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

Protocol and guidelines for recommendations on NDSS placement, sampling, and specifications for *in situ* analysis of food products

Keywords: Non-destructive spectral sensors, NIR, NIRS, Analytical Technology, Near Infrared Spectroscopy, in situ analysis.

Table of contents

1. INTRODUCTION

There is an increasing need for the food industry to provide information on their products in order to satisfy quality standards and to protect their products from food fraud. Recent developments in technology, and advances in big data analytics, provide the opportunity for step-changes that can transform the role of food integrity assurance from one of just strictly conformance to one that addresses a wide range of business-critical concerns, including quality, safety, and authenticity solutions. Non-destructive Spectroscopic Sensors (NDSS), such as NIR Spectroscopy, Fluorescence, Raman, or Hyperspectral imaging, enable rapid, non-destructive, and environmentally safe assessment of multiple parameters in a variety of food products. Most applications of these technologies in the food industry are made at laboratory scale. Industry requires them to be deployed *in situ* along different points in the entire food chain (on-farm, in-field or on-line/in-line for full process control). These requirements introduce constraints on sensor design and calibration development which do not normally apply to laboratory-based instruments. Long-term stability of instruments, robustness of the calibrations, sensor integration in production environments, transferability of data and the building of real-time decision-making systems are critical issues to be considered.

The first step when using NDSS is the development of a correct analysis methodology, which should allow fast, non-destructive measurement of materials with little or no sample preparation, for a wide range of food products and processes. This is complex and time-consuming, therefore, the existence of established protocols that help to simplify this process is essential. This guideline aims to serve as a basis for establishing analytical methodologies for specific applications in the food industry.

Although in this protocol more emphasis is going to be placed on NIRS technology, the chemometric principles described within this guideline may also be applicable to other analytical techniques.

2. NDSS PLACEMENT

NDSS instruments can be placed at different distances from the process, so a distinction is made between laboratory, in-line, in-line or in-line measurements (Figure 1), but in-situ measurements with handheld devices could also be considered.

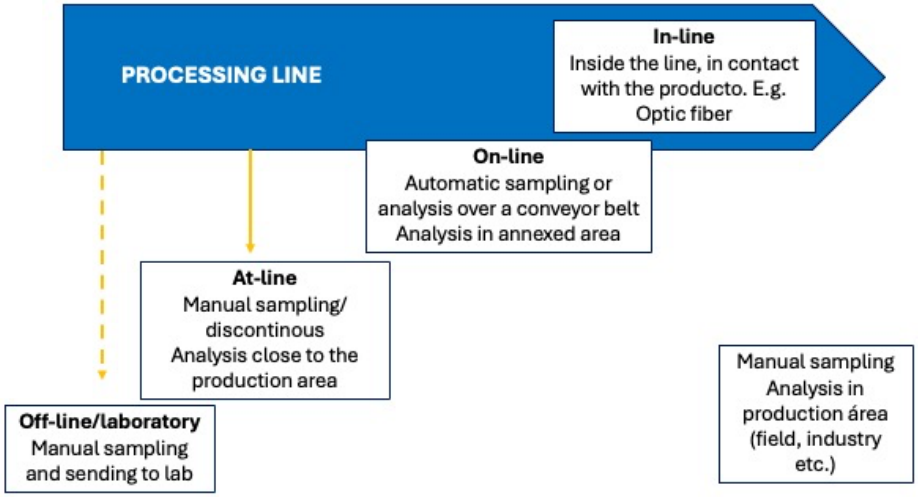

Figure 1. Location of NIRS instrumentation according to distance to the process.

The in-line and on-line instruments are located on the production line, normally working continuously, connected to the SCADA industrial control and automation system of each industry, which allows full control of the production process in real time. The main difference between them is that the in-line ones are in direct contact with the product (an immersion probe, for example). These devices can be classified into two types, on the one hand, those that have the optics in the measurement head itself, so they must be robust and completely sealed to avoid any source of external interference, such as vibrations, dust, etc. On the other hand, there are the instruments located away from the process, isolated and well protected, a setup which also allows the connection of different measurement accessories (measurement heads for solids, liquid probes, etc.) which can be placed at different points in the plant by means of fibre optic cables of up to 100 m length.

