FOODINTEGRITY
Ensuring the Integrity of the European food chain

613688: Collaborative Project

Seventh Framework Programme
KBBE.2013.2.4-01: Assuring quality and authenticity in the food chain

Deliverable: 19.7

Training Guidelines and wider exploitations of NIRS
Author(s): Dolores Pérez-Marín¹, Ana Garrido -Varo¹, Cecilia Riccioli¹, Tom Fearn²

Date of preparation: December 2018
Period covered: Months 27-55
Status: version 1

<table>
<thead>
<tr>
<th>Dissemination level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>PP</td>
<td>Restricted to other participants</td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only members of the consortium</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

The project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 613688.
1 Department of Animal Production, Non-destructive Sensor Unit, Faculty of Agricultural and Forestry Engineering, University of Córdoba, Campus Rabanales, N-IV, km 396, Córdoba 14014, Spain

2 Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK
TABLE OF CONTENTS

1. Introduction ..................................................................................................................................... 3
2. Training guidelines .......................................................................................................................... 5
   2.1. Selection of the instrument and the analysis mode ................................................................. 6
   2.1. Spectral data acquisition ......................................................................................................... 7
   2.2. Reference data ........................................................................................................................ 9
   2.3. Calibration/training population ............................................................................................... 9
   2.4. Data preprocessing ................................................................................................................ 10
   2.5. Calibration development using multivariate analysis ........................................................... 11
   2.6. Statistics used in evaluating and optimizing calibration models ........................................... 12
   2.7. Outlier detection during calibration development ................................................................. 13
   2.8. Validation of the models ....................................................................................................... 14
   2.9. Routine analysis and recalibration ....................................................................................... 15
   2.10. Sources of error in NIRS analysis ....................................................................................... 16
3. Information on current and future courses and workshops ............................................................ 18
4. Wider applications of NIR for other uses in food authentication .................................................. 20
5. NIRS implementation in companies ............................................................................................. 21
6. References ..................................................................................................................................... 21
Annex I. Graphical guidelines for the use of the App for a fast and instantaneous prediction of quality parameters of Iberian pig carcasses for monitoring the production process ........................................... 23
Annex II. Graphical guidelines for the use of the App for consumers ............................................. 34
Deliverable 19.7 Training Guidelines and wider exploitations of NIRS: report of the training guidelines and wider applications of NIR for other uses in food authentication

1. Introduction

Food and feed integrity, including quality, safety and traceability, is a primary concern for European and international policies, since are keys requirements to avoid frauds and ensure consumer health and successful national and international trade.

Currently, the increase in complexity of the food chain requires its control as a whole, covering all the steps that make up it, i.e. from raw material production, to processing and distribution. In fact, the food chain is vulnerable to a range of food hazards that may arise at any stage of the chain. Thus, an integrated approach for control, inspection and monitoring of the food chain and efficient analytical methods, able to be implemented along the whole food chain, are essential tools in preventing or avoiding new food crises and frauds. In addition, the analytical needs for the food industry are linked not only to compliance with regulations, but also to the need to control their processes for increasing their competitiveness. Inaccurate or uninformative quality and safety assessment methodologies are detrimental to producers, processors and, ultimately, consumers.

Currently, food quality and safety are based on discontinuous at-line measurements and traditional analytical methods, while most of the traceability systems are based on documentary tracking and effective labeling. As can be deducted, this is not enough to answer the described problems of the food industry, nowadays more consumer-focused, which needs faster response times to deal with food frauds and crises.

Therefore, there is a clear the need for rapid and efficient analytical procedures to monitor and verify food integrity, and to trace food efficiently from the origin to the consumption, including the labeling of the final product. Moreover, aspects related to prove product authenticity, to control adulteration or to confirm the provenance of food, in particular for specific quality labels, are important issues for the food sector, increasing the added value of the products. All of these include demands for more credible labeling systems, requiring extensive sampling and non-destructive analysis of individual items, which cannot be met by traditional analytical methods.

Near Infrared Spectroscopy (NIRS) has been demonstrated to be ideally suited for this purpose, enabling rapid, non-destructive, accurate, and economical analysis of large volume of samples and environmentally safe assessment (not use chemical reagent and not produce waste chemical residues) of multiple parameters in a variety of products and processes. It must be highlighted that one of its main advantages is related with the big amount of product analysed when it works in continuous, reducing the sampling error and permitting the individual control.
Near Infrared Spectroscopy has been involved in an exciting evolution process during the last thirty years, exhibiting nowadays an excellent reputation as an accurate and reproducible technique for qualitative and quantitative analysis in many fields like agrí-food, pharmaceutical and chemical industries, human health, environmental studies, etc. Focusing on food products, most of the current applications are implemented at-line, but NIRS can be used, in-line, on-line, in situ, from farm to supermarket, in a large range of products (oils, milk, meat, grains, fruits, processed food, etc.).

NIRS analysis is nowadays conceived as a synthesis of spectroscopy, mathematics, statistics and instrumentation, being the progress in this technology closely related to the advances produced in these fields and, particularly, in the development of chemometric algorithms. Chemometrics procedures are usually applied to NIR spectra in order to extract relevant information, as the spectra are influenced by features like temperature, moisture content, particle size, texture, light scattering, sample manipulation, and other sources of error. Precision and accuracy in NIRS analysis is subjected to the availability of algorithms or spectral data pre-treatments which have been developed to reduce the above mentioned effects, particularly the scatter effect, one of the most relevant factors affecting to the spectra of agro-food products.

