Section 2.1, ensuring sample stability in relation to chemical/biological variability over time, accounting for matrix effects, maintaining instrument stability, and accounting for potential variations introduced by different operators or laboratories. It is also important to replicate the conditions under which future samples will be gathered. Validation helps to identify and mitigate

overfitting, which occurs when a model learns the noise or specific characteristics of the training data too well but fails to generalize to new data.

3 Combining Multiple Sensors

As described in the previous sections, each NDSS has its own advantages and disadvantages, making sometimes difficult the choice of the proper technique for a specific application. Different sensors provide different information about the sample, and not always a single technique provides enough information to achieve good results. To overcome this problem, it is advisable to take measurements and to integrate data from multiple NDSS, as it can significantly enhance the depth and accuracy of *in situ* analysis of food products. Different spectral sensors may capture distinct aspects of a sample's composition or properties. Understanding the complementary nature of sensors is crucial for maximizing the information gained from the combined dataset. Moreover, evaluating the spectral profiles of each sensor allows to identify areas of overlap and gaps in information. It is useful to select sensors that provide unique insights or additional context to complement the overall analysis. The best-known example of spectral sensor fusion is the combination of imaging with NIR or other spectral regions, to provide not only spectral information, but also spatial information, essential for specific applications where, for example, determining the presence of the studied analyte is not sufficient and its homogeneous distribution is a key issue.

Performing the measurements in the same sample order for each technique is advisable, as it reduces the variability introduced by changes in sample handling procedures and allows for direct comparison of results obtained from different techniques. This facilitates meaningful interpretations and conclusions about the sample properties. In general, it helps control for potential confounding factors or sources of variability, a particularly important aspect in experiments where factors like sample aging or environmental conditions may affect the measurements. Moreover, consistent sample order facilitates data analysis, as it allows for straightforward pairing of measurements obtained from different techniques, simplifying tasks like data fusion, calibration, and interpretation.

Using different NDSS also leads to face different challenges. Integrating multiple spectral sensors can be technically complex, as different sensors may have different specifications, sensitivities, and data output formats. Ensuring that data from multiple

spectral sensors are synchronized is crucial for combining information from different NDSS and obtaining a comprehensive analysis of the sample. Moreover, these sensors may have overlapping spectral bands, leading to redundancy in the information collected and making the process of managing and processing this data efficiently essential to avoid information overload. Another thing to consider is that effective sampling protocols need to be developed to ensure that each sensor receives representative samples. Lastly, acquiring and maintaining multiple spectral sensors can be costly, requiring a significant investment in terms of equipment, training, and maintenance.

In the next sections, it will be illustrated which are the methods to handle data coming from different spectroscopic techniques.

3.1 Data Fusion Techniques

Data fusion involves combining information from multiple sensors to create a unified dataset. As the name suggests, data fusion leverage information derived from different analytical techniques or types of data and combine them with the aim of maximizing obtainable information. They constitute an extension of the multivariate chemometric approach, not only to individual variables but also to the actual analysis techniques. Different fusion techniques, such as low, mid and high-level data fusion, offer various approaches to merging data. Choose a fusion method based on the specific objectives of the analysis, the nature of the data, and the compatibility of the sensors [73].

Low level is the simplest and most immediate data fusion method. It involves merging different datasets to obtain an "augmented" new one with the same number of rows as the individual datasets and a number of columns equal to the sum of variables present in each profile. If needed, different pre-treatments can be applied to different components of the dataset. Although this approach is more easily interpretable (since model parameters, such as loadings, directly relate to the measured variables), it usually does not yield the best results because the chemical information remains enclosed in the original data along with the noise, which is not reduced and contributes to the spurious variance of the dataset.