The at-line and off-line devices are also similar, depending on the remoteness of the analysis site. Normally they are instruments located in laboratories near to (at-line) or far (off-line) from the production line, from where the sample must be taken, then analysed manually and later

usually subjected to a traditional wet analysis. They are generally high-performance devices but need more care in their use.

Finally, on-site hand-held instruments, which are devices that, thanks to their low weight and high manoeuvrability, can be used to analyse the product at any point in the production chain, from the field, through supplier control to the final product. Specifically, NIRS portable instruments, which will be covered in this protocol, are perhaps the ones that have evolved the most in recent years. The first handheld NIR devices available in the 1990s had some limitations for generalisation to other agri-food products, such as a limited wavelength range, low resolution (10-12 nm), poor reproducibility of spectra, limited size of the optical window, not adapted to the analysis of liquid samples and dependent on an external computer. However, in the first decade of the 21st century we have witnessed enormous advances in portable instrumentation and due to the rapid progress of technologies such as Linear Variable Filters (LVF) or Microelectromechanical Systems (MEMS) and their integration with Microoptics (MOEMS).

3. MEASUREMENT MODES AND SAMPLING TECHNIQUES

There are many different NDSS that also have different types of accessories. Which to use in each application depends on a multitude of factors such as:

- The physical chemical state of the sample, meaning, if it is solid, liquid or paste
- The type of product and the presentation form, ground, grain, fresh products etc.
- The analysis mode, that at the same time, depends on the type of product and its physical-chemical state. These modes can be transmission, transflectance or diffuse reflection.

For each analysis mode, there are different factors that we must consider and that we will mention below regarding the use of portable handheld NIRS devices.

3.1. Transmission.

If the product we are going to analyse is a clear liquid, perhaps the most commonly used method is transmission, although it can also be used for turbid liquids. In this case, the light coming from the source will pass through the sample and we will measure the light that reaches the detector, located behind the sample.

First, we must take into account the path length. It greatly influences the quality of the spectral information depending on the type of liquid that we are analysing and its interaction with light. The longer the path length, the more molecules there are in the path of the beam of radiation,

therefore the absorbance goes up. But if the pathlength is too big, the spectra may not be adequate. It is therefore necessary to determine, by testing, the correct pathlength for the application that we are developing. For example, a pathlength of 2 mm works well for the analysis of olive oils, but for milk it may be too high, being recommended 1mm or less.

The material of the cuvettes, when using this type of accessory, is also an important factor to consider. If we prioritize spectral quality, perhaps the best option is quartz cuvettes, which are more stable than glass or plastic, but they are expensive, delicate, and cleaning them for later reuse is a tedious and time-consuming process. Therefore, the use of disposable cuvettes may be advisable due to their easy handling, although it is also important to maintain consistent glass/plastic quality and path length over time to keep calibrations valid.

Depending on the type of liquid, the homogenization of the sample is also a key aspect, avoiding in any case the analysis of a sample with bubbles, as they interfere with the incident light and distort the spectrum.

For instruments working in this analysis mode, it is recommended to carry out the calibration using an empty cuvette.

3.2. Transflectance.

This analysis mode could be used for clear, opaque, turbid, and viscous liquids but also for semisolid with high water content, emulsions, and slurries.

If we have an appropriate pathlength, the light will pass through the sample, then will be reflected back by a reflective material, and will pass through the sample again until it reaches the detector, which is equivalent to a double transmission. The most common reflective materials are gold coated surfaces, spectralon, and stainless steel. Although gold coating is more reflective, it may degrade more with use. It should be noted that if the path length is very low, meaning lower than 0.5mm, the measurement accessories will be more difficult to manufacture consistently.

It is important to always use the same sample quantity in this analysis mode, so that the repeatability and reproducibility of the spectra is correct.

Regarding the calibration of the instrument, the dark reference is usually measured with the lamp off or pointing at a fixed point in the room without the reflective material, while the white reference is usually measured with the reflective material and the lamp on. Previous analysis tests

must be carried out to avoid the appearance of negative absorbances in the spectrum, which, although they can be corrected by means of spectral pre-treatments, are not advisable.

3.3.Diffuse reflectance.