Its implementation is not always easy, since there are several factors associated to the nature of the food products in particular and technical impediments that difficult the use of NIR for massive control of products and processes for this type of products. The main impediments are the heterogeneity of many food matrices; the existence of multiple agro-food ingredients in a single formula (sometimes more than ten in a dairy feed or at least four in a baby food for example); the chemical and physical variability of each ingredient caused by multiple agronomic factors such as soil, variety, irrigation, etc.; the ingredients behavior when are heated, pressed, wet, etc.; and the perishable nature and high moisture content of some these food products. The tremendously complex nature of the food and feed products needs more sophisticated instrumentation, sampling accessories and mathematical data processing than those needed for quality control and monitoring of other kind of products, such as the chemical and pharmaceutical ones. The large variability found between the food products, which different sizes, shapes, textures, composition and properties in general, requires almost the development of specific and product-dedicated applications for each type of products. Nevertheless, there are also common treatments, math tools, technologies, procedures and ultimately ways for deal the development and implementation of NIRS for food and feed control.

Work Package 19, framed within the Topic “Rapid, on-site, cost-effective methods for feed/food fraud detection”, aims at designing a system of “voluntary labelling” based on Near Infrared Spectroscopy (NIRS) in combination with information and communication technologies (ICTs) to be used in the Iberian Pig industry sector.

This deliverable reports on the activities carried out as part of tasks 19.5 aiming at:

- providing NIRS training guidelines
- exploring and implementing new training itineraries and methodologies, with the particular view to widening the career prospects of highly specialized researchers and professionals in the development and use of innovative NIRS sensors and ICT for the European agrofood industry.
- studying the potentialities of NIRS for other uses in food authentication, in particular focused to other applications in the meat industry.
2. Training guidelines

Guidelines for food industries that want to implement NIRS technology are outlined below. The described procedure includes: selection of the optimum NIRS instrumentation, selection of the most suitable analysis mode and building of predictive models. This is intended to be valid for the analysis of food product in general.

The most traditional application of NIRS, i.e. the quantitative analysis, requires the development of calibrations or prediction models. For quantitative applications, calibration involves ascertaining the optimum mathematical relationship between spectral data and data provided by the reference method. The model may be fairly complex, since the NIRS spectrum is highly variable and contains physical/chemical information for the sample that may be redundant. As a result, multivariate calibration is required, based on a set of absorption values from several wavelengths (Pérez-Marin et al., 2007). Figure 1 shows the main steps to consider in the development of an application using NIR based sensors, once the instrument decision have been taken.

Figure 1. Procedure for the development of NIRS applications.
2.1. Selection of the instrument and the analysis mode.

In generic terms, the main components of any NIRS instrument are: a source of radiation, a device to distinguish wavelengths, a sample presentation system, one or more detectors to convert radiation into an electrical signal, and a processor or signal emitter. The Near Infrared region of the electromagnetic spectrum covers from 780 nm to 2500 nm.

NIRS instruments have constantly been improving, offering a wider variety of systems of interaction with the sample, providing a broad range of accessories and fibre optics, and especially in recent years – developing instruments without moving parts, together with handheld and miniature devices suitable for use on-site and on-line (Sánchez and Pérez, 2012). Currently there are more than one hundred NIRS instrument manufacturers; each instrument has specific features that determine its suitability for various applications, its ease of use and its cost.

The main aspects to consider in the election of a NIRS instrument are:

1.- The optical design: discrete wavelengths (more used in the past) or continuous scanning. The last ones can be based on several technologies: grating or diffraction grating, Fourier Transform FT-NIR, AOTF, diode array instruments, MEMS, LVF (Linear Variable Filters), etc.

2.- The modes of analysis (figure 2), depending on the arrangement of the detectors with regard to sample presentation: transmittance/transmission, reflectance/reflection, transfectance and interaction-reflectance. Usually, reflectance and transmittance are used for solid samples; transmission or transfectance for liquids. For solids, interactance-reflectance is used as well, when a fibre optics probe is embedded in the instrument, and contact is required for taken the measurement.

3.- The place where the analysis are going to be made: lab instruments for at-line analysis (mainly monochromator and FT-NIR devices); instrument for on-line and in-line analysis (mainly diode array, FT-NIR, LVF devices); instruments for in situ analysis (on farm, in field) with portable and miniature devices (mainly based on LFT, diode array and MEMS, technologies).

4.- Others aspects to consider: contact or non-contact analysis, accessories available (cups with different sizes, fibre optic probes, liquids vials etc.); resolution, noise and stability of the signal; scan rates (minutes to milliseconds and microseconds); software; optical range (from 780 to 2500 nm); detector type; prices (from 3,000 to 80,000-90,000 $).

Therefore, the decision making regarding the purchase of the most appropriate instrument, for a given application, is not easy and it needs external advise in addition of the one provides by instruments sales staff.