In the mid-level data fusion, the information concentrated in a few latent variables (typically determined through PCA and PLS) is extracted from each original matrix, and these undergo the fusion process. The new matrix will thus be composed of fewer columns than the sum of the original variables, while practically maintaining the same

chemical variance and drastically reducing the variance associated with noise. The use of block scaling allows for "weighing" the components in the final dataset to prevent one technique from dominating over another solely because it is represented by a greater number of LVs. The mid-level approach is generally the most suitable as it starts from data already filtered from noise, and the algorithm (whatever it may be) is applied to the entirety of the data (and therefore to the entirety of the real information present in the original datasets).

High-level data fusion is an *a posteriori* approach. It is not a true data fusion technique, but a response fusion technique. The treatment is applied to the original matrices to acquire predictions on unknown samples, and it is these that undergo the fusion process. Since this technique does not benefit from the simultaneous integration of all information but starts from data already processed and potentially affected by modelling errors, it is often used for confirmation or in cases where lower levels of fusion were not able to provide meaningful answers. It is not as widely used as the other two approaches.

Choosing the best data fusion strategy depends on the data structure, size and quality. To reduce noise or unwanted variability in the data, mid-level data fusion can be a valuable choice, as the extracted features should contain just relevant information. However, in that case, operators should know *a priori* which the most important features to extract are. In the context of quality control applications, there is a growing trend towards employing a comprehensive and unbiased approach centered on the thorough characterization (fingerprinting) of products, and various analytical techniques are utilized for this purpose. The key question at hand is to discern the overlapping and distinct information provided by each platform. This assessment is crucial in determining which techniques are truly necessary for establishing an effective quality monitoring protocol. The ultimate goal is to achieve optimal product characterization while minimizing costs. Therefore, it is imperative to aggregate both shared and unique information contributions from each data block. In this regard, a low-level data fusion framework appears to be a better approach [73].

3.2 Multiblock Methods

Multiblock analysis methods enable the simultaneous analysis of data from multiple spectral sensors, making their use advised. They can combine information from

different sources or data sets, which may have varying types, structures, or units, allowing for a more comprehensive analysis. By considering multiple data sets simultaneously, multiblock methods provide a more holistic view of complex systems, as they can uncover relationships and patterns that may not be apparent when analyzing individual blocks of data. In this way, different data blocks may contain complementary information about the same samples or processes. Furthermore, the integration of multiple data sets can help filter out noise and remove redundant or irrelevant information, leading to more accurate and reliable models [74]. One of the most used multiblock method is Multiblock-Partial Least Squares (MB-PLS) [75], that aims to find latent variables that explain the maximum covariance between the input and output blocks. It is widely used mainly because its single block version (PLS) is the most used chemometric technique to perform multivariate regressions. This is mainly due to its implementation in many instruments and statistical software. Regarding sequential methods, the most used ones are Sequential and Orthogonalized Partial Least Squares (SO-PLS) [76] and Response-Oriented Sequential Alternation (ROSA) [71], a sequential algorithm based on PLS that are invariant with respect to block-scaling and, in case of ROSA, of block ordering. All these methods can be extended to classification problems through the combination with Linear Discriminant Analysis (LDA). Another well-known method is Common Dimension (ComDim) [77] that aim to identify common components shared across different blocks of data. The problem of choosing the proper multiblock method can be solved considering factors like the number of blocks, their sizes, if they are time-ordered (in case of a process monitoring issue) and whether they have different scales or units. For instance, ROSA deals perfectly with high number of large data blocks thanks to its high computational efficiency.

4 Conclusions

This comprehensive protocol and guideline for combining NDSS in the *in situ* analysis of food products provides a structured framework to enhance the accuracy and reliability of analytical results. The main NDSS have been inspected and compared, describing which the most common applications for each of them are. By systematically considering factors such as sensor selection, acquisition modes, preprocessing techniques, and good practices in using chemometric methods, this protocol ensures a robust approach to spectral data fusion. In conclusion, NDSS are a powerful tool for food industry, as they allow for non-destructive, fast, green and in situ analyses. Their

combined use, if properly performed, allows to greatly enhance results obtained for all the possible tasks and objectives.