For solid products, the most used method is diffuse reflectance, although it can also be applied for semi-solids, emulsions, and slurries. When the light reaches the sample, it is reflected in different ways and directions depending on the characteristics of the analysed product.

The heterogeneity of the sample is of paramount importance when establishing a correct analysis methodology. When the sample is finely ground and homogeneous, repeatability should be excellent. The higher the heterogeneity, the more we have to take into account the analysis window of the device to obtain representative spectral information of the sample (Figure 2). To achieve this, we can take several measurements of each sample or analyse dynamically, i.e. moving the sensor along the sample. The method chosen depends on the morphology of the product and its physical and/or chemical heterogeneity. For example, if it is a fresh product with an irregular surface, e.g., bell peppers, it will be convenient to take several measurements at different points of the sample (Figure 2, a). Alternatively, if we have a heterogeneous sample, such as a meat meal with the presence of bone pieces, we could try to collect spectra in dynamic mode (Figure 2, b), thus reducing the influence of the window size due to the averaging effect.

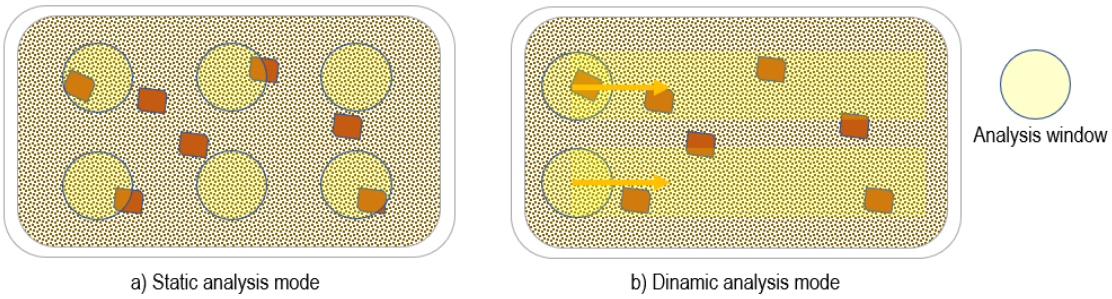

Figure 2. Static and dynamic analysis modes in a heterogeneous sample.

To obtain consistent spectral information, it is important to keep the handling of portable devices as constant as possible, that is, distance to the sample, angle with the surface and position where the sample is analysed. Especially for applications in the field, a reproducible handling during measurement is one of the most important aspects for best possible performance. Normally, in portable instruments, it is usual to analyse in direct contact with the product. Again, if the product is physically irregular, we must analyse on an area that facilitates the analysis in contact, so that the light escapes as little as possible. If the sample is irregular, but due to its granularity, such as

small fruits (olives, for example), it is advisable to use a glass container such as a petri dish and group them together, trying to cover the entire analysis surface. In the latter case, it is advisable to place a black plastic behind the petri dish, so that the light that filters through the gaps is absorbed by it.

4. SPECIFICATIONS

4.1. Instrument selection

Depending on the application, an analysis mode must be chosen, then, a device that allows it must be selected, and then, the following aspects must be considered:

- Spectral range: Portable instruments tend to have a narrower wavelength range than benchtop instruments, so, depending on the parameters to be predicted, we should select the instrument that, based on the literature, contains the NIR range where the components of greatest interest for our application are located, if possible.
- Analysis window: We have already mentioned above that the more heterogeneous a sample is, the more sample area must be analysed to obtain representative information from it. Therefore, we must also select a device with a suitable analysis window for the application in question. For homogeneous samples, small analysis windows may be sufficient.
- Operability: This is very important for routine application. If the analysis procedure is time-consuming, we can obtain good results in the laboratory during the feasibility study, but probably when using the handheld device on site, on the production line, the throughput will be much lower. Therefore, it is important to select more ergonomic devices, if they are to be used more intensively.
- Calibration transferability: spectrometers, even of the same model manufactured by the same company, may differ slightly. For this reason, it is important to consider whether databases or calibration models can be easily transferred from one instrument to another. In the specific case of portable instruments, this is a key issue, as several devices will operate in the same industry with chemometric models developed from a database collected with a master instrument. These databases/calibration models must therefore be easily transferable from one

device to another in order to achieve the best possible performance. This aspect has a major influence on the decision to select a NIR technology or a supplier.