When defining instrument performance for a specific NIR spectroscopic application, the terms detection limit, sensitivity or signal-to-noise are used. Detection limit can be loosely approximated for any NIR spectroscopic method as equal to three times the Standard error of calibration (SEC) for the specified application. This value approximates a 99% confidence that a difference in concentration between two samples can be correctly determined (Workman, 2007).
Figure 2. NIRS analysis modes (adapted from Kawano, 2004).

**2.1. Spectral data acquisition**

Instrument performance should be check, using the control tests provided by the instrument manufacturer. Moreover, references must be taken at least once at the beginning of the analysis (black
reference and white reference, usually Spectralon). It is also recommended to analyse daily a check-sample — if it is possible unalterable, packed in a sealed cup, made of a material similar to the product analysed — to can monitor any change in the instrument and model performances, enabling to correct any deviation.

Optimal analysis mode (see section 2.1) must be selected according to the nature, size and shape of the products. Samples for NIRS analysis should be representative of the product analysed, covering the variability to be expected.

Before taking spectra, it is recommended to fit the spectral repeatability of the specific application. For this, it can be used the Root Mean Squared (RMS) statistic from a sample analysed 10 times or from 10 samples analysed by duplicate each one. The RMS statistic is defined as the averaged root mean square of differences between the different subsamples scanned at n wavelengths (Shenk & Westerhaus, 1995). This statistic indicates the similarity between different spectra of a single sample. The RMS for an individual subsample (j) of the sample (k), and the MEAN and STD values for a given k sample were calculated according to these formulæ.

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

\[
MEAN_k = \frac{\sum_{j=1}^{N} (RMS_j)^2}{N}
\]

\[
STD_k = \sqrt{\sum_{j=1}^{N} (RMS_j)^2 / (N - 1)}
\]

where \(y_{ij}\) is \(\log(1/R)\) at wavelength \(i\) for subsample \(j\), and \((y_{i})_j\) 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 NIRS instrument. The RMS value obtained in each case was multiplied by 106 to facilitate value management and processing.

A STD limit to establish the RMS cutoff for each product and analysis mode can be calculated for comparing the RMS values of subsamples (Pérez-Marín et al. 2019).

\[
STD_{\text{limit}} = 1.036 \sqrt{\sum_{k=1}^{m} STD_k^2 / m} = 1.036 \sqrt{\overline{STD^2}}
\]

where \(\overline{STD^2}\) is the standard deviation per sample and \(m\) is the number of samples.

Once the repeatability of the method is stabilish, it is recommended to take between 2 to 10 spectra per sample, depending on the heterogeneity and shape of the product analysed, and considering the
size of the analysis windows available for the NIRS instrument used. Quality of spectral data is very important to have models with a high accuracy and precision.

2.2. Reference data

For development NIRS prediction models, both quantitative or qualitative, it is required to have available information of the sample. This information could be quantitative — contents of moisture, protein, fat, starch, aminoacids, sugars, texture values, etc.— or qualitative — as commercial category, variety, species, etc. But always it is compulsory to have reliable and authentic data.

The reference method used for calibration should follow internationally accepted methods, replicates by samples must be considered (at least duplicates) and the results must be reproducible and repeatable using statistical control, as the standard error of differences between replicates. It is essential to know the precision of the reference method.

2.3. Calibration/training population

The population of used for the development of the calibrations or discriminant models and its distribution, what is called calibration or training sets, is one of the main aspects to have a successful application.

This population should cover all potential variation that may be encountered in routine production regarding the values of the parameter studied and the different factors affecting the products studied (water content, soil, variety, climate conditions, industrial processing, etc.). Moreover, it should contain a sufficient number of samples to statistically define the relationships between the spectral variables and the component concentrations or properties to be modeled. In applications related to food products, both fresh and processed, depending on the complexity of the samples being analysed, a minimum number of samples for calibration development should be more than 120-150 samples, and have enough variability. Nevertheless, most of the times, robust applications require having available hundreds of samples.

Selection of samples for the calibration set may be performed using mathematical population-structuring tools and choosing the most representative samples (Shenk and Westerhaus, 1991). For this goal, it can be applied Principal Component Analysis (PCA) and then the calculation of distances of each sample to the center of the population, based on the statistic GH (Global Mahalanobis distance). At the same time, this procedure enables the detection of samples displaying different behaviour (known as “outliers”), in which the anomaly may be caused by either spectroscopic or chemical information.

Usually, samples with GH values higher than 3 can be consider outliers i.e. not belonging to the population. It is important to justify the reasons or characteristics that make these samples outliers, and when these are confirmed, removing them from the training set. Thus, the detection, interpretation and possible removal of outlier samples represents a critical stage in the development of a calibration equation, since they can have a marked influence on the performance of the final model (Williams and Norris, 2001).
2.4. Data preprocessing

To extract relevant chemical information from each sample, it is necessary to use spectral-signal pretreatments which allow purely chemical information to be separated from variations of physical origin (e.g. texture, size, particle geometry). The mathematical pre-treatment of the NIR raw data is an essential step in the development of a successful NIR application. However, the right selection of the best spectral pre-treatment for an specific application is not easy, since there is not a standard procedure to decide which pre-treatment will give the best results when developing NIRS calibrations.

The most commonly-used pretreatments are derivation and scatter correction (Shenk and Westerhaus, 1995; Naes et al., 2002).