5 References

1. Li, X., Zhang, L., Zhang, Y., Wang, D., Wang, X., Yu, L., Zhang, W., & Li, P. (2020). Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils. *Trends in Food Science & Technology*, *101*, 172-181.
2. Yakubu, H. G., Kovacs, Z., Toth, T., & Bazar, G. (2022). The recent advances of near-infrared spectroscopy in dairy production—A review. *Critical Reviews in Food Science and Nutrition*, *62*(3), 810-831.
3. Vitalis, F., Tjandra Nugraha, D., Aouadi, B., Aguinaga Bósquez, J. P., Bodor, Z., Zaukuu, J. L. Z., Kocsic, T., Zsom-Muha, V., Gillay, Z., & Kovacs, Z. (2021). Detection of monilia contamination in plum and plum juice with NIR spectroscopy and electronic tongue. *Chemosensors*, *9*(12), 355.
4. Xu, J., Huang, F., Li, Y., Chen, Z., & Wang, Y. (2015). Rapid detection of total nitrogen content in soy sauce using NIR spectroscopy. *Czech Journal of Food Sciences*, *33*(6), 518-522.
5. Valinger, D., Longin, L., Grbeš, F., Benković, M., Jurina, T., Kljusurić, J. G., & Tušek, A. J. (2021). Detection of honey adulteration–The potential of UV-VIS and NIR spectroscopy coupled with multivariate analysis. *Lwt*, *145*, 111316.
6. Li, X., He, Y., Wu, C., & Sun, D. W. (2007). Nondestructive measurement and fingerprint analysis of soluble solid content of tea soft drink based on Vis/NIR spectroscopy. *Journal of Food Engineering*, *82*(3), 316-323.
7. Arslan, M., Tahir, H. E., Zareef, M., Shi, J., Rakha, A., Bilal, M., ... & Xiaobo, Z. (2021). Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques. *Trends in Food Science & Technology*, *107*, 80-113.
8. Zareef, M., Arslan, M., Hassan, M. M., Ahmad, W., Ali, S., Li, H., Wu, X., Hashim, M. M., & Chen, Q. (2021). Recent advances in assessing qualitative and quantitative aspects of cereals using nondestructive techniques: A review. *Trends in Food Science & Technology*, *116*, 815-828.
9. Zhang, S., Liu, S., Shen, L., Chen, S., He, L., & Liu, A. (2022). Application of near-infrared spectroscopy for the nondestructive analysis of wheat flour: A review. *Current Research in Food Science*.
10. Wang, Y., Peng, Y., Zhuang, Q., & Zhao, X. (2020). Feasibility analysis of NIR for detecting sweet corn seeds vigor. *Journal of Cereal Science*, *93*, 102977.
11. Genis, H. E., Durna, S., & Boyaci, I. H. (2021). Determination of green pea and spinach adulteration in pistachio nuts using NIR spectroscopy. *Lwt*, *136*, 110008.
12. Zhu, M. Z., Wen, B., Wu, H., Li, J., Lin, H., Li, Q., Li, Y., Huang, J., & Liu, Z. (2019). The quality control of tea by near-infrared reflectance (NIR) spectroscopy and chemometrics. *Journal of Spectroscopy*, *2019*.
13. Alessandrini, L., Romani, S., Pinnavaia, G., & Dalla Rosa, M. (2008). Near infrared spectroscopy: An analytical tool to predict coffee roasting degree. *Analytica Chimica Acta*, *625*(1), 95-102.