4.2. Device configuration

Once the instrument has been decided, configuration tests must be carried out to try to obtain spectral information with the highest possible signal to noise ratio, while maintaining operability and considering the final use that the sensor will have. Some of the parameters that are configured in portable sensors are:

Integration time: the higher the integration time, the longer the detector is exposed to light, and therefore we will obtain more signal intensity. Care must be taken as a high integration time can cause the detector to saturate. The longer the integration time, the longer it will take to do a scan.

Number of scans: the greater the number of scans, the more spectral information we will obtain from the product, but the longer it will take to analyse a sample. We must find the middle ground so that the information is representative while keeping the analysis time consistent with the future use of the instrument. Typically, 2 or 3 replicates are recommended, and, if possible, consideration should also be given to performing the analysis in dynamic mode, as mentioned in section 3.3, to collect more representative spectra.

Spectral resolution: This is normally not critical for the NIR application itself. Very high resolutions can cause noise problems. If the instrument allows you to select the resolution, it will be convenient to select the highest resolution possible, while maintaining an appropriate signal to noise ratio.

Reference measurement: any type of spectral sensor is sensitive to temperature and environmental changes, whatever its spectral range. Calibration of the instrument by background measurements helps to correct for such effects if performed frequently enough. In portable instruments that will be working in the field, where conditions will be much more variable than in a laboratory, the background measurement should be performed much more frequently, e.g. before each sample analysis or every 10 minutes.

4.3. Spectral repeatability.

Once the device configuration is established, the feasibility study can begin. To ensure the consistency of the spectral information, it is advisable to perform the calculation of the root mean square (RMS) statistic. The RMS statistic is defined as the averaged root mean square of the differences between different subsamples scanned at n wavelengths. This statistic indicates similarities or differences in absorbance values between several spectra of the same sample, thus being a measure of the difference between spectra that allows spectral comparison.

It is advisable that, from the first 10 samples analysed in a new project, between 5 and 10 spectra per sample are collected for the calculation of the RMS. Then, the RMS for an individual subsample (*j*) of the sample (*k*), and the MEAN and STD values for a given k sample must be calculated according to the following formulas:

$$RMS_{j,k} = \sqrt{\frac{\sum_{i=1}^{n} D_{ij}^2}{n}};\ D_{ij} = y_{ij} - \overline{y}_i$$

$$MEAN_k = \sqrt{\frac{\sum_{j=1}^{N} (RMS_j)^2}{N}}$$

$$STD_k = \sqrt{\frac{\sum_{j=1}^{N} (RMS_j)^2}{(N-1)}}$$

where y_{ij} is log (1/R) at wavelength i for subsample j, and $\overline{y}_i$ is log (1/R) at wavelength i for the average spectrum of N subsamples of a sample k; n is the number of data points collected by the instrument. It is advisable to multiply by 10^6 each RMS value obtained to facilitate management and processing. An STD_{limit} can then be calculated for comparing the RMS values of subsamples, following this formula:

$$STD_{limit} = 1.036 \sqrt{\frac{\sum_{k=1}^{k=m} STD_k^2}{m}} = 1.036 \sqrt{STD^2}$$

where STD is the standard deviation per sample and m is the number of samples.

The value 1.036 corresponds to a probability level of 85%. The STD_{limit} values can then be used to establish the RMS_{cutoff} for each instrument and sample presentation. Therefore, when we collect spectra of a new sample and calculate its RMS, if it exceeds this limit, we will have to repeat the spectral acquisition and recalculate the RMS until it reaches values below the limit. Then, the

mean spectrum of each sample can be calculated. It should be noted that the RMS_{limit} statistic offers an indicative value, and that we should not take it as a strict limit, so that those spectra that exceed the cutoff must be studied separately to see if it could have been due to errors in the spectra acquisition, or other factors related to the heterogeneity of the product and whether they should be eliminated or not.

No recommendations can be made about the advisable values of this statistic, since they depend on the instrument optics, the morphology of the product, its heterogeneity, the operator, and the analysis mode, so that they must be calculated for each application.