Scatter effects cause significant differences in their spectra due to uncontrollable physical variations such as non-homogeneous distribution of the particles, changes in refractive index, particle size distribution, sample packing/density variability, and sample morphology (surface roughness/shape). This contributes to varying the effective pathlength and result in additive, multiplicative and wavelength-dependent effects. Scatter correction methods, as Multiplicative Scatter Correction (MSC), Standard Normal Variate (SNV), Detrending (DT) and their combinations, are very useful in products where the scatter effect could affects the results; in particular this effect is important mainly in solid samples, with big particle size (whole grains, pellets, whole fruits, etc.). The application of pre-treatments to remove scatter effect is only recommended when you need to minimise the influence of particle size, heterogeneity or density differences. When the application developed is related to physical parameters, not always is recommended to apply these signal pre-treatments.

In the case of derivation, first and second derivatives are very popular. These signal pre-treatments are used to enhance spectral features, to reduce baseline offsets, to solve nearby peaks and, in summary, to improve spectral resolution. The recommendation is apply always a first or second derivative pre-treatment during calibration development.

It is important to consider that software packages operating with NIRS instruments offer the possibility of using several mathematical transformations, although sometimes, it does not exist a clear definition of the chemometric algorithms applied or clear recommendations for their right application. This circumstance could cause a chaotic situation for the beginners in NIRS calibration development, obtaining uncertain results with the performed predictions models.

Furthermore, the conclusions obtained by different authors regarding the use of certain spectral pre-treatments cannot be generalized to other data sets. Some previous works used an automatic routine to obtain calibrations, including 49 combinations of pre-treatments (first and second derivatives, an auto scaling procedure, Detrending and two versions of Multiplicative Scatter Correction). However, in spite of the fact that validation statistics did not show a clear distribution pattern, some combinations of pre-treatments provided consistent better results. Thus, the conclusions indicate that calibrations should be developed at least with a first or second alone or plus some scatter and baseline corrections algorithms as SNV, DETRENDING and Multiplicative scatter correction (Fernández-Cabanás et al. 2006 and 2007).
2.5. Calibration development using multivariate analysis

Calibration may be defined as the development of a prediction model that relates spectral data for the samples comprising the training set to the values provided by the reference method for the parameter in question. The calibration equation thus obtained enables prediction of the content in other samples similar to those included in the calibration set. The model may be fairly complex, since the spectrum obtained in the NIR region is highly variable, and contains physical/chemical information on the sample. Because it is therefore difficult to find specific wavelengths in the near-infrared region (780-2500 nm), there is a need for multivariate techniques, based on a set of absorption values from several wavelengths (Pérez-Marín et al., 2007).

A good calibration method must satisfy a number of requirements. First, the mathematical models applied must provide an efficient estimation and at the same time address the problem of colinearity, which is particularly marked when dealing with NIRS spectroscopic information, and can lead to unstable predictions (Pérez-Marín et al., 2007). At the same time, as it is reported previously (section 2.2), the reference data obtained using laboratory methods must be as precise as possible, since the precision of the results obtained using a secondary method such as NIRS will always be similar to those of the conventional reference method on which the calibration is based. A major factor in developing NIRS equations is the availability of a calibration set providing as much variance as possible (see section 2.3). Variance, ideally, should be both physico-chemical and spectral, and should represent the variance that one might subsequently expect to encounter during routine daily analysis of new samples. Thus, the defining features of the calibration set used will determine both the stability and the accuracy of the equations developed.

Most mathematical techniques applied for the development of NIRS calibrations are based on linear fitting methods, mainly multiple linear regression (MLR), principal component regression (PCR) and, especially, partial least squares regression (PLSR). MLR provides model-fitting using a small number of spectral variables selected from the whole NIR spectrum, in particular, those which display the strongest correlation between the absorption value and the corresponding reference datum. PCR and PLSR are calibration methods that work with the whole spectrum, although synthetising it into a series of linearly-independent variables in order to avoid data colinearity.

PLSR assumes that spectral information is a function of a small number of linearly-independent variables; calculation of those variables is based on both spectral data and reference values for the parameter measured in each case. PLS regression is the most widely used method in NIRS food applications.

When working with PLS, and to avoid overfitting, it is essential to set an optimum number of factors or terms for the adjusted model. For this purpose, cross validation is applied; this consists in dividing the calibration set into a series of groups (4, 5, 6, etc., depending on the number of calibration samples, usually 4 if you have available more than 120 samples), each of which is used to validate the equation developed with the others. Cross validation enables all the samples to be used for both calibration and validation, and prevents overfitting of the model.

In many current and potential applications of NIR spectroscopic measurement, the relationship to be modelled is not always linear. The source of non-linearity may vary widely, and is difficult to identify. In these cases, it is required the use of special non-linearity adjustment tools, as Local Regressions,
ANN (Artificial Neural Network), Support Vector Machines, Bayesian approaches, etc. (Pérez-Marín et al., 2007).

Nevertheless, a general recommendation is to apply always the simplest method possible for the development of the models, and only sophisticates the method when the application requires it.

2.6. Statistics used in evaluating and optimizing calibration models

Several statistics are used to evaluate and optimize the performance of multivariate calibration models. These statistics are generally applied only to data in the calibration set; they should not be confused with the statistics that are used to validate the model, that are calculated based on a separately external validation set. Sometimes statistics resulting from cross validation procedure are used to evaluate the performance of the calibrations, since they provide more precise information about how the models will work later in routine. The main statistic used for evaluating and optimizing calibrations are listed below.