14. Beć, K. B., Grabska, J., & Huck, C. W. (2021). NIR spectroscopy of natural medicines supported by novel instrumentation and methods for data analysis and interpretation. *Journal of Pharmaceutical and Biomedical Analysis*, *193*, 113686.
15. Shah, S. S. A., Zeb, A., Qureshi, W. S., Arslan, M., Malik, A. U., Alasmary, W., & Alanazi, E. (2020). Towards fruit maturity estimation using NIR spectroscopy. *Infrared Physics & Technology*, *111*, 103479.
16. Pu, Y., Pérez-Marín, D., O'Shea, N., & Garrido-Varo, A. (2021). Recent advances in portable and handheld NIR spectrometers and applications in milk, cheese and dairy powders. *Foods*, *10*(10), 2377.
17. Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., & Elliott, C. (2020). Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. *Infrared Physics & Technology*, *104*, 103127.
18. Tanzilli, D., D'Alessandro, A., Tamelli, S., Durante, C., Cocchi, M., & Strani, L. (2023). A Feasibility Study towards the On-Line Quality Assessment of Pesto Sauce Production by NIR and Chemometrics. *Foods*, *12*(8), 1679.
19. da Paixao Teixeira, J. L., dos Santos Carames, E. T., Baptista, D. P., Gigante, M. L., & Pallone, J. A. L. (2021). Rapid adulteration detection of yogurt and cheese made from goat milk by vibrational spectroscopy and chemometric tools. *Journal of Food Composition and Analysis*, *96*, 103712.
20. Núñez-Sánchez, N., Martínez-Marín, A. L., Polvillo, O., Fernández-Cabanás, V. M., Carrizosa, J., Urrutia, B., & Serradilla, J. M. (2016). Near Infrared Spectroscopy (NIRS) for the determination of the milk fat fatty acid profile of goats. *Food Chemistry*, *190*, 244-252.
21. Wang, W., Peng, Y., Sun, H., Zheng, X., & Wei, W. (2018). Spectral detection techniques for non-destructively monitoring the quality, safety, and classification of fresh red meat. *Food Analytical Methods*, *11*, 2707-2730.
22. Mureșan, V., Danthine, S., Mureșan, A. E., Racolța, E., Blecker, C., Muste, S., Socaciu, C., & Baeten, V. (2016). In situ analysis of lipid oxidation in oilseed-based food products using near-infrared spectroscopy and chemometrics: The sunflower kernel paste (tahini) example. *Talanta*, *155*, 336-346.
23. Contal, L., León, V., & Downey, G. (2002). Detection and quantification of apple adulteration in strawberry and raspberry purees using visible and near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, *10*(4), 289-299.
24. Müller-Maatsch, J., Bertani, F. R., Mencattini, A., Gerardino, A., Martinelli, E., Weesepoel, Y., & van Ruth, S. (2021). The spectral treasure house of miniaturized instruments for food safety, quality and authenticity applications: A perspective. *Trends in Food Science & Technology*, *110*, 841-848.
25. Hong, T., Yin, J. Y., Nie, S. P., & Xie, M. Y. (2021). Applications of infrared spectroscopy in polysaccharide structural analysis: Progress, challenge and perspective. *Food chemistry: X*, *12*, 100168.
26. Li, Q., Chen, J., Huyan, Z., Kou, Y., Xu, L., Yu, X., & Gao, J. M. (2019). Application of Fourier transform infrared spectroscopy for the quality and safety analysis of fats and oils: A review. *Critical reviews in food science and nutrition*, *59*(22), 3597-3611.
27. López-Lorente, Á. I., & Mizaikoff, B. (2016). Mid-infrared spectroscopy for protein analysis: potential and challenges. *Analytical and bioanalytical chemistry*, *408*, 2875-2889.

28. Valand, R., Tanna, S., Lawson, G., & Bengtström, L. (2020). A review of Fourier Transform Infrared (FTIR) spectroscopy used in food adulteration and authenticity investigations. *Food Additives & Contaminants: Part A*, *37*(1), 19-38.
29. Gorla, G., Mestres, M., Boque, R., Riu, J., Spanu, D., & Giussani, B. (2020). ATR-MIR spectroscopy to predict commercial milk major components: A comparison between a handheld and a benchtop instrument. *Chemometrics and Intelligent Laboratory Systems*, *200*, 103995.
30. Aykas, D. P., Rodrigues Borba, K., & Rodriguez-Saona, L. E. (2020). Non-destructive quality assessment of tomato paste by using portable mid-infrared spectroscopy and multivariate analysis. *Foods*, *9*(9), 1300.
31. Aykas, D. P., & Menevseoglu, A. (2021). A rapid method to detect green pea and peanut adulteration in pistachio by using portable FT-MIR and FT-NIR spectroscopy combined with chemometrics. *Food Control*, *121*, 107670.
32. Pan, L., Zhang, P., Daengngam, C., Peng, S., & Chongcheawchamnan, M. (2022). A review of artificial intelligence methods combined with Raman spectroscopy to identify the composition of substances. *Journal of Raman Spectroscopy*, *53*(1), 6-19.
33. Xu, Y., Zhong, P., Jiang, A., Shen, X., Li, X., Xu, Z., Shen, Y., Sun, Y., & Lei, H. (2020). Raman spectroscopy coupled with chemometrics for food authentication: A review. *TrAC Trends in Analytical Chemistry*, *131*, 116017.
34. Arroyo-Cerezo, A., Jimenez-Carvelo, A. M., González-Casado, A., Koidis, A., & Cuadros-Rodríguez, L. (2021). Deep (offset) non-invasive Raman spectroscopy for the evaluation of food and beverages–A review. *LWT*, *149*, 111822.
35. Dib, O. H., Assaf, A., Pean, A., Durand, M. J., Jouanneau, S., Ramanathan, R., & Thouand, G. (2022). Raman Spectroscopy Application in Food Waste Analysis: A Step towards a Portable Food Quality-Warning System. *Sustainability*, *15*(1), 188.
36. Sourdaine, M., Guenther, D., Dowgiallo, A. M., Harvey, C., Mattley, Y., Guckian, A., & Lischtschenko, O. (2015). Protecting the food supply chain: Utilizing SERS and portable Raman spectroscopy. *tm-Technisches Messen*, *82*(12), 625-632.
37. Sakare, P., Giri, S. K., Mohapatra, D., & Tripathi, M. K. (2022). UV-Vis spectroscopic investigation on color change kinetics of lac dye as influenced by some food spoilage metabolites: validation for milk quality monitoring. *Pigment & Resin Technology*.
38. Kourani, K., Kapoor, N., Badiye, A., & Shukla, R. K. (2020). Detection of synthetic food color “Metanil yellow” in sweets: A systematic approach. *JPC–Journal of Planar Chromatography–Modern TLC*, *33*(4), 413-418.
39. Zhang, Z. H., Zeng, X. A., Brennan, C. S., Ma, H., & Aadil, R. M. (2019). Preparation and characterisation of novelty food preservatives by Maillard reaction between ε-polylysine and reducing sugars. *International Journal of Food Science & Technology*, *54*(5), 1824-1835.
40. Zang, S., Tian, S., Jiang, J., Han, D., Yu, X., Wang, K., Li, D., Lu, D., Yu, A., & Zhang, Z. (2017). Determination of antioxidant capacity of diverse fruits by electron spin resonance (ESR) and UV–vis spectrometries. *Food Chemistry*, *221*, 1221-1225.