- **Number of variables in a model**: Determining how many variables (wavelengths in principal components or PLS latent variables) to use in a model is a critical step in the model development. There are no hard and fast rules upon which to make this determination. In general, if too few variables are used, a less precise model will result. If too many variables are used, the estimates from the model may be unstable, that is, small changes in the spectrum on the order of the spectral noise may produce statistically significant changes in the estimates, i.e. the model will be overfitted. Thus, the number of samples used to develop the calibration model should be much greater than the number of variables or PLS factors. A minimum standard of a 6:1 ratio of sample number to factors is proposed by the ASTM (1997). However, the most commonly used informal rule of thumb is a 10:1 ratio.

As it has been mentioned before, cross validation procedures are used to estimate the optimum number of variables that should be included in a model. In cross validation, one or more sample spectra are removed from the data matrix, their corresponding reference values are removed from the reference value vector, and a model is built on the remaining samples. The model is then used to estimate the value for the samples that were left out. This process is repeated until each sample has been left out once. A complete description of the cross validation procedure (algorithm) can be found in the bibliography (Draper & Smith 1998).

- **Standard Error of Calibration (SEC)**: SEC is the standard deviation for the differences between the reference and NIR estimated values for the calibration set. It is an indication of the total residual error due to the particular regression equation to which it applies. The SEC will generally decrease when the number of independent variables used in the model increases, indicating that increasing the number of terms will allow more variation in the data to be explained, or "fitted". But too many terms could produce the overfitting of the models. The SEC is used for estimating the expected agreement between values predicted using the calibration models and values measured by the reference method, using the following formula.
\[
SEC = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2}{N - p - 1}}
\]

being \(y_i\) the reference data for the calibration set; \(\bar{y}_i\) the predicted data for the calibration set; \(N\) the number of samples in the calibration set; \(P\) the number of terms for the regression.

It is not a well estimator of the calibration performance, just an estimator of the adjustment of the model.

- **Standard Error of Cross Validation (SECV)**: SECV is one of the best estimators of the predictive capacity of an equation, being equivalent to the mean standard error of prediction (SEP) of 10 randomly selected groups of samples (Shenk and Westerhaus, 1996). It can be calculated as follow, in a similar way to the SEC.

\[
SECV = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2}{N}}
\]

- **Coefficient of determination for calibration** (\(R^2\)). The \(R^2\) represents the percentage of the population variability explained by the model developed. Its value ranges from 0 to 1, and the goal is to be closed to 1.

- **Coefficient of determination for cross validation** (\(r^2\)). The same concept that the previous one but for the cross validation set.

- **Additional statistics for evaluating the mathematical models**: There are other very useful statistics enable SEC or SECV to be standardised or converted to comparable units, facilitating the comparison of results obtained using sets with different means, standard deviations and ranges (Williams and Norris, 2001). The most used are: The Residual Predictive Deviation (RPD), calculated as the ratio of the standard deviation (SD) of the reference data to the SEC or SECV (recommended values higher than 3, although currently there is not agreement in the scientific community about these recommended values, since several authors have shown that depends on the type of sample, population variability, etc.); The Range Error Ratio (RER), defined by the ratio between the range of the reference data to the SEC or SECV (recommended values higher than 10. Same comment than for RPD); The coefficient of variation (CV), which relates the SEC or SECV to the mean (the goal is to have the lower CV).

### 2.7. Outlier detection during calibration development

During calibration passes, it can be detected outliers, i.e. samples that have unusually high residuals (difference between reference and predicted values) or spectrally different, based on the used of two types of statistics, student T statistic and the GH distance, enabling to identify samples which could represent an extrapolation of the model. Percentages between 10-15% of sample outliers can be considered natural. Higher outlier percentages can indicate problems with the method repeatability or model adjustment.
Student statistic is calculated as follow, considering as outliers samples with t value higher than 2.5.

\[ T = \frac{\text{Residual}}{\text{SEC}} = \frac{y_i - \bar{y}_i}{\text{SEC}} \]

being \( y_i \) the reference data for the calibration set; \( \bar{y}_i \) the predicted data for the calibration set; SEC, the standard error of calibration.

T outliers could be of two main types. One of then is referred to samples representing an extreme composition regarding the calibration set. These samples have very high leverage on the regression results, determining at least one of the regression coefficient values. Their inclusion in the calibration may lead to erroneous estimations of similar samples if the reference value for the high leverage sample is erroneous. The second type is one for which the estimated value differs from the reference value by a statistically significant amount. Such outliers could indicate an error in the reference measurement or a failure of the model.

Therefore, if a sample is identified as an outlier based on studentized residuals, the reference value could be erroneous. These samples should be re-analysed using the reference method. The exclusion of any such samples should be documented and justified by the applicant. If the reference value is correct, then the large studentized residuals may indicate a basic failure in the model. Removing outliers of this type without evidence of error in the reference value should be avoided whenever possible, since these samples may provide the only indication that the model is not applicable to a certain class of materials.

### 2.8. Validation of the models

The basis of the validation is the comparison of results obtained by analysis of the same set of samples by the NIRS and reference methods. Therefore, the best-fitting equations should be subsequently evaluated by a validation procedure, where the predictive ability of an equation is tested with an external sample set (validation set), representative of the population to be analyzed, which has not been used for calibration.

For the validation of a multivariate model, an ideal validation sample set will:

- Contain samples for which the range of variation of the studied parameter is comparable to the range of variation expected for samples to be predicted in routine with the model.

- Contain samples for which the concentrations of the studied parameters are uniformly distributed over their total range of variation

- Contain a sufficient number of samples to statistically test the relationships between the spectral variables and the component concentrations or properties that were modeled.

Spectra of validation samples should be collected using exactly the same procedures used to collect spectra of the calibration samples. Reference values for the validation samples should be obtained using the same reference method used for the calibration samples. After that, comparison between reference and predicted values based on statistic criteria should be accomplished.
There are different statistical validation protocols. One of the most used with NIRS data is the one proposed by Windham et al. (1989), based on the determination of a known significant error, termed “bias”, and an unexplained significant error, termed SEP(c) (standard error of prediction, bias-corrected). Generally, for calibration groups comprising 100 or more samples, and validation groups containing nine or more samples, the following control limits are assumed (Shenk et al., 1989):

Limit Control SEP(c) = 1.30 x SEC (standard error of calibration).

Limit Control bias = ± 0.60 x SEC (standard error of calibration).

The SEP (standard error of prediction) is the most important and most widely-used statistic to estimate the predictive ability of a NIRS calibration equation. The SEP represents the uncertainty of a NIRS prediction (Mark & Workman 2003). Generally with NIRS analysis, it is usually considered the 95% of probability as the uncertainty value. In this case, the constituent to be measured will be in the range: (\( y_{\text{pred}} - 2 \times \text{SEP}, y_{\text{pred}} + 2 \times \text{SEP} \)).

The total error variance in an NIR application includes the variance errors of the spectral analysis, the adjustment of the model and the reference method. If the errors associated to NIRS are controlled the target for the SEP should be the error associated to the reference method (SEL, standard error of laboratory). Thus, values of the SEP similar to SEL indicate an excellent accuracy of the model, SEP values ranges from 2 or 3 times the SEL indicate good accuracy. In the other hand, if the SEP is 5 times the SEL, the accuracy of the model is considered low.

### 2.9. Routine analysis and recalibration

Once the models developed have been validated and the results are positive according to the statistic protocol criteria, these models can be used for routine analysis for the prediction of new unknown samples.

When the calibrations are applied in routine analysis, to know if the model response is acceptable, it is recommended to obtain some control statistics together with the predictions. For this purpose, there are several statistics available.

Among them, it can be highlighted the following ones: 1.- The \( t \) Student statistic (described in section 2.7) which informs if the predicted values are included in the range of the calibration set for each parameter. 2. The distance GH (Global Mahalanobis distance, defined in section 2.3) informs if the sample is similar spectrally to the samples belonged to the calibration set, measuring its distance to the centre of the calibration population. 3. The distance NH (Neighbourhood Mahalanobis distance), which indicates if the sample has neighbours, i.e. similar samples around it; this is important since if a sample has more neighbours means that is better represented, identified by the model, and that can be predicted with more accuracy and precision. The systematic recommended is that when a predicted sample has \( t \) values >2.5, GH >3 or NH>1, must be sent to the laboratory to be confirmed by the reference method.

This kind of samples should be added later on to the existing model for a recalibration process, in order to make the calibrations more robust and to improve their generalisation and prediction capacity by increasing the range of applicability of the model. Recalibration is required whenever the instrument...
changes, samples exhibit many large GH and NH values, or if during monitoring of NIRS analyses using the primary reference method, the test results continue to fall outside the control limits (Shenk et al., 2007). Thus, the calibration process with NIR sensors is dynamic; therefore, it is recommended a recalibration each year or each two years in the case of applications more robust and stable.

The procedure followed in routine analysis for a NIRS application is shown in Figure 3.

Figure 3. Procedure for NIRS routine analysis.

2.10. Sources of error in NIRS analysis

Since routine use of NIRS technology is simple (limited or no sample preparation), fewer errors are introduced than in traditional analytical techniques. However, despite the excellent results achieved using NIRS, a number of sources of error have been identified, and it is important to know in order to minimise or remove their effect in the measurements. Williams and Norris (2001) classified sources of error into three major groups:

1) Instrument sources (e.g. signal/noise ratio, wavelength accuracy and precision, signal linearity, capsule type and size).
2) Sample sources (e.g. homogeneity, density, texture, granulometry, sensitivity to temperature).

3) Operational sources (e.g. sample preparation, operator errors, statistical strategies and procedures used).

Table 1 shows some possible sources of error that can occur during spectral analysis and some potential solutions for these problems.

**Table 1. General sources of error in spectral measurements.**

<table>
<thead>
<tr>
<th>Source of spectral error</th>
<th>Possible solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor Instrument performance</td>
<td>Conduct instrument performance tests to monitor changes in instrument performance. Analyze a check-cell sample diary to determine the instrument performance.</td>
</tr>
<tr>
<td>Absorbance exceeds linear response range</td>
<td>Determine linear response range for instrument. Choose pathlengths to keep bands of interest in range.</td>
</tr>
<tr>
<td>Optical polarization effects</td>
<td>Use depolarizing elements. Improve sample presentation methods. Investigate commercially available sample presentation equipment.</td>
</tr>
<tr>
<td>Optical component contamination</td>
<td>Inspect windows, lens, etc., for contamination.</td>
</tr>
</tbody>
</table>

Table 2 shows errors arising from sampling problems and possible solutions to these problems.

**Table 2. Sampling-related errors.**

<table>
<thead>
<tr>
<th>Sampling Error</th>
<th>Possible solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-homogeneity of samples</td>
<td>Improve mixing guidelines or grinding procedures, or both. Average replicate repacks. Rotate sample cups. Measure multiple aliquots from large sample volume.</td>
</tr>
<tr>
<td>Physical variation in solid samples</td>
<td>Improve sample mixing during sample preparation. Diffuse light before it strikes the sample using a light diffusing plate. Use signal pre-treatments. Average multiple repacks of each sample Rotate sample or average five sample measurements</td>
</tr>
</tbody>
</table>
### Sampling Error

<table>
<thead>
<tr>
<th>Chemical variation in sample with time</th>
<th>Freeze/dry sample for storage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Identification of kinetics of chemical change and avoidance of rapidly changing spectral regions</td>
</tr>
</tbody>
</table>

3. **Information on current and future courses and workshops**

Regarding the future courses and workshops, it is worth to mention the face-to-face course “Espectroscopia De Infrarrojo Cercano (NIRS). Aplicaciones En El Control De Calidad Y Trazabilidad De Productos Y Procesos” (in Spanish language). This specific training is offered annually in February by the University of Cordoba (Spain) and supervised by the Prof. Ana Garrido-Varo and Dolores Pérez-Marín. Participants per course edition are limited to 20. Next course edition (19th edition) will be held on 11th-14th February 2019. More information can be found here: [https://www.uco.es/servicios/scai/novedades/curso_nir_2018.html](https://www.uco.es/servicios/scai/novedades/curso_nir_2018.html).

The objective of the course is to provide basic information on scientific and technical aspects related to NIRS technology. The profile of the attendees includes teachers, researchers, collaborators, undergraduate and postgraduate students and technical staff from different private companies and public/private research centres. Regarding origin, the participants in the course come mainly from Spain, Portugal and South America. The beneficiaries of the course acquire knowledge on the instantaneous and non-destructive analysis, which will allow them to benefit from the experience accumulated by the Engineering Production Systems Research Group (Professors Ana Garrido-Varo and Dolores Pérez-Marín belong to this research group) and the instrumentation and software available in the NIR/MIR spectroscopy unit of the SCAI (Central Service for the Research Support in UCO) and in the Sensor Unit coordinated by Prof. Garrido and Prof. Pérez-Marín.

The theoretical content includes:

1. Immersion and introduction in the course: program, objectives, teaching methodology, sources of information.

2. Introduction to NIRS technology. Brief historical development.

3. Theoretical bases of diffuse reflectance. Modes of interaction between radiation and sample. Instrumentation. Accessories and software NIRS.

4. The process of development of quantitative and qualitative NIRS applications. Critical stages in the process of developing NIRS applications. Sources of error in the NIRS analysis.

5. Development and evaluation of calibrations. Spectral signal pre-treatments. Regression methods. Study of spectral populations: algorithms of knowledge of the structure of the collective of calibration (Center) and of selection of representative samples (Select).

7. Qualitative analysis. Discriminant methods: Principal component analysis; Discriminant factor analysis; SIMCA; PLS discriminant.

8. Applications of NIRS technology to quality control and traceability of products and processes. NIRS analytical services in the world. Future of NIRS technology.

Practical sections include:

1. Practical exercises to obtain spectroscopic data, with different equipment and modes of analysis.


3. Development and evaluation of qualitative applications. Statistics and interpretation of the results

A second very interesting international NIRS course is the titled “Fundamentals and Applications of Near Infrared Spectroscopy”, completely on-line and taught in English language. The course has been selected as the first joint initiative of the IVPTL-NIRS (International Virtual Platform for Teaching and Learning Near Infrared Spectroscopy), a cooperative online educational project launched by the International Council of Near Infrared Spectroscopy (ICNIRS) and the University of Córdoba (UCO).

It is the first NIRS virtual course that is taught entirely in English and is offered worldwide through the Moodle platform of the University of Córdoba. For this reason, one of the main objectives in its design has been to have teachers (professors and professionals of the sector) of recognized international prestige.

At this moment, two editions have been held in 2017 and 2018. Next edition will be foreseeably held at the end of 2019. More information can be found here: [http://www.uco.es/nirsplatform/index.php/9-news-announcements/36-2nd-edition-of-fundamentals-and-applications-of-near-infrared-technology](http://www.uco.es/nirsplatform/index.php/9-news-announcements/36-2nd-edition-of-fundamentals-and-applications-of-near-infrared-technology). The profiles of the participants in the first two editions of this international course on NIRS Technology comprise 51 attendees from many countries around the world as Spain, Canada, Egypt, Australia, Colombia, Ecuador, Philippines, The Nederland’s, Iran, New Zealand, Portugal, South Korea, UK, Argentina, Italia, Denmark or Brazil.

It also worth to mention that during the two editions of both courses, the supervisors (Prof. Garrido and Prf. Pérez-Marín), in their lectures, have presented and described results obtained in the frame of the FOODINTEGRITY project as an example of successful NIRS applications.

It is also important to highlight the International Council of NIRS organized biannually International Conference about the technology [http://www.icnirs.org](http://www.icnirs.org). This is am important date for the NIR scientific community and users. Next edition will be held in Australia in September 2019 [http://www.nir2019.com/](http://www.nir2019.com/)
4. Wider applications of NIR for other uses in food authentication

Due to its speed and precision, near-infrared reflectance spectroscopy (NIRS) has become a widely used analytical technique in many food products and industries (specific information can be found in numerous reviews paper and specific books). It offers, moreover, a number of other advantages which make it ideal for meeting current demands in terms of integrity, control, safety and traceability: low cost per sample analysed; little or no need for sample preparation; ability to analyse a wide range of products and parameters; environmental friendly (no reagents are required, so produces no wastes); a high degree of reproducibility and repeatability and the possibility to analyse more amount of product, improving the sampling and permitting the individual control. NIRS can be used at-line, but also in-line, on-line or in situ.

The specific development of a portable miniaturised NIR sensor and the methodology designed and tuned in WP19 & WP21 of the FOODINTEGRITY project could be of application for other fields and food products. Thus, over the last two years, the NIRS system (MicroNIR Onsite, Viavi Solutions, USA) used to collect spectra from pigs’ fat tissue in the frame of the FOODINTEGRITY project, has been used by members of the UCO (University of Cordoba) Research Team to analyse other agri-food products. In particular, this sensor has been successfully tested in meat samples of vacuum packaged sliced cured ham to discriminate “Premium” vs “Non-Premium” categories and to determine fatty acid profile, demonstrating the feasibility of NIRS for the quality assessment of this product.

In addition, in vegetables, this NIRS instrument has been used to quantify nitrates, sugars and dry matter in spinach (Spinacia oleracea), peppers (Capsicum annuum) and squash (Cucurbita pepo). Furthermore, an ongoing study is trying to relate spectral features with maturity of oranges. Acidity and sugars have been demonstrated to have a strong correlation with NIRS spectra, and other characteristics like limonin content are currently being explored.

This sensor is being applied in transmittance mode for the analysis of olive oil for the prediction of several quality parameters and for its authentication.

In animal feeding, the instrument is being evaluated for the analysis of fresh silage and total mixed rations, with the goal to characterize the feeds and design and fit the diets for the cows, looking for efficiency in what now is called livestock precision farming. In addition, the instrument is being evaluated for the analysis of meat and bone meal for species authentication.

All the above-mentioned research works have been carried out thanks to the incentive to invest in new technologies shown by national and local companies, together with by national project funds. Some of the results are published in specialized high impact research journals (Pérez-Marín et al., 2019; Sánchez et al. 2018), and others have been included in final degree projects or master’s thesis of UCO’s students supervised by Professors Ana Garrido Varo and/or Dolores Pérez Marín, now being processing to be submitted for publication.
5. NIRS implementation in companies.

The level of implementation of NIRS technology in the food industry is not homogeneous around the world and in the different fields. At laboratory level, the implementation in the industry started several decades ago, and now in general it is well implemented in animal feeding, cereals, flourmills, oils or biscuits. Some implementation exists in meat industry, processed food and dairy. Fewer examples around the world are found in the vegetable and fruits sectors.

Nowadays, the trend in the implementation of NIR sensors is directed to the use of these devices in the control of products and processes but in a continuous way, i.e. on-line, in-line or in situ analysis. Thus, currently the interest of the worldwide industry is to can incorporate spectral sensors that enable them to automatically control and monitor their production. This is not easy, since NIR sensors are not simple analysers. Because of that, an implementation NIRS project is require during at least 1-2 years for the development of robust prediction models, and some technical consulting during this period. After that, the use in routine of the NIRS technology and the models developed is easy and friendly for the technical staff.

Nevertheless, it is important to mention that the demand of the industry for the use of NIR sensors is increasing everyday. In Spain, UCO team has everyday new demands of different companies. It is important to stress that IBESA, the company involved in the project as representative of the traditional Iberian pig ham producers, is willing to integrate the NIRS technology developed in the frame of the FOODINTEGRITY project to their traceability system.

Finally, it is important to mention that in the FOODINTEGRITY project, it was dealt the combination of the NIR spectral signal with information and communication technologies (ICTs) to be used in the Iberian Pig industry sector. The App was designed to access to a web virtual environment through cloud computing from mobile/tablet devices for the instantaneous monitoring of Iberian pig quality. In this way, non-expert users could make, analyse and interpret measurements in a simple and fast way. A specific guideline for the use of an App described in D19.5 and D19.6 (for both monitoring and consumers) can be found in Annexes I and II in this report.

6. References


Annex I. Graphical guidelines for the use of the App for a fast and instantaneous prediction of quality parameters of Iberian pig carcasses for monitoring the production process
DATA
Run Analysis is disabled.
ANALYSIS
Annex II. Graphical guidelines for the use of the App for consumers
INFO
Welcome!
(1) According to NIRS qualitative model
(2) r > 2.5, possible reference outlier
Authorised Members only