41. Saha, S., Singh, J., Paul, A., Sarkar, R., Khan, Z., & Banerjee, K. (2020). Anthocyanin profiling using UV-Vis spectroscopy and liquid chromatography mass spectrometry. *Journal of AOAC International*, *103*(1), 23-39.
42. Semitsoglou-Tsiapou, S., Meador, T. B., Peng, B., & Aluwihare, L. (2022). Photochemical (UV–vis/H2O2) degradation of carotenoids: Kinetics and molecular end products. *Chemosphere*, *286*, 131697.
43. Feseha, A., & Gholap, A. V. (2020). Determination of chlorophylls and carotenes content in some vegetables of ethiopiaby using uv-vis spectroscopy. *Annals: Food Science & Technology*, *21*(4).
44. Woo, K. K., Ngou, F. H., Ngo, L. S., Soong, W. K., & Tang, P. Y. (2011). Stability of betalain pigment from red dragon fruit (Hylocereus polyrhizus). *American Journal of Food Technology*, *6*(2), 140-148.
45. Yu, X., Zhao, M., Hu, J., Zeng, S., & Bai, X. (2012). Correspondence analysis of antioxidant activity and UV–Vis absorbance of Maillard reaction products as related to reactants. *LWT-Food Science and Technology*, *46*(1), 1-9.
46. Tian, W., Chen, G., Gui, Y., Zhang, G., & Li, Y. (2021). Rapid quantification of total phenolics and ferulic acid in whole wheat using UV–Vis spectrophotometry. *Food Control*, *123*, 107691.
47. Aberásturi, F., Jiménez, A. I., Jiménez, F., & Arias, J. J. (2001). UV-visible first-derivative spectrophotometry applied to an analysis of a vitamin mixture. *Journal of Chemical Education*, *78*(6), 793.
48. Gonçalves, T. R., Rosa, L. N., Gonçalves, R. P., Torquato, A. S., Março, P. H., Marques Gomes, S. T., Matsushita, M., & Valderrama, P. (2018). Monitoring the oxidative stability of monovarietal extra virgin olive oils by UV–Vis spectroscopy and MCR–ALS. *Food Analytical Methods*, *11*, 1936-1943.
49. Bartoszek, M., & Polak, J. (2016). A comparison of antioxidative capacities of fruit juices, drinks and nectars, as determined by EPR and UV–vis spectroscopies. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *153*, 546-549.
50. Valinger, D., Longin, L., Grbeš, F., Benković, M., Jurina, T., Kljusurić, J. G., & Tušek, A. J. (2021). Detection of honey adulteration–The potential of UV-VIS and NIR spectroscopy coupled with multivariate analysis. *Lwt*, *145*, 111316.
51. Ahmad, M. H., Sahar, A., & Hitzmann, B. (2017). Fluorescence spectroscopy for the monitoring of food processes. *Measurement, modeling and automation in advanced food processing*, 121-151.
52. Christensen, J., Becker, E. M., & Frederiksen, C. S. (2005). Fluorescence spectroscopy and PARAFAC in the analysis of yogurt. *Chemometrics and Intelligent Laboratory Systems*, *75*(2), 201-208.
53. Cao, J., Li, C., Liu, R., Liu, X. R., Fan, Y., & Deng, Z. Y. (2017). Combined application of fluorescence spectroscopy and chemometrics analysis in oxidative deterioration of edible oils. *Food Analytical Methods*, *10*, 649-658.
54. Ji, R., Ma, S., Yao, H., Han, Y., Yang, X., Chen, R., ... & Bian, H. (2020). Multiple kinds of pesticide residue detection using fluorescence spectroscopy combined with partial least-squares models. *Applied Optics*, *59*(6), 1524-1528.
55. Székely, D., Csighy, A., Stéger-Máté, M., & Monspart-Sényi, J. (2014). Analysis of heavy metal accumulation food with X-Ray fluorescence spectometry. *Analecta Technica Szegedinensia*, *8*(1), 1-5.

56. Zhang, G., & Ma, Y. (2013). Mechanistic and conformational studies on the interaction of food dye amaranth with human serum albumin by multispectroscopic methods. *Food chemistry*, *136*(2), 442-449.
57. Gu, H., Lv, R., Huang, X., Chen, Q., & Dong, Y. (2022). Rapid quantitative assessment of lipid oxidation in a rapeseed oil-in-water (o/w) emulsion by three-dimensional fluorescence spectroscopy. *Journal of Food Composition and Analysis*, *114*, 104762.
58. Ma, J., Sun, D. W., Pu, H., Cheng, J. H., & Wei, Q. (2019). Advanced techniques for hyperspectral imaging in the food industry: Principles and recent applications. *Annual review of food science and technology*, *10*, 197-220.
59. Feng, C. H., Makino, Y., Oshita, S., & Martín, J. F. G. (2018). Hyperspectral imaging and multispectral imaging as the novel techniques for detecting defects in raw and processed meat products: Current state-of-the-art research advances. *Food Control*, *84*, 165-176.
60. Wu, D., & Sun, D. W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part II: Applications. *Innovative Food Science & Emerging Technologies*, *19*, 15-28.
61. Zhu, S., Feng, L., Zhang, C., Bao, Y., & He, Y. (2019). Identifying freshness of spinach leaves stored at different temperatures using hyperspectral imaging. *Foods*, *8*(9), 356.
62. Crichton, S. O., Kirchner, S. M., Porley, V., Retz, S., von Gersdorff, G., Hensel, O., Weygandt, M., & Sturm, B. (2017). Classification of organic beef freshness using VNIR hyperspectral imaging. *Meat science*, *129*, 20-27.
63. Yu, X., Tang, L., Wu, X., & Lu, H. (2018). Nondestructive freshness discriminating of shrimp using visible/near-infrared hyperspectral imaging technique and deep learning algorithm. *Food analytical methods*, *11*, 768-780.
64. Zhuang, Q., Peng, Y., Yang, D., Wang, Y., Zhao, R., Chao, K., & Guo, Q. (2022). Detection of frozen pork freshness by fluorescence hyperspectral image. *Journal of Food Engineering*, *316*, 110840.
65. Bonah, E., Huang, X., Aheto, J. H., & Osae, R. (2019). Application of hyperspectral imaging as a nondestructive technique for foodborne pathogen detection and characterization. *Foodborne pathogens and disease*, *16*(10), 712-722.
66. Jiang, J., Qiao, X., & He, R. (2016). Use of Near-Infrared hyperspectral images to identify moldy peanuts. *Journal of Food Engineering*, *169*, 284-290.
67. Saeidan, A., Khojastehpour, M., Golzarian, M. R., Mooenfard, M., & Khan, H. A. (2021). Detection of foreign materials in cocoa beans by hyperspectral imaging technology. *Food Control*, *129*, 108242.
68. Benelli, A., Cevoli, C., Fabbri, A., & Ragni, L. (2022). Ripeness evaluation of kiwifruit by hyperspectral imaging. *Biosystems Engineering*, *223*, 42-52.
69. Garillos-Manliguez, C. A., & Chiang, J. Y. (2021). Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation. *Sensors*, *21*(4), 1288.
70. Cocchi, M. (2006). Chemometrics for food quality control and authentication. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, 1-29.

71. Liland, K. H., Næs, T., & Indahl, U. G. (2016). ROSA—A fast extension of partial least squares regression for multiblock data analysis. *Journal of Chemometrics*, *30*(11), 651-662.
72. Roger, J. M., Palagos, B., Bertrand, D., & Fernandez-Ahumada, E. (2011). CovSel: variable selection for highly multivariate and multi-response calibration: application to IR spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, *106*(2), 216-223.
73. Cocchi, M. (2019). Introduction: Ways and means to deal with data from multiple sources. In *Data handling in science and technology* (Vol. 31, pp. 1-26). Elsevier.
74. Mishra, P., Roger, J. M., Jouan-Rimbaud-Bouveresse, D., Biancolillo, A., Marini, F., Nordon, A., & Rutledge, D. N. (2021). Recent trends in multi-block data analysis in chemometrics for multi-source data integration. *TrAC Trends in Analytical Chemistry*, *137*, 116206.
75. MacGregor, J. F., Jaeckle, C., Kiparissides, C., & Koutoudi, M. (1994). Process monitoring and diagnosis by multiblock PLS methods. *AIChE Journal*, *40*(5), 826-838.
76. Næs, T., Måge, I., & Segtnan, V. H. (2011). Incorporating interactions in multi-block sequential and orthogonalised partial least squares regression. *Journal of chemometrics*, *25*(11), 601-609.
77. Nielsen, J. P., Bertrand, D., Micklander, E., Courcoux, P., & Munck, L. (2001). Study of NIR spectra, particle size distributions and chemical parameters of wheat flours: a multi-way approach. *Journal of Near Infrared Spectroscopy*, *9*(4), 275-